Assessing the Cost of Assortment Complexity in Consumer Goods Supply Chains by Reconfiguration of Inventory and Production Planning Parameters in Response to Assortment Changes

Dissertation
zur Erlangung der Würde eines
Doktors der Wirtschaftswissenschaften
(Dr. rer. pol.)
der Universität Paderborn

vorgelegt von

Dipl.-Wirt.-Inf. Christoph Danne
33102 Paderborn

Paderborn, September 2009

Dekan:	Prof. Dr. Peter F. E. Sloane
Referent:	Prof. Dr.-Ing. habil. Wilhelm Dangelmaier
Korreferentin:	Prof. Dr. Leena Suhl

Vorwort

Die vorliegende Arbeit entstand während meiner Zeit als Stipendiat der International Graduate School Paderborn und im Rahmen meiner Tätigkeit am Lehrstuhl Wirtschaftsinformatik, insb. CIM und bei der Freudenberg Haushaltsprodukte KG. Das Vorhaben hatte viele Wegbegleiter, denen ich an dieser Stelle danken möchte.

Mein erster Dank gilt meiner Frau Nathalie. Du warst stets mein größter Rückhalt, hast mit mir in der Zeit "so manchen Streifen mitgemacht" und viel Verständnis für verlorene gemeinsame Stunden und meine Launen aufgebracht. Die größten Herausforderungen während dieser Zeit waren nicht wissenschaftlicher Natur. Durch deine Unterstützung hast du zur Entstehung dieses Buches sehr viel beigetragen.

Ich danke allen fachlichen Wegbegleitern, insbesondere meinem Doktorvater Prof. Dr.-Ing. habil. Wilhelm Dangelmaier für die wissenschaftliche Betreuung, die stets offen stehende Tür und die zahlreichen Diskussionen. Frau Prof. Dr. Leena Suhl möchte ich für die konstruktiven Anmerkungen bei Vorträgen während der Promotionszeit und für die Übernahme des Zweitgutachtens danken. Herrn Prof. Dr. Eckhard Steffen danke ich für das Mitwirken in der Prüfungskommission. Den Kollegen in der Arbeitsgruppe Wirtschaftsinformatik, insb. CIM sei für alle fachlichen Diskussionen und für den selten guten Humor am Lehrstuhl gedankt. Wie beruhigend zu wissen, dass man auch "am Rande des Wahnsinns" noch in guter Gesellschaft ist.

Bei der Freudenberg Haushaltsprodukte KG danke ich Herrn Dr. Alexander Moker, der mir mit einem großen Vertrauensvorschuss die Möglichkeit für diese Praxiskooperation gab und als Ansprechpartner und Mitglied der Prüfungskommission stets mit wertvollem Feedback zur Seite stand. Mein besonderer Dank gilt Herrn Oliver Neubert, der die Arbeit auf praktischer Seite mit großem Einsatz federführend betreute und sich immer wieder Zeit für Diskussionen nahm. Frau Dr. Petra Häusler möchte ich für viele Diskussionen und für die Durchsicht der Dissertation danken. Herrn Frank Mönikes von FHP Augsburg danke ich für die Einblicke in die Tücherproduktion. Allen Kollegen bei der FHP danke ich für die angenehme und offene Arbeitsatmosphäre.

Ich danke meinen Eltern, die mich während meiner gesamten Ausbildung gefördert und in meinem Weg bestärkt haben. Schließlich möchte ich mich bei meiner ganzen Familie und meinen Freunden bedanken, die mir stets Rückhalt gegeben haben. Bewusst oder unbewusst habt ihr mir gerade während der "Tiefs" immer wieder gezeigt, dass sich die Welt aus gutem Grund nicht nur um diese Arbeit dreht.

Paderborn, im November 2009

Christoph Danne

Contents

List of Figures

List of Tables

Acronyms

ABC	activity based costing
BCU	basic consumer unit
BOM	bill of material
CLSP	capacitated lot-sizing problem
ERP	enterprise resource planning
FMCG	fast moving consumer goods
GSA	guaranteed service approach
GUI	graphical user interface
MAD	mean absolute deviation
MIP	mixed integer programming
MPS	master production schedule
MRP	material requirements planning
PDN	production and distribution network
PPS	production process step
RCP	rich client platform
RLT	replenishment lead time
SA	simulated annealing
SCOR	supply chain operations reference model
SKU	stock keeping unit
SSA	stochastic service approach
TS	tabu search
TSU	trade sales unit

List of symbols

Production and distribution network structure

$\mathcal{L}$	set of locations, indexed $l \in \mathcal{L}$
$\mathcal{M}$	set of materials, indexed $m \in \mathcal{M}$
$\mathcal{N}$	set of items, indexed $i, j \in \mathcal{N}$
$\mathcal{N}^{PROC}$	set of procurement items
$\mathcal{N}^{PROD}$	set of production items
$\mathcal{N}^{DIST}$	set of distribution items
$mat(i)$	material of item i
$loc(i)$	location of item i
$\mathcal{V}$	set of links v connecting two items, $\mathcal{V} \subset \mathcal{N} \times \mathcal{N}$
$w_{i,j}$	quantity relationship for material flow from i to j
$PR(i)$	set of direct predecessor items of i
$SC(i)$	set of direct successor items of i
$UP(i)$	set of direct and indirect predecessor items of i
$DN(i)$	set of direct and indirect successor items of i
$LS(i)$	set of end product items at sales locations that i goes into
$\mathcal{S}$	set of process steps
$\mathcal{S}^{PROD}$	set of production process steps, indexed $s \in \mathcal{S}^{PROD}$
$\mathcal{S}^{TRANS}$	set of transport process steps, indexed $s \in \mathcal{S}^{TRANS}$
$p_{i,s}$	Binary indicator showing if i is processed on s
$\mathcal{N}_s$	set of items processed on s
$\mathcal{S}_i$	set of processes related to i
$\mathcal{B}_i$	bill of material of i
$PDN(\mathcal{N}, \mathcal{V}, \mathcal{S})$	production and distribution network

Time model

$\mathcal{T}$	set of discrete mid-term time periods $t \in \mathcal{T} = \{1, \ldots, T\}$
T^S	number of short-term periods that constitute one mid-term period
t_i^{trans}	transport time of i
TT_i	throughput time of i
RLT_i	replenishment lead time of i
ST_i	service time that i presents to its successors
Δt_i	coverage time that has to be covered with inventory at i

External processes - demands and service levels

$d_{i,t}^{ext}$	expected primary demand for i in period t
$d_{i,t}$	expected total demand for i in period t
$D_{i,t}$	random variable describing the actual demand for i in period t
$FD_{i,t}$	random variable describing the forecast deviation for i in period t
$\sigma_{i,t}^d$	standard deviation of forecast deviations at i in period t
$\sigma_{i,t}^{dext}$	standard deviation of forecast deviations for primary demands at i in period t
FD_i^{mad}	relative mean absolute deviation of forecasts for item i
α_i^{SL}	α-service level for i
β_i^{SL}	β-service level for i
ST_i^{max}	maximum delivery time for i if i represents an end product sold to customers

Internal processes

K_s	capacity provided by s in one short-term period
k_s	capacity coefficient for production of i on production process step s
Q_i^{rnd}	lot-size rounding value for production quantities of $i \in \mathcal{N}^{PROD}$
PB_{s,pb_s}^{pen}	penalty costs for choosing planning buffer pb_s on s

Scenario definition

$\mathcal{A}$	set of new item additions
$\mathcal{R}^{fin}$	set of material replacement definitions for end products
$\mathcal{R}^{rs}$	set of material replacement definitions for raw and semi-finished materials
$\mathcal{M}^{new}$	set of all materials that are added via the elements of $\mathcal{A}$
d_k^{ratio}	demand ratio for an item replacement
cv_k	quantity conversion factor for an item replacement

Cost assessment

C_i	internal accounting value for one basic unit of i
$\overline{C}^{inv}$	inventory holding cost rate for one period $t \in \mathcal{T}$
C_l^{whsg}	warehousing cost for one storage unit at location l over one period $t \in \mathcal{T}$
C_i^{inv}	total inventory cost for i in one period $t \in \mathcal{T}$
$\overline{C}_{i,s,pb_s}^{stp}$	average setup cost incurred by production of i on production process step s with a planning buffer of pb_s
$scrp_i$	quantity of scrap produced during a production run of i on s
qpp_i	number of basic units of i that can be stored on one storage unit at $loc(i)$

Decision variables

$\mathcal{SP}$	set of stockpoints
$z_{i,t}$	safety factor to determine safety stock levels as multiples of lead time demand variation for i in period t
$RP_{i,t}$	reorder point for i in period t
$I_{i,t}^{ss}$	safety stock level for i in period t
$I_{i,t}^{cs}$	cycle inventory for i in period t
$Q_{i,s,t}$	planned production quantity of i in time period t
pb_s	planning buffer for production scheduling on $s \in \mathcal{S}^{PROD}$
$X_{i,s,t}$	binary variable indicating planned production of i on s in period t

CHAPTER 1

Introduction

What's it going to be then, eh?
Anthony Burgess

Assortment complexity has a major impact on the complexity of the entire supply chain, as the number of products and product variants affects the complexity of production and distribution systems in several ways. Especially for repetitive manufacturing companies like consumer goods manufacturers, it greatly affects planning of production processes and material management.

Managerial decisions on changes of assortment are driven by the trade-off between additional benefits in terms of sales and additional costs incurred by the increased complexity. This has led to much research effort in the area of complexity management and assortment variety in particular. However, the assessment of the effects of assortment-related decisions on the underlying production and distribution network has not yet been systematically undertaken.[1] As the assortment of a company evolves continuously by introducing new or discontinuing existing products, the most important question is *what effects on the configuration of the production and distribution network and related costs can be expected if the assortment is changed in a particular way.* The answer to this question would give valuable decision support for assortment-related decisions.

Methods from costing, especially activity-based costing, seek to assign costs fairly to single product variants according to the input involved. This approach is not suitable for the intended *what-if* analysis, even if the cost assignment were perfectly fair, because of interdependencies between the single product variants that cannot be mapped into a single cost value per product. For example, if a certain degree of standardisation of packaging options across several market regions is reached, it becomes favourable to

1 See Ramdas (2003, p. 49).

keep inventories at a central warehouse instead of local stocks. Thus, statements about potential cost effects derived from assortment changes can only be made for complete assortment change scenarios rather than on a per product basis.

In order to assess these effects monetarily, the question *what changes in the configuration of the production and distribution network can be expected in response to the assortment changes* must be answered first. Only on the basis of a production and distribution network adapted to the new assortment statements about cost effects of these changes can be made. Hence the requirement for methods to adapt the planning and control parameters to the new situation in order to assess the optimisation potential offered by assortment reduction exists. Accordingly, the method presented in this work first seeks a cost-optimal configuration of the most relevant assortment-dependent parameters in the production and distribution network, in order to assess the cost effects of assortment changes on the basis of this adapted network configuration. Thereby the above-mentioned *what-if* analyses can be performed to support assortment-related decisions.

Assortment complexity incurs costs in almost all areas of a company's operation and has therefore been analysed extensively for distinct functional areas and by different methods.[2] This work seeks to evaluate the most relevant assortment-dependent cost positions for the entire production and distribution network, focusing on the areas of inventory management and production execution. This is due to the fact that cost effects in response to assortment changes are particularly expected as a result of changes in inventory requirements and related inventory holding costs as well as setup costs and scrap in the production area. As a consequence, the parameter optimisation for the alternative assortment scenarios adapts the inventory allocation as well as the material requirements planning (MRP) production planning parameters for the production process steps of the network.

In accordance with the statement above that any *what-if* analysis can only be conducted for concrete assortment change scenarios and on the basis of an optimally adapted production and distribution network, this work presents a method to assess the cost effects of assortment complexity in inventory management and production operation in consumer goods supply chains by adjustment of inventory allocations and production parameters in response to assortment changes.

2 For example, Kestel (1995) analyses the effects on logistics operations like transport and picking, Rathnow (1993) describes the cost effects along the entire supply chain and Bräutigam (2004), Heina (1999), Köhler (1988), Schuh (2005) all focus on production operations and analyse the effects of product variety for different types or production systems with different methods.

CHAPTER 2

Problem Statement

> Had no objections, sir, my only questions were 'Where do I go and will I know when I'm there?'
>
> The Thermals

The subject-matter of this work is the quantitative relationship between assortment complexity and related costs in consumer goods supply chains. This chapter defines the problem to be solved in three parts: Firstly, consumer goods supply chains as the system under consideration are described with all relevant notions and concepts (Section 2.1). Secondly, assortment complexity as the phenomenon to be analysed is described and its effects on different aspects of consumer goods supply chains are presented (Section 2.2). Thirdly, the problem is defined and decomposed into smaller subproblems that have to be solved. For each subproblem, the aim and requirements for its solution are defined and a possible solution approach is outlined (Section 2.3). The problem decomposition and corresponding solution approaches together constitute the concept for the required cost assessment.

2.1 Production and Distribution in Consumer Goods Supply Chains

Analysis of assortment-related cost can be seen in the context of supply chain management research.[1] This huge research field has emerged due to the fact that most

1 For a condensed introduction to the subject matters of supply chain management, see the fundamental works of Davis (1993) and Lambert and Cooper (2000).

companies nowadays form part of a larger network consisting of suppliers, manufacturers, logistic service providers and retailers, among others. Such a network is called a supply chain[2].

Definition 2.1 (Supply Chain): *The network of all parties involved in satisfying customer demand for a certain product assortment. These parties comprise component and raw material suppliers, manufacturers, logistic service providers, sales organisations and retailers. They are involved in a variety of partly value-adding processes and are interconnected by upstream and downstream flows of material and information. All supply chain members pursue the aim of producing and delivering end products that satisfy customer demand.*

With respect to the subject-matter of this work, Definition 2.1 captures the most relevant elements from the multitude of definitions presented in the literature and accords with the common understanding of a *supply chain*[3].

In this work we restrict our view to supply chains that produce physical final products and exclude the special case of supply chains for non-material services. In particular, this work deals with consumer goods supply chains, which implies several characteristics for the organisation of the supply chain, the products it produces and the production processes involved.[4] Consumer goods are products purchased by individuals for final personal use. In contrast to durable goods and major appliances, consumer goods are comparably inexpensive and consumed shortly after purchase. This is why they are usually purchased repeatedly at smaller intervals and therefore often referred to as fast moving consumer goods (FMCG). They are often classified according to the three major categories of packaged foods, cosmetics and toiletries, and household products. The corresponding supply chains share most characteristics and the greater part of this work can be applied to all three types of supply chains. This work focuses however on non-food products and household products in particular. It does not take into account the additional peculiarities of food supply chains, like the increased perishability and very limited shelf life of products.

2 Note that although the term supply *chain* suggests a set of sequentially interacting organisations, supply chains usually have a network structure of interconnected organisations (Lambert and Cooper, 2000, p. 65). Although alternative terms like *supply networks* can be found in literature, none of them prevailed. We shall therefore use the commonly accepted notion of a *supply chain* in this work, despite its slight imprecision.

3 Elements of this definition can also be found in the definitions given by Lee and Billington (1995), Chopra and Meindl (2004), Christopher (2005), Sahin and Robinson (2002), Stevens (1989), Hopp (2006) and the Supply Chain Council (2007). The reader is referred to Van der Vorst (2000, p. 22) for a more comprehensive list of supply chain definitions in the literature.

4 For general descriptions of the characteristics of consumer goods supply chains, the reader is referred to Soman et al. (2004, pp. 227-229) and Fleischmann and Meyr (2003, pp. 463-465).

The remainder of this section follows the definition of a supply chain as the system under consideration. Section 2.1.1 describes the participants in consumer goods supply chains as well as the organisation of material and information flows between them. Section 2.1.2 describes the characteristics of the value-adding production processes required to produce the end products.

2.1.1 Organisation and Material Coordination Processes

This work considers the supply chain from the manufacturer's point of view. We distinguish five types of supply chain participants (in the order found upstream in the supply chain):

Customers For consumer goods manufacturers, customers are retailers and wholesalers. Note that the term *customer* is not a synonym for *consumer* here. While the customers are retailers, consumers are those who purchase the products from a retail store and eventually consume them.

Sales companies Customers order at sales companies that ship the products from distribution warehouses to the customers' stores or central warehouses.

Production facilities A number of specialised production facilities each produce a certain product range and supply the sales companies.

Suppliers External suppliers supply the production facilities or sales companies with raw and packaging material as well as selected semi-finished components or finished products.

Logistic service providers The supply chain actors usually make use of the transport and warehousing services offered by third-party logistic providers. These service providers may carry out transports, provide physical storage space and perform order picking and dispatching tasks.

For consumer goods manufacturers, the retailers or retail chains are considered the customers. All customers order the required finished products at sales companies and manage the delivery of these products to the actual consumers.

Sales companies usually supply customers in a defined region, e.g. a single country. These additional storage and transshipment points in terms of sales companies are required for two main reasons: customers like retail stores usually expect very short order fulfilment within 1-2 days from the manufacturers. As shipments directly from the production sites would usually take too long, finished goods are shipped from local distribution warehouses, which may e.g. be located at the sales company sites. Transport costs are, of course, an important consideration in consumer goods distribution. Given the comparatively low value and large volume of the products, transport costs are particularly high and constitute a considerable fraction of total costs. As customer orders contain products that are likely to be produced at several production sites, an additional point for transshipment and consolidation of goods is required to achieve efficient transport.

Considering that consumer goods are fairly inexpensive, production facilities are comparatively capital intensive. They are thus often specialised in production of a certain product range and supply many, if not all, sales companies with these types of product. Consequently, sales companies procure finished goods for their sales activities from a number of production facilities at various locations to supply all customers in their sales region.

Many consumer good supply chains comprise a focal company, typically the main manufacturer and owner of the product's consumer brands.[5] This focal company has to oversee the entire supply chain and manage all relations with raw and packaging material suppliers and customers. The sales companies typically belong to the focal company as well and manage all relations with the customers, i.e. the retail stores. Thus we consider a decentralised supply chain in which central coordination and control can be assumed to exist between the production sites and sales companies.

The central control is implemented as a common material coordination system that integrates the production and distribution processes. If there is the possibility of employing a central material control system, it is reasonable and common to use program orientated material coordination, like the well-known MRP[6] system. The starting point of MRP-based planning is a master production schedule (MPS) that defines the required production quantities.

The finished products have to be available from stock in local warehouses at the sales companies to meet the short delivery times and high service levels. Alternatively, the replenishment lead times at the sales companies must be extremely short to allow procurement from production sites only on customer orders. As the total lead time for procurement and production operations is generally much longer than the delivery times accepted by customers, the material coordination is mainly based in forecasts issued by the sales companies. These forecasts contain the expected sales quantities

5 For more detail on supply chains with focal companies, see Lambert and Cooper (2000).
6 See Orlicky (1975). Modern material requirements planning systems usually employ the extended MRP II concept (Vollmann et al., 2004).

for a certain period so that procurement and production activities can be planned with the required lead time based on these forecasts. The MPS required for MRP based demand planning is then derived from the forecasts issued by the sales companies.

It must always be expected that forecasts are not totally correct and actual requirements differ from the forecasted values.[7] The fact that most production and procurement activities are based on uncertain forecasts requires some means of buffering against the uncertainties that are inherent in forecasts in order to meet the required service levels and delivery times to customers. A variety of strategies can be employed to hedge against these uncertainties.[8] The most widely used means of achieving this is the installation of safety stocks at determined points in the supply chain.

Definition 2.2 (Safety stock): *Safety stock is stock used to protect against uncertainty that arises from internal processes like production lead time, from unknown customer demand and from uncertain supplier lead times.*[9]

From Definition 2.2 it follows that the main drivers for safety stock requirements are production and transport disruptions, forecasting errors, and lead time variations. Safety stock is held to improve customer service and to avoid lost sales and loss of goodwill[10]. The decision where to place which amount of safety stock in order to meet all customer requirements at minimal cost is an important tactical decision problem in supply chain management.

Figure 2.1 depicts the material coordination and planning processes described so far, together with the effects of safety stock placement decisions. At the very final stage, sales companies are faced with uncertain customer demand for all products of the assortment offered. As these demands may change daily, forecasts are made for a mid-term period to allow the calculation of planned requirements for the upstream stages[11]. As the replenishment lead times (RLTs) agreed with preceding stages have

7 As Nahmias states: *"the main characteristic of forecasts is that they are usually wrong"* (Nahmias, 1997, p. 61). For a more detailed analysis of problems in forecasting consumer demand in the fast-moving consumer goods industry, the reader is referred to the empirical study of Adebanjo (2000), who investigates the demand-forecasting systems of some leading UK food companies.

8 Extensive discussion of competing techniques and literature reviews on uncertainty handling in MRP-driven manufacturing systems are provided by Koh et al. (2002), Whybark and Williams (1976), Guide and Srivastava (2000), Buzacott and Shanthikumar (1994) and Enns (2002).

9 See Stadtler and Kilger (2005, p. 61).

10 See ibid., p. 61.

11 Since common in supply chain literature, we shall use the terms *upstream* and *downstream* throughout this work to refer to directions of material flow in the supply chain. Downstream refers to the flow from external procurement to the customer stages, while upstream refers to the opposite direction.

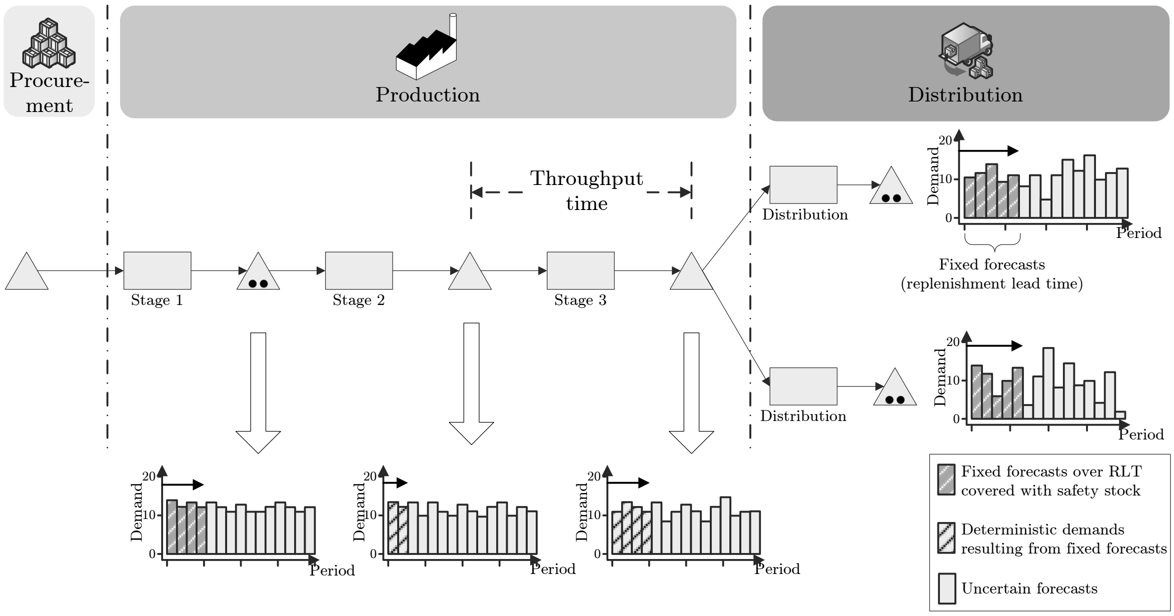


Figure 2.1: Overview of supply chain planning processes

to be respected, these forecasts have to be transferred to fixed orders according to these replenishment lead times. If the required delivery time is shorter than the replenishment lead times, the stage has to keep safety stocks to buffer against demand fluctuations over this time interval.

This principle is especially clear for the interaction of sales companies with production facilities, as delivery times to customers are short and replenishment lead times are longer due to the lead time required by production and transport. However, this principle can be applied to any stage of the supply chain, including production stages. Some stages can operate without inventories since they operate on orders that successor stages have issued by fixing uncertain forecast values over their replenishment lead times. These successor stages then need to keep safety stock to buffer against the variations of actual demands from these forecasted values.

If a stage has fixed its forecasts over a certain replenishment lead time, this fixed time period is reduced at each preceding stage by that stage's throughput time. As long as a positive difference remains, preceding stages can operate deterministically. The fixed time period is consumed successively and eventually upstream stages are required to keep inventories again as the demands they receive are forecasts that are still subject to change. The trade-off in such a system is as follows: the longer the replenishment lead times of the inventory carrying stages, the more preceding stages can operate on fixed orders, as their throughput times are covered by these fixed planning intervals.

However, longer replenishment lead times also increase the safety stock required at the inventory carrying stages.

This high-level description of the planning and material coordination processes reveals that there are many interrelations between decisions on the required inventories and agreed throughput and replenishment lead times of all stages in a supply chain, which leads to complex decision tasks, especially if complex assortments are considered.

2.1.2 Production Processes

As production capacities are comparably expensive in relation to the product value, production systems for consumer goods are designed to efficiently produce large quantities of certain goods. Production processes for consumer goods accordingly follow a serial production oriented layout[12]. Typical characteristics of such production systems are that[13]

- product development and engineering activities are independent of customer orders
- end products are produced to stock based on demand forecasts, independently of concrete customer orders
- customer orders are filled immediately from stocks of finished goods
- stocks of finished goods cause high levels of working capital and a high risk of obsolescence due to imprecise forecasts of future demands
- precise forecasts are of crucial importance

In such production systems large batches of certain goods are produced with lot sizes $\gg 1$ as each changeover from the production of given goods to another requires some expensive setup procedures. The reduction of these unproductive setup times and thereby the increase of the production system's utilisation rate is one of the primary goals of production planning in serial production systems.

The production process of any finished goods typically comprises several distinguishable production process steps. A production process step (PPS) takes a set of input components and transforms them into a certain quantity of an output material. Both

12 See Hopp and Spearman (2000, pp. 8-10) for an overview of production layout types. Sometimes, the term *batch production* is used synonymously.
13 See Dangelmaier and Warnecke (1997, pp. 10-12).

the input and output materials may be physically stored as inventories. A production process step typically represents the production process on one major production line.

Production of consumer goods can be subdivided into the two major parts of base production and converting. Base production comprises all activities required to produce the base products in high volumes. Typically, base production stages use volume, weight or length as the unit of measure in contrast to the quantities in discrete manufacturing. The output of this part of the production process is large units of the actual final goods that have not yet been split up into smaller sellable units and that therefore can be stored efficiently.

Consequently the main activity of the converting stages is packaging, possibly after some smaller assembly and converting procedures. Single production steps of the converting stage comprise the division of the batch quantities received from base production into smaller units as sold to the consumer and the subsequent packaging procedures. Components for converting processes comprise the output of base production and mostly externally procured packaging materials like boxes, paper and foil for end consumer units and larger boxes to form batches for shipment to retail stores.

> **Example 2.1** In cloth production, fibres and chemicals are first mixed and compacted at the base production stage. The resulting raw cloth is further refined by applying additional coatings, prints and sometimes imprinting structures. The converting steps then use the wide cloth rolls that form the output of the base production part as input components and the rolls are cut into smaller lanes and single small cloth. In the final step, cloths are folded, packed into foil and then into boxes in larger batches for distribution to the sales companies.

Base production stages are capital intensive and require little labour. Converting stages are typically more labour intensive, as assembly and packaging activities may require some manual work, depending on the degree of automation of the production process. Often there are highly automated processes for some products and manual operations for others, depending on their value and volume. There may also be various ways of producing a single product via alternative routings that vary in their degree of automation.

All production stages face long and often sequence-dependent setup and cleaning times. The degree to which setups are critical and sequence-dependent varies between the single process steps and with the machines available. Generally, both the absolute importance and also the sequence dependence of setups is more critical in base production than in converting stages since the setups required for each changeover are much longer and comprise cleaning activities that depend on the predecessor product in the production sequence. However, the principle of sequence-dependent setups also holds in the converting stages and causes a similar decrease in the machine's utilisation rates, although their absolute value may not be as high due to the less capital-intensive production resources.

Production stages also produce scrap, either as a fixed amount during the start and end of production runs or as scrap that occurs as a fraction of the volumes produced. Unlike the former scrap type, the latter cannot be avoided or reduced by production planning and sequencing decisions and therefore is not of interest in this work.

The cost implications of these setup and scrap related characteristics result in the requirement for big production lots, as each changeover and start of a new production run incurs setup and scrap cost. Apart from this economical minimum lot size, there are technical restrictions to lot sizes on different production lines. Depending on the type of production output, there exist rounding values for the lot sizes, so that each production run must produce a multiple of this rounding value. Such restrictions result from the physical storage characteristics of the output, which e.g. require full pallets.

The sequence-dependent setup costs are the main reason for installing a *planning buffer*, i.e. a time interval in which the actual production order can be shifted. This gives the planner the flexibility to create sequence-optimised production plans and thereby reduce setup costs.

Definition 2.3 (Production planning buffer): *The planning buffer of a production process step determines the calculation of requirement dates for all materials processed at that stage. It is the time between the provision of all components (earliest production start date) and the latest production start date of a production order. Within this time interval the actual production start can be shifted in order to create cost-optimal production sequences.*

Figure 2.2 illustrates the calculation of requirement dates for a simple example with 2 production stages and one procurement stage. With a given requirement date for the finished product, the requirement dates for all components are determined by calculating the latest possible production start date with the actual processing time for that order. In addition, the production planning buffer of the corresponding production process step is added to determine the requirement dates for all required components.

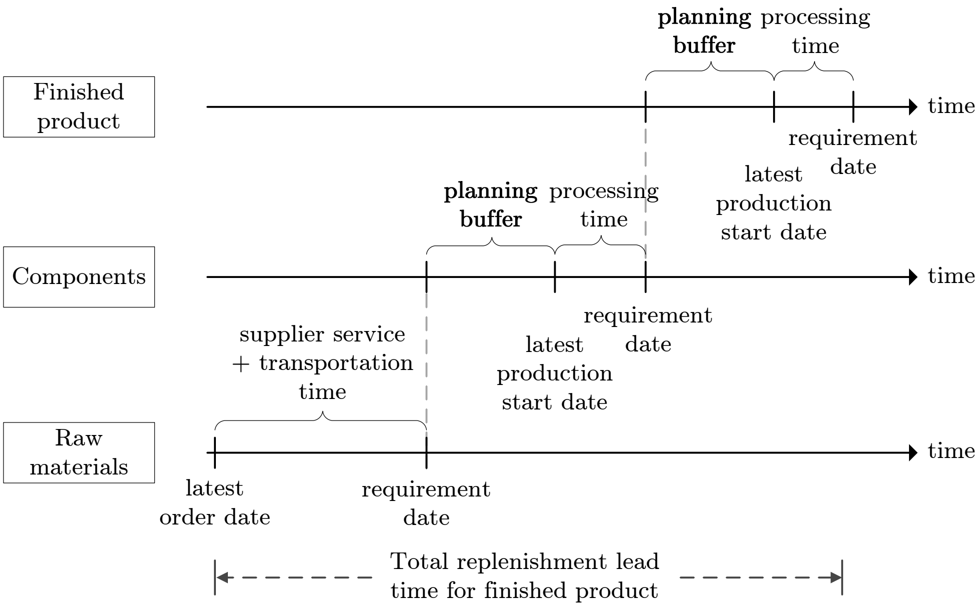


Figure 2.2: The planning buffer concept

Likewise, the requirement dates for all externally procured raw materials are determined at the second production stage. With the delivery time of the supplier and the required transport time, the latest order date for the raw materials can be determined. The difference between this latest order date and the requirement date of the finished product represents the total replenishment lead time for that product.

This simple example demonstrates the trade-off related to the determination of the planning buffers: the shorter the buffer interval, the fewer possibilities to create sequence-optimised production plans. Therefore longer planning buffers can reduce the average setup costs incurred by the changeovers required to process a certain set of production orders. However, this relationship is non-linear, as the additional benefit of longer planning buffer decreases. On the other hand, increasing planning buffers also increase the throughput times for that production stage. This also affects subsequent stages, as their replenishment lead times increase linearly with that throughput time. For those stages that hold safety stock, these increased replenishment lead times may result in longer coverage times and increasing safety stock levels. Figure 2.3 illustrates these relationships.

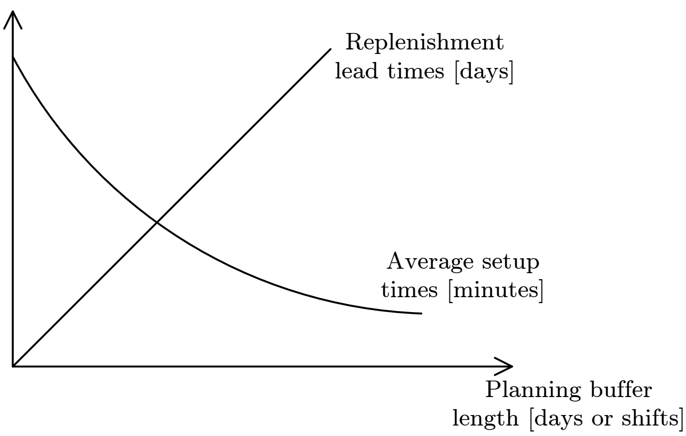


Figure 2.3: Planning buffer, replenishment lead times and average setup times

2.2 Assortment Complexity

2.2.1 Notions and Concepts

The aim of this work is to assess the effects of assortment complexity in consumer goods supply chains. The characteristics of the latter having been described in the preceding section, this section discusses the concept of assortment complexity in more detail. First of all it must be noted that many works on this topic use the term *product variety* rather than *assortment complexity.* While these terms are sometimes used synonymously, we favour the latter throughout this work[14] since both the terms *product* and *variety* may be misleading.

Firstly, the term 'product' is generally used to refer to finished goods requested by customers to serve a certain purpose. However, identical products may still vary in their packaging, which is an important complexity driver for consumer goods and therefore also considered in this work. We thus prefer the term 'assortment' as defined below.

Definition 2.4 (Assortment): *The set of packed end products distributed to a set of sales locations to satisfy customer demand.*

Secondly, the term 'variety' suggests that only the mere number of different product variants is relevant. It can be stated however that relations between materials, both where-used relations in production and transport links in distribution, affect supply

14 As there is no commonly accepted definition for either of these terms, references made to other works also include works that use a different notion.

chain cost as well. These relations thus have to be considered since they directly result from the design of the assortment. This leads to the concept of complexity.

The concept of *complexity* lacks a commonly adapted definition in literature. It is discussed in many disciplines, including system theory, computer science, philosophy and economics and thus different definitions are given.[15] Approaches to a definition of *complexity* in general agree that it is best defined via the notions of elements and relations, as described in the comprehensive discussion of complexity given by Luhmann[16]. One general definition that is applicable to all types of systems and consistent with most of the specialised definitions from different disciplines is given by Buhr and Klaus[17]:

Definition 2.5 (Complexity): *The complexity of a system is determined by the number of elements in the system and the relations between these elements.*

Definitions 2.4 and 2.5 underline that assortment complexity refers both to the size of the assortment in terms of all end products with their packaging variants as well as the relations between materials in terms of where-used relations in production and transport relations at the distribution stages.

The relevance of assortment complexity is supported by the fact that is has a predominant effect on the total supply chain complexity. This supply chain complexity may be defined in terms of the dimensions structure, products and processes.[18] Due to its generality, Definition 2.5 may be used to describe complexity in the structure and process dimensions as well. The structural complexity is captured e.g. via the number of suppliers, production stages, warehouses and customers and the relations between them. Process complexity may be considered as the number and design of material flow or information flow processes. Supply chain complexity management has attracted increasing attention of both scientists and practitioners, as this complexity is assumed to diminish total supply chain performance.[19] With an increasing complexity

15 An overview of literature on complexity theory is given by Anderson (1999).
16 Luhmann (1980).
17 Buhr and Klaus (1975). A similar definition is also given by Ulrich and Probst (1988).
18 See the *internal complexity* in the framework presented by Blecker et al. (2005) and the framework given by Danne et al. (2008). Adam (2004) also claims that complexity can hardly be formalised since it comprises many dimensions. He argues that e.g. the number of products, parts, customers, suppliers and value adding processes are all interrelated factors that make up the complexity of a supply chain (Adam, 2004, p. 20). Further classifications of supply chain complexity can be found in Wilding (1998), Meier and Hanenkamp (2004) and Perona and Miragliotta (2004).
19 In a comprehensive empirical study in the household appliances industry Perona and Miragliotta (2004) provide evidence for the claim that a direct relation between logistical complexity and supply chain performance holds.

of the supply chain, its size and the number of interactions grow and planning[20] and controlling[21] becomes more and more difficult.

Definition 2.1 states that the ultimate goal of all supply chain activities is to *satisfy customer demand for a certain product assortment.* Decision on size and structure of the assortment directly affects the supply chain in its products, structures and processes and thereby the complexity in the product dimension mainly determines the complexity in all other dimensions as well. This observation is especially true for consumer goods supply chains, where assortment complexity has already been identified as a chief determinant of complexity.[22] In contrast to other industries that suffer from high product variety, each variant of consumer goods is not only a theoretical variant that the production system must be able to produce in case of a given customer order, but a real variant that has to be produced, distributed and kept available for customers at possibly different locations at all times. Therefore assortment-related decisions directly affect inventory management[23] and the production process.

Depending on the industry and the types of product considered, different drivers of assortment complexity can be identified.[24] For consumer goods, the most important drivers at the design level are variations in colour, size and materials used. Another important driver especially found in the case of consumer goods is variety in packaging options. Packaging takes place in several steps and has to be considered on various levels: The logistical units shipped between production and sales locations are not identical to those purchased by consumers. The latter, called basic consumer units (BCUs), are packed units of single or few products that are purchased by the consumers. The so-called trade sales units (TSUs) constitute the logistical unit shipped from production locations to sales companies and to customers. These TSUs are larger

20 *Planning* is defined as the process of solving the problem that a current or anticipated state deviates from the desired goal state by anticipating future actions that create the desired goal state (Klein and Scholl, 2004, p. 1). For an overview of the multitude of planning tasks in a supply chain and their classification according to the time horizon affected (short, mid and long-term planning) and the related supply chain process (procurement, production, distribution and sales) see the *supply chain planning matrix* (Rohde et al. (2000), Stadtler and Kilger (2005, p. 87)).

21 *Controlling* is a process in a system in which one or more input values influence output values according to the system's principles (DIN19226, 1968).

22 Hoole reports that *"Product proliferation is the leading driver of supply chain complexity"*, according to the management executives responding to a supply chain complexity survey conducted by the consulting firm PRTM (Hoole, 2006, p. 3).

23 Inventory management comprises the tasks of *"Deciding which inventories are to be held at which stage, in what quantity and for what purpose [...]. It is basically a resource allocation decision - establishing the overall inventory 'budget' that firms can afford and setting targets for how much will be allocated to each inventory type"* (Warner, 2001, p. 2395).

24 For an elaborate list of product variation features, see Ramdas (2003, p. 5).

batches of several BCUs. Thus variety in packaging can first be introduced in the individual packaging of the BCUs, e.g. via localised packaging with country-specific imprints. Secondly, it can be introduced in the batch sizes and case dimensions of the TSUs.

There is an undoubted trend to a continuously increasing number of stock keeping units (SKUs), which is often referred to as *product proliferation*. Baker notes that this *"trend is most noticeable in the consumer packaged goods industry"*[25]. This observation is supported by various empirical studies on different types of consumer goods that have indicated a strong product proliferation and accompanying increase of assortment complexity in recent years.[26] The number of products available in large supermarkets has increased from the order of 1000 in the 1950s to 30,000 in a modern supermarket[27].

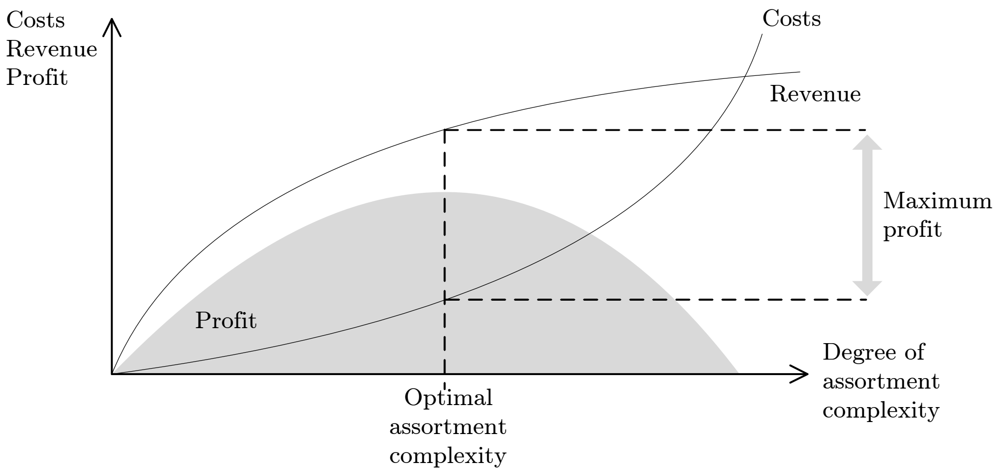


Figure 2.4: Cost and benefit trade-off of assortment complexity[28]

Assortment complexity creates both potential benefits as well as increasing costs via various levers. The trade-off results from the fact that increasing assortment complexity allows the exploitation of *economies of scope*, while reduced assortment complexities allows the exploitation of *economies of scale*, especially in the areas of production and inventory management.[29] This leads to the question of an optimal assortment, i.e. the one that yields the highest profit, as depicted in Figure 2.4. As assortment complexity affects both costs as well as revenue, there is at least one *optimal* assortment complex-

25 See the entry *product proliferation* in Baker (2002)
26 See Quelch and Kenny (1994) and the recent study conducted by The Economist Intelligence Unit (2008).
27 See Thonemann and Bradley (2002, p. 594).
28 Adapted from Herrmann and Peine (2007, p. 654) and Rathnow (1993, p. 44).
29 See Lancaster (1998, p. 8), Lancaster (1990, p. 191).

ity where the difference of revenue and costs is at maximum. The following sections briefly outline the factors involved in this trade-off.

2.2.2 Benefits of Assortment Complexity

Benefits of assortment complexity are reflected in gains in market share and increasing revenue. They can be summarised as follows:

Meeting heterogeneous consumer preferences If we describe each product via the combination of characteristics or attributes it provides, then individual consumers have different preferences with regard to these combinations. A large product portfolio better meets the preferences of different consumers and can increase turnovers, both due to larger sales quantities as well as better acceptance of higher prices for the preferred products.[30]

Price discrimination Different consumers are willing to pay different prices for a certain type of products. Offering a wide range of product variants at different prices thus allows the seller to address a larger consumer target group.[31]

Variety seeking Consumers seek diversity in their choice of goods from time to time. A wide product portfolio can therefore help to prevent consumers from changing to competitor products.[32]

Scatter shot approach Offering a great variety of products may be used as a way of gaining information about the consumers' preferences. As consumer preferences are difficult to foresee, companies may use the so-called scatter shot approach and launch more products that they expect to sustain in the long-term, just to see how consumers accept the different variants.[33]

Demarcation from competitors Product variants are used as means of demarcating products of a certain brand from their competitors. This reduces competitive pressure and avoid direct price competition.[34] At the same time, the active occupation of market segments raises market entry barriers to prevent the emergence of new competitors in that area.

---

30 See Kahn and Morales (2001, p. 64).

31 See Ulrich (2006, p. 7).

32 An extensive review of the different type of variety seeking with explanations from both psychology and marketing literature is given by Kahn (1998).

33 See Lancaster (1998, pp. 15-16).

34 See Lancaster (1990).

Globalisation Given that most consumer goods manufacturers conduct sales activities in different countries, product differentiation due to different languages and legal regulations becomes necessary.

Shelf-space effects Consumer goods manufacturers or brand owners compete for shelf-space in retail stores, as the shelf-space occupied by products of one brand is assumed to correlate with the corresponding turnover. It is thus alluring to widen the product portfolio to occupy more shelf-space.

Trade pressure Direct customers in terms of retail stores and trade may demand a high variety in end products. Such requirements may comprise different packaging sizes or styles to fit their logistical or marketing needs, as well as customised product variants that are unique to their store and also impede a direct comparison with competitor offerings by the consumer.[35]

All the above-mentioned potential benefits are hardly measurable. Eventually, all approaches that try to assess the benefits of assortment complexity have to answer the question how consumers change their behaviour and purchasing habits when confronted with more or fewer product variants. Research approaches mainly comprise qualitative methods and empirical studies mostly discussed in marketing literature.[36]

2.2.3 Costs of Assortment Complexity

There is consensus that assortment complexity is a major influencing factor for supply chain wide costs and supply chain performance.[37] Kluge et al.[38] show empirical evidence that successful companies are characterised by a stronger focus within their assortment, compared with less successful companies. Quelch and Kenny[39] believe that many manufacturing companies have allowed their assortments to proliferate too much. Indeed, some major companies including large consumer goods manufacturers have already taken measures in order drastically to reduce assortment complexity

35 See Quelch and Kenny (1994, pp. 154-155).
36 See e.g. Riemenschneider (2006) and Kahn (1998) and the references given there.
37 Homburg and Daum provide a systematic overview of complexity-induced cost and corresponding levers to influence them (Homburg and Daum, 1997). More evidence for this hypotheses is given by Klaus (2005) and Hoole (2006).
38 See Kluge et al. (1994, p. 41).
39 See Quelch and Kenny (1994).

and obtain considerable cost savings[40]. The assessment of cost incurred by assortment complexity interrelates with almost all aspects of a company's logistics and production activities.[41] Figure 2.5 illustrates how assortment complexity affects the areas of supply chain uncertainty, inventory management and production planning and related costs.

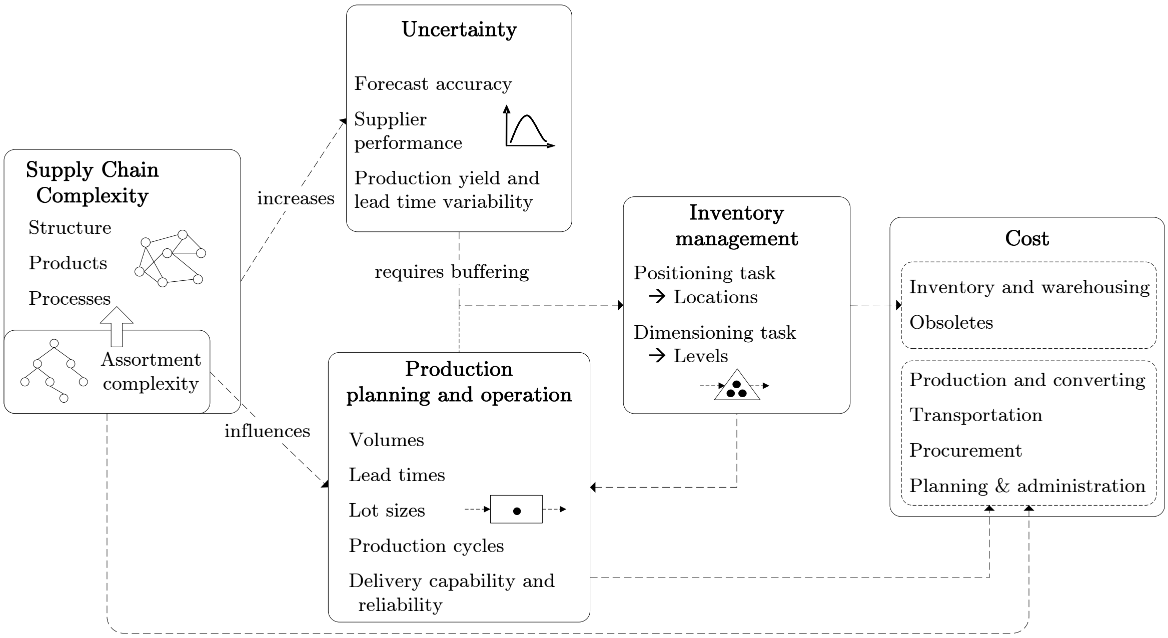


Figure 2.5: Assortment complexity influences supply chain cost via its effects on uncertainty, inventory management and production planning

Assortment complexity directly affects costs in the areas of procurement, planning and administration. External suppliers usually offer quantity-dependent price discounts. The more variants of an externally procured material exist, the smaller the procurement quantities and the less these price discounts can be exploited. Furthermore, all materials in an assortment have to be planned and managed, i.e. there is considerable effort in master data maintenance, administration and the periodic planning tasks. Planning tasks include demand planning, revision of inventory levels and planning of replenishment to ensure material availability.

40 For example, Schiller et al. (1996) report concrete measures taken by Procter & Gamble to reduce the existing assortment complexity and limit further product proliferation. Hoole (2005) propose several complexity reduction measures along the processes of the Supply chain operations reference model (SCOR) (Supply Chain Council, 2006) and reports on their application in practice.

41 Rathnow provides a graphical overview of complexity-incurred costs along the value processes 'research and development', 'procurement', 'production', 'sales' and 'customer service' (Rathnow, 1993, p. 24).

Besides these direct cost impacts, there are several levers that indirectly incur additional cost. Figure 2.5 shows how these costs are created via effects in the area of supply chain uncertainty and via changes in the area of production planning and operation as well as inventory management, which we discuss in the following sections.[42]

2.2.3.1 Effects on Uncertainty

The more complex an assortment and the corresponding supply chain become, the more possibilities for unforeseen events are introduced and the harder it gets to forecast the effects of the actions taken. Generally, this is called *uncertainty*: Van der Vorst defines supply chain uncertainty as *"decision-making situations in the supply chain in which the decision-maker lacks effective control actions or is unable to accurately predict the impact of possible control actions on system behaviour"*[43]. It is an inherent property of complex systems that the multitude of interrelations between the elements makes it hard to foresee the ultimate effects of certain actions.

In the context of production and logistic processes, different types of uncertainty can be distinguished: Uncertainty in

Demand Future customer requirements are generally unknown.

Supply Suppliers or supplying production and distribution stages do not always fulfil their deliveries as expected.

Processes No production or distribution process is deterministic in its outcome.

All these types of uncertainty become apparent as fluctuations in *quantity* and *time*. In order to alleviate the problems caused by uncertainty, companies have to install costly buffers in terms of material, capacity and time. An increasing complexity thus increases uncertainty and thereby costs via various levers.[44] Davis[45] states that inherent uncertainty is one of the major problems in planning and controlling supply chains and greatly affects their performance.

42 Similar categorisations of costs incurred by assortment complexity can be found in literature: Randall and Ulrich distinguish *production costs* and *market mediation costs*, where the former comprise all costs of *"materials, labour, manufacturing overhead, and process technology investments"*. The latter consist of *"inventory holding costs and product mark-down costs occurring when supply exceeds demand, and the costs of lost sales when demand exceeds supply"*, which occur due to *"presence of demand uncertainty"* (Randall and Ulrich, 2001, p. 1588).

43 Van der Vorst (2000, p. 74).

44 For a more detailed discussion of this causal relationships, see Wilding (1998).

45 Davis (1993).

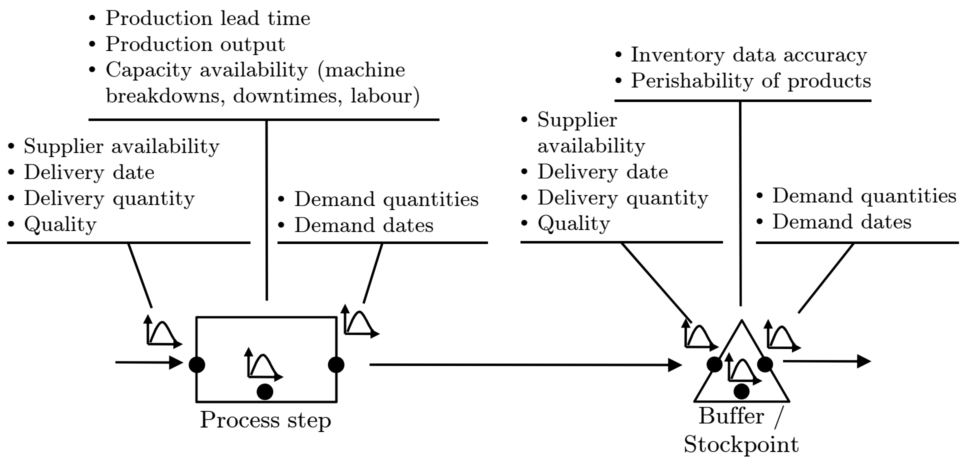


Figure 2.6: Types of uncertainty at demand, supply and process points

A more complex assortment and related supply chain complexity increase uncertainty in the above-mentioned dimensions of demand (forecast accuracy), supply (supplier performance) and internal processes (production yield and lead time variability). These impacts are either direct or via the increasing total supply chain complexity.

Sales forecasts for single variants and sales regions are generally less accurate than aggregate forecasts from different regions and for products with commonality. The average sales quantity for each product decreases and natural fluctuations in demand or deviations in forecasts cannot be compensated for by corresponding fluctuations in other market regions. A high assortment complexity impedes the exploitation of risk pooling effects[46] and thereby increases demand uncertainty, as all different variants are held on stock at different locations and forecasts are made on a per material and per location basis.

Process uncertainties become apparent in larger variations of production yield, transport times, information availability and accuracy. With an increasing number of materials produced, the learning effects and experience for the production of single materials decrease and create a larger variety in production yield and processing times. Considering the entire production process, structural features like the number of production steps also have an impact, as these effects accumulate over the various steps. In distribution stages, the uncertainty in transport times are linked to structural features like transport links and facility locations. With respect to information availability and accuracy, it may be expected that the error-proneness of a business information system increases with the number of materials and locations involved.

46 See Nahmias (1997, p. 61) and Zipkin (2000). A more detailed discussion of risk pooling is presented in Section 3.2.4.

Supply risks regarding supply availability and reliability are introduced via structural complexity with each additional supplier. Increasing assortment complexity in terms of materials procured externally also increases supply uncertainty since suppliers face greater challenges to comply with the required service levels and delivery times if the number of materials ordered increases.

2.2.3.2 Cost Effects in Inventory Management

Increasing assortment complexity affects inventory management by creating increasing inventory requirements. The determination of the total inventory requirements comprises a positioning and a dimensioning task, which are both affected by a changing assortment.

The positioning of stockpoints[47] in the supply chain changes due to the creation of new potential stockpoints with each additional material. Each amplification of the assortment with additional end products may require additional stockpoints to assure the timely delivery of these products to customers at the required service level. For each additional variant, stock of end products and possibly of components and packaging materials on upstream stages has to be built.

With respect to the dimensioning task, it must be noted that inventory requirements are mainly driven by replenishment lead times, demand quantities and demand uncertainty. As inventories in terms of safety stocks are one of the major means of buffering against the uncertainties mentioned above[48], we can expect an increase in the total amount of inventory distributed along the supply chain to provide the same service to customers, despite the increased levels of uncertainty in demand, supply and processes.

These changes in the required inventory levels directly affect costs. Firstly, costs are incurred in terms of opportunity costs for working capital bound in inventories as well as warehousing costs for the provision of physical storage space. Secondly, inventories always risk obsolescence due to seasonality or limited product lifecycles which then incurs costs of scrapping and write-offs.

47 We again refer to stockpoints at the individual material level, as defined in Section 2.1.1.
48 See Section 2.1.

2.2.3.3 Cost Effects in Production Execution

In the production area, we consider flow shop oriented batch production systems that do not produce customer individual variants with lot size 1, but have to plan production orders of reasonable lot sizes for each variant.[49] In this setting, changes in the assortment and thereby the set of materials produced on one production process step as well as the demand volumes of these materials heavily affect the costs incurred for their production.

A decrease in required production volumes per material is the logical consequence of increasing assortment complexity. A large number of materials with small demand volumes conflicts with the requirements for large production lots and avoidance of setup and scrap cost. Due to a larger number of materials processed, lead times on the corresponding production process step become larger and / or production cycles[50] increase. Smaller production volumes tend to result in smaller average lot sizes. This leads to more frequent changeovers in the production sequence, each of which incurs cost for setup operations in terms of labour and opportunity costs for the unused production capacity during the changeover. The decrease in lot sizes may be alleviated by an increase in production cycles, which means that demands of longer time periods have to be aggregated to form larger production lots. However, this results in increasing replenishment lead times of successor stages, as the agreed service time for a certain material depends on the maximum time it may take to schedule a production order of that material. Figure 2.7 illustrates the interdependency of lot sizes, replenishment lead times and production cycles.

The set of materials processed on a certain production process step influences the planning buffer required to enable production planners to generate production sequences with reasonable setup costs. The larger the number of production orders to be scheduled, the more difficult it becomes to create optimal production sequences and longer planning buffers are required to allow the generation of cost-efficient production sequences.

These production planning and operation decisions are also interrelated with the area of inventory management in a variety of ways, as the decision which materials to produce to stock influences production parameters such as production cycles. Furthermore, the lead times that result from the selection of planning buffers influence the amount of

49 See Section 2.1.2

50 For this chain of reasoning, we use the concept of fixed production cycles per material. Even if such production cycles may not be used in practice, the rationale remains valid for the average time between two production orders of the same material.

