

TECHNISCHE UNIVERSITÄT MÜNCHEN  
LEHRSTUHL FÜR TECHNISCHE ELEKTRONIK  
FACHGEBIET HALBLEITERPRODUKTIONSTECHNIK

# Optimierungsmodelle in der Halbleiterproduktionstechnik

Kirsten Hilsenbeck

Vollständiger Abdruck der von der Fakultät für  
Elektrotechnik und Informationstechnik der Technischen Universität München  
zur Erlangung des akademischen Grades eines  
**Doktor-Ingenieurs**  
genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr.-Ing. Ulf Schlichtmann  
Prüfer der Dissertation: 1. Univ.-Prof. Dr.-Ing. Walter Hansch  
2. Univ.-Prof. Dr. rer. nat. Ignaz Eisele,  
Universität der Bundeswehr München

Die Dissertation wurde am 20.10.2004 bei der  
Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik  
und Informationstechnik am 25.01.2005 angenommen.







# Abstract

The progress in semiconductor industry is based on a reduction of cost per function by 35% a year. Half of this reduction is reached by shrinking, enlarging wafer size and yield enhancement. Because of the advanced technology stage, it is necessary that in future 50% of cost reduction will be obtained by optimising productivity in manufacturing. The field of semiconductor manufacturing technology (product engineering) deals with methods of describing, valuating and optimising the productivity of semiconductor manufacturing.

In this work, at first the technical structures of semiconductor manufacturing are described and an overview of methods to measure and increase productivity is given.

For the field of planning, the subject of delivery reliability is examined. Key data of delivery reliability is used for performance assessment as well as target setting. To examine the interdependencies of targets within the inner supply chain (frontend, backend, delivery) of semiconductor manufacturing a time discrete model was built. The results show that targets are interdependent and can not be set independently. By changing strategy the resulting performance can be improved. To achieve better planning it is suggested to change delivery weeks to daily due dates.

For the field of control and logistics two dispatching systems are developed and examined. A dispatch-own-system synchronises the manufacturing resources and supports the lot dispatching to the machines. Such a system was developed and examined in real manufacturing, where it could shorten the preparation time for processes for more than 28%. Additionally, this is not only optimising the performance but it is also the first step towards a paperless fabrication.

The so called dispatch-ahead-function is controlling the order of production corresponding to the needs of the following workcenter. To examine its characteristics a simulation model was created. For this reason, the simulation software AutoMod was tested for its usability in semiconductor manufacturing. After several modifications the model was built and simulated. The results show that the machine utilisation can be increased and inventory can be reduced at the same time. So the dispatch-function offers a tool to optimise bottleneck systems.

The investigation of the three aspects could show clear optimisation potentials for semiconductor manufacturing.



# Kurzzinhalt

Der Fortschritt in der Halbleiterindustrie basiert u.a. auf einer Kostenreduktion pro Bauteil von 35% pro Jahr. Die Hälfte davon wird durch Shrinking, Wafervergrößerung und Ausbeuteverbesserung erreicht. Da die Technologie schon sehr weit fortgeschritten ist, müssen zukünftig 50% der Kostenreduktion durch Produktivitäts-Optimierung der Herstellung erzielt werden. Das Fachgebiet der Halbleiterproduktionstechnik beschäftigt sich mit den entsprechenden Methoden zur Beschreibung, Bewertung und Verbesserung der Produktivität einer Halbleiterfertigung.

In dieser Arbeit werden zunächst die produktionstechnischen Strukturen einer Halbleiterfertigung formuliert und ein Überblick über die Methoden der Produktivitätsbestimmung und -steigerung gegeben.

Für den Bereich der Planung wurde das Thema Liefertreue untersucht. Kennzahlen der Liefertreue werden als Leistungsgrößen und als Plan-Vorgaben verwendet. Um die Wechselwirkungen der Plan-Vorgaben in der internen Supply-Chain (Frontend, Backend, Delivery) einer Halbleiterfertigung zu untersuchen, wurde ein zeitdiskretes Modell entwickelt. Die Ergebnisse zeigen, dass sich die einzelnen Leistungs-Vorgaben gegenseitig beeinflussen und somit nicht unabhängig festgelegt werden können. Durch Veränderungen der Strategien konnten Verbesserungen in der resultierenden Performance erzielt werden. Um eine genauere Planung erreichen zu können, wird der Übergang von Lieferwochen hin zu tagesfeinen Zielgrößen vorgeschlagen.

Für den Bereich der Steuerung und Logistik wurden Dispatch-Systeme entwickelt und untersucht. Ein Dispatch-Own-System, das die Synchronisation der Produktionsfaktoren verbessert und die Verteilung der Lose auf die Anlagen vereinfacht, wurde entwickelt und in der Praxis erprobt, wobei Prozess-Vorbereitungszeiten um mehr als 28% verkürzt werden konnten. Die Realisierung trägt nicht nur zur Verbesserung der Performance und Erhöhung der Kapazität des Workcenters bei. Sie ist darüber hinaus auch ein wichtiger Schritt in Richtung der papierlosen Fertigung (Paperless Fab).

Die so genannte Dispatch-Ahead-Funktion steuert die Produktionsreihenfolge einer Arbeitsstation entsprechend der Bedürfnisse der nachfolgenden Station. Um ihre Eigenschaften zu untersuchen, wurde sie in einer Simulationsumgebung modelliert. Hierzu wurde die Software AutoMod auf ihre Verwendbarkeit in der Halbleiterfertigung getestet und nach entsprechender Anpassung Modellrechnungen durchgeführt. Es zeigte sich, dass gleichzeitig die Maschinenauslastungen erhöht und die Lagerbestände verringert werden konnten. Somit bietet die Dispatch-Funktion ein Werkzeug, um Bottleneck-Anlagen zu optimieren.

Die Analysen der drei untersuchten Systeme konnten jeweils deutliche Verbesserungspotentiale aufzeigen.





# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Formulierung der produktionstechnischen Strukturen einer Halbleiterfertigungslinie</b>	<b>7</b>
2.1	Einleitung .....	7
2.2	Moore's Law und die Entwicklung der Halbleiterindustrie .....	8
2.3	Charakteristiken der Chip-Herstellung .....	12
2.3.1	Halbleiterfab .....	12
2.3.2	Schichtherstellung und Produktvielfalt .....	13
2.3.3	Fablayout .....	15
2.3.4	Der Faktor Mensch .....	16
2.3.5	Die Fototechnik .....	17
2.3.6	Produktionsfluss und Liniensteuerung .....	17
2.4	Supply Chain Management .....	18
2.5	Performance-Kennzahlen der Anlagen .....	20
2.5.1	Standards .....	20
2.5.2	Maschinenzustände .....	22
2.5.3	Overall Equipment Efficiency OEE .....	23
2.5.4	Das Vier-Partner-Modell .....	26
2.6	Performance-Kennzahlen der Produktion .....	29
2.6.1	Flussfaktor und Little's Law .....	29
2.6.2	Taktrate .....	30
2.6.3	Dispatch-Regeln .....	32
2.6.4	Dynamische Kennzahlen .....	33
2.6.5	Liefertreue - Confirmed Line Item Performance CLIP .....	36
2.6.6	Ausbeute - Yield .....	37
2.6.7	Weitere Kennzahlen .....	39
2.6.8	Betriebskennlinie und Variabilität .....	40
2.6.9	Overall Fab Efficiency OFE .....	45
<b>3</b>	<b>Liefertreue in mehrstufigen Produktionen</b>	<b>47</b>
3.1	Einleitung .....	47
3.2	Statistische Grundlagen .....	48
3.2.1	Wahrscheinlichkeitsrechnung .....	48
3.2.2	Die Normalverteilung .....	49
3.2.3	Diskretisierung kontinuierlicher Zufallsvariablen .....	51

3.2.4	Faltungssatz für diskrete Zufallsvariablen .....	53
3.2.5	Notation .....	53
3.3	Confirmed Line Item Performance und On-Time-Delivery .....	54
3.3.1	Einstufige Betrachtung der On-Time-Delivery .....	55
3.3.2	Näherung zur Abschätzung der On-Time-Delivery .....	68
3.4	On-Time-Delivery mit Due Date Weekly und Release Backend .....	70
3.4.1	Frontend .....	71
3.4.2	Backend .....	73
3.4.3	Manufacturing .....	75
3.4.4	Delivery und Over All .....	81
3.4.5	Ergebnisse .....	83
3.5	On-Time-Delivery mit Due Date Daily und Release Backend .....	97
3.5.1	Voraussetzungen .....	97
3.5.2	Modell .....	99
3.5.3	Bewertung .....	99
3.6	On-Time-Delivery mit Due Date Daily ohne Release Backend .....	101
3.6.1	Voraussetzungen .....	101
3.6.2	Modell .....	104
3.6.3	Bewertung .....	105
3.7	Rechenkomplexität und Implementierung .....	108
3.8	Zusammenfassung .....	108
3.9	Ausblick .....	109
<b>4</b>	<b>Stepper-Dispatcher</b> .....	<b>111</b>
4.1	Einleitung .....	111
4.2	Dispatch Own .....	112
4.2.1	Vorbereitung und Durchführung von Belichtungsoperationen .....	112
4.2.2	Synchronisation .....	114
4.2.3	Performancesteigerung durch Dispatch Own .....	118
4.3	Dispatch Ahead .....	124
4.3.1	Steuerung der Voroperation .....	124
4.3.2	Simulation zweier Arbeitsstationen .....	125
4.3.3	Simulation eines Dispatch Ahead Szenarios .....	129
4.3.4	Performancesteigerung durch Dispatch Ahead .....	131
4.3.5	Rechenkomplexität und Implementierung .....	133
4.4	Zusammenfassung und Ausblick .....	133
<b>5</b>	<b>Resumé</b> .....	<b>135</b>
	Abbildungsverzeichnis .....	137
	Tabellenverzeichnis .....	143
	Literaturverzeichnis .....	145
	Formelzeichen und Abkürzungen .....	149
	Tabelle der Standardnormalverteilung .....	155
	Danksagung .....	157

# Kapitel 1

## Einleitung

Das Wachstum der Halbleiterbranche sowie die ständig steigende Chip-Performance beruht auf einer jährlichen Kostenreduktion von ca. 30% pro Transistorfunktion. Dennoch bleiben die Preise auf gleichem Niveau. Bisher wurde diese Kostenreduktion hauptsächlich durch das Verringern der Bauteilgröße, das so genannte Shrinking, Vergrößerung der Wafer und Steigerung der Ausbeute (Yield) erreicht. Die Möglichkeiten dazu nehmen jedoch immer stärker ab, wodurch das vierte Mittel der Kostenreduktion, die Produktivitätssteigerung, immer wichtiger für das Überleben der einzelnen Halbleiterhersteller wird. Abbildung 1.1 [SEMA] zeigt, wie sich die Anteile der einzelnen Faktoren in den vergangenen Jahren verändert haben und was in Zukunft zu erwarten ist. So wird im Jahr 2010 die Steigerung der Produktivität 50% der Kostenreduktion erbringen müssen. Optimierung der Logistik und Verbesserung der Maschinenauslastung sind dazu zwei der bekanntesten Schlagworte.

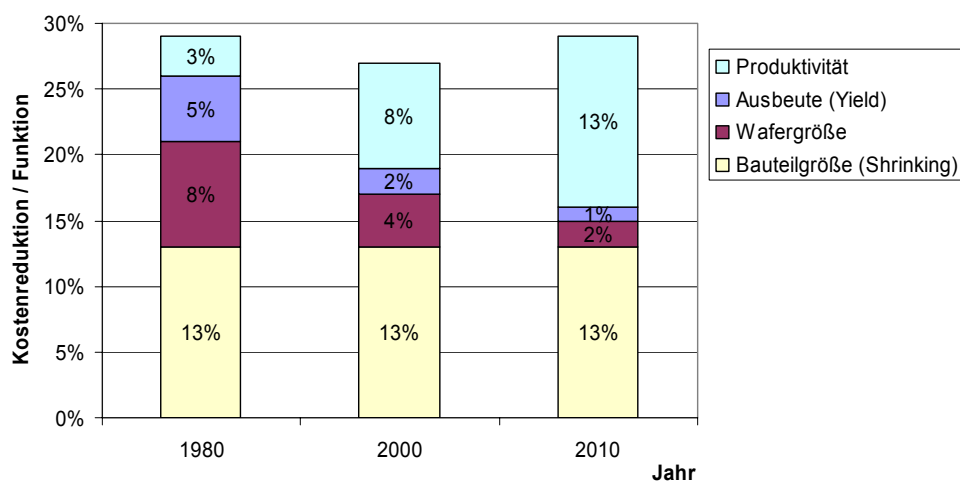


Abbildung 1.1: Pro Jahr und Transistorfunktion werden 30% Kostenreduktion benötigt

Um dieser Entwicklung Rechnung zu tragen, stiftete die Infineon AG 1999 an der Technischen Universität München die Professur „Halbleiterproduktionstechnik“, auf die Dr. W. Hansch vom Research Institute of Nanodevices and Systems, Hiroshima, Japan, berufen wurde. Die Professur wurde in der Fakultät für Elektrotechnik und Informationstechnik, am Lehrstuhl für Technische Elektronik, angesiedelt. Aufgaben dieser ersten Professur für dieses Fachgebiet in Deutschland sind die wissenschaftliche Bearbeitung entsprechender Themen aus der Halbleiterproduktionstechnik sowie die Ausbildung von Studenten in diesen Bereichen.

Mit dieser Dissertation soll festgestellt werden, mit welchen Fragestellungen das Thema wissenschaftlich behandelt werden kann. Ist es möglich, von einer, zu einem Unternehmen als extern zu betrachtenden Universität aus, produktionstechnische Theorien aufzustellen und zu optimieren und inwieweit müssen Kontakte zu realen Produktionsstätten hergestellt werden.

Zunächst jedoch ist die Frage zu klären, was Halbleiterproduktionstechnik ist und in welche Teilgebiete sie untergliedert werden kann. Dazu ist es hilfreich, in einer Abstraktionsebene höher, das Gebiet der Produktionstechnik zu definieren.

An der Fachhochschule Köln wird ein eigener Studiengang Produktionstechnik angeboten [FHK]. Aus den Lehrinhalten (siehe Abbildung 1.2) können Parallelen zu den klassischen Studiengängen Elektrotechnik, Maschinenbau und Betriebswirtschaftslehre gezogen werden. Das zeigt bereits deutlich, wie interdisziplinär die Produktionstechnik ist. Im Hauptstudium kann zwischen den beiden Studienschwerpunkten Produktionsorganisation und Fertigungstechnologie gewählt werden.

Fachhochschule Köln:  
Studiengang Produktionstechnik

**Grundstudium**

- Datenverarbeitung
- Industriebetriebswirtschaftslehre
- Konstruktionslehre
- Mathematik
- Physik
- Steuerungs- und Regelungstechnik
- Technische Mechanik
- Werkstoffkunde

**Hauptstudium**

- Arbeitswissenschaft
- Betriebsorganisation
- Fertigungsverfahren
- Grundlagen Fertigungsmittel
- Grundlagen Produktionsplanung und -steuerung
- Kosten- und Optimierungsrechnung
- Unternehmensführung und Wirtschaftsrecht

**Produktionsorganisation**

- Angewandte Statistik
- Fabrikplanung
- Produktionsplanung und -steuerung
- Unternehmensforschung

**Fertigungstechnologie**

- Automatisierungstechnik
- Fertigungsmittel
- Produktionslogistik
- Qualitätsmanagement

Wahlpflichtfächer für beide Studienschwerpunkte  
Montagetechnik oder Ökologie in der Produktion  
Projekt

Abbildung 1.2: Grund- und Hauptstudium sowie Studienrichtungen des Studiengangs Produktionstechnik an der Fachhochschule Köln

Eine andere fachliche Aufteilung des Begriffs der Produktionstechnik kann aus den Tätigkeitsfeldern der Wissenschaftlichen Gesellschaft für Produktionstechnik gewonnen werden. Die Wissenschaftliche Gesellschaft für Produktionstechnik ist ein Zusammenschluss

von über 30 Lehrstühlen und Instituten der Produktionstechnik. „Sie versteht sich als Organ zur Vertretung der Belange von Forschung und Lehre auf dem Gebiet der Produktionstechnik im wissenschaftlichen, gesellschaftlichen und politischen Raum der Bundesrepublik Deutschland“ [WGP]. Aus Abbildung 1.3 ist die starke maschinenbauliche Orientierung der Gesellschaft zu erkennen. Das spiegelt den Sachverhalt wieder, dass die Produktionstechnik ursprünglich aus dem Maschinenbau stammt, da es in erster Linie Maschinen sind, die für eine industrielle Produktion charakteristisch sind.

Interessant ist auch, dass der Bereich des Qualitätsmanagements hier dem Produktionsmanagement zugeordnet wird, während in der Fachhochschule Köln das Qualitätsmanagement in die Fertigungstechnologie fällt. Das lässt schon erahnen, dass die Aufteilung in (und Zuordnung zu) einzelnen Teilgebieten nicht einfach und bisweilen auch nicht möglich ist. Veränderungen eines Bereiches haben immer auch Auswirkungen auf andere Bereiche. So darf bei der Auseinandersetzung mit einer detaillierten Aufgabenstellung der Blick für das Ganze, also die vollständige Produktion, nie verloren gehen.

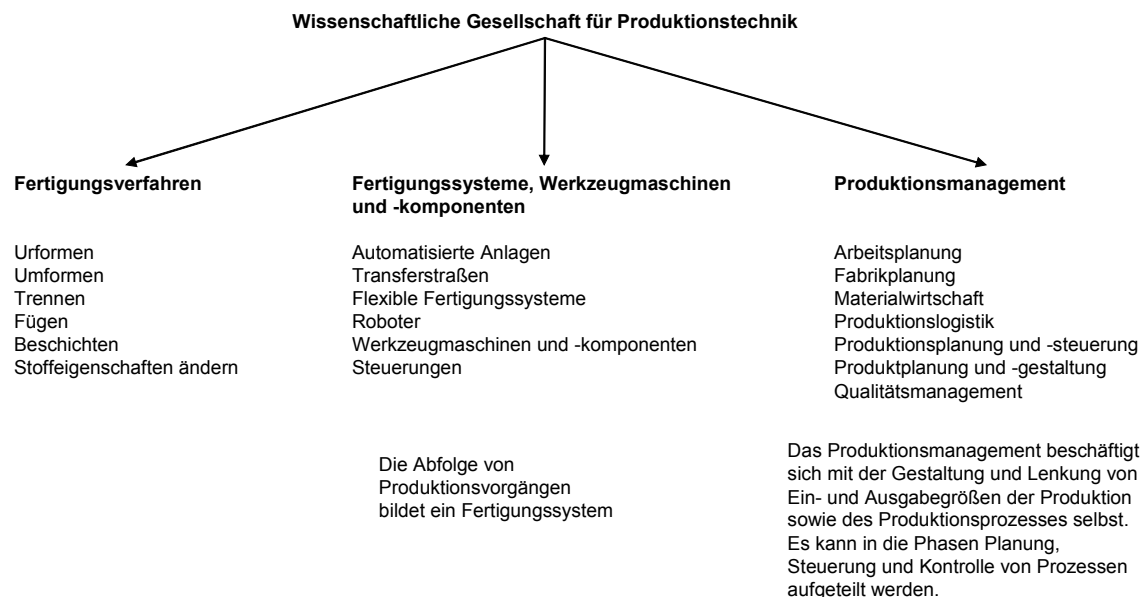


Abbildung 1.3: Tätigkeitsfelder der Wissenschaftlichen Gesellschaft für Produktionstechnik

Eine weitere interessante Feststellung ist, dass sich zwar jeder etwas unter dem Begriff Produktionstechnik vorstellt, bei weiteren Nachfragen stellt sich jedoch heraus, dass jede Person an etwas anderes dachte. Entsprechend des jeweiligen Fachgebietes aus dem eine Person stammt, füllt sie den Begriff Produktionstechnik mit unterschiedlichen Inhalten. Auch die Spezialisierung in Halbleiterproduktionstechnik führt nicht zu einem einheitlicheren Begriffsverständnis. Während ein Elektrotechniker nur an die Entwicklung von Bauteilen und die Funktionsweise der Produktionsmaschinen, also an die Technologie denkt, wird im Maschinenbau bereits das Zusammenspiel von Maschinen in einem Produktionsverbund betrachtet. Bei den Wirtschaftswissenschaften wiederum liegt der Schwerpunkt auf Kundenorientierung und Entwicklung von Produktspektren. Abbildung 1.4 stellt die unterschiedlichen Betrachtungsweisen den unterschiedlichen Abstraktionsebenen zur Produktionsbetrachtung gegenüber. Jedoch kann keine der drei Arbeitsebenen ohne

Kenntnis über den Entwicklungsstand der anderen optimale Ergebnisse erzielen. Im klassischen Fall muss für die Erstellung eines Produktdesigns die Fähigkeit der Herstellung bekannt sein. Fähigkeiten einer Produktion basieren auf technologischem Können. In der Halbleiterherstellung dreht sich diese Reihenfolge bisweilen um: Kundenanforderungen und Produktdesign sind Vorgaben für die Entwicklung der entsprechenden technologischen Prozesse. Meist gibt es aber keine Kenntnisse über logistische Zusammenhänge einer Produktion innerhalb der elektrotechnischen Disziplinen. Umso wichtiger ist es, die Produktionstechnik auch im Gebiet der Elektrotechnik zu etablieren. Da dort die Herstellung von Halbleiterbauteilen die komplexeste Produktion darstellt, ist sie besonders geeignet, um Methoden der Produktionstechnik anzuwenden und zu untersuchen.

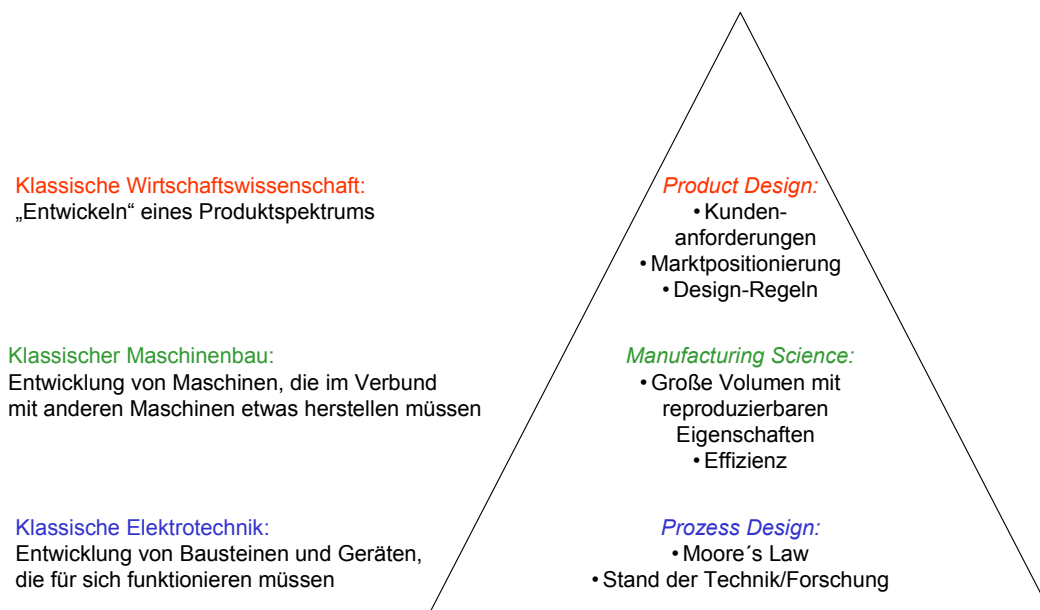


Abbildung 1.4: Betrachtungsebenen einer Produktion

Mit der Kenntnis der vorangegangenen Überlegungen und unter Berücksichtigung der technischen Möglichkeiten, soll nun das Fachgebiet Halbleiterproduktionstechnik in die Bereiche Technologie und Management/Organisation aufgeteilt werden. Abbildung 1.5 zeigt die Feingliederung in Arbeitsgebiete, die realisiert werden sollen. Die vorliegende Dissertation untersucht Umsetzungsmöglichkeiten des Bereiches Management/Organisation, das wiederum in Planung, Steuerung und Logistik, Performance-Evaluierung und Qualitätsmanagement unterteilt ist.

Für sinnvolle Modellbildungen ist es notwendig, Zugang zu realen Daten zu erhalten. Die Zusammenarbeit mit der Infineon AG in München Neuperlach hat sich hier als sehr vorteilhaft erwiesen. Bei der Fertigung in Neuperlach handelt es sich um die Produktion von Logikbausteinen, wobei hauptsächlich niedrigvolumige Produkte hergestellt werden. Dies hat eine nicht unerhebliche Auswirkung auf die Produktionsstruktur: Der Automatisierungsgrad ist relativ gering und die Vorhersagbarkeit von Ereignissen und Zuständen ist sehr schwierig.

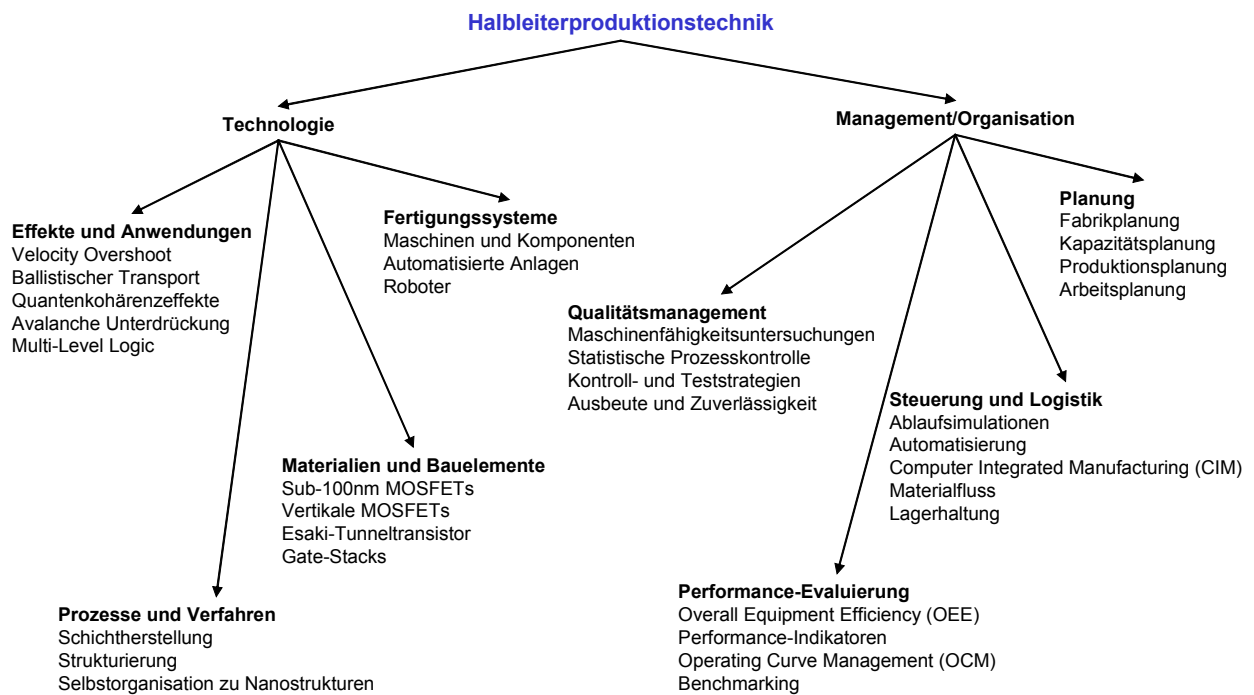


Abbildung 1.5: Ausgewählte Arbeitsgebiete der Halbleiterproduktionstechnik

Diese Arbeit beschäftigt sich im Bereich Planung mit dem Thema Liefertreue (Kapitel 3), für die Steuerung und Logistik wurden ein Dispatch-Systeme entwickelt und untersucht (Kapitel 4). In weiteren Arbeiten sollen für die Performance Evaluierung neue Bewertungssysteme entwickelt werden und im Bereich des Qualitätsmanagement für den lehrstuhleigenen Reinraum ein Qualitätsmanagementsystem aufgestellt werden. Somit sind die Arbeitsgebiete des Bereiches Management/Organisation aus Abbildung 1.5 abgedeckt. Die notwendigen Grundlagen und theoretischen Hintergründe werden zunächst in Kapitel 2 formuliert.





# Kapitel 2

## Formulierung der produktionstechnischen Strukturen einer Halbleiterfertigungslinie

### 2.1 Einleitung

Die Herstellung von Halbleiterchips ist ein komplexes System aus vielen Einzelschritten. Ihre Steuerung gliedert sich somit in viele Teilbereiche und ist daher nicht leicht überschaubar. Vor allem die vielen Möglichkeiten, das System und seine Teilbereiche aus unterschiedlichen Blickwinkeln zu betrachten, führt dazu, dass es keine geschlossene Darstellung der Halbleiterfertigung gibt.

Seit Mitte der 90er Jahre werden verstärkt Bemühungen unternommen, die Effizienz einer Halbleiterfabrik zu bewerten und zu verbessern. Die unterschiedlichen Ansätze, das System zu beschreiben, beginnen im Kleinen mit der Zeitmessung von unterschiedlichen Maschinenzuständen. Aus den Fertigungsfortschritten der Wafer können weitere Parameter kreiert werden, die Eigenschaften einzelner Bereiche oder der gesamten Fab bewerten. Der nächste Schritt wäre die Beschreibung der Gesamteffizienz einer Halbleiterfabrik, also der Übergang von der Beschreibung der Effizienz eines einzelnen Tools (Overall Equipment Efficiency, OEE) zur Overall Fab Efficiency OFE. Dazu gibt es jedoch noch keine allgemein anerkannte Methode.

Bei der Beschäftigung mit Fragestellungen der Performance einer Halbleiterfab fällt auf, dass der Kostenbegriff nur sehr selten bis gar nicht verwendet wird. Während anfallende Kosten über Kostenstellen ermittelt und mittels mathematischer Verfahren unterschiedlichen Bezugskategorien zugeordnet werden, ist die umgekehrte Vorgehensweise oft nicht möglich. Für Verbesserungen in Steuerung und Performance der Halbleiterproduktion können Kosten bzw. Kosteneinsparungen meist nur abgeschätzt, jedoch nicht errechnet werden. Was ist z.B. die Reduzierung einer Durchlaufzeit oder das Einhalten eines Liefertermins wert? Wie kann eine daraus resultierende Kundenzufriedenheit monetär bewertet werden? Entscheidungen für entsprechende Maßnahmen werden dadurch erschwert. Dennoch darf die Bedeutung dieser Maßnahmen nicht unterschätzt werden. Eine kontinuierliche Kostensenkung kann nur durch kontinuierliche Performance-Steigerungen erreicht werden.

In diesem Kapitel werden die wesentlichen Parameter und produktionstechnischen Strukturen einer Halbleiterfertigungslinie zusammengetragen, um einen Überblick und das Verständnis für Fragestellungen der Halbleiterproduktionstechnik zu entwickeln.

## 2.2 Moore's Law und die Entwicklung der Halbleiterindustrie

1948 entdeckten Walter Brattain, John Bardeen und William Shockley den Bipolartransistor, der die relativ großen Röhren in der Computertechnik ablöste. In den Jahren 1958 bis 1960 wurde die Planartechnologie erfunden, womit sich MOSFETs (Metal Oxide Feldeffekttransistoren) und daraus integrierte Schaltungen (ICs) realisieren ließen. Die Vorteile gegenüber der Bipolartechnologie liegen in:

- der Strukturgröße: Bei gleicher minimaler Strukturgröße benötigt ein MOSFET nur 1/10 des Platzes eines Bipolartransistors, was die Entwicklung von integrierten Schaltungen ermöglicht.
- der Skalierbarkeit: Verringerungen der Bauteilabmessungen im gleichen Verhältnis führen zu gleich bleibenden elektrischen Eigenschaften, ohne dass ein neues Design entwickelt werden muss. Die Möglichkeit, die Eigenschaften der Bauteile so vorherzusagen, führt zur Erstellung von Roadmaps. Durch die Skalierbarkeit wird die Kostenreduktion mittels Shrinking erst ermöglicht.
- der CMOS-Technologie: Verbrauchsarme Schaltungen in CMOS-Technologie sind wiederum besonders für integrierte Schaltungen geeignet.

Nachdem 1965 bereits fünf Jahre MOSFETs produziert wurden, gelangte Gordon E. Moore zu einer Aussage, die später als Moore's Law bekannt wurde: Jedes Jahr verdoppelt sich die Anzahl der Transistorfunktionen pro Chip. Die Kosten pro Transistor werden dabei halbiert [Moo65]. Seit 1975 hatte sich diese Verdoppelung der Chip-Komplexität verlangsamt und fand in einem Zeitintervall von eineinhalb Jahren statt, so dass Gordon E. Moore für die darauf folgende Zeit eine weitere Verlangsamung zu einem Zeitintervall von zwei Jahren vorhersagte [Moo75]. Zwischen 1975 und 2001 erfolgte die Verdoppelung tatsächlich jeweils innerhalb von zwei Jahren (siehe Tabelle 2.3) [Hut05].

Jahr	Gesamt	MPU	DRAM
1959-1956	12		
1966-1975	17	33	17
1976-1985	22	22	25
1986-1995	32	38	25
1996-2001	22	31	15
1976-2001	24	27	20

Tabelle 2.1: Durchschnittliche Zeit in Monaten, in denen die Bauteilkomplexität verdoppelt wurde, für Prozessoren (MPU) und Speicher (DRAM)

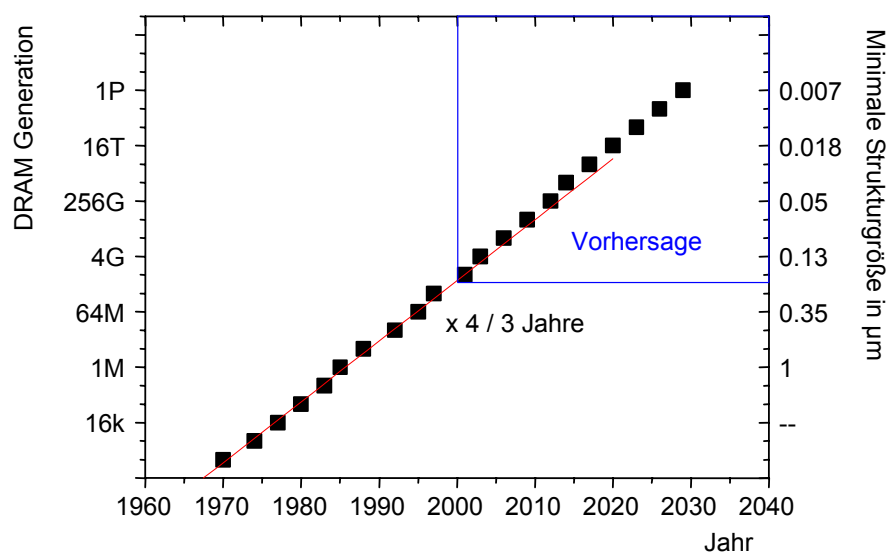


Abbildung 2.1: Moore's Law [Han03]

Das Mooresche Gesetz ist dabei kein physikalisches Gesetz im eigentlichen Sinne, vielmehr eine Beobachtung, die zu einer Roadmap extrapoliert werden kann. Schließlich hat sich herausgestellt, dass das Mooresche Gesetz hauptsächlich eine ökonomische Bedeutung hat: Halten sich die Hersteller daran, können sie ihren Gewinn maximieren, da sie so genau das Potential des Marktes (mit all seinen Nebenbedingungen wie Abschreibungen, Amortisation, etc.) treffen. Darüber hinaus erleichtert die Vorhersagbarkeit den Firmen ihre Planungen.

Seit 1992 gibt es eine detaillierte Technologieroadmap, ITRS, International Technology Roadmap for Semiconductors. Sie beschreibt für MOSFETs die technisch notwendigen Eigenschaften, die für die Einhaltung des Mooreschen Gesetzes erforderlich sind. Dabei enthält sie keinerlei Aussagen darüber, wie die Anforderungen technologisch erfüllt werden können. Schon mehrfach wurde in den letzten Jahren befürchtet, dass die Mikrochip-Entwicklung durch physikalische Grenzen gestoppt werden könnte. Bisher konnte jedoch immer wieder mit unvorhergesehenen Entdeckungen neuer Technologien Moore's Law eingehalten werden. Momentan erscheint die „Red Brick Wall“ (Anforderungen ohne bekannte Realisierungsmöglichkeiten werden in der ITRS rot hinterlegt) im Jahr 2008 [Mil04].

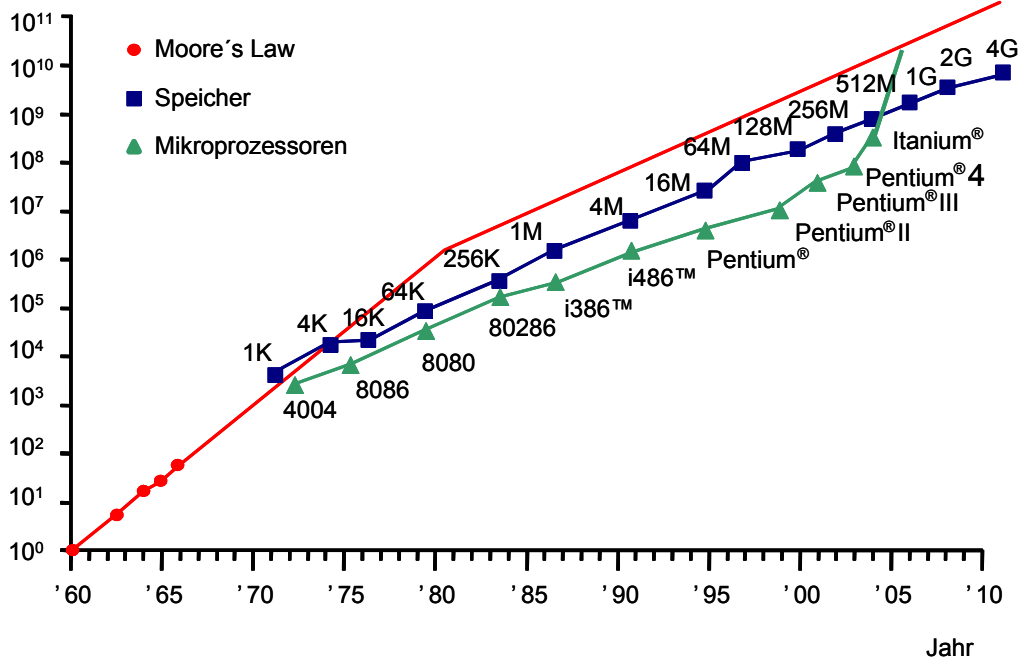
Die Halbierung der Kosten pro Funktion in 1,5 Jahren ist gleichbedeutend mit einer Kostenreduktion von 30% pro Jahr. Da dies gleichzeitig mit einer Funktionalitätsverdoppelung pro Chip erfolgt, bleiben die Preise pro Chip in etwa stabil. Oder anders ausgedrückt: Wenn die Preise gleich bleiben, müssen die Kosten pro Jahr um 30% sinken.

Die Kostenreduktion pro Funktion wird mit Hilfe unterschiedlicher Methoden erzielt. Den größten Anteil hatte bisher die Verkleinerung der Strukturgröße (Shrinking). Etwa alle zwei Jahre wird bei den Strukturgrößen ein neuer Technologieknoten erreicht. Dadurch können mehr Chips pro Wafer produziert werden. Tabelle 2.2 zeigt Zusammenhänge zwischen Strukturgrößen und Preis pro Funktion. Die Entwicklung der Transistorenanzahl pro Chip ist in Abbildung 2.2: dargestellt.

Jahr	Minimale Strukturgröße	DRAM Preis/Bit in USD millicents	DRAM Performance in Bitkapazität	Prozessoren Preis/Transistor in USD cents	Prozessoren Performance in Systemgeschwindigkeit
1970	10 µm	492	1 Kbit	50	0,5 MHz
1980	1,3 µm	15	64 Kbit	3,5	5 MHz
1990	0,5 µm	0,48	4 Mbit	0,25	50 MHz
2000	0,18 µm	0,015	256 Mbit	0,018	1 GHz
2010	0,05 µm	0,00047	4 Gbit	0,0013	50 GHz

Tabelle 2.2: Preis pro Funktion in Bezug auf Strukturgröße, Speicherkapazität und Prozessorgeschwindigkeit [Woj02]

Transistor/Chip



(Quelle: Intel)

Abbildung 2.2: Entwicklung der Transistorenanzahl pro Chip [Man02]

Der gleiche Effekt (mehr Chips pro Wafer) wird durch größere Waferdurchmesser erreicht. So bringt der Übergang von 200 mm auf 300 mm eine Flächenerhöhung um den Faktor 2,25 (quadratisch), während die Kosten nur linear um den Faktor 1,5 steigen. Da eine Erhöhung der Wafergrößen mit hohen Investitionen verbunden ist, wird hier nur etwa alle 8-10 Jahre ein Technologiesprung vorgenommen. Abbildung 2.2 stellt die Entwicklung der Chipanzahl pro Wafer dar.

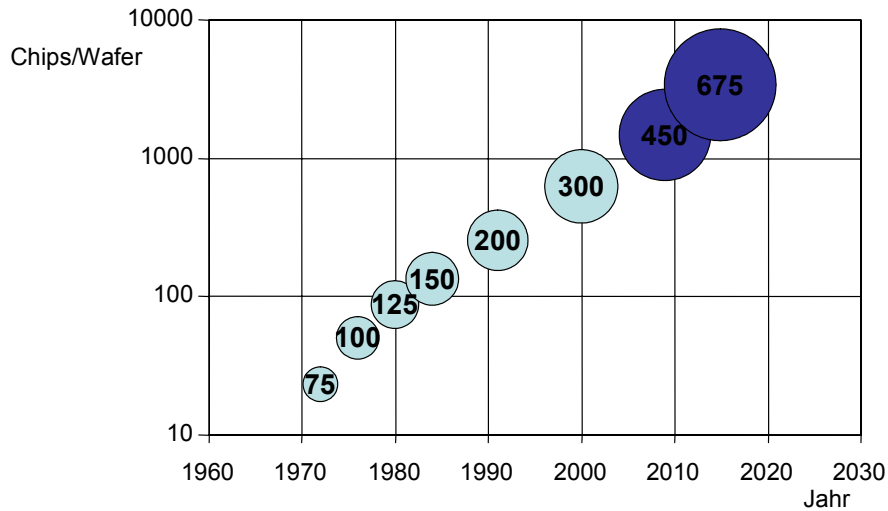


Abbildung 2.3: Entwicklung der Chipmenge pro Wafer [Han04]

Schließlich hängen die Kosten pro Chip auch von der Ausbeute (Yield) der Produktion ab. Yield ist eine Funktion von Chipgröße (je größer der Chip desto größer die Wahrscheinlichkeit eines Defekts), Strukturgröße (kleinere Strukturen sind „anfälliger“) und Schichtanzahl (jede Schicht eröffnet eine neue Defekt-Möglichkeit). Der Yield steigt mit zunehmender Prozess Erfahrung. Wenn die Produktion eines Produktes nach der Entwicklungsphase mit einer Ausbeute von 30% beginnt, muss der Yield innerhalb des ersten Produktionsjahres auf ca. 80% gesteigert werden [EIK95] [ITRS03a]. In günstigen Fällen sind auch Ausbeuten um die 95% möglich.

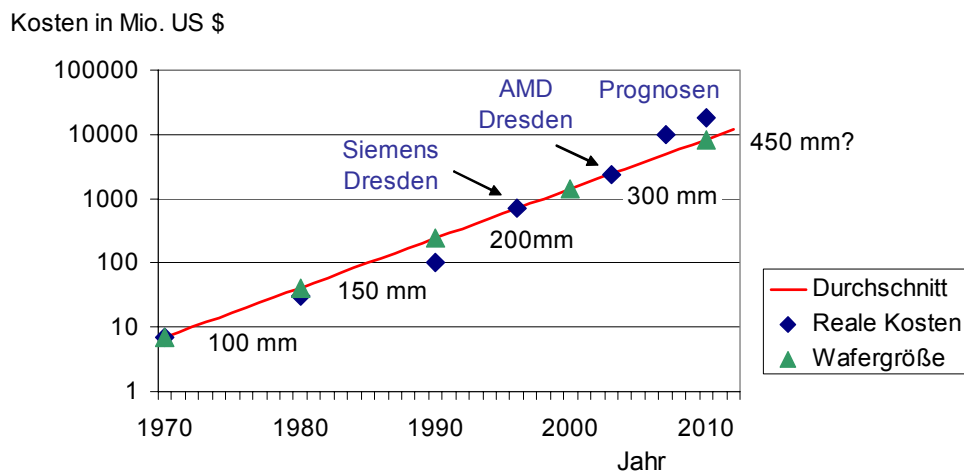


Abbildung 2.4: Entwicklung der Kosten für Halbleiterfabs (Quelle: Texas Instruments)

Die fortschreitende Entwicklung der Technologie führt dabei aber nicht nur zur erwünschten Kostenreduktion, sie verursacht ihrerseits neue Kosten. Immer kleinere Strukturen müssen mit immer teureren Geräten produziert werden. Die Baukosten für

Halbleiterfabriken steigen deshalb mit Faktor 10 in 15 Jahren (siehe Abbildung 2.4). Allein 70% davon sind Kosten von Produktionsanlagen. Das bedeutet, die Geräte müssen möglichst effizient sein. Die steigenden Kosten müssen durch weitere Möglichkeiten der Kostenreduktion aufgefangen werden. Während in den 80er Jahren die Ausbeute deutlich verbessert werden konnte, begann Mitte der 90er Jahre die Optimierung der Equipment-Nutzung. Seit 2000 wird versucht, die Produktivität des gesamten Herstellungsprozesses zu verbessern.

## 2.3 Charakteristiken der Chip-Herstellung

### 2.3.1 Halbleiterfab

In einer Halbleiterfabrik, auch Fab genannt, werden Halbleiterchips hergestellt. Auf Halbleiterchips befinden sich entweder **Logikbausteine** (Integrated Circuits, ICs) oder **Speicherbausteine** (z.B. DRAMs). Für die Herstellung ist es ein entscheidender Unterschied, welche Halbleiterchips produziert werden, da die unterschiedliche Produktvielfalt Auswirkungen auf die logistischen Anforderungen hat. Im Allgemeinen haben Speicher bessere produktionstechnische Eigenschaften, weil sie in größeren Stückzahlen hergestellt werden. Deswegen können Verbesserungsmaßnahmen besonders auf dem Logiksektor wirken. Die Ausführungen dieser Arbeit beziehen sich auf die Herstellung von Logikbausteinen mit der zusätzlichen Besonderheit einer hohen Produktvielfalt bei geringem Volumen pro Produkt.

Die Herstellung von Halbleitern findet in Reinräumen statt. Je nach Strukturgröße der Bauelemente werden unterschiedliche Anforderungen an die Reinheit der Produktionsumgebung gestellt. Bereitstellung und Unterhalt der **Reinraumtechnik** sind sehr kostspielig. Zur Zeit verursacht das Errichten einer Halbleiterfab Kosten von ca. 2 Milliarden Euro, wovon zwei Drittel für Geräte aufgewendet werden müssen. Um die enormen Investitionen möglichst gut ausnutzen zu können, wird rund um die Uhr, d.h. 24 Stunden am Tag und sieben Tage pro Woche, produziert. Die Unterhaltskosten für die Produktion liegen dabei bei einigen Millionen Euro pro Tag.

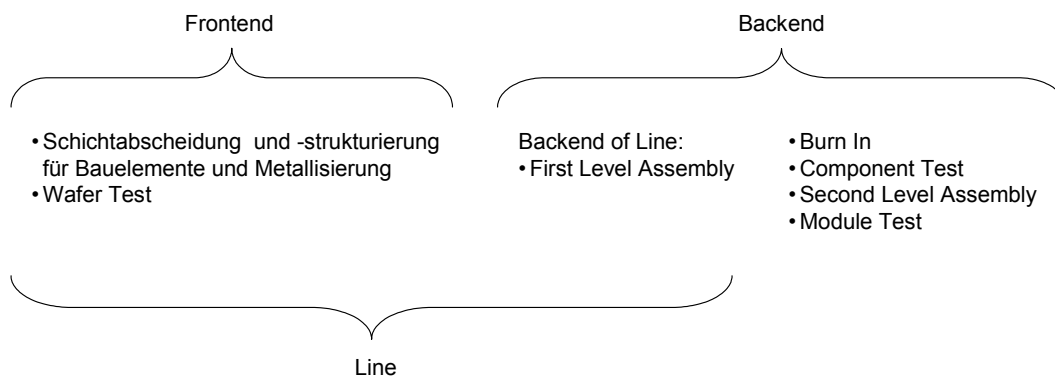


Abbildung 2.5: Aufteilung der Halbleiterproduktionslinien in Frontend und Backend

Die Produktion wird grob in zwei Bereiche unterteilt: **Frontend** und **Backend** (Abbildung 2.5). Im Frontend befindet sich die vollständige Prozessierung der Scheiben bis zum Wafer Test, bei dem die defekten Chips von den funktionierenden unterschieden werden. Im Backend werden die Chips zusammengefügt, eingebrannt, zu Modulen montiert und wiederum getestet.

Der erste Teil des Backend befindet sich noch innerhalb der so genannten Line und wird daher auch Backend of Line genannt. Hier werden die Chips im First Level Assembly vereinzelt, gebondet und gekapselt. Danach stehen sie zum Verkauf zur Verfügung oder können in einem Second Level Assembly auf Panels zu Modulen weiterverarbeitet werden.

### 2.3.2 Schichtherstellung und Produktvielfalt

Halbleiterchips werden nicht einzeln nacheinander gefertigt. Vielmehr werden auf einer Siliziumscheibe (**Wafer**) mehrere **Chips** (auch **Dies** genannt) gleichzeitig hergestellt. Je nach Größe des Wafers und Größe der Chips können so bis zu einige hundert Chips auf einem Wafer Platz finden. Innerhalb der Fab bewegen sich die Scheiben in Losen von 25 Wafers (in manchen Fabs auch von 50 Wafers), die sich während des Transportes in einer gemeinsamen Box befinden [Ath95]. Der Begriff **Los** wird in der Halbleiterfertigung immer als diese feste Transporteinheit von 25 Wafers verstanden. Sie ist nicht zu verwechseln mit dem Begriff Los aus der klassischen Produktionstechnik, womit eine im Umfang beliebige Bestellmenge gemeint ist. Werden für Prozesse, wie z.B. Ofenprozesse, mehrere Lose gemeinsam bearbeitet, so wird diese Einheit als **Batch** bezeichnet. Mit dem Begriff **Horden** können mehrere Lose wie auch mehrere Wafer beliebiger Menge bezeichnet werden.

Alle Halbleiterchips sind aus Schichten (Layer) aufgebaut, die auf eine Siliziumscheibe aufgebracht werden und unterschiedliche Strukturen erhalten. Die einzelnen Bereiche der Strukturen unterscheiden sich in ihrer Materialzusammensetzung und somit in ihren elektrischen Eigenschaften. Dadurch entstehen die Schaltungen auf den Chips. Um die Strukturen und Zusammensetzungen der einzelnen Schichten zu erzeugen, ist eine Vielzahl von Arbeitsschritten notwendig, die mit unterschiedlichen Maschinen erbracht werden. Die Herstellung einer Schicht erfolgt im Prinzip nach folgendem Ablauf (siehe Abbildung 2.6):

Zunächst wird eine Oxidschicht erzeugt, die nach ihrer Strukturierung als Maske dient, um das darunter liegende Material entsprechend dieser Struktur zu bearbeiten. Die Oxidschicht erhält ihre Maskenstruktur, indem ein Fotolack aufgebracht wird, welcher durch eine weitere externe Maske belichtet und anschließend entwickelt wird. Danach existieren nur noch die unbelichteten Bereiche des Lacks. Durch die Struktur des Fotolacks kann nun geätzt werden, wodurch sich die Struktur auf die Oxidschicht überträgt. Ein weiterer Prozess, z.B. eine Dotierung, verändert nun die Eigenschaften der Bereiche, die nicht durch die verbleibende Oxidschicht geschützt sind. Schließlich wird das restliche Oxid entfernt. Das Ergebnis sind Bereiche im Grundmaterial, die durch die Dotierung andere elektrische Eigenschaften haben.

Diese prinzipielle Prozessfolge wird in modernen CMOS-Bauelementen bis zu 30-mal wiederholt, wobei bis zu 1000 einzelne Prozessschritte durchlaufen werden. Für die unterschiedlichen Prozessschritte werden unterschiedliche Maschinen benötigt, wobei gleichartige Prozessschritte, die sich schichtweise wiederholen, von den gleichen Maschinen ausgeführt werden können.

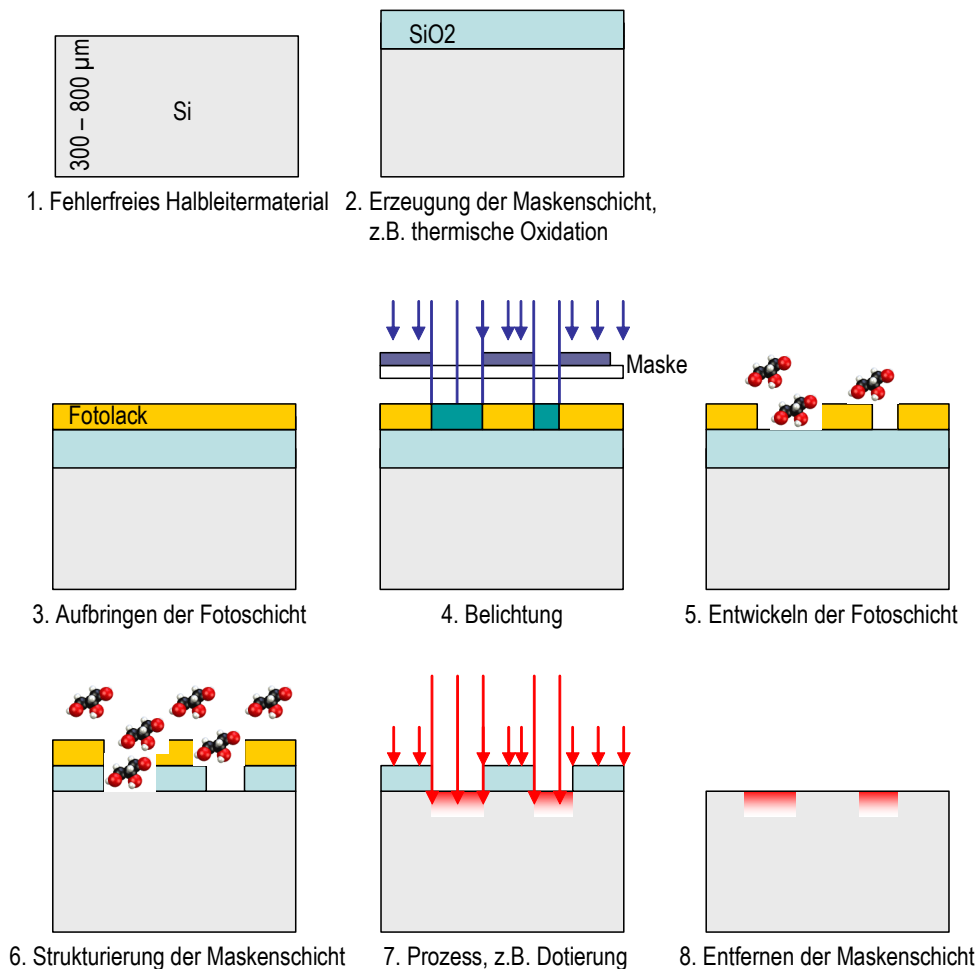


Abbildung 2.6: Herstellung einer Schicht einer Halbleiterscheibe [Han04]

Die Scheiben bzw. Lose, die auf diese Weise in der Produktionshalle unterwegs sind, sind aber nicht unbedingt gleicher Art. Die Lose können nach ihrer Zugehörigkeit zu den unterschiedlichen Technologien, die jeweils unterschiedliche technische Fortschritte darstellen, unterschieden werden. Die feinste Unterscheidung, nach denen Lose eingeteilt werden können, bezeichnet die einzelnen **Produkte**, die kundenspezifisch sind. Wenn in der Bearbeitungsreihenfolge einer Maschine unterschiedliche Produkte aufeinander folgen, so kann das einen Umrüstaufwand an der Maschine zur Folge haben, der Materialaufwand, Personeneinsatz, Maschinenstillstand und weitere Zuarbeiten mit sich bringt. Werden in einer Produktion Speicherchips (DRAM) hergestellt, so gibt es gleichzeitig ca. 5 verschiedene Produkte im Sortiment. Werden jedoch Logikbausteine produziert, so können es bis zu 400 oder auch mehr unterschiedliche Produkte sein. Somit ergibt sich für Produktionsstätten von Logikbausteinen von vorne herein ein größerer Umrüstaufwand als für Hersteller von Speicherbausteinen. Aufgrund dieser Tatsache ist die Automatisierung in Logikfabriken wesentlich komplexer und tatsächlich nicht so weit fortgeschritten wie in Fabs für Speicherbausteine.

Die einzelnen Produkte haben unterschiedliche Volumina und auch unterschiedliche Lebensdauer bezüglich der Zeitspanne in der sie hergestellt, d.h. dem Kunden angeboten werden. Neue Produkte können nicht sofort in großer Anzahl hergestellt werden, da zunächst die einzelnen Prozesse mit den vorhandenen Maschinen qualifiziert und so



eingestellt werden müssen, dass eine möglichst große Ausbeute erzielt werden kann. Erst allmählich kann das volle angestrebte Volumen produziert werden. Die Zeitphase bis dahin wird Ramp-Up genannt. Danach erfolgt eine Produktionsphase von einigen Monaten. Die Endphase der Produktion eines Produktes erfährt analog dazu ein Ramp-Down.

Jedes Produkt hat einen eigenen Arbeitsplan, der die Route bestimmt, nach der sich ein Los durch die Linie bewegt und die einzelnen Prozessschritte erfährt. Die Abfolge der Prozessschritte kann sich von Produkt zu Produkt ändern.

### 2.3.3 Farmlayout

Von Fertigungslinien, wie z.B. in der Automobilindustrie, ist die Fließbandherstellung bekannt. Eine derartige Fertigung ist in der Halbleiterbranche nicht möglich. Das Errichten einer Produktionsstätte kostet einige Milliarden Euro und deren Unterhalt täglich ca. eine Million Euro. Deshalb muss mit Platz und Ausrüstung sparsam umgegangen werden. Dies bedeutet wiederum, dass das Layout einer Fab sich nicht nach dem Materialfluss richten kann, sondern möglichst Platz sparend ausgerichtet werden muss. Eine Anordnung der Maschinen nach dem Materialfluss wäre auch aufgrund des bereits erwähnten Produktmixes und den daraus resultierenden unterschiedlichen und von Generation zu Generation wechselnden Arbeitsplänen nicht möglich, da unterschiedliche Produkte unterschiedliche Prozessschritte in unterschiedlichen Reihenfolgen benötigen.

Die Maschinen werden nach dem Werkstattprinzip ihrem jeweiligen Prozess entsprechend zu **Modulen** zusammengefasst und bilden das für die Halbleiterindustrie so typische **Farm-Layout** (auch **Job Shop** genannt). Der große Nachteil zum Work Flow Layout (Flow Shop), in dem die Maschinen entsprechend dem Materialfluss angeordnet sind, sind die stark verlängerten Wege der Produkte innerhalb der Fertigung und die damit verbundene Unübersichtlichkeit. Gleichartige Maschinen innerhalb eines Moduls werden als **Workcenter** bezeichnet. Teure Maschinen werden in möglichst geringer Anzahl angeschafft und stellen somit ein Bottleneck der Produktion dar.

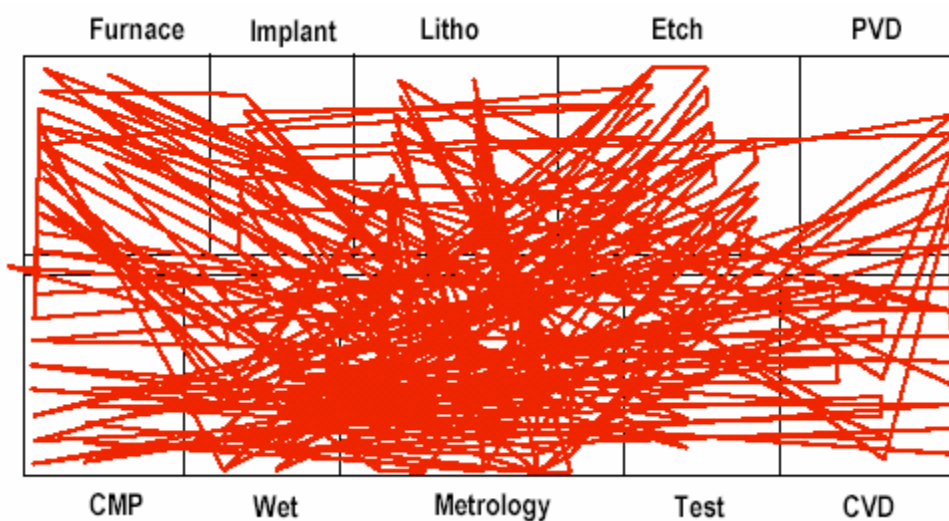


Abbildung 2.7: Weg einer einzelnen Scheibe (bzw. Los) während der Produktion in einer Halbleiterfab [Hei02]

Das bedeutet, dass eine Scheibe sich mehrmals auf ähnlichen Wegen durch die Fab und die einzelnen Module bewegt, wobei sie einige (mehr als 20) Kilometer hinter sich lässt. Abbildung 2.7 zeigt den Weg einer Scheibe, der z.B. folgende typische Reihenfolge hat: Einschleusung, Beschriftung, Ofentechnik, Fototechnik, Ätztechnik, (...), Fototechnik, (...), usw. Zur gleichen Zeit befinden sich einige tausend Scheiben, zusammengefasst in Losen, in einer Fab auf ihrer Route. Dabei kann es zu zufälligen Überlastungen bzw. Leerläufen von einzelnen Maschinen kommen. Die Kunst der Planung und Logistik liegt daher u.a. darin, den Ablauf bzw. den Fluss der Produktion möglichst gleichmäßig und konstant zu halten.

### 2.3.4 Der Faktor Mensch

Trotz des großen technischen Aufwandes ist in einer Halbleiterfab immer noch der **Operator**, der Mensch, der die Maschinen bedient, eine entscheidende Größe. Zu den Aufgaben eines Operators gehören Transport der Lose, Bestückung der Anlagen, Parametereingabe und Laden von Maschinenprogrammen sowie Beschaffung und Einsatz zusätzlicher Werkzeuge. Dabei hängt die Einflussmöglichkeit der Operatoren vom Automatisierungsgrad der Fab ab.

In den momentan existierenden Logik-Fertigungen sieht der Stand der Technik noch Laufzettel bzw. Protokolle in Papierform für jedes einzelne Los vor. Die Protokolle begleiten die Lose und beinhalten alle Prozessschritte sowie Informationen über Parameter-einstellungen an den Maschinen und eventuell zusätzlich benötigte Werkzeuge wie z.B. Belichtungsmasken. Je weiter die Automatisierung in einer Fab vorangeschritten ist, desto mehr Informationen sind in elektronischer Form bereits an den Maschinen vorhanden. Das Übertragen oder Eingeben von Daten entfällt und vermindert die Fehlermöglichkeiten. Werden Fabs neu errichtet, so geht der Trend zur papierlosen Fertigung (**Paperless Fab**). Allerdings ist es gerade für die Logikherstellung besonders schwierig, jedem Los von vorne herein eine Route inklusive der zu benutzenden Maschinen zuzuweisen. Je starrer ein System Arbeitsabläufe vorschreibt, desto schlechter kann auf unvorhergesehene Ereignisse (z.B. Maschinenausfälle) reagiert werden. Eine vollständig automatisierte Fertigung ohne Operators (Lights Out Fab, [Hei02]) wird wohl kaum zu realisieren sein.

Die Verteilung der Lose auf einzelne Maschinen kann eine große Auswirkung auf die Produktionsgeschwindigkeit einzelner Anlagengruppen sowie auf die Quantität und Qualität der hergestellten Chips haben. Jede Maschine hat andere Eigenschaften, sogar baugleiche Anlagen unterscheiden sich in ihrer Prozessgenauigkeit und den daraus resultierenden Schwankungen. Manche Lose dürfen deshalb nur auf bestimmten Maschinen prozessiert werden bzw. sie müssen von Schicht zu Schicht immer die gleichen Maschinen durchlaufen. Z.B. muss für zwei Lithografieschritte zur Strukturierung zweier aufeinander folgender Layer dieselbe Belichtungsanlage verwendet werden. Diese **Dedizierungen** ändern sich im Laufe der Zeit, wie sich auch die Maschineneigenschaften während des Betriebes ändern. Je nach Automatisierungsgrad der Fab und seines Steuerungssystems sind die notwendigen Daten einfach (z.B. elektronische Übertragung an die Maschine) oder mühsam (z.B. durch Nachschlagen in einer Liste) für den Operator zugänglich. Erfahrene Operatoren haben die notwendigen Daten auswendig parat und können schneller die Lose verteilen und die Prozesse starten. Somit hat die Erfahrung der Operators großen Einfluss auf die Performance der Produktion.

Meist bedient ein Operator mehrere Anlagen. Die Reihenfolge, in der er sie bedient, unterliegt dabei seiner eigenen Entscheidung. Durch geschickte Zeiteinteilung kann ein Operator Maschinenstandzeiten reduzieren.

### 2.3.5 Die Fototechnik

Da die Lithografie (Fototechnik) eine entscheidende Rolle für die Performance einer Halbleiterfab spielt und Ausgangspunkt des Stepper-Dispatchers in Kapitel 4 darstellt, soll sie hier näher erläutert werden.

In der Fototechnik werden die Belichtungen durchgeführt, mit deren Hilfe die benötigten Strukturen auf den Wafern erzeugt werden. Dazu wird zunächst ein Fotolack auf dem Wafer aufgebracht. In den so genannten **Steppern** wird mittels einer komplexen Linsenoptik und kurzwelligem Licht (zur Zeit 193 nm Wellenlänge, siehe auch [Zel02]) die benötigte Struktur von einer Maske (**Reticle**) auf den belackten Wafer übertragen. Bei der Belichtung wird dabei sukzessive vorgegangen. Da für große Wafer keine fehlerfreie Linsenabbildung möglich ist, werden kleinere Felder (ein oder mehrere Chips) schrittweise (Stepfield für Stepfield) belichtet. Dazu existieren sehr präzise Vorgaben bezüglich der Justierung der einzelnen Belichtungsebenen auf darunter liegende Schichten (Toleranz) und bezüglich der Auflösung zum Erzeugen der Minimalstrukturen (Strukturbreite). Zur Herstellung von 400 unterschiedlichen Produkten werden ca. 30.000 verschiedene Reticles benötigt, was wiederum eine logistische Herausforderung darstellt. Nach der Belichtung erfolgt die Entwicklung des Fotolacks, so dass nur noch (je nach Fotolacktyp) der belichtete bzw. der unbelichtete Bereich auf dem Wafer zurückbleibt. Der verbleibende Lack bildet eine Art Schutzschild für folgende Prozessschritte.

Die Lithografie ist an jeder Chip-Ebene beteiligt, d.h. pro Chip wird 5- bis 35-mal belichtet. 50% aller Einzelprozesse erfolgen so in der Fototechnik. Die Durchsatzleistung muss in diesem Bereich wesentlich höher sein als in anderen Bereichen. Daher hat eine Optimierung in der Lithografie besonders großes Potential zur Effektivitätssteigerung einer Halbleiterproduktion.

Darüber hinaus stellen Stepper aufgrund ihrer hohen Investitionskosten das größte Bottleneck einer Linie dar. Eine schlechte Performance in der Fototechnik hat sofort eine bedeutende negative Auswirkung auf das Produktionsergebnis der gesamten Linie.

