

# **Crystal Ball**

## **Die Gewinnung von verwertbarer Information aus Datenobjekten mit unscharfem Zusammenhang**

**Vom Fachbereich Elektrotechnik und Informatik der  
Universität Siegen**

zur Erlangung des akademischen Grades

**Doktor der Ingenieurwissenschaften  
(Dr.-Ing.)**

genehmigte Dissertation

von

**Dipl. Phys. Ralf Montino**

1. Gutachter: Prof. Dr. Madjid Fathi
  2. Gutachter: Prof. Dr. Rainer Brück
- Vorsitzender: Prof. Dr. Udo Kelter

Tag der mündlichen Prüfung 7. Juli 2009

urn:nbn:de:hbz:467-3965

gedruckt auf alterungsbeständigem holz- und säurefreiem Papier

# Kurzfassung

In den Zeiten der Globalisierung steigen für die Unternehmen nicht nur die Chancen, sondern auch die Herausforderungen. Der ständig vorhandene Wettbewerb fordert von jedem Unternehmen, seine Produkte und Prozesse kontinuierlich zu verbessern. Differenzieren können sich die Anbieter über eine bessere Qualität der Produkte. Neben der Steigerung der Kundenzufriedenheit führt eine bessere Produktqualität durch höhere Planungssicherheit und geringere Ausschüsse zu reduzierten Kosten und steigert damit die Wettbewerbsfähigkeit des Unternehmens.

Insbesondere bei der Herstellung von Produkten für die Automobilzulieferindustrie sind die Qualitätsanforderungen bereits seit langem sehr hoch. Am Beispiel der Halbleiterindustrie wird der Stand der Technik beschrieben. Die heute eingesetzten Werkzeuge und Abläufe dienen nahezu durchgängig der Sicherstellung der hinreichend identischen Reproduktion von einzelnen Prozessschritten, die zuvor als richtig definiert wurden. Insbesondere bei den sehr komplexen Herstellungsprozessen sind die Auswirkungen von Variationen bei einzelnen Prozessschritten auf das Endprodukt häufig nicht klar.

In den vergangenen Jahren haben sich viele Unternehmen unterschiedlichster Größe mit Unterstützung von Hochschulen und Instituten mit der Suche nach Lösungen für die Handhabung der sehr großen vorhan-

denen Datenmengen und Gewinnung von verwertbarer Information aus diesen beschäftigt.

Der Fokus der hier vorliegenden Arbeit liegt auf der Suche nach den Ursachen für Abweichungen in einem hochkomplexen Fertigungsumfeld. Aufgrund der Vielzahl der Einflussfaktoren und dem langen Fertigungsdurchlauf führt eine analytische Lösung nicht zum Ziel.

Abweichend vom in der Literatur beschriebenen Einsatz von Neuronalen Netzwerken zur Klassifizierung oder Prognose des Verhaltens von komplexen Systemen wird hier ein solches Netzwerk eingesetzt, um die Ursachen für das Systemverhalten zu identifizieren. Ebenfalls neu ist die Reduktion der Datenmenge durch den Einsatz der Feature Selection, wobei das Neuronale Netz jeweils als Validierungsinstanz der Feature Selection genutzt wird. Um keine möglichen Lösungen auszuschließen, wird keinerlei Wissen über die technischen Zusammenhänge für diese Verfahren genutzt.

Durch diese Kombination aus datenreduzierendem (Feature Selection) und wissensgenerierendem (Neuronale Netz) Verfahren gelingt es, die für den untersuchten Effekt relevanten Eingangsparameter zu isolieren. Darüber hinaus liefern die Regeln des Neuronalen Netzes Hinweise für die Prozessoptimierung oder Fehlerverhütung.

Dieses neu entwickelte Verfahren wurde zunächst erfolgreich mit synthetischen Daten validiert. Anschließend erfolgte der Einsatz des Systems mit Daten aus der Fertigung. Hier konnten überzeugende Ergebnisse erreicht werden.

In einem aufgezeigten Beispiel wurde für ein Produkt bei mehr als 1000 Datensätzen aus einer Menge von nahezu 100 Parametern eine kleine Anzahl von Parametern identifiziert, die für den untersuchten unerwünschten Effekt verantwortlich waren.

Aktuell befindet sich das System in der Einführung bei einem Halbleiterhersteller. Weitere Unternehmen haben Interesse an dem Verfahren geäußert. Darüber hinaus scheint ein Einsatz des Verfahrens auch in anderen Bereichen der Wirtschaft sinnvoll, insbesondere wenn es sich um Geschäftsprozesse handelt, die sehr hohe Qualitätsanforderungen mit einer guten Dokumentation der durchgeführten Prozesse verbinden.

# Abstract

In times of globalization opportunities but also challenges are rising for the enterprises. The permanent competition requires continuous improvement of products and processes from all companies. The companies can set themselves apart from the market by better quality. Beside the increase in customer satisfaction, improvement of quality leads to reduced costs by higher reliability in planning and reduced scrap. A rising competitiveness of the company is the result.

The production of parts for the automotive industry has been faced with very high quality demands for a long time. Taking the semiconductor industry as example, the state of the art will be described.

Reproducing single process steps –which have been defined as correct before– as identical as necessary is the main target of tools and procedures which are currently used. Especially in very complex production processes, the impact of variations in a single process step for the final product is often not clear.

Many companies –from small ones to the very big ones– made a lot of efforts together with universities and institutes. They searched for solutions to deal with the huge amounts of data and they tried to gain useful information from this data.

The focal point of this work is the retrieval of root causes for deviations in a very complex production environment. Because of the huge amount

of influencing parameters and the long production time, an analytical solution does not lead to the target.

As described in the literature, Neural Networks are often used for classification purposes or for prediction of the behavior of complex systems. In this elaboration, a Neural Network is used to find the root cause for the system behavior. Also new is data reduction by Feature Selection, whereas the Neural Network acts as validation instance for the selected parameters. No knowledge about the technical dependencies is being used in this process. So, no possible solutions are excluded in advance.

With this combination of data reduction (Feature Selection) and knowledge generation (Neural Network), the relevant input parameters for the analyzed effect can be identified successfully. Furthermore, the rules of the Neural Network provide hints for process optimization and error prevention.

At first, this newly developed method has been validated successfully with synthetic data. After that, the system has been used with real production data obtaining convincing results.

Analyzing one problem as an example, more than 1000 datasets from one product were selected. The system identified a small number of parameter responsible for the effect from a bunch of nearly 100.

Actually, the method is being implemented at a semiconductor manufacturer. Other companies show interest in this method.

Beyond that, an implementation of this method seems to be reasonable in other parts of the economy, especially when there are business processes with a combination of high quality requirements and a very good documentation of these processes.

---

## Inhaltsverzeichnis

<b>1 EINLEITUNG .....</b>	<b>1</b>
1.1 GESELLSCHAFTLICHER HINTERGRUND.....	3
<b>2 HALBLEITERINDUSTRIE.....</b>	<b>8</b>
2.1 HERSTELLUNGSPROZESS.....	10
2.1.1 Vielschichtigkeit des Herstellungsprozesses.....	13
2.2 FERTIGUNGSANLAGEN .....	16
2.2.1 Cluster.....	16
2.3 MIXED-SIGNAL-ASIC .....	18
2.4 SYSTEMS IN PACKAGE (SIP) / MULTICHIP MODULES (MCM).....	21
2.5 NANOTECHNOLOGIE .....	23
2.5.1 Leading Edge und Nanotechnologie.....	23
2.5.2 Mixed-Signal und Nanotechnologie.....	26
2.6 MEMS .....	28
2.6.1 Zerstäuber .....	29
2.6.2 Drucksensoren.....	30
2.6.3 Beschleunigungssensoren.....	31
2.6.4 Integration.....	33
2.6.5 Zusammenfassung MEMS.....	34
2.7 SOLARZELLEN.....	35
2.8 FERTIGUNGSLEITSYSTEME .....	36
2.8.1 Funktionalität von Fertigungsleitsystemen.....	39
2.9 AUTOMATISIERUNG .....	45
2.9.1 Automatisierung im Halbleiterumfeld .....	45
2.9.2 Stand der Technik bei der Automatisierung .....	46
2.9.3 Schlüsselfaktoren für die Automatisierung .....	48
<b>3 KUNDENANFORDERUNGEN / QUALITÄT .....</b>	<b>50</b>
3.1 PROCESS CONTROL MONITOR.....	52
3.2 FUNKTIONSTEST .....	55
3.2.1 Bauteilidentifikation.....	56
3.2.2 Guard Banding: Messgenauigkeit.....	58

3.2.3 Wafertest .....	59
3.2.4 Final Part Test.....	60
3.2.5 In-Strip-Test.....	61
3.2.6 Guard Banding: Temperaturgang .....	64
3.2.7 Einordnung Funktionstest in den Gesamtfertigungsprozess .....	65
3.3 TRACEABILITY .....	67
3.4 AUFBEWAHRUNGSFRISTEN .....	68
<b>4 STAND DER TECHNIK .....</b>	<b>70</b>
4.1 „SHORT LOOP“ .....	71
4.1.1 Statistical Process Control.....	71
4.1.2 Statistical Equipment Control.....	72
4.1.3 Fault Detection and Classification.....	74
4.1.4 Advanced Process Control.....	75
4.2 AUTOMATISCH SELEKTIERENDE VERFAHREN .....	77
4.2.1 Statistical Yield Analysis / Statistical Bin Analysis .....	77
4.2.2 Part Average Test.....	79
4.3 OPTISCHE KONTROLLE / BEWERTUNG .....	80
4.4 ENGINEERING DATA ANALYSIS.....	81
<b>5 GRENZEN DER ETABLIERTEN VERFAHREN.....</b>	<b>84</b>
5.1 EINFLÜSSE EQUIPMENT .....	85
5.2 EINFLÜSSE UMGEBUNG.....	87
5.2.1 Luftfeuchte und Temperatur.....	87
5.2.2 Defektdichte .....	87
5.2.3 Luftdruck .....	88
5.3 LANGE DURCHLAUFZEITEN / HOHE KOSTEN.....	89
5.4 UNBEKANNTE ZUSAMMENHÄNGE.....	89
5.4.1 Variabilität des Versuchsaufbaus .....	89
5.4.2 Optimierung .....	90
5.5 GRENZEN DER SIMULATION.....	90
<b>6 AUSWEGE AUS DER KRISE.....</b>	<b>92</b>
6.1 LEBENSZYKLUS VON FERTIGUNGSLEITSYSTEMEN .....	93
6.2 INFORMATIONSGENERIERUNG AUS DATEN.....	95

---

6.3 NUTZUNG DER MASCHINENDATEN .....	97
6.4 EREIGNISVERARBEITUNG.....	100
6.5 STAND DER WISSENSCHAFT .....	102
6.5.1 <i>Massendaten und Neuronale Netze</i> .....	105
<b>7 CRYSTAL BALL .....</b>	<b>108</b>
7.1 MATHEMATISCHE PROBLEMBESCHREIBUNG.....	110
7.2 NEUER ANSATZ: NN ALS VALIDATOR DER FEATURE SELECTION .....	112
7.2.1 <i>Ablauf des Verfahrens</i> .....	115
7.2.2 <i>Systemübersicht: Datenkonzentration</i> .....	119
7.3 BENUTZERSCHNITTSTELLE .....	121
7.4 DATENQUELLE.....	122
7.4.1 <i>Voraussetzung für „Generalisierten NN Ansatz“</i> .....	123
7.4.2 <i>Granularität der Analyse</i> .....	123
7.4.3 <i>Chargensplits</i> .....	126
7.4.4 <i>Koordinatensysteme</i> .....	127
7.5 DATENTYPEN .....	127
7.5.1 <i>PCM Daten</i> .....	128
7.5.2 <i>Maschinendaten</i> .....	128
7.5.3 <i>Messwerte Maschinen</i> .....	128
7.5.4 <i>Messwerte Funktionstest</i> .....	129
7.5.5 <i>Ausbeutedaten</i> .....	130
7.6 VORBEREITUNG DER DATEN .....	131
7.6.1 <i>Selektion</i> .....	131
7.6.2 <i>Normalisierung / Ausreißer-Selektion</i> .....	132
7.6.3 <i>Topologische und chronologische Zuordnung</i> .....	135
7.6.4 <i>Bereinigung</i> .....	138
7.6.5 <i>Abschluss der Vorbereitung</i> .....	139
7.7 FEATURE SELECTION .....	140
7.7.1 <i>Sequential Forward Floating Search</i> .....	142
7.7.2 <i>K Nearest Neighbor als Klassifikator</i> .....	143
7.7.3 <i>Bewertungsfunktion</i> .....	146
7.7.4 <i>Abbruchbedingung</i> .....	147

---

7.7.5 Klasseneinteilung der Eingangsdaten.....	147
7.7.6 Verifikation .....	148
7.8 NEURONALES NETZ .....	150
7.8.1 Fuzzy ARTMAP.....	151
7.8.2 Implementierung .....	154
7.8.3 Parametrisierung .....	156
7.8.4 Verifikation .....	157
<b>8 ERGEBNISSE .....</b>	<b>160</b>
8.1 ÜBERPRÜFUNG / PARAMETRISIERUNG NEURONALES NETZ.....	161
8.1.1 Vigilance.....	162
8.1.2 Learning Rate.....	166
8.2 ÜBERPRÜFUNG FEATURE SELECTION .....	170
8.3 ERGEBNISSE MIT REALEN DATEN.....	174
8.3.1 Datenvorbereitung reale Daten .....	174
8.3.2 Klassifizierung reale Daten .....	175
8.3.3 Bestimmung Vigilance mit realen Daten.....	175
8.3.4 Bestimmung Learning Rate mit realen Daten .....	176
8.3.5 SFFS/KNN und SFAM mit realen Daten .....	178
<b>9 AUSBLICK .....</b>	<b>181</b>
<b>10 ANHANG .....</b>	<b>183</b>
10.1 OTC DATEIFORMAT .....	183
10.2 ÜBERSICHT SOFTWAREWERKZEUGE.....	186
10.3 ABKÜRZUNGSVERZEICHNIS.....	187
10.4 ABBILDUNGSVERZEICHNIS .....	194
10.5 LITERATURVERZEICHNIS .....	197
10.5.1 Eigene Veröffentlichungen bezogen auf die vorliegende Arbeit .....	197
10.5.2 Sonstige Veröffentlichungen.....	202
10.6 DANKSAGUNG.....	220

# 1 Einleitung

Durch die Globalisierung und die stark steigende Transparenz der Märkte und Produkte wachsen neben den Chancen für einen Hersteller auch die Herausforderungen. So kann heute –maßgeblich getrieben durch das Internet– praktisch jeder Anbieter seine Waren und Dienstleistungen weltweit anbieten, ohne im Zielmarkt in irgendeiner Weise präsent zu sein. Auf der anderen Seite kann jeder Kunde einfach und in kurzer Zeit mittels des Internets Preise und Produkteigenschaften vergleichen[ALL02]. Zudem kann er über die vielfältigen Angebote unmittelbar Bewertungen von anderen Kunden erfahren [VOT07]. Einige Anbieter, insbesondere Händler, integrieren die Produktinformation und die Kundenbewertung [AMAZ07].

Diese Möglichkeiten werden von den Kunden in stark steigendem Maßstab genutzt, inzwischen auch und gerade für IT – ferne Produkte wie Reisen und Kraftfahrzeuge [ALL02]. Daher befindet sich praktisch jeder Anbieter in einem dauernden weltweiten Wettbewerb. Daraus ergeben sich erhebliche, neue Anforderungen an die Qualität und Effizienz der Produktentwicklung und Fertigung.

Diese Globalisierung wird aktuell bei dem Geschäft mit Endverbrauchern nur durch die verbliebenen Hürden Sprache, Zoll, Verkaufsberatung und Gewährleistung gebremst.

## 1 Einleitung

---

Gegenüber dem beschriebenen B2C<sup>1</sup> Markt hat die Globalisierung bei B2B<sup>2</sup> Geschäftsbeziehungen wesentlich früher begonnen und ist wesentlich weiter fortgeschritten. Insbesondere bei mittleren und größeren Unternehmen ist die weltweite Beschaffung von Waren und auch Dienstleistungen seit Jahren absoluter Standard. Die für den Endkunden beschriebenen Hemmnisse haben die Unternehmen entweder durch Aufbau geeigneter Ressourcen beseitigt oder sie sind gegenüber dem erzielbaren kommerziellen Vorteil schlicht nachrangig. Natürlich muss jeder Anbieter trotz der gestiegenen Anforderungen an die Flexibilität und des beständigen Preisdruckes Erträge erwirtschaften. Anderenfalls kann kein Unternehmen überleben.

Damit ergibt sich für den Hersteller von Produkten die Anforderung, mindestens vier sich teilweise widersprechende Anforderungen gleichzeitig zu erfüllen (siehe Abbildung 1).

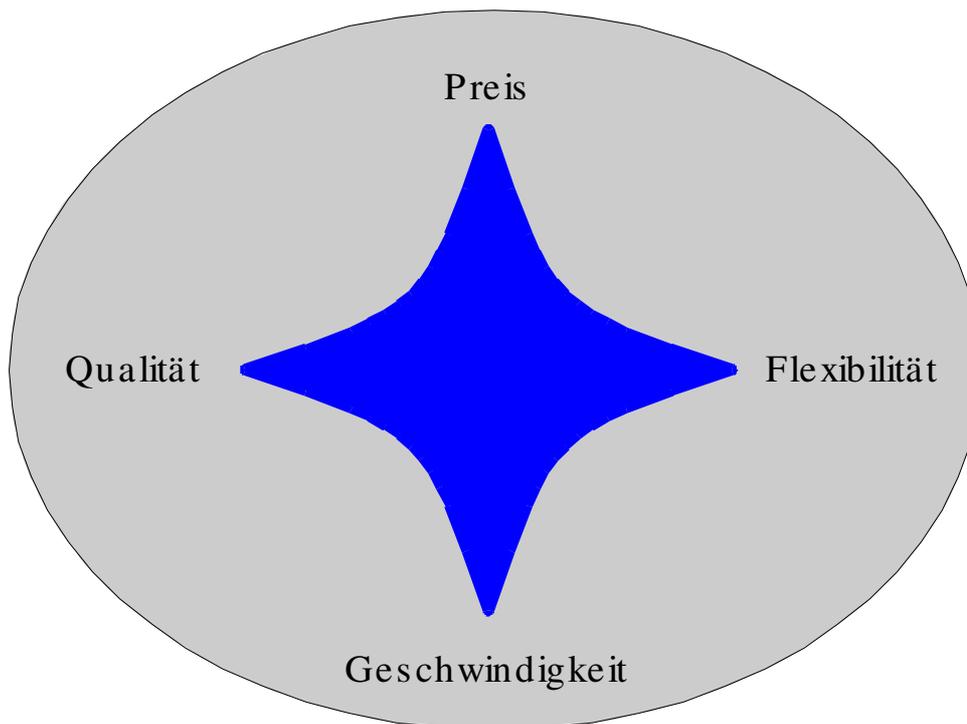


Abbildung 1: Anforderungsviereck modernes Fertigungsunternehmen

Je nach dem Marktsegment, das von dem Unternehmen bedient wird, liegen die Prioritäten natürlich leicht unterschiedlich.

---

<sup>1</sup> Business to Customer: Endkundengeschäft

<sup>2</sup> Business to Business: Kunden- Lieferantenbeziehung zwischen Unternehmen

Die Optimierung auf Preis und Geschwindigkeit haben die Unternehmen bereits lange und intensiv betrieben, so dass hier die Option für weitere große Schritte nicht erkennbar ist. Die Flexibilität der Fertigung ergibt sich aus der Notwendigkeit, wegen der höheren Transparenz der Märkte, wesentlich schneller als in der Vergangenheit auf Änderungen der Kundenanforderungen zu reagieren. Flexibilität muss also als voraussetzende Bedingung für einen Erfolg im Markt betrachtet werden.

Eine relevante Steigerung der Produktqualität führt neben dem indirekten Effekt der Kundenzufriedenheit und der Vermeidung eventueller Regresse durch höhere Planungssicherheit und geringere Ausschüsse zu reduzierten Kosten.

Diese Anforderung ist mit den Methoden nach dem Stand der Technik nur unzureichend umzusetzen. Dies gilt insbesondere für moderne, komplexe Fabriken die sich in einem Marktumfeld bewegen, in dem der Qualitätsanspruch bereits sehr hoch ist.

### ***1.1 Gesellschaftlicher Hintergrund***

Die Steigerung der Produktqualität hat –gerade vor dem Hintergrund der Globalisierung und der gerade aktuellen Krise der Weltwirtschaft– eine entscheidende Bedeutung. Nur die Differenzierung gegenüber dem weltweiten Wettbewerb sichert gerade in der Zulieferkette der Automobilindustrie den Erfolg und das Überleben der hiesigen Standorte und damit auch den Wohlstand und sozialen Frieden in der Gesellschaft.

Gerade unter dem Eindruck der Krise ist es entscheidend, die Qualität der Produkte nicht aus den Augen zu lassen. Schließlich ist die Bereitschaft der Menschen zu konsumieren und zu investieren in einer Krisensituation mit ungewisser Zukunft sowieso nur eingeschränkt vorhanden. Wenn überhaupt wird eher für langlebige Produkte Geld ausgegeben (investiert) als für sehr kurzlebige (Verbrauch).

Die Elektronik im Automobil wird hier an mehreren Stellen entscheidend beeinflusst:

- Positiv wirkt sich aus, dass in Zeiten nachlassenden Wirtschaftswachstums die Automobilhersteller (OEM<sup>3</sup>) typischerweise versuchen, den Verkauf mit besonders gut ausgestatteten Fahrzeugen anzukurbeln. Je höher die Ausstattung der Fahrzeuge, desto höher auch der Anteil an Elektronik im Fahrzeug, so dass Umsatzverluste aus einem nachlassenden Fahrzeugabsatz kompensiert werden können.
- Die weiter ansteigenden Forderungen hinsichtlich Sicherheit und Umweltschutz sind ohne hochentwickelte Elektronikkomponenten nicht erfüllbar. Dieses führt zu einem weiteren Anstieg des Elektronikanteils im Kfz.
- Negativ würde sich ein Trend zu kleineren und preiswerteren Fahrzeugen auswirken, der leicht eine Folge einer Wirtschaftskrise sein kann. Der oben genannte Punkt hinsichtlich Sicherheit und Umweltschutz dürfte den bisherigen niedrigeren Elektronikanteil in preiswerteren Fahrzeugen weiter stark ansteigen lassen und so die Folgen einer Krise mindestens abmildern.

In den letzten etwa 20 Jahren ist der Einsatz von elektronischen Systemen in der Automobilindustrie kontinuierlich angestiegen. Der Halbleitereinsatz legte innerhalb des Elektronikanteils überproportional zu. In Abbildung 2 ist deutlich zu erkennen, dass inzwischen auch der wertmäßige Anteil der elektronischen Komponenten in Fahrzeugen sehr erheblich ist (Daten nach [ZVEI08]). Für die Zukunft wird, trotz oder vielleicht sogar verstärkt durch die aktuelle Krise dieser voraussichtlich noch weiter zunehmen.

Dieser an und für sich erst einmal positive Effekt führt leider auch zu erhöhtem Kostendruck, da die Elektronik aus der Marktnische kommt und zum normalen Zulieferprodukt geworden ist.

Aufgrund des extrem hohen Fixkostenanteils in der Halbleiterfertigung ist es unabdingbar, die Auslastung der Fabriken hoch zu halten. Bei typischen Personalkostenanteilen unterhalb eines Drittels oder gar eines Viertels des Umsatzes ist unmittelbar klar, dass die üblichen Managementmethoden bei nachlassenden Umsätzen (Reduzierung der Perso-

---

<sup>3</sup> Original Equipment Manufacturer: In der Automobilindustrie übliche Bezeichnung für den Fahrzeughersteller.

nalkosten) hier das Problem nicht lösen können. Durch die sehr investitionsintensive Fertigung sind –je nach Unternehmenspolitik– die Abschreibungen oder die langfristigen Verbindlichkeiten aus Leasingverträgen der größte Kostenblock. Hinzu kommen die Energiekosten, die größtenteils durch die notwendige aufwändige Klimatisierung verursacht werden. Die Klimatisierung kann man aber ebenfalls nicht reduzieren oder gar abschalten, da die dann anschließend notwendige Reinigung mehr Ressourcen benötigt, als man vorher eingespart hat. Dieser dauerhafte Zwang zur Volllauslastung der Fabrik und der effizienten Produktion ist der wesentliche Antrieb für die Ausweitung der Informationstechnik in den Unternehmen.

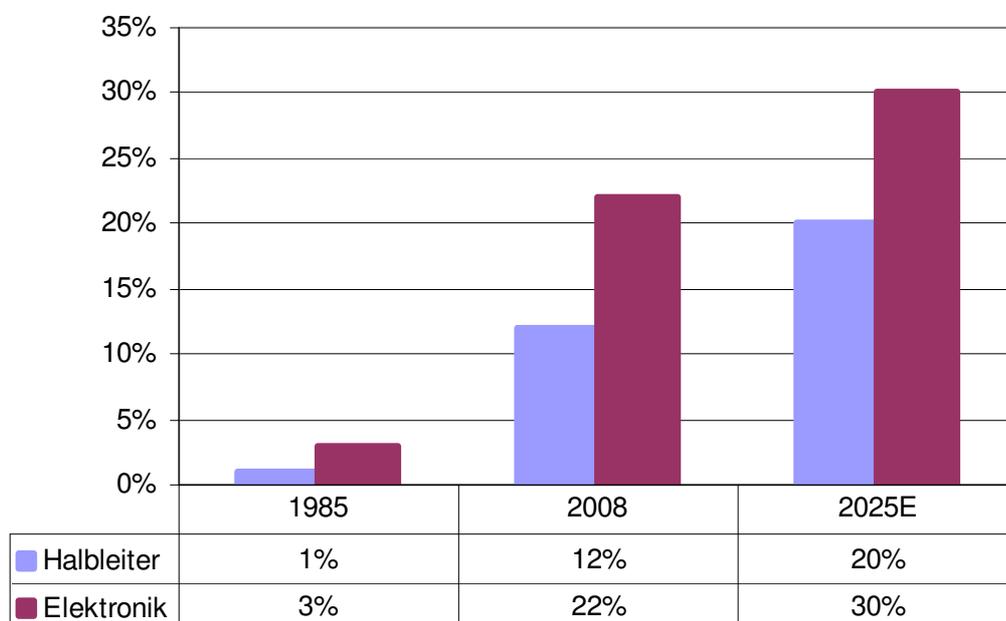


Abbildung 2 Anteil der Kosten elektronischer Systeme an den Produktionskosten eines mittleren PKW, Zahlen nach [ZVEI08]

Diese Arbeit zeigt einen neuen Weg zur Steigerung der Produktqualität gerade bei bereits bestehenden hohen Anforderungen an Qualität, Preis und Flexibilität auf. Als Musterbeispiel wird hier die Halbleiterindustrie betrachtet, insbesondere mittelgroße, flexible Fabriken. Der besondere Fokus auf die Qualität ergibt sich aus der Lieferung in die automobiler Zulieferkette.

Der Bedarf in der Industrie macht es gemeinsam mit der sehr starken Leistungssteigerung bei den Computern sinnvoll und möglich, eine Verbesserung der Analysemöglichkeiten auf Basis einer Kombination aus

Feature Selection und Künstlichen Neuronalen Netzen(ANN)<sup>4</sup> einzusetzen. Die Verfahren wurden vormals aus Komplexitätsgründen als nur eingeschränkt praxistauglich betrachtet. Zudem setzt der erfolgreiche Einsatz hohe Rechenleistungen und eine sehr weit ausgebaute Daten Infrastruktur voraus.

### **Die vorliegende Arbeit gliedert sich wie folgt:**

Beginnend auf Seite 8 gibt Kapitel 2 einen Überblick über die Halbleiterherstellung unter besonderer Betrachtung der Anforderungen an die Informationstechnik. Insbesondere wird auf kleinere bis mittlere Fabriken mit hohen Anforderungen an die Flexibilität der Fertigung eingegangen.

Im folgenden Kapitel 3 wird ab Seite 50 aufgezeigt, welche Kundenanforderungen insbesondere in der automobilen Zulieferkette bestehen und welche unmittelbaren Konsequenzen sich daraus ableiten. Dies betrifft neben der geforderten Produktqualität auch die Anforderungen an die Dokumentation und Rückverfolgbarkeit der Fertigungsprozesse sowie die Aufbewahrungsfristen der erfassten Daten.

Der aktuelle Stand der Technik in den Unternehmen und die Grenzen der etablierten Verfahren sind ab Seite 70 in Kapiteln 4 und 5 aufgeführt. Hier werden anhand von Beispielen Lücken in den etablierten Verfahren aufgezeigt.

Die Grenzen der etablierten Verfahren sind in der Industrie durchaus bekannt. Entsprechend gab es über die letzten Jahre einige Versuche, die damit einhergehenden Probleme zu lösen. In Kapitel 6 werden ab Seite 92 einige Ansätze aufgezeigt, in denen Unternehmen gemeinsam mit Instituten und Universitäten versuchten, einige der bekannten Probleme zu beseitigen.

Kapitel 7 zeigt ab Seite 108 den neuen Ansatz zur Verbesserung der Produktanalyse und der Fehlerverhütung in komplexen Fertigungen auf. Im Einzelnen wird hier der gekoppelte Einsatz von Neuronalen Netzen zur Datenanalyse und vorgeschalteten Systemen zur Feature Selection beschrieben. Neben der mathematischen Beschreibung des Problems

---

<sup>4</sup> ANN: Artificial Neural Network

werden hier neben der Feature Selection und dem Neuronalen Netz die entscheidenden Verbindungen zwischen den Verfahren und zur Datenquelle dargestellt. Schließlich werden in Kapitel 8 die Validierungen des Systems dargestellt. Weiterhin sind hier die Ergebnisse der erfolgreichen Anwendung des Verfahrens in einer industriellen Fertigung dargestellt. Ergänzt wird die Arbeit durch den Ausblick in Kapitel 9.

# 2 Halbleiterindustrie

Die Halbleiterindustrie ist mit einem Beginn vor etwa 40 Jahren ein vergleichsweise junger Industriezweig. Hauptsächlich unterscheiden sich die Produkte in analoge Komponenten wie beispielsweise Operationsverstärker und digitale Komponenten wie Speicher und Mikroprozessoren. Weiterhin gibt es Mischformen von analogen und digitalen Komponenten (Mixed Signal) und in der jüngeren Vergangenheit zunehmende Fertigung von Sensorprodukten auf Halbleiterbasis. Zudem unterscheiden sich die Produktarten in Standardprodukte, die im Wesentlichen unabhängig von konkreten Aufträgen eines einzelnen Kunden entwickelt und gefertigt werden und ASICs<sup>5</sup>, die typischerweise im Kundenauftrag entwickelt und gefertigt werden.

Gemeinsam ist den verschiedenen Produkten die Herstellung auf kreisrunden dünnen Scheiben (Wafer). Diese bestehen –von Nischenanwendungen abgesehen– aus hochreinem, monokristallinen Silizium. Diese Scheiben werden von weltweit sehr wenigen Spezialunternehmen hergestellt.

Die Herstellung der Produkte auf den Siliziumscheiben erfolgt heute praktisch durchgehend in Fertigungslosen a 25 Wafer. Die verwendeten Scheibendurchmesser sind typischerweise innerhalb einer Fabrik ein-

---

<sup>5</sup> Application Specific Integrated Circuit: Applikations- und damit typischerweise kundenspezifische integrierte Schaltung

heitlich. Wegen der geringeren Randeffekte und der effizienteren und damit billigeren Fertigung gibt es einen anhaltenden Trend zu größeren Scheibendurchmessern. Allerdings steigen die sinnvollen minimalen Stückzahlen pro Produkt überproportional stark an und nicht jeder Fertigungsprozess ist sinnvollerweise in einer 300mm Fabrik durchzuführen. Dieses wird im Weiteren noch präzisiert. Abseits von speziellen Prozessen sind heute Fertigungen von 150mm, 200mm und 300mm aktiv. Einen Überblick über die Scheibengrößen gibt Abbildung 3. Für den Neubau von Fabriken spielt der Scheibendurchmesser von 150mm etwa seit dem Beginn dieses Jahrtausends praktisch keine Rolle mehr.

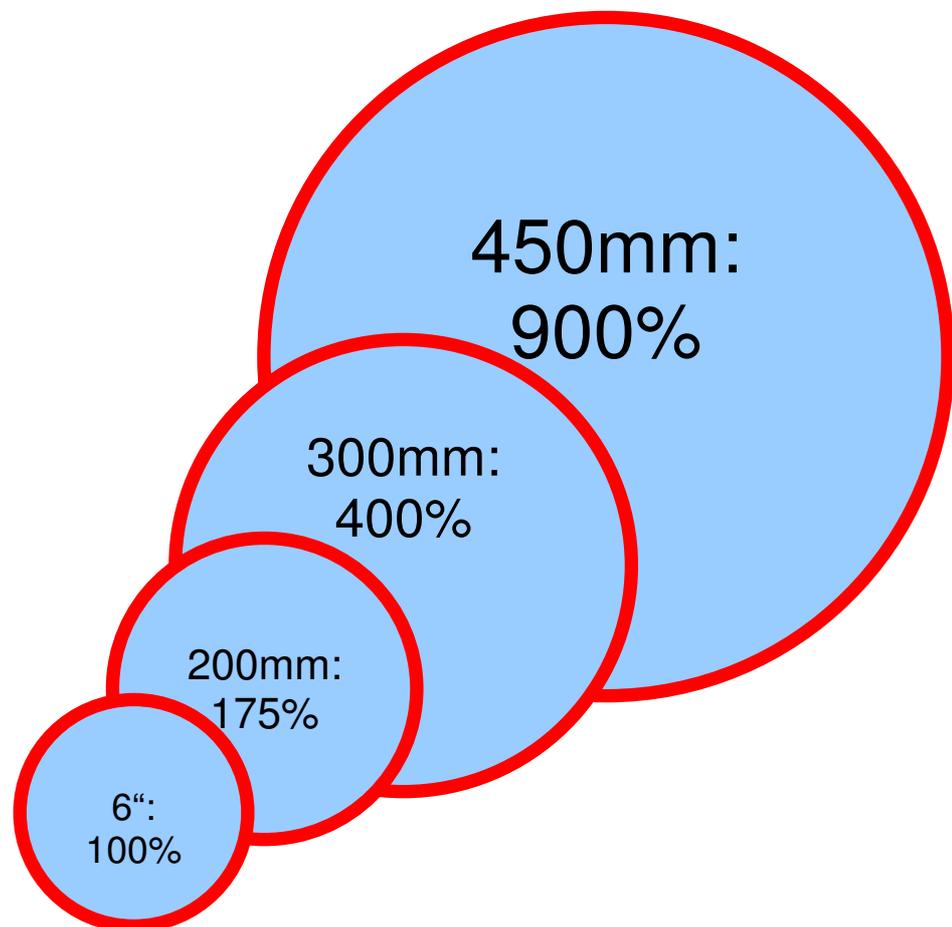


Abbildung 3 Übersicht Scheibendurchmesser

Neben den Anforderungen an die Lieferqualität führt schon der enorme Investitionsaufwand (Tabelle 1) für die Erstellung einer Halbleiterfabrik unmittelbar zu dem Zwang, verlässlich und schnell sehr gute Ausbeuten zu erzeugen. Damit ist es notwendig, zeitnah und zuverlässig möglichst alle relevanten Informationen aus der Fertigungslinie für die Steuerung derselben bereitzustellen.

## 2 Halbleiterindustrie

Scheiben- durchmesser	Investition	Beispiel
200mm	550 Mio Euro	Bosch [BOSCH06]
300mm	3 Mrd. Dollar	Intel [IN05A]
450mm	>>10 Mrd. Dollar (geschätzt in 2005 für Bau in 2012. Schätzung bestätigt in 2008)	Semiconductor International [SIN05, EET07] Samsung, Intel, TSMC [IST08] <sup>6</sup>

Tabelle 1: Investitionskosten für Halbleiterfabriken

Nicht enthalten sind in obiger Kostenschätzung die Kosten für die nachfolgenden Prozessschritte wie Assembly und den Test der Bauelemente. Wie in 3.2.7 ab Seite 65 noch ausgeführt werden wird, sind diese im Wesentlichen proportional zur Anzahl der Bauelemente in einer integrierten Schaltung. Da durch die fortschreitende Reduzierung der Strukturgrößen in der Halbleiterfertigung die Kosten pro Einzelbauelement (Typischerweise pro Transistor) kontinuierlich sinken, verändert sich das Verhältnis der Kosten zwischen Fertigung und Test. Auch wenn die in Tabelle 1 dargestellten Kostenentwicklung auf den ersten Blick etwas anderes erwarten lassen würde sinkt bislang sehr dominierende Anteil der Fertigung an den Gesamtkosten der Herstellung einer integrierten Schaltung, während gleichzeitig die anteiligen Testkosten ansteigen.

### ***2.1 Herstellungsprozess***

Die Herstellung von Halbleiterbauelementen kann schon aufgrund der unvermeidlichen mikroskopischen Defekte durch Partikel generell nicht vollständig fehlerfrei sein. Zwar versuchen alle Hersteller mit enormem

---

<sup>6</sup> Angesichts der Ende 2008 beginnenden Wirtschaftskrise ist die genannte Terminplanung sehr fraglich. Aktuelle Presseveröffentlichungen und Äußerungen in der Industrie deuten auf eine Verschiebung hin.

Aufwand die Anzahl der Partikel in der für die jeweilige Technologie kritischen Größe zu minimieren, offensichtlich ist eine Partikeldichte null aber nicht erreichbar. Damit ergibt sich aus der Defektdichte auf dem Material während der Fertigung ein Ausbeuteverlust, der nur durch entsprechende Prüfverfahren erkannt werden kann. Klarerweise hängt die maximal erzielbare Ausbeute von der Defektdichte und der Chipgröße ab, da sich aus diesen beiden Faktoren die Wahrscheinlichkeit für das Vorliegen eines Defektes auf dem einzelnen Chip errechnet. Für diese Berechnung gibt es verschiedene Modellierungen [MISKO00]. Für kleine Defektdichten und –im Verhältnis zum Scheibendurchmesser– kleine Schaltungen konvergieren die verschiedenen Verfahren.

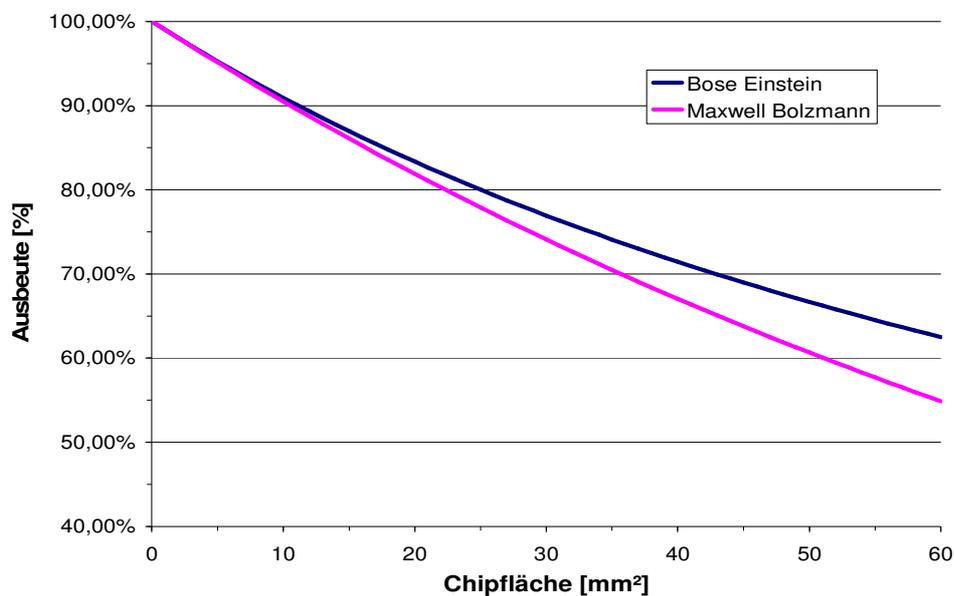


Abbildung 4 Maximalausbeute abhängig von der Chipfläche

In Abbildung 4 ist die theoretische Maximalausbeute abhängig von der Chipfläche dargestellt. Ausgegangen wurde bei der Berechnung von einer Defektdichte von einem Defekt pro Quadratmillimeter (Defektdichte eins). Die Kurve „Bose Einstein“ zeigt den Ausbeuteverlauf unter Zugrundelegung ununterscheidbarer Defekte und damit des Bose-Einstein-Modells. Analog zeigt die Kurve „Maxwell Boltzmann“ den Ausbeuteverlauf bei unterscheidbaren Defekten und damit des Maxwell-Boltzmann-Modells.

Die genauere Betrachtung der Kurve zeigt unmittelbar einige Herausforderungen für Halbleiterunternehmen. Bei steigender Komplexität der Schaltkreise steigt die Fläche des Produktes und damit sinkt direkt die

Ausbeute dieses Produktes in der Fertigung. Eine deutliche Reduzierung der Defektdichte ist in einer existierenden Fertigung normalerweise nur mit großem finanziellem Aufwand möglich, da die Reinraumgüte eine konstruktive Rahmenbedingung für eine Fabrik darstellt und Basis für viele Entscheidungen bei der Planung derselben ist.

Eine Verkleinerung der Strukturbreite führt bei ansonsten identischen Randbedingungen natürlich zu einer Reduktion der Fläche. Leider reduziert sich damit auch die Minimalgröße der kritischen Defekte, was –wiederum bei ansonsten gleichen Randbedingungen– zu einer Erhöhung der Defektdichte führt, schlicht weil mehr Defekte mitgezählt werden.

Die Unterscheidung in unterscheidbare und in nicht unterscheidbare Defekte spielt in der Realität keine Rolle. Für typische Ausbeuten oberhalb von 80% ist offensichtlich die Wahl des Berechnungsverfahrens irrelevant. Für Produkte aus dem Marktsegment Consumer ist das Thema der Ausbeuteverluste rein kommerziell zu entscheiden. Hier ist eine Abwägung des Wertes der nicht nutzbaren Ware gegenüber der (unter Umständen erheblichen) Steigerung der Produktionskosten durch einen Umbau oder Neubau einer Fabrik möglich.

Für Produkte, die in sehr qualitätskritischen Bereichen eingesetzt werden, besteht das Kernproblem vielmehr darin, dass sich ein Partikel auf einem einzelnen Bauteil in einer Maskenebene befindet und dieses keinen Ausfall beim Test erzeugt. Zum Beispiel kann ein Partikel, der vor der Metallisierung auf das Bauteil gelangt, ohne weiteres dazu führen, dass während der anschließenden Prozesse eine Metallbahn lokal wesentlich schmaler wird als geplant. Das sich daraus ergebende Zuverlässigkeitsrisiko lässt sich durch entsprechende optische Kontrolle reduzieren. Allerdings sind diese Verfahren nicht 100% sicher. Weiterhin ist eine vollständige optische Kontrolle der Produkte –gegebenenfalls mehrfach in jeder Belichtungsebene– unter kommerziellen Aspekten nicht durchführbar. Der Preis pro Produkt würde um ein Vielfaches steigen, ohne dass ein Defekt wirklich absolut sicher ausgeschlossen werden kann. Daraus ergibt sich ein unabwendbares Zuverlässigkeitsrisiko.

Die von den Herstellern benutzten Gegenmaßnahmen bestehen –neben angemessenen Fertigungsbedingungen– primär in der Schaffung redundanter Strukturen auf dem Chip und der Abschaltung der nicht funktio-

nierenden Schaltungsteile. Dies geschieht entweder im Funktionstest der Bauteile oder sogar im Feld im Zuge des Selbsttests (BIST<sup>7</sup>) von Bauteilen [LI04,GEHRM06]. Der Erfolg dieser Maßnahmen kann nicht 100%ig sein, da es schaltungstechnisch bedingt immer Single Points of Failure auf dem Bauteil gibt. Zudem eignet sich dieses Verfahren nur gut für Produkte mit vielen identischen Strukturen wie Speicher. Diese wiederum sind vergleichsweise groß, so dass aus der Defektdichte analog Abbildung 4 ohne diese Maßnahmen ein erheblicher Ausbeuteverlust zu erwarten wäre. Daher werden die redundanten Strukturen primär mit dem Ziel der Ausbeuteverbesserung eingesetzt. Die Erhöhung der Zuverlässigkeit beim Kunden ist hier nicht zu erwarten.

### **2.1.1 Vielschichtigkeit des Herstellungsprozesses**

Der Fertigungsprozess besteht aus vielen Hundert Prozessschritten die sich nach 10 bis weit jenseits der 30 Maskenebenen gruppieren. In jedem dieser Prozessschritte kann es zu Fehlern, sowohl technischen als auch menschlichen, kommen. Übliche Beispiele solcher Fehler sind:

- Anlagenfehlfunktion
- Auswahl eines falschen Rezeptes für das gerade aktuelle Material an der jeweiligen Maschine
- Überspringen eines Prozessschrittes
- Doppelte Durchführung eines Prozessschrittes
- Umgebungsbedingungen (Temperatur, Luftdruck, Partikel)

Die Konsequenzen derartiger Fehler sind durchaus vielfältig. Je nach Art des Fehlers und Prozessschritt, in dem der Fehler geschieht, reichen die Konsequenzen von der Nacharbeit über die Verschrottung einzelner Wafer oder ganzer Chargen bis zu verdeckten Produktfehlern die sich schlimmstenfalls als lebensdauerverkürzend darstellen und damit erst beim Endkunden auffallen. Die sich daraus möglicherweise ergebenden Rückrufkosten können für ein Unternehmen durchaus existenzgefähr-

---

<sup>7</sup> Built-In Self-Test: Selbsttest von Bauteilen typischerweise nach Anlegen der Versorgungsspannung, auch in der Applikation.

dend sein. Dies gilt insbesondere deshalb, weil es praktisch nicht möglich ist, sich zuverlässig gegen diese Art des Risikos zu versichern.

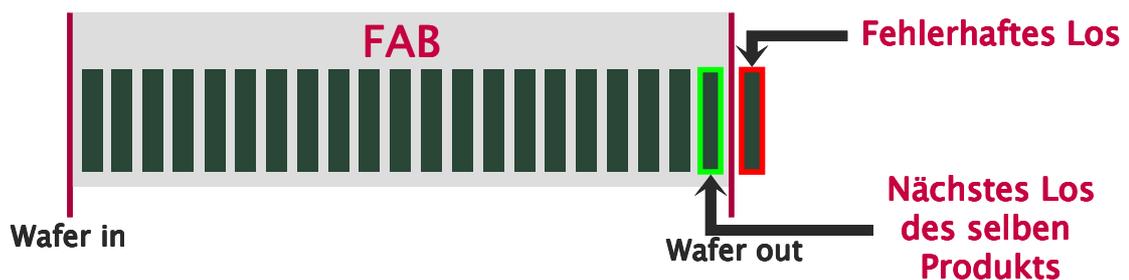
Se- ver- ity	Fehler- art	Entde- ckung	Konsequenz			Beispiel
			Logis- tisch	Kosten	Quali- tät	
1	Rework	Nach Prozess- schritt	Gering	Gering	Keine	Fehlbelich- tung in der Fototechnik
2	Scrap_A	Während oder nach Prozess- schritt	Mittel	Mittel	Keine	Zerbroche- ne Scheiben durch Handling- fehler
3	Scrap_B	Ende der Ferti- gung	Erheb- lich	Erheb- lich	Keine	Rezeptfeh- ler in der Implantati- on
4	Scrap_C	Prüffeld	Erheb- lich	Sehr er- heblich	Mittel	Fehler in Hochtem- peratur- prozess
5	DOA <sup>8</sup>	Beim Kunden	Sehr er- heblich	Extrem	Extrem	Fehler in der Passi- vierung
6	Lifetime	Im Feld	Gering	Existenz gefähr- dend	Extrem	Fehler in Prozessre- zept

Tabelle 2 Strukturierung Fertigungsfehler

<sup>8</sup> Dead on Arrival: Direktausfall bei der Inbetriebnahme

Hinzu kommt noch das Risiko aus dem Produkthaftungsgesetz. Hier kommt erschwerend hinzu, dass sich aus dem Produkthaftungsgesetz eine verschuldensunabhängige Haftung ergeben kann. Eine Fehlervermeidung entsprechend dem Stand der Technik, ist dann nicht ausreichend, um sich vor Forderungen zu schützen. Um die möglichen Fehler und deren Konsequenzen klar zu machen, habe ich in Tabelle 2 eine Klassifizierung der vorkommenden Fehler vorgenommen.

### Massenproduktion: Kein Materialengpass nach Verschrottung



### SMEs (z.B. ASIC Produktion): Materialengpass für Tage oder Wochen

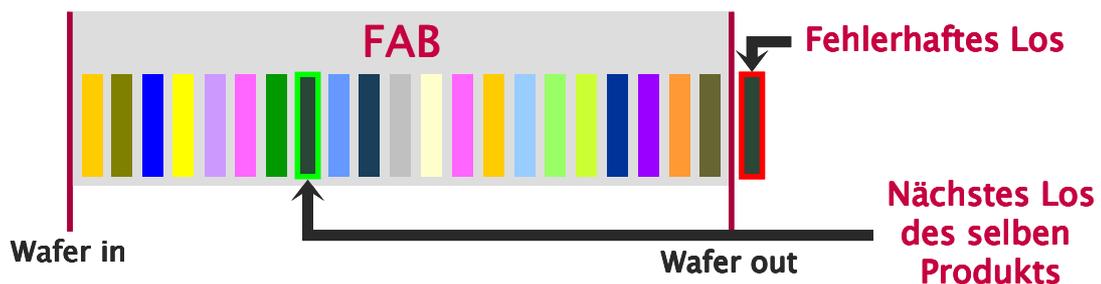


Abbildung 5 Verlust eines Produktionsloses [nach MWK04, AMA05]

Beispielsweise haben erkannte Fehler der Klasse 4 und 5 aus Tabelle 2 in der Fertigung von kundenspezifischen Produkten erhebliche Konsequenzen. Neben dem wirtschaftlichen Verlust durch die notwendige Verschrottung von Ware entsteht das Problem auf der Kundenseite. Da typischerweise nur auftragsbezogen gefertigt wird, steht dann für den Kunden kein Material zur Verfügung (s. Abbildung 5). Insbesondere bei den langen Fertigungsdurchlaufzeiten macht das aufwändige und teure Sonderaktionen notwendig um den Kunden dennoch rechtzeitig zu beliefern.

### 2.2 Fertigungsanlagen

Unter dem Gesichtspunkt der notwendigen Regelung des gesamten Fertigungsprozesses kommt hinzu, dass die einzelnen Fertigungsanlagen bereits hochkomplexe Gebilde mit intrinsischen Regelschleifen darstellen. Diese Regelschleifen optimieren das Verhalten der Anlage, erschweren damit aber die übergeordnete Prozessregelung, da die intrinsischen Regelkreise im Allgemeinen nicht brauchbar beschrieben sind. Schon gar nicht lassen sich diese von außerhalb der Maschine parametrieren oder anderweitig beeinflussen.



Abbildung 6 Canon Stepper mit Lackstraße [OEE02]

#### 2.2.1 Cluster

Moderne Fertigungsanlagen verfügen typischerweise über mehrere Prozesskammern sowie mehr oder minder aufwändige Mechanismen zum Pre- und Postconditioning des Materials. Dies führt dazu, dass die Anlagen neben der bereits beschriebenen Prozessregelung auch noch eine logistische Optimierung durchführen. Damit ändern sich der Maschinenzustand und der Ablauf abhängig von dem aktuellen WIP<sup>9</sup> der entsprechenden Anlage. Diese Abläufe lassen sich ebenfalls normalerweise

---

<sup>9</sup> Work in Process: Bezeichnet den Zustand und die Position des angearbeiteten Materials

nicht beeinflussen, sondern bestenfalls über die Anlagenautomatisierung auslesen.

Üblicherweise werden mehrere Anlagen (Beispiel: Fotocluster aus Lackstraße und Belichter) gekoppelt und kommunizieren dann auch direkt. Für die übergeordnete Fertigungssteuerung können diese gekoppelten Anlagen sinnvollerweise nur noch als Einheit betrachtet werden.

Damit muss sowohl die Fertigungssteuerung als auch die Prozessregelung jedes System als Cluster betrachten. Eine ausführliche Untersuchung zur notwendigen Abbildung der Anlagen und den Konsequenzen aus dieser Abbildung wurde im Rahmen des OEE Förderprojektes durchgeführt [OEE02b]. Hier wurden drei unterschiedliche Clusterarten definiert:

- Lineare Cluster (Beispiel Lackstraße, Stepper<sup>10</sup>, Abbildung 6)
- Heterogene Cluster (Beispiel Mehrkammer-Ätz- und/oder Depositionsanlage, Abbildung 7)
- Virtual Machine

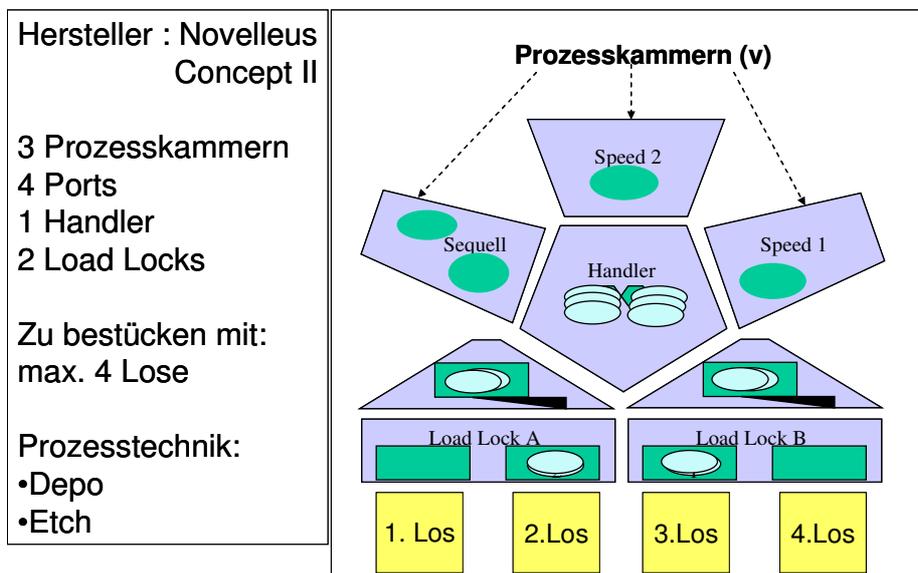


Abbildung 7 Beispiel heterogener Cluster [OEE02b]

Letztendlich ist sowohl das Einzelequipment als auch –am anderen Ende– die gesamte Halbleiterfertigung jeweils nur ein Sonderfall eines all-

<sup>10</sup> Eine Belichtungsanlage in der Halbleiterfertigung belichtet normalerweise nacheinander Teile der Wafer. Daher werden diese Geräte als Stepper bezeichnet.

gemeinen Clusters. Wie schon bei den Einzelmaschinen ist der Ansatz der Prozessregelung hier nur eingeschränkt umsetzbar. Die Fertigungsrezepte werden durch die betreuenden Ingenieure erstellt. Der präzise interne Ablauf ergibt sich dann aus der Kombination der produktspezifischen Rezepte und weiteren Einflussgrößen wie dem Wartungszustand der Maschine sowie den ansonsten im Prozess befindlichen Produkten. Aus Sicht der Prozessregelung sind dies zusätzliche systematische Einflüsse auf den Fertigungsprozess eines Produktes.

### 2.3 Mixed-Signal-ASIC

Halbleiterfabriken, die sich auf kundenspezifische Schaltungen mit analogen und digitalen Schaltungsteilen spezialisiert haben, sind weder beim Scheibendurchmesser noch bei der minimalen Strukturgröße an der Spitze des aktuell Machbaren. So werden aktuelle Mixed-Signal-ASIC-Fabriken typischerweise mit 150mm Wafern betrieben. Einige Hersteller haben bereits auf 200mm Scheibendurchmesser umgestellt oder planen dies.

Die geringen Stückzahlen pro Projekt und die hohe notwendige Flexibilität machen die Nutzung von Foundries<sup>11</sup>, die auf 300mm Scheibendurchmesser produzieren, teuer. Der Bau einer eigenen 300mm Fabrik verbietet sich neben den finanziellen Gründen (siehe Tabelle

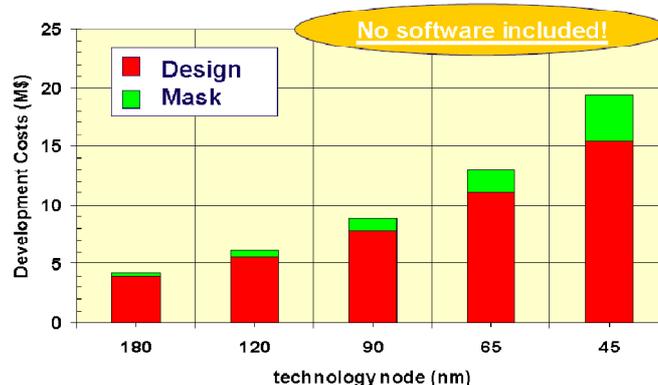


Abbildung 8 Entwicklungskosten für Halbleiterprodukte [FRIES04]

---

<sup>11</sup> Halbleiterhersteller, der kundenentwickelte Produkte mit eigenen Prozessen fertigt. Typischerweise werden die prozessierten Scheiben nach Bewertung der PCM Ergebnisse verkauft. Die Verantwortung für die Funktion der Schaltung trägt der Auftraggeber.

---

1: Investitionskosten für Halbleiterfabriken) schon aus der nicht erreichbaren Stückzahl pro Projekt, für die eine solche Fertigung effizient ist.

Ähnlich ist die Situation bei der minimalen Strukturgröße. Die Kosten für das Equipment, die Werkzeuge (Fotomasken) und die Prozessführung steigt mit der Reduzierung der minimalen Strukturgröße stark an. Abbildung 8 zeigt eine Übersicht über die Aufwände für die Entwicklung eines Produktes in den verschiedenen Technologiegenerationen. Bei großvolumigen Fertigungen von Speicherchips, Prozessoren oder Logikbausteinen wird dieser Effekt durch die mit der Reduzierung der minimalen Strukturgröße quadratisch anwachsende Packungsdichte überkompensiert. In Mixed-Signal-Bauelementen, die zudem noch typischerweise bei höheren Spannungen als die inzwischen üblichen etwa 1 V eingesetzt werden, können nur kleine Anteile der Schaltung in minimaler Strukturbreite ausgeführt werden. Zudem müssen die Anschlußpads der Bauteile für den Bondprozess eine minimale Größe aufweisen. Bei vielen Verbindungen zur Außenwelt sind dann vielfach die Produkte „pad limited“. Das bedeutet, das sich die Größe des Dies aus der benötigten Anzahl der Pads ergibt. Eine Reduzierung der Strukturbreite verkleinert dann das Produkt nicht. Damit ist ein Shrink<sup>12</sup> des Produktes betriebswirtschaftlich nicht mehr interessant.

Entsprechend bieten auch die existierenden Fertigungsprozesse für die kleinsten Strukturen keine Prozessoptionen für hohe Spannungen, hohe Ströme oder ähnliche –dem Umfeld der Analogelektronik entstammenden– Anforderungen.

Für Mixed-Signal-ASICs werden aktuell typisch Strukturbreiten bis hinunter zu 0,25µm eingesetzt. Damit sind die Anforderungen an die Fertigung für derartige Produkte mit Fertigungsmaschinen abzudecken, die vor mehr als 10 Jahren entwickelt wurden. Da der Entwicklungsschwerpunkt der Equipmenthersteller bei der aktuellen Technologie liegt, werden diese Anlagen praktisch nicht weiterentwickelt.

Dieser Umstand ist für alle Anforderungen der Prozessregelung und Maschinensteuerung von entscheidender Bedeutung. Alle neuen Anforde-

---

<sup>12</sup> Die Umstellung eines Produktes auf eine kleinere minimale Strukturbreite ohne funktionale Modifikationen am Produkt wird als Shrink bezeichnet.

rungen, die sich aus dem Marktumfeld und den Kundenwünschen ergeben, müssen mit den existierenden Maschinen erfüllt werden. Insbesondere bei der informationstechnischen Integration der Fertigungen ergibt sich ständig die Notwendigkeit, Datenverarbeitungssysteme aus verschiedenen Dekaden produktionsstauglich miteinander zu verbinden.

Konkret werden für die Steuerung- und Prozessregelung heute genutzter Fertigungsmaschinen Rechner eingesetzt, die vor deutlich mehr als 10 Jahren Stand der Technik waren. Das bedeutet typischerweise den Einsatz von 8 oder 16 Bit-CPU's mit entsprechenden Betriebssystemen. Von den Herstellern der Computer sowie der Betriebssysteme sind keine Ersatzteile und keine Unterstützung mehr verfügbar.

Entsprechend sind die Schwierigkeiten bei der Integration dieser Systeme in eine moderne Produktionsumgebung. Viele der Systeme unterstützen eine Vernetzung gar nicht. Andere, wie z.B. MS-DOS können mit einigem Aufwand in ein TCP/IP<sup>13</sup>-Netzwerk integriert werden. Allerdings kommen die Systeme mit dem hohen Datenverkehr (insbesondere Broadcast<sup>14</sup>) auf dem Netzwerk nicht zurecht und müssen daher in ein eigenes Netzwerk isoliert werden. Weiterhin steht der erfolgreichen Netzwerkintegration der stark limitierte adressierbare Hauptspeicher entgegen. Bei den inzwischen üblichen Speicherausstattungen selbst preiswerter Heimcomputer wird leicht vergessen, dass z.B. MS-DOS als sehr verbreitetes System insgesamt selbst mit speziellen Zugriffstechniken nur 640 KB Hauptspeicher adressieren konnte. Das ist nur etwa 0,5% des heute typischen Wertes. Erschwerend kommt noch hinzu, dass MS-DOS als 16 Bit-Betriebssystem nur kontinuierliche adressierbare Speicherblöcke von maximal 64 KB bereitstellen konnte.

Aktuelle Methoden und insbesondere Werkzeuge für die Softwareentwicklung können mit derartigen Beschränkungen typischerweise nicht einmal näherungsweise umgehen. Damit ist es auch sehr problematisch, Mitarbeiter für diese Tätigkeiten zu finden, die bereits Erfahrung in diesem Umfeld haben. Stattdessen müssen neue Mitarbeiter hier mit hohem Aufwand ausgebildet werden.

---

<sup>13</sup> Transmission Control Protocol/Internet Protocol

<sup>14</sup> Adressierungsverfahren im TCP/IP-Netzwerk bei dem alle Systeme angesprochen werden, die sich im selben Subnetz befinden.

Entsprechend kompliziert sind Softwareänderungen oder gar Neuentwicklungen in diesem Umfeld.

## ***2.4 Systems in Package (SIP<sup>15</sup>) / Multichip Modules (MCM)***

Bei fortschreitender Integration von Funktionen auf einem Chip kommt es unter anderem zu folgenden Konflikten:

1. Die benötigten Funktionen setzen unterschiedliche Technologien und/oder Materialien voraus (Silizium / III/V-Halbleiter).
2. Die Anforderungen können nicht oder nicht sinnvoll durch die Fertigungsmöglichkeiten einer Fabrik erfüllt werden. (Kombination von Analogtechnik mit großen Mengen an Flash-Speicher).
3. Die Anforderungen setzen Produkte von verschiedenen Herstellern voraus.

Klassisch wurde dieses Problem durch die Montage von mindestens zwei Bauteilen auf einer Platine gelöst.

Die Vorteile der Integration in einem Package anstelle der Integration auf der Platine liegen in der höheren Qualität, da

- die Verbindungstechnik besser gegen Umwelteinflüsse geschützt ist.
- die Verbindungstechnik durch die nun geringere Anzahl an externen Verbindungen auf der Platine weniger komplex ist.

