MASTER

Similarity based forecasting of semiconductor technology demand
a case study at NXP Semiconductors

Marcelis, J.H.

Award date:
2020

Similarity Based Forecasting of Semiconductor Technology Demand

A Case Study at NXP Semiconductors

Joris Marcelis
0909548

A thesis presented for the degree of Master of Science

TUe Supervisors:
dr. ir. R.M. Dijkman
dr. A.M. Wilbik

NXP Supervisors:
J. Snijder
R.J. Westerduin

Department of Industrial Engineering & Innovation Sciences
Eindhoven University of Technology
Netherlands
12-05-2020

Abstract

In the semiconductor industry, long-term estimates of demand are required for capacity planning. This thesis evaluates procedures for forecasting the demand of semiconductor technologies, which are characterized by long life cycles. To this end, similarity based forecasting procedures are applied. These procedures forecast the demand of a technology utilizing the historical demand time series of preceding similar technologies.

This thesis proposes that these procedures can be structured into two steps: (1) finding similar technologies and (2) predicting demand using similar technologies. A literature review led to the identification of various approaches for both steps. Similar technologies could be found using clustering, selection algorithms or business experts. Forecasts could then be made with growth models, regression models or a prototype curve approach. All methods except the prototype curve approach were evaluated with a case study at NXP. In the experiment, the demand of 11 recent technologies was forecasted and a database of 55 historical technologies was used to find similar technologies.

The best forecasting accuracy was achieved when business experts identified similar technologies and forecasts were made with growth models. Moreover, the growth models resulted in a higher accuracy than regression models regardless of the approach used for finding similar technologies. The clustering and selection approach provided similar performance. The best forecasting procedure resulted in lower errors than the traditional demand estimates of NXP. The proposed procedure obtained a mean absolute percentage error of 38% , while the traditional estimates resulted in an error of 52%.

The forecasting procedures can be used to support the long-term planning process of NXP. The traditional estimation of demand is a labor intensive process which relies on many information sources that are difficult to incorporate in a model. Alternatively, the similarity based forecasting procedures provide a quick and accurate method using alternative demand information sources. Thus, they present a valuable tool to evaluate the traditional demand estimates.

Executive Summary

At NXP, long term estimates of wafer demand are required for making long term capacity decisions. The wafer production volumes of NXP are divided into wafer diffusion technology groups, which refer to the production process used to produce the wafers. Each diffusion technology has a life cycle characterized by a growth, mature and declining phase. The length of the life cycles varies between 3 and 30 years. The estimates of the demand of diffusion technologies contain some inaccuracies which possibly have large consequences for the effectiveness of planning decisions. This thesis hypothesises that the life cycle patterns of previous diffusion technologies could be used to improve the prediction of demand of a diffusion technology. Consequently, the goal of this project is:

"Find life cycle patterns in the wafer demand of semiconductor wafer technologies and use these patterns to improve long term wafer demand predictions"

A literature review was conducted to create an overview of relevant forecasting procedures. Next, a number of procedures were selected and implemented in the case study of NXP. The demand of 11 recent diffusion technologies was forecasted to evaluate the procedures and a database of 55 earlier technologies was used to search for similar technologies. The next segment of this summary discusses the answers to the research questions.

(1) How can similar technologies be identified by either a clustering or selection approach?

The clustering approach consists of two steps. First the procedures are clustered, then a procedure, such as a classification model, is applied which can assign new products to clusters. In the literature the following clustering methods were applied: k-means clustering, hierarchical clustering and fuzzy clustering. In addition, the following classification methods were used: decision tree classification, probabilistic neural networks and nearest neighbours search.

Alternatively, a selection algorithm was applied. One procedure used a correlation coefficient threshold to select similar technologies based on demand patterns. Another applied a nearest neighbour search which compared technologies based on descriptive criteria. Last, some studies let business experts identify similar technologies.

(2) How can the demand of a technology be predicted, using a similar technology or cluster of similar technologies?

Three prediction approaches were identified. One approach forecasted with growth models. Some studies estimated the growth models by taking the weighted parameters of earlier similar technologies. Others regressed the growth model parameters of earlier technologies against their descriptive criteria. Another procedure first estimated the growth model with the historical demand of the prediction target. Subsequently, another growth model was estimated with the data of similar technologies. These estimates were then combined with a Bayesian updating procedure.

The second approach applied regression models. One procedure trained an auto-regressive model on each cluster and forecasted the demand of the prediction target with the model of the most similar cluster. In another study, a direct relationship between the prediction target and a time-lagged similar technology was estimated with a regression model. Later periods of the time-lagged similar technology could then be used to predict demand. The last approach assigned technologies to a cluster of similar technologies and used the average time series of the cluster as the demand prediction.

(3) Which combination of techniques identified for RQ1 and RQ2 provides the most accurate predictions?

Answering research question 1 and 2 led to the construction of the conceptual overview of similarity based sales forecasting procedures presented in figure 1. All three approaches for finding similar technologies were evaluated in the case study. The clustering approach was implemented with k-means time series clustering and a decision tree classifier based on descriptive criteria. The second approach applied a nearest neighbour search based on descriptive criteria. Last, business experts were asked to identify similar technologies.

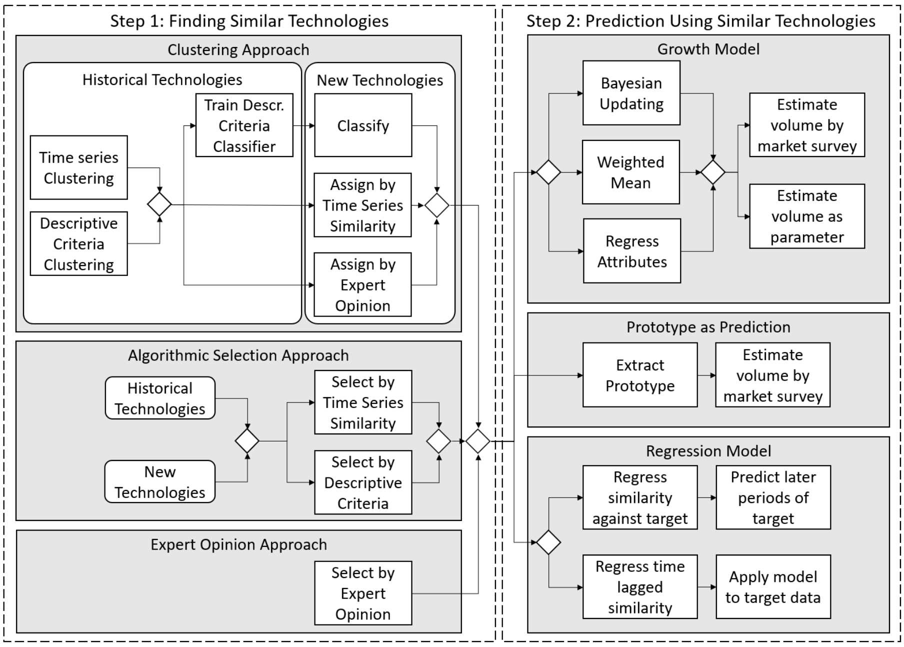


Figure 1: Conceptual overview of procedures in the Literature

For the second step, the regression and growth model approaches were evaluated. In the first approach, the demand of the prediction target was regressed against the time-lagged similar technology to predict later periods of demand. In the second approach, growth models were estimated with the time series data of the target and with target data extended with the time series of the similar technology. Subsequently, these estimates were combined in a Bayesian updating procedure.

The best accuracy was achieved if similar technologies by the combination of the expert approach and growth models. In addition, for all combinations with approaches for finding similar technologies, growth models outperformed regression models. Last, a clustering or selection approach provided similar performance combined with the growth models. The accuracy for each combination of techniques is shown in table 1.

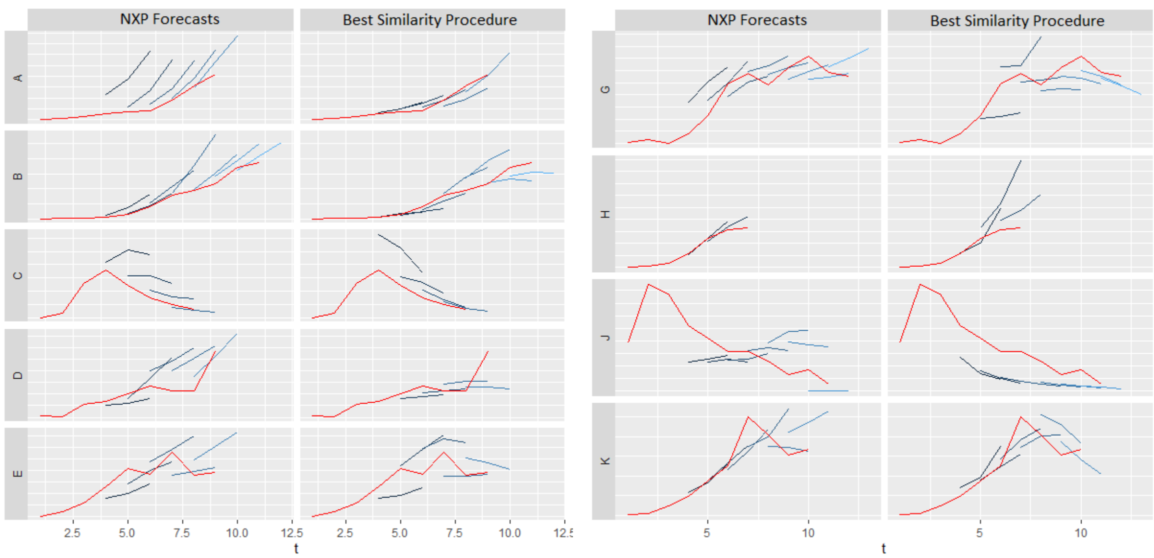
(4)To what extent can the current market demand estimates be improved with the prediction techniques?

Table 1: Accuracy by prediction model and approach for finding similar technologies (MAPE)

	Clustering	Selection	Expert
Regression Model	51%	145%	62%
Growth Model	48%	47%	38%

On average, the procedure combining growth models with the expert opinion approach outperformed the traditional forecasts of NXP. The mean absolute percentage error of the proposed procedure equaled 38% and of the traditional forecasts equaled 52%. If either clustering or a selection approach were used, a similar accuracy to traditional forecasts was achieved. In contrast, the regression model sometimes produced good results; however, it could also lead to forecasts of poor quality.

Figure 2 shows the forecasts of the best performing procedure and the forecasts of NXP. In this figure, each red line represents the actual production volume and each blue line represents a forecast for the next three years. In each subfigure, the diffusion technologies are divided into rows where the first column shows the NXP forecasts and the second column shows the growth model forecasts.



(a) Technologies A, B, C, D and E

(b) Technologies G, H, J and K

Figure 2: Forecasts of NXP and Growth Models with Expert Opinion Approach

This project contributes to the research area with a conceptual overview of the procedures for similarity based sales forecasting. In addition, it evaluated three different approaches for finding similar technologies and two approaches for prediction in a single case study. Last, the procedures were applied to long life cycle technologies, where previous research focused on short life cycle products.

Ultimately, similarity based sales forecasting procedures cannot replace the forecasts of NXP. Firstly, it does not utilize all information sources available to NXP. Secondly, the procedure might lead to poor results for some cases. Thirdly, similarity based forecasts do not include sufficient argumentation for the estimate to support major planning decisions. Still, the procedure can lead to large performance gains. These procedures could best be deployed in the review process of the estimates provided by the product lines.

Preface

This thesis presents the results of my graduation project at NXP and concludes my master degree in Operations Management and Logistics at the Eindhoven University of Technology. It was a great experience to apply the knowledge and skills I have gained during my study to the project at NXP. During the project, I have received great support from my university and company supervisors.

First, I would like to thank Remco Dijkman for guiding me in this project. You have introduced me to this opportunity at NXP and your feedback and suggestions have vastly improved my work. I would also like to thank Anna Wilbik, my second university supervisor. You have provided valuable insights and suggestions regarding my forecasting methods and useful feedback to improve my thesis.

Jark Snijder, you have been a great company supervisor. I have enjoyed the many discussions we had while exploring the direction of the project. You have also helped me connect to a large number of people who were interested in my work.

During my project, I gained another supervisor at NXP: Robbert-Jan Westerduin. I could only have analyzed about half of the data if it were not for Robbert-Jan. You have invested a lot of time in finding, checking and correcting the data. Thank you for your support.

Joris Marcelis
May 12, 2020

Contents

1 Introduction **10**
1.1 Long-term Planning . . . 10
1.2 Case Study . . . 10
1.3 Research Goal . . . 12
1.4 Research Questions . . . 12
1.5 Research Design . . . 13
1.6 Thesis Outline . . . 13

2 State of the Art in Similarity Based Sales Forecasting Techniques **14**
2.1 Review Methodology . . . 14
2.2 Forecasting Procedures . . . 15
2.2.1 Growth Models . . . 15
2.2.2 Regression Models . . . 16
2.2.3 Prototype Curves . . . 17
2.3 Coding of Approaches . . . 18
2.3.1 Framework for Finding Similar Technologies . . . 18
2.3.2 Framework for Predicting Using Similar Technologies . . . 19
2.3.3 Conceptual Overview of Procedures . . . 20

3 Selection of Similarity Based Forecasting Techniques **22**
3.1 Conceptual Design . . . 22
3.1.1 Procedures for Finding Similar Technologies . . . 22
3.1.2 Procedures for Prediction of Demand . . . 23
3.1.3 Overview of Conceptual Design . . . 23
3.2 Statistical and Machine Learning Techniques . . . 24
3.2.1 Growth Models for Curve Fitting and Prediction . . . 24
3.2.2 K-Means Time Series Clustering . . . 25
3.2.3 Decision Tree Classification . . . 25
3.2.4 Selection of Similar Technologies by Nearest Neighbours . . . 26
3.2.5 Selection of Similar Technologies by Business Experts . . . 26
3.2.6 Estimating the Regression Model . . . 26
3.2.7 Bayesian updating of Growth Models . . . 27

4 Case Study **29**
4.1 Data Extraction and Preparation . . . 29
4.1.1 Data Integration and Cleaning . . . 29
4.1.2 Selection of Diffusion Technologies . . . 30
4.1.3 Constructing Descriptive Criteria . . . 30
4.1.4 Data Characteristics . . . 30
4.2 Results . . . 32
4.2.1 Curve Fitting . . . 32

4.2.2 Clustering . . . 33
4.2.3 Forecasting Experiment Setup . . . 36
4.2.4 Regression Model Forecasts . . . 36
4.2.5 Growth Model Similarity Forecasts . . . 38
4.2.6 Growth Model Similarity and LTP Forecasts . . . 39
4.2.7 Prediction Intervals . . . 40

5 Conclusion **41**
5.1 Revisiting the Research Questions . . . 41
5.2 Relevance . . . 44
5.2.1 Scientific Relevance . . . 44
5.2.2 Company Relevance . . . 45
5.3 Limitations and Suggestions for Further Research . . . 46
5.3.1 Additional Predictor Variables . . . 46
5.3.2 Bias of Similar Technologies . . . 46
5.3.3 Uncertainty of Total Volume Estimate . . . 47
5.3.4 Overlap Training and Testing Data Split . . . 47
5.3.5 Dissaggregation of Forecasts . . . 48

References **49**

A Proof of equation 3.11 and 3.12 **53**

B Curves fitted to training technologies **55**

List of Figures

1	Conceptual overview of procedures in the Literature	3
2	Forecasts of NXP and Growth Models with Expert Opinion Approach	4
1.1	Demand forecasts for a new, mature and legacy wafer diffusion technology group	12
1.2	Research Design	13
2.1	Clustering and nearest neighbours procedure of Ilonen et al. (2006)	16
2.2	Clustering procedure of Basallo-Triana et al. (2017)	17
2.3	Clustering procedure of Thomassey (2010) and others	17
2.4	Clustering procedure Hu et al. (2009)	18
2.5	Conceptual overview of procedures in the Literature	21
3.1	Implemented procedures	23
4.1	Histogram of total volume and life cycle length	31
4.2	Number of technologies of the training set launched and ended per year	31
4.3	Clustered time series of diffusion technologies	32
4.4	Seven cluster validity index scores	33
4.5	Number of technologies in each cluster	34
4.6	Clustered time Series of diffusion technologies	34
4.7	Clustered time Series of diffusion technologies	35
4.8	Forecasts of NXP and Regression Models with Clustering Approach	37
4.9	Forecasts of Growth Models with Expert Opinion Approach	39
4.10	Forecasts of Growth Models with Expert Opinion Approach and Integration of LTP	40
5.1	Conceptual overview of procedures in the Literature	43

List of Tables

1 Accuracy by prediction model and approach for finding similar technologies (MAPE) 4

2.1 Coding form clustering and selection of similar products 19
2.2 Coding form predicting life cycle and total volume 20

4.1 Success rate of fitting models to diffusion technologies 32
4.2 MAPE of regression model 37
4.3 MAPE of growth model forecast updated with similar technology 38
4.4 MAPE of growth model forecast with similarity and LTP 39
4.5 Percentage of demand inside prediction interval 40

5.1 Accuracy by prediction model and approach for finding similar technologies (MAPE) 43

Chapter 1

Introduction

This thesis presents procedures for forecasting the demand of semiconductor technologies using the historical demand of similar technologies. The similarity based forecasting procedures first identify technologies similar to a new technology and subsequently use their historical demand to create forecast for the new technology. Three approaches for finding similar technologies and two approaches for predicting demand using similar technologies are implemented. The performance of the procedures is evaluated with a case study conducted at NXP Semiconductors.

This chapter provides an introduction to the thesis. Section 1.1 provides a motivation for forecasting semiconductor demand. In section 1.2, the case study environment is described. The research goal is presented in section 1.3 and the research questions formulated to reach the project goal are presented in section 1.4. Next, the research design is described in section 1.5. Last, section 1.6 outlines the structure of this thesis.

