**Delft University of Technology**
**Faculty Electrical Engineering, Mathematics and Computer Science**
**Delft Institute of Applied Mathematics**

# Optimal "here-and-now" decisions in multi-stage robust optimisation

**(Nederlandse titel: Optimale "hier-en-nu" beslissingen in multi-stadia robuuste optimalisatie)**

A thesis submitted to the
Delft Institute of Applied Mathematics
in partial fulfillment of the requirements

for the degree

**BACHELOR OF SCIENCE**
**in**
**APPLIED MATHEMATICS**

**by**

**Janne Heslenfeld**

**Delft, the Netherlands**
**July 2021**

**Thesis committee:**
Dr. K.S. Postek (supervisor)
Dr. W.G.M. Groenevelt

## Abstract

Robust optimisation is shown to be extremely important in a wide range of applications including real life. Many research projects are dedicated to this relatively young and active research field and show the significant value of robust optimisation. Since many researchers have developed complex methods dedicated to robust optimisation, the aim of this project is to find out whether or not these complex methods were helpful or that the same or even better results could have been obtained with rather simpler methods.
This is done by analysing one famous paper about multi-stage robust optimisation and comparing the "here and now" decisions of this approach with the "here and now" decisions of a much simpler method.

# Contents

# 1 Introduction

Mathematical optimisation is of great importance in finance, engineering, healthcare, scheduling and logistics. All these different fields of research encounter complex optimisation problems. An example is the task to assign crews to different airline flight segments to minimize the total costs, but making sure that a crew member begins and ends in the same city.

There are optimisation problems that only uses data that is defined and/or estimated before the problem is solved, we refer to this as certain data. However, most of the time, real life optimisation problems consist of a lot of uncertain data. The weather for example, is a rather uncertain parameter, but a schedule maker of the airlines needs to take the weather into account as well. What if the weather delays the one airplane such that the crew member cannot make the transfer to the next flight?
Robust optimisation is the field in optimisation where this type of uncertainty occurs. In robust optimisation it is assumed that the uncertainty resides in a so called uncertainty set. This set is defined as a set of 'reasonable' outcomes for the uncertain data. This set is defined before the problem is solved. Solving robust optimisation problems is most of the time complex.

The aim of this report is to analyse a robust optimisation method, used in another famous paper, and see if this rather complex method is superfluous since the same or even better "here and now" decisions could be obtained with a much simpler method. Is the sophistication worth the effort? During my thesis I only implement the "here and now" decisions since these are the ones that need to be implemented right away. The other decisions can be defined/altered at a later time stage.
The complex method I consider is a method by A. Ben-Tal in 2004, called "Adjustable robust solutions of uncertain linear programs", [1].

The remainder of the report is organised as follows: in Chapter 2, a brief introduction is given about optimisation, robust and multi-stage optimisation. In this Chapter the main concepts and assumptions are mentioned and illustrated. In Chapter 3 the choice of the paper is made and substantiated and a concise explanation of the affinely adjustable robust counterpart is given, in accordance with the method of the chosen paper. In Chapter 4, the model given in the chosen paper is presented and explained. Chapter 5 provides a comparative study of the "here and now" decisions of the 'easy' and complex methods and statistical tests are done to explain relations between variables of the nominal model, which is used for the 'easy' method.
In Chapter 6, I produce a regression on the "here and now" decisions of the nominal model and the demand trajectory. This is done to make a prediction of the "here and now" decision with a given demand trajectory without having to solve an optimisation problem with a method that takes into account all possible outcomes. I close my report with a clear conclusion of my research and I discuss and reflect on my thesis with some recommendations for further research.

# 2 Optimisation

## 2.1 Introduction

Mathematical optimisation or mathematical programming is the way of solving a problem such that the outcome or solution is optimal. The complexity of optimisation lays in the uncertainty whether an optimal solution exists. Also an important question that arises when one uses optimisation methods is, how could one solve these problems in the most efficient way.

The latter is often complicated. The goal of this report is to analyse if the approaches people use to solve complex optimisation problems are the most efficient way of doing so. Aren't there easier methods to obtain the same or maybe even better results?

A basic optimisation problem is defined as follows. First the goal of the problem, for example "maximize the profit" or "minimize the cost" has to be defined. Secondly, some restrictions have to be taken into account when solving the problem. These restrictions are called constraints. These constraints have to be true no matter what the solution will be. Lastly, the model also needs to provide where the variables lay in the space. For example in the $\mathbb{R}, \mathbb{N}, \mathbb{Z}$ etc.

## 2.2 Robust optimisation

Real-life optimisation often contains uncertain data. This data can be random or uncertain due to made errors in real life, for example, estimation errors or forecasts that have to be made. Robust and stochastic optimisation are the fields in optimisation that deal with optimisation that contains such uncertainty. Robust optimisation (RO) distinguishes itself from stochastic optimisation by means of assumptions. Stochastic optimisation assumes that the real probability distribution of the uncertain data is known or estimated. RO, on the other hand, assumes that the uncertain data lays in a so called uncertainty set, which consist of reasonable outcomes for the uncertain data.

An uncertain linear optimisation problem has a "general" formulation as follows [2]:

$$\min_{x}\{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{d}\}_{(\boldsymbol{c},\boldsymbol{A},\boldsymbol{d})\in\mathcal{U}}, \tag{1}$$

where $\boldsymbol{c} \in \mathbb{R}^n$, $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and $\boldsymbol{d} \in \mathbb{R}^m$ denote the uncertain coefficients, and $\mathcal{U}$ denotes the specified uncertainty set.

Decisions that should get values as a result of solving a problem **before** the actual data is known, are called "here and now" decisions. The decisions have to be made here and now, and cannot wait until the data reveals itself.

The assumptions of RO mentioned as follows are stated in [2]:

**Assumptions by (Gorissen, 2015) about robust optimisation**

1. All decision variables $\mathbf{x} \in \mathbb{R}^n$ represent "here and now" decisions;
2. The decision maker is fully responsible for consequences of the decisions to be made when, and only when, the actual data is within the prespecified uncertainty set $\mathcal{U}$;

3. The constraints of the uncertain problem in question are "hard", i.e. the decision maker cannot tolerate violations of constraints when the data is in prespecified uncertainty set $\mathcal{U}$.

Furthermore, in RO it is assumed that the uncertainty set is a non-empty compact convex set.

For the remaining part of this report the first assumption can be relaxed. This means that the assumption that all variables are "here and now" decisions will be flexible. This will be explained in Section 3.3.

The robust counterpart is defined as follows [2]:

$$\min_{x}\left\{\sup_{(\boldsymbol{c},\boldsymbol{A},\boldsymbol{d})\in\mathcal{U}}(\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}) : \boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}-\boldsymbol{d}\leq 0\ \ \forall(\boldsymbol{c},\boldsymbol{A},\boldsymbol{d})\in\mathcal{U}\right\} \tag{2}$$

This robust counterpart is often necessary to solve a problem with a certain solution method. A lot more can be said about uncertain sets and properties of RO [2], but for the sake of this report, it will not be discussed here.

## 2.3 Multi-stage optimisation

In two-stage optimisation, you have to make decisions in two different stages. So first some decisions have to be made. Then, one observes what happened and makes the second stage decisions about what to do. An example is the inventory problem. The difficulty of this problem lays in the uncertainty of the demand and how to optimize the problem without knowing the exact demand.

**Example 2.1. Inventory model** A company has to make sure that every warehouse has enough inventory according to their demand. First, the factory of the company distributes the inventory across warehouses. This is called the first stage. Then the company observes the demand of the shops and the second stage decisions, last minute transport between the shops to meet the demand, are made.

As one can already guess, multi-stage optimisation, is the optimisation problem where decisions have to be made in multiple stages. I reuse the inventory model example.

**Example 2.2. Inventory model 2** Imagine that a company has two factories that need to produce a particular product. Every month the company counts the demand of the previous month and makes a decision how many products each factory has to make for the coming month. The factory thus has to make decisions in multiple time stages. The decision that has to be made in the first time stage, is called a "here and now" decision. These decisions have to be implemented right away and can not depend on observed demands. The decisions that the company makes in other time stages are not "here and now" decisions. These decisions can wait. The company observes what the demand was in the previous time stages and based on these observations, the company makes a decision.

# 3 The considered research paper

## 3.1 Introduction

For this report a research paper about robust multi-stage optimisation will be used to analyse the considered method to see if this method is worth the effort. Many papers are already dedicated to the robust optimisation field. To make a choice I will narrow my search area to a limited number of papers and shortly analyse these to make a choice.

## 3.2 Choice of research paper

I narrowed my search area to two different papers. Namely, "Adjustable robust solutions of uncertain linear programs" by Ben-Tal et al. [1] and "$K$-adaptability in two-stage mixed-integer robust optimisation" by Wieseman et al. [3].

Both papers focus on linear programs where some parameters lay in some pre-described uncertainty set. Ben-Tal et al. focuse on the problems where part of the variables must be determined before the realization of the uncertain parameters, those are called "non-adjustable variables", i.e. "here and now" decisions, while the other part are variables that can be chosen after some realization, those are called "adjustable variables". This method of solving optimisation problems by defining adjustable variables is called adjustable affine RO. The term affine indicates that I model the adjustable decisions as affine functions of the uncertainty set. This is done since most of the time the adjustable RC is computationally intractable, i.e. NP-hard. The adjustable affine RC (AARC) on the other hand is shown to be most of the time tractable. The paper uses this AARC approach on a multi-stage inventory problem, which I will analyse. [1]

The second paper uses a $K$-adaptability approach. "This approach selects $K$ candidate recourse policies before observing the realization of the uncertain parameters and that implements the best of these policies after the realization is known." [3] The paper uses the $K$-adaptability approach on multiple different problems, such as vehicle routing, project management and shortest paths.

The main difference between the two papers is that $K$-adaptability considers a small number of 'back up' plans to address uncertainty. The AARC approach constructs an infinite set of back up plans because they consist of affine functions of the infinite uncertainty set. This means that $K$-adaptability is in some sense more limited. However, it is more conform human decision making since it considers a limited number of options.

After analysing both papers shortly, I have decided to choose the first paper about AARC. This choice is underpinned by, in my opinion, a more interesting application problem.

### 3.3 The method: Adjustable affine robust optimisation

As mentioned above the chosen paper uses an adjustable affine robust counterpart (AARC). This approach uses two kind of variables, namely adjustable and non-adjustable variables. What differs from the RO approach is that with RO one has to make a decision before the actual realization of the uncertain data is known. So RO treats all the variables as non-adjustable variables, i.e. "here and now" decisions.
However, in many real life cases, only part of the decision variables are "here and now" decisions.
Other types of variables are "wait and see" variables and analysis variables that are used to model the problem as an LP. Since analysis variables do not mean actual decisions, they can adjust themselves to different data.
To illustrate what is meant with a "wait and see" decision, an example is given.