Die Lithografie der Fototechnik ist durch ihre Komplexität besonders schwer zu synchronisieren (siehe zu Synchronisation Kapitel 2.5.4). Die Eigenschaften der einzelnen Prozessschritte variieren von Produkt zu Produkt viel stärker als in anderen Bereichen der Produktion. So kann nicht jedes Produkt auf jeder Belichtungsanlage bearbeitet werden. Es gibt diverse produktspezifische Freigaben und Sperrungen, die von Zeit zu Zeit immer wieder dem Stand der Technik und den sich verändernden Eigenschaften der Anlagen angepasst werden müssen.

### 2.3.6 Produktionsfluss und Liniensteuerung

Im Frontend der Halbleiterproduktion ergibt sich ein Szenario, bei dem Lose verschiedener Produkte in unterschiedlichen Fertigungsstufen auf unterschiedlichen Routen um die gleichen Ressourcen (Maschinen) konkurrieren. Die Bearbeitungsreihenfolge der Lose und die Wahl der bearbeitenden Maschinen hängen von vielen Randbedingungen und unvorhersehbaren Ereignissen ab und können somit für ein Los bei dessen Einschleusung nicht vorbestimmt werden.

Dass bei einem derartigen System nicht immer ein gleichmäßiger Produktionsfluss durch die Fertigungshalle erreicht werden kann, ist einsichtig. Es wird immer Situationen geben, in denen manche Maschinen nicht produzieren, weil gerade kein Material (Lose) vorhanden ist,

während vor anderen Maschinen die Pufferlager geradezu überquellen. Um eine möglichst effiziente und somit Kosten minimierende (= Gewinn maximierende) Produktion zu erreichen, müssen Maßnahmen getroffen werden, die den Produktionsfluss möglichst optimal innerhalb der Fab steuern. Allerdings stehen diese Zielvorgaben in Konflikt mit anderen. So erfordern kurze **Durchlaufzeiten** und hohe **Liefertreue** einen geringen **Warenbestand** innerhalb der Produktion und stehen damit im direkten Widerspruch zu hohen Maschinenauslastungen, großem Durchsatz und attraktiven Verkaufspreisen, die nur mit hohem Warenbestand in der Produktion erreichbar sind. Die Produktions- bzw. Firmenleitung muss bei ihren Entscheidungen immer Präferenzen setzen und zwischen unterschiedlichen, konkurrierenden Zielen Kompromisse schließen.

Die Aufgabe der **Liniensteuerung** ist das möglichst effiziente Managen von Auslastung, Durchlaufzeiten, Liefertreue und Warenbestand in der Fertigung, ohne die Betriebsmittel zu ändern. Ist dies nicht mehr möglich, muss über Neuinvestitionen entschieden werden. Eine ideale Steuerung wäre dann erreicht, wenn vorhergesagt bzw. vorherbestimmt werden kann, wann welche Einheit mit welcher Anlage prozessiert wird. Das ist aber aus den oben genannten Gründen (noch?) nicht möglich. Dennoch ist die Liniensteuerung mit den entsprechenden Regulierungen und Methoden entscheidend für die Performance einer Halbleiterfabrik.

Es können mehrere Steuerungsbereiche unterschieden werden. In erster Linie werden Vorgaben über die zu erbringenden Arbeitsleistungen pro Zeiteinheit (pro Tag) festgelegt. Die Vorgaben beziehen sich auf die Leistung der gesamten Linie wie auch auf einzelne Workcenter oder Maschinen. Zur Steuerung einer Linie gehören ebenso die Bereiche: **Sequencing**, **Batching**, **Dispatching** und **Resource Dedication**. Unter Sequencing werden alle Regeln verstanden, die die Reihenfolge der zu bearbeitenden Lose festlegen. Batching behandelt die möglichst optimale Zusammenfassung mehrerer Lose für gemeinsame Prozesse. Dispatching legt die Verteilung der Lose auf die einzelnen Maschinen fest und Resource Dedication sind ständige oder zeitweise Vorgaben, nach denen bestimmte Lose bzw. Produkte auf einigen Maschinen nicht bzw. ausschließlich produziert werden dürfen. Darüber hinaus ist es notwendig, die Arbeitsschritte an sich zu vereinfachen bzw. informationstechnisch zu unterstützen.

Eine effektive Liniensteuerung kann nur unter Berücksichtigung der Ergebnisse der **Linienbewertung** stattfinden. Die Bewertung einer Fabrik erfolgt meist mittels Kennzahlen, sog. Performance-Indikatoren. Werden die Performance-Indikatoren immer auf gleiche Weise erhoben, so können mit ihnen einzelne Fabs verglichen werden (Benchmarking). Die Kennzahlen, die die tatsächliche Leistung einer Fertigung beschreiben, können auch von der Liniensteuerung als Soll-Werte vorgegeben werden. Werden sie nicht erreicht, sind geeignete Maßnahmen zu erarbeiten, die zum Erreichen der Zielwerte führen.

Eine Vielzahl von Kennzahlen ist bereits definiert worden und kann zum größten Teil von den elektronischen Datenerfassungen der Fabs automatisch berechnet werden.

## 2.4 Supply Chain Management

Unter Supply Chain wird im allgemeinen die Lieferkette, Versorgungskette oder Wertschöpfungskette eines Produktes entlang der beteiligten Unternehmen, wie Lieferanten, Hersteller, Warenlager, Händler und Verkäufer verstanden. Bei mehrstufigen Produktionen gibt es auch eine interne Supply Chain innerhalb eines Unternehmens.

In [Sta02] wird Supply Chain Management als die Aufgabe definiert, die Organisationseinheiten entlang einer Wertschöpfungskette zu integrieren und Material, Informationen und Finanzflüsse zu koordinieren, um eine Kundenanforderung zu erfüllen und das Ziel zu verfolgen, die Wettbewerbsfähigkeit der Lieferkette als Ganzes kontinuierlich zu verbessern.

Wettbewerbsfähigkeit kann auf viele verschiedene Arten verbessert werden: durch Kostenreduktion, größere Flexibilität bezüglich Kundenanforderungen, Verbesserung der angebotenen Qualität, Erhöhung des Services oder Verbesserung der Liefertreue. Dies alles kann zusammengefasst werden mit dem Begriff „Verbesserung des Kundenservices“. Um den Kundenservice zu messen, werden Kennzahlen, so genannte Key-Performance-Indicators (KPI), verwendet. KPIs beschreiben und strukturieren komplexe Systeme, wobei die Informationsvielfalt oft vereinfacht wird. Sie erfüllen dabei drei Funktionen:

- Information. Der hauptsächliche Zweck von Key-Performance-Indicators ist, das Management zu informieren. Sie sind Grundlage für Entscheidungen und Hilfsmittel zur Identifizierung von Problemen. Die tatsächlichen Ist-Kennwerte können mit Standard-Werten oder Zielvorgaben verglichen werden.
- Steuerung. Performance-Indicators sind die Basis für Zielsetzungen (Target Setting). Die einzelnen Zielvorgaben sollen die Produktion so steuern, dass die gewünschten Ergebnisse erreicht werden.
- Kontrolle. Performance-Kennwerte sind gut geeignet, um Operationen und Prozesse zu überwachen.

Die Kennzahlen unterliegen dabei einer Hierarchie, entsprechend der Hierarchie des Managements, das sie zu bewerten hat. Nach [Tho03] ist der Service Level eine der KPIs in der obersten Hierarchieebene, adressiert an die Geschäftsführung. Er beschreibt die Liefertreue gegenüber dem Endkunden. Der Service Level ist der Prozentsatz aller Lieferungen einer bestimmten Lieferperiode, die mit richtiger Menge und rechtzeitig beim Kunden ausgeliefert wurden.

Die Schwierigkeiten des Supply Chain Managements bestehen zum einen in der Koordination und Optimierung verschiedenartiger Organisationseinheiten, die an unterschiedlichen Orten angesiedelt sind (globale Optimierung). Zum anderen müssen Entscheidungen trotz Unsicherheit getroffen werden. Das betrifft Unsicherheiten über die Nachfrage, Lieferzeiten, Ausbeute, Transportzeiten und Komponenten-Verfügbarkeiten [Sim03].

Die Themen des Supply Chain Managements lassen sich in drei Kategorien einteilen:

- Strategische Ebene. Hier werden Entscheidungen getroffen, die eine Firma auf lange Sicht betreffen. Das beinhaltet die Planung von Lager-, Verteilungs- oder Produktionsstätten sowie des Materialflusses durch die Wertschöpfungskette.
- Taktische Ebene. In ihr werden Entscheidungen alle viertel bis ganze Jahre getroffen. Dabei handelt es sich um Fragen des Einkaufs, der Produktionsplanung, der Lagerhaltung, der Transportstrategien, etc.
- Operative Ebene. Sie bezieht sich auf Entscheidungen wie Ablaufplanungen, Lieferzeitkalkulationen, Routenplanungen, etc.

Die Supply Chain der Halbleiterindustrie (Abbildung 2.8) wird durch einige besondere Charakteristiken beeinflusst. Zum einen ist das Verhältnis von Wert und Volumen der Produkte sehr groß, so dass Transporte über große Distanzen auch mittels Flugzeug rentabel sind. Die benötigte Ausrüstung dagegen ist teuer und schlecht bewegbar. Der

Produktionsprozess kann somit über mehrere Orte aufgeteilt werden und die jeweiligen Standortvorteile, wie Kosten für Grundbesitz, Maschinen und Gehälter, ausgenutzt werden. Tatsächlich sind bei den meisten Herstellern das Frontend in Europa oder USA angesiedelt, während sich das Backend in Asien befindet.

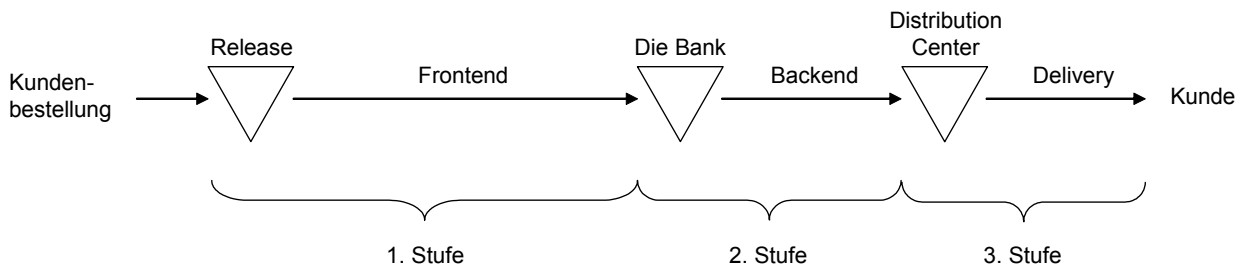


Abbildung 2.8: Die interne Supply Chain der Halbleiterindustrie kann in drei Stufen eingeteilt werden

Die Kunden der Halbleiterhersteller sind wiederum meist große Industriekonzerne, so dass sie Einfluss auf den Herstellungsprozess der Halbleiter nehmen können. Das äußert sich z.B. in Forderungen nach speziellen Herstellungsorten oder Benutzung bestimmter Maschinen.

Die Nachfrage nach Halbleiterchips ist nicht nur stark schwankend, sie unterliegt auch einer hohen Unsicherheit in ihrer Vorhersagegenauigkeit. Das erfordert die Fähigkeit, rasch auf große Veränderungen der Produktionsmengen reagieren zu können.

## 2.5 Performance-Kennzahlen der Anlagen

### 2.5.1 Standards

Die Einführung der SEMI-Standards löste die firmeneigenen Definitionen ab. Zuvor waren sämtliche Begriffe uneinheitlich besetzt, was mitunter dazu führte, dass Formulierungen existierten, die zu Betriebszeiten von über 100% gelangten [Vad03]. Anbieter und Kunden hatten Probleme, mit ihren unterschiedlichen Formeln Verträge über Akzeptanzkriterien, Garantiefälle, technischen Support, etc. zu schließen. Außerdem war es unmöglich, die Performance verschiedener Hersteller zu vergleichen.

Um die technologischen Herausforderungen besser verwirklichen und kostengünstiger produzieren zu können, haben sich Anbieter und Kunden des Zuliefermarktes der Halbleiterindustrie 1970 im Wirtschaftsverband SEMI zusammen geschlossen. SEMI steht für Semiconductor Equipment and Materials International und repräsentiert ca. 2.500 Gerätehersteller und Materiallieferanten der Halbleiterindustrie. SEMI erarbeitet u.a. Standards für die Halbleiterproduktion. Bereits 1986 wurde eine Richtlinie zur Zeitmessung der Maschinenzustände veröffentlicht, die seit 1996 als SEMI E10-Standard bekannt ist. In den seitdem vergangenen 10 Jahren konnte die Industrie erste Maschinen mit der entsprechenden Software und den benötigten Schnittstellen ausrüsten. Somit ist es jetzt möglich, in den neueren Fabs mit diesen Maschinenzuständen zu arbeiten. Im Februar 1999,

also erst kurz vor Beginn dieser Arbeit, wurde der Standard SEMI E79 zur Messung der Anlagenproduktivität veröffentlicht.

Weitere Standards beschreiben u.a. Performanceverluste, Zustandsübergänge und Maschinenkommunikation. In einer vorläufigen Ausgabe wird bereits der erste Versuch unternommen, eine Kennzahl zur Overall Fab Efficiency zu berechnen.

Tabelle 2.3 gibt eine Übersicht zu den SEMI-Standards zur Optimierung der Anlagenperformance und Produktivität von Halbleiterfabs.

<b>Standard</b>	<b>Titel und Inhalt</b>
SEMI E10	<i>Standard for Definition and Measurement of Equipment Reliability, Availability and Maintainability (RAM)</i> Der Standard führt eine allgemeine Basis zur Kommunikation zwischen Anwendern und Anbietern von Halbleiterproduktions-Equipment ein, indem Standards zur Messung von Maschinenzeiten und somit Zuverlässigkeit, Verfügbarkeit und Wartbarkeit gegeben werden.
SEMI E58	<i>Automated Reliability, Availability and Maintainability Standard (ARAMS)</i> ARAMS bietet einen Standard zur Implementierung und Sammlung von Zustandsänderungen nach SEMI E10 auf der Anlagenebene in einem automatisierten Umfeld.
SEMI E79	<i>Standard for Definition and Measurement of Equipment Productivity (OEE)</i> Der Standard definiert Berechnungen und Analysen für Equipment-Produktivitäts-Kennzahlen.
SEMI E116	<i>Provisional Specification for Equipment Performance Tracking (EPT)</i> Der (vorläufige) Standard bietet Spezifikationen, um eine grundlegende Equipment-Performance-Verfolgung zu implementieren.
SEMI E124	<i>Provisional Guideline for Definition and Calculation of Overall Factory Efficiency (OFE) and Other Associated Factory-Level Productivity Metrics</i> Der (vorläufige) Standard schlägt eine Maßzahl zur Berechnung der gesamten Fab-Performance vor.

Tabelle 2.3: SEMI-Standards zur Optimierung der Anlagenperformance und Produktivität von Halbleiterfabs nach [Piv02] und [Cim03]

SEMI E10, SEMI E79 und SEMI E124 werden im folgenden näher beschrieben. Da SEMI E58 die Codes von automatisch generierten Nachrichten beschreibt, ist es für die Untersuchung von Performanceverbesserungen nur als Werkzeug zur Datenerzeugung notwendig. SEMI E116 beschreibt die Generierung von Maschinenzuständen, unabhängig von SEMI E10 und ohne Bedienung eines Operators. SEMI E124 bezieht sich auf die

gesamte Fab und wird daher in Kapitel 2.6, Performance-Kennzahlen der Produktion, näher beschrieben. Standards bieten für Zulieferer und Hersteller definierte Schnittstellen zur Kommunikation. Mit ihnen können vergleichbare Werte mit der Möglichkeit zur Klassifizierung und Bewertung ermittelt werden. Sie liefern dadurch auch einen Beitrag zur Qualitätssicherung: Garantien können somit auf bestimmte, definierte Eigenschaften gegeben werden.

### 2.5.2 Maschinenzustände

Um Performance-Verluste ermitteln und Parameter wie Zuverlässigkeit, Auslastung und Wartbarkeit berechnen zu können, müssen zunächst unterschiedliche Maschinenzustände definiert und gemessen werden. Der SEMI-Standard SEMI E10 spezifiziert sechs Maschinenzustände [SEMI]. Abbildung 2.9 stellt den Zusammenhang der Maschinenzustände dar. Während der geplanten Nichtbetriebszeit (Non-Scheduled Time) findet auf Grund von Ferien oder Produktionsstillstand während Fabrikumbauten keine Produktion statt. In der Halbleiterproduktion existiert dieser Zustand praktisch nicht. Ungeplante Ausfallzeiten (Unscheduled Downtime) entstehen durch Maschinenausfälle, unerwartete oder unerwartet lange Wartungsarbeiten, etc. Zu den geplanten Ausfallzeiten (Scheduled Downtime) zählen planmäßige Wartungsarbeiten oder Umbaumaßnahmen. Während der Engeneering-Zeit (Engeneering Time) werden Maschinentests durchgeführt, um die Maschine zu qualifizieren oder neue Prozesse einzufahren. Während Wartezeiten (Standby Time) könnte Produktion stattfinden, die Maschine produziert jedoch nicht wegen fehlendem Material, nicht verfügbarem Operator etc. Die Produktionszeit (Productive Time) ist die Zeit, in der für den Verkauf bestimmte Produkte prozessiert werden.

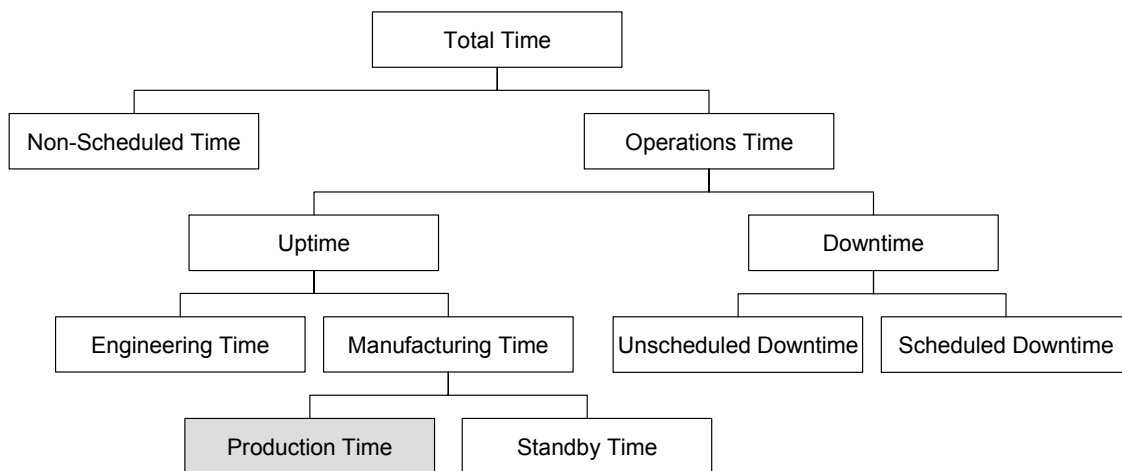


Abbildung 2.9: Maschinenzustände nach SEMI E10

Mit den Zeitmessungen der Maschinenzustände können die grundlegenden Performance-Parameter der einzelnen Maschinen ermittelt werden [Tra04].

Die durchschnittliche Zeit, in der eine Anlage produziert, wird Mean Time Between Failures *MTBF* genannt und aus der produktiven Zeit und der Anzahl der aufgetretenen Fehler berechnet.



$$MTBF = \frac{\text{productive time}}{\text{number of failures}} \quad (2.1)$$

Wird die produktive Zeit ins Verhältnis zur Anzahl der Eingriffe eines Operators gesetzt, so ergibt sich die mittlere Zeit zwischen zwei Eingriffen, Mean Time Between Assists *MTBA*.

$$MTBA = \frac{\text{productive time}}{\text{number of assist}} \quad (2.2)$$

Ausfallzeiten können für die Produktion schwer wiegende Folgen haben. Das Verhältnis der Ausfallzeiten zur Anzahl der Ausfälle wird als Mean Time Off Line *MTOL* ermittelt.

$$MTOL = \frac{\text{downtime}}{\text{number of interrupts}} \quad (2.3)$$

Um die Wartungsfähigkeit einer Maschine zu bestimmen, muss zunächst auch noch die Zeit während den Reparaturen gemessen werden. Die Mean Time To Restore *MTTR* ist dann das Verhältnis der Reparaturzeiten zu der Anzahl der Fehler.

$$MTTR = \frac{\text{repair time}}{\text{number of failures}} \quad (2.4)$$

Die Beschreibung der Maschinenzustände sowie die Ermittlung der Parameter für Zuverlässigkeit, Auslastung und Wartbarkeit sind reine Visualisierungswerkzeuge. Eine Performance-Verbesserung lässt sich mit ihnen nicht erreichen. Sie können lediglich den Erfolg einer Verbesserungsmaßnahme messen, sofern er Auswirkungen auf die Maschinenzustände hat.

### 2.5.3 Overall Equipment Efficiency OEE

Der SEMI E79-Standard „Standard for Definition and Measurement of Equipment Productivity“ wurde 1999 veröffentlicht [Piv02]. Er beinhaltet Maßzahlen und Formeln, mit denen die Produktivität der Anlagen berechnet wird. Die Overall Equipment Efficiency *OEE* wird dabei vollständig durch zeitbezogene Ausdrücke beschrieben.

Die vorhergehende Version des Standards beschrieb die Overall Equipment Effectiveness, die zusätzlich zu den zeitbezogenen Maßzahlen auch die Anzahl produzierter Einheiten berücksichtigt. Diese Berechnungsvorschrift ist jedoch für Cluster-Tools schlecht verwendbar. Cluster-Tools sind Maschinen, in denen mehrere Prozessschritte integriert sind. In [Kie03] wird Effektivität mit Wirksamkeit (d.h. das Richtige machen) und Effizienz mit Wirtschaftlichkeit (d.h. etwas richtig machen) beschrieben.

Die Overall Equipment Efficiency *OEE* wird als Produkt aus Verfügbarkeit (Availability), Betriebseffizienz (Operational Efficiency), Durchsatzeffizienz (Rate Efficiency) und Qualitätseffizienz (Quality Efficiency) definiert.

$$OEE = \text{availability} \cdot \text{rate efficiency} \cdot \text{operational efficiency} \cdot \text{quality efficiency} \quad (2.5)$$

Die Verfügbarkeit (Availability) berechnet sich aus dem Verhältnis von Uptime (Zeit, in der die Maschine verfügbar ist) zu Gesamtzeit (Total Time).

$$\text{availability} = \frac{\text{uptime}}{\text{total time}} \quad (2.6)$$

Für die Betriebseffizienz (Operational Efficiency) wird die Production Time (die Zeit, in der die Maschine für den Verkauf arbeitet) zur Uptime in Beziehung gesetzt.

$$\text{operational efficiency} = \frac{\text{production time}}{\text{uptime}} \quad (2.7)$$

Das Verhältnis von theoretischer Produktionszeit einer zu produzierenden Einheit (Theoretical Production Time) zur tatsächlich benötigten Produktionszeit der zu produzierenden Einheit (Actual Production Time), ergibt die Durchsatzeffizienz (Rate Efficiency).

$$\text{rate efficiency} = \frac{\text{theoretical production time}}{\text{actual production time}} \quad (2.8)$$

Das Produkt aus Betriebseffizienz und Durchsatzeffizienz wird auch als Performanceeffizienz (Performance Efficiency) bezeichnet.

Wird die Zeit, die theoretisch für brauchbare Einheiten benötigt wird, durch die Zeit, die theoretisch für die aktuell produzierten Einheiten verwendet wird, geteilt, so ergibt sich die Qualitätseffizienz (Quality Efficiency).

$$\text{quality efficiency} = \frac{\text{theoretical time for effective units}}{\text{theoretical production time for actual units}} \quad (2.9)$$

Die vorhergehende Version zur Berechnung der Overall Equipment Effectiveness sah die Berechnung einer Quality Rate als Quotient aus Gesamtzahl produzierter Wafer minus zurückgewiesenen Wafern und Gesamtzahl produzierter Wafer vor.

Die Maßzahlen zur Ermittlung der OEE nach SEMI E79 benutzen die Definitionen des Standards SEMI E10 zur Erfassung der Maschinenzeiten. Abbildung 2.10 zeigt die Maschinenzeiten nach SEMI E10 und ihre resultierenden Verluste [Tra04].

Die Berechnung der Overall Equipment Efficiency OEE beruht somit darauf, Verluste zu messen. Ursachen, die zu den einzelnen Verlusten führen können, sind in Abbildung 2.11 dargestellt.

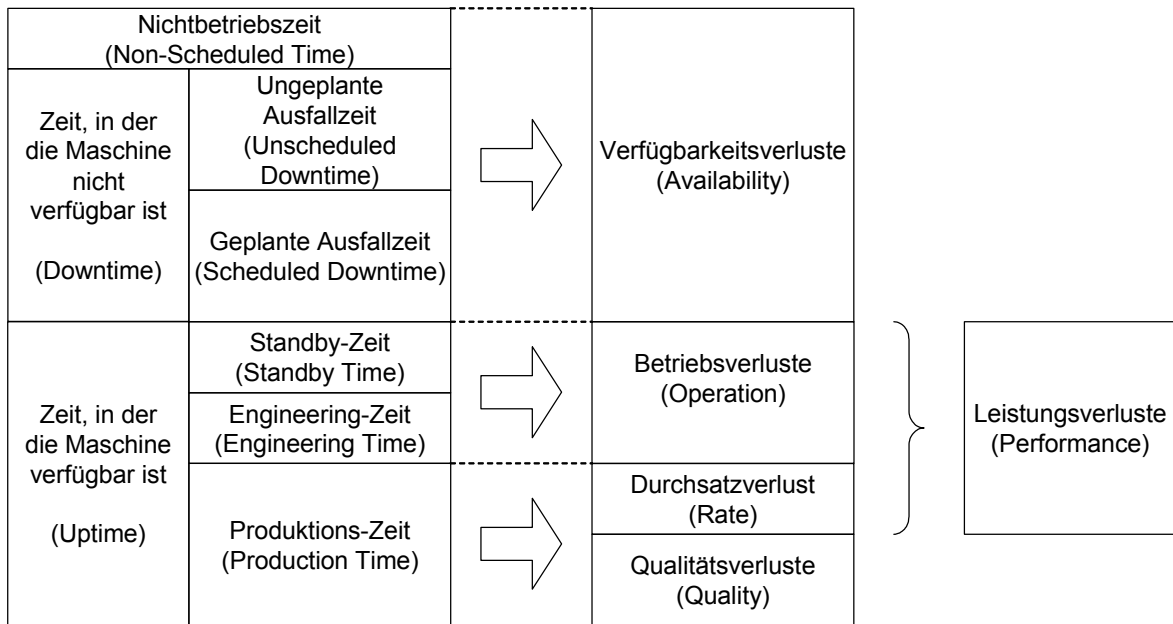


Abbildung 2.10: Verluste der Maschinenzeiten

Die Verfügbarkeit wird durch ungeplante und geplante Ausfallzeiten (Unscheduled Down Time, Scheduled Down Time) reduziert. Dies geschieht durch Reparaturen ungeplanter Ausfälle und geplante Wartungen. Standby-Zeiten und Engineering-Zeiten führen zu Betriebsverlusten und verringern die Betriebseffizienz. Diese Verluste entstehen durch fehlendes Material, fehlende Operatoren, Maschinentests, etc. Die Durchsatzeffizienz sinkt durch langsame Produktion, die u.a. durch Equipmentfehler hervorgerufen werden kann. Nacharbeit und Ausschuss reduzieren die Qualitätseffizienz.

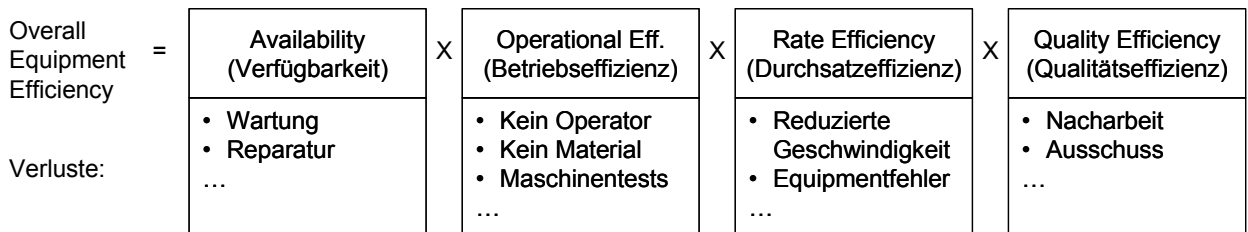


Abbildung 2.11: Verluste der Overall Equipment Efficiency

Die Untersuchung der einzelnen Zeitverluste können unter Umständen direkte Hinweise auf die Ursachen geben, womit konkrete Verbesserungsmaßnahmen eingeleitet werden können. Eine Automatisierung der Auswertung könnte für die einzelnen Parameter geeignete Eingriffsmethoden ermitteln. Eine vollständige Automatisierung scheiterte bisher jedoch immer noch daran, dass in den bestehenden Geräten keine Software und keine Schnittstellen vorhanden sind, die die SEMI E10-Zeiten erfassen und zu Datenbanken weiterleiten können.

### 2.5.4 Das Vier-Partner-Modell

Jedem Einzelprozess innerhalb der unterschiedlichen Module kann ein Produktionssystem zugeordnet werden. Unter einem Produktionssystem wird die Gesamtheit der Produktionspartner Material, Mensch, Maschine und Methode verstanden, auch „die vier Ms“ genannt (Abbildung 2.12). Material und Maschine bilden dabei die Produktionsanlage, zusammen mit dem Faktor Mensch die Produktionseinheit. Material sind alle Lose (Work In Progress, WIP), die auf ihre Bearbeitung vor dem Produktionssystem warten oder sich bereits in Bearbeitung befinden. Zum Begriff Maschine gehören die Maschinen sowie die für den jeweiligen Prozess benötigten Rohstoffe. Die Menschen, die das Material bewegen und die Maschinen bedienen, werden als Operatoren bezeichnet. Unter Methode ist die Gesamtheit der technischen Daten und Maschineneinstellungen wie auch der Programmcode zu verstehen.

Ein Produktionssystem kann redundant ausgeführt sein, d.h. es können mehr als eine Maschine und ein Mensch existieren. Dabei ergibt sich bereits das erste Logistikproblem. Welcher Operator bearbeitet wann und in welcher Reihenfolge an welcher Maschine welches Material.

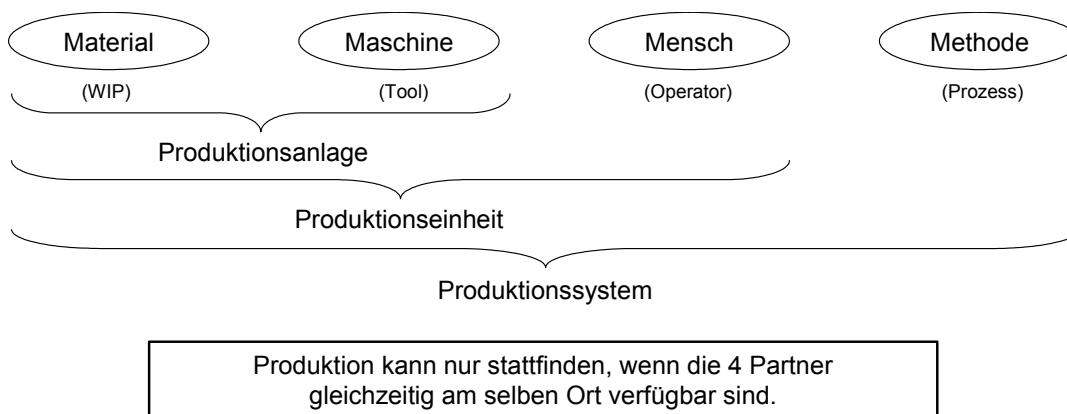


Abbildung 2.12: Ein Produktionssystem besteht aus den vier Ms: Material, Maschine, Mensch und Methode

An einem Produktionssystem kann nur dann eine Produktion stattfinden, wenn alle vier Partner gleichzeitig am selben Ort verfügbar sind [Bau01]. Fehlt nur ein Partner (z.B. Mittagspause des Operators, Reparatur einer Maschine, Fehlen der Prozessparameter oder Abbruch des Materialflusses) steht das gesamte Produktionssystem still. D.h. das Produktionssystem ist nur dann verfügbar, wenn alle Produktionsfaktoren verfügbar sind. Die Verfügbarkeiten der Produktionsfaktoren sind die Wahrscheinlichkeiten der Partner, zu einem beliebigen Zeitpunkt verfügbar zu sein. Im Allgemeinen handelt es sich hier um statistisch unabhängige Größen. Damit ergibt sich die Verfügbarkeit des Gesamtsystems als Produkt der Einzelverfügbarkeiten (Gleichung 2.10). Das bedeutet, dass jede Verfügbarkeit, die nicht 100% beträgt, die Gesamtverfügbarkeit verringert. Oder anders: Eine 5%ige Steigerung der Gesamtverfügbarkeit lässt sich durch 5% mehr Verfügbarkeit eines einzelnen Faktors erreichen. Die vier Partner sind gleichwertige Einsatzfaktoren. Gelingt es, das System zu **synchronisieren**, d.h. die Produktionspartner sind immer zur gleichen Zeit nicht verfügbar, so beschränkt nur der Partner mit der geringsten Verfügbarkeit die mögliche Produktionszeit. Die Gesamtverfügbarkeit ist dann gleich der geringsten Einzelverfügbarkeit

(Gleichung 2.11). Die Partner sind nun nicht mehr gleichberechtigt. Nur die Anhebung der minimalen Verfügbarkeit führt zu einer Erhöhung der Gesamtverfügbarkeit.

Unabhängigkeit der Produktionsfaktoren:

$$A_{\text{Produktionssystem}} = A_{\text{Material}} \cdot A_{\text{Maschine}} \cdot A_{\text{Mensch}} \cdot A_{\text{Methode}} \quad (2.10)$$

Synchronisation der Produktionsfaktoren:

$$A_{\text{Produktionssystem}} = \text{Min}\{A_{\text{Material}} \cdot A_{\text{Maschine}} \cdot A_{\text{Mensch}} \cdot A_{\text{Methode}}\} \quad (2.11)$$

Das Beispiel aus Abbildung 2.13 kann dies eindrucksvoll erläutern. Nur zu den Zeiten, bei denen alle übereinander liegenden Zeilen schraffiert sind, kann produziert werden. Zunächst wird die produktive Zeit eines Tages für unabhängige Verfügbarkeiten berechnet. Trotz der hohen Einzelverfügbarkeit ist die Gesamtverfügbarkeit deutlich niedriger. Durch Synchronisation wird bei gleicher Einzelverfügbarkeit die Gesamtverfügbarkeit deutlich höher. Die produktive Zeit kann von 12,5 Stunden auf 19,2 Stunden erhöht werden, was eine Steigerung von mehr als 53% bedeutet.

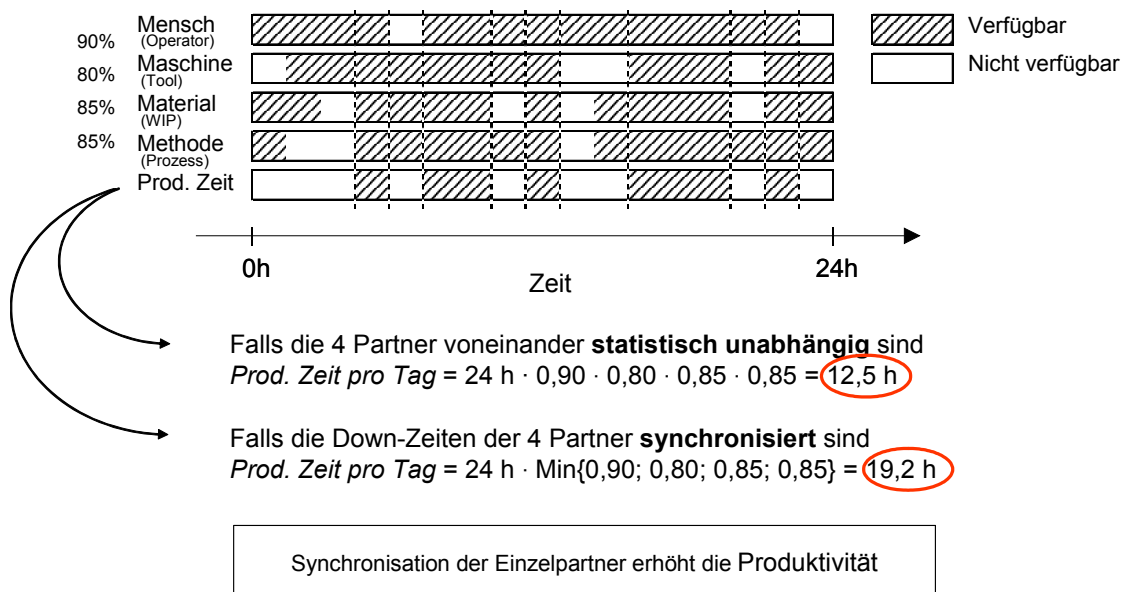


Abbildung 2.13: Erhöhung der produktiven Zeit durch Synchronisation der Produktionsfaktoren

Die Schwierigkeit besteht nun darin, die zunächst statistisch unabhängigen Ausfälle der Produktionsfaktoren zu synchronisieren. Bei regelmäßig vorgesehenen Pausen und Wartungen ist das entsprechend planbar. So ist es sinnvoll, Wartungsarbeiten an Maschinen während der Pausenzeiten des Operators durchzuführen. Unvorhergesehene Ausfälle sind jedoch nur bedingt synchronisierbar.

Wie die Warteschlangentheorie zeigt (siehe Kapitel 2.6.8) gibt es einen weiteren Aspekt, der die Performance einer Produktionslinie entscheidend beeinflusst: die Variabilität. Die vier Partner beeinflussen nicht nur durch ihre Verfügbarkeit, sondern auch durch die Variation ihrer Verfügbarkeit und durch die Variation der zeitlichen Übereinstimmung ihrer Verfügbarkeit die dynamische Gesamtleistung des Produktionssystems.

Somit sind nicht nur die statischen Betrachtungen der Verfügbarkeiten wichtig, sondern auch das dynamische Verhalten. Es kann gezeigt werden, dass häufige kurze Ausfälle für das Gesamtsystem weniger schädlich sind als seltene lange Ausfälle [Sch99]. Um Durchsatzausfälle zu vermeiden, muss für die Zeit des Ausfalls Material (Zwischenprodukte) bevorratet sein. Wenn für lange Ausfälle nicht genügend Material für die Folgestationen zwischengelagert ist, führt dies zu Produktionsstillstand. Kürzere Ausfälle können mit weniger Lagerbeständen aufgefangen werden.

Das in einem Produktionssystem zu erwartende Optimierungspotential hängt vom bereits realisierten Synchronisationsgrad ab: Je geringer der Synchronisationsgrad, desto höher das Potential zur Produktivitätssteigerung. In Abbildung 2.14 sind Maßnahmen zur Verbesserung der Verfügbarkeiten und des Synchronisationsgrades der vier Partner zusammengestellt. Für eine kontinuierliche Verbesserung des Gesamtsystems ist es notwendig, immer wieder die einzelnen Optimierungsstrategien zu überprüfen und mit geeigneten Maßnahmen die vier Partner zu beeinflussen.

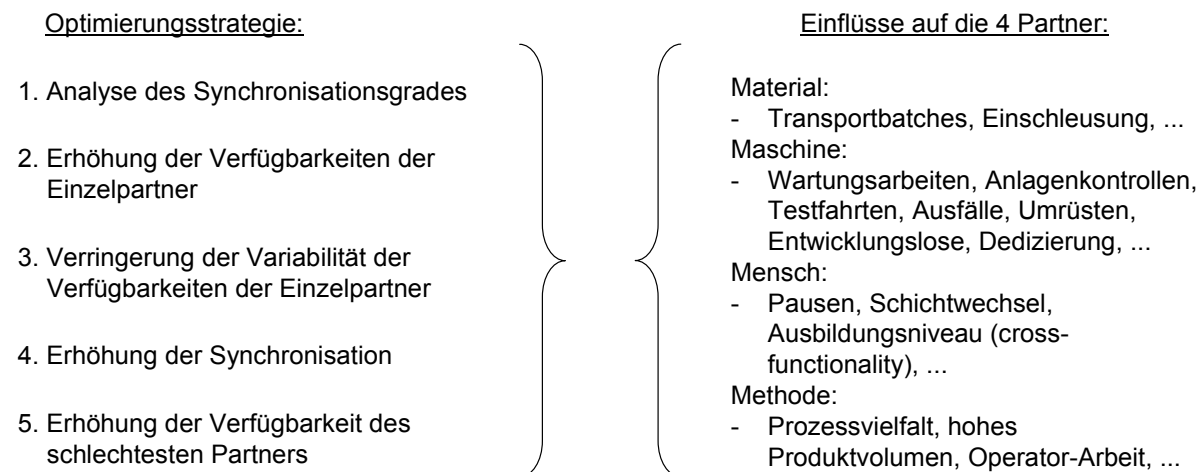


Abbildung 2.14: Maßnahmen zur Verbesserung der Verfügbarkeiten und des Synchronisationsgrades

## 2.6 Performance-Kennzahlen der Produktion

### 2.6.1 Flussfaktor und Little's Law

Die Grundlagen der Liniensteuerung wie auch der Linienbewertung stellen Flussfaktor und Little's Law dar. Der Flussfaktor beschreibt das Verhältnis von Durchlaufzeit (Cycle Time, *CT*) und der Rohprozesszeit (Raw Process Time, *RPT*). Die Rohprozesszeit ist die minimal notwendige Produktionszeit einer Halbleiterscheibe ohne Warte- und Maschinendownzeiten. Mit dem Flussfaktor kann die Schnelligkeit einer Produktionslinie beschrieben werden (tatsächlicher Flussfaktor). Er wird ebenso als Vorgabewert zur Steuerung der Linie verwendet.

In Little's Law ist der Zusammenhang von Durchlaufzeit *CT*, Warenbestand (Work in Progress, *WIP*) und Durchsatz (Throughput, hier: Going Rate, *GR*) beschrieben. Dabei geht Little's Law aus der Umstellung der Definition des Durchsatzes (Durchsatz = Material pro Zeit) hervor; die Bezeichnung als Gesetz (Law) ist demnach nicht physikalisch zu interpretieren. Bei gleich bleibendem Durchsatz steigt die Durchlaufzeit mit steigendem Warenbestand. An dem Beispiel einer einfachen Linie wird der Zusammenhang leicht deutlich. In Abbildung 2.15 kann das markierte Los in der sonst leeren Linie 1 seine Rohprozesszeit erreichen, während in Linie 2 mit einigen Wartezeiten zu rechnen ist. In Linie 3 werden Flussfaktor und Durchlaufzeit am höchsten sein, auch wenn der Durchsatz bei gleichen Werten bleibt.

$$\text{Flussfaktor: } FF = \frac{CT}{RPT} \quad (2.12)$$

$$\text{Little's Law: } CT = \frac{WIP}{GR} \quad (2.13)$$

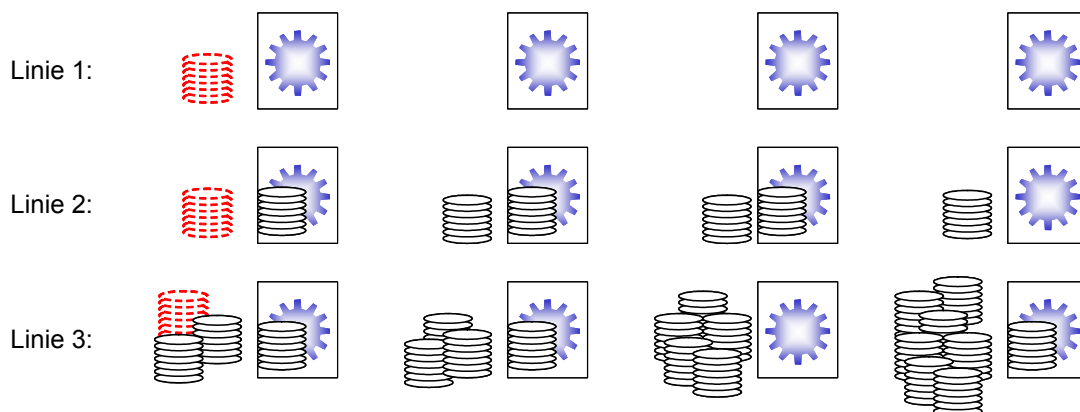


Abbildung 2.15: Wichtige Kennzahlen: Flussfaktor und Little's Law

Wird der Flussfaktor als Steuermittel verwendet, so bestimmt er die Soll-Durchlaufzeit  $SollCT$ , da die Rohprozesszeit ein fester Wert pro Produkt ist. Dabei gibt es pro Produkt für

jede einzelne Operation einen eigenen Flussfaktor (siehe Abbildung 2.16). Bei einer kurzen Operation kann eine Störung eine prozentual gesehen größere Zeitverzögerung bewirken als bei einer langen Operation. Innerhalb der Soll-Durchlaufzeit ist somit für jede einzelne Operation eine definierte Zeitdauer vorgesehen. Damit kann auch ermittelt werden, am wievielten Tag nach Einschleusung der Scheibe jede einzelne Operation fertig sein soll. Diese Tage werden Ranges genannt [She97]. Durch die mögliche unterschiedliche Vorgabe der Flussfaktoren können sich unterschiedliche Aufteilungen der Operationen auf die Ranges ergeben. Dieser Einfluss wird dynamischer Effekt genannt.

$$SollICT = RPT \cdot SollFF \tag{2.14}$$

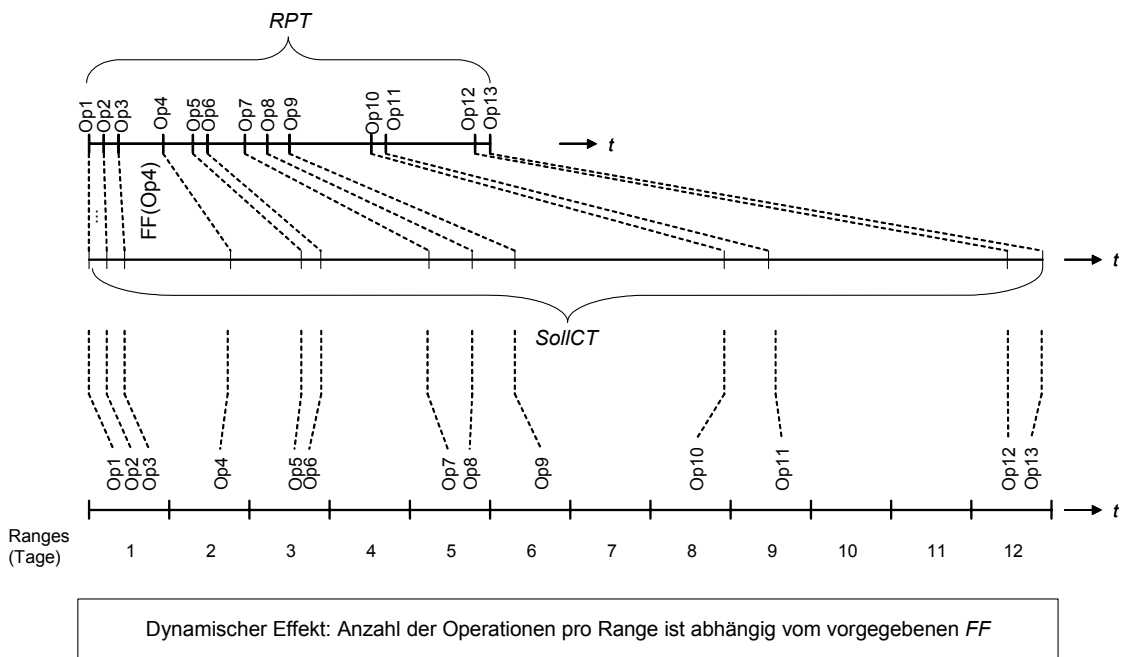


Abbildung 2.16: Aufteilung der Operationen in Ranges

### 2.6.2 Taktrate

Ist die Aufteilung der Operationen aller Produkte auf ihre Ranges bekannt, so können die Operationen ermittelt werden, die an jedem Tag in der Produktion zu erbringen sind, um die Einschleusung gleichmäßig abzuarbeiten (siehe Abbildung 2.17). Je nachdem wann ein Los eingeschleust wurde, befindet es sich in einem seiner Ranges. Die Summe der Operationen aller Lose in ihrem jeweiligen aktuellen Range ergibt alle zu erbringenden Operationen eines Produktionstages. Die Anzahl der zu erbringenden Operationen wird Taktrate bezeichnet. Es gibt jeweils eine Taktrate pro Operation.



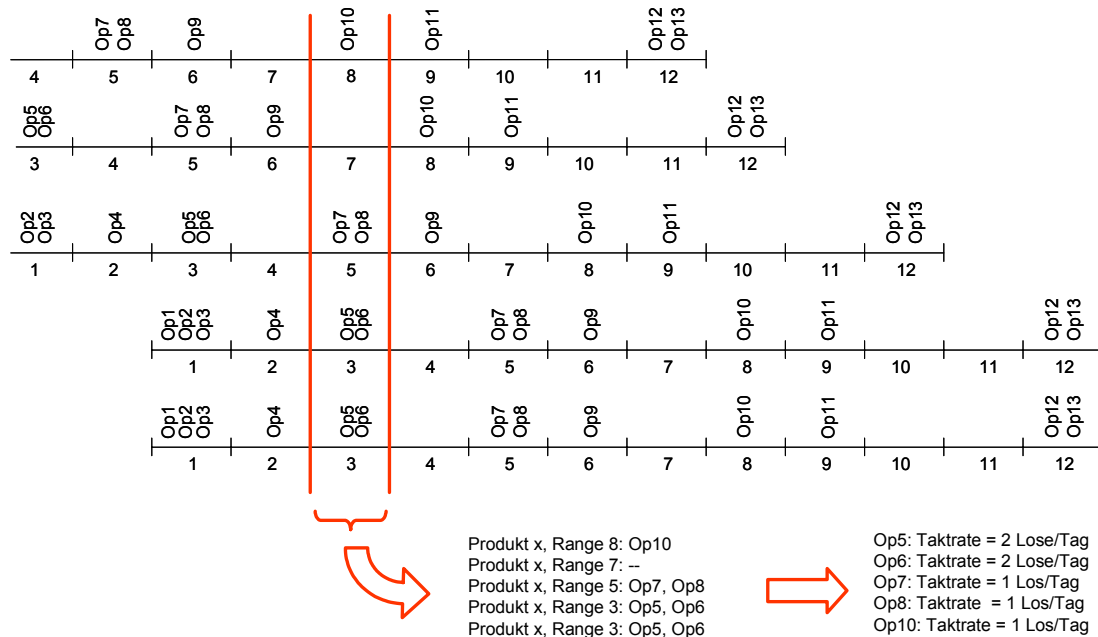


Abbildung 2.17: Taktrate je Operation und Produkt am Beispiel von fünf Losen mit gleicher Prozessfolge aber unterschiedlicher Einschleusung

Jede abgeschlossene Operation ist eine Bewegung. Bewegungen (**Moves**) werden pro prozessierter Scheibe gezählt. Die Soll-Bewegungen pro Operation und Tag wiederum werden mittels der Taktrate und einiger Korrekturfaktoren ermittelt. Der Speedfactor ist ein Beschleunigungsfaktor, um den vorgegebenen Wert der Taktrate leicht zu variieren. Um realistische Vorgaben zu erreichen, wird die Taktrate mit dem Losbestand, der in den nächsten 24 Stunden vor der Operation zu erwarten ist (*WIP24*), verglichen. Entsprechend ergeben sich unterschiedliche Soll-Durchlaufzeiten [Sar99b]. Erreichen zu wenige Lose die Operation, wird mit dem Kompensationsfaktor  $\kappa$  die Unstimmigkeit, also die Differenz zwischen vorhandenem Material und eigentlich verlangter Taktrate, ausgeglichen. Der Intensivierungsfaktor  $\gamma$  beschleunigt bei zu großem Losaufkommen die Produktion. Mit einem Geschwindigkeitsfaktor  $\beta$  können Schwankungen der Produktionsmenge berücksichtigt werden.

Somit ergibt sich für jeden Tag eine Anzahl von Soll-Bewegungen (*SollDGR*) pro Operation sowie als Summe davon die Soll-Bewegungen für die gesamte Produktion.

$$\begin{aligned}
 \text{Taktrate} > \text{WIP24} &\Rightarrow \text{SollDGR} = \text{WIP24} + (\text{Taktrate} \cdot \beta - \text{WIP24}) \cdot \kappa \\
 \text{Taktrate} \leq \text{WIP24} &\Rightarrow \text{SollDGR} = \text{Taktrate} \cdot \beta \cdot \gamma
 \end{aligned}
 \tag{2.15}$$

Beispiel: Für Intensivierungsfaktor  $\gamma = 1,25$ , Kompensationsfaktor  $\kappa = 0,25$  und Geschwindigkeitsfaktor  $\beta = 1$  ergeben sich folgende Soll-Durchsätze:

$$\begin{aligned}
 \text{Taktrate} > \text{WIP24} &\Rightarrow \text{SollDGR} = \text{WIP24} + (\text{Taktrate} \cdot 1 - \text{WIP24}) \cdot 0,25 \\
 &= \text{WIP24} + 0,25 \cdot \text{Taktrate} - 0,25 \cdot \text{WIP24} \\
 \text{Taktrate} \leq \text{WIP24} &\Rightarrow \text{SollDGR} = \text{Taktrate} \cdot 1 \cdot 1,25 \\
 &= 1,25 \cdot \text{Taktrate}
 \end{aligned}
 \tag{2.16}$$

Bewegen sich alle eingeschleusten Scheiben mit ihrer Solldurchlaufzeit *SollCT* durch die Linie, dann würden die Bewegungen jeder Operation ihrer Taktrate entsprechen. Es ist aber möglich, dass durch den Produktmix der Einschleusung die technischen Gegebenheiten und die Kapazitäten einzelner Operationen an einigen Tagen dazu nicht ausreichen. Die Einschleusung darauf einzustellen wäre extrem aufwendig, da die einzelnen Schwankungen innerhalb einer Produktionslinie nie genau beschreibbar oder vorhersagbar sind. Es wird nie vermeidbar sein, dass einige Maschinen leer stehen, während sich vor anderen Maschinen eine Warteschlange bildet. Die Vorgabe der Soll-Bewegungen bewirkt eine Entzerrung und Glättung des Materialflusses.

### 2.6.3 Dispatch-Regeln

Eine weitere Steuerungs-Vorgabe sind Dispatch-Regeln, die die Reihenfolge der zu bearbeitenden Lose festlegen (siehe Tabelle 2.4). Sie können in statische und dynamische Regeln unterteilt werden [Ask93]. Statische Prioritätsindizes bleiben während des Produktionsvorgangs gleich, dynamische ändern sich im Verlauf der Produktion.

Name	Beschreibung
SPT	Shortest Processing Time, wählt das Los mit der kürzesten Bearbeitungszeit
EDD	Earliest Due Date, wählt das Los mit dem frühesten Liefertermin
FCFS	First Come First Served, wählt das Los, das als erstes eingetroffen ist; auch als First In First Out FIFO bekannt
FISFS	First In System First Served, wählt das Los, das als erstes eingeschleust wurde
S/RO	Slack per Remaining Operation, wählt das Los mit geringstem Slack; Slack = Liefertermin - aktuelle Zeit - Zeit der verbleibenden Arbeit
LTWK	Least Total Work, wählt das Los mit der kürzesten Gesamtbearbeitungszeit
LWKR	Least Work Remaining, wählt das Los mit der kürzesten Restbearbeitungszeit
MOPNR	Most Operations Remaining, wählt das Los mit der größten Zahl verbleibender Operationen
MWKR	Most Work Remaining, wählt das Los mit der höchsten Restbearbeitungszeit
RANDOM	Random, wählt ein Los zufällig aus

Tabelle 2.4: Dispatch-Regeln

Darüber hinaus können Loseigenschaften, die nicht alle Lose haben, dazu führen, dass einzelne Lose vorgezogen werden. Eine mögliche Reihenfolge ist in Abbildung 2.18 dargestellt. Blitz- und Eillose haben erste Priorität. Hot Lots werden gebildet, wenn vor einer Maschine bzw. einem Workcenter eine längere Warteschlange steht. Danach haben Lose,

deren Produktionsfortschritt eine Verspätung aufweist, Vorrang vor Losen, die zur Anlagenkontrolle dienen. An fünfter Stelle werden Lose bearbeitet, die eine Nacharbeit benötigen, bevor der Erfüllungsgrad der Bewegungen des Produktes berücksichtigt wird. Alle restlichen Lose werden nach dem FIFO-Prinzip (First In First Out) entsprechend ihrer Ankunftsreihenfolge bearbeitet.

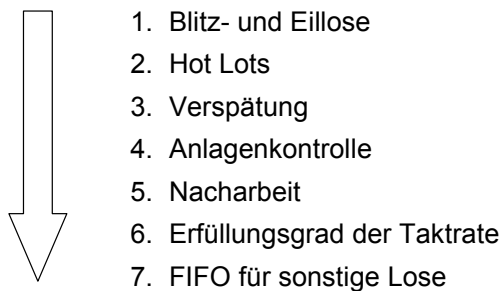


Abbildung 2.18: Beispiel zur Reihenfolge von Dispatch-Regeln

Während unter Verwendung der FIFO-Regel der Produktionsfluss immer stabil bleibt, kann die Verwendung des Fälligkeitstermins der Lieferung (Due Date) als Dispatch-Kriterium nach [Ros03] zu instabilen Zuständen führen. Die Wahl der Dispatch-Regeln hat also nicht nur Auswirkungen auf die Bearbeitungsreihenfolge, sondern unter Umständen auch auf weitere Bereiche der Produktion.

#### 2.6.4 Dynamische Kennzahlen

Die Ermittlung der Kennzahlen einer Halbleiterproduktion beruht auf der Aufnahme von Zeitmarken (Time-Stamps) an den einzelnen Operationen, die in entsprechenden Datenbanken verarbeitet und gespeichert werden. Time-Stamps sind Datensätze, die Informationen über das jeweilige Los, die Zeit, die Operation, etc. enthalten. Ist ein Los an einer Operation fertig prozessiert, wird es von dieser Operation abgemeldet, was einen Time-Stamp erzeugt (siehe Abbildung 2.19). Dies erfolgt meist durch Barcodelesegeräte. Dabei werden die Barcodes von Los, Operation (auf dem Laufzettel vermerkt) und Maschine eingelesen und zu einem Datenbankeintrag verarbeitet. Der Transport zur nächsten Operation ist eine Bringschuld, d.h. sie muss von den Operatoren der abmeldenden Station erbracht werden. Ab einer gewissen Anzahl von fertig prozessierten Losen (z.B. drei) muss ein Transport erfolgen. Die Zeit, die für diesen Transport benötigt wird, zählt allerdings bereits zur nachfolgenden Operation. Wenn die Lose nicht sofort an die Maschinen der nachfolgenden Operation gebracht werden können, werden sie in ein Regal gestellt. Das wird mit einer Einbuchung, ebenfalls mittels Barcodeleser, festgehalten. Ist eine entsprechende Maschine frei, so wird das Los angemeldet, wobei ein Move-In-Time-Stamp entsteht. Ist der Bearbeitungsvorgang beendet und das Los wird aus der Maschine herausgenommen, so entsteht wiederum ein Move-Out-Time-Stamp.

Die Zeitdifferenz zwischen den Move-Out-Time-Stamps zweier aufeinander folgender Operationen ist die Durchlaufzeit (Cycle Time) der zweiten Operation.

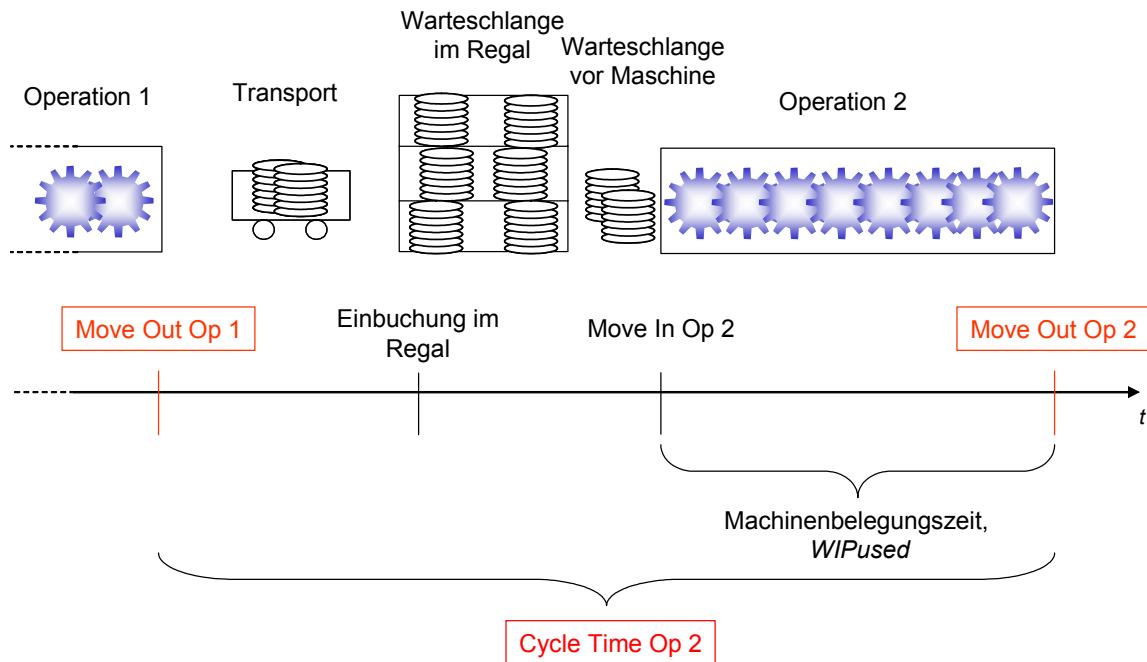


Abbildung 2.19: Aufnahme der Time-Stamped

Es kann also zu jeder Zeit über die Datenbank ermittelt werden, welches Los sich wann in welcher Entwicklungsstufe, an welcher Operation, in welchem Status (wartend, produktiv, etc.) und an welchem physikalischen Ort befindet oder befunden hat.

Die einzelnen Move-Out-Time-Stamped entsprechen jeweils einer Bewegung (Move). Sie können nach unterschiedlichen Kriterien sortiert und zusammengefasst werden (siehe Abbildung 2.20). Werden die Bewegungen eines Tages betrachtet (Daily Going Rate,  $DGR$ ), so können jeweils der Tagesdurchsatz einer Operation bezüglich eines Produktes (Arbeitsplan)  $DGR_{op}$ , der Tagesdurchsatz einer Operation  $DGR_o$  oder der Tagesdurchsatz der gesamten Linie  $DGR_{line}$  ermittelt werden.

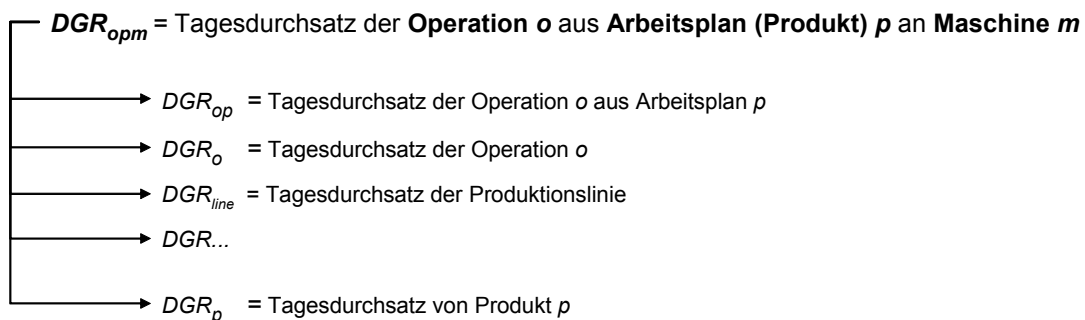


Abbildung 2.20: Formen der Going Rate

Der Tagesdurchsatz eines Produktes  $DGR_p$  kann auf zwei Arten ermittelt werden. So gibt es den tatsächlichen Durchsatz und den dynamischen Durchsatz. Der tatsächliche Durchsatz besteht aus den tatsächlich fertig produzierten Scheiben eines Produktes, also aus den Bewegungen der letzten Operation. Dieser Wert beinhaltet damit eine Aussage über

die Performance der Fertigungslinie während der gesamten Durchlaufzeit dieser Lose. Soll eine Aussage über die Produktivität der Fertigungslinie nur bezüglich des aktuellen Tages getroffen werden, wird der dynamische Durchsatz herangezogen.

Für den dynamischen Durchsatz  $dynDGR$  werden die Bewegungen aller Lose eines Produktes an allen Operationen aufsummiert und zur Anzahl aller benötigten Operationen für die Herstellung einer Scheibe des Produktes ins Verhältnis gesetzt. Für eine exakte Berechnung dieses Wertes müssten die einzelnen Operationen noch entsprechend ihrer unterschiedlichen Zeitdauer gewichtet werden, was aus Gründen der Vereinfachung nicht erfolgt. In den folgenden Formeln stellt  $O_p$  die Anzahl der benötigten Operationen des Produktes  $p$  dar, während  $M_{op}$  die Anzahl der Maschinen enthält, die für Operation  $o$  von Produkt  $p$  verwendbar sind.  $P$  ist die Anzahl der produzierten Produkte.

$$\begin{aligned}
 dynDGR_p &= \frac{1}{O_p} \sum_{o=1}^{O_p} \sum_{m=1}^{M_{op}} DGR_{opm} \\
 &= \frac{\text{alle Moves von Produkt } p}{\text{Anzahl Operationen von Produkt } p}
 \end{aligned}
 \tag{2.17}$$

$$dynDGR_{line} = \sum_{p=1}^P dynDGR_p = \sum_{p=1}^P \frac{moves_p / Tag}{\text{Anzahl Operationen in } p}
 \tag{2.18}$$

Die Kennzahl Cycle Time kann ebenfalls als absoluter oder dynamischer Wert ermittelt werden. Zunächst werden aus allen Move-Out-Time-Stamped die Durchlaufzeiten für jedes Los pro Operation, Arbeitsplan und Maschine  $CT_{n,opm}$  errechnet (siehe Abbildung 2.21). Durch Summenbildung und Durchschnittsberechnung ergeben sich die unterschiedlichen Durchlaufzeiten bezüglich eines Produktes, einer Operation oder der gesamten Linie.

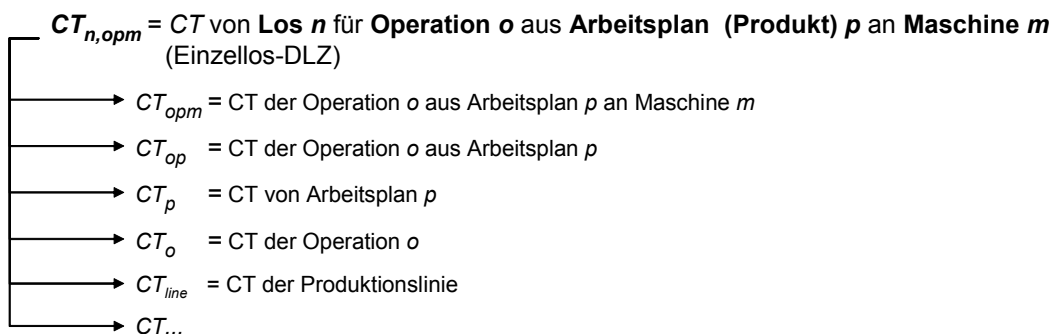


Abbildung 2.21: Formen der Cycle Time

Für die Berechnung der dynamischen Durchlaufzeit  $dynCT$  wird der Zusammenhang aus Little's Law verwendet und der Warenbestand zum dynamischen Durchsatz ins Verhältnis gesetzt. Während die gemessene tatsächliche Durchlaufzeit eine kundenrelevante Größe und wichtig für den Bereich Vertrieb ist, liefert die dynamische Durchlaufzeit eine Aussage über die Liniendynamik.

$$\text{dynCT} = \frac{WIP}{\text{dynDGR}} \quad (2.19)$$

Aus der dynamischen Durchlaufzeit  $\text{dynCT}$  folgt der dynamische Flussfaktor  $\text{dynFF}$ , der zunächst für jedes Produkt und anschließend für die gesamte Linie ermittelt wird. Der dynamische Flussfaktor ist in der Theorie der Betriebskennlinie für die Berechnung der Variabilität  $\alpha$  notwendig. Zunächst wird der dynamische Flussfaktor für jedes Produkt ermittelt:

$$\text{dynFF}_p = \frac{\text{dynCT}_p}{\text{RPT}_p} = \frac{\frac{WIP_p}{\text{dynDGR}_p}}{\text{RPT}_p} = \frac{\frac{WIP_p}{\text{moves}_p / \text{Tag}}}{\frac{\text{Anzahl der Operationen in } p}{\text{RPT}_p}} \quad (2.20)$$

Der dynamische Flussfaktor der gesamten Linie kann nicht über den arithmetischen Mittelwert der dynamischen Flussfaktoren der einzelnen Produkte berechnet werden. Dieser Wert würde die Performance der Linie nur ungenügend wiedergeben, da die einzelnen Produkte unterschiedliche Volumina haben. Vielmehr würde ein arithmetisches Mittel der Flussfaktoren aller produzierten Scheiben benötigt werden. Dieser Wert ist aber nur statisch und nicht dynamisch ermittelbar.

