Development of a Cost Price Model to Support Bid Price Estimation

UNIVERSITY
OF TWENTE.

Bachelor Thesis

Industrial Engineering and Management

Development of a Cost Price Model to Support Bid Price Estimation

July 2021

Author:

B.G. Benneker (Bram)

b.g.benneker@student.utwente.nl

Supervisors University of Twente:

Dr. P.C. Schuur (Peter)

B. Gerrits, MSc (Berry)

Supervisor SCS Logistics:

E. Damink (Erik)

University of Twente

Drienerlolaan 5

7522 NB Enschede

(053) 489 9111

Preface

For the completion of my Bachelor's degree in Industrial Engineering and Management, I am doing my Bachelor assignment at SCS Logistics in Ootmarsum. The thesis is about developing a cost price model to support bid price estimation. During the whole process of carrying out my Bachelor assignment, I received help and support from various people. Therefore I would like express my gratitude towards these people.

First, I would like to thank all colleagues at SCS Logistics for helping me during my graduation project and for making my time at SCS Logistics so enjoyable. In particular I would like to thank Erik Damink for giving me the opportunity to execute my Bachelor assignment at SCS Logistics and the good guidance during the process. Besides that I would like to thank SCS Logistics that I was allowed to work at the office during the COVID-19 pandemic.

Moreover, a special thanks to my supervisor Peter Schuur, who guided me during the research. I would like to thank you for the enjoyable but above all instructive meetings. Your enthusiasm and critical view enabled me to take my Bachelor Thesis to a higher level. I would also like to thank Berry Gerrits for being willing to join as second supervisor and the time you invested in this project. On top of this, I want to thank Ipek Seyran Topan for her guidance during the preparation phase of thesis.

Finally, I would like to express my gratitude to my family and friends for their support during the execution of my Bachelor assignment. I want to thank Matthijs ten Cate in particular. He gave me good advice and feedback and kept me motivated during graduation.

Management Summary

This research has been performed at SCS Logistics in Ootmarsum. SCS Logistics is a freight forwarder that focuses on transport and logistics that deviates from the 'day to day' logistics. The market in which SCS Logistics operates is very competitive. An appropriate bid price and the speed with which the bid price is issued are therefore essential for winning customers. However, the time pressure is often high, which sometimes leads to inaccurate bid prices being issued. That is why there is still room for improvement at SCS Logistics in terms of speed and accuracy when determining the bid prices. Currently, the employees of SCS Logistics have no guidance and support when they have to determine the bid prices. This is because there is no standard procedure. To solve this problem, this research focuses on developing a cost price model that offers guidance in determining the bid prices. Based on the purchase prices calculated by this cost price model, the employees can create a bid price. Therefore, the main research question addressed in this thesis is formulated as follows:

How to give the employees of SCS Logistics a standard procedure in determining the bid price?

In order to answer this research question, we first analysed the current situation of the purchasing and sales process. By executing interviews with the planners and an observation study of the TMS, we were able to create better insights into the process. Because we provided insight into the current situation, we were able to better understand where the problems and challenges are within the process. Based on this, the desired situation has been mapped out. Using the model of the desired situation, it becomes clear where the solution must be implemented. Both the current and desired situation were visualized using business process models. These models were the first step towards solving the core problem. Ultimately, we concluded from the desired situation that estimating the purchase price must be automated in the TMS using a cost price model.

After the desired situation was known, the research focused on the cost price model for calculating the purchase price. In order to determine the purchase price, it is important to properly map out the cost drivers. With the use of interviews and an observational study of the TMS, we determined, classified and described the relevant cost drivers. Ultimately, we decided to divide the cost drivers into three categories. These categories are: shipment details, service level and additional costs.

To develop the cost price model, we needed to know which cost estimation methods we were going to use for calculating the purchase price. By executing a literature study, we found several cost estimation methods. Subsequently, we determined the strengths and weaknesses of the various methods using the following four criteria: accuracy, simplicity, customizability and speed. Based on the data availability of the cost drivers and how the methods score on each criterion, we chose the relevant cost estimation methods. The *case-based reasoning method* scored the highest. That is why it was chosen as the framework for the cost price model. In addition, the case-based reasoning process is reinforced with the methods *regression analysis models* and *expert systems*.

Because both the case-based reasoning and regression analysis models require input data we created a final data set from initial raw data. The initial raw data was retrieved from the database of SCS Logistics and consists of historical transactions. In order to create the final data set we removed all transactions with missing values and the transactions with inconsistent values have been corrected or removed. This is because it is important that the input data is of good quality. Furthermore, to get familiar with the data set and discover initial insights, we did an exploratory data analysis.

After the final data set was created, we developed the cost price model that forms the basis for determining the bid prices. The three methods chosen from the literature review were combined and merged into a cost price model. In order to make the cost price model usable for the employees, we implemented the model in a *user-friendly tool*. In this way, we were able to test the cost price model. We programmed the tool using *Microsoft Excel* and *Visual Basic for Applications*. The cost price model

with the tool ultimately formed the solution to the core problem. The left side of figure 0 shows the home page of the tool. When the employees press the ‘New Order’ button, the ‘Enter Order’ form pops up, which can be seen on the right of Figure 0. In this form the employees can fill in the shipment details, after which they can calculate the purchase price and save the order in the database.

Figure 0: Tool of the Cost Price Model

We tested the performance of the cost price model with selected forecast error measurements from the literature and a survey completed by four involving stakeholders. The accuracy of the calculated purchase prices of the cost price model has been assessed using new incoming data. The results are that the MAE = 15.15 and the RMSE = 20.70. This means that the average deviation between the calculated purchase price of the tool and the actual purchase price is approximately €15.15. And when the outliers are weighted more heavily, the deviation is €20.70. Besides that the relative deviation of the cost price model is 7.9%. Furthermore, we found that for the 33 test transactions, the tool more often quoted a purchase price that was too high than too low. Lastly, we concluded that employees should be extra alert for shipments to and from the top of the Netherlands, since the calculated purchase prices of the cost price model often deviate slightly more from the actual purchase price for these type shipments.

According to the outcomes of the survey, the deviation of the cost price model is accurate enough for determining the bid prices at SCS Logistics. The employees found the calculated purchase prices issued by the tool a good target price. It was also concluded from the survey that the employees create bid prices faster and more confidently when using the cost price model. Therefore, they would like to have the cost price model implemented in the TMS. In that case, the model can provide a good indication and support in determining the bid prices.

Based on the performed research and regarding the implementation of the cost price model in the TMS, recommendations are made to SCS Logistics. The main recommendations are as follows:

- We recommend to rearrange the service levels, so that all employees know exactly which shipment belongs to which service level.
- When this cost price model is going to be used for other countries and/or service levels, it is very important to keep in mind that the transactions in the current database contain a lot of errors. Therefore, we propose to remove all outliers per country or to start over by filling in the database correctly.

- When more transactions become available, we recommend to divide the countries into smaller zones. This makes the calculated freight rates more accurate.

Lastly, we recommend using the calculated purchase prices of the cost price model as an indication and a guideline. It is important that the employees always have a critical look at the purchase price calculated with the tool.

For the successful integration of the cost price model within SCS Logistics, we have compiled a roadmap with various actions and associated persons who are responsible. The roadmap is shown in Table 0.

Table 0: A Roadmap for Integrating the Cost Price Model

Order	Actions	Actor(s)
1.	First of all, it must be ensured that the new orders are entered correctly in the TMS. It is therefore important that all employees are aware of the consequences if they fill in the order incorrectly in the TMS and the benefits if they fill it in correctly. The TMS must also alert the employees if they enter something incorrectly or not at all. This can be done, for example, with the help of a message box.	CEO, Me, Software developer
2.	Implement the developed cost price model in the TMS.	CEO, Me, Head of the operations departments, Software developer
3.	When the cost price model has been implemented in the system, a clear manual must be made. This manual should be about the cost price model and the new adjustments that have been made in the TMS.	Me
4.	Evaluate after 10 weeks and improve possible errors within the cost price model.	CEO, Employees, Me

When the orders are eventually entered correctly in the system and the cost price model has been successfully integrated into the TMS, SCS Logistics can focus on expanding and making the cost price model more accurate. This can be done in the following ways:

- Expand the tool with new countries and service levels.
- Make the zones of the countries smaller when possible.
- Use neural networks to weight seasonal periods and holidays. In addition, implement a check box in the TMS, so that purchase prices can increase during certain seasonal periods and holidays.
- Determine the purchase price of different suppliers with the help of neural networks. Based on this, we already have an estimated purchase price for various suppliers that is close to the actual purchase price.

Contents

List of Figures

List of Tables

Reader's Guide

We briefly introduce the eight chapters of the research at SCS Logistics.

Chapter 1: Introduction
In this chapter an introduction to the research is given. The company will be introduced shortly. In addition, the problem identification is described and the research methodology explained.

Chapter 2: Current and Desired Situation
The current and desired situation are explored in the second chapter. Both situations are visualized using business process models. These models help to better understand the problems and challenges within the process, and where the solution must be implemented.

Chapter 3: Cost Drivers for Bid Prices
This chapter maps out the cost drivers that determine the bid prices. These cost drivers are determined, classified and described

Chapter 4: Cost Estimation: Literature Study
The literature study to find relevant cost estimation methods is described in the fourth chapter. To develop the cost price model, we needed to know which cost estimation methods we were going to use for calculating the purchase price. Therefore different cost estimation methods are described and evaluated in order to arrive at the best methods that are suitable for SCS Logistics.

Chapter 5: Data Exploration and Preparation
In the fifth chapter a final data set is created from initial raw data. In order to create the final data set we removed all transactions with missing values and the transactions with inconsistent values have been corrected and removed. After that the data of the historical transactions is explored.

Chapter 6: Design and Development of Cost Price Model and Tool
The design and development of the cost price model and corresponding tool is described in Chapter 6. The tool is used to make the cost price model usable for the employees. How the case-based reasoning, regression analysis and expert systems are executed and combined within the cost price model is explained.

Chapter 7: Evaluation, Implementation and Recommendation
In this chapter we evaluated on the cost price model and tool, and provided the implementation and corresponding recommendations. Forecast error measurements are used to assess the accuracy of the cost price model. In addition we performed a survey for five stakeholders at SCS Logistics to asses the tool for long term use. Based on the results of the survey and interviews with the employees we described how the cost price model should be implemented and gave recommendations regarding this implementation.

Chapter 8: Conclusion, Discussion and Future Research
Conclusions and limitations of the performed research are given in this last chapter. Furthermore, potential future research is explained.

List of Acronyms

ADR Accord européen relatif au transport international des marchandises Dangereuses par Route

BPNN Back-Propagation Neural-Network

BPMN Business Process Model and Notation

CBR Case-based Reasoning

CR Cost Revenue

CRM Convention Relative au Contrat de Transport International de Marchandises par Route

DSRM Design Science Research Methodology

IMO International Maritime Organisation

LDM Loading Meter

MAE Mean Absolute Error

MSE Mean-Squared Error

RRSE Root Relative Squared Error

RSE Relative Squared Error

RAE Relative Absolute Error

RMSE Root Mean-Squared Error

TMS Transportation Management System

1 Introduction

For the completion of my Bachelor's degree in Industrial Engineering and Management, I am doing my Bachelor assignment at SCS Logistics. The focus point of this research is the development of a cost price model that supports the bid price estimation at SCS Logistics. By designing a user-friendly tool, the cost price model is made usable for the employees of SCS Logistics. First of all Section 1.1 introduces the reader to the company. In addition, Section 1.2 gives the motivation for the research. Section 1.3 contains the problem identification. Then, in Section 1.4 the key constructs are described. In Section 1.5 the problem-solving approach, with the research design is discussed. And Section 1.6 reflects on reliability and validity.

1.1 The Company: SCS Logistics

SCS Logistics is located in Ootmarsum and focuses on transport and logistics that deviates from the standard type of transport. This are transport and logistics activities that require flexibility, transparency, control, dedication and know-how, resulting in a high service standard. SCS Logistics offers services for the transport of a machine, project cargo, products with abnormal measurements (out of gauge cargo), cargo to be shipped to a remote location, urgent shipments or anything else which is different from the 'day to day' logistics. SCS Logistics is a freight forwarder, which belongs to the category of 4PL providers. This means that it manages transport flows, but does not have its own vehicles or a physical warehouse. Therefore it uses external parties. The below pyramid in Figure 1, represents one of the most common ways of looking at and defining the role of a 4PL provider.

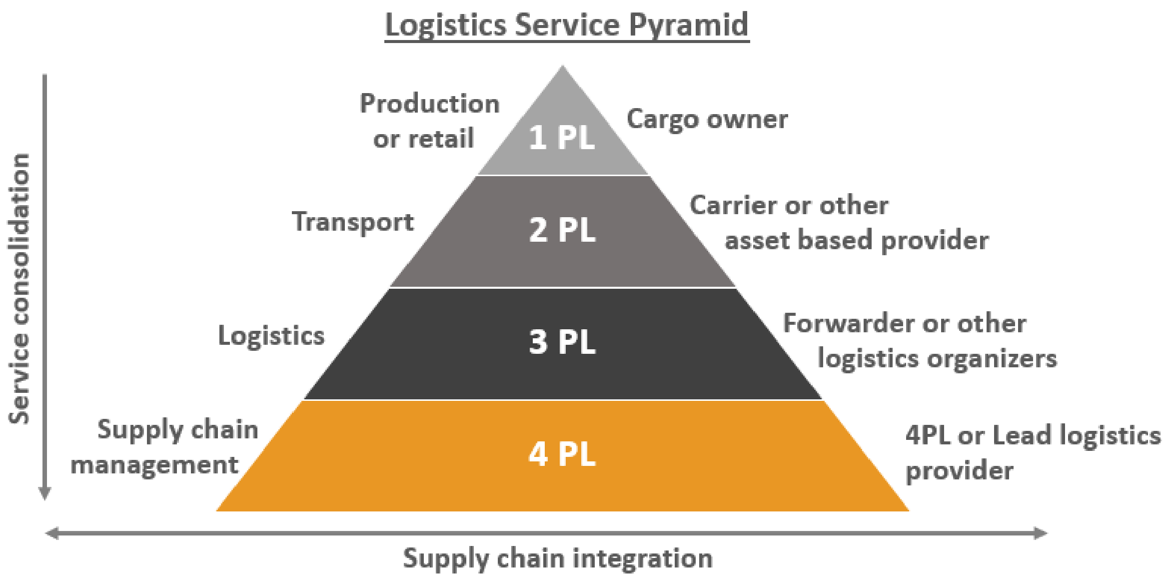


Figure 1: Logistics Service Pyramid (Allyn International, n.d.)

Within SCS Logistics I will mainly focus on the optimization of the purchasing and sales process.

1.2 Research motivation

The market in which SCS Logistics operates is very competitive. An appropriate bid price and the speed with which the bid price is issued are therefore essential for winning customers. However, issuing an accurate bid price can sometimes take a lot of time. In addition, time pressure is often high, which means that prices are not always given accurately. According to the CEO of SCS Logistics, there is still much room for improvement in terms of speed and accuracy.

Therefore, the goal of the research is:

Helping the employees of SCS Logistics, so that they can determine the bid price faster and more accurately.

1.3 The problem

This section investigates the problems SCS Logistics encounters. Section 1.3.1 describes the problems at SCS Logistics. Section 1.3.2 displays the problem cluster of the described problems. And Section 1.3.3 provides the core problem and the motivation behind the choice.

1.3.1 Problem Context

The purchasing and sales process of SCS Logistics is currently not working optimal. The process faces a number of bottlenecks. These bottlenecks cause speed and accuracy of the purchasing and sales process to decline. The bottlenecks encountered are discussed in this section.

Within the purchasing and sales process of SCS Logistics there is no standard procedure for determining the bid prices. All planners estimate the bid prices manually based on feeling and experience. This causes that the bid price is very sensitive to errors and that the bid price calculation differs per employee. In addition, the prices of the suppliers can fluctuate. These fluctuations are caused, for example by exchange rate fluctuations or seasonal products. For instance, there is the melon season in Spain. In this season, import loads from Spain are more expensive, because many trucks from Spain are full of fruit. The trucks coming from Spain are then more scarce. The planners must be alert to all these changes, because the fluctuations in prices of suppliers can be overlooked, leading to an incorrect bid price. Furthermore, the customers' requirements can also change during the process, as a result of which the issued prices are no longer correct. The bottom line is that determining the bid price depends on many factors. And because there is currently no standard procedure for determining the bid price, errors and incorrect quotations arise. This ultimately causes that there are sometimes low or no profits or even losses.

Besides the fact that the bid prices are sensitive to errors, it sometimes also happens that determining the price takes a long time. Currently everything is done manually. To determine the bid price, different price lists from different suppliers or historical transactions are compared with each other in order to arrive at the cheapest and/or most suitable bid price for the customer. This process is time consuming and sensitive to errors. In addition, information about the prices of some suppliers may not be known. These must first be requested. Requesting prices takes a lot of time, which means that the customer has to wait. The longer the customer has to wait, the greater the chance that this customer will go to the competitor. Especially the complex quotations take a lot of time. Complex quotations are quotations for orders with many different loads and many different sizes. A lot of different information and therefore research work is required for these quotations.

To conclude the two action problems are:

- Sometimes low or no profits or even losses
- Sometimes it takes a long time to determine the bid prices

1.3.2 Problem Cluster

To bring order to the problem context in Section 1.3.1, and to identify the core problem a problem cluster is made. All the problems in the problem cluster have been indicated by executing interviews with different stakeholders. The problems are mapped along with their connections in Figure 2, which gives insight in the causes and effects.

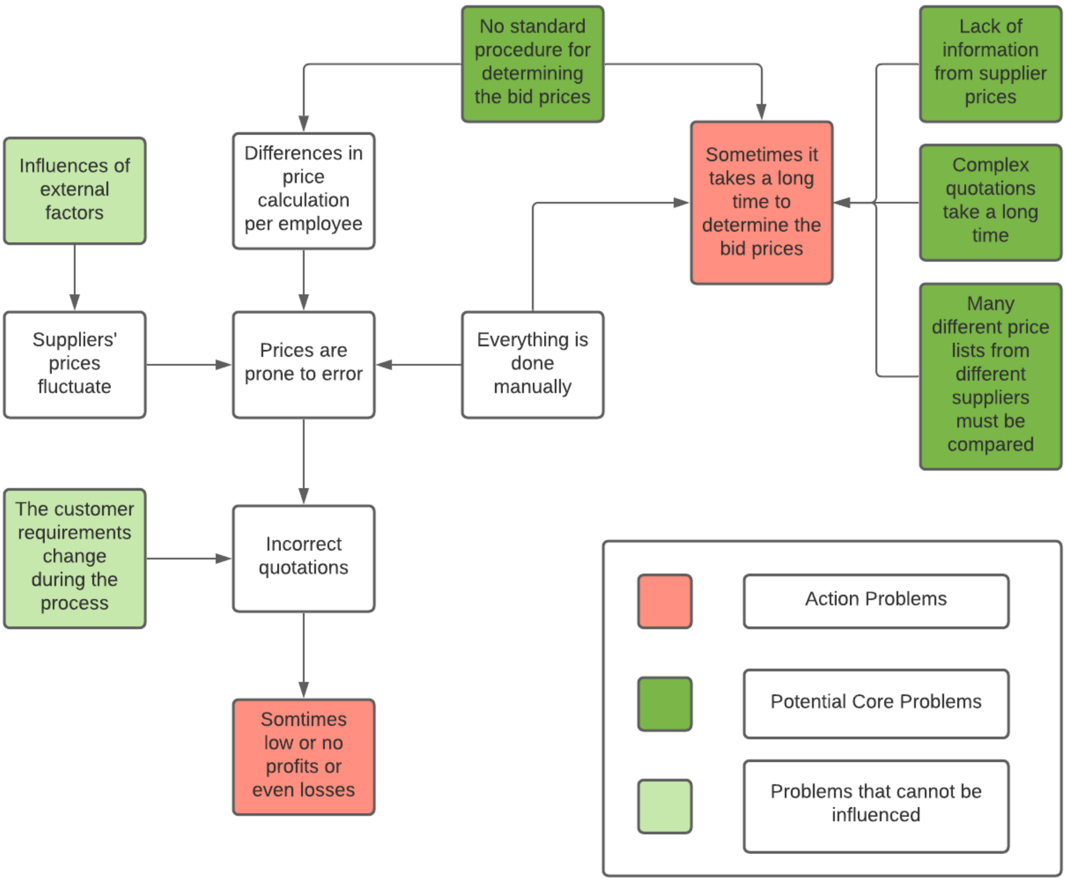


Figure 2: Problem Cluster of the Purchasing and Sales Process SCS Logistics

1.3.3 Core Problem

When looking at the problem cluster, there are six potential core problems, which do not have direct causes. The six potential core problems are shown in green in Figure 2. A core problem can never be a problem that you cannot influence (Heerkens & van Winden, 2017). The problems 'influences of external factors' and 'the customer requirements change during the process' cannot be influenced. Therefore these problems cannot be a core problem.

There are now four potential core problems left. When more than one potential core problem in the cluster remains, the most important one should be solved. The most important problem is the one whose solution would have the greatest impact effect at the lowest cost (Heerkens & van Winden, 2017).

The two action problems are: 'sometimes low or no profits or even losses' and 'sometimes it takes a long time to determine the bid prices' and are given in red in Figure 2. Ultimately it comes to two things that the price must be determined more accurately and faster. The problem cluster actually only represents one problem that has an impact on both of these starting problems. This problem is 'no standard procedure for determining the bid prices'. The solution to this problem therefore has the greatest impact. When there is a standard procedure for determining the bid prices, prices become more accurate and profits become higher. In addition, it takes less time to determine the prices, which increases customer satisfaction and more customers can be served per employee.

Another factor that plays a role in choosing the core problem is that the ultimate goal SCS Logistics strives for is a customer portal for regular customers. In that portal, the customers then receive a price based on their requirements. Partly for this reason, a standard procedure for determining the bid prices is very important, because it is the basis for the customer portal.

To conclude, the following core problem has been selected:

The employees of SCS Logistics have no standard procedure for determining the bid prices

To express the core problem in norm and reality, we state that in reality (the current situation) the employees of SCS Logistics have no standard procedure for determining the bid prices and the norm (the desired situation) is that there must be a cost price model for the employees of SCS Logistics that offers guidance in determining the bid prices. Based on the purchase prices calculated by this cost price mode, the employees can create a bid price.

1.4 Key Constructs

For clarifying the direction of the research it is important to elaborate on the key constructs related to the research. The key constructs discussed in this section are: freight forwarder, 4PL provider, cost estimation, and business process modelling.

Freight Forwarder
SCS Logistic is a freight forwarder. Freight forwarders are the link between the owner of the goods and the carrier, and provide forwarding or clearing services. They play the role of the intermediary in the transportation sector (Saeed, 2013).

4PL Provider
SCS Logistics is a freight forwarder, which belongs to the category of 4PL providers. According to Bade and Mueller (as cited in Büyüközkan et al., 2009) a fourth party logistics (4PL) is: "a supply chain integrator who assembles and manages the resources, capabilities, and technology of its own organization with those of complementary service providers to deliver a comprehensive supply chain solution". In other words, 4PL providers manage transport flows, but does not have its own vehicles or physical warehouses. Therefore it uses external parties.

Cost estimation
In order to estimate the cost price of SCS Logistics cost estimation methods are used. Cost estimation is: "the process of predicting the costs required to perform the work within the scope of the project" (Elfaki et al. 2014).