1.1 Long-term Planning

A manufacturing company aims to match its production capacity to the demand of its customers. An oversupply of production capacity results in a lower utilization of resources which increases the production cost per unit. Thus, an oversupply of production capacity affects the ability to make a profit (Kiran, 2019b). Alternatively, a shortage of capacity prevents a company to fulfill their customer demand resulting in lost revenues.

Matching the production capacity and customer demand is especially important in capital intensive industries. A company in a capital intensive industry has a high operating leverage, which means that a large proportion of a company's costs are fixed costs. In this case, the company earns a relatively large profit on each additional sale. Alternatively, if sales turn out lower than expected, the company will still have to pay its large fixed costs (Milgram, Spector, & Treger, 1999). Thus, the ability of a company with a high operating leverage to make a profit is heavily influenced by changes in the sales quantity.

Increasing production capacity requires a large financial commitment. This requires resources to be in service for a long time to produce a volume which is large enough to turn the investment into a profit. To sum up, long term investments in machinery and equipment are required in capital intensive industries. As a consequence, investments in production capacity have to be made long before the customer demand is known. In these cases, long term demand forecasts inform the decisions about planning of production capabilities (Kiran, 2019a). Inaccurate forecasts can lead to an oversupply or shortage of capacity which effect the final profits.

1.2 Case Study

NXP Semiconductors is a leading semiconductor manufacturer in the automotive, secure identification and digital networking industry. In 1953 the company was founded as Philips Semiconductors as a part of the electronics firm Philips. The company was sold to private investors in 2006 and subsequently changed its name to NXP. The company employs 30,000 employees worldwide and its corporate headquarters is situated in Eindhoven, the Netherlands.

NXP determines their long-term capacity requirements in a yearly long-term planning process (LTP). Important decisions such as investments in production capabilities and sourcing strategies depend on the output of the LTP. Especially decisions related to the fabrication of wafers require estimates of the long-term capacity requirements, because increasing the wafer production capacity is very expensive and takes one to three years to implement. In the LTP process, the wafer production quantities in the next three to five years are estimated.

The product lines of NXP provide the input for the long-term planning process. A product line is a hierarchical level which performs the business functions marketing, development, finance and operations for a specific product group and is the lowest level accountable for sales, profit and loss. The product lines estimate wafer demand based on market expectations, long term agreements with customers and design-wins, which are situations where a customer designs an NXP's product into their own product. In addition, strategic directions from the top management regarding sales targets and portfolio changes influence the volume estimates. A standard process of estimating wafer quantities does not exists. The product lines developed their own methods which they considered most appropriate for their market area. Most product lines define their long-term estimations at a detailed product level and specify five year quantities for each end product.

NXP uses a product and a process structure to categorize the produced products. At NXP there are multiple levels of product categories of which the highest hierarchical level is called the Main Article Group (MAG). The MAGs are created by grouping products with similar characteristics and which are sold to the same market. An example MAG is "In Vehicle Networks", the products of this group are sold in the automotive market and deliver networking functionality.

The highest level in the process structure is the wafer diffusion technology group. This level defines the type of diffusion process used to produce the wafers. The demand of a diffusion technology is characterized by a product life cycle curve. A diffusion technology generally first experiences a growth stage with increasing demand (ramp up). Next, the sales growth stabilises in the mature stage. Last, the diffusion technology reaches the end of its life cycle in the decline stage. The length of the life cycles of diffusion technologies can be between 3 to 30 years, depending on the corresponding market area. For example, wafer diffusion technologies applied in the agile mobile industry generally have shorter life cycles than those applied in the automotive industry characterized by strict quality requirements (Ahari, Viehl, Bringmann, & Rosenstiel, 2018).

Each product line estimates the demand of the products in one or more Main Article Groups. While a MAG is assigned to a specific product line, a diffusion technology is used for products in multiple MAGs. In other words, the same diffusion technology is often used in the product portfolio of multiple product lines. As a result, the aggregated demand of a diffusion technology is calculated by combining estimations of a number of product lines.

The estimates of the demand of diffusion technologies contain some inaccuracies. Figure 1.1 shows the difficulty experienced at NXP in forecasting demand. In the graphs each line represents a long-term forecasts of demand for a diffusion technology. The first graph shows forecasts of a new diffusion technology, the second graph of a group in the mature stage and the last graph of a legacy diffusion technology. In the example of the new technology, the forecasts are adjusted downwards each subsequent year. Consequently, NXP expected the rapid increase of new product demand to have occurred at an earlier point in time. The forecasts of the mature and legacy diffusion technologies show a consistent underestimation of demand. These examples demonstrate the challenge for NXP to predict the product life cycle of diffusion technologies.

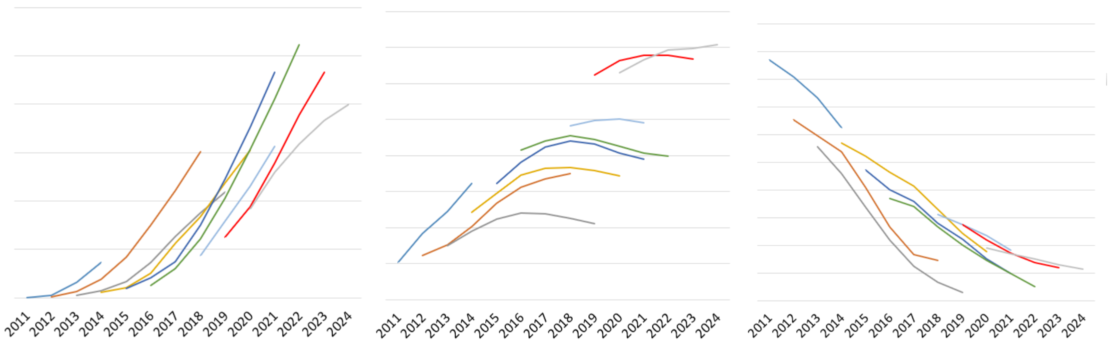


Figure 1.1: Demand forecasts for a new, mature and legacy wafer diffusion technology group

1.3 Research Goal

The inaccuracies of the estimations have large consequences for the effectiveness of planning decisions. Thus, NXP wants to investigate whether these estimates can be improved. Additionally, NXP would like to know the uncertainty of forecasts or different possible scenarios of demand. Based on this information, NXP could implement capacity capabilities considering different realizations of demand. This thesis hypothesises that the life cycle patterns of other diffusion technologies could be used to improve the prediction of demand of a diffusion technology. Consequently, the goal of this project is:

"Find life cycle patterns in the wafer demand of semiconductor wafer technologies and use these patterns to improve long term wafer demand predictions"

This project focuses on the prediction of wafer demand for semiconductor products. However, the project aims to produce results which can be generalized to other industries where similar demand patterns can be found.

1.4 Research Questions

A preliminary review of the literature revealed that clustering can be applied to find relevant technologies (Hu, Acimovic, Erize, Thomas, & Van Mieghem, 2019) as well as a method which directly identifies similar technologies (Aytac & Wu, 2013). In addition, different types of prediction models can be applied (Basallo-Triana et al., 2017). For both the step of finding similar technologies and the step of creating predictions, a research question is formulated to identify appropriate procedures. A third question evaluates the performance of each combination of techniques for finding similar technologies and creating predictions. The fourth question compares the performance of the proposed procedures to the traditional forecasts of NXP. Thus, the following research questions are answered in this report:

1. How can similar technologies be identified by either a clustering or selection approach?
2. How can the demand of a technology be predicted, using a similar technology or cluster of similar technologies?
3. Which combination of techniques identified for RQ1 and RQ2 provides the most accurate predictions?
4. To what extent can the current market demand estimates be improved with the prediction techniques?

1.5 Research Design

This section discusses the steps performed to answer the research questions. An overview of these steps is given in figure 1.2. First, the literature was reviewed for relevant techniques. Relevant articles were identified using a search query applied to four academic databases. The resulting articles were filtered using selection criteria. The filtered list was then expanded by forward and backwards reference searching. Next, the techniques were extracted from the articles and analyzed. The review provides answers to research question one and two. The next step was to select methods of the literature review to evaluate in the case study. The forecasting procedures were chosen based on the case study context and additional literature.

In the data preparation step, the data of NXP was prepared for modelling. First, the data was extracted from the sources at NXP and integrated into a single dataset. Next, this integrated dataset was cleaned by identifying and correcting inaccurate or incomplete data. After the data was extracted, integrated and cleaned, the relevant data points were selected for analysis. The last step was to analyze the descriptive statistics of the data.

The prepared dataset could then be used as input for the proposed forecasting procedures. A forecasting experiment was formulated to test all combinations of selected techniques of research question 1 and 2. The results of the implementation were then evaluated to answer research questions 3 and 4.

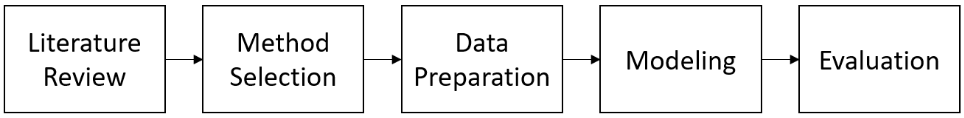


Figure 1.2: Research Design

1.6 Thesis Outline

The literature review is reported in chapter 2. Chapter 3 provides a description of the chosen methods and the argumentation for their selection. Next, chapter 4 examines the case study. In this chapter, section 4.1 describes the preparation of data and section 4.2 presents the results of the implemented procedures. Last, the conclusions of this report are provided in chapter 5.

Chapter 2

State of the Art in Similarity Based Sales Forecasting Techniques

This chapter provides an overview of the techniques found in literature relevant to the research goal. The first section outlines the methodology used for finding and analyzing relevant articles. The next section reports the procedures found in the literature. The third section proposes a structure which defines different aspects of each procedure.

2.1 Review Methodology

This section presents the methodology of the literature review. First, a search query was constructed based on the research goal. This query was then applied to four academic databases. From the results, relevant articles were selected using a number of criteria. The list of relevant articles was then expanded by forward and backward reference searching and exploring other works of the authors. Last, the forecasting procedures were extracted and analyzed. The remainder of this section discusses each component in more detail.

The search query was formulated by identifying the search therms relevant to the research goal. To facilitate this process, the research goal (section 1.3) was divided into three components: (1) the life cycle patterns, (2) finding the life cycle patterns and (3) predicting demand. Relevant procedures in the literature should include each component. The procedure should utilize life cycle patterns (first component), provide a method to identify those patterns (second component) with the aim to predict demand (third component). Thus, search terms were defined for each component. Alternative terms for a component were connected with an "OR" operator and the components were connected by an "AND" operator.

The terms "demand pattern" and "life cycle" were used for the first aspect. The terms "cluster" and "analog" were used for the second aspect. The term "cluster" was included because a preliminary review showed that demand patterns could be clustered (Hu et al., 2019). The second term "analog" was used in the literature in the context of "analogous forecasting" or "analogue products", which refers to the practice of forecasting demand of a product using the historical data of earlier products (Goodwin, Dyussekeneva, & Meeran, 2013). Last, the term "forecast" was used to represent predicting demand. Thus, the following search query was formulated:

("demand pattern" OR "life cycle") AND (cluster OR analog*) AND (forecast)

The query was applied to four databases: ACM Digital, IEEE Explore, Scopus and Web of Science. The resulting articles had to be written in English and full access had to be provided. Only journals, conference papers and book chapters are included, because the review processes of these publications guarantee a high standard of quality and validity. Applying these criteria led to a long list of articles.

Two criteria were applied to create a shortlist of articles. The articles had to search for patterns in other products or technologies than the prediction target. This was required because the forecasting problem analyzed in this report is characterized by a lack of historical data of the prediction target. Secondly, the patterns in data were life cycle patterns and not other patterns such as stock prices, electricity load or water demand. These type of patterns differ greatly from life cycle patterns, and thus methods predicting these type of patterns were deemed irrelevant.

The shortlist was extended by forward and backward reference searching and exploring other works of the authors. Forward reference searching is the method to analyze the articles published which have cited the original after it was published. Backwards reference searching is the method to analyze articles cited by the original article. Relevant articles identified by these methods were subjected to the same selection criteria.

The next step was to analyze the procedures reported in the literature. The prediction procedure of each articles was extracted from the the articles. Each procedure is summarized in section 2.2. The procedures were then analyzed for a common structure. A coding framework was established which defined the different modelling choices made in each procedure. Section 2.3 reports this framework.

2.2 Forecasting Procedures

This section describes the procedures found in the literature. The procedures are discussed in three subsections. Each subsection discusses procedures which apply a similar prediction model.

2.2.1 Growth Models

In the literature, growth models are frequently applied to forecast the life cycles of new technologies (Radojičić & Marković, 2009). Growth models are used for forecasting the diffusion of an innovation. The models describe the proportion of the relevant population that has adopted the innovation at a certain moment in time. Generally, the diffusion of an innovation starts with a slow phase, followed by rapid growth and ending with a declining adoption rate. Plotting the total degree of adoption against time results in the S-shaped or sigmoid curve (Meade & Islam, 1998). The derivative produces a bell-shaped curve representing the growth or sales per period.

Tuning the parameters of a growth model is limited by the availability of historical data of the product. A number of articles address this issue by using the historical demand of similar products launched earlier in time. These products have sufficient data available to estimate the parameters of growth models. The procedures in the literature differ in how they find and group similar products and how they use the data of similar products to predict demand.

Zhu and Thonemann (2004) implemented an approach forecasting semiconductor demand using the Bass Diffusion Model and data of similar products. The authors estimated the Bass parameters for each similar product. Next, business experts assigned to each similar product a probability that the new product would follow its demand pattern. In addition, the available data of the new product was used to estimate the model parameters. A Bayesian procedure combined the weighted parameters of similar products with the parameters of the new product into a single forecast for the new product. (Jung & Lim, 2016) extended this approach with a method based on Winter's model. Their method provided short-term forecasts of temporary increases or decreases in demand in the life cycle.

Aytac and Wu (2013, 2011) proposed a comparable approach. The authors selected seven growth models which outperformed most alternative growth models in a study of Meade and Islam (1998). They considered the forecast at time t as a random variable. First, a prior distribution of the forecast was estimated based on the historical data of the new product. The data of the new product was then extended with data of similar products. The growth models were estimated again using the extend data, which resulted in a sample distribution of the forecast. The authors applied a Bayesian procedure to update the prior distribution with the sample distribution to form a posterior distribution of the forecast. The similar products were identified using a correlation coefficient for time series data. If the coefficient of a historical product exceeded a specified threshold, it was regarded as a similar product and used to extend the data of the new product. Compared to Zhu and Thonemann (2004), the authors did not make an estimate of the parameter for the total life cycle volume. Some authors have argued that the estimates of this parameter are inaccurate and are better made by traditional market surveys (Heeler & Hustad, 1980; Tigert & Farivar, 1981).

Some articles assumed that the target product was completely new, meaning that the model parameters could only be estimated from similar products. Goodwin et al. (2013) applied a nearest neighbor search on descriptive criteria to identify similar products. The nearest neighbor search weighted the similar products and a forecast was made with a growth model using the weighted parameters. Similarly, Ilonen, Kamarainen, Puumalainen, Sundqvist, and Kälviäinen (2006) applied a nearest neighbor analysis to find similar countries where a product was already introduced at an earlier point in time. The assumption was that the product's growth process would behave similar in countries with the same characteristics. This procedure clustered the countries using Self Organizing Maps (SOM) before applying the nearest neighbor algorithm to select the most similar countries. A training set of technologies was used to determine the optimal k number of nearest neighbors to select. Figure 2.1 shows the clustering and nearest neighbours procedure of Ilonen et al. (2006).

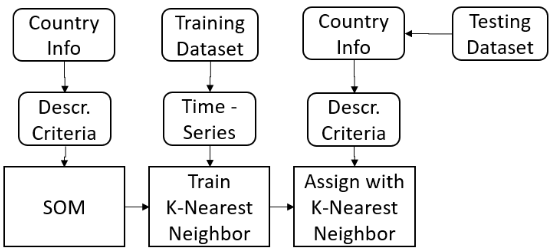


Figure 2.1: Clustering and nearest neighbours procedure of Ilonen et al. (2006)

In addition to a nearest neighbor search, Goodwin et al. (2013) and Ilonen et al. (2006) implemented an alternative method to estimating growth model parameters. Instead of using the growth model parameter values of similar products, the parameter values were regressed against the product attributes. Subsequently, the trained regression model was used on the attributes of a new product to estimate its growth model parameters. A similar approach was proposed by Lee, Kim, Park, and Kang (2014), who compared the performance of linear regression, k-nearest neighbor regression, classification and regression tree, artificial neural networks, support vector regression and Gaussian process regression for estimating growth model parameters based on product attributes.

2.2.2 Regression Models

Procedures applying alternative forecasting models can be found in the works of Wu, Aytac, Berger, and Armbruster (2006) and Basallo-Triana et al. (2017). These authors estimated regression models instead of growth models. The first article found similar products by using a correlation coefficient threshold. This method for finding similar products was also used by Aytac and Wu (2013, 2011) and is described in the previous subsection. Wu et al. (2006) regressed a section of the time series of the target against a time-lagged series of the similar product. The remaining time-series of the target was predicted with the regression model applied to the later periods of the similar product.

Basallo-Triana et al. (2017) clustered time-series with Fuzzy Clustering and calculated distances with the Gustafson-Kessel Algorithm. First, an auto-regressive model was trained on each cluster. Second, the demand of the new product was predicted with each auto-regressive model. Last, the cluster of which the auto-regressive model created predictions with the smallest distance to the cluster's prototype was selected as the cluster with the most similar products. A cluster's prototype, also referred to as the centroid, is defined as the average time series in a cluster. A schematic overview of the procedure can be found in figure 2.2.