**Example 3.1.** The amount a factory will produce next month is not a "here and now" decision but a "wait and see" decision that will be made based on the amount sold in the previous and current months. Those "wait and see" decisions can be made when part of the data is known, in this example the amount sold in the previous months. [2]

Both the analysis and "wait and see" decisions can adjust themselves to a corresponding part of the realized data, and thus these variables are called adjustable variables.

As seen in Section 2.2, an uncertain LP problem is defined as (1). When treating adjustable and non-adjustable variables, we rewrite the equivalent (2) to the form [1]:

$$\min_{\boldsymbol{x},\boldsymbol{y}(\cdot)} \left\{ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} : \boldsymbol{A}(\boldsymbol{\zeta})\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y}(\boldsymbol{\zeta}) \leq \boldsymbol{d} \;\; \forall \boldsymbol{\zeta} \in \mathcal{U} \right\}, \tag{3}$$

with $\boldsymbol{x} \in \mathbb{R}^n$ is the first-stage "here and now" decision that is made before $\zeta \in \mathbb{R}^L$ is realized, $\boldsymbol{y} \in \mathbb{R}^k$ is the second stage "wait and see" decision (this one can be adjusted) and $\boldsymbol{B} \in \mathbb{R}^{m \times k}$ is a fixed recourse matrix.
As mentioned in Section 3.2, this optimisation over general $y(\cdot)$ is not conservative enough and therefore I restrict myself to affine adjustable robust optimisation. I approximate $\boldsymbol{y(\zeta)}$ by affine decision rules:

$$\boldsymbol{y}(\boldsymbol{\zeta}) := \boldsymbol{y}^{\boldsymbol{0}} + \boldsymbol{Q}\boldsymbol{\zeta} \tag{4}$$

This leads to a reformulation of (3) [1]:

$$\min_{\boldsymbol{x},\boldsymbol{y}^{\boldsymbol{0}},\boldsymbol{Q}} \left\{ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} : \boldsymbol{A}(\boldsymbol{\zeta})\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y}^{\boldsymbol{0}} + \boldsymbol{B}\boldsymbol{Q}\boldsymbol{\zeta} \leq \boldsymbol{d} \;\; \forall \boldsymbol{\zeta} \in \mathcal{U} \right\}, \tag{5}$$

For more details, see [2] and [1]. For simplicity on the notation, this is the case where the "wait and see" variables $\boldsymbol{y(\cdot)}$ are allowed to depend on the whole $\boldsymbol{\zeta}$. Let's consider the "wait and see" variables as the amount a factory produces and that the uncertainty lays in the demand. This ordering decision can depend on the demand over the whole time horizon that is considered, so on the whole uncertainty set. However most of the time these decisions can only depend on the demands that are known, i.e. the past demands. Further in the report, the "wait and see" variables have their own information basis, i.e. these variables depend on a certain part of the realized data.

# 4 The examined model of the research paper

## 4.1 Introduction

In this Chapter I introduce the experiment introduced in [1] and considered in this report. The paper [1] uses an inventory management problem. This problem considers a single product inventory system, which consists of a warehouse and $I$ factories. The goal of this problem is to minimize the total production cost. This must be minimal for all factories over the entire planning period. First I define the variables and parameters of the model, then I attach values to the parameters in an illustrative example. Lastly, this illustrative example is solved with the AARC approach and the optimal values are given for different values of the uncertainty level.

## 4.2 The variables and parameters

For the planning horizon $T$ the following holds at time stage $t$:

- $d_t$ is the demand for the product. All the demand must be satisfied;
- $v(t)$ is the amount of the product in the warehouse at the beginning of the period ($v(1)$ is given);
- $p_i(t)$ is the $i$-th order of the period - the amount of the product to be produced during the period by factory $i$ and used to satisfy the demand of the period (and, perhaps, to replenish the warehouse);
- $P_i(t)$ is the maximal production capacity of factory $i$;
- $c_i(t)$ is the cost of producing an unit of the product at a factory $i$.

Other parameters of the problem are:

- $\mathrm{V}_{\min}$ - the minimal allowed level of inventory at the warehouse;
- $\mathrm{V}_{\max}$ - the maximal storage capacity of the warehouse;
- $Q_i$ - the maximal cumulative production capacity of i'th factory throughout the planning horizon.

When all the data is known on beforehand, a LP is defined as follows [1]:

$$\begin{aligned}
\min_{p_i(t),F} \quad & F \\
\text{s.t.} \quad & \sum_{t=1}^{T}\sum_{i=1}^{I} c_i(t)p_i(t) \le F \\
& 0 \le p_i(t) \le P_i(t), \quad i = 1,...,I, \quad t = 1,...,T \\
& \sum_{t=1}^{T} p_i(t) \le Q(i), \quad i = 1,...,I \\
& \mathrm{V}_{\min} \le v(1) + \sum_{s=1}^{t}\sum_{i=1}^{I} p_i(s) - \sum_{s=1}^{t} d_s \le \mathrm{V}_{\max}, \quad t = 1,...,T.
\end{aligned} \tag{6}$$

Not always all the data is known on beforehand so we assume that uncertainty occurs in the model. When specifying our supply policies, which have to be done before the planning period starts ($t = 0$), one does not know exactly the future demands, all one knows is that this demand exists in a certain uncertainty set. This uncertainty depends on the positive nominal demand $d_t^*$ and positive $\theta$. Another aspect that occurs in the uncertain model is that when deciding the amount of produced products ($p_i(t)$) which have to be decided at the beginning of period $t$, one is allowed to look at the demands $d_r$, where $r \in I_t$, with $I_t$ a given subset of {1,....,t}. [1]. In this report the *information basis* $I_t$ is defined as follows: $I_t = \{1, ..., t-1\}$.
This *information basis* is for the sake of this report chosen as the most natural one, the past is known and the present and future are unknown.
When applying the AARC methodology, I use affine decision rules defined by [1]:

$$p_i(t) = \pi_{i,t}^0 + \sum_{r \in I_t} \pi_{i,t}^r d_r \tag{7}$$

with $\pi_{i,t}^r$ are the new non-adjustable variables.
With this additional information approach, I obtain an uncertain LP model of the following form [1]:

$$\begin{aligned}
\min_{\pi,F} \quad & F \\
\text{s.t.} \quad & \sum_{t=1}^{T}\sum_{i=1}^{I} c_i(t)\left(\pi_{i,t}^0 + \sum_{r\in I_t}\pi_{i,t}^r d_r\right) \le F \\
& 0 \le \pi_{i,t}^0 + \sum_{r\in I_t}\pi_{i,t}^r d_r \le P_i(t), \quad i = 1, ..., I, \quad t = 1, ..., T \\
& \sum_{t=1}^{T}\left(\pi_{i,t}^0 + \sum_{r\in I_t}\pi_{i,t}^r d_r\right) \le Q(i), \quad i = 1, ..., I \\
& \mathrm{V}_{\min} \le v(1) + \sum_{s=1}^{t}\left(\sum_{i=1}^{I}\pi_{i,s}^0 + \sum_{r\in I_s}\pi_{i,s}^r d_r\right) - \sum_{s=1}^{t} d_s \le \mathrm{V}_{\max} \quad t = 1, ..., T. \\
& \forall\{d_t \in [d_t^* - \theta d_t^*, d_t^* + \theta d_t^*], \quad t = 1, ..., T\}.
\end{aligned} \tag{8}$$

To convert model (8) into an equivalent LP model which I can put into a solver, I define certain additional analysis variables mentioned in [1] and obtain a model seen in (39) in [1].

## 4.3 An illustrative example

Now, our next step is to attach values to the parameters such that the problem can be solved. The paper attaches the following values to the parameters, for details see [1]:

- $I = 3$;
- $T = 24$ periods;
- $d_t^* = 1000\left(1 + \frac{1}{2}\sin\left(\frac{\pi(t-1)}{12}\right)\right)$; $\qquad t = 1,.........,24$
- $d_t \in [(1-\theta)d_t^*, (1+\theta)d_t^*]$;
- $c_i(t) = \alpha_i\left(1 + \frac{1}{2}\sin\left(\frac{\pi(t-1)}{12}\right)\right)$; $\qquad t = 1,.........,24$
  $\alpha_1 = 1$
  $\alpha_2 = 1.5$
  $\alpha_3 = 2$
- $P_i(t) = 567$ units;
- $Q_i = 13600$;
- $V_{\text{min}} = 500$ units;
- $V_{\text{max}} = 2000$ units.

The initial inventory $v(1)$ is known beforehand. However the paper [1] did not mention the value of this initial inventory. Therefore, I set this initial inventory in the middle of the minimum and maximum inventory. Thus:

$$v(1) = 1250 \text{ units.}$$

## 4.4 Results of the AARC approach

The AARC model (8) with the parameters defined in the previous Section, is solved by the commercial solver GUROBI [4].
In Figure 9 in the Appendix, the solutions of the AARC model are plotted against different values of $\theta$. As seen in the Figure, there exists a strong linear relation between the value of $\theta$ and the solution of the model. If the value for $\theta$ increases, the uncertainty set increases, and that leads to an increase of the costs. Indeed, when the model has to take into account more uncertainty of the demands, the factories need to have a wider range of inventory, this leads to higher costs since the factories have to be prepared for different sizes of demands.
In table 1, one can see the optimal production costs of some values of $\theta$.

**Table 1:** Uncertainty level vs. total production cost.

| AARC model | |
|---|---|
| Uncertainty value $\theta$ | Total minimum production cost |
| 0.00 | 32074 |
| 0.02 | 33099 |
| 0.05 | 34642 |
| 0.10 | 37258 |
| 0.20 | 42569 |
| 0.25 | 45259 |

# 5 Comparative study of the "here and now" decisions

## 5.1 Introduction

This Chapter is used to answer the main question of the thesis: "Do the optimal "here and now" decisions differ when we solve the problem with uncertainty and affine decision rules included i.e. the AARC approach, vs. solving it for different single-trajectory realizations of demand?". This question is raised to investigate if the sophistication of the complex AARC method is really worth the effort. If it is the case that a simpler method, i.e. the single-trajectory demand approach, obtained the same or even better "here and now" decisions, performing complicated methods could be superfluous.

A single-trajectory of demands is a set of demands for every time period $t$ of the period horizon. Every demand in time period $t$ has to lie in an earlier described set depending on the aforementioned variable $d_t^*$ and the uncertainty variable $\theta$.

