Improving the container-planning at Royal Grolsch

Student information

Student:	R.M. Spanninga
Study:	Industrial Engineering and Management
Specialization:	Production and Logistics Management

Organizational information

Company Supervisor:	MSc K. Kamp
University lead supervisor:	Dr. P.C. Schuur
University second supervisor:	Dr. E. Topan

Public version

Sensitive information is left out

Management summary

Annually, Grolsch has to buy new containers to make up for losses in the market, losses during production or to support sales growth. Currently, the planning for the procurement of new containers (injection) is done using a long-term container-planning model. This model forecasts container returns and then decides based on the production plan when new bottles are needed. However, the injection plan is not always accurate and how the input parameters for this model are calculated is questioned. A more accurate injection plan can lead to a reduction in overstocking- and understocking costs. These costs include investment costs, holding costs, changeover costs and stockout costs. We have started this research with the intention to only improve the calculation of the input parameters of the container-planning model for two bottle types: the 30cl green bottle (Apollo) and the 30cl brown bottle (BNR). These input parameters are:

- Trade Loss (TL), the percentage of containers lost in the market
- Internal Loss (IL), the percentage of containers lost at the brewery
- Trade Population (TP), the number of containers currently in the market
- Weeks in Trade (WiT), the number of previous weeks of sales that are currently still in the market and have not yet returned
- Days of Cover (Doc), the working days of production that the empty container stock should cover (safety stock)

After analyzing the current situation of how Apollo and BNR returns are currently forecasted, we have decided to also improve the return forecasting model. After making a parameter calculation model in Excel and improving the return forecast, we have extended the research further by also proposing a different purchasing policy for new Apollo and BNR bottles. We propose to calculate the safety stock of empty bottles with a well-known formula instead of the current way based on experts opinions. The main research question of this research is therefore:

"How can the injection planning for Apollo and BNR be improved, by improving the long-term container-planning model's input parameter calculation, return forecast and purchasing policy?"

Current situation

The first part of the research is the analysis of the current container return process and the description of how the current container-planning model works. We describe the inputs, the outputs and how these different inputs and outputs of the current model are currently calculated. Afterwards, the performance of the current container-planning model is measured by the KPI "Return forecast accuracy". The Mean Absolute Percentage Errors (MAPE) of the weekly return forecasts are considered high with 25% for Apollo and 44% for BNR. We conclude that the way Grolsch currently forecasts returns (using WiT) can be improved.

Literature review

The literature review is used to see what is written about return forecasting and procurement in reverse logistics. The first takeaway from the literature is that there is a difference between general sales forecasting and return forecasting. With return forecasting, the variable's values should depend on another explanatory variable (sales) instead of just the variable's own past values. This makes standard forecasting methods as moving-average, exponential smoothing or Holt's method less appropriate.

Because Grolsch does not register the production code of each individual returned bottle, the real time that a container stays in the market is unknown. There are only aggregate data available: total

returns per container type per period. This makes return forecasting for Grolsch different than return forecasting for companies in other industries, where an individual item's time in the market is known. There are several methods and models proposed in literature that work with only aggregate return data. In our opinion, the return forecast for Grolsch can best be modeled as a finite Distributed Lag Model (DLM), that can be solved using time series analysis techniques.

For the injection planning problem (order policy) there is one particular model that is interesting. This is the model of Kelle and Silver (1989b), which can deal with uncertainty in demand and supply (returns) while these uncertainties can also be correlated with each other. With this model, the safety stock (of empty bottles) can be calculated using a well-known formula.

Parameter calculation and return forecast results

First, the parameter calculation of TL is improved. This is done by updating the calculation of the realized returns and looking at the TL stability to choose a period over which to calculate the TL. The updated values are X% for Apollo and X% for BNR. The TL for Apollo is assumed to be stable. For determining the TL of BNR, human input is necessary to make sure sudden market changes are incorporated in the TL calculation.

The proposed return forecasting model for Grolsch is a finite Distributed Lag Model (DLM), which we have solved with a lognormal distribution structure for the time in the market of bottles. This results in a Time in Trade (TiT) distribution, the probability that a container returns in a certain week after it is sold. The parameters of the lognormal distribution are estimated in Excel by including 52 weeks in the calculation after which bottles can still return, and scaling the remaining probability mass after the 52 weeks over the included 52 weeks. After we have seen the cumulative error of using one TiT-distribution for the whole year, we have concluded that a two period seasonality needs to be incorporated in the model. The average TiT we found for Apollo is X weeks and for BNR X weeks. The improvements in terms of MAPE for 2020, with the TiT-distribution fitted on data from 2018-2019 are:

	Return forecast MAPE Apollo 2020 (based on realized sales)	2020 (based on forecasted sales)
Current return forecast	24.0%	26.1%
Improved return forecast	12.6%	14.3%

Table 0.1: Return forecast MAPE Apollo

	Return forecast MAPE BNR 2020 (based on realized sales)	2020 (based on forecasted sales)
Current return forecast	35.7%	38.7%
Improved return forecast	14.8%	16.1%

Table 0.2: Return forecast MAPE BNR

We conclude that the improvement to the return forecast is significant and would likely lead to a more accurate injection planning. After the TiT-distributions are estimated, the TP can be estimated because it is known how many sales of each week are still expected to be in the market on each point in time. In Chapter 5 we used the improved return forecasting model in a new injection planning model based on the model of Kelle and Silver (1989b).

Expected costs

The new injection planning model outputs an injection plan based on the well-known formula for safety stock. In the end we have calculated the expected holding-, changeover- and stockout costs for three scenarios by simulating sales and returns with an expected normally distributed sales forecast error. These three scenarios are:

1) The current injection planning model with the current return forecast (CIP+CRF)
2) The current injection planning model with the new return forecast (CIP+NRF)
3) The new injection planning model with the new return forecast (NIP+NRF)

The results for Apollo are as follows:

Scenario	Expected total costs	Expected holding costs	Expected changeover costs	Expected stockout costs	Injection
1.CIP+CRF	X	X	X	X	X
2.CIP+NRF	X	X	X	X	X
3.NIP+NRF	**X**	**X**	**X**	**X**	**X**

Table 0.3: Expected costs Apollo

The results for BNR are:

Scenario	Expected total costs	Expected holding costs	Expected changeover costs	Expected stockout costs	Injection
1.CIP+CRF	X	X	X	X	X
2.CIP+NRF	X	X	X	X	X
3.NIP+NRF	**X**	**X**	**X**	**X**	**X**

Table 0.4: Expected costs BNR

Stockout costs and holding costs both decrease from scenario 2 to 3 for BNR. This is the case because the timing of the injection was too late in Scenario 2 compared to Scenario 3. Sales for BNR can be volatile and the current safety stock does not protect against this. We conclude that the total costs for Apollo and BNR together can be reduced with 5%.

Recommendations

The first recommendation for Grolsch is to use the updated parameters in the current container-planning model. These parameters can be updated annually with the parameter calculation tool. Second, we recommend Grolsch to incorporate the improved return forecasting model in the current container-planning model. Not only is the improved method more realistic and accurate, the old method is very prone to changes in sales and a WiT-profile cannot just be copied to a next year. We also recommend to look into the new injection planning model as holding costs can be saved. Finally, we recommend Grolsch to keep track of where returns crates came from (from which customer), so a more accurate analysis can be done and uncertainty in the return forecast can be further reduced.

Roadmap

The container-planner of the Supply Chain Planning department can update the input parameters of the container-planning model annually by using the parameter calculation tool. The improved return forecasting model can be implemented in the current container planning model of Grolsch, which has already been tested during the research. For the improved injection planning model, a simulation needs to be done in Microsoft Excel which can also be done by the container-planner annually.

Preface

This thesis is the final part of my studies Industrial Engineering and Management at the University of Twente. I would like to thank Grolsch for the opportunity to do my master assignment at their company. At the start of my assignment, it looked as if I could work at the brewery as the numbers regarding Covid-19 looked brighter. But because of the second wave of Covid-19 I had to work from home instead. I am very happy I could join the daily operational meetings from which I learned a lot about how the production planning is done in a real company environment. I therefore want to thank my colleagues from the Supply Chain Planning department who made me feel at home in their team right from the start. Even though I have not met some colleagues in real life because of Covid-19, I always felt welcome and they were always available to answer my questions as well.

Special thanks to my supervisor from Grolsch, Kristian Kamp, who helped me a lot during the research. I really enjoyed the weekly meetings in which we discussed the project. With your knowledge of the brewing industry and also your background with the University of Twente you were the perfect supervisor for this research. You always knew where I could find the data I needed and which people I needed to ask for answers to my questions.

I would like to thank my supervisors from the University of Twente, Peter Schuur and Engin Topan for their useful feedback and nice ideas on how I could take the research further. Your feedback has helped me to improve the quality of the research as well as the readability of the report. Even though you were busy with lectures, research and giving feedback to (many) other students, you always found the time for regular meetings which helped me greatly to progress the assignment.

R.M. Spanninga,

Enschede, May 2021

List of figures

List of tables

Glossary

Word or abbreviation	Meaning
Apollo	The main green Grolsch bottle (30cl)
BNR	Brown Dutch Returnable Bottle (30cl)
CIP	Current injection planning model
Containers	Materials that can “contain” something Includes: bottles, crates and kegs
CRF	Current return forecasting model
DLM	Distributed Lag Model
DoC	Days of Cover: how many extra days of production needs to be covered with stock of empty containers
FIFO	First In First Out
hl	hectoliter
IL	Internal Loss (%): the percentage of the containers that is lost during production or in outside storage because of bad weather
Injection	Buying new containers from the supplier and move them from the supplier’s stock to the brewery
MAPE	Mean Absolute Percentage Error
Maximum lag length	Maximum amount of weeks that containers can return in (in the return forecast model)
NIP	New injection planning model
NRF	New return forecasting model
Overfitting	Fitting a model too tightly to data points
Overstocking	Buying too many new containers so holding costs are high
RFID	Radio Frequency Identification
SKU	Stock Keeping Unit
TiT	Time In Trade: how long it takes between a container leaving the brewery and being returned to the brewery
TL	Trade Loss (%): the percentage of the sold containers that does not return to the brewery
TP	Trade Population: the amount of containers that is in the market
Understocking	Buying too few new containers so risks of production inefficiencies and stockouts are high
WACC	Weighted Average Cost of Capital
WiT	Weeks in Trade: the amount of previous weeks of sales that the TP consists of

Table of contents

1. Introduction

Grolsch is a more than 400 years old Dutch beer brewery and is characterized by its taste, strength and its unique own character. Grolsch produces a wide range of beers at its brewery in Enschede, from normal beer to many special beers as Radler, Kornuit, and Seasonbok. Even though currently the most demand comes from within the Netherlands, the export market is also a big part of Grolsch. Within Grolsch there are a lot of departments that together make sure enough beer is brewed and filled to fulfill the demand.

As we have been working in the team of the Supply Chain Planning department (SCP) during this research, we briefly describe the roll of this department within Grolsch:
SCP is responsible for planning the production on the eight production lines, material planning and the brew planning. SCP therefore has a central role within Grolsch. If SCP changes the planning this has consequences for other departments (for example operations and warehouse) as well. Planning is split in long-term (tactical planning) and short term (scheduling).

At the SCP, two people are responsible for the tactical planning and they create a production plan on week level for up to 78 weeks. After the tactical planners release their plan, the two schedulers plan the output amounts of the tactical plan on the production lines. In practice this plan changes a lot, because of issues on the production lines or problems with material availability. It is the goal to change the planning as little as possible.

Another important part of SCP is the material planning. No material means no production. The two material planners use sophisticated tools to help them determine which materials are needed at what moment in what amount. They then communicate this to the suppliers and make sure that the suppliers deliver the materials on time. If there are any problems the material planners communicate with the other team members so they might change up the production plan.

The last part within the SCP department is the brewing and filtration. The location of Grolsch is called a brewery for a reason: there is not just a filling line, Grolsch brews its own beers as well. One person is responsible for planning this process that takes approximately X weeks per beer depending on the type of beer that is brewed.

The part of the SCP that we have been working on is container-planning, which is a subpart of the material planning. Grolsch has planning models for determining the returns of bottles and crates and is unsure if the input parameters for these models are correct. Because a lot of money is spent each year on new bottles and crates it is expected that savings can be realized when the models are optimized.

1.1. Problem statement

Annually, Grolsch has to procure new containers to make up for the container losses in the market, the breakage of containers on the filling lines and possibly sales growth. The procurement decisions are made the year before the containers are needed for production, as the supplier needs time to plan and produce the containers. Currently, these decisions are supported by a long-term container-planning model, that forecasts how many new containers (injections) are needed based on the returns of previously sold containers. Grolsch faces high costs for planning too little or too much injection.

First, we look at the costs of planning too much container injection. If Grolsch expects to need too many containers, the suppliers have to hold a big amount of stock. This is not only costly for the suppliers, but after some time Grolsch has to decide what to do with the stock: inject (buy the containers and move them to storage at the brewery), relocate to other beer brewing companies (possible for some bottles) or destroy. Costs that could have been saved here are: investment costs that could have been postponed (including depreciation of the assets), stock costs at the supplier (that Grolsch might pay for) and costs for relocating or destroying.

In the case of planning too little container injection, the suppliers don't hold enough stock. They might hold some safety stock, but this is often limited. The lead times on new containers from order are too long to flexibly handle additional demand. The feasibility of Grolsch' production plan depends on the number of containers that comes back from the market when suppliers are unable to deliver more containers.

The returns from the market are influenced by customer return behavior, but also by the logistics of the pick up at the client. For example: When a client of Grolsch has a promotion period, Grolsch brings many full crates to the client. In these periods Grolsch gets back more returned containers, as the policy is to take a full truck of returned containers back to the brewery after a delivery. But when the promotion has ended, there is an accumulation at the client. Now the containers that were sold during promotion get returned but Grolsch has no logistical capacity to pick them up.

Currently, Grolsch also faces high costs for the manual sorting of containers, which is needed when there are no empty containers left in storage and the automated sorting line has no capacity. This manual sorting is inefficient, costly and should be prevented if possible. Another possibility when there are too few empty containers available is to change up the planning and postpone the production for which these containers are needed. This option is also limited as changing up the planning has consequences for many departments and should be avoided if possible. Eventually, when Grolsch is unable to realize the production plan this can lead to stock-outs, lost sales and unsatisfied clients.

It all comes down to an accurate planning of the injections. In the first place so that not too much new material is ordered and in the second place that suppliers are still able to have sufficient stock to flexibly respond to higher demand of Grolsch.

Grolsch currently uses a long-term container-planning model as a decision support tool for planning the injections. However, the performance of this model is questioned as there are regularly problems with the availability of some containers. The following input parameters go into the current model:

- Trade Loss (TL), the percentage of the containers lost in the market
- Internal Loss (IL), the percentage of the containers lost at the brewery
- Trade Population (TP), the number of containers currently in the market

- Weeks in Trade (WiT), the number of previous weeks of sales that are currently in the market
- Days of Cover (DoC), the working days of production that the empty container stock should cover (safety stock)

The problem is that these input parameters are not updated for some years, and the way in which they are calculated is questioned. Besides, the way the current model forecast returns and plans injections based on these parameters can likely also be improved. The procurement decision based on this return forecast is now a simple order-up-to policy, with safety stock that is manually determined. The expectation is that with improvements to (1) the parameter calculation, (2) the return forecast and (3) the purchasing policy or safety stock calculation, the annual injection plan will be more accurate and overstocking- and understocking costs can be reduced.

The abovementioned problems are shown in the problem cluster in Figure 1. At the top we see the main problem. Below every problem is the cause of that problem. This results in possible “core problems” at the end of the tree. The light red problems are the problems we (indirectly) want to better, dark red problems are the core problems we tackled and yellow problems are our of scope for this research.

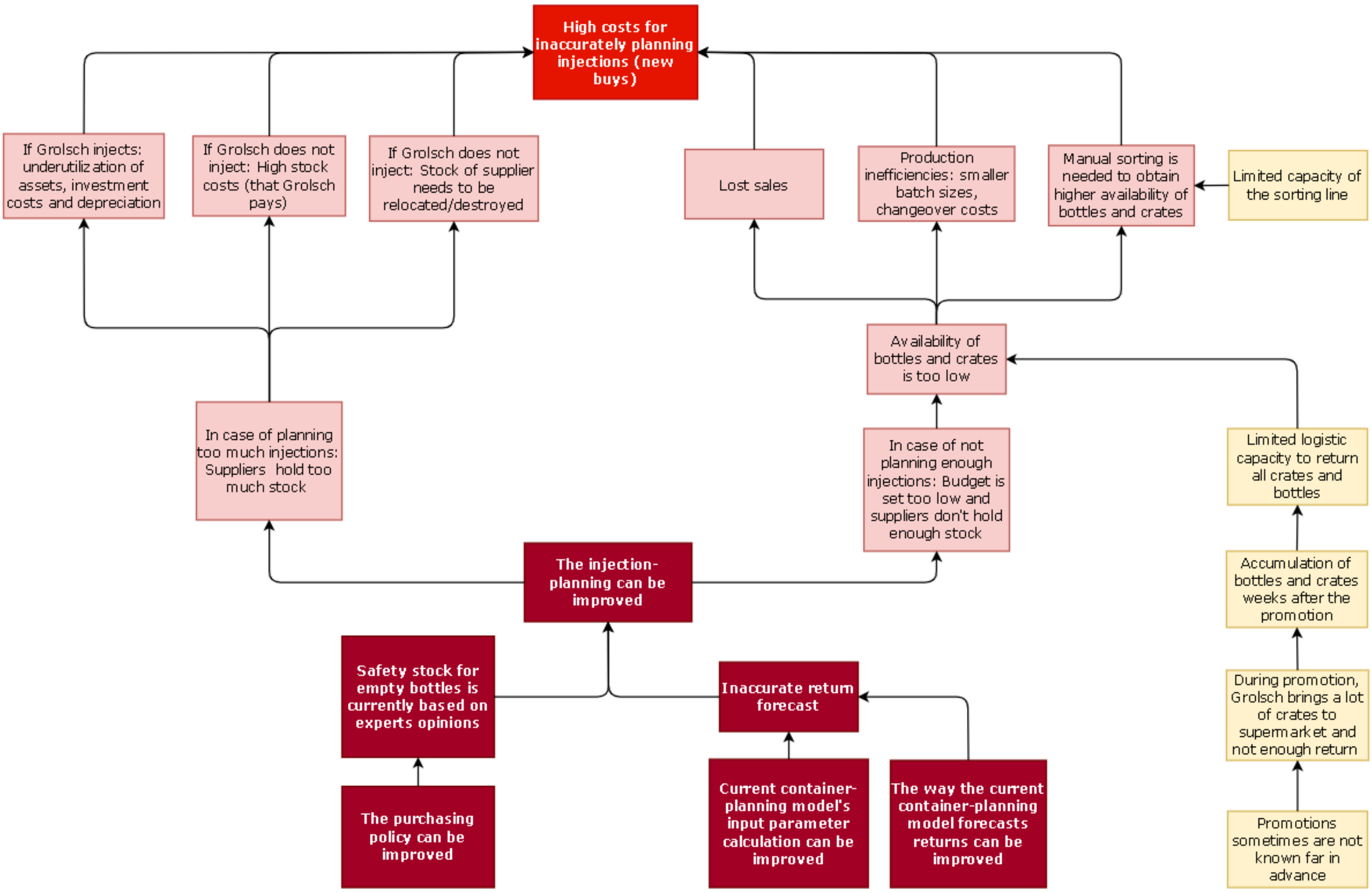

Figure 1: Problem cluster

1.2. Scope

In this research we improve Grolsch' current long-term container-planning model. We are going to improve the input parameter calculation, the long-term return forecast and the long-term injection planning for the two main bottle types: the trademark green Grolsch bottle 30cl (Apollo) and the Dutch Brown Returnable Bottle 30cl (BNR).

Improving other processes in the return process, such as: improvements to the sorting line, pick-up logistics of empty bottles, market research on why consumers return their crates and bottles in a certain time or trying to influence their behavior is out of scope. Some ideas about these topic are described in Chapter 8. These topics remain interesting for further research.

Because the returns depend on the sales, the return forecast accuracy depends on the sales forecast (in)accuracy. The focus lies on improving the return forecast and improving the sales forecast accuracy is therefore out of scope for this research.

1.3. Research questions

This section is about the research questions and the way to solve them. From the problem statement follows the main research question:

"How can the injection planning for the two main bottle types be improved, by improving the long-term container-planning model's input parameter calculation, return forecast and purchasing policy? "

As this main research question is too comprehensive to answer at once we split the question up in multiple sub research questions regarding current situation, literature review and solution design:

Current Situation

Before a parameter calculation model can be build, we need to develop a deeper understanding of the container return process and how the long-term container-planning model works. In order to come up with improvements for the calculation of the input parameters for this long-term model, a detailed analysis of the last parameter determination (done in 2018) is needed. Besides, it is unknown how good the current long-term model performs in terms of accurate planning of container injections (based on the current input parameters). We use the following research to describe the current situation:

1) *What is meant with "containers"?*

2) *How is the current container return process set up?*

3) *How does the current long-term container-planning model forecast returns and plan injections?*

4) *How are the current input parameters of this model (TL, IL, TP, WiT and DoC) calculated?*

5) *How did the model perform over the last years in terms of accuracy?*

Literature review

Grolsch does not register the time in the market of each returned container. The only data that are available (after some cleaning) are: the total amount of sales per container type per week and the total amount of returns per container type per week. It is unknown when the returned containers are sold. The literature review serves as means to see if there are solutions to similar problems where returns need to be forecasted (using a loss percentage and a time in the market) based on aggregate data. In this part we also look for papers in which ideas are presented to calculate the return parameters. Finally, some background on useful models is needed as a guide for the later parts of this report. The research questions for the literature part include:

6) *How are container returns related to returns in different industries?*

7) *How are return parameters determined in comparable industries?*

8) *Which methods and models are proposed for forecasting container returns?*

9) *Which methods and models are proposed for procurement of new materials in reverse logistics?*

10) *What is the theory behind the methods and models that are interesting for this research?*

Solution design
When it is clear which methods are available, we describe them in detail and determine which method to use to calculate the parameters TL, IL, WiT and TP. After updating the parameters, the performance of the updated long-term container-planning model can be compared to the current situation in terms of forecast accuracy and cost savings. The research questions for the solution design are:

11) *How can the parameter calculation of TL, IL, TiT and TP be improved?*

12) *How can the return forecast be improved?*

13) *How can the purchasing policy be improved?*

14) *How accurate is the updated model?*

15) *What are the expected savings per year when using the improved container-planning model over the current model?*

1.4. Research design

The research questions about the current situation are mainly answered by interviews with experts from Grolsch that are involved in the container return process and use the long-term container-planning model in their daily activities. To determine the performance of the long-term model the sales data, return data and data about the sorting process are required. These data are provided and validated by Grolsch. How to measure forecast accuracy is based on literature.

The literature review includes papers found in the scientific databases that are connected to the University of Twente.

The solution design is based on the useful methods found in literature, Grolsch' experts opinions and own ideas. Especially the practicality of the solution is important as Grolsch is part of a bigger organization and any changes to the current container-planning model should be a clear improvement, user friendly and usable for other breweries that are part of the same organization. Interviews with the people that will use the updated container-planning model are needed to make sure the solution meets these standards.

The deliverables of this research are:

- An analysis of the performance of the current long-term container-planning model.
- A literature review on container return forecasting and purchasing.
- Main deliverable: A parameter calculation model that can be used to (annually) update the input parameters of Grolsch' current long-term container-planning model.
- An improved return forecasting method, that can be built into the current long-term container-planning model of Grolsch.
- An improved purchasing policy.
- An implementation plan.

Chapter outline

We start in Chapter 2 by analyzing the current situation. We describe the current container return process, describe the current container-planning model and measure its performance in terms of return forecast accuracy. In Chapter 3 we review the literature on reverse logistics to find possibilities for improving the return forecasting model of Grolsch and to see if there are methods and models for the procurement decision in reverse logistics. In Chapter 4 we calculate the input variables for the container-planning model of Grolsch and propose a new return forecasting model. The improved return forecast of Chapter 4 is used as input in Chapter 5 in which we propose a new injection planning model. In Chapter 6 we review the injection plan output by the injection planning model of Chapter 5 in terms of holding, changeover and stockout costs. Afterwards, we describe the implementation plan in Chapter 7. In the final chapters we give our conclusions, recommendations and discuss topics for further research.

2. Current situation

In this chapter we describe the current situation of the container return process and the performance of the current long-term container-planning model. The first section is about what is meant with containers and the different types of containers. Then the return process is described from when a container is sold till it is available again for production. The explanation of the current long-term model is found in Section 2.3 and the performance of this model is the end of this chapter.

2.1. Containers

Containers are all materials that “contain” something. The returnable containers are the returnable bottles and crates and also the returnable kegs. As mentioned in the scope, the focus lies on the two main bottle types in this research: the 30cl green bottle, and the 30cl brown. Important to note that the same bottle type is used for the filling of different types of beer. Specific bottles are needed in specific crates for production. The different bottle-types that Grolsch uses are:

Figure 2: Bottle types. (From left to right) Apollo bottle 30cl, Bruine Nederlandse Retourfles (BNR) 30cl, Kornuit bottle 30cl, Brown Swingtop bottle 45cl, Green Swingtop bottle 45cl.

The standard green 30cl Grolsch bottle, the Apollo bottle, is used for normal beer and the Radler variants. BNR, or in English “Brown Dutch Returnable Bottle” is a bottle that is used among more breweries and is also used for the most sorts of beers within Grolsch. The fourth bottle in Figure 2 is the characteristic Swingtop bottle with the famous “plop” sound upon opening. This bottle also has a brown variant and both have 1.5L variants. The percentage of sales of brown compared to green is small and the brown and 1.5L variants are not included in the current container-planning model. The Kornuit bottle is introduced in 2018, but has already earned his spot in the container-planning model. In this research we focus on the Apollo and BNR bottles.

Bottles are sold and returned in the following types of crates:

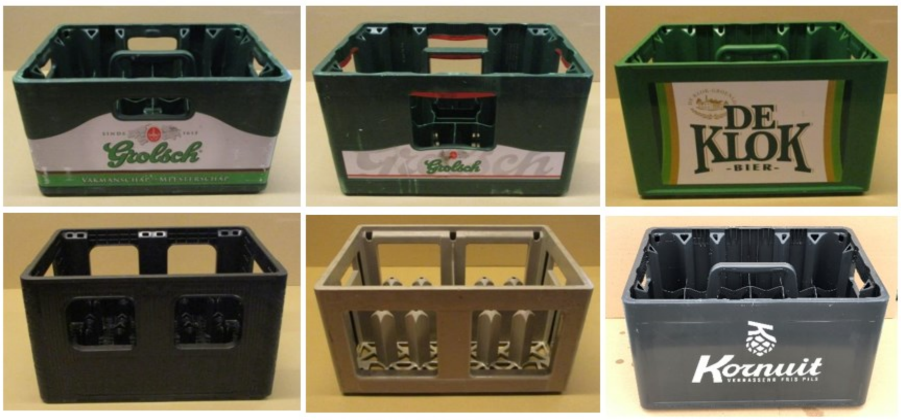

Figure 3: Crate types. (From left to right) Top: Eagle crate, Swingtop Crate, De Klok crate. Bottom: Pinolen crate, Pelican crate, Kornuit crate. Bron: Grolsch' handboek emballage artikelen (2019)

The Pinolen and Pelican crates are also used by other brewing companies. As quickly mentioned, a bottle can be needed for production in different types of crates.

2.2. Return process

The return process consists of many steps with multiple parties involved. To get a good understanding of the process, we first describe these steps in detail. The process is mapped in the flowchart in Figure 4 and the different steps are described in the rest of this section. Green lines are return streams, red lines are waste streams and black lines are general flows.

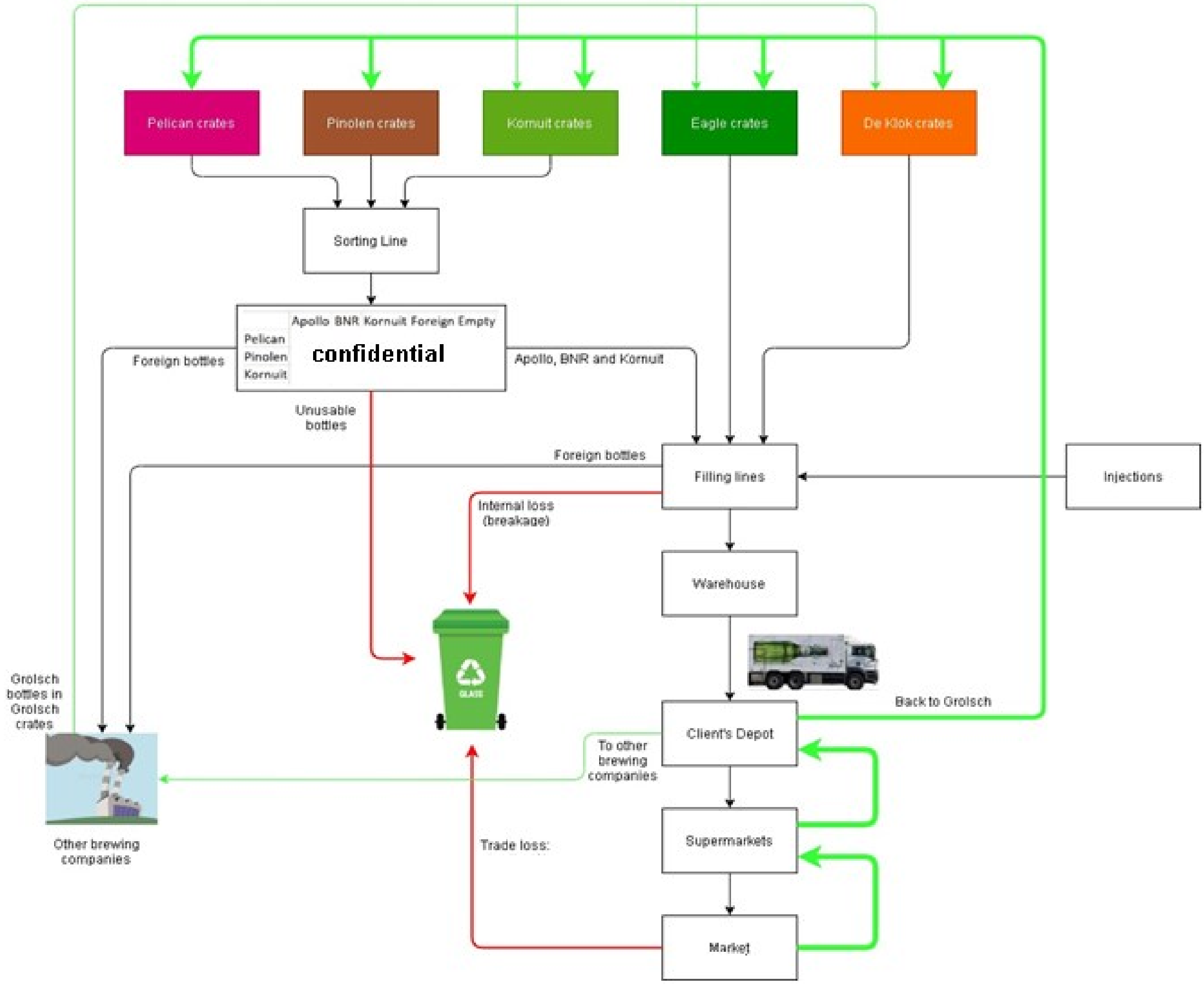

Figure 4: Container return process

From filling the beer till return to the brewery

After the beer has been filled on a filling line it gets stored in the warehouse, waiting for departure to a client. The beer is moved out in a First In First Out (FIFO) manner to reduce the amount of obsoletes. The time in the warehouse is different per type of beer as it is mainly determined if the type of beer is a slow-moving or fast-moving. A general assumption is that the beer stays in the warehouse approximately one full week.