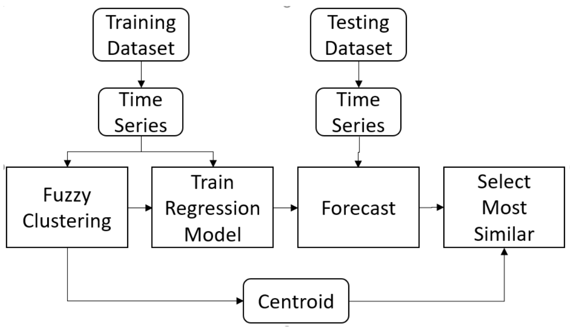


Figure 2.2: Clustering procedure of Basallo-Triana et al. (2017)

2.2.3 Prototype Curves

Another class of procedures forecasted demand by assigning the prototypes of clusters as predictions to the target product. Thomassey and Fiordaliso (2006) clustered the time series with the K-Means clustering algorithm. Next, they trained a decision tree on the descriptive criteria of the data in the clusters. The decision tree was used to assign a new product to a cluster. The prototype of the assigned cluster was used as the sales profile of the new product. The sales profile could then be transformed into a forecast by using traditional market survey data on volume and life cycle length. Thomassey (2010) extended the procedure with a fuzzy inference system (FIS). This model was used on the similar products to remove the influence of other variables. After a prototype was assigned to a new product, the influence of the variables was added to the new product using the FIS model trained on similar products. Another paper by Thomassey and Happiette (2007) replaced the decision tree of the procedure of Thomassey and Fiordaliso (2006) with a probabilistic neural network (PNN). A schematic overview of the procedures of Thomassey and Fiordaliso (2006); Thomassey and Happiette (2007); Thomassey (2010) can be found in figure 2.3

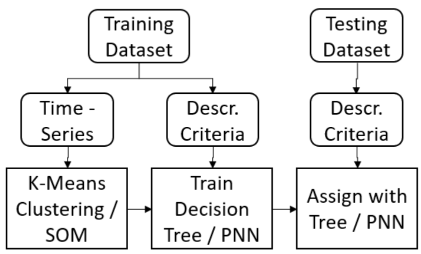


Figure 2.3: Clustering procedure of Thomassey (2010) and others

A similar approach to forecasting with prototypes was proposed by Hu et al. (2019). The authors included a prior step of fitting product life cycle curves to the historical products. This step ensured that demand patterns unrelated to the product life cycle were removed from the historical products. Subsequently, the authors applied a hierarchical clustering procedure on these fitted curves. A combined correlation coefficient and euclidean distance measure was used for the clustering procedure. Business experts assigned the target product to a cluster in their case study, because a small number of data points did not allow the training of a classification model which could assign new products to clusters. Hu et al. (2019) scaled their forecasts to match the company's traditional estimation of the volume in the initial weeks. The clustering procedure of Hu is shown in figure 2.4.

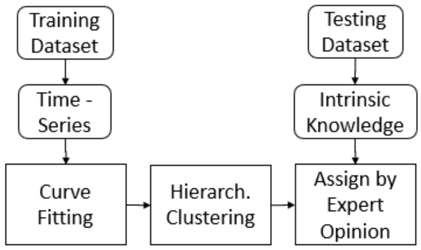


Figure 2.4: Clustering procedure Hu et al. (2009)

2.3 Coding of Approaches

This review concludes that the approaches reported in the previous section can be divided into two steps. The first step is to identify which technologies are similar to a new technology. The second step is to use the identified similar technologies to predict the demand of the new technology. Each step is further divided into components. The first two sections reports the components identified for each step. The last sections provides a conceptual overview of the procedures.

2.3.1 Framework for Finding Similar Technologies

Some articles first cluster technologies and subsequently select a cluster with similar technologies. Other articles do not apply clustering, instead they directly identify similar technologies either using and algorithm or the opinion of experts. Thus, all approaches apply a selection step and some approaches apply clustering as a prepossessing step. Moreover, both the clustering and algorithm based selection approach utilize a similarity measure. The articles differ in the data type used for this similarity measure. The data types used are descriptive criteria, time series and curves fitted to time series. Table 2.1 reports for each article the clustering method, the data type used by the clustering method, the selection method and the data type used by the selection method. If no method is applied, the cell in the table reports n/a. Additionally, the data type column reports n/a, if either the expert opinion based selection method is used or all data is selected

Table 2.1: Coding form clustering and selection of similar products

Author	Clustering	Data Type	Selection	Data Type
Aytac & Wu (2011)	n/a	n/a	Correlation	Time Series
Aytac & Wu (2013)	n/a	n/a	Correlation	Time Series
Basallo-Triana et al. (2017)	Fuzzy	Time Series	Best Model	Time Series
Goodwin et al. (2013) [1]	n/a	n/a	NNS	Descriptive
Goodwin et al. (2013) [2]	n/a	n/a	All Data	n/a
Hu et al. (2019)	Hierarchical	Curves	Expert Opinion	n/a
Ilonen et al. (2006) [1]	SOM	Descriptive	NNS	Descriptive
Ilonen et al. (2006) [2]	n/a	n/a	All Data	n/a
Jung & Lim (2016)	n/a	n/a	Expert Opinion	n/a
Lee et al. (2014)	n/a	n/a	All Data	n/a
Thomassey (2010)	K-Means	Time Series	Decision Tree	Descriptive
Thomassey et al. (2006)	K-Means	Time Series	Decision Tree	Descriptive
Thomassey et al. (2007)	SOM	Time Series	PNN	Descriptive
Wu et al. (2006)	n/a	n/a	Correlation	Time Series
Zhu & Thonemann (2004)	n/a	n/a	Expert Opinion	n/a

2.3.2 Framework for Predicting Using Similar Technologies

Three components are defined for the prediction step. The first component is the model used to create the predictions. The second component defines how the parameters of the model are estimated using the similar technologies. The last component defines the method for estimating the total volume of the life cycle. Table 2.2 reports the model, parameter estimation method and volume estimation method used by each article.

Three types of models were used to create a forecast. One class of models applied was the growth model. Two articles implemented seven growth models, the other articles applied the Bass Diffusion Model. Three methods of estimating the parameters of growth models were identified. Some articles regressed the parameters against the product attributes. This regression model was then used to estimate the growth model parameters of a new product based on the values of its attributes. Another option was to take the parameters of one similar product, the weighted parameters of multiple similar products or the parameters of a cluster. Last, some procedures combined the parameters of similar products and of the target product with a Bayesian updating procedure.

The next class of models used the prototype of a cluster as the demand prediction. Assigning a prototype as the prediction does not use require the estimation of parameters, thus, the estimation cell reports n/a for these articles. The last class trained regression models on similar technologies or clusters of similar technologies and subsequently choose the best model to forecast the target product.

A number of the identified procedures only estimated a sales pattern, without estimating the total volume of the sales. Some authors argued that estimating the total volume with a growth model leads to inaccurate results. However, one paper included the total volume parameter in its procedure. All the procedures which choose a prototypes as prediction used the traditional estimations of a company to scale the volume of the forecasts. Notably, one paper scaled the prototype using traditional estimates of sales in a number of periods; thus, the procedure used only an estimate of a portion of the total sales volume. Some authors did not apply methods to find the total volume. These methods only evaluated the ability of methods to predict the shape of the product life cycle. N/a is reported in the volume column, when only the normalized life cycle is evaluated in the article.

Table 2.2: Coding form predicting life cycle and total volume

Author	Model	Estimation	Volume
Aytac & Wu (2011)	7 Diffusion	Bayesian Updating	Market survey
Aytac & Wu (2013)	7 Diffusion	Bayesian Updating	Market survey
Basallo-Triana et al. (2017)	Regression	n/a	n/a
Goodwin et al. (2013) [1]	Bass	Weighted Mean	n/a
Goodwin et al. (2013) [2]	Bass	Regression attributes	n/a
Hu et al. (2019)	Prototype	n/a	Market survey
Ilonen et al. (2006) [1]	Bass	Weighted Mean	Country Population
Ilonen et al. (2006) [2]	Bass	Regression attributes	n/a
Jung & Lim (2016)	Bass	Bayesian Updating	n/a
Lee et al. (2014)	Bass	Regression attributes	Market survey
Thomassey (2010)	Prototype	n/a	Market survey
Thomassey et al. (2006)	Prototype	n/a	Market survey
Thomassey et al. (2007)	Prototype	n/a	Market survey
Wu et al. (2006)	Regression	n/a	n/a
Zhu & Thonemann (2004)	Bass	Bayesian Updating	Estimated in Model

2.3.3 Conceptual Overview of Procedures

Figure 2.5 provides a conceptual overview of the components defined in the previous section. The outer boxes with dashed lines refer to the two modeling steps. The solid lined boxes refer to the alternative approaches for each step. Last, the flow chart elements represent the methods used in each approach.

Three approaches for finding similar technologies were identified: a clustering approach, a algorithmic selection approach and a expert opinion approach. The clustering approach first clusters the historical technologies using either descriptive criteria, raw time-series or fitted curves. Next, a classification model is trained on the clusters. Again, a choice is made for a data type to use for defining the similarity between the clusters and new technologies. Last, a new technology, to be predicted, is assigned to a cluster with the trained classification model.

The algorithmic selection approach applies an algorithm to define the degree of similarity between technologies and subsequently selects the most similar technologies. The procedures differ in the selection algorithm and the data type utilized by the algorithm. Some compare the time-series of technologies and others compare descriptive criteria. Alternatively, the expert opinion approach is used. In this approach, similar products were selected by business experts.

Three types of approaches to creating predictions using similar technologies were applied in the literature. Some articles estimated growth models to create predictions. Three methods of parameter estimation were applied. The first method was to use a weighted mean of the parameters of similar technologies. The second method used a Bayesian updating procedure to combine estimations created with data of the target technology and with data of similar technologies. The third method estimated the relationship between attributes of similar technologies and growth model parameters. Consequently, the attributes of the prediction target could be used to estimate the growth model parameter values.

The second prediction approach is based on extracting a prototype time series from similar technologies or a cluster of technologies. This prototype was then either scaled in volume and length to fit the prediction target or directly used as a prediction. The third prediction approach applied regression models. Two types of regression models were identified: auto-regressive models and models regressing a new technology against a time-lagged similar technology.

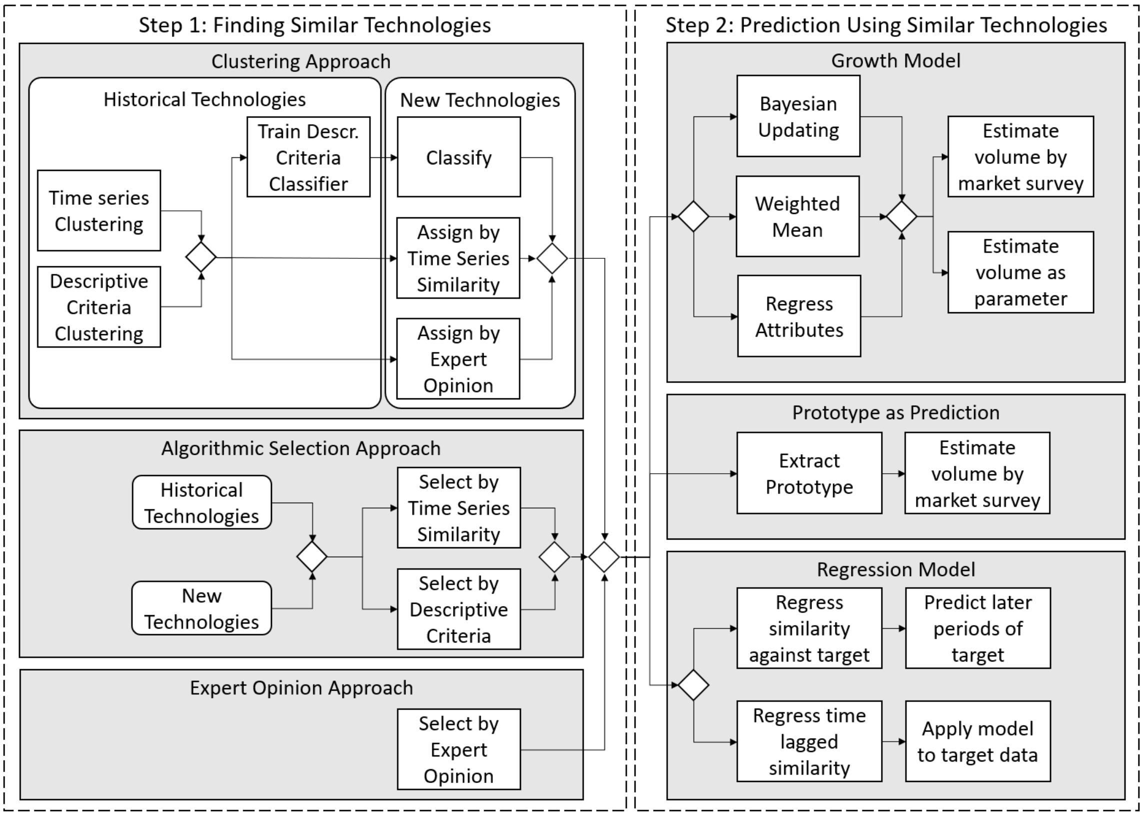


Figure 2.5: Conceptual overview of procedures in the Literature

Chapter 3

Selection of Similarity Based Forecasting Techniques

This chapter discusses the forecasting procedures formulated for reaching the project's goal. The first section outlines the conceptual design of the forecasting procedure. The conceptual design defines which approaches and methods identified in section 2.3 were implemented. The remaining sections describe the applied machine learning and statistical techniques.

3.1 Conceptual Design

The first subsection discusses the procedures for the first modeling step of finding similar technologies. The next subsection considers the procedures for predicting demand using the similar technologies. The last subsection provides an overview of the conceptual design.

3.1.1 Procedures for Finding Similar Technologies

This project implemented all three approaches for finding similar technologies to evaluate which approach would be most appropriate in the case study at NXP. The clustering approach could provide an advantage because it reduces noise and complexity which assists the next step of classification (Witten, Frank, Hall, & Pal, 2017). Thus, linking a new technology to a group of similar technologies might be easier than linking individual technologies. This hypothesis could be evaluated by comparing the performance with the algorithmic selection approach. In addition, the expert opinion approach could then be used to evaluate whether data mining methods would provide better performance than human judgement. In this approach, five business experts were asked to provide as many examples of wafer diffusion technology groups which are both similar to each other and launched at different times.

The clustering and algorithmic selection approach required the selection of data types to use for the similarity measure (see section 2.3.1). This project implemented time series clustering and used descriptive criteria, such as market focus and technology type, for classifying a new technology to a clusters or algorithmic selection of similar technologies. Time series clustering was chosen instead of clustering based on descriptive criteria, because the aim was to identify clusters of technologies which displayed the same life cycle patterns. However, new technologies could not be compared to earlier technologies or clusters of earlier technologies based on time series data, because new technologies do not have sufficient data points available for comparison. Thus, descriptive criteria were used to link new technologies with clusters and earlier technologies.

The goal of the clustering procedure is to group clusters with similar life cycle patterns. Thus, demand fluctuations caused by other effects should be removed. For example, economic downturns or inventory corrections could create temporary deviations from the life cycle pattern of a technology (Cheong, 2016). To mitigate this effect, curve fitting was applied as a preprocessing step to the clustering procedure. The curves were fitted to the technologies by estimation of growth models.

3.1.2 Procedures for Prediction of Demand

Two approaches for predicting demand were implemented: the regression model approach and the growth model approach. The prototype curves approach was not included. This approach does not utilize the historical data of the prediction target and thus does not allow for a rolling-window analysis which simulates NXP's long term planning forecasting process.

The Bayesian updating procedure was chosen for estimating the parameters of the growth models. This procedure enabled the combination of the three demand information sources: historical data of technology, historical data of the similar technology and the LTP forecasts of NXP. The alternative method of weighting the similar technologies and LTP forecasts was rejected because it would have been a labour intensive and subjective process. The last method, the estimation of parameters with regression, was rejected because it does not utilize the historical time series of the prediction target.

The regression method which regressed the new technology against the time lagged similar technology was selected. The advantage of this method is its simplicity in transforming the similar technology into a forecast via regression. As mentioned in the previous subsection, the life cycles of technologies could be influenced by other effects. It is undesirable to transfer these fluctuations from the similar technology to the new technology. Moreover, it might hamper the estimation of the regression parameter. Thus, curve fitting was applied as a preprocessing step to estimating the regression model. This step was omitted when a prototype of a cluster was used as the input, because other influences were already removed in the clustering procedure.

3.1.3 Overview of Conceptual Design

A schematic overview of the implemented procedures can be found in figure 3.1. The left hand side shows the clustering, algorithmic selection and expert opinion approach. The right hand side shows the regression approach and the growth model approach. For each modelling step in the figure, the subsection which discusses the applied techniques is indicated .