Deshalb werden die dynamischen Flussfaktoren der einzelnen Produkte verwendet und nach ihrem Anteil des gesamten Liniendurchsatzes gewichtet bevor sie addiert werden.  $P$  gibt wieder die Anzahl der produzierten Produkte an.

$$\begin{aligned} \text{dynFF}_{\text{line}} &= \frac{\sum_{p=1}^P \text{dynFF}_p \cdot \text{dynDGR}_p}{\text{dynDGR}_{\text{line}}} = \frac{\sum_{p=1}^P \frac{\text{dynCT}_p}{\text{RPT}_p} \cdot \text{dynDGR}_p}{\text{dynDGR}_{\text{line}}} \\ &= \frac{\sum_{p=1}^P \frac{\text{dynCT}_p}{\text{RPT}_p} \cdot \frac{WIP_p}{\text{dynCT}_p}}{\text{dynDGR}_{\text{line}}} = \frac{\sum_{p=1}^P \frac{WIP_p}{\text{RPT}_p}}{\text{dynDGR}_{\text{line}}} \end{aligned} \quad (2.21)$$

### 2.6.5 Liefertreue - Confirmed Line Item Performance CLIP

Mit *CLIP* (Confirmed Line Item Performance ~ Erfüllungsgrad der zugesagten Einzelposten) wird der akkumulierte Erfüllungsgrad der geplanten Auslieferungen einer Geschäftswoche bezeichnet. Der *CLIP* wird zunächst pro Produkt erfasst und anschließend für den Produktionsstandort berechnet. Somit eignet sich die *CLIP*-Betrachtung auch für Zwecke des Benchmarking.

Um den so genannten *Wochen-CLIP* eines Produktes zu berechnen, werden zunächst zwei Effektivwerte ermittelt: die Effektivanforderung und die Effektivlieferung. Die Effektivanforderung eines Produktes  $p$  für die betrachtete Geschäftswoche  $GW$  besteht aus

den aktuellen Aufträgen dieser Woche (Auftragsbestand) zuzüglich den bis zu dieser Woche aufsummierten Lieferrückständen (Auftragsrückstand). Ein Lieferrückstand besteht dann, wenn die Summe der ausgelieferten Chips geringer ist als die Summe der bestellten Chips.

$$\text{Effektivanforderung}_{p,GW} = \text{aktuelle Aufträge}_{p,GW} + \text{Lieferrückstände}_{p,GW} \quad (2.22)$$

Die Effektivlieferung eines Produktes  $p$  für die betrachtete Geschäftswoche  $GW$  ergibt sich aus den Auslieferungen der für diese Woche aktuellen Aufträge und der aufsummierten Mehrlieferung des Produktes bis zu dieser Woche. Eine Mehrlieferung besteht aus Vorauslieferungen und Überlieferungen. Vorauslieferungen ergeben sich, wenn die Summe der gelieferten Chips größer ist als die Summe der bestellten Chips und weitere Bestellungen für dieses Produkt in der Zukunft bestehen. Um Überlieferungen handelt es sich, wenn für dieses Produkt keine Bestellungen in der Zukunft bekannt sind.

$$\text{Effektivlieferung}_{p,GW} = \text{aktuelle Lieferung}_{p,GW} + \text{Mehrlieferung}_{p,GW} \quad (2.23)$$

Der *CLIP* von Produkt  $p$  für die Geschäftswoche  $GW$  wird als Quotient von Effektivlieferung zu Effektivanforderung berechnet. Sind mehr Chips ausgeliefert worden als von den Anforderungen nötig gewesen wären, so beträgt der *CLIP* = 100%. Somit ist der maximal mögliche *CLIP* = 100%.

$$\text{Wochen-CLIP}_{p,GW} = \min \left\{ \frac{\text{Effektivlieferung}_{p,GW}}{\text{Effektivanforderung}_{p,GW}}, 1 \right\} \quad (2.24)$$

Der *Wochen-CLIP* der gesamten Produktionsstätte wird als Durchschnitt aller Produkte mit einer Effektivanforderung berechnet. Dazu wird jedes Produkt einzeln betrachtet. Eine Gewichtung nach Auftragsvolumen wird nicht berücksichtigt.

$$\text{Wochen-CLIP}_{line} = \frac{\sum_{p=1}^{\text{Anzahl bestellter Produkte}} \text{Wochen-CLIP}_{p,GW}}{\text{Anzahl bestellter Produkte}} \quad (2.25)$$

### 2.6.6 Ausbeute - Yield

Während in anderen Produktionen Ausschuss gemessen wird, ist für die Halbleiterfertigung die Ausbeute (Yield,  $Y$ ) ausschlaggebend. Ein weiterer Unterschied zu anderen Fertigungen, wie z.B. des Automobilbaus, ist die Tatsache, dass Fehler der Produktion nur bedingt repariert werden können. Ein Automobil, das nicht vollständig funktioniert, erfährt eine Nacharbeit, so dass letzten Endes jedes produzierte Auto das Werk funktionstüchtig verlässt. Funktioniert dagegen ein Chip nicht, so kann ein Fehler in einer der Schichten nachträglich nicht mehr behoben werden.

Ursachen zur Ausbeuteverringeringung können zufälliger oder systematischer Natur sein. Defekte sind auf zufällige Ereignisse zurückzuführen. Falsche elektrische Parameter, fehlerhafte Layer, schlechte Ausrichtung der Layer oder eine schlechte Abstimmung des Chipdesigns auf die Technologie sind systematische Fehlerquellen.

Defekte, die während der Produktion zu Ausfällen führen, werden meist von Partikeln auf der Chipstruktur verursacht, die Kurzschlüsse oder Verbindungsunterbrechungen erzeugen. Dabei sind alle Partikel „Killer“, deren Durchmesser gleich oder größer als die Hälfte der Strukturgröße ist. Unter der Annahme, dass sich pro Flächeneinheit eines Wafers eine bestimmte Anzahl Partikel findet, ist die Chipgröße eine Einflussgröße für die Ausbeute. Je größer die Chipfläche ist, desto größer ist auch die Wahrscheinlichkeit, dass sich dort ein kritisches Partikel findet und den Chip funktionsunfähig werden lässt. Die Ausbeute hängt ebenso von der verwendeten Technologie sowie dem Schaltungsdesign ab, die die kritische Partikelgröße bestimmen.

Werden Defekte rechtzeitig im Produktionsablauf erkannt, so können sie unter bestimmten Bedingungen korrigiert werden. Ist z.B. bei der Belichtung eines Fotolackes die Ausrichtung nicht innerhalb der Toleranzen, so kann der gesamte Fotolack entfernt und der Produktionsschritt erneut durchgeführt werden. Ein so genannter Rework-Schritt beeinträchtigt dann allerdings die Performance der Produktion.

Wird ein Produkt neu in die Fertigung eingeführt, ist seine Ausbeute zunächst gering. Prozesse und Produkte müssen so angepasst werden, dass ein großer Yield möglich ist. Fehler in der Produktion müssen schnell beseitigt werden. Dieser Korrekturprozess ist als Yield Learning bekannt. Während der laufenden Produktion ist es wichtig, nicht reparable, sich neu entwickelnde systematische Fehler schnell zu erkennen, da sonst alle in der Zwischenzeit prozessierten Lose den gleichen Fehler aufweisen können und somit den Ausbeuteverlust stark vergrößern. So wird z.B. nach jeder Belichtung eine optische Kontrolle mittels Mikroskopie durchgeführt, um die Genauigkeit der Ausrichtung und Verlauf der belichteten Strukturen zu überprüfen.

Werte des Yields werden hauptsächlich an zwei Stellen der Fertigung erfasst [Cun90]. Bevor die Lose in den Wafer Test gelangen, wird die Linienausbeute  $Y_{line}$  (auch: Prozessyield, Scheibenyield oder Flächenyield) ermittelt. Sie ergibt sich aus dem Verhältnis aller eingeschleusten zu allen fertig prozessierten Wafern. Vollständige Scheibenverluste, die dazu führen, dass Wafer nicht mehr in den Wafer Test gelangen, können durch fehlerhafte Handhabung beim Transport oder Maschinendefekte während des Prozesses verursacht werden.  $P$  bezeichnet die Anzahl der Produkte und  $L_p$  die Anzahl der Lose für Produkt  $p$ .

$$Y_{line} = \frac{\sum_{p=1}^P \sum_{l=1}^{L_p} wafers\ out\ line_{lp}}{\sum_{p=1}^P \sum_{l=1}^{L_p} wafers\ in\ line_{lp}} \quad (2.26)$$

Im Wafer Test werden die einzelnen Chips auf ihre Funktionstüchtigkeit untersucht. Stimmen die elektrischen Werte eines Chips nicht mit den Anforderungen überein, wird er markiert und nach dem Zersägen des Wafers verworfen. Die Ausbeute im Wafer Test  $Y_{die}$  (Wafer Test-Yield, Die-Yield) ergibt sich aus dem Verhältnis aller funktionierenden zu allen getesteten Chips.



$$Y_{die} = \frac{\sum_{p=1}^P \sum_{l=1}^{L_p} \text{passed chips out wafer test}_{lp}}{\sum_{p=1}^P \sum_{l=1}^{L_p} \text{chips per wafer}_p} \cdot \frac{\sum_{p=1}^P \sum_{l=1}^{L_p} \text{wafers in wafer test}_{lp}}{\sum_{p=1}^P \sum_{l=1}^{L_p} \text{wafers in wafer test}_{lp}} \quad (2.27)$$

Da die Ausbeute ein statistischer Parameter ist, werden die Ausbeutewerte der einzelnen Produktionsbereiche multipliziert, um die gesamte Ausbeute zu erhalten. Der gesamte Yield  $Y$  einer Fertigung entspricht dann dem Verhältnis von brauchbaren Einheiten nach der vollständigen Fertigung zur Anzahl der potentiell brauchbaren Einheiten bei Beginn der Fertigung.

$$Y = Y_{line} \cdot Y_{die} \quad (2.28)$$

## 2.6.7 Weitere Kennzahlen

Für die Bewertung und Steuerung einer Halbleiterfabrik gibt es noch eine Vielzahl von weiteren Kennzahlen, die für die aktuelle Arbeit nicht im Vordergrund standen. Sie sollen hier nur zusätzlich erwähnt werden, um einen Eindruck ihrer Vielfalt zu vermitteln:

- *WIPende*: Anzahl des WIP am Ende einer Zeitperiode, meist am Ende des Tages
- Stop-Lose: sie geben einen Hinweis auf Prozess-Probleme
- Flussfaktor pro Prozessgruppe
- Layer Out Per Week
- Layer Start Per Week
- Wafer Out Per Week
- Wafer Start Per Week
- Manufactured Layer Per Week
- Transportzeiten
- Moves Per Operator
- Cycle Time Per Layer
- Maschinen- bzw. Workcenterzustände
- ...

Sämtliche Kennzahlen werden in einem festen Turnus (z.B. wöchentlich) ermittelt und dienen als Grundlage für Management-Entscheidungen und zu Benchmarking- Zwecken.

### 2.6.8 Betriebskennlinie und Variabilität

Die Betriebskennlinie (BKL, oder auch: Operating Curve Management, OCM) ist eine Theorie, mit deren Hilfe die Performance einer Fertigungslinie bewertet werden kann [Boe96]. Sie wird vor allem in Produktionen mit Farm Layout (siehe Kapitel 2.1) verwendet.

Die Grundlagen der Theorie der Betriebskennlinie liegen in der Warteschlangentheorie. Das einfachste Warteschlangensystem besteht aus einer Bedieneinheit (z.B. eine Maschine, ein Postschalter, etc.), sowie einem Ankunfts- und einem Bedienprozess. Die Ankunfts- und Bedienprozesse sind meist statistischer Natur, deren Zeiten durch eine Verteilungsfunktion mit Mittelwert (bzw. Ankunfts- und Bedienrate) und Standardabweichung beschrieben werden können, wie in Abbildung 2.22 gezeigt.

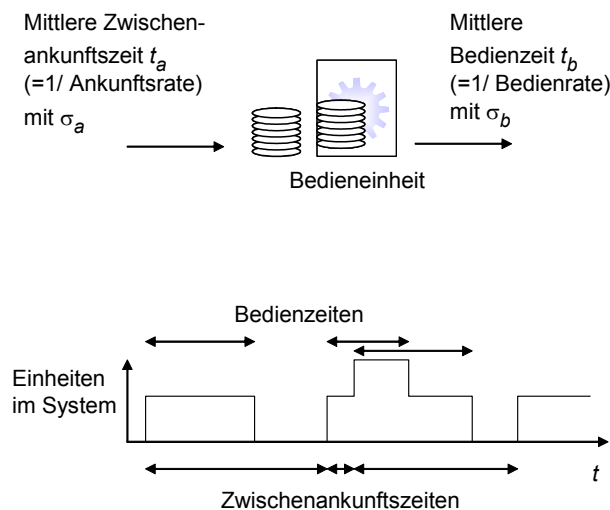


Abbildung 2.22: Größen der Warteschlangentheorie

Um unterschiedliche Wartesysteme beschreiben zu können, wird meist die Kendall-Notation (siehe Abbildung 2.23) verwendet [Gro98]. Von Ankunfts- und Bedienprozess werden jeweils die Typen ihrer Verteilung angegeben. Verwendet werden häufig:

- Exponentialverteilung, d.h. die Zufallsvariable Zwischenankunftszeit bzw. Bedienzeit ist exponentialverteilt, somit ist die Zufallsvariable „Ankünfte pro Zeiteinheit“ bzw. „Bediente Einheiten pro Zeiteinheit“ poissonverteilt. Es handelt sich hier um einen Markov-Prozess, der mit M abgekürzt wird,
- Erlangverteilung, die mit E abgekürzt wird,
- Hyperexponentialverteilung, H,
- mit G wird der allgemeine Fall einer beliebigen Verteilung angegeben, dessen Ereignisse statistisch unabhängig sind.

Da eine Warteschlange auch von mehreren Einheiten bedient werden kann, wird die Anzahl der Bedieneinheiten angegeben. Der Sonderfall „Anzahl der Warteplätze  $m = 0$ “ beschreibt ein reines Verlustsystem, bei  $m = \infty$  handelt es sich um ein reines Wartesystem. Häufig wird die Bezeichnung zu A|B|n verkürzt, wobei die Anzahl der Warteplätze als unendlich und die Bedienregel mit FIFO (First In First Out) angenommen werden.

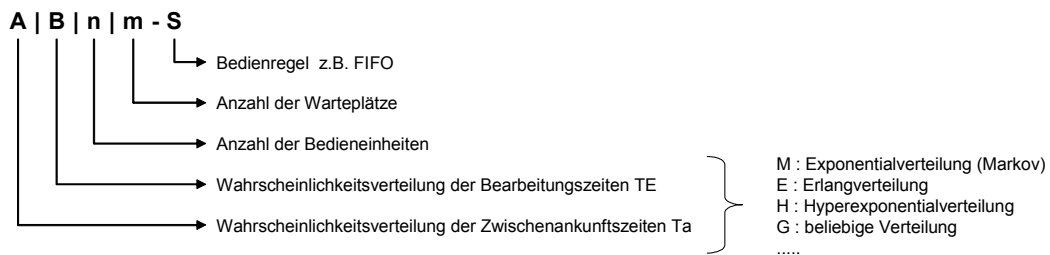


Abbildung 2.23: Kendall-Notation für Wartesysteme

Für die mittlere Wartezeit in einem System mit einer Bedieneinheit und beliebig verteilten Ankunfts- und Bedienzeiten sowie mittlerer Auslastung  $u$  wurde von Kingman eine Näherung ermittelt (siehe Abbildung 2.24). Danach ergibt sich die mittlere Wartezeit  $T_W$  für ein Element in der Warteschlange als Produkt dreier Faktoren: Variation, Auslastungsverhältnis und mittlere Bedienzeit  $t_b$ . Die Variation wird aus den Variationskoeffizienten der Ankunfts- und Bedienzeiten berechnet. Die Variationskoeffizienten  $c_a$  und  $c_b$  wiederum werden jeweils aus dem Mittelwert ( $t_a$  und  $t_b$ ) und der Standardabweichung ( $\sigma_a$  und  $\sigma_b$ ) der Zwischenankunftszeiten bzw. der Bedienzeiten errechnet. Die Auslastung  $u$  ergibt sich aus der Division von mittlerer Bedienzeit  $t_b$  und mittlerer Zwischenankunftszeit  $t_a$ .

Die Betriebskennlinie einer Produktion geht direkt aus der Kingman-Gleichung hervor, wobei die Durchlaufzeit  $CT$  aus der Wartezeit der Kingman-Gleichung und der Rohprozesszeit besteht [Han03]. Der Faktor der Variationskoeffizienten wird zu  $\alpha$  zusammengefasst, der Faktor der Auslastung bleibt bestehen und die mittlere Bedienzeit entspricht der Rohprozesszeit  $RPT$ .

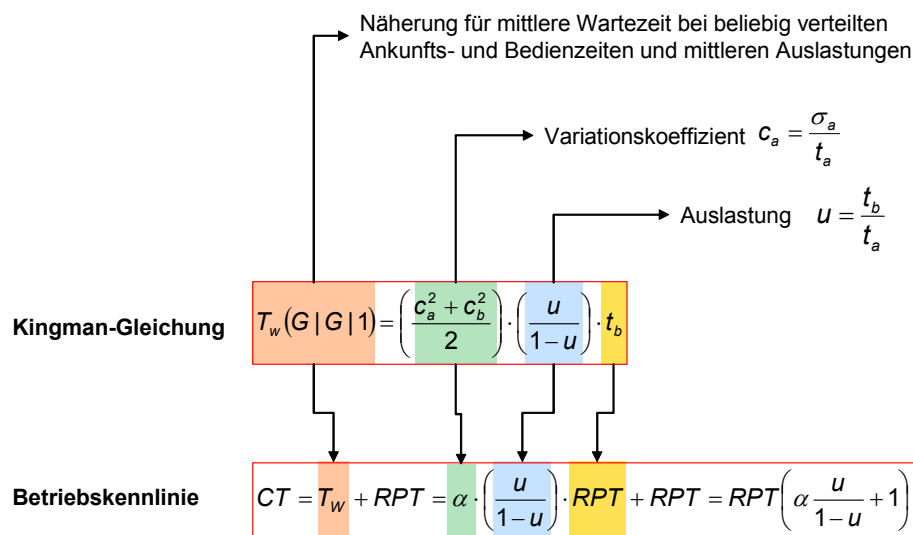


Abbildung 2.24: Von der Kingman-Gleichung zur Betriebskennlinie

Die Betriebskennlinie stellt somit den Zusammenhang zwischen Durchlaufzeit ( $CT$ , Cycle Time) und Durchsatz ( $GR$ , Going Rate) dar (siehe Gleichung (2.29)).

$$CT = RPT \left( \alpha \frac{U_{line}}{1 - U_{line}} + 1 \right) \quad (2.29)$$

Abbildung 2.25 veranschaulicht die Zusammenhänge. *Capa* ist dabei die Kapazität der Produktionslinie und in der Auslastung  $U_{line}$  enthalten (siehe unten).

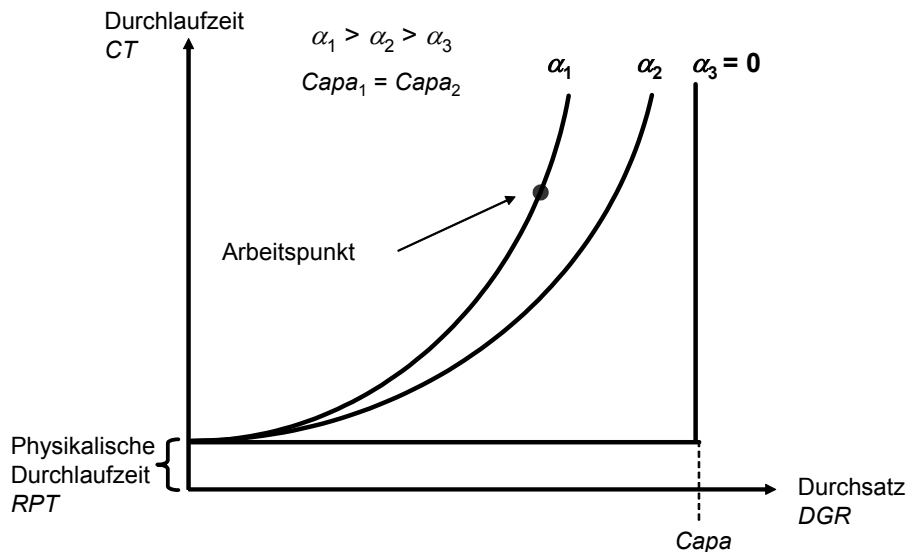


Abbildung 2.25: Die Betriebskennlinie (BKL, Operating Curve Management, OCM) [Eng00]

Für eine Produktionslinie mit bestimmtem Maschinenpark und ihren spezifischen Eigenschaften existiert theoretisch genau eine Betriebskennlinie. Die momentane Auslastung bzw. der Durchsatz bestimmt, an welchem Arbeitspunkt sich die Produktion gerade befindet. Der Arbeitspunkt beschreibt die momentane dynamische Performance eines Produktionssystems. Durch Veränderung der Einschleusung kann der Arbeitspunkt auf der gegebenen Betriebskennlinie verschoben werden. Eine Änderung der Leistungsfähigkeit des Produktionssystems findet dabei aber nicht statt. Die Entscheidung, ob schnelle Durchlaufzeiten bei geringerem Durchsatz einem größeren Durchsatz mit längeren Durchlaufzeiten vorzuziehen sind oder umgekehrt, ist eine klassische Management-Entscheidung.

Die Lage der Betriebskennlinie wird von der Variabilität  $\alpha$  bestimmt [Eng00]. Sie stellt die Gesamtheit aller in der Fertigung vorkommenden Schwankungen dar. Mit ihr wird die Qualität der Nicht-Einheitlichkeit von Prozesszeiten, Ausfallzeiten (Downtimes), Operatorverfügbarkeiten, Umrüstungszeiten, Transportstrategien, Wartungsstrategien, Entwicklungsstrategien, Produktmix, Lose in Halteoperationen, das Auftreten von beschleunigten Losen, etc. beschrieben [Bau01].

Für eine Variabilität von  $\alpha = 0$  verläuft die Betriebskennlinie auf der Höhe der Rohprozesszeit (Physikalische Durchlaufzeit,  $RPT$ ) bis zur Kapazität der Fertigungslinie. Dort geht sie schlagartig ins Unendliche über. Für eine Produktion bedeutet dies: Solange in allen Prozessen und Vorgängen keine zeitlichen Schwankungen auftreten, entspricht die Durchlaufzeit der Rohprozesszeit. Diese Durchlaufzeit bleibt für jeden beliebigen Durchsatz

erhalten bis er die Kapazitätsgrenze erreicht hat. Dann wird das System instabil und die Durchlaufzeit würde unendlich werden. In der Realität jedoch sind fast alle Zeitabläufe mit statistischen Fluktuationen behaftet. Darüber hinaus sind die einzelnen Ereignisse voneinander abhängig.

Ein Ausgleich der Schwankungen kann daher nicht stattfinden. Die Abhängigkeit der Prozesse (eine Maschine kann nur produzieren, wenn Material vorhanden ist) führt dazu, dass „verlorene Zeiten“ nicht mehr aufgeholt werden können. Vielmehr findet eine Akkumulation der Fluktuationen statt, die das Gesamtsystem verlangsamen, weil die gegenseitigen Abhängigkeiten eine zwischenzeitliche Beschleunigung stark einschränken.

Die Ermittlung der Variabilität  $\alpha$  der gesamten Linie erfolgt mittels dynamischem Flussfaktor  $dynFF_{line}$  und Auslastung (Utilisation)  $U_{line}$ . Unter Verwendung von Gleichung (2.29) und der Definition des Flussfaktors (2.12) kann die Variabilität wie folgt berechnet werden:

$$\alpha_{line} = \frac{(dynFF_{line} - 1)(1 - U_{line})}{U_{line}} \quad (2.30)$$

Die Variabilität  $\alpha$  ist eine Kennzahl, mit der auch die Performance einer gesamten Halbleiterproduktion beschrieben werden kann. Daher eignet sie sich auch für Zwecke des Benchmarking [Aur97]. Durch die starke Informationsverdichtung, die bei der Berechnung der Variabilität entsteht, ist es allerdings nicht möglich, mit der Theorie der Betriebskennlinie die Ursachen für eine „gute“ oder „schlechte“ Produktionsperformance zu erkennen. Dazu werden andere Kennzahlen benötigt, wie u.a. oben beschrieben.

Um die Bedeutung der Variabilität herauszustreichen formulieren Hopp und Spearman [Hop01] folgend Aussage:

“Pay me now or pay me later!”

Was zu bedeuten hat: Wenn man nicht für die Reduzierung der Variabilität zahlt, dann wird man auf die ein oder andere der folgenden Arten zu zahlen haben: verlorener Durchsatz, verschwendete Kapazität, erhöhte Durchlaufzeiten, höhere Lagerbestände, lange Lieferzeiten und/oder schlechter Kundenservice.

Zur Berechnung der Auslastung wird die Linienkapazität benötigt, für die der durchsatzbasierte Ansatz verwendet werden soll. Hat ein Modul einen geringeren Durchsatz als die Lithografie, so gilt es als Bottleneck und sein maximal möglicher Durchsatz wird als Kapazität der Linie bezeichnet. Sonst stellt typischerweise die Lithografie das Bottleneck dar.

$$Capa_{line} = \text{maximaler Durchsatz des Bottlenecks} \quad (2.31)$$

Manchmal findet auch eine zeitbasierte Berechnung für die Linienkapazität Verwendung, die den Anteil der produktiven Zeit an der Uptime über alle Maschinen aufsummiert. Dieser Ansatz ergibt einen größeren aber unrealistischeren Wert.

Die Auslastung der Linie berechnet sich dann aus Durchsatz und Auslastung zu:

$$U_{\text{line}} = \frac{DGR}{Capa_{\text{line}}} \quad (2.32)$$

Sollen gleichzeitig geringere Durchlaufzeiten und ein größerer Durchsatz erzielt werden, muss eine Performance-Steigerung der Fertigungslinie erreicht werden. Dazu muss eine neue optimierte Kennlinie realisiert werden. Grundsätzlich sind zwei Optimierungsansätze möglich [Eng00]: Eine Reduzierung der Schwankungseinflüsse und somit der Variabilität  $\alpha$  führt zu einer tiefer liegenden Kennlinie (Abbildung 2.26). Eine Erhöhung der Kapazität  $Capa$  (z.B. durch Zukauf weiterer Maschinen) verschiebt die Betriebskennlinie nach rechts (Abbildung 2.27). In beiden Fällen ergibt sich ein Bereich mit möglichen neuen optimierten Arbeitspunkten. Umgangssprachlich wird dieser Zusammenhang gerne als „Manta-Prinzip“ bezeichnet: Die niedrigere Kennlinie ist die bessere.

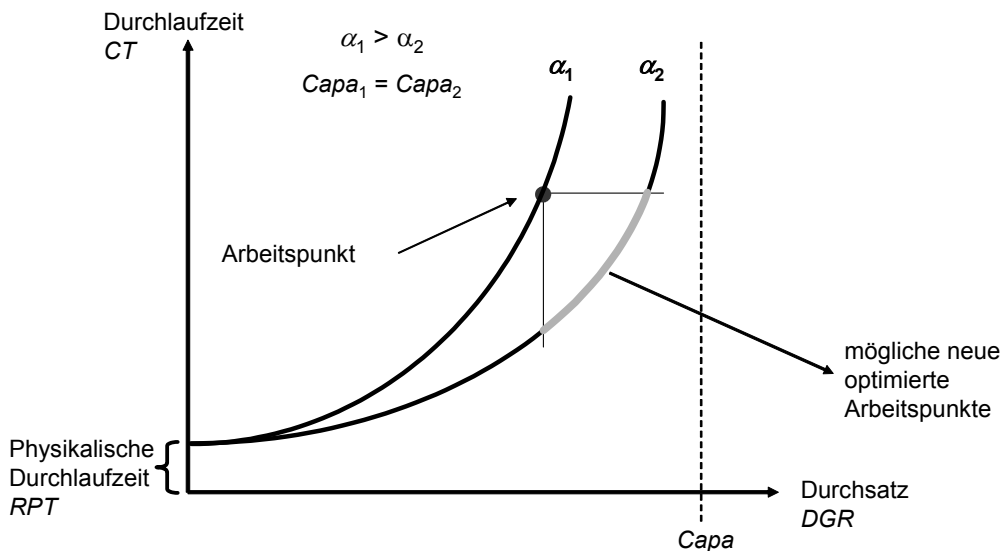
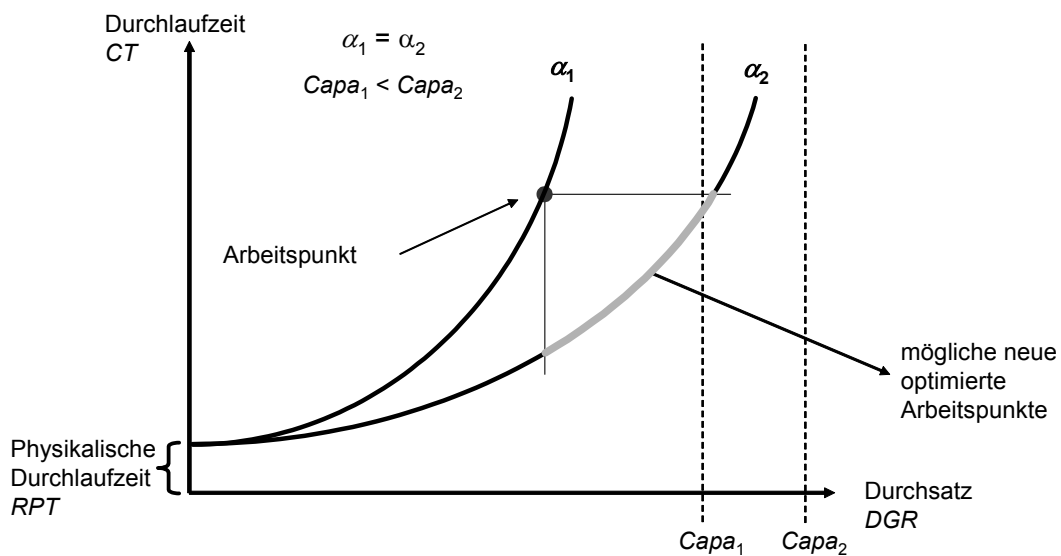


Abbildung 2.26: Optimierte Arbeitspunkte durch Verringerung der Variabilität  $\alpha$

In der Praxis wirken Verbesserungsmaßnahmen meist auf beide Kenngrößen, die Variabilität  $\alpha$  und die Kapazität  $Capa$ . Dabei sollte vor Einführung der Maßnahme untersucht werden, ob die Verbesserung der einen Größe nicht mit der Verschlechterung der anderen „erkauft“ wird. Letzten Endes soll nach einer Verbesserungsmaßnahme ein Arbeitspunkt erreicht werden, der gleichzeitig bei einer geringeren Durchlaufzeit und einem größeren Durchsatz liegt.

Geeignete Maßnahmen für eine Performance-Verbesserung können mit der Theorie der Betriebskennlinie nicht ermittelt werden. Dazu wird meist vom Vier-Partner-Modell ausgegangen. Zur Überprüfung des Erfolges einer Verbesserungsmaßnahme ist die Betriebskennlinie jedoch sehr gut geeignet.

Abbildung 2.27: Optimierte Arbeitspunkte durch Erhöhung der Kapazität  $Capa$ 

### 2.6.9 Overall Fab Efficiency OFE

Der Standard SEMI E79 bietet eine Möglichkeit, die Overall Equipment Efficiency  $OEE$  zu berechnen. Dieser Standard findet in der Industrie inzwischen breite Anerkennung und Verwendung. In einer Halbleiterfab jedoch agieren die einzelnen Anlagen nicht unabhängig voneinander. Vielmehr arbeiten sie in einer vernetzten und komplexen Umgebung. Somit erscheint es sinnvoll, die Performance einer gesamten Halbleiterfab zu bewerten.

In SEMI E124-0703 „Provisional Guideline for Definition and Calculation of Overall Factory Efficiency (OFE) and Other Associated Factory-Level Productivity Metrics“ wird die Berechnung einer Overall Factory Efficiency vorgeschlagen.

Die Overall Fab Efficiency OFE soll angeben, wie gut eine Fab arbeitet im Vergleich dazu, wie gut sie für den entsprechenden Produktmix arbeiten könnte. Sie berechnet sich als Produkt aus Produktivitätseffizienz (Productivity Efficiency) und Ausbeuteeffizienz (Yield Efficiency).

$$\text{overall fab efficiency} = \text{productivity efficiency} \cdot \text{yield efficiency} \quad (2.33)$$

Die Produktivitätseffizienz wird dabei aus einer Durchsatzrate, einer Durchlaufzeiteffizienz und einem „normalisierenden“ Exponenten (Idealitätsfaktor) gebildet.

$$\text{productivity efficiency} = \left( \frac{\text{throughput rate and}}{\text{cycle time efficiency}} \right)^{\text{normalizing exponent}} \quad (2.34)$$

Die Ausbeuteeffizienz ist das Produkt aus Linienausbeute (Line Yield) und Testausbeute (Test Yield), was gleichbedeutend wie das Verhältnis von funktionierenden Wafern zur Gesamtzahl der produzierten Wafer ist.

$$\text{yield efficiency} = \text{line yield} \cdot \text{test yield} = \frac{\text{good wafer equivalents out}}{\text{total wafers out}} \quad (2.35)$$

Wie bei allen bisher erörterten Kennzahlen finden auch hier Kosten keine Berücksichtigung. In [Oec02] werden deshalb Maßzahlen zur Charakterisierung einer Overall Fab Effectiveness gefordert, die einen Wert für die Kosten pro geliefertem Chip (Cost per Die Out) ergeben.



# Kapitel 3

## Liefertreue in mehrstufigen Produktionen

### 3.1 Einleitung

Die Halbleiterindustrie hat eine der längsten und komplexesten Herstellungsprozesse und Supply Chains. Der schnell wechselnde Markt, die rasch voranschreitende Technologie und der kontinuierliche Preisverfall üben einen enormen Druck auf die Hersteller aus. Die Lieferzeiten sollen verringert werden, während die Kosten niedrig gehalten werden müssen. Als Folge davon kommen der Steuerung und Leistungsbewertung der internen Supply Chain eine immer größere Bedeutung zu [Lev04]. Die größte Herausforderung liegt in ihrer Struktur mit weltweiten Fertigungsstätten, Beziehungen mit Zulieferern und Verflechtungen mit ausgelagerten, externen Fertigungsschritten. Der sich hieraus ergebende Materialfluss ist hinsichtlich seiner logistischen und steuerungstechnischen Eigenschaften sehr heterogen.

Um eine Halbleiterproduktion steuern und bewerten zu können, werden daher viele so genannte Performance-Parameter verwendet. Sie werden teilweise als Soll-Werte vorgegeben und dann mit den tatsächlich erreichten Ist-Werten verglichen. Während die Ist-Werte auf Messergebnissen beruhen und berechenbar sind und somit als technische Performance-Kennzahlen bezeichnet werden können, gehören die Vorgabewerte in den Bereich des Managements und der Betriebswirtschaft. Sie beschreiben nicht die tatsächlichen Zustände, sondern werden als Steuerungselemente verwendet. Zu den Werten, die in beiden Formen verwendet werden, gehören u.a. die Confirmed Line Item Performance (der *CLIP*, Erfüllungsgrad der zugesagten Einzelposten, siehe Kapitel 2) und der Service Level (*SL*, Liefertreue gegenüber dem Kunden). Dabei hängen diese Leistungsgrößen von anderen Parametern ab, bzw. ergeben sich aus ihnen. So ist der *CLIP* abhängig von den Durchlaufzeiten der Produkte, der angesetzten Plandurchlaufzeit sowie der Ausbeute (Yield, *Y*).

Es scheint, dass sich die für Frontend und Backend vorgegebenen *CLIP*-Werte sowie der Service Level des Kundenendtermins gegenseitig beeinflussen und im Grunde nicht unabhängig festgelegt werden können. Mangels einfacher Berechnungsmöglichkeiten wird dies in der Praxis jedoch oft so durchgeführt.

Um die Wechselwirkungen und gegenseitigen Abhängigkeiten der einzelnen Liefertreu-Parameter darzustellen, wurde das in diesem Kapitel beschriebene Modell entwickelt, das den vollständigen Produktionsablauf von der Kundenbestellung bis zur Auslieferung abbildet.

Wenn die existierende Struktur durch das Modell erfasst ist und ihre Mechanismen verstanden werden, ist es möglich, Überlegungen zu Verbesserungsmaßnahmen anzustellen [Hil05].

## 3.2 Statistische Grundlagen

### 3.2.1 Wahrscheinlichkeitsrechnung

Ist eine Datenmenge vollständig bekannt, können Methoden der beschreibenden Statistik angewandt werden. Die Beschreibung von Zufallsvariablen geschieht meist mittels ihrer Dichte-, Wahrscheinlichkeits- oder Verteilungsfunktionen. Es gibt diskrete und kontinuierliche Verteilungen.

Bei einer kontinuierlichen Zufallsvariablen tritt an die Stelle der Wahrscheinlichkeitsfunktion die so genannte Dichtefunktion (siehe Abbildung 3.1). Aus der Wahrscheinlichkeitsfunktion einer diskreten Zufallsvariablen können die Wahrscheinlichkeiten der einzelnen Werte abgelesen werden. Dagegen können aus der Dichtefunktion einer kontinuierlichen Zufallsvariablen nur Wahrscheinlichkeiten für Intervalle berechnet werden. Die Verteilungsfunktion beschreibt die Summenwahrscheinlichkeit, also die Wahrscheinlichkeit dafür, dass ein Wert kleiner oder gleich eines bestimmten Wertes  $x$  eintritt.

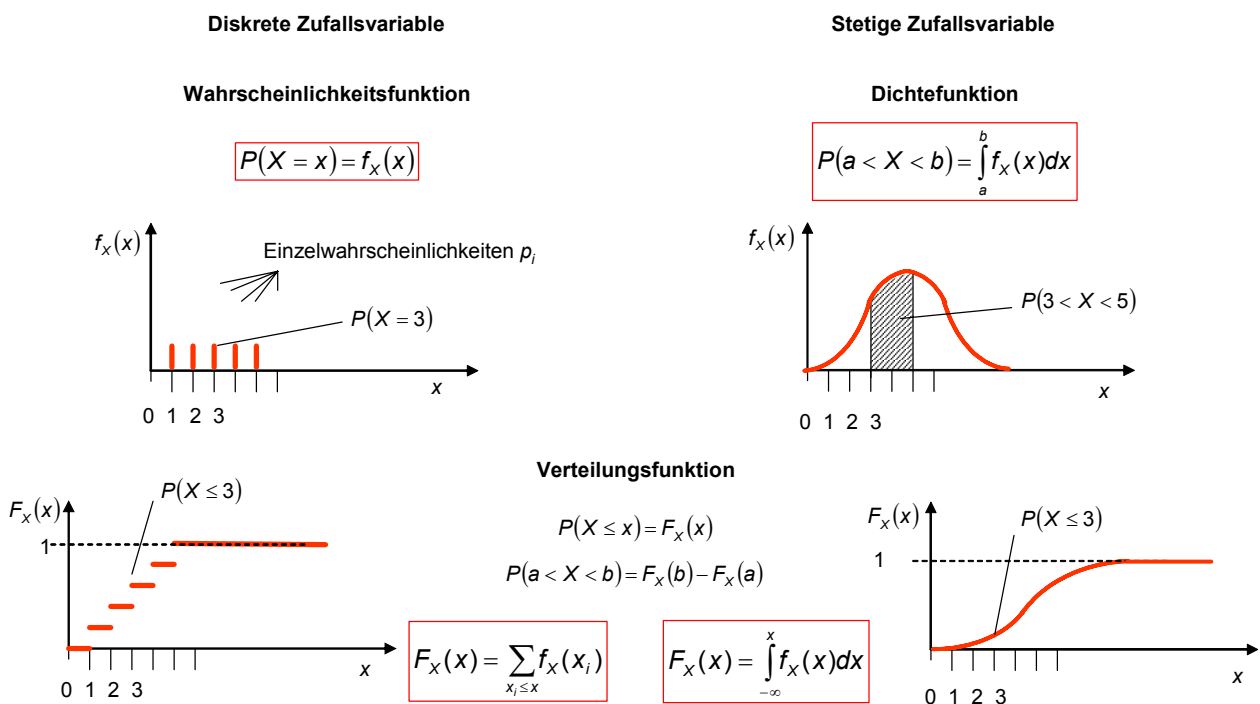


Abbildung 3.1: Darstellung kontinuierlicher und diskreter Zufallsvariablen

Um einzelne Wahrscheinlichkeiten zu beschreiben werden folgende Schreibweisen verwendet:

Mit  $P(X = 3) = 0,2$  wird die Wahrscheinlichkeit, dass die Zufallsvariable  $X$  den Wert 3 annimmt, mit 0,2 angegeben.

$P(X \leq 3) = F_X(3) = 0,2$  steht für die Wahrscheinlichkeit von 0,2, dass die Zufallsvariable einen Wert kleiner oder gleich 3 annimmt. Da dies genau der Summenwahrscheinlichkeit für  $x = 3$  entspricht, kann auch die Bezeichnung der Summenwahrscheinlichkeit benutzt werden.

$P(2 \leq X \leq 3) = F_X(3) - F_X(2) = 0,1$  gibt die Wahrscheinlichkeit dafür, dass die Zufallsvariable einen Wert zwischen 2 oder 3 annimmt, mit 0,1 an.

Die Summenwahrscheinlichkeit ist die Wahrscheinlichkeit dafür, dass die Realisierung einer Zufallsvariablen kleiner oder gleich eines bestimmten Wertes ist. Die umgekehrte Zuordnung eines Variablenwertes zu einer Wahrscheinlichkeit wird mit Quantilen beschrieben (siehe Abbildung 3.2). Quantile werden mit dem Kleinbuchstaben der Zufallsvariablen und der Wahrscheinlichkeit als Index bezeichnet. Das  $1/3$ -Quantil  $x_{1/3}$  ist also der Variablenwert für den gilt:  $P(X \leq x_{1/3}) = 1/3$ .

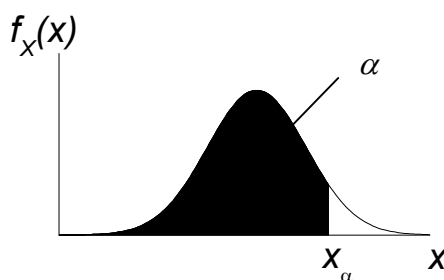


Abbildung 3.2: Das  $\alpha$ -Quantil  $x_\alpha$  ist der  $x$ -Wert für den gilt:  $P(X \leq x_\alpha) = \alpha$

### 3.2.2 Die Normalverteilung

Die Normalverteilung ist die wichtigste stetige Verteilung. Sie findet immer dann Verwendung, wenn additive Überlagerungen von Zufallsgrößen auftreten.

Der Gipfel der Dichtefunktion der Normalverteilung liegt über dem Mittelwert  $\mu$ . Die Kurve ist symmetrisch, fällt nach beiden Seiten glockenförmig ab und nähert sich der Abszissenachse asymptotisch. Die Kurve hat zwei Wendepunkte, deren Abstand vom Mittelwert  $\mu$  als Standardabweichung  $\sigma$  bezeichnet werden (Abbildung 3.3). Die Normalverteilung ist durch ihren Mittelwert (= Erwartungswert) und die Standardabweichung vollständig beschrieben. Mathematisch ist sie durch ihre Dichtefunktion  $g(x, \mu, \sigma)$  gegeben mit:

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (3.1)$$

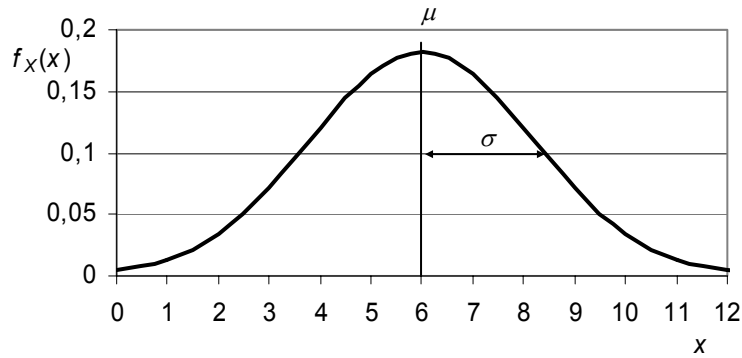


Abbildung 3.3: Dichtefunktion einer normalverteilten kontinuierlichen Zufallsvariablen mit  $\mu = 6$  und  $\sigma = 2,2$

Die Verteilungsfunktion der Normalverteilung ist nicht mehr mit Hilfe elementarer Funktionen darstellbar.

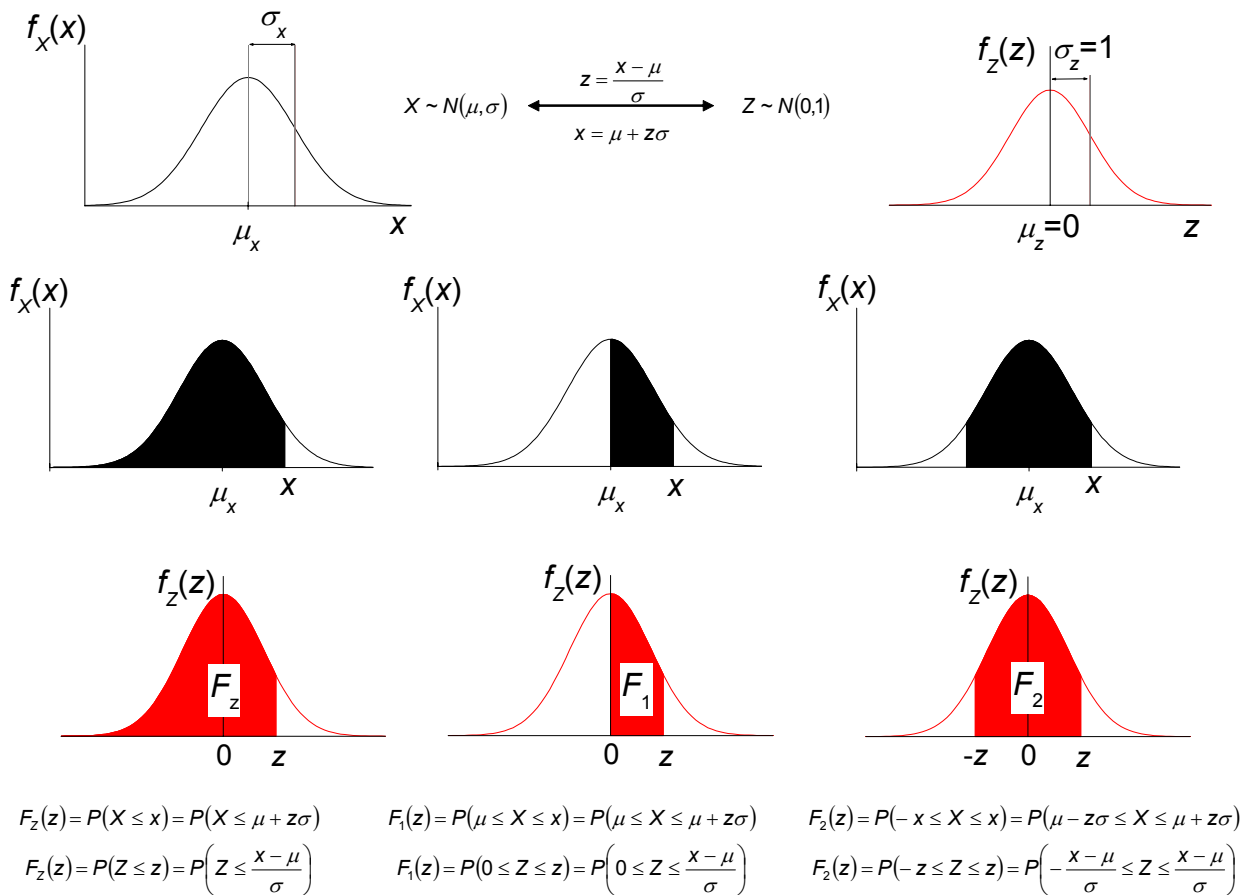


Abbildung 3.4: Umrechnungsschema zum Arbeiten mit der Tabelle der Standardnormalverteilung

Für die praktische Berechnung wird die standardisierte Dichtefunktion der Normalverteilung mit dem Erwartungswert 0 und der Varianz 1 verwendet, deren Werte den Tabellen im Anhang zu entnehmen sind. Dazu sind die realen Werte zu transformieren. Die Zufallsvariable  $X$  wird transformiert in die standardisierte Zufallsvariable  $Z$ . Jeder Wert  $x$  wird als Abstand vom Mittelwert in Vielfachen der Standardabweichung angegeben und mit  $z$  bezeichnet:

$$Z = \frac{X - \mu}{\sigma} \quad (3.2)$$

Die Funktion der Wahrscheinlichkeitsdichte der Standardnormalverteilung lautet:

$$g(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \quad (3.3)$$

Die Fläche unter der Kurve, d.h. die Wahrscheinlichkeitssumme für  $z = -\infty$  bis  $z = +\infty$  beträgt 1 bzw. 100% und wird mathematisch durch das Integral der Wahrscheinlichkeitsdichte beschrieben. Abbildung 3.4 zeigt das Umrechnungsschema zwischen Normalverteilung und Standardnormalverteilung.

### 3.2.3 Diskretisierung kontinuierlicher Zufallsvariablen

Zeiten, wie die zu untersuchende Liefertreue, sind an sich kontinuierliche Werte. Die analytische Untersuchung von kontinuierlichen Zufallsvariablen ist jedoch mathematisch sehr aufwendig bzw. nicht berechenbar. Deshalb werden die kontinuierlichen Zufallsvariablen, gemäß der im Folgenden beschriebenen Diskretisierung, in zeitdiskrete Zufallsvariablen umgewandelt.

Eine zeitkontinuierliche Zufallsvariable  $A$  wird mit ihrer Dichtefunktion  $f_A(t)$  oder ihrer Verteilungsfunktion  $F_A(t)$  beschrieben. Dabei gilt: Die Wahrscheinlichkeit dafür, dass eine Realisierung der Zufallsvariablen  $A$  in einem Intervall  $]t_1, t_2]$  liegt, ist gleich dem Integral über der Dichtefunktion innerhalb der entsprechenden Grenzen. Die Verteilungsfunktion gibt die Wahrscheinlichkeit an, mit der die Zufallsvariable einen Wert kleiner oder gleich des Variablenwertes  $t_1$  annimmt.

$$P(t_1 < A \leq t_2) = \int_{t_1}^{t_2} f_A(t) dt \quad (3.4)$$

$$P(A \leq t_1) = F_A(t_1) = \int_{-\infty}^{t_1} f_A(t) dt \quad (3.5)$$

Eine diskrete Zufallsvariable  $X$  kann nur bestimmte diskrete Werte annehmen. Sie besitzt eine Wahrscheinlichkeitsfunktion und ebenfalls eine Verteilungsfunktion. Die Wahrscheinlichkeitsfunktion gibt an, mit welcher Wahrscheinlichkeit die Zufallsvariable einen bestimmten Wert  $k$  annimmt, die Verteilungsfunktion, mit welcher Wahrscheinlichkeit sie kleiner oder gleich eines bestimmten Wertes  $k$  ist.

$$P(X = k) = f_x(k) \tag{3.6}$$

$$P(X \leq k) = F_x(k) = \sum_{i=-\infty}^k f_x(i) \tag{3.7}$$

In zeitdiskreten Modellen wird die Zeitkomponente in äquidistanten Abständen der Länge  $\Delta t$  diskretisiert. Realisierungen von Zufallsvariablen betragen somit ein ganzzahliges Vielfaches der Zeiteinheit  $\Delta t$ .

Die zeitkontinuierliche Dichtefunktion der Zufallsvariable  $A$  geht über in die zeitdiskrete Wahrscheinlichkeitsfunktion der Zufallsvariable  $X$ . Da die Betrachtungen immer zu einem Zeitpunkt  $t = 0$  Tage beginnen, nimmt die diskretisierte Zufallsvariable nur Werte größer oder gleich  $k = 1$  an.

$$f_x(k) = P(X = k) = P((k - 1) \cdot \Delta t \leq A < k \cdot \Delta t); \quad \text{mit } k = 1, 2, 3, \dots \tag{3.8}$$

In Abbildung 3.5 ist die Diskretisierung anhand zweier Beispiel mit unterschiedlicher Diskretisierungskonstante graphisch dargestellt. Je größer die Diskretisierungskonstante gewählt wird, um so größer werden die Einzelwahrscheinlichkeiten und um so geringer ist die Anzahl der Wahrscheinlichkeiten. Die Werte der einzelnen Wahrscheinlichkeiten sind somit nur in Zusammenhang mit der Diskretisierungskonstante interpretierbar.

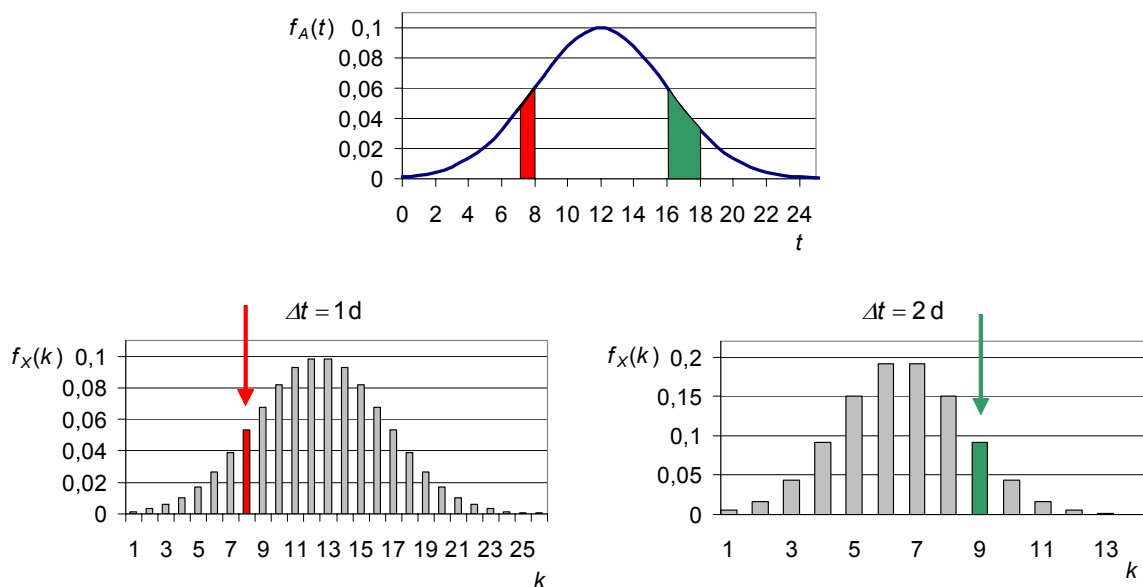


Abbildung 3.5: Diskretisierung einer Dichtefunktion mit unterschiedlichen Diskretisierungskonstanten

### 3.2.4 Faltungssatz für diskrete Zufallsvariablen

Die Addition zweier diskreter statistisch unabhängiger Zufallsvariablen ergibt wieder eine Zufallsvariable. Die dazugehörige Wahrscheinlichkeitsfunktion ergibt sich dabei aus der Faltung der beiden Wahrscheinlichkeitsfunktionen der Summanden. Die Summe

$$Z = X + Y \quad (3.9)$$

besitzt die Wahrscheinlichkeitsfunktion

$$\begin{aligned} f_z(k) &= f_x(k) \otimes f_y(k) \\ &= \sum_{j=1}^k f_x(k+1-j) \cdot f_y(j), \end{aligned} \quad (3.10)$$

wobei das Symbol „ $\otimes$ “ die diskrete Faltungsoperation repräsentiert [Tra96a].

Das Beispiel in Abbildung 3.6 zeigt die Faltung zweier Wahrscheinlichkeitsfunktionen in Vektorschreibweise und als Graphik.

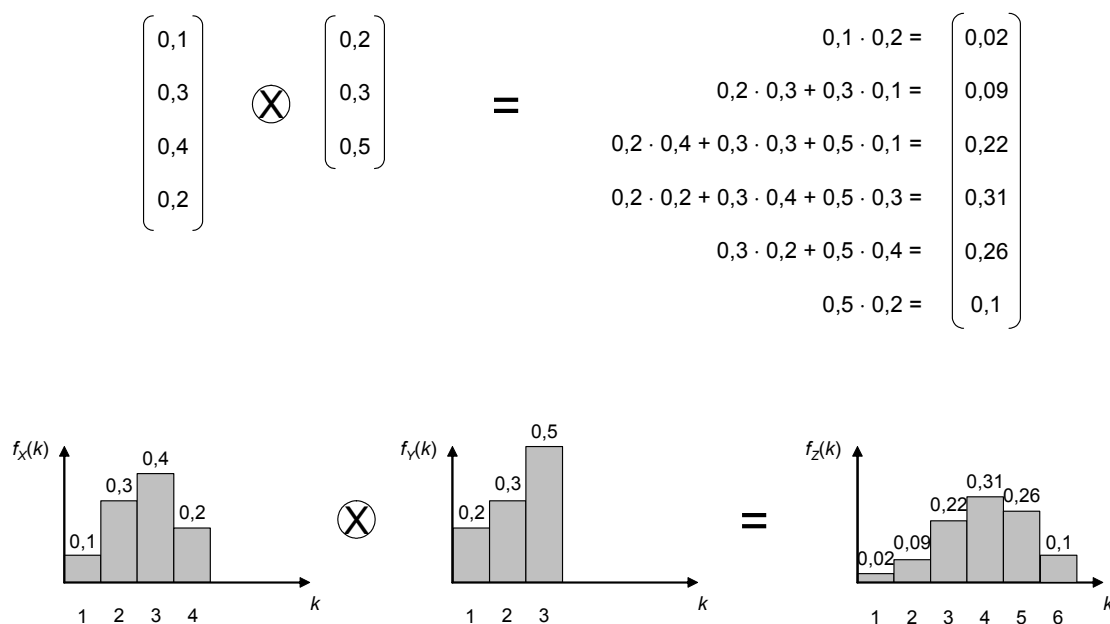


Abbildung 3.6: Faltung zweier Wahrscheinlichkeitsfunktionen

### 3.2.5 Notation

In der Statistik wird die Unterscheidung zwischen Zufallsvariable und einer ihrer möglichen Realisierungen üblicherweise dadurch dargestellt, dass die Zufallsvariable mit Großbuchstaben und die Realisierungen mit Kleinbuchstaben bezeichnet werden. Um die ohnehin schon große Vielfalt der Parameter im folgenden Kapitel nicht noch zu vergrößern

und somit nicht zu einer ungewollten Verwirrung beizutragen, wird im Folgenden von dieser Nomenklatur abgewichen. Beide Größen werden mit den gleichen Großbuchstaben bezeichnet, jedoch erhält die Zufallsvariable zur Unterscheidung einen Stern. Auch wird die Bezeichnung der an sich kontinuierlichen Zufallsvariablen nach ihrer Diskretisierung beibehalten.

### 3.3 Confirmed Line Item Performance und On-Time-Delivery

Mit Kunden werden Liefertermine vereinbart. Deshalb ist es interessant, Aussagen darüber zu treffen, ob und in welchem Maß diese Liefertermine eingehalten werden können. Um die Liefertreue zu bewerten, können Kennzahlen definiert werden, wie z.B. die Confirmed Line Item Performance *CLIP*. Abbildung 3.7 zeigt schematisch den Weg einer Bestellung, von der Einschleusung (Release) in die Fertigung bis zur Auslieferung (Delivery) beim Kunden. Dieser Weg kann in der Halbleiterproduktion in drei Stufen untergliedert sein: Frontend, Backend und Delivery [Sch00].

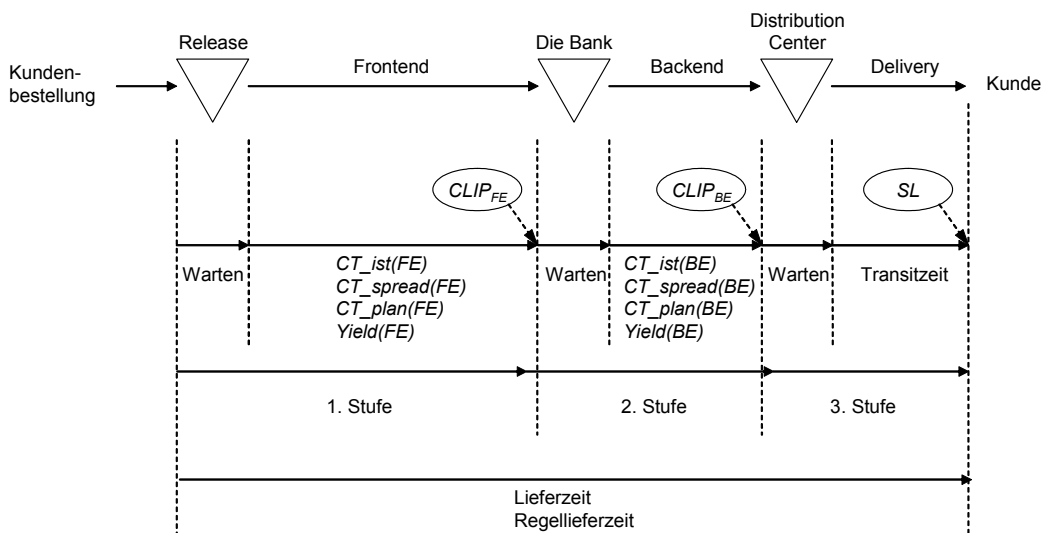


Abbildung 3.7: Weg einer Bestellung von der Einschleusung bis zur Auslieferung beim Kunden

In der ersten Stufe wird das Produkt eingeschleust und erfährt die Waferherstellung im Frontend. Dabei kann der Vertrieb die Woche der Einschleusung bestimmen. Tag und genauer Zeitpunkt der Einschleusung wird von der Produktionssteuerung festgelegt. So ist es möglich, dass ein Los bis zu sieben Tage auf seine Einschleusung innerhalb der Einschleuswoche zu warten hat. Die Waferherstellung ist ein sehr aufwendiger Prozess mit vielen Teilschritten (siehe Kapitel 2), bei der die Vorhersage der genauen Durchlaufzeit nicht möglich ist. Die tatsächliche Durchlaufzeit (Cycle Time, *CT*) und die Ausbeute müssen jedoch bestimmten Planwerten entsprechen, damit das Produkt rechtzeitig für die Weiterverarbeitung das Zwischenlager, die so genannte Die Bank, erreicht. Der  $CLIP_{FE}$  ist ein Maß für die rechtzeitige Lieferung des Frontends. Aus der Die Bank wird die Bestellung in das Backend eingeschleust, was wiederum innerhalb einer Einschleuswoche geschieht. Im Backend werden die vereinzelt Chips gebondet, gepackt und getestet. Die rechtzeitige



Lieferung an das Distribution Center wird wiederum mit einem *CLIP*-Wert, dem *CLIP<sub>BE</sub>* bewertet. Der Anteil der Bestellungenpositionen, die vollständig und rechtzeitig beim Kunden ausgeliefert werden, wird durch den Service Level *SL* ausgedrückt.

Bevor das Gesamtsystem betrachtet werden kann, werden zuerst an einem einfachen Modell die Eigenschaften und Zusammenhänge der einzelnen Parameter beispielhaft für die erste Stufe (Frontend) untersucht. Als Maß für die Liefertreue *CLIP* wird der Parameter On-Time-Delivery *OTD* definiert und eingeführt.

### 3.3.1 Einstufige Betrachtung der On-Time-Delivery

Der *CLIP* berechnet sich aus dem Verhältnis von gelieferter Menge zu bestellter Menge und wird für jedes Produkt getrennt ermittelt. Lieferrückstände aus Vorwochen sollen in folgender Betrachtung nicht berücksichtigt werden. Für die einstufige Betrachtung des Frontends setzt sich die gelieferte Menge aus der Summe der funktionierenden Chips aller Lose eines Produktes zusammen, die rechtzeitig die Die Bank erreichen. Der zu erwartende *CLIP* ergibt sich aus dem Erwartungswert der gelieferten Menge im Verhältnis zur bestellten Menge. Der Erwartungswert der gelieferten Menge wiederum ist die Summe über alle *m* Lose eines Produktes aus den jeweiligen Losgrößen *Lg*, gewichtet mit der Ausbeute (Yield, *Y*) und der Wahrscheinlichkeit  $P(L \leq 0)$  dafür, dass das Los rechtzeitig geliefert wird.

$$E(CLIP) = \frac{E(\text{gelieferte Menge})}{\text{bestellte Menge}} = \frac{\sum_{i=1}^m Y(i) \cdot Lg(i) \cdot P(L \leq 0)}{n} \quad (3.11)$$

Die Wahrscheinlichkeit für ein Los, rechtzeitig geliefert zu werden, wird durch den Ausdruck  $P(L \leq 0)$  beschrieben (Wahrscheinlichkeit für einen Lieferverzug ist kleiner oder gleich 0 Wochen; *L* bezeichnet dabei die Differenz von Planlieferwoche und tatsächlicher Lieferwoche). Für die Wahrscheinlichkeit einer rechtzeitigen Lieferung wiederum sind die drei Größen Plan-Durchlaufzeit, tatsächliche Durchlaufzeit und Tag der Einschleusung relevant. Ist die tatsächliche Durchlaufzeit größer als die Plan-Durchlaufzeit, so hängt es vom Tag der Einschleusung ab, ob sich der Auslieferungstermin verschiebt. Die Plan-Durchlaufzeit wird üblicherweise in Wochen angegeben. Die Einschleusung findet innerhalb einer durch die Planung vorgegebenen Einschleuswoche statt. Heutzutage wird in den meisten Firmen eine Wocheneinteilung verwendet, bei der der Wochenbeginn der Samstag und das Ende der Woche der Freitag ist.

In einer tagesfeinen Darstellung zeigt Abbildung 3.8 für eine Plandurchlaufzeit *CTP* (Cycle Time Planed) von vier Wochen die maximal mögliche Durchlaufzeit *CT* (Cycle Time), die, abhängig von der Einschleusung, eine rechtzeitige Lieferung ermöglicht. Wird das Los am Samstag eingeschleust, so muss seine Durchlaufzeit kleiner oder gleich 34 Tage sein, damit es in der geplanten Woche geliefert werden kann. Wird das Los an einem Sonntag eingeschleust, so muss seine Durchlaufzeit weniger oder gleich 33 Tage betragen, bei einer Einschleusung an einem Montag weniger oder gleich 32 Tage, usw.

*Anmerkung:* Zu der hier verwendeten Notation der Zufallsvariablen siehe Kapitel 3.2.