Business process modelling
The current and desired purchasing and sales process at SCS Logistics will be visualized with business process modelling. Business process modelling produces descriptions of business processes that are central to organizations as they enable its specification, documentation, analysis, and engineering through multiple paradigms, languages and techniques (Caetano et al., 2012).

1.5 Problem Solving Approach

In this section the problem solving approach is explained. Section 1.5.1 describes the methodological framework. Section 1.5.2 shows the corresponding knowledge questions. Section 1.5.3 displays the research design. And Section 1.5.4 describes the main deliverables.

1.5.1 Methodological Framework

This section describes the methodological framework that will be used and how it fits within this research. The Design Science Research Methodology (DSRM) will be used to conduct the research.

Design science is the creation and evaluation of artifacts, with the intention to solve identified organizational problems (Peffers et al., 2007). According to Peffers et al. (2007) the artifacts are "any designed object with an embedded solution to an understood research problem". Figure 3 visualizes the process model with the six phases of the Design Science Research Methodology.

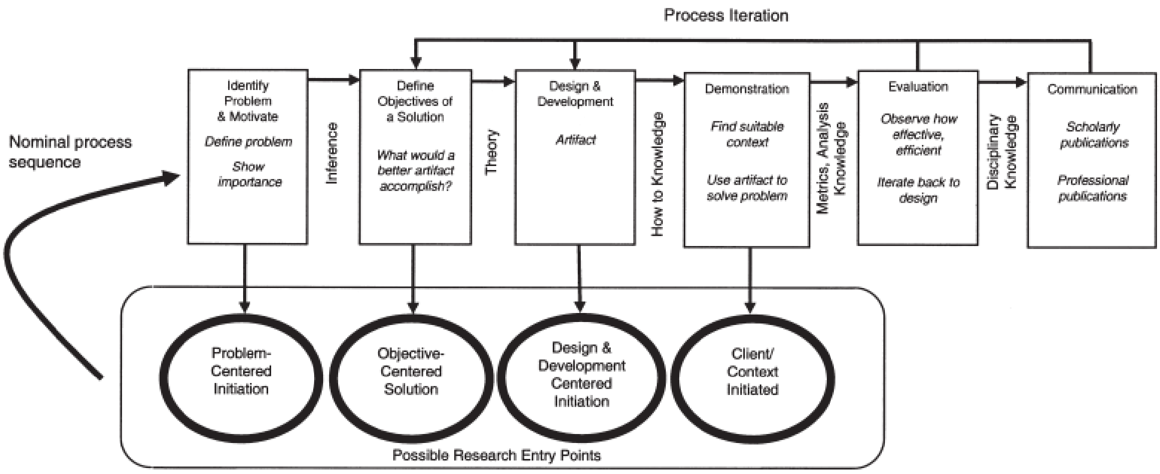


Figure 3: Process Model of the Design Science Research Methodology (Peffers et al., 2007)

In the next section, the six phases with the corresponding knowledge questions are discussed.

1.5.2 Research Questions

The core problem stated in Section 1.3.3 lead to the main research question of this thesis:

How to give the employees of SCS Logistics a standard procedure in determining the bid price?

In order to answer the main research question, research is conducted. The six phases of the DSRM, with the associated knowledge questions that are presented below, form the step-by-step approach for answering the main research question.

Phase 1: Problem identification and motivation

1. What does the current and desired purchasing and sales process at SCS Logistics look like?

In this phase the problem is identified. This is done in Section 1.3. The core problems are identified with the help of the problem cluster. After that we will determine and motivate on which core problem we will focus on. Moreover, the current and desired situation will be mapped out in order to answer the first knowledge question. The current and desired purchasing and sales process will be visualized with a business process model.

Phase 2: Define the objectives for a solution

2. Which cost drivers determine the bid prices of SCS Logistics?
3. What are relevant methods for estimating the cost prices for SCS Logistics?

Peffers et al. (2007) describe this phase as follows: "infer the objectives of a solution from the problem definition and knowledge of what is possible and feasible". First of all, by means of interviews with the employees and an observation study of the TMS, the cost drivers that determine the final bid prices of SCS Logistics are determined, classified and described. When the cost drivers are known, different cost estimation methods are identified by conducting a systematic literature review. After the various cost estimation methods have been clearly described, it is determined which method(s) are best applicable within the company. These method(s) form the basis for the solution/artifact to be developed.

Phase 3: Design and Development

4. What is the value of the available data of SCS Logistics?
5. How should the cost prices be calculated?
6. How should the cost price model be implemented at SCS Logistics?

In the design and development phase, a solution to the core problem is developed with the knowledge of the cost drivers, literature and the desired situation. First of all, the value of the data of the historical transactions is considered. Furthermore, an approach must be developed for determining the cost prices. This approach is based on the answers of the first four knowledge questions. Finally, when the approach for determining the cost prices has been developed, a user-friendly tool must be designed for the employees of SCS Logistics.

Phase 4: Demonstration

7. How can we test the performance of the solution?

When the design and development phase is finished, a way must be developed to test the performance of the solution. The way of testing answers the seventh knowledge question.

Phase 5: Evaluation

8. What recommendations can be given to SCS regarding the solution and its implementation?

This eighth knowledge question is answered by means of a survey. Here the opinions of the employees about the functioning of the tool are very important. Employees can try out the new cost price determination tool. After trying the tool, the employees of SCS Logistics complete the survey I made. This survey ensures that the opinion of the employees about the new tool and thus the impact of the tool is made clear. Furthermore, with the help of the results of the survey, the performance of the tool, evaluation, recommendations, limitations, conclusions, and future research will be explained.

Phase 6: Communication

In the communication phase, the tool, the entire research and the recommendations are communicated to SCS Logistics, so that they can continue to work with the developed tool. Moreover this will also be presented during the public defence and the report of the research will be published as well. A schematic overview of the activities performed for every knowledge question can be found below, in Table 1 under Section 1.5.3, Research Design.

1.5.3 Research Design

Table 1: Research Design

How to give the employees of SCS Logistics a standard procedure in determining the bid price?							
Knowledge Problem	**Type of Research**	**Research Population**	**Subjects**	**Research Strategy**	**Method of data gathering**	**Method of data processing**	**Activity Plan**
1. How does the current and desired purchasing and sales process at SCS Logistics look like?	Descriptive	SCS Logistics	Employees, Suppliers, Customers	Deep Qualitative	Semi-structured interviews, Observation (cross-sectional)	Visual representation using BPMN, Qualitative	Interviewing employees → observation of the TMS → gathering overview with business process model
2. Which cost drivers determine the bid prices of SCS Logistics?	Explanatory	SCS Logistics	Employees Suppliers Customers	Deep Qualitative	Semi-structured interviews and Observation (cross-sectional)	List of cost drivers, Qualitative	Interviewing employees → observation of the TMS → overview of cost drivers → classification of cost drivers
3. What are relevant methods for estimating the cost prices for SCS Logistics?	Descriptive	Literature	Database, Researchers	Broad Qualitative	Literature study (cross-sectional)	Descriptive text about the methods, Qualitative	Research → literature study → overview of methods → choosing the best methods
4. What is the value of the available data of SCS Logistics?	Descriptive	Software system – Financial department	Dataset	Deep Quantitative	Data can be retrieved from the database of TMS (cross-sectional)	Data analyses, Graphs and Tables, Statistical analyses, Qualitative and Quantitative	Data gathering → Data preparation → Data analyses → Conclude
5. How should the cost prices be calculated?	Descriptive	Operations department	Employees, Dataset	Deep Quantitative	Using the answers of the first 4 knowledge questions + the dataset, Semi-structured interviews (longitudinal)	Development of an approach to determine the cost prices, with mathematical equations; Quantitative	Combine answers of the first 4 knowledge questions → Use dataset → Interview employees →Develop approach and mathematical models
6. How should the cost price model be implemented at SCS Logistics?	Descriptive	Operations department	Employees	Deep Qualitative	Semi-structured interview, Observation (cross-sectional)	List of requirements	Interviewing employees → Observation of Planning department and TMS→ Design a solution
7. How can we test the performance of the solution?	Descriptive	Operations department	Cost price calculation tool, Employees	Deep Quantitative	Outcomes retrieved from developed calculation model, Semi-structured interviews, Experiment (longitudinal)	Statistical analyses of outcomes + Description of outcomes	Gather new (test) data → Gather outcomes of the tool → Compare Outcomes → Analyse outcomes → Conclude
8. What recommendations can be given to SCS regarding the solution and its implementation?	Descriptive	Operations department	Employees, Cost price calculation tool	Deep Quantitative	Survey and Observation (cross-sectional)	Table of outcomes, Evaluation story	Making survey → Hand out survey → Gathering outcomes → Observation of current TMS → Evaluate → Make recommendations

Limitation Research Design

To solve the core problem the research design in Table 1 is used. But when using this research design it is important to state the limitations of the research. The following three limitations have been taken into account:

- **Time constraints**
  There are 10 weeks available for the project. A solution must be found within these 10 weeks. If the solution has not been worked out perfectly within 10 weeks or the solution has not yet been applied, advice for a future study will have to be mentioned.

- **Lack of previous research studies on the topic**
  Not much literature can be found on determining the bid prices for 4PL providers. It has therefore been decided to compare different cost estimation methods within different sectors. In order to arrive at suitable cost estimation methods that are applicable to situation in which SCS Logistics operates.

- **Feeling and emotion cannot be included in a tool**
  The feeling, emotion and experience that the employees of SCS Logistics have cannot be included in a tool. It is therefore important to give the employees a suggested purchase price. This price must always be checked by the employees and adjusted if necessary. In this way, prices can be adjusted to the external conditions. This limitation should be properly communicated to the employees of SCS Logistics.

1.5.4 Deliverables

This report will deliver the following main deliverables:

1. Business process model of both the current and desired purchase and sales process
2. Cost price model and tool that supports bid price estimation for SCS Logistics
3. Implementation and recommendations

1.6 Validity and Reliability

In this section the validity and reliability of the research is discussed. Reliability means, that when the research, with the same methods, is conducted at a later date the results should be the same. In other words, reliability is concerned with the stability of the research results. Validity is concerned with the contents. Are you measuring what you want to measure? There are many potential threats that can compromise the validity of the research (Heerkens & van Winden, 2017). The validity will be discussed with internal and external validity.

Internal validity is concerned with whether the research design is set up to measure what is intended to measure (Heerkens & van Winden, 2017). The biggest threat to internal validity in this research is the historical data. The data will be discussed extensively with the employees to make sure that the dataset is properly displayed and all units are correct. In addition, if necessary, the system can be used to provide insight into how a certain value has been determined. Finally, the results of the calculations with the data will be extensively discussed with the employees. In this way the internal validity is guaranteed.

External validity concerns about whether the research can applied to other groups than your research population (Heerkens & van Winden, 2017). Although the research is focused on the situation of SCS Logisitcs, the research could be used for the same type of companies. That is why the way of thinking and the approach must be well documented in a report. In this way, the research can be carried out within other companies or situations.

Reliability concerns about whether the results are the same, if the same research is conducted at a later date. This is a very difficult one, because interviews are conducted with the employees and a survey will be filled out by the employees. The data gathered through interviews and a survey contains the opinion of the employees, which can change over time. Furthermore, the results of the cost price model only remain the same when the input data remains the same.

2 Current and Desired Situation

This chapter contains a description of the current and desired situation. By creating business process models, we visualize both the current situation and the desired situation. In Section 2.1 we map out the current situation. After which we map out the desired situation in Section 2.2. Finally, the conclusion is described in Section 2.3. The main input for mapping the process is provided by conducting interviews with the planners and an observation study of the TMS (Transportation Management System).

2.1 Mapping the Current Situation

This section describes the current purchasing and sales process with the help of a business process model. We create these models after the interviews and the observation study are done. In order to arrive at a suitable solution for the problem, it is important that the process in which the problem occurs is clearly mapped out. The business process model of the current situation is shown in Figure 4. Section 2.1.1 describes the current sales process and Section 2.1.2 the current purchasing process.

2.1.1 Current Sales Process

We start with the sales process. The sales process takes place at the beginning of the business process model, and can be seen on the left side of Figure 4. The sales process is the process in which the activities focus on the sale of transportation and logistics services. The ultimate goal of the sales process is to secure as many transport orders as possible. This is achieved by issuing tailor-made quotations.

The sales activities start with an incoming transportation request from the customer. If the customer is not an existing customer, a credit check is used to check whether the customer is financially reliable. If the customer is not financially reliable, no further actions are taken and the process stops. If the customer is financially reliable, this new customer is created in the system. Existing customers are already known in the system. Subsequently, the customer’s transport requirements are analysed. The wishes can consist of various criteria, for example: number of loading meters, loading/unloading date, transit time, fixed loading or unloading time, ADR, truck-mounted forklift, customs documents, etc. These transport details are included when offering the quotation. The prices vary per request. Currently, all planners estimate purchase prices manually based on feeling and experience. After that, they based the bid price on the estimated purchase price. The bid price consists of the purchase price plus a profit margin. When the bid price has been determined, the quotation is sent to the customer. Based on the customer's response to the quote, the sales process can end in two ways. If the customer does not agree with the quotation, the entire process stops. If the customer does agree with the quotation, the sales process stops and the purchasing process starts.

2.1.2 Current Purchasing Process

The purchasing process starts after the sales process is completed and can be seen on the right side of Figure 4. The purchasing process is the process in which the activities focus on the procurement of transportation and logistics services. The ultimate goal of the purchasing process is to select a suitable carrier for the customer’s request.

The output of the sales process is the information for the first step of the purchasing process. The information obtained, which is an agreement on the quotation with the associated transport details, is analysed. Subsequently, it is checked which carrier is suitable for the customer order. The carrier with the best price, conditions, quality and reliability is selected to carry out the customer order. When there is an agreement between SCS Logistics and the carrier, this is confirmed with a transport order. This order is created in the TMS. When the order is created, it is stored in the database of SCS Logistics.

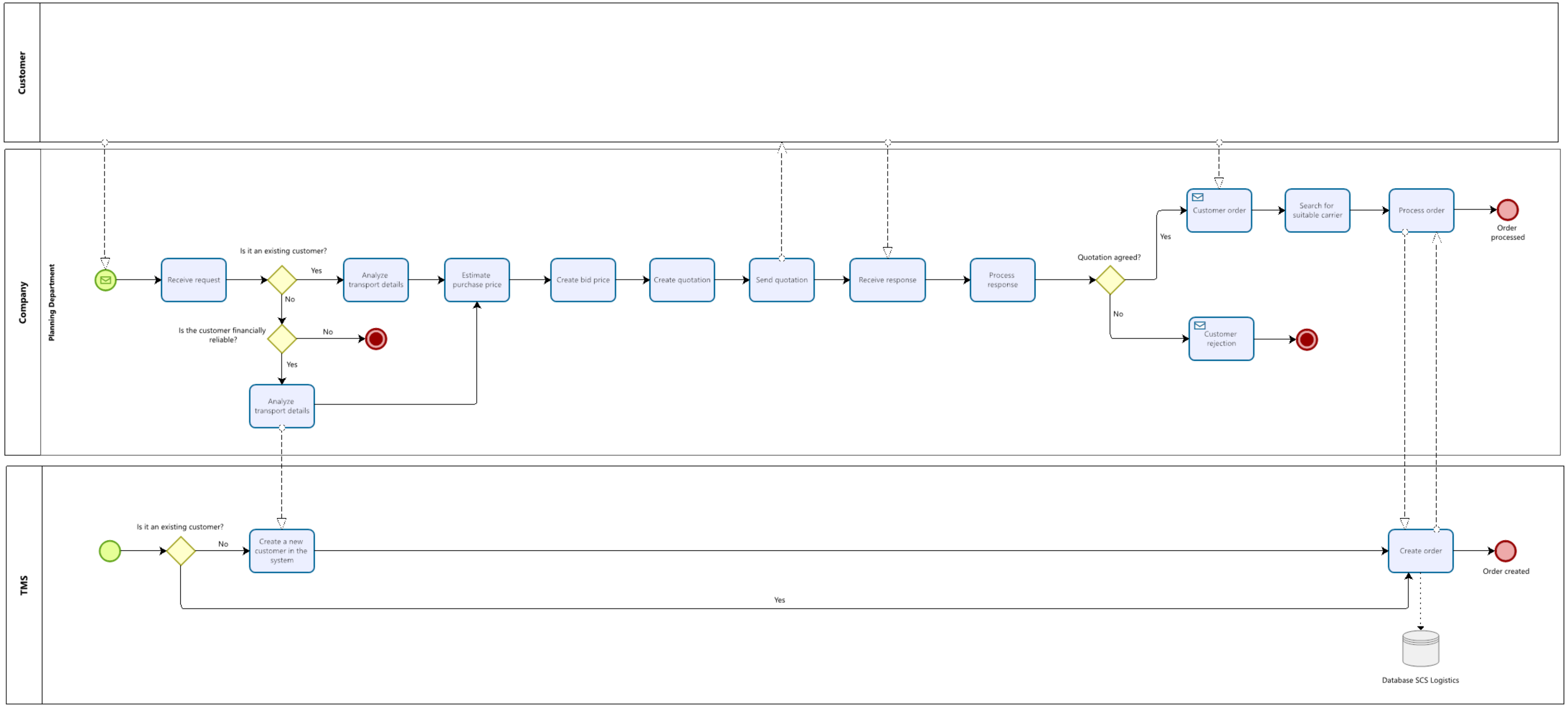


Figure 4: Current Purchasing and Sales Process

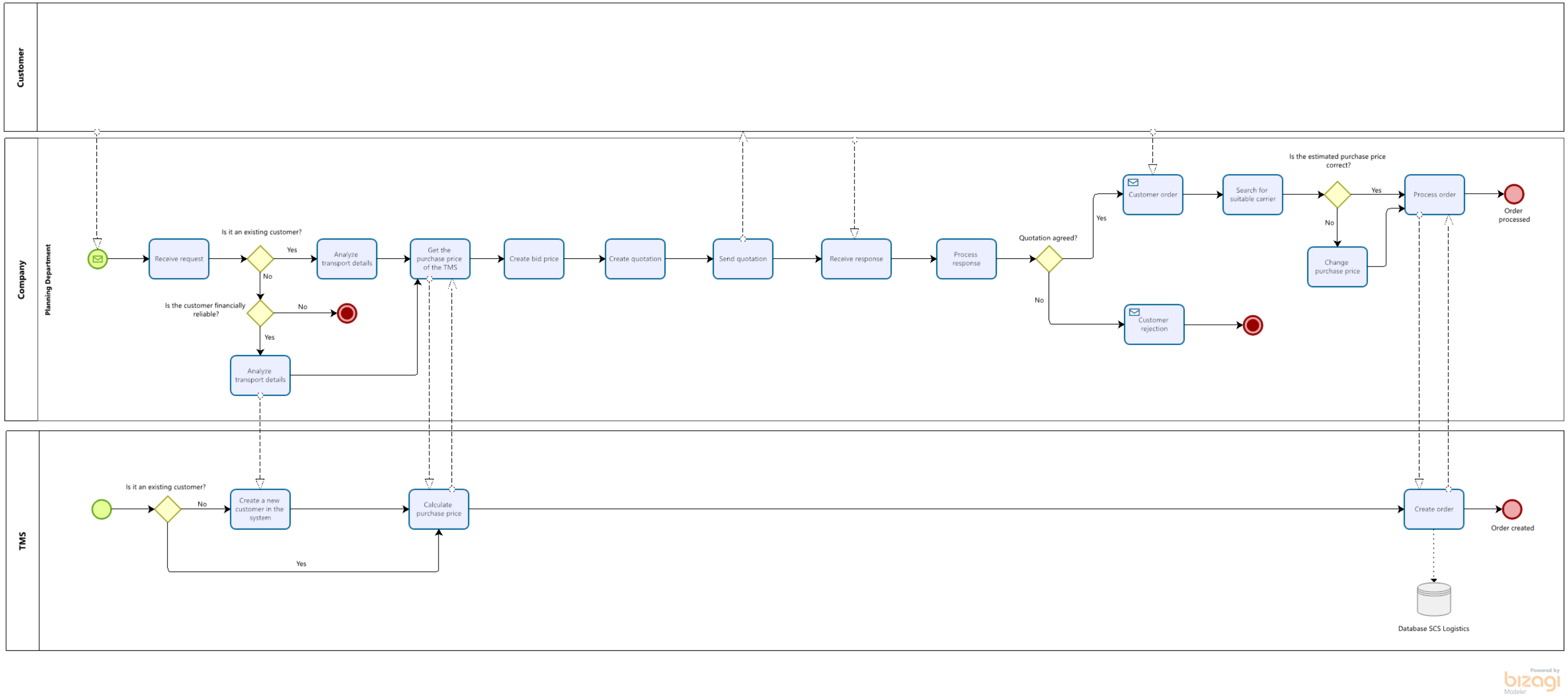


Figure 5: Desired Purchasing and Sales Process

2.2 Mapping the Desired Situation

Now that the current situation has been mapped out, this section describes the desired purchasing and sales process with the help of a business process model. We create this model after conducting interviews and an observation study of the TMS. The business process model of the desired situation is shown in Figure 5.

As can be seen in Figure 5, the estimation of the purchase price within the sales process is improved. The purchase price is now not only estimated on the basis of feeling and experience, but also with the support of a *cost price model*. This model is based on past transactions. When the transport details of the customer have been analysed, they are entered in the TMS. The transport details are the input for the cost price model. The cost price model is integrated within the TMS. This allows the employees of SCS Logistics to easily calculate the purchase price with the cost price model. After the cost price model has calculated a purchase price, the bid price can be created. Later, when the quotation is approved by the customer and a suitable supplier is selected, the actual purchase price is known. It may be possible that the actual purchase price differs from the estimated purchase price. In that case, the planner can adjust the estimated purchase price, after which he or she can process the order. In this way, the accuracy of the estimated purchase prices is improved and ensured. However, if the estimated purchase price is exactly the same as the actual purchase price, the order can be processed immediately.

How the cost price model in the TMS works is discussed in Chapter 6. In addition, a process overview of the cost price model is shown in Figure A.1 in Appendix A.

2.3 Conclusion

By creating business process models, we were able to get a better understanding of the current and desired purchasing and sales process at SCS Logistics. We determined the business process models by executing interviews with the planners and an observation study of the TMS. Because the current situation is now clearly visualized, we understand better where the problems and challenges are within the process. Based on this, the desired situation has been mapped out. Using the business process model of the desired situation, it becomes clear where the solution must be implemented. These visualizations are the first step towards solving the core problem. Ultimately, estimating the purchase price within the sales process will have to be automated in the TMS, using a *cost price model*. This conclusion will be taken into account in the remaining research.

3 Cost Drivers for Bid Prices

Now that the desired situation is known, the research focuses on the cost price model for calculating the purchase price. In order to determine the purchase price, it is important to properly map out the cost drivers. Cost drivers are factors which have a cause-effect relationship with costs. The cost drivers are any factors which cause a change in the costs of work performed in an organisation or process (Bokor, 2010). In this chapter the cost drivers are determined, classified and described. The cost drivers for the bid prices have been determined by conducting interviews with the employees of SCS Logistics and an observation study of the TMS. The cost drivers are divided into three types, which are shown in the following sections:

- 3.1 *Shipment Details*
- 3.2 *Service Level*
- 3.3 *Additional Costs*

After the cost drivers of each type have been described, we discuss in Section 3.4 with which cost type each type of cost driver is related. The discussion and conclusion are described in Section 3.5.