Die Nachteile liegen primär in der schlechteren Testbarkeit, da das einzelne Bauteil auf Grund der nicht mehr vorhandenen Zugänglichkeit der internen Verbindungen für Messungen nicht oder nicht vollständig zur Verfügung steht. Zur Verdeutlichung dieses Umstandes ist in Abbildung 9 ein solches MCM dargestellt. Es ist deutlich zu erkennen, dass die beiden Schaltungen direkte Verbindungen aufweisen, die von extern nicht zugänglich sind. Auch für die Analyse im Falle von Kundenreklamationen ist dieses ein zusätzliches Problem, da die Bauteile erst aufwändig

---

<sup>15</sup> Kombination von mehreren Halbleiterprodukten in einem Gehäuse. Häufig werden hier ASICs, Speicher (FLASH) und / oder Sensoren kombiniert.

chemisch geöffnet werden müssen. Zu berücksichtigen ist, dass das Beispiel ein Keramikgehäuse zeigt, bei dem der Deckel entfernt wurde. Keramikgehäuse spielen in der industriellen Anwendung wegen des im Vergleich zu Plastikgehäusen exorbitanten Preises und der vergleichsweise schlechteren Zuverlässigkeit nur dann eine Rolle, wenn die Anforderungen nicht mit einem Kunststoffgehäuse erfüllbar sind (sehr hoher Bedarf der Wärmeableitung, MIL<sup>16</sup> Temperaturbereich bis  $-55^{\circ}\text{C}$ )

Der sicher relevante Punkt des Kostenvergleichs zwischen MCMs und der klassischen Methode ist leider nicht einfach und schon gar nicht allgemein zu entscheiden. Bei der Montage von zwei getrennten Bauteilen fallen höhere Aufwendungen für die Montage auf dem Board sowie höhere Kosten für das Board selber an. Dem ent-

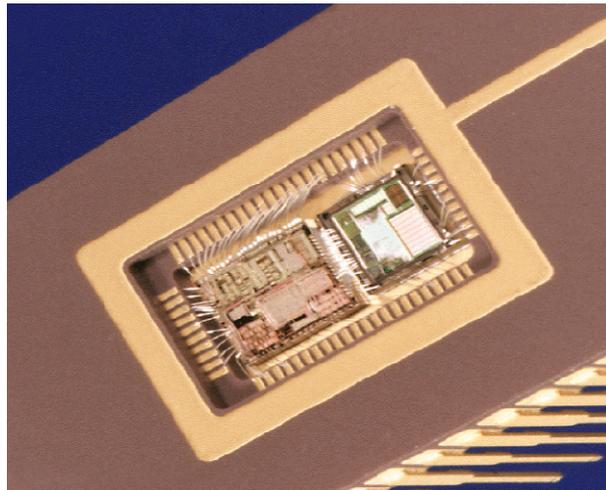


Abbildung 9 Zwei Dies in einem Gehäuse [FRIES04]

gegen stehen höhere Kosten für die Ausbeuteverluste beim abschließenden Funktionstest, da ja bei der Fehlfunktion eines Chips immer das gesamte MCM defekt ist. Weiterhin ist nur ein Assemblyvorgang mit weniger Verbindungen notwendig. Allerdings ist dieses ein Spezialprozess beim Assembler, für den dieser wesentlich höhere Kosten als für ein Standard JEDEC<sup>17</sup>-Gehäuse in Ansatz bringt. Ob ein MCM gegenüber einer kointegrierten Lösung oder einer Mehrchiplösung vorzuziehen ist, muss für jedes geplante Projekt unter Beachtung der genannten Randbedingungen neu bewertet werden.

---

<sup>16</sup> Mit MIL oder ausführlicher MIL-STD werden die US Verteidigungs- oder Militärischen Standards bezeichnet

<sup>17</sup> Ursprünglich Joint Electron Devices Engineering Council, heute Solid State Technology Association [JEDEC08], führender Entwickler von Standards in der Halbleiterindustrie

## ***2.5 Nanotechnologie***

Die Nanotechnologie wird von vielen Seiten als Schlüsseltechnologie des 21. Jahrhunderts bezeichnet. Die am häufigsten aufzufindende Begriffsklärung bezeichnet die Untersuchung und Anwendung von sehr kleinen Strukturen unterhalb von 100 Nanometern [BMU07]. Da der Begriff in der Presse und der Politik große Beachtung findet, wird entsprechend ein sehr großer Bereich von Begriffen abgedeckt. Ob die Halbleiterindustrie der Nanotechnologie zugerechnet werden muss, wird in den Veröffentlichungen durchaus widersprüchlich gesehen. So wird beispielsweise im Aktionsplan Nanoinitiative des BMBF<sup>18</sup> [BMBF06] die Mikroelektronik nicht erwähnt. Auch bei der Diskussion des BMU<sup>19</sup> über die Chancen und Risiken der Nanotechnik findet keine Erwähnung statt [BMU06]. Auf der anderen Seite wird in einer aktuellen Studie für das BMVBS<sup>20</sup> [BMVBS09] die Mikroelektronik in Teilen der Nanoelektronik zugeordnet.

Diese Unterscheidung zwischen Nanotechnologie und Mikroelektronik wird bei oberflächlicher Betrachtung durch die Strukturgröße Mikro, die ja in der Bezeichnung Mikroelektronik enthalten ist, nahe gelegt. Bei der Betrachtung der industriellen Anwendung wird schnell klar, dass dieser Unterschied so nicht existiert.

### **2.5.1 Leading Edge und Nanotechnologie**

Bei der Herstellung von Standardprodukten, insbesondere von Prozessoren und Speichern ist die Leading Edge der aktuellen Fertigungsindustrie bereits bei einer kommunizierten Strukturweite von 45nm angekommen. Allerdings werden die Prozesse –analog der Bezeichnung Mikroelektronik– nach wie vor typischerweise in  $\mu\text{m}$  angegeben. So wird der 45nm Prozess häufig als 0,045 $\mu\text{m}$  Prozess bezeichnet.

---

<sup>18</sup> Bundesministerium für Bildung und Forschung

<sup>19</sup> Bundesministeriums für Umwelt, Naturschutz und Reaktorsicherheit

<sup>20</sup> Bundesministerium für Verkehr, Bau und Stadtentwicklung

Es ist hierbei zu berücksichtigen, dass die von den jeweiligen Unternehmen kommunizierte Zahl nur einen Anhaltspunkt für die produzierten Strukturgrößen darstellt und man bei der Betrachtung, ob es sich um Nanotechnologie handelt, nicht umhin kommt, sich genauer mit dem Thema auseinanderzusetzen. So ist zum Beispiel die minimale Gatelänge beim Intel 65nm Prozess nur 50nm [IN05].

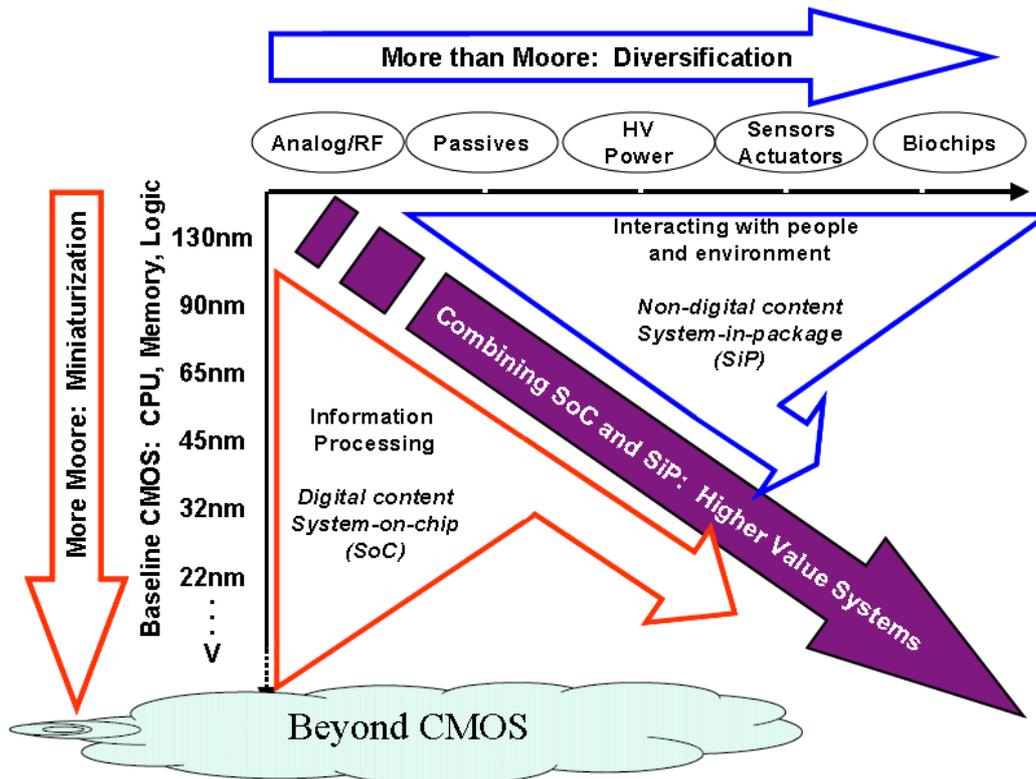


Abbildung 10 More than Moore / More Moore: [ITRS05], S. 9

Die ITRS<sup>21</sup> (z.B. in [ITRS06]) stellt die tatsächlichen physikalischen Größen für die verschiedenen Anforderungen der unterschiedlichen Produkte sehr detailliert dar. In früheren Jahren wurde die aktuelle und zukünftige Leistungsfähigkeit der Fertigungen durch eine einzige Zahl nahezu vollständig beschrieben. Bei dieser Zahl handelte es sich um die minimale Strukturbreite bei der Leitungsführung in einer Speicherzelle (DRAM<sup>22</sup>). Dies war sinnvoll, da die Speicherhersteller bedingt durch die Größe des Marktes und die Härte des Preiskampfes immer an der Spitze

<sup>21</sup> International Technology Roadmap for Semiconductors

<sup>22</sup> Dynamic Random Access Memory: Dynamischer, wahlfrei adressierbarer Speicher

der Prozesstechnologie standen. Zudem waren die Variationen der Anforderung an Bauelemente noch nicht so weit reichend wie heute. Zwischenzeitlich hat sich –abseits der in den Massenmedien veröffentlichten Meinung– die Überzeugung durchgesetzt, dass diese eine Zahl als so genannten Prozesskennziffer nicht mehr ausreicht. Entsprechend sind Kennzahlen für die Gatelängen bei Mikroprozessoren und ASICs sowie eine Kennzahl für die Strukturgröße bei Flash hinzugekommen.

Bedingt durch die Herstellungstechnik unterscheidet sich die gezeichnete Strukturbreite von der sich nachher tatsächlich auf der Schaltung befindlichen durchaus erheblich. So wird die gezeichnete Strukturbreite für Gatelängen bei Mikroprozessoren für 2007 mit 42nm angegeben, was dem kommunizierten 45nm Prozess entspricht. Die tatsächlich entstehende Gatelänge liegt mit 25nm um nahezu 70% unter diesem Wert.

Standort	Partner	Wafergröße	Produktiv
Crolles2, France	ST, Motorola, Philips, (TSMC <sup>23</sup> )	300mm	2003
East Fish- kill, NY, USA	IBM, (AMD, SONY <sup>24</sup> )	300mm	2005
Tokio, Ja- pan	Toshiba, Fujitsu, NEC	300mm	Geplant 2010
Taichung, Taiwan	Elpida, Powerchip Semiconductor Corp.	300mm	2007
SSMC, Sin- gapore <sup>25</sup>	NXP, TSMC	200mm	2000

Tabelle 3 Leading-Edge Kooperationen

<sup>23</sup> TSMC ist nur für die Prozessentwicklung beteiligt, nicht für die Fertigung.

<sup>24</sup> AMD und SONY investieren in die Prozessschrittentwicklung in der IBM Fabrik

<sup>25</sup> SSMC ist ein Joint Venture von NXP B.V. (NXP) und Taiwan Semiconductor Manufacturing Company Ltd (TSMC). Es wurde ursprünglich gegründet von Philips, TSMC und dem Economic Development Board Investments Pte Ltd (EDBI) [SSMC07].

Die weitere Entwicklung der Mikroelektronik wird durch weitere Miniaturisierung in den Fabriken der Großserienhersteller gekennzeichnet sein [s. Abbildung 10]. Durch den bereits beschriebenen enormen Investitionsbedarf wird es wohl zu einer weiteren Konzentration der Hersteller, mindestens aber zu weiterem Ausbau der auch heute schon üblichen Kooperationen beim Betrieb einer aktuellen Großserienfabrik kommen. Beispiele für bereits heute existierende Kooperationen werden in Tabelle 3 dargestellt. Die aufgeführten Fabriken werden aktuell weiter aufgerüstet.

### 2.5.2 Mixed-Signal und Nanotechnologie

Für Produkte, die neben den digitalen Schaltungsblöcken auch analoge Komponenten beinhalten, ist diese primäre Fokussierung auf die Miniaturisierung und den Scheibendurchmesser nicht wirtschaftlich. Wie bereits beschrieben, steigen der Investitionsaufwand und damit auch die Prozesskosten mit der Miniaturisierung stark an. Bei reinen Digitalschaltungen mit hohen Anforderungen an die Verarbeitungsgeschwindigkeit und sehr starken Einschränkungen hinsichtlich des Umgebungstemperaturbereiches und der Versorgungsspannungen werden diese Mehraufwände durch die Flächensparnis und die mit der Miniaturisierung einhergehende Steigerung der Verarbeitungsgeschwindigkeit überkompensiert.

Für Anwendungen in rauen Umgebungen (z.B. Kraftfahrzeuge, Industrie) sind die genannten Einschränkungen bei der Versorgungsspannung und dem Temperaturbereich nicht akzeptabel.

Die notwendige Verwendung von Analogkomponenten und höherer Versorgungsspannung macht klarerweise die Verwendung von extrem geringen Strukturbreiten unmöglich. Damit können –wie bereits dargestellt– bestenfalls kleinere Teile der Schaltungen von der möglichen hohen Integration profitieren, während große Flächen der Schaltung die aufwändigen Prozessschritte bekommen, ohne den Vorteil davon nutzen zu können. Hinzu kommt, dass die initialen Werkzeugkosten (Maskensätze) für die genannten kleinen Strukturbreiten sehr teuer sind und sich nur bei sehr großen Stückzahlen rechnen.

Zusammenfassend wird sehr wahrscheinlich auch auf lange Sicht ein Markt für 200mm-Fabriken mit minimalen turbreiten deutlich oberhalb von 100 nm ( $0,1\mu\text{m}$ ) existieren. Typische Werte heutiger Mixed-Signal Fabriken sind 150mm oder 200mm Scheibendurchmesser und minimale Strukturbreiten zwischen  $0,25\mu\text{m}$  und  $0,6\mu\text{m}$ .

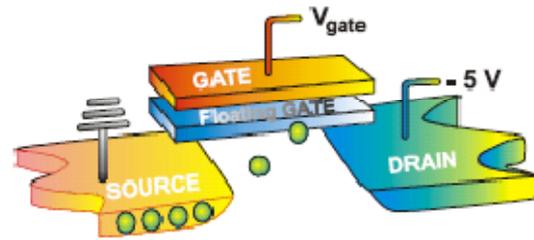


Abbildung 11 Schematische Darstellung einer Flash Zelle [PAUL03]

Doch auch bei diesen Applikationen handelt es sich ohne Zweifel um Nanotechnologie. Betrachtet man die Schichtdicken, so ist bereits ein typisches Gateoxid in einer Digitalschaltung deutlich unterhalb von 50nm. Das bedeutet, hier muss die gesamte Struktur der Nanotechnologie zugeordnet werden.

Darüber hinaus werden heute in Mixed-Signal-Bauteilen häufig nichtflüchtige Speicher eingesetzt. Bei den hier eingesetzten Technologien wie EEPROM<sup>26</sup> oder FLASH<sup>27</sup> wird der quantenmechanische Tunneleffekt für die Programmierung der Speicher ausgenutzt [SCHU99]. Dieser führt nur bei sehr geringen Abständen zu makroskopischen Effekten. Der Tunnelstrom ist exponentiell von diesem Abstand abhängig. Damit muss allein schon für dieses Bauelement eine Schicht von typisch 10nm aufgebracht werden. Die maximal zulässigen Dickschwankungen betragen wenige Atomlagen. Bei der in Abbildung 11 gezeigten schematischen Darstellung einer solchen Speicherzelle befindet sich dieses Tunneloxid zwischen Source und dem Floating Gate. Durch Anlegen einer entsprechenden Programmiervspannung (typisch 10–20 V) an  $V_{\text{gate}}$  wird ein elektrisches Feld erzeugt. Damit tunneln Elektronen aus der Source

<sup>26</sup> Electrically Erasable Programmable Read Only Memory: Halbleiterspeicher der nach der Programmierung die Daten auch ohne Versorgungsspannung hält (PROM). Durch geeignete Verfahren ist diese Programmierung elektrisch lösbar.

<sup>27</sup> Eigentlich vollständig FLASH – EEPROM. FLASH ist eine Weiterentwicklung der EEPROMS. FLASH – EEPROMS lassen sich Byteweise beschreiben, allerdings nur blockweise löschen.

auf das floating Gate und verändern so dauerhaft den Schaltzustand des Transistors. Dieser dünne Isolator (Tunneloxid) muss für die geforderte Zuverlässigkeit über eine sehr homogene Kristallstruktur verfügen.

Dies ist wiederum äußerst herausfordernd, sobald derartige Speicher in qualitativ anspruchsvollen Umgebungen eingesetzt werden. Gerade bei Flash-Speichern für das Consumer-Umfeld (Stichwort: USB-Stick) hat in der jüngeren Vergangenheit der Preiskampf der Anbieter zu einer dramatischen Preisreduzierung geführt. Hohe Anforderungen an die Zuverlässigkeit und die Fehlerfreiheit sind da natürlich nicht zu erfüllen. Flash-Speicher für hohe Qualitätsanforderungen sind dem gegenüber dramatisch aufwändiger und teurer. So werden bei Consumer Flash-Bausteinen typischerweise über verschiedene Spannungswerte in der Speicherstelle mehrere Bit pro Speicherstelle gespeichert, um Fläche zu sparen und damit preiswerter fertigen zu können. Für Automotive Flash hingegen werden zusätzliche Speicherstellen investiert, um Fehler erkennen oder sogar korrigieren zu können. Hier werden die schon aus der DRAM-Nutzung bekannten Verfahren (Parity, ECC<sup>28</sup>) eingesetzt. Zudem ist die künstliche Voralterung (Burn-In) auch für die Großserie Standard. Faktisch handelt es sich bei Consumer Flash und Automotive Flash um vollkommen getrennte Produktlinien, zwischen denen es keinen Austausch gibt.

Die Hersteller, die über diese Technologie verfügen, setzen dieses dünne Oxid typischerweise auch als Gateoxid in den digitalen Schaltungsteilen ein. Tatsächlich findet heute Mikroelektronik ohne Nanotechnologie gar nicht statt.

### ***2.6 MEMS<sup>29</sup>***

Neben der Fertigung von analogen und/oder digitalen Schaltungen werden ähnliche Prozesse auf Siliziumscheiben auch für die Fertigung von

---

<sup>28</sup> Error Correcting Code: Verfahren zur Speicherung von Daten derart, dass Speicherfehler korrigiert oder mindestens erkannt werden können.

<sup>29</sup> Micro-Electro-Mechanical Systems: Mikromechanische Produkte wie Druck- oder Beschleunigungssensoren.

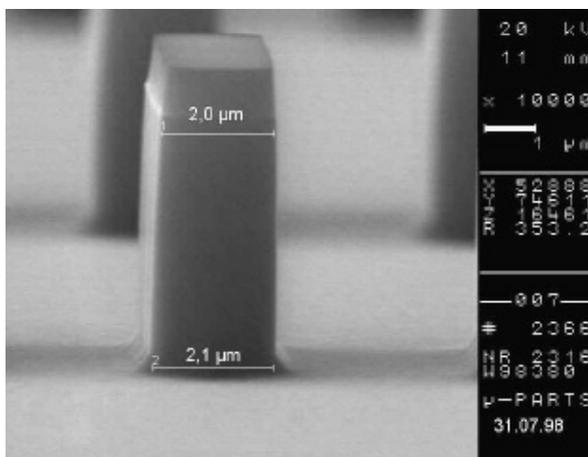
mikromechanischen Produkten eingesetzt. Dieser –auch im Vergleich zur Halbleiterfertigung– junge Industriezweig nutzt Prozesse und Fertigungsanlagen aus der Halbleiterindustrie. Hinzu kommen spezielle Rezepte und Fertigungsverfahren sowie einige spezielle Anlagen.

Bei rein mechanischen Produkten wird im mikroskopischen Maßstab konstruiert. Daneben existieren auch verschiedenste Mischformen, die mit der Aufbringung von Leiterbahnen und Widerständen bei Drucksensoren beginnen und bei kointegrierten Produkten enden. Bei diesen kointegrierten Produkten durchlaufen die Siliziumscheiben neben den für die Mikromechanik notwendigen Lithografie- und Ätzprozessen auch noch einen vollständigen Halbleiterfertigungsprozess, so dass sich schließlich auf dem fertigen Silizium neben einem Sensor auch noch die vollständige Auswerteschaltung befinden kann.

Mikromechanische Produkte oder Produkte der Mikrosystemtechnik bilden ein sehr weitgehendes Spektrum ab und weisen Wachstumsraten deutlich oberhalb der in der Vergangenheit schnell wachsenden Halbleiterindustrie auf. Als Beispiele seien hier genannt:

### 2.6.1 Zerstäuber

In der Medizintechnik ist für die Dosierung von Medikamenten eine möglichst feine Zerstäubung des Medikamentes anzustreben. Hier werden inzwischen mikromechanisch hergestellte Zerstäuber eingesetzt [s. Abbildung 12]. Diese erlauben es, durch die feinere Zerstäubung des Medikamentes die Dosis herabzusetzen [MIPA07].



chip size 2 x 3 mm  
 feature size minimum 2 μm  
 etch vertical sidewalls  
 smooth surfaces  
 no residues  
 mask thermal oxide  
 wafer 150 mm diameter  
 etch depth 5 μm ± 0.15

Abbildung 12 Nozzle Etch process specifications [MIPA01]

Das hier beispielhaft genannte Produkt funktioniert rein mechanisch. Entsprechend ist der Herstellungsprozess auf die Strukturierung des Basismaterials (Silizium) beschränkt. Allerdings sind die Anforderungen an diese Strukturierung sehr hoch, was wiederum erhebliche Anforderungen an die Prozessführung und Prozessregelung stellt. Für eine zuverlässige und effiziente Prozessregelung wird hier mittels einer Transformation in den Phasenraum in Kombination mit einer Regressionsrechnung der optimale Ätzprozess gesteuert [MIPA01].

### 2.6.2 Drucksensoren

Mittels spezieller Prozesse wird durch eine Ätzung von der Rückseite einer Siliziumscheibe eine dünne Membran hergestellt. Diese wird mit als Dehnungsmessstreifen ausgebildeten Widerständen belegt. Je nach Ausführung des Gehäuses kann dieser Sensor dann als Differenz- oder Absolutdrucksensor betrieben werden. Diese Produkte sind sowohl als

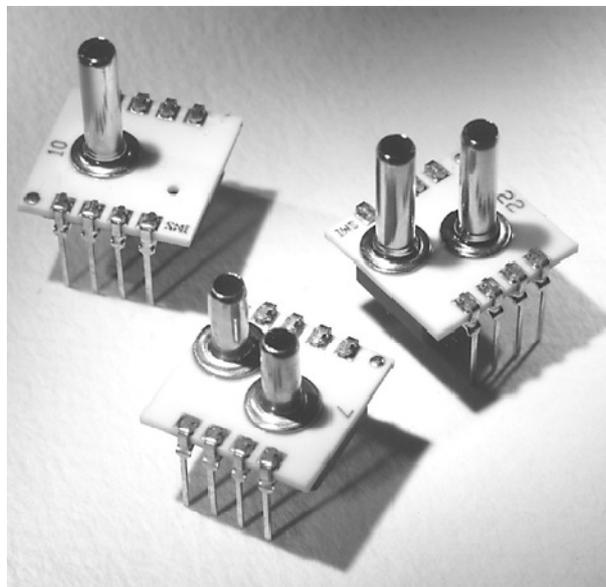


Abbildung 13 SM5651/SM5652, SMI

Standard- als auch als kundenspezifische Produkte von verschiedenen Herstellern verfügbar. Anwendungsgebiete sind hier neben der vergleichsweise unkritischen Unterhaltungsindustrie auch Anwender aus der automobilen Zulieferkette sowie dem Healthcare Bereich. Die Anforderungen sind hier deutlich höher und die Zahl der möglichen Anbieter ist entsprechend kleiner. An dem gezeigten Beispiel eines Niederdruck-sensors (Abbildung 13, [SMI07]) ist direkt erkennbar, dass neben der Fertigung auch die Gehäusetechnologie produktspezifisch ist. Die Erreichung der –für die Zielmärkte notwendigen– Effizienz und das Qualitätslevel sind hier durch die speziellen Gehäuse aus mehreren Gründen erschwert:

- Die speziellen Gehäuse lassen sich mit standardisierten automatischen Handlingsystemen nicht transportieren. Damit ist der funktionale Test der gehäuseten Bauelemente wegen der manuellen Schritte aufwändiger, teurer und unsicherer.
- Die Gehäuse selber müssen ebenfalls hohe Qualitätsanforderungen erfüllen. Dazu gehören mechanische Toleranzen, Robustheit gegenüber Umgebungseinflüssen und Dichtigkeit. Diese Anforderungen sind für Spezialgehäuse natürlich wesentlich schwieriger zu erfüllen als für Standardgehäuse.

### 2.6.3 Beschleunigungssensoren

Von der Fertigungstechnik den Drucksensoren verwandt sind die Beschleunigungssensoren. Beim Drucksensor muss für eine angemessene Funktion die Masse der Membran vernachlässigbar sein. Beim Beschleunigungssensor wird nun gerade auf dieser Membran eine definierte Masse erzeugt. Zudem wird die Membran geeignet durchbrochen, so dass die Beschleunigung der Masse zu einer Durchbiegung des

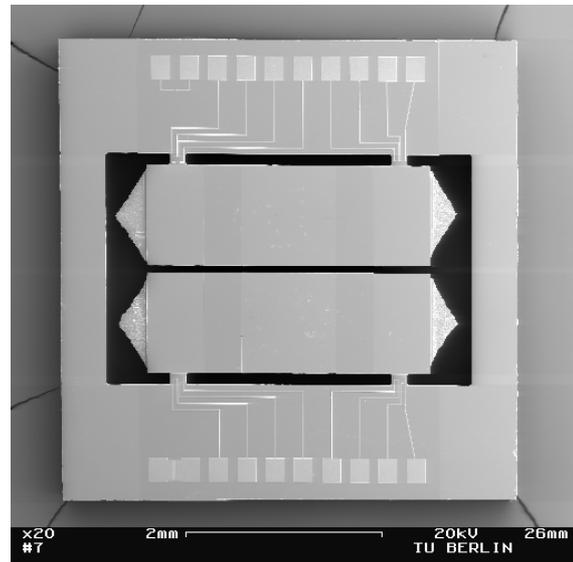


Abbildung 14 2D-Sensorchip [VELT00]

Steges führt. Über die analog zu den Drucksensoren aufgebrachtten Widerstände ist diese Biegung mess- und auswertbar. Abbildung 14 zeigt die Struktur eines solchen Sensors. Es handelt sich hier um einen reinen Sensor, ohne die ebenfalls mögliche Integration der Auswertelektronik.

Beschleunigungssensoren werden ebenfalls in weiten Bereichen der Industrie eingesetzt. Neben dem Einsatz in ESP Systemen in Fahrzeugen erstrecken sich die Einsatzgebiete von der Funktion als Schalter (Beispiel: Reifendrucksensor beim Kfz, Zentrifugalkraftsensor als Ein-/Ausschalter für die Auswerte- und Sendeinheit zur Verlängerung der

Batterielebensdauer im Reifen) bei der Bewegungserkennung bis zur präzisen Beschleunigungsmessung in der Robotik.

Verpackt werden diese Sensoren in Spezialgehäuse, da die Sensoren selbst natürlich keinen mechanischen Kontakt zum Gehäuse haben dürfen. Verglichen mit Drucksensoren ist die Frage des Gehäuses trotzdem weniger kritisch, da Gehäuse geschlossen sind. Dafür ist der Test der Sensoren sehr problematisch. Das automatische Handlingsystem muss das sehr kleine, leichte und vor allem sehr empfindliche Bauteil sicher vereinzeln und zuverlässig kontaktieren, ohne beispielsweise die Anschlussbeine zu verbiegen oder auch nur zu verkratzen.



Abbildung 15 Handlingsystem Beschleunigungsmessung [MUL08]

Diese hochkomplexe Kontaktiereinheit muss nun definiert beschleunigt werden, da das Bauteil während der Messung mit einer wohldefinierten Beschleunigung beaufschlagt werden muss. Weiterhin muss diese Messung zwischen Testprogramm, Testsystem und Handlingsystem sehr präzise zeitlich koordiniert werden. Die normale Handlerkommunikation kommt ohne zeitliches Limit seitens des Handlingsystems aus. Der Handler meldet dem Testsystem die erfolgreiche Kontaktierung des Bauteils, und die Testmaschine meldet dem Handler das Prüfergebnis nach Abschluss der elektrischen Messungen, worauf der Handler das

Bauteil in das entsprechende Ausgabemagazin sortiert. Wie lange die Messung dauert, ist im Normalfall für die Kombination Handler / Tester nicht von Belang. Das Aufbringen der definierten Beschleunigung wird durch eine Erweiterung des Handlingsystems durchgeführt (siehe Abbildung 15). Klarerweise ist der Zeitraum, in dem eine definierte Beschleunigung innerhalb bekannter Toleranzgrenzen aufrechterhalten werden kann, eng eingegrenzt.

### 2.6.4 Integration

Wie bereits beschrieben, wird an der Spitze der Technologie heute die Integration der Sensorfertigung und der Halbleiterfertigung betrieben. Dies geschieht entweder wie beschrieben in Multichip-Modulen (siehe 2.4), oder es wird die Vollintegration auf einem Stück Silizium betrieben. Gegenüber den Multichipmodulen sind bei der Vollintegration der Aufwand und das erhöhte Qualitätsrisiko durch das Multichipassembly nicht vorhanden. Dafür ist der Prozessaufwand stark erhöht und die Sensorfläche ist nicht anderweitig nutzbar. Damit wird die für Schaltungen nutzbare Fläche auf der Scheibe reduziert. Zudem sinkt die Ausbeute, da sowohl Yieldverluste im Sensorprozess als auch Yieldverluste im CMOS-Prozess zur Fehlfunktion des fertigen Produktes führen. Welche Lösung für die jeweilige Anwendung die geeignete ist, hängt von der Einzelfallprüfung ab.

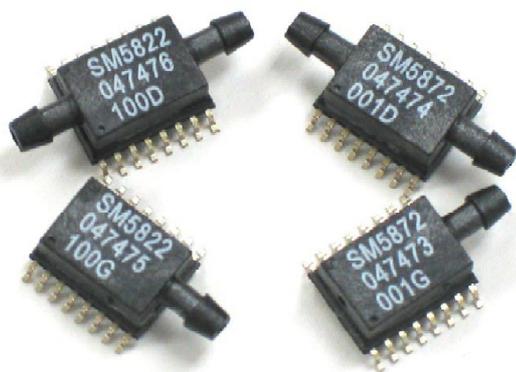


Abbildung 16 SM5822/SM5872,  
Kointegrierter Drucksensor [SMI07]

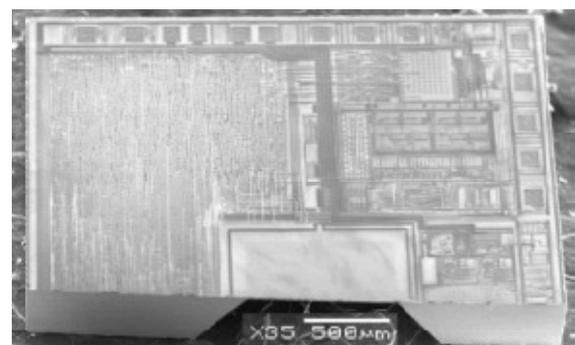


Abbildung 17 SM5822/SM5872  
REM Bild des kointegrierten  
Sensors [SMI07]

Zum Beispiel finden sich Produkte, die neben dem Drucksensor auch schon die Kompensationsschaltungen für Temperatur sowie die Kalibrationsdaten und deren Auswertung beinhalten. Abbildung 16 zeigt ein kommerziell verfügbares Produkt als gehäustes Bauteil. Gut zu sehen sind die unterschiedlichen Varianten des Gehäuses. In Abbildung 17<sup>30</sup> ist das Siliziumkristall dieses Produktes dargestellt. Zur besseren Darstellung ist der Chip in der Mitte durchtrennt. Deutlich erkennbar ist die Druckmembran von etwa 1 mm<sup>2</sup>. Umgeben wird diese von den Bauteilen der Steuerungs- und Auswertelektronik.

### 2.6.5 Zusammenfassung MEMS

Die Anforderungen an die Strukturbreite sind bei der MEMS-Fertigung erneut geringer als bei der ASIC-Fertigung. Die Anforderungen an die Flexibilität sind hingegen erneut höher. Insbesondere ist die Trennung zwischen Produkt und Prozess hier nicht mehr gegeben. Das heißt, praktisch jedes Produkt wird in einem angepassten Prozess gefertigt. Für die Erfüllung der Qualitätsanforderungen der Kunden hat dies erhebliche Relevanz. Insbesondere beim Einsatz von MEMS-Produkten in Industriebereichen wie der Automobilindustrie kollidieren die etablierten Verfahren zur Qualitätssteigerung mit den Fertigungsanforderungen. Damit können bereits die –für die aktuellen Anforderungen nicht ausreichenden– etablierten Verfahren hier nicht eingesetzt werden. Aktuell gibt es gemeinsame Bestrebungen von Wissenschaft und Industrie, diese Probleme durch Definition und Entwicklung von geeigneten Verfahren zu reduzieren [COR08].

Typische MEMS-Fabriken fertigen durchaus noch auf 100mm-Wafern, und nur besonders moderne Fabriken sind auf Wafer mit dem Durchmesser 150mm umgestiegen. Die hier notwendigen minimalen Strukturbreiten liegen nicht im Sub- $\mu$ -Bereich.

---

<sup>30</sup> Rasterelektronenmikroskop, englisch SEM, Scanning Electron Microscope

## 2.7 Solarzellen

Die Fertigung von Solarzellen hat –gerade in Deutschland– aufgrund der wachsenden Bedeutung regenerativer Energien stark zugenommen. Die Anlagen und die Prozessschritte entsprechen im Wesentlichen denen der Halbleiterfertigungsindustrie.

Allerdings sind die Anforderungen an die Fertigung, die Maschinen und nicht zuletzt die Rohstoffe und Basismaterialien zum Teil deutlich geringer. So wurden zu Beginn der Solarzellenfertigung häufig für die Halbleiterfertigung nicht mehr brauchbare Siliziumwafer genutzt. Dieses ressourcen- und kostenschonende Verfahren ist aufgrund der Industrialisierung der Fertigung und der damit einhergehenden Entwicklung der Fertigungsmengen nicht mehr durchzuhalten, da die Bedarfe der Solarzellenproduzenten durch die zur Verfügung stehenden Wafer nicht näherungsweise gedeckt werden können.

Der starke Nachfrageanstieg bei den Solarzellen hat allerdings die Fertigungskapazitäten der Hersteller von Rohsilizium erschöpft. Das für die Fertigung von monokristallinen Solarzellen notwendige Basismaterial muss von geringerer Reinheit sein als für die Halbleiterfertigung und ist damit natürlich preiswerter. Da hier aber nicht ausreichend Material angeboten wird, haben die Produzenten von monokristallinen Solarzellen auch das höherwertige Basismaterial für die Halbleiterfertigung zugekauft. Durch den hieraus entstehenden Nachfrageüberhang sind die Preise für die Rohwafer in der jüngeren Vergangenheit stark angestiegen. Besserung ist hier –insbesondere im Bereich der 150mm–Scheiben– nicht zu erwarten, da die Neuinvestitionen der Waferproduzenten sich auf die 300mm (und demnächst wohl die 450mm) Scheiben konzentrieren. Die Rückwirkung auf die Halbleiterindustrie besteht damit in einem weiter steigenden



Abbildung 18 Polykristalline Solarzellen [SLIK06]

Kostendruck, da die erheblichen Mehrkosten für das Rohmaterial typischerweise nicht an die Kunden weitergegeben werden können.

Abseits der Solarzellenfertigung trifft dies auch auf weitere Rohstoffpreise, insbesondere die Nichteisenmetalle wie Kupfer und Gold zu, die während des Assemblyprozesses benötigt werden. Die in anderen Industriezweigen üblichen Rohstoffzuschläge (Stichwort Kupferzuschlag) sind in der Halbleiterindustrie nicht etabliert.

Abbildung 18 zeigt ein Fotovoltaik-Modul aus polykristallinen Solarzellen. Diese werden nicht wie die monokristallinen Zellen aus einem aus der Schmelze gezogenen Einkristall hergestellt sondern gegossen. Dieses Herstellungsverfahren ist wesentlich einfacher und damit preiswerter. Zudem können die Zellen einfacher nahezu quadratisch hergestellt werden, was den Flächenverlust gegenüber den runden Zellen bei der Anordnung in einem Solarpanel reduziert.

Allerdings ist auch der Wirkungsgrad gegenüber den monokristallinen Zellen wesentlich geringer. Bei den polykristallinen Zellen gibt es keine Konkurrenz um die Rohstoffe mit den Halbleiterherstellern, da das Basismaterial für die Halbleiterfertigung nicht brauchbar ist.

Für die in diesem Dokument beschriebenen Analyseverfahren ist bei der Solarzellenfertigung aktuell keine relevante Anwendungsmöglichkeit ersichtlich. Daher wird das Thema Solarzellenfertigung im Weiteren nicht behandelt.

### ***2.8 Fertigungsleitsysteme***

Die beschriebene Kapitalintensität der Halbleiterfertigung führt unmittelbar zu dem Zwang, die extrem teure Infrastruktur möglichst intensiv zu nutzen. Zudem ist der Fertigungsablauf sehr komplex, vielschichtig, fehlerträchtig und langwierig.

Die sich daraus ergebenden Anforderungen an die Planung und die Transparenz der Fertigung machen die Nutzung von spezialisierten Fertigungsleitsystemen(MES<sup>31</sup>) notwendig.

Der Standardisierungsgrad ist hier sehr niedrig. Es sind weltweit Produkte wie Promis [POPP05,APP09], Workstream [APP06] oder Factory Works [APP08] verfügbar. Diese sind primär auf die Anforderungen von Fabriken ausgelegt, die mit hohen Volumina Standardprodukte wie Speicher oder Prozessoren herstellen.

Selbst für diese Anforderungen sind diese Systeme immer sehr stark auf die jeweiligen lokalen Anforderungen angepasst und haben sich typischerweise weit vom Standardsystem entfernt. Diese Anpassungen gehen weit über ein Einstellen der Software hinaus. Sehr häufig werden –durch die Anforderungen erzwungen– undokumentierte Schnittstellen der Software für die Programmierung von zusätzlichen Funktionalitäten genutzt. Klarerweise führt dieses zu stark erhöhten Aufwendungen für die Wartung der Systeme.

Das Einspielen einer neuen Softwareversion der Fertigungssteuerungssoftware ist in einer Halbleiterfabrik ein großes und risikoreiches Projekt. Ein Softwareupdate zur Laufzeit wird von den verwendeten Technologien nicht unterstützt. Zudem gibt es keine wirklich zufrieden stellende Testmöglichkeit, da natürlich keine „Testfabrik“ bereitsteht. Darüber hinaus gibt es die oben beschriebenen Modifikationen der Vorgängerversion und die Sicherheit, dass neue Fehler in der neuen Software verborgen sind. Entsprechend zurückhaltend sind die Anwender bei der Nutzung von Softwareupdates.

Wegen der beschriebenen besonderen Anforderungen von kleineren und flexibleren Fabriken sind die notwendigen Aufwendungen für die Anpassung der Standardsysteme in diesen Fällen noch weiter gehend. Da zudem diese Systeme wie oben beschrieben für den Einsatz in großen Fabriken ausgelegt sind, sind die Lizenzkosten sehr erheblich. Auch die Anforderungen an die Hardware sind aufwändig zu erfüllen. Werden ältere Versionen der Software eingesetzt, unterstützen diese häufig auch

---

<sup>31</sup> Manufacturing Execution System: Fertigungsleitsysteme, die typischerweise die Verbindung zwischen der übergeordneten Unternehmensteuerung (ERP) und der Maschinenebene der Fertigung darstellen. Häufig als Shop Floor Control System bezeichnet.

keine aktuelle Hardware, so dass alte Versionen von Hardware, Betriebssystem und gegebenenfalls Datenbank mit hohem Aufwand weiterbetrieben werden müssen. Zudem werden –aufgrund der Herkunft der Systeme– typischerweise Großrechner oder bestenfalls Systeme der mittleren Datentechnik unterstützt. Diese Systeme konnten sich dem allgemeinen Preisverfall der Rechentechnik weitestgehend entziehen und stellen nach wie vor für Beschaffung und Betrieb erhebliche Kostenfaktoren dar.

Erschwerend kommt hinzu, dass die genannten marktgängigen Produkte inzwischen durch einen einzigen Hersteller gekauft wurden. Für 300mm-Fabriken gibt es neben den oben genannten Produkten mit Si-View noch eine Alternative von IBM [IBM08]. Für 200mm-Fabriken hat sich durch die Aquisitionsstrategie von Applied Materials eine faktische Monopolstellung ergeben.

Software-basis	Kommerzielle Lösung	Datenbank / Browserzentrierte Lösung	3-Tier Lösung mit Pascal	Entwicklung innerhalb ERP <sup>32</sup> (SAP / ABAP)
Hardware-basis	Open VMS / ALPHA	Unix (AIX, Solaris) / Sun Sparc, IBM P-Series	IBM Z-Series	Windows Server / X86
Datenbanken	Proprietär	Informix	Oracle	DB2

Tabelle 4 Heterogenität verwendeter MES

Die genannten Punkte führen dazu, dass in vielen Fabriken auf die Beschaffung der sogenannten Standardprodukte zugunsten einer Speziallösung verzichtet wird.

---

<sup>32</sup> Enterprise Resource Planning: Verallgemeinerte Bezeichnung für Unternehmenssteuerungssoftware, teilweise einschließlich Logistik und Planung.

Bei den im Einsatz befindlichen Speziallösungen zeigt sich ein sehr weites Feld. Es finden sich sowohl intern entwickelte Lösungen als auch Auftragsentwicklungen durch Softwarehersteller [CIM99]. Einen Einblick in die Heterogenität der verwendeten Lösungen alleine in Deutschland zeigt Tabelle 4. Die Spalten sind hierbei unabhängig voneinander und es kommen durchaus auch Mischformen vor. Unmittelbar zu erkennen ist, dass es praktisch keine Gemeinsamkeiten gibt. Für Datenanalysewerkzeuge, die vom MES Daten benötigen, ist damit jeweils ebenfalls eine Speziallösung notwendig.

### 2.8.1 Funktionalität von Fertigungsleitsystemen

Aufgrund der oben geschilderten Heterogenität kann die Beschreibung der Funktionalität natürlich nicht einheitlich sein. Die existierenden Systeme besitzen als Kern immer ein WIP-Modul. Dieses Modul dient dazu, die aktuelle Materialverteilung in der Fabrik sowie den Anarbeitungsstand des Materials jederzeit zu kennen. Nur mit dieser Information macht der Versuch einer Fertigungssteuerung in einer Halbleiterfabrik überhaupt Sinn. Je nach Anforderungen der Fabrik, Möglichkeiten der verwendeten Software und Möglichkeiten der verwendeten Fertigungsmaschinen wird dieses Modul durch weitere Module ergänzt.

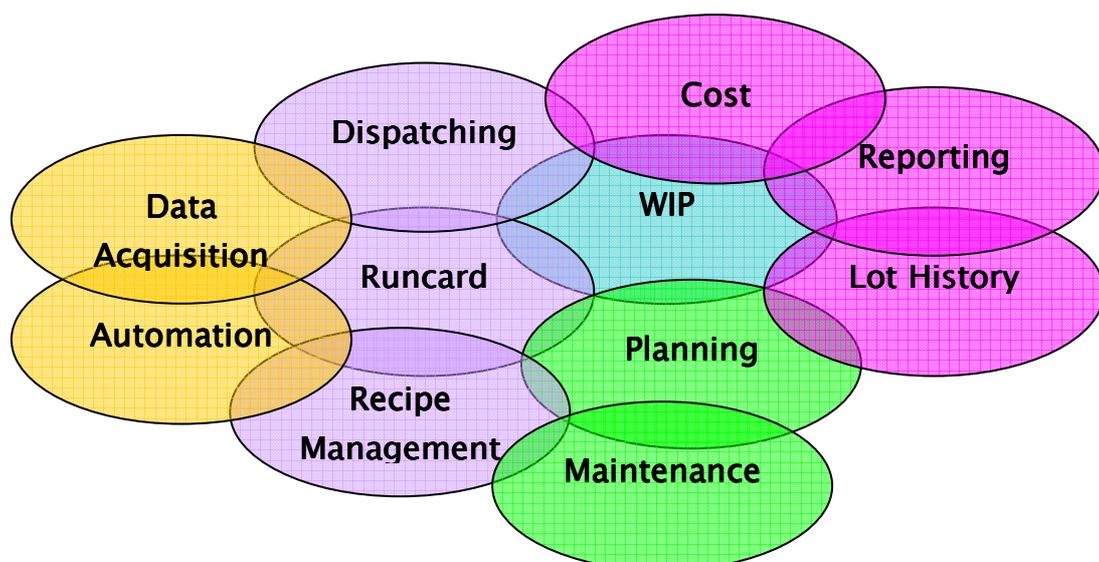


Abbildung 19 MES Übersicht

Abbildung 19 zeigt eine strukturierte Übersicht der eingesetzten Module. Zur besseren Übersicht sind die Module in verschiedenen Farben dargestellt. Die Bedeutung der Farben ist wie folgt:

- Grün Planung: Die Planung und die Wartung sind in die Zukunft gerichtete Komponenten.
- Rot Berichtswesen: Die Schnittstellen zu ERP und MIS<sup>33</sup> Systemen sind rötlich dargestellt.
- Blau WIP: Die WIP Komponente
- Lila Materialfluss: Neben der zentralen WIP Komponente wird die Feinsteuerung in der Fabrik mit den Lila dargestellten Modulen durchgeführt.
- Gelb Messdatenerfassung: Diese Komponenten betreffen die Auswertung der Fertigungsdaten.

Im Folgenden werden die dargestellten Komponenten erläutert.

Komponente	Funktionsumfang
Planung: Planning	Die Planungskomponente betrachtet die Fabrik (oder mehrere Fabriken) ganzheitlich. Die Planung erfolgt normalerweise kunden- oder vertriebsgetrieben, d.h. aus der Vorausschau der benötigten Fertigprodukte nach Produkt, Stückzahl und Termin wird die Einsteuerung in die Fabrik nach Kapazität, Produktkomplexität und aktueller Auslastung geplant.
Planung: Maintenance	Halbleiterfertigungsmaschinen sind hochkomplexe Maschinen, die mit hoher Präzision im 7*24 Betrieb funktionieren müssen. Das setzt natürlich voraus, dass die Anlagen angemessen gewartet werden. Da in dieser Zeit natürlich nicht produziert werden kann und es eine generelle produktionsfreie Zeit nicht gibt, ist die Abstimmung der Wartungszyklen mit dem Planungssystem eine kontinuierliche, sehr komplexe und sehr aufwändige Aufgabe. Erschwerend kommt hinzu, dass je nach vorgenommener Wartung nach Abschluss derselben eine Freigabeprozedur durchlaufen werden muss, die sicherstellen soll, dass

---

<sup>33</sup> Management Information System: Berichtssysteme für die Leitungsebene

Komponente	Funktionsumfang
	die Anlage nach wie vor spezifikationsgemäß funktio- niert.
Berichtswesen: Cost	Natürlich ist bei dem oben gezeigten Investitionsvolumen für eine Halbleiterfabrik eine Schnittstelle zum ERP System in vielen Installationen in Betrieb. Gerade in Fabriken, die viele verschiedene Produkte in verschiedenen Prozessen herstellen, ist eine Ermittlung des Ertrages eines Produktes ohne eine solche Schnittstelle kaum möglich.
Berichtswesen: Reporting	Die langen Durchlaufzeiten des Materials durch die Fertigung von typisch 8 – 10 Wochen machen ein aufwändiges Reporting notwendig. Schwierigkeiten müssen frühzeitig erkannt und abgestellt werden, da sich immer eine sehr erhebliche Menge an Material in der Fertigung befindet. Zudem ist eine kontinuierliche Messung der Leistung und der Qualität der Produktion unabdingbar. Diese wird typischerweise zielgruppenorientiert aufgebaut. Hochentwickelte Systeme sind zur zentralen Visualisierung von Schlüsselkennzahlen mehrerer Fabriken in der Lage. Abbildung 20 zeigt eine derartige Visualisierung. Hier wird der Vergleich zwischen Soll- und Istzahlen für vier Fertigungslinien an drei verschiedenen Standorten aktuell dargestellt. Diese Darstellung wurde vom Autor auf der ZuE2008 anlässlich des eingeladenen Vortrags über Zuverlässigkeit in der Entwicklung, Anforderung an Prozesse und IT Systeme publiziert [ZUE08].
Berichtswesen: Lot History	Aufgrund der Komplexität der Materie ist dem Material sehr selten direkt anzusehen, was die Ursache für eine Auffälligkeit ist. Um hier überhaupt eine Chance zur Ursachenermittlung zu haben, ist eine Lot History unabdingbar. Zudem kann man aus der Lot History Hinweise für ein zukünftiges Verhalten gleichartigen Materials ableiten, was die Planung verbessert. Schließlich ist die Lot History notwendige Voraussetzung für jede Form der Rückverfolgbarkeit.

Komponente	Funktionsumfang
Materialfluss: Recipe Ma- nagement	Das Rezeptmanagement sorgt dafür, dass auf Basis der Laufkarte die korrekten Rezepte für die Fertigungsmaschine bereitgestellt werden. Das ist insbesondere deshalb von Bedeutung, als für jeden Prozessschritt in einer Fertigung mehrere gleiche oder gleichartige Fertigungsanlagen zur Verfügung stehen. Diese –gewollten oder ungewollten– Unterschiede der Anlagen bei demselben Prozessschritt machen es notwendig, abhängig vom der mittels Dispatching festgelegten konkreten Anlage das jeweils für diese Anlage und den gewünschten Prozessschritt benötigte Rezept zu identifizieren.
Materialfluss: Runcard	Die Laufkarte ist das zentrale Dokument der Fertigungssteuerung in einer Halbleiterfabrik. Vor der Automatisierung der Fabriken stellte diese die alleinige und vollständige Fertigungsanweisung für das Material dar. Da die Durchführung der Prozessschritte in diesem Papierdokument dokumentiert wurde und die Ergebnisse von Inline-Messungen auch dort eingetragen wurden, war diese Laufkarte nach Ende des Fertigungsdurchlaufs zugleich die vollständige Fertigungsdokumentation. Durch die Automatisierung der Fabriken sind mindestens wesentliche Teile der Fertigungsanweisung auch oder ausschließlich in elektronischer Form im MES vorhanden. In „green field“-Fabriken ist typischerweise keine papierene Laufkarte vorhanden. Die nachträgliche Entfernung des Papiers aus gewachsenen Fabriken gestaltet sich sehr langwierig. Neben technischen Hindernissen wie nicht automatisierbare Maschinen sind die „weichen“ Faktoren nicht zu unterschätzen. Ähnlich der Situation in Büros gibt es bei sehr vielen Mitarbeitern eine starke Bindung an die papierbasierten Prozesse. Änderungen an dieser Stelle erfordern viel Zeit und die Überzeugung der Mitarbeiter vor Ort.

---

**Materialfluss:** Das Dispatching unterstützt das Fertigungspersonal bei der Feinplanung in der Fabrik. Üblicherweise wird bei der Steuerung von Halbleiterfabriken zwischen Grob- und Feinplanung unterschieden, da die Komplexität eine hinreichend präzise Planung der gesamten Fabrik oder gesamten Fabrikclusters nicht zulässt. Dieses Dispatching wird üblicherweise als RBD<sup>34</sup> ausgeführt. Die hier vorgegebenen Regeln sind meist pragmatisch und betrachten nur das direkte Umfeld. Dafür ist die Regelung aber schnell und durchschaubar. Zusammenfassend ist RBD für Fabriken mit im wesentlichen konstanten Materialfluss und wenigen signifikant unterschiedlichen Produkten ein erfolgreiches Konzept. Für die kleineren, flexiblen Fabriken sind die Störungen so groß, dass das System aus Regelung, Ergebnis und manueller Nachbesserung der Regeln nicht gut konvergiert.

---

**Messdaten-  
erfassung:  
Data Acqui-  
sition** Die Erfassung und Auswertung von Parametern betrifft sowohl Werte von Inline-Messungen an Fertigungsprodukten als auch Daten aus den Fertigungsmaschinen oder weitere vorhandene Daten etwa aus der Haustechnikinstallation. Klassisch ist dies keine Kernfunktion eines MES, sondern ist eher in EDA<sup>35</sup>-Systemen beheimatet. Die Datenaufnahme erfolgt aber –zumindest zum Teil– durch das MES, da man die Anzahl der in der Fabrik laufenden Systeme natürlich möglichst minimal halten will. Zudem unterstützt die typischerweise verwendete Automatisierungsschnittstelle nur eine Punkt-zu-Punkt-Kommunikation, so dass ein zusätzliches System hier nur mit erheblichen Problemen integrierbar ist<sup>36</sup>.

---

---

<sup>34</sup> Rule Base Dispatch: Beschreibung für typischerweise automatische Auftragsverwaltungssysteme in Fabriken. Das nächste zu fertigende Los wird an Hand von vorher festgelegten Regeln ausgewählt.

<sup>35</sup> Engineering Data Analysis

<sup>36</sup> In der Praxis findet man vereinzelt entsprechende Lösungen. Diese hängen als dritter Partner zwischen den beiden anderen und „belauschen“ die Kommunikation auf der

---

## 2 Halbleiterindustrie

Messdate-  
nerfassung:  
Automation

Die Automatisierung ist für die Erreichung der Qualitäts-  
ziele in einer flexiblen Fabrik unabdingbar. Für ein MES,  
das die Vermeidung von logistischen Verlusten zum Ziel  
hat, ist das eher ein Randthema. Unter anderem dieser  
Umstand führt zu erheblichen Problemen wie später noch  
gezeigt wird.



Abbildung 20 Zentrale Statusvisualisierung eines Fabrikclusters

seriellen Schnittstelle. Bei Kommunikationspausen setzen sie eigene Kommandos ab. Diese Lösung ist entsprechend kritisch und wird dann eingesetzt, wenn das verwendete MES keine anderen Möglichkeiten bereitstellt.

## ***2.9 Automatisierung***

Die Automatisierung folgt der Industrialisierung in den Industrieländern. Die Triebfeder der Automatisierung ist in den klassischen Industriezweigen üblicherweise der Versuch, die Arbeitskosten und die Zahl der Mitarbeiter zu reduzieren oder zumindest unterproportional mit dem gefertigten Material wachsen zu lassen. Entsprechend ersetzt häufig die Automatisierung manuelle Tätigkeiten, durchaus auch in gefährlichen oder sonst für den Menschen eher ungeeigneten Umgebungen.

### **2.9.1 Automatisierung im Halbleiterumfeld**

Abweichend davon ist die Triebfeder der Automatisierung<sup>37</sup> in der Halbleiterindustrie primär die Erhöhung der Qualität und die Reduzierung von Materialverlust durch menschliche Fehler.

Angesichts der beschriebenen Komplexität des Herstellungsprozesses sind viele Halbleiterhersteller dazu übergegangen, die Fertigungsanlagen mit der Fertigungssteuerung zu verbinden. Die ersten Festlegungen in diesem Umfeld wurden bereits 1980 durch die SEMI<sup>38</sup> getroffen. Der in den entsprechenden Normen beschriebene Kommunikationsstandard SECS<sup>39</sup>–GEM<sup>40</sup> [SEM4, SEM5, SEM30] beschreibt eine textbasierte Kommunikation über eine serielle Schnittstelle oder eine Netzwerkschnittstelle. Die Kommunikation über Ethernet wird auch als HSMS<sup>41</sup> bezeichnet. Es ist wichtig zu erwähnen, dass auch HSMS genau wie die Kommunikation über die serielle Schnittstelle eine reine Punkt-zu-Punkt-Kommunikation ist. Gleichzeitige Verbindungen zu mehreren Partnern,

---

<sup>37</sup> Automatisierung im Halbleiterumfeld beschreibt die Verbindung der Fertigungsmaschinen untereinander und mit der Fabriksteuerung. Wie bereits dargestellt, handelt es sich bei den installierten Fertigungsanlagen um hochkomplexe Maschinen. Der Fertigungsablauf in der Fertigungsmaschine funktioniert üblicherweise schon seit langem vollautomatisch. Die Automatisierung der Maschinen selber ist daher kein Thema.

<sup>38</sup> Semiconductor Equipment and Materials Institute, [www.semi.org](http://www.semi.org)

<sup>39</sup> SEMI Equipment Communications Standard

<sup>40</sup> Generic Model for Communications and Control of Manufacturing Equipment

<sup>41</sup> High-Speed SECS Message Services

wie sie in der TCP/IP-Kommunikation normalerweise absolut üblich sind, werden hier nicht unterstützt.

Trotz der bereits seit langem definierten Kommunikation ist die Automatisierung in einer Halbleiterfabrik immer noch ein sehr aufwändiges Projekt. Dies liegt zum einen daran, dass die Wurzeln der heute im Einsatz befindlichen MES ebenfalls sehr weit in der Vergangenheit liegen und diese Systeme als „Black Box“ entwickelt wurden. Weiterhin sind die marktgängigen Maschinen typischerweise nur teilweise konform zur Norm. Dies trifft für die oben genannten Mixed Signal- und MEMS-Fabriken in besonderem Maße zu, da die Entwicklung der Fertigungsmaschinen, die heute für diese Anforderungen geeignet sind, wie gezeigt viele Jahre zurück liegt. Die Relevanz der Automatisierung war zu jener Zeit für die Anlagenkäufer nicht wirklich zwingend und entsprechend niedrig war die Priorität der Anlagenhersteller bei diesem Thema.

Zusammengefasst führt das dazu, dass ein Automatisierungsprojekt für einen neuen Maschinentyp in einer Halbleiterfabrik selten unterhalb einer sechsstelligen Summe in Euro kostet. Das kann erfahrungsgemäß soweit gehen, dass eine Fertigungsanlage auf ein legales SECS-Kommando mit einem Not Aus reagiert was durch die dann zerstörten Turbopumpen viele tausend Euro an Schaden erzeugt.

Da die Fertigungsmaschinen letztendlich dem Sondermaschinenbau zuzuordnen sind und wenn überhaupt nur in Kleinserien entstehen, sind selbst identisch bezeichnete Maschinen typischerweise nicht gleich. Insbesondere unterschiedliche Softwarestände führen zu unterschiedlichem Verhalten an der SECS-Schnittstelle und damit zu erhöhtem Integrationsaufwand auch bei der Integration einer weiteren Maschine eines bereits integrierten Typs.

### **2.9.2 Stand der Technik bei der Automatisierung**

Trotz der beschriebenen Schwierigkeiten ist es auch in kleineren, modernen Fabriken Stand der Technik, die korrekte Durchführung der vorgesehenen Prozessschritte durch die beschriebene Automatisierung der Fertigungsanlagen zu steuern. Durch die Anforderungen des Marktes an Lieferfähigkeit, Qualität und nicht zuletzt der Rückverfolgbarkeit der

Produkte bleibt keine Alternative zur Automatisierung zumindest eines Teils der Fertigungsmaschinen.

Konkret wird im Allgemeinen der Name des korrekten Rezeptes für den nächsten Prozessschritt an die Anlage übergeben. Das Rezept befindet sich typischerweise auf der Anlage und wird über den Rezeptnamen ausgewählt. Der Abgleich des Rezeptes mit dem Arbeitsplan des aktuellen Loses erfolgt durch die Fertigungssteuerung.

Der offensichtliche Schwachpunkt dieses Konzeptes ist die Unsicherheit bezüglich des Rezeptes. Eine Veränderung des Rezeptes direkt an der Maschine bei Beibehaltung des Rezeptnamens oder unterschiedliche Rezepte mit identischen Namen auf mehreren Maschinen desselben Typs können Fehlerquellen sein.

Die vermeintlich sichere Lösung, das Rezept serverseitig zu speichern und bei Bedarf auf die Maschine zu laden scheitert an dem bereits beschriebenen sehr kritischen Verhalten der marktgängigen Fertigungsmaschinen.

Eine bisweilen gewählte Alternative ist der „smart download“. Hier werden die Rezepte vor dem Start von der Produktionsmaschine auf einen Server geladen und mit dem Sollzustand verglichen. Da es sich bei den Rezeptinformationen der Maschinen im Allgemeinen um undokumentierte –häufig binäre– Formate handelt, kann die Überprüfung nur rein formal sein.

Final sind immer die Risiken gegen den Fertigungsdurchsatz zu bewerten. Schließlich ist die Übertragung über die häufig vorzufindende serielle Schnittstelle vergleichsweise langsam. Hinzu kommt, dass –wie schon beschrieben– die Steuerrechner der Fertigungsmaschinen normalerweise nicht Stand der Technik sind. Für die Behandlung der seriellen Schnittstelle bedeutet das häufig eine Beschränkung der Übermittlungsgeschwindigkeit auf 9600bd. Schwerwiegender als diese Beschränkung ist die Ursache für diese. Üblicherweise muss die Kommunikation nämlich durch die Zentraleinheit selber durchgeführt werden. Die heute üblichen intelligenten Schnittstellensysteme sind nicht vorhanden. Damit steigt mit der Geschwindigkeit und der Länge der Übermittlung das Risiko eines Pufferüberlaufs oder einer durch die Zentraleinheit nicht behandelten Unterbrechungsanforderung durch ein anderes Subsystem der Fertigungsanlage stark an. Beide Fälle führen unter Umständen zu Fehl-

funktionen der Anlage, was wiederum schwerwiegende Konsequenzen für das gesamte Material haben kann, das sich zu dem Zeitpunkt in der Anlage befindet. Es muss also die Erhöhung der Sicherheit gegen Fehlprozessierung eines Loses gegen die Erhöhung des Verschrottungsrisikos mehrere Lose abgewägt werden.

Weiterhin ist zu bedenken, dass es eine absolute Sicherheit nicht geben kann, da ja auch an der Hardware der Fertigungsanlage Veränderungen durchgeführt werden. Die Rezeptänderungen werden durch das betreuende Engineering durchgeführt und können im Allgemeinen gut durch organisatorische Maßnahmen abgesichert werden. Daher ist das oben beschriebene Verfahren der Übertragung des Rezeptnamens eine gute und weithin angewandte Lösung.

### 2.9.3 Schlüsselfaktoren für die Automatisierung

In 300mm-Fabriken ist ein manuelles Handling der Lose nicht sinnvoll durchführbar. In vielen Ländern verbieten Arbeitsschutzgesetze das regelmäßige manuelle Handling wegen des erheblichen Gewichtes. Zudem kann der Wert eines Loses sehr leicht mehrere einhundert tausend Euro betragen. Ein manuelles Handling geht immer mit der Gefahr des Fallens einher und schon wegen dieses Risikos ist der manuelle Transport der Lose in derartigen Fabriken kein vorgesehener Arbeitsschritt. Üblich ist hier eine Vollautomatisierung des Materialhandlings in den Fabriken (AMHS<sup>42</sup>). In diesem Fall gibt es keine Alternative zur Automatisierung der Fertigungsanlagen.

In den bereits beschriebenen kleineren ASIC-Fabriken ist ein Transportsystem sowie die durchgängige Benutzung von Minienvironments mit entsprechenden SMIF<sup>43</sup>-Systemen im Allgemeinen wegen der Kosten und

---

<sup>42</sup> Automated Material Handling Systems, typischerweise automatisches Schienentransportsystem an der Decke des Fertigungsbereiches. Praktisch immer mit SMIF kombiniert.

<sup>43</sup> Standard Mechanical InterFace: Standard für Waferfertigung mit sog. Minienvironments. Der Transport zwischen den Maschinen erfolgt in sog. SMIF PODS, hermetisch geschlossenen Behältern. Diese werden an den Fertigungsmaschinen automatisch geöffnet und das Material entnommen. Das Ziel ist besondere Sauberkeit.

der mit der Vollautomatisierung erschwerten Flexibilität nicht umgesetzt. Die Automatisierung wird aber typischerweise trotz der beschriebenen Risiken und Kosten durchgeführt, sobald die notwendigen Bedingungen erfüllt sind.

Treibender Faktor ist hier die Fehlervermeidung. Die bessere Transparenz der Fertigungsanlagen und die damit einhergehenden wesentlich verbesserten Möglichkeiten zur Prozesssteuerung werden zwar auch gesehen, sind aber sehr schwer im Voraus quantifizierbar. Die Vermeidung der Verschrottung von Material durch den automatisierten Check, ob der richtige Prozessschritt und das richtige Rezept für das gerade für die Prozessierung vorgesehene Material gewählt ist, ergibt einen direkt darstellbaren ROI<sup>44</sup> und ist damit praktisch immer der Schlüssel für ein Automatisierungsprojekt.

Voraussetzungen für die Durchführung eines Automatisierungsprojektes ist die bereits beschriebene technische Integrationsfähigkeit der Anlage, typischerweise mit SECS II. Weiterhin notwendig ist eine präzise elektronische Repräsentation der Fertigungsanweisungen. Üblicherweise sind im Rahmen einer Automatisierung sehr umfangreiche Erweiterungen und Änderungen am Fertigungsleitsystem notwendig.

---

<sup>44</sup> Return On Invest: Zeitpunkt, ab dem eine Investition einen positiven Ergebnisbeitrag erzeugt sowie die Höhe des Beitrages.