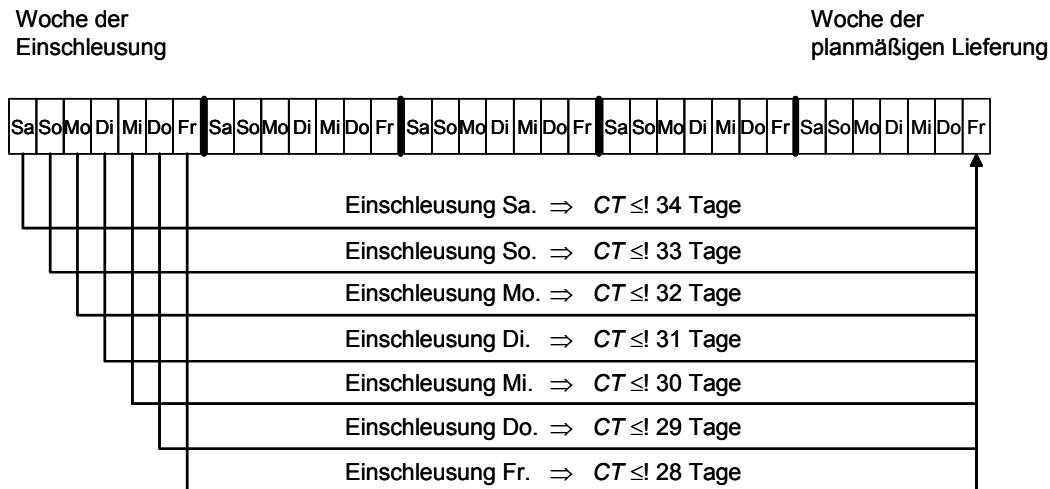


Abbildung 3.8: Notwendige Durchlaufzeiten zum Erreichen der Planlieferwoche bei einer Plandurchlaufzeit von 4 Wochen

Wird dieser Zusammenhang für beliebige Plan-Durchlaufzeiten verallgemeinert, so ergibt sich folgende Vorschrift für das Erreichen des planmäßigen Liefertermins:

$$\begin{aligned}
 \text{Einschleusung Samstag} &\Rightarrow CT^* \leq CTP + 6 \text{ d} \\
 \text{Einschleusung Sonntag} &\Rightarrow CT^* \leq CTP + 5 \text{ d} \\
 \text{Einschleusung Montag} &\Rightarrow CT^* \leq CTP + 4 \text{ d} \\
 \text{Einschleusung Dienstag} &\Rightarrow CT^* \leq CTP + 3 \text{ d} \\
 \text{Einschleusung Mittwoch} &\Rightarrow CT^* \leq CTP + 2 \text{ d} \\
 \text{Einschleusung Donnerstag} &\Rightarrow CT^* \leq CTP + 1 \text{ d} \\
 \text{Einschleusung Freitag} &\Rightarrow CT^* \leq CTP
 \end{aligned}
 \tag{3.12}$$

Sind die Wahrscheinlichkeitsfunktionen der tatsächlichen Durchlaufzeit und der Einschleusung bekannt, so kann die Wahrscheinlichkeit angegeben werden, mit der ein Los in der geplanten Lieferwoche geliefert werden kann.

Die Durchlaufzeit eines Loses des betrachteten Produktes wird mit der Zufallsvariable  $CT^*$  und deren zeitdiskreten Wahrscheinlichkeitsfunktion  $f_{CT^*}(k)$  beschrieben. Die Wahrscheinlichkeitsfunktion beschreibt die Wahrscheinlichkeit, mit der der Wert der entsprechenden Zufallsvariable innerhalb eines bestimmten Intervalls liegt.

$$f_{CT^*}(k) = P((k - 1) \cdot \Delta t < CT^* \leq k \cdot \Delta t) \quad \text{mit } k = 1, 2, 3, \dots \tag{3.13}$$

Für die entsprechende Verteilungsfunktion (auch oft als Summenfunktion bezeichnet) gilt:

$$F_{CT^*}(k) = P(CT^* \leq k \cdot \Delta t) \tag{3.14}$$

Sie beschreibt die Wahrscheinlichkeit, mit der die Realisierung der Zufallsvariable kleiner oder gleich zu einem bestimmten Wert ist.

Der Kaufmann, der die Kundenbestellung entgegen nimmt und in der Produktion einplant, kann die Einschleuswoche der entsprechenden Lose bestimmen. Der Tag und der genaue Zeitpunkt der Einschleusung wird von der Liniensteuerung festgelegt. Dabei werden produktionstechnische Aspekte zugrunde gelegt. Somit ist jeder Zeitpunkt gleich möglich, deshalb wird der Zeitpunkt der Einschleusung innerhalb der Einschleuswoche als gleichverteilt angenommen. Der Tag der Einschleusung (Release) eines Loses sei mit der Zufallsvariable  $R^*$  und deren Wahrscheinlichkeitsfunktion  $f_{R^*}(k)$  beschrieben. Für eine Diskretisierungskonstante von  $\Delta t = 1$  Tag steht  $R = 1$  für eine Einschleusung am Samstag,  $R = 2$  für Sonntag,  $R = 3$  für Montag, usw. Auch hier gelten folgende Zusammenhänge:

$$f_{R^*}(k) = P((k-1) \cdot \Delta t < R^* \leq k \cdot \Delta t) \quad \text{mit } k = 1, 2, 3, \dots \quad (3.15)$$

$$F_{R^*}(k) = P(R^* \leq k \cdot \Delta t) \quad (3.16)$$

Ebenso ist die oben definierte Differenz der tatsächlichen zur geplanten Lieferwoche  $L^*$  eine Zufallsvariable für die das gleiche gilt:

$$f_{L^*}(k) = P((k-1) \cdot \Delta t < L^* \leq k \cdot \Delta t) \quad \text{mit } k = 1, 2, 3, \dots \quad (3.17)$$

$$F_{L^*}(k) = P(L^* \leq k \cdot \Delta t) \quad (3.18)$$

Die Wahrscheinlichkeitsfunktion der Lieferwochendifferenz  $L^*$  kann zusammen mit der Plandurchlaufzeit aus den Wahrscheinlichkeitsfunktionen der Durchlaufzeit und der Einschleusung berechnet werden.

Es wird wieder eine tagesfeine Darstellung betrachtet, womit  $\Delta t = 1$  Tag gilt. Die Wahrscheinlichkeit einer Fertigstellung in der planmäßigen Lieferwoche  $P(L^* = 0)$  berechnet sich zu:

$$\begin{aligned} P(L^* = 0) = f_{L^*}(0) = f_{R^*}(1) \cdot (f_{CT^*}(CTP) + \dots + f_{CT^*}(CTP + 6 \text{ d})) \\ + f_{R^*}(2) \cdot (f_{CT^*}(CTP - 1 \text{ d}) + \dots + f_{CT^*}(CTP + 5 \text{ d})) \\ + f_{R^*}(3) \cdot (f_{CT^*}(CTP - 2 \text{ d}) + \dots + f_{CT^*}(CTP + 4 \text{ d})) \\ + f_{R^*}(4) \cdot (f_{CT^*}(CTP - 3 \text{ d}) + \dots + f_{CT^*}(CTP + 3 \text{ d})) \\ + f_{R^*}(5) \cdot (f_{CT^*}(CTP - 4 \text{ d}) + \dots + f_{CT^*}(CTP + 2 \text{ d})) \\ + f_{R^*}(6) \cdot (f_{CT^*}(CTP - 5 \text{ d}) + \dots + f_{CT^*}(CTP + 1 \text{ d})) \\ + f_{R^*}(7) \cdot (f_{CT^*}(CTP - 6 \text{ d}) + \dots + f_{CT^*}(CTP)) \end{aligned} \quad (3.19)$$

Für die Wahrscheinlichkeit, dass ein Los zu spät geliefert wird, ergibt sich folgender Ausdruck:

$$\begin{aligned}
 P(L^* > 0) = 1 - F_L(0) = & f_R(1) \cdot (1 - F_{CT}(CTP + 6 \text{ d})) \\
 & + f_R(2) \cdot (1 - F_{CT}(CTP + 5 \text{ d})) \\
 & + f_R(3) \cdot (1 - F_{CT}(CTP + 4 \text{ d})) \\
 & + f_R(4) \cdot (1 - F_{CT}(CTP + 3 \text{ d})) \\
 & + f_R(5) \cdot (1 - F_{CT}(CTP + 2 \text{ d})) \\
 & + f_R(6) \cdot (1 - F_{CT}(CTP + 1 \text{ d})) \\
 & + f_R(7) \cdot (1 - F_{CT}(CTP))
 \end{aligned} \tag{3.20}$$

Analog kann die Wahrscheinlichkeit berechnet werden, dass ein Los bereits vor der geplanten Lieferwoche seinen Durchlauf beendet hat:

$$\begin{aligned}
 P(L^* < 0) = F_L(-1) = & f_R(1) \cdot F_{CT}(CTP - 1 \text{ d}) \\
 & + f_R(2) \cdot F_{CT}(CTP - 2 \text{ d}) \\
 & + f_R(3) \cdot F_{CT}(CTP - 3 \text{ d}) \\
 & + f_R(4) \cdot F_{CT}(CTP - 4 \text{ d}) \\
 & + f_R(5) \cdot F_{CT}(CTP - 5 \text{ d}) \\
 & + f_R(6) \cdot F_{CT}(CTP - 6 \text{ d}) \\
 & + f_R(7) \cdot F_{CT}(CTP - 7 \text{ d})
 \end{aligned} \tag{3.21}$$

Für die folgenden Beispiel-Berechnungen werden normalverteilte Durchlaufzeiten angenommen. Die Wahrscheinlichkeitsfunktion der Durchlaufzeit mit einem Mittelwert von  $\mu = 28$  Tagen und einer Standardabweichung von  $\sigma = 5$  Tagen bei einer Diskretisierungskonstante von  $\Delta t = 1$  Tag zeigt Abbildung 3.9.

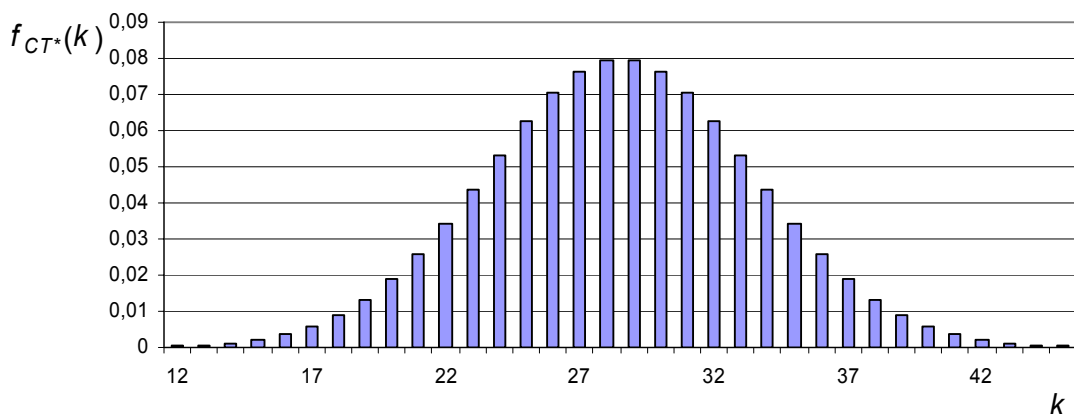


Abbildung 3.9: Wahrscheinlichkeitsverteilung der Durchlaufzeit mit einem Mittelwert von  $\mu = 28$  Tagen und einer Standardabweichung von  $\sigma = 5$  Tagen bei einer Diskretisierungskonstante von  $\Delta t = 1$  Tag

Da bei der Einschleusung ins Frontend nur die Woche, jedoch nicht der genaue Tag geplant werden kann, ist für die Einschleusung eines Loses jeder Tag gleich wahrscheinlich. Für die Wahrscheinlichkeitsfunktion der Einschleusung wird somit eine Gleichverteilung zu Grunde gelegt.

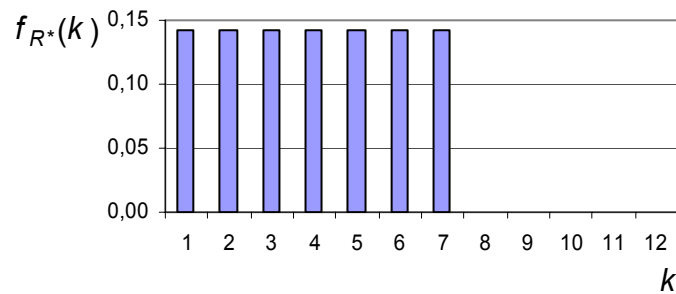


Abbildung 3.10: Wahrscheinlichkeitsverteilung der Einschleusung

In Abbildung 3.11 sind drei Beispiele für die Wahrscheinlichkeitsfunktion der Lieferwochendifferenz  $L^*$  zusammengefasst und dargestellt.

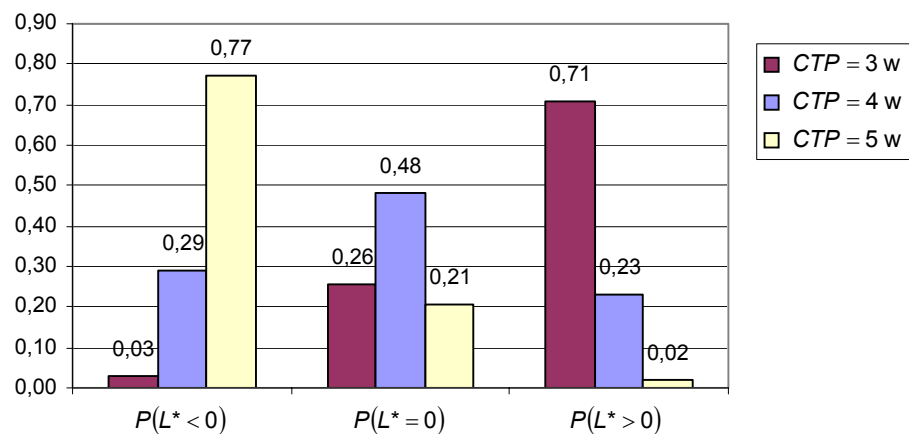


Abbildung 3.11: Wahrscheinlichkeiten zum Erreichen bzw. Über- und Unterschreiten der Lieferwoche bei gleichmäßiger Einschleusung,  $CT^* \sim N(\mu = 28 \text{ d}, \sigma = 5 \text{ d})$  und einer variierenden CTP

So erfolgt die Fertigstellung eines Loses mit einer normalverteilten Durchlaufzeit mit einem Mittelwert von 28 Tagen und einer Standardabweichung von 5 Tagen bei einer Plandurchlaufzeit von vier Wochen mit einer Wahrscheinlichkeit von 48% genau in der geplanten Lieferwoche. Mit einer Wahrscheinlichkeit von 29% wird das Los früher und mit

23% zu spät geliefert. Bei einer Plandurchlaufzeit von drei Wochen erfolgt die Fertigstellung mit einer Wahrscheinlichkeit von 26% in der geplanten Lieferwoche, bei einer Plandurchlaufzeit von fünf Wochen mit 21%. Die Unsymmetrie der berechneten Werte ist eine Folge der Diskretisierungsvorschrift und der großen Diskretisierungskonstante.

Für die CLIP-Betrachtung ist es jedoch unbedeutend, ob ein Los zu früh oder rechtzeitig lieferbar ist, in beiden Fällen findet keine CLIP-Verletzung statt. Die Wahrscheinlichkeit, keine CLIP-Verletzung zu verursachen, ergibt sich aus der Summe dieser beiden Wahrscheinlichkeiten, die sich wie folgt zusammen fassen lässt:

$$\begin{aligned}
 P(L^* \leq 0) = F_L(0) = & f_R(1) \cdot F_{CT}(CTP + 6 \text{ d}) \\
 & + f_R(2) \cdot F_{CT}(CTP + 5 \text{ d}) \\
 & + f_R(3) \cdot F_{CT}(CTP + 4 \text{ d}) \\
 & + f_R(4) \cdot F_{CT}(CTP + 3 \text{ d}) \\
 & + f_R(5) \cdot F_{CT}(CTP + 2 \text{ d}) \\
 & + f_R(6) \cdot F_{CT}(CTP + 1 \text{ d}) \\
 & + f_R(7) \cdot F_{CT}(CTP)
 \end{aligned} \tag{3.22}$$

Das gleiche Beispiel wie oben ergibt somit folgendes Bild, das in Abbildung 3.12 dargestellt ist. Ein Los, dessen Durchlaufzeit normalverteilt ist mit einem Mittelwert von 28 Tagen und einer Standardabweichung von 5 Tagen, wird bei einer Plandurchlaufzeit von vier Wochen mit einer Wahrscheinlichkeit von 77% rechtzeitig für die geplante Lieferwoche fertig gestellt. Dies ist die Summe aus zu früher und geplanter Lieferung.

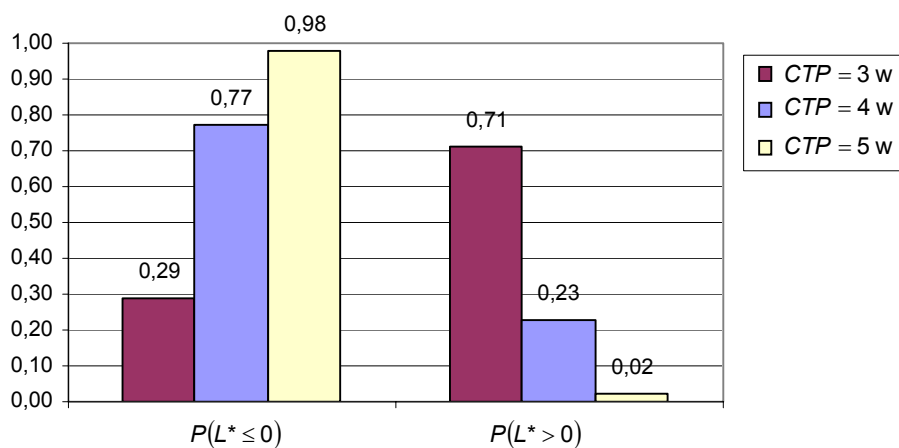


Abbildung 3.12: Wahrscheinlichkeiten zur rechtzeitigen bzw. verspäteten Lieferung bei gleichmäßiger Einschleusung für  $CT^* \sim N(\mu = 28 \text{ d}, \sigma = 5 \text{ d})$  und einer variierenden CTP

Die bisherigen Überlegungen und Rechenbeispiele erfolgten auf der Basis einer tagfeinen Einteilung der Werte. Sowohl die Einschleusung als auch Beginn und Ende der

Durchlaufzeiten können zu beliebigen Zeitpunkten erfolgen. Um genauere Ergebnisse zu erzielen ist es daher notwendig, eine feinere Diskretisierung zu verwenden.

Dazu wird zunächst die Wahrscheinlichkeit der rechtzeitigen Lieferung eines Loses verallgemeinert zu:

$$P(L^* \leq 0) = P(R^* + CT^* \leq CTP + 7 \text{ d}) \quad (3.23)$$

Dabei ist der Stichtag, zu dem eine Lieferung als rechtzeitig oder zu spät bewertet wird, der letzte Tag der geplanten Lieferwoche. Da die Zeitmessung am ersten Tag der Einschleuswoche beginnt, berechnet sich dieser Tag genau aus der Plandurchlaufzeit plus sieben Tage, also  $(CTP + 7 \text{ d})$ . Oder anders: Die Einschleusung  $R$  beschreibt die Zeitdauer von Beginn der Einschleuswoche bis zur tatsächlich erfolgten Einschleusung. Addiert mit der tatsächlichen Durchlaufzeit darf sie nicht größer als die Plandurchlaufzeit plus sieben Tage dauern, um eine rechtzeitige Lieferung sicher zu stellen. Für die Summe aus Einschleusung und Durchlaufzeit wird eine neue Zufallsvariable  $C^*$ , die Fertigstellung (Completion,  $C$ ) definiert.

$$C^* = R^* + CT^* \quad (3.24)$$

Die Wahrscheinlichkeitsfunktion der Summe zweier nicht-negativer statistisch unabhängiger Zufallsvariablen errechnet sich aus der Faltung der beiden Wahrscheinlichkeitsfunktionen. Für die Fertigstellung eines Loses ergeben sich Wahrscheinlichkeitsfunktion und Verteilungsfunktion der diskreten Zufallsvariable  $C^*$  zu:

$$f_{C^*}(k) = f_{R^*}(k) \otimes f_{CT^*}(k) = \sum_{j=1}^k f_{R^*}(k+1-j) \cdot f_{CT^*}(j) \quad (3.25)$$

$$F_{C^*}(k) = \sum_{i=1}^k f_{C^*}(i) \quad (3.26)$$

Aus der Verteilungsfunktion der Fertigstellung kann dann die Wahrscheinlichkeit für eine rechtzeitige Lieferung abgelesen werden. Sie ist gleich dem Wahrscheinlichkeitswert für den letzten Tag der Lieferwoche. Dieser Stichtag (Due Date,  $DD$ ) ergibt sich immer dann, wenn die Einschleusung über eine Woche verteilt ist, aus der Plandurchlaufzeit plus sieben Tagen.

$$DD = CTP + 7 \text{ d} \quad (3.27)$$

Die Wahrscheinlichkeit für eine rechtzeitige Lieferung (siehe Gleichung (3.23)) ist nicht die einzige Größe, die den  $CLIP$  beeinflusst. Unter der Annahme, dass die Bestellungen eines Produktes mit einem Los erfüllt werden können, d.h. es wird nur ein Los betrachtet, ergibt sich der erwartete  $CLIP$  aus der Multiplikation der Wahrscheinlichkeit einer rechtzeitigen Lieferung mit der Ausbeute.

$$E(CLIP_{1\text{Los}}) = Y \cdot P(L^* \leq 0) \quad (3.28)$$

Der Yield ist ebenfalls eine Zufallsvariable mit einer eigenen Wahrscheinlichkeitserteilung. Wird jedoch angenommen, dass der tatsächliche Yield dem geplanten Yield entspricht, also dem Wert, mit dem die zu produzierende Menge hochgerechnet wurde, um den Auftrag erfüllen zu können, so entspricht der *CLIP* der Wahrscheinlichkeit einer rechtzeitigen Lieferung. Sie ist die Wahrscheinlichkeit dafür, dass die Fertigstellung bis zum Stichtag *DD* erfolgt und kann aus der Verteilungsfunktion der Fertigstellung abgelesen werden.

$$\begin{aligned} E(\text{CLIP}_{1\text{Los}, Y=YP}) &= P(L^* \leq 0) \\ &= P(C^* \leq DD) \\ &= F_{C^*}(k = DD / \Delta t) \end{aligned} \tag{3.29}$$

Da die Normalverteilung keine Ober- und Untergrenzen hat, soll noch ein Korrekturfaktor eingefügt werden. Für den Fall, dass für die Durchlaufzeit eine zu große Standardnormalverteilung gewählt wird oder der Mittelwert sehr nah an den Grenzen des Berechnungszeitraumes (bzw. am Nullpunkt) liegt, ist die Summe aller Wahrscheinlichkeiten der Wahrscheinlichkeitsfunktion kleiner als 1. Damit erfüllt sie nicht mehr die Voraussetzung einer Wahrscheinlichkeitsfunktion. Die daraus berechneten Werte der Wahrscheinlichkeitsfunktion der Fertigstellung sind dann ebenso zu gering, womit die Wahrscheinlichkeit einer rechtzeitigen Lieferung auch zu gering ausfällt.

Werden die Werte der Wahrscheinlichkeitsfunktion bzw. der Verteilungsfunktion auf den höchsten Wert der Verteilungsfunktion normiert, so ergibt sich wieder eine korrekte Wahrscheinlichkeitsfunktion, deren Summe aller Werte gleich 1 beträgt. Ebenso beträgt der größte Wert der Verteilungsfunktion gleich 1.

Trotz Korrekturfaktor muss natürlich immer darauf geachtet werden, dass das der Rechnung zu Grunde liegende Zeitintervall groß genug ist, damit die Wahrscheinlichkeitsfunktionen am Rande nicht abgeschnitten werden.

Mit den oben getroffenen Annahmen und Herleitungen sowie dem Korrekturfaktor wird die On-Time-Delivery (*OTD*) als Wahrscheinlichkeit für eine rechtzeitige Lieferung definiert.

$$\text{OTD} = \frac{F_{C^*}(k = DD / \Delta t)}{F_{C^*}(k_{\max})} \tag{3.30}$$

Abbildung 3.13 zeigt den Zusammenhang der Zeitgrößen an einem Beispiel. Die Zeitachse ist kontinuierlich dargestellt, um die Übersicht zu vereinfachen. Für die Berechnungen wird sie mittels der Diskretisierungskonstante in äquidistante Bereiche eingeteilt.

Die Einschleusung beginnt mit dem Zeitpunkt Null der Einschleuswoche. Sie kann maximal sieben Tage betragen. Der Stichtag der Fertigstellung ist deshalb sieben Tage länger als die Plandurchlaufzeit. Findet die Einschleusung z.B. zum Zeitpunkt von drei Tagen nach Beginn der Einschleuswoche statt und dauert die Durchlaufzeit der Fertigung 27 Tage, so erfolgt die Fertigstellung zum Zeitpunkt von 30 Tagen nach Beginn der Einschleuswoche. Einschleusung, Stichtag und Fertigstellung sind Zeitgrößen, die immer vom Zeitpunkt Null der Einschleuswoche ab erfasst werden.



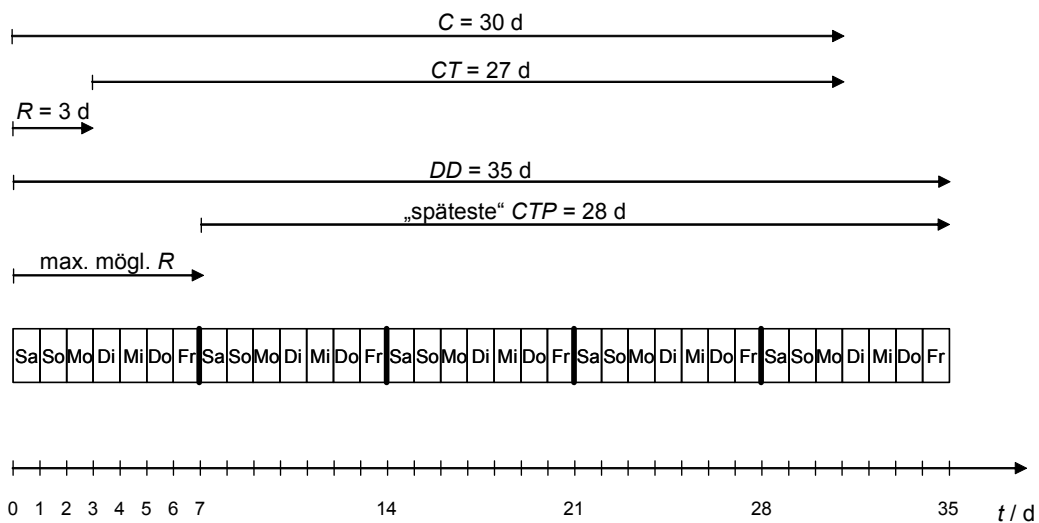


Abbildung 3.13: Beispiel von zur Ermittlung der OTD notwendigen Zeitgrößen

Als Diskretisierungskonstante bietet sich  $\Delta t = 0,001$  Tage an. 0,001 Tage entsprechen 1,44 Minuten. Eine feinere Diskretisierungskonstante wäre nur dann sinnvoll, wenn die erfassten Daten ebenfalls in so feiner Auflösung vorhanden wären. Durch teilweise manuelle Zeiterfassung ist dies nicht gewährleistet. Darüber hinaus erfolgt bei einer weiteren Verfeinerung zu 0,0001 Tagen lediglich eine Veränderung der OTD in der fünften Nachkommastelle (siehe Tabelle 3.1). Da CLIP-Werte nur in voller Prozentzahl angegeben werden, ist eine Diskretisierungskonstante von  $\Delta t = 0,001$  Tage ausreichend.

$\Delta t$	OTD
1 Tag	0,77059
0,1 Tage	0,74423
0,01 Tage	0,74154
0,001 Tage	0,74127
0,0001 Tage	0,74124

Tabelle 3.1: On-Time-Delivery für  $CT^* \sim N(\mu = 28 \text{ d}, \sigma = 5 \text{ d})$ ,  $CTP = 4$  w mit unterschiedlicher Diskretisierungskonstante

Von der sonst für diskrete Wahrscheinlichkeitsfunktionen üblichen Darstellung in Säulendiagrammen wird, um die Übersichtlichkeit der Graphik zu verbessern, abgesehen. Stattdessen werden Polygonzüge verwendet, deren verbundene Punkte horizontal am linken Rand der Klassengrenze ausgerichtet sind (siehe Abbildung 3.14). Durch die Verwendung kleiner Schrittgrößen ist die Verschiebung allerdings nicht merklich. Um die Darstellungen unabhängig von der gewählten Zeiteinheit  $\Delta t$  ablesen zu können, wird die Abszisse mit der entsprechenden kontinuierlichen Zeit  $t$  beschriftet. Da unterschiedliche Zeitschritte  $\Delta t$  zu unterschiedlich großen Wahrscheinlichkeiten führen, muss die jeweilige Diskretisierungsgröße bei der Interpretation der graphischen Darstellungen berücksichtigt werden.

Bei der Interpretation der einzelnen Werte einer Wahrscheinlichkeitsfunktion ist zu beachten, dass ihre Höhe von der Schrittgröße, also der Diskretisierungskonstante abhängt. Vergleiche der Höhe von Wahrscheinlichkeitsfunktionen, die auf unterschiedliche Schrittgrößen beruhen, sind somit nur durch Umrechnung möglich. Verteilungsfunktionen sind jedoch unabhängig von der gewählten Schrittgröße. Sie werden nur in der Glätte der Funktion beeinflusst.

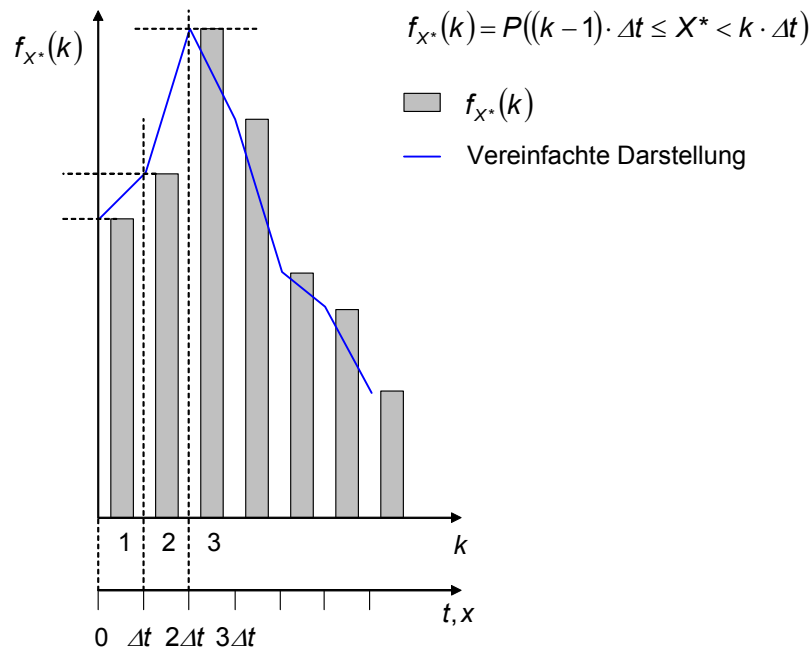


Abbildung 3.14: Vereinfachte Darstellung der diskretisierten Wahrscheinlichkeitsfunktion

Eine Darstellung der so ermittelten Fertigstellung und On-Time-Delivery zeigt Abbildung 3.15. Erfolgt die Einschleusung  $R^*$  gleichverteilt über die Einschleuswoche und ist die Durchlaufzeit  $CT^*$  normalverteilt mit  $CT^* \sim N(\mu = 43, \sigma = 7)$ , so ergibt sich daraus die Wahrscheinlichkeitsverteilung für die Fertigstellung  $C^*$ . Eine Plandurchlaufzeit von  $CTP = 49$  d ergibt einen Stichtag von  $DD = 56$  d. Bis zu diesem Termin ist das Los mit einer Wahrscheinlichkeit von  $OTD = 90,38\%$  fertig gestellt. Dieser Wert entspricht der Fläche unter der Kurve der Fertigstellung  $C^*$ , „links“ von der Due Date  $DD$ .

Für unterschiedliche Wahrscheinlichkeitsfunktionen der Durchlaufzeit sind in Tabelle 3.2 die Werte der On-Time-Delivery angegeben und in Abbildung 3.16 dargestellt. Sowohl steigende Mittelwerte als auch steigende Standardabweichungen der Durchlaufzeit verringern die rechtzeitige Lieferung, was auch intuitiv verständlich erscheint.

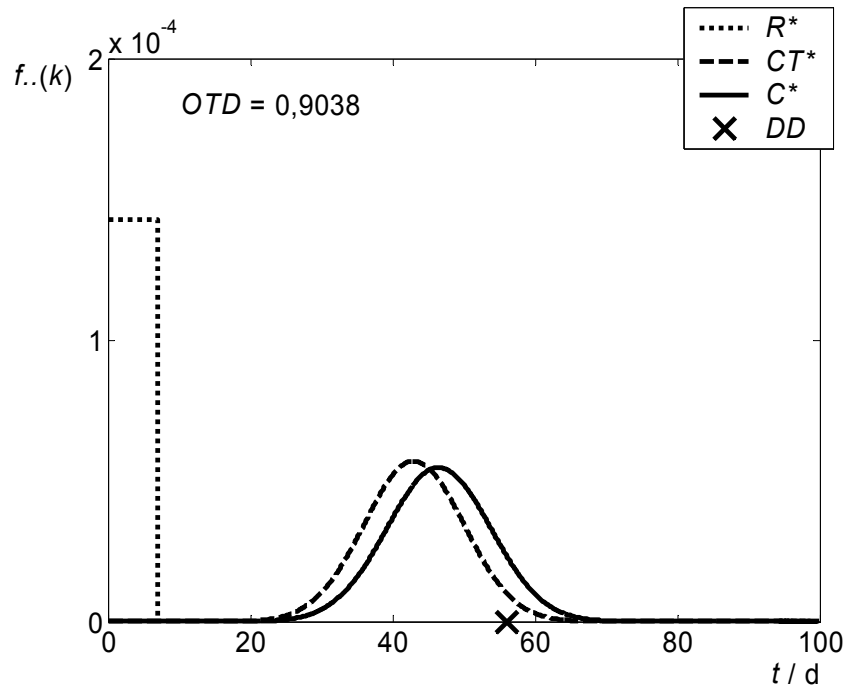


Abbildung 3.15: Wahrscheinlichkeitsfunktionen von Einschleusung  $E^*$ , Durchlaufzeit  $CT^*$  und Fertigstellung  $C^*$  mit On-Time-Delivery für  $CT^* \sim N(\mu = 43 \text{ d}, \sigma = 7 \text{ d})$ ,  $CTP = 49 \text{ d}$

$\mu/d \setminus \sigma/d$	5,0	5,5	6,0	6,5	7,0	7,5	8,0	8,5	9,0
41,0	0,9837	0,9753	0,9654	0,9544	0,9428	0,9306	0,9183	0,9059	0,8937
41,5	0,9795	0,9698	0,9589	0,9469	0,9344	0,9216	0,9087	0,8960	0,8834
42,0	0,9744	0,9635	0,9514	0,9385	0,9252	0,9117	0,8984	0,8852	0,8725
42,5	0,9682	0,9561	0,9429	0,9291	0,9150	0,9010	0,8872	0,8738	0,8608
43,0	0,9610	0,9475	0,9332	0,9185	0,9038	0,8893	0,8751	0,8615	0,8484
43,5	0,9524	0,9377	0,9223	0,9069	0,8916	0,8766	0,8622	0,8485	0,8354
44,0	0,9424	0,9265	0,9102	0,8940	0,8782	0,8630	0,8485	0,8346	0,8216
44,5	0,9309	0,9138	0,8967	0,8799	0,8638	0,8484	0,8338	0,8200	0,8071
45,0	0,9176	0,8995	0,8817	0,8646	0,8482	0,8328	0,8182	0,8046	0,7919

Tabelle 3.2: On-Time-Delivery für normalverteilte Durchlaufzeiten mit unterschiedlichem Mittelwert und Standardabweichung bei einer Plandurchlaufzeit von 49 Tagen

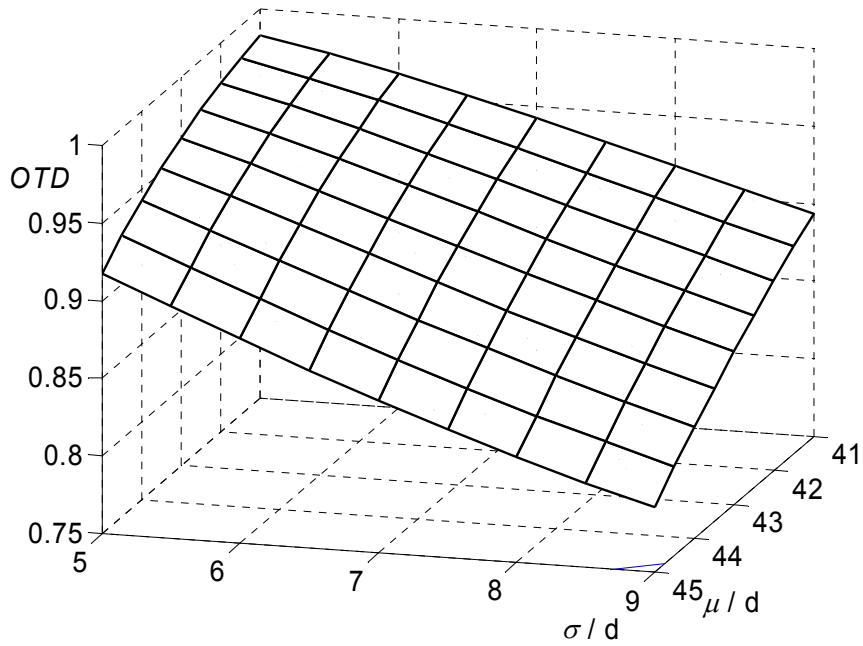


Abbildung 3.16: On-Time-Delivery für normalverteilte Durchlaufzeiten mit unterschiedlichem Mittelwert und Standardabweichung bei einer Plandurchlaufzeit von 49 Tagen

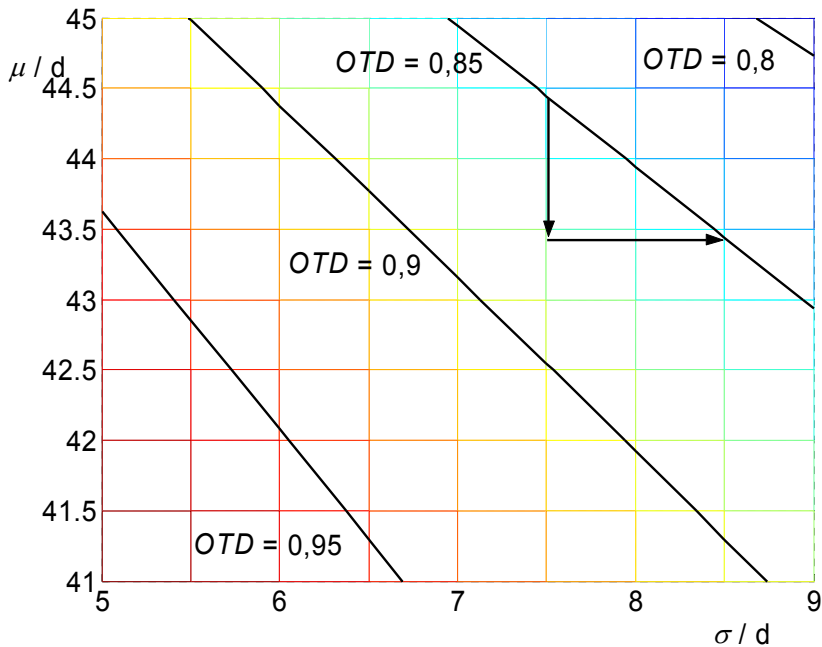


Abbildung 3.17: Linien gleicher On-Time-Delivery für normalverteilte Durchlaufzeiten mit unterschiedlichem Mittelwert und Standardabweichung bei einer Plandurchlaufzeit von 49 Tagen

Die Darstellung der Höhenlinien in Abbildung 3.17 erleichtert den Vergleich der Effekte durch Änderung des Mittelwertes und Änderung der Standardabweichung. So führt z.B. im Punkt ( $\mu = 44,5$  d,  $\sigma = 7,5$  d) die Erhöhung der Standardabweichung um einen Tag zu einer OTD-Verschlechterung, die aber durch die gleichzeitige Verringerung des Mittelwertes um einen Tag kompensiert werden kann. Produktionstechnisch bedeutet dies, dass durch konstanten Prozessfluss (bessere Steuerung) höhere Durchlaufzeiten in Kauf genommen werden können.

Je nach Arbeitspunkt kann entweder die Veränderung des Mittelwertes oder die Veränderung der Standardabweichung eine größere Auswirkung auf die On-Time-Delivery bewirken. Die getrennte Darstellung in Abhängigkeit von jeweils dem Mittelwert und der Standardabweichung der Durchlaufzeit lässt diesen Sachverhalt besser erkennen. Abbildung 3.18 zeigt dies für das oben genannte Beispiel. Die größte Auswirkung wird dort erzielt, wo die Steigung der Kurven am größten ist. Das ist hier bei der geringsten Standardabweichung zusammen mit dem höchsten Mittelwert bei Änderung der Standardabweichung der Fall. Bei großer Standardabweichung und großem Mittelwert jedoch hat die Änderung des Mittelwertes eine größere Auswirkung als die der Standardabweichung. Wird die Änderung der On-Time-Delivery allerdings bezogen auf die relative Änderung der Eingangsgrößen betrachtet, so erreicht die relative Änderung des Mittelwertes immer einen größeren relativen Effekt als die der Standardabweichung (siehe Tabelle 3.3 und Tabelle 3.4).

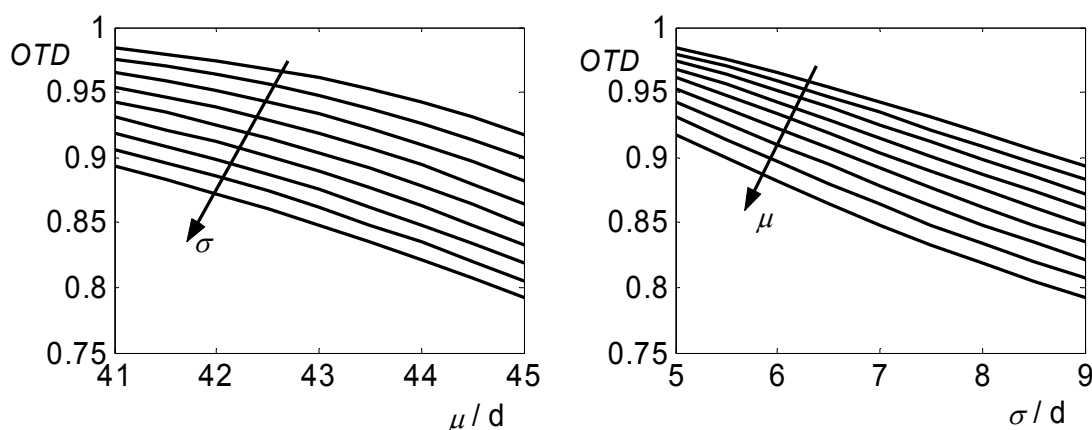


Abbildung 3.18: OTD in Abhängigkeit von Mittelwert oder Standardabweichung der Durchlaufzeit bei einer Plandurchlaufzeit von 49 Tagen

$\sigma / d$	$\mu_1 / d$	$\mu_2 / d$	$OTD_1$	$OTD_2$	Absolute Änderung $OTD_2 - OTD_1$	Relative Änderung $\frac{(OTD_2 - OTD_1) / OTD_1}{(\mu_2 - \mu_1) / \mu_1}$
5,0	41,0	41,5	0,9837	0,9795	-0,0042	-0,3501
9,0	44,5	45,0	0,8071	0,7919	-0,0152	-1,6761

Tabelle 3.3: Auswirkung auf die OTD bei Veränderung des Mittelwertes der Durchlaufzeit

$\mu / d$	$\sigma_1 / d$	$\sigma_2 / d$	$OTD_1$	$OTD_2$	Absolute Änderung $OTD_2 - OTD_1$	Relative Änderung $\frac{(OTD_2 - OTD_1)}{OTD_1}$ $\frac{(\sigma_2 - \sigma_1)}{\sigma_1}$
41,0	5,0	5,5	0,9837	0,9753	-0,0084	-0,0854
45,0	8,5	9,0	0,8046	0,7919	-0,0127	-0,2683

Tabelle 3.4: Auswirkung auf die  $OTD$  bei Veränderung der Standardabweichung der Durchlaufzeit

### 3.3.2 Näherung zur Abschätzung der On-Time-Delivery

Die On-Time-Delivery wird aus der Wahrscheinlichkeitsfunktion ermittelt, die aus der Faltung der Rechteckfunktion der Einschleusung und der Normalverteilung der Durchlaufzeit hervorgeht. Ihre Form sieht der einer Normalverteilung sehr ähnlich. Dies gilt besonders dann, wenn die Breite der Wahrscheinlichkeitsfunktion der Durchlaufzeit ( $2\sigma$ ) groß genug gegenüber einer Woche ist. In Abbildung 3.19 ist für zwei Fälle die Wahrscheinlichkeitsfunktion der Fertigstellung und eine Normalverteilung mit gleichem Mittelwert und gleicher Standardabweichung dargestellt. Während im linken Bild die Wahrscheinlichkeitsverteilung der Fertigstellung  $C$  deutliche Unterschiede zu einer Normalverteilung zeigt, sind im rechten Bild die Unterschiede so gering, dass sie optisch nicht mehr erkennbar sind.

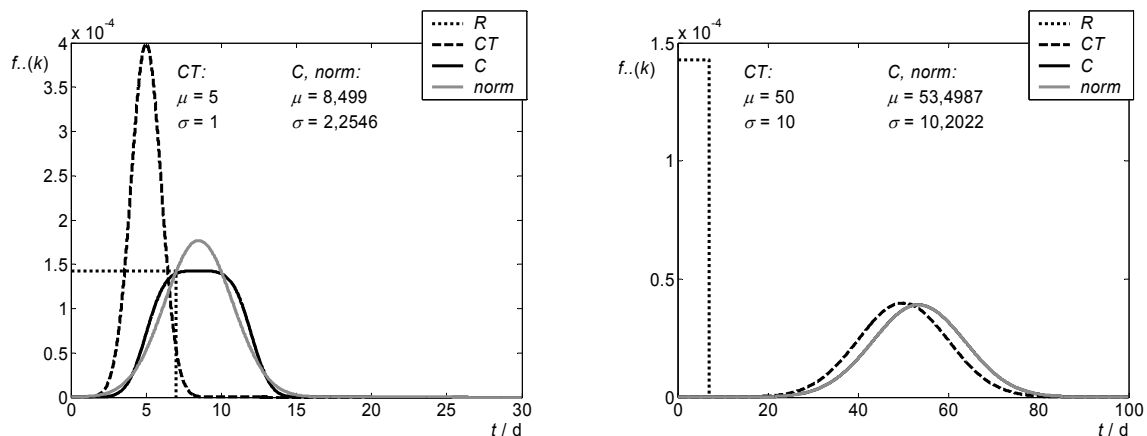


Abbildung 3.19: Vergleich der Wahrscheinlichkeitsfunktionen von Fertigstellung und einer Normalverteilung mit gleichen Parametern

Wäre die Einschleusung nicht gleichverteilt sondern normalverteilt, dann wäre die Fertigstellung ebenfalls normalverteilt, da die Summe zweier unabhängiger normalverteilter Zufallsvariablen eine normalverteilte Zufallsvariable ergibt, wobei sich Mittelwerte und Varianzen addieren.

Wird die gleichverteilte Einschleusung mit ihrem Mittelwert von 3,5 Tgen und ihrer Standardabweichung von 4,08 Tagen an eine Normalverteilung mit einem gleichen Mittelwert von 3,5 Tagen und einer gleichen Standardabweichung von 4,08 Tagen angenähert, kann eine einfache Abschätzung für die Fertigstellung eines Loses und seiner On-Time-Delivery

ermittelt werden. Die Zufallsvariable Näherung Fertigstellung  $NC^*$  ergibt sich somit aus der Summe der Zufallsvariablen Näherung Einschleusung  $NR^*$  und Durchlaufzeit  $CT^*$ . Sie ist normalverteilt, wobei sich Mittelwert und Varianz aus den Mittelwerten und Varianzen der Einschleusung und der Durchlaufzeit ergeben.

$$NC^* = NR^* + CT^* \quad (3.31)$$

$$NC^* \sim N(3,5 d + \mu, \sqrt{4,08 d + \sigma}) \quad (3.32)$$

Die Näherung für die On-Time-Delivery entspricht dem Wert der Verteilungsfunktion der Näherung Fertigstellung für den Zeitpunkt des Stichtages  $DD$ .  $F_z$  ist die Verteilungsfunktion der Standardnormalverteilung. Durch die Umrechnung in den  $z$ -Wert kann das Ergebnis leicht in Tabellen der Standardnormalverteilung abgelesen werden. Eine Diskretisierung ist nicht notwendig; auf eine Stetigkeitskorrektur kann ebenso verzichtet werden.

$$OTD = P(C^* \leq DD) \approx F_{NC^*}(t = DD) = F_z \left( z = \frac{DD - \mu - 3,5 d}{\sqrt{\sigma^2 + 4,08 d}} \right) \quad (3.33)$$

Die Beispielrechnungen in Tabelle 3.5 zeigen, dass die Abschätzung zu Ergebnissen führt, die sehr nah an den Ergebnissen des zeitdiskreten Modells liegen bzw. in manchen Fällen sogar identisch sind.

In allen Fällen, in denen die Breite der Wahrscheinlichkeitsfunktion der Durchlaufzeit groß genug gegenüber der Einschleuswoche ist, ergibt die Abschätzung identische Werte wie das zeitdiskrete Modell (ab Nr. 12). Für Durchlaufzeiten mit geringen Standardabweichungen ( $\sigma < 9$ ) können die Abschätzungen als Richtwerte verwendet werden. Dabei kann nicht gesagt werden, ob die Schätzung zu gering, zu groß oder zufällig richtig (Nr. 2, 6, 7) ausfällt.

Nr.	$\mu / d$	$\sigma / d$	$DD / d$	$z$	$F_{NC^*}(DD)$	<b>OTD</b>
1.	5	1	10	0,67	0,7486	0,7131
2.	5	1	11,8	1,46	0,9279	0,9276
3.	5	1	13	2,00	0,9772	0,9881
4.	10	3	20	1,8	0,9641	0,9643
5.	41	5	59	2,69	0,9964	0,9965
6.	41	5	63	3,43	0,9997	0,9997
7.	44	5,5	59	1,96	0,9752	0,9753
8.	35	6	56	2,76	0,9971	0,9972
9.	41	6	56	1,82	0,9656	0,9654
10.	41	7	56	1,58	0,9429	0,9428
11.	41	8	56	1,39	0,9177	0,9183
12.	41	9	59	1,57	0,9420	0,9420
13.	45	9	59	1,14	0,8725	0,8725
14.	45	9	63	1,576	0,9420	0,9420
15.	35	16	56	1,09	0,8611	0,8591
16.	41	16	56	0,71	0,7621	0,7610

Tabelle 3.5: Abschätzung und Berechnung der On-Time-Delivery

### 3.4 On-Time-Delivery mit Due Date Weekly und Release Backend

Um das System im Ganzen untersuchen zu können und die Zusammenhänge der einzelnen Parameter darzustellen, ist ein mathematisches Modell notwendig, das die vollständige Produktion beschreibt.

Es wird wieder die Produktion eines einzelnen Loses betrachtet. Dieses Los soll vollständig unabhängig von anderen möglichen Losen des gleichen Produktes bearbeitet werden.

Im Frontend ist es üblich, Produkte, deren Bestellung öfters zu erwarten ist, auch ohne Bestellung auf Vorrat zu produzieren und in der Die Bank auf Lager zu legen. Derartige Replishment-Vorgänge müssen nicht gesondert betrachtet werden, da die Plan-Durchlaufzeiten sowohl eine Bevorratung wie auch eine Entnahme bevorrateter Chips im entsprechenden Fall bereits berücksichtigen.

Prinzipiell ist es natürlich auch immer möglich, durch Veränderungen von Daten Performance-Kennzahlen zu manipulieren. So könnte durch die Verschiebung des geplanten Liefertermins der CLIP-Wert erhöht werden, auch wenn der eigentliche Liefertermin nicht



erreicht wurde. Das zeigt, dass durch Performance-Indikatoren nicht immer zwangsweise das eigentliche Geschehen erfasst wird. Die Verbesserung der Aussagekraft von Performance-Parametern soll jedoch nicht Inhalt dieser Arbeit sein.

In den drei Stufen der Produktion können fünf unterschiedliche Werte der On-Time-Delivery ermittelt werden. Um die Performance von Frontend und Backend zu bewerten werden  $CLIP_{FE}$  und  $CLIP_{BE}$  ermittelt. Für den Fortgang der Produktion im Hinblick auf den Kundenendtermin ist es jedoch sinnvoll, nach dem Backend auch die Übereinstimmung mit dem vorgegebenen Gesamtzeitplan zu bewerten. Der Service Level  $SL$  bewertet die Liefertreue gegenüber dem Kunden. Darin spiegelt sich die gesamte Performance aller drei Stufen. Jedoch ist es auch sinnvoll, die Leistung der Auslieferung unabhängig von der Vorgeschichte zu betrachten.

So werden die unabhängigen Leistungen der drei Stufen mit On-Time-Delivery Frontend, On-Time-Delivery Backend und On-Time-Delivery Delivery bezeichnet. Die entsprechenden Performanewerte nach jeder Stufe, bezogen auf das Gesamtergebnis, werden durch die On-Time-Delivery Manufacturing und die On-Time-Delivery Over All beschrieben (siehe Abbildung 3.20).

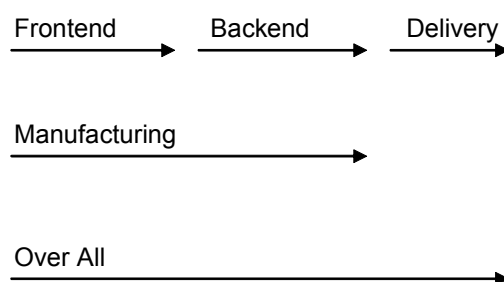


Abbildung 3.20: Die fünf Werte der On-Time-Delivery

### 3.4.1 Frontend

Der Produktionsablauf im Frontend entspricht der oben beschriebenen einstufigen Betrachtung der On-Time-Delivery. Deshalb sollen hier nur noch die Rechenvorschriften des diskreten Modells zusammengefasst werden.

Die Einschleusung ins Frontend wird durch die Zufallsvariable  $Rf^*$  mit gleicher Wahrscheinlichkeit der Einschleusung zu jedem diskreten Zeitpunkt der Einschleuswoche beschrieben. Ihre Wahrscheinlichkeitsfunktion ergibt sich zu:

$$f_{Rf^*}(k) = \begin{cases} \frac{1}{7 d/\Delta t} & , \quad 1 \leq k < 7 d/\Delta t \\ 0 & , \quad k \geq 7 d/\Delta t \end{cases} \quad \text{mit } k = 1, 2, 3, \dots \quad (3.34)$$

Die Durchlaufzeit der Lose eines Produktes im Frontend  $CTf^*$  wird als normalverteilt angenommen, mit einem Mittelwert  $\mu f$  und einer Standardabweichung  $\sigma f$ . Ihre Wahrscheinlichkeitsfunktion ergibt sich somit zu:

$$f_{CTf^*}(k) = \int_{(k-1) \cdot \Delta t}^{k \cdot \Delta t} g(t) dt = G(k \cdot \Delta t) - G((k-1) \cdot \Delta t) \quad (3.35)$$

Dabei steht  $g(t)$  für die Dichtefunktion der Gaußschen Normalverteilung und  $G(t)$  für ihre Verteilungsfunktion.

$$g(t) = \frac{1}{\sigma f \sqrt{2\pi}} \cdot e^{-\left(\frac{t - \mu f}{\sigma f}\right)^2} \quad (3.36)$$

$$G(t) = \int_{-\infty}^t g(\xi) d\xi \quad (3.37)$$

Die Zufallsvariable Fertigstellung Frontend  $Cf^*$  (Completion Frontend) ergibt sich aus der Addition der Zufallsvariablen für den Tag der Einschleusung  $Rf^*$  und der Durchlaufzeit  $CTf^*$ . Ihre Wahrscheinlichkeitsfunktion berechnet sich somit aus der Faltung der beiden Wahrscheinlichkeitsfunktionen.

$$Cf^* = Rf^* + CTf^* \quad (3.38)$$

$$f_{Cf^*}(k) = f_{Rf^*}(k) \otimes f_{CTf^*}(k) \quad (3.39)$$

Der Stichtag für den Liefertermin ergibt sich aus der Plandurchlaufzeit plus sieben Tage aufgrund der siebentägigen Einschleuswoche. Wenn die Plandurchlaufzeit in Wochen angegeben wird, ist ihr Wert immer ein Vielfaches von sieben Tagen.

$$DDf = CTPf + 7 \text{ d} \quad (3.40)$$

Die Wahrscheinlichkeit für das rechtzeitige Erreichen der Die Bank, also die On-Time-Delivery des Frontends kann in der Verteilungsfunktion der Fertigstellung Frontend am Zeitpunkt des Stichtages abgelesen werden. Abbildung 3.21 zeigt ein Beispiel.

$$OTDf = P(Cf^* \leq DDf) = \frac{F_{Cf^*}(k = DDf / \Delta t)}{F_{Cf^*}(k_{\max})} \quad (3.41)$$

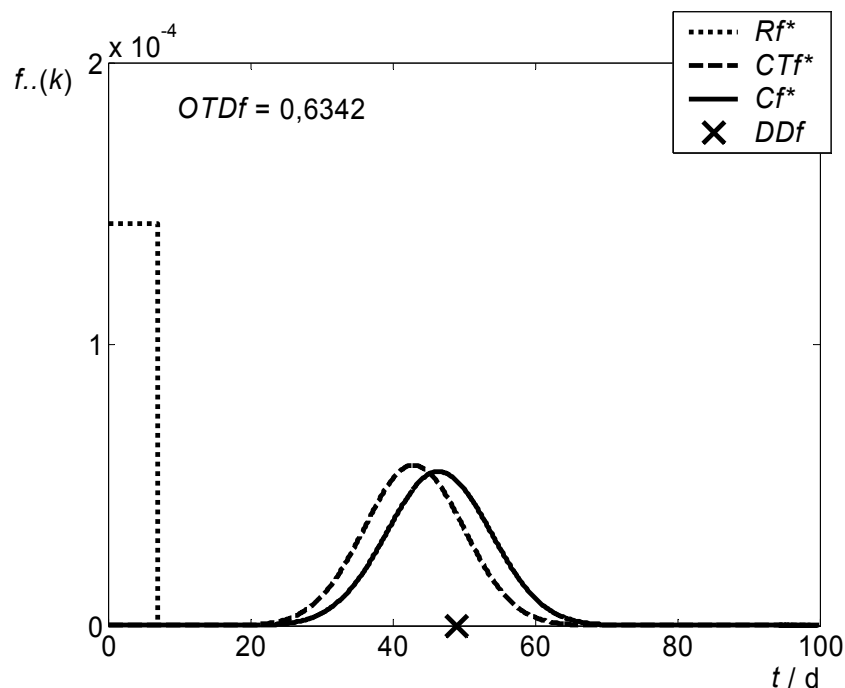


Abbildung 3.21: Berechnung der OTDf für  $CTf^* \sim N(\mu f = 43 \text{ d}, \sigma f = 7 \text{ d})$  und  $CTPf = 42 \text{ d}$

### 3.4.2 Backend

Da sich in der Praxis bei mehrstufigen Produktionen die Verantwortlichen der einzelnen Stufen nicht gerne die Verzögerungen der Vorstufe anrechnen lassen wollen, ist jede Stufe bemüht, ihre eigene Sollerfüllung nachzuweisen. Daher ist es notwendig, sowohl kombinierte als auch unabhängige Performance-Parameter zu ermitteln. Der  $CLIP_{BE}$  bewertet die Performance des Backends unabhängig von der Vorgeschichte der Produktion. D.h. alle Chips, die nicht rechtzeitig in der Die Bank zur Verfügung standen, werden aus der bestellten Menge zur Berechnung des  $CLIP_{BE}$  herausgenommen. In unserer Modellrechnung ist eine derartige Umrechnung nicht notwendig, weil die Kenntnis der Durchlaufzeit eine einfache Berechnung des unabhängigen  $CLIPs$  erlaubt. Zur Ermittlung der Performance des Backends (Backend Only) werden deshalb, analog zur Einschleusung Frontend, gleichmäßig verteilte Wahrscheinlichkeiten der Einschleusung innerhalb der Einschleuswoche verwendet. Die Wahrscheinlichkeitsfunktion Release Backend Only ergibt sich somit zu:

$$f_{Rbo^*}(k) = \begin{cases} \frac{1}{7 \text{ d}/\Delta t} & , \quad 1 \leq k < 7 \text{ d}/\Delta t \\ 0 & , \quad k \geq 7 \text{ d}/\Delta t \end{cases} \quad \text{mit } k = 1, 2, 3, \dots \quad (3.42)$$

Die On-Time-Delivery Backend wird ebenso analog zur On-Time-Delivery Frontend ermittelt. Der Stichtag  $DDb$  ergibt sich aus der Plandurchlaufzeit des Backends plus sieben Tage. Die Plandurchlaufzeit des Backends beträgt wie die Plandurchlaufzeit des Frontends immer ein Vielfaches von sieben Tagen.

$$DDb = CTPb + 7 \text{ d} \quad (3.43)$$

Die Durchlaufzeit der Lose eines Produktes im Backend  $CTb^*$  wird wiederum als normalverteilt angenommen mit einem Mittelwert  $\mu b$  und einer Standardabweichung  $\sigma b$ .

$$\begin{aligned} f_{CTb^*}(k) &= \int_{(k-1) \cdot \Delta t}^{k \cdot \Delta t} g(t) \\ &= G(k \cdot \Delta t) - G((k-1) \cdot \Delta t) \end{aligned} \quad (3.44)$$

$G(t)$  und  $g(t)$  stehen für die Verteilungsfunktion und die Dichtefunktion der Gaußschen Normalverteilung.

Die Wahrscheinlichkeitsfunktion der Fertigstellung für die alleinige Betrachtung des Backends (Backend Only) ergibt sich aus der Faltung der optimalen Einschleusung und der Durchlaufzeit.

$$f_{Cbo^*}(k) = f_{Rbo^*}(k) \otimes f_{CTb^*}(k) \quad (3.45)$$

Die On-Time-Delivery Backend ist der Funktionswert der Verteilungsfunktion der Fertigstellung Backend Only für den Zeitpunkt der Fälligkeit Backend.

$$OTDb = P(Cbo^* \leq DDb) = \frac{F_{Cbo^*}(k = DDb / \Delta t)}{F_{Cbo^*}(k_{\max})} \quad (3.46)$$

Das Beispiel in Abbildung 3.22 zeigt eine typisch geringe Durchlaufzeit des Backends. Mit einem Mittelwert von sieben Tagen und einer Standardabweichung von zwei Tagen ergibt die Durchlaufzeit mit einer (idealen) gleichverteilten Einschleusung und einer Plandurchlaufzeit von sieben Tagen eine On-Time-Delivery von  $OTDb = 88,6\%$ . Der Stichtag (Due Date) liegt dabei bei 14 Tagen.

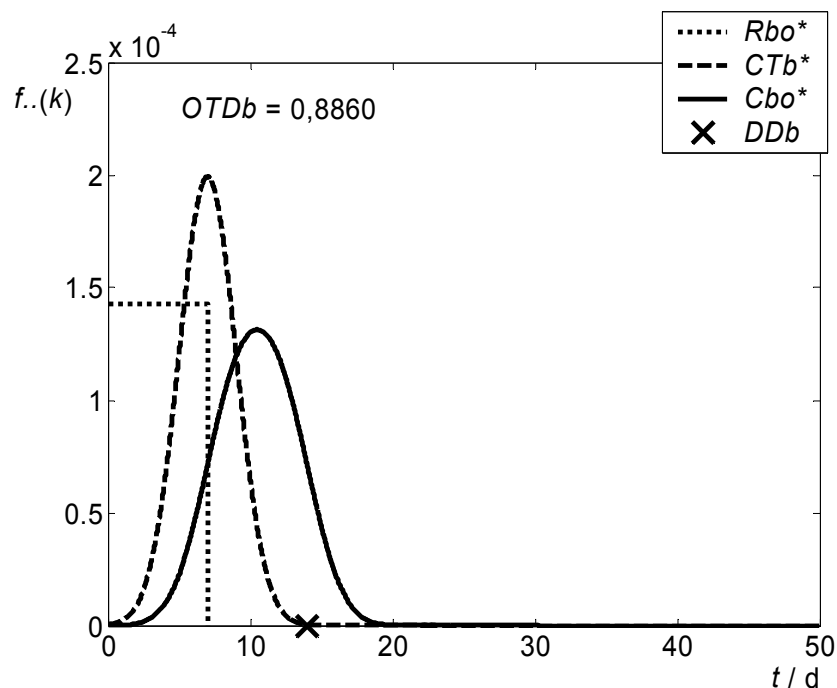


Abbildung 3.22: Berechnung der  $OTDb$  für  $CTb^* \sim N(\mu_b = 7 \text{ d}, \sigma_b = 2 \text{ d})$  und  $CTP_b = 7 \text{ d}$

### 3.4.3 Manufacturing

Zur Ermittlung der Liefertreue des Backends bezogen auf das Gesamtsystem muss zunächst die Wahrscheinlichkeitsfunktion der Einschleusung mit Berücksichtigung des Vorgeschehens ermittelt werden. Die Einschleusung in das Backend erfolgt nach der planmäßigen Lieferwoche des Frontends. Auch Lose, die schon früher im Frontend fertig gestellt wurden, warten bis zum Beginn der Einschleuswoche des Backends. Verspätete Bestellungen werden im Backend bevorzugt weiterverarbeitet. Der Zeitpunkt der Einschleusung in das Backend liegt also immer nach dem Fälligkeitstermin des Frontends.

$$Rb^* > DDf \quad (3.47)$$

Die Due Date Frontend ist hier also sowohl der Stichtag für die Liefertreue im Frontend als auch der Starttermin für die frühest mögliche Einschleusung in das Backend.

Die Wahrscheinlichkeitsverteilung der Einschleusung ins Backend hängt von der Form der Wahrscheinlichkeitsverteilung der Fertigstellung Frontend sowie von der planmäßigen Lieferwoche für das Frontend ab. Ist ein Los im Frontend rechtzeitig fertig gestellt worden, so hat es innerhalb der planmäßigen Einschleuswoche des Backends eine für jeden Zeitpunkt gleich große Wahrscheinlichkeit, eingeschleust zu werden. Ist ein Los im Frontend nicht rechtzeitig fertig gestellt worden, seine Lieferung fällt aber noch in die Einschleuswoche, so hat es für die restlichen Tage der Einschleuswoche eine gleichverteilte Einschleuswahrscheinlichkeit. Lose, die erst nach Ablauf der Einschleuswoche des Backends vom Frontend fertig gestellt werden, werden sofort in das Backend eingeschleust und weiterverarbeitet.

