Delft University of Technology

FACULTY OF MECHANICAL, MARITIME AND MATERIALS ENGINEERING
Department of Marine and Transport Technology

Mekelweg 2
2628 CD Delft
the Netherlands
Phone +31 (0)15-2782889
Fax +31 (0)15-2781397
www.mtt.tudelft.nl

Specialization: Transport Engineering and Logistics

Report number: 2015.TEL.7931

Title: **Improving operations at a fulfilment operation at PostNL E-Commerce Services by improving the order volume forecasts and controlling planning parameters**

Author: I.C. Gelens

Title (in Dutch) Operationele verbetering van een fulfilment operatie bij PostNL E-Commerce Services door het verbeteren van de order volume voorspellingen en het aanpassen van planning parameters.

Assignment: Master's thesis

Confidential: yes (until March 23, 2025)

Initiator (university): prof.dr.ir. G. Lodewijks

Initiator (company): ir. J.X. Jolink (PostNL E-Commerce Services, Houten)

Supervisor: dr.ir. H.P.M. Veeke

Date: March 10, 2015

Delft University of Technology

FACULTY OF MECHANICAL, MARITIME AND MATERIALS ENGINEERING
Department of Marine and Transport Technology

Mekelweg 2
2628 CD Delft
the Netherlands
Phone +31 (0)15-2782889
Fax +31 (0)15-2781397
www.mtt.tudelft.nl

Student:	I.C. Gelens
Supervisor (TUD):	Dr. ir. H.P.M. Veeke
Supervisor (Company):	Ir. J.X. Jolink

Assignment type:	Master project
Creditpoints (EC):	35
Specialization:	TEL
Report number:	2015.TEL.7931
Confidential:	Yes
	until: March 23, 2025

Subject: **Improving operations at a fulfilment operation at PostNL E-Commerce Services by improving the order volume forecasts and controlling planning parameters**

Introduction

PostNL E-Commerce Services is a part of PostNL that operates in fulfilment services. This means that they provide warehouse and logistic planning and control for web shops of companies in the retail industry. Even though the post market is declining, package distribution is growing rapidly due to a continuous increase in online sales. PostNL ECS handles around 25 000 - 30 000 orders per week. Two thirds of these orders are generated by the two biggest customers. The rest is divided amongst 20 - 30 different web shops.

Problem definition

One of the major customers sells Senseo-like coffee machines, accessories and coffee refills to customers in all surrounding European countries. PostNL ECS handles order fulfilment for this entire market. With the growing volumes the question arises whether current processes are still adequate to handle the order load. Therefore, the assignment is to analyse the process to identify improvement opportunities to increase the capacity of this operation and increase the quality of service.

Research Goal

The goal of this research is to measure the influence of different planning parameters on the packaging operation in order to identify improvement opportunities. The planning parameters of which the influence will be measured are the quality of the forecast, the scheduling decisions resulting from these

Delft University of Technology

FACULTY OF MECHANICAL, MARITIME AND MATERIALS ENGINEERING
Department of Marine and Transport Technology

Mekelweg 2
2628 CD Delft
the Netherlands
Phone +31 (0)15-2782889
Fax +31 (0)15-2781397
www.mtt.tudelft.nl

forecasts and the shift times of the packaging personnel. Performance will be measured in late orders and cost.

Research structure

- Analyse the packaging processes according the Delft Systems Approach
- Collect and analyse the data necessary to identify improvement opportunities
- Identify different forecasting approaches
- Construct a simulation model in which the different forecasting approaches can be investigated
- Show the influence of forecasting, scheduling and shift times
- Develop an implementation strategy
- Study relevant literature

The Professor,

b.a.

Prof. dr. ir. G. Lodewijks

Summary

For a web shop, PostNL E-Commerce Services handles distribution of their Senseo-like coffee concept. Not only is the order volume substantial, but this is also expected to increase over the coming years. Therefore, the question from PostNL E-Commerce Services was how the activities deployed on this packaging operation can be expanded.

From the problem analysis it followed that capacity increase over the following years is not necessary, because in case of the maximal order growth possible, current capacity will only be insufficient several days a year. Furthermore, this capacity shortage could be easily solved by adding a nightshift. However, during this analysis a large misalignment was discovered between the orders being placed and the production capacity available to process these orders. The result is that in spite of contractual obligations of delivering 98% of orders on time, currently only 67% actually is handled on time. Reducing this misalignment is potentially beneficial to the operation of every web shop at PostNL E-Commerce Services. Therefore, this is the opportunity that has been investigated further.

In the final problem definition the research question has been formulated as follows:

> *Can the amount of late orders be reduced to 2% of total handled orders by changing the planning without an increase in cost?*

To find a solution to this problem the following sub questions needed to be answered:

> *What influence do different forecasting methods have?*
>
> *What influence does changing the schedule and shift times have?*

To answer these questions a simulation model was created where 28 days of operation were simulated starting at May 16th, 2014. This period of 28 days has been chosen, because during this period order volumes fluctuate, reaching almost peak capacity as well as a very low capacity. Even though the simulation model is fairly simple, the use of neural network forecasting and a large amount of combinations of different model parameters made discrete event simulation with Matlab® the simulation method of choice. During this simulation every day an amount of orders was processed. These orders were generated according to the actual incoming order distribution of those days taken from historic data. A baseline run was created by running the simulation with parameters as they have been estimated by discussing the operation with the Planner. The simulation using the web shop forecast as an input for planning, a shift starting time of 1 pm and a max duration of 9 hours per shift yielded a late order percentage of 33% and a cost of 4106 man-hours. This was used as a starting point for further improvement.

To map the influence of forecasting, three forecasting methods were tested over 5 months of historic data, namely the naïve method, the simple moving average and an artificial neural network which yielded a Mean Average Percentage Error (MAPE) of 50%, 36% and 33% respectively. Compared to

the web shop forecast with a MAPE of 67% this seemed a big improvement. However, when the simple moving average and artificial neural network forecasting models were tested as an input to the simulation model the late order percentage increased from the baseline percentage of 33% to 34% and 42% respectively. The cost decreased from 4106 to 4018 and 3917 man-hours respectively.

The influence of the schedule and shift times was researched by analysing the planning process. Every day, the forecast is translated in a schedule, which consists of an amount of packaging lines to plan. This translation was defined as a function of x and x was varied to measure the influence of the amount of lines. From this experiment it followed that structurally planning more packaging lines decreases late orders. However, no clear relation between x and cost has been found. Apart from adapting the schedule, the influence of two other planning parameters was researched, namely changing the starting time and changing the duration. For this test the starting time was advanced by an s amount of hours whenever the forecast exceeded 3 times the line capacity for one shift. In other words, whenever it was expected that the work would not be finished before 9 pm and therefore orders would be late, the starting time was advanced by s hours. The other factor that was varied is the shift duration d. This represented whether people are allowed to work overtime to clear the order load. Advancing the starting time of shift decreased late orders up to 13% of total orders handled. The maximal duration did not influence results as long as it is not shorter than 10 hours.

To promote implementation of the results, the influence of varying different parameters on the system was mapped consistently. By doing this, a set of desirable options for the different planning parameters was created which can be used by the Planner as an input for future planning decisions.

Returning to the sub questions of the problem definition, it is found that in general:

- Forecasting hardly has any influence
- As expected, on average, decreasing the amount of lines planned (a higher x) increases late orders. The influence of x on cost does not show this clear relation
- Extending the shifts hardly has any influence beyond 10 hours, which means that it should be possible to let workers stay for two extra hours when necessary
- Conditionally advancing the starting time greatly reduces the amount of late orders

Concluding the thesis, analysis of the full set of simulation outcomes showed that, the lowest late order percentage without an increase in cost is 13% and the lowest cost without an increased late order percentage is 3727 man-hours (9% decrease). This means that the answer to the main research question is no. It is not possible to reduce the amount of late orders to less than 2% without an increase in cost.

For this research a neural network model was tested, but not a lot of data was available to train this model. It is expected that the results of this forecasting method will improve as more data becomes available. Therefore, this could be researched further. For the company in general using available data better during the daily processes is recommended.

List of Figures

Contents

1 Introduction

E-Commerce is a volatile world where sales rise and fall with the day. Within this market, PostNL E-Commerce Services operates and tries to provide order fulfilment services against a competitive rate. In order to achieve this goal the company has to be able to flexibly handle whatever amount of orders comes its way fast and efficiently. To help with this problem a master's thesis is written on how to improve the operations on a single web shops operation. Within this research project an attempt will be made at gaining insight into the processes, order flows and possible future sales of this operation and to use this information to support daily business decisions.

The goal of this research is to adjust daily schedules and forecasting models in such a way that a better result is achieved in sense of cost and the amount of orders that is handled late on a day to day basis. The question that will be researched to reach this goal is: "*Can the amount of orders handled late be reduced to 2% of total orders handled by changing the planning without an increase in cost?"* A simulation will be made to map the effect of different planning decisions on the overall system. This will not only enable answering of the research question, but can also function as a roadmap to better daily planning for PostNL E-Commerce Services.

Chapter 2 will start with outlining the company. In this outline the market, daily operation and request from the company will be covered. In Chapter 3 the problem will be analysed in detail. The system boundary will be set and the processes analysed. The chapter will conclude with choosing a topic of research. The final problem definition and approach will be given in Chapter 4. A simulation model will be created in Chapter 5, after which Chapter 6 and 7 will deal with finding a good forecasting regime and selecting the proper process parameters to provide a solution. Chapter 8 will summarize the results of the different experiments and models. In Chapter 9 a possible implementation strategy will be proposed. The thesis will conclude with Chapter 10 containing conclusions and recommendations.

2 Company Outline

PostNL E-Commerce Services was founded as a fulfilment centre for print media under the name TopPak. Amongst others, they packed and distributed books for Readers Digest and the Disney Book Club. With the coming of the internet age, the demand for print media declined and TopPak had to look for new ways to attract business. With their experience as a fulfilment centre they found a place within the rapidly expanding e-commerce market providing warehouse management and fulfilment services for web shops of major retailers as well as small businesses. They were acquired by PostNL in 2010 and under the name PostNL E-Commerce Services they continue to serve a variety of national as well as internationally operating customers. Moving into e-commerce turned out to be a good decision and business for PostNL E-Commerce Services has been rapidly expanding ever since.

2.1 The E-Commerce Market

The e-commerce market has been growing rapidly over the last decade. New web shops are started continuously. Building a website for a certain market has become easier with the development of standard web shop modules. However, this is only the start of an e-retail business. On the back-end goods need to be fabricated, stored and distributed. The latter two are the field of an e-fulfilment business.

Building and maintaining your own warehouse can be a business involving considerable capital and risk. Fulfilment specialists cope with this by facilitating multiple web shops under a single roof. This is the playing field of PostNL E-Commerce Services or PostNL ECS. The services PostNL ECS offers include storage, picking, packaging and sending goods to consumers throughout Europe as well as handling the returns for a selected group of web shops. Since PostNL ECS already has storage and packaging facilities available, they can offer a fixed price per order to the customer. This way no initial investments are needed from the web shop and PostNL can make a margin over each fulfilled order.

Today PostNL ECS handles around 25 000 - 30 000 orders per week. Two thirds of these orders are generated by the two biggest customers. The rest is divided amongst 20 - 30 different web shops.

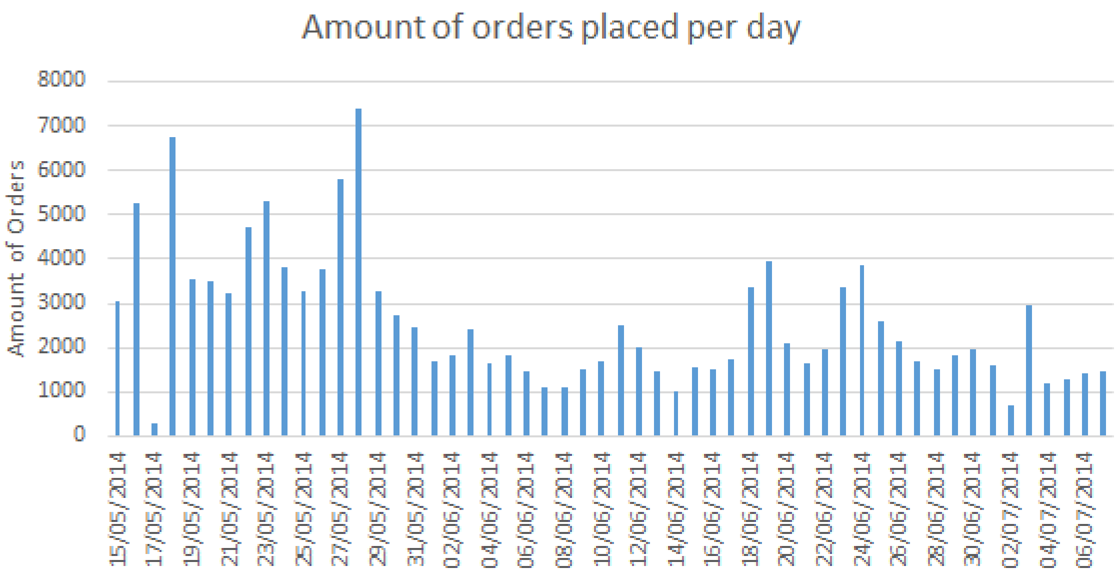


Figure 1: The amount of orders coming in per day for one of the web shops

The main challenge of this e-commerce model is that demand fluctuates greatly. Figure 1 shows the amount of orders placed per day for one of the web shops. As can be seen the amount of orders can double in a fortnight. Responding to this fluctuating demand of goods requires a very flexible production planning and creates challenges in stock keeping and production capacity optimization.

2.2 Daily Operation

Like a lot of businesses, the operations of PostNL E-Commerce Services can be modelled by the process performance (PROPER) model (Veeke, Ottjes, & Lodewijks, 2008). This model is based on the philosophy that a production process is in essence a transformation step with three in/outflows, namely orders, products and resources. Coordination control, in this case management, sets standards for this transformation process based on measured results within the process. This is the first control loop. The second control loop is with the external environment. The environment sets requirements for the process and evaluates the performance of the production system. These requirements mainly consist of requirements set by customers, but also include for example certification standards and regulations set by the government.

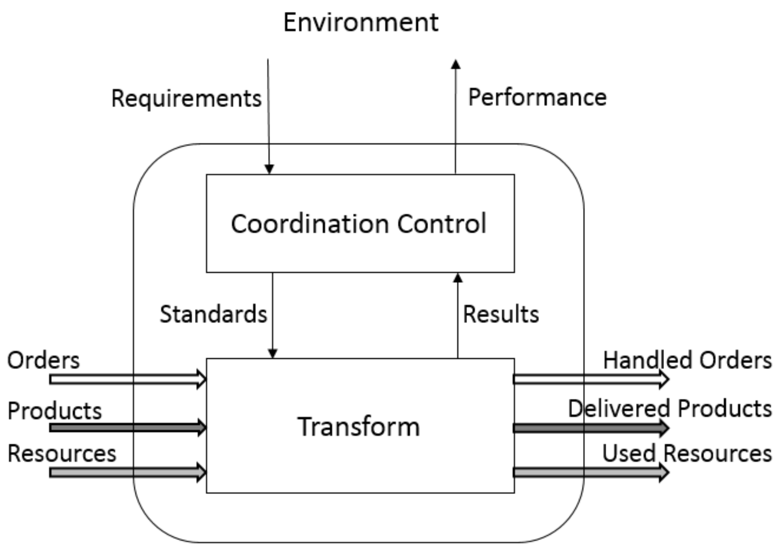


Figure 2: The Process Performance (PROPER) model (Veeke, Ottjes, & Lodewijks, 2008)

Since PostNL E-Commerce Services plays a facilitating role for web shops instead of dealing with consumers directly, the requirements set to the process consist of contractual obligations towards the different web shops. The performance is measured by Key Performance Indicators (KPIs) on which PostNL is reporting weekly or monthly. Based on these KPIs, PostNL ECS sets standards for their operation.

The inflows consists of orders placed by consumers which are encoded through the web shops interface. Based on these orders, products are taken from the warehouse, packaged and expedited. Hardly any automation exists in this system. All moves of pallets are done by employees with fork-lifts. Products are picked by hand from different parts of the warehouse and packaging is done by hand. Therefore, used resources mainly consist of man-hours of the packaging employees. Using this strategy, PostNL ECS has been able to secure several large players in the e-commerce industry and has grown rapidly in a short period.

However, rapid growth combined with a change in business model has come at a price. When PostNL ECS moved from print media to e-commerce a big effort was made to rapidly attract new customers. In order to facilitate these web shops a lot of concessions were made on the side of PostNL E-Commerce Systems. This resulted in different approaches in the handling of orders, different packaging requirements and different IT-systems per web shop. Furthermore, there are web shops that are not selling the projected amount of goods. Pallet places have a lower margin than fulfilled orders so slow moving web shops fill up the building against a low return. Right now there is no room in the building for new web shops. Concluding, the low level of standardization combined with slow moving web shops in an overfull building has caused financial results of this company to be not as good as they could be.

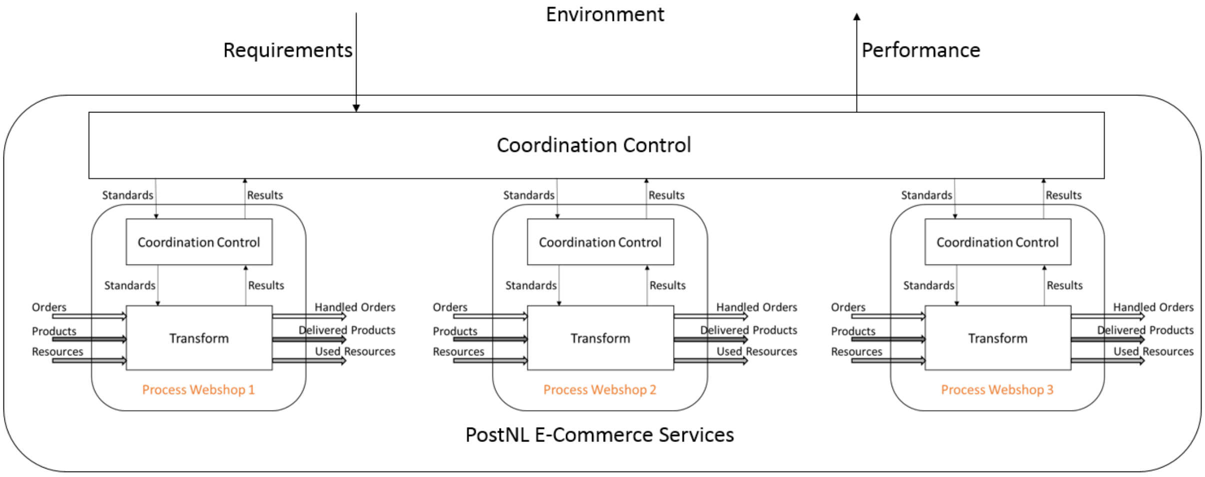


Figure 3: The operations at PostNL E-Commerce Services zoomed in one layer

On the operations side, this means that the amount of standardization is low. Some clusters of customers exist, but each of these clusters is handled separately with their own packaging line. When this is modelled within the PROPER-model framework it looks like Figure 3.

During the past year an effort was made to further align the different business processes, but up until this day different IT systems are in operation and each customer is handled as a different process with its designated place in the building. The IT-systems are quite rigid causing a lot of employees to cut corners in order to make their lives easier, especially under stressful conditions. To quote an employee: "There are systems, but on a busy day they go overboard and everybody is just focussing on getting the orders out in time". Misalignment of the different systems and an overfull building increase the amount of ad hoc decisions that are made on the work floor and the amount of firefighting on a day-to-day basis is relatively high.

However, a lot of work is being done to improve the results at PostNL E-Commerce Services. A project to replace the different IT-solutions by one overall system has been launched. Projects also exist to improve the internal reporting and new web shop implementations and Lean Black Belts are starting improvement projects for the different lines as well. The graduation assignment as laid out by the company consists of looking at one of the production lines to see if improvements can be made there. This will be outlined in the next section.

2.3 Company Question

One of the major customers sells Senseo-like coffee machines, accessories and coffee refills to customers in all surrounding European countries. PostNL ECS handles order fulfilment for this entire market. Even though this is one of the best organised lines of PostNL ECS the orders are still being picked by hand. With the growing volumes the question arises whether this is still the way to go. In other words, a growth strategy for this process is needed.

The coffee-machine is not heavily branded in the Netherlands but has been growing steadily ever since its launch. In the first half of this year, 581858 orders have been realised. This equals 65% of total sales in 2013. Because the peak season is in December, it is expected that in 2014 1.2 million orders will be sold. This represents a 34% increase compared to 2013. With these growing volumes the question arises how to handle the future order volumes for this process. Therefore, the question posed by PostNL E-Commerce Services can be defined as follows:

How can production of this operation be increased to cope with an increase in order volume?

3 Problem Analysis

This chapter contains the problem analysis from which the final research question will be formulated. Section one will contain a definition of the system boundary. Section two will contain a detailed process analysis. Section three will offer a conclusion introducing the research questions posed in the next chapter.

3.1 System Boundary

In order to properly analyse the system the system boundary has to be defined. As mentioned in the previous chapter, each web shop has its own process within the hall. Therefore this report shall be limited to the process of the coffee web shop. Since PostNL ECS does not control the incoming orders as well as the logistical process after the parcel has been loaded onto the truck, the system boundary shall be placed only around the picking and packaging operation. As soon as the parcels are in the truck, the product is delivered and the order handled.

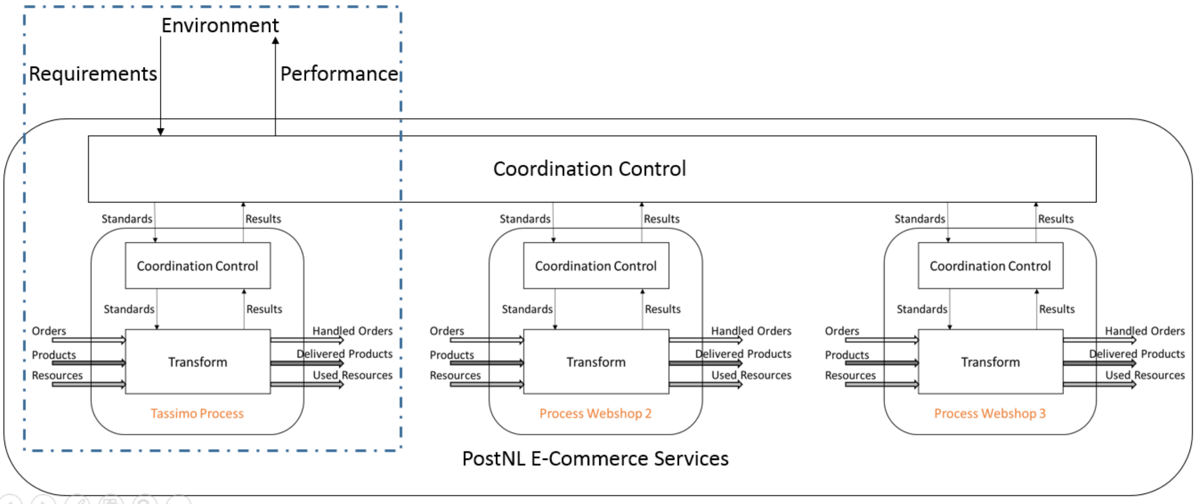


Figure 4: This research will focus on the coffee web shop process

3.2 Process Analysis

In this section this system will be analysed in detail. Firstly, the requirements and performance indicators of this process have to be identified. Secondly, the effect of these requirements on the standards of the production process will be covered. Lastly, the different parts of the detailed process will be analysed.

3.2.1 Requirements and Performance

Every process within PostNL ECS has requirements and performance standards introduced by the Dutch government as well as certain certification agencies. However, in this report, only the requirements posed by the contractual agreements with the web shop shall be covered, because these are the requirements that are relevant for this research.

Several contractual agreements were made with the web shop on the performance of order handling. These agreements are:

- (Production Planning) The web shop will provide PostNL ECS with weekly forecasts
- (Accuracy) At least 98.5% of orders will be shipped containing the correct content with the correct packing material
- (Production Capacity) At least 98% of all orders will be processed before cut-off times based on jointly agreed max capacity per day
- (Production Capacity) Extra capacity will be available the next day to catch up if the amount of orders exceeds the forecast by more than 10%
- (Production Capacity) Extra capacity will be available within two days if the amount of orders exceeds the forecast by more than 20% and order volume is above max capacity

Noticeable about these contractual agreements is they contain a lot of duties for PostNL ECS and not a lot of rights. The only obligation that the web shop has is to provide a forecast. However, no agreements have been made on the required accuracy of this forecast. The rest of the agreements are all performance requirements for PostNL ECS.

The performance on these requirements is evaluated in meetings between the web shop and PostNL. Up until now most other performance criteria are met and the contractual obligations have been fulfilled. Based on the complaints register of the web shop it is clear that less than 1.5% of orders have incorrect contents. As for the production capacity constraints, extra capacity has always been available whenever asked for. However, within the company it is not clear whether the 98% of all orders where processed in time. This is an observation which shall be returned to further in this analysis.

3.2.2 Coordination Control (Standards and Results)

Using these performance requirements, standards are created for the production process. Within the process this happens on two levels. The top level is where the sales department and Operations Manager set the global standards for the process. This level consists of decisions concerning expansion, training personnel and expanding the decision set of operational management. The lower level is where the daily operational decisions are made. Here the shifts are planned and production capacity is increased or decreased. In summary, one could say that in the lower coordination control block operational decisions are made, while in the upper coordination control block the focus lies on tactical and strategic decisions.

3.2.3 Transformation

With the information gathered on the work floor, a detailed process diagram of the transformation process of the system can be made. This process diagram has been depicted in Figure 5. Note that this figure only contains one coordination control block. In this analysis the focus lies mainly on evaluating daily operation. In this section the different parts of the process will be covered separately.

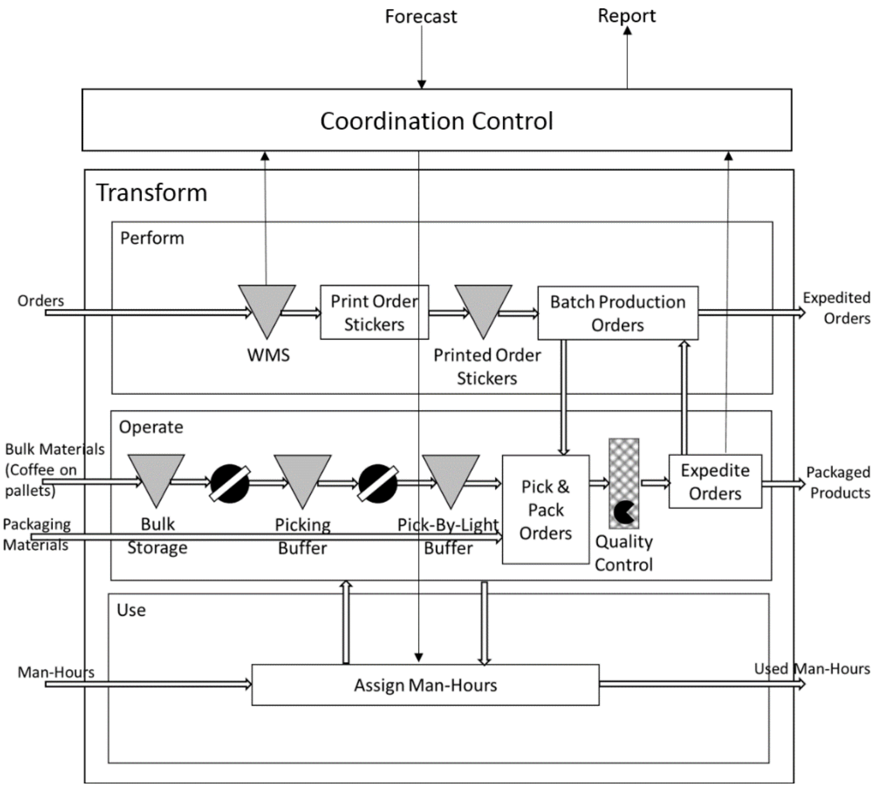


Figure 5: A map of the transformation process

Orders

Firstly, let us take a closer look at the incoming orders. The orders enter the system through a connection with the web shop. Orders do not come in at the same rate over the year. When looking at the distribution of orders over the year as plotted in Figure 6, one notices that the amount of orders coming in at December and January is considerably higher than throughout the rest of the year.