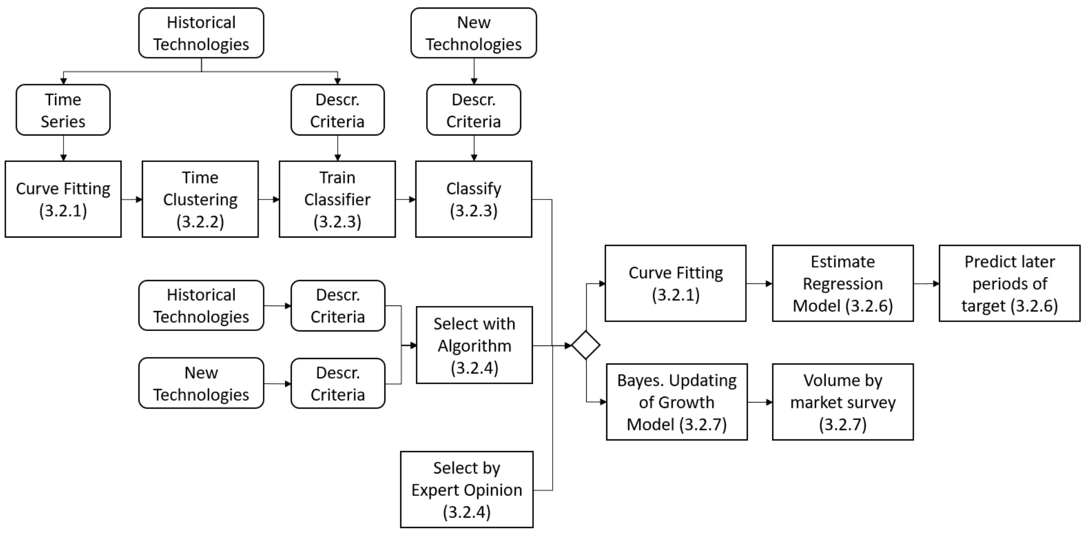


Figure 3.1: Implemented procedures

3.2 Statistical and Machine Learning Techniques

In subsection 3.2.1, four growth models are described which were used for fitting curves as a preprocessing step and for creating predictions. Next, subsection 3.2.2 argues the selection of the k-means time series clustering algorithm. The application of a decision tree classification model is discussed in subsection 3.2.3. Subsection 3.2.4 outlines the application of a nearest neighbour search for finding similar technologies. The alternative approach which had business experts identify similar technologies is discussed in section 3.2.5. The last two subsections explain the prediction models. The application of regression models is discussed in subsection 3.2.6 and the procedure for Bayesian updating of growth models in subsection 3.2.7.

3.2.1 Growth Models for Curve Fitting and Prediction

Four growth models were applied for fitting curves as a preprocessing step and for creating predictions. The decision was made to implement multiple models to be able to capture different types of growth processes of the technologies. The fitting of a growth model might fail if the life cycle pattern of the time series differs from the general shape of the model (Meade & Islam, 1998). Thus, the average of the growth models was calculated to form the prototype life cycle of a diffusion technology. This method reduced the bias induced by each model and solves the problem of non-fitting models.

The Bass Diffusion Model is a highly influential model used for forecasting the diffusion or adoption of an innovation. Bass (1969) made a distinction between innovators and imitators. He proposed that innovators decide to adopt an innovation independently of the decisions of others. The decision of imitators to adopt is influenced by the decision of others to adopt. Bass (1969) defined the parameters p, q and m. Parameter p reflects the adoptions of innovators, q the adoptions of imitators and m equals the total number of adoptions. The Bass Diffusion Model can be used in the form of equation 3.1 to calculate the growth of adoptions, denoted $f(t)$, at time t.

$$f(t) = m\frac{(p+q)^2}{p}\frac{e^{-(p+q)t}}{(1+\frac{q}{p}e^{-(p+q)t})^2} \tag{3.1}$$

Three additional growth models were implemented: Simple Logistic, Gompertz and Weibull. These models were identified to be high performing models in a study by Meade and Islam (1998). In addition, each model can produce differently shaped curves. The difference in shape can be defined by the timing of the point of inflection on the cumulative curve (Meade & Islam, 1998). In the case of sales forecasting, the inflection point represents the period of peak sales. The simple logistic model has the inflection point at 50% of total volume and thus is a symmetric curve. The Gomptertz model is asymmetric and has an inflection point before 50%. Last, the inflection point of the Weibull model has a range which includes 50%. Thus, the Weibull model is the most flexible regarding the timing of the sales peak. Equations 3.2, 3.3 and 3.4 show the formulas of the Simple Logistic, Gompertz and Weibul models, respectively. Again, the growth of adoptions at time t is denoted $f(t)$ and parameter m equals the total volume of the curve. In addition, parameters b and c determine the shape of the curve.

$$f(t) = m\frac{b.c.e^{-b.t}}{(1+c.e^{-bt})^2} \quad b>0, c>0 \tag{3.2}$$

$$f(t) = m.b.c.e^{-c.e^{-bt}-b.t} \quad b>0, c>0 \tag{3.3}$$

$$f(t) = m\frac{b}{c}\Big(\frac{x}{c}\Big)^{b-1}e^{-(x/c)^b} \tag{3.4}$$

Model parameters were estimated by non-linear least squares estimation (NLS). A study by Srinivasan and Mason (1986) estimated Bass Diffusion Models by ordinary least squares, non-linear least squares and maximum likelihood estimation. The article concluded that NLS provided the best performance and proposed that the NLS estimation was also appropriate for other growth models.

3.2.2 K-Means Time Series Clustering

As a preprocessing step to clustering, curves were fitted to the time series. These fitted curves were normalized by dividing the series by the total volume of the curve. Consequently, the total area under each curve equals one. Normalization was required to enable the clustering of life cycles based on their shape irrespective of their volumes.

Selecting the appropriate time series clustering method depends on the objectives of the modeling method. Relevant aspects of the life cycles are the shape, scale and location. The scale refers to how stretched out or concentrated the life cycle is. The location determines the timing of the peak of the cycle. Last, the shape refers to other moments such as skewness or kurtosis.

The K-means clustering algorithm utilizes the euclidean distance to calculate the distance between time series in each time step (Aghabozorgi, Seyed Shirkhorshidi, & Ying Wah, 2015). Thus, this method is the appropriate for identifying differences in the scale and location of a life cycle. An alternative method is applying Dynamic Time Warping (DTW) for clustering. The DTW method allows for warping the time dimension to identify similar shapes which do not occur the same speed or at exactly the same time step (Keogh, 2002). However, the time series already share a general life cycle shape and the difference in timing and speed of the life cycle are precisely the aspects which define (dis)similarity between diffusion technologies. Dynamic Time Warping is likely to place time series with different peaks in the same cluster. Thus, the K-means method is considered more appropriate.

A partitional clustering algorithm such as the k-means method requires a specification of the number of clusters to be generated. Consequently, the optimal number of clusters had to be determined. No labels of the desired output were available, since unsupervised learning was applied with the goal of identifying hidden patterns in the data. Thus, it was necessary to rely on internal validity indices to determine the optimal number of clusters. Internal validity indices score the clustering performance based on the cluster purity, while external indices evaluate the performance with information about the desired clustering output (Theodoridis & Koutroumbas, 2009).

The prototype curve of a cluster of time series was extracted to represent the average life cycle pattern of a group of diffusion technologies. Multiple methods exist for extracting prototypes (Aghabozorgi et al., 2015). This project applied the arithmetic mean in combination with the euclidean distance to extract a prototype curve. The same arguments for the selection of k-means clustering instead of more advanced methods apply to the selection of the arithmetic mean. The advanced prototype extraction methods are appropriate for preserving more complicated patterns with difference in speed. Similar life cycle patterns do not contain these characteristics; thus, the arithmetic provides a simple and effective method for averaging the time series.

3.2.3 Decision Tree Classification

A decision tree was applied to determine which cluster of diffusion technologies is most similar to the diffusion technology to be predicted. The assignment of a technology to a cluster represents a supervised classification problem. The clusters of the diffusion technologies included in the clustering step are known. Thus, it is possible to model the relationship between descriptive criteria and clusters. This model can subsequently be used to determine the relevant cluster for new diffusion technologies which lack the historical data required to be included in the time series clustering procedure.

Decision trees are a frequently used as classification models (Theodoridis, 2020). A decision tree consists of a number of questions organized in a treelike structure. At each node the data is split according to a value of a variable. Each node has two leafs which lead to the next node or endpoint (Nisbet, Miner, & Yale, 2018).

3.2.4 Selection of Similar Technologies by Nearest Neighbours

The Nearest Neighbours Algorithm was implemented to identify similar technologies based on descriptive criteria. This classification algorithm chooses the label of a data point based on the most common label of the k nearest neighbours (Yahyaoui's, Yahyaoui, & Yumuşak, 2018). In this implementation, k was set to 1 and the nearest neighbour was selected as a similar technology. The method requires a distance measure to identify the technology nearest to the prediction target. The euclidean distance measure is most common and thus applied in this procedure.

3.2.5 Selection of Similar Technologies by Business Experts

Five business experts were consulted to identify preceding technologies which were similar to the technologies to be predicted. The business experts performed different functions at NXP. Two experts had strategic roles, two experts worked in operations and one expert worked in technical project management. Together they determined which technologies were similar. Thus, knowledge of multiple disciplines at NXP was used to identify similar technologies. The results of the expert opinion approach for finding similar technologies were used to evaluate the benefits of the clustering and algorithmic selection approach.

3.2.6 Estimating the Regression Model

In the regression model approach, the historical demand time series of a technology is regressed against a time-lagged similar technology. As explained in section 3.1, first, a prototype curve is extracted from the similar technology using the growth models presented in subsection 3.2.1. The relationship between the prototype curve and the available demand data of the prediction target is formalized in equation 3.5. The demand of the prediction target at time t is denoted n_t and the demand of the prototype curve at time t is denoted p_t. Furthermore, ε_t represents the residuals at time t. The intercept of the regression model was set to 0, because the inclusion of an intercept would change the shape of the life cycle curve of the similar technology. Adding or subtracting a constant to a curve influences the ratio of two values at different points on the curve. For example, a positive intercept would cause the regression model to produce a flatter curve and a negative intercept would produce a more concentrated curve.

If we have historical sales data till time L and apply the least squares estimation, the value for β can determined with equation 3.6. Consequently, using the prototype and an estimated value of β forecasts can be made with equation 3.7 for periods t > L. In this equation, d_t denotes the demand forecast at time t and ϵ_t is the forecasting error at time t.

$$n_t = \beta.p_t + \varepsilon_t \quad for \quad 1 \leq t \leq L \tag{3.5}$$

$$Min \sum_{n=1}^{L} \varepsilon_t^2 = Min \sum_{n=1}^{L} (n_t - \beta.p_t)^2 \tag{3.6}$$

$$d_t = \beta.p_t + \epsilon_t \quad for \quad t > L \tag{3.7}$$

3.2.7 Bayesian updating of Growth Models

The alternative forecasting method applied growth models and Bayesian updating. The four growth models described in section 3.2.1 were used for forecasting. Each model required the estimation of two parameters that determine aspects of the curve such as the shape, scale and location. An estimate of the total volume was supplied as input to the procedure. The procedure considered the demand estimate at time t of the growth model as a random variable. A prior distribution of the random variable was estimated by fitting a growth model to the historical data of the technology. Next, the historical data was extended with the data of a similar technology. A growth model was fitted to the extended data. This resulted in a sample life cycle projection. Last, a Bayesian updating procedure was applied to form a posterior distribution of the demand estimate. This project applied a Bayesian updating procedure similar to the procedure described in Aytac and Wu (2013) and Aytac and Wu (2011).

Prior distribution

The estimate of growth model m with data θ_T up time T for time $T+S$ is represented by $\widehat{X}_m(T+S|\theta_T)$. The estimation error is assumed to be normally distributed. Thus, the random variable $\tilde{X}_m(T+S)$ corresponding to the estimation of actual demand by model m at time $T+S$ equals a normal distribution with the following mean and variance:

$$\tilde{X}_m(T+S) \sim \ N(\widehat{X}_m(T+S|\theta_T), \sigma_m^2)$$

The mean $\widehat{X}_m(T+S|\theta_T)$ can be calculated with the estimated model. The variance is more difficult to derive. If we assume that the estimate of the growth model and the estimation error are independent, the variance σ_m^2 equals the sum of the variance of the model's estimate and the variance of the estimation error.

$$VAR(\tilde{X}_m(T+S)) = VAR(\widehat{X}_m(T+S|\theta_T)) + VAR(\epsilon(T+S|\theta_T)) \tag{3.8}$$

Estimation of growth model variance

Meade and Islam (1995) proposed three methods for approximating the uncertainty of the model's estimate. Of these approaches, the explicit density approach performed best if a low number of data points was available. This method defines the error in estimating a growth curve function as a function of the error of its parameters (ε_b and ε_c):

$$\epsilon_t = f(b, c, t) - f(b + \varepsilon_b, c + \varepsilon_c, t) \tag{3.9}$$

Meade and Islam (1995) noted that growth models are too complex for analytical derivation of the density function of ϵ_t. Thus, they proposed a numerical method which produces a probability mass function for ϵ_t. The first step is to calculate the probabilities of different combinations of parameter values. Next, the error of each combination is calculated using equation 3.9 and combined with the probabilities to form the probability mass function of ϵ_t. This probability mass function is combined with the distribution of the error term ε to form the distribution of $\tilde{X}_m(T+S)$.

To aid the estimation of the probabilities of the combinations of parameter values, it was assumed that the error vector of ε_b and ε_c was a multivariate normal random variable. The joint probability density function of a bivariate normal distribution is given by 3.10. The values of σ_x, σ_y and ρ were taken from the covariance matrix of ε_b and ε_c, which is estimated as a byproduct in non linear least squares estimation.

$$f_{X,Y}(x, y) = ce^{-q(x,y)} \tag{3.10}$$

with the normalizing constant:

$$c = \frac{1}{2\pi\sqrt{1-\rho^2}\sigma_X\sigma_Y}$$

and the exponent term as a quadratic function of c and y:

$$q(x,y) = \frac{\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} - \rho\frac{xy}{\sigma_x\sigma_y}}{2(1-\rho^2)}$$

Sampling distribution

The actual demand of a similar technology is used to estimate the demand of the target technology. A growth model was fitted to the historical data of the target extended with the data of the similar technology. This extended data is represented by θ_{T+L}. If a similar technology is an unbiased estimation of the demand of the prediction target, the prior distribution can be updated with the sampling distribution. The estimate the variance of the model estimate and the error term of the sampling distribution was found with the method described in the previous subsection.

Posterior distribution

The prior distribution was updated with the sample distribution to give a posterior distribution of the demand estimation. The posterior mean and variance can be calculated with equation 3.11 and 3.12, respectively. The posterior distribution is a weighted function of the variance σ_m^2 (equation 3.8) of the estimate of model m fitted to the historical data of the target technology and variance τ_m^2 of the estimate of the model fitted with the data extended by the similar technology. The proof, formulated by Aytac and Wu (2013), can be found in appendix A.

$$\mu'_m = \frac{1/\sigma_m^2}{1/\sigma_m^2 + 1/\tau_m^2}\widehat{X}_m(T+S|\theta_T) + \frac{1/\tau_m^2}{1/\sigma_m^2 + 1/\tau_m^2}\widehat{X}_m(T+S|\theta_{T+L}) \tag{3.11}$$

$$\sigma_m'^2 = \frac{\sigma_m^2\tau_m^2}{\sigma_m^2 + \tau_m^2} \tag{3.12}$$

A combined forecast is formed by taking the average of the posterior mean and variance of each growth model. In formal representation: $\mu' = \sum_{m=1}^4 \mu'_m$ and $\sigma'^2 = \sum_{m=1}^4 \sigma_m'^2$. The one, two and three step ahead forecasts are created; thus, the procedure is repeated for $S\epsilon\{1,2,3\}$.

Extension by Updating with NXP forecasts

In addition to Bayesian updating with similar technologies, the procedure is extended with Bayesian updating with NXP's forecast. The similarity based sales forecasting procedure only uses the actual demand of the prediction target and the similar technology as demand information sources. NXP relies on many additional sources of information such as market outlook and customer estimates. Thus, itt might be beneficial to include the demand estimations of NXP in the forecasting procedure.

The NXP forecasts were included in the updating scheme using the same approach applied to similar technologies. First, the posterior distribution found by using a similar technology as sampling data was taken as the prior distribution. Next, a growth model is estimated to the historical data extended with NXP's forecasts. Last, a posterior distribution of the estimate is calculated using the equations described in this section.

Chapter 4

Case Study

This chapter presents the case study conducted at NXP Semiconductors. The first section discusses the extraction and preparation of data for the modelling step. The second section provides the results of the implementation of the forecasting procedures described in chapter 3.

4.1 Data Extraction and Preparation

A considerable number of diffusion technologies at NXP has a long life cycle. Some technologies remain in use for over 20 years. To include the longer life cycles in the analysis, the records of wafer demand starting from 1999 till 2019 were gathered. In addition, the records from 2008 till 2019 of the LTP forecasts of NXP were gathered. The integration and cleaning of the demand and LTP data files is reported in section 4.1.1. Next, a selection was made of diffusion groups to include in the analysis. This step is discussed in section 4.1.2. In section 4.1.3, the construction of descriptive criteria of diffusion technologies is presented. Last, the data characteristics of the selected diffusion technologies are presented in section 4.1.4

4.1.1 Data Integration and Cleaning

NXP stored the data on actual demand and LTP forecasts in a large number of data files. These data files do not apply the same tabular formats and were inconsist in naming objects. Four causes for the lack of agreement between data files were identified. First, this project analyzed historical production quantities and forecasts of a time period spanning multiple years back. NXP rarely analyzed actual demand and LTP forecasts data of earlier years. Thus, there had not existed an incentive to create a database of actual quantities or forecasts spanning multiple years. Second, in 2015 NXP merged with Freescale Semiconductor. As a result, integration problems existed between the data files originating from the two previous companies. Third, the LTP process is conducted using excel without the implementation of a formal process for controlling data consistency over the yearly iterations of the LTP process. Fourth, the LTP forecasts and actual quantities were rarely analyzed in the same context. Consequently, the consistency of object names and data formats between actual demand and LTP forecasts had not been controlled. As a result, planners regrouped wafer diffusion technologies according to their own needs. These four factors led to the task of integrating around 30 data files with several data quality issues that had to be resolved.

Data quality can be assessed with the following criteria: validity, accuracy, completeness, consistency and uniformity. The validity criterion evaluates the degree to which values match to constraints and business rules. In the data files, constraint violations were found in the column of the wafer diffusion technologies. Some types of production at NXP does not relate to a wafer diffusion technology. This type of validity issue is referred to as a set-membership constraint.

With the help of business experts, 22 groups of alternative demand types were identified and removed from the dataset.

Accuracy refers to whether the measured data values match the true value. In this project's context, accuracy refers to whether the true historical demand is measured. It is difficult to evaluate the accuracy, because there is no external data source available that contains the true values. Thus, the accuracy was evaluated by a thorough examination of the yearly demand figures of each diffusion technology with a panel of business experts. To evaluate the yearly figures, the panel relied on its knowledge of past business development. The completeness of the data and consistency in the naming of wafer diffusion technology groups was also evaluated in this process.

Last, the uniformity of data refers to whether the same units of measure are applied. NXP uses a standard wafer size as the unit of measure to report wafer demand. The standard wafer size facilitates the comparison of production processes that produce wafers of unequal size. In one yearly demand file, the application of an alternative wafer size was identified. All diffusion technologies showed a similar demand fluctuation which was corrected by converting the demand to the standard wafer size.

4.1.2 Selection of Diffusion Technologies

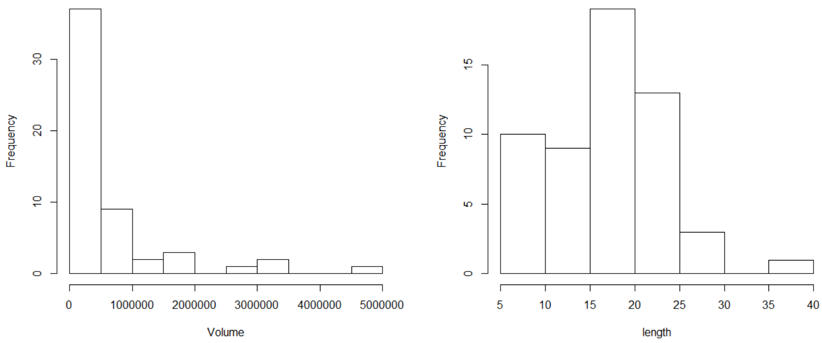
The integration and cleaning of data led to a dataset of 327 wafer diffusion technology groups. Of these technologies, 134 could not be used for analysis because they were launched before 1999 which was out of scope of the data set. Likewise, 37 technologies were removed because they were launched in the last three years and thus lacked data points for analysis. Last, 90 diffusion technologies were removed from the set which lacked significant production volume or had an unusually short life cycle length. The business experts involved in the deletion process referred to these technologies as failed or experimental technologies. Predicting the failure of technologies is outside the scope of this project. The selection process resulted in a database of 66 diffusion technologies. These technologies were divided into a training set of 55 historical technologies and testing set of 11 recent technologies. A technology was considered recent if it was launched somewhere in the last 10 years.

4.1.3 Constructing Descriptive Criteria

The clustering and algorithmic selection procedures required the definition of descriptive criteria of diffusion technologies. A group of business experts was consulted to formulate the descriptive criteria. The experts proposed that the diffusion technologies can be categorized by its markets focus and general technology type. It was not possible to directly extract these descriptive criteria from the dataset. Thus, they were added by business experts. In addition, the total length and volume of diffusion technologies were added to the data as descriptive criteria. Estimates of the volume and length were used for the diffusion technologies in the testing set, because these technologies had not yet completed their life cycle.

4.1.4 Data Characteristics

The diffusion technologies of the training set varied greatly in volume and length. Figure 4.1 shows a histogram of the length and total volume of the diffusion technologies. For most of the diffusion technologies, less than 500,000 wafers are produced in their complete life cycle. A minority of diffusion technologies have a volume between 500,000 and 2,000,000. Some exceptionally large groups have volumes of more than 2,000,000 wafers. In contrast, the distribution of life cycle lengths is less skewed. The life cycle lengths are centered around 15 years, with most of the diffusion technologies between 5 and 25 years. One exceptionally long diffusion technology has a life cycle between 35 and 40 years.

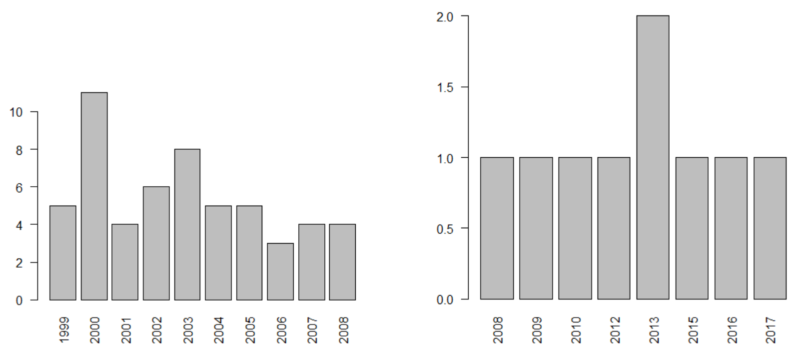


(a) Histogram of total volume (b) Histogram of length of life cycle

Figure 4.1: Histogram of total volume and life cycle length

The historical technologies of the training set were launched between 1999 and 2008. Sub-figure 4.2a shows for each year the number of historical technologies launched. In general, the number of launches equally distributed over the years. However, the number of historical technologies launched decreases slightly with the year.

The last year of demand was not included in the dataset for most diffusion technologies. The long average length of the life cycles meant that 46 of the 55 historical technologies continued to be in use in the present day. The life cycle of the remaining 9 technologies ended within the scope of the dataset. Sub-figure 4.2b shows the number of historical technologies that ended in each year.



(a) Technologies launched per year (b) Technologies ended per year

Figure 4.2: Number of technologies of the training set launched and ended per year

The 11 recent technologies of the testing set were launched between 2009 and 2013. Since yearly demand figure were available up to and including 2019, between 8 and 11 data points were available per diffusion technology. The combined volume of these testing technologies in 2019 represented roughly a quarter of the total volume of 2019.

4.2 Results

This section presents the results of applying the proposed procedures to the diffusion technologies of NXP. In subsection 4.2.1, the result of the preprocessing step of curve fitting is discussed. This step was implemented before clustering and forecasting with the regression model (see section 3.1). Subsection 4.2.2 analyses the results of the clustering step. Next, subsection 4.2.3 outlines the forecasting experiment used to evaluate the proposed procedures. Subsection 4.2.4 presents the results of forecasting with the regression model. The results of similarity based forecasting using growth models and Bayesian updating with similar technologies are presented in subsection 4.2.5. The extension of this procedure by updating the growth models with the LTP forecasts is presented in subsection 4.2.6. Last, the ability of growth models to create prediction intervals is evaluated in section 4.2.7.

4.2.1 Curve Fitting

The curve fitting step was applied as a preprocessing step before the clustering procedure and the regression model (see section 3.1). Four growth models were fitted and the resulting curves were averaged to form the life cycle curve of a technology. The objective of this step was to remove the effect of other influences on the demand of the diffusion technologies.

Table 4.1 presents the success rate of fitting the growth models to the diffusion technologies of the training dataset. The logistic distribution was able to fit to all diffusion technologies. The Weibull distribution was the least successful, which is a surprising result because the Weibull distribution has a the most flexibile shape. It can be used to generate both asymmetric curves like a Gompertz curve and symetric curves like the logistic distribution. Meade and Islam (1998) presented similar results and argued that the logistic model is easier to fit to shorter data sets than more complex flexible models.

Table 4.1: Success rate of fitting models to diffusion technologies

Model	Success	Failed	Success Rate
Bass	45	10	82%
Gompertz	46	9	84%
Logistic	55	0	100%
Weibull	33	22	60%

Figure 4.3 illustrates the curve fitting step with six examples. Each graph presents the results of one diffusion technology. The yearly actual production volumes are represented by dots and the solid line represents the curve fitted by averaging the results of estimating75 four growth models.

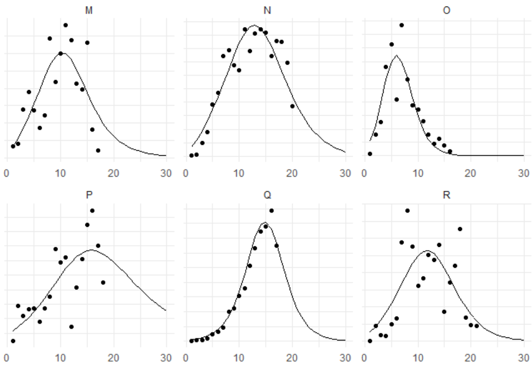


Figure 4.3: Clustered time series of diffusion technologies

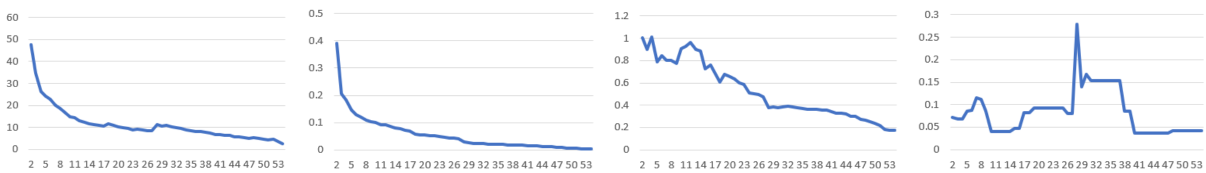
The technologies of figure 4.3 were selected to represent technologies of the training set with varying characteristics. The time series of technology M demonstrate a clear life cycle pattern, while simultaneously having large residuals. Technology N has almost completed its life cycle and shows minimal deviations from the fitted life cycle curve. Next, technology O has a sharp peak and decline. The demand of technology P contains large amount of noise. In contrast, technology Q strongly follows a life cycle pattern. However, this technology has not progressed far past the peak of its life cycle. Last, technology R demonstrates a noisy but complete life cycle. The complete results of fitting curves to the technologies of the training set can be found in appendix B.

4.2.2 Clustering

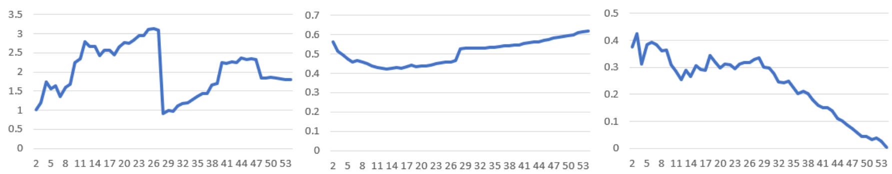
The first task of clustering was to determine the optimal number of clusters. Figure 4.4 shows seven internal validity measures plotted against the number of clusters. The Chalinski-Harabasz and the COP index show an L-shaped curve. The elbow method can be used to determine the optimal number of clusters based on these graphs. These figures suggest an optimal number of three and four clusters, respectively. In addition, the L-shape can be found in the first section of the Score Function. This validity index points to six as the optimal number of clusters. Averaging the results of these three indices leads to four as the optimal number of clusters.

Alternatively, the validity indexes c to g point to 28 as the optimal number of clusters. Index c, Davies-Bouldin, is an index that should be minimized. It displays a decreasing trend till 28 clusters after which the trend stabilizes. Thus, adding more than 28 clusters would increase the model performance relatively less. Next, the Dunn index shows a peak at 28 clusters. A higher Dunn score suggests better performance. In sub figure 4.4e, the Modified Davies Bouldin index decreases sharply at 28 cluster which suggests 28 being the optimal number of clusters. Moreover, the Score Function, which should be maximized, displays a moderate increase in score at 28 clusters. Last, the Silhoutte index should be maximized. This index is relatively stable between 11 and 28 clusters, after which the score starts to decrease sharply. Thus, it suggests that adding more clusters than 28 decreases performance.

Five validity indices suggest 28 clusters as the optimal number of clusters, while combining three indices suggest four clusters. One index is included in both counts, because it suggests two different optimal number of clusters. Further analysis was required to determine which number of clusters was more appropriate. Thus, the k-means clustering approach was implemented once with k set to 4 and once with k set to 28.



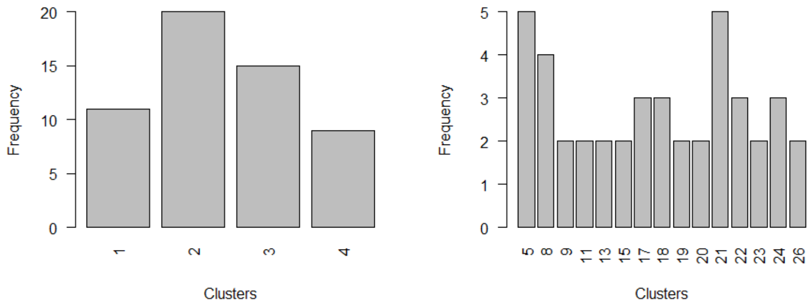
(a) Calinski-Harabasz (b) COP (c) Davies-Bouldin (d) Dunn



(e) Modified DB (f) Score Function (g) Silhoutte

Figure 4.4: Seven cluster validity index scores

Applying k-means clustering with k set to 4 and 28 led to the clusters shown in figure 4.6 and 4.7. The colored lines represent the clustered time series and the grey dashed lines represent the cluster prototype. In the clustering result with k set to 28, a large proportion of the clusters consists of only one time series. This is to be expected when grouping 55 time series into 28 clusters. The number of time series in the 4 and 28 clusters can be found in figure 4.5. For readability, only clusters consisting of more than one time series are included in subfigure 4.5b.



(a) Cluster population k=4 (b) Cluster population k=28

Figure 4.5: Number of technologies in each cluster

Clustering k=4

The performance of the clusters was evaluated by visual inspection. The four clusters in figure 4.6 display different degrees of homogeneity. The curves in clusters one and three are mostly centered around their prototypes. On the contrary, the second cluster contains a large variety of curves. The cluster's prototype, the dashed grey line, does not accurately represent the average curve of all the curves. Moreover, the fourth cluster contains a significant number of curves with large deviations from the prototype. Consequently, if a new technology is assigned to either the second or fourth cluster by the classification model, it would be unlikely that its demand would follow the cluster's prototype curve. Thus, the result of the k-means clustering procedure with k set to four was not used for the prediction models.

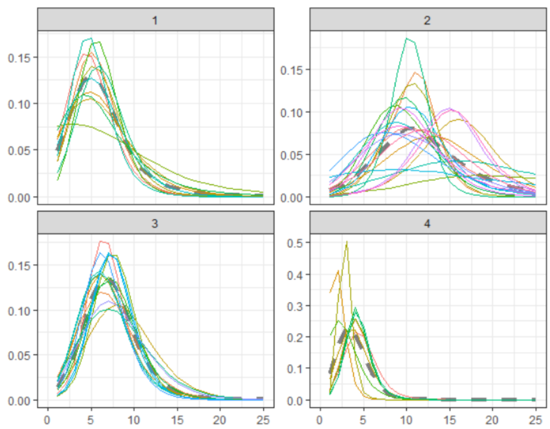


Figure 4.6: Clustered time Series of diffusion technologies

Clustering k=28

Clustering with k set to 28 leads to highly homogeneous clusters. Only cluster 18 does not consist of similar curves. The other clusters of multiple time series consist of time series with minimal differences. A high degree of homogeneity is expected when grouping 55 time series into 28 clusters.

Besides homogeneity in clusters, the heterogeneity between clusters was visually inspected. Large similarities are visible between clusters. One example is the similarity between cluster 4, 5 and 6. These clusters have an average curve of short length and an early peak. Moreover, cluster 10 and 11 have similar average curves of long length and are slightly skewed to the right. Another example can be found in cluster 21 and 22. Both curves are concentrated around their peak and have a medium length.

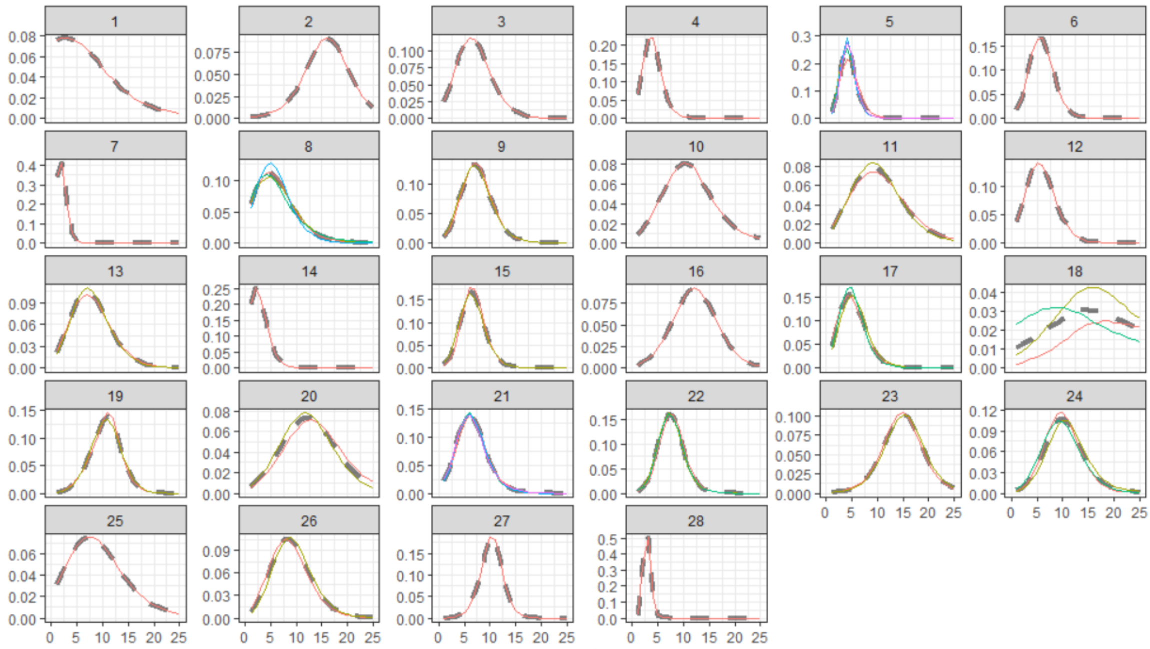


Figure 4.7: Clustered time Series of diffusion technologies

Overall, the time series in a cluster are highly similar and the differences between clusters are minimal. However, slight differences between clusters were identified by the clustering algorithm. While clusters might be similar in degree of spread or concentration, location of the peak and length, other subtle difference can be seen in the slopes and starting values of the curves. The second aspect might be important for the prediction procedures. Some diffusion technologies had very low demand in the first periods compared to some other technologies which had a more rapid ramp up. The technology curve might differ in the first periods and be similar on other aspects. An example would be cluster 4 and 5, where the curve of cluster 4 has a high starting point than the curve of cluster 5.

Classification

The forecasting procedure assigned new products to clusters based on descriptive criteria with a classification model (see section 3.1). Thus, homogeneity within clusters and heterogeneity between clusters of descriptive criteria would facilitate classification. In addition, it might provide valuable insights into the variables determining the life cycle of a technology. Since the clustering results with k set to four were not used for prediction, a classification model was solely trained on the clustering results with k set to 28.

The technologies in each cluster are highly similar in respect to the descriptive criterion of the length of the life cycle. This was expected because curves of similar length are more likely to be grouped together. In contrast, the clusters are not very homogeneous in volume. The differences between the volumes of technologies in a cluster are relatively large. On average, the largest volume difference in a cluster was 1,000,000 wafers. Next, each cluster contains technologies with largely the same market focus. However, it was not possible to identify significant differences between clusters based on this criterion. Last, the general technology type does not relate with the formed clusters. In conclusion, it was not possible to identify a relationship with the clustering of diffusion technologies for any descriptive criteria.

The classification model is able to predict the correct cluster of a technology in 33% of the cases. The accuracy is expected to be even lower for the testing dataset. Thus, the model is more often wrong than right. The poor performance can partly be attributed to the large number of clusters. To illustrate, a random classifier would correctly predict the cluster of 3.5% of the technologies. However, the classification results does not suggest a strong relation between the descriptive criteria and clusters. The next step is to evaluate prediction results of the forecasting models combined with the clustering approach.

4.2.3 Forecasting Experiment Setup

As explained in chapter 3, this project evaluated three approaches for finding similar technologies and two approaches for creating forecasts using similar technologies. Similar technologies were found by a clustering approach, an algorithmic selection approach and an expert opinion approach. For the prediction step, regression models and growth models were applied. All combinations of the three approaches for finding similar technologies and the two prediction approaches were implemented; thus, six procedures were evaluated.

Each diffusion technology of the testing set was forecasted using a rolling window. For each year, a one, two and three step ahead forecast was made. The first forecast was made in the third year of demand, because the procedures require some data points of the prediction target to estimate their parameters. This rolling window forecast simulates the LTP forecasting process, which is a yearly process to make forecasts for the next three years.

The accuracy of the similarity based sales forecasting procedure depends on the level of similarity between the prediction target and the identified diffusion technology or cluster. As a result, the performance of the procedure may vary for each diffusion technology. Thus, the accuracy is reported separately for each diffusion technology of the testing set. Moreover, large differences exist in the total volumes of the diffusion technologies. This led to the selection of the mean absolute percentage error (MAPE) to report the performance.

4.2.4 Regression Model Forecasts

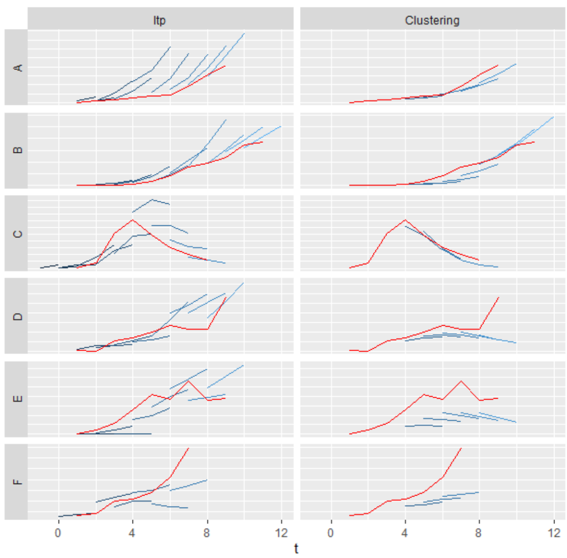
The results of the regression model approach, described in section 3.2.6, are presented in table 4.2. Each row shows the results for one technology and the last row shows the mean performance for all technologies. The first column reports the diffusion technology. The second column reports the accuracy for the LTP forecasts of NXP. The remaining columns report the accuracy of the regression model combined with each of the three approaches for finding similar technologies. The experts were unable to identify similar technologies for technology F and I; thus, no errors are reported in the last column for these technologies.

Overall, the performance of the clustering approach matches the LTP forecasts of NXP. The expert opinion approach performs slightly worse and the selection approach results in considerably worse performance than the clustering approach. The regression model is especially successful for technology A for which it produces low errors in combination with each approach for finding similar technologies. In contrast, the LTP forecasts for this technology are relatively inaccurate. Furthermore, the poor mean performance of the selection approach can be attributed to the highly inaccurate results for technology C, F and J. Last, the expert opinion approach results in an inaccurate forecast for technology C, while producing similar results to the clustering approach for other technologies.

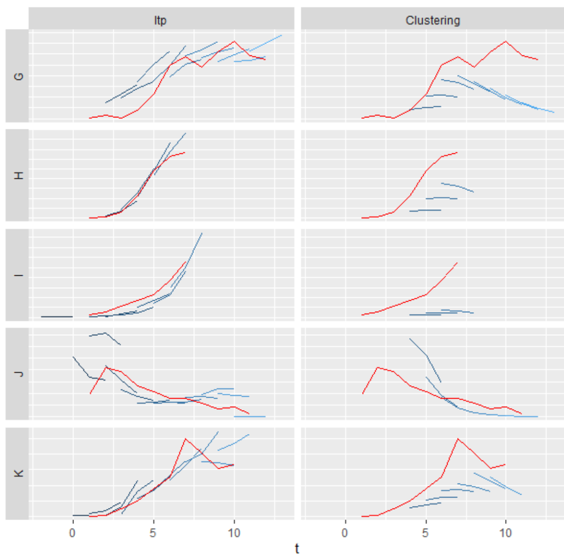
Table 4.2: MAPE of regression model

Technology	LTP	Clustering	Selection	Expert
A	142%	22%	27%	20%
B	49%	41%	34%	87%
C	58%	27%	399%	171%
D	57%	38%	29%	29%
E	29%	60%	96%	48%
F	37%	49%	190%	
G	34%	49%	33%	31%
H	12%	64%	37%	81%
I	68%	86%	69%	
J	58%	75%	637%	74%
K	25%	48%	45%	21%
Mean	52%	51%	145%	62%

These findings confirm the expectation that performance could vary greatly for each diffusion technologies. Thus, further analysis is warranted to identify where performance was gained by the forecasting procedure. The forecasts of NXP and the combination of the regression model with the clustering approach are shown in figure 4.8. The clustering approach is shown in this figure, because it produces the best results. The first subfigure includes technologies A to F and the second subfigure includes technologies G to K. In the subfigures, each row consists of two graphs about one wafer diffusion technology. The first graph shows the forecasts of NXP and the second graph shows the forecasts of regression model with the clustering approach. The red line represents the actual demand and each blue line represents a forecasts for the next three periods. The y-axis indicates demand volume. The values on this axis were removed for confidentially. The x-axis represents the time axis in years, where the first period of demand is set to 1.



(a) Technologies G, H, J and K



(b) Technologies G, H, I, J and K

Figure 4.8: Forecasts of NXP and Regression Models with Clustering Approach

The figure demonstrates that the regression model results in large performance gains for the first three technologies. In contrast, for technology E, H, I and K it produces forecast which do not resemble so much as the general trend of demand. The accuracy of the procedure is comparable to NXP forecasts for the remaining technologies. To sum up, the regression models is able to improve long term demand estimates in some cases; however, it also leads to unexpected results.

4.2.5 Growth Model Similarity Forecasts

In addition to regression models, forecast were created with the approach applying growth models and Bayesian updating. This approach first estimated growth models with the time series of the prediction target's demand. Next, growth models were estimated with the target's demand extended by the demand of a similar technology. Subsequently, these estimates were combined with a Bayesian updating procedure. A full description of this approach can be found in section 3.2.7.

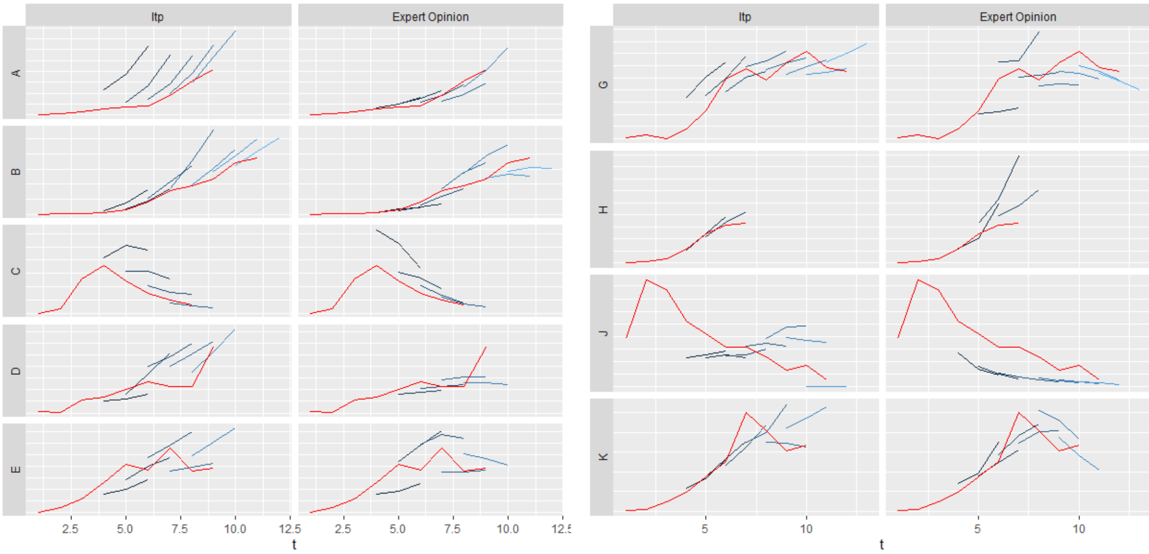
Table 4.3 presents the results of the growth models in the same format as section 4.2.4. The first observation is that the forecasting accuracy has been vastly improved. Whereas the regression models results in large errors for some technologies, the growth models produced reasonable forecasts in all cases. In particular, the results selection and expert opinion approach showed large performance gains. As a consequence, the best performance of growth models was achieved with the expert opinion approach for finding similar technologies. It even provided a considerable advantage over the LTP forecasts. The clustering and selection approach produced a comparable accuracy to the LTP forecasts.

Table 4.3: MAPE of growth model forecast updated with similar technology

Technology	LTP	Clustering	Selection	Expert
A	142%	40%	38%	32%
B	49%	42%	37%	32%
C	58%	51%	70%	54%
D	57%	41%	28%	27%
E	29%	28%	36%	33%
F	37%	53%	41%	
G	34%	55%	62%	27%
H	12%	82%	70%	48%
I	68%	40%	52%	
J	58%	67%	61%	68%
K	25%	25%	22%	22%
Mean	52%	48%	47%	38%

The results are analyzed in further detail with the same figure presented in the previous subsection. As the expert opinion approach performs best with growth models, it is compared with the LTP forecasts. In figure 4.9, the first column of each subfigure again contains the LTP forecasts and the second column contains the forecasts of the combination of the growth model with expert opinion approach. Technologies F and I are excluded from this figure, because the experts could not identify technologies similar to these technologies.

Similar to the regression model, the largest accuracy improvement is realized for the technologies A and B. The technologies H and J proved more difficult to predict. To sum up, the procedure produces comparable forecasts for the remaining technologies. This evidence suggests that growth models are more robust methods than regression models. They are less likely to produce unusable results. A possible explanation is that the inclusion of total volume estimate as input to the growth model prevents large forecasting errors like those produced by the regression model. Another likely explanation is that growth models are able to adapt the shape of the curves to the data of the prediction target and the similar technology, whereas regression models only transform the volume of the curve of the similar technology to produce a forecast.



(a) Technologies A, B, C, D and E

(b) Technologies G, H, J and K

Figure 4.9: Forecasts of Growth Models with Expert Opinion Approach

4.2.6 Growth Model Similarity and LTP Forecasts

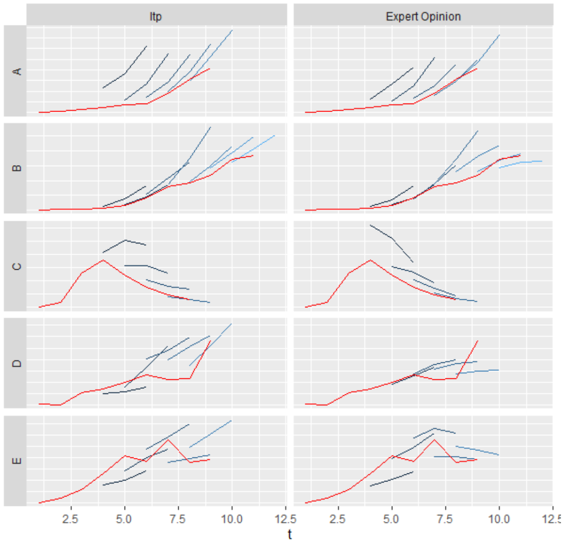
As explained in section 3.2.7, NXP's LTP forecasts are based on many additional demand information sources not used by the similarity based sales forecasting procedure. Consequently, the LTP estimates were included in the growth model approach. The results of this integration can be found in table 4.4 which presents the results in the same manner as the previous two subsections.

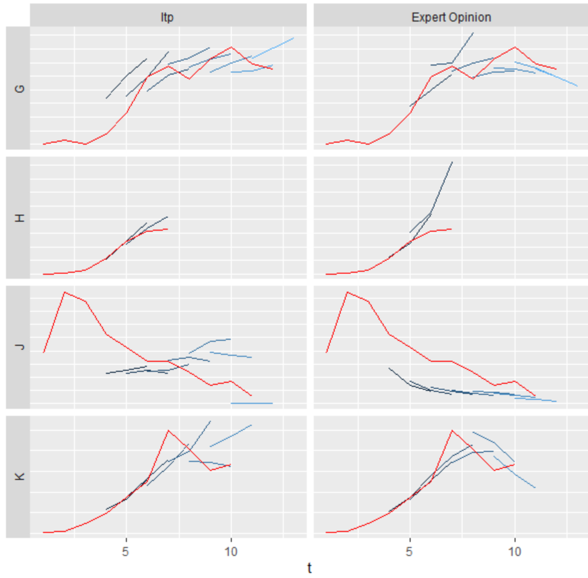
Table 4.4: MAPE of growth model forecast with similarity and LTP

Technology	LTP	Clustering	Selection	Expert
A	142%	60%	112%	98%
B	49%	38%	45%	42%
C	58%	52%	64%	54%
D	57%	53%	34%	37%
E	29%	27%	27%	28%
F	37%	39%	32%	
G	34%	30%	36%	17%
H	12%	54%	51%	46%
I	68%	48%	49%	
J	58%	56%	58%	59%
K	25%	16%	16%	16%
Mean	52%	43%	48%	44%

The mean accuracy changes slightly compared to the previous approach. Part of the improvements are lost for technology A and B, for which the forecast becomes more similar to the inaccurate forecasts of NXP. For the other technologies, the accuracy generally improves a few percentage points. Moreover, the inclusion of the LTP forecast leads to small differences between the approaches for finding similar technologies. Notably, the growth models provide on average slightly more accurate forecasts than NXP.

The forecasts of the expert opinion approach are shown in figure 4.10. In general, the





(a) Technologies A, B, C, D and E

(b) Technologies G, H, J and K

Figure 4.10: Forecasts of Growth Models with Expert Opinion Approach and Integration of LTP

forecasts made with the regression model are more similar to the LTP forecasts. However, significant differences remain for some technologies. In particular, the growth model forecasts of technology J are not similar at all to the LTP forecast. In addition, it does not change significantly from the forecast of the growth models which do not integrate the LTP forecasts (see figure 4.9).

4.2.7 Prediction Intervals

The Bayesian updating procedure required the calculation of the variance of the growth model forecasts. The variance could be used to calculate the 90% prediction intervals. Table 4.5 compares the performance of the prediction intervals. The performance is measured by the fraction of actual demand that fell inside the prediction interval. The first column reports the approach used for finding similar technologies. The second column reports whether the growth model was updating with only the similar technology or with both the similar technology and NXP's LTP forecasts. The percentage is reported for 1, 2 and 3 step ahead forecasts.

Overall, the performance of the prediction intervals is rather poor. In addition, the performance of the prediction intervals consistently decreased with the number of steps ahead. Another consistent result is the prediction intervals of growth models estimated with only similar technologies outperformed the intervals estimated with similar technologies and LTP forecasts combined.

Table 4.5: Percentage of demand inside prediction interval

Approach	Update	1-Step Ahead	2-Step Ahead	3-Ahead
Clustering	Similarity	48%	40%	37%
Clustering	LTP	42%	40%	33%
Selection	Similarity	58%	52%	51%
Selection	LTP	47%	44%	30%
Expert Opinion	Similarity	65%	59%	51%
Expert Opinion	LTP	52%	44%	33%

Chapter 5

Conclusion

This chapter concludes the report. The first section answers the research questions formulated in section 1.4. The second section discusses the relevance of this work for the research area and NXP. The last section addresses the limitations of this study and provides suggestions for further research.

5.1 Revisiting the Research Questions

As stated in section 1.3, NXP wanted to investigate whether the long term estimates of production volumes could be improved. They expected that patterns could be found in the historical production volumes and long-term plans. Thus, the following research goal was formulated:

"Find life cycle patterns in the wafer demand of semiconductor wafer technologies and use these patterns to improve long term wafer demand predictions"

Four research questions were formulated in section 1.4 for reaching the research goal. The remainder of this section discusses the answer to each of the questions.

(1) How can similar technologies be identified by either a clustering or selection approach?

Relevant techniques were identified by conducting a literature review. It was discovered that a clustering approach required a method for clustering the technologies and a classification method for assigning a new technologies to a cluster of similar technologies. The following clustering approaches were implemented in the literature:

- K-means time series clustering with a decision tree classifier
- K-means time series clustering with a probabilistic neural networks classifier
- Hierarchical time series clustering with classification by business expert
- Fuzzy time series clustering with a custom classifier
- Self-organizing maps with a nearest neighbours search

The clustering and classification methods used a measure defining the similarity between two technologies. The first two approaches clustered based on similarity of time series. Next, a classification model was trained on these clusters based on descriptive criteria. New technologies, which lacked historical time series data, could then be assigned to a cluster based on their descriptive criteria. Alternatively, the third approach used business experts to link new technologies to clusters.

The fourth approach again clustered based on time series similarity. However, it also used time series similarity to assign a new technology to a cluster. First, an auto-regressive model was fitted to each cluster. Next, the demand of a new technology was forecasted with each model. Subsequently, a technology was linked to the cluster of which the model resulted in predictions with the smallest distance to the cluster's average time series. The last clustering method created clusters and assigned technologies to clusters based on descriptive criteria. Thus, this method assumed that a relationship existed between descriptive criteria and time series patterns. In addition to the clustering approaches, the following selection approaches were identified:

- Correlation coefficient threshold procedure
- Nearest neighbour search
- Selection by expert Opinion

The first approach calculated a correlation coefficient to measure the similarity between two technologies. If the coefficient exceeded a threshold, the technologies were considered similar. This method identifies technologies similar to the prediction target directly based on their demand patterns. Granted, this method requires sufficient historical data of the prediction target to compare with earlier technologies. The second approach does not require the availability of historical time series data on the prediction target. The nearest neighbour search identifies technologies which are most similar in their descriptive criteria. The last approach lets business experts identify similar technologies and does not rely on any algorithms or similarity measures.

2. How can the demand of a technology be predicted, using a similar technology or cluster of similar technologies?

A review of the literature led to the identification of three relevant prediction approaches:

- Growth Models
- Regression Models
- Prototype curves as predictions

Three methods of estimating growth model were found. Some procedures used the weighted mean of the model parameters of similar technologies. Another procedure estimated growth models first using the historical data of the prediction target and then using the historical data of similar technologies. These two estimates were combined with a Bayesian updating procedure. The last method of estimating the parameters concerned regression models. In this procedure, the growth model parameters of earlier technologies were regressed against their descriptive criteria. Next, the parameters of a new technology were found by applying the trained regression model on its descriptive criteria.

Two types of regression models were identified. The first type regressed the target technology against a time lagged similar technology. The demand of later periods of the similar technology could then be used to predict the target technology. The second type estimated auto-regressive models on similar technologies. Subsequently, the estimated model of the similar technology was used to predict demand. The last prediction method assigned prototype curves of clusters as the forecast. In this approach, the method first assigned technologies to a cluster of most similar technologies. Next, the prototype curve of the cluster was scaled by a total volume estimate to create a forecast. To summarise, a conceptual overview of all the approaches identified with research question one and two is given in figure 5.1.

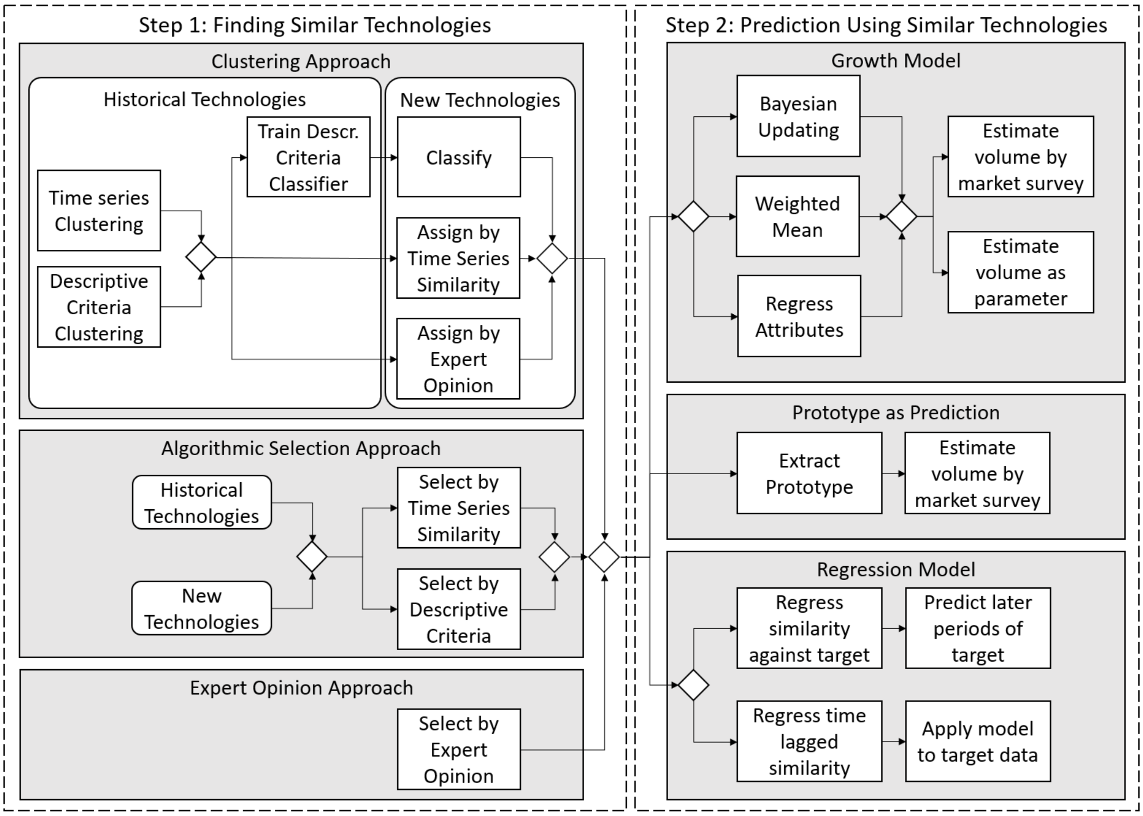


Figure 5.1: Conceptual overview of procedures in the Literature

3. Which combination of techniques identified for RQ1 and RQ2 provides the most accurate predictions?

A number of techniques identified for research question one and two were selected and implemented in a case study at NXP. Regarding the first research question, the clustering, algorithmic selection and expert opinion approach were evaluated. The clustering approach was implemented with k-means time series clustering. Next, the algorithmic selection approach was implemented with a nearest neighbour search. Last, five business experts were asked to identify similar technologies. Regarding the second research question, the growth model and regression model approaches were applied. The estimation of growth models was implemented in combination with a Bayesian updating procedure. The second approach was implemented with the time-lagged regression model which regressed the demand of the target technology against the demand of the time-lagged similar technology.

The best performance was achieved if the forecasts were made with growth models and similar technologies were identified by business experts. In addition, for all approaches for finding similar technologies, the growth model approach resulted in more accurate forecasts than the regression models. The performance gain of the growth models over the regression models was considerable. The regression models resulted in forecasts of poor quality for some technologies, while the growth models produced decent forecasts for all technologies. The accuracy for each combination of techniques is shown in table 5.1.

Table 5.1: Accuracy by prediction model and approach for finding similar technologies (MAPE)

	Clustering	Selection	Expert
Regression Model	51%	145%	62%
Growth Model	48%	47%	38%

It is difficult to determine whether the clustering approach or selection approach was more appropriate. While the approaches produced similar results in combination with the growth models, the results of the clustering approach considerably outperformed the selection approach in combination with regression models. Thus, in this case study the most appropriate approach for finding similar technologies depends on the prediction approach.

4. To what extent can the current market demand estimates be improved with the prediction techniques?

The growth model prediction approach produces on average more accurate predictions than the NXP's LTP forecasts. In combination with the clustering or selection approach, the difference in accuracy with the LTP forecasts is insignificant. A considerable accuracy increase is achieved when growth models were combined with the expert opinion approach. The mean absolute percentage error of the proposed procedure equals 38% and of the LTP forecasts 52%. Overall, the growth models provides adequate forecasts for each technology and is less likely to lead to poor results. Thus, the growth models constitute a more stable procedure and were on average able to improve accuracy in combination with the expert opinion approach.

The regression model in combination with the clustering approach results on average in demand estimates with a similar accuracy to the LTP forecasts. For some technologies, the procedure provides significant improvements, while for other technologies it results in forecasts of poor quality. This finding is consistent for all combinations with approaches for finding similar technologies. To sum up, the regression models present an unreliable method; although it could potentially lead to a large increase in accuracy.

5.2 Relevance

This section discusses the relevance for the research area and the company. The first subsection examines the contributions of this project to the research area. Next, the relevance of this research for NXP is discussed by evaluating how the proposed techniques and its results can be used.

5.2.1 Scientific Relevance

The research area of similarity based sales forecasting is poorly documented. There exist a lack of literature describing the various approaches to similarity based sales forecasting. In the literature review, there were not any review articles or textbooks chapters on the subject identified. The forecasting method was only discussed in textbooks as a qualitative method for new product forecasting. Thus, the literature review contributed to the research area by creating a conceptual overview of the forecasting procedures. The conceptual overview provides a model for the classification of different procedures in the literature. In addition, it can be used as a blueprint for formulating similarity based sales forecasting procedures.

The construction of the conceptual overview led to the insight that the procedures in the literature applied either a clustering procedure, a selection algorithm or relied on the opinions of experts for finding similar technologies. This project implemented all three approaches in a case study to evaluate which approach provided the best performance in the setting of this case study. Notably, the clustering of demand patterns did not lead to clusters which were similar in descriptive criteria or logical to business experts. However, it still produced a decent forecasting performance. Furthermore, it was not possible to conclude that either clustering or a selection approach was more appropriate in the case study setting. Nevertheless, the results showed that the expert opinion approach provided the best performance. Consequently, the evidence suggests that the best approach to similarity forecasting of long life cycle semiconductor technologies makes use of business experts to identify similar technologies. That is, the advanced time series clustering and classification procedure could not identify relationships unknown to the business experts.

In addition to the multiple approaches for finding similar technologies, the project evaluated forecasting with regression models and Bayesian updating of growth models. While a few previous studies investigated multiple methods for parameter estimation of the Bass Diffusion Model, they did not evaluate the application of different forecasting models. This projected demonstrated the clear advantage of growth models over regression models. This result suggests that it is preferable to formulate a complex procedure combining Bayesian updating and growth models instead of a linear regression model.

Another aspect on which this project differs from previous research is the application area. Previous studies have applied similarity based sales forecasting to short life cycle products. This project adds to this research by providing evidence that similarity based sales forecasting can be applied to long life cycle semiconductor technologies. This was not a foregone conclusion because a longer time horizon might include additional factors influencing demand which change the relationship between similar technologies.

5.2.2 Company Relevance

This project analyzed three approaches for finding similar technologies: a clustering, algorithmic selection and expert opinion approach. The most appropriate approach for NXP is to let experts identify similar technologies. This approach resulted in the forecast with the highest accuracy. However, a more important argument for this approach is the ability to interpret of the results. The validity of the forecast can be supported by the expert's arguments for selecting the similar technology. The selection of a similar technology or a cluster of technologies by an algorithm is more difficult to justify. Since the LTP forecasts inform major planning decisions, it is important to provide arguments for a forecast instead of applying black box forecasting models.

Besides the approaches for finding similar technologies, two forecasting models were evaluated. The result of the case study suggested that growth models are more appropriate than regression models. In addition to the superior accuracy, they estimate demand with additional information sources. Whereas the regression model determines a linear relationship between a similar technology and the prediction target, the growth model estimates the demand curve based on the historical data of the target, the historical data of a similar technology and a total volume estimate. Thus, the application of growth models is the most appropriate for NXP.

The added value of the forecasting procedure is that it creates decent predictions without a lot of human effort, while making use of historical data of the prediction target and a similar technology. These demand information sources are underutilized at NXP. Ultimately, the similarity based sales forecasting procedure cannot be used to replace the forecasts of the long-term planning process. Firstly, it does not utilize all information sources available to NXP. Secondly, it could lead to unexpected results. Thirdly, similarity based forecasts do not include sufficient argumentation for their estimates to support major planning decisions. Still, the procedure can lead to large performance gains. Still, the case study demonstrated that the procedure in some cases leads to large improvements of the forecasting accuracy. Thus, an appropriate place for this modelling exercise would be in the review process of the estimations constructed by the product lines. The product lines' estimates could be compared to similarity based growth model forecasts as a tool to discuss the validity of the product lines' estimates.

Two additional useful properties of growth models are its ability to create prediction intervals and scenario forecasts. Stakeholders of NXP have stated these capabilities could assist flexible capacity planning. However, the results showed that the prediction intervals could not be estimated with a high reliability. On the contrary, the demand scenarios could be generated by varying the value for the total volume estimate. The growth model would produce different demand curves which can be used as different demand scenarios.

To summarise, the estimation processes of the product lines include more information sources than were modelled in the forecasting procedure. Besides, growth models cannot match the interpretabilitiy of NXP's forecast. The advantage of the growth model procedure is that it presents a data based method which includes the historical data of the prediction target and a similar technology as demand information sources. As a consequence, a suitable application of these tools would be to facilitate discussions about the validity of the product lines' demand estimates.

5.3 Limitations and Suggestions for Further Research

This section evaluates the limitations of the study and provides suggestions for further research. Each subsection discusses a limitation and provides research suggestions relevant to the limitation.

5.3.1 Additional Predictor Variables

This study made use of four information sources for forecasting demand: historical demand of similar technologies, historical demand of prediction target, an estimate of the total volume and NXP's forecasts. Including additional demand information sources could result in more accurate models. Potential demand predictors could be macroeconomic effects, market effort by NXP, inventory corrections and market indicators such as design-win data or customer estimates. Besides impacting the forecasting performance, the omission of other variables may have caused the poor clustering result. To clarify, time series clustering may have led to clusters of technologies where the similarity of their patterns was caused by other predictors and not by the similarity of the technologies.

The only consideration given to other influences was the application of a preprocessing step of fitting curves which removed periodic fluctuations from life cycle curves. However, it is entirely possible that other factors influence the shape of the life cycle rather than causing only periodic demand fluctuations. For example, a technology launched during an economic upturn might have a significantly larger early growth of sales than if it was launched during an economic downturn.

Additional prediction variables were not included because the long length of the life cycles made it difficult to derive and test relationships between the predictors and actual demand. A prediction model requires a training dataset to derive relationships and a testing set to evaluate the predictive power of these relationships. Ideally, earlier technologies are used to train a model with predictor variables. Later technologies are then used to evaluate the model. This approach would require values for the predictor variables dating back many years. Depending on the specific predictor, it was only be feasible to obtain values for recent years.

In addition to the difficulty of obtaining values for the predictors, a different type of model must be constructed to include the other demand information sources. Some studies propose extensions to the Bass Diffusion model which incorporate other explanatory variables. In their review of forecasting the diffusion of innovation, Meade and Islam (2006) propose that explanatory variables are typically incorporated into the total volume parameter, the probability of adoption or both. To forecast new technologies lacking historical data, they suggest to forecast by analogy, in other words, similarity based forecasting. Thus, a suggestion for further research would be to evaluate procedures for estimating the variables of complex growth models for new semiconductor technologies which lack historical data points.

5.3.2 Bias of Similar Technologies

The forecasting procedure combining Bayesian updating and growth models (described in section 3.2.7) assumed that similar technologies are unbiased estimates of the real demand. It is possible that this demand information source is biased. For example, market characteristics of the semiconductor industry could change over time which affects the rate of adoption of new technologies. These effects are likely to have an influence if a similar technology is launched a many years prior the the technology to predict.

The existence of a bias can be measured with a t-test evaluating whether the estimation errors of the similarity based estimates have a zero mean (Aytac & Wu, 2013). The future demand of the prediction target is unknown; thus, we can only evaluate the bias over the available periods of demand of the target. However, reliable estimation of bias requires many data points on the target which are not available for new products. Secondly, this assumes that the bias will replicate in future periods.

Zhu and Thonemann (2004) proposed an alternative procedure combining growth modeling and Bayesian updating. They had experts assign probabilities that a prediction target would follow the demand of each similar product. These prior beliefs were updated after each period when new data became available. The estimates of demand were weighted combinations of the demand of similar technologies. This approach reduces over time the influence of similar technologies which demand pattern does not match the prediction target. Another study by Chien, Chen, and Peng (2010) performed a test to check whether parameters of the Bass Diffusion Model changed in substituting technologies. Both studies used historical data of the prediction target to evaluate the bias. A direction for further research would be to create a model which is able to predict if a similar technology is biased before demand data of the prediction target is available.

5.3.3 Uncertainty of Total Volume Estimate

An estimation by business experts of the total volume of a diffusion technology is supplied as input to the growth model forecasting procedure. As a consequence, the forecasting performance might suffer from inaccurate estimations of the total volume. Moreover, the uncertainty of this input parameter is not considered in the Bayesian updating procedure. Including this parameter might have resulted in larger model variances. Consequently, the forecast made with the target's data and the similar technology's data would have been combined with different weights (see equation 3.11). In addition, wider prediction intervals would have been constructed.

It is possible to include the total volume as one of the growth model parameters to be estimated. However, previous studies concluded that including the total volume in the estimation of growth models leads to inaccurate results. Thus, the decision was made to provide the total volume as an input to the modelling procedure.

Further research could consider the uncertainty of the total volume estimate by evaluating the sensitivity of the results to the total volume estimate. In this analysis, various values of the total volume are supplied as input to the forecasting procedure to measure its effect on forecasting accuracy. Alternatively, the total volume estimate could be considered as a random variable. In this approach, business experts would formulate a belief about the uncertainty of the total volume estimate in addition to its mean. The distribution of the volume estimate is then included in the explicit density approach for approximating the growth models variance.

5.3.4 Overlap Training and Testing Data Split

The diffusion technologies were split into a training set and testing set. The most recent technologies were selected for the testing set and the remainder was used for the training set. The life cycles of some diffusion technologies in the training set ended after the life cycle of a diffusion technology in the training had already begun. In other words, the periods of demand of training technologies could overlap with testing technologies. The overlapping periods of demand of training technologies were ignored except in the clustering approach. This approach first fitted a curve to all of the available data of each diffusion technology in the training set step.

Subsequently, clusters were constructed based on the similarity of the full life cycles of diffusion technologies (see section 3.1 for a detailed description). As a consequence, the strict division in years used for training data and testing data was not made in the clustering approach.

The inclusion of the overlapping data points improved the fitting of life cycle curves to the diffusion technologies. Excluding these data points, would have led to an approximation of the right hand side of the life cycle curve. Alternatively, removing the diffusion technologies with overlap from the training set would have led to large decrease in the number technologies in the training set. In addition, it would have removed relatively more technologies with longer life cycles.

Nevertheless, in real forecasting scenario these data points would not be available. Further research could only select technologies with a life cycle completed before the training and testing split. This approach would still suffer from a small training dataset, while leaving out valuable information of more recent technologies. An alternative approach could be to cluster either the first periods of demand of each technology or only specific parts of life cycles. The second approach leads to the additional task of defining individual parts in a life cycle.

5.3.5 Dissaggregation of Forecasts

As mentioned in section 1.2, wafer demand is reported on several levels of aggregation at NXP. This project considered demand of diffusion technologies which is the highest level in the process structure. Wafer demand can also be split into product categories or individual products. A better forecasting result might have been achieved if demand of product categories or individual products were forecasted. However, the demand at these levels could only reliably be extracted of the last five years. This time horizon is not enough to split into a training and testing dataset.. Although the studies identified in the literature review applied similarity forecasting at a lower hierarchical levels, they did forecast products with a short life cycle. Thus, a suggestion for further research would be to apply similarity based sales forecasting to long-life cycle products.

References

Aghabozorgi, S., Seyed Shirkhorshidi, A., & Ying Wah, T. (2015). Time-series clustering – A decade review. *Information Systems*, *53*, 16–38. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0306437915000733` doi: https://doi.org/10.1016/j.is.2015.04.007

Ahari, A., Viehl, A., Bringmann, O., & Rosenstiel, W. (2018). Mission profile-based assessment of semiconductor technologies for automotive applications. *Microelectronics Reliability*, *91*(February), 129–138. Retrieved from `https://doi.org/10.1016/j.microrel.2018.08.008` doi: 10.1016/j.microrel.2018.08.008

Aytac, B., & Wu, S. D. (2011, 3). Modelling high-tech product life cycles with short-term demand information: a case study. *Journal of the Operational Research Society*, *62*(3), 425–432. Retrieved from `https://www.tandfonline.com/doi/full/10.1057/jors.2010.89` doi: 10.1057/jors.2010.89

Aytac, B., & Wu, S. D. (2013, 3). Characterization of demand for short life-cycle technology products. *Annals of Operations Research*, *203*(1), 255–277. Retrieved from `http://link.springer.com/10.1007/s10479-010-0771-5` doi: 10.1007/s10479-010-0771-5

Basallo-Triana, M. J., Rodríguez-Sarasty, J. A., Benitez-Restrepo, H. D., Jose Basallo-Triana, M., Rodriguez-Sarasty, J. A., & Dario Benitez-Restrepo, H. (2017, 4). Analogue-based demand forecasting of short life-cycle products: a regression approach and a comprehensive assessment. *International Journal of Production Research*, *55*(8), 2336–2350. Retrieved from `https://www.tandfonline.com/doi/full/10.1080/00207543.2016.1241443` doi: 10.1080/00207543.2016.1241443

Bass, F. M. (1969). A New Product Growth for Model Consumer Durables. *Management Science*, *15*(5), 215–227. Retrieved from `https://doi.org/10.1287/mnsc.15.5.215` doi: 10.1287/mnsc.15.5.215

Cheong, H.-W. (2016). Management of Technology Strategies Required for Major Semiconductor Manufacturer to Survive in Future Market. *Procedia Computer Science*, *91*, 1116–1118. Retrieved from `http://www.sciencedirect.com/science/article/pii/S1877050916313564` doi: https://doi.org/10.1016/j.procs.2016.07.163

Chien, C.-F., Chen, Y.-J., & Peng, J.-T. (2010). Manufacturing intelligence for semiconductor demand forecast based on technology diffusion and product life cycle. *International Journal of Production Economics*, *128*(2), 496–509. Retrieved from `http://www.sciencedirect.com/science/article/pii/S092552731000263X` doi: https://doi.org/10.1016/j.ijpe.2010.07.022

Goodwin, P., Dyussekeneva, K., & Meeran, S. (2013). The use of analogies in forecasting the annual sales of new electronics products. *IMA Journal of Management Mathematics*, *24*(4), 407–422. doi: 10.1093/imaman/dpr025

Heeler, R. M., & Hustad, T. P. (1980, 10). Problems in Predicting New Product Growth for Consumer Durables. *Management Science*, *26*(10), 1007–1020. Retrieved from https://doi.org/10.1287/mnsc.26.10.1007 doi: 10.1287/mnsc.26.10.1007

Hu, K., Acimovic, J., Erize, F., Thomas, D. J., & Van Mieghem, J. A. (2019, 1). Forecasting New Product Life Cycle Curves: Practical Approach and Empirical Analysis. *Manufacturing & Service Operations Management*, *21*(1), 66–85. Retrieved from https://doi.org/10.1287/msom.2017.0691 http://pubsonline.informs.org/doi/10.1287/msom.2017.0691 doi: 10.1287/msom.2017.0691

Ilonen, J., Kamarainen, J.-K., Puumalainen, K., Sundqvist, S., & Kälviäinen, H. (2006). Toward automatic forecasts for diffusion of innovations. *Technological Forecasting and Social Change*, *73*(2), 182–198. Retrieved from http://www.sciencedirect.com/science/article/pii/S0040162504001507 doi: https://doi.org/10.1016/j.techfore.2004.11.005

Jung, C., & Lim, D.-E. (2016). Development of an Adaptive Forecasting System: A Case Study of a PC Manufacturer in South Korea. *Sustainability*, *8*(3). Retrieved from https://www.mdpi.com/2071-1050/8/3/263 doi: 10.3390/su8030263

Keogh, E. (2002). Exact Indexing of Dynamic Time Warping. In P. A. Bernstein, Y. E. Ioannidis, R. Ramakrishnan, & D. B. T. Papadias (Eds.), *Vldb '02: Proceedings of the 28th international conference on very large databases* (pp. 406–417). San Francisco: Morgan Kaufmann. Retrieved from http://www.sciencedirect.com/science/article/pii/B9781558608696500433 doi: https://doi.org/10.1016/B978-155860869-6/50043-3

Kiran, D. (2019a). Capacity planning. In *Production planning and control* (pp. 293–301). Elsevier. Retrieved from http://www.sciencedirect.com/science/article/pii/B9780128183649000202 https://linkinghub.elsevier.com/retrieve/pii/B9780128183649000202 doi: 10.1016/B978-0-12-818364-9.00020-2

Kiran, D. (2019b). Cost of production. In *Production planning and control* (pp. 207–221). Elsevier. Retrieved from http://www.sciencedirect.com/science/article/pii/B9780128183649000159 https://linkinghub.elsevier.com/retrieve/pii/B9780128183649000159 doi: 10.1016/B978-0-12-818364-9.00015-9

Lee, H., Kim, S. G., Park, H.-w., & Kang, P. (2014). Pre-launch new product demand forecasting using the Bass model: A statistical and machine learning-based approach. *Technological Forecasting and Social Change*, *86*, 49–64. Retrieved from http://www.sciencedirect.com/science/article/pii/S0040162513001881 doi: https://doi.org/10.1016/j.techfore.2013.08.020

Meade, N., & Islam, T. (1995, 9). Prediction intervals for growth curve forecasts. *Journal of Forecasting*, *14*(5), 413–430. Retrieved from https://doi.org/10.1002/for.3980140502 doi: 10.1002/for.3980140502

Meade, N., & Islam, T. (1998, 8). Technological Forecasting—Model Selection, Model Stability, and Combining Models. *Management Science*, *44*(8), 1115–1130. Retrieved from https://pubsonline.informs.org/doi/abs/10.1287/mnsc.44.8.1115 doi: 10.1287/mnsc.44.8.1115

Meade, N., & Islam, T. (2006). Modelling and forecasting the diffusion of innovation – A 25-year review. *International Journal of Forecasting*, *22*(3), 519–545. Retrieved from http://www.sciencedirect.com/science/article/pii/S0169207006000197 doi: https://doi.org/10.1016/j.ijforecast.2006.01.005

Milgram, L., Spector, A., & Treger, M. (1999). Understanding Operating Leverage. In *Managing smart* (pp. 387–387). Boston: Elsevier. Retrieved from http://www.sciencedirect.com/science/article/pii/B9780884157526503155 https://linkinghub.elsevier.com/retrieve/pii/B9780884157526503155 doi:

10.1016/B978-0-88415-752-6.50315-5

Nisbet, R., Miner, G., & Yale, K. (2018). Classification. In *Handbook of statistical analysis and data mining applications* (pp. 169–186). Boston: Elsevier. Retrieved from `http://www.sciencedirect.com/science/article/pii/B9780124166325000098` `https://linkinghub.elsevier.com/retrieve/pii/B9780124166325000098` doi: 10.1016/B978-0-12-416632-5.00009-8

Radojičić, V. D., & Marković, G. Z. (2009). New technology forecasting using the bass model. In *2009 9th international conference on telecommunication in modern satellite, cable, and broadcasting services* (pp. 277–280). doi: 10.1109/TELSKS.2009.5339534

Srinivasan, V., & Mason, C. H. (1986, 3). Nonlinear Least Squares Estimation of New Product Diffusion Models. *Marketing Science*, *5*(2), 169–178. Retrieved from `http://www.jstor.org/stable/183671`

Theodoridis, S. (2020). Classification: a Tour of the Classics. In *Machine learning* (2nd ed., pp. 301–350). Elsevier. Retrieved from `http://www.sciencedirect.com/science/article/pii/B9780128188033000167` `https://linkinghub.elsevier.com/retrieve/pii/B9780128188033000167` doi: 10.1016/B978-0-12-818803-3.00016-7

Theodoridis, S., & Koutroumbas, K. (2009). Cluster Validity. In *Pattern recognition* (4th ed., pp. 863–913). Boston: Elsevier. Retrieved from `http://www.sciencedirect.com/science/article/pii/B9781597492720500189` `https://linkinghub.elsevier.com/retrieve/pii/B9781597492720500189` doi: 10.1016/B978-1-59749-272-0.50018-9

Thomassey, S. (2010). Sales forecasts in clothing industry: The key success factor of the supply chain management. *International Journal of Production Economics*, *128*(2), 470–483. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0925527310002598` doi: https://doi.org/10.1016/j.ijpe.2010.07.018

Thomassey, S., & Fiordaliso, A. (2006). A hybrid sales forecasting system based on clustering and decision trees. *Decision Support Systems*, *42*(1), 408–421. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0167923605000114` doi: https://doi.org/10.1016/j.dss.2005.01.008

Thomassey, S., & Happiette, M. (2007). A neural clustering and classification system for sales forecasting of new apparel items. *Applied Soft Computing*, *7*(4), 1177–1187. Retrieved from `http://www.sciencedirect.com/science/article/pii/S156849460600010X` doi: https://doi.org/10.1016/j.asoc.2006.01.005

Tigert, D., & Farivar, B. (1981, 9). The Bass New Product Growth Model: A Sensitivity Analysis for a High Technology Product. *Journal of Marketing*, *45*(4), 81–90. Retrieved from `https://doi.org/10.1177/002224298104500411` `http://journals.sagepub.com/doi/10.1177/002224298104500411` doi: 10.1177/002224298104500411

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2017). Beyond supervised and unsupervised learning. In *Data mining* (pp. 467–478). Elsevier. Retrieved from `http://www.sciencedirect.com/science/article/pii/B9780128042915000118` `https://linkinghub.elsevier.com/retrieve/pii/B9780128042915000118` doi: 10.1016/B978-0-12-804291-5.00011-8

Wu, S. D., Aytac, B., Berger, R. T., & Armbruster, C. A. (2006, 6). Managing Short Life-Cycle Technology Products for Agere Systems. *INFORMS Journal on Applied Analytics*, *36*(3), 234–247. Retrieved from `https://pubsonline.informs.org/doi/abs/10.1287/inte.1050.0195` doi: 10.1287/inte.1050.0195

Yahyaoui's, A., Yahyaoui, I., & Yumuşak, N. (2018). Machine Learning Techniques for

Data Classification. In I. B. T. A. i. R. E. Yahyaoui & P. Technologies (Eds.), *Advances in renewable energies and power technologies* (pp. 441–450). Elsevier. Retrieved from `http://www.sciencedirect.com/science/article/pii/B9780128131855000097` `https://linkinghub.elsevier.com/retrieve/pii/B9780128131855000097` doi: 10.1016/B978-0-12-813185-5.00009-7

Zhu, K., & Thonemann, U. W. (2004, 8). An adaptive forecasting algorithm and inventory policy for products with short life cycles. *Naval Research Logistics (NRL)*, *51*(5), 633–653. Retrieved from `https://doi.org/10.1002/nav.10124` doi: 10.1002/nav.10124

Appendix A

Proof of equation 3.11 and 3.12

The prior distribution of $X_k(T+S)$ is $N(\widehat{X}_m(T+S|\theta_T, \sigma^2)$ with the probability density function:

$$p(X_m(T+S)) = (2\pi\sigma_m^2)^{-1/2} \exp\left[-\frac{1}{2\sigma_m^2}\Big(X_m(T+S) - \widehat{X}_m(T+S|\theta_T)\Big)^2\right]$$

The sample distribution is $N(\widehat{X}_m(T+S|\theta_{T+L}, \tau^2)$ with the probability density function:

$$p(\widehat{X}_m(T+S|\theta_{T+L})|X_m(T+S)) = (2\pi\tau_m^2)^{-1/2} \exp\left[-\frac{1}{2\sigma_m^2}\Big(\widehat{X}_m(T+S|\theta_{T+L}) - X_m(T+S)\Big)^2\right]$$

Following Bayes' theorem the posterior probability density function of $X_k(T+S)$ is:

$$p(X_m(T+S)|\widehat{X}_m(T+S|\theta_{T+L}) = \frac{p\Big(\widehat{X}_m(T+S|\theta_{T+L})|X_m(T+S)\Big)p\Big(X_m(T+S)\Big)}{\int p\Big(\widehat{X}_m(T+S|\theta_{T+L})|X_m(T+S)\Big)p\Big(X_m(T+S)\Big)dX_m(T+S)}$$

where the denominator is constant. Thus, the left hand side is proportional to the right hand side:

$$p(X_m(T+S)|\widehat{X}_m(T+S|\theta_{T+L}) \propto p\Big(\widehat{X}_m(T+S|\theta_{T+L})|X_m(T+S)\Big)p\Big(X_m(T+S)\Big)$$

Substituting with the probability density functions gives:

$$p(X_m(T+S)|\widehat{X}_m(T+S|\theta_{T+L}) \propto$$
$$\exp\left[-\frac{1}{2}\left(\frac{\Big(X_m(T+S) - \widehat{X}_m(T+S|\theta_T)\Big)^2}{\sigma_m^2} + \frac{\Big(\widehat{X}_m(T+S|\theta_{T+L}) - X_m(T+S)\Big)^2}{\tau_m^2}\right)\right]$$

Subsequently, this is rearranged to:

$$p(X_m(T+S)|\widehat{X}_m(T+S|\theta_{T+L}) \propto \left[\frac{1}{2}\frac{(X_m(T+S) - \mu'_m)^2}{\sigma_m'^2}\right]$$

where

$$\mu'_m = \frac{1/\sigma_m^2}{1/\sigma_m^2 + 1/\tau_m^2}\widehat{X}_m(T+S|\theta_T) + \frac{1/\tau_m^2}{1/\sigma_m^2 + 1/\tau_m^2}\widehat{X}_m(T+S|\theta_{T+L})$$

and

$$\sigma_m'^2 = \frac{\sigma_m^2 \tau_m^2}{\sigma_m^2 + \tau_m^2}$$

To conclude, the posterior distribution of the estimate at time $T + S$ for growth model m has mean μ_m' and variance $\sigma_m'^2$.

Appendix B

Curves fitted to training technologies