

Qualitätsmanagement in der Bosch-Gruppe | Technische Statistik

# 11. Versuchsmethodik (DoE)



**BOSCH**  
Technik fürs Leben





**Inhaltsverzeichnis:**

1.	Einführung in die Versuchsmethodik .....	3
1.1.	Bedeutung von Versuchen .....	3
1.1.1.	Erkenntnisgewinn durch Versuche .....	4
1.1.2.	Versuche und Erprobung .....	8
1.1.3.	Versuche und Qualität .....	9
1.2.	Die Rolle der Statistik beim Versuch .....	10
1.2.1.	Die Quellen von Streuung .....	10
1.2.2.	Prinzipien der statistischen Versuchsplanung .....	11
1.3.	Chancen und Risiken der Versuchsmethodik .....	12
1.3.1.	Vorteile, Erfolgsfaktoren und Stärken .....	12
1.3.2.	Grenzen und Gefahren .....	13
1.4.	Versuchsmethodik bei Bosch .....	14
1.4.1.	Anwendung der Versuchsmethodik im Produktentstehungsprozess (PEP) .....	14
1.4.2.	Qualifizierung und Ansprechpartner .....	15
1.5.	Elementare Regeln der Versuchsplanung und -durchführung .....	16
2.	Aufgabenanalyse .....	19
2.1.	Auftrag, Projektmanagement .....	19
2.2.	Analyse der Ausgangssituation .....	19
2.3.	Festlegung der Versuchsstrategie .....	19
3.	Systemanalyse .....	21
4.	Versuchsplanung .....	23
4.1.	Festlegung der Zielgrößen .....	23
4.2.	Festlegung der Faktoren .....	24
4.3.	Festlegung des Variationsbereichs der Faktoren .....	24
4.4.	Wahl des Modellansatzes .....	25
4.5.	Festlegung der Faktorstufen .....	27
4.6.	Festlegung der Faktorstufenkombinationen .....	28
4.6.1.	Trial-and-error Verfahren .....	28
4.6.2.	One-factor-at-a-time (OFAT) Verfahren .....	28
4.6.3.	Faktorielle Versuchspläne .....	31
4.7.	Festlegung der Anzahl der Wiederholungen .....	32
4.8.	Festlegung der Reihenfolge und Gruppierung .....	33
4.9.	Planung des Versuchsdurchlaufs und der Auswertung .....	33
4.10.	Planung der Versuchseinrichtungen und der Prüflinge .....	34
4.11.	Aufwandsabschätzung .....	34
5.	Versuchsdurchführung und Dokumentation .....	35
5.1.	Physische Versuche .....	35
5.2.	Computereperimente .....	35
6.	Auswertung .....	37
6.1.	Plausibilitätsprüfung .....	37
6.2.	Bestimmung der Modellgleichung .....	37
6.2.1.	Beispiel Einfaktorversuch auf 2 Stufen .....	37
6.2.2.	Beispiel Zweifaktorversuch auf 2 Stufen .....	38
6.2.3.	Versuche mit k-Faktoren auf 2 Stufen .....	40
6.2.4.	Regressionsanalyse .....	41
6.3.	Validierung der Modellgleichung .....	42
6.3.1.	Einfache Signifikanzbewertung bei 2 <sup>k</sup> -Plänen .....	42
6.3.2.	Validierung der Regression .....	43
6.3.3.	Validierung durch Varianzanalyse .....	47
6.4.	Grafische Darstellung .....	50
6.5.	Interpretation der Ergebnisse .....	52
6.6.	Schlussfolgerungen und weiteres Vorgehen .....	53
7.	Anwendungen .....	54
7.1.	Prognose .....	54
7.2.	Optimierung .....	54
7.2.1.	Gradientbasierte Verfahren .....	55
7.2.2.	Simplex Verfahren .....	56
7.2.3.	Evolutionäre Algorithmen .....	57
7.3.	Robustheit- und Zuverlässigkeitsanalysen .....	57
8.	Weiterführende Ansätze .....	59



8.1.	Teilfaktorische Pläne.....	59
8.1.1.	Motivation.....	59
8.1.2.	Effekte und Vermengungen.....	59
8.1.3.	Handlungsempfehlung.....	61
8.2.	Pläne für nicht-lineare Zusammenhänge.....	61
8.2.1.	Pläne vom Typ $3^k$ .....	62
8.2.2.	Zentral zusammengesetzte Pläne.....	62
8.2.3.	D – optimale Pläne.....	63
8.2.4.	Handlungsempfehlung.....	64
8.3.	Alternative Modellansätze.....	64
8.4.	Stochastische Versuchspläne.....	67
8.4.1.	Plain Monte Carlo (PMC).....	67
8.4.2.	Latin Hypercube (LHC).....	68
8.5.	Spezielle Erprobungspläne.....	68
8.6.	Versuchspläne nach Taguchi.....	69
8.6.1.	Motivation: robuste Produkte und Prozesse.....	69
8.6.2.	Versuchspläne nach Taguchi.....	69
8.6.3.	Auswertung und Ergebnis.....	70
8.7.	Shainin® Methode.....	71
8.7.1.	Motivation: Die Suche nach dem Red X®.....	71
8.7.2.	Systematische Beobachtung.....	71
8.7.3.	Einfache Versuche.....	72
8.7.4.	Weitere Tools.....	75
9.	Anhang.....	77
9.1.	Grundlegende Konzepte der Statistik.....	77
9.1.1.	Datenreihen und deren Kennwerte.....	77
9.1.2.	Verteilungen und deren Kennwerte.....	78
9.1.3.	t-Test, Vergleich von zwei Mittelwerten.....	81
9.1.4.	Mindeststichprobenumfang.....	82
9.1.5.	Varianzanalyse.....	84
9.1.6.	Regressionsanalyse.....	87
9.2.	Softwaretools.....	90
9.3.	Tabellen.....	91
10.	Literatur.....	96
11.	Index.....	97

2020-04-06 - SOCCOS



## 1. Einführung in die Versuchsmethodik

Ziel dieses Heftes ist es, Mitarbeitern aus Forschung, Entwicklung und Fertigung einen Überblick über das Thema Versuchsmethodik bei Bosch zu geben. Unter Versuchsmethodik (auf Englisch: Design of Experiments, DoE) wird dabei die systematische Vorgehensweise zur Planung, Durchführung und Auswertung von Versuchen unter Minimierung der dafür erforderlichen Ressourcen verstanden. Die in der Literatur unter der Bezeichnung „DoE“ üblicherweise beschriebene „statistische Versuchsplanung“ wird als Teilaspekt der Versuchsmethodik hier nur soweit behandelt, wie es im Rahmen der üblichen technischen Fragestellungen notwendig ist. Für weitergehende Studien sei auf die einschlägige umfangreiche Literatur verwiesen. Auf eine mathematische Herleitung der Zusammenhänge wird weitestgehend verzichtet.

Der Inhalt des vorliegenden Heftes deckt wesentliche Aspekte der Versuchsmethodik für die ingenieurmäßige Praxis ab, z.B. Versuchsführung und die Rolle der Modellbildung. Darüber hinaus soll es als Leitfaden zur Planung, Vorbereitung und Auswertung von Versuchen dienen. Zur selbständigen Behandlung komplexer Fragestellungen muss dies jedoch durch ausreichende Praxis - Erfahrung ergänzt werden.

Aufbau des Heftes:

- Kapitel 1: Überblick und wichtige Aspekte zum Verständnis der Methode, deren Bedeutung und insbesondere Anwendung bei Bosch.
- Kapiteln 2...7: Detaillierte Beschreibung der Versuchsmethodik in chronologischer Abfolge der Einzelschritte.
- Kapitel 8: Weiterführende Ansätze (Details zu relevanten statistischen Grundlagen siehe Anhang).

### 1.1. Bedeutung von Versuchen

Aufgabe eines Versuchs oder Experimentes ist die, meist empirische, Überprüfung einer Annahme (Lateinisch experimentum: Beweis, Prüfung, Probe). Am realen Objekt wird durch Beobachtung oder Messung festgestellt, ob die Annahme zutrifft oder nicht. Letztendlich ist das Ziel eines jeden Versuches die Erweiterung der Erkenntnis. Einen Versuch durchzuführen, bevor die Frage gestellt wurde, welche beantwortet werden soll, ist nicht zielführend.

Der Versuch ist immer eine Untersuchung des kausalen Zusammenhangs von Ursache und Wirkung, wobei stets der Gedanke eines Prozesses zugrunde liegt, mit Einflüssen am Eingang und Reaktionen am Ausgang. Größen, welche die Einflüsse beschreiben, werden als Einflussgrößen, diejenigen, welche die Reaktionen beschreiben, als Zielgrößen bezeichnet.

Ein Versuch ist die empirische Untersuchung über den kausalen Zusammenhang zwischen Einfluss- und Zielgrößen eines Systems auf der Grundlage von Beobachtung und Messung mit dem Ziel der Erkenntnisgewinnung.

Je nachdem, ob die Einflussgrößen beim Versuch gezielt eingestellt werden können oder nicht, unterscheidet man:

- a) **Versuche unter kontrollierten Bedingungen**, z.B. Laborversuche, bei denen die Einflussgrößen nach dem vorgesehenen Schema eingestellt werden können.
- b) **Versuche unter unkontrollierten Bedingungen**, oft auch Feldversuche genannt, bei denen die Einstellung entweder physikalisch nicht möglich oder aus wirtschaftlichen, ethischen etc. Gesichtspunkten nicht vertretbar erscheint.

Der erste Typ ist nach Möglichkeit immer vorzuziehen, da er ein gezieltes Eingreifen für einen schnellen und wirtschaftlichen Erkenntnisgewinn erlaubt, im zweiten Fall ist dies bedeutend schwieriger.

Eine weitere Unterscheidung ist zwischen den Begriffen **Versuch** und **Versuchsrealisierung** notwendig. Falls die Untersuchung streuende Größen betrifft, sind in der Regel mehrere Versuchsrealisierungen erforderlich. Diese werden in ihrer Gesamtheit dann als Versuch bezeichnet.



### 1.1.1. Erkenntnisgewinn durch Versuche

In der Systemtheorie spielen die Begriffe System und Modell eine zentrale Rolle. Systeme bezeichnen zunächst ganz abstrakt die unterschiedlichsten Objekte der Betrachtung, z.B. eine Kraftstoff-Einspritzanlage, Fertigungsanlage, ein Haus oder aber auch einen Organismus. Modelle dienen als vereinfachte Abbildungen realer Systeme zur Beschreibung von wesentlichen Eigenschaften des Systems wie Struktur und Funktion. Bei der Aufstellung von Modellen spielen Versuche eine entscheidende Rolle.

#### 1.1.1.1. Grundbegriffe, System und Modell

##### Definition System:

BES-PE-Glossar (Ausgabe 3-2009/11): Ein System ist eine Gesamtheit von Elementen, die so aufeinander bezogen sind und in einer Weise wechselwirken, dass sie als eine aufgaben-, sinn- oder zweckgebundene Einheit angesehen werden können und sich in dieser Weise gegenüber der sie umgebenden Umwelt abgrenzen.

Ein System weist eine Systemstruktur auf und erfüllt eine bestimmte Funktion, vgl. [13], S.35. EN ISO 9000: System = Satz von in Wechselbeziehung oder Wechselwirkung stehenden Elementen.

Da die Aspekte Abgrenzung zur Umgebung und Funktion jedoch für die Versuchsmethodik wichtig sind, soll im Folgenden die erste Definition verwendet werden.

##### Definition Modell:

BES-PE-Glossar (Ausgabe 3-2009/11): Ein Modell ist eine auf relevante Eigenschaften reduzierte Abbildung der Wirklichkeit.

Ein Modell ist die vereinfachte, beschränkt gültige Abbildung eines Systems, welche für einen bestimmten Modellzweck gültige Aussagen über das System vermitteln soll, vgl. auch [13], S.51. Dabei ist wichtig zu verstehen, dass ein Modell nicht richtig oder falsch ist, sondern zweckmäßig oder nicht, je nach Aufgabe, die das Modell erfüllen soll. Für verschiedene Fragestellungen können unterschiedliche Modelle desselben Systems erforderlich sein.

Mit der Entwicklung von Modellen beschäftigt sich die **Systemanalyse** als eine der Methoden der Systemtheorie, dieser Prozess wird oft **Modellbildung** genannt.

##### Gründe für eine Modellbildung:

- Reduzierung der Komplexität der Zusammenhänge in einem System
- erhöhtes Verständnis und Erklärung des Systemverhaltens durch Simulation, u.a. dort, wo Studien am realen System nicht möglich, nicht wirtschaftlich oder nicht vertretbar sind
- Prognose oder Optimierung des Systemverhaltens bevor das System physisch existiert.

##### Generelle Vorgehensweise bei der Modellbildung:

1. **Erstellung eines Modells:** meist iteratives, heuristisches Vorgehen, bei dem das Modell, ausgehend vom einfachst-möglichen Modellansatz, durch Erkenntnisse aus der Realität solange angepasst wird, bis alle aus Sicht der zu lösenden Aufgabe wesentlichen Aspekte berücksichtigt sind.
2. **Validierung des Modells:** Vergleich der prognostizierten mit den beobachteten Eigenschaften und Verhaltensweisen des Systems.
3. **Simulation:** Erkenntnisgewinn über das System anhand des Modells.
4. ggf. **Weiterentwicklung des Modells**, anschließend Fortsetzung mit Schritt 2.

Abhängig von den Eigenschaften Struktur und Funktion des Systems, welche das Modell wiedergibt, spricht man von strukturtreuen oder verhaltensgültigen Modellen [13], S.61.

##### Die Abbildung der Systemstruktur erfolgt durch:

- die Systemgrenze (Abgrenzung zur Umgebung),
- die Systemelemente mit Eigenschaften, wobei die Elemente wiederum Systeme (Subsysteme) sein können,



- Wechselwirkungen inkl. Rückkopplung zwischen den Elementen und zwischen System und Umwelt durch Stoff-, Energie- und Informationsflüsse.

Die Systemstruktur hängt entscheidend von der Betrachtungstiefe ab. Zum einen müssen die wesentlichen Einflüsse (z.B. Wechselwirkung zwischen den Elementen) erfasst werden können, zum anderen darf der Detaillierungsgrad das Modell nicht zu komplex werden lassen. Der **Grundsatz** lautet somit: So grob wie möglich, so detailliert wie nötig.

Die Funktion eines (Sub-)Systems beschreibt nicht „wie etwas ist“, sondern „was es tut“, d.h. es werden Verhaltensweisen, nicht Bestandteile nachgebildet [14], S.55.

Die **Abbildung der Funktion** erfolgt durch:

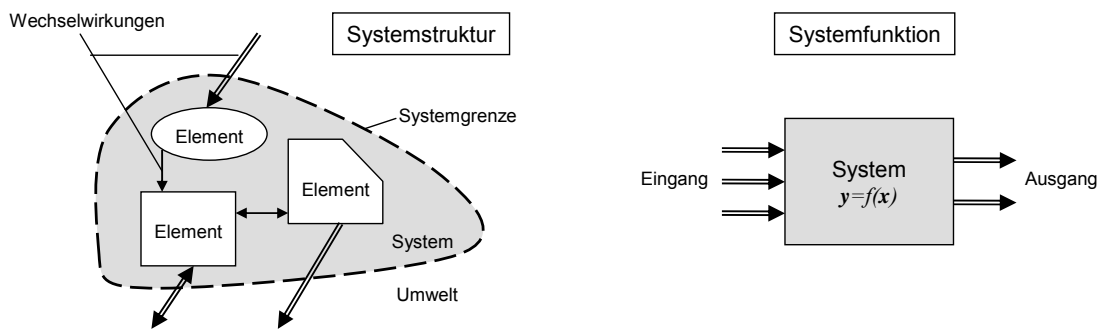
- eine Beschreibung der Interaktion des Systems mit seiner Umwelt durch Eingangs- und Ausgangsgrößen.

*ANMERKUNG: Eingangsgrößen beschreiben die Wirkung der Umwelt auf das System, Ausgangsgrößen seine Reaktion auf die Umwelt. Diese Größen müssen nicht mit den Einfluss- und Zielgrößen des Versuchs übereinstimmen, sondern beschreiben die Beziehung des Systems mit seiner Umwelt. Einfluss- und Zielgrößen beschreiben hingegen den Versuch im Rahmen eines Prozessmodells.*

- eine Beschreibung des funktionalen Zusammenhangs zwischen Eingangs- und Ausgangsgrößen durch eine Transferfunktion.

*ANMERKUNG: Die Beschreibung kann auf der höchsten (Super-) Systemebene als auch auf einer beliebigen (Sub-)Systemebene erfolgen.*

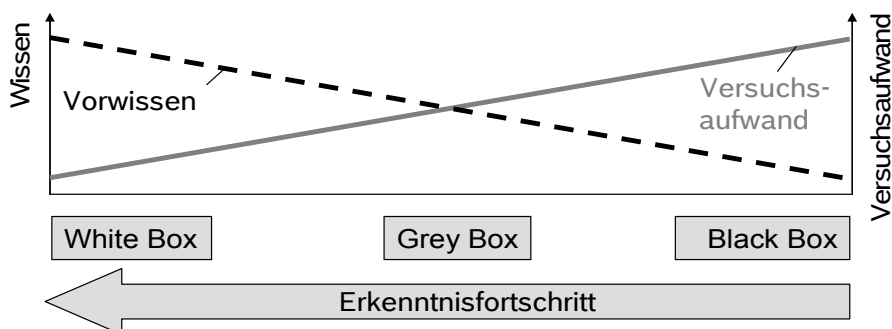
Die Darstellung der Struktur und Funktion eines Systems kann als Graph erfolgen.



**Bild 1.1:** Systemstruktur und -funktion

### 1.1.1.2. Arten der Modellbildung

Man unterscheidet zwei verschiedene Wege der Modellbildung sowie Mischformen davon [13], S. 52ff.:



**Bild 1.2:** Arten der Modellbildung



### Physikalische Modellbildung (White Box)

Die physikalische Modellbildung (auch white- oder glass-box Modellbildung genannt) bezeichnet eine deduktive Vorgehensweise (vom Allgemeinen zum Speziellen) auf der Basis von allgemeinen Naturgesetzen und Prinzipien (z.B. Erhaltungssätze, Gesetze aus Chemie, Physik oder Thermodynamik) sowie Strukturwissen über das System. Die Funktion der Systemelemente muss hier auf der gewählten Betrachtungsebene vollständig bekannt sein, so dass aus der Systemstruktur die Funktion des Gesamtsystems eindeutig abgeleitet werden kann. Modellparameter müssen durch physikalische Konstanten und die *bekannt*en Eigenschaften der Systemelemente vorgegeben sein. Das Modell wird oft als Satz komplexer partieller Differentialgleichungen (PDGL's) dargestellt.

Der Modellansatz ist hier nur durch die Gültigkeit der zugrundeliegenden Gesetzmäßigkeiten eingeschränkt, so dass ein Einsatz auf breiter Basis möglich ist, was einen großen Vorteil dieser Vorgehensweise darstellt. Man spricht von einem strukturtreuen Modell, welches das Systemverhalten *erklärt*.

Auch bei einer physikalischen Modellbildung muss das Modell experimentell validiert werden.

*BEISPIEL: Als Beispiel einer physikalischen Modellbildung seien strukturmekanische Modelle von Bauteilen erwähnt, deren Deformationsverhalten unter Belastung mittels der Methode der Finiten Elemente untersucht wird. Das zu modellierende Bauteil wird hier als System „finiter Elemente“ betrachtet. Die Systemstruktur ist durch die Aufteilung des Bauteilvolumens in z.B. Tetraeder-Elemente und deren Anordnung vorgegeben. Das Verhalten (Funktion) der einzelnen Elemente wird durch ein System partieller Differentialgleichungen beschrieben, die Interaktion zwischen den Elementen bilden globale Gleichgewichtsbedingungen. Aufgrund der bekannten Gesetzmäßigkeiten der Elastizitätstheorie und des Systemwissens (z.B. über die Geometrie und Beschaffenheit des Systems) ist die Ableitung des globalen Systemverhaltens möglich. Die Eingangsgrößen bilden die Lasten und sonstige Randbedingungen. Das System reagiert darauf mit einer bestimmten Deformation (Ausgangsgrößen). Die Modellparameter sind durch E-Modul und Querkontraktionszahl des Werkstoffs vorgegeben, diese müssen bekannt sein. Das Modell ist durch die Gültigkeit der zugrundeliegenden elastizitätstheoretischen Gesetze begrenzt, bei plastischer Deformation unter großen Lasten ist diese nicht mehr gegeben. Deswegen ist eine experimentelle Validierung erforderlich, z.B. durch Messung der tatsächlichen Deformation mit Hilfe eines Dehnungsmessstreifens und dem anschließenden Vergleich mit dem rechnerischen Wert.*

White-box-Modellbildung ist in Reinform selten anzutreffen. Die Anwendung physikalischer Modellbildung ist bei technischen Fragestellungen möglich, falls der Stand der Technik dafür ausreicht sowie ein umfassendes Systemwissen vorhanden ist. Dies trifft für quantitative Aussagen oft nicht zu, da ausreichendes Wissen über Struktur, Gesetze oder Elementeigenschaften entweder in der erforderlichen Detailtiefe nicht verfügbar oder aus Zeit- und Kostengründen nicht beschaffbar ist. Streuende Größen erschweren die Modellbildung zusätzlich. Mögliche Abhilfe verschafft hier die grey-box-Modellierung. Qualitative Analysen der Abhängigkeiten und Wechselwirkungen durch Parameterstudien sind jedoch möglich.

### Empirische Modellbildung (Black Box)

Die empirische Modellbildung, auch Systemidentifikation oder black-box- Modellbildung genannt, ist eine induktive Vorgehensweise (vom Speziellen zum Allgemeinen), bei der nur durch eine experimentelle Erfassung der Eingangs- und Ausgangsgrößen des Systems in endlich vielen, diskreten Fällen (bzw. in einer endlichen Zeit), z.B. anhand von Messung und Beobachtung, ein Modell aufgestellt wird, welches die Beziehung zwischen Eingangs- und Ausgangsgrößen durch eine Transferfunktion bestmöglich beschreibt. Man spricht in diesem Zusammenhang auch von einem *verhaltensbeschreibenden* Ansatz. Basis für die Transferfunktion ist ein allgemeiner mathematischer Ansatz, welcher a-priori vorgegeben werden muss, meistens als einfache algebraische Gleichung, z.B. als Polynom ersten oder zweiten Grades. Anschließend Systemidentifikation erfolgt durch Bestimmung der freien Modellparameter (Koeffizienten in der Gleichung) so, dass die Transferfunktion im Rahmen der Modellgenauigkeit das Systemverhalten bestmöglich beschreibt.

#### Vorteil:

- Strukturwissen sowie Wissen über physikalische Gesetze und Elementeigenschaften des Systems müssen nicht vorhanden sein (Wissen ist tatsächlich nicht vorhanden oder nicht beschaffbar).

#### Nachteile:

- Chancen hinsichtlich eines tieferen Verständnisses des Systemverhaltens werden eventuell nicht genutzt.





- das Modell kann nur in engen Grenzen eingesetzt werden, da ein komplexes globales Systemverhalten nicht durch die begrenzte Ordnung der Modellgleichung angenähert werden kann. Modellierung ist somit nur zur Interpolation geeignet.

*ANMERKUNG: In diesen engen Grenzen kann man jedoch durch Taylor-Entwicklung zeigen, dass das Modell jedes stetige Systemverhalten bis auf einen Restfehler höherer Ordnung wiedergeben kann, auch wenn es von den (unbekannten) tatsächlichen physikalischen Zusammenhängen abweicht. Dieser Sachverhalt beantwortet auch die Frage, warum überhaupt ein allgemeiner Modellansatz in der Lage sein soll, unterschiedliches Systemverhalten zu beschreiben, und stellt einen wesentlichen Vorteil dieses Ansatzes dar.*

Die Verifikation eines black-box-Modells ist bereits in der Modellbildung enthalten, die Validierung (ob der a-priori gewählte Modellansatz für die Fragestellung sinnvoll ist) muss separat geführt werden.

Die black-box-Modellbildung dient bei vollkommen unbekanntem Systemen als Basis für eine physikalische Modellbildung, indem Fakten und Informationen gesammelt werden, woraus durch Hypothesenbildung die physikalischen Theorien abgeleitet werden können. Diese Art der Modellbildung ist auch bei technischen Fragestellungen ohne Vorwissen in Anwendung und Gegenstand der klassischen statistischen Versuchsplanung, z.B. mittels faktorieller Prüfpläne.

*BEISPIEL: Ein Beispiel einer empirischen Modellbildung ist das Korrosionsverhalten nichtrostender Stähle im ethanolhaltigen Benzin, in Abhängigkeit von der Material- und Mediumzusammensetzung, der Oberfläche, Temperatur und Expositionszeit. Nähere Zusammenhänge sind hier nicht bekannt, so dass zu einer black-box-Modellbildung gegriffen werden muss. In einer Versuchsreihe wird die Abhängigkeit der Intensität der Korrosionserscheinungen, gemessen an den Abmessungen der Korrosionslöcher (Systemausgang), von der Expositionsdauer und allen weiteren oben erwähnten Parametern (Systemeingang) untersucht.*

### Gemischte Modellbildung (Grey Box)

Neben den ersten beiden Modellierungsansätzen existieren diverse Mischformen, diese werden als grey-box-Modellierungen bezeichnet.

Falls die Systemstruktur bekannt ist, nicht aber die Funktion der Systemelemente, kann auf Ebene der Elemente die Technik der empirischen Modellbildung angewendet werden, um das Systemverhalten zu beschreiben. Man spricht in diesem Zusammenhang von verhaltensbeschreibenden Elementen in einem verhaltensklärenden Modell, [13], S.57.

Andererseits können Strukturwissen und bekannte physikalische Zusammenhänge, soweit vorhanden, bei der Aufstellung des Modellansatzes verwendet werden. Ein Polynom kann z.B. durch eine Exponential- oder Logarithmus-Funktion aus einem physikalischen Modell ersetzt werden. Bei der klassischen Versuchsplanung mittels faktorieller Prüfpläne (die auf Polynomansätzen basieren) kann die Verwendung von solchen Ansätzen durch Variablentransformation, erfolgen. Dies würde der Berücksichtigung verhaltensklärender Ansätze in einem beschreibenden Modell entsprechen.

Eine weitere Art von grey-box-Modellierungen ergibt sich aus der Tatsache, dass eine physikalische Modellbildung meistens auf gekoppelte Systeme partieller Differentialgleichungen führt. Die Eigenschaften der Abhängigkeit zwischen Einfluss- und Zielgrößen sind somit nicht direkt ersichtlich, Lösungen ergeben sich hier erst durch komplexe numerische Berechnungsverfahren wie die Methode der Finiten Elemente. Damit liegt die Modellgleichung aber nicht als explizite analytische Funktion sondern implizit als „Rechenvorschrift“ vor, d.h. ein computerbasiertes Berechnungstool (Solver) kann für diskrete Einstellungen der Einflussgrößen die zugehörige Zielgröße auf Basis des komplexen physikalischen Modells liefern. Dieser Vorgang wird im Kontext der Versuchsmethodik analog zu den physikalischen Versuchen als Computereperiment bezeichnet. Hier werden nicht direkte Untersuchungen an einem realen System durchgeführt, sondern Simulationen an einem bereits existierenden Modell. Die dabei entstandenen Modelle werden dementsprechend als Meta-Modelle bezeichnet, sie können so einfach wie nötig gestaltet werden (z.B. oft als Polynom ersten oder zweiten Grades), um somit direkt Aufschluss über die Art der Abhängigkeit zwischen Einfluss- und Zielgrößen liefern zu können.

Bei der Bestimmung der freien Parameter eines physikalisch-basierten Modells spricht man von einer experimentellen Parameteridentifikation oder gelegentlich davon, dass das Modell „kalibriert“ wird. Dies kann bei komplexen Modellen nicht trivial sein, da keine explizite Darstellung der Modellgleichung mehr möglich ist, was zu sogenannten inversen Problemen führt. Da es nicht mehr möglich ist, den Versuch so zu planen, dass die freien Parameter der Modellgleichung unmittelbar aus den Versuchsergebnissen bestimmt werden können, ist eine Lösung hier nur durch Optimierung möglich. Bei diesem Modellierungsansatz ist die Parameteridentifikation gleichzeitig die Verifikation des



Modells. Eine Validierung (ob der Modellansatz für die Fragestellung geeignet ist) muss allerdings separat erfolgen.

Dieser Ansatz entspricht weitestgehend dem typischen ingenieurmäßigen Vorgehen bei komplexen technischen Fragestellungen, bei dem ein grobes physikalisches Modell durch Versuche über vorhandenes Systemwissen angepasst wird. Das Modell kann in begrenztem Umfang auch zur Extrapolation verwendet werden.

*BEISPIEL: Beispiele für grey-box-Modelle sind die nicht-linearen Materialmodelle der Strukturmechanik, welche auf Basis rheologischer Gesetzmäßigkeiten konstruiert wurden und mehrere Dutzend freie Modellparameter besitzen. Diese werden bestimmt, indem durch Optimierung eine Anpassung des Modells an komplexe Versuchsergebnisse vorgenommen wird.*

### Handlungsleitlinie:

Es muss so weit wie möglich eine white-box-Modellbildung angestrebt werden. Hier sind Versuche lediglich zur Verifikation und Validierung erforderlich.

Jedem muss jedoch bewusst sein, dass

- in den seltensten Fällen eine reine white-box-Modellbildung möglich ist,
- in vielen Fällen eine mehr oder weniger davon abweichende grey-box-Modellbildung notwendig ist,
- und manchmal sogar eine reine black-box-Modellbildung je nach Aufgabenstellungen erforderlich sein kann.

Bei den letzten beiden Fällen ist das theoretische Wissen nicht ausreichend verfügbar (z.B. bei neuen Produkten und Technologien) oder kann nicht rechtzeitig beschafft werden, somit sind Versuche außer zur Validierung auch zur Parameter- oder zur Systemidentifikation erforderlich.

### 1.1.2. Versuche und Erprobung

Versuche sind auch zur Validierung der Qualität von Produkten notwendig. Praktisch alle relevanten Normen, Regelwerken, Richtlinien etc. z.B. [6,7,8,9] fordern, dass während der Entwicklung Produkte *erprobt* werden.

Unter Erprobung wird dabei die Prüfung an einem technischen Erzeugnis zur Verifizierung und Validierung von Qualitäts-Merkmalen verstanden.

Nach DIN EN ISO 9000 ist die Verifizierung die Bestätigung durch Bereitstellung eines objektiven Nachweises, dass festgelegte Forderungen erfüllt worden sind.

Die Validierung ist nach obiger Norm hingegen die Bestätigung durch Bereitstellung eines objektiven Nachweises, dass die Forderungen für einen spezifischen beabsichtigten Gebrauch oder Anwendung erfüllt worden sind.

Mit Erprobung soll nachgewiesen werden, dass die gestellten Anforderungen, d.h. die Funktionalität, Zuverlässigkeit, Robustheit und Sicherheit des Erzeugnisses erfüllt werden. Darüber hinaus soll die Erprobung Erkenntnisse über diverse Einflüsse (z.B. Umwelteinflüsse) auf Produkte und Prozesse bringen und insbesondere die Aussagesicherheit erhöhen.

Folgende **Prinzipien** stellen die Grundlagen der Erprobung dar:

- **Frontloading:**  
bedeutet, dass so früh wie möglich mit der Erprobung begonnen werden muss.
- **Feldrelevanz:**  
bedeutet, dass die Prüflinge selbst sowie die Bedingungen der Erprobung so weit wie möglich den späteren Einsatzbedingungen entsprechen müssen.
- **Standardisierung:**  
Wichtiges Grundprinzip ist die Arbeit mit Standard-Prüfverfahren, welche in Regelwerken beschrieben werden, z.B. DIN IEC 68 „Grundlegende Umweltprüfverfahren“ oder VDI 4005 „Einflüsse von Umweltbedingungen auf die Zuverlässigkeit technischer Erzeugnisse“.

Zwei **Erprobungsphilosophien** mit entsprechenden Vor- und Nachteilen stehen obigen Prinzipien gegenüber:



1. **“Test to pass”**: Die geforderte Qualität bzw. Zuverlässigkeit des Produktes wird durch Bestehen einer vorgegebenen Erprobung ohne Beeinträchtigung des Erzeugnisses nachgewiesen. Oft basieren solche Anforderungen auf übergeordneten Regelwerken, so dass das Erzeugnis als **„fit-for-standard“** angesehen werden kann.