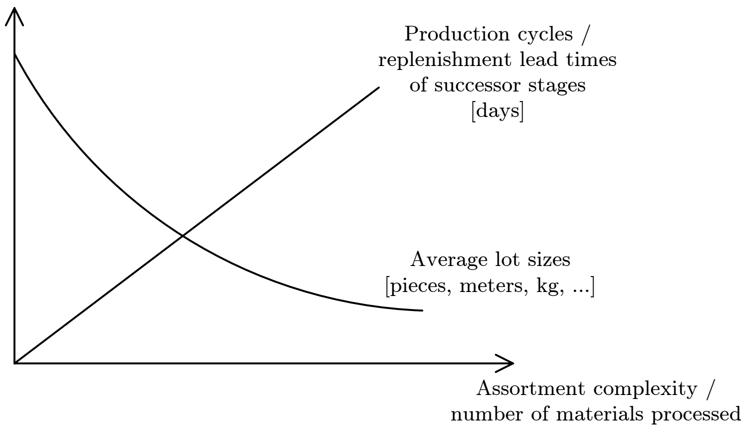


Figure 2.7: Interdependency of lot sizes, replenishment lead times and production cycles

inventory required at different locations and can thereby change the optimal inventory allocation.

2.3 Assessing the Cost of Assortment Complexity in Production and Distribution Networks

This work seeks to evaluate the most relevant assortment-dependent cost positions and therefore focuses on the areas of inventory management and production execution. Accordingly, we also use the term production and distribution network (PDN) instead of supply chain, as this implies that only production and distribution processes and their related costs are the subject-matter of this analysis. This focus is selected for the following reasons: firstly, cost effects particularly are expected as a result of changing inventory requirements and related inventory holding costs as well as changing setup times and scrap quantities in the production area. Secondly, the effects of assortment complexity on these costs are particularly difficult to assess by conventional cost-accounting methods and therefore often remain disregarded when it comes to assortment-related decisions. Quelch and Kenny[51] find that the cost factors not considered correctly and / or sufficiently when taking assortment-related decisions comprise:

- increased production complexity resulting from shorter production runs and more frequent line changeovers

51 Quelch and Kenny (1994, p. 156).

- more errors in forecasting demand
- increased logistics complexity, resulting in increased remnants and larger buffer inventories to avoid stockouts

The problem of assessing and foreseeing these costs correctly results from the interdependencies between single variants. These interdependencies make it impossible to have a single cost value per material reflecting the exact cost for its production and distribution and thus the savings that can be expected if this single material is discontinued.[52] To underline this point, the following list gives examples where a simple summation of the costs assigned to each individual discontinued material does not represent the real savings potential of e.g. an assortment reduction.

- If end product packaging is standardised, e.g. by introducing multilingual packaging, a central warehouse with finished products, e.g. at the production site, may become preferable to decentralised stocks at local sales warehouses.
- If the set of semi-finished materials for one category of end products can be reduced, it may become preferable to keep inventories of semi-finished products instead of raw materials. Furthermore, it may be possible to reduce the planning buffer of the corresponding production process step without significant increases in the average setup costs due to less complex production sequencing.
- If the set of raw and / or packaging materials is standardised to a certain degree, it may become preferable to keep inventories of these materials to reduce the total replenishment lead times for the end products.

As the cost effects depend on such concrete assortment changes, they cannot be reflected by traditional cost-accounting systems, even if they allocate the costs fairly according to the inputs involved. In order to provide a reliable decision support for assortment-related decisions, this work also answers the question

How do costs in inventory management and production execution change in response to assortment changes?

With this guiding research question, the methodic approach of this work is mainly *quantitative*. In order to allow the type of *what-if* analysis implied by this question, a method of assessment of costs incurred in a production and distribution network by a

52 Quelch and Kenny also point out that *"the costs of line-extension proliferation remain hidden"*, as *"traditional cost-accounting systems allocate overheads to items in proportion to their sales. These systems [...] overburden the high sellers and undercharge the slow movers"* (ibid., p. 156).

certain assortment is required. Given such a method, arbitrary possible assortments can be assessed and compared. Figure 2.8 shows how the task of carrying out such a cost assessment is broken down according to the three cost areas to be assessed.

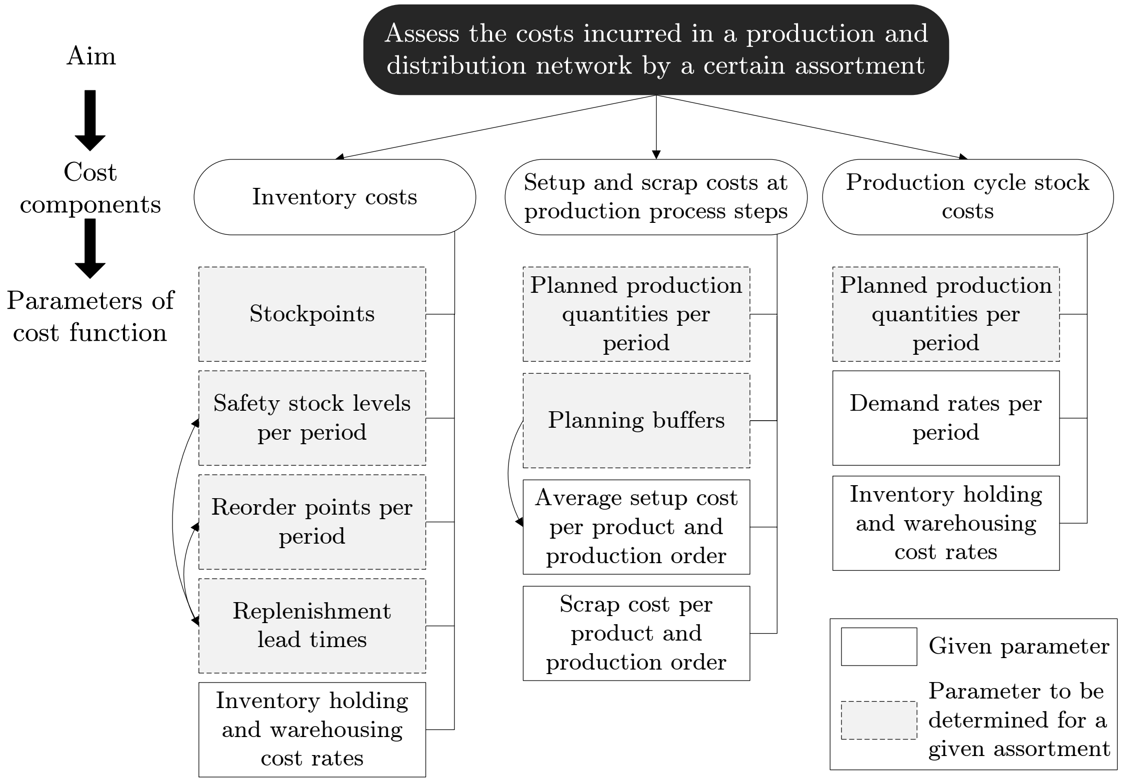


Figure 2.8: Cost model: components and parameters

In accordance with the focus on inventory management and production execution, the cost areas can be summarised as inventory cost, setup and scrap cost, and cycle stock costs. For each such cost component, the figure lists the factors that determine the individual cost positions.

For the assessment of inventory costs, we can distinguish between cost incurred by general inventories and costs incurred by production cycle stocks. For the former, both the positioning of stockpoints in the network and the required inventory levels at each such stockpoint have to be known. For the latter, the position of the reorder points and the corresponding replenishment lead times have to be known to estimate the average cycle stock levels. These factors are partly interdependent, as the replenishment lead times influence the values of both the reorder point as well as the required inventory levels for the considered material. At the same time the chosen safety stock level influences the replenishment lead times of successor materials, which is why this is a bidirectional dependency. Given that these factors are known, inventory costs can be easily calculated with the known inventory holding and warehousing cost rates.

For the assessment of setup and scrap costs, planned production quantities per period have to be known for each material that is produced internally in order to derive the number of production runs per period. With the number of production runs, the corresponding setup and scrap costs can be estimated via the known cost parameters for scrap and average setup costs. As described in Section 2.1.2, the average setup cost parameter also depends on the choice of the planning buffer, which makes the decision variables interdependent.

For the assessment of production cycle stock costs, the difference of planned production quantities and actual demand quantities has to be calculated for each period. Thus relevant factors comprise planned production quantities, demand rates and the inventory holding and warehousing cost rates already used for the assessment of general inventory costs.

These factors can be subdivided into two groups: firstly, there are fixed parameters in terms of cost rates or externally given demands, represented by white boxes in Figure 2.8. These parameters either remain constant across assortments or change in a known way. While cost parameters generally remain constant, changes of demand quantities may be defined on the basis of expected customer behaviour.

Secondly, there are several variable factors that have to be adapted and determined for each individual assortment, represented by grey shaded boxes in Figure 2.8. These variables are the means of taking into account the various ways in which the complexity of the given assortment influences the costs mentioned above. As the aim is to provide a method to assess these costs for an arbitrary assortment, we cannot rely on actual values from the existing production and distribution system: an optimisation method that determines these variables for any given production and distribution network is required.

By providing such optimisation method that is able to determine these variables for arbitrary assortments, this work not only provides means of analysing the cost effects, but also identifies the effects on the production and distribution processes by answering the question

What changes in the configuration of production and inventory management parameters may be expected in response to the assortment changes?

The fact that there are several variable factors in the cost function requires a multi-stage analysis process. Figure 2.9 summarises the input, steps of the analysis process and output of the the method developed.

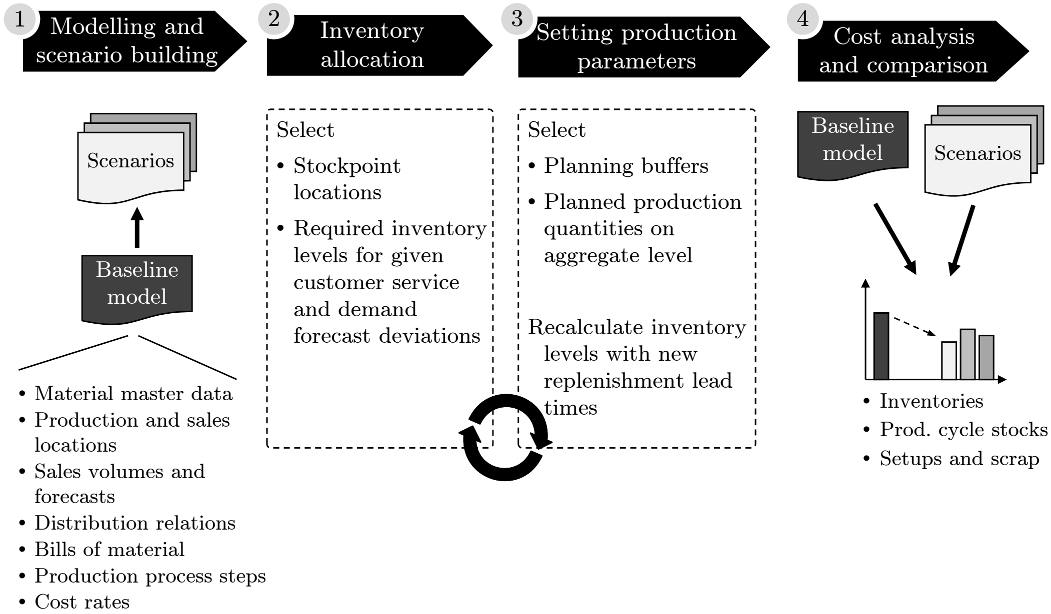


Figure 2.9: Input, output and steps of the analysis process

In the first step, a model[53] of the system[54] under consideration has to be built. This model has to represent the production and distribution network of a particular assortment with the entire product structure. From this baseline model, which is most likely to represent the current as-is situation, scenarios representing theoretical alternative scenarios are derived. They differ from the baseline models in terms of the set of materials in the assortment, the product structures, the distribution relations and / or assumptions about the demand distribution and demand uncertainty. Both the baseline models and all scenarios can be represented with the same formal model.

53 *"A model is an abstract description of the real world giving an approximate representation of more complex functions of physical systems"* (Papalambros and Wilde, 2000, p. 4). A similar definition is also given by Buhr and Klaus (1975, p. 805).

54 A *system* are *"groups of interacting, interrelated, or interdependent elements forming a complex whole"* (Pickett, 2000). Common characteristics of a system are to

1. consists of interacting components
2. be hierarchically structured
3. be associated with a function it is intended to perform.

From the fact that there are numerous types of systems that each require a specific definition, Cassandras and Lafortune (2007, p. 2) conclude that *"'system' is one of those primitive concepts [...] whose understanding might be left to intuition rather than an exact definition"*.

In the second step, the inventory allocation in both the baseline model and all relevant scenarios has to be determined. We define an *inventory allocation* as the positioning of stockpoints and the determination of inventory levels and corresponding replenishment lead times for each material.

In the third step, planning buffers and planned production quantities have to be set. As planning buffers are also part of the throughput times of each production process step, the second and third step are interrelated: after the throughput times have changed due to the adapted planning buffers, replenishment lead times may change for a set of affected materials and the required inventory levels and reorder points have to be adapted. Accordingly, steps 3 and 4 may either have to be solved in an integrated way or may have to be solved repeatedly.

In the fourth and final step the required cost analysis can be performed on the basis of the adapted configurations of the production and distribution networks represented by the baseline model and the scenarios. This step enables the assessment of the effects of assortment complexity by a pairwise comparison of the assortments and cost changes of the baseline model and the various scenarios.

Complex problems are solved by decomposing them into a set of smaller subproblems that have a clear relationship to each other.[55] The structuring of the cost assessment and analysis process allows us to identify three major subproblems to be solved in the context of this work:

1. Modelling of assortments, assortment scenarios and production and distribution related cost
2. Inventory allocation in production and distribution networks
3. Determination of planning buffers and planned production quantities

The following sections discuss these subproblems by describing the related tasks to be addressed and solution requirements for each of them.

55 See Pärli (1980).

2.3.1 A Model to Assess the Cost Effects of Assortment Complexity Changes

This subproblem may be summarised by the question

How can the production and distribution structure of a certain assortment be represented by a formal model that serves as a basis for what-if analyses?

The specific tasks that result from this question may be summarised as follows:

1. to provide a model that serves as a basis for the analysis of assortment complexity-related cost
2. to provide a formalism to define assortment scenarios on that model
3. to enable the assessment of setup, scrap, inventory and cycle stock costs on the basis of any given model instance with all parameters set

The solution developed to address these tasks must fulfil the following requirements:

Model completeness for the aspired analysis The model has to serve as a basis for the entire analysis process. Its level of abstraction must be defined to explicitly represent all the information required for the analysis and optimisation steps. In particular, this information comprises

- the assortment structure with all materials, their where-used relations and distribution relations between various locations,
- the production process steps required to produce the assortment under consideration, i.e. a mapping of single materials to production resources together with their resource requirements,
- the characteristics of internal and external processes: external supply, external customer demand, production and transport processes,
- demand and demand uncertainty data over a certain period of time and at an aggregate level. As it should be possible to define scenarios that also change the distribution of demands over the end products and materials, this information should be defined as a demand distribution over aggregate time periods. A stochastic model for the uncertainty inherent in the demand forecasts is also required and must be available at all elements of the PDN.

At the same time, the model must remain manageable in its complexity.[56]

Manageable effort for model building Models of real-world assortments easily become complex themselves. In order to be able to create such models for practical use, this process must be automated as fully as possible. This seems especially reasonable as almost all companies can be expected to use some IT based enterprise resource planning (ERP) system that already contains most of the required data. It must be possible automatically to create network models from this information based on the definition of a set of end products.

Configurable definition of consistent scenarios The definition of alternative assortment scenarios must allow addition of new materials as well as discontinuation and / or replacement of existing materials both at end product level and on the level of raw and semi-finished materials. It must be possible to also include the expected effects of the assortment changes on demand volumes and distributions. The application of a scenario definition to an existing baseline model must ensure that the resulting scenario is consistent in terms of structure as well as demand information.

Assessment of relevant cost areas The cost assessment must consider all relevant areas defined in Section 2.3 and allow their assessment for a concrete assortment model or scenario.

An approach to attaining the aims stated above may be based on a mathematical model[57]. Existing approaches to representations of assortment complexity are analysed in Section 3.1. Once such a model is defined, algorithms can be developed to derive such a model from standard data available in ERP systems such as bills of materials (BOMs), routings and procurement information. Based on such a model, a formalism for definition of assortment scenarios and a cost function that includes all relevant cost areas can be defined.

56 This trade-off between the level of detail and model complexity holds for all types of model and is discussed e.g. by Dangelmaier (2003, p. 41).

57 A mathematical model is *"a model that represents a system by mathematical relations."* (Papalambros and Wilde, 2000).

2.3.2 Inventory Allocation in Production and Distribution Networks

This subproblem may be summarised by the question

What is an optimal inventory allocation in a production and distribution network of a certain assortment?

The term *inventory allocation* here refers to both the inventory positions in terms of stockpoints in a production and distribution network and the corresponding inventory levels. The answer to the question above thus provides the values for the variables required to assess inventory costs as defined in Section 2.3.

The solution developed to find optimal inventory allocations must fulfil the following requirements:

Make decisions on the detailed level of individual materials Inventory allocation problems can be considered at various levels of abstraction. Nodes and potential stockpoints can represent physical locations, products at different levels of aggregation, or combinations of these.[58] Given that the inventory allocation is used to compare assortments that differ in the set of materials they contain, it has to operate on the very detailed level of individual materials at different locations.

Work on variable demand data on aggregate level Demand for consumer goods cannot be assumed to follow a certain known probability distribution, as factors like seasonality, promotions and generally short product lifecycles result in extremely volatile demand. The method therefore must be able to make inventory allocation decisions on the basis of expected demand quantities that vary over time and are known for a set of aggregate discrete time periods.[59]

Usage and integration of domain knowledge Given that the model on which the inventory allocation decisions are made cannot contain every detail about the underlying production and distribution system, it must be possible to integrate further domain knowledge into the allocation decisions to obtain solutions that are feasible in practice. This might be necessary to adequately represent situations in which certain materials cannot be put on stock due to their physical characteristics or the characteristics of the underlying production process.

58 See Zipkin (2000, p. 107).
59 See Section 2.3.1

Computational efficiency for realistic network sizes The combinatorial complexity of the optimisation problem considered requires the application of computational optimisation methods. As this problem has to be solved for each alternative assortment scenario that is to be analysed, the optimisation procedure must be able to find solutions even for large assortments in acceptable time. Considering the combinatorial complexity for large networks and the fact that the result of the optimisation serves as the basis for a cost analysis in the first place, the guaranteed optimality of the solution is not a primary requirement.

Figure 2.10 illustrates how the inventory allocation problem can be seen as the selection of stockpoint nodes in a network together with the determination of the corresponding inventory levels and delivery times between adjacent stages.

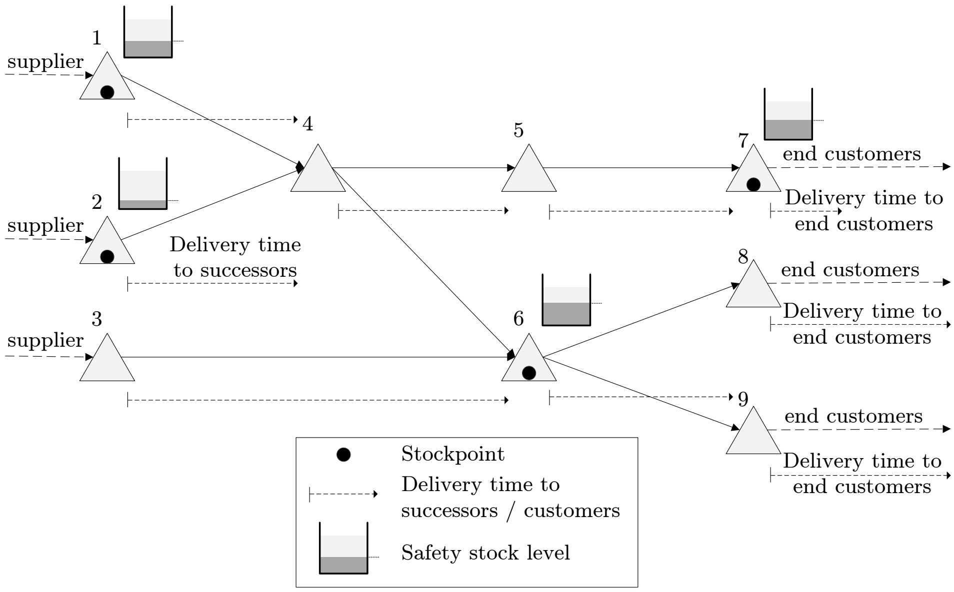


Figure 2.10: Inventory allocation problem

On this basis, an optimisation model is defined to determine the cost-minimal inventory allocation. Available approaches to similar problems are evaluated and analysed in Section 3.2 with respect to the questions how their models and assumptions match this application scenario. A solution method is developed based on existing techniques for computational optimisation. Heuristic techniques may have to be considered in this context due to the complexity of the optimisation problems.

2.3.3 Determination of Planning Buffers and Planned Production Quantities

This subproblem may be summarised by the question

What are optimal planning buffers and planned production quantities such that total setup, scrap and inventory costs are minimal?

The trade-offs to be solved in this context are illustrated in Figure 2.11.

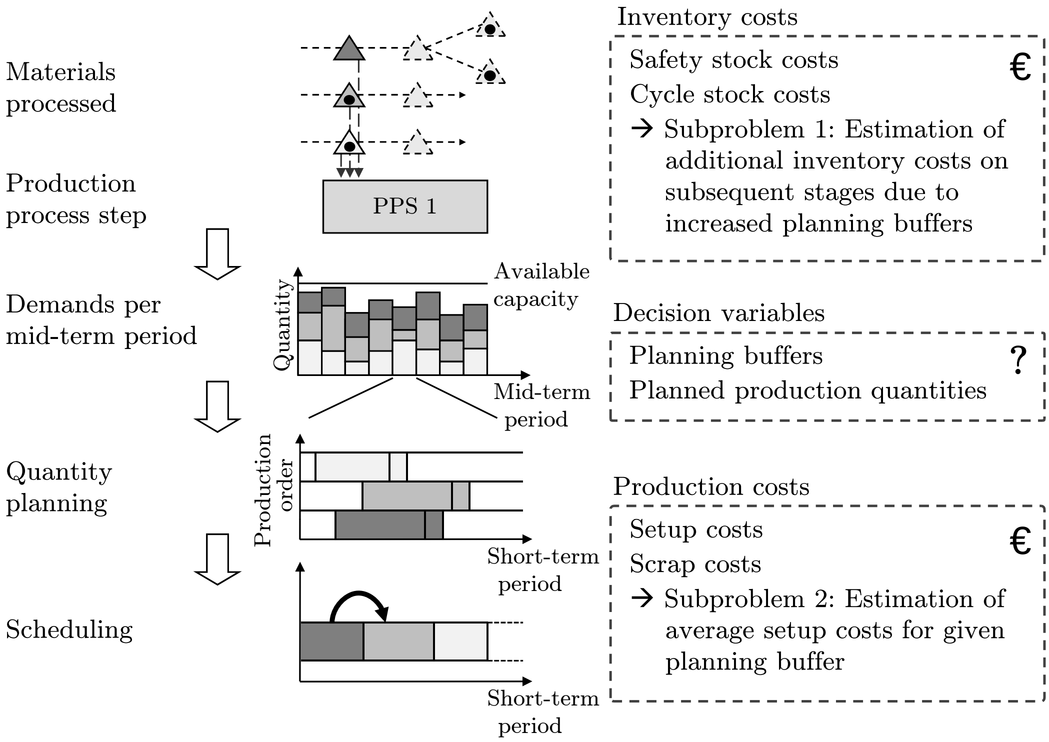


Figure 2.11: Rationale for the determination of planning buffers and planned production quantities

For each production process step, the set of materials processed and their corresponding demands are known, giving a set of demands for each period that have to be covered by corresponding production quantities. In a real-world setting, concrete production orders are the result of a quantity planning and scheduling process that takes into account the requirement dates of the orders from which the demands originate. The planning buffers define the set of potential production start dates for the orders and thereby the optimisation potential for the sequencing task. The actual production cost that should be assessed is incurred when such a detailed production plan is carried out and the setup and scrap cost that depend on the lot sizes and the sequence of the production orders are actually known.

One problem is that we cannot expect to have sufficiently detailed demand information to derive concrete requirement dates. As outlined in Section 2.3.1, demand information for the assortment models and especially the theoretical alternative scenarios is only available at an aggregate level. The assessment of operational cost factors would actually require detailed production planning, which is not possible with the information available. An appropriate handling of this situation will be recorded as a requirement for the solution to this subproblem.

Both the planning buffers and planned production quantities that should be determined have cost effects which conflict. This creates two trade-offs to be solved: firstly, increasing the length of the planning buffer reduces the average setup costs incurred in the production execution. At the same time, safety stock requirements for materials on the same or subsequent production step increase due to increased replenishment lead times created by the planning buffers. Secondly, large planned production quantities lead to bigger lot sizes and consequently fewer changeovers that incur setup and scrap cost. At the same time, the large planned production quantities increase the average cycle stocks as the produced quantities are only consumed successively.

A solution to the problem of finding the optimal trade-offs in this setting must:

Assess operational production cost from demand data on aggregate level As described above, the solution has to assess operational production cost without any exact knowledge of the detailed production plans.

Consider interrelation of setup, scrap and cycle stock costs The determination of planned production quantities has to consider the cost trade-off between setup, scrap and cycle stock cost and find solutions where these total costs are minimal.

Consider effects of planning buffers on inventory requirements The determination of suitable planning buffers has to consider that larger planning buffers also increase inventory requirements. The planning buffers must be chosen such that the total costs are minimal.

Computational efficiency for realistic model sizes Although number of potential configuration increases rapidly with the number of materials, time periods considered and potential planning buffers, the solution method must be able to determine suitable variable values in acceptable time even for the large model instances expected in practical applications.

Consider the interrelation to the inventory allocation subproblem Figure 2.12 shows how the optimisations of the different variables relate to each other. As indicated

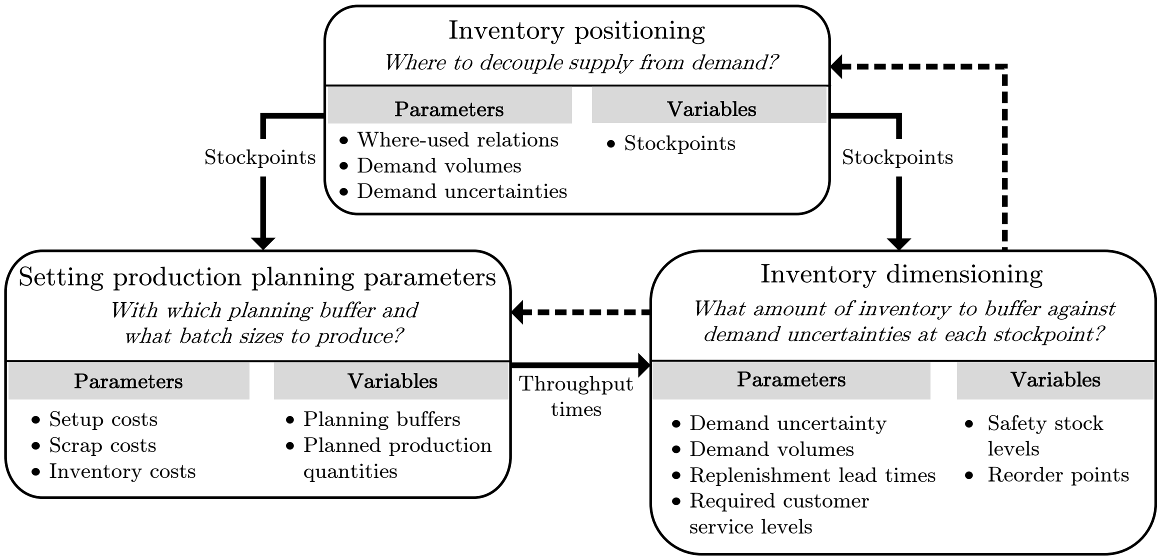


Figure 2.12: Relations between subproblems

by the solid arrows, the results of the inventory placement are input parameters for both the optimisation of planning buffers and for the dimensioning of inventory levels. In the former case, the positions of the stockpoints influence the effect of planning buffers on inventory costs. In the latter case, stockpoints have to be determined before the required inventory levels can be calculated. As the determination of planning buffers also increases the throughput times of production stages and thereby the replenishment lead times of successor stages, these results are also a required input for the dimensioning of inventory levels. In addition, the results of the dimensioning of inventories are a required input to evaluate the solutions of the other subproblems. A solution to this dependency problem is another requirement for the overall optimisation procedure.

To find an optimal configuration for the variables described, optimisation methods from mixed integer programming can be used and thus a mathematical optimisation model is formulated. In particular, the determination of planned production quantities shows parallels to big-bucket lot-sizing models. In order to comply with the requirement of efficient solvability for large model instances, the model is defined so that it can be solved with standard optimisation software. This implies that the model should be formulated as a linear optimisation problem and that linearisation is applied where necessary and applicable. In order to handle the aggregate demand data, the relations between some of the cost factors may have to be estimated where applicable.

2.4 Structure of this Work

Figure 2.13 summarises how the structure of the remainder of this work corresponds to the subproblems described above. Existing works from the relevant fields mentioned

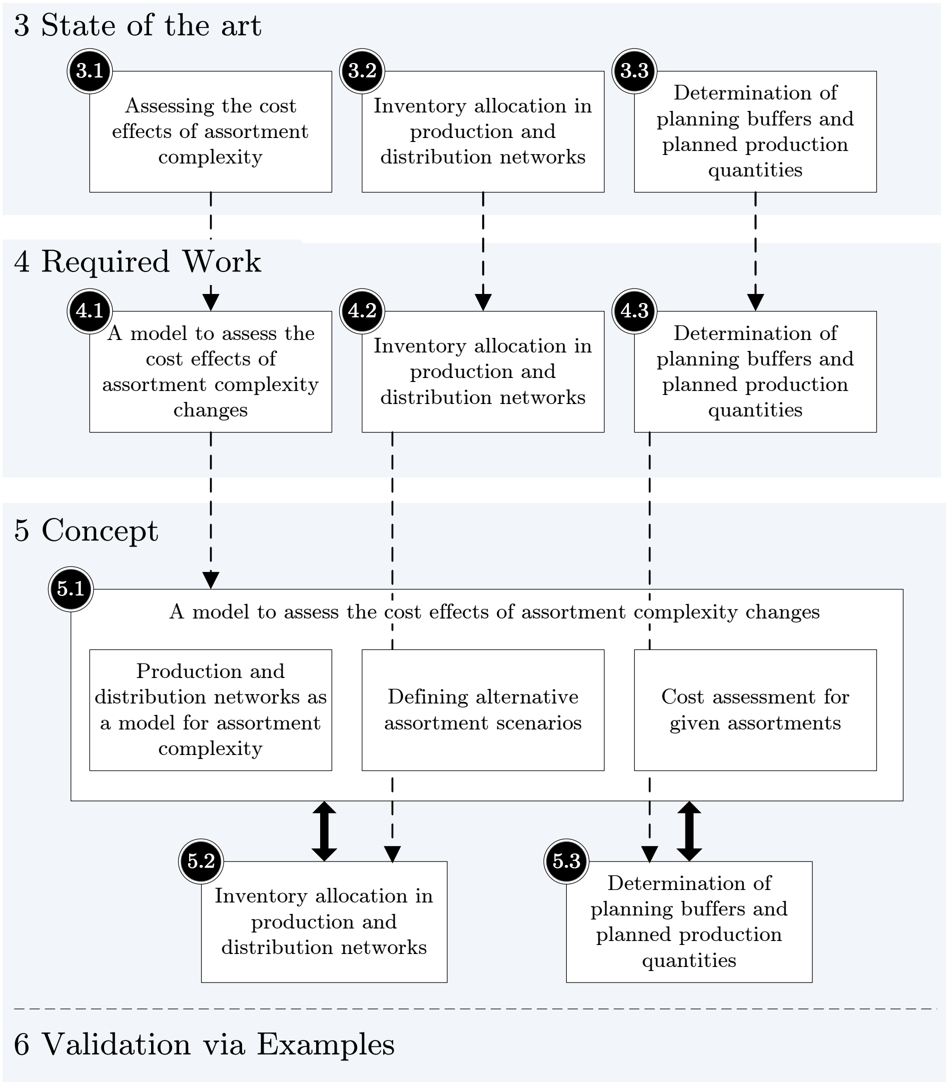


Figure 2.13: Structure of this work

in Sections 2.3.1 to 2.3.3 are evaluated with respect to their applicability in the corresponding sections of Chapter 3. Upon these findings, Chapter 4 summarises the remaining tasks to be addressed. Chapter 5 then develops the individual solutions in the structure depicted in Figure 2.13. The validation in Chapter 6 shows how the software implementation of the methods developed can be used to apply these methods to real-world problems and shows both the applicability and validity of the solution.

CHAPTER 3

State of the Art

All in all you're just another
brick in the wall.

Pink Floyd

This chapter summarises the state of the art in all areas relevant to this work. Section 3.1 first analyses existing work on the assessment of the cost effects of assortment complexity in order to see if the intended *what-if* analysis is a viable approach. Section 3.2 then focuses on work on optimisation methods for the inventory allocation problem. Finally, Section 3.3 provides a literature review on the optimisation of selected production planning parameters that are similar to the planning buffers and planned production quantities considered in this work.

3.1 Assessing the Cost Effects of Assortment Complexity

Since assortment complexity is omnipresent in virtually all manufacturing companies and due to its far reaching implications, its management has gained considerable attention in both literature and management practice. Against the background of this work, this section first reviews formal models of assortment complexity and then evaluates existing work on the assessment of complexity-related costs.

3.1.1 Models of Assortment Complexity

Different models of assortment complexity can be grouped according to their view on the assortment. Figure 3.1 illustrates these different views and dimensions of assortment complexity. The first view on assortment complexity is to consider the end

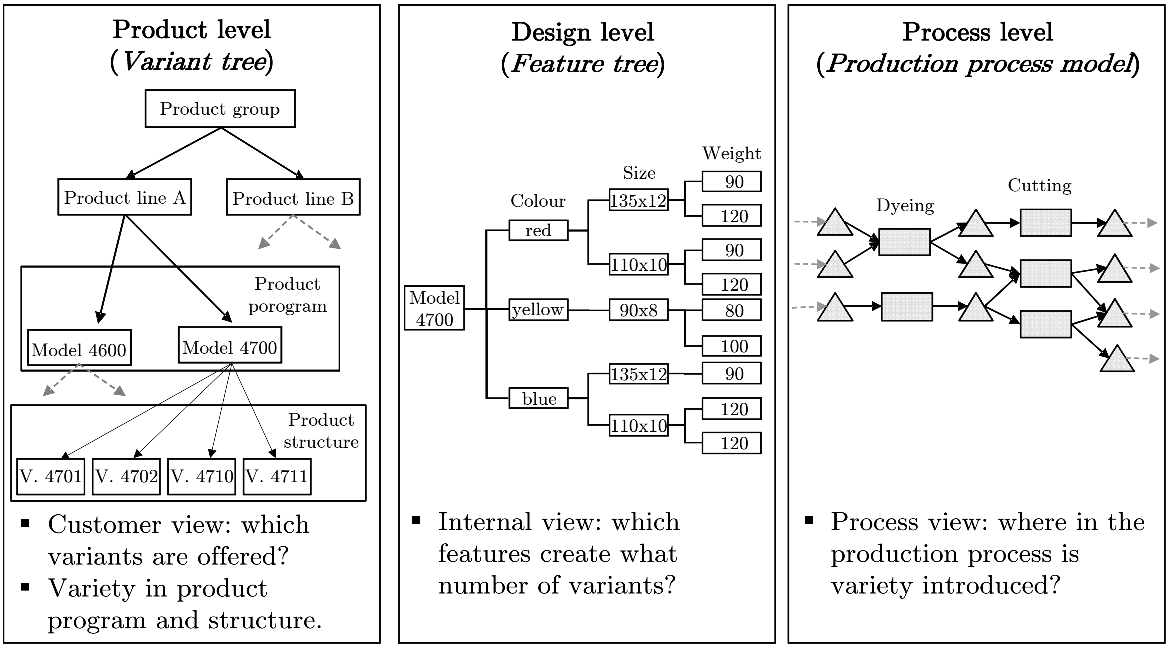


Figure 3.1: Views and dimensions of assortment complexity

product level with all product groups, product lines, the product program for an individual product line and finally the individual variants. This can be seen as the 'customer view', as the individual variants of finished products represent the variety as it is visible to the customer. The design level describes the individual product features and possible combinations of these features. The feature tree helps to determine which features create what number of finished product variants and is the most common way to represent assortment complexity.[1] The production process view combines these features with the actual production processes required to manufacture a given set of product variants. Depending on the organisation of these production processes, variants may emerge at different production stages. For a given product structure, there are many possibilities to organise the production process so that variety may be introduced at different steps.

One particular class of formal models of assortment complexity are market oriented models that describe the variety of products offered by a company within a certain market. The aim of these models is to analyse the suitability of an assortment in a given market setting. The two basic models that can be distinguished in this context are those given by Chamberlin[2] with its formalisation by Dixit and Stiglitz[3] and the

1 See e.g. Moore et al. (1986).
2 Chamberlin (1933).
3 Dixit and Stiglitz (1977).

model by Hotelling[4] with its adaptations by Lancaster[5]. The Chamberlin-Dixit model analyses how product differentiation may lead to markets with monopolistic competition, as customers do not perceive the products of different brands as substitutes due do their differentiating characteristics. The Hotelling-Lancaster model describes each product via its characteristics in a virtual space of product characteristics. This way similarities and substitution effects between products can be represented. Customer preferences in terms of desired product characteristics can be modelled analogously by defining points in that space and measuring their distance to existing products. Chen et al.[6] build on this model to address product line design and product pricing problems. For a more detailed description of these models, the reader is referred to the work of Lancaster[7]. Although these models have been used and enhanced extensively, they do not provide a suitable approach in the context of this work since they do not contain any information about the cost of the product variety, but take a market oriented view to describe suitable assortments given assumptions about customer preferences.

In their analysis of the effects of brand width on brand market share, Chong et al.[8] use a simple *product tree* to represent the assortment of a company. This tree roughly corresponds to the feature tree depicted in Figure 3.1. On the basis of this tree, they derive different measures for brand width, comprising the number of SKUs, the number of feature levels and the number of distinct product variants. While this model is intuitive and sufficient for their analysis, it does not contain sufficient information about the involved production and distribution processes to be used for the purpose of this work.

Schuh[9] proposes the *variant tree* as an extension of the product tree to include information about the assembly sequences of car variants. This variant tree also serves as the basis for an accounting method developed by the same author to assess complexity-related cost, as described in Section 3.1.2.3. While this is a first approach to include information about the production processes into a model of assortment complexity, it lacks a representation of the required production facilities, subsequent distribution structures and is especially tailored to the automotive industry.

4 Hotelling (1929).
5 Lancaster (1979).
6 Chen et al. (1998).
7 Lancaster (1998).
8 Chong et al. (1998).
9 Schuh (2005, pp. 158-159).

3.1.2 Quantitative Assessment of Complexity-Related Cost

Several reasons have led to the large number of works that deal with the effects of assortment complexity. Firstly, the topic has already been discussed since the early 19th century.[10] Secondly, research in this area has gained increasing attention during the last decades due to an increasing need for complexity management methods.[11] As this chapter cannot give an exhaustive literature review, we classify relevant works into three categories and describe some selected examples for each of these categories in the following sections. These exemplary works have been chosen to match the problem addressed in this work as closely as possible.

There exist several more comprehensive reviews on this topic. For example, Ramdas[12] provides a *"framework for managerial decisions about variety"* to structure the large amount of literature available. This framework distinguishes between four decision themes in variety *creation* (dimensions of variety, product architecture, degree of customisation and timing) and three decision themes in variety *implementation* (process and organisational capabilities, points of variation and day-to-day decisions). Further comprehensive literature surveys are provided by Lancaster[13], Herrmann and Peine[14], Bräutigam[15] and Heina[16]. The current state of the art is also summed up in the edited volumes by Adam[17] and by Ho and Tang[18], who provide collections of relevant papers with a focus on controlling and operations research methods, respectively.

3.1.2.1 Descriptive and Empirical Works

The majority of relevant publications do not provide quantitative models to directly assess the costs of assortment complexity. Moreover, the effects of an increasing number of product variants are described on an abstract level and / or analysed in empirical studies. This section briefly describes some selected works from this category.

10 See e.g. Wolter (1937).
11 See the explanations in Section 2.2.
12 Ramdas (2003).
13 Lancaster (1990).
14 Herrmann and Peine (2007).
15 Bräutigam (2004).
16 Heina (1999).
17 Adam (1998).
18 Ho and Tang (1998).

For further examples, the reader is referred to the works of Marti[19], Rathnow[20] and Kirchhof[21].

Kestel[22] describes the effects of increasing assortment complexity in logistic systems in a detailed analysis of its implications on warehousing, picking, transport and information flow processes. Furthermore, he considers the impact on several functional areas like production and procurement. While all considerations are made on a detailed level, no quantitative measures for the relationship between the number of variants in the logistic system and its performance are provided.

Davila and Wouters[23] provide a model to assess the effects of product standardisation and show its application in a case study. The performance model comprises key indicators for the areas of inventory reduction, service improvement and quality improvements. This model is used in a case study with a major hard disk drive manufacturer. Despite the relevance of the performance indicators used, it has to be noted that the entire model is designed to assess the effects of standardisation based on information from the post-standardisation phase. It is thus unsuitable to predict the effects of potential standardisation as the basis for assortment-related decisions.

Prillmann[24] discusses assortment complexity as a strategic managerial task and describes the levers that determine both assortment and total supply chain complexity. He formulates several hypotheses to describe the success factors for effective complexity management. These hypotheses are validated with several empirical studies from the electronics industry to show that companies which consider these success factors appropriately are generally more successful.

Randall and Ulrich[25] describe the effects of assortment complexity on supply chain performance via a case study in the U.S. bicycle industry. They distinguish production costs and market mediation costs as the main performance metrics. Production costs also include the fixed investments associated with providing additional product variants. Market mediation costs are incurred by all measures required to match supply and demand despite the presence of demand uncertainty. These costs include the variety-related inventory holding cost, obsolescence cost as well as lost sales. The

19 Marti (2007).
20 Rathnow (1993).
21 Kirchhof (2002).
22 Kestel (1995).
23 Davila and Wouters (2003).
24 Prillmann (1996).
25 Randall and Ulrich (2001), Ulrich et al. (1998).

particular interest of the authors is to investigate how companies can match their assortment complexity with the structure of their supply chain to reduce these costs.

Many authors conjecture that complexity-induced costs increase progressively with the degree of assortment complexity. This progressive cost behaviour leads into the well-known *complexity trap* described by Adam and Johannwille[26]. Likewise, Klaus[27] claims that there is a certain degree of complexity beyond which an *over-complexity* phenomenon arises and complexity-related costs even increase exponentially. However, the analytical considerations of Bräutigam[28] suggest that this progressive cost increase does not always hold and that degressive cost functions are possible in certain settings as well. This disagreement about the *general* nature of the cost function shows that it is almost impossible to define a generic, monovariable cost function for the effects of assortment complexity, an argument also supported by the review of quantitative approaches in the following section.

3.1.2.2 Cost Estimation with Simplifying Assumptions

Adam[29] points out that complexity-related costs have some unpleasant characteristics that impede a precise assessment. Firstly, they are overhead costs and thus cannot be assigned precisely to the causing product variant. Within certain intervals of numbers of variants, complexity-induced costs can be fixed and increase only when the provided management and coordination capacities do not suffice any more.[30] This way, slow-moving items are generally subsidised by the standard products with high volumes. Secondly, complexity consists of many interdependent factors and drivers[31], whose interrelations have to be known in order to measure the effects of a change. To solve this cost assignment problem, some authors propose a monovariable approach and only consider the *number* of product variants, assuming the ceteris paribus assumption for all other factors. This may lead to problems, as this assumption generally does not hold in practice.[32]

Bräutigam[33] analyses the cost behaviour in production systems with a high degree of product variety. He argues that the *number* of product variants offered is a strate-

26 Adam and Johannwille (1998).
27 Klaus (2005).
28 Bräutigam (2004).
29 Adam (2004).
30 See ibid., pp. 22-23.
31 See Section 2.2.
32 Adam (2004, p. 21).
33 Bräutigam (2004).

gic decision parameter of uttermost importance for all manufacturing companies and therefore pursues the aim to derive a generic production cost function that depends only on this single parameter. The main determinants of this cost function are setup costs caused by changeovers between production runs of different variants. In order to estimate the number of changeovers required for a certain set of product variants, the author requires that the occurrences of demand for the variants are classified in the dimensions *time* (uniform, random, accumulated) and *quantity* (uniform, accumulated).[34] Given a certain number of variants and classifications of their demand according to this scheme, he then derives expected production sequences and corresponding estimator functions for the number of changeovers.

Although such a generic estimator function would provide a simple way to support assortment-related decisions, its simplicity can only be obtained at the expense of far reaching assumptions about all other influencing factors. Since the estimator function is monovariable, i.e. only depends on the number of variants, it requires that all other influencing factors have to be assumed constant and known a priori.[35] In the work of Bräutigam, these assumptions include

Identical value of all variants If only the *number* of variants to add or discontinue is considered instead of a concrete set of variants, statements about cost changes can only be interpreted if all variants are assumed to have the same value.

Fixed production sequences In order to estimate the number of changeovers required to produce a certain number of variants, fixed and cyclic production sequences have to be assumed, based on the demand characteristics of the variants.

Identical setup costs The total setup costs incurred are only estimated via the expected number of changeovers caused by a given number of variants, which in turn implies that all changeovers are assumed to incur identical setup costs.

These assumptions can be questioned and they probably bias the results of applications of this method in practice.

Thonemann and Bradley[36] provide a model to assess the effects of product variety on supply chain performance in a setting with a single manufacturer and multiple retailers. Supply chain performance is measured in terms of setup costs and inventory costs at the retail stores. The authors consider a setting where setup *times* are important and focus on the effects of increased lead times due to more frequent changeovers caused by

34 See ibid., pp. 101-104.
35 See ibid., pp. 96-99.
36 Thonemann and Bradley (2002).

increasing variety. They conclude that the lead time effect of product variety on cost is *"substantially greater than that suggested by the risk-pooling literature"*.[37] While the aim of their work coincides in many aspects with the aim of this work, they obtain their general results on the expense of assumptions about the production system and inventory policies. Demands are all assumed to be identically Poisson distributed. Production planning is done according to a simple batching strategy where production of a certain good is only initiated when orders for a certain minimum quantity are present. Inventory management is done according to a simple $(S-1, S)$ strategy at the retail stores. They only consider one-stage production systems with direct distribution relations to retail stores. The complex closed-form analysis makes it hard to adapt this approach to other settings.

In the context of assortment complexity, the concept of *postponement* becomes relevant. Lee and Tang[38] provide a quantitative model to assess the costs and benefits obtained by shifting the point of differentiation downstream in the production process by making more options common. The costs comprise required investments, additional material and processing costs. A more qualitative analysis of the effects of postponement on supply chains is provided by Yang and Burns[39]. Venkatesh and Swaminathan[40] discuss postponement as a key strategy to manage product variety and provide a case study that shows a successful implementation at Benneton. However, no model for such an assessment in the general case is provided.

3.1.2.3 Activity Based Costing and Related Approaches

One approach to assess the cost of assortment complexity is to precisely calculate the actual costs incurred by the production of each single product variant. On that basis, the cost incurred by a given variant can be compared to the revenue it generates in order to assess its profitability. Adam and Johannwille[41] argue that the results of traditional accounting approaches based on overhead costing may be misleading when they are used for variety-related decisions. New variants and slow movers are usually cross-subsidised by the standard products since the fixed overhead cost rates are applied to all products. There two main reasons for this are:

37 Ibid., p. 563. We discuss risk pooling in Section 3.2.4.
38 Lee and Tang (1997).
39 Yang and Burns (2003) and partly their more recent works (Yang et al., 2004a,b).
40 Venkatesh and Swaminathan (2003).
41 Adam and Johannwille (1998, pp. 16-19).

- Additional overhead costs created by the introduction of new variants are equally split up over all products, including the existing variants. Such additional overhead costs should only be assigned to the products that cause them in order to make them only appear profitable if they actually cover these expenses.
- Overhead costs are not calculated according to the actual resource consumption of each individual product, but proportionally to the single unit costs. A cause-fair distribution of overhead costs must consider the time and quantity requirements for resources of each product to avoid cross-subsidisation.

The inability of traditional accounting approaches to correctly assess assortment complexity-induced costs has led to much research in the development of alternative accounting systems that circumvent these drawbacks. Most of these approaches come from the area of activity based costing (ABC), an accounting principle mainly developed by Johnson and Kaplan[42], Cooper and Kaplan[43] and Horváth and Mayer[44].

Activity based costing tries to improve the assignment of overhead costs by focussing on the resource utilisation of each cost unit. Within each cost centre, the corresponding processes are analysed, documented and later aggregated over cost centres to several main processes. For each main process, the costs incurred by the required resources are derived. Process costs can be quantity-invariant or quantity-dependent. For the latter type a process cost rate is derived that reflects the resource costs for the output quantity of a single process execution.[45]

The usage of process cost rates in place of fixed overhead cost rates make it possible to consider the fact that the materials produced have different characteristics and different resource requirements. In the context of a correct assessment of assortment complexity-induced costs, three different effects can be observed:[46]

Allocation effect In contrast to traditional accounting systems, overhead costs are not fixed, but assigned according to the utilisation of operational resources.

Complexity effect More complex products are drivers of overhead costs as they increase the effort for indirect activities related to their production. The applica-

42 Johnson and Kaplan (1991).
43 Cooper and Kaplan (1988).
44 Horváth and Mayer (1989), who established activity based costing especially in the German speaking countries.
45 A more detailed description is beyond the scope of this summary and can be found e.g. in Coenberg and Fischer (1991).
46 See Heina (1999, pp. 66-67).

tion of activity based costing thus leads to lower cost rates for standard products and higher rates for slow movers.

Degression effect Unit costs decrease with higher production volume due to the fact that constant process cost are split up. This leads to a degressive cost function depending on the order volumes, whereas traditional accounting systems imply constant overhead costs per unit.

Due to these advantages, activity based costing has been widely used in the context of variety management. Horváth and Mayer[47] provide several examples of how variant costs can be calculated with the help of activity based costing. Kaiser[48] first uses it to develop a more comprehensive variety management method based on a profitability analysis for each variant. Such applications have shown that activity based costing tends to result in much better estimates for variant costs, i.e. slow movers are assigned a higher share of the overhead costs, while standard products get cheaper, compared to the cost rates that result from traditional overhead calculation.

The basic idea of activity based costing has been extended in many ways. Schuh[49] presents an extension called *resource oriented* activity based costing. The main differences are that there is no aggregation of single processes to main processes, but a complete process hierarchy is built such that each sub-process is assigned exactly one resource. Furthermore, he distinguishes between technical and value-related resource utilisation. The core element of this approach is the functional description of the resource utilisation by a certain sub-process. These relationships are defined in so-called *nomograms*, which combine functions to describe the resource utilisation and cost developments. The resource oriented activity based costing has also been combined with target costing approaches[50], especially by Horváth and Seidenschwarz[51] and Schuh[52].

Heina[53] presents an integrated method to determine the optimal product variety. In the context of the required cost assessment, he mainly builds on activity based costing methods and adapts them in several ways to make them *decision-oriented.* Firstly, he proposes to reduce the required effort by assigning costs to individual product characteristics rather than product variants. Secondly, he splits the entire cost calculation

47 Horváth and Mayer (1989, p. 217).
48 Kaiser (1995).
49 Schuh (2005).
50 For a description of target costing, see Seidenschwarz (1991).
51 Horváth and Seidenschwarz (1992).
52 Schuh (2005).
53 Heina (1999).

into a basic and variant calculation such that only the relevant characteristics that differ between products are considered in the detailed analysis.

Adam and Johannwille[54] argue that both traditional accounting approaches as well as activity based costing approaches and its extensions cannot fulfil the requirements of an analysis of complexity-related costs. Traditional overhead calculations suffer from the fixed and equal overhead cost rates assigned to each variant. The activity based costing approaches discussed here suffer from various drawbacks as well:

- The definition of the processes requires standardised and repetitive activities. Especially the production of exotic variants often requires non-standard activities for which no general processes can be defined in the first place.
- It is difficult to correctly define the relation between the materials and required resources, especially for those activities that are only indirectly related to the value adding processes. Direct relationships are usually only defined for production processes in terms of material routings.
- Process complexity and costs increase with the number of variants. Therefore, the problem that standard products are assigned overhead costs that they do not cause still remains, as these products now face increased unit cost rates.
- There is no way to consider fixed step costs. Additional complexity may require to create additional coordination resources which cannot be included in the calculations since only currently existing resources are considered.

A more detailed discussion of these existing drawbacks is provided by Adam and Johannwille[55] and Glaser[56].

3.2 Inventory Allocation in Production and Distribution Networks

Inventory management is an important pillar of supply chain management research. Neale et al.[57] analyse the impact of inventory management on supply chain performance and stress its importance. Lee and Billington[58] describe common pitfalls and

54 Adam and Johannwille (1998, pp. 14-22).
55 Ibid.
56 Glaser (1993).
57 Neale et al. (2003).
58 Lee and Billington (1992).

opportunities in inventory management and give practical suggestions how logistics managers can identify optimisation potential in this area. Simchi-Levi et al.[59] present a case study showing how the optimisation of the *"location of inventory across the various stages of the manufacturing and assembly process"* and the *"quantity of safety stock for each component at each stage"* can significantly reduce total cost.

The question what savings can be achieved in inventory management by altering the product assortment also leads to the question of optimal stock positions and dimensions in the production and distribution network under consideration. Chung et al.[60] define this problem as *"the task [...] to deploy, possibly subject to constraints, inventory at a number of stages in a supply network through the exploitation of economies of scale and flexibility so as to serve customer demand optimally"*. An important distinction for inventory models is between single installation (i.e. single product and location) and multi-echelon systems.[61] As any multi-echelon model builds on some model for the single installations, we will briefly summarise these models before discussing mutli-echelon systems relevant in the context of this work.

3.2.1 Single Installation Systems

The ultimate cause of inventory are imbalances between supply and demand processes. These imbalances may be desired or not. Desired or intentional inventory results from the exploitation of economies of scale, where batching production or order quantities yields savings in setup and ordering cost as well as supplier discounts. Other reasons comprise technological restrictions like minimum lot sizes. These inventories are called *cycle stocks*. Another class of inventory is *pipeline* or *work in progress* stock caused by finite production and transport rates. Further *anticipation stocks* may be caused by capacity limits, production smoothing or expected price fluctuations. Besides these intentionally planned inventories installed to decrease costs of production and distribution, *safety stocks*[62] are required as means of buffering against uncertainties present in demand, supply and transformation processes.[63]

For the context of this work, models for inventory positioning and dimensioning are relevant. The base stock model is one of the most important inventory models for single

59 Simchi-Levi et al. (2005, pp. 305–313).
60 See Chung et al. (2006, p. 178).
61 See the discussions of different system types in Zipkin (2000) and Axsäter (2006).
62 See the explanations on safety stocks in Section 2.2.
63 This is a generally accepted classification of inventory types in literature, e.g. in Chopra and Meindl (2004, pp. 57-59).

installation systems and serves as the basis for many multi-echelon models. Given a single installation that faces stationary stochastic demand, the stock is increased to a base stock level y in each discrete time period. Demand is assumed to follow a probability distribution with mean μ and standard deviation σ. The base stock level can then be written as

$$y^* = \mu + z \cdot \sigma \tag{3.1}$$

The base stock is set to the expected demand μ plus a safety stock, measured in multiples of the demand standard deviation. This multiple is called the *safety factor*, denoted z. If a positive lead time T exists, the base stock level has to cover the entire *lead time demand* over T, which results in the base stock formula

$$y^* = \mu \cdot T + z \cdot \sigma \cdot \sqrt{T} \tag{3.2}$$

Given that μ, σ and T are known, the safety factor needs to be specified. As the safety factor determines the degree to which the inventory buffers against demand fluctuations, it has to be set in accordance with some objective. Such objectives comprise required service levels that must be met or the minimisation of inventory holding and backordering costs.