# 3 Kundenanforderungen / Qualität

Wie bereits dargestellt, ist der Herstellungsprozess in der Halbleiterfertigung prinzipbedingt fehlerbehaftet. Neben den durch Defekte verursachten Fehlfunktionen gibt es auch Fehlfunktionen oder Spezifikationsverletzungen der fertigen Bauelemente durch Bearbeitungsfehler des Bedienpersonals, Fehlfunktionen der Fertigungsanlagen etc.

Fehlfunktionen in der Größenordnung von Promille sind heutzutage auch in der Konsumgüterindustrie oder der Unterhaltungselektronik nicht mehr akzeptabel. Daraus ergibt sich der unmittelbare Zwang für jeden Hersteller, integrierte Schaltungen mindestens einmal elektrisch zu testen, um die Funktion sicherzustellen. Stichprobenverfahren, wie sie in anderen Industriezweigen üblich sind, reichen hier nicht aus. Für besondere Qualitätsanforderungen (z.B. Automotive) ist ein mehrfacher 100%-Test bei verschiedenen Ecktemperaturen sowie die künstliche Voralterung der Produkte (Burn In) nicht ungewöhnlich.

Der von den Kunden geforderte Temperaturbereich für automobiler Anwendungen reicht im Standard bereits von  $-40^{\circ}\text{C}$  bis  $+125^{\circ}\text{C}$ . Je nach Umfeld der Anwendung im Fahrzeug sind  $125^{\circ}\text{C}$  als maximale Umgebungstemperatur nicht ausreichend (Beispiele: Anwendungen im Ölkreislauf, in der Nähe der Bremsen). In diesen Fällen muss die obere

Temperaturgrenze noch angehoben werden. Dies ist insbesondere wegen der verwendeten Kunststoffe für das Assembly der Gehäuse problematisch. Bei ähnlichen Anforderungen werden im militärischen Einsatz häufig Keramikgehäuse verwendet, die über einen größeren Temperaturbereich funktionieren. Allerdings ist die Zuverlässigkeit von Bauteilen in Keramikgehäusen deutlich schlechter als die von Bauteilen in Kunststoffgehäusen. Zudem verhindert der exorbitante Preis den Einsatz von Keramikgehäusen in der Großserie.

Zur Absicherung der Zuverlässigkeit ist darüber hinaus eine künstliche Voralterung der Bauteile mittels dynamischem BurnIn üblich. Hier werden die Bauteile über einen definierten Zeitraum (typisch 12 - 48 Stunden) bei möglichst hoher Umgebungstemperatur betrieben. Ursprünglich war das Ziel dieser Voralterung das Aussortieren der Frühausfälle.

Im Rahmen der stark gestiegenen Qualitätsanforderungen wird inzwischen das Vorhandensein von Frühausfällen durch die Kundschaft nicht mehr akzeptiert, so dass aus dem ursprünglich qualitätsverbessernden Fertigungsschritt mit anschließender Messung ein rein qualitätsüberwachender Messschritt geworden ist. Ein Überblick über aktuelle Qualitätsanforderungen und den Umgang damit findet sich in [KMCM07].

Der elektrische Test von integrierten Schaltungen teilt sich in den Parametertest (PCM<sup>45</sup>) und den Funktionstest auf.

Üblicherweise ist der Ablauf wie folgt:

1. Waferfertigung
2. PCM, Parametertest
3. Funktionstest auf Waferlevel, teilweise mehrfach bei verschiedenen Temperaturen
4. Assembly
5. Funktionstest der Final Parts, teilweise mehrfach bei verschiedenen Temperaturen
6. Gegebenenfalls BurnIn
7. Gegebenenfalls weitere Funktionstests auf Final Parts

---

<sup>45</sup> Process Control Monitor: Parametertest von prozessabhängigen Teststrukturen

### ***3.1 Process Control Monitor***

Der Parametertest ist die Ausgangskontrolle der Halbleiterfertigung. Auf jedem gefertigten Wafer werden neben den Produkten für diese Kontrolle Einzelbauelemente gefertigt. Diese Einzelbauelemente sind für den gewählten Herstellungsprozess typischerweise einheitlich und unterscheiden sich nicht bei verschiedenen Produkten, die im selben Prozess gefertigt werden. Die Anzahl der Messungen pro Messpunkt sind üblicherweise im dreistelligen Bereich. Da Einzelbauelemente vermessen werden, ist der Temperaturverlauf normalerweise klar. Daher wird in der laufenden Fertigung PCM nicht bei verschiedenen Temperaturen durchgeführt.

Diese Einzelbauelemente (z.B. verschiedene Transistoren, Kontaktketten) werden dann nach Abschluss der Fertigung mit spezialisierten automatischen Testsystemen elektrisch vermessen. Diese Messergebnisse werden gegen die entsprechenden Spezifikationsgrenzen des Prozesses geprüft. Bei Auffälligkeiten werden weitergehende Analyse-schritte durchgeführt. Je nach Ergebnis der Analyse werden Teile einer Scheibe, eine oder mehrere Scheiben oder die gesamte Charge verworfen.



Abbildung 21 S680 DC/RF Parametric-Test-System [KEITH08]

Vorrangig dient der Parametertest der Erzeugung von Regelgrößen für die Fabrik. Anhand der hier ermittelten Messergebnisse werden gegebenenfalls Fertigungsparameter nachgeführt. Damit dient diese Prüfung der Verbesserung und Stabilisierung der Halbleiterfabrik.

Produziert werden diese Bauelemente im Ritzrahmen auf den Halbleiterscheiben. Wie beschrieben werden die Bauteile im weiteren Verlauf der Herstellung vereinzelt. Dazu werden diese mittels geeigneter Sägen entlang vorher festgelegter Linien gesägt. Diese für die Breite des Säge-

blatts notwendigen Flächen werden als Ritzrahmen bezeichnet und stehen klarerweise für die Produkte nicht zur Verfügung.

Nur in Ausnahmefällen wird über die Einblendung eines Testinserts verwertbare Fläche für die parametrische Messung genutzt. Dies geschieht dann, wenn die Flächen (insbesondere die Breiten) im Ritzrahmen für spezielle Untersuchungen nicht ausreichen.

Vom theoretischen Ansatz ergibt sich aus den PCM-Ergebnissen direkt die Qualität der Scheibenfertigung. Unter Berücksichtigung der Defektdichte ergibt sich damit direkt die Ausbeute der gefertigten Produkte. Leider gibt es in der industriellen Fertigung weitere zu berücksichtigende Dinge, so dass sich die Realität leider sehr komplex darstellt und die „einfache“ Korrelation: PCM-Ergebnisse ↔ Ausbeute verhindern.

Die Messungen sind zeitaufwändig und stellen hohe Anforderungen an die verwendeten Messgeräte und den sonstigen Aufbau. Die für den Funktionstest der Bauteile genutzten automatischen Testsysteme sind für den Parametertest nicht brauchbar, da diese Systeme die hier notwendige Messgenauigkeit und Empfindlichkeit nicht aufweisen. Umgekehrt sind die PCM Testsysteme für den Funktionstest nicht zu gebrauchen, da die Messgeschwindigkeit für den Funktionstest aufgrund der Empfindlichkeit der Messgeräte nicht ausreichen kann. Zudem verfügen PCM Testsysteme nicht über die notwendigen Geräte für den Test von digitalen Schaltungsteilen. Für den PCM Test werden also Spezialsysteme eingesetzt, die weltweit nur von sehr wenigen Herstellern angeboten werden. Ein Beispiel für ein derartiges System befindet sich in Abbildung 21. Die Anforderungen an den Messbereich eines PCM Systems sind mit bis zu 15 Größenordnungen sehr weitgehend (s. Abbildung 22).

Die Auswertung der Ergebnisse wird in existierenden Fabriken von erfahrenen Spezialisten durchgeführt, da es einige problematische Randbedingungen gibt:

- Unter Produktionsbedingungen wird keine flächendeckende parametrische Messung vorgenommen, sondern man beschränkt sich aus Kostengründen auf eine vergleichsweise kleine Anzahl von Messpunkten. Diese hängt natürlich von der Scheibengröße ab und bleibt im einstelligen oder kleinen zweistelligen Bereich pro Scheibe.

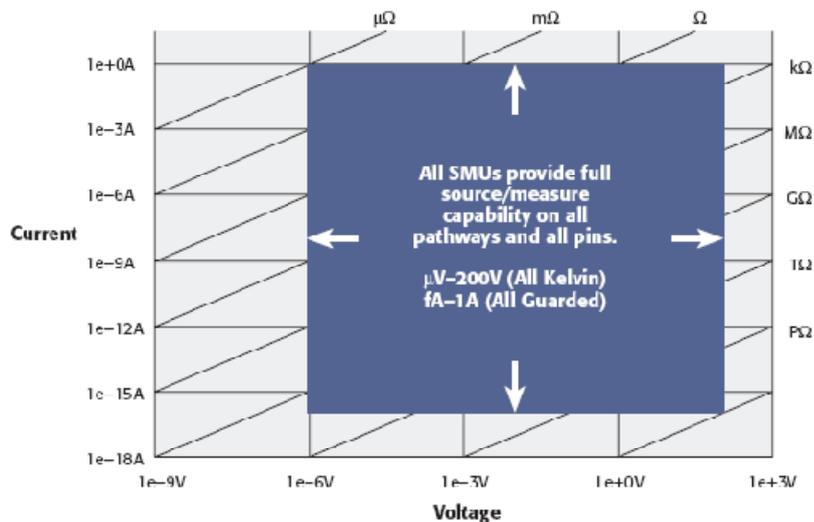


Abbildung 22 Messbereich PCM-System [KEITH08]

- Viele der gemessenen Parameter verändern sich über die Scheibe, da die zu Grunde liegenden Strukturen wie z.B. Schichtdicken ebenfalls nicht konstant über die gesamte Fläche herstellbar sind.
- Durch fehlerhafte Kontaktierung oder durch aufgrund von Defekten nicht funktionierende Bauelemente sind nicht immer alle Messwerte der Parametermessung einer Scheibe vorhanden.
- Da hier einzelne Bauelemente vermessen werden, sind die gemessenen Werte sehr klein. Typische gemessene Werte sind Femtoampere oder darunter, Bruchteile von Millivolt oder Femtofarad. Damit ist der Messfehler einer einzelnen Messung nicht mehr zu vernachlässigen. Standard ist eine 4-Punkt-Messung (Kelvin) ebenso wie eine dynamische Nachführung des Potentials auf dem Masseschirm der Leitungen (Guard).
- Eine sichere Entscheidung bezüglich der Validität der Messwerte ist nicht immer ohne den Kontext möglich. Insbesondere bei Leckstrommessungen ist die Entscheidung, ob die Nadeln überhaupt sauber kontaktiert wurden, teilweise nicht zu treffen, ohne die Bewertung von anderen Messungen am selben Bauelement zu berücksichtigen.

### **3.2 Funktionstest**

Nach der bestandenen Ausgangskontrolle der Halbleiterfabrik (PCM) ist klar, dass die Produktion der Wafer innerhalb der vorgegebenen Prozess toleranzen erfolgreich war, wobei eine sich aus der Größe der gewählten Stichprobe und der Verteilung der Messpunkte auf der Scheibe ergebende Restunsicherheit bleibt.

Nun werden die produzierten Bauteile auf ihre spezifikationsgemäße Funktion überprüft. Bei entsprechenden Qualitätsanforderungen des Kunden werden alle spezifizierten Parameter und noch einige mehr gemessen. Typischerweise enthält das Bauteil auch digitale Baugruppen wie Logik oder Speicher. Diese Baugruppen werden ebenfalls auf vollständige Funktion überprüft. In der Summe kommen hier häufig deutlich mehr als 100 Messwerte pro gemessenes Bauteil und pro Prüfschritt zusammen.

Dies ist aus mehreren Gründen unabdingbar:

- Wie beschrieben, kommt es durch Defekte oder allgemeiner durch lokale Prozessabweichungen zu Fehlfunktionen einzelner Bauelemente. Diese lokalen Effekte können durch PCM nicht erkannt werden.
- Die Parameter der Bauteile in der Schaltung des Produktes schwanken natürlich genauso wie die Parameter der Bauteile, die der PCM unterworfen werden. Insbesondere bei Mixed-Signal- und Anlogschaltungen ist es nicht sicher auszuschließen, dass es Kombinationen von Parameterlagen gibt, bei denen die Schaltung nicht oder nicht vollständig spezifikationskonform funktioniert. Zwar wird während der Entwicklung der Schaltung vieles simuliert, aber alle Kombinationen von allen Parametern sind klarerweise nicht in akzeptabler Zeit simulierbar.
- Wird für die Herstellung der Scheiben ein Dienstleister gewählt (Silicon Foundry<sup>46</sup>), findet zwischen dem Parametertest und dem

---

<sup>46</sup> Halbleiterhersteller der Produkte seiner Kunden mit eigenen Prozessen fertigt. Typischerweise werden die prozessierten Scheiben nach Bewertung der PCM Ergebnisse verkauft. Die Verantwortung für die Funktion der Schaltung trägt der Auftraggeber.

Funktionstest häufig ein Wechsel der Umgebung statt. Die PCM wird durch den Waferfabrikanten mit Regeln und Rezepten des Waferfabrikanten durchgeführt. Der Funktionstest setzt hingegen intensives Wissen über die internen Strukturen der Produkte voraus. Das entsprechende Testprogramm entsteht parallel mit der Produktentwicklung. Bei kleineren Serien und im Produktanlauf wird der Funktionstest typischerweise durch das entwickelnde Unternehmen durchgeführt. Zudem ist die Durchführung des Funktionstestes durch Dritte auch wegen dem potentiellen Verlust von geistigem Eigentum des entwickelnden Unternehmens problematisch.

Abhängig von den Qualitätsanforderungen des Kunden und der aktuellen Situation des Produktes im jeweiligen Produktlebenszyklus werden verschiedenste Prozesse mit dem Material durchgeführt.

#### 3.2.1 Bauteilidentifikation

Beim Funktionstest auf Waferlevel [s. Abbildung 23 für ein Kontaktersystem] sind noch die topologischen Informationen vorhanden. Damit können die Messwerte noch der Scheibenummer sowie der Position auf der Scheibe zugeordnet werden. Mit dem Assemblyprozess ergibt sich ein Systembruch. Die Teile sind nach Assembly normalerweise nicht mehr unterscheidbar. Dieser Umstand erschwert alle Analysen klarerweise erheblich.



Abbildung 23 Waferprober [EGLAS07]

Die Erzeugung einer eindeutigen Identifikationsnummer auf jedem Bauteil ist eine wirkungsvolle Gegenmaßnahme gegen diesen Systembruch. Allerdings setzt die Erzeugung dieser Nummer, deren Speicherung und die zugehörige Verwaltung viele Dinge voraus:

1. Erzeugung der Nummer: Damit jedes Bauteil eine eigene Nummer bekommen kann, ist ein programmierbarer, nichtflüchtiger Speicher zwingende Voraussetzung. Je nach vorhandenen Prozessoptionen und den Anforderungen an das Bauteil kann dieses als OTP<sup>47</sup> oder NV-Ram<sup>48</sup> ausgelegt sein. Ist sowieso nichtflüchtiger Speicher auf dem Produkt geplant, ist für die Seriennummer des Bauteils zusätzliche Chipfläche notwendig. Im Allgemeinen sind die wenigen notwendigen Bytes hier allerdings nicht entscheidend. Anders stellt sich die Situation dar, wenn das Produkt ansonsten keinen nichtflüchtigen Speicher benötigt. In diesem Fall können neben der zusätzlichen Chipfläche weitere Bearbeitungsschritte oder sogar Prozessoptionen sowie zusätzliche Maskenkosten notwendig werden. Angesichts des allgegenwärtigen Preisdrucks können diese Umstände leicht dazu führen, dass ein Produkt keine Identifikationsnummer bekommen kann. Hinzu kommt das zusätzliche Risiko durch den notwendigen zusätzlichen Betriebsmodus des Produkts. Da ja keine zusätzlichen Verbindungen für Testmodi bereitstehen, muss das Bauteil über die vorhandenen Kontakte in einen Modus gebracht werden können, in dem die Seriennummer ausgelesen werden kann. Da das Produkt in diesem Testmodus die normalen Funktionen nicht zeigt, muss ein Wechsel in den Testmodus in der Applikation sicher ausgeschlossen werden. Natürlich darf dieser Testmodus daher auch nicht persistent ausgeführt werden.
2. Speicherung der Nummer: Die Kennung kann sinnvollerweise nur als nicht sprechender Schlüssel ausgeführt werden. Die Kodierung von Information im Schlüssel selbst würde die Menge der für die Speicherung notwendigen Speicherstellen unnötigerweise erhöhen. Hinzu kommt, dass der Hersteller kein Interesse daran haben kann, dass der Kunde die topologischen Informationen ebenfalls dekodiert. Handelt es sich bei der benutzten Speichertechnologie

---

<sup>47</sup> One Time Programmable: Einmal programmierbare Halbleiterstruktur.

<sup>48</sup> Non-Volatile Random Access Memory: Allgemeine Bezeichnung für nicht flüchtigen Speicher. Die Inhalte bleiben auch nach Abschalten der Versorgungsspannung erhalten.

um NV-RAM, ist es notwendig, die Überschreibung der Seriennummer nach der ersten Programmierung zuverlässig zu unterbinden. Dies gilt insbesondere für die zu verhindernde Änderung der Seriennummer durch den Anwender.

3. Für die Verwaltung der Nummern und die Rückverfolgbarkeit ist ein hoch entwickeltes Verwaltungssystem notwendig. Dieses System muss sicherstellen, dass
  - a. sich die Nummern innerhalb einer sinnvollen Periode nicht wiederholen.
  - b. die richtigen Nummern zum richtigen Zeitpunkt in den jeweiligen Chip geschrieben werden. Dies ist nur während des Wafertests möglich, wobei das Verwaltungssystem die Koordinaten des Chips auf der Scheibe sowie Scheiben und Chargennummer kennen und speichern muss.
  - c. sämtliche Prüfergebnisse mit diesem Produkt über die Seriennummer zusammenführbar sind.

Der Hersteller muss gegenüber dem Kunden den vollständigen Spezifikationsumfang sowohl über die Lebensdauer des Produktes als auch über den Temperaturbereich gewährleisten.

Um diese geforderte Zuverlässigkeit zu gewährleisten, muss bei den Überprüfungen entsprechender Vorhalt (Guard Banding) eingebaut werden. D.h. es wird schärfer geprüft, als es die Spezifikation vorsieht und damit, auf den ersten Blick, spezifikationskonformes Material verschrottet. Um diese qualitätsbedingten Verluste zu reduzieren, wird erheblicher Aufwand getrieben.

#### **3.2.2 Guard Banding: Messgenauigkeit**

Klarerweise ist die Messgenauigkeit der verwendeten Messgeräte beschränkt. Weiterhin muss bei der Beurteilung des Messfehlers der gesamte Messaufbau einschließlich des gerade geprüften Bauteils berücksichtigt werden. Die hier unter industriellen Bedingungen durchgeführten Prüfungen unterscheiden sich deutlich von einer Messung eines einzelnen Bauteils unter Laborbedingungen. Auch gegenüber dem bereits beschriebenen Verfahren PCM sind beim Funktionstest wesentlich ver-

änderte Bedingungen vorhanden. Neben den im automatischen Testsystem typischerweise integrierten Messgerät besteht der Messaufbau noch aus anderen Systemen des Testsystems (Relaismatrizen, Kabel, 4-Quadranten Quellen, Spannungsquellen, Sequencer etc.) und dem produktabhängigen Adapter (DUT<sup>49</sup>-Board). Hinzu kommt noch beim Test auf Waferebene die Kontaktierung der Einzeldevices auf der Scheibe über die Probecard.

### 3.2.3 Wafertest

Da die Bondpads aus Aluminium bestehen, muss das darauf immer vorhandene Aluminiumoxid durch die Probenadel durchstoßen werden. Das Bondpad selber muss beim Bonden der Bauteile während des Assemblyvorgangs noch intakt sein, da hier die Verbindung zur Außenwelt durchgeführt wird. Daher muss der Kontakt der Probenadel zum Bondpad von Nadelform über die Aufsetzposition bis zum Anpressdruck ein gut kontrollierter Prozess sein.

Abbildung 24 zeigt ein Beispiel für eine Probecard zur Kontaktierung von Bauteilen beim Funktionstest. Sichtbar ist hier eine Probecard in Blade-Technologie. Das bedeutet, die Nadeln werden auf einen entsprechenden Träger aufgelötet. Bei höheren Anforderungen an die Präzision werden aufwändiger zu fertigende und daher

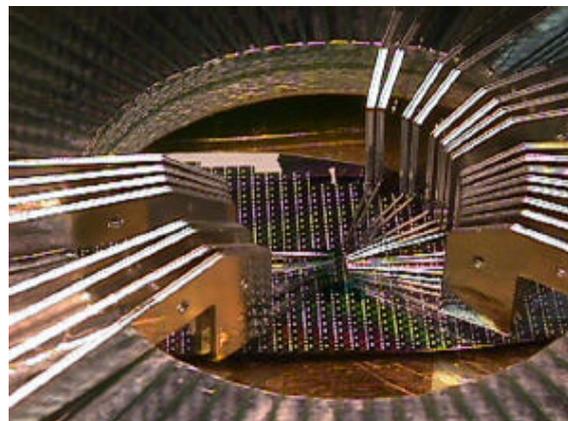


Abbildung 24 Probecard [APR07]

teurere Probecards in Epoxy-Technologie benutzt. Hier werden die Nadeln aufgeklebt. Die Grenzen dieser Technologie sind dann erreicht, wenn die notwendige Kontaktierung aus topologischen Gründen nicht mehr so möglich ist, weil mehrere Produkte zur selben Zeit kontaktiert werden müssen um Zeit zu sparen oder wenn die Kontaktpads in meh-

---

<sup>49</sup> Device Under Test: Bezeichnung des Bauteils, das sich gerade im Testaufbau für die elektrische Prüfung befindet.

ren Reihen angeordnet sind. Ebenfalls überschritten sind die Grenzen dieser Technologie wenn die Kapazität der Nadeln für die notwendigen Messungen zu groß ist. In diesen Fällen werden sogenannte V-Probes eingesetzt. Hier sind die Probenadeln gerade und werden durch präzise gefertigte Führungen aus nichtleitendem Material an das Bauteil gebracht. Eine unmittelbare mikroskopische Kontrolle des korrekten Sitzes der Nadeln und der Nadelabdrücke auf den Bondpads ist dann allerdings nicht mehr möglich.

#### 3.2.4 Final Part Test

Beim Test der gehäusten Bauelemente kommt noch das Handlingsystem und hier besonders die Kontaktiereinheit als Einflussfaktor hinzu. Abbildung 25 zeigt einen geöffneten Bauteilhandler. Die zu prüfenden Bauelemente werden im oberen (Input-)Magazin temperiert und dann nach Vereinzelung der Kontaktiereinheit zugeführt. Nach dem Signal an das Testsystem wird das DUT dann elektrisch getestet, und der Handler bekommt vom Testsystem das Prüfergebnis. Abhängig vom Ergebnis werden die getesteten Bauteile ins Output-Magazin sortiert.

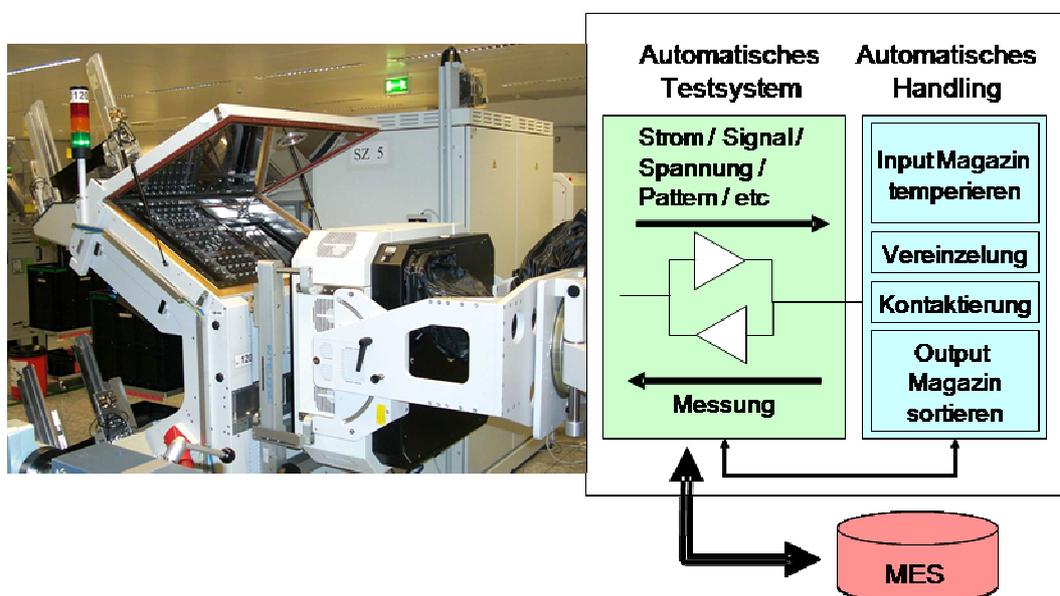


Abbildung 25 Final Part Testzelle [nach OEE02]

Weiterhin ist der Prüfling selbst zu berücksichtigen. Es darf an dieser Stelle nicht vergessen werden, dass das zu prüfende Bauteil ein komplexes System aus mindestens vielen tausend Einzelbauteilen darstellt das

mittels eines produktspezifischen Testprogramms in den jeweils für die Messung geeigneten Zustand versetzt wird.

Ersichtlich ist, dass eine Ableitung des Messfehlers über die Berechnung der Fehlerfortpflanzung hier nicht zum Ziel führt da, spätestens die durch das getestete Produkt selbst induzierten Messfehler nicht mit der notwendigen Genauigkeit vorliegen können.

Die aktuell benutzte Lösung besteht hier darin, über geeignete statistische Verfahren gekoppelt mit einer hinreichend oft wiederholten Prüfung (R&R Analyse) die Reproduzierbarkeit der Messungen festzustellen und aus der Schwankungsbreite auf den Gesamtmessfehler und damit auf den notwendigen Vorhalt bei der Bewertung der Messergebnisse zu schließen.

### **3.2.5 In-Strip-Test**

Das Handling der Bauelemente ist eine kontinuierlich größer werdende Herausforderung. Trotz der vorhandenen Toleranzen der Gehäuse müssen diese bei den Testtemperaturen zuverlässig und vor allem beschädigungsfrei transportiert, vereinzelt und kontaktiert werden. Hier ist zu berücksichtigen, dass insbesondere die Anschlusspins der Bauteile sehr empfindlich auf mechanische Belastungen reagieren. Jegliche Veränderungen an den Kontaktpins des Bauteils müssen unterbleiben. Dies gilt für alle benutzten Gehäuseformen, also beispielsweise auch für Gehäuse wie PLCC, die auf allen vier Seiten über Kontakte verfügen. Der Trend zu kleineren Gehäusen und geringerem Abstand der Kontakte erhöht die Anforderungen. Hinzu kommt der erhebliche Druck die Indexzeiten, also die Zeiten für den Wechsel der Bauteile in der Kontaktiereinheit, zu reduzieren.

Ein Ansatz zur Verbesserung der Nutzung und damit zur Reduzierung der Testkosten ist die gleichzeitige Kontaktierung von mehreren Bauelementen. Dieses als Multi - Site Testing bekannte Verfahren reduziert die pro Bauteil benötigte Indexzeit. Zudem erlaubt dieses den teilweise gleichzeitigen Test von Bauteilen. Der gleichzeitige Test hängt davon ab, dass das Testsystem entsprechend viele Ressourcen für den jeweiligen Test zur Verfügung hat. Problematisch ist der mechanische Aufwand an den Handlingsystemen und der Integrationsaufwand. Schließlich müssen

auch beim Multi – Site Test die Messwerte und vor allem die Prüfergebnisse jeweils korrekt dem Prüfling zugeordnet werden.

Eine neuere Entwicklung bei der Funktionsprüfung von Halbleiterbauelementen ist der sogenannte In-Strip-Test. Hier wird der klassische Ablauf aus Waferfertigung, Funktionstest auf der Scheibe, Assembly und dann Funktionstest beim fertigen Bauelement durchmischt. Der erste Test der Bauteile findet bereits während des Assemblyvorgangs statt.

Der Funktionstest findet unmittelbar vor der Vereinzelung der Bauteile statt. Die Bauteile sind gebondet und mit Mold Compound umschlossen. Sie befinden sich noch auf dem Lead Frame. Die Verbindung der Kontakte wird mechanisch getrennt, die Verbindungen zum Gehäuse sind noch intakt. Eine Übersicht dieser Veränderungen ist in Abbildung 26 dargestellt. Ein Test zu diesem Zeitpunkt hat einige wesentliche Vorteile:

- Das Handling der Komponenten beim Test ist vergleichsweise einfach, da die Positionen der Bauelemente sehr genau bekannt sind.
- Gegebenenfalls für Hoch- oder Tieftemperaturprüfungen notwendige Wartezeiten fallen durch die gleichzeitige Temperaturänderung einer größeren Anzahl Devices weniger ins Gewicht.

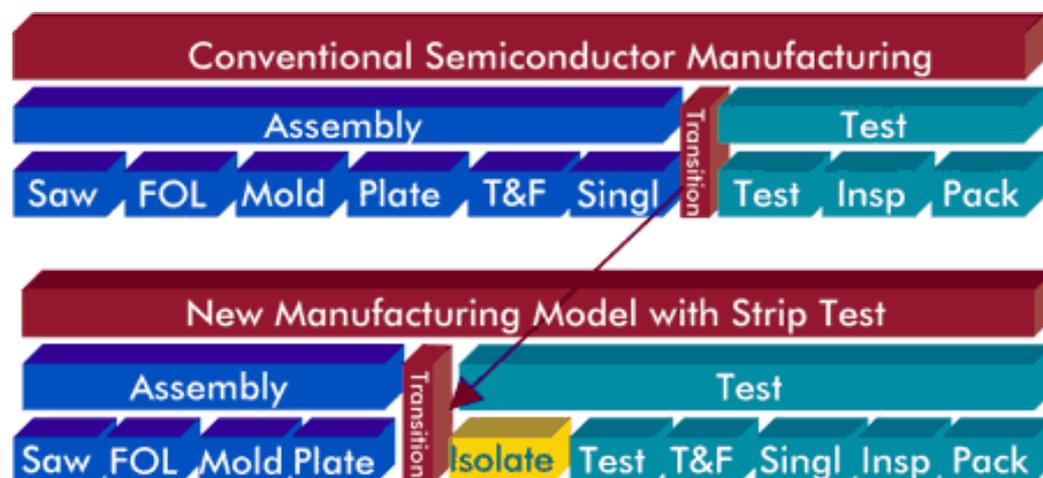


Abbildung 26 Arbeitsablauf In-Strip-Test [AMK08]

Dass sich dieser Ablauf noch nicht durchgesetzt hat, hängt mit den leider ebenfalls erheblichen Nachteilen des Verfahrens zusammen:

- Nach dem Test findet noch eine mechanische Bearbeitung, nämlich das Trimmen der Kontaktbeine statt. Für Produkte mit einer sehr hohen Anforderung an die Qualität ist damit ein weiterer (100%) Test unabwendbar.

- Test und Assembly werden häufig durch unterschiedliche Unternehmen durchgeführt. Beim assemblierenden Unternehmen ist häufig nicht das notwendige Know-How für den Test in ausreichendem Masse vorhanden.
- Der Assemblyprozess wird wegen der Arbeitskosten häufig durch Lohnveredler im Fernen Osten durchgeführt. Für den In-Strip-Test sind Investitionen in Testsysteme und Handlingsysteme notwendig. Abbildung 27 zeigt ein aktuelles In-Strip-Handlingsystem.
- Die Fertigungsleitsysteme sind zwischen dem Testbereich und dem Assemblybereich getrennt. Typischerweise gibt es keine automatisierbaren Verbindungen der Systeme. Für den In-Strip-Test muss nun entweder das Fertigungsleitsystem des Final-Tests zumindest punktuell in der Assembly Fabrik installiert oder das Assemblysystem um die notwendigen Funktionalitäten erweitert werden.



Abbildung 27 In-Strip-Handlingsystem, [MUL08]

Der wohl wesentliche Vorteil des In-Strip Tests ist jedoch in der häufig nur hier noch möglichen Rückverfolgbarkeit der Produkte zu suchen. Im Falle von mehreren Produkten in einem Gehäuse, die nicht alle über eine auslesbare Seriennummer verfügen, ist die Position auf dem (nummerierten) Leadframe die Verbindung zum Ursprung der hier positionierten Bauteile. An dieser Stelle sind die fertigen Bauteile noch un-

terscheidbar, so dass hier der Rückschluss auf die Chargennummer, die Scheibenummer sowie die Position auf der einzelnen Scheibe für alle Produkte im Package identifizierbar ist. Anschließend ist das nicht mehr der Fall.

Es sollte an dieser Stelle jedoch deutlich gemacht werden, dass eine derartige Integration sehr erhebliche finanzielle und personelle Aufwände voraussetzt. Insbesondere die Rückverfolgbarkeit im Assemblyprozess ist sehr schwierig zu etablieren. Ursache hierfür ist, dass der Assemblyprozess in der Prozesskette der Halbleiterfertigung zu den weni-

ger automatisch kontrollierten Abläufen gehört. Die in einer Halbleiterfertigung praktisch immer vorhandene weitgehende Durchdringung der Fertigung mit Datenverarbeitungssystemen ist hier typischerweise noch nicht so weit fortgeschritten.

Für die Auswertung der Fertigungsdaten ist dieses ein weiteres Problem, da aufgrund der mangelhaften Datennahme und Datenkommunikation die Einflüsse durch das Assembly schwer zu analysieren sind.

#### **3.2.6 Guard Banding: Temperaturgang**

Der Temperaturgang eines Halbleiterbauelementes ist ein in der Theorie gut verstandener Vorgang. Entsprechend kann ein gemessener Parameter eines einzelnen Bauelements bei einer Temperatur vergleichsweise einfach für eine andere Temperatur berechnet werden. Der minimal notwendige Vorhalt bei der Bewertung einer Messung in Bezug auf die Spezifikation dieses Parameters ergibt sich dann aus den spezifizierten Betriebstemperaturgrenzen sowie der Messtemperatur und dem Messwert.

Leider ist die industrielle Realität auch hier wiederum deutlich komplexer als nach dem ersten Anschein zu vermuten ist. Im Allgemeinen sind gemessene Werte wie Versorgungsströme eine Überlagerung von mehreren Effekten, die üblicherweise einen unterschiedlichen Temperaturgang aufweisen. Da die Einzelwerte der verschiedenen Effekte nicht bekannt sein können, ist eine hinreichend genaue Vorhersage des Temperaturgangs dieses Parameters normalerweise nicht möglich. Wegen der typischerweise exponentiellen Zusammenhänge lassen sich auf dem Wege der elektrischen Simulation der Produkte auch keine hinreichend genauen Voraussagen machen.

Zudem wird das Bauteil während des Tests natürlich betrieben. Die hier notwendigerweise anfallende Verlustleistung führt zur Erwärmung einzelner Bereiche des Siliziumkristalls. Diese inhomogene Temperaturverteilung ist während der Messung nicht bekannt.

Der Ausweg besteht darin, das Bauteil beim Serienstart möglichst umfassend zu charakterisieren und daraus die notwendigen Vorhalte für die Serienproduktion zu ermitteln. Implizit wird hierbei vorausgesetzt, dass sich die Fertigung nicht für das Bauteil relevant verändert. Wiede-

rum ist nicht sichergestellt, dass die Abhängigkeiten linear sind, daher können sich kleine Schwankungen in der Fertigung hier durchaus negativ auswirken. Entsprechend kann dieses bei kritischen Bauteilen auch dauerhaft zu der Notwendigkeit einer Prüfung bei verschiedenen Temperaturen führen. Dieses vervielfacht offensichtlich den Aufwand und die Datenmenge.

### 3.2.7 Einordnung Funktionstest in den Gesamtfertigungsprozess

In der Vergangenheit war die Kostenverteilung Fertigung (Frontend) zu den Kosten für die nachfolgenden Prozessschritte ganz eindeutig durch die Kosten des Waferprozesses definiert, so dass Aufwendungen und Forschungen für die Optimierung der Ausbeuten und der Fertigungskosten sich primär auf die Waferfabrikation bezogen.

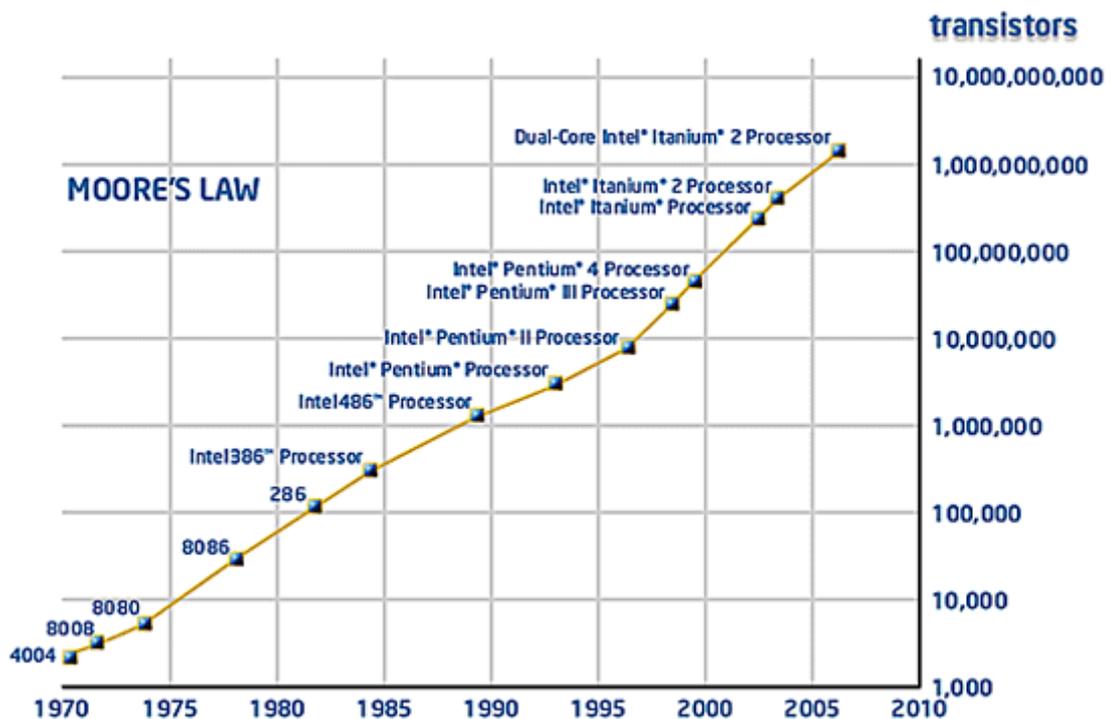


Abbildung 28 Moore's Law mit aktuellen Produkten [IN07B]

Durch die sehr erfolgreiche Weiterentwicklung der Halbleiterfertigungen weltweit ist die Anzahl der Funktionen auf einem Halbleiterbauelement dramatisch gestiegen. Nach wie vor verdoppelt sich die Anzahl von Transistoren auf einem Schaltkreis alle 18 bis 24 Monate [vgl. MOO65].

Eine logarithmierte Darstellung dieser Entwicklung zeigt Abbildung 28. Die damit einhergehenden dramatischen Reduktionen der Fertigungskosten pro Transistor können jedoch beim Test der produzierten Bauteile nicht in gleicher Weise erreicht werden. Klarerweise wird auch hier durch verbesserte Prüfverfahren für existierende Produkte und verbesserte Testfähigkeit für neue Produkte die Effizienz verbessert.

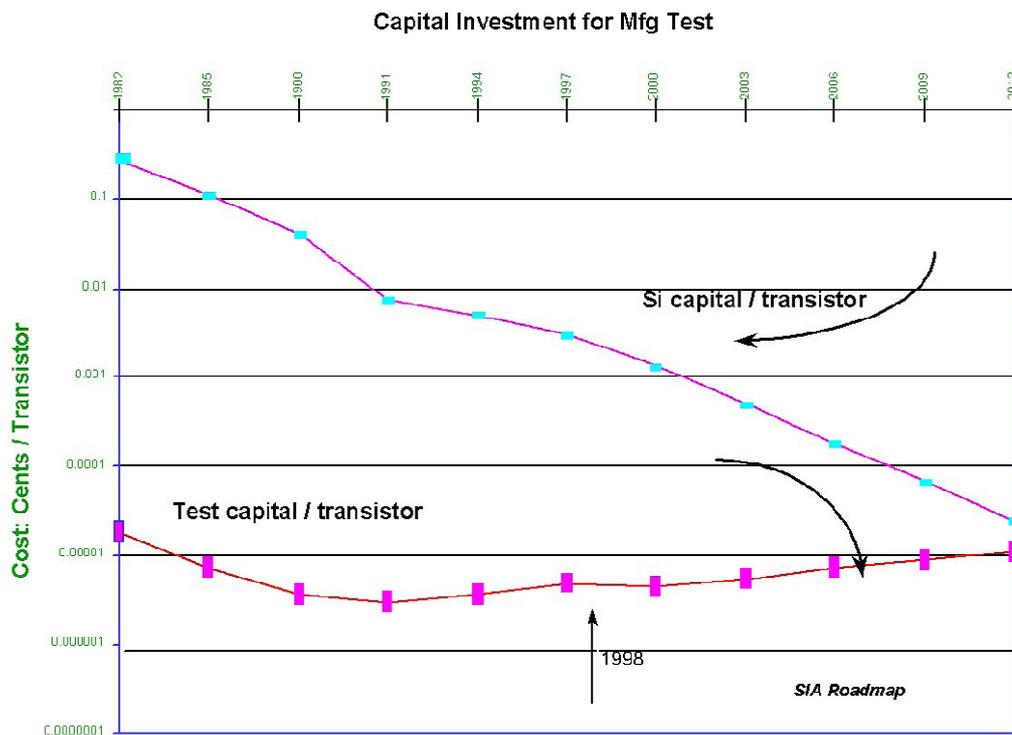


Abbildung 29 Trends bei Fertigungs- und Testkosten [IN99]

Entsprechende Skaleneffekte wie bei der Halbleiterfertigung, wo aus einer Reduzierung der Strukturbreite eine direkte sprunghafte Erhöhung der pro Flächeneinheit produzierten aktiven Bauelemente resultiert, gibt es beim Funktionstest nicht. Basierend auf einer Prognose von der SIA<sup>50</sup> wird der Break Even für Digitalschaltungen 2020 erreicht. Abbildung 29 zeigt, basierend auf Zahlen der Firma Intel, die Gegenüberstellung der Testkosten und der Fertigungskosten pro Transistor über die Zeit. Eine Fortschreibung der hier gezeigten Kurven legt den Break Even bereits vor dem Jahr 2015 nahe.

<sup>50</sup> Semiconductor Industry Association

Gegenüber den Digitalschaltungen primär für die Konsumgüterindustrie steigt der Testaufwand durch die Qualitätsanforderungen der Kunden (Stichwort: 0-Fehler Strategie) –je nach Marktsegment– sehr stark an. Daher verschiebt sich die Kostenverteilung bei der Halbleiterproduktion zu noch größeren Anteilen bei den Testkosten was dort unmittelbar den Kostendruck weiter verschärft.

Hinzu kommt, dass bei neueren Produkten wie den bereits beschriebenen kointegrierten Produkten oder Multi - Chip Lösungen ein elektrischer Test nach wie vor notwendig aber nicht mehr hinreichend ist. Neben den unveränderten Aufwendungen für den Funktionstest sind nun auch noch Druck- oder Beschleunigungsmessungen bei den gehäuteten Bauelementen notwendig.

### ***3.3 Traceability***

Schon aus Haftungsgründen ist es notwendig, die bei der Fertigung anfallenden Daten zu archivieren. Kundenanforderungen gehen häufig noch sehr weit über die gesetzlichen Anforderungen hinaus. Hier ist das Ziel, bei eventuellen Problemen mit einzelnen Produkten im Feld herausfinden zu können, ob mit dem betreffenden Fertigungslos irgendwelche Auffälligkeiten verbunden sind. Gegebenenfalls muss ein Rückruf ausgelöst werden, der sowohl das Image des Herstellers beschädigt als auch kommerziell sehr nachteilig ist.

Auch für diesen Rückruf ist eine möglichst detaillierte Rückverfolgbarkeit der Ware in der Fertigung von erheblichem Vorteil. Je präziser die Rückverfolgbarkeit des Materials ist und je kleiner die Einheiten sind, die rückverfolgbar sind, desto besser lässt sich ein Problem eingrenzen und desto geringer fällt der wirtschaftliche Schaden aus, wenn doch einmal Bauteile vom Kunden oder vom Endkunden zurückgerufen werden müssen. Auch wenn der Kunde hier keine besonderen Anforderungen stellt, ist der Hersteller gut beraten, die möglichst präzise Rückverfolgbarkeit zu sichern, da er damit seine Haftung erheblich begrenzen kann, auch wenn in den nachfolgenden Prozessschritten durch die folgende Supply Chain diese nicht durchgehalten wird.

### ***3.4 Aufbewahrungsfristen***

Die Anforderung der Rückverfolgbarkeit ist primär eine Archivfunktion. Damit sind die Anforderungen an die Analysierbarkeit der Daten überschaubar. Allerdings sind die geforderten Aufbewahrungsfristen sehr lang. Aktuell resultiert die maximal geforderte Zeitspanne aus dem Produkthaftpflichtgesetz. Die hier geforderten Aufbewahrungsfristen betragen mindestens 10 Jahre. Da sich diese 10 Jahre allerdings auf die Auslieferung des Endproduktes an den Kunden beziehen, haben sich 15 Jahre als Industriestandard für Aufbewahrungsfristen aktuell durchgesetzt.

Angesichts der Entwicklungsgeschwindigkeit in der Informationstechnologie kann hier nicht auf etablierte Verfahren zurückgegriffen werden. Es gibt keine Datenträger, die eine akzeptable Kapazität bereitstellen und für die es Erfahrungen über derartig lange Zeiträume gibt. Hinzu kommt die zu berücksichtigende Verfügbarkeit der Lesesysteme. Beispielsweise wurden in den 90'er Jahren des vergangenen Jahrhunderts viele Archivdaten auf MO-Disks (Magneto-Optical) geschrieben, da diese eine zugesicherte Lebensdauer von 30 Jahren hatten. Diese Technologie ist zwischenzeitlich überholt, so dass keine Lesesysteme mehr verfügbar sind.

Entsprechend werden hier mit hohem Aufwand die Archive regelmäßig auf neuere Medien umkopiert. Zudem muss natürlich sichergestellt werden, dass die Daten von den Sicherungsmedien nicht nur formal sondern auch inhaltlich lesbar sind. Daher müssen die Aufzeichnungsformate auf den Archivmedien voll dokumentiert sein. Weiterhin ist es notwendig, dass die Dateiformate archivgeeignet sind. Klarerweise verbieten sich hier jegliche nicht vollständig dokumentierten proprietären Dateiformate.

Wegen des rasanten Wachstums der Festplattenkapazitäten ist es aktuell ein alternativer Ansatz, die Daten alle online zu halten und nicht zu archivieren oder spezielle Archivsysteme zu nutzen, die die Daten ebenfalls auf Standard Festplatten speichern. Damit steigt natürlich der Aufwand für den Betrieb des Systems. Zudem verlängert sich gegebenenfalls die Zugriffszeit auf die benötigten aktuellen Daten.

In Kapitel 2 habe ich die spezielle Situation der Halbleiterindustrie beschrieben. Nach der Beschreibung der Qualitätsanforderungen des Marktes und den sich daraus ergebenden Konsequenzen in diesem Kapitel wird im Folgenden beschrieben, welche Abläufe und Werkzeuge heute eingesetzt werden, um die Qualitätsanforderungen zu erfüllen ohne dabei zu große Kostennachteile in Kauf nehmen zu müssen.

# 4 Stand der Technik

Die beschriebenen Prüfverfahren für die gefertigten Produkte sind zur Erreichung der Qualitätsanforderung der Kunden notwendig aber nicht hinreichend. Zudem sind Ausbeuteverluste über das in Kapitel 2.1 beschriebene Maß hinaus in der Kalkulation der Produkte typischerweise nicht berücksichtigt. Damit reduzieren diese Ausbeuteverluste den für die notwendigen Investitionen notwendigen Ertrag und gefährden damit die Wettbewerbsfähigkeit des Unternehmens.

Durch die Komplexität des Herstellungsprozesses ist ein deduktierbarer Zusammenhang der Prüfergebnisse mit Fertigungsparametern im Allgemeinen nicht herstellbar. Die typische Vorgehensweise besteht daher darin, im Rahmen einer Prozessentwicklung einen Herstellungsprozess zu entwickeln und im Folgenden die Randbedingungen möglichst konstant zu halten und derart eine stabile und reproduzierbare Fertigung zu garantieren (Prozessregelung).

Zudem werden über verschiedene Verfahren Abweichungen identifiziert und die betroffenen Produkte verworfen oder nachgearbeitet.

## 4.1 „Short Loop“

In der Fertigung werden Messungen durchgeführt, um Prozessparameter so regeln zu können, dass diese möglichst in der Mitte des Prozessfensters liegen.

### 4.1.1 Statistical Process Control

Die Statistische Prozessregelung (SPC<sup>51</sup>) ist ein etabliertes Verfahren, um Prozessabweichungen zu erkennen und zu analysieren.

Dazu wird abhängig von der Fragestellung sowie den technischen Möglichkeiten und dem Grad der Automatisierung der jeweilige Messwert oder ein Mittelwert aus mehreren Messwerten in einem SPC-Chart dargestellt. Abweichungen von den vorherigen Schritten an derselben Anlage und mit demselben Rezept werden so unmittelbar nach dem Abschluss

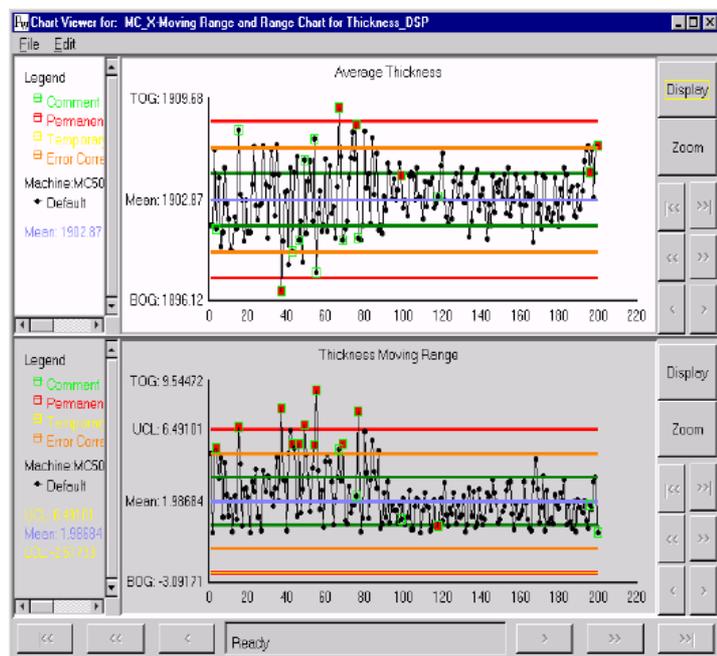


Abbildung 30 SPC Werkzeug [ADV2000]

des Fertigungsschrittes erkennbar. Das Verfahren wird eingesetzt, um dem Maschinenbediener eine direkte Rückkopplung über den Erfolg oder Misserfolg der aktuell durchgeführten Aufgabe zu geben. Handelt es sich nicht um eine Fehlmessung, werden das Material und/oder die Maschine für weitere Schritte in der Produktion gesperrt.

Weiteres Einsatzgebiet von SPC ist die kontinuierliche Überwachung der Prozessgüte durch den verantwortlichen Prozessingenieur. Hierzu analy-

<sup>51</sup> Statistical Process Control

siert der Verantwortliche in regelmäßigen Abständen den Verlauf der Kurve und trifft gegebenenfalls Gegenmaßnahmen, um einen Trend der Messwerte umzukehren oder eine zu große Schwankung der Werte zu reduzieren. Ein Beispiel für ein SPC Werkzeug zeigt Abbildung 30.

SPC kann nur dazu dienen, den vorher festgelegten Zielwert möglichst gleichmäßig zu reproduzieren. Sie kann helfen, die Stabilität des Prozesses zu erhöhen und die Schwankungen zu minimieren. Als Werkzeug für die Suche nach Fehlerursachen oder für ein besseres Verständnis der Prozesse sowie deren Verbesserung ist sie prinzipbedingt wenig geeignet. Weiterhin ist SPC nicht geeignet, mit Abhängigkeiten von mehreren Parametern umzugehen.

Schließlich ergeben sich erhebliche Probleme, wenn verschiedene Produkte mit verschiedenen Fertigungsrezepten in der Fabrik gefertigt werden. In diesem Falle sind die Messwerte mit fremden systematischen Effekten „verunreinigt“, und die Schwankungen der Messwerte steigen. Unter Umständen wird die Regelung dann kontraproduktiv, da auf Basis von systematisch abweichenden Messwerten unangemessene Gegenmaßnahmen getroffen werden.

Eine Einengung der Betrachtungsweise auf das jeweils identische Produkt und Rezept schmälert die statistische Basis stark und erschwert insbesondere bei Produkten, die in geringen Stückzahlen laufen, sowie beim Anlauf von Neuprodukten die Prozessregelung.

### 4.1.2 Statistical Equipment Control

Die statistische Regelung der Fertigungsanlagen (SEC<sup>52</sup>) entspricht von der Darstellung und den benutzten Verfahren der SPC. Im Unterschied zur SPC werden hier aber keine Parameter beim gefertigten Material gemessen sondern Parameter der Produktionsmaschinen wie Druck, Gasflüsse oder ähnliches. Auch diese Methode muss im Verbund mit den beschriebenen Verfahren zur Stabilisierung der Fertigung gesehen werden. Die Maschinen sind derartig komplex, dass nicht alle gegebenenfalls auftretenden Seiteneffekte bekannt sind. Entsprechend werden von

---

<sup>52</sup> Statistical Equipment Control

der internen Regelung der Maschine auch nicht alle relevanten Parameter überwacht geschweige denn geregelt.

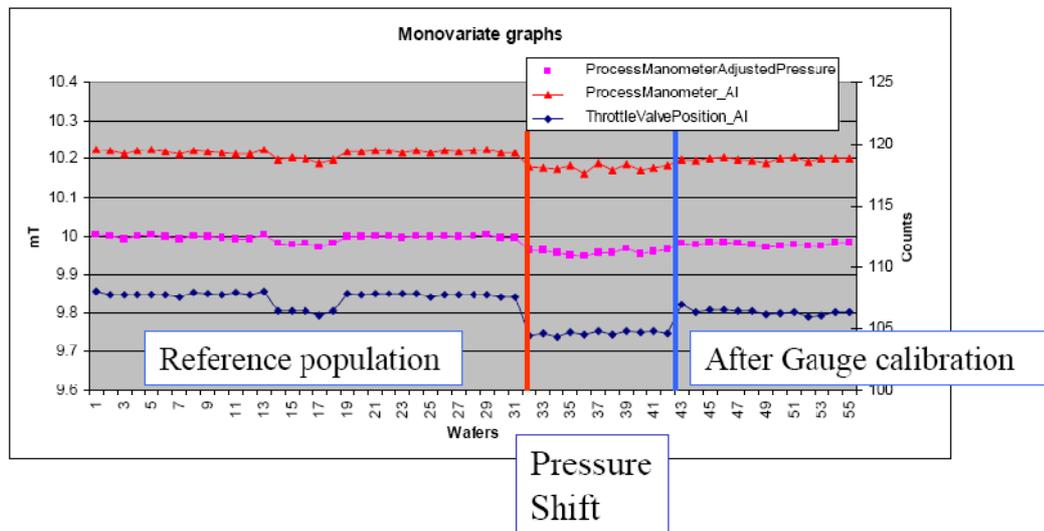


Abbildung 31 Gasdruck innerhalb einer Maschine [HERK2004]

SEC wird typischerweise im Zuge von aufgetretenen Fehlern in den Fertigungen aufgesetzt. Ein sehr signifikantes Beispiel für die Komplexität findet sich in Abbildung 31 und Abbildung 32. In Abbildung 31 werden bei den beiden oberen Kurven einige der gemessenen Drücke innerhalb der Fertigungsanlage dargestellt. Die untere Kurve stellt eine Ventilposition innerhalb der Maschine dar. Eine Betrachtung der Drücke zeigt zunächst kein deutliches Ergebnis. Erst nach der Hauptachsentransformation in Abbildung 32 ist die Grenzverletzung deutlich sichtbar und ist auch per SEC überwachbar.

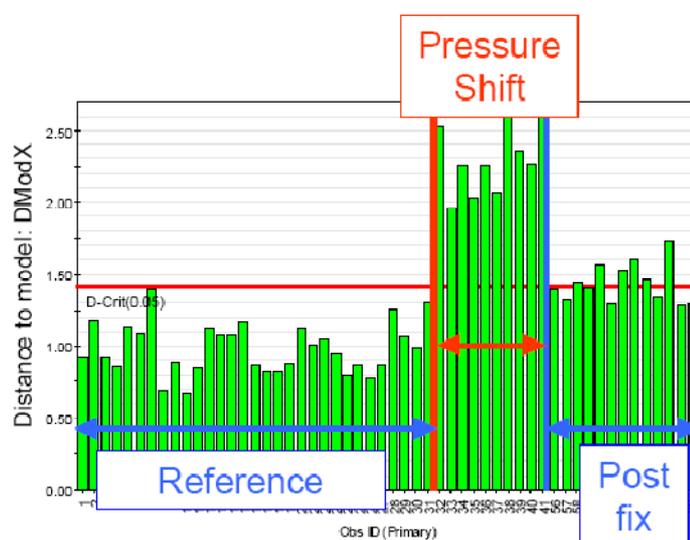


Abbildung 32 Messwerte auf dem Material [HERK04]

An diesem Beispiel wird deutlich, dass erfolgreiche SEC-Regelungen sehr häufig ein Ergebnis von zuvor aufgetretenen Fehlern und deren erfolgreicher Analyse sind.

Wie bei der bereits beschriebenen SPC ist auch die SEC ein symptomatisches Verfahren. Man versucht, anhand von gemessenen Parametern die Fertigung möglichst konstant zu halten, und die bei SPC genannten Einschränkungen gelten hier in gleichem Maße.

### 4.1.3 Fault Detection and Classification

Eine weitere in jüngerer Vergangenheit hinzugekommene Technik ist die Fehlererkennung und Klassifizierung (FDC<sup>53</sup>). Final handelt es sich hier um eine Erweiterung und Verallgemeinerung der SEC. Ziel der unter dem Kürzel FDC zusammengefassten Verfahren ist die Erkennung von jeglicher Veränderung an den Fertigungsverfahren und -maschinen. Anders als bei SPC wird hier nicht gewartet, bis sich an dem produzierten Material eine Veränderung zeigt. Es ist nicht einmal relevant, ob sich überhaupt eine identifizierbare Veränderung am Material zeigt.

Diese Veränderungen werden dann so schnell wie möglich rückgängig gemacht. Schließlich ist das bereits beschriebene Verfahren optimiert worden: Ein einmal als funktionierend erkanntes Setup wird möglichst konstant gehalten. Dahinter

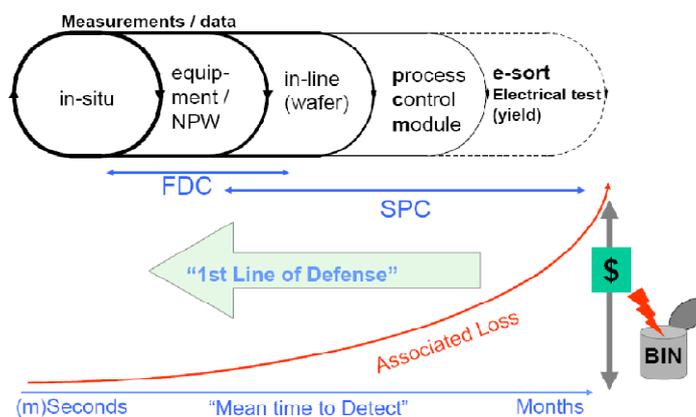


Abbildung 33 Ziel FDC: Frühe Erkennung von Verlusten[CROL05]

steckt die These, dass die Verfahren so weit optimiert sind, dass jegliche Veränderung mit sehr hoher Wahrscheinlichkeit eine Verschlechterung darstellt. Einer Evolution der Verfahren wird hier bewusst kein Raum gegeben.

In Abbildung 33 ist die Abhängigkeit der Folgekosten einer Prozessabweichung von dem zeitlichen Abstand der Entdeckung desselben nach

<sup>53</sup> Fault Detection & Classification, manchmal auch Fault Detection & Control

seinem Auftreten dargestellt. Ebenfalls dargestellt ist für FDC und SPC die Erkennungsgeschwindigkeit von Prozessabweichungen. Hier ist direkt erkennbar, warum FDC in den heutigen Fabriken trotz der beschriebenen Einschränkungen ein erfolgreiches Konzept ist, da aufgetretene Fehler vergleichsweise viel schneller gefunden werden können.

#### 4.1.4 Advanced Process Control

Bei den unter 4.1.1 –4.1.3 beschriebenen Verfahren ist die Rückkopplungsschleife auf einen Prozessschritt, gegebenenfalls mit nachgeschalteter Messung an einem anderen Gerät begrenzt. Zudem sind die Regel Schleifen normalerweise nicht geschlossen, da sich eine veränderte Regelgröße erst auf folgende Fertigungslose auswirkt. Dieses –auch als Run to Run (R2R) Control bezeichnete und in Abbildung 34 skizzierte– Vorgehen ist gut brauchbar in einer Fabrik, die große Stückzahlen nahezu identischer Produkte fertigt.

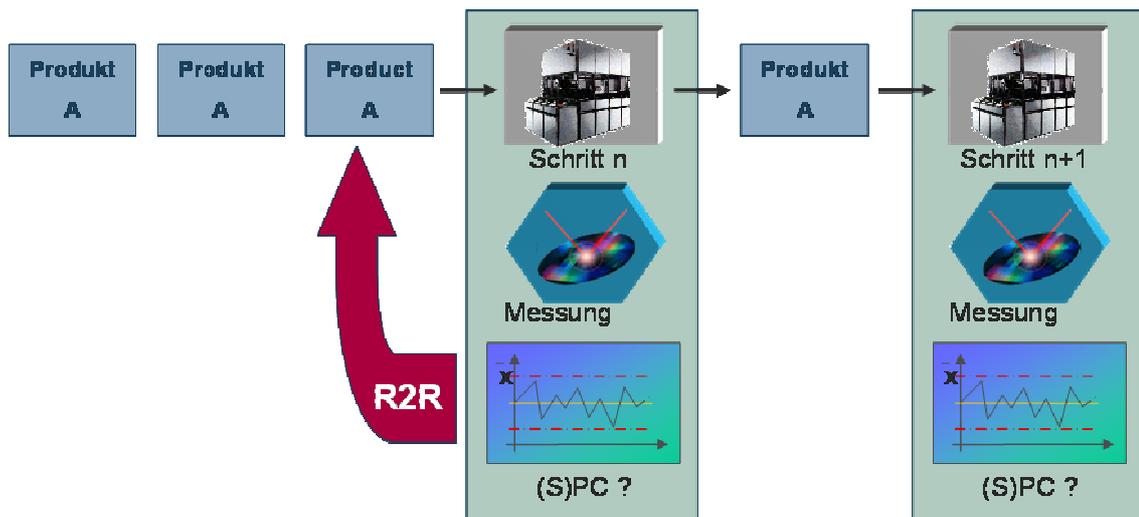


Abbildung 34 Prozessregelung als offene Rückkopplung[MOBE05]

Mit steigender Anzahl unterschiedlicher Produkte sinkt der Erfolg der Verfahren stark ab, da sich die Vorgeschichte des Materials und die Reaktion auf die Prozesse ebenfalls unterscheiden. Weiterhin wird das gerade produzierte Material nicht mehr beeinflusst, so dass ein Gegensteuern nicht möglich ist. Dieses ist besonders problematisch, wenn sich die Fertigung an der Grenze des technisch Machbaren befindet. Entsprechend wurde beginnend bei den Leading-Edge Fabriken in der Fototechnik eine prozessschrittübergreifende Prozessregelung entwickelt.

Schlüsselursache war hier die Schwankung der Strukturbreiten des durch Belacken, Belichten und Entwickeln aufgetragenen Fotolacks. Die nachfolgenden Prozessschritte, hier insbesondere das Ätzen, erhöhen mit den ihnen innewohnenden Toleranzen die Fertigungsungenauigkeit weiter.

Daher wird hier das Rezept des nachfolgenden Prozessschrittes abhängig von den Messwerten des vorangegangenen Schrittes modifiziert. Abbildung 35 zeigt eine Übersicht über ein solches System. Eine Übersicht über die etablierten Verfahren und deren Probleme hat der Autor in [MWK04] und [MOBE05] dargestellt.