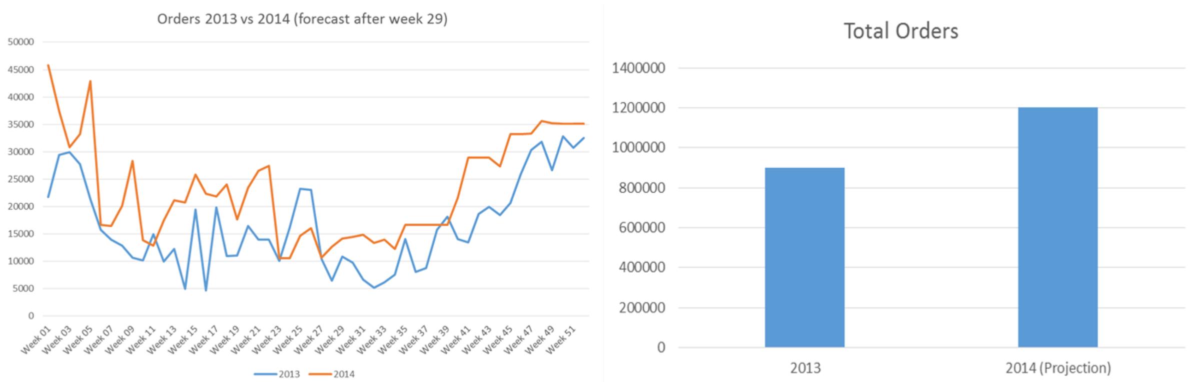


Figure 6: The distribution of the projected order growth

As mentioned in the previous chapter, the annual volume of this line has been 898 415 orders in 2013. This amount is expected to increase to approximately 1.2 million orders by the end of 2014. Based on this projection, let us assume that order volumes increase by 300 000 every year. This means that in 2014 and 2015 the amount of orders will be 1.2 and 1.5 million respectively. Assuming that working weeks consist of 5 days this means that the amount of orders per day in 2015 is on average 5769 orders. This has been depicted in Table 1.

YEAR	ANNUAL VOLUME	AVERAGE DAILY VOLUME
2013	898 415	3455
2014 (PROJECTION)	1 200 000	4615
2015 (ASSUMPTION)	1 500 000	5769

Table 1: Projected sales volume development

PostNL ECS receives a fixed payment per handled order. Therefore, the more efficient they are able to handle orders, the more money is made. The most important indicator is the productivity indicator which is measured in handled orders per man-hour. PostNL ECS uses a productivity norm of 22.5 orders per man-hour. When the process is at full capacity it employs 23 people. PostNL ECS runs a maximum of two 8 hour shifts per day.

A closer look to the daily distribution of orders is depicted in Figure 7. What can be learned from this graph is that the amount of orders placed on a day to day basis is pretty unevenly distributed. When taking in mind that the maximum capacity based on two shifts is 8280 orders, it is found that the amount of orders on a day does not exceed the maximum capacity at any point. However, these averages give a skewed view of the current situation, because as discussed in the previous chapter, the amount of orders varies greatly per day and can double in a fortnight. Furthermore, the amount of orders placed in December and January is about 35 % higher than in the rest of year. This means that in these busy months the amount of orders would exceed the maximum capacity one day of the week.

Figure 7: The amount of orders coming in per day for the web shop between May 15th and July 6th of 2014

Is this a problem? An argument can be made that it is not. If the maximum capacity is only exceeded a couple of days per year there is no need to expand the line. Extra capacity would be added for a few busy days in December and January which would be useless throughout the rest of the year. Also, this extra capacity could be added by adding an extra shift on busy days. Therefore, expanding the line does not seem necessary.

Buffers

When looking at production systems, the stock levels are usually a good indicator on how efficient a process is. In Figure 8 there are 3 buffers leading up to the packaging process. These buffers are stages of unpacking the bulk materials.

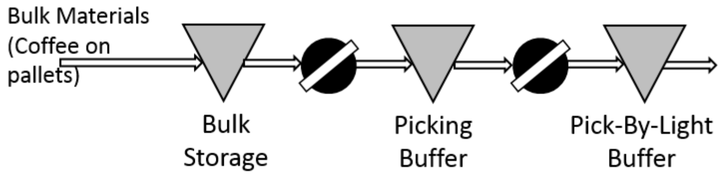


Figure 8: The different buffers in the process

When stock arrives at PostNL ECS it is stored in bulk on the pallets they came on. Whenever a type of coffee is needed the pallet is taken from the bulk storage and stored next to the packaging lines so they are accessible for the line fillers. When the small packages are running out at the line, the line fillers unpack the pallets into the Pick-By-Light system (the Pick-By-Light shall be covered more in-depth below).

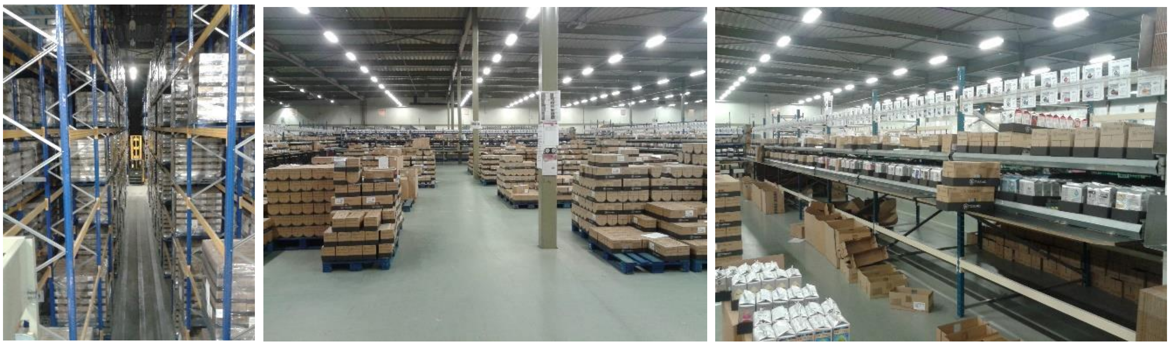


Figure 9: The different buffers of the process

There are quite a few different places where stock is kept in this system, so it might be a good idea to find a way to reduce this. However, PostNL ECS is not in charge of procurement of goods. This means that all the products are bought by the web shop who pays for all the pallet places that are used. Therefore, this is a part that will not be looked into further.

Packaging

The packaging process is the centre of the operation. The packaging operation will be described in detail in this section. A floor map of the process has been depicted in Figure 10.

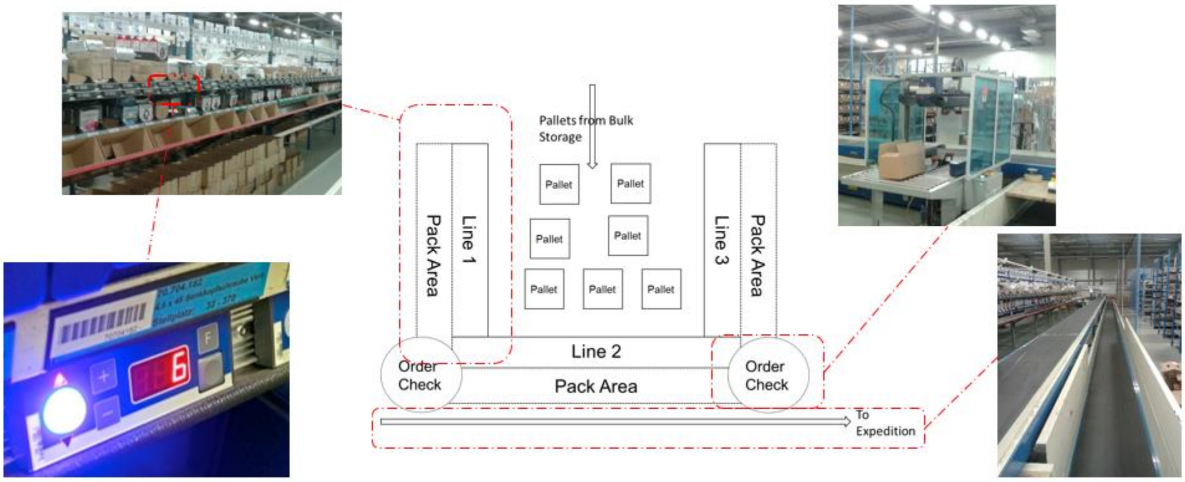


Figure 10: A map of the process containing the different elements (image bottom left courtesy of fastec.de)

When consumer orders arrive in the system they are printed out as stickers at the planning station of the lines. These stickers contain coded information. This includes what products need to be in the pack, consumer address and a bar-code for internal use as well as Track-and-Trace. These stickers are batched based on the destination country of the order because different countries have different cut-off times. Batches of these stickers on roles are brought to the line. At the beginning of a line, a sticker is attached to a box. The different boxes move from station to station where the lights of the Pick-By-Light system indicate which product needs to go into the box. Please note that no automation is involved here. The conveyor system is only used to transport scrap packaging material to the bin. The shipping boxes are pushed to the next station by hand by the employees at the line. After the boxes are filled with products they are verified and sealed at the quality control station. All boxes are checked whether they contain the right quantity and every fifth box is checked completely for content. When the boxes are checked and sealed they are placed on a conveyor system that transports the boxes to expedition.

AMOUNT OF LINES	EMPLOYEES ON LINE	EMPLOYEES SUPPORT	TOTAL	CAPACITY PER SHIFT
1	5	5	10	1400
2	10	7	17	2800
3	14	9	23	4200

Table 2: Capacity per shift for the different amount of lines

The process is organised in three packaging lines. Dependent on demand either 1, 2, or 3 lines are utilised. The lay-out is depicted in Figure 10. As can be seen the temporary storage is in the middle of the three different lines. Based on how many lines are used an amount of supporting staff is required. The amount of staff needed in the different modes of operation is listed in Table 2. The packaging personnel is provided by an in-house employment agency. This employment agency plans the people on the different shifts. The general planning is handed to this employment agency one week in advance, but shifts can be cancelled up to three hours before the shift starts. A shift normally starts at 1 pm.

After this time there are either one or two 8 hour shifts based on requirement. This means that the first shift will work from 1 to 9 pm and the second shift will work from 9 pm to 5 am. Somewhere between 9 and 10 pm the last truck to the PostNL distribution centre leaves. All orders processed after this time will not make it to the distribution centre until the next day and can be considered as late orders. This means that in the normal situation the windows in which orders will be handled is between 1 and 9 pm.

As established in the previous section the amount of orders placed per day varies greatly. To give PostNL ECS a sense of what to expect the web shop sends a weekly forecast with the amount of orders to be expected each day. Based on this forecast a daily workforce planning is made. The problem with this, is that the forecast is really inaccurate. Figure 11 shows a highly fluctuating deviation of the estimated amount against the actual amount. These errors can be as big as 200% of total orders.

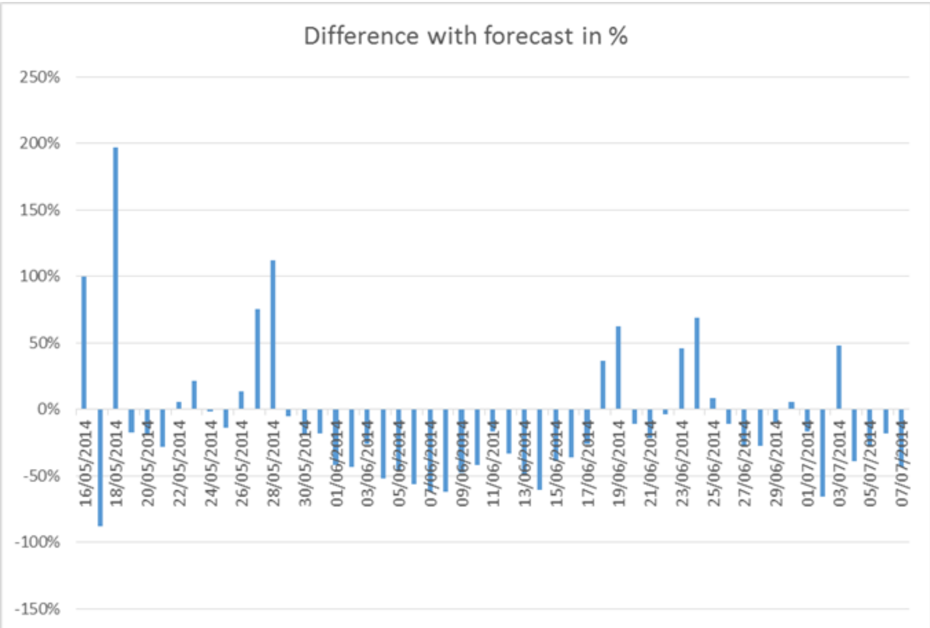


Figure 11: The error of the forecast provided by the web shop

Up until now, a forecast provided by the web shop is used to make planning decisions. The web shop forecasts weekly orders for the coming year and uses this weekly forecast to produce a day forecast which it shares with production planning. This day forecast is the basis for the weekly schedule. What happens is that according to some distribution a part of the week forecast is expected every day. For example, on Monday 15% of the week forecast is expected, on Tuesday 13% and so on. In theory this forecast is also used to make adjustments during the week, however, going through the email correspondence it is found that the web shop exerts a great influence on the packaging process capacity. Basically, the web shop tells PostNL ECS how much capacity to plan and does not provide updated information. This, together with the fact that the day forecasts used for decision making are actually derived from the week forecasts, could be an explanation for the big forecasting error.

The production planning at PostNL ECS is off. Figure 12 shows the graph of orders placed, but this time it included the amount of orders handled per day. Apart from the high production on Monday, which is to be expected after a free weekend, there are large productivity peaks further into the week as well.

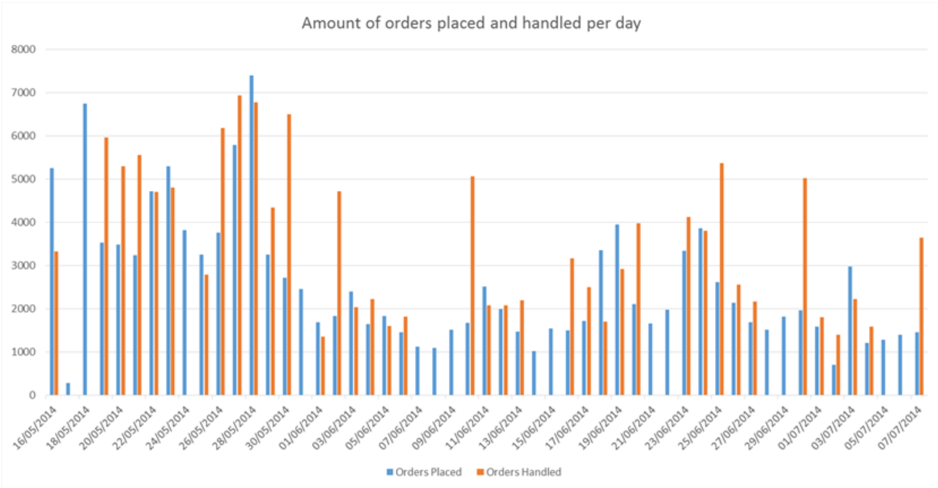


Figure 12: Orders placed versus orders handled

Part of the contractual agreements with every customer are the so called "cut-off times". Cut-off times can be defined as the latest time at which an order can enter the system and still be processed on the same day. For example, whenever a web shop states on their website that orders placed before 8 pm will be delivered the next day, this means that the cut-off time at PostNL ECS is set at 8 pm. This means that the amount of orders that needs to be processed every day is the amount that is still open from yesterday plus the amount of orders that comes in during the day before cut-off times. The web shop has multiple cut-off times for the different countries which are displayed in Table 4.

COUNTRY	CUT-OFF TIME
PORTUGAL, SPAIN	12:00
GREECE	15:00
FRANCE, AUSTRIA, GERMANY, UK, NORWAY, SWEDEN, DANMARK	17:00

Table 3: Cut-off times for the different countries

In the database of the order system the "receive time" of an order is registered. If the receive time is before cut-off time, the order needs to be processed that same day. By analysing the "send times" an overview can be made of whether orders are processed on time.

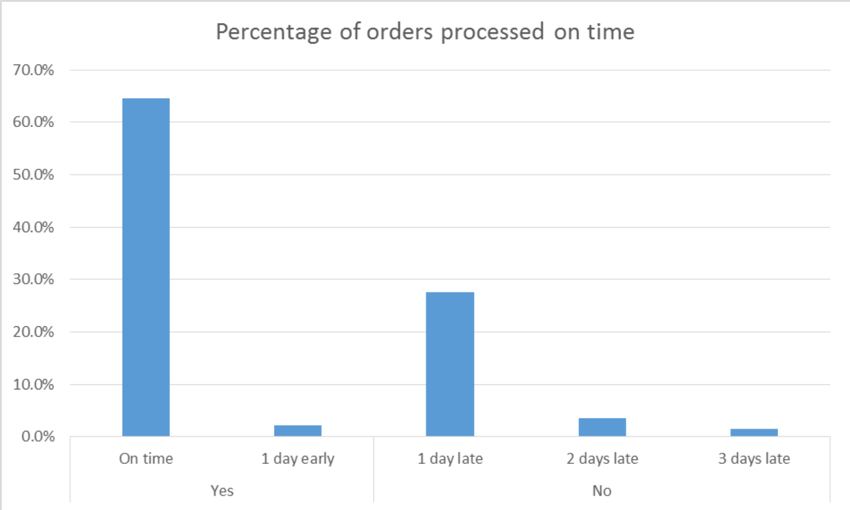


Figure 13: Data analysis shows that 33% of orders is processed too late

When analysing the data on the send times of the orders placed at the web shop, it becomes clear that 33% of orders does not make it to the truck in time as depicted in Figure 13. Going through the email correspondence providing the production forecasts it turns out that in practice the web shop exerts a lot of influence on the production process. They keep a close tab on how many orders are in the system and demand extra man power when the amount of backlog grows too large. A logical response to 33% of orders being processed too late

Summarizing, the discrepancy between forecasted demand and actual demand of production capacity causes problems for all parties involved. For the web shops and consumers the service level is low because 33% of orders are late. For PostNL ECS the discrepancy means that they either have too little capacity causing orders being too late or that they have excess capacity causing the operation to increase in cost. Both effects are undesirable.

3.3 Conclusion

Returning to the original question as posed by PostNL ECS, it is found that increasing the capacity is not necessary, because the maximum capacity is only exceeded several times per year. This difference can easily be compensated by employing an extra shift of people. If capacity would be increased this excess capacity would be useless in the months February to November.

Having established that capacity expansion is not necessary the question arises where this perceived capacity shortage comes from. This is a result of the planning difficulties arising from a highly fluctuating demand and an inaccurate forecast. Therefore, capacity planning could be improved in two ways, namely by increasing the forecasting accuracy and/or by improving the decision policy of daily operations management. This could not only lead to more efficient use of resources, but could also decrease the amount of late orders. Therefore, improvement in this field would not only benefit PostNL, but also the web shops and customers.

4 Final Problem Definition

As outlined in the problem analysis the largest gain can be expected by improving the capacity planning. Currently, the forecasts received by web shops are inaccurate. Increasing this accuracy will benefit PostNL as well as all customers. For PostNL ECS it will be possible to plan resources more accurately. Because supply and demand can be better aligned, the amount of late orders as well as cost are expected to decrease. This is not only a benefit for this web shop, but is a potential benefit for every web shop that is using the services of PostNL ECS now and in the future.

4.1 Problem Statement

Based on the problem analysis in the previous question, the assignment has been formulated as follows:

> *Can the amount of late orders be reduced to 2% of total handled orders by changing the planning without an increase in cost?*

In order to complete this assignment the following questions need to be answered:

> *What influence do different forecasting methods have?*
>
> *What influence does changing schedule and shift times have?*

The next chapters will answer these questions and propose a solution for the outlined problem.

4.2 Approach

To provide a solution for the problem the following approach will be used:

1) Simulation Model. A simulation model will be created to test the effect of different parameters on the system.
2) Forecast. This will be a search for a forecasting model that is able to predict the incoming flow of orders more accurately and can be used by someone without any knowledge on forecasts to make predictions. This has been researched extensively. The impact of different forecasting models will be tested using the simulation model.
3) Planning. The effect of different scheduling regimes will be tested using the simulation model as well as the effect of changing around the shift times.

These steps will be followed by the results, an implementation strategy, conclusions and recommendations.

5 Simulation Model

In this chapter the composition of the simulation model will be further outlined. The first section will motivate the choice for simulation as a modelling tool. Section 2 will outline the goal of the model. Section 3 and 4 will deal with the simulation plan and modelling parameters respectively. This will be followed by Section 5 containing a general simulation description. The paper model will be presented in Section 6. Section 7 will contain the model verification runs. This all will be concluded by generating a baseline in Section 8.

5.1 Why Simulation?

There are several reasons why a simulation is preferred over other possible calculation methods, namely:

- The impact of planning decisions on late orders could span multiple days. For example, when not all orders for a certain day are handled on that day they are moved to the next day thus impacting the order load on that day. These effects are hard to measure without a simulation.
- The impact of planning decisions on cost are not linear. Contractual obligations with the employment agency require that staff works a minimum amount of hours. Also, when too little lines are planned for a day it is not possible to increase the capacity anymore.
- A simulation allows for testing over a longer period of time. This makes long term effects visible.
- It is an efficient way of testing the impact of multiple decisions when a lot of factors are interrelated.
- As a forecasting method, neural network forecasting is proposed further in this report. Matlab® contains a neural network toolbox which makes programming a neural network a lot less tedious. Combining these forecasts with the SimEvents toolbox provides an integrated framework for all experiments in this report.

5.2 Goal of the Model

As introduced in the last chapter the goal of this model is to test different forecasting models as well as planning policies on a simulation that resembles the actual situation. In this way it will be possible to come up with an optimal planning policy and to select a forecasting method that will improve the performance of the system.

5.3 Simulation Plan

To reach this goal a simulation plan must be made. The fundament of the analysis in Chapter 3 was 53 days of detailed order information downloaded from the order database. This detailed information consists of the exact times orders entered the system over the period between May 16th and July 7th 2014 and will be used as an input for the simulation. To reduce the simulation time 4 weeks or 28 days of this data will be used to run a full discrete event simulation. The first 28 days of this dataset are taken because they contain all ranges of order volumes of the set. With this simulation it is possible to accurately describe the distribution of the orders going in and to test the impact of different responses to this order flow on system performance. Work times and order processing capacity will closely resemble real life. Every day a forecast and production planning will be made one day in advance. By monitoring the effects of these decisions on the systems performance an optimal strategy can be selected.

5.4 Modelling Parameters

5.4.1 Performance Measurement

To measure the performance of the model two results will be observed. Since late orders was identified as a big problem in the problem analysis, this is obviously a parameter that will be measured and optimized. Using logic reasoning it is possible to deduce that the amount of late orders will approach zero as the capacity goes to infinite. Unfortunately, as capacity goes to infinite, so will cost. Therefore, the second parameter that will be measured is the operating cost, which will be measured in total man-hours deployed. The aim is to modify the current system in a way that eliminates late orders at a decreased cost.

5.4.2 Assumptions

To construct a model of this systems several assumptions about the system need to be made. These assumptions are outlined below.

- The schedule is made the day before the operation. Even though it is suggested that in real life schedules are made on a weekly basis, the schedule may be changed up to three hours before a shift starts. Therefore, the schedule will be based on a forecast of tomorrow's order volumes together with what is left from today. In this way it will use the maximum amount of information available.
- From the analysis in Section 3.2 it followed that 33% of orders are processed late, of which 6% was more than one day late. For the simulation it will be assumed that late orders are always handled first on the next day thus eliminating orders that are late by more than one day.
- Even though manual labour means that productivity might fluctuate from day to day, for this simulation a fixed capacity of 175 orders per hour per line will be used.
- At whatever day in of the year the simulation starts, there will be no orders left in the system from previous days. A clean start will be assumed.
- All hours worked are equally expensive. The difference that might exist in overtime or second shifts will not be taken into account.

5.4.3 Constraints

From the actual work environment there are some constraints which should also be incorporated

- Orders handled after 9 pm will be considered late orders, since the last truck to the distribution centre will have left.
- Whenever an amount of people is planned for a shift, they will work for at least 5 hours due to contractual obligations towards the employment agency.
- If the amount of orders coming in is above expectations, it is possible to make a shift work up to 1 hour of overtime.
- The maximum amount of shifts planned for a day is two eight hour shifts.

5.4.4 Free Parameters

The next step is to identify the parameters which will be changed to influence the system in sense of operating cost and late orders. During these simulation sets there are four system parameters which will be varied. First a new way to forecast the amount of incoming orders will be researched. This will be covered in detail in Chapter 6. The variable x which regulates the amount of lines planned, the starting time s which regulates the amount of hours to start early when a busy day is expected and d which regulates the shift duration will also be researched. This will be further outlined in Chapter 7.

5.4.5 Input Data

As input data to the model the actual order data will be used. When looking at the average of the cumulative orders coming in during the day as depicted in Figure 14 , one notices that between 9 am and 12 pm the inflow of orders follows a highly linear pattern.

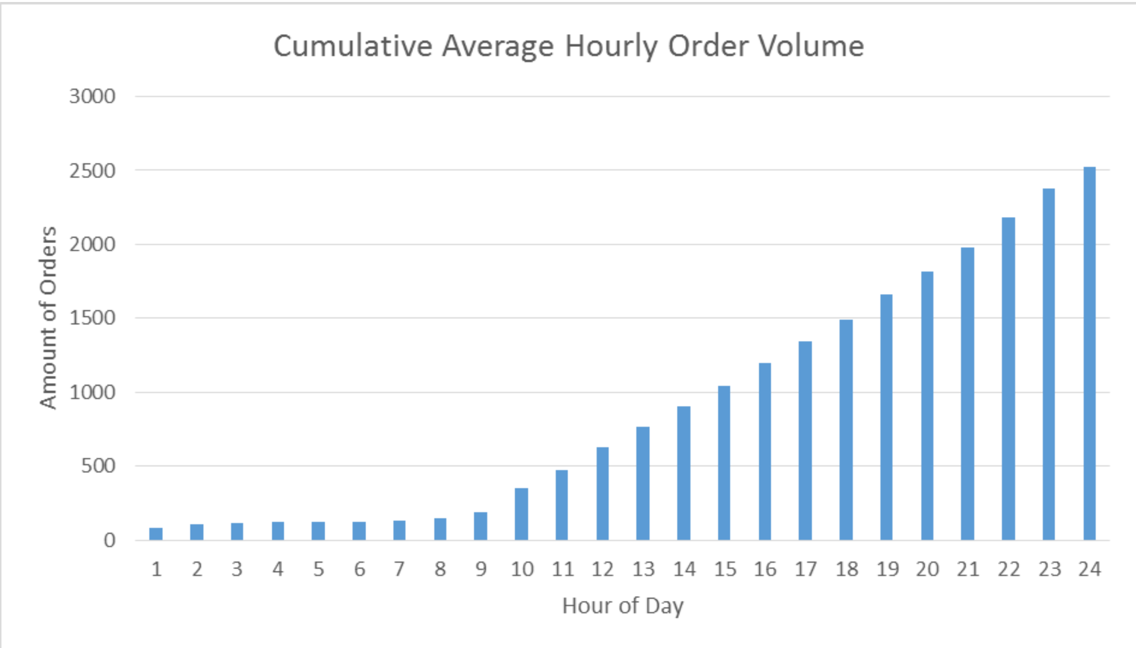


Figure 14: Cumulative Average Hourly Order Volume

On average about 92 % of all orders is placed between 9 am and 12 pm and the behaviour during this time is highly linear. However, because not all orders coming in have the same cut-off times, it is still desirable to use the actual data and not for example a constant inter-arrival time.

5.5 General Simulation Description

To understand how this simulation was performed the process model of Chapter 3 (Figure 15) needs to be taken as a starting point.