After the beer is delivered to a client it takes some time before the consumer buys the beer from the client. In a supermarket the beer is waiting on the shelves to be bought by a consumer. In bars and restaurants it takes time before the beer is ordered and consumed. On top of the beer price Grolsch' clients, or the end consumer who buys the beer in the supermarket, need to pay a deposit as a motivation to return the bottles and crates. How long a crate or bottle stays in the market depends on the return-behavior of the consumer and the pickup logistics of Grolsch. Empty bottles could be available at the supermarket, but they don't return to Grolsch if there is no logistical capacity to pick them up.

The time that the bottle stays at the consumer is dependent on multiple factors. The first factor is the type of beer as some types of beer are consumed faster than others (standard beer vs special beers). Another factor is the time of the year, as in the summer more beer is consumed than in the winter period. A certain percentage of the containers does not return at all, leaving the need for yearly new injections to make up for this "trade loss".

When the consumer returns the bottles and crates, the crates are filled by the workers at the bars, restaurants or supermarkets. Because crates are not filled completely with the right bottles, this results in the following types of pollution in the crates:

- Empty spots
- Unusable bottles
- Bottles that Grolsch needs in other crates
- Bottles from other brewing companies

Some crate types have more pollution than others. Eagle and Swingtop crates (with normal beer) are often purchased by the consumer as a full crate. They are therefore often returned as a full crate as well, resulting in little pollution. However, special beers are usually purchased in smaller quantities and without a crate. When these bottles return, they are put together in a Pinolen or Pelican crate by the workers at the supermarkets, resulting in more pollution.

There is not always a good balance between bottles and crates in the market. During periods of six-pack and gift-pack promotions, when a lot of bottles get sold without a crate, Grolsch may send empty crates into the market to restore the balance. The crates are brought to the client's depot by Grolsch itself.

Grolsch is responsible for returning the bottles and crates from the customers to the brewery. Most of the time a delivery of filled containers to the customer's depot gets combined with taking available empty containers back. But when Grolsch desperately needs a certain container-type for production, Grolsch may actively get containers back from the market by sending empty trucks as well. Some Grolsch bottles end up in crates that go to other breweries and the other way around. Since a few years there is an agreement to trade these "lost" bottles with each other.

When the bottles and crates get back to the brewery, the number of crates of each type is registered together with the date and time. The production codes of the bottles are not registered. Millions of bottles get returned to the brewery each week, so to (manually) register all production codes of the bottles is too much work. It is therefore unknown how long each bottle really stayed in the market as a returned crate can contain bottles from different production batches.

Sorting process

As the returned crates can contain different bottle types, even bottles from other breweries, most crates need to get sorted before they can be used on the filling lines. Grolsch has one automated sorting line that can output around X sorted crates per hour. A few workers at the sorting line make sure the machines correctly sort the correct bottles into a wanted crate type. These workers make sure there are no problems during the sorting process. When a truck arrives with returned crates, the crates can be automatically unloaded from the truck onto the sorting line.

When there is not enough capacity on the sorting line to fulfill the demand of the filling line, containers can be sorted by manual sorting. This is however a time consuming process and should only be used when really necessary. The manual sorting can also get to the output per hour of X crates if there are enough workers available. This manual sorting is costly and inefficient.

Storage of empty containers

If there is no capacity to unload trucks right away, or the types of crates are not needed yet, the unsorted containers are placed in storage in the crate park outside of the brewery. Unsorted containers are always stored outside, except for the crates with little pollution (empty spots, bottles from other breweries). To make it clear for the warehouse personnel to visually see which containers are already sorted, unsorted containers are placed wide (A) in storage and sorted empty containers are placed small (B):

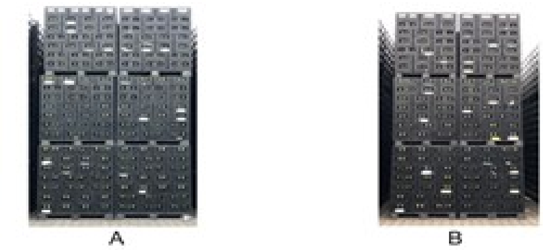

Figure 5: Wide (A) and Small (B) pallet placement

Some crates have more pollution than other crates. Eagle crates are for example relatively “clean”. These crates are not sorted before production but can be used on the filling line right away. The few spots in the crate that are polluted are corrected at the filling line.

There is a limited capacity of 2.378 pallets for empty containers inside the brewery. This area is called “De Hoogbouw” and as a pallet can hold 70 crates (60 for Swingtop crates), the total capacity of this area is 167.160 crates. De Hoogbouw is a storage area inside the brewery where only sorted containers are stored that are ready to use for production. Because of the limited capacity of De Hoogbouw it is important to consider which bottles and crates are stored here.

Sorted containers for which there is no space in De Hoogbouw are stored outside at the brewery or at the harbor of Enschede. Initially it was the idea to store everything inside the brewery, but it is clear that this is impossible looking at the size of the crate park outside. There is a lot of storage capacity outside, but bottles that are stored outside have a change of breaking in bad weather conditions such as frost. Even if they don’t break outside, cold bottles still have a change to break when they are moved into the hot washer on the filling line. These losses are considered the internal losses. Wet bottles also give problems during production, as they are too slippery to grab by the machines. Newly purchased bottles are also stored outside but are generally well packaged against damage. New bottles are not stored in crates, and therefore take less space in storage. New bottles are put on the production line together with empty crates.

From De Hoogbouw the containers are automatically picked up and moved to the production lines. On the production lines the unusable bottles are sorted (as the sorting line is not perfect and some crates were not sorted before) and go into the glass waste bin. Bottles that can be used in other production batches are put in other Grolsch crates. Bottles from other breweries are filtered as well. Because other breweries also filter bottles from Grolsch, since a few years there are arrangements between (some of the) brewing companies to trade these “lost” bottles with each other.

On average a bottle goes through the return process X times before it is (taken) out of roulation. Because of the market loss, production loss and possibly sales growth, every year new containers need to be injected.

2.3. Current long-term container-planning model

To support the decisions regarding the procurement of new containers, Grolsch uses a long-term container-planning model. In this section we describe the inputs and outputs of the model and how the model goes from input to output. In Section 2.4. we describe how the inputs and outputs are currently calculated.

It plans the expected injections per week for each container for about one and a half years ahead. This is also the model that we will focus on during this research. The output of the model is used to distribute the budget across the different container types and to order new containers from suppliers. The container return forecast is currently based on the forecasted weekly sales, the tactical production plan and the input parameters TL, IL, TP, WiT and DoC. The inputs and outputs of the current model are illustrated in Figure 6:

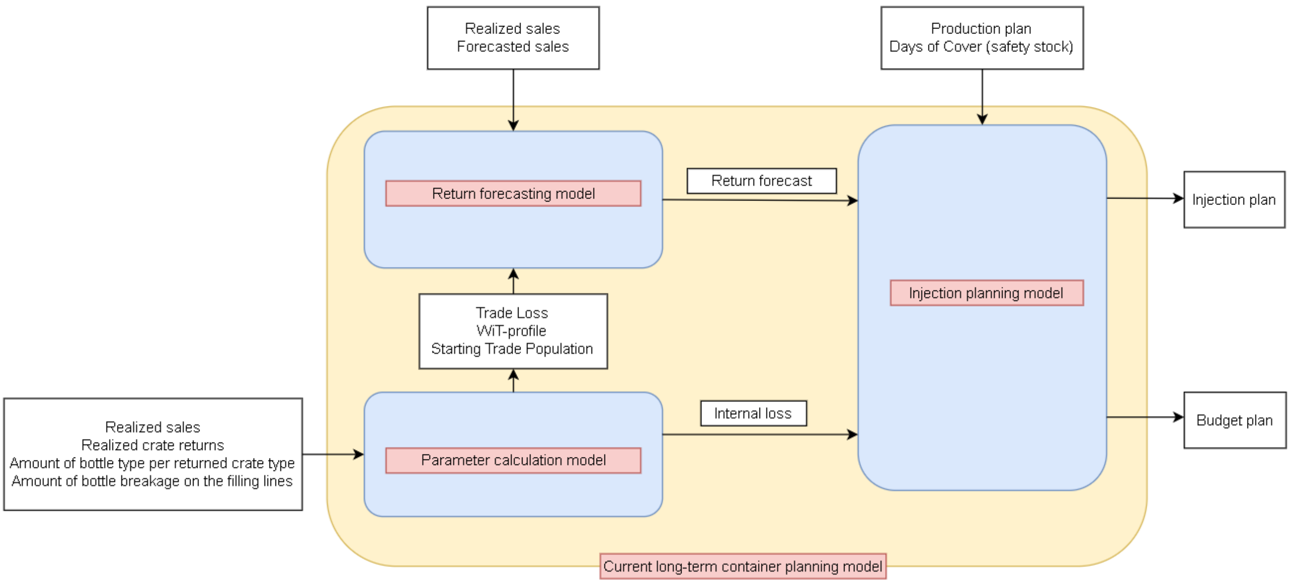

Figure 6: Current container-planning model

2.3.1. Current model inputs

“**Trade Loss**” is the percentage of the sold containers that does not return to the brewery. It is hard to determine when containers are specified as lost, as they can still be returned even after years in the market. Trade loss should be stable over the years. If it is not stable, the events should be identified that could have caused the change in trade loss. How this parameter is currently determined is described in Section 2.4.

"**Internal Loss**". Containers are not only lost in the market, but also at the brewery in outside storage and during production. In outside storage bottles are mainly lost because of bad, cold weather. Glass expands when it gets cold and can also break when put into the hot washer at the filling line. During production there can also be breakage of glass due to machine failure or breakage caused by the workers.

"**Trade Population**" means how many containers are in the market. This is an important parameter for the finance department as the containers remain property of Grolsch while they are in the market. What finally ends up on the balance sheet is the "total population" which consists of empty container stock at the brewery, full container stock at the brewery and the number of the containers in the market. The input for the model is a starting Trade Population, based on an estimation for the starting period of the model.

"**Weeks in Trade**" is the number of weeks of sales that the TP consists of. A WiT of 20 means that the model expects that the sales of the last 20 weeks are still in trade and that all containers sold before 20 weeks ago are either returned to the brewery or lost in the market. WiT should not be confused with Time in Trade (TiT), the time in the market. TiT is currently not used in the return forecast.

"Days of Cover" (Safety Stock)

In the current long-term container-planning model, uncertainties in demand and returns are not directly included in the calculation of the injection planning. The model uses some safety stock to protect against these uncertainties, but the determination of this safety stock is based on experts opinions and it is filled in manually. In current model terms, the safety stock is called Days of Cover (DoC). DoC is the number of working days of planned production that the stock of empty bottles should cover. Most of the year the standard value of DoC is five working days (one production week).

In the summer period, when sales are high, the DoC for Apollo is set to one to three days to limit the amount of injection that the model plans. This does not directly mean that the absolute values of safety stock are lower, as Grolsch also produces more in peak season. Namely, the amount of production during five production days in peak season is more than during five production days in off-season. But it still seems counterintuitive to have the same or lower safety stock in peak season, as that is the period with high volatility in weekly sales volumes. Our feeling is that in this period the safety stock should be higher instead of the same or lower than in other periods of the year. In Figures 7 and 8, we see the current required safety stock can change quite a lot from week to week.

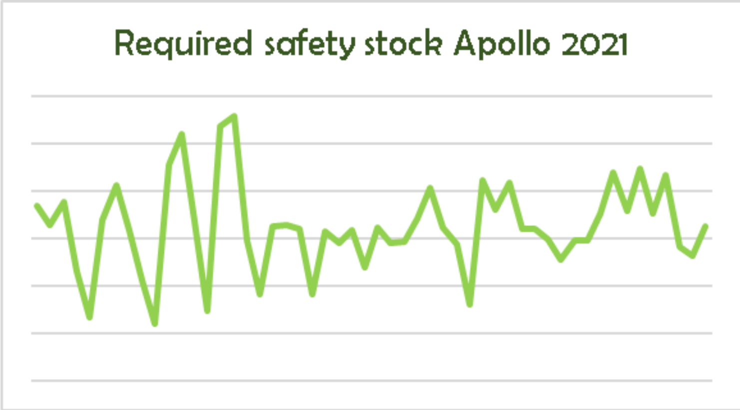

Figure 7: Current safety stock Apollo 2021

Sometimes there is less money for injection available than is needed based on the model's injection planning. DoC is then one of the first variables that is played with. Lowering the DoC by two days for some weeks can significantly reduce the amount of planned injection. This results in a bigger risk of production postponements, especially if the DoC is lowered in periods with high volatility in sales. As these peak periods of sales are the periods where injections are usually planned, the risk of production postponements and stockouts is real.

For BNR the safety stock is sometimes even 0, as there are some weeks for which no production is planned:

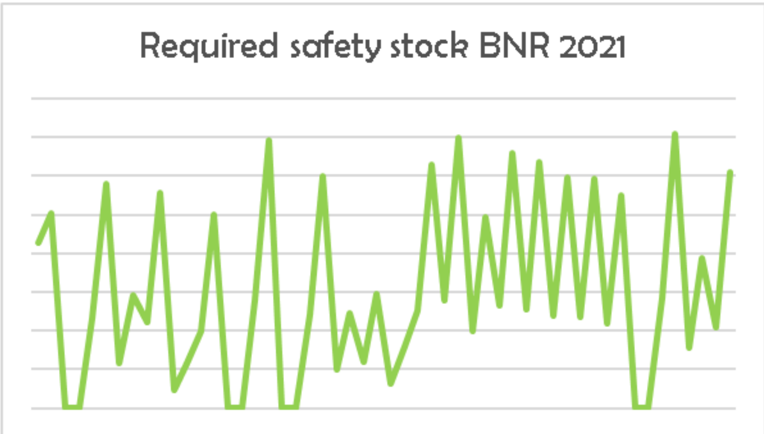

Figure 8: Current safety stock BNR 2021

Sales forecast

The sales forecast is made by the Demand Planning department (DP). Every week the long-term and short-term sales forecasts are sent to the SCP. These forecasts are input for the SCP to come up with production plans. The sales forecast can be uncertain and volatile (especially if measured per week) as Grolsch' sales are mainly based on uncertain price promotions of customers. Every day the DP gives an update to the SCP how the sales are going in comparison with the forecasted sales so the SCP can make changes to the production plan if necessary. The long-term sales forecast is input for the container-planning model.

Production plan

The tactical planners of the SCP come up with a long-term production plan every week. This production plan is based on the long-term sales forecast, inventory capacity, minimum batch sizes and safety stock. In the long-term container-planning model, the production plan is seen as demand that needs to be fulfilled. The safety stock measure DoC is also based on days of production that the empty bottle stock needs to cover.

2.3.2. Current model outputs

In Figure 6 the outputs of the current model are shown: the return forecast, injection planning and budget plan for the next year. First, the return forecast is made based on the input parameters described in the previous section. In this section we describe how the model currently forecasts returns. When the returns are forecasted, the model can make the injection planning. In this section we also describe the current purchasing policy and how this translates to a budget for the next year.

Return forecast

In this part we explain how the model currently forecasts returns. This is based on the parameters TL, WiT and TP. In Figure 9 can be seen that historic WiT-values are found using historic data. Every week of the year has its own WiT value. So, how many previous weeks of sales are present in the market in a certain week. As an example: If in week 10 the previous 3 weeks of sales are still in trade (the sales of week 10, 9 and 8), then the WiT value of week 10 is 3.The assumption is that these values for WiT should be the same in the next year. This means that in week 10 in the coming year, the expected TP also consists of the sales (forecast) of week 10,9 and 8.

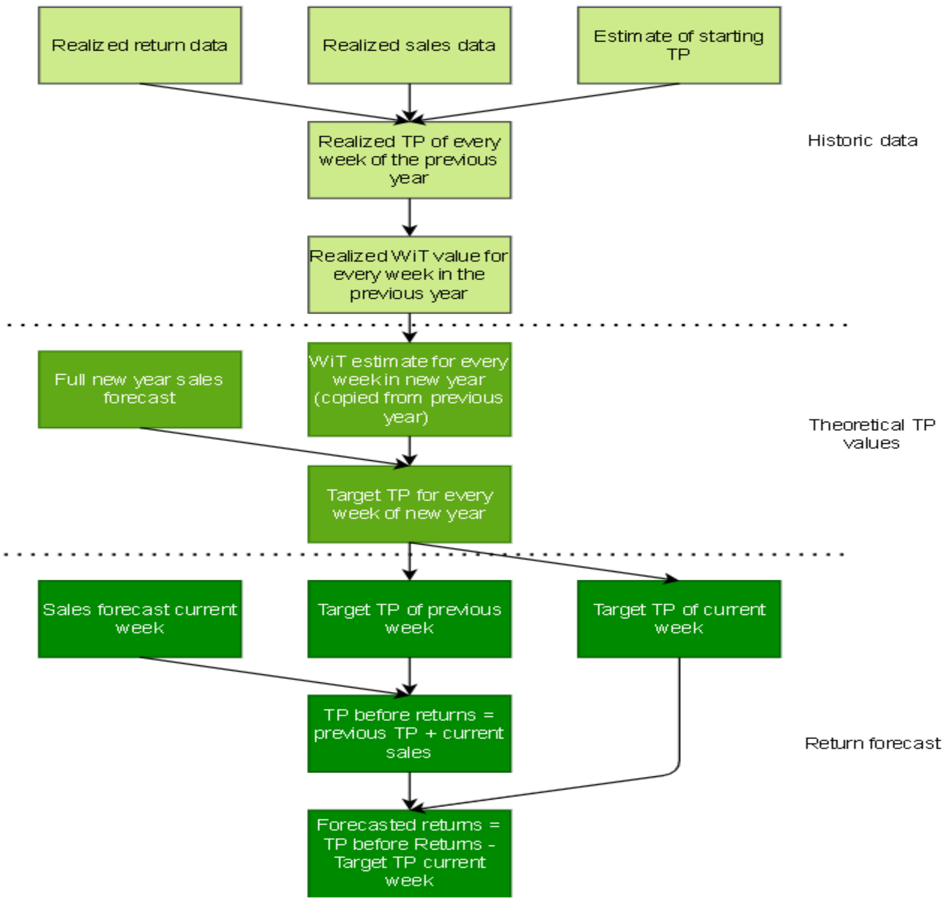

Figure 9: Current return forecast

To obtain the historic WiT-values, first an estimate of an historic starting TP needs to be made. This estimate is currently made based on the average sales, expected average time in the market, estimated losses and realized injections. One the starting TP is estimated for the starting week of the historic data, the historic TP of every week of the previous year can be estimated with the following formula:

$$\textbf{\textit{Trade Population}} = \textbf{\textit{Previous Trade Population}} + \textbf{\textit{Sales}} - \textbf{\textit{Returns}} - \textbf{\textit{Trade Loss}}$$

In Table 1 and the explanation underneath, we show how the WiT-values translate to forecasted returns.

Week	WiT	Sales	Target TP	End TP previous week	TP before returns	Returns	End TP
1	3	300	300	300	600	300	300
2	3	300	300	300	600	300	300
3	2.5	250	250	300	550	300	250
4	2	200	450	250	450	0	450
5	2.5	250	250	450	700	450	250

Table 1: Current forecasting of the returns

Target TP

With the values of WiT and values of sales of previous weeks, an estimate of TP can made for every week. *In blue in Table 1: The TP of week 4 is believed to consist of the previous two weeks of sales, so 200+250=450.* This 450 is the target value for TP for week 4: the expected value of TP based on the historic values of WiT. For every week, the model makes sure that the TP is equal to the target TP by modifying the returns.

TP before returns

So in week 4, the starting TP is the target TP of week 3 (which is the ending TP of week 3). The amount of sales (that are put in trade) of week 4 are then added to the starting TP of week 4 to obtain the TP before returns of week 4:

$$\textbf{\textit{TP before returns}} = \textbf{\textit{Target TP of previous week}} + \textbf{\textit{Sales of current week}}$$

Returns

Because the model makes sure the ending TP of week 4 has the value of the target TP of week 4, the difference between the TP before returns of week 4 and the target TP of week 4 are the expected returns.

$$\textbf{\textit{Expected returns}} = \textbf{\textit{TP before returns}} - \textbf{\textit{Target TP}}$$

It is an interesting method, but as we will see the historic values of WiT cannot just be copied to the next year. The change of WiT over time (also present Table 1) does also not mean that containers are returning faster. This is a hard thing to grasp, and has to do with how the WiT-values are calculated. We explain this in Section 2.4.

Injection plan and budget for next year

The current container injection planning is made by using a simple order-up to policy. For each week in the planning horizon, the planned injections are as big as needed to get the stock on hand to the required level described by the DoC value. The injection plan does not indicate when to order the new containers, but when the containers are needed. The supplier then has to make sure the containers are available at these times. You could say the order is made right after the model is run, and the replenishment lead time is the time from the order (t=0) till when the containers are needed. The budget plan is a direct consequence of the planned injections. However, the amount of money needed to do the planned injections may not available. If this is the case, the DoC is lowered for some weeks in which injections are planned. This results in less planned injections, but the risks of production problems because of bottle unavailability increase.

We have seen that the safety stock can fluctuate quite a lot from week to week for both Apollo and BNR. What service level is implied with the current values for DoC (and thus the current injection plan) is not known and is worth investigating. In Chapter 6 we calculate what Grolsch currently implies as a service level.

2.4. Current input parameter calculation

One of the main goals of this research is to improve the input parameter calculation of the current model. In this section we therefore look at how the parameters of the current model were calculated the last time, which was in 2019.

“**Trade Loss**”

First Grolsch determined how many of Eagle- and Pelican crates returned to the brewery in the years 2016, 2017 and 2018. These are the unsorted crates. The number of crates is multiplied by 24 bottles to obtain the total number of (potential) bottles returned. This is the number of bottles returned before sorting so pollution is not taken into account yet. This is a questionable method because returned crates are not always sorted right away and sorting input (and sorting losses) can differ a lot from week to week.

As Eagle crates are not sorted on the sorting line, the sorting loss for the Apollo bottle only includes the sorting loss of pelican and other crates. The sorting data includes: amount of empty positions, amount of other bottles used by Grolsch and amount of bottles from other breweries.

After the sorting line losses per week are subtracted from the returns per week, the inline sorting losses are subtracted as well to find the realized returns per week. Inline sorting is the sorting that is done on the filling lines. Crates that are sorted on the filling lines are Eagle and DeKlok crates, as these crates have generally very little pollution. The amount of total returns finally is determined by taking the sum over all weeks included in the Trade Loss calculation. Finally, the TL is determined by the following formula:

$$\boldsymbol{Trade\ loss_{i,y}} = \frac{\boldsymbol{Total\ sales_{i,y} - Total\ returns_{i,y}}}{\boldsymbol{Total\ sales_{i,y}}}$$

with:

- i = container type and y = year(s)

A period of three years is used to determine Trade Loss. As a rule, the minimum period to calculate the TL over is one full year. The assumption is that TL is stable during the year and shows no seasonality.

"Internal Loss"

The breakage of the bottles on the filling lines and the amount of breakage in outside storage is the amount considered as "Internal Loss". Now, this loss is determined by dividing the amount of breakage by the total filling line input. These data are taken from the inline sorting data, that the Warehouse departments sends to the SCP every week.

"Trade Population"

As mentioned in the model explanation, the TP is the amount of containers in the market. In the last parameter calculation, the TP can be calculated in three different ways:

Formula 1:

$$\boldsymbol{Trade\ Population = Previous\ Trade\ Population\ + Sales - Returns - Trade\ Loss}$$

Formula 2:

$$\boldsymbol{Trade\ Population = Previous\ Trade\ Population\ + Injections - Trade\ Loss}$$

Formula 3:

$$\boldsymbol{Trade\ Population = WIT * Sales}$$

Grolsch is not sure which of these formulas is the best, but Grolsch uses Formula 1 in its current model. An interesting thing to notice is that Formula 3 includes WiT while WiT is based on Formula 1. It is also clear that with using Formula 3 the TP is very unstable as the TP varies greatly from week to week. This may be the case because the method is in no way based on the previous trade population.

"Weeks in Trade"

As mentioned in the model explanation, WiT is the amount of weeks of sales that the TP consists of. It is one of the most important parameters in the current model as it is currently used for the timing of the container returns. In the current parameter calculation, the WIT estimation is made based on the TP (method 1) and the sales. WiT is determined for each week of the year. The most important thing to mention is that WiT is not the same as TiT, because WiT does not say how long containers stay in the market. WiT and TiT are confusing terms and are easily mixed up.

Seasonality of WiT

When you see the WiT-values change over time, the feeling is that this means the time in the market is assumed to be seasonal. But that is not necessarily the case. Recall that WiT is not the time in market, but the amount of weeks of sales that the TP consists of. In the example below we show that WiT is heavily dependent on differences in sales from week to week, and that the values will change even when the time in the market (TiT-distribution) stays the same.

After week	Return %
1	20%
2	40%
3	40%

Table 2: TiT-profile

Sales	Returns	TP	WiT
10	10	50	5
10	10	50	5
10	10	50	5
10	10	50	5
10	10	50	5
10	10	50	5
10	10	50	5

Table 3: WiT-calculation with constant sales

Sales	Returns	TP	WiT
10	10	50	5
100	10	140	5
10	28	122	3.2
10	46	86	2.66
10	46	50	3.2
10	10	50	4.1
10	10	50	5

Table 4: WiT-calculation with sales peak

In Table 3, the TP keeps being the sum of 5 weeks of sales. In Table 4 the TP changes because of the change in sales (10→100→10). Note that the TP of 140 is still 5 weeks of sales (100+10+10+10+10). When the returns of these 100 sales come, the WiT is going to change even if the TiT-distribution stays the same.

Because this WiT parameter changes so much with sales differences, a WiT-profile cannot just be copied to a new year to obtain an accurate target TP-estimate. In the new year, sales peaks may come in different weeks. We will see in Section 2.5 that the return forecast is inaccurate, mainly because of the usage of the parameter WiT.

2.5. Performance of the current model

Before we make improvements to the parameter calculation, it is important to know how accurate the current model is using the current parameters. In this section we will discuss the performance of this current model in the past few years.

2.5.1. Key Performance Indicators (KPIs)

The goal of the current model is to provide an injection plan of new containers for the next budget year. As lead times on new containers are long and suppliers need to know Grolsch' injection plan long in advance, the budget is made around May-June for one and a half years ahead. We are interested in how accurate the injection planning was in the last years. However, we don't use this injection accuracy directly as a Key Performance Indicator (KPI), as the injection planning is a decision based on the return forecast. The return forecast accuracy is therefore the main KPI for this research, as a high accuracy of the return forecast will translate to an accurate injection planning.

Uncertainty in the sales forecast (not considered influenceable in this research) can contribute to inaccuracy of the return forecast. We therefore let the current model forecast returns based on forecasted sales data as well as on realized sales data. With the realized sales data, the forecasted returns by the current model should in theory be (almost) the same as the realized returns. If this is not the case it is says something about possible improvements to the current model's- and current input parameters' validity. To conclude, we measure the performance of the current container-planning model by the following KPIs:

1) *Accuracy of the return forecast (based on realized sales data)*
2) *Accuracy of the return forecast (based on forecasted sales data)*

To measure the accuracy of the forecasted returns, we use the Mean Absolute Percentage Error (MAPE). The MAPE is a widely used method for determining accuracy of forecasts (Silver, Pyke & Thomas, 2017). The formula for the MAPE is:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|$$

with:

- A_t = the realized values at time t
- F_t = the forecasted values at time t
- n = the amount of observations

The MAPE gives only the absolute errors and does therefore not say something about under- or over-forecasting. Thus, we also want the (non-absolute) percentage errors to see if the model is generally under-forecasting, over-forecasting or a mixture of the two. This is important as there are different costs involved for under- and over-forecasting. It also gives an indication if the used values for TL are in the right range. An important note: under-forecasting container returns will lead to planning too much injection.

2.5.2. Calculation of the realized returns per week

To measure the return forecast accuracy, we first need to determine the realized returns. Obtaining these realized returns per container type is not an easy task. Returns are only registered on crate level. So, the only thing that is registered at return is how many crates of each crate type are returned at which date. At the registry stage, it is still unclear how much bottles of each bottle type are present in these returned crates. In the current parameter calculation, the realized returns per week are calculated by taking the crate returns of the specific week and subtracting the sorting losses of that week. The problems with this approach is that not all crates are sorted right after they return. So the sorting loss of a certain week is actually the sorting loss from crates that returned weeks earlier. And, in some weeks more crates are sorted than in other weeks, resulting in differences in sorting losses between weeks. This will become clear in next the part: “sorting delay”.

Sorting delay

The returned crates are not always sorted right away. The sorting output per week can therefore not be taken as the realized returns per week. Also the current approach of subtracting the sorting losses per week from the crate returns is considered inaccurate. Figure 10 and Figure 11 shows the delay between the unsorted crate returns and when these returned crates are input to the sorting line. In some weeks no sorting is done at all, and in other weeks sorting peaks exist for a specific crate type. Using this sorting loss per week results in spiky returns per week. Because these (spiky) realized returns are used to calculate the WiT- and TP values, it is the question how accurate the WiT- and TP-estimates currently are.