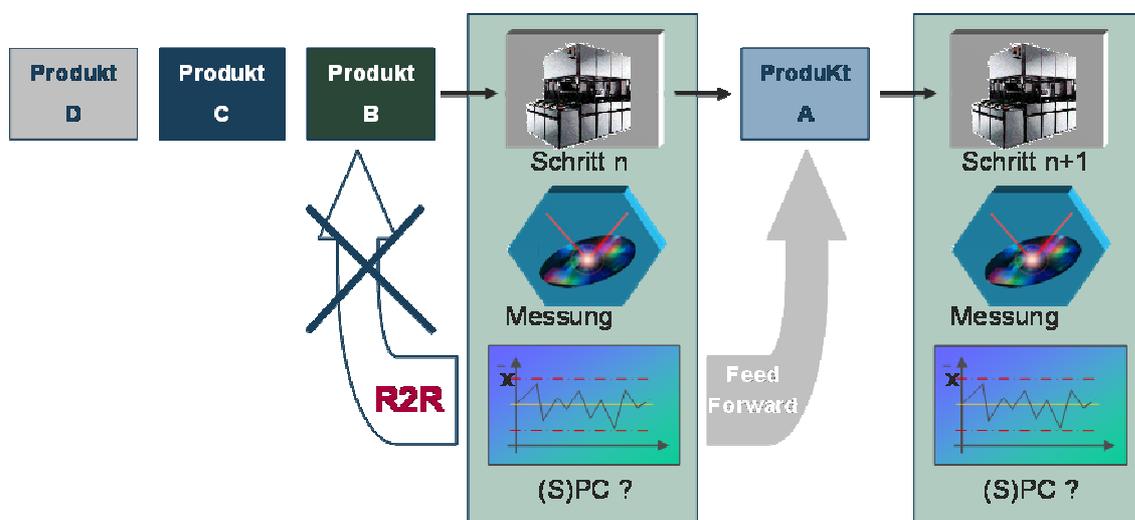


Abbildung 35 Allgemeine APC<sup>54</sup> Lösung [MOBE05]

Im Grundsatz scheint dieses Vorgehen ein einfacher und naheliegender Ansatz zu sein. Im Detail ist das Verfahren wegen der vielen zu beachtenden Nebenbedingungen sehr kompliziert und damit auch teuer zu implementieren. In den Fertigungen, die an der Grenze des aktuell technisch Machbaren operieren, gibt es trotz des erheblichen Aufwandes für die Einführung keine Alternative. Insbesondere in der Lithografie ist APC unterhalb von  $0,25\mu\text{m}$  als Standard zu betrachten. Für Fabriken bei deutlich größeren Strukturbreiten kann die Einführung von APC aber ebenfalls zum Thema werden. Sind nämlich die aktuell zu fertigenden Strukturen nicht an der Grenze des aktuell Machbaren, wohl aber an der Grenze der aktuell eingesetzten Fertigungsmaschinen, ist im Allgemei-

<sup>54</sup> APC: Advanced Process Control

nen die Implementierung einer APC-Lösung gegenüber der Neuanschaffung von entsprechendem Equipment die effizientere Lösung.

Die Implementierung einer APC-Lösung setzt anders als bei den bisher beschriebenen Lösungen eine mathematisch abbildbare Funktion der Zusammenhänge voraus. Dies gelingt mit erheblichem Aufwand durchaus[z.B. ADV00], allerdings nur beschränkt auf zwei von vielen Hundert Prozessschritten und auch nur dann, wenn produktspezifische Besonderheiten die Regelung nicht stören.

## ***4.2 Automatisch selektierende Verfahren***

Die bisher beschriebenen Verfahren finden direkt nach dem zugehörigen Fertigungsschritt oder sogar in diesen integriert ihre Anwendung. Mit der Ausnahme APC geben die Verfahren Hinweise auf notwendige weitere Untersuchungen oder allgemeiner auf eine Sonderbehandlung des Materials. Abweichend davon finden die folgenden Verfahren im Anschluss an mindestens eine 100%-Prüfung des Materials statt und selektieren oder klassifizieren das vorliegende Material automatisch.

### **4.2.1 Statistical Yield Analysis / Statistical Bin Analysis**

Die Statistical Bin<sup>55</sup> Analysis (SBA) und die Statistical Yield Analysis (SYA) sind vom Automotive Electronic Council definierte Verfahren [AECQ002]. Im Kern geht es, ganz analog zum SPC-Verfahren, darum, Veränderungen an den hergestellten Produkten zu erkennen und entsprechende Maßnahmen zu ergreifen. Basis für dieses Verfahren sind die Klassifikationsergebnisse vom Funktionstest der Bauteile. Hier wird ab einer Abweichung von drei Sigma vom Mittelwert der letzten sechs Fertigungslose dieses Produktes eine technische Analyse der Ursachen für die Abweichung verlangt. Ab einer Abweichung von vier Sigma von diesem Mittelwert wird eine Kundeninformation vor der Lieferung der Teile vorgeschrieben.

---

<sup>55</sup> Die Klassifikation von Bauteilen beim Test wird in der Halbleiterindustrie generell als Binning bezeichnet.

Es ist hier wichtig zu berücksichtigen, dass die Definition von einer Kundenvereinigung stammt. Entsprechend ist diese aus Kundensicht formuliert und damit nicht notwendigerweise im Interesse der Hersteller.

Auch die Güte der vorliegenden Beschreibung hält hohen Anforderungen kaum stand. So wird zwar die zulässige Abweichung beschrieben, die zugrunde liegende Verteilungsfunktion hingegen wird nicht erwähnt. Damit obliegt die Entscheidung über die angewendete Verteilungsfunktion dem Anwender und behindert damit die Vergleichbarkeit der Ergebnisse und damit auch den Erfolg des Einsatzes.

Entsprechend werden diese Verfahren in der Industrie unterschiedlich eingesetzt. Zwar kann für jede beliebige Verteilungsform eine Standardabweichung (Theorem von Tschebyscheff) definiert werden, aber in der angewandten Realität wird praktisch immer vereinfachend die Normalverteilung vorausgesetzt. Da es sich sowohl bei SBA als auch bei SYA um zweiseitig begrenzte Ergebnisräume handelt und die Mittelwerte typischerweise nahe an den Grenzen liegen, kann strenggenommen die Berechnung der Standardabweichung schon aufgrund der nicht gegebenen Voraussetzungen nicht so vorgenommen werden.

Final ist das Verfahren nur beschränkt wirksam, da es sich nur auf historische Daten bezieht und keine Vergleichswerte anderer Hersteller oder Fabriken berücksichtigt. Wenn der Hersteller seine Prozesse unter Kontrolle hat, wird entsprechend die Standardabweichung kleiner. Das führt dann dazu, dass schon aus statistischen Gründen mehr als 10% der Ware einer technischen Analyse unterzogen werden müssen, obwohl es dafür keine technische Ursache gibt. Die Ursache für dieses scheinbare Missverhältnis zwischen der Forderung „3 Sigma“ und der obengenannten Ausfallrate liegt darin begründet, dass die Untersuchung sich auch auf Einzelwafer bezieht. Angehalten wird aber immer das ganze Los (25 Wafer), sobald auch nur einer der Wafer betroffen ist.

Hat ein Hersteller seine Prozesse schlechter unter Kontrolle, ist aufgrund der verborgenen systematischen Einflüsse die berechnete Standardabweichung größer und damit sowohl die Entdeckungswahrscheinlichkeit als auch der zusätzliche Aufwand für den Hersteller geringer.

## 4.2.2 Part Average Test

Wie die SBA / SYA ist auch der sog. Part Average Test (PAT) durch das AEC definiert. Ähnlich den oben genannten Verfahren werden auch hier wieder Ausreißer mittels statistischer Verfahren identifiziert. Abweichend von dem unter 4.2.1 beschriebenen Verfahren, wird hier die Ausreißer Selektion auf Bauteilebene durchgeführt. Damit ist dieses Verfahren wesentlich präziser aber auch aufwändiger zu handhaben. Es werden relevante Tests definiert und alle Bauteile, deren Messwerte mehr als drei Sigma vom Mittelwert dieses Tests abweichen, werden verworfen. Da sich diese Bewertung auf einzelne Tests von einzelnen Bauelementen bezieht, ist die Aussage hier wesentlich klarer. Auch ist hier die implizite Annahme der Normalverteilung für die Einzelmesswerte wesentlich besser gerechtfertigt als bei der Bewertung ganzer Lose.

Auch dieses Verfahren dient weder der Analyse noch der nachhaltigen Beseitigung von Normabweichungen. Das heißt, auch der Part Average Test hilft nicht bei der Verbesserung der Fertigungsprozesse. Er kann allerdings erfolgreich angewandt werden, um die

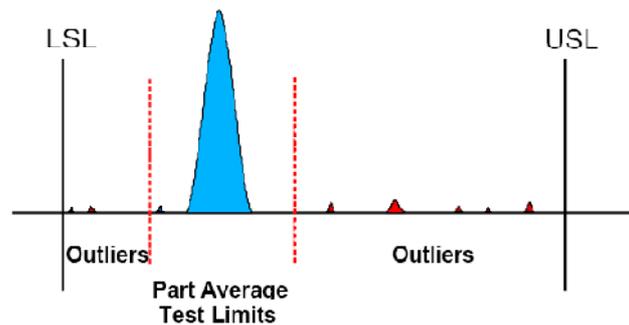


Abbildung 36 Darstellung von PAT Grenzen [AECQ001]

Ausfallzahlen der Bauteile beim Kunden zu reduzieren, indem Ausreißer nicht ausgeliefert werden, auch wenn diese vollständig der Spezifikation entsprechen. Klar ist, dass ein derartiges Verfahren nur eingesetzt wird, wo der Kunde extrem hohe Qualitätsanforderungen hat. In Abbildung 36 ist das Vorgehen skizziert. LSL (Lower Spec Limit) und USL (Upper Spec Limit) stellen hier die Spezifikationsgrenzen eines einzelnen Messwertes dar. Die roten senkrechten Linien zeigen die aus der Verteilung der Blau dargestellten Messwerte errechneten PAT Grenzen. Damit werden die in der Grafik als „Outliers“ bezeichneten Bauteile verschrottet, obwohl diese der Spezifikation entsprechen.

### **4.3 Optische Kontrolle / Bewertung**

Klarerweise sind die kleinen und vor allem dünnen Strukturen, die bei der Halbleiterfertigung hergestellt werden, gegenüber mechanischen Beschädigungen äußerst empfindlich. Mechanische Beschädigungen oder fehlerhafte Herstellung der letzten, schützenden Passivierungsschicht sind mit den beschriebenen Prüfverfahren nicht zu entdecken, da die Funktion der Produkte im Allgemeinen nicht beeinträchtigt ist. Da aber gerade Schäden in der Passivierung zu Frühausfällen der Produkte führen können, ist eine zusätzliche optische Kontrolle der Produkte notwendig. Dieser Schritt findet nach dem Funktionstest auf Wafer und vor dem Assembly der Bauteile statt.

Auch bei diesem manuellen Prüfprozess ist eine nahtlose Integration in die Datenverarbeitung unabdingbar. Die Bewertung von einzelnen Produkten auf der Scheibe oder von zusammenhängenden Strukturen (Bildfelder<sup>56</sup>) muss konsistent erfolgen. Die Ausbeutenachverfolgung muss natürlich wei-

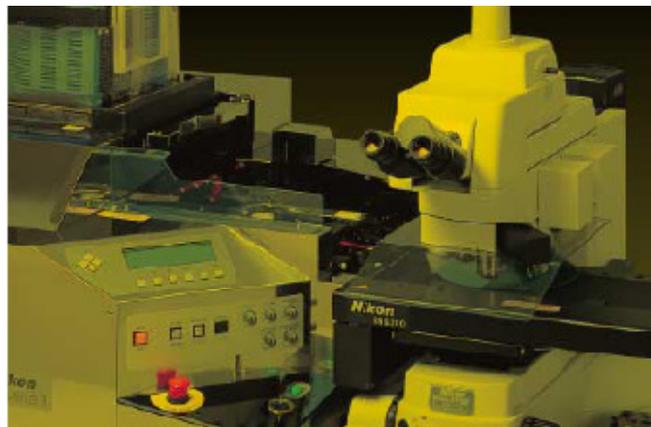


Abbildung 37 Inspektionsmikroskop  
NWL-860 [NIK08]

terhin stimmen. Zudem wird die Kennzeichnung der weiter zu bearbeitenden Produkte im Folgeschritt Assembly typischerweise nicht mehr durch die Aufbringung eines Farbkennzeichens auf dem Chip (inken) erreicht. Aus Qualitäts- und Effizienzgründen werden die Daten elektronisch übermittelt. Fehler hätten hier schwerwiegende Folgen, daher müssen diese Daten vollständig konsistent mit den Produkten sein.

Aus diesem Grund werden für diesen Prüfschritt Mikroskope mit automatisierter Verfahren des Objektträgers eingesetzt (Beispiel: Abbildung 37). Damit ist die Automatisierung jederzeit darüber informiert, welcher

---

<sup>56</sup> Ein Stepper bildet die Fotomasken auf den gerade belichteten Wafer ab. Das verkleinerte Abbild der Masken wird als Bildfeld bezeichnet.

Chip gerade angesehen wird. Die Bewertung des Bildes erfolgt allerdings weiterhin manuell. Entsprechend integrierte IT Lösungen ermöglichen die konsistente Beurteilung der Ware. Einschränkend ist zu bemerken, dass die optische Kontrolle normalerweise nur die letzte Fertigungsebene beurteilen kann. Fehler in darunterliegenden Schichten sind nicht erkennbar. Weiterhin ist dieses Verfahren wegen des manuellen Aufwands im Produktionseinsatz immer ein Stichprobenverfahren.

#### ***4.4 Engineering Data Analysis***

Die Qualitätsbewertung der geprüften Bauteile findet direkt bei Ende des Tests statt. Kommt es nun zu Abweichungen bei den Ausbeuten, bei den Messwerten oder bei der Verteilung der Fehler auf die verschiedenen Fehlerklassen wird versucht, die Ursache für diese Abweichungen zu finden und diese für zukünftige Lose desselben oder vergleichbarer Produkte zu korrigieren. Die verantwortlichen Engineeringabteilungen stützen sich hier auf die bereits genannten Auswertemöglichkeiten. Darüber hinaus gehend wird eine große Vielfalt von Werkzeugen zur Auswertung großer Datenmengen eingesetzt.

Die hier eingesetzten Werkzeuge reichen von Standard-Office-Anwendungen über Business Intelligence-Produkte (z.B. SAS) bis zu spezialisierten Werkzeugen zur statistischen Auswertung von Messwerten (Genesis, Statistica, Cornerstone, Spotfire, JMP). Insbesondere die Spezialprodukte bringen häufig eigene Datenbanken mit. Damit wird die Auswertung der Daten erleichtert und beschleunigt. Nachteilig ist allerdings, dass damit zusätzlich redundante Datenspeicher erzeugt werden.

Abbildung 38 zeigt ein Beispiel für eine Hauptachsentransformation. Die bereits beschriebenen Analyseverfahren für Messergebnisse (zum Beispiel SPC) erfordern funktionale Abhängigkeiten von einem Parameter. Durch geeignete Drehung der Punktwolke im Koordinatensystem kann man die Dimensionalität des Problems reduzieren (hier von zwei auf eins)

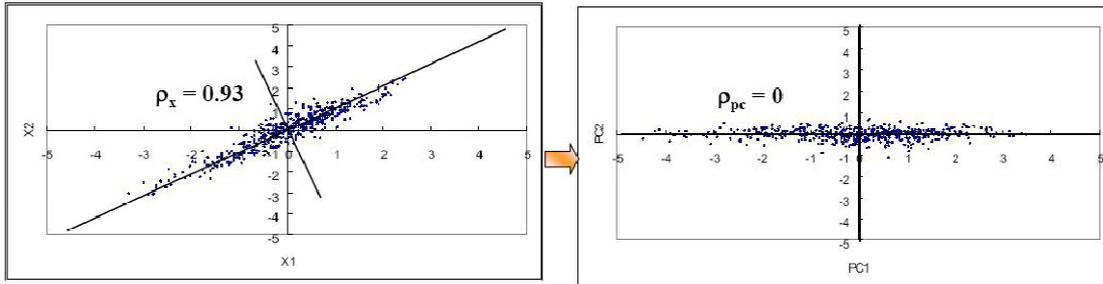


Abbildung 38 Principal Component Analysis [SPZ07]

Ein weiterer häufig erfolgreicher Ansatz ist die multivariante Statistik. Die geeignete Zusammenfassung verschiedener, ursächlich zusammenhängender Messgrößen zu einer Auswertung kann Auffälligkeiten identifizieren, die bei der getrennten Darstellung unerkennbar sind. Dieses ist in Abbildung 39 gut direkt erkennbar. Bei linear zusammenhängenden Parametern zeigt die  $T^2$ -Verteilung eine gute Sensitivität und kann für die mathematische Identifikation der Ausreißer verwendet werden.

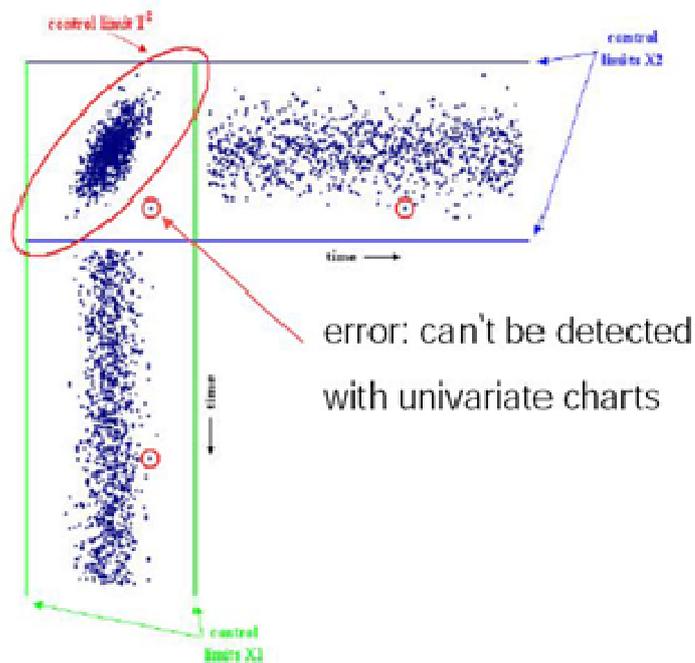


Abbildung 39 Hotelling  $T^2$  [SPZ07]

Der Ansatz ist jeweils identisch: Aufgrund der gesehenen Probleme entwickelt der bearbeitende Ingenieur Ideen, was als Ursache für die Effekte in Frage kommt. Anschließend wird versucht, mit den jeweils vorhandenen Werkzeugen in den Messdaten Belege für diese Hypothese zu finden oder sie zu widerlegen. Im Anschluss wird dann gegebenenfalls auf Basis dieser Hypothese eine Anpassung in der Fertigung oder ein Versuch

in der Fertigung durchgeführt. Die Resultate dieser Veränderungen sollen dann die Eingangshypothese verifizieren oder falsifizieren.

Leider bleibt hier typischerweise eine erhebliche Restunsicherheit bestehen. Auf der einen Seite wird der Herstellungsprozess wie beschrieben von sehr vielen Parametern beeinflusst. Dieser Einfluss ist im Allgemeinen nicht vollständig bestimmt. Zudem ist der Einfluss unter Umständen nicht stetig sondern chaotisch. Zusammen mit der Unmöglichkeit alle anderen Parameter vollständig identisch zu halten, ist hier eine verbleibende Unsicherheit in final unbekannter Höhe unabwendbar.

Hinzu kommt das jeder Änderung der Abläufe und Prozesse innewohnende Risiko, andere Eigenschaften des Produktes ungewollt zu verändern.

Die aufgezeigte starke Durchdringung einer modernen Halbleiterfabrik mit Informationstechnik führt zu enormen Anforderungen an den Datenhaushalt für diese Tätigkeiten. Die typischerweise in verschiedenen Systemen in unterschiedlichem Kontext gespeicherten Daten müssen zusammengeführt werden, um überhaupt hinreichend signifikante Ergebnisse erzeugen zu können.

Wie beschrieben werden für die Regelung der Fertigungsprozesse und die Analyse von Abweichungen erhebliche Aufwände in den heutigen Fabriken erbracht. Trotzdem ist das Ergebnis nicht zufriedenstellend. Häufig verschwinden die Abweichungen während der Analyse, ohne dass die Ursache sicher gefunden wurde, einfach weil nicht alle Rahmenbedingungen konstant gehalten werden können. Weiterhin kommt es vor, dass es final unerklärliche Instabilitäten bei einzelnen Produkten gibt. Die ist unter anderem mit den Beschränkungen den etablierten Verfahren zu begründen, die im folgenden Kapitel dargelegt werden.

# 5 Grenzen der etablierten Verfahren

Die beschriebenen Verfahren werden in den produzierenden Unternehmen eingesetzt, wobei SPC Standard ist und seit langem eingesetzt wird. Die restlichen Verfahren (4.1.2 – 4.1.4 , 4.2) sind noch nicht so etabliert und befinden sich bei vielen Unternehmen in oder vor der Einführung. Die Implementierung wird erschwert durch die hohe Komplexität der Fabriken. Selbst vergleichsweise kleine Änderungen und/oder Erweiterungen erhöhen das Risiko einer Qualitätsverschlechterung. Hinzu kommt noch die kontinuierliche Fertigung die die Einführung oder –schlimmer– Ablösung von Werkzeugen während der Produktionszeit notwendig macht. Regelmäßige Fertigungsstillstände gibt es nur in wenigen Fabriken und auch dort nur ein bis zweimal im Jahr (Weihnachten/Sylvester und/oder Ostern) für typischerweise wenige Stunden.

Einen geplanten Fertigungsstillstand für ein Upgrade eines bestehenden Werkzeuges oder die Einführung eines neuen verhindert im Regelfall der nicht darstellbare Return on Invest insbesondere unter Betrachtung der damit einhergehenden Risiken. Doch auch wenn die aufgeführten Verfahren vollständig implementiert sind, bleiben relevante nicht abgedeckte Felder übrig.

- Die Verfahren sind fast ausschließlich reaktiv.
- Die Regelung betrifft nicht die aktuellen Produkte sondern nur die zeitlich folgenden Produkte in der Fertigung.
- Die Verfahren sind primär auf die Symptome ausgerichtet und nicht auf die Ursachen.
- Abhängigkeiten werden nur im Zuge der technischen Datenanalyse untersucht. Angesichts der enormen Vielzahl der Parameter ist die Lösungsquote endlich.
- Einflüsse von Fertigungsmaschinen werden üblicherweise nicht berücksichtigt.
- Sonstige Umgebungseinflüsse werden üblicherweise nicht berücksichtigt.

Diese Unwägbarkeiten der Maschinen zusammen mit der Medienversorgung und den sonstigen Umwelteinflüssen wie dem Klima sind sehr komplex und final undurchschaubar. Der größte Halbleiterhersteller der Welt, die Firma Intel, kopiert bei neuen Fabriken bereits bestehende exakt [IN07].

### ***5.1 Einflüsse Equipment***

Wie schon bei der Prozessregelung beschrieben (siehe 2.2) handelt es sich bei den verwendeten Werkzeugen zur Halbleiterherstellung um sehr komplexe Maschinen, die dem Sondermaschinenbau zuzurechnen sind. Neben den beschriebenen Komplikationen für die Prozessregelung gibt es innerhalb der Systeme durchaus vielschichtige Ausfallmöglichkeiten.

Diese Veränderungen innerhalb der Fertigungsanlagen werden nicht alle gemessen geschweige denn geregelt. In wie weit sich Schwankungen von internen Maschinenparametern auf die Eigenschaften der produzierten Produkte auswirken, kann typischerweise nicht systematisch untersucht werden.

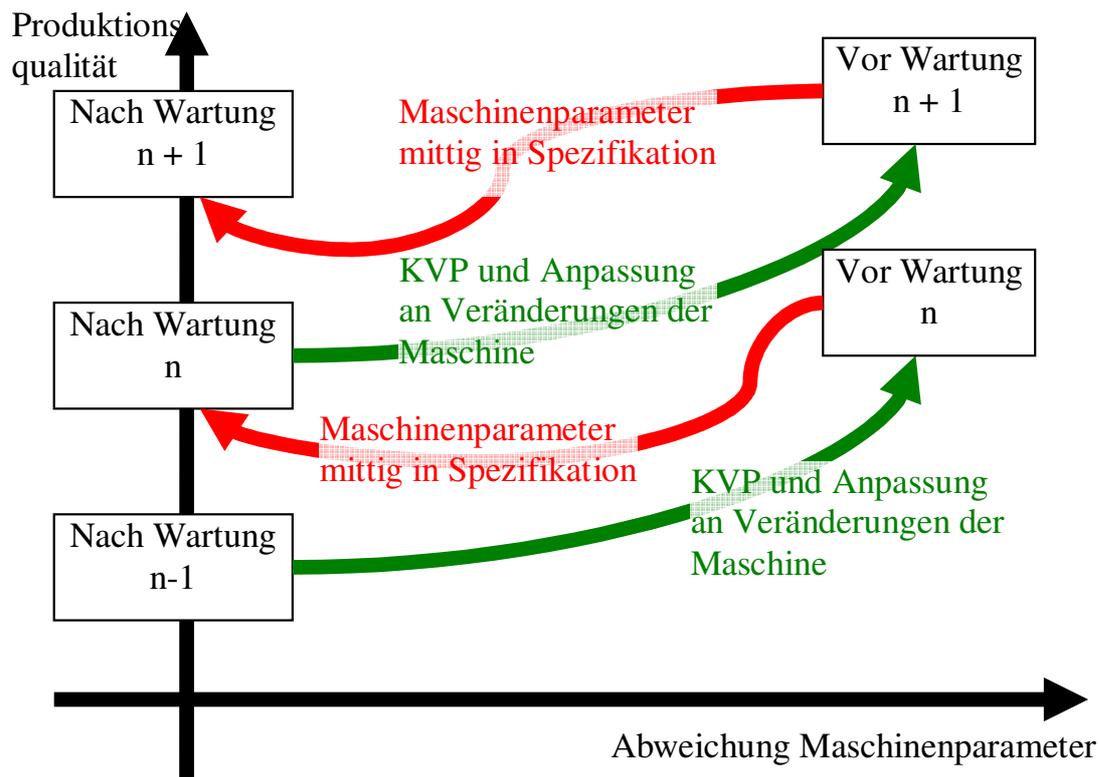


Abbildung 40 Optimierungsspirale für Fertigungsmaschinen

Üblicherweise gibt es beispielsweise direkt nach einer durchgeführten Wartung einer Maschine größere Schwierigkeiten mit der Prozessführung als vor der Wartung. Erklärbar ist dieses eigentlich unverständliche, für die Mitarbeiter vor Ort aber völlig „normale“ Phänomen mit kontinuierlichen symptomatischen Anpassungen der Fertigungsumgebung, zum Beispiel im Zuge der SPC-Nachregelung. Funktioniert die Maschine dann nach der Wartung wieder wie spezifiziert und in der Mitte des Prozessfensters, müssen die unter Umständen vielen durchgeführten Anpassungen wieder rückgängig gemacht werden. Da die Prozessführung weiterentwickelt wird, sind nicht alle Anpassungen durch Änderungen der Maschine bedingt. Daher ist eine Rückgängigmachung der Anpassungen nicht möglich. Stattdessen ist eine neue Einregelung des aktuellen Prozessfensters auf die Maschinenparameter notwendig. Abbildung 40 zeigt einen Überblick über den beschriebenen Ablauf.

## ***5.2 Einflüsse Umgebung***

Neben den beschriebenen ungeplanten und geplanten Einflüssen des Fertigungsequipments gibt es darüber hinaus trotz den aufwändigen Installationen Einflüsse der Umgebung.

### **5.2.1 Luftfeuchte und Temperatur**

Luftfeuchte und Temperatur werden in Halbleiterfabriken innerhalb vorgegebener Toleranzen geregelt. Diese sind zum Teil außerordentlich eng (innere Kammer Stepper:  $\sim 0,1$  K). Wie beim Equipment wird vorausgesetzt, dass Schwankungen innerhalb des Toleranzbereiches keinen Einfluss auf die Produktqualität haben. Beide Parameter können allerdings nicht als lokal invariant bezeichnet werden. Die vielen vorhandenen elektrischen Verbraucher erhöhen lokal die Temperatur, was unmittelbaren Einfluss auf die Luftfeuchte hat. Über die durch die Klimatechnik vorgegebenen Luftwege wird die Abluft der Verbraucher möglichst direkt abgesaugt. Dieses ist in „green field“-Fabriken, die direkt mit Zielkapazität und Zielprozess errichtet werden, wesentlich besser der Fall als in kontinuierlich gewachsenen, flexiblen Fabriken.

### **5.2.2 Defektdichte**

Die Defektdichte ist ein Ergebnis des Reinraummanagements. Die Untergrenze wird mit der Konstruktion der Fabrik faktisch vorgegeben. Abweichungen nach oben (zu höheren, also schlechteren Defektdichten) werden über das Reinraummanagement, primär durch die Klimatechnik verhindert. Wegen des unmittelbaren Zusammenhangs (siehe Abbildung 4) zwischen Defektdichte und Ausbeute und dem bereits beschriebenen Zuverlässigkeitsrisiko (siehe Kapitel 2.1) bei Vorliegen von Defekten ist das Thema von außerordentlichem Interesse in jeder Halbleiterfabrik. Die Defektdichte kann allerdings nicht direkt gemessen werden, sondern wird retrograd aus den Ausbeuten der Fertigung ermittelt. Damit ist eine direkte Regelung in der Fabrik faktisch unmöglich. Daher wird –faktisch als Steuerung– erheblicher Aufwand mit Minienvironments (SMIF), loka-

lem laminar flow und engen Überwachungs- und Wartungszyklen für die Klimatechnik getrieben, um die Defektdichte zu minimieren.

### 5.2.3 Luftdruck

Der Luftdruck nimmt eine Sonderstellung ein. Er gehört zu den einfach zu messenden und darüber hinaus lokal im Wesentlichen invarianten Größen. Allerdings ist er im Gegensatz zu anderen Größen wie Temperatur und Feuchte in Fertigungsumgebungen nur –auch im Vergleich zu den ansonsten notwendigen Maßnahmen für die Stabilisierung der Umgebung– sehr aufwändig zu beeinflussen. Das heißt, es handelt sich beim Luftdruck um eine unbeeinflussbare äußere Größe für die Prozessführung.

Erschwerend kommt hinzu, dass die hierfür besonders empfindlichen Ofenprozesse eine in Relation zu Luftdruckänderungen lange Laufzeit haben. Sehr fortschrittliche Öfen verfügen über geschlossene Prozesskammern und halten den Luftdruck innerhalb dieser Kammern konstant. Wegen der starken Temperaturschwankungen während des Prozesses muss es sich natürlich um eine aktive Regelung handeln. Ein einfaches Abdichten ist nicht ausreichend.

In aktuellen Fabriken ist durchaus –wie schon mehrfach angesprochen– älteres Equipment im Einsatz. In diesem Falle bedeutet das, dass häufig eben keine Öfen mit aktiver Druckregelung und geschlossener Prozesskammer vorhanden sind. In diesen Fällen werden typischerweise angepasste Rezepte für verschiedene Luftdruckszenarien vorgehalten. Das Rezept wird aber beim Start des Prozesses selektiert und kann nicht auf während des Prozesses erfolgende Luftdruckschwankungen reagieren.

Eine Optimierungsmöglichkeit besteht darin, die Rezeptauswahl von der Wettervorhersage abhängig zu machen. Inzwischen ist es sinnvoll möglich, eine kurzfristige, lokale Wettervorhersage zu empfangen und diese automatisiert für die Prozessregelung einzusetzen.

### ***5.3 Lange Durchlaufzeiten / hohe Kosten***

Verifikation oder Falsifikation der vermuteten Ursache für eine Abweichung kann unter Umständen einfach durch eine Nachmessung von elektrischen oder mechanischen Parametern bei dem vorliegenden Material erfolgen. Häufig ist es aber notwendig, entsprechende Versuche in der Fertigungslinie durchzuführen. Das bedeutet, dass für die Überprüfung der Vermutung Material nachproduziert werden muss. Dieses dauert mit einigen Wochen sehr lange und bedeutet wegen des Einsatzes von Material auch das Anfallen von erheblichen Kosten. Typischerweise ist Material, mit dem Versuche durchgeführt wurden, nicht verkaufsfähig.

### ***5.4 Unbekannte Zusammenhänge***

Der Erfolg der technischen Analyse der Ursache für eine Abweichung ist abhängig vom durchführenden Ingenieur. Neben dieser Abhängigkeit von der Qualifikation und Kreativität eines Einzelnen oder einer kleinen Gruppe hängt die Entdeckungswahrscheinlichkeit der Ursache in hohem Maße von der Ursache selbst ab. Abhängigkeiten von mehreren Parametern, deren Einfluss auf die vorliegende Abweichung nicht klar ersichtlich ist, können kaum entdeckt werden. Hinzu kommt, dass es bei der enormen Vielzahl der Parameter auch immer Raum für Optimierungen gibt. Diese verstellen dann bei der Analyse die Sicht auf den kritischen Effekt, der unbedingt identifiziert werden muss.

Schließlich sind die vielen Hundert verschiedenen Parameter auch für einen Spezialisten nicht überblickbar.

#### **5.4.1 Variabilität des Versuchsaufbaus**

Angesichts der sehr vielfältigen Abhängigkeiten der Fertigung und der sehr hohen Anzahl von Prozessschritten ist es nicht möglich, alle möglichen Einflussgrößen bis auf eine konstant zu halten. Eine zuverlässige Verifikation der These ist damit nur sehr eingeschränkt möglich. Damit

ist das klassische Verfahren zur Durchführung von Versuchen für das Ziel „Null-Fehler-Produktion“ nicht hinreichend.

### 5.4.2 Optimierung

Die beschriebenen Probleme bei der Analyse sind aber nur ein Teil der offenen Fragen. Mindestens genau so gravierend ist die Selektivität der menschlichen Wahrnehmung. Ein sich langsam verändernder Wert wird physiologisch bedingt sehr schlecht wahrgenommen. Dies führt dazu, dass einzelne Effekte erst dann wahrgenommen werden, wenn es zu einer Spezifikationsverletzung oder Fehlfunktion des Produktes kommt. Aufgrund der dann zu betrachtenden langen Zeiträume, ist die Identifikation der Ursache sehr schwierig.

### *5.5 Grenzen der Simulation*

Der Versuch der Herleitung der Parameter des fertigen Produktes aus den Fertigungsparametern im Zuge einer Simulation ist für die Entwicklung unverzichtbar und wird dort auch intensiv genutzt. Allerdings findet hier ein mehrstufiges Verfahren statt. Üblicherweise werden die Einzelnen verwendbaren Bauelemente innerhalb eines Fertigungsprozesses im Zuge der Entwicklung dieses Prozesses entwickelt. Für diese Bauelemente wird mit speziellen Werkzeugen der Fertigungsprozess simuliert, also das Aufbringen der verschiedenen Schichten, die Implantation etc. Sodann werden diese Bauteile in drei Dimensionen elektrisch simuliert. Nach der Fertigung der Bauteile werden diese manuell vermessen. Bei Erreichung des gewünschten Ergebnisses werden diese Einzelbauteile – häufig zu Baugruppen zusammengefasst – den Produktentwicklern zur Verfügung gestellt. Für die Werkzeuge zur Schaltungsentwicklung werden dann die Schlüsselparameter bereitgestellt. Diese Schlüsselparameter werden dann für die elektrische Simulation des entwickelten Produktes genutzt.

Einen Durchgriff von einem gefertigten Produkt zu den Simulationsergebnissen in der Produktentwicklung gibt es durchaus. Eine weitere

Verbindung zu den Konsequenzen aus den physikalischen Randbedingungen der Fertigung gibt es in der Anwendung nicht.

Hinzu kommt, dass bereits eine vollständige analoge Simulation eines digitalen oder Mixed-Signal-Produktes aus Ressourcengründen nicht durchführbar ist. Eine noch wesentlich weiter gehende mechanisch-elektrische Simulation aller Bauteile in einem fertigen IC würde jeden Rahmen sprengen.

Damit ist die Herleitung der Ursachen für Produktabweichungen aus Fertigungsparametern mit einer Simulation nur in seltenen Ausnahmefällen brauchbar.

In der Vergangenheit hat es einige Ansätze gegeben, die Steuerungs- und Analysemöglichkeiten in den Fabriken wesentlich zu verbessern. Neben der kontinuierlichen Verbesserung der eingesetzten Verfahren und Prozesse gab es gemeinschaftliche Projekte von Wissenschaft und Industrie, um neue Verfahren zu entwickeln. Einige dieser Ansätze werden im folgenden Kapitel beschrieben.

# 6 Auswege aus der Krise

Die in den komplexen High-Tech-Fabriken vorhandene Software-Infrastruktur hat sicher zum Erfolg der Technologie in der Vergangenheit beigetragen. Angesichts gestiegener Anforderungen treten allerdings Defizite der installierten Lösungen zu Tage:

- Für die verschiedenen Kernfunktionen (z.B. MES, SPC, Maintenance, EDA) werden häufig verschiedene Systeme eingesetzt, die mangelhaft oder gar nicht integriert sind.
- Die kommerziellen primär logistischen Systeme (MES/ERP/PPS<sup>57</sup>) erfüllen nicht einmal näherungsweise die heutigen Anforderungen. Entsprechend werden diese massiv modifiziert und mit vielen „Rucksäcken“ versehen, die häufig auf undokumentierte Schnittstellen angewiesen sind.
- Die verfügbaren Werkzeuge sind fast durchgängig reaktiv.
- Die Schnittstellenproblematik zwischen den installierten Systemen und zwischen diesen und den Fertigungsanlagen ist kaum noch beherrschbar.

---

<sup>57</sup> Produktionsplanungs- und Steuerungssystem

- Ein deduktiver Zusammenhang der vorhandenen Daten ist nicht darstellbar. Systeme die aus den vorhandenen unstrukturierten Daten Informationen erzeugen sind nicht vorhanden.

Viele Ansätze wurden gemacht um die genannten Themen zu adressieren. Wegen der enormen Komplexität der Themen und der unbekanntem Lösung waren das häufig Kooperationen mehrerer Industriepartner mit wissenschaftlichen Institutionen.

Andere Industriezweige haben ebenfalls Anstrengungen unternommen, um die IT Unterstützung für die Prozessprobleme signifikant zu verbessern. Wie beispielsweise im Bericht des DFG Sonderforschungsbereiches 281 zu erkennen ist, sind viele der in der Halbleiterindustrie bereits installierten Systeme in anderen Industriebereichen nicht selbstverständlich [PFT02].

### ***6.1 Lebenszyklus von Fertigungsleitsystemen***

Bei der Betrachtung des Lebenszyklus von Fertigungsleitsystemen werden erhebliche Dilemmata deutlich:

- Das System entspricht schon bei der Installation nicht richtig den Anforderungen und Bedürfnissen.
- Die Anforderungen in modernen Fabriken ändern sich schnell während die Standardsoftware sehr statisch ist.
- Die –typischerweise alte– Basis der Software gibt die notwendigen Änderungen gar nicht her.
- Trotzdem wird bei der Beschaffungsentscheidung für ein MES in einer neuen Fabrik häufig ein bereits im jeweiligen Konzern bekanntes MES gewählt. Die Ursache ist das erhebliche Risiko bei der Einführungsentscheidung.
- Ein Austausch des MES in einer laufenden Fabrik findet praktisch nicht statt.
- Die Systeme basieren typischerweise auf proprietären Datenablagensystemen ohne hinreichende dokumentierte Schnittstellen.

Um die Basis der Fertigungsleitsysteme zu vereinheitlichen und zu öffnen, erstellte Sematech mit starker Förderung durch die USA ein Framework für eine Halbleiterfabrik [SEM98]. Da dieses Framework –analog

der Entwicklung der marktgängigen Fertigungsleitsysteme- für sehr große Fabriken mit wenigen verschiedenen Produkten optimiert ist, wurde im Rahmen eines vom BMBF geförderten Projektes nach einer Lösung für die in Europa vorherrschenden flexiblen und kleineren Fabriken gesucht [SFCIM98, CIM99]. Neben der Definition eines allgemeinen objektorientierten Datenformats für MES in Halbleiterfabriken war wesentliches Thema des Projektes die mögliche Umstellung der bei den teilnehmenden Partnern vorhandenen MES auf die neue Version auf Basis von offenen Systemen[KRAM02]. Die Umstellungszeiträume sind aus den genannten Gründen sehr lang. Entsprechend sind die hier beschriebenen Arbeiten nach wie vor aktuell.

Die Definition eines verallgemeinerten Ansatzes sowie die Definition der notwendigen Technologien und Schnittstellen war sehr erfolgreich. Das entwickelte Framework ist sowohl für die Anbieter als auch für die Nutzer von Anlagen und Software ein sehr hilfreicher Rahmen. Beispielsweise wurde ein generelles Modell für die Datenanalyse entwickelt und auf dem zweiten offenen Workshop des Projektes publiziert [MSK97]. Die Definition fand dann anschließend Eingang in den entsprechenden Normungsvorschlag.

„Migration ist ein permanenter Prozess und damit ein Lösungsweg, der das Dilemma zwischen informationstechnischer Entwicklung, der Organisation der Geschäftsprozesse, der Organisation der Informationsverarbeitung und den ständig sich entwickelnden Nutzerforderungen in vertretbaren Grenzen hält.“

Definition Migration nach SF-CIM

Eines der relevanten Ergebnisse dieses mehrjährigen Projektes ist die Erkenntnis, dass neben den technischen und organisatorischen Verfahren auch die Migrationsstrategie selbst kontinuierlicher Veränderung unterliegt.

Daraus folgt, dass man Migrationsprojekte in überschaubare kleine Module zerlegen muss, die trotzdem dem weiteren Wandel gewachsen sind. Eine schlagartige, vollständige Migration (Big Bang) einer prozessieren-

den Fabrik (365 Tage/24 Stunden) ist wegen des ökonomischen Risikos angesichts des ungewissen Ausgangs nicht vertretbar.

Entsprechende IT-Projekte (mitsamt den Migrationsproblemen) unterliegen einem enormen Kosten/Nutzen-Druck, der durch stark schwankende Wirtschaftsverhältnisse in der Halbleiterindustrie sowie unsicheren DV-technologischen Entwicklungsperspektiven beeinflusst wird [SUN99].

In [EIT07] wird ein vergleichendes Klassifikationsschema für Umstellungsprojekte von kritischen und integrierten Softwarewerkzeugen vorgeschlagen. Damit könnten sinnvollerweise die Umstellungsrisiken der einzelnen Module strukturiert werden.

## ***6.2 Informationsgenerierung aus Daten***

Neben den beschriebenen, sich aus den vorhandenen Systemen ergebenden Schwierigkeiten ist die sehr große anfallende Datenmenge eine Herausforderung für die Auswertung derselben.

Diese Daten entstammen verschiedenen Datenquellen und sind typischerweise nicht miteinander verbunden geschweige denn normalisiert. Zudem sind die Zusammenhänge von Ursache und Wirkung bei diesen Daten, wie in den bisherigen Kapiteln ausführlich dargelegt, üblicherweise nicht bekannt. Dieser Umstand ist sowohl von der Industrie als auch von wissenschaftlichen Instituten erkannt und es gibt und gab vielfältige Ansätze, diese Situation zu verbessern.

Die deutschen Halbleiterhersteller haben sich gemeinsam mit der Fraunhofer Gesellschaft um Fördermittel für das Projekt ak@de.mi [AKA99] bemüht. Fokus dieses Projektes war –neben der notwendigen Qualifikation der Mitarbeiter– das Auffinden und Darstellen der Zusammenhänge zwischen den vorliegenden Datenmengen aus modernen Fabriken.

Dazu ist zunächst die Beherrschung der Datenmenge vonnöten. Weiterhin war das Ziel des Projektes die Reduktion der Daten und final die Erzeugung von Wissen (s. Abbildung 41). Das Projekt wurde wegen der Kürzung von Fördermitteln vom Zuwendungsgeber final nicht bewilligt.

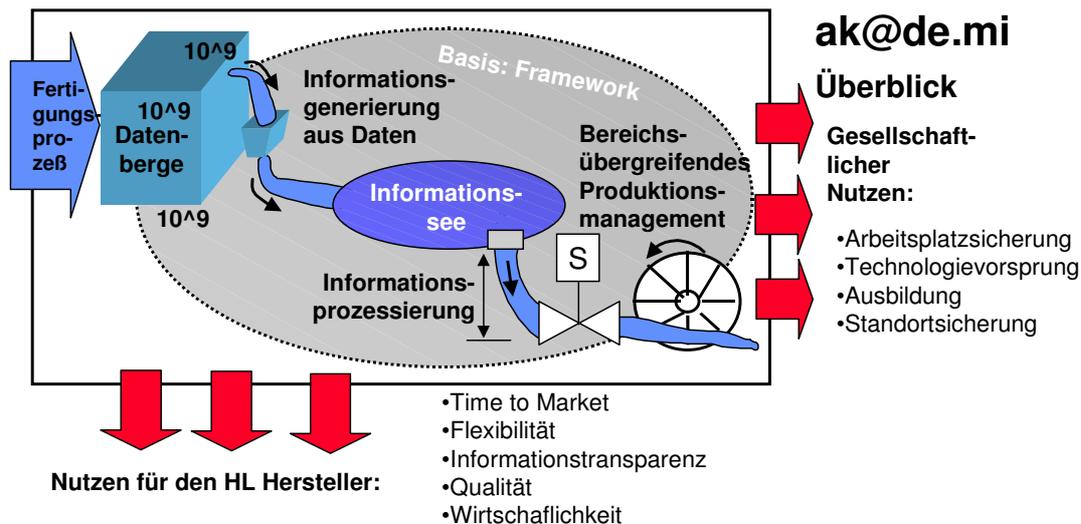


Abbildung 41 Informationsgewinnung aus Daten [AKA99]

Eine Lösung für die beschriebenen Probleme ist den Unternehmen und den Instituten allerdings auch in der Folgezeit nicht gelungen. Entsprechend finden sich weitere Ansätze mit einer ähnlichen Fragestellung.

Auch die Projektskizze „Flexible Advanced Process Control: fAPC, Neuartige Flexible Prozesskontrollverfahren“ [FAPC04] zielt auf die Verbesserung der Analyse und Regelungsmöglichkeiten in aktuellen High-Tech Fabriken ab. Auch hier wurde eine Kooperation der deutschen Halbleiterhersteller [Bosch, Elmos, Infineon, Micronas, Philips] mit Hochschulen [UNI Paderborn], Forschungseinrichtungen [Fraunhofer IAO Stuttgart, Fraunhofer IISB Erlangen] und Softwareherstellern [Abakus, AIS, camLine, TIP] angestrebt um auf dem beschriebenen komplexen Weg voranzukommen. Die Beschreibung der Situation und die notwendigen Forschungs- und Entwicklungstätigkeiten sind in [FAPC04] zusammengefasst.

Die Projektskizze fokussiert klar auf den APC (siehe 4.1.4 , Seite 75) bezogenen Teil der Problemstellung. Deutlich wird, dass die Suche nach den Ursachen von außergewöhnlichen und im Allgemeinen unerwünschten Effekten nach wie vor in Forschung und Industrie einen großen Raum einnimmt. Allerdings gibt es –auch zwischenzeitlich– keine veröffentlichten Lösungen zu den in Tabelle 5 dargestellten Fragestellungen, so dass die Themen nach wie vor als im Wesentlichen ungelöst betrachtet werden müssen.

Stand der Technik	F&E-Bedarf
Weitgehend SPC-basierte Fertigungskontrolle	Automatische Regelschleifen basierend auf Prozessmodellen
Starres APC-Framework zur Datenverteilung	Konfigurierbares, nach Fabrikfordernissen skalierbares Framework
Festverdrahtete APC-Software	Offenes System, das für plug-in der bestmöglichen Komponenten zugänglich ist
Häufig mehrere spezifische Prozesskontroll-Lösungen basierend auf einzelnen Softwareprodukten mit den Defiziten: <ul style="list-style-type: none"> <li>• zueinander inkompatible Systemkomponenten</li> <li>• unnötig redundante Funktionen (nicht eindeutig bestimmte Systeme)</li> <li>• Gefahr von Inkonsistenzen</li> </ul>	Generisches fAPC-Framework: <ul style="list-style-type: none"> <li>• kompatible, plug-in Systemkomponenten</li> <li>• fAPC-Standard mit eindeutiger Definition der APC - Komponenten</li> <li>• Anwender-unabhängig nutzbare, offene, standardisierte Schnittstellen</li> </ul>
Implementierung von Prozess- und Regelungs-Modellen meist im „Trial-and Error“-Verfahren, mit Risiken für die laufenden Prozesse	Simulationsgestützte off-line Implementierung - Testmöglichkeiten, APC-Modelle für Simulation und Funktionsvalidierung

Tabelle 5 Stand der Technik [FAPC04]

### 6.3 Nutzung der Maschinendaten

Die stark steigenden Investitionen für Halbleiterfabriken sind primär durch den Kostenanstieg der Fertigungsmaschinen verursacht. Mittels Statistic Equipment Control (siehe 4.1.2 ) wurden bereits zuvor Maschinendaten überwacht und gegebenenfalls nachgeregelt. Für die zwi-

schenzeitlich gewachsenen Anforderungen reichte dieses jedoch nicht aus.

Um den Return on Invest und damit die Wettbewerbsfähigkeit zu steigern mussten die Betriebszeiten der Anlagen erhöht werden. Gleichzeitig sollte durch die Reduktion der ungeplanten Unterbrechungen die Materialverschrottungsrate reduziert und die Qualität erhöht werden.

Insbesondere die Nutzung der prinzipiell vorhandenen Maschinendaten und deren zielgruppenorientierte Filterung und Aufbereitung war nicht angemessen gelöst. Daher haben die Halbleiterhersteller Atmel, Bosch, Elmos, Philips, X-Fab und ZMD mit den Softwareherstellern Camline und TIP sowie der Fraunhofer Gesellschaft und der Universität Tübingen das Projekt: „Overall Equipment Effectiveness, neues Konzept zur Effizienzsteigerung in der Halbleiterfertigung (OEE)“ [OEE05] durchgeführt.

Neben der geordneten Sammlung der Daten von den Fertigungsmaschinen wurden in diesem Projekt umfangreiche Untersuchungen zur Datenverdichtung und zur Datenaufbereitung durchgeführt. Die sehr große Menge an verfügbaren Equipmentdaten wurde selektiert und nach Relevanz für die verschiedenen Zielgruppen in der Fabrik klassifiziert. Das Thema war, auf der einen Seite den entsprechenden Funktionseinheiten genau die für die jeweils aktuelle Aufgabenstellung relevanten Daten zur Verfügung zu stellen. Eine Informationsüberflutung mit dem damit einhergehenden Verlust von Performance und dem steigenden Risiko etwas Relevantes zu übersehen wurde als genau so kritisch eingestuft wie ein Mangel an Information.

Die hier durch eigene Arbeiten entstandene Basis zur Datenerhebung und Datenverarbeitung dient als Grundlage für weitere Arbeiten auf dem Feld der Datenanalyse und Reporterzeugung.

Auf der in Tabelle 6 beschriebenen Grundlage wurde im Rahmen dieses Projektes eine allgemeine Lösung entwickelt. Weitere erfüllte Anforderungen waren die einfache Integrationsfähigkeit in existierende Landschaften und der weitgehende Verzicht auf herstellereigene Standards.

Persistenz	Die Inhalte einer Auswertung dürfen nicht vom Zeitpunkt der Auswertung abhängen. Das heißt wenn noch einzelne Status unklar sind, ist das in den Auswertungen entsprechend zu kennzeichnen.
Automatische Auswertung / Verdichtung	Sämtliche Protokolldateien, die erzeugt werden, müssen automatisch ausgewertet und / oder verdichtet werden.
Transparenz	Die Benutzerführung über die verschiedenen Benutzergruppen muss durchgängig einheitlich sein.
Sicherheit	Die Abgrenzbarkeit der Informationsempfänger gegeneinander ist notwendig. Nicht jeder darf jede Information sehen.
Administrierbarkeit	Die Konfiguration und Vergabe der Zugriffsrechte muss möglichst einfach möglich sein. Eine Ankopplung an die in den Organisationen vorhandenen Systeme ist zwingend (z.B. LDAP <sup>58</sup> )
Performance	Interaktive Aktionen sollen die noch nicht störenden Reaktionszeiten (2 s) nicht überschreiten.

Tabelle 6 Allgemeine Anforderungen an Reportingstrukturen

Gegenüber der Statistical Equipment Control ergeben sich aus der großflächigen Einführung von OEE Vorteile hinsichtlich der nutzbaren Betriebszeit und damit dem Return On Invest. Allerdings beschränkt sich das Modell auf die Maschinenseite und lässt die Materialseite außer Acht. Damit sind die Themen Ausbeute und Produktqualität mit den hier gefundenen Methoden nur mittelbar zu beeinflussen.

---

<sup>58</sup> Lightweight Directory Access Protocol: Kommunikationsprotokoll zur Abfrage und Modifikation eines Verzeichnisdienstes.

Die wesentliche Einschränkung der hier untersuchten Methodik besteht jedoch darin, dass hier wiederum regelbasiert vorgegangen werden muss. Damit kann das Verfahren nur für vorher schon bekannte oder erwartete Effekte brauchbare Antworten liefern.

### 6.4 Ereignisverarbeitung

Neben den genannten Punkten, die die Verbreiterung der Datenbasis und deren bessere Nutzung zum Thema haben, gibt es auch erheblichen Handlungsbedarf bei der Frage, wie auftretende Ereignisse weiterverarbeitet werden. Bislang sind in existierenden Installationen primär starre Arbeitsabläufe definiert worden. Dieses hat den Vorteil, zuverlässige Abläufe sicher zu stellen und für deren angemessene Dokumentation zu sorgen. Auf der anderen Seite findet eine Weiterentwicklung der Abläufe auf Basis der neu gewonnenen Erkenntnisse bestenfalls in größeren Abständen und manuell statt.

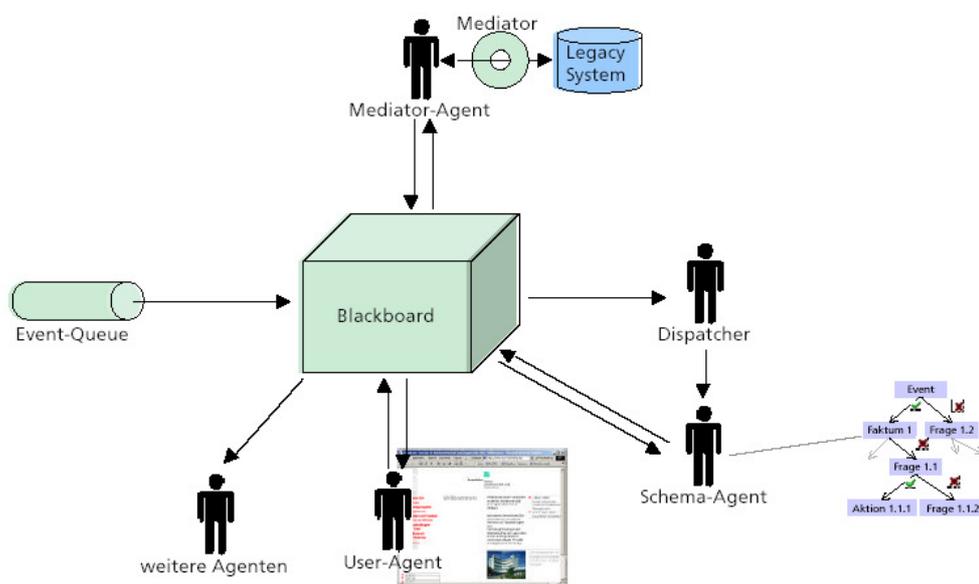


Abbildung 42 Flexible Ablaufverwaltung [OEE02]

Im Rahmen des obengenannten OEE Projektes wurde daher ein flexibles und dynamisches Verfahren zu Behandlung von Ausnahmesituationen in Fabriken (aber nicht nur dort) identifiziert. Abbildung 42 zeigt die Übersicht über dieses Verfahren. Ein Blackboard-System [FATHI94] wird über

spezifische Mediatoren (Vermittlerprozesse) an das jeweilig installierte System gekoppelt. Eingehende Events werden über die Event Queue serialisiert und auf das Blackboard gestellt. Dann können beliebige Agenten beliebige Aktionen auf den Events ausführen. Das Blackboard-System stellt die Rückverfolgbarkeit und die Zugriffssicherung sicher.

Eingebunden in die vielfältigen Steuerungssysteme einer Fabrik kann mit dem gefundenen Verfahren ein deutlich leistungsfähigerer Arbeitsablauf dargestellt werden, als das mit gängigen installierten Verfahren der Fall ist. Ein Beispiel für einen solchen Ablauf zeigt Abbildung 43. Hier wird eine Statusmeldung von einer Fertigungsanlage an das Blackboard gemeldet. In diesem Falle handelt es sich um eine „down“ Meldung, das heißt die Fertigungsanlage produziert nicht mehr. Wichtig ist es nun, möglichst schnell, zuverlässig und robust die entsprechenden Organisationseinheiten des Unternehmens zu benachrichtigen und sie mit den notwendigen Informationen zu versehen. Umgesetzt ist hier ein flexibles Eskalationsszenario vom Operator über den Service bis zum Hersteller der Anlage. Die notwendige Dokumentation der Vorgänge sowie die flexible Steuerung der Eskalation wird durch das Blackboard-System bereitgestellt.

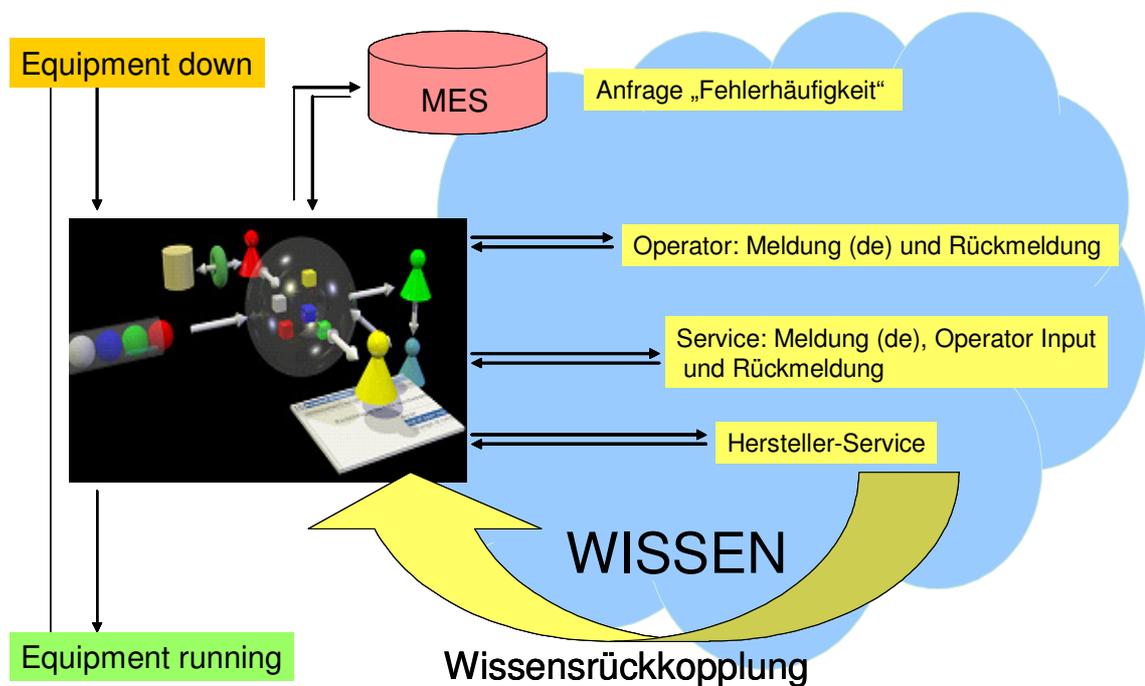


Abbildung 43 Wissensverarbeitung [OEE02]

Das Verfahren kann helfen, die Konsequenzen aus einer Ausnahmesituation zu reduzieren. Da es sich hier wiederum um ein reaktives Ver-

fahren handelt, ist auch dieser Ansatz keine Lösung für die Frage nach der besseren Erkennung und Beseitigung von Fehlern und deren Ursachen.

In Kapitel 4 wurde der Stand der Technik beschrieben, wie er aktuell in den existierenden Fabriken genutzt wird. Wie aufgezeigt, haben diese Verfahren noch Lücken, unter anderem in Bezug auf die Suche nach Ursachen für Abweichungen von den zu erwartenden Ergebnissen der Fertigungsprozesse. Bisher wurden in diesem Kapitel Beispiele aufgezeigt für Versuche der Industrie und der Wissenschaft diese Lücken zu schließen. Da diese Versuche insbesondere in Hinblick auf die Ursachensuche nicht erfolgreich waren, wird nun ein neuer Ansatz versucht. Mit dem in diesem Dokument vorgestellten Ansatz wird ein neuer Weg beschritten. Dieser besteht darin, dass die Ursachenanalyse im Kern ohne das Know How eines menschlichen Experten Ursachen für Prozessabweichungen in den vorhandenen Daten sucht. Neu ist hier, dass eine gekoppelte Kombination eines datenreduzierenden Prozesses (Feature Selection) mit einem wissensgenerierenden Prozess (Neuronales Netz) zur Analyse hochdimensionaler Massendaten eingesetzt wird. Dazu wird zunächst untersucht, ob für diese Fragestellung bereits wissenschaftliche Ergebnisse veröffentlicht wurden.

### ***6.5 Stand der Wissenschaft***

Im Umfeld Zero Defect Produktion und Automobilindustrie gibt es national und international aktuelle Veröffentlichungen. Die Themen der Veröffentlichungen beziehen sich häufig auf die Verbesserung der bereits beschriebenen Verfahren [z.B. ZVEI06, ZVEI07, DEBA07]. Weiterhin werden die notwendigen Veränderungen im Entwicklungsprozess zur Erreichung der Qualitätsziele diskutiert [z.B. MEMU08, LAMB06].

Zum Einsatz von wissensverarbeitenden Systemen und insbesondere von neuronalen Netzen und Feature Selection sind kaum aktuelle Veröffentlichungen zu finden. Viele Arbeiten zum Thema Feature Selection datieren aus den Jahren 1997 bis 2002. Die Veröffentlichungen beziehen sich primär auf die Basis des Verfahrens während Veröffentlichungen über den erfolgreichen Einsatz der Verfahren in der Industrie sehr selten sind.

In [SRMU03] wird eine Kombination aus einem Neuronalen Netz und einem Klassifikator (in diesem Falle eine SVM<sup>59</sup>) für ein Intrusion-Detection-System genutzt. Das System dient zur Klassifizierung von Daten, leider aber nicht zur Erzeugung von Regeln für die Klassifizierung.

Zur an der Realität überprüften Kombination aus Feature Selection und Neuronalen Netzen konnten kaum passende Veröffentlichung gefunden werden.

Die vorherrschenden Veröffentlichungen hinsichtlich des Einsatzes von Neuronalen Netzen betreffen die Modellierung von komplexen Systemen zur Vorhersage des Systemverhaltens [NAC08, JIND08, FIGA08, HAN08, TSAI08, ANNA08, LAW99]. Die durchgeführten Forschungen gehen durchaus mit vielen Datensätzen um. Allerdings ist die Dimensionalität der untersuchten Datenstrukturen klein gegenüber der in der Halbleiterfertigung anfallenden Parameteranzahl. Zudem wird nicht auf die dem Netzwerk zugrundeliegenden Regeln zurückgegriffen, das Ziel ist nur die Vorhersage.

Weitere untersuchte Veröffentlichungen betreffen den Einsatz von Neuronalen Netzen in der Auswertung von unstrukturierten Daten [WAN08, KARIM08, HWAN08, HEIG08]. Hierbei ist vor allem die Suche nach Inhalten bei Texten [REDD08] und Bildern [SHEI08, CHO08, AMIN98, FU03, BELLE08] im Zentrum des Interesses. In den genannten Untersuchungen werden keine Messwerte ausgewertet und es wird auch nicht versucht, die Regeln des Netzwerkes zu extrahieren. Die vorgestellten Untersuchungen zeigen aber die Eignung von Künstlichen Neuronalen Netzen zur Untersuchung von sehr großen Datenmengen.

Die untersuchten Veröffentlichungen zur Feature Selection befassen sich primär mit der Datenreduktion bei der Untersuchung von Texten [SHEN08, KUM08, KORY08, TAHA99, KIT99, DOHO05]. Wie schon bei den Veröffentlichungen über Künstliche Neuronale Netze ist hier zu erkennen, dass auch Feature Selection Verfahren für sehr große Datenmengen geeignet sind. In [CXW08] werden große Datenmengen mit einer SVM

---

<sup>59</sup> Bei der Support Vector Machine handelt es sich um einen nichtlinearen numerischen Klassifikator. Dieser ist besonders für Zeitreihenanalysen geeignet.

reduziert. Dieses zeigt die grundsätzliche Eignung von Feature Selection Verfahren auch für sehr große Mengen an Messwerten.