3.1 Shipment details

Shipment details are the specifications of the shipment. The following cost drivers belong to the shipment details category:

- **Loading and unloading address**
  The loading address is the address where the cargo is loaded. The unloading address is the address where the cargo is unloaded. The location of both the loading and unloading address is considered in the bid price determination. The distance between the loading and unloading address and the accessibility of the addresses are then considered.

- **Loading meter (LDM)**
  Loading meter is a standardized measure in road transport. A loading meter is 1 meter in length from the loading space of a truck. The formula for calculating the loading meters is:

$$(Length\ x\ Width)\ /\ 2.4\ (in\ meters) \tag{1}$$

  The 2.4 is the standard width of a truck in meters. The loading meter of a standard truck is 13.6.

  <u>Calculation example:</u>

  You want to transport a EUR-pallet. The standard dimensions of a EUR-pallet are:
  Length: 1.20 meter
  Width: 0.80 meter
  So in total: (1.20 * 0.80) / 2.4 = 0.4 loading meter.

- **Weight (in tons)**
  This is the actual weight (in tons) of the cargo that is being transported. The maximum payload of a standard truck is 24 tons. One loading meter may weigh a maximum of 1.75 tons. When the freight weighs more than 1.75 tons per loading meter, more loading meters will be charged. Only then the loading meters are calculated using the following equation:

$$Loading\ meters\ to\ be\ paid = loading\ meters\ of\ freight * \frac{weight(in\ tons)}{1.75} \tag{2}$$

Calculation example:

You want to transport a machine with the following dimensions and weight:
Length: 2.40 meter
Width: 1.00 meter
Loading meter: (2.40 * 1.00) / 2.4 = 1 LDM
Weight: 2.8 tons
But because the maximum weight of a loading meter is 1.75 tons;
1 * (2.8 / 1.75) = 1.6 LDM is charged.

3.2 Service Level

This section is about the different service levels that are available at SCS Logistics. The service level identifies the type of shipment. In Section 3.2.1 the service levels are described. And Section 3.2.2 explains the relation between the service levels and costs.

3.2.1 Different Service Levels

At SCS Logistics, a distinction is made between the following service levels:

- **Standard**
  This means that the chosen service level is not special. The cargo has no specifics or extra additions. The standard service level includes both full truck load and groupage shipments. In general, these are the standard shipments within Europe. For standard shipments only tautliners are used, which can be found in Figure 6.

Figure 6: Tautliner (Three Horses Trailer, n.d.)

- **Express**
  These are shipments where the transit time is important. The transit times of express shipments are faster than those of standard shipments. Therefore the load of the truck only consists of the freight to be carried. For this cargo, mainly vans and double-manned cars are used.

- **Exceptional**
  Exceptional shipments are shipments with different dimensions and/or specifications. Exceptional includes all shipments that are driven with special transport. An example of an exceptional shipment is the transport of a dragline machine. This machine does not fit in a normal truck. That is why a suitable vehicle is selected based on the dimensions of the product to be transported. Vehicles for special transport are:

- Flatbed trailer

Figure 7: Flatbed Trailer (TIP Trailer Services, n.d.)

- Lowbed trailer

Figure 8: Lowbed Trailer (Nooteboom, n.d.)

- Semi-Lowbed trailer

Figure 9: Semi-Lowbed Trailer (Pacton, n.d.)

- Extendable trailer
- Crane car
- Container chassis

Figure 10: Container Chassis (TIP Trailer Services, n.d.)

- Other vehicles for special transport

- **Temperature control**
  With temperature control shipments, the temperature of the product to be transported must be taken into account. The temperature of the cargo must remain within certain values.

- **Container**
  These are all sea freight shipments that are transported between two ports. With these types of shipments, the before and after transport is sometimes included in the price.

- **Air freight**
  These are all air freight shipments that are transported between two airports. With these types of shipments, the before and after transport is sometimes included in the price.

- **Other service level**
  All other shipments that do not belong to the above-mentioned types fall under the category ‘other service level’.

3.2.2 Relation between Service Level and Costs

The relationship between service level and costs can be expressed using ranking. The three modalities are ranked from most expensive to least expensive:

1. Air freight
2. Road transport
3. Container

All of these three modalities have the following sub-rank, which is ranked from most expensive to least expensive:

1. Exceptional
2. Express
3. Temperature control
4. Standard

In general when all other conditions are the same, the relation between the service level and the costs can be ranked like this, but there can always be exceptions. The category other service level cannot be related to the costs since the properties of these shipments are not known. These properties differ a lot per shipment.

3.3 Additional Costs

Additional costs are all extra additions for a certain shipment. The following cost drivers fall under the additional costs:

- **ADR**
  ADR is the abbreviation of the French title of the European Convention on the International Carriage of Dangerous Goods by Road: "Accord européen relatif au transport international des marchandises Dangereuses par Route". So, ADR includes the transport of dangerous goods. The ADR specifications are examined for each shipment. With an ADR value of more than 1000 points, an ADR certified driver with an ADR equipped truck must be deployed. Values below 1000 points do not require an ADR driver and ADR equipped truck, but an adapted fire extinguisher must be present (no extra costs). If the shipment is transported over water, an IMO certificate is required in both cases, which sometimes entails additional costs.

- **Tail lift and pallet truck**
  If the truck used has a tail lift and pallet truck at, at least one of the loading or unloading addresses, additional cost will be charged.

- **Truck-mounted forklift**
  If the truck used has a truck-mounted forklift at, at least one of the loading or unloading addresses, additional costs will be charged.

- **Double manned**
  If the truck has two drivers, additional costs will be charged.

- **Waiting times**
  These are the times the driver is waiting. Normally, groupage shipments take 30 minutes for loading and 30 minutes for unloading. With full truck load shipments, 2 hours are taken for loading and 2 hours for unloading. These waiting times are included in the price. But when a truck has a longer loading or unloading time, additional costs are charged.

- **Administration costs**
  Additional administrative actions are charged.

- **Cancelation costs**
  These are the costs for cancelling the freight. Customers can cancel free of charge 24 hours in advance, unless costs have already been incurred for the transport. When the freight is cancelled on the day of loading, 70% of the freight price will be charged.

- **Customs and transport documents**
  These are additional costs for the preparation of customs and transport documents.

- **Additional (un)loading address**
  This can be a customs stop, but also an extra loading and/or unloading address. The amount of the costs depends on whether the location is on the route or not.

- **Fixed times at which the cargo must be (un)loaded**
  If the cargo has to be loaded or unloaded at a fixed time, additional costs will be charged

- **Storage charges**
  There are additional costs when the cargo has to be stored in a warehouse. The handling costs within the warehouse also fall under storage charges. For example, a forklift driver must usually be present to transfer the load.

- **Permits**
  There is an additional charge when permits must be applied for.

- **Transport insurance**
  All customers are insured under the CRM conditions (these costs are included in the freight price). But when the customer wants to insure the freight extra, additional costs are charged.

- **Other additional costs**
  The other additional costs that do not fall under any of the above categories.

3.4 Relation between Types of Cost Drivers and Types of Costs

The total costs of a shipment at SCS Logistics consist of the following two cost types:

- The freight rate
- Additional costs

The total cost of a shipment is calculated using the following equation:

$$Total\ Costs\ =\ freight\ rate\ +\ additional\ costs \tag{3}$$

The freight rate is determined by the cost drivers of the following two types:

- Shipment details
- Service level

The additional costs are determined by the cost drivers of the following type:

- Additional costs

3.5 Discussion and Conclusion

In the sections above we described all relevant cost drivers. The cost drivers ‘other service level’ and ‘other additional costs’ are described very generally. A discussion can be made whether these cost drivers should be divided into more specifically described cost drivers. For the cost driver ‘other service level’ this discussion point is included in the recommendations given in Chapter 7.2.2. But with the cost driver ‘other additional costs’, a broad concept has been deliberately chosen. This is because these costs are uncommon and can be anything. They may be costs that have never occurred before and it is therefore difficult to describe all these costs separately. For simplicity and convenience, it was decided to describe these costs as ‘other additional costs’. Furthermore, there is something to criticize about the calculation of the loading meters in Section 3.1. Because this calculation is only used for freights with standardized dimensions. When the freight has a non-standardized dimension, the number of loading meters is not calculated using Equation 1. In such a situation, the planners of SCS Logistics issue a matching loading meter. This is because non-standardised dimensions of the freight can lead to lost space in the truck. And this lost space cannot be filled by other shipments. The last point of discussion is the ranked service levels as discussed in section 3.2.2. When all circumstances are the same, this ranking can be maintained. But there can always be exceptions. For example, it can happen that with a standard shipment there is space left in a tautliner that cannot be used. This unused space will be charged. As a result, it can happen that a standard shipment is more expensive than, for example, a temperature control shipment.

With the use of interviews and an observational study of the TMS, it was ultimately decided to divide the cost drivers into three categories. The categories are: *shipments details*, *service level* and *additional costs*. The cost drivers within these categories and the service levels have been discussed, after which we explained the relation between the types of cost drivers and the types of cost. Ultimately, we use the cost drivers for the choice of the cost estimation method in Chapter 4, the preparation of the data set in Chapter 5 and the development of the cost price model in Chapter 6. Furthermore, in Chapter 6 the relationships between the cost price and the cost drivers are determined and the cost drivers are quantified.

4 Cost Estimation: Literature Study

Now that the relevant cost drivers for the bid prices are determined, we have to conduct a literature study to find relevant cost estimation methods, that are applicable in the situation of SCS Logistics. Different cost estimation methods are described and evaluated in order to arrive at the best methods that are suitable for SCS Logistics. In Section 4.1 the theoretical perspective for the literature study is given. Section 4.2 classifies the various cost estimation methods. The categories that fit in the situation of SCS Logistics are discussed in Section 4.3. Section 4.4 shows intuitive cost estimation methods. And Section 4.5 shows analogical cost estimation methods. In Section 4.6 the choice between the methods is discussed. Finally the conclusion is described in Section 4.7.

4.1 Theoretical Perspective

To determine the bid price, it is very important to know what the cost price is. The cost price has the most influential impact on determining the price of a product or service (Huang et al., 2012). When the costs are properly estimated, competitive and accurate bid prices can be issued. This is crucial in the very competitive market in which SCS Logistics operates. Hence, it is important for SCS Logistics to adapt and apply appropriate cost estimating methods. Cost estimation methods help predict the costs for a particular product, project or service. Within this literature study, different cost estimation methods with their strength and weaknesses in various domains, for example in aerospace and construction areas, has been examined in order to find the best methods that are applicable to the situation of SCS Logistics.

4.2 The Classification of Cost Estimation Methods

Niazi et al. (2006) developed a comprehensive hierarchical classification of the different cost estimation methods. The classification divides the methods with similar features into various categories and can be found in Figure 11. The product cost estimation methods are first divided between qualitative and quantitative methods. The qualitative methods investigate whether a new product or service has similarities with previously sold products or services. These similarities can then be used to determine the cost of the new product or service (Huang et al., 2012). This saves time and effort as the costs do not have to be determined from scratch. The requirements for implementing qualitative methods are historical design and manufacturing data, and/or previous cost estimating experience from experts (Huang et al., 2012). The qualitative cost estimation methods can be further categorised into intuitive and analogical methods, which are discussed in detail in sections 4.3 and 4.4. In contrast to the qualitative methods, the quantitative methods are based on a detailed analysis of a product design, its features, and corresponding manufacturing processes instead of simply relying on the past data or knowledge of an estimator (Niazi et al., 2006). In quantitative cost estimation methods, costs are estimated using a mathematical formula, taking different product or resources parameters during whole product life cycle into account (Huang et al., 2012). The quantitative cost estimation methods can be further categorised into parametric and analytical methods.

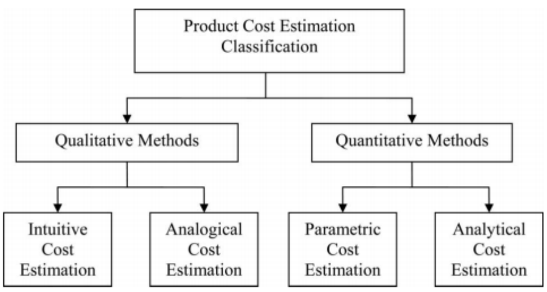


Figure 11: The Classification of Cost Estimation Methods (Huang et al., 2012)

4.3 The Categories that fit in the Situation of SCS Logistics

To determine which types of cost estimating methods fit in the situation of SCS Logistics it is important to know the strengths and weaknesses, and the aim of the categories into which the methods are divided. The quantitative estimation methods focus on a detailed analysis of a product design, its features, and corresponding manufacturing processes. This approach would generally obtain more accurate costing results than the qualitative method (Huang et al., 2012).

This is because the qualitative method estimates costs based on historical data and/or previous experience of an expert. All costs involved in producing a product or providing a service are not determined in detail and are not divided over different layers. The advantage of qualitative methods is that they help to obtain quick and rough estimates during the estimation of the cost prices. Hence, qualitative estimation methods are better suited to a situation in which the estimating time is limited, past data and/or experts' knowledge is available, and the estimating accuracy requirement is moderate (Huang et al., 2012). In contrast, in situations where the estimating time is sufficient, relationships of different cost variables are identified, and the cost estimation accuracy need to be high, the quantitative methods are preferable (Huang et al., 2012).

In the competitive market in which SCS Logistics operates, time in which the bid price is issued is very important. SCS Logistics strives to issue a quotation within 30 minutes after request. Furthermore, they do not have their own trucks or warehouses and are dependent on the sales price from different transport companies. That is why SCS Logistics cannot map out in detail all the costs of the services they purchase for selling their own services. The only data available at SCS Logistics is historical data from past transactions and expert knowledge. In addition, SCS Logistics sometimes has more than 180 requests per day. It is therefore not possible to determine the costs in detail for each request, because this takes too much time and the detailed cost data is not available. Since speed is a strong quality of qualitative methods and the cost price can be determined fairly accurately, qualitative methods are chosen. The cost estimation methods under this category are discussed below.

4.4 Intuitive Cost Estimation Methods

Intuitive methods relies on expert experiences. The experience can use their knowledge directly or rules can be established to build expert systems that store and computerize expertise and automate heavy calculations (Salmi et al., 2016). Intuitive cost estimation methods can be categorized into case-based reasoning and decision support systems, which are now discussed.

4.4.1 Case-Based Reasoning

Case-based reasoning solves a new problem by searching for a prior similar situation and by reusing information and knowledge of that situation (Hu et al., 2016). This approach assumes that similar products have similar costs and estimates the cost of that new product based on the cost of the past cases (Huang et al., 2012). The case base reasoning process is shown in Figure 12. The main process of case-based reasoning consists of retrieving, reusing, revising, and retaining. The retrieving process is retrieving past cases in order to solve new cases. Similar past cases are retrieved with the help of key attributes (Elmousalami, 2021). When missing data occur it has to be collected and implemented. This improving process has to be carried out until the design conforms the specifications (Busachi et al., 2017). In the reusing process, the old cases are used to determine the new cases. The revising process evaluates the suggested solution to the problem. When the current case and the past cases are sufficiently similar the solution can be used directly. Otherwise the difference between the new and retrieved cases should be taken into account and the solution should be revised (Hu et al., 2016). Finally, in the retaining the existing case-base will be updated with the new case. According to Huang et al. (2012) the main advantage is that the design approach of case-based reasoning tends to be very creative. In addition, the disadvantage is that it relies on past data and previous experience.

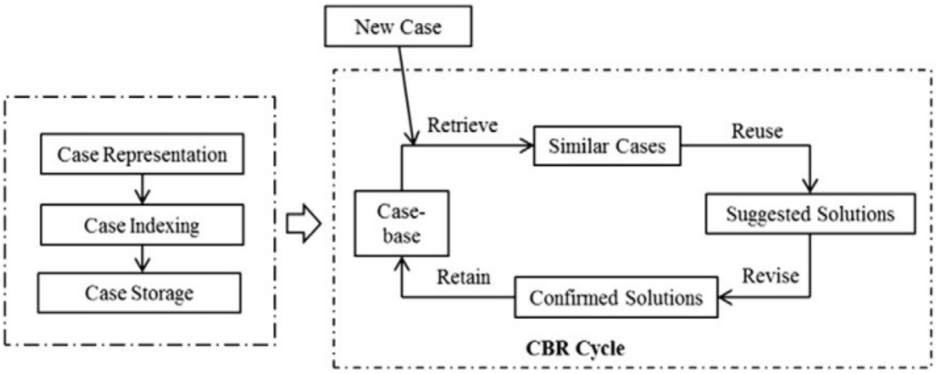


Figure 12: Case-Based Reasoning Process (Hu et al., 2016)

4.4.2 Decision Support Systems

Decision support systems are used to help cost estimators in selecting the most appropriate design solution by utilising information from the knowledge database (Huang et al., 2012). Niazi et al. (2006) defines three main categories: *Fuzzy-Logic Approach*, *Rule-Based systems* and *Expert Systems*. According to Huang et al. (2012) the *fuzzy logic approach* is not applicable to services, since it is designed for using in design and production. The *fuzzy logic system* focuses on the estimation of tangible features, therefore this method will not be discussed in further detail. *Rule-based systems* establish a set of design and/or manufacturing constraints to give the manufacture a guide in selecting the appropriate manufacturing process. Then on the basis of the process time and the cost calculations of the different processes the price of the product can be determined (Niazi et al., 2006). Also *rule based systems* might be very difficult, time consuming to adopt for service cost estimation, and for some types of services not applicable. According to Niazi et al. (2006) *expert systems* are based on storing knowledge of experts in a database and manipulating it on demand to infer quicker, more consistent, and more accurate results, through rule-based programming. *Expert systems* are ideal for minimising human errors and different results between estimators. However the drawback is that this technique often requires complex programming (Huang et al., 2012).

4.5 Analogical Cost Estimation Methods

Analogical methods estimates the proposed project, by comparing it with already developed projects that has the same requirements. The achieve this, data is extracted from the previous projects and compared with the proposed project (Saleem et al., 2019). Analogical cost estimation methods can be categorized into regression analysis models and back-propagation neural-network, which are now discussed.

4.5.1 Regression Analysis Models

Regression models uses historical cost data to forecast the cost of a new design, by adjusting the regression parameters to map a mathematical relationship based on the given data (Elmousalami, 2020). The goal of regression is to investigate the contribution of each variable to the overall cost (Busachi et al., 2017). Therefore, the major advantage of regression analysis models is the ability to interpret a relation between variables and costs. In addition, regression methods are relative easy to use (Huang et al., 2012), but the drawbacks are that these methods are sensitive to outliers (Yu Hao et al. 2020) and related to data availability.

4.5.2 Back-Propagation Neural-Network (BPNN) Models

As an alternative to regression models neural network can be used. According to Boehm et al., (2000) neural networks are estimation models that can be trained using historical data to produce ever better results by automatically adjusting their algorithmic parameter values to reduce the delta between known actuals and model predictions. Neural networks are able to deal with nonlinearity issues and detailed cost information is not necessary. The most common type of neural networks is the back-propagation neural network (BPNN) and it suits best the product cost estimation, according to Niazi et al. (2006). The limitation of the BPNN is that the accuracy and performance highly depends on the data collected. If the data in not complete or inaccurate, the network is unable to generate suitable algorithms to train the model for cost estimations. For an accurate model big data and expertise is required and this requires a long processing time (Yu Hao et al., 2020). Besides that neural networks have a high establishment cost according to Niazi et al. (2006).

4.6 Method Comparison and Choice

After identifying all relevant cost estimation methods we have to evaluate and choose which methods are suitable to the situation of SCS Logistics. In Section 3.4 we identified the two types of costs that make up the total cost of a shipment and their relation with the cost drivers. The two cost types are: *freight rate* and *additional costs*. The possible cost estimation methods for the freight rate and additional costs are determined in Section 4.6.1 and 4.6.2. This selection is based on the data availability. After the possible cost estimation methods are known, the correct method(s) are chosen based on the required criteria from SCS Logistics. These criteria were obtained through interviews and are:

- *Accuracy*: the method should be as accurate as possible
- *Simplicity*: the method should be easy to use
- *Customizability*: the method must be adaptable to the situation of SCS Logistics
- *Speed*: the method must be able to quickly calculate an estimated purchase price

4.6.1 Method choice for Freight Rate

As stated in Section 3.4, the cost drivers: *shipment details* and *service level* determine the freight rate. Historical data of these cost drivers is available. This data consists of past shipments. In addition, the price of shipments that have not been arranged in the past must be estimated. Therefore the possible cost estimation methods for determining the freight rate are:

- Case-based reasoning
- Expert systems
- Regression analysis models
- Back-Propagation Neural-Network (BPNN) Models

In Table 2, these cost estimation methods are assessed based on the criteria.

Table 2: Different Cost Estimation Methods with their Score per Criterion

	Accuracy	Simplicity	Customizability	Speed
Case-based reasoning	+ +	+ +	+ + +	+ + +
Expert systems	+ +	+	+ + +	+ +
Regression analysis models	+ +	−	+	+
BPNN	+ +	− −	+	+

Based on the results in Table 2, we come to the conclusion to combine two cost estimation methods. The case-based reasoning method is in fact very suitable to the situation of SCS Logistics. Because the process tends to be creative, fast and easy to use, and the past data and/or previous expertise are available. In addition, there is almost no difference in the accuracy of the four methods. Furthermore, there are some missing shipments within the available data. This is the reason we combine the case-based reasoning method with the regression analysis method. Both regression analysis models and BPNN models are made for predicting missing values, but the BPNN models are deliberately not chosen as these methods requires a high expertise level and a long processing time. The Fuzzy-Logic Approach and Rule-Based systems are not even used in the assessment, since these methods are not well applicable in the situation of SCS Logistics as stated in Section 4.4.2.

4.6.2 Method choice for Additional Costs

No or very little historical data is available for the cost drivers that determine the additional costs. For this reason, all cost estimation methods except expert systems are excluded. Using the expertise of the employees, we will determine the additional costs. This knowledge is then stored in the cost price model and will be retrieved on request.

4.7 Conclusion

In this chapter we executed a systematic literature review to find relevant cost estimation methods for calculating the purchase price at SCS Logistics. From the literature reviewed, we found that the qualitative cost estimation methods fit best within the situation of SCS Logistics. Therefore intuitive and analogical cost estimation methods have been explored. The methods that fall under the intuitive category are case-based reasoning and decision support systems. In addition, the methods regression analysis models and back-propagation neural network models fall under the analogical category. After evaluating these methods based on the required criteria of SCS Logistics and the data availability of the cost drivers, the following three methods were chosen:

- Case-based reasoning
- Regression analysis models
- Expert systems

These three methods are combined and merged into a cost price model, which is described in Chapter 6.

5 Data Exploration and Preparation

In this chapter we explore and prepare the data of the historical transactions. Because both the case-based reasoning method and the regression analysis models, chosen in Chapter 4, require input data, we create a final data set from the initial raw data. This final data set is used during the design and development phase of the cost price model described in Chapter 6. In Section 5.1 we collect the raw data. Section 5.2 describes the collected data set. After the data set is collected and described, the variables are selected in Section 5.3. It is important that the data used as input is of good quality. Therefore we clean the data in Section 5.4. To get familiar with the data and discover initial insights, an exploratory data analysis is given in Section 5.5. After that we discuss the quality of the data in Section 5.6. Finally, the final data set is described in Section 5.7.