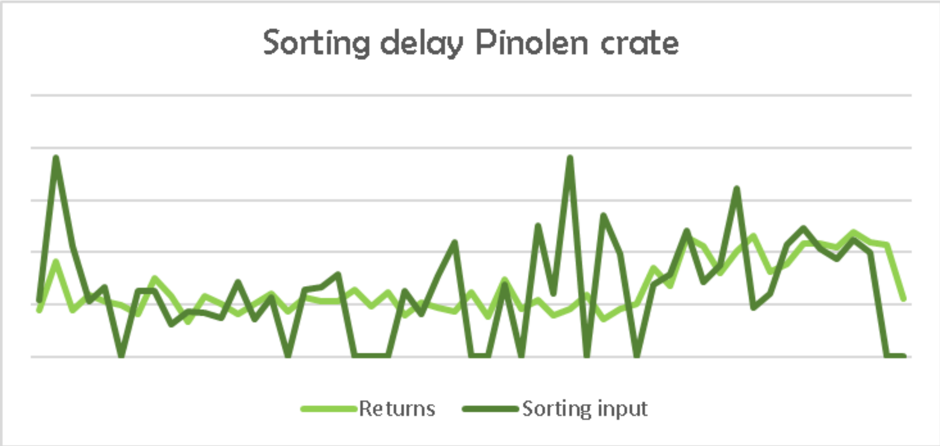

Figure 10: Sorting delay Pinolen

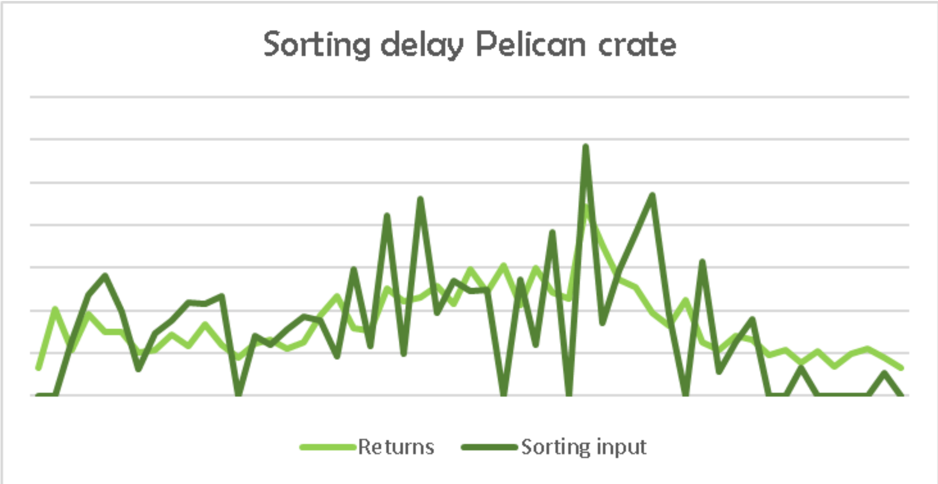

Figure 11: Sorting delay Pelican

Sorting percentages

Instead of the current method, we calculate sorting percentages (for each year) that say how many of each bottle type is present in each crate type. The crate returns per week can be multiplied by these percentages to obtain the realized container returns per week. The formula to calculate these sorting percentages is:

$$\boldsymbol{Sorting_{b,c,y}}\ \% = \frac{\mathbf{24} * \boldsymbol{Sorting\ input_{c,y}}}{\boldsymbol{Sorting\ output_{b,c,y}}}$$

with:

- *b* = the bottle type
- *c* = the crate type
- *y* = the year

The sorting input is measured in amount of crates and therefore needs to be multiplied with 24 so we have the amount of bottle spots. The sorting output is the amount of bottles that were actually present in the returned crates. The standard of the sorting process is that it should be 99% accurate.

Filling line sorting

As mentioned before, not all crate types get sorted on the sorting line. And even the crates that do get sorted are not always sorted 100% correctly. The pollution that (still) exists in the crates that go to the filling line (empty spots, foreign bottles and unusable bottles) is removed there. Unusable bottles go into the glass bin. Incorrect bottles (bottles of the wrong type) that can be used during other production batches are generally filtered out and stored, as long as the amount of filtered bottles can be handled. Sometimes the filling line cannot handle the amount of filtered bottles. In that case not only the unusable bottles go into the glass bin, but also the filtered reusable bottles to prevent line breakdown.

The input of the filling line consists of returned containers and injected containers. Injected containers can be considered as "sorted" 100% correctly. The filling line sorting loss percentage should therefore be calculated by using only the returned container input, and not the full input. To calculate the percentage correctly we use the following formula:

$$\boldsymbol{Filling\ line\ sorting\ loss_{i,y}}\ \% = \frac{\boldsymbol{Total\ input\ returns_{i,y}}}{\boldsymbol{Filling\ line\ sorting\ loss_{i,y}}}$$

We end up with a filling line sorting loss percentage per container type for each year between 2017 and 2021. The assumption here is that all sorted returns are used for production. This seems a fine assumption as every year the returns alone are not enough to fulfill the demand.

Realized returns per week

To end up with the realized returns per week, we first take the unsorted crate returns and multiply with the sorting percentages of the specific bottle from the type of crate. We then add up all incoming bottles from the specific bottle type from all different crates. Then we multiply this outcome with the filling line sorting loss percentage to end up with the realized returns per week:

$$\boldsymbol{Realized\ returns} = \boldsymbol{Crate\ returns} * \boldsymbol{24} * \boldsymbol{Sorting}\ \% * \boldsymbol{Filling\ line\ sorting\ loss}\ \%$$

2.5.3. Return forecast accuracy

For determining the accuracy of the forecasted returns, we use the realized sales and calculated realized return data from 2017 onwards.

Apollo

We start with analyzing the performance of the current model for the Apollo bottle. The Apollo bottle is the standard green 30cl Grolsch bottle and is the most important bottle in terms of production, sales and budget for new injection. For the Apollo bottle, we calculated the realized returns as outlined in the procedure in Section 2.5.2. In Figure 12 the comparison between the realized returns and the forecasted returns is shown. Because the forecasted returns are here based on realized sales data, the two lines in the graph should in theory be close to each other (the uncertainty of the sales forecast is taken out of the equation). The difference between the lines is in theory influenceable by adjusting the input parameters and the way the model forecasts the returns.

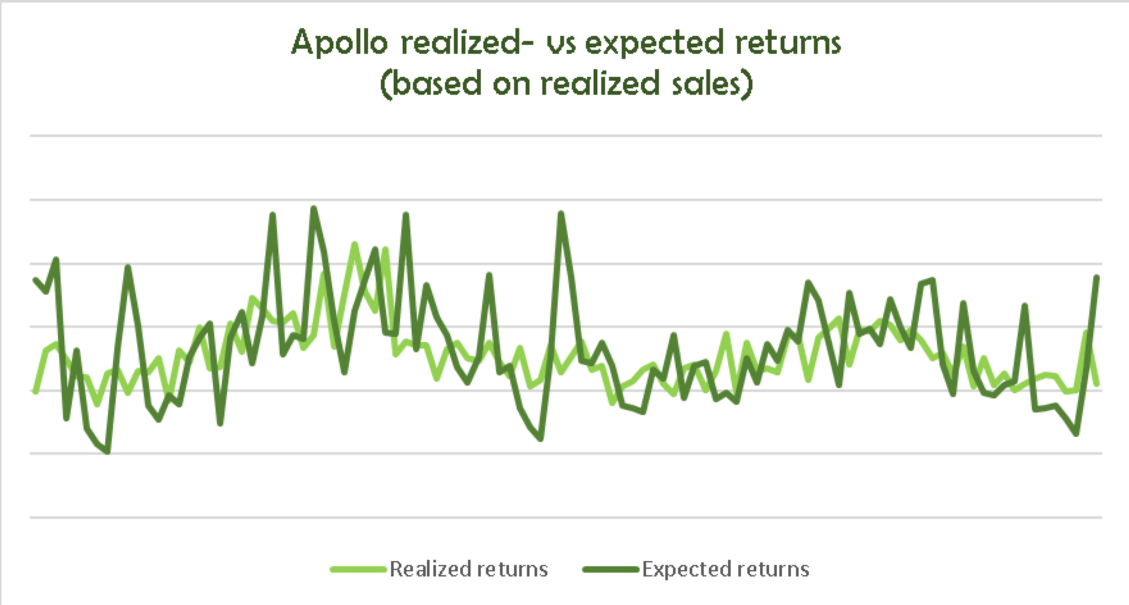

Figure 12: Apollo realized returns vs expected returns

Year	Expected hl	Realized hl	%	Trade loss used	MAPE
2018	X	X	X	X	29%
2019	X	X	X	X	21%
2020	X	X	X	X	26%
Total	X	X	X	X	25%

Table 5: Apollo current return forecast accuracy

The expected returns are not in sync with the realized returns considering the total amounts per year. Even though the error percentages are smaller than X%, this is an indication that values for TL should be higher, and the error of the total returns per year should not be higher than X%. Also recall that these return expectations are based on the realized sales.

The next thing to notice is that the expected return line is less smooth than the line of the realized returns. A week of few forecasted container returns is often followed by a week of many forecasted container returns. This is indication that the input parameter WiT is not yet optimized as we expect a smoother line even when the values of TL are not optimal. We notice seasonality in the returns with a peak around weeks X and less returns during the autumn- and winter months, highly correlated with respectively the higher consumption of beer in these periods.

A high return forecast error per week does not immediately translate to lost production, stock-outs or lost sales. If the under-forecasting for one week is accounted for by over-forecasting for the next week, the two-week total is still accurate. In practice it can be assumed that the planning can be changed up if the two-week totals are accurate. However, problems occur when the under- and over-forecasting gets out of hand, even when the two-week totals are accurate.

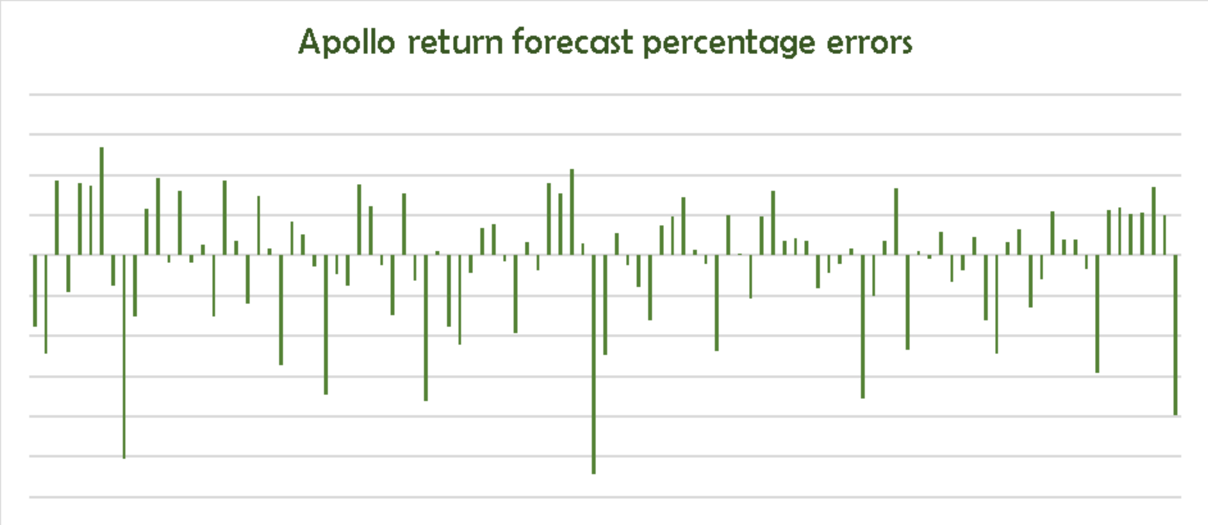

Figure 13: Percentage errors return forecast

In Figure 13 we show the percentage errors of the return forecast for Apollo. The main thing to take away from this graph is that the under-forecasting (percentages > 0) of one week is not directly compensated by over-forecasting (percentages <0) in the next week. We notice many consecutive weeks of either over- or under-forecasting. This poses problems for the container-planning. If the percentage errors over two-week periods were minimal, SCP would have some flexibility to change the planning. Because the current model under- or over forecasts returns for consecutive periods, the injection planning is likely to be less accurate.

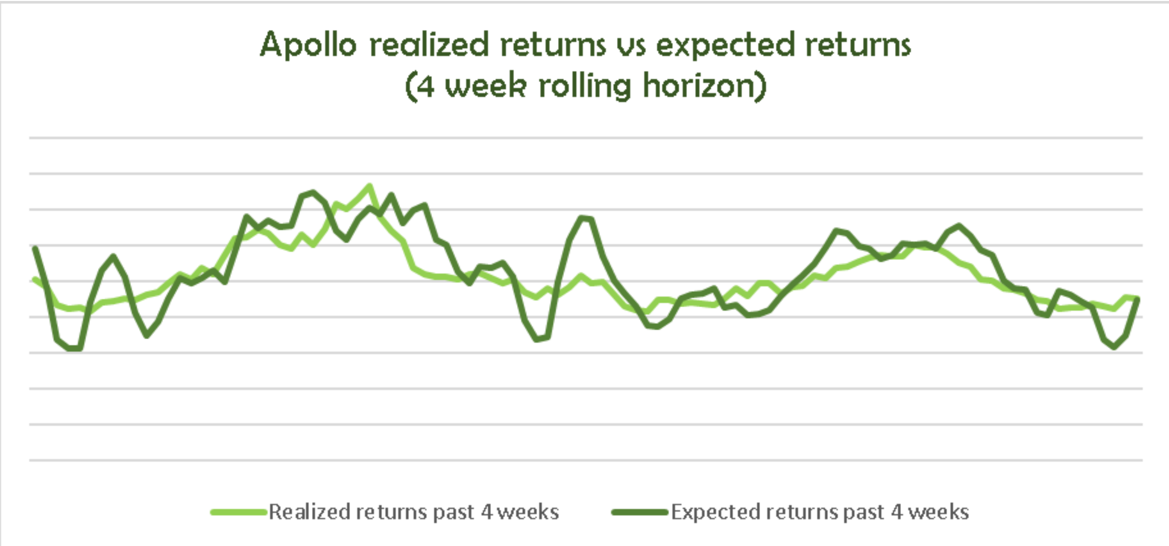

Figure 14: 4 week sum - Apollo realized returns vs expected returns

We also take a look at the case with a 4 week rolling horizon, which is shown in Figure 14. The line is more smooth but the same pattern is present as before.

The MAPE of this period (2018 and 2019) is X% per 4 week sum, which seems high as the peaks and lows are still not smoothened enough over a period of a month. This X% of the amount of hectoliter (hl) that is returned is a large amount and the money involved in the injections of these amounts cannot be disregarded. The error should be close to 0% as the model should forecast the returns accurately when based on realized sales data (especially over longer periods like four weeks).

The feeling is that the MAPE is too high and can likely be improved by calculating the input parameters in a better way and changing the way the model forecasts returns.

To see the impact of the sales forecast error on the expected returns we now use the forecasted sales instead of the realized sales. Figure 12 shows the graph of the expected returns based on realized sales of 2019, and Figure 15 shows the graph based on forecasted sales.

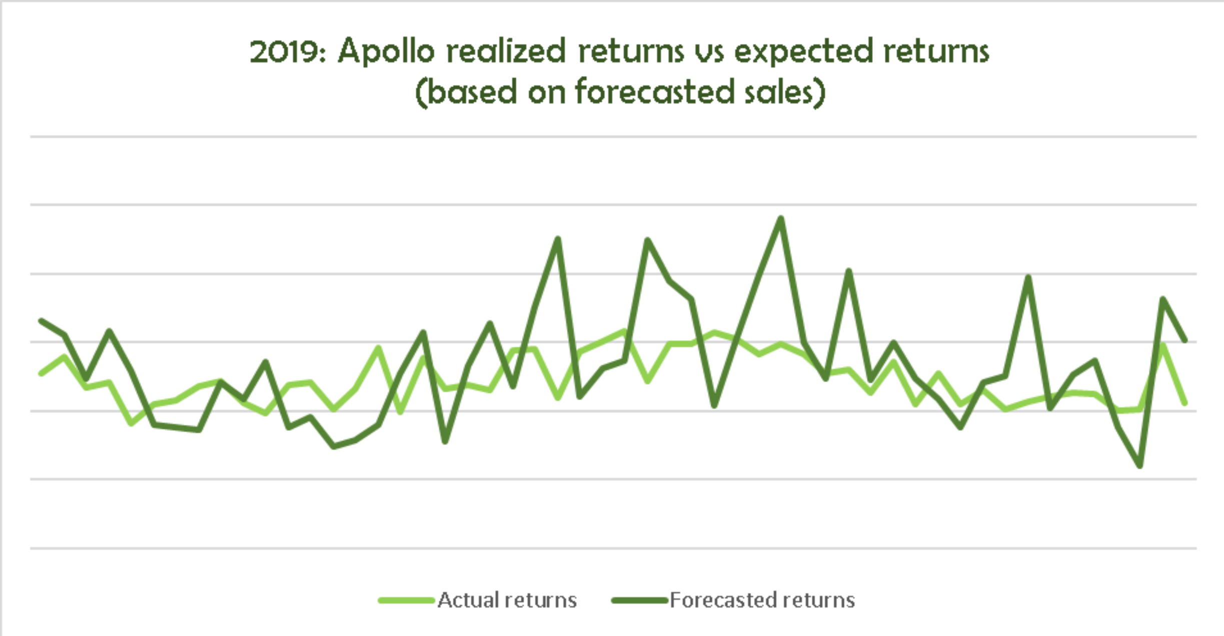

Figure 15: Apollo returns based on forecasted sales

We are convinced that the error of the forecasted returns compared to the realized returns will get a lot smaller when the error of the expected returns based on realized sales compared to the realized returns is minimized. The return forecast based on forecasted sales has a slightly lower accuracy, but the shape of the graph is similar to the graph of the forecasted returns based on realized sales. We therefore think the improvement based on realized returns will also translate to improved performance based on forecasted sales.

BNR

The returns of the BNR are harder to forecast as BNR is a bottle shared with other breweries, which has an impact on the loss in the market. Some years the loss in the market is big (X%) and some years the loss in the market can even be negative, when Grolsch gets more BNR back than was sold. Overall we see the same pattern as we have seen with Apollo. We again notice decent accuracy (apart from 2019) of annual returns, but high inaccuracy of the timing of these returns:

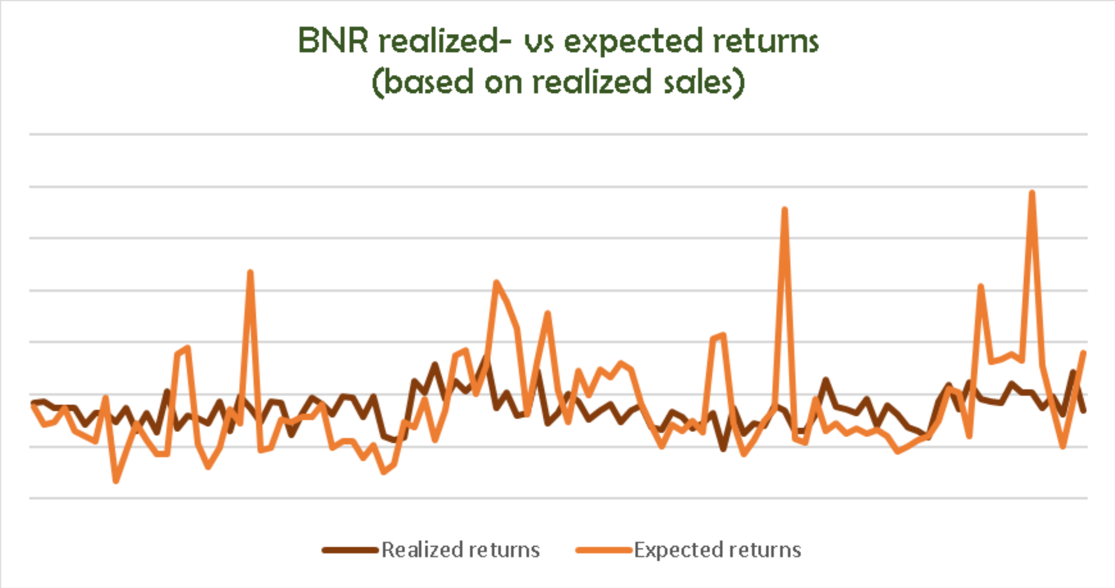

Figure 16: BNR realized returns vs expected returns

Year	Expected hl	Realized hl	%	Trade loss used	MAPE
2018	X	X	X	X	41%
2019	X	X	X	X	40%
2020	X	X	X	X	43%
Total	X	X	X	X	44%

Table 6: BNR current return forecast accuracy

The expected returns are based on X% TL in the market for 2018 and X% for 2019 and 2020. These TL values are checked with the supply chain manager of Grolsch. This lead to the drop in the TL for 2019 and 2020. Looking at the realized returns in 2019 compared to the expected returns, we conclude that the loss percentage drop from X% to X% might have been too big. For 2020 the X% looks a better estimate and Grolsch expects the loss percentage in the market to slowly rise again for the coming years. The peaks in expected returns can be explained by the fact that the model expects a week of high sales to return almost all at once (Figure 17).

Figure 17: BNR expected returns vs sales (confidential)

We notice that a peak in sales is often followed by a spike in the expected returns. As we have seen, the expected return line should be more smooth. We again question the usage of the parameter WiT, which makes the return forecast very receptive to spikes in sales.

2.5.4. Accuracy of the planned injections

As the goal of the current long-term container-planning model is to provide an estimate of the amount of injection needed for the year ahead, it is important to see if the model's expectations in the past years are in line with what Grolsch actually injected. Injection means Grolsch buys the bottles from the supplier's stock and becomes owner of the stock.

Apollo

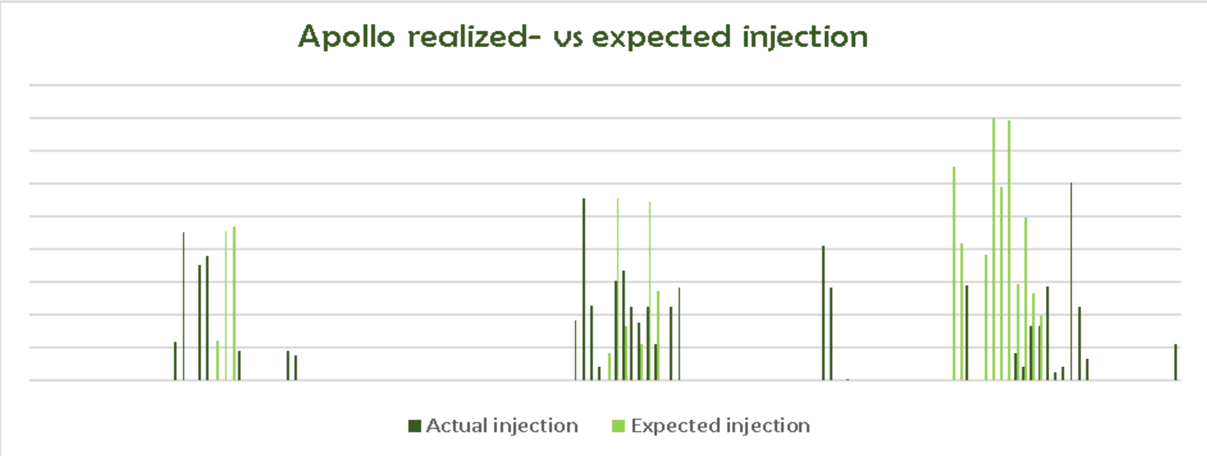

Figure 18: Apollo realized injection vs expected injection

Year	Expected hl	Realized hl
2017	X	X
2018	X	X
2019	X	X
Total	X	X

Table 7: Apollo expected injection vs actual injection

The first thing to notice is that the timing of the realized injections is roughly in line with the timing of the expected injections. Realized injections are often done in weeks X to X, which is peak season for the Apollo bottle. The amount of injection is a different story as the model expected only X% of the realized injections over the course of 3 years, which is a difference of roughly X euros. For the past years it was possible to redistribute the budget over different containers during the year, so more Apollo injection could be made than were expected in advance. Grolsch expects that this redistribution is unlikely to be possible in the same way in the coming years, so it is important to get the budgeting right the first time.

For 2018, the realized production and sales were X% lower than forecasted back in 2017. In this case we would expect that the amount of realized injection is also lower or in line with expected injection. The model expected around X hl injection in 2018, but Grolsch actually injected around X hl. When we let the model forecast injections in 2018 based on realized production and sales data, the model suddenly expects X hl injection. The timing of production is dictating the injection planning, together with the measure of safety stock DoC. Even if the total production over a year is less than another year, there could be still more bottles needed if there is a lot of production planned in the peak period.

If we then fill in the realized returns instead of letting the model forecast returns based on the realized sales data, the planned injection becomes 0. This again shows the importance of the timing of the returns, as the difference in total hl is minimal while the difference in expected injections is significant. It is strange that the realized returns were around X hl higher than expected, while Grolsch injected more than the model expected. In this case we would expect Grolsch to inject less than the model initially planned as Grolsch got more back than expected.

For 2019 it is a different story. The total Apollo production (filling) was X% below what was expected a year before, but the amount of realized injection is higher than the model forecasted. Even if we base the injection forecast on realized production sales data for 2019 the model still only predicts around X hl injection. The difference with the X hl realized injection is for a big part (X hl) additional production at the supplier and Grolsch decided to inject even when it was not needed (yet). We conclude that this injection was not needed based on the fact that the model does not plan injection when we remove the X hl realized injection. If these X hl were needed, the model would have planned them after we removed X hl incoming injection.

One reason for additional injection could be that there are bottles available at Grolsch, but in a different type of crate than that is needed for production. So if Apollo is needed in Eagle crates, but is only present in Pelican crates, the model says there are enough bottles but in practice Grolsch needs to repack these bottles in different crates which is a time consuming and cost inefficient procedure. Grolsch can therefore decide to already inject new additional Apollo in Eagle crates, despite the investment costs.

In 2020 less injection was needed than was forecasted because of Covid-19. The total production and sales in the first 8 months were X% lower than expected. And the more bottles Grolsch expects to sell, the more bottles are expected to be lost in the market (and need to be replaced). When we take the injection planning of 2020 based on realized production and sales data, we see that the model forecasts the injections with good accuracy: X/X = X% . This means that the model would have calculated the injections quite well if the production and sales forecast for 2020 would have been accurate.

BNR

For BNR we see that the total amount of realized injections in the last two and a half years is lower than the total amount of expected injections. This can be the consequence of injecting more Apollo than planned and therefore distributing less money to BNR. The experience is that BNR does not have a good availability in some periods of the year. Recently, the planning has been changed up quite some times because there was simply too little BNR available and Grolsch actively had to get BNR back from the market. It is unknown to how much backorders and lost sales this has led, but at least there have been production inefficiencies. Batch sizes were lowered and changeover costs rose because of the unavailability of the BNR bottle. A good thing to mention is that an out of stock at Grolsch not directly means that there is an empty shelf at the customer.

The overall feeling is that more BNR injections should be planned so more initial budget can be spent on BNR. Another reason for the gap between realized- and expected injections could be that the TL of 2018 was estimated too high with X%. The model expected that much BNR bottles needed to be replaced, while this was not the case (explaining X hl of the X hl difference).

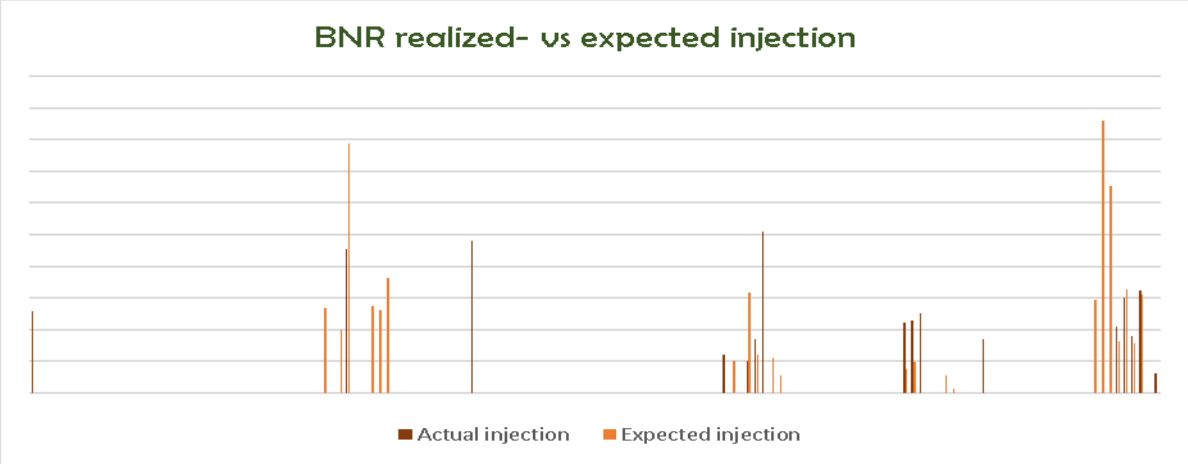

Figure 19: BNR actual injection vs expected injection

Year	Expected hl	Realized hl
2018	X	X
2019	X	X
2020	X	X
Total	X	X

Table 8: BNR expected injection vs actual injection

Supplier flexibility

In this section we have seen that the realized injections can deviate from the expected injections. The expected injections are done 1 to 1.5 years before the bottles are needed, and apparently Grolsch can update this forecast up to 6 months before the bottles are needed. The order lead time of new containers is in general very long (several months). The production plan of Grolsch' glass suppliers is often full. Therefore there are very few changes possible when Grolsch gets the injection planning wrong. For Apollo there are no more changes possible from six months before the bottles are needed.

For BNR there is some flexibility with the supplier. Because BNR is a shared bottle, the supplier can sell bottles allocated to Grolsch to other breweries and the other way around (for example if one brewing company over-forecasts and another under-forecasts). The limit of extra available BNR for Grolsch is around X hl, as this was tested last year when Grolsch had problems with BNR availability. In that period production batches were postponed because there were simply too few BNR bottles available.

The general feeling is: If there is little possible with suppliers, why does Grolsch not keep a lot of safety stock of empty bottles when the depreciation costs are minimal and there is enough storage capacity? Grolsch does simply not want to invest in new bottles when they are not expected to be needed as the money can be used for other important production related investments. It is necessary to make the injection planning as accurate as possible to avoid costs for production postponements, holding costs and costs for injecting less than forecasted.

2.6. Conclusion

We have started this chapter by explaining the term “container” and showed which containers we focus on in this research. We then described the return process from when the filled beer leaves the warehouse to when the returned containers can be used again for production. Next we explained the models that are used for forecasting the returns and planning the procurement of new containers.