	Vorhersage	Text	Audio	Bilder	Datenanalyse
Japan	3%	12%	0%	0%	0%
Europa	6%	3%	3%	12%	0%
USA	18%	6%	0%	3%	3%
Sonst	21%	3%	0%	9%	0%
Gesamt	48%	24%	3%	21%	3%

Tabelle 7 Veröffentlichungen nach Thema und Region

Notwendig ist noch die Untersuchung, ob sich diese Eignung auch auf die Verfahren erstreckt, die für die dieser Arbeit zu Grunde liegende Fragestellung geeignet sind und ob die Verfahren auch in der Kombination funktionieren. Von zentraler Bedeutung ist jedoch die notwendige Extraktion der Regeln des Neuronalen Netzes. Zu diesem Punkt konnte keine geeignete Veröffentlichung gefunden werden. In Tabelle 7 sind die beschriebenen Veröffentlichungen zusammengefasst dargestellt

Abbildung 44 zeigt den Inhalt der Tabelle noch einmal in grafischer Darstellung. Wie schon im Text beschrieben ist deutlich erkennbar, dass kaum Veröffentlichungen über die Auswertung von Messdaten aufgefunden werden konnten.

Aufgrund der besonderen Konzentration der automobilen Zulieferkette in Europa und hier insbesondere in Deutschland wurden zusätzlich die Forschungsschwerpunkte von Hochschulen auf hier signifikante Veröffentlichungen gesondert untersucht.

An der TU München findet sich mit dem statistischen Herangehen an das Lernen von Regeln ein interessanter Ansatz. Nach Rückert [UR08] sind Systeme für maschinelles Lernen häufig auf die Vorhersagegenauigkeit optimiert, worunter häufig die Interpretierbarkeit der Regeln leidet. In der Arbeit „A Statistical Approach to Rule Learning“ wurde nun durch die Kombination von statistischen Methoden mit denen des maschinellen Lernens versucht, die Interpretationsfähigkeit der Regeln zu erhöhen ohne die Vorhersagegenauigkeit zu reduzieren.

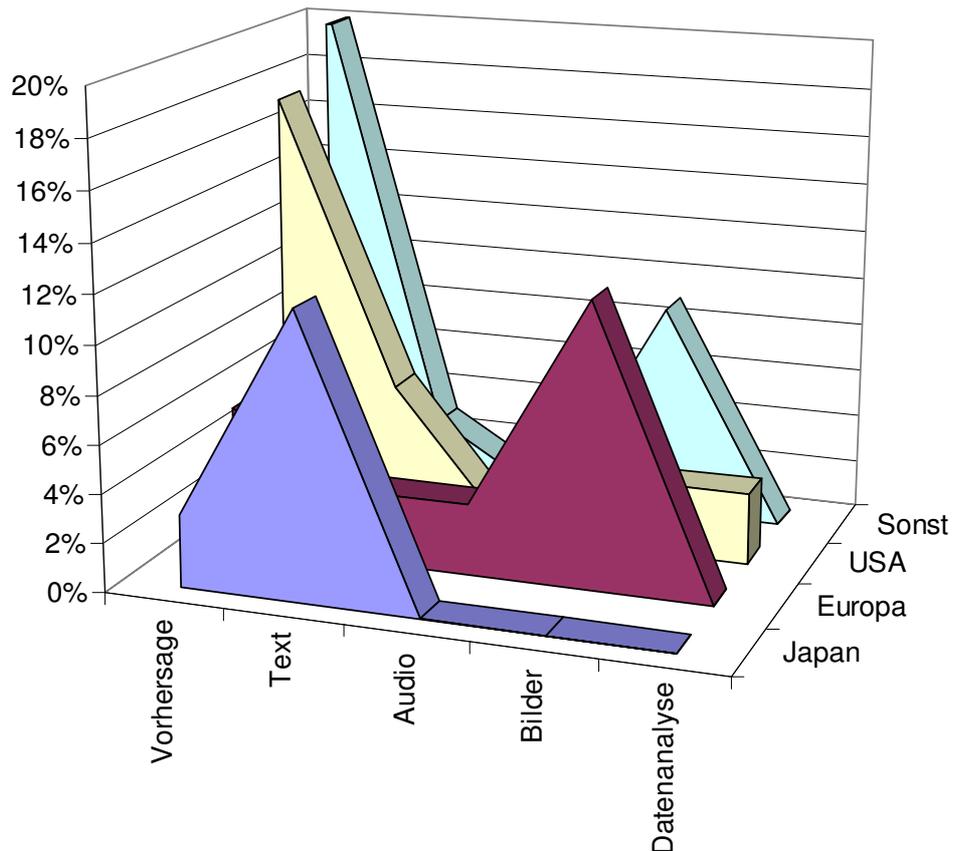


Abbildung 44 Veröffentlichungen nach Thema und Region

Für die bei Rückert beschriebene Statistische Auswertung der Regeln ist die Reproduzierbarkeit der Vorgänge und damit die dauerhafte Verfügbarkeit hinreichender Datenmengen Voraussetzung. Wie bereits aufgezeigt, ist das hier zu lösende Problem immer unterbestimmt. Damit kann dieses Verfahren auf die dieser Arbeit zu Grunde liegende Fragestellung nicht sinnvoll angewendet werden.

### 6.5.1 Massendaten und Neuronale Netze

In der großen Mehrzahl der Veröffentlichungen über Neuronale Netze ist die auszuwertende Datenbasis im Vergleich zu der hier untersuchten Datenmenge vergleichsweise klein. Wie schon im letzten Absatz ausgeführt, ist sind insbesondere kaum Veröffentlichungen zu finden, die die Regelextraktion bei großen Datensätzen zum Thema haben. Um den Stand der Wissenschaft weiter zu überprüfen, wurde neben der durchgeführten Untersuchung der Veröffentlichungen noch eine weitere Untersuchung ausgehend von bereitgestellten Datenquellen durchgeführt.

Die University of California stellt Datensätze für die Forschung maschinellen Lernens bereit [UCI08]. Die Untersuchung der zahlenbasierten Datensätze neueren Datums zeigt, dass insbesondere Anzahl der Attribute mit einer Ausnahme eher klein ist. Für die dort bereitgestellten Datensätze mit größeren Datenmengen sind keine Veröffentlichungen angegeben. Die Untersuchung der am häufigsten für Veröffentlichungen genutzten Datensätze zeigt, dass hier die Zahl der Attribute immer kleiner als 50 ist.

Aktuelle Datensätze	Datum	Zeilen	Spalten	Fehlende Datensätze	Veröffentlichungen
Ozon Detect.	2008	2536	73	Ja	0
Abscisic Acid Sig. Network	2008	300	43	n/a	0
Hill-Valley	2008	606	101	n/a	0
Forest Fires	2008	517	13	n/a	0
Dorothea	2008	1950	100000	n/a	0
Populärste Datensätze	Datum	Zeilen	Spalten	Fehlende Datensätze	Veröffentlichungen
Iris	1988	150	4	Nein	>100
Adult Data	1996	48842	14	Ja	>50
Wine	1991	178	13	Nein	>40
Breast Cancer Wisconsin	1995	569	32	Nein	>35

Tabelle 8 Datensätze für KDD Forschung von der University of California, Irvine

Die Parameteranzahl der in dieser Arbeit untersuchten Daten liegt typischerweise bei mehreren Hundert. Die Anzahl der Datensätze, entsprechend den Zeilen der Formel, kann durchaus sechsstellig werden. Übliche Werte in dieser Anwendung liegen im vier- bis fünfstelligen Bereich.

Die Zeilenanzahl erhöht die Anforderungen an die Software und die Rechenleistung und Speicherausstattung der verwendeten Systeme. Inhaltlich schwieriger ist der Umgang mit der großen Spaltenanzahl. Für eine sinnvolle Extraktion der Regeln nach denen das Neuronale Netz funktioniert ist es –wie in Kapitel 7 ausführlich belegt wird– notwendig, die Parameteranzahl zu reduzieren. Trotzdem ist es zwingend erforderlich, dass sowohl die Feature Selection als auch das Neuronale Netz mit den sehr großen Datenmengen sicher funktionieren.

Den Stand der Wissenschaft und Technik –wie in den bisherigen Kapiteln dargestellt– zusammenfassend sind bei der Behandlung von Ausnahmesituationen in der Fertigung und bei dem gefertigten Material viele Prozesse und Systeme in der industriellen Anwendung in Betrieb. Diese Systeme sind, wie gezeigt, fast durchgängig reaktiv und symptomatisch. Das Auffinden der Ursachen für Abweichungen von Prozessen wird durch die vorhandenen Werkzeuge nur sehr unvollständig abgedeckt. Da diese Werkzeuge im Wesentlichen die möglichst genaue Wiederholung von als richtig definierten Abläufen unterstützen, sind sie für die zielgerichtete Verbesserung des Gesamtprozesses nicht wirklich geeignet. Es gibt vielfältige Werkzeuge, die eine lokale Optimierung nach Fertigungsparametern unterstützen. Für die Suche nach dem globalen Optimum einer Fabrik gibt es praktisch keine Werkzeuge und Verfahren. In der jüngeren Vergangenheit wurde in mehreren herstellerübergreifenden Projekten versucht, die vorhandenen Werkzeuge zu verbessern oder durch bessere zu ersetzen. Wie in diesem Kapitel gezeigt wurde, wurden für die Fragestellung der Ursachenfindung und der Qualitätsverbesserung keine vollständig erfolgreichen Lösungen präsentiert. Auch die Untersuchung der Wissenschaftlichen Veröffentlichungen zeigt keine Lösung für Suche nach den Ursachen für die Prozessabweichungen in den zur Verfügung stehenden Daten auf. Daher ist es notwendig, einen neuen Weg zu beschreiten. Es wird im Folgenden ein System vorgeschlagen, dass aus einer Kombination eines Klassifikationsverfahrens zur Datenreduktion (Feature Selection) sowie einem Neuronalen Netzwerk besteht. Dieses Verfahren soll die Ursachen in den beobachteten Prozessabweichungen in den bereitstehenden Daten formulieren, ohne dass dem System Informationen über die Fertigungsverfahren und die physikalischen und technischen Zusammenhänge zwischen den verschiedenen Parametern bereitgestellt werden.

# 7 Crystal Ball

Die bisher beschriebenen Verfahren zielen primär darauf ab, bereits erfolgreich durchgeführte Schritte in hinreichend enger Toleranz zu reproduzieren. Häufig aber keinesfalls immer sind die Auswirkungen von Grenzwertüberschreitungen auf das fertige Produkt bekannt. Noch weniger klar sind typischerweise die Auswirkungen von Prozessschwankungen innerhalb des zulässigen Bereiches. Hier ist auch zu berücksichtigen, dass eine derartige Abhängigkeit –so diese identifiziert wurde– zu einer produktbezogenen Veränderung der Spezifikation für einzelne Herstellungsschritte führt. Derartige Vorgänge haben mehrere nachteilige Konsequenzen:

**Aufwand I:** Klarerweise wird ein Toleranzbereich für Fertigungsschritte so groß wie möglich gewählt. Eine Verkleinerung des Toleranzbereiches führt zu höherem Ausschuss und / oder zu höherem Fertigungsaufwand. Gegebenenfalls müssen zusätzliche Maschinen angeschafft werden, die Wartungsintervalle verkürzt werden oder die Fertigung einzelner Produkte wird auf Teile des Maschinenparks eingeschränkt.

**Aufwand II:** Veränderung der Prozessschritte, auch wenn diese nur gering sind, erhöhen den Fertigungsaufwand, da dann pro Produkt unterschiedliche Rezepte auf die Fertigungsmaschinen geladen werden müssen. Zudem führt das zu Produktivitätsverlusten da je nach Maschinentyp mehrere verschieden Lose in einer Anlage (Beispiel: Horizontalofen)

gemeinsam prozessiert werden können. Unterscheiden sich die Rezepte ist dieses eben nicht möglich und der Anlagennutzungsgrad sinkt.

**Aufwand III:** Jedes Rezept muss überprüft und freigegeben werden. Zudem muss jede neue Erkenntnis auf mögliche Einflüsse für die verwendeten Herstellungsrezepte geprüft werden. Gibt es durch diese neue Erkenntnis Anpassungsbedarf in einzelnen Prozessschritten, so müssen alle betroffenen Rezepte angepasst und wiederum freigegeben werden.

**Risiko:** Jede Sonderablauf erhöht die Komplexität der Fertigung. Damit steigt auch das Risiko der Fehlbearbeitung an.

**Regelung:** Die beschriebenen Abläufe der Prozessregelung setzen auf statistischen Verfahren auf. Durch eine produktbezogene Veränderung des Fertigungsprozesses werden systematische Effekte in die für die Regelung benutzten Parameter eingebracht. Damit wird die Regelungsgüte wesentlich verschlechtert, da die Eingangsparameter mit systematischen Effekten verunreinigt sind. Alternativ muss für die unterschiedlichen Verfahren eine getrennte Regelung aufgesetzt werden. Dieses verschlechtert aufgrund der geringeren Anzahl von Parametern aber ebenfalls die Regelungsgüte. Hinzu kommt, dass bei Produkten die in geringen Stückzahlen produziert werden, schon die minimale Anzahl von Parametern einen großen Zeitraum überspannt. Die in diesem Zeitraum durchgeführten Weiterentwicklungen am gesamten Fertigungsprozess sind ebenfalls wieder systematische Störungen der Regelung.

Die beschriebenen Lücken in der gegenwärtigen Prozessführung sind nicht einfach schließbar. Eine vollständige Simulation des Fertigungsprozesses einschließlich aller Parameter ist nicht durchführbar. Also ist der Ausweg die Verbesserung der Analysemöglichkeiten. Die bisher eingesetzten Verfahren (siehe 4.4, Seite 81) zur Analyse und zur Prozessverbesserung setzen mindestens implizit den klassischen Versuchsansatz voraus: Es gibt eine These über Zusammenhänge und diese wird dann verifiziert oder falsifiziert. Dieses gilt klarerweise für die Analyse durch den menschlichen Spezialisten.

Aber auch die vielfältigen benutzten statistischen Verfahren dienen der Überprüfung einer vorher aufgestellten These, auch wenn die handelnden Personen sich das nicht in allen Fällen bewusst machen. Dieser beschriebene Ansatz ist erwiesenermaßen in vielen aber leider nicht in allen Fällen zielführend. Die aufgezeigten Anstrengungen von Forschung

und Industrie haben das beschriebene Problem bislang nicht lösen können.

### **7.1 Mathematische Problembeschreibung**

Mathematisch betrachtet, kann man den Herstellungsprozess eines Schaltkreises als allgemeine Funktion darstellen. Die Parameter dieser Funktion sind alle Fertigungsdaten. Das Resultat dieser Funktion sind entsprechend relevante Ergebnisse der Fertigung wie Ausbeute, Durchlaufzeit, Langzeitqualität.

$$\vec{\Lambda} \begin{pmatrix} \textit{Materialien} \\ \textit{Rezepte} \\ \textit{Maschinen} \\ \textit{Parameter} \\ \textit{etc} \end{pmatrix} = \vec{\Theta} \begin{pmatrix} \textit{Ausbeute} \\ \textit{Qualität} \\ \textit{Durchlaufzeit} \\ \textit{etc} \end{pmatrix}$$

Abbildung 45 Allgemeine Herstellungsfunktion einer einzelnen Fertigungseinheit

Einen Sonderfall stellen hier die Prüfergebnisse der einzelnen Bauelemente sowie der Ausgangskontrolle der Fertigung dar. Einerseits sind diese die Ergebnisse des Fertigungsprozesses, gehören also in den Ergebnisvektor  $\Theta$ . Andererseits ergibt sich aus den Messwerten die Ausbeute und sie geben Hinweise auf mögliche qualitative Probleme. Damit gehören diese in die Abbildungsfunktion  $\Lambda$ . Damit können diese Werte auf beiden Seiten der Gleichung auftreten.

Weitere Betrachtung zeigt, dass ja die Bauteile nicht einzeln gefertigt werden. Je nach Fokus der Untersuchung müssen

- die Zusammenfassung von mehreren Bauteilen zu einem Bildfeld,
- mehreren Bildfeldern zu einem Wafer,
- mehreren (typischerweise 25) Wafern zu einem Fertigungslos (Charge),
- mehreren Losen eines Produktes oder sogar
- vielen Losen unterschiedlicher Produkte

betrachtet werden. In allen Fällen gibt es Abhängigkeiten mit anderen gefertigten Produkten. Deshalb muss die in Abbildung 45 gezeigte Funktion dahingehend erweitert werden, dass die genannten Abhängigkeiten abgebildet werden.

$$A \begin{pmatrix} \vec{\Lambda}_1 \\ \vec{\Lambda}_2 \\ \vdots \\ \vec{\Lambda}_n \end{pmatrix} = \begin{pmatrix} \vec{\Theta}_1 \\ \vec{\Theta}_2 \\ \vdots \\ \vec{\Theta}_n \end{pmatrix}$$

Abbildung 46 Allgemeine Funktion mehrerer gefertigten Einheiten

Mit Abbildung 46 lässt sich das Problem formal durch eine Abbildung einer Matrix mittels der Funktion A auf eine Ergebnismatrix beschreiben.

Zieht man nun die Funktionalitäten in  $\Lambda$  mit den Funktionalitäten in A zusammen, lässt sich die verallgemeinerte Abbildung wie folgt darstellen. Alle Parameter werden im Weiteren mit P, alle Ergebnisse mit R bezeichnet.

$$M \begin{pmatrix} P_{11} & P_{12} & \cdots & P_{1m} \\ P_{21} & P_{22} & \cdots & P_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ P_{n1} & P_{n2} & \cdots & P_{nm} \end{pmatrix} = \begin{pmatrix} R_{11} & R_{12} & \cdots & R_{1o} \\ R_{21} & R_{22} & \cdots & R_{2o} \\ \vdots & \vdots & \ddots & \vdots \\ R_{n1} & R_{n2} & \cdots & R_{no} \end{pmatrix}$$

Abbildung 47 Allgemeine Funktion für die Fertigung

Für die in Abbildung 47 dargestellte Funktion muss das Optimum bestimmt werden.

Nachteilig ist:

- Es ist zu keiner Zeit klar, ob alle relevanten Parameter  $P_{nm}$  für die Funktion bekannt sind.
- Eine geschlossene analytische Darstellung dieser Funktion gibt es nicht.
- Die Ergebnisse  $R_{no}$  sind auch nur in Teilen direkt bekannt. Insbesondere Fragen zur Lebensdauer sind nicht direkt entscheidbar.

In der Anwendung geht es um die Lösung eines konkreten Problems oder um die Optimierung eines Wertes. Für die obige Gleichung bedeutet dieses, dass für die Anwendung der Ergebnistensor auf einen Vektor

reduziert wird. Damit lässt sich ebenfalls ohne Beschränkung der Allgemeinheit voraussetzen, dass die Parameter des Ergebnisvektors  $\vec{R}$  nicht ebenfalls Parameter der Funktion  $\vec{F}$  sind.

Damit stellt sich die Funktion wie in Abbildung 48 gezeigt dar:

$$\vec{F} \begin{pmatrix} P_{11} & P_{12} & \cdots & P_{1m} \\ P_{21} & P_{22} & \cdots & P_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ P_{n1} & P_{n2} & \cdots & P_{nm} \end{pmatrix} = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{pmatrix}$$

Abbildung 48 Konkrete Funktion

Die Lösung des beschriebenen Problems ist damit eine Optimierungsaufgabe für die dargestellte Funktion. Eine analytische Darstellung dieser Funktion ist nicht erreichbar. Für die konkrete Situation in einer komplexen Fabrik ist die Gleichung immer unterbestimmt, da wie be-

$$\vec{C} \begin{pmatrix} P_{11} & P_{12} & \cdots & P_{1m} \\ P_{21} & P_{22} & \cdots & P_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ P_{n1} & P_{n2} & \cdots & P_{nm} \end{pmatrix} = \begin{pmatrix} C_1 \\ C_2 \\ \vdots \\ C_n \end{pmatrix}$$

Abbildung 49 Abbildungsfunktion auf Klassen

reits ausführlich gezeigt nicht alle Einfluss nehmenden Faktoren bekannt sein können. Hinzu kommt noch die Unschärfe aus den Messfehlern. Um das Problem einer Lösung näher zu bringen, werden die genannten Ergebnisse  $R_1$  bis  $R_n$  klassifiziert. Damit ergibt sich das in Abbildung 49 dargestellte Klassifizierungsproblem.

Dieses wird im Weiteren betrachtet.

## 7.2 Neuer Ansatz: NN als Validator der Feature Selection

In der High Tech Industrie werden enorme Anstrengungen unternommen, um die Qualität der hergestellten Produkte zu steigern und die Effizienz und damit den wirtschaftlichen Erfolg der Unternehmen zu si-

chern. Trotz der unbestreitbaren Erfolge der in den bisherigen Kapiteln beschriebenen Methoden und Verfahren ist nach wie vor Bedarf an weiterer Optimierung vorhanden. Insbesondere die Identifikation von Ursachen anstelle der Korrektur von Symptomen ist häufig nicht hinreichend erfolgreich. Diese Identifikation der Ursachen und damit deren Beseitigung ist sicher ein wesentlicher Schlüssel für den zukünftigen Erfolg der Unternehmen. Nur die tatsächliche Beseitigung der Ursachen verhindert Wiederholungen von unerwünschten Effekten zuverlässig. Weiterhin erhöht jede gefundene Ursache für ein Problem das Verständnis der in den Unternehmen durchgeführten Produktionsprozesse und ermöglicht damit auch eine bessere Fehlervermeidung. Sowohl die Qualitätsanforderungen der Kunden als auch der beständige Kostendruck zwingt die Unternehmen, die eigenen Prozesse nachhaltig zu verbessern.

In Kapitel 6 wurden verschiedene Ansätze von Hochschulen, Instituten und der Industrie beschrieben, die die Verbesserung der Prozessführung sowie die Steigerung von Qualität und Effizienz der Unternehmen zum Ziel hatten. Auch diese Schritte zeigten Erfolge, blieben aber wie beschrieben primär reaktiv und symptombezogen.

Daher wird in dieser Arbeit ein alternativer Weg vorgeschlagen. Der neue Ansatz besteht darin, auf Basis der vorhandenen Daten mit einem automatischen System konkrete Vorschläge für Ursachen von Abweichungen zu erzeugen.

Für einen erfolgreichen Aufbau eines solchen Systems ist es zunächst notwendig, brauchbare und möglichst umfassende Datenquellen bereitzustellen. Die auszuwertenden Daten müssen in belastbarer Relation zueinander stehen (s. beispielsweise 7.4.4). Eine Darstellung analog der in Abbildung 49 dargestellten Formel ist für die erfolgreiche Auswertung notwendig.

Das vorgeschlagene System besteht im Kern aus der Kopplung eines künstlichen Neuronalen Netzwerks und einer Feature Selection. Die Ziel-spezifikation der Software Architektur wurde 2007 [KMIS07] vorgestellt. Weitere Publikationen des Autors über diesen neuen Ansatz erfolgten unter anderem auch in [KMCM07, ZUE08, DMIN08].

Möglich ist dieser Ansatz, da die Leistungsfähigkeit von preiswerten Computern sehr große Fortschritte gemacht hat. Insbesondere bei den sehr umfangreichen Datenmengen wird für die Operationen auf den Da-

ten sehr viel Hauptspeicher benötigt. Weiterhin gibt es sowohl bei den Hochschulen als auch bei den Unternehmen die Bereitschaft, gemeinsam neue Dinge im Umfeld des Wissensmanagements anzugehen. Dieses wurde beispielsweise in auf der Panel Discussion der IRI deutlich [IRI07].

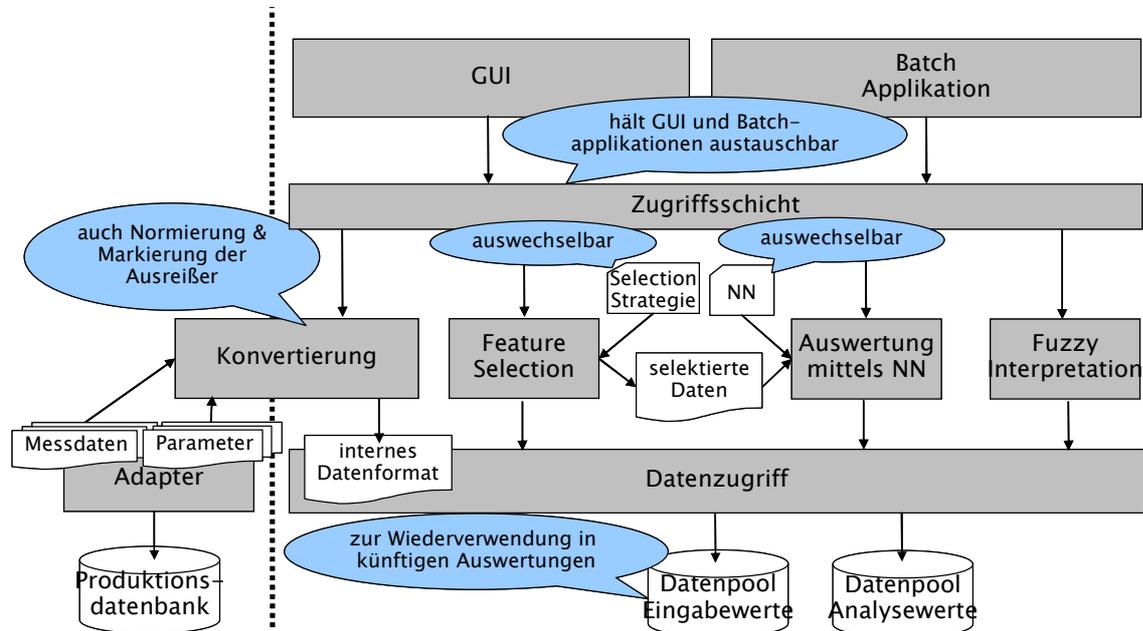


Abbildung 50 Zielspezifikation Software Architektur [nach KMIS07]

Wie in Abbildung 50 dargestellt ist die ursprüngliche Datenquelle die Produktionsdatenbank. Nach Datenkonvertierung finden dann Feature Selection und die Auswertung mittels Neuronaler Netze statt. Die Datenzugriffsebene steht den Werkzeugen ebenso gemeinsam zur Verfügung wie die Zugriffsschicht. Die Zugriffsschicht ermöglicht die Abstraktion der Benutzerinteraktion. Damit ist das Werkzeug sowohl interaktiv als auch automatisiert im Batchbetrieb nutzbar. Der Prozessablauf und Details zu den einzelnen Punkten wird im Folgenden näher erläutert.

Alternativ zu den bisher verfolgten Ansätzen wird hier keinerlei Wissen über die physikalisch-technischen Zusammenhänge der Abläufe integriert. Dies geschieht, um mögliche neue Lösungen zu ermöglichen, die sonst durch unter Umständen falsche oder unvollständige Annahmen ausgeschlossen würden. Weiterhin können völlig beliebige Daten für die Analyse herangezogen werden. Die Auswertung ist also nicht beschränkt auf Messwerte sondern kann auch alle anderen in der Fertigung und deren Umfeld vorhandene Daten nutzen. Der prinzipielle Ablauf des Verfahrens wurde vom Autor auf der DMIN'08 vorgetragen [DMIN08].

## 7.2.1 Ablauf des Verfahrens

In der folgenden Abbildung 51 ist der Ablauf der Analyse in einer Übersicht dargestellt. Eine Beschreibung der einzelnen Schritte befindet sich auf den folgenden Seiten.

Ausgehend von der Definition des zu lösenden Problems (①) wird durch den Bearbeiter festgelegt, welches die zu analysierenden Objekte sind(②). Dabei kann es sich beispielsweise um einzelne Dies, einzelne Bildfelder, ganze Scheiben oder ganze Lose handeln(siehe Abschnitt 7.4.2 )

### A Datenvorbereitung

Mit diesen Voreinstellungen werden die Daten aus den entsprechenden Datenquellen extrahiert (③). Je nach benötigten Datentypen und dem Ausbau der in der jeweiligen Fabrik vorhandenen Fertigungsleitsysteme kann dieser Schritt einfach bis sehr aufwändig sein (siehe Abschnitt 7.5).

Im besten Fall stehen in den Vorsystemen die ungefilterten Rohdaten zur Verfügung. Dann müssen diese nun von offensichtlichen Fehlmesungen oder ansonsten falschen Daten bereinigt werden(④). Stehen die Rohdaten zur Verfügung, können die Kriterien für die Ausreißer-Selektion im weiteren Verlauf der Analyse gegebenenfalls modifiziert werden um die Ergebnisse zu verbessern. Stehen nur bereits verarbeitete Daten zur Verfügung, sind die Möglichkeiten entsprechend eingeschränkt.

Bei der folgenden Normalisierung (⑤) der Daten werden diese auf rationale Zahlen im Wertebereich 0 - 1 abgebildet. (siehe Abschnitt 7.6.2 ).

Der vorletzte Schritt der Vorbereitung der Daten für die Analyse ist die Bereinigung der Daten (⑥). Hier werden widersprüchliche Daten entfernt und -abhängig von der gewählten Granularität und den gewählten Datentypen- die gegebenenfalls leeren Positionen  $P_{nn}$  in Abbildung 49 sinnvoll gefüllt (siehe 7.6.4 ).

Schließlich werden die vorliegenden Ergebnisdaten in Klassen eingeteilt (⑦). Hier ist letztmalig eine inhaltliche Vorgabe des Bearbeiters notwendig, da die sinnvolle Klasseneinteilung von dem gesuchten Problem abhängt (siehe 7.7.5 ).

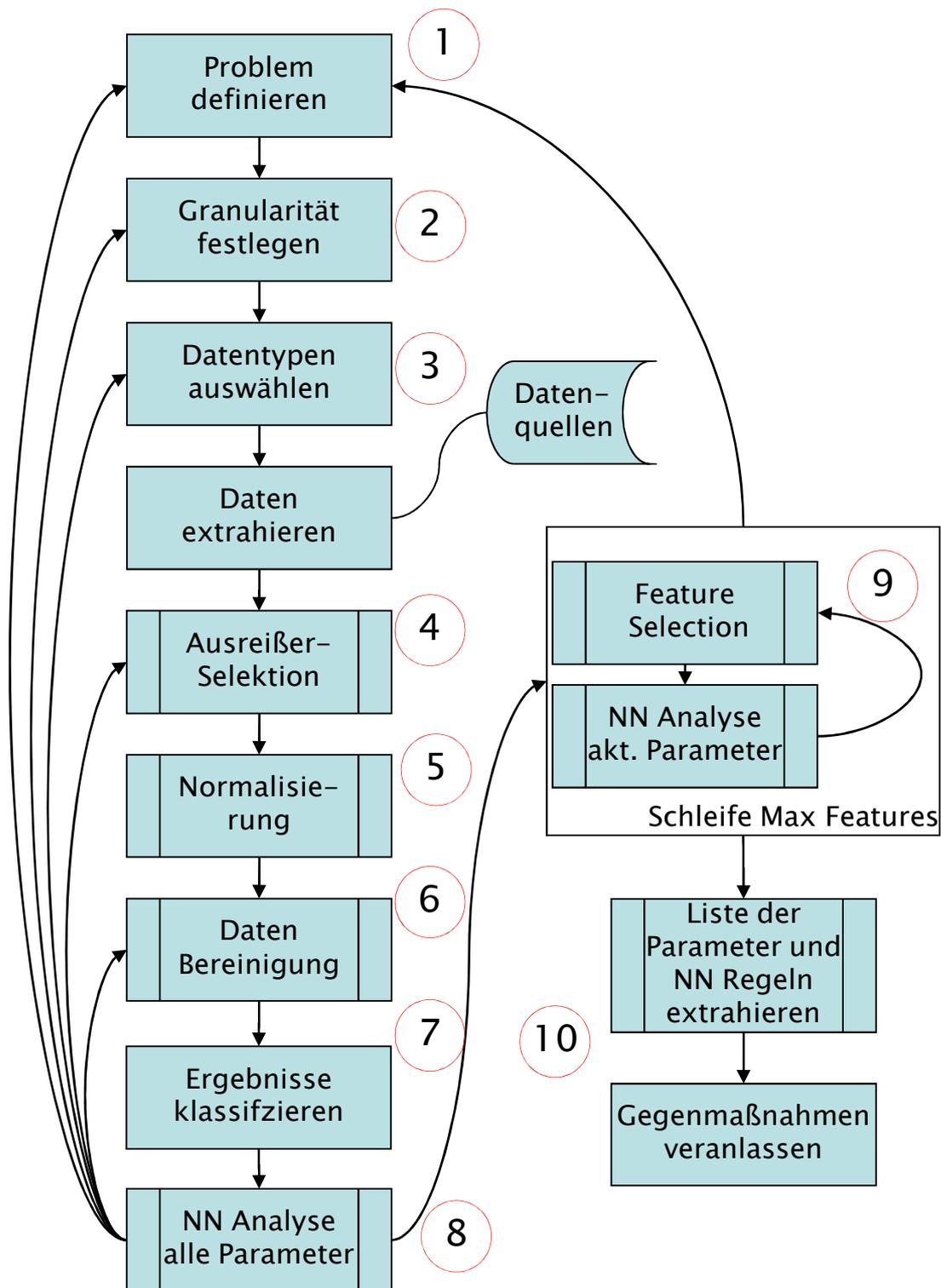


Abbildung 51 Darstellung des Ablaufes

## B Inhaltliche Validierung der Daten

Damit ist die Vorbereitung der Daten beendet. Die Daten liegen entsprechen Funktion in Abbildung 49 vor. Der nächste Schritt ist die Analyse der gesamten Daten mit dem Neuronalen Netz (Abbildung 51, ⑧). Dazu werden die Daten zeilenweise und typischerweise unsortiert in einen Trainings und einen Validierungsdatensatz getrennt. Das Neuronale Netz wird mit den Trainingsdaten (einschließlich der Ergebnisklassen) angelernt. Danach werden mit dem trainierten Netzwerk für die Validierungsdaten jeweils Ergebnisklassen  $V_n$  ermittelt. Details zum verwendeten Neuronalen Netz und zum Einsatz des Netzes finden sich in Abschnitt 7.8 ab Seite 150.

Über den Vergleich mit den bekannten tatsächlichen Ergebnisklassen  $C_n$  ergibt sich die Güte der Abbildung der Realität im Netz. Nur wenn die Validierung positiv endet, ist der gesuchte Effekt mit hoher Wahrscheinlichkeit in den betrachteten Daten enthalten. Nur in diesem Fall ist die Weiterführung des Ablaufs zu Punkt ⑨ in Abbildung 51 sinnvoll. Andernfalls ist –wie in der Abbildung über die Pfeile gekennzeichnet– eine Änderung der Datenaufbereitung und ein erneuter Durchlauf durch die jeweils folgenden Schritte so lange notwendig, bis die Validierung positiv abgeschlossen werden kann.

## C Datenabstraktion durch Feature Selection

Nach erfolgreichem Abschluss von Punkt ⑧ ist die Ursache für den untersuchten Effekt mit hoher Wahrscheinlichkeit in den Daten enthalten und in den Regeln des Neuronalen Netzwerkes hinterlegt. Wie in Abschnitt 7.8 weiter ausgeführt wird, wird hier ein Netzwerktyp verwendet, dessen Regeln extrahiert werden können. Wie mehrfach dargelegt, werden hier sehr hochdimensionale Daten verarbeitet. So kann ein Datensatz leicht mehrere hundert Parameter (Spalten des Eingangstensors) umfassen. Eine extrahierte Regel zu diesem Zeitpunkt enthält entsprechend mehrere hundert Terme. Das Ziel der Erkennung von Ursachen für Abweichungen ist mit einer derartigen Regel nicht zu erreichen, da nicht klar ist, welche von den Termen ursächlich für den gesuchten Effekt sind und welche nicht.

Um den Parametersatz zu reduzieren ohne relevante Information für die aktuelle Untersuchung zu verlieren, wird ein zweites Verfahren mit diesen Daten genutzt.

Mit der hier eingesetzten Feature Selection wird die Spaltenanzahl des Eingangstensors einstellbar reduziert. Neu ist hier, dass das Ergebnis der Feature Selection mit dem neuronalen Netz überprüft wird. Die Entscheidungsinstanz in diesem Verfahren ist, wie schon unter ⑧ das Neuronale Netz. Die Feature Selection ermöglicht durch die reduzierte Anzahl von Parametern die Nutzung der durch das Neuronale Netz gefundenen Regeln.

Unter ⑨ wird in Abbildung 51 der Kernprozess des Verfahrens gezeigt. Die Feature Selection wird –wie in 7.7 beschrieben– mit verschiedenen Werten für die maximale Featureanzahl gestartet. Für jeden dieser Werte gibt es als Ergebnis einen Satz von Parametern, die die Feature Selection als relevant ausgewählt hat. Für jeden dieser Sätze wird die in 7.2.1 .B beschriebene Analyse der Daten mit dem Neuronalen Netz durchgeführt. Da es sich in jedem Falle um einen anderen Datensatz anderer Dimension handelt, muss der Ablauf Training / Validierung jedes Mal vollständig durchlaufen werden. Für alle diese Parametersätze wird die Quote der korrekt zugeordneten Klassen bestimmt und ausgegeben.

Neu an diesem Ansatz ist die direkte Kopplung zwischen Neuronalem Netz und Feature Selection, insbesondere für diese sehr großen Datenmengen und die sehr hohen Qualitätsansprüche in der Anwendung.

Mit diesem Ansatz ist es möglich, ähnlich der Näherungsrechnung, sich iterativ der richtigen Parameteranzahl zu nähern. Durch die Kontrolle des zweiten Verfahrens gibt es jederzeit die Information, ob der gerade beschrittene Weg noch der Richtige ist. Sind auf diesem Wege keine brauchbaren Ergebnisse zu erzielen, obwohl die Vorabprüfung mit dem vollen Datensatz gute Ergebnisse gezeigt hat, muss das Verfahren bei der Fragestellung(①) neu gestartet werden. In diesem Fall ist zu vermuten, dass sich die ursächlichen Parameter nicht im Datensatz befinden und das Verfahren viele abgeleitete Parameter findet.

## D Nutzung der Ergebnisse

Der Bearbeiter wählt nun anhand der Quoten der korrekt zugeordneten Klassen eine Feature Anzahl, die eine hohe Trefferquote hat (Ⓢ). Das wird im Allgemeinen nicht das absolute Maximum der Trefferquoten sein, da diese für den vollständigen Parametersatz zu erwarten ist. Real existierende Problem hängen aber normalerweise nicht von einer hohen Anzahl von Parametern ab, daher wird der gewählte Wert im kleinen einstelligen Bereich liegen. Für den Bearbeiter interessant ist aber durchaus, welche Parameter hinzugenommen wurden und welche entfernt wurden.

Für den gewählten Parametersatz werden dann die Regeln extrahiert und die Gegenmaßnahmen in den Prozessen eingeleitet. Damit ist der Ablauf erfolgreich beendet.

### 7.2.2 Systemübersicht: Datenkonzentration

Abbildung 52 zeigt eine Übersicht über das Gesamtsystem hinsichtlich des Datenflusses. Die für die Analyse zur Verfügung stehenden Daten werden in den existierenden Fabriken häufig von unterschiedlichen Organisationseinheiten mit unterschiedlichen Systemen für unterschiedliche Auswertungen erhoben.

Engineering Data: Wie schon in Kapitel 4.4 auf Seite 81ff ausgeführt, werden die während der Fertigung und im Prüffeld anfallenden Messwerte typischerweise mit spezialisierten Werkzeugen ausgewertet. Häufig werden dabei auch mit den Werkzeugen gelieferte Datenbanken genutzt.

Logistic Data: Wie in Kapitel 2.8 ab Seite 36 beschrieben ist die Bereitstellung des aktuellen Materialstandes und damit auch die Dokumentation bereits am Material durchgeführter Schritte Kernfunktionalität eines jeden MES und in jeder Halbleiterfabrik vorhanden. In vielen Fällen sind die zu Grunde liegenden Datenstrukturen, wie bei den anderen genannten Systemen auch, proprietär.

Service Data: Durch die in sehr vielen Installationen vorhandene Anlagenautomatisierung (s. 2.9, 45ff) werden mindestens die Rezeptdaten häufig gemeinsam mit den Logistikdaten verwaltet. Für die Überwa-

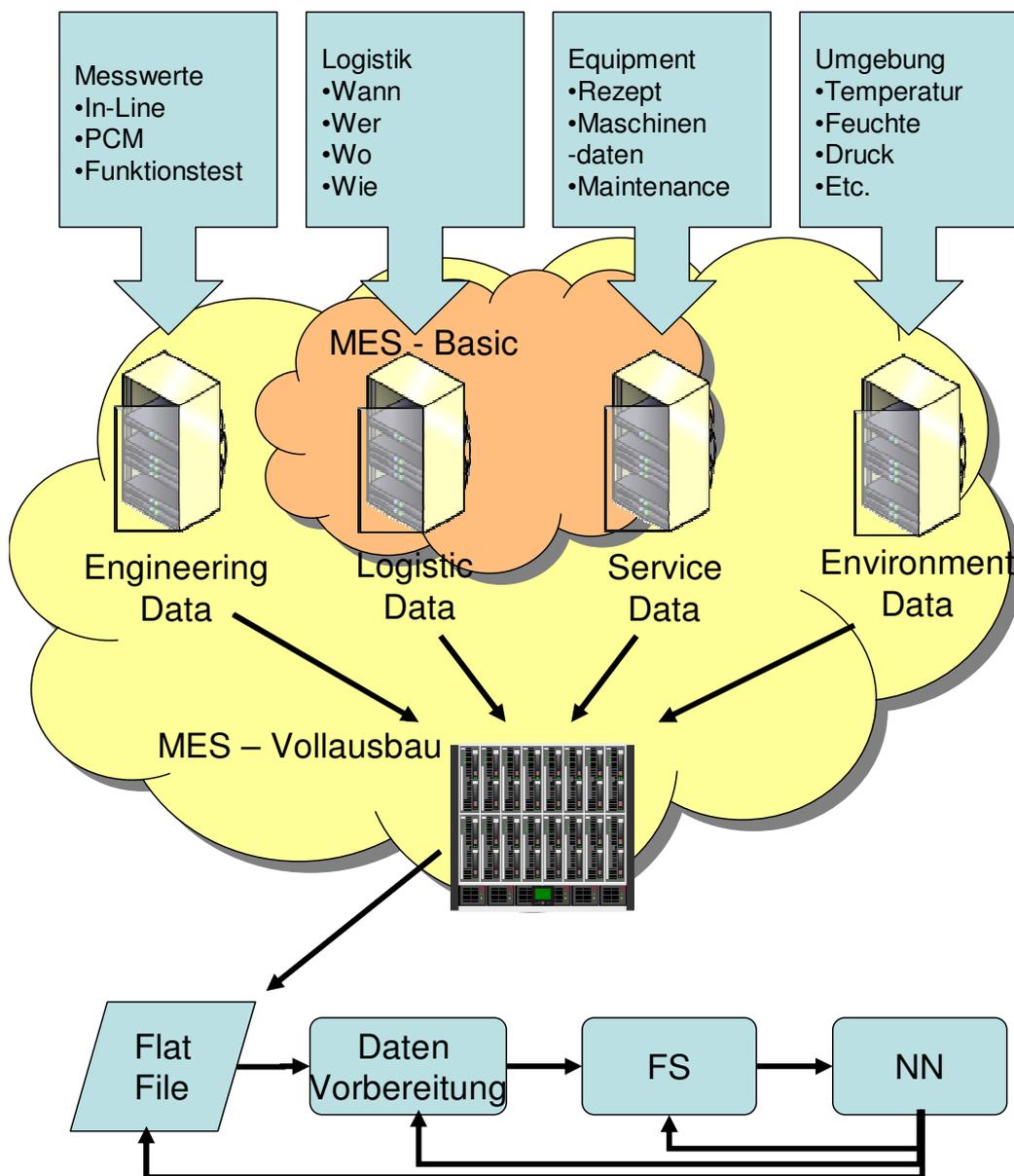


Abbildung 52 Übersicht Gesamtsystem

chung der Maschinenparameter und die Wartung der Maschinen werden häufig wiederum spezielle Systeme mit eigener Datenverwaltung eingesetzt.

Environment Data: Die Kontrolle und Regelung der Umgebungsparameter in der Fertigung wird –ähnlich wie im Falle der Messdaten– von anderen Organisationseinheiten verantwortet. Entsprechend sind hier häufig wiederum proprietäre Systeme –diesmal aus dem Umfeld der Gebäudeleittechnik– im Einsatz.

Um übergreifende Auswertungen zu ermöglichen, müssen die Daten zusammengeführt werden.

Die vorhandene Datenintegration wird mit den Wolken skizziert. Bei vollständiger Konsolidierung der Datensysteme beschränkt sich der Aufwand der Zusammenführung im Idealfall auf Datenbankabfragen (MES-Vollausbau). Sind sehr viele verschiedene Systeme im Einsatz (MES-Basic), ist der Aufwand für die Zusammenführung der Daten unter Umständen sehr erheblich. Dies ist insbesondere dann der Fall, wenn diese Systeme jeweils eigene Stammdaten pflegen.

### ***7.3 Benutzerschnittstelle***

Die Wahl der Benutzerschnittstelle für das System erfolgte unter folgenden Kriterien:

- Alle Systemteile sollten möglichst einfach automatisierbar sein.
- Die Systemteile sollten ebenfalls einfach mit einer grafischen Benutzeroberfläche versehen werden können.
- Das zugrunde liegende Computersystem muss wegen der hohen Anforderungen an die Systemleistung und insbesondere den Speicherplatz für die Verwaltung großer Speichermengen geeignet sein.
- Auf Grund des vorgesehenen Produktionseinsatzes ist eine sehr lange Lebensdauer des Systems notwendig.
- Die Nutzung sollte ohne relevante Einschränkungen auch Remote über LAN<sup>60</sup> oder WAN<sup>61</sup> möglich sein.

Unter Berücksichtigung der Randbedingungen fiel die Entscheidung auf Unix als Betriebssystem. Im ersten Schritt wurde die Entwicklung auf einem 32-Bit Linux System durchgeführt.

---

<sup>60</sup> Local Area Network: auf einen Standort beschränktes, typischerweise IP basiertes, kabelgebundenes Netzwerk.

<sup>61</sup> Wide Area Network, Weitverkehrsnetz: Verbindung von mehreren LANs über größere Entfernungen, sowohl kabelgebunden als auch drahtlos.

Aus den vorgenannten Gründen wurde die Benutzerschnittstelle als reine ASCII<sup>62</sup> Schnittstelle ausgeführt. Bei der Entwicklung wurden die Paradigmen der Softwareentwicklung unter Unix eingehalten, so dass eine Automatisierung der Nutzung über Scripten genauso möglich ist wie die Ansteuerung durch eine grafische Benutzeroberfläche, bevorzugt aus einem Webbrowser.

### ***7.4 Datenquelle***

Das Gesamtsystem setzt auf einem relationalen Datenbanksystem auf. Da, wie schon aufgezeigt, die Konsistenz der Daten für jede Auswertung essentiell ist, ist dieses Datenbanksystem die einzige unterstützte Datenquelle. Konkret handelt es sich bei der Quelle für zu analysierenden Daten um ein Data Warehouse basierend auf einer Oracle Datenbank (Version 10). Die Datenbank läuft auf einem mittelgroßen Server (4 Dual Core CPU, 8 G Speicher) unter dem Betriebssystem Sun Solaris (Ebenfalls Version 10).

Für das System wurde ein eigenes Schema innerhalb der Datenbank bereitgestellt. Die für die jeweils zu beantwortende Fragestellung relevanten Daten werden vom Benutzer selektiert.

Zentrale Eigenschaft des Verfahrens ist die Möglichkeit, beliebige Daten zur Auswertung zu nutzen. Wichtig ist natürlich, dass diese hinsichtlich Ihrer Zugehörigkeit zur jeweils untersuchten Ergebnisspalte zugeordnet werden kann.

Ebenso wie auf der Eingangsseite ist das System auch auf der Ergebnisseite praktisch beliebig flexibel. Auch hier können zum Beispiel Messwerte, Ausbeuten aber auch logistische Daten genutzt werden.

---

<sup>62</sup> American Standard Code for Information Interchange: 7 Bit Code für Zeichenkodierung.

### 7.4.1 Voraussetzung für „Generalisierten NN Ansatz“

Das Ziel der möglichst vollständigen Analyse der Eingangsdaten setzt eine homogenisierte Datenquelle voraus. Klarerweise haben Fehler in den Eingangsdaten für die Ergebnisse gegebenenfalls sehr erhebliche Konsequenzen.

Daher dürfen als betrachtete Parameter nur Daten genutzt werden, die die folgenden Bedingungen erfüllen:

- Messfehler und Ausreißer müssen entfernt werden / entfernbar sein.
- Die sichere Zuordnung der Daten untereinander muss gegeben sein.
- Die Daten müssen normalisierbar sein, d.h. es muss eine bijektive Abbildung des Wertebereichs der Daten auf Fließkommazahlen zwischen 0 und 1 möglich sein.

Die notwendigen Vorgänge werden im Folgenden noch beschrieben.

### 7.4.2 Granularität der Analyse

Für eine erfolgreiche Analyse ist es zunächst notwendig, sich für eine Bedeutung einer Zeile aus Abbildung 48 zu entscheiden. Wie beschrieben, liegen die Daten mit sehr unterschiedlicher Granularität vor. Einzelne Messwerte –beispielsweise aus dem Funktionstest– beziehen sich auf ein einzelnes Bauteil auf einer Scheibe einer Charge. Andererseits gibt es –zum Beispiel mit den Umgebungswerten von Temperatur und Druck– ebenfalls Messwerte bei denen nur ein Datum für eine ganze Anzahl von Chargen vorliegt. Es erscheint naheliegend, immer mit der höchstmöglichen Granularität in die Analyse zu gehen. Leider ist dieser naheliegende Ansatz nicht in jedem Fall der beste Weg zum Ziel. Neben den genannten Granularitäten gibt es nämlich durchaus Messwerte, die nur stichprobenhaft erhoben werden (z.B. PCM) Geht man nun mit maximaler Granularität in die Analyse, muss man entscheiden, wie die wenigen parametrischen Messwerte auf der Scheibe (siehe 3.1, Seite 52) den vergleichsweise sehr vielen Ergebnissen aus dem Funktionstest zugeordnet werden sollen. Da die PCM-Teststrukturen Bauelemente sind,

wie alle anderen, haben auch diese eine Ausbeute. Entsprechend kommen natürlich auch bei der PCM Fehlmessungen vor. Dieses verkompliziert die Zuordnung von Funktionstestmesswerten zu PCM-Werten zusätzlich und erheblich. Es ist notwendig möglichst weitgehend alle systematischen Störungen zu unterlassen. Das hier angewandte Verfahren funktioniert wie geplant ohne Zuhilfenahme jeglicher Prozesskenntnis. Entsprechend gering sind die Fähigkeiten, systematische Störungen der Daten zu erkennen. Eine systematische Störung liegt bei fehlerhaften Annahmen und sich daraus ableitenden Aktionen bei der Auswahl und Abbildung der Messdaten in jedem Fall vor.

In diesem Falle kann die unzutreffende Zuordnung von PCM zu Funktionstestdaten durchaus für die folgenden Verarbeitungsschritte falsche Zusammenhänge nahelegen und damit gegebenenfalls sogar richtige Zusammenhänge verdecken.

Es kann also, wie an dem Beispiel PCM-Daten – Funktionstestdaten gezeigt wurde, durchaus sinnvoll sein, nicht mit Messwerten von einzelnen Dies sondern mit den Ergebnissen oder Mittelwerten einer Scheibe in die Analyse zu gehen. Analog gilt dieses auch für Mittelwerte oder Ergebnisse ganzer Chargen, nämlich immer dann, wenn Parameter eine Rolle spielen, die nur für einige Scheiben einer Charge erhoben wurden (Beispiel: Messungen auf Produktionsscheiben in der Fabrik).

Ebenfalls kann es sinnvoll sein, bei der Granularität ein einzelnes Bildfeld zu wählen. Das Bildfeld ist die kleinste einzeln parametrisierte Fertigungsstruktur in einer Halbleiterfabrik.

Die topologische Struktur auf einem Wafer zeigt Abbildung 53. Der blaue Rand stellt den Waferrand dar. Die türkisfarbenen Linien markieren jeweils ein einzelnes Bildfeld, also den Bereich eines Wafers der bei der Belichtung vom Reticle auf die Scheibe abgebildet wird. Die dünnen schwarzen Linien markieren ein einzelnes Bauteil (Die). Wie schon beschrieben, wird der Ritzrahmen für die Erzeugung von PCM-Strukturen genutzt. Typischerweise sind die Inhalte aller Bildfelder identisch. Damit sind in diesem Beispiel die zu messenden PCM-Strukturen in jedem Bildfeld in dem zweiten vertikalen Ritzrahmen innerhalb des Bildfeldes. mit den gelben Linien sind die ausgewählten Messpositionen für die PCM gekennzeichnet.

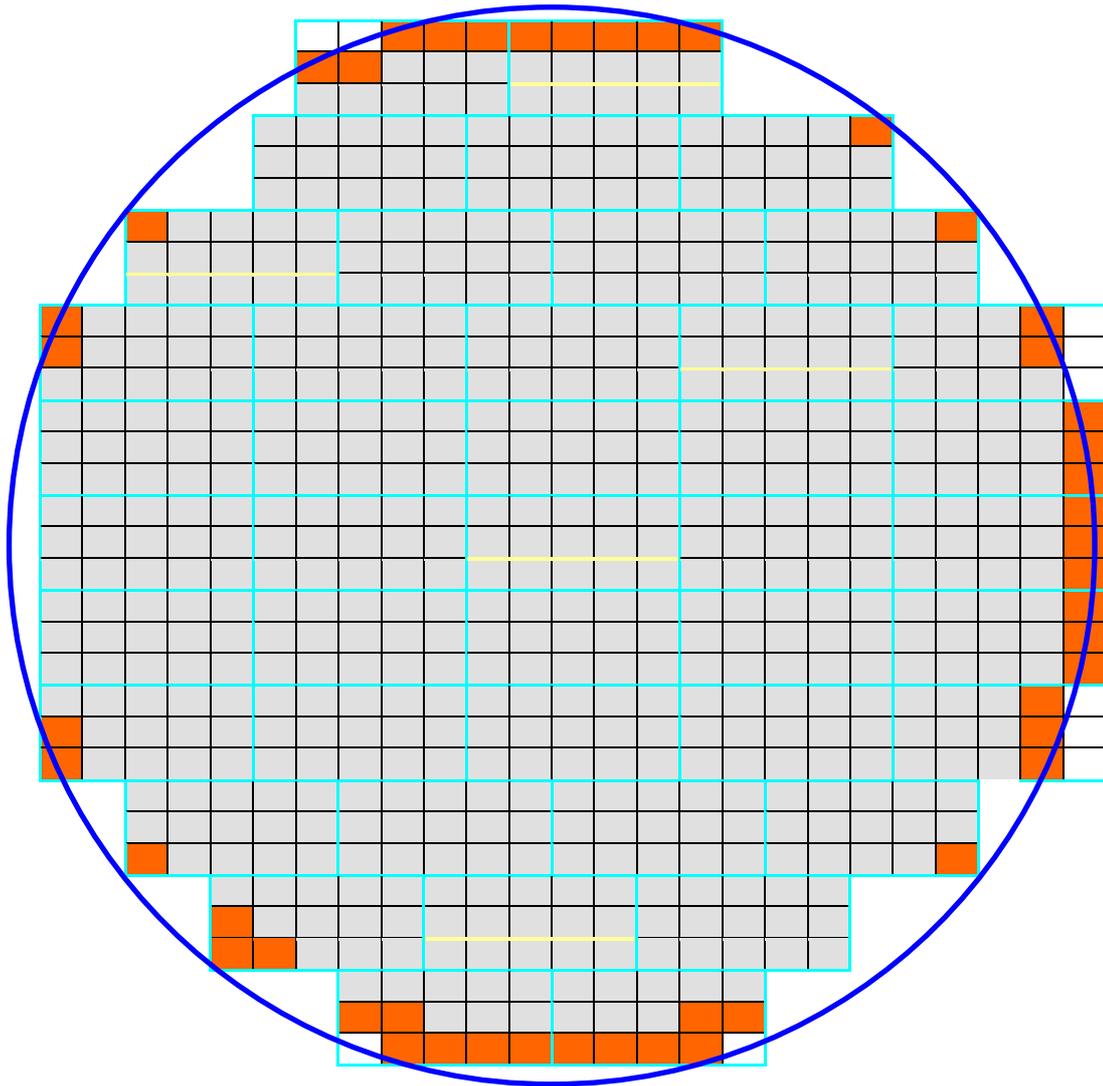


Abbildung 53 Wafermap mit Bildfeldern und PCM Positionen

Zu beachten ist, dass es sich bei dieser Abbildung nicht um eine reale Wafermap aus einer Fabrik handelt. Die tatsächlich in den Fabriken benutzen Layouts sind das Ergebnis vielfältiger Optimierungen die in diesem Beispiel nicht vollständig vorgenommen wurden. So gibt es typischerweise eine Randzone, in der die Bauteile wegen der dort auftretenden Prozessschwankungen nicht verkaufsfähig sind. Weiterhin kommt es, wie auch in dem Beispiel erkennbar, vor, dass einzelne Bauteile nicht ganz vollständig auf der Scheibe abgebildet werden. In beiden Fällen müssen die Bauteile vor der Assemblierung als nicht verkaufsfähig gekennzeichnet werden (geinkt), auch wenn diese funktionsfähig wären. In dem Beispiel sind diese Bauteile orange gekennzeichnet.

Zudem kann normalerweise, wie auch im Beispiel sichtbar, nicht die gesamte Waferfläche genutzt werden. Der Belichter muss seine Justage-

marken auf der Waferoberfläche finden können. Dieses setzt einen gewissen Anteil des Steppingfeldes auf der Waferoberfläche voraus. Zudem kostet jede Belichtung Zeit und damit Geld. Damit muss die Anzahl der Belichtungen auf einer Scheibe und der Optimierung der Verfahrenswege zwischen den Belichtungen der Anzahl der potentiell guten Dies entgegengestellt werden.

### 7.4.3 Chargensplits

Es muss an dieser Stelle darauf hingewiesen werden, dass immer dann, wenn die Zuordnungseinheit Charge erwähnt wird, tatsächlich ein Chargensplit gemeint ist. Ein Fertigungslos besteht in der Halbleiterindustrie –unabhängig vom Scheibendurchmesser– normalerweise aus 25 Wafern. Aus verschiedensten Gründen kommt es während des Fertigungsprozesses vor, dass diese Charge geteilt wird und die beiden Teile im Folgenden unterschiedliche Behandlung erfahren. Solche Splits werden durchaus auch mehrfach an derselben Charge durchgeführt. Minimale Splitgrößen sind häufig 6 Wafer pro Split, wobei in Einzelfällen auch kleinere Splits von 3 Wafern pro Split verarbeitet werden.

Gründe für die Durchführung solcher Splits gibt es verschiedene, sowohl technischer als auch logistischer Art. Beispiele:

- Bei Fertigungsmaschinen mit Einzelscheibenprozessen hängt die Fertigungsdauer von der Anzahl der zu prozessierenden Scheiben ab. Bei extrem eiligen Produkten wird bisweilen ein –dann schneller laufender Teil– abgetrennt.
- Für Neu- oder Requalifikationen von Fertigungsanlagen werden typischerweise Fertigungslose gesplittet. Ein Teil läuft dann über eine freigegebene Fertigungsanlage, der andere Teil über die freizugebende Anlage. Die Prüfergebnisse dieser Lose sind dann die Basis für die Freigabeentscheidung.
- Im Zuge der Prozessentwicklung weiterentwickelte Rezepte werden ebenfalls typischerweise über Splits von Fertigungslosen untersucht. Zur Maximierung der Validität der Ergebnisse werden in diesem und dem vorangegangenen Fall üblicherweise die dem Split folgenden Schritte möglichst identisch durchgeführt um

mögliche Abweichungen auf den zu untersuchenden Schritt zurückführen zu können.

In der industriellen Anwendung ist das Splitten einer Charge immer das Ergebnis einer Abwägung. Durch die dann kleineren Losgrößen erhöht sich der logistische Aufwand in der Fertigung. Zudem sinkt der Anlagennutzungsgrad. Daher wird die Fertigungsleitung immer versuchen, die Splitzanzahl zu minimieren.

In Fabriken mit hohen Anforderungen an Qualität und Rückverfolgbarkeit ist ein Wiederausammenführen (Merge) von Material mit unterschiedlicher Vorgeschichte typischerweise untersagt und wird vom Fertigungsleitsystem konsequent unterbunden.

#### **7.4.4 Koordinatensysteme**

Um die verschiedenen vorliegenden Daten erfolgreich in Zusammenhang bringen zu können, ist es offensichtlich notwendig, diese in ein einheitliches Koordinatensystem zu transformieren. Der Fertigungsprozess beginnt mit unstrukturierten Wafern. Die verschiedenen eingesetzten Systeme in der Fabrik und im Prüffeld nutzen unterschiedliche Koordinatensysteme. Da es keine ausgezeichneten Punkte auf der Waferoberfläche gibt, ist die Koordinatentransformation ein komplizierter Vorgang der hinsichtlich des Nullpunkts des Koordinatensystems nicht vollkommen fehlerfrei sein kann. Die verbleibenden Unschärfe von +/- einem Die muss bei kritischen Erkenntnissen über die Materialqualität berücksichtigt werden. Final bedeutet das, das eine Reihe Dies mehr verworfen wird, um die Auslieferung von potentiell fehlerhaftem Material sicher auszuschließen.

#### ***7.5 Datentypen***

Bisher wurden bei der Arbeit mit dem System die folgenden Daten genutzt.

### 7.5.1 PCM Daten

Beim PCM werden die Messdaten direkt in die Produktionsdatenbank geschrieben. Wegen vorhandenen Beschränkungen des Messsystems werden die Daten nicht relational normalisiert abgelegt sondern in einer breiten Tabelle mit einer Spalte pro Wafernummer. Als erster Schritt für die weitere Auswertung wird eine Normalisierung der Daten durchgeführt so dass die Wafernummer ebenso wie das Bildfeld, die Parameterid und der Messwert ein Parameter ist. Dazu muss die Tabelle aufwändig transponiert werden. Dieses wird durch das System automatisch durchgeführt.

### 7.5.2 Maschinendaten

Im MES (siehe 2.8, S. 36ff) werden die logistischen Daten der Fertigung im Waferprozess erfasst. Durch geeignete Filterung werden der Name des aktuellen Teilschritts, der Name der Fertigungsmaschine, der Name des benutzten Rezeptes sowie Datum und Uhrzeit des Prozessstarts extrahiert. Wie in den anderen Fällen auch, ist es notwendig auf eine saubere und belastbare Stammdatenbasis zurückgreifen zu können. Ohne eine strukturierte Namensgebung der Fertigungsmaschinen und ein zuverlässige Versionsverwaltung bei den Fertigungsrezepten gehen gegebenenfalls wesentliche Informationen verloren oder das System wird durch fehlerhafte Angaben in eine unzutreffende Richtung geleitet.

Sicher gelten die genannten Punkte auch für die klassische Analyse durch den Spezialisten. Für die Auswertung durch einen Automaten sind die Anforderungen allerdings wesentlich höher da dieser die durch den Menschen automatisch vorgenommene Berücksichtigung von Kontextinformationen nicht vornimmt. Für den menschlichen Bearbeiter offensichtliche Fehler kann die Maschine nicht erkennen. Im Gegenzug werden nur tatsächlich vorhandene Daten berücksichtigt.

### 7.5.3 Messwerte Maschinen

Neben den Messwerten am Produkt im Prüffeld werden auch Messwerte in der Halbleiterfertigung erfasst. Dieses sind zum Teil (Defektdaten,

Linienbreiten) Messwerte am Material die diesem auch zugeordnet werden können. Allerdings handelt es sich bei diesen In-Line Messungen praktisch durchgängig um Stichproben. Weiterhin sind auch interne Prozessdaten der Maschine verfügbar (s. 4.1.2 , S. 72). Da diese, wie aufgeführt auch bei Schwankungen innerhalb der zugelassenen Toleranzen Einfluss auf das Material haben können, wurden diese mit in die Auswertung einbezogen. Die Anzahl der internen Parameter der Fertigungsmaschinen ist typischerweise mit teilweise mehr als 10.000 sehr groß. Gleichzeitig sind diese Parameter durch den Anlagenhersteller typischerweise mehrheitlich nicht dokumentiert und schon gar nicht spezifiziert. Daher wurden in die durchgeführten Auswertungen primär die Messwerte am Material und die als relevant bekannten internen Maschinendaten berücksichtigt. Die Nutzung der weiteren erreichbaren Daten steht wie beschrieben zur Verfügung sollte die tiefere Analyse des Verhaltens einer Fertigungsmaschine notwendig sein.