Vorteile:

- breite Akzeptanz der Ergebnisse
- gewisse Sicherheit, da man sich stets auf das „Bestehen“ einer vorgeschriebenen Erprobung berufen kann.

Nachteile:

- Feldrelevanz oft unklar
- Mögliche Feldprobleme trotz bestandener Erprobung.

2. **“Test to failure”**: Durch eine (meist) verschärfte Erprobung werden die Grenzen des Erzeugnisses in einem end-of-life-Test (deutliche Beeinträchtigung bis zum Ausfall) ermittelt. Die Erprobung ist hier eine wichtige Entwicklungstätigkeit, durch welche Risiken erkannt und Verbesserungen herbeigeführt werden sollen. Die „Bewertung“ des Erzeugnisses erfolgt nicht direkt aus der Erprobung, sondern in einem separaten Arbeitsschritt, bei dem die Ergebnisse der Erprobung und die feldrelevanten Einsatzbedingungen einfließen. Da auf Feldrelevanz bei dieser Betrachtung großen Wert gelegt wird, kann nach positiver Bewertung das Erzeugnis hier als **„fit-for-use“** eingestuft werden.

Vorteil:

Bei richtiger Anwendung kann es hier nicht zu Feldproblemen kommen, da die Grenzen des Erzeugnisses und die feldrelevanten Einsatzbedingungen bekannt, und der Abstand zwischen Beiden als ausreichend eingestuft wurde.

Schwierigkeit:

Ermittlung der feldrelevanten Einsatzbedingungen.

### 1.1.3. Versuche und Qualität

Die Herausforderung heutiger Unternehmen ist, qualitativ hochwertige Produkte (Null-Fehler-Philosophie) zunehmender Komplexität in kürzester Zeit zu konkurrenzfähigen Preisen auf den Markt zu bringen.

**Zwei generelle Ansätze** können als Antwort auf diese Herausforderung angesehen werden, vgl. [21], S.5:

- **Six-Sigma** als eigenständige **Verbesserungsmethodik** zur Reduzierung bzw. Eliminierung bestehender Fehler/Abweichungen auf der Basis von bereits bekannten Qualitäts- und Lean-Management-Ansätzen (reaktiver Ansatz)
- **Design-For-Six-Sigma (DFSS)** als **Entwicklungsmethodik**, um durch ein fundiertes Produktverständnis und unter Einsatz adäquater Entwicklungsmethoden von Anfang an fehlerfreie, robuste und zuverlässige Produkte zu entwickeln (präventiver Ansatz) und dabei die eingesetzten Ressourcen möglichst effizient zu nutzen, [21] S. 14.

**Versuche** spielen sowohl in der klassischen Six-Sigma Qualitätsphilosophie als auch bei DFSS eine wesentliche Rolle:

- **bei Six-Sigma** hauptsächlich in der Analyse-Phase, wo es darum geht, die Wirkzusammenhänge zwischen Problemen und Ursachen zu finden, zu analysieren und statistisch abzusichern,
- **bei DFSS** als Werkzeug, um unbekannte Wirkzusammenhänge zu ermitteln und als Basis für die Gestaltung, Bewertung und Optimierung von Robustheit und Zuverlässigkeit zu verwenden.



## 1.2. Die Rolle der Statistik beim Versuch

Statistische Ansätze sind dann erforderlich, wenn bei der Untersuchung technischer Systeme streuende Größen auftreten. Je nach Ursache der Streuung können unterschiedliche Strategien angewendet werden, um sinnvoll mit dieser Streuung umzugehen.

### 1.2.1. Die Quellen von Streuung

Allgemein können bei einem Versuch drei verschiedene Quellen von Streuungen identifiziert werden:

- die untersuchten Prüflinge selbst
- die Messung der untersuchten Einfluss- und Zielgrößen
- die Versuchsanordnung.

Streuungen treten bei technischen Anwendungen auf weil die **Werkstoffeigenschaften**, die **Fertigungsbedingungen**, sowie die **Nutzungsbedingungen** von Erzeugnissen bei Stückzahlen > 1 nicht absolut identisch sein können.

Der Wert einer Messung unterliegt ebenfalls einer Streuung. Die Abweichung des (korrigierten) Messwertes vom prinzipiell unbekanntem wahren Wert wird durch die **Messunsicherheit** beschrieben (s. [15] Heft 8).

Da auch unter Laborbedingungen nicht sichergestellt werden kann, dass nur die untersuchten Einflüsse auf das Versuchsobjekt wirken, sondern auch andere **unbekannte Störgrößen** auftreten können, müssen diese Streuungen mit in Betracht gezogen werden (Einfluss der Versuchsanordnung selbst inkl. Umgebungsbedingungen).

#### 1.2.1.1. Streuende Prüflinge

Bei streuenden Prüflingen können zwei unterschiedliche Strategien angewendet werden:

- Streuung reduzieren:** Zunächst ist eine Design-Lösung anzustreben, bei der die betrachteten Effekte bedeutend größer werden als die Streuung, z.B. durch gezielte Streuungsreduktion. Solche Lösungen sind immer zu bevorzugen, denn sie zeichnen sich durch eine hohe Robustheit aus. Die Streuung tritt dann nach wie vor auf, spielt jedoch für die Anwendung eine untergeordnete Rolle. Beim Versuch können in diesem Fall auch ohne Anwendung statistischer Verfahren gute Ergebnisse erzielt werden. Für quantitative Angaben über die Güte der Ergebnisse (z.B. Vertrauensbereiche und Signifikanz) sind allerdings wieder statistische Ansätze erforderlich.
- Streuung beherrschen:** Falls man sich an der technologischen Grenze bewegt, so dass eine weitere Reduktion der Streuung nicht möglich oder nicht wirtschaftlich ist und diese vergleichbar mit den betrachteten Effekten wird, muss bereits in der Designphase sichergestellt werden, dass die Streuung beherrschbar (d.h. bekannt, stabil und zielführend) ist. In diesem Fall ist mit einem erhöhten Aufwand und einem steigenden Risiko für das Produkt-Konzept zu rechnen. Hier kann auf die Anwendung statistischer Verfahren bei der Versuchsführung nicht verzichtet werden.

Grundsätzlich ist immer die erste Strategie anzustreben. Denn selbst wenn durch statistische Ansätze ein Effekt im Mittel nachgewiesen wird, bleibt die Frage offen, ob dieser Effekt für die spätere Anwendung ausreichend, und die auftretende Streuung zielführend ist.

*BEISPIEL 1: In einer Einspritzkomponente soll ein Bauteil eingesetzt werden, welches mittels eines wirtschaftlichen, bisher jedoch nicht verwendeten formgebenden Fertigungsverfahrens hergestellt werden soll. Das Design des Bauteils ist vollständig auf dieses Fertigungsverfahren ausgerichtet.*

*Am Anfang der Entwicklung liegen keine weiteren Informationen über die Streuung der Materialeigenschaften dieses Bauteils vor. Bald stellt sich jedoch heraus, dass diese einer erheblichen Streuung unterliegen, bedingt durch die hohe Sensitivität des Fertigungsverfahrens gegenüber diversen Einflussgrößen.*

*Da kein anderes Fertigungsverfahren bei dem vorliegenden Design wirtschaftlich erscheint, wurden erhebliche Anstrengungen unternommen, die Fertigungsstreuung zu reduzieren. An dieser Stelle wäre der Ansatz „mit der Streuung zu leben“ z.B. durch statistisch gut abgesicherte Versuche zur Bestimmung der Materialeigenschaften höchst problematisch gewesen, denn die hohe Streuung war weder stabil noch zielführend im Hinblick auf eine spätere Serienproduktion. Ziel muss sein, vor der Festlegung des Designs nach einem stabilen Fertigungsverfahren zu suchen und bei der Betrachtung der Kosten das Kostenrisiko der weniger robusten Lösung mit einzubeziehen.*



*BEISPIEL 2: Die Schwingfestigkeit von höchstfesten Stählen für den Einsatz bei Hochdruckeinspritzkomponenten ist u. a. durch Brüche an nichtmetallischen Einschlüssen begrenzt.*

*Die Reduzierung der Streuung der Größe von Einschlüssen und damit der Streuung der Beanspruchbarkeit hat einen hohen Stand erreicht und ist derzeit technisch nicht weiter möglich. Somit muss ein Konzept zur Bewertung des Einflusses der Streuung der Beanspruchbarkeit auf die Zuverlässigkeit entwickelt werden. Dieses Konzept beinhaltet Versuche mit hoher statistischer Absicherung zur Bestimmung der Beanspruchbarkeit.*

*Damit wird die hohe Streuung beherrschbar und die Einhaltung der hohen Anforderungen bezüglich Zuverlässigkeit solcher Komponenten kann sichergestellt werden.*

Im Falle streuender Prüflinge, bei denen die Streuung nicht weiter reduzierbar, jedoch beherrschbar ist, können objektive Aussagen bei einem Versuch erst durch die Anwendung statistischer Ansätze ermöglicht werden. Dies gilt besonders dann, wenn die Streuung nicht wesentlich kleiner als der untersuchte Effekt ist. Was „wesentlich kleiner“ heißt, und insbesondere, ob der erzielte Effekt sowie die Streuung für die technische Lösung ausreichend ist, kann nur im Kontext der Anwendung entschieden werden. Nach [1], S.3 wird als Richtwert empfohlen: wenn die untersuchten Effekte kleiner als das Fünffache der Standardabweichung sind, *oder* der Effekt quantitativ ermittelt werden soll, *müssen* statistische Ansätze angewendet werden.

In allen Fällen sind die Vertrauensbereiche durch Anwendung statistischer Ansätze zu ermitteln, um die Unsicherheit der Aussage abschätzen zu können.

### 1.2.1.2. Streuende Messergebnisse, Störgrößen

Bzgl. streuender Messergebnisse und Störgrößen gilt beim Versuch die Anforderung der Objektivität: Versuchsanordnung und Messung dürfen keinen wesentlichen Einfluss auf das Versuchsergebnis haben. Dabei können zwei Fehlerarten auftreten, vgl. [2], S.21:

- **Systematische Fehler:** Abweichung der ermittelten Mittelwerte vom tatsächlichen Mittelwert. Diese beeinträchtigen die **Genauigkeit** des ermittelten Wertes. Ursachen: dejustierte Messgeräte, nicht richtig kalibrierte Versuchsanordnungen, systematische Einflüsse der Vorrichtung oder des Experimentators etc. Aussagen über die Genauigkeit können durch Vergleiche mit einer anderen (Referenz-) Messmethode oder einer Standardprobe gemacht werden. Systematische Fehler können durch eine sorgfältige Versuchsplanung und Durchführung entsprechend Kapitel 1.5 vermieden werden.
- **Zufällige Fehler:** Streuung der Einzelmessungen vom Mess-Mittelwert. Diese beeinträchtigen die **Präzision** des ermittelten Wertes. Ursachen: ungeeignete Messverfahren, Einwirkung von Störgrößen, zufällige Ablese- und Rechenfehler etc. Aussagen über die Präzision liefert die reine Versuchsstreuung.

Der Nachweis über die Erfüllung obiger Anforderung ist durch Kalibrierung der Versuchsanordnung, Messfähigkeitsanalysen etc. zu führen, Details hierzu sind in [15,16] zu finden.

Eine statistische Versuchsführung ist keine Abhilfe für unzureichende Messgenauigkeit. Insbesondere systematische Fehler müssen vermieden werden. Durch Anwendung statistischer Ansätze können jedoch zufällige Störgrößen kontrolliert und die Aussagegenauigkeiten erhöht, sowie Aussagen über die Signifikanz der Ergebnisse getroffen werden.

Statistik darf auch nicht als Ersatz für fehlende Sorgfalt bei der Versuchsdurchführung dienen. Insofern ist z.B. eine bekannte Größe, welche mutmaßlich einen Einfluss auf das Ergebnis haben könnte (z.B. Werkzeugverschleiß) immer als Versuchsfaktor zu erfassen (z.B. durch Blockbildung) und nicht als Störgröße zu behandeln.

### 1.2.2. Prinzipien der statistischen Versuchsplanung

Begriffsdefinition:

Unter Versuchsplanung wird die Aufstellung eines Versuchsplans (inkl. der Planung von Durchführung und Auswertung) im Rahmen der Versuchsmethodik verstanden.

Mit statistischer Versuchsplanung bezeichnet man eine besondere Versuchsplanung, welche unter Einsatz von Mitteln der Statistik auch bei streuenden Versuchsergebnissen eine rationale Aussage über die Fragestellung erlaubt.



Folgende Prinzipien der statistischen Versuchsführung erhöhen die Aussagesicherheit bei streuenden Größen [2]:

- **Wiederholung** zur Verbesserung der Genauigkeit (Mittelwertstreuung ist kleiner als die Einzelwertstreuung) und Aussagen über die Signifikanz der Ergebnisse.
- **Abdeckung des Versuchsraums**, z.B. durch eine symmetrische Anordnung der Versuchspunkte, für eine höhere Effektivität und zur Erfassung von Wechselwirkungen. Die gezielte Variation von mehreren Faktoren, welche die Anzahl der Versuche bei gleicher Aussagesicherheit minimiert, ist ein wichtiger Ansatz der statistischen Versuchsplanung.
- **Randomisierung** (Zufallsreihenfolge) zur Ausschaltung von unbekanntem Störfaktoren. Falls ein unbekannter Trend, z.B. in der Beschaffenheit der Prüflinge aufgrund von Werkzeugverschleiß vorliegt, kann dieser dadurch ausgeschaltet werden, dass Prüflinge aus der Fertigung zufällig für den Versuch entnommen werden.
- **Blockbildung** zur Erfassung von nicht einstellbaren Faktoren. Beispielsweise müssen Versuche, die nicht alle zeitnah realisiert werden können, in Blöcke aufgeteilt werden, um eventuelle Störeinflüsse wie den Werkzeugverschleiß zu erfassen. Falls Unterschiede zwischen den Blöcken auftreten, können diese erkannt und eliminiert werden.
- Vermengung von Einflüssen zur begrenzten Reduktion des Aufwandes. Diese Technik ist allerdings mit Nachteilen verbunden, siehe dazu auch Kapitel 8.1.
- Sequentielles Vorgehen, so dass entweder Fortführung oder aber Abbruch der Aktivitäten möglich ist, wenn ausreichende Ergebnisse vorliegen.

Die Vorgehensweise in Kapitel 4 berücksichtigt die meisten dieser Prinzipien. Weiterführende Ansätze finden sich in Kapitel 8, Grundlagen siehe [11].

### 1.3. Chancen und Risiken der Versuchsmethodik

#### 1.3.1. Vorteile, Erfolgsfaktoren und Stärken

Vorteile:

- **Objektives Vorgehen:** DoE ermöglicht objektive und präzisierte Aussagen, inkl. einer quantitativen Modellierung und Optimierung auf Basis der Versuchsdaten, sowie Aussagen über deren Signifikanz, vgl. [2], S.40.
- **Effektives Vorgehen:** Anwender berichten von einer Verkürzung der Projektlaufzeiten und einer deutlichen Senkung der Versuchskosten [1]. Der Gesamtaufwand kann frühzeitig abgeschätzt werden, statt unplanbarer Taktik „nur noch ein Versuch“. Der Stichprobenumfang wird der zu lösenden Aufgabe angepasst, was zu einer ausreichenden Genauigkeit des Modells führt. Sowohl eine unzureichende Genauigkeit (Verpassen technischer Chancen) als auch eine übertriebenen Genauigkeit (unwirtschaftliches Vorgehen) werden vermieden. Außerdem wird ein Maximum an Informationen aus den Untersuchungen erreicht z.B. über Wechselwirkungen und der Signifikanz der Ergebnisse.
- **Systematisches Vorgehen:** Anwender sind häufig damit überfordert, aus Massendaten Erkenntnisse zu ziehen, DoE bietet dafür ein klares und gut vermittelbares Vorgehen für Planung und Auswertung. Das Verdichten der Daten erleichtert die Darstellung und Kommunikation und ermöglicht die Auswertung auch zu einem späteren Zeitpunkt sowie die Einbeziehung der Daten in zukünftige Untersuchungen.

Erfolgsfaktoren:

- **Experten Know-how:** Der Einsatz von DoE benötigt ausreichendes Wissen und Erfahrung sowohl bezüglich der zu lösenden Aufgabe, als auch bezüglich der Methode selbst, einschließlich eines ausreichenden Wissens über Statistik und der eingesetzten Software. Da erfahrungsgemäß diese Anforderungen selten in einer Person vereint werden können, ist die Aufstellung eines Versuchsteams zu empfehlen.
- **Geeignete Software:** Falls anhand der Versuchsdaten Modellbildung mit mehreren Faktoren durchgeführt werden soll, sind sowohl das Aufstellen der Versuchspläne als auch die meisten Auswertungen ohne Softwarehilfe nicht zu bewerkstelligen. Über geeignete Softwaretools wird im Anhang berichtet.



- **Frühzeitiger Einstieg bei der Versuchsplanung:** Falls bereits Daten (oft in unklarer Qualität) vorhanden sind, ist eine nachträgliche Bewertung zwar möglich, aber schwierig, z.B. wegen unzureichender Abdeckung des Versuchsraumes. In diesem Fall steigen die Anforderungen an den Methoden-Experten, um die Aussagekraft der Ergebnisse einzuschätzen und Schlussfolgerungen zu ziehen. Die Vergleichbarkeit und Aktualität historischer Daten muss sichergestellt sein, bevor mit der Auswertung begonnen wird.
- **Ausreichende Ressourcen:** Obwohl DoE eine ressourcenoptimierte Vorgehensweise darstellt, können die Untersuchungen u.U. beträchtliche Ressourcen in Anspruch nehmen. Reichen die Ressourcen für eine DoE-Untersuchung nicht aus, führt wahrscheinlich auch kein anderer Versuchsansatz zum Ziel. Abhilfe kann hier nur eine Anpassung des Versuchsziels liefern.

Die Stärke von DoE liegt bei Aufgabenstellungen, bei denen die Abhängigkeit streuender Zielgrößen von mehreren Einflussgrößen untersucht werden muss. Hier kann mit einer überschaubaren Anzahl von Versuchen ein komplexes Systemverhalten quantitativ beschrieben werden, auch wenn eine physikalische Modellbildung schwierig ist.

### 1.3.2. Grenzen und Gefahren

Was kann von DoE erwartet werden und was nicht?

DoE bietet Lösungen für typische Fragestellungen u. a. (siehe auch 2.3):

- Modellierung: gezielter Erkenntnisgewinn über Wirkzusammenhänge im System (Produkt oder Prozess), falls physikalische Modellbildung schwierig
- Prognose und Optimierung von Produkt- und Prozesseigenschaften,
- Screening: Ermittlung der wesentlichen aus einer Menge von möglichen Einflussgrößen.

Die Methode braucht aber an entscheidenden Stellen **Input von Experten**, diese bestimmen entscheidend über die Qualität der Ergebnisse:

- **Definition der Faktoren und deren Variationsbereiche sowie der Zielgrößen:** Fehlen entscheidende Einflussgrößen, so kann sich dies lediglich durch eine erhöhte Streuung der Residuen bemerkbar machen, was nur dann auffällt, falls Erfahrungswerte bekannt sind. In der Regel gibt es keinen Ansatz, mit dem man entscheiden kann, ob ein wichtiger Faktor vergessen wurde. Falls die Residuenvarianz (RMS-Error, bzw. Bestimmtheitsmaß) nicht akzeptabel erscheint, müssen weitere Faktoren berücksichtigt werden.
- **Wahl des Modellansatzes:** Eine Validierung des Modellansatzes kann nur in begrenztem Umfang durch Überprüfung mit einem Ansatz höherer Ordnung erfolgen, somit ist die Wahl des Modellansatzes entscheidend.

*ANMERKUNG: Dieser Punkt darf allerdings auch nicht überbewertet werden: Es kann durch Taylor-Entwicklung gezeigt werden, dass jede stetige Funktion durch einen Polynom-Ansatz bis auf ein Restglied höherer Ordnung gut angenähert werden kann. Außerdem bieten alternative Modellansätze, welche nicht auf Polynome basieren, hier gewisse Abhilfe, siehe Kapitel 8.3.*

- **Sorgfalt bei der Versuchsplanung, Durchführung und Auswertung:** Dies entscheidet über die Qualität der Daten, welche für die Modellierung verwendet werden, vgl. Kapitel 1.5.
- **Plausibilisierung und physikalische Deutung der Ergebnisse:** Die Ergebnisse der Modellierung müssen durch Experten physikalisch interpretiert werden. Der Nutzen entsteht durch die physikalische Deutung, ansonsten bleiben die Versuche eine Ansammlung von Daten und Informationen.
- **Konsequenzen und Festlegung von weiteren Maßnahmen:** DoE bietet eine objektive Basis für Entscheidungen auf der Grundlage von Versuchsergebnissen, kann jedoch nicht die Entscheidungen selbst, sowie die Diskussion über die Folgen für die eigentliche Entwicklungsaufgabe ersetzen.

Grenzen:

- **Zu hoher Aufwand:** Auch unter dem Einsatz von DoE kann es dazu kommen, dass zu viele oder zu teure Versuche gebraucht werden, z.B.
  - unterschiedliche Ausführungen von Prüflingen erfordern teure Herstellverfahren
  - nicht ausreichende Anzahl von Prüflinge im Musterbau herstellbar,



- die Prüflinge sind in der Eigenschaftskombination, wie es das Versuchsschema vorsieht, nicht gezielt herstellbar,
- nicht ausreichende Zeit, um den Versuchsplan durchzuführen, etc.
- **Begrenzte Erkenntnis:** DoE ist ein effektiver Ansatz, um aus Versuchsergebnissen ein Maximum an Information herauszuholen. Die Möglichkeiten, den Versuchsaufwand im begrenzten Maße zu reduzieren können in bestimmten Fällen, z.B. bei Screening-Untersuchungen, durchaus sinnvoll sein, siehe dazu auch Kapitel 8. Diese sind jedoch immer auch mit einem Informationsverlust, bzw. mit erforderlichen a-priori Annahmen verbunden, welche die Gefahr von Fehlinterpretationen bergen, z.B. Abschwächung/Auslöschung bis hin zur Umkehr von Effekten bei teilfaktoriellen Plänen niedriger Auflösung. Zu entscheiden, wo die **Grenzen des Sinnvollen** liegen, ist immer ein wichtiger Beitrag des Methodenexperten. Beispielsweise werden für Untersuchungen der Wirkzusammenhänge (Response-Surface-Untersuchungen) Pläne mit mindestens der Auflösung V empfohlen, was für die meisten praxisrelevanten Fälle höchstens eine Halbierung des Aufwands bedeutet, siehe Kapitel 8.1. Insofern führt jeder Versuch, darüber hinaus Aufwand einzusparen zu deutlichen Erkenntniseinbußen.
- **Ansatzbedingte Schwächen:** DoE ist begrenzt geeignet für Untersuchungen, bei denen Wert auf die Extrapolierbarkeit der Ergebnisse gelegt wird. Solche Vorgaben können nur dann erfüllt werden, wenn mindestens eine grey-box-, besser eine white-box-Modellbildung vorliegt. Hier können physikalische Ansätze z.B. durch Variablentransformation berücksichtigt werden, siehe Kapitel 4.4. Gleiches gilt bei Systemen mit unstetigem oder periodischem Verhalten bzw. bei der Untersuchung eines transienten Zeitverhaltens (z.B. Einschwingverhalten), bedingt durch die Lokalität des Modellansatzes.

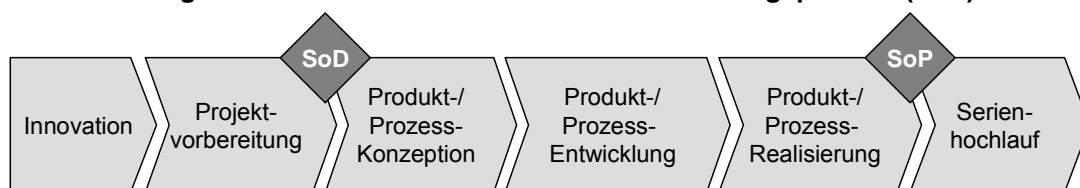
Gefahren:

- Erfahrene Versuchsexperten können dazu verleitet werden, auch dann eine black-box-Modellbildung vorzunehmen, wenn physikalische Zusammenhänge bekannt sind. Diese müssen, soweit wie möglich, immer bei der Modellbildung berücksichtigt werden, siehe Kapitel 1.1.
- Die experimentellen Untersuchungen können Lösungen „an der falschen Stelle“ liefern, obwohl zielführende Lösungen auf einer ganz anderen Ebene erforderlich wären, z.B. beim Design des Erzeugnisses, siehe dazu auch Kapitel 1.2.
- DoE kann zu einem falschen Sicherheitsgefühl verleiten. Bei streuenden Größen besteht immer ein kleines, aber endliches und genau definiertes Restrisiko, dass die getroffenen Aussagen falsch sind. Das Ausschalten dieses Restrisikos würde unendlich große Stichproben erfordern und kann somit nicht in Betracht gezogen werden.
- Anwender berichten von einer gewissen Unsicherheit, da der Weg zum Ergebnis insbesondere bei der Modellbildung von komplexen Systemen nicht immer offensichtlich ist. Die Ergebnisse werden manchmal als „mysteriös“ empfunden, obwohl sie nach einer klaren Systematik entstehen. Abhilfe bietet hier die Vermittlung eines ausreichenden methodischen Wissens.

**1.4. Versuchsmethodik bei Bosch**

Die Vorteile der Anwendung einer Versuchsmethodik wurden bei Bosch frühzeitig erkannt. Seit 1988 wurden technische Untersuchungen zunehmend mit Hilfe von DoE durchgeführt (QS-Information 01/1990). Dabei wurde festgestellt, dass „sich die in der Literatur genannten Vorteile (z.B. minimaler Aufwand, maximale Information) ohne Einschränkung bestätigen.“ Deswegen wurde bereits zu diesem Zeitpunkt die Empfehlung gegeben, sowohl im Entwicklungs- als auch im Fertigungsbereich verstärkt von den Vorteilen von DoE Gebrauch zu machen.

**1.4.1. Anwendung der Versuchsmethodik im Produktentstehungsprozess (PEP)**



**Bild 1.3:** Der Produktentstehungsprozess





DoE ist eine unterstützende Methode im Engineering und kann zu jedem Zeitpunkt des PEP's in jedem Geschäftsbereich und für jedes Produkt (und jedem Prozess) angewendet werden, falls es die jeweilige Fragestellung erfordert. Der Einsatz von DoE orientiert sich an der Fragestellung, welche bearbeitet werden soll.

### **Typische Einsatzbeispiele:**

#### **Phase Innovation:**

- Grundlagenuntersuchungen von neuartigen physikalischen Effekten, neuen Technologien oder Produktinnovationen durch black-box-Modellierung der Wirkzusammenhänge zwischen relevanten Einfluss- und Zielgrößen, typischerweise in der Forschung und Vorausbildung.

#### **Phase Projektvorbereitung:**

- Ermittlung von wesentlichen Einflussgrößen mit signifikanten Effekten auf die Zielgrößen (Screening).
- Modellbildung von unbekanntem Wirkzusammenhängen bei ausgewählten Entwicklungsschwerpunkten von alternativen Produktkonzepten.
- Untersuchung unterschiedlicher Werkstoffe in Bezug auf geeigneten Eigenschaften (magnetisch, elektrisch, Festigkeit, Korrosionsbeständigkeit etc.) im Rahmen einer Plattformentwicklung.

#### **Phase Produkt- und Prozesskonzeption:**

- Experimentelle Ermittlung komplexer funktionaler Zusammenhänge (Kennfelder) des favorisierten Designkonzepts
- Parameteridentifikation von physikalisch begründeten Modellen.
- Modellbasierte Prognose des Verhaltens und der Robustheit von Produkten mittels Computereperimenten.

#### **Phase Produkt- und Prozessentwicklung:**

- Prognosen des Verhaltens, Robustheitsanalysen und modellbasierte Optimierungen des Designs mittels Computereperimenten.
- Physische Versuche bei der Verifikation und Validierung von Produktanforderungen im Rahmen einer Erprobung (Erprobungsplan).

#### **Phase Produkt- und Prozessrealisierung:**

- systematische Untersuchung des Einflusses von Fertigungsparametern auf Produkteigenschaften mit anschließender versuchs- oder modellbasierter Optimierung im Hinblick auf Robustheit und Minimierung von Streuungen.

#### **Phase Serienhochlauf und Serienfertigung:**

- systematische Aufdeckung von Stör-Einflüssen in der Fertigung,
- versuchsbasierte Optimierung zur Ausnutzung von Kostenpotentialen.

### **1.4.2. Qualifizierung und Ansprechpartner**

Für den Einsatz von DoE ist ausreichendes Wissen und Erfahrung erforderlich, sowohl im jeweiligen Fachgebiet der zu bearbeitenden Aufgabe, als auch bezüglich der Versuchsmethodik selbst, einschließlich eines ausreichenden Hintergrundwissens über Statistik und der eingesetzten Software.

Qualifizierung bzgl. Versuchsmethodik:

1. Studium anwendungsorientierter Literatur z.B. [1] und das vorliegende Heft.
2. Seminare:
  - Allgemein: Seminar TQ009 „Versuchsmethodik“ (3 Tage). Voraussetzung: Kenntnis statistischer Grundlagen wie z.B. TQ002 „Grundlagen der Technischen Statistik“ und TQ003 „Auswerten von Messreihen“.
  - Aspekte der Erprobung: Seminar TQ023 „Erprobung neu entwickelter Erzeugnisse der Kfz-Ausrüstung“ (3 Tage).



- Spezielle Techniken der Versuchsführung: Seminare TQ007 „Grundlagen der Zuverlässigkeit“, TQ008 „Zuverlässigkeit von Steuergeräten“ und TF024 „Bauteil-Schwingfestigkeit“. Die Inhalte gehen allerdings deutlich über das Thema Versuchsmethodik hinaus.
3. Zur selbständigen Lösung komplexer Fragestellungen ist eine ausreichende Praxiserfahrung notwendig. Hier empfiehlt es sich die Mitwirkung bei DoE-Untersuchungen, die von einem erfahrenen Kollegen geleitet werden.
  4. Fragen, welche während der weiteren Tätigkeit aufgeworfen werden, sollten mit anderen Experten diskutiert werden. Generelle Ansprechpartner sind in erster Linie die Experten aus dem Kompetenznetzwerk DoE (aktuelle Liste im BGN unter <http://rb-kgw.bosch.com/>).
  5. Zu jedem Zeitpunkt ist das Studium der weiterführenden Literatur sinnvoll, welche die interessierenden Aspekte der Versuchsmethodik detaillierter behandelt, z.B. [4,5].

### 1.5. Elementare Regeln der Versuchsplanung und -durchführung

Die nachfolgenden elementaren Regeln sind Erkenntnisse aus jahrelanger praktischer Erfahrung und daher besonders zu beachten.

Der generelle Ablauf einer experimentellen Untersuchung besteht aus 5 wesentlichen Schritten (s. Bild 1.4):

1. Aufgabenanalyse
2. Systemanalyse
3. Versuchsplanung
4. Versuchsdurchführung und Dokumentation
5. Versuchsauswertung

**Genereller Grundsatz:** Versuche sollen nur zur quantitativen Ermittlung von Modellparametern und zur Validierung des Modells durchgeführt werden.

**Folgende Hinweise sind zu beachten** (vgl. Bild 1.4):

**Schritt 1** (Einzelheiten siehe Kapitel 2):

- Versuchsziel ist so festgelegt, dass der erwartete Wirkzusammenhang im Arbeitsbereich und darüber hinaus quantitativ ermittelt werden kann
- Die Erreichbarkeit / Machbarkeit des Versuchsziels ist zwischen Auftraggeber und Auftragnehmer (Versuchsteam) geklärt

**Schritt 2** (Einzelheiten siehe Kapitel 3):

- Einfluss- und Zielgrößen sind soweit bekannt, dass der zu ermittelnde Wirkzusammenhang daraus qualitativ oder quantitativ abgeleitet werden kann.
- Definition der für das Versuchsziel notwendigen Zielgrößen einschließlich Messprinzip (Messfähigkeitsnachweis), Messunsicherheit, Auflösung, Abtastrate und Abschaltkriterium:
  - möglichst direkte Messgröße während des Versuchs wählen
  - falls nicht möglich: nach Messgrößen mit bekannter Korrelation zur Zielgröße suchen
  - falls nicht möglich (bei Versuchen zu Alterung/Schädigung von Bauteilen):
    - a) Versuch mit definierten Entnahme-Zeitpunkten der Muster gestuft fahren
    - b) Zwischenentnahme zur Prüfung/Messung unter Berücksichtigung des Einflusses auf den Versuch (Achtung: Unterbrechung Wirkparameter!)
- Klärung und Erfassung zusätzlich wirkender Störgrößen z.B. durch Handling, Lagerung, Zwischenentnahme, Analyse und Wiederanfahen.