We explained the input parameters that the model requires (TL, IL, WiT, TP and DOC) and described the steps and calculations the model uses to get to the output: the planned injections for the upcoming year.

The goal of this research is to improve the performance of the injection planning model, so to set a standard we analyzed the performance of this model in the past few years.

Planned injection can deviate from realized injection for many reasons:

1) Too much realized injection to prevent repacking to other crates
2) Too much realized injection because supplier has too much stock
3) Too little realized injection because of budget constraints
4) Sales forecast uncertainty

Because planned injection is a decision based on forecasted returns, we use the accuracy of the return forecast as Key Performance Indicator (KPI). The Mean Absolute Percentage Errors (MAPE) per week of the return forecasts are considered high with 25% for Apollo and 40% for BNR. The conclusion is that the accuracy of the return forecast can significantly be improved as the returns are forecasted based on realized sales. This means based on the realized sales, when does the model expect returns? In this way, the inaccuracy of the sales forecast is let out of the equation and the inaccuracy of the return forecast is purely based on the return forecasting model. Especially the timing of the returns, which has a big impact on when to do injections, can be improved. This holds for Apollo and BNR.

In the next chapter we review the literature on reverse logistics to find possibilities for improving the return forecasting model.

3. Literature review

The main conclusion of the previous chapter is that the injection planning is heavily dependent on the weekly returns, and that the forecast of these weekly returns is inaccurate. The focus in this research will therefore lie on container parameter estimation and improving the return forecast to indirectly improve the injection planning. In this chapter, we start with a short introduction to reverse logistics and then present a literature review on return forecasting in general and container return forecasting in particular. Other interesting literature sources that are included in this chapter are about calculation of return parameters such as time in the market, trade loss and trade population. In the end of this chapter we give theoretical background on the most promising method for this research.

3.1. Reverse logistics

Before we move to return forecasting, we give some background information on the research field reverse logistics. Reverse logistics is about "activities associated with the handling and management of equipment, products, components, materials or even entire technical systems to be recovered" (De Brito and Dekker, 2002). The research field of reverse logistics is still widely unexplored (Mahmoudi and Parviziomran, 2020), but more and more companies try to become sustainable enterprises that reduce their environmental impact with their activities. In the coming years companies (and researchers) are therefore expected to put more emphasis on reusing products. With optimizing the return process of these products "efficient resource utilization and environmental protection could be achieved" (Senthil & Sridharan, 2014). Reverse logistics provides a "fertile and attractive research area in the field of operations management" (Mahmoudi and Parviziomran, 2020). This operations research field of reverse logistics and in particular reusable packaging is split into four areas in a literature review by Mahmoudi and Parviziomran (2020) (Figure 20):

1. Inventory management
2. Scheduling and routing
3. Procurement and repairing
4. Performance measures

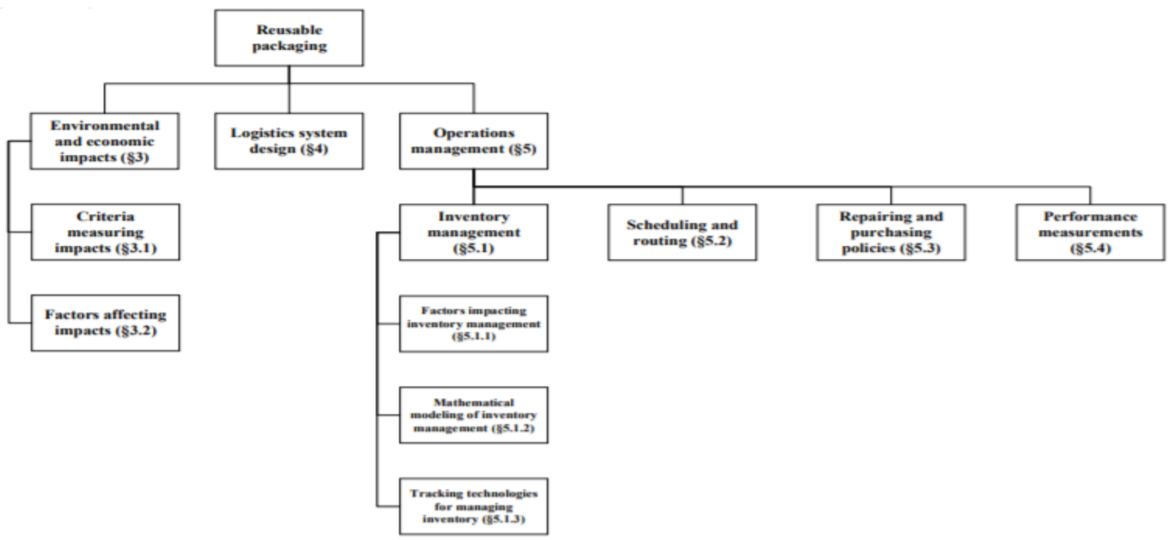

Figure 20: Research fields reusable packaging (from Mahmoudi and Parviziomran, 2020)

From these four areas, particularly procurement and inventory management (including return forecasting) are interesting for our research problem. It is stated by Toktay, Van der Laan and De Brito (2003) that forecasting returns is an essential part in optimizing the reverse logistics supply chain and in this research we focus on this forecasting of returns (which is a part of inventory management), purchasing policies (procurement) for new containers and performance measures as circulation time and asset utilization.

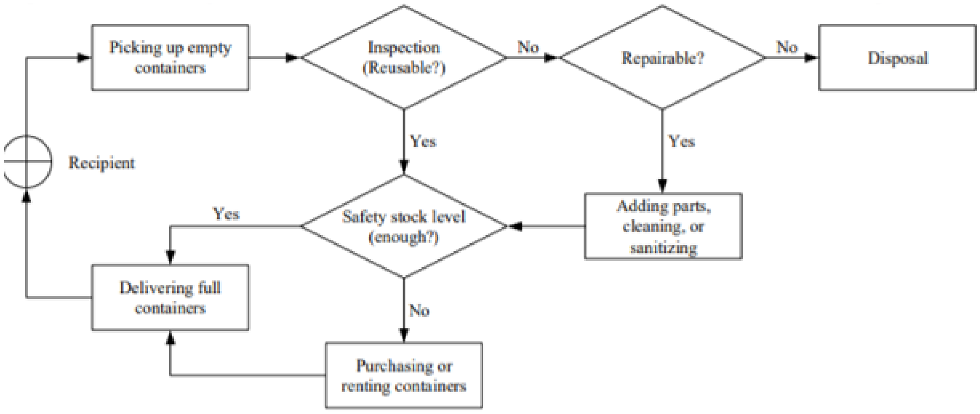

Figure 21: Reverse logistics process (from Mahmoudi and Parviziomran, 2020)

3.2. Return forecasting

Return forecasting versus traditional sales forecasting

The first thing to note is that return forecasting is very different from sales forecasting as returns are not only dependent on its own past values but also on another variable's past values: sales. This means returns should in principle be forecasted based on historic sales data in combination with historic return data, rather than only on historic return data (Kiesmüller and Van der Laan, 2001).

Kiesmüller and Van der Laan (2001) modelled an inventory system in a reverse logistics supply chain with returns dependent on sales with a fixed lag as well as with returns independent of sales. They conclude that modelling returns independent of sales gives a worse performance than letting the returns depend on sales. In literature multiple standard forecasting techniques based on solely return data, like moving average models, exponential smoothing and Holt's method have been found inaccurate for forecasting returns ((Ma & Kim, 2016);(Krapp, Nebel & Sahamie, 2013b);(Geda & Kwong, 2019)). Besides, these methods often focus on updating parameters (after observations of realized returns) which makes them less suitable for long-term forecasts. The augmentation of the data on realized returns cannot be utilized for long-term forecasts (Toktay, Van der Laan and De Brito, 2003) that are only made a few times per year mainly to the determine the procurement budget. Toktay, Van der Laan and De Brito (2003) also mention that for short-term forecasting it does make sense to use these methods with updating as they can utilize the augmentation of new data that come available each week. A nice overview of the forecasting modeling process is given in the book of Silver, Pyke and Thomas (2017):

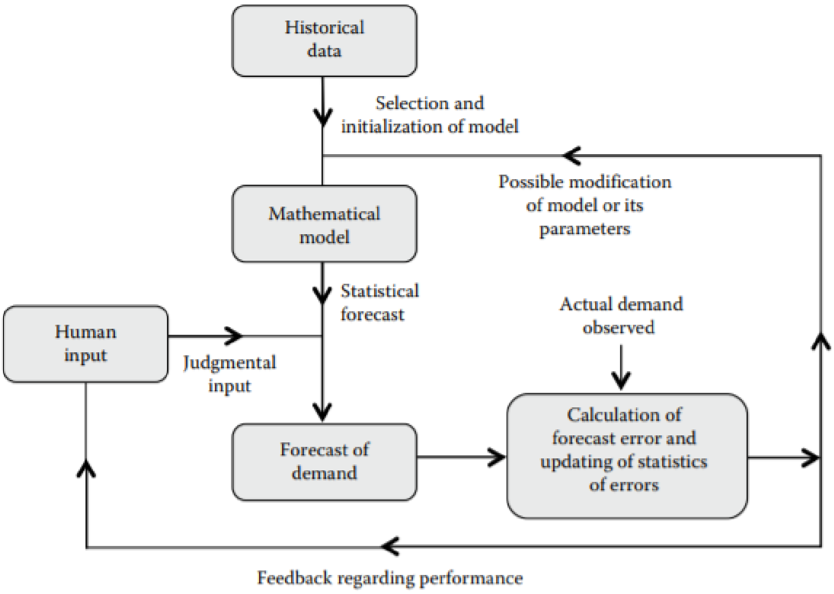

Figure 22: Forecasting framework, retrieved from Silver, Pyke and Thomas (2017)

Returns in other industries versus container returns

In other industries (like for example e-sales, clothing stores or toy stores) each individual item return is usually registered. Besides, return forecasts are made for product categories instead of individual items (Shang, McKie, Ferguson & Galbreth; 2019). For container returns it is usually the other way around: only aggregate return data are available (Kelle and Silver, 1989a), but forecasts are made for each individual container type. The return data in our case are on container-type level, so it is not registered which beer was filled in which bottle. Items in other industries are generally returned relatively quickly after purchase (Shang, McKie, Ferguson & Galbreth; 2019), while it may take several months for containers to return in the brewing industry ((Van Dalen, Van Nunen & Wilens; 2005); (Widi, 2009)). Item-level return data of containers may however “represent a valuable resource that allows for production cost savings or additional revenues” (Fleischmann & Minner; 2004). For low demand items, Toktay, Van der Laan and De Brito (2004) state that it might be beneficial to collect item-level information as there are less items to keep track off than with high demand items. For now, Grolsch finds it too costly to register every individual container upon return, so the focus in this literature review lies on methods and models that use aggregate return data to forecast future returns.

3.3. Past contributions on container returns

In this section we describe different methods and models that are proposed in literature on container returns.

Distributed lag model (DLM)

Goh and Varaprasad (1986) were one of the first to successfully implement a model that uses the relationship between sales and aggregate returns. They used a Box-Jenkins transfer function (Box & Jenkins, 2016) to estimate the distribution of the turnaround time for Coca-Cola bottles and the proportion that is lost in market. The main idea is that the input, sales, undergoes a transformation in a black box resulting in the output, the returns (see Figure 23 below).

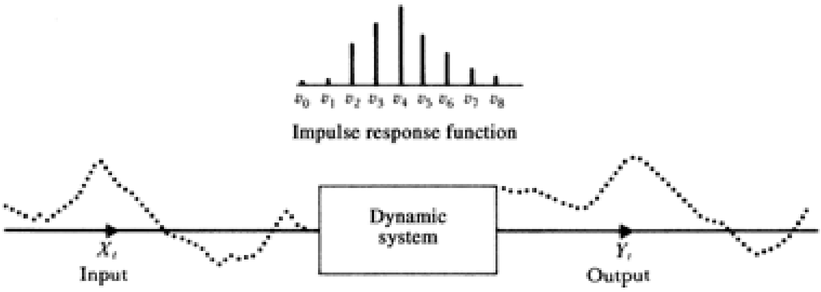

Figure 23: Time series transformation with a distributed lag (from Box & Jenkins, 2016)

Their model estimates the proportions of the past months of sales that contribute to this months' returns. The main drawback of their model is the data requirement, as they suggest to use at least 50 data points while using monthly data. Daily data are too prone to short-term disturbances and gives problems estimating (too many) parameters. Using weekly or two-weekly data can still be appropriate, when the type of industry and the design of the supply chain don't allow spurious changes in the return process from week to week. They also mention to be careful with calculating the recovery rate by dividing the total sales in a year by the returns in a year. The problem is that last year's sales are not taken into account that contributed to the returns of this year. Similarly, the last weeks of sales of this year are included, for which the returns don't come until early next year.

Ma and Kim (2016) added to the model of Goh and Varaprasad (1986) and showed by using the same dataset that their proposed "mixed model" performs better than a Distributed Lag Model (DLM) when there is a weak relationship between sales and returns. A DLM is used to model the relationship between an independent variable and a dependent variable based upon a lag that follows a distribution (Ravines, Schmidtz, Migon; 2006). The model of Ma and Kim (2016) is "mixed" in the sense that it uses past sales as well as past returns to forecast future returns instead of only estimating the weights of the lags based on sales like in a traditional DLM. They also state that the problem with solving a DLM is the multicollinearity (correlated input variables) (see also Geda and Kwong (2019)), not allowing for the traditional Ordinary Least Squares technique with linear regression. However, the time series in a transfer function approach used by Goh & Varaprasad are first made stationary (by removing the correlation of a variable with its own past values) to eliminate the problem of multicollinearity (Box & Jenkins, 2016). This makes the transfer function approach appropriate for modeling the relationship between sales and returns.

Kelle and Silver (1989a) proposed four methods for estimating container returns. The difference between the methods is based on differences in data availability. One method that is particularly interesting is the method that only requires sales and aggregate return data. This method is based on using binary variables for if and when a container is returned. Seeing returns as binary variables can be interesting as sums of independent (but not necessarily evenly distributed) binary variables are approximately normally distributed (Krapp, Nebel, and Sahamie (2013)) if there are enough observations and the percentages are not too low (Chevallier, 2006) (in general n*p >10). This normal distribution can then be used to construct confidence intervals and to say something about the reliability of the amount of returns as done by Kelle and Silver (1989a). However, Clottey (2016) talked about interesting statistical methods for forecasting returns and mentions that using a DLM for forecasting product returns “is considered superior to the normal approximation approach (from Kelle and Silver (1989a)) due to its dynamic nature”.

Kelle and Silver (1989b) added to their previous paper (Kelle and Silver, 1989a) by implementing the forecasting of returns in a purchasing policy for new containers. They opted to minimize holding costs while maintaining high service levels. It is a very useful model and works with cumulative net demand (cumulative sum of demand-returns up to each time period). The cumulative net demand is the difference between demand and returns and needs to be fulfilled with new bought bottles. Because this cumulative net demand is approximately normally distributed for each time period, a standard formula for safety stock can be used to calculate order-up to levels. This could be very useful model for Grolsch, as the current measure of safety stock is determined by experts’ opinions.

Carrasco-Gallego and Ponce-Cueto (2009) proposed a dynamic regression model that can be estimated by using transfer functions (using historical time series data) or other distributed lag techniques. They used transfer functions to solve the dynamic regression model to forecast the returns of LPG-cylinder containers. They also mention that besides estimating the turnaround time distribution in the way of Goh and Varaprasad (1986) and Kelle and Silver (1989a) one could try to fit an existing theoretical distribution on the data instead. Commonly used distributions for this purpose are geometrical and negative binomial distributions. However, Clottey and Benton (2014) found the gamma distribution superior to these distributions because of more flexibility and therefore higher accuracy.

A possible addition to these distributions is the lognormal distribution (Widi, 2009). Widi (2009) analyzed the returns of beer bottles with item-level data in her master thesis, in her case a known day of return per bottle. She fitted different distributions to the data using multiple statistical tests and the lognormal distribution provided the best fit. She then used this distribution in a simulation model of an inventory system to determine optimal order quantities and inventory levels. Geda and Kwong (2019) state that fitting of the parameters of the return lag distributions is not yet studied sufficiently, and they propose to use maximum likelihood estimators for this purpose. A drawback of this idea is that a large sample size is needed when calculating parameters with maximum likelihood estimators (Geda & Kwong, 2019). They used a negative binomial distributed lag in their own forecasting model, which gave good results for the short-term but got more inaccurate further in the time horizon.

Widi (2009) also found that return time significantly differs per season and per product (same bottle, different products). This was also found in a case study of Heineken described in the book of Van Dalen, Van Nunen and Wilens (2005) and expected by Bierkens et al. (2013) who set up a statistical procedure based on item-level data collection to test this. In the case study of Heineken (Van Dalen, Van Nunen & Wilens; 2005), RFID chips were placed in the crates so the time between two production batches with the same crate could be precisely measured. The returns were found to follow a distribution similar to a lognormal distribution: only few returns in the first three weeks, followed by a large amount of returns in the next seven weeks and then diminishing returns from week ten onwards. However, the time of a return in this case study includes sorting and storage time and this might be the reason there are few returns during first weeks after production. The average return time of the crates was 12.5 weeks. The actual time in the market was estimated by determining the average time at the brewery and subtracting it from the total circulation time. But as “FIFO rules were probably not strictly followed” (Van Dalen, Van Nunen & Wilens; 2005), the average time at the brewery of four weeks of warehouse time could be unreliable, especially with crates already being returned in the first week after production.

Markov Chain

Bierkens et al. (2013) tried to find the trade population of beer bottles by modelling the return process of these bottles as a Markov-process using four states: full, empty, broken and returned. They see the transition time from full to empty as the time the consumer takes to consume the beer. But as the storage and delivery times are included in the transition time from full to empty, we question their approach of using only these states. Because of the poison arrivals in the Markov-chain, they used exponentially distributed times for the transition from full to empty and from empty to returned. They then estimated the parameters of the hypo exponential distribution of the total return time by fitting the distribution to observed total return time data. Based on the graph we think the hypo exponential distribution with only two parameters is a poor fit. They do address the problems of seasonality, unreadable bottle labels and the difference in circulation time per distribution channel, bars/restaurants versus supermarkets. The trade population should be big enough for the company to be able to fulfill demand, but not too big as the company will then experience holding costs and underutilization of assets (Carrasco-Callego, Ponce-Cueto & Dekker, 2012).

Grey box modeling

Ene and Özturk (2017) proposed a so called "grey box model" for the forecasting of the returns of end-of-life vehicles. A grey model uses a theoretical formulation together with only a small amount of data. This is different than black box modeling where no theoretical formulation is used at all. The method of Ene and Özturk (2017) could be useful for forecasting returns for newly introduced container types, as there are not much data of these container types available yet. Besides, a theorethical form of other container types could be included.

Wave function

Xiaofeng and Tijun (2009) forecasted the returns of products with a wave function, incorporating cyclical behavior in the function using a cosine. They compared their results in terms of forecast accuracy with a grey model and a Markov-chain model. Their wave function provided the best results. This type of method is useful for modelling seasonality, but has a big drawback of not using sales data in the return forecasting. Another potential problem of using a wave function is overfitting the return curve. Overfitting is trying to explain every data point of the return curve in a function, while these points are also made up out of random noise (Box & Jenkins, 2016). This random noise does not describe the actual process and the function inaccurately forecasts the future of the time series if the noise is incorporated in fitting the function. However, it could be interesting to see if a wave function can be incorporated with statistical techniques such as transfer functions to account for the seasonality between two time series. We did not find any contributions on explaining seasonality between two time series (time series having stronger cross-correlation in different periods of the year). The transfer function approaches used by Goh and Varaprasad (1986) and Carrasco-Callego and Ponce-Cueto (2009) assume a distributed lag that does not change over time.

Neural network

There are also some contributions for forecasting returns based on machine learning. These methods are particularly useful when there are many explanatory variables or when there is vagueness about how variablesinteract with each other and models can be set up by experts' rules (Cui, Rajagopalan & Ward, 2020). This is also done by Gomez et al. (2002) who proposed a neuro-fuzzy approach to forecast the returns of photocopiers. The input variables included sales, life-cycle stage, usage intensity, life-expectancy and reasons why photocopiers do not return. Experts' rules are needed to set up the relationship between these variables and the return rate and timing. The system is then trained with historic data. Neural networks need a vast amount of training data and in practice these data are not available for all product-types that need forecasting (Ene & Özturk, 2017).

3.4. Theoretical background

In the first part of the literature review, we have described methods that are used in the past to determine the important parameters for container returns. In this part we describe the theory behind the most interesting approach for our case: modelling the return forecast problem as a DLM and solving the model with time series analysis.

Time series analysis

A time series is a set of data points (observations) taken sequentially in time (*Holmes, Scheuerell & Ward; 2020). The analysis of* time series has two main goals:

1. To understand the structure of the time series (how it depends on time, itself, and other time series)
2. To forecast/predict future values of the time series

There are many methods and models present for time series analysis ranging from quite simple to very sophisticated. We start by describing some simple regression techniques to give some background on time series analysis. We then describe the transfer function approach that is particular interesting for our research as seen in the previous section about container return forecasting.

Simple linear regression

In the simple case of having one explanatory variable x_t , the output y_t can be explained by the simple linear regression formula (*Hyndman & Athanasopoulos, 2018*):

$$\boldsymbol{y_t = \beta_0 + \beta_1 x_t + \varepsilon_t}$$

With B_0 the intersect value at X=0, B_1 the slope of the linear line and E_t the error term. In order to find the values of B_0 and B_1, the ordinary least squares technique can be used which minimizes:

$$\sum_{t=1}^{T} \varepsilon_t^2 = \sum_{t=1}^{T} (y_t - \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \dots + \beta_k x_{k,t})^2$$

This will result in the best fit linear line that explains the points, as shown in Figure 24:

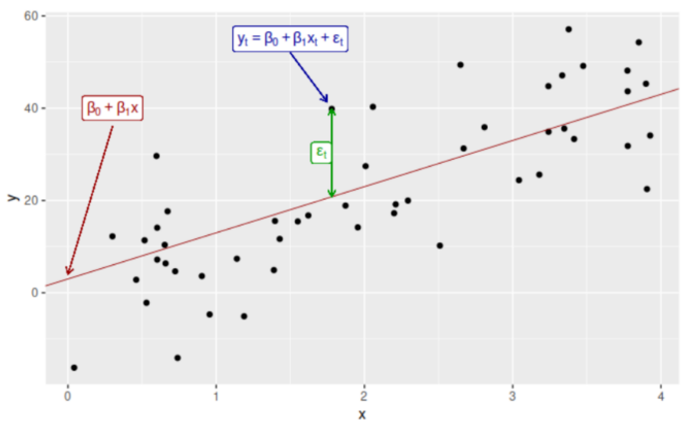

Figure 24: Simple linear regression (From Hyndman and Athanasopoulos, 2018)

Multiple linear regression

In multiple linear regression, the output y_t is explained using multiple explanatory variables (*Hyndman & Athanasopoulos, 2018*):

$$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \ldots + \beta_k x_{k,t} + \varepsilon_t$$

The problem with applying multiple linear regression to our case is that the explanatory variables are probably correlated with each other, which will give poor regression results. This phenomenon is called multicollinearity. As Goh and Varaprasad (1986) mention: "more sophisticated techniques are necessary".

Autocorrelation

In order to find out if the explanatory variables are correlated we can make an autocorrelation plot. Autocorrelation is the correlation of a variable's value with its own past values. It is a very important measure for seasonality within a time series and it is needed to identify if multiple linear regression techniques can be used without encountering multicollinearity (Box & Jenkins, 2016). The autocorrelation at lag k is determined by dividing the autocovariance s_k by the sample variance s_0:

$$r_k = \frac{s_k}{s_0}$$

with:

$$s_k = \frac{1}{n}\sum_{t=1}^{n-k}(y_t - \bar{y})(y_{t+k} - \bar{y}) = \frac{1}{n}\sum_{t=k+1}^{n}(y_t - \bar{y})(y_{t-k} - \bar{y})$$

A plot of r_k against the lag k is called a correlogram. When a time series shows no autocorrelation, the time series is called "stationary".

Cross-correlation

Cross correlation is the correlation of a variable's value with another variable's past values. Cross-correlation $r_{xy,k}$ is calculated using the following formula (Box & Jenkins, 2016):

$$r_{xy,k} = \frac{s_{xy,k}}{\sigma_x \sigma_y}$$

with:

$$s_{xy,k} = E[(y_t - \mu_t)(y_{t+k} - \mu_x)]$$

Cross-correlations are used in the transfer function approach to estimate the delay distribution (Box & Jenkins, 2016).

Box-Jenkins transfer functions

A delayed change in a variable caused by a change in another variable is called a dynamic response. A dynamic response can be modeled by a transfer function model with the general form:

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 x_{t-1} + \ldots + \beta_k x_{t-k-1} + \varepsilon_t$$

Where y_t is the output at time t, x_t is the input at time t, B_t is the weight of the input and E_t is the error or "noise". The form of formula is very similar to the multiple linear regression formula, except that the output is not dependent on multiple variables but on lagged values of one variable. Besides, the transfer function approach makes up for correlation between the independent variables.

The process of modelling such a dynamic response by a transfer function model consists of three steps: identification of the transfer function, fitting and diagnostic checking (Box & Jenkins, 2016). We give a summary of these steps as the total procedure is too long to describe here. A detailed description of transfer functions can be found in the book of Box and Jenkins (2016). Another paper with a clear description of the transfer function approach is a paper by Helmer and Johansson (1977).

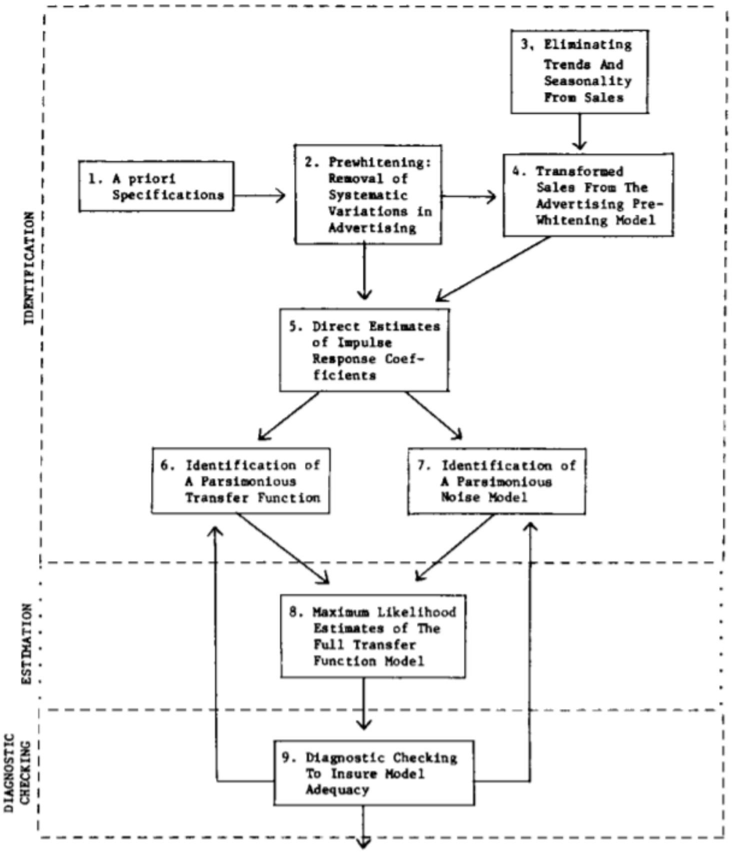

Figure 25: The steps in identification, fitting and diagnostically checking a transfer function (from Helmer and Johansson, 1977)

They studied the relationship between the independent variable advertising (which would be sales in case of Grolsch) and the dependent variable sales (which would be returns in case of Grolsch) by the means of a transfer function. They give a nice overview of the process and divided the identification into multiple steps (see Figure 25).

Safety stock

Safety stock is the amount of extra inventory to protect against uncertainties in both supply and demand (Silver, Pyke & Thomas, 2017). How much safety stock is needed, is based on what level of customer service is desired (Cycle Service Level (CSL)). The safety factor can be determined as follows:

$$Safety\ factor\ k = \Phi^{-1}(CSL)$$

The safety stock can be calculated with the following formula:

$$Safety\ Stock\ = k * \sigma_L$$

with:

- k = the safety factor
- σ_L = the standard deviation of the demand during the lead time.

The reorder point is the inventory point at which a new order has to be made to be able to fulfill the demand during the lead time:

$$Reorder\ point\ =\ \hat{X}_L\ +\ Safety\ stock$$

with:

- $\hat{X}_L$ = the mean demand during the lead time

3.5. Conclusion

For long-term forecasting, updating approaches such as moving average, exponential smoothing, Holt's method and Bayesian methods are considered inappropriate. The main reason these methods are undesirable is the weekly augmentation of sales and return data can't be utilized to update the parameters as the long-term forecast is only made a few times per year, mainly to determine the procurement budget. (These methods make more sense for short-term forecasting.) Besides, univariate methods seem inappropriate for this research as they fail to take sales into account to forecast the returns and sales can differ significantly from week to week. For long-term return forecasting the better approach is to perform time series analysis, which is the analysis of trends, seasonality (cycles) and correlation between sales and returns. The DLM is in our opinion the best way to model the return forecast for Grolsch, because it only needs aggregate data and can be solved using time series analysis.

Other methods such as machine learning (for example neural networks) are typically used for forecasting with more explanatory variables (features) than we have in our case (Cui, Rajagopalan & Ward, 2020). This makes them more suitable for return forecasting in other industries, but less suitable for container return forecasting. The purchasing model of Kelle and Silver(1989b) can be adapted as purchasing policy. Although Kelle and Silver do not use restrictions on production capacity, these restrictions could be added to the model to make it realistic for Grolsch. It is a useful model because the safety stock formula can be used as the cumulative net demand in a time period (the cumulative sum of demand-return up to the specific time period) is approximately normally distributed. This could be an improvement to Grolsch current measure of safety stock, the standard of five working days of production.