**The single-trajectory demand approach** For this approach we solve a problem multiple times for different single-trajectory demands with a model that does not include the whole uncertainty set, as the AARC approach does, but solves it for 1 particular demand trajectory. To solve the problem for different single-trajectory demands, we use the nominal model.
In the nominal model you think you know the entire demand realization and you optimize for fixed decisions in each time period, i.e. a single-trajectory of demands. The nominal model I use is defined as follows [1]:

$$
\begin{aligned}
\min_{p_i(t),F} \quad & F \\
\text{s.t.} \quad & \sum_{t=1}^{24}\sum_{i=1}^{3}\left(1+\frac{(i-1)}{2}\right)\left(1+\frac{1}{2}\sin\left(\frac{\pi(t-1)}{12}\right)\right)p_i(t) \leq F \\
& 0 \leq p_i(t) \leq 567, \quad i=1,2,3, \quad t=1,...,24 \\
& \sum_{t=1}^{24} p_i(t) \leq 13600, \quad i=1,2,3 \\
& 500 \leq 1250 + \sum_{s=1}^{t}\sum_{i=1}^{3} p_i(s) - \sum_{s=1}^{t} d_s \leq 2000, \quad t=1,...,24.
\end{aligned} \tag{9}
$$

The single-trajectory realization of demands $\{d_1, ..., d_s, ..., d_T\}$ I make by randomly computing a column-vector with $T$ entries, each entry is a demand at time step $t \in [1, T]$. This means that every time I run this model, the outcome will not be the same, since the demand realization is different every time.

This Chapter analyses this nominal model and then compares the made "here and now" decisions of the single-trajectory demand approach with the "here and now" decisions of the AARC approach.

For a reliable comparative study of the "here and now" decisions, a big sample has to be developed. This big sample consists of 4000 observations of the calculated optimal ordering decisions of the nominal model, and the 4000 different demand trajectories for which it is constructed.

Solving the nominal model with the constructed single-trajectory demand, is having a realistic forecast of demands, rather than knowing exactly what the demand is going to be. This means that in time, things change and the realization differs from our realistic good forecast.
Thus, our "here and now" decision is the ordering decision only at time step $t = 1$, because these are the only things that need to be implemented immediately. The ordering decisions for the other time steps, can be "re-optimized" when we know the actual demand of time step $t = 1$.

There are 2 samples constructed, one with an uncertainty level of $\theta = 0.2$ and one with an uncertainty level of $\theta = 0.05$. Shortly analysing the big sample of 4000 observations with an uncertainty level $\theta = 0.2$, I notice a couple of things:

1. The ordering decision at every time step except $t = 24$, factory 1 produces the maximal production capacity (567).
2. Particular, the optimal "here and now" decision for both factory 1 and 2 are always 567 (maximal production capacity).
3. When looking at the ordering decisions of factory 3 at all observations, most of the time steps factory 3 produces 0 products.
4. At the last time period $t = 24$, all factories produce 0 product.
5. Most of the time the "here and now" decision of factory 3 is 567, but some are below 567.

Analysing the big sample with uncertainty level $\theta = 0.05$ I notice the following things:

1. The ordering decision at every time step except $t = 24$, factory 1 produces the maximal production capacity (567).
2. Particular, the optimal "here and now" decision for both factory 1 and 2 are always 567 (maximal production capacity).
3. When looking at the ordering decisions of factory 3 at all observations, most of the time steps factory 3 produces 0 products.
4. At the last time period $t = 24$, all factories produce 0 product.
5. The "here and now" decision of factory 3 is almost always 567, so factory 3 produces at time $t = 1$ a lot more compared to the sample with $\theta = 0.2$

So for both the samples I notice the same things. However, the second sample shows bigger "here and now" decisions of factory 3 compared to the first sample.

## 5.2 Statistical analysis

In this Section I do statistical tests to verify my observations mentioned in the previous Section. I do this to get more understanding of what is happening in the nominal model before moving on to the comparative analysis.

The first and third observations go hand in hand. From the definition of $c_i(t)$, we can see that the production cost per time period $t$ is the most expensive for factory 3 and less expensive for factory 1. Therefore, the warehouse assigns first factory 1 the maximal production capacity, then goes to factory 2, sees how many demand still need to be satisfied, and most of the time that's more than the maximal production capacity, so the warehouse assigns to factory 2 the maximal production capacity as well. The remaining demand that needs to be satisfied is left for factory 3, since factory 3 is the most expensive factory. If factory 1, factory 2 and the inventory already satisfy the demand and the minimal inventory, factory 3 does not have to produce any products.

This explains partly also the second observation. The initial inventory compared to the demand is very low, therefore both factory 1 and 2 have to produce the maximal production capacity (567) at time step $t = 1$.

### 5.2.1 Correlation between variables

The last observation is explained by the amount of demand at time step $t = 1$. When the demand is large, factory 3 has to produce more of the product. To see the relation between the demand at $t = 1$ and the "here and now" decision of factory 3, I compute the correlation coefficient between the 2 variables. I do this for the sample with $\theta = 0.2$ and for the sample with $\theta = 0.05$.

**Uncertainty level $\theta = 0.2$** For the sample with the largest uncertainty set, I obtain the following correlation coefficient:

$$r = 0.6826$$

This correlation coefficient is a numerical measure of the linear relationship between the response and explanatory variables.
Since we have an $r$ of 0.6826, it indicates that there is a positive correlation between the 2 variables. This means that when the demand increases, the "here and now" decision of factory 3 increases as well. The strength of the correlation is between a moderate and strong (linear) relationship.
To see how the relationship looks like we use regression. For this I use the features of the statistics software R [7].

Figure 1 shows a scatterplot of the "here and now" decisions of factory 3 and the demands at time step $t = 1$. The line is the regression line that shows the best fit of the data. One thing I notice right away are the two thick black lines on the top of the plot. The horizontal line is easily explained since the maximal production capacity is 567, which is exactly this horizontal line. No fitted "here and now" decisions will go above this line. The thickness of this line is explained by the fact that we noticed that a lot "here and now" decisions of factory 3 are equal to 567. So a lot of points lay on this line.
The points that do not lie on these 2 lines, are "here and now" decisions that correspond to a relatively high demand at $t = 1$ compared to demands at time periods $t \in [2, T]$. Indeed, when

the demands at time periods $t \in [2, T]$ are relatively low, factory 3 does not have to produce as much as when the demands are larger.

List 1 in the Appendix shows the output of a simple linear regression model, with response variable: "here and now" decision of factory 3 and the explanatory variable: the demand at time step $t = 1$.
The F-statistic is 3487 on 1 and 3996 degrees of freedom. Since $3487 \gg 1$, we can confidently reject the null hypothesis that there is no relationship between the two variables. This confirms that there is indeed a relationship between the "here and now" decision of factory 3 and the demand at time step $t = 1$ which was also confirmed by the correlation coefficient as well.
From the output we can see that the linear model has an adjusted $R^2$ of 0.4657, this indicates that roughly $46,6\%$ of the variance found in the "here and now" decisions can be explained by the amount of demand.
From this linear regression we obtain a formula of the form:

$$\widehat{HAN}_3 = -48.7975 + 0.5467d_1 \tag{10}$$

With $\widehat{HAN}_3$ the fitted value of the "here and now" decision of factory 3, expressed in terms of the demand at time step $t = 1$.

**Uncertainty level $\theta = 0.05$** For the sample with uncertainty level $\theta = 0.05$ I obtain a lower correlation coefficient in comparison with the first sample. The following correlation coefficient is obtained:

$$r = 0.4272$$

This means that the relationship between the "here and now" decision and the demand at time step $t = 1$ is less strong. Despite the fact that this correlation is lower than the 0.6826 of the first model, this model has also a significant relationship between the "here and now" decision of factory 3 and the demand at time step $t = 1$. This correlation coefficients indicates that it's a positive relationship that's between weak and moderate.

In Figure 2 one can see the scatterplot of this sample, with the regression line. Again, we can see a horizontal thick line at $y = 567$. The same argument as in the previous Section applies here. Again we can see some dots below this line and again the same argument as in the previous Section applies here.

List 2 in the Appendix shows the output of plotting a linear model to fit the data. The response variable is the "here and now" decision at factory 3 with an explanatory variable the demand at time step $t = 1$. Again, the F-statistic is $\gg 1$, and I can confidently say that the null hypothesis $H_0$ that there exist no relationship, can be rejected. Furthermore, the model produced an adjustable $R^2$ of 0.1823, which is lower than the 0.4657 produced in the first model.

From this linear regression we obtain a formula of the form:

$$\widehat{HAN}_3 = 419.0 + 0.1446d_1 \tag{11}$$

Due to the purpose of the report, further analysing will not be done here. However, a short look into the QQplot of the residuals for the first sample, which can be seen in Figure 10 in the appendix, tells us that the residuals are likely not to be normally distributed. This argument is confirmed by the low P-value of the Shapiro-Wilk and Kolmogorov-Smirnov tests. That means that we reject the null hypothesis that the residuals come from a normal distribution. Further analysis can be done here, by for example transforming the data, removing big outliers or investigating other forms of regression.

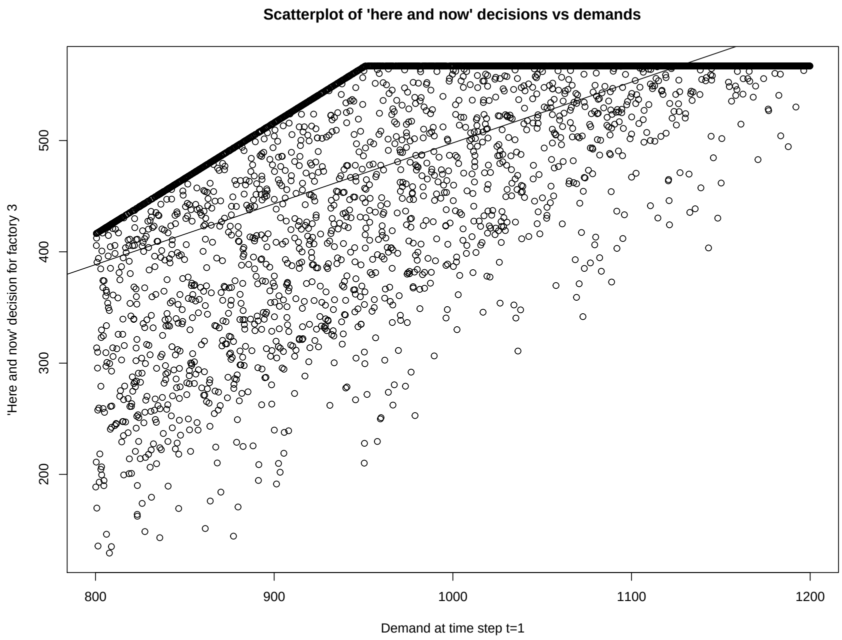


**Figure 1:** A scatterplot of the "here and now" decisions of factory 3 and the demands at time step $t = 1$. Calculated with uncertainty level $\theta = 0.2$.