5.1 Data Collection

The software system of SCS Logistics is a 'customized' transportation management system (TMS) with a Microsoft SQL database for the storage of the data. When a request comes in from a customer, the planners of SCS Logistics will make a quotation. If the customer agrees with the quotation, the shipment is sold. After sale, the planners enter all the specifications of the freight, with the associated purchase and sales price, into the software system. The newly entered data is immediately stored in the Microsoft SQL database. The data from this database is therefore collected for the research. The collected data includes all transactions from 1-1-2020 to 2-4-2021. This is because the transactions before 1-1-2020 were not properly filled in by the planners. The planners filled in a value because they had to, for example, a 1 was always entered for loading meter. The process of creating the final data set is described below.

5.2 Data Set Description

In order to create the input data for the cost price model we examine the collected data set of total 39,473 transactions and 18 variables. These 18 variables are the specifications of the transactions and some are the cost drivers explored in Chapter 3. The 18 variables with the associated sub-variables are described in Table 3.

Table 3: Description of Initial Data Set

Variables	Sub-variables	Description	Example
Date		This is the date that the shipment was entered into the system	1-1-2020
CRType (CostRevenueType)	ADMK	Administrative expenses	-
	ANNK	Cancelation costs	-
	DIVERS	All other additional costs that do not fall under a category in the system	-
	DOCU	Customs and transport documents	-
	EXTL	Additional (un)loading address	-
	FIX	If the cargo has to be loaded or unloaded at a fixed time	-
	OPSL	Storage charges	-
	PERM	Permits	-
	TRANS	Freight rate	-
	TRVZ	Transport insurance	-
	WACH	Delay expenses	-
OrderNo		This is the number of the order	95341
SalesDebtor		This variable contains the abbreviation of the customer	POLCOEVO
PurchaseCreditor		This variable contains the abbreviation of the supplier	SPLIAMST
LoadingCountryCode		The code of the country where loading has taken place	NL
UnloadingCountrycode		The code of the country where unloading has taken place	DE
LoadingPostalCode		The postal code where loading has taken place	3197 LH
LoadingDestination		The destination where loading has taken place	ROTTERDAM
UnloadingPostalCode		The postal code where unloading has taken place	13500
UnloadingDestination		The destination where unloading has taken place	Martiques
LoadingMeter		The loading meters with one decimal place accuracy	13.6
Weight (TONS)		This the weight of the load in tonnes with one or two decimal places accuracy	6.5
ServiceLevel	Standard	A detailed description of the different service levels can be found in Section 3.2.1	-
	Express		-
	Exceptional		-
	Temperature control		-
	Container		-
	Air freight		-
	Other		-
ADR		This variable is a 'J' if it is an ADR freight and an 'N' if it is not an ADR freight.	J
		The 'J' stands for 'ja' and means 'yes' in English	
		The 'N' stands for 'nee' and means 'no' in English	
LoadingDate		This is the date the cargo was shipped	27-1-2021
DeliveryDate		This is the date the cargo was delivered	22-2-2021
PurchasePrice		This is the actual purchase price for the sold shipment	560.00

5.3 Selection of Variables

After the data has been collected and described we now select the relevant variables that we will use as input for the cost price model. The overlapping variables from the data set, in Table 3 in Section 5.2, are described in Section 5.3.1. To determine the purchase price, it is important that the data set contains the relevant cost drivers. In Chapter 3, we determined the relevant cost drivers. The selected variables with variable and data type are shown in Table 4.

Table 4: The Selected Variables

Variable	Variable type	Data type
PurchasePrice	Dependent	Continuous
CRType	Independent	Categorical
LoadingCountryCode	Independent	Categorical
UnloadingCountryCode	Independent	Categorical
LoadingPostalCode	Independent	Categorical
UnloadingPostalCode	Independent	Categorical
LoadingMeter	Independent	Continuous
ServiceLevel	Independent	Categorical
ADR	Independent	Categorical
OrderNo	-	-

The order number is used to look up transactions in the TMS. When an inconsistent value occurs, the associated transaction can easily be looked up in the TMS using the order number. It can then be determined why this transaction deviates and whether it should be removed.

5.3.1 Overlapping Variables

Some variables from the data set described in Table 3 in Section 5.2 contain the same information and overlap. Since these overlapping variables are not relevant cost drivers, they are not selected. The following overlapping variables have not been selected:

- **Weight (TONS)**
  The variable 'Weight (TONS)' is not selected, since the price is determined by the number of loading meters and the weight is already included in the loading meters. As mentioned in Section 3.1 a loading meter may weigh a maximum of 1.75 tons. When the freight of a loading meter weighs more than 1.75 tons, more loading meters will be charged. For example, if a load is 1.0 LDM but the weight is 3.5 tons, 2 LDM will be charged and entered in the TMS. It is therefore not necessary to include the weight as an independent variable in the data set.

- **Loading and Delivery Date**
  The loading and delivery date are also not selected. These dates are not important as the speed of the transport is already included in the type of service level.

- **Loading and Unloading Destination**
  Since the loading and unloading postal codes clearly indicate where the cargo must be loaded and unloaded, it is not necessary to include the name of the destination in the data set.

5.4 Data Cleaning

Now that the variables have been selected, we will clean the data set. First, we remove the missing values in Section 5.4.1. Thereafter, we remove the inconsistent values in Section 5.4.2. The data set consists of eight independent variables (the cost drivers), one dependent variable (the purchase price) and the order numbers. The variables containing missing and/or inconsistent values are shown in Table 5.

Table 5: Number of Missing and Inconsistent Values in the Data Set

Variables	# Missing values	# Inconsistent values
OrderNo	29	-
CRType	0	2
LoadingCountryCode	2514	2
UnloadingCountryCode	2671	4
LoadingPostalCode	163	1
UnloadingPostalCode	446	4
LoadingMeter	29	-
ServiceLevel	29	-
ADR	30	-
PurchasePrice	2334	956

For the development of a cost price model it is important that the data used as input is correct. Therefore, we correct or remove the transactions with missing and/or inconsistent values from the data set.

5.4.1 Missing Values

The transactions with missing values are removed with the following approach:

1. From the variable ‘PurchasePrice’ all *2334* transactions with missing values are removed. Transactions of which the purchase price is not known cannot be used. These transactions have no added value for determining the purchase price.
2. Also the transactions for which no value has been entered for the variable ‘ADR’ or ‘ServiceLevel’ are not usable. Because not all information that determines the purchase prices is then known. Therefore *28* transactions that have missing values for the variables ‘ADR’ and/or ‘ServiceLevel’ are removed. The other transactions were already removed in step 1.
3. After that, we delete the transactions where no loading or unloading postal code has been filled in. Because, then it is not clear where the shipment went or where it came from. Therefore we remove *496* transactions from the data set.
4. Finally, there are still 855 transactions with a missing value for the variable ‘LoadingCountryCode’ and 1001 transactions with a missing value for the variable ‘UnloadingCountryCode’. We could look up the loading and unloading countries for these transactions using the postal codes given in the system. But since this takes a lot of time and there are enough transactions left for the development of a cost price model, these transactions are removed. The number of deleted transactions is *1063*.

After removing the missing values 35,552 transactions remain.

5.4.2 Inconsistent Values

To assess whether the transactions that contain inconsistent values should be removed from the data set, additional information is needed. The transactions are looked up in the TMS using the order number. The software system contains the name of the planner with all associated information about the shipment. In consultation with the planner and the available information, it is examined whether the inconsistent values can be corrected or should be removed.

The following inconsistent values were found:

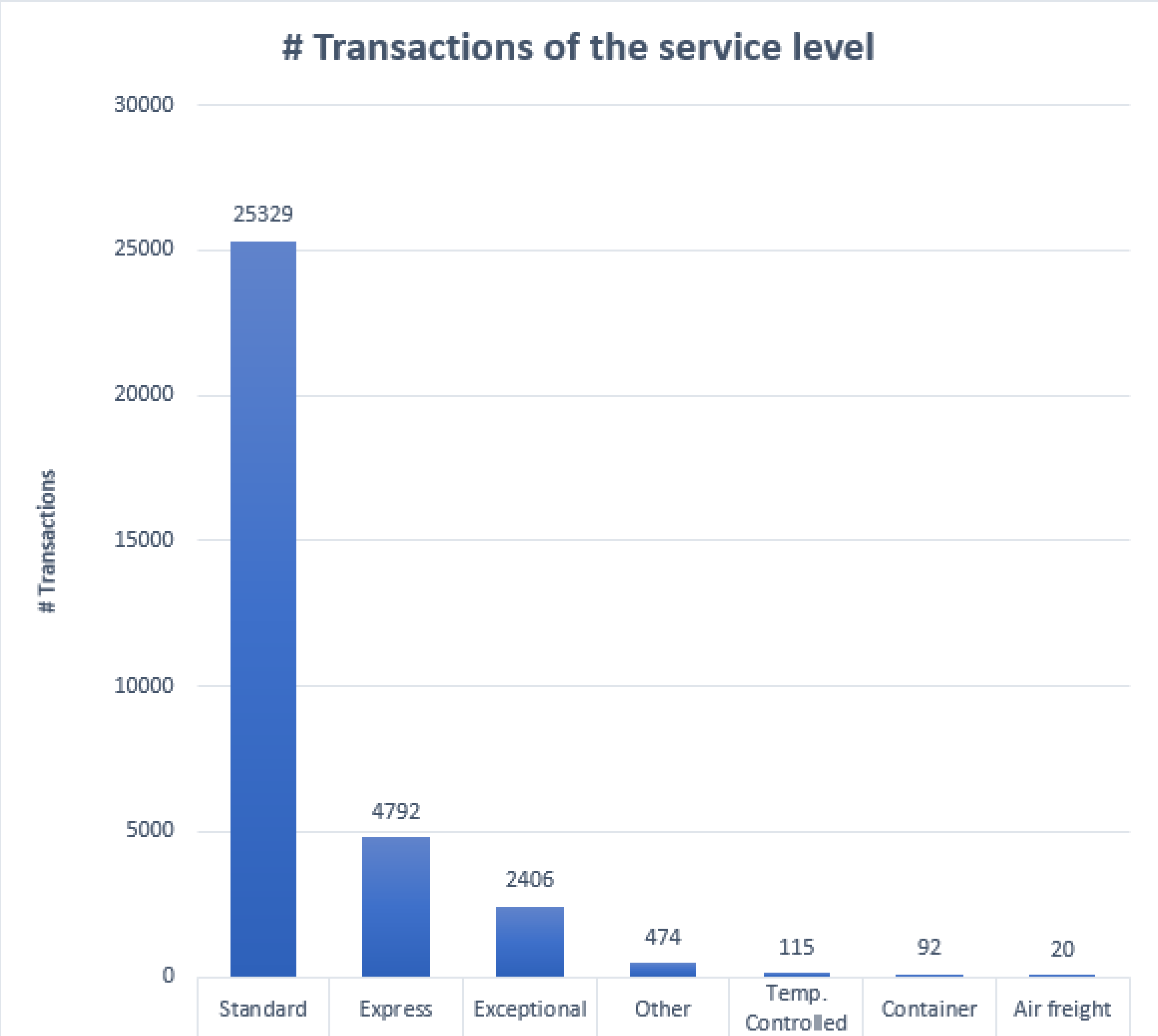
- Purchase prices with a 0 or a minus number are incorrect. This is because shipments with a purchase price of 0 euros or with a minus amount are not possible. Often these shipments have been entered incorrectly or the cargo of the shipments have been damaged. We remove *72* transactions with a negative purchase price and *779* transactions with a zero price. The other transaction with an inconsistent purchase price were already removed in Section 5.4.1.
- The variable 'CRType' contains two incorrectly entered values. These two values are '2' and '3450'. When looking in the software system we discovered that the shipment with CRType '3450' could be replaced by 'TRANS', but that the value 2 cannot be corrected.
- The variable 'LoadingCountryCode' contains two incorrectly entered values. These two values are '08:' and 'T'. The transaction with value '08:' turns out to be loaded in the town Bad Iburg, so the country code is corrected to DE. The transaction with value 'T' turns out to have been loaded in Italy. The country code is corrected to IT.
- The variable 'UnloadingCountryCode' contains four incorrectly entered values. There are two transactions with an unloading country code 'DDK'. This code means Denmark and therefore 'DDK' is corrected to 'DK'. The unloading country code 'NL' has also been entered incorrectly in two different ways. Therefore the codes 'N' and 'BNL' are corrected to 'NL'.
- The variable 'LoadingPostalCode' contains one incorrectly entered value. This is the postal code ')X29 7PL'. This value is not corrected and is removed from the data set for accuracy.
- The variable 'UnloadingPostalCode' contains four times the incorrectly entered postcode '541 AA'. This value is not corrected and the *4* transactions are removed from the data set for accuracy.

After correcting and removing the inconsistent values from the data set, 34,695 transactions remain.

5.5 Exploratory Data Analysis

Now that the data set has been created and cleaned, we will explore the data set. To get some insight into the types of transactions and shipments that SCS Logistics perform, bar graphs are created using *Microsoft Excel*. Because we want to gain insight into the transactions of the shipments and not the transactions of the additional costs, the CRType is filtered on 'TRANS'. This results in a total of 33,228 shipments.

Figure 13 shows that by far the most shipments for which SCS Logistics arranges the transport are standard shipments. Between 1-1-2020 and 2-4-2021, 76.23% of all shipments fell under the service level standard.



	Standard	Express	Exceptional	Other	Temp. Controlled	Container	Air freight
Percentage	76.23%	14.42%	7.24%	1.43%	0.35%	0.28%	0.06%

Figure 13: Number of Transactions of the Service Level (n=33228)

In total, the data set contains 42 different loading countries and 61 different unloading countries. Figure 14 shows the top 5 countries with the most shipments. Most shipments arranged by SCS Logistics are to and from the Netherlands. This is because a large part of SCS Logistics' customers are located in the Netherlands. Germany is number 2 with about half fewer shipments than the Netherlands. After Germany the countries Great Britain, France and Belgium follow.

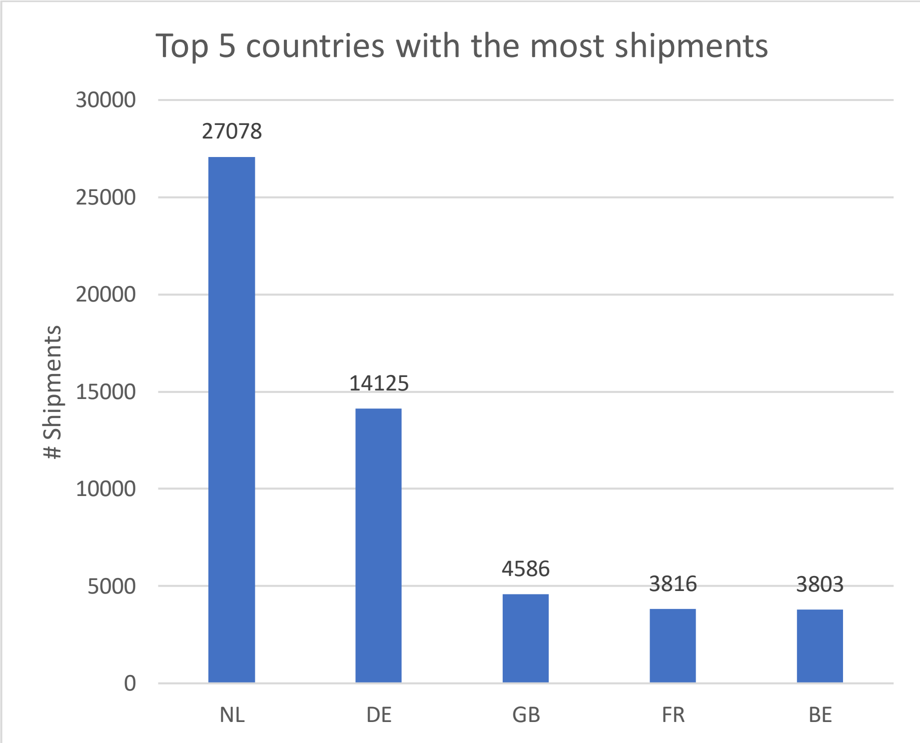


Figure 14: Top 5 Countries with the Most Shipments

Figure 15 shows the shipments arranged by SCS Logistics in more detail and contains the top 5 most shipments between countries. Most shipments for which SCS Logistics has arranged transport are from Germany to the Netherlands. Followed by shipments from the Netherlands to Germany. It can therefore be concluded that most shipments take place between the Netherlands and Germany.

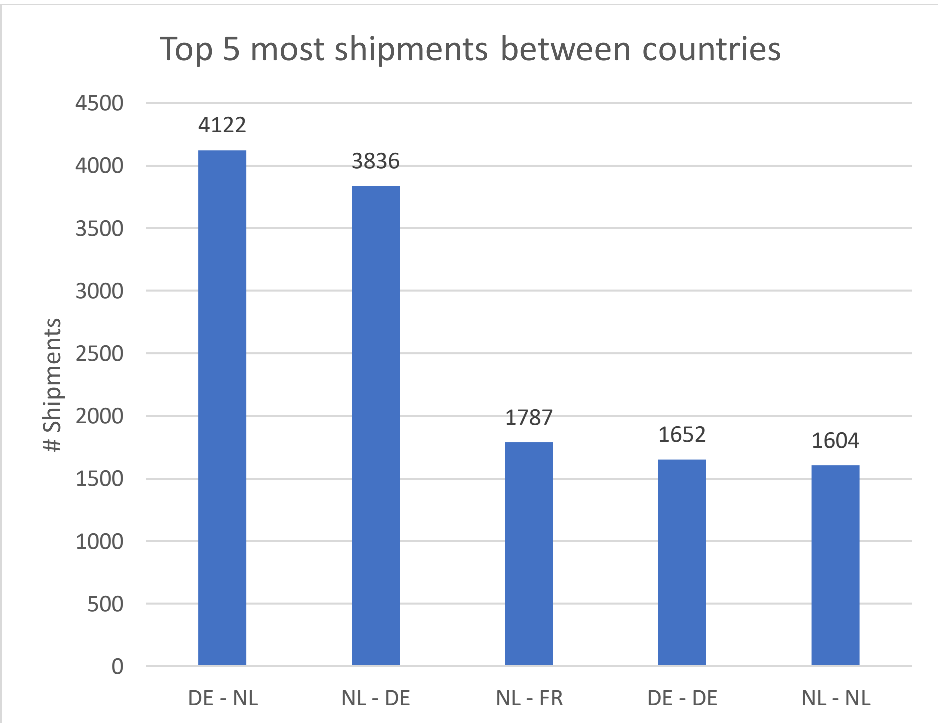


Figure 15: Top 5 Most Shipments between Countries

The exploratory data analysis showed the most common transactions/shipments. These transactions/shipments has to be taken into account in the design and development phase of the cost price model. This will be discussed in more detail in Chapter 6.

5.6 Data Quality

For the development of a cost price model it is important that the input data is of good quality. The majority of the transactions accurately reflect the purchase price with accompanying specifications of the shipment. However, there are also a number of transactions in the data set that have been entered incorrectly. Table 6 shows an example of an incorrectly filled in shipment. The purchase price of € 389.00 are the storage costs of the cargo. However, the cost revenue type (CRType) contains TRANS. These types of incorrectly filled in shipments give a distorted picture of the purchase price. Because this purchase price does not belong to the transport of a load with 0.4 loading meter and a standard service level, but to the storage of the load. The data set does contain more of these types of incorrectly entered shipments. For example, the service level is sometimes entered incorrectly. It is important that these types of incorrectly entered shipments are filtered out of the data set. This point is therefore taken into account during the development of the cost price model in Chapter 6, where we look more specifically at different shipments and the incorrect transactions will be removed. The outliers and remarkable shipments can easily be found in the TMS using the order number. Together with the planners, it can then be determined whether the shipment has been filled in correctly or not.

Table 6: Example Incorrectly Filled In Shipment (in this case storage costs are reported as transport costs)

CRType	Loading CountryCode	Unloading CountryCode	Loading PostalCode	Unloading PostalCode	Loading Meter	Service Level	ADR	Purchase Price
TRANS	DE	NL	30855	7575 EC	0.4	Standaard	N	€ 389.00

5.7 Final Data Set Description

After preparation of the data, the final data set contains 34,695 historical transactions. The data set consists of eight independent variables (the cost drivers), one dependent variable (the purchase price) and the order numbers as shown in Table 4 in Section 5.3. Almost all independent variables are categorical variables. Furthermore, all transactions with missing values are removed from the data set and the transactions with inconsistent values that stood out have been corrected or removed. But as described in Section 5.6, the data set contains many more transactions with inconsistent values. These cannot all be removed from the data set yet. But when we look more specifically at the data during the development of the cost price model in Chapter 6, these inconsistent transactions will be manually removed in consultation with the employees.

6 Design and Development of Cost Price Model and Tool

In this chapter the *cost price model*, which forms the basis for determining the bid prices, is developed. To make the cost price model usable for the employees, we have implemented the model in a user-friendly tool. In this way, we can test the cost price model. We programmed the tool using *Microsoft Excel* and *Visual Basic for Applications*. The cost price model in combination with the tool ultimately forms the solution to the core problem. First, we establish the requirements for the cost price model in Section 6.1. In Section 6.2 we describe the data used for the cost price model. And in Section 6.3 we give a concise description of the model and the corresponding tool. After that, follows a detailed description of how the cost price model and tool works. In Section 6.4 we describe the user interface of the tool. Section 6.5 explains how we implement the case-based reasoning process. The case-based reasoning process forms the framework for the cost price model. Then in section 6.6 we discuss how we calculate the freight rate using regression analysis, when similar shipments are missing from the database. Furthermore, we determine the cost prices of the additional costs using the expertise of the employees in section 6.7. And in Section 6.8 we provide a summary of the entire cost price model process.

6.1 Requirements for the Cost Price Model

The cost price model must be easy to use, so that the employees receive a purchase price quickly and can therefore quickly issue a bid price. The requirements of the cost price model were established by conducting an interview with the CEO of SCS Logistics. Therefore, the following requirements are taken into account for designing the cost price model tool:

1. The purchase price can be calculated based on input parameters.
2. Each newly entered transaction must be automatically added to the data used (database).
3. The model should be easy to use.
4. The model must be able to be implemented in the current software system.

To implement the cost price model in the software system of SCS Logistics, we contacted the software developer. From this is has become clear that the software system of SCS Logistics is a ‘customized’ transportation management system (TMS) with a Microsoft SQL database for the storage of the data. To make it possible for the software developer to eventually implement the cost price model within the TMS, they require that we develop the cost price model in *Microsoft Excel*. Therefore a tool is being developed using *Microsoft Excel* and *Visual Basic for Applications*. Using Macros, several calculations and tasks can be performed at once. Buttons are used to execute these Macros. The tool is used for testing the cost price model (the solution method, for determining an appropriate purchase price). When the cost price model works well, the tool is used as a basis for the software developer. In this way, the software developer knows how the cost price model works and can implement it in the TMS. Recommendations for the implementation are given in Chapter 7.

6.2 Data for the Cost Price Model

The data used for the cost price model is based on the relevant cost drivers defined in Chapter 3. The final variables selected are described in Section 5.3. The exploratory data analysis in Section 5.5 shows that 76.23% (25,329) of the transactions consist of shipments with a standard service level. In addition, most shipments are from Germany to the Netherlands (4,122) and from the Netherlands to Germany (3,836) as shown in Figure 15 in Section 5.5.

As mentioned in Section 5.4, many transactions with missing and/or inconsistent values have been removed. However, the final data set as described in Section 5.7 contains many more incorrectly entered transactions, as stated in section 5.6. These incorrect transactions cannot be quickly found and deleted.