4. Improved return forecast and parameter calculation

This chapter describes the proposed return forecasting model and the improved calculation of the parameters Trade Loss (TL), Internal Loss (IL) and Trade Population (TP). The parameters TL and IL are input for the improved return forecast, the expected TP is output. We did explicitly not mention the parameter Weeks in Trade (WiT). As we have seen in Chapter 2, forecasting returns with the parameter WiT is inaccurate because of its sensitivity to large changes in sales from week to week. The realized WiT-values of a previous year can also not be copied to the next year, which is currently done. We have seen in the literature review in Chapter 3 that the container returns can be best modeled by using a distributed time in the market. In this chapter we therefore propose to forecast returns using a different parameter: Time in Trade (TiT). TiT is the time that a container stays in the market. Grolsch is already familiar with the term TiT but currently does not use TiT in its return forecast. The TiT-distribution is the probability distribution of the time that a container returns in.

In Section 4.1 the proposed return forecasting model is outlined and in Section 4.2 the improved parameter calculations and the improved return forecast are described. In Section 4.4 we try to validate the improved return forecast model.

4.1. Proposed return forecasting model

In Chapter 3 we described several methods and models that can be used to forecast container returns. As a summary, these are shown in Figure 26:

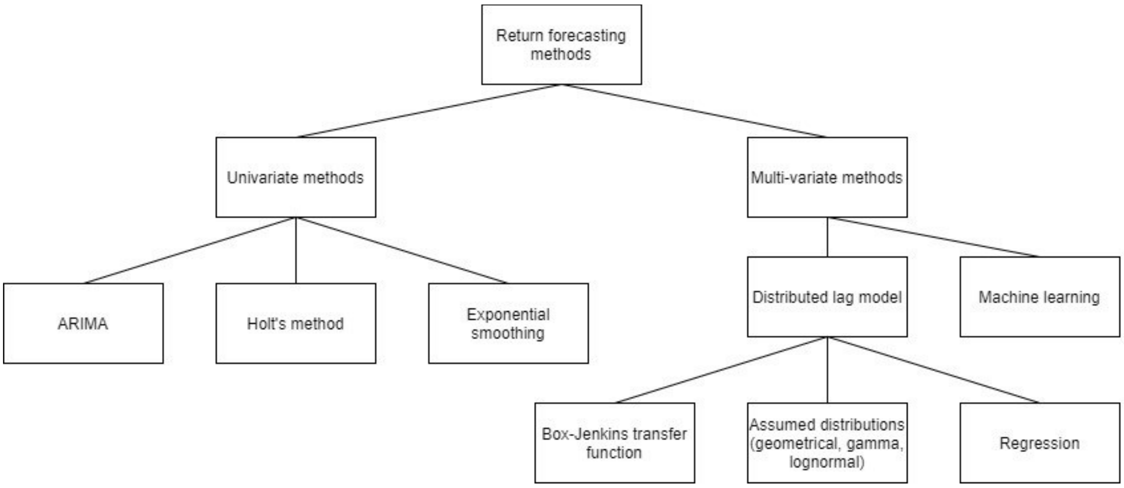

Figure 26: Alternative methods and models

In Chapter 3 we came to the conclusion that multivariate time series modeling is a good method to forecast returns for the long term with aggregate data and with only two variables (sales and returns).

The problem can be modeled as a Distributed Lag Model (DLM), but there are some choices to be made on how to exactly do this. First there is the choice between an infinite- and finite DLM. The difference between the infinite and finite DLM is the number of parameters (lags) that is included in the return forecasting model.

With infinite distributed lag models a certain structure needs to be assumed to solve the model as no technique is able to determine an infinite number of parameters. In literature we found several distribution structures that are generally used for this purpose: the geometrical distribution, the gamma distribution and the lognormal distribution.

Another option to find a solution to the infinite distributed lag model is the transfer function approach of Box and Jenkins (2016). This method uses correlation between a variable's own values together with the cross-correlation between the two variables to say something about the maximum amount of significant lags, while allowing an infinite amount of lags. To drawback of this method is that many years of data are needed and the validity of the method is determined by the quality of the data. The data should for example not show spurious effects. Time periods longer than weeks are therefore preferred, probably a monthly lag seems appropriate. Especially because the feeling is that containers are still returning after longer than a year. As Grolsch needs the return forecast per week, as this is used organization-wide, we need to see if this approach will be accurate. Grolsch has some idea about a maximum significant lag length and we therefore propose to model the return forecast as a finite distributed lag model.

The proposed return forecasting model is:

$$R_t \ = \ Y(S_t - S_{t-1}) + b_1(S_{t-1} - TL) \ + b_2(S_{t-2} - TL) + \cdots + b_n(S_{t-n} - TL)$$

with:

- R_t = Forecasted returns in week *t*
- S_t = Forecasted or realized sales in week *t,* depending on for which *t* the return forecast is calculated. (The return forecast of four weeks in the future depends on the sales forecast of the next four weeks as well as on the realized sales of past weeks.) The realized sales are known. The sales forecasts of Apollo and BNR for 78 months ahead are updated on a weekly basis by the Demand Planning department and are also available for our research.
- TL = Trade loss
- b_t = Lag weight of week *t* (b_1 = % of bottles that return one week after they are sold). Together, the b_t values should sum up to 100% as the sold bottles (S_t) minus the bottles lost in the market (TL) are the returns (R_t).
- *n* = The maximum amount of weeks before week *t,* from which the sales still contribute to the return forecast of week *t.*
- Y = Extra parameter (index) indicating extra returns if more is sold than a week earlier and less returns if less is sold than a week earlier. We have to note that an accurate return forecast is not just based on the TiT-distribution of an individual bottle. There is likely a high correlation between sales in a week and returns in the same week, because Grolsch gets more back the more customers are visited (more is sold). This can however not be the same bottles that are sold this week as they have to go through selling, returning and transport phases that together take at least a week. This parameter is explained in Section 4.2.3.

To illustrate how the return forecast is found, we give a small example:

Consider the sales of week 1 are 10 and of week 2 are 20,TL is 10% and we want the return forecast of week 3. The calculation becomes: $R_3 = Y(18\text{-}9) + b_1{*}18 + b_2{*}9$*. Because the realized returns are known, the Excel solver can be used to find the best fit parameters Y, M and V. M and V are the mean and variance of the lognormal distribution of TiT.* b_1 *and* b_2 *are probabilities of that lognormal distribution at x=1 and x=2 respectively.* b_1 *is the probability of return after one week and* b_2 *the probability of return after two weeks.*

Maximum lag length

With a finite DLM, the maximum lag length needs to be determined. In our case the maximum lag length is the number of weeks in the market after which Grolsch does not expect any containers back. This maximum lag length is crucial for determining the TP. The cumulative effect of containers staying in market gets bigger the longer the maximum lag length is. This is the case because the longer a container is expected to be in the market, the more the TP grows before it returns. We solve our DLM with a maximum lag length of 52 weeks, as this is the best guess by Grolsch and a sample of 2019 showed very few bottles returning after 52 weeks.

With a maximum lag length of 52 weeks, we need to specify what happens to the remaining probability mass after 52 weeks. As 100% of sales-TL is expected back, all remaining probability mass has to be scaled over the 52 lags that are included in the model. How this is modeled becomes clear in the next part: "Finding the parameters of the return forecast model".

Finding the parameters of the return forecast model

There are different techniques to solve a finite DLM, of which the most standard one is regression analysis. Standard regression is however not a possible solution to our model as both the sales and return data are auto-correlated. From literature it becomes clear that using regression in this case will lead to multicollinearity which gives poor lag weight estimates.

In Figure 26 there are two other methods present that can be used to solve this model:

1) Assuming a lognormal distribution structure
2) Using a Box-Jenkins impulse response estimate

Both methods are interesting for our case, even though the Box-Jenkins approach might not be optimal with weekly time intervals. These methods differ from each other in terms of how much the data can dictate the shape of the return distribution. The feeling is that letting the data dictate the shape is a good thing. However, this is not always the best approach. Especially with weekly data compared to monthly data, overfitting can become an issue because of short-term disturbances in the return process. Besides, many parameters have to be estimated and the accuracy of this method is unknown beforehand. With weekly disturbances in the data, assumed structures could provide better results.

If we assume a maximum lag length of 52 weeks, 52 parameters need to be estimated. Assuming logical structures drastically reduces this amount of parameters that needs to be estimated. The lognormal distribution has been identified to accurately describe the return distribution in other cases of beer container returns and makes sense to Grolsch as well. We tested multiple distributions, but the lognormal distribution indeed gave the best fit. Instead of estimating every lag weight, the model now only has to estimate the two parameters of the lognormal distribution: the mean and the variance. We choose to test both methods. We now move to the inputs of our return forecasting model:

Inputs

1) Realized sales per container type per week
2) Realized returns per container type per week (determined in the same way as outlined in Section 2.5.2)
3) Sorting line output data: % of each bottle type present in each crate type per year
4) Filling line sorting loss and bottle breakage per bottle type per year

5) Trade Loss which is based on inputs 1), 2) , 3) and 4) as explained in Section 4.2.
6) Internal loss based on inputs 3) and 4) as explained in Section 4.2.

Our parameter calculation model translates these inputs to the following outputs (for each container type) for the coming year:

Outputs

1) A TiT-distribution of the returns (Section 4.2.2)
2) Return forecast (Section 4.3)
3) Expected TP per week (Section 4.2.4)
   (When the TiT-distribution is estimated, it is also clear how much percent of the sales of each week are still expected to be in the market. Needs an appropriate maximum lag length to ensure the cumulative aspect of trade population is accounted for.)

In Section 4.2 the parameter calculation for Trade Loss, Internal Loss, Time in Trade, Trade Population and the impact of the difference in total sales is explained.

4.2. Parameter calculations

In this section is explained how the parameters Trade Loss, Internal loss, Time in Trade and Trade Population are calculated. Trade loss is needed as input to calculate TiT. Internal loss does not influence the return forecast, but does influence the purchasing decisions in Chapter 5 and 6.

4.2.1. Trade Loss and Internal Loss

In this section we describe the calculation for Trade Loss and Internal Loss. We start with Trade Loss. Trade loss is currently determined by Grolsch using the following formula:

$$Trade\ Loss_{i,p} = \frac{Total\ sales_{i,p} - Total\ returns_{i,p}}{Total\ sales_{i,p}}$$

with:

- i = container type
- p = period over which to take TL

Realized returns

The first thing that is important for the TL-calculation is that the realized returns are calculated correctly. We have seen that the current calculation of TL is based on a biased calculation of the realized returns. The realized returns are currently calculated by subtracting sorting losses per week. This results in spiky realized returns as some weeks more is sorted than in other weeks. Above all, not all containers that are returned in a period are sorted in the same period. This results in a slightly too low estimate of TL because too few losses are subtracted. We therefore propose to use the realized returns calculation of Section 2.5.2 for TL-calculation that uses sorting percentages instead. As a small recap, the number of returned crates of each type is known and with sorting data we work out how many of each bottles are returned in these crates.

Time period

The second thing to consider is that it is not set in stone over how long of a period to calculate TL and how often it should be updated. Grolsch may not use a period of less than a full year for the reason that returns follow later than sales.

For example: Consider two years of sales and returns, 2018 and 2019. The TL is going to be calculated for 2019 so we would take the total sales and returns of 2019 in the TL formula. However, the returns of the first weeks of 2019 are based on sales of the final weeks of 2018, which are not included in the calculation. Besides, the sales of the final weeks of 2019 are included, but they have not yet been returned. If there is no clear sales trend present, the sales of the final weeks of 2019 correspond with returns in the first weeks of 2019. This is the case because the sales in the final weeks of 2019 are comparable to sales in the final week of 2018. This is why TL should be calculated over full years.

There is a trade-off for the time period over which to calculate the TL. If TL is calculated over a too short period, the values will be very volatile. If TL is calculated over a too long period, the market might have been changed. For example: Grolsch might have introduced SKUs that are sold without a crate, which could lead to an increase in Trade Loss.

In Figure 27 the sales- and return sums of 1 year are shown with a rolling horizon. The difference between the lines in each graph are the TL values calculated over a period of one year, starting in the week that is indicated on the x-axes. The values for TL that are calculated over a period of 1 year, fluctuate a lot based on which starting week is taken. This is the case because sales and returns per week can differ a lot from year to year as shown in Figure 29.

For example: If the TL for Apollo is calculated from week 1 to week 52 of 2018, the TL is X%. If the TL for Apollo is calculated over week 2 of 2018 to week 1 of 2019, the TL is X%. This is a significant difference as X% difference equals X hl and translates to X euros that might be unrightfully included in the injection budget.

Figure 27: Trade loss 1 year rolling

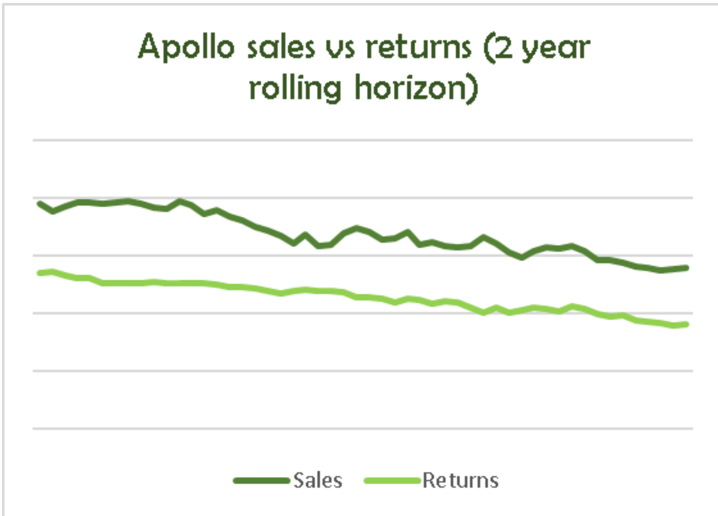

Figure 28: Trade loss 2 year rolling

The TL calculated over a period of two years is already more stable as shown in Figure 28. Currently a three-year average TL is used. However, the longer the period, the more changes in the market can have occurred. For example: if more Radler beer is sold without a crate over the years, more could be lost in the market. For Apollo the TL should be relatively stable throughout multiple years as this bottle is only used by

Figure 29: Sales differences year to year (confidential)

Grolsch and not much changed in the market landscape. The differences in sales and returns in the same week of different years can be significant.

Then there is the delay between sales and returns. When returns come on average X to X weeks later than the sales, taking a period less than 52 weeks can lead to serious inaccuracies. For this reason the minimum period is one year.

Then there are periods in which more is returned compared to sales than other periods. If we for example take a period of 1.5 years, ending with a peak season, this includes a high amount of sales while the returns are not yet followed. For this reason Grolsch always uses full years as periods. Returns in the first weeks of the period over which TL is taken, can be roughly compared with the returns of the sales of the last weeks of the period if the sales remains stable.

Stability of Trade Loss

Stability is not directly mentioned in the formula of Trade Loss. We are concerned with stability because Grolsch uses Trade Loss as a constant value, while in reality the market can change and TL needs to be updated. Also if there is a trend in sales present, the Trade Loss calculation is biased because returns follow later than the sales. If sales are increasing, Trade loss is increasing because the increase in returns comes on average X weeks later. So we consider two types of stability: stability of TL based on the market and stability of TL based on trends in sales.

There is a big difference in market stability between Apollo and BNR. Apollo is only used by Grolsch and BNR is shared among different brewing companies. For that reason, for Apollo the Trade Loss can be assumed to be stable throughout multiple years. For BNR Grolsch' experience is the TL differs more from year to year. One reason is that the amount of BNR that is put into Grolsch crates at the supermarkets can differ from year to year. This is also a reason that new brewing companies have a hard time breaking into the market. They need to invest in BNR and the bottles may end up at Grolsch. The stability of TL for BNR in terms of sales is shown in Figure 30. We see that the trend is upwards and that the TL values are increasing, even if TL is taken over a relatively long period of two years. This is mainly due to the fact that BNR returns follow on average X weeks after the sales (which is calculated later on in Section 4.2.2.). When more is sold, the additional returns are expected X weeks later. We need to shift the returns with X weeks to get a more stable TL. This is needed because the data indicate an increasing TL, while in reality the returns corresponding with the increased sales are not taken into account.

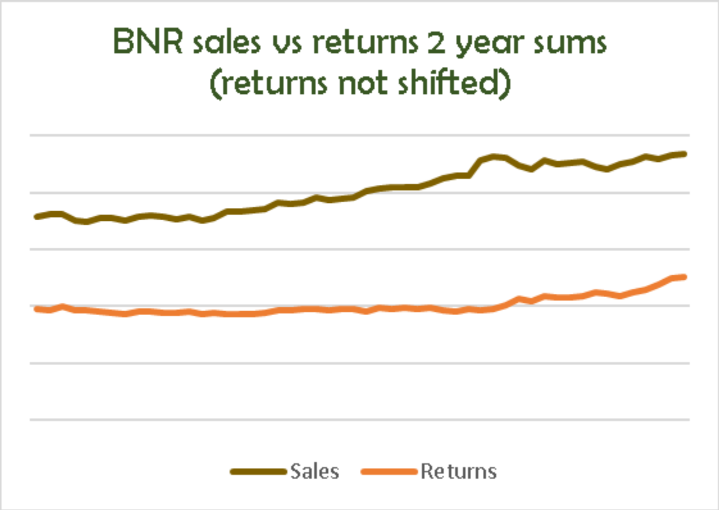

Figure 30: Trade loss BNR returns not shifted

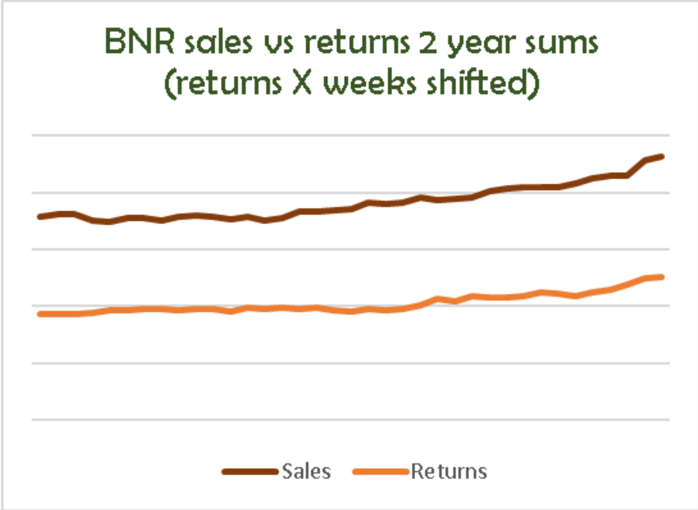

Figure 31: Trade loss BNR returns shifted

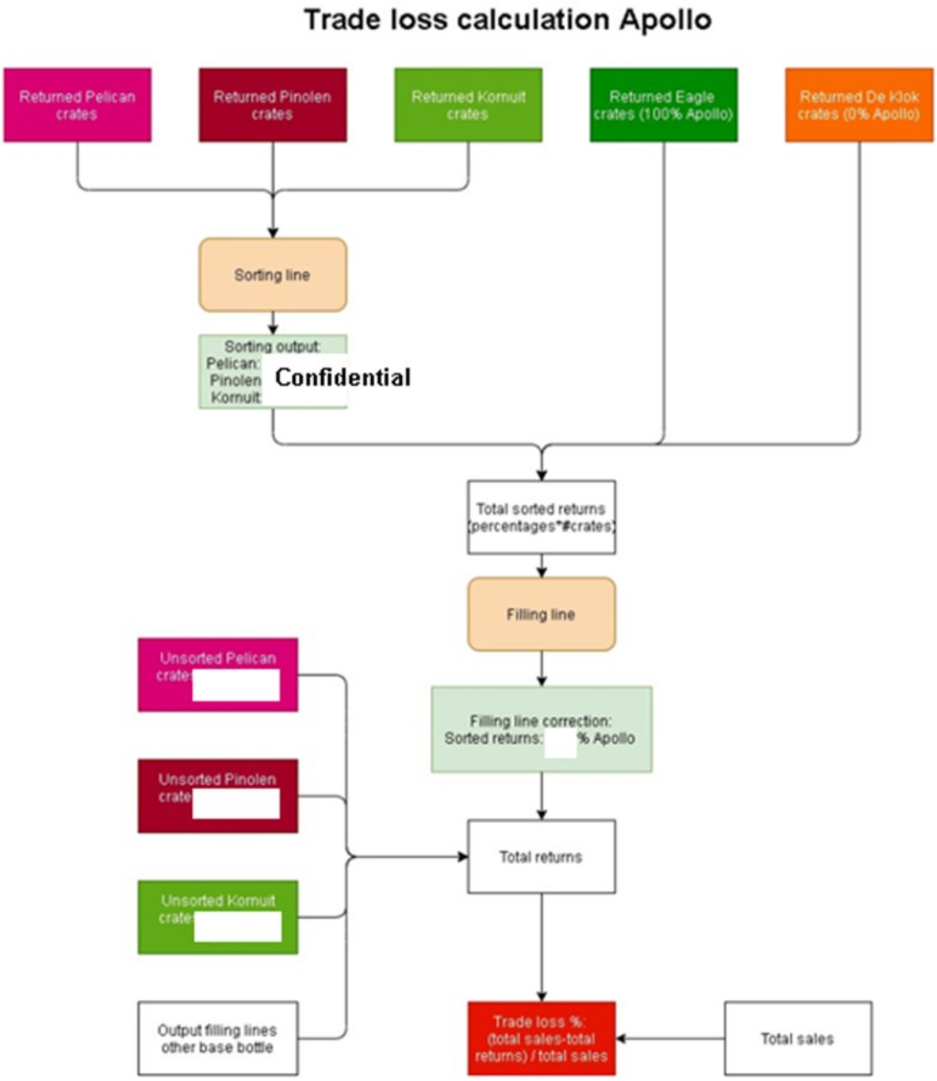

Figure 32: Trade loss calculation

For Apollo we propose to use an average TL, based on the previous three full years. For BNR we also propose to use a three-year average, but with a correction for the sales trend by shifting the returns by X weeks. Besides, for BNR there always needs to be human input to react to changes in the market. For example: When a brewing company stops producing with BNR, the TL value needs to be adjusted accordingly (lowered). For Apollo this results in a TL value of X% and for BNR X%.

A small validity check is done to see if the amount of injection is in line with the expected TL. The amount of Apollo injection in the years 2018, 2019 and 2020 together was around X hl. The amount of sales in these years was around X hl. This indicates a maximum TL of X%. As background information from Grolsch, we know that at the end of 2020 more was injected than needed as the supplier had the stock available. In this X% an Internal Loss of around X% needs to be taken into account, which means the TL for Apollo should indeed be around X%.

Internal loss

For internal loss we propose to keep using the current method of taking the average production line breakage of the last three years. There is no direct data of how many containers are lost during storage, but broken bottles may stay in the crates till they arrive at the production line. Overall the current approach seems appropriate. The IL is calculated with the following formula:

$$\boldsymbol{Internal\ Loss_{i,p}}\ \% = \frac{\boldsymbol{Production\ line\ breakage_{i,p}}}{\boldsymbol{Production\ line\ input_{i,p}}}$$

with:

- i = container type
- p = period over which to take IL

The calculation using the formula above results in X% and X% IL for Apollo and BNR respectively. Unlike Trade Loss, Internal loss is stable.

4.2.2. Time in Trade

The TiT-distribution is the distribution of when a container is expected to be returned (% per week up to 52 weeks). When the TL value is estimated, the Sales-TL is the total amount of returns of a certain week of sales. The TiT-distribution says how these returns are divided over the weeks after the containers are sold. TiT can be estimated using a discretized lognormal distribution. The estimated TiT is the best possible fit in terms of MAPE. With the lag limit of 52 weeks, the part of the returns that is expected after these 52 weeks needs to be scaled over the 52 weeks that are included. The scaling is done by assuming the same distribution. So if X% is expected to come back after 52 weeks, this X% is scaled over the 52 weeks in the same way the 52 weeks are distributed (most of the X% is added to the week where most % of returns are expected. This keeps the TiT-distribution smooth. This does result in a difference between the average TiT of the 52 weeks and the actually implied TiT with no time limit. This is important to keep in mind for calculation the TP in Section 4.2.3. The percentage that is allowed after 52 weeks is set to X% for Apollo, deliberated with Grolsch and validated by a small sample from 2019 that showed very few bottles returning after 52 weeks.

There is a trade-off for allowing too much or too little % returns after 52 weeks. The model has more freedom with more % allowed after 52 weeks, as more shapes of the lognormal distribution are allowed. However, because the part after 52 weeks has to be scaled, more % allowed means a bigger difference in the average TiT and implied average TiT for TP estimation. For BNR the TiT is in general longer than for Apollo and we therefore set the percentage for BNR to X%.

Excel calculates the best fit lognormal TiT-distribution, by minimizing the MAPE. The interface in Excel looks as follows:

Confidential

Figure 33: Interface TiT-distribution

This fitting results in the following TiT-distribution for Apollo and BNR, fit over 2018-2019:

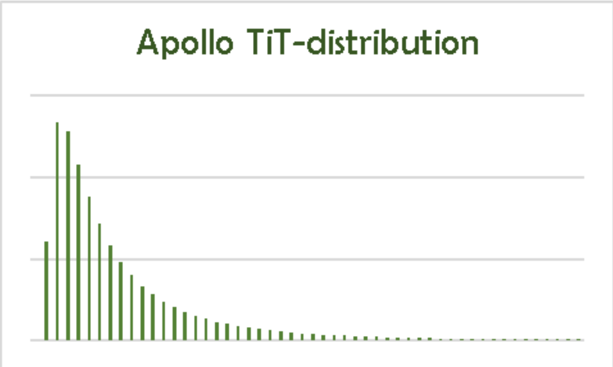

Figure 34: TiT-distribution Apollo

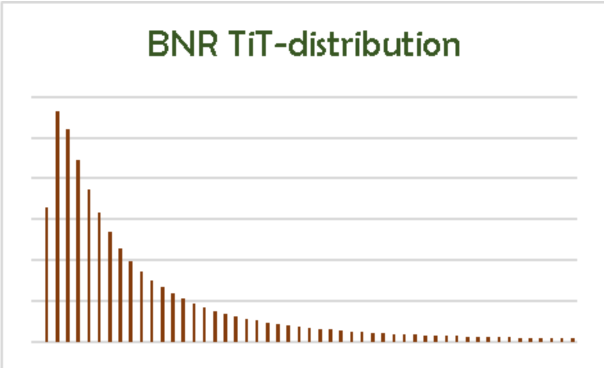

Figure 35: TiT-distribution BNR

We notice that both have generally the same shape (which is not necessarily the case as the lognormal distributions can have many shapes), but BNR is returning slower than Apollo. After three weeks the total expected returned amount for Apollo is around X%, while for BNR this is around X%. The average implied TiT for Apollo is X weeks and for BNR X weeks.

Seasonality of TiT

Till now we used one distribution for TiT, which does not change over the year. We minimized the MAPE with the fitting of the TiT-distribution. The return forecast MAPE of the fitted data is already significantly reduced by using one TiT-distribution compared to the current method using WiT However, the weekly **absolute** return forecast errors may be small, but if there is constant over- or under-forecasting of returns, this can have an impact on the injection planning. As an example:

If the return forecast errors for weeks 1 and 2 are -3000 hl and +3000 hl respectively, the MAPE over these two weeks is 3000 hl, but the two week total error is 0 hl. If the return forecast errors for weeks 1 and 2 are +2000 hl and +2000 hl respectively, the MAPE is reduced to 2000 hl. However, the two week total error is 4000 hl, compared to 0 hl in the first case.

We have the feeling that bottles return faster in the peak-season, because people drink more beer in this period. To check this, we plot the cumulative hl-error of the fitted TiT-distributions of Apollo and BNR in Figure 36 and Figure 37 respectively. Note that at the peak the underforecasted returns is at X hl. As this peak also lies the period in which injections are generally necessary, around X euros on injections might be unrightfully included in the budget. There will always be some error and the current model has errors of more than X hl in one week, but the seasonality needs to be adressed in our proposed return forecasting model.

The main split we can identify in the “One TiT-distribution” lines in Figure 36 and Figure 37 is a split between two periods. Apparently, Apollo and BNR are returning faster in one period than the other seen the increasing and decreasing of the cumulative error. Seasonality of TiT needs to be included to make the forecast more accurate. Note that the peaks for Apollo are around weeks 36 of 2018 and 2019 and that the low is around week 10 of 2019. Week 36 is the end of the peak-season for Apollo, with the summer and holidays coming to an end. Till week 10, with a small peak in demand for Apollo bottles in week 49 before Christmas. From week 10, the beginning of Spring, the beer consumptions rises again.

This validates our feeling that the main split should be two periods of 26 weeks. A peak in week 36 in Figure 36 means that from week 10 till week 36 containers are in reality returning faster than is currently modeled with one distribution for TiT. From week 37 till week 9 containers are in reality returning slower than currently modeled.

For BNR we see in Figure 37 that there are no clear peaks, but rather some horizontal parts in the graph. This means that the hl-errors of the weeks 36 to 46 are around the same as those of weeks 10 till 20. This would call for a split in three different distributions for BNR:

- One TiT-distribution for weeks 21-35 (faster BNR return than average, based on increased sales over all containers)
- One TiT-distribution for weeks 36-46 and weeks 10-20 (start peak season of BNR with introduction of Herfstbok in week 36 and beginning of Spring in week 10)
- One TiT-distribution for weeks 47- 9 (slower BNR return than average, based on lower sales over all containers)

An optimal model might even include a different distribution for every week in the year, but the risk of overfitting becomes large. We see this as a topic for further research and stick to a two period split of TiT for Apollo as well as for BNR because Grolsch does not want different solutions for different container types if the improvements are expected to be minimal. This makes the solution more interpretable and easier to implement for Grolsch. However, even this split in two periods makes the model more complicated than using one distribution for TiT. Summarized: For Apollo we split the year in weeks 10 till 36 and weeks 37 till 9. For BNR we split the period in weeks 15 till 40 and 41 till 14. This is only done for TiT. The TL and IL are assumed constant throughout the year.

If we have a different distribution for peak season and off-season, how do we make the switch between these periods in the return forecast? It has to do with looking backward or forward. Looking backward means: In a certain week, how many returns do we expect from the sales of previous weeks? Looking forward means: When do we expect to get the sales of the current week back? If we look backward with a different distribution in let’s say week 20 and week 21, the sales of week 10 (10 weeks before) might already be expected back in the week 20 and expected to be in trade again in week 21. This results in an invalid return forecast. If we make a split in distribution, we need to look forward and keep track of each week of sales how much percent returns in which week after the sales. Then for each week, we distribute when the sales are returned.