If demand is assumed to follow a certain probability distribution and the safety stock should be dimensioned to guarantee a given stockout probability (α-service level)[64] or fill rate (β-service level)[65], standard procedures to determine the required safety factor exist. If no predefined service level exists and if both inventory holding and backordering costs can be quantified, the safety factor can be determined to minimise the sum of inventory and backordering costs incurred. If excess inventory incurs a holding cost of h monetary units per period and unmet demand is backordered at cost of p monetary units per period, then the optimal α-service level can be calculated as

$$\alpha^{SL} = \frac{p}{p+h} \tag{3.3}$$

Equation 3.3 is also called the *newsboy ratio* because the single period, single installation base stock model is often referred to as the *newsboy problem* in which a news vendor has to determine the optimal stock level for newspapers, given that unmet demand can be quantified as lost sales and excess newspapers incur penalty costs.[66]

64 The α-service level is defined as the probability of having sufficient inventory to meet all demand in a given period. In other words, it is the proportion of periods where no stockout occurs.

65 The β-service level is defined as the proportion of demand that is met from stock and can be delivered in time

66 See Axsäter (2006, pp. 114-116).

With this ratio, the cost minimisation problem can be solved with the methods for α-service levels.

For the common cases where demand is assumed to follow a normal distributed $N(\mu, \sigma)$, the required safety factor can be determined straightforward. If the required service level is given as an α-service level, the density function $F(\cdot)$ of this probability distribution can be used to calculate the corresponding safety factor[67], while the standard loss function $L(\cdot)$ is used for the case of the β-service level criterion[68].

Extensions to these models include the integration of other sources of uncertainty, like processing delays or imperfect supplier service. If these additional uncertainties can also be described via probability distributions, the distribution of the aggregate uncertainties is determined as the convolution of these individual distributions. Generally, this distribution has to be known to derive an appropriate safety factor. The convolution of several distributions easily becomes statistically and computationally complex. Even if an aggregate distribution can be derived, the determination of appropriate safety factors remains a difficult task, which is why the majority of works on safety stock models either considers one type of uncertainty only or assumes all uncertainties to be normally distributed. For the latter case, the aggregate uncertainties are again normally distributed and the above-mentioned models can be used.

3.2.2 Multi-Echelon Systems with Various Products

For the extension of single installation systems to multi-echelon systems with various products and stochastic demand, two main branches of research can be identified. The most active research on inventory allocation is built on the work of Simpson Jr.[69]. The second branch of research is based on the work of Clark and Scarf[70]. Both works propose a modelling approach for serial production and distribution systems to analyse possible inventory allocations and have been extended successively.

The approaches differ in terms of modelling the replenishment mechanism and the resulting service time characteristics. In the stochastic service approach (SSA) proposed by Clark and Scarf, the service (=delivery) times at one stage are stochastic and vary based on the material availability at supplying stages. In the guaranteed

67 See Axsäter (2006, p. 96), (Nahmias, 1997, pp. 272 et seqq.) and Zipkin (2000).
68 See van Ryzin (2001).
69 Simpson Jr. (1958).
70 Clark and Scarf (1960).

service approach (GSA) proposed by Simpson Jr., each stage quotes a service time to its successor which it can always satisfy.[71]

Both approaches work on network models in which each node i represents an item in a supply chain that performs some operation like a production or transport process. Each such item has a known and deterministic throughput time and is a potential stockpoint that can hold safety stock after the processing is finished. The only real source of uncertainty is stochastic customer demand, which is represented as some probability distribution with known demand and standard deviation. None of the original models considers any capacity constraints, i.e. production stages are assumed to produce arbitrary quantities within one period and external suppliers can always deliver any quantity ordered. The inventory cost is assumed to be linear in the quantities held on stock.

3.2.2.1 Stochastic Service Models

The stochastic service approach makes the following additional assumptions:

- serial network structure
- linear backordering cost rates known for each item
- internal shortages immediately affect service time to succeeding location

The third assumption is the main model property: The replenishment lead time of an item becomes stochastic due to the stochastic service time of its predecessor. This makes the determination of the required safety stock levels much harder since the safety stock has to buffer against demand variability over a stochastic replenishment lead time. While the external service level is given as a parameter, the internal service levels are an additional decision variable in the optimisation problem.

Clark and Scarf[72] introduce the concept of *echelon stock*, defined as *"the stock at any given installation plus stock in transit to or on hand at a lower installation"*. They show that under the above-mentioned assumptions, the optimal stock policy is an *order-up-to policy* with order-up-to levels $(Y_1^*, \ldots, Y_n^*)$. This means that each item takes into account the amount of stock available at all items further downstream the network and fills orders such that the echelon stock at that location i reaches its order-up-to-level Y_i^*. The calculation of the optimal order-up-to-levels is computationally

71 The notions *stochastic service approach* and *stochastic service approach* were not coined in the original works but by other authors, e.g. Klosterhalfen and Minner (2006).
72 Clark and Scarf (1960, p. 1784).

very complex, partly since it requires numerical integration. Clark and Scarf present a dynamic programming algorithm to determine these levels sequentially, beginning with the most downstream stockpoint.

The stochastic service approach has some practical drawbacks. Firstly, the analytical model is mathematically difficult due to the stochastic modelling of service times. This makes it very hard to extend this model to more complex, especially non-serial networks. Secondly, the model does not distinguish explicit stockpoints, but assumes that each stage holds stock and periodically fills this stock up to its order-up-to level. This may become a problem depending on the level of abstraction of the network model. It is reasonable if the network items represent physical production or storage locations, but it is unrealistic if the network is a detailed model at the single product level. In this case, this approach prescribes an order-up-to-level for all products at all stages, including raw and semi-finished materials. Thirdly, the echelon stock concept assumes central control, as planners at each installation have to consider stock at downstream locations.

There are some works that extend the approach of Clark and Scarf in different ways. As the suitability of this model for the problem considered in this work is limited, the reader is referred to van Houtum et al.[73] and Minner[74] for an overview of such extensions. Among other extensions, the approach has been adapted to some more realistic network structures like distribution and assembly structures. However, there is currently no extension to general networks in which each item can have arbitrary numbers of predecessors and successors. Other works present adaptations to different service measures, e.g. the fulfilment of a certain α-service-level instead of backordering costs.[75]

3.2.2.2 Guaranteed Service Models

The area of most active research on inventory placement is built on the work of Simpson Jr.[76], who proposes a model for serial production and distribution systems to analyse possible stockpoint allocations. In contrast to the approach of Clark and Scarf, Simpson makes several simplifying assumptions that facilitate the analysis:

73 van Houtum et al. (1996).
74 Minner (2000, pp. 123-125).
75 See Lagodimos and Anderson (1993).
76 Simpson Jr. (1958).

- Each item quotes a constant and deterministic service time to its successor that it can always satisfy.
- Demand covered by safety stock is bound by a *maximum reasonable demand* determined via the α-service level required for that item. The α-service level defines the maximum reasonable demand via the probability that actual demand exceeds this value. With assumptions about the demand distribution, the maximum reasonable demand can be computed.[77]
- All demand that exceeds this maximum reasonable demand is assumed to be handled by *operating flexibility*, e.g. emergency orders, accelerated or overtime production and rescheduling of the existing production sequence, so that stockouts do not affect the agreed delivery time at successor stages.
- The total replenishment lead time for each stage is deterministic and equals the service time of its successor plus its own processing time.

The assumption of operational flexibility and the resulting deterministic service times are discussed controversially. While replenishment lead times are surely not always deterministic in reality, there probably is some operational flexibility that can be used to react to unexpected demand fluctuations. In this sense, the guaranteed service approach is more realistic as it does not consider safety stocks as the only mean to react to demand uncertainty. This has led to a bias towards GSA models, as their assumptions are justifiable and tremendously facilitate the analysis.

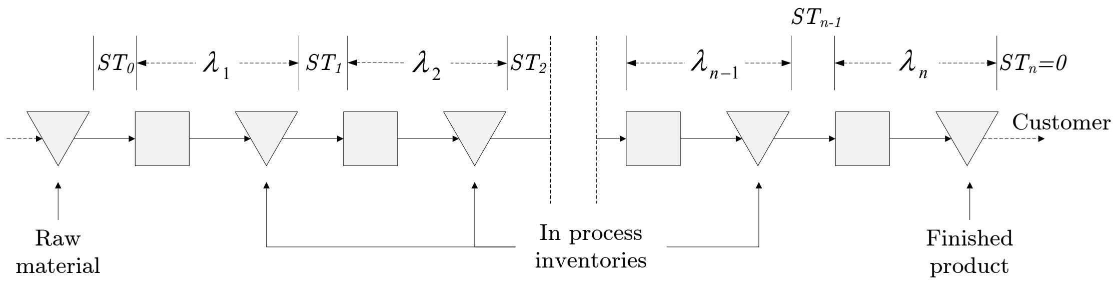


Figure 3.2: Basic model of guaranteed service approaches

Figure 3.2 depicts the basic model and the main parameters for a serial network. Each stage i has a processing time λ_i and quotes a service time ST_i to its successor. The replenishment lead time of i can thus be computed as $ST_{i-1}+\lambda_i$. The time interval that

77 Simpson Jr. uses this approach in his work for normally distributed demand.

has to be covered with safety stocks equals this replenishment lead time, diminished by i's service time[78]: $ST_{i-1} + \lambda_i - ST_i$.

The aim to find combinations of service times and corresponding safety stock levels that minimise the total safety stock cost[79] can now be written as

$$\begin{aligned} \min \quad & \sum_{i=1}^{n} c_i \cdot \sigma_i \cdot z_i \cdot \sqrt{ST_{i-1} + \lambda_i - ST_i} \\ \text{s.t.} \quad & ST_i \leq ST_{i-1} + \lambda_i \qquad \forall i \in \{1, \ldots, n\} \\ & ST_i \geq 0 \qquad \forall i \in \{1, \ldots, n\} \\ & ST_0 = 0 \\ & ST_n = 0 \end{aligned}$$

The objective function seeks to minimise the safety stock cost required to guarantee the required service level, which determines the safety factors z_i. The constraints assure that the service times are none-negative and that a stockpoint's service time is not longer than its replenishment lead time. The central result of Simpson's work is the *extreme point property*, which states that in all optimal solutions

$$ST_i \in \{0, ST_{i-1} + \lambda_i\} \tag{3.4}$$

holds for all $i = 1, \ldots, n$. Each item either covers its entire replenishment lead time with safety stock (and has a service time of 0) or does not hold any safety stock and passes its entire replenishment lead time plus the throughput time to its successor (and has a service time of $ST_{i-1} + \lambda_i$). This is also referred to as the *all or nothing policy* since the service times are set to cover either the maximum or minimum feasible values.

While the original work of Simpson only addresses serial systems and does not propose any solution technique for the optimisation problem apart from testing all possible combinations, many extensions have been proposed in past years. If some restrictive assumptions regarding the structure of the network are made, dynamic programming can be used to solve the optimisation problem efficiently. The above-mentioned extreme point property is the basis for all such approaches that make use of the fact that the optimal service time of each item can be expressed via the service times of

78 The same modelling approach can be found in Simchi-Levi et al. (2005, pp. 194-196), who calls this time interval the *net lead time*.

79 See Simpson Jr. (1958), Minner (2000, p. 93).

the downstream and upstream items.[80] The safety stock levels can be calculated via the replenishment lead times that are not yet covered by safety stock at upstream items.[81]

Inderfurth[82] proposes such approaches for convergent and divergent network structures and later extends these models to different service criteria[83]. Graves and Willems[84] propose a special algorithm for the case that the network is a spanning tree. They also extend the problem formulation to include strategic supply chain configuration decisions.[85] These decisions comprise the selection of suppliers, parts, processes and modes of transport, given that for each stage there are different options that can be distinguished by their lead times and costs added. The optimisation model then chooses a sourcing option for each stage as to minimise the total cost, comprising safety stock costs.[86] Minner[87] provides an overview of how the original model can be extended to different special network types and how dynamic programming algorithms are used to solve the safety stock allocation problem optimally.

Despite the progress in this area of research, no optimal algorithm for general networks exists so far. General networks are those networks where each node can have an arbitrary number of predecessors and successors, which is generally the case if the network is used to model an inventory system on the single product level.

3.2.3 Inventory Allocation as a Combinatorial Optimisation Problem

The extreme point property from Equation 3.4 makes the inventory allocation problem a *combinatorial optimisation problem.* In this subclass of general optimisation problems the decision variables are discrete, i.e. *"the solution is a set, or a sequence of integers or other discrete objects"*[88]. According to Blum and Roli[89], a combinatorial optimisation problem is defined by

- a set of variables $X = \{x_1, \ldots, x_n\}$
- variable domains $D_1, \ldots, D_n$

80 See Minner (1997).
81 See Inderfurth (1992, p. 23).
82 Inderfurth (1991).
83 Inderfurth and Minner (1998).
84 Graves and Willems (1996, 2000).
85 Graves and Willems (2003).
86 Ibid., p. 121.
87 Minner (2000).
88 Reeves (1993, p. 2).
89 Blum and Roli (2003, p. 269).

- constraints among variables
- an objective function $f : D_1 \times D_1 \cdots \times D_n \mapsto \mathbb{R}^+$ to be minimised.

The set of feasible solutions is called the *solution space* and is given by

$$\mathcal{S} = \{s = \{(x_1, v_1), \ldots, (x_n, v_n)\} \mid v_i \in D_i, s \text{ satisfies all the constraints}\} \tag{3.5}$$

Among this set of candidate solutions, there is a subset of optimal solutions $\mathcal{S}^* \subseteq \mathcal{S}$ with minimum objective function values: $\mathcal{S}^* = \{s^* | \ f(s^*) \leq f(s) \forall s \in \mathcal{S}\}$.

Given that each item in the network either covers its entire replenishment lead time with safety stock or does not hold any safety stock and passes its entire replenishment lead time plus the processing time to its successor[90], the remaining decision is a binary *stockpoint* or *no stockpoint* decision. The safety stock allocation problem described in Section 2.3.2 can thus be mapped to a combinatorial optimisation problem. The set of decision variables is $X = \{sp_i, \ldots, sp_n\}$ with binary stockpoint indicators $sp_i \ \forall i \in \mathcal{N}$ as the decision variables. All domains are $D_i = \{0, 1\} \ \forall i \in \mathcal{N}$ and the solution space is $\mathcal{S} = \{s = \{(sp_i, v_1), \ldots, (sp_n, v_n)\} \mid v_i \in D_i, s \text{ satisfies all the constraints}\}$. The set of stockpoint nodes $\mathcal{SP}$ then is $\mathcal{SP} = \{i \in \mathcal{N} | \ sp_i = 1\}$.

For each item, the only decision that has to be taken is whether or not it holds safety stock. Simpson states that due to this property, there are $2^n - 1$ possible combinations for a serial network with n items and a predefined service time of $ST_n = 0$. More generally, the problem complexity is $\mathcal{O}(2^n)$ for general networks regardless of the assumptions made with respect to the service times of items without successors. The exponential increase of the problem size makes the computation of optimal solutions for realistic problem sizes with *complete* optimisation methods[91] impossible in reasonable time.

Combinatorial optimisation problems are particularly suited to be addressed with approximate or *heuristic*[92] solution techniques. While each heuristic has to be prob-

90 See Section 3.2.2.2.

91 *"Complete algorithms are guaranteed to find for every finite size instance of a combinatorial optimisation problem an optimal solution in bounded time"* (Blum and Roli, 2003, p. 269).

92 *"A heuristic is a technique which seeks good (i.e. near-optimal) solutions at reasonable computational cost without being able to guarantee either feasibility or optimality, or even in many cases to state how close to optimality a particular feasible solution is"* (Reeves, 1993, p. 6). The term heuristic is derived from the Greece *heuriskein* (ευρισκειν), which means *to find.*

lem specific to some extend, several high-level strategies called *meta-heuristics*[93] have emerged that are aimed at efficiently and effectively exploring the solution space of a combinatorial optimisation problem. Among these, the most relevant are evolutionary computation including genetic algorithms, ant colony systems, iterated local search, scatter search, greedy randomized adaptive search, simulated annealing and tabu search.[94]

The only application of meta-heuristics to the inventory allocation problem is presented by Minner[95], who uses the local search heuristics tabu search and simulated annealing to determine stockpoints in a general network.

Local search algorithms start from some initial solution and iteratively try to improve the current solution by replacing it with a solution from an appropriately defined neighbourhood. Given a current feasible solution s, a local search heuristic examines a neighbourhood $U(s)$ and may select one of its members to be the new current solution. The neighbourhood is defined as the set of solutions that can be reached by applying a single *move* or *operation* to the current solution. These moves usually comprise changes to the values of certain elements in the current solution. The main problem of local search heuristics is the risk of getting stuck in so-called local optima, i.e. in a solution $s \notin \mathcal{S}^*$ whose neighbourhood does not contain any solution with a better objective value: $f(s) \leq f(\hat{s}) \; \forall \hat{s} \in U(s)$.

Meta-heuristics employing local search provide different strategies to prevent the optimisation from getting stuck in such local optima. The simplest approach is the one of iterated local search, which restarts the search with a modified initial solution after a local optimum has been found. The above-mentioned meta-heuristics simulated annealing (SA) and tabu search (TS) pursue more sophisticated strategies.

The basic idea of SA is to accept inferior neighbourhood solutions $\hat{s}$ with a certain probability. The difference $f(\hat{s}) - f(s)$ together with the current *temperature* value determine the probability at which an inferior neighbourhood solution $\hat{s}$ is accepted. The temperature value decreases during the search process according to a temperature

93 *"A meta-heuristic is an iterative master process that guides and modifies the operations of subordinate heuristics to efficiently produce high-quality solutions. It may manipulate a complete (or incomplete)single solution or a collection of solutions at each iteration. The subordinate heuristics may be high (or low) level procedures, or a simple local search, or just a construction method"* (Voß, 2001, p. 5). For a description of common characteristics of meta-heuristics, see Blum and Roli (2003, pp. 270-271).

94 For more detailed descriptions of the above-mentioned meta-heuristics and their applications, see Dréo et al. (2006), Reeves (1993), Blum and Roli (2003) and Voß (2001). Specifically, we refer to Dorigo et al. (1999) for ant colony systems and Holland (1975) for genetic algorithms.

95 Minner (2000, pp. 154-169).

function, so that large deteriorations are accepted during the first iterations, while the probability that a non-improving solution is accepted is successively reduced.[96]

Like SA, tabu search is not a pure improvement procedure, but always accepts the best solution in the current neighbourhood, even if it is worse than the current solution. The best solution found so far is stored during the entire process and returned when a stopping criterion is met. Tabu search is a memory-based approach that uses information about the search history to determine future moves. This is done via a tabu list, which defines a set of forbidden moves that must not be executed for a certain number of iterations, called tabu tenure, in order to prevent cyclic computations. In many implementations, this constraint may be violated if an aspiration criterion is met, allowing the move to be performed despite its tabu status. In order to avoid getting stuck in local optima, TS may also make use of diversification strategies that encourage the search process to examine regions of the solution space that have not been visited so far by generating new initial solutions that differ significantly from the solutions visited before. On the other hand, intensification strategies make the search examine the neighbourhood of the best solutions found so far or generate new solutions by combining the best solutions visited.[97]

Minner uses both meta-heuristics for the inventory allocation problem, as they only differ in the strategy to avoid local optima and can use the same solution representation and neighbourhood definition. He defines the solution representation for the safety stock allocation problem as a binary vector whose entries indicates whether or not safety stock is held at an item $i \in \mathcal{N}$. The neighbourhood comprises all solutions that can be reached by any zero/one switch for each single bit within the binary vector.[98]

$$U(s) = \{\hat{s} \mid \hat{s}_i = 1 - s_i \text{ for one } i \in \mathcal{N}, \hat{s}_j = s_j \forall j \in \mathcal{N} \setminus \{i\}\} \tag{3.6}$$

A move thus corresponds to changing the stockpoint status of one node from non-stockpoint to stockpoint or vice versa. With this neighbourhood definition, the author uses both the simulated annealing and the tabu search meta-heuristic to escape local optima in the search process. Minner tests both approaches and shows that near optimal results can be obtained for some problem instances. No significant differences between the performance of TS and TS can be observed.

While this use of meta-heuristics for the inventory allocation problem appears successful, the size of the networks used for testing is not clearly stated. Since the author

96 For a more elaborate tutorial on SA, see Eglese (1990).
97 See Glover (1989, 1990) for a comprehensive tutorial on TS.
98 See Minner (2000, p. 154).

claims that the comparison is done against optimal solutions that *"have been determined by enumeration"*[99], the network size n must have been very limited to allow the full enumeration of all 2^n possible configurations. Therefore, this can only be considered a first proof of the general suitability of local search based meta-heuristics to the safety stock allocation problem. In order to be able to perform this optimisation in large networks, the search process should employ more sophisticated moves than all possible binary switches on the stockpoint statuses.

3.2.4 Risk Pooling Effects in Inventory Management

Some of the expected positive effects of assortment reduction and standardisation are based on the concept known as *risk pooling*. The statistical phenomenon underlying this concept has been summarised by Nahmias[100] as follows: *"the variance of the average of a collection of independent identically distributed random variables is lower than the variance of each of the random variables; that is, the variance of the sample mean is smaller than the population variance"*.

This concept helps to understand many economic phenomena, e.g. in banking and insurance, as well as in supply chain management. Particularly, the effects obtained by consolidating multiple random demands have been observed and analysed in the context of inventory management.[101] This consolidation may take many forms, either a geographical consolidation of multiple inventories at one physical location, or the consolidation at the product level by rationalising product lines.[102] Hopp[103] characterises risk pooling as follows: *"the combination of sources of variability [...] reduces the total amount of buffering required to achieve a given level of performance"*.

In the context of this work, risk pooling occurs as the reduction in demand variability and forecast deviations. Sobel[104] states that *"the standard deviation of a sum of interdependent random demands can be lower than the sum of the standard deviations of the component demands"*. With respect to forecast accuracy, Nahmias[105] finds that *"aggregate forecasts are more accurate. [...] On the percentage basis, the error made*

99 Minner (2000, p. 166).
100 Nahmias (1997, p. 61).
101 A detailed overview of risk pooling application areas and especially risk pooling in inventory management for different stock policies is provided by Sobel (2008).
102 Hopp describes the four applications areas warehouse centralisation, product standardisation, postponement and worksharing (Hopp, 2006, pp. 121-126).
103 Hopp (2006, p. 120).
104 Sobel (2008, p. 155).
105 Nahmias (1997, p. 61).

in forecasting sales for an entire product line is generally less than the error made in forecasting sales for an individual item". Simchi-Levi et al.[106] make an important point by underlining that risk pooling is also obtained by commonality in product structure and therefore *"demand for a component used by a number of finished products has smaller variability and uncertainty than that of the finished goods".* These insights have led to a great research interest quantitative models of the costs and benefits of risk pooling.

Research on the effects of risk pooling in inventory management was initiated by Eppen[107], who analyses the effects of consolidating normally distributed demands from several facilities in a multi-location newsboy model, assuming linear holding and back-ordering costs[108]. He provides a model to derive expected inventory holding and back-ordering penalty costs for different demand parameters. Eppen's results have proven to be valid in several works and remain the key insights with respect to risk pooling in inventory management. He calls these findings the *statistical economies of scale*, which can be summarised as follows:

1. Total cost of centralised systems are lower compared to decentralised systems.
2. The magnitude of these savings depends on the correlation between demands.
3. For identical and uncorrelated demands, cost decreases by the factor $\frac{1}{\sqrt{n}}$ when consolidating n demands.

Finding 1 can be traced back to effects of pooling multiple random demands. Such pooling can be achieved via various means. Apart from consolidation of physical warehouses, a reduction of assortment complexity can yield the same effects, as either end customer demand is concentrated on fewer end products or individual end product demands can be pooled as aggregated demand for semi-finished products and raw materials. As this concept applies both to the end product level as well as to all intermediate components, demand for a component used by a number of finished products is less volatile and uncertain than that for finished goods. Accordingly, postponement strategies[109] are also seen as one way to achieve the desired risk pooling effects.[110]

106 Simchi-Levi et al. (2005, p. 312).
107 Eppen (1979).
108 See Section 3.2.1.
109 See Venkatesh and Swaminathan (2003) for a general introduction to the concept of postponement.
110 Yang et al. (2004a,b) analyse the implications of postponement on various types of uncertainty in supply chains and also provide an extensive literature review on this topic.

Finding 2 captures the fact that the magnitude of the risk pooling effect depends on the correlation between the stochastic factors. For example, if two demands are highly correlated, their aggregation hardly affects the total variability. If we consider normal demands, this can be shown easily, as the folding of two normal random variables $N_1(\mu_1, \sigma_1)$ and $N_2(\mu_2, \sigma_2)$ with a correlation coefficient of $\rho \in [-1, 1]$ is defined as

$$N_1(\mu_1, \sigma_1) \oplus N_2(\mu_2, \sigma_2) = N\left(\mu_1 + \mu_1, \sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}\right) \tag{3.7}$$

For fully correlated variables ($\rho = 1$), there clearly is no risk pooling effect since the fluctuations of the two stochastic elements cannot compensate each other. On the other extreme, risk pooling effects are biggest if the variables are fully negatively correlated ($\rho = -1$). Sobel[111] concludes that *"the advantage grows as the correlation shifts from strongly positive to strongly negative"*. For examples and graphical illustrations of this concept, the reader is referred to Alicke[112].

Finding 3 is closely linked to the above-mentioned conclusions about correlation and has become a well-known rule for both researches as well as practitioners. The so-called *square root formula* has become a standard technique in inventory management and can be found in any textbook on this topic[113]. When aggregating demands or consolidating warehouses while maintaining the same service level and service times, total system costs decrease as a strictly monotonically decreasing convex function of the number of demands or warehouses. So the relative benefit of consolidation decreases with the number of items or warehouses consolidated.

In order to quantitatively evaluate risk pooling effects for a wider range of realistic application scenarios, the model used by Eppen has been extended successively. Eppen and Schrage[114] introduce positive lead times. Chang and Lin[115] extend the model with the consideration of transport costs. This is one of the most important extensions, as the implementation of physical inventory pooling in central warehouses goes along with an increase in transport costs since the distances from the centralised warehouses to the customers are increased.[116] These additional costs have to be offset by the savings achieved via inventory reduction.

111 Sobel (2008, p. 159).
112 Alicke (2005, pp. 160-162).
113 See Chopra and Meindl (2004, pp. 313–317), Tempelmeier (2005, pp. 156 et seq.).
114 Eppen and Schrage (1981).
115 Chang and Lin (1991).
116 See Simchi-Levi et al. (2007, p. 232).

Recently, Corbett and Rajaram[117] generalised the findings to almost arbitrary multivariate-dependent demand distributions. They show how pooling of inventories can be analysed without the need to resort to assumptions of independence and normal distributions. For these generalised assumptions, they prove that centralisation or pooling of inventories is more valuable when demands are less positively correlated.

Alfaro and Corbett[118] discuss how risk pooling effects change if the system under consideration does not operate under an optimal inventory policy. They argue that in practice, an optimal policy is normally impossible to find and also consider non-normal demand distributions in this context. They provide a model to compare the relative value of implementing some form of inventory pooling against the value of improving a suboptimal inventory policy. The authors conclude that there is always a uniquely defined interval within which pooling leads to greater cost reductions than optimizing inventory policy, while the reverse is true outside that interval. Apart from this concrete model, the work also presents an elaborate literature review on inventory pooling.

Gerchak and He[119] show that the benefits of risk pooling increase with the variability of the original demands. This goes in accordance with intuition, as the positive effects of risk pooling increase with the risks that are aggregated. They prove this observation for certain mean-preserving variations of demand variability. Simchi-Levi et al.[120] draw similar conclusions and use the coefficient of variation of the demand process as a metric to evaluate the effects of risk pooling. The higher the coefficient of variation, the greater the effects of risk pooling and the inventory reduction. As a consequence, they suggest to evaluate the opportunities to foster consolidation for slow-moving items.

Kulkarni et al.[121] investigate risk pooling effects on a strategic level for the question of network configurations and evaluate the trade-off between logistic costs and risk pooling benefits in production networks. The alternatives considered are *product networks* with component manufacturing being spread over all plant and *process networks* with component manufacturing being consolidated in a single plant. They show that the process plant networks offer significant risk pooling advantages under a wide range of conditions, even without accounting for the benefits of economies of scale. They conclude that companies should consider this network configuration due to the risk pooling benefits offered.

117 Corbett and Rajaram (2006).
118 Alfaro and Corbett (2003).
119 Gerchak and He (2003).
120 Simchi-Levi et al. (2007, pp. 318–319).
121 Kulkarni et al. (2004, 2005)

Benjaafar et al.[122] extend the analysis of risk pooling in pure inventory systems to production-inventory systems, in which lead time demands are not exogenous, but influenced by the lead times that result from the sharing of the production supply process. Due to the consideration of the production stages with their characteristics like utilisation, there can be significant correlation in the lead-time demands of the different items, even if the individual demand processes are independent. This is due to the correlation between the lead times, which is particularly significant if the production system's utilisation is high. As correlation influences the potential benefits of risk pooling, they use the model to analyse how factors like utilisation, demand and service time variability and supply structure affect the benefits of risk pooling. Finally, they use the model to compare the benefits of inventory pooling to those of capacity pooling at the production stages.

As supply chains often do not have any central control, Hartman and Dror[123] analyse the fair allocation of benefits gained from inventory centralisation among various participants. They consider a system of retail stores that should be supplied from a central inventory location and use a game theoretic approach to share the savings among all participants such that no participant or subset of participants (called *coalitions*) has any incentive to order separately, even if the holding and backordering penalty costs are not the same at all stores and for all coalitions.

3.3 Determination of Planning Buffers and Planned Production Quantities

3.3.1 Determination of MRP Production Parameters

The planning buffer as defined in Section 2.1.2 can be seen as one particular parameter for production planning in MRP systems. The corresponding literature review therefore focuses on approaches that analyse the optimal determination of such planning parameters and their impact on production costs. The literature review shows that there is a certain set of parameters that have frequently been analysed. These parameters can be summarised as

- MPS frozen interval,

122 Benjaafar et al. (2005).
123 Hartman and Dror (2005).

- MPS replanning frequency,
- MPS planning horizon,
- product structure,
- forecast error,
- safety stock and
- lot-sizing rules.

Table 3.1 provides an overview of works that analyse the performance of MRP production systems under variation of different parameters. This comparative composition of relevant works is based on the extensive literature review conducted by Yeung and Wong[124]. Table 3.1 has been updated and extended by more recent works reviewed for this purpose.

Although the planning buffer is an important planning parameter that is also implemented in commonly used ERP systems[125], no work exists to the best of our knowledge that explicitly addresses its optimal determination. In particular, the trade-offs described in Section 2.1.2 have not yet been addressed in the literature.

3.3.2 Suitability of Standard Lot-Sizing Models

The determination of planned production quantities for all materials and several time periods shows many similarities to standard lot-sizing problems. Lot-sizing models determine production lots based on primary demand and under consideration of available resources, setup times and several cost components. The application context of this work does not require the determination of production lots and sequences on the detailed planning level. Moreover, we consider planned production quantities for each production stage and the respective materials over a larger time horizon. These conditions coincide with the characteristics of existing big-bucket lot-sizing models that allow production of different products within one time period without making any statements about their production sequence. Suerie[126] summarises the basic assumptions of the Capacitated lot-sizing problem (CLSP) as the most common multi-item basic big-bucket model as:

124 Yeung and Wong (1998).
125 For example, SAP R/3 allows the definition of this parameter as the *scheduling margin key*, which determines float periods before and after production orders (Dickersbach et al., 2005, p. 259).
126 Suerie (2005, p. 14).

	Frozen interval	Replanning frequency	Planning horizon	Product structure	Forecast error	Safety stock	Lot-sizing rules
Kunreuther and Morton (1973, 1974)	○	○	●	○	○	○	○
Whybark and Williams (1976)	○	○	○	○	○	●	○
Baker (1977)	○	○	●	○	○	○	○
Carlson et al. (1979)	○	○	○	○	○	○	●
Kropp et al. (1979)	○	○	○	○	○	○	●
Baker and Peterson (1979)	○	○	●	○	○	○	○
Blackburn and Millen (1980)	○	○	○	○	○	○	●
Carlson et al. (1982)	○	○	●	○	○	○	○
Biggs and Campion (1982)	○	○	○	○	●	○	○
Blackburn and Millen (1982a,b)	○	○	○	○	○	○	●
DeBodt and Wassenhove (1983)	○	○	○	○	●	●	●
Wemmerlov and Whybark (1984)	○	○	○	○	○	○	●
Chung and Krajewski (1984)	○	○	●	○	○	○	○
Benton and Srivastava (1985)	○	○	○	●	○	○	○
Wemmerlov (1985)	○	○	○	○	●	○	○
Yano and Carlson (1985)	○	●	○	○	○	●	○
Lee and Adam (1986)	○	○	○	●	●	○	○
Chung and Krajewski (1986)	○	●	○	○	○	○	○
Carlson and Yano (1986)	○	○	○	○	○	●	○
Wemmerlov (1986)	○	○	○	○	●	○	○
Sridharan et al. (1987)	●	○	○	○	○	○	○
Yano and Carlson (1987)	○	●	○	○	○	○	○
Sridharan and Berry (1990a,b)	●	●	○	○	○	○	○
Ristroph (1990)	○	○	○	○	○	○	●
Barrett and LaForge (1991)	○	●	○	○	○	○	○
Benton (1991)	○	○	○	○	○	●	○
Bregman (1991a,b)	○	○	○	○	○	○	●
Lin and Krajewski (1992)	●	●	●	○	○	○	○
Lee et al. (1993)	○	○	○	○	○	○	●
Zhao and Lee (1993)	●	●	○	○	●	○	○
Russell and Urban (1993)	○	○	●	○	○	○	○
Lin et al. (1994)	○	●	●	●	○	○	○
Sridharan and LaForge (1994a,b)	●	○	○	○	○	○	○
Kadipasaoglu (1995)	●	○	○	○	○	○	●
Zhao et al. (1995)	●	○	○	○	○	○	●
Zhao and Lee (1996)	●	●	●	○	○	○	○
Molinder (1997)	●	○	○	○	○	●	●
Enns (2001)	●	○	○	○	○	○	●

Table 3.1: Classification of studies regarding MRP parameter settings

- Several products are produced on one shared resource with limited capacity.
- The planning horizon is finite and divided into discrete periods.
- All products face a deterministic dynamic demand.
- If a product is produced in a certain period, the resource has to be set up for this product in this period.
- Setups consume resource capacity and incur a setup cost.
- The aim is to minimise the sum of holding costs and setup costs.

All these assumptions coincide with the characteristics of the corresponding problem considered in this work. For each production process step with a limited capacity, planned production quantities have to be determined for a sequence of time periods, where multiple materials are produced within one period. As setup costs should be estimated via the number of products with strictly positive production volumes in that periods, the assumptions regarding setup costs also holds. Based on these assumptions, the CLSP can be formulated as a mathematical optimisation problem:

$$\begin{array}{lll} \min & \sum_{j \in J} \sum_{t \in T} h_{i,t} \cdot I_{j,t} + \sum_{j \in J} \sum_{t \in T} sc_j \cdot Y_{j,t} & \\ \text{s.t.} & I_{j,t-1} + X_{j,t} = d_{j,t} + I_{j,t} & \forall j \in J, t \in T \\ & \sum_{j \in J} a_j \cdot X_{j,t} + \sum_{j \in J} st_j \cdot Y_{j,t} \leq c_t & \forall t \in T \\ & X_{j,t} \leq b_{j,t} \cdot Y_{j,t} & \\ & X_{j,t} \geq 0,\ I_{j,t} \geq 0,\ I_{j,0} = 0, Y_{j,t} \in \{0,1\} & \forall j \in J, t \in T \end{array}$$

Following the notation used by Suerie, the model determines optimal lot sizes $X_{j,t}$ for each product $j \in J$ and period $t \in T$. The objective function seeks to minimise total setup and holding costs by incurring holding costs $h_{j,t}$ for the inventories $I_{j,t}$ present in each period and incurring setup costs sc_j if product j is produced in period t, as indicated by $Y_{j,t}$. In the order of occurrence, the restrictions ensure the inventory balance, consideration of available capacities and enforce the binary production indicators $Y_{j,t}$ to be equal to 1 where required. For a more in-depth discussion of this basic model and its extensions, the reader is referred to the works of Suerie[127] and Tempelmeier[128]. Due to the similarities in the underlying assumptions, the CLSP provides a good basis to build the required optimisation model for planned production quantities upon.

127 Ibid.
128 Tempelmeier (2005).

CHAPTER 4

Required Work

Promise less or do more.
The Whitest Boy Alive

In this chapter we deduce the conceptual tasks to be addressed in Chapter 5 via a comparative analysis of the findings from Chapter 3 and the requirements identified for each subproblem in Section 2.3. We also show how the solutions to these tasks can contribute to and extend the current state of the art.

4.1 A Model to Assess the Cost Effects of Assortment Complexity Changes

Having reviewed existing approaches to assessing the costs of assortment complexity, we may conclude that these approaches fall into two categories. Firstly, some use accounting methods to fairly assign the costs to the product variants according to the input involved. Some drawbacks of these approaches have already been discussed in Section 3.1.2.3. But even if it were possible to assign all overhead costs perfectly accurately to each variant, the interrelations between the single variants are always neglected. Consequently, these approaches cannot evaluate cost effects that occur if certain combinations of assortment change decisions are evaluated. None of these methods tries to analyse the changes to the optimal configuration of the production and distribution network that an assortment change might bring. Secondly, some approaches try to derive general cost functions with the number of variants as the only independent variable. While this approach is in principle suited for what-if analyses, using the *number* of product variant as the only input variable requires many simplifying assumptions that make this type of analysis very imprecise. If concrete assortment changes form the input of the analysis, potential changes to the configuration of the production and distribution network are not considered. There is currently no approach that allows *what-if* analyses in response to assortment changes to evaluate the cost effects of concrete changes in detail.

Against this background, this work defines a new model for production and distribution networks to represent assortments and assortment scenarios. For its practical usability, algorithms to generate these models from existing ERP system data are developed to facilitate model generation. On the basis of this model, operations to derive scenarios are defined and algorithms to apply such scenario definitions to a model are developed. A cost function to assess the relevant costs for an arbitrary assortment model or scenario is defined. These elements form the basis for the novel approach to use *what-if* analyses quantitatively to assess assortment complexity. Reviewing the analysis process in Figure 2.9, they support the steps 1 and 4. The optimisation problems posed by steps 2 and 3 are treated as separate subproblems.

4.2 Inventory Allocation in Production and Distribution Networks

Considering existing works in the area of inventory allocation, the basic model of Simpson Jr.[1] is identified as a suitable approach that can be adapted and extended for use in this work. The assumptions made coincide well with the actual practice of MRP-based material planning in the production and distribution networks under consideration. These commonalities comprise a similar network structure and service times as the main scheduling parameter between adjacent stages. The assumption of operating flexibility to guarantee constant replenishment lead times can also be justified.

On this basis, the existing model has to be adapted to accurately fulfil all differing requirements. These comprise a different demand representation, where individual demand volumes and consequently safety stock levels are considered for each planning period. This eliminates the need to assume strictly normally distributed demands, which cannot be expected to exist in practical applications. Further extensions are required to handle different service level measures like the β-service level as well as more complex network structures. In particular, networks cannot only be assumed to be linear, divergent or convergent, but have to allow arbitrary non-cyclic structures to represent arbitrary assortments.

1 See Section 3.2.2.2.

It has to be pointed out that there is no known approach that is computationally feasible in large networks like those we have to expect when considering the inventory allocation problem at the product level. Given the results of Chapter 3.2, the use of heuristic approaches to determine optimal inventory allocations in real-world scenarios is necessary and has proven to be promising. This is especially true if strict optimality is not the primary goal, as is the case here, where the resulting parameters rather serve as a basis for a cost analysis. The task therefore is to define a heuristic solution method that is able to make use of domain knowledge to find reasonable and near-optimal solutions in acceptable time. The considerations in Section 3.2.3 indicate that tabu search is a proven meta-heuristic where domain knowledge can easily be incorporated into the definition and selection of moves.

4.3 Determination of Planning Buffers and Planned Production Quantities

The subproblem of setting planning buffers in MRP manufacturing environments has not yet been addressed to the best of our knowledge. While several works treat the optimal determination of different MRP parameters, none addresses any parameter similar to the planning buffers described in Section 2.1.2. By contrast, the determination of planned production quantities shows certain similarities to known big-bucket lot-sizing problems. Against the background of the current state of the art, the task of determining optimal planning buffers and planned production quantities remains as described in Section 2.3.3. The trade-offs described in that section may be modelled as an optimisation problem, which should use elements of existing big-bucket lot-sizing models where possible. These models must be required to include all relevant cost components and especially integrate the planning buffers and their inventory cost effects as additional elements. Despite these profound extensions, the overall aim is to guarantee the solvability of the model with standard optimisation software.

CHAPTER 5

Configuration of Inventory Management and Production Planning Parameters to Assess the Effects of Assortment Complexity in Consumer Goods Supply Chains

I keep up with the racing rats
and do my best to win.
Editors

5.1 A Model to Assess the Effects of Assortment Complexity Changes

To assess the effects of assortment complexity changes, Section 5.1.1 first defines an assortment model on which to base the cost assessment. On the basis of this model, a formalism to define assortment scenarios is derived in Section 5.1.2. A cost model for the actual assessment is then defined in Section 5.1.3.

5.1.1 Production and Distribution Networks as a Model for Assortment Complexity

This section defines the common basic model of a production and distribution network that serves as the basis for all subordinate optimisation and cost models. Accordingly, only the common elements are introduced here, while additional elements are introduced in later sections along with the optimisation or evaluation problems where they are used. As a convention, we use calligraphic letters for sets, while the corresponding capital letters denote the cardinality of the set. General variables and parameters are denoted both by upper or lowercase letters. Superscripts are used to further define the variable or parameter as required.

5.1.1.1 Model Elements

The definition of model elements is structured according to the categories product and distribution structure, time, internal processes (material transformation and transport) and external processes (customer demand).

Product and distribution structure An assortment defines a certain product and distribution structure which we represent in a network model. Table 5.1 summarises the notations used for the structural elements of that network model.

Table 5.1: Product and distribution structure model elements

Symbol	Description
$\mathcal{L}$	set of physical locations of production sites and sales locations
$\mathcal{M}$	set of materials
$\mathcal{N}$	set of *items* each representing a material at a certain location
$\mathcal{V}$	set of links connecting two items, $\mathcal{V} \subset \mathcal{N} \times \mathcal{N}$
$w_{(i,j)}$	quantity relationship for material flow from i to j
$\mathcal{N}^{PROC}$	set of first-stage items procured externally
$\mathcal{N}^{PROD}$	set of intermediate-stage production items
$\mathcal{N}^{DIST}$	set of intermediate or last-stage distribution items
$\mathcal{SP}$	set of stockpoint items $\mathcal{SP} \subseteq \mathcal{N}$ where inventory is held

The set of locations is denoted $\mathcal{L}$. A location $l \in \mathcal{L}$ refers to the physical location and organisational unit of a production or distribution site of one of the supply chain actors. According to the function of the organisational unit, we distinguish production and sales locations. We only distinguish physical locations at the level of different production or sales sites, while different storage locations within the same plant or warehouse are considered the same physical location.

The set of materials is denoted $\mathcal{M}$. A material $m \in \mathcal{M}$ is goods at any production or distribution stage and can be procured, produced, consumed in the production process or be distributed. Thus we can distinguish between finished, semi-finished, raw and packaging materials. Finished materials are also called *end products* and describe the subset of materials that are not processed any further and sold to customers. Semi-finished materials are all intermediate goods produced in the production process of the end products. Raw and packaging materials are used to produce the semi-finished and sometimes finished materials and are procured externally.

The central element of model are the *items* $\mathcal{N}$, which are valid combinations of materials at their respective physical locations, and thus $\mathcal{N} \subset \mathcal{M} \times \mathcal{L}$. Valid here means that the material is considered for planning at the respective location. An item $i = (p, l)$ thus indicates that product $p \in \mathcal{M}$ is procured or produced at or distributed to location $l \in \mathcal{L}$. The distinction between materials $\mathcal{M}$ and items $\mathcal{N}$ is necessary as the resulting model should represent both the assortment and distribution structure and so each material may be relevant in more than one location.

The set $\mathcal{V} \subset \mathcal{N} \times \mathcal{N}$ represents the links between pairs of items $i, j \in \mathcal{N}$ to represent the product and distribution structure. Such a link $(i, j) \in \mathcal{V}$ may either represent a where-used relation as defined in an entry of j's bill of material (BOM), or a distribution relation between two locations. In the first case, the locations of i and j are identical and their products differ. Item i then is a component in the bill of material of j. In the latter case, demand for j is filled by ordering the required quantities from i, thus the items' locations differ while their products are identical and there is a transport relation between the two locations. The weight $w_{(i,j)}$ represents the quantity relation for material flows from i to j. Depending on whether the link $v = (i, j)$ represents a where-used or distribution relation, w_v is the production coefficient that describes the quantity of i required to produce one basic unit of j, or it is 1, respectively.

In order to achieve the change of either material characteristics or physical locations between adjacent items, each item $i \in \mathcal{N}$ may have some type of process related to it. These processes may either be a production process step that transforms a set of input components into a resulting product, or a transport process that transfers a material from one location to another. More details on the characteristics of these processes are given in Section 5.1.1.1.

With this definition of a network structure and the processes attached to single items, the set of items $\mathcal{N}$ can be further classified into disjunctive subsets. Firstly, set $\mathcal{N}^{PROC}$ contains all items that are procured externally and thereby mark the system boundary at one side of the network. All products and related production and transport processes further upstream in the supply chain are not considered in the model. Consequently, nodes $i \in \mathcal{N}^{PROC}$ have no predecessors in the network representation, which we write as $PR(i) = \emptyset$. Any practical analysis will focus on a limited part of the production and distribution network and consider some materials externally procured. Although it is theoretically possible to extend the model to all suppliers and sub-suppliers, such a limitation may be required to limit model complexity or due to information availability about the structures and processes at the suppliers.

Secondly, the set $\mathcal{N}^{PROD}$ contains all intermediate stage items that have a production process related to them. Such a node can have an arbitrary number greater or equal to 1 of predecessors and successors.

Thirdly, the set $\mathcal{N}^{DIST}$ contains all items with distribution processes attached. With respect to their position in the network, it can only be assured that all last-stage items i with no successors $SC(i) = \emptyset$ are elements of $\mathcal{N}^{DIST}$, while not all elements of $\mathcal{N}^{DIST}$ are necessarily last-stage nodes. Those items with no successors generally represent the end products at the sales locations that are requested by and shipped to customers from there. But there may be distribution structures that span more than one sales location, as not all sales locations necessarily procure their end products directly from the production locations. Moreover, they can procure products from other sales locations if the quantities procured by a sales location do not suffice to organise full truck load transports from the production location. In this case, it may be reasonable to use a nearby sales location as a transshipment point. In any case, we assume that each distribution node $i \in \mathcal{N}^{DIST}$ has exactly one predecessor and thus $|PR(i)| = 1$. This means that the material at a distribution item is *either* procured from the corresponding production location, *or* from another sales location. It is obvious that these three sets form the set of all items $\mathcal{N} = \mathcal{N}^{PROC} \cup \mathcal{N}^{PROD} \cup \mathcal{N}^{DIST}$.

A given set of materials and corresponding links suffices to represent a certain assortment with its product and distribution structure. This network forms the basis for the PDN. Each item $i \in \mathcal{N}$ is a potential stockpoint, which can be interpreted as the decision to keep inventory of the respective item at the respective physical location to uncouple the supply and demand processes.[1] The set of all stockpoints is denoted $\mathcal{SP} \subset \mathcal{N}$.

Time In order to describe any sequence of events (i.e. state changes) in a system, a time model is required.[2] The analyses carried out in this work must always refer to a

1 Hopp defines stockpoints as *"locations in the supply chain where inventories are held"* (Hopp, 2006, p. 2). This definition describes stockpoints on the higher aggregation level of physical location that does not allow to distinguish *which* materials and products are held in stock at these locations.

2 See Dangelmaier (2003, p. 224).

certain period of time, as they cannot use information from an infinite past nor can they extend endlessly to the future. All assumptions made and all conclusions derived refer to a certain period of time that has a sound interpretation in the real world. Time can be modelled as continuous or discrete. As we do not consider states of the system at a single moment, but only make statements about events in certain time intervals[3], we consider discrete time periods of equal length. The set of all discrete time periods is denoted $\mathcal{T} = \{1, \ldots, T\}$. The analysis then always refers to the time interval of lengths T.

Table 5.2: Time model elements

Symbol	Description
$\mathcal{T}$	set of mid-term time periods under consideration $\mathcal{T} = \{1, \ldots, T\}$
T^S	number of short-term periods that constitute one mid-term period in their interpretation in the real world

In order to discretise a time interval, one has to determine the lengths of the single time periods $t \in \mathcal{T}$. For the application considered here we find that some operations like demand forecasting and tracing are made on a timely aggregated level, while planning parameters like planning buffers and replenishment lead times are defined at a more detailed level, i.e. in units of smaller time periods. The discrete time periods $t \in \mathcal{T}$ refer to the aggregated level and represent the *mid-term periods* used to forecast and trace demand. In order to be able to relate the *short-term periods* used to define planning parameters to the mid-term periods, the parameter T^Sdefines the number of short-term planning periods within one mid-term period. That is, T^Sshort-term periods describe the same time span as one mid-term period $t \in \mathcal{T}$ in the real world.[4]

Internal processes: Material transformation and transport In order to describe the processes that transform adjacent items in a network, we define the set $\mathcal{S}$ of production and transport processes. Each production process step $s \in \mathcal{S}^{PROD} \subseteq \mathcal{S}$ describes an arbitrary set of consecutive operations required to produce any related material from its input components. In flow production systems such a process step might describe

3 For example demand in a certain month, production quantities in single shifts etc.

4 In practice, e.g. months or weeks may be chosen as the mid-term period, while all operational parameters are defined in days as the short-term periods. The set of all time periods $\mathcal{T} = \{1, \ldots, 12\}$ may then describe a year and the number of short-term periods per mid-term period would be $T^S = 7$ or $T^S = 30$ respectively. If there are no operations on weekend days, T^Smust be adjusted accordingly.

Table 5.3: Material transformation, transport and coordination model elements

Symbol	Description
$\mathcal{S}$	set of process steps
$\mathcal{S}^{PROD}$	set of production process steps $\mathcal{S}^{PROD} \subseteq \mathcal{S}$ to produce items in $\mathcal{N}^{PROD}$
$\mathcal{S}^{TRANS}$	set of transport process steps $\mathcal{S}^{TRANS} \subseteq \mathcal{S}$ to distribute items in $\mathcal{N}^{DIST}$
$\mathcal{N}_s$	set of items processed on $s \in \mathcal{S}$
K_s	capacity available for production process step $s \in \mathcal{S}^{PROD}$ in one short-term period
$k_{i,s}$	fraction of K_s required to produce one basic unit of $mat(i)$, $i \in \mathcal{N}_s$ on s
$Q_{i,t}$	planned production quantity of i in time period t
$Q^{rnd}_{i,s}$	lot-size rounding value for production quantities of i
pb_s	planning buffer for production scheduling on $s \in \mathcal{S}^{PROD}$

processing on a particular production line. We assume that the set of resources required by different production process steps are disjunctive, such that each production process step can be assigned an independent capacity. In flow production systems the bottleneck resources are the production lines. Separate lines with identical processing capabilities are modelled as separate production process steps. Additional machines required for pre or postprocessing on more than one line should not be any bottleneck resources, so that the assumption of disjunctive resource requirements can be justified.

The processing of an item $i \in \mathcal{N}^{PROD}$ on a production process step $s \in \mathcal{S}^{PROD}$ is further defined in terms of capacities. The capacity available on a production process step $s \in \mathcal{S}^{PROD}$ in one short-term period is denoted K_s, defined in the common basic unit of measure for all items assigned to s. While the units of measure used to quantify production quantities usually differ between production process steps, it can be assumed that the different items processed on one production process step can be measured with the same unit, e.g. pieces, kilogrammes, metres etc. The total capacity in one midterm period is therefore $K_s \cdot T^S$. The capacity coefficient $k_{i,s}$ of the mapping from i to s provides information about the capacities required to process i and is defined as the fraction of K_s required to process one unit of i.

In order to fulfil demands for production items, these quantities have to be produced at a given time. Accordingly, planned production quantities $Q_{i,s,t}$ are defined for each item $i \in \mathcal{N}^{PROD}$ and related production process step s. These planned production quantities are an assignment of demand quantities to a production process step in a certain period, which is either the period in which the demand occurs or any earlier period. These quantities are defined at the aggregate level of mid-term periods, as this is the aggregation level on which demand information is available. Keeping in mind that this model is used to asses cost for real or theoretical assortments, we cannot assume that demand information in terms of concrete order dates and quantities is available, but only at such an aggregate level. The definition of planned production quantities for all demand quantities describes a theoretical production plan at the aggregate level of the time periods $t \in \mathcal{T}$. The problem of assessing operational production cost without any knowledge of the operational production plan is addressed in Section 5.1.3.

Each production process step $s \in \mathcal{S}^{PROD}$ has a planning buffer[5] pb_s, defined as the time buffer between the provision of components and the requirement date of an order, measured in short-term periods.

A transport process step $s \in \mathcal{S}^{TRANS} \subseteq \mathcal{S}$ is characterised by its start and end locations as well as a transport time t_s^{trans}, measured in short-term periods. The required resources of means of transport are not modelled explicitly and are assumed to have unlimited capacity, as they are usually provided by third-party logistic service providers. Additional capacity can therefore be ordered from these service providers at the aggregate planning level considered here.

Each item i is that is not a procurement item is assigned to at least one process step. If the item is a production item $i \in \mathcal{N}^{PROD}$, this process is a production process step $s \in \mathcal{S}^{PROD}$, while it is a transport process step $s \in \mathcal{S}^{TRANS}$ if $i \in \mathcal{N}^{DIST}$. Production items can be assigned more than one production process step as there may be alternative routings that allow processing an item on various machines.[6] Distribution items $i \in \mathcal{N}^{DIST}$ are assigned exactly one transport process step $s \in \mathcal{S}^{TRANS}$.

The relations between the items and the process steps are modelled via binary indicator variables $p_{i,s}$, which is 1 if item i is assigned to process step s and 0 otherwise. We can then define the set of all processes assigned to one item as

$$\mathcal{S}_i = \{s \in \mathcal{S} \mid p_{i,s} = 1\}, \tag{5.1}$$

5 See Section 2.1.2 for a definition and description of planning buffers.

6 Alternative routings sometimes require alternative BOMs, as different machines require slightly different input components. However, these differences are negligible in this context and therefore only one BOM per item is considered.

and analogously the set of all items assigned to one process step as

$$\mathcal{N}_s = \{i \in \mathcal{N} \mid p_{i,s} = 1\} . \tag{5.2}$$

In this way, each node can be assigned to a subset of available processes. For model consistency, we require that

- $\mathcal{S}_i \subseteq \mathcal{S}^{PROD}$ and $|\mathcal{S}_i| \geq 1$ if $i \in \mathcal{N}^{PROD}$: Each production item is mapped to one or more production process steps
- $\mathcal{S}_i \in \mathcal{S}^{TRANS}$ and $|\mathcal{S}_i| = 1$ if $i \in \mathcal{N}^{DIST}$: Each distribution item is mapped to exactly one transport process step.

There are no restrictions in the sets $\mathcal{N}_s$, as an arbitrary number of items can be assigned to all process steps.

In order to describe material flows in the PDN, we need some notion of the duration of activities carried out at the items and their related processes. We define a throughput time TT_i for each item $i \in \mathcal{N}$, which is the time required to carry out all operations of the related process steps, plus supplier lead times for procurement items and any in-plant transports from the physical storage location of all predecessor items to the physical location for production items. Transport times, both of in-plant and inter-location transport, are assumed to include any time required at the destination for goods receipt processing. In total, TT_i is the time required to make the respective materials available for planning at i, given that all predecessors have the respective materials available for planning at their respective locations. Depending on the type of item i, it is defined as

$$TT_i = \begin{cases} \text{supplier delivery time (order lead time)} + \text{transport time} & i \in \mathcal{N}^{PROC} \\ \text{max. in-plant transport of components} + |\mathcal{S}_i|^{-1} \cdot \sum_{s \in \mathcal{S}_i} pb_s & i \in \mathcal{N}^{PROD} \\ t_s^{trans} \text{ with } s \in \mathcal{S}_i & i \in \mathcal{N}^{DIST} \end{cases}$$

For procurement nodes, the throughput time comprises the supplier order lead time as well as the transport time from the supplier to the destination location. As distribution nodes form the system boundary, the suppliers and the respective transport processes are not modelled explicitly and are only included in these throughput times of procurement nodes.