For this reason, the cost price model is developed for shipments from the Netherlands to Germany and from Germany to the Netherlands with a standard service level, as these shipments are the most common. In addition, all shipments with ADR are filtered out, because the purchase price of these shipments gives a distorted picture of the normal purchase price. The final data set consist of 6,650 transactions.

6.3 Concise Description of the Cost Price Model and the Corresponding Tool

As noted in Chapter 4, a combination of the following three cost estimation methods will be used to develop the cost price model:

- Case-based reasoning
- Regression analysis models
- Expert systems

We made the cost price model usable by developing a tool with *Microsoft Excel* and *Visual Basic for Applications*. The *case-based reasoning process* and *regression analysis models* are used for calculating the freight rate. With the case-based reasoning process, we search for the similar shipments in the database based on input parameters. The estimated freight rate is then calculated by taking the average of the similar shipments. However, it is also possible that no similar shipments can be found in the database. In that case, *regression analysis* is used to estimate the freight rate. To apply this, the Netherlands and Germany are first divided into zones. After that, we determine the mathematical equations in order to calculate the freight rates for the combinations between the different zones. We do this by fitting curves through the data points of the shipments between the different zones. Furthermore, the additional costs are determined based on the *expertise* of the employees. When the additional costs are determined, they are added into the cost price model. Finally, the total purchase price is calculated by adding the freight rate and additional costs together. After all necessary costs have been calculated and checked for correctness, the employees can add the shipment with the costs to the SCS Logistics database. In this way, this freight rate of the new shipment could be used for calculating the future shipments. In the following sections a detailed explanation of how the combination of the cost price model and tool works is given. Additionally, the explanation is supported with examples and images of the tool.

6.4 User Interface of the Tool

By developing a tool, the employees of SCS Logistics are able to use the cost price model in a user-friendly way and we are able to test the cost price model. The tool interacts with the cost price model, which plays on the background of the tool. So using the tool, the cost price model will be activated. The output data that the tool gives is described in Section 6.4.1. Figure 16 shows the user interface of the tool.

Figure 16: User Interface of the Tool

When the employees of SCS logistics want to generate a new shipment, they have to click on the button 'New Order'. When employees click on that button a new form, which is shown in Figure 17, pops up on the screen.

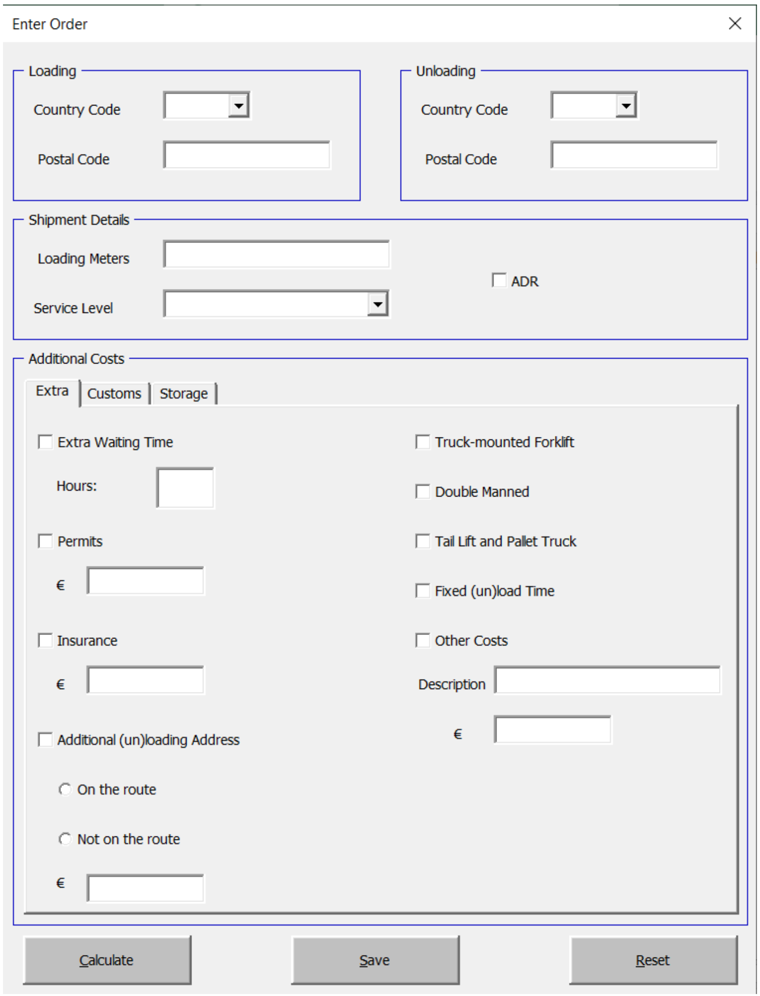


Figure 17: Pop-up Form when Clicking on the Button 'New Order'

6.4.1 Output Data

When using the 'calculate' button, shown in Figure 17, the model will process the following output data:

- Freight rate
- Additional costs
  - Extra costs
  - Customs document costs
  - Storage costs
- Total costs

As discussed in Section 4.6 the freight rate is calculated using *case-based reasoning* and *regression analysis models*. Furthermore, the additional costs are determined based on the *expertise* of the employees. The total costs are the freight rate and the additional costs added together. How the tool, with the associated calculations of the cost price model works is explained in the following sections.

6.5 Implementation of Case-Based Reasoning

According to the literature review described in Chapter 4, we select the case-based reasoning method as the framework for the cost price model. This section provides the implementation of the case-based reasoning process. By using *Microsoft Excel* and *Visual Basic for Applications* we create the case-based reasoning model. According to Hu et al. (2016), the case-based reasoning cycle consists of the following four phases and is shown in Figure 12 in Section 4.4.1:

1. Retrieving (6.5.1)
2. Reusing (6.5.2)
3. Revising (6.5.3)
4. Retaining (6.5.4)

With the help of the case-based reasoning process we determine the freight rates. We are now going to explain the four phases.

6.5.1 Retrieving Process

The case-based reasoning process adapts past design information collected from a database to the new design (Busachi et al., 2017). So to determine the correct freight rate, the same type of shipments that were arranged in the past must be retrieved. The similar past cases can be identified with the help of key attributes (Elmousalami, 2021). These key attributes are the input parameters that the employees of SCS logistics must enter in the tool.

<u>*Remark:*</u> the retrieving process can only be used when there are similar shipments in the database. If that is not the case, regression analysis is used. How we calculate the freight rate using regression analysis is explained in Section 6.6.

Input Parameters

The input parameters are the characteristics of the shipment. In Figure 18 an example of the retrieving process is given. The employees have to fill in the following input parameters to ensure that the cost price model can search for similar shipments:

- Loading country code
- Loading postal code
- Unloading country code
- Unloading postal code
- Loading meters
- Service level

For the 'loading postal code' and 'unloading postal code', the employees of SCS Logistics do not necessarily have to fill in the full postal code. They often use only the first two digits of the postal code to determine the price. This is because the first two digits of a postal code indicate a postal code area, as shown in Figure 21 & 22 in Section 6.6.1.

All input parameters, listed above, must be filled in by the employees otherwise the tool will not be able to search for similar shipments. When all input parameters have been filled in, the employee can calculate the freight rate by clicking on the 'Calculate' button, shown in Figure 17. But if not all input parameters have been filled in, the employee will receive a message box and the box(es) of the input parameter(s) that have not been filled in will turn red. In order to calculate the freight rate, the red coloured box(es) must still be filled in.

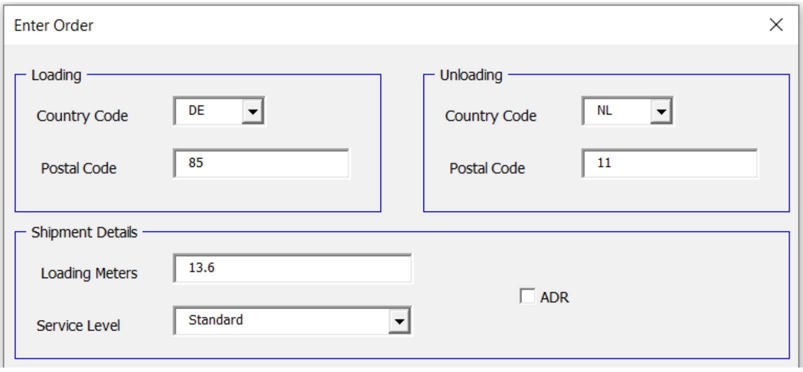


Figure 18: Example Retrieving Process

Freight Rate Calculation

When the employee presses the calculate button, the cost price model is activated. Using the entered shipment details and the cost revenue type (CRType) 'TRANS', the model searches in the database of SCS Logistics for similar shipments. After the similar shipments have been found, the freight rate is calculated. The freight rate is the average purchase price of all similar shipments and is calculated using Equation 4.

$$\overline{freightrate} = \frac{\sum_{i=1}^{n} freight\,rate_i}{n} \tag{4}$$

When the shipment is an ADR shipment, an additional 10% will be charged on top of the freight rate. The employee can indicate an ADR shipment by checking the ADR box, shown in Figure 18.

6.5.2 Reusing and Revising Process

If the current case and retrieved cases are sufficiently similar, solutions of retrieved cases can be reused directly without any modification (Hu et al., 2016). But when the freight rate is not correct, the employees must revise the price by taking the differences between the new case and the retrieved cases into account. Using the 'Calculate' button, the employees can calculate the purchase prices and the freight rate will be displayed in the form on the left of the home page, which is shown in Figure 16 in Section 6.4 and in Figure 19. When the employees have issued the bid prices based on the estimated purchase price and the customer has agreed, a suitable supplier is selected and the actual purchase price is known. It may be possible that the actual purchase price differs from the estimated purchase price. In that case, the employees must adjust the estimated purchase prices in the form in Figure 19 to the actual purchase price. In this way, the accuracy of the estimated purchase prices is improved and ensured

Purchase:		
Full Loading Postal Code:		85232
Full Unloading Postal Code:		1119 NR
Type of Costs:	Amount:	
Freight Rate		€ 810.00
Total Costs:		€ 810.00

Figure 19: Example Reusing and Revising Process

6.5.3 Retaining Process

The last process of the case-based reasoning method is the retaining process. In this process the new case and its solution can be retained in the database for future reuse. This makes case-based reasoning a self-learning system (Hu et al., 2016). When all costs are correct according to the employee, he or she can add the new shipment to the database using the 'Save' button, shown in Figure 17 in Section 6.4. Before the shipment is added to the database, the employee still has to fill in the 'Full loading postal code' and the 'Full unloading postal code'. Employees can do this at the top of the form on the home page, as shown in Figure 19 in Section 6.5.2. Once the shipment has been saved, it is added to the database, as shown in Figure 20. The shipment receives a new order number and a matching cost revenue type (CRType). Furthermore, the shipment gets a 'Loading zone' and 'Unloading zone', but this will not shown in the database. The function of these zones will be explained in Section 6.6. The next time the employees receive a request for the same kind of shipment, this newly added shipment is included in the calculation of the average. Finally when the shipment has been saved, the employee can use the 'Reset' button, shown in Figure 17 in Section 6.4, to clean all input data and the home page.

OrderNO	CRType	LoadingCountryCode	UnloadingCountryCode	LoadingPostalCode	UnloadingPostalCode	LoadingMeter	ServiceLevel	ADR	PurchasePrice
6301	TRANS	NL	DE	7559 SH	52372	0.4	Standard	N	100.00
6302	TRANS	DE	NL	75417	7575 EC	1.6	Standard	N	150.00
6303	TRANS	DE	NL	59597	7575 EC	0.5	Standard	N	70.93
6304	TRANS	DE	NL	37520	7575 EC	1.5	Standard	N	200.00
6305	TRANS	DE	NL	88239	1437 EK	5.5	Standard	N	500.00
6306	TRANS	DE	NL	94577	1118 LA	13.6	Standard	N	950.00
6307	TRANS	DE	NL	85232	1119 NR	13.6	Standard	N	810.00

Figure 20: Example Retaining Process

6.6 Calculation of Freight Rate when there are no Similar Shipments

According to Busachi et al. (2017) it also occurs that the old and new case does not have strong similarities. This is also the case with this case-based reasoning process. The database does not contain all possible combinations between the two-digit postal code areas. For example, there is no shipment from the postal code area 44 in Germany to the postal code area 75 in the Netherlands. And even if the database will contain all possible combinations between the two-digit postal code areas, then there will not be at least one shipment known for every loading meter. Therefore, when no similar shipments can be found missing information and data has to be collected and implemented. We calculated these missing data (freight rate) by fitting curves to the data points of the shipments that are known. According to

Brown (2001) *Microsoft Excel* is a good programme for fitting curves. It offers a user-friendly interface, with built-in mathematical functions, flexible data manipulation and instantaneous graphing of data. Therefore we use *Microsoft Excel* to fit mathematical equations to the data.

The plan of approach we used for calculating the missing data (freight rate) is as follows:

1. The Netherlands and Germany are divided into different zones (Section 6.6.1)
2. Removing the outliers (Section 6.6.2)
3. Curve fitting to the data points (Section 6.6.3)
4. Implementing the mathematical equations in the tool (Section 6.6.4)

These four phases are explained in the following chapters.

6.6.1 Dividing the Netherlands and Germany into Different Zones

We first divide the Netherlands and Germany into different zones. We do this for the following two reasons:

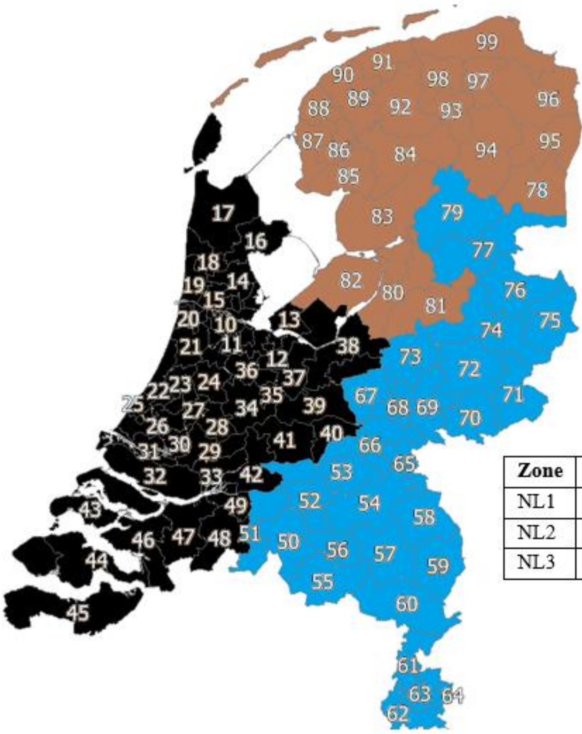
1. For fitting curves to the data enough data points must be available.

2. And we would like to know the relationship between the loading meter and the freight rate. Therefore, all other variables that determine the freight rate must be kept constant. Since all variables except the distance are already constant, we try to keep the distance as constant as possible with the zones.

With the help of the expertise of the workers, the zones are determined on the basis of distance and accessibility. First we divided Germany into nine zones, which can be seen in Figure 21, and the Netherlands into three zones, which can be seen in Figure 22.

Zone	Postal code	Colour
DE1	40-42 & 44-49	
DE2	50-53 & 57-59	
DE3	26-28	
DE4	20-22 & 29-33	
DE5	34-38 & 54-56 & 60-69	
DE6	19 & 23-25 & 70-71 & 74-76	
DE7	39 & 72-73 & 77 & 89-91 & 95-99	
DE8	01-16 & 78-81 & 85-86	
DE9	17-18 & 82-84 & 87-88 & 92-94	

Figure 21: Germany Divided into Nine Zones



Zone	Postal code	Colour
NL1	10-49	
NL2	80-99	
NL3	50-79	

Figure 22: The Netherlands Divided into Three Zones

Table 7: Number of Shipments between Zones, when the Netherlands is Divided into Three Zones

Export NL (From NL to DE)	DE1	DE2	DE3	DE4	DE5	DE6	DE7	DE8	DE9	Total
NL1	140	125	157	255	162	186	177	222	342	**1766**
NL2	12	5	6	18	21	19	8	7	9	**105**
NL3	210	74	62	246	127	93	175	143	82	**1212**
Total	**362**	**204**	**225**	**519**	**310**	**298**	**360**	**372**	**433**	
Import NL (From DE to NL)	**DE1**	**DE2**	**DE3**	**DE4**	**DE5**	**DE6**	**DE7**	**DE8**	**DE9**	**Total**
NL1	214	123	101	241	239	123	153	205	164	**1563**
NL2	22	10	0	9	13	4	8	8	6	**80**
NL3	167	243	27	188	368	176	221	447	87	**1924**
Total	**403**	**376**	**128**	**438**	**620**	**303**	**382**	**660**	**257**	

As can be seen in Table 7, there are very few shipments to and from zone NL2. Because there are too few data points between NL2 and the zones in Germany, it is not possible to fit a mathematical equation to the data points. Too many values are missing. That is why we are forced to divide the Netherlands into two zones. Figure 23 shows how the Netherlands is ultimately divided. Table 8 shows the number of transactions between the zones.

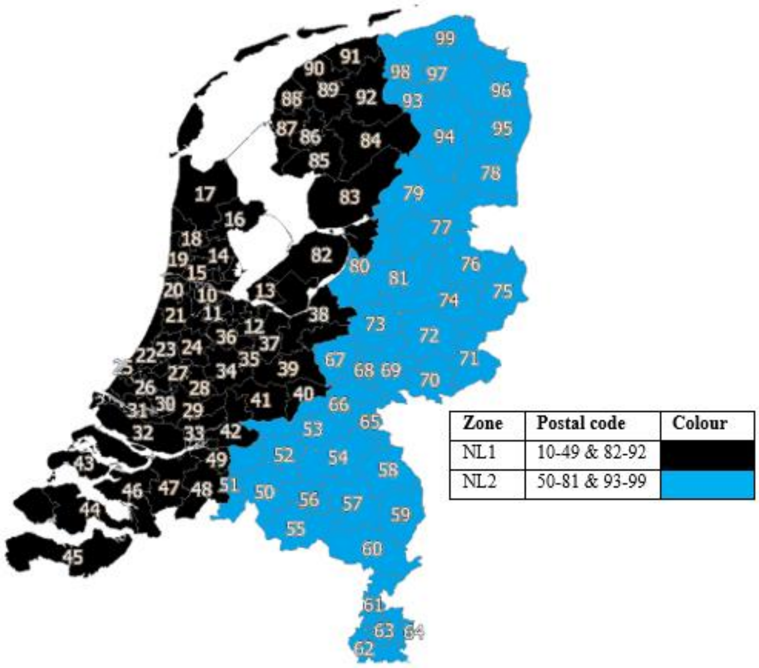


Figure 23: The Netherlands Divided into Two Zones

Table 8: Number of Shipments between Zones, when the Netherlands is Divided into Two Zones

Export NL (from NL to DE)	**DE1**	**DE2**	**DE3**	**DE4**	**DE5**	**DE6**	**DE7**	**DE8**	**DE9**	**Total**
NL1	140	128	163	272	173	199	184	226	347	**1832**
NL2	222	76	62	247	137	99	176	146	86	**1251**
Total	**362**	**204**	**225**	**519**	**310**	**298**	**360**	**372**	**433**	
Import NL (from DE to NL)	**DE1**	**DE2**	**DE3**	**DE4**	**DE5**	**DE6**	**DE7**	**DE8**	**DE9**	**Total**
NL1	222	131	101	246	250	126	160	206	165	**1607**
NL2	181	245	27	192	370	177	222	454	92	**1960**
Total	**403**	**376**	**128**	**438**	**620**	**303**	**382**	**660**	**257**	

6.6.2 Removing Outliers

According to Yu Hao et al. (2020) regression analysis models are sensitive to data outliers. When outliers are present, the accuracy of the regression models decreases. It is therefore necessary to identify and remove or correct the outliers. Outlier identification and removal or correction is part of the data screening process which should be done routinely before the statistical analyses (Laurikkala et al., 2000). To clarify the approach for identifying and removing or correcting the outliers, the shipments from zone NL1 to zone DE9 are given as an example. After removing or correcting the first outliers in Section 5.4, we plotted the shipments between the different zones in separate scatter plots. The scatter plot of the shipments from zone NL1 to DE9 can be seen in Figure 24.

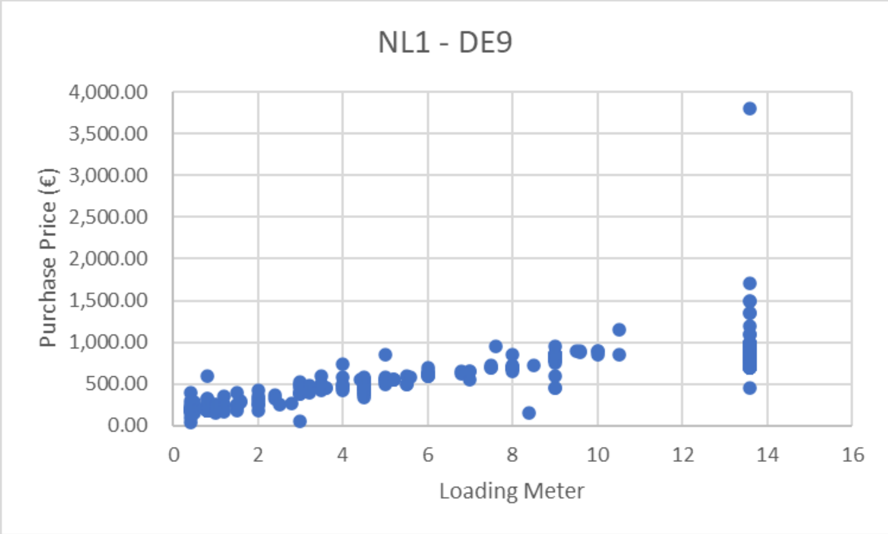


Figure 24: Scatterplot of the Shipments from Zone NL1 to DE9 before Removing Outliers

As can be seen in Figure 24, the data point at 13.6 loading meter with a purchase price of 3,800.00 is a possible outlier. To visualize the possible outliers, we have plotted the purchase prices in a box plot, which can be seen in Figure 25. According to Laurikkala et al. (2000) box plots are an easy and informal method for outlier detection. The box plot is a simple display of the five-number summary:

- Lower extreme: the smallest data value excluding any outliers
- Lower quartile: median of the lower half of the dataset
- Median: middle value of the dataset
- Upper quartile: median of the upper half of the dataset
- Upper extreme: the largest data value excluding any outliers

The outliers are defined with the thresholds, which are defined as follows (Laurikkala et al., 2000):

- Lower threshold = lower quartile – 1.5 IQR → 250 – 1.5 (800 – 250) = -575
- Upper threshold = upper quartile + 1.5 IQR → 800 + 1.5 (800 – 250) = 1625
- IQR = upper quartile – lower quartile

Therefore the datapoint x is a lower outlier, if x < lower threshold and an upper outlier, if x > upper threshold.