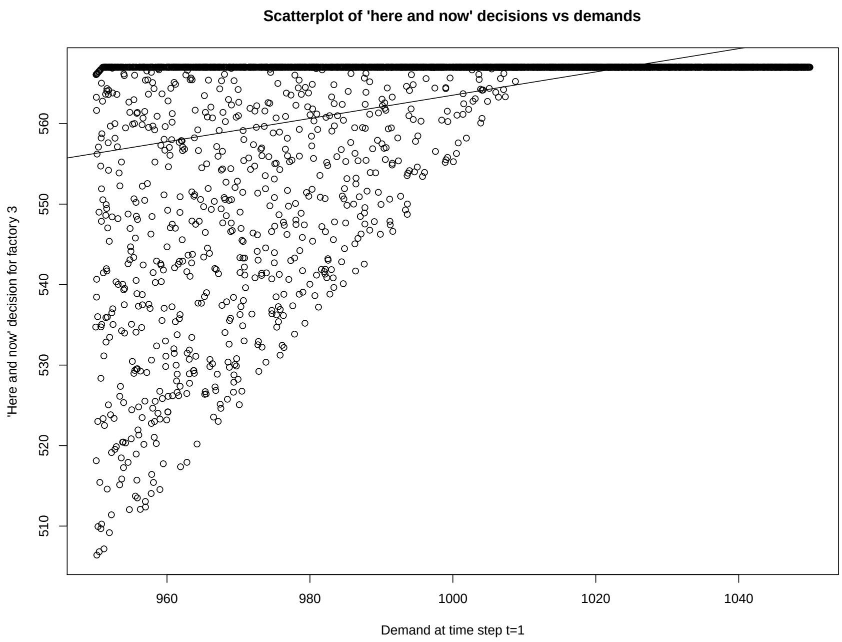


**Figure 2:** A scatterplot of the "here and now" decisions of factory 3 and the demands at time step $t = 1$. Calculated with uncertainty level $\theta = 0.05$.

## 5.3 The single-trajectory approach vs the AARC approach

In this Section I analyse the difference between the "here and now" decisions of the AARC approach vs the "here and now" decisions of the approach that solves the problem for multiple single-trajectory demands with the nominal model.

### 5.3.1 Uncertainty level $\theta = 0.2$

Solving the AARC model, with

$$\theta = 0.2.$$

We obtain the following optimal "here and now" decisions for factory 1,2, and 3:

**Table 2:** Optimal "here and now" decisions for the AARC approach.

| AARC model | | |
|---|---|---|
| Factory 1 | Factory 2 | Factory 3 |
| 567.0 | 567.0 | 416.0 |

I begin with factory 1 and 2. Due to the fact that I noticed in the beginning of the Chapter that the optimal "here and now" decision of the nominal model of factory 1 and 2 are always 567 and the optimal "here and now" decisions of the AARC of factory 1 and 2 are 567 as well, making a histogram and analysing results is cumbersome and we disregard the 2 factories for now and focus on factory 3 only.

Plotting the "here and now" decisions of factory 3 of the first sample ($\theta = 0.2$) in a histogram, I obtain the following Figure:

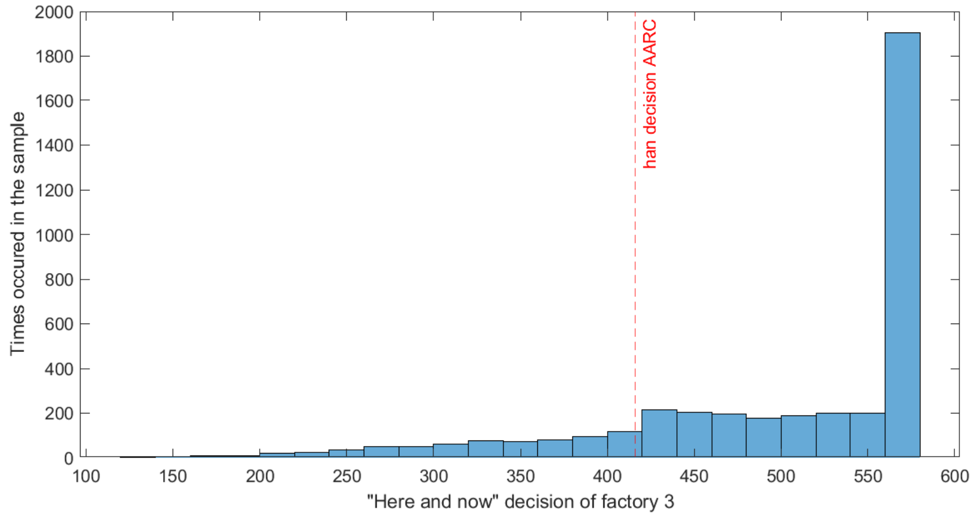


**Figure 3:** The optimal "here and now" decisions of factory 3. $\theta = 0.2$.

For a more detailed overview of the exact heights of the bins, I refer to table 10 in the appendix.

The vertical dotted red line is the "here and now" decision of factory 3 of the affine decision model (AARC approach). From this plot we see the following, the nominal model with a single demand trajectory, produces most of the time higher "here and now" decisions for factory 3 compared to the AARC approach. Almost half of the observations had a "here and now" decision of factory 3 of 567 units. This is far higher than the "here and now decision" of the AARC model of 416. So at this point, we see that considering affine decisions and considering uncertainty, did contribute to determining the "here and now" decision of factory 3.

### 5.3.2 Uncertainty level $\theta = 0.05$

In this Section instead of $\theta = 0.2$, I use $\theta = 0.05$. This means that the uncertainty set is smaller. So there will be less uncertainty and the realization of the demands will be closer to the nominal demand $d_t^*$. Solving the AARC model with this value of $\theta$, we obtain the following optimal "here and now" decisions for factor 1,2 and 3:

**Table 3:** Optimal "here and now" decisions for the AARC approach.

| AARC model | | |
|---|---|---|
| Factory 1 | Factory 2 | Factory 3 |
| 567.0 | 567.0 | 566.0 |

Again factory 2 and 3 will be disregarded due to the same arguments as mentioned in the previous Section. The "here and now" decision of factory 3 has become larger compared to the decision of the previous Section. Indeed, when the uncertainty decreases, the optimal decisions become more similar to the ones of the nominal model. There is less uncertainty to consider and take into account.

Plotting the "here and now" decisions of factory 3 of the second sample ($\theta = 0.05$) in a histogram, I obtain the following Figure:

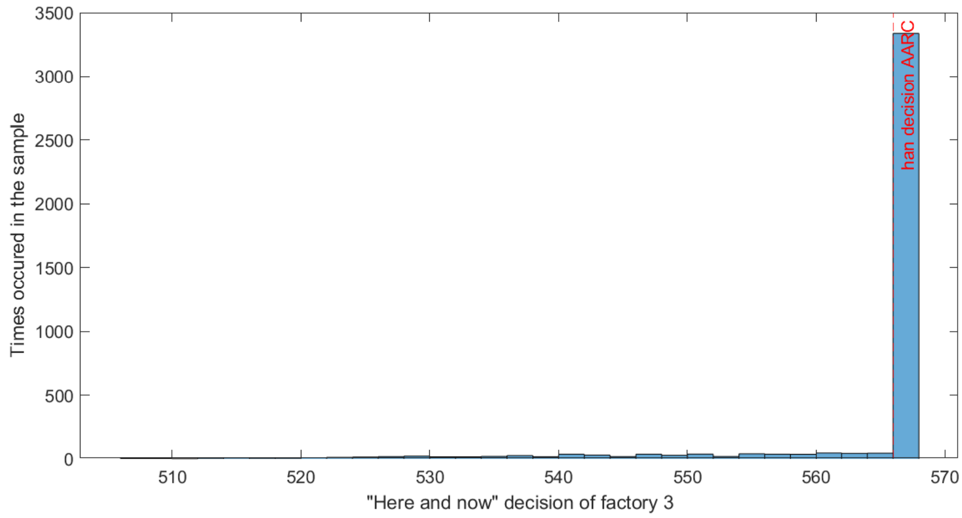


**Figure 4:** The optimal "here and now" decisions of factory 3. $\theta = 0.05$.

For a more detailed overview of the exact heights of the bins, I refer to Table 11 in the appendix.

As expected we can indeed see that the optimal decision of the AARC approach is closer to the large peak in the histogram. So the optimal "here and now" decision of the AARC model is closer to the most frequent optimal "here and now" decision of the nominal model.

I notice then, that when the degree of uncertainty decreases, the producer tends to produce more in factory 3 at time step $t = 1$. Indeed, when plotting the values of $\theta$ versus the "here and now" decision of factory 3, we obtain a negative linear relationship, so when $\theta$ increases, and thus the uncertainty increases, the producer decides to produce less product at time step $t = 1$ at factory 3.
The horizontal line from 0 to 0.05 in Figure 5 is explained by the maximal production capacity of 567.

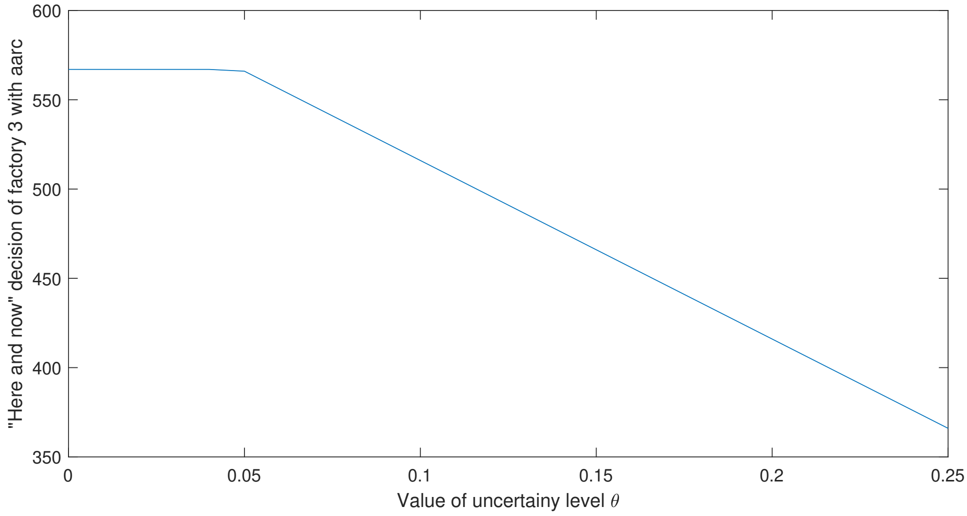


**Figure 5:** The optimal "here and now" decisions of factory 3 versus values of $\theta$'s.

We investigate Figure 5 and 9 together and we see that an increase in $\theta$ leads to an increase in optimal values for the total production cost and a decrease in "here and now" decisions. This leads to a negative relationship between the "here and now" decisions of factory 3 and the optimal total production cost, solved with the AARC approach. Indeed, figure 6 shows this relationship.

The "here and now" decisions of this model are the highest when the optimal total production cost is lowest. Since the production cost per time step of factory 3, $c_3(t)$, has low cost in the first time step $t = 1$, this means that when the factory 3 has to produce, it is the most beneficial to do this in the first time period. Thus 'better' "here and now" decisions can be interpreted as higher "here and now" decisions since this leads to lower minimum total production costs.