**Schritt 3** (Einzelheiten siehe Kapitel 4):

- Sicherstellung, dass ausschließlich die selektierten Faktoren (Einflussgrößen, die im Versuch gezielt variiert werden) wie vorgesehen wirken, alle anderen Einflussgrößen definiert konstant gehalten bzw. ausgeschlossen werden



- Definition einer geeigneten Versuchseinrichtung, Messverfahren und Messeinrichtung ausgerichtet auf die Erfordernisse des Wirkzusammenhangs
- Definition und Festlegung der Beschaffenheit der Prüflinge
  - gezielte Einstellung der Designparameter, Fertigungsverfahren entsprechend Versuchserfordernis und Dokumentation aller relevanten Parameter
  - Klärung Herstellbarkeit und Definition Analyseverfahren der Prüflinge
- Festlegung der Anzahl der Prüflinge
- Planung der Archivierung von Prüflingen und Rückstellungsmustern (Referenz)
- Versuchsdurchlauf je Prüfling planen inkl. einem Störfallplan: Ereignis, Maßnahmen (z. B. Erfassung, gezieltes Herunterfahren auf sicheren Zustand und Hochfahren)
- Prognose der erwarteten Versuchsergebnisse und Vergleich mit dem Versuchsziel (Wirkzusammenhang, Signifikanz)
- Planung der Auswertung und Dokumentation

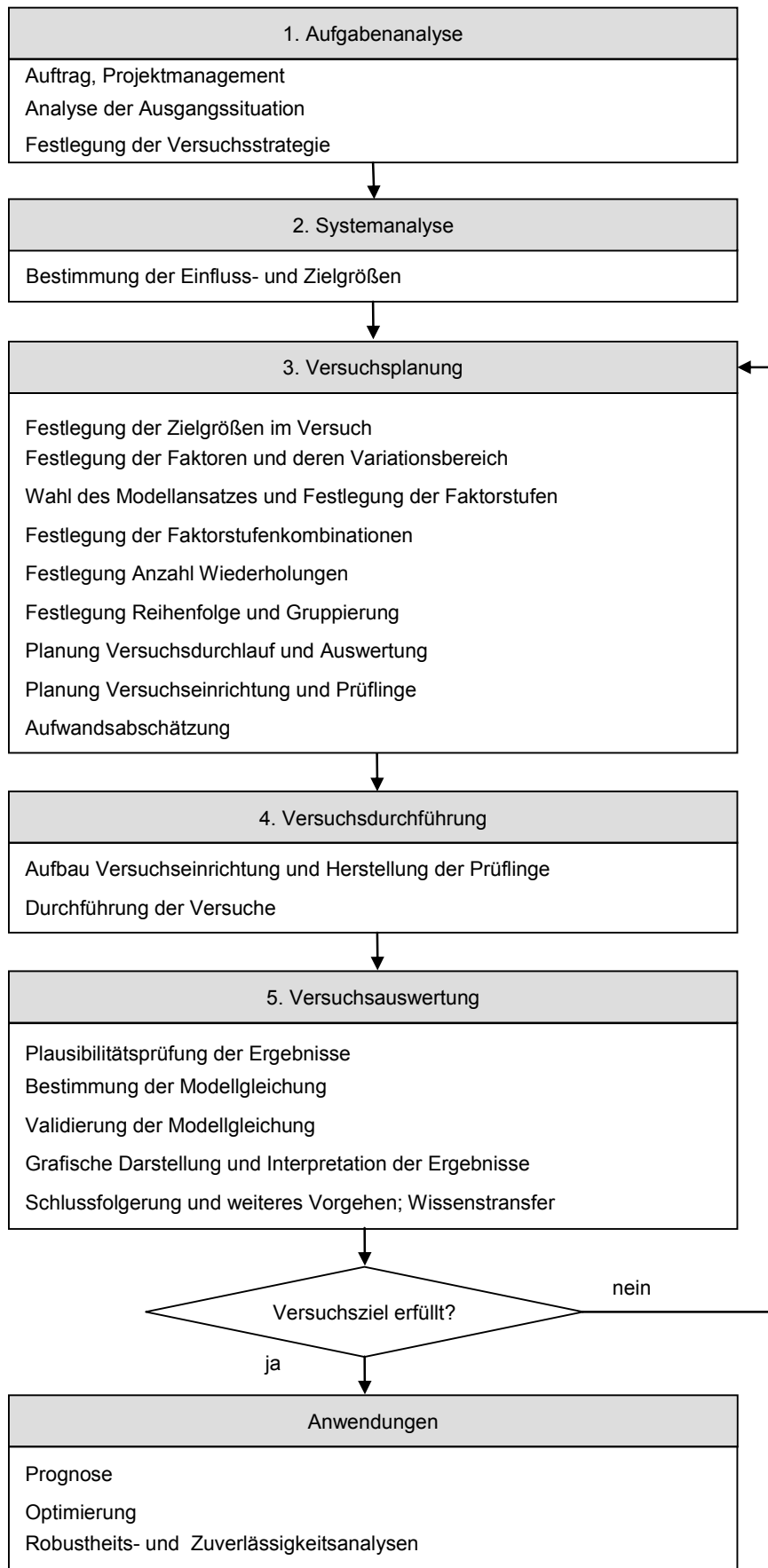
### **Schritt 4** (Einzelheiten siehe Kapitel 5):

- Aufbau und Überprüfung der Versuchseinrichtungen, Sicherstellung der Versuchsfähigkeit (Messgenauigkeit, Toleranzen, ...)
- Realisierung Messverfahren der festgelegten Zielgrößen
- Überprüfung Beschaffenheit und Kennzeichnung der Prüflinge
- Abfahren des Versuchsplans
  - Realisierung und Einhaltung der Vorgaben für die festgelegten Faktoren, Sicherstellung durch geeignete Messungen und Mitschriebe (auch Prüfmedien)
  - Sicherstellung, dass Faktoren am Wirkort gemäß Vorgabe anliegen
  - Erfassen und Protokollieren des Versuchsablaufs je Prüfling: Alle Größen entsprechend Versuchsplan einschl. Störfallplan
- Konsequenzen aus nicht vorgesehenen Störfällen ziehen, dokumentieren

### **Schritt 5** (Einzelheiten siehe Kapitel 6):

- Darstellung der Istwerte der Versuchsdurchführung
  - Versuchseinrichtung, verwendete Messtechnik
  - Beschaffenheit der Prüflinge
  - Durchgeführte Versuche
  - Besonderheiten im Versuchsablauf
- Darstellung der Versuchsergebnisse
- Auswertung der Daten
- Vergleich mit erwartetem Ergebnis, Bei Abweichungen Suche nach weiteren wirkenden Einflussgrößen und Bewertung der Konsequenzen
- Bewertung der Erreichung des Versuchsziels
- Zusammenfassung und Schlussfolgerungen
- Wissenstransfer





**Bild 1.4:** Ablauf einer experimentellen Untersuchung



## 2. Aufgabenanalyse

In diesem und den nachfolgenden Kapiteln werden die typischen Schritte der Versuchsmethodik chronologisch beschrieben.

### 2.1. Auftrag, Projektmanagement

Zunächst muss seitens des Auftraggebers ein Koordinator mit Verantwortlichkeit für Termine, Kosten und Qualität der Ergebnisse ernannt werden. Der Koordinator muss einen klaren Auftrag für die Untersuchung erhalten, in dem

- die zu lösende Aufgabenstellung sowie
- die dafür vorläufig in Aussicht gestellten Ressourcen wie Zeit, Geld, Maschinen- und Personalkapazität benannt sind.

Hierbei handelt es sich um erste grobe Schätzungen. Nur in den seltensten Fällen wird sich vermeiden lassen, dass Anpassungen sowohl bei der Aufgabenstellung, als auch bei den Lösungsansätzen vorgenommen werden.

In der Regel wird auch die Installation eines Versuchsteams aus Fachexperten, Methodenspezialisten und Versuchsfachleuten erforderlich werden, da für den Einsatz von DoE ausreichendes interdisziplinäres Wissen und Erfahrung erforderlich ist, sowohl im jeweiligen Fachgebiet als auch bezüglich der Versuchsmethodik selbst, einschließlich Statistik und der eingesetzten Software. Die Zuständigkeiten im Team müssen klar und eindeutig geregelt werden.

Das Versuchsteam kann je nach Aufgabenstellung auch später einberufen werden. In jedem Fall muss jedoch frühzeitig geklärt werden, welche Stellen von der Untersuchung betroffen sind. Deren Vertreter müssen bei der Lösung mit einbezogen oder zumindest informiert werden.

### 2.2. Analyse der Ausgangssituation

Das Versuchsteam muss sich am Anfang mit folgenden Fragen auseinandersetzen:

- Was ist das Ziel der Untersuchung, d.h. welche Frage soll mit dieser Untersuchung beantwortet werden? Dieser Punkt darf nicht unterschätzt werden, denn Klarheit über die Aufgabenstellung ist Voraussetzung für jede erfolgreiche Lösung und die Zufriedenstellung des Auftraggebers. Hier ist eine detaillierte Abstimmung erforderlich.
- Was ist bereits bekannt und wie kann dieses Vorwissen in den Untersuchungen mit einbezogen werden. Nur selten ist man mit völlig neuen Problemstellungen konfrontiert, bei denen kein Vorwissen existiert. Außerdem müssen physikalische Zusammenhänge, so weit wie bekannt, bei der späteren Modellbildung verwendet werden.
- Wie könnten Ergebnisse des Versuchs aussehen und wie werden sich die neuen Erkenntnisse auf die ursprüngliche Fragestellung auswirken. Hier muss man sich ernsthaft mit der Frage auseinandersetzen, ob der Versuch überhaupt in der Lage ist, eine zielführende Lösung der ursprünglichen Aufgabenstellung zu liefern. Es lohnt sich nur selten mit umfangreichen Versuchen den Nachweis zu erbringen, dass das derzeitige Konzept nicht zielführend ist.
- Wie passt die Untersuchung in die sonstige Versuchslandschaft, bzw. in die langfristige Gesamtstrategie. Manchmal bringt ein begrenzter Zusatzaufwand langfristige Aufwandsminimierung oder ermöglicht überhaupt erst eine aussagekräftige Analyse der Ergebnisse. Typische Beispiele aus der Praxis sind vorzeitig abgebrochene Untersuchungen, welche zwar „nur“ 80% des Aufwandes konsumiert, aber auch nur 20 % Erkenntnisgewinn gebracht haben.

### 2.3. Festlegung der Versuchsstrategie

Die passende Versuchsstrategie muss definiert werden, mit welcher eine zufriedenstellende Lösung der Aufgabe gefunden werden kann. DoE bietet für eine Reihe typischer Fragestellungen geeignete Ansätze, welche in den nachfolgenden Kapiteln näher erläutert werden:

- Gezielter Erkenntnisgewinn über die Wirkzusammenhänge im System (Produkt oder Prozess), falls eine physikalische Modellbildung nicht vollständig möglich ist.



## Versuchsmethodik (DoE)

*ANMERKUNG: Diese Aufgabenstellung bildet oft einen Schwerpunkt in der Produktentwicklung. Je nach Faktorenanzahl kommen hier Versuchspläne für einen oder mehreren Faktoren zum Einsatz.*

- Prognose von Produkt- und Prozesseigenschaften in frühen Entwicklungsstadien.

*ANMERKUNG: DoE-Ansätze können sowohl für eine empirische Modellbildung durch physische Versuche als auch bei der Erstellung von Meta-Modellen mit Hilfe von Computerexperimenten verwendet werden, falls noch keine physischen Prototypen existieren.*

- Optimierung von Produkt- und Prozessparametern.

*ANMERKUNG: Geeigneter Lösungsansatz ist hier die Erstellung von Meta-Modellen mit Hilfe von Computerexperimenten mit anschließender modellbasierter Optimierung (z.B. durch Gradientenverfahren) oder alternativ die experimentelle Parameter- bzw. Systemidentifikation mit anschließender modellbasierter Optimierung oder der Anwendung versuchsbasierter Optimierungsstrategien wie der Simplex-Algorithmus.*

- Robustheitsanalysen.

*ANMERKUNG: Falls keine expliziten Modelle vorliegen, kommen hier vorzugsweise Monte-Carlo-Verfahren in Verbindung mit Computerexperimenten und stochastischen Samplingverfahren zum Einsatz. Alternativ können ggf. zunächst explizite Meta-Modelle auf der Basis von Computer-Experimenten abgeleitet werden und anschließend analytische Robustheitsanalysen durchgeführt werden.*

- Robustheitsoptimierung.

*ANMERKUNG: Für die Reduzierung der Streuung mit dem Ziel eines robusten Produktes oder Prozesses bilden die Taguchi- oder verwandte Methoden einen geeigneten Lösungsansatz.*

- Ermittlung der wesentlichen aus einer Menge von möglichen Einflussgrößen und Wechselwirkungen.

*ANMERKUNG: Man spricht hier auch von Screening-Versuchen oder Sensitivitätsuntersuchungen. Als Lösung bieten sich teilfaktorielle Prüfpläne an.*

- Schnelles Erkennen und Beseitigen wichtiger Störgrößen in der Fertigung.

*ANMERKUNG: Hier bietet sich die Methode nach D. Shainin an.*

- Verifikation und Validierung von Qualitätsmerkmalen technischer Erzeugnisse z.B. durch den Nachweis von Funktionalität, Dauerhaltbarkeit, Robustheit oder Sicherheit.

*ANMERKUNG: Hier sind spezielle Versuchstechniken für Erprobung von Bauteilen erforderlich.*

Nachdem die Strategie für die Untersuchung ausgewählt wurde, muss anschließend in einer Abstimmung die endgültige Einigung mit allen Beteiligten und dem Auftraggeber über das Ziel der Untersuchung und der Versuchsstrategie erzielt werden.



### 3. Systemanalyse

Der erste Schritt in Richtung Umsetzung des Versuchs beinhaltet die Analyse des Systems bzgl. seiner Einfluss- und Zielgrößen, wobei die Betrachtung sowohl von Produkten als auch von Prozessen möglich ist.

Der Versuch ist eine empirische Untersuchung über den kausalen Zusammenhang zwischen Einflüssen und Verhalten eines Systems. Zunächst müssen die Größen identifiziert werden, welche das Systemverhalten beschreiben (Zielgrößen) sowie all diejenigen, welche einen Einfluss darauf haben (Einflussgrößen). Je nach Fragestellung ist dies ein zeitaufwendiger, aber ganz wesentlicher Schritt. Falls er übersprungen wird sind spätere Probleme vorprogrammiert, z.B. weil wichtige Einflussgrößen nicht berücksichtigt wurden. Welche von den Einflussgrößen in einem Versuch tatsächlich variiert werden hängt entscheidend von der jeweiligen Fragestellung ab und wird deswegen zu einem späteren Zeitpunkt entschieden.

Die Systemanalyse muss soweit wie möglich auf der Basis einer physikalischen Modellbildung erfolgen, siehe dazu auch Kapitel 1.1.1. Das vorhandene Wissen über das System und die physikalischen Gesetzmäßigkeiten muss für die Identifikation von Ziel- und Einflussgrößen und der Einschätzung über deren Relevanz genutzt werden. Gleichzeitig liefern alle bekannten Zusammenhänge wertvollen Input für die Festlegung des zu verwendenden Modellansatzes, siehe dazu auch Kapitel 4.4.

Zur Identifizierung der Ziel- und Einflussgrößen ist zunächst, soweit bekannt, das Funktionsmodell des betrachteten Systems heranzuziehen. Der Bezug ist allerdings sehr vielfältig, allgemeine Regeln sind kaum aufstellbar.

Hier zwei mögliche Aufgabenstellungen:

- Falls die Zielsetzung des Versuchs die Modellbildung selbst ist, z.B. als Grundlage einer späteren Simulation, sind die Einflussgrößen des Versuchs typischerweise die Eingangsgrößen des Systems, und die Zielgrößen des Versuchs die Ausgangsgrößen des Systems.

*BEISPIEL: Für weitergehende Analysen ist das Emissionskennfeld eines Verbrennungsmotors zu bestimmen. Dazu wird der Verbrennungsmotor als ein System modelliert, welches einen Kraftstoff- /Luftfluss unter einer vorgegebenen Last (Systemeingang) Abgase mit einer bestimmten Zusammensetzung „erzeugt“ (Systemausgang). Durch den Versuch kann die Konzentration von Kohlenstoffmonoxid und den Stickoxyden im Abgas (Zielgrößen) in Abhängigkeit der vorgegebenen Last und dem Kraftstoff-Luft-Mischverhältnis (Einflussgrößen) bestimmt werden.*

- Bei der Optimierung eines Systems tritt typischerweise ein anderer Bezug auf. Einflussgrößen können hier Systemparameter sein, welche verändert werden können. Zielgrößen können sowohl die Eingangs- als auch Ausgangsgrößen des Systems sein, z.B. falls diese als Schnittstelle zur Umwelt unverändert bleiben sollen oder optimiert werden müssen.

*BEISPIEL: Im obigen Beispiel eines Verbrennungsmotors kann z.B. gefordert werden, eine Wirkungsgradverbesserung durch Optimierung der Verbrennung zu erzielen, ohne dabei die Konzentration von Stickoxyden im Abgas zu erhöhen. Dazu werden verschiedene Kolben- und Einspritzdüsengeometrien (Parameter des Systems = Einflussgrößen im Versuch) entworfen, und deren Einfluss auf den Wirkungsgrad und der Konzentration von Stickoxyden (Ausgangsgrößen im System = Zielgröße im Versuch) untersucht.*

Falls die Wirkzusammenhänge nicht ausreichend bekannt sind und somit ein detailliertes Funktionsmodell nicht verfügbar ist, bzw. erst durch den Versuch erstellt werden muss, wird für die Bestimmung der relevanten Einfluss- und Zielgrößen generell folgendes zweistufiges Vorgehen empfohlen:

- Zunächst müssen Systemgrößen von einem Fachexpertenteam bestimmt und anschließend je nach Fragestellung in Einfluss- und Zielgrößen aufgeteilt werden. Hier ist Fachwissen und Erfahrung gefragt. Dieser erste Schritt dient als Basis für die spätere Bewertung der Relevanz der fraglichen Größen und kann methodisch durch verschiedene Ansätze (z.B. Ishikawa-Diagramm) unterstützt werden. Größen, die an dieser Stelle unbeachtet bleiben haben auch nicht die Chance später als relevant eingestuft zu werden.
- Der zweite wichtige Schritt besteht darin, die Einflussgrößen nach Wichtigkeit im Hinblick auf deren möglichen Einfluss auf die Zielgrößen zu bewerten. Dies muss in erster Linie durch das Heranziehen bekannter physikalischer Zusammenhänge erfolgen. Falls solche Zusammenhänge nicht bekannt sind, kann die Technik des Paarweise-Vergleichs oder die Cause & Effect Matrix (C&E Matrix) als Hilfsmittel verwendet werden.

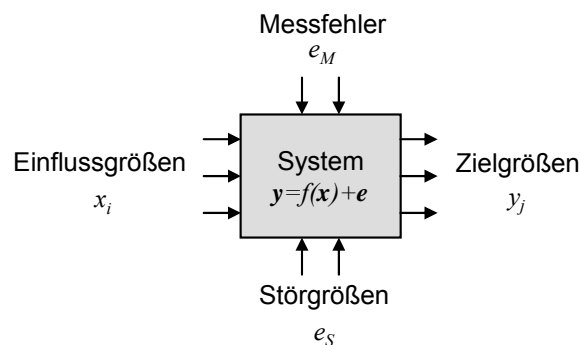


Bei der C&E Matrix wird die Intensität des Einflusses typischerweise mit den Zahlen 0, 1, 3 und 9 für „nicht vorhanden, schwach, deutlich und stark“ bewertet. Die Bewertung muss begründet werden. Die Summe der Bewertungen einer bestimmten Einflussgröße wird als ihre Aktivsumme bezeichnet. Ein Ranking der Einflussgrößen kann anhand ihrer Aktivsummen vorgenommen werden. Die Summe der Bewertungen einer Zielgröße wird als ihre Passivsumme bezeichnet. Diese liefert einen Hinweis auf jene Zielgrößen, welche für die aktuelle Untersuchung relevant sein könnten. Nachfolgend als Beispiel eine C&M-Matrix, aus der geschlossen werden kann, dass vermutlich nur die Wirkung von Einflussgrößen 1 und 3 auf Zielgröße 1 experimentell untersucht werden muss.

	Zielgröße 1	Zielgröße 2	Aktivsumme
Einflussgröße 1	9	1	10
Einflussgröße 2	1	0	1
Einflussgröße 3	3	1	4
Passivsumme	13	2	

**Tabelle 3.1:** Beispiel Cause & Effect Matrix

Das Ergebnis wird oft auch in Form eines P-(Prozess)-Diagramms grafisch dargestellt, als allgemeine Beschreibung der kausalen Beziehung zwischen Einfluss- und Zielgrößen. Die Darstellung umfasst auch die (unbekannten) Messfehler welche die Zielgrößen beeinflussen. Ein zusätzliches Problem stellen die in der Regel streuenden Zielgrößen dar, weil die Kausalität zwischen Einfluss- und Zielgrößen so nicht ohne Weiteres aufrechterhalten werden kann. Abhilfe schafft hier die Einführung zusätzlicher unbekannter Einflussgrößen, genannt Störgrößen, welche ebenfalls die Zielgröße beeinflussen, ähnlich wie die unbekanntes Messfehler. So kann ohne Verlust des Kausalitätsprinzips und im Rahmen einer deterministischen Denkweise die Streuung der Zielgrößen durch das Auftreten von unbekanntes Störgrößen und Messfehlern erklärt werden.



**Bild 3.1:** P-Diagramm eines Versuchs





#### 4. Versuchsplanung

Die Versuchsplanung bildet den Kern der Versuchsmethodik. Zwar können auch historische Daten ausgewertet werden, in diesem Fall ist zu Kapitel 6 überzugehen. Nur systematisch aufgestellte Versuchspläne decken jedoch den Versuchsraum gut ab und sind robust gegen Fehldeutungen. Die Vergleichbarkeit und Aktualität historischer Daten muss sichergestellt sein, bevor mit der Auswertung begonnen werden kann. Folgende Schritte werden in diesem Kapitel näher erläutert:

1. Festlegung der Zielgrößen
2. Festlegung der Faktoren
3. Festlegung des Variationsbereichs der Faktoren
4. Wahl des Modellansatzes
5. Festlegung der Faktorstufen
6. Festlegung der Faktorstufenkombinationen
7. Festlegung der Anzahl der Wiederholungen
8. Festlegung der Reihenfolge und Gruppierung
9. Planung des Versuchsdurchlaufs und der Auswertung
10. Planung der Versuchseinrichtungen und der Prüflinge
11. Aufwandsabschätzung

##### 4.1. Festlegung der Zielgrößen

Zunächst müssen aus allen Zielgrößen diejenigen bestimmt werden, die im Versuch betrachtet werden sollen. Dabei ist folgendes zu beachten:

- Deren Anzahl muss möglichst klein sein, um den Aufwand zu reduzieren. Redundante Zielgrößen verkomplizieren unnötig die Auswertung.
- Es müssen alle Zielgrößen erfasst werden, welche für die Fragestellung relevant sind. Ein Anhaltspunkt, welche der Zielgrößen relevant sein könnten, liefern z.B. die Passivsummen der im Kapitel 3 beschriebenen C&E-Matrix.
- Es dürfen nicht nur die momentan „problematischen“ Größen berücksichtigt werden, da sich z.B. bei einer Optimierung auch andere Zielgrößen verändern können, welche durch Nebenbedingungen eingeschränkt werden müssen.
- Die betrachteten Zielgrößen müssen quantitativen Charakter haben, denn nur solche können sinnvoll ausgewertet werden. Zielgrößen vom Typ „gut-schlecht“ können z.B. durch Einführung einer Bewertungsskala quantifiziert werden.
- Die Größen müssen nach Möglichkeit direkt messbar sein. Falls eine Zielgröße nicht direkt messbar ist, kann:
  - eine korrelierte Größe gesucht werden, um diese direkt zu messen, oder
  - der Versuch zur zerstörungsfreien Zwischenprüfung unterbrochen werden.  
*ANMERKUNG: Dies ist meist nicht rückwirkungsfrei möglich z.B. wegen zwischenzeitliches Handling, Lagerung, Analyse oder Wiederanfahren der Versuchseinrichtung.*
  - Der Versuch kann auch abgestuft mit zerstörender Zwischenprüfung gefahren werden. Hierbei muss die Streuung verschiedener Prüflinge beachtet werden.

Falls mehrere Zielgrößen existieren, können diese separat ausgewertet werden. Schwierigkeiten treten hier bei der Optimierung auf, da unterschiedliche (möglicherweise auch widersprüchliche) Zielgrößen zu berücksichtigen sind. Eine mögliche Lösung bietet die sogenannte Pareto-Optimierung, vgl. dazu auch Kapitel 7.2.

Der Zeitpunkt der Erfassung der Zielgröße muss eindeutig festgelegt werden. Oft soll die Zielgröße nach Beendigung der jeweiligen Versuchsrealisierung erfasst werden. In diesem Fall ist festzulegen, wann die einzelnen Versuchsrealisierungen als beendet gelten. Dies ist für manche Versuche wichtig, bei denen das Versuchsende eventuell in Abhängigkeit eines Parameters festgelegt wird, z.B. bei der



Festlegung eines Abbruchkriteriums von Lebensdauerversuchen. Ggf. muss eine periodische Abtastrate der Zielgröße festgelegt werden.

Nicht zuletzt muss für die Zielgrößen wie für die Faktoren eine Festlegung des Messverfahrens, des Messmittels sowie der Messauflösung erfolgen, mit denen die einzelnen Zielgrößen erfasst werden sollen. Die Frage, ob die Genauigkeit und Präzision der Messmittel und des Messverfahrens auch zur Erreichung der Versuchsziele ausreichen, muss geklärt werden (Messfähigkeitsnachweis).

#### 4.2. Festlegung der Faktoren

Faktoren sind diejenigen Einflussgrößen, welche im Versuch gezielt verändert werden sollen.

Faktoren müssen:

- mutmaßlich einen bedeutenden Einfluss auf die Zielgrößen haben

*ANMERKUNG: Einen Hinweis dazu liefern bekannte physikalische Zusammenhänge oder die Aktiv-Summen der in Kapitel 3 vorgestellten C&E-Matrix. Zu berücksichtigen sind alle Faktoren, welche vermutlich einen Einfluss haben können. Der Versuch selbst liefert im Nachhinein den eindeutigen Beweis, ob ein Faktor tatsächlich relevant ist oder nicht, solange dieser auch tatsächlich im Versuch berücksichtigt wurde. Das Nachträgliche Ausschließen eines Faktors aus dem Modell ist ohne weiteres möglich (die entsprechenden Versuche sind auch nicht „verloren“, sondern können für die bessere statistische Absicherung verwendet werden), nicht jedoch die Aufnahme einer zusätzlichen Einflussgröße. Im Zweifelsfall muss also nach Möglichkeit eher „für“ die Berücksichtigung einer Einflussgröße als Faktor im Versuch entschieden werden. Fehlende entscheidende Einflussgrößen können sich lediglich durch eine erhöhte Streuung der Residuen bemerkbar machen, was nur dann auffällt, falls Erfahrungswerte bekannt sind. In der Regel gibt es also keinen Ansatz, mit dem man entscheiden kann, ob ein wichtiger Faktor vergessen wurde. Falls die Residuenstreuung nicht akzeptabel erscheint, müssen weitere Faktoren berücksichtigt werden.*

*Nach [1] werden für physische Versuche typischerweise 3-6, selten mehr als 10 Faktoren berücksichtigt,*

- direkt oder indirekt messbar sein
- kontrollierbar sein

*ANMERKUNG: Nicht-kontrollierbare Einflussgrößen (beispielsweise Außentemperatur, Luftdruck, aber auch Fertigungsmaße, Passungen) müssen beherrscht, d.h. entweder erfasst oder experimentell bzw. statistisch ausgeschaltet werden.*

- möglichst genau und reproduzierbar eingestellt werden können

*ANMERKUNG: Andernfalls sind spezielle Ansätze wie die D-optimalen Pläne erforderlich, bei denen existierende Punkte im Versuchsraum als Randbedingung für die Planerstellung vorgegeben werden können.*

- den Einstell-Aufwand begrenzen
- unabhängig voneinander sein und unabhängig eingestellt werden können (Orthogonalitätsbedingung)
- nach Möglichkeit quantitativen Charakter haben

*ANMERKUNG: Qualitative Faktoren bringen Schwierigkeiten mit sich (Modellgleichung ist diskontinuierlich, Interpolation zwischen den Stufen ist sinnlos, Auswertung erfolgt durch Varianzanalyse und nicht durch Regression), können aber durchaus ebenfalls betrachtet werden.*

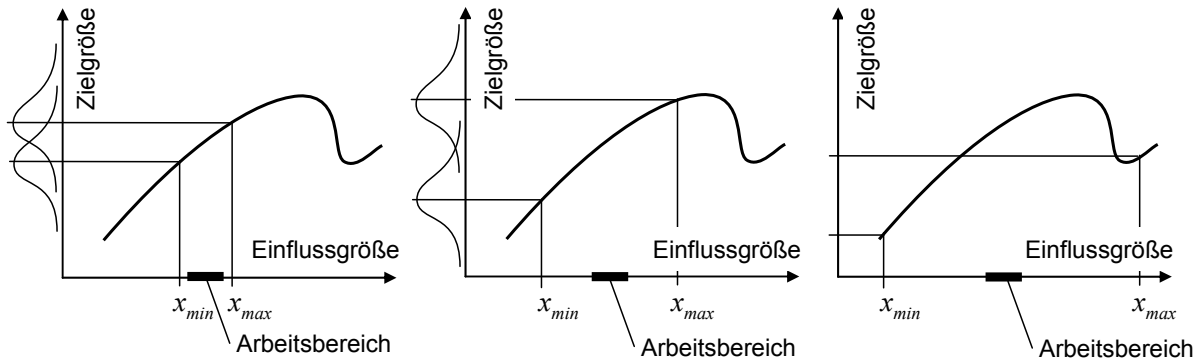
Alle anderen Einflussgrößen, die keine Faktoren sind, müssen beobachtet, dokumentiert und möglichst konstant gehalten werden.

#### 4.3. Festlegung des Variationsbereichs der Faktoren

Der interessierende Arbeitsbereich der Faktoren muss ermittelt und auf dieser Basis der Variationsbereich für den Versuch festgelegt werden. Der Variationsbereich deckt mindestens den Arbeitsbereich ab. Dabei tritt allerdings auch die Problematik „kleine Stufe - kleiner Effekt“ auf: Der Effekt kann manchmal undeutlich sein, weil die Variationsstufe klein ist. Somit muss die Stufe so groß gewählt werden, dass der Effekt sichtbar wird. In der Regel geht der Variationsbereich also deutlich über den Arbeitsbereich hinaus, insbesondere bei erhöhter Streuung. Allerdings darf der Variationsbereich auch nicht zu groß gewählt werden, denn dann können lokale Effekte das Ergebnis beeinflussen.



Zu Beachten sind mögliche Einschränkungen im Versuchsraum, z.B. Bereiche die unsinnig, nicht machbar, gefährlich oder anderweitig unzulässig sind. Falls solche Bereiche auftreten sind spezielle Versuchspläne erforderlich, z.B. die D-optimalen Pläne, siehe Kapitel 8.



**Bild 4.1:** Arbeits- und Variationsbereich ( $x_{min}$ ,  $x_{max}$ ) eines Faktors (schematisch). Links: Variationsbereich ist zu klein (Effekt bei der vorliegenden Streuung zu undeutlich). Mitte: Variationsbereich ist richtig gewählt. Rechts: Variationsbereich ist zu groß (lokale Effekte am rechten Rand sind für den Arbeitsbereich nicht von Interesse).