For production nodes the throughput time comprises the maximum in-plant transport time of all components and the average planning buffer of all related production process steps. Taking the average planning buffer introduces some imprecision, as actual throughput times may be both longer and shorter in particular cases. There are some possible alternatives: Firstly, one could use the maximum planning buffer of all related processes $\max_{s \in \mathcal{S}_i}(pb_s)$ instead of the average. This would result in a throughput time that represents the maximum time required to make the materials available for planning, but would overestimate that time in general. Secondly, the planning buffers can be weighted with the probabilities $w_{i,s}$ that process step s will be used to produce a concrete production order of i. While this approach is the most exact, it will generally be difficult to define these probabilities. If the probabilities are unknown, they may be approximated with the ratio of the capacity provided by s over the capacity provided by all related process steps:

$$w_{i,s} = \frac{K_s}{\sum_{p \in \mathcal{S}_i} K_p} \tag{5.3}$$

The precision of this approximation increases as capacity utilisation approaches 100%, as other factors like different machine cost rates then become negligible and the distribution of the production volumes over the production process steps follows the distribution of capacities.

For distribution nodes, the throughput time only comprises the transport time of the related transport process step.

Table 5.4: Customer demand model elements

Symbol	Description
$d_{i,t}^{ext}$	expected external market demand (primary demand) for i in period t
$d_{i,t}$	expected total demand for i in period t
FD_i^{mad}	relative mean absolute deviation of forecasts for i
$D_{i,t}$	random variable describing the actual demand for i in period t
$\sigma_{i,t}^{d}$	standard deviation of forecast deviation distribution of total demand for i in period t
$\sigma_{i,t}^{dext}$	standard deviation of forecast deviation distribution of primary demand for i in period t
α_i^{SL}	α-service level required by customers for i
β_i^{SL}	β-service level required by customers for i
ST_i^{max}	maximum delivery time for i if i represents an end product sold to customers

External factors: Material supply and customer demand Customer demand is the main external factor considered in the model. Both the cost assessment as well as optimisation problems described in Section 2.3 require information on the demand volumes and distribution at each item. We assume that information about the demands $D_{i,t}$ for item i is available at the aggregation level of the mid-term periods $t \in \mathcal{T}$. This demand information may be derived both from historical data or assumptions about theoretical demand developments. Practically, historical data can serve as a starting point to define realistic demand quantities for each item.

For end products there is primary demand $d_{i,t}^{ext}$ that originates from the marked demand in terms of customer orders, for which no explicit model element exists. Furthermore, demand at an item may be dependent demand originating from successor items within the network. In most cases, we can expect that an item is *either* an end product at a sales location that faces primary demand *or* represents an arbitrary material at a production location, which faces dependent demand only from its successors, be it from the adjacent distribution items for end products or from the next production step in case of semi-finished and raw materials. However, we cannot explicitly assume that, as there may be cases where both types of demand are present, e.g. when a sales location supplies another sales location that does not procure its materials directly from the production location, of if semi-finished products are also sold on the marked.

As the expected demand $d_{i,t}$ for an item i is based on forecasts, a probabilistic model of the uncertainty related to the forecast is required. Demand is forecasted for the mid-term periods, i.e. for each period $t \in \mathcal{T}$, demand is forecasted in period $t-1$. If we consider the forecast error $FD_{i,t}$, i.e. the deviation of forecasted from actual demand for item i in period t as a random variable, we can describe the uncertainty via the probability distribution of this variable. While there are no assumptions about the distribution of the actual demand over the mid-term periods, we do assume that the forecast error is normally distributed $FD_{i,t} \sim N(0, \sigma_{i,t}^d)$ with mean 0 and standard deviation $\sigma_{i,t}^d$. The zero mean is reasonable as the long-term error should not be biased to any side. This assumptions holds for most statistical forecasting methods and is also reasonable if the forecasts are made or at least corrected by human planners, as they should neither over nor underestimate the demand on the long-term average. The actual measure of demand uncertainty is the standard deviation $\sigma_{i,t}^d$, which indicates how the forecast errors are distributed around their zero mean. If an item faces primary demand, we denote the corresponding forecast error standard deviation with $\sigma_{i,t}^{dext}$.

With the expected demands $d_{i,t}$ and a model of their variation, the actual demand in a single period $t \in \mathcal{T}$ can be considered a random variable $D_{i,t}$, whose realisations consist of the deterministic expected demand $d_{i,t}$ and an error distributed according to $FD_{i,t}$. Consequently, $D_{i,t}$ also follows a normal distribution $D_{i,t} \sim N(d_{i,t}, \sigma^d_{i,t})$ with mean $d_{i,t}$ and standard deviation $\sigma^d_{i,t}$. It is important to note that this model does not assume that all $D_{i,t}$ are identically distributed for different t. The actual demand is normally distributed *within in each period*, which is why the corresponding random variable $D_{i,t}$ is indexed over the periods $t \in \mathcal{T}$. The sequence of demands for one item i over all periods is therefore a stochastic process D_t with parameter $t = (1, \ldots, T)$, rather than a single random variable that follows one particular distribution.

Production and distribution network model With all the model elements defined in the preceding sections, we can define the production and distribution network model as $PDN(\mathcal{N}, \mathcal{V}, \mathcal{S})$[7]. Figure 5.1 shows a small example of the structure of a PDN.

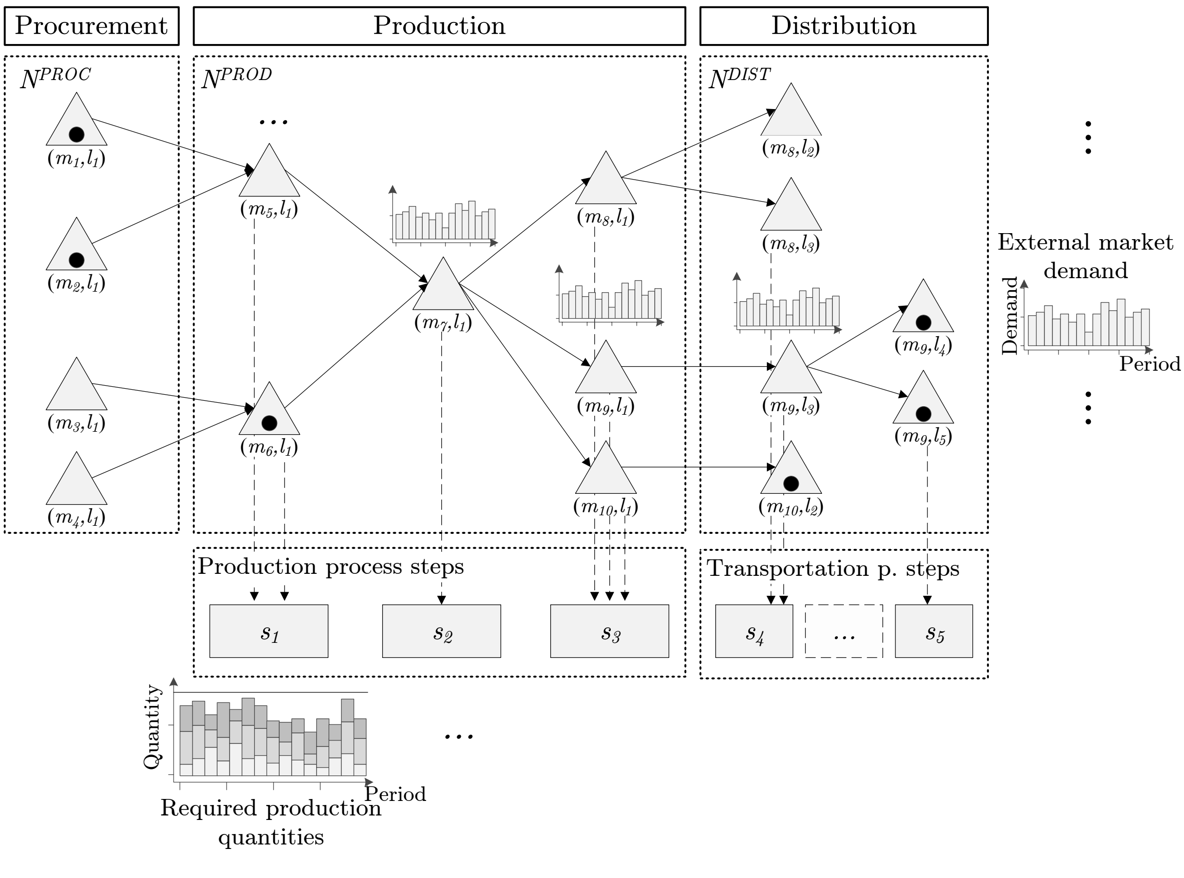


Figure 5.1: Production and distribution network example

This simplified example shows the production and distribution structure for 3 end products $\{m_8, m_9, m_{10}\}$ which are all produced at the production site l_1 and distributed to

7 We shall use $PDN(\cdot)$ as a shorthand for this full notation.

the sales locations l_2, l_3, l_4 and l_5. The two last mentioned locations are not supplied directly from the production site but procure m_9 via l_3 as a transhipment point. All items are either procurement, production or distribution items. All demands originate from the external market demands at the sales locations and are propagated to all items of the network. Each production item is assigned to one production process step (e.g. $\mathcal{S}_{(m_6,l_1)} = \{s_1\}$), while each distribution items is assigned to one transport process step (e.g. $\mathcal{S}_{(m_8,l_3)} = \mathcal{S}_{(m_9,l_3)} = \{s_4\}$). For one particular production process step like s_1, the required production quantities for each period can be calculated by aggregating the demands of all related items $\mathcal{N}_{s_1} = \{(m_5, l_1), (m_6, l_1)\}$. In this example the set of stock-point is selected to be $\mathcal{SP} = \{(m_1, l_1), (m_2, l_1), (m_6, l_1), (m_9, l_4), (m_9, l_5), (m_{10}, l_2)\}$.

5.1.1.2 Model Generation

Considering the intended usage of the model, it is obvious that the representation of realistic assortments results in large networks that easily contain hundreds or thousands of nodes. For its practical application it is desirable that model generation is at least semi-automatic. The widespread use of ERP systems makes it reasonable to use the available data to generate the network structures with less effort. This section shows how parts of the PDN models can be built from data available in standard ERP systems. We firstly show how the network structure can be generated and secondly how demand and forecast deviation data can be calculated consistently over the network.

Network structure The following list defines the required input data that can be obtained from most ERP systems.

Bills of material Each item $i \in \mathcal{N}^{PROD}$ that is produced at location $l \in \mathcal{L}$ has a BOM $\mathcal{B}_i$ that defines a set of a of input components and the respective quantities required of each component j to produce one basic unit of i. A BOM entry $b \in \mathcal{B}_i$ thus is a tuple (c_1, q_1) defining that $c_1 \in \mathcal{M}$ is a required input component for i and that q_1 basic units of c_1 are required to produce one basic unit of i.

Distribution relations Each material $i \in \mathcal{N}^{DIST}$ defines a location $src(i) = l \in \mathcal{L}$ as its procurement source. This may be interpreted as the production or distribution site that supplies i's location $loc(i)$ with the respective material $mat(i)$. Such distribution relations are usually part of the masterdata maintained for each material to allow the automatic generation of purchase proposals by the ERP system.

Routings A routing for $i \in \mathcal{N}^{PROD}$ defines a set of operations carried out on different work centres to produce the respective material.

As described in Algorithm 5.1, the structure of the production and distribution network can be derived from the documents mentioned above. As input parameters, the algorithm requires a set of starting materials $\mathcal{M}^* \subset \mathcal{M}$ that contains only end products and thus finished materials. The selection of this set determines for which part of the assortment the PDN is built. For practical applications this allows invocation of the generation of an assortment model e.g. for a certain product category by simply providing a list of the corresponding end product material numbers as an input.

Furthermore, it requires that all processes, i.e. production and transport process steps, are defined a priori. A production process step is not necessarily a single physical production resource and thus may comprise several work centres. If the routings from an ERP system are used to map the items to the production process steps, the production process steps have to be defined such that each $s \in \mathcal{S}^{PROD}$ comprises all work centres required by the operations in each routing for an item i if $s \in \mathcal{S}_i$. This implies that for each routing of i, there is at most one $s \in \mathcal{S}^{PROD}$ assigned to i. This correspondence of routings to production process steps guarantees that if multiple production process steps are assigned to an item, they always represent alternative routings and thus alternative production possibilities.

The set $\overline{\mathcal{N}}$ contains all items that still have to be processed, i.e. added to the network and possibly expanded if there are any supplying items or component predecessors to be defined. Initially, this set is filled with all items that represent any material from $\mathcal{M}^*$ at all the locations (lines 3 to 6) where it is produced, distributed to and / or sold: $\overline{\mathcal{N}} = \{i \in \mathcal{N} \mid mat(i) \in \mathcal{M}^*\}$. All items in this set are then iteratively processed (line 7). Each item is added to the network and then handled differently depending on the type of item encountered.

For distribution items (lines 10 to 18), a new item with the same material at the supplying plant is added to the set of nodes and to the set of items to process if necessary. The link between the two items is added with a weighting of one, as the materials are not altered during the transport process. The process mapping is extended by the mapping of i to the respective transport process.

For production items (lines 19 to 30), i's BOM is exploded and for each of its components, a new item with the material at the same production location is added to the set of nodes and to the set of items to process if necessary. The corresponding links

Algorithm 5.1: Generate a production and distribution network

Input: $\mathcal{M}^*$, $\mathcal{S}$
Result: $PDN(\mathcal{N}, \mathcal{V}, \mathcal{S})$

$\mathcal{N} \leftarrow \emptyset$
$\overline{\mathcal{N}} \leftarrow \emptyset$
foreach $m \in \mathcal{M}^*$ **do**
    foreach *location l where m is defined* **do**
        $\overline{\mathcal{N}} \leftarrow \overline{\mathcal{N}} \cup (m, l)$

while $\overline{\mathcal{N}} \neq \emptyset$ **do**
    select $i \in \overline{\mathcal{N}}$
    $\mathcal{N} \leftarrow \mathcal{N} \cup i$
    if *i is procured from a production location* **then**
        $j \leftarrow (mat(i), src(i))$
        if $j \notin \mathcal{N}$ **then**
            $\mathcal{N} \leftarrow \mathcal{N} \cup j$
            $\overline{\mathcal{N}} \leftarrow \overline{\mathcal{N}} \cup j$
        $\mathcal{V} \leftarrow \mathcal{V} \cup (j, i)$
        $w_{(j,i)} \leftarrow 1$
        get $s \in \mathcal{S}^{TRANS}$ that represents transport from $loc(j)$ to $loc(i)$
        $p_{i,s} \leftarrow 1$
    if *i is produced at l* **then**
        foreach *BOM entry $(c, q) \in \mathcal{B}_i$* **do**
            $j \leftarrow (c, l)$
            if $j \notin \mathcal{N}$ **then**
                $\mathcal{N} \leftarrow \mathcal{N} \cup j$
                $\overline{\mathcal{N}} \leftarrow \overline{\mathcal{N}} \cup j$
            $\mathcal{V} \leftarrow \mathcal{V} \cup (j, i)$
            $w_{(j,i)} \leftarrow q$
        foreach *routing r defined for i* **do**
            get $s \in \mathcal{S}^{PROD}$ that comprises all work centres referenced in r
            $p_{i,s} \leftarrow 1$

    if *i is procured from a external supplier* **then**
        `// nothing to do`
    remove i from $\overline{\mathcal{N}}$

between the component items and i are added and weighted with the production coefficients from the corresponding BOM entry. Finally, the routings defined for i are used to determine the set of production process steps to be assigned to i via the indicators $p_{i,s}$.[8]

For procurement items (lines 31) no changes have to be made to the network since the network is constructed *upstream*, i.e. for each item predecessors are determined and the corresponding links are created. Procurement items do not have any predecessors and are connected to their successors during the processing of these successors, which by definition are either production or distribution items.

After an item has been processed, it is removed from the set of items to be processed (line 32). In this way the set of items to be processed will eventually be empty, as each successive explosion of BOMs will result in procurement items that are not further exploded and will only be removed from $\overline{\mathcal{N}}$. This guarantees the termination of the algorithm due to the loop's exit condition $\overline{\mathcal{N}} \neq \emptyset$.

Demands and forecast deviations In addition to the network structure, the demand and forecast data of each item usually can be obtained from an ERP system. Clearly, all this data can also be defined "manually" based on assumptions about the development of customer demand and forecast accuracy for each period. However, as this requires huge effort[9], we propose an alternative method that uses historical data to automatically derive these assumptions. It takes into account two practical considerations:

1. If demand information is to be based on historical data, it can only be based on the primary demands. Dependent demands, which mainly occur for end products at production locations and semi- as well as raw materials within the production stages are influenced by historical planning decisions. Orders from sales locations are aggregated to obtain full truck load transports, and the material flow in the production area is influenced by lot-sizing decisions.

2. The same applies to information about forecast deviations. Forecast deviations are only traced for those items that face primary demands, as the forecasts are only made for these items and then propagated to all predecessors in the material requirements planning process.

8 By definition there is only one production process step required to perform the operations defined in one routing. See Section 5.1.1.1 for the definition of production process steps.

9 For example, consider a network with 500 items, for which an analysis should be made over a period of one year. If we select months as the mid-term periods, a total of $500 \cdot 12 = 6000$ values as demand estimations have to be defined.

Due to these observations, we propose a method to calculate demands and forecast deviations *only* from information about primary demands $d_{i,t}^{ext}$ and the related forecast deviations. Both demand and forecast deviations are propagated upstream the network, aggregating at each item the potential primary demand, where applicable, and the dependent demands of its successors. With this logic, the $d_{i,t}$ are calculated as

$$d_{i,t} = d_{i,t}^{ext} + \sum_{j \in SC(i)} w_{(i,j)} \cdot d_{j,t} \tag{5.4}$$

While it may be reasonable to use historical demand data in terms of the expected demands in each period, this is unreasonable for stochastic values like the forecast deviations. As we consider the demand deviations $FD_{i,t}$ in each period to be separate random variables, we cannot assume that these single distributions are known. Moreover, we want to derive them from an *average* forecast error observed over the entire time horizon T. We therefore assume that a long-term measure FD_i^{mad} for the relative mean absolute forecast deviation is available for each item that faces primary demand. This measure is defined as the long-term average of

$$\frac{|\text{actual primary demand in period } t - \text{primary demand forecasted in period } (t-1)|}{\text{actual primary demand in period } t} \tag{5.5}$$

This is a much more intuitive and commonly-used long-term measure of the forecast deviations, as it expresses the long-term mean deviation as a percentage of demand.[10]

As defined in Section 5.1.1.1, the demand for an item i is represented as a series of random variables $D_{i,t} \sim N(d_{i,t}, \sigma_{i,t}^d)$ over the periods $t \in \mathcal{T}$. As expected demand can be calculated as described above, the remaining question is to derive the standard deviations $\sigma_{i,t}^d$ for a normally distributed forecast error from the relative mean absolute error FD_i^{mad}. For a normally distributed random variable $X \sim N(\mu, \sigma)$, there is a constant relation between the mean absolute deviation $\delta = \sum_i |x_i - \mu|$ of realisations x_i of X from its mean and the standard deviation σ. Burrows and Talbot[11] show that

$$\delta = \sigma \sqrt{\frac{2}{\pi}} \tag{5.6}$$

and consequently the standard deviation can be obtained from the mean absolute deviation as

$$\sigma = \delta \sqrt{\frac{\pi}{2}} \tag{5.7}$$

10 For example, *"forecasts deviate from the actual demand by $FD_i^{mad} = 30\%$ on average"*.

11 Burrows and Talbot (1985, p. 89).

Transferred to our model, we can now calculate the standard deviations of the forecast errors on the basis of the relative mean absolute deviations. As the latter refer to primary demands, we obtain the standard deviations of forecast errors for the external market demands as

$$\sigma_{i,t}^{dext} = \sqrt{\frac{\pi}{2}} \cdot FD_i^{mad} \cdot d_{i,t}^{ext} \tag{5.8}$$

similar to the demand calculation in Equation 5.4, we have to aggregate forecast deviations from primary demands and dependent demands at an item. This aggregation is conducted by forming the convolution of the relevant forecast deviation distributions. For normally distributed random variables, the convolution equals the sum of these variables.[12] As the means of forecast deviation distributions are always 0, we only have to consider the summation of the standard deviations related to the primary and successors' demands:

$$\sigma_{i,t}^{d} = \sqrt{\left(\sigma_{i,t}^{dext}\right)^2 + \sum_{j \in SC(i)} \left(\sigma_{j,t}^{d}\right)^2} \tag{5.9}$$

As expected, Equation 5.9 yields $\sigma_{i,t}^{d} = \sigma_{i,t}^{dext}$ if $SC(i) = \emptyset$, i.e. demand uncertainty only originates from external market demands if an item has no successors.

Algorithm 5.2: Propagating demands and forecast deviations over a PDN

Input: a directed graph $G(\mathcal{N}, \mathcal{V})$ as part of a PDN
Result: updated graph G with demand and forecast information at each $i \in \mathcal{N}$

$L \leftarrow$ `TopologicalSorting(`G`)`
`Invert(`L`)`
foreach *item* i *in* L **do**
  with successors $SC(i)$
  begin
    foreach *period* $t \in \mathcal{T}$ **do**
      calculate $d_{i,t}$ according to Equation 5.4
      calculate $\sigma_{i,t}^{d}$ according to Equation 5.9
      $D_{i,t} \leftarrow N(d_{i,t}, \sigma_{i,t}^{d})$
  end

Algorithm 5.2 shows the propagation of demands and forecast deviations in a PDN via the iteration over the items of the graph $G(\mathcal{N}, \mathcal{V})$. The most important operations

12 The sum $Y = X_1 + X_2$ of two normally distributed random variables X_1 and X_2 with means μ_1 and μ_2 and standard deviations σ_1 and σ_2 is $Y \sim N(\mu_1 + \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2})$.

are the topological sorting[13] and the inversion of the result list in lines 1 and 2, which guarantees that no item is processed before the calculations are finished for all its successors. This is important as the calculation of the $\sigma^d_{i,t}$ for an item i requires that the $\sigma^d_{j,t}$ have already been calculated for all $j \in SC(i)$ (compare Equation 5.9). Note that the results of these calculations are only determined by the external primary demands $d^{ext}_{i,t}$ and the relative mean absolute forecast deviations FD^{mad}_i observed for these demands. This implies that Algorithm 5.2 can also be used to update the demands and forecast deviations at all items if primary demands change, a feature we can also use for the scenario generation, as described in the Section 5.1.2.2.

5.1.2 Alternative Assortment Scenarios

An assortment scenario (or scenario) is an assortment model that is derived from an existing assortment model (called *baseline model*) by changing the assortment is represents. All changes are made via one of the operations

1. Adding new materials and items
2. Replacing and / or discontinuing existing materials
3. Changing demand and forecast deviation information.

Operation 3 is trivial and not described in more detail, as it only changes some parameters in the model. This section describes how scenarios based on changes with operations 1 and 2 are defined and applied to a baseline model.

5.1.2.1 Definition of Alternative Assortment Scenarios

The changes made to the assortment of the baseline model are encoded in a scenario definition, which consists of

- $\mathcal{A}$: a set of new item additions each describing a possibly new material at a respective production or sales location
- $\mathcal{R}^{fin}$: a set of material replacement definitions for end products
- $\mathcal{R}^{rs}$: a set of material replacement definitions for raw and semi-finished materials

13 A topological sorting of a directed acyclic graph is a linear ordering of its nodes in which each node comes before all its successors in the graph. Algorithms to compute such orderings are well known and described e.g. by Cormen et al. (2001, pp. 549–551).

The set $\mathcal{A}$ contains item additions, each of which describes an item that is to be added to the network. If the item refers to a material that does not yet exist in the assortment, it first has to be added at a production location and be integrated into the product structure and production process steps. If the corresponding material already exists, the addition defines that it is to be distributed to a new location and the item has to be integrated into the distribution structure. The set of all materials that are added via the elements of $\mathcal{A}$ is denoted $\mathcal{M}^{new}$. In general, each material addition $a = \left(m, l_m, \mathcal{B}_{(m,l_m)}, r_m, src((m,l_m)), d^{ext}_{(m,l_m)}, FD^{mad}_{(m,l_m)}\right) \in \mathcal{A}$ contains the definition of

- m: the material to be added to the assortment
- l_m: the production or distribution location where the material is to be added
- $\mathcal{B}_{(m,l_m)}$: a bill of material if l_m is a production location, otherwise not specified
- r_m: a routing if l_m is a production location, otherwise not specified
- $src((m,l_m))$: a procurement source if l_m is a sales location, otherwise not specified
- $d^{ext}_{(m,l_m),t}$: primary demands for m at location l_m for each $t \in \mathcal{T}$
- $FD^{mad}_{(m,l_m)}$: relative mean absolute forecast deviation for m at l_m

For the definition we have to distinguish between new materials at production and sales locations. For production locations, a BOM has to be specified and the primary demands and related forecast deviations are likely all to be 0. New materials at sales locations require the definition of a procurement source and will probably have positive demands and forecast deviations. The definition of material additions at sales locations is only required if the material really is an *additional* material in the assortment and does not serve as a replacement for any other material. In the latter case, it is sufficient to define one material addition for the new material at the production location and define appropriate material replacements that affect the products at the sales locations, as described below.

The set $\mathcal{R}^{fin} \subset \mathcal{M} \times (\mathcal{M} \cup \mathcal{M}^{new})^n \cup \circ \times \mathbb{R}^n \cup \circ \times \mathbb{R}^n \cup \circ$ contains material replacement definitions for end products, each of which specifies how a finished material should be partly replaced and possibly discontinued in the scenario. Such a replacement definition $r = (m_r, M^{rep}_r, D^{ratio}_r, CV_r)$ is given via

- $m_r \in \mathcal{M}$: finished material to be replaced and possibly discontinued
- M^{rep}_r: tuple of replacement materials
- D^{ratio}_r: tuple of replacement material demand ratios

- CV_r: tuple of replacement materials conversion factors

An end product can be discontinued without replacement, indicated as $M_r^{rep} = \circ$, or it can be replaced by $n \geq 1$ replacement materials $M_r^{rep} = \left(m_1^{repl}, \ldots, m_n^{repl}\right)$. The replacement materials can either be existing products or new products defined in $\mathcal{M}^{new}$, which leads to $M_r^{rep} \in (\mathcal{M} \cup \mathcal{M}^{new})^n \cup \circ$. The placeholder element $\circ$ indicates that there is no replacement and the corresponding product is just removed from the assortment. If $n \geq 1$, the tuple $D_r^{ratio} = (d_1^{ratio}, \ldots, d_n^{ratio})$ specifies for each replacement material, what percentage d_i^{ratio} of primary demand of m_r is transferred to each m_i^{repl} and added to its primary demand. These percentages need not sum up to 1 (or 100%) over all replacement materials. If $\sum_i d_i^{ratio} < 1$, the remaining demand ratio is interpreted as lost sales, which means that customers are expected to buy less of the new material. Analogously, if $\sum_i d_i^{ratio} > 1$, the surplus is interpreted as additional demand expected as a result of the replacement. The conversion factors $CV_r = (cv_1, \ldots, cv_n)$ specify how demand quantities of m_r are converted to demand quantities of each m_i^{rep}. Such a conversion factor is $\neq 1$ if the finished products represented by these materials m_r and m_i^{rep} are packed end products with different packaging sizes. This is relevant both for products on the BCU and TSU packaging level.[14]

For end products we do not have to pose any restrictions on the number of replacement definitions per material due to the structure of the subgraphs spanned by all direct and indirect successors $DN(i)$ of any end product item i. By definition, i is not processed any further on any production process step and is only distributed to different locations. Therefore the structure of the sub-network of all items in $DN(i)$ is always strictly divergent[15] for any and end product item i, so that one item can be replaced by multiple items without creating any ambiguities with respect to the material flows represented by that subgraph. Nodes and links are duplicated according to number of replacement definitions.

For all materials that are not end products, a substitution of one material by several replacement materials causes ambiguities with respect to their use as input components on a production process step. If such a material m_r^1 were to be replaced with multiple materials $m_{r_1}^{repl}, m_{r_2}^{repl}$, there would be no sound interpretation of how these materials are used as input components for the production of other materials. For this reason, the set of material replacement definitions $\mathcal{R}^{rs} \subset \mathcal{M} \times \mathcal{M} \cup \mathcal{M}^{new} \times \mathbb{R}$ is more restrictive with respect to the definition of replacements for one material, as it requires that

14 For example, if the materials are both packed at TSU level and the material that is to be replaced is a batch of 20 single BCUs, while the replacement material only contains 10 BCUs, the conversion factor would be 2. See Section 2.2 for a description of different packaging levels.

15 A directed graph is called divergent if each node has at most one predecessor.

there is at most one replacement definition $r = (m_r, m_r^{rep}, cv_r)$ for one particular raw or semi-finished material:

$$m_{r_1} \neq m_{r_2} \forall r_1, r_2 \in \mathcal{R}^{rs} \tag{5.10}$$

In contrast to the case of end products, an indicator for discontinued materials ($\circ$) is not allowed and each replacement definition requires a specific replacement material $m_r^{rep} \in \mathcal{M} \cup \mathcal{M}^{new}$. As demand for raw and semi-finished products is allways dependent demand, no demand has to be transferred from m_r to m_r^{rep} and the d_r^{ratio} can be omitted. The conversion factor cv_r defines how production coefficients of the BOM entries that specify m_r as the input component have to be adapted.

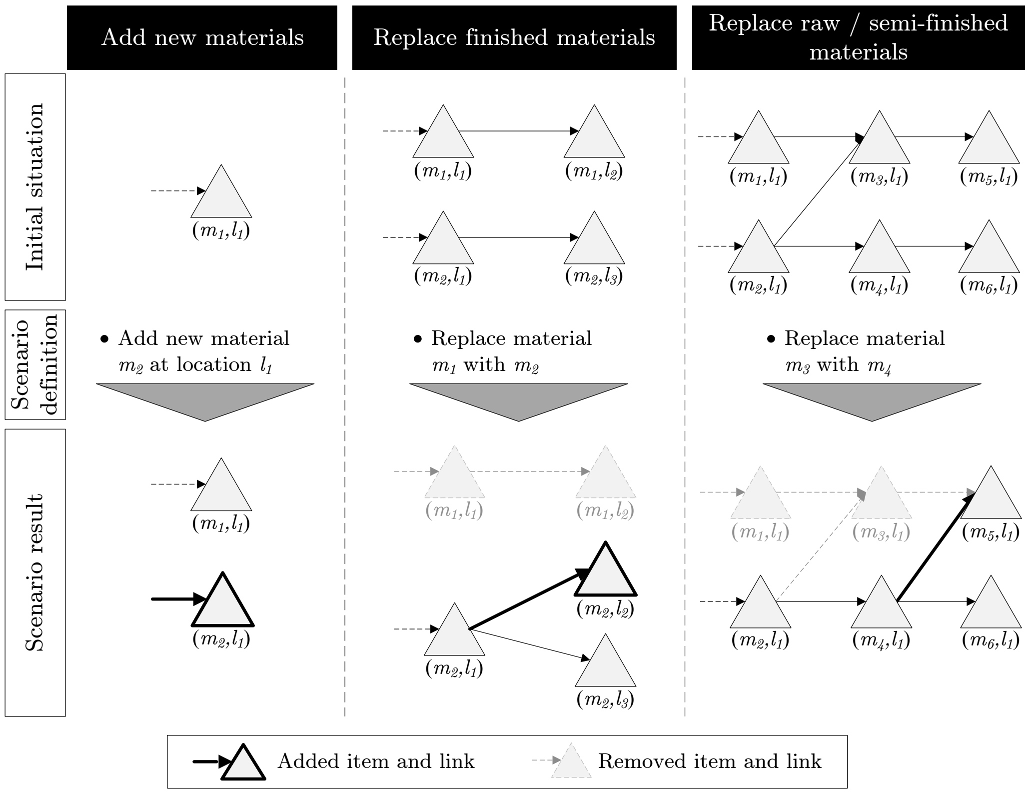


Figure 5.2: Simple examples of scenario definition operations

Figure 5.2 illustrates a simple example of each of the operations that can be used to define a scenario. For the sake of simplicity of illustration the depicted examples are the simplest possible cases, with single replacements and simple distribution structures only.

5.1.2.2 Application to a Baseline Model

To obtain a scenario, a scenario definition is applied to a baseline model. In order to maintain the consistency of the model the sequence in which changes are applied to the baseline model is important and must follow the steps:

1. Add the new materials as defined in $\mathcal{A}$. Assure that those additions with new materials at production locations are made first.
2. Replace end products as defined in $\mathcal{R}^{fin}$
3. Replace raw and semi-finished products as defined in $\mathcal{R}^{rs}$
4. Update all demands and forecast deviations by propagating external (primary) demands and forecast deviations over the graph $G(\mathcal{N}, \mathcal{V})$ according to Algorithm 5.2.

We define the algorithms for each of these steps in the next sections. Step 4 is an exception, as it is already fully defined via Algorithm 5.2 in Section 5.1.1.2.

Adding new materials The addition of new items described in Algorithm 5.3 processes each item addition and first creates the new item i and adds it to the network (lines 1 to 3). For production items, the corresponding BOM $\mathcal{B}_i$ is exploded and the entries are used to determine the predecessor items in the production network and to add the required links (lines 5 to 8). For each routing defined for i, the corresponding production process steps are determined and the mapping of i to these process steps is created (lines 9 to 11). If i is a distribution item, the supplying predecessor is determined and linked to the new item. The items in $\mathcal{A}$ are ordered to process those item additions at production locations first to ensure that the predecessors already exist even if $mat(i) \in \mathcal{M}^{new}$. On the basis of this distribution relation, the corresponding transport process step is assigned to the new item (lines 13 to 16).

Algorithm 5.3 assumes that for new production items, both a BOM as well as at least one routing are defined for each new material. For practical applications, the tedious task of explicitly specifying all this information can be alleviated by deriving each new material from a *template* material. All information required for the assignments of predecessors and production process steps is copied from that material and adapted as required. This approach is especially suitable if the assortment changes represent a standardisation of several items to one generic variant. In this case, the new generic

Algorithm 5.3: Scenario generation: adding new materials

Input: a baseline model $PDN(\mathcal{N}, \mathcal{V}, \mathcal{S})$, a set $\mathcal{A}$
Result: the scenario model $PDN^*(\mathcal{N}, \mathcal{V}, \mathcal{S})$ including the new items defined in $\mathcal{A}$

foreach $a \in \mathcal{A}$ **do**
    create item $i = (m, l_m)$
    $\mathcal{N} \leftarrow \mathcal{N} \cup i$
    if l_m *is a production location* **then**
        foreach *BOM entry* $(c, q) \in \mathcal{B}_i$ **do**
            $j \leftarrow (c, l_r)$
            $\mathcal{V} = \mathcal{V} \cup (j, i)$
            $w_{(j,i)} \leftarrow q$
        foreach *routing* r_m *defined for* i **do**
            select $s \in \mathcal{S}^{PROD}$ that comprises all work centres referenced in r_m
            $p_{i,s} \leftarrow 1$
    else
        $j \leftarrow (m, src(i))$
        $\mathcal{V} \leftarrow \mathcal{V} \cup (j, i)$
        select $s \in \mathcal{S}^{TRANS}$ that represents transport from $loc(j)$ to l_m
        $p_{i,s} \leftarrow 1$

variant can be defined as a new material with all information copied from that existing variant that is most *similar* to it.

Discontinuing or replacing end products Each end product replacement defined in $\mathcal{R}^{fin}$ specifies the replacement for one material and therefore affects all items that represent that material at any location. For each replacement Algorithm 5.4 first processes all items that represent the corresponding material at a sales location. For each such item several replacements can be specified, for each of which the following operations are performed (lines 6 to 19):

1. Add the replacement item and predecessor link to the network as required.
2. Update the demand information at the replacement item.
3. Update the forecast deviation information at the replacement item.
4. Remove the item and related links from the network.

Algorithm 5.4: Scenario generation: replacing end products

Input: a baseline model $PDN(\mathcal{N}, \mathcal{V}, \mathcal{S})$, a set $\mathcal{R}^{fin}$
Result: a scenario model $PDN^{*}(\mathcal{N}, \mathcal{V}, \mathcal{S})$

foreach $r \in \mathcal{R}^{fin}$ **do**
  foreach *Sales location l where m_r is defined* **do**
    $i \leftarrow (m_r, l)$
    if $M_r^{repl} \neq \circ$ **then**
      for $idx = 1, \ldots n$ **do**
        $j \leftarrow (m_{idx}^{repl}, l)$
        if $j \notin \mathcal{N}$ **then**
          $\mathcal{N} \leftarrow \mathcal{N} \cup j$
          $v \leftarrow (PR(i), j)$
          $w_v \leftarrow 1$
          $\mathcal{V} \leftarrow \mathcal{V} \cup v$
          foreach $t \in \mathcal{T}$ **do**
            $d_{j,t}^{ext} \leftarrow d_{i,t}^{ext} \cdot d_{idx}^{ratio} \cdot cv_{idx}$
          $FD_j^{mad} \leftarrow FD_i^{mad}$
        else
          foreach $t \in \mathcal{T}$ **do**
            $\overline{d_{i,t}^{ext}} \leftarrow d_{i,t}^{ext} \cdot d_r^{ratio} \cdot cv_r$
            $d_{j,t}^{ext} \leftarrow \overline{d_{i,t}^{ext}} + d_{j,t}^{ext}$
          $FD_j^{mad} \leftarrow \frac{\sqrt{\left(FD_i^{mad} \cdot \sum_{t \in \mathcal{T}} \overline{d_{i,t}^{ext}}\right)^2 + \left(FD_j^{mad} \cdot \sum_{t \in \mathcal{T}} d_{j,t}^{ext}\right)^2}}{\sum_{t \in \mathcal{T}} \left(\overline{d_{i,t}^{ext}} + d_{j,t}^{ext}\right)}$.
    $\mathcal{N} \leftarrow \mathcal{N} \setminus \{i\}$
    $\mathcal{V} \leftarrow \mathcal{V} \setminus \{(k, i) \in \mathcal{V} | k \in PR(i)\}$
  foreach *Production location l where m_r is defined* **do**
    $k \leftarrow (m_r, l)$
    if $|SC(k)| = 0$ **then**
      `RemoveObsoleteItems(`k`)`

Demands are calculated separately for each period $t \in \mathcal{T}$. The relative mean absolute forecast deviations FD_i^{mad} are used to aggregate forecast deviation information for all periods. As there is a defined relationship between the single primary demand distribution standard deviations σ_t^{dext} and the relative mean absolute forecast deviation FD_i^{mad}, it is sufficient to update the latter value. The single demand distribution standard deviations $\sigma^d_{i,t}$ can then be calculated according to Equations 5.8 and 5.9.

The algorithm distinguishes two cases, depending on whether the replacement item j already exists in the network or not. If it does not exist, the replacement material was not sold at that sales location before and the corresponding item j is added to the network. A distribution link to this item is added, where the source of the link is determined by the predecessor of i and thus the distribution structure of the old product is maintained. If m_r was procured directly from the corresponding production location, m_r^{repl} is procured from the same production location. The case where products are procured via another sales location is handled analogously.[16] The external demands at the new item j directly result from the demands at i, adapted with the demand rate and conversion factor. The mean relative forecast deviation in this case can be taken from i without modifications as it is a percentage value that does not require any quantity adaptations.

If the material was already sold at the sales location, no structural changes have to be made to the network. However, the adaptation of demands and forecast deviations now has to consolidate the information from both i and the existing replacement item j. In the notation of the algorithm, $\overline{d^{ext}_{i,t}}$ holds the demand quantity that is transferred from i to j in period t, considering demand ratios and conversion rates. It is added to j's previous demand to yield the new demand quantity $d^{ext}_{j,t}$. In order to aggregate relative the relative mean absolute forecast deviations one has to take into account that different demand quantities determine the degree to which a certain forecast deviation influences the aggregate deviation. Thus FD^{mad}_i and FD^{mad}_j are aggregated by weighting them with the relevant total demands $\sum_{t\in\mathcal{T}} \overline{d^{ext}_{i,t}}$ and $\sum_{t\in\mathcal{T}} d^{ext}_{j,t}$, respectively (line 19).

Procedure `RemoveObsoleteItems(`k`)`

foreach $j \in PR(k)$ **do**
  if $|SC(j) = 1|$ **then**
    `RemoveObsoleteItems(`j`)`

$\mathcal{N} \leftarrow \mathcal{N} \backslash i$
$\mathcal{V} \leftarrow \mathcal{V} \backslash (j, i)$

After all replacements for an item i have been finished, the item is removed from the network, together with all links that have i as their target (lines 22 to 22). After

16 Note that as we consider distribution items here, they have only one predecessor by definition. See Section 5.1.1.1 on page 76.

all items that represent the material of a given material replacement definition at a sales location have been processed, the material can only be found at a production location. As the material is no longer sold at any sales location, there is no dependent demand at the production location and the item can be removed from the network. If an end product is not produced any more, there may be some semi-finished, raw and packaging materials that were only used for that product and consequently can be removed from the assortment now. Therefore all direct and indirect predecessors of the item in terms of production and procurement items can be removed as long as they do not form input components of any other materials. To achieve this each such item k is removed via the function `RemoveObsoleteItems(`k`)` if it does not have any successors any more. This function then removes k from the network and is called recursively for all predecessors $j \in PR(k)$ if k is their only successor and thus they are not input components to any other materials. The recursive invocation guarantees that all unnecessary items are removed from the assortment.

Algorithm 5.6: Scenario generation: replacing raw and semi-finished products

Input: a baseline model $PDN(\mathcal{N}, \mathcal{V}, \mathcal{S})$, a set $\mathcal{R}^{rs}$
Result: a scenario model $PDN^{*}(\mathcal{N}, \mathcal{V}, \mathcal{S})$

foreach $m_r \in \mathcal{R}^{rs}$ **do**
  foreach *production location l where m_r is defined* **do**
    $i \leftarrow (m_r, l)$
    $j \leftarrow (m_r^{rep})$
    foreach $k \in SC(i)$ **do**
      $\mathcal{V} \leftarrow \mathcal{V} \cup (j, k)$
      $w_{(j,k)} \leftarrow w_{(i,k)} \cdot cv_r$
      $\mathcal{V} \leftarrow \mathcal{V} \backslash (i, k)$
    `RemoveObsoleteItems(`i`)`

Replacing materials in procurement and production stages The replacement of raw and semi-finished products defined in Algorithm 5.6 is similar to the replacement of end products described above. The main differences are simplifications that result from the more restrictive material replacement definitions[17] and the fact that no primary demands exist at the replaced items. In particular, these differences are:

17 See Section 5.1.2.1.

1. No items have to be added to the network, as they either already exist or have been added to the network during the addition of new materials.[18]
2. There is no processing of multiple replacement items per replacement definition.
3. Due to the absence of primary demands, no demands and forecast information has to be transferred between the discontinued item and its replacement. This information will result from the propagation of demands and forecast deviations made when all structural changes have been applied.

For a production item i that is to be replaced with an item j, all links that have i as their source have to be replaced by links with j as their source. Thus any material that had m_r as an input component now has m_r^{rep} in that place. Changes in the quantities required are considered via the conversion factor cv_r and included by adapting the corresponding link weights. Finally, the item i is removed from the network via the `RemoveObsoleteItems(`i`)` function already introduced in the previous section, which guarantees that all direct and indirect predecessors are also removed recursively.

5.1.3 Cost Assessment for a Given Assortment

Based on the model from Section 5.1.1.1, we can define the cost model used to assess the cost incurred by a certain assortment, represented as a production and distribution network. As described in Section 2.3, we distinguish inventory, scrap and setup cost as well as cycle stock costs, such that the total cost C can be written as

$$C = C^{INV} + C^{STSCRP} + C^{CYST} \tag{5.11}$$

Table 5.5 summarises the notation for the cost parameters used in this section.

Each item has a value C_i that represents the internal accounting value of the material $mat(i)$ at location $loc(i)$. This value generally increases with each production or distribution stage, due to the value added over these stages. For a production item $i \in \mathcal{N}^{PROD}$, C_i is at least as large as the sum of values of its input components $C_i \geq \sum_{j \in PR(i)} C_j \cdot w_{j,i}$ and generally greater than this, as the production process itself adds more value to the material and the cost incurred by the production is reflected in C_i. For distribution items $i \in \mathcal{N}^{DIST}$ the cost incurred by the transport is generally reflected in a value increase for the material at the target location $C_i \geq C_j \forall j \in PR(i)$.

18 See Section 5.1.2.2: new items are added to the network at their corresponding production location and connected to their predecessors via explosion of their BOM.

Table 5.5: Cost parameters

Symbol	Description
C_i	internal accounting value for one basic unit of i
$\overline{C}^{inv}$	inventory holding cost rate to account for capital binding over one period $t \in \mathcal{T}$
C_l^{whsg}	warehousing cost for one storage unit at location $l \in \mathcal{L}$ over one period $t \in \mathcal{T}$
qpp_i	number of basic units of i that can be stored on one storage unit at $loc(i)$
$\overline{C}_{i,s,pb_s}^{stp}$	average setup cost incurred by production of i on production process step s with a planning buffer of pb_s
$scrp_{i,s}$	quantity of scrap produced during start and end of a production run of i on s

Inventory costs comprise inventory holding and warehousing costs. The inventory holding cost rate $\overline{C}^{inv}$ is an interest rate used to express the opportunity cost incurred by the capital tied up in inventory over one single period $t \in \mathcal{T}$. Warehousing costs are incurred for the provision of physical space used to store the inventory. They are given for the space required for one storage unit, typically a pallet in the case of consumer goods. In order to break these costs down to one basic unit of a certain item, we denote the number of basic units of i that can be stored on one storage unit qpp_i. This value is maintained on the item level rather than per material, as different locations may use different types and sizes of storage units and thus these values may differ between locations. These costs can be aggregated into a single total inventory cost rate C_i^{inv} for one basic unit of i over one period t:

$$C_i^{inv} = C_i \cdot \overline{C}^{inv} + \frac{C_{loc(i)}^{whsg}}{qpp_i} \tag{5.12}$$

Within a single period inventory levels are not constant. For our cost assessment, it is sufficient to derive average inventory levels per period. According to the requirements from Section 2.3, inventory costs are assessed on the basis of fixed safety stock levels and reorder points per periods, which determine the average inventory level at an item i in a given period t. The reorder point $RP_{i,t}$ is set so that at the moment a replenishment order is placed, the remaining quantity on stock covers the *lead time demand*, i.e. the demand over the replenishment lead time RLT_i.[19] Given that demands are given on

19 This implies that there are no outstanding orders at the time the net inventory is compared with the reorder point. We thus assume that an item does not place new orders while there are outstanding orders. If this assumption does not hold, the reorder point is compared to the inventory position rather than the net inventory on hand. For the concept of inventory position, see Zipkin (2000, pp. 30-32) or van Ryzin (2001).

the level mid-term periods, while replenishment lead times are measured in short-term periods, it is given as

$$\text{lead time demand of } i \text{ in } t = \frac{RLT_i}{T^S} \cdot d_{i,t} \tag{5.13}$$

If this lead time demand is uncertain, i is a stockpoint $i \in \mathcal{SP}$ and holds an additional quantity $I^{ss}_{i,t}$ of safety stock. The reorder point is given by the lead time demand plus the defined safety stock.

$$RP_{i,t} = \underbrace{\frac{RLT_i}{T^S} \cdot d_{i,t}}_{\text{demand over } RLT_i} + I^{ss}_{i,t} \tag{5.14}$$

Figure 5.3 shows how the reorder point and safety stock level relate and determine the average inventory level.

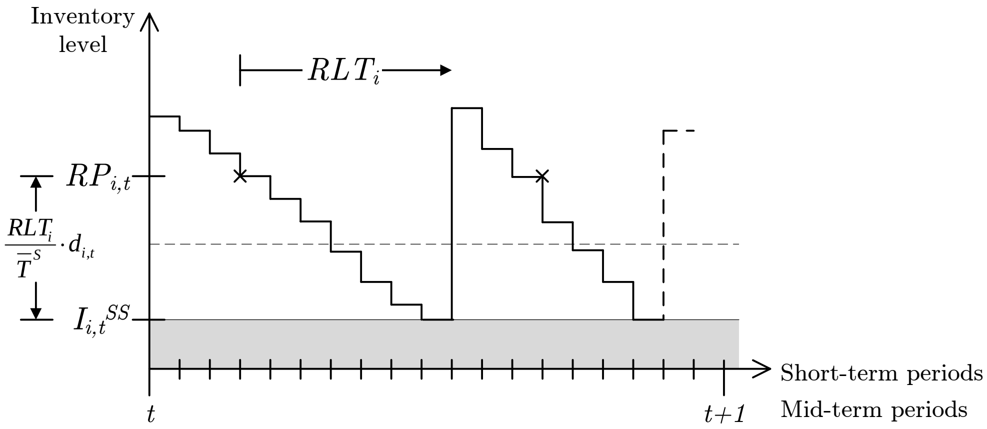


Figure 5.3: Reorder point, safety stock level and average inventory level

The cycle stock build to cover the expected demand over the replenishment lead time is assumed to be half the demand quantity on average. The total inventory cost can then be calculated on the basis of safety and average cycle stock as

$$C^{INV} = \sum_{i \in \mathcal{N}} \sum_{t \in \mathcal{T}} \left(\frac{RLT_i}{2 \cdot T^S} \cdot d_{i,t} \cdot C^{inv}_i \right) + \sum_{i \in \mathcal{SP}} \sum_{t \in \mathcal{T}} I^{ss}_{i,t} \cdot C^{inv}_i \tag{5.15}$$

Setup and scrap cost are operational cost incurred during the execution of a concrete production plan. Given that such an operational production plan is not known[20], estimations of average cost rates to be used in the assessment are required. Let $scrp_i$

20 See the discussion of planned production quantities as a theoretical production plan on an aggregate level in Section 5.1.1.1.

denote the average quantity of scrap produced during the start and end phase of each production run of i. We do not consider additional scrap produced during the entire production execution as it only represents a loss caused by technical characteristics of the production process which cannot be influenced by any decisions related to the distribution of planned production quantities or the assortment and are thus irrelevant to this analysis.

As setup costs are generally sequence-dependent, a precise assessment is only possible on the basis of a concrete production plan with sequence information. As we do not assume any knowledge of production sequences within the periods, we base the assessment of setup costs on values $\overline{C}^{stp}_{i,s,pb_s}$ that express the average setup cost for one production run of i on production process step s with a planning buffer of pb_s. Given that planned production quantities $Q_{i,s,t}$ are defined for each item, production process step and period, these average scrap quantities and setup costs per production run can be used to calculate setup and scrap costs as

$$C^{STSCRP} = \sum_{s \in \mathcal{S}^{PROD}} \sum_{i \in \mathcal{N}_s} \sum_{t \in \mathcal{T}} \left(\overline{C}^{stp}_{i,s,pb_s} + scrp_i \cdot C_i \right) \cdot X_{i,s,t} \tag{5.16}$$

where $X_{i,s,t}$ is a binary indicator to control whether or not a production run is required for the indicated combination of item, production process step and period, defined as

$$X_{i,s,t} = \begin{cases} 1 & \text{if } Q_{i,s,t} > 0 \\ 0 & \text{if } Q_{i,s,t} = 0 \end{cases}$$

With this definition of $X_{i,s,t}$, we assume that if a positive planned production quantity is determined for a certain period, the entire quantity is produced in a single batch. While this is reasonable to keep setup and scrap costs within the periods low, there may be reasons to split up these quantities into more than one batch within a period, which would result in higher setup and scrap costs than those calculated in Equation 5.16. One such reason are capacity bottlenecks, which lead to the inability or unwillingness to block a certain production resource for the time required to produce the entire planned production quantity. The cost function used here thus is conservative and only incurs setup and scrap costs once per period. Alternatives are possible but are highly application-dependent and therefore must be defined individually. Ideally, rules can be defined to obtain more precise estimation of the number of production runs required to produce a certain planned production quantity.

Following the traditional lot-sizing logic, it may be reasonable to produce some materials in advance, i.e. before the period in which the demand originates in order to reduce setup and scrap costs. If planned production quantities do not match the demand quantities exactly but exceed them in any period, cycle stock $I^{cs}_{i,t}$ is built up. The costs for these cycle stocks are assessed analogously to the general inventory cost in Equation 5.15, which yields

$$C^{CYST} = \sum_{s \in \mathcal{S}^{PROD}} \sum_{i \in \mathcal{N}_s} \sum_{t \in \mathcal{T}} \frac{I^{cs}_{i,t-1} + I^{cs}_{i,t}}{2} \cdot C^{inv}_i \tag{5.17}$$

The evaluation of Equations 5.11 to 5.17 for an arbitrary baseline model or scenario represents the ultimate goal of this work. To make this evaluation possible, the following Sections 5.2 and 5.3 describe optimisation methods to set those parameters that are not known in advance and have to be adapted for each model and scenario.

5.2 Inventory Allocation in Production and Distribution Networks

This section presents an optimisation method that yields a near optimal inventory allocation in terms of a set of stockpoints $\mathcal{SP}$, replenishment lead times RLT_i, corresponding safety stock levels $I^{ss}_{i,t}$ and reorder points $RP_{i,t}$ on the basis of a given model $PDN(\mathcal{N}, \mathcal{V}, \mathcal{S})$. The method addresses all requirements described in Section 2.3.3.

5.2.1 Inventory Model and Optimisation Objective

The basic prerequisite to finding an optimal inventory allocation in a production and distribution network is an inventory model that determines the required amount of safety stocks and average inventory levels at a stockpoint depending on the inventory positions and levels at the adjacent items. The key concept to connect the inventory requirements at a stockpoint with the decisions about inventory placements on adjacent items are the replenishment lead times.

As all items $i \in \mathcal{N}$ have positive throughput times[21] TT_i, demand for these items cannot in general be met immediately. Accordingly, each item i quotes a service time ST_i to its successors. This order lead time between adjacent items is the time required

21 See Section 5.1.1.1 for the definition of throughput times for different item types.

for i to fill orders from successor items (*reaction time*). We assume that there are no fixed order cycles such that orders can be placed in any short-term period. For each item the sum of the maximum service time of any predecessor plus the throughput time of that item corresponds to its total replenishment lead time RLT_i:

$$RLT_i = \max_{j \in PR(i)} ST_j + TT_i \tag{5.18}$$

At a given item i it takes replenishment lead time short-term periods from the placement of a replenishment order until the ordered quantity is available for planning at the respective location.[22] This means that an item i has to transfer its forecasted demand volumes into fixed orders to all its predecessors at least RLT_i short-term periods before the requirements date. This replenishment lead time thus corresponds to a frozen forecast interval in which the available quantities at the end of this interval cannot be influenced any more.

Given the short delivery times requested by customers, the service times of the final distribution stages are generally not long enough to cover the total throughput times of all upstream stages. It follows that there must be some items in the network that cannot pass their replenishment lead time to their successor as the service time, which results in a positive *coverage time* $\Delta t_i = RLT_i - ST_i$ at these items. Over this coverage time, demand information is uncertain. Any production or distribution operations for an item i can only be planned with fixed order quantities over time ST_i. In order to be able to deliver in time ST_i if $ST_i < RLT_i$, this item has to replenish its inventory upon forecasts and buffer against demand uncertainty over the time interval Δt_i with safety stock. Figure 5.4 illustrates with a small example how replenishment lead, service and coverage times interrelate.

> **Example 5.1** In the depicted example, n_1 is selected as a stockpoint and covers its entire replenishment lead time with safety stock to ensure a service time of $ST_{n_1} = 0$. The replenishment lead time of the successor item n_2 therefore consists only of its own throughput time, which is passed as the service time $ST_{n_2} = RLT_{n_2}$ to the successors n_3 and n_4. Item n_3 adds its throughput time to that service time and again passes the sum further

22 Analogously to the approach of Simpson Jr. (1958) we assume the existence of operating flexibility to ensure deterministic and constant replenishment lead times. This is a justifiable assumption as it is practically irrelevant to what extent this flexibility actually exists. Without any demand bound stocks would grow infinitely. Demand thus has to be bounded by some service level and any excessive demand can either be handled by operating flexibility or leads to stockouts. If such excessive demand originates from a single customer, other measures like agreements on longer delivery times are also possible.

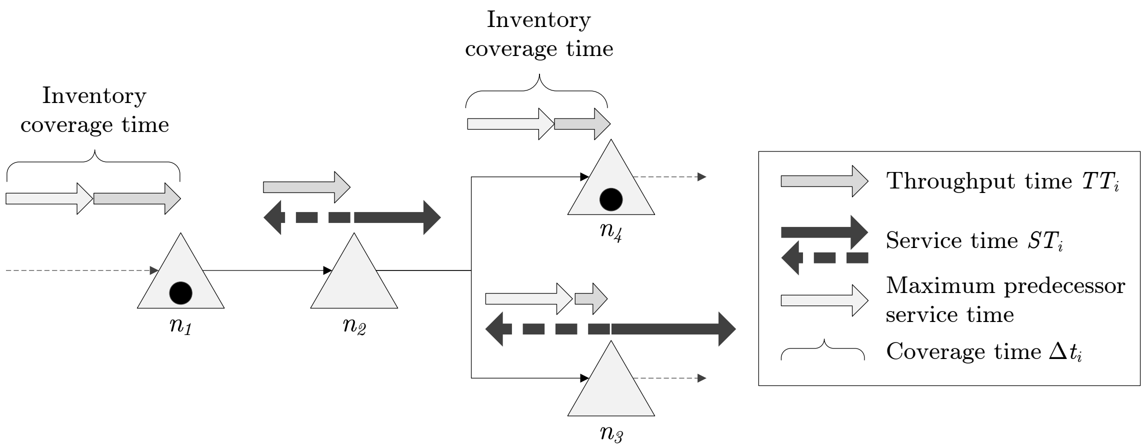


Figure 5.4: Inventory allocation and times between adjacent items

downstream as its own service time $ST_{n_3} = ST_{n_2} + TT_{n_3}$. By contrast, n_4 is selected as a stockpoint and covers n_2's service time plus its own throughput time with safety stock to present a service time $ST_{n_4} = 0$ to its successors.