#### **7.5.4 Messwerte Funktionstest**

Hier können sowohl die Messwerte aus dem Funktionstest auf Wafer als auch die Ergebnisse des Funktionstests von Final Parts genutzt werden. Wichtig ist hier zu beachten, dass für die Messwerte von der Final Part Messung üblicherweise keine Topologieinformationen mehr zugeordnet werden können. Es ist also nicht mehr nachzuvollziehen, von welcher Scheibe und von welcher Position auf der Scheibe das aktuell getestete Device stammt. Darüber hinaus ergibt sich auf Grund der Ununterscheidbarkeit des Materials wiederum beim Test der Final Parts die Schwierigkeit endgültige Messwerte zu identifizieren. Üblicherweise werden bei vermuteten Fehlmessungen beispielsweise durch nicht korrekte Kontaktierung die Bauteile nach dem Abschluss der Messung erneut getestet. Die ursprünglichen Messdaten können im Datensatz aber nicht identifiziert werden und müssen daher in diesem verbleiben.

Für den ersten Ansatz der Analyse mit diesem System haben wir die Auswertung üblicherweise mit Chargen oder Wafermittelwerten begonnen und den als kritisch betrachteten Parameter als Ergebniswert definiert.

### 7.5.5 Ausbeutedaten

Die Ausbeutedaten sind natürlich ein sehr wichtiger Punkt in der Betrachtung. Entsprechend wurden diese primär als Ergebnisvektor genutzt. Wie bereits in den vorangegangenen Abschnitten beschrieben, ist auch hier eine sehr solide Datenbasis unabdingbar.

Wenn beispielsweise die Auswertung auf Bildfeldebene stattfinden soll um positionsabhängige Effekte auf der Scheibe zu erkennen, ist es eine zwingende Voraussetzung, die potentiell guten Devices pro Bildfeld zu kennen, Wie im Beispiel in Abbildung 53 deutlich wird, ist dieser Wert keinesfalls konstant. Weiterhin ist es bei einer Ausbeuteanalyse wesentlich, welche möglichen Fehlerquellen gerade im Fokus der Untersuchung stehen. Typischerweise wird in der Halbleiterindustrie das Prüfergebnis eines einzelnen Chips klassifiziert. Diese Klassifizierung wird innerhalb eines Unternehmens über die verschiedenen Produkte weitgehend vereinheitlicht. So ist häufig die Binklasse 1 für die fehlerfreien Bauteile festgelegt. Ebenso typischerweise werden Binklassen für Kontaktierung, Versorgungsstrom, Leckstrom sowie Fehler im Digitalteil definiert. Neben dieser Klassifizierung von Messergebnissen aus dem Funktionstest werden auch sonstige erkannte Fehler oder erwartete Zuverlässigkeitsrisiken über dafür speziell festgelegte Binklassen aussortiert. Hierzu gehören beispielsweise im Rahmen der optischen Inspektion erkannte mechanische Beschädigungen, auch wenn diese nicht zu elektrischer Fehlfunktion führen. Weiterhin werden bereits während des Waferprozesses erkannte Fehler wie Dejustage einzelner Bildfelder, Probleme am Wafer- rand oder ähnliches entsprechend klassifiziert.

Bei der Analyse kann es nun sinnvoll sein, einzelne Fehlerklassen zu betrachten oder andere Fehlerklassen von der Analyse auszuschließen um den Blick für das Wesentliche nicht zu verstellen. Auch hier ist wieder ein sehr präziser Umgang mit den Daten notwendig um keine Artefakte zu erzeugen.

Die mit dem System durchgeführten Analysen betrafen primär die Ergebnisse des elektrischen Tests. Daher wurden die Binklassen für die optische Inspektion und die Klassifizierung aufgrund von Auffälligkeiten in der Fertigung vor der Analyse ausgeblendet.

## ***7.6 Vorbereitung der Daten***

Um die Daten in den folgenden Schritten weiter verarbeiten zu können, müssen zunächst die relevanten Datensätze selektiert werden. Anschließend erfolgt die Normalisierung einschließlich der Ausreißerselektion. Als dritter Schritt folgt dann die Bereinigung der Daten.

### **7.6.1 Selektion**

Aus der Gesamtheit der zur Verfügung stehenden Daten müssen für die Analyse die relevanten Datensätze selektiert werden. Für die Anwendung des Systems haben sich die folgenden Selektionskriterien als sinnvoll herausgestellt. Die folgenden Kriterien sind alle jeweils einzeln optional:

- **Testprogramm:** Sowohl die Messwerte als auch die Ausbeuten werden durch das benutzte Prüfprogramm im Funktionstest beeinflusst. Zudem wird durch die Wahl des Testprogramms zugleich das Produkt festgelegt.
- **Datum:** Da sich über längere Zeiträume wie bereits beschrieben im Zuge der Weiterentwicklung durchaus Dinge verändern, ist es eher sinnvoll, die Analyse auf einen kleineren Zeitbereich zu beschränken.
- **Chargennummer:** Halbleiterfabriken fertigen die Produkte in Chargen. Um Identifikation und Rückverfolgbarkeit zu gewährleisten, werden diese Chargen nummeriert. Entsprechend ist die Chargennummer das primäre Materialidentifikationsmerkmal in einer Halbleiterfabrik.
- **Produktionsstandort:** Wird ein Produkt in mehreren Produktionsstandorten gefertigt, ist es typischerweise sinnvoll, die Auswertung auf einen Fertigungsstandort zu beschränken.
- **Wafergröße:** Der Scheibendurchmesser kann sich zwischen verschiedenen Fertigungsstandorten aber auch zwischen verschiedenen Linien eines Standorts unterscheiden. Da der Gerätepark sich normalerweise wesentlich unterscheidet, ist es sinnvoll, die Auswertung nach Wafergrößen zu trennen.

- **Chargentyp:** Neben den Produktionschargen gibt es in den Fertigungsbetrieben noch weitere spezielle Chargen. Zur Weiterentwicklung des Fertigungsprozesses sind Versuche notwendig. Diese werden häufig mit Serienprodukten durchgeführt, um die Ergebnisse vergleichen zu können. Ohne besondere Freigabe darf derartige Material nicht zum Kunden geliefert werden und wird daher gesondert gekennzeichnet. Ähnliches gilt für Modifikationen am Produkt zur Produktweiterentwicklung. Typischerweise beschränkt man sich bei der Auswertung auf Produktionsmaterial.

Sollten für bei der künftigen Nutzung des Systems weitere Selektionskriterien benötigt werden, ist die Erweiterung der Software einfach möglich.

### 7.6.2 Normalisierung / Ausreißer-Selektion

Vor der Verarbeitung der Daten ist es notwendig, diese in einen einheitlichen Wertebereich zu überführen. Damit die nachfolgenden, informationsverarbeitenden, Schritte funktionieren können, muss der Wertebereich der auszuwertenden Daten auf das Intervall von  $[0..1]$  abgebildet werden.

#### **A Motivation**

Wird hier eine einfache Galilei Transformation durchgeführt, dann wird nur den Ursprung des Koordinatensystems verschoben und eine konstante Skalierung eingeführt. Aufgrund des bereits beschriebenen Vorkommens von Fehlmessungen und fehlenden Messungen führt diese einfach umkehrbare Funktion häufig dazu, dass die typischerweise weit-ab liegenden Fehlmessungen den Wertebereich der realen Messungen sehr stark einschränken und damit dazu führen, dass tatsächlich vorliegenden und auch gesuchte Effekte im Rauschen untergehen. Ein analoges Problem liegt mit fehlenden Messwerten vor, die typischerweise in den Messergebnissen mit einem definierten, in der Realität nicht vorkommenden, Messwert dargestellt werden. Diese Fakten führen unmittelbar zu der Notwendigkeit, irreguläre Messwerte aus den Daten zu entfernen bevor diese normalisiert werden.

Offensichtlich besteht hier das Risiko, den eigentlich gesuchten Effekt direkt mit zu entfernen. Daher muss der Prozess der Ausreißererken-  
nung und der Normalisierung sorgfältig und reproduzierbar erfolgen.  
Nur dann besteht die Möglichkeit im nachfolgenden Iterationsprozess  
die Ausreißereselektion noch zu modifizieren.

## B Verfahren

Als sehr gutes Verfahren hat sich die Bewertung anhand der Berechnung  
von Quartilen<sup>63</sup> herausgestellt. Quartile sind gegenüber der Berechnung  
von Mittelwert und Standardabweichung wesentlich robuster gegen sehr  
weit abseits liegende Messwerte. Klarerweise unterstellt auch dieses  
Verfahren eine einheitliche Fertigung und im Kern eine Verteilung die  
der Normalverteilung mindestens stark ähnelt. Durch die hier benutzten  
Grenzen von  $\pm 6\sigma$  fällt diese nicht in allen Fällen zutreffende Voraus-  
setzung typischerweise nicht relevant ins Gewicht. Daher wurde dieses  
Verfahren als Standardverfahren hinterlegt.

Mit  $Z$  als Symbol für das einzelne Datum und  $L_{\min}$  als untere sowie  $L_{\max}$   
als obere Grenze der glaubwürdigen Messwerte ergibt sich folgende Re-  
chenvorschrift:

Median:

$$\bar{Z} = Q_{.5}$$

Die Schwankungsbreite wird hier wie folgt festgelegt:

$$\sigma = \frac{Q_{.75} - Q_{.25}}{1,35}$$

Die Grenzen für die Verteilung des Datensatzes ergeben sich damit zu:

$$L_{\min} = \bar{Z} - 6\sigma$$

$$L_{\max} = \bar{Z} + 6\sigma$$

Die Transformationsfunktion für die Daten jedes Datensatzes ergibt sich  
damit wie in Abbildung 54 dargestellt.

---

<sup>63</sup> Das Quartil ist eine Sonderform des Quantils. Jedes Quartil beschreibt die Grenze  
zwischen jeweils 25% der nach Größe selektierten Daten. Damit ist das zweite Quartil  
der Median.

$$\dot{Z} = \frac{Z - L_{min}}{L_{max} - L_{min}} = \frac{Z - L_{min}}{12\sigma}$$

Abbildung 54 Transformationsfunktion zur Normierung der Daten

Für einen Analysevorgang werden entsprechend Abbildung 48 viele verschiedene Datensätze benötigt. Die allgemeine Transformationsfunktion besteht daher aus Grenzen, Medianen und Schwankungsbreiten für jede einzelne Spalte in der genannten Funktion. Die Abbildungsfunktion für jede Zeile aus Abbildung 49 lautet daher wie in Abbildung 55 gezeigt.

$$\vec{Z} = \frac{\vec{Z} - \vec{L}_{min}}{\vec{L}_{max} - \vec{L}_{min}} = \frac{\vec{Z} - \vec{L}_{min}}{12\vec{\sigma}}$$

Abbildung 55 Allgemeine Normalisierungsfunktion

Unter Nutzung dieser Funktion werden die betrachteten Daten innerhalb des Datenbanksystems normalisiert. Ergebnisse der Transformationsfunktion außerhalb des zulässigen Intervalls werden verworfen. Damit werden die betroffenen Daten als Ausreißer behandelt und nicht weiter betrachtet.

Bei vorliegender Normalverteilung liegen im obengenannten Fall 0,002ppm oder 2 Messwerte von einer Milliarde außerhalb des Fensters. Die in der Literatur häufig [z.B. BARRY08, SKF08 und andere] genannte Anzahl von 3,4ppm resultiert aus der Annahme der Verschiebung des Mittelwertes bezüglich der Messgrenzen in der Anwendung. Da hier die Grenzen keine bereits definierten Spezifikationsgrenzen sind sondern im Rahmen der Normalisierung symmetrisch berechnet werden, trifft diese „Verschlechterung“ hier nicht zu. Damit liegen wir hier -bei hinreichender Ähnlichkeit der Verteilung der Messwerte zur Gauss'schen Normalverteilung, auf der sicheren Seite der Ausreißerererkennung.

Technisch wird hier die in der Oracle Datenbank vorhandene Funktion PERCENTILE\_CONT [ORA05] genutzt, so dass keine individuelle Kodierung notwendig war.

Die als Ausreißer bewerteten Messwerte werden in der Tabelle der normalisierten Daten mit NULL gekennzeichnet. Wichtig ist hier zu erwähnen, das NULL bei einer relationalen Datenbank einen nicht vorhandenen Datensatz anzeigt und dieser Zustand vom Inhaltswert Null unterschied-

lich ist. Diese Unterscheidung ist essentiell und wird im weiteren Verlauf der Auswertung benötigt.

Sollte es im Rahmen der Nutzung des Systems erforderlich sein, für einzelne Werte andere Verfahren der Ausreißererkennung zu verwenden, so ist dieses ohne Weiteres möglich. Sowohl die Quelldaten als auch die normalisierten Daten liegen in entsprechenden Datenbanktabellen vor so dass es vielfältige Möglichkeiten gibt, alternative Verfahren einzusetzen. Dieses war bislang allerdings nicht erforderlich.

### **C Persistenz der Ausreißererkennung**

Wie im Weiteren noch genauer aufgezeigt werden wird, ist es ein Betriebsmodus des Systems, das erfolgreich trainierte Neuronale Netz für nachfolgende Fertigungslose zu verwenden. Dieses setzt natürlich voraus, dass die Abbildungsfunktion (Abbildung 55) für die neuen Daten unverändert genutzt wird.

Da in der industriellen Fertigung die Parameterlage von Los zu Los leicht schwankt, ändert sich entsprechend der Anteil der aussortierten Ausreißer. Schwankt darüber hinaus, gegebenenfalls wegen eines Problems an einer Fertigungsmaschine, die Parameterverteilung, erhöht sich der Anteil der aussortierten Ausreißer unter Umständen erheblich.

Der Anteil zulässiger Ausreißer wird aktuell manuell vorgegeben. Überschreitet ein Datensatz diese gesetzten Grenzen, wird der Benutzer durch entsprechende Fehlermeldung auf diesen Umstand hingewiesen und hat dann die Möglichkeit, auf Basis der nun vorliegenden Daten eine Normalisierung durchzuführen. In diesem Falle sind die Trainingsdaten des neuronalen Netzes natürlich nicht mehr brauchbar und das Netz muss erneut angelernt werden.

### **7.6.3 Topologische und chronologische Zuordnung**

Wie schon in den vergangenen Kapiteln dargestellt, liegen die Daten in sehr unterschiedlicher Dichte vor. Zur Nutzung der Auswertung müssen die Daten in der in Abbildung 49 beschriebenen Form vorliegen. Je nach gewählter Granularität der Analyse (siehe 7.4.2 ) müssen die Daten verdichtet oder geeignet zugeordnet werden.

Am einfachsten wird das notwendige Vorgehen an zwei Beispielen deutlich. In beiden Fällen werden Transaktionen für die Verwendung der Eingangsdaten beschrieben.

### **A Beispiel I: Auswertegranularität Wafer**

**PCM-Daten:** Wie beschrieben werden in der Serienfertigung auf jeder produzierten Scheibe PCM Daten an mehreren Stellen ermittelt. Für die Auswertung auf Scheibenebene werden diese nach der Ausreißer-selektion gemittelt.

**Maschinendaten:** Die logistischen Maschinendaten aus der Halbleiterfertigung liegen häufig pro Charge vor. Für diese Fertigungsmaschinen werden die entsprechenden Kennungen von Maschine, Fertigungsrezept sowie weitere Daten wie Bediener, Start- und Endzeitpunkt, etc. für die einzelnen Scheiben übernommen. Soweit bei Einzelscheibenprozessen Informationen auf Scheibenbasis vorliegen, werden diese –gegebenenfalls gemittelt– der jeweiligen Scheibe zugeordnet.

**Messwerte Maschinen(Material):** Die Messungen am Produktionsmaterial werden in der Fertigungslinie nur an wenigen Scheiben und dort auch an wenigen Punkten durchgeführt. Es ist davon auszugehen, dass der Stichprobenumfang der durch die Spezialisten vor Ort festgelegt wird, signifikant für das gesamte Fertigungslos ist. Daher werden hier die Messdaten sinnvollerweise für das ganze Los gemittelt und diese dann jeder einzelnen Scheibe zugeordnet.

**Messwerte Maschinen(Equipment):** Die benötigten internen Daten der Fertigungsmaschinen sind typischerweise nicht unmittelbar mit dem gefertigten Material verbunden. Die Zuordnung muss hier über den zeitlichen Zusammenhang hergestellt werden, wobei wiederum typischerweise die Werte einer Charge zugeordnet werden und diese dann für alle Wafer in gleicher Weise gelten.

**Messwerte Funktionstest/Ausbeutedaten:** Die Ausbeutedaten liegen wie die Messwerte vom Funktionstest pro Die vor. Für die Auswertung auf Waferlevel werden diese pro Wafer gemittelt und die Ergebnisse diesem Datensatz zugeordnet.

## **B Beispiel II: Auswertegranularität Die**

**PCM-Daten:** Hier muss eine schwierige Entscheidung getroffen werden. Prinzipiell ist davon auszugehen, dass die gemessenen Werte signifikant für die ganze Scheibe sind. Entsprechend kann man die gültigen Werte mitteln und die Ergebnisse jedem Die zuweisen. Andererseits ist es aus der Technologie klar, dass es einige Parameter gibt, die über die Scheibe eben nicht konstant sind (z.B. Schichtdicken wenn Horizontalöfen verwendet werden). Zudem kann gerade die Veränderung von Parametern über eine Scheibe hinweg sehr wichtige Hinweise auf mögliche Fehlerursachen liefern. Die Nutzung der topologischen Informationen der PCM ist allerdings automatisiert nicht sinnvoll machbar. Eine Ausgleichsrechnung zwischen den Messpositionen und Zuweisung der ermittelten Werte an die jeweiligen Dies setzt Linearität der Änderung voraus. Zudem kann in diesem Fall mit fehlenden Messwerten nicht mehr umgegangen werden. Im Normalfall wird an dieser Stelle als Auswertegranularität eben nicht das einzelne Die sondern das Bildfeld genutzt. Das führt dann typischerweise dazu, dass für die meisten Bildfelder keine PCM Messwerte vorliegen und diese Datensätze während der Datenvorbereitung entfernt werden. Übrig bleiben dann die typischerweise vergleichsweise wenigen Messwerte von Bauteilen die in Bildfeldern liegen, für die die PCM Daten vorhanden sind. Alternativ werden die PCM Daten über die Scheibe gemittelt und allen Funktionstestdaten der jeweiligen Scheibe identisch zugeschrieben. Hier ist also die Entscheidung zwischen der gewünschten großen Menge an Datensätzen und der Verlässlichkeit der einzelnen Messwerte zu treffen. Diese Entscheidung muss jeweils bei der einzelnen Analyse getroffen werden. In Spezialfällen müssen zur Absicherung der Analyse zeitweise flächendeckende PCM durchgeführt werden. In diesem Fall kann aufgrund der deutlich dichteren Messpunkte in guter Näherung von einer linearen Änderung ausgegangen werden. Damit wird dann der mittels Ausgleichsrechnung errechnete Wert dem Die zugeordnet.

**Maschinendaten:** Wie schon beim ersten Beispiel beschrieben, müssen die pro Los oder pro Scheibe vorliegenden Daten gegebenenfalls gemittelt für alle Dies verwendet werden.

**Messwerte Maschinen(Material):** Ebenfalls wie bei der Auswertung auf Scheibenbasis werden die gemittelten Messdaten jedem Auswertektor, hier einem einzelnen Die, zugeordnet.

**Messwerte Maschinen(Equipment):** Auch diese Werte liegen nur pro Charge vor und werden allen Dies identisch zugewiesen.

**Messwerte Funktionstest/Ausbeutedaten:** Die Ausbeutedaten liegen wie die Messwerte vom Funktionstest pro Die vor. Diese werden nach der Normalisierung entsprechend unverändert übernommen.

### 7.6.4 Bereinigung

Die Daten liegen nun nach den bisher beschriebenen Schritten in einem der Zielfunktion entsprechenden Format vor. Bevor nun die wissensgenerierenden Schritte gestartet werden können, müssen diese Daten noch bereinigt werden.

#### **A Widerspruchsfreiheit**

Da nur mit einer Abstraktion der Wirklichkeit analysiert werden kann, kommt es vor, dass die vorhandenen Daten widersprüchlich sind. Die Wahrscheinlichkeit des Auftretens ist am höchsten, wenn bei den Eingangsparametern nur logistische Informationen benutzt werden. Wenn dann bei identischen Eingangsparametern ein unterschiedliches Klassifizierungsergebnis vorliegt, ist damit kein funktionaler Zusammenhang mehr möglich.

In diesem Fall werden die entsprechenden Zeilen in der Matrix entfernt. Wie in den anderen datenreduzierenden Fällen ist auch dies wieder Teil eines Näherungsverfahrens. Sollte kein qualifiziertes Ergebnis mehr erzielt werden, müssen weitere Daten hinzugefügt werden. Selbiges gilt, wenn ein erheblicher Teil der vorliegenden Daten aufgrund dieses Kriteriums verworfen werden muss.

Werden Messwerte analysiert, sind die Daten eher kontinuierlich verteilt. Damit sind mathematisch widersprüchliche Zusammenhänge praktisch ausgeschlossen. Inhaltlich können die Daten allerdings nach wie vor widersprüchlich sein, wenn beispielsweise sich ein einzelner Messwert nur im Rahmen der Messtoleranz von einem anderen unterscheidet und die

Klassifizierung unterschiedlich ist. In diesem Fall sind die Daten mathematisch widerspruchsfrei. Trotzdem wird keiner der nachfolgenden Prozesse ohne die Hinzunahme weiterer Daten ein sinnvolles Ergebnis hervorbringen können, da die Daten eben inhaltlich doch widersprüchlich sind.

Das erstellte System erkennt mathematisch widersprüchliche Werte automatisch und entfernt diese. Inhaltlich widersprüchliche Werte können bestenfalls symptomatisch erkannt werden. Die Beseitigung ist sehr schwierig da die Erkennung bereits kompliziert ist.

### **B Konstante Spalten**

Weiterhin kommt es –wiederum primär bei logistischen Daten– vor, dass nach der Normalisierung eine Spalte der Eingangsmatrix konstant ist. Diese enthält dann keinerlei verwertbare Information mehr und wird entfernt.

### **C Mindestbelegung**

Schließlich hat sich die Definition einer Mindestbelegung als sehr sinnvoll herausgestellt. Wie beschrieben, kommt es aus verschiedenen Ursachen vor, dass einzelne Messwerte fehlen oder als ungültig gekennzeichnet werden. Unter Umständen führt das dazu, dass einzelne Spalten der Eingangsmatrix sehr schwach besetzt sind. In diesen Fällen wird das Ergebnis der Analyse insgesamt verschlechtert. Daher gibt es innerhalb des Systems eine definierte Mindestfüllquote, sowohl für Spalten als auch für Zeilen. Zeilen oder Spalten unterhalb dieses Wertes werden wie die konstanten Spalten entfernt.

## **7.6.5 Abschluss der Vorbereitung**

Damit ist die Vorbereitung der Daten für die Auswertung abgeschlossen. Die unter Vorbereitung der Daten ausführlich beschriebenen Abläufe sind im System abgebildet und laufen automatisch ab. Damit stehen die zu untersuchenden Daten mit folgenden Eigenschaften zur Verfügung:

- Ein Datensatz pro Zeile der Matrix
- Zugehöriges Ergebnis vorhanden
- Ausreißer sind entfernt

- Die Daten sind auf das Intervall [0..1] beschränkt
- Konstante und widersprüchliche Terme sind eliminiert.

Bis zu diesem Zeitpunkt, wurden die Daten im relationalen Datenbanksystem gehalten. Für die folgenden Schritte werden diese nun extrahiert und in einem definierten Textformat [s.10.1, S. 183] gespeichert. Durch diesen Wechsel werden folgende Vorteile erreicht:

- Verschiedene Versionen und Selektionen der Daten sind leichter handhabbar.
- Die Daten sind sehr einfach zu archivieren.
- Durch die ausführliche und lesbare Kennung im Header der Datei ist direkt aus den OTC Dateien heraus erkennbar, worum es sich handelt.

Die zumindest potentiellen Nachteile sind dem gegenüber zu vernachlässigen:

- Durch die Ablage der Dateien auf einem gesicherten Dateisystem ist die nun nicht mehr automatisch über die Datenbank garantierte Datensicherung nicht kritisch.
- Durch die Zusammenführung der beschreibenden Informationen und der Daten in einer Datei wird die Identifizierbarkeit sichergestellt.
- Da auf die OTC Dateien nach ihrer Erzeugung nur von einem Prozess und nur lesend zugegriffen wird, ist die nicht vorhandene Transaktionssicherheit kein Problem.

Damit sind die notwendigen Datenvorbereitungen beendet und die wissensgenerierenden Schritte können beginnen.

### ***7.7 Feature Selection***

In dieser Arbeit ist das Thema die Analyse von sehr hochdimensionalen Daten. Mit Hilfe der Feature Selection sollen die Anzahl der Parameter um die für das aktuell untersuchte Problem irrelevanten reduziert werden. Die übrigbleibenden relevanten Parameter werden in den folgenden Schritten mit dem Neuronalen Netz weiterverarbeitet.

Der prinzipielle Ablauf von Feature Selection ist beispielsweise in [DASH97] beschreiben.

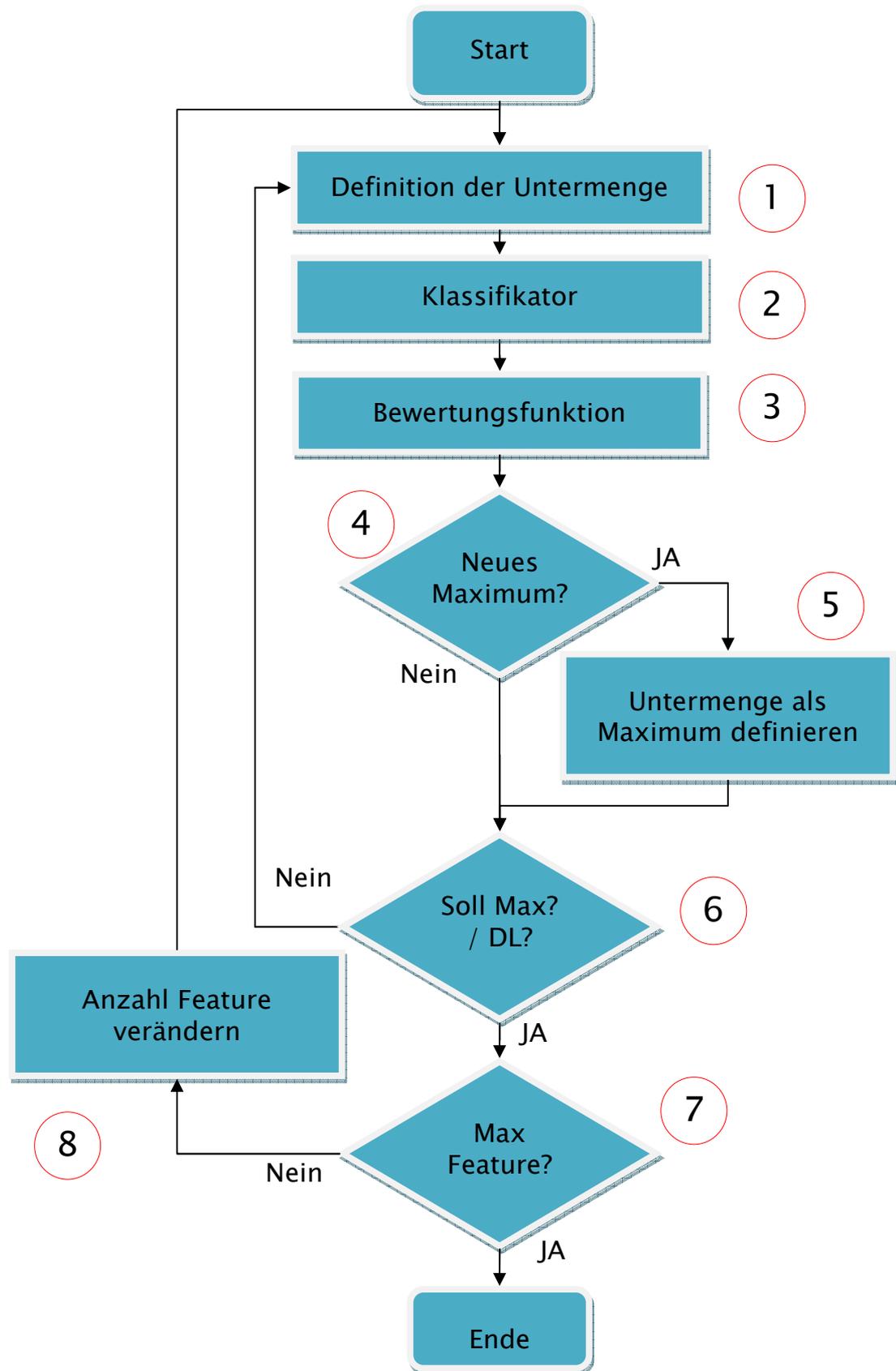


Abbildung 56 Ablauf Feature Selection

Nach der beschriebenen Vorbereitung der Daten ist der nächste Schritt die Feature Selection. Der Ablauf ist in Abbildung 56 dargestellt. Die äußere Schleife in Abbildung 56 (① – ⑦) wird für jede untersuchte Parameterkombination (und damit auch Anzahl) mehrfach durchlaufen. Generell kann man die Untersuchung mit allen Parametern beginnen und mit jedem Schritt ein Feature<sup>64</sup> entfernen. Alternativ beginnt man ohne Parameter und fügt sukzessive weitere Parameter hinzu. Dieses Verfahren wird in der Literatur als Sequential Forward Search (SFS) bezeichnet [z.B. LIMO98]. Entsprechend wird das Verfahren bei dem Features entfernt werden als Sequential Backward Search (SBS) bezeichnet.

### 7.7.1 Sequential Forward Floating Search

Die Anzahl der bei den hier untersuchten Datensätzen zur Verfügung stehenden Parameter ist mit bis zu mehreren Hundert sehr groß. Die für den analysierten Effekt relevanten Parameter müssen für eine erfolgreiche Analyse und mögliche Gegenmaßnahmen im Vergleich sehr wenige sein. Daher wird hier das vorwärtsgerichtete Verfahren, also der Beginn mit einer leeren Menge an Parametern, betrachtet.

Das einfachere Sequential Forward Search (SFS) hat den Nachteil des sogenannten „nesting effect“. Ein einmal zu einer Untermenge hinzugefügter Parameter kann im weiteren Verfahren nicht wieder entfernt werden. Daher wurde hier eine Erweiterung des Standardverfahrens eingesetzt. Die Definition von neuen Subsets wird nach dem Sequential Forward Floating Search (SFFS) Verfahren nach [PUDI94] durchgeführt. Dieses Verfahren evaluiert nach der Hinzunahme eines Parameters nach dem SFS Algorithmus einen SBS Schritt (Abbildung 56,⑧). Wird die Bewertung durch die Entfernung eines Parameters nicht wesentlich reduziert, wird dieser wieder entfernt. Damit werden auch die bereits selektierten Parameter mehrfach überprüft. Ein Festhalten an einmal selektierten Parametern, die sich im weiteren Verlauf als weniger relevant erweisen, wird damit verhindert.

---

<sup>64</sup> Die Begriffe Parameter und Feature werden hier synonym gebraucht.

Um den richtigen Parameter hinzuzufügen werden verschiedene Untermengen mit der gewählten Anzahl von Parametern gebildet.

### 7.7.2 K Nearest Neighbor als Klassifikator

Die in Abbildung 56 unter ① gebildete Untermenge an Parametern wird nun klassifiziert. Für Feature Selection-Verfahren sind mehrere verschiedene Klassifizierungsalgorithmen möglich, der Ablauf ist allerdings jeweils gleich:

Die zur Verfügung stehenden Datensätze werden in Trainingsdatensatz und Testdatensatz aufgetrennt. Auf Basis des Trainingsdatensatzes wird eine Struktur aufgebaut. Innerhalb dieser Struktur werden dann die Testdatensätze klassifiziert.

Der KNN (k Nearest Neighbor) Algorithmus wählt die Klassenzugehörigkeit des Testvektors anhand der Majorität der Klassenzugehörigkeit der k nächsten Nachbarn. Bei den Nachbarn handelt es sich um den von den Trainingsdatensätzen aufgespannten euklidischen Raum. Die Dimension des Raumes entspricht der aktuell betrachteten Featureanzahl.

Zur Absicherung des Ergebnisses wird hier zudem eine Cross Validierung eingesetzt. In [KOHA95] werden verschiedene Verfahren zur Validierung vorgeschlagen. Wir setzen hier konkret das mit „Leave-one-out Cross Validation“ bezeichnete Verfahren ein. Es wird jeweils ein einzelner Datensatz als Testvektor genutzt. Der Rest der Daten dient jeweils als Trainingsvektoren. Dies wird für alle Datensätze des Eingangsdatensatzes durchgeführt.

#### **A Warum KNN Algorithmus?**

Der hier vorgeschlagene Lösungsweg soll neue Methoden im Auffinden von Ursachen für Besonderheiten bei existierenden Fertigungsprozessen aufzeigen. Damit muss das Verfahren ein System abbilden, das nach physikalischen Gesetzmäßigkeiten funktioniert, auch wenn wir wie mehrfach beschrieben nicht in der Lage sind, die Abhängigkeit vollständig mathematisch zu beschreiben.

Unter der Annahme, dass die Ergebnisfunktion hinreichend glatt ist, ist es wahrscheinlich, dass ein Datenpunkt dieselbe Klasse hat wie seine

Nachbarn. Diese Annahme wurde aus mehreren Gründen getroffen und hat sich auch für die Analyse bewährt:

- Ist die Ergebnisfunktion an vielen Stellen unstetig, ist die Wahrscheinlichkeit einer erfolgreichen Analyse –egal ob mit konventionellen oder den hier vorgestellten Methoden– sehr gering.
- Die große Mehrzahl der hier betrachteten Parameter sind Messwerte die sich aus Fertigungsprozessen ergeben. Bei ausreichender Vergrößerung sind diese fast immer hinreichend glatt.

Entsprechend wurde für das hier vorliegende System eine KNN (K Nearest Neighbors) Evaluierungsfunktion ausgewählt. Bei einstellbarer Anzahl der nächsten Nachbarn weist diese den Datenpunkten dieselbe Klasse wie der Mehrzahl der umliegenden (K) Datenpunkte zu. Da der zu Grunde liegende Vektorraum eine Metrik hat, wird hier euklidische Distanz als Abstandsmaß verwendet.

### **B Erweiterung KNN-Algorithmus**

Die kritische Würdigung des KNN-Verfahrens zeigt, dass das beschriebene Verfahren in der Basiskonfiguration für die vorliegende Fragestellung nicht hinreichend robust ist.

Abbildung 57 zeigt einen einfacher Datensatz mit zwei Dimensionen und zwei Klassen. Die Kreise stellen jeweils Datenpunkte dar deren Farbe die Klassenzugehörigkeit signalisiert. Der weiße Punkt ist der jeweils untersuchte. Leicht zu erkennen ist, dass die mittels KNN ermittelte Klassenzugehörigkeit vom gewählten K abhängt. Für  $K=3$  ergibt sich eine Zugehörigkeit zur blauen Klasse. Für  $K=5$  ist das Ergebnis Grün während für  $K=7$  wiederum Blau das Ergebnis ist. Diese Abhängigkeit vom gewählten Parameter K lässt keine optimalen Ergebnisse erwarten. Diese Erwartung wurde in Versuchen auch bestätigt.

Weiterhin von Nachteil bei diesem Verfahren ist die Pattsituation die immer dann eintritt, wenn keine Klasse in der Treffermenge die Mehrheit hat. Für den o.g. trivialen Fall mit zwei Klassen kann dieses durch Beschränkung von K auf ungerade Zahlen vermieden werden. Sobald die Klassenanzahl aber größer als zwei wird, ist dieses nicht mehr möglich.

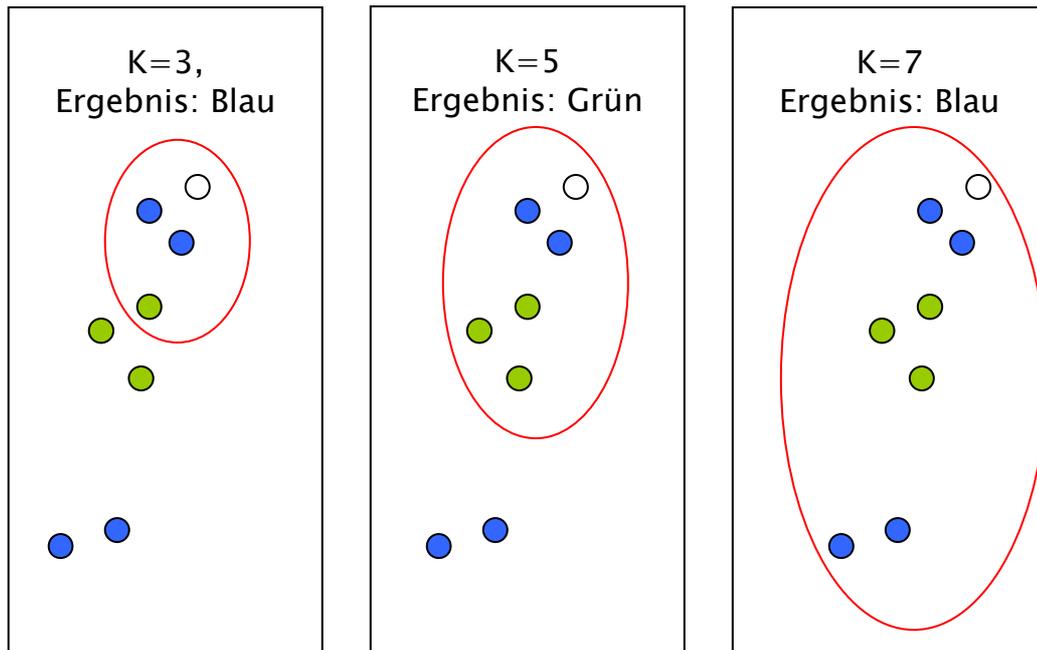


Abbildung 57 Beispiel KNN Standard

Analog dem in [HECH04] beschriebenen und unter dem Namen  $wkNN$  (Weighted  $k$ -Nearest-Neighbors) veröffentlichten Ansatz haben wir das Verfahren um einen Wichtungsfaktor erweitert, der sich aus der Entfernung der Punkte ergibt. Hier sind verschiedene Bewertungen des Abstands ( linear, quadratisch etc. ) denkbar. Die einfache Wichtung umgekehrt proportional zur Entfernung hat in den Versuchen sehr gute Erfolge gezeigt und wird daher im Weiteren verwendet.

Dies ist auch durchaus nachvollziehbar, da die Werte für  $K$  in den hier untersuchten Fällen klein sind und daher auch lokale Phänomene im Raum untersucht werden. Hier ist ein linearer Ansatz häufig zielführend. Abbildung 58 zeigt schematisch die vorgenommene Wichtung über die Fläche der Punkte. Der, hier wiederum weiß dargestellte, aktuell untersuchte Punkt wird in allen dargestellten Fällen der blauen Klasse zugeordnet. Dieses entspricht auch dem Ergebnis einer manuell durchzuführenden Klassifikation.

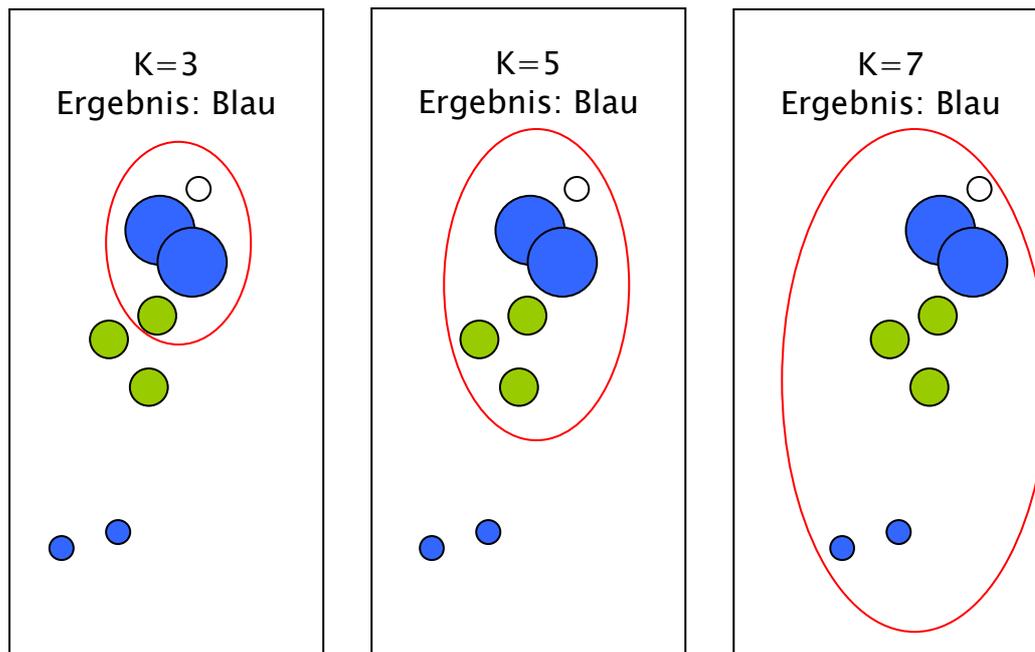


Abbildung 58 KNN mit linearem Distanzmaß

### 7.7.3 Bewertungsfunktion

Wie beschrieben, erfolgt die Klassifikation für jeden Datensatz und jede Untermenge durch den Klassifikator. Wie in Abbildung 56 ersichtlich, folgt nun die Bewertung für die jeweils aktuelle Untermenge (Abbildung 56, ③).

Im hier behandelten Umfeld liegen beim Start der Analyse sowohl die Eingangsdaten als auch die Klassifikationsergebnisse vor. Wie befinden uns also in einer Problemstellung des überwachten Lernens (Supervised Learning). Daher ist die –im Falle des nichtüberwachten Lernens (Un-supervised Learning)– vorhandene Schwierigkeit einer geeigneten Bewertungsfunktion vergleichsweise einfach zu lösen. Die Bewertungsfunktion ist –wie auch bei der Bewertung des Neuronalen Netzes im Folgenden– die Quote der korrekt vorgesagten Klassifizierung für die einzelnen Datensätze. Übertrifft diese Quote für die gerade aktuelle Untermenge das bisherige Maximum, wird die neue Untermenge als aktuell beste definiert (Abbildung 56, ④ / ⑤).

### 7.7.4 Abbruchbedingung

Aufgrund der bereits erwähnten sehr großen Anzahl von Parametern im Eingangsdatensatz ist es notwendig eine Abbruchbedingung zu formulieren, da alle Kombinationen nicht in sinnvollen Zeiträumen überprüfbar sind. Hier wird nach dem Erreichen einer vorher eingestellten Trefferquote die Schleife beendet. Weiterhin wird eine maximale Anzahl von Durchläufen vorgegeben, um eine endlose Schleifenausführung zu vermeiden (Abbildung 56,⑥). Sobald die voreingestellte maximale Anzahl von Parametern erreicht (Abbildung 56,⑦) wird, endet der Prozess der Feature Selection.

### 7.7.5 Klasseneinteilung der Eingangsdaten

Neben der Anzahl der Feature ist die Klasseneinteilung des Ergebnisvektors ein notwendiger Input für die Feature Selection. Während die optimale Anzahl der Feature iterativ automatisiert bestimmt wird, ist die Entscheidung über die Klasseneinteilung notwendige Vorarbeit durch den Benutzer.

Das hier vorgestellte System erlaubt die automatische oder manuelle Klasseneinteilung. In der vorgesehenen Anwendung wird die Feature Selection in Kombination mit dem Neuronalen Netz betrieben. Dafür ist eine identische Klasseneinteilung in beiden Systemen und damit eine manuelle Vorgabe der Klassengrenzen notwendig. Da mittels dieses Systems die Ursache für unerwünschte Effekte in Prozessen gesucht wird, ist das auch gleichzeitig der normalerweise sinnvolle Fall, da die Klassengrenzen manuell zwischen die erwünschten und die unerwünschten Ergebnisse gelegt werden können.

Die automatische Klassifizierung kann sinnvoll sein, wenn der Bediener sich beim Beginn der Analyse ein Bild machen will und wenn mittels der manuellen Vorgabe der Grenzen keine sinnvollen Ergebnisse erzielt werden.

Bei der automatischen Klasseneinteilung wird die Anzahl der Klassen durch den Benutzer vorgegeben. Weiterhin hat der Benutzer die Auswahl zwischen folgenden Verteilungen:

1. Klassengröße identisch: Das System teilt den Wertebereich der Ergebnisse  $[0..1]$  in  $N$  gleich große Bereiche auf.
2. Klassenbelegung identisch: Das System wählt die Klassengrenzen so, dass in allen Klassen gleich viele Datensätze enthalten sind.

Die zweite Option ist aus softwaretechnischer Sicht die bessere, da der Algorithmus bessere Ergebnisse liefert, wenn alle Klassen etwa gleich besetzt sind. Allerdings ist man in diesem Fall typischerweise weit vom gewünschten Analyseergebnis entfernt da die Gleichverteilung der Messwerte auf die Klassen „erwünscht“ und „unerwünscht“ die Ausnahme darstellt. Damit wird für eine erfolgreiche Analyse die Option eins die häufigere sein. Diese Annahme hat sich auch bei den durchgeführten Versuchen bestätigt.

### 7.7.6 Verifikation

Der Einsatz der Feature Selection in diesem Umfeld erfolgt um die sehr große Menge an Eingangsdaten auf einen handhabbaren Umfang zu reduzieren und dabei die relevante Information möglichst zu behalten. Die These, dass dies sinnvoll möglich ist muss nun verifiziert oder falsifiziert werden. Zudem ist es notwendig, die Güte des Systemkonzepts und die Güte der Implementierung zu überprüfen.

Die Güte des Systemkonzeptes und der Implementierung wurde an konstruierten Testdaten überprüft. Die Verifikation der o.g. These setzt den Nachweis voraus, dass mit durch die Feature Selection ausgewählten Parametern gegenüber dem Monte Carlo Ansatz signifikant höhere Validierungsquoten erreicht werden. Dieser Nachweis kann final erst im Zusammenspiel mit dem Neuronalen Netz erfolgen, so dass auf für die finale Verifikation der These auf die folgenden Kapitel Neuronales Netz (S. 150) und vor allem Ergebnisse (S. 160) hingewiesen wird. Jedoch kann auch der alleinige Einsatz der Feature Selection in einigen Fällen bereits sehr hilfreiche Ergebnisse liefern. Dazu wird im Folgenden ein Fallbeispiel beschrieben:

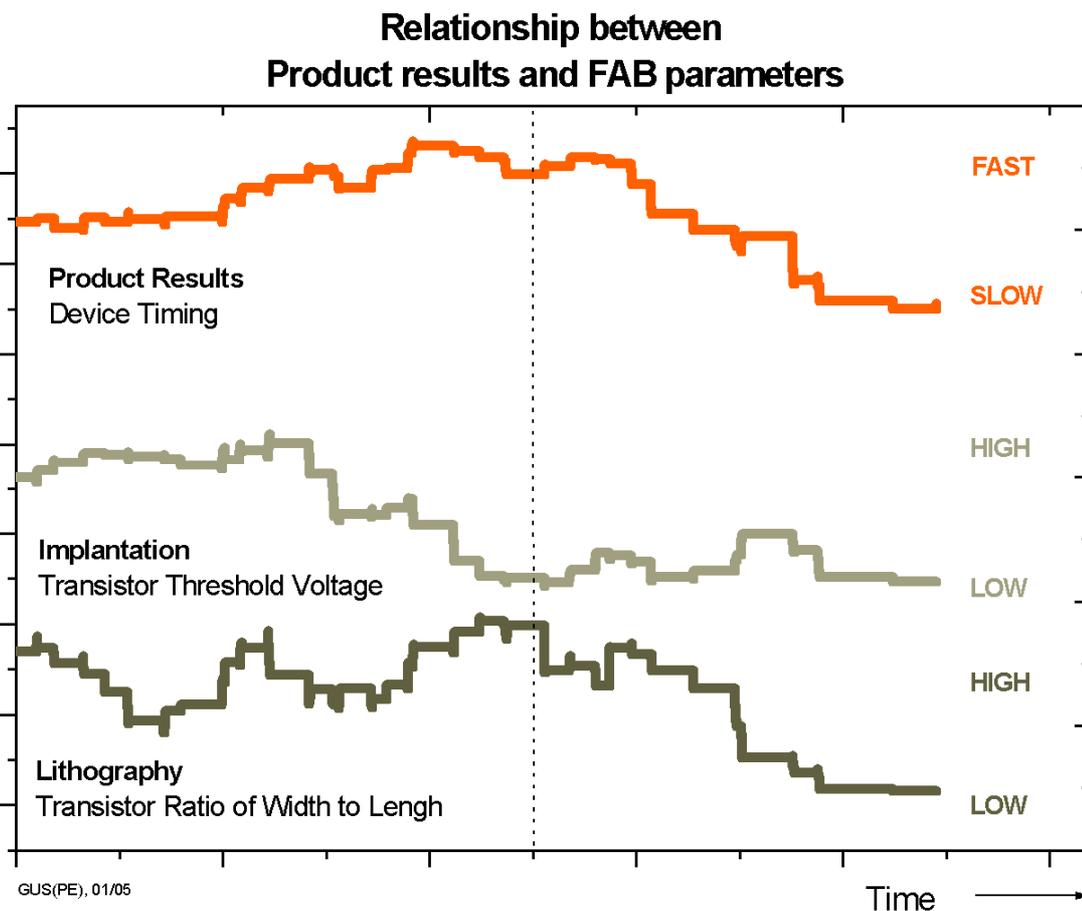


Abbildung 59 Fallbeispiel Feature Selection [MOBE05]

Abbildung 59 zeigt die Ergebnisse der Anwendung von Feature Selection mit dem SFFS Algorithmus und der KNN Klassifizierung. Das auslösende Problem war das Messergebnis eines Parameters beim Funktionstest der Bauteile auf Wafer Ebene. Ein Timing Messwert des Produktes wurde von der statistischen Prozesskontrolle als auffällig gemeldet. Die Analyse der Experten lieferte zunächst keine plausible Erklärung für dieses Phänomen. Daraufhin wurden die Ergebnisse des PCM und des Funktionstests auf Wafer Ebene für 280 Chargen zu je 25 Wafern extrahiert. Die Feature Selection lieferte fünf signifikante Parameter aus einer Gesamtparameteranzahl von 85. Die weitere Untersuchung ergab, dass zwei Parameterpaare direkt technologisch voneinander abhingen, also final denselben Effekt beschrieben. Damit blieben noch drei relevante Parameter übrig, die in der Abbildung dargestellt wurden. Die Rot dargestellte Kurve zeigt den per SPC auffällig gewordenen Messwert. Der als „SLOW“ bezeichnete Wert ist ein unerwünschter Zustand. Betrachtet man die beiden anderen dargestellten Kurven, ist jeweils einzeln kein direkter Zu-

sammenhang mit der abfallenden Geschwindigkeit des Produktes erkennbar. In der kombinierten Darstellung ist ersichtlich, dass das Timing dann zu langsam tendiert, wenn sich sowohl Implantation als auch das Weiten zu Längenverhältnis der Transistoren in Richtung der unteren Grenze bewegt. Obwohl hier die beschriebene Erweiterung des KNN Klassifikators (siehe 7.7.2 B) für diese Auswertung noch nicht genutzt wurde, konnten hier schon überzeugende Ergebnisse erzielt werden. Die gefundene Abhängigkeit von zwei verschiedenen Parametern wäre sonst unter Umständen lange nicht entdeckt worden. Zumal die Messwerte aus verschiedenen Bereichen der Fabrik stammen und der direkte Zusammenhang nicht naheliegend ist. Als Konsequenz dieses Ergebnisses wurden Fertigungsparameter in der Fabrik angepasst. Diese Ergebnisse wurden vom Autor auf der AEC/APC Konferenz im Jahr 2005 [MOBE05] publiziert. Eine weitere Veröffentlichung erfolgte auf der ESANN [ESAN05].

### ***7.8 Neuronales Netz***

Die Validierung der mit der Feature Selection vorgenommenen Parameterauswahl wird durch ein Neuronales Netz vorgenommen. Zudem soll durch die Extraktion der Regeln Wissen über die Zusammenhänge erzeugt werden. Dieses Wissen soll den Fachmann vor Ort darin unterstützen, die Prozesse zu verbessern und damit qualitativ besseres Material und/oder eine höherer Ausbeute zu erzeugen.

Wie bereits eingehend beschrieben, lässt sich die als Beispiel herangezogene Halbleiterfertigung nicht funktional abbilden. Dies gilt auch, wenn Know-how über die technologischen Zusammenhänge der Halbleitertechnologie genutzt wird.

Daher ist hier der Ansatz, die sehr große Menge an zur Verfügung stehenden Daten mittels automatischer Verfahren zu analysieren. Neu ist es, keinerlei Wissen über die –vermuteten oder tatsächlichen– Zusammenhänge der vorliegenden Daten für die Auswertung derselben zu nutzen. Das Expertenwissen über die Zusammenhänge ist –wie beschrieben– bei der Selektion und Vorbereitung der auszuwertenden Daten unverzichtbar. Für den eigentlichen wissensgenerierenden Prozess

wird dieses vorsätzlich nicht genutzt. Hintergrund dieser Entscheidung des Autors ist der Ansatz, eine weiterführende Alternative zu den etablierten Methoden der technischen Datenanalyse bereitzustellen. Eine Hinterlegung des vorhandenen Expertenwissens geht mit dem Risiko einher, die eingefahrenen Wege nicht zu verlassen und damit gegebenenfalls vorhandene Abhängigkeiten zu unterdrücken.

Die Akquisition von neuem, regelbasiertem Wissen spielt in diesem Kontext keine Rolle, da das System in der Wissensbasis nur Faktenwissen enthält. Entsprechend werden Methoden der Wissensakquise hier nicht betrachtet.

### **7.8.1 Fuzzy ARTMAP<sup>65</sup>**

Künstliche Neuronale Netze sind ein Teilbereich des Maschinellen Lernens der bei den Wissensbasierten Systemen eingeordnet ist. Sie sind ein Versuch, die Datenverarbeitung und Wissensgenerierung durch Computer dem menschlichen Gehirn nachzubilden. Die kleinste Einheit eines solchen Netzes ist –analog zum Gehirn– ein Neuron. Dieses verfügt über Verbindungen zu anderen Neuronen und kann Signale an diese weitergeben. Diese Signalweitergabe kann abhängig von einem komplexen Eingangsmuster verstärkt oder abgeschwächt erfolgen. Zudem können –wiederum analog zur Biologie– neue Verbindungen erzeugt werden.

Künstliche Neuronale Netze werden –wie bereits gezeigt– in vielen Anwendungen eingesetzt. Dabei kommen –je nach Fragestellung– verschiedenste Kodierungen mit unterschiedlichen Stärken und Schwächen zum Einsatz. Wie bereits in 6.5 und 6.5.1 beschrieben, ist der hier vorgesehene Einsatz von Neuronalen Netzen vom üblichen Verfahren abweichend. Bisherige Einsatzgebiete von Neuronalen Netzen sind die Erstellung von Prognosen über das zukünftige Verhalten von komplexen Systemen. Weiterhin werden derartige Systeme für die Klassifikation unstrukturierter Daten (Bilder, Texte, Audio) genutzt.

---

<sup>65</sup> Erweiterung des ART Verfahrens um Supervised Learning zu unterstützen. Unterstützt wie ART nur binäre Eingangsdaten.

In dieser Arbeit ist das Ziel, das Verhalten eines Systems zu verstehen, also die Ursachen für eine Verhaltensabweichung zu erkennen. Dazu muss das eingesetzte Netzwerk in der Lage sein, das Systemverhalten zu lernen. Weiterhin ist es aber unabdingbar, die Regeln des Netzes zu extrahieren. Diese Regeln sind dann Basis einer weiteren Analyse durch den Bearbeiter mit dem Ziel, das Systemverhalten zu verbessern. Im Fokus steht die Suche nach Zusammenhängen, die bislang unerkannt in den zur Verfügung stehenden Datenmengen vorhanden sind.

Zunächst liegen die Klassifikationsergebnisse für die Eingangsdaten bereits vor. Netztypen, die überwachtes Lernen (Supervised Learning) nicht unterstützen, fallen also direkt aus der Auswahl heraus. Weiterhin sind die automatischen Klassifikationsverfahren Neuronaler Netze in dieser Arbeit nicht relevant. Die Ergebnisse liegen vor und die Klasseneinteilung der Ergebnisse wird vom Bediener vorgegeben, da diese sich aus dem zu untersuchenden Problem ergeben muss (siehe 7.7.5 ).

Die zur Verfügung stehenden Daten bestehen aus einer großen Anzahl von numerischen Parametern oder sonstigen Daten, die in ein numerisches Format überführbar sind. Die entsprechende Datenvorbereitung wurde bereits in Kapitel 7.6 beschrieben.

Fuzzy ARTMAP Netze bieten gegenüber vielen anderen Netztypen die Extraktion der gefundenen Regeln an und scheinen daher für die Lösung dieser Fragestellung geeignet. Für eine ausführliche Beschreibung der Fuzzy ARTMAP Netze und der Eignung für die hier untersuchte Fragestellung sei auf [SAPO04] verwiesen. Im Folgenden erfolgt eine kurze Erläuterung der Besonderheiten dieses Netztyps.

ART<sup>66</sup> Netze wurden von Grossberg[GRO76] vorgeschlagen. Bei ART handelt es sich um eine Erweiterung der bereits existierenden Verfahren des kompetitiven Lernens (Competitive Learning, CL). CL Algorithmen bilden eine Eingangsmenge über gewichtete Verbindungen auf eine Ausgangsmenge ab. Alle Ausgänge werden verglichen, der höchste Wert gewinnt. Anschließend werden die Gewichte aller Verbindungen erhöht, die zum Gewinner führen. Alle anderen Verbindungen bleiben unverän-

---

<sup>66</sup> Adaptive Resonance Theory

dert (WTA<sup>67</sup> Prinzip). ART erweitert dieses Verfahren um eine Rückkopplung von der Ausgangs- auf die Eingangsmenge. Eine dann auftretende positive Rückkopplung (Resonanz) erhöht die Stabilität des Lernverhaltens, insbesondere bei neuen, bisher nicht aufgetretenen Mustern [GRO87]. Carpenter, Grossberg und Reynolds [CGR91] haben mit ARTMAP ein Verfahren aus zwei gekoppelten ART Netzen vorgeschlagen. Dieses ist für Supervised Learning geeignet. Der ursprüngliche Vorschlag [GRO76] war nur für Unsupervised Learning geeignet.

Eine weitere Erweiterung des ARTMAP Verfahrens betrifft den Umgang mit kontinuierlichen Eingangsdaten. ARTMAP kann nur mit binären Eingangsdaten umgehen, was für die Analyse von Messdaten ungeeignet wäre. Unschärfe Mengen (Fuzzy Sets) wurden zuerst durch Zadeh [ZAD65] vorgeschlagen, um die Ungenauigkeit der Wirklichkeit mit dem bis dahin binären Mengenbegriff der Mathematik zu verbinden.

Carpenter, Grossberg und Rosen haben in [CGRO91] mit Fuzzy ART eine Kombination des ART Verfahrens mit den Möglichkeiten der Fuzzy Logik vorgeschlagen. Die internen Strukturen des ART Netzes wurden um die unscharfe Klassenzugehörigkeit der Fuzzy Logik erweitert.

Entsprechend der Erweiterung von ART Netzen zu ARTMAP Netzen wurden auch die Fuzzy ART Netze zu Fuzzy ARTMAP Netzen (FAM) erweitert, um Systeme abzubilden, bei denen sowohl die Eingangs- als auch die Ausgangsdaten vorliegen. Fuzzy ARTMAP wurde von Carpenter et al. in [CGMRR92] vorgeschlagen.

Der FAM Algorithmus kann noch vereinfacht werden, da durch die Kopplung von zwei Fuzzy ART Netzen zu einem Fuzzy ARTMAP Netz Symmetrien in beiden Netzen auftreten, die nicht erneut berechnet werden müssen. Dieses von T. Kasuba als Simplified Fuzzy ARTMAP (SFAM) vorgeschlagene Verfahren [KASU93] wird im Weiteren genutzt.

Die Fuzzy ARTMAP Netze haben also für die hier betrachtete Fragestellung folgende Vorteile:

- ARTMAP Netze gehören zu den Selbstorganisierenden Netzen, können also abhängig von den Eingangsdaten neue Neuronen bil-

---

<sup>67</sup> Winner Takes It All

den. Daher sind diese für die Suche nach neuen, bisher unbekanntem Zusammenhängen gut geeignet.

- Anders als bei vielen anderen Netztypen können die gefundenen Regeln extrahiert werden.
- Durch die Rückkopplung innerhalb des Netzes wird die Stabilität beim Lernprozess erhöht, insbesondere dann, wenn bisher unbekannte Kombinationen von Eingangsparametern vorliegen.
- Durch die Integration der Fuzzy Logik können analoge Eingangsdaten verarbeitet werden.
- Die grundsätzliche Eignung von FAM Netzen für die Analyse hochdimensionaler Messdaten wurde bereits veröffentlicht [LSLR00]

### 7.8.2 Implementierung

Der Algorithmus wurde in C++ unter Linux mit der Entwicklungsumgebung Eclipse implementiert. Dabei wurde auf Erfahrungen aus dem Förderprojekt OEE [OEE05] zurückgegriffen. Im Rahmen dieses Projektes wurden erste Versuche der Behandlung hochdimensionaler Eingangsdaten mit Fuzzy ARTMAP Implementierungen gesammelt [SAPO04].

Die hier neue Lösung der direkten Kopplung von Feature Selection und Neuronalem Netzwerk war mit der beschriebenen Lösung nicht umsetzbar, was eine Neuimplementierung und Erweiterung notwendig machte.

Die folgenden Bedingungen wurden vor der Implementierung definiert und bei der Implementierung eingehalten:

- Die Software muss vollständig automatisierbar sein. Entsprechend ist die Applikation von der Benutzerführung zu trennen. Weiterhin sind die üblichen Unix Konventionen für die Ausgabe und die Fehlerbehandlung einzuhalten [RITH78].
- Eine durchgängige Versionskontrolle ist notwendig.
- Die Datenschnittstelle besteht aus dem OTC Dateiformat (s. 10.1)
- Das System muss industriell einsetzbar sein. Insbesondere sind Reproduzierbarkeit und Robustheit zwingende Voraussetzungen.

- Jedes Systemteil muss reentrant<sup>68</sup> kodiert sein.
- Die Implementierung soll sehr lange nutzbar sein. Dazu ist die Abhängigkeit von proprietären Betriebssystemen zu minimieren. Zudem sollten möglichst keine lizenzpflichtigen Softwareteile enthalten sein, insbesondere nicht solche, die nicht im Quelltext vorliegen.
- Ein Ergebnis der Analyse soll spätestens nach 24 Stunden vorliegen.
- Die Software muss unter einem SystemV kompatiblen Unix ablauffähig sein.
- Ergebnisse der Analyse sind mit der Version des erzeugenden Systems gekennzeichnet.
- Das System muss an neue Anforderungen anpassbar sein.

## **A Umfang des Systems**

Der das Neuronale Netz betreffende Teil des Systems besteht aus dem Kernprogramm, in dem die grundlegenden Algorithmen kodiert sind. Daneben ist ein Befehlsinterpreter notwendig, der die interaktiven Benutzeranforderungen sowie die Parameter aus den Automatisierungsskripten extrahiert und das Kernprogramm geeignet steuert.

## **B Automatisierter Softwaretest**

Um die notwendigen Tests effizient und vor allem reproduzierbar durchführen zu können, wurden die Testfälle per Script automatisiert. Wie alle anderen Softwarekomponenten auch sind die Testfälle im Versionskontrollsystem eingecheckt und damit jederzeit rückverfolgbar.

---

<sup>68</sup> Für die Nutzung in multitasking Umgebungen geeigneter Programmcode muss von verschiedenen Instanzen gleichzeitig ohne Seiteneffekte aufgerufen werden können. Dazu dürfen beispielsweise keinerlei globale Daten geändert werden. Auch Schreibzugriffe auf Dateien, Geräte oder Prozesse unterliegt harten Beschränkungen.

### 7.8.3 Parametrisierung

#### A Vigilance

Beim Training des Neuronalen Netzes werden bestehende Verbindungen zwischen den Neuronen gestärkt, wenn ein bereits existierendes Muster erneut auftritt. Gerade im hier beschriebenen Kontext mit sehr vielen Eingangsdaten die darüber hinaus Abbildungen von im Wesentlichen kontinuierlichen Messwerten sind, ist dieses aber die Ausnahme. Der regelmäßig auftretende Fall ist das Auftreten eines noch nicht bekannten Datensatzes. Damit ist die Entscheidung innerhalb des Neuronalen Netzes notwendig, ob die bereits vorhandenen Neuronen für den neuen Datensatz angepasst werden oder ob ein neues Neuron erstellt wird. Diese Entscheidung wird mittels des Vigilance<sup>69</sup> Parameters beeinflusst. Ein hoher vorgegebener Wert fördert die Erzeugung neuer Neuronen, ein niedriger Wert dämpft diese Erzeugung entsprechend.