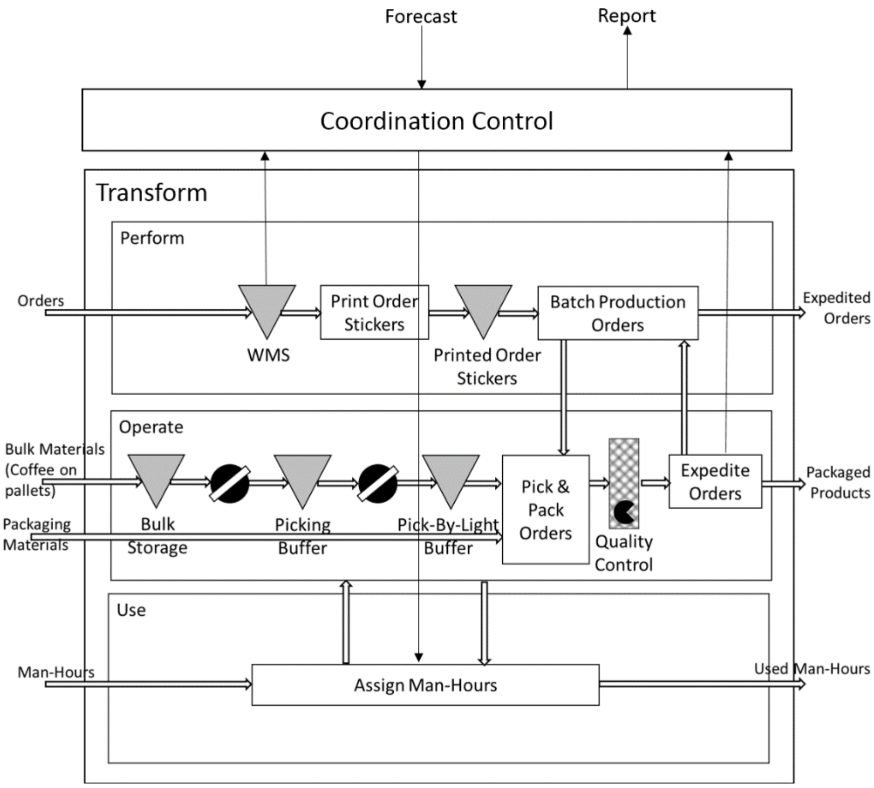


Figure 15: A map of the transformation process

The part of this model that is relevant here is coordination control. The goal of this research is to try different forecasting methods and planning policies to uncover a better way to plan the capacity of this system. The starting point for this simulation is the model depicted in Figure 16.

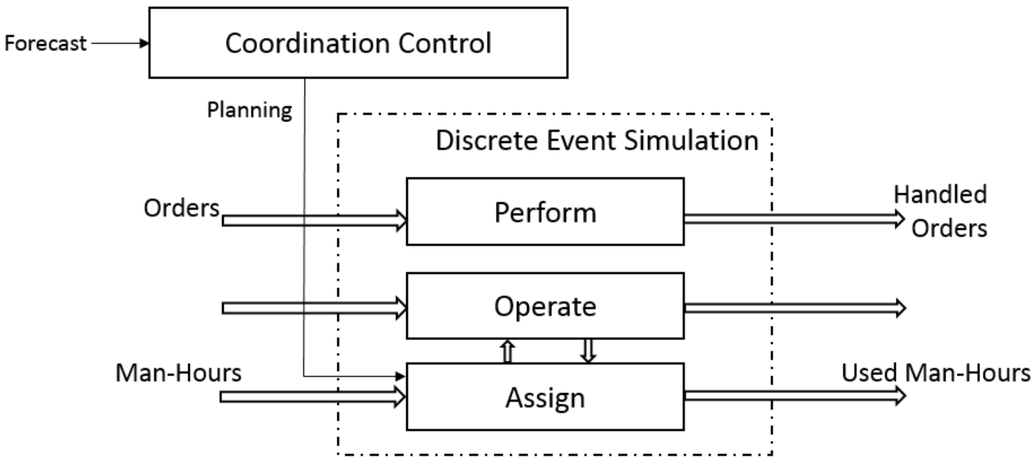


Figure 16: The simulation model

Figure 16 shows the current situation. A forecast is delivered by an external party. This forecast is used by coordination control to generate a planning for the assign block of the system. Within the simulation this assignment of man-hours happens in the form of selecting a certain amount of lines and shifts. These assignments results in a certain system performance which is measured by monitoring the handled orders (late order percentage) and used man-hours (cost). A schematic view of the discrete event simulation is presented in Figure 17.

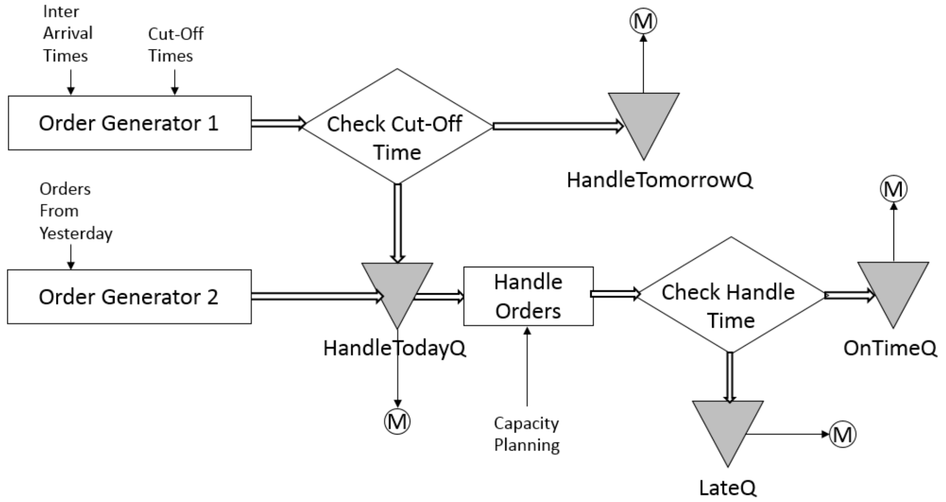


Figure 17: A schematic overview of the discrete event model

The discrete events model is fairly simple. However, the goal is not to optimize the size of the system, but to optimize the coordination controls influence and response to the system. This can be done by changing the forecast, or changing the planning resulting from the forecast. Figure 18 shows this in more detail. As explained in Section 5.4.4 the forecast will be varied by exploring new ways to forecast orders based on the results of the system. Results that are relevant in this case consist of historic data and other factors that might influence the forecast. The planning will be regulated by the scheduling parameter x, starting time s and duration d. These parameters will be explained in detail in Chapter 7.

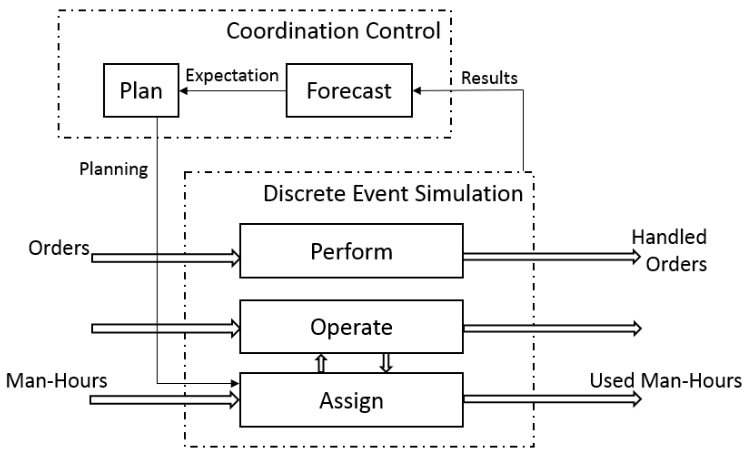


Figure 18: Varying what happens in the coordination control block is key in this simulation

In summary, a rough description of the models workings is as follows:

- Set forecasting method
- Set amount of days to simulate
- Repeat for every day in the simulation:
    - Make forecast
    - Add forecast to orders left from yesterday to form expectation
    - Based on the forecast set the planning parameters
    - Run a day of simulation using the planning parameters (run the model in Figure 17)
- At the end add all results together to determine late order percentage and cost in man-hours

5.6 Paper Model in Program Description Language

```
//Definition Section
Order: (SimElement)
MyCutOffTime                        //Cut Off Time of this particular order

OrderGenerator: (SimElement)
Process: Method

OrderGenerator2: (SimElement)
Process: Method

CutOffTimeChecker: (SimElement)
CheckerQ:Set                        //Check Orders Q
HandleTomorrowQ: Set                //Orders to handle tomorrow
Process: Method

PackagingOperation: (SimElement)
Capacity                            //Amount of time
Start_of_workday                    //Time that the workday starts
End_of_workday                      //Time the workday ends
Process: Method

LateOrderChecker: (SimElement)
HandledOrderQ: Set                  //OrdersHandled
OnTimeQ: Set                        //Orders handled on time go here
LateQ: Set                          //Orders handled late go here
Process: Method

//Process Section
OrderGenerator1.process
While True
        Create NewOrder
        Orders_in = Orders_in + 1
        CutOffTime = from Distribution
        Add NewOrder to CutOffTimeChekcer.CheckerQ
Wait(InterArrivalTime)

OrderGenerator2.process
For i = 1 to the amount of orders left from yesterday
        Create NewOrder
        Add NewOrder to HandleTodayQ
```

CutOffTimeChecker.process

```
Create CheckerQ
Wait until there is an order in the CheckerQ
        MyOrder = CheckerQ.FirstElement
        If TNow<CutOffTime
                Move MyOrder to HandleTodayQ
        Else
                Move MyOrder to HandleTomorrowQ
```

PackagingOperation.process

```
Standby while TNow < Start_of_workday
Packaging_start_time = TNow
While TNow < End_of_workday AND HandleTodayQ is not empty
        Move First Order from HandleTodayQ to LateOrderChecker.HandledOrderQ
        Wait(packaging_speed)
Packaging_end_time = TNow
```

LateOrderChecker.process

```
Create HandledOrderQ
Create OnTimeQ
Create LateQ
Standby until Start_of_workday
MyOrder = HandledOrderQ.FirstElement

If TNow<last_truck_leaves
        Move MyOrder from HandledOrderQ to OnTimeQ
else
        Move MyOrder from HandledOrderQ to LateQ
```

//Model Control Section

```
MAIN
//Set global variables
Tend = 1440                               //24*60 minutes
linecapacity_hour = 175                   //orders per hour
linecapacity = linecapacity_hour/60       //orders per minute
last_truck_leaves = 21                    //time the truck leaves in hours
forecast_mode = 1                         //1 for baseline, 2 for sma, 3 for neural network
end_of_forecasting_period = 28            //The amount of days to forecast

//Set starting variables
Orders_in = 1
Orders_left = 0

//Set global Queues
Create HandleTodayQ
Create HandleTomorrowQ


//Read Data
InterArrivalTimes = read inter arrival times from Excel file
CutOffTImes = read cut-off times from Excel file
forecast = read the web shop forecast from Excel file
actual_data = read actual daily order volumes from Excel file

For i = 1 to end_of_forecasting_period
```

```
//Make Order Expectation
Make forecast                    // Can be either The web shop, Moving Average or NN
Order_expectation = forecast + orders_left

        //Set Parameters
Based on Forecast select start_of_workday, end_of_workday, amount_of_lines and
amount_of_people

        //Run a day of simulation
        StartSimulation
        Advance(Tend)
        Finish

        //Collect results
        Late_order_amount = LateOrderChecker.LateQ.length + HandleTodayQ.length
        Total_late_orders = Total_late_orders + Late_order_amount
        orders_left =  HandleTomorrowQ.length + HandleTodayQ.length
        running_time = Packaging_end_time - Packaging_start_time

If running_time < 5            //Minimum amount of time planned is 5 hours for 1 shift
                Cost = 5*amount_of_people
        Elseif running_time < 9 or running_time > 8+5
                Cost = running_time*amount_of_people
        Elseif running_time < 8+5       // Minimum amount of working time is 8+5 hours for 2 shifts
                Cost  = 8+5*amount_of_people

Total_cost = total_cost + Cost
```

5.7 Model Verification

5.7.1 Verification Run of the SimEvents model

To verify that the model is working correctly, a verification run is executed with the following input parameters:

Orders left from yesterday	500
Orders coming in	1 per minute
Cut-off time	900 (3 pm)
Last Truck Leaves (Orders Late Time)	960 (4 pm)
Starting Time	780 (1 pm)
Ending Time	1260 (21 pm)

Looking at the output of the two order generators, it is found that indeed 500 orders are generated at t=0. These orders represent the orders that were left over from yesterday. The other order generator generates 1 order per hour as planned.

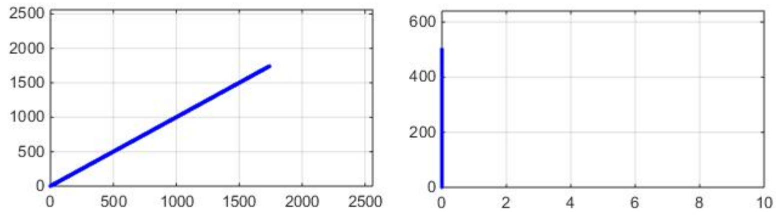


Figure 19: Left to right, Order Generator 1 (Orders Coming In), Order Generator 2 (Orders From Yesterday)

Looking further at the HandleTodayQ it can be seen that 500 orders are inserted at t=0. After this the HandleTodayQ fills with 1 order per minute. At starting time 1 pm or simulation time 780 the Handle Orders process is taking orders out of the Q, but orders still come in. At Cut-off time 900 the amount of orders in the queue starts to decline faster because at this point the orders generated by Generator1 are put into the HandleTomorrowQ.

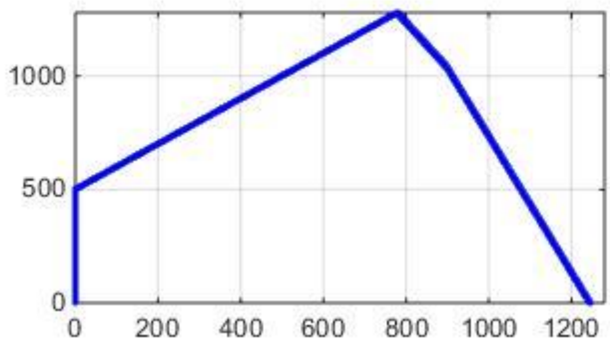


Figure 20: HandleTodayQ

Verifying the level of the HandleTomorrowQ it is found that indeed from t=900 onwards, orders start flowing into this queue.

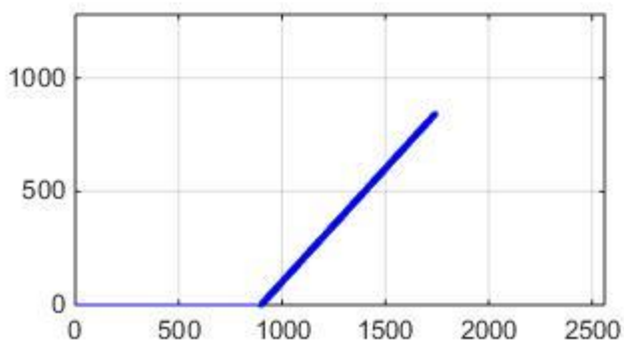


Figure 21: HandleTomorrowQ

Finally, because the Orders Late Time is set early at 960, it is found that from 960 onwards, orders are flowing into the LateQ instead of the OnTimeQ after being handled.

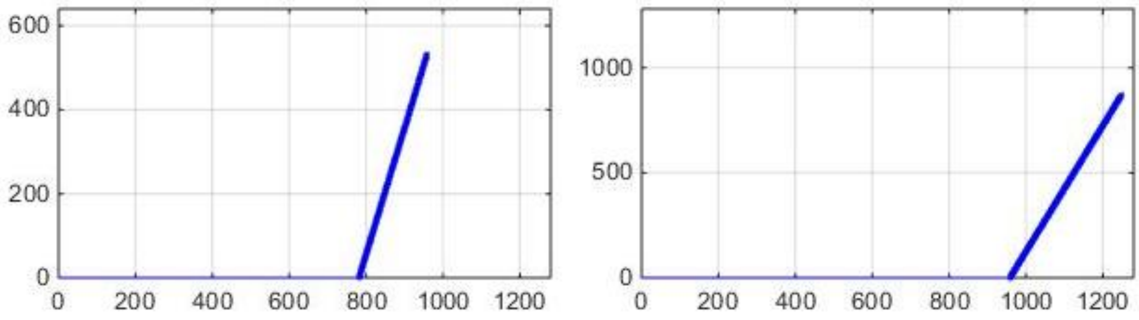


Figure 22: Left to right, OnTimeQ, LateQ

5.7.2 Further checks and balances

The model runs a simulation for every day. After every day it collects the amount of orders that are still in the system and uses this as input for the next day. To make sure this process works properly it must be verified that these orders are transferred to the next day properly. That the results do add up can be verified in Table 4: "Orders received" equals orders received before and after cut-off timeTable 4, Table 5 and Table 6.

	DAY 1	DAY 2	DAY 3	DAY 4
ORDERS RECEIVED	5257	277	6754	3535
ORDERS RECEIVED BEFORE CUT-OFF TIME	3244	277	4897	1540
ORDERS RECEIVED AFTER CUT-OFF TIME	2013	0	1857	1995

Table 4: "Orders received" equals orders received before and after cut-off time

	DAY 1	DAY 2	DAY 3	DAY 4
ORDERS TO HANDLE TODAY	3245	2290	7187	10584
ORDERS RECEIVED BEFORE CUT-OFF TIME	3244	277	4897	1540
ORDERS LEFT FROM YESTERDAY	1	2013	2290	9044

Table 5: "Orders to handle today" equals "orders received before cut-off time" combined with "orders left from yesterday"

	DAY 1	DAY 2	DAY 3	DAY 4
ORDERS TO HANDLE TODAY	3245	2290	7187	10584
ORDERS HANDLED	3245	0	0	8640
ORDERS LEFT	2013	2290	9044	3939

Table 6: During weekends "orders to handle today" equals "orders left"

One final note is that the model generates 1-2 orders more than it should. This follows from the generator blocks used in this simulation. The order generator that generates incoming orders always generates one order too much due to that it is based on inter arrival times. This means that it always needs an order at t=0 for reference. The order generator that generates orders left from the previous day does not accept a zero input. Therefore, whenever zero orders are left from the last day, this block will still generate an order. This means that at most two orders too many are generated. On an average of 2558 orders generated per day this is a negligible difference.

5.8 Baseline Run

In order to generate a reference point a baseline run is performed. By interviewing the Planner about daily operation the following baseline parameters are established:

- The forecast of the web shop is used as the sole basis for planning
- The Planner plans the lines a little conservatively due to pressure to minimize costs. Therefore, the factor x that establishes the schedule is set at 9. This factor will be discussed in more detail in Chapter 7.
- Work always starts at 1 pm.
- The maximal duration of a shift is 9 hours thus allowing for 1 hour of overtime.

These baseline parameters yield the following results when running the simulation:

	BASELINE
ORDERS RECEIVED	85752
TOTAL MAN-HOURS	4106
LATE ORDERS (PERCENTAGE OF TOTAL)	33%

Table 7: The baseline from which the system will be improved.

This is the baseline performance. These results shall be improved upon.

6 Forecasting

In this chapter the effect of the forecast will be examined. Where PostNL ECS depended on a web shop forecast in the past, this chapter is an overview of the possibilities of using the data collected by the company itself as a starting point to create an expectation of how many orders will be coming in and therefore, indirectly, what capacity to plan. Section 1 will cover the goals of this simulation. In Section 2 the dataset will be further examined. Section 3 will focus on forecasting model selection. After that some experiments and analyses are conducted in Section 4. Finally, Section 5 will present the results obtained by running the simulation model with different forecasting methods.

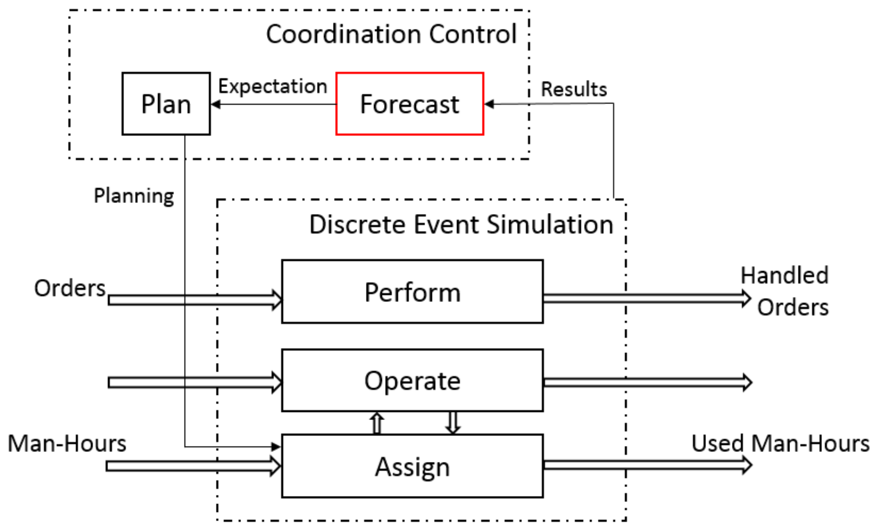


Figure 23: This chapter will focus on improving the forecast

6.1 Goals

The goal of this chapter is to produce a forecast that is better than the web shop forecast and measure the impact of this forecast on production. A proper forecasting model in this situation will consist of two parts.

The first part is the historic part. For this part access to all the historic data of PostNL ECS is required. The goal is to extract as much information from the historic data as possible. This information shall serve as a basis for the model. In online sales the same rules apply as to most things. Results made in the past do not guarantee results in the future. However, to deny the correlation of past results to the future completely is also incorrect. When sales have been steady at 10 000 for the past week, it is unlikely that they will suddenly be 100 000 tomorrow. The trick is to find whatever law is available in the historic data and use this law to forecast future sales.

The second part consists of using as much information about the future as can be gathered. This part consists of studying the factors of which it is known that they affect consumer behaviour in a way that impacts sales. For example, free delivery offers, discounts and package deals will all boost sales. The key of this part is to find out how much these different offers affect sales. To reduce the forecasting error in this part, at least a heads up from the web shop is needed when they have any promotions coming up. Further data that could be incorporated could be market trends and strategic developments at the web shop or PostNL itself.

Important to note is that the best forecasting model will never be the model that fits the historic data best, but the model that minimizes the error between the forecast and the actual value. Obviously, some form of function approximation will be performed, but the model is only as good as its forecasting error is small. The experiment design will support this important point.

6.2 Examining the Dataset

A series of observations of a development of some variable over time is knows as a time-series (Hillier & Lieberman, 2001). Extracting patterns from these observations is known as time-series analysis. This specific branch of forecasting problems is not focussed at finding relations between different variables of which it is suspected that they are dependent, but looks mainly at the development of a certain variable over time. The goal is to extract a model that resembles the development of the actual variable. The business use of time-series analysis is mainly focussed on price and sales forecasting, but the methods are also widely used in econometric models to forecast stock prices. Even though mainly historic data is used, obviously factors that influence prices can also be taken into account. This will be the main starting point for the forecasting model.

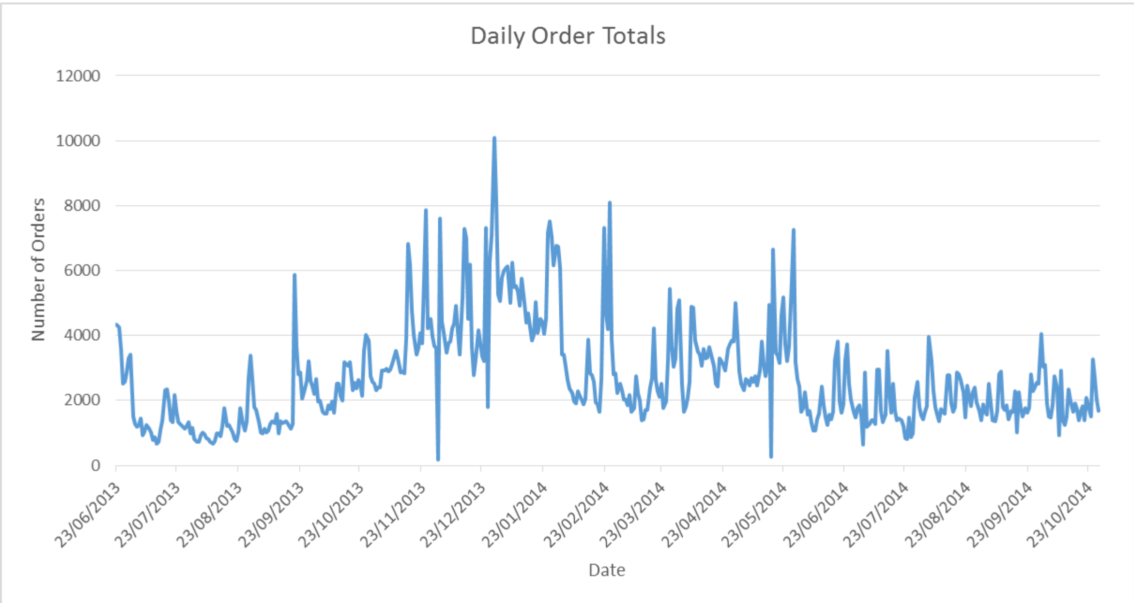


Figure 24: Daily order volumes for the product

To gain an understanding on what has to be done to construct a proper forecasting model first the dataset shall be analysed. The dataset that shall be used in this part of the research is the dataset which consists of a daily account of order volumes as received by PostNL ECS. Note that the amount of orders received and not the amount of orders handled is used. This is because the handled orders depends on the production capacity available on that particular day. The amount of orders received is the sales volume in its purest form.

The development of sales over time is the perfect example of a univariate time-series. Franses, van Dijk, & Opschoor (2014) wrote an excellent book on time-series analysis in which they describe the different parts that can be extracted from a dataset. Let us begin with the three fundamental parts of datasets. The first part is the trend. This is definitely not the most complicated part. In short, the trend indicates wheter datapoints in general tend to move up or down. The second part is seasonality. A lot of datasets are dependent on seasonal effects. Every recurring pattern in data can be seen as seasonality. The third part is the random part. Real-life data is never perfect. There will always be noise on the data. This is typically modelled by a white noise function or left out deliberately. The three basic parts are depicted in Figure 25.

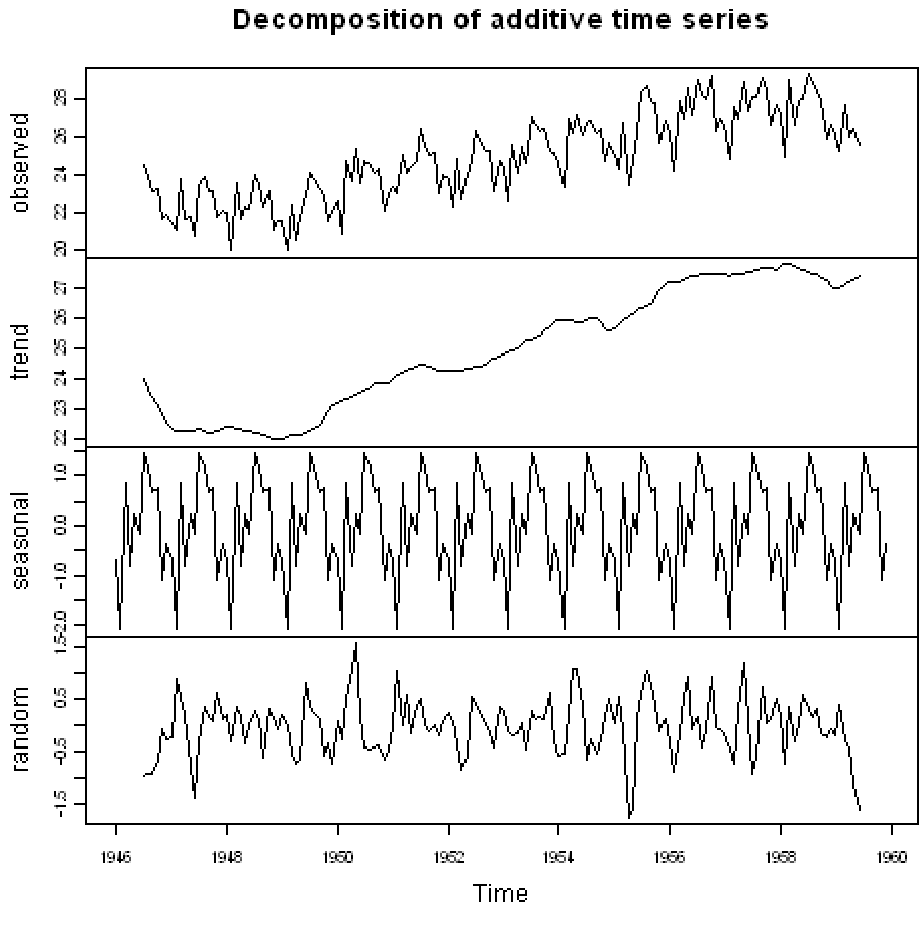


Figure 25: The decomposition of time-series data (Coghlan, 2014)

Two other important concepts are heteroskedasticity and non-linearity. Heteroskedasticity means that data can become more volatile over time. In order words, the variance around the mean will have different values for different time periods. This can be important for the validity of some models. Obviously, most time series that contain any of the parts above can be considered non-linear in the traditional sense of the word. Within time-series however, the definition is slightly different. Here, nonlinear behaviour is when the first differences $y_t - y_{t-1}$ take different average values across different parts of the dataset (Franses, van Dijk, & Opschoor, 2014).