Now that the interrelations between the nodes can be expressed depending on the selected inventory allocation, a model to determine the required safety stock levels at a single item has to be defined. As demands for an item $i \in \mathcal{N}$ may fluctuate considerably over the periods $\mathcal{T}$, we consider safety stock levels $I^{ss}_{i,t}$ for each single mid-term period $t \in \mathcal{T}$. This is reasonable as we assume that these periods have been chosen to represent the demand planning and forecasting periods[23] and thus the information required to determine safety stock levels is available at this level of aggregation.

The determination of safety stock levels $I^{ss}_{i,t}$ has to consider the replenishment lead time as well as the service time at an item. Depending on the length of the resulting coverage time Δt_i, the safety stock levels have to be dimensioned for each period. The reorder points $RP_{i,t}$ are incremented by this amount of safety stock.[24] The safety stock levels are measured as multiples of the demand standard deviation over the coverage time in which demand is uncertain:[25]

$$I^{ss}_{i,t} = z_{i,t} \cdot \sigma^d_{i,t} \cdot \sqrt{\frac{\Delta t_i}{T^S}} \tag{5.19}$$

23 See Section 5.1.1.1.
24 See Equation 5.14 on page 101.
25 Compare the models for single installation systems discussed in the state of the art in Section 3.2.1.

The task is to determine a value for $z_{i,t}$ that yields a safety stock level that guarantees an average β-service level[26] of β_i^{SL}. In order to determine the proportion of demand that cannot be met with a certain safety stock level, we note that actual demand in each period $D_{i,t}$ is a random variable and use the *partial expectation* of the corresponding probability distribution. The partial expectation $E(X-x)^+$ of a random variable X with respect to a threshold value x is defined as the long run average of the positive part of the difference $X-x$. It tells us how much, on average, the realisations of X exceed the threshold value x.

Transferred to our problem, this enables us to evaluate what proportion of demand cannot be met on average and what β-service level can be achieved with a given safety stock level $I_{i,t}^{ss}$ and a given demand distribution over the replenishment lead time RLT_i. For this evaluation, we consider the partial expectation of the lead time demand random variable with respect to the reorder point as the threshold value. The reorder point $RP_{i,t}$ results from a certain safety stock level, replenishment lead time and expected demand.[27] Knowing the demand distribution $D_{i,t}$ for one mid-term period$t$[28], the lead time demand for an arbitrary lead time interval within that period can be expressed as a random variable $D_{i,t}^{rlt}$, defined as

$$D_{i,t}^{rlt} \sim N\left(d_{i,t}\frac{RLT_i}{T^S}, \sigma_{i,t}^d \cdot \sqrt{\frac{\Delta t_i}{T^S}}\right) \tag{5.20}$$

We use the partial expectation to calculate how much, on average, the lead time demand exceeds the reorder point. As this is the amount of unmet demand, we can set it in relation to the expected lead time demand to obtain the percentage of unmet demand, which then relates to the β-service level as our target criterion as:

$$\beta_i^{SL} = 1 - \frac{E\left(D_{i,t}^{rlt} - RP_{i,t}\right)}{d_{i,t} \cdot \frac{RLT_i}{T^S}} \tag{5.21}$$

The partial expectation is hard to compute for normally-distributed random variables with arbitrary means and standard deviations. However, there are approximations as well as tabular values for the case of a standard normal random variable $Z \sim N(0,1)$ with zero mean and a standard deviation of one. The partial expectation $E(Z-z)^+$ of such a standard normal variable is denoted $L(z)$ and L is called the *standard loss*

26 For the proportion of demand that can be met within a defined time interval, see Section 3.2.1.
27 See Equation 5.14.
28 See the definition of demand distributions $D_{i,t}$ in Section 5.1.1.1.

function. The partial expectation of any normally distributed variable $X \sim N(\mu, \sigma)$ can be expressed via this standard loss function as[29]

$$E(X - x)^+ = \sigma L(z) \quad \text{with} \quad z = \frac{x - \mu}{\sigma} \tag{5.22}$$

If we express the safety stock level $I_{i,t}^{ss}$ via the reorder point and the lead time demand (see Equation 5.14), we can rearrange Equation 5.19 to

$$z_{i,t} = \frac{RP_{i,t} - d_{i,t} \cdot \frac{RLT_i}{T^S}}{\sigma_{i,t}^d \sqrt{\frac{\Delta t_i}{T^S}}} \tag{5.23}$$

This shows that $z_{i,t}$ fits into the normalisation scheme shown in Equation 5.22 and thus Equation 5.21 can be rewritten as

$$\beta_i^{SL} = 1 - \frac{\sigma_{i,t}^d \cdot \sqrt{\frac{\Delta t_i}{T^S}} \cdot L(z_{i,t})}{d_{i,t} \cdot \frac{RLT_i}{T^S}}, \tag{5.24}$$

With the partial expectation, the numerator represents the expected amount of unmet demand, while the denominator represents the expected demand volume over the considered time period. Subtracting this ratio from 1 equals the desired β-service level. Rearranging this equation yields

$$L(z_{i,t}) = \frac{(1 - \beta_i^{SL}) d_{i,t} \frac{RLT_i}{T^S}}{\sigma_{i,t}^d \sqrt{\frac{\Delta t_i}{T^S}}} \tag{5.25}$$

This equation gives a relation between the parameters service level, replenishment lead time, expected demand, demand variation and the standard loss function value of the corresponding safety factor $z_{i,t}$. To obtain the value for $z_{i,t}$, we have to inverse the standard loss function. However, as tabular values for $L(\cdot)$ are available, the corresponding $z_{i,t}$ values can easily be looked up and we do not go into detail of the computation of this function and its inverse. With a method to calculate the $z_{i,t}$ values, Equation 5.19 can be used to calculate the required safety stock levels $I_{i,t}^{ss}$. We can now write the total inventory cost equation used in the cost assessment as a function of the coverage times at each stockpoint:

$$C_i^{INV}(\Delta t_i) = \sum_{t \in \mathcal{T}} \left(\frac{RLT_i}{2 \cdot T^S} \cdot d_{i,t} + z_{i,t} \cdot \sigma_{i,t}^d \cdot \sqrt{\frac{\Delta t_i}{T^S}} \right) \cdot C_i^{inv} \tag{5.26}$$

29 For a proof of this equality, see van Ryzin (2001, p. 12).

The complete inventory allocation problem can now be formulated as

$$\min \quad \sum_{i \in \mathcal{N}} \sum_{t \in T} \left(\frac{RLT_i}{2 \cdot T^S} \cdot d_{i,t} + z_{i,t} \cdot \sigma^d_{i,t} \cdot \sqrt{\frac{\Delta t_i}{T^S}} \right) \cdot C_i^{inv} \tag{5.27a}$$

$$\text{s.t.} \quad RLT_i \geq \max_{j \in PR(i)} ST_j + TT_i \qquad \forall i \in \mathcal{N} \tag{5.27b}$$

$$ST_i \leq ST_i^{max} \qquad \forall i \in \mathcal{N} \tag{5.27c}$$

$$ST_i \leq RLT_i \qquad \forall i \in \mathcal{N} \tag{5.27d}$$

$$ST_i \geq 0 \qquad \forall i \in \mathcal{N} \tag{5.27e}$$

The objective function seeks to minimise the total cost incurred by the average inventory held by all nodes over all time periods. It corresponds to the sum of the inventory costs according to Equation 5.26 over all items. As Equation 5.26 is a function of the coverage times Δt_i, the actual decision variables in the model are the service times ST_i, which determine the values of the Δt_i. It is notable that the required customer service levels do not appear in any restriction, as they are already used to determine the safety factors $z_{i,t}$ in the objective function according to Equation 5.25.

Restriction 5.27b assures that the replenishment lead times are defined according to Equation 5.18. For all items that represent end products and thus are shipped to customers, this service time cannot be decided on independently but is bounded by service agreements with these customers. Therefore the service time of each last-stage distribution item i must not be greater than ST_i^{max} and thus is bounded accordingly by restriction 5.27c. For items where there is no ST_i^{max} defined, this restriction may be omitted or ST_i^{max} may be set to a sufficiently large value. Restrictions 5.27d and 5.27e require the values of ST_i to be positive and at most as long as i's replenishment lead time.

A closer analysis of this optimisation problem shows that it is non-linear. Both the safety stock factors $z_{i,t}$ and the coverage times Δt_i depend on the service times ST_i. Since the ST_i are decision variables and the safety stock factors and the square-root term with the coverage times are multiplied in the objective function, the problem is non-linear. The revision of the state of the art in Section 3.2.3 showed that for this type of inventory model, the extreme point property holds and can be used to transform this problem into a combinatorial optimisation problem. The solution space grows exponentially with the problem size, i.e. the number of items. Therefore we cannot expect to find an efficient exact solution algorithm and must consider a heuristic approach. In the next section we first discuss the use of domain knowledge to find good inventory allocations and then present a concrete heuristic based on these considerations in Section 5.2.3.

5.2.2 Domain Knowledge for Heuristic Inventory Allocation

An important consideration for the development of any heuristic is the integration of as much problem-specific domain knowledge as possible in order to avoid the evaluation of practically irrelevant or infeasible solutions during the optimisation process. For the given problem, there are several such considerations. From the inventory model presented in the previous section, it results that some items are better suited as stockpoints than others. Here, *suited* means that they contribute to covering throughput times with inventory levels that incur comparably little cost. *Comparably* means compared with other items that could be used to cover part of or the same throughput time.

The aim of incorporating domain knowledge is to identify these items in advance, i.e. without testing all possibilities to determine those favourable. The domain knowledge is used via a set of quantitative criteria that can be divided into the two categories

- item characteristics and
- network structure.

For each particular criterion, a key indicator for the eligibility of the considered item as a stockpoint with respect to this criterion is derived. A key indicator is a measurable value that contains aggregated information and is used to put different values into relation with each other for purposes of comparison. Finally, all these key indicators are aggregated to obtain a single measure of an item's stockpoint eligibility, which can be used in a heuristic for inventory allocation. With the aim to have one standardised measure of stockpoint eligibility, all key indicators have to be defined on the common interval $[0, 1]$ and are dimensionless.

5.2.2.1 Item Characteristics

The characteristics of the demand for a certain item mainly determine the total inventory requirements and especially safety stock requirements in case the item is chosen to be a stockpoint. Accordingly, we base the assessment of an item's eligibility as a stockpoint on the following characteristics of the demand process:

- demand volume
- demand variation over time
- forecast deviation

For each of these criteria, the corresponding information already exists in our model. In order to derive a normalised key indicator on the interval $[0, 1]$, we have to put these values of each item into relation with the values of other items. Here it is not always reasonable to compare the values of one particular item with the values of *all* other items in the network, as the following problems may occur:

- For demand volumes, the units of measure may differ and thus quantities are not directly comparable. Comparing valuated quantities, i.e. demand volumes measured in a monetary unit, does not help either, as materials on later production stages naturally have higher volumes and thus bias the comparison.
- A single model may contain products and related materials from different product categories and ranges, for which demand quantities and values are not comparable.

To circumvent these problems, we define a partition[30] $\Pi = \{M_1, \ldots, M_n\}$ over the set of items $\mathcal{N}$ and assess each item $i \in M_k$ only relative to all other items in M_k. The definition of such a partition must follow a reasonable logic to ensure that items within each subset can be compared unbiasedly. Possible approaches to define the subsets M_k are:

- Group production items that are processed on the same production process step: $\Pi = \left\{\mathcal{N}_s \mid s \in \mathcal{S}^{PROD}\right\}$.
- Group procurement items according to the type of material, e.g. raw materials and packaging materials.
- Group distribution items according to product range or category of the end products they represent.
- Generally group items according to their basic unit of measure, such that all items in one $M_k \in \Pi$ are comparable with respect to their quantities.

30 A partition $P = \{S_1, \ldots, S_n\}$ of a set S is a family of subsets of S that are mutually exclusive and jointly exhaustive. That is, no element of S is present in more than one of the subsets, and all the subsets together contain all the members of the original set: $\bigcup_{i=1\ldots n} S_i = S \wedge S_i \cap S_j = \emptyset$ for $i \neq j$.

With these subsets, we can derive normalised key indicators on the interval $[0, 1]$ for each criterion by calculating the ratio of the value of the considered item $i \in M_k$ and a reference value from all items in M_k. These reference values are maximum or minimum values, depending on the definition of the key indicator. However, this approach has to consider potential outliers in the comparison. As the model data may be derived practically from historical data, there may be items with exceptionally high or low key indicator values, e.g. high demand quantities due to promotions or forecast deviations of 100% for items that have not been forecasted at all. If such an item is present in one subset M_k, all other items receive particular low ratings. Thus reference values are adjusted by taking a high percentile ≤ 100, e.g. the 95th percentile of the relevant value.[31] This approach is general enough to allow the use of the maximum (or minimum value) by taking the 100th percentile. At the same time it bears the risk that values > 1 become possible for those items that do not fall below the percentile boundary. For these items, stockpoint eligibility for the corresponding criterion is bounded by definition to 1.

The first criterion is the total expected demand quantity over the periods $\mathcal{T}$, denoted $d_i^{\Sigma} = \sum_{t \in \mathcal{T}} d_{i,t}$. Items with high demands are considered more eligible as stockpoints than items with lower demands. This rationale is also commonly used in approaches to deciding about inventory strategies based on ABC-analyses.[32] With $P_{M_k}^{dem}$ denoting the chosen percentile of the values $\{d_j^{\Sigma} \mid j \in M_k\}$ we measure the stockpoint eligibility of item i with respect to the demand quantity as

$$SE_i^{dem} = \min\left(\frac{d_i^{\sum}}{P_{M_k}^{dem}}, 1\right) \qquad i \in M_k \tag{5.28}$$

The second criterion is the variation of the demand distribution over the time period. As the corresponding key indicator, we use the sample coefficient of variation as the common measure of variation of a time series.[33] The sample coefficient of variation for

31 The Nth percentile of a set of values is defined as the value below which N percent of observations fall. So here e.g. the 95th percentile of the demand volumes would be the smallest value below which 95% of all observed demand quantities may be found. Percentiles are often used in descriptive statistics as a means of avoiding the inclusion of outliers in an analysis.

32 See Tempelmeier (2003, pp. 31-33).

33 The sample coefficient of variation is a measure of the volatility of a time series, defined as the ratio of the sample standard deviation over the sample mean:

$$CV(x_1, \ldots, x_n) = \frac{\sqrt{n^{-1} \sum_i (x_i - \overline{x})^2}}{\overline{x}}$$

with $\overline{x} = n^{-1} \sum_i x_i$ (Brockwell and Davis, 1991, p. 212).

the expected demand over the time periods is defined as

$$CV_i^{dem} = \frac{\sqrt{T^{-1} \sum_{t \in \mathcal{T}} \left(d_{i,t} - \overline{d_i} \right)^2}}{\overline{d_i}} \tag{5.29}$$

where $\overline{d_i} = T^{-1} \sum_{t \in \mathcal{T}} d_{i,t}$. Analogous to the definitions above, $P_{M_k}^{demvol}$ denotes the chosen percentile of the values $\{CV_j^{dem} \mid j \in M_k\}$. As items with smaller demand volatility are better suited as stockpoints than those with high demand volatility, the corresponding key indicator is defined as the distance to 1 and therefore has to be bound to 0 for those items that do not fall below the percentile boundary:

$$SE_i^{demvar} = \max \left(1 - \frac{CV_i^{dem}}{P_{M_k}^{demvol}}, 0 \right) \qquad i \in M_k \tag{5.30}$$

The third criterion is forecast uncertainty as measured by the relative mean absolute forecast deviations FD_i^{mad}. The better demand of an item can be forecasted, the smaller the safety stock requirements of that item in case it is chosen as a stockpoint, which indicates that it is comparatively suitable as a stockpoint in the sense defined above. With $P_{M_k}^{fd}$ defining the chosen percentile of the values $\{FD_j^{mad} \mid j \in M_k\}$, we can define the corresponding key indicator as

$$SE_i^{fd} = \min \left(\frac{FD_i^{mad}}{P_{M_k}^{fd}}, 1 \right) \qquad i \in M_k \tag{5.31}$$

The criteria presented above are interdependent. It can generally be expected that there is correlation between their values, as

1. items with high demand volumes generally have less volatile demand distributions and
2. items with stable demand can be forecast with greater precision and have smaller relative mean absolute forecast deviations.

In summary, the first two criteria are expected to determine the value of the third criterion, which raises the question why the consideration of forecast deviations is not sufficient. We argue that all three criteria have to be evaluated and considered, as there are reasonable exceptions to the relations described above. Sporadic demand that occurs e.g. for promotion items is interpreted as volatile, but may be forecast very precisely due to fixed promotion quantities. Furthermore, demand quantities and volatility are purely external factors, while the forecast deviations result from past planning decisions if working with historical data. Thus relying only on the forecast deviations may lead to wrong assumptions about the eligibility of that item due to particular good or bad forecasts in the past. Accordingly, only the combination of all three factors provides a strong indicator of stockpoint eligibility, as it identifies items with high and steady demands that can be forecast with high precision.

5.2.2.2 Network Structure

The second class of criteria is derived from the structure of the assortment and the distribution relations. Considering a single item, we analyse

- the number of direct successors and end products that the corresponding material goes into,
- the number of direct predecessors and
- whether there is only a single option for further use of that item.

Items of high commonality are more eligible as stockpoints than items of low commonality. As described in Section 3.2.4, item commonality plays an important role in inventory allocation, as external demand uncertainties can be pooled and thus reduced by distributing finished products to multiple locations or using raw and semi-finished materials as input components for production of different materials. These effects have already been quantified in Equation 5.9, which shows how the standard deviations of the demand distributions at an item result from the aggregation of the deviations of all successor items. Another reason is that multiple successor items are dependent on the service time of the considered item, which means that holding inventory at this single central location can reduce the service time to a large number of successor items at comparatively low cost. In the key indicator for this concept, we consider two types of item commonality:

- Item commonality in the strict sense of the number of successors $|SC(i)|$ of an item i.
- Item commonality in the wider sense of the number of distribution items representing end products that an item is related to. Each item in the network is connected, directly of indirectly, to a set of ≥ 1 items that represent end products at sales locations.[34] We denote the set of these end items to which a particular item i is related $LS(i)$.

34 Otherwise, the item could be removed from the network as it is not required for any end product and thus cannot face any dependent demands.

When considering network structures, there is no comparison bias due to outliers as in the case of the demand characteristics. Thus we can safely use the minimum and maximum values from the reference set to calculate the key indicators, which for item commonality is calculated from the indicators for the two different types of commonality as

$$SE_i^{comm} = \frac{1}{2}\left(\frac{|SC(i)|}{\max_{j \in M_k}|SC(j)|} + \frac{|LS(i)|}{\max_{j \in M_k}|LS(j)|}\right) \qquad i \in M_k \tag{5.32}$$

Analogous considerations can be made for the number of predecessors of an item. Placing inventory at an item with a high number of predecessors makes the item less dependent on the material availability of all the supplying items. Accordingly, stockpoint eligibility is also evaluated based on the number of predecessors of an item i as

$$SE_i^{pred} = \frac{|PR(i)|}{\max_{j \in M_k}|PR(j)|} \qquad i \in M_k \tag{5.33}$$

Apart from the advantages of inventory placements at item with many predecessors or successors, we also consider single usage structures, where an item i has only one successor j.

In this case, the downstream item j should be considered a preferred stockpoint, as the absolute inventory requirements are not higher for that item. This can be assured because demand quantities and forecast deviations are directly propagated from j to i.[35] The only causes that might make inventory at i more economical are increased inventory costs at j. These increased costs may result either from higher inventory holding costs due to the value added during a corresponding production stage, or from higher warehousing cost rates at $loc(j)$ if the link between the two items represents a distribution relation. In either case, the increased inventory costs may outweigh the savings generated by reduced replenishment lead times of items further downstream. In contrast to all the criteria described above, this consideration is not a soft criterion to be expressed in an eligibility rating, but gives a direct hint to try to move inventory positions downstream if there is only a single usage relation.

5.2.2.3 Aggregation to a Single Indicator of Stockpoint Eligibility

The eligibility of an item as a possible stockpoint depends on all the criteria expressed via the corresponding key indicators described in the preceding sections that have

35 Compare Section 5.1.1.2.

to be combined into a single metric. As all key indicators are already dimensionless and normalised to $[0,1]$, we can use the straightforward approach of weighting each individual key indicator and using the weighted sum as the overall key indicator SE_i for the stockpoint eligibility of an item i:

$$SE_i = w^{dem} SE_i^{dem} + w^{demvar} SE_i^{demvar} + w^{fd} SE_i^{fd} + w^{comm} SE_i^{comm} + w^{pred} SE_i^{pred} \quad (5.34)$$

where the weights must sum up to one $w^{dem} + w^{demvar} + w^{fd} + w^{comm} + w^{pred} = 1$ to ensure $SE_i \in [0,1]$.

5.2.3 A Tabu Search Heuristic for Inventory Allocation

Tabu search is a proven neighbourhood search meta-heuristic[36] that gained popularity in recent years. Neighbourhood search heuristics define a neighbourhood as the set of solutions that can be reached by applying a single *move* or *operation* to the current solution. These moves usually consist of replacing a certain number of elements in the current solution with new ones. Tabu search seems especially suited to the problem at hand as it is well proven in general and has already yielded promising results in its application to similar problems.[37] Moreover, the possibility of defining moves on a given solution to improve it makes it easy to incorporate the domain knowledge described above into the optimisation process.

Algorithm 5.7 shows the tabu search[38] implementation proposed for this problem.

A solution is represented by a set of stockpoints $\mathcal{SP}$. The current best solution is denoted $\mathcal{SP}^*$ and is set to be the initial solution when the optimisation starts. The most important step of each iteration is the generation of the moves based on the current solution $\mathcal{SP}$ (lines 4 to 6). The definition of the three neighbourhoods used here is discussed in detail in Section 5.2.3.2. From the set $M(\mathcal{SP})$ of all generated moves that are not tabu at that moment, the algorithm selects the move $\overline{m}$ that yields the best objective value. The objective values are evaluated by a function $\overline{f}(\mathcal{SP}, m)$, which gives the objective value of the solution $\mathcal{SP}$ after move m has been applied to it. If the solution in the current iteration is better than the best solution found so far, the latter is updated. At the end of an iteration, the selected solution is set as the new current solution for the next iteration and the tabu list is updated. This update adds

36 For the definition and description of meta-heuristics, see Section 3.2.3.
37 See Section 3.2.3.
38 For a more detailed description of the tabu search approach, see e.g. Glover (1989, 1990).

Algorithm 5.7: Inventory allocation tabu search heuristic

Input: a model $PDN(\mathcal{N}, \mathcal{V}, \mathcal{S})$, an initial solution $\mathcal{SP}$
Result: a near optimal set of stockpoints $\mathcal{SP}^*$

$TL \leftarrow \emptyset$
$\mathcal{SP}^* \leftarrow \mathcal{SP}$
while *stopping criterion is not met* **do**
  generate moves $M_1(\mathcal{SP})$ based on stockpoint eligibility ratings
  generate moves $M_2(\mathcal{SP})$ moving inventory downstream where appropriate
  generate moves $M_3(\mathcal{SP})$ by randomly switching the stockpoint status
  $M(\mathcal{SP}) \leftarrow \{M_1(\mathcal{SP}) \cup M_2(\mathcal{SP}) \cup M_3(\mathcal{SP})\} \setminus TL$
  select $\overline{m} \in M(\mathcal{SP})$ as $\overline{m} = \arg\min_{m \in M(\mathcal{SP})} \left\{ \overline{f}(\mathcal{SP}, m) \right\}$
  $\overline{\mathcal{SP}} \leftarrow$ apply move $\overline{m}$ to $\mathcal{SP}$
  if $f(\overline{\mathcal{SP}}) < f(\mathcal{SP}^*)$ **then**
    $\mathcal{SP}^* \leftarrow \overline{\mathcal{SP}}$
  $\mathcal{SP} \leftarrow \overline{\mathcal{SP}}$
  update TL
return $\mathcal{SP}^*$

the selected move $\overline{m}$ to the tabu list and removes all moves that have been in the tabu list for the length of the tabu tenure.

Further implementation details which are not of conceptual interest are omitted here and described in the validation Chapter 6. Such details include the selection of an appropriate stopping criterion, the length of the tabu list and the usage of an aspiration criterion.

5.2.3.1 Definition of an Initial Solution

Algorithm 5.7 requires an initial solution $\mathcal{SP}$, i.e. an initial set of stockpoints as an input parameter. There are different ways to define such an initial solution:

Random selection For each item, decide randomly if it is a stockpoint in the initial solution or not. The ratio of stockpoints to the total number of items can be controlled via the probability that an item is decided to be a stockpoint. This is an advisable strategy if no assumptions about favourable inventory distributions can be made. The configuration via the probabilities also allows us to obtain extreme cases where there are no stockpoints at all or all items are stockpoints initially.

Last-stage nodes only Make all finished products at sales locations stockpoints. This guarantees service times of 0 at all sales location items and immediate fulfilment of customer orders. The optimisation procedure then searches for possibilities to reduce total inventory cost by introducing additional stockpoints at upstream stages. This is an advisable strategy if material commonality in the assortment represented by the model is very low and therefore inventories are primarily held at the end product level.

Based on stockpoint eligibilities Use the calculated stockpoint eligibilities to determine the stockpoints in the initial solution. This can either be done by definition of a threshold value above which items are set to be stockpoints, or by definition of a ratio of items that should be stockpoints in the initial solution. In the latter case, the required number of stockpoints with highest eligibility ratings are selected. This strategy incorporates most of the domain knowledge and provides good starting solutions. However, it also risks quickly getting stuck in local optima if the whole optimisation process also relies on the eligibility ratings.

While many more alternatives are possible, only these three strategies are used in this work, as they already provide the possibility to define many different starting solutions via their parameterisation.

5.2.3.2 Definition of Moves and Neighbourhoods

In each iteration, the neighbourhood comprises all solutions that can be reached by applying a move $m \in M(\mathcal{SP})$ to the current solution. For our application, a move should be able to represent the following operations:

- add inventories at a stockpoint
- remove inventories from existing stockpoints
- move inventories from one stockpoint to another

In the formal representation, each move $m = (SP^+, SP^-)$ consists of a set of items SP^+ that are made stockpoints in the new solution, and a set of items SP^- that are made non-stockpoints in the new solution. In this way, a move can express an arbitrary change to the solution, including the above-mentioned operations. A move is applied to a solution by updating the set of stockpoints as

$$\mathcal{SP} \leftarrow \left\{ \mathcal{SP} \cup SP^+ \right\} \backslash SP^- \tag{5.35}$$

One effective means of reducing the search space is to fix the stockpoint status of items where possible. For example, there may be items for which inventories are always required, as service times of zero to customers must be guaranteed. On the other hand, there may be items where keeping inventory is impossible due to technical characteristics of the production processes that impede the installation of physical buffers for these items. This may be the case in highly-automated production systems where the output of one production process step is directly passed to the next step. For the sake of simplicity in the description of the neighbourhoods, we do not explicitly consider these fixed stockpoint status, as they do not present any conceptual problem. If such status is maintained for each item, the optimisation procedure has only to ensure that these items are excluded during the generation of moves so that their stockpoint status is never changed.

The set of available moves consists of distinct subsets that pursue different strategies. The moves in M_1 are based on the domain knowledge encoded in the stockpoint eligibility ratings described in Section 5.2.2.3. It contains moves that either

- make an item a stockpoint that is a non-stockpoint in the current solution and has a high stockpoint eligibility rating, or
- make an item a non-stockpoint that is a stockpoint in the current solution and has a low stockpoint eligibility rating.

To define the set of moves, we consider a list of all items non-stockpoint items $\mathcal{N}\backslash\mathcal{SP}$ in descending order of their stockpoint eligibility ratings. We denote the position of an item i in that list as $r^+(i)$. Analogously, we denote the position of an item in the list of all stockpoints $\mathcal{SP}$ in ascending order of their stockpoint eligibility ratings as $r^-(i)$. We can then define the first set of moves as $M_1 = M_1^+ \cup M_1^-$ with

$$M_1^+ = \Big\{ m = (SP^+, SP^-) | SP^- = \emptyset, SP^+ = \{i\} \Leftrightarrow i \notin \mathcal{SP} \wedge r^+(i) > r^+(j) - m_1 \, \forall j \in \mathcal{N}\backslash\mathcal{SP} \Big\}$$
$$M_1^- = \Big\{ m = (SP^+, SP^-) | SP^- = \emptyset, SP^+ = \{i\} \Leftrightarrow i \in \mathcal{SP} \wedge r^-(i) > r^-(j) - m_1 \, \forall j \in \mathcal{SP} \Big\}$$

This definition yields moves that select the most eligible stockpoints and deselect the most ineligible stockpoints under consideration of their current stockpoint status. The size of the neighbourhood created by M_1 is determined by the parameter m_1 and equal to $2 \cdot m_1$, as both M_1^+ and M_1^- include the m_1 changeable items with the highest and lowest ratings, respectively.

The set of moves M_2 is defined as

$$M_2 = \Big\{ m = (SP^+, SP^-) | i \in SP^+ \Leftrightarrow |SC(j)| = 1 \forall j \in PR(i), SP^- = PR(i) \Big\}$$

The moves $m \in M_2$ pursue the strategy to move the inventory positions downstream as described in Section 5.2.2.2. If all predecessors of an item i have i as their only successor, inventory is placed at $i$and removed from all of i's predecessors.

The neighbourhoods that result from the two above-mentioned sets of moves are fully dependent on the appropriateness of the stockpoint eligibility ratings or limited in their scope as they only consider certain network structures. Thus the moves in M_3 add a stochastic element and randomly switch the stockpoint status of single items. Each move only changes one item, as there is no reason to assume that moves with multiple, randomly selected items create any beneficial combinations. The set of all possible single item switches is

$$\overline{M_3} = \left\{ m = (SP^+, SP^-) | i \notin \mathcal{SP} \Rightarrow SP^+ = \{i\} \wedge SP^- = \emptyset, i \in \mathcal{SP} \Rightarrow SP^- = \{i\} \wedge SP^+ = \emptyset \right\}$$

The resulting neighbourhood contains all solutions in which the stockpoint status of one item has been changed.[39] It is our aim to avoid this full evaluation of *all* possible moves as it requires considerable effort to evaluate such a big neighbourhood in each iteration. Accordingly, M_3 only contains a subset of m_3 randomly selected elements from $\overline{M_3}$.

$$M_3 = \left\{ m_3 \text{ moves randomly selected from } \overline{M_3} \right\} \tag{5.36}$$

The full neighbourhood used in the tabu search is defined by the union of the single neighbourhoods. The sizes of neighbourhoods defined via M_1 and M_3 are controlled via the parameters m_1 and m_3. The definition of appropriate neighbourhood combinations and sizes significantly affects the performance of the optimisation and will be discussed in the validation Chapter 6. It has to be noted that the neighbourhoods cannot be expected to be disjunctive, i.e. a certain move can appear twice in different neighbourhoods during a single iteration of the tabu search. In order to avoid duplicate evaluations of the same neighbourhood solution, all evaluated moves are stored during one iteration and then checked before any move is evaluated. The next section discusses how relevant neighbourhood solutions are actually evaluated.

5.2.3.3 Evaluation of Inventory Cost for a Given Solution

In line 8, the tabu search algorithm has to select the move that yields the solution with the best (i.e. minimal) objective value, which requires that all solutions of a neighbourhood are evaluated. Each solution is evaluated with an objective function

39 Note that this also comprises the moves from M_1.

$f : \mathcal{P}(\mathcal{N}) \mapsto \mathbb{R}$ that assesses the total inventory cost $f(\mathcal{SP})$ caused with a certain configuration $\mathcal{SP}$. This assessment is made according to Equation 5.15, which requires that each stockpoint $i \in \mathcal{SP}$ has a replenishment lead time RLT_i and safety stock levels $I^{ss}_{i,t}$ defined.

Algorithm 5.8: $f(\mathcal{SP})$: full evaluation of a solution

Input: a model $PDN(\mathcal{N}, \mathcal{V}, \mathcal{S})$, a current solution $\mathcal{SP}$
Result: total inventory cost for $PDN(\mathcal{N}, \mathcal{V}, \mathcal{S})$ with solution $\mathcal{SP}$

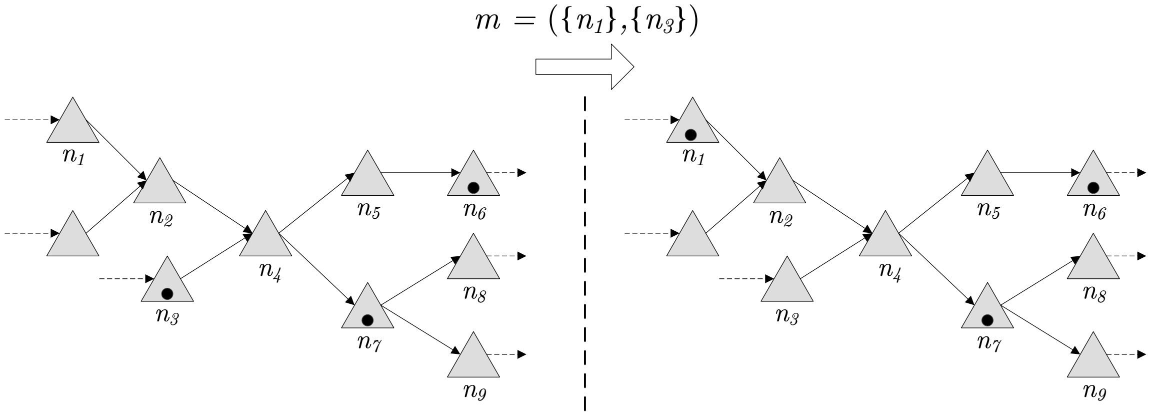
$L \leftarrow$ `TopologicalSorting(`G = ($\mathcal{N}$, $\mathcal{V}$)`)`
foreach *item* i *in* L **do**
    calculate RLT_i according to Equation 5.18
    if $i \in \mathcal{SP}$ **then**
        $ST_i \leftarrow 0$
        calculate safety stock levels $I^{ss}_{i,t}$ for each $t \in \mathcal{T}$ according to Equation 5.19
    else
        $ST_i \leftarrow RLT_i$
        set safety stock levels $I^{ss}_{i,t} = 0$ for each $t \in \mathcal{T}$
return *total inventory cost calculated according to Equation 5.15*

Algorithm 5.8 shows how replenishment lead times, service times and safety stock levels are calculated successively for each item to prepare the evaluation of Equation 5.15. For each item the replenishment lead time is first updated. Depending on the stockpoint status of the current item, the algorithm sets its service time to zero and calculates the required safety stock levels, or it sets the service times at equal to its replenishment lead time and the safety stock levels to zero.

This full evaluation computes the replenishment lead times and service times for all items in the network. The topological sorting is used to ensure that the replenishment lead times and service times of an item are not computed before they have been computed for all predecessors of that item.[40] However, moves generally only change the stockpoint status of a small subset of the items. These changes often do not affect the entire network, but only a small number of items compared to the total number of items in the network. Thus, the evaluation of a neighbourhood solution can be sped up if we only calculate the changes in the inventory levels and related cost of these particular items.

Example 5.2 Consider the part of a network depicted in Figure 5.5. The table shows the throughput, replenishment lead and service times of all

40 See Section 5.1.1.2 for a definition and a similar application of a topological sorting.



Item	TT	RLT	ST
n_1	2	5	5
n_2	2	7	7
n_3	3	5	0
n_4	1	8	8
n_5	2	10	10
n_6	2	12	0
n_7	1	9	0
n_8	2	2	2
n_9	1	1	1

Item	TT	RLT	ST
n_1	2	5	0
n_2	2	2	2
n_3	3	5	5
n_4	1	6	6
n_5	2	8	8
n_6	2	10	0
n_7	1	7	0
n_8	2	2	2
n_9	1	1	1

Figure 5.5: Example of incremental solution evaluation

items. The applied move $m = (\{n_1\}, \{n_3\})$ adds a stockpoint at item n_1 and removes the stockpoint at n_3. The replenishment lead times and service times of all direct and indirect predecessors in the network remain unchanged, and so neither are their inventory requirements affected. Only direct or indirect successor items that can be reached from n_1 or n_3 via a directed path might be affected in terms of changing replenishment lead times. From this set, we can further exclude all items that can only be reached via a path that contains a stockpoint on an intermediate item in the path. In the example, items n_8 and n_9 are not affected, as the increased replenishment lead time is fully covered by increased inventory levels at n_7. The stockpoints n_6 and n_7 are reached on paths that start in n_1 or n_3, and reflect the changes in their replenishment lead times in changing inventory levels, while all their successors are no longer affected due to the defined service time of 0 for all stockpoints. The table on the right side shows the adapted throughput, service and replenishment lead times after the application of the move, where only the highlighted values have actually changed.

Algorithm 5.9: $\overline{f}(\mathcal{SP}, m)$: incremental evaluation of a move on a solution

Input: a model $PDN(\mathcal{N}, \mathcal{V}, \mathcal{S})$, a current solution $\mathcal{SP}$ and a move $m = (SP^+, SP^-)$
Result: Total inventory cost for network $PDN(\mathcal{N}, \mathcal{V}, \mathcal{S})$ after applying move m to $\mathcal{SP}$

apply move m to $\mathcal{SP}$
foreach *item* $i \in SP^+ \cup SP^-$ *changed by* m **do**
  recalculate RLT_i according to Equation 5.18
  if $i \in \mathcal{SP}$ **then**
    $ST_i \leftarrow 0$
    calculate safety stock levels $I^{ss}_{i,t}$ for each $t \in \mathcal{T}$ according to Equation 5.19
  else
    $ST_i \leftarrow RLT_i$
    set safety stock levels $I^{ss}_{i,t} = 0$ for each $t \in \mathcal{T}$
  foreach $j \in SC(i)$ **do**
    `RecalculateInventory` (j)
return *total inventory cost calculated according to Equation 5.15*

Procedure `RecalculateInventory`(i)

recalculate RLT_i according to Equation 5.18
if $i \in \mathcal{SP}$ **then**
  $ST_i \leftarrow 0$
  calculate safety stock levels $I^{ss}_{i,t}$ for each $t \in \mathcal{T}$ according to Equation 5.19
else
  $ST_i \leftarrow RLT_i$
  set safety stock levels $I^{ss}_{i,t} = 0$ for each $t \in \mathcal{T}$
if $i \notin \mathcal{SP}$ **then**
  foreach $j \in SC(i)$ **do**
    `RecalculateInventory` (j)

Algorithm 5.9 together with procedure `RecalculateInventory`(i) shows how the insights from the example can be used to define an *incremental evaluation* of a neighbourhood solution. The objective function $\overline{f}(\mathcal{SP}, m)$ represented by this algorithm evaluates the application of a given move m to a current solution $\mathcal{SP}$. In contrast to the full evaluation of a given solution in Algorithm 5.8, this approach evaluates the inventory cost in the PDN after a given move has been applied, which is just what the tabu search algorithm has to do for each neighbourhood solution (line 8). Starting from all items that are directly changed by the move, their service times and inventory levels are updated and `RecalculateInventory(`i`)` is invoked on all their successors. This procedure updates the replenishment lead time, service time and inventory levels of the processed item and invokes the procedure recursively on all successors until a stockpoint is encountered on the path or no more successors are found. With this procedure, it is perfectly possible that one item is re-evaluated several times, like items n_4 to n_7 in the example given above. This does not cause any inconsistency, as the replenishment lead time of each item always is determined by the maximum predecessor service time after the entire evaluation[41], independently of the sequence in which they are re-evaluated. At the end of the recursion, the relevant values have been updated for all affected items and the costs can be evaluated with Equation 5.15. This limitation in terms of the number of items considered in the evaluation significantly reduces the number of recalculations required to evaluate a single move.

5.3 Determination of Planning Buffers and Planned Production Quantities

This section presents an optimisation method that yields optimal production parameters in terms of planning buffers for all $s \in \mathcal{S}^{PROD}$ and planned production quantities for all $i \in \mathcal{N}^{PROD}$ on the basis of a given model $PDN(\mathcal{N}, \mathcal{V}, \mathcal{S})$.

5.3.1 Optimisation Model

The problem is formulated as a mixed-integer mathematical programming model. This model is designed to solve to optimality the trade-offs described in Section 2.3.3 and can be written as:

41 See Equation 5.18.

$$\min \quad \sum_{s\in\mathcal{S}^{PROD}} \sum_{i\in\mathcal{N}_s} \sum_{t\in\mathcal{T}} \left(\overline{C}^{stp}_{i,s,pb_s} + scrp_i \cdot C_i\right) \cdot X_{i,s,t} \tag{5.37a}$$

$$+ \sum_{s\in\mathcal{S}^{PROD}} \sum_{i\in\mathcal{N}_s} \sum_{t\in\mathcal{T}} \frac{I^{cs}_{i,t-1} + I^{cs}_{i,t}}{2} \cdot C^{inv}_i \tag{5.37b}$$

$$+ \sum_{s\in\mathcal{S}^{PROD}} \sum_{t\in\mathcal{T}} PB^{pen}_{s,pb_s} \tag{5.37c}$$

$$\text{s.t.} \quad \sum_{t=1}^{t=t^*} \sum_{s\in\mathcal{S}_i} Q_{i,s,t} \geq \sum_{t=1}^{t=t^*} d_{i,t} \qquad \forall t^* \in \mathcal{T}, i \in \mathcal{N}^{PROD} \tag{5.38a}$$

$$M \cdot X_{i,s,t} \geq Q_{i,s,t} \qquad \forall i \in \mathcal{N}^{PROD}, s \in \mathcal{S}^{PROD}, t \in \mathcal{T} \tag{5.38b}$$

$$I^{cs}_{i,t-1} + \sum_{s\in\mathcal{S}^{PROD}} Q_{i,s,t} - d_{i,t} = I^{cs}_{i,t} \qquad \forall i \in \mathcal{N}^{PROD}, t \in \mathcal{T} \tag{5.38c}$$

$$\sum_{i\in\mathcal{N}_s} k_{i,s} \cdot Q_{i,s,t} \leq K_s \qquad \forall s \in \mathcal{S}^{PROD}, t \in \mathcal{T} \tag{5.38d}$$

$$Q_{i,s,t} \cdot \overline{Q^{rnd}_{i,s}} = k \cdot Q^{rnd}_{i,s} \qquad k \in \mathbb{N}_0, \forall i \in \mathcal{N}^{PROD}, s \in \mathcal{S}^{PROD}, t \in \mathcal{T} \tag{5.38e}$$

$$\overline{Q^{rnd}_{i,s}} \cdot M \geq Q^{rnd}_{i,s} \qquad \forall i \in \mathcal{N}^{PROD}, s \in \mathcal{S}^{PROD} \tag{5.38f}$$

$$\overline{Q^{rnd}_{i,s}} \in \{0,1\} \qquad \forall i \in \mathcal{N}^{PROD}, s \in \mathcal{S}^{PROD} \tag{5.38g}$$

$$pb_s \in \left\{pb^{min}_s, \ldots, pb^{max}_s\right\} \qquad \forall s \in \mathcal{S}^{PROD} \tag{5.38h}$$

$$X_{i,t} \in \{0,1\}, Q_{i,s,t} \in \mathbb{R}, I^{cs}_{i,t} \in \mathbb{R} \qquad \forall i \in \mathcal{N}^{PROD}, s \in \mathcal{S}^{PROD}, t \in \mathcal{T} \tag{5.38i}$$

The model is intended to determine the parameters that serve as the basis of the cost assessment. The objective function comprises the corresponding setup, scrap and cycle stock costs (5.37a and 5.37b), analogously to the calculations of the cost components

C^{STSCRP} and C^{CYST} in Equations 5.16 and 5.17, respectively. In addition, penalty costs for inventories incurred via the length of the chosen planning buffer are also added to the objective function (5.37c). In this way the relevant trade-off between the length of the planning buffer and the amount of setup cost incurred is encoded in the objective function. With increasing planning buffers pb_s, the average setup costs $\overline{C}^{stp}_{i,s,pb_s}$ and therefore the value of 5.37a decrease. At the same time the penalty costs in 5.37c increase with pb_s.

Restriction 5.38a assures that the planned production quantities cover the actual demand quantities. The planned production quantities have to be scheduled at latest in the periods where the corresponding demand occurs. Demand quantities for an item i can be covered by allocating planned production quantities on any of the production process steps assigned to i. This is implicitly assured by restriction 5.38a, as for any production item i, only planned production quantities on related production process steps $s \in \mathcal{S}_i$ are summed. Thus any planned production quantities allocated to other infeasible process steps are automatically set to 0. The binary production indicator variable $X_{i,s,t}$ is used in the objective function to incur setup and scrap cost where appropriate. It is therefore forced to a value of 1 by restriction 5.38b in case there are positive production quantities defined. The parameter M here refers to an arbitrary but sufficiently large number $M \gg Q_{i,s,t}$ for any planned production quantity.

The inventory balance restriction 5.38c assures that the variables $I^{cs}_{i,t}$ hold the correct level of cycle stock for item i at the end of period t. For the cost assessment it is sufficient to determine these cycle stocks for each item and period, which is why the planned production quantities are summed up over all relevant production process steps.

The planned production quantities describe a theoretical production plan at an aggregate level. This plan has to be defined in a way that makes it possible to transfer it into an operational production plan by scheduling production orders for all planned production quantities assigned to one period and sequencing their execution. Thus planned production quantities must already consider all capacity and lot-sizing constraints. Restriction 5.38d assures that planned production quantities are distributed over the available production process steps so that the capacity available on each production process step is not exceeded. We assume that the total capacity available suffices to produce the required quantities, which is particularly reasonable if the model data is derived from historical data where observed demands did not exceed the capacity of the production resources.

Further lot-sizing restrictions may result from technical characteristics of the production process that only allows economical production of multiples of certain quantities.[42] Therefore, $Q_{i,s}^{rnd}$ defines the rounding value for actual lot sizes and consequently also the planned production quantities. Restriction 5.38e assures that all production quantities for an item i on a production process step s must be multiples of this rounding value $Q_{i,s}^{rnd}$. This restriction can be made ineffective via the additional binary indicator $\overline{Q_{i,s}^{rnd}}$ (5.38g), which is set to 0 if there are no specific rounding values and $Q_{i,s}^{rnd}$ is 0. Otherwise, it is forced to 1 by restriction 5.38f, where M again denotes an arbitrary but sufficiently large value. The additional indicator $\overline{Q_{i,s}^{rnd}}$ is necessary to avoid making the model infeasible by impeding positive planned production quantities on production process steps without such rounding constraints in restriction 5.38e.

Restrictions 5.38g to 5.38i define the domains of all variables. Potential planning buffers are restricted to a maximum and minimum value. This is required as the average setup cost for each material must be available for each potential planning buffer, so that a finite set of potential planning buffers has to be defined a priori. This is also reasonable from a practical point of view, as the additional benefit of increasing planning buffers decreases. A production planner should be able define what minimum planning buffer is possible and beyond which planning buffer length no additional benefit is to be expected.

One requirement for the solution of the subprobem addressed in this section is the possibility to solve the optimisation problem with standard software. The presented model contains several binary and integer variables that make it a mixed-integer optimisation problem, which does not impede the use of standard solvers. However, the objective function part 5.37a is non-linear due to the relation between average setup costs $\overline{C}_{i,s,pb_s}^{stp}$ and production indicators $X_{i,s,t}$. These are all decision variables since the former depend on the choice of a planning buffer, while the latter are determined by the corresponding planned production quantities.

As this is a problem for the use of standard solvers, we propose two approaches to deal with this particular non-linearity. Both are based on the idea of changing the status of the planning buffers pb_s from decision variables to fixed parameters in order to make the objective function part 5.37a linear. The resulting problem instances can then be solved separately. The remaining questions are what combinations of planning buffers to choose to obtain this set of optimisation problem instances and how the best solution from this set relates to the optimal solution of the original model.

42 This is especially the case in base production stages where the output cannot be measured in pieces.

The first possibility is to try all combinations of planning buffer candidates for all production process steps. This is possible as both the number of production process steps $\left|\mathcal{S}^{PROD}\right|$ and the number of planning buffer candidates $npb_s = |\{pb_s^{min}, \ldots, pb_s^{max}\}|$ are finite and yields the same exact results as the original problem formulation. However, the number of resulting subproblems then is $npb_s^{\left|\mathcal{S}^{PROD}\right|}$ and grows exponentially with the number of production process steps. This approach is thus only applicable if either the number of planning buffer candidates or the number of considered production process steps is very small. Considering practical applications, this cannot be taken for granted.

One way to alleviate this complexity problem is to constrain the set of combinations of planning buffers tested by explicitly imposing restrictions on these combinations. For instance, a production process step s_1 always works with the same planning buffer as another production process step s_2, or if pb_{s_1} is larger than a certain value, pb_{s_2} must not be greater than a certain value. The biggest simplification that can be made using this approach is to require that all production process steps operate with the same planning buffer. This results in only $npb_s \cdot \left|\mathcal{S}^{PROD}\right|$ separate problems to be solved, but at the same time this equality of planning buffers is a strong assumption that may lead to suboptimal solutions. Whether or not such rules can be formulated to reduce the number of tested combinations sufficiently is highly application-dependent.

An alternative approach to reduce the number of problem instances is to formulate the problem only for a single production process step. Considering the critical objective function part 5.37a, there is no reason that requires all production process steps being optimised in an integrated model. With this approach, a separate problem instance is solved for each planning buffer candidate of each production process step. This decoupling of the production process steps allows to test each planning buffer candidate on each production process step with reasonable effort, since the total number of problem instances to be solved again is $npb_s \cdot \left|\mathcal{S}^{PROD}\right|$. However, there are two points to consider.

Firstly, it may lead to suboptimal results if there are alternative production possibilities and some items are assigned to more than one production process step $|\mathcal{S}_i| > 1$ for some $i \in \mathcal{N}^{PROD}$. In this case, demand quantities of these items have to be split up and distributed over the corresponding production process steps a priori. This may make the overall result suboptimal, as the optimisation model cannot distribute required production quantities arbitrarily over the available production process steps to find assignments at minimum setup and scrap costs.

Secondly, the precise determination of the penalty cost parameters PB^{pen}_{s,pb_s} may become impossible since the required penalties for a given production process step and planning buffer may depend on the choice of planning buffers for other production process steps. Objective function part 5.37c thus requires the integrated optimisation of all production process steps. We discuss this problem in Section 5.3.3 and present two approaches to solve this interdependency by approximation of the planning buffer penalties.

In summary, we conclude that the optimisation model can be transformed into a linear model by solving it either for a subset of planning buffer combinations or separately for each production process step. If the number of production process steps and planning buffers is too big to try all combinations and if no reasonable subset of these combinations can be defined, the second approach described above represents the best trade-off between the approximations made and the practical solvability of the problem. Methods for the approximation of planing buffer penalties required for this approach are presented in Section 5.3.3.

5.3.2 Estimating Average Sequence-Dependent Setup Cost

The relation between the length of the planning buffer and the average setup costs on a certain production process step is represented by the parameters $\overline{C}^{stp}_{i,s,pb_s}$. As these values are generally not known, this section presents a method of deriving an estimation based on the data available from ERP systems. The observation that average setup costs can be reduced with larger planning buffers, i.e. the series $(\overline{C}^{stp}_{i,s,pb^{min}_s}, \ldots, \overline{C}^{stp}_{i,s,pb^{max}_s})$ is monotonically decreasing, is due to the fact that the sequencing procedure has more possibilities to bring forward or postpone production orders so they can be combined sequenced after orders for similar materials. The method of generating estimations for the $\overline{C}^{stp}_{i,s,pb_s}$ uses this logic.

We assume that setup cost information is available in terms of setup matrices for each production process step. In such a setup matrix, the entries $c^{stp}_{s,i,j}$ represent the setup costs incurred by a changeover from $i \in \mathcal{N}_s$ to $j \in \mathcal{N}_s$ on production process step s. We propose a simulation approach to generate the estimated values from these setup matrices by evaluating a large number of setup sequences. Algorithm 5.11 shows the steps of the calculation and is invoked for each $s \in \mathcal{S}^{PROD}$ and each potential planning buffer candidate $pb_s \in \{pb^{min}_s, \ldots, pb^{max}_s\}$.

Algorithm 5.11: Simulation approach to estimate average setup costs

Input: $s \in \mathcal{S}^{PROD}$, $pb_s \in \{pb_s^{min}, \ldots, pb_s^{max}\}$, z_{pb_s}
Result: average setup costs $\overline{C}_{i,s,pb_s}^{stp}$ $\forall i \in \mathcal{N}_s$

while *not* $n_i^{sim} \geq \overline{n}^{sim} \forall i \in \mathcal{N}_s$ **do**
  $\overline{\mathcal{N}} \leftarrow z_{pb_s}$ items randomly selected from $\mathcal{N}_s$
  Create cost-optimal production sequence for items in $\overline{\mathcal{N}}$
  foreach $i \in \overline{\mathcal{N}}$ *that is not the first item in the optimised sequence* **do**
    $\overline{c}_i^{stp} \leftarrow$
    $\overline{c}_i^{stp} + c_{s,j,i}^{stp}$ where j is the predecessor of i in the optimised production sequence
    $n_i^{sim} \leftarrow n_i^{sim} + 1$

foreach $i \in \mathcal{N}_s$ **do**
  $\overline{C}_{i,s,pb_s}^{stp} \leftarrow \overline{c}_i^{stp} / n_i^{sim}$

The algorithm makes use of the fact that the average setup costs incurred by a set of production orders depends in the size of this set. At any given time, there is a certain number of production orders available for which all required components have been provided by the predecessor stages and which are therefore available for release to production. This number depends on the length of the planning buffer: The longer the planned time span between the provision of all components and the requirement date of a production order, the more production orders are available for release to production at any given time. This expected number of production orders within the planning buffer interval pb_s is denoted z_{pb_s}.[43]

The basic idea of this approach is to imitate this situation that a production planner faces when generating production sequences for a certain production process step with a given planning buffer. The algorithm randomly selects z_{pb_s} items from the set of items processed on the considered production process step. This subset represents the set of production orders that have to be executed within the planning buffer interval. For this set an optimal production sequence with respect to the total setup cost incurred is created. Due to the very limited number of orders to be scheduled, this optimisation problem does not pose a major problem. For realistic values of z_{pb_s}, the problem can still be solved optimally, either by full enumeration of the possibilities or with standard solvers. If the number of orders within a single planning buffer interval should really

43 If this number is not known, an historical average can be used or it can be estimated. For example, if the planning buffer interval comprises 6 shifts, and in each shift 1,5 different production orders are processed on average, we can use $z_{pb_s} = 9$.

get too large in a practical application, there are plenty of well-established heuristics for the sequencing problem[44] and potentially suboptimal solutions of such heuristics do not cause major problems in this context of parameter estimation.

For each item in that sequence, the observed setup costs are taken as a single sample for the estimation of the average setup costs. Over the iterations of the simulation, the total setup cost incurred for a certain item is recorded as $\overline{c}_i^{stp}$ (line 5). The cost incurred for each item is the actual setup cost taken from the setup matrix, given the predecessor in the optimised production sequence. For the first item in the optimised sequence no sample is recorded since we only consider a limited section of an actual production plan. The number of samples observed for each item is recorded as n_i^{sim}. This number serves as the stopping criterion and the simulation stops when $\overline{n}^{sim}$ samples have been recorded for each item. It furthermore is used to calculate the estimated average setup costs from the total setup costs observed over the iterations of the simulation (line 9).

5.3.3 Calculating Penalty Cost for Inventories Incurred by Planning Buffers

One integral part of determining appropriate planning buffers is the integration of penalty costs for the length of the planning buffer. The task addressed in this section is the determination of appropriate values for the planning buffer penalty costs PB_{s,pb_s}^{pen}. These penalty costs for planning buffers are included in the objective function (5.37c) to ensure that not all planning buffers are set to their maximum values pb_s^{max}, which would ensure the minimal setup costs (5.37a), but not necessarily minimum overall cost if we also consider inventory costs in the network. The planning buffer penalty PB_{s,pb_s}^{pen} must therefore represent the additional inventory cost incurred in the *entire* PDN if s operates with a planning buffer of pb_s, compared to the configuration where the planning buffer is 0.