So we may conclude that with an uncertainty level of $\theta = 0.2$, the AARC approach produces lower "here and now" decisions, i.e. AARC produces worse "here and now" decisions.
With an uncertainty of $\theta = 0.05$, the "here and now" decisions are comparable.

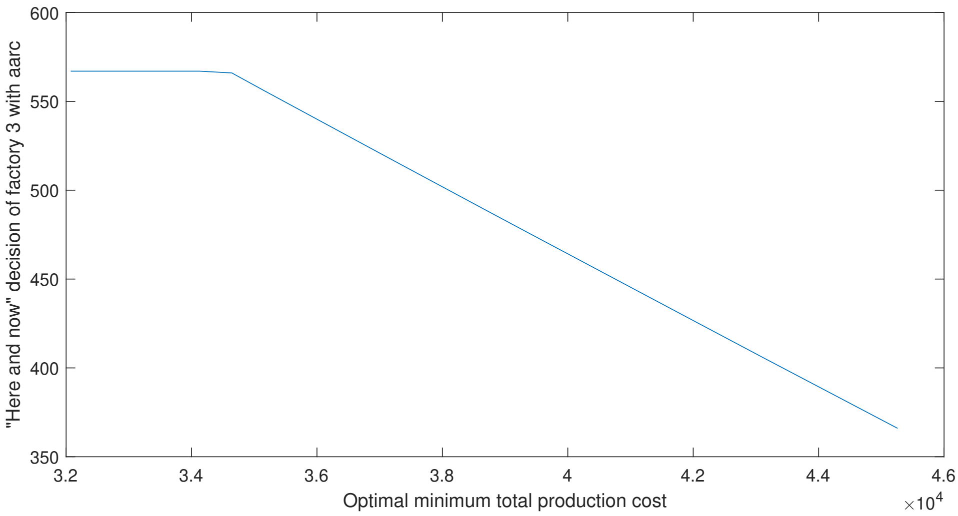


**Figure 6:** The optimal "here and now" decisions of factory 3 versus the optimal minimum total production costs solved with the AARC approach.

# 6 Regression of the "here and now" decisions and the demand trajectory

## 6.1 Introduction

This Chapter analyses if there exists a relationship between the "here and now" decisions of the nominal model versus the whole demand trajectory. I will perform this analysis because I want to investigate if finding "here and now" decisions of factory 3 requires the whole demand trajectory. This analysis may lead to the conclusion that for determining the "here and now" decisions of factory 3, I do not need the whole demand trajectory but just a minor subset of this. If I discover that I can predict the "here and now" decisions of factory 3 by a part of the demand trajectory, I save a lot of effort since I do not have to solve the whole optimisation problem.

As well as in the previous Chapter, factory 1 and factory 2 will be disregarded and I will focus on factory 3 only.
In Section 5.2 I analysed and discussed the relationship between the "here and now" decision of factory 3 and the demand at time step $t = 1$. However, this model is solved for not one demand, but T demands, that means that the "here and now" decisions do not have to depend only on the demand at time step $t = 1$. It is possible that the "here and now" decision also depends on the second,third, and/or last demand.

## 6.2 Uncertainty level $\theta = 0.2$

To check this, I first determine the coefficient estimates of them separately, as seen in table 4. The intercept is

$$\alpha = -574.3719$$

**Table 4:** The coefficient estimates for the "here and now" decision of factory 3 of the demands at every time step $t$.

| Coefficient estimates | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $t=1$ | $t=2$ | $t=3$ | $t=4$ | $t=5$ | $t=6$ | $t=7$ | $t=8$ | $t=9$ | $t=10$ | $t=11$ | $t=12$ |
| 0.5539 | 0.3581 | 0.0875 | -0.0022 | -0.0051 | 0.0050 | -0.0054 | 0.0127 | -0.0001 | -0.0027 | -0.0061 | 0.0189 |
| $t=13$ | $t=14$ | $t=15$ | $t=16$ | $t=17$ | $t=18$ | $t=19$ | $t=20$ | $t=21$ | $t=22$ | $t=23$ | $t=24$ |
| 0.0042 | -0.0021 | 0.0102 | -0.0049 | -0.0213 | 0.0199 | 0.0231 | -0.0271 | -0.0139 | 0.0048 | -0.0022 | -0.0109 |

The first 2 time step coefficients are larger compared to the others. To see which time steps are significant, we use multiple linear regression. Again we use the software R [7]. As seen in list 3 in the appendix the output of the multiple linear regression is shown. The response variable is the optimal "here and now" decision of factory 3 and the explanatory variables are all the demands at time step $t$.

Multiple assumptions are made in multiple linear regression. One of them is the assumption that the residuals are normally distributed. Otherwise the errors are not consistent across your whole data and the explanatory variables can mean different things at different levels of the response variable. Non-normal distributed residuals can thus lead to regression models that are not reliable.
Examining the residuals, I see that they are not strongly symmetrical distributed. This means that the residuals might not be normal distributed. When looking at the residual standard

error, i.e. the difference between the observed and predicted "here and now" decision, the value is 47.34. This is approximately between 8% and 12% of the observed decision.

Our adjusted R-squared is 0.7367, so 73,7% of variation of the "here and now" decision of factory 3 can be explained by the model (with the whole demand trajectory).

The corresponding F-value is equal to 457.3 on 24 and 3975 degrees of freedom. This means that we can safely reject the null hypothesis $H_0$ that no explanatory variable is significantly related to the response variable, i.e. all the model coefficients are 0. This argument is also supported by the value of the overall P-value: $< 2.2e - 16$. To determine which do and which do not, we investigate the values of "Pr(> $|t|$)".
In table 5 the P-values for all the explanatory variables are shown.

The rows that are colored have a P-value $< \alpha$. For this report I use an significance level $\alpha = 0.05$, i.e. P-value $< 0.05$. This means that for those variables that are colored, we reject the null hypothesis $H_0$ that the significance of those variables is zero. For the remaining variables we see a P-value larger than the significance level 0.05. However, we cannot remove these variables right away.

| The explanatory variable | P-value |
|---|---|
| Intercept | < 2e-16 |
| $d_1$ | < 2e-16 |
| $d_2$ | < 2e-16 |
| $d_3$ | < 2e-16 |
| $d_4$ | 0.650498 |
| $d_5$ | 0.263634 |
| $d_6$ | 0.241306 |
| $d_7$ | 0.220712 |
| $d_8$ | 0.003616 |
| $d_9$ | 0.990762 |
| $d_{10}$ | 0.579966 |
| $d_{11}$ | 0.236942 |
| $d_{12}$ | 0.000926 |
| $d_{13}$ | 0.516674 |
| $d_{14}$ | 0.784077 |
| $d_{15}$ | 0.238192 |
| $d_{16}$ | 0.626441 |
| $d_{17}$ | 0.066072 |
| $d_{18}$ | 0.112972 |
| $d_{19}$ | 0.074616 |
| $d_{20}$ | 0.032649 |
| $d_{21}$ | 0.224869 |
| $d_{22}$ | 0.632714 |
| $d_{23}$ | 0.797319 |
| $d_{24}$ | 0.140905 |

**Table 5:** P-values corresponding to the explanatory variables of the multiple linear regression.

To choose which variables to delete, I use the so called backward selection. One at a time I remove the variables with the highest P-value and re-estimate the model until only variables with P-value $< 0.05$ remain. This is done since the P-values are adjusted by other terms in the model. As seen in table 5 the first variable to delete is $d_9$. This results in still too many high P-values. Following the pseudocode written in algorithm 1, I obtain a list of explanatory variables seen in table 6.

**Algorithm 1** Algorithm to remove insignificant variables

**Result:** List of significant variables for multiple linear model
$variables \leftarrow [intercept, d_1, d_2, ..., d_{24}]$
$pvalues \leftarrow [pvalue_{intercept}, pvalue_1, ..., pvalue_{24}]$
**while** $max\{pvalues\} \geq 0.05$ **do**
  remove variable with highest pvalue from *variables*. Running this updated model again with *variables* and updating *pvalues*.
**end**
return variables and their pvalues

| The explanatory variable | P-value |
|---|---|
| Intercept | < 2e-16 |
| $d_1$ | < 2e-16 |
| $d_2$ | < 2e-16 |
| $d_3$ | < 2e-16 |
| $d_8$ | 0.003616 |
| $d_{12}$ | 0.000926 |
| $d_{20}$ | 0.032649 |

**Table 6:** P-values corresponding to the explanatory variables after backward elimination.

The output of this 'new' multiple linear regression is seen in list 4. Analysing this output shortly we see that again the adjusted $R^2$ is 0.737. Furthermore the residuals are still not very symmetric distributed. However, the F-value is still $\gg 1$ and thus we can safely reject the null hypothesis that no explanatory variable is significantly related to the response variable. The overall P-value is also $< 2.2e - 16$ which confirms the latter statement.

This means I finally obtain a linear relationship of the form:

$$\widehat{HAN}_3 = -593.7 + 0.5536d_1 + 0.3586d_2 + 0.0873d_3 + 0.0122d_8 + 0.0180d_{12} - 0.0270d_{20} \quad (12)$$

With $\widehat{HAN}_3$ the fitted optimal "here and now" decision of factory 3 expressed in terms of the demand at time periods $t$.

### 6.3 Uncertainty level $\theta = 0.05$

For the second sample with uncertainty level $\theta$ equal to 0.05, I do the same procedure. So I first determine the coefficient estimates separately, as seen in table 7. The intercept is estimated at

$$\alpha = 243.2806$$

**Table 7:** The coefficient estimates for the "here and now" decision of factory 3 of the demands at every time step $t$.

| Coefficient estimates | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $t = 1$ | $t = 2$ | $t = 3$ | $t = 4$ | $t = 5$ | $t = 6$ | $t = 7$ | $t = 8$ | $t = 9$ | $t = 10$ | $t = 11$ | $t = 12$ |
| 0.1481 | 0.1365 | 0.0053 | -0.000 | -0.0041 | 0.0035 | -0.0007 | 0.0073 | -0.0033 | -0.0023 | 0.0008 | 0.0091 |
| $t = 13$ | $t = 14$ | $t = 15$ | $t = 16$ | $t = 17$ | $t = 18$ | $t = 19$ | $t = 20$ | $t = 21$ | $t = 22$ | $t = 23$ | $t = 24$ |
| -0.0018 | -0.0001 | 0.0160 | -0.0035 | -0.0069 | 0.0121 | 0.0151 | -0.0196 | -0.0113 | 0.0089 | -0.0011 | -0.0080 |

The output of the multiple linear model is shown in list 5 of the appendix. For this sample the residuals are again not completely symmetric as well. The residual standard error is 7.594, which is somewhat better than the 47.34 of the first sample. However, again non-normality can occur. The adjusted R-squared of this sample is 0.3891 which is lower than the adjusted R-squared of the first sample. The generated F-value is 107.1 on 24 and 3975 degrees of freedom. Together with it's P-value of $<$ 2.2e-16, we are allowed to reject the null hypothesis $H_0$ that no explanatory variable is significantly related to the response variable.