Using this information, it is possible to look at the dataset in Figure 24 with new knowledge. Within the dataset, it is hard to spot any trends. This is because there is no more than 1.5 years of data available. However, based on the observations made in Section 3.2 and Figure 6 an upward trend in sales volumes is expected. The same goes for the seasonality component. Based on the data it is hard to identify seasonality components for the same reason it is hard to identify trends. From the retail business and data of other clients, it is known that the data contains a strong seasonality component as well. Apart from these fundamental parts, what can be seen is that this is a noisy dataset with strong heteroskedasticity and non-linear components. All this should be taken into account when selecting a model in the next sections.

6.3 Model Selection

For this report an extensive literature review of forecasting methods was conducted. The results of this review have been enclosed in 0. This section will further explore their use as a forecasting method for PostNL E-Commerce Services.

The first class of methods are the simple methods. Since these models hardly use any data and therefore are easily tested as well as implemented in the future organisation, these models will be tested in an experiment to see if they can outperform the current status quo. The models that shall be tested are the (seasonal) naïve method as well as the simple moving average.

The second class of methods are classical time-series models. Hillier & Lieberman (2001) state in their book that: "Although the Box-Jenkins (ARIMA) method is complicated, the resulting forecasts are extremely accurate and, when the time horizon is short, better than most other forecasting methods." Overall it seems that most time-series models can produce more accurate forecasts than the simple methods when applied correctly. However, when further investigating this possibility several problems came up.

One of the problems is the amount of data available. To properly identify and incorporate seasonal effects at least a couple of years of data is needed. When going through the database it turned out that the amount of data available was little over a year. This is too little to properly quantify seasonal effects even though it is clear that they exist. Another problem is that when a classical time-series model would be fitted to the data available, due to the little amount of data available, there is no way to tell whether this model would still be valid next year or the year after. The model should be refitted when more data is available. Since the chance is small that at PostNL ECS forecasting expertise will be available at that time, this is not a good idea. The risk that some non-performing tool will be in use even though it is not valid anymore is too great.

The third class of methods are the Artificial Neural Networks. ANN has several advantages over time-series models. As long as the amount of hidden layers, neurons and delays is correct, the Network can be trained to a dataset without any expertise at all. All that is needed is to input the proper data and the network will train itself based on all inputs and generate a forecast. Since Matlab® has a neural network tool available which can be built into a forecasting tool using the royalty free distribution package, this tool could be used and will keep adapting itself to the presented data. In other words, the tool will keep using all the data that is available to produce the forecast. This seems like a promising option and will be investigated further.

The final class of models are the combination models. In literature, the approach of combining classical time-series models with Artificial Neural Networks has been researched. This approach is dropped as well due to the same reasons as the classical time-series models.

In summary, models that could be useful for this particular application are:

- Naïve Method
- Moving Average
- Artificial Neural Network

These models will be evaluated in the experiment covered in the next section.

6.4 Experiment

To see how the forecasts relate to the ones done at the web shop a baseline needs to be created out of the forecasts that are received on a day to day basis. Collecting data on this is a tedious task because even though the web shop sends emails almost daily, none of the information is recorded systematically. Therefore, data for the baseline had to be extracted by going through the emails of a sales rep one by one.

By doing this a sample of about two months was created from which both the forecasts and the actual values where known. For this sample the indicators were calculated. The results for the baseline are depicted in Table 8. An explanation of the performance metrics used as well as a motivation for using them is offered in 0.

DAILY DATA	MAE	MAPE
WEB SHOP FORECAST	1085	67%

Table 8: Accuracy of the daily web shop forecasts

It does not need any further explanation that when the forecast used as a base for planning deviates by 67% on average, an efficient production planning becomes virtually impossible. These numbers serve as a baseline for the forecasts that are made further in the report.

6.4.1 Fitting the data

In this section the experiment used to test the different models will be explained. Because the Artificial Neural Network needs to be trained to a dataset the dataset is divided in two parts. The first part, which shall be called the training part from now on, will be used for training the model. The second part will be used for testing the forecasts and shall be called the testing part. Training the model is not necessary for the simple models (naïve and moving average), however, in order to create a fair comparison, the result of these models are still only calculated over the testing part. Both the training part and the testing part of the experiment will be discussed separately.

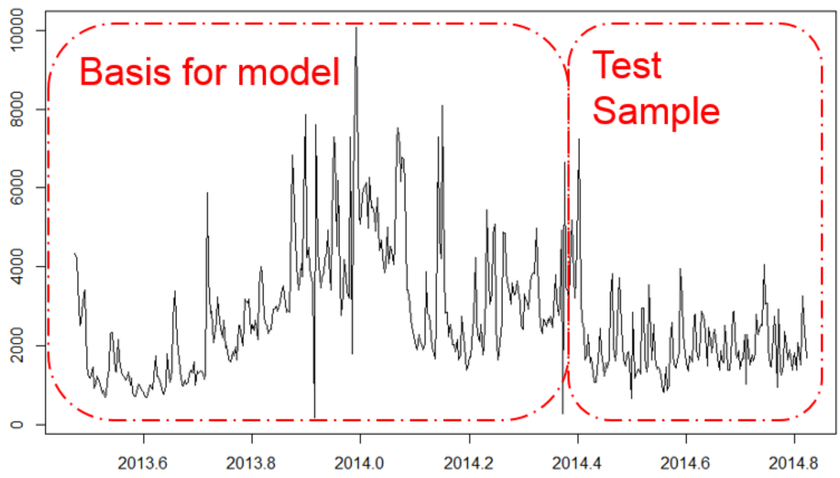


Figure 26: The dataset divided into two parts

The Artificial Neural Network needs to be trained based on the data. By feeding it with the known in and outputs, the model can be trained and will learn relations between the input and output data automatically. The training stage has been depicted in Figure 27. The inputs consist of coded days to find seasonal effects when they occur grouped in the vector u_t and the last known value of the amount of realised orders in scalar y_t. Note that u_t should also have contained some distribution to estimate promotional actions. However, promotional actions are not recorded systematically within PostNL ECS. Therefore, it was not possible to include this data in the research.

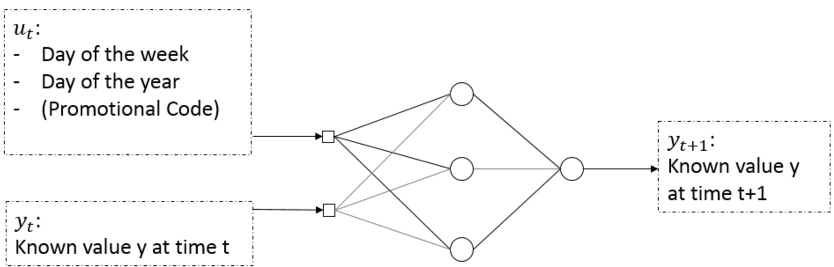


Figure 27: Training the ANN with known in- and outputs

Using the known value for y_{t+1} the network weights are adapted as explained in Appendix D.4. The optimization algorithm which is used in this particular case is the Levenberg-Marquardt algorithm which is the standard used by the Matlab® Neural Network toolbox.

Even though the weight optimization is fully automated, the neural network architecture still needs to be selected. This consists of selecting the appropriate amount of neurons, layers and delays for the network to actually start generating proper results.

As explained in 0 the weights of the network are adapted according to the error between the known inputs and outputs. Adding hidden layers, neurons and delays to a neural network increases the capacity to approximate complex functions. However, adding too much detail to the network will cause it to perfectly fit the historic data, but not make accurate predictions since it starts finding patterns that are missing a deterministic underlying mechanism. A problem that is known as over fitting. Therefore, the goal is not to find the network architecture that fits the data best, but the network that makes the best predictions.

Using the time-series neural network tool in Matlab® the dataset can be loaded and will be divided into three parts: the training part, the validation part and the testing part. The training part is used to train the network, the validation part is used by the optimization algorithm to know when the weights are optimal and the testing part is used to test the predictions of the model. By monitoring the error on these different parts it is possible to select a right amount of layers, neurons and delays. When starting with the minimal amount of two neurons, one hidden layer and one delay the training stage optimizes the weights. All errors will be large since the network is really crude. By adding neurons, layers or delays the errors made by the network will be reduced. When the errors on the training and validation set are still improving, but the error on the test set is growing larger, the network is becoming over fitted. By adding layers, neurons and delays until the neural network is over trained the optimal architecture of the network for this application is found. This turned out to be a neural network of one hidden layer, five neurons and one delay.

When the network is trained it can be used to make forecasts which will be explained in the next section.

6.4.2 Testing the performance

After the different models are fitted to the data it is time to test how these different models perform on the test sample. Every day a forecast is made. For the test, the first prediction is made based on the first data point of the test part of the dataset. From this point the model is used to forecast. These predictions are recorded as a results table.

Simultaneously, a table of actual values for the forecasts is constructed. Note that this information is not used in the forecast. Obviously, this information cannot be used since it will not be available when forecasts are made in real life situations. By subtracting the actual value table from the results table an error table is constructed. From this error table, the MAE and MAPE can be calculated. The Matlab® code for this operation as well as the training phase has been enclosed in Appendix C.

	Web Shop Forecast	*Seasonal Naïve*	*Moving Average*	*ANN*
MAE	1085	1046	634	575
MAPE	67%	50%	36%	33%

Table 9: The Mean Absolute Error (MAE) and Mean Average Percentage Error (MAPE) for the different models

By generating the error matrix as explained, a table of the results can be created. The Mean Average Error and Mean Average Percentage Error are shown in Table 9. When looking at the Mean Absolute Error of the models it can be noticed that all models outperform the original web shop forecast. Apart from that the moving average and the Artificial Neural Network perform in a similar way.

6.4.3 Forecasting in Detail

To see whether a more accurate forecast could be made using more detailed information, the daily order flow is analysed further. In Section 5.4.5, the linear relation of the average hourly order volumes was introduced. Of course it is possible that this pattern is an effect of taking the average of all data, but further insight into the data shows that a similar linear pattern also emerges on separate days for example on May 19th, 2014 as depicted in Figure 28.

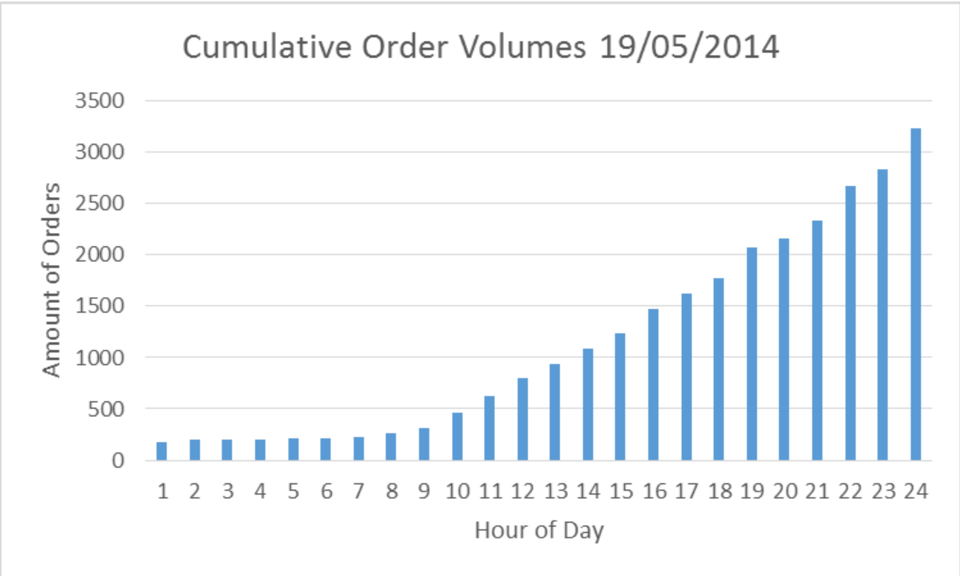


Figure 28: Cumulative hourly order volume on 19/05/2014

To see if making use of this linear relation could be used in the decision making process at PostNL ECS, a model is constructed. To assess the feasibility of this approach the following simple linear model is used to forecast the amount of orders that came in before 5 pm.

$$Day\ forecast = orders\ at\ t + \ orders\ per\ hour * hours\ left\ until\ midnight$$

$$Day\ forecast\ = Ord_t + \frac{Ord_t - Ord_{9:00}}{t-9} * (17-t)$$

This model was applied to a week of data from the detailed dataset. The goal of the analysis was to see at what time forecasts could be useful for decision making. The errors as a percentage of the actual value are depicted in Figure 29. In this graph it can be seen that for all days monitored, at 1 pm the amount of orders coming in until 5 pm can be predicted with an error of less than 10%. However, this is too late to be useful for capacity planning as well as changing shifts around. Therefore, this is a road that will not be pursued further.

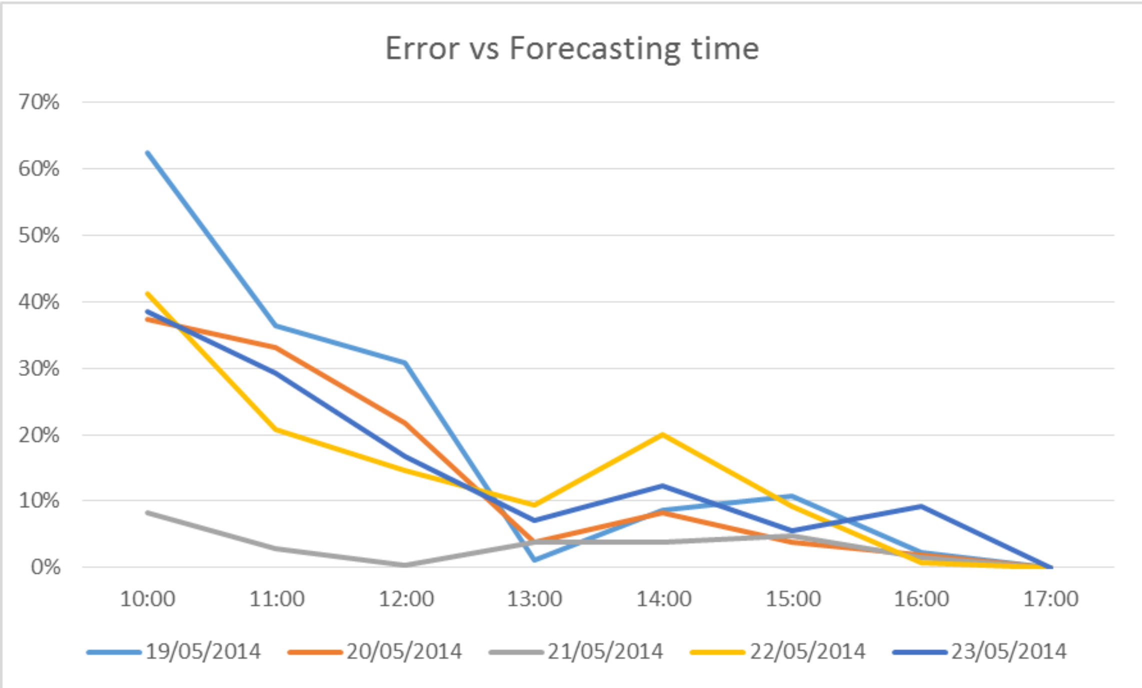


Figure 29: Forecasting error versus the time at which the forecast was made

6.5 Simulation Results

To test the promising results of the error calculation on the actual outcome of the system, the simulation model was run while varying the forecasting model. During the 28 days of operation the late order percentage and the cost were monitored. During this test it was found that it does not really matter if a simple moving average or the web shop forecast is used. Both methods produce similar results. Looking at the Neural Network it can be seen that on average the cost will be reduced. However, this cost reduction is created at the expense of late orders. This holds also for the full solution space presented in Appendix B.

	WEB SHOP FORECAST	**SIMPLE MOVING AVERAGE**	**NEURAL NETWORK**
LATE ORDERS	33%	34%	42%
COST	4106	4018	3917

Table 10: Results of the simulation while varying the forecast

7 Planning

In this chapter the planning decisions will be discussed. The production planning roughly consists of two elements, namely the line schedule and the shift times. Both will be discussed in this chapter. The first section will cover the scheduling and how this has been researched further. In section 2, varying the shift times will be explored. Finally, in section 3, the simulation results of this part of the research will be discussed.

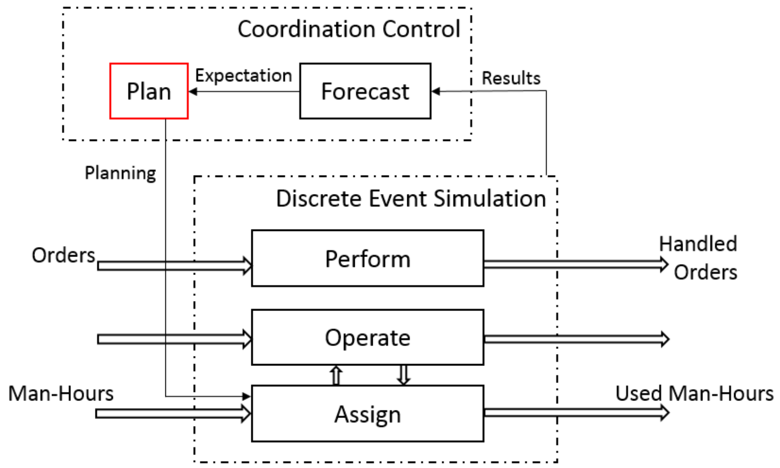


Figure 30: In this chapter, the planning will be discussed

7.1 Scheduling

In the last chapter the different forecasting techniques were discussed. Combining the orders that are left every day with the forecasts for the next day form the expectation for the next day. This expectation is translated in a schedule, which consists of an amount of lines to plan and an amount of shifts. This is done by the Planner. Based on interviews with the Planner the baseline schedule is estimated as the baseline entry in the table below.

EXPECTED AMOUNT OF ORDERS	AMOUNT OF SHIFTS PLANNED	AMOUNT OF LINES PLANNED
<1575	1 shift	1 line
<3150	1 shift	2 lines
<4725	1 shift	3 lines
<6300	2 shifts	2 lines
>6300	2 shifts	3 lines

Table 11: An example of how the amount of lines is planned

For example, in the baseline situation, when 3000 orders are expected this will result in 2 lines and 1 shift being planned. Changing this planning could be beneficial. In theory, heightening the thresholds for the different amount of lines would mean that more often a lower amount of lines is planned. This will probably impact cost positively and late orders negatively. To test this assumption, these threshold values should be varied.

To vary the different threshold variables in a systemic manner, the variable x is introduced. This variable regulates the threshold values according to the formula $threshold = amount\ of\ lines * amount\ of\ shifts * line\ capacity\ per\ hour * x$. This means that basically the value of x represents the shift time. Setting x to 8 would mean that the threshold for a single line is exactly the capacity of a single line for 8 hours. By varying these thresholds the selected amount of lines will vary. Examples of this are in Table 12.

THRESHOLD EXPECTED AMOUNT OF ORDERS	EXAMPLE THRESHOLD (X=6)	EXAMPLE THRESHOLD (X=9)	EXAMPLE THRESHOLD (X=12)
1*1*175*X	<1050	<1575	<2100
1*2*175*X	<2100	<3150	<4200
1*3*175*X	<3150	<4725	<6300
2*2*175*X	<4200	<6300	<7400
1*3*175*X	>4200	>6300	>7400

Table 12: Examples of different threshold values as a function of x

Looking at Table 12, it can be seen that increasing x will increase the threshold value for a certain amount of lines. This means that when x is higher a lower amount of lines will be planned, which could reduce the cost. Vice versa, a lower x would mean that more capacity is available, possibly reducing the amount of late orders.

7.2 Varying Shift Times

There are two possibilities in varying shift times, namely changing the starting time or changing the duration. For varying the starting time the following rule was used: "if the amount of orders exceeds full capacity between 1 and 9 pm, start s hours early. A deliberate choice was made to use a fixed amount of hours instead of a variable amount based on the forecast for example. This was done to eliminate forecasting accuracy from influencing this parameter. In this research s ranges from 0 to 4 hours early. The duration was varied from 6 to 14 hours. Note that this is not a fixed duration. Shifts can still end sooner if the amount of capacity planned proves too large. The duration is a maximal duration. If this exceeds 8 hours this means that overtime is desirable.

7.3 Simulation Results

Simulation results for varying x are presented in Table 13. The assumption that late orders decreases for lower x holds. However, the cost increase is not proved with these results

	x = 5	x = 6	x = 7	x = 8	x = 9	x = 10	x = 11	x = 12	x = 13	x = 14
Late Orders	30%	31%	31%	31%	33%	40%	48%	48%	53%	56%
Cost	3945	4102	4157	4206	4106	3914	3885	3885	4055	4202

Table 13: Simulation result of varying the amount of lines planned

For this test the starting time was advanced by an s amount of hours whenever the forecast exceeded 3 times the line capacity for one shift. In other words, whenever it was expected that the work would not be finished before 9 pm and therefore orders would be late, the starting time was advanced by s hours. The other factor that was varied is the shift duration d. This represents whether people are allowed to work overtime to clear the order load. Results are depicted below.

		d = 6	d = 7	d = 8	d = 9	d = 10	d = 11	d = 12	d = 13	d = 14
start at 13:00	Late Orders	51%	39%	34%	33%	33%	33%	33%	33%	33%
	Cost	4032	3928	4021	4106	4106	4106	4106	4106	4106
start at 12:00	Late Orders	44%	33%	29%	28%	28%	28%	28%	28%	28%
	Cost	4032	3928	4021	4124	4124	4124	4124	4124	4124
start at 11:00	Late Orders	38%	28%	23%	24%	24%	24%	24%	24%	24%
	Cost	4032	3928	4021	4066	4066	4066	4066	4066	4066
start at 10:00	Late Orders	32%	23%	19%	20%	19%	19%	19%	19%	19%
	Cost	4043	3943	4032	4077	4099	4099	4099	4099	4099
start at 9:00	Late Orders	27%	19%	15%	15%	14%	14%	14%	14%	14%
	Cost	4069	3984	4057	4103	4129	4129	4129	4129	4129

Table 14: Simulation results of varying starting time s and duration d

These results show that the duration has no effect on late orders and cost given the fact that duration is large enough. Advancing the starting time has a lot of effect possibly cutting the late order percentages in half.

8 Results

By combining all the scenarios presented a full solution space can be generated. Varying results over three different forecasting scenarios and using 5 different start-times generates 15 tables of data. This full solution space is presented in Appendix B.

By collecting the averages of all options tried in the last chapters more general relations between the different forecasting and planning parameters can be found. When calculating the averages on the cost results the results seem very similar.

		Mondelez Forecast			Simple Moving Average			Neural Network		
Starting Time	Variables	Min Result	Average Result	Max Result	Min Result	Average Result	Max Result	Min Result	Average Result	Max Result
13:00	0	3767	3995	4206	3727	4013	4180	3778	4059	4557
12:00	-1	3767	4017	4252	3727	4041	4226	3775	4086	4663
11:00	-2	3767	4020	4261	3727	4050	4244	3760	4065	4651
10:00	-3	3786	4027	4250	3735	4059	4256	3746	4057	4625
09:00	-4	3815	4087	4276	3734	4113	4248	3770	4062	4751

Table 15: Cost results of all simulation runs in man-hours

When the average late order percentage it becomes clear that the amount of late orders decreases as the starting time is advanced further.

		Mondelez Forecast			Simple Moving Average			Neural Network		
Starting Time	Starting Time	Min Result	Average Result	Max Result	Min Result	Average Result	Max Result	Min Result	Average Result	Max Result
13:00	0	30%	42%	95%	30%	42%	98%	31%	46%	118%
12:00	-1	24%	36%	89%	24%	36%	92%	25%	40%	116%
11:00	-2	19%	30%	82%	19%	31%	86%	20%	35%	113%
10:00	-3	16%	26%	76%	16%	26%	80%	16%	31%	104%
09:00	-4	13%	22%	70%	13%	22%	74%	13%	27%	105%

Table 16: Late order percentage results of all simulation runs

With the tables generated from the simulation it is possible to look at the simulation in more detail. To map the feasible solutions the table has been coloured. The baseline is represented as the blue entry. Whenever a combination of parameters improves on the baseline in terms of cost and late order percentage the solution is marked green. An orange solution means that the late order percentage has improved, but that in terms of cost, this solution performs worse. For a yellow solution the results are the other way around. Cost has improved, but more late orders are generated in the process. A red solution performs worse on all measured variables.

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	43%	33%	30%	30%	30%	30%	30%	30%	30%
	Cost	3991	3840	3945	3945	3945	3945	3945	3945	3945
x=6	Late Orders	48%	36%	31%	31%	31%	31%	31%	31%	31%
	Cost	4126	3909	4017	4102	4102	4102	4102	4102	4102
x=7	Late Orders	48%	36%	32%	31%	31%	31%	31%	31%	31%
	Cost	4183	3964	4072	4157	4157	4157	4157	4157	4157
x=8	Late Orders	48%	37%	32%	31%	31%	31%	31%	31%	31%
	Cost	4161	4013	4121	4206	4206	4206	4206	4206	4206
x=9	Late Orders	51%	39%	34%	33%	33%	33%	33%	33%	33%
	Cost	4032	3928	4021	4106	4106	4106	4106	4106	4106
x=10	Late Orders	64%	44%	39%	40%	40%	40%	40%	40%	40%
	Cost	3976	3783	3866	3914	3914	3914	3914	3914	3914
x=11	Late Orders	70%	56%	49%	48%	48%	47%	47%	47%	48%
	Cost	3932	3796	3842	3885	3878	3878	3878	3878	3882
x=12	Late Orders	76%	55%	49%	48%	48%	49%	49%	48%	48%
	Cost	3930	3767	3842	3885	3878	3895	3888	3880	3882
x=13	Late Orders	84%	63%	56%	53%	50%	50%	40%	37%	36%
	Cost	3977	3767	3828	4055	3947	3997	3903	3882	3906
x=14	Late Orders	95%	65%	57%	56%	55%	46%	42%	38%	37%
	Cost	3997	3808	3895	4202	4085	4131	4046	3961	3986

Table 17: The results for the web shop forecast with s = 0 (13:00)

When comparing Table 17 to Table 18, it is found that for Table 18, which represents the solutions for when the operation is started 3 hours earlier on an expected busy day, the amount of improving solutions increases considerably. This observation holds for all forecasting methods.

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	27%	18%	16%	16%	16%	16%	16%	16%	16%
	Cost	4042	3906	4010	3993	3916	3916	3916	3916	3916
x=6	Late Orders	30%	21%	17%	16%	16%	16%	16%	16%	16%
	Cost	4176	3974	4082	4151	4073	4073	4073	4073	4073
x=7	Late Orders	30%	21%	17%	16%	16%	16%	16%	16%	16%
	Cost	4194	3990	4097	4166	4088	4088	4088	4088	4088
x=8	Late Orders	30%	21%	17%	16%	17%	17%	17%	17%	17%
	Cost	4172	4024	4132	4201	4199	4199	4199	4199	4199
x=9	Late Orders	32%	23%	19%	20%	19%	19%	19%	19%	19%
	Cost	4043	3943	4032	4077	4099	4099	4099	4099	4099
x=10	Late Orders	48%	30%	25%	24%	22%	22%	22%	22%	22%
	Cost	4003	3846	3931	4014	4030	4030	4030	4030	4030
x=11	Late Orders	53%	39%	35%	32%	29%	29%	28%	28%	27%
	Cost	3948	3846	3907	4090	3979	3971	3962	3954	3957
x=12	Late Orders	60%	42%	35%	32%	29%	29%	28%	27%	26%
	Cost	3949	3786	3907	4090	3979	3971	3962	3932	3934
x=13	Late Orders	65%	50%	45%	34%	28%	24%	23%	19%	19%
	Cost	3989	3786	3848	4112	4159	4138	4013	3890	3915
x=14	Late Orders	76%	52%	45%	38%	28%	24%	23%	19%	19%
	Cost	3997	3813	3903	4250	4188	4168	4042	3919	3944

Table 18: The results for the web shop forecast with s = -3 (10:00)

Analysing the full set of simulation outcomes presented in Appendix B shows that:

- The lowest late order percentage possible without an increase in cost is 13%
- The lowest cost without an increased late order percentage is 3723 man-hours (9% decrease)

9 Implementation

With the results generated with the simulation the solution space is large. The tables generated as a result of this simulation can be used to support a choice for a certain decision system. To support this use the absolute values in the tables have been translated to a percentage relative to the baseline situation. An example of this is presented in Table 19. The full conversion is also given in Appendix B.