Additional inventory costs are incurred at items which are stockpoints and which are affected by a change of the planning buffer of production process step s. An item i is affected if changes of the considered planning buffer pb_s change its replenishment lead time RLT_i. The logic to find these affected items is similar to that logic used to update the replenishment lead times and inventory levels in the incremental evaluation presented in Section 5.2.3.3 and we can reuse part of that logic. All affected items can be found on a *critical path* downstream in the network. A critical path is a maximum sequence $(n_1, \ldots, n_k)$ of items with an item $n_1 \in \mathcal{N}_s$ that is processed on s at its beginning and which contains at most one stockpoint at the end of the path, i.e.

44 See e.g. Fleischmann and Meyr (1997), Haase (1996) and Gupta and Magnusson (2005).

$n_1, \ldots, n_{k-1} \notin \mathcal{SP}$. The replenishment lead times of all items processed on s are changed directly by the increase of the planning buffer. This increased lead time is passed downstream as an increased service time until a stockpoint is reached that covers the additional replenishment lead time with an increased inventory level.

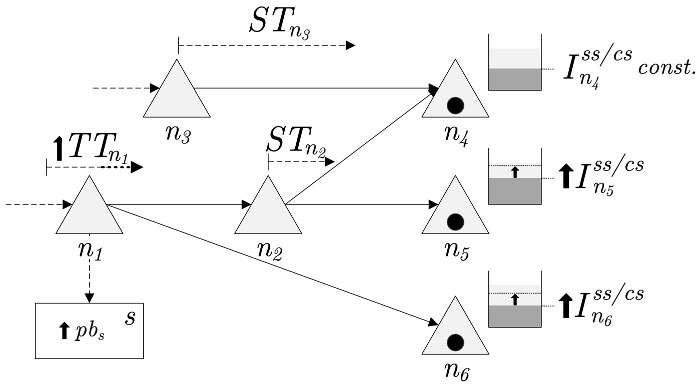


Figure 5.6: Additional inventory cost due to increasing planning buffers

> **Example 5.3** Figure 5.6 shows which items are affected in a network if the planning buffer of production process step s is increased. Firstly, the throughput times and thereby the replenishment lead times of all items processed on s are increased, which in this case applies only to n_1. As n_1 is not a stockpoint itself, it propagates this increased throughput time as an increased service time to its successors. This logic is applied recursively, until a stockpoint item is reached, in this case n_4, n_5 and n_6. These items may be affected by the changing service time of one of their predecessors, but only if that service time is the maximum predecessor service time that directly affects their replenishment lead time.[45] In this case, n_4 and n_5 are affected as they only have one predecessor that determines their replenishment lead time. Item n_4, however, though reachable on a path from n_1, is not affected as its replenishment lead time is determined by the service time of n_3 due to $ST_{n_3} > ST_{n_4}$ and therefore $\max_{j \in PR(n_4)} ST_j = ST_{n_3}$.

For the calculation of PB^{pen}_{s,pb_s} for each production process step s and planning buffer candidate pb_s, two characteristics have to be considered:

Nonlinearity As the inventory cost function is non-linear, inventory costs do not increase linearly with the planning buffers and replenishment lead times.

45 See Equation 5.18.

Interdependence The degree to which the replenishment lead time of a particular item i is increased depends on the planning buffer decisions on *all* production process steps related to i via any direct or indirect predecessor.

The combination of these two characteristics impedes an exact evaluation of the inventory penalty costs. Firstly, the nonlinearity is a problem due to the interdependencies. It suggests evaluation of the PB^{pen}_{s,pb_s} explicitly for each planning buffer candidate. This would however be necessary for all possible combinations of planning buffer configurations for the remaining production process steps and thus lead to the same complexity already discussed in the context of the nonlinearity of the optimisation model in Section 5.3.1. We already proposed to circumvent this problem by solving the optimisation model for each production process step separately, which renders the exact assessment with consideration of all combinations impossible. Secondly, the interdependencies are a problem due to the nonlinearity of the inventory cost function. With a linear cost function, the configuration for the remaining planning buffers would be irrelevant when considering one particular production process step, as costs would increase linearly anyway. With the combination of both characteristics, none of these approaches suffices.

This analysis suggests two feasible approaches to approximately assess penalty costs. Firstly, the interdependencies may be neglected and penalty costs are evaluated for each planning buffer candidate of each individual production process step separately. Secondly, the inventory cost function can be linearly approximated, which renders the interdependencies irrelevant and allows assessment of the penalties for a single production process step without consideration of the configurations of remaining planning buffers.

The first approach is defined in Algorithm 5.12. For each production process step and planning buffer candidate the cumulative penalty costs for all *critical paths* that start in one of the related production items $\mathcal{N}_s$ are calculated. During the iteration over all related production items, two cases have to be distinguished: If the considered item i is a stockpoint itself, the penalty costs can be calculated directly as the difference of inventory costs (Equation 5.26) with and without the considered planning buffer of length pb_s. If it is not a stockpoint, the penalty cost for all paths that start in i are calculated by summing the results of procedure `CalcPBPenalty`(i, j, pb_s) over all successor items j. This procedure first checks if an increase of i's service time would affect the replenishment lead time of the considered successor item j. If this is not the case, the penalty costs amount to 0 and no further successor has to be considered from here. Otherwise the same logic as already described above applies. If j is a stockpoint,

Algorithm 5.12: Calculation of planning buffer penalties disregarding interdependencies

Input: a model $PDN(\mathcal{N}, \mathcal{V}, \mathcal{S})$ with an inventory allocation $\mathcal{SP}$
Result: the planning buffer penalty costs PB^{pen}_{s,pb_s}

foreach $s \in \mathcal{S}^{PROD}$ **do**
    foreach $pb_s \in \{pb_s^{min}, \ldots, pb_s^{max}\}$ **do**
        foreach $i \in \mathcal{N}_s$ **do**
            if $i \in \mathcal{SP}$ **then**
                $PB^{pen}_{s,pb_s} \leftarrow PB^{pen}_{s,pb_s} + \left(C_i^{INV}(\Delta t_i + pb_s) - C_i^{INV}(\Delta t_i)\right)$
            else
                $PB^{pen}_{s,pb_s} \leftarrow PB^{pen}_{s,pb_s} + \sum_{j \in SC(i)}$ `CalcPBPenalty(`i, j, pb_s `)`

Procedure `CalcPBPenalty(`i, j, pb_s`)`

if $\arg\max_{p \in PR(j)} ST_p = i$ **then**
    if $j \in \mathcal{SP}$ **then**
        return $C_j^{INV}(\Delta t_j + pb_s) - C_j^{INV}(\Delta t_j)$
    else
        return $\sum_{k \in SC(j)}$ `CalcPBPenalty(`j, k, pb_s `)`
else
    return 0

the penalty costs are calculated and returned immediately, while the calculation is invoked recursively over the successors of j if j is not a stockpoint.

The second feasible approach is linearly to approximate the inventory cost function. If inventory costs are assumed to increase linearly with replenishment lead times, a single penalty factor that is independent of the configuration of planning buffers at the remaining production process steps can be calculated. For fixed safety stock factors $z_{i,t}$ the inventory cost function is concave due to the square root of the coverage time interval.[46] Concave functions can be linearly approximated by introducing 2 reference

46 See Equation 5.27a on page 108.

points that mark the relevant interval of that function.[47]. For the coverage times Δt_i, we can limit this interval to

$$0 \leq \Delta t_i \leq \Delta t_i^{max}, \qquad \text{with } \Delta t_i^{max} = \max_{w \in W(\mathcal{N}^{PROC}, i)} \sum_{j \in w} TT_j \tag{5.39}$$

The set $W(\mathcal{N}^{PROC}, i)$ denotes all paths from a procurement node to i. The upper bound Δt_i^{max} is the maximum replenishment lead time that can be expected, based on the maximum sum of throughput times of the nodes on any path $w \in W(\mathcal{N}^{PROC}, i)$. The convex part of the inventory cost formula is given by the square root term, which can now be approximated over the interval $[0, \Delta t_i^{max}]$ as[48]

$$\sqrt{\frac{\Delta t_i}{T^S}} \quad \approx \quad \frac{1}{\sqrt{\Delta t_i^{max}}} \cdot \frac{\Delta t_i}{\sqrt{T^S}} \tag{5.40}$$

As expected, this approximation yields exact results only for the two references points 0 and Δt_i^{max} at the extremes of the interval. Applying this approximation to the inventory cost Equation 5.26 yields

$$\tilde{C}_i^{INV}(\Delta t_i) = \sum_{t \in \mathcal{T}} \left(\frac{RLT_i}{2 \cdot T^S} \cdot d_{i,t} + \frac{z_{i,t} \cdot \sigma_{i,t}^d}{\sqrt{\Delta t_i^{max}}} \cdot \frac{\Delta t_i}{\sqrt{T^S}} \right) \cdot C_i^{inv} \tag{5.41}$$

which in contrast to the original equation is linear in Δt_i. This allows us to derive exactly one planning buffer penalty value per item that approximately shows how inventory costs increase if the replenishment lead time and thereby the coverage time Δt_i increase by 1 short-term period. For any $\Delta t_i \in [0, \Delta t_i^{max}]$, this value is constant and equal to

$$\tilde{C}_i^{INV}(\Delta t_i + 1) - \tilde{C}_i^{INV}(\Delta t_i) \tag{5.42}$$

Algorithm 5.14 defines the calculations of the second approach and shows how to determine the planning buffer penalties based on the presented linearisation.

As this algorithm is similar to Algorithm 5.12, we limit explanation to the differences. As only one *linear* penalty cost factor is derived, the entire procedure is only carried out once per production process step. For each affected item at which additional

47 Minner also proposes such a linearisation approach to transform a inventory allocation problem to a linear optimisation problem (Minner, 2000, pp. 152-154).

48 This is the simplest form of such a linearisation using only two reference points. More exact approximations can be obtained by increasing the number of reference points and thereby making the approximation a piecewise linear function. For a description of the required additional variables and constraints, see Domschke and Drexl (2005, pp. 206-208).

Algorithm 5.14: Calculation of planning buffer penalties using linearisation

Input: a model $PDN(\mathcal{N}, \mathcal{V}, \mathcal{S})$ with an inventory allocation $\mathcal{SP}$
Result: the planning buffer penalty costs PB^{pen}_{s,pb_s}

1 **foreach** $s \in \mathcal{S}^{PROD}$ **do**
2   **foreach** $i \in \mathcal{N}_s$ **do**
3   **if** $i \in \mathcal{SP}$ **then**
4             $PB^{pen}_s \leftarrow PB^{pen}_s + \left(\tilde{C}^{INV}_i(\Delta t_i + 1) - \tilde{C}^{INV}_i(\Delta t_i)\right)$
5   **else**
6             $PB^{pen}_s \leftarrow PB^{pen}_s + \sum_{j \in SC(i)}$ `LinearPBPenalty(`i, j`)`
7
8   **foreach** $pb_s \in \{pb^{min}_s, \ldots, pb^{max}_s\}$ **do**
9         $PB^{pen}_{s,pb_s} \leftarrow PB^{pen}_s \cdot pb_s$
10
11

Procedure `LinearPBPenalty(`i, j`)`

1 **if** $\arg\max_{p \in PR(j)} ST_p = i$ **then**
2   **if** $j \in \mathcal{SP}$ **then**
3   **return** $\tilde{C}^{INV}_i(\Delta t_i + 1) - \tilde{C}^{INV}_i(\Delta t_i)$
4   **else**
5   **return** $\sum_{k \in SC(j)}$ `LinearPBPenalty(`j, k`)`
6
7 **else**
8   **return** 0
9

costs are incurred, the linear approximation for the penalty cost is used according to Equation 5.42 (compare line 4 in the algorithm and line 3 in the procedure). After all items of the considered production process step have been processed, the value PB^{pen}_{s,pb_s} contains the penalty cost for the increase of the corresponding planning buffer by 1. As we assume linear cost increases in this approach, the corresponding penalty costs for each individual planning buffer PB^{pen}_{s,pb_s} can be determined by multiplication with each planning buffer candidate $pb_s \in \{pb^{min}_s, \ldots, pb^{max}_s\}$ (lines 8 to 9).

5.3.4 Integration with the Inventory Allocation Problem

This section addresses the remaining open task of integrating the two optimisation models presented. According to the requirements formulated, this integration also has to solve the problem of interdependencies between the variables of both models as described in Section 2.3.3.

The only viable solution is a repeated sequential invocation of both optimisation methods, where each execution uses the results of the previous solution of the *other* optimisation problem. Figure 5.7 shows what the intended sequence of invocations looks like.

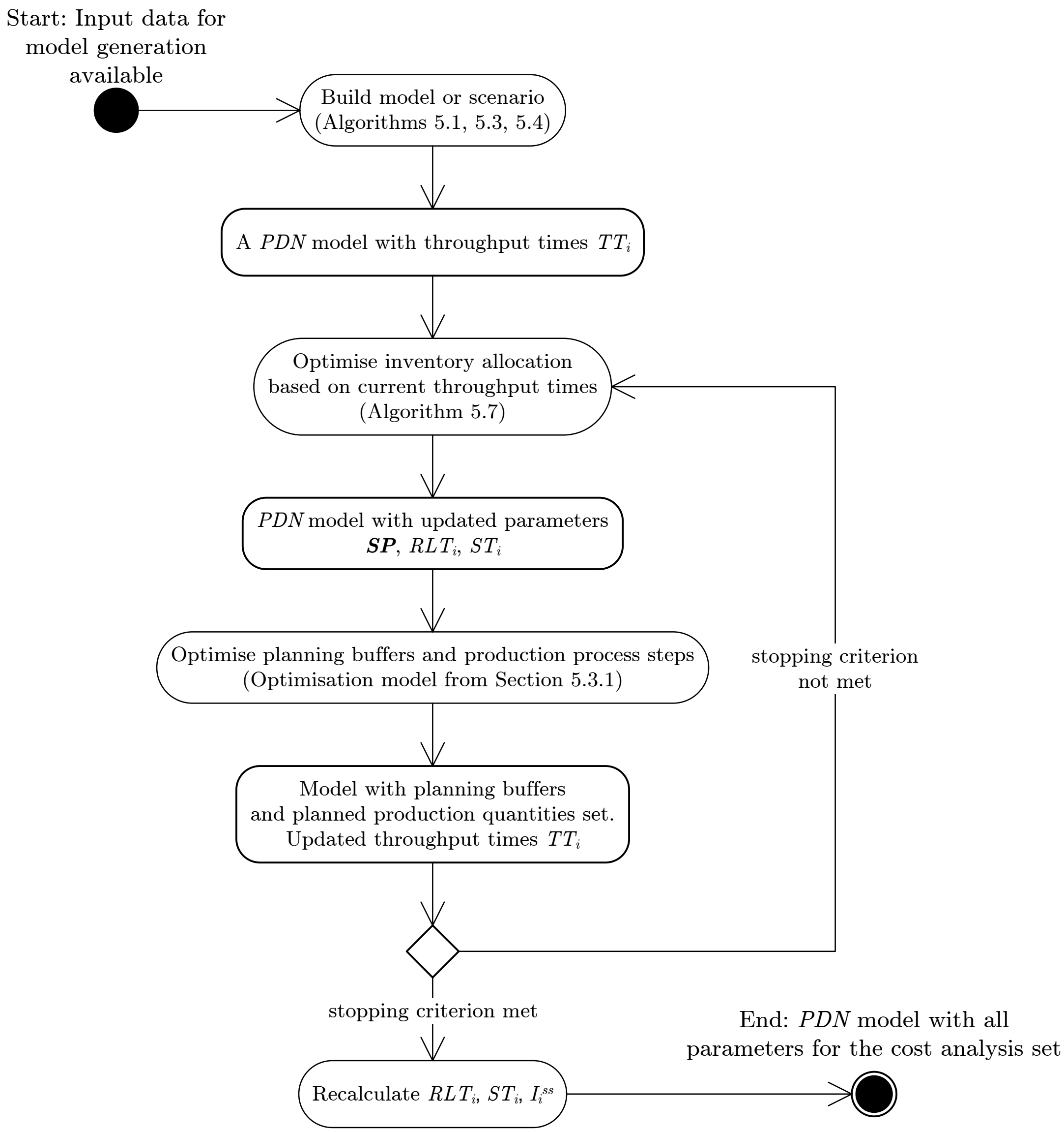


Figure 5.7: Integration of the optimisation methods

Once a model or scenario has been built, the inventory allocation method is used to determine the stockpoints $\mathcal{SP}$, replenishment lead times RLT_i and service times ST_i for each item. The set of stockpoints and the determined replenishment lead times are required as input parameters for the parameter estimation (Algorithms 5.12 and 5.14) in the context of the optimisation of planning buffers and planned production quantities. After this second optimisation step has been performed, the network model contains defined planning buffers and planned production quantities. The changes of planning buffers also affect the throughput times TT_i. The stockpoint allocation determined previously may therefore no longer be optimal, as it was based on different input parameters.

On the basis of the updated network, the sequence provides a loop back to the inventory allocation optimisation method, which is then invoked again to consider the updated throughput times. In any repeated invocation, the previous stockpoint allocation can be used as a starting solution. The fact that only a small part of the input parameters in terms of planning buffers has actually changed suggests that the new optimal solution is closely similar to the previous one. These repeated evaluations are thus much faster than the initial optimisation. The same applies to the subsequent repeated solution of the second optimisation model where the determination of planned production quantities is not affected by the parameters changed during the repeated inventory allocation optimisation. Therefore the planned production quantities determined in the initial optimisation can be fixed and only planning buffers are optimised in subsequent optimisation runs, which reduces the problem complexity enormously.

As indicated in the flow chart, some stopping criterion is required to limit the number of repeated invocations and guarantee an eventual termination. There are several stopping criteria possible, e.g.

No more changes in decision variables No repeated invocation is required if no more changes in the optimal configuration of stockpoints and planning buffers are observed. This stopping criterion is desirable, as it guarantees that an overall optimal configuration is found despite the sequential solution of the optimisation problems. Practically, the fulfilment of this criterion is even likely since the stockpoint allocation is only relevant for the second optimisation model precisely to estimate the penalty costs for increasing planning buffers. If these cost rates slightly change, it does not have too big an influence on the overall objective value and optimal planning buffers might well remain constant. Without changing planning buffers, the throughput times are also constant and there is no need to invoke the inventory allocation again.

Convergence of total cost As a relaxed version of the first stopping criterion, we do no longer require that no changes are made to the decision variables any more, but that the sum of both objective function values converges. *Conversion* can be defined in different ways, e.g. if the value remains constant over a certain number of iterations or does not improve beyond a certain percentage threshold.

Maximum number of iterations In the unlikely case that none of the stopping criteria mentioned is ever met, a maximum number of iterations should be defined to guarantee the termination after a finite period of time.

This sequence shows how the two separate optimisation models and solution approaches presented in this work can be connected to obtain a basis for the assessment of assortment-related costs. Following the sequence described in the flowchart, a production and distribution network model or scenario can be built and optimised so that all parameters required for the cost assessment are determined nearly optimally. The interdependencies between the single steps are resolved by a repeated invocation, which has been shown to lead to a globally near-optimal configuration after a limited number of iterations.

CHAPTER 6

Application and Validation via Examples

We have the facts and
we're voting 'yes'.
Death Cab for Cutie

The next Section 6.1 gives a short overview of the software developed to test the concepts presented in Chapter 5. A more detailed description of the technical implementation can be found in appendix A. The remainder of this chapter provides a validation and exemplary application of these concepts in three steps: Firstly, Section 6.2 defines the example assortments and scenarios used. Secondly, Section 6.3 then describes and discusses results of the analyses performed. Finally, Section 6.4 provides some details on the performance of the optimisation methods as observed during the example application.

6.1 Implementation

For this exemplary application and validation, the entire concept presented in Chapter 5 was implemented in a the software tool *Complana* (COMPLexityANAlyser). It provides a graphical user interface that supports the entire analysis process with the following steps:

Data import and management The data required to build the production and distribution network models can be imported into the Complana database from an SAP ERP system. Import functionality is provided for material master and accounting data, production and sales locations with relevant customer service and cost parameters, demand and forecast accuracy data, bills of material, routings and production process steps with the corresponding data on setup and scrap times and quantities. Existing data can be viewed and edited to allow quick adaptations of the relevant parameters.

Model building and management The automatic generation of production and distribution networks can be invoked with a set of end products as the main input. Generated models are stored in the database along with descriptive meta-information. Existing models can be edited in terms of adapting demands and forecast deviations, either individually for single items or collectively for sets of selected items.

Scenario building and management For a selected baseline model, scenarios can be generated from a scenario definition. The software tool supports the user by automatically calculating the conversion factors from material master data and taking care of peculiarities like currency conversions in case the replacement material is valued in a currency other than the original material. To facilitate the selection of sets of materials to be replaced, standard ABC/XYZ analyses can be carried out on a selected model. Scenarios can also be edited just like the corresponding models.

Visualisation Models and scenarios can be visualised as networks. This provides the means visually to inspect the production and distribution relations and the data at each individual item. Furthermore, it helps to compare the complexity of different assortments by visualising their items and relations between them.[1]

Optimisation The optimisation methods can be parameterized and invoked for all models and scenarios in the database. The optimised models and scenarios are stored back into the database afterwards and a detailed result report is written into a spreadsheet file.

Comparative analysis Optimised models and scenarios can be used for comparative analyses. For pairs of an optimised model and corresponding scenario, a cost comparison is carried out and the results are visualised in different types of charts. Information about the changes in the number of materials at different locations, on different production process steps and of different material types is calculated.

1 See Definition 2.5.

6.2 Specification of Example Models

In order to show the practical applicability for real-world business problems, this validation uses example cases based on the assortment of an international household product manufacturer. The considered company produces, distributes and sells mechanical household products in different parts of the world. Several specialised production sites each produce a certain product range and supply a number of sales locations, which in turn manage sales to customers in their market region. Such a market region usually comprises a certain country or set of adjacent countries. Customer orders are fulfilled by shipment of goods from dedicated sales warehouses that are typically situated close to the sales locations. Customers include retail stores and wholesalers as well as cleaning companies and hotels, thus separate products for consumer and professional business divisions are available. The supply chain employs decentralised control, i.e. sales and production locations each manage stock-keeping, purchasing and distribution autonomously, though integrated via a central material coordination system.

6.2.1 Product Assortments

We consider two baseline assortment models M_1 and M_2, which are both based on the household cloth products of an international household product manufacturer.

M_1 - **window cloths** The first model was built from all window cloths end products offered at any sales location.

M_2 - **complete cloths assortment** The second model was built from all cloths end products offered at any sales location. The first model is therefore a subset of this model.

This particular product category was chosen as it has grown for many years and therefore suffers from particularly high complexity. The fact that both the entire cloth assortment as well as a smaller subset are considered separately allows validation of different aspects of the methods developed. Due to its manageable size, the first model M_1 allows a detailed inspection of the results to detect any unexplainable output as well as a detailed sensitivity analysis to ensure the model reacts as expected. The second example model M_2 is in turn used to show the feasibility of the methods for large problem instances. Table 6.1 shows an overview of the assortments considered. Clearly model M_2 comprises many more elements in all respects; in particular it also comprises some locations and production process steps that are not used for the production and distribution of window cloths in M_1.

Table 6.1: Overview of example assortment models

	M_1	M_2
No. of materials	293	2006
thereof end products	78	588
No. of items	419	3330
finished material items thereof	195	1810
semi-finished material items thereof	66	649
packaging material items thereof	115	764
raw material items thereof	43	107
No. of locations	23	27
No. of PPS	10	15

Almost all cloth end products are produced at a single production location that externally procures raw and packaging materials and performs all required steps to produce the end products. The end products are then distributed to the sales locations from where they are sold to customers. In some cases semi-finished products, bulk goods or finished products on a non-TSU packaging level are shipped to sales companies and simple converting activities are carried out at these sales locations. The production costs of these converting activities are not considered in our model since the converting processes mainly comprise manual assembly and packaging activities that incur only negligible setup and scrap costs. Moreover, their consideration would even further increase the data requirements with information about the production processes at each of the converting sites. Technically this means that there are no production process steps defined at any location except the considered production site. We do, however, consider the related BOMs at these locations and use them to determine product structures just as for the actual production location.

Figure 6.1 illustrates the distribution structure of the cloths assortment as derived from model M_2. At each location, the number of cloth materials handled at that location is shown in a small bar graph, separated according to material type.[2] From the production location in Germany, mainly finished products are distributed to a number of sales locations abroad. In some cases further transshipment is carried out at the sales location to consolidate transport for countries with small sales volumes.

2 Note that these bars have a different scaling for the production and sales locations.

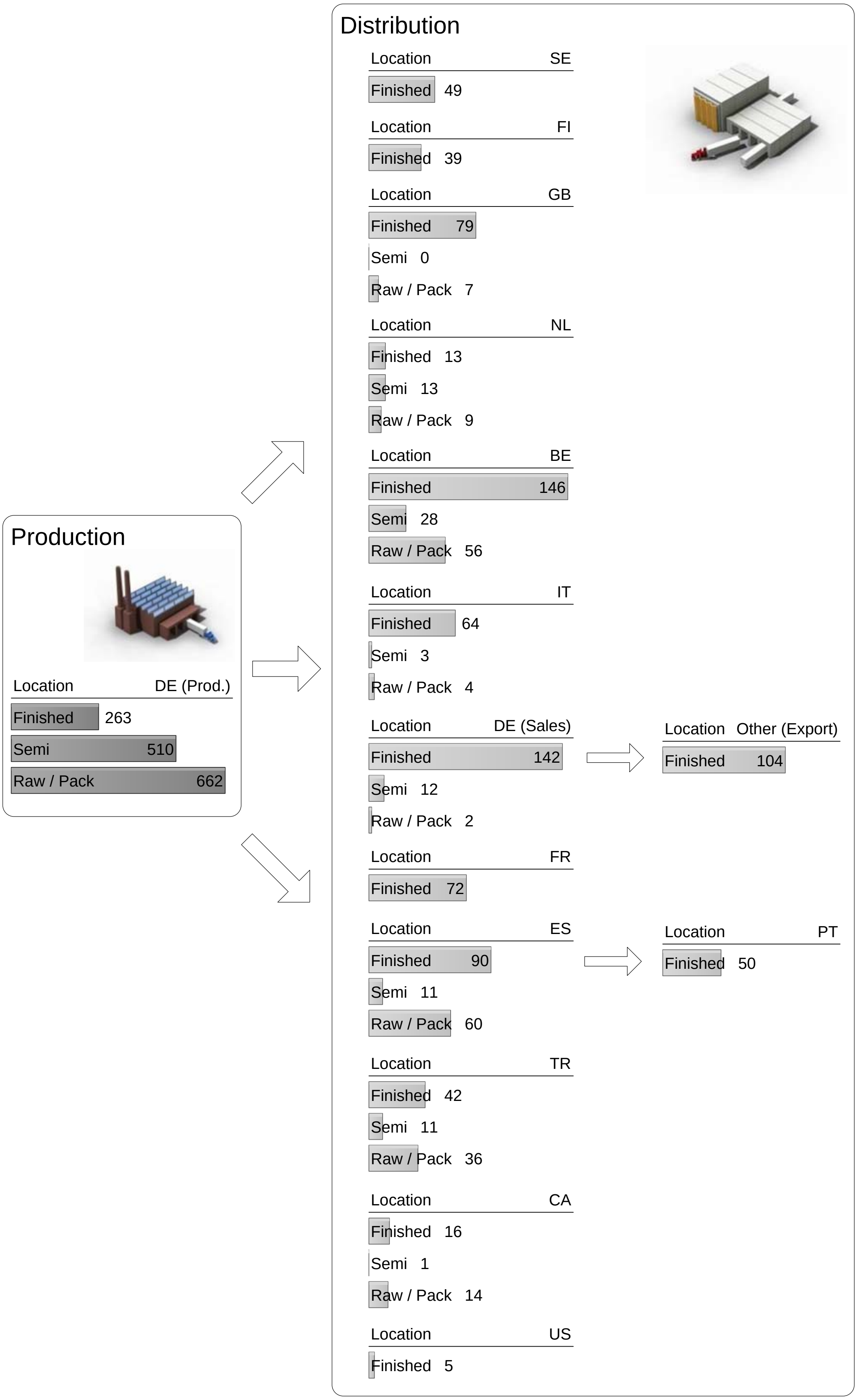


Figure 6.1: Distribution structure of cloth assortment (model M_2)

6.2.2 Alternative Assortment Scenarios

For each assortment model, several alternative assortment scenarios are defined. Table 6.2 summarises the scenario definitions and shows the changes made to the network by the scenario application in terms of materials discontinued. For each scenario, the number of discontinued end products and the resulting numbers of discontinued items of different types are shown both as absolute and relative figures, the latter compared with the corresponding baseline model.

The scenarios for M_1 represent different assortment reduction strategies on the end product level, based on the demand characteristics of the products. As a basis for scenario definitions, the demand series of end products at the sales locations are analysed with standard ABC and XYZ analyses. The parameters for the ABC analysis are chosen such that the A, B and C products contribute 80%, 15% and 5% respectively to the total cumulative demand. For the XYZ analysis, the classification of an item i to one of the classes is made on the basis of the coefficient of variation cv_i of its demand distribution. The intervals are chosen as $cv_i \leq 0.5$, $0.5 < cv_i \leq 1.0$ and $1 < cv_i$, for classes X, Y and Z respectively. The scenarios for M_1 are defined with the goal of finding out what items cause the highest complexity-related costs. Scenario S_1 discontinues all products from class C. Scenario S_2 additionally discontinues the class B products with medium demand volumes. S_3 analyses the effects of standardisation of A products with high demands. Analogously to the scenarios on the basis of the ABC analysis, S_4 and S_5 are based on the XYZ analysis and discontinue those products with volatile and thus hardly predictable demands. S_6 describes the largest standardisation possible in that model with of a reduction to a single end product which is then sold at all locations.

For the full cloth assortment model M_2 it is to be noted that it comprises cloths from different categories like window cloths, floor cloths and micro fibre cloths, among others. Since cloths of different types may not substitute each other, standardisation to one end product is made only within these categories in the first scenario S_1, leaving one standard product per category. While this may still be a reasonable example from practice, scenario S_2 finally considers the hypothetical case of standardisation to a single multi-purpose cloth. Although it is practically unrealistic to reduce such a complex assortment with different types of product to one single product, it provides insights into what fraction of total costs is caused by the existing assortment complexity.

The term *replace* in the scenario definitions means that all affected products are discontinued and replaced by one newly-added replacement product. For this newly-

Table 6.2: Scenario definitions

Name	Scenario description	Input: discontinued end products	Scenario application result: total no. discontinued items	discontinued finished material items	discontinued semi-finished material items	discontinued raw and pack. material items
M_1/S_1	Replace all C end products with the C end product with the highest total demand	53 67.95%	173 41.29%	103 52.82%	17 25.70%	53 33.54%
M_1/S_2	Replace all C and B end products with the B end product with the highest total demand	66 84.62%	256 61.10%	138 70.77%	33 50.00%	85 53.80%
M_1/S_3	Replace all A end products with the A end product with the highest total demand	9 11.54%	46 10.98%	25 12.82%	4 6.06%	17 11.04%
M_1/S_4	Replace all Z end products with the Z end product with the highest total demand	44 56.41%	110 26.25%	77 39.49%	6 9.09%	27 17.09%
M_1/S_5	Replace all Z and Y end products with the Y end product with the highest total demand	62 79.49%	218 52.03%	134 68.72%	21 31.82%	63 39.87%
M_1/S_6	Replace all end products with the end product with highest total demand	77 98.72%	346 82.58%	176 90.26%	46 69.70%	124 78.48%
M_2/S_1	Replace end products within each product category with the end product with highest total demand in that category	570 96.94%	2788 83.72%	1533 84.70%	511 78.74%	744 85.42%
M_2/S_2	Replace all end products with the end product with highest total demand	587 99.83%	3287 98.71%	1788 98.78%	645 99.38%	854 98.05%

introduced product, all master data and the production BOM are taken from the material with highest total demand in the set of replaced products. In the scenarios all demands of discontinued products are assumed to be transferred to the respective replacement products, i.e. we do not consider any lost sales. This is reasonable to assure a direct comparability of the absolute costs indicated by the cost assessment.

During the calculation of demand for the replacement products we also consider different packaging sizes for discontinued and replacement products. The software tool automatically calculates the conversion factor[3] for each replacement relation to transfer the correct demand quantities based on the number of individual products, i.e. on BCU level[4].

The possibility of visualising all models and scenarios provides the means for a mere visual comparison of the network complexity. Appendix B contains various visualisations of the example models used. For example, Figure B.1 shows the graphical representation of the smaller model M_1, where one can distinguish procurement, base production, converting and distribution stages in the network structure. As a comparison, Figure B.7 shows the visualisation of the related scenario M_1/S_6, where the complexity reduction becomes clearly visible. See the appendix for more such visualisation examples.

6.2.3 Production Processes

The products of the cloth product category are mainly produced at a specialised production facility in Germany. For both models, the same production system at that location is considered. It comprises a number of production process steps that can roughly be grouped into the categories base production and converting.[5] Production costs in terms of setup, scrap and production cycle stock costs are assessed for all of these production process steps. Figure 6.2 shows the considered production process steps, their categorisation as well as their logical sequence in the production process.

The base production starts with two production process steps that create the paint and fibre mixtures from the raw materials. These production steps are of less interest when considering assortment complexity, as they only do the preliminary work for the subsequent production lines in terms of composing the raw materials according to the corresponding recipes and thus operate fully in line with the demands of these subsequent process steps. At the actual base production stage these fibres and chemical mixtures are compacted to a basic raw cloth produced on large rolls. For these operations three production lines (Line 1, 2, 3) are used, each mainly for one particular cloth type. Some of the resulting raw cloths are further refined by applying additional

3 See Section 5.1.2.1.
4 See Section 2.2 for a description of packaging levels.
5 See Section 2.1.2.

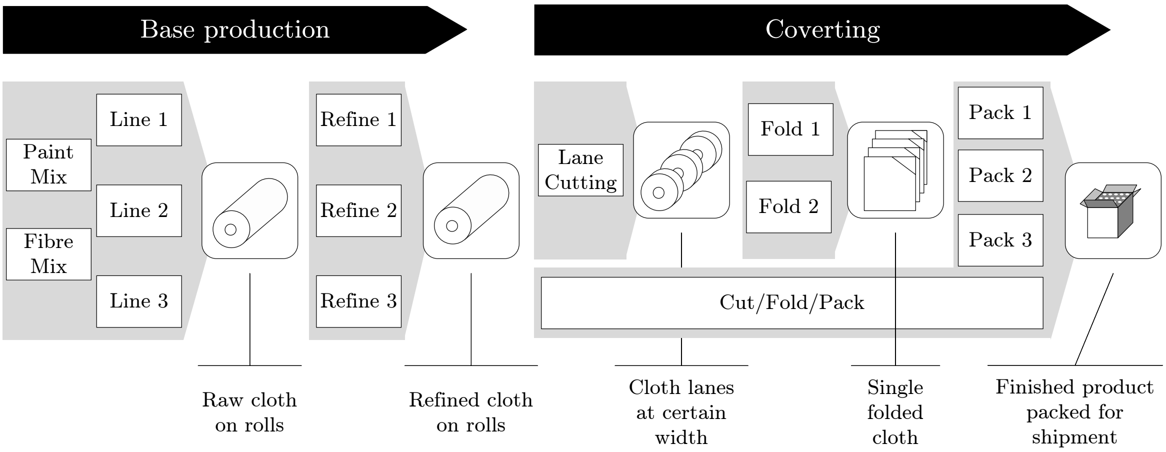


Figure 6.2: Production process steps

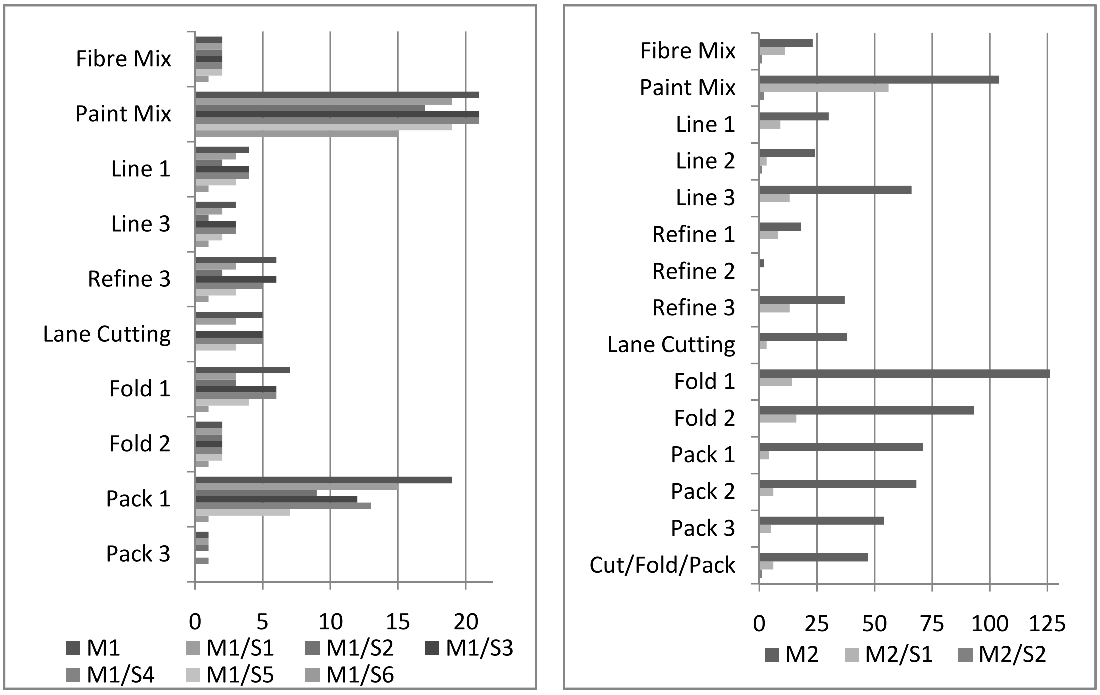
coatings, prints and imprinting structures into them. Depending on the type of refinement, one of the three alternative machines (Refine 1, 2, 3) is used. The result of this step then is the finished cloth material, again produced in large rolls.

In the converting stage, the wide rolls are first cut into smaller lanes on a separate production process step (Lane Cutting). These lanes then form the input components to two cutting and folding machines (Fold 1, 2), which output the finished unpacked cloth. In the final packaging step, these cloths are then packed into foil and then into cases in certain batches for shipment to the sales locations, which is done on three packaging machines (Pack 1, 2, 3). Apart from the described separate cutting, folding and packaging steps, one additional fully automated production process step (Cut/Fold/Pack) is considered, which is capable of performing all converting operations in-line, i.e. from the refined cloth on rolls to the packaged end products on TSU level.

For each of these production process steps, setup, scrap and cycle stock costs are assessed on the basis of optimised planning buffers and planned production quantities. The observed cost developments are based on the changing number of materials processed in each production process step in each scenario. Table 6.3 shows the number of materials processed on each individual production process step in the baseline models and all related scenarios. Figures 6.3(a) and 6.3(b) visualise these numbers for models M_1 and M_2 and their scenarios, respectively. It can be seen that for the window cloth assortment, only 10 out of the 15 production process steps are required.

Table 6.3: Number of materials processed on each production process step

PPS	M_1	M_1/S_1	M_1/S_2	M_1/S_3	M_1/S_4	M_1/S_5	M_1/S_6	M_2	M_2/S_1	M_2/S_2
Fibre Mix	2	2	2	2	2	2	1	23	11	1
Paint Mix	21	19	17	21	21	19	15	104	56	2
Line 1	4	3	2	4	4	3	1	30	9	-
Line 2	-	-	-	-	-	-	-	24	3	1
Line 3	3	2	1	3	3	2	1	66	13	-
Refine 1	-	-	-	-	-	-	-	18	8	-
Refine 2	-	-	-	-	-	-	-	2	-	-
Refine 3	6	3	2	6	5	3	1	37	13	-
Lane Cutting	5	3	-	5	5	3	-	38	3	-
Fold 1	7	3	3	6	6	4	1	126	14	-
Fold 2	2	2	2	2	2	2	1	93	16	-
Pack 1	19	15	9	12	13	7	1	71	4	-
Pack 2	-	-	-	-	-	-	-	68	6	-
Pack 3	1	1	1	-	1	-	-	54	5	-
Cut/Fold/Pack	-	-	-	-	-	-	-	47	6	1



(a) M_1 and related scenarios (b) M_2 and related scenarios

Figure 6.3: Number of materials processed on each production process step

6.2.4 Parameter Settings

Table 6.4 summarises the most important parameters used in the model and scenario generation and explains from where the corresponding values are derived. In order to base the analysis on real-world data and realistic assumptions, most data were taken from the practice of the household product manufacturer considered. The parameters are set according to company policy or are derived from historical data as described below.

Table 6.4: Model generation parameter data sources

Parameter	Data source
Demands	Historical demands observed at the sales locations for each month from January to December 2008.
Forecast deviations	Historical forecast deviations observed at the sales locations for each month from January to December 2008. The forecast deviations are available as mean absolute deviations, from which forecast error distributions are derived as described in Section 5.1.1.2.
Service levels	The service levels for all items are set to the internal target service level aspired for customer orders at the sales locations.
Inventory cost rates	Inventory costs are derived for each material at each location on the basis of the internal average cost of capital and the warehousing cost rates for one pallet at each location.
Setup cost rates	Setup cost rates are calculated based on setup times, labour and machine cost at the production site. On some occasions, average costs had to be used due to insufficient data availability. This does not affect the validity of the presented results or the general correctness of the presented methods.
Scrap cost rates	Scrap cost rates are derived from average scrap quantities per production run on the different production process step and the corresponding material values per basic unit.

6.3 Application and Results

6.3.1 Cost Effects in Inventory Allocation

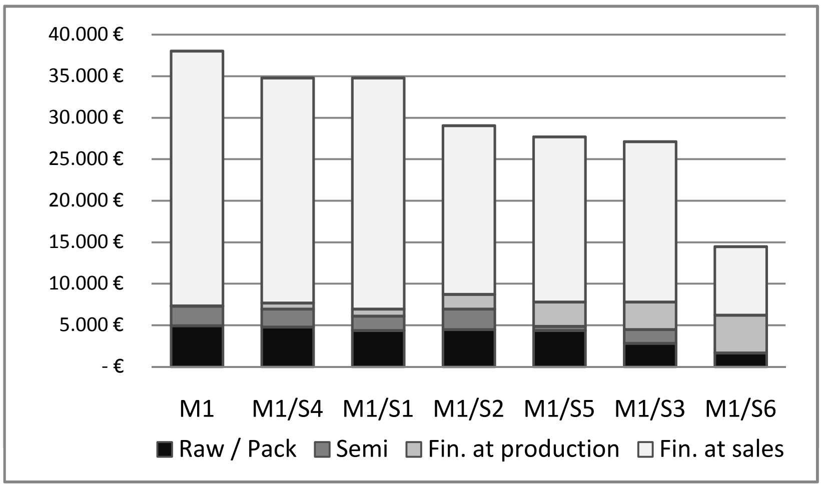
This section presents the results of the inventory allocation for the two baseline models and their corresponding scenarios and discusses the cost changes per scenario on an aggregate level. For one selected scenario a more detailed analysis is provided. The complete numerical results are shown in Table 6.5. For each baseline model or scenario, it shows the inventory costs in the optimised production and distribution network, additionally broken down according to material type. For each scenario, it also provides the relative change compared with the cost of the corresponding baseline model, again for each material type and scenario.

Table 6.5: Comparison of inventory cost

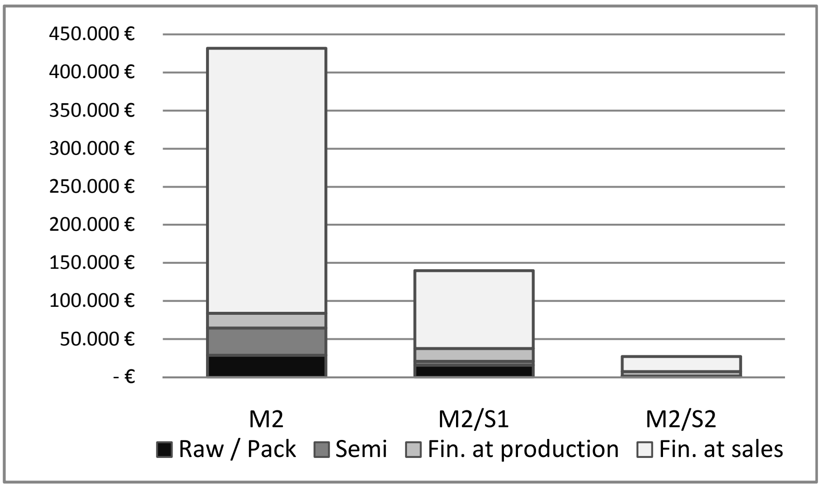
	Inventory cost [€]				Relative change [%]			
Model	Finished	Semi	R./P.	**Total**	Finished	Semi	R./P.	**Total**
M1	30,715.14	2,388.37	4,935.87	**38,039.38**				
M1/S1	28,657.01	1,734.58	4,378.80	**34,770.40**	-6.70	-27.37	-11.29	**-8.59**
M1/S2	22.085,66	2.475,29	4,472.50	**29.033,44**	-28,10	3.64	-9.39	**-23,68**
M1/S3	22,616.20	1,635.71	2,852.00	**27,103.91**	-26.37	-31.51	-42.22	**-28.75**
M1/S4	27,817.16	2,189.80	4,784.21	**34,791.17**	-9.44	-8.31	-3.07	**-8.54**
M1/S5	22,840.53	504.11	4,363.92	**27,708.56**	-25.64	-78.89	-11.59	**-27.16**
M1/S6	12.797,18	0,00	1.666,36	**14.463,54**	-58,34	-100.00	-66.24	**-61,98**
M2	367.329,11	35.770,81	28.867,32	**431.967,25**				
M2/S1	119.212,57	5.199,66	15.701,32	**140.113,55**	-67,55	-85,46	-45,61	**-67,56**
M2/S2	25.601,97	84,77	1.661,19	**27.347,93**	-93,03	-99,76	-94,25	**-93,67**

Figure 6.4(a) visualises the inventory cost data for M_1 and all corresponding scenarios, sorted by their total cost value. It first can be noted that all forms of standardisation lead to cost savings in the area of inventory management. The scenarios S_4 (Z materials) and S_1 (C materials) lead to similar changes in total inventory costs. This can be explained by closer comparison of these product sets, which reveals that they coincide at $74.\overline{074}\%$[6]. Analogously, scenarios S_2 and S_5 with the combined replacements of all

6 That is, 40 out of the 45 Z products are also part of the set of C products with 54 elements.



(a) Comparison for M_1



(b) Comparison for M_2

Figure 6.4: Inventory cost comparisons

C+B and Z+Y materials show similar total savings, where in this case the material sets coincide at 85.07%[7].

Despite the fact that there are considerably fewer A products than C and B products combined, assortment reduction by all materials in that class yields comparable and even slightly higher savings. The same applies in the comparison with the combined replacement of all Z and Y products. This is mainly due to the higher demand quantities for the discontinued A products, which result in much higher theoretically required safety stocks and thus higher savings potential, even for a small number of discontinued A products. However, it has to be noted that in practice, average inventories for C and

7 That is, 57 out of 63 Z and Y products are also part of the set of C and B products with 67 elements.

sometimes B products often turn out to be higher than the theoretically calculated levels from the model. In this case the savings potential in a real-world application may be even greater than indicated here. Reasons for higher inventory levels for C products include additional restrictions like minimum order sizes or psychological factors that cause planners keep a higher level of base stock. Furthermore, our model assumes that safety stock levels can be adapted on the level of mid-term periods, which can be a practical problem for products with minimum order sizes or production batch sizes and sporadic demand.

Analogously to the analysis for M_1, figure 6.4(b) depicts the results of the inventory allocation for M_2 and the corresponding scenarios. Notably, the cost change of scenario M_2/S_1 coincides very well with the changes calculated for the corresponding scenario M_1/S_6 of the window cloth model (67.56% versus 61.98%), which supports the validity of this figure. It indicates that the window cloths scenario is representative also of the other cloth product categories, as the joint standardisation within each such category yields approximately the same savings ratio. In the scenario with only one multi-purpose cloth for all application types (M_2/S_2), almost all inventory costs can be eliminated. While this change is admittedly unlikely to be implemented in practice, it shows that the model correctly represents the fact that a production and distribution system with only one standard, fast-moving product can operate almost without any inventory.

It is also notable that the allocation and type of inventory vary with different degrees of standardisation. Generally, increasing standardisation also increases the ratio of inventory held at production sites, which manifests itself in increasing inventory costs for raw and packaging materials, semi-finished materials and especially finished materials held at production locations. In the baseline models there are practically no or very little finished material inventories at the production site, as there is almost no item commonality in the diversified assortment and therefore inventories are better kept locally at the sales locations to reduce service time to the customers. This ratio changes with increasing standardisation. In scenario M_1/S_6, 35.59% of finished material inventories are held at the production location and inventories at the production location account for 50.01% of total inventory costs.

This effect can be traced back to the risk pooling effects of assortment standardisations.[8] With increasing item commonality, forecast deviations decrease upstream in the network and inventories at early stages become preferable. A detailed analysis

8 See Section 3.2.4 for a description of risk pooling and Section 5.1.1.2 for a description of how these effects are quantified in the underlying model.

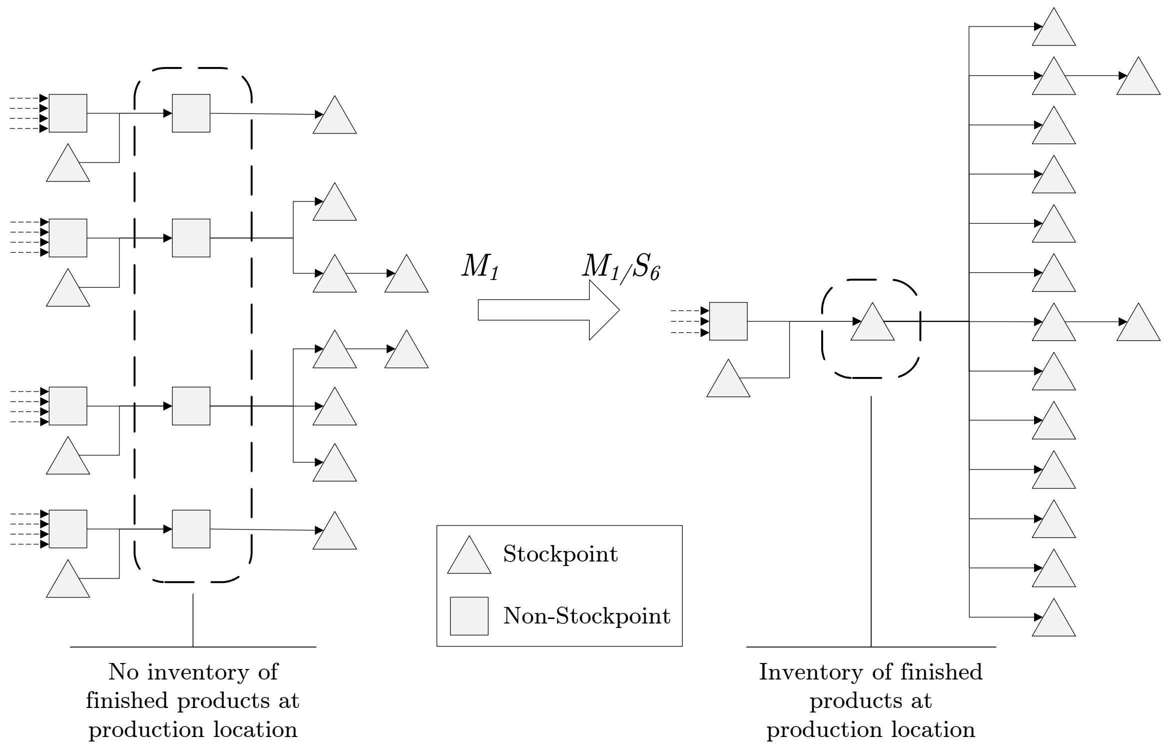


Figure 6.5: Introducing additional stockpoints for finished products

shows that this is especially the case for end products. Replacing various variants of a certain product with one standard product makes it favourable to keep inventories of the latter at the production location to decrease replenishment lead times to sales locations. The additional costs for inventory at the production location are outweighed by the savings generated at sales locations. Figure 6.5 shows a detail of the network visualisations of M_1 and M_1/S_6. On the left, inventory for each specialised product variant is only held at the sales locations, whereas on the right, an additional stockpoint is introduced for the newly-added finished product at the sales location, which reduces replenishment lead times at the sales locations and allows them to lower their safety stock levels.

Further analysis of the cost changes in the network supports this conclusion. Figure 6.6 illustrates the inventory cost changes for each location. All sales locations experience a cost decrease, partly due to standardisations in their local assortments, but mainly due to shorter replenishment lead times for the new replacement product. The only location that faces an increase in inventory cost is the main production location. However, the figure shows that this is by far outweighed by the savings generated, such that a total cost decrease of 61.98% remains. Furthermore, this cost increase at the production location is practically negligible $(4,45\%)$. This is due to the fact that the increased

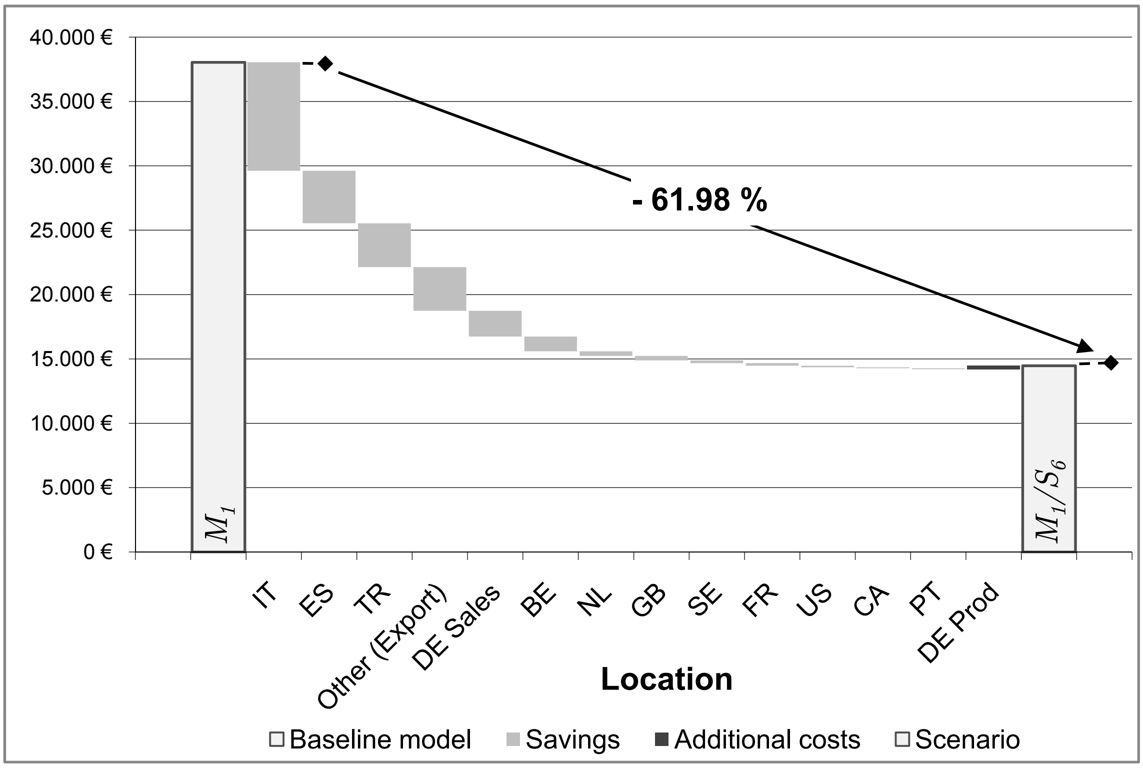


Figure 6.6: Inventory cost changes on location level for M_1/S_6

cost for additional finished material inventory are almost completely outweighed by the inventory reductions for semi-finished, raw and packaging materials in the production area, which result from the fact that many of these materials are no longer present in the scenario.

6.3.2 Cost Effects in Production Execution

For all baseline models and scenarios, the optimisation of planned production quantities and planning buffers provides insight into the development of production costs as a response to assortment changes. Table 6.6 shows the resulting setup, scrap and cycle stock costs for all models and scenarios in their optimal configuration. For each scenario, the relative changes in comparison with the corresponding baseline model are also indicated.