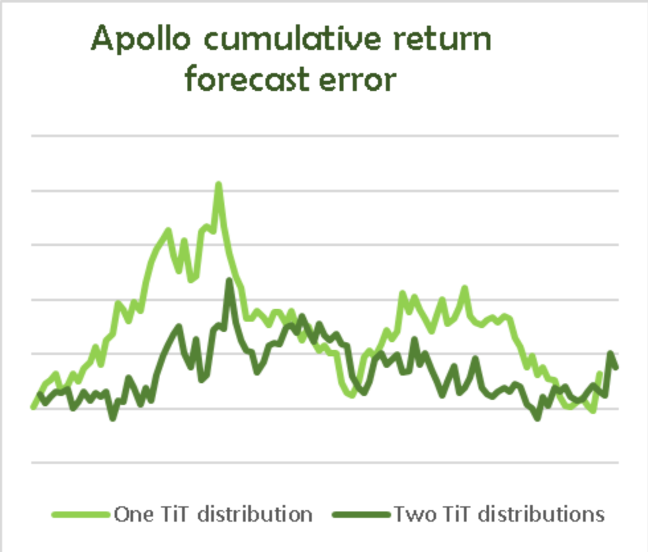

Figure 36: Apollo cumulative return forecast error

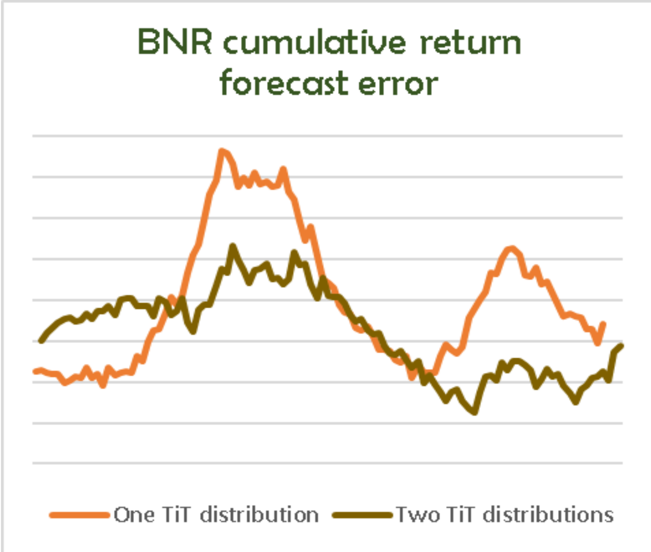

Figure 37: BNR cumulative return forecast error

Most seasonality is indeed captured by using two different distributions for peak-season and off-season. There is no systematic over- and under-forecasting anymore in the same periods in both years. The pattern that still exists is likely because TL varies over the year instead of the assumed constant TL. Especially for BNR the TL can variate throughout the year. Overall the cumulative error is staying in a good range and will not cause problems for the injection planning.

If returns are under-forecasted this may result in a too big amount of planned injection. This is only the case if there is less stock of empty bottles available than the amount of returns that is under-forecasted. If the empty bottle stock is big enough, under-forecasting of returns does still not lead to planned injections. If returns are over-forecasted, this may result in a too small amount of planned injection. But if too much injection is planned, this amount is added to the expected stock and is used (needed) at one point or another. This is why the range of the cumulative error is important. As long as the cumulative error is relatively small, the impact on the injection planning will be small.

4.2.3. Extra return parameter Y: difference in total sales in consecutive periods

Before the final improved return forecast is made, we determine the return parameter Y from the DLM of Section 4.1. The reason why this parameter was included in the DLM of Section 4.1. is that weekly returns follow a very spiky pattern, similar to weekly sales. Grolsch gets more bottles back if customers' depots are visited more often. So in periods of more total sales (sales sum over all bottle types), more total returns are expected as well. To go one step further, if Grolsch has more total sales in the current week compared to the previous week, most of the spare returns that are available at the customers' depots are returned to the brewery. The next period (even when sales are also high), the returns will likely be lower. As an indication that this may be true we plot the graph of the difference in sales versus the differences in returns in Figure 38.

For 2020, X% of the time the sign of difference in total sales (+ or -) was the same as the difference in total returns. One would expect this value to be 50% if this parameter's influence was completely random. As returns are influenced by way more variables, this X% is certainly a reason to check the impact on the return forecast accuracy by including the total sales as input parameter to the return forecast. However, extra returns in the current week caused by more sales in that week are not the same bottles that are sold in that week. These bottles first have to go through transportation, selling and returning phases that together usually take at least one week.

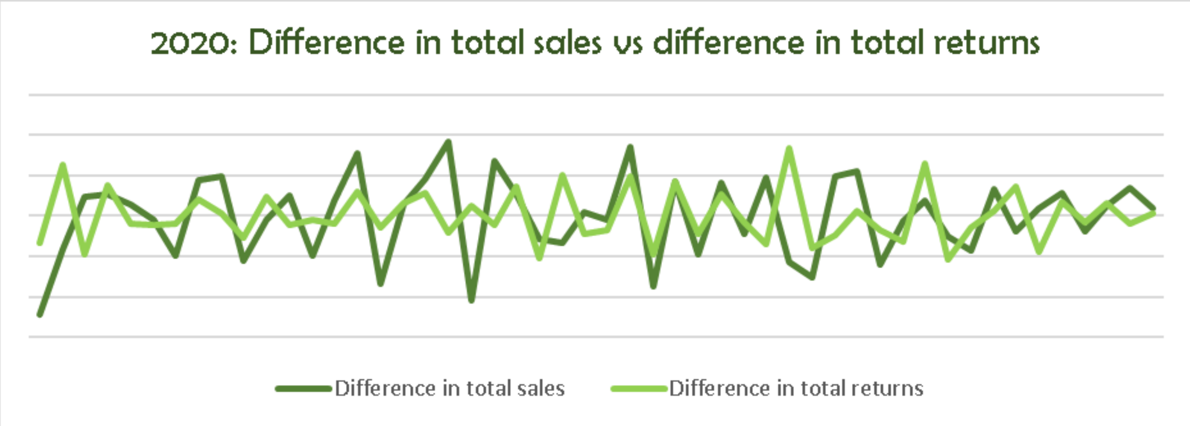

Figure 38: Difference in total sales vs difference in total returns

Keeping the return forecast model valid

Instead of working with correlations that are hard to interpret for Grolsch (which would be the case when using for example linear regression), we want to keep working with Time in Trade of a bottle. Recall that the updated return forecast expects every bottle of the sales- trade loss back according to a lognormal distribution. When in a certain week the returns are higher (or lower) than expected based on total sales, we need to define when these bottles were sold so they cannot return for a second time later based on the lognormal distribution. Several policies can be implemented to determine when these extra or fewer returns were sold. For example: on FIFO basis (if there are extra returns these are the containers that are in the market for the longest amount of time) or working backwards with the lognormal distribution of returns (If there are extra returns they are proportionally returned from the weeks of sales that are currently in the market). The return forecasting model should save how many of each week of sales are still expected to return. The solution therefore becomes harder to implement in the current container-planning model of Grolsch, but it might be worth it. The MAPE only decreases by around 0.5% for both Apollo and BNR compared to using only the seasonal lognormal distribution. Probably, the sales differences are already caught by the two period seasonality of the implemented lognormal distribution.

4.2.4. Trade Population

Grolsch currently uses a TP starting estimate of week 1 in 2017 which is based on average sales, expected TL, WiT and injections. This starting estimate was calculated years ago and how this is exactly done is unknown even for the people that currently work with the container-planning model.

Because we estimated for each week of sales when the sold bottles are expected to return (TiT-distribution), we also have the number of bottles that are expected to still be in the market. Note that we included a maximum lag length of 52 weeks in the estimated TiT-distribution. The returns that were expected after these 52 weeks based on the lognormal distribution, are scaled over the 52 weeks that are included. This results in an underestimation of the average TiT, which is the sumproduct of the weeks and their respective return percentages. The average TiT that is actually implied with the used lognormal distribution can be calculated as the limit of this distribution. As the percentage of the returns that is allowed after 52 weeks is an input and set to only X% for Apollo, the difference between the average used TiT of the 52 weeks and the implied average TiT is minimal. In general: The larger this allowed percentage after 52 weeks becomes, the larger the part that is scaled and the larger the difference between the used average TiT and the implied average TiT.

The approach to deal with this issue is to calculate the expected average TP based on the implied average TiT. This is done by multiplying the average sales over the same period that is input to the TP with the average implied TiT. This average is then compared with the average TP based on the used

TiT of the 52 weeks and the difference is added to the latter. This results in the TP-estimates for Apollo and BNR in Figures 39 and 40:

The TP of Apollo seems to remain stable throughout the years 2018 to 2020. The trend is only slightly downwards, which matches with the slightly downward trend of Apollo sales in the last years.

For BNR the trend of the TP is clearly up. This also matches with the sales trend of the last years.

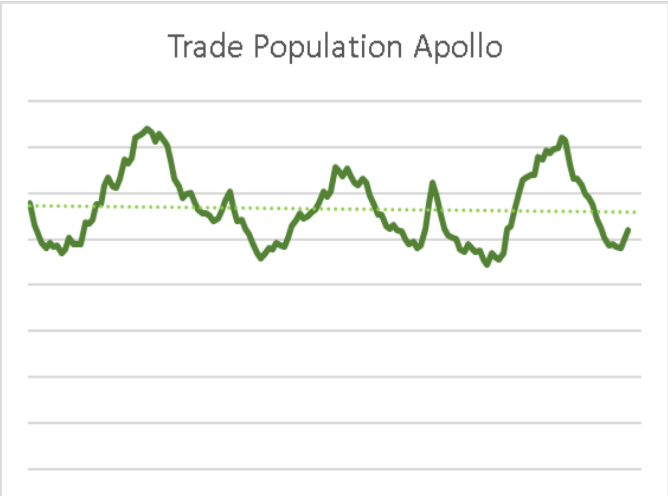

Figure 39: Trade Population Apollo

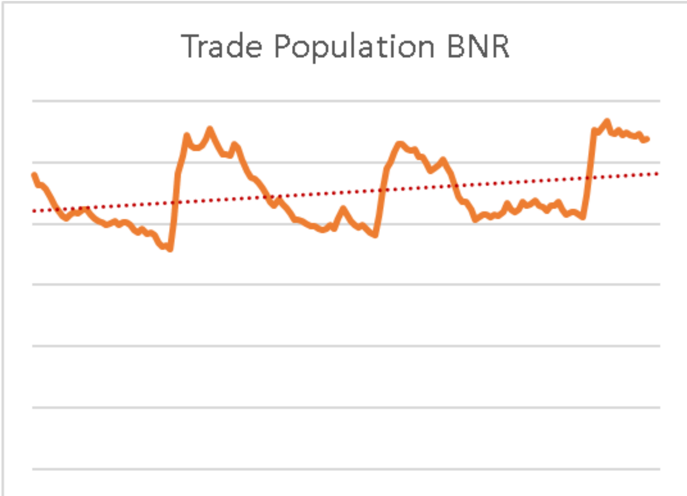

Figure 40: Trade Population BNR

The estimate can be used as a starting TP in the current model of Grolsch. The TP can then be updated with the formula that is currently already used:

$$\boldsymbol{Trade\ Population = Previous\ Trade\ Population\ + Sales - \ Returns - Trade\ Loss}$$

4.3. Improved return forecast

In the previous section we talked about the different parameters that are included in the return forecasting model. In this section the calculation of the improved return forecast is described. The inputs to fit the return forecast are (as mentioned in Section 4.1):

- Realized returns
- Realized sales
- Improved TL

Before the distribution of the time in the market and the impact of differences in sales can be estimated, first the Trade Loss must be subtracted from the sales of each week. This is needed to find how many bottles that are sold are expected to return of each week of sales. Then the following parameters have to be determined at once with the Microsoft Excel solver while minimizing the MAPE:

- Mean and variance of two lognormal distributions
- Extra returns caused by difference of total sales (Y) of Section 4.2.3
  (more total sales results in more returns and less total sales in less returns).

An extended example on how the return forecast is found

*Consider the sales of week 1 are 10 and of week 2 are 20,TL is 10% and we want the return forecast of week 3. The calculation becomes: $R_3 = Y(18\text{-}9) + b_1*18 + b_2*9$. Because the realized returns are known, the Excel solver can be used to find the best fit parameters Y, M and V by minimizing the MAPE. M and V are the mean and variance of a log-normal distribution. b_1 and b_2 follow from this lognormal distribution, as b_1 is the probability of return after one week and b_2 the probability of return after two weeks (probabilities at t=1 and t=2 from lognormal distribution with mean M and variance V). The remaining probability after two weeks needs to be scaled over the amount of weeks included in the calculation (b_1 and b_2) in the same proportions. So if $b_1 = 5\%$ and b_2 is 15% before scaling than 80% needs to be scaled and b_1 becomes 5%*1/0.2 = 25% and b_2 =15%*1/0.2 = 75%.*

This procedure results in the following fit on the data for Apollo and BNR in Figure 41 and Figure 42:

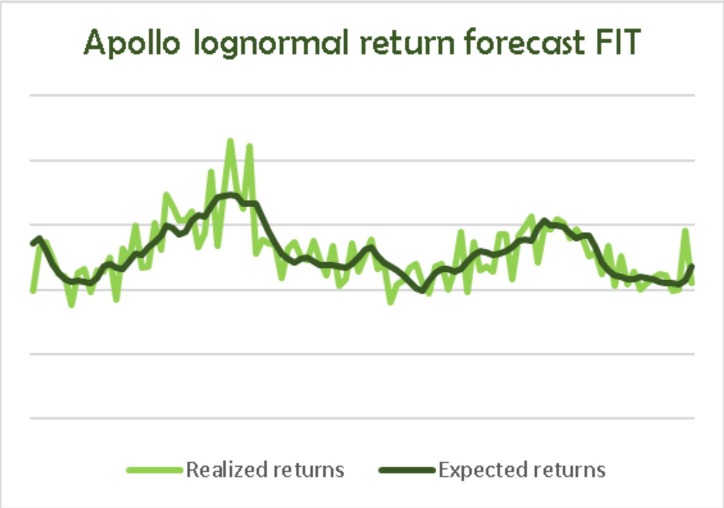

Figure 41: Apollo lognormal return forecast FIT

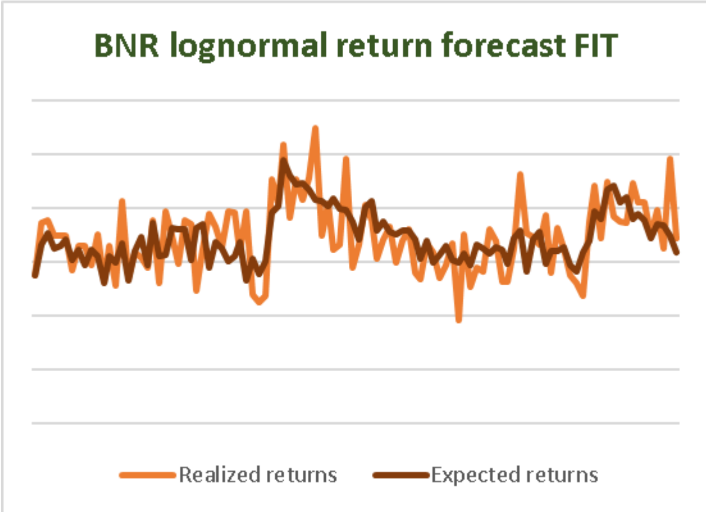

Figure 42: BNR lognormal return forecast FIT

In Table 9 the MAPE of return forecast fit on 2018-2019 is shown, which is the average absolute percentage error of the weekly returns. The return forecast fit is based on realized sales data. Grolsch’ current return forecast is explained in Section 2.3.2. Including two period seasonality in the return forecast is not only effective for eliminating the cumulative return forecast error but also the MAPE of the fitted data reduces further compared to the case where no seasonality is used.

	Return forecast FIT MAPE 2018-2019	
	Apollo	BNR
Current return forecast	25.1%	40.1%
Improved return forecast without seasonality	10.2%	14.8%
Improved return forecast with seasonality	**9.3%**	**12.5%**

Table 9: Return forecast FIT MAPE 2018-2019

The MAPE-fit is reduced significantly compared to the current method for both Apollo and BNR. This will likely lead to more accurate return forecasts and injection plans, but first we try to validate the return forecast in Section 4.4.

4.4. Validation

In this section we try to validate our return forecast assumptions and test the fitted TiT-distributions on test data to see if they are realistic. In this way we can see if the improved fit from Section 4.3 does indeed lead to a higher forecast accuracy. We do this by comparing our results from Section 4.3 with a small sample that was done in 2019 and with benchmarks found in the literature review.

4.4.1. Comparison with realized returns

To validate the improved return forecast we calculate the return forecast accuracy in 2020, with the TiT-distribution and parameter Y fitted on 2018-2019 respectively. Just like in Section 2.5, we forecast returns based on the realized sales and the forecasted sales. In this way the impact of the sales forecast inaccuracy can be seen. Are the returns accurately forecasted if the sales are accurately forecasted?

	Return forecast MAPE Apollo 2020 (based on realized sales)	2020 (based on forecasted sales)
Current return forecast	24.0%	26.1%
Improved return forecast	12.6%	14.3%

Table 10: Apollo return forecast MAPE

	Return forecast MAPE BNR 2020 (based on realized sales)	2020 (based on forecasted sales)
Current return forecast	35.7%	38.7%
Improved return forecast	14.8%	16.1%

Table 11: BNR return forecast MAPE

As can be seen in Table 10 and Table 11, the improvement from the fit also translates to improvements to the unfitted data. Even with possible changes in the return timing because of Covid-19, the improved forecast still manages to get a good weekly MAPE.

4.4.2. Benchmark comparison

In this chapter we are going to discuss the validity of the updated return forecasting model. We are interested in how the parameter values change if we change our assumptions about the underlying distribution structure and the maximum lag length.

Comparison to sample 2019

An indication that the return forecasting model is in the right direction could be to compare the model with a small sample from 2019. In 2019 a small sample of returned Apollo bottles were checked for their production batch code (which is present on every bottle). For every bottle an estimate for Time in Trade could be made. But the time in the warehouse was set to one week which is also questionable for all bottles.

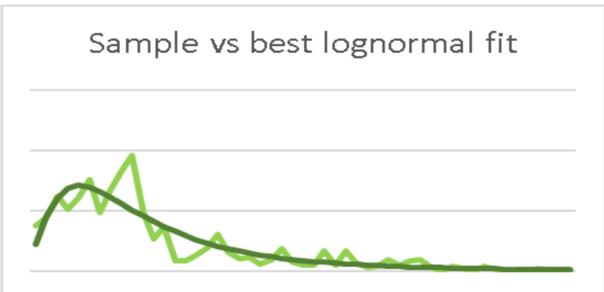

Figure 43: Sample TiT vs best lognormal TiT-fit

We conclude that this sample is too small to base conclusions on as this sample size is only X hl/X hl = X% of total average Apollo returns per week. We would expect a more smooth realized TiT-line. Using the exact return distribution of the sample (light green line in Figure 44), even for the period in 2019 that the sample was taken in, did not results in better performance than the proposed lognormal distribution.

Based on this graph we conclude that the real average TiT is shorter than the average TiT found in the sample (X weeks), as expected return line should be more towards the left.

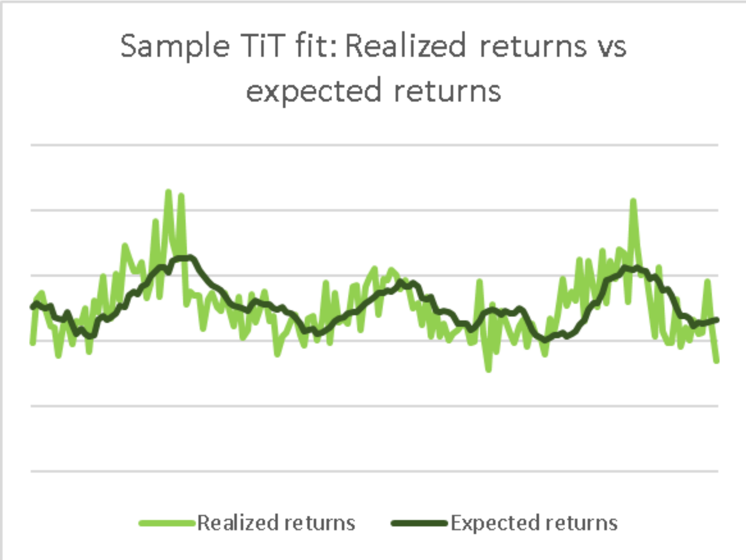

Figure 44: Return forecast fit with sample TiT-distribution

One explanation for the difference could be that the assumed fixed warehouse time before distribution of one week is too short. Another explanation is that it is unknown if this sample is representative in terms of customers and the frequency that they are visited. Besides, in several studies from the literature review the main bottle of a brewery turns out to have an average Time in Trade of around 8-9 weeks (Van Dalen, Van Nunen & Wilens (2005); Widi(2009)).

4.5. Conclusion

In this chapter we first concluded that the return forecast for Grolsch can be well modeled as a finite Distributed Lag Model, with lognormal distributed lag weights. This means that the time in the market (Time in Trade) of a bottle is assumed to be lognormally distributed.

We then calculated the parameters Trade Loss and Internal Loss, from which Trade Loss is needed as input for the return forecast. Before the return forecast can be made, the Trade Loss per week needs to be subtracted from the sales per week. Then the expected returns **from** that week of sales are known and can be distributed **to** the weeks after the sales. First we used one TiT-distribution for the whole year. This means bottles have the same expected time in the market during the whole year. We concluded assuming only one TiT-distribution for the whole year results in a biased forecast. The returns are structurally over-forecasted in the off-season (quarters 1 and 4) and under-forecasted in peak-season (quarters 2 and 3) for both Apollo and BNR. We therefore included seasonality by using two separate TiT-distributions for peak- and off-season.

Because Grolsch gets more bottles back when more clients are visited, we added another parameter to the Distributed Lag Model. This parameter is about the difference in total sales, which means that extra returns are expected in the current week if more is sold in the current week compared to the previous week. These parameters are found using the Microsoft Excel solver.

The MAPE of the improved return forecast is significantly reduced in comparison to the current return forecast. The MAPE for Apollo decreased from 25% to 13% and the MAPE of BNR from 40% to 16%.

As the timing of the returns is assumed to follow the TiT-distributions, the amount that is not yet returned is assumed to be known as well. So after we found the TiT-profiles, we were able to calculate the expected Trade Population for both Apollo and BNR.

The improved return forecast of this chapter is used as input to a new injection planning calculation in the next chapter.

5. Improved injection planning

In the previous chapter we improved the return forecast for Apollo and BNR. The goal of this research is to improve the planning of the procurement of new bottles. In this chapter we therefore use the improved return forecast from Chapter 4 as input for an improved injection planning model. We model the injection planning problem of Grolsch based on the purchasing policy for container purchasing proposed by Kelle and Silver (1989b). First we describe the goal of the improved injection planning model, then we give our assumptions and describe how we calculated the different inputs that are needed for this injection planning model. In the final part of this chapter we give the mathematical formulation of the model.

5.1. Goal of the new injection planning model

Grolsch currently does not directly use uncertainty in the planning of new buys (injections). There is a measure of safety stock in place (Days of Cover), but it is based on by experts' opinions and the values are low for peak season.

The improved parameter calculation and return forecast can be incorporated in the current container-planning model of Grolsch to calculate an improved injection plan. However, the safety stock (of empty bottles) that is included in the current injection planning model of Grolsch is based on experts' opinions. After the updating of the parameters and improving the container return forecast, there is still a sales forecast error present and therefore also the returns are uncertain. Grolsch does not know what service level is implied by using the current values for DoC. The risks of lowering the DoC in the peak season (so that the model plans less injections and less budget is needed) are also unknown. As mentioned in the chapter introduction, we model the injection planning problem based on the purchasing policy model proposed by Kelle and Silver (1989b). Because their model uses the concept of "cumulative net demand", which is approximately normally distributed, the standard formula for service level and safety stock (Section 3.4) can be used. Cumulative net demand of a certain week is the sum of demand-returns up to this week. An example to make the term cumulative net demand more clear:

In period 1 the sales are expected to be 100 and the returns 60, then 40 new bottles are needed to fulfill the sales forecast (when we assume that returned bottles can be used for production in the same week). The net demand of period 1 is 40. If in period 2 the sales are 200 and the returns 120, the net demand of period 2 is 80. The cumulative net demand up to period 2 is 40+80=120. This means that up till period 2, 120 new bottles need to be injected. The injections of 120 bottles total can be planned in every period up to period 2.

In the model of Kelle and Silver(1989b) the cumulative net demand up to every period must be satisfied. This means that the net demand of week 1 from the example above needs to be injected in week 1. The net demand of week 2 can be injected in week 1, or in week 2.

The advantage of the model of Kelle and Silver is that it is relatively simple, but can be extended with constraints of production capacity, service level or by using different assumptions.

Safety stock

In Figure 45 is shown that unlike in standard inventory management, the stock that is replenished from procurement is the "raw material" stock (empty bottle stock) instead of the "serviceable inventory" (full bottle stock). The standard scenario is having demand and supply for only the serviceable inventory. In our case we have to think about how to model the translation from demand to when bottles are needed.

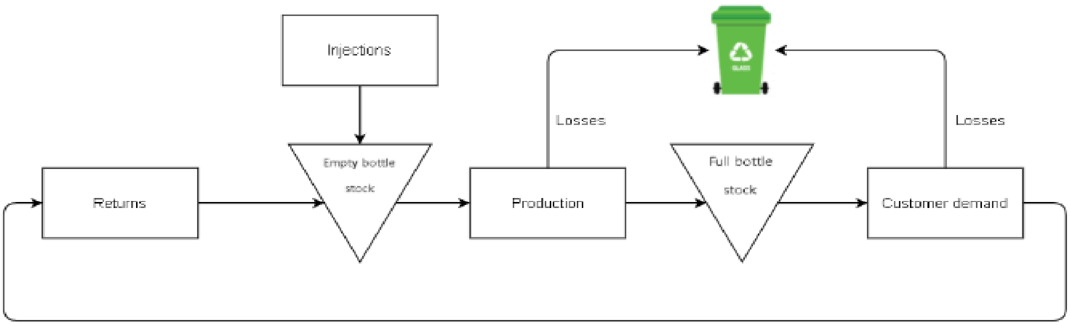

Figure 45: Difference Grolsch with remanufacturing

One thing that is important is that the safety stock at the serviceable inventory is in reality included in the production plan. On average the safety stock covers one week of demand at the serviceable inventory. This means that the production has to be done on average one week before the expected demand. Note that safety stock at the serviceable inventory is to protect against uncertainty in demand but also against problems during production.

Safety stock at the empty container stock is for protection against uncertainties in supply. The empty container stock also includes protection against internal losses, but the internal loss percentage is assumed to be known and does not contribute to uncertainty. The uncertainty in supply is assumed to only include uncertainty in returns as suppliers are considered reliable on the long term. What makes it complex is that uncertainty in returns is based on return forecast inaccuracy but also on uncertainty in demand. Because if more bottles are sold, more bottles are expected to return. The model of Kelle and Silver uses the cumulative net demand which includes both the uncertainty in supply and demand. We can assume that the demand is known one week before it has to be satisfied and that the production plan can be adjusted accordingly. The net demand of a certain week then becomes the demand of the next week minus the returns of the current week.

5.2. Assumptions and injection planning model inputs

In this section we describe our assumptions and the inputs that go into the proposed injection planning model. We make the following assumptions:

1) Amounts are expressed in Grolsch' main calculation unit: hl. One bottle = 0.003 hl.
2) To determine the cumulative net demand, the return times of Apollo and BNR are assumed to follow the TiT-distribution solutions of the DLM with seasonality fit on 2018 and 2019 outlined in Chapter 4.
3) The time horizon is finite (78 weeks) and split up in deterministic periods of one week. The impact of this is that it might be hard to get a highly accurate injection plan. It is easier to get the return forecast and injection planning more accurate if longer periods

are taken, but Grolsch uses weekly periods across the whole organization so we stick to weeks.

4) The returns in a certain week can be used to fulfill production in that week. This assumption means there is no sorting delay for returned bottles. Bottles could be sorted with priority. This is also currently assumed in the current injection planning model and is also realistic. The impact of this assumption is that bottles can be filled earlier than compared to when there is a sorting delay. Therefore: the longer the time before a bottle can be filled again, the bigger the total population of that bottle type needs to be.
5) Production in a certain week can be used to fulfill customer demand one week later. This is a realistic assumption because the current safety stock of filled bottles is on average five days and there is a small buffer against problems during production. The impact of this assumption is that we do not need to translate changes in sales in changes in the production plan. The past three years the sales per week did not exceed the theoretical production capacity per week. The production demand becomes the external demand shifted by one week. As this assumption has a big impact on when bottles are needed, we validate this assumption in Section 6.1.
6) The bottles are ordered one and a half years before the bottles are needed. With this time horizon, the supplier has enough flexibility to plan their own production accordingly. Grolsch does not directly work with replenishment lead times in their current container-planning model. The outcome of Grolsch' model is when the bottles are needed, not when they should be ordered.
   In the model we see the lead time as the difference between t=0 and the period in which the new bottles are needed(t=t). Lead time is therefore increasing with t.
7) Reused bottles are as good as new ones.
8) Storage capacity of empty bottles is unrestricted, as they can be stored outside and there is enough space available. This impacts the holding costs as there is no real risk of purchasing more bottles apart from investment costs. The new bottles are also well packaged against damage, so breakage is also no big risk.
9) More bottles than the demand are needed to fulfill the demand because of internal losses. The internal loss percentages are assumed to be X% for Apollo and X% for BNR (Section 4.2.1).

Normal approximation of the cumulative net demand of bottles

As showed by Kelle and Silver (1989b) the cumulative net demand can be approximated by a normal distribution when returns per week are sums of binomials and mixed binomials and demand is also normally distributed (as mentioned in the literature review). By using the normal distribution the safety stock can be calculated using the standard formula from Section 3.4. It is however hard to theoretically obtain the parameters for the normal distributions, that need to be estimated for each period in the time horizon. The injection planning for an upcoming year is already determined in summer of the current year. The return forecast for the upcoming year heavily depends on the sales forecast for the upcoming year.

Silver, Pyke and Thomas (2017) mention that "for purposes of establishing the safety stock of an individual item (to provide an appropriate level of customer service), we will need an estimate of the standard deviation (σL) of the errors of forecasts of total demand over a period of duration L (the replenishment lead time)". We have the sales forecast for 2021 available. To check if we can indeed assume a normal distribution for the cumulative net demand, we first check if the sales-forecast errors can assumed to be normally distributed.