Faktisch liegt hier eine Analogie zur klassischen Ausgleichsrechnung vor. Eine große Anzahl von Neuronen bildet die Trainingsdatensätze sehr präzise nach, ist also analog der Interpolation zu sehen. Wie auch bei der Interpolation besteht hier das Risiko der Überanpassung (Overfitting). Jeder Einzelfall wird exakt „auswendig gelernt“. Die gesuchten Zusammenhänge zwischen den Daten werden dann nicht erkannt.

Eine zu geringe Anzahl von Neuronen erhöht das Risiko, die signifikantesten Punkte nicht zu erkennen. Der für das jeweilige Problem geeignete Wert für die Vigilance muss experimentell ermittelt werden.

#### B Learning Rate

Anders als die Vigilance, die ja ein kontinuierlich einstellbarer Parameter ist, handelt es sich bei der Learning Rate um eine Mischung aus einstellbarem Parameter und Schalter. Das Setzen von Learning Rate = 1 schaltet das System in den „fast learning mode“. In diesem Modus ist nur ein Durchlauf der Daten notwendig, die Neuronen werden direkt angepasst. Bei Werten für die Learning Rate  $< 1$  steuert die Learning Rate die Ände-

---

<sup>69</sup> Vigilance: Wachsamkeit

rungsrate des Netzes. Da hier kein Bedarf an einem Echtzeitbetrieb ist, wird das System mit ausgeschaltetem „fast learning mode“ betrieben.

Der optimale Wert für die Learning Rate muss wiederum experimentell ermittelt werden.

### **C Class Delimiter**

Wie die Feature Selection ist auch das Neuronale Netz ein Klassifikator. Damit müssen die Ergebnisse in Klassen eingeteilt werden. Diese – manuell durchzuführende– Klasseneinteilung ist ein Prozess, der sehr wesentlichen Einfluss auf die Qualität des Ergebnisses hat. Technisch kann das System sowohl identisch große als auch identisch bevölkerte Klassen über die automatische Berechnung geeigneter Klassengrenzen erzeugen. In der Anwendung spielt die manuelle Definition der Klassengrenzen aber die größere Rolle. Die Ursache dafür ist, dass die Klassengrenzen sinnvollerweise abhängig von der zu lösenden Fragestellung gewählt werden.

#### **7.8.4 Verifikation**

Wir erwarten eine gute Vorhersagequalität des Netzes bei den genannten Problemstellungen. Nur diese gute Vorhersagequalität führt zu ausreichender Sicherheit, dass die extrahierten Regeln valide sind.

Klarerweise muss die gewünschte Funktionalität der Implementierung und des Algorithmus überprüft werden. Hierzu werden Datensätze genutzt, deren Ergebnis bekannt ist und bei denen die funktionalen Zusammenhänge zwischen den Eingangsparametern und dem Ergebnis klar sind. Wichtig ist es hier, die Aussagekraft der Ergebnisse kritisch zu bewerten. Es muss sorgfältig darauf geachtet werden, ob die gefundenen Ergebnisse überhaupt signifikant sind. Dazu ist es jeweils nötig, die korrekt klassifizierten Datensätze mit dem Anteil dieser Klasse am gesamten Datenpool zu bewerten.

### **A Die N–dimensionale Kugel**

Zur Verifikation des Lernerfolges eines Neuronalen Netzes wird häufig das „Kreis im Quadrat“ Problem genutzt [SAPO04]. Hier wird für zwei Parameter die Lernfähigkeit überprüft. Dazu werden entsprechende

(Pseudo) Zufallszahlen im Bereich [0..1] erzeugt und als Koordinaten betrachtet. Klassifiziert werden die Wertepaare dann anhand der Lage innerhalb oder außerhalb eines Kreises um (0,5; 0,5) mit dem Radius 0,5. Nach dem Anlernen mit einem Teil der erzeugten Daten kann dann mit den restlichen Daten direkt geprüft werden, wie das Lernverhalten des Netzes ist.

Wie bereits ausgeführt, ist die hier betrachtete Problemstellung gekennzeichnet durch eine sehr große Anzahl von Dimensionen. Daher wurde das „Kreis im Quadrat“ Problem auf beliebige Dimensionen erweitert.

Für die Überprüfung der korrekten Lernfähigkeit der Implementierung haben wir die N-dimensionale Verallgemeinerung der Kugel genutzt. Dieser Körper umfasst in einem N-dimensionalen euklidischen Raum die Menge aller Punkte, deren Abstand vom Mittelpunkt kleiner gleich dem Radius R ist. Das Volumen der N-Dimensionalen Einheitskugel berechnet sich nach Abbildung 60.

$$V_N = r^N \frac{\pi^{\frac{N}{2}}}{\Gamma(\frac{N}{2} + 1)}$$

Abbildung 60 Volumen N - dimensionale Kugel<sup>70</sup>

Bei einer gegebenen Dimension des Raumes kann für jeden zufällig im Raum erzeugten Punkt einfach festgestellt werden, ob dieser Punkt innerhalb oder außerhalb der Einheitskugel liegt. Nimmt man dieses als Klassifikationsergebnis und lernt mit einer geeigneten Anzahl von Vektoren das Netz an, muss dieses Netz anschließend anhand der Koordinaten die Lage innerhalb oder außerhalb der Kugel voraussagen können.

Bei der Bewertung der Ergebnisse ist zu berücksichtigen, dass das Volumenverhältnis der Einheitskugel zum Einheitsquader bis zur Dimension 5 steigt und dann sehr schnell gegen Null strebt. Die dann sehr unterschiedliche Anzahl der Datensätze in beiden Klassen erschwert die Auswertung erheblich.

Im folgenden Kapitel werden die Ergebnisse des entwickelten Verfahrens vorgestellt. Für die Überprüfung der korrekten Funktion und die initiale

---

<sup>70</sup> Die Gammafunktion ist die eulersche Erweiterung der Fakultätsfunktion für positive, reelle Zahlen.

Parametrisierung des Gesamtsystems wird das „Kreis im Quadrat“ Problem in zwei und drei Dimensionen konkret ausgewertet. Versuche mit einer wesentlich höheren Anzahl von Dimensionen brachten wegen der dramatischen Reduzierung des Kugelvolumens im Verhältnis zum Volumen des Quaders keine signifikante Verbesserung der Ergebnisse.

# 8 Ergebnisse

In diesem Kapitel werden die Ergebnisse der durchgeführten Untersuchungen mit den beschriebenen Werkzeugen vorgestellt. Beginnend mit dem Neuronalen Netz wird zunächst die Funktionsüberprüfung zusammen mit einer initialen Parametrisierung des Netzwerkes anhand von synthetischen Daten durchgeführt. Sodann erfolgt die Analyse von realen Daten. Anschließend wird die Feature Selection überprüft, wobei das Neuronale Netz hier als Validierungsinstanz genutzt wird.

Schließlich erfolgt eine Bewertung der Ergebnisse.

Gegenüber den an anderer Stelle in dieser Arbeit beschriebenen typischen Anwendungsfällen von Neuronalen Netzen ist hier die Analyse der Signifikanz der gefundenen Ergebnisse aus den folgenden Gründen besonders wichtig:

- Das primäre Ziel des hier beschriebenen Verfahrens ist die Identifizierung der Ursachen für Abweichungen und nicht die Vorhersage der Ergebnisse.
- Durch die sich daraus ergebende typischerweise geringe Anzahl von Ergebnisklassen ist die Anzahl der aus Gründen der Statistik korrekten Vorhersagen nicht vernachlässigbar.
- Die Qualitätsanforderungen an die Produkte von den Kunden sind außerordentlich hoch (siehe Kapitel 3).

- Die Daten liegen in unterschiedlicher Granularität (siehe 7.4.2 ) vor. Damit ist nicht in jedem Fall wirklich klar, ob bei der Betrachtung eines einzelnen Datensatzes jedes einzelne Datum für gerade diesen Datensatz uneingeschränkt gültig ist.
- Zudem ist wegen der prinzipbedingten unvollständigen Abbildung der Realität eine nahezu vollständig korrekte Vorhersage des Netzes nicht zu erwarten.

Daher wird bei den Erkennungsquoten der Anteil der „automatisch“ korrekten Klassifizierungen vor der Bewertung des Ergebnisses abgezogen. Es handelt sich hier um die Abschätzung des ungünstigsten Falls, so dass die tatsächlichen Erkennungsquoten keinesfalls kleiner sein können als der angegebene Wert.

### ***8.1 Überprüfung / Parametrisierung Neuronales Netz***

Wie bereits ausgeführt, sind die Parameter Vigilance und Learning Rate für die verwendete Implementierung des Neuronalen Netzes einstellbar. Um einen sinnvollen Startpunkt für diese beiden Parameter zu finden, haben wir mit dem unter 7.8.4 beschriebenen Verfahren die Lernfähigkeit des Netzes analysiert. Dazu wurden Dateien für den zwei- und für den dreidimensionalen Fall der Kugel erstellt. In beiden Fällen wurden jeweils 1000 Datensätze erstellt. Wichtig ist es hier zu bemerken, dass die verwendeten Daten 87 Datenspalten enthalten. Neben den für das Ergebnis relevanten Koordinaten wurden noch verschiedenste Störgrößen hinzugefügt. Das Neuronale Netz soll schließlich in der Anwendung in der Lage sein, relevante Parameter zu identifizieren.

Für das Training des Netzes wurden 80% der Datensätze verwendet. 20% der Datensätze wurden entsprechend für die Validierung verwendet.

Die beiden genannten Parameter sind jeweils pro Datensatz einstellbar. Dabei wird jeweils zuerst das Optimum für die Vigilance gesucht. Sodann wird mit diesem Vigilance Wert das Optimum für die Learning Rate gesucht. Für die Suche des Optimums für die Vigilance wird das Netz in den Fast Learning Mode geschaltet, die Learning Rate wird also auf 1 gesetzt.

Bei unklarer Datenlage wird der kleinere Vigilance Wert gewählt. Wie bereits beschrieben, steigt mit dem Vigilance Wert auch die Interpolation der Datenstrukturen durch das Neuronale Netz. Das Ziel ist jedoch nicht, jeden angelernten Datensatz abzubilden, sondern die den Datensätzen zugrunde liegenden Regeln zu entdecken. Daher handelt es sich hier eher um ein Regressionsproblem, und deshalb wird im Zweifel der Übertragbarkeit der Ergebnisse ein größerer Wert beigemessen als der präzisen Abbildung der einzelnen Datenpunkte. Hinzu kommt noch, dass –wie vielfach in dieser Arbeit beschrieben– praktisch alle vorhandenen Daten mit einer Unschärfe versehen sind und schon deswegen eine Interpolation nicht der richtige Weg sein kann.

### 8.1.1 Vigilance

#### A Kreis

Bei dem verwendeten zweidimensionalen Datensatz liegen 65% der Daten innerhalb des Kreises, 35% liegen außerhalb.

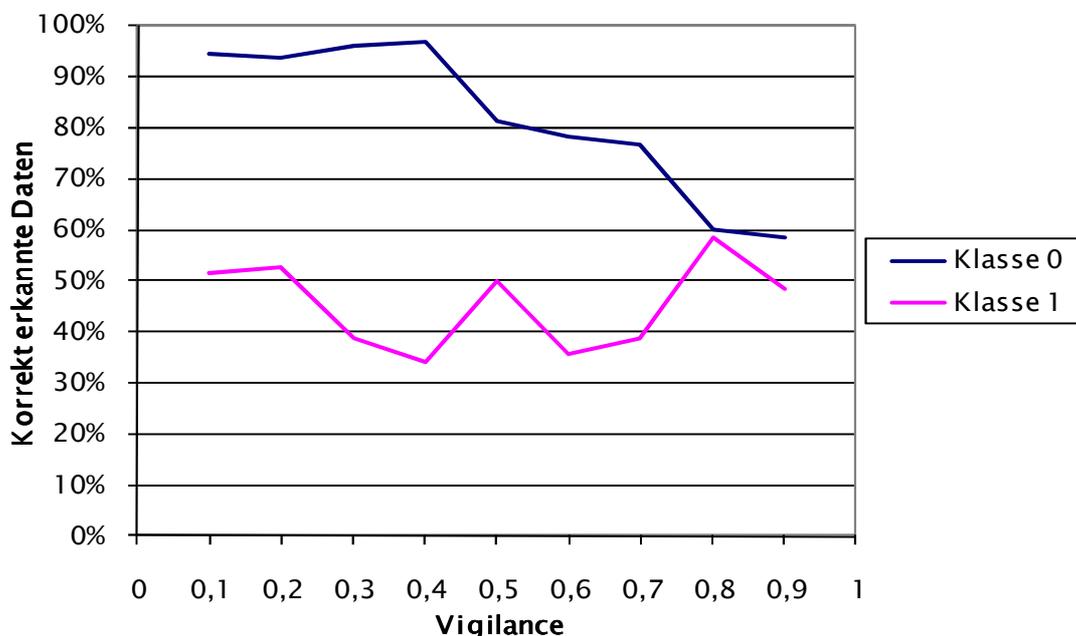


Abbildung 61 Ermittlung der Vigilance für das Kreisproblem

Für neun verschiedene Werte für die Vigilance (0,1; 0,2; .. 0,8; 0,9) wurde jeweils das Netz mit den Trainingsdaten angelernt. Anschließend

wurde die Validierung durchgeführt. Hier wurde der prozentuale Anteil der korrekt durch das Netzwerk zugeordneten Datensätze ermittelt. Dabei bezeichnet die Klasse 0 Punkte, die innerhalb des Kreises liegen. Die Klasse 1 bezeichnet entsprechend die Punkte außerhalb des Kreises.

Abbildung 61 zeigt die korrekt erkannte Klassenzugehörigkeit der Validierungsdatensätze für die verschiedenen Werte für die Vigilance. Bei der Auswertung dieser Daten ist folgendes zu berücksichtigen:

1. Das Ziel der hier durchgeführten Arbeiten ist die Erzeugung von Wissen über die Ursachen der beobachteten Effekte, nicht die möglichst gute Vorhersage einer einzelnen Klasse. Daher ist für die Optimierung der Einstellung hier die Summe der beiden Klassen maßgeblich und nicht der Maximalwert einer Klasse.
2. Weiterhin ist es für die Bewertung der Güte des Algorithmus notwendig, die Signifikanz der Ergebnisse zu untersuchen. Dazu wird der Prozentsatz korrekt erkannte Daten um den Wert reduziert, der sich durch ein zufälliges Würfeln schon aus der Statistik ergibt. Da die Aufteilung des ursprünglichen Datensatzes in Lern- datensätze und Validierungsdatensätze randomisiert erfolgt, wird für die Validierungsdatensätze dieselbe Klassenverteilung angenommen wie die für die Lerndatensätze vorliegende.

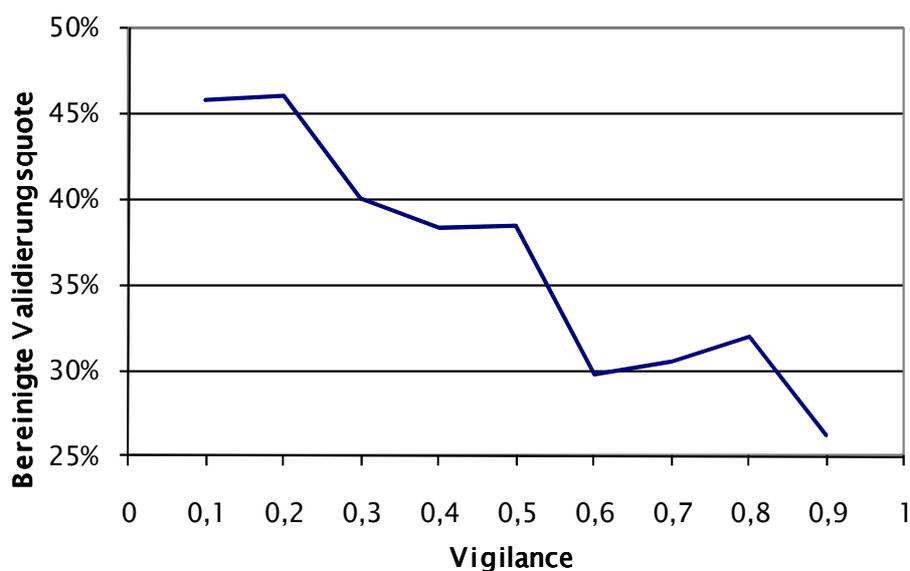


Abbildung 62 Bestimmung Vigilance im zweidimensionalen Fall

## 8 Ergebnisse

---

Aus der Klassenverteilung von 65% in der Klasse 0 und 35% in der Klasse 1 ergeben sich die schon aus Gründen der Statistik korrekten Daten zu 42% für die Klasse 0 und 12% für die Klasse 1.

Die um die automatisch korrekten Daten bereinigte Summe der Validierungsquoten ist in Abbildung 62 dargestellt. Aus dem Kurvenverlauf ergibt sich für diese Implementierung und diese Daten ein Optimum für den Vigilance Parameter von 0,2. Dieser wird im Folgenden verwendet.

Relevant für die Bewertung ist das Maximum der Kurve. Die absolute Quote der korrekt erkannten Klassen wird hier tendenziell zu niedrig bewertet. Die vorgenommene Korrektur um die automatisch richtigen Daten gilt für den ungestörten Fall. Hier wurden aber durch zusätzliche (irrelevante) Datenspalten gestörte Daten analysiert.

### B Kugel

Die Rahmenbedingungen für den Datensatz entsprechen den schon beim Fall Kreis beschriebenen Bedingungen. Wiederum waren es 1000 Datensätze und wiederum waren die Daten durch Hinzufügen von weiteren Spalten verrauscht. Bei dem hier verwendeten Datensatz liegt das Ergebnis von 45,5% der Datensätze in der Klasse 0, 54,5% der Ergebnisse liegen in der Klasse 1.

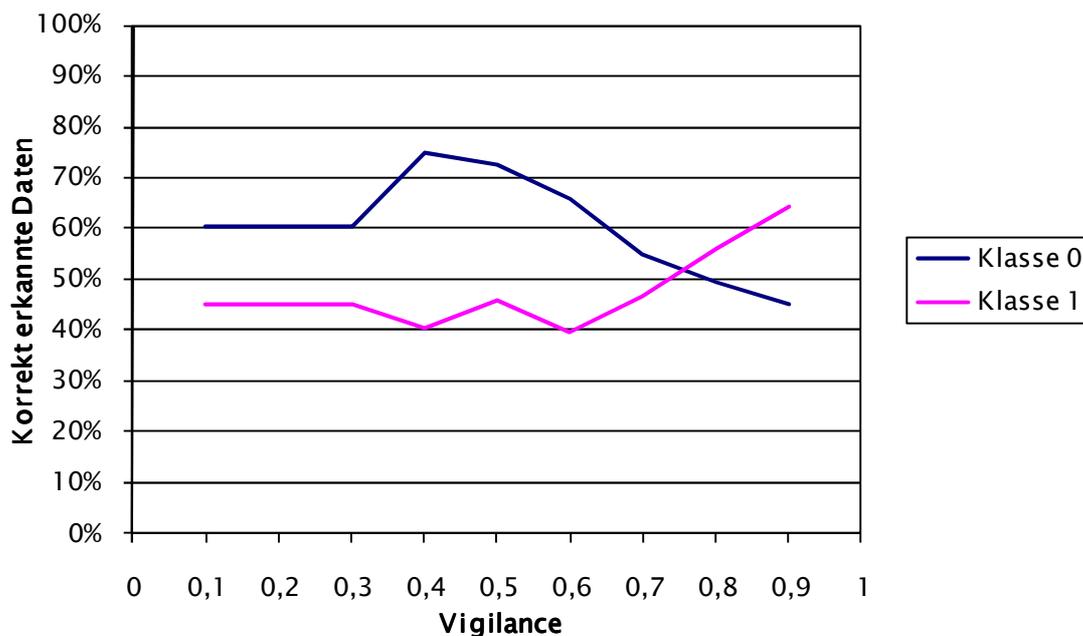


Abbildung 63 Ermittlung Vigilance für die Kugel

Wiederum wurden neun verschiedene Werte für die Vigilance für Training und Validierung des Netzes verwendet. Die Klasse 0 bezeichnet Punkte, die innerhalb der Kugel liegen. Punkte, die zur Klasse 1 gehören, liegen außerhalb der Kugel. Die Ergebnisse für die korrekt erkannte Klassenzugehörigkeit nach der Trainingsphase zeigt Abbildung 63.

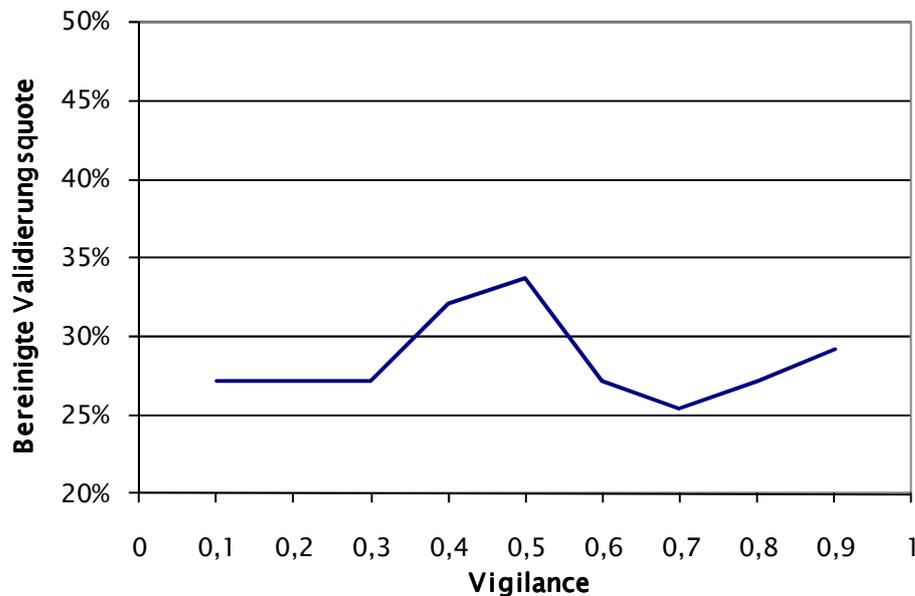


Abbildung 64 Bestimmung Vigilance im dreidimensionalen Fall

Analog zum bei der Analyse des zweidimensionalen Falles beschriebenen Vorgehen wird auch hier wieder die Summe der beiden Klassen gebildet und um die Quote der „zufällig korrekten“ Datensätze vermindert. Diese Quoten ergeben sich hier aus der Klassenverteilung zu 21% für die Klasse 0 und zu 30% für die Klasse 1. Die in Abbildung 64 aufgetragenen Ergebnisse zeigen eine wesentlich geringere Abhängigkeit der korrekten Zuordnung der Klassen von der Vigilance als beim zweidimensionalen Fall, der in Abbildung 62 dargestellt ist.

### C Ergebnisse Vigilance

Die Voruntersuchungen anhand der Daten für das zwei- und dreidimensionale Kugelproblem haben folgende Ergebnisse erbracht:

- Die Implementierung funktioniert wie gewünscht. Das Netz ist nachweislich in der Lage, relevante Informationen aus den Eingangsdaten zu filtern.

- Weiterhin werden Zusammenhänge korrekt erfasst. Für jeden Wert für die Vigilance liegt die Korrektheit der erkannten Daten mit diesen generierten Datensätzen (Abbildung 61, Abbildung 63) deutlich oberhalb des statistischen Wertes. Damit ist deutlich, dass mit dem Parameter für die Vigilance die Güte der Abbildung optimiert werden kann. Es werden aber auch bei einem ungünstigen Wert keine falschen Daten erzeugt.
- Ebenfalls bewiesen ist, dass der optimale Wert für die Vigilance von den verwendeten Daten abhängt. Offensichtlich gibt es Datensätze –wie hier im zweidimensionalen Fall– bei denen die Wahl des Vigilance Parameters einen sehr starken Einfluss hat. Hingegen ist für den dreidimensionalen Fall der Einfluss des Vigilance Parameters wesentlich geringer bis fraglich.
- Nach diesen Untersuchungen ist ein Startwert von 0,2 für die Nutzung von realen Daten naheliegend. Dieser wird einheitlich für die Bestimmung der Learning Rate genutzt.

### 8.1.2 Learning Rate

#### A Kreis

Der bereits unter 8.1.1 für den zweidimensionalen Fall verwendete Datensatz wurde in identischer Form wiederum benutzt. Allerdings wurde hier –nun mit festem Wert von 0,2 für die Vigilance– der Wert für die Learning Rate variiert.

Für zehn verschiedene Werte für die Learning Rate (0,1; 0,2; ... ; 0,9; 1) wurde jeweils das Netz mit den Trainingsdaten angelern. Anschließend wurde wiederum die Validierung durchgeführt und der prozentuale Anteil der korrekt zugeordneten Datensätze ermittelt. Die Klasse 0 bezeichnet weiter die Punkte innerhalb des Kreises, die Klasse 1 bezeichnet die Punkte außerhalb des Kreises [s. Abbildung 65].

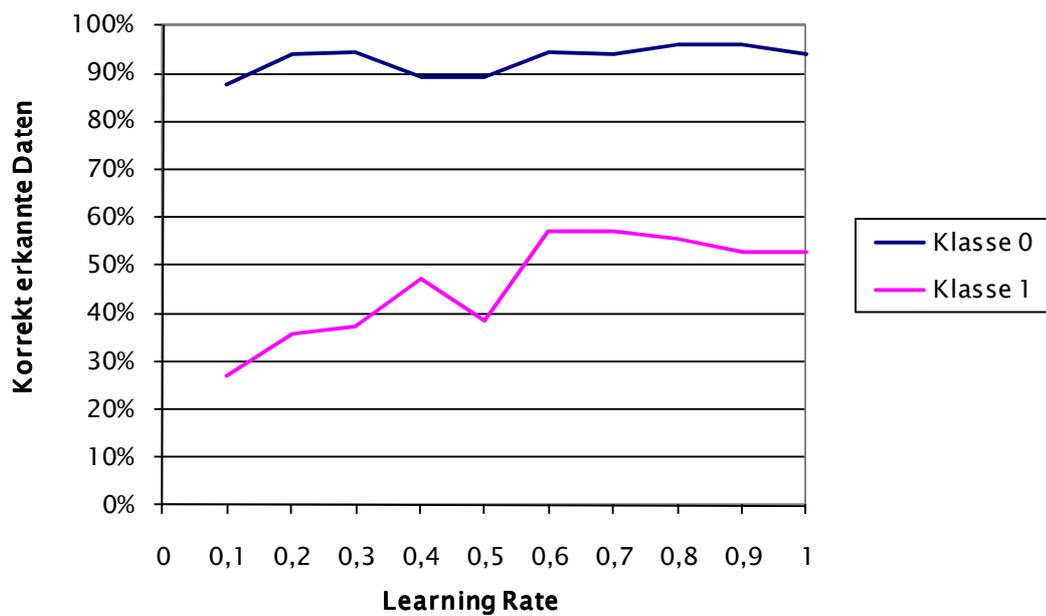


Abbildung 65 Ermittlung Learning Rate für den Kreis

Wie schon bei der Ermittlung des Wertes für die Vigilance wurde auch bei der Bestimmung des Wertes für die Learning Rate die Summe der Prozentzahlen der korrekt erkannten Werte um den Anteil der zufällig richtigen Werte reduziert. Das Ergebnis ist in Abbildung 66 dargestellt.

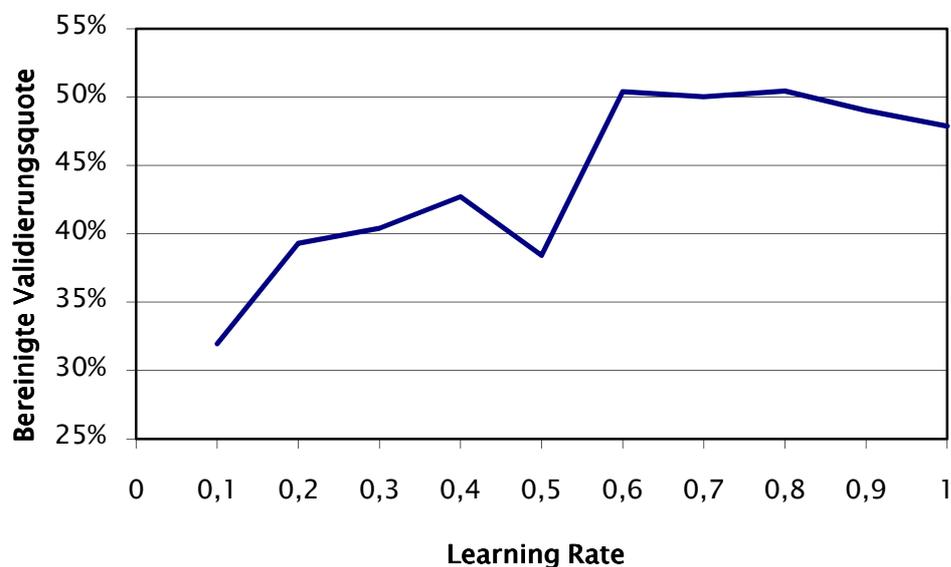


Abbildung 66 Bestimmung Learning Rate im zweidimensionalen Fall

Wiederum ist das Maximum der Kurve für die Bewertung relevant. Auch hier ist wieder die Quote der korrekt erkannten Klassen tendenziell zu

niedrig bewertet, da die Quelldaten durch zusätzliche (irrelevante) Datenspalten verunreinigt waren.

### B Kugel

Auch diese Untersuchung wird analog zu der Bestimmung der Vigilance durchgeführt. Wie dort beschrieben, befinden sich hier 45,5% der Datensätze in der Klasse 0, die restlichen 54,5% befinden sich in der Klasse 1. Allerdings wurde hier der Wert für die Learning Rate variiert.

Für zehn verschiedene Werte für die Learning Rate (0,1; 0,2; ... ; 0,9; 1) wurde jeweils das Netz mit den Trainingsdaten angelernt. Anschließend wurde wiederum die Validierung durchgeführt und der prozentuale Anteil der korrekt zugeordneten Datensätze ermittelt. Die Klasse 0 bezeichnet weiter die Punkte innerhalb der Kugel, die Klasse 1 bezeichnet die Punkte außerhalb der Kugel [s. Abbildung 67].

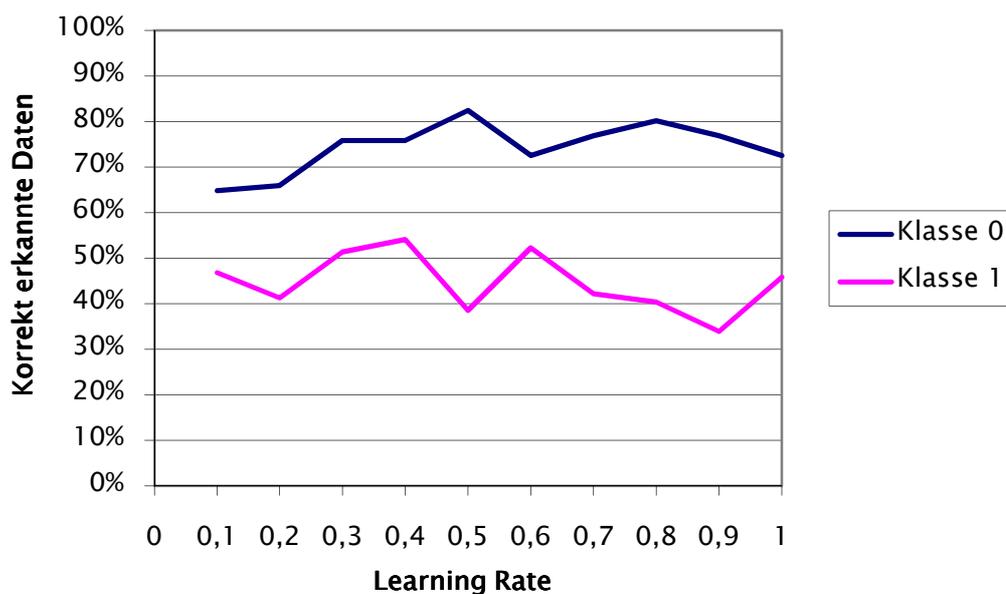


Abbildung 67 Ermittlung Learning Rate für die Kugel

Wie in den bereits beschriebenen Fällen geschehen, werden auch diese Daten wiederum summiert und um die zufällig korrekten Daten vermindert.

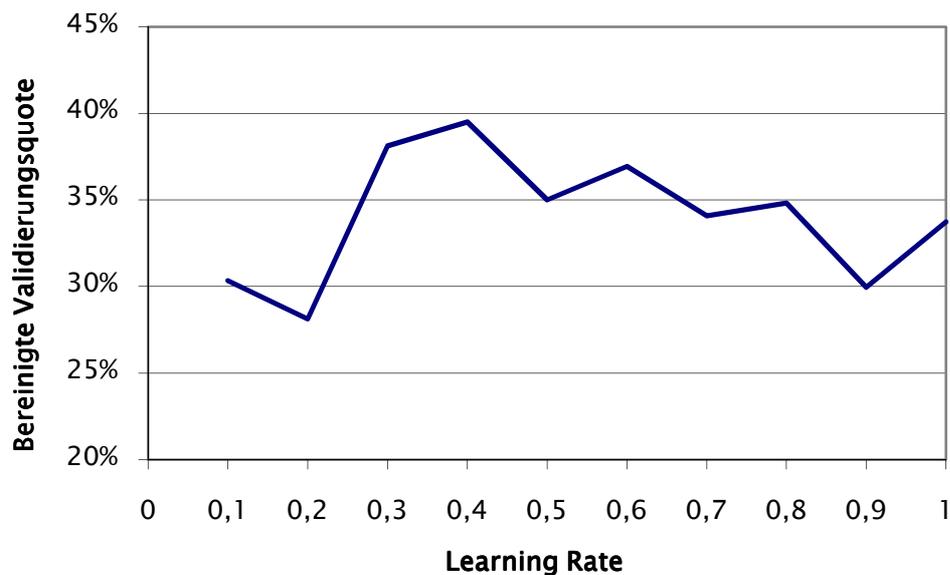


Abbildung 68 Bestimmung Learning Rate im dreidimensionalen Fall

Abbildung 68 zeigt erneut, dass die optimale Wahl beider Parameter von den zu analysierenden Daten abhängt. Wiederum wird deutlich, dass der Einfluss der Parametrisierung ebenfalls stark schwankt.

Anhand synthetischer Daten wurde die von uns durchgeführte Implementierung des Neuronalen Netzes überprüft. In allen untersuchten Fällen hat das Netz die relevanten Informationen aus den bereitgestellten Datensätzen identifiziert und das richtige Verhalten gelernt.

Weiterhin wurden für die Parameter Vigilance und Learning Rate günstige Startbedingungen für die Analyse realer Daten gesucht und gefunden. Die wichtigsten Ergebnisse sind:

- Die Größe des Einflusses der Parameter hängt stark von den Eingangsdaten ab. Jedoch ist in allen Fällen der Lernerfolg vorhanden. Damit beeinflusst die Wahl der Parameter die Güte des Lernerfolges, nicht aber den Lernerfolg an sich.
- Das vorhandene Neuronale Netz kann aus mehr als 80 Parametern die zwei oder drei Parameter korrekt identifizieren, die für das Ergebnis relevant sind.
- Wenn die Parameter einen deutlichen Einfluss auf den Lernerfolg haben, so sind anhand der vorliegenden Ergebnisse folgende Werte naheliegend:

$$\text{Vigilance} = 0,2 / \text{Learning Rate} = 0,8$$

Diese werden für den weiteren Ablauf als Startwerte benutzt.

### ***8.2 Überprüfung Feature Selection***

Die bereits bei der Überprüfung der Funktion des Neuronalen Netzes in Abschnitt 8.1 verwendeten Datensätze für den Kreis und die Kugel werden nun für die Validierung der Implementierung und des Verfahrens der Feature Selection genutzt. Bei der Überprüfung des Neuronalen Netzes war das Ziel herauszufinden, ob das Neuronale Netz die Zusammenhänge korrekt vorhersagt, ohne die zugrundeliegende Funktion zu kennen. Über den in der Literatur üblichen Test hinausgehend, wurden auf der Eingangsseite neben den für die Klassifizierung relevanten Koordinaten noch viele weitere verschiedenen Datenspalten hinzugefügt. Wie gezeigt, hat das gewählte Neuronale Netz die Aufgabe sehr gut gelöst. Bei der nun durchgeführten Untersuchung des Feature Selection Algorithmus werden die identischen Daten genutzt. Im idealen Fall sollte der Feature Selection Algorithmus schnell die für das Ergebnis relevanten Parameter identifizieren. Auch in diesem Fall ist es wieder relevant, die Signifikanz des Ergebnisses sorgfältig zu untersuchen und zu bewerten. Daher sind auch hier wiederum die „zufällig richtigen“ Werte bei der Klassifizierung abzuziehen. Das Ergebnis wird wiederum –wie schon unter 8.1.1 und 8.1.2 – bereinigt und summiert dargestellt.

#### **A Kreis**

Bei dem verwendeten zweidimensionalen Datensatz liegen 65% der Daten innerhalb des Kreises, 35% liegen außerhalb. Der Datensatz besteht, wie auf Seite 161 beschrieben, aus 1000 Zeilen und den um Störgrößen angereicherten Koordinaten.

Der KNN-Algorithmus wurde mit  $K=3$  durchgeführt, wie in Abbildung 56 dargestellt. Abbildung 69 zeigt die korrekt klassifizierte Daten für jede Featureanzahl. Wiederum bezeichnet die Klasse 0 die Punkte innerhalb des Kreises, während die Klasse 1 die Punkte außerhalb des Kreises enthält.

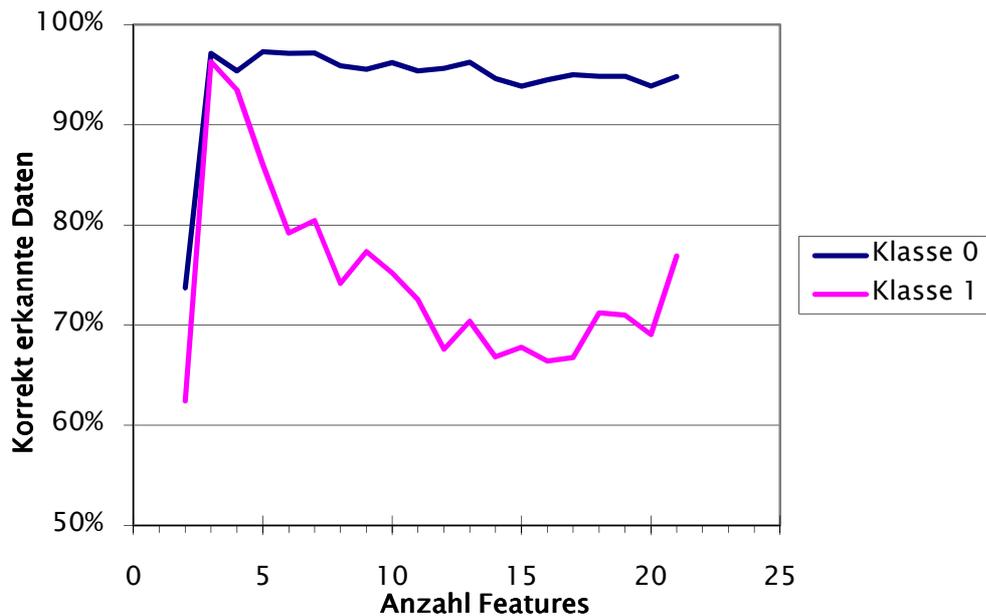


Abbildung 69 KNN Ergebnisse Kreis in Quadrat

Bereits bei den noch nicht bereinigten Daten ist klar erkennbar, dass der Algorithmus sehr schnell die relevanten Features findet. Bereits bei drei selektierten Features werden praktisch alle Datenpunkte korrekt klassifiziert.

Ebenfalls gut zu erkennen ist, dass die „Floating“ Erweiterung gegenüber dem SFS-Verfahren (s. 7.7.1) gut funktioniert. Es kommt vor, dass die bereits korrekt erkannten Parameter wieder entfernt werden. Durch den zusätzlichen SBS Schritt werden wie geplant bereits gefundene Parameter erneut in Frage gestellt. Dieses dient dazu, gefundene Maxima durch eine gezielte Störung zu verlassen, um ein Verharren in einem lokalen Maximum zu vermeiden. Natürlich wird damit auch das globale Maximum in Frage gestellt. Bei einem geringen Anteil von Datenpunkten in einer Klasse steigt damit die Wahrscheinlichkeit, dieses Maximum wieder zu verlassen. Entsprechend geringer ist die Wahrscheinlichkeit für eine Klasse mit einem großen Anteil an Datenpunkten. Zudem sind bei diesem Datensatz aufgrund des Anteils von 65% der Daten in der Klasse 0 schon 42% der Klassifizierungsergebnisse automatisch richtig. Damit ist das Ergebnis der Klasse 0 in jedem Fall stabiler. Für die Klasse 1 liegt der Anteil der aus Gründen der Statistik richtigen Daten wie beschrieben bei 12%.

Die um die automatisch korrekten Daten bereinigte Summe der Validierungsquoten ist in Abbildung 70 dargestellt. Das bereits in Abbildung

## 8 Ergebnisse

---

69 deutliche Maximum bei der Featureanzahl drei wird bestätigt. Damit ist die korrekte Systemfunktion für diesen Datensatz belegt.

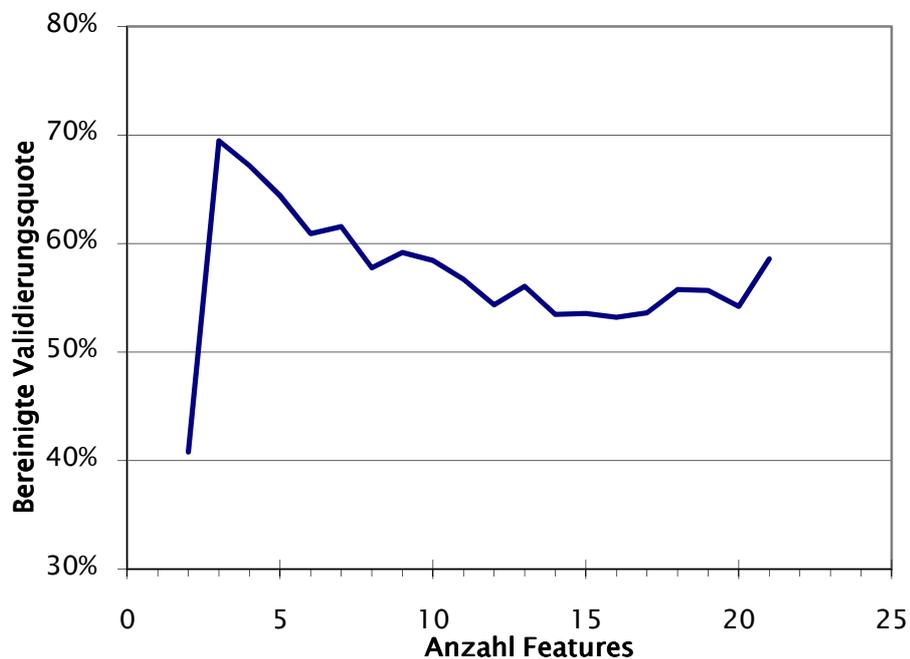


Abbildung 70 Bereinigte KNN Ergebnisse Kreis im Quadrat

### B Kugel

Analog der Durchführung der Feature Selection für den zweidimensionalen Fall wird nun die Untersuchung für den dreidimensionalen Fall durchgeführt. Wie auch schon beim Kreis werden auch hier wieder die unter 8.1 beschriebenen Daten genutzt. Abbildung 71 zeigt die Ergebnisse der durchgeführten Feature Selection, wie zuvor mit einem Wert von  $K=3$ . Wiederum zeigen bereits die noch nicht bereinigten Daten eine sehr schnelle Erkennung der relevanten Parameter. Erneut ist erkennbar, dass der SFFS-Algorithmus wie gewünscht funktioniert. Hier sind beide Klassen etwa gleich groß. Entsprechend zeigen auch beide Klassen ähnliches Verhalten.

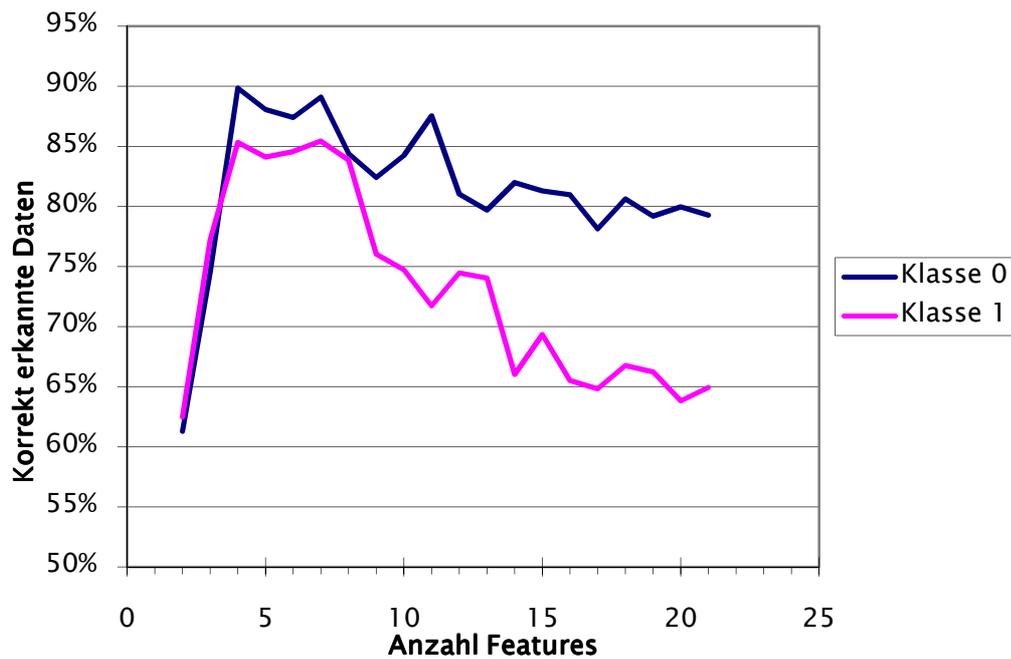


Abbildung 71 KNN Ergebnisse Kugel im Würfel

Abbildung 72 zeigt die bereinigten Daten. Erneut ist der Prozentsatz der automatisch korrekten Werte von den korrekt erkannten Werten subtrahiert worden. Die Summe beider Klassen bestätigt die gute Erkennungsleistung der Feature Selection für diesen Datensatz.

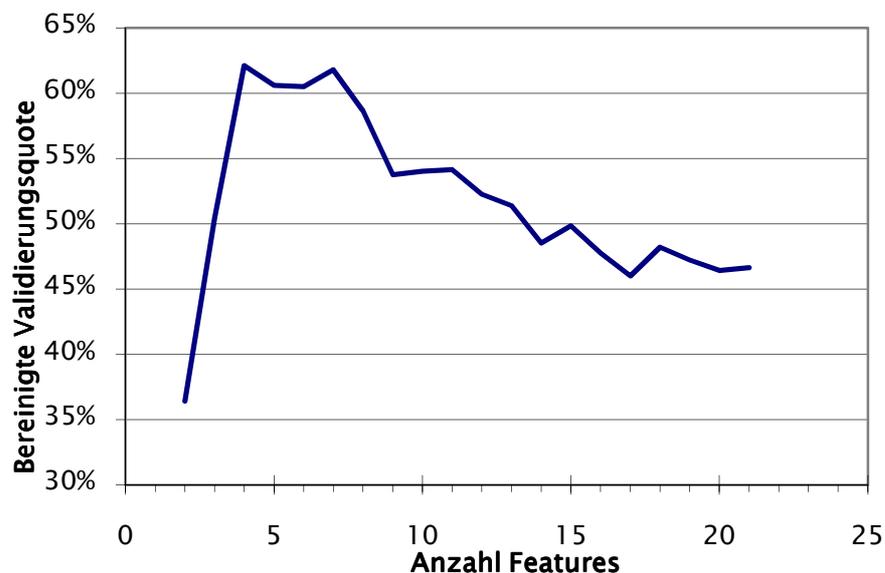


Abbildung 72 Bereinigte KNN Ergebnisse Kugel im Würfel

Nach dem Beleg der Eignung des SFFS/KNN Verfahrens für die Analyse hochdimensionaler Daten (s. Abbildung 59 ) wurde nun hier die korrekte

Funktion des Algorithmus und der Implementierung anhand synthetischer Daten nachgewiesen. Damit ist die Eignung des umgesetzten KNN Algorithmus für die zu untersuchenden Datentypen belegt.

### ***8.3 Ergebnisse mit realen Daten***

Es ist nun belegt, dass sowohl die eingesetzte Sequential Forward Floating Feature Selection (SFFS) mit KNN Klassifikator als auch das Neuronale Netz in der Ausprägung Simplified Fuzzy ARTMAP (SFAM) jeweils einzeln mit synthetischen Daten funktionieren. Der finale Beweis der Eignung des beschriebenen Ansatzes muss nun mit realen Datensätzen aus der Produktion geführt werden. Der in Abbildung 51 auf Seite 116 dargestellte Ablauf wird nun für Daten aus der Realität durchgeführt. Im konkreten Fall liegt für ein Produkt eine Ausbeuteanomalie vor. Diese Ausbeuteanomalie betrifft einzelne Wafer dieses Produktes. Die klassischen Analyseverfahren geben keinen Hinweis auf die Ursache für die Ausbeuteverluste bei einzelnen Scheiben. Daher wurde versucht, mit den hier beschriebenen wissensgenerierenden Verfahren der Lösung des Problems näher zu kommen. Konkret ging es darum zu verstehen, welcher Prozessparameter überhaupt für die Ausbeuteschwankungen verantwortlich war. Um diese Analyse durchführen zu können, werden 83 verschiedene Produktparameter für 1345 Datensätze extrahiert. Konkret handelt es sich um PCM Daten, die jeweils auf Scheibenbasis gemittelt wurden. Zielfunktion ist die Ausbeute, also die Ergebnisse des Funktionstests.

#### **8.3.1 Datenvorbereitung reale Daten**

Es werden also die PCM Daten und die Funktionstestergebnisse für 1345 Scheiben desselben Produktes aus der zentralen Datenbank extrahiert. Anschließend wird entsprechend den in Kapitel 7.6 beschriebenen Verfahren die Datenvorbereitung durchgeführt. Zunächst erfolgen die Elimination der Ausreißer und die Normalisierung. Bei der topologischen Zuordnung werden in diesem Fall die Scheibenmittelwerte gebildet. Schließlich müssen die Daten noch bereinigt werden, um beispielsweise konstante Spalten oder zu gering belegte Zeilen zu entfernen.

### 8.3.2 Klassifizierung reale Daten

Die Ergebnisse wurden in zwei Klassen aufgeteilt. Die Klasse 0 enthält 1275 Datensätze, die Klasse 1 enthält 70 Datensätze. Die Klasse 1 ist der unerwünschte Zustand, und das Ziel der Analyse ist es herauszufinden ob sich diese Einheiten signifikant von den anderen unterscheiden. Auffällig ist zunächst, dass gegenüber den synthetischen Daten die Verteilung auf die Klassen deutlich unsymmetrischer ist.

Dies darf aber den erfolgreichen Einsatz des Werkzeuges nicht behindern. Schließlich ist das Ziel, die Ursache für Prozessabweichungen (durchaus auch zum „Positiven“) zu finden. Das Umfeld ist –wie ebenfalls schon gezeigt– eine vergleichsweise gut kontrollierte Fertigungsumgebung. Daher werden die Abweichungen häufig in der deutlich kleineren Klasse zu finden sein.

### 8.3.3 Bestimmung Vigilance mit realen Daten

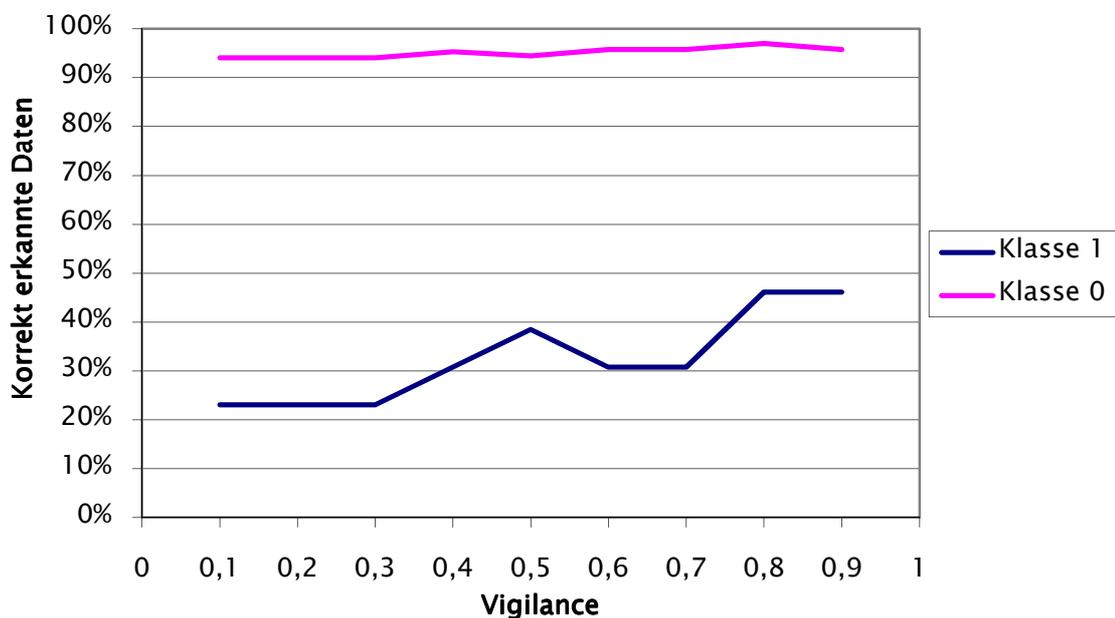


Abbildung 73 Bestimmung Vigilance reale Daten

Wie schon detailliert in den Abschnitten 7.8.3 und 8.1.1 beschrieben, ist für die vorhandenen Daten vor dem sinnvollen Einsatz des Neuronalen Netzes zunächst die Vigilance zu bestimmen. Dazu wird im Fast Learning Mode, also mit dem Wert 1 für die Learning Rate, der Trai-

## 8 Ergebnisse

---

nings/Validierungsablauf des Neuronalen Netzes für verschieden Werte für die Vigilance durchgeführt.

Abbildung 73 zeigt die korrekt klassifizierten Daten abhängig von dem Wert für die Vigilance. Dabei wurde der vollständige Parametersatz genutzt. Analog dem bereits beschriebenen Ablauf, wird nun wiederum der Anteil der automatisch korrekten Werte subtrahiert. Dieser beträgt für die Klasse 0 90%. Der Wert für die Klasse 1 ist aufgrund der geringen Population dieser Klasse mit 0,3% vernachlässigbar.

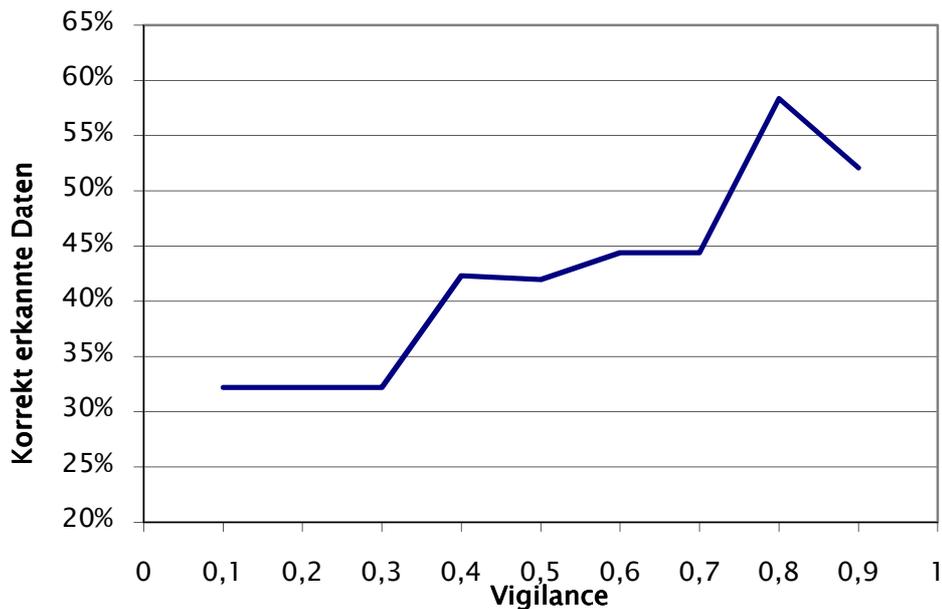


Abbildung 74 Bereinigte Ergebnisse Vigilance mit realen Daten

Wegen dieser stark unterschiedlichen Verteilung der Daten auf die beiden Klassen wurde nach der Subtraktion der automatisch korrekten Werte der restliche Wertebereich wiederum auf 100% normiert. Abbildung 74 zeigt den daraus resultierenden Kurvenverlauf. Offensichtlich verspricht für diese Problemstellung ein Wert von 0,8 für die Vigilance die besten Ergebnisse.

### 8.3.4 Bestimmung Learning Rate mit realen Daten

Mit dem nun identifizierten Vigilance Wert wird die korrekte Einstellung für die Learning Rate ermittelt. Dazu wird mit nun festem Wert für die Vigilance ( 0,8 ) wiederum für den vollständigen Parametersatz die korrekte Klassifizierung durch das Neuronale Netz für verschiedene Werte

für die Learning Rate ermittelt. Das Vorgehen wurde bereits detailliert in den Abschnitten 7.8.3 und 8.1.2 beschrieben. Die Ergebnisse sind in Abbildung 75 dargestellt. Wie bereits mehrfach gezeigt, ist der nächste Schritt die Prüfung der Signifikanz.

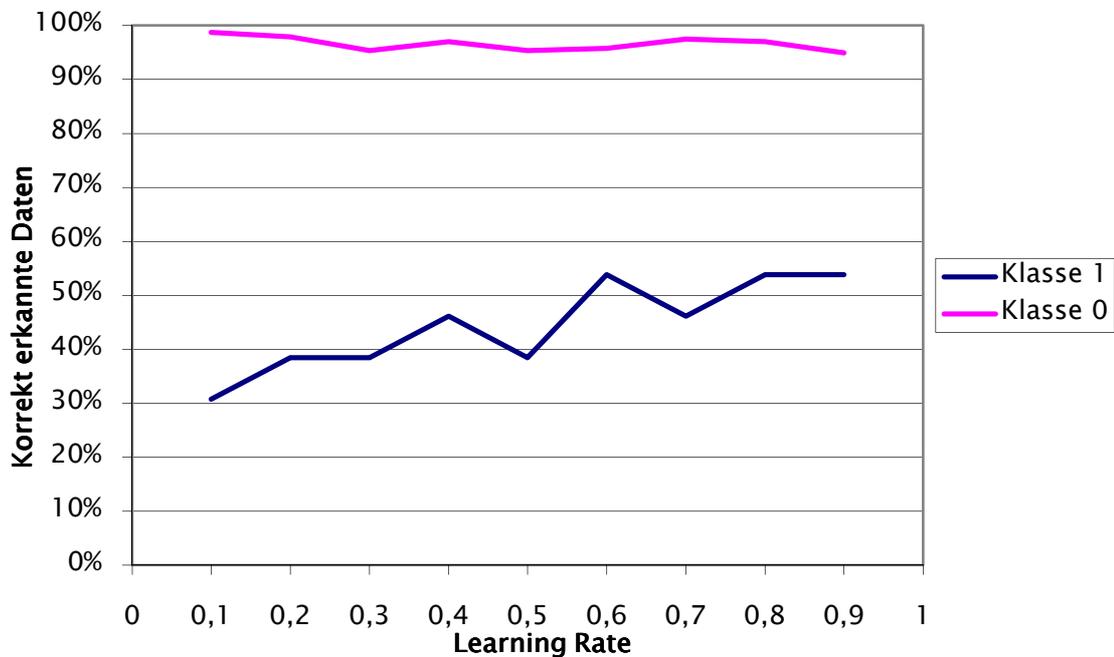


Abbildung 75 Bestimmung Learning Rate reale Daten

Wie bei der Bestimmung des optimalen Wertes für die Vigilance wird wiederum der bereits statistisch richtige Anteil der Daten von den Klassifikationsergebnissen subtrahiert. Wie zuvor beträgt dieser für die Klasse 0 90% und kann für die Klasse 1 vernachlässigt werden.

Wiederum wird der verbleibende Wertebereich auf den ursprünglichen Bereich renormiert. Die Ergebnisse sind in Abbildung 76 aufgetragen. Wie schon bei den synthetischen Daten scheint die Abhängigkeit der Ergebnisse von der Learning Rate weniger signifikant als die Abhängigkeit von der Vigilance. Dennoch legt der Kurvenverlauf eine Festlegung der Learning Rate auf 0,8 nahe. Neben der Tatsache, dass hier tatsächlich das Maximum der Kurve vorliegt, ist hier auch die Schwankung der Kurve vergleichsweise gering.

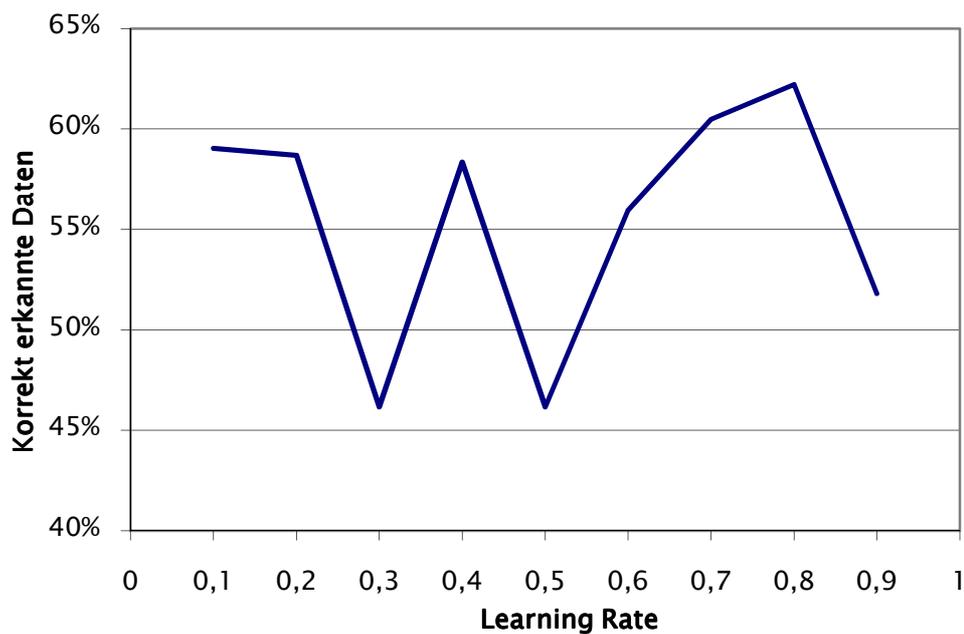


Abbildung 76 Bereinigte Ergebnisse Learning Rate mit realen Daten

Wie bereits bei den synthetischen Daten ist auch hier wieder die Kurve kontinuierlich oberhalb des Nullwertes. Damit ist nun ebenfalls für reale Daten gezeigt, dass das Neuronale Netz auch für nicht optimale Werte für die Parameter keine falschen Ergebnisse erzeugt.

Damit ist die Parametrisierung des Neuronalen Netzes für diesen Datensatz abgeschlossen. Im Weiteren wird für die Analyse dieses Datensatzes sowohl für die Vigilance als auch für die Learning Rate ein Wert von 0,8 festgelegt.

### 8.3.5 SFFS/KNN und SFAM mit realen Daten

Wie schon bei der Überprüfung des Verfahrens mit den synthetischen Daten beschrieben, ist nun der zentrale Schritt des in Abbildung 51, S. 116, © skizzierten Ablaufes an der Reihe. Für ansteigende Festlegungen der maximalen Featureanzahl wird jeweils das SFFS Verfahren sowie anschließend der Trainings/Validierungszyklus des Neuronalen Netzes durchlaufen. Die Überprüfung durch das Neuronale Netz ist notwendig um zu erkennen, ob die für das Ergebnis relevanten Daten sich überhaupt noch in der aktuellen Featureliste befinden.

Dieser Vorgang wurde für Werte von zwei bis 43 für Max Features durchgeführt. Die jeweils korrekt klassifizierten Ergebnisse sind in Abbildung 77 dargestellt. Wiederum ist es wichtig, jeweils die Signifikanz der Ergebnisse zu überprüfen. Zum Vergleich wurde anstelle der Feature Selection eine zufällige Parameterwahl durchgeführt. Diese wird in der Abbildung als „Validation RND“ dargestellt.