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	-18%	-45%	-52%	-52%	-52%	-52%	-52%	-52%	-52%
	Cost	-2%	-5%	-2%	-3%	-5%	-5%	-5%	-5%	-5%
x=6	Late Orders	-9%	-36%	-48%	-52%	-52%	-52%	-52%	-52%	-52%
	Cost	2%	-3%	-1%	1%	-1%	-1%	-1%	-1%	-1%
x=7	Late Orders	-9%	-36%	-48%	-52%	-52%	-52%	-52%	-52%	-52%
	Cost	2%	-3%	0%	1%	0%	0%	0%	0%	0%
x=8	Late Orders	-9%	-36%	-48%	-52%	-48%	-48%	-48%	-48%	-48%
	Cost	2%	-2%	1%	2%	2%	2%	2%	2%	2%
x=9	Late Orders	-3%	-30%	-42%	-39%	-42%	-42%	-42%	-42%	-42%
	Cost	-2%	-4%	-2%	-1%	0%	0%	0%	0%	0%
x=10	Late Orders	45%	-9%	-24%	-27%	-33%	-33%	-33%	-33%	-33%
	Cost	-3%	-6%	-4%	-2%	-2%	-2%	-2%	-2%	-2%
x=11	Late Orders	61%	18%	6%	-3%	-12%	-12%	-15%	-15%	-18%
	Cost	-4%	-6%	-5%	0%	-3%	-3%	-4%	-4%	-4%
x=12	Late Orders	82%	27%	6%	-3%	-12%	-12%	-15%	-18%	-21%
	Cost	-4%	-8%	-5%	0%	-3%	-3%	-4%	-4%	-4%
x=13	Late Orders	97%	52%	36%	3%	-15%	-27%	-30%	-42%	-42%
	Cost	-3%	-8%	-6%	0%	1%	1%	-2%	-5%	-5%
x=14	Late Orders	130%	58%	36%	15%	-15%	-27%	-30%	-42%	-42%
	Cost	-3%	-7%	-5%	4%	2%	2%	-2%	-5%	-4%

Table 19: The relative results for the web shop forecast with s = -3 (10:00)

So how should these tables be used? These tables provide a guideline on how to set the planning parameters for the system. For example, from Table 19 it follows that late orders could decrease by 52 percent at the same cost by:

- Using the same forecast as before (web shop)
- Starting 3 hours early when a busy day is expected (s = -3)
- Setting the max duration of a shift to 10 hours (d = 10)
- Setting x to 7 resulting in the threshold values presented in Table 20

EXPECTED AMOUNT OF ORDERS (X=7)	AMOUNT OF SHIFTS PLANNED	AMOUNT OF LINES PLANNED
<1225	1 shift	1 line
<2450	1 shift	2 lines
<3675	1 shift	3 lines
<4900	2 shifts	2 lines
>4900	2 shifts	3 lines

Table 20: An example of how the amount of lines is planned

Obviously, the actual result will not be a decrease in late orders of precisely 52%. The result is more indicative. However, from these tables it is clear that for example these planning parameters generate a better result on long term than for example using x=14 and a max duration of 6 hours.

That is how these solutions should be used. They do not guarantee a certain outcome, but can be used to assess the different options relative to each other. Therefore, by using these tables a long term strategy for planning orders can be established. Doing this would require the manager to pick a target late order percentage and cost and set the variables accordingly. The strength of generating all solutions like this is that it provides management with an overview of the bigger picture, allowing the selection of not just a technically desirable solution, but also one that takes secondary factors that might exist into account.

10 Conclusion and Recommendations

In this thesis the effect of multiple planning and forecasting decisions on the late orders and cost of the web shop packaging operation were tested. In this chapter the conclusions and recommendations of this research will be outlined.

10.1 Conclusions

To start with the first sub-question: What influence has using different forecasts? The conclusion that can be drawn from this thesis is that in the current set up, forecasts do not have a lot of influence. Even though the accuracy of a self-created forecast (MAPE 33%) is way better than that of the web shop (MAPE 67%), it turns out that the effect on the outcome of the simulation in terms of late orders and cost is minimal. However, it is expected that using a more accurate forecast should generate better results in the long run. Perhaps the amount of runs was insufficient.

What influence has the amount of lines planned? To measure the influence of planning, the schedule was defined as a result of parameter x. This parameter x was varied to see whether the schedule has influence. The expectation was that when a higher amount of lines was planned, the late orders would decrease and cost would increase. This relation turned out not to hold. An explanation for this is that multiple optima exist in the cost due to the effect of different planning agreements with the employment agency.

The influence of changing around the shift times was measured as well. The starting times where adapted whenever a lot of orders where expected. This resulted in the largest measurable influence of all parameters. Doing so potentially cuts late orders in half while at equal costs. Starting earlier was already done to catch up on late orders. The results of this thesis indicate that it is a good idea to do this preventively from now on. By varying the duration it was found that this should not be too short. Therefore, it is advised to negotiate the possibility of two hours overtime for every shift when order volumes are high.

Regarding the implementation of the results, a full solution space was created using all possible outcomes of the situation. Using these solution tables individual cases where the system performs well can be identified. This allows management to select the solutions that are not only desirable for the amount of late orders, but also allows them to take into account any other external factors that might play a role when selecting a decision strategy.

To return to the main question of this thesis. Can the amount of late orders be reduced to 2% of total handled orders by changing the planning without an increase in cost? According to the results of this thesis it cannot. However, either the late order percentage can be reduced from 33% to 13% or the cost could be reduced by approximately 9%.

10.2 Recommendations

One of the problems identified in the problem analysis was the dependence of an inaccurate forecast for daily operation. Seeing that the influence of forecasts on operation is minimal and that the forecasting result with a simple moving average over the last week is already a lot better than the web shop generated forecast, there definitely seems no need to depend on the web shop forecasts as an input for daily operation. Much more is gained from communicating with the web shops about their promotional actions and structurally documenting this communication for modelling purposes.

For this research a neural network model was tested, but not a lot of data was available to train this model. The amount of historic order volumes available turned out to be a little over a year. This was combined with the day of the week and the day of the year to enable the neural network to identify any seasonal effects. These effects could be identified better if several years of data were available. Data on promotions and other actions on client side was not available at all. It is suspected that these actions greatly influence daily order volumes and that the results of this forecasting method will improve as more data becomes available. Therefore, this should be researched further.

For the company in general it was found during this thesis that a lot of data remains in the system unused during the actual packaging process. This could be due to either aging IT systems or limited staff knowledge about these systems. It is recommended to use the available data better during the daily processes. In e-commerce, using the available data is what could give a company the competitive edge.

11 References

Babuška, R. (2010). *Knowledge Based Control Systems (Lecture Notes).* TU Delft.

Coghlan, A. (2014, January 24). *A Little Book of R for Time Series.* Retrieved September 17, 2014, from readthedocs.org: http://a-little-book-of-r-for-time-series.readthedocs.org/

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems, 2*(4), 303-314.

Franses, P., van Dijk, D., & Opschoor, A. (2014). *Time Series Models for Business and Economic Forecasting* (2nd ed.). Cambridge: Cambridge University Press.

Hillier, F., & Lieberman, G. (2001). *Introduction to Operations Research* (7th ed.). New York: McGraw-Hill.

Ho, S., Xie, M., & Goh, T. (2002). A comparative study of neural network and Box-Jenkins ARIMA modeling in time-series prediction. *Computers & Industrial Engineering*, 371-375.

Hyndman, R., & Athanasopoulos, G. (2013). *Forecasting: principles and practice.* OTexts.org. Retrieved November 25, 2014, from http://otexts.org/fpp/

Khashei, M., & Bijari, M. (2011). A novel hybridization of artificial neural networks and ARIMA forecasting for time-series forecasting . *Applied Soft Computing*, 2664-2675.

Tang, Z., de Almeida, C., & Fishwick, P. (1991). Time series forecasting using neural networks vs. Box-Jenkins methodology. *Simulation 51*, 303-310.

Veeke, H., Ottjes, J., & Lodewijks, G. (2008). *The Delft Systems Approach.* London: Springer-Verlag.

Wong, W., Xia, M., & Chu, W. (2010). Adaptive neural network model for time-series forecasting. *European Journal of Operations Research, 207*, 807-816.

Appendix A Scientific Research Paper

Improving operations at a fulfilment operation at PostNL E-Commerce Services by improving the order volume forecasts and controlling planning parameters

I.C. Gelens BSc., Dr. ir. H.P.M. Veeke, Prof. Dr. Ir. G. Lodewijks
Faculty of mechanical Engineering
Delft University of Technology
Delft, the Netherlands
i.c.gelens@student.tudelft.nl

Abstract— **PostNL E-Commerce Services is an e-fulfilment centre, which means that they keep stock, pick, package and send out goods for web shops. For a web shop they handle the European distribution. On this packaging operation it turned out that 33% of orders were handled late. A discrepancy was found between the amount of orders placed and handled every day possibly resulting in late orders and increased cost. To reduce this discrepancy a simulation model was created to test the impact of different planning decisions on the performance. By experimenting with different forecasting models, the forecasting accuracy was greatly increased. However, the effect on the amount of late orders was minimal. Other planning parameters that were researched were the schedule resulting from the forecast, the starting time s and the maximal shift duration d. From these parameters, starting time s turned out to have the greatest impact. By changing the planning either late orders can be reduced from 33% to 13% of total handled orders or cost could be reduced by 9%. The full results are mapped and can be used as a guidance for planning on this operation.**

Index Terms— **fulfilment, e-commerce, planning, forecasting**

INTRODUCTION

PostNL E-Commerce Services was founded as a fulfilment centre for print media under the name TopPak. With most printed media going out of business TopPak changed its business to e-fulfilment, which means keeping stock, packaging and sending out products for web shops. This is a successful business model for all parties involved because it saves the web shops the initial investment in a distribution centre and allows TopPak to make a margin on every parcel. Because these activities complement the logistic activities of PostNL, TopPak was acquired in 2010 and changed its name to PostNL E-Commerce Services.

For a web shop, PostNL ECS handles the European distribution of their Senseo-like coffee concept. Not only is the order volume substantial, but this is also expected to increase over the coming years. Therefore, the question from PostNL ECS is how the activities deployed on this packaging operation can be expanded.

When analysing the packaging operation, a process description was formulated of which the key points are:

- All packaging is done by hand using a pick-by-light order picking system.
- The packaging process consists of three packaging lines which can be deployed based on demand.
- Production capacity is planned on a weekly basis, but shifts can be adjusted one day in advance.
- The packaging staff is hired at an employment agency and paid per hour, so hours are fairly flexible. However, a shift takes a minimum of 5 hours due to contractual obligations.
- A maximum amount of two eight hour shifts is planned every day. So the first shift generally starts at 1 pm and ends at 9 pm. When it is needed to catch up on late orders, workdays can start sooner.
- All orders that arrive before cut-off time (5 pm) need to be handled the same day before 9 pm, so daily order load consists of the amount of orders placed the previous day after 5 pm and whatever comes in between the morning and 5 pm.

To assess this problem, the packaging operation was analysed with the Delft Systems Approach [1] and looking at data from the company database. From this problem analysis it followed that capacity increase over the following years is not necessary, because in case of the maximal order growth possible, current capacity will only be insufficient several days a year. Furthermore, this capacity shortage could be easily solved by adding a nightshift.

However, during the process analysis a large misalignment was discovered between the orders being placed and the production capacity available to process these orders. The result is that in spite of contractual obligations of delivering 98% of orders on time, currently only 67% actually is handled on time. Reducing this misalignment is potentially beneficial to the operation of every web shop at PostNL ECS. Therefore, this is the opportunity that has been investigated further.

The final research question has been formulated as: "*Can the amount of orders handled late be reduced to 2% of total orders handled by changing the planning without an increase in cost?*" To research this further, the different responses to planning strategies were tested in two fields, namely:

- Forecasting. An effort was made to improve the forecasts to have a better basis for decision making.
- Planning. The different planning decisions based on this forecast were explored. Also, the effect of shift start- and end-times on the operation was measured.

METHOD

Because the focus of this research lies mainly on the impact of forecasts and planning decisions, a relatively simple discrete event simulation mimicking the daily operations was created. This operation was influenced by either changing the forecast or changing the planning resulting from this forecast. The performance was measured by collecting the amount of late orders and the cost in man-hours afterwards.

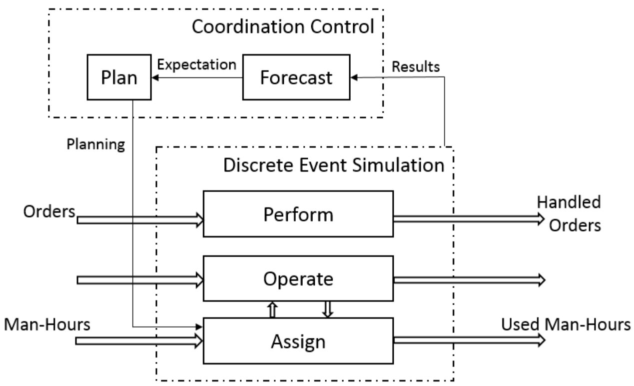


Figure 31: The simulation model

To test the impact of the different measures 28 days of operation was simulated starting at May 16th, 2014. This period of 28 days has been chosen, because during this period order volumes fluctuate reaching almost peak capacity as well as a very low capacity. The use of neural network forecasting and a large amount of combinations of different model parameters made discrete event simulation with Matlab® the simulation method of choice. During this simulation every day an amount of orders was processed. These orders were generated according to the actual incoming order distribution of those days taken from historic data. For the simulation the following constraints were established.

- Orders handled after 9 pm are late
- When people are planned, they work for at least 5 hours (minimal shift duration is 5 hours)
- Maximum amount of shifts is two eight-hour shifts
- People will work 1 hour of overtime unannounced (maximal shift duration is 9 hours)

Also a set of assumptions have been added.

- Cost in man-hours is the only cost considered
- Even though all orders are packaged manually and therefore the capacity may fluctuate, capacity per line is fixed at 175 orders per hour
- Planning is made 1 day in advance
- The simulation will start empty
- No difference in hourly rates of personnel, no matter what time.

The simulation works as follows:

- Use forecast method X to predict orders for the following day
- Based on this forecast, set an amount of lines and shifts
- Set a shift starting time and ending time
- Run the next day of simulation using the set parameters
- Repeat this for 28 days

After 28 days observe total cost and late order percentage

A baseline run was created by running the simulation with parameters as they have been estimated by discussing the operation with the Planner. The simulation using the web shop forecast as an input for planning, a shift starting time of 1 pm and a max duration of 9 hours per shift yielded a late order percentage of 33% and a cost of 4106 man-hours. This was used as a starting point for further improvement.

Forecasting

To decide on what forecasting methods could be used to improve the basis for planning, a literature research was conducted from which roughly four classes of methods can be identified [2] [3] [4] [5], namely:

- Simple methods
- Classical time-series models
- Artificial neural networks
- Combinations between classical time-series and artificial neural networks

Classical time-series modelling generates a rigid model. Based on the data available, it cannot be guaranteed that this model will stay valid over time without intervention of forecasting experts. Therefore, methods based on this approach were dropped. Three methods were tested over 5 months of historic data, namely the naïve method, the simple moving average and an artificial neural network which yielded a Mean Average Percentage Error (MAPE) of 50%, 36% and 33% respectively. Compared to the web shop forecast with a MAPE of 67% this is a big improvement in sense of forecasting accuracy.

Planning

Every day, the forecast is translated in a schedule, which consists of an amount of lines to plan. This is done by the Planner by translating the forecasted amount of orders to an amount of lines and shifts. Based on interviews with the Planner the baseline situation has been estimated as the entry in the table below.

Expected Amount of Orders	Amount of Shifts Planned	Amount of Lines Planned
<1575	1 shift	1 line
<3150	1 shift	2 lines
<4725	1 shift	3 lines
<6300	2 shifts	2 lines
>6300	2 shifts	3 lines

Table 21: The table used to translate the forecast into a planning

For example, in the baseline situation, when 3000 orders are expected this will result in 2 lines and 1 shift being planned. Changing this planning could be beneficial. It was expected that planning more lines on average would have a negative effect on cost and a positive effect on late orders. To test this, the production schedule was defined as a formula with one dependent variable x:

$$threshold = amount_{lines} * amount_{shifts} * (\frac{capacity_{line}}{hour}) * x$$

In this formula, a higher x means that on average a lower amount of lines will be planned, which could reduce the cost. Vice versa, a lower x would mean that more capacity is available, possibly reducing the amount of late orders. An example of threshold values resulting from varying x is given in Table 22.

Threshold	Example (x=6)	Example (x=9)	Example (x=12)
1*1*175*x	<1050	<1575	<2100
1*2*175*x	<2100	<3150	<4200
1*3*175*x	<3150	<4725	<6300
2*2*175*x	<4200	<6300	<7400
2*3*175*x	>4200	>6300	>7400

Table 22: Examples of threshold values resulting from x

Apart from adapting the amount of lines resulting from the forecast, there are two other planning parameters that were tested, namely changing the starting time and changing the duration. For the starting time test, the starting time was advanced by an s amount of hours whenever the forecast exceeded 3 times the line capacity for one shift. In other words, whenever it was expected that the work would not be finished before 9 pm and therefore orders would be late, the starting time was advanced by s hours. The other test was varying the maximal shift duration d. Workers are sent home when the amount of orders is less than the expected amount (with a minimum of 5 hours). However, the max duration d is a measure for whether workers are allowed to work overtime as well. This variable could influence the amount of late hours, because whatever amount of orders remains after the workers are sent home will have to be handled the next day, adding to the order load.

RESULTS

The simulation model from the previous section was run using 28 days of data for every different forecasting method, planning parameter x, starting time s and max duration d. The results will be discussed in this section.

Three different forecasting models were selected for further testing. The original web shop forecast, the simple moving average and the artificial neural network. This yielded the following results.

	Web Shop Forecast	Simple Moving Average	Neural Network
Late Orders	33%	34%	42%
Cost	4106	4018	3917

Table 23: The influence of different forecasting models

What can be seen from these results is that the simple moving average performs similar to the web shop forecast. In general, the neural network performs a little better on cost, but this performance increase is at the expense of the late orders.

The planning parameter x was also varied. Partial results of varying x are presented below.

	x = 7	x = 8	x = 9	x = 10	x = 11
Late Orders	31%	31%	33%	40%	48%
Cost	4157	4206	4106	3914	3885

Table 24: A part of the influence of x

From these results it followed that planning more lines on average (decrease x) resulted in less late orders. The effects on cost are not as clear.

Finally, the influence of planning parameters s and d was measured. A part of these results is given in Table 25.

		d = 6	d = 7	d = 8	d = 9
start at 13:00	Late Orders	51%	39%	34%	33%
	Cost	4032	3928	4021	4106
start at 12:00	Late Orders	44%	33%	29%	28%
	Cost	4032	3928	4021	4124
start at 11:00	Late Orders	38%	28%	23%	24%
	Cost	4032	3928	4021	4066
start at 10:00	Late Orders	32%	23%	19%	20%
	Cost	4043	3943	4032	4077
start at 9:00	Late Orders	27%	19%	15%	15%
	Cost	4069	3984	4057	4103

Table 25: A part of the solution space for starting time s and maximal duration d

The results in terms of max duration d mainly show that the max duration should not be too short. Having people available to work overtime for 1-2 hours of overtime reduces the amount of late orders. There is no need to keep anyone longer and overtime only needs to be done when the order load is great.

Another result that follows from Table 25 is that starting earlier greatly reduces the amount of late orders. It is known that starting early was already done to catch up on late orders. However, these results show that it is very beneficial to start early preventively. This greatly reduces the amount of late orders and has no negative effect on cost in general.

A part of the full solution space is given in Table 26. A table like this was generated for every forecasting method and every starting time bringing the total to 15 tables. These tables show all the different outcomes of adjusting the planning parameters. The green marking means that the solution resulting from those parameters performs better in terms of both cost and late orders. Yellow and orange solutions perform better on either cost or late orders. Red solutions perform worse on all performance indicators. These tables can be used to decide on a planning strategy for this particular operation. Doing this would require the manager to pick a target late order percentage and cost and set the variables accordingly. In this way the manager can consider all options while taking into account other desirable factors not taken into account in this study.

x\d		d=6	d=7	d=8	d=9	d=10
x=6	Late Orders	36%	26%	21%	19%	19%
	Cost	4142	3926	4033	4102	4101
x=7	Late Orders	36%	26%	21%	19%	19%
	Cost	4183	3964	4072	4141	4139
x=8	Late Orders	36%	26%	21%	19%	19%
	Cost	4161	4013	4121	4190	4189
x=9	Late Orders	38%	28%	23%	24%	24%
	Cost	4032	3928	4021	4066	4066
x=10	Late Orders	52%	34%	30%	28%	28%
	Cost	3988	3800	3883	3965	3965

Table 26: A part of the full solution space

Analysing the full set of outcomes of the simulation shows that:

- The lowest late order percentage without an increase in cost is 13%
- The lowest cost without an increased late order percentage is 3723 man-hours (9% decrease)

Conclusion

Returning to the focus areas of the problem definition, it is found that in general:

- As a forecasting model, the simple moving average performs similar to the original web shop forecast. The neural network performs slightly better on cost, but increases the amount of late orders.
- As expected, on average, decreasing the amount of lines planned (a higher x) increases late orders. The influence of x on cost does not show this clear relation
- Extending the shifts hardly has any influence beyond 10 hours, which means that it should be possible to let workers stay for two extra hours when necessary
- Conditionally advancing the starting time greatly reduces the amount of late orders
- Individual cases where the system performs well can be identified using the outcomes of the simulation.

Can the amount of orders handled late be reduced to 2% of total orders handled by changing the planning without an increase in cost? According to the results of this thesis it cannot. However, either the late order percentage can be reduced from 33% to 13% or the cost could be reduced by approximately 9%.

DISCUSSION

For this research a neural network model was tested, but not a lot of data was available to train this model. Because data about promotional actions of the web shop was not recorded systemically, this could not be used for the forecasting models. This information is expected to significantly influence the order volumes. Therefore, it is expected that the results of this forecasting method will improve as more data becomes available. This could be researched further.

REFERENCES

[1] Veeke, H., Ottjes, J., & Lodewijks, G. (2008). The Delft Systems Approach. London: Springer-Verlag.

[2] Hillier, F., & Lieberman, G. (2001). Introduction to Operations Research (7th ed.). New York: McGraw-Hill.

[3] Franses, P., van Dijk, D., & Opschoor, A. (2014). Time Series Models for Business and Economic Forecasting (2nd ed.). Cambridge: Cambridge University Press.

[4] Ho, S., Xie, M., & Goh, T. (2002). A comparative study of neural network and Box-Jenkins ARIMA modeling in time-series prediction. Computers & Industrial Engineering, 371-375.

[5] Khashei, M., & Bijari, M. (2011). A novel hybridization of artificial neural networks and ARIMA forecasting for time-series forecasting . Applied Soft Computing, 2664-2675.

Appendix B Full Results

B.1 Absolute Results

B.1.1 Web Shop Forecast

The results for s = 0 (13:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	43%	33%	30%	30%	30%	30%	30%	30%	30%
	Cost	3991	3840	3945	3945	3945	3945	3945	3945	3945
x=6	Late Orders	48%	36%	31%	31%	31%	31%	31%	31%	31%
	Cost	4126	3909	4017	4102	4102	4102	4102	4102	4102
x=7	Late Orders	48%	36%	32%	31%	31%	31%	31%	31%	31%
	Cost	4183	3964	4072	4157	4157	4157	4157	4157	4157
x=8	Late Orders	48%	37%	32%	31%	31%	31%	31%	31%	31%
	Cost	4161	4013	4121	4206	4206	4206	4206	4206	4206
x=9	Late Orders	51%	39%	34%	33%	33%	33%	33%	33%	33%
	Cost	4032	3928	4021	4106	4106	4106	4106	4106	4106
x=10	Late Orders	64%	44%	39%	40%	40%	40%	40%	40%	40%
	Cost	3976	3783	3866	3914	3914	3914	3914	3914	3914
x=11	Late Orders	70%	56%	49%	48%	48%	47%	47%	47%	48%
	Cost	3932	3796	3842	3885	3878	3878	3878	3878	3882
x=12	Late Orders	76%	55%	49%	48%	48%	49%	49%	48%	48%
	Cost	3930	3767	3842	3885	3878	3895	3888	3880	3882
x=13	Late Orders	84%	63%	56%	53%	50%	50%	40%	37%	36%
	Cost	3977	3767	3828	4055	3947	3997	3903	3882	3906
x=14	Late Orders	95%	65%	57%	56%	55%	46%	42%	38%	37%
	Cost	3997	3808	3895	4202	4085	4131	4046	3961	3986

The results for s = -1 (12:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	37%	28%	25%	24%	24%	24%	24%	24%	24%
	Cost	3991	3840	3945	3991	3991	3991	3991	3991	3991
x=6	Late Orders	42%	31%	26%	24%	24%	24%	24%	24%	24%
	Cost	4126	3909	4017	4148	4148	4148	4148	4148	4148
x=7	Late Orders	42%	31%	26%	24%	24%	24%	24%	24%	24%
	Cost	4183	3964	4072	4203	4203	4203	4203	4203	4203
x=8	Late Orders	42%	31%	27%	25%	25%	25%	25%	25%	25%
	Cost	4161	4013	4121	4252	4252	4252	4252	4252	4252
x=9	Late Orders	44%	33%	29%	28%	28%	28%	28%	28%	28%
	Cost	4032	3928	4021	4124	4124	4124	4124	4124	4124
x=10	Late Orders	57%	39%	35%	35%	35%	35%	35%	35%	35%
	Cost	3976	3783	3866	3940	3940	3940	3940	3940	3940
x=11	Late Orders	63%	51%	44%	39%	38%	40%	39%	39%	38%
	Cost	3932	3796	3842	3906	3906	3902	3894	3886	3887
x=12	Late Orders	70%	51%	44%	42%	42%	42%	41%	41%	40%
	Cost	3930	3767	3842	3933	3926	3918	3911	3902	3904
x=13	Late Orders	78%	58%	53%	46%	43%	37%	34%	31%	30%
	Cost	3977	3767	3828	4088	3966	3971	4008	3906	3930
x=14	Late Orders	89%	60%	53%	47%	44%	38%	35%	30%	29%
	Cost	3997	3808	3895	4168	4171	4051	4088	4006	4039

The results for s = -2 (11:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	32%	23%	20%	19%	19%	19%	19%	19%	19%
	Cost	4008	3857	3962	3944	3944	3944	3944	3944	3944
x=6	Late Orders	36%	26%	21%	19%	19%	19%	19%	19%	19%
	Cost	4142	3926	4033	4102	4101	4101	4101	4101	4101
x=7	Late Orders	36%	26%	21%	19%	19%	19%	19%	19%	19%
	Cost	4183	3964	4072	4141	4139	4139	4139	4139	4139
x=8	Late Orders	36%	26%	21%	19%	19%	19%	19%	19%	19%
	Cost	4161	4013	4121	4190	4189	4189	4189	4189	4189
x=9	Late Orders	38%	28%	23%	24%	24%	24%	24%	24%	24%
	Cost	4032	3928	4021	4066	4066	4066	4066	4066	4066
x=10	Late Orders	52%	34%	30%	28%	28%	28%	28%	28%	28%
	Cost	3988	3800	3883	3965	3965	3965	3965	3965	3965
x=11	Late Orders	57%	46%	40%	35%	34%	33%	33%	32%	32%
	Cost	3933	3812	3858	4018	3996	3987	3979	3970	3972
x=12	Late Orders	64%	46%	40%	35%	34%	33%	33%	32%	32%
	Cost	3931	3767	3858	4018	4009	4001	3992	3984	3985
x=13	Late Orders	72%	53%	49%	37%	36%	31%	29%	23%	22%
	Cost	3977	3767	3828	4063	4026	3999	3981	3988	4021
x=14	Late Orders	82%	55%	49%	41%	33%	28%	27%	23%	22%
	Cost	3997	3808	3895	4236	4261	4162	4082	4052	4085