Um die Wahrscheinlichkeitsfunktion der Einschleusung ins Backend ermitteln zu können, wird zunächst die gestutzte Wahrscheinlichkeitsfunktion der Fertigstellung Frontend benötigt. Dazu wird der zeitdiskrete Operator der Stutzung [Tra96b] verwendet:

$$\pi_m(x(k)) = \begin{cases} 0 & , k < m \\ \sum_{l=-\infty}^m a(l) & , k = m \\ x(k) & , k > m \end{cases} \quad (3.48)$$

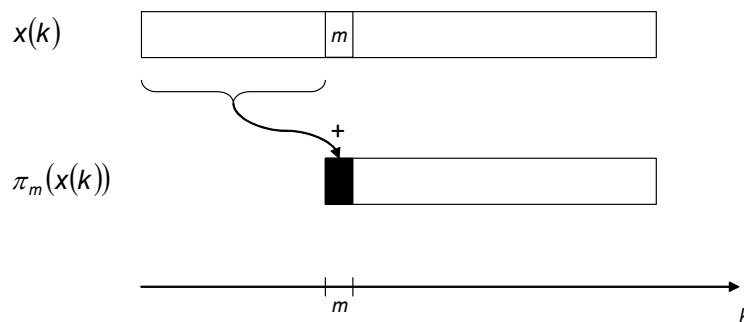


Abbildung 3.23: Wirkungsweise des zeitdiskreten Operators der Stutzung

Der Stutzungs-Operator  $\pi_m$  addiert alle Werte der Wahrscheinlichkeitsfunktion für Variablenwerte kleiner oder gleich  $m$  und weist sie als Funktionswert genau diesen Variablenwertes zu. Die Wahrscheinlichkeitswerte aller Variablenwerte kleiner als  $m$  werden zu Null gesetzt, die Wahrscheinlichkeitswerte aller Variablenwerte größer als  $m$  behalten ihren ursprünglichen Wert.

Die auf das erste diskrete Zeitintervall der Einschleuswoche des Backends gestutzte Wahrscheinlichkeitsfunktion der Fertigstellung Frontend wird somit durch folgenden Ausdruck beschrieben:

$$\pi_{\frac{DDf}{M}+1} (f_{Cf^*}(k)) \quad (3.49)$$

Sie gibt die Wahrscheinlichkeitsfunktion der theoretisch erst möglichen Einschleusung des Loses in das Backend an.

Abbildung 3.24 zeigt die auf die *DDf* gestutzte Wahrscheinlichkeitsfunktion der Fertigstellung Frontend für das Beispiel aus Abbildung 3.21. Der Funktionswert für den ersten Zeitabschnitt nach der *DDf* entspricht dabei der *OTDf*, die hier den Wert 0,6342 hat und weit außerhalb der Darstellung liegt. Für die Zeiten nach der *DDf* liegt die gestutzte Kurve genau über der Wahrscheinlichkeitsfunktion der Fertigstellung Frontend *Cf\**.

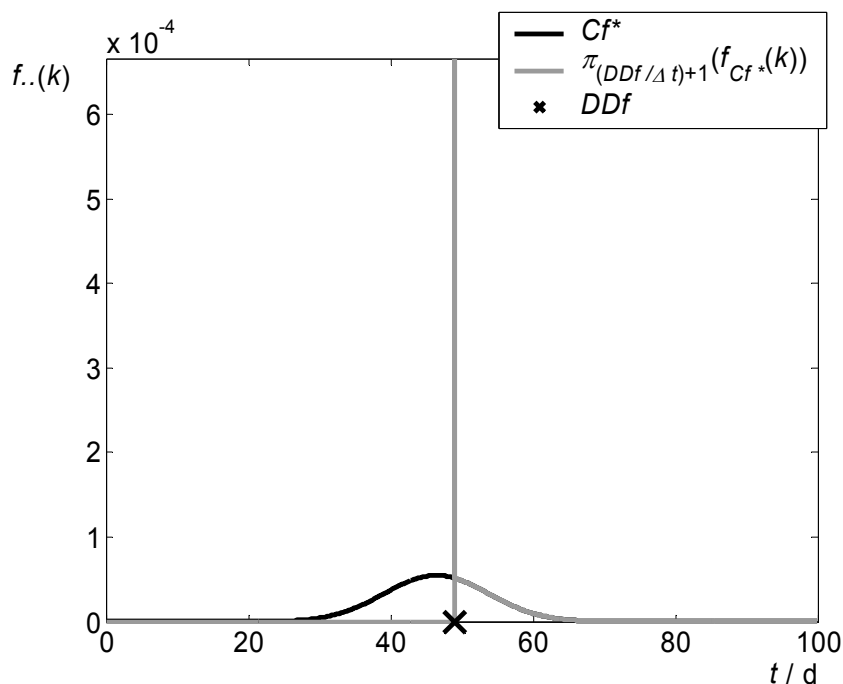


Abbildung 3.24: Wahrscheinlichkeitsfunktionen der Fertigstellung Frontend sowie der auf die  $DDf$  gestutzten Fertigstellung für  $CTf^* \sim N(\mu f = 43 \text{ d}, \sigma f = 7 \text{ d})$  und  $CTPf = 42 \text{ d}$

Aus dieser auf das erste Zeitintervall der Einschleuswoche des Backends gestutzte Wahrscheinlichkeitsfunktion der Fertigstellung Frontend wird die Wahrscheinlichkeitsfunktion für die Einschleusung Backend ermittelt. Abbildung 3.25 zeigt die Berechnung der Wahrscheinlichkeitsfunktion für die Einschleusung in das Backend anhand eines Beispiels mit einer Diskretisierungskonstante von  $\Delta t = 1$  Tag für die Werte innerhalb der Einschleuswoche. Ausgegangen wird von der gestutzten Wahrscheinlichkeitsfunktion der Fertigstellung Frontend. Sie gibt die Wahrscheinlichkeit an, an welchem Tag ein Los frühestens zur Einschleusung ins Backend bereit steht. Ist ein Los am ersten Tag der Einschleuswoche verfügbar, so ist für dieses Los die Wahrscheinlichkeit der Einschleusung an jedem Tag der Einschleuswoche gleich groß, nämlich  $P = 1/7$ . In Abbildung 3.25 ist die dazu gehörige Einzelfunktion mit  $f_1(k)$  bezeichnet. Steht ein Los erst am zweiten Tag der Einschleuswoche zur Verfügung, so kann es nur an diesem und den folgenden Tagen eingeschleust werden. Für jeden Tag gilt die gleiche Wahrscheinlichkeit von  $P = 1/6$ , hier durch die Einzelfunktion  $f_2(k)$  beschrieben. Für die weiteren Tage gilt analog das Gleiche.

Die Wahrscheinlichkeitsfunktion für die Einschleusung in das Backend ergibt sich dann aus der Summe der Einzelfunktionen, die mit dem jeweiligen Wert der gestutzten Wahrscheinlichkeitsfunktion der Fertigung Frontend multipliziert wurden.

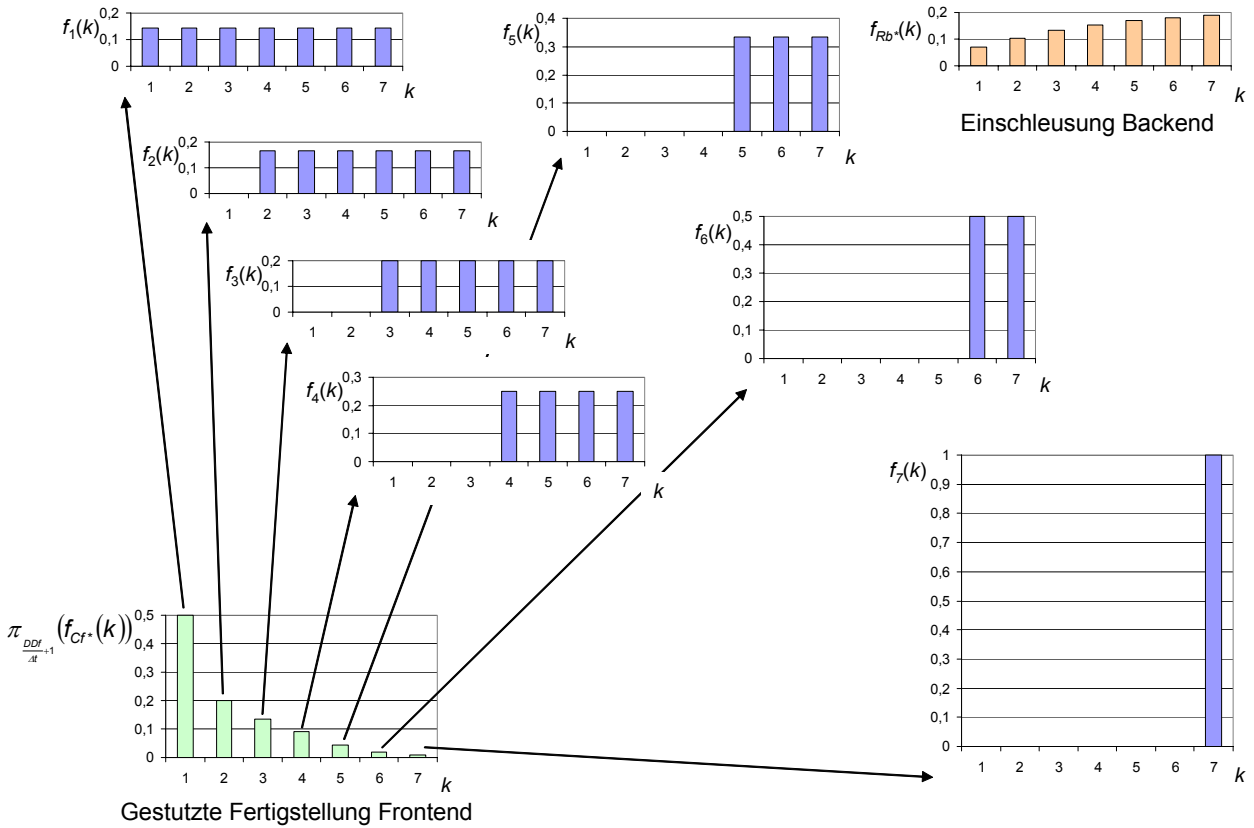


Abbildung 3.25: Ermittlung der Wahrscheinlichkeitsfunktion der Einschleusung Backend aus der gestutzten Wahrscheinlichkeitsfunktion der Fertigstellung Frontend für die Einschleuswoche

Werden Lose im Frontend so spät fertig gestellt, dass sie die Die Bank erst nach ihrer planmäßigen Einschleuswoche für das Backend erreichen, ist es sinnvoll, sie sofort in das Backend einzuschleusen. Deshalb sind für die Zeit nach der Einschleuswoche des Backends die Wahrscheinlichkeitswerte der Fertigstellung Frontend und der Einschleusung Backend identisch. Mathematisch wird dieser gesamte Sachverhalt in folgender Gleichung beschrieben.

$$f_{Rb^*}(k) = \begin{cases} 0 & , \quad k \leq \frac{DDf}{\Delta t} \\ \sum_{i=\frac{DDf}{\Delta t}+1}^k \pi_{\frac{DDf}{\Delta t}+1} (f_{Cf^*}(i)) \cdot \frac{1}{\frac{7d}{\Delta t} - i + \frac{DDf}{\Delta t} + 1} & , \quad \frac{DDf}{\Delta t} < k \leq \frac{DDf + 7d}{\Delta t} \\ f_{Cf^*}(k) & , \quad k > \frac{DDf + 7d}{\Delta t} \end{cases} \quad (3.50)$$

Die aus der gestutzten Wahrscheinlichkeitsfunktion Fertigstellung Frontend ermittelte Wahrscheinlichkeitsfunktion Einschleusung Backend zeigt für obiges Beispiel Abbildung 3.26. Die Länge der Einschleuswoche des Backends ist an der grauen Kurve der Wahrscheinlichkeitsfunktion der Einschleusung Backend  $Rb^*$  gut erkennbar. Die



Einschleusung ist am Ende der Woche wahrscheinlicher als zu Beginn der Woche. Nach der Einschleuswoche liegt die Wahrscheinlichkeitsfunktion der Einschleusung genau über der Wahrscheinlichkeitsfunktion der Fertigstellung.

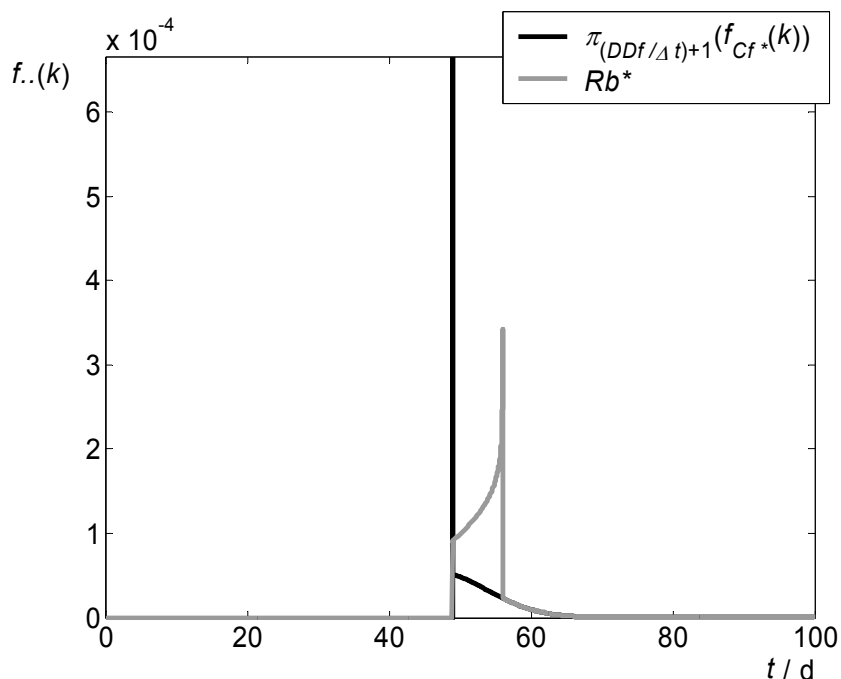


Abbildung 3.26: Wahrscheinlichkeitsfunktion Einschleusung Backend  $Rb^*$  für  $CTf^* \sim N(\mu f = 43 \text{ d}, \sigma f = 7 \text{ d})$  und  $CTPf = 42 \text{ d}$

Die Durchlaufzeit der Lose eines Produktes im Backend  $CTb^*$  wird wiederum als normalverteilt angenommen, mit einem Mittelwert  $\mu b$  und einer Standardabweichung  $\sigma b$ .

$$\begin{aligned}
 f_{CTb^*}(k) &= \int_{(k-1) \cdot \Delta t}^{k \cdot \Delta t} g(t) \\
 &= G(k \cdot \Delta t) - G((k-1) \cdot \Delta t)
 \end{aligned}
 \tag{3.51}$$

$G(t)$  und  $g(t)$  stehen für die Verteilungsfunktion und die Dichtefunktion der Gaußschen Normalverteilung.

Die Zufallsvariable Fertigstellung Backend  $Cb^*$  ergibt sich wiederum aus der Addition der Zufallsvariablen Einschleusung Backend  $Rb^*$  und Durchlaufzeit Backend  $CTb^*$ , womit sich die Wahrscheinlichkeitsfunktion aus der Faltung der beiden Wahrscheinlichkeitsfunktionen ergibt.

$$f_{Cb^*}(k) = f_{Rb^*}(k) \otimes f_{CTb^*}(k)
 \tag{3.52}$$

Die On-Time-Delivery Manufacturing wird am Ende der planmäßigen Lieferwoche des Backends ermittelt. Daraus ergibt sich der Stichtag der Lieferung. Er ergibt sich aus den Fälligkeitsterminen des Frontends und des Backends.

$$DDm = DDf + DDb \tag{3.53}$$

Die Wahrscheinlichkeit für das rechtzeitige Erreichen des Distribution Centers, die On-Time-Delivery Manufacturing, kann in der Verteilungsfunktion der Fertigstellung Backend am Zeitpunkt des Stichtages abgelesen werden.

$$OTDm = P(Cb^* \leq DDm) = \frac{F_{Cb^*}(k = DDm / \Delta t)}{F_{Cb^*}(k_{max})} \tag{3.54}$$

Für das bisher gezeigte Beispiel sind die Wahrscheinlichkeitsfunktionen von Release Backend, Cycle Time Backend, Completion Backend sowie Due Date Backend und die daraus ermittelte On Time Delivery Backend in Abbildung 3.27 dargestellt. Die Graphik zeigt, dass ein Großteil der verspätet aus dem Frontend gelieferten Lose durch die Einschleuswoche des Backends aufgefangen werden kann, so dass nach einer  $OTDf = 63,42\%$  und einer  $OTDb = 88,6\%$  eine  $OTDm = 78,55\%$  erreicht werden kann.

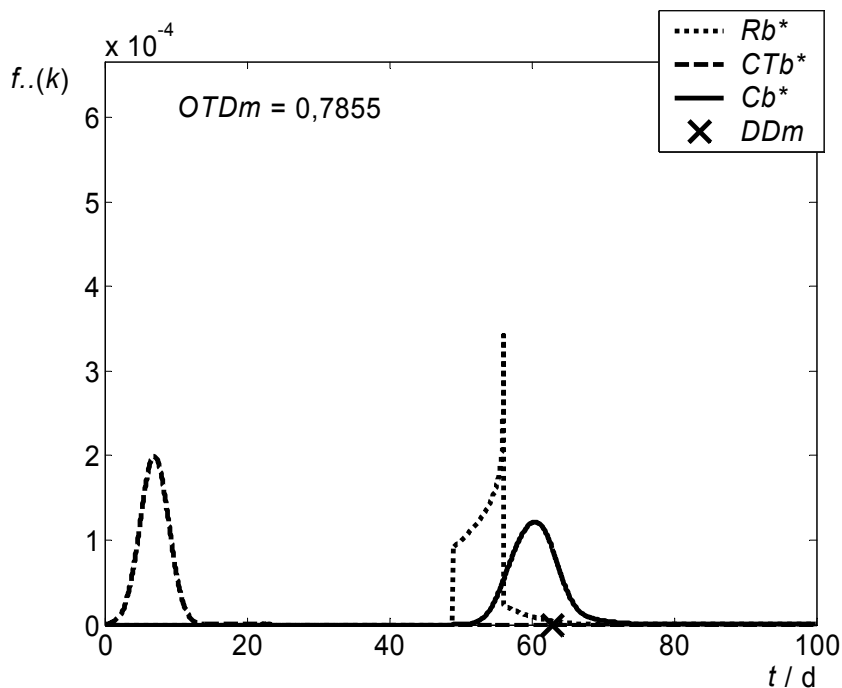


Abbildung 3.27: Berechnung der  $OTDm$  für  $CTf^* \sim N(\mu f = 43 \text{ d}, \sigma f = 7 \text{ d})$ ,  $CTP f = 42 \text{ d}$ ,  $CTb^* \sim N(\mu b = 7 \text{ d}, \sigma b = 2 \text{ d})$  und  $CTP b = 7 \text{ d}$ ;

### 3.4.4 Delivery und Over All

Unter Delivery werden alle Vorgänge zusammengefasst, die nach der Lieferung eines Loses in das Distribution Center erfolgen. Um in der Kontinuität der bisherigen Beschreibung zu bleiben, werden wiederum eine „Einschleusung“ und eine Durchlaufzeit ermittelt. Analog wie zuvor erfolgt ebenfalls eine alleinige Betrachtung dieser dritten Stufe sowie eine Gesamtbetrachtung.

Die Bewertung der Performance der Auslieferung ohne Berücksichtigung der Vorgeschichte erfolgt nur aufgrund der Wahrscheinlichkeitsfunktion der Durchlaufzeit der Auslieferung, da es keine Einschleuswoche gibt. Der Liefertermin entspricht dann der Plandurchlaufzeit.

$$DDd = CTPd \quad (3.55)$$

Die On-Time-Delivery Delivery ist somit die Wahrscheinlichkeit dafür, dass die Durchlaufzeit in der Auslieferung kleiner oder gleich dem Liefertermin der Auslieferung ist.

$$OTDd = P(CTd^* \leq DDd) = \frac{F_{CTd^*}(k = DDd / \Delta t)}{F_{CTd^*}(k_{\max})} \quad (3.56)$$

Der Service Level  $SL$  wird tagfein betrachtet. Er ist der Prozentsatz der Lieferungen, die komplett in der Regellieferzeit ( $RLZ$ ) an den Kunden ausgeliefert wurden. Da in diesem Modell nur die Produktion eines einzelnen Loses abgebildet wird, kann nur eine On-Time-Delivery für einen derartigen Einzelauftrag berechnet werden. Sie wird On-Time-Delivery Over All bezeichnet.

Um in der gleichen Systematik wie bisher zu bleiben, wird der Beginn der Auslieferung als Einschleusung Delivery bezeichnet. Die Einschleusung Delivery erfolgt unmittelbar nach dem Liefertermin Manufacturing, sofern das Los bereits fertig gestellt wurde und sich im Delivery Center befindet. Die Wahrscheinlichkeitsfunktion der Einschleusung Delivery ist die auf das erste Zeitintervall nach dem Manufacturing-Stichtag gestützte Wahrscheinlichkeitsfunktion der Fertigstellung Backend (siehe auch Abbildung 3.28).

$$f_{Rd^*}(k) = \pi_{\frac{DDb}{\Delta t} + 1}(f_{Cb^*}(k)) \quad (3.57)$$

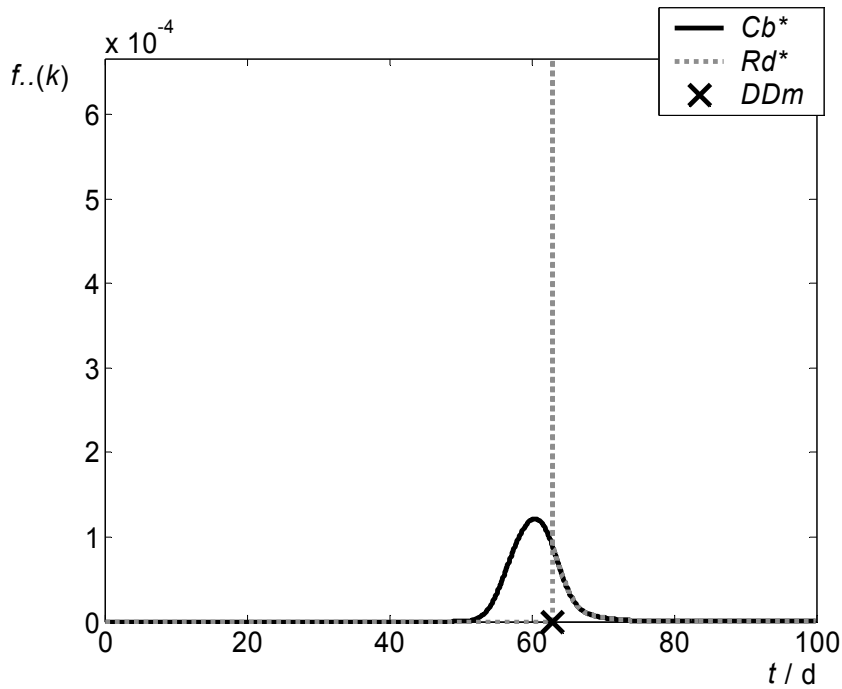


Abbildung 3.28: Wahrscheinlichkeitsfunktion der Einschleusung Delivery für  $CTf^* \sim N(\mu f = 43 \text{ d}, \sigma f = 7 \text{ d})$ ,  $CTPf = 42 \text{ d}$ ,  $CTb^* \sim N(\mu b = 7 \text{ d}, \sigma b = 2 \text{ d})$  und  $CTPb = 7 \text{ d}$

Die Zeitdauer zwischen Entnahme aus dem Distribution Center und der Auslieferung an den Kunden ist die Durchlaufzeit der Delivery  $CTd^*$ . Sie ist die Auslieferungszeit, entsprechend der Fahrt eines Lkws von A nach B. Für sie wird wiederum eine Normalverteilung angenommen mit einem Mittelwert  $\mu d$  und einer Standardabweichung  $\sigma d$ .

$$f_{CTd^*}(k) = \int_{(k-1) \cdot \Delta t}^{k \cdot \Delta t} g(t) dt = G(k \cdot \Delta t) - G((k-1) \cdot \Delta t) \quad (3.58)$$

Die Wahrscheinlichkeitsfunktion der Fertigstellung Delivery ergibt sich aus der Faltung der Wahrscheinlichkeitsfunktionen von Einschleusung und Durchlaufzeit.

$$f_{Cd^*}(k) = f_{Rd^*}(k) \otimes f_{CTd^*}(k) \quad (3.59)$$

Der Stichtag der Lieferung beim Kunden ergibt sich aus den Lieferterminen des Backends und dem Liefertermin Delivery.

$$DDo = DDm + DDd \quad (3.60)$$

Die On-Time-Delivery Over All wird aus der Verteilungsfunktion der Fertigstellung Delivery am Zeitpunkt des Liefertermins Over All ermittelt.

$$OTDo = P(Cd^* \leq DDo) = \frac{F_{Cd^*}(k = DDo / \Delta t)}{F_{Cd^*}(k_{\max})} \quad (3.61)$$

Abbildung 3.29 zeigt ein Beispiel für die Wahrscheinlichkeitsfunktionen von Release Delivery, Cycle Time Delivery, Completion Delivery sowie Due Date Over All und die daraus resultierende On Time Delivery Over All. Die Auslieferung geschieht im Allgemeinen sehr schnell und unterliegt nur geringen Schwankungen. Verspätungen aus der Lieferung des Backends können deshalb nur sehr geringfügig kompensiert werden.

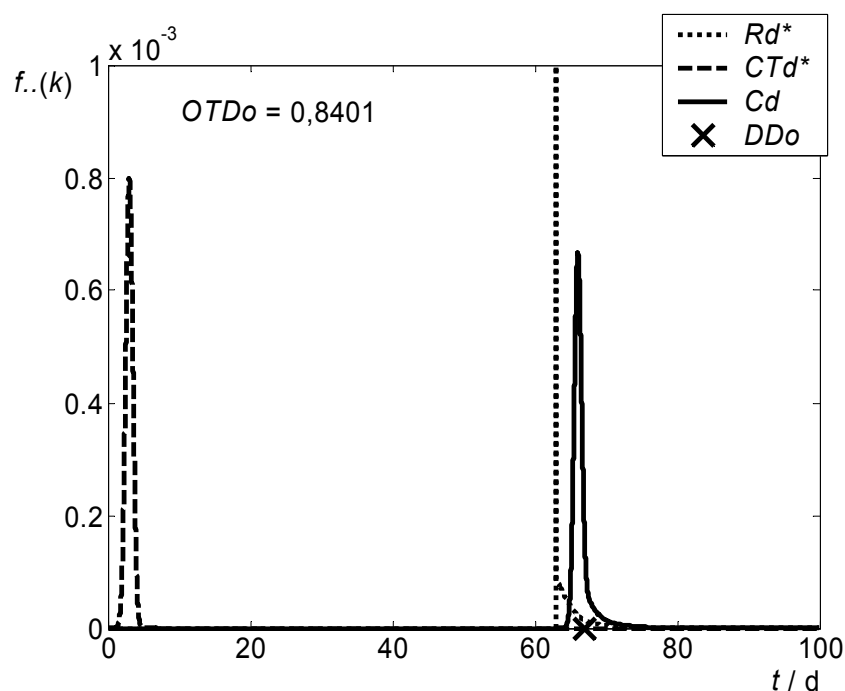


Abbildung 3.29:  $OTDo$  für  $CTf^* \sim N(\mu f = 43 \text{ d}, \sigma f = 7 \text{ d})$ ,  $CTPf = 42 \text{ d}$ ,  
 $CTb^* \sim N(\mu b = 7 \text{ d}, \sigma b = 2 \text{ d})$ ,  $CTPb = 7 \text{ d}$ ,  
 $CTd^* \sim N(\mu d = 3 \text{ d}, \sigma d = 0,5 \text{ d})$  und  $CTPd = 4 \text{ d}$

### 3.4.5 Ergebnisse

Das oben beschriebene Modell berechnet mit den Daten aus Tabelle 3.6 die in Abbildung 3.30 gezeigten Ergebnisse. Die Wahrscheinlichkeitsfunktion der Einschleusung Frontend  $Rf^*$  gefaltet mit der Wahrscheinlichkeitsfunktion der Durchlaufzeit Frontend  $CTf^*$  ergibt die Wahrscheinlichkeitsfunktion der Fertigstellung Frontend  $Cf^*$ . Daraus wird die On-Time-Delivery Frontend  $OTDf$  zum Zeitpunkt Due Date Frontend  $DDf$  berechnet und die Wahrscheinlichkeitsfunktion der Einschleusung Backend  $Rb^*$  ermittelt. Aus der Faltung mit der Wahrscheinlichkeitsfunktion der Durchlaufzeit Backend  $CTb^*$  ergibt sich die Wahrscheinlichkeitsfunktion der Fertigstellung Backend  $Cb^*$ , woraus für den Zeitpunkt der Due Date Manufacturing  $DDm$  die On-Time-Delivery Manufacturing  $OTDm$  berechnet und die Wahrscheinlichkeitsfunktion der Einschleusung Delivery  $Rd^*$  ermittelt wird. Gefaltet mit der Wahrscheinlichkeitsfunktion der Durchlaufzeit Delivery  $CTd^*$  ergibt sich die

Wahrscheinlichkeitsfunktion der Fertigstellung Delivery  $Cd^*$ . Aus ihr wird mit der Due Date Over All  $DDo$  die On-Time-Delivery Over All  $OTDo$  berechnet.

OTDmRb-002	$\mu$	$\sigma$	CTP	DD	DDm/o	OTD	OTDm/o
Frontend	43 d	7 d	56 d	63 d	--	0,9883	--
Backend	7 d	2 d	7 d	14 d	77 d	0,8860	0,8849
Delivery	3 d	0,5 d	4 d	4 d	81 d	0,9773	0,9199

Tabelle 3.6: Daten zu Abbildung 3.30

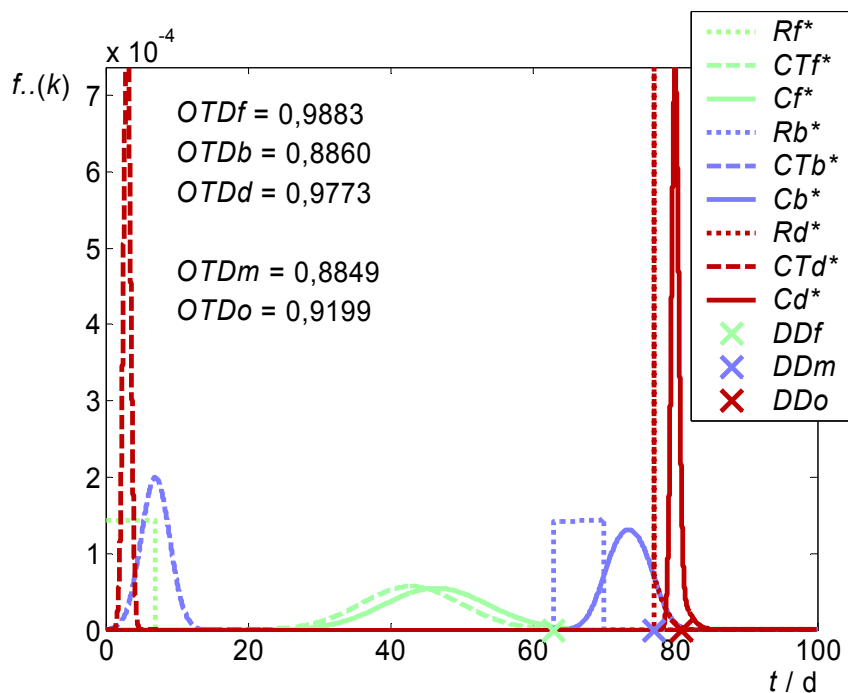


Abbildung 3.30:  $OTD$ -Berechnungen für  $CTf^* \sim N(\mu_f = 43 \text{ d}, \sigma_f = 7 \text{ d})$ ;  $CTP_f = 56 \text{ d}$ ;  
 $CTb^* \sim N(\mu_b = 7 \text{ d}, \sigma_b = 2 \text{ d})$ ;  $CTP_b = 7 \text{ d}$ ;  
 $CTd^* \sim N(\mu_d = 3 \text{ d}, \sigma_d = 0,5 \text{ d})$ ;  $CTP_d = 4 \text{ d}$

In Abbildung 3.30 ist gut erkennbar, dass die Plandurchlaufzeit Frontend für die gegebenen Durchlaufzeit-Wahrscheinlichkeiten groß genug gewählt ist, um eine On-Time-Delivery von  $OTD_f = 98,83\%$  zu erreichen. Als Folge davon ergeben sich für die Einschleusung Backend Wahrscheinlichkeitswerte, die der Wahrscheinlichkeitsfunktion einer idealen Einschleusung (entsprechend der Einschleusung Frontend) sehr nahe kommen. Die On-Time-Delivery Manufacturing ist dadurch mit  $OTD_m = 88,49\%$  auch nur unwesentlich geringer als die On-Time-Delivery Backend. Da die Auslieferung mit einer sehr hohen On-Time-Delivery Delivery von  $OTD_d = 97,73\%$  ohne Einschleuswoche direkt an die Due Date Manufacturing anschließt, kann mit ihr ein Teil der Verspätungswahrscheinlichkeit kompensiert werden, so dass sich eine On-Time-Delivery Over All von  $OTD_o = 91,99\%$  ergibt.

Für das Beispiel, das bereits im Kapitel oben gezeigt wurde (siehe auch Tabelle 3.7 und Abbildung 3.31) wurde bei sonst gleichen Werten eine zwei Wochen geringere Plandurchlaufzeit Frontend gewählt. Sie führt dazu, dass ein Los nur noch mit einer Wahrscheinlichkeit von  $OTDf = 63,42\%$  rechtzeitig an die Die Bank geliefert wird. Somit ergeben sich auch andere Wahrscheinlichkeiten für die Einschleusung Backend. Die Wahrscheinlichkeit einer Einschleusung am Anfang der Einschleuswoche ist wesentlich geringer als die Wahrscheinlichkeit einer Einschleusung am Ende der Einschleuswoche. Mit der gleichen Wahrscheinlichkeit, mit der das Los die Die Bank erst nach Beendigung der Einschleuswoche Backend erreicht, wird es dann auch in das Backend eingeschleust. Damit ergibt sich eine geringere On-Time-Delivery Manufacturing, die mit  $OTDm = 78,55\%$  um über zehn Prozentpunkte geringer ist als die On-Time-Delivery Backend von  $OTDb = 88,60\%$ . Die Auslieferung kann wiederum einen Teil der Verspätung kompensieren, allerdings nur bis zu einer On-Time-Delivery Over All von  $OTDo = 84,01\%$ .

OTDmRb-026	$\mu$	$\sigma$	CTP	DD	DDm/o	OTD	OTDm/o
Frontend	43 d	7 d	42 d	49 d	--	0,6342	--
Backend	7 d	2 d	7 d	14 d	63 d	0,8860	0,7855
Delivery	3 d	0,5 d	4 d	4 d	67 d	0,9773	0,8401

Tabelle 3.7: Daten zu Abbildung 3.31

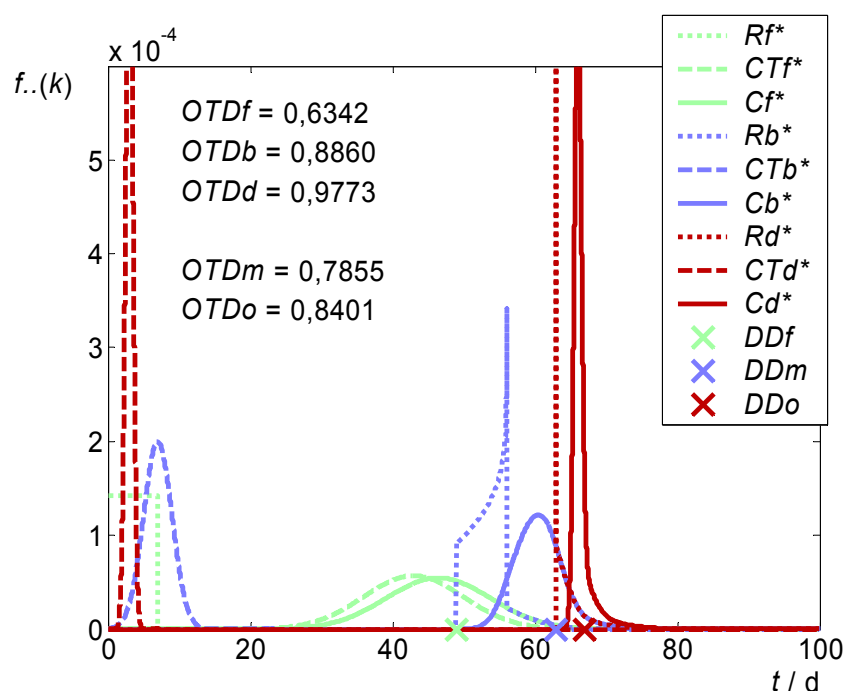


Abbildung 3.31: OTD-Berechnungen für  $CTf^* \sim N(\mu_f = 43 \text{ d}, \sigma_f = 7 \text{ d})$ ;  $CTPf = 42 \text{ d}$ ;  $CTb^* \sim N(\mu_b = 7 \text{ d}, \sigma_b = 2 \text{ d})$ ;  $CTPb = 7 \text{ d}$ ;  $CTd^* \sim N(\mu_d = 3 \text{ d}, \sigma_d = 0,5 \text{ d})$ ;  $CTPd = 4 \text{ d}$

Die Durchlaufzeit für die Auslieferung (Delivery) wurde in beiden Beispielen mit einer Normalverteilung mit sehr geringer Standardabweichung berechnet, um die relativ kurze Dauer darzustellen. Im Normalfall kann davon ausgegangen werden, dass die Auslieferung nur sehr geringen bis gar keinen Schwankungen unterliegt. Verspätungen bei der Lieferung in das Distribution Center können durch den Vorgang der Auslieferung meist nicht mehr kompensiert werden. Wird jedoch die Auslieferung als eine konstante Größe ohne den Charakter einer streuenden Zufallsvariable betrachtet, so ist die On-Time-Delivery Manufacturing bereits ein Maß für die Liefertreue der gesamten Supply Chain von der Einschleusung ins Frontend bis zur Auslieferung beim Kunden und somit für den Service Level. Im Folgenden werden daher Zusammenhänge von On-Time-Delivery Frontend, On-Time-Delivery Backend und On-Time-Delivery Manufacturing untersucht.

Bevor weitere Ergebnisse diskutiert werden, sollen zunächst zwei Annahmen widerlegt werden, die in der Diskussion über die interne Liefertreue und die CLIP-Berechnung immer wieder auftauchen, jedoch nicht zutreffen.

Vielfach wird davon ausgegangen, dass sich die Gesamtperformance von On-Time-Delivery Frontend und On-Time-Delivery Backend aus einer Multiplikation beider Werte ermitteln lässt. Da beide Werte Wahrscheinlichkeiten sind, also immer kleiner als eins, würde somit immer eine Gesamtperformance erreicht werden, die geringer als der kleinere der beiden Werte ist. Wie die oben dargestellten Ergebnisse zeigen, trifft dies nicht zu. In Abbildung 3.32 sind On-Time-Delivery-Werte für variierende Mittelwerte der Durchlaufzeit Frontend dargestellt. Die On-Time-Delivery Manufacturing liegt dabei für einen Mittelwert der Durchlaufzeit Frontend von 38 Tagen unterhalb beider Werte von On-Time-Delivery Frontend und On-Time-Delivery Backend. Für höhere Mittelwerte liegt sie zwischen den einstufigen Werten. Werte, die sich aus der Multiplikation aus On-Time-Delivery Frontend und On-Time-Delivery Backend ergeben, liegen weit unter der tatsächlichen Kurve der On-Time-Delivery Manufacturing.

Wie sich hier schon vermuten lässt, ist die Lage der On-Time-Delivery Manufacturing eine Folge der Lage von On-Time-Delivery Frontend und On-Time-Delivery Backend zueinander. Diese Zusammenhänge werden weiter unten ausführlich betrachtet.

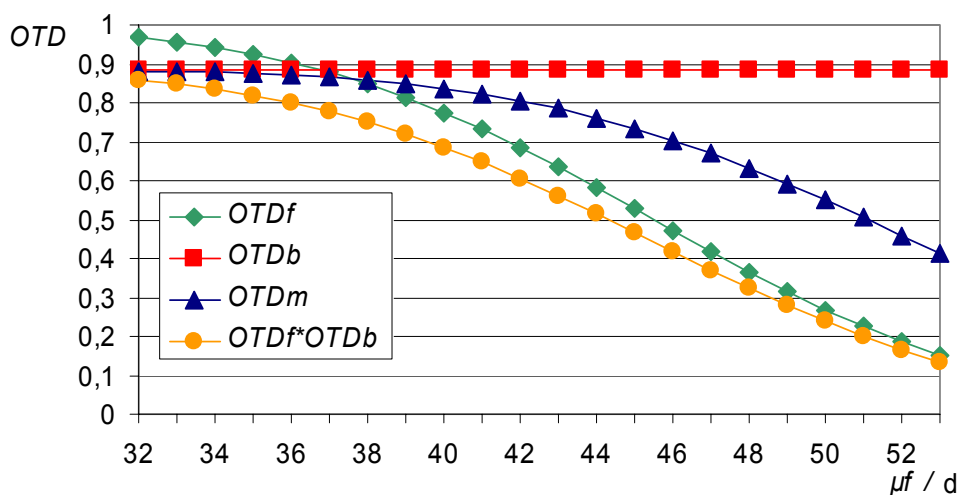


Abbildung 3.32: On-Time-Delivery für  $CTf \sim N(\mu_f, \sigma_f = 7 \text{ d})$ ;  $CTPf = 42 \text{ d}$ ;  
 $CTb \sim N(\mu_b = 7 \text{ d}, \sigma_b = 2 \text{ d})$ ;  $CTPb = 7 \text{ d}$



Bei der zweiten irrtümlich getroffenen Annahme handelt es sich um den Versuch, eine Tabelle aufzustellen, in der den Werten von On-Time-Delivery Frontend und On-Time-Delivery Backend eindeutig Werte der On-Time-Delivery Manufacturing zugeordnet werden. Dies ist jedoch nicht möglich, wie die Ergebnisse in Tabelle 3.8 belegen.

Nr.	$\mu f / d$	$\sigma f / d$	$CTPf / d$	$OTDf$	$\mu b / d$	$\sigma b / d$	$CTPb / d$	$OTDb$	$OTDm$
1.	40 d	5 d	35 d	0,3908	7 d	2 d	7 d	0,8860	0,7218
2.	48 d	10 d	42 d	0,4033	7 d	2 d	7 d	0,8860	0,5848
3.	39 d	7 d	42 d	0,8137	7 d	2 d	7 d	0,8860	0,8485
4.	43 d	10 d	49 d	0,8241	7 d	2 d	7 d	0,8860	0,8326
5.	49 d	2 d	49 d	0,8861	7 d	2 d	7 d	0,8860	0,8826
6.	36 d	12 d	49 d	0,9123	7 d	2 d	7 d	0,8860	0,8592

Tabelle 3.8: Beispiele gegenläufiger OTD-Entwicklungen

Obwohl die On-Time-Delivery Frontend in der Berechnung 2 um einen Prozentpunkt größer ist als in Berechnung 1, ist die On-Time-Delivery Manufacturing bei gleicher On-Time-Delivery Backend in Berechnung 2 um 14 Prozentpunkte geringer als in Berechnung 1. Bei den insgesamt höheren Werten in Berechnung 3 und 4 sind die Unterschiede nicht mehr so groß, sie sind aber in gleicher Weise gegeben. Hier ergibt sich die größere On-Time-Delivery Manufacturing aus der niedrigeren On-Time-Delivery Frontend bei gleicher On-Time-Delivery Backend.

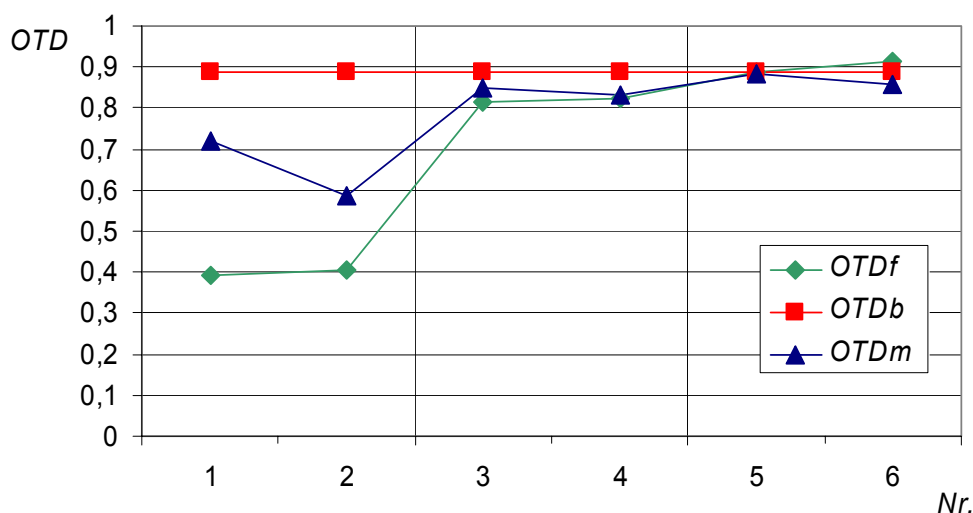


Abbildung 3.33: Gegenläufige OTD-Entwicklungen

Auch nahe der Obergrenze der On-Time-Delivery Manufacturing, also der On-Time-Delivery Backend, können Beispiele für diesen Sachverhalt gefunden werden. Die

Berechnungen zeigen, dass die unterschiedlichen Entwicklungen der On-Time-Delivery Manufacturing eine Folge der unterschiedlichen Standardabweichungen der Durchlaufzeiten im Frontend sind. Die in Tabelle 3.8 berechneten Beispiele sind in Abbildung 3.33 graphisch dargestellt. Zur besseren Lesbarkeit sind die Einzelwerte mit Linien verbunden.

Wie die Berechnungen zeigen, ist die On-Time-Delivery Manufacturing immer kleiner als die On-Time-Delivery Backend. Das liegt daran, dass die Berechnung der On-Time-Delivery Backend nur die Lose berücksichtigt, die bereits in der Die Bank vorhanden sind. Somit wird eine optimale Einschleusung in das Backend erreicht, bei der die einzelnen Wahrscheinlichkeiten für jeden Zeitpunkt der Einschleuswoche gleich sind. Wird die Vorgeschichte mit berücksichtigt, so kann es auch Lose geben, die bei Beginn der Einschleuswoche des Backends die Die Bank noch nicht erreicht haben. Diese Lose können dann nur nach ihrer Lieferung durch das Frontend ins Backend eingeschleust werden. Damit ergibt sich eine Wahrscheinlichkeitsverteilung für die Einschleusung Backend, die für spätere Zeitpunkte höhere Wahrscheinlichkeiten hat (siehe Abbildung 3.31 im Vergleich zu Abbildung 3.30). Diese ungünstigere Einschleusung führt dann zwangsläufig dazu, dass die On-Time-Delivery Manufacturing geringer ist als die On-Time-Delivery Backend. In Abbildung 3.34 bewirkt eine On-Time-Delivery Backend von  $OTDb = 56,56\%$  Werte der On-Time-Delivery Manufacturing die weit unter den Werten der teilweise sehr hohen On-Time-Delivery Frontend liegen.

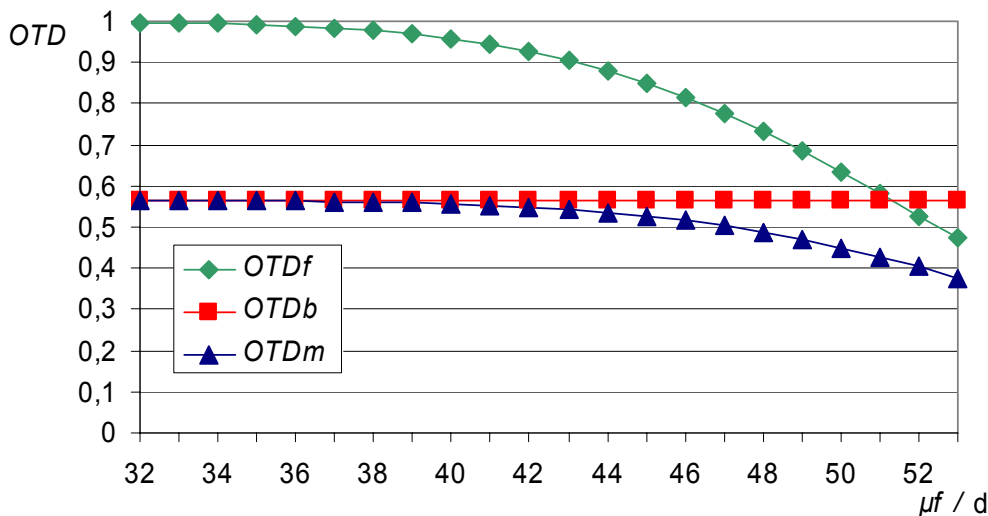


Abbildung 3.34: On-Time-Delivery für  $CTf^* \sim N(\mu_f, \sigma = 7 \text{ d})$ ,  $CTPf = 49 \text{ d}$ ,  $CTb^* \sim N(\mu = 10 \text{ d}, \sigma = 2 \text{ d})$  und  $CTPb = 7 \text{ d}$

Sowohl die Wahl der Plandurchlaufzeiten wie auch Veränderungen der Wahrscheinlichkeitsfunktionen der Durchlaufzeiten haben Auswirkungen auf die Performance-Werte der On-Time-Delivery. In Abbildung 3.35 sind die On-Time-Deliverys von Frontend, Backend und Manufacturing für unterschiedliche Mittelwerte der Durchlaufzeit Frontend dargestellt. Die Plandurchlaufzeit Frontend von  $CTPf = 49$  Tagen wird zu einer Due Date von  $DDf = 56$  Tagen. Mit steigender mittlerer Durchlaufzeit wird die Wahrscheinlichkeit für ein Los, die Die Bank rechtzeitig zu erreichen, also die On-Time-Delivery Frontend geringer. Mit sinkender mittlerer Durchlaufzeit wird sie immer größer. Auf die On-Time-

Delivery Manufacturing hat dieser Anstieg jedoch nur dann eine deutliche Wirkung, solange die On-Time-Delivery Frontend geringer ist als die On-Time-Delivery Backend.

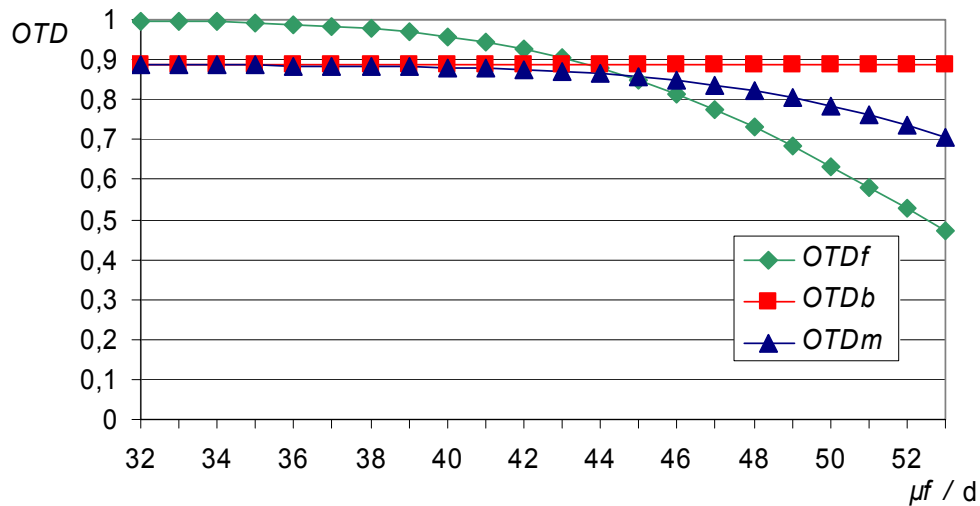


Abbildung 3.35: On-Time-Delivery für  $CTf^* \sim N(\mu_f, \sigma = 7 \text{ d})$ ,  $CTPf = 49 \text{ d}$ ,  $CTb^* \sim N(\mu = 7 \text{ d}, \sigma = 2 \text{ d})$  und  $CTPb = 7 \text{ d}$

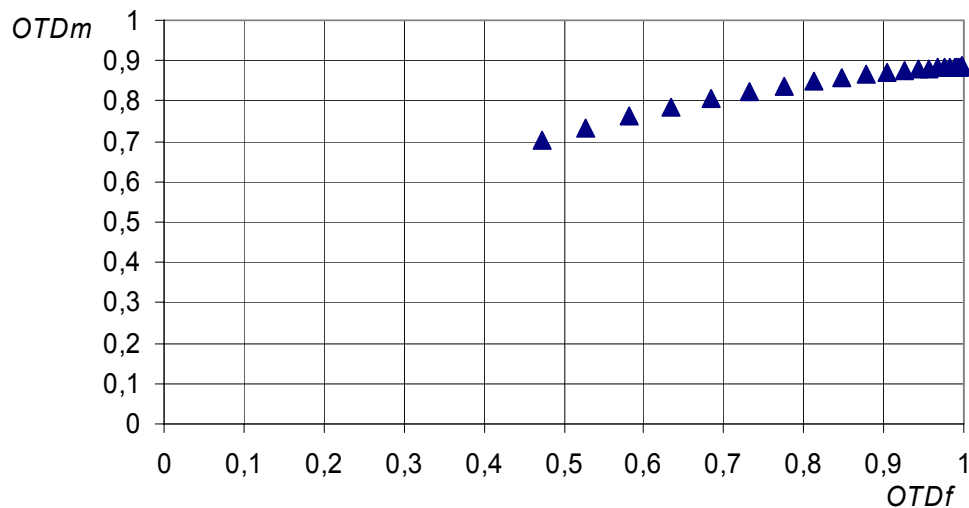


Abbildung 3.36: On-Time-Delivery für  $CTf^* \sim N(\mu_f, \sigma = 7 \text{ d})$ ,  $CTPf = 49 \text{ d}$ ,  $CTb^* \sim N(\mu = 7 \text{ d}, \sigma = 2 \text{ d})$  und  $CTPb = 7 \text{ d}$

Abbildung 3.36 zeigt bei gleichem Sachverhalt die On-Time-Delivery Manufacturing als Funktion der On-Time-Delivery Frontend bei konstanter On-Time-Delivery Backend von  $OTDb = 88,60\%$ . Deutlich zu erkennen ist nicht nur, dass jeder Tag Verkürzung der mittleren Durchlaufzeit von Tag zu Tag eine immer geringere Auswirkung auf die On-Time-Delivery

Frontend hat, sondern auch, dass die Verbesserung der On-Time-Delivery Manufacturing mit steigenden Werten der On-Time-Delivery Frontend abnimmt.

Für die Werte in Abbildung 3.37 und Abbildung 3.38 wurde eine geringere Plandurchlaufzeit Frontend von  $CTPf = 42$  Tage verwendet. Die Kurven der On-Time-Delivery Frontend und Manufacturing erfahren dadurch eine horizontale Verschiebung entlang der mittleren Durchlaufzeit des Frontends, die Interpretation bleibt jedoch die gleiche.

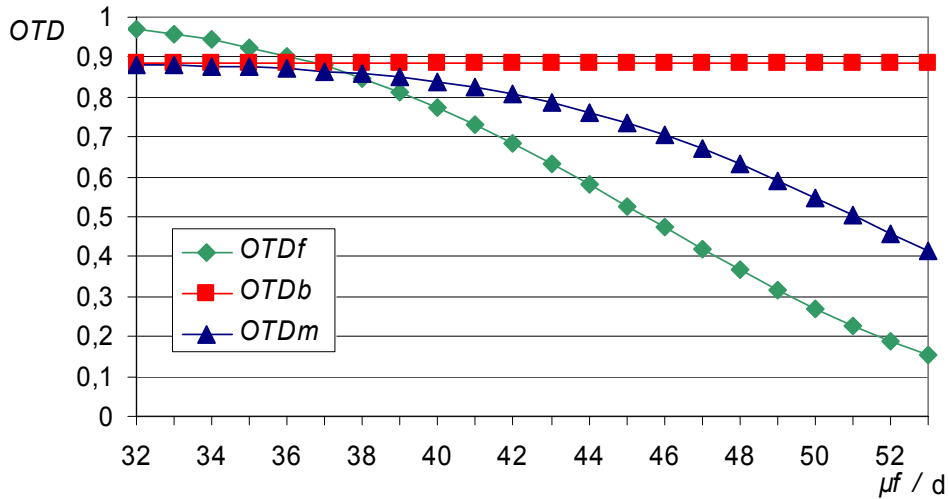


Abbildung 3.37: On-Time-Delivery für  $CTf^* \sim N(\mu_f, \sigma = 7 \text{ d})$ ,  $CTPf = 42 \text{ d}$ ,  $CTb^* \sim N(\mu = 7 \text{ d}, \sigma = 2 \text{ d})$  und  $CTPb = 7 \text{ d}$

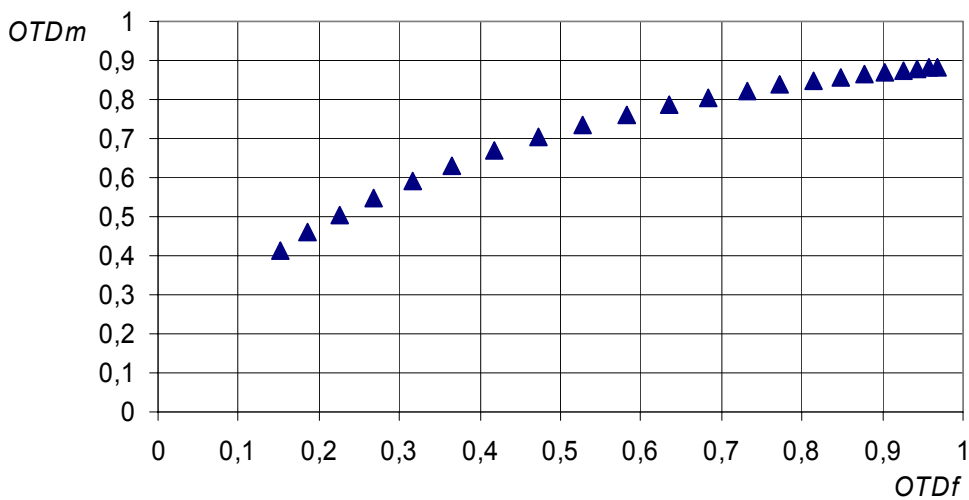


Abbildung 3.38: On-Time-Delivery für  $CTf^* \sim N(\mu_f, \sigma = 7 \text{ d})$ ,  $CTPf = 42 \text{ d}$ ,  $CTb^* \sim N(\mu = 7 \text{ d}, \sigma = 2 \text{ d})$  und  $CTPb = 7 \text{ d}$

Die Darstellung der On-Time-Delivery Manufacturing als Funktion der On-Time-Delivery Frontend in Abbildung 3.38 erfährt keine Verschiebung der Graphik. Hier scheint die Kurve

gleich zu bleiben, lediglich eine andere Auswahl der Werte lässt die Kurve weiter links erscheinen. Die 22 Werte liegen tatsächlich auf der gleichen Kurve, es handelt sich lediglich um 22 weiter links liegende Werte.

Die Verschiebung der Ergebniswerte durch unterschiedliche Plandurchlaufzeiten entspricht genau der Änderung der Plandurchlaufzeit, in diesem Beispiel also sieben Tage. Abbildung 3.35 und Abbildung 3.37 stellen identische Kurven dar, die lediglich um 7 Positionen entlang der Abszisse verschoben sind. Beispielsweise sind die On-Time-Deliverys für  $\mu f = 40$  Tage mit  $CTPf = 42$  Tage in Abbildung 3.37 und für  $\mu f = 47$  Tage mit  $CTPf = 49$  Tage in Abbildung 3.35 identisch. Das zweite Wertepaar ist gegenüber dem ersten bezüglich des Mittelwertes der Durchlaufzeit Frontend um 7 Tage verschoben. Somit ergibt sich in der Berechnung eine Verschiebung der Wahrscheinlichkeitsfunktion der Durchlaufzeit Frontend gleichzeitig mit einer Verschiebung der Due Date Frontend. Auch die weiteren Wahrscheinlichkeitsfunktionen von Fertigstellung Frontend, Einschleusung Backend und Fertigstellung Backend verschieben sich entsprechend um 7 Tage, ebenso die Due Date Manufacturing. Die relativen Lagen aller Wahrscheinlichkeitsfunktionen zueinander bleiben gleich. Dadurch ergeben sich identische Werte für die On-Time-Delivery Frontend und die On-Time-Delivery Backend.

Die Differenz zwischen Plandurchlaufzeit Frontend und Mittelwert der Durchlaufzeit Frontend beträgt für beide oben genannten Beispielwerte 2 Tage. Das gilt auch für alle anderen übereinstimmenden Werte der On-Time-Deliverys: Dort, wo die Differenz zwischen Plandurchlaufzeit Frontend und mittlerer Durchlaufzeit Frontend gleich ist, ergeben sich die gleichen On-Time-Delivery-Werte. Das führt zu dem Schluss, dass der Abstand zwischen dem Mittelwert der Wahrscheinlichkeitsfunktion der Durchlaufzeit Frontend und der Plandurchlaufzeit Frontend ein entscheidender Parameter ist.

Allgemein gilt folgendes: Die Standardabweichung bestimmt die Form der Wahrscheinlichkeitsfunktion. Zusammen mit der Differenz von Plandurchlaufzeit und Mittelwert der Durchlaufzeit ist dann die On-Time-Delivery eindeutig bestimmt. Dabei ist die Lage der Wahrscheinlichkeitsfunktion, also deren Mittelwert, unerheblich. Das gilt sowohl für das Frontend als auch für das Backend.

Die On-Time-Delivery Manufacturing wiederum ist eindeutig durch die Standardabweichung der Durchlaufzeit Frontend, die Differenz von Plandurchlaufzeit Frontend und Mittelwert der Durchlaufzeit Frontend, die Standardabweichung der Durchlaufzeit Backend und die Differenz von Plandurchlaufzeit Backend und Mittelwert der Durchlaufzeit Backend bestimmt. Kurz: die  $OTDm$  ist eindeutig durch  $\sigma f$ ,  $(CTPf - \mu f)$ ,  $\sigma b$  und  $(CTPb - \mu b)$  bestimmt. Auch hier gilt: die Lage der Wahrscheinlichkeitsfunktionen der Durchlaufzeiten ist unbedeutend für die Größe der On-Time-Delivery Manufacturing.

Abbildung 3.39 zeigt die On-Time-Deliverys über einen weiten Bereich der Differenz von Plandurchlaufzeit Frontend und Mittelwert der Durchlaufzeit Frontend. Abbildung 3.35 und Abbildung 3.37 sind für die entsprechenden Werte von  $CTPf$  und  $\mu f$  jeweils ein Ausschnitt aus Abbildung 3.39.

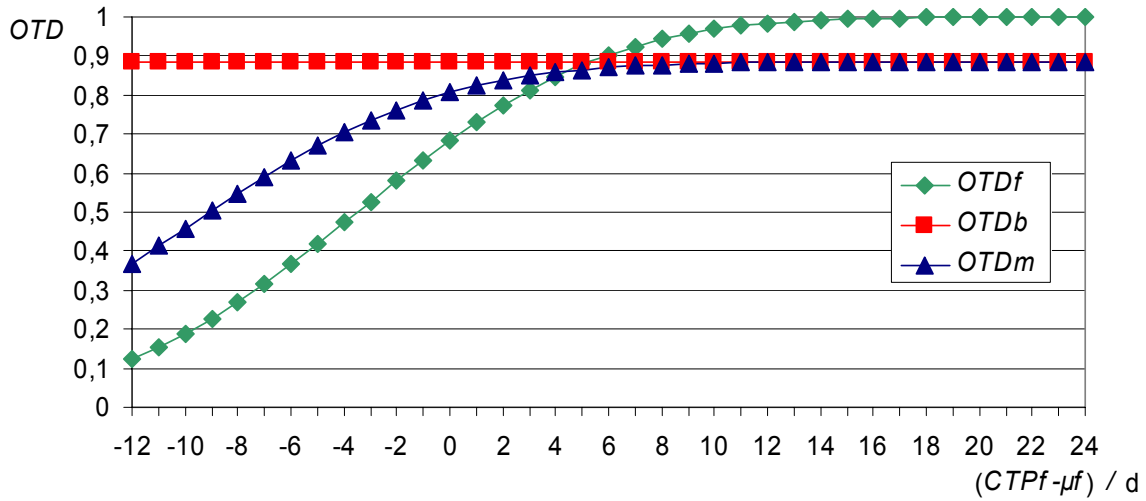


Abbildung 3.39: On-Time-Delivery für  $CTf^* \sim N(\mu f, \sigma f = 7 \text{ d})$ ,  $CTb^* \sim N(\mu b = 7 \text{ d}, \sigma b = 2 \text{ d})$  und  $CTPb = 7 \text{ d}$

Für die Berechnung einer einstufigen On-Time-Delivery wird die Angabe der Eingangsparameter somit von drei auf zwei reduziert. Dadurch kann eine Tabelle angegeben werden, aus der die einzelnen Wahrscheinlichkeiten für eine rechtzeitige Lieferung abgelesen werden können. In Tabelle 3.9 ist diese Tabelle in vier Bereiche aufgeteilt. Der untere Bereich beinhaltet alle On-Time-Delivery-Werte über 0,9. Danach folgen alle Werte zwischen 0,7 und 0,9. Im dritten Bereich sind alle Werte von 0,5 bis 0,7 enthalten. Der obere Bereich ergibt Werte unter 0,5, die darüber hinaus auch ein entgegengesetztes Verhalten bezüglich der Standardabweichung der Durchlaufzeit aufzeigen.

$\sigma / d$	1	2	3	4	5	6	7	8	9	10	11	12
-5	0,2870	0,3090	0,3420	0,3699	0,3908	0,4065	0,4186	0,4279	0,4355	0,4416	0,4467	0,4510
-4	0,4287	0,4346	0,4462	0,4560	0,4632	0,4686	0,4727	0,4759	0,4784	0,4805	0,4822	0,4836
-3	0,5715	0,5656	0,5539	0,5441	0,5369	0,5315	0,5274	0,5242	0,5216	0,5196	0,5178	0,5164
-2	0,7131	0,6911	0,6581	0,6302	0,6092	0,5935	0,5815	0,5721	0,5646	0,5585	0,5534	0,5491
-1	0,8453	0,8008	0,7519	0,7103	0,6780	0,6533	0,6342	0,6190	0,6068	0,5968	0,5884	0,5814
0	0,9430	0,8861	0,8305	0,7813	0,7413	0,7096	0,6844	0,6648	0,6478	0,6342	0,6228	0,6132
1	0,9881	0,9435	0,8916	0,8413	0,7974	0,7611	0,7315	0,7072	0,6871	0,6704	0,6563	0,6442
2	0,9988	0,9762	0,9354	0,8894	0,8456	0,8072	0,7747	0,7474	0,7245	0,7051	0,6886	0,6744
3	1	0,9916	0,9644	0,9262	0,8856	0,8475	0,8137	0,7845	0,7595	0,7380	0,7194	0,7034
4	1	0,9976	0,9819	0,9529	0,9176	0,8817	0,8482	0,8182	0,7919	0,7688	0,7488	0,7312
5	1	0,9994	0,9915	0,9713	0,9424	0,9102	0,8782	0,8485	0,8216	0,7976	0,7764	0,7576
6	1	0,9999	0,9964	0,9834	0,9610	0,9332	0,9038	0,8751	0,8484	0,8241	0,8022	0,7825
7	1	1	0,9986	0,9908	0,9744	0,9514	0,9252	0,8984	0,8725	0,8483	0,8261	0,8059
8	1	1	0,9995	0,9952	0,9837	0,9654	0,9428	0,9183	0,8937	0,8702	0,8481	0,8277
9	1	1	0,9998	0,9976	0,9899	0,9759	0,9569	0,9351	0,9123	0,8897	0,8681	0,8478
10	1	1	1	0,9989	0,9940	0,9836	0,9681	0,9491	0,9283	0,9071	0,8863	0,8664
11	1	1	1	0,9995	0,9965	0,9891	0,9768	0,9606	0,9420	0,9224	0,9026	0,8833
12	1	1	1	0,9998	0,9981	0,9929	0,9834	0,9699	0,9536	0,9357	0,9171	0,8986
13	1	1	1	0,9999	0,9990	0,9955	0,9883	0,9773	0,9632	0,9471	0,9299	0,9124
14	1	1	1	1	0,9995	0,9972	0,9919	0,9831	0,9711	0,9569	0,9412	0,9248
15	1	1	1	1	0,9997	0,9983	0,9945	0,9876	0,9776	0,9651	0,9510	0,9358

Tabelle 3.9: On-Time-Delivery-Werte für Frontend und Backend, abhängig von der Standardabweichung der Durchlaufzeit und der Differenz von Plandurchlaufzeit

Durch die Faltung mit der Wahrscheinlichkeitsfunktion der gleichverteilten Einschleusung liegt der Mittelpunkt der Fertigstellung Frontend immer bei  $CTPf-\mu f = -3,5$  Tagen. Ist der Differenzwert geringer, so liegt der Stichtag für die On-Time-Delivery vor dem Mittelpunkt der Wahrscheinlichkeitsfunktion der Fertigstellung. Wie Tabelle 3.9 zeigt, ergeben sich für die On-Time-Delivery dann Werte unter 0,5 und steigende Standardabweichungen der Durchlaufzeit führen zu steigenden On-Time-Delivery-Werten. Differenzwerte über -3,5 führen zu On-Time-Deliverys über 0,5, die bei steigenden Standardabweichungen sinken.