Once more we use the values of "Pr($>|t|$)". These values are represented in table 8. Again the colored variables are the variables that have a P-value $\geq 0.05$. We see that $d_3$ has a higher P-value and $d_{15}$ a much lower P-value. At this stage we can reject the null hypothesis $H_0$ that the significance of $d_1$, $d_2$, $d_8$, $d_{12}$, $d_{15}$ and $d_{20}$ is zero.

| The explanatory variable | P-value |
|---|---|
| Intercept | < 2e-16 |
| $d_1$ | < 2e-16 |
| $d_2$ | < 2e-16 |
| $d_3$ | 0.11468 |
| $d_4$ | 0.99277 |
| $d_5$ | 0.16245 |
| $d_6$ | 0.20810 |
| $d_7$ | 0.79533 |
| $d_8$ | 0.00892 |
| $d_9$ | 0.25495 |
| $d_{10}$ | 0.46272 |
| $d_{11}$ | 0.82138 |
| $d_{12}$ | 0.01339 |
| $d_{13}$ | 0.66306 |
| $d_{14}$ | 0.97582 |
| $d_{15}$ | 0.00362 |
| $d_{16}$ | 0.58755 |
| $d_{17}$ | 0.35730 |
| $d_{18}$ | 0.13673 |
| $d_{19}$ | 0.06964 |
| $d_{20}$ | 0.01509 |
| $d_{21}$ | 0.12507 |
| $d_{22}$ | 0.17376 |
| $d_{23}$ | 0.84083 |
| $d_{24}$ | 0.09333 |

**Table 8:** P-values corresponding to the explanatory variables of the multiple linear regression.

To find out if we can safely remove the remaining variables with a P-value $\geq 0.05$, we use Algorithm 1 again. We obtain eventually the significant variables shown in Table 9.

| The explanatory variable | P-value |
|---|---|
| Intercept | < 2e-16 |
| $d_1$ | < 2e-16 |
| $d_2$ | < 2e-16 |
| $d_8$ | 0.01704 |
| $d_{12}$ | 0.01833 |
| $d_{15}$ | 0.00346 |
| $d_{20}$ | 0.01612 |

**Table 9:** P-values corresponding to the explanatory variables after backward elimination.

The output of this 'new' multiple linear regression is seen in list 6. One more time analysing this output shortly we see that again the adjusted $R^2$ is 0.3886. Furthermore the residuals are still not very symmetrically distributed. However, the F-value is still $\gg 1$ and thus we can safely reject the null hypothesis that no explanatory variable is significantly related to the response variable. The overall P-value is also $< 2.2e - 16$ which confirms the latter statement.

To conclude, we obtain a linear relationship of the form:

$$\widehat{HAN_3} = 239.1645 + 0.1483d_1 + 0.1368d_2 + 0.0067d_8 + 0.0087d_{12} + 0.0161d_{15} - 0.0193d_{20} \quad (13)$$

### 6.4 Checking assumptions of multiple linear regression

This Section is used to check all the assumptions of multiple linear regression and improving my linear model. I already mentioned that the residuals do seem non-normally distributed. This Section elaborates more on this argument and introduces how to fix this flaw in the model.

For multiple linear regression, the following assumptions are made [5]:

1. Linear relationship between dependent and independent variable
2. Multivariate normality of the errors
3. No multicollinearity

**Uncertainty level** $\theta = 0.2$ The first assumption is met. Plotting every independent variable against the dependent variable, will lead in 6 scatterplots with a linear form.
I check the second assumption again but now by generating the QQPlot and histogram of the residuals. If the residuals are normally distributed, the qqplot shows points nicely along it's qqline. The histogram has to be in the form of a normal distribution, i.e. bell-shaped.

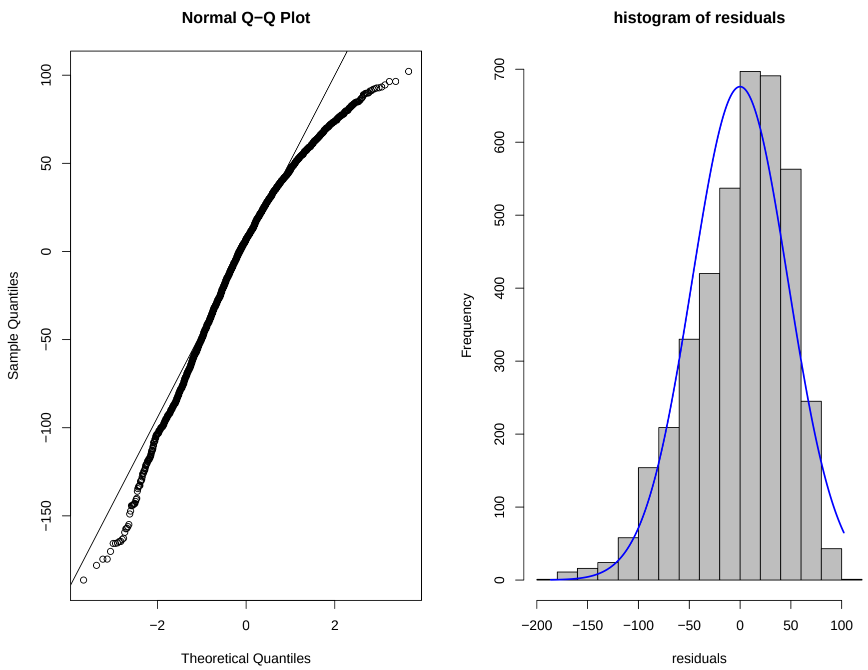


**Figure 7:** Histogram and QQ plot of the residuals for sample with $\theta = 0.2$

As seen in Figure 7, both the Figures do not show normality. The qq-plot has light tails and the histogram does not look like a nice bell-shaped histogram. This means that the second assumption is not met. This argument is also confirmed by plotting the residuals, Figure 8. The residuals show a clear pattern.
As seen in Figure 7 and 8 the residuals thus do not come from a normal distribution. This means there is room to improve our model.

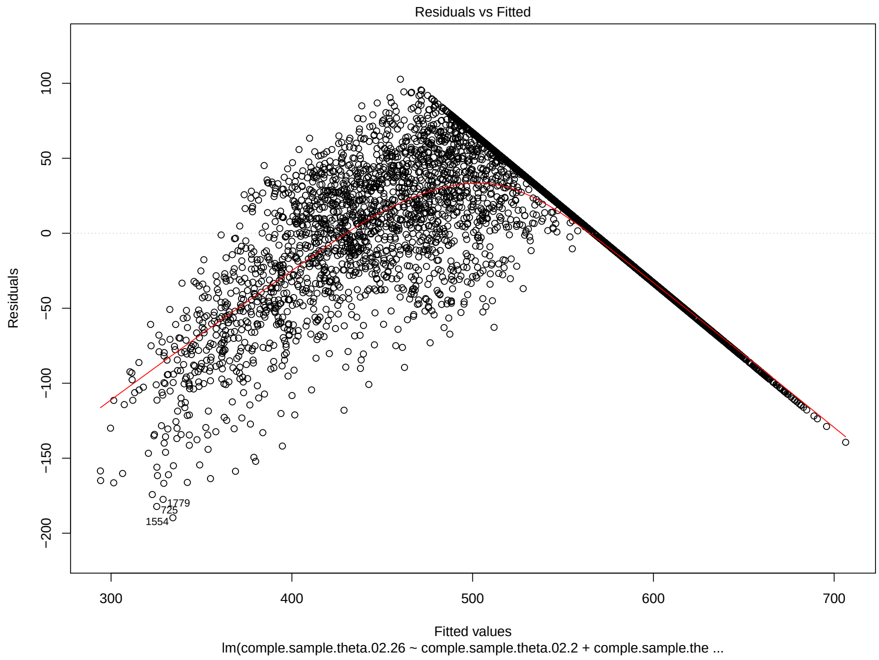


**Figure 8:** The residuals plotted against predicted values for the "here and now" decision of factory 3.

Before adjusting the regression or model we check the last assumption. All the independent variables must not be too highly correlated with each other. It is assumed that no variables can have a correlation coefficient $\geq 0.8$. To check this we use a so called correlation matrix. This correlation matrix shows the correlation between the different variables at once. Investigating this 25x25 matrix, apart from the diagonal, every entry is $< 0.8$. This means that the third assumption is met.

Since the assumptions of multiple linear regression are not met, we focus on other type of regressions. A famous regression that does not take normality of residuals into account, is called quantile regression [8]. This is done by using the conditional quantile equal to $\tau = 0.5$, which corresponds to median regression.
Again using the software R [7] and using algorithm 1, I obtain a linear model of the form:

$$\widehat{HAN}_3 = -479.92 + 0.52434d_1 + 0.31262d_2 + 0.06590d_3 + 0.01138d_8 + 0.02181d_{12} - 0.02542d_{17} \quad (14)$$

Calculating the goodness of the fit by pseudo-$R^2$ introduced for quantile regression [6], it produces a value of 0.4518 which indicates a decent goodness of fit.
Be aware that this pseudo-$R^2$ is a local measure of fit since it is computed for a specific value of $\tau$.

This means that we cannot 100% guarantee a good fit by looking at this value of pseudo-$R^2$. This pseudo-$R^2$ gives only a rough estimation. It can be that your model fits in the tails but does not anywhere else.

To come to an conclusion, it is debatable if the quantile regression or the regression in Section 6.2 is a better fit. The adjusted $R^2$ of the multiple linear regression is not comparable with the pseudo $R^2$ of the quantile regression.
However, since the residuals are not normally distributed, and the assumption of multiple linear regression aims that it is, quantile regression is in this respect an improvement.

# 7 Conclusion

During this project I investigated the affine adjustable RO approach, introduced by [1]. This approach is used for optimisation problems that include uncertainty. It splits the variables in adjustable and non-adjustable variables, where adjustable variables can alter themselves to realized data. Reflecting on the purpose of the report, I compared the "here and now" decisions of the AARC approach that considers all demand trajectories simultaneously with the "here and now" decisions of the nominal model solved for various simulations of demand trajectory. From this comparison I concluded that with a relatively large uncertainty set, $\theta = 0.2$, the "here and now" decisions of the AARC approach were smaller than those of the nominal model. This means that considering uncertainty and affine decision rules did in fact contribute to the "here and now" decisions. However smaller "here and now" decisions are interpreted as worse "here and now" decisions for this particular problem. This means that the AARC approach did not generate better "here and now" decisions compared to the nominal model.
Performing another comparative study, but then with less uncertainty freedom, $\theta = 0.05$, lead to "here and now" decisions that were more comparable. When the uncertainty freedom was low, the "here and now" decision of the AARC was close to the most frequent "here and now" decision of the nominal model. Thus given a low uncertainty level, the AARC approach gives "here and now" decisions close to the "here and now" decisions of the simpler method which only considers a single demand trajectory and no uncertainty.
Furthermore, what can be seen from these comparative studies is the ordering policy of the producer. When the uncertainty increases, the producer tends to produce less at factory 3 at time step $t = 1$.