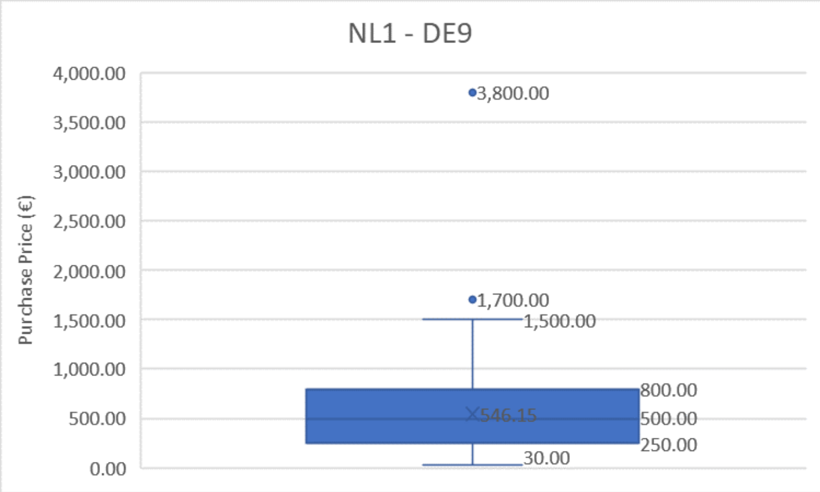


Figure 25: Boxplot of the Shipments from Zone NL1 to DE9

In the case of the shipments form zone NL1 to zone DE9 the thresholds are:

- Lower threshold: 250 – 1.5 (800 – 250) = -575
- Upper threshold: 800 + 1.5 (800 – 250) = 1625

Therefore the shipments with a purchase price of 1,700.00 and 3,800.00 are outliers. But to be sure that these two transactions are outliers we do a subjective evaluation. The transactions are looked up in the TMS using the order number. The software system contains the name of the planner with all associated information about the shipment. In consultation with the planner and the available information, it is examined whether the outlier can be removed or corrected. In the end we come to the conclusion that both transactions can be removed. The reason for this is described in Table 9.

After removing the two shipments, the scatterplot looks like this, as can be seen in Figure 26.

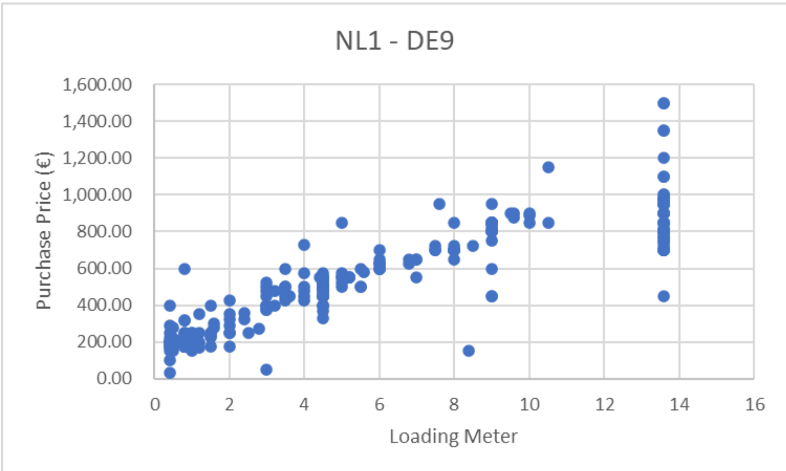


Figure 26: Scatterplot of the Shipments from Zone NL1 to DE9 after Removing Outliers with Boxplot

Figure 26 shows that there may still be possible outliers between data. That is why the deviating data points are searched for in the system using the order number. The result is that 13 more outliers have been removed or corrected. These outliers are described in Table 9.

Table 9: Outliers of the Shipments from Zone NL1 to Zone DE9

OrderNO	Loading Meter	Purchase Price (€)	The Mistake
103721	13.6	3,800.00	This is a project, with 4 trucks
126575	13.6	1,700.00	Express shipment
109763	3	50.00	Customs costs
119930	0.4	30.00	Extra cost
106249	8.4	150	Error freight costs
107299	13.6	1,500.00	Express shipment
108504	13.6	1,500.00	Express shipment
117283	13.6	1,350.00	Price should be € 1,000.00
117283	13.6	1,350.00	Price should be € 1,000.00
106922	13.6	450.00	Price should be € 800
103488	9	450.00	2 x € 450 is entered in the system, but the total price should be € 900
132675	0.8	320	Extra cost
129522	0.8	600.00	Express shipment
109617	0.4	400.00	Express shipment
132467	5	850.25	Price should be € 500

Figure 27 shows the scatter plot of all shipments from zone NL1 to zone DE9 after the process of identifying and removing or correcting the outliers.

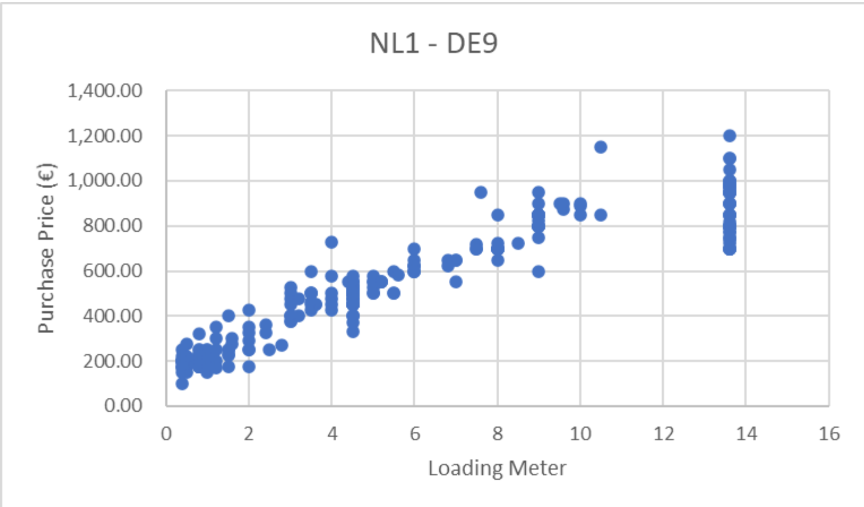


Figure 27: Scatterplot of the Shipments from Zone NL1 to DE9 after Removing all Outliers

To conclude, the approach of identifying and removing or correcting the outliers has been explained with the shipments from zone NL1 to zone DE9. We used this approach for all shipments between the different zones. How many outliers we removed per combination between the zones can be seen in Figure B.1 in Appendix B. But in total we removed 344 outliers, causing that 6,306 transactions remain. At the end we have 36 different scatterplots, which we have to fit a curve through. This process is explained in the chapter below.

6.6.3 Curve Fitting Process

Now that all outliers have been removed from the database and the scatterplots of the shipments between the different zones have been plotted, we start with the curve fitting process. The goal of this curve fitting process is to eventually calculate the freight rate when no similar shipments can be found in the database. With the exception of the loading meter, we have made all variables that determine the freight rate constant. These variables are:

- Service level = standard
- ADR = only shipments without ADR
- The distance:
  Because we have divided the Netherlands and Germany into zones, the distance of the shipments between certain zones remains approximately the same. For example, the distance of all shipments from zone NL1 to DE4 is approximately the same.

Because all variables, except for the loading meter, are constant, we can determine the relationship between the loading meter and the purchase price. We do this by determining mathematical equations for all shipments between the different zones. These mathematical equations are the functions of the fitted lines to the data points. With *Microsoft Excel* there are two ways for fitting curves to data points (John, 1998):

1. <u>Standard curve fitting function:</u>
   This is the Trendline function, within *Microsoft Excel*. The function permits the user to obtain a least squares fit to a given data series using any of the following models (John, 1998):
   - Linear (y = mx + c)
   - Logarithmic (y = a * ln(x) + b)
   - Polynomial ($y = b + a_1x + a_2x^2 + a_3x^3 + a_4x^4 \ldots$)
   - Power ($y = ax^b$)
   - Exponential ($y = ae^{bx}$)
2. <u>The Solver add-in:</u>
   The Solver is an optimisation procedure, which can be used to generate solutions to a wide spectrum of linear, non-linear and integer problems (John, 1998). To fit the curves, the Solver finds the minimum value for the sum of the squared errors, by changing the parameters of the chosen model. The only difficulty of the Solver is which model should be chosen (John, 1998).

As mentioned above, both methods use the least squares fit method. This method is based on the principal that the magnitude of the difference between the data points and the curve is a good measure of how well the curve fits the data (Brown, 2000). The least squares fitting method minimizes the sum of the squares (SS), which is described by Equation 5 (Brown, 2000).

$$SS = \sum_{i=1}^{n} (y - y_{fit})^2 \tag{5}$$

Where y is the data point, and y_{fit} the value of the curve at point y.

One can see in Figure 27 that when the number of loading meters increases, the purchase price flattens more and more. The data points in the scatterplot increases strongly, where in the end they flatten more and more. This is because the employees often quote a freight rate of a full truck load (13.6 loading meters) from 10 loading meters. For this reason, all scatterplots flatten off from 10 loading meters. We searched for a curve that fits this characteristic best. We expect that the models: logarithmic, power and quadratic polynomial, with only one side of the U shape, fits the data best. Since both methods of curve fitting in *Microsoft Excel* use the least squares fit method and the Trendline function contains the expected models, we use the Trendline function.

To assess the goodness of fit of the equations, we calculate the R^2 value. The R^2 value expresses the proportion of variance in the dependent variable explained by the independent variable (Brown, 2000). The R^2 value is a value between 0 and 1. The closer the value is to 1, the more accurately the equation fits the data.

Using the following calculation example, we are going to explain the curve fitting process.

<u>Calculation example:</u>

Suppose we want to know the freight rate of a shipment with the following shipment details:

- From NL postal code 3147 PA (= Zone NL1) to DE postal code 83104 (= Zone DE9)
- Loading meter = 4
- Service level = Standard
- ADR = No

This shipment is not in our database. Therefore, no freight rate can be calculated using the case-based reasoning process. Since the freight rate cannot be calculated using the case-based reasoning process, the freight rate must be calculated using a mathematical equation. To determine this mathematical equation we need to fit a line through the data points. Figure 28 shows a fitted curve through the datapoints of the shipments from NL1 to DE9.

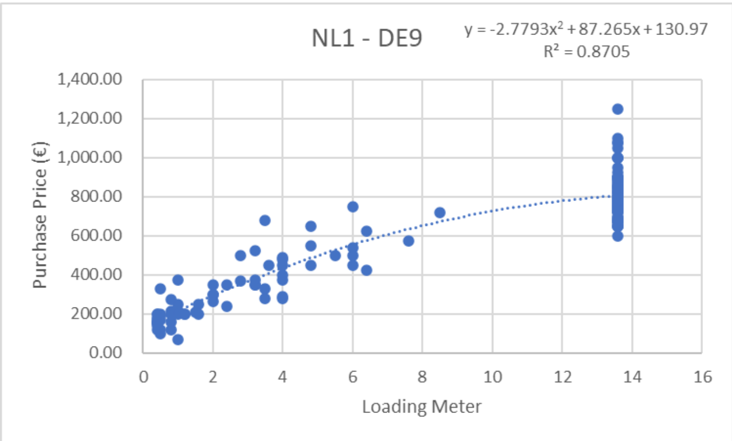


Figure 28: Curve Fitting to the Datapoints of the Shipments from NL1 to DE9

The regression type of the fitted curve is polynomial, with the following mathematical equation:

$$y = -2.7793x^2 + 87.265x + 130.97 \quad (6)$$

The calculated $R^2 = 0.8705$. This is close to 1 and means that approximately the model fits the data for 87%.

Now the equation is known we can calculate the freight rate of the shipment from the example by filling in the loading meters in the equation.

The freight rate = -2.7793 * 4^2 + 87.265 * 4 + 130.97 = € 435.56

This curve fitting process is done for all combinations between the zones. In this way, the freight rate can still be calculated for every shipment that is not in the database. The characteristics of all the fitted curves can be found in Figure B.1 in Appendix B.

6.7 Additional Costs

Now that the approach for calculating the freight rate is known, the additional costs are determined using the expertise of the employees. We determined the cost prices by conducting interviews with the employees of SCS Logistics. We examined that the additional costs can be divided into three categories.

The three categories with the associated costs are:

- **<u>Extra costs: (the form for the input parameters can be found in Figure 29)</u>**
    - Truck-mounted Forklift = € 125,-
    - Double manned = € 600,-
    - Tail lift and pallet truck = € 75,-
    - Fixed (un)load Time = € 100,-
    - Permits = Depends on the situation. The planner must be able to enter this cost price.
    - Insurance = Depends on the situation. The planner must be able to enter this cost price.
    - Extra waiting time = € 55,- per hour
    - Additional (un)loading address on the route = € 75,-
    - Additional (un)loading address not on the route = Depends on how far from the route. Therefore the planner must be able to enter this cost price.
    - Other additional costs = If there are other additional costs, which do not fall under any of the above costs, the planner must be able to enter this cost price with clear description about the costs.

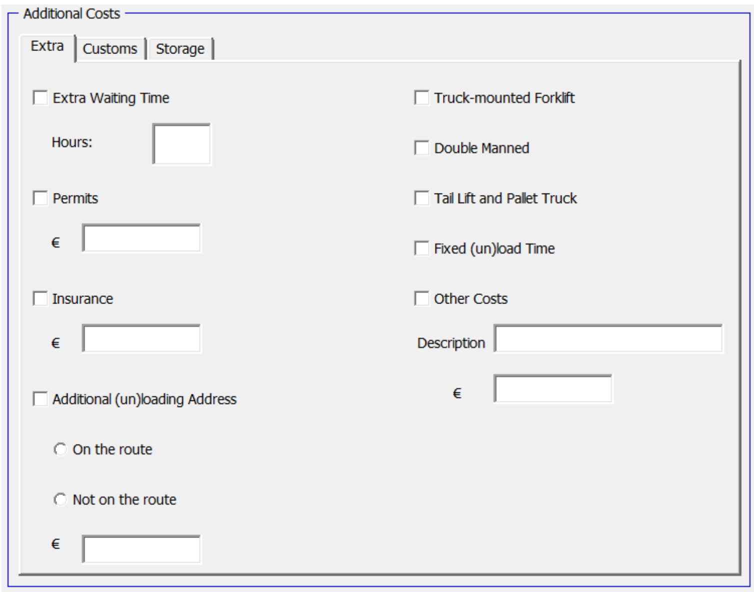


Figure 29: Input Form for Extra Costs

- **<u>Customs costs: (the form for the input parameters can be found in Figure 30)</u>**
  - Export document = € 55,-
  - Import document = € 55,-
  - T1/T2 document = € 85,-
  - CFO document = € 70,-
  - EUR1 document = € 70,-
  - T2L document = € 40,-

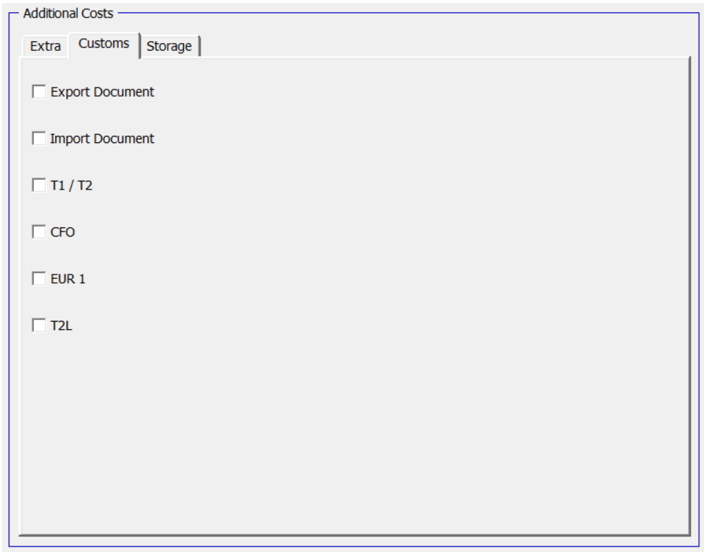


Figure 30: Input Form for Customs Costs

- **<u>Storage costs: (the form for the input parameters can be found in Figure 31)</u>**
  - Pallet = € 5.50,-
  - Container = € 2,-
  - Other storage costs = Depends on the situation. The planner must be able to enter this cost price, with a clear description.

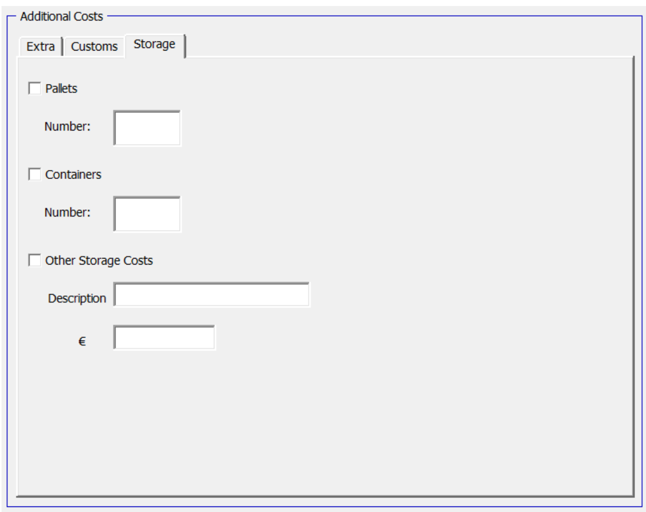


Figure 31: Input Form for Storage Costs

Finally, all additional costs are added to the freight rate. This are the total costs. The formula for the total costs is Equation 3 in Section 3.4. When the employee presses the 'Calculate' button, all costs with the corresponding cost price and description are displayed on the home page. In addition, the total purchase price can be seen at the bottom of the form. An example of what the form on the home page looks like after a shipment with additional costs has been calculated can be seen in Figure 32.

Purchase:		
Full Loading Postal Code:		**85232**
Full Unloading Postal Code:		**1119 NR**
Type of Costs:	**Amount:**	
Shipping Costs with ADR		€ 891.00
Additional (un)loading Address		€ 75.00
Truck-mounted Forklift		€ 125.00
Double Manned		€ 600.00
Customs Costs:		
T1/T2 Document		€ 85.00
Storage Costs:		
Pallets Storage	5	€ 27.50
Total Costs:		**€ 1,803.50**

Figure 32: Example of what the Form on the Home Page looks like after a Shipment with Extra Costs has been Calculated

When the employee presses the 'Save' button, separate lines with the corresponding CRType are created for the additional costs. Figure 33 shows how the example of Figure 32 is stored in the database.

OrderNO	CRType	LoadingCountryCode	UnloadingCountryCode	LoadingPostalCode	UnloadingPostalCode	LoadingMeter	ServiceLevel	ADR	PurchasePrice
6301	TRANS	NL	DE	7559 SH	52372	0.4	Standard	N	100.00
6302	TRANS	DE	NL	75417	7575 EC	1.6	Standard	N	150.00
6303	TRANS	DE	NL	59597	7575 EC	0.5	Standard	N	70.93
6304	TRANS	DE	NL	37520	7575 EC	1.5	Standard	N	200.00
6305	TRANS	DE	NL	88239	1437 EK	5.5	Standard	N	500.00
6306	TRANS	DE	NL	94577	1118 LA	13.6	Standard	N	950.00
6307	TRANS	DE	NL	85232	1119 NR	13.6	Standard	J	891.00
6307	EXTL	DE	NL	85232	1119 NR	13.6	Standard	J	75.00
6307	TMK	DE	NL	85232	1119 NR	13.6	Standard	J	125.00
6307	DM	DE	NL	85232	1119 NR	13.6	Standard	J	600.00
6307	DOCU	DE	NL	85232	1119 NR	13.6	Standard	J	85.00
6307	OPSL	DE	NL	85232	1119 NR	13.6	Standard	J	27.50

Figure 33: Example Storage of Additional Costs

6.8 Summary of the Cost Price Model Process

Now that all the different parts of the cost price model have been explained separately, this chapter provides a summary of the cost price model process as a whole. Figure A.1 in Appendix A shows the business process model of the cost price model process. The three different lanes in the model are the three buttons of the cost price model tool. We are now going to describe this business process model.

Calculate

When the employee presses the 'Calculate' button, the model checks whether all input parameters required for calculating the cost price have been entered. If not all parameters have been entered, the tool will show a message box with the following text: "Please enter a value in the text box(es) that are colored Red" and the text boxes that have not been filled in will colour red. However, if all parameters are entered, the calculating process starts. First, the case-based reasoning process is activated. Based on the input parameters, the cost price model searches for similar shipments in the database of SCS Logistics. If the database contains similar shipments, the average freight rate of these shipments is calculated. If the database does not contain any shipment that is the same, the cost price model links the right zones to the loading and unloading postal codes. With the help of these zones the correct mathematical equation can be found that belongs to the combination of the loading and unloading zone. After the mathematical equation is known, the loading meter is entered in the equation. The solution of the equation is the freight rate. Now that the freight rate is known, the cost price model checks whether input parameters for calculating the additional costs have been entered. If there are additional costs they will be added to the freight rate, and so the total purchase price is calculated. Now all costs have been calculated and a total purchase price is issued. The employees can use this purchase price to set up the bid price.

Save

When the employee presses the 'Save' button, the cost price model checks whether the total purchase price has been calculated or not. If the total purchase price has not been calculated, the tool will display a message box with the following text: "The Purchase Price is not Calculated – First press the Calculate Button". However, if the total purchase price has been calculated, the cost price model checks whether the full loading and unloading postal code has been entered. If these have not been entered, the tool will display a message box with one of the two texts: "The Full Loading Postal Code has not been Entered" or "The Full Unloading Postal Code has not been entered been Entered". Which text the message box shows depends on which postal code is not filled in. When both postal codes are entered, the cost price model creates a new order number for the shipment and the matching CRTypes are added to the costs. After that, the different cost rules are stored in the SCS Logistics database. Finally, all completed forms will be cleaned, so that the employees can create the next order.

Reset

When the employee presses the 'Reset' button, all boxes in the forms that have been filled in are cleaned.

7 Evaluation, Implementation and Recommendations

In this chapter we are going to evaluate on the cost price model and tool, and provide the implementation and corresponding recommendations. Section 7.1 evaluates on the performance of the cost price model and tool. In addition we show how the cost price model could be implemented within the current TMS and give recommendations about the implementation in Section 7.2. Lastly, in Section 7.3 the conclusion is given.

7.1 Evaluation

In this section we evaluate on the performance of the tool. According to Peffers et al. (2007) we must first demonstrate the artifact by solving one or more instances of the problem. We do this by testing the tool with new incoming data and having four employees of SCS Logistics use the tool for one week. After that, we observe and measure how well the artifact supports a solution to the problem. Evaluation involves comparing the objectives of a solution to actual observed results from use of the artifact in the demonstration (Peffers et al., 2007). In order to test the performance of the solution we evaluate on the following two things:

1. Forecast accuracy
2. A survey performed for four stakeholders at SCS Logistics.

In Section 7.1.1 we assess the accuracy of the cost price model with the use of forecast error measurements. After that we conduct a survey for the four employees who have used the tool. The survey with the outcomes is discussed in Section 7.1.2.

7.1.1 Forecast Accuracy

In order to estimate the quality of the forecast methods and for choosing the best forecasting mechanism in case of multiple objects, forecast error measurements are used (Shcherbakov et al., 2013). We executed a literature study to get more insights in the various forecast error measurements. First, we discuss and choose different forecast error measurements. After that we assess the accuracy of the cost price model using these forecast error measurements.