Die Werte in Tabelle 3.9 wurden mit Wahrscheinlichkeitsfunktionen berechnet, die weit genug im positiven Bereich der Zeitachse liegen, so dass jeweils die Summe aller Einzelwahrscheinlichkeiten genau 1 ergibt (Genauigkeit: 5 Nachkommastellen). Bei manchen konkreten Beispielen ist dies nicht der Fall. Zum Beispiel liegt die Kurve der Normalverteilung mit  $\mu = 7$  Tage und  $\sigma = 2$  Tage etwas zu nahe am Nullpunkt, so dass die Summe aller Einzelwahrscheinlichkeiten 0,99977 ergibt. Zwar ist für solche Fälle ein Korrekturfaktor (siehe Gleichung (3.30)) vorgesehen, dennoch ergibt sich ein geringer Unterschied zu den Werten der Tabelle. Mit einer Plandurchlaufzeit von  $CTP = 7$  Tagen wird eine  $OTD = 0,8860$  ermittelt, während in Tabelle 3.9 für  $CTP-\mu = 0$  und  $\sigma = 2$  eine  $OTD = 88,61\%$  berechnet wurde. Im Allgemeinen liegen Wahrscheinlichkeitsverteilungen der Normalverteilung weit genug im positiven Bereich, wenn die Standardabweichung ein fünftel des Mittelwertes nicht überschreitet.

$CTPf-\mu f / d$	$\sigma f / d$	$OTDf$	$CTPb-\mu b / d$	$\sigma b / d$	$OTDb$	$OTDm$
2	8	0,7474	2	8	0,7474	0,7157
2	8	0,7474	2	5	0,8456	0,8041
2	8	0,7474	8	8	0,9183	0,8988
2	8	0,7474	8	5	0,9837	0,9683
2	5	0,8456	2	8	0,7474	0,7382
2	5	0,8456	2	5	0,8456	0,8347
2	5	0,8456	8	8	0,9183	0,9137
2	5	0,8456	8	5	0,9837	0,9814
8	8	0,9183	2	8	0,7474	0,7402
8	8	0,9183	2	5	0,8456	0,8365
8	8	0,9183	8	8	0,9183	0,9142
8	8	0,9183	8	5	0,9837	0,9810
8	5	0,9837	2	8	0,7474	0,7468
8	5	0,9837	2	5	0,8456	0,8449
8	5	0,9837	8	8	0,9183	0,9180
8	5	0,9837	8	5	0,9837	0,9836

Tabelle 3.10: Kombination aus Eingangswerten und On-Time-Deliverys

Tabelle 3.10 zeigt beispielhaft einige Kombinationen aus Eingangswerten und On-Time-Deliverys für Frontend, Backend und das Manufacturing-Ergebnis. Zur eindeutigen Beschreibung des gesamten Produktionsvorgangs genügen bereits zwei Werte aus Frontend und Backend.

Der Einfluss der Standardabweichungen der Durchlaufzeiten zeigt eine weitere Eigenschaft. Je größer die Standardabweichung einer Durchlaufzeit in Frontend oder Backend ist, desto näher rückt die On-Time-Delivery Manufacturing an die entsprechende On-Time-Delivery Frontend oder Backend. Abbildung 3.40 zeigt für unterschiedliche Standardabweichungen der Durchlaufzeit im Frontend die On-Time-Delivery Frontend und die sich mit der On-Time-Delivery Backend ergebende On-Time-Delivery Manufacturing. Für die größeren Werte der Standardabweichung der Durchlaufzeit Frontend liegen die Werte dicht über den Werten der On-Time-Delivery Frontend. Bei den niedrigeren Standardabweichungen wird die On-Time-Delivery Manufacturing von der On-Time-Delivery Backend begrenzt.

Damit wird wiederum die Forderung nach einer gleichförmigen Produktion bestärkt. Alle Schwankungen sollten möglichst gering sein. Um dies zu erreichen, ist die Vier-Partner-Analyse mit ihren Optimierungsstrategien (siehe Kapitel 2.5.4) ein passendes Mittel. Da der Nutzen einer verbesserten On-Time-Delivery nicht leicht in monetärer Form auszudrücken ist, ist ein Kosten-Nutzen-Vergleich für mögliche Maßnahmen oftmals schwer durchzuführen.

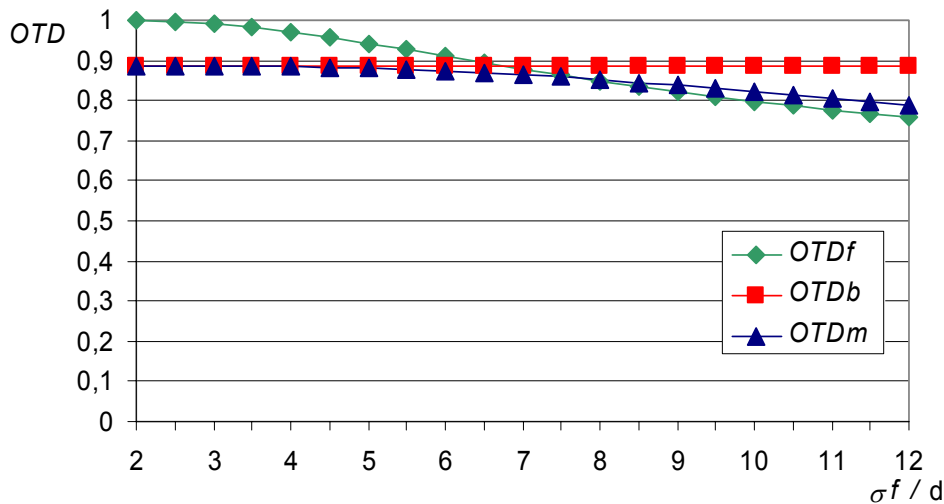


Abbildung 3.40: On-Time-Delivery für  $CTf^* \sim N(\mu f = 44 \text{ d}, \sigma f)$ ,  $CTPf = 49 \text{ d}$ ,  $CTb^* \sim N(\mu b = 7 \text{ d}, \sigma b = 2 \text{ d})$  und  $CTPb = 7 \text{ d}$

In Abbildung 3.41 wurde ein geringerer Abstand (nur ein Tag anstatt fünf Tage) zwischen Plandurchlaufzeit Frontend und mittlerer Durchlaufzeit Frontend gewählt, um geringere Werte der On-Time-Delivery Frontend zu erreichen und den Einfluss der Standardabweichung der Durchlaufzeit Frontend auf die Größe der On-Time-Delivery Manufacturing besser darzustellen. Hier ist zu erkennen, dass die On-Time-Delivery Manufacturing immer näher an die On-Time-Delivery Frontend im Vergleich zur On-Time-Delivery Backend rückt, je größer die Standardabweichung der Durchlaufzeit Frontend wird.



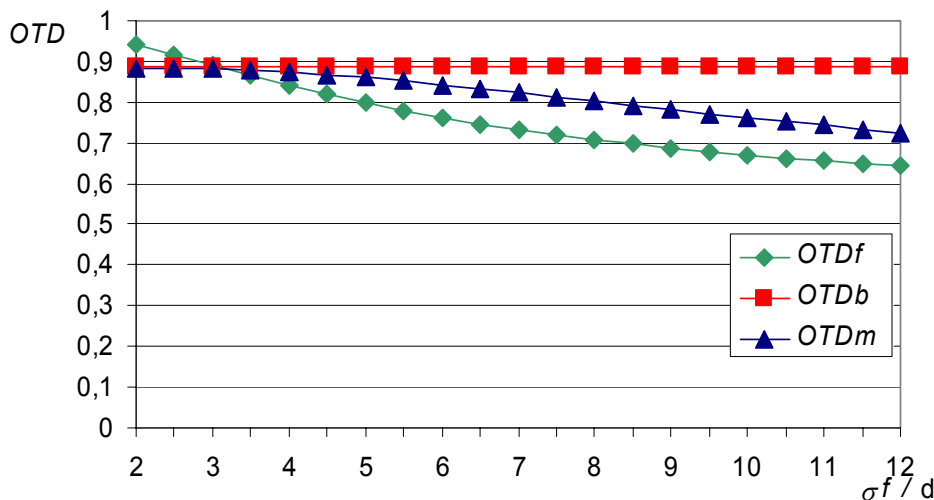


Abbildung 3.41: On-Time-Delivery für  $CTF^* \sim N(\mu_f = 41 \text{ d}, \sigma_f)$ ,  $CTPf = 42 \text{ d}$ ,  $CTb^* \sim N(\mu_b = 7 \text{ d}, \sigma_b = 2 \text{ d})$  und  $CTPb = 7 \text{ d}$

Während in Abbildung 3.42 eine konstante Standardabweichung der Durchlaufzeit Frontend von  $\sigma_f = 10$  Tagen verwendet wurde, beträgt sie in Abbildung 3.43  $\sigma_f = 4$  Tage. Deutlich zu erkennen ist der schnellere Übergang von geringen zu höheren Werten der On-Time-Delivery Frontend bei der geringeren Standardabweichung der Durchlaufzeit Frontend. Die Werte der On-Time-Delivery Manufacturing liegen jedoch bei der geringeren Standardabweichung nicht nur relativ zur On-Time-Delivery Frontend näher an den Werten der On-Time-Delivery Backend, sondern zum überwiegenden Teil auch absolut betrachtet. Dies spricht dafür, im Frontend eine möglichst geringe Standardabweichung der Durchlaufzeit zu erreichen.

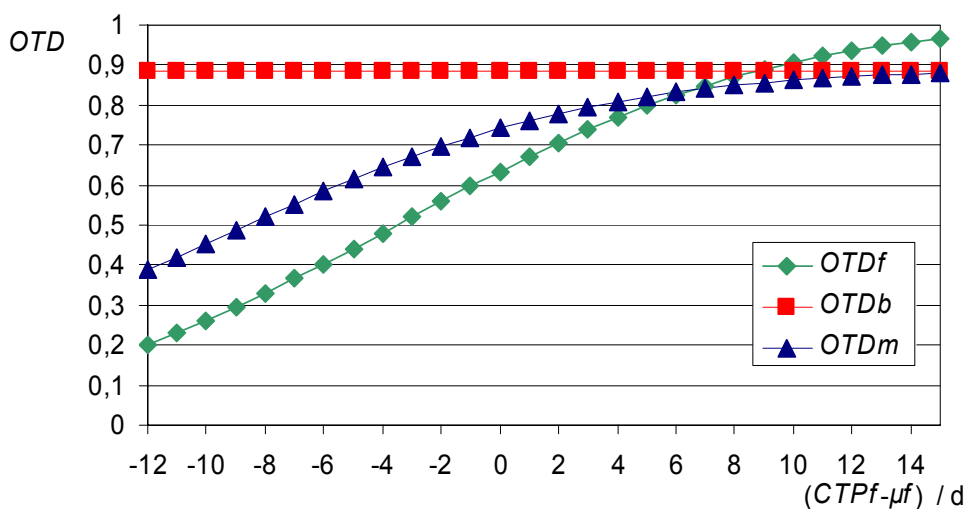


Abbildung 3.42: On-Time-Delivery für  $CTF^* \sim N(\mu_f, \sigma_f = 10 \text{ d})$ ,  $CTb^* \sim N(\mu_b = 7 \text{ d}, \sigma_b = 2 \text{ d})$  und  $CTPb = 7 \text{ d}$

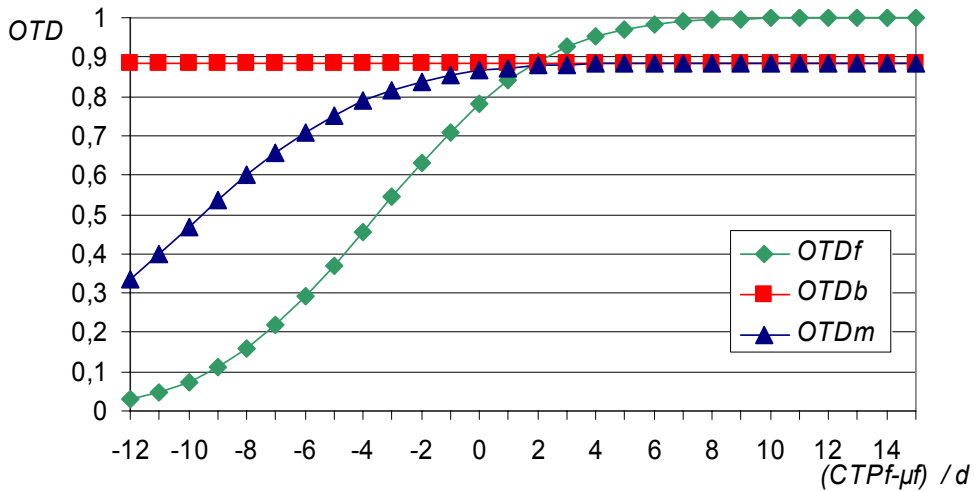


Abbildung 3.43: On-Time-Delivery für  $CTf \sim N(\mu f, \sigma f = 4 \text{ d})$ ,  $CTb \sim N(\mu b = 7 \text{ d}, \sigma b = 2 \text{ d})$  und  $CTPb = 7 \text{ d}$

Hat die Durchlaufzeit des Backends eine Standardabweichung von  $\sigma b = 7$  Tagen, rückt die On-Time-Delivery Manufacturing näher an die On-Time-Delivery Backend (siehe Abbildung 3.44). Der Effekt wird noch verstärkt, wenn die Standardabweichung der Durchlaufzeit im Frontend wie in Abbildung 3.45 auf vier Tage reduziert wird.

Zu einer hohen Standardabweichung der Durchlaufzeit im Backend zu raten, erscheint allerdings nicht sinnvoll, da Variabilitäten immer negative Auswirkungen auf Performance-Parameter wie Auslastung der Maschinen, Warenbestand (Work in Progress) etc. bewirken.

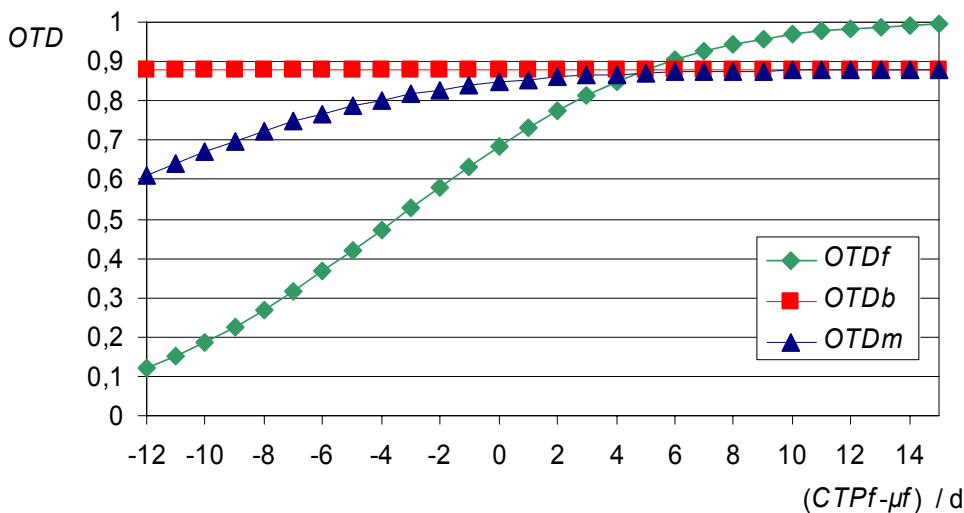


Abbildung 3.44: On-Time-Delivery für  $CTf^* \sim N(\mu f, \sigma f = 7 \text{ d})$ ,  $CTb^* \sim N(\mu b = 30 \text{ d}, \sigma b = 7 \text{ d})$  und  $CTPb = 35 \text{ d}$

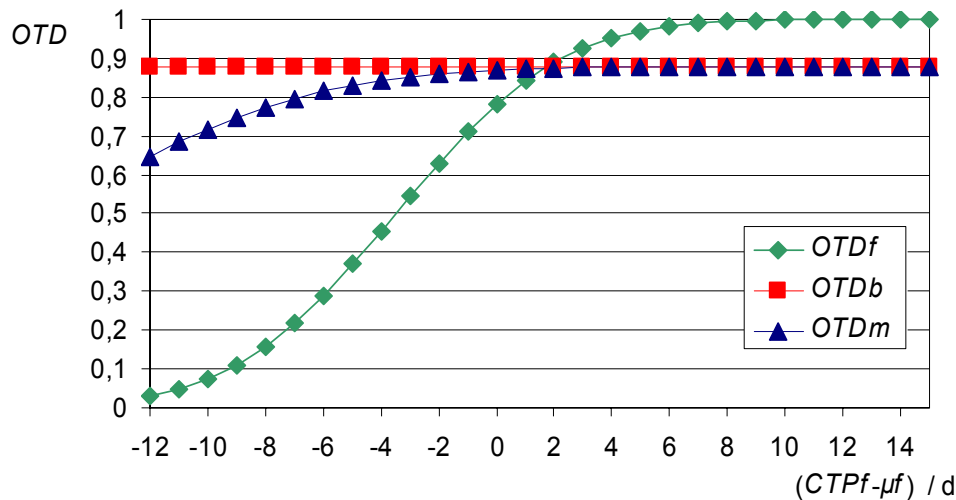


Abbildung 3.45: On-Time-Delivery für  $CTf \sim N(\mu f, \sigma f = 4 \text{ d})$ ,  
 $CTb \sim N(\mu b = 30 \text{ d}, \sigma b = 7 \text{ d})$  und  $CTPb = 35 \text{ d}$

## 3.5 On-Time-Delivery mit Due Date Daily und Release Backend

### 3.5.1 Voraussetzungen

Eine Veränderung der mittleren Durchlaufzeit im Frontend kann durch eine entsprechende Veränderung der Plandurchlaufzeit Frontend kompensiert werden. Da die Plandurchlaufzeiten aber nur in Wochen bzw. in Vielfachen von 7 Tagen angegeben werden können, ist eine solche Kompensation im Modell der „On-Time-Delivery mit Due Date Weekly und Release Backend“ nur in Schritten von 7 Tagen möglich.

Durchlaufzeiten bzw. Mittelwerte der Wahrscheinlichkeitsverteilungen von Durchlaufzeiten und tatsächliche Durchlaufzeiten sind Werte, die in Tagen angegeben werden. Dass sie genau das Vielfache einer Woche betragen, ist dann nur zufällig. Wie Abbildung 3.35 und Abbildung 3.37 aus Kapitel 3.4.5 zeigen, kann daher mit einer wochenfeinen Vorgabe der Plandurchlaufzeit Frontend nicht für jede Durchlaufzeit in Verbindung mit der gegebenen On-Time-Delivery Backend eine optimale On-Time-Delivery Manufacturing erreicht werden. Z.B. führt für eine Frontend-Durchlaufzeit, deren Wahrscheinlichkeitsfunktion einen Mittelwert von  $\mu f = 40$  Tagen und eine Standardabweichung von  $\sigma f = 7$  Tagen besitzt, bei gegebenen Backend-Werten weder eine Plandurchlaufzeit von 49 noch von 42 Tagen zu einer optimalen Ausnützung der Ressourcen. In beiden Fällen wird die jeweils gute Performance einer Stufe durch die andere Stufe im Gesamtergebnis verringert. Es erscheint also sinnvoll, auch die Plandurchlaufzeiten tagesfein anzugeben, um eine genauere Steuerung und somit eine bessere Bewertung einer Halbleiterfab zu erreichen.

Wird die Plandurchlaufzeit in Tagen angegeben, so stellt sich die Frage, wie daraus die Lieferwoche bzw. der Stichtag des Liefertermins ermittelt wird. Aufgrund der Einschleuswoche ist die Due Date um 7 Tage größer als die Plandurchlaufzeit, die bisher ein Vielfaches von 7 Tagen war (siehe Abbildung 3.46 a). Somit lag die Due Date immer am letzten Tag einer so genannten Lieferwoche. Dies entspricht dem Stichtag an dem der CLIP

gemessen wird. Der *CLIP* gibt dabei an, zu welchem Prozentsatz die für die Lieferwoche bestellte Menge an Chips bis zum letzten Tag der Lieferwoche geliefert wurde. Die Messung des *CLIPs* am Ende einer Lieferwoche ist die Ursache für die Angabe der Plandurchlaufzeit in Wochen.

Wird die bisherige *CLIP*-Messung mit Stichtag am Ende einer Lieferwoche für alle Bestellungen der Lieferwoche beibehalten, ist die Angabe von tagesfeinen Plandurchlaufzeiten überflüssig. Dies wäre dann nur eine Mogelpackung, die niedrigere Plandurchlaufzeiten vortäuscht, in Wirklichkeit aber mit effektiven Plandurchlaufzeiten arbeitet, die auf das nächste Vielfache von 7 aufgerundet sind (siehe Abbildung 3.46 b).

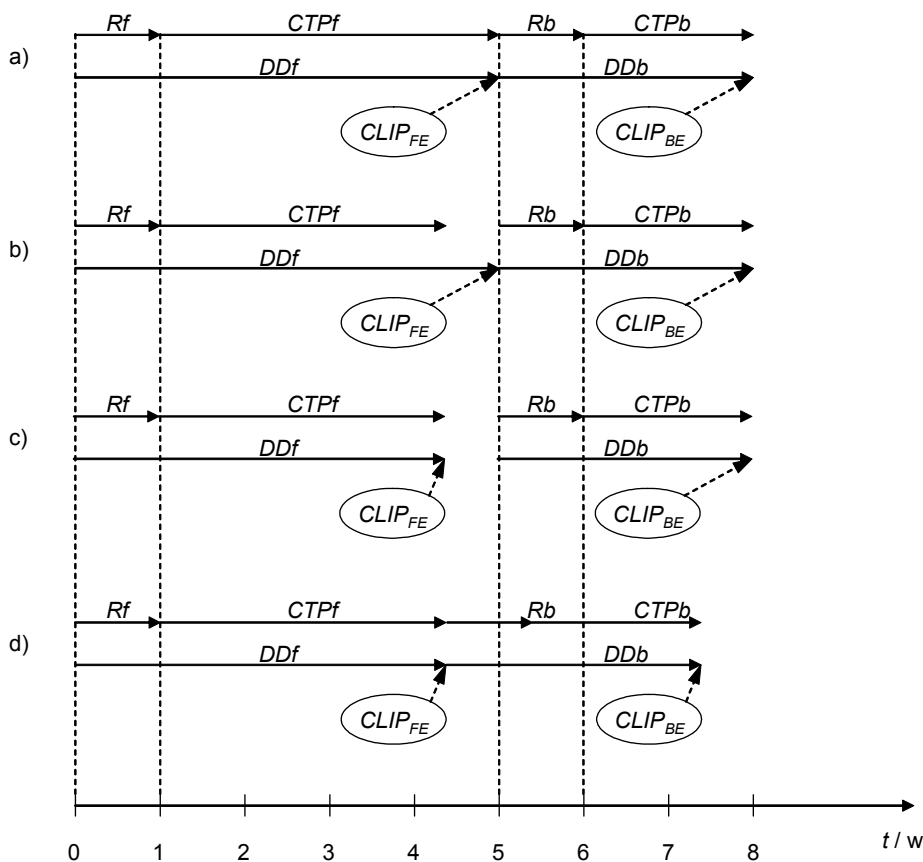


Abbildung 3.46: Wird von wochenfeinen Plandurchlaufzeiten im Frontend a) zu tagesfeinen Plandurchlaufzeiten übergegangen, so müssen sich auch der Stichtag sowie die Einschleusung ins Backend vom Wochenraster lösen d)

Der Übergang zu tagesfeinen Plandurchlaufzeiten kann nur dann eine genauere Steuerung und Bewertung der Halbleiterfab bewirken, wenn der Stichtag für die *CLIP*-Messung der einzelnen Bestellungen jeweils ein Tag, unabhängig vom Wochenraster, ist. Da die Einschleusung zwischen dem ersten und siebten Tag der Einschleuswoche erfolgen kann, muss sich der Stichtag weiterhin aus der Plandurchlaufzeit plus 7 Tage ergeben. Die On-Time-Delivery und der *CLIP* einer Produktionsstufe müssen sich aus der Erfüllung der individuellen Stichtage jeder einzelnen Bestellung ermitteln. Ein Überschreiten dieses Stichtages wird als *CLIP*- bzw. *OTD*-Verletzung gewertet, auch wenn die Lieferung nach

dem Stichtag, aber noch innerhalb der zu Grunde gelegten Lieferwoche erfolgt (siehe Abbildung 3.46 d).

Die On-Time-Delivery mit Due Date Daily und Release Backend kann somit auch als Methode mit gleitenden Lieferwochen interpretiert werden, deren Anfang und vor allem ihr Ende unabhängig von den Wochentagen sind. Bleibt die Durchlaufzeit unter der Plandurchlaufzeit, so wird jedes Los unabhängig von seiner Einschleusung rechtzeitig fertig gestellt.

Bezüglich der Gesamtperformance, also der On-Time-Delivery Manufacturing bzw. des Service Levels, bewirken tagesfeine Plandurchlaufzeiten im Frontend nur dann eine Verbesserung, wenn auch die Einschleusung in das Backend zu gleitenden Einschleuswochen übergeht (siehe Abbildung 3.46 d). Würde ein Los, das rechtzeitig zu seinem Stichtag innerhalb einer beliebigen Woche fertig gestellt wird, auf eine Einschleuswoche warten müssen, die erst zur nächsten vollen Woche beginnt, dann wäre zwar die On-Time-Delivery Frontend genauer ermittelt worden, aber die On-Time-Delivery Manufacturing hätte wieder den gleichen Wert und die gesamte Plandurchlaufzeit wäre die Gleiche (siehe Abbildung 3.46 c).

Das bedeutet, dass im Frontend für jedes Los eine individuelle Lieferwoche bzw. ein individueller Liefertag existiert, an den direkt eine individuelle Einschleuswoche in das Backend folgt.

#### 3.5.2 Modell

Das Modell zur Berechnung der On-Time-Delivery mit Due Date Daily und Release Backend ist identisch mit dem Modell zur Berechnung der On-Time-Delivery mit Due Date Weekly und Release Backend. Der Übergang zu tagesfeinen Plandurchlaufzeiten wird allein durch die Eingabe entsprechender Werte der Plandurchlaufzeit erreicht. Die Due Date Frontend ergibt sich aus der tagesfeinen Plandurchlaufzeit für das Frontend plus sieben Tage.

Die in der Praxis durchgeführte *CLIP*-Ermittlung muss dieser Vorgehensweise angepasst werden. Es ist nach wie vor sinnvoll, einen *CLIP*-Wert zu ermitteln, der sich auf eine Lieferwoche bezieht. Dieser muss allerdings aus einzelnen *CLIP*-Werten hervorgehen, die für jeden einzelnen Tag der Lieferwoche den Erfüllungsgrad der bestellten Chipmenge bewerten. Bestellungen, die nach ihrer Due Date aber vor Ende der Lieferwoche geliefert werden, dürfen nicht als rechtzeitig geliefert betrachtet werden, sonst würde kein Unterschied zum ursprünglichen Modell bestehen. Ebenso ist es notwendig, dass die Einschleuswoche des Backends individuell für jedes Los direkt auf den Liefertag des Frontends folgt.

#### 3.5.3 Bewertung

Die Auswahl einer passenden tagesfeinen Plandurchlaufzeit ermöglicht für jedes Produkt die Realisierung sowohl einer möglichst hohen On-Time-Delivery Manufacturing als auch einer möglichst geringen Gesamtdurchlaufzeit.

Für oben genanntes Beispiel sind die Auswirkungen unterschiedlicher Plandurchlaufzeiten Frontend in Tabelle 3.11 und Abbildung 3.47 aufgeführt. Während eine Plandurchlaufzeit von  $CTPf = 42$  Tage zu einer relativ geringen Liefertreue von  $OTDm = 83,70\%$  nach einer relativ geringen Durchlaufzeit von  $DDm = 63$  Tage führt, wird für eine Plandurchlaufzeit von  $CTPf = 49$  Tage nach einer relativ langen Durchlaufzeit von  $DDm = 70$  Tage eine relativ hohe Liefertreue von  $OTDm = 88,05\%$  erreicht. Hier erscheint der Kompromiss mit einer

Plandurchlaufzeit von  $CTPf = 46$  Tage sinnvoll. Mit einer nur einen Prozentpunkt geringeren Liefertreue gegenüber der Plandurchlaufzeit von  $CTPf = 49$  Tage kann die Durchlaufzeit um drei Tage verkürzt werden. Im Vergleich zur Plandurchlaufzeit von  $CTPf = 42$  Tage wird die Durchlaufzeit zwar um 4 Tage verlängert, die Liefertreue kann aber gleichzeitig um drei Prozentpunkte verbessert werden.

$\mu_f$ / d	$\sigma_f$ / d	$CTPf$ / d	$DDf$ / d	$OTDf$	$\mu_b$ / d	$\sigma_b$ / d	$CTPb$ / d	$DDb$ / d	$OTDb$	$DDm$ / d	$OTDm$
40	7	42	49	0,7747	7	2	7	14	0,8860	63	0,83696
40	7	43	50	0,8137	7	2	7	14	0,8860	64	0,84847
40	7	44	51	0,8482	7	2	7	14	0,8860	65	0,85769
40	7	45	52	0,8782	7	2	7	14	0,8860	66	0,86495
40	7	46	53	0,9038	7	2	7	14	0,8860	67	0,87058
40	7	47	54	0,9252	7	2	7	14	0,8860	68	0,87488
40	7	48	55	0,9428	7	2	7	14	0,8860	69	0,8781
40	7	49	56	0,9569	7	2	7	14	0,8860	70	0,88048

Tabelle 3.11: Beeinflussung der On-Time-Delivery durch die Plandurchlaufzeit Frontend

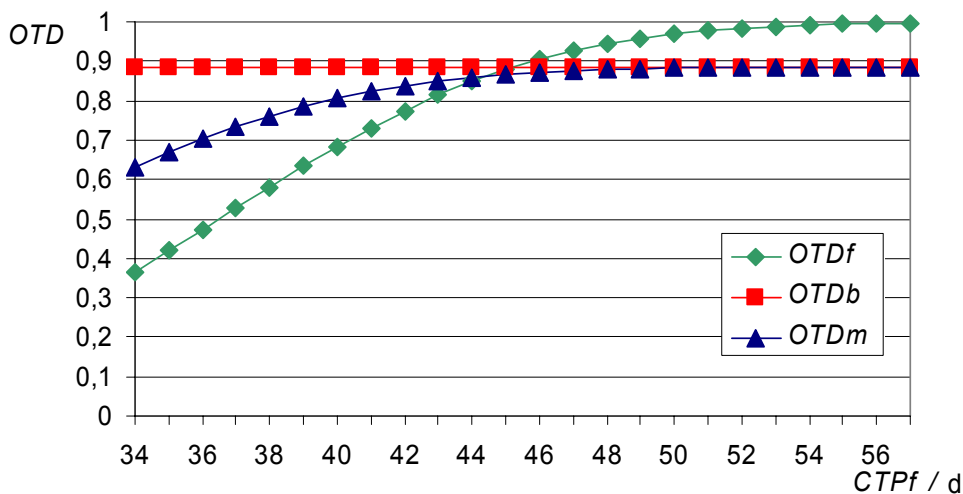


Abbildung 3.47: On-Time-Delivery für  $CTf^* \sim N(\mu_f = 40 \text{ d}, \sigma = 7 \text{ d})$ ,  $CTb^* \sim N(\mu = 7 \text{ d}, \sigma = 2 \text{ d})$  und  $CTPb = 7 \text{ d}$

In Abbildung 3.47 sind die Werte der On-Time-Deliverys über der Plandurchlaufzeit Frontend aufgetragen. Diese Darstellung erleichtert ein schnelles Auffinden der optimalen Plandurchlaufzeit. Bei Veränderungen des Mittelwertes der Durchlaufzeit Frontend müsste die Graphik jedoch neu erstellt werden. Deshalb bietet es sich an, auch hier eine Darstellung

über der Differenz aus Plandurchlaufzeit und Mittelwert der Durchlaufzeit  $CTPf-\mu$  zu wählen. So ist Abbildung 3.47 ein Ausschnitt aus Abbildung 3.39. Ob der Mittelwert erhöht oder die Plandurchlaufzeit um den gleichen Betrag verringert wird, oder umgekehrt, ist gleichbedeutend und hat die gleiche Wirkung.

Prinzipiell bewirken längere Plandurchlaufzeiten immer höhere On-Time-Delivery-Werte. Da die  $CTP$  eine bestimmbare Größe ist, birgt das auch immer die Gefahr der „Augenwischerei“. Gerade im Frontend muss hinsichtlich des Verhaltens des Backends ein sinnvoller Kompromiss gefunden werden. Für dieses Beispiel gilt: Eine Verlängerung der Plandurchlaufzeit zu Werten über 47 Tagen bewirkt zwar eine Verbesserung der On-Time-Delivery Frontend, führt aber für die Gesamtperformance nur zu einer Verlängerung der Durchlaufzeit, ohne die On-Time-Delivery Manufacturing zu beeinflussen und stellt daher nur eine Verschwendung des Performance-Potentials dar.

## 3.6 On-Time-Delivery mit Due Date Daily ohne Release Backend

### 3.6.1 Voraussetzungen

Die Einschleusung während einer Einschleuswoche hat den Effekt, dass Lose, die am Ende der Woche eingeschleust werden, genau mit ihrer Plan-Durchlaufzeit produziert werden müssen, um nicht eine  $CLIP$ -Verletzung zu erzeugen. Lose, deren Durchlaufzeit größer als die Plan-Durchlaufzeit ausfällt, haben dann eine Chance, rechtzeitig fertig zu werden, wenn ihre größere Durchlaufzeit durch eine frühe Einschleusung kompensiert werden kann.

Um diese Planungsungenauigkeit zu vermindern, könnten drei Methoden eingesetzt werden:

1. Der Stichtag wird erst nach erfolgter Einschleusung ermittelt und ergibt sich aus der Addition von Plandurchlaufzeit und tatsächlicher Einschleusung.
2. Da der Erwartungswert der Einschleusung (bei gleich verteilter Einschleusung innerhalb der Woche) in der Mitte des vierten Tages liegt, ergibt sich der Stichtag aus der Plan-Durchlaufzeit plus 3,5 Tage.
3. Eine Einschleusung wird von vornherein direkt geplant, die Einschleuswoche entfällt.

Die erste Methode erscheint zunächst am genauesten, hat jedoch eine andere Interpretation der somit errechneten Liefertreue zur Folge. Die Zufallsvariable Einschleusungstag hat keinen Einfluss mehr auf den  $CLIP$  bzw die On-Time-Delivery; es wird nur noch die eigentliche Durchlaufzeit bewertet. Dazu müssten aber keine Berechnungen zur Liefertreue durchgeführt werden. Zur Bewertung der Durchlaufzeiten genügen die Durchlaufzeiten an sich. Unter der Annahme, die Durchlaufzeiten seien normalverteilt, ergibt sich die Wahrscheinlichkeit  $P$ , mit der die tatsächliche Durchlaufzeit eines Loses kleiner oder gleich einem vorgegebenen Liefertermin ist, aus der Verteilungsfunktion der Standardnormalverteilung (siehe Tabelle 3.12). Dieser Liefertermin soll, um Verwechslungen auszuschließen, mit  $lt$  bezeichnet werden. Die Quantile  $z$  errechnen sich aus der Differenz von Liefertermin  $lt$  und Mittelwert  $\mu$  der Durchlaufzeit, gemessen in Vielfachen der Standardabweichung  $\sigma$  der Durchlaufzeit.

$$z = \frac{lt - \mu}{\sigma} \quad (3.62)$$

$z$	$P(CT \leq lt)$ $= P\left(\frac{CT - \mu}{\sigma} \leq z\right)$
0,00	0,50
0,67	0,75
0,84	0,80
1	0,84
1,28	0,90
1,64	0,95
1,75	0,96
1,88	0,97
2	0,977
2,05	0,98
2,33	0,99

Tabelle 3.12: Bewertung der Durchlaufzeit mit Hilfe der Verteilungsfunktion der Normalverteilung

Sollen also z.B. 95% der Lose einen Liefertermin  $lt$  nicht überschreiten, so darf  $z$  den Wert 1,65 nicht überschreiten.

Die Wahrscheinlichkeit dafür, dass die Durchlaufzeit kleiner als die Plandurchlaufzeit ist, ist allerdings immer geringer als die Wahrscheinlichkeit der rechtzeitigen Lieferung unter Berücksichtigung der Einschleusung. Denn während die Due Date genau um sieben Tage gegenüber der Plandurchlaufzeit verschoben ist, sind die Werte der Verteilungsfunktion der Fertigstellung erst für  $F_C(k) = 1$  um sieben Tage gegenüber der Verteilungsfunktion der Durchlaufzeit verschoben. D.h. die Faltung einer Wahrscheinlichkeitsfunktion mit einer Rechteckfunktion  $[0; a]$  führt zur Verbreiterung der Wahrscheinlichkeitsfunktion um die Breite der Rechteckfunktion. Der Mittelwert der neuen Funktion ist bereits um die Hälfte dieser Breite (Mittelwert der Rechteckfunktion) verschoben, die vollen 7 Tage sind aber erst am „Ende“ der neuen Funktion erkennbar. Liegt der Stichtag nicht an diesem Ende, so bewirkt er eine Vergrößerung der zu errechnenden Wahrscheinlichkeit.

$$P(C^* \leq DD) > P(CT^* \leq CTP) \tag{3.63}$$

Gegen diese Methode spricht jedoch, dass einzelne Durchlaufzeiten aus Sicht der Kunden uninteressant sind. Für den Kunden ist entscheidend, wann er die Lieferung seiner Bestellung erhält. Das gilt sowohl für das Backend als „Kunde“ des Frontends wie auch für den Kunden als Käufer des Endproduktes.

Der zweiten Methode liegt eine Mittelwertbetrachtung zu Grunde. Mittelwertbetrachtungen jedoch verführen zu Vorgaben, die immer dann, wenn Variationen auftreten, mit Sicherheit schlechte Ergebnisse zur Folge haben. Würde ein Los genau mit seiner Plan-Durchlaufzeit



gefertigt, so wird es bei Anwendung der zweiten Methode nur dann rechtzeitig fertig gestellt werden können, wenn es (zufällig) in der ersten Hälfte der Woche eingeschleust worden wäre. Zu spät fertig gestellte Lose können bei einer *CLIP*-Bewertung nicht durch zu früh gelieferte kompensiert werden. Generell gilt: Ohne Berücksichtigung von Schwankungen bzw. Varianzen sind Mittelwerte immer vorsichtig einzusetzen.

Bei der dritten Methode muss zwischen Frontend und Backend unterschieden werden.

Die Einschleusung eines Loses ins Frontend kann von außen (durch den Vertrieb) nur wochenfein vorgegeben werden. Die tagesfeine Einschleusung wird von der Produktion vorgenommen. Dabei ist es sinnvoller, produktionstechnische Aspekte zu berücksichtigen, als die Durchlaufzeit durch einen fest vorgegebenen Einschleustermine vorhersagbarer zu gestalten. Der Einschleuszeitpunkt hängt dabei vom aktuellen Produktmix ab. Die Einschleusung von Losen erfolgt u.a. nach folgenden Einschleusregeln: Lose, die gleiche Masken benötigen, sollen nicht zusammen gestartet werden. Lose, die einen gleichen Ofenprozess benötigen, sollen so gestartet werden, dass sie gleichzeitig diesen Ofenprozess erreichen. Für das Frontend kommt ein Wegfall der Einschleuswoche somit nicht in Frage.

Werden die Plan-Durchlaufzeit und der geplante Liefertermin im Frontend tagesfein angegeben, so könnte auch die Einschleusung der Weiterverarbeitung tagesfein vorgegeben werden. Einschleus- und Lieferwochen führen zu Zwischenlagerzeiten, die die gesamte Durchlaufzeit bzw. die Lieferzeit unnötig erhöhen. Im Modell „On-Time-Delivery Weekly mit Release Backend“ wird erst nach Ablauf der Lieferwoche mit der planmäßigen Weiterverarbeitung der in der Lieferwoche fertig gestellten Lose begonnen. Natürlich können nicht alle Lose, die innerhalb einer Woche im Frontend fertig gestellt wurden, gleichzeitig in das Backend eingeschleust werden. Deshalb ist die Aufteilung innerhalb einer Einschleuswoche notwendig. Liegen die Liefertage, wie im Modell „On-Time-Delivery Daily mit Release Backend“ auch innerhalb des Wochenrasters, verteilen sie sich von vornherein gleichmäßig über jede einzelne Woche. Lose, die zu ihrem Liefertag fertig gestellt sind, könnten somit sofort nach dem Liefertag planmäßig in die Weiterverarbeitung eingeschleust werden. Wartezeiten hängen dann nur noch von der Durchlaufzeit in Bezug auf die Plan-Durchlaufzeit ab und nicht mehr von der Plan-Durchlaufzeit in Bezug zum Wochenraster.

Andererseits stellt die Einschleuswoche im Backend ebenso wie die ursprüngliche Lieferwoche im Frontend eine Puffergröße dar, die unvorhergesehene Unregelmäßigkeiten ausgleichen kann. Die Puffergröße, die für ein Los zur Verfügung steht, hängt aber sowohl vom Tag der Einschleusung als auch von der Durchlaufzeit sowie der Plan-Durchlaufzeit ab. Größere Plan-Durchlaufzeiten für jedes Produkt können auch einen Pufferzweck erfüllen, würden aber für alle Lose gleichermaßen zur Verfügung stehen, nicht nur für früh eingeschleuste und im Frontend schnell bearbeitete Lose.

Die Ausführungen haben gezeigt:

1. Eine Kennzahl, die die Liefertreue bewertet, muss immer alle Faktoren, die in die Lieferzeit eingehen, enthalten. D.h. wenn eine Einschleuswoche existiert, dann muss sie auch in *CLIP* bzw. *OTD* einfließen.
2. Der Liefertermin als maximal erlaubte Lieferzeit muss sich aus maximal erlaubter Durchlaufzeit (= Plandurchlaufzeit) und maximal erlaubter Wartezeit bis zur Einschleusung (im Frontend = 7 Tage) ergeben.
3. Die Einschleuswoche im Frontend kann aus produktionstechnischen Gründen nicht entfallen.
4. Die Angabe der Plan-Durchlaufzeiten in Tagen ermöglicht einen tagfeinen Liefertermin und somit eine unmittelbar folgende Einschleusung zur Weiter-

verarbeitung. Wartezeiten können somit verringert und die Lieferzeit bis zum Kunden verkürzt werden.

Im Modell „On-Time-Delivery Daily ohne Release Backend“ soll daher die Einschleuswoche im Backend entfallen. Der verloren gegangene Puffereffekt soll durch eine längere Plan-Durchlaufzeit berücksichtigt werden.

### 3.6.2 Modell

Das Modell „On-Time-Delivery Daily mit Release Backend“ kann für das Modell „On-Time-Delivery Daily ohne Release Backend“ zum großen Teil übernommen werden. Da die Einschleuswoche ins Backend wegfällt, ist für die Berechnung der On-Time-Delivery Backend nur noch die Durchlaufzeit im Backend zu berücksichtigen. Die Due Date Backend ist gleich der Plandurchlaufzeit im Backend.

$$DDb = CTPb \tag{3.64}$$

$$OTDb = P(CTb^* \leq DDb) = \frac{F_{CTb^*}(k = DDb / \Delta t)}{F_{CTb^*}(k_{\max})} \tag{3.65}$$

Die Berechnung der Wahrscheinlichkeitsfunktion der Zufallsvariablen für die Einschleusung in das Backend  $Rb^*$  wird nun genauso ermittelt wie für die Einschleusung in den Kundentransit. Sie ist in folgender Gleichung beschrieben.

$$f_{Rb^*}(k) = \pi_{\frac{DDf}{\Delta t} + 1}(f_{Cf^*}(k)) \tag{3.66}$$

Wie Abbildung 3.48 zeigt, gibt es keine Einschleuswoche des Backends mehr. Die Wahrscheinlichkeitsfunktion der Einschleusung Backend entspricht zum Zeitpunkt der  $DDf$  der  $OTDf$  und ist danach identisch mit der Wahrscheinlichkeitsfunktion der Fertigstellung Frontend. Um im Vergleich zum Beispiel aus Abbildung 3.31 nicht zu schlechte OTD-Werte zu erhalten, wird versucht, den Wegfall der Einschleuswoche Backend mit einer verlängerten Plandurchlaufzeit des Backends zu kompensieren. Statt sieben Tage beträgt sie hier zehn Tage. Dadurch erhöht sich zunächst die On-Time-Delivery Backend von 88,60% auf 93,32%. Dies reicht jedoch nicht aus, um gleiche  $OTDm$ - und  $OTDo$ -Werte zu erzielen. Die On-Time-Delivery Manufacturing fällt von 78,55% auf 72,73% und die On-Time-Delivery Over All von 84,01% auf 77,29%. Dafür hat sich jedoch die gesamte Durchlaufzeit verringert. Die  $DDm$  beträgt nun 59 Tage (zuvor 63 Tage) und die  $DDo$  liegt bei 63 Tagen (zuvor 67 Tage). Die gesamte Kurve der Fertigstellung Delivery liegt bei geringeren Durchlaufzeiten.

Für die Realität bedeutet dies, dass die Inventarliste der Die Bank viel dynamischer wird und die einzelnen Produkte sich dort viel kürzer aufhalten.

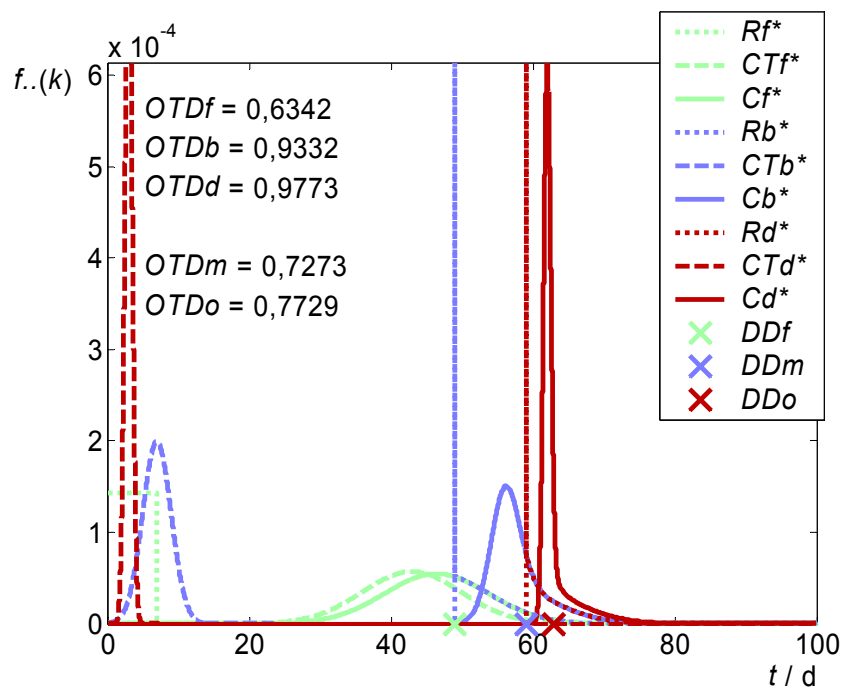


Abbildung 3.48: OTD ohne Release Backend für  $CTf^* \sim N(\mu f = 43 \text{ d}, \sigma f = 7 \text{ d})$ ;  
 $CTPf = 42 \text{ d}$ ;  $CTb^* \sim N(\mu b = 7 \text{ d}, \sigma b = 2 \text{ d})$ ;  $CTPb = 10 \text{ d}$ ;  
 $CTd^* \sim N(\mu d = 3 \text{ d}, \sigma d = 0,5 \text{ d})$ ;  $CTPd = 4 \text{ d}$

### 3.6.3 Bewertung

Wird das Modell „On-Time-Delivery Daily ohne Release Backend“ mit den gleichen Vorgabewerten betrieben wie das Modell „On-Time-Delivery Daily mit Release Backend“, so ergeben sich deutlich geringere Werte für die On-Time-Delivery Backend, da die Pufferwirkung der Einschleuswoche und die dadurch längere Due Date Backend wegfällt.

So ergeben die gleichen Eingabewerte wie in Abbildung 3.39 die in Abbildung 3.49 dargestellte On-Time-Delivery Backend von  $OTDb = 0,5$  im Gegensatz zur ursprünglichen  $OTDb = 0,89$ . Die resultierende On-Time-Delivery Manufacturing ist entsprechend reduziert.

Um den verlorenen Puffereffekt der Einschleuswoche zu kompensieren, muss die Plandurchlaufzeit Backend entsprechend vergrößert werden. Dazu bietet sich eine  $CTPb = 10$  Tage an, da sie eine  $OTDb = 0,9332$  ergibt, die die geringste  $OTDb$  über dem ursprünglichen Wert darstellt (siehe Tabelle 3.13).

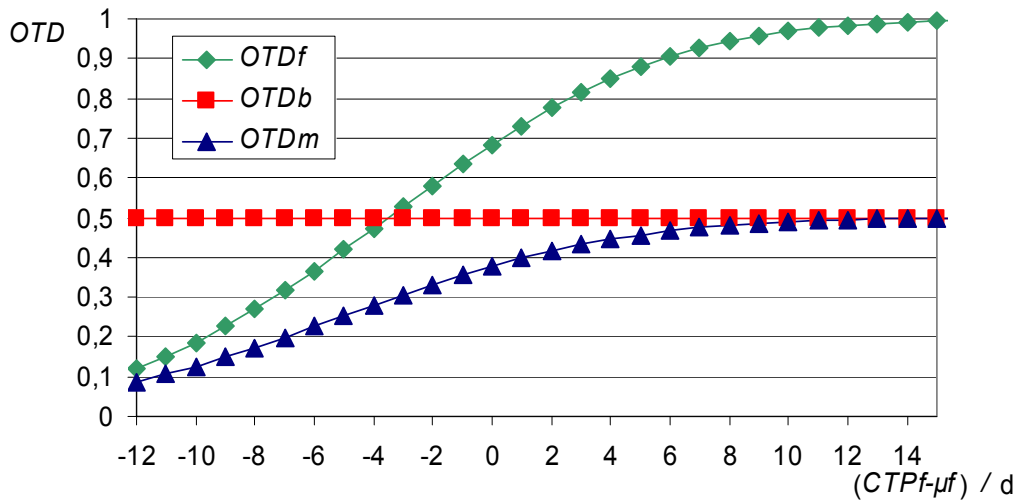


Abbildung 3.49: On-Time-Delivery Daily ohne Release Backend für  $CTf^* \sim N(\mu f, \sigma f = 7 \text{ d})$ ,  $CTb^* \sim N(\mu b = 7 \text{ d}, \sigma b = 2 \text{ d})$  und  $CTPb = 7 \text{ d}$

$CTPb / \text{d}$	$OTDb$
7	0,5
8	0,6915
9	0,8413
10	0,9332
11	0,9772
12	0,9938

Tabelle 3.13: On-Time-Delivery Backend für  $CTb^* \sim N(\mu = 7 \text{ d}, \sigma = 2 \text{ d})$

Abbildung 3.50 zeigt die nun möglichen *OTD*-Werte des Modells ohne Release Backend in Abhängigkeit der Differenz  $(CTPf - \mu f)$ . Die ursprünglichen Produktionsleistungen sind gleich geblieben. Die Veränderungen wurden nur durch ein neues Einschleusverfahren im Backend und eine veränderte Plan-Durchlaufzeit Backend erreicht.

Der Wegfall der Einschleuswoche führt aber auch dazu, dass die On-Time-Delivery Manufacturing näher an der On-Time-Delivery Frontend liegt. Für niedrige *OTDf*-Werte verringert sich die *OTDm* während sie für hohe *OTDf*-Werte steigt (vergleiche Tabelle 3.11 und Tabelle 3.14). In allen Fällen wird jedoch eine Reduzierung der gesamten Durchlaufzeit von vier Tagen erzielt.

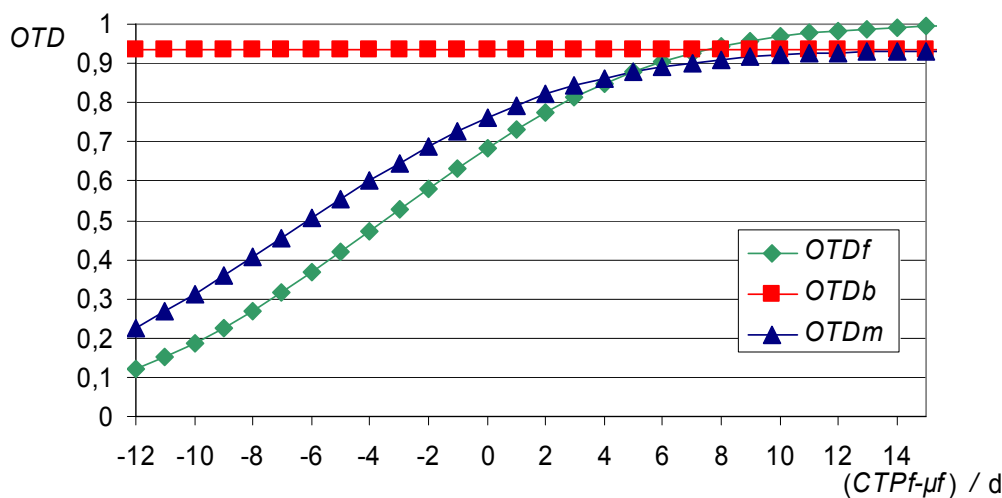


Abbildung 3.50: On-Time-Delivery Daily ohne Release Backend für  $CTf^* \sim N(\mu, \sigma = 7 \text{ d})$ ,  $CTb^* \sim N(\mu = 7 \text{ d}, \sigma = 2 \text{ d})$  und  $CTPb = 10 \text{ d}$

$\mu f$ / d	$\sigma f$ / d	$CTPf$ / d	$DDf$ / d	$OTDf$	$\mu b$ / d	$\sigma b$ / d	$CTPb$ / d	$DDb$ / d	$OTDb$	$DDm$ / d	$OTDm$
40	7	42	49	0,7747	7	2	10	10	0,9332	59	0,8204
40	7	43	50	0,8137	7	2	10	10	0,9332	60	0,8436
40	7	44	51	0,8482	7	2	10	10	0,9332	61	0,8631
40	7	45	52	0,8782	7	2	10	10	0,9332	62	0,8791
40	7	46	53	0,9038	7	2	10	10	0,9332	63	0,8922
40	7	47	54	0,9252	7	2	10	10	0,9332	64	0,9025
40	7	48	55	0,9428	7	2	10	10	0,9332	65	0,9106
40	7	49	56	0,9569	7	2	10	10	0,9332	66	0,9169
40	7	50	57	0,9681	7	2	10	10	0,9332	67	0,9216
40	7	51	58	0,9768	7	2	10	10	0,9332	68	0,9251
40	7	52	59	0,9834	7	2	10	10	0,9332	69	0,9276

Tabelle 3.14: OTD-Werte im Modell „On-Time-Delivery ohne Release Backend“

Werden dem Vergleich die Durchlaufzeiten zu Grunde gelegt, ergeben sich immer höhere Werte der Liefertreue. Bei einer Gesamtdurchlaufzeit von 64 Tagen z.B. ergibt sich eine Verbesserung der  $OTDm$  um 5,41 Prozentpunkte, was einer relativen Änderung von 6,38% entspricht (siehe Tabelle 3.15).

<b><i>DDm / d</i></b>	<b><i>OTDm mit Release Backend</i></b>	<b><i>OTDm ohne Release Backend</i></b>
63	0,8370	0,8922
64	0,8485	0,9026
65	0,8577	0,9106
66	0,8650	0,9168
67	0,8706	0,9216
68	0,8749	0,9251

Tabelle 3.15: Verbesserung der Liefertreue bei gleicher Produktionszeit durch Wegfall der Einschleuswoche im Backend für das Beispiel  $CTf^* \sim N(\mu = 40 \text{ d}, \sigma = 7 \text{ d})$ ,  $CTb^* \sim N(\mu = 7 \text{ d}, \sigma = 2 \text{ d})$

### 3.7 Rechenkomplexität und Implementierung

Für die Berechnungen der Wahrscheinlichkeit zur rechtzeitigen Lieferung in tagesfeiner Betrachtung in Kapitel 3.3.1 wurde Microsoft Excel verwendet. Mit der Funktion der Normalverteilung wurden die Wahrscheinlichkeitsfunktionen der Durchlaufzeit generiert. Die entsprechenden Werte wurden mit den Werten der Einschleusung multipliziert und zu Lieferwahrscheinlichkeiten zusammengefasst.

Die detaillierten Berechnungen wurden mit MATLAB, Version 6.1.0.450 Release 12.1 unter Verwendung der Statistik-Toolbox durchgeführt. Die Dauer einzelner Berechnungsdurchgänge ist abhängig von der Größe der Diskretisierungskonstante und den vorhandenen Systemvoraussetzungen. Für diese Arbeit stand ein AMD Athlon XP 2400+ mit 2 GHz und 992 MB RAM zur Verfügung, der mit dem Betriebssystem Microsoft Windows XP Professional betrieben wird. Für Diskretisierungskonstanten von  $\Delta t = 0,001$  Tage benötigen die einzelnen Durchgänge (Berechnung zusammen gehörender  $OTDf$ ,  $OTDb$ ,  $OTDm$ ) bis zu 40 Minuten. Eine Diskretisierungskonstante von  $\Delta t = 0,0001$  Tage würde den Zeitaufwand auf bis zu zwei Tage erhöhen.

### 3.8 Zusammenfassung

Die Steuerung der Fertigung von Chips in der Halbleiterindustrie wird aus Vorgabewerten von so genannten Performance-Indikatoren abgeleitet. Die tatsächlich erzielten Werte dienen dann zur Leistungsbewertung. Ein wichtiger Parameter ist dabei die Liefertreue. Sie wird für das Frontend, das Backend und den Kundenendtermin vorgegeben und ermittelt. Das Erreichen der Vorgaben hängt dabei von Werten und Verfahren wie Durchlaufzeit, Plandurchlaufzeit, Einschleusmodus und Zwischenlagerstrategie ab.

Basierend auf zeitdiskreten Analysen wurde ein Berechnungsverfahren entwickelt, das die oben genannten Parameter in Beziehung setzt und mit ihnen den vollständigen Produktionsablauf von Kundenbestellung bis Auslieferung abbildet. Dem stochastischen Charakter des komplexen Produktionssystems wird durch die Modellierung verschiedener

Einflussgrößen Rechnung getragen. Als Zielfunktionen ergeben sich Wahrscheinlichkeitsverteilungen der On-Time-Delivery.

Mit der Entwicklung eines zeitdiskreten Modells konnte gezeigt werden, dass sich die einzelnen Leistungs-Vorgaben gegenseitig beeinflussen und somit nicht unabhängig festgelegt werden können. Die Untersuchung der Wechselwirkungen der Leistungsgrößen der Liefertreue für die interne Supply Chain einer Halbleiterfabrik ergaben folgende Erkenntnisse:

- Die  $OTD_m$  ist immer kleiner als die  $OTD_b$ . Für eine  $OTD_f = 1$  ist die  $OTD_m$  gleich der  $OTD_b$ .
- Die  $OTD_m$  geht nicht aus der Multiplikation von  $OTD_f$  und  $OTD_b$  hervor.
- Die Abhängigkeit der  $OTD_f$  und der  $OTD_m$  von  $\mu_f$  ist gekoppelt mit der  $CTP_f$ . Eine Änderung von  $\mu_f$  wird durch die entsprechende Änderung von  $CTP_f$  bei gleich bleibenden übrigen Werten kompensiert. Es erfolgt keine Veränderung der  $OTD$ -Werte. Das Gleiche gilt für eine gleichmäßige Änderung von  $\mu_b$  und  $CTP_b$ . Da die  $CTP_f$  eine Plangröße ist, kann auf dem Papier immer eine große Erfüllung der Liefertreue hergestellt werden.
- Der Einfluss der Standardabweichung der Durchlaufzeit im Frontend auf die  $OTD_m$  ist je nach  $\mu_f$  und  $CTP_f$  unterschiedlich. Für einen gleichen Abstand zwischen  $CTP_f$  und  $\mu_f$  ergeben gleiche Standardabweichungen die gleichen  $OTD$ -Werte.
- Im Frontend sollte auf eine möglichst geringe Standardabweichung der Durchlaufzeit geachtet werden.
- Die Angabe von  $OTD_f$  und  $OTD_b$  genügt nicht, um eine eindeutige  $OTD_m$  zu ermitteln. Eine eindeutige Zuordnung einer  $OTD_m$  kann gegeben werden, wenn die Angaben je um eine zusätzliche Information über die Wahrscheinlichkeitsverteilung der Durchlaufzeit des Frontends und des Backends ergänzt werden. Dabei kann es sich wahlweise um die Standardabweichung der Durchlaufzeit oder die Differenz von Plandurchlaufzeit und Mittelwert der Durchlaufzeit handeln.
- Tagesfeine Plandurchlaufzeiten im Frontend bewirken keinerlei Änderung zu wochenfeinen Plandurchlaufzeiten, solange die Berechnung von  $OTD$  bzw.  $CLIP$  nicht unabhängig von Lieferwochen erfolgt und solange die Einschleusung ins Backend im Raster von Kalenderwochen erfolgt.
- Durch tagesfeine Liefertermine und Modifizierung der Liefer- und Einschleuswochen können bei gleicher Performance sowohl die Durchlaufzeit als auch die  $OTD_m$  verbessert werden.
- Ein Wegfall der Einschleuswoche in das Backend erzielt weitere Verbesserungen für Durchlaufzeit und  $OTD_m$ .

### 3.9 Ausblick

Das entwickelte Modell betrachtet die Fertigstellung eines einzelnen Loses. Manche Aufträge benötigen jedoch mehrere Lose. Dazu ist die Frage zu untersuchen, wie weitere Lose in das Modell integriert werden können. Damit kann ebenso untersucht werden, ob durch den gleichzeitigen Start weiterer Lose die Wahrscheinlichkeit einer  $CLIP$ -Verletzung verringert, bzw. die Wahrscheinlichkeit einer rechtzeitigen und genügenden Lieferung erhöht werden kann.

Wird mehr als ein Los betrachtet, müssen die Einschleusregeln beachtet werden, die die Reihenfolge der Lose in der Einschleusung vorgeben. Außerdem ist zu klären, ob die Durchlaufzeiten der Lose eines gleichen Produktes voneinander statistisch unabhängig sind oder ob sie voneinander abhängig produziert werden. D.h. das Ergebnis hängt davon ab, ob die beiden Lose des gleichen Produktes immer beieinander bleiben und direkt nacheinander die einzelnen Prozessschritte durchlaufen, oder ob sie im Laufe der Produktion voneinander getrennt und individuell bearbeitet werden.

Bei einer abhängigen Bearbeitung erreichen beide Lose mit hoher Wahrscheinlichkeit gemeinsam das Ende der Produktion. Somit haben sie dieselbe Durchlaufzeit. Eine Verbesserung bezüglich der Durchlaufzeit ist dann nicht möglich. Wenn das erste Los zu spät fertig gestellt wird, so ist das zweite Los ebenso zu spät. Dennoch kann in diesem Fall eine zu geringe Ausbeute aufgefangen werden. Ist die Ausbeute geringer als geplant, so kann eine dadurch hervorgerufene Verletzung des *CLIP* mit Hilfe eines zusätzlichen Loses ausgeglichen werden.

Ist die Bearbeitung und somit die Durchlaufzeit der Lose statistisch unabhängig, so kann über den Ausgleich einer zu geringen Ausbeute hinaus auch die Wahrscheinlichkeit einer rechtzeitigen Fertigstellung des Auftrages verbessert werden. Wird ein zusätzliches Los eingeschleust, ist die Wahrscheinlichkeit, dass ein Los den Liefertermin erreicht, höher als wenn nur ein Los produziert wird.

Der gesamte *CLIP* einer Produktionsstufe ergibt sich aus den Performance-Werten mehrerer Produkte. In einer Erweiterung des Modells könnte dies berücksichtigt werden.

Wie stark das Modell der Realität angeglichen werden sollte, hängt davon ab, welche Fragestellungen untersucht werden. Einfache Modelle können grundsätzliche Zusammenhänge aufzeigen. Sollen Vorhersagen über tatsächliche Szenarien getroffen werden, müssen auch Details im Modell berücksichtigt werden. So gibt es z.B. die Möglichkeit, dass auch Teilaufträge vom Backend weiterverarbeitet werden oder, wenn eine Bestellung zu spät die Die Bank erreicht, dann können statt dessen Chips aus der Die Bank weiterverarbeitet werden, die sie zu früh erreicht haben. Ein sinnvoller Schritt zu größerer Realitätsnähe ist auch eine genauere Modellierung der Durchlaufzeiten. Dazu können Messungen und Auswertungen tatsächlicher Werte hilfreich sein.



# Kapitel 4

## Stepper-Dispatcher

### 4.1 Einleitung

Mit der Fragestellung, in welcher Reihenfolge Produkte prozessiert werden, beschäftigen sich sowohl die Planung als auch die Steuerung einer Halbleiterfab. Während die Einschleuswoche eines Produktes durch den Vertrieb vorgegeben wird, wird der genaue Zeitpunkt der Einschleusung von der Liniensteuerung bestimmt. Da aber nicht jedes Produkt die gleichen Maschinen in gleicher Reihenfolge benötigt, wird die Reihenfolge der Produkte nach der Einschleusung schnell verändert. Die vor einer Maschine bzw. Anlagengruppe wartenden Lose befinden sich meist in unterschiedlichen Produktionsfortschritten. Dadurch kann die Wahl der Reihenfolge eine starke Auswirkung auf das Verhalten der gesamten Produktion haben. Bei mehreren gleichartigen Maschinen kommt noch die Aufgabe dazu, die Produkte zu verteilen. Mögliche Vorschriften zur Reihenfolgenbildung sind in Kapitel 2 beschrieben.

In einer realen Produktion gibt es jedoch auch ganz einfache und grundlegende Argumente für bzw. gegen die Produktion eines Loses zu einem bestimmten Zeitpunkt. Sind noch nicht alle notwendigen Betriebsmittel bereitgestellt, so führt die Wahl eines Loses nur zur Blockade der entsprechenden Maschine. Ebenso kann ein unerwarteter Maschinenausfall eine veränderte Vorgehensweise erfordern. Die tatsächliche Einteilung der Bearbeitungsreihenfolge von Losen muss deshalb vor Ort entschieden werden und immer wieder den veränderten Umständen angepasst werden.

Um die Performance einer Linie zu steigern, ist es darüber hinaus notwendig, die Bestimmung von Produktionsreihenfolgen unter einem größeren Blickwinkel zu betrachten und weitere Arbeitsschritte mit einzubeziehen. Die Reihenfolge, in der Lose eine Arbeitsstation verlassen, bestimmt die Wahlmöglichkeiten, mit der mögliche Bearbeitungsreihenfolgen der Folgestationen gebildet werden können. Um z.B. die Auslastung der Maschinen einer Anlagengruppe zu verbessern, ist es entscheidend, welche Auswahl an Produkten zur Verfügung steht. Diese Auswahl kann durch die Produktionsreihenfolge der Vorstation beeinflusst werden.