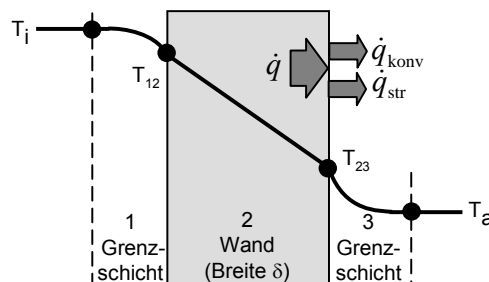
#### 4.4. Wahl des Modellansatzes

Der Modellansatz muss so ausgewählt werden, dass er in der Lage ist die Wirkzusammenhänge ausreichend genau zu beschreiben. Dies stellt einen wichtigen Input der Versuchsplanung dar. Das Aufstellen des Modellansatzes ist ein wichtiger Schritt, welcher nicht etwa aus Zeitgründen übersprungen werden darf, denn Fehler in dieser Phase können durch die Versuchsführung später nicht aufgehoben werden. Ein fehlerhaftes Modell führt zu einem **Modellfehler**, zusätzlich zu den in Kapitel 1.2.1 erwähnten Fehlerarten.

Um Missverständnisse vorzubeugen: lediglich der **Modellansatz**, d.h. der Typ der mathematischen Transferfunktion muss hier festgelegt werden und nicht die Transferfunktion selbst. Die Ermittlung der kompletten Funktion geschieht erst bei der Versuchsauswertung. Falls ein einfacher Polynomansatz als Transferfunktion gewählt wird, stellt sich lediglich die Frage nach der Polynom-Ordnung. Die Bestimmung der freien Koeffizienten des Polynoms erfolgt durch Auswertung der späteren Versuche, z.B. mittels Regression.

Bei der Aufstellung des Modellansatzes muss soweit wie möglich eine physikalische Modellbildung erfolgen, siehe dazu auch Kapitel 1.1.1. Das vorhandene Wissen über das System und die physikalischen Gesetzmäßigkeiten muss genutzt werden, indem alle bekannten Zusammenhänge in dem Modellansatz eingeordnet werden. Die Mindestanforderung ist, dass das Modell nicht den Erfahrungen widersprechen darf. Ist z.B. eine exponentielle Abhängigkeit wie bei der Bevölkerungsentwicklung zu erwarten, darf kein logarithmischer Ansatz verwendet werden. Das folgende Beispiel soll das Zusammenspiel zwischen physikalische und empirische Modellbildung verdeutlichen.

*BEISPIEL:* Die Verlustleistung eines Ofens soll in Abhängigkeit der Ofentemperatur bei gegebener Umgebungstemperatur bestimmt werden. Bild 4.2 zeigt eine schematische Darstellung der Ofenwand mit den angrenzenden Bereichen, zusammen mit dem Temperaturverlauf vom Inneren des Ofens ( $T_i$ ) bis zur äußeren Umgebung ( $T_a$ ).



**Bild 4.2:** Temperaturverlauf in der Ofenwand

2020-04-06 - SOCCOS



Die Verlustleistung des Ofens entsteht durch Konvektion und Strahlung. Diese physikalischen Effekte können durch entsprechende Wärmestromdichten auf Basis thermodynamischer Zusammenhänge wie folgt beschrieben werden:

$$\dot{q}_{\text{konv}} = \alpha_3 \cdot (T_{23} - T_a); \quad \dot{q}_{\text{str}} = \sigma \varepsilon \cdot (T_{23}^4 - T_a^4). \quad (4.1)$$

Mit  $\alpha_3$  wurde der Wärmeübergangskoeffizient in der Grenzschicht 3 bezeichnet,  $\sigma$  ist eine bekannte Naturkonstante (Stefan-Boltzmann-Konstante),  $\varepsilon$  bezeichnet den Emissionsgrad der Wand. Eine Wärmestrombilanz an der Außenfläche 2-3 der Ofenwand ergibt:

$$\dot{q} = \dot{q}_{\text{konv}} + \dot{q}_{\text{str}} = \alpha_3 \cdot (T_{23} - T_a) + \sigma \varepsilon \cdot (T_{23}^4 - T_a^4), \quad (4.2)$$

wobei  $\dot{q}$  die Wärmestromdichte aus dem Ofeninneren bezeichnet. Diese muss durch Wärmeleitung über die Ofenwand und durch Konvektion über die Grenzschicht 1 im Inneren transportiert werden:

$$\dot{q} = \frac{\lambda}{\delta} \cdot (T_{12} - T_{23}), \quad (4.3)$$

$$\dot{q} = \alpha_1 \cdot (T_i - T_{12}). \quad (4.4)$$

Mit  $\lambda$  wurde die Wärmeleitfähigkeit der Wand (eine Materialkonstante) bezeichnet. Einsetzen von Gl. 4.4 in Gl. 4.3 führt zu

$$T_{23} = T_i - \dot{q} \cdot \left( \frac{1}{\alpha_1} + \frac{\delta}{\lambda} \right), \quad (4.5)$$

daraus folgt folgender Ausdruck für die Wärmestromdichte:

$$\dot{q} = \alpha_3 \cdot \left( T_i - \dot{q} \cdot \left( \frac{1}{\alpha_1} + \frac{\delta}{\lambda} \right) - T_a \right) + \sigma \varepsilon \cdot \left( \left( T_i - \dot{q} \cdot \left( \frac{1}{\alpha_1} + \frac{\delta}{\lambda} \right) \right)^4 - T_a^4 \right). \quad (4.6)$$

Für die Lösung der Aufgabe anhand obiger Gleichung ergeben sich nun folgende Schwierigkeiten:

- Die gesuchte Wärmestromdichte steht sowohl auf der linken, als auch auf der rechten Seite der Gleichung (hier sogar in der 4. Potenz), was die symbolische Auflösung unmöglich macht. Der gewünschte Zusammenhang kann also nur in impliziter Form angegeben werden:

$$F(\dot{q}, T_i, \dots) = 0 \quad (4.7)$$

- Die Gleichung enthält unbekannte Konstanten wie  $\alpha_1$ ,  $\alpha_3$  und  $\varepsilon$ , welche experimentell bestimmt werden müssen. Die ersten beiden hängen von den Konvektions-Strömungsverhältnissen in der Grenzschicht, letztere von der Beschaffenheit (z.B. Rauheit und Farbe) der Wandfläche ab.

An dieser Stelle setzt die empirische Modellbildung unter Berücksichtigung der bekannten physikalischen Zusammenhänge an. Folgender Modellansatz würde sich dafür anbieten, da in Gleichung 4.6 die Innentemperatur in der 4. Potenz enthalten ist:

$$\dot{q} = a_0 + a_1 T_i + a_2 T_i^2 + a_3 T_i^3 + a_4 T_i^4 \quad (4.8)$$

Die Koeffizienten  $a_i$  werden experimentell durch Messungen des Wärmestroms bei unterschiedlichen Innentemperaturen im stationären Zustand bestimmt. Sie hängen implizit von den oben erwähnten Konstanten und von der Umgebungstemperatur ab.

Falls eine white-box-Modellbildung nicht möglich ist, kann also ein genereller Ansatz als Polynom n-ten Grades für die empirische Modellbildung verwendet werden. Hintergrund ist die Tatsache, dass durch Taylor-Entwicklung im betrachteten Bereich gezeigt werden kann, dass jedes stetige Verhalten lokal durch einen Polynom bis auf ein Restglied höherer Ordnung angenähert werden kann.

Ein multilinearer Ansatz berücksichtigt lineare Effekte und darüber hinaus auch Wechselwirkungen durch die gemischten Terme der Faktoren. Für 1, 2 bzw. 3 Faktoren sieht ein multilinearer Ansatz wie folgt aus:

$$y = f(x_1) = a_0 + a_1 x_1 \quad (4.9)$$

$$y = f(x_1, x_2) = a_0 + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2 \quad (4.10)$$

$$y = f(x_1, x_2, x_3) = a_0 + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2 + a_3 x_3 + a_{13} x_1 x_3 + a_{23} x_2 x_3 + a_{123} x_1 x_2 x_3 \quad (4.11)$$

Ansätze höherer Ordnung berücksichtigen darüber hinaus auch quadratische, kubische usw. Effekte der Faktoren. Als Beispiel sei hier ein Polynom 2.-Grades erwähnt:

$$y = f(x_1) = a_0 + a_1 x_1 + a_{11} x_1^2 \quad (4.12)$$

$$y = f(x_1, x_2) = a_0 + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2 + a_{11} x_1^2 + a_{22} x_2^2 \quad (4.13)$$

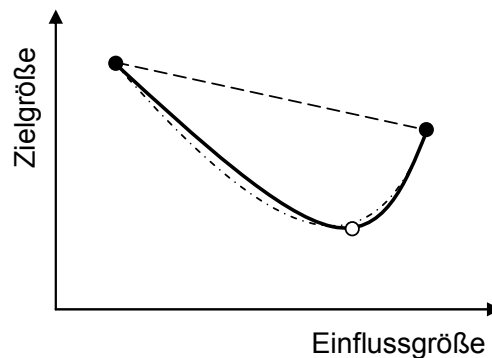


$$y = f(x_1, x_2, x_3) = a_0 + a_1x_1 + a_2x_2 + a_{12}x_1x_2 + a_3x_3 + a_{13}x_1x_3 + a_{23}x_2x_3 + a_{123}x_1x_2x_3 + \dots$$

$$\dots + a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 \quad (4.14)$$

Es kann erforderlich werden, den Ansatz weiter zu vereinfachen. Dies erfolgt z.B. durch eine Linearisierung oder durch eine Variablentransformation, falls das Modell nichtlinear in den Variablen, jedoch linear in den Koeffizienten ist. Beispielsweise könnte anhand physikalischer Überlegungen festgestellt worden sein, dass die Zielgröße exponentiell von einer Einflussgröße abhängt. Das einfachste Vorgehen besteht nun darin, ein lineares Modell mit der Einflussgröße und dem Logarithmus der Zielgröße aufzubauen oder eine neue Einflussgröße durch Substitution einzuführen, welche exponentiell von der ursprünglichen Einflussgröße abhängt. Weitere Details dazu siehe [2], S.34.

Es muss überprüft werden, ob die Erwartungen an Aussagen und Genauigkeit sich mit dem Modell decken. So würde z.B. eine Optimierungsaufgabe generell gegen ein lineares Modell sprechen, da in der Umgebung des Optimums die Krümmung nicht vernachlässigt werden kann, siehe nachfolgendes Bild. Der Modellansatz muss also die Umgebung des interessierenden Bereichs (z.B. in der Umgebung von Extremwerten der Zielgrößen) möglichst gut und detailliert wiedergeben. Andererseits steht der Modellansatz generell in Bezug zum Versuchsumfang, da die Berechnung der Modellparameter mit den geplanten Versuchen möglich sein muss. Ein Ansatz höherer Ordnung kann den Versuchsumfang derart erhöhen, dass eine praktische Umsetzung nicht mehr möglich ist. Ziel ist es also, von allen möglichen Modellen das einfachste auszuwählen.



**Bild 4.3:** Schematischer Verlauf einer Zielgröße im Bereich des Optimums (durchgezogene Linie). Ein linearer Ansatz (gestrichelte Linie) wäre nicht in der Lage das Optimum zu finden. Besser geeignet ist ein quadratischer Ansatz (Strich-Punkt-Linie).

#### 4.5. Festlegung der Faktorstufen

Auf Basis des Modellansatzes sind nun die Faktorstufen nach Anzahl und Wert festzulegen. Hier wurde die meiste Arbeit bereits bei der Festlegung des Modellansatzes getan (siehe auch Kapitel 4.4), denn jeder Ansatz erfordert eine bestimmte Mindestanzahl von Faktorstufen. Ein linearer Ansatz erfordert beispielsweise mindestens 2 Faktorstufen, ein quadratischer Ansatz 3 Faktorstufen etc. Hintergrund dabei: Die Anzahl freier Parameter im Modell muss mit der vorhandenen Versuchsanzahl bestimmt werden können.

Obwohl also die Aufgabe nicht unabhängig von der Modellansatzwahl gelöst werden kann, sind gewisse Freiräume bei der Wahl der Faktorstufen vorhanden. Zunächst kann über die erforderliche Mindestanzahl Stufen hinausgegangen werden, um die Zweckmäßigkeit des gewählten Ansatzes zu überprüfen: Bei einem linearen Ansatz kann z.B. zusätzlich ein Versuchspunkt in der Mitte des Variationsbereiches gelegt werden, um die Abweichung des tatsächlichen Verhaltens von der Linearität zu überprüfen. Bei quadratischen Ansätzen kann z.B. entschieden werden, ob zentral-zusammengesetzte oder flächenzentrierte Pläne verwendet werden, siehe dazu auch Kapitel 8.2.

Zu berücksichtigen ist außerdem, dass qualitative Faktoren nur bestimmte von der Fragestellung definierte Werte einnehmen können. Bei einem Faktor „Werkstoff“ können z.B. keine Zwischenwerte zwischen Metall und Kunststoff eingenommen werden. Gleiches gilt bei einem Faktor Lieferant, Bediener oder Prüfer.



#### 4.6. Festlegung der Faktorstufenkombinationen

An dieser Stelle werden nur 3 grundlegende Vorgehensweisen erwähnt, weiterführende Ansätze sind in Kapitel 8 zu finden.

##### 4.6.1. Trial-and-error Verfahren

Damit wird oft die einfachst-mögliche Vorgehensweise bezeichnet, bei der nicht zielgerichtet vorgegangen, sondern einfach „ausprobiert“ wird, bis eine zufriedenstellende Lösung gefunden wurde. So verlockend der Ruf dieser Vorgehensweise ist, muss beachtet werden, dass experimentelle Untersuchungen in der Regel teuer sind und die Verschwendung von Ressourcen vermieden werden muss.

An dieser Stelle soll auch deutlich darauf hingewiesen werden, dass das Argument, so würde man über die Fragestellung viel lernen, nicht wirklich greift. Der Erkenntnisgewinn hält sich in Grenzen, wenn keine Systematik bei der Untersuchung angewendet wurde. Man kann beispielsweise niemals sicher sein, ein Optimum gefunden zu haben, auch wenn in einer Versuchsrealisierung dieses zufällig getroffen wurde (was extrem unwahrscheinlich ist). Im Unterschied dazu kann man bei einer modellbasierten Optimierung das Optimum auch dann gut vorhersagen, wenn es nicht Teil einer Versuchsrealisierung war.

Eine nicht zielgerichtete Vorgehensweise ist somit nicht sinnvoll. Aufgrund der erwähnten Nachteile kann dieser Ansatz auch nicht wirklich als Bestandteil einer Versuchsmethodik angesehen werden und wird hier nicht weiter betrachtet.

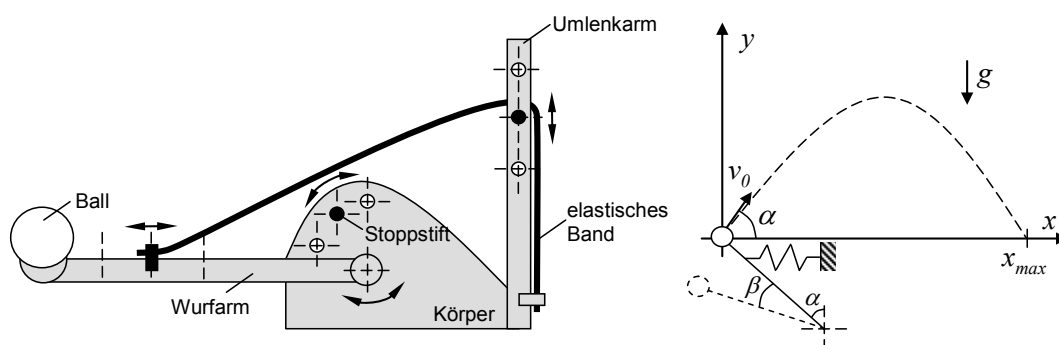
##### 4.6.2. One-factor-at-a-time (OFAT) Verfahren

Mit OFAT werden Versuche bezeichnet, bei denen zunächst eine Basis-Einstellung der Faktoren definiert und anschließend jeweils nur ein Faktor variiert wird, während alle anderen ihre Basis-Einstellung beibehalten. Bei der Variation des nächsten Faktors wird der vorhergehende wieder auf seine Basis-Einstellung zurückgebracht usw. So werden nacheinander alle Faktoren variiert. Als Begründung dieser Vorgehensweise dient oft das Argument, dass so die Wirkung einer einzigen Ursache, nämlich jenem Faktor zugeschrieben werden kann, der aktuell variiert wurde.

Zunächst kann festgestellt werden, dass falls nur ein einziger Faktor vorkommt, dies sicherlich die natürlichste Vorgehensweise darstellt.

Darüber hinaus kann diese Vorgehensweise auch dann verwendet werden, wenn zwar mehrere Faktoren vorkommen, aber die Effekte und Wechselwirkungen aller Faktoren bis auf einen anhand physikalischer Überlegungen prognostiziert werden können. Der Versuch mit dem verbleibenden Faktor wäre quasi ein OFAT-Versuch.

*BEISPIEL: Wir betrachten das System „Katapult“, welches Bälle in eine gewisse Entfernung schleudern kann.*



**Bild 4.4:** Das System Katapult

Das Katapult besteht aus einem Körper, einem drehbaren Wurfarm, einem fixen Umlenkarm und einem elastischen Band, welches den Wurfarm über den Umlenkarm mit der vorderen Seite des Katapultkörpers verbindet. Am oberen Ende des Wurfarms befindet sich eine Mulde, welche einen Ball aufnehmen kann. Die Funktionsweise ist recht einfach: der drehbare Wurfarm wird aus seiner Ruheposition ausgelenkt, dabei wird das elastische Band gespannt. Beim Loslassen wird der Wurfarm durch das elastische Band zurück zu seiner Ruheposition gezogen und beschleunigt so den Ball. Der Wurfarm prallt beim Erreichen der Ruheposition auf einen Stift und wird so gestoppt. Der Ball wird in diesem Moment nach vorne weggeschleudert. Folgende Einstellungen beeinflussen die Wurfweite:



- Auslenkungswinkel  $\beta$
- Stoppposition (Ruheposition) des Wurfarms, charakterisiert durch Winkel  $\alpha$
- Einspannstelle am Wurfarm
- Einspannstelle am Umlenkarm

Die Aufgabe besteht nun darin, eine Modellierung des Katapults vorzunehmen, um so in Zukunft die für eine bestimmte Wurfweite erforderlichen Einstellungen ermitteln zu können.

Als Erstes müssen bekannte physikalische Zusammenhänge ermittelt und bei der Modellierung berücksichtigt werden. Bei diesem Beispiel ist dies auch ohne weiteres möglich: Das Anspannen des elastischen Bandes bei der Auslenkung des Wurfarms kann durch eine Feder modelliert werden, welche angespannt wird und eine Rückstellkraft und somit auch die beschleunigte Bewegung des Balls bis zum Moment des Abschusses erzeugt. Die Bahn des Balls nach dem Abschuss kann durch die Gesetze des schiefen Wurfs beschrieben werden:

$$x(t) = v_0 \cos \alpha \cdot t \quad (4.15)$$

$$y(t) = -g \frac{t^2}{2} + v_0 \sin \alpha \cdot t \quad (4.16)$$

Die Wurfweite  $x_{\max}$  kann aus der Überlegung abgeleitet werden, dass zum Zeitpunkt  $t_{\max}$  des Aufschlagens des Balls auf dem Boden  $y(t_{\max}) = 0$  ist. Mit Gl. 4.16 folgt, dass

$$t_{\max} = 2 \frac{v_0}{g} \sin \alpha \quad (4.17)$$

und nach einsetzen in Gl. 4.15

$$x_{\max} = \frac{v_0^2}{g} 2 \sin \alpha \cos \alpha = \frac{v_0^2}{g} \sin(2\alpha) \quad (4.18)$$

Allerdings fehlen für eine komplette white-box-Modellierung noch einige Informationen. So sind Eigenschaften der Systemelemente nicht bekannt, z.B. die Elastizitätskonstante des elastischen Bandes, so dass damit auch  $v_0$  unbekannt ist. Die bekannten physikalischen Zusammenhänge können aber in eine grey-box-Modellbildung einfließen und durch Versuche zur Parameteridentifikation ergänzt werden.

Es stellt sich die Frage, welche der vier oben erwähnten Einflussgrößen als Faktor in einem Versuch variiert werden müssen. Zunächst ist klar, dass falls der Auslenkwinkel  $\beta$  Null ist, die Wurfweite ebenfalls Null sein wird, und zwar unabhängig von allen anderen Einstellungen. Bei maximaler Auslenkung  $\beta = 90^\circ - \alpha$  wird auch die Wurfweite maximiert; welche Wurfweite erreicht werden kann, hängt allerdings von den übrigen Einstellungen ab. Aus physikalischen Überlegungen ist z.B. klar, dass beide Einspannstellen in eindeutiger Weise die Rückstellkraft der Feder „elastisches Band“ und somit die Beschleunigung und Abwurfgeschwindigkeit des Balls beeinflussen. Maximale Vorspannung bedeutet maximale Abwurfgeschwindigkeit  $v_0$  und nach Gl. 4.18 bedeutet dies auch maximale Wurfweite  $x_{\max}$ . Bleibt noch die Stoppposition des Wurfarms, hier kann anhand obiger Gleichung einfach abgeleitet werden, dass die Wurfweite bei einer Einstellung der Stoppposition von  $\alpha = 45^\circ$  maximiert wird, bei der  $\sin(2\alpha) = 1$  ist.

Zusammenfassend lässt sich also — ohne einen Versuch gemacht zu haben — sagen, dass mit den Einstellungen der Einspannstellen auf maximale Vorspannung und einer Stoppposition auf  $\alpha = 45^\circ$

- bei einer Auslenkung  $\beta = 45^\circ$  die maximal erreichbare Wurfweite erzeugt wird,
- bei einer Auslenkung von  $\beta = 0^\circ$  die Wurfweite 0 erzeugt wird und
- jede Wurfweite dazwischen nur durch Variation des Auslenkwinkels  $\beta$  erreicht werden kann.

Insofern liegt eine Aufgabe vor, bei der  $x_{\max}$  nur in Abhängigkeit des Faktors Auslenkwinkel  $\beta$  zu untersuchen ist. Die Zuordnung, welcher Auslenkwinkel für welche Wurfweite erforderlich wäre, ist allerdings nur experimentell zu klären, da wie bereits erwähnt wichtiges Systemwissen über Eigenschaften des Katapults fehlt. Dies kann natürlicherweise in einem OFAT-Versuch erfolgen.

Der Versuch kann so ausgeführt werden, dass in mehreren Stufen der Auslenkungswinkel  $\beta$  zwischen 0 und  $45^\circ$  variiert und die zugehörige Wurfweite  $x_{\max}$  registriert wird. Vorzugsweise würden 2 Stufen erforderlich sein, nämlich die Maximaleinstellung und die Hälfte davon, denn im konkreten Fall führt  $\beta = 0$  trivialerweise zu  $x_{\max} = 0$ . Nur die Maximaleinstellung zu untersuchen würde nicht ausreichen, denn das würde eine lineare Abhängigkeit zwischen Wurfweite und Auslenkungswinkel bedeuten. Eine solche Annahme wäre zum gegenwärtigen Wissensstand sicherlich gewagt. Falls die Versuche aber sehr aufwändig wären, würde es sich sicherlich lohnen darüber nachzudenken, ob so eine Annahme eventuell zutreffen könnte.



Mit den Versuchspunkten bei halber und maximaler Auslenkung (und dem Wissen über die Auslenkung Null) wäre man in der Lage, ein quadratisches Modell der Beziehung zwischen Auslenkung und Wurfweite aufzustellen:

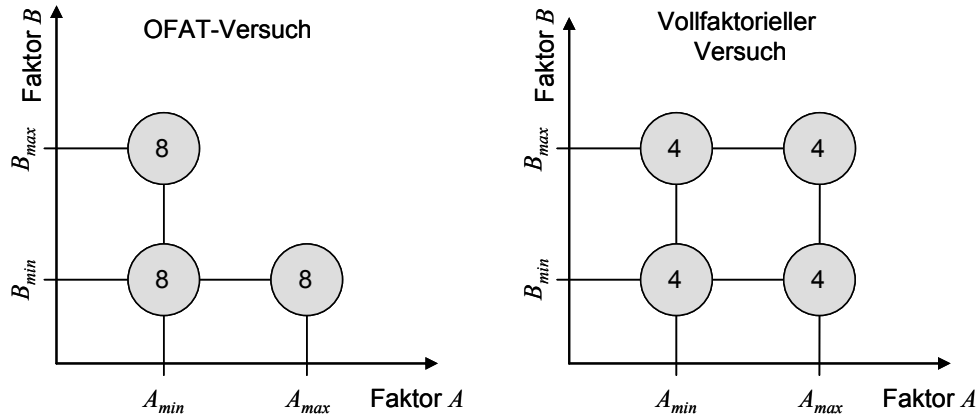
$$x_{\max} = a_0 + a_1\beta + a_2\beta^2 \quad (4.19)$$

Ansätze höherer Ordnung wären wahrscheinlich nicht zielführend, denn diese würden das Interesse an einem Wendepunkt voraussetzen, welcher bei der vorliegenden monotonen Beziehung (größere Auslenkung führt zu größeren Wurfweiten) wahrscheinlich aufgrund der begrenzten Genauigkeit der Untersuchungen nicht ermittelt werden kann. Über die Anzahl der Versuchsrealisierungen in beiden Versuchspunkten muss man sich aber je nach Streuung der Wurfweite gesondert Gedanken machen; dies wird in Kapitel 4.7 näher erläutert.

Obwohl OFAT-Versuche beim Vorliegen eines einzigen (oder quasi eines einzigen) Faktors die natürliche Vorgehensweise darstellen, weist diese Art der Versuchsführung bei mehreren Faktoren zwei wesentliche Nachteile auf:

- 1) Wechselwirkungen zwischen den Faktoren werden nicht erfasst. Mit Wechselwirkungen ist dabei gemeint, dass die Wirkung eines Faktors auf die Zielgröße, entscheidend von der Einstellung der restlichen Faktoren abhängen kann. Beispielsweise kann bei einer bestimmten Einstellung der restlichen Faktoren die Zielgröße mit steigendem Faktor ebenfalls steigen. Bei einer anderen Einstellung aber kann dieser Effekt verstärkt, abgeschwächt oder gar umgekehrt werden, so dass nun mit steigendem Faktor die Zielgröße plötzlich sinkt. Um diesen Effekt zu erfassen reicht es also nicht, den Faktor nur bei einer Basis-Einstellung der restlichen Faktoren zu untersuchen, wie in einem OFAT-Versuch.
- 2) Bei einem OFAT-Versuch entsteht unnötig hoher Aufwand, wenn bei streuenden Zielgrößen eine Mehrfachrealisierung der Versuchspunkte erforderlich ist, um die erforderliche Genauigkeit zu erreichen.

Um beides zu verdeutlichen, betrachten wir eine Fragestellung, bei der die Wirkung zweier Faktoren  $A$  und  $B$  — jeweils zwischen einem Minimal- und Maximalwert — auf eine Zielgröße untersucht werden soll.



**Bild 4.5:** OFAT-Versuche im Vergleich zum vollfaktoriellen Versuchsplan

Aus Erfahrung ist bekannt, dass die Zielgröße streut und deswegen eine 8-fach-Realisierung je Versuchsvariante erforderlich ist, um die Effekte richtig zu erfassen (Details dazu werden später diskutiert). In einem OFAT-Versuch mit der Basiseinstellung  $(A_{\min}, B_{\min})$  müssten zunächst der Faktor  $A$  und anschließend der Faktor  $B$  auf das jeweilige Maximum eingestellt werden, es wären also zusammen mit der Basisvariante 3 Versuchspunkte mit je 8 Realisierungen und somit insgesamt 24 Einzelversuche erforderlich. Mit jedem weiteren Faktor würden weitere 8 Versuche fällig.

Trotz des relativen hohen Versuchsaufwandes bleiben aber wichtige Fragen unbeantwortet. So ist die Wirkung des Faktors  $B$  auf die Zielgröße im gesamten interessierenden Bereich bekannt, falls der Faktor  $A$  seine Einstellung bei  $A_{\min}$  hat, aber bei der Einstellung  $A_{\max}$  liegt Information nur bei einem Randpunkt des Versuchsraums, nämlich bei  $B_{\min}$  vor. Daraus wird deutlich, dass der Basiskombination bei einem OFAT-Versuch auf der einen Seite eine herausragende Bedeutung zugemessen wird. Auf der anderen Seite erfolgt deren Wahl aber oft willkürlich, da es keine objektive Grundlage dafür gibt. Warum soll  $(A_{\min}, B_{\min})$  besser dafür geeignet sein als  $(A_{\min}, B_{\max})$  oder  $(A_{\max}, B_{\min})$ ?

Ein optimierter Ansatz wird im folgenden Kapitel vorgestellt.





### 4.6.3. Faktorielle Versuchspläne

Für die Erfassung von Wechselwirkungen ist ein zusätzlicher Punkt bei  $A_{max}$  und  $B_{max}$  notwendig. Damit kann die Wirkung von  $B$  komplett beschrieben werden, sowohl bei  $A_{min}$  als auch bei  $A_{max}$ . Diese Vorgehensweise hat auch weitere Vorteile, insbesondere im Hinblick auf den erforderlichen Versuchsaufwand. Bei einer 4-fach-Realisierung jeder Variante ergibt sich für jede Faktorstufe dieselbe statistische Absicherung wie bei dem OFAT-Versuch. Bei  $A_{min}$  existieren wieder insgesamt 8 Einzelversuche, nämlich 4 bei  $(A_{min}, B_{min})$  und 4 bei  $(A_{min}, B_{max})$ . Dasselbe gilt auch für  $A_{max}$ . Trotzdem werden statt 24 Einzelversuche bedeutend weniger, nämlich 16 Einzelversuche benötigt (4 Versuchspunkte  $\times$  4 Realisierungen). Der zusätzliche Versuchspunkt bei  $(A_{max}, B_{max})$  wird in gewisser Hinsicht mehrfach verwendet. Mit wachsender Faktorenanzahl erhöht sich dieser Effekt. So wären bei einem OFAT Versuch mit 4 Faktoren  $(4+1) \times 8 = 40$  Einzelversuche erforderlich, beim obigen Schema würden aber immer noch  $2^4 = 16$  Einzelversuche ausreichen, mit einfacher Realisierung in jedem Versuchspunkt. Unsymmetrische Faktorkombinationen verhindern die Vergleichbarkeit der Ergebnisse. Eine symmetrische Abdeckung des Versuchsraums bringt hingegen nur Vorteile.

Fazit:

Die beschriebene Vorgehensweise braucht deutlich weniger Versuche und ermöglicht gleichzeitig die Erkennung von Wechselwirkungen der Faktoren. Pläne, bei denen alle Faktor-Kombinationen variiert werden, heißen vollfaktorielle Versuchspläne. Sie bilden die Grundlage der statistischen Versuchsplanung.

2020-04-06 - SOCCOS

2 <sup>4</sup> -Plan				
2 <sup>3</sup> -Plan				
2 <sup>2</sup> -Plan				
Versuch	A	B	C	D
1	-	-	-	-
2	+	-	-	-
3	-	+	-	-
4	+	+	-	-
5	-	-	+	-
6	+	-	+	-
7	-	+	+	-
8	+	+	+	-
9	-	-	-	+
10	+	-	-	+
11	-	+	-	+
12	+	+	-	+
13	-	-	+	+
14	+	-	+	+
15	-	+	+	+
16	+	+	+	+

**Tabelle 4.1:** Vollfaktorieller Plan mit 4-Faktoren auf zwei Stufen

Bei einem vollfaktoriellen Plan mit  $k$ -Faktoren auf zwei Stufen treten  $m = 2^k$  Kombinationen auf. Solche Pläne werden deswegen als  $2^k$ -Pläne bezeichnet. Tabelle 4.1 zeigt die Kombinationen für 2 bis 4 Faktoren auf zwei Stufen. Die Bezeichnungen „+/-“, die üblicherweise verwendet werden, kennzeichnen die obere bzw. untere Stufe des jeweiligen Faktors. Das Schema kann leicht auf eine beliebige Faktorenanzahl erweitert werden, indem der Plan für die untere und für die obere Stufe des nächsten Faktors wiederholt wird.