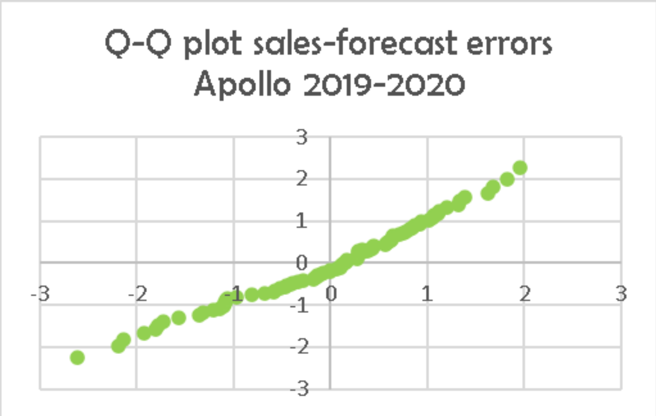

Figure 46: Q-Q plots sales forecast error Apollo

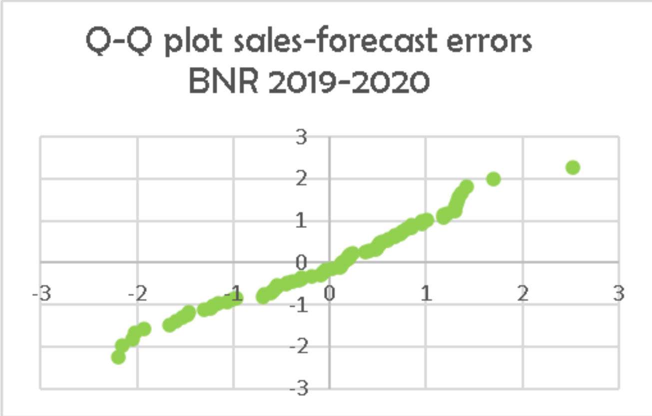

Figure 47: Q-Q plots sales forecast error BNR

Goodness-of-fit	Apollo	BNR
Df	10	10
Mean	-X%	-X%
Std. dev.	X%	X%
p	0.05	0.05
Critical value	3.94	3.94
Test statistic	3.31	3.72

Table 12: Goodness-of-fit

Based on the QQ-plots and the goodness-of-fit test, we assume the forecast errors to be normally distributed. One thing to note is that the mean forecast error for Apollo is -X%. This mean error is based on the sales forecast for 2020. In the summer of 2020, when the Covid-19 situation was looking brighter, the new sales forecast for 2021 was made. It was expected that at the start of 2021 things would be close to the old normal. We assume the sales forecast for 2021 will likely also be too high. The sales forecast for Apollo over 2021 is more than X hl, which is more than the realized sales in 2019 and 2020. As the mean weekly forecast error of Apollo was -X% in the second half of 2019 and 2020, we assume the weekly forecast error of the second half of 2020 and 2021 to be equally distributed (~N(-X%,X%)). Besides, the last 20 weeks of 2020 are included in the new sales forecast, and the average forecast error over these 20 weeks (comparing this new sales forecast with realized sales) was also close to -X%, namely -X %. However for normal years without Covid-19, the expected sales forecast error should be close to 0%.

With the sales forecast for 2021 and the forecast errors, the sales for 2021 and corresponding returns can be simulated. De Brito and Van der Laan (2009) used simulation to estimate the mean and variance of the net demand during the lead time with the formulas proposed by Kelle and Silver (1989a). As Kelle and Silver (1989a) mentioned: “Neither the distribution nor even the standard deviation of the relative error could be expressed in appropriate analytic form”. Simulation of demand and returns therefore seems an appropriate method to calculate the distribution parameters of the normally distributed cumulative net demand and gave De Brito and Van der Laan (2009) clear and useful results. These distribution parameters (mean and variance) of the normally distributed cumulative net demand need to be determined for every period in the time horizon. This is needed because the sales forecast is different for every period, and every period has its own cumulative net demand normal distribution. These normal distributions are then used to determine safety factors per period, that are used in the new injection planning model’s constraints.

Simulation setup

Here we simulate the sales and corresponding returns for week 33 of 2020 till week 52 of week 2021.We obtain the cumulative net demand for each period in each simulation run and can estimate the parameters of the approximately normally distributed cumulative net demand. As input we need the sales forecast and error distribution.

Sales

The injection planning period runs from week 33 in 2020 to week 52 of 2021, while the budget is purely determined for 2021. As the returns for 2020 still depend on the sales from the same week a year before, we include the realized sales from week 33 of 2019 to week 33 of 2020 as simulation input. For the forecast period we use the sales forecast as input. The sales from week 33 of 2020 till week 52 of 2021 are simulated based on the sales forecast and the normally distributed sales forecast errors.

Returns

We use the updated TL values for Apollo and BNR together with the TiT-distributions obtained in Chapter 4 to determine how much and when the sold bottles come back to the brewery.

The forecasted sales per week, even with an negative expected sales forecast- error, are sums of millions of bottles. We don't simulate the return arrivals, but just use the TiT-distribution as it is expected that with millions of bottles the simulated return curve is the same as the distribution they are taken from. If sales would have been much lower we would have simulated the returns as well.

Even with an expected error of 0 and some standard deviation, some simulation iterations will have infeasible returns based on the sales. The total returns over a period cannot be more than the sales they come from, and the loss is also not expected to be lower. The uncertainty in returns should come from timing and not totals.

Simulation results From theory we know the cumulative net demand can be approximated by a normal distribution when the demand is normally distributed and the returns are seen as mixed binomials (Kelle and Silver, 1989a). When we plot the distributions of 10000 simulation runs, we see that we can indeed assume a normal distribution for the cumulative net demand (shown for t=1, but this is the case for every t in the time horizon).

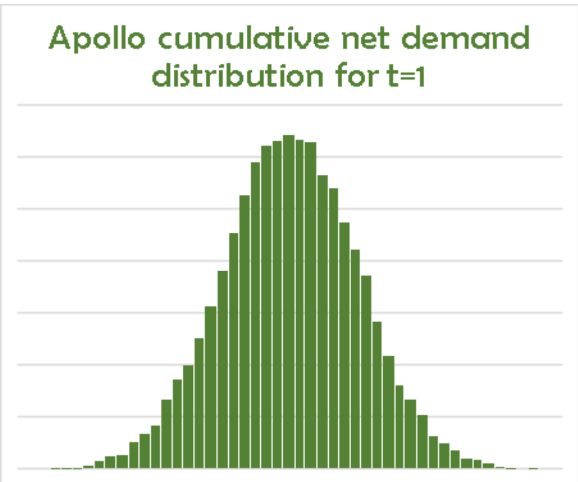

Figure 48: Normal distribution cumulative net demand of Apollo

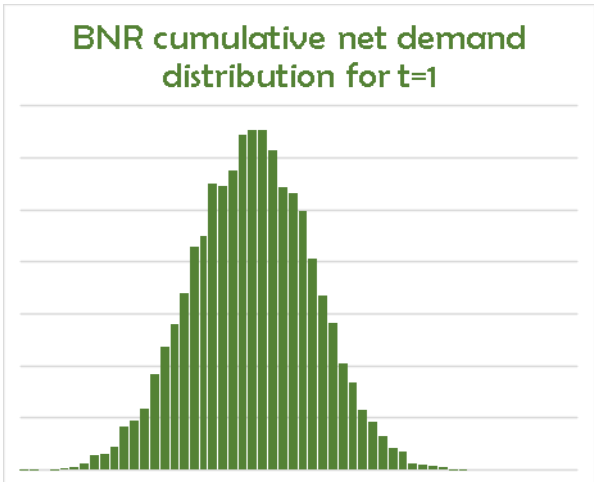

Figure 49: Normal distribution cumulative net demand of BNR

Cumulative net demand of bottles over the time horizon:

In Figure 50 and 51 we show the cumulative net demand for Apollo and BNR respectively. The line indicates in each time period how much new bottles are needed to be injected up to that period. For example: the peak lies at t=50 (week 30 of 2021), with the corresponding value X hl. This means that up to t=50 X hl needs to be injected and after t=50 no more injections are needed.

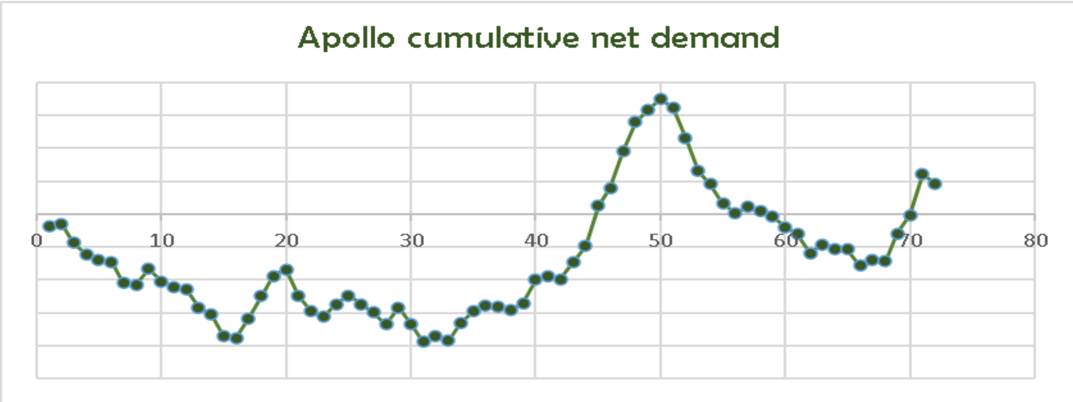

Figure 50: Apollo cumulative net demand

For Apollo we see that the mean cumulative net demand is negative for every period up to t=45, which is week 25 of 2021. This means no new bottles are expected to be needed yet. This is logical as peak season (in terms of sales and production) is already coming to an end at the start of the time horizon. The next peak season starts from t=40 onwards, which is why the graph is increasing from that point on. More bottles are needed for production while there is a significant amount lost in the first 40 weeks of the time horizon. Also the returns of the sales during peak season follow later and cannot be used to fulfill demand on average nine weeks for Apollo. The maximum of the line at t=50 is the total amount of injection needed over the time horizon of 72 weeks. This total amount has to be injected before t=51. (An period's cumulative net demand needs to be satisfied in any period up to or in the specific time period.)

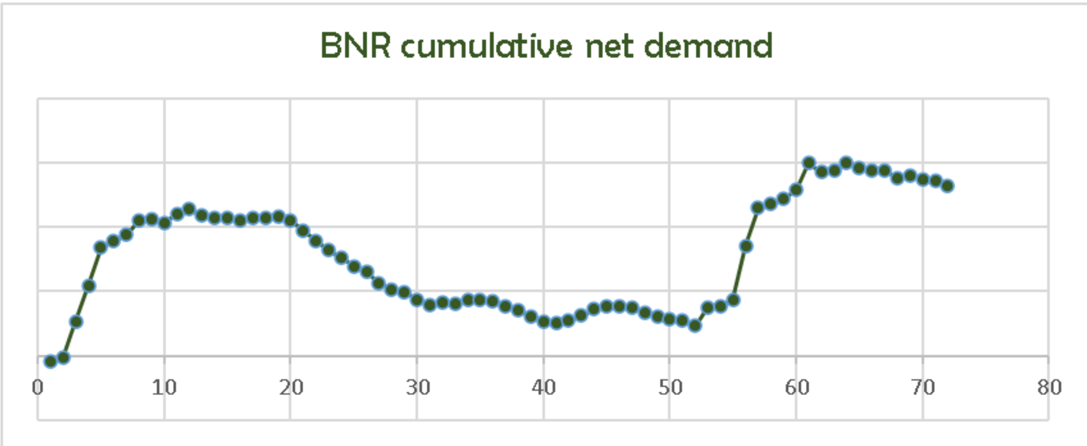

Figure 51: BNR cumulative net demand

For BNR we notice that the peak season is early in the time horizon, because the graph increases to around X hl right from t=0. This is logical as t=0 is week 33 and the peak season for BNR starts around week 36 (t=56) with the introduction of the specialty beer Herfstbok. Sales are expected to increase for BNR as Grolsch is increasing its specialty beer segment. A bigger amount injection is needed to support this sales growth and the cumulative net demand does not drop under 0 hl again.

Service level

Another input to the model is the service level. The service level is the probability that the demand in a replenishment cycle can be fully satisfied (no stockout occurs). Grolsch uses a service level of X% for final products. In our lot sizing model this means that for each period in the time horizon the probability that the cumulative net demand (the net demand up to the specific period) can be fully satisfied should be X%.

Safety stock calculation

Figures 50 and 51 only show the mean cumulative net demand. Because the cumulative net demand in each period can be approximated with a normal distribution, we can use the standard deviation to calculate the safety factor and safety stock for each period.

Capacity constraints

The model of Kelle and Silver (1989a) can be used, but it is quite simple and does not consider capacity restrictions: "*We are effectively assuming that container requirements match the demand requirements period by period. In other words, we are ignoring the complicating effects of production (container filling) scheduling logic, including capacity considerations.*"

Generally the model of Kelle and Silver (1989b) would work fine as the service level constraint should be met up to every period. In practice however, production (container filling) with new bottles is restricted as there is only one production line that can use new 30cl bottles. We need to extent the model of Kelle and Silver with some form of production capacity constraint, to make sure the model does not implicitly plan too much production by ordering containers for later in the time horizon to save holding costs.

Production Line 2:

On Production Line 2 both Apollo and BNR can be filled. The filling capacity of Production Line 2 is X hl per week. Production Line 2 is the only production line that can use new 30cl bottles. The X hl is therefore also the maximum amount of hl of new 30cl bottles that can be used in a week. (The production with new bottles of Apollo and BNR can *together* not exceed X hl.)

Storage capacity

As new- as well as returned bottles can be stored in the large crate park outside of the brewery, we can assume there is no storage capacity restriction. Besides, new bottles are not stored in crates, which means that there is less space needed to store them in comparison to returned bottles.

5.3. Costs

Before the new injection planning model can be implemented and evaluated the costs need to be specified. In the new injection planning model we only need holding costs as we work with a service level as input. The holding costs are needed because otherwise the injection planning model will plan all injections at the start of the time horizon so the bottles are already stored and available. The injection planning model outputs an injection plan that fulfills the required service level with minimal holding costs. After the plan is output by the model, we calculated the total expected costs of the plan in a simulation in Chapter 6. In this simulation we also take into account changeover- and stockout costs. In this section we describe the different costs and how they are calculated.

Holding costs

Holding costs are the costs associated with holding a unit on stock for a certain period. There are many factors that should be included in the holding costs, but most of them are hard to put a value to. For example risks of breaking (by bad weather or handling by workers) , costs of space and opportunity costs of the investment.

In outside storage there is (still) a lot of space for empty (newly purchased) bottles. As mentioned before, new bottles are generally well packaged against damage. They are not put in crates either, which reduces the used storage space by X% compared to returned bottles.

Because space is not limited for empty bottles (they can be stored outside), the main cost for ordering too much material is the cost of capital. Money that is used for investments that are not needed yet cannot be used for other investments. For example it cannot be invested against the risk free rate.

Another part of the costs of ordering too much material are the depreciation costs. For (extra) empty bottles the depreciation costs can however considered to be minimal when they are only stored and not used for production yet. With extra bottles, we mean injected bottles that are not needed for short-term production. Also, when these new bottles are used for production instead of returned bottles, the depreciation of the new bottles would start, but the depreciation of the returned bottles slows down. And new bottles, unlike new cars, are valued the same for years to come as long as the bottle design does not change. As new bottles are also well packaged against damage, we assume that the depreciation costs of extra new bottles are zero.

For holding costs per week we use the Weighted Average Cost of Capital (WACC). The formula for holding costs is:

$$h = WACC\ per\ week * purchasing\ price\ of\ one\ hl\ of\ containers * Inventory\ level\ in\ hl$$

Grolsch uses a weekly WACC rate of X%.

The purchasing price of one hl of new 30cl bottles is X euros for Apollo and X euros for BNR.

Production inefficiency costs

In the final months of 2020, Grolsch had problems with the BNR bottle and had to produce more frequently with smaller batch sizes. This resulted in higher changeover costs and the overall production volume that can be filled was lower in the same amount of production time. The average batch size on Production Line 2 is X hl. When there are less empty bottles available than needed to fulfill the demand, the batch sizes generally become smaller. Grolsch has to produce more frequently because the bottles need to come back first. This results in an increase in changeover costs.

The average changeover time on Production Line 2 is X minutes and the average costs per hour are X euros. So we can assume the costs of moving X hl to the next week costs X euros. These costs can be used in the costs savings calculation in Chapter 6.

Stockout costs

It can be the case that the production cannot be done on time to prevent a stockout occurrence. Recall that a stockout at Grolsch does not necessarily means a stockout at the customer. The general rule is that one week of stock out of a certain Stock Keeping Unit (SKU) is not yet a problem. It is hard to determine the exact value stockout costs as the impact at the customer is hard to estimate. Stockouts are always communicated with the customer and often mean backorders and not lost-sales. In deliberation with Grolsch we assume that demand can still be satisfied one week later. If there are not enough bottles in the next week to add this additional production in the next week, the amount that cannot be satisfied is considered lost sales. This means there are changeover costs if the backordered demand can be fulfilled in the next week, and stockout costs for the amount that cannot be satisfied.

5.4. Mathematical formulation

Below we describe the proposed injection planning model of Section 5.1 and 5.2 in terms of sets, input parameters, objective function, variables and constraints. We use the model of Kelle and Silver (1989b) and extended it with capacity constraints. We program the model in AIMMS, a solving tool we are familiar with.

Sets

$\boldsymbol{C\ =\ Containers\ types\ \ (c=1,2)}$

The container types that are included are Apollo (1) and BNR (2). More container types could be added to the set in the future.

$\boldsymbol{T\ =\ Time\ periods\ in\ weeks\ (t=1,2,..,78)}$

The time horizon is now set to 78 periods (weeks), but in some years this time horizon may be a little shorter or longer depending on when the container budget for the upcoming year is made.

Input parameters

$\boldsymbol{h\ =\ Inventory\ holding\ cost\ of\ one\ hl\ per\ week}$

This input parameter is calculated with the WACC in Section 5.3 and equals X euros per hl per week. Holding costs are needed so that the model does not plan all injections at the start of the time horizon so they are already available for whenever they are needed in the time horizon. This parameter is used in the objective function.

$\boldsymbol{ei_{c,0}\ =\ Initial\ inventory\ of\ empty\ containers\ of\ type\ c}$

This input parameter is needed because there is already a stock of empty containers of Apollo and BNR present at the brewery. Less new bottles are needed if the initial stock of empty bottles is larger.

$\boldsymbol{fi_{c,0}\ =\ Initial\ inventory\ of\ filled\ containers\ of\ type\ c}$

The same goes for full bottle stock. Less extra production is needed to fulfill demand because there is already a stock of filled bottles present at the brewery.

$\boldsymbol{nd_{c,t} = Net\ demand\ for\ container\ type\ i\ in\ period\ t}$

Net demand is the demand minus the returns.

$\boldsymbol{x_{c,t} = Mean\ cumulative\ net\ demand\ for\ container\ type\ i\ up\ to\ and\ including\ period\ t}$

The mean cumulative net demand is found using the simulation explained in Chapter 5. The sales and returns are simulated and the amount of new bottles that are needed to be injected up to and including each period to fulfill the demand are found. The cumulative net demand can be assumed to be normally distributed as seen in Section 5.2. Each period in the time horizon has its own cumulative net demand normal distribution.

$k_{c,t} \ = \ Safety\ factor\ for\ cumulative\ net\ demand\ of\ container\ type\ c\ in\ period\ t$

Just making sure the mean cumulative net demand of bottles is met does not protect against uncertainties in supply and demand. The safety factor is based on a X service level.

$$z_{c,t} = E[x_{c,t}] + k_{c,t} * Var[x_{c,t}] \qquad For\ c = 1,2\ \&\ t = 1,2,..,78$$

$Z_{c,t}$ is the order-up-to level. The amount that needs to be ordered for period *t* is the order-up-to level minus the initial inventories and the stock in the pipeline (planned injections in the time periods before *t*). Why do we use the order-up-to level? The injection plan is a plan when bottles are needed. The order-up-to level makes sure that enough injection is planned in each period. The order-up-to level is calculated by taking the mean cumulative net demand of period t plus safety stock to protect against uncertainties in demand and supply. The mean cumulative net demand follows from the simulation outlined in 5.2.

$f \ = Filling\ capacity\ of\ new\ bottles = X\ hl$

Only X hl per week can be injected because new 30cl bottles can only be used on Production Line 2 and the capacity of this line is X hl.

Decision variables

$Q_{c,t} \ = \ Planned\ injection\ of\ containers\ of\ type\ c\ in\ period\ t$

Planned injection is when the bottles are needed, not when they are ordered. The injection plan is already communicated to the bottle suppliers a year before the bottles are needed. The supplier then makes sure the bottles are available when Grolsch needs them.

Other variables

$EI_{c,t} \ = \ Inventory\ of\ empty\ new\ containers\ of\ type\ c\ at\ the\ start\ of\ period\ t$

This variable is kept track of by the model over the time horizon and is used to calculate the holding costs.

Objective function

$$Min\ Z \ = \sum_{t=1}^{78} (h * (EI_{1,t} + EI_{2,t}))$$

This objective function makes sure the model minimizes the sum of inventory holding costs of Apollo and BNR. Backorder costs are not included here as this injection planning model is based on a service level and some backorders are allowed. Besides, sales have the be simulated in order to find the backorder costs. This model outputs an injection plan based on a service level of X, and the corresponding backorder costs are found in Section 6.3.

Constraints

$$Q_{c,t} \geq 0 \qquad For\ c = 1,2\ \&\ t = 1,2,..,78$$

Planned injection should be non-negative. Especially because net demand (demand-returns) can be negative if returns exceed the demand, this constraint is necessary to make sure the injection planning model does not plan negative injection.

$$\sum_{j=1}^{t} \boldsymbol{Q}_{c,j} + \boldsymbol{ei}_{c,0} + \boldsymbol{fi}_{c,0} \geq \boldsymbol{z}_{c,t} \qquad \boldsymbol{For}\ c = 1,2\ \&\ t = 1,2,..,78$$

Cumulative net demand and safety stock should be satisfied up to or in every period t. Demand can also be met out of the initial inventories.

$$\boldsymbol{Q}_{1,t} + \boldsymbol{Q}_{2,t} \leq \boldsymbol{f} \qquad \boldsymbol{For}\ \ t = 1,2,..,78$$

Filling capacity (hl) per week of new bottles. More new bottles cannot be filled so if more new bottles are needed to fulfill the demand, they are needed earlier in the time horizon.

Balance equations:

$$\boldsymbol{EI}_{c,t} = \boldsymbol{EI}_{c,t-1} + \boldsymbol{Q}_{c,t} - \boldsymbol{nd}_{c,t} \qquad \boldsymbol{For}\ c = 1,2\ \&\ t = 1,2,..,78$$

Calculation of inventory with new containers available and net demand goes out for each period t. Needed to calculate holding costs.

The injection planning model is a mixed linear program and can be solved in the program AIMMS that uses the well-known CMPLX solver. The exact solution of the model without capacity constraints can also be found using the Wagner-Whitin algorithm (Kelle and Silver, 1989b). The single item capacitated lot-sizing problem is already NP hard as the running time is of the order O(4^t) (Bitran & Yanasse, 1981). The multiple item case under capacity restrictions like our case, is also NP- hard. However, the problem for Grolsch is not that big as the number of container types and capacity constraints is low. It only takes a total of 15 minutes for the whole process from simulating the cumulative net demand to calculating safety stock and running the injection planning model. The model can be solved exactly in a reasonable amount of time. As this only has to be done once in a year this is considered acceptable.

The output of this model is an improved injection plan for the next 78 weeks based on a service level of X. We use this improved injection plan in Section 6.3 to calculate the expected costs and compare these costs to the current injection planning model to see if savings are possible.

5.5. Conclusion

In this chapter we proposed a new container purchasing policy, which is based on the standard safety stock formula (for the empty bottle stock) instead of the current method of a fixed amount of safety stock (which covers 5 days of production demand). We used the model of Kelle and Silver (1989b) for this purpose, extended with some production capacity constraints. The main thing to take away from this chapter is the concept of cumulative net demand. The net demand of a certain week is the demand for empty bottles minus the returned bottles. So the net demand is the amount of new bottles that is needed to fulfill demand. The cumulative net demand of a certain week is the amount of new bottles that needs to be injected up to this week. The injections can be planned in any period up to when they are needed. The model is described mathematically in Section 5.4. and can be solved using the program AIMMS.

The over- and understocking costs are also explained in this chapter. The main costs are: the holding costs, calculated with the Weighted Average Cost of Capital (WACC), the changeover costs for delaying a production and stockout costs. All these costs are used in the next chapter in which we simulate sales and returns to find the expected costs of three different injection plans. The injection plan based on the injection planning model of this Chapter 5 is one of these three. This new plan is compared with two other injection plans to see if there is a costs reduction possible. One of these other plans is output by the current injection planning model of Grolsch with the current return forecast, and the other plan is based on the current injection planning model with the new return forecast.

6. Improved injection planning results

The new injection planning model has been described in previous chapter. In this chapter we present the results that can be achieved by adapting this model. To make sure our model is in line with reality, first the assumption of when bottles are needed is validated. We then see what service level Grolsch currently implies with the used safety stock (Days of Cover) so it can be compared to the required service level. Afterwards, three injection plan scenarios (based on current- and proposed return forecasting- and injection planning models) are tested in a simulation to obtain the expected costs of each scenario.

6.1. Validation

Because the translation from sales error to production plan changes is hard to make, we made assumptions about when bottles are needed. To validate these assumptions in the simulation model we check if the simulated demand for bottles is realistic compared to the actual production plan.

To see how the production is implied, we can plot the (empty bottle) stock on hand. The realized stock on hand based on the production plan can be compared with the stock on hand values from the simulation model. To test our assumption of producing one week before demand, we compare the stock on hand output with the production plan for 2021. We don't compare with realized production, because the realized production can be changed in various ways for example because of bottle unavailability. While we assume safety stock always has to be in place, which is also the case in the production plan of Grolsch, the realized production may be lower because demand was fulfilled from safety stock. Even if bottles where available, production problems might have contributed

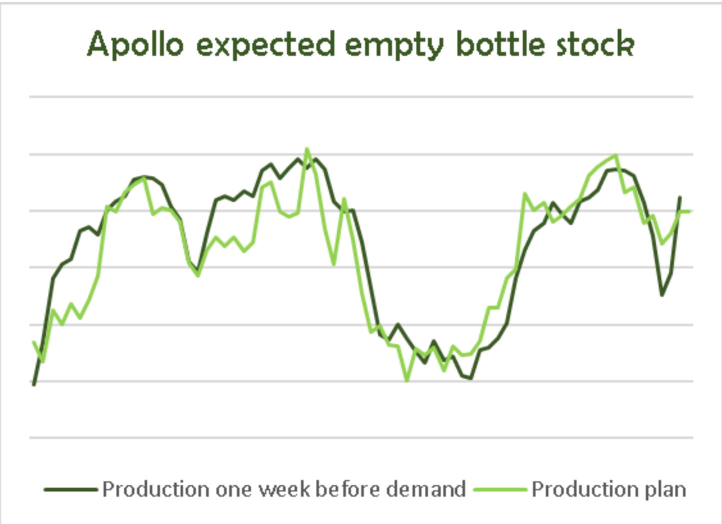

Figure 52: Empty bottle stock Apollo

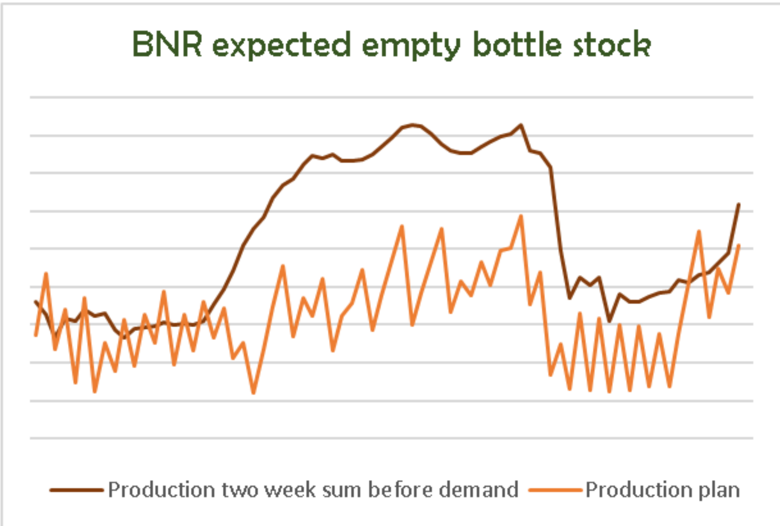

Figure 53: Empty bottle stock BNR

towards this usage of safety stock. We therefore use the production plan for validation. For Apollo we consider the fit reasonable and the holding costs are only slightly higher than using the production plan. There are no periods with over and underproduction in the same period in both years.

For BNR we consider the fit not good enough. The first point of attention is that production with BNR is only done in uneven weeks. Instead of producing one week up front, for BNR this should be two weeks. Besides, from the data of the last three year we notice that in the beginning of a new year around X hl extra is produced with BNR than needed based on sales. Around week 40 around X hl less is produced. With these new assumptions included in the simulation model the fit becomes reasonable as shown in Figure 54:

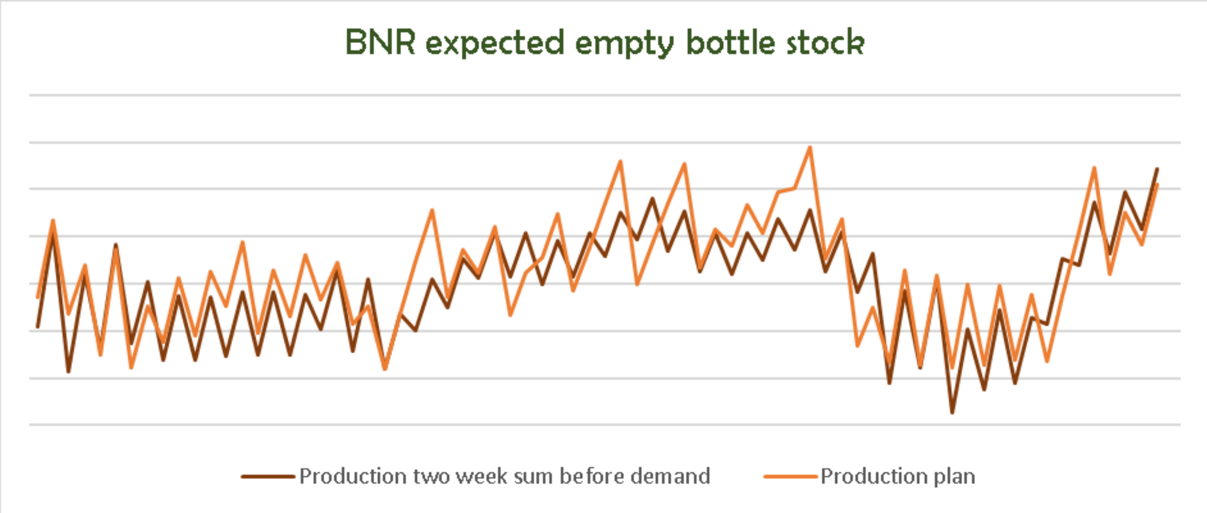

Figure 54: BNR empty bottle stock new assumption

6.2. Current implied service level

The service level that Grolsch implies with the current injection planning model is unknown. Recall that the current injection planning is based on the safety stock measure DoC. Currently a standard value of five working days of planned production should be possible with the stock of empty bottles. If less stock is available than needed to fulfill the required five working days of planned production, the current injection planning model plans an injection so that the safety stock is met again. It was unknown what service level this current policy implies. In Section 5.2 we simulated sales and returns and found the approximate normal distributions of the cumulative net demand. Because we know the required safety stock of the current model, we can calculate the implied service level when we compare the required safety stock with the cumulative net demand.