Im Folgenden werden zwei Systeme vorgestellt, die eine Reihenfolgenbildung unterstützen bzw. bestimmen. Im Dispatch-Own-System werden die aktuellen Umstände innerhalb einer Produktionsstufe zusammengetragen und für eine sinnvolle Auswahl der Lose ausgewertet. Das Dispatch-Ahead-System bestimmt eine Produktionsreihenfolge, um die Maschinenauslastungen der Folgestation zu verbessern. Da die Lithografie das Bottleneck einer Halbleiterlinie ist, werden die Systeme zunächst dort eingeführt bzw. am Beispiel der Lithografie untersucht.

Veränderungen der Produktionsreihenfolge, die das Erreichen von Mengenvorgaben erleichtern und somit die Performance der Linie unter Umständen verringern, können von einem Dispatch-System natürlich nicht verhindert werden. So kann es vorkommen, dass zur Erfüllung der Planvorgabe „Soll-Moves pro Tag“ Lose vorgezogen werden, die „leicht“ bzw. schnell zu bearbeiten sind. Für die Folgestationen kann eine daraus resultierende Reihenfolge eventuell zu Performance-Verlusten führen. Dies soll hier als Beispiel dafür erwähnt werden, dass unterschiedliche Sichtweisen immer zu unterschiedlichen, konkurrierenden „optimalen“ Handlungsvorgaben führen.

## 4.2 Dispatch Own

### 4.2.1 Vorbereitung und Durchführung von Belichtungsoperationen

Das Vier-Partner-Modell (siehe Kapitel 2.5.4) besagt, dass Produktion nur dann möglich ist, wenn alle vier Produktionspartner am selben Ort zur selben Zeit verfügbar sind. Durch Synchronisation der Produktionspartner können große Produktivitätssteigerungen erreicht werden. Beispielhaft für ein Modul der Lithografie soll durch Verbesserung von Verfügbarkeiten und Synchronisationsgrad die Produktivität der Operation Belichtung gesteigert werden. Da es sich bei der im Folgenden beschriebenen Maßnahme um eine Unterstützung des Operators beim Verteilen der Lose auf die Belichtungsanlagen handelt, wird das System Dispatch Own genannt [Sal01a].

Um ein Los belichten zu können, ist eine Belichtungsanlage, in diesem Fall sind es so genannte Stepper, eine Belichtungsmaske (Reticle) und ein passendes Maschinenprogramm notwendig. Jedes Los wird während seiner Produktion von einem Protokoll, dem so genannten Laufzettel, begleitet. In diesem Laufzettel sind die einzelnen Operationen und die jeweils benötigten Zusatzinformationen beschrieben.

Abbildung 4.1 zeigt die Vorbereitungsschritte, die von einem Operator für eine Belichtung durchzuführen sind. Zunächst muss der Operator ein Los aus dem Lager vor der Belichtungsoperation auswählen. Eine Software hilft ihm dabei, ein Los mit einer hohen Priorität auszuwählen. Die Priorisierung erfolgt dabei über Dispatch-Regeln (siehe Kapitel 2). Der Laufzettel enthält Angaben über Operation, Reticle und Maschinenparameter wie Dosierung und Justierung. In Dedizierungstabellen, die sich täglich ändern können, muss der Operator überprüfen, an welchem Stepper die entsprechende Operation durchgeführt werden darf. Danach muss ein passender Stepper gewählt werden, der frei ist oder bald frei wird.

Der Lagerort des entsprechenden Produkts und die Schichtnummer des benötigten Reticles kann im Reticle Management System ermittelt werden. In vielen Fällen befindet sich die Maske im angegebenen Reticle-Lager, das aus Platzgründen in einer anderen Reinraumhalle untergebracht ist. Der Operator muss das Reticle von dort holen und zum

ausgewählten Stepper bringen. Die Maske kann aber ebenso bereits an einem Stepper vorhanden sein. Ist dies dem Operator nicht bekannt, geht er dennoch in die Lagerhalle bevor er die Maske in der Lithografie suchen muss. Wird die Maske bereits von einem Stepper einer anderen Steppergeneration verwendet, so muss die Maske auch noch in einen anderen Aufnahmebehälter (Box) umgebettet werden.

Wurde die Operation am ausgewählten Stepper schon längere Zeit nicht mehr oder noch nie durchgeführt, muss das Maschinenprogramm, das die Maschine steuert, geladen werden. Bei neuen Produkten ist es möglich, dass das Maschinenprogramm noch nicht geschrieben oder zumindest noch nicht in der Produktionshalle verfügbar ist.

Bevor der Belichtungsprozess gestartet werden kann, muss der Operator verschiedene Prozessparameter, wie Dosis und Fokus, von Hand oder über Barcodeleser an der Maschine einstellen.

Durch die Vielzahl der Systeme, aus denen die jeweiligen Informationen bezogen werden, ergibt sich ein entsprechender Zeitaufwand. Darüber hinaus ist nicht vermeidbar, dass die eine oder andere Eingabe fehlerhaft durchgeführt wird, was eine Nacharbeit oder im schlimmsten Fall einen Losverwurf zur Folge haben kann.

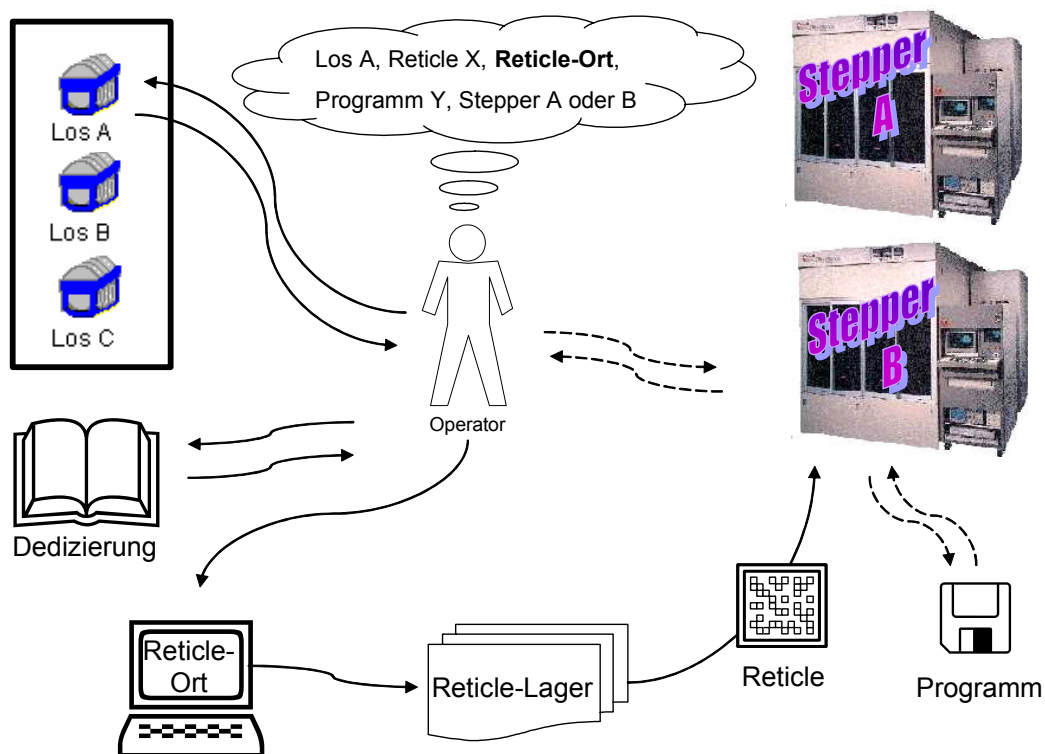


Abbildung 4.1: Vorbereitung einer Belichtungsoperation

Um die Zeiten für die Vorbereitung einer Belichtungsoperation zu ermitteln, wurden mit einem erfahrenen Operator Zeitmessungen während der Bearbeitung mehrerer Standardlose durchgeführt. Zu den Standardfällen gehören Lose, deren Reticle aus dem Reticle-Lager geholt werden müssen und deren Maschinenprogramm (der Job) bereits im richtigen Stepper vorhanden ist. Der erfahrene Operator wusste bereits, welche Stepper für die einzelnen Lose freigegeben waren und konnte auf die Informationen der

Dedizierungstabellen verzichten. Die Aufenthaltsorte der Masken waren dem Operator ebenfalls bekannt, so dass eine Reticle-Suche entfiel. Da nicht immer nur ein Los gleichzeitig bearbeitet wird, wurde die Bearbeitung mehrerer Lose aufgenommen und zur Anzahl der Lose in Beziehung gesetzt. Die gemittelten Werte für die reinen Wegezeiten zur Vorbereitung eines Loses ergaben sich zu 1,74 Minuten. Die einzelnen Arbeitsschritte können in Bereitstellung des Loses, Bereitstellung des Reticles und Beschaffung des Reticles unterteilt werden. Die Werte sind in Tabelle 4.1 angegeben.

Vorbereitungszeit	Bereitstellung des Loses	Bereitstellung des Reticles	Weg für Reticle aus Lagerhalle	Summe
Ohne Dispatch-Own (erfahrener Operator)	0,06 min	0,26 min	1,42 min	1,74 min
Durchschnittlich erfahrener Operator				3,48 min
Worst-Case-Fall				15 min

Tabelle 4.1: Durchschnittliche Wegezeit pro Los zur Vorbereitung einer Belichtungsoperation

Ein weniger erfahrener Operator benötigt darüber hinaus zusätzlich Zeit, um in den Dedizierungstabellen die aktuell freigegebenen Stepper nachzuschlagen und das Reticle zu suchen, wenn es sich nicht im Reticlelager befindet. Damit werden die Wegezeiten zur Vorbereitung einer Belichtung auf durchschnittlich 3,48 Minuten pro Los erhöht.

Für den Fall, dass das Maschinenprogramm noch nicht am benötigten Stepper aufgespielt ist, erhöht sich die Vorbereitungszeit nochmals. Der Worst-Case-Fall tritt ein, wenn der Job erst angelegt und in die Produktion gebracht werden muss. Dazu werden durchschnittlich zusätzlich 15 Minuten benötigt.

#### 4.2.2 Synchronisation

Für den Vorgang der Belichtungsvorbereitung ist auffällig, dass die Verfügbarkeiten der Produktionspartner vom Arbeitsfortschritt des Operators abhängen und somit immer erst nacheinander eintreten. Aus dem Vier-Partner-Modell ist bekannt, dass nicht nur die Höhe der Verfügbarkeiten, sondern auch die Synchronisation der Einzelverfügbarkeiten ausschlaggebend für die Produktivität des Produktionssystems ist. Dabei ist entscheidend, dass die Zustandsänderungen von „nicht verfügbar“ auf „verfügbar“ und umgekehrt möglichst gleichzeitig erfolgen. Produktivitätsverluste entstehen genau dann, wenn Partner, die bereits verfügbar sind, auf die Verfügbarkeit eines oder mehrerer anderer Partner warten müssen.

An dieser Stelle wird auch die Unterscheidung zwischen Verfügbarkeit im Allgemeinen und Verfügbarkeit für einen bestimmten Bearbeitungsschritt notwendig. Entscheidend dafür, ob ein Prozess gestartet werden kann, ist die gleichzeitige Verfügbarkeit aller Produktionspartner am selben Ort für den betrachteten Prozess. Auch wenn die absolute Zeit der Verfügbarkeit eines Partners nicht verbessert wird, so führt das Verringern von Wartezeiten, während der der Partner zwar verfügbar aber nicht produktiv ist, zu einer

größeren Verfügbarkeit des Partners für andere Prozesse. Wird also von Erhöhung der Verfügbarkeit gesprochen, so kann dies auch durch Reduzierung der Beteiligung eines Partners an einem Prozess erreicht werden, da somit der entsprechende Produktionspartner für andere Prozesse eine höhere und schneller zu erreichende Verfügbarkeit gewinnt.

Für die Vorbereitung von Belichtungsoperationen kommen somit folgende Optimierungsansätze in Betracht: Verringerung von Wartezuständen, in denen ein Produktionspartner verfügbar aber nicht produktiv ist und möglichst gleichzeitiges Erreichen der Verfügbarkeit aller Produktionspartner. Können die einzelnen Arbeitsschritte des Operators bis zum Start der Belichtung zeitlich verkürzt werden, so werden Wartezustände verringert und Verfügbarkeiten entstehen zeitnäher. Eine Möglichkeit dazu bietet das Zusammenfassen aller für die Belichtung benötigten Informationen und deren Darstellung an einem Medium. Damit reduzieren sich die Wegezeiten des Operators, die Belichtung kann schneller gestartet werden und Wartezeiten reduzieren sich.

Als Medium wird das bisher verfügbare Dispatch-System verwendet, das bereits in allen Modulen verwendet wird und eine Priorisierung der Lose nach den bekannten Dispatch-Regeln vornimmt. Für jedes vor der Belichtungsoperation wartende Los sollen folgende, für die anstehende Belichtung notwendigen Informationen angezeigt werden: alle freigegebenen Stepper, das benötigte Reticle mit tatsächlichem Aufenthaltsort und das benötigte Programm mit Information über dessen Existenz und dem jeweiligen Ladezustand in den Steppern.

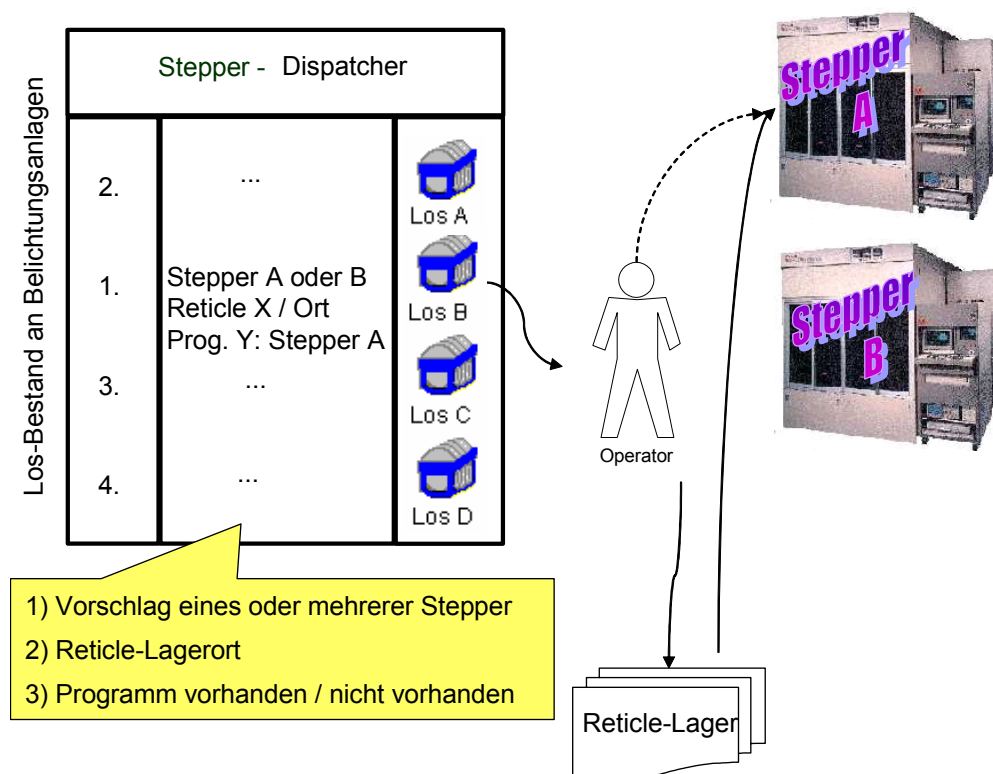


Abbildung 4.2: Vorbereitung einer Belichtungsoperation mit Dispatch-Own-Stepper-Dispatcher

Abbildung 4.2 zeigt die deutlich verkürzten Wege, die zur Vorbereitung einer Belichtungsoperation mit Hilfe der Dispatch-Own-Funktionalität des Stepper-Dispatchers notwendig sind. Die Wege zu den Dedizierungstabellen und zum Terminal des Reticle

Management Systems entfallen vollständig. Ist auf einem Stepper das benötigte Programm bereits vorhanden, so kann diesem Stepper der Vorzug gegeben werden. Ein Los, dessen Programm noch nicht vorhanden ist, wird noch nicht vorbereitet.

Die softwaretechnische Implementierung kann in [Sa101a], [Sa101b] und [Sa101c] nachgelesen werden. Hier soll nur ein kurzer Überblick über die Arbeitsschritte zur Realisierung des Stepper-Dispatchers gegeben werden.

Zunächst mussten Recherchen angestellt werden, ob und wo die benötigten Daten elektronisch vorhanden sind und wie sie in eine neu anzulegende Datenbank (Dispatch-DB) übertragen werden können. Abbildung 4.3 gibt einen vereinfachten Überblick über die Zusammensetzung der benötigten Datenbanken; Details der Datawarehouse-Struktur sind in [Sar99a] dargestellt. Für den Datenübertrag mussten unterschiedliche Einspielprozesse programmiert werden. So mussten z.B. die Parameter der Protokolldaten (Ebene, Reticle, Dosistabelle, Justierung) anhand von Schlüssel-Zeichenfolgen erkannt und extrahiert werden, da sie nur als Zeichenketten innerhalb des Laufzetteltexes zur Verfügung standen.

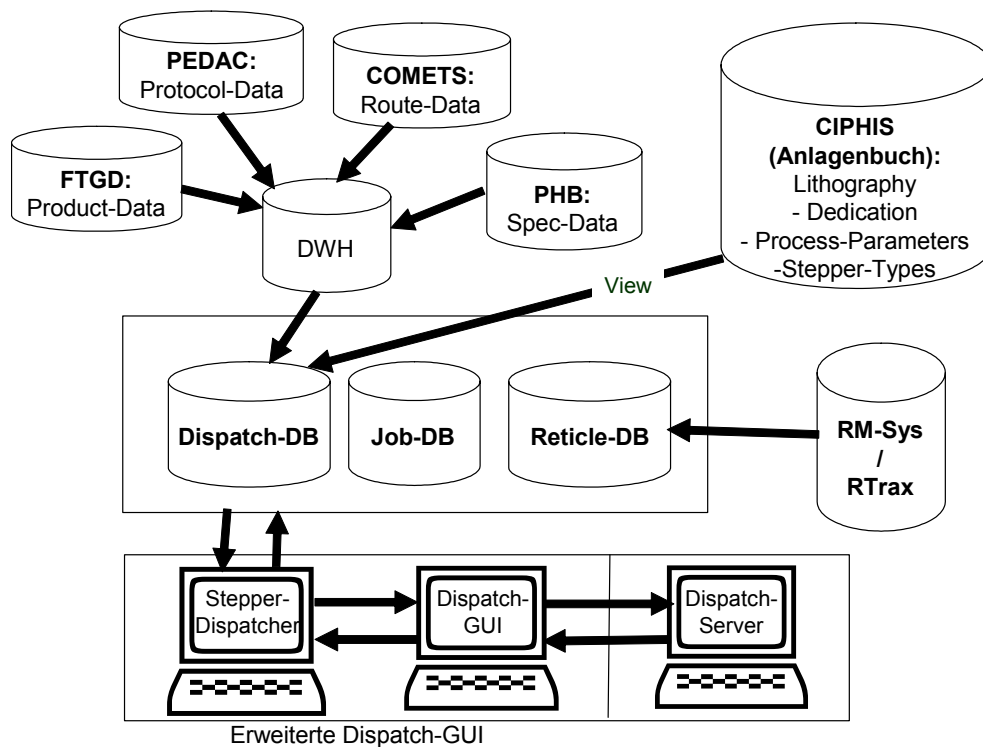


Abbildung 4.3: Datenherkunft

Das bisher verfügbare Dispatch-System, das eine einfache Priorisierung der Lose vornimmt, wird in allen Modulen der Produktion gleichermaßen verwendet. Die neu zu entwickelnde Software musste deshalb konfigurierbar und pluginfähig gestaltet werden. Die Konfigurationsmöglichkeiten wurden durch schematische Views auf andere Datenbanken und durch generische Funktionsaufrufe realisiert. Die bereits bestehende grafische Oberfläche Dispatch-GUI (Graphical User Interface, Benutzeroberfläche), die als Server-Client-Anwendung zur Verfügung stand, diente dazu, die Software zu einer pluginfähigen Applikation zu erweitern. Dazu wurde eine allgemeine Bibliothek eingerichtet, die alle Eigenschaften und Funktionen modulübergreifend zur Verfügung stellt.

Die erweiterte Dispatch-GUI mit Dispatch-Own-Funktionalität kann in zwei Versionen verwendet werden: Version 1 für Los-Verteiler-Regale und Version 2 für Anlagen-PCs. Am Verteiler-Regal werden alle Lose der Anlagen angezeigt, die mit Losen aus diesem Regal beliefert werden können. D.h. am Verteilerregal der Lithografie werden alle Lose angezeigt, die auf eine Belichtung warten. Mit Hilfe der weiteren Angaben über freigegebene Stepper, Reticlelagerort und Job-Status kann ein Los ausgewählt werden, das alle Voraussetzungen zu einer schnellen Weiterverarbeitung erfüllt.

Zur besseren Auslastung der Maschinen ist es allerdings sinnvoll, Lose, bezogen auf den Arbeitszustand der Anlagen, auszuwählen. Wird eine Maschine frei, so sollte zuerst ein Los weiterverarbeitet werden, das mit dieser Maschine bearbeitet werden kann. In der Dispatch-Own-GUI eines Anlagen-PCs werden alle Lose ausgeblendet, die nicht an diesem Stepper belichtet werden dürfen.

Abbildung 4.4 zeigt einen Screenshot der Dispatch-Own-GUI am Stepper Nummer 0560. Im Warteregal befinden sich drei Lose, die auf diesem Stepper belichtet werden können. Neben der Losnummer sind alle Laufzetteldaten (Protokollparameter) und alle für die jeweilige Belichtungsoperation freigegebenen Stepper angegeben. Die Gruppenliste gibt eine kompaktere Anzeige der Dedizierung mit Informationen zum Vorhandensein des Programms und Status der Anlagengruppe. Für das benötigte Reticle werden Lagerort, Status und eventuell Aufenthalt in einem Stepper angezeigt. Der Detaildialog in der Mitte der Anzeige enthält für das ausgewählte Los weitere detailliertere Daten und Informationen.

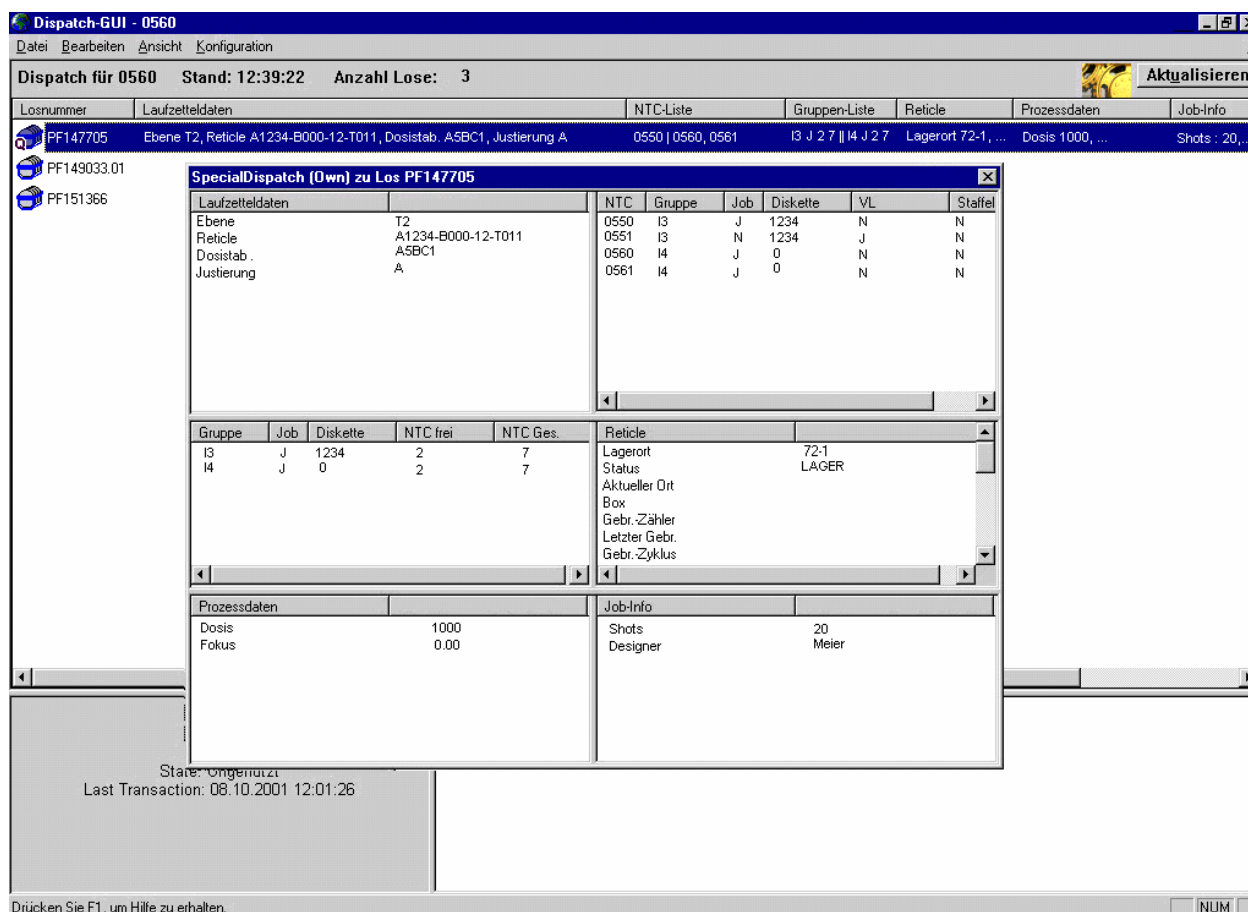


Abbildung 4.4: Erweiterte Dispatch-GUI (Dispatch Own) mit Detaildialog

### 4.2.3 Performancesteigerung durch Dispatch Own

Nachdem die Dispatch-GUI mit Dispatch-Own-Funktionalität in der Fertigung eingeführt war, wurden Zeitaufnahmen durchgeführt und mit den ursprünglichen Zeiten verglichen.

Abbildung 4.5 stellt die Ergebnisse der Zeitmessungen vor und nach der Einführung der Dispatch-Own-Funktionalität nebeneinander. Die gemessenen Zeiten sind wiederum in die Bereiche Bereitstellung des Loses, Bereitstellung des Reticles, Weg für Herbeischaffen des Reticles aus der Lagerhalle und „Job nicht vorhanden“ unterteilt. Neu hinzu kommt ein Zeitabschnitt für das Auswerten der Daten am Dispatch-PC.

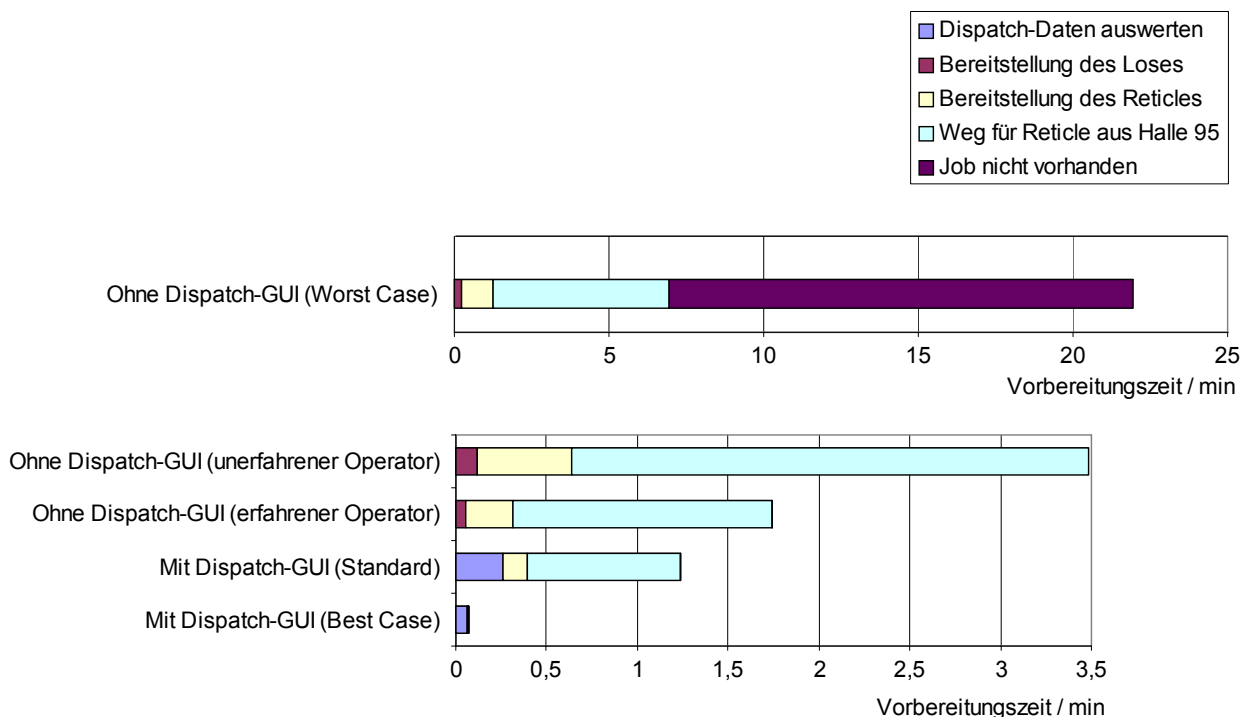


Abbildung 4.5: Vorbereitungszeiten vor und nach der Einführung der Dispatch-GUI mit Dispatch-Own-Funktionalität

Die ersten drei Darstellungen zeigen die Vorbereitungszeiten ohne Dispatch-GUI. Im ungünstigsten Fall (Worst Case) wird ein Los vorbereitet, das nach der Vorgabeliste ganz oben steht, dessen Job aber nicht auf der Festplatte des ausgewählten Steppers vorhanden ist. Ein bereits vorbereitetes Reticle muss dann durchschnittlich 15 Minuten warten, bis die Belichtung beginnen kann. In günstigen Fällen, wenn der Aufenthaltsort des Reticles bekannt ist und der Job bereits geladen ist, benötigt ein unerfahrener Operator durchschnittlich 3,48 Minuten, während ein erfahrener Operator die Vorbereitung einer Belichtungsoperation in der Hälfte der Zeit erledigen kann.

Mit der neuen Dispatch-GUI ergibt sich eine neue Vorgehensweise für die Vorbereitung der Lose zur Belichtung. Die ersten Messungen (mit Standard bezeichnet) ergaben folgende Zeitaufteilung: Zunächst ist eine gewisse Zeit notwendig, um aus der Dispatch-GUI ein Los auszuwählen und die weiteren Informationen zu erfassen. Vor der Einführung der Dispatch-Own-GUI wurde noch ein Bereitstellungsregal verwendet, in dem die Lose solange



untergebracht waren, bis der richtige Stepper ermittelt und das passende Reticle vorbereitet war. Mit Hilfe der Dispatch-GUI und des eingebundenen Lot-Location-Managements wird dieses Regal nicht mehr benötigt, da das Los sofort zum passenden Stepper transportiert wird. Die Zeit für das Bereitstellen des Loses wird damit so gering, dass es in der graphischen Darstellung nicht mehr als eigenständiger Bereich erkennbar ist. Mit den neuen Informationen über den Aufenthaltsort des Reticles verringert sich der Suchaufwand und die Zeiten für das Bereitstellen des Reticles werden geringer. Ebenso verringern sich die Zeiten für die Wege in die Lagerhalle. Da die Verfügbarkeit des Jobs bekannt ist, werden die Lose nur noch zu Steppern gebracht, an denen der Job bereits vorhanden ist. Ist ein Job noch nicht geschrieben, so kann das Erstellen des Jobs und das Laden des Programms auf den benötigten Stepper veranlasst werden, bevor die Vorbereitung zur Belichtung für das Los begonnen wird. Mit Hilfe der Dispatch-Own-GUI können die Fälle, in denen der Job während der Vorbereitung des Loses besorgt werden muss, vollständig verhindert werden. Eine Standard-Vorbereitung mit Dispatch-GUI benötigt durchschnittlich nur noch 1,24 Minuten.

Im optimalen Fall (Best Case) werden nur noch Zeiten für das Auswerten der Dispatch-Daten und zum Bereitstellen des Loses benötigt. Ein erfahrener Operator findet die von ihm benötigten Informationen in der Dispatch-GUI sehr schnell. Das Reticle kann bereits im voraus beschafft werden, so dass die dazu benötigte Zeit nicht mehr in die Zeiterfassung für die Vorbereitung der Belichtung fällt. Für einen derart optimalen Fall kann sich die Vorbereitungszeit auf 5 Sekunden reduzieren.

Der Erfolg der Dispatch-Own-Funktion kann mit Hilfe der Zeitmessungen konkret in Zahlen ausgedrückt werden. Bei einem Vergleich der Zeiten ohne Dispatch-GUI mit erfahrenem Operator mit den Zeiten mit Dispatch-GUI im Standard-Fall ergibt sich eine Verringerung der reinen Wegezeiten um 28,74% (Abbildung 4.6). Verglichen mit den Zeiten eines unerfahrenen Operators oder des Worst Cases ergeben sich Verbesserungen von 64,35% und 94,35%. Im Folgenden werden die Zeiten eines erfahrenen Operators ohne Dispatch-GUI mit der Standard-Situation mit Dispatch-GUI verglichen. Die Verbesserung der Performance stellt dabei immer die Mindestverbesserung dar, da die ungünstigen Fälle vor der Verbesserungsmaßnahme nicht berücksichtigt werden.

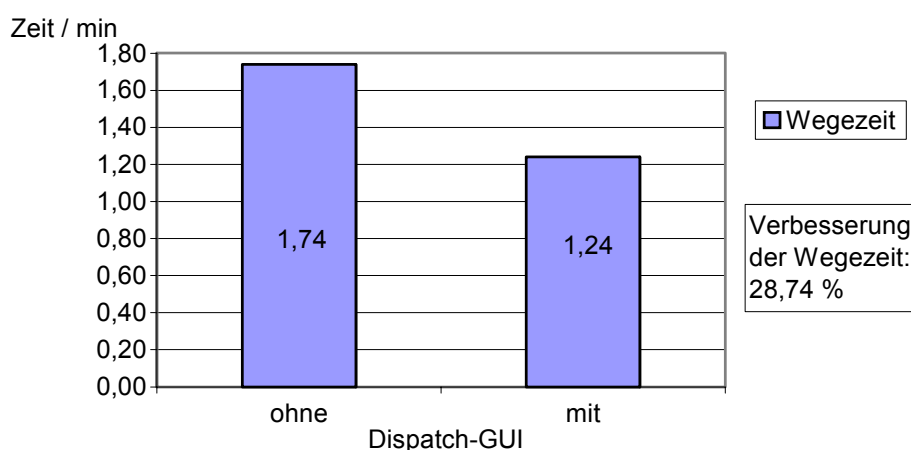


Abbildung 4.6: Veränderung der reinen Wegezeit

Für die Belichtungsoperation ist die Vorbereitungszeit, je nach ihrer Länge, ein geringer Teil der gesamten Losbearbeitungszeit. Vor Einführung der Dispatch-GUI betrug die mittlere Losbearbeitungszeit inklusive Belichtung bei Bearbeitung durch einen erfahrenen Operator 43,54 Minuten. Durch die Reduzierung der Wegezeiten wird die Bearbeitungszeit im Standard-Fall auf 42,48 Minuten verringert. Somit wird eine Verbesserung der Durchlaufzeit der Lose um 2,44% erreicht (Abbildung 4.7).

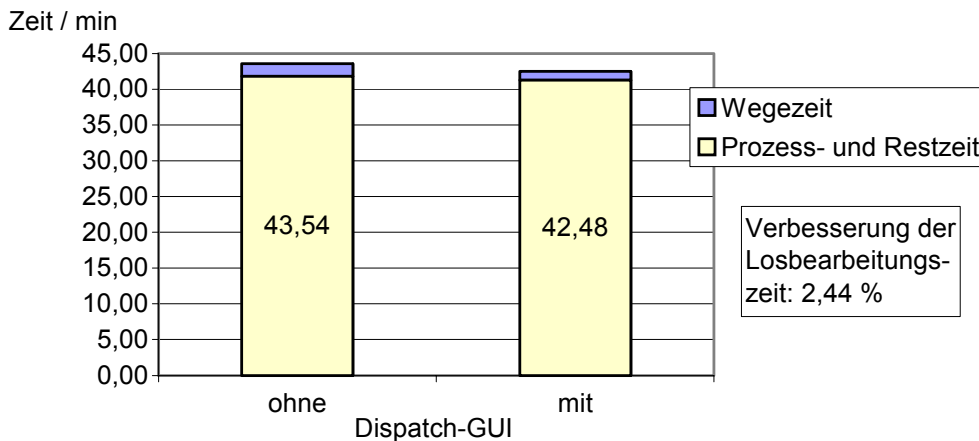


Abbildung 4.7: Veränderung der Losbearbeitungszeit

Unter Beachtung der Tatsache, dass für den Unterhalt einer Halbleiter-Fabrik pro Tag mindestens eine Million Euro aufgewendet werden müssen, wird deutlich, welche Ausmaße eine Produktivitätssteigerung von 2,5% annimmt. Da die Fotolithografie das Bottleneck einer Halbleiterfab ist, bedeutet diese Steigerung auch 2,5% Mehrauslastung der Fabrik, was mindestens einer Einsparung von zusätzlichen 25.000 Euro pro Tag für eine ansonsten um zweieinhalb Prozent zu vergrößernde Fertigungshalle entspricht. Auf der anderen Seite bedeutet diese Zeitersparnis auch eine Umsatzsteigerung, da 2,5% mehr Scheiben pro Tag bearbeitet werden können. Bei einem Durchsatz von 15.000 Scheiben pro Tag, was in etwa dem Produktionsvolumen der Fototechnik entspricht, können somit ca. 375 Scheiben zusätzlich produziert werden.

Zusammen mit einem Flussfaktor  $FF = 1,97$  und einer Variabilität  $\alpha = 1,23$  ergeben ein Durchsatz von 15.000 Scheiben pro Tag und eine Durchlaufzeit von 43,2 Minuten die in Abbildung 4.8 dargestellte Betriebskennlinie BKL 1. Können Flussfaktor und Variabilität gleich gehalten werden, so bewirkt die Durchsatzsteigerung um 375 Scheiben pro Tag eine Erhöhung der Kapazität, was zur Betriebskennlinie BKL 2 führt. Abbildung 4.9 zeigt den für die Arbeitspunkte relevanten Bereich der Kennlinien. Der Übergang auf eine tiefer liegende Kennlinie kann entweder zu einer maximal möglichen Verringerung der Durchlaufzeit bei gleich bleibendem Durchsatz oder zu einer maximal möglichen Erhöhung des Durchsatzes bei gleich bleibender Durchlaufzeit oder zur Verbesserung beider Werte innerhalb der Maxima dieser Grenzwerte genutzt werden. Der hier in Frage kommende Bereich ist in Abbildung 4.9 gekennzeichnet. Die kleinste mögliche Durchlaufzeit beträgt 41,47 Minuten bei gleich bleibendem Durchsatz. Der größte mögliche Durchsatz beträgt 15.921 Scheiben pro

Tag bei gleich bleibender Durchlaufzeit. Tatsächlich realisiert wurde in Arbeitspunkt 2 eine Erhöhung des Durchsatzes auf 15.375 Scheiben pro Tag bei einer Verringerung der Durchlaufzeit auf 42,12 Minuten.

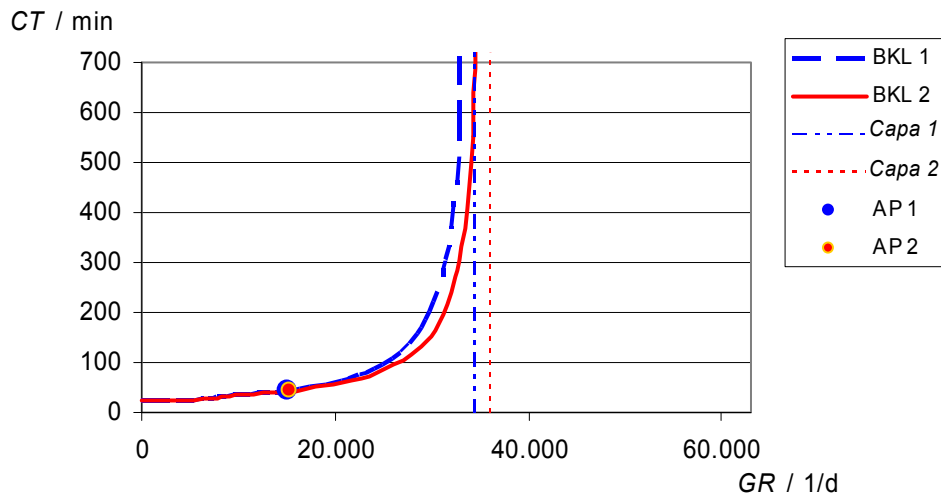


Abbildung 4.8: Betriebskennlinie vor und nach Einführung der Dispatch-GUI

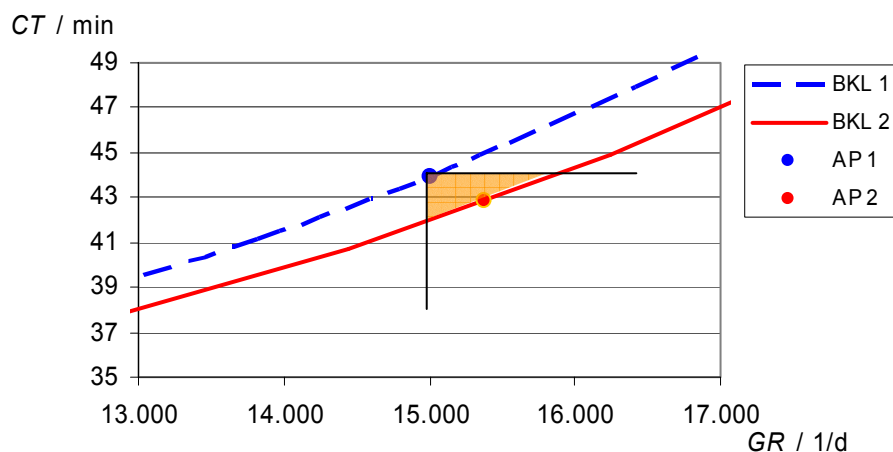


Abbildung 4.9: Arbeitspunkte vor und nach Einführung der Dispatch-GUI

Noch nicht in Zahlen messbar, aber dennoch nicht zu vernachlässigen ist ein weiterer Effekt, der enormes Potential zur Optimierung der Produktionsperformance bietet. Vor Einführung der Dispatch-GUI wurden die Lose nach den Dispatch-Regeln abgearbeitet, die im Wesentlichen aus First-In-First-Out-Vorgaben bestehen. Die Produktion wird somit aus Sicht der Lose gesteuert. Es bleibt dem Einsatz und der Erfahrung des Operators überlassen, Lose vorzuziehen, wenn die dazu benötigte Maschine leer steht. Um eine

möglichst hohe Auslastung der Maschinen zu erzielen, muss die Produktion jedoch aus Sicht der Maschinen gesteuert werden. Bevor eine Maschine keinen Arbeitsvorrat mehr hat, muss ein passendes Los ausgewählt und vorbereitet werden. Die Auslastung der Maschinen steuert die Reihenfolge der Bearbeitung der Lose und zieht somit das Material durch die Bearbeitungsschritte. Abbildung 4.10 stellt die „ziehende“ Produktionssteuerung der „schiebenden“ gegenüber.

Die Dispatch-GUI mit der Dispatch-Own-Funktionalität eröffnet die Möglichkeit, eine ziehende Produktionssteuerung zu verfolgen. Ist der Arbeitsvorrat eines Steppers aufgebraucht, so kann am Anlagen PC gezielt nach einem Los hoher Priorität gesucht werden, das auf dieser Anlage bearbeitet werden kann. Auf die Optimierung der Maschinenauslastung und die Produktionssteuerung aus Sicht der Maschinen wird im folgenden Kapitel 4.3 Dispatch Ahead näher eingegangen.

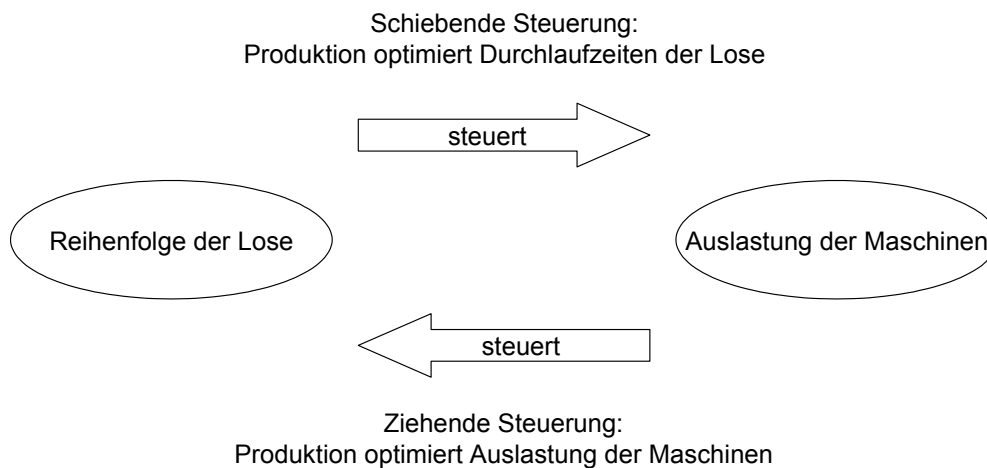


Abbildung 4.10: Schiebende und ziehende Produktionssteuerung

Beim Vier-Partner-Modell erscheint es sinnvoll, dieses für die Bewertung der Dispatch-Own-Funktionalität zu modifizieren. Die Verfügbarkeit des Operators wurde, bezogen auf seine absolut verfügbare Zeit, nicht verbessert. Die Verringerung des Zeitbedarfs für die Vorbereitung eines Loses führt jedoch dazu, dass der Operator schneller für die Vorbereitung eines weiteren Loses zur Verfügung steht. Es erscheint somit zweckmäßiger, eine Verfügbarkeit bezogen auf Arbeitspakete anzugeben, die wie folgt aussehen sollte: Die Anzahl der Arbeitspakete, die ein Operator an einem Tag bearbeiten kann, ist das Verhältnis aus der Gesamtzeit, die dem Operator zur Verfügung steht, und der benötigten Zeit pro Arbeitspaket  $t_{AP}$ .

$$A_{AP} = \frac{A \cdot 24 \text{ h}}{t_{AP}} \tag{4.1}$$

Hat ein Operator eine Verfügbarkeit von 90%, so erreichte er vor der Einführung der Dispatch-GUI für die Vorbereitung von Belichtungslosen eine Arbeitspakete-Verfügbarkeit von  $A_{AP} = 745$  und nach deren Einführung von  $A_{AP} = 1.045$ . Das entspricht einer Erhöhung um 40%.

Für die vier Partner können folgende Auswirkungen zusammengefasst werden:

- **Mensch:**  
Die Reduktion von Such- und Wegezeiten führt zu einer Erhöhung der Arbeitspakete-Verfügbarkeit. Fehlermöglichkeiten bezüglich Fehlbelichtungen werden reduziert. Auch weniger erfahrene Operatoren können Lose schneller vorbereiten.
- **Maschine:**  
Die Maschinen werden mit Hilfe der Dispatch-GUI zuverlässiger mit Material versorgt. Leerlaufzeiten können durch konsequentes Verfolgen der ziehenden Produktionssteuerung auf ein Minimum reduziert werden. Zum anderen können Maschinen, die zu stark belegt sind, entlastet werden, indem alternativ mögliche Anlagen gleichzeitig angezeigt werden. Der Aufbau von Losbeständen kann gleichmäßig über die Anlagen verteilt werden.
- **Methode:**  
Durch die Anzeige der freigegebenen Stepper, der vorhandenen Jobs und des Aufenthaltsortes des Reticles wird die Verwendung einer fehlerhaften Methode ausgeschlossen. Die produkt- und prozessgruppenfeine Dedizierung auf elektronischer Basis ist immer auf dem aktuellsten Stand. Versionsverwechslungen sind nicht mehr möglich.
- **Material:**  
Die Verteilung der Lose wird beschleunigt. Durch eindeutige Los-Stepper-Zuordnungen werden die Lose nicht nur aufgrund eigener Prioritäten weiterverarbeitet. Geringe Auslastungen geeigneter Maschinen führen zum Vorziehen von Losen und somit zur Beschleunigung ihrer Durchlaufzeiten.

Für alle Partner gilt: Die Bereitstellung der Informationen darüber, welche Partner verfügbar gemacht werden können, reduziert die Situationen, in denen ein Partner nicht verfügbar ist, auf ein Minimum. Die anderen Partner werden somit nicht belegt und stehen für andere Prozesse zur Verfügung.

Die Einführung der Dispatch-Own-GUI zeigt, dass durch die Reduzierung des Arbeitsaufwandes eines Produktionspartners, die Synchronisation aller Partner erhöht werden kann, was die Verfügbarkeit für andere Operationen steigert und somit eine Erhöhung der Kapazität verursacht. Verfügbarkeit und Synchronisationsgrad sind also zwei Größen, die nicht voneinander getrennt betrachtet werden können. Ebenso beeinflussen sich die Auswirkungen des Vier-Partner-Modells und der Betriebskennlinientheorie gegenseitig. Darüber hinaus verändert die Aufbereitung von Daten, also reines Informationsmanagement, die Richtung der Produktionssteuerung. Die Dispatch-Own-GUI hat den Übergang von einer schiebenden zu einer ziehenden Produktionssteuerung ermöglicht, womit die Auslastung der Maschinen gesteigert werden kann.

## 4.3 Dispatch Ahead

### 4.3.1 Steuerung der Voroperation

Ca. 70% der Gesamtkosten einer Fab entfallen auf die Produktionsmaschinen. Daher ist eine möglichst hohe Auslastung anzustreben. Die tatsächliche Auslastung der Maschinen liegt, je nach Maschinentyp bei 40-80%. D.h. nur 40-80% der gesamten verfügbaren Zeit arbeitet die Maschine für die Produktion [Han03]. Für die Jahre nach 2006 fordert die „International Technology Roadmap für Semiconductors 2003“ [ITRS03b] Auslastungen für Bottleneck-Anlagen von über 90%. Dabei werden diese Werte als noch nicht realisierbar eingestuft.

Um eine möglichst hohe Auslastung der Stepper zu erzielen, ist eine Vorgehensweise aus Sicht der Belichtungsanlagen sinnvoller als aus Sicht der Lose. Nicht allein die Dringlichkeitsliste der Lose soll deren Bearbeitungsreihenfolge bestimmen. Die Verfügbarkeit der Stepper muss den Ausschlag zur Bearbeitung passender Lose geben, die dann wiederum nach den bekannten Dispatch-Regeln abgearbeitet werden. Dabei genügt es nicht, erst dann zu reagieren, wenn die Auslastung eines oder einiger Stepper bereits stark abgesunken ist und aufgrund fehlender Lose in nächster Zeit nicht verbessert werden kann. Derartige Situationen treten meist zusammen mit einem Überangebot an Losen für bereits ausgelastete Stepper auf, so dass die Ursache der geringen Auslastung nicht sofort ersichtlich wird. Gerade dann, wenn viel Material bewegt werden muss, ist eine optimierte Auslastung aller Anlagen notwendig. Materialanhäufungen erhöhen die Variabilitäten in der Produktion und haben eine negative Auswirkung auf die Performance der gesamten Linie.

Die Auslastung von Anlagen eines Workcenters hängt somit auch von der Verfügbarkeit des Materials, dem WIP ab. Ob genügend WIP vorhanden ist, hängt wiederum von den Bewegungen der Voroperation ab. Da bei den Belichtungsanlagen eine detaillierte Dedizierung besteht, kommt es für deren Auslastung auch darauf an, welche Lose aus der Voroperation ankommen. Soll die Auslastung der Belichtungsanlagen gesteigert werden, so ist es notwendig, auf die Voroperation Einfluss zu nehmen, um die Lose in einer Reihenfolge zu bearbeiten, die eine optimale Auslastung der nachfolgenden Anlagen sicher stellt. Das heißt vor allem, dass verhindert wird, dass Anlagen wegen fehlenden WIPs leer laufen.

Um auf die Voroperation Einfluss zu nehmen, muss sie entweder von der folgenden Operation gesteuert werden oder Kenntnisse über die Situation in der Folgeoperation sowie die Prozessdaten der eigenen Lose bezüglich der Folgeoperation haben. Werden die Dispatch-Own-Daten der Belichtungsoperation an der Belackungsanlage angezeigt, so kann die Belackung Lose für spezielle Stepper in der vorgelagerten Bearbeitung vorziehen. Mit Kenntnis der aktuellen Auslastung der einzelnen Stepper kann somit die Steuerung der Voroperation Belackung realisiert werden. Diese Steuerung der Voroperation wird Dispatch Ahead genannt (siehe Abbildung 4.11).

Um mögliche Auswirkungen eines Dispatch-Ahead-Systems auf die Auslastung der Folgeoperation zu untersuchen, wird ein System mit zwei Arbeitsstationen in einer Simulationsumgebung modelliert.

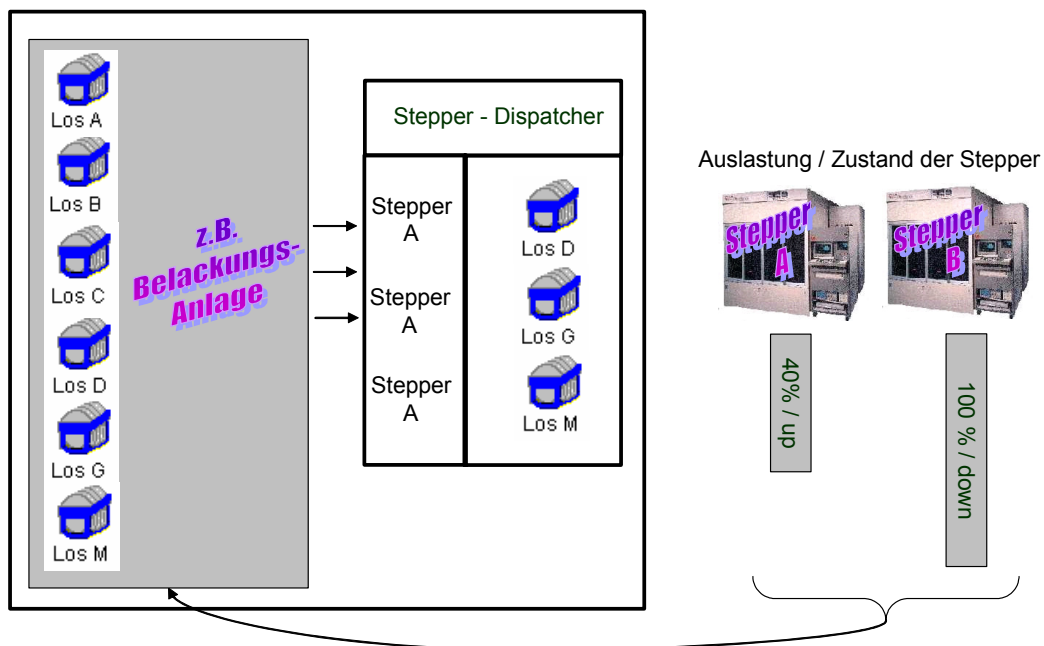


Abbildung 4.11: Steuerung der Voroperation durch Dispatch Ahead

### 4.3.2 Simulation zweier Arbeitsstationen

Nach [VDI96] ist ein Modell „eine vereinfachte Nachbildung eines existierenden oder gedachten (bzw. vergangenen) Systems mit seinen Prozessen in einem anderen begrifflichen oder gegenständlichen System. Es unterscheidet sich hinsichtlich der untersuchungsrelevanten Eigenschaften nur innerhalb eines vom Untersuchungsziel abhängigen Toleranzrahmens vom Vorbild. Es wird genutzt, um eine bestimmte Aufgabe zu lösen, deren Durchführung mittels direkter Operation am Original nicht mehr möglich oder zu aufwendig wäre.“

Die Untersuchung soll auf das Verhalten der Maschinen fokussiert werden, deshalb wird ein vereinfachtes Modell erstellt. Das vereinfachte Modell ist dabei schon stark idealisiert. So werden keine Transportwege implementiert, Operatoren sind nicht vorhanden, eventuelle Rüstzeiten sowie Ausfall- und Wartungszeiten werden nicht berücksichtigt.

Das Modell besteht aus einem Eingangslager vor Maschine 1 und drei Zwischenlager nach Maschine 1, die gleichzeitig als Eingangslager vor den drei Maschinen der Stufe 2 dienen (siehe Abbildung 4.12) [Now04]. Nach den Maschinen der Stufe 2 befinden sich drei Ausgangslager. Drei Zwischenlager wurden gewählt, um eine bessere graphische Übersicht zu erreichen.

Die Unterscheidung der Maschinen wird durch Farben (gelb, rot, blau) gekennzeichnet. Dieselbe Einteilung erhalten auch die zu verarbeitenden Lose, Loads genannt. Um eine einfache Dedizierung zu realisieren, können die Lose nur von den passenden Maschinen bearbeitet werden.

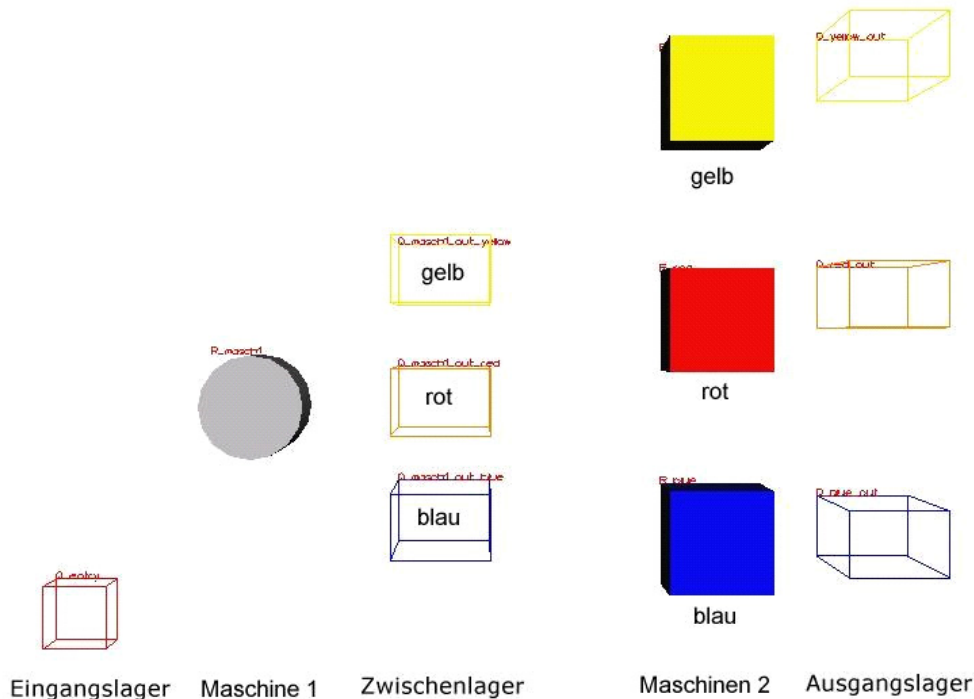


Abbildung 4.12: Modellgeometrie zur Simulation zweier Arbeitsstationen

Das Eingangslager wird nach und nach mit Losen gefüllt, abhängig von der gewählten Befüllungsmethode und der Zwischenankunftszeit. Die Befüllung des Eingangslagers kann nach zwei Methoden erfolgen. Entweder wird die Reihenfolge der Losfarben zufällig ermittelt, wobei das Verhältnis der einzelnen Farben angegeben werden kann (z.B. rot:blau:gelb entspricht 1:1:1). Oder die Reihenfolge erfolgt nach einem festen Schema, wobei die Anzahl der gleichfarbigen Lose angegeben wird (z.B. 10 rote, 10 blaue, 10 gelbe usw. in Folge). Aus dem Eingangslager wird nach dem FIFO-Prinzip das erste Los entnommen und zu Maschine 1 zur Verarbeitung geschickt. Die Verarbeitung in Maschine 1 erfolgt nach einer einstellbaren Verarbeitungszeit, die entweder aus einem fixen Wert oder aus einem, nach einer Wahrscheinlichkeitsverteilung variierenden Wert bestehen kann aber gleich für alle Lossorten ist. Nach der Verarbeitung in Maschine 1 wird das Los im entsprechenden Zwischenlager aufgenommen. Die Zwischenlager können mit einer Kapazitätsgrenze versehen werden, so dass sich nur eine maximal mögliche Anzahl von Losen in ihnen aufhalten kann. Ist in einem Zwischenlager die Kapazitätsgrenze erreicht, so kann kein weiteres Los der gleichen Farbe von Maschine 1 produziert werden. Durch das FIFO-Prinzip ist die Produktion in Maschine 1 dann so lange gestoppt, bis das entsprechende Zwischenlager wieder geringer als die Kapazitätsgrenze gefüllt ist. Die Lose in den Zwischenlagern werden wiederum nach dem FIFO-Prinzip behandelt. Ist die passende Maschine der Stufe 2 frei, wird das erste Los des Zwischenlagers weiterverarbeitet. Auch die Verarbeitungszeit der Maschinen der Stufe 2 können fix oder verteilt sein. Nach der Verarbeitung in einer Maschine der Stufe 2 wird das Los in das Ausgangslager geschickt und aus dem System entfernt.

Um die Einschwingphase des Systems zu verkürzen, befinden sich bei Beginn jeder Simulation jeweils 3 Lose in den Zwischenlagern.



Bei Verwendung von verteilten Zeiten werden jeweils vier Simulationen durchgeführt und von den Ergebnissen Mittelwerte berechnet. Versuche haben gezeigt, dass weitere Durchläufe nur geringfügige Veränderungen bewirken.

Die Auslastung der Maschinen wird vom Simulationstool AutoMod berechnet, wobei folgende Zusammenhänge verwendet werden: Die Anzahl der produzierten Lose (*total*) multipliziert mit der mittleren benötigten Zeit pro Los (*average time*) im Verhältnis zum Produkt aus Kapazität der Maschine (*capacity*) und gesamter Simulationszeit (*total simulation time*) ergibt die Auslastung (*utilisation*) der Anlage. Die Kapazität der Maschine ist dabei die Anzahl der Lose, die gleichzeitig in einer Maschine bearbeitet werden können.

$$\begin{aligned} utilisation &= \frac{total \cdot average\ time}{capacity \cdot total\ simulation\ time} \\ &= \frac{average}{capacity} \end{aligned} \quad (4.2)$$

$$average = \frac{total \cdot average\ time}{total\ simulation\ time} \quad (4.3)$$

Die Anzahl produzierter Lose gewichtet mit dem Verhältnis aus mittlerer Bearbeitungszeit eines Loses zur gesamten Simulationszeit ergibt die mittlere Anzahl der Lose in Bearbeitung. Somit ist die Auslastung gleich der mittleren Anzahl von Losen in Produktion pro Maschinenkapazität. Für eine Maschine mit der Kapazität *capacity* = 1 ist die durchschnittliche Anzahl der Lose in Bearbeitung gleich der Auslastung.

Zunächst werden die Eigenschaften des aufgestellten Modells ohne Dispatch Ahead untersucht. Dazu erfolgt eine Reihe von Durchläufen mit unterschiedlichen Einstellungen. Folgende Variationsmöglichkeiten werden in unterschiedlichen Kombinationen verwendet:

- Befüllungsmethode des Eingangslagers (Bef.)
  - 10 rote Lose, 10 gelbe Lose, 10 blaue Lose, usw. (10)
  - Zufällige Füllung mit gleichem Verhältnis der Farben (zuf)
- Zwischenankunftszeiten der Lose (ZAZ)
  - alle 10 Minuten (10)
  - alle 20 Minuten (20)
  - exponentiell verteilt mit Mittelwert  $\mu = 10$  Minuten (exp 10)
  - exponentiell verteilt mit Mittelwert  $\mu = 20$  Minuten (exp 20)
- Verarbeitungszeit Maschine 1 (VZ M1)
  - 20 Minuten (20)
  - normalverteilt mit  $\mu = 20$  Minuten und  $\sigma = 4$  Minuten (n 20,4)
- Verarbeitungszeit der Maschinen der Stufe 2 (VZ M2)
  - 60 Minuten (60)
  - normalverteilt mit  $\mu = 60$  Minuten und  $\sigma = 12$  Minuten (n 60,12)

- Kapazitätsgrenze der Zwischenlager (ZL)
  - bei 4 Stück pro Lager (4)
  - bei 7 Stück pro Lager (7)
  - bei 10 Stück pro Lager (10)
  - unbegrenzt (ub)

Tabelle 4.2 zeigt einige Simulationsergebnisse. Die ersten Spalten geben die Einstellungen wieder, die letzten vier Spalten enthalten die Ergebnisse in Form der Auslastung jeder einzelnen Maschine. „U M1“ bezeichnet die Auslastung der Maschine 1, „U M2r“ die Auslastung der Maschine der 2. Stufe für die roten Lose, „U M2b“ für die blauen Lose und „U M2g“ für die gelben.

Für fixe Zeiten (Nr. 1-8) ergeben sich identische Auslastungen sowohl für Zwischenankunftszeiten von 10 Minuten als auch von 20 Minuten. Sind die Zeiten verteilt, ist nur eine exponentiell verteilte Zwischenankunftszeit mit einem Mittelwert von  $\mu = 10$  Minuten sinnvoll, da sonst für das Dispatch-Ahead-Szenario keine Optimierungsmöglichkeit besteht. Befinden sich im Eingangslager zu wenig Lose, kann kein sinnvoller Einfluss auf die Reihenfolge ihrer Bearbeitung genommen werden.

Die Simulationen zeigen, dass die stärkste Auswirkung auf die Auslastung der Maschinen von der Begrenzung des Zwischenlagers ausgeht. Mit sinkenden Kapazitäten der Zwischenlager sinken auch die Auslastungen der Maschinen (Nr. 1-4).

Wird das Zwischenlager nicht systematisch sondern zufällig gefüllt (Nr. 5-8), erfolgt eine Auslastungsverbesserung bei geringer Lagerkapazität, während sich die Auslastung bei höherer Lagerkapazität verschlechtert. Erst bei unbegrenztem Zwischenlager ergibt sich eine nahezu 100%ige Auslastung. Dort steigen allerdings die Lagerbestände innerhalb der Simulationszeit bis auf über 60 Lose, mit steigender Tendenz für längere Simulationen. Hohe Lagerbestände sind jedoch nicht wünschenswert. Eine zufällige Befüllung ist also nur dann vorteilhaft, wenn sie eine bessere bzw. gleichmäßigere Auffüllung der Lager bewirkt als die systematische Befüllung. Ist das Lager groß genug, stellt die systematische Befüllung zu jeder Zeit sicher, dass von jeder Farbe ein genügend großer Vorrat vorhanden ist. Aufgrund des immer gleich bleibenden Schemas ist diese Befüllung einheitlicher als die zufällige Befüllung.

In realen Produktionssystemen der Halbleiterindustrie unterliegen alle Zeiten Schwankungen. Auch wenn das Eintreffen der Lose in einer realen Halbleiterfab keine statistisch unabhängigen Ereignisse darstellt, so sollen die Zwischenankunftszeiten als exponentiell verteilt angenommen werden. Für die Bearbeitungszeiten der Maschinen werden normalverteilte Zeiten verwendet. Die gleichen Simulationen werden erneut mit verteilten Zeiten durchgeführt.

Im Großen und Ganzen verringern sich die Auslastungen bei Verwendung von verteilten Zeiten (Nr. 9-16), weil es in gewissem Maß zu Maschinenwartezeiten kommen kann. Da sich gleichzeitig auch der Lagerbestand erhöht, lässt das darauf schließen, dass die Schwankungen der Maschinen der Stufe 2 dazu führen, dass die Zwischenlager langsamer geleert werden und somit schneller an ihre Kapazitätsgrenze stoßen. Dadurch wird Maschine 1 in ihrer Produktion unterbrochen, was zu geringerem Durchsatz und somit zu geringerer Auslastung der Maschinen führt.

Nr.	Bef.	ZAZ	VZ M1	VZ M2	ZL	U M1	U M2r	U M2b	U M2g
1.	10	10/20	20	60	4	0,622	0,624	0,622	0,622
2.	"	"	"	"	7	0,999	1	0,999	0,998
3.	"	"	"	"	10	1	1	1	0,998
4.	"	"	"	"	ub	1	1	1	0,998
5.	zuf	"	"	"	4	0,868	0,877	0,882	0,845
6.	"	"	"	"	7	0,918	0,909	0,913	0,931
7.	"	"	"	"	10	0,941	0,932	0,935	0,954
8.	"	"	"	"	ub	1	0,990	0,975	0,980
9.	10	exp 10	n 20, 4	n 60, 12	4	0,623	0,625	0,621	0,623
10.	"	"	"	"	7	0,974	0,971	0,974	0,970
11.	"	"	"	"	10	0,992	0,990	0,989	0,990
12.	"	"	"	"	ub	1	0,997	0,998	0,993
13.	zuf	"	"	"	4	0,848	0,844	0,847	0,850
14.	"	"	"	"	7	0,907	0,908	0,911	0,907
15.	"	"	"	"	10	0,929	0,920	0,931	0,927
16.	"	"	"	"	ub	1	0,984	0,986	0,984

Tabelle 4.2: Maschinenauslastungen unterschiedlicher Simulationen ohne Dispatch Ahead

Die insgesamt hohen Auslastungen zeigen, dass es sich um ein stark idealisiertes System handelt. Es ist offensichtlich, dass Performancesteigerungen nur in geringem Maße möglich sein können. Dennoch soll eine Auslastungssteigerung durch die Dispatch-Ahead-Systematik erreicht werden.

### 4.3.3 Simulation eines Dispatch Ahead Szenarios

Um ein Dispatch-Ahead-Szenario zu modellieren wird das oben beschriebene Modell modifiziert. Für die Lose im Eingangslager wird eine so genannte Orderliste eingerichtet. In Orderlisten können Lose unabhängig ihres physikalischen Aufenthaltsortes eingetragen werden. Sie verbleiben so lange in der Orderliste und auch an ihrem physikalischen Ort, bis sie geordert werden („wait to be ordered“). Nachdem ein Los geordert wurde, fährt es in seinem Ablaufplan fort („order load to continue“). Jedes Los, das nun das Eingangslager erreicht, wird auf diese Orderliste gesetzt und wartet so lange, bis es geordert wird. Eine spezielle Order-Routine überprüft laufend die Situation und legt fest, welches Los geordert werden soll. Da diese Order die zu ordernden Lose entsprechend des Warenbestandes vor den Maschinen der Stufe 2 bestimmen, ist diese Vorgehensweise eine Realisierung eines Dispatch-Ahead-Systems.

Immer wenn Maschine 1 leer und das Eingangslager mit mindestens einem Los gefüllt ist, wird ein Los geordert. Dabei wird ein Los derjenigen Farbe geordert, deren Zwischenlager am geringsten befüllt ist. Kann kein Zwischenlager als am geringsten befüllt ermittelt werden, wird der Reihe nach eine Farbe nach der anderen geordert.

Kann eine Order nicht erfüllt werden, weil sich kein passendes Los im Eingangslager befindet, wird nach der nächsten passenden Farbe gesucht oder ein Lostyp nach einer festen Reihenfolge gewählt.

Die Anzahl der Lose in den Zwischenlagern kann wieder begrenzt werden. Dabei erfolgt eine Order nur dann, wenn der Lagerbestand eines Lostyps (einer Farbe) unterhalb einer Ordergrenze liegt. Diese Ordergrenze führt dann automatisch dazu, dass die Zwischenlager eine entsprechende Kapazitätsgrenze erhalten.

Das Simulations-Szenario in Abbildung 4.13 zeigt das Eingangslager mit einer zufälligen Füllung von Losen. Diese Lose stehen auf der virtuellen Orderliste für eine Order zur Verfügung. In Maschine 1 wird gerade ein rotes Los bearbeitet, das zuvor geordert wurde. Der Grund für diese Order ist in der Füllung der Zwischenlager ersichtlich. Während im gelben Lager zwei und im blauen Lager drei Lose warten, ist das rote Lager leer. Alle drei Maschinen der Stufe 2 bearbeiten gerade ein entsprechendes Los. Da während der Bearbeitungszeit der Maschine 1 die Maschinen der Stufe 2 maximal ein Los fertig stellen können (ihre Bearbeitungszeit ist im Durchschnitt drei mal so lange wie die der Maschine 1), wird als nächstes Los wieder ein rotes geordert werden.

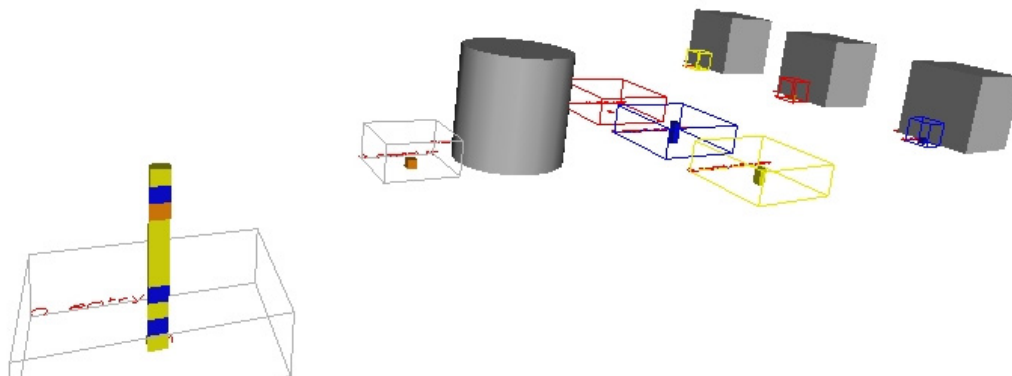


Abbildung 4.13: Simulations-Szenario des Dispatch Ahead

Tabelle 4.3 zeigt die Simulationsergebnisse für die gleichen Eingangswerte unter Verwendung der Dispatch-Ahead-Funktion.

Auffallend ist, dass bei festen Zeiten (Nr. 1-8) nahezu immer 100% Auslastung erreicht wird. Dadurch, dass die Zeiten von Maschine 1 und den Maschinen der Stufe 2 genau zueinander passen (wie es von den Fab-Designern angelegt sein sollte) und keinerlei Schwankungen im Zeitablauf passieren, kann Maschine 1 innerhalb der gleichen Zeit, in denen die Maschinen der Stufe 2 ihre Lose bearbeiten, genügend Nachschub herbeibringen. Die Dispatch-Ahead-Funktion sorgt dafür, dass für jede Maschine der Stufe 2 gleichmäßig Lose bereitgestellt werden. Sind genügend Lose im Eingangslager, führt dies zu einer 100%igen Auslastung.

Bei verteilten Zeiten (Nr. 9-16) sind die Auslastungen ebenfalls kontinuierlich hoch, wenn auch nicht ganz 100%.

Nr.	Bef.	ZAZ	VZ M1	VZ M2	ZL	U M1	U M2r	U M2b	U M2g
1.	10	10	20	60	4	0,999	1	1	1
2.	"	"	"	"	7	1	1	1	1
3.	"	"	"	"	10	1	1	1	1
4.	"	"	"	"	ub	1	1	1	1
5.	zuf	"	"	"	4	1	1	1	1
6.	"	"	"	"	7	1	1	1	1
7.	"	"	"	"	10	1	1	1	1
8.	"	"	"	"	ub	1	1	1	1
9.	10	exp 10	n 20, 4	n 60, 12	4	0,993	0,995	0,995	0,996
10.	"	"	"	"	7	0,995	0,997	0,998	0,997
11.	"	"	"	"	10	0,996	0,997	0,997	0,996
12.	"	"	"	"	nb	0,996	0,997	0,997	0,996
13.	zuf	"	"	"	4	0,994	0,995	0,995	0,995
14.	"	"	"	"	7	0,996	0,996	0,995	0,996
15.	"	"	"	"	10	0,996	0,994	0,993	0,994
16.	"	"	"	"	ub	0,996	0,994	0,993	0,994

Tabelle 4.3: Maschinenauslastungen unterschiedlicher Simulationen mit Dispatch Ahead

#### 4.3.4 Performancesteigerung durch Dispatch Ahead

In Tabelle 4.4 sind die Differenzen der Auslastungen ohne und mit Dispatch Ahead sowie die daraus resultierenden Verbesserungen in % angegeben.

Für Simulation Nr. 1 bedeutet die Verwendung der Dispatch-Ahead-Funktion eine Verbesserung der Maschinenauslastung um bis zu 38 Prozentpunkte, was einer Erhöhung der Auslastung um 61% entspricht.

Besonders für die niedrige Auslastung bei geringem Zwischenlager und ungleichmäßiger Befüllung des Eingangslagers kann die Order-Routine große Verbesserungen erzielen. Durch die Order werden die Lose weiterverarbeitet, die gerade benötigt werden. So muss, im Gegensatz zur Situation ohne Dispatch Ahead, keine Maschine mit leerem Zwischenlager auf ein Los warten, während das Zwischenlager einer anderen Maschine vollständig gefüllt

ist und Maschine 1 nicht weiterarbeiten kann. Dies gilt auch für die zufällige Füllung der Eingangslager und verteilte Bearbeitungszeiten.

Nr.	Differenz der Auslastung				Verbesserung in %			
	M1	M2r	M2b	M2g	M1	M2r	M2b	M2g
1.	0,377	0,376	0,378	0,378	60,61	60,26	60,77	60,77
2.	0,001	0	0,001	0,002	0,10	0,00	0,10	0,20
3.	0	0	0	0,002	0,00	0,00	0,00	0,20
4.	0	0	0	0,002	0,00	0,00	0,00	0,20
5.	0,132	0,123	0,118	0,155	15,21	14,03	13,38	18,34
6.	0,082	0,091	0,087	0,069	8,93	10,01	9,53	7,41
7.	0,059	0,068	0,065	0,046	6,27	7,30	6,95	4,82
8.	0	0,01	0,025	0,02	0,00	1,01	2,56	2,04
9.	0,37	0,37	0,374	0,373	59,39	59,20	60,23	59,87
10.	0,021	0,026	0,024	0,027	2,16	2,68	2,46	2,78
11.	0,004	0,007	0,008	0,006	0,40	0,71	0,81	0,61
12.	-0,004	0	-0,001	0,003	-0,40	0,00	-0,10	0,30
13.	0,146	0,151	0,148	0,145	17,22	17,89	17,47	17,06
14.	0,089	0,088	0,084	0,089	9,81	9,69	9,22	9,81
15.	0,067	0,074	0,062	0,067	7,21	8,04	6,66	7,23
16.	-0,004	0,01	0,007	0,01	-0,40	1,02	0,71	1,02

Tabelle 4.4: Performancesteigerung durch Dispatch Ahead

Bei verteilten Zeiten und unbegrenztem Zwischenlager sind zwar die Verbesserungen der Auslastungen nicht sehr hoch, dafür sorgt das Dispatch-Ahead-System für geringere Mengen im Zwischenlager. Während in Simulation Nr. 12 ohne Order im Zwischenlager Bestände zwischen durchschnittlich 4 bis 11 Lose (Maximalwert 33 Lose) aufgebaut werden, erreichen die Lager mit Order nur Füllungen von durchschnittlich 3 Lose (Maximalwert 9 Lose). Bei zufälliger Befüllung des Eingangslagers (Nr. 16) ist die Lagerreduzierung noch größer: Von durchschnittlich 31 bis 45 Losen (Maximalwert 122 Lose) erfolgt eine Reduzierung auf durchschnittlich 3 Lose (Maximalwert 9 Lose).

Durch Dispatch Ahead wird die Auslastung der Maschinen unabhängig von der Größe des Zwischenlagers und der Gleichmäßigkeit der Befüllung des Eingangslagers erhöht. Bis auf drei Werte (Auslastungen bei Nr. 12 und 16) konnten alle Auslastungen verbessert werden. Darüber hinaus erfolgt eine Verringerung der Lagerbestände.

Für reale Produktionssysteme gilt: Ist das Zwischenlager groß genug, um durch Vorratshaltung eine einseitige Anlieferung kompensieren zu können, kann auch ohne Dispatch Ahead eine hohe Auslastung der Maschinen erzielt werden. Da große

Lagerbestände aber die Gesamtperformance einer Fab verschlechtern, sollten sie abgebaut werden. Mit Hilfe einer ausgewogenen Anlieferung bzw. eines Dispatch-Ahead-Systems wird eine hohe Auslastung bei gleichzeitig geringen Lagerbeständen gewährleistet.

Eine Performancesteigerungs-Maßnahme ist um so erfolgreicher, je „schlechter“ ein System ist. Je näher ein Produktionssystem dem Optimum kommt, um so weniger ist es möglich, die Performance zu steigern. Das hier untersuchte Modell ist bereits stark vereinfacht und optimiert. Dennoch konnten mit Hilfe des Dispatch-Ahead-Verfahrens Verbesserungen in der Auslastung der Maschinen erzielt werden. Um so größer wird das Potential des Dispatch Ahead für eine reale Produktion mit weitaus schlechterer Anfangsperformance sein.

### **4.3.5 Rechenkomplexität und Implementierung**

Als Simulationstool wurde AutoMod gewählt. AutoMod ist eine Software, die zur Erstellung von Produktionsumgebungen verwendet werden kann. Sie arbeitet nach dem Prinzip der prozessorientierten Simulation. Dabei wird die Reihenfolge der Schritte, die ein Objekt in einem System durchläuft, betrachtet. Der Status eines Objektes wird Schritt für Schritt simuliert.

Als Studentenversion kann AutoMod kostenlos unter [www.brooks.com](http://www.brooks.com) inklusive Anleitung bezogen werden. Im Gegensatz zur lizenzierten Version ist die Anzahl unterschiedlicher Einheiten auf 100 Stück begrenzt. Die Simulationen wurden mit AutoMod, Version 11.0, auf einem Pentium 4 (Samsung X10 Notebook mit Intel Centrino) durchgeführt. Für die einzelnen Simulationen wird jeweils ein Zeitraum von 100 Tagen zu Grunde gelegt. Bei ausgeschalteter graphischer Darstellung benötigen die einzelnen Durchläufe ca. 1 Minute.

## **4.4 Zusammenfassung und Ausblick**

Um die Synchronisation der Produktionsfaktoren zu erhöhen und die Verteilung der Lose auf die Anlagen (hier: Stepper) zu vereinfachen, wurde ein Dispatch-Own-System entwickelt und in der Praxis erprobt. Eine Dispatch-Ahead-Funktion steuert die Produktionsreihenfolge einer Arbeitsstation entsprechend der Bedürfnisse der nachfolgenden Station. Um ihre Eigenschaften zu untersuchen, wurde sie in einer Simulationsumgebung modelliert.

Die Realisierung der Dispatch-Own-Funktion trägt nicht nur zur Verbesserung der Performance und Erhöhung der Kapazität der Fotolithografie bei. Sie ist darüber hinaus auch ein wichtiger Schritt in Richtung der papierlosen Fertigung (Paperless Fab). Sind alle benötigten Informationen in elektronischer Datenform vorhanden, kann auch ein automatischer Übertrag von Los- und Prozessdaten an die Anlagen eingerichtet werden. Damit wird die Bearbeitung schneller und die Fehlermöglichkeiten weiter verringert. Es muss allerdings auch bedacht werden, dass sich manche Erfahrungen und Geschicklichkeiten eines Menschen nicht immer problemlos in Soft- und Hardware umsetzen lassen.

Darüber hinaus gibt es in der Produktionstechnik immer konkurrierende Ziele, die beachtet werden müssen und somit zu Kompromissen führen. Eine derartige Fragestellung könnte lauten: Ist es günstiger, ein Los warten zu lassen und somit Wegezeit zu sparen oder sollte es besser gleich prozessiert werden, da sonst eine CLIP-Verletzung riskiert wird.

Die Dispatch-Ahead-Funktion bietet die Möglichkeit, Maschinenauslastungen zu erhöhen bei gleichzeitig verringerten Lagerbeständen. Somit bietet sie ein Werkzeug, um Bottleneck-Anlagen zu optimieren. Zusätzlich könnte ein automatisches Beschaffungssystem für Material, Werkzeuge und Maschinen eingerichtet werden. Schon im voraus könnten Anlagen freigehalten, Programme geschrieben oder Werkzeuge herbeigebracht werden, noch bevor das Los den Prozess erreicht.

Moderne Belichtungsanlagen sind so genannte Cluster-Tools, in denen die Operationen Belackern, Belichten und Entwickeln in direkter Folge ausgeführt werden. Damit erübrigt sich eine Anwendung des Dispatch Ahead, in dem die gesteuerte Voroperation die Belackung ist. Für jede andere Folge von Operationen bleiben die gewonnenen Erkenntnisse natürlich bestehen. Die Anwendung der Dispatch-Ahead-Methode wird für jede Arbeitsstation mit Engpasscharakter Vorteile für die Auslastung und Bewegungsleistung erbringen.



# Kapitel 5

## Resumé

Die Arbeiten der letzten Jahre haben gezeigt, dass die Halbleiterproduktionstechnik ein interdisziplinäres Fachgebiet ist. Kenntnisse aus der Elektrotechnik, der Physik, der Mathematik, der Informatik und den Wirtschaftswissenschaften werden benötigt. Dabei ist auch das Gebiet des Operations Research zu nennen. Deshalb ist es hilfreich und notwendig, dass Personen aus den unterschiedlichen Fachrichtungen zusammenarbeiten. In allen diesen Fachbereichen finden sich daher auch Institute und Forschungseinrichtungen, die sich u.a. mit Fragestellungen aus der Halbleiterproduktionstechnik befassen. Die spezielle Ausrichtung alleine auf die Halbleiterproduktionstechnik, ist dabei jedoch neu.

Die wissenschaftliche Behandlung der Halbleiterproduktionstechnik eröffnet ein großes Gebiet mit interessanten Aufgabenstellungen. Dabei ist es sinnvoll, Kontakte zu einer realen Produktionsstätte zu haben. So kann vermieden werden, dass Forschungen und Entwicklungen an der Realität vorbei erfolgen. Schon Unterschiede in kleinen Details können zu Ergebnissen führen, die in der Praxis nicht mehr anwendbar sind.

Der Abstand zum alltäglichen Tagesgeschäft einer Halbleiterfirma aus der Sicht einer Universität kann also einerseits hinderlich sein, wenn eine industriennahe Forschung betrieben werden soll. Andererseits erlaubt er jedoch die Behandlung eines Themas aus einer viel grundlegenderen Sichtweise wie sie in einer Firma aus gegebenen Termin- und Zeitgründen oft nicht möglich ist. Somit ist es geradezu notwendig, Inhalte der Halbleiterproduktionstechnik an einer Universität zu etablieren und die Verbindung zur Grundlagenforschung herzustellen.

Die Einführung der Halbleiterproduktionstechnik in der Lehre bereichert nicht nur die Themenauswahl des Vorlesungsverzeichnisses. Vielmehr ist es für die Ausbildung der Studenten unerlässlich, produktionstechnische Themen zu bearbeiten. Gerade für die Ingenieure der Elektro- und Informationstechnik gab es bisher keine Veranstaltung mit entsprechenden Inhalten.

In dieser Arbeit wurden konkrete Themen für das Gebiet Management/Organisation bearbeitet. Im Bereich der Planung wurden Zusammenhänge der Liefertreue untersucht, für die Steuerung und Logistik wurden Dispatch-Systeme entwickelt. Das zeigt, dass in jedem Bereich praxisrelevante Aufgabenstellungen vorhanden sind. Darüber hinaus können

Simulationswerkzeuge angewandt werden. Aufgrund der Komplexität der Halbleiterproduktion scheinen die Fragestellungen sehr detailliert und spezifisch zu sein. Die hier bearbeiteten Beispiele zeigen jedoch, dass grundlegende Aussagen zur Produktionsverbesserung erarbeitet werden können.

Die Halbleiterindustrie befindet sich bereits in ihrer Konsolidierungsphase [Woj02]. Das abgeschwächte Trendwachstum, die Zersplitterung der Wertschöpfungskette (Foundries, Fabless Designers) und die Standardisierung der Produkte führen zu einem höheren Konkurrenzdruck. Die Behandlung von Themen zur Optimierung der Supply Chain und zur Steigerung der Produktivität werden dadurch immer wichtiger. Die Aufgaben der Halbleiterproduktionstechnik werden somit umfangreicher und bedeutender.

# Abbildungsverzeichnis

Abbildung 1.1: Pro Jahr und Transistorfunktion werden 30% Kostenreduktion benötigt.....	1
Abbildung 1.2: Grund- und Hauptstudium sowie Studienrichtungen des Studiengangs Produktionstechnik an der Fachhochschule Köln.....	2
Abbildung 1.3: Tätigkeitsfelder der Wissenschaftlichen Gesellschaft für Produktionstechnik .....	3
Abbildung 1.4: Betrachtungsebenen einer Produktion .....	4
Abbildung 1.5: Ausgewählte Arbeitsgebiete der Halbleiterproduktionstechnik.....	5
Abbildung 2.1: Moore´s Law [Han03] .....	9
Abbildung 2.2: Entwicklung der Transistorenanzahl pro Chip (Quelle: Intel).....	10
Abbildung 2.3: Entwicklung der Chipmenge pro Wafer [Han04] .....	11
Abbildung 2.4: Entwicklung der Kosten für Halbleiterfabs (Quelle: Texas Instruments).....	11
Abbildung 2.5: Aufteilung der Halbleiterproduktionslinien in Frontend und Backend .....	12
Abbildung 2.6: Herstellung einer Schicht einer Halbleiterscheibe [Han04].....	14
Abbildung 2.7: Weg einer einzelnen Scheibe (bzw. Los) während der Produktion in einer Halbleiterfab [Hei02] .....	15
Abbildung 2.8: Die interne Supply Chain der Halbleiterindustrie kann in drei Stufen eingeteilt werden .....	20
Abbildung 2.9: Maschinenzustände nach SEMI E10.....	22
Abbildung 2.10: Verluste der Maschinenzeiten .....	25
Abbildung 2.11: Verluste der Overall Equipment Efficiency .....	25
Abbildung 2.12: Ein Produktionssystem besteht aus den vier Ms: Material, Maschine, Mensch und Methode .....	26
Abbildung 2.13: Erhöhung der produktiven Zeit durch Synchronisation der Produktionsfaktoren.....	27
Abbildung 2.14: Maßnahmen zur Verbesserung der Verfügbarkeiten und des Synchronisationsgrades .....	28
Abbildung 2.15: Wichtige Kennzahlen: Flussfaktor und Little´s Law .....	29

Abbildung 2.16: Aufteilung der Operationen in Ranges .....	30
Abbildung 2.17: Taktrate je Operation und Produkt am Beispiel von fünf Losen mit gleicher Prozessfolge aber unterschiedlicher Einschleusung .....	31
Abbildung 2.18: Dispatch-Regeln.....	33
Abbildung 2.19: Aufnahme der Time-Stamps .....	34
Abbildung 2.20: Formen der Going Rate .....	34
Abbildung 2.21: Formen der Cycle Time.....	35
Abbildung 2.22: Größen der Warteschlangentheorie .....	40
Abbildung 2.23: Kendall-Notation für Wartesysteme .....	41
Abbildung 2.24: Von der Kingman-Gleichung zur Betriebskennlinie.....	41
Abbildung 2.25: Die Betriebskennlinie (BKL, Operating Curve Management, OCM) .....	42
Abbildung 2.26: Optimierte Arbeitspunkte durch Verringerung der Variabilität $\alpha$ .....	44
Abbildung 2.27: Optimierte Arbeitspunkte durch Erhöhung der Kapazität <i>Capa</i> .....	45
Abbildung 3.1: Darstellung kontinuierlicher und diskreter Zufallsvariablen.....	48
Abbildung 3.2: Das $\alpha$ -Quantil $x_\alpha$ ist der x-Wert für den gilt: $P(X \leq x_\alpha) = \alpha$ .....	49
Abbildung 3.3: Dichtefunktion einer normalverteilten kontinuierlichen Zufallsvariablen mit $\mu = 6$ und $\sigma = 2,2$ .....	50
Abbildung 3.4: Umrechnungsschema zum Arbeiten mit der Tabelle der Standardnormalverteilung .....	50
Abbildung 3.5: Diskretisierung einer Dichtefunktion mit unterschiedlichen Diskretisierungskonstanten .....	52
Abbildung 3.6: Faltung zweier Wahrscheinlichkeitsfunktionen .....	53
Abbildung 3.7: Weg einer Bestellung von der Einschleusung bis zur Auslieferung beim Kunden.....	54
Abbildung 3.8: Notwendige Durchlaufzeiten zum Erreichen der Planlieferwoche bei einer Plandurchlaufzeit von 4 Wochen.....	56
Abbildung 3.9: Wahrscheinlichkeitsverteilung der Durchlaufzeit mit einem Mittelwert von $\mu = 28$ Tagen und einer Standardabweichung von $\sigma = 5$ Tagen bei einer Diskretisierungskonstante von $\Delta t = 1$ Tag .....	58
Abbildung 3.10: Wahrscheinlichkeitsverteilung der Einschleusung .....	59
Abbildung 3.11: Wahrscheinlichkeiten zum Erreichen bzw. Über- und Unterschreiten der Lieferwoche bei gleichmäßiger Einschleusung, $CT^* \sim N(\mu = 28 \text{ d}, \sigma = 5 \text{ d})$ und einer variierenden <i>CTP</i> .....	59
Abbildung 3.12: Wahrscheinlichkeiten zur rechtzeitigen bzw. verspäteten Lieferung bei gleichmäßiger Einschleusung für $CT^* \sim N(\mu = 28 \text{ d}, \sigma = 5 \text{ d})$ und einer variierenden <i>CTP</i> .....	60
Abbildung 3.13: Beispiel von zur Ermittlung der <i>OTD</i> notwendigen Zeitgrößen .....	63
Abbildung 3.14: Vereinfachte Darstellung der diskretisierten Wahrscheinlichkeitsfunktion.....	64

Abbildung 3.15: Wahrscheinlichkeitsfunktionen von Einschleusung $E^*$ , Durchlaufzeit $CT^*$ und Fertigstellung $C^*$ mit On-Time-Delivery für $CT^* \sim N(\mu = 43 \text{ d}, \sigma = 7 \text{ d})$ , $CTP = 49 \text{ d}$ .....	65
Abbildung 3.16: On-Time-Delivery für normalverteilte Durchlaufzeiten mit unterschiedlichem Mittelwert und Standardabweichung bei einer Plandurchlaufzeit von 49 Tagen .....	66
Abbildung 3.17: Linien gleicher On-Time-Delivery für normalverteilte Durchlaufzeiten mit unterschiedlichem Mittelwert und Standardabweichung bei einer Plandurchlaufzeit von 49 Tagen .....	66
Abbildung 3.18: <i>OTD</i> in Abhängigkeit von Mittelwert oder Standardabweichung der Durchlaufzeit bei einer Plandurchlaufzeit von 49 Tagen .....	67
Abbildung 3.19: Vergleich der Wahrscheinlichkeitsfunktionen von Fertigstellung und einer Normalverteilung mit gleichen Parametern .....	68
Abbildung 3.20: Die fünf Werte der On-Time-Delivery .....	71
Abbildung 3.21: Berechnung der <i>OTDf</i> für $CTf^* \sim N(\mu f = 43 \text{ d}, \sigma f = 7 \text{ d})$ und $CTPf = 42 \text{ d}$ .....	73
Abbildung 3.22: Berechnung der <i>OTDb</i> für $CTb^* \sim N(\mu b = 7 \text{ d}, \sigma b = 2 \text{ d})$ und $CTPb = 7 \text{ d}$ .....	75
Abbildung 3.23: Wirkungsweise des zeitdiskreten Operators der Stutzung .....	76
Abbildung 3.24: Wahrscheinlichkeitsfunktionen der Fertigstellung Frontend sowie der auf die <i>DDf</i> gestutzten Fertigstellung für $CTf^* \sim N(\mu f = 43 \text{ d}, \sigma f = 7 \text{ d})$ und $CTPf = 42 \text{ d}$ .....	77
Abbildung 3.25: Ermittlung der Wahrscheinlichkeitsfunktion der Einschleusung Backend aus der gestutzten Wahrscheinlichkeitsfunktion der Fertigstellung Frontend für die Einschleuswoche .....	78
Abbildung 3.26: Wahrscheinlichkeitsfunktion Einschleusung Backend $Rb^*$ für $CTf^* \sim N(\mu f = 43 \text{ d}, \sigma f = 7 \text{ d})$ und $CTPf = 42 \text{ d}$ .....	79
Abbildung 3.27: Berechnung der <i>OTDm</i> für $CTf^* \sim N(\mu f = 43 \text{ d}, \sigma f = 7 \text{ d})$ , $CTPf = 42 \text{ d}$ , $CTb^* \sim N(\mu b = 7 \text{ d}, \sigma b = 2 \text{ d})$ und $CTPb = 7 \text{ d}$ ; .....	80
Abbildung 3.28: Wahrscheinlichkeitsfunktion der Einschleusung Delivery für $CTf^* \sim N(\mu f = 43 \text{ d}, \sigma f = 7 \text{ d})$ , $CTPf = 42 \text{ d}$ , $CTb^* \sim N(\mu b = 7 \text{ d}, \sigma b = 2 \text{ d})$ und $CTPb = 7 \text{ d}$ .....	82
Abbildung 3.29: <i>OTDo</i> für $CTf^* \sim N(\mu f = 43 \text{ d}, \sigma f = 7 \text{ d})$ , $CTPf = 42 \text{ d}$ , $CTb^* \sim N(\mu b = 7 \text{ d}, \sigma b = 2 \text{ d})$ , $CTPb = 7 \text{ d}$ , $CTd^* \sim N(\mu d = 3 \text{ d}, \sigma d = 0,5 \text{ d})$ und $CTPd = 4 \text{ d}$ .....	83
Abbildung 3.30: <i>OTD</i> -Berechnungen für $CTf^* \sim N(\mu f = 43 \text{ d}, \sigma f = 7 \text{ d})$ ; $CTPf = 56 \text{ d}$ ; $CTb^* \sim N(\mu b = 7 \text{ d}, \sigma b = 2 \text{ d})$ ; $CTPb = 7 \text{ d}$ ; $CTd^* \sim N(\mu d = 3 \text{ d}, \sigma d = 0,5 \text{ d})$ ; $CTPd = 4 \text{ d}$ .....	84
Abbildung 3.31: <i>OTD</i> -Berechnungen für $CTf^* \sim N(\mu f = 43 \text{ d}, \sigma f = 7 \text{ d})$ ; $CTPf = 42 \text{ d}$ ; $CTb^* \sim N(\mu b = 7 \text{ d}, \sigma b = 2 \text{ d})$ ; $CTPb = 7 \text{ d}$ ; $CTd^* \sim N(\mu d = 3 \text{ d}, \sigma d = 0,5 \text{ d})$ ; $CTPd = 4 \text{ d}$ .....	85
Abbildung 3.32: On-Time-Delivery für $CTf^* \sim N(\mu f, \sigma f = 7 \text{ d})$ ; $CTPf = 42 \text{ d}$ ; $CTb^* \sim N(\mu b = 7 \text{ d}, \sigma b = 2 \text{ d})$ ; $CTPb = 7 \text{ d}$ .....	86
Abbildung 3.33: Gegenläufige <i>OTD</i> -Entwicklungen .....	87

Abbildung 3.34: On-Time-Delivery für $CTf^* \sim N(\mu f, \sigma = 7 \text{ d})$ , $CTP_f = 49 \text{ d}$ , $CTb^* \sim N(\mu = 10 \text{ d}, \sigma = 2 \text{ d})$ und $CTP_b = 7 \text{ d}$ .....	88
Abbildung 3.35: On-Time-Delivery für $CTf^* \sim N(\mu f, \sigma = 7 \text{ d})$ , $CTP_f = 49 \text{ d}$ , $CTb^* \sim N(\mu = 7 \text{ d}, \sigma = 2 \text{ d})$ und $CTP_b = 7 \text{ d}$ .....	89
Abbildung 3.36: On-Time-Delivery für $CTf^* \sim N(\mu f, \sigma = 7 \text{ d})$ , $CTP_f = 49 \text{ d}$ , $CTb^* \sim N(\mu = 7 \text{ d}, \sigma = 2 \text{ d})$ und $CTP_b = 7 \text{ d}$ .....	89
Abbildung 3.37: On-Time-Delivery für $CTf^* \sim N(\mu f, \sigma = 7 \text{ d})$ , $CTP_f = 42 \text{ d}$ , $CTb^* \sim N(\mu = 7 \text{ d}, \sigma = 2 \text{ d})$ und $CTP_b = 7 \text{ d}$ .....	90
Abbildung 3.38: On-Time-Delivery für $CTf^* \sim N(\mu f, \sigma = 7 \text{ d})$ , $CTP_f = 42 \text{ d}$ , $CTb^* \sim N(\mu = 7 \text{ d}, \sigma = 2 \text{ d})$ und $CTP_b = 7 \text{ d}$ .....	90
Abbildung 3.39: On-Time-Delivery für $CTf^* \sim N(\mu f, \sigma = 7 \text{ d})$ , $CTb^* \sim N(\mu b = 7 \text{ d}, \sigma b = 2 \text{ d})$ und $CTP_b = 7 \text{ d}$ .....	92
Abbildung 3.40: On-Time-Delivery für $CTf^* \sim N(\mu f = 44 \text{ d}, \sigma f)$ , $CTP_f = 49 \text{ d}$ , $CTb^* \sim N(\mu b = 7 \text{ d}, \sigma b = 2 \text{ d})$ und $CTP_b = 7 \text{ d}$ .....	94
Abbildung 3.41: On-Time-Delivery für $CTf^* \sim N(\mu f = 41 \text{ d}, \sigma f)$ , $CTP_f = 42 \text{ d}$ , $CTb^* \sim N(\mu b = 7 \text{ d}, \sigma b = 2 \text{ d})$ und $CTP_b = 7 \text{ d}$ .....	95
Abbildung 3.42: On-Time-Delivery für $CTf^* \sim N(\mu f, \sigma f = 10 \text{ d})$ , $CTb^* \sim N(\mu b = 7 \text{ d}, \sigma b = 2 \text{ d})$ und $CTP_b = 7 \text{ d}$ .....	95
Abbildung 3.43: On-Time-Delivery für $CTf \sim N(\mu f, \sigma f = 4 \text{ d})$ , $CTb \sim N(\mu b = 7 \text{ d}, \sigma b = 2 \text{ d})$ und $CTP_b = 7 \text{ d}$ .....	96
Abbildung 3.44: On-Time-Delivery für $CTf^* \sim N(\mu f, \sigma f = 7 \text{ d})$ , $CTb^* \sim N(\mu b = 30 \text{ d}, \sigma b = 7 \text{ d})$ und $CTP_b = 35 \text{ d}$ .....	96
Abbildung 3.45: On-Time-Delivery für $CTf \sim N(\mu f, \sigma f = 4 \text{ d})$ , $CTb \sim N(\mu b = 30 \text{ d}, \sigma b = 7 \text{ d})$ und $CTP_b = 35 \text{ d}$ .....	97
Abbildung 3.46: Wird von wochenfeinen Plandurchlaufzeiten im Frontend a) zu tagesfeinen Plandurchlaufzeiten übergegangen, so müssen sich auch der Stichtag sowie die Einschleusung ins Backend vom Wochenraster lösen d) .....	98
Abbildung 3.47: On-Time-Delivery für $CTf^* \sim N(\mu f = 40 \text{ d}, \sigma = 7 \text{ d})$ , $CTb^* \sim N(\mu = 7 \text{ d},$ $\sigma = 2 \text{ d})$ und $CTP_b = 7 \text{ d}$ .....	100
Abbildung 3.48: OTD ohne Release Backend für $CTf^* \sim N(\mu f = 43 \text{ d}, \sigma f = 7 \text{ d})$ ; $CTP_f = 42 \text{ d}$ ; $CTb^* \sim N(\mu b = 7 \text{ d}, \sigma b = 2 \text{ d})$ ; $CTP_b = 10 \text{ d}$ ; $CTd^* \sim N(\mu d = 3 \text{ d}, \sigma d = 0,5 \text{ d})$ ; $CTP_d = 4 \text{ d}$ .....	105
Abbildung 3.49: On-Time-Delivery Daily ohne Release Backend für $CTf^* \sim N(\mu f, \sigma f = 7 \text{ d})$ , $CTb^* \sim N(\mu b = 7 \text{ d}, \sigma b = 2 \text{ d})$ und $CTP_b = 7 \text{ d}$ .....	106
Abbildung 3.50: On-Time-Delivery Daily ohne Release Backend für $CTf^* \sim N(\mu, \sigma = 7 \text{ d})$ , $CTb^* \sim N(\mu = 7 \text{ d}, \sigma = 2 \text{ d})$ und $CTP_b = 10 \text{ d}$ .....	107
Abbildung 4.1: Vorbereitung einer Belichtungsoperation .....	113
Abbildung 4.2: Vorbereitung einer Belichtungsoperation mit Dispatch-Own-Stepper- Dispatcher .....	115
Abbildung 4.3: Datenherkunft .....	116

Abbildung 4.4: Erweiterte Dispatch-GUI (Dispatch Own) mit Detaildialog.....	117
Abbildung 4.5: Vorbereitungszeiten vor und nach der Einführung der Dispatch-GUI mit Dispatch-Own-Funktionalität .....	118
Abbildung 4.6: Veränderung der reinen Wegezeit.....	119
Abbildung 4.7: Veränderung der Losbearbeitungszeit .....	120
Abbildung 4.8: Betriebskennlinie vor und nach Einführung der Dispatch-GUI .....	121
Abbildung 4.9: Arbeitspunkte vor und nach Einführung der Dispatch-GUI.....	121
Abbildung 4.10: Schiebende und ziehende Produktionssteuerung .....	122
Abbildung 4.11: Steuerung der Voroperation durch Dispatch Ahead.....	125
Abbildung 4.12: Modellgeometrie zur Simulation zweier Arbeitsstationen .....	126
Abbildung 4.13: Simulations-Szenario des Dispatch Ahead .....	130





# Tabellenverzeichnis

Tabelle 2.1:	Durchschnittliche Zeit in Monaten, in denen die Bauteilkomplexität verdoppelt wurde, für Prozessoren (MPU) und Speicher (DRAM) .....	8
Tabelle 2.2:	Preis pro Funktion in Bezug auf Strukturgröße, Speicherkapazität und Prozessorgeschwindigkeit [Woj02] .....	10
Tabelle 2.3:	SEMI-Standards zur Optimierung der Anlagenperformance und Produktivität von Halbleiterfabs nach [Piv02] und [Cim03] .....	21
Tabelle 2.4:	Dispatch-Regeln .....	32
Tabelle 3.1:	On-Time-Delivery für $CT^* \sim N(\mu = 28 \text{ d}, \sigma = 5 \text{ d})$ , $CTP = 4 \text{ w}$ mit unterschiedlicher Diskretisierungskonstante .....	63
Tabelle 3.2:	On-Time-Delivery für normalverteilte Durchlaufzeiten mit unterschiedlichem Mittelwert und Standardabweichung bei einer Plandurchlaufzeit von 49 Tagen .....	65
Tabelle 3.3:	Auswirkung auf die <i>OTD</i> bei Veränderung des Mittelwertes der Durchlaufzeit .....	67
Tabelle 3.4:	Auswirkung auf die <i>OTD</i> bei Veränderung der Standardabweichung der Durchlaufzeit .....	68
Tabelle 3.5:	Abschätzung und Berechnung der On-Time-Delivery .....	70
Tabelle 3.6:	Daten zu Abbildung 3.30 .....	84
Tabelle 3.7:	Daten zu Abbildung 3.31 .....	85
Tabelle 3.8:	Beispiele gegenläufiger <i>OTD</i> -Entwicklungen .....	87
Tabelle 3.9:	On-Time-Delivery-Werte für Frontend und Backend, abhängig von der Standardabweichung der Durchlaufzeit und der Differenz von Plandurchlaufzeit .....	92
Tabelle 3.10:	Kombination aus Eingangswerten und On-Time-Deliverys .....	93
Tabelle 3.11:	Beeinflussung der On-Time-Delivery durch die Plandurchlaufzeit Frontend .....	100
Tabelle 3.12:	Bewertung der Durchlaufzeit mit Hilfe der Verteilungsfunktion der Normalverteilung .....	102
Tabelle 3.13:	On-Time-Delivery Backend für $CTb^* \sim N(\mu = 7 \text{ d}, \sigma = 2 \text{ d})$ .....	106

Tabelle 3.14:	OTD-Werte im Modell „On-Time-Delivery ohne Release Backend“ .....	107
Tabelle 3.15:	Verbesserung der Liefertreue bei gleicher Produktionszeit durch Wegfall der Einschleuswoche im Backend für das Beispiel $CTf^* \sim N(\mu = 40 \text{ d}, \sigma = 7 \text{ d}), CTb^* \sim N(\mu = 7 \text{ d}, \sigma = 2 \text{ d})$ .....	108
Tabelle 4.1:	Durchschnittliche Wegezeit pro Los zur Vorbereitung einer Belichtungsoperation.....	114
Tabelle 4.2:	Maschinenauslastungen unterschiedlicher Simulationen ohne Dispatch Ahead.....	129
Tabelle 4.3:	Maschinenauslastungen unterschiedlicher Simulationen mit Dispatch Ahead.....	131
Tabelle 4.4:	Performancesteigerung durch Dispatch Ahead .....	132

# Literaturverzeichnis

- [Ask93] Askin, R.G., Standrige, C.R.: *Modeling and Analysis of Manufacturing Systems*. John Wiley & Sons, Inc., New York, 1993.
- [Ath95] Atherton, L.F., Atherton, R.W.: *Wafer Fabrication - Factory Performance and Analysis*. Kluwer Academic Publishers, Boston, Dordrecht, London, 1995.
- [Aur97] Aurand, S., Miller, P.: *The Operating Curve - A Method to Measure and Benchmark Manufacturing Line Productivity*. In: Proceedings of the IEEE/SEMI Advanced Semiconductor Manufacturing Conference 1997. Cambridge, 10.-12- Sept. 1997, S. 391-397.
- [Bau01] Baur, D., Nagel, W., Berger, O.: *Systematics and Key Performance Indicators to Control a GaAs Volume Production*. In: The International Conference on Compound Semiconductor Manufacturing Technology, GaAs Mantech, Inc. Las Vegas Nevada, 21.-24. Mai 2001.
- [Boe96] Boebel, F.G., Ruelle O.: *Cycle time reduction programm at ACL*. In: Proceedings of the IEEE/SEMI Advanced Semiconductor Manufacturing Conference 1996. Cambridge, 12.-14-Nov. 1996, S. 165-168.
- [CIM03] Cimetrix, Inc.: *Introduction to the SEMI 300mm Standards*. Revision C, August 2003, [www.cimetrix.com](http://www.cimetrix.com).
- [Cun90] Cunningham, J.A.: *The Use and Evaluation of Yield Models in Integrated Circuit Manufacturing*. In: IEEE Transactions on Semiconductor Manufacturing. Volume 3, No. 2, Mai 1990.
- [EIK95] El-Kareh, B., Ghatalia, A., Satya, A.V.S.: *Yield Management in Microelectronic Manufacturing*. In Proceedings of the 45th Electronic Components Conference. San Diego, 1.-4. Juni 1995, S. 58-63.
- [Eng00] Engelhardt, C.: *Betriebskennlinien*. Carl Hanser Verlag, München, Wien, 2000.
- [FHK] Fachhochschule Köln, Institut für Produktion, [www.pt.fh-koeln.de/](http://www.pt.fh-koeln.de/)
- [Gro98] Gross, D., Harris, C.M.: *Fundamentals of queuing theory*. 3rd edition. John Wiley & Sons, Inc., New York, 1998.

- [Han03] Hansch, W., Hilsenbeck, K.: *Halbleiterproduktionstechnik*. Unterlagen zur Vorlesung, Technische Universität München, 2003.
- [Han04] Hansch, W.: *Advanced MOSETs and Novel Devices*. Unterlagen zur Vorlesung, Technische Universität München, 2004.
- [Hei02] Heinlein, F., Leitermann, T.: *Semiconductor-Production*. In: Unterlagen zur 3. Dresdner Sommerschule Mikroelektronik, 2002.
- [Hil05] Hilsenbeck, K., Schömig, A., Hansch, W.: *Zeitdiskrete Modellierung der Wechselwirkungen der Plan-Vorgaben bei Verwendung der Liefertreue als Leistungsgröße für die interne Supply Chain in der Halbleiterindustrie*. In: Operations Research Proceedings 2004. Tilburg, Niederlande, 1.-3. Sept. 2004, Hrsg: Fleuren, H., et al., Springer Verlag, Berlin, 2005, S. 93 - 101.
- [Hop01] Hopp, W.J., Spearman, M.L.: *Factory Physics - Foundations of Manufacturing Management*. Second Edition. Irwin/McGraw-Hill, New York, 2001.
- [Hut05] Hutcheson, G.D.: *The Economic Implications of Moore's Law*. In: Huff, H.R. (Hrsg.), *High Dielectric Constant Materials*. Springer-Verlag, Berlin, 2005.
- [ITRS03a] *International Technology Roadmap for Semiconductors 2003 Edition, Yield Enhancement*, <http://public.itrs.net>
- [ITRS03b] *International Technology Roadmap for Semiconductors 2003 Edition, Executive Summary*, <http://public.itrs.net>
- [Kie03] Kiesel, J.: *Fachwörterbuch Logistik und Supply Management*. 13. Auflage. Publics Corporate Publishing, Erlangen, 2003.
- [Lev04] Levine, B.: *Today's Supply Chain Management Challenge: Keeping the Recovery On Track. Supplying the Stress during the Upturn*. In: *Semiconductor Manufacturing*, Vol. 5, Issue 2. February 2004, S. 39-52.
- [Man02] McManus, T.J.: *Moore's Law & PC Power*. Presentation Slides, Tulane Engineering Forum 2002, New Orleans, 13. Sept. 2002. <http://www.eng.tulane.edu/Tef/tef2002.htm>.
- [Mil04] Miller, J.: *Upholding Moore's Law*. In: Mass High Tech. Boston 07.06.2004.
- [Moo65] Moore, G.E.: *Cramming more components onto integrated circuits*. In: *Electronics*, Volume 38, Number 8. April 19, 1965.
- [Moo75] Moore, G.E.: *Progress in Digital Integrated Electronics*. IEDM, 1975.
- [Now04] Nowack, L.: *Modellierung und Untersuchung eines Dispatch-Ahead-Systems in einer Simulationsumgebung*. Diplomarbeit am Lehrstuhl für Technische Elektronik, Fachgebiet Halbleiterproduktionstechnik, Technische Universität München, München, 2004.

- [Oec02] Oechsner, R., Pfeffer, M., et al.: *From overall equipment efficiency (OEE) to overall Fab effectiveness (OFE)*. In: Materials Science in Semiconductor Processing, Volume 5, Issue 4-5. August - October 2002, S. 333-339.
- [Piv02] Pivin, L.: *SEMI Standards and Equipment Performance Improvements*. In: SEMI Industry News. October 2002.
- [Ros03] Rose, O.: *Comparison of Due-date Oriented Dispatch Rules in Semiconductor Manufacturing*. In: Proceedings of the 2003 Industrial Engineering Research Conference. Portland, OR, 18.-20 Mai 2003.
- [Sar99a] Saroia, M., Halmel, S.: *Production Datawarehouse and Software Toolset to support Productivity Improvement Activities*. In: IEEE/SEMI Advanced Semiconductor Manufacturing Conference Proceedings 1999. Boston, Massachusetts, 8.-10. Sept. 1999, S. 183-187.
- [Sar99b] Saroia, M.: *Synchronizing Productivity Enhancement and Line Control with a new method of target setting for logic plants*. In: IEEE/ISSM International Symposium on Semiconductor Manufacturing Conference Proceedings 1999. Santa Clara, California, 11-13 Oct. 1999, S. 349 - 351.
- [Sal01a] Saller, F.: *Stepper-Dispatcher*. Diplomarbeit am Lehrstuhl für Technische Elektronik, Fachgebiet Halbleiterproduktionstechnik, Technische Universität München, München, 2001.
- [Sal01b] Saller, F.: *SpecialDispatch - Technische Dokumentation*. Infineon-interne Dokumentation des dem SpecialDispatch-System zugrunde liegenden Datenbank- und Software-Systems, München, 2001.
- [Sal01c] Saller, F., Kirf, M., Röhl, I.: *Lastenheft Stepper-Dispatcher*. Infineon-interne Spezifikation der Erweiterung der Dispatch-GUI um SpecialDispatch-Funktionen, München, 2001.
- [Sch99] Schömig, A.K.: *On the Corrupting Influence of Variability in Semiconductor Manufacturing*. Proceedings of the 1999 Winter Simulation Conference. Phoenix, 5.-8. Dez. 1999, S. 837-842.
- [Sch00] Schömig, A.K., Fowler, J.W.: *Modelling Semiconductor Manufacturing Operations*. In: Proceedings of the 9th ASIM Dedicated Conference Simulation in Production und Logistics. Berlin, März 2000, S. 56-64.
- [SEMA] SEMATECH Inc., [www.sematech.org](http://www.sematech.org).
- [SEMI] Semiconductor Equipment and Material International, [www.semi.org](http://www.semi.org).
- [She97] Shea, S.J., LaFreniere, J.A.: *Development and Implementation the Range Managemant System in a Multi-Flow Fabricator*. In: Advanced Semiconductor Manufacturing Conference and Workshop. IEEE/SEMI, Cambridge, 10.-12. Sept. 1997, S. 398-404.
- [Sim03] Simchi-Levi, D., Kaminsky, P., Simchi-Levi, E.: *Designing and Managing the Supply Chain*. McGraw-Hill/Irwin, New York, 2003.

- [Sta02] Stadtler, H., Kilger, C.: *Supply Chain Management and Advanced Planning*. Second Edition. Springer-Verlag, Berlin, Heidelberg, New York, 2002.
- [Tho03] Thonemann, U. et al.: *Supply Chain Champions*. Betriebswirtschaftlicher Verlag Dr. Th. Gabler/GWV Fachverlag GmbH, Wiesbaden 2003.
- [Tra96a] Tran-Gia, P.: *Analytische Leistungsbewertung verteilter Systeme*. Springer-Verlag, Berlin, 1996.
- [Tra96b] Tran-Gia, P., Schömig, A., Dümmler, M.: *Discrete-time Analysis of Batch Servers with Bounded Idle Time*. In: Proceedings of the European Simulation Multiconference 1996. Budapest.
- [Tra04] TransFact: *Lexikon der Produktionskennzahlen und Normen*. [www.transfact.de](http://www.transfact.de)
- [Vad03] Vadas, D.: *Power To Optimize Seminar*. Semiconductor Manufacturing Productivity Seminar Series, Brooks Automation, Inc., 2003.
- [VDI96] VDI (Hrsg.): *Simulation von Logistik-, Materialfluss- und Produktionssystemen - Begriffsdefinitionen*. VDI-Richtlinie 3633, Verein deutscher Ingenieure (VDI), Berlin 1996. a.a.O., VDI-Richtlinie 3633.
- [WGP] Wissenschaftliche Gesellschaft für Produktionstechnik, [www.wgp.de/](http://www.wgp.de/)
- [Woj02] Wojahn, O.: *Sektoranalyse Halbleiter - Stückweiser Aufschwung*. Studie der Barenberg Bank Hamburg, Juli 2002.
- [Zel02] Zell, T.: *Lithography*. In: Unterlagen zur 3. Dresdner Sommerschule Mikroelektronik, 2002.

# Formelzeichen und Abkürzungen

Die Vielfalt der Abkürzungen und Formelzeichen lässt leicht den Überblick verlieren. Um dies in Grenzen zu halten, sind die Indizes im Folgenden meist extra aufgeführt.

Symbol, Abk.	Bedeutung	Einheit	Kap.
*	Kennzeichnung einer Zufallsvariable	-	3
$\alpha$	Variabilität	-	2
$a$	Index für Ankunftsprozesse	-	2
$A$	Availability, Verfügbarkeit (z.B. eines Produktionspartners)	-	2
ANKO	Anlagenkontrolle, Losbezeichnung	-	2
AP	Arbeitspunkt	-	4
$AP$	Index für Arbeitspaket	-	4
ARAMS	Automated Reliability, Availability and Maintainability Standard, SEMI-Standard, siehe Seite 21	-	2
$\beta$	Geschwindigkeitsfaktor	-	2
$b$	Backend, zweite Stufe in der Halbleiterproduktion	-	3
$b$	Index für Bedienprozesse	-	2
BE	Backend, zweite Stufe in der Halbleiterproduktion, auch als Index verwendet	-	3
BKL	Betriebskennlinie	-	2, 4
$bo$	Backend Only	-	3
$c$	Variationskoeffizient	-	2
$C$	Completion, Fertigstellung	d	3
$C^*$	Zufallsvariable Completion (Fertigstellung)	d	3
$Capa$	Kapazität einer Produktionslinie	$h^{-1}, d^{-1}, ..$	2

Symbol, Abk.	Bedeutung	Einheit	Kap.
<i>CLIP</i>	Confirmed Line Item Performance, Performance-Parameter für gelieferte Bestellmengen	-	3
<i>CLIP<sub>BE</sub></i>	CLIP Backend	-	3
<i>CLIP<sub>FE</sub></i>	CLIP Frontend	-	3
<i>CT</i>	Cycle Time, Durchlaufzeit	d, w, ..	2, 3
<i>CT*</i>	Zufallsvariable Cycle Time (Durchlaufzeit)	d, w, ..	3
<i>CTP</i>	Cycle Time Planned, Plandurchlaufzeit	d, w, ..	3
$\Delta t$	Diskretisierungskonstante	d, ..	3
<i>d</i>	Delivery, Auslieferung, 3. Stufe der Produktion	-	3
DB	Die Bank, Chip-Zwischenlager nach dem Frontend	-	3
DB	Datenbank	-	4
DC	Distribution Center, Chiplager nach dem Backend	-	3
<i>DD</i>	Due Date, Stichtag für rechtzeitige Lieferung	d, w, ..	3
<i>DGR</i>	Daily Going Rate, Durchsatz pro Tag	d <sup>-1</sup>	2
<i>die</i>	Index bei Werten, die für einen Chip (Die) gelten		2
DWH	Data Warehouse, Zusammenfassung aller Datenbanken	-	
<i>dynCT</i>	Dynamische Durchlaufzeit	d, w, ..	2
<i>dynDGR</i>	Dynamic Daily Going Rage, Dynamischer Durchsatz pro Tag	d <sup>-1</sup>	2
<i>dynFF</i>	dynamischer Flussfaktor	-	2
E	Erlangverteilung	-	2
<i>E(..)</i>	Erwartungswert von ..	entspr.	3
EDD	Earliest Due Date, Dispatch-Regel, siehe Seite 32	-	2
EPT	Equipment Performance Tracking, SEMI-Standard, siehe Seite 21	-	2
<i>f</i>	Frontend, erste Stufe der Halbleiterproduktion	-	3
<i>f<sub>..</sub>(k)</i>	Wahrscheinlichkeitsfunktion einer diskreten Zufallsvariable	-	3
<i>F<sub>..</sub>(k)</i>	Verteilungsfunktion einer diskreten Zufallsvariable	-	3
<i>f<sub>..</sub>(t)</i>	Dichtefunktion einer kontinuierlichen Zufallsvariable		3
<i>F<sub>..</sub>(t)</i>	Verteilungsfunktion einer kontinuierlichen Zufallsvariable	-	3
<i>F<sub>Z</sub>(z)</i>	Verteilungsfunktion der Standardnormalverteilung	-	3



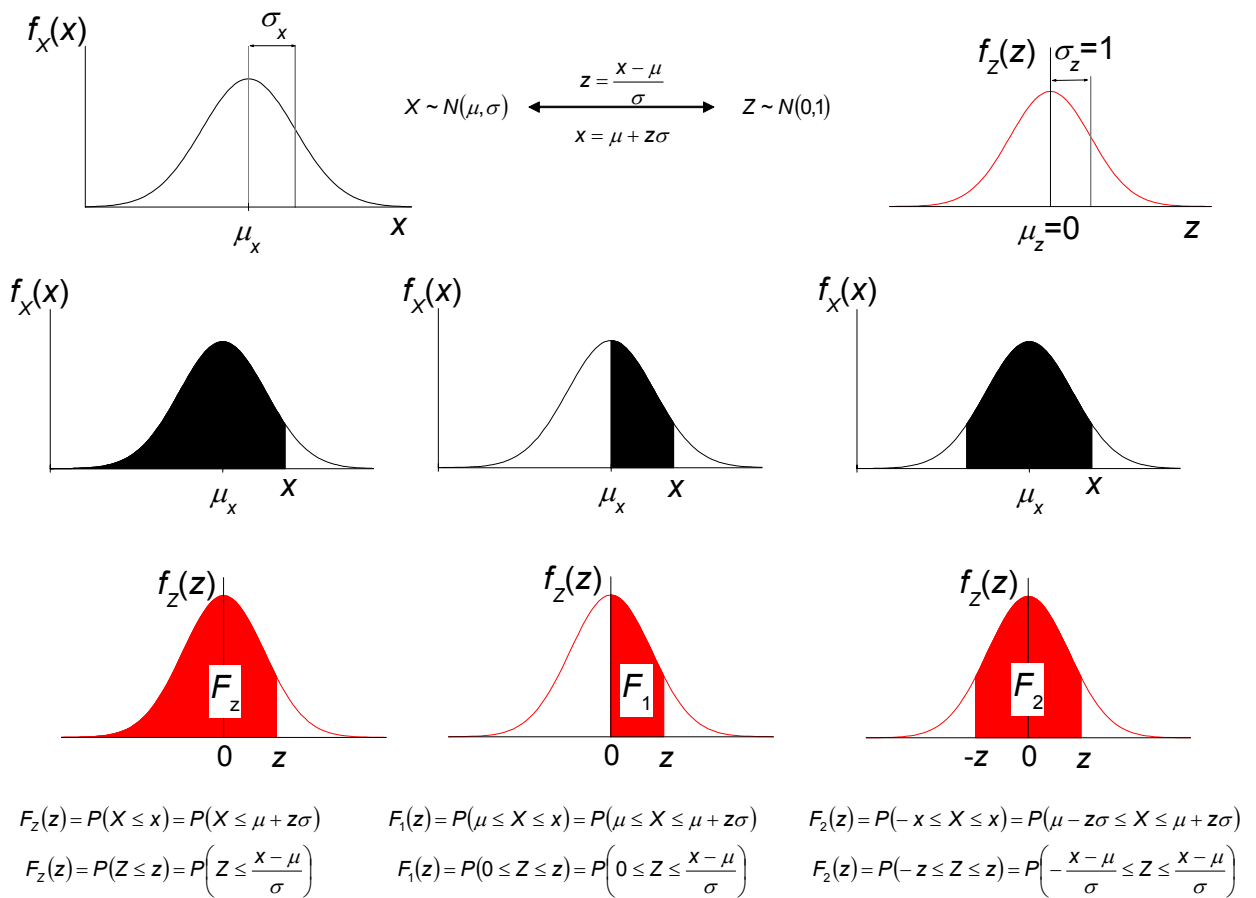
Symbol, Abk.	Bedeutung	Einheit	Kap.
FCFS	First Come First Served, Dispatch-Regel, siehe Seite 32	-	2
FE	Frontend, erste Stufe in der Halbleiterproduktion, auch als Index verwendet	-	3
<i>FF</i>	Flow Factor, Flussfaktor	-	2
FIFO	First In First Out, Bearbeitungsreihenfolge entsprechend der Ankunftsreihenfolge	-	2
$\gamma$	Intensivierungsfaktor	-	2
G	beliebige Verteilung, deren Ereignisse statistisch unabhängig auftreten	-	2
$g(t)$	Dichtefunktion der Gaußschen Normalverteilung	-	3
$G(t)$	Verteilungsfunktion der Gaußschen Normalverteilung	-	3
GR	Going Rate, Durchsatz	$h^{-1}, d^{-1}, ..$	2
GUI	Graphical User Interface, Bedienoberfläche	-	4
GW	Geschäftswoche		2
H	Hyperexponentialverteilung	-	2
$\kappa$	Kompensationsfaktor	-	2
$k$	Diskrete Zeiteinheit	-	3
$L$	Differenz von Planlieferwoche und tatsächlicher Lieferwoche	w	3
$L^*$	Zufallsvariable: Differenz von Planlieferwoche und tatsächlicher Lieferwoche	w	3
$L_p$	Anzahl der Lose für Produkt $p$	-	2
$L_g$	Losgröße	-	3
<i>line</i>	Index bei Werten, die für die ganze Linie gelten	-	2
$lt$	Liefertermin	d, w	3
LTWK	Least Total Work, Dispatch-Regel, siehe Seite 32	-	2
LWKR	Least Work Remaining, Dispatch-Regel, siehe Seite 32	-	2
$\mu$	Mittelwert	entspr.	3
$m$	Index für Maschine $m$	-	2
$m$	Anzahl der Warteplätze in einem Wartesystem	-	2
$m$	Anzahl der Lose eines Produktes innerhalb einer bestimmten Zeitdauer	-	3
$m$	Index für Manufacturing = Frontend + Backend	-	3

Symbol, Abk.	Bedeutung	Einheit	Kap.
M	Exponentialverteilung (Markov-Ereignis)	-	2
MOPNR	Most Operations Remaining, Dispatch-Regel, siehe Seite 32	-	2
MTBA	Mean Time Between Assists, Zeit zwischen zwei Eingriffen	h, w, ..	2
MTBF	Mean Time Between Failures, Zeit zwischen zwei Fehlern	h, w, ..	2
MTOL	Mean Time Off Line, mittlere Ausfallzeit	h, w, ..	2
MTTR	Mean Time To Restore, mittlere Reparaturzeit	min, h, ..	2
MWKR	Most Work Remaining, Dispatch-Regel, siehe Seite 32	-	2
$n$	Anzahl der Bedieneinheiten in einem Wartesystem	-	2
$N$	Normalverteilung	-	3
$NC^*$	Zufallsvariable: Näherung Completion (Fertigstellung)	d	3
$NR^*$	Zufallsvariable: Näherung Release (Einschleusung)	d	3
$o$	Over All = Frontend + Backend + Delivery	-	3
$o$	Index für Operation $o$	-	2
OCM	Operating Curve Management, Betriebskennlinie	-	2, 4
OEE	Overall Equipment Efficiency, Kennzahl für die Produktivität einer Maschine; auch: SEMI-Standard, siehe Seite 21	-	2
OFE	Overall Fab Efficiency, Kennzahl für die Produktivität einer Fab; auch: SEMI-Standard, siehe Seite 21	-	2
$O_p$	Anzahl der benötigten Operationen des Produktes $p$	-	2
Op	Operation	-	2
OTD	On-Time-Delivery, Wahrscheinlichkeit einer rechtzeitigen Lieferung	-	3
$\pi$	Konstante $\pi = 3,1416$	-	3
$\pi_m$	Stützungsoperator	-	3
$p$	Index für Arbeitsplan (Produkt) $p$	-	2
$P(\dots)$	Wahrscheinlichkeit von ...	-	3
$R$	Release, Einschleusung	d, w	3
$R^*$	Zufallsvariable Release (Einschleusung)	d, w	3

Symbol, Abk.	Bedeutung	Einheit	Kap.
RAM	Reliability, Availability and Maintainability, SEMI-Standard, siehe Seite 21	-	2
RANDOM	Random, Dispatch-Regel, siehe Seite 32	-	2
RLZ	Regellieferzeit	d, w, ..	3
RPT	Raw Process Time; Rohprozesszeit, auch physikalische Durchlaufzeit	d, ..	2
$\sigma$	Standardabweichung	entspr.	2, 3
SEMI	Semiconductor Equipment and Materials International, Wirtschaftsverband von Geräteherstellern der Halbleiterindustrie, entwickelt Standards (SEMI E10, ...)	-	2
SL	Service Level, Performance-Parameter für die Liefertreue beim Endkunden	-	3
SollCT	Soll-Cycle-Time, Solldurchlaufzeit	h, d, ..	2
SollDGR	Soll-Daily Going Rate; Soll-Tagesdurchsatz	d <sup>-1</sup>	2
SollFF	Soll-Flussfaktor	-	2
SPT	Shortest Processing Time, Dispatch-Regel, siehe Seite 32	-	2
S/RO	Slack per Remaining Operation, Dispatch-Regel, siehe Seite 32	-	2
$t$	Zeit	d, w, ..	3
$t_a$	Zwischenankunftszeit	min, h, ..	2
$t_b$	Bedienzeit	min, h, ..	2
$T_w$	mittlere Wartezeit	min, h, ..	2
$u, U$	Utilisation, Auslastung	-	2
WIP	Work in Progress, Warenbestand innerhalb der Fertigung	-	2
WIP24	Warenbestand, der innerhalb der nächsten 24 Stunden (vor einer Anlage) erwartet wird	-	2
Y	Yield, Ausbeute in der Chipherstellung	-	3
$Y_F$	Flächenausbeute	-	3
$Y_{SPF}$	Ausbeute im Scheibenprüffeld	-	3
z	Werte der Standardnormalverteilung	-	3



# Tabelle der Standardnormalverteilung





# Danksagung

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftliche Mitarbeiterin am Lehrstuhl für Technische Elektronik, Fachgebiet Halbleiterproduktionstechnik, an der Technischen Universität München in Zusammenarbeit mit der Infineon Technologies AG München. An dieser Stelle möchte ich all denen, die zum Gelingen dieser Arbeit beigetragen haben, meinen aufrichtigen Dank aussprechen.

Mein ganz besonderer Dank gilt Herrn Prof. Dr.-Ing. Walter Hansch. Sein großes Vertrauen machte diese Arbeit erst möglich. Mit Weitsicht und dem Blick für das Essenzielle half er, die richtige Richtung nie aus den Augen zu verlieren. Sein Stil schafft eine Atmosphäre von gegenseitiger Unterstützung und Heiterkeit, die dem Arbeiten eine besondere Qualität verleiht.

Bei Frau Prof. Dr. rer. nat. Doris Schmitt-Landsiedel möchte ich mich für die freundliche Aufnahme am Lehrstuhl für Technische Elektronik und ihr persönliches Wohlwollen bedanken.

Für die Übernahme des Korreferats und die damit verbundenen Mühen danke ich herzlichst Herrn Prof. Dr. rer. nat. Ignaz Eisele von der Universität der Bundeswehr München.

Herrn Dr. Alexander Schömig (Infineon Technologies AG, CL SPC PPC SCM) verdanke ich meine Einsichten in zeitdiskreter Analyse und den Zusammenhängen firmeninterner Lieferketten. Seine Kenntnisse und Überlegungen haben in zahlreichen spannenden Diskussionen den Verlauf der Untersuchungen in hervorragender Weise angestoßen und geprägt.

Herrn Lothar Kratzer (Infineon Technologies AG, AI PL PFM PER M8 PS) danke ich für sein stets unkompliziertes Entgegenkommen. Mit unendlicher Geduld klärte er immer wieder alle produktionstechnischen Fragen und Geheimnisse.

Für die freundliche Offenheit und die ansteckende Begeisterung bedanke ich mich bei Herrn Hans Lutwitz (Infineon Technologies AG, AI PL PFM PER M2), bei Herrn Ignaz Röhl darüber hinaus auch für seine informationstechnischen Unterrichtungen.

Bei meinen Diplomanden Frau Linda Nowack und Herrn Florian Saller bedanke ich mich für ihre besonders engagierte Arbeit.

Allen Mitarbeitern der Infineon Technologies AG möchte ich für die selbstverständliche und freundliche Aufnahme sowie die entgegengebrachte Unterstützung danken.

Den Kollegen am Lehrstuhl sei für die angenehme Zeit gedankt; für die Hilfe bei kleinen und großen Rechnerproblemen Herrn Werner Kraus, Herrn Jürgen Fischer und Herrn Rainer Emling.

Ein letztes großes Danke an alle, die der HLPT-Gruppe ihren besonderen Charme gegeben haben: alle Kollegen (auch aus dem 1. Stock), den vielen Studenten und vor allem Michael Oswald und Christoph Stepper. Der große Zusammenhalt trug einen nicht unwesentlichen Anteil zum Gelingen der Arbeit bei. Doch wenn´s am schönsten ist...