All in all, I conclude that for the AARC approach and the considered inventory model, with a high enough uncertainty level, the effort of the paper is superfluous since the paper delivers lower, and thus worse "here and now" decisions compared to the "here and now" decisions of the simpler method.
With an adequate low uncertainty level, the simpler method produces comparable "here and now" decisions, and thus the approach of the paper of Ben-Tal et al. might be superfluous as well.

# 8 Discussion

As already mentioned in Section 5.2.1, it is possible to deeper investigate the regressions made in the report. The regressions made in this report often contained non-normal residuals and variables. For the scope of this report this is omitted, however for a perfect fit/prediction of the "here and now" decisions, further research on these residuals is recommended since the regression could give you unreliable results.

Another aspect that could be improved is the simulations for the nominal model. My code for this nominal model had a running time of 0.8 seconds approximately. Therefore a sample of only 4000 observations is used for 2 different values of $\theta$. The sample collection could be expanded to multiple samples each with different values of $\theta$ and for example 10.000 observations. The v(1) remained consistent in the whole project but solving the AARC model and nominal model for different values of v(1) might lead to different results.

In the conclusion I mentioned that the producer tends to produce less at time step $t = 1$ as the uncertainty increases. This statement can be investigated more to see how this relationship is developed exactly. If this holds if factory 1, 2 and 3 will have the same production costs. In this thesis I focused on one particular approach and example. In further research, the search area can be expanded and one could ask the same question for different approaches on RO, apart from AARC. For example investigating the "here and now" decisions of the $K$-adaptability approach of [3].
Also other examples of AARC can be investigated and see if the same phenomena happen there as well. For example a project management problem where multiple activities need to satisfy certain constraints that depend on other activities. This problem can be represented by direct graphs and nodes.

## References


[1] E. Guslitzer. A. Ben-Tal A. Goryashko. "Adjustable robust solutions of uncertain linear programs". In: *Math. Program.* 99 (2004), pp. 351–376. DOI: https://doi.org/10.1007/s10107-003-0454-y.

[2] D. Den Hertog B. L. Gorissen I. Yanıkoğlu. "A Practical Guide to Robust Optimization". In: *Omega* 53 (2015), pp. 124–131. DOI: https://doi.org/10.1016/j.omega.2014.12.006.

[3] W. Wiesemann G. A. Hanasusanto D. Kuhn. "K-adaptability in two-stage distributionally robust binary programming". In: *Operations Research Letters* 44 (2016), pp. 6–11. DOI: https://doi.org/10.1016/j.orl.2015.10.006.

[4] LLC Gurobi Optimization. *Gurobi Optimizer Reference Manual.* 2021. URL: http://www.gurobi.com.

[5] O.w. Jason and E. Waters. "Four assumptions of multiple regression that researchers". In: *Practical Assessment, Research, and Evaluation* 8 (2002), pp. 1–8. DOI: https://doi.org/10.7275/r222-hv23.

[6] R. Koenker and J. A. F. Machado. "Goodness of Fit and Related Inference Processes for Quantile Regression". In: *Journal of the American Statistical Association* 94.448 (1999), pp. 1296–1310. DOI: http://www.jstor.org/stable/2669943.

[7] RStudio Team. *RStudio: Integrated Development Environment for R.* RStudio, PBC. Boston, MA, 2020. URL: http://www.rstudio.com/.

[8] Great Learning Team. *What is Quantile Regression? — Introduction to Quantile Regression.* URL: https://www.mygreatlearning.com/blog/what-is-quantile-regression/.

# A Appendix

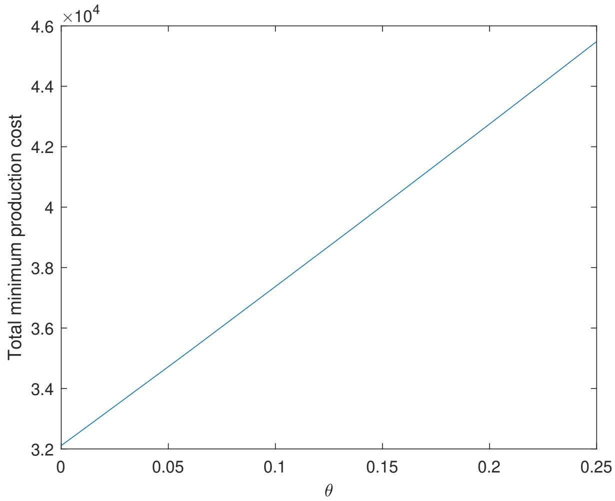


**Figure 9:** The relationship between the solution of the AARC model and the value of $\theta$. A larger value of $\theta$ leads to a bigger uncertainty set.

**Listing 1:** Regression of "here and now" decision factory 3 vs demand at first timestep with $\theta$ equal to 0.2.

```
Call:
lm(formula = data_st_sample$sample.3 ~ data_st_sample$sample.4)

Residuals:
    Min      1Q  Median      3Q     Max
-286.25  -31.12   11.35   50.25   95.90

Coefficients:
                          Estimate  Std. Error  t value  Pr(>|t|)
(Intercept)             -48.797457    9.324040   -5.234  1.75e-07 ***
data_st_sample$sample.4   0.546666    0.009257   59.053   < 2e-16 ***
---
Signif. codes:  0  ***  0.001  **  0.01  *  0.05  .  0.1

Residual standard error: 67.44 on 3998 degrees of freedom
Multiple R-squared:  0.4659,    Adjusted R-squared:  0.4657
F-statistic:  3487 on 1 and 3998 DF,  P-value: < 2.2e-16
```

**Listing 2:** Regression of "here and now" decision factory 3 vs demand at first time step with $\theta$ equal to 0.05.

```
Call:
lm(formula = data_st_sample2$sample2.3 ~ data_st_sample2$sample2.4)

Residuals:
    Min      1Q  Median      3Q     Max
-49.930  -1.876   1.431   5.031  10.549
```

```
Coefficients:
                             Estimate Std. Error t value Pr(>|t|)
(Intercept)                 4.190e+02  4.841e+00   86.55   <2e-16 ***
data_st_sample2$sample2.4  1.446e-01  4.839e-03   29.88   <2e-16 ***
---
Signif. codes:  0  ***  0.001  **  0.01  *  0.05  .  0.1

Residual standard error: 8.785 on 3998 degrees of freedom
Multiple R-squared:  0.1825,    Adjusted R-squared:  0.1823
F-statistic: 892.7 on 1 and 3998 DF,  P-value: < 2.2e-16
```

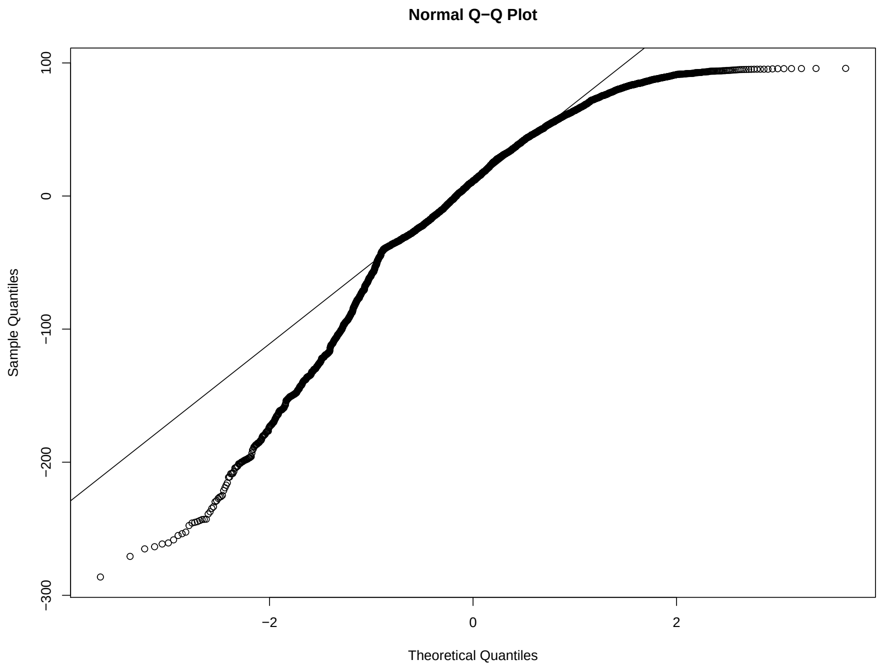


**Figure 10:** The normal QQ plot of the residuals of the "here and now" decision of factory 3 versus the demand at time step $t = 1$, with $\theta = 0.2$.

**Table 10:** The bin counts of the histogram for uncertainty level $\theta = 0.2$. The first row explains the width of the bar and the second row the height.

| Histogram bin counts | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 120 and 140 | 140 and 160 | 160 and 180 | 180 and 200 | 200 and 220 | 220 and 240 | 240 and 260 | 260 and 280 | 280 and 300 | 300 and 320 | 320 and 340 | 340 and 260 |
| 3 | 5 | 8 | 11 | 19 | 25 | 37 | 49 | 52 | 61 | 77 | 73 |
| 360 and 380 | 380 and 400 | 400 and 420 | 420 and 440 | 440 and 460 | 460 and 480 | 480 and 500 | 500 and 520 | 520 and 540 | 540 and 560 | 560 and 580 | |
| 80 | 94 | 119 | 214 | 203 | 197 | 179 | 188 | 201 | 199 | 1906 | |

**Table 11:** The bin counts of the histogram for uncertainty level $\theta = 0.05$. The first row explains the width of the bar and the second row the height.

| Histogram bin counts second sample | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 506 and 520 | 520 and 522 | 522 and 524 | 524 and 526 | 526 and 528 | 528 and 530 | 530 and 532 | 532 and 534 | 534 and 536 | 536 and 538 | 538 and 540 | 540 and542 | 542 and 544 |
| 34 | 10 | 12 | 15 | 18 | 22 | 16 | 16 | 20 | 26 | 17 | 37 | 30 |
| 544 and 546 | 546 and 548 | 548 and 550 | 550 and 552 | 552 and 554 | 554 and 556 | 556 and 558 | 558 and 560 | 560 and 562 | 562 and 564 | 564 and 566 | 566 and 568 | |
| 18 | 37 | 29 | 38 | 19 | 40 | 37 | 36 | 47 | 43 | 45 | 3338 | |