Wie aus obiger Formel ersichtlich, steigt der Versuchsaufwand mit wachsender Faktoranzahl stark an. Bereits bei 6 Faktoren sind somit 64 Versuchspunkte erforderlich, welche gegebenenfalls auch mehrfach realisiert werden müssen. Dabei haben wir lediglich 2 Stufen, auf denen jeder Faktor variiert



werden muss, und somit nur einen linearen Modellansatz zur Beschreibung seiner Wirkung auf die Zielgröße gewählt. Falls z.B. quadratische Effekte berücksichtigt werden sollen, sind für einen vollfaktoriellen Plan 3 Stufen je Faktor erforderlich; der Aufwand steigt auf  $3^k$  Versuchspunkte. Wie mit solchen Aufgabenstellungen sinnvoll umzugehen ist, zeigen weitergehende Ansätze in Kapitel 8. Diese können den Prüfaufwand in begrenztem Maße reduzieren.

#### 4.7. Festlegung der Anzahl der Wiederholungen

In Abhängigkeit von der erwarteten Streuung des Versuchs und der geforderten Genauigkeit, mit welcher der untersuchte Effekt bestimmt werden soll, muss die Gesamtanzahl  $N$  der Einzelversuche festgelegt werden.

Nach [1], S. 78 gilt

$$N = 60 \left( \frac{\sigma}{\Delta} \right)^2, \tag{4.20}$$

wobei  $\sigma$  die Standardabweichung als Maß für die Versuchsstreuung ist und  $\Delta$  den Effekt bezeichnet, welcher sicher nachgewiesen werden soll. Die Wahrscheinlichkeit, dass ein tatsächlich vorhandener Effekt nicht erkannt wird, wurde hier zu 10% angenommen (Fehler 2. Art). Die Wahrscheinlichkeit, dass ein Effekt erkannt wird, obwohl dieser nicht vorhanden ist, beträgt 1% (Fehler 1. Art). Es handelt sich hierbei um eine zweiseitige Fragestellung, da sowohl positive als auch negative Effekte erkannt werden sollen, vgl. auch Kapitel 9.1.4

Falls der Versuch  $m$  Faktorstufenkombinationen umfasst, können  $n = N / m$  Versuchsrealisierungen auf jeder Faktorstufe durchgeführt werden. Vorteil beim Einsatz faktorieller Versuchspläne ist dabei, dass mehr Faktoren gleichzeitig untersucht werden können, fast ohne Zunahme des Versuchsumfangs.

Tabelle 4.2 soll dies verdeutlichen. Falls ein Effekt in der Größenordnung  $1\sigma$  erkannt werden soll, sind  $N = 60$  Einzelversuche erforderlich. Abhängig von der Faktoranzahl erhält man damit:

Anzahl Faktoren	Kombinationen	Realisierungen pro Kombination $n$
1	$2^1 = 2$	30
2	$2^2 = 4$	15
3	$2^3 = 8$	8
4	$2^4 = 16$	4

**Tabelle 4.2:** Mögliche Realisierungen bei konstanter Anzahl Einzelversuche

Falls die Anzahl der Faktoren keine Mehrfachrealisierung erlaubt ( $n=1$ ) ergeben sich Schwierigkeiten bei der späteren Signifikanzanalyse. Folgende Möglichkeiten bestehen:

- Der Versuchsumfang wird über das Erforderliche hinaus erhöht. Dies muss der „Königsweg“ bleiben, da einige zusätzliche Freiheitsgrade „der Statistik immer gut tun“.
- Die Varianz ist bekannt oder wird über Mehrfachrealisierung einer einzigen Kombination getrennt ermittelt, z.B. im Zentralpunkt. Dies ist ebenfalls eine sinnvolle Alternative.
- Es werden teilfaktorielle Pläne verwendet, bei denen nicht alle  $2^k$  Faktorstufenkombinationen betrachtet werden. Dies ist nur dann sinnvoll, wenn die Auflösung des teilfaktoriellen Plans ausreichend für die gestellte Aufgabe ist, siehe auch Kapitel 8.1.
- Die Varianz wird bei der Auswertung durch Pooling geschätzt. Damit wird ein Vorgang bezeichnet, bei dem nicht-signifikante, meistens höhere Wechselwirkungen im Normalverteilungsplot der Effekte erkannt, aus dem Modell ausgeschlossen und somit als zusätzliche Freiheitsgrade zur statistischen Absicherung verwendet werden. Diese Vorgehensweise muss immer die letzte Wahl bleiben und ist mit äußerster Vorsicht zu verwenden, da ein subjektiver Einfluss und Möglichkeiten der Fehlinterpretation hier nicht von der Hand zu weisen sind, vgl. auch Kapitel 6.3.1.

#### Fazit:

Es ist sinnvoll, gerade so viele Faktoren zu berücksichtigen, dass die erforderliche Anzahl von Einzelversuchen realisiert werden kann und weitere Freiheitsgrade für die Bestimmung der Signifikanz verbleiben.



#### 4.8. Festlegung der Reihenfolge und Gruppierung

Ziel ist an dieser Stelle, die Störgrößen zu beherrschen. Dies kann geschehen durch:

- Ausschalten, entweder experimentell, indem diese konstant gehalten werden oder statistisch durch eine Zufallsreihenfolge oder
- Miterfassen, entweder experimentell durch Blockbildung oder statistisch durch eine Kovarianzanalyse.

Versuche können in Blöcke aufgeteilt werden, innerhalb derer die Faktorstufenkombinationen möglichst gleich häufig auftreten und die Zufallsschwankungen möglichst klein sind. Falls Unterschiede zwischen den Blöcken auftreten, können diese erkannt und eliminiert werden. Beispielsweise müssen Versuche, die nicht alle zeitnah realisiert werden können, in Blöcke aufgeteilt werden, um eventuelle Störeinflüsse zu erfassen.

Um zu verhindern, dass ein unbekannter Trend (z.B. sich systematisch ändernde Prüflingsabmessungen aufgrund von Werkzeugverschleiß) die Ergebnisse verfälscht, werden die Versuchsrealisierungen in einem Block in zufälliger Reihenfolge durchgeführt. Dieser Prozess wird Randomisieren genannt.

Allerdings kann im Einzelfall ein Niveauwechsel eines Faktors sehr aufwendig sein, z.B. bei Anlagenumbau oder langen Wartezeiten. Deshalb kann es erforderlich sein, die Versuche nach diesem Niveau zu sortieren. Man muss sich dabei Gedanken machen, wie Trends erkannt und vermieden werden können.

#### 4.9. Planung des Versuchsdurchlaufs und der Auswertung

Manchmal scheitern teure Versuche nur deswegen, weil eine für den Bediener der Versuchseinrichtung nicht offensichtliche Anforderung nicht explizit festgelegt oder nicht deutlich genug kommuniziert wurde. Beispielsweise ist die Reihenfolge der Versuche in einem DoE-Versuchsplan (inkl. Blockbildung und Zufallsreihenfolge) manchmal umständlich umzusetzen. Allen Beteiligten muss deutlich gemacht werden, dass die Reihenfolge zu beachten ist.

In diesem Zusammenhang geht es bei diesem Schritt darum:

- den **Versuchsdurchlauf je Prüfling zu planen**,

*ANMERKUNG: Dies schließt die Aufstellung eines Störfallplans mit ein, in dem das jeweilige Ereignis sowie geeignete Maßnahmen zur Beseitigung benannt werden. Dies sind beispielsweise die Erfassung, das gezielte Herunterfahren im sicheren Zustand und ein anschließendes Hochfahren der Versuchseinrichtung, und zwar so, dass nach Möglichkeit die Aussagekraft des Versuchs nicht beeinträchtigt wird.*

- die **erforderliche Dokumentation** bei der Durchführung der Versuche **festzulegen**,

*ANMERKUNG: Sie beinhaltet die Erfassung (ggf. durch Mitschriften) und Protokollierung*

- *des Versuchsablaufs inkl. auftretender Sonderereignisse,*
- *der jeweiligen Einstellung der Faktoren, z.B. durch Prüflingsbezeichnung und sonstiger Randbedingungen des Versuchs, sowie*
- *der Ergebnisse.*
- die **Archivierung der Prüflinge und Rückstellmuster zu planen**, was bei umfangreichen Versuchen einen gewissen logistischen Aufwand bedeutet,
- und alles zu **dokumentieren** und denjenigen zu **kommunizieren**, welche mit der Umsetzung der Versuche beauftragt wurden.

Spätestens jetzt muss die Frage beantwortet werden, wie die Versuchsergebnisse ausgewertet werden sollen, und ob der Versuchsplan mit der geplanten Auswertung kompatibel ist. Die statistische Versuchsplanung hat an dieser Stelle einen wesentlichen Vorteil, da auf ein standardisiertes Vorgehen bei der Auswertung zurückgegriffen werden kann. Die geplante Auswertung ist zu dokumentieren, insbesondere dann, wenn eine andere Stelle im Unternehmen damit beauftragt werden soll oder die Auswertung wesentlich später erfolgt. In diesen Fällen muss davon ausgegangen werden, dass Hintergrundinformationen zum Zeitpunkt der Auswertung bereits verloren gegangen sind, was zu Fehlern führen kann.



Die Zweckmäßigkeit des Versuchsplans im Hinblick auf das Versuchsziel ist ebenfalls zu überprüfen. Dazu müssen die erwarteten Versuchsergebnisse inkl. Aussagen über deren Signifikanz mit dem Ziel der Untersuchung verglichen werden.

#### 4.10. Planung der Versuchseinrichtungen und der Prüflinge

Bei diesem Schritt geht es darum, geeignete Versuchseinrichtungen auszuwählen oder, falls solche nicht vorhanden sind, zu konzipieren und zu entwerfen. Dabei muss die Versuchsfähigkeit einer solchen Einrichtung (z.B. durch Wahl geeigneter Toleranzen) im Vordergrund stehen und sichergestellt werden.

Darüber hinaus müssen die Prüflinge entworfen werden. Die Merkmale der Prüflinge sowie deren Anzahl ergeben sich zunächst aus dem Versuchsplan.

Zu berücksichtigen ist:

- Eine ausreichende Anzahl Prüflinge muss für den Versuchsplan, für Rückstellmuster und eine gewisse Reserve vorhanden sein. Falls zukünftige Untersuchungen absehbar sind, kann eine vorbeugende Herstellung aus derselben Material- und Herstellcharge sinnvoll sein, damit Störgrößen das Ergebnis nicht verfälschen können.
- Die Herstellbarkeit muss geklärt werden, insbesondere, ob alle Merkmalskombinationen, welche sich aus dem Versuchsplan ergeben, herstellbar sind. Dies ist oft schwierig, da es sich im gewissen Sinne um Grenzmuster handelt, welche entweder physikalisch nicht möglich, nicht relevant oder unsinnig sein können.
- Auch diejenigen Eigenschaften der Prüflinge, welche nicht als Faktoren untersucht werden, müssen festgelegt und dokumentiert werden. Es kann sich im Nachhinein vielleicht herausstellen, dass diese Eigenschaften einen deutlichen Einfluss haben und deshalb als Faktoren zu berücksichtigen sind. Beispielsweise kann sich herausstellen, dass die Oberflächenrauigkeit eines Erzeugnisses bei der Untersuchung - anders, als zunächst vermutet - eine Rolle spielt.
- Die Rückwirkung des Herstellungsverfahrens z.B. auf Oberfläche oder Eigenspannungen muss berücksichtigt werden.
- Eine Beschriftung der Prüflinge ist vorzusehen, damit Ergebnisse dokumentierbar bleiben. Kleinteile können z.B. in beschrifteten Behältern aufbewahrt werden.

#### 4.11. Aufwandsabschätzung

Die erforderlichen Ressourcen wie Zeit, Geld, Maschinen- und Personalkapazität für den so geplanten Versuch müssen abgeschätzt und mit den vorhandenen Ressourcen und Anforderungen verglichen werden. Oft wird es notwendig sein, den Versuchsplan anzupassen und einen gesunden Kompromiss zwischen Aufwand und Erkenntnisgewinn zu finden.

Man muss jedoch stets darauf achten, dass die Aussagefähigkeit der Untersuchung noch den Anforderungen entspricht. Manchmal ist erst eine kritische Mindestanzahl an Einzelversuchen erforderlich, bevor überhaupt ein aussagefähiges Ergebnis vorgelegt werden kann.



## 5. Versuchsdurchführung und Dokumentation

Bei der Durchführung und Dokumentation der Versuche kommen zwei unterschiedliche Vorgehensweisen zum Einsatz, je nachdem ob für die Ermittlung der Ergebnisse reale physische Versuche oder Computereperimente vorgesehen wurden. Letzteres gewinnt immer mehr an Bedeutung wenn ein validiertes Modell des untersuchten Systems vorliegt, welches zwar das Verhalten des Systems auf einer tieferen Ebene komplett beschreibt, bei dem aber die Wirkung diverser Einflüsse auf Grund der Komplexität des Modells nicht offensichtlich ist und erst in einer Simulation ermittelt werden muss.

### 5.1. Physische Versuche

Bei realen physischen Versuchen muss zunächst die Versuchseinrichtung, so wie in Kapitel 4.10 geplant aufgebaut und anschließend für die Durchführung der Versuche vorbereitet werden. Dies hat sich in der Praxis oft als besonders zeitraubend erwiesen, deshalb muss möglichst frühzeitig damit begonnen werden.

Die Vorbereitung umfasst vor allem die Überprüfung der Versuchsfähigkeit der Einrichtung und deren Kalibrierung. Als Kalibrierung wird dabei der Vergleich der Ergebnisse eines mithilfe der Einrichtung durchgeführten Versuchs mit einer bekannten Referenz bezeichnet. Dabei muss sichergestellt werden, dass die Abweichung zur Referenz im Hinblick auf die Zielsetzung des Versuchs in akzeptablen Grenzen liegt.

Insbesondere für jene Faktoren, welche direkt an der Prüfeinrichtung eingestellt werden, muss die korrekte Funktionsweise der Einstellung überprüft werden. Dies ist besonders dann kritisch, wenn die Einstellung später im Versuch nicht durch online-Messung der tatsächlichen Werte überprüft werden kann. Falls eine solche Messung durchgeführt wird, muss die Messprozessfähigkeit überprüft bzw. sichergestellt werden. Gleiches gilt für die Messeinrichtungen, mit deren Hilfe die Zielgrößen ermittelt werden. Näheres zu diesem Thema findet sich z.B. in [15,16].

Parallel zur Vorbereitung der Prüfeinrichtung kann mit der Herstellung der Prüflinge begonnen werden. Die meisten Faktorkombinationen des Versuchsplans werden durch entsprechende Merkmale der Prüflinge realisiert, deshalb ist bei der Herstellung eine besondere Sorgfalt gefragt. Dies muss auch mit dem Hersteller der Prüflinge besprochen werden. Merkmale wie z.B. Maße, Oberfläche oder mechanische Eigenschaften müssen nach der Herstellung kritisch überprüft und dokumentiert werden. Damit werden Versuche mit Prüflingen vermieden, die sich im Nachhinein als untauglich für die Untersuchung erweisen.

Nachdem die Einrichtung und die Prüflinge hergestellt und überprüft wurden, können die Versuche wie in Kapitel 4 geplant durchgeführt werden. Der konkrete Ablauf inkl. eines Störfallplans sowie die erforderliche Dokumentation wurde bereits, wie in Kapitel 4.9. beschrieben, festgelegt.

Konsequenzen aus nicht vorgesehenen Ereignissen wie Störfällen müssen für zukünftige Untersuchungen gezogen werden.

### 5.2. Computereperimente

Falls ein validiertes Modell des untersuchten Systems vorliegt (z.B. ein strukturmechanisches FE-Modell eines komplexen Erzeugnisses), können alternativ auch „Versuche“ am Modell durchgeführt werden, was üblicherweise als Computereperiment bezeichnet wird. Die einzelnen Versuchsrealisierungen stellen in diesem Fall Simulationen dar, deren Erkenntnisse auf das reale System übertragen werden können.

Der Vorteil von Computereperimenten liegt darin, dass in der Regel mehr Simulationen als reale Versuche durchgeführt werden können. Allerdings sind Simulationen insbesondere an komplexen Modellen ebenfalls kostspielig und zeitaufwendig. Deshalb ist es sinnvoll, hier ebenfalls mit den Ansätzen der Versuchsmethodik Ressourcen zu sparen.

Das generelle Vorgehen unterscheidet sich nicht von dem eines realen physischen Versuchs: die Versuchspläne werden wie in Kapitel 4 verdeutlicht aufgestellt und wie in Kapitel 6 beschrieben ausgewertet. Die Ergebnisse liefert kein physisches Experiment, sondern ein Simulationstool, genannt Solver.

Ein wichtiger Unterschied zu den realen physischen Versuchen besteht darin, dass Computereperimente keinen statistischen Fehler aufweisen, dies muss bei der Wahl des



Versuchsplantyps berücksichtigt werden. Allerdings können hier nicht nur deterministische, sondern auch stochastische Modelle verwendet werden, bei denen das Auftreten von Störfaktoren durch überlagerte Zufallswerte simuliert werden kann, so dass die Anwendung statistischer Ansätze durchaus Sinn ergeben kann. Darüber hinaus können streuende Faktoren in Robustheits- und Zuverlässigkeitsanalysen durch stochastische Versuchspläne wie z.B. Plain Monte Carlo berücksichtigt werden.

Typische Anwendungen bestehen aus drei Schritten:

1. Durch Sensitivitätsuntersuchungen werden die relevanten unter vielen Einflussgrößen ermittelt.
2. Eine modellbasierte Optimierung der Zielgrößen unter Berücksichtigung der relevanten Einflussgrößen und weiterer Nebenbedingungen wird durchgeführt.
3. Eine Robustheitsanalyse des optimalen Designs wird durchgeführt. Dabei wird die Frage beantwortet, wie sich kleine, meist zufällige und unvermeidbare Abweichungen von den optimalen Werten der Faktoren auf die Zielgrößen auswirken und welcher Anteil von allen möglichen Kombinationen von Merkmalen noch die Anforderungen des Designs erfüllt. Analog dazu kann auch eine Zuverlässigkeitsanalyse durchgeführt werden, bei der ebenfalls der Anteil der zulässigen unter allen möglichen Merkmalskombinationen ermittelt wird. Der Unterschied zur Robustheitsanalyse besteht lediglich darin, dass dort sehr wahrscheinliche und bei der Zuverlässigkeitsanalyse sehr unwahrscheinliche Zustände untersucht werden.

Durch diese Vorgehensweise können Fragestellungen untersucht werden, welche durch reale Versuche nicht zu bewältigen sind. Anwender berichten z.B. über Gewichtsoptimierung eines kompletten Schiffs mit 30.000 Variablen [18].



## 6. Auswertung

Ziel der Auswertung ist die quantitative Bestimmung der Modellgleichung, welche die Abhängigkeit zwischen Faktoren und Zielgrößen beschreibt, sowie deren Validierung.

### 6.1. Plausibilitätsprüfung

Für eine erste Plausibilitätsprüfung ist zunächst eine einfache Darstellung der wesentlichen Versuchsergebnisse d.h. der Zielgrößen in Abhängigkeit der Faktoren sinnvoll. Hier müssen die verwendete Versuchseinrichtung, die Messtechnik sowie eventuelle Besonderheiten im Versuchsablauf aufgeführt werden. So fallen bei der Diskussion der Ergebnisse im Team schnell Unstimmigkeiten auf, welche eventuell näher untersucht werden müssen. Bei mehr als 2 Faktoren wird dies allerdings schwierig. Hier kann eventuell eine parametrische Darstellung oder eine Darstellung lediglich der Versuchspunkte im Versuchsraum (ohne Zielgrößen) helfen.

Aus dieser ersten Visualisierung muss auf jeden Fall die Abdeckung des Versuchsraums ersichtlich werden. Hintergrund ist die Tatsache, dass nur systematisch aufgestellte Versuchspläne den Versuchsraum gut abdecken und robust gegen Fehldeutungen sind, nicht jedoch Pläne welche nachträglich in vorhandene Daten mit unklarer Qualität „hineininterpretiert“ werden. Darüber hinaus muss die Vergleichbarkeit und Aktualität historischer Daten sichergestellt werden, bevor mit der Auswertung begonnen werden kann.

Es muss analysiert werden, ob Auffälligkeiten, wie z.B. unplausible Werte der Zielgröße zu finden sind. Generell muss die Ursache für solche Auffälligkeiten gesucht werden um ggf. das Ergebnis korrigieren zu können. Am einfachsten gelingt dies, falls es sich hierbei um Übertragungs- oder Tippfehler handelt. Manchmal können auch einfache Ablesefehler korrigiert werden. Falls die Ursache unklar bleibt oder ein Faktor falsch eingestellt wurde, muss der entsprechende Einzelversuch wiederholt werden. Ist dies nicht möglich wird dieser Einzelversuch aus der weiteren Analyse ausgeschlossen, denn Ausreißer können das Ergebnis der Untersuchung maßgeblich verfälschen und dürfen nicht verwendet werden. Dies muss allerdings mit größter Vorsicht geschehen, insbesondere dann, wenn nur eine Versuchsrealisierung vorliegt. In diesem Fall könnte ein vermeintlicher Ausreißer auf einen unvermuteten Effekt hindeuten. Statistische Ausreißertests können bei der Entscheidung helfen.

### 6.2. Bestimmung der Modellgleichung

Durch Darstellung der Vorgehensweise anhand zweier einfacher Beispiele wird zunächst in die Thematik eingeführt. Anschließend wird das allgemeine Vorgehen mittels Regression erläutert. Mathematische Hintergründe sind im Anhang zu finden.

#### 6.2.1. Beispiel Einfaktorversuch auf 2 Stufen

Aufgabe: Die Längenausdehnung einer Legierung soll experimentell bestimmt werden. Es werden mehrere Versuchsrealisierungen bei 2 Temperaturen durchgeführt. Tabelle 6.1 enthält die Mittelwerte der Ergebnisse der Realisierungen.

Versuch	$T, ^\circ C$	$L, cm$
1	25	100,04
2	100	100,16

**Tabelle 6.1:** Längenausdehnung einer Legierung

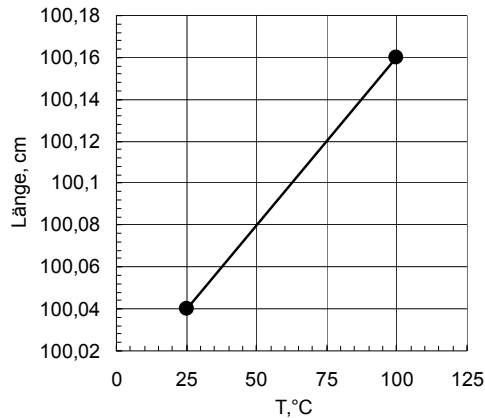
Man geht davon aus, dass zwischen der Längenänderung und der Temperatur ein linearer Zusammenhang besteht und möchte die Gleichung der Geraden berechnen, um beliebige Zwischenwerte rechnerisch bestimmen zu können. Diese lautet:

$$L = A_0 + A_1 T \quad (6.1)$$

Durch eine Koordinatentransformation wird das Wertepaar  $(T_1, T_2)$  formal in  $(-1, +1)$  überführt. Die Transformationsgleichung lautet dabei:

$$x = \frac{2}{T_2 - T_1} (T - T_2) + 1. \quad (6.2)$$





**Bild 6.1:** Längenausdehnung einer Legierung

Dabei kann leicht überprüft werden, dass  $x(T=T_1)=-1$  und  $x(T=T_2)=+1$ . Im transformierten Koordinatensystem lautet die Geradengleichung:

$$L = a_0 + a_1 x . \quad (6.3)$$

Aus Gleichung (6.3) und den beiden Versuchsergebnissen folgt:

$$\text{für } x = +1 : 100,16 = a_0 + a_1 , \quad (6.4)$$

$$\text{für } x = -1 : 100,04 = a_0 - a_1 . \quad (6.5)$$

An dieser Stelle wird der Grund für die Koordinatentransformation offensichtlich; die Koeffizienten  $a_0$  und  $a_1$  sind nämlich durch Addition bzw. Subtraktion der Gleichungen (6.4) und (6.5) sehr leicht zu berechnen:

$$a_0 = \frac{L_2 + L_1}{2} = \frac{100,16 + 100,04}{2} = 100,1 \quad a_1 = \frac{L_2 - L_1}{2} = \frac{100,16 - 100,04}{2} = 0,06 . \quad (6.6)$$

Allgemein gilt also, dass der Koeffizient  $a_0$  der Mittelwert der beiden Längen, der Koeffizient  $a_1$  hingegen die halbe Differenz der Längen ist.

Die Differenz der Längen bei unterer und oberer Temperatur wird *Effekt* des Faktors Temperatur genannt. Damit kann der Koeffizient  $a_1$  gleichzeitig aus dem halben Effekt der Temperatur berechnet werden. Dies stellt eine wichtige Beziehung zwischen dem Effekt des Faktors und der Koeffizienten der Modellgleichung dar, mit der Berechnung der Effekte werden somit automatisch auch die Koeffizienten der Modellgleichung bestimmt und umgekehrt.

Im transformierten System lautet die Gleichung (6.3) der Geraden:

$$L = 100,1 + 0,06x . \quad (6.7)$$

Die Geradengleichung (6.1) im ursprünglichen System erhält man durch die Rücktransformation (6.2)

$$L = 100,1 + 0,06 \frac{T - 62,5}{37,5} = 100 + 0,0016 \cdot T . \quad (6.8)$$

### 6.2.2. Beispiel Zweifaktorversuch auf 2 Stufen

Dieses Beispiel soll anhand des Ohmschen Gesetzes die Vorgehensweise bei der Auswertung von Versuchen verständlich machen. Wir versetzen uns dazu in die Lage eines Experimentators, der den Zusammenhang zwischen Spannung, Strom und Widerstand mit Hilfe eines einfachen Versuchs ermitteln möchte. Gesucht ist

$$U = f(R, I) . \quad (6.9)$$

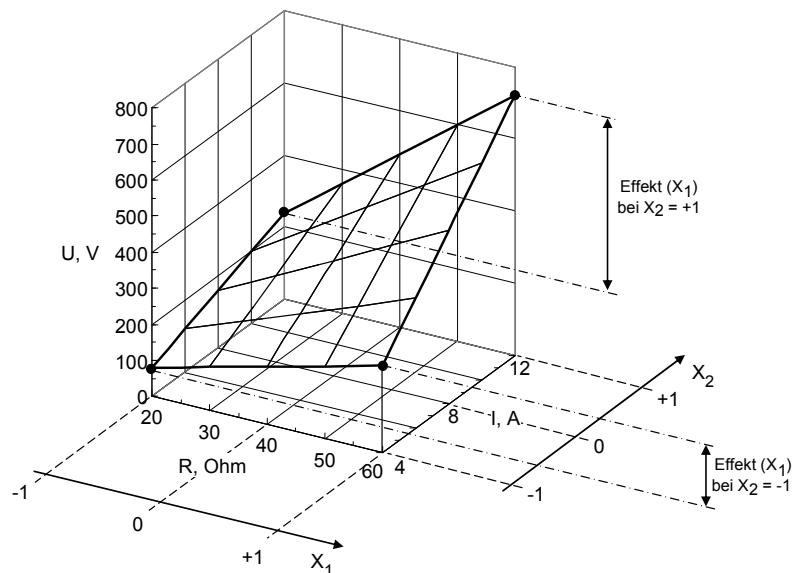
Wir nehmen an, er habe Einzelversuche mit mehrmaliger Versuchsrealisierung durchgeführt und folgende mittlere Spannung gemessen:





Versuch	$R, \text{ Ohm}$	$I, \text{ A}$	$U, \text{ V}$
1	20	4	80
2	60	4	240
3	20	12	240
4	60	12	720

**Tabelle 6.2:** Das Ohmsche Gesetz



**Bild 6.2:** Das Ohmsche Gesetz

Zur Ermittlung der Beziehung wird zunächst ein multilinearer Lösungsansatz für zwei Faktoren benutzt:

$$U = a_0 + a_1x_1 + a_2x_2 + a_{12}x_1x_2. \quad (6.10)$$

Folgende Transformation kann die Auswertung wieder vereinfachen:

$$x_1 = \frac{R - \frac{R_2 + R_1}{2}}{\frac{R_2 - R_1}{2}} = \frac{R - 40}{20}, \quad x_2 = \frac{I - \frac{I_2 + I_1}{2}}{\frac{I_2 - I_1}{2}} = \frac{I - 8}{4}. \quad (6.11)$$

Damit ergibt sich ein Gleichungssystem, aus dem die freien Koeffizienten des Lösungsansatzes leicht bestimmt werden können (rechts stehen die bei den einzelnen Versuchskombinationen erhaltenen Spannungen  $U$ ):

$$\begin{aligned} x_1 = -1, \quad x_2 = -1, \quad a_0 - a_1 - a_2 + a_{12} &= 80, \\ x_1 = +1, \quad x_2 = -1, \quad a_0 + a_1 - a_2 - a_{12} &= 240, \\ x_1 = -1, \quad x_2 = +1, \quad a_0 - a_1 + a_2 - a_{12} &= 240, \\ x_1 = +1, \quad x_2 = +1, \quad a_0 + a_1 + a_2 + a_{12} &= 720. \end{aligned} \quad (6.12)$$

Daraus ergibt sich:

$$\begin{aligned} a_0 &= \frac{80 + 240 + 240 + 720}{4} = 320, \\ a_1 &= \frac{720 - 240}{4} + \frac{240 - 80}{4} = 160, \end{aligned}$$



$$a_2 = \frac{720 - 240}{4} + \frac{240 - 80}{4} = 160, \tag{6.13}$$

$$a_{12} = \frac{720 - 240}{4} + \frac{80 - 240}{4} = 80.$$

Einsetzen von (6.13) in den Lösungsansatz (6.10) ergibt:

$$U = 320 + 160x_1 + 160x_2 + 80x_1x_2. \tag{6.14}$$

Durch Rücktransformation von (6.14) anhand der Gleichung (6.11) erhält man die Modellgleichung in den untransformierten Koordinaten:

$$U = 320 + 160 \frac{R-40}{20} + 160 \frac{I-8}{4} + 80 \left( \frac{R-40}{20} \right) \left( \frac{I-8}{4} \right) = RI. \tag{6.15}$$

In diesem konstruierten Beispiel ergibt sich zwangsläufig die richtige Lösung (Ohmsches Gesetz), weil mit der Multilinearform gerade der richtige Ansatz gewählt wurde. Ein komplizierterer funktionaler Zusammenhang mit Quotienten oder Potenzen der Einflussfaktoren wäre mit diesem Ansatz nur näherungsweise zu beschreiben.

### 6.2.3. Versuche mit k-Faktoren auf 2 Stufen

Falls ein vollfaktorieller Versuch mit  $k$ -Faktoren auf jeweils 2 Stufen mit insgesamt  $N$  Einzelversuchen und  $m=2^k$  Faktorstufenkombinationen durchgeführt wurde, können die Effekte und Wechselwirkungen sowie die Koeffizienten der Modellgleichung einfach anhand der Auswertematrix des Versuchsplans berechnet werden.

Wir betrachten dazu folgenden Versuchsplan:

Versuch Nr.	A	B	...	Ergebnisse $y_{ij}$	Mittelwert $\bar{y}_i$
1	-	-	...	$y_{11}, y_{12}, \dots, y_{1n}$	$\bar{y}_1$
2	+	-		$y_{21}, y_{22}, \dots, y_{2n}$	$\bar{y}_2$
...					
m	+	+		$y_{m1}, y_{m2}, \dots, y_{mn}$	$\bar{y}_m$

**Tabelle 6.3:** Vollfaktorieller Versuchsplan mit  $k$ -Faktoren auf jeweils 2 Stufen

Bei der mehrmaligen Ausführung jeder Faktorenkombination kann bei der Auswertung zunächst zeilenweise jeweils der Mittelwert der Einzelergebnisse berechnet werden, die entsprechende Spalte wurden in der obigen Matrix ergänzt.

Die multilineare Modellgleichung lautet in diesem Fall:

$$y = a_0 + \sum_{i=1}^k a_i x_i + \sum_{i=1}^{k-1} \sum_{j=i+1}^k a_{ij} x_i x_j + \dots + a_{12\dots k} (x_1 x_2 \dots x_k). \tag{6.16}$$

Diese enthält insgesamt  $2^k$ -Koeffizienten und genauso viele Terme: einen freien Term,  $k$ -Terme mit den Faktoren sowie gemischte Terme mit 2, 3 usw. Faktoren bis hin zu einem gemischten Term mit allen  $k$  Faktoren. Alle gemischten Terme repräsentieren 2-fach, 3-fach usw. Wechselwirkungen.