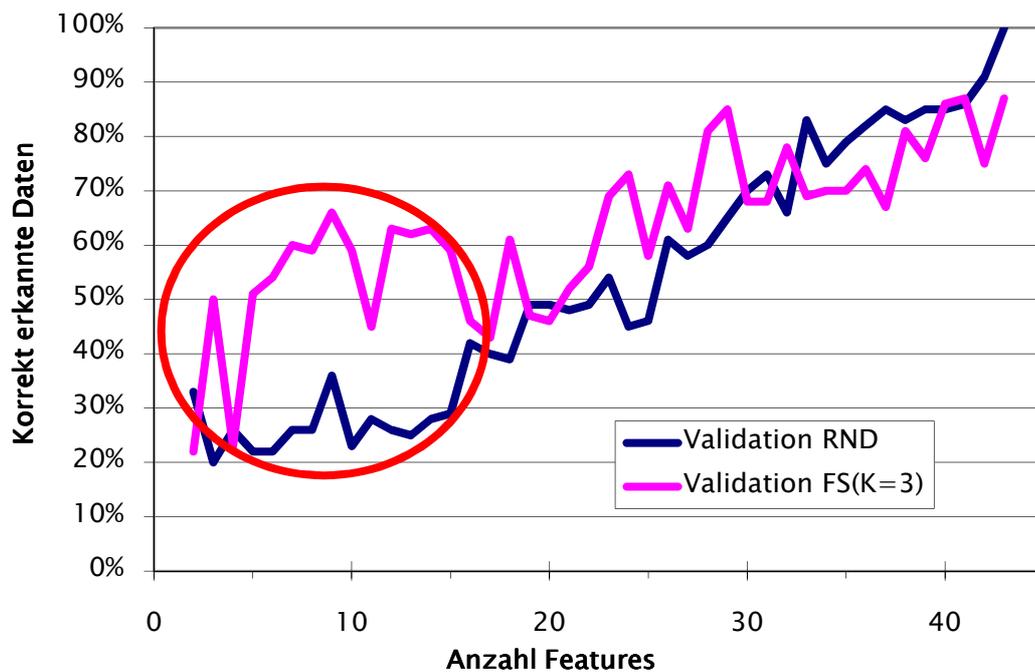


Abbildung 77 SFFS/KNN und FAM mit realen Daten [ISDT09]

Bei der Analyse der Kurve zeigt sich unmittelbar die erwartungsgemäße Funktion des Algorithmus. Die mit „Validation FS“ bezeichnete Kurve stellt die Ergebnisse des kombinierten Ansatzes von Feature Selection und Neuronalem Netz dar. Gerade bei den interessanten kleinen Parameterzahlen ist die Quote der korrekt erkannten Daten mit Feature Selection deutlich besser als mit dem Monte Carlo Vorgehen. Diese Ergebnisse wurden 2009 auf der ISDT vorgestellt [ISDT09].

Auch absolut sind die Ergebnisse besser als erwartet. Es sollte an dieser Stelle nicht vergessen werden, dass das System immer unterbestimmt sein wird und entsprechend nicht über alle notwendigen Informationen verfügt.

Die vorhandenen Messfehler und die ebenfalls potentiell vorhandenen Fehlmessungen sind eine weitere Fehlerquelle.

## 8 Ergebnisse

---

Zusammenfassend ist es hier gelungen, aus einer Menge von über 80 verschiedenen Parametern die Parameter zu identifizieren, die einen Effekt verursachen, der nur in unter 5% der Fälle auftritt. Bereits bei sieben Parametern wird die korrekte Ergebnisklasse mit 60%iger Sicherheit vorhergesagt. Damit ist die Datenmenge, die der bearbeitende Ingenieur untersuchen muss, dramatisch reduziert. Das System kann also die Suche nach den Ursachen für Prozessabweichungen erfolgreich unterstützen.

# 9 Ausblick

Mit der in dieser Arbeit beschriebenen Kombination aus datenreduzierendem (Feature Selection) und wissensgenerierendem (Neuronales Netzwerk) Prozess konnte erfolgreich ein neuer Weg zur Identifikation von Ursachen für Prozessabweichungen in hochkomplexen Fertigungsprozessen aufgezeigt werden. Sowohl jeweils einzeln, als auch in der Kombination zeigen die Untersuchungen sehr gute Erfolge. In allen untersuchten Fällen sind die Ergebnisse signifikant besser als ungerichtete Verfahren. Damit kann das System die Wettbewerbsfähigkeit des Unternehmens wie geplant steigern, da nun ein Werkzeug zur Verfügung steht, das die Ingenieure bei der Suche nach den wirklichen Ursachen der Effekte wirksam unterstützt.

Sowohl der Bedarf [MWK04, FAPC04] für ein derartiges System als auch das Konzept [AMA05, MOBE05, OEE05, KMIS07, KMCM07] und die Ergebnisse [DMIN08, ZUE08, ISDT09] wurden mehrfach durch den Autor publiziert. Zudem ist eine Veröffentlichung akzeptiert [EIT09] und ein weiteres Paper eingereicht [ETFA09].

Zur Vereinfachung der Bedienung wird gerade an der Implementierung einer grafischen Benutzeroberfläche für das System gearbeitet. Mit einer einfacher zu bedienenden grafischen Oberfläche sinken die Hürden für die Benutzung des Systems durch die Ingenieure in der Fertigung.

Das Verfahren funktioniert völlig unabhängig von den technologischen Zusammenhängen. Daher wird es auch auf andere Fragestellungen

übertragbar sein. Erfolgversprechend ist der Einsatz insbesondere dann, wenn eine zuverlässige und zusammenhängende Dokumentation der Abläufe vorhanden ist und die Suche nach der Ursache für eine Abweichung zentrales Ziel der Analyse ist. Beispielsweise könnte die medizinische Diagnostik insbesondere im klinischen Umfeld mit diesem System unterstützt werden. Ein weiteres denkbares Einsatzgebiet könnten Zuverlässigkeitsuntersuchungen im Maschinenbau, insbesondere in der Luftfahrttechnik sein.

Die Integration einer Wichtung der Parameter sollte untersucht werden, sofern die Datenqualität bei der Nutzung in einem anderen Kontext nicht einheitlich ist. Diese Wichtung sollte davon abhängen, wie verlässlich der jeweilige Parameter ist. Als Kriterien für die Wichtung käme beispielsweise der relative Messfehler bei erfassten Messwerten in Frage. Aber auch die Verlässlichkeit der Datenquelle an sich oder die Art der Datennahme (automatisch/manuell) sollte als Kriterium in Betracht gezogen werden.

# 10 Anhang

## *10.1 OTC Dateiformat*

Um die Identifizierbarkeit und Nachverfolgbarkeit sicherzustellen, sind die Datendateien mit zusätzlichen Informationen versehen:

```
# Version : <version>
# User : "<username>"
# Date : <date>
# Columns : <columncount>(a<addresscount>)
# Rows : <rowcount>
# Set_id : <setid>
# Norm_set_id : <normsetid>
#@k1 @k2 ... @kn =e p1 p2 ... pm
kv1 kv2 ... ev e pv1 pv2 ... pvm
.
.
```

wobei gilt

## 10 Anhang

<version>	Version des Otc Formates, stellt die Rückwärtskompatibilität sicher.
<username>	Name des Erstellers, gekapselt in ""
<date>	Erstellungsdatum (YYYY/MM/DD)
<columncount>	Anzahl der Spalten
<addresscount>	Anzahl der Koordinaten für die Identifikation des Datensatzes
<rowcount>	Anzahl der Datensätze
<setid>	Eindeutiger Bezeichner des Selektionsvorgangs
<normsetid>	Eindeutiger Bezeichner des Normierungsvorgangs
k[1-n]	Koordinaten-Namen, wie Charge, Wafer
e	Name des Ergebnisvektors
p[1-n]	Name der Parameter
kv[1-n]	Koordinaten-Werte
ev	Wert für den Ergebnisvektor
pv[1-m]	Normierte Messwerte

### *Beispiel:*

```
# Version : 1.1
# User : "User17"
# Date : 2008/03/28
# Columns : 5(a2)
# Rows : 25
# Set_id : 328
#@Addr1 @Addr1 =FT//234//curr PT//456//vth1 PT//345//vth2
123 1b3 12.35 0.7 0.3
123 1b4 15.45 0.8 0.2
```

---

123 1b5 12.38 0.9 0.1  
123 1b6 12.31 0.7 0.32  
123 1b7 12.32 0.2 0.34  
123 1b8 12.38 0.4 0.35  
123 1b9 12.37 0.6 0.32  
123 1b10 12.45 0.8 0.31  
123 1b11 12.25 0.8 0.38  
123 1b12 12.95 0.8 0.39  
123 1b13 12.65 0.9 0.34  
123 1b14 12.75 0.7 0.32  
124 1b15 12.38 0.5 0.333  
124 1b16 12.39 0.6 0.37  
124 1b17 12.45 0.4 0.38  
124 1b18 12.32 0.8 0.39  
124 1b19 12.75 0.7 0.366  
124 1b20 12.35 0.4 0.35  
124 1b21 12.95 0.5 0.357  
124 1b22 12.34 0.6 0.3552  
124 1b23 12.55 0.7 0.32  
124 1b34 12.36 0.8 0.32  
124 1b56 12.65 0.7 0.301  
124 1b33 12.31 0.9 0.32  
124 1b78 12.30 0.7 0.34

## 10.2 Übersicht Softwarewerkzeuge

Das System ist wie folgt aufgebaut:

Modul	VC	OS	Aufgabe	Sprache
IOBU	SVN	Linux / Solaris	Datenextraktion Ausreißer- Selektion Normierung	Java / SQL
NIODA	CVS	Linux	Feature Selection Neuronale Netze	C++ (GCC)
	SVN	Linux	Testdatengenerator	Java
	CVS	Linux	Automatisierung	BASH

Es wurde eine 32 Bit Variante des Betriebssystems Linux eingesetzt. Bei Solaris handelt es sich um ein 64 Bit Betriebssystem.

Die Spalte VC bezeichnet das eingesetzte Versionskontrollsystem, hier entweder Subversion (SVN) oder das Concurrent Versions System (CVS). Ziel ist in jedem Falle die Absicherung des Softwareentwicklungsprozesses durch Rückverfolgbarkeit und der Zwang zur eindeutigen Versionierung.

### ***10.3 Abkürzungsverzeichnis***

AMHS	Automated Material Handling System, automatisches Schienentransportsystem häufig an der Decke des Fertigungsbereiches.
ANN	Artificial Neural Network: Künstliches Neuronales Netz
APC	Advanced Process Control
ART	Adaptive Resonance Theory: Netztyp der Künstlichen Neuronalen Netze. Unterstützt nur Unsupervised Learning und nur binäre Eingangsdaten.
ARTMAP	Erweiterung des ART Verfahrens um Supervised Learning zu unterstützen. Unterstützt wie ART nur binäre Eingangsdaten.
ASCII	Ursprünglich: American Standard Code for Information Interchange, 7 Bit Code für Zeichenkodierung. Wird heute häufig -unpräzise- für reine Textdateien benutzt, auch dann wenn diese auf erweiterten Kodierungen beruhen.
ASIC	Application Specific Integrated Circuit: Applikations- und damit häufig kundenspezifische integrierte Schaltung
B2B	Business to Business: Kunden- Lieferantenbeziehung zwischen Unternehmen
B2C	Business to Customer: Endkundengeschäft
Bildfeld	Ein Stepper bildet die Fotomasken auf den gerade belichteten Wafer ab. Das verkleinerte Abbild der Masken wird als Bildfeld bezeichnet.
Binning	Die Klassifikation von Bauteilen beim Test wird in der Halbleiterindustrie generell als Binning bezeichnet.

BIST	Built-In Self-Test: Selbsttest von Bauteilen typischerweise nach Anlegen der Versorgungsspannung, auch in der Applikation.
BMBF	Bundesministerium für Bildung und Forschung
BMU	Bundesministeriums für Umwelt, Naturschutz und Reaktorsicherheit
BMVBS	Bundesministerium für Verkehr, Bau und Stadtentwicklung
DOA	Dead on Arrival: Direktausfall bei der Inbetriebnahme durch den Kunden
DRAM	Dynamic Random Access Memory: Dynamischer, wahlfrei adressierbarer Speicher
DUT	Device Under Test: Bezeichnung des Bauteils, das sich gerade im Testaufbau für die elektrische Prüfung befindet.
ECC	Error Correcting Code: Verfahren zur Speicherung von Daten derart, dass Speicherfehler korrigiert oder mindestens erkannt werden können.
EDA	Engineering Data Analysis
EEPROM	Electrically Erasable Programmable Read Only Memory: Halbleiterspeicher der nach der Programmierung die Daten auch ohne Versorgungsspannung hält (PROM). Durch geeignete Verfahren ist diese Programmierung elektrisch löschtbar.
ERP	Enterprise Ressource Planning: Verallgemeinerte Bezeichnung für Unternehmenssteuerungssoftware die im Allgemeinen neben den klassischen Themen wie Finanz- und Personalbuchhaltung auch logistische und planerische Komponenten umfasst.
FAM	Fuzzy ARTMAP: Kombination aus Fuzzy Logik und ARTMAP zur Unterstützung von analogen Eingangsdaten.

---

FDC	Fault Detection & Classification, manchmal auch Fault Detection & Control
FLASH	Eigentlich vollständig FLASH – EEPROM. FLASH ist eine Weiterentwicklung der EEPROMS. Häufig werden zwischenzeitlich über unterschiedliche Spannungszustände (Ladungsmengen) auf dem Floating Gate mehrere Bits pro Zelle gespeichert. FLASH – EEPROMS lassen sich Byteweise beschreiben, allerdings nur blockweise löschen.
Foundry	Halbleiterhersteller der Produkte seiner Kunden mit eigenen Prozessen fertigt. Typischerweise werden die prozessierten Scheiben nach Bewertung der PCM Ergebnisse verkauft. Die Verantwortung für die Funktion der Schaltung trägt der Auftraggeber.
GEM	Generic Model for Communications and Control of Manufacturing Equipment
Guard	Zusätzliche (aktive) Schirmung bei der Messung kleiner Ströme.
HSMS	High-Speed SECS Message Services
ITRS	International Technology Roadmap for Semiconductors
JEDEC	Ursprünglich Joint Electron Devices Engineering Council, heute Solid State Technology Association[JEDEC08], führender Entwickler von Standards in der Halbleiterindustrie
Kelvin	Die Vierpunktkontaktierung für die hochpräzise Spannungsmessung wird auch Kelvin Kontaktierung genannt.
LAN	Local Area Network: auf einen Standort beschränktes, typischerweise IP basiertes, kabelgebundenes Netzwerk.

LDAP	Lightweight Directory Access Protocol: Kommunikationsprotokoll zur Abfrage und Modifikation eines auf einer hierarchischen Datenbank basierenden Verzeichnisdienstes. LDAP ist weit verbreitet. Neben der häufigen direkten Implementierung ist zum Beispiel das Microsoft Active Directory ein abgewandeltes LDAP System.
MCM	Multi Chip Module: Alternative Bezeichnung für SIP
MEMS	Micro-ElectroMechanical Systems: Mikromechanische Produkte wie Druck- oder Beschleunigungssensoren.
MES	Manufacturing Execution System: Fertigungsleitsysteme die typischerweise die Verbindung zwischen der übergeordneten Unternehmensteuerung (ERP) und der Maschinenebene der Fertigung darstellen.
MIL	Mit MIL oder ausführlicher MIL-STD werden die US Verteidigungs- oder Militärischen Standards bezeichnet
MIS	Management Information System: Berichtssysteme für die Leitungsebene
NV-Ram	Non-Volatile Random Access Memory: Allgemeine Bezeichnung für nicht flüchtigen Speicher. Die Inhalte bleiben auch nach Abschalten der Versorgungsspannung erhalten.
OEM	Original Equipment Manufacturer: In der Automobilindustrie übliche Bezeichnung für den Fahrzeughersteller.
OTP	One Time Programmable: Einmal programmierbare Halbleiterstruktur.
PAT	Part Average Test
PCM	Process Control Monitor: Parametertest von prozessabhängigen Teststrukturen auf Wafern.
PPS	Produktionsplanungs- und Steuerungssystem

---

PROM	Programmable Read Only Memory: Elektrisch programmierbarer Halbleiterspeicher, der die Daten nach der Programmierung auch ohne Versorgungsspannung behält.
Quartil	Das Quartil ist eine Sonderform des Quantils. Jedes Quartil beschreibt die Grenze zwischen jeweils 25% der nach Größe selektierten Daten. Damit ist das zweite Quartil der Median.
RBD	Rule Base Dispatch: Beschreibung für typischerweise automatische Auftragsverwaltungssysteme in Fabriken. Das nächste zu fertigende Los wird an Hand von vorher festgelegten Regeln ausgewählt.
Reentranz	Für die Nutzung in multitasking Umgebungen geeigneter Programmcode muss von verschiedenen Instanzen gleichzeitig ohne Seiteneffekte aufgerufen werden können. Dazu dürfen beispielsweise keinerlei globale Daten geändert werden. Auch Schreibzugriffe auf Dateien, Geräte oder Prozesse unterliegt harten Beschränkungen.
REM	Rasterelektronenmikroskop, englisch SEM, Scanning Electron Microscope
ROI	Return On Invest: Zeitpunkt, ab dem eine Investition einen positiven Ergebnisbeitrag erzeugt sowie die Höhe des Beitrages.
SBA	Statistical Bin Analysis
SBS	Sequential Backward Search: Verfahren bei der Feature Selection, es wird jeweils ein Parameter entfernt.
SEC	Statistical Equipment Control
SECS	SEMI Equipment Communications Standard
SEMI	Semiconductor Equipment and Materials Institute
SFAM	Simplified Fuzzy ARTMAP: Um Redundanzen bereinigter Fuzzy ARTMAP Algorithmus.

---

SFS	Sequential Forward Search: Verfahren bei der Feature Selection, es wird jeweils ein Parameter hinzugefügt.
Shop Floor Control	Sowohl in der Literatur als auch in der Anwendung werden die Begriffe Shop Floor Control System und MES synonym gebraucht.
Shrink	Die Umstellung eines Produktes auf eine kleinere minimale Strukturbreite ohne funktionale Modifikationen am Produkt wird als Shrink bezeichnet. In der Theorie sind dazu keine Änderungen am Produkt notwendig. Durch veränderte Entwicklungsregeln und unterschiedliche Bauteile in verschiedenen Prozessen gilt das in der Anwendung zumindest bei nicht reinen Digitalschaltungen nicht vollständig.
SIA	Semiconductor Industry Association
SIP	System In Package: Kombination von mehreren Halbleiterprodukten in einem Gehäuse. Häufig werden hier ASICs, Speicher (FLASH) und / oder Sensoren kombiniert.
SMIF	Standard Mechanical InterFace: Standard für Waferfertigung mit sog. Minienvironments. Der Transport zwischen den Maschinen erfolgt in sog. SMIF PODS, hermetisch geschlossenen Behältern. Diese werden an den Fertigungsmaschinen automatisch geöffnet und das Material entnommen. Das Ziel ist besondere Sauberkeit.
SPC	Statistical Process Control
Stepper	Eine Belichtungsanlage in der Halbleiterfertigung belichtet normalerweise nacheinander Teile der Wafer. Daher werden diese Geräte als Stepper bezeichnet.
SVM	Bei der Support Vector Machine handelt es sich um einen nichtlinearen numerischen Klassifikator. Dieser ist besonders für Zeitreihenanalysen geeignet.
SYA	Statistical Yield Analysis

TCP/IP	Transmission Control Protocol / Internet Protocol
WAN	Wide Area Network, Weitverkehrsnetz: Verbindung von mehreren LANs über größere Entfernungen, sowohl kabelgebunden als auch drahtlos.
WIP	Work in Process: Bezeichnet den Zustand und die Position des angearbeiteten Materials in der Fabrik

## 10.4 Abbildungsverzeichnis

<i>Abbildung 1: Anforderungsviereck modernes Fertigungsunternehmen ..2</i>	
<i>Abbildung 2 Anteil der Kosten elektronischer Systeme an den Produktionskosten eines mittleren PKW, Zahlen nach [ZVEI08].....5</i>	5
<i>Abbildung 3 Übersicht Scheibendurchmesser.....9</i>	9
<i>Abbildung 4 Maximalausbeute abhängig von der Chipfläche .....11</i>	11
<i>Abbildung 5 Verlust eines Produktionsloses [nach MWK04, AMA05] ....15</i>	15
<i>Abbildung 6 Canon Stepper mit Lackstraße [OEE02] .....16</i>	16
<i>Abbildung 7 Beispiel heterogener Cluster [OEE02b].....17</i>	17
<i>Abbildung 8 Entwicklungskosten für Halbleiterprodukte [FRIES04].....18</i>	18
<i>Abbildung 9 Zwei Dies in einem Gehäuse [FRIES04].....22</i>	22
<i>Abbildung 10 More than Moore / More Moore: [ITRS05], S. 9.....24</i>	24
<i>Abbildung 11 Schematische Darstellung einer Flash Zelle [PAUL03] .....27</i>	27
<i>Abbildung 12 Nozzle Etch process specifications [MIPA01].....29</i>	29
<i>Abbildung 13 SM5651/SM5652, SMI.....30</i>	30
<i>Abbildung 14 2D-Sensorchip [VELT00] .....31</i>	31
<i>Abbildung 15 Handlingsystem Beschleunigungsmessung [MUL08].....32</i>	32
<i>Abbildung 16 SM5822/SM5872, Kointegrierter Drucksensor [SMI07]...33</i>	33
<i>Abbildung 17 SM8522/SM5872 REM Bild des kointegrierten Sensors [SMI07].....33</i>	33
<i>Abbildung 18 Polykristalline Solarzellen [SLIK06].....35</i>	35
<i>Abbildung 19 MES Übersicht.....39</i>	39
<i>Abbildung 20 Zentrale Statusvisualisierung eines Fabrikclusters.....44</i>	44
<i>Abbildung 21 S680 DC/RF Parametric-Test-System [KEITH08] .....52</i>	52
<i>Abbildung 22 Messbereich PCM-System [KEITH08].....54</i>	54
<i>Abbildung 23 Waferprober [EGLAS07] .....56</i>	56
<i>Abbildung 24 Probecard [APR07] .....59</i>	59
<i>Abbildung 25 Final Part Testzelle [nach OEE02].....60</i>	60
<i>Abbildung 26 Arbeitsablauf In-Strip-Test [AMK08] .....62</i>	62
<i>Abbildung 27 In-Strip-Handlingsystem, [MUL08] .....63</i>	63
<i>Abbildung 28 Moore's Law mit aktuellen Produkten [IN07B] .....65</i>	65
<i>Abbildung 29 Trends bei Fertigungs- und Testkosten [IN99].....66</i>	66
<i>Abbildung 30 SPC Werkzeug [ADV2000] .....71</i>	71
<i>Abbildung 31 Gasdruck innerhalb einer Maschine [HERK2004] .....73</i>	73
<i>Abbildung 32 Messwerte auf dem Material [HERK04].....73</i>	73

---

<i>Abbildung 33 Ziel FDC: Frühe Erkennung von Verlusten[CR0L05].....</i>	<i>74</i>
<i>Abbildung 34 Prozessregelung als offene Rückkopplung[MOBE05].....</i>	<i>75</i>
<i>Abbildung 35 Allgemeine APC Lösung [MOBE05].....</i>	<i>76</i>
<i>Abbildung 36 Darstellung von PAT Grenzen [AECQ001].....</i>	<i>79</i>
<i>Abbildung 37 Inspektionsmikroskop NWL-860 [NIK08].....</i>	<i>80</i>
<i>Abbildung 38 Principal Component Analysis [SPZ07].....</i>	<i>82</i>
<i>Abbildung 39 Hotelling <math>T^2</math> [SPZ07].....</i>	<i>82</i>
<i>Abbildung 40 Optimierungsspirale für Fertigungsmaschinen .....</i>	<i>86</i>
<i>Abbildung 41 Informationsgewinnung aus Daten [AKA99].....</i>	<i>96</i>
<i>Abbildung 42 Flexible Ablaufverwaltung [OEE02].....</i>	<i>100</i>
<i>Abbildung 43 Wissensverarbeitung [OEE02].....</i>	<i>101</i>
<i>Abbildung 44 Veröffentlichungen nach Thema und Region .....</i>	<i>105</i>
<i>Abbildung 45 Allgemeine Herstellungsfunktion einer einzelnen Fertigungseinheit .....</i>	<i>110</i>
<i>Abbildung 46 Allgemeine Funktion mehrerer gefertigten Einheiten ...</i>	<i>111</i>
<i>Abbildung 47 Allgemeine Funktion für die Fertigung .....</i>	<i>111</i>
<i>Abbildung 48 Konkrete Funktion .....</i>	<i>112</i>
<i>Abbildung 49 Abbildungsfunktion auf Klassen.....</i>	<i>112</i>
<i>Abbildung 50 Zielspezifikation Software Architektur [nach KMIS07]... </i>	<i>114</i>
<i>Abbildung 51 Darstellung des Ablaufes .....</i>	<i>116</i>
<i>Abbildung 52 Übersicht Gesamtsystem.....</i>	<i>120</i>
<i>Abbildung 53 Wafermap mit Bildfeldern und PCM Positionen .....</i>	<i>125</i>
<i>Abbildung 54 Transformationsfunktion zur Normierung der Daten ...</i>	<i>134</i>
<i>Abbildung 55 Allgemeine Normalisierungsfunktion .....</i>	<i>134</i>
<i>Abbildung 56 Ablauf Feature Selection .....</i>	<i>141</i>
<i>Abbildung 57 Beispiel KNN Standard .....</i>	<i>145</i>
<i>Abbildung 58 KNN mit linearem Distanzmaß .....</i>	<i>146</i>
<i>Abbildung 59 Fallbeispiel Feature Selection [MOBE05].....</i>	<i>149</i>
<i>Abbildung 60 Volumen N – dimensionale Kugel.....</i>	<i>158</i>
<i>Abbildung 61 Ermittlung der Vigilance für das Kreisproblem.....</i>	<i>162</i>
<i>Abbildung 62 Bestimmung Vigilance im zweidimensionalen Fall.....</i>	<i>163</i>
<i>Abbildung 63 Ermittlung Vigilance für die Kugel.....</i>	<i>164</i>
<i>Abbildung 64 Bestimmung Vigilance im dreidimensionalen Fall .....</i>	<i>165</i>
<i>Abbildung 65 Ermittlung Learning Rate für den Kreis.....</i>	<i>167</i>
<i>Abbildung 66 Bestimmung Learning Rate im zweidimensionalen Fall</i>	<i>167</i>
<i>Abbildung 67 Ermittlung Learning Rate für die Kugel.....</i>	<i>168</i>
<i>Abbildung 68 Bestimmung Learning Rate im dreidimensionalen Fall .</i>	<i>169</i>

---

*Abbildung 69 KNN Ergebnisse Kreis in Quadrat..... 171*  
*Abbildung 70 Bereinigte KNN Ergebnisse Kreis im Quadrat..... 172*  
*Abbildung 71 KNN Ergebnisse Kugel im Würfel ..... 173*  
*Abbildung 72 Bereinigte KNN Ergebnisse Kugel im Würfel ..... 173*  
*Abbildung 73 Bestimmung Vigilance reale Daten..... 175*  
*Abbildung 74 Bereinigte Ergebnisse Vigilance mit realen Daten..... 176*  
*Abbildung 75 Bestimmung Learning Rate reale Daten..... 177*  
*Abbildung 76 Bereinigte Ergebnisse Learning Rate mit realen Daten.. 178*  
*Abbildung 77 SFFS/KNN und FAM mit realen Daten [ISDT09]..... 179*

## 10.5 Literaturverzeichnis

### 10.5.1 Eigene Veröffentlichungen bezogen auf die vorliegende Arbeit

Kennung	Autor(en)	Titel	Jahr	Quelle
AKA99	Absalon, Ehm, Montino, Reichelt, Schmalfuß, Schmelzle, Schmidt, Suntrup	Ausweitung der Informations- und Kommunikationssysteme mit der deutschen Mikroelektronik.	1999	Persönliche Kommunikation, Projektskizze ak@de.mi zur Vorlage beim BMBF. Beteiligte Unternehmen: Temic, Infineon, Elmos, ZMD, Thesys, SMST, Bosch, FHG IPA
AMA05	Ralf Montino	Weiterentwicklungen bei Prozessanforderungen in der Mikrofertigung von Siliziumbauteilen	2005	AMA Fachverband für Sensorik e.V., Fachauschuß Mikrosystemtechnik, 27. April 2005, Frankfurt
CIM99	Ralf Montino	Smart Fabrication CIM Abschlussbericht	1999	BMBF Förderprojekt: Neue Produktionskonzepte für die Halbleiterfertigung Förderschwerpunkt SMART FABRICATION SF-CIM FKZ: 01 M 2964 C

Kennung	Autor(en)	Titel	Jahr	Quelle
COR08	Project Consortium: IVAM, Coventor, Elmos, ITE, PR, Theon, USI, UCA, X-FAB	Customer Oriented Product Engineering of Micro an Nano Devices	2008	A NMP project in FP7 funded by the European Commission NMP2-SL-2008-213969, <a href="http://www.corona-mnt.eu/index.html">http://www.corona-mnt.eu/index.html</a> , zuletzt gesehen 20090501, Elmos Beiträge durch den Autor
DMIN08	R. Montino , C. Sassenberg, C. Weber, M. Fathi, A. Holland	Feature Selection as an Instrument to Improve the Usability of Classification Results of High-Dimensional Data	2008	DMIN'08, The 2008 International Conference on Data Mining, Las Vegas, Nevada, 14. - 17. Juli 2008
EIT07	M. Fathi, A. Holland, R. Montino, T. Schmidt, H. Peuser	Calculating risk of integration relations in application landscapes	2007	2007 IEEE International Conference on Electro/Information Technology, 17-20 May 2007, pp 210-214, Chicago, IL, ISBN: 978-1-4244-0941-9
EIT09	C. Sassenberg, C. Weber, M. Fathi, R. Montino	A Data Mining based Knowledge Management Approach for the Semiconductor Industry	2009	2009 IEEE INTERNATIONAL CONFERENCE on ELECTRO/INFORMATION TECHNOLOGY, June 7-9, 2009, Windsor, Canada
ESAN05	Bensch, Bogdan, Czerner, Linke, Montino, Schröder, Rosenstiel, Soberger,	Feature Selection for High-Dimensional Industrial Data	2005	ESANN 2005, 13th European Symposium On Artificial Neural Networks, Bruges (Belgium), 27-28-29 April 2005

Kennung	Autor(en)	Titel	Jahr	Quelle
ETFA09	Schmidt F. Ansari, C. Sassenberg, R. Montino	Integration of Knowledge Discovery Techniques in the Quality Management Model to Achieve Higher Target Quality	2009	ETFA 2009 – 14th IEEE International Conference on Emerging Technologies and Factory Automation, 22.–26.9 2009, Mallorca, Spain
FAPC04	Barth, Bichlmeier, Frikking, Hettwer, Katte, Krumm, Nutsch, Maciej, Montino, Mörsch, Öchsner, Reuß, Röder, Rudolph, Schneider, Schellenberger, Schuler, Slodowski, Wagner, Wirbeleit, Wolters, Zakrewski, Zwicker	Advanced Process Control-Systeme für die flexible Halbleiterfertigung (fAPC)	2004	Persönliche Kommunikation: Projektskizze fAPC zur Vorlage beim BMBF. Beteiligte Organisationen: Bosch, Elmos, Infineon, Micronas, Philips, Abakus, AIS, Camline, TIP, FHG IISB, FHG IAO, Universität Paderborn
IRI07	Fathi, Zadeh, Silverberg, Montino, Reza, Garmestani	Panel: Knowledge Management and Reuse in Industrial Approaches:	2007	IEEE International Conference on Information Reuse and Integration, IRI, 13. – 15. 8.2007 Las Vegas, NV, USA, ISBN: 1-4244-1500-4,

Kennung	Autor(en)	Titel	Jahr	Quelle
ISDT09	Ralf Montino	Know-how Reuse and Transfer Combined Usage of different knowledge generating processes in the high tech industry	2009	3rd International Conference on Integrated Systems Design and Technology, 11. - 14. Mai 2009, Anchorage, Alaska
KMCM07	Ralf Montino	Fulfill Today And Tomorrow Customer Needs With New Ways Of Data Analysis	2007	KMCM 2007, International Joint Conference on Knowledge Management for Composite Materials, Düsseldorf, Germany 3-6 July, 2007
KMIS07	Ralf Montino	Hochschulkooperation: Wissensmanagement in Mittelständischen Unternehmen	2007	Center for Knowledge Management & Intelligent Systems, Auftaktveranstaltung, Siegen, 14. Mai 2007, Teile der Veröffentlichung gemeinsam mit C. Weber, C. Sassenberg erstellt.
MOBE05	Montino, Bensch, Bogdan, Schröder, Rosenstiel, Czerner, Soberger, Linke, Schmidt	Neural network algorithms for generalised online tool control in medium size semiconductor fabs	2005	6th European Advanced Equipment Control/Advanced Process Control (AEC/APC) Conference, Dublin, Ireland, April 6-8, 2005
MSK97	Montino, Schäfer, Schmelzle, Suntrup, Kumar	Analyse Engine	1997	Smart Fabrication Verbund CIM : Zweiter Offener Workshop München, 10./11. Juli 1997, S. 9-19

---

<b>Kennung</b>	<b>Autor(en)</b>	<b>Titel</b>	<b>Jahr</b>	<b>Quelle</b>
MWK04	R. Montino, K. Weyer, R. Krumm	APC for SMEs	2004	5th European Advanced Equipment Control/ Advanced Process Control (AEC/APC) Confer- ence, Dresden, 18.-20. April 2004
OEE05	R. Montino, R. Krumm	Abschlussbericht	2005	BMBF Verbundprojekt Overall Equipment Effec- tiveness, neues Konzept zur Effizienzsteige- rung in der Halbleiterfertigung (OEE) FKZ: 01 M 2994 L/8
OEE02	H. Raaf, R. Montino, P. Thies	Corrective Action Framework (CAF)	2002	Zweiter offener Workshop Verbundprojekt OEE, BMBF FKZ: 01 M 2994, November 2002, Stutt- gart
ZUE08	Ralf Montino	Zero Defect: Anforde- rungen an Prozesse und IT Systeme	2008	2. GMM/GI/ITG-Fachtagung 29. September – 1. Oktober 2008, Ingolstadt, ISBN 978-3- 8007-3119-0, S. 115ff, Veröffentlichung in Tagungsband 2009.

## 10.5.2 Sonstige Veröffentlichungen

Kennung	Autor(en)	Titel	Jahr	Quelle
ADV00	Fiorletta, Carl	Implementation of APC in a Production Fab	2000	1st European Advanced Equipment Control / Advanced Process Control (AEC/APC) Workshop, Dresden
AECQ001	Automotive Electronics Council	GUIDELINES FOR PART AVERAGE TESTING	2003	AEC – Q001 – Rev-C, <a href="http://www.aecouncil.com">http://www.aecouncil.com</a> , zuletzt gesehen 20090215
AECQ002	Automotive Electronics Council	GUIDELINES FOR STATISTICAL YIELD ANALYSIS	2000	AEC – Q002 – Rev A, <a href="http://www.aecouncil.com">http://www.aecouncil.com</a> , zuletzt gesehen 20090215
ALL02	INSTITUT FÜR DEMOSKOPIE ALLENSBACH	ACTA 2002	2002	MARKTPLATZ INTERNET Ergebnisse aus der neuen Allensbacher Computer- und Telekommunikations-Analyse (ACTA 2002), <a href="http://www.ifd-allensbach.de/news/prd_0220.html">http://www.ifd-allensbach.de/news/prd_0220.html</a> , zuletzt gesehen 20090215
AMAZ07	Amazon EU S.à.r.l.	Website	2007	<a href="http://www.amazon.de">www.amazon.de</a> , zuletzt gesehen 20090215
AMIN98	Adnan Amin et al.	Advances in Pattern Recognition	1998	Proceedings of Joint IAPR International Workshop SSPR'98 and SPR'98, Sydney, Australia, August 1998, ISBN 3540648585 <a href="http://www.amkor.com/services/test/striptest.cfm">http://www.amkor.com/services/test/striptest.cfm</a> , zuletzt gesehen 20090215
AMK08	Amkor Technology, Inc.	Amkor Strip Test Overview		
ANNA08	Australian Government	The Artificial Neural Network for Australia	2008	<a href="http://www.coastalwatch.com/AboutANNa.htm">www.coastalwatch.com/AboutANNa.htm</a> , zuletzt gesehen 20090217

Kennung	Autor(en)	Titel (ANNA)	Jahr	Quelle
APP06	Applied Materials	Applied WorkStream	2006	<a href="http://www.appliedmaterials.com/products/workstream.html">http://www.appliedmaterials.com/products/workstream.html</a> , zuletzt gesehen: 20090215
APP08	Applied Materials	Product List	2008	<a href="http://www.appliedmaterials.com/products/factoryworks_2.html">http://www.appliedmaterials.com/products/factoryworks_2.html</a> , zuletzt gesehen 20090215
APP09	Applied Materials	Applied Promis	2009	<a href="http://www.appliedmaterials.com/products/promis_2.html">http://www.appliedmaterials.com/products/promis_2.html</a> , zuletzt gesehen 20090215
APR07	Accuprobe, Inc.	METAL BLADE PROBES	2007	<a href="http://www.accuprobe.com/probes.php">http://www.accuprobe.com/probes.php</a> , zuletzt gesehen 20090215
BARRY08	E. Barry	Validation and Six Sigma	2008	Medical device technology, November/Dezember 2008
BELLE08	Vaishak Belle, Thomas Deselaers, Stefan Schiffer	Randomized Trees for Real-Time One-Step Face Detection and Recognition	2008	Proceedings of IAPR 2009, nineteenth conference of the International Association for Pattern Recognition, 8.-11. Dezember 2008, Tampa, Florida, ISBN: 978-1-4244-2175-6
BMBF06	BMBF	Nanotechnologie – eine Zukunftstechnologie mit Visionen	2006	Hightech Strategie BMBF, <a href="http://www.bmbf.de/de/nanotechnologie.php">http://www.bmbf.de/de/nanotechnologie.php</a> , zuletzt gesehen 20090321
BMU06	BMU	Nanotechnologie: Chancen für die Umwelt – Schutz von Ressourcen	2006	Veröffentlichung des BMU, <a href="http://www.bmu.de/gesundheits_und_umwelt/nanotechnologie/chancen_fuer_gesundheit/umwelt/doc/38364.php">http://www.bmu.de/gesundheits_und_umwelt/nanotechnologie/chancen_fuer_gesundheit/umwelt/doc/38364.php</a> , zuletzt gesehen 20090321

Kennung	Autor(en)	Titel	Jahr	Quelle
BMU07	BMU	Info Nanotechnologien	2007	<a href="http://www.bmu.de/gesundheit_und_umwelt/nanotechnologie/kurzinfo/doc/37260.php">http://www.bmu.de/gesundheit_und_umwelt/nanotechnologie/kurzinfo/doc/37260.php</a> , zuletzt gesehen 20090215
BMVBS09	A. Braun, S. Kies, W. Luther, A. Zweck	Status Quo und Entwicklungsperspektiven der Nanotechnologie in Ostdeutschland		Kurzfassung der Studie zur Zukunftskonferenz Nanotechnologien am 19.2.2009, VDI Technologiezentrum GmbH, Eine Studie im Auftrag des Bundesministeriums für Verkehr, Bau und Stadtentwicklung
BOSCH06	Robert Bosch GmbH	Bosch to build new semiconductor fab in Reutlingen	2006	Press release June 14, 2006, <a href="http://www.bosch-presse.de/TBWebDB/en-US/PressText.cfm?id=2751">http://www.bosch-presse.de/TBWebDB/en-US/PressText.cfm?id=2751</a> , zuletzt gesehen 20090215
CGMRR92	Carpenter, G.A., Grossberg, S., Markuzon, N., Reynolds, J.H., Rosen, D.B.	Fuzzy ARTMAP: A neural network architecture for incrementalsupervised learning of analog multidimensional maps	1992	IEEE Transactions on Neural Networks, Sep 1992, Volume: 3, Issue: 5, pp 698–713, ISSN: 1045–9227
CGR91	Carpenter, G. A., Grossberg, S., and Reynolds, J. H	ARTMAP: Supervised real-time learning and classification of nonstationary data by self-organizing neural network.	1991	Neural Networks, vol. 4, pp. 565– 588, 1991, Pergamon Press plc.

Kennung	Autor(en)	Titel	Jahr	Quelle
CGRO91	G. A. Carpenter, S. Grossberg, D. B. Rosen	Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system.	1991	Neural Networks, vol. 4, issue 6, pp. 759–771, 1991, ISSN:0893–6080
CHO08	Siu–Yeung Cho, Teik–Toe Teoh	A Real–time Facial Emotion Recognition System Using Fuzzy Neural Networks	2008	ICAI'08, The 2008 International Conference on Artificial Intelligence, Las Vegas, Nevada, 14. – 17. Juli 2008
CROL05	van Herk, Doclot, Labourier, Henderson, Martins	Data Quality Improvements & Future Needs for AEC APC in Crolles2	2005	6th European Advanced Equipment Control/Advanced Process Control (AEC/APC) Conference/ Jury's Hotel Ballsbridge Dublin, Ireland, April 6–8, 2005
CXW08	Jair Cervantes, Xiaoou Li, Wen Yu	Multi–Class SVM for Large Data Sets Considering Models of Classes Distribution	2008	DMIN'08, The 2008 International Conference on Data Mining, Las Vegas, Nevada, 14. – 17. July 2008
DASH97	M. Dash, H. Liu	Feature Selection for Classification	1997	Intelligent Data Analysis 1, 1997, pp131–156 ,Elsevier Science B.V. PII: S1088–467X(97)00008–5
DEBA07	Denner, Bauer et al.	Qualität bestimmt die Wertschöpfung in der	2007	11. EUROFORUM–Jahrestagung „Elektronik–Systeme im Automobil“, 12. bis 15. Februar

Kennung	Autor(en)	Titel	Jahr	Quelle
		Auto-Elektronik		2007, München
DOHO05	Son Doan, Susumu Horiguchi	Improving Text Categorization by Multicriteria Feature Selection	2005	Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol.9, No.5 pp. 570-575, 2005
EET07	Rick Merrit	Chiphersteller planen Roadmap für 450-mm-Wafer	2007	<a href="http://eetimes.eu/semi//showArticle.jhtml?articleID=177104489">http://eetimes.eu/semi//showArticle.jhtml?articleID=177104489</a> , zuletzt gesehen 20090215
EGLAS07	Electrogilas, Inc.	4090 $\mu$ + Fast Probe	2007	<a href="http://www.electrogilas.com/products/waferpr ober_4090ufp.shtml">http://www.electrogilas.com/products/waferpr ober_4090ufp.shtml</a> , zuletzt gesehen 20090215
FATHI94	Fathi-Torbagan / Höffmann	Fuzzy Logik und Blackboard Modelle in der technischen Anwendung	1994	ISBN 3-486-22650-9
FIGA08	C.J. Figueroa and Jaime E. Araya	A Neural Mining Approach for Predicting Customer Loyalty in Large Retail Databases	2008	DMIN'08, The 2008 International Conference on Data Mining, Las Vegas, Nevada, 14. - 17. July 2008
FRIES04	R. de Fries	The Future of Process Control in Addressing Nano-Technology Challenges	2004	5th European Advanced Equipment Control/Advanced Process Control (AEC/APC) Conference, Dresden, 18.-20. April 2004
FU03	Zhouyu Fu, Terry	Boosted Band Ratio Fea-	2006	icpr,pp.1059-1062, 18th International Confe-

Kennung	Autor(en)	Titel	Jahr	Quelle
	Caelli , Nianjun Liu, Antonio Robles-Kelly	ture Selection for Hyper- spectral Image Classifi- cation		rence on Pattern Recognition (ICPR'06) Volume 1, 2006
GEHRM06	Valentin Gherman	Scalable Deterministic Logic Built-In Self-Test	2006	Dissertation, Institut für Technische Informatik, Fakultät Informatik, Elektrotechnik und Infor- mationstechnik, Universität Stuttgart
GRO76	S. Grossberg	Adaptive pattern classi- fication and universal recoding, I: Parallel de- velopment and coding of neural feature detectors.	1976	Biological Cybernetics, vol. 23, pp. 121-134, 1976, Springer Verlag.
GRO87	S. Grossberg	Competitive learning: From interactive activa- tion to adaptive reson- ance.	1987	Cognitive Science, vol. 11, Issue 1, pp. 23-63, 1987
HAN08	D. Han	Neuro-fuzzy models in real-time flood forecast- ing	2008	DMIN'08, The 2008 International Conference on Data Mining, Las Vegas, Nevada, 14. - 17. July 2008
HECH04	Hechenbichler, Schliep:	Weighted k-Nearest- Neighbor Techniques and Ordinal Classification	2004	LMU München, Institut für Statistik, Sonderfor- schungsbereich 386, Paper 399

Kennung	Autor(en)	Titel	Jahr	Quelle
HEIG08	Georg Heigold, Thomas Deselaers, Ralf Schlüter, Hermann Ney	Modified MMI/MPE: A Direct Evaluation of the Margin in Speech Recognition	2008	Proceedings of ICML 2008, 25th International Conference on Machine Learning, 5. – 9. July 2008, Helsinki, Finland
HERK04	Joost van Herk	Fault Detection and Classification	2004	5th European Advanced Equipment Control/Advanced Process Control (AEC/APC) Conference, Dresden, April 14–16, 2004
HWAN08	Yuan-Chu Hwang	Pervasive Knowledge Reusability for Leisure e-Service Provision	2008	IKE'08, The 2008 International Conference on Information and Knowledge Engineering, Las Vegas, Nevada, 14. – 17. July 2008
IBM08	IBM	SiView	2008	<a href="http://www-03.ibm.com/industries/electronics/doc/content/solution/322070212.html">http://www-03.ibm.com/industries/electronics/doc/content/solution/322070212.html</a> , zuletzt gesehen 20080705
IN05	Paolo Gargini	Silizium-Roadmap – 8 nm im Jahr 2017	2005	Intel Developer Forum 2005, <a href="http://www.tecchannel.de/test_technik/news/422150/">http://www.tecchannel.de/test_technik/news/422150/</a> , zuletzt gesehen 20090215
IN05A	Intel Cooperation	Intel To Build New 300 mm Wafer Factory In Arizona	2005	Intel news release, July 25, 2005, <a href="http://www.intel.com/pressroom/archive/releases/20050725corp.htm">http://www.intel.com/pressroom/archive/releases/20050725corp.htm</a> , zuletzt gesehen 20090215
IN07	Intel	"Copy Exactly" Factory Strategy	2007	<a href="http://www.intel.com/pressroom/kits/manufacturing/copy_exactly_bkgrnd.htm">http://www.intel.com/pressroom/kits/manufacturing/copy_exactly_bkgrnd.htm</a> , zuletzt gesehen 20090215

Kennung	Autor(en)	Titel	Jahr	Quelle
IN07B	Intel	„Moore's Law“ today	2007	<a href="http://www.intel.com/technology/mooreslaw/index.htm">http://www.intel.com/technology/mooreslaw/index.htm</a> , zuletzt gesehen 20090215
IN99	Sanjay et. Al.	Defect-Based Test: A Key Enabler for Successful Migration to Structural Test	1999	Intel Technology Journal Q1'99
IST08	Intel, Samsung, TSMC	Übergang zu 450 mm Wafer Herstellung	2008	<a href="http://www.intel.com/cd/corporate/pressroom/emea/deu/391579.htm">http://www.intel.com/cd/corporate/pressroom/emea/deu/391579.htm</a> Zuletzt gesehen 20090215/ <a href="http://www.samsung.com/us/business/semiconductor/newsview.do?news_id=913.0">http://www.samsung.com/us/business/semiconductor/newsview.do?news_id=913.0</a> Zuletzt gesehen 20080608/
ITRS05	Arden, Coge, Graef, Goltz, Ishiuchi, Osada, Moon, Sohn, Liang, Lin, Lu, Apte, Doering, Gargini	INTERNATIONAL TECHNOLOGY ROADMAP FOR SEMICONDUCTORS: 2005 Edition	2005	<a href="http://www.itrs.net/Links/2005ITRS/Home2005.htm">http://www.itrs.net/Links/2005ITRS/Home2005.htm</a> , zuletzt gesehen 20090215
ITRS06	Arden, Coge, Graef, Ishiuchi, Osada, Moon, Sohn, Liang, Lin, Lu, Apte, Doering, Gargini	INTERNATIONAL TECHNOLOGY ROADMAP FOR SEMICONDUCTORS: 2006 Update	2006	<a href="http://www.itrs.net/Links/2006Update/2006UpdateFinal.htm">http://www.itrs.net/Links/2006Update/2006UpdateFinal.htm</a> , zuletzt gesehen 20090215
JEDEC08	JEDEC Solid State	JEDEC is the leading de-	2008	<a href="http://www.jedec.org">http://www.jedec.org</a> , zuletzt gesehen

Kennung	Autor(en)	Titel	Jahr	Quelle
	Technology Association	veloper of standards for the solid-state industry.		20090215
JIND08	Priyanka Jindal et al.	A NEURAL NETWORK FOR PROGNOSIS OF CHRONIC MYELOID LEUKEMIA	2008	ICAI'08, The 2008 International Conference on Artificial Intelligence, Las Vegas, Nevada, 14. - 17. Juli 2008
KARIM08	Maryam Karimzadehgan1, Gevena G. Belford, Farhad Oroumchian	Expert Finding by Means of Plausible Inferences	2008	IKE'08, The 2008 International Conference on Information and Knowledge Engineering, Las Vegas, Nevada, 14. - 17. July 2008
KASU93	T. Kasuba	Simplified Fuzzy ART-MAP	1993	AI Expert, November 1993, pp. 18-25
KEITH08	Keithley Instruments, Inc.	Series S600 Parametric Test Systems	2008	<a href="http://www.keithley.com/products/semiconductor/aptsystems/?mn=S680">http://www.keithley.com/products/semiconductor/aptsystems/?mn=S680</a> , zuletzt gesehen 20090215
KIT99	Kitauchi Akira, Utsuro Takehito, Matsumoto Yuji	Probabilistic Model Learning for Japanese Morphological Analysis by Error-driven Feature Selection	1999	Transactions of Information Processing Society of Japan, ISSN:0387-5806 VOL.40;NO.5;PAGE.2325-2337
KOHA95	Ron Kohavi	A study of cross-validation and bootstrap	1995	Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI),

Kennung	Autor(en)	Titel	Jahr	Quelle
		for accuracy estimation and model selection		1995, pp. 1137–1143
KORY08	D. Korytina, P. A. Graf, R. King, W.B. Jones	Expressing high-throughput data in a re-structurable, integrated form for knowledge extraction	2008	IKE'08, The 2008 International Conference on Information and Knowledge Engineering, Las Vegas, Nevada, 14. – 17. July 2008
KRAM02	James R. Kramer	Importance of an Open System Architecture in the Semiconductor Industry	2002	Technology Review Journal · Spring/Summer 2002 / ISBN 1–929313–06–3. pp 123–136
KUM08	Anuj Kumar, Anupam Shukla	Improving Document Classification Using Anchor Text	2008	IKE'08, The 2008 International Conference on Information and Knowledge Engineering, Las Vegas, Nevada, 14. – 17. July 2008
LAMB06	Lamberg	Model-Based Testing of Automotive Electronics	2006	9th Design Automation and Test in Europe Conference, 6. – 9.3. 2006, Munich
LAW99	Rob Law, Norman Au	A neural network model to forecast Japanese demand for travel to Hong Kong	1999	Tourism Management, Volume 20, Number 1, February 1999 , pp. 89–97(9), Elsevier
LI04	Hongzhi, Li	A BIST (built-in self-test) strategy for mixed-	2004	Dissertation Technische Fakultät Universität Erlangen-Nürnberg

Kennung	Autor(en)	Titel	Jahr	Quelle
LIMO98	Huan Liu, Horoshi Motoda	signal integrated circuits Feature Selection for Knowledge Discovery and Data Mining	1998	Kluwer Academic Publishers, Boston–Dordrecht–London, ISBN 079238198X
LSLR00	L. Ludwig, E. Sapojnikova, V. Lunin, W. Rosenstiel	Error Classification and Yield Prediction of Chips in Semiconductor Industry Applications	2000	Neural Computing & Applications, ISSN 0941–0643, Volume 9, Number 3 / Oktober 2000, Seiten 202 – 210
MEMU08	Melzner, Müller–L	Quantitative Assessment of Wiring Layout Quality	2008	2. GMM/GI/ITG–Fachtagung 29. September – 1. Oktober 2008, Ingolstadt, ISBN 978–3–8007–3119–0
MIPA01	Kadel, Reinecke	In Situ Profile Depth Control during Deep Dry Etching for Biomedical Application	2001	2nd European Advanced Equipment Control / Advanced Process Control (AEC/APC) Conference – Dresden, Germany, April 18 – 20, 2001
MIPA07	Boehringer Ingelheim microParts GmbH	Mikrosysteme für die Medizin–, Bio– und Pharmatechnik	2007	<a href="http://www.boehringer-ingelheim.de/unternehmenprofil/standorte/dortmund.jsp">http://www.boehringer-ingelheim.de/unternehmenprofil/standorte/dortmund.jsp</a> , zuletzt gesehen 20090215
MISKO00	Peter Miskowiec	Schaltungsbezogene Modellierung der Ausbeute und des Ausfallrisikos mikroelektronik–	2000	Dissertation Fachbereich Elektrotechnik Gerhard–Mercator–Universität–Gesamthochschule Duisburg

Kennung	Autor(en)	Titel	Jahr	Quelle
		scher Schaltkreise unter Berücksichtigung defektinduzierter Ausfallmechanismen		
MOO65	G. Moore	Cramming more components onto integrated circuits	1965	Electronics, Volume 38, Number 8, April 19, 1965
MUL08	Multitest Elektronische Systeme GmbH	Acceleration MEMS Test Module / Shaker In-Strip-Test	2008	<a href="http://www.multitest.de/media/media_0092720.pdf">http://www.multitest.de/media/media_0092720.pdf</a> , Zuletzt gesehen 20090215 <a href="http://www.multitest.de/eng/page_002036.aspx">http://www.multitest.de/eng/page_002036.aspx</a> , zuletzt gesehen 20090215
NAC08	Anatoli Nachevand Seamus Hill	Business Failure Prediction Using Distributed ARTMAP Neural Networks	2008	ICAI'08, The 2008 International Conference on Artificial Intelligence, Las Vegas, Nevada, 14. - 17. Juli 2008
NIK08	NIKON INSTRUMENTS INC.	NWL-860 IC inspection wafer loader	2008	<a href="http://www.nikoninstruments.com/products.php?n=NWL-860">http://www.nikoninstruments.com/products.php?n=NWL-860</a> , zuletzt gesehen 20090215
OEE02b	G. Wagner, K. Bader, D. Hornung et al.	Cluster Tools	2002	Zweiter offener Workshop Verbundprojekt OEE, BMBF FKZ: 01 M 2994, November 2002, Stuttgart
ORA05	Oracle	Oracle® Database SQL Language Reference 10g	2005	Oracle part number: B14200-02 <a href="http://download-uk.oracle.com/docs/cd">http://download-uk.oracle.com/docs/cd</a>

Kennung	Autor(en)	Titel	Jahr	Quelle
				/B19306_01/server.102/b14200/functions110.htm, zuletzt gesehen 20090217
PAUL03	Manfred Paul	Bauelemente der Technischen Informatik	2003	Vorlesung SS2003, Fachbereich IV, Universität Trier
PFT02	Addy, Buchholz, Franke, Middendorft	Einsatz produktbegleitender Informationssysteme und ihre Auswirkungen auf die Produktionstechnik	2002	DFG Sonderforschungsbereich 281, Förderkennzeichen: 32800-02-SFB281,
POPP05	Jens Popp	Ein Konzept für die Datenhaltung in der fertigungsnahen Mikrosystemtechnik	2005	Fachbereich Elektrotechnik und Informatik, Universität Siegen, Dissertation
PUDI94	P. Pudil et al.	Floating Search Methods for Feature Selection with Nonmonotonic Criterion Functions	1994	Proceedings of the 12th IAPR International Conference on Pattern Recognition, ISBN 0818662654, 0818662700, 0818662751, pp. 297-283
REDD08	R. Reddy, L. Wang, S. Reddy, S. Devalapalli, G. Sasanaka, S. Macha, S. Teja, R. Doppalapudi, J. Yu	Vijjana: A Pragmatic Model for Collaborative, Self-organizing, Domain Centric Knowledge Networks	2008	IKE'08, The 2008 International Conference on Information and Knowledge Engineering, Las Vegas, Nevada, 14. - 17. July 2008

Kennung	Autor(en)	Titel	Jahr	Quelle
RITH78	Dennis M. Ritchie, Ken Thompson	The UNIX Time-Sharing System	1978	The Bell System Technical Journal, Vol. 57, July-August 1978, No. 6, Part 2, S. 1897-2312
SAPO04	E. Sapojnikova	ART-based Fuzzy Classifiers: ART Fuzzy Networks for Automatic Classification	2004	Dissertation Universität Tübingen, Technische Informatik, ISBN 3865370578
SCHU99	F. Schuler	Modellierung von nicht-flüchtigen Speicherbauelementen	1999	Dissertation, Institut für Elektronik EIT4, Universität der Bundeswehr, München,
SEM98	Sematech, Factory Integration Division	Computer Integrated Manufacturing (CIM) Framework Specification Version 2.0	1998	<a href="http://ismi.sematech.org/docubase/abstracts/1697jeng.htm">http://ismi.sematech.org/docubase/abstracts/1697jeng.htm</a> , zuletzt gesehen 20090215
SEM30	SEMI©	SEMI E30-0307 – Generic Model for Communications and Control of Manufacturing Equipment	2007	Semi Standard, <a href="http://www.semi.org">www.semi.org</a>
SEM4	SEMI©	SEMI E4-0699 (Reapproved 0307) – SEMI Equipment Communications Standard 1 Mes-	2007	Semi Standard, <a href="http://www.semi.org">www.semi.org</a>

Kennung	Autor(en)	Titel	Jahr	Quelle
SEM5	SEMI©	sage Transfer (SECS-I) SEMI E5-1106E – SEMI Equipment Communica- tions Standard 2 Mes- sage Content (SECS-II)	2006	Semi Standard, www.semi.org
SFCIM98	M. Suntrup	Smart Fabrication CIM Framework: ein rich- tungweisendes Konzept	1998	BMBF Förderprojekt: SMART FABRICATION SF-CIM FKZ: 01 M 2964: Dritter Offener Work- shop, München 29.10.–30.10.1998
SHEI08	Eman M., El-Sheikh1, Bradley A. Swain, Mo- hamed A. Khabou	A Comparison of Neural Network Architectures for Handwritten Digit Recognition	2008	ICAI'08, The 2008 International Conference on Artificial Intelligence, Las Vegas, Nevada, 14. – 17. Juli 2008
SHEN08	Haiying Shen, Ting Li, Ze Li	Exploring Efficient and Effective LSH-based Methods for Data Pros- pecting	2008	IKE'08, The 2008 International Conference on Information and Knowledge Engineering, Las Vegas, Nevada, 14. – 17. July 2008
SIN05	Scotten W. Jones,	A Simulation Study of the Cost and Economics of 450 mm Wafers	2005	Semiconductor International, 8/1/2005: <a href="http://www.semiconductor.net/article/CA632496.html">http://www.semiconductor.net/article/CA6324 96.html</a> , zuletzt gesehen 20090215
SKF08	Manfred Franzen	Six Sigma als Dienstleis- tung: Von der Prozess- optimierung zur Effizi-	2008	MAINTAIN09, Trendforum, 15. Oktober 2008, München

Kennung	Autor(en)	Titel	Jahr	Quelle
SLIK06	Georg Slickers	enzsteigerung Polykristalline Silizium- Solarzellen in Photovol- taik-Modul, Q-Cells 5"- Zelle	2006	<a href="http://de.wikipedia.org/wiki/Bild:Polycrystalline-silicon-wafer_20060626_568.jpg">http://de.wikipedia.org/wiki/Bild:Polycrystalline-silicon-wafer_20060626_568.jpg</a> zuletzt gesehen 20090215
SMI07	Silicon Microstructu- res Inc.	Products	2007	<a href="http://www.si-micro.com/">http://www.si-micro.com/</a> , zuletzt gesehen 20090215
SPZ07	Gerhard Spitzlsperger	Fault Detection in Semi- conductor Manufactur- ing	2007	KMCM 2007 International Joint Conference on Knowledge Management for Composite Materials, Düsseldorf, Germany, 3-6 July, 2007
SRMU03	Srinivas Mukkamala1, Andrew H. Sung	Feature Selection for In- trusion Detection using Neural Networks and Support Vector Machines	2003	Journal of the Transportation Research Board of the National Academics, Transportation Research Record No 1822, pp. 33-39, 2003
SSMC07	SSMC	Company profile	2007	<a href="http://www.ssmc.com.sg/companyprofile.html">http://www.ssmc.com.sg/companyprofile.html</a> , zuletzt gesehen 20090215
SUN99	Suntrup, M., Dorner, J.	MES Integration Assess- ment and Influences on IT Projects	1999	Semiconductor FABTECH, Vol.9 (1999), pp.63-67
TAHA99	Hirotoishi Taira, Ma- sahiko Haruno	Feature selection in SVM text categorization	1999	AAAI '99/IAAI '99: Proceedings of the 16th national conference on Artificial intelligence and the 11th Innovative applications of artificial

Kennung	Autor(en)	Titel	Jahr	Quelle
TSAI08	Tsung-Hsien Tsai	Case Based Reasoning Models to Predict Final Sales: A Test of Railway Passenger Arrivals	2008	intelligence conference innovative applications of artificial intelligence, isbn 0-262-51106-1, pages 480-486, Orlando, Florida, DMIN'08, The 2008 International Conference on Data Mining, Las Vegas, Nevada, 14. - 17. July 2008
UCI08	Asuncion, A. Newman, UCI	UCI Machine Learning Repository	2008	<a href="http://www.ics.uci.edu/~mlearn/MLRepository.html">http://www.ics.uci.edu/~mlearn/MLRepository.html</a> , zuletzt gesehen 20090215
UR08	Ulrich Rückert	A Statistical Approach to Rule Learning	2008	Dissertation: Technische Universität München, Fakultät für Informatik, 2008
VELT00	T. Velten	Mikromechanisch gefertigter 3D-Beschleunigungssensor für die Hand-Gebärdenerfassung	2000	Dissertation TU Berlin, FB 12, Elektrotechnik
VOT07	ADENION GmbH	Website	2007	<a href="http://www.votello.de">www.votello.de</a> , zuletzt gesehen 20090215
WAN08	Huanjing Wang, Guangming Xing, Kairui Chen	Categorical Data Transformation Methods for Neural Networks	2008	IKE'08, The 2008 International Conference on Information and Knowledge Engineering, Las Vegas, Nevada, 14. - 17. July 2008
ZAD65	Zadeh, L	Fuzzy sets	1965	Symposium on System Theory, Polytechnic Institute of Brooklyn, April 20, 21, 22 1965,

---

Kennung	Autor(en)	Titel	Jahr	Quelle
				Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers: Selected Papers by Lofti A. Za- deh, ISBN: 9810224214
ZVEI06	ZVEI	General Automotive Quality Agreement for Electronic Components	2006	ZVEI – Zentralverband Elektrotechnik- und Elektronikindustrie e.V.
ZVEI07	ZVEI	Zero Defect Strategy – a joint approach all along the value chain	2007	ZVEI – Zentralverband Elektrotechnik- und Elektronikindustrie e.V.
ZVEI08	ZVEI	Mikroelektronik- Trendanalyse	2008	ELEKTRONIKPRAXIS NR. 6 – 19. März 2008, S 36 –39

## ***10.6 Danksagung***

An dieser Stelle möchte mich bei denjenigen bedanken, deren Unterstützung wesentlich zum Gelingen dieser Arbeit beigetragen hat.

Zunächst ist hier Prof. Dr. Madjid Fathi zu nennen, der diese Arbeit durch die Übernahme der Betreuung erst ermöglichte. Prof. Fathi war jederzeit ein sehr verlässlicher Ansprechpartner und seine Hinweise waren eine wertvolle Hilfe.

Herrn Prof. Dr. Rainer Brück danke ich für die bereitwillige Übernahme des Koreferats. Prof. Brück hat sich freundlicherweise intensiv mit der vorliegenden Arbeit auseinandergesetzt und viele sehr hilfreiche Anregungen beigesteuert. Da das Thema der Arbeit die Fachbereiche beider Gutachter betrifft, war die intensive Diskussion mit beiden Gutachtern für den erfolgreichen Abschluss des Vorhabens außerordentlich hilfreich.

Ferner gilt mein Dank Herrn Prof. Dr. Udo Kelter für die Übernahme des Vorsitzes der Prüfungskommission sowie dem Dekan Prof. Dr. Elmar Griese und dem Prorektor Prof. Dr. Peter Haring Bolivar für die Bereitschaft, trotz der vollen Terminkalender, an der Prüfung als Prüfer teilzunehmen.

Weiterhin gilt mein Dank den Mitarbeitern am Institut für Wissensbasierte Systeme für die Unterstützung, die gerade bei einer externen Promotion sehr hilfreich ist. Mein besonderer Dank gilt hier den Herren C. Sassenberg und C. Weber für die Mitarbeit bei der Erzeugung der Ergebnisse.

Vor allem aber danke ich meiner Frau für die Unterstützung während der Durchführung der Arbeiten.

Dortmund, im Juli 2009

Ralf Montino