Forecast Error Measurements

According to Witten & Frank (2005) several measures can be used to evaluate the success of numeric prediction. The principal and most commonly used error measure is the *Mean-Squared Error* (MSE) (Witten & Frank, 2005). The formula of the MSE is:

$$Mean - Squared\ Error = \frac{(p_1 - a_1)^2 + \cdots + (p_n - a_n)^2}{n} \quad (7)$$

Where $p_1, p_2, \ldots, p_n$ are the predicted values on the test instances and $a_1, a_2, \ldots, a_n$ are the actual values. To give the error measure the same dimensions as the predicted value itself, the square root is taken. This is also called the *Root Mean-Squared Error* (RMSE). Both the MSE and the RMSE tend to exaggerate the effect of outliers. The outliers are the instances whose prediction error is larger than the others (Witten & Frank, 2005). However, the *Mean Absolute Error* (MAE) does not have this effect. With the MAE all sizes of error are treated evenly according to their magnitude. The MAE averages the magnitude of the individual errors without taking account of their sign. The formula used for the MAE is:

$$\text{Mean Absolute Error} = \frac{|p_1 - a_1| + \cdots + |p_n - a_n|}{n} \quad (8)$$

In some situations it is the *relative* error rather than *absolute* error values that are of importance. According to Witten & Frank (2005), such situations are for example when a 10% error is equally important whether it is an error of 50 in a prediction of 500 or an error of 0.2 in a prediction of 2. In these kind of situations the averages of absolute errors will be meaningless and relative errors are appropriate.

The *Relative Squared Error* (RSE) refers to something quite different. It is made relative to what it would have been if a simple predictor had been used (Witten & Frank, 2005). This simple predictor is just the average of the actual values from the training data. So the RSE takes the total squared error and normalizes it by dividing by the total squared error of the default predictor. The RSE is computed as follows:

$$Relative\ Squared\ Error = \frac{(p_1 - a_1)^2 + \cdots + (p_n - a_n)^2}{(a_1 - \bar{a})^2 + \cdots + (a_n - \bar{a})^2}\ , where\ \bar{a} = \frac{1}{n}\sum_i a_i \qquad (9)$$

The *Root Relative Squared Error* (RRSE) is just the square root of the RSE. The next error measure mentioned by Witten & Frank (2005) is the *Relative Absolute Error* (RAE) and is just the total absolute error, with the same kind of normalization. The formula used for the RAE is:

$$Relative\ Absolute\ Error = \frac{|p_1 - a_1| + \cdots + |p_n - a_n|}{|a_1 - \bar{a}| + \cdots + |a_n - \bar{a}|} \qquad (10)$$

In all the three relative error measures, the errors are normalized by the error of the simple predictor that predicts average values (Witten & Frank, 2005).

According to Shcherbakov et al. (2013) it is impossible to choose only one measure. This is due to the fact that each error measure has the disadvantages that can lead to inaccurate evaluation of the forecasting results. In the case of the same scale and the data pre-processing procedures were performed (data cleaning, anomaly detection), it is reasonable to choose MAE and RMSE (Shcherbakov et al., 2013). In addition, we would also like to know the relative error of the cost price model. Therefore, we choose the MAE, RMSE and RAE to assess the accuracy of the cost price model. Where the smaller the error measure, the more accurate the forecasting method is.

Assessing Accuracy

To assess the accuracy of the cost price model using the MAE, RMSE and RAE, we need new data. Therefore we took the data from week 22 in 2021. These are the transactions made from May 31 to June 6. This week was deliberately chosen because there were many public holidays in the weeks before, such as Ascension Day and Pentecost. That is why the purchase prices in the month of May are always more expensive. From week 22, the purchase prices reflect the normal circumstances again. For testing the accuracy of the tool, only the transactions from the Netherlands to Germany and from Germany to the Netherlands with a standard service level are taken. This is because the cost price model has now only been developed for such kind of shipments. In total there are 198 transactions, but all these transactions are only from the Netherlands to Germany. To make the assessment as reliable as possible, we use two random transactions per combination between the zones from the Netherlands to Germany. These transactions are then checked for correctness by looking them up in the TMS. After all transactions have been checked, it appears that the combinations NL2 – DE3, NL2 – DE6 and NL2 – DE9 have only one correct transaction. Because there are 18 combinations (9 Zones in DE & 2 Zones in NL), 33 transactions are used to calculate the MAE, RMSE and RAE. The details of these transactions can be found in Table C.1 in Appendix C. After calculating the MAE, RMSE and RAE, we get the following outcomes:

$$MAE = 15.15$$

$$RMSE = 20.70$$

$$RAE = 0.079$$

The MAE is the average of the absolute values of the forecast errors. This means that the average price deviation of the tool compared to the 33 test transactions is €15.15. At the MAE all individual differences between the prediction and actual observation have equal weight. In contrast to the RMSE. This is because with the RMSE the errors are squared before they are averaged. As a result, the RMSE gives relatively high weight to large errors. So it makes sense that the deviation given by the RMSE is slightly higher than that of the MAE. Furthermore, the RAE = 0.079, which is a relative deviation of 7.9%.

In addition to these forecast error measurements, we would also like to know whether a trend is noticeable in the prices calculated by the cost price model. For example, are the predicted prices always higher or lower, or is there a balance? To find out, we have shown the forecast errors of the 33 test transactions in a bar chart in Figure 34.

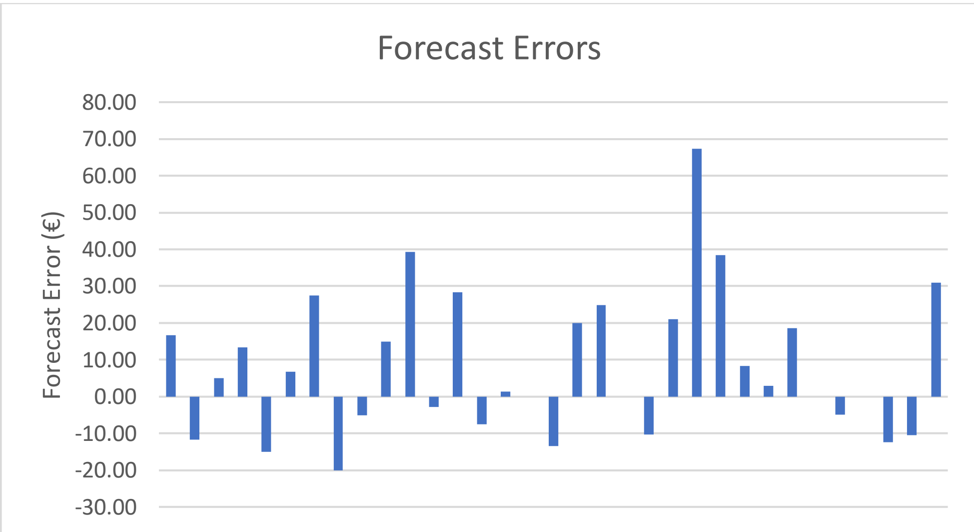


Figure 34: Forecast Errors of the 33 Test Transactions

As can be seen in Figure 34, the cost price model generally gives a purchase price that is too high than too low. Of the 33 test transactions, the predicted purchase prices are 18 times higher and 11 times lower than the actual purchase prices. In addition, the predicted and actual purchase prices correspond exactly 4 times. Figure 34 shows that there is one reasonable outlier. Here the forecast error is € 67.30. This is because this shipment was loaded in the Netherlands with postal code 96 (Zone 2) and unloaded in Germany with postal code 26 (Zone 3). There are no shipments in the database of the cost price model that go from the top in the Netherlands to Zone 3 in Germany. That is why the cost price model gives a purchase price that is more appropriate for shipments from the middle of Zone 2 in the Netherlands to Zone 3 in Germany. However, the distance from the middle of Zone 2 in the Netherlands to Zone 3 in Germany is longer. For this reason, the cost price model gives a higher purchase price than the actual purchase price.

During the division of the zones in the Netherlands, it appeared that there were few shipments from the top of the Netherlands to Germany. That is why the employees have to be extra alert with these types of shipments, since the purchase prices issued by the cost price model are not always accurate for these shipments.

7.1.2 Survey

Now that we have tested the cost price model for accuracy, it is important to assess the tool for long-term use. The results of this survey can therefore be used for the recommendations and implementation of the cost price model within the TMS, which is described in Chapter 7.2. We select four stakeholders who will use the tool in week 22 of 2021. These four employees have been chosen in such a way that we have participants from both departments where the tool will be used, and the participants have different functions. After the four stakeholders have used the tool for a week, they have to evaluate by filling in the survey. The four stakeholders who completed the survey are from the following departments:

- **Operations department:**
  Two people of the operations department, including one planner and head of the operations departments

- **Sales department:**
  Two people from the sales department, including a sales employee and the CEO

Focus Points

To assess the performance of the tool, the survey focuses on the following points:

- Applicability
- User-friendliness
- Speed
- Reliability

Statements in the Survey

The survey consists of the following statements:

1. The prices correspond to what I had in mind
2. When I use the tool I feel more confident in giving the prices to the customers
3. With the help of the tool I can give the bid prices to the customers faster
4. I will use the tool when it is implemented within the TMS
5. I would like to be able to use the tool for more countries and service levels in the future
6. I still miss the following parts within the tool

For statements 1 to 5, the stakeholders must rate on a scale from 1 to 5. The grades respectively represents the following definitions: 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree. In addition, the stakeholder can provide an additional remark for each statement. If the stakeholder is still missing things in the tool, he or she can describe this in statement 6. In Appendix D the survey with the complete answers of the statements are provided.

Outcomes

After the surveys have been completed, we analyse the outcomes. The results of the first five statements are shown in Table 10. For the first five statements, we display the maximum and minimum values, and calculate the average grade and the standard deviation between the grades. By calculating the standard deviation, it can be shown how much the opinions of the stakeholders differ. The lower the standard deviation, the less the opinions of the employees differ.

The first statement indicated whether the calculated prices of the cost price model correspond to the prices that the employees had in mind. Three out of four employees answered the statement with ‘agree’ and one employee answered with the opinion ‘strongly agree’. These answers indicates that the calculated prices of the cost price model correspond well with the prices that the employees would use. The employees stated that the cost price model can work well as support. In the test week the tool worked as a good indication/target price. However, the employees were also aware that the input data had to be more accurate and that they had to be extra careful with the seasonal products.

With the second statement, we would like to know whether the employees feel more confident in giving the prices to the customers when using the tool. All employees answered the statement with a 4, so the standard deviation is zero. From this we can conclude that the tool improves the self-confidence of the employees. An employee also indicated that the tool is also very useful for new employees with little experience. With the help of the tool, the new employee gets an idea of what the purchase prices are faster.

The average score of the third statement is a bit higher than the second one. This is because one employee gave a 5 instead of a 4. With the third statement, we wanted to know whether the tool ensured that the employees could create bid prices more quickly. One of the action problems was namely, ‘sometimes it takes a long time to determine the bid prices’. Since the average answer to this question is a 4.25, this action problem can certainly be solved with the tool.

With the fourth statement we wanted to find out whether the employees will actually use the cost price model if it is implemented within the TMS. The results show that two employees answered the question with ‘strongly agree’, one with ‘agree’, one with ‘neutral’. The employee with the answer ‘neutral’ wondered how much it will cost to implement and whether it is really necessary to implement the model. The average answer for this question is a 4.25. From this we can conclude that the employees will start using the cost price model when it is implemented within the TMS.

Currently, the cost price model has only been developed for shipments from the Netherlands to Germany and from Germany to the Netherlands with a standard service level. With the fifth statement we want to measure whether the employees would like to use the cost price model for more countries and service levels. The average answer to the fifth statement is 4.25. The employees indicated that they would like to use it for more countries, but not necessarily for every country. They stated that every country is different. For example, an employee indicated that he would like to use the model for the countries Germany, Italy, Sweden and Denmark.

The last statement is an open statement. The statement is, ‘I still miss the following parts within the tool’. Almost every stakeholder indicated that, besides more countries and service levels, nothing else was missing. However, the CEO of SCS Logistics wants to further develop the tool, so that a full quotation can be automatically made per country and postal code.

To conclude, the survey shows that the prices of the cost price model correspond well to the prices that the employees had in mind. In addition, the employees would like the tool to be implemented within the TMS. Because the cost price model, helps the employees to create bid prices faster and more confidently. Therefore the model offers the employees good support in determining the bid price. Table 10 shows that the average score of the statements is 4.2, which represents an opinion between ‘agree’ and ‘strongly agree’. The standard deviation is 0.58. This means that the answers of the stakeholders differ on average 0.58 points from the average 4.2. So on average there are no big variations between the opinions of the employees.

Table 10: Outcomes survey performed by four stakeholders at SCS Logistics

Statement	Min Value	Max Value	Average	Standard Deviation
1	4	5	4.25	0.5
2	4	4	4	0
3	4	5	4.25	0.5
4	3	5	4.25	0.96
5	3	5	4.25	0.96
Average Performance	**3.6**	**4.8**	**4.2**	**0.58**

7.2 Implementation and recommendations

In this section recommendations are made for implementing the cost price model at SCS Logistics. As described in the desired situation in Section 2.2 and in Section 6.1, one of the requirements of SCS Logistics is that the cost price model must be implemented within the current TMS. The results of the survey also show that the employees would like to see this. That is why, after conducting interviews with the employees, we indicate in paragraph 7.2.1 where in the system the cost price model can best be implemented. After that we provide recommendations for the implementation in Section 7.2.2. In this way we try to ensure that both SCS Logistics and the software developer know which adjustments must be made within the TMS and the business operations.

7.2.1 Implementation

The planning of the transport orders is all arranged with the TMS of SCS Logistics. The home page of the TMS can be seen in Figure 35. By conducting interviews with the employees, we came to the conclusion that calculating the purchase price should be possible directly on the homepage. In this way the purchase price can be calculated the fastest with the fewest possible actions. To implement this, a button ‘Inkoopprijs’ must be added on the home page. When the employee presses the ‘Inkoopprijs’ button, the developed form for calculating the purchase price will appear, as shown in Figure 35. However, in this form the ‘Save’ button is missing, which does contain the form in Figure 17 in Section 6.4. This is because in the TMS the transport orders are registered via the ‘Order’ button at the top of Figure 35. When the orders are registered, the shipment details are immediately added to the database and the order is given an order number. The recommendations regarding the implementation are discussed in the next section.

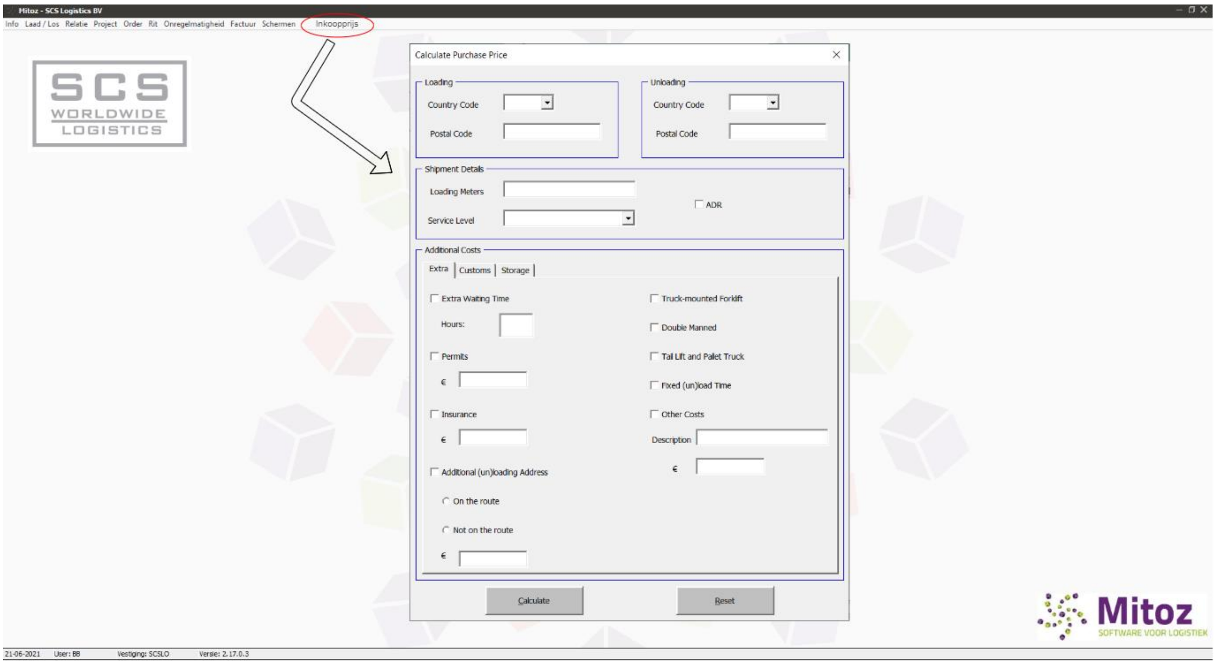


Figure 35: Home Page of the TMS, with the New Button 'Inkoopprijs'

7.2.2 Recommendations

In this section, recommendations regarding the implementation and overall findings during the research are given.

- During the research, especially while exploring and preparing the data in Chapter 5, we found that sometimes there are disagreements between employees regarding the service levels. In order for the cost price model to provide purchase prices as accurately as possible, the employees must all be aligned. For this reason, we have scheduled a meeting with several employees. We came to the conclusion that the service levels had to be rearranged, so that everyone knows exactly which shipment belongs to which service level. In consultation with the employees, we propose to divide the service levels into the following types, as shown in Table 11.

Table 11: The New Service Levels

Service level	
Standard (complete or groupage)	This means that the chosen service level is not special. In general, these are the standard and groupage shipments within Europe and Great Britain.
Groupage-express	Groupage express means that the shipment is transported as a partial batch, but within a faster transit time than normal groupage. This has mainly to do with the price-technical aspect.
Express	These are shipments for which one complete truck or van has been rented. This truck or van therefore contains no other shipments. This is because the transit time is very important. For express shipments, think of small vans or double-manned vans/trucks.
Exceptional	Exceptional means that all shipments are driven with special transport. So, with flatbed trailer, lowbed trailer, semi-lowbed trailer, extendable trailer, crane car, container chassis and other vehicles for special transport.

Temperature control	These are shipments that must be transported within a certain temperature. This can be both groupage and express, but falls under the service level temperature control
Air freight	These are all air freight shipments that are transported between two airports.
Sea freight	These are all sea freight shipments that are transported between two ports.
Rail transport	These are all freights that are transported by train.

- We suggest showing message boxes within the system when certain shipment details, needed to calculate the purchase price, are not filled in when saving the order. In this way you avoid missing certain values in the database.

- In addition, we recommend that SCS Logisitcs automates the linking of a CRType to the costs, as is done in the developed tool. This ensures that, for example, there are no more storage costs under the CRType of transport costs. Currently, linking the CRType to the costs is done manually leading to unnecessary mistakes.

- When the cost price model is going to be used for other countries and/or service levels, it is very important to keep in mind that the transactions in the current database contain a lot of errors. Therefore, we propose to remove all outliers per country according to the approach described in Chapters 5 and 6 or to start over by filling in the database correctly.

- When more transactions becoming available, the zones of the countries can be made smaller. This makes the calculated freight rates more accurate.

- We recommend using the calculated purchase prices of the cost price model as an indication and a guideline. It is important that the employees always have a critical look at the purchase price calculated with the tool. This is because they sometimes have to deal with seasonal periods and holidays, which means that the purchase prices are higher than normal. In addition, the feeling and experience of the employees cannot be captured in a tool, but this is precisely the strength of SCS Logistics.

7.3 Conclusion

In this chapter we evaluated the cost price model and discussed the implementation with corresponding recommendations. In order to assess the accuracy of the cost price model, we tested the tool with new incoming data. Out of literature review, we concluded that we use the forecast error measurements MAE, RMSE and RAE to assess the accuracy of the model. The results are that the MAE = 15.15, the RMSE = 20.70 and the RAE = 0.079. In addition, we found that for the 33 test transactions, the tool more often quoted a purchase price that was too high than too low. We also concluded that employees should be extra alert for shipments to and from the top of the Netherlands, since the calculated purchase prices of the cost price model often deviate slightly more from the actual purchase price for these type shipments.

After the cost price model was tested for accuracy, we assessed the tool for long term use. We did this by having four employees use the tool for a week, after which they had to fill in a survey. The survey showed that the cost price model helps the employees to create the bid prices faster and more confidently. For these reasons, the employees would like to have the cost price model implemented in the TMS. To ensure a smooth implementation, we discussed how the tool can best be implemented within TMS. In addition, several recommendations regarding the implementation were discussed.

8 Conclusions, Discussion and Future Research

In this chapter the conclusions, discussion and future research is given. First, we conclude in Section 8.1 on the process of developing the cost price model and its contribution to the bid price estimation. Second, Section 8.2 discusses the limitations of the cost price model. Lastly, we give suggestions for possible future research in Section 8.3.

8.1 Conclusions

In this section, we answer the main research question this research addresses:

"How to give the employees of SCS Logistics a standard procedure in determining the bid price?"

The answer to this research question is based on the answers of the eight sub-research questions given in Section 1.5.2.

After it was clear that the employees of SCS Logistics do not have a standard procedure for determining the bid prices, we started to analyse the current purchasing and sales process, by means of an observational study of the TMS and conducting interviews. Both processes were made transparent with the help of business process models. The model of the current situation gave us a better picture of where the problems and challenges were within the process. Based on this, the desired situation was mapped out. By using the business process model of the desired situation, it became clear that estimating the purchase price within the sales process should be automated in the TMS, using a cost price model.

For the development of the cost price model, it is important to map out the different cost drivers that determine the purchase price. Therefore, we determined, classified and described the cost drivers using interviews with the employees and an observational study of the TMS. The cost drivers were classified into the following types: shipment details, service level and additional costs. Out of the literature review the cost estimation methods: case-based reasoning process, regression analysis models and expert systems were selected. The combination of these three methods forms the cost price model. The input for the cost price model was created from initial raw data from the database of SCS Logistics. But during the data preparation we found out that the data contains a lot of transactions with missing and/or inconsistent values, which are not easy to find and therefore to remove. So it takes a lot of time to correct or delete all these transactions. For this reason, the cost price model is developed for shipments from the Netherlands to Germany and from Germany to the Netherlands with a standard service level, since these shipments are the most common. In order to make this cost price model usable for the employees, we implemented the model in a user-friendly tool, using *Microsoft Excel* and *Visual Basic for Applications*.

We tested the performance of the cost price model with selected forecast error measurements from the literature and a survey completed by four involving stakeholders. The accuracy of the calculated purchase prices of the tool has been assessed using new incoming data. The results are that the MAE = 15.15, RMSE = 20.70 and the RAE = 0.079. This means that the average deviation between the calculated purchase price of the tool and the actual purchase price is approximately €15.15. And when the outliers are weighted more heavily, the deviation is €20.70. Besides that the relative deviation of the cost price model is 7.9%. Furthermore, we found that for the 33 test transactions, the tool more often quoted a purchase price that was too high than too low. Lastly, we concluded that employees should be extra alert for shipments to and from the top of the Netherlands, since the calculated purchase prices of the cost price model often deviate slightly more from the actual purchase price for these type shipments.

According to the outcomes of the survey, the deviation of the cost price model is accurate enough for determining the bid prices at SCS Logistics. The employees found the calculated purchase prices issued by the tool a good target price. It was also concluded from the survey that the employees creates bid prices faster and more confidently when using the cost price model. Therefore, they would like to have the cost price model implemented in the TMS. In that case, the model can provide a good indication and support in determining the bid prices.

It can finally be concluded that if the recommendations are followed, the cost price model can solve the two action problems in the future. Determining the bid prices can therefore be done faster and more accurately using the *cost price model*.

8.2 Discussion

This section discusses the limitations of the research. It should be noted that the cost price model provides estimated purchase prices based on previously quoted prices. This results in the following limitations of the cost price model:

- The cost price model calculates the freight rates by taking averages of the freight rates of similar historical shipments and if that is not possible it uses regression analysis. As a result, the cost price model does not give higher purchase prices during seasonal periods and holidays, when prices are higher than normal.