The results for s = -3 (10:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	27%	18%	16%	16%	16%	16%	16%	16%	16%
	Cost	4042	3906	4010	3993	3916	3916	3916	3916	3916
x=6	Late Orders	30%	21%	17%	16%	16%	16%	16%	16%	16%
	Cost	4176	3974	4082	4151	4073	4073	4073	4073	4073
x=7	Late Orders	30%	21%	17%	16%	16%	16%	16%	16%	16%
	Cost	4194	3990	4097	4166	4088	4088	4088	4088	4088
x=8	Late Orders	30%	21%	17%	16%	17%	17%	17%	17%	17%
	Cost	4172	4024	4132	4201	4199	4199	4199	4199	4199
x=9	Late Orders	32%	23%	19%	20%	19%	19%	19%	19%	19%
	Cost	4043	3943	4032	4077	4099	4099	4099	4099	4099
x=10	Late Orders	48%	30%	25%	24%	22%	22%	22%	22%	22%
	Cost	4003	3846	3931	4014	4030	4030	4030	4030	4030
x=11	Late Orders	53%	39%	35%	32%	29%	29%	28%	28%	27%
	Cost	3948	3846	3907	4090	3979	3971	3962	3954	3957
x=12	Late Orders	60%	42%	35%	32%	29%	29%	28%	27%	26%
	Cost	3949	3786	3907	4090	3979	3971	3962	3932	3934
x=13	Late Orders	65%	50%	45%	34%	28%	24%	23%	19%	19%
	Cost	3989	3786	3848	4112	4159	4138	4013	3890	3915
x=14	Late Orders	76%	52%	45%	38%	28%	24%	23%	19%	19%
	Cost	3997	3813	3903	4250	4188	4168	4042	3919	3944

The results for s = -4 (9:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	22%	14%	13%	13%	13%	13%	13%	13%	13%
	Cost	4088	3975	4079	4062	4033	4033	4033	4033	4033
x=6	Late Orders	25%	16%	13%	13%	13%	13%	13%	13%	13%
	Cost	4222	4043	4151	4220	4190	4190	4190	4190	4190
x=7	Late Orders	25%	16%	13%	13%	13%	13%	13%	13%	13%
	Cost	4219	4038	4146	4215	4217	4217	4217	4217	4217
x=8	Late Orders	25%	16%	13%	13%	13%	13%	13%	13%	13%
	Cost	4197	4050	4158	4227	4262	4262	4262	4262	4262
x=9	Late Orders	27%	19%	15%	15%	14%	14%	14%	14%	14%
	Cost	4069	3984	4057	4103	4129	4129	4129	4129	4129
x=10	Late Orders	44%	26%	22%	20%	20%	20%	20%	20%	20%
	Cost	4033	3875	3995	4083	4107	4107	4107	4107	4107
x=11	Late Orders	48%	31%	28%	28%	24%	24%	23%	23%	23%
	Cost	3978	3889	3961	4182	4059	4052	4052	4019	4029
x=12	Late Orders	49%	39%	32%	28%	24%	24%	23%	23%	23%
	Cost	3996	3815	3974	4182	4037	4029	4022	4019	4029
x=13	Late Orders	59%	44%	39%	30%	24%	22%	19%	16%	16%
	Cost	4012	3815	3891	4181	4249	4050	4011	3977	4010
x=14	Late Orders	70%	50%	42%	35%	24%	22%	19%	16%	16%
	Cost	3997	3821	3924	4276	4235	4036	4028	3995	4028

B.1.2 Simple Moving Average

The results for s = 0 (13:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	43%	33%	30%	30%	30%	30%	30%	30%	30%
	Cost	4086	3935	4040	4040	4040	4040	4040	4040	4040
x=6	Late Orders	48%	36%	32%	31%	31%	31%	31%	31%	31%
	Cost	4114	3895	4003	4088	4088	4088	4088	4088	4088
x=7	Late Orders	48%	37%	32%	31%	31%	31%	31%	31%	31%
	Cost	4132	3987	4095	4180	4180	4180	4180	4180	4180
x=8	Late Orders	49%	37%	32%	31%	31%	31%	31%	31%	31%
	Cost	4084	3956	4064	4149	4149	4149	4149	4149	4149
x=9	Late Orders	64%	44%	35%	34%	34%	34%	34%	34%	34%
	Cost	3976	3802	3933	4018	4018	4018	4018	4018	4018
x=10	Late Orders	69%	45%	42%	41%	41%	41%	41%	41%	41%
	Cost	4059	3868	3930	3999	3999	3999	3999	3999	3999
x=11	Late Orders	80%	45%	42%	41%	41%	41%	41%	41%	41%
	Cost	3983	3844	3907	3977	3977	3977	3977	3977	3977
x=12	Late Orders	86%	55%	50%	49%	49%	49%	46%	45%	45%
	Cost	4024	3813	3881	4176	4062	4029	4004	3996	3998
x=13	Late Orders	98%	54%	52%	54%	51%	48%	45%	44%	44%
	Cost	3897	3746	3822	4126	3995	3962	3939	3931	3933
x=14	Late Orders	98%	73%	59%	54%	51%	48%	41%	40%	36%
	Cost	3897	3727	3813	4126	3995	3966	3956	3925	4037

The results for s = -1 (12:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	37%	28%	25%	24%	24%	24%	24%	24%	24%
	Cost	4086	3938	4042	4088	4088	4088	4088	4088	4088
x=6	Late Orders	42%	31%	26%	24%	24%	24%	24%	24%	24%
	Cost	4114	3897	4005	4136	4136	4136	4136	4136	4136
x=7	Late Orders	42%	31%	27%	25%	25%	25%	25%	25%	25%
	Cost	4132	3987	4095	4226	4226	4226	4226	4226	4226
x=8	Late Orders	42%	31%	27%	25%	25%	25%	25%	25%	25%
	Cost	4084	3956	4064	4195	4195	4195	4195	4195	4195
x=9	Late Orders	57%	39%	30%	29%	29%	29%	29%	29%	29%
	Cost	3976	3804	3935	4021	4021	4021	4021	4021	4021
x=10	Late Orders	62%	39%	37%	35%	35%	35%	35%	35%	35%
	Cost	4059	3871	3932	4027	4027	4027	4027	4027	4027
x=11	Late Orders	73%	40%	37%	36%	36%	36%	36%	36%	36%
	Cost	3986	3846	3909	4006	4006	4006	4006	4006	4006
x=12	Late Orders	80%	50%	45%	46%	43%	40%	38%	38%	37%
	Cost	4027	3813	3883	4117	4078	4039	4010	4001	4003
x=13	Late Orders	92%	50%	48%	45%	42%	42%	38%	37%	35%
	Cost	3897	3746	3822	4050	4011	4054	4059	4050	4081
x=14	Late Orders	92%	69%	55%	49%	45%	37%	34%	33%	30%
	Cost	3897	3727	3813	4202	4096	4100	4064	4055	4083

The results for s = -2 (11:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	32%	23%	20%	19%	19%	19%	19%	19%	19%
	Cost	4086	3961	4065	4048	4048	4048	4048	4048	4048
x=6	Late Orders	36%	26%	22%	19%	19%	19%	19%	19%	19%
	Cost	4114	3920	4028	4097	4096	4096	4096	4096	4096
x=7	Late Orders	36%	26%	22%	19%	19%	19%	19%	19%	19%
	Cost	4132	3987	4095	4164	4163	4163	4163	4163	4163
x=8	Late Orders	37%	26%	21%	20%	20%	20%	20%	20%	20%
	Cost	4084	3956	4064	4163	4163	4163	4163	4163	4163
x=9	Late Orders	51%	34%	25%	24%	24%	24%	24%	24%	24%
	Cost	3976	3827	3958	4035	4035	4035	4035	4035	4035
x=10	Late Orders	55%	34%	32%	28%	28%	28%	28%	28%	28%
	Cost	4059	3894	3955	4059	4059	4059	4059	4059	4059
x=11	Late Orders	67%	34%	32%	28%	28%	28%	28%	28%	28%
	Cost	4005	3869	3932	4007	4006	4006	4006	4006	4006
x=12	Late Orders	73%	45%	41%	38%	37%	35%	32%	29%	29%
	Cost	4046	3815	3906	4144	4112	4070	4092	4129	4139
x=13	Late Orders	86%	45%	44%	40%	37%	31%	30%	29%	29%
	Cost	3897	3748	3822	4158	4117	4133	4133	4133	4143
x=14	Late Orders	86%	64%	51%	37%	35%	29%	27%	26%	23%
	Cost	3897	3727	3813	4180	4244	4136	4126	4102	4129

Results for s = -3 (10:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	27%	18%	16%	16%	16%	16%	16%	16%	16%
	Cost	4097	4009	4114	4097	4020	4020	4020	4020	4020
x=6	Late Orders	31%	21%	17%	16%	16%	16%	16%	16%	16%
	Cost	4125	3969	4077	4146	4068	4068	4068	4068	4068
x=7	Late Orders	31%	21%	17%	16%	16%	16%	16%	16%	16%
	Cost	4143	4002	4109	4178	4154	4154	4154	4154	4154
x=8	Late Orders	31%	21%	17%	17%	16%	16%	16%	16%	16%
	Cost	4095	3971	4078	4177	4123	4123	4123	4123	4123
x=9	Late Orders	45%	29%	21%	21%	19%	19%	19%	19%	19%
	Cost	3987	3856	4007	4084	4105	4105	4105	4105	4105
x=10	Late Orders	49%	29%	28%	24%	22%	22%	22%	22%	22%
	Cost	4059	3912	3987	4096	4080	4080	4080	4080	4080
x=11	Late Orders	61%	30%	28%	24%	22%	22%	22%	22%	22%
	Cost	4006	3887	3964	4045	4037	4037	4037	4037	4037
x=12	Late Orders	66%	42%	37%	34%	30%	26%	26%	25%	25%
	Cost	4047	3824	3938	4182	4214	4236	4136	4043	4045
x=13	Late Orders	80%	42%	41%	36%	29%	26%	26%	25%	25%
	Cost	3897	3768	3842	4192	4256	4195	4095	4001	4003
x=14	Late Orders	80%	59%	47%	34%	29%	25%	24%	23%	20%
	Cost	3897	3735	3825	4194	4238	4223	4083	3969	3989

Results for s = -4 (9:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	22%	14%	13%	13%	13%	13%	13%	13%	13%
	Cost	4120	4078	4183	4166	4137	4137	4137	4137	4137
x=6	Late Orders	25%	16%	14%	13%	13%	13%	13%	13%	13%
	Cost	4148	4038	4146	4215	4185	4185	4185	4185	4185
x=7	Late Orders	25%	16%	14%	13%	13%	13%	13%	13%	13%
	Cost	4166	4042	4149	4218	4221	4221	4221	4221	4221
x=8	Late Orders	26%	16%	13%	13%	13%	13%	13%	13%	13%
	Cost	4118	4011	4118	4217	4222	4222	4222	4222	4222
x=9	Late Orders	40%	18%	17%	17%	16%	16%	16%	16%	16%
	Cost	4010	3980	4056	4153	4179	4179	4179	4179	4179
x=10	Late Orders	42%	27%	23%	20%	18%	18%	18%	18%	18%
	Cost	4059	3972	4013	4142	4170	4170	4170	4170	4170
x=11	Late Orders	55%	27%	23%	20%	21%	21%	21%	21%	21%
	Cost	4010	3947	3990	4091	4145	4145	4145	4145	4145
x=12	Late Orders	60%	38%	33%	30%	25%	23%	22%	22%	21%
	Cost	4051	3897	3965	4228	4243	4125	4110	4107	4117
x=13	Late Orders	74%	39%	38%	33%	25%	23%	22%	22%	21%
	Cost	3897	3788	3884	4238	4196	4077	4062	4060	4070
x=14	Late Orders	74%	55%	44%	31%	25%	23%	20%	19%	17%
	Cost	3897	3734	3844	4234	4248	4080	4055	4027	4055

B.1.3 Neural Network

The results for s = 0 (13:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	43%	34%	31%	32%	31%	31%	31%	32%	31%
	Cost	4086	3990	4081	4118	4111	4094	4094	4148	4064
x=6	Late Orders	43%	35%	31%	32%	32%	31%	32%	31%	31%
	Cost	4018	3988	4026	4213	4111	4044	4092	4064	4125
x=7	Late Orders	52%	36%	33%	33%	32%	33%	31%	32%	32%
	Cost	3955	3829	3954	4036	3987	4049	4082	3887	3887
x=8	Late Orders	53%	44%	37%	38%	32%	32%	34%	36%	37%
	Cost	4033	3778	3929	3974	3971	3971	3934	4011	3905
x=9	Late Orders	69%	44%	36%	42%	37%	38%	37%	36%	37%
	Cost	3976	3778	3802	3917	3907	3974	3912	3962	3966
x=10	Late Orders	81%	53%	44%	45%	44%	38%	43%	42%	40%
	Cost	3920	3779	3860	4081	4118	4050	4038	3917	4167
x=11	Late Orders	86%	63%	53%	48%	43%	41%	44%	42%	43%
	Cost	3926	3785	3837	4216	4143	4130	4205	4196	4216
x=12	Late Orders	92%	69%	53%	49%	47%	47%	45%	44%	44%
	Cost	3920	3801	3795	4543	4277	4214	4317	4346	4300
x=13	Late Orders	105%	86%	62%	62%	60%	48%	44%	50%	39%
	Cost	3884	3798	3870	4393	4414	4399	4358	4249	4177
x=14	Late Orders	118%	93%	72%	69%	61%	53%	56%	40%	41%
	Cost	3884	3780	3838	4557	4510	4427	4349	4291	4230

The results for s = -1 (12:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	37%	29%	25%	25%	25%	25%	25%	25%	25%
	Cost	4099	3990	3993	4127	4127	4194	4106	4164	4177
x=6	Late Orders	42%	30%	28%	25%	25%	25%	25%	25%	25%
	Cost	4093	3998	4128	4153	4149	4177	4149	4177	4140
x=7	Late Orders	46%	32%	28%	26%	28%	26%	25%	29%	26%
	Cost	4030	3893	3967	4011	3952	4015	4145	4040	4015
x=8	Late Orders	47%	33%	32%	31%	31%	31%	31%	31%	31%
	Cost	3959	3792	3926	3945	4012	3945	4012	3945	3945
x=9	Late Orders	59%	39%	34%	32%	35%	29%	29%	29%	29%
	Cost	3929	3778	3840	4109	3934	4062	3995	3999	3995
x=10	Late Orders	72%	46%	36%	38%	39%	38%	35%	37%	36%
	Cost	3929	3775	3789	4110	4165	4144	4202	4027	4202
x=11	Late Orders	80%	55%	44%	41%	40%	38%	43%	37%	34%
	Cost	3926	3790	3845	4411	4240	4184	4068	4215	4180
x=12	Late Orders	86%	65%	52%	51%	41%	40%	41%	38%	40%
	Cost	3920	3780	3830	4496	4307	4336	4255	4318	4236
x=13	Late Orders	96%	72%	59%	46%	50%	45%	44%	32%	31%
	Cost	3920	3787	3870	4568	4309	4318	4428	4285	4357
x=14	Late Orders	116%	89%	65%	53%	45%	40%	36%	35%	32%
	Cost	3884	3780	3838	4663	4546	4321	4286	4167	4377

The results for s = -2 (11:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	32%	24%	21%	21%	21%	20%	20%	20%	20%
	Cost	4139	3977	4094	4196	4101	4077	4094	4077	4034
x=6	Late Orders	32%	24%	22%	21%	20%	20%	22%	21%	20%
	Cost	4030	3939	4089	4176	4026	4004	4125	4075	4026
x=7	Late Orders	40%	24%	23%	23%	23%	22%	22%	23%	23%
	Cost	4030	3902	3954	4052	4040	3968	3983	3901	3966
x=8	Late Orders	53%	28%	27%	26%	22%	24%	24%	24%	26%
	Cost	3956	3792	3926	4003	3985	3999	3932	3932	3937
x=9	Late Orders	56%	34%	29%	26%	25%	26%	25%	25%	28%
	Cost	4050	3778	3841	3937	3905	3914	3905	3905	3929
x=10	Late Orders	66%	41%	35%	34%	33%	34%	34%	32%	31%
	Cost	3929	3775	3855	4162	4076	4135	4152	4009	4063
x=11	Late Orders	71%	50%	40%	34%	33%	31%	32%	30%	31%
	Cost	3920	3830	3845	4297	4040	4139	4148	4188	4237
x=12	Late Orders	84%	59%	47%	37%	36%	29%	27%	28%	32%
	Cost	3879	3760	3828	4494	4272	4328	4258	4196	4198
x=13	Late Orders	89%	66%	54%	40%	34%	32%	32%	26%	26%
	Cost	3920	3808	3830	4559	4445	4368	4339	4282	4305
x=14	Late Orders	113%	86%	64%	56%	43%	34%	46%	29%	48%
	Cost	3925	3780	3870	4647	4651	4417	4342	4280	4327

Results for s = -3 (10:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	27%	19%	17%	17%	16%	17%	17%	17%	16%
	Cost	4056	3977	4092	4088	4017	4093	4011	4000	3987
x=6	Late Orders	28%	20%	17%	17%	19%	19%	19%	18%	19%
	Cost	4085	3989	4047	4093	4091	4052	4063	4081	4078
x=7	Late Orders	29%	21%	19%	18%	20%	19%	19%	20%	19%
	Cost	4068	3826	3952	4115	3959	4009	3927	4033	3942
x=8	Late Orders	36%	23%	22%	20%	19%	21%	21%	19%	21%
	Cost	3972	3859	3859	3932	3966	3908	3904	3899	3904
x=9	Late Orders	50%	31%	25%	21%	22%	22%	24%	21%	21%
	Cost	3976	3774	3838	4073	4030	4030	3985	3908	3975
x=10	Late Orders	60%	43%	32%	30%	28%	30%	26%	28%	27%
	Cost	3929	3746	3865	4134	4172	4151	4031	4033	4016
x=11	Late Orders	69%	51%	41%	35%	30%	28%	24%	26%	26%
	Cost	3884	3824	3798	4313	4195	4261	4073	4125	4125
x=12	Late Orders	78%	56%	37%	35%	30%	27%	28%	25%	24%
	Cost	3879	3800	3822	4503	4383	4332	4227	4219	4173
x=13	Late Orders	87%	72%	48%	33%	33%	31%	25%	38%	22%
	Cost	3961	3808	3820	4505	4463	4355	4278	4334	4251
x=14	Late Orders	104%	80%	60%	44%	35%	27%	29%	23%	26%
	Cost	3884	3780	3830	4580	4625	4452	4172	4183	4313

Results for s = -4 (9:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	22%	15%	13%	13%	15%	14%	13%	15%	13%
	Cost	4086	4041	4175	4034	4092	4037	4000	4148	4030
x=6	Late Orders	22%	18%	15%	14%	15%	14%	15%	15%	16%
	Cost	4073	3999	4128	4075	4167	4057	4118	4089	4062
x=7	Late Orders	26%	16%	14%	16%	13%	14%	15%	14%	13%
	Cost	4090	3839	3917	4134	3956	3987	3838	3922	3901
x=8	Late Orders	44%	24%	18%	17%	14%	17%	14%	14%	14%
	Cost	3997	3778	3926	4027	3989	3927	3927	3922	3922
x=9	Late Orders	44%	28%	23%	17%	17%	16%	22%	19%	20%
	Cost	3976	3773	3837	3975	3931	3981	4080	3972	3964
x=10	Late Orders	54%	33%	27%	26%	24%	20%	24%	22%	24%
	Cost	3929	3781	3886	4117	4014	4133	4018	4042	4030
x=11	Late Orders	57%	42%	33%	36%	27%	24%	26%	21%	24%
	Cost	3920	3790	3879	4215	4255	4136	4197	4201	4206
x=12	Late Orders	73%	52%	33%	32%	30%	23%	22%	23%	23%
	Cost	3961	3801	3779	4458	4411	4285	4245	4095	4155
x=13	Late Orders	83%	62%	43%	32%	31%	27%	31%	30%	21%
	Cost	3931	3770	3798	4572	4444	4224	4298	4338	4333
x=14	Late Orders	105%	76%	53%	52%	31%	39%	26%	29%	30%
	Cost	3801	3780	3830	4751	4617	4400	4284	4309	4277

B.2 Relative Results (Relative to Baseline)

B.2.1 Web Shop Forecast

The results for s = 0 (13:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	30%	0%	-9%	-9%	-9%	-9%	-9%	-9%	-9%
	Cost	-3%	-6%	-4%	-4%	-4%	-4%	-4%	-4%	-4%
x=6	Late Orders	45%	9%	-6%	-6%	-6%	-6%	-6%	-6%	-6%
	Cost	0%	-5%	-2%	0%	0%	0%	0%	0%	0%
x=7	Late Orders	45%	9%	-3%	-6%	-6%	-6%	-6%	-6%	-6%
	Cost	2%	-3%	-1%	1%	1%	1%	1%	1%	1%
x=8	Late Orders	45%	12%	-3%	-6%	-6%	-6%	-6%	-6%	-6%
	Cost	1%	-2%	0%	2%	2%	2%	2%	2%	2%
x=9	Late Orders	55%	18%	3%	0%	0%	0%	0%	0%	0%
	Cost	-2%	-4%	-2%	0%	0%	0%	0%	0%	0%
x=10	Late Orders	94%	33%	18%	21%	21%	21%	21%	21%	21%
	Cost	-3%	-8%	-6%	-5%	-5%	-5%	-5%	-5%	-5%
x=11	Late Orders	112%	70%	48%	45%	45%	42%	42%	42%	45%
	Cost	-4%	-8%	-6%	-5%	-6%	-6%	-6%	-6%	-5%
x=12	Late Orders	130%	67%	48%	45%	45%	48%	48%	45%	45%
	Cost	-4%	-8%	-6%	-5%	-6%	-5%	-5%	-6%	-5%
x=13	Late Orders	155%	91%	70%	61%	52%	52%	21%	12%	9%
	Cost	-3%	-8%	-7%	-1%	-4%	-3%	-5%	-5%	-5%
x=14	Late Orders	188%	97%	73%	70%	67%	39%	27%	15%	12%
	Cost	-3%	-7%	-5%	2%	-1%	1%	-1%	-4%	-3%

The results for s = -1 (12:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	12%	-15%	-24%	-27%	-27%	-27%	-27%	-27%	-27%
	Cost	-3%	-6%	-4%	-3%	-3%	-3%	-3%	-3%	-3%
x=6	Late Orders	27%	-6%	-21%	-27%	-27%	-27%	-27%	-27%	-27%
	Cost	0%	-5%	-2%	1%	1%	1%	1%	1%	1%
x=7	Late Orders	27%	-6%	-21%	-27%	-27%	-27%	-27%	-27%	-27%
	Cost	2%	-3%	-1%	2%	2%	2%	2%	2%	2%
x=8	Late Orders	27%	-6%	-18%	-24%	-24%	-24%	-24%	-24%	-24%
	Cost	1%	-2%	0%	4%	4%	4%	4%	4%	4%
x=9	Late Orders	33%	0%	-12%	-15%	-15%	-15%	-15%	-15%	-15%
	Cost	-2%	-4%	-2%	0%	0%	0%	0%	0%	0%
x=10	Late Orders	73%	18%	6%	6%	6%	6%	6%	6%	6%
	Cost	-3%	-8%	-6%	-4%	-4%	-4%	-4%	-4%	-4%
x=11	Late Orders	91%	55%	33%	18%	15%	21%	18%	18%	15%
	Cost	-4%	-8%	-6%	-5%	-5%	-5%	-5%	-5%	-5%
x=12	Late Orders	112%	55%	33%	27%	27%	27%	24%	24%	21%
	Cost	-4%	-8%	-6%	-4%	-4%	-5%	-5%	-5%	-5%
x=13	Late Orders	136%	76%	61%	39%	30%	12%	3%	-6%	-9%
	Cost	-3%	-8%	-7%	0%	-3%	-3%	-2%	-5%	-4%
x=14	Late Orders	170%	82%	61%	42%	33%	15%	6%	-9%	-12%
	Cost	-3%	-7%	-5%	2%	2%	-1%	0%	-2%	-2%

The results for s = -2 (11:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	-3%	-30%	-39%	-42%	-42%	-42%	-42%	-42%	-42%
	Cost	-2%	-6%	-4%	-4%	-4%	-4%	-4%	-4%	-4%
x=6	Late Orders	9%	-21%	-36%	-42%	-42%	-42%	-42%	-42%	-42%
	Cost	1%	-4%	-2%	0%	0%	0%	0%	0%	0%
x=7	Late Orders	9%	-21%	-36%	-42%	-42%	-42%	-42%	-42%	-42%
	Cost	2%	-3%	-1%	1%	1%	1%	1%	1%	1%
x=8	Late Orders	9%	-21%	-36%	-42%	-42%	-42%	-42%	-42%	-42%
	Cost	1%	-2%	0%	2%	2%	2%	2%	2%	2%
x=9	Late Orders	15%	-15%	-30%	-27%	-27%	-27%	-27%	-27%	-27%
	Cost	-2%	-4%	-2%	-1%	-1%	-1%	-1%	-1%	-1%
x=10	Late Orders	58%	3%	-9%	-15%	-15%	-15%	-15%	-15%	-15%
	Cost	-3%	-7%	-5%	-3%	-3%	-3%	-3%	-3%	-3%
x=11	Late Orders	73%	39%	21%	6%	3%	0%	0%	-3%	-3%
	Cost	-4%	-7%	-6%	-2%	-3%	-3%	-3%	-3%	-3%
x=12	Late Orders	94%	39%	21%	6%	3%	0%	0%	-3%	-3%
	Cost	-4%	-8%	-6%	-2%	-2%	-3%	-3%	-3%	-3%
x=13	Late Orders	118%	61%	48%	12%	9%	-6%	-12%	-30%	-33%
	Cost	-3%	-8%	-7%	-1%	-2%	-3%	-3%	-3%	-2%
x=14	Late Orders	148%	67%	48%	24%	0%	-15%	-18%	-30%	-33%
	Cost	-3%	-7%	-5%	3%	4%	1%	-1%	-1%	-1%

The results for s = -3 (10:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	-18%	-45%	-52%	-52%	-52%	-52%	-52%	-52%	-52%
	Cost	-2%	-5%	-2%	-3%	-5%	-5%	-5%	-5%	-5%
x=6	Late Orders	-9%	-36%	-48%	-52%	-52%	-52%	-52%	-52%	-52%
	Cost	2%	-3%	-1%	1%	-1%	-1%	-1%	-1%	-1%
x=7	Late Orders	-9%	-36%	-48%	-52%	-52%	-52%	-52%	-52%	-52%
	Cost	2%	-3%	0%	1%	0%	0%	0%	0%	0%
x=8	Late Orders	-9%	-36%	-48%	-52%	-48%	-48%	-48%	-48%	-48%
	Cost	2%	-2%	1%	2%	2%	2%	2%	2%	2%
x=9	Late Orders	-3%	-30%	-42%	-39%	-42%	-42%	-42%	-42%	-42%
	Cost	-2%	-4%	-2%	-1%	0%	0%	0%	0%	0%
x=10	Late Orders	45%	-9%	-24%	-27%	-33%	-33%	-33%	-33%	-33%
	Cost	-3%	-6%	-4%	-2%	-2%	-2%	-2%	-2%	-2%
x=11	Late Orders	61%	18%	6%	-3%	-12%	-12%	-15%	-15%	-18%
	Cost	-4%	-6%	-5%	0%	-3%	-3%	-4%	-4%	-4%
x=12	Late Orders	82%	27%	6%	-3%	-12%	-12%	-15%	-18%	-21%
	Cost	-4%	-8%	-5%	0%	-3%	-3%	-4%	-4%	-4%
x=13	Late Orders	97%	52%	36%	3%	-15%	-27%	-30%	-42%	-42%
	Cost	-3%	-8%	-6%	0%	1%	1%	-2%	-5%	-5%
x=14	Late Orders	130%	58%	36%	15%	-15%	-27%	-30%	-42%	-42%
	Cost	-3%	-7%	-5%	4%	2%	2%	-2%	-5%	-4%