Der Effekt des  $j$ -ten Faktors kann einfach berechnet werden als

$$\Delta_j = \frac{2}{m} \sum_{i=1}^m \text{sgn}_{ij} \cdot \bar{y}_i, \tag{6.17}$$

wobei „sgn“ das Vorzeichen des entsprechenden Faktors in der Auswertematrix bedeutet. Wechselwirkungen können analog behandelt werden, deren Vorzeichenspalte wird als Produkt der entsprechenden Spalten der beteiligten Faktoren berechnet. Beispielsweise wird bei der 2-fach Wechselwirkung  $x_1x_2$  eine zusätzliche Spalte gebildet, bei der die Spalten A und B miteinander



multipliziert werden. Anschließend wird mit dieser Vorzeichenspalte nach obiger Vorschrift die Wechselwirkung berechnet.

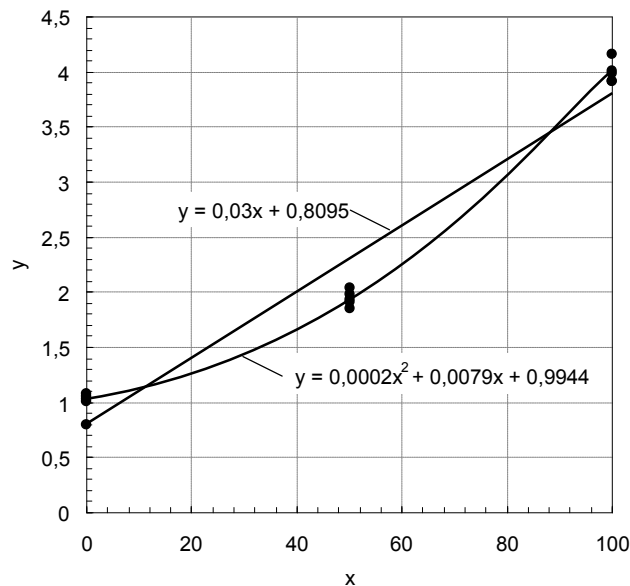
Die Koeffizienten in der Multilinearform ergeben sich dann zu:

$$a_0 = \frac{1}{m} \sum_{i=1}^m \bar{y}_i, \quad a_j = \frac{\Delta_j}{2}, \quad j = 1 \dots 2^k - 1. \quad (6.18)$$

#### 6.2.4. Regressionsanalyse

Für die Berechnung der Modellgleichung nach den bisherigen Ausführungen lagen genau so viele Versuchspunkte wie unbekannte Koeffizienten vor. So konnte ein Gleichungssystem mit  $m$  Gleichungen (aus den Versuchsergebnissen der  $m$  Faktorstufenkombinationen) und  $m$  Unbekannten (die Koeffizienten) aufgestellt werden, aus dem diese Koeffizienten bestimmt werden konnten. Falls bei einem Versuch mehr Versuchspunkte (nicht jedoch unbedingt Versuchsrealisierungen) als Koeffizienten vorliegen, würden mehr Gleichungen als Unbekannte vorhanden sein, so dass das Gleichungssystem überbestimmt und obiges Vorgehen nicht mehr möglich wäre. In diesem allgemeinen Fall ist die Durchführung einer Regressionsanalyse erforderlich. Die Durchführung von mehr Versuchen als für die Bestimmung der unbekannt Koeffizienten der Modellgleichung unbedingt erforderlich hat große Vorteile. Damit kann überprüft werden, ob der gewählte Modellansatz für die betrachtete Aufgabenstellung sinnvoll war oder nicht.

Bei einer Regression werden die Parameter eines bereits vorgegebenen mathematischen Ansatzes so angepasst, dass eine im Rahmen des gewählten Ansatzes bestmögliche Beschreibung der Versuchsdaten möglich ist. Dies geschieht durch die Methode der kleinsten Quadrate, bei der die Koeffizienten des Modells so bestimmt werden, dass die Summe der Abweichungen der Versuchspunkte von der Geraden minimiert wird. Auf eine formale Ableitung wird an dieser Stelle verzichtet, mehr mathematischer Hintergrund zu diesem Thema ist im Anhang zu finden.



**Bild 6.3:** Lineare und quadratische Regression eines Einfaktor-Versuchs auf 3 Stufen mit je 5 Realisierungen pro Stufe

Bei einem nichtlinearen Zusammenhang zwischen Ziel- und Einflussgrößen kann in den meisten Fällen versucht werden, durch geeignete Transformationen der Variablen das Problem durch eine quasi-lineare Regression zu beheben. Erwartet man z.B. eine Potenz-Anhängigkeit vom Typ

$$y = a_0 + a_1 t^2, \quad (6.19)$$

kann durch eine Substitution  $x = t^2$  die anfänglich nichtlineare Aufgabenstellung in einem linearen Zusammenhang überführt werden:

$$y = a_0 + a_1 x, \quad (6.20)$$

welcher durch eine lineare Regression behandelt werden kann.



Soll ein Ansatz durch Regression an Versuchsergebnisse angepasst werden, bei dem die Zielgröße von mehr als einer Einflussgröße abhängt, spricht man von einer multiplen Regression. Die Auswertung erfolgt nach denselben Prinzipien, wie bei der einfachen Regression.

Die Durchführung einer Regressionsanalyse ist ein komplexes Unterfangen, insbesondere bei mehr als einer Einflussgröße, deshalb wird an dieser Stelle die Verwendung geeigneter Softwaretools empfohlen. Näheres zu diesem Thema ist im Anhang zu finden.

### 6.3. Validierung der Modellgleichung

Im Allgemeinen werden die Zielgrößen eines Versuchs bei Mehrfachrealisierung (wiederholter Ausführung) aufgrund von Störeinflüssen mehr oder weniger stark variieren. Die auf der Basis der Versuchsergebnisse berechneten Effekte und Koeffizienten der Modellgleichung werden auch gewissen Schwankungen unterliegen und im Allgemeinen immer von Null verschieden sein, auch wenn der wahre aber unbekannte Effekt gleich Null ist. Somit stellt sich die Frage, ob die berechneten Effekte sich nur zufällig von Null unterscheiden oder nicht. Im ersten Fall wären diese nicht signifikant, falls die Effekte jedoch nicht zufällig von Null verschieden sind, werden sie als signifikant bezeichnet.

Ziel dieses Kapitels ist es zu zeigen, wie ein Effekt (bzw. Koeffizient der Modellgleichung) als signifikant oder nicht eingestuft werden kann.

#### 6.3.1. Einfache Signifikanzbewertung bei 2<sup>k</sup>-Plänen

Falls ein vollfaktorieller Versuch mit *k*-Faktoren auf jeweils 2 Stufen mit insgesamt *N* Einzelversuchen und *m* = 2<sup>k</sup> Faktorstufenkombinationen durchgeführt wurde, kann der Mittelwertvergleich für die Untersuchung der Signifikanz von Effekten und Wechselwirkungen verwendet werden, vgl. auch [1], S.99. Das Verfahren basiert auf der Tatsache, dass die Effekte als Differenz von jeweils zwei Mittelwerten eines Faktors berechnet werden können.

Wir betrachten im Folgenden dazu folgende Auswertematrix eines Versuchsplans:

Versuch Nr.	A	B	...	Ergebnisse $y_{ij}$	Mittelwert $\bar{y}_i$	Varianz $s_i^2$
1	-	-		$y_{11}, y_{12}, \dots, y_{1n}$	$\bar{y}_1$	$s_1^2$
2	+	-		$y_{21}, y_{22}, \dots, y_{2n}$	$\bar{y}_2$	$s_2^2$
...						
m	+	+		$y_{m1}, y_{m2}, \dots, y_{mn}$	$\bar{y}_m$	$s_m^2$

**Tabelle 6.4:** Auswertematrix eines Versuchsplans

Bei einer mehrmaligen Ausführung einer jeden Faktorenkombination kann bei der Auswertung zeilenweise je der Mittelwert und die Varianz der Einzelergebnisse berechnet werden, entsprechende Spalten wurden in der obigen Matrix ergänzt.

Der Effekt des *j*-ten Faktors kann berechnet werden als

$$\Delta_j = \frac{2}{m} \sum_{i=1}^m \text{sgn}_{ij} \cdot \bar{y}_i, \tag{6.21}$$

wobei „sgn“ das Vorzeichen des entsprechenden Faktors in der Auswertematrix bedeutet, vgl. auch Kapitel 6.2.3. Die Varianz

$$\bar{s}_y^2 = \frac{1}{m} \sum_{i=1}^m s_i^2, \tag{6.22}$$

welche als Mittelwert der Einzelvarianzen berechnet werden kann, stellt eine Schätzung für die Versuchsstreuung dar. Sie wird manchmal auch Pure Error genannt.

Da jeder Effekt die Differenz zweier Mittelwerte von jeweils *N*/2 Einzelwerten bildet, kann seine Standardabweichung geschätzt werden durch



$$s_{\Delta}^2 = \frac{4}{N} \overline{s_y^2}. \tag{6.23}$$

Die Signifikanz eines jeden Effekts kann wie im Kapitel 9.1.3 beschrieben nun durch einen Mittelwertvergleich auf Basis des t-Tests angegeben werden. Die zugehörige Testgröße kann berechnet werden als

$$t_j = \frac{\Delta_j}{s_{\Delta}}. \tag{6.24}$$

Ist  $t_j > t(f, 99\%)$ , so ist die Nullhypothese abzulehnen und der zugehörige Effekt signifikant (die Größe  $t(f, 99\%)$  ist im Anhang tabelliert). Der Freiheitsgrad beträgt  $f = N - 2^k$  ( $k$  - Anzahl Faktoren,  $N$  - Gesamtanzahl Einzelversuche). Eine abgestufte Beurteilung der Signifikanz des Effekts kann anhand Tabelle 6.5 erfolgen:

Untere Grenze	Parameter	Obere Grenze	Bewertung
	$t \leq$	$t_{f;99\%}$	nicht signifikant
$t_{f;99\%}$	$< t \leq$	$t_{f;99,9\%}$	signifikant
$t_{f;99,9\%}$	$< t$		hoch signifikant

**Tabelle 6.5:** Signifikanzbewertung durch t-Test

### Signifikanz bei Versuchen mit einfacher Realisierung

Falls ein Versuch mit Einfachrealisierung  $n=1$  vorliegt, ergeben sich Schwierigkeiten bei der Schätzung der Varianz der Einzelwerte und damit auch der Varianz der Effekte und deren Signifikanz. Die beste Möglichkeit einer solchen Situation zu begegnen ist, wenn die Streuung aus anderweitigen Untersuchungen bereits bekannt ist, oder über Mehrfachrealisierung einer einzigen Kombination (oft im Zentralpunkt) getrennt ermittelt werden kann.

Falls obige Möglichkeiten nicht bestehen, kann die Varianz durch sogenanntes Pooling geschätzt werden, wobei die zufälligen Effekte und Wechselwirkungen aus dem Modell entfernt werden. Hintergrund: Wenn die tatsächlichen Effekte und Wechselwirkungen Null sind, werden die berechneten zufälligen Effekte und Wechselwirkungen normalverteilt mit Mittelwert Null sein. Alle Ausreißer sind hier mit hoher Wahrscheinlichkeit tatsächlich auftretende Effekte oder Wechselwirkungen. Alle zufälligen Effekte können nun aus dem Modell ausgeschlossen werden und zur besseren statistischen Absicherung benutzt werden. Die „Zufälligkeit“ wird daran erkannt, dass diese Effekte und Wechselwirkungen auf einer Geraden im Wahrscheinlichkeitsplot liegen.

Diese Vorgehensweise muss immer die letzte Wahl bleiben und ist mit äußerster Vorsicht zu verwenden, denn ein subjektiver Einfluss und Möglichkeiten der Fehlinterpretation sind hier nicht von der Hand zu weisen. Die Empfehlung ist, diese Technik nur bei mehr als 5 Faktoren anzuwenden und dafür nur Wechselwirkungen höherer Ordnung zu benutzen, vgl. auch [1], S.107.

### 6.3.2. Validierung der Regression

Im Allgemeinen wird es erforderlich sein, die Koeffizienten der Modellgleichung durch eine Regressionsanalyse zu ermitteln. Die Qualität der Regressionsanalyse kann durch einige Auswertungen und Kennwerte analysiert werden, welche im Folgenden näher erläutert werden.

#### 6.3.2.1. Bewertung der Anpassungsgüte

Zunächst muss visuell überprüft werden, wie gut die berechnete Modellgleichung mit den Versuchsergebnissen übereinstimmt. Dazu ist es erforderlich, über mehr Versuchspunkte (nicht Versuchsrealisierungen) zu verfügen als freie Koeffizienten in der Modellgleichung vorhanden sind. Beispielsweise kann bei einem Einfaktorversuch auf zwei Stufen auch die Mitte des Intervalls als zusätzlicher Versuchspunkt erfasst werden. Dadurch kann die „Abweichung“ des Versuchsergebnisses von der Annahme einer linearen Beziehung überprüft werden. Falls lediglich zwei Versuchspunkte vorliegen ist eine solche Überprüfung nicht möglich, da 2 Punkte immer eine Gerade definieren.



Die Vorteile dieses zusätzlichen Versuchspunktes können auch so interpretiert werden, dass bei 3 Punkten die Auswertung durch eine lineare Regression möglich ist, bei der durch die im Folgenden beschriebenen Kennwerte auch die Güte der linearen Anpassung quantifiziert werden kann. Bei lediglich 2 Versuchspunkten wäre das nicht möglich, siehe dazu auch Kapitel 9.1.6.

Darüber hinaus ist es sinnvoll, die **Qualität der Regression** durch folgende Kennwerte anzugeben und quantitativ zu bewerten:

- Ein übliches Maß stellt der **RMS-Error**  $S_R$  (root mean square error) dar, auch Standardfehler bzw. Standardabweichung der Regression genannt. Dieser wird als Wurzel aus dem Mittelwert der Residuenquadrate berechnet. Residuen sind die Abweichungen der Zielgröße aus dem Versuch zu der Prognose der Modellgleichung

$$r_i = y_i - \hat{y}_i, \quad (6.25)$$

wobei  $\hat{y}_i = f(x_i)$  den Wert der mittels Regression berechneten Funktion an der Stelle  $x_i$  bezeichnet, im Unterschied zum Versuchswert  $y_i$  an dieser Stelle. Der RMS-Error ist somit eine skalare Größe, welche alle Residuen zusammenfassend charakterisiert und als mittlere Abweichung der gemessenen Zielgrößenwerte von der Regressionsfunktion zu verstehen ist. Ein niedriger Wert ist ein Indiz dafür, dass die Regressionsfunktion gut in der Lage ist, die Versuchsdaten zu beschreiben. Beim Standardfehler der Regression handelt es sich um eine absolute Größe, welche nicht direkt interpretiert werden kann, insbesondere dann nicht, wenn keine früheren vergleichbaren Untersuchungen existieren. Abhilfe schafft hier ein anderer wichtiger Kennwert, das sogenannte Bestimmtheitsmaß.

- Das **Bestimmtheitsmaß**  $R^2$  kann als jener Anteil der Streuung der Zielgröße interpretiert werden, welcher durch die Regression „erklärt“ wird, vgl. auch Kapitel 9.1.6 im Anhang. Hintergrund ist die Tatsache, dass Zusammenhänge, die nicht im Modell vorhanden sind (z.B. unbekannte Einflussgrößen) auch nicht durch das Modell erklärt werden können. Das Bestimmtheitsmaß nimmt Werte zwischen 0.0 und +1.0 ein. Somit ist ein  $R^2$  von oder nahe an 1.0 ein Zeichen dafür, dass der Regressionsansatz geeignet war die Versuchsdaten zu erklären.
- Alternativ dazu, kann auch der **Korrelationskoeffizient**  $R$  benutzt werden.  $R$  nimmt Werte zwischen -1.0 und +1.0 ein, wobei beide Grenzwerte nur dann auftreten, wenn die Daten komplett durch den Regressionsansatz erklärt werden können. Bei einer linearen Regression müssten alle Versuchspunkte beispielsweise komplett auf einer geraden Linie liegen. Falls Abweichungen davon auftreten muss beurteilt werden, wie groß diese sind und welche Ursachen sie haben können. Denn ein Korrelationskoeffizient von oder nahe an 0.0 heißt nicht automatisch, dass eine hohe Streuung vorliegt, sondern, dass der Regressionsansatz nicht in der Lage war ausreichend genau die Daten zu erklären. Dies kann an der hohen Streuung liegen oder aber am Ansatz selbst. So wird der Versuch eine im Versuchsraum symmetrisch liegende Parabel (also ein Ansatz 2. Ordnung) durch lineare Regression an eine Gerade anzunähern auch bei fehlender Streuung mit einem  $R$  von 0.0 quittiert, vgl. dazu auch Kapitel 9.1.6 im Anhang.

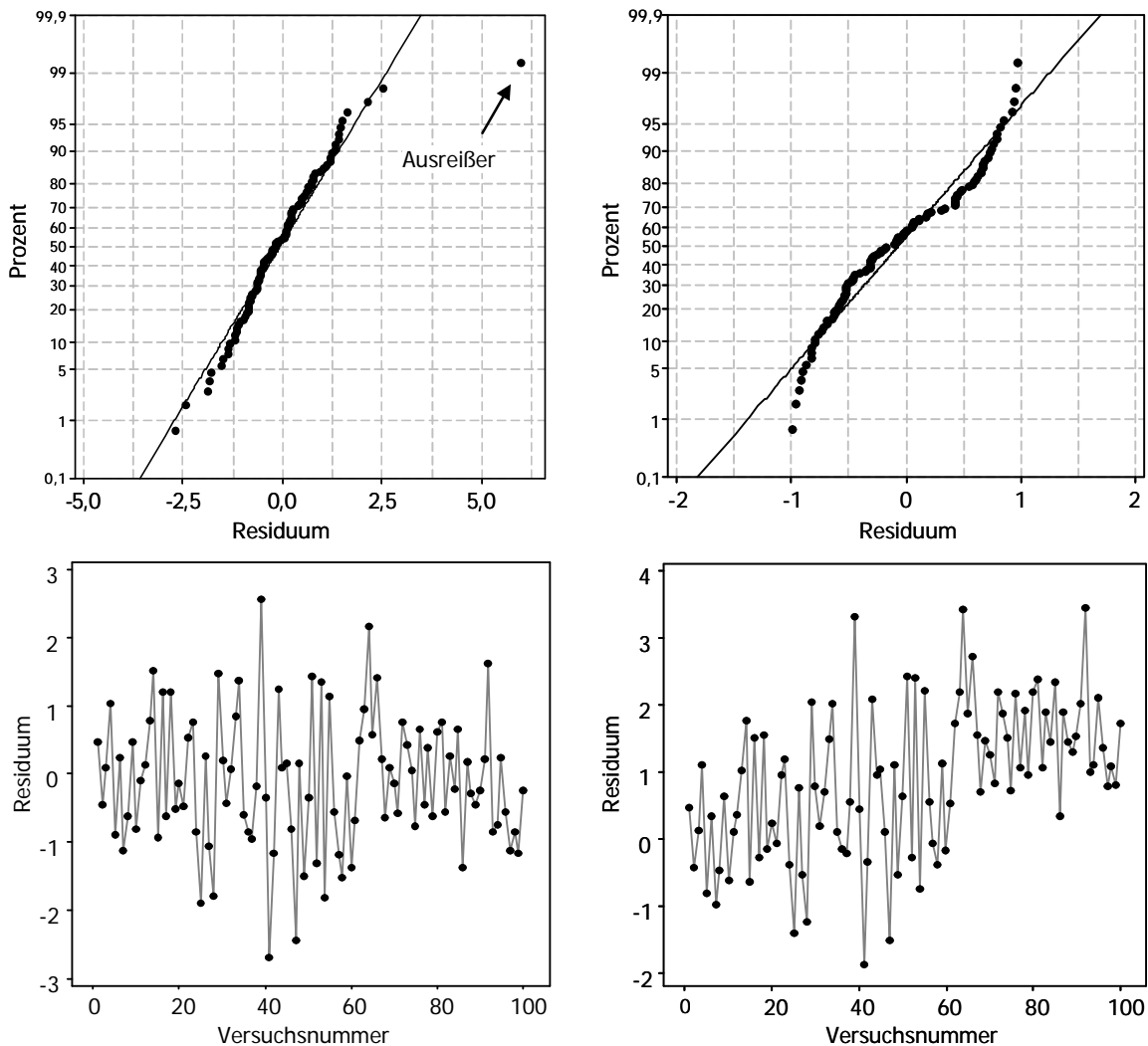
### 6.3.2.2. Residuenanalyse

Die Residuen geben wichtige Hinweise zur Güte der Anpassung der Modellgleichung an die Versuchsergebnisse. Sie charakterisieren die Abweichung zwischen den im Versuch ermittelten und den Anhand der Modellgleichung prognostizierten Ergebnissen. Falls der Modellansatz gut geeignet ist die Versuchsergebnisse zu beschreiben, müssen diese Abweichung nur *zufällig* sein. Dies muss nach jeder Regressionsanalyse überprüft werden, wobei eine grafische Darstellung immer von Vorteil ist, da sie unmittelbar prägnante Hinweise auf die diskutierten Zusammenhänge gibt. Gute Software erlaubt die Darstellung automatisch, vgl. auch Kap. 9.2. Folgende Auswertungen werden empfohlen:

- **Residuen im Normalverteilungsplot:** Diese Darstellung soll sicherstellen, dass die Residuen normalverteilt sind, denn in diesem Fall liegen sie im Wahrscheinlichkeitsnetz näherungsweise auf einer Geraden. Fallen deutliche Krümmungen auf, muss eventuell eine Transformation der Variablen vorgenommen werden. Die Normalverteilung kann ggf. auch durch geeignete statistische Tests untermauert werden.
- **Residuen in Abhängigkeit von der Versuchsnummer:** Diese Auswertung soll die Frage beantworten, ob ein Trend oder Sprung in den Versuchsergebnissen vorliegt. Auch hier muss keine systematische Abhängigkeit existieren, die Residuen müssen zufällig in einem Band streuen. Falls ein Trend existiert, können die Residuen z.B. systematisch zu- oder abnehmen. Dies wäre ein Anzeichen für die Existenz nicht erfasster Störgrößen (z.B. Werkzeugverschleiß),



welche das Ergebnis beeinflussen. Wichtig ist an dieser Stelle zu verstehen, dass ein solcher Trend nur dann in den Residuen sichtbar wird, wenn eine zufällige Reihenfolge bei der Versuchsdurchführung (Randomisierung) verwendet wurde. Andernfalls wird der Trend der Abhängigkeit der Zielgröße von den Faktoren hinzugerechnet und ist nicht mehr in den Residuen sichtbar.



**Bild 6.4:** Oben: Residuen im Normalverteilungsplot bei normal-verteilten Daten mit Ausreißer (links), bzw. bei gleich-verteilten Daten im Bereich  $[-1,+1]$  (rechts). Unten: Residuen in Abhängigkeit von der Versuchsnummer ohne Trend (links), bzw. mit Trend (rechts).

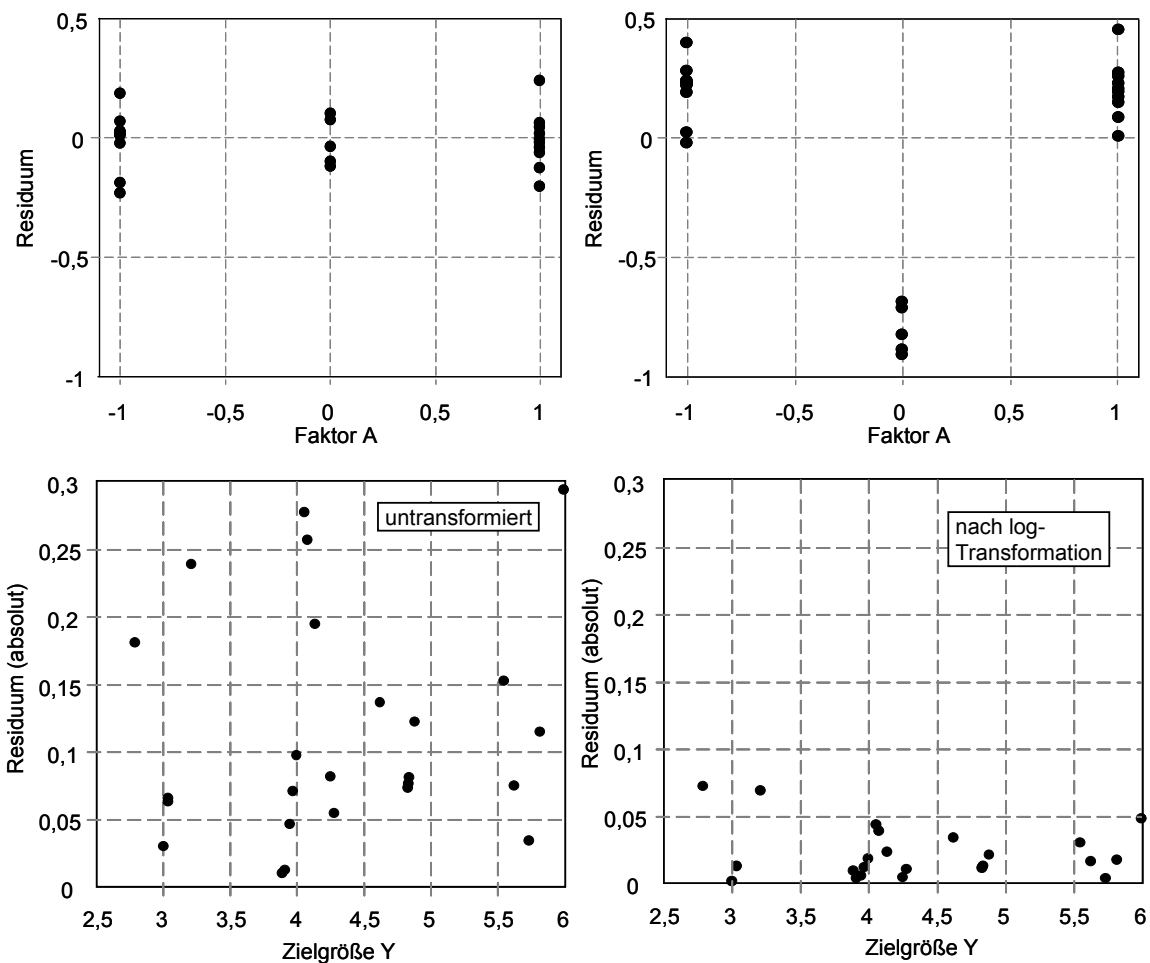
- **Residuen in Abhängigkeit von den Faktoren:** Diese Auswertung dient der Überprüfung, ob der verwendete Modellansatz konsistent mit den Versuchsergebnissen ist. Dies kann prinzipiell auch durch die Darstellung der Zielgröße in Abhängigkeit der Faktoren analysiert werden, beispielsweise kann die Validierung eines linearen Modellansatzes durch Vergleich mit Versuchsergebnissen in der Mitte des Bereichs erfolgen. Allerdings sind die Residuen normalerweise kleiner als die Zielgröße, so dass eine Darstellung der Residuen in Abhängigkeit von den Faktoren besser dafür geeignet ist. Dabei ist darauf zu achten, dass keine systematische Abhängigkeit existiert, sondern die Residuen zufällig in einem Band um die berechnete Zielgröße streuen.

Die Verteilung der Residuen auf der  $x$ -Achse spielt hingegen keine Rolle. Bei systematisch geplanten Versuchen werden die Residuen nur an den im Versuch verwendeten Stufen des Faktors vorliegen, bei ungeplanten Versuchen (oder Auswertung historischer Daten) werden die Faktorwerte beliebig sein.

- **Residuen in Abhängigkeit der berechneten Zielgröße:** Eine Voraussetzung für die Anwendung vieler statistischer Analysen ist die Forderung, dass die Standardabweichung nicht von der Zielgröße abhängt. Diese Voraussetzung kann überprüft werden, indem die Residuen gegenüber der berechneten Zielgröße aufgetragen werden. Falls diese nicht zufällig in einem Band liegen,



sondern ein Trend sichtbar ist (oft z.B. trichterförmige Zu- oder Abnahme) muss eine Transformation der Variablen (z.B. logarithmisch) in Erwägung gezogen werden.



**Bild 6.5:** Oben: Residuen in Abhängigkeit eines Faktors bei einem linearen Modell. Die Daten links bestätigen den linearen Verlauf. Die Daten rechts weisen auf einen nicht-linearen Verlauf hin. Unten: Residuen (Absolutwert) in Abhängigkeit von einer exponentiellen Zielgröße, ohne log-Transformation (links), mit log-Transformation (rechts).

### 6.3.2.3. Konfidenzintervall und Signifikanz von Regressionskoeffizienten

Die Koeffizienten des Regressionsmodells wurden anhand der Daten des Versuchs bestimmt, welche einer Zufallsstreuung unterliegen. Somit ist anzunehmen, dass die Koeffizienten selbst Zufallsvariablen sind. Man kann zeigen, dass unter bestimmten Voraussetzungen (siehe dazu Kapitel 9.1.6 im Anhang) die mittels der Methode der kleinsten Quadrate bestimmten Werte eine gute Schätzung der wahren Werte darstellen.

Mit dem Standardfehler des  $i$ -ten Regressionskoeffizienten  $S_i$  kann mittels der Technik des Mittelwertvergleichs ein zweiseitiger Vertrauensbereich zum Vertrauensniveau  $(1 - \alpha)$  angegeben werden, auf Details der Herleitung wird an dieser Stelle jedoch verzichtet, vgl. dazu auch Kapitel 9.1.6 im Anhang. Die Ergebnisse werden üblicherweise zusammen mit den Werten der Koeffizienten dargestellt, siehe Tabelle 6.6.

Die Vertrauensbereiche für den Mittelwert der Zielgröße  $\mu_Y(x)$  hängen von  $x$  ab und werden grafisch als „trompetenartige“ Umhüllende der Regressionskurve dargestellt. Die Signifikanz der berechneten Steigung kann analog zum Vorgehen beim t-Test beurteilt werden, indem eine Testgröße berechnet und mit tabellarischen Werten verglichen wird.

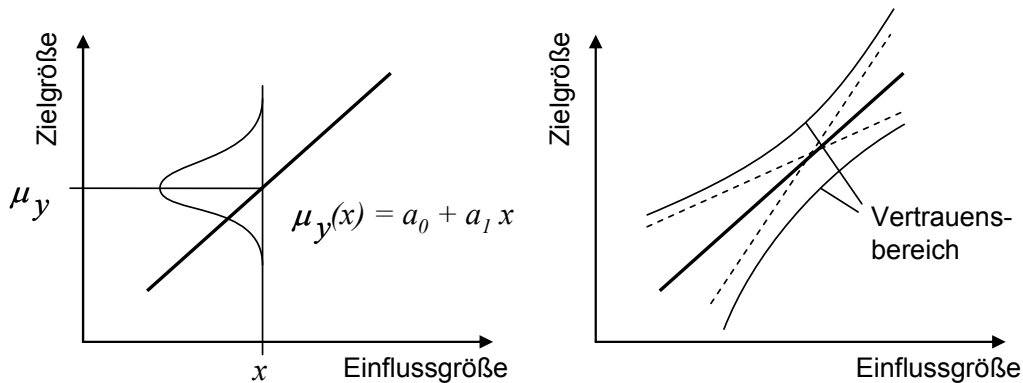




Variable	Koeffizient	Std. Fehler	t-Wert	p-Wert
$x_1$	$b_1$	$S_1$	$t_1$	$p_1$
...				

**Tabelle 6.6:** Koeffiziententabelle einer Regression

Sowohl die Darstellung des Vertrauensbereichs des Mittelwertes als auch die Angabe eines Vertrauensintervalls für die Steigung machen deutlich, dass die berechnete Gerade nur eine Schätzung der wahren aber unbekanntenen Gerade ist. Die Vertrauensbereiche stellen im Gewissen Sinne die Fülle möglicher Lagen der wahren Gerade dar. Bei der Interpretation muss beachtet werden, dass eine signifikante Steigung nicht den Beweis liefert, dass die Abhängigkeit auch linear ist, lediglich dafür, dass ein linearer Anteil existiert. Die beschriebenen Sachverhalte werden im folgenden Bild 6.6 verdeutlicht. Weitere Details zu diesem Thema finden sich in [1].