- When using the case-based reasoning process, it can be possible that no similar shipments are found in the database. In that case, regression analysis is used to calculate the freight rate. In order to use regression analysis enough data points must be available. For this reason, and for the reason that the distance must remain approximately constant to determine the relationship between the loading meter and purchase price, the Netherlands and Germany are divided into zones. However, the Netherlands cannot be divided into more than 2 zones. Therefore a discussion can be made whether the distance is still kept constant.

- Another point of discussion about the zone division in the Netherlands is that from the 1804 shipments to zone 2 in the Netherlands 1404 shipments go to the 75 postal code. This is because SCS Logistics stores many shipments from Germany in the warehouse in Oldenzaal. These are mainly small shipments where the transit time is not important. Therefore, it gives the employees more time to find a cheap supplier. So because many shipments go to zone 2 in the Netherlands with a generally lower purchase price, the average purchase price in zone 2 is reduced. As a result, the cost price model may give a slightly too low purchase price for shipments in the north or south of the Netherlands in Zone 2.

- While identifying and removing or correcting the outliers, we found out that the data contains a lot of outliers that are difficult to find and therefore also to remove or correct. So even though the input data has been thoroughly identified, it is always possible that the input data still contains a few outliers.

8.3 Future Research

This research has been executed within 10 weeks. Therefore extensive research was not possible. In this section we describe possible future research, so that the cost price model can be more accurate and expanded in the future.

- The survey shows that the employees would also like to use the tool for other countries and service levels. It could be investigated for which countries the cost price model should be expanded. If sufficient and correct data is available, the cost price model can be applied to any country. Besides that the cost price model could be expanded with new countries, the model could also be expanded with new service levels.

- In the current cost price model, seasonal periods and holidays are not taken into account. Further research can be done in order to investigate when all these periods and holidays take place. Using neural networks, the influence of certain holidays and seasonal periods on the purchase price could be determined. Weightings can then be attached to the different periods and holidays by means of neural networks. By building a check box for holidays and seasonal influences in the TMS, purchasing prices can be increased based on this weights.

- Using neural networks, it could be estimated for various suppliers how high their purchasing price is approximately. Because the neural network is a self-learning system, these estimated purchase prices are becoming increasingly accurate. Based on this, we already have an estimated purchase price for various suppliers that is close to the actual purchase price before requesting.

References

Brown, A. M. (2001). A step-by-step guide to non-linear regression analysis of experimental data using a Microsoft Excel spreadsheet. *Computer Methods and Programs in Biomedicine*, *65*(3), 191–200. https://doi.org/10.1016/s0169-2607(00)00124-3

Bokor, Z. (2010). Cost drivers in transport and logistics. *Periodica Polytechnica Transportation Engineering*, *38*(1), 13. https://doi.org/10.3311/pp.tr.2010-1.03

Boehm, B., Abts, C., & Chulani, S. (2000). Software development cost estimation approaches – A survey. *Annals of Software Engineering*, *10*(1/4), 177–205. https://doi.org/10.1023/a:1018991717352

Busachi, A., Erkoyuncu, J., Colegrove, P., Martina, F., Watts, C., & Drake, R. (2017). A review of Additive Manufacturing technology and Cost Estimation techniques for the defence sector. *CIRP Journal of Manufacturing Science and Technology*, *19*, 117–128. https://doi.org/10.1016/j.cirpj.2017.07.001

Büyüközkan, G., Feyzioğlu, O., & Şakir Ersoy, M. (2009). Evaluation of 4PL operating models: A decision making approach based on 2-additive Choquet integral. *International Journal of Production Economics*, *121*(1), 112–120. https://doi.org/10.1016/j.ijpe.2008.03.013

Caetano, A., Pereira, C., & Sousa, P. (2012). Generation of Business Process Model Views. *Procedia Technology*, *5*, 378–387. https://doi.org/10.1016/j.protcy.2012.09.042

Elfaki, A. O., Alatawi, S., & Abushandi, E. (2014). Using Intelligent Techniques in Construction Project Cost Estimation: 10-Year Survey. *Advances in Civil Engineering*, *2014*, 1–11. https://doi.org/10.1155/2014/107926

Elmousalami, H. H. (2020). Artificial Intelligence and Parametric Construction Cost Estimate Modeling: State-of-the-Art Review. *Journal of Construction Engineering and Management*, *146*(1). https://doi.org/10.1061/(asce)co.1943-7862.0001678

Elmousalami, H. H. (2021). Comparison of Artificial Intelligence Techniques for Project Conceptual Cost Prediction: A Case Study and Comparative Analysis. *IEEE Transactions on Engineering Management*, *68*(1), 183–196. https://doi.org/10.1109/tem.2020.2972078

Heerkens, H. and Winden, A. V. (2017). Solving Managerial Problems Systematically. Noordhoff, 1 edition. isbn: 987-90-01-88795-7.

Hu, X., Xia, B., Skitmore, M., & Chen, Q. (2016). The application of case-based reasoning in construction management research: An overview. *Automation in Construction*, *72*, 65–74. https://doi.org/10.1016/j.autcon.2016.08.023

Huang, X. X., Newnes, L. B., & Parry, G. C. (2012). The adaptation of product cost estimation techniques to estimate the cost of service. *International Journal of Computer Integrated Manufacturing*, *25*(4–5), 417–431. https://doi.org/10.1080/0951192x.2011.596281

John, E. G. (1998). Simplified Curve Fitting using Spreadsheet Add-ins. *International Journal of Engineering Education*, *14*(5), 375–380.

Laurikkala, J., Juhola, M., & Kentala, E. (2000). Informal identification of outliers in medical data. In N. Lavrac, S. Miksch, & K. Branko (Eds.), *Workshop Notes of the 14th European Conference on Artificial Intelligence (ECAI-2000): The Fifth Workshop on Intelligent Data Analysis in Medicine and Pharmacology (IDAMAP-2000)* (pp. 20-24) http://www.ifs.tuwien.ac.at/~silvia/idamap-2000/

Niazi, A., Dai, J. S., Balabani, S., & Seneviratne, L. (2006). Product Cost Estimation: Technique Classification and Methodology Review. *Journal of Manufacturing Science and Engineering*, *128*(2), 563–575. https://doi.org/10.1115/1.2137750

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, *24*(3), 45–77. https://doi.org/10.2753/mis0742-1222240302

Saeed, N. (2013). Cooperation among freight forwarders: Mode choice and intermodal freight transport. *Research in Transportation Economics*, *42*(1), 77–86. https://doi.org/10.1016/j.retrec.2012.11.005

Salmi, A., David, P., Blanco, E., & Summers, J. (2016). A review of cost estimation models for determining assembly automation level. *Computers & Industrial Engineering*, *98*, 246–259. https://doi.org/10.1016/j.cie.2016.06.007

Saleem, M. A., Ahmad, R., Alyas, T., Idrees, M., A., Farooq, A., Shahid, A., & Ali, K. (2019). Systematic Literature Review of Identifying Issues in Software Cost Estimation Techniques. *International Journal of Advanced Computer Science and Applications*, *10*(8), 341–346. https://doi.org/10.14569/ijacsa.2019.0100844

Shcherbakov, M. V., Brebels, A., Shcherbakova, N. L., Tyukov, A. P., Janovsky, T. A., & Kamaev, V. A. (2013). A survey of forecast error measuers. *World Applied Sciences Journal*, *24*, 171–176. https://doi.org/10.5829/idosi.wasj.2013.24.itmies.80032

Witten, I. H., & Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques, Second Edition (Morgan Kaufmann Series in Data Management Systems)* (2nd ed.). Morgan Kaufmann.

Yu Hao, A. C., Lai, C. H., Su, H. T., Then, Y. L., Tay, F. S., Hui Then, P. H., & Chen Yi, N. L. (2020). Cost estimation methods for internet infrastructure deployment in Rural Sarawak: a review. *Indonesian Journal of Electrical Engineering and Computer Science*, *19*(3), 1375–1383. https://doi.org/10.11591/ijeecs.v19.i3.pp1375-1383

Websites

Allyn International. (n.d.). What is a 4PL, 4th Party Logistics provider. Retrieved 18 March 2021, from https://logisticselearning.com/what-is-a-4pl/

Flatbed trailer | Flatbed trailer huren en leasen | TIP Trailer Services. (n.d.). Retrieved 8 June 2021, from https://www.tipeurope.nl/onze-vloot/trailers/flatbed-trailer

Low Bed Trailer with various combination options. (n.d.). Retrieved 8 June 2021, from https://www.nooteboom.com/low-bed-trailer-suitable-heavy-transport/

Pacton. (n.d.). Semi dieplader Archieven. Retrieved 8 June 2021, from https://www.pacton.nl/producten/semidieplader/

What Is a Tautliner Truck. (n.d.). Semi-Trailers Manufacturer, Semi Trucks for Sale | Three Horses Trailer. Retrieved June 9, 2021, from https://www.3horsestrailer.com/what-is-a-tautliner-truck/

A Process overview of the Cost Price Model

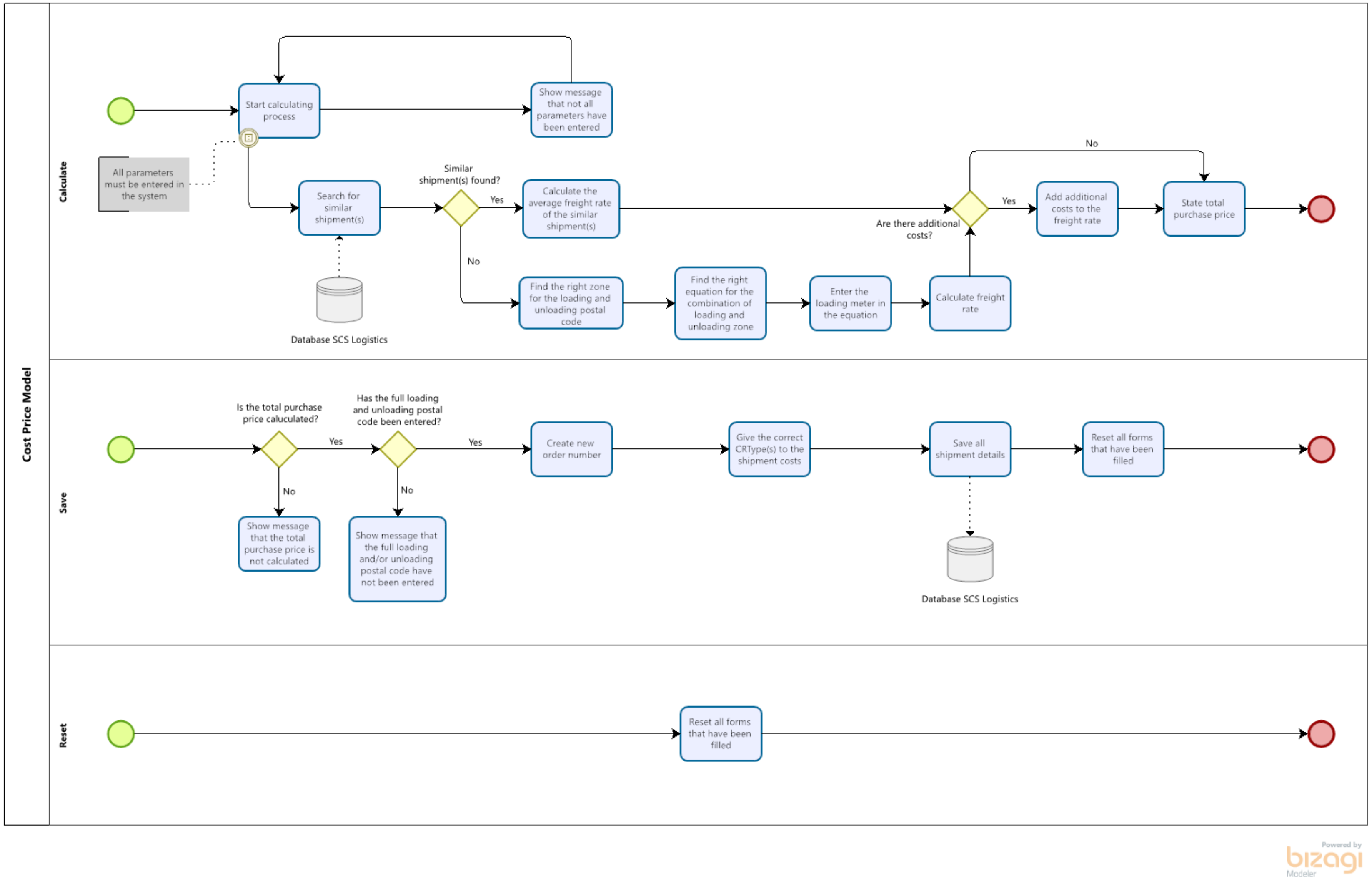


Figure A. 1: Process Overview of the Cost Price Model

B Description of the Fitted Lines between the Zones

Table B. 1: Description of the Fitted Lines between the Zones

Loading zone	Unloading zone	# Outliers Removed	# Transactions Remaining	Regression Type	Equation	R^2
NL1	DE1	5	135	Polynomial	$y = -0.9719x^2 + 34.104x + 111.41$	0.7575
NL1	DE2	6	122	Polynomial	$y = -0.4748x^2 + 35.415x + 119.54$	0.8287
NL1	DE3	11	152	Polynomial	$y = -2.0087x^2 + 57.492x + 122.88$	0.8251
NL1	DE4	11	261	Polynomial	$y = -2.2522x^2 + 68.106x + 126.58$	0.8605
NL1	DE5	7	166	Polynomial	$y = -2.0397x^2 + 63.282x + 141.44$	0.848
NL1	DE6	5	194	Polynomial	$y = -3.947x^2 + 103.72x + 116.58$	0.8253
NL1	DE7	1	183	Polynomial	$y = -1.2596x^2 + 71.954x + 141.97$	0.8963
NL1	DE8	3	223	Polynomial	$y = -1.8496x^2 + 81.635x + 143.27$	0.8711
NL1	DE9	11	336	Polynomial	$y = -3.6203x^2 + 105.97x + 115$	0.9237
NL2	DE1	7	215	Polynomial	$y = -1.5268x^2 + 40.49x + 101.23$	0.6302
NL2	DE2	1	75	Polynomial	$y = -2.2724x^2 + 54.098x + 84.166$	0.8487
NL2	DE3	0	62	Polynomial	$y = -0.9422x^2 + 31.581x + 162.07$	0.5836
NL2	DE4	4	243	Polynomial	$y = -2.4425x^2 + 57.659x + 131.01$	0.7814
NL2	DE5	5	132	Polynomial	$y = -3.0925x^2 + 77.389x + 110.54$	0.6744
NL2	DE6	2	97	Polynomial	$y = -4.1609x^2 + 101.78x + 119.9$	0.7053
NL2	DE7	5	171	Polynomial	$y = -2.9593x^2 + 87.991x + 95.661$	0.7853
NL2	DE8	4	142	Polynomial	$y = -5.7327x^2 + 116.25x + 118.2$	0.7398
NL2	DE9	5	81	Polynomial	$y = -3.6472x^2 + 95.573x + 139.05$	0.8455
DE1	NL1	16	206	Polynomial	$y = -1.265x^2 + 36.341x + 121.84$	0.7729
DE2	NL1	13	118	Polynomial	$y = -1.5551x^2 + 46.259x + 114.19$	0.9017
DE3	NL1	1	100	Polynomial	$y = -1.5485x^2 + 46.372x + 121.46$	0.77
DE4	NL1	12	234	Polynomial	$y = -1.7549x^2 + 52.543x + 110.98$	0.8558
DE5	NL1	27	223	Polynomial	$y = -1.3914x^2 + 49.555x + 138.21$	0.8478
DE6	NL1	12	114	Polynomial	$y = -2.2162x^2 + 61.387x + 128.56$	0.9066
DE7	NL1	4	156	Polynomial	$y = -0.6781x^2 + 48.07x + 156.88$	0.8946
DE8	NL1	9	197	Polynomial	$y = -2.7704x^2 + 80.971x + 134.61$	0.8992
DE9	NL1	1	164	Polynomial	$y = -2.7793x^2 + 87.265x + 130.97$	0.8705
DE1	NL2	7	174	Polynomial	$y = -1.8398x^2 + 44.107x + 72.532$	0.8269
DE2	NL2	16	229	Logarithmic	$y = 93.839\ln(x) + 155.96$	0.8579
DE3	NL2	1	26	Polynomial	$y = -2.3517x^2 + 58.667x + 66.285$	0.9502
DE4	NL2	9	183	Polynomial	$y = -0.678x^2 + 32.839x + 98.402$	0.7836
DE5	NL2	21	349	Polynomial	$y = -2.4932x^2 + 56.776x + 74.858$	0.8386
DE6	NL2	12	165	Polynomial	$y = -2.6106x^2 + 67.218x + 78.351$	0.8627
DE7	NL2	19	203	Polynomial	$y = -4.5101x^2 + 91.206x + 59.858$	0.8834
DE8	NL2	63	391	Polynomial	$y = -1.0668x^2 + 64.354x + 75.635$	0.8497
DE9	NL2	8	84	Polynomial	$y = -1.6015x^2 + 68.932x + 75.36$	0.9217

C Details of the Test Data

Table C. 1: Details of the Test Data

Loading Zone	Unloading Zone	Loading PostalCode	Unloading PostalCode	Loading Meter	ServiceLevel	ADR	Actual PurchasePrice (A)	Forecast (F)	Forecast Error (FE = (F - A))	Absolute FE	FE^2
NL1	DE1	3542 AR	48455	13.6	Standard	N	€ 275.00	€ 291.67	16.67	16.67	277.89
NL1	DE1	1118 DT	49401	0.5	Standard	N	€ 140.00	€ 128.22	-11.78	11.78	138.77
NL1	DE2	3147 PA	51149	0.8	Standard	N	€ 120.00	€ 125.00	5.00	5	25.00
NL1	DE2	3984 LR	51545	13.6	Standard	N	€ 500.00	€ 513.37	13.37	13.37	178.76
NL1	DE3	3147 PA	26655	8.0	Standard	N	€ 415.00	€ 400.00	-15.00	15	225.00
NL1	DE3	1119 MB	28197	0.5	Standard	N	€ 100.00	€ 106.67	6.67	6.67	44.49
NL1	DE4	1118 DT	22335	4.5	Standard	N	€ 360.00	€ 387.45	27.45	27.45	753.50
NL1	DE4	3029 AA	20457	0.4	Standard	N	€ 120.00	€ 100.00	-20.00	20	400.00
NL1	DE5	2964 LN	63762	13.6	Standard	N	€ 630.00	€ 624.81	-5.19	5.19	26.94
NL1	DE5	3147 PA	37434	0.8	Standard	N	€ 260.00	€ 275.00	15.00	15	225.00
NL1	DE6	3984 LR	24601	6.8	Standard	N	€ 600.00	€ 639.37	39.37	39.37	1550.00
NL1	DE6	4421 SH	75031	13.6	Standard	N	€ 800.00	€ 797.13	-2.87	2.87	8.24
NL1	DE7	3147 PA	89567	6.0	Standard	N	€ 500.00	€ 528.35	28.35	28.35	803.72
NL1	DE7	3147 PA	39638	0.8	Standard	N	€ 200.00	€ 192.50	-7.50	7.5	56.25
NL1	DE8	4421 SH	86633	13.6	Standard	N	€ 910.00	€ 911.40	1.40	1.4	1.96
NL1	DE8	3945 PB	12487	0.8	Standard	N	€ 200.00	€ 200.00	0.00	0	0.00
NL1	DE9	2676 LS	93107	13.6	Standard	N	€ 900.00	€ 886.58	-13.42	13.42	180.10
NL1	DE9	3147 PA	88348	1.0	Standard	N	€ 175.00	€ 195.00	20.00	20	400.00
NL2	DE1	7575 EJ	49811	4.4	Standard	N	€ 225.00	€ 249.83	24.83	24.83	616.53
NL2	DE1	7651 AX	48341	5.0	Standard	N	€ 250.00	€ 250.00	0.00	0	0.00
NL2	DE2	6546 AS	50126	13.6	Standard	N	€ 410.00	€ 399.60	-10.40	10.4	108.16
NL2	DE2	7559SH	52372	0.4	Standard	N	€ 88.79	€ 109.85	21.06	21.06	443.52
NL2	DE3	9636 HT	26759	13.6	Standard	N	€ 350.00	€ 417.30	67.30	67.3	4529.29
NL2	DE4	7575 CA	21261	13.6	Standard	N	€ 425.00	€ 463.41	38.41	38.41	1475.33

NL2	DE4	7741 MJ	21107	5.0	Standard	N	€ 350.00	€ 358.24	8.24	8.24	67.90
NL2	DE5	5482 TK	63303	2.8	Standard	N	€ 300.00	€ 302.98	2.98	2.98	8.88
NL2	DE5	7547 RD	63303	3.5	Standard	N	€ 325.00	€ 343.52	18.52	18.52	342.99
NL2	DE6	7575 CA	24817	13.6	Standard	N	€ 600.00	€ 600.00	0.00	0	0.00
NL2	DE7	8153 RB	90765	13.6	Standard	N	€ 750.00	€ 744.99	-5.01	5.01	25.10
NL2	DE7	5047 RC	72108	2.0	Standard	N	€ 200.00	€ 200.00	0.00	0	0.00
NL2	DE8	7575 CA	16866	13.6	Standard	N	€ 575.00	€ 562.50	-12.50	12.5	156.25
NL2	DE8	7772 BJ	86653	13.6	Standard	N	€ 600.00	€ 589.42	-10.58	10.58	111.94
NL2	DE9	7141 JP	83339	1.0	Standard	N	€ 200.00	€ 230.98	30.98	30.98	959.76

D Survey to test the Performance of the Tool

In this Appendix the survey with the complete answers of the statements are provided. The average scores are calculated and the additional remarks are combined. The four stakeholders who completed the survey are from the following departments:

- **Operations department:**
  Two people of the operations department, including one planner and head of the operations departments

- **Sales department:**
  Two people from the sales department, including a sales employee and the CEO

Goal:
The aim of this survey is to get feedback on the performance of the tool. The ultimate goal is that the tool will be implemented within the TMS. The answers can be given in a range of 1 to 5 according to the following definitions:

1 = strongly disagree
2 = Disagree
3 = Neutral
4 = Agree
5 = Strongly Agree

Statement 1
The prices correspond to what I had in mind

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Average grade: *4.25*

Additional remark:

All employees agree that calculated prices from the cost price model correspond with the prices they had in mind. In addition, the CEO indicated that the input data must be more accurate. The sales employee noted that employees should be alert with seasonal products. These products can cause the purchase price to rise.

Statement 2
When I use the tool I feel more confident in giving the prices to the customers

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Average grade: *4*

Additional remark:

All employees really like the tool for support and indication. In addition, the head of the operations department noted that the tool can also be very useful for new employees with little experience.

Statement 3
With the help of the tool I can give the bid prices to the customers faster

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Average grade: *4.25*

Additional remark:

Everyone agreed on this

Statement 4
I will use the tool when it is implemented within the TMS

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Average grade: *4.25*

Additional remark:

The planner indicated that he was very much looking forward to using the tool. The tool is very useful for checking, he notes.

Statement 5
I would like to be able to use the tool for more countries and service levels in the future

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Average grade: *4.25*

Additional remark:

The head of the operations department stated that he would like to see the tool for the countries Germany, Italy, Sweden and Denmark.

Statement 6
I still miss the following parts within the tool

All employees would like the tool to be further expanded with more countries and service levels. In addition, the CEO stated that he wants to further develop the tool, so that a quote can be automatically made per country and postal code