The results for s = -4 (9:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	-33%	-58%	-61%	-61%	-61%	-61%	-61%	-61%	-61%
	Cost	0%	-3%	-1%	-1%	-2%	-2%	-2%	-2%	-2%
x=6	Late Orders	-24%	-52%	-61%	-61%	-61%	-61%	-61%	-61%	-61%
	Cost	3%	-2%	1%	3%	2%	2%	2%	2%	2%
x=7	Late Orders	-24%	-52%	-61%	-61%	-61%	-61%	-61%	-61%	-61%
	Cost	3%	-2%	1%	3%	3%	3%	3%	3%	3%
x=8	Late Orders	-24%	-52%	-61%	-61%	-61%	-61%	-61%	-61%	-61%
	Cost	2%	-1%	1%	3%	4%	4%	4%	4%	4%
x=9	Late Orders	-18%	-42%	-55%	-55%	-58%	-58%	-58%	-58%	-58%
	Cost	-1%	-3%	-1%	0%	1%	1%	1%	1%	1%
x=10	Late Orders	33%	-21%	-33%	-39%	-39%	-39%	-39%	-39%	-39%
	Cost	-2%	-6%	-3%	-1%	0%	0%	0%	0%	0%
x=11	Late Orders	45%	-6%	-15%	-15%	-27%	-27%	-30%	-30%	-30%
	Cost	-3%	-5%	-4%	2%	-1%	-1%	-1%	-2%	-2%
x=12	Late Orders	48%	18%	-3%	-15%	-27%	-27%	-30%	-30%	-30%
	Cost	-3%	-7%	-3%	2%	-2%	-2%	-2%	-2%	-2%
x=13	Late Orders	79%	33%	18%	-9%	-27%	-33%	-42%	-52%	-52%
	Cost	-2%	-7%	-5%	2%	3%	-1%	-2%	-3%	-2%
x=14	Late Orders	112%	52%	27%	6%	-27%	-33%	-42%	-52%	-52%
	Cost	-3%	-7%	-4%	4%	3%	-2%	-2%	-3%	-2%

B.2.2 Simple Moving Average

The results for s = 0 (13:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	30%	0%	-9%	-9%	-9%	-9%	-9%	-9%	-9%
	Cost	0%	-4%	-2%	-2%	-2%	-2%	-2%	-2%	-2%
x=6	Late Orders	45%	9%	-3%	-6%	-6%	-6%	-6%	-6%	-6%
	Cost	0%	-5%	-3%	0%	0%	0%	0%	0%	0%
x=7	Late Orders	45%	12%	-3%	-6%	-6%	-6%	-6%	-6%	-6%
	Cost	1%	-3%	0%	2%	2%	2%	2%	2%	2%
x=8	Late Orders	48%	12%	-3%	-6%	-6%	-6%	-6%	-6%	-6%
	Cost	-1%	-4%	-1%	1%	1%	1%	1%	1%	1%
x=9	Late Orders	94%	33%	6%	3%	3%	3%	3%	3%	3%
	Cost	-3%	-7%	-4%	-2%	-2%	-2%	-2%	-2%	-2%
x=10	Late Orders	109%	36%	27%	24%	24%	24%	24%	24%	24%
	Cost	-1%	-6%	-4%	-3%	-3%	-3%	-3%	-3%	-3%
x=11	Late Orders	142%	36%	27%	24%	24%	24%	24%	24%	24%
	Cost	-3%	-6%	-5%	-3%	-3%	-3%	-3%	-3%	-3%
x=12	Late Orders	161%	67%	52%	48%	48%	48%	39%	36%	36%
	Cost	-2%	-7%	-5%	2%	-1%	-2%	-2%	-3%	-3%
x=13	Late Orders	197%	64%	58%	64%	55%	45%	36%	33%	33%
	Cost	-5%	-9%	-7%	0%	-3%	-4%	-4%	-4%	-4%
x=14	Late Orders	197%	121%	79%	64%	55%	45%	24%	21%	9%
	Cost	-5%	-9%	-7%	0%	-3%	-3%	-4%	-4%	-2%

The results for s = -1 (12:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	12%	-15%	-24%	-27%	-27%	-27%	-27%	-27%	-27%
	Cost	0%	-4%	-2%	0%	0%	0%	0%	0%	0%
x=6	Late Orders	27%	-6%	-21%	-27%	-27%	-27%	-27%	-27%	-27%
	Cost	0%	-5%	-2%	1%	1%	1%	1%	1%	1%
x=7	Late Orders	27%	-6%	-18%	-24%	-24%	-24%	-24%	-24%	-24%
	Cost	1%	-3%	0%	3%	3%	3%	3%	3%	3%
x=8	Late Orders	27%	-6%	-18%	-24%	-24%	-24%	-24%	-24%	-24%
	Cost	-1%	-4%	-1%	2%	2%	2%	2%	2%	2%
x=9	Late Orders	73%	18%	-9%	-12%	-12%	-12%	-12%	-12%	-12%
	Cost	-3%	-7%	-4%	-2%	-2%	-2%	-2%	-2%	-2%
x=10	Late Orders	88%	18%	12%	6%	6%	6%	6%	6%	6%
	Cost	-1%	-6%	-4%	-2%	-2%	-2%	-2%	-2%	-2%
x=11	Late Orders	121%	21%	12%	9%	9%	9%	9%	9%	9%
	Cost	-3%	-6%	-5%	-2%	-2%	-2%	-2%	-2%	-2%
x=12	Late Orders	142%	52%	36%	39%	30%	21%	15%	15%	12%
	Cost	-2%	-7%	-5%	0%	-1%	-2%	-2%	-3%	-3%
x=13	Late Orders	179%	52%	45%	36%	27%	27%	15%	12%	6%
	Cost	-5%	-9%	-7%	-1%	-2%	-1%	-1%	-1%	-1%
x=14	Late Orders	179%	109%	67%	48%	36%	12%	3%	0%	-9%
	Cost	-5%	-9%	-7%	2%	0%	0%	-1%	-1%	-1%

The results for s = -2 (11:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	-3%	-30%	-39%	-42%	-42%	-42%	-42%	-42%	-42%
	Cost	0%	-4%	-1%	-1%	-1%	-1%	-1%	-1%	-1%
x=6	Late Orders	9%	-21%	-33%	-42%	-42%	-42%	-42%	-42%	-42%
	Cost	0%	-5%	-2%	0%	0%	0%	0%	0%	0%
x=7	Late Orders	9%	-21%	-33%	-42%	-42%	-42%	-42%	-42%	-42%
	Cost	1%	-3%	0%	1%	1%	1%	1%	1%	1%
x=8	Late Orders	12%	-21%	-36%	-39%	-39%	-39%	-39%	-39%	-39%
	Cost	-1%	-4%	-1%	1%	1%	1%	1%	1%	1%
x=9	Late Orders	55%	3%	-24%	-27%	-27%	-27%	-27%	-27%	-27%
	Cost	-3%	-7%	-4%	-2%	-2%	-2%	-2%	-2%	-2%
x=10	Late Orders	67%	3%	-3%	-15%	-15%	-15%	-15%	-15%	-15%
	Cost	-1%	-5%	-4%	-1%	-1%	-1%	-1%	-1%	-1%
x=11	Late Orders	103%	3%	-3%	-15%	-15%	-15%	-15%	-15%	-15%
	Cost	-2%	-6%	-4%	-2%	-2%	-2%	-2%	-2%	-2%
x=12	Late Orders	121%	36%	24%	15%	12%	6%	-3%	-12%	-12%
	Cost	-1%	-7%	-5%	1%	0%	-1%	0%	1%	1%
x=13	Late Orders	161%	36%	33%	21%	12%	-6%	-9%	-12%	-12%
	Cost	-5%	-9%	-7%	1%	0%	1%	1%	1%	1%
x=14	Late Orders	161%	94%	55%	12%	6%	-12%	-18%	-21%	-30%
	Cost	-5%	-9%	-7%	2%	3%	1%	0%	0%	1%

Results for s = -3 (10:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	-18%	-45%	-52%	-52%	-52%	-52%	-52%	-52%	-52%
	Cost	0%	-2%	0%	0%	-2%	-2%	-2%	-2%	-2%
x=6	Late Orders	-6%	-36%	-48%	-52%	-52%	-52%	-52%	-52%	-52%
	Cost	0%	-3%	-1%	1%	-1%	-1%	-1%	-1%	-1%
x=7	Late Orders	-6%	-36%	-48%	-52%	-52%	-52%	-52%	-52%	-52%
	Cost	1%	-3%	0%	2%	1%	1%	1%	1%	1%
x=8	Late Orders	-6%	-36%	-48%	-48%	-52%	-52%	-52%	-52%	-52%
	Cost	0%	-3%	-1%	2%	0%	0%	0%	0%	0%
x=9	Late Orders	36%	-12%	-36%	-36%	-42%	-42%	-42%	-42%	-42%
	Cost	-3%	-6%	-2%	-1%	0%	0%	0%	0%	0%
x=10	Late Orders	48%	-12%	-15%	-27%	-33%	-33%	-33%	-33%	-33%
	Cost	-1%	-5%	-3%	0%	-1%	-1%	-1%	-1%	-1%
x=11	Late Orders	85%	-9%	-15%	-27%	-33%	-33%	-33%	-33%	-33%
	Cost	-2%	-5%	-3%	-1%	-2%	-2%	-2%	-2%	-2%
x=12	Late Orders	100%	27%	12%	3%	-9%	-21%	-21%	-24%	-24%
	Cost	-1%	-7%	-4%	2%	3%	3%	1%	-2%	-1%
x=13	Late Orders	142%	27%	24%	9%	-12%	-21%	-21%	-24%	-24%
	Cost	-5%	-8%	-6%	2%	4%	2%	0%	-3%	-3%
x=14	Late Orders	142%	79%	42%	3%	-12%	-24%	-27%	-30%	-39%
	Cost	-5%	-9%	-7%	2%	3%	3%	-1%	-3%	-3%

Results for s = -4 (9:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	-33%	-58%	-61%	-61%	-61%	-61%	-61%	-61%	-61%
	Cost	0%	-1%	2%	1%	1%	1%	1%	1%	1%
x=6	Late Orders	-24%	-52%	-58%	-61%	-61%	-61%	-61%	-61%	-61%
	Cost	1%	-2%	1%	3%	2%	2%	2%	2%	2%
x=7	Late Orders	-24%	-52%	-58%	-61%	-61%	-61%	-61%	-61%	-61%
	Cost	1%	-2%	1%	3%	3%	3%	3%	3%	3%
x=8	Late Orders	-21%	-52%	-61%	-61%	-61%	-61%	-61%	-61%	-61%
	Cost	0%	-2%	0%	3%	3%	3%	3%	3%	3%
x=9	Late Orders	21%	-45%	-48%	-48%	-52%	-52%	-52%	-52%	-52%
	Cost	-2%	-3%	-1%	1%	2%	2%	2%	2%	2%
x=10	Late Orders	27%	-18%	-30%	-39%	-45%	-45%	-45%	-45%	-45%
	Cost	-1%	-3%	-2%	1%	2%	2%	2%	2%	2%
x=11	Late Orders	67%	-18%	-30%	-39%	-36%	-36%	-36%	-36%	-36%
	Cost	-2%	-4%	-3%	0%	1%	1%	1%	1%	1%
x=12	Late Orders	82%	15%	0%	-9%	-24%	-30%	-33%	-33%	-36%
	Cost	-1%	-5%	-3%	3%	3%	0%	0%	0%	0%
x=13	Late Orders	124%	18%	15%	0%	-24%	-30%	-33%	-33%	-36%
	Cost	-5%	-8%	-5%	3%	2%	-1%	-1%	-1%	-1%
x=14	Late Orders	124%	67%	33%	-6%	-24%	-30%	-39%	-42%	-48%
	Cost	-5%	-9%	-6%	3%	3%	-1%	-1%	-2%	-1%

B.2.3 Neural Network

The results for s = 0 (13:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	30%	3%	-6%	-3%	-6%	-6%	-6%	-3%	-6%
	Cost	0%	-3%	-1%	0%	0%	0%	0%	1%	-1%
x=6	Late Orders	30%	6%	-6%	-3%	-3%	-6%	-3%	-6%	-6%
	Cost	-2%	-3%	-2%	3%	0%	-2%	0%	-1%	0%
x=7	Late Orders	58%	9%	0%	0%	-3%	0%	-6%	-3%	-3%
	Cost	-4%	-7%	-4%	-2%	-3%	-1%	-1%	-5%	-5%
x=8	Late Orders	61%	33%	12%	15%	-3%	-3%	3%	9%	12%
	Cost	-2%	-8%	-4%	-3%	-3%	-3%	-4%	-2%	-5%
x=9	Late Orders	109%	33%	9%	27%	12%	15%	12%	9%	12%
	Cost	-3%	-8%	-7%	-5%	-5%	-3%	-5%	-4%	-3%
x=10	Late Orders	145%	61%	33%	36%	33%	15%	30%	27%	21%
	Cost	-5%	-8%	-6%	-1%	0%	-1%	-2%	-5%	1%
x=11	Late Orders	161%	91%	61%	45%	30%	24%	33%	27%	30%
	Cost	-4%	-8%	-7%	3%	1%	1%	2%	2%	3%
x=12	Late Orders	179%	109%	61%	48%	42%	42%	36%	33%	33%
	Cost	-5%	-7%	-8%	11%	4%	3%	5%	6%	5%
x=13	Late Orders	218%	161%	88%	88%	82%	45%	33%	52%	18%
	Cost	-5%	-8%	-6%	7%	8%	7%	6%	3%	2%
x=14	Late Orders	258%	182%	118%	109%	85%	61%	70%	21%	24%
	Cost	-5%	-8%	-7%	11%	10%	8%	6%	5%	3%

The results for s = -1 (12:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	12%	-12%	-24%	-24%	-24%	-24%	-24%	-24%	-24%
	Cost	0%	-3%	-3%	1%	1%	2%	0%	1%	2%
x=6	Late Orders	27%	-9%	-15%	-24%	-24%	-24%	-24%	-24%	-24%
	Cost	0%	-3%	1%	1%	1%	2%	1%	2%	1%
x=7	Late Orders	39%	-3%	-15%	-21%	-15%	-21%	-24%	-12%	-21%
	Cost	-2%	-5%	-3%	-2%	-4%	-2%	1%	-2%	-2%
x=8	Late Orders	42%	0%	-3%	-6%	-6%	-6%	-6%	-6%	-6%
	Cost	-4%	-8%	-4%	-4%	-2%	-4%	-2%	-4%	-4%
x=9	Late Orders	79%	18%	3%	-3%	6%	-12%	-12%	-12%	-12%
	Cost	-4%	-8%	-6%	0%	-4%	-1%	-3%	-3%	-3%
x=10	Late Orders	118%	39%	9%	15%	18%	15%	6%	12%	9%
	Cost	-4%	-8%	-8%	0%	1%	1%	2%	-2%	2%
x=11	Late Orders	142%	67%	33%	24%	21%	15%	30%	12%	3%
	Cost	-4%	-8%	-6%	7%	3%	2%	-1%	3%	2%
x=12	Late Orders	161%	97%	58%	55%	24%	21%	24%	15%	21%
	Cost	-5%	-8%	-7%	9%	5%	6%	4%	5%	3%
x=13	Late Orders	191%	118%	79%	39%	52%	36%	33%	-3%	-6%
	Cost	-5%	-8%	-6%	11%	5%	5%	8%	4%	6%
x=14	Late Orders	252%	170%	97%	61%	36%	21%	9%	6%	-3%
	Cost	-5%	-8%	-7%	14%	11%	5%	4%	1%	7%

The results for s = -2 (11:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	-3%	-27%	-36%	-36%	-36%	-39%	-39%	-39%	-39%
	Cost	1%	-3%	0%	2%	0%	-1%	0%	-1%	-2%
x=6	Late Orders	-3%	-27%	-33%	-36%	-39%	-39%	-33%	-36%	-39%
	Cost	-2%	-4%	0%	2%	-2%	-2%	0%	-1%	-2%
x=7	Late Orders	21%	-27%	-30%	-30%	-30%	-33%	-33%	-30%	-30%
	Cost	-2%	-5%	-4%	-1%	-2%	-3%	-3%	-5%	-3%
x=8	Late Orders	61%	-15%	-18%	-21%	-33%	-27%	-27%	-27%	-21%
	Cost	-4%	-8%	-4%	-3%	-3%	-3%	-4%	-4%	-4%
x=9	Late Orders	70%	3%	-12%	-21%	-24%	-21%	-24%	-24%	-15%
	Cost	-1%	-8%	-6%	-4%	-5%	-5%	-5%	-5%	-4%
x=10	Late Orders	100%	24%	6%	3%	0%	3%	3%	-3%	-6%
	Cost	-4%	-8%	-6%	1%	-1%	1%	1%	-2%	-1%
x=11	Late Orders	115%	52%	21%	3%	0%	-6%	-3%	-9%	-6%
	Cost	-5%	-7%	-6%	5%	-2%	1%	1%	2%	3%
x=12	Late Orders	155%	79%	42%	12%	9%	-12%	-18%	-15%	-3%
	Cost	-6%	-8%	-7%	9%	4%	5%	4%	2%	2%
x=13	Late Orders	170%	100%	64%	21%	3%	-3%	-3%	-21%	-21%
	Cost	-5%	-7%	-7%	11%	8%	6%	6%	4%	5%
x=14	Late Orders	242%	161%	94%	70%	30%	3%	39%	-12%	45%
	Cost	-4%	-8%	-6%	13%	13%	8%	6%	4%	5%

Results for s = -3 (10:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	-18%	-42%	-48%	-48%	-52%	-48%	-48%	-48%	-52%
	Cost	-1%	-3%	0%	0%	-2%	0%	-2%	-3%	-3%
x=6	Late Orders	-15%	-39%	-48%	-48%	-42%	-42%	-42%	-45%	-42%
	Cost	-1%	-3%	-1%	0%	0%	-1%	-1%	-1%	-1%
x=7	Late Orders	-12%	-36%	-42%	-45%	-39%	-42%	-42%	-39%	-42%
	Cost	-1%	-7%	-4%	0%	-4%	-2%	-4%	-2%	-4%
x=8	Late Orders	9%	-30%	-33%	-39%	-42%	-36%	-36%	-42%	-36%
	Cost	-3%	-6%	-6%	-4%	-3%	-5%	-5%	-5%	-5%
x=9	Late Orders	52%	-6%	-24%	-36%	-33%	-33%	-27%	-36%	-36%
	Cost	-3%	-8%	-7%	-1%	-2%	-2%	-3%	-5%	-3%
x=10	Late Orders	82%	30%	-3%	-9%	-15%	-9%	-21%	-15%	-18%
	Cost	-4%	-9%	-6%	1%	2%	1%	-2%	-2%	-2%
x=11	Late Orders	109%	55%	24%	6%	-9%	-15%	-27%	-21%	-21%
	Cost	-5%	-7%	-8%	5%	2%	4%	-1%	0%	0%
x=12	Late Orders	136%	70%	12%	6%	-9%	-18%	-15%	-24%	-27%
	Cost	-6%	-7%	-7%	10%	7%	6%	3%	3%	2%
x=13	Late Orders	164%	118%	45%	0%	0%	-6%	-24%	15%	-33%
	Cost	-4%	-7%	-7%	10%	9%	6%	4%	6%	4%
x=14	Late Orders	215%	142%	82%	33%	6%	-18%	-12%	-30%	-21%
	Cost	-5%	-8%	-7%	12%	13%	8%	2%	2%	5%

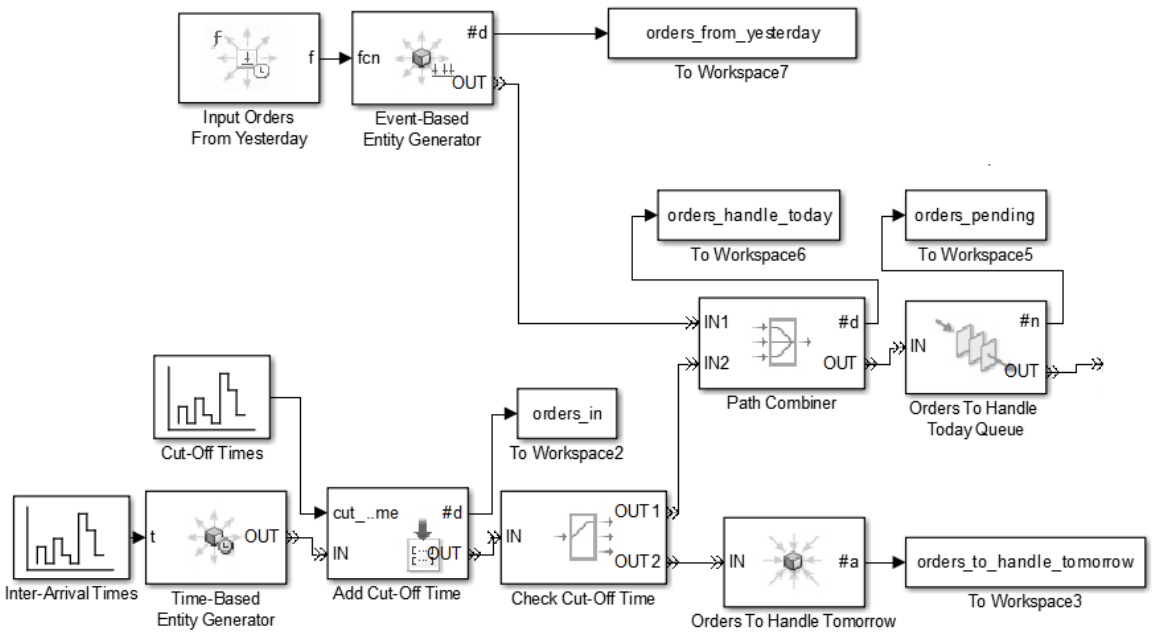
Results for s = -4 (9:00)

x\d		d=6	d=7	d=8	d=9	d=10	d=11	d=12	d=13	d=14
x=5	Late Orders	-33%	-55%	-61%	-61%	-55%	-58%	-61%	-55%	-61%
	Cost	0%	-2%	2%	-2%	0%	-2%	-3%	1%	-2%
x=6	Late Orders	-33%	-45%	-55%	-58%	-55%	-58%	-55%	-55%	-52%
	Cost	-1%	-3%	1%	-1%	1%	-1%	0%	0%	-1%
x=7	Late Orders	-21%	-52%	-58%	-52%	-61%	-58%	-55%	-58%	-61%
	Cost	0%	-7%	-5%	1%	-4%	-3%	-7%	-4%	-5%
x=8	Late Orders	33%	-27%	-45%	-48%	-58%	-48%	-58%	-58%	-58%
	Cost	-3%	-8%	-4%	-2%	-3%	-4%	-4%	-4%	-4%
x=9	Late Orders	33%	-15%	-30%	-48%	-48%	-52%	-33%	-42%	-39%
	Cost	-3%	-8%	-7%	-3%	-4%	-3%	-1%	-3%	-3%
x=10	Late Orders	64%	0%	-18%	-21%	-27%	-39%	-27%	-33%	-27%
	Cost	-4%	-8%	-5%	0%	-2%	1%	-2%	-2%	-2%
x=11	Late Orders	73%	27%	0%	9%	-18%	-27%	-21%	-36%	-27%
	Cost	-5%	-8%	-6%	3%	4%	1%	2%	2%	2%
x=12	Late Orders	121%	58%	0%	-3%	-9%	-30%	-33%	-30%	-30%
	Cost	-4%	-7%	-8%	9%	7%	4%	3%	0%	1%
x=13	Late Orders	152%	88%	30%	-3%	-6%	-18%	-6%	-9%	-36%
	Cost	-4%	-8%	-8%	11%	8%	3%	5%	6%	6%
x=14	Late Orders	218%	130%	61%	58%	-6%	18%	-21%	-12%	-9%
	Cost	-7%	-8%	-7%	16%	12%	7%	4%	5%	4%

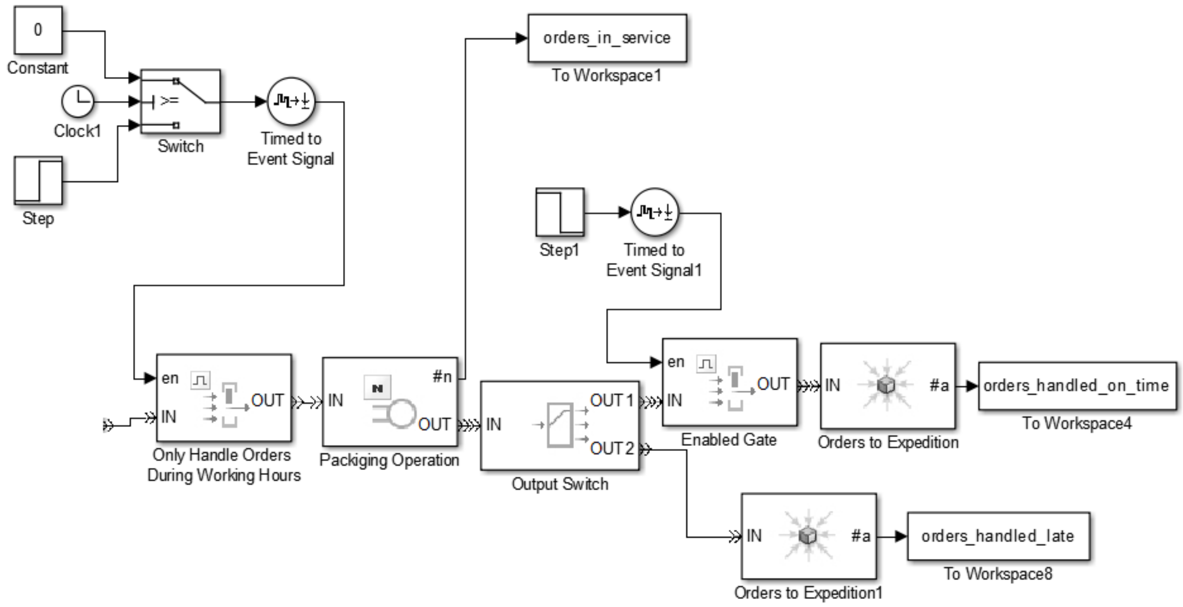
Appendix C Simulation Matlab® Code

C.1 Simulink Model

Part 1



Part 2



C.2 Model Code

C.2.1 Discrete Event Simulation

The following code is a Discrete Event Simulation of the web shop operation at PostNL E-Commerce Services. From day to day the packaging operation is simulated with an actual order distribution as input. The three types of forecasts can all be tested on this simulation to see how they perform.

Set Initial Parameters

In this part the initial parameters for the simulation are set. The simulation time is in minutes. For this simulation 53 days of data were available and have been used. A line capacity of 175 orders per hour is assumed.

```
simulation_time = 1980; %in minutes
starting_day = 327; %starting day of simulation (327 for real data)
simulation_period = 28; %amount of days to simulate (for real data max = 53)
linecapacity_hour = 175; %orders per hour
linecapacity = linecapacity_hour/60; %orders per minute
service_time = 3; %minutes
last_truck_leaves = 21; %time the truck leaves in hours
forecast_mode = 1; %1 for baseline, 2 for sma, 3 for neural network
orders_left = 0; %parameter to store orders left from yesterday
```

Read Data

Three sets of data are loaded. The excel-file simulation_data contains the distribution of incoming orders for 53 days. The other two files contain information as a basis for forecasts.

```
simulation_data = xlsread('simulation_data.xlsx','Sheet3');
forecast_mdlz = xlsread('Forecast_Mondelez.xlsx','Sheet1');
forecast_data = xlsread('OrderTotalenPerDag.xlsx','Blad1');
```

Simulate Packaging Lines

This is the heart of the simulation. The for-loop below makes a forecast everyday and uses this forecast to plan production. With this production planning as parameters a simulink model is executed for every day of operation. Weekends are simulated as well, but no orders are processed just like in real life.

```
for i = 1:simulation_period
```

```
% Select day of data and set simulation parameters
day = starting_day+i;
idx = (simulation_data(:,1)==i);
day_data = simulation_data(idx,:);
inter_arrival_times = day_data(:,4);
cut_off_times = transpose(day_data(:,6));
weekday = day_data(1,7);
```

Generate Forecast

Based on the forecast mode selected in simulation.m a forecast is made in this file. Mode 1 is the baseline, which is the forecast as delivered by Mondelez International. Mode 2 is the simple moving average. This means taking the average over the last 7 days as a forecast for the next day. Mode 3 is training a neural network on the dataset and using the network to make a forecast.

```
%Mode 1: Baseline
if forecast_mode == 1
    forecast_result = round(forecast_mdlz(i));
    collect_forecast_results(i) = forecast_result;

%Mode 2: Simple Moving Average
elseif forecast_mode == 2
    lastweek = forecast_data(day-8:day-1,2);
    forecast_result = round(mean(lastweek));
    collect_forecast_results(i) = forecast_result;

%Mode 3: Neural Network
else
    % Prepare data
    dataset = transpose(forecast_data(1:(day),2));
    days = transpose(forecast_data(1:(day),3:4));
    inputSeries  = con2seq(days);
    targetSeries = con2seq(dataset);

    % Set Neural Network parameters
    trainFcn = 'trainlm';  % Levenberg-Marquardt
    delays = 1;
    hiddenLayerSize = 5;
    net = narxnet(1:delays,1:delays,hiddenLayerSize,'open',trainFcn);
    [Xs,Xi,Ai,Ts] = preparets(net,inputSeries,{},targetSeries);

    %Train the Neural Network
    net = train(net,Xs,Ts,Xi,Ai);
    net = train(net,Xs,Ts,Xi,Ai);
    Y = net(Xs,Xi,Ai);

    % Make a prediciton for the next day
    nets = removedelay(net);
    [Xs,Xi,Ai,Ts] = preparets(nets,inputSeries,{},targetSeries);
    Ys = nets(Xs,Xi,Ai);
    forecast_result = round(cell2mat(Ys(end))*(17/24));
    collect_forecast_results(i) = forecast_result;
    %perf1 = perform(net,Ts,Y);
end


% Make a forecast for the amount of orders for this day
if weekday == 1
    order_expectation = collect_forecast_results(i-3)*2 + collect_forecast_results(i-
3)*(17/24) + collect_orders_left(i-3);
else
    order_expectation = forecast_result*(17/24) + orders_left;
end
```

Daily Simulation

This part simulates a day of operation. First the amount of lines is selected based on the forecast. Then the Simulink model is loaded and executed. The final part of this file collects the results from the Simulink model.

```
if order_expectation > 3*8*linecapacity_hour
    start_of_workday = (13*60) - (s*60);
else
    start_of_workday = 13*60;
end
```

Decide How Many Lines to Use Based on the Forecasting Result

```
if weekday > 5
    amount_lines = 1;
    amount_persons = 0;
    start_of_workday = 0;
    end_of_workday = 0;
elseif order_expectation < x*linecapacity_hour
    amount_lines = 1;
    amount_persons = 10;
    end_of_workday = start_of_workday+d*60;
elseif order_expectation < 2*x*linecapacity_hour
    amount_lines = 2;
    amount_persons = 17;
    end_of_workday = start_of_workday+d*60;
elseif order_expectation < 3*x*linecapacity_hour
    amount_lines = 3;
    amount_persons = 23;
    end_of_workday = start_of_workday+d*60;
elseif order_expectation < 4*x*linecapacity_hour
    amount_lines = 2;
    amount_persons = 17;
    end_of_workday = start_of_workday+2*d*60;
else
    amount_lines = 3;
    amount_persons = 23;
    end_of_workday = start_of_workday+2*d*60;
end
```

Set Some Daily Parameters

The amount of servers needs to be set. For this level of detail it is not really important how many servers there are, therefore, this value is a direct result from the required capacity.

```
servers = ceil(linecapacity*service_time*amount_lines);

% The function-block that generates the orders left from yesterday cannot
% be zero. This means that when zero orders are left from yesterday, one
% order too many will be processed in the simulation. This is not
% significant with order volumes in the thousands.
if orders_left == 0
```

```
    orders_left = 1;
end
```

Run simulink model

```
sim('model');
```

Collect results from simulation

Read relevant data from simulation time-series

```
orders_received_today = orders_in.data(end);
orders_to_handle_today = orders_handle_today.data(end);
orders_left_from_yesterday = orders_from_yesterday.data(end);
orders_received_before_cot = orders_to_handle_today - orders_left_from_yesterday;
orders_received_after_cot = orders_to_handle_tomorrow.data(end);
orders_handled_today = orders_handled_on_time.data(end) + orders_handled_late.data(end);
orders_handled_on_time_today = orders_handled_on_time.data(end);
orders_handled_late_today = orders_handled_late.data(end);

% Calculate the amount of late orders and time in operation
if weekday > 5
    late_orders = 0;
    orders_left = orders_received_today + orders_left_from_yesterday;
else
    late_orders = orders_to_handle_today - orders_handled_on_time_today;
    orders_left = orders_received_after_cot + late_orders - orders_handled_late_today;
end
collect_orders_left(i) = orders_left;
time_in_operation = orders_in_service.time(end) - start_of_workday;

% Calculate hours in operation
if weekday > 5
    hours_in_operation = 0;
else
    if time_in_operation < 5*60
        hours_in_operation = 5;
    elseif time_in_operation < 9*60
        hours_in_operation = time_in_operation / 60;
    elseif time_in_operation < 13*60
        hours_in_operation = 13;
    else
        hours_in_operation = time_in_operation / 60;
    end
end
% Calculate cost in man-hours
cost = amount_persons * hours_in_operation;

% Aggregate simulation data
accumulated_amount_lines(i) = amount_lines;
accumulated_orders_received(i) = orders_received_today;
accumulated_orders_to_handle(i) = orders_to_handle_today;
accumulated_orders_received_before_cot(i) = orders_received_before_cot;
accumulated_orders_received_after_cot(i) = orders_received_after_cot;
accumulated_orders_handled(i) = orders_handled_today;
accumulated_orders_left(i) = orders_left;
```

```
accumulated_late_orders(i) = late_orders;
accumulated_hours_in_operation(i) = hours_in_operation;
accumulated_daily_cost(i) = cost;
accumulated_orders_left_from_yesterday(i) = orders_left_from_yesterday;

end
```

Calculate totals over period

This part gathers the data for a quick overview of the performance.

```
total_orders_received = sum(accumulated_orders_received);
total_orders_handled = sum(accumulated_orders_handled);
total_late_orders = sum(accumulated_late_orders);
total_hours_in_operation = sum(accumulated_hours_in_operation);
total_cost = sum(accumulated_daily_cost);
total_late_orders_percentage = (total_late_orders/total_orders_handled)*100;

% All the results are put into a table for easy reference.
result = vertcat(total_orders_received, total_orders_handled, total_late_orders,
total_hours_in_operation, total_cost, total_late_orders_percentage);
result = table(result, 'RowNames', {'Total Orders Received', 'Total Orders Handled', 'Total
Late Orders', 'Total Hours in Operation', 'Total Cost', 'Total Late Orders Percentage'})
```

Published with MATLAB® R2014b

C.2.2 Runfile

This file sets the different planning and forecasting paramters and runs the simulation for a different forecasting mode. It saves the results to an excel file.

```
forecast_mode = 1;
for s = 0:4;
    for x = 5:14;
        for d = 5:14;
            run('simulation.m');
            save_orders_received(x,d) = total_orders_received;
            save_orders_handled(x,d) = total_orders_handled;
            save_late_orders(x,d) = total_late_orders;
            save_hours_in_op(x,d) = total_hours_in_operation;
            save_total_cost(x,d) = total_cost;
            save_late_orders_percentage(x,d) = total_late_orders_percentage;
            clear -regexp ^accumulated ^total ^collect
            clear result i idx day
        end
    end
filename1 = strcat(int2str(forecast_mode),int2str(s),'precentage_results.xlsx');
filename2 = strcat(int2str(forecast_mode),int2str(s),'cost_results.xlsx');
result_percentages = table(save_late_orders_percentage);
result_cost = table(save_total_cost);
writetable(result_percentages, filename1);
writetable(result_cost, filename2);
```

```
clear -regexp ^result ^filename ^save;
end
```

Published with MATLAB® R2014b

C.3 Forecasting Experiment

In this part of code the results from the different models are aggregated into a single table containing the performance metrics by model and forecasting horizon.

Read Data and Set Parameters

```
input1 = xlsread('OrderTotalenPerDag2.xlsx','Blad1');   %order totals
input2 = xlsread('OrderTotalenPerDag2.xlsx','Sheet2');  %day of the week
N=7;    %forecasting horizon is 7 days
training_percentage = 65;   %65% of the data is used for training
hiddenLayerSize = 5;    %set the amount of neurons in the hidden layer
```

Baseline

```
%Reading the forecast provided by the web shop from an excel file
forecast_ws = xlsread('Forecast_ws.xlsx');

%Calculate the actual order volumes for this sample of forecasts
actual_ws=input1(328:(328+length(forecast_ws)-1));

%Calculate Results
error_ws = forecast_ws - actual_ws;           %Calculate forecasting errors
results_MSE(1:N) = mean(error_ws.^2);                       %Calculate Mean Squared Error
results_MAE(1:N) = mean(abs(error_ws));                           %Calculate Root of MSE
results_MAPE(1:N) = mean(100*(abs(error_ws)./actual_ws)); %Calculate Mean Average Percentage
Error
```

Seasonal Naive Model

```
%s_ = Interval of seasonality (7 for weekly)
s=7;

%Generate a forecast using the seasonal Naïve Model
samplesize = length(input1)-(s+1);                  %Calculate Sample Size
for i = 1:samplesize
    forecast_snm(i)= input1(i);                     %Forecast is the ith value of sample of
actual values
    actual_snm(i)= input1(i+s);                     %Actual is the i+7th value of sample of
actual values
end

%Calculate Results
error_snm = forecast_snm - actual_snm;
MSE = mean(error_snm.^2);
MAE = mean(abs(error_snm));
MAPE = mean(100*(abs(error_snm)./actual_snm));
MSE(1:N) = MSE;
MAE(1:N) = MAE;
```

```
MAPE(1:N) = MAPE;
results_MSE = vertcat(results_MSE,MSE);
results_MAE = vertcat(results_MAE,MAE);
results_MAPE = vertcat(results_MAPE,MAPE);
```

Simple Moving Average

```
samplesize = round(((100-training_percentage)/100)*length(input1));
y = length(input1)-samplesize;
q = 7;

for i = 1:samplesize
    forecast_ma(i)= mean(input1((i+y-q-h):(i+y-h))); %y_fc
    actual_ma(i) = input1(y+i);      %y
end
error_ma = forecast_ma - actual_ma;
SE_ma = error_ma.^2;
MSE(h) = mean(SE_ma);
MAE(h) = mean(abs(error_ma));
MAPE(h) = mean(100*(abs(error_ma)./actual_ma));
results_MSE = vertcat(results_MSE,MSE);
results_MAE = vertcat(results_MAE,MAE);
results_MAPE = vertcat(results_MAPE,MAPE);
```

Artificial Neural Network

```
dataset = transpose(input1);
days = transpose(input2);

training_amount = round((training_percentage/100)*length(dataset));
samplesize = round(((100-training_percentage)/100)*length(dataset));


%Prepare Data
inputSeries  = con2seq(days(:,1:training_amount));
targetSeries = con2seq(dataset(1:training_amount));

inputSeriesTest  = con2seq(days(:,(training_amount+1):end-(N+1)));
targetSeriesTest = con2seq(dataset((training_amount+1):end-(N+1)));

%Generate Network Architecture
trainFcn = 'trainlm';  % Levenberg-Marquardt
delays = 1;
net = narxnet(1:delays,1:delays,hiddenLayerSize,'open',trainFcn);

%Train Network
[Xs,Xi,Ai,Ts] = preparets(net,inputSeries,{},targetSeries);
net = train(net,Xs,Ts,Xi,Ai);
net = train(net,Xs,Ts,Xi,Ai);
Y = net(Xs,Xi,Ai);
perf1 = perform(net,Ts,Y);

%Multi-step ahead prediction
netc = closeloop(net);
for i = (training_amount+1):length(dataset)-(N+7)
    inputSeriesPred  = [inputSeries(end-delays+1:end),inputSeriesTest((i-training_amount):(i-
```

```
training_amount)+(N-1))];
    targetSeriesPred = [targetSeries(end-delays+1:end), con2seq(nan(1,N))];
    [Xs,Xi,Ai,Ts] = preparets(netc,inputSeriesPred,{},targetSeriesPred);
    yPred = netc(Xs,Xi,Ai);
    result_nn((i-training_amount),:) = yPred;       %add the predictions to the result matrix
    inputSeries = horzcat(inputSeries, inputSeriesTest(i-training_amount));
    targetSeries = horzcat(targetSeries,dataset(i));       %add the actual value y to the
targetSeries
end
for i= 1:samplesize-(N+7)
    resultVal(i,:) = dataset((training_amount+i):(training_amount+(N-1)+i));
end

%Calculate Results
error_matrix = minus(resultVal,cell2mat(result_nn));
squared_error_matrix = error_matrix.^2;
MSE = mean(squared_error_matrix);
MAE = mean(abs(error_matrix));
MAPE = mean(100*(abs(error_matrix)./resultVal));
results_MSE = vertcat(results_MSE,MSE);
results_MAE = vertcat(results_MAE,MAE);
results_MAPE = vertcat(results_MAPE,MAPE);
```

Published with MATLAB® R2014b

Appendix D Literature Review on Forecasting Methods

This section will focus on the different models that exist to forecast time-series data. Firstly, de simple methods will be discussed. These contain the very basic models of forecasting like the naïve method and the simple moving average. Secondly, some classical time-series modelling techniques will be covered. The third section focusses on artificial intelligence based methods, which in the case of forecasting will mainly be an explanation of artificial neural networks or ANN. The final section will discuss some of the more advanced methods that were found in literature.

D.1 Evaluating Forecasts

To evaluate how well the predictions are the performance metrics MAE, MSE and MAPE can be used. The Mean Absolute Error (MAE) is also known as the Mean Average Deviation (MAD) and is formally defined as (Hillier & Lieberman, 2001):

$$MAE = \frac{\sum_{i=1}^{n} e_i}{n}$$

The Mean Squared Error (MSE) is almost the same as the MAE but in this case the error is squared (Hyndman & Athanasopoulos, 2013).

$$MSE = \frac{\sum_{i=1}^{n} e_i^2}{n}$$

The Mean Average Percentage Error (MAPE) gives the absolute error as a percentage of the forecasted values (Hyndman & Athanasopoulos, 2013).

$$MAPE = 100\frac{\sum_{i=1}^{n} e_i/y_i}{n}$$

All these metrics have their advantages. The MAE is the most natural measure because it returns the error in the same order of magnitude as the forecast, which makes it easy to comment on the results. This metric shall be used as the main driver of our discussion. The MAPE is powerful because it also gives information how big the percentage error is in relation to the actual values. This is a good metric to keep an eye on as well. The MSE could be better than the MAE in some applications because in this metric big errors are punished more severely than small errors. However, since the final information will be used to drive human decision making, and not some automated system, the MAE is a much more intuitive measure than the MSE. Therefore, the MSE will not be used in this report.

D.2 Simple Methods

Simple methods are the most basic methods of forecasting that exist. In this report two methods will be discussed, namely the last-value forecasting method (which shall be called the naïve method) and the moving average method.

The naïve method is the simplest forecasting method available. Forecasting with this method means taking the last recorded value of y as a forecast for y_{t+1} (Hillier & Lieberman, 2001). So:

$$\hat{y}_{t+1} = y_t$$

If seasonal effects are expected the naïve method can also be adapted to incorporate this (Hyndman & Athanasopoulos, 2013). This changes the formula to:

$$\hat{y}_{t+1} = y_{t-s}$$

Where s is the interval of seasonality. In words this means that in case of weekly seasonal effects, instead of taking today as a forecast for tomorrow, last Tuesday is taken as a forecast for next Tuesday.

Another simple forecasting method is to just take the average of the data. However, this will only work for data that has a similar mean throughout the entire dataset. As soon as trend patterns exist this method does not work anymore. It is better to only take the average of the last n data points. This way the forecast moves along with trend and seasonal developments. The simple moving average is defined as follows (Hillier & Lieberman, 2001):

$$\hat{y}_{t+1} = \sum_{i=t-n+1}^{t} \frac{y_i}{n}$$

These are all really simple methods which hardly extract any data from the historic set at all. Also, incorporating information about future developments into this model in a mathematically sound manner is not really sensible. Incorporating information about promotions can only be used by means of an educated guess whether the actual amount will surpass the forecasted amount. By how much the forecasted amount will be surpassed is guesswork as well. The next sections will cover some more advanced methods to forecast the amount of orders that will be realised, in which it is possible to incorporate information that will impact future sales.

D.3 Classical Time Series Modelling Techniques

If simple models are not enough, there are also more complicated techniques to consider. This section will try to give an overview of the more advanced forecasting techniques that are available today.

The first model to discuss is the Auto Regressive or AR(p) model. This model is defined mathematically as (Hyndman & Athanasopoulos, 2013):

$$\hat{y}_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + \varepsilon_t$$

In plain text this model states that the forecast for y_t equals a combination of p linear functions of the previous values of y, a constant c and an unobserved white noise error function. So this model is basically a linear combination of the previous values. The parameters are a measure of how much the forecast depends on its previous values and can be estimated by using least squares regression amongst others (Franses, van Dijk, & Opschoor, 2014).

Another model in the linear category is the Moving Average or MA(q) model. Where the AR(p) model is a linear correlation of the previous values of y, the MA(q) model is a linear combination of the unobserved errors of y which can be stated as:

$$\hat{y}_t = \mu + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \cdots + \theta_p\varepsilon_{t-q}$$

Even though it looks as though this model does not depend on previous values of y it can be rewritten to show that it does (Franses, van Dijk, & Opschoor, 2014).

When combining these two models, the ARMA(p, q) model is created. Obviously, an ARMA($p, 0$) model equals an AR(p) model and an ARMA($0, q$) model equals a MA(q) model. With these models it is possible to describe a dataset more accurately than with the other approaches. However, these types of models are only applicable when the dataset is stationary. This means that the mean and variance are constant over time, or mathematically (Franses, van Dijk, & Opschoor, 2014):

$$E[y_t] = \mu$$

$$E[(y_t - \mu)^2] = \gamma_0$$

$$E[(y_t - \mu)(y_{t-k} - \mu)] = \gamma_k$$

For all $t = 0,1,2 \ldots T$ and all $k = \ldots, -2, -1, 0, 12, \ldots$. In essence, this means that the data cannot contain any trends or seasonality. To overcome this problem the ARIMA(p, d, q) model was introduced. The I in this case stands for Integration. The ARIMA model is exactly the same as the ARMA model, with the difference that the dataset is differentiated d times to eliminate any trend effects. By doing this the ARIMA model is applicable to trend data, but not yet to seasonal data. It is possible to adjust the ARIMA model to accommodate seasonal data. For this purpose, several seasonal terms need to be added to the model (Hyndman & Athanasopoulos, 2013). Doing so makes the ARIMA model applicable to a more wide range of problems and it is a widely popular approach to forecasting time series (Hillier & Lieberman, 2001).

When a model is selected, what is selected is a certain polynomial function of which the parameters still need to be estimated from the data. This can be done through certain methodologies of which the Box-Jenkins method turns up in every text on basic forecasting examined for this report (Franses, van Dijk, & Opschoor, 2014) (Hillier & Lieberman, 2001) (Hyndman & Athanasopoulos, 2013). The Box-Jenkins method is an iterative process to fit an ARIMA model to the data. First the auto-correlation function (ACF) and partial auto-correlation function (PACF) of the data are plotted. Based on these values, a selection of models with the associated parameters is selected. With these models, predictions are made and compared to the known data. If the residuals (forecasting errors) behave like white noise functions with $\mu = 0$ the process is complete. This is because if the residuals are not white noise functions there is still data left to extract. The Box-Jenkins method is complicated (Hillier & Lieberman, 2001) but doable with modern computer software.

Up until know models were covered to forecast data that contains seasonality and trends. However, when periods with different variances occur in the data, these models cannot be used. For this type of data, the Auto-Regressive Conditional Heteroskedasticity (ARCH) model was invented. The general ARCH(q) model is given by (Franses, van Dijk, & Opschoor, 2014):

$$h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \alpha_q \varepsilon_{t-q}^2$$

Just like the ARIMA model, the ARCH model can be fitted to the data as well. However, the methods to do so seem to be less developed than the Box-Jenkins method, since it is way harder to find proper documentation on this.

D.4 Artificial Neural Networks

Up until now all methods are based on manually extracting information from a dataset. This means that patterns need to be found in datasets where little of the underlying mechanisms are known. However, modern research is moving more in the direction of artificial intelligence. Pattern recognition algorithms are potentially way better in finding patterns in data that has no apparent relation. For time-series a widely researched topic is artificial neural networks. Information for this section is largely taken from the lecture notes from the course SC4081 Knowledge Based Control Systems (Babuška, 2010). Other material will be cited when used.

Artificial Neural Networks (ANN) are based on synapses in the human brain. A synapse sends out an impulse if the threshold value exceeds a certain value. In an ANN artificial neurons (Figure 32) consist of a nonlinear activation function $\sigma(z)$. Whenever the sum of weighted inputs exceeds the activation function of the artificial neuron an output is given. The activation functions can be all kinds of functions. The most used are threshold, piece-wise linear and sigmoidal functions.

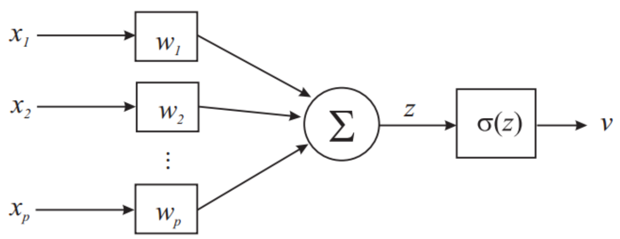


Figure 32: An artificial neuron (Babuška, 2010)

With these artificial neurons a network can be constructed. This network will consist of an input layer, one or more hidden layers containing multiple artificial neurons and an output layer. The connection between the different layers are weighted. In this way a network can be constructed which is able to approximate all kinds of (nonlinear) functions. The approximation of the function is done by adjusting the different weights based on known in- and outputs of the network using an error minimizing optimization algorithm. An operation which is known as training the network.

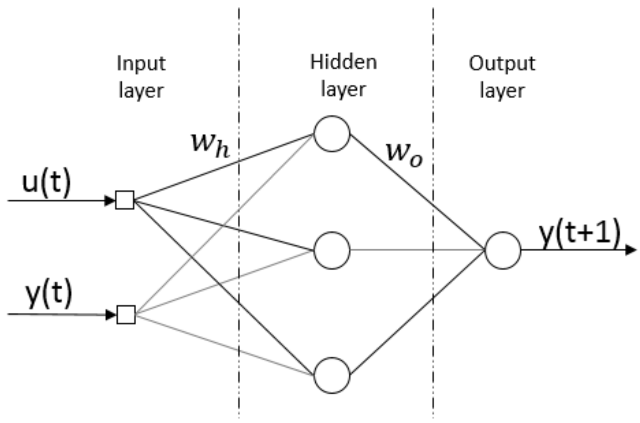


Figure 33: An Artificial Neural Network with one hidden layer of artificial neurons

This training operation starts by feeding the network with known inputs and outputs. How that works in the time series case is as follows. Every day the network produces an estimate for the next day $\hat{y}_{t+1}$ based on today y_t. The estimate $\hat{y}_{t+1}$ is compared to the actual value y_{t+1}. This produces an error which will be quite large the first time. The optimization algorithm is designed to minimize the error by adjusting the weights of the network. For multiple steps the network will keep on producing an estimate and compare this to the actual value while the weights will keep on being adjusted. In this way the forecasting error is minimized step by step. So by adjusting the weights of the network a forecasting model will be created and all relations between the future and the past are extracted from the data.

Obviously, it would be way better to add other known relational data to the network as well. For example, if the network would only consider the order amounts, no seasonal relations will ever be found. This is where u_t comes in. Through u_t all other information that could have an impact can be inserted into the model. For example, if a function is added to u_t that takes the value 1-7 for every day of the week, the network will automatically use this information in the forecast. If there are weekly seasonality effects present, the optimization algorithm will set the weights to scale the values 1-7 in such a way that they introduce these effects in the output. If no weekly seasonality effects are present, the optimization algorithm will gradually set the weights to zero and the information will be ignored.

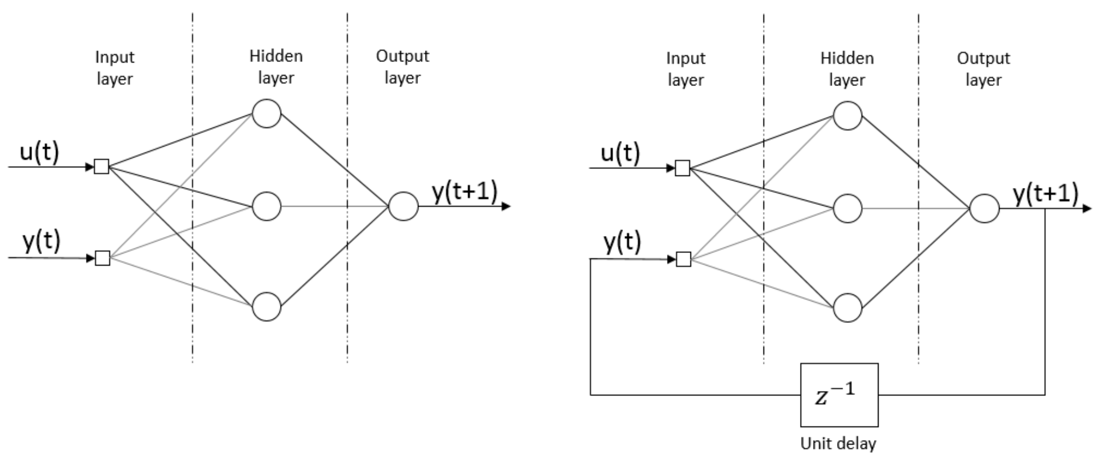


Figure 34: The feedforward state used for training (left) and the closed-loop state used for prediction (right)

So instead of finding a model to fit a dataset, the ANN approach gradually learns a model from the dataset, giving more weight to the information that is important and less weight to the information that does not matter. By closing the loop as depicted in Figure 34, this model can be used to make predictions several steps in the future. In this way the amount of orders can also be forecasted several days into the future building along on the same weights using one forecast as an input for the next.

The last topic considering ANN that still needs to be discussed is the configuration. For the ANN all kind of different configurations are possible which have influence on how well the ANN performs in minimizing the forecasting error. Basically, there are three parameters that can be set for an ANN. Firstly, the amount of hidden layers can be set. For simple applications, one is enough, but multiple layers can be used as the function grows in complexity. Secondly, the amount of neurons in the layers can be adjusted. More neurons will enhance the fit up until some point, but too many neurons will cause the network to over fit the data, increasing the forecasting error. Lastly, the amount of delays can be adjusted. This impacts whether the ANN uses data from this day or also data from yesterday and even further back to forecast the amount of orders that comes in tomorrow.

So how well will these type of networks be able to forecast the data? A feedforward neural net with at least one hidden layer with sigmoidal activation functions can approximate any continuous nonlinear function $R^p \to R^n$ arbitrarily well on a compact set, provided that sufficient neurons are available (Cybenko, 1989). However, ANN has a tendency to over fit the data as the amount of neurons grows too large (Khashei & Bijari, 2011) (Wong, Xia, & Chu, 2010).

D.5 Further Literature Research

In recent literature, a lot can be found on improving forecasting accuracy by combining some models discussed in the previous section. At first traditional time-series methods and ANN are only compared (Tang, de Almeida, & Fishwick, 1991) (Ho, Xie, & Goh, 2002), but later they are also combined to form hybrid models which are claimed to improve accuracy even further (Khashei & Bijari, 2011). However, due to the complexity, this is out of scope for this research.