Figures 6.7 summarise the cost changes in a bar chart, separated by cost types and ordered by their total cost value. Comparing the order in figure 6.7(a) with the corresponding figure 6.4(a) obtained for the cost assessment in the area of inventory management, we note that it is identical except that M_1/S_3 has moved and now appears

Table 6.6: Comparison of production execution costs

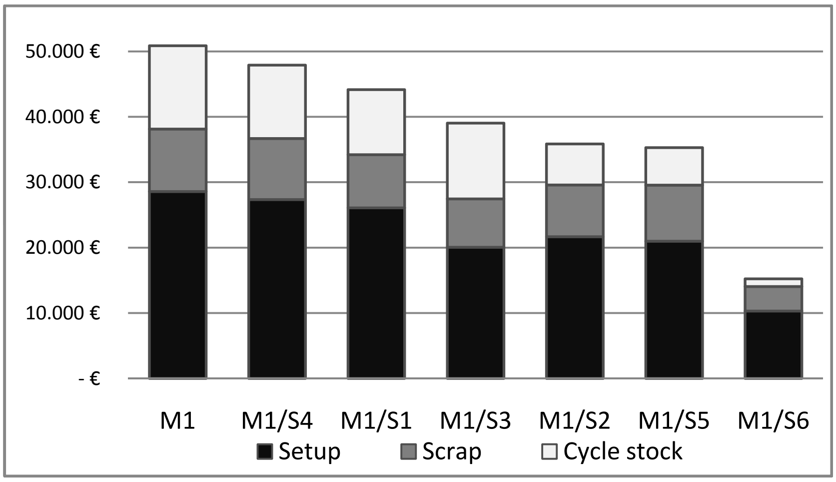
	Production cost [€]				Relative change [%]			
Model	Setup	Scrap	Cycle	**Total**	Setup	Scrap	Cycle	**Total**
M1	28,608.94	9,546.90	12,735.13	**50,890.97**				
M1/S1	26,100.67	8,144.05	9,946.22	**44,190.94**	-8.77	-14.69	-21.90	**-13.17**
M1/S2	21,684.11	7,917.24	6,286.56	**35,887.91**	-24.21	-17.07	-50.64	**-29.48**
M1/S3	20,071.41	7,399.14	11,591.51	**39,062.06**	-29.84	-22.50	-8.98	**-23.24**
M1/S4	27,377.88	9,338.92	11,212.30	**47,929.10**	-4.30	-2.18	-11.96	**-5.82**
M1/S5	21,011.14	8,580.18	5,716.77	**35,308.09**	-26.56	-10.13	-55.11	**-30.62**
M1/S6	10,315.17	3,760.30	1,161.29	**15,236.77**	-63.94	-60.61	-90.88	**-70.06**
M2	285,906.71	110,178.83	226,426.34	**622,511.88**				
M2/S1	96,415.46	37,971.70	29,457.20	**163,844.36**	-66.28	-65.54	-86.99	**-73.68**
M2/S2	2,259.51	804.92	231.76	**3,296.19**	-99.21	-99.27	-99.90	**-99.47**

before M_1/S_2 and M_1/S_5, i.e. now yields lower savings than these two scenarios. This is one indicator for the observation that in the area of production costs, the relative importance of the two factors *number of materials discontinued* and *demand value of the discontinued materials* shifts towards the factor *number of materials discontinued*. While inventory costs are heavily influenced by the demand rates of the materials, demand-invariant fixed costs for setups and scrap have a higher importance here and result in higher savings obtained by a standardisation of both C and B products than those obtained by a standardisation of the high-volume A products. Furthermore, figure 6.7(a) shows that the commonalities in the scenario inputs again lead to a similar cost behaviour for scenarios S_1/S_4 and S_2/S_5.

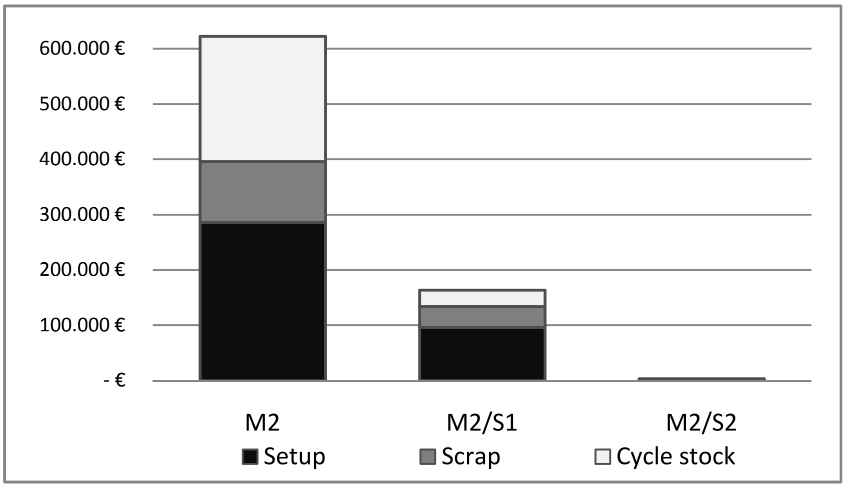
All costs shown here are based on the optimal choices for planning buffers for the production process steps considered in each respective model and scenario.[9] For each production process step s, the set of candidate planning buffers was defined as $pb_s \in \{0, \ldots, 6\}$. From this set, the planning buffer that yields minimal total cost was chosen and consequently also used in the calculation of the throughput times TT_i of all materials processed on s.

In order to illustrate the cost behaviour when changing the planning buffer at a single production process step, Figure 6.8 shows the production costs and planning buffer penalties for production process step 'Refine 3' in the base production area for the large baseline model M_2. Each bar indicates the total cost as well as the proportion of each

9 Compare Table 6.3 for an overview of the usage of the production process steps in the models and scenarios.



(a) Comparison for M_1



(b) Comparison for M_2

Figure 6.7: Comparison of setup, scrap and production cycle stock costs

constituent. One can clearly see the anti-proportional developing of setup costs and planning buffer penalties as planning buffers increase. This developing quantitatively represents the trade-off between long planning buffers to reduce setup costs by sequence optimisation versus the additional inventories made necessary by the longer throughput times. In this case, the total cost is minimal for $pb_s = 5$, which is therefore selected as the optimal planning buffer for this production process step in baseline model M_2. The same consideration was made in the optimisation for each production process step in each model and scenario.

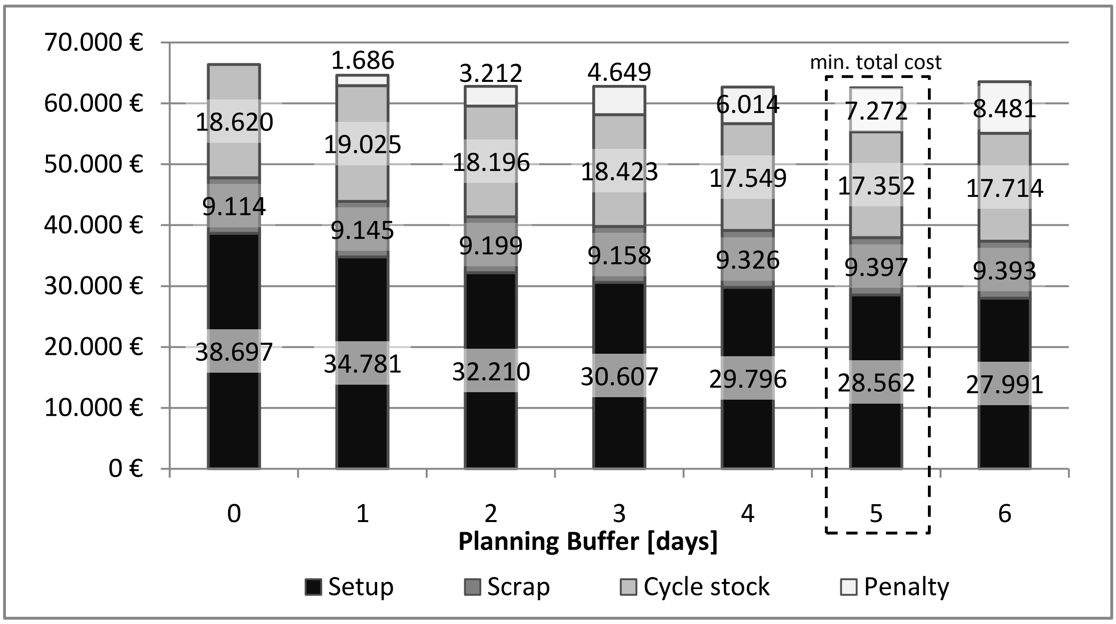


Figure 6.8: Development of production costs for production process step 'Refine 3' in model M_2

6.3.3 Conclusions: Cost Effects of Assortment Changes

From this experimental evaluation, we can derive some general conclusions about the cost effects of assortment changes. In the area of inventory management, we can identify considerable differences in the ratio of the number of products discontinued and the obtained savings when considering products with different demand volumes. This ratio is, according to the model, much higher for large volume products. However, in practical applications it might well be that the savings in inventory costs turn out to be higher due to comparably higher safety stock levels before standardisation. The advantage of the replacement of products with highly volatile demand distributions may be reduced by the likelihood that these products will also be low in demand and thus offer little potential for inventory reduction. Considering the analysis of execution production cost, the requirement to address the high volume products to obtain considerable savings is not as big as in the area of inventory management. This is mainly because setup costs are fixed costs per production order, which are also incurred for the production of low volumes.

Figure 6.9 combines the results obtained from the two analysis areas, showing the total relative cost savings per scenario and the composition of these savings. For each scenario the pie chart indicates what ratio of the total savings is due to what cost area. It can be seen that these distributions vary among the scenarios. However,

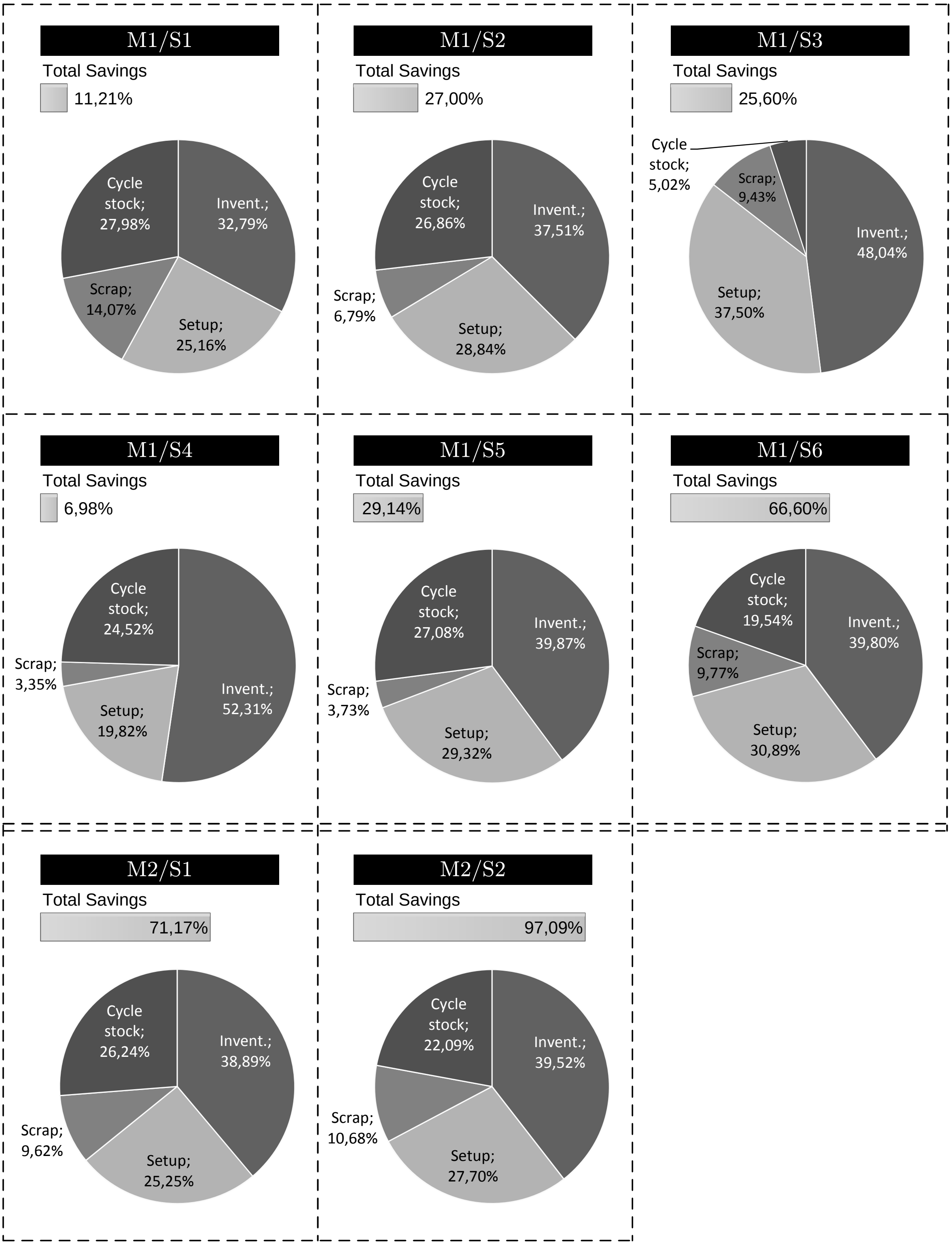


Figure 6.9: Total cost changes for all scenarios

these variations can be explained via the set of products replaced in each scenario. For example, the particularly small fraction of savings in cycle stock costs for M_1/S_3 may be explained by the fact that A products do not cause high cycle stock cost in the baseline model and therefore there is only little potential for any savings in this area.

In summary, it has to be noted that the multitude of influencing factors make it impossible to draw universal conclusions about the actual cost effects of assortment changes. Consequently, there exist no general statements about recommended assortment changes either. However, the results of the exemplary applications show how the methods developed in this work help to derive recommendations for concrete change scenarios. This goes in accordance with the initial motive for this approach, stating that such individual assessments provide much better decision support than general statements about cost developments with increasing numbers of product variants.

6.4 Performance of the Optimisation Methods

6.4.1 Inventory Allocation Heuristic

6.4.1.1 Configurations

The analyses and experiments were carried out on a laptop computer equipped with a Dual Core processor at 1.66 GHz and 2 GB RAM. The two processor kernels were never used concurrently, as the underlying software does not support parallel computations. Statements about the configurations and performance of the optimisation methods all refer to the optimisation runs carried out with the baseline models M_1 and M_2. This is sufficient, as all scenarios generally comprise smaller networks and therefore cannot represent more complex optimisation problems as the corresponding baseline models. The described configurations were used identically for the corresponding scenarios of each baseline model.

One of the most crucial decisions for the tabu search presented in Section 5.2.3 is the definition of neighbourhood sizes and compositions. Extensive tests with various configurations have shown that the best results are obtained with a neighbourhood composed of a limited number of solutions based on all the presented strategies. Table 6.7 shows the parameters for neighbourhood generation chosen for the two baseline models and all related scenarios.

Table 6.7: Neighbourhood definitions

Model	Random switches	S.E. switches	move downstream?
M_1	50	50	Yes
M_2	100	100	Yes

The neighbourhoods for both models use the stockpoint eligibility ratings to construct part of the neighbourhood solutions. While this helps quickly to find good first moves, it risks getting stuck in local optima, as the generated moves are all based on the static stockpoint eligibility ratings. Accordingly, randomly selected moves are also added to keep the search process exploring new areas of the solution space. The third set of moves for the neighbourhoods is based on the strategy of moving inventories downstream if there is no item commodity in the network structures. As the network structure remains unchanged during the optimisation, these items are identified only once before the optimisation starts and then stored in memory to avoid a repeated search in each iteration.

As a meta-heuristic, tabu search leaves many additional degrees of freedom to fine-tune the search process. For all optimisations described in this section, the following configuration was used.

Initial solution From the set of possible starting solutions described in Section 5.2.3.1, the one that worked best for the considered model was to preselect all end products as stockpoints. This can be reasonably explained from the application context since the sales locations provide immediate order fulfilment to their customers, which requires inventories of the finished products in most cases.

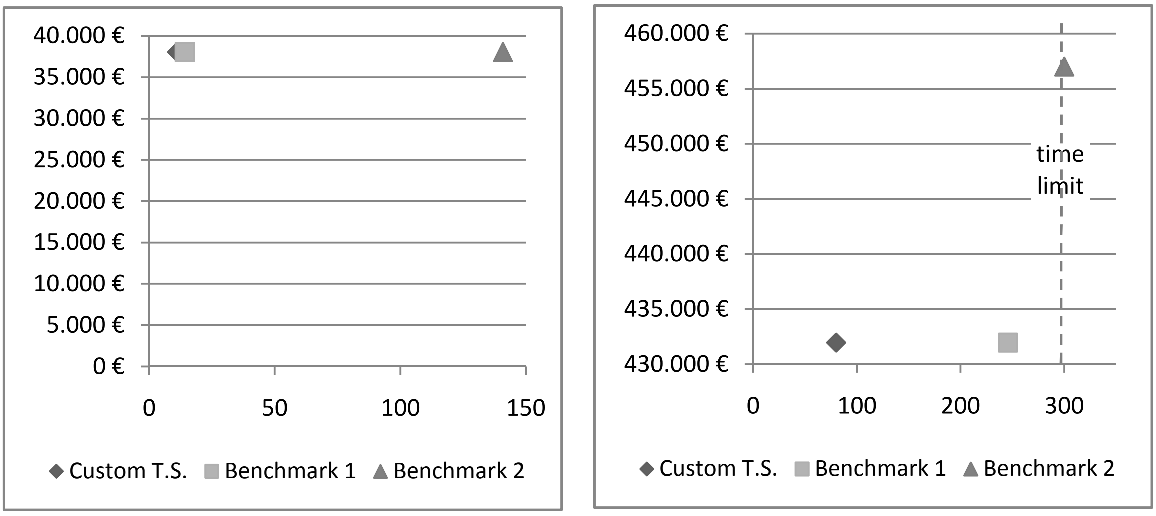
Tabu list The tabu list is defined to forbid moves that have already been performed during the tabu tenure. To assure this, the changed items and the direction of the stockpoint switches are stored in the tabu list for each selected move. The length of the tabu tenure were set to 100 and 200 for models M_1 and M_2, respectively.

Aspiration criterion To assure that the tabu list does not impede the discovery of improved solutions, the search employs the *best ever* aspiration criterion. When a move that is tabu would lead to a solution which is better than the current best solution, the move is permitted despite its tabu status.

Stopping criterion The search is bounded by a time limit, which was set to 300 seconds for all optimisation runs. However, in many applications the search converged long before this limit was reached. Accordingly, the convergence of the search process was used as an alternative stopping criterion and the search terminated if no new best solution was found during the latest 500 tabu search iterations.

6.4.1.2 Computational Performance

Figure 6.10 shows the runtimes and obtained objective values for the inventory allocation optimisation of the baseline models M_1 and M_2. For presentation purposes we restrict the description to the performance on the baseline models here, as all scenarios can be seen as subsets of these models[10] and thus do not pose any additional challenges for the optimisation methods. The figures compare the performance of the heuristic developed in Section 5.2 with a comparable tabu-search heuristic based on a simple one-switch neighbourhood. This alternative heuristic was developed as an implementation of the state of the art with respect to the optimisation methods used so far for similar inventory problems as described in Sections 3.2.2.2 and 3.2.3. Thereby, the effects of the enhanced neighbourhood definition is made visible. In addition, the effect of the partial solution recalculation as the second major enhancements of the new heuristic is analysed as we test the benchmark heuristic both with and without this feature.



(a) Performance for M_1 (b) Performance for M_2

Figure 6.10: Performance of the inventory allocation heuristic

10 Except for the newly-added replacement products.

The results show the effects of the network size on the performance of the different optimisation methods. For baseline model M_1, figure 6.10(a) shows that all methods find an equally good solution within the time limit. The advances of the new tabu search are reflected in the fact that it achieves faster convergence. This time advance is quite small for the benchmark heuristic that also uses the solution recalculation, which can be explained by the small network where both neighbourhoods always almost comprise all eligible items and thus the *intelligent* selection of stockpoint items cannot make a big difference. The consideration of the benchmark without the solution recalculation shows the runtime advantage obtained by this feature, as its runtime is more than 12 times longer than the runtime of our heuristic.

Figure 6.10(b) shows the analog results for the large model M_2. The larger network clearly increases runtimes, as the neighbourhood generation as well as recalculation of the objective value has to consider far more items. Our heuristic converges after 80.15 seconds with the best known objective value. The benchmark heuristic that employs the partial solution recalculation manages to find a solution of the same quality, but requires 245.82 seconds to converge, which is more than 3 times as long. This speedup is fully attributable to the enhanced neighbourhood definition. Considering the benchmark heuristic without the partial solution recalculation, we see that no solution of equal quality can be found within the time limit of 300 seconds and the best objective value deteriorates to 457,008.89.

In summary, the heuristic developed in Section 5.2 has proven to fulfil its requirements[11] during its use for the various optimisation instances. It provides acceptable solutions and runtimes for all tested practical instances. In particular, it generates good solutions even for extremely big networks like M_2 within few seconds or at most minutes. With respect to the absolute solution quality, we cannot guarantee its optimality, as a total enumeration of all solutions is impossible within reasonable time[12] and standard solvers cannot easily be used due to the problem's nonlinearity. However, the fact that the benchmark heuristic with an entirely different neighbourhood structure yields the same objective value may be interpreted as an indicator of the good quality, if not optimality of these solutions. This is evidence that the solutions are of sufficient quality to serve as the basis for a cost analysis.

The predominance of the new heuristic becomes more apparent with increasing network sizes. While the benchmark with the simple one switch neighbourhood can still

11 See Section 2.3.2.

12 As noted earlier, the solution space is of dimension 2^n where n is the number of items in the network. For the example models under consideration M_1 and M_2, this results in $2^{419} \approx 1.3538 \times 10^{126}$ and $2^{3330} \approx 2.6908 \times 10^{1002}$ possible combinations, respectively.

compete for M_1, its inferiority becomes visible in instances with larger networks. As intended, the heuristic is particularly suited to large networks that describe complex assortments. The configurations described in the previous section have proven efficient for the considered example models. It has to be noted however that the configuration of such heuristics is always problem-dependent and may have to be adapted for different problem instances to obtain competitive results.

6.4.2 Production Parameter Optimisation

6.4.2.1 Configurations

For the optimisation of planning buffers and planned production quantities, standard solvers for linear optimisation problems are used on a computer equipped with an Intel Core 2 Duo processor at 3.33 GHz and 8 GB of RAM. In particular, the commercial solvers CPLEX[13] version 11.0.1 and the free alternative GLPK[14] version 4.34 have been tested. The model is implemented using the MathProg modelling language[15], which provides a subset of the commercial AMPL language[16]. This subset is sufficient to represent the optimisation model defined in Section 5.3.1. This modelling language is used to maintain the flexibility easily to use different solvers. The GLPK solver natively supports MathProg models as its input. For the usage of CPLEX, concrete instances of the optimisation problem are generated in the CPLEX LP format with the help of GLPG and written to disk. The CPLEX solver is then invoked from within the Complana tool and instructed to read these files and solve the corresponding problem instances. In both cases the optimisation results are written to a file and parsed again by the Complana software. As one would expect, the commercial CPLEX solver showed significantly better performance and therefore only these numbers are presented in this validation. The solver was invoked with its default settings for mixed integer optimisation problems. In addition, a time limit of 15 minutes was set for each single optimisation problem instance, i.e. for each production process step and planning buffer candidate.

13 CPLEX is a state of the art solver for linear problems, developed and distributed by ILOG, an IBM company. See http://www.ilog.com/cplex/ for further details.

14 The GLPK solver is part of the Gnu Linear Programming Kit, which is an open-source project licensed under the Lesser Gnu Public Licence (LGPL).

15 Like the GLPK solver, MathProg is part of the Gnu Linear Programming Kit.

16 AMPL is a formal modelling language used to describe mathematical optimisation problems. See Fourer et al. (1993).

6.4.2.2 Computational Performance

Table 6.8 shows the average solution times for each model or scenario and the contained production process steps. The runtimes are averaged over the solution times for each planning buffer candidate and shown in milliseconds.

Table 6.8: Runtime of optimisation, averaged over all planning buffer candidates (in [ms])

PPS	M_1	M_1/S_1	M_1/S_2	M_1/S_3	M_1/S_4	M_1/S_5	M_1/S_6	M_2	M_2/S_1	M_2/S_2
Fibre Mix	2	3	2	2	2	2	2	12	6	1
Paint Mix	1	1	1	1	1	1	0	2	1	1
Line 1	25	80	10	92	33	6	2	3.44%	2,332	-
Line 2	-	-	-	-	-	-	-	0.46%	17	2
Line 3	13	22	2	10	23	2	2	9.42%	6,453	-
Refine 1	-	-	-	-	-	-	-	294	9	-
Refine 2	-	-	-	-	-	-	-	10	-	-
Refine 3	46	30	3	50	34	6	2	3.42%	4,734	-
Lane Cutting	28	13	-	35	31	16	-	1.04%	3	-
Fold 1	103	19	23	15	109	8	2	14.07%	3,564	-
Fold 2	3	3	3	3	3	3	2	10.08%	368	-
Pack 1	306	214	33	185	133	19	2	0.01%	12	-
Pack 2	-	-	-	-	-	-	-	1.18	7	-
Pack 3	2	2	2	-	2	-	-	1558	7	-
Cut/Fold/Pack	-	-	-	-	-	-	-	1.20%	13	2
Sum	**529**	**387**	**79**	**393**	**371**	**63**	**14**		**17,526**	**6**

The table shows that the majority of the optimisation problems are solved to optimality almost instantaneously. All problems related to the smaller model M_1 are solved on average in less that half a second. The total of the sums in the last row can be seen as the average solution time for the entire model (optimising all contained production process steps). Even for a large number of planning buffer candidates, total optimisation time would still be in the range of a few seconds for these problem instances. It however has to be noted that this solution time is required for each planning buffer candidate and thus has to be multiplied with the number of planning buffer candidates to obtain an estimate for the total computation time.

Considering the problem instances related to the larger model M_2, one can clearly see the effect of an increasing number of materials on the production process step. The extreme scenario M_2/S_2 with only one end product is solved instantaneously due to the very small number of materials. The scenario M_2/S_2 shows runtimes of up to 6.45 seconds for some production process steps, which is still a very quick response time. Only for the full model M_2 with the entire cloths assortment the solver does not find a proven optimal solution within the short time limit for some of the production process steps. In these cases, the table cells are marked grey and show the maximum gap of the solution to the optimal solution. For example, it shows that for production process step 'Line 1' the objective value of the best solution found within the time limit is at most 3.44% worse than that of the optimal solution. Considering these cases, we can see that in most cases the solutions found are very close ($< 5\%$) to the optimal solution. Only for 3 production process steps we observe gaps of 9.42%, 10.07% and 14.07%. For this exemplary analysis, we state that these solutions are sufficient to serve as the basis for the cost analysis. If more exact solutions are required, the time limit can be further extended.

It can be concluded that for practically relevant problem sizes, solutions can be obtained with the use of standard solvers in acceptable times. Even for the largest problem instances, the solutions show a sufficiently high quality within comparably short time limits. Given that this type of analysis is performed for non time-critical, strategic or tactical rather than operational decisions, there is also the possibility of accepting longer runtimes to obtain proven optimal solutions.

CHAPTER 7

Conclusions and Future Research

So keep moving, onward,
run through that open door.
Dredg

7.1 Summary and Conclusions

Assortment complexity is an ubiquitous topic in any industrial enterprise. Product proliferation and assortment complexity have increased impressively in the past as markets become more global and customer-oriented. This raises the importance of the managerial task of deciding the breadth and width of assortment.

Approaches to address this task in its entirety all fall short, as a true *optimal* assortment depends on far too many factors to be assessed precisely. Restricting the focus to the cost of assortment complexity in the areas of inventory management and production execution, we identify the lack of an integrated method to assess the cost of assortment complexity based on concrete assortment change scenarios. This approach promises to deliver more precise assessments than any monovariable model, defined e.g. on the number of products. At the same time it has more practical relevance since any assortment-related decision starts with some status quo and decisions are to be taken on concrete changes rather than on an entirely new assortment.

Such a method was presented in this work. The conceptual solution starts with the development of a formal model of a production and distribution network which captures all relevant characteristics of product and distribution structures as well as production and distribution processes. To facilitate practical usage, algorithms were formulated automatically to build such networks from data typically available in modern ERP systems. On the basis of this model, assortment change scenarios can be defined to describe theoretical assortment changes in terms of material additions and replacements at all production and distribution stages.

An inherent advantage of the cost assessment on the basis of concrete assortment change scenarios is that important parameters in the areas of inventory management and production operations can be adjusted to the new situation. This permits a more precise cost assessment since it can be based on a production and distribution system adapted to the new situation. This approach requires to address the determination of stockpoints with corresponding inventory levels as well as the determination of production planning buffers and planned production quantities as related optimisation problems.

For comparison of inventory costs, an inventory allocation model was formulated as a combinatorial optimisation problem on the basis of the production and distribution network model. This optimisation model captures the interrelation between the positions of stockpoints in the network and the replenishment lead times observed at each individual item. For the solution of this optimisation problem, a tabu search heuristic was developed that makes use of special domain knowledge to select promising subsets of items as stockpoint candidates to guide the search process.

For the comparison of production execution costs, a mathematical optimisation model was formulated to decide the cost-optimal planning buffers and to determine planned production quantities to minimise setup, scrap and production cycle stock costs. Since the resulting model is non-linear, approaches to transfer it into a linear model were discussed to allow the use of standard software. Aware of possible obstacles to the practical applicability of this method, we also addressed the question of cost parameter definitions in this context and provided methods to estimate setup cost developments for given planning buffers.

In order to validate both the individual methods as well as the overall approach, the entire concept was implemented in a software tool and validated in cooperation with an international household product manufacturer. This practical background of the validation assures that the input parameters are near to real-world application scenarios. On the basis of this data two assortments were analysed considering their current as-is status as well as various alternative scenarios. The developed optimisation models and solution methods proved to be feasible for real-world problem instances. The results have also proven to be valid, as far as this can be assured by sensitivity analyses. In this context, the experience of logistics and production managers could be used to assure the correspondence to reality of the numerical results and ensure that the model responds as expected to parameter changes. The validations showed how the developed method can be used to support managerial decisions on concrete assortment changes.

7.2 Limitations and Outlook on Potential Extensions

Any work that addresses a far-reaching problem like the effects of assortment complexity naturally has its limitations. First of all, this work does not provide a method to calculate an optimal assortment, as this is not believed to be a very promising approach at the current state of research. The question of finding an optimal assortment complexity is addressed indirectly by providing decision support via cost estimations for theoretical assortment changes.

As stated in Chapter 2, assortment complexity incurs cost in virtually every functional unit and every process of a company. Accordingly, the view on assortment-related costs had to be limited for the scope of this work, which was imposed by focusing on inventory management and production execution as functional areas with generalisable and quantifiable cost components. Further extension may include other general areas like product development (R&D), quality assurance or master data maintenance. However, the corresponding processes in these areas are probably harder to generalise to obtain a generic cost assessment, making it a rather case-specific task to describe the corresponding processes, assess the related costs and describe their relation to assortment complexity.

All results in terms of cost estimations can only be approximate. This is inevitable given that much data that the estimations rely on are assumptions about future developments, e.g. expected demand volumes and their distribution over time. Furthermore, the assessment of effects in production planning makes some simplifying assumptions about the production planning and scheduling processes. This opens up opportunities for further research by making the underlying models more detailed, i.e. either include additional information or decrease the aggregation level. For example, the assessment of production-related costs could be refined if it were based on more detailed demand data and also included production sequence planning to avoid the usage of average cost rates. Furthermore, there is additional loss of accuracy due to the usage of heuristic optimisation methods. However, the practical applications have shown that they can be assumed to be negligible in this application context.

This work focuses on consumer goods supply chains. The transferability to other industries is only partly given. Of the methods developed, the assessment of inventory costs may easily be transferred to other industries as the underlying assumptions are rather weak and apply to almost any production and distribution network with inventories in it. However, the assessment of production-related cost assumes certain characteristics of the production layout as found in consumer goods supply chains, making it less transferable to other industries. In summary, the method and results are limited to a certain type of supply chain or industry as outlined in Section 2.1, which definitely covers all consumer goods supply chains, both in food and non-food industries. Its extension to other industries forms another field of further research.

Bibliography

Adam, D., editor (1998). *Komplexitätsmanagement*, volume 61 of *Schriften zur Unternehmensführung*, Wiesbaden.

Adam, D. (2004). Controlling bei Komplexität. In Mertens, P., Bensberg, F., Grob, H. L., vom Brocke, J., and Schultz, M. B., editors, *Trendberichte zum Controlling*. Physica-Verlag, second edition.

Adam, D. and Johannwille, U. (1998). Die Komplexitätsfalle. In Adam (1998), pages 5–28.

Adebanjo, D. (2000). Identifying problems in forecasting consumer demand in the fast moving consumer goods sector. *Benchmarking: An International Journal*, 7(3):223–230.

Alfaro, J. A. and Corbett, C. J. (2003). The value of SKU rationalization in practice (the pooling effect under suboptimal inventory policies and nonnormal demand). *Production & Operations Management*, 12(1):12–29.

Alicke, K. (2005). *Planung und Betrieb von Logistiknetzwerken - Unternehmensübergreifendes Supply Chain Management*. VDI-Buch. Springer, second edition.

Anderson, P. (1999). Complexity theory and organization science. *Organization Science*, 10(3):216–232.

Axsäter, S. (2006). *Inventory Control*. International Series In Operations Research & Management Science. Springer.

Baker, K. (1977). An experimental study of the effectiveness of rolling schedules in production planning. *Decision Sciences*, 8(1):19–27.

Baker, K. and Peterson, D. W. (1979). An analytical framework for evaluating rolling schedules. *Management Science*, 25(4):341–351.

Baker, M. J., editor (2002). *The Westburn Dictionary of Marketing*. Westburn Publishers.

Barrett, R. T. and LaForge, R. L. (1991). A study of replanning frequencies in a material requirements planning system. *Computers Operations Research*, 18(6):569–578.

Benjaafar, S., Cooper, W. L., and Kim, J.-S. (2005). On the benefits of pooling in production-inventory systems. *Management Science*, 51(4):548–565.

Benton, W. C. (1991). Safety stock and service levels in periodic review inventory systems. *Journal of Operational Research Society*, 42(12):1087–1095.

Benton, W. C. and Srivastava, R. (1985). Product structure complexity and multilevel lotsizing using alternative costing policies. *Decision Sciences*, 16:357–369.

Biggs, J. R. and Campion, W. M. (1982). The effect and cost of forecast error bias for multistage production-inventory systems. *Decision Sciences*, 13:570–584.

Blackburn, J. D. and Millen, R. A. (1980). Heuristic lot-sizing performance in a rolling schedule environment. *Decision Sciences*, 11:691–701.

Blackburn, J. D. and Millen, R. A. (1982a). The impact of a rolling schedule in a multilevel MRP system. *Journal of Operations Management*, 2(2):125–135.

Blackburn, J. D. and Millen, R. A. (1982b). Improved heuristics for multistage requirements planning systems. *Management Science*, 28(1):44–56.

Blecker, T., Kersten, W., and Meyer, C. (2005). Development of an approach for analyzing supply chain complexity. In Blecker, T. and Friedrich, G., editors, *Mass Customization. Concepts, Tools, Realization. Proceedings of the International Mass Customization Meeting 2005*, pages 47–59, Austria. Gito Verlag.

Blum, C. and Roli, A. (2003). Metaheuristics in combinatorial optimization: Overview and conceptual comparison. *ACM Comput. Surv.*, 35(3):268–308.

Bräutigam, L.-P. (2004). *Kostenverhalten bei Variantenproduktion.* Deutscher Universitäts-Verlag.

Bregman, R. (1991a). Selecting among MRP lot-sizing methods for purchased components when the planning horizon is limited. *Production and Inventory Management Journal*, Q2:32–39.

Bregman, R. L. (1991b). An experimental comparison of MRP purchase discount methods. *Journal of Operations Research Society*, 42(3):235–245.

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods.* Springer Series in Statistics. Springer, second edition.

Buhr, M. and Klaus, G. (1975). *Philosophisches Wörterbuch*. Verlag das europäische Buch, Berlin, eleventh edition.

Burrows, B. L. and Talbot, R. F. (1985). The mean deviation. *The Mathematical Gazette*, 69(448):87–91.

Buzacott, J. and Shanthikumar, J. (1994). Safety stock versus safety time in MRP controlled production systems. *Management Science*, 40(12):1678–1689.

Carlson, R. C., Beckman, S., and Kropp, D. (1982). The effectiveness of extending the horizon in rolling production scheduling. *Decision Sciences*, 13:129–146.

Carlson, R. C., Jucker, J. V., and Kropp, D. H. (1979). Less nervous MRP systems: A dynamic economic lot-sizing approach. *Management Science*, 25(8):754–761.

Carlson, R. C. and Yano, C. A. (1986). Safety stocks in MRP systems with emergency setups for components. *Management Science*, 32(4):403–412.

Cassandras, C. G. and Lafortune, S. (2007). *Introduction to Discrete Event Systems*. Springer, second edition.

Chamberlin, E. H. (1933). *Theory of Monopolistic Competition*. Harvard University Press, Cambridge.

Chang, P.-L. and Lin, C.-T. (1991). On the effect of centralization on expected costs in a multi-location newsboy problem. *The Journal of the Operational Research Society*, 42(11):1025–1030.

Chen, F., Eliasberg, J., and Zipkin, P. (1998). Customer preferences, supply-chain costs, and product-line design. In Ho and Tang (1998).

Chong, J.-K., Ho, T.-H., and Tang, C. S. (1998). Product structure, brand width and brand share. In Ho and Tang (1998).

Chopra, S. and Meindl, P. (2004). *Supply Chain Management: Strategy, Planning and Operation*. Prentice Hall, second edition.

Christopher, M. (2005). *Logistics and Supply Chain Management*. Prentice Hall, London, third edition.

Chung, C. and Krajewski, L. J. (1984). Planning horizons for master production scheduling. *Journal of Operations Management*, 4(4):389–406.

Chung, C. and Krajewski, L. J. (1986). Replanning frequencies for master production schedules - notes and recommendations. *Decision Sciences*, 17:263–273.

Chung, C.-S., Flynn, J., Kuik, R., Nunen, J. v., and Stalinski, P. (2006). Inventory placement problems in dynamic supply chains. *ZfB - Zeitschrift für Betriebswirtschaft*, 76(2):177–198.

Clark, A. J. and Scarf, H. (1960). Optimal policies for a multi-echelon inventory problem. *Management Science*, 6(4):1782–1790.

Coenberg, A. G. and Fischer, T. M. (1991). Prozesskostenrechnung - Strategische Neuorientierung der kostenrechnung. *Die Betriebswirtschaft*, 1:21–38.

Cooper, R. and Kaplan, R. S. (1988). Measure costs right: Make the right decisions. *Harvard Business Review*, pages 96–103.

Corbett, C. J. and Rajaram, K. (2006). A generalization of the inventory pooling effect to nonnormal dependent demand. *Manufacturing & Service Operations Management*, 8(4):351–358.

Cormen, T. T., Leiserson, C. E., and Rivest, R. L., editors (2001). *Introduction to Algorithms*. MIT Press, second edition.

Dangelmaier, W. (2003). *Produktion und Information: System und Modell*. Springer.

Dangelmaier, W. and Warnecke, H.-J. (1997). *Fertigungslenkung. Planung und Steuerung des Ablaufs der diskreten Fertigung*. Springer.

Danne, C., Blecken, A., and Dangelmaier, W. (2008). Complexity-induced uncertainty in supply chains - a framework and case studies. In *Robuste und sichere Logistiksysteme - 4. Wissenschaftssymposium Logistik der Bundesvereinigung Logistik (BVL)*, Munich.

Davila, T. and Wouters, M. (2003). Measuring the benefits of product standardization and postponement of configuration in a supply chain. In Harrison et al. (2004).

Davis, T. (1993). Effective supply chain management. *Sloan Management Review*, 34(4):35–46.

DeBodt, M. and Wassenhove, L. N. V. (1983). Cost increases due to demand uncertainty in MRP lot sizing. *Decision Sciences*, 14:345–361.

Dickersbach, J., Keller, G., and Weihrauch, K. (2005). *Production Planning and Control with SAP*. SAP Press.

DIN19226 (1968). *Regelungstechnik und Steuerungstechnik. Begriffe und Benennungen*. Beuth, Berlin.

Dixit, A. K. and Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity. *The American Economic Review*, 67(3):297–308.

Domschke, W. and Drexl, A. (2005). *Einführung in Operations Research*. Springer, Berlin - Heidelberg, sixth edition.

Dorigo, M., Caro, G. D., and Gambardella, L. M. (1999). Ant algorithms for discrete optimization. *Artificial Life*, 5(2):137–172.

Dréo, J., Siarry, P., Pétrowski, A., and Taillard, E. (2006). *Metaheuristics for Hard Optimization*. Springer, Berlin Heidelberg.

Eglese, R. W. (1990). Simulated annealing: A tool for operational research. *European Journal of Operational Research*, 46(3):271–281.

Enns, S. T. (2001). MRP performance effects due to lot size and planned lead time settings. *International Journal of Production Research*, 39(3):461–480.

Enns, S. T. (2002). MRP performance effects due to forecast bias and demand uncertainty. *European Journal of Operational Research*, 138(1):87–102.

Eppen, G. and Schrage, L. (1981). Centralized ordering policies in a multi-warehouse system with leadtimes and random demand. In Schwarz, L. B., editor, *Multi-Level Production/Inventory Systems: Theory and Practice*, pages 51–67. North-Holland, New York.

Eppen, G. D. (1979). Effects of centralization on expected costs in a multi-location newsboy problem. *Management Science*, 25(5):498–501.

Fleischmann, B. and Meyr, H. (1997). The general lotsizing and scheduling problem. *OR Spectrum*, 19:11–21.

Fleischmann, B. and Meyr, H. (2003). Planning hierarchy, modeling and advanced planning systems. In de Kok, A. and Graves, S., editors, *Supply Chain Management: Design, Coordination and Operation*, volume 11 of *Handbooks in Operations Research and Management Science*, pages 457–524. Elsevier.

Fourer, R., Gay, D., and Kernighan, B. (1993). *AMPL : A Modelling Language for Mathematical Programming*. Boyd & Fraser, Danvers, Massachusetts.

Gerchak, Y. and He, Q.-M. (2003). On the relation between the benefits of risk pooling and the variability of demand. *IIE Transactions*, 35(11):1027 – 1031.

Glaser, H. (1993). Zur entscheidungsrelevanz prozessorientierter stückkosten. *Kostenrechnungspraxis*, 2:43–47.

Glover, F. (1989). Tabu search - part I. *ORSA Journal on Computing*, 1(3):190–206.

Glover, F. (1990). Tabu search - part II. *ORSA Journal on Computing*, 2(1):4–32.

Graves, S. C. and Willems, S. P. (1996). Strategic safety stock placement in supply chains. In *Proceedings of the 1996 MSOM Conference*, pages 299–304, Dartmouth College.

Graves, S. C. and Willems, S. P. (2000). Optimizing strategic safety stock placement in supply chains. *Manufacturing & Service Operations Management*, 2(1):68–83.

Graves, S. C. and Willems, S. P. (2003). Supply chain design: Safety stock placement and supply chain configuration. In de Kok, A. and Graves, S., editors, *Supply Chain Management: Design, Coordination and Operation*, volume 11 of *Handbooks in Operations Research and Management Science*, pages 95–132. Elsevier.

Guide, V. D. R. and Srivastava, R. (2000). A review of techniques for buffering against uncertainty with MRP systems. *Production Planning and Control*, 11(3):223–233.

Gupta, D. and Magnusson, T. (2005). The capacitated lot-sizing and scheduling problem with sequence-dependent setup costs and setup times. *Computers and Operations Research*, 32(4):727–747.

Haase, K. (1996). Capacitated lot-sizing with sequence dependent setup costs. *OR Spectrum*, 18:51–59.

Harrison, T. P., Lee, H. L., and Neale, J. J., editors (2004). *The Practice of Supply Chain Management: Where Theory and Application Converge*, International Series In Operations Research & Management Science, New York. Springer.

Hartman, B. C. and Dror, M. (2005). Allocation of gains from inventory centralization in newsvendor environments. *IIE Transactions*, 37(2):93–107.

Heina, J. (1999). *Variantenmanagement: Kosten-Nutzen-Bewertung zur Optimierung der Variantenvielfalt*. Dt. Univ.-Verl., Wiesbaden.

Herrmann, A. and Peine, K. (2007). Variantenmanagement. In Albers, S. and Herrmann, A., editors, *Handbuch Produktmanagement*, pages 649–679. Gabler.

Ho, T.-H. and Tang, C. S., editors (1998). *Product Variety Management*, volume 10 of *International Series in Operations Research & Management Science*. Springer.

Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, MI.

Homburg, C. and Daum, D. (1997). Wege aus der Komplexitätskostenfalle. *Zeitschrift für wirtschaftlichen Fabrikbetrieb*, 92:333–337.

Hoole, R. (2005). Five ways to simplify your supply chain. *Supply Chain Management: An International Journal*, 10:3–6(4).

Hoole, R. (2006). Drive complexity out of your supply chain. *Harvard Business Review Supply Chain Strategy*.

Hopp, W. (2006). *Supply Chain Science*. McGraw-Hill, New York.

Hopp, W. and Spearman, M. (2000). *Factory Physics: Foundations of Manufacturing Management*. McGraw-Hill/Irwin, second edition.

Horváth, P. and Mayer, R. (1989). Preozesskostenrechnung - Konzeption und Entwicklungen. *Kostenrechnungspraxis*, 4:214–219.

Horváth, P. and Seidenschwarz, W. (1992). Zielkostenmanagement. *Controlling*, 3:142–150.

Hotelling, H. H. (1929). Stability in competition. *Economic Journal*, 39:41–57.

Inderfurth, K. (1991). Safety stock optimization in multi-stage inventory systems. *International Journal of Production Economics*, 24(1-2):103–113.

Inderfurth, K. (1992). Mehrstufige Sicherheitsbestandsplanung mit dynamischer Programmierung. *OR Spektrum*, 14(1):19–32.

Inderfurth, K. and Minner, S. (1998). Safety stocks in multi-stage inventory systems under different service measures. *European Journal of Operational Research*, 106(1):57–73.

Johnson, H. T. and Kaplan, R. S. (1991). *Relevance lost: the rise and fall of management accounting*. Harvard Business Press.

Kadipasaoglu, S. N. (1995). The effect of freezing the master production schedule on cost in multilevel MRP systems. *Production and Inventory Management Journal*, Q3:30–36.

Kahn, B. (1998). Variety: From the consumer's perspective. In Ho and Tang (1998).

Kahn, B. E. and Morales, A. C. (2001). Choosing variety. In Hoch, S. J., Kunreuther, H. C., and Gunther, R. E., editors, *Wharton on Making Decisions*, pages 63–77. New York.

Kaiser, A. (1995). *Integriertes Variantenmanagement mit Hilfe der Prozesskostenrechnung*. PhD thesis, Universität St. Gallen.

Kestel, R. (1995). *Variantenvielfalt und Logistiksysteme: Ursachen, Auswirkungen, Lösungen*. Dt. Univ.-Verl., Wiesbaden.

Kirchhof, R. (2002). *Ganzheitliches Komplexitätsmanagement - Grundlagen und Methodik des Umgangs mit Komplexität im Unternehmen*. Dt. Univ.-Verl., Wiesbaden.

Klaus, P. (2005). Die Frage der optimalen Komplexität in Supply-Chains und Supply-Netzwerken. In *Perspektiven des Supply Management*, pages 361–375. Springer, Berlin/Heidelberg.

Klein, R. and Scholl, A. (2004). *Planung und Entscheidung: Konzepte, Modelle und Methoden einer modernen betriebswirtschaftlichen Entscheidungsanalyse*. Vahlen, München.

Klosterhalfen, S. and Minner, S. (2006). Comparison of stochastic- and guaranteed-service approaches to safety stock optimization in supply chains. In Waldmann, K.-H. and Stocker, U. M., editors, *Operations Research Proceedings*, pages 485–490. Springer.

Kluge, J., Stein, L., Krubasik, E., Beyer, I., Düsedau, D., Huhn, W., Schmidt, E., and Deger, R. (1994). *Wachstum durch Verzicht. Schneller Wandel zur Weltklasse: Vorbild Elektronikindustrie*. Schäffer-Poeschel, Stuttgart.

Koh, S. C. L., Saad, S. M., and Jones, M. H. (2002). Uncertainty under MRP-planned manufacture: review and categorization. *International Journal of Production Research*, 40:2399–2421.

Köhler, A. (1988). *Beitrag zur Verbesserung der Fertigungskostentransparenz bei Großserienfertigung mit Produktvielfalt*. Springer, Berlin Heidelberg.

Kropp, D. H., Carlson, R. C., and Jucker, J. V. (1979). Use of dynamic lot-sizing to avoid nervousness in material requirements planning systems. *Production and Inventory Management Journal*, Q3:49–59.

Kulkarni, S. S., Magazine, M. J., and Raturi, A. S. (2004). Risk pooling advantages of manufacturing network configuration. *Production & Operations Management*, 13(2):186 – 199.

Kulkarni, S. S., Magazine, M. J., and Raturi, A. S. (2005). On the trade-offs between risk pooling and logistics costs in a multi-plant network with commonality. *IIE Transactions*, 37(3):247–265.

Kunreuther, H. C. and Morton, T. E. (1973). Planning horizons for production smoothing with deterministic demands. *Management Science*, 20(1):110–125.

Kunreuther, H. C. and Morton, T. E. (1974). General planning horizons for production smoothing with deterministic demands. *Management Science*, 20(7):1037–1047.

Lagodimos, A. and Anderson, E. (1993). Optimal positioning of safety stocks in MRP. *International Journal of Production Research*, 31(8):1797–1813.

Lambert, D. and Cooper, M. (2000). Issues in supply chain management. *Industrial Marketing Management*, 29:65–83.

Lancaster, K. (1979). *Variety, Efficiency and Equity*. Columbia University Press, New York.

Lancaster, K. (1990). The economics of product variety: A survey. *Marketing Science*, 9(3):189–206.

Lancaster, K. (1998). Markets and product variety management. In Ho and Tang (1998).

Lee, H. L. and Billington, C. (1992). Managing supply chain inventory: Pitfalls and opportunities. *Sloan Management Review*, 33(3):65–73.

Lee, H. L. and Tang, C. S. (1997). Modelling the costs and benefits of delayed product differentiation. *Management Science*, 43(1):40–53.

Lee, J., Ristroph, J. H., Zhu, Z., , and Ruangdet, M. (1993). Performance evaluation of lotsizing method with multiple quantity discounts in a rolling horizon environment. *Production and Inventory Management Journal*, Q1:50–57.

Lee, L. and Billington, C. (1995). The evolution of supply chain management: Models and practice at Hewlett-Packard. *Interfaces*, 25(5):42–63.

Lee, T. S. and Adam, E. E. J. (1986). Forecasting error evaluation in MRP production inventory systems. *Management Science*, 39(9):1186–1205.

Lin, N. and Krajewski, L. J. (1992). A model for master production scheduling in uncertain environments. *Decision Sciences*, 23:839–861.

Lin, N., Krajewski, L. J., Leong, G. K., and Benton, W. C. (1994). The effects of environmental factors on the design of master production scheduling systems. *Journal of Operations Management*, 11:367–384.

Luhmann, N. (1980). Komplexität. In Grochla, E., editor, *Handwörterbuch der Organisation*, pages 1064–1070. Poeschel Verlag, Stuttgart.

Marti, M. (2007). *Complexity Management: Optimizing Product Architecture of Industrial Products*. PhD thesis, University of St. Gallen.

Meier, H. and Hanenkamp, N. (2004). Komplexitätsmanagement in der Supply Chain. In Busch, A. and Dangelmaier, W., editors, *Integriertes Supply Chain Management*. Gabler.

Minner, S. (1997). Dynammic programming algorithms for multi-stage safety stock optimization. *OR Spektrum*, 19(4):261–271.

Minner, S. (2000). *Strategic Safety Stocks in Supply Chains*, volume 490 of *Lecture Notes in Economics and Mathematical Systems*. Springer.

Molinder, A. (1997). Joint optimization of lot-sizes, safety stocks and safety lead times in an MRP system. *International Journal of Production Research*, 35(4):983–994.

Moore, W. L., Lehmann, D. R., and Pessemier, E. A. (1986). Hierarchical representations of market structures and choice processes through preference trees. *Journal of Business Research*, 14(5):371–386.

Nahmias, S. (1997). *Production and Operations Analysis*. McGraw-Hill, third edition.

Neale, J., Tomlin, B., and Willems, S. (2003). The role of inventory in superior supply chain performance. In Harrison et al. (2004).

Orlicky, J. (1975). *Material requirements planning*. McGraw-Hill, New York.

Papalambros, P. Y. and Wilde, D. J. (2000). *Model building in mathematical programming*. Cambridge University Press, second edition.

Pärli, H. (1980). Problemanalyse. In Grochla, E., editor, *Handwörterbuch der Organisation*, pages 1912–1917. Poeschel Verlag, Stuttgart.

Perona, M. and Miragliotta, G. (2004). Complexity management and supply chain performance assessment. A field study and a conceptual framework. *International Journal of Production Economics*, 90(1):103–115.

Pickett, J. P., editor (2000). *The American Heritage Dictionary of the English Language*. Houghton Mifflin Company, fourth edition.

Prillmann, M. (1996). *Management der Variantenvielfalt*. Lang, Frankfurt am Main.

Quelch, J. A. and Kenny, D. (1994). Extend profits, not product lines. *Harvard Business Review*, 72.

Ramdas, K. (2003). Managing product variety: An integrative review and research directions. *Production and Operations Management*, 12(1):79–101.

Randall, T. and Ulrich, K. (2001). Product variety, supply chain structure, and firm performance: Analysis of the U.S. bicycle industry. *Management Science*, 47(12):1588–1604.

Rathnow, P. J. (1993). *Integriertes Variantenmanagement: Bestimmung, Realisierung und Sicherung der optimalen Produktvielfalt*. Vandenhoeck & Ruprecht.

Reeves, C. R. (1993). *Modern heuristic techniques for combinatorial problems*. McGraw-Hill.

Riemenschneider, M. (2006). *Der Wert von Produktvielfalt: Wirkung grosser Sortimente auf das Verhalten von Konsumenten*. Deutscher Universitäts-Verlag.

Ristroph, J. H. (1990). Simulation of ordering rules for single level MRP with a rolling horizon and no forecast error. *Computers and Industrial Engineering*, 19(1-4):155–159.

Rohde, J., Meyr, H., and Wagner, M. (2000). Die Supply Chain Planning Matrix. *PPS Management*, 5(1):10–15.

Russell, R. A. and Urban, T. L. (1993). Horizon extension for rolling production schedules: length and accuracy requirements. *International Journal of Production Economics*, 29:111–122.

Sahin, F. and Robinson, E. (2002). Flow coordination and information sharing in supply chains: review, implications and directions for future research. *Decision Sciences*, 33(4):505–536.

Schiller, Z., Burns, G., and Miller, K. L. (1996). Make it simple. *BusinessWeek*.

Schuh, G. (2005). *Produktkomplexität managen: Strategien - Methoden - Tools*. Hanser, München.

Seidenschwarz, W. (1991). Target costing. *Controlling*, 4:198–203.

Simchi-Levi, D., Chen, X., and Bramel, J. (2005). *The Logic of Logistics: Theory, Algorithms, and Applications for Logistics and Supply Chain Management*. Springer Series in Operations Research and Financial Engineering. Springer.

Simchi-Levi, D., Kaminsky, P., and Simchi-Levi, E. (2007). *Designing & Managing the Supply Chain*. McGraw-Hill, third edition.

Simpson Jr., K. F. (1958). In-process inventories. *Operations Research*, 6(6):863–873.

Sobel, M. J. (2008). Risk pooling. In Chhajed, D. and Lowe, T. J., editors, *Building Intuition*, volume 115 of *International Series in Operations Research & Management Science*, pages 155–174. Springer US.

Soman, C. A., van Donk, D. P., and Gaalman, G. (2004). Combined make-to-order and make-to-stock in a food production system. *International Journal of Production Economics*, 90(2):223–235.

Sridharan, V. and Berry, W. L. (1990a). Freezing the MPS under demand uncertainty. *Decision Sciences*, 21:97–120.

Sridharan, V. and Berry, W. L. (1990b). Master production scheduling make-to-stock products: a framework for analysis. *International Journal of Production Research*, 28(3):541–558.

Sridharan, V., Berry, W. L., and Udayabhanu, V. (1987). Freezing the MPS under rolling planning horizons. *Management Science*, 33(9):1137–1149.

Sridharan, V. and LaForge, R. L. (1994a). Freezing the master production schedule: implications for fill rate. *Decision Sciences*, 25:461–469.

Sridharan, V. and LaForge, R. L. (1994b). A model to estimate service levels when a portion of the master production schedule is frozen. *Computers Operations Research*, 21(5):477–486.

Stadtler, H. and Kilger, C., editors (2005). *Supply Chain Management and Advanced Planning - Concepts, Models, Software and Case Studies*. Springer-Verlag, third edition.

Stevens, G. C. (1989). Integrating the supply chain. *International Journal of Physical Distribution & Logistics Management*, 19(8):3–8.

Suerie, C. (2005). *Time Continuity in Discrete Time Models - New Approaches for Production Planning in Process Industries*, volume 552 of *Lecture Notes in Economics and Mathematical Systems*. Springer.

Supply Chain Council (2006). Supply-chain operations reference-model, overview of SCOR version 8.0. available online at http://www.supply-chain.org, Washington DC. accessed 25.05.2009.

Supply Chain Council (2007). Supply Chain Council homepage. http://www.supply-chain.org. accessed 25.05.2009.

Tempelmeier, H. (2003). *Material-Logistik – Modelle und Algorithmen für die Produktionsplanung und -steuerung und das Supply Chain Management*. Springer, Berlin, fifth edition.

Tempelmeier, H. (2005). *Bestandsmanagement in Supply Chains*. Books on Demand.

The Economist Intelligence Unit (2008). World consumer goods: White paper - managing the challenge of product proliferation.

Thonemann, U. W. and Bradley, J. R. (2002). The effect of product variety on supply-chain performance. *European Journal of Operational Research*, 143(3):548 – 569.

Ulrich, H. and Probst, G. J. B. (1988). *Anleitung zum ganzheitlichen Denken und Handeln*. Haupt, Bern / Stuttgart.

Ulrich, K. T. (2006). Variety. In *Design: Creation of Artifacts in Society*. online book chapter: http://opim.wharton.upenn.edu/ ulrich/documents/ulrichvariety.pdf, accessed 31.03.2009.

Ulrich, K. T., Randall, T., Fisher, M. L., and Reibstein, D. (1998). Managing product variety. In Ho and Tang (1998).

Van der Vorst, J. G. (2000). *Effective food supply chains; generating, modelling and evaluating supply chain scenarios*. PhD thesis, Wageningen University.

van Houtum, G. J., Inderfurth, K., and Zijm, W. H. M. (1996). Materials coordination in stochastic multi-echelon systems. *European Journal of Operational Research*, 95(1):1–23.

van Ryzin, G. J. (2001). Analyzing inventory cost and service in supply chains.

Venkatesh, S. and Swaminathan, J. (2003). Managing product variety through postponement: concepts and applications. In Harrison et al. (2004).

Voß, S. (2001). Meta-heuristics: The state of the art. In *Local Search for Planning and Scheduling*, volume 2148 of *Lecture Notes in Computer Science*, pages 1–23, London, UK. Springer.

Vollmann, T. E., Berry, W. L., Whybark, D. C., and Jacobs, F. R., editors (2004). *Manufacturing Planning and Control for Supply Chain Management*. McGraw-Hill, fifth edition.

Warner, M. (2001). *International Encyclopedia of Business and Management*. Thomson Learning, second edition.

Wemmerlov, U. (1985). Comments on cost increases due to demand uncertainty in MRP lot sizing: a verification of ordering probabilities. *Decision Sciences*, 16:410–419.

Wemmerlov, U. (1986). A time-phase order-point system in environments with and without demand uncertainty. *International Journal of Production Research,*, 24(2):343–358.

Wemmerlov, U. and Whybark, D. C. (1984). Lot-sizing under uncertainty in a rolling schedule environment. *International Journal of Production Research*, 22(3):467–484.

Whybark, D. C. and Williams, J. G. (1976). Material requirements planning under uncertainty. *Decision Sciences*, 7(4):595–606.

Wilding, R. (1998). The supply chain complexity triangle: Uncertainty generation in the supply chain. *International Journal of Physical Distribution & Logistics Management*, 28(8):599–616.

Wolter, A. (1937). Das Problem der Wirtschaftlichkeit bei der industriellen Sortenproduktion. *Zeitschrift für handelswissenschaftliche Forschung*, 31:331–356, 374–402, 426–443.

Yang, B. and Burns, N. (2003). Implications of postponement for the supply chain. *International Journal of Production Research*, 41(9):2075–2090(16).

Yang, B., Burns, N. D., and Backhouse, C. J. (2004a). Management of uncertainty through postponement. *International Journal of Production Research*, 42(6):1049–1064(16).

Yang, B., Burns, N. D., and Backhouse, C. J. (2004b). Postponement: a review and an integrated framework. *International Journal of Operations & Production Management*, 24(5):468–487.

Yano, C. A. and Carlson, R. C. (1985). An analysis of scheduling policies in multiechelon production systems. *IIE Transactions*, 17(4):370–377.

Yano, C. A. and Carlson, R. C. (1987). Interaction between frequency of rescheduling and the role of safety stock in MRP systems. *International Journal of Production Research*, 25(2):221–232.

Yeung, J. H. Y. and Wong, W. C. K. (1998). Parameters affecting the effectiveness of MRP systems: a review. *International Journal of Production Research*, 36(2):313–332.

Zhao, X., Goodale, J. C., and Lee, T. S. (1995). Lot-sizing rules and freezing the MPS in material requirements planning systems under demand uncertainty. *International Journal of Production Research*, 33(8):2241–2276.

Zhao, X. and Lee, T. S. (1993). Freezing the MPS for MRP systems under demand uncertainty. *Journal of Operations Management*, 11:185–205.

Zhao, X. and Lee, T. S. (1996). Freezing themaster production schedule in multi-level material requirements planning systems under deterministic demand. *Production Planning and Control*, 7(2):144–161.

Zipkin, P. H. (2000). *Foundations of Inventory Management*. McGraw-Hill/Irwin.

APPENDIX A

Implementation

Figure A.1 provides an overview of the Complana functional modules and their interactions. The software support for the single steps of the analysis process are described in more detail in the following sections. Table A.1 shows the separate packages that make up the Complana software and summarises the functionality provided by each of them.

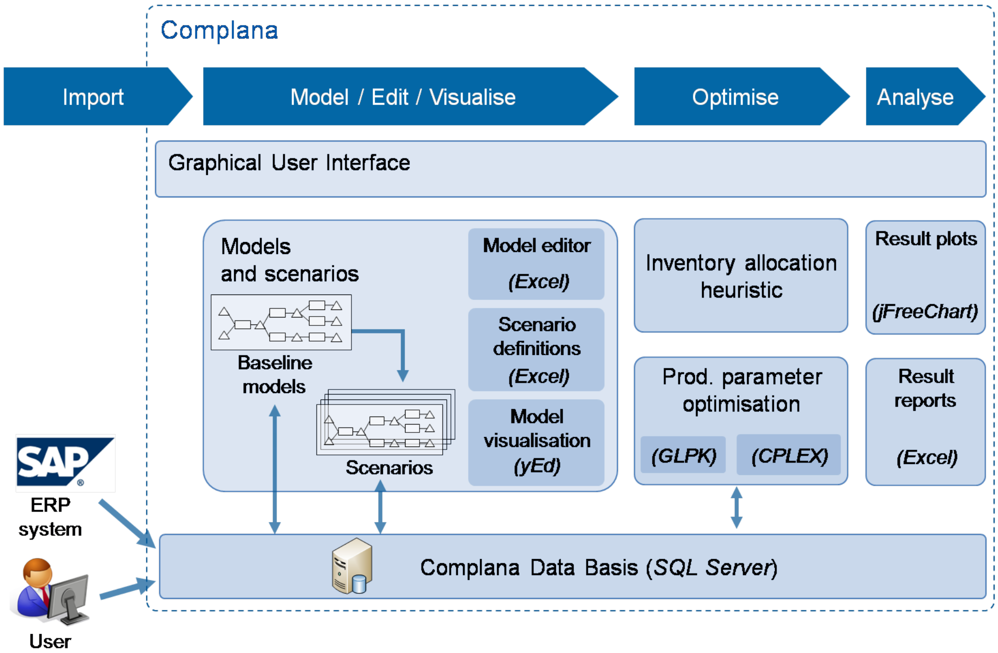


Figure A.1: Overview of the Complana software tool

The Complana tool was developed using the Java programming language and the Eclipse Rich client platform (RCP) as the underlying development framework[1]. Further third-party technologies used in the implementation are

1 The Eclipse RCP is an open tools platform for building and deploying rich client applications. See http://www.eclipse.org/rcp/ for details.

- Microsoft SQL Server for data persistence[2]
- Hibernate framework for object-relational mapping[3]
- yEd for network visualisation[4]
- jFreeChart for results visualisation[5]
- OpenTS as a framework for the development of the tabu search heuristic[6]
- Standard solvers and modelling languages for the mixed-integer optimisation problems, as described in Section 6.4.2.1.

2 See http://www.microsoft.com/sql/ for details.
3 See http://www.hibernate.org for details.
4 See http://www.yworks.com/yed/ for details.
5 See http://www.jfree.org/jfreechart/ for details.
6 See http://www.coin-or.org/ots/ for details.

Table A.1: Complana software packages

Package	Functional description
complana.app	This package contains all elements of the rich client application. In particular, this includes all GUI elements like perspectives, view and user dialogs.
complana.imports	This package provides the functionality to import the major part of the required input data from existing ERP systems. In particular, the import from the SAP ERP system was implemented and tested.
complana.model	This package contains all classes that represent data elements in the application domain, like materials, plants, production process steps etc. Furthermore, these classes use special annotations to specify all the information needed to configure their object relational mapping to and from the database via the Hibernate framework.
complana.network	This package comprises the network data model, i.e. all classes used to represent production and distribution networks with all their elements. Furthermore, it provides several features to work with these network models, like their transformation into the GraphML format for visualisation and their serialisation for permanent storage on the filesystem or in the underlying database.
complana.prodeval.optimisation	This package contains all classes required to generate MIP optimisation problems from existing production and distribution networks to optimise planning buffers and planned production quantities. After generation of these problems, the standard solvers CPLEX and GLPK can be invoked from the classes provided.
complana.ssplacement	This package contains all classes required to implement the inventory allocation tabu search heuristic. As it uses the OpenTS framework, the classes are implemented according the interfaces defined by that framework and thus provide custom implementation for the objective function, solutions, moves and neighbourhood managers.
complana.util	This package contains general utility classes used at different points in the entire software. This comprises classes that provide logging, time measuring, and reading and writing Microsoft Excel spreadsheet files.

A.1 Model Building and Management

The automatic generation of new assortment models from imported master data can be invoked via the GUI. Figure A.2 shows the corresponding dialog with its input parameters. The main inputs comprise the specification of the time interval for which demands should be considered, the definition of a set of finished products to start with as well as an optional list of locations to be considered in the model. If not specified, materials at all locations are included.

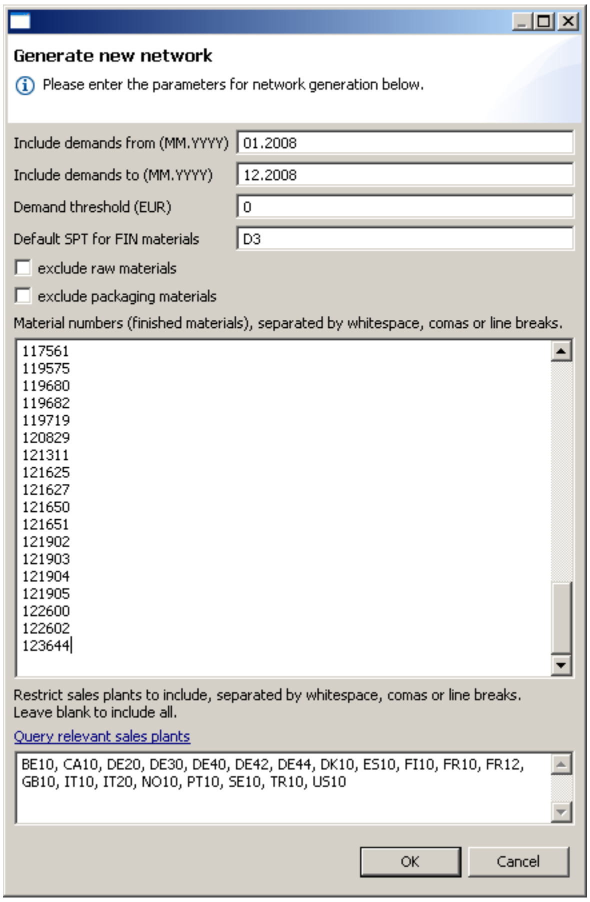


Figure A.2: Network generation dialog

Once a model has been generated, it is stored in the underlying database along with optimal descriptive meta data. Via the models management perspective shown in Figure A.3, all models stored in the database can be viewed, edited, visualised and deleted. As all models are stored in a central database, this also allows to easily share models and thus facilitates collaboration.

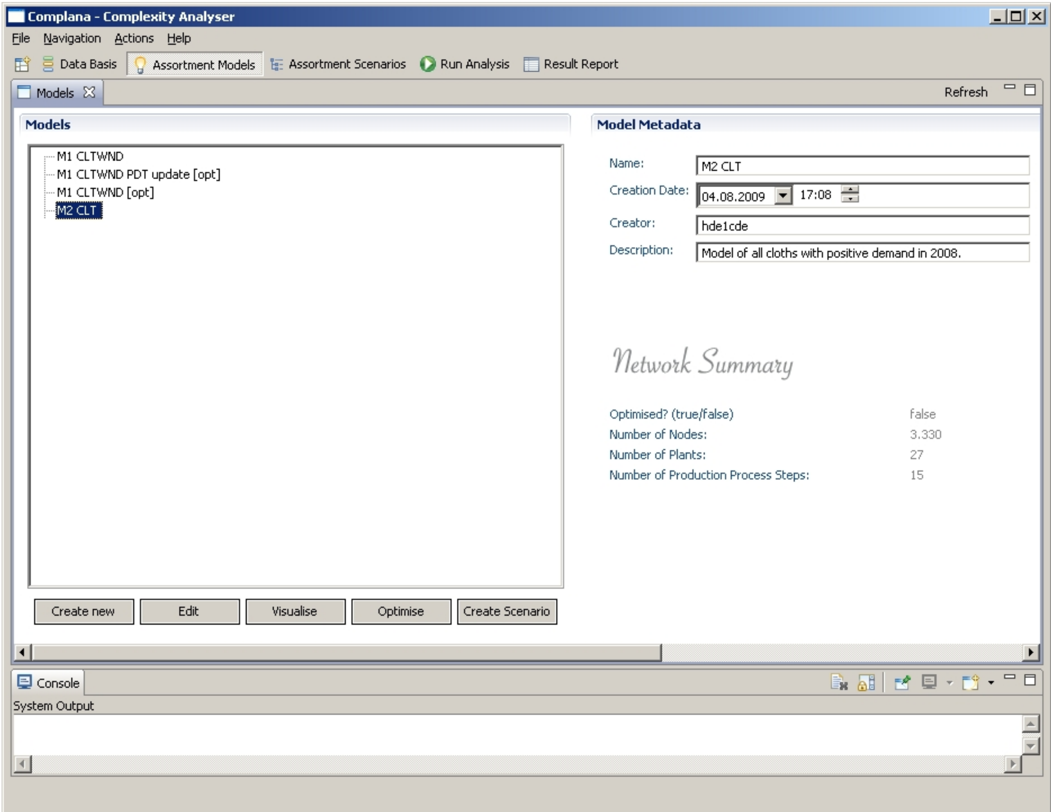
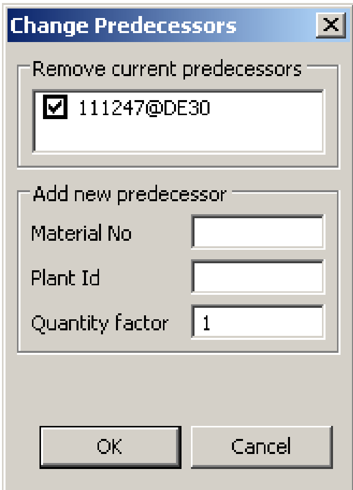


Figure A.3: Model management

Models can be edited in terms of the external demands and forecast deviations stored at each item as well as in terms of their structure. In order to allow easy editing of demands and forecast deviations for a large number of items, a list of items together with their demands and forecast deviations is written to a spreadsheet file. In this file, the demand parameters can be adjusted, before the file is parsed again by the Complana tool and the changes defined in the spreadsheet are applied to the model. Figure A.4(a) shows a detail of such a model editing spreadsheet file, while figure A.4(b) shows how the network structure can be changed by adapting the predecessors of single items.

Plant	Material No.	Material Type	Node Type	Sum External Demands	New Total external Demand	Adjust demand percentage	External forecast deviation	New external forecast deviation	Forecast Deviation Multiplier
ES10	112006	FIN	DIST	25	25	100%	100%	100%	1
ES11	112006	FIN	PROD	0	0	100%	0%	0%	1
DE40	126250	FIN	DIST	728	728	100%	100%	100%	1
DE30	126250	FIN	DIST	0	0	100%	0%	0%	1
DE31	126250	FIN	PROD	0	0	100%	0%	0%	1
BE10	112441	FIN	DIST	2394	2394	100%	33%	33%	1
BE11	112441	FIN	PROD	0	0	100%	0%	0%	1
BE10	112444	FIN	DIST	115	115	100%	74%	74%	1
BE11	112444	FIN	PROD	0	0	100%	0%	0%	1
BE10	120829	FIN	DIST	1379	1379	100%	28%	28%	1
BE11	120829	FIN	PROD	0	0	100%	0%	0%	1
BE10	121311	FIN	DIST	2312	2312	100%	58%	58%	1
BE11	121311	FIN	PROD	0	0	100%	0%	0%	1
TR10	121904	FIN	PROD	153	153	100%	307%	307%	1
TR10	121902	FIN	PROD	1269	1269	100%	151%	151%	1
TR10	121903	FIN	PROD	467	467	100%	186%	186%	1

(a) Model editing spreadsheet file



(b) Changing predecessor relations of an item

Figure A.4: Editing possibilities for existing models

Another useful feature for working with production and distribution networks is the visualisation component. Existing models can be visualised by automatically trans-forming their network structure into the GraphML file format and opening that file with the third-party tool yEd. Figure A.5 shows the corresponding dialog in the Complana tool. The visualisation can be further parameterised, e.g. by omitting raw and packaging materials to reduce the network complexity. Furthermore, an arbitrary set of material numbers can be specified such that only items related to these materials over some connection are included in the visualisation. This opens the possibility to use the visualisation as a graphical where-used list, which can be used e.g. to determine the set of end products and locations that would be affected by a change in a base material or some packaging.

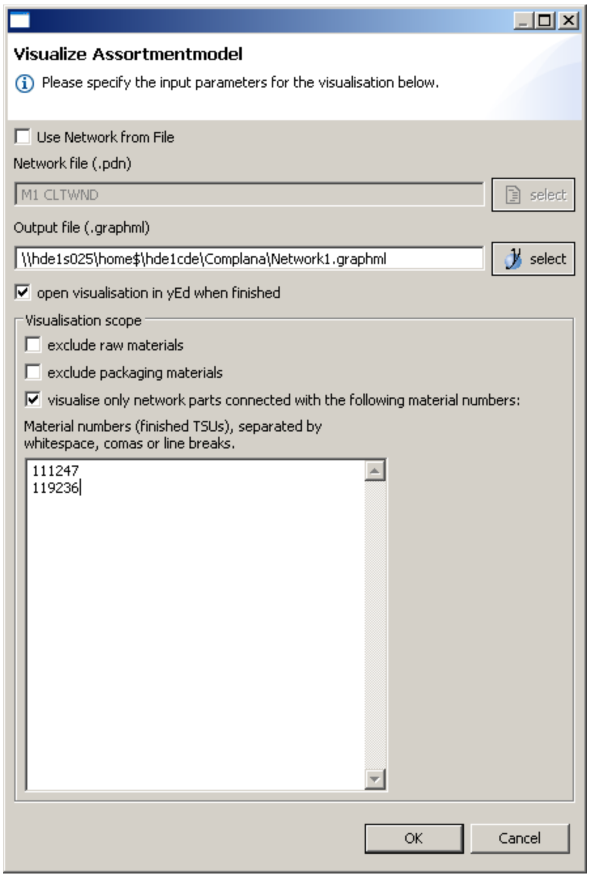


Figure A.5: Network visualisation dialog

A.2 Scenario Building and Management

Scenarios are defined in a spreadsheet file based on a certain template. In three distinct sections new material additions, finished material replacements as well as replacements of raw/packaging/semi-finished materials are defined. While especially the replacement of finished products is highly configurable, it is always possible to only define the material numbers of the discontinued and of the replacement materials, leaving all other options to their defaults. Conversion factors for replacements can be calculated automatically from within the software tool. Figure A.6 shows a detail of a scenario file for replacement of finished materials.

Plants [optional]	Material No.	Description [auto fill]	# Products [auto fill]	Replacement Material No.	Description [auto fill]	# Products [auto fill]	Procurement Source SPT [optional]	Demand Percentage [optional]	Conversion Factor [optional / auto fill]
	121311	CLTWND/_/T/_12Bx__2P/390x360/200x180/YW	24	119999	CLTWND/P/T/_2Sx_20Bx__2P/395x365/200x185	40	D3		1.67
	126250	CLTWND/P/T/_30Bx__2P/395x365/200x185/YW	60	119999	CLTWND/P/T/_2Sx_20Bx__2P/395x365/200x185	40			0.67
	121902	CLTWND/_/T/_4Tx_80Bx__1P/390x340/195x170	80	119999	CLTWND/P/T/_2Sx_20Bx__2P/395x365/200x185	40			0.50
	115021	CLTWND/_/T/_20Bx__5P/550x430/270x190/BN	100	119999	CLTWND/P/T/_2Sx_20Bx__2P/395x365/200x185	40			0.40
	125847	CLTWND/P/T/_2Sx_20Bx__2P/395x365/200x185	40	119999	CLTWND/P/T/_2Sx_20Bx__2P/395x365/200x185	40			1.00
	120829	CLTWND/P/T/_12Bx__2P/390x360/200x180/YW	24	119999	CLTWND/P/T/_2Sx_20Bx__2P/395x365/200x185	40			1.67
	100072	CLTWND/_/T/_6Sx60Bx1P/550x395/235x200/YW	60	119999	CLTWND/P/T/_2Sx_20Bx__2P/395x365/200x185	40			0.67
	112441	CLTWND/_/T/_15Bx__1P/360x390/180x200/YW	15	119999	CLTWND/P/T/_2Sx_20Bx__2P/395x365/200x185	40			2.67
	121905	CLTWND/_/T/_20Bx__2P/390x340/195x170/YW	40	119999	CLTWND/P/T/_2Sx_20Bx__2P/395x365/200x185	40			1.00
	112439	CLTWND/_/T/_10Bx__1P/360x490/YW	10	119999	CLTWND/P/T/_2Sx_20Bx__2P/395x365/200x185	40			4.00
	112431	CLTWND/_/T/10Bx_30Bx__1P/395x365/200x185	300	119999	CLTWND/P/T/_2Sx_20Bx__2P/395x365/200x185	40			0.13
	113571	CLTWND/_/T/100P/395x365/YW/C	100	119999	CLTWND/P/T/_2Sx_20Bx__2P/395x365/200x185	40			0.40
	115332	CLTWND/_/T/_30Bx__1P/550x395/270x190/YW	30	119999	CLTWND/P/T/_2Sx_20Bx__2P/395x365/200x185	40			1.33

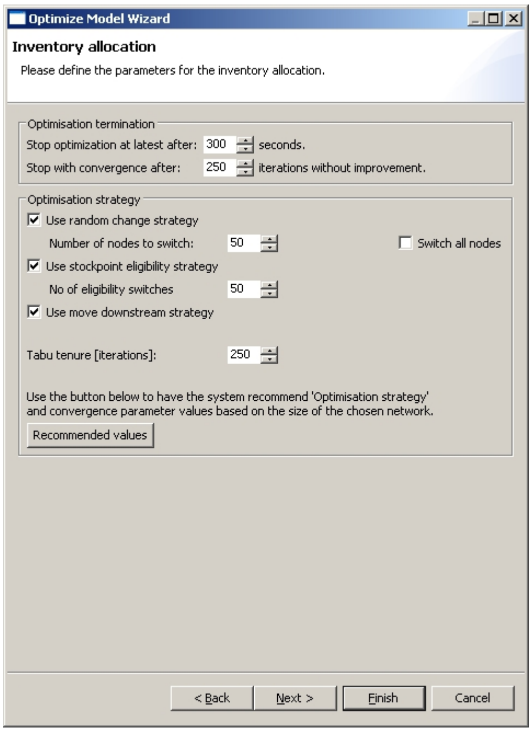
Figure A.6: Scenario definition in a spreadsheet file

Once a scenario definition is applied to a baseline model, the same viewing, editing and visualisation features can be used as already described for the baseline models in the preceding section.

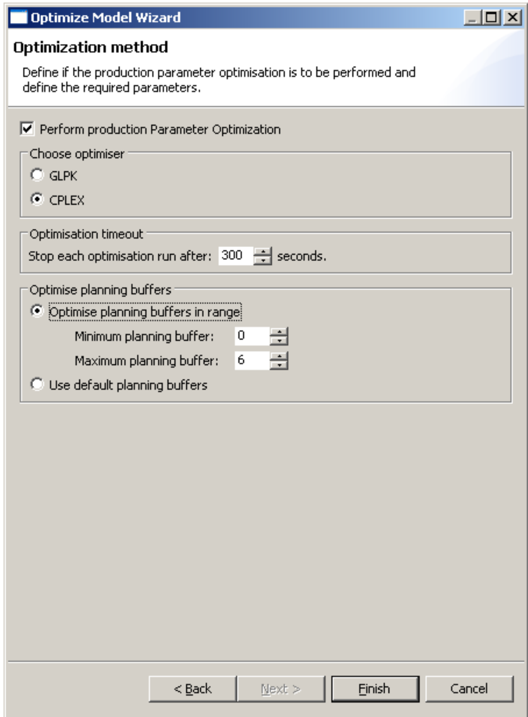
A.3 Optimisation Methods

For all generated models and scenarios, the optimisation of inventory allocation, planning buffer and planned production quantities can be invoked. Figure A.7 shows the user dialogs in the GUI with the possible parameter settings for each optimisation run.

When the optimisation is finished, the resulting optimised networks can be stored in the database as new models or the existing models can be updated with the optimised values. Furthermore, the results of the optimisation are written to a spreadsheet file to be used for more in-depth analysis.



(a) Inventory allocation parameters



(b) Production optimisation parameters

Figure A.7: Optimisation dialog wizard

A.4 Cost Comparison

For all generated models and scenarios, a comparative cost analysis can be invoked. The Complana tool compares the selected model and scenario with respect to the number of items for the different material types, inventory and production cost. Figures A.8 shows the selection of models and scenarios to be compared, as well as an example of a chart that compares their inventory costs.

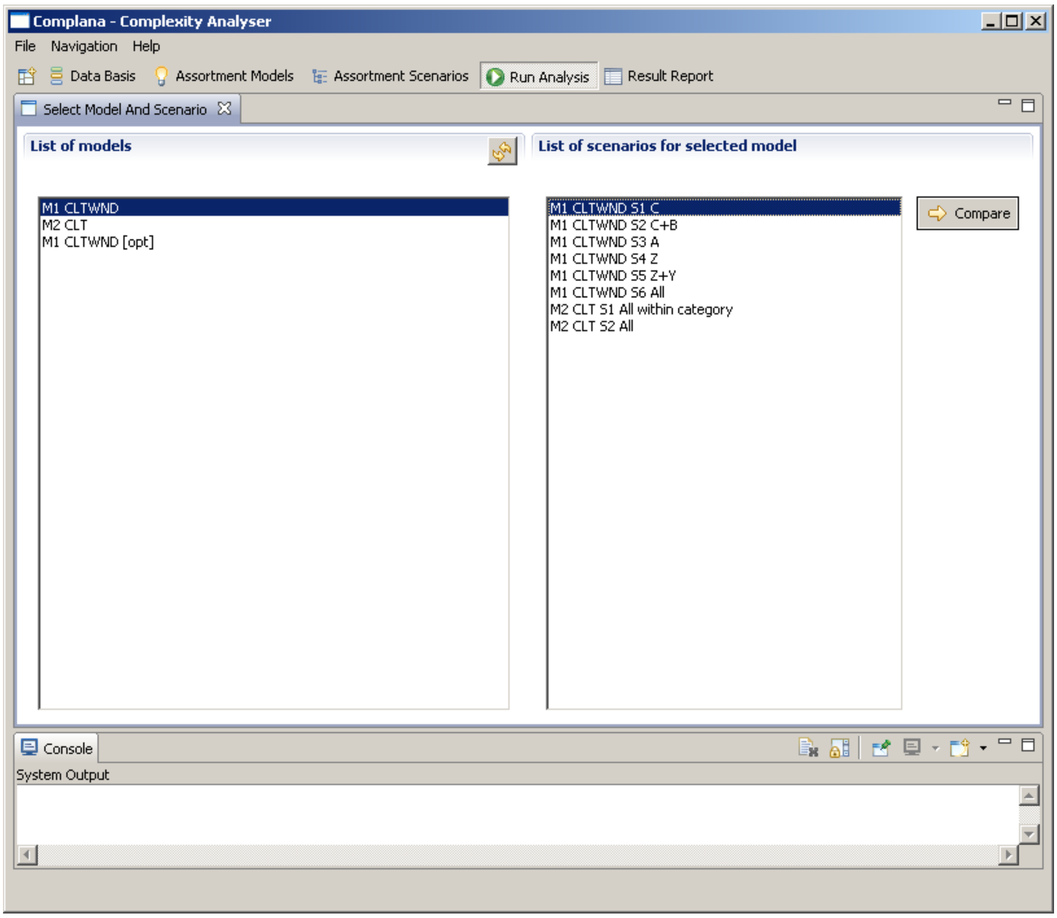


Figure A.8: Selection of models for cost comparison

APPENDIX B

Example Networks

This section provides some exemplary visualisations of the assortment models and scenarios used for the validation in Chapter 6. The visualisations of model M_2 and its related scenarios have been skipped due to their mere size, as their visual representation is far too large to be scaled on a single page. These visualisation examples should serve the purpose to allow a visual comparison of the complexity of the models and the corresponding scenarios.

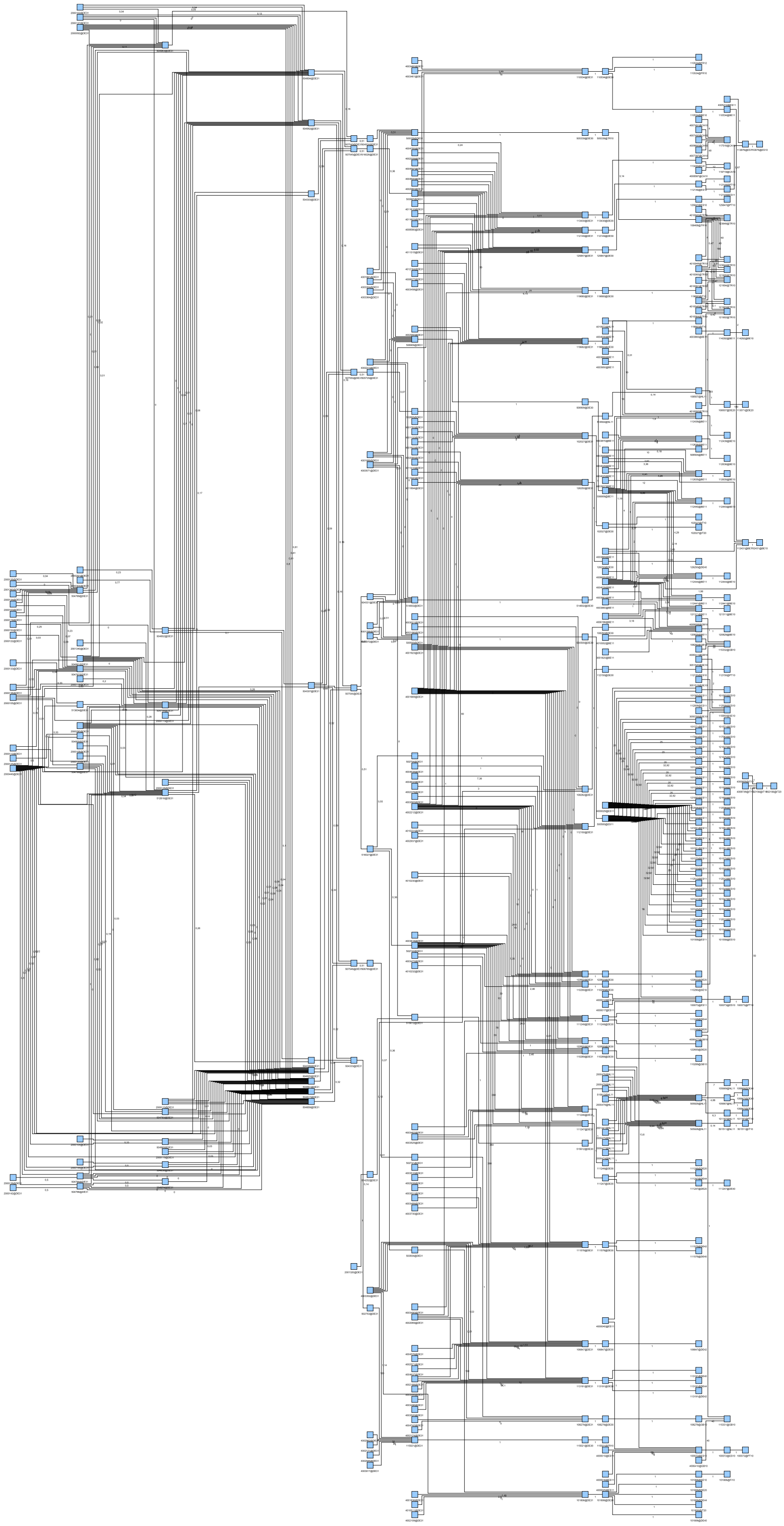


Figure B.1: Network visualisation for M_1

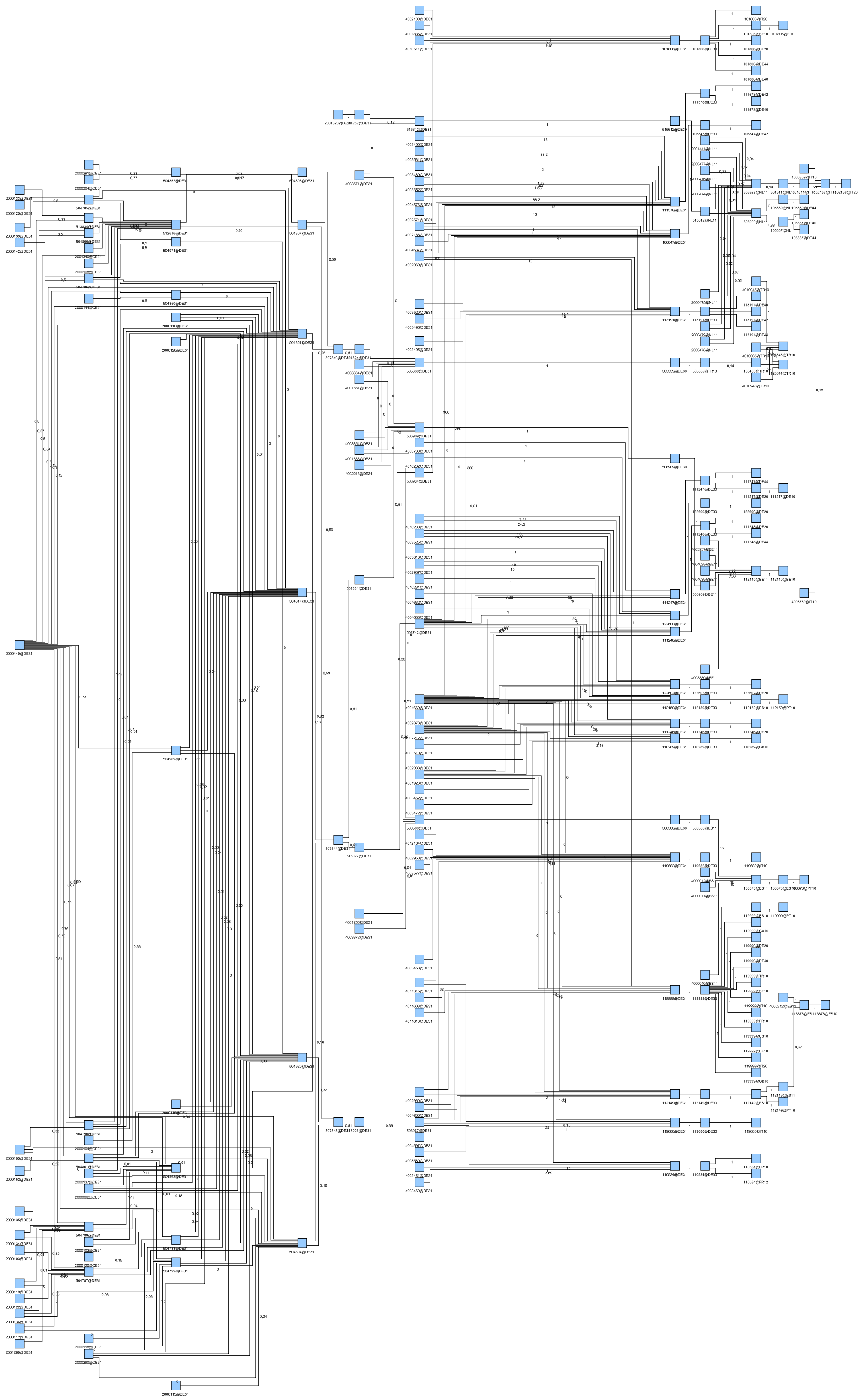


Figure B.2: Network visualisation for M_1/S_1

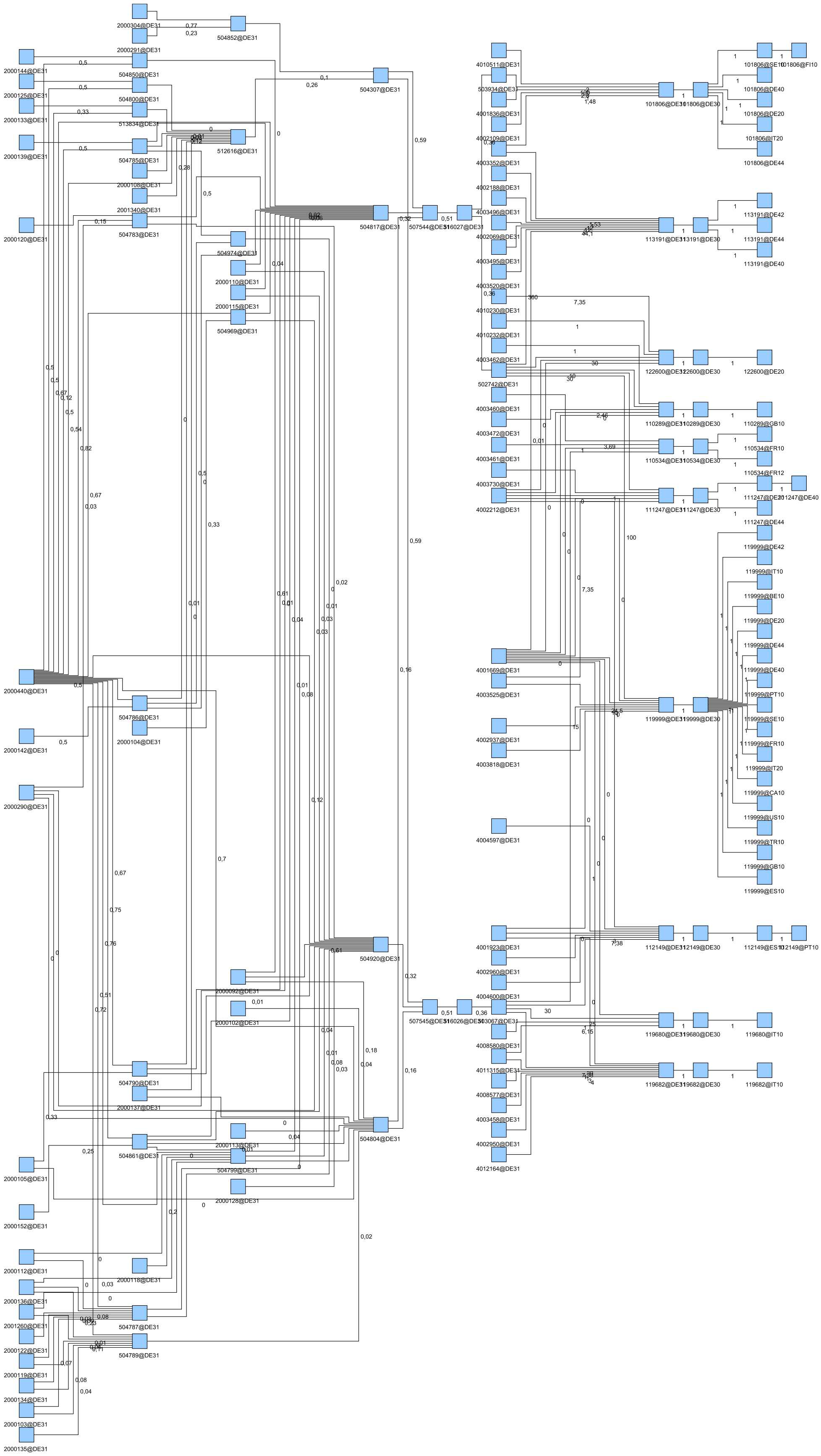


Figure B.3: Network visualisation for M_1/S_2

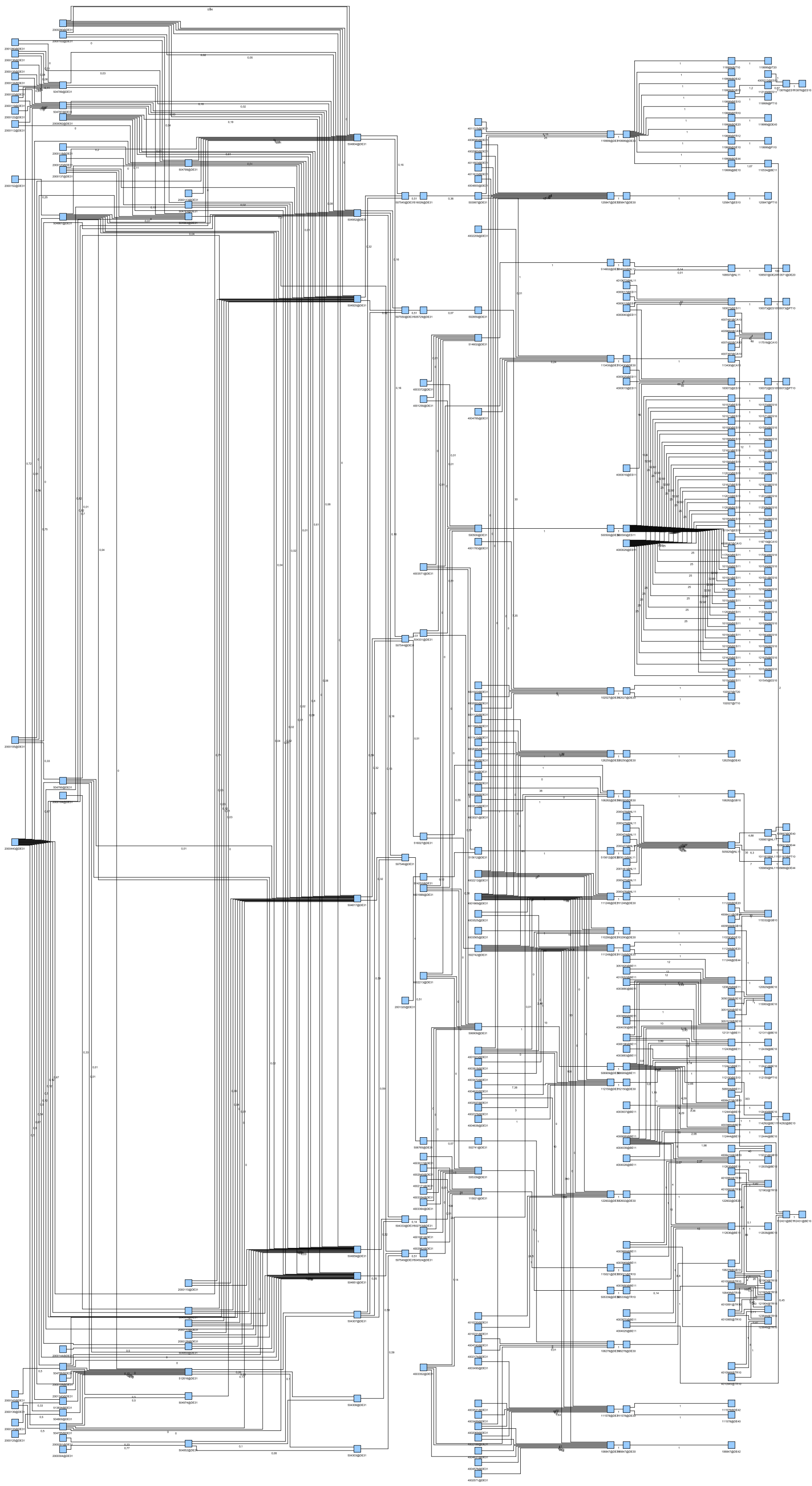


Figure B.4: Network visualisation for M_1/S_3

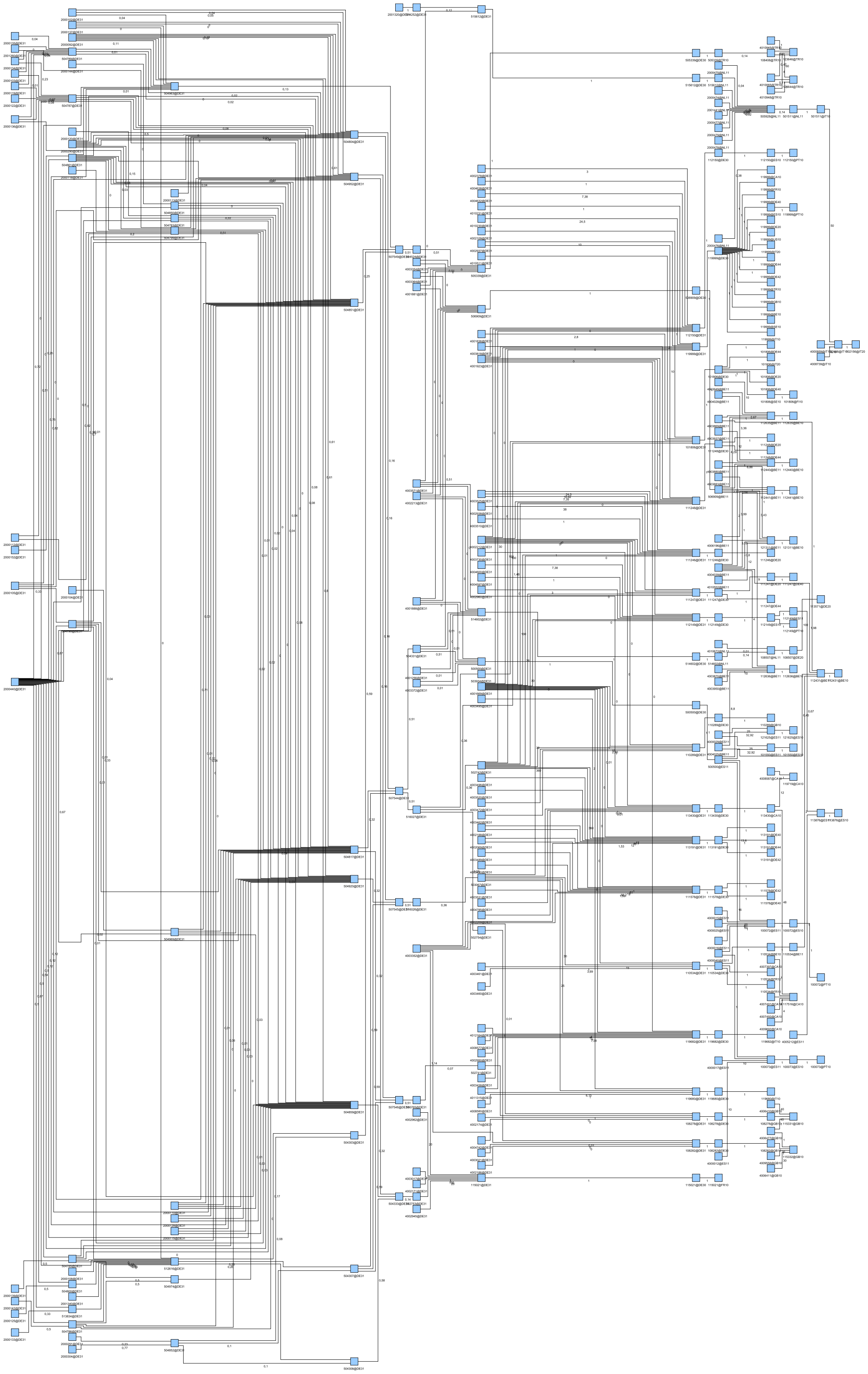


Figure B.5: Network visualisation for M_1/S_4

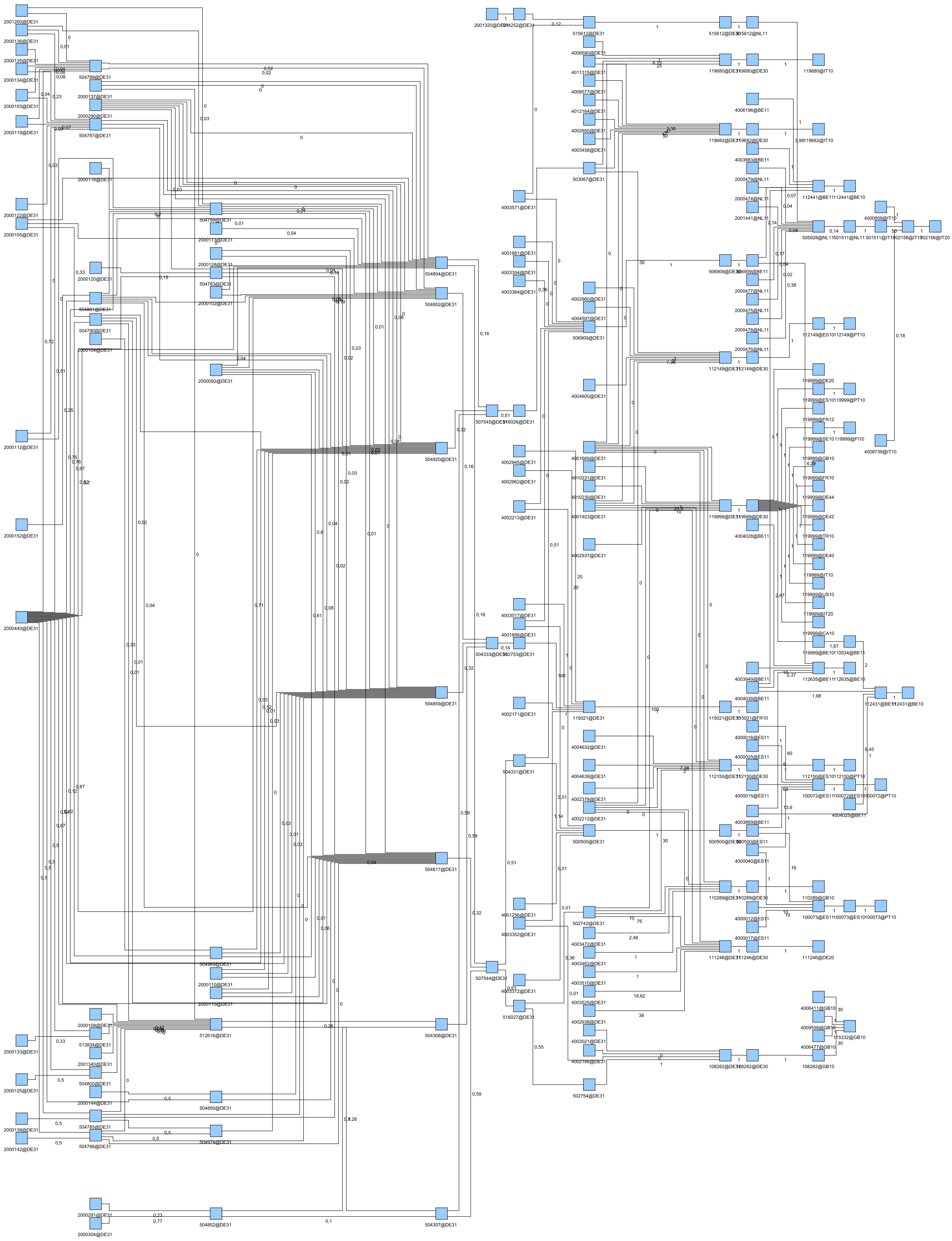


Figure B.6: Network visualisation for M_1/S_5

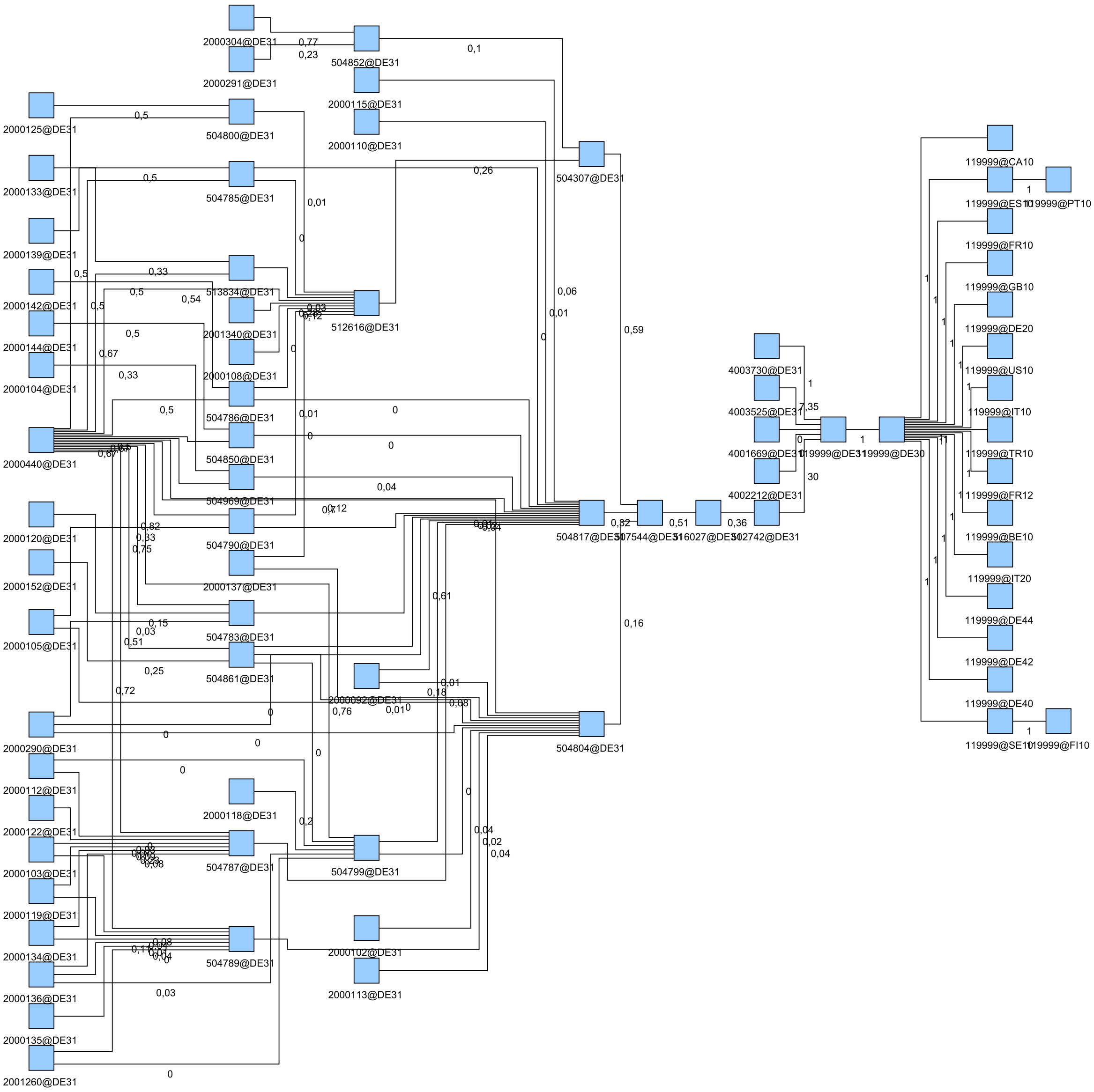


Figure B.7: Network visualisation for M_1/S_6