**Listing 3:** Multiple linear regression of the han decision of factory 3 vs the whole demand trajectory with $\theta$ equal to 0.2.

```
Call:
lm(formula = comple.sample.theta.02.26 ~ comple.sample.theta.02.2 +
    comple.sample.theta.02.3 + comple.sample.theta.02.4 +
    comple.sample.theta.02.5 + comple.sample.theta.02.6 +
    comple.sample.theta.02.7 +  comple.sample.theta.02.8 +
    comple.sample.theta.02.9 + comple.sample.theta.02.10 +
    comple.sample.theta.02.11 + comple.sample.theta.02.12 +
    comple.sample.theta.02.13 +  comple.sample.theta.02.14 +
    comple.sample.theta.02.15 + comple.sample.theta.02.16 +
    comple.sample.theta.02.17 + comple.sample.theta.02.18 +
    comple.sample.theta.02.19 +  comple.sample.theta.02.20 +
    comple.sample.theta.02.21 + comple.sample.theta.02.22 +
    comple.sample.theta.02.23 + comple.sample.theta.02.24 +
    comple.sample.theta.02.25, data = data_st_complete_sample)

Residuals:
     Min       1Q   Median       3Q      Max
-186.433  -29.881    6.985   35.688  102.147

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)              -5.744e+02  3.223e+01 -17.819  < 2e-16 ***
comple.sample.theta.02.2  5.539e-01  6.517e-03  85.006  < 2e-16 ***
comple.sample.theta.02.3  3.581e-01  5.853e-03  61.186  < 2e-16 ***
comple.sample.theta.02.4  8.754e-02  5.188e-03  16.874  < 2e-16 ***
comple.sample.theta.02.5 -2.176e-03  4.802e-03  -0.453 0.650498
comple.sample.theta.02.6 -5.058e-03  4.524e-03  -1.118 0.263634
comple.sample.theta.02.7  5.045e-03  4.305e-03   1.172 0.241306
comple.sample.theta.02.8 -5.356e-03  4.373e-03  -1.225 0.220712
comple.sample.theta.02.9  1.274e-02  4.377e-03   2.912 0.003616 **
comple.sample.theta.02.10 -5.208e-05 4.498e-03  -0.012 0.990762
comple.sample.theta.02.11 -2.664e-03 4.812e-03  -0.553 0.579966
comple.sample.theta.02.12 -6.111e-03 5.167e-03  -1.183 0.236942
comple.sample.theta.02.13  1.891e-02 5.705e-03   3.315 0.000926 ***
comple.sample.theta.02.14  4.221e-03 6.509e-03   0.649 0.516674
comple.sample.theta.02.15 -2.056e-03 7.504e-03  -0.274 0.784077
comple.sample.theta.02.16  1.019e-02 8.635e-03   1.180 0.238192
comple.sample.theta.02.17 -4.880e-03 1.003e-02  -0.487 0.626441
comple.sample.theta.02.18 -2.126e-02 1.157e-02  -1.838 0.066072 .
comple.sample.theta.02.19  1.993e-02 1.257e-02   1.585 0.112972
comple.sample.theta.02.20  2.310e-02 1.295e-02   1.783 0.074616 .
comple.sample.theta.02.21 -2.712e-02 1.269e-02  -2.137 0.032649 *
comple.sample.theta.02.22 -1.392e-02 1.147e-02  -1.214 0.224869
comple.sample.theta.02.23  4.846e-03 1.014e-02   0.478 0.632714
comple.sample.theta.02.24 -2.189e-03 8.522e-03  -0.257 0.797319
comple.sample.theta.02.25 -1.090e-02 7.405e-03  -1.473 0.140905
---
```

```
Signif. codes:  0  ***  0.001  **  0.01  *  0.05  .  0.1

Residual standard error: 47.34 on 3975 degrees of freedom
Multiple R-squared:  0.7383,   Adjusted R-squared:  0.7367
F-statistic: 467.3 on 24 and 3975 DF,  P-value: < 2.2e-16
```

**Listing 4:** Multiple linear regression of the han decision of factory 3 vs the adjusted demand trajectory with $\theta$ equal to 0.2.

```
Call:
lm(formula = comple.sample.theta.02.26 ~ comple.sample.theta.02.2 +
comple.sample.theta.02.3 + comple.sample.theta.02.4 +
comple.sample.theta.02.9 + comple.sample.theta.02.13 +
comple.sample.theta.02.21, data =data_st_complete_sample)

Residuals:
     Min       1Q   Median       3Q      Max
-189.679  -29.500    7.048   35.766  102.688

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)              -5.937e+02  1.588e+01 -37.398  < 2e-16 ***
comple.sample.theta.02.2  5.536e-01  6.507e-03  85.073  < 2e-16 ***
comple.sample.theta.02.3  3.586e-01  5.838e-03  61.420  < 2e-16 ***
comple.sample.theta.02.4  8.732e-02  5.175e-03  16.874  < 2e-16 ***
comple.sample.theta.02.9  1.222e-02  4.362e-03   2.801  0.00512 **
comple.sample.theta.02.13 1.795e-02  5.696e-03   3.152  0.00163 **
comple.sample.theta.02.21 -2.693e-02 1.267e-02  -2.125  0.03365 *
---
Signif. codes:  0  ***  0.001  **  0.01  *  0.05  .  0.1

Residual standard error: 47.36 on 3993 degrees of freedom
Multiple R-squared:  0.7369,   Adjusted R-squared:  0.7365
F-statistic:  1864 on 6 and 3993 DF,  P-value: < 2.2e-16
```

**Listing 5: Multiple linear regression of the han decision of factory 3 vs the whole demand trajectory with $\theta$ equal to 0.05.**

```
Call:
lm(formula = comple.sample.theta.03.26 ~ comple.sample.theta.03.2 +
    comple.sample.theta.03.3 + comple.sample.theta.03.4 +
    comple.sample.theta.03.5 + comple.sample.theta.03.6 +
    comple.sample.theta.03.7 + comple.sample.theta.03.8 +
    comple.sample.theta.03.9 + comple.sample.theta.03.10 +
    comple.sample.theta.03.11 + comple.sample.theta.03.12 +
    comple.sample.theta.03.13 + comple.sample.theta.03.14 +
    comple.sample.theta.03.15 + comple.sample.theta.03.16 +
    comple.sample.theta.03.17 + comple.sample.theta.03.18 +
    comple.sample.theta.03.19 + comple.sample.theta.03.20 +
    comple.sample.theta.03.21 + comple.sample.theta.03.22 +
    comple.sample.theta.03.23 + comple.sample.theta.03.24 +
    comple.sample.theta.03.25, data = data_st_complete_sample)

Residuals:
    Min      1Q  Median      3Q     Max
-43.607  -3.044   1.389   5.110  11.451

Coefficients:
                            Estimate Std. Error t value Pr(>|t|)
(Intercept)                2.433e+02  2.111e+01  11.522  < 2e-16 ***
comple.sample.theta.03.2   1.481e-01  4.194e-03  35.300  < 2e-16 ***
comple.sample.theta.03.3   1.365e-01  3.750e-03  36.386  < 2e-16 ***
comple.sample.theta.03.4   5.262e-03  3.335e-03   1.578  0.11468
comple.sample.theta.03.5  -2.798e-05  3.088e-03  -0.009  0.99277
comple.sample.theta.03.6  -4.068e-03  2.911e-03  -1.397  0.16245
comple.sample.theta.03.7   3.465e-03  2.752e-03   1.259  0.20810
comple.sample.theta.03.8  -7.298e-04  2.813e-03  -0.259  0.79533
comple.sample.theta.03.9   7.320e-03  2.798e-03   2.616  0.00892 **
comple.sample.theta.03.10 -3.297e-03  2.896e-03  -1.139  0.25495
comple.sample.theta.03.11 -2.265e-03  3.084e-03  -0.734  0.46272
comple.sample.theta.03.12  7.503e-04  3.323e-03   0.226  0.82138
comple.sample.theta.03.13  9.102e-03  3.679e-03   2.474  0.01339 *
comple.sample.theta.03.14 -1.826e-03  4.191e-03  -0.436  0.66306
comple.sample.theta.03.15 -1.463e-04  4.826e-03  -0.030  0.97582
comple.sample.theta.03.16  1.603e-02  5.508e-03   2.911  0.00362 **
comple.sample.theta.03.17 -3.491e-03  6.437e-03  -0.542  0.58755
comple.sample.theta.03.18 -6.863e-03  7.455e-03  -0.921  0.35730
comple.sample.theta.03.19  1.211e-02  8.135e-03   1.488  0.13673
comple.sample.theta.03.20  1.509e-02  8.315e-03   1.815  0.06964 .
comple.sample.theta.03.21 -1.960e-02  8.064e-03  -2.431  0.01509 *
comple.sample.theta.03.22 -1.128e-02  7.350e-03  -1.534  0.12507
comple.sample.theta.03.23  8.870e-03  6.520e-03   1.360  0.17376
comple.sample.theta.03.24 -1.107e-03  5.511e-03  -0.201  0.84083
comple.sample.theta.03.25 -7.976e-03  4.752e-03  -1.678  0.09333 .
---
```

```
Signif. codes:  0   ***   0.001   **   0.01   *   0.05   .   0.1

Residual standard error: 7.594 on 3975 degrees of freedom
Multiple R-squared:  0.3928,    Adjusted R-squared:  0.3891
F-statistic: 107.1 on 24 and 3975 DF,  P-value: < 2.2e-16
```

**Listing 6:** Multiple linear regression of the han decision of factory 3 vs the adjusted demand trajectory with $\theta$ equal to 0.05.

```
Call:
lm(formula = comple.sample.theta.03.26 ~ comple.sample.theta.03.2 +
    comple.sample.theta.03.3 + comple.sample.theta.03.9 +
    comple.sample.theta.03.13 + comple.sample.theta.03.16
    + comple.sample.theta.03.21, data = data_st_complete_sample)

Residuals:
    Min      1Q  Median      3Q     Max
-43.191  -3.031   1.431   5.097  10.976

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)              239.164519  10.442334  22.903  < 2e-16 ***
comple.sample.theta.03.2   0.148280   0.004189  35.400  < 2e-16 ***
comple.sample.theta.03.3   0.136812   0.003742  36.558  < 2e-16 ***
comple.sample.theta.03.9   0.006650   0.002786   2.387  0.01704 *
comple.sample.theta.03.13  0.008668   0.003673   2.360  0.01833 *
comple.sample.theta.03.16  0.016095   0.005502   2.925  0.00346 **
comple.sample.theta.03.21 -0.019384   0.008053  -2.407  0.01612 *
---
Signif. codes:  0   ***   0.001   **   0.01   *   0.05   .   0.1

Residual standard error: 7.597 on 3993 degrees of freedom
Multiple R-squared:  0.3895,    Adjusted R-squared:  0.3886
F-statistic: 424.6 on 6 and 3993 DF,  P-value: < 2.2e-16
```