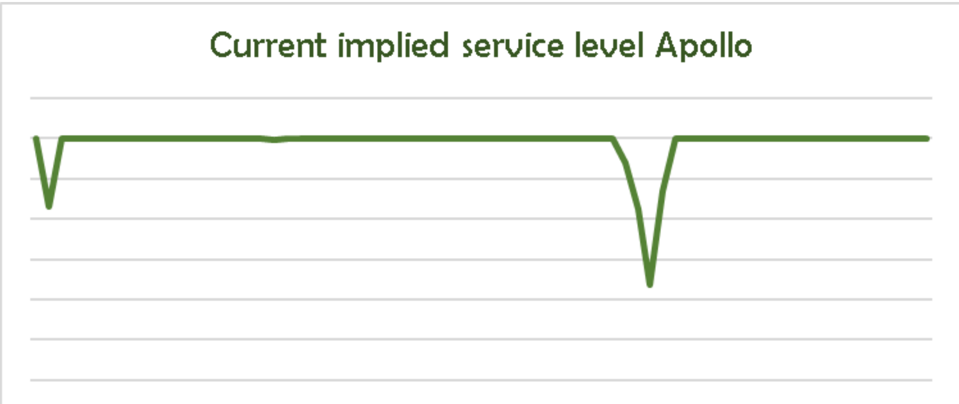

Figure 55: Current implied service level Apollo

As shown in Figure 55, Grolsch implies a service level (X) with the current injection plan. The feeling is that this current policy results in high holding costs, and a slightly lower service level might decrease holding costs significantly. The current injection planning model has safety stock in periods were there it is certain that there are enough bottles available. In these non-peak sales periods, less safety stock is needed based on the new injection planning model of Chapter 5. In the next section we see if cost savings are indeed possible.

6.3. Costs savings

In this section we calculate the expected holding-, changeover- and stockout costs based on three scenarios (three different injections plans). These three scenarios are:

1) The current injection planning model with the current return forecast from Section 2.4.4. (CIP + CRF)
2) The current injection planning model with the new return forecast from Chapter 4 (CIP + NRF)
3) The new injection planning model from Chapter 5 with the new return forecast from Chapter 4 (NIP + NRF)

We again simulate the sales for 2021 based on the sales forecast and the normally distributed forecast errors. For all three injections plans, we evaluate the overstocking- and understocking costs. Each simulation run has expected sales, returns and stock. Because each injection plan is put into the simulation, the total costs (holding costs +changeover costs) can be calculated.

The expected costs for Apollo over the time horizon of 78 weeks are:

Scenario	Expected total costs	Expected holding costs	Expected changeover costs	Expected stockout costs	Injection
1.CIP+CRF	X	X	X	X	X
2.CIP+NRF	X	X	X	X	X
3.NIP+NRF	**X**	**X**	**X**	**X**	**X**

Table 13: Expected costs Apollo

For Apollo using the updated return forecast with the current safety stock measure of a DoC of 5 already results in a holding costs saving of X euros. The old return forecast can have large errors and the old model is likely to plan more injections if the returns are under-forecasted in certain weeks. The updated return forecast is more smooth and more accurate as well. Likely less injections are needed and the timing of the injections is more accurate, resulting in less holding costs.

One point for discussion is that the simulation to determine the costs uses the new return forecast, but this is the most realistic return simulation that we can use. Even though the return forecast errors are normally distributed, simulating with the standard deviation of the error can result in invalid return forecasts. This is the case as more returns are forecasted than the amount that is sold, and forecasting too few returns (when the simulated errors together sum up to a negative amount) is also not realistic.

Planning the injections based on a DoC of 5 (the current method) results in very few problems with empty bottle availability. The implied service level by Grolsch is X. Therefore the holding costs are currently high, and with a slightly lower service level (X) the holding costs can be reduced while the stockout costs are not likely to increase much.

With the new purchasing policy based on a cycle service level of X the amount of injection is lower, the holding costs can be reduced with 8% while the expected stockout costs only increase with X euros. Total savings for Apollo are expected to be 7%.

The expected costs for BNR over the time horizon of 78 weeks are:

Scenario	Expected total costs	Expected holding costs	Expected changeover costs	Expected stockout costs	Injection
1.CIP+CRF	X	X	X	X	X
2.CIP+NRF	X	X	X	X	X
3.NIP+NRF	**X**	**X**	**X**	**X**	**X**

Table 14: Expected costs BNR

For BNR the holding costs are also reduced as less injections are planned, and the timing will be more accurate with the updated improved forecast. However, we see that the stockout costs increase with the injection plan based on the old model with the new return forecast. Because the standard deviation of the weekly sales forecast is almost X%, injecting less is risky. In some periods the safety stock measure DoC needs to be higher than 5. We conclude this based on the fact that the proposed method with a service level of X injects even less, while the stockout costs are lower. The timing of these injections is better. We notice that both the holding costs and the stockout costs decrease in Scenario 3 compared to Scenario 2. In Scenario 2 more injection is done compared to Scenario 3, but the timing was slightly worse. If more injection is done, but too late, stockouts first occur and then too much is injected after the stockouts, resulting in higher holding costs. The current safety stock, 5 Days of Cover, is not enough for weeks where the sales can deviate much. Overall we see that the new proposed purchasing policy saves 5% of total costs. This is mainly a saving in holding costs, as the expected understocking costs increased.

6.4. Conclusion

In this chapter we first validated the assumption that bottles are filled one week before the demand. We conclude that Apollo has a reasonable fit and for BNR the assumption can also be used with a slight alteration. BNR is only filled once every two weeks and at the start of the year slightly more BNR is filled than in the end of the year.

Then we simulated the sales and returns for week 33 of 2020 till week 52 of 2021 and calculated the expected costs for three scenarios:

1) The current injection planning model with the current return forecast from Section 2.4.4. (CIP+CRF)
2) The current injection planning model with the new return forecast from Chapter 4 (CIP+NRF)
3) The new injection planning model from Chapter 5 with the new return forecast from Chapter 4 (NIP+NRF)

One conclusion of this chapter is that the total amount of planned injection is lower with the improved return forecast. The current return forecasting method of Grolsch is very sensitive to sales differences from week to week. In a week where only few bottles are expected to return, it is likely that more injection is planned, which may or may not be needed later in the time horizon. With more accurate return forecast the injections are expected to be more accurate as well. However, a smaller amount of injection may result in more changeover- and stockout costs because of the uncertainty in the weekly sales. The total costs savings for Apollo are X euros and the costs savings for BNR are X euros. These savings are a reduction in holding costs as stockout costs slightly increased for both bottle types. The total costs savings of using the new injection planning model with the new return forecasting model compared to the current injection planning model with the current return forecasting model are 5% for Apollo and BNR combined. In Chapter 7 we describe the plan to implement the new parameter calculation, the new return forecasting model and the new injection planning model.

7. Implementation

In Chapters 4 to 6 we updated the input parameters, improved the return forecasting and the injection planning. Throughout this research we have been focusing on making the proposed solutions to the research problem interpretable and easy to implement for Grolsch. We chose to implement most problem solutions in Microsoft Excel, as Grolsch' existing models are also made in Excel and it's a widely used program in the whole organization. In this chapter we describe how Grolsch can implement our solutions.

The implementation is split into three parts:

1) Updating of parameters Trade Loss, Internal Loss, Time in Trade and Trade Population
2) Implementing the new return forecasting model
3) Implementing the new injection planning model

In this chapter each part of the implementation is described.

7.1. Updating of parameters TL, IL, TiT and TP

The wish of Grolsch was that the input parameters are not only recalculated, but can also be updated each year. The parameters should be annually updated to include new market data and trends. For this purpose we built an Excel tool to make this updating easy when new data becomes available. The dashboard of the tool is shown in Figure 56:

Confidential

Figure 56: Dashboard Excel tool

There are some manual inputs needed for this tool to work. These inputs are:

- Sorting input and output needs to be input manually into the parameter calculation model. The Warehouse department sends an Excel-file containing sorting input and output and production losses every week to the Supply Chain Planning department. The Excel-file that is send at the start of each year includes the sorting input and output data of the full previous year. This needed tabs can be manually copied to the parameter updating tool. Once the sorting input and output are manually input, the parameter calculation tool will automatically update the sorting percentage table (% per crate of each bottle type).
- Return data on crate level need to be manually input as well. These data can be easily exported from the ERP system SAP and it is outlined in the Excel File how to exactly do this.

Within the Excel tool, we wrote an explanation on how to use the model and how the model calculates. The tool is made and improved based on feedback from Grolsch and can be used right away.

7.2. Implementing the new return forecasting model

In the current container-planning model of Grolsch the return forecasting model and injection planning model are built in. Grolsch wants to keep using the current container-planning model, because this model is also used among other breweries in the Asahi group that Grolsch is part of. We had to think how we can include our proposed solutions in the current container-planning model. It is good to say that we already implemented the new return forecasting model in the container-planning model of Grolsch during the research. Two things needed to be altered:

1) The current container-planning model needed one extra column (containing the TiT-distribution that followed from the parameter calculation tool)
2) The return forecast column needed to be altered so the forecasted returns of a week are based on the TiT-distribution and the sales instead of on Weeks in Trade and TP.

The mother organization needs to be convinced that the updated method should be used in practice. To do this we need to present the improved solution and show the return forecast becomes more accurate and holding costs can be saved. A meeting will be planned in the near future.

7.3. Implementing the new injection planning model

To improve the current injection planning model, we need to implement our new injection planning model in the container-planning model of Grolsch just as with the improved return forecast. To be able to do this we need to know the required safety stock. The required safety stock is calculated in Chapter 5. This is done as follows: the cumulative net demand is currently simulated in Excel and this Excel-File can be used by Grolsch with a clear explanation on how to do this. The Supply Chain Planning department of Grolsch can run this simulation every year (or when significant changes to the sales forecast are made) to calculate new levels of safety stock. The safety stock follows from the standard formula from Section 3.4 and the calculation is incorporated in this Excel-File. This is however a separate file from the current container-planning model, so the safety stock output has to be copied to the current model for budget analysis. Based on the improved return forecast and the improved required safety stock of empty bottles, the improved injection plan is output by the current container-planning model.

8. A look towards the future

In the previous chapters we improved the return forecast and injection planning of Apollo and BNR for Grolsch. However, reverse logistics remains a relatively unexplored research area. Literature on reverse is logistics is still scarce and the data on returns that Grolsch currently has available could be extended. Technology is ever evolving and environmental issues are becoming more and more In the coming years which raises more interest for recycling of materials. In this chapter we give our ideas on how these trends could lead to further improvements of the return process for Grolsch.

RFID-chips

The first thing that comes to mind is Radio-Frequency Identification (RFID) tracking. This technology is becoming cheaper with the years, and its benefits are already seen in the pilot at Heineken in Section 3.3. RFID chips can be put into the crates and with a sensor at the production line the time between successive filling batches can be accurately measured. If Grolsch registers which production batches go to which clients, the time the crates stayed in the market is known while no manual reading of labels is needed for this purpose.

Reading labels at the sorting line

But not all bottles are sold in crates. Besides, when the time in the market of a crate is known the time in the market of the bottles is still not precisely known. So instead of RFID tracking, there is another possibility of improving the data availability of returned containers namely: reading the labels of the bottles at the sorting/ filling line. Currently, at the sorting line a picture is made from above a crate to identify how many bottles of each type are in the crate. This could first be extended to making pictures of all individual bottles and reading them with machine learning with a computer. Our feeling is that the return process of bottles can be improved a lot more when real circulation times are known. The time in the warehouse needs to be tracked as well.

New beers and deposit on cans

One trend in the beer market is that customers want to have more and more choice. Grolsch currently introduces new beers every year. New beers sold without crates can contribute to more Trade Loss in comparison to bottles sold in crates. In the near future, deposits on cans becomes a reality. It is not yet known how consumers will react to this. Consumers might choose more for bottles instead of cans. This might be the case when cans get more expensive and consumers find it harder to save the empty cans than empty bottles in a crate.

Transport optimization

Currently Grolsch tries to combine full bottle deliveries to clients with taking empty bottles back. However, sometimes more bottles are needed in the short term so trucks are sent into the market to collect more bottles. When more accurately can be forecasted where bottles are in the supply-chain, better decisions can be made regarding pickup of extra bottles in times of need. For example when it is known at which clients' depots bottles of a certain type are present, trucks can be send there to pick these bottles up. In this way there may be a smaller bottle population needed overall, which can save purchasing costs.

9. Conclusions and recommendations

In this chapter we provide the answer to our main research question and give our recommendations to Grolsch. The main research question was:

"How can the injection planning for the two main bottle types be improved, by improving the long-term container-planning model's input parameter calculation, return forecast and purchasing policy? "

We have split the question into sub-research questions based on three parts: Current situation, Literature review and Solution design. The questions of these three parts are answered in this chapter.

Current situation

In this part we described the current return process of containers and talked about how the current container-planning model works. We described the inputs, the outputs and how these different inputs and outputs of the model are currently calculated. The most important questions this part is:

1) *How did the model perform over the last years in terms of return forecast accuracy?*

The model is considered inaccurate in terms of total amounts of returns forecasted (mainly based on TL) as well as the timing of the returns (mainly based on WiT and TP). The MAPE per week is high for both Apollo (25%) and BNR (40%), with the return forecast based on realized sales.

The conclusion from the current situation is that the accuracy of the return forecast can significantly be improved and that the calculation of the parameters TL, WiT, TP and DoC can also be improved. Especially the timing of the returns, which has a big impact on when to do injections, is inaccurate.

Literature review

2) *How are container returns related to returns in different industries?*

Grolsch does not register the production code of every returned bottle and does therefore not know how long each bottle stayed in the market. In other industries each individual item return is usually registered. Besides, return forecasts are made for product categories instead of individual items. For container returns it is usually the other way around: only aggregate return data are available, but forecasts are made for each individual container type. Items in other industries are generally returned relatively quickly after purchase, while it may take several months for containers to return in the brewing industry.

3) *Which methods and models are proposed for forecasting container returns?*

There are different methods and models proposed in literature: Univariate methods (ARIMA, exponential smoothing and Holt's method) and multivariate methods (DLM, regression and machine learning). Univariate methods seem inappropriate for this research as they fail to take sales into account to forecast the returns and sales can differ significantly from week to week. For long-term return forecasting the better approach is to perform time series analysis, which is the analysis of trends, seasonality (cycles) and correlation between sales and returns. The DLM is in our opinion the best way to model the return forecast for Grolsch, because it only needs aggregate data and can be solved using time series analysis.

Other methods such as machine learning (for example neural networks) are typically used for return forecasting with more explanatory variables (features). These methods are more suitable for return forecasting in other industries, but less suitable for container return forecasting. We only have aggregate return data available (total returns per period) and do not know the exact times that bottles stay in the market. We have more detailed sales data (for example on SKU level), but we do not know what SKU-number a returned bottle belongs to. It is therefore hard to train a model and to find accurate relationships between sales and returns.

4) *Which methods and models are proposed for lot-sizing in reverse logistics?*

We have seen multiple models for purchasing but most were models for remanufacturing, which differs from Grolsch' situation because the "raw material" stock is filled with purchasing at Grolsch instead of the serviceable inventory in typical remanufacturing. Besides, there is no disposal decision or workforce scheduling in our research problem. However, we found one particular model on container purchasing decisions that can indeed be used. It uses cumulative net demand (demand-returns) to determine how many new containers are needed up to each period in the time horizon.

Solutions

5) *How can the parameter calculation of TL, IL, TiT and TP be improved?*

We conclude that the parameter calculation for TL and TP can be improved and that the usage of TiT instead of WiT improves the accuracy of the timing of the forecasted returns. In the TL-calculation the unsorted returns need to be taken into account. For BNR specifically, the upwards sales trend needs to be included. The updated values for TL are X% and X% for Apollo and BNR respectively. The current IL calculation seems appropriate resulting in updated values of X and X% for Apollo and BNR respectively. The expected TP can be calculated using the TiT-distribution as it is estimated how many percent of the sales are still expected in the market each week.

6) *How can the return forecast be improved?*

The finite DLM is in our opinion the best way to model the return forecast for Grolsch, because it only needs aggregate data and can be solved using time series analysis. TiT-distribution can be estimated with the lognormal distribution in Excel. Seasonality needed to be included to prevent under-forecasting of returns in peak periods and over-forecasting in the off-season. The average TiT for Apollo is X weeks and for BNR X weeks.

7) *How can the purchasing policy be improved?*

We used the model of Kelle and Silver (1989b) to improve the purchasing policy. The standard formula for safety stock can be used because of the normally distributed cumulative net demand. This results in different amounts of safety stock than the current usage of DoC, which indicates a service level of X.

8) *How accurate is the improved return forecasting model?*

The improvements in terms of MAPE for 2020, with the return forecasting model fitted on data of 2018-2019, are:

	Return forecast MAPE Apollo	
	2020 (based on realized sales)	2020 (based on forecasted sales)
Current return forecast	24.0%	26.1%
Improved return forecast	12.6%	14.3%

Table 15: Expected costs Apollo

	Return forecast MAPE BNR 2020 (based on realized sales)	2020 (based on forecasted sales)
Current return forecast	35.7%	38.7%
Improved return forecast	14.8%	16.1%

Table 16: Expected costs BNR

We conclude that the improvement to the return forecast is significant and will likely lead to a more accurate injection planning.

9) *What are the expected savings per year when using the improved container-planning model over the current model?*

We calculated the expected holding-, changeover- and stockout costs based on three scenarios (three different injections plans). These three scenarios are:

1) The current injection planning model with the current return forecast
2) The current injection planning model with the new return forecast
3) The new injection planning model with the new return forecast

The results for Apollo are as follows:

Scenario	Expected total costs	Expected holding costs	Expected changeover costs	Expected stockout costs	Injections
1.CIP+CRF	X	X	X	X	X
2.CIP+NRF	X	X	X	X	X
3.NIP+NRF	**X**	**X**	**X**	**X**	**X**

Table 17: Expected costs Apollo

The results for BNR are:

Scenario	Expected total costs	Expected holding costs	Expected changeover costs	Expected stockout costs	Injections
1.CIP+CRF	X	X	X	X	X
2.CIP+NRF	X	X	X	X	X
3.NIP+NRF	**X**	**X**	**X**	**X**	**X**

Table 18: Expected costs BNR

We conclude that total costs (Apollo + BNR) can be reduced with 5%.

Recommendations

We recommend Grolsch to use updated parameters and return forecast in their current long-term container-planning model. The accuracy is significantly improved in comparison with the old method based on WiT. The parameters should be updated annually with the Excel tool.

The next recommendation Is to keep track of the return data per customer. This is easier than checking the production code of all bottles and can still give good insights in how containers are returning. These data are mainly interesting to analyze returns for example based on the different sold SKUs, delivery frequency and geographical location. With more explanatory variables the proposed return forecasting model can still be used. For example to find an Time in Trade distribution of a specific region. Other methods such as machine learning become suitable as well.

In 2019 a small sample of Apollo bottles is checked for their production code. If sampling is done again, there are some considerations. The main concern is obtaining a representative sample in terms of size, customers from which the containers are selected and time in the year that the sampling is done. Another consideration is for which container to take a sample. A bigger impact for Grolsch is knowing how the distribution looks for the main bottle type Apollo. However, it might be hard to get a representative sample, while for Kornuit the sample needed to get a good estimate may be way lower.

With the new injection planning model 5% of total costs can be saved, so although stockouts may occur more frequently because of lower safety stock of empty bottles, holding costs can be reduced significantly.

10. Discussion and further research

This chapter is about the limitations of the research and the possibilities to extend it in the future.

Trade Loss is assumed to be constant during the year. In reality, the TL might vary throughout the year. Trade Loss can be dependent on the SKUs that are sold. Bottles that are sold without a crate might have a lower possibility to return than bottles sold in a crate. In different times of the year different SKUs are sold. For BNR the Trade Loss is also dependent on how well the workers at the supermarkets put BNR bottles in Grolsch crates as BNR bottles also go to other brewing companies.

The main point for discussion of the return forecast is the amount that is still allowed to come back after 52 weeks, and how this amount is scaled. This needs to be further investigated, but it was the best guess of experts from Grolsch and our own insights. Over time Grolsch might get a better feeling about how much comes back in each week. The percentage that comes back after 52 weeks may for example be estimated with greater precision. Another example could be that not more than 10% is expected back in the first week. All this knowledge can be put into the model to hopefully optimize the return forecast year by year.

We now used two Time in Trade distributions to capture the seasonality and we have seen that this is reasonable. In reality the seasonality over the weeks may be more smooth and in the future it can be researched if a smart algorithm can be made to determine how the seasonality over the weeks can be found, even without data of realized times in the market of containers.

Over-stocking and understocking costs in the simulation model are based on Apollo and BNR on aggregate level instead of on SKU level. The changeover and stockout costs may differ from SKU to SKU. In reality however, the SKUs that have a big impact on are generally prioritized with production. This means that the stockout costs in reality might be slightly lower. In the future the injection planning model could be extended with a production planning to further optimize the process of procurement of new containers and then Grolsch exactly knows when they are needed instead of procurement based on assumptions.

We have briefly described new technologies in Chapter 8. But if there will not be an technological idea to help determine the TiT of bottles like RFID, it will be very hard to determine TiT based on samples. A sample of one period does not say anything about another period as TiT is expected to be seasonal over the year. Also TiT is very much based on where the bottles came from. Supermarkets will have lower TiT than Liquor stores. We do not advise on doing more sampling for these reasons. Grolsch has to keep data about where each crate came from, so a better and more accurate analysis can be done to further reduce the uncertainty in the return process.

When the return quantities per time period are known per customer instead of only on aggregate level, the return forecast can be improved on a deeper level. Analysis on how different customers are clustered together in deliveries and returns. How much do you get back based on the frequency of visiting customers and the kind of products that you sell them?

References

Bierkens, J., Blok, H., Heydenreich, M. O., Núñez Queija, R., Meurs, van, P. J. P., Spieksma, F. M., & Tuitman, J. (2013). Estimates on returnable packaging material. In M. O. Heydenreich, S. C. Hille, V. Rottschäfer, F. Spieksma, & E. Verbitskiy (Eds.), Proceedings of the 90th European Study Group Mathematics in Industry (SWI 2013, Leiden, The Netherlands, January 28-February 1, 2013) (pp. 31-49). Lorentz Center.

Box, E. P. G. & Jenkins, M. (2016). Time Series Analysis: Forecasting and Control, 5th Edition, Published by John Wiley and Sons Inc., Hoboken, New Jersey, pp. 712. ISBN: 978-1-118-67502-1. Journal of Time Series Analysis. 37. n/a-n/a. 10.1111/jtsa.12194. Ravines, Schmidtz, Migon; 2006

De Brito, M.P. & Dekker R. (2002). A framework for Reverse Logistics. in R. Dekker, K. Inderfurth, L. van Wassenhove and M. Fleischmann (eds.), Quantitative approaches to reverse logistics (forthcoming)

De Brito, M.P. & Van der Laan, E. (2009). Inventory control with product returns: The impact of imperfect information, European Journal of Operational Research, 194 (1), pp. 85- 101.

Carrasco-Callego, R. & Ponce-Cueto, E. (2009). Forecasting the returns in reusable containers' closed-loop supply chains. A case in the LPG industry. 3rd International Conference on Industrial Engineering and Industrial Management XIII Congreso de Ingeniería de Organización Barcelona-Terrassa, September 2nd-4th 2009

Carrasco-Callego, R., Ponce-Cueto, E. & Dekker, R. (2012). Closed-loop supply chains of reusable articles: A typology grounded on case studies. International Journal of Production Research - INT J PROD RES. 50. 1-15. 10.1080/00207543.2011.649861.

Chevallier (2006). Law of the sum of Bernoulli random variables. Université de Haute Alsace, 4, rue des fréres Lumiére 68093 Mulhouse

Clottey, T. (2016). Development and evaluation of a rolling horizon purchasing policy for cores, International Journal of Production Research, 54:9, 2780-2790 .

Clottey, T. & Benton, W.C. (2014). Determining core acquisition quantities when products have long return lags, IIE Transactions, 46:9, 880-893

Cui, H, & Rajagopalan, S. & Ward, A R., (2020). Predicting product return volume using machine learning methods, European Journal of Operational Research, Elsevier, vol. 281(3), pages 612-627.

Van Dalen J., van Nunen J.A. & Wilens C.M. (2005). The chip in crate: the Heineken case. In: Flapper S.D.P., van Nunen J.A., Van Wassenhove L.N. (eds) Managing Closed-Loop Supply Chains. Springer, Berlin, Heidelberg.

Ene & Özturk (2017). Grey modelling based forecasting system for return flow of end-of-life vehicles, Technological Forecasting and Social Change, Volume 115, 2017, Pages 155-166, ISSN 0040-1625

Fleischmann, M. & Minner, S. (2004). Inventory Management in Closed Loop Supply Chains. 10.1007/978-3-540-24815-6_6.

Geda, M. W. & Kwong, C.K. (2019). Forecasting of Used Product Returns for Remanufacturing. 889-893. 10.1109/IEEM.2018.8607362.

Goh, T. N. and Varaprasad, N. (1986). A statistical methodology for the analysis of the Life-Cycle of Reusable Containers, IIE Transactions, 18, pp. 42-47.

Gómez, J. & Rautenstrauch, C. & Nürnberger, A. & Kruse, R. (2002). Neuro-fuzzy approach to forecast returns of scrapped products to recycling and remanufacturing. Knowl.-Based Syst.. 15. 119-128. 10.1016/S0950-7051(01)00128-9.

Helmer R.M., Johansson J.K. (1977). An Exposition of the Box-Jenkins Transfer Function Analysis with an Application to the Advertising-Sales Relationship. Journal of Marketing Research. 1977;14(2):227-239.

Holmes, E. E., M. D. Scheuerell, and E. J. Ward. (2020). Applied time series analysis for fisheries and environmental data. NOAA Fisheries, Northwest Fisheries Science Center, 2725 Montlake Blvd E., Seattle, WA 98112

Hyndman, R.J., & Athanasopoulos, G. (2018). Forecasting: principles and practice, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2.

Kelle, P. & Silver, E.A. (1989a), Forecasting the Returns of Reusable Containers, Journal of Operations Management, Vol.8, No.1, pp. 17-35.

Kelle, P. & Silver, E.A. (1989b). Purchasing Policy of New Containers Considering the Random Returns of Previously Issued Containers. IIE Transactions, 21(4), pp.349-354.

Kiesmüller, G. & Van der Laan, E. A. (2001). An Inventory Model with Dependent Product Demands and Returns (27 2001 3,).

Krapp, M., Nebel, J. and Sahamie, R. (2013). Forecasting product returns in closed-loop supply chains, International Journal of Physical Distribution & Logistics Management, Vol. 43 No. 8, pp. 614-637.

Van der Laan, E. & De Brito, M.P.(2003). Managing Product Returns: The Role of Forecasting. In: Dekker, R., Fleischmann, M., Inderfurth, K. and Van Wassenhove, L.N. (eds.), Reverse Logistics. Quantitative Models for Closed-Loop Supply Chains. Springer.

Ljung, G. and Box, G.E.P. (1978). On a measure of a lack of fit in time series models, Biometrika, Vol. 65 No. 2, pp. 297-303.

Ma, J., and Kim, H. M. (2016). Predictive Model Selection for Forecasting Product Returns, ASME. *J. Mech. Des*. May 2016; 138(5): 054501.

Mahmoudi, M & Parviziomran, I. (2020). Reusable packaging in supply chains: A review of environmental and economic impacts, logistics system designs, and operations management, International Journal of Production Economics, Volume 228, 2020, 107730, ISSN 0925-5273.

Moroke, N. (2015). BOX-JENKINS TRANSFER FUNCTION FRAMEWORK APPLIED TO SAVING-INVESTMENT NEXUS IN THE SOUTH AFRICAN CONTEXT. *Journal of Governance and Regulation, 4*, 63-77.

Pierce, D. A. (1972). Least Squares Estimation in Dynamic-Disturbance Time Series Models, Biometrika, 59, 73-8.

Ravines, R. & Schmidt, M. A. & Migon, H. (2006). Revisiting distributed lag models through a Bayesian perspective. Appl. Stochastic Models Bus. Ind., 22: 193-210.

Senthil, S. (2014). REVERSE LOGISTICS: A REVIEW OF LITERATURE. International Journal of Research in Engineering and Technology. 3. 140-144. 10.15623/ijret.2014.0323031.

Silver, E.A., Pyke, D.F., & Thomas, D.J. (2016). Inventory and Production Management in Supply Chains (4th ed.). CRC Press. https://doi.org/10.1201/9781315374406

Toktay, L. B., van der Laan, E. A., & De Brito, M. P. (2004). Managing product returns: Informational issues and forecasting methods. In R. Dekker, M. Fleischmann, K. Inderfurth, & L. N. Van Wassenhove (Eds.), Reverse logistics: Quantitative methods for closed- loop supply chains. Berlin: Springer.

Widi, S. (2009). Inventory Management of Returnable Bottles of Brewery. Thesis Erasmus University Rotterdam

Xiaofeng, X. & Tijun, F. (2009). Forecast for the Amount of Returned Products Based on Wave Function. In Proceedings of the 2009 International Conference on Information Management, Innovation Management and Industrial Engineering - Volume 02 (ICIII '09). IEEE Computer Society, USA, 324–327.12.