**Bild 6.6:** Vertrauensbereich des Mittelwertes bei einer linearen Regression (schematisch)

#### 6.3.2.4. Schrittweise Regression

Bei der Schrittweise- (oder Stufenweise-) Regression wird die Regressionsanalyse iterativ in mehreren Schritten wiederholt, wobei bei jedem Schritt Effekte und Wechselwirkungen, welche als nicht signifikant eingestuft werden von dem Modell entfernt und die verbliebenen Freiheitsgrade zur besseren statistischen Absicherung benutzt werden. Dies wird solange wiederholt, bis nur noch signifikante Effekte und Wechselwirkungen im Modell verbleiben, welches dann die einfachst-mögliche Form besitzt.

Zu empfehlen ist folgende Vorgehensweise, vgl. auch [19], S.74:

- zunächst muss die Signifikanz höherer (z.B. quadratischer) Effekte überprüft werden, falls solche im Modell enthalten sind. Diese können bei Bedarf als erstes entfernt werden.
- Danach kommen höhere Wechselwirkungen an der Reihe.
- Als letztes wird die Signifikanz der Zweifachwechselwirkungen und der Effekte überprüft, wobei die hierarchische Struktur des Modells beachtet werden muss. Ein Effekt muss demnach nur dann entfernt werden, wenn auch alle Terme höherer Ordnung sowie alle seine Wechselwirkungen bereits entfernt wurden. Dies stellt die Invarianz des Modells gegen Reskalierung der Einflussgrößen sicher.

Bei jedem Schritt der Regression, müssen die Änderungen des Bestimmtheitsmaßes  $R^2$  verfolgt werden,  $R^2$  muss im Laufe der Regression steigen.

#### 6.3.3. Validierung durch Varianzanalyse

In den vorangegangenen Kapiteln wurde erläutert, wie durch eine auf den t-Test basierende Vorgehensweise des Mittelwertvergleichs für Pläne von  $k$ -Faktoren auf 2 Stufen die Signifikanz der Effekte beurteilt werden kann. Bei mehrstufigen Faktoren konnten im Rahmen einer Regressionsanalyse ebenfalls auf dem t-Test basierende Signifikanzaussagen gemacht werden, allerdings nur bei quantitativen Faktoren und quasi-linearen Abhängigkeiten. Die Varianzanalyse stellt



nun ein Hilfsmittel dar, um Signifikanzaussagen auch für jene Fälle machen zu können, bei denen die Technik des Mittelwertvergleichs versagt, typischerweise im Falle mehrerer Stufen eines qualitativen Faktors. Da es sich hierbei um ein allgemeingültiges Vorgehen handelt, das bei statistischer Software weit verbreitet ist, soll an dieser Stelle kurz darauf eingegangen werden. Weitere Hintergründe zur Varianzanalyse finden sich im Anhang und [27].

### 6.3.3.1. Einfache Varianzanalyse (ANOVA)

Einfach bedeutet in diesem Zusammenhang, dass mehrere Stufen eines einzigen Faktors betrachtet werden. Dieses Vorgehen kann für die Analyse eines Einfaktorversuchs auf mehreren Stufen verwendet werden. Als erster Schritt in einem multifaktoriellen Versuch wird zunächst überprüft, ob die Varianz der Ergebnisse in den Versuchszeilen sich signifikant von der Versuchsstreuung unterscheidet, d.h. ob irgendein Faktor einen Einfluss auf die Zielgröße hat.

Wir betrachten folgenden Versuchsplan mit  $n$  Versuchsrealisierungen auf  $m$  Stufen:

Stufe Nr.	Ergebnisse $y_{ij}$	Mittelwert $\bar{y}_i$	Varianz $s_i^2$
1	$y_{11}, y_{12}, \dots, y_{1n}$	$\bar{y}_1$	$s_1^2$
2	$y_{21}, y_{22}, \dots, y_{2n}$	$\bar{y}_2$	$s_2^2$
...			
m	$y_{m1}, y_{m2}, \dots, y_{mn}$	$\bar{y}_m$	$s_m^2$

**Tabelle 6.7:** Versuchsplan mit  $n$  Versuchsrealisierungen auf  $m$  Stufen.

Die einfache Varianzanalyse kann in drei Stufen durchgeführt werden:

1. Berechnung der mittleren Streuung der Einzelwerte (diese Größe ist ein Maß für die Versuchsstreuung):

$$\bar{s}_y^2 = \frac{1}{m} \sum_{i=1}^m s_i^2. \quad (6.26)$$

2. Berechnung der Varianz der Mittelwerte

$$s_{\bar{y}}^2 = \frac{1}{m-1} \sum_{i=1}^m (\bar{y}_i - \bar{\bar{y}})^2 \quad \text{mit} \quad \bar{\bar{y}} = \frac{1}{m} \sum_{i=1}^m \bar{y}_i \quad (\text{Gesamtmittelwert}). \quad (6.27)$$

3. F-Test mit der Prüfgröße  $F = \frac{ns_{\bar{y}}^2}{s_y^2}$ . (6.28)

Falls  $F$  größer ist als der Schwellenwert  $F(f_1, f_2, 99\%)$  der im Anhang tabellierten F-Verteilung mit den Freiheitsgraden  $f_1=m-1$ ,  $f_2=(n-1)m$ , so liegt ein signifikanter Unterschied der Versuchsergebnisse vor ( $n$  = Anzahl der Realisierungen pro Stufe,  $m$  = Anzahl Stufen).

Die Signifikanzaussagen können analog zum Mittelwertvergleich vom Intervall abhängig gemacht werden, in dem sich die berechnete Prüfgröße befindet:

Untere Grenze	Prüfgröße	Obere Grenze	Bewertung
	$F \leq$	$F(f_1, f_2, 99\%)$	nicht signifikant
$F(f_1, f_2, 99\%)$	$< F \leq$	$F(f_1, f_2, 99.9\%)$	signifikant
$F(f_1, f_2, 99.9\%)$	$< F$		hoch signifikant

**Tabelle 6.8:** Signifikanzbewertung durch F-Test



### 6.3.3.2. Faktorielle Varianzanalyse

Mit der faktoriellen Varianzanalyse wird für jeden Faktor eines multifaktoriellen Versuchsplans entschieden, ob er einen signifikanten Einfluss auf das Versuchsergebnis hat. Folgende Vorgehensweise bietet sich an:

1. Berechnung der Stufenmittelwerte des Faktors  $\bar{y}_i$ ,  $i=1..m$ . In der Mittelwertbildung gehen alle Versuchswerte mit ein, bei denen der ausgewertete Faktor die jeweilige Stufe aufweist. Alle übrigen Faktoren sind somit mit Werten über alle Stufen in der Mittelwertbildung enthalten.
2. Berechnung der Versuchsstreuung  $\overline{s_y^2}$ . Dies ist die mittlere Varianz der Versuchspunkte mit Mehrfachrealisierung.
3. Berechnung der Varianz  $s_y^2$  der nach Punkt 1 berechneten Mittelwerte des Faktors.
4. F-Test mit der Prüfgröße  $F = \frac{ns_y^2}{s_y^2}$ . (6.29)

Falls die Prüfgröße  $F$  größer ist als der Schwellenwert  $F(f_1, f_2, 99\%)$  der im Anhang tabellierten F-Verteilung mit den Freiheitsgraden  $f_1=m-1$ ,  $f_2$ , so hat der Faktor einen signifikanten Einfluss auf das Versuchsergebnis, mit  $n$  = Anzahl der Versuchswerte pro Stufe,  $m$  = Anzahl Stufen.

Die Signifikanzaussagen können von dem Intervall abhängig gemacht werden, in dem sich die berechnete Prüfgröße befindet, so wie in Tabelle 6.8 dargestellt.

### 6.3.3.3. Varianztabelle, multiple Varianzanalyse

Die Varianzanalyse kann als Zerlegung der Summe der quadratischen Abweichungen (S.d.q.A.) der Einzelwerte vom Gesamtmittelwert in einem Anteil  $Q_A$  und einem Rest  $Q_R$  aufgefasst werden, wobei der erste Anteil durch die Stufen des Faktors erklärt wird und der verbleibende Anteil für die Abschätzung der Zufallsstreuung verwendet werden kann. Eine Ableitung der Zerlegung findet sich im Anhang Kapitel 9.1.5, darauf soll hier nicht weiter eingegangen werden. Das Ergebnis der Zerlegung wird üblicherweise in einer Tabelle zusammengefasst, welche folgendes Aussehen hat:

	Freiheitsgrad	S.d.q.A.	Varianz	F	p-Wert
Faktor $A$	$f_A = m-1 = f_1$	$Q_A$	$s_A^2 = \frac{Q_A}{f_A} = n \cdot s_y^2$	$F_A = \frac{s_A^2}{s_{Rest}^2}$	$p_A$
...	...	...	...	...	...
Rest	$f_R = m(n-1) = f_2$	$Q_R$	$s_{Rest}^2 = \frac{Q_R}{f_R} = \overline{s_y^2}$		
Gesamt	$f = f_A + \dots + f_R$	$Q = Q_A + \dots + Q_R$			

**Tabelle 6.9:** MANOVA-Tabelle ( $n$  = Anzahl der Versuchswerte pro Stufe,  $m$  = Anzahl Stufen)

Die Signifikanzaussagen können von dem Intervall abhängig gemacht werden, in dem sich die berechnete Testgröße  $F_A$  befindet, so wie in Tabelle 6.8 dargestellt. Bei manchem statistischen Tool wird zusätzlich zur Testgröße  $F_A$  eine Größe  $p_A$  zur Beurteilung der Signifikanz angegeben, bei der gilt

$$F_A = F(1 - p_A, f_1, f_2). \tag{6.30}$$

Diese Größe wird als „Irrtumswahrscheinlichkeit“ (Fehler 1. Art) interpretiert: je kleiner diese Wahrscheinlichkeit ist, desto signifikanter wird der Einfluss des Faktors sein. Die Bewertung erfolgt analog zu Testgröße  $F_A$ : ein signifikantes Ergebnis liegt vor, wenn  $p_A \leq 1\%$ , ein hoch-signifikantes, wenn  $p_A \leq 0,1\%$ .

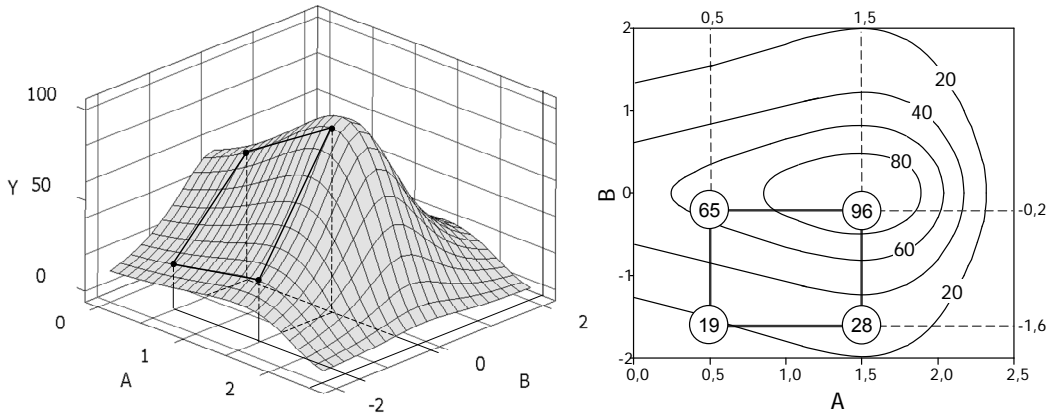
Wie in Tabelle 6.9 bereits angedeutet, kann die Zerlegung auch für mehrere Einflussfaktoren und Wechselwirkungen durchgeführt werden, man spricht von einer multiplen Varianzanalyse. Die mathematische Behandlung ist hier deutlich aufwendiger, lässt sich aber gut formalisieren, so dass diese Analyse in den meisten statistischen Softwarepaketen mit enthalten ist.



### 6.4. Grafische Darstellung

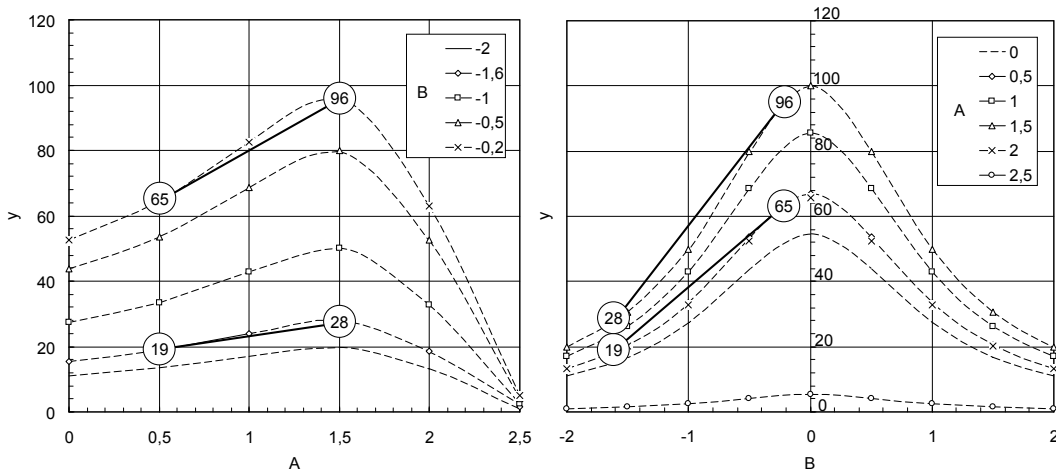
Die berechnete Modellgleichung und damit auch die Ergebnisse der Versuchsauswertung werden üblicherweise in einer grafischen Darstellung verdeutlicht. Zielgrößen, die von zwei oder mehr Faktoren abhängig sind, werden meist in Form von Höhenlinien, Kennfeldern und Kurvenscharen dargestellt. Dieser Zusammenhang soll anhand der nachfolgenden Bilder veranschaulicht werden.

Im Bild 6.7 links ist die 3D-Darstellung eines „Berges“ wiedergegeben. Die Darstellung im Bild 6.7 rechts zeigt in einer „Ansicht von Oben“ die Höhenlinien dieses „Berges“ wie sie auf topographischen Karten zu finden sind. Im gezeigten Beispiel entspricht der Sprung von einer Linie zu einer benachbarten Linie einem definierten Höhenunterschied von 20 Einheiten. Dicht beieinanderliegende Höhenlinien bedeuten einen steilen Anstieg in einer Richtung senkrecht zu den Höhenlinien. Bleibt man auf einer geschlossenen Höhenlinie, so bewegt man sich - bildlich gesprochen - in konstanter Höhe um den Berg herum.



**Bild 6.7:** Funktionsverlauf in einer 3D- (links) und 2D-Darstellung (rechts)

Löst man sich nun von dem Bild des „Berges“ und betrachtet statt seiner Höhe allgemein eine Funktion  $y$ , die von den Variablen  $A$  und  $B$  abhängt:  $y = f(A, B)$ . Diese Gleichung kann als Modell für eine Zielgröße  $y$  dienen, deren Wert durch die Einstellung der Einflussfaktoren  $A$  und  $B$  bestimmt wird. Jeder Einstellung  $(A, B)$  entspricht dann ein Wert  $y$ .



**Bild 6.8:** Parametrische Darstellung des Funktionsverlaufs aus Bild 6.7

Das Bild 6.8 zeigt eine weitere Darstellungsmöglichkeit für diese Ergebnisse. Die Zielgröße  $y$  ist als Funktion von  $A$  mit  $B$  als festem Parameter aufgetragen, man spricht von einer „parametrischen Darstellung“. Die punktierten Kurven im Bild 6.8 links stellen jeweils die Funktion  $y$  bei festem  $B$  dar. Sie sind sozusagen die Schnittlinien eines senkrechten Schnittes durch die Bergoberfläche bei festem  $B$ . Analog dazu stellen die punktierten Linien im Bild 6.8 rechts jeweils die Funktion  $y$  bei festem  $A$  dar.

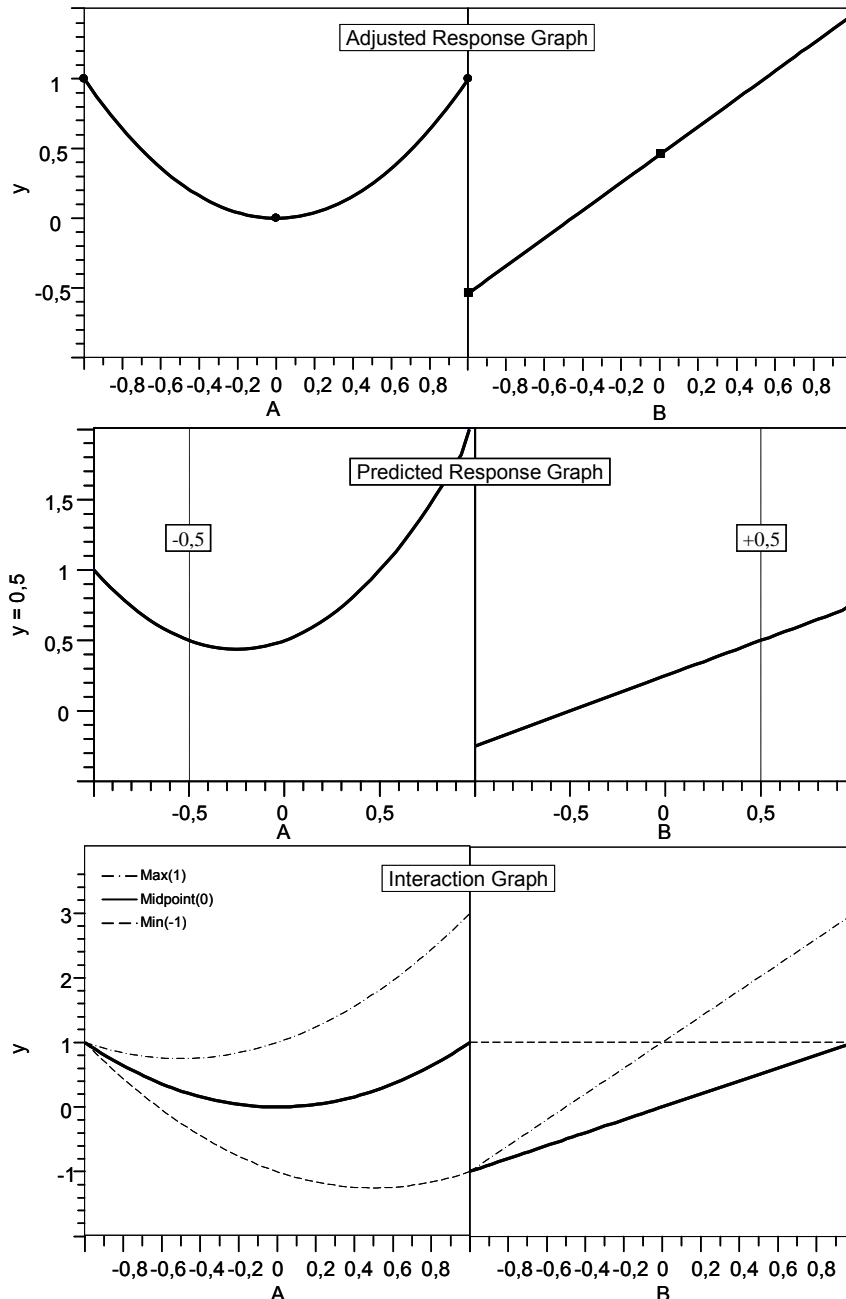
Prinzipiell macht man sich diese Methoden auch zur Darstellung von Versuchsergebnissen zunutze. Typische Darstellungen sind im Bild 6.9 wiedergegeben. Man unterscheidet drei Darstellungen:

- Beim **Adjusted Response Graph** wird der mittlere Effekt eines Faktors dargestellt. Damit ist gemeint, dass die Effekte aller anderen Faktoren durch Mittelwertbildung eliminiert wurden. Falls



der dargestellte Faktor signifikante Wechselwirkungen mit anderen Faktoren aufweist, kann diese Darstellung irreführend sein, da der dargestellte Verlauf stark von der jeweiligen Einstellung der übrigen Faktoren abhängen kann.

- Der **Predicted Response Graph** ist eine parametrische Darstellung der Zielgröße in Abhängigkeit eines Faktors bei fest vorgegebenen Einstellungen aller anderen Faktoren.
- Beim **Interaction Graph** wird jeder Faktor bei jeder Stufe aller übrigen Faktoren dargestellt. Signifikante Wechselwirkungen können hier anhand eines nicht-parallelen Verlaufs der Kurvenscharen erkannt werden. Der nicht-parallele Verlauf bedeutet, dass der Effekt eines Faktors signifikant von der Einstellung eines anderen Faktors anhängt, was als Wechselwirkung bezeichnet wird.



**Bild 6.9:** Typische Darstellungen der Versuchsergebnisse der Funktion  $y = A^2 + B + AB$

Für die Kommunikation mit Kunden wird der „predicted response graph“ empfohlen. Einzelne Effekte können hier diskutiert werden, ohne zu tief in DoE - spezifische Details einzutauchen.

Diese Darstellungsform ist auch dann verwendbar, wenn es sich bei einem (oder mehreren) der untersuchten Faktoren nicht um eine quantitative, einstellbare Variable handelt, sondern um eine



qualitative Variable mit festen Stufen (z.B. Material 1 - Material 2). Natürlich ist in diesem Fall eine Interpolation von Zwischenwerten nicht sinnvoll.

Die graphische Darstellung des Versuchsraumes ist bei drei Einflussfaktoren in Form eines Würfels möglich. Jeder Eckpunkt entspricht dabei einer Stufenkombination der Faktoren  $A$ ,  $B$  und  $C$ . Bei mehr als drei Faktoren können lediglich zwei- oder dreidimensionale Projektionen des mehrdimensionalen Versuchsraums dargestellt werden.

### 6.5. Interpretation der Ergebnisse

Die Ergebnisse müssen mit Fachexperten diskutiert, physikalisch interpretiert, und in Form eines Vergleichs mit dem erwarteten Ergebnis plausibilisiert werden. Unerwartete Ergebnisse sind immer ein Grund für weitere Nachforschungen und eine Erweiterung des Wissens, beispielsweise um unbekannte Einflussfaktoren oder Wechselwirkungen zu finden. Generell darf man keinem Ergebnis trauen, welches man nicht versteht.

Zunächst muss überprüft werden, welche Faktoren den größten signifikanten Einfluss auf die Zielgröße haben, dies sind die Faktoren mit den größten Effekten. Falls eine Veränderung (z.B. Optimierung) der Zielgröße erwünscht ist, müssen genau diese Faktoren dafür verwendet werden. Falls die Zielgröße jedoch möglichst konstant oder in engen Grenzen gehalten werden muss, müssen besondere Anstrengungen unternommen werden um genau diese Faktoren konstant zu halten oder deren Streuung zu begrenzen. Es ist allerdings zu betonen, dass die Effekte maßgeblich von der Wahl der Variationsbereiche der Faktoren abhängen. Insofern kann ein kleiner Effekt daran liegen, dass der Variationsbereich des Faktors zu klein war.

Größere Effekte äußern sich bei einem linearen Ansatz in einer größeren Neigung der Effekt-Geraden. Sofern sich die Faktoren additiv verhalten, ergibt sich ein paralleler Verlauf der Kurven eines Faktors bei allen Einstellungen der übrigen Faktoren. Hängt dagegen der Effekt eines Faktors von der Einstellung (der Stufe) eines anderen ab, so liegt eine *Wechselwirkung* dieser beiden Faktoren vor, sie verhalten sich nicht additiv. Die Effekt-Gerade des Faktors weist unterschiedliche (auch gegenläufige) Steigungen auf, je nach Einstellung der übrigen Faktoren.

Falls signifikante Wechselwirkungen ermittelt wurden bedeutet dies, dass alle in der Wechselwirkung vorkommenden Faktoren zusammen betrachtet werden müssen. Der Einfluss der Wechselwirkung kann dazu führen, dass sich der Effekt eines Faktors verstärkt, vermindert oder sogar umkehrt. Es dürfte darum intuitiv klar sein, dass eine Vernachlässigung der Wechselwirkung zu völlig falschen Ergebnissen bzw. Schlussfolgerungen führen kann. Darüber hinaus kann bei vorliegender starker Wechselwirkung  $AB$  der mittlere Effekt beider Faktoren  $A$  und  $B$  Null werden, obwohl jeder Faktor durchaus großen Einfluss besitzen kann. Bei Versuchsplänen mit mehr als 2 Faktoren können auch höhere Wechselwirkungen auftreten. Dreifach-Wechselwirkungen bedeuten z.B., dass eine Kombination von 2 Faktoren auf einen dritten Faktor einen Einfluss hat etc. Höhere Wechselwirkungen haben selten einen großen Einfluss und sind generell schwer zu interpretieren. Sie können sich z.B. in der Krümmung der Zielfunktion eines Faktors  $A$  bemerkbar machen, welche von der Einstellung eines anderen Faktors  $B$  abhängt. Dieses Verhalten kann nur dann erfasst werden, wenn ein Wechselwirkungsterm  $A^2B$  in dem Modell aufgenommen wird.

*BEISPIEL: Wir betrachten als Beispiel einen Plan mit 2 Faktoren auf zwei Stufen und folgenden Ergebnissen:*

Nr.	A	B	AB	$\bar{y}$
1	-	-	+	$\bar{y}_1$
2	+	-	-	$\bar{y}_2$
3	-	+	-	$\bar{y}_3$
4	+	+	+	$\bar{y}_4$

**Tabelle 6.10:** Versuchsplan mit 2 Faktoren auf zwei Stufen

Die Effekte der Faktoren können berechnet werden als (vgl. auch Kapitel 6.2.3):

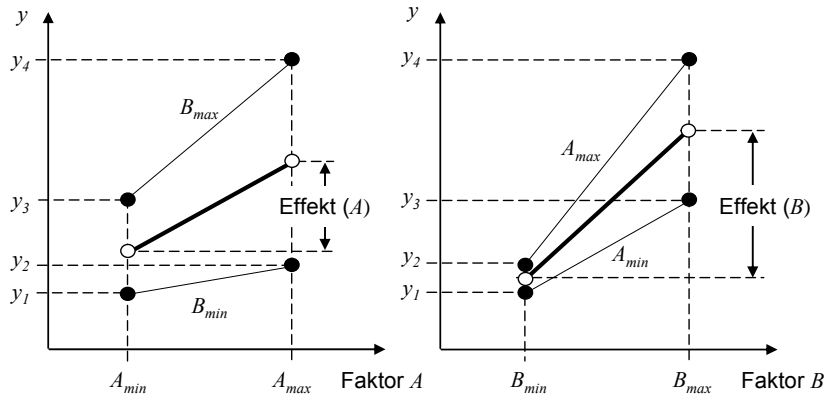
$$\Delta_A = \frac{\sum \bar{y}(A_+)}{2^{k-1}} - \frac{\sum \bar{y}(A_-)}{2^{k-1}} = \frac{\bar{y}_2 + \bar{y}_4}{2} - \frac{\bar{y}_1 + \bar{y}_3}{2}$$



$$\Delta_B = \frac{\sum \bar{y}(B_+)}{2^{k-1}} - \frac{\sum \bar{y}(B_-)}{2^{k-1}} = \frac{\bar{y}_3 + \bar{y}_4}{2} - \frac{\bar{y}_1 + \bar{y}_2}{2} \quad (6.31)$$

$$\Delta_{AB} = \frac{\sum \bar{y}(AB_+)}{2^{k-1}} - \frac{\sum \bar{y}(AB_-)}{2^{k-1}} = \frac{\bar{y}_1 + \bar{y}_4}{2} - \frac{\bar{y}_2 + \bar{y}_3}{2}.$$

Die graphische Darstellung der Effekte ist für das Beispiel des Zweifaktorenversuchs im Bild 6.10 gezeigt.



**Bild 6.10:** Graphische Darstellung der Effekte in diesem Beispiel

Aus obiger Darstellung ist ersichtlich (gleiche y-Skalierung in beiden Bildern vorausgesetzt), dass Faktor B einen größeren Effekt auf die Zielgröße hat, als Faktor A. Außerdem ist eine deutliche Wechselwirkung zwischen beiden Faktoren zu sehen, da die Neigung der Geraden eines Faktors offensichtlich von der Einstellung des jeweils anderen Faktors abhängt.

## 6.6. Schlussfolgerungen und weiteres Vorgehen

Nachdem das Modell berechnet und validiert sowie die Ergebnisse dargestellt und interpretiert wurden, sind Schlussfolgerungen aus den Untersuchungen zu ziehen. An dieser Stelle kann auch die Erreichung des Versuchsziels bewertet werden.

Falls Ergebnisse nicht plausibel oder zufriedenstellend sind, können eventuell weitere Untersuchungen durchgeführt werden, wie z.B. Versuche mit weiteren Faktoren, Faktorstufen, Faktorstufenkombinationen oder Zielgrößen. Gegebenenfalls ist auch ein erweitertes Modell zu betrachten.

Nicht zuletzt muss eine zusammenfassende Dokumentation und ein Wissenstransfer erfolgen.



## 7. Anwendungen

Ziel der in Kapitel 6 dargestellten Auswertung war die Anpassung des Modellansatzes an die Versuchsergebnisse. In diesem Kapitel werden kurz die wichtigsten Anwendungen für das Modell dargestellt, welche einen Erkenntnisgewinn über das reale System mittels Simulation ermöglichen.

### 7.1. Prognose

Nachdem ein validiertes Modell vorliegt, besteht die einfachste Aufgabe darin, den Wert der Zielgröße bei vorgegebenen Werten der Faktoren zu berechnen, d.h. das Verhalten des realen Systems in Abhängigkeit der Einflussgrößen zu prognostizieren.

Dabei ist zu beachten, dass die Prognose bei black-box-Modellen sinnvollerweise nur als Interpolation vorzunehmen ist. Extrapolationen können zu komplett falschen Aussagen führen, da keinerlei experimentelle Information über das Verhalten der Zielgröße außerhalb des untersuchten Versuchsraums vorliegt.

### 7.2. Optimierung

Eine weitere wichtige Aufgabe eines Modells besteht darin, eine Optimierung der Zielgröße zu ermöglichen. Es handelt sich hierbei um die der Prognose komplementären Aufgabe: es sollen jene Einstellungen der Faktoren gefunden werden, bei denen die Zielgrößen vorgegebene Werte annehmen. Hierbei handelt es sich oft um Maxima oder Minima. Gegebenenfalls sind auch Randbedingungen vorgegeben, welche nicht verletzt werden dürfen. Die Einflussfaktoren werden hier üblicherweise als Designvariablen bezeichnet.

Für die Optimierung können unterschiedliche Strategien angewendet werden:

- a) Werden Faktorstufenkombinationen nach einem bestimmten Schema generiert, können deterministische Verfahren verwendet werden. Werden sie zufällig ausgewählt kommen stochastische Verfahren zum Einsatz. Typische Vertreter der deterministischen Verfahren sind der Simplex-Algorithmus und die gradientbasierten Verfahren. Bei den stochastischen Verfahren sind vor allem der Genetische Algorithmus (GA) und die Evolutionäre Strategie (EV) sowie die Künstlichen Neuronalen Netze (KNN) zu erwähnen.
- b) Es können Verfahren danach unterschieden werden, ob für die Berechnungen nur die zu optimierende Funktion oder auch deren 1. Ableitung oder gar zusätzlich deren 2. Ableitung erforderlich ist. Hintergrund ist die Tatsache, dass die Berechnung von Ableitungen aufwendig bis unmöglich sein kann. Typische Vertreter der ersten Kategorie sind alle stochastischen Verfahren sowie das Simplex-Verfahren. Eine erste Ableitung der Funktion in Form eines Gradienten benötigt hingegen das klassische Gradientenverfahren. Das quasi-Newton-Verfahren benötigt sogar die zweite Ableitung der Funktion, hier gibt es allerdings Varianten welche ohne diese auskommen.
- c) Manche Verfahren kommen für eine modellbasierte, andere für eine versuchsbasierte Optimierung in Frage und es gibt auch welche, die in beiden Fällen angewendet werden können. Versuchsbasierte Verfahren benötigen nur die Ermittlung der wahren Antwort des Systems bei diskreten Einstellungen der Einflussgrößen. Diese Systemantwort kann direkt durch Versuche ermittelt werden. Modellbasierte Verfahren hingegen setzen zunächst die Berechnung einer Modellgleichung voraus, d.h. die wahre Antwort des Systems aus dem Versuch wird zunächst durch eine Näherung, genannt Antwortfläche (Response-Surface) approximiert. Anschließend wird eine Optimierung der Näherung durchgeführt. Diese Verfahren werden oft Response-Surface-Methoden genannt. Für versuchsbasierte Optimierung kommen alle Verfahren wie z.B. die Simplex-Methode in Frage, welche nur Funktionswerte an diskreten Stellen des Versuchsraums benötigen. Sind hingegen Ableitungen wie bei den gradientbasierten Verfahren erforderlich, muss zunächst eine Modellgleichung aufgestellt werden mit deren Hilfe die erforderlichen Gradienten berechnet werden können. Alternativ können Gradienten allerdings auch durch Differenzialquotienten angenähert werden, dadurch wäre nur die Ermittlung von Funktionswerten erforderlich.
- d) Je nachdem, ob für eine modellbasierte Optimierung eine globale oder eine lokale (eventuell iterativ zu verbessernde) Modellgleichung verwendet wird, werden Response-Surface-, bzw. Adaptive-Response-Surface-Methoden unterschieden. Global bedeutet in diesem Zusammenhang, dass die Modellgleichung den gesamten Versuchsraum abdeckt. Lokale Modelle decken hingegen nur Teilbereiche des Versuchsraums ab.





- e) Manchmal müssen nicht nur eine, sondern mehrere Zielfunktionen optimiert werden. Dies kann dadurch auf den einfachen Fall zurückgeführt werden, dass alle Zielfunktionen gewichtet und superponiert in einer neuen Zielfunktion zusammengefasst werden. Nachteil dieser Vorgehensweise ist, dass willkürliche Gewichtungsfaktoren benötigt werden, welche den „Kompromiss“ zwischen den einzelnen Zielgrößen eindeutig festlegen. Manchmal ist es aber erforderlich, alle Kombinationen von Einflussgrößen im Versuchsraum zu finden, welche „gleichwertig“ sind, d.h. dass keine Verbesserung einer Zielgröße ohne die Verschlechterung einer anderen möglich ist. Aus dieser Menge kann anschließend anhand weiterer Kriterien ein geeignetes Optimum ermittelt werden. Diesem Ansatz widmet sich die Pareto-Optimierung.

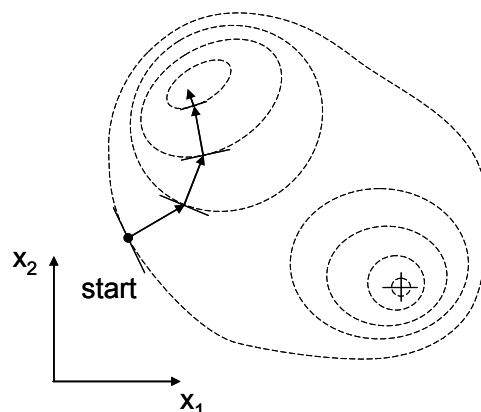
Im Folgenden werden einige der erwähnten typischen Verfahren näher erläutert, für eine detaillierte Beschreibung wird auf die einschlägige Literatur verwiesen. Nachdem mit einem geeigneten Ansatz ein Optimum gefunden wurde, muss in der Regel auch ein Validierungsversuch durchgeführt werden, welcher die Ergebnisse der Optimierung bestätigt.

### 7.2.1. Gradientbasierte Verfahren

Die am häufigsten verwendete Vorgehensweise wird eine modellbasierte Optimierung sein, z.B. durch ein Gradientenverfahren. Hierbei müssen mindestens quadratische Ansätze verwendet werden, da das Extremum eines linearen Ansatzes stets auf dem Rand des Versuchsraums liegt und nicht geeignet ist, eine Antwortfläche mit einem innerhalb des Versuchsraums liegenden Extremum angemessen anzunähern. Im linearen Fall können viel einfachere Ansätze der sogenannten linearen Programmierung verwendet werden.

Beim klassischen **Gradientenverfahren** handelt es sich um ein grundlegendes Verfahren der nichtlinearen Optimierung, welches wie die meisten Verfahren auf einer lokalen Suche basiert. Eine nichtlineare Funktion kann mehrere lokale Extrema besitzen und das Optimierungsverfahren wird eines davon identifizieren. Dieses lokale Extremum muss aber nicht zwangsläufig auch das globale Optimum darstellen. Dies ist auch einer der größten Nachteile gradientenbasierter Verfahren. Ausweg aus dieser Situation bietet die mehrmalige Wiederholung des Verfahrens mit unterschiedlichen Startwerten.

Das Gradientenverfahren basiert auf der Tatsache, dass der Gradient einer skalaren Funktion ein Vektor ist, welcher die Richtung des steilsten Anstiegs in der gegenwärtigen Position vorgibt und dessen Länge ein Maß für die Steigung an dieser Stelle ist. Diese Richtung, welche immer senkrecht zu den Höhenlinien der Funktion liegt, muss man einschlagen, wenn man schnellstmöglich zum Maximum der Funktion gelangen möchte. (Im Übrigen gelten die folgenden Überlegungen auch für die Suche eines Minimums, die Methode kann analog angewendet werden.) Da der Gradient eine lokale Größe ist und von Ort zu Ort variiert, ergibt sich eine iterative Vorgehensweise, bei der in der aktuellen Position der Gradient berechnet und damit die Richtung des Aufstiegs bestimmt wird und anschließend ein Schritt in diese Richtung getan wird, bevor in der neuen Position erneut der Gradient berechnet und das Vorgehen wiederholt wird. Für die Bestimmung der Schrittweite gibt es verschiedene Ansätze, eine sinnvolle Vorgehensweise ist, den Schritt umgekehrt proportional zum Gradientenbetrag zu wählen, so dass der Schritt je kleiner ausfällt, je steiler der Anstieg ist. Die Suche ist beendet, wenn im Rahmen der vorgegebenen Genauigkeit kein benachbarter Ort gefunden wird, welcher höher als die momentane Position liegt – das Maximum ist damit gefunden.

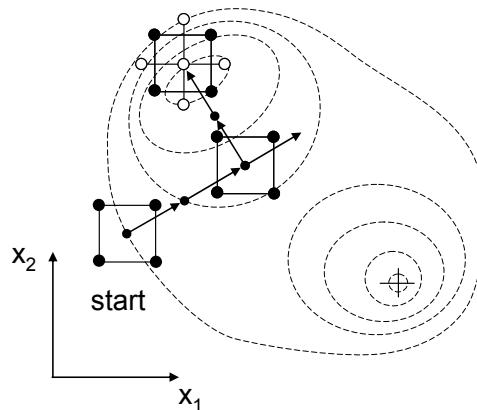


**Bild 7.1:** Gradientenverfahren



Dieses Verfahren ist eines der am weitesten verbreiteten Optimierungsalgorithmen, denn es konvergiert sehr schnell gegen ein lokales Extremum. Nachteile des Verfahrens sind, dass Konvergenz-Schwierigkeiten auftreten, sobald die Antwortfläche durch das Auftreten von Störfaktoren verrauscht ist. Eine weitere Einschränkung ist, dass nur glatte Funktionen mit wohldefiniertem Gradient und damit nur kontinuierliche (quantitative) Designvariablen behandelt werden können. Daraus ergibt sich auch der typische Einsatz dieses Verfahrens bei der modellbasierten Optimierung von Zielgrößen welche mittels Computerexperimenten bestimmt werden, wobei oft ein FEM-Programm als Solver dient.

Ein ähnliches Verfahren, welches auch für eine experimentelle Optimierung geeignet ist, ist die **Methode des steilsten Anstiegs**. Die Grundidee hier ist, den Gradienten der Funktion näherungsweise aus dem Differenzenquotienten der Zielgröße bezüglich der Designvariablen zu bestimmen, so dass keine globale Antwortfläche vorliegen muss. Das Vorgehen besteht darin, im Ausgangspunkt einen typischerweise vollfaktoriellen Plan auf zwei Stufen aufzulegen und daraus die Richtung des steilsten Anstiegs (bzw. Abfalls) näherungsweise zu bestimmen. Gegebenenfalls können auch bei mehr als 5 Faktoren teilfaktorielle Pläne verwendet werden. In dieser Richtung werden nun in gewissen Schritten Einzelversuche solange durchgeführt, bis es wieder zu einem Abstieg kommt. An dieser Stelle wird erneut ein Versuchsplan aufgestellt und das Vorgehen solange wiederholt, bis ein lokales Maximum (bzw. Minimum) im Rahmen einer vorgegebenen Genauigkeit gefunden wurde. In der Nähe des Optimums müssen Versuchspläne für nichtlineare Zusammenhänge verwendet werden, damit dessen Position auch genau genug bestimmt werden kann.



**Bild 7.2:** Methode des steilsten Anstiegs

Eine an den Anforderungen der Fertigung angepasste Variante der Methode des steilsten Anstiegs sind die sogenannten **Evolutionary Operations (EVOP)**. Das Verfahren wird in der laufenden Fertigung verwendet, deshalb sind die Variationsbereiche der Faktoren kleiner und das generelle Vorgehen vorsichtiger. Kernidee von EVOP ist, dass im Rahmen der ständigen Verbesserung der Produktion gleichzeitig produziert und optimiert werden kann. Das Vorgehen wird in [1] näher erläutert.

### 7.2.2. Simplex Verfahren

Grundidee des Simplexverfahrens ist, die Richtung für eine Verbesserung in der gegenwärtigen Position ohne Gradientenberechnung direkt aus einem Versuch mit möglichst wenigen Faktorstufenkombinationen zu bestimmen. Der Vorschlag dabei ist, bei  $k$  Faktoren eine Struktur, den so genannten Simplex, aus  $k+1$  Kombinationen zu verwenden. Bei 2 Faktoren ist dies ein Dreieck.

Das Vorgehen besteht aus folgenden Schritten, vgl. [2], S.308f:

- Zunächst werden im aktuellen Arbeitspunkt die Versuchsergebnisse in den Ecken des Simplex 1 bestimmt.
- Der Punkt mit dem schlechtesten Ergebnis wird an der durch die restlichen Punkte bestimmten Hyperebene (Gerade bei 2 Faktoren) gespiegelt, so entsteht eine neue Faktorstufenkombination. Diese bildet mit den restlichen Punkten des alten Simplex 1 einen neuen Simplex 2 usw.
- So wird das Ergebnis Schritt für Schritt verbessert. Das Vorgehen wird solange wiederholt, bis das angestrebte Optimum gefunden wurde.

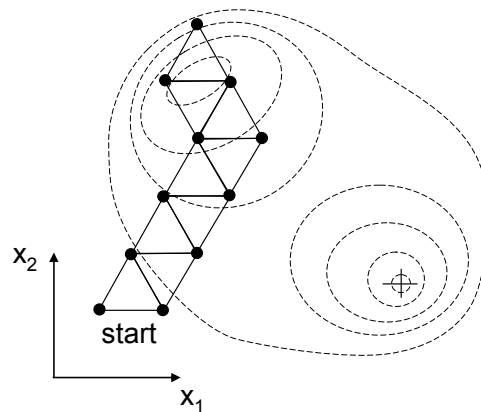
Die **Vorteile** dieses einfachen Vorgehens bestehen darin, dass keine aufwendige Auswertung erforderlich ist (insbesondere keine Gradientenbestimmung) und dass jeder weitere Arbeitspunkt besser



ist als der aktuelle, so dass man immer mit einer Verbesserung rechnen kann, egal wie weit man sich vom tatsächlichen Optimum befindet. Diese Vorteile machen das Verfahren besonders beliebt in der experimentellen Optimierung.

Demgegenüber stehen allerdings auch gewisse **Nachteile**:

- Das Verfahren funktioniert wie beschrieben nur, falls die Effekte deutlich größer sind als die Streuung. Falls verrauschte Antwortflächen vorliegen kann eine Mehrfachrealisierung als Abhilfe verwendet werden, allerdings wird das Verfahren dann schnell aufwendig, da keine Mehrfachverwendung der Versuche möglich ist.
- Das Verfahren konvergiert schlecht in der Nähe des Optimums, was für ein Optimierungsverfahren zunächst unerwartet erscheint. Eine Analyse zeigt jedoch, dass das Verhalten dort nichtlinear ist, so dass der Simplex keine wesentliche Verbesserung der Zielgröße liefern kann.
- Der Simplexalgorithmus ist auch nicht anwendbar bei der Optimierung mit mehreren Zielgrößen.



**Bild 7.3:** Simplex Verfahren

### 7.2.3. Evolutionäre Algorithmen

Diese stochastischen Optimierungsverfahren wurden für solche Fälle entwickelt, bei denen gradientenbasierte Optimierungsverfahren mehr oder weniger große Konvergenzschwierigkeiten aufweisen. Dementsprechend können sowohl stark verrauschte Antwortflächen, sowie qualitative und quantitative Designvariablen behandelt werden. Es handelt sich dabei außerdem um globale Suchverfahren, welche tatsächlich die globalen Optima unter vielen lokalen Extrema ermitteln können. Eine Berechnung der Antwortfläche ist dabei nicht erforderlich, diese wird lediglich an diskreten Stellen durch Einzelversuche ermittelt.

Die Ansätze orientieren sich stark am Naturprinzip dass „nur die Angepassten überleben werden“, denn die Natur muss ebenfalls dem Problem größter Variabilität ohne aufwändige Berechnungsansätze begegnen. Der Preis, der dafür bezahlt wird ist die Konvergenzgeschwindigkeit, d.h. diese Verfahren brauchen für die Suche eines Extremums in der Regel deutlich länger.

Details über eine konkrete Implementierung finden sich in [20].

### 7.3. Robustheit- und Zuverlässigkeitsanalysen

BES-PE-Glossar (Ausgabe 3-2009/11): Robustheit ist die Fähigkeit eines Systems, seine Funktion auch unter veränderlichen Randbedingungen beizubehalten.

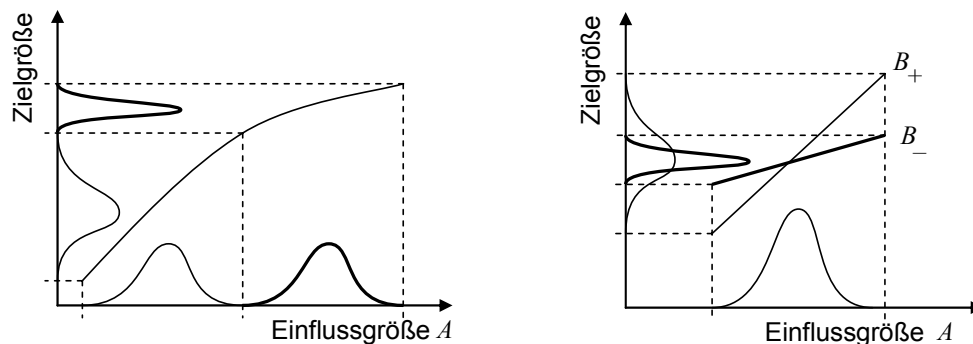
Solche Randbedingungen sind typischerweise zufällig streuende geometrische Größen wie Form- und Lagetoleranzen, Materialparameter wie E-Modul, Umgebungsbedingungen wie Temperatur und nicht zuletzt streuende Lasten. Die Wirkung dieser streuenden Parameter auf designrelevante Zielgrößen wird allgemein mit einer Robustheitsanalyse bewertet. Sie ist integraler Bestandteil der Designgestaltung in der Konzept- und Entwurfsphase.

Die zentrale Frage welche im Rahmen einer Robustheitsanalyse gestellt wird ist, inwieweit streuende Einflussgrößen die Zielgröße beeinflussen können. Ein Prozess ist robust, wenn dessen Zielgrößen möglichst wenig von unvermeidbaren Schwankungen in den Einfluss- und eventueller Störgrößen abhängen. Analog ist das Verständnis von robusten Produkten, deren Eigenschaften möglichst wenig



von den Fertigungs- und Einsatzbedingungen abhängen sollen. Dabei können unterschiedliche Parameter die Robustheit deutlich beeinflussen, so dass gewisse Einstellungen für die Streuung der Zielgröße günstig und andere wiederum ungünstig sein können. Bei der Robustheitsoptimierung geht es letztendlich darum, günstige Werte für jene Einflussparameter zu finden, welche deutlichen Einfluss auf die Streuung haben, ohne den Mittelwert des Prozesses zu verändern.

Die prinzipiellen Zusammenhänge sind im Bild 7.4 verdeutlicht. Im Bild links ist eine nichtlineare Abhängigkeit der Zielgröße von einem Faktor  $A$  dargestellt. Höhere Werte des Faktors vermindern offensichtlich die Streubreite der Zielgröße und sind somit für die Robustheit günstig. Allerdings verschiebt sich der Mittelwert der Zielgröße mit steigendem  $A$  ebenfalls deutlich, was den Designanforderungen widersprechen kann. Im Bild 7.4 rechts ist die Abhängigkeit der Zielgröße von zwei Faktoren  $A$  und  $B$  dargestellt, welche eine deutliche Wechselwirkung aufweisen. Durch eine günstige Einstellung des Faktors  $B$  auf der Stufe  $B(-)$  kann nun die Zielgröße wesentlich robuster gegenüber dem Einfluss des Faktors  $A$  gemacht werden, als bei der Einstellung  $B(+)$ , ohne dabei den Mittelwert der Zielgröße zu verändern. Wechselwirkungen spielen also bei der Robustheitsoptimierung eine ganz wesentliche Rolle.



**Bild 7.4:** Robuste Prozesse (fette Linie) im Vergleich zu weniger robuste Prozesse (dünne Linie)

Eine Robustheitsanalyse wird typischerweise in drei Schritten durchgeführt:

- Zunächst werden jene Parameter bestimmt, welche einen besonders hohen Einfluss auf designrelevante Zielgrößen haben und somit in engen Toleranzgrenzen geführt werden müssen um die Streuung der Zielgröße einzuschränken. Diese können entweder aus bekannten Wirkzusammenhängen zwischen Einfluss- und Zielgrößen abgeleitet oder durch experimentelle Sensitivitätsuntersuchungen (vgl. auch Kapitel 8) bestimmt werden.
- Den Kern der Analyse bildet die quantitative Bewertung der Robustheit des Designs durch Berechnung der Verteilung der Zielgröße in Abhängigkeit der bekannten Verteilungen relevanter Einflussgrößen. Als Maß für die Robustheit wird üblicherweise der Anteil zulässiger unter allen möglichen Zuständen verwendet, welcher aus der Verteilung der Zielgröße innerhalb der aus den Anforderungen abgeleiteten Toleranzen abgeleitet werden kann. Als robust werden z.B. in [20] jene Designs bezeichnet, bei denen das Intervall der Zielgröße  $\mu_y \pm 2\sigma_y$  innerhalb der Toleranzanforderung liegt, vgl. [20], S. 236.
- Anschließend kann ggf. eine Optimierung des Designs im Hinblick auf Robustheit erforderlich werden, bis die gestellten Anforderungen erfüllt sind.

Die Verfahren zur Analyse von Robustheit und Zuverlässigkeit ähneln sich in dem Sinne, dass beide Verfahren den Anteil zulässiger unter allen möglichen Zuständen bewerten. Der Unterschied besteht aus Sicht der Berechnung darin, dass bei der Robustheitsanalyse relativ häufige Ereignisse auftreten, welche bei implizit vorgegebenen Abhängigkeiten zwischen Ziel- und Einflussgrößen mit Hilfe der klassischen stochastischen Samplingmethoden (z.B. Monte-Carlo oder Latin-Hypercube, vgl. auch Kapitel 8) behandelt werden können. Bei der Zuverlässigkeitsanalyse werden hingegen sehr seltene Ereignisse betrachtet, welche den Einsatz spezieller Verfahren erfordern und hier nicht weiter betrachtet werden sollen. Weitere Details zu diesem Thema finden sich in [22] Kapitel 7 und 8.



## 8. Weiterführende Ansätze

### 8.1. Teilfaktorielle Pläne

Teilfaktorielle Pläne werden hauptsächlich verwendet, um bei Fragestellungen, bei denen relativ wenig über die Wirkzusammenhänge bekannt ist, aus einer Menge von potentiell relevanten Faktoren jene mit der höchsten Wirkung auf die Zielgröße zu bestimmen. Man spricht in diesem Zusammenhang auch von „Screening“-Untersuchungen. In zweiter Linie sind solche Pläne auch geeignet, den Versuchsaufwand bei Untersuchungen mit mehreren Faktoren im begrenzten Umfang zu reduzieren.

#### 8.1.1. Motivation

Betrachten wir zunächst Aufgabenstellungen mit 3, 5 und 8 Faktoren auf jeweils 2 Stufen und analysieren wie viele von den Termen eines kompletten multilinearen Ansatzes die Haupteffekte sowie die 2-fach, 3-fach und höheren Wechselwirkungen (WW) beschreiben. Unter den Termen ist zusätzlich auch eine freie Konstante, welche nicht explizit aufgeführt wird.

Faktoren	Anzahl Terme			
	Summe	Haupt-Effekte	2-fach-WW	3-fach und höhere WW
3	8	3	3	1
5	32	5	10	16
8	256	8	28	219

**Tabelle 8.1:** Anzahl Effekte und Wechselwirkungen in Abhängigkeit der Faktoranzahl

Wie aus Tabelle 8.1 ersichtlich, steigt der Anteil höherer Wechselwirkungen mit steigender Faktorenanzahl stark an. Bei einem Plan mit 8 Faktoren auf 2 Stufen mit  $m=2^8=256$  Termen sind bereits ein Großteil davon, nämlich 219, höhere Wechselwirkungen, bei denen vermutlich nur ein Bruchteil tatsächlich als Effekte höherer Ordnung eine Rolle spielen wird. Es ist klar, dass auch genauso viele Versuche eines vollfaktoriellen Plans zu deren Bestimmung erforderlich wären, so dass der Versuchsaufwand mit steigender Faktorenanzahl zunehmend zur Bestimmung höherer Wechselwirkungen verwendet wird.

Die Idee der teilfaktoriellen Pläne ist nun anstelle einiger dieser höheren Wechselwirkungen weitere Faktoren zu untersuchen und dabei die Anzahl der Versuche nicht zu erhöhen.

#### 8.1.2. Effekte und Vermengungen

Betrachten wir dazu ein Beispiel mit 2 Faktoren auf 2 Stufen. Der multilineare Ansatz enthält insgesamt 4 Terme, inklusive einer 2-fach-WW:

$$y = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2 . \quad (8.1)$$

Es sind somit auch 4 Versuche erforderlich, um die unbekanntenen Koeffizienten eindeutig zu bestimmen. Diese können in einem vollfaktoriellen Plan durchgeführt werden. Die Idee eines teilfaktoriellen Plans ist nun, anstelle der 2-fach-WW einen dritten Faktor zu untersuchen und dafür folgenden Ansatz zu verwenden:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 . \quad (8.2)$$

Vor allem sollen wieder nur 4 Versuche verwendet werden, um die 4 unbekanntenen Koeffizienten des Ansatzes zu bestimmen. Dieser teilfaktorielle Plan wird üblicherweise mit  $2^{3-1}$  bezeichnet und soll verdeutlichen, dass 3 Faktoren mit den Versuchen eines 2-faktoriellen Plans untersucht werden sollen.

Allerdings sieht der vollständige multilineare Ansatz mit drei Faktoren wie folgt aus:

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3 + a_{123}x_1x_2x_3 . \quad (8.3)$$

Dieser enthält 8 Koeffizienten die 8 Versuche zu deren Bestimmung erfordern. Welche Konsequenzen die Verwendung von weniger als 8 Versuchen hat, soll am Beispiel des teilfaktoriellen Plans  $2^{3-1}$  erläutert werden.



Die Planmatrix des vollfaktoriellen Plans  $2^3$  ist unten angezeigt. Da für unseren teilfaktoriellen Plan  $2^{3-1}$  nur 4 Versuche zur Verfügung stehen, sollen nur die ersten 4 Versuche dafür verwendet werden.

	A	B	C	AB	AC	BC	ABC
1	-	-	+	+	-	-	+
2	+	-	-	-	-	+	+
3	-	+	-	-	+	-	+
4	+	+	+	+	+	+	+
5	-	-	-	+	+	+	-
6	+	-	+	-	+	-	-
7	-	+	+	-	-	+	-
8	+	+	-	+	-	-	-

**Tabelle 8.2:** Planmatrix eines vollfaktoriellen Plans 2<sup>3</sup>

Die Spalten der zugehörigen Wechselwirkungen  $AB$ ,  $AC$ ,  $BC$  und  $ABC$  lassen sich als Produkt der entsprechenden Spalten für  $A$ ,  $B$  und  $C$  berechnen. Vergleicht man diese Spalten, so zeigt sich, dass  $AB$  mit  $C$ ,  $AC$  mit  $B$  und  $BC$  mit  $A$  jeweils übereinstimmt. Im vollfaktoriellen Plan würde der Unterschied erst durch hinzuziehen der Ergebnisse aus den Versuchen 5 bis 8 erfolgen, welche hier nicht durchgeführt werden. Dementsprechend sind die Spalten  $A$  und  $BC$ ,  $B$  und  $AC$ ,  $C$  und  $AB$  in der Auswertematrix nicht unterscheidbar. Man sagt, der Faktor  $A$  ist mit der Wechselwirkung  $BC$  *vermengt*, gleiches gilt für den Faktor  $B$  und der Wechselwirkung  $AC$  sowie für den Faktor  $C$  und der Wechselwirkung  $AB$ . Darüber hinaus müssten in der Auswertematrix eine weitere Spalte aufgeführt werden, aus der die Vermengung zwischen Absolutglied und der Dreifachwechselwirkung  $ABC$  deutlich wird, darauf wird aber der Einfachheit halber hier nicht näher eingegangen.

Es ist also nicht möglich, beispielsweise den Effekt des Faktors  $A$  getrennt vom Effekt der Wechselwirkung  $BC$  zu berechnen, was einen wesentlichen Nachteil darstellt. Fehlende Versuche führen also zur Vermengung mit dem Risiko der Fehlinterpretation. Insbesondere können entgegengesetzte Effekte und/oder WW eliminiert, Scheineffekte vorgetäuscht oder auch tatsächlich vorliegende Haupteffekte kompensiert werden. Falls aus übergeordneten (z.B. physikalisch begründeten) Überlegungen entschieden werden kann, welcher der vermengten Effekte das Ergebnis bestimmt, hat man das gewünschte Ergebnis und gleichzeitig Versuche eingespart. Sind die Überlegungen falsch oder man ignoriert die Vermengung, hat man ein falsches Ergebnis!

Das Risiko der Fehlinterpretation hängt stark davon ab, welche Effekte vermengt werden. Falls 2-fach-WW mit den Effekten der Faktoren vermengt werden ist dies als kritisch zu betrachten, denn Effekte und 2-fach Wechselwirkungen spielen in den meisten Anwendungen eine wichtige Rolle. Solche Pläne haben die Auflösung III. Weniger kritisch ist eine Vermengung von 2-fach-WW untereinander bei Plänen der Auflösung IV und meistens unkritisch ist die Vermengung der 2-fach-WW mit 3-fach-WW bei Plänen der Auflösung V. Eine Zusammenfassung ist in Tabelle 8.3 vorgegeben.

In der Fachliteratur wird häufig die Anwendung sogenannter Aussiebepläne nach Plackett und Burman empfohlen, um aus einer größeren Anzahl möglicher Einflussfaktoren die wesentlichen auszusieben. Dabei handelt es sich um besondere teilfaktorielle Pläne der Auflösung III und IV, bei denen die erwähnten Nachteile teilweise aufgehoben wurden. Die Versuchsanzahl (Anzahl der Zeilen) ist gegenüber den Faktorenplänen ein ganzzahliges Vielfaches von vier, also  $N = 4, 8, 12, 16, 20, 24, \dots$ . Zudem werden diese Aussiebepläne nach anderen Regeln aufgebaut als die Faktorenpläne.

Trotzdem handelt es sich bei den Plackett-Burman-Plänen um hoch-vermengte Anordnungen, die die bereits angesprochenen Probleme mit sich bringen. Es wird nachdrücklich betont, dass wegen dieser Vermengungsstruktur Plackett-Burman-Pläne nur erfolgreich anwendbar sind, wenn keine Wechselwirkungen vorhanden sind. Andernfalls können sich völlig falsche Schlussfolgerungen ergeben.



Faktoren	Anzahl Versuche		Aufretende Vermengung beim teilfaktoriellen Plan	Auflösung	Bewertung
	voll-faktorieller Plan	teil-faktorieller Plan			
3 5	8 32	$4 = 2^{3-1}$ $8 = 2^{5-2}$	Faktor mit 2fWW	III	kritisch
4 6	16 64	$8 = 2^{4-1}$ $16 = 2^{6-2}$	Faktor mit 3fWW 2fWW mit 2fWW	IV	weniger kritisch
5 8	32 256	$16 = 2^{5-1}$ $64 = 2^{8-2}$	Faktor mit 4fWW 2fWW mit 3fWW	V	weitgehend unkritisch
6 9	64 512	$32 = 2^{6-1}$ $128 = 2^{9-2}$	Faktor mit 5fWW 2fWW mit 4fWW 3fWW mit 3fWW	VI	unkritisch

**Tabelle 8.3:** Auflösung teilfaktorieller Pläne

### 8.1.3. Handlungsempfehlung

Als allgemeine Handlungsempfehlung für Screening-Untersuchungen gilt (vgl. [1], S.145):

- Es genügt, jeden Faktor auf 2 Stufen zu untersuchen.
- Pläne der Auflösung III sollten nicht verwendet werden, es sei denn man kann aus physikalischen Überlegungen die Vermengungsstruktur auflösen.
- Den besten Kompromiss zwischen Aufwand und Vermengung bieten Pläne der Auflösung IV. Diese haben eine klare Vermengungsstruktur ohne Vermengung von Faktoren und 2-fach-Wechselwirkungen und sind somit besonders geeignet für Screening-Untersuchungen.
- Pläne der Auflösung V und höher können zur Ermittlung von Wirkzusammenhängen verwendet werden, denn sie weisen weder eine Vermengung von Faktoren und 2-fach-Wechselwirkungen noch eine Vermengung von 2-fach-Wechselwirkungen untereinander auf.

Falls man nicht mehr als 32 Versuchspunkte untersuchen kann (was durchaus realistisch erscheint, wenn man bedenkt, dass diese u.U. auch mehrfach realisiert werden müssen) folgt daraus

- Untersuchungen der Wirkzusammenhänge sind mit maximal 6 Parametern bei 32 Versuchen (bzw. 5 Parametern bei 16 Versuchen) möglich,
- Screening-Untersuchungen können hingegen mit bis zu 12 Parametern bei 32 Versuchen (bzw. 8 Parametern bei 16 Versuchen) durchgeführt werden.

### 8.2. Pläne für nicht-lineare Zusammenhänge

In den vorangegangenen Abschnitten wurde gezeigt, dass Faktorenpläne vom Typ  $2^k$  stets ein lineares Modell voraussetzen. Extremwerte (Maxima oder Minima) liegen dabei jeweils am Rand des Versuchsraums. Die Richtigkeit des Modells kann nur durch zusätzliche Versuche überprüft werden. Liegt beispielsweise innerhalb des Versuchsraums ein lokales Maximum, so wird man nur durch mindestens einen Versuch innerhalb dieses Bereichs einen Hinweis auf dessen Existenz erhalten können.

Allgemein können Nichtlinearitäten nur erkannt werden, wenn man für jeden Faktor mehr als zwei Stufen wählt. Wegen der exponentiell anwachsenden Versuchszahl stellen drei Stufen für praktische Anwendungen meist auch die Obergrenze dar, so dass für den Anwender im Wesentlichen die quadratischen Ansätze interessant sein dürften. Der vollständige Modellansatz lautet in diesem Fall:

$$y = a_0 + \sum_{i=1}^k a_i x_i + \sum_{i=1}^{k-1} \sum_{j=i+1}^k a_{ij} x_i x_j + \sum_{i=1}^k a_{ii} x_i^2 \quad (8.4)$$

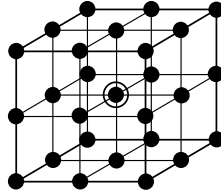


Mit diesem Polynom zweiten Grades können auch gekrümmte Flächen dargestellt werden.

Einige der gängigsten Pläne für nichtlineare Zusammenhänge werden in den folgenden Kapiteln dargestellt. Solche Pläne sollten nur mit wenigen signifikanten Faktoren angewendet werden. Falls die Signifikanz unklar ist, müssen zunächst Screening-Untersuchungen mit linearen Ansätzen durchgeführt werden.

### 8.2.1. Pläne vom Typ $3^k$

Das Naheliegendste ist, für die Ermittlung des quadratischen Modellansatzes je 3 Stufen pro Faktor zu verwenden. Für 3 Faktoren ergibt sich damit auch automatisch der Plan  $3^3$ , wobei die drei Faktorstufen im transformierten Koordinatensystem üblicherweise mit -1, 0, +1 bezeichnet werden.



**Bild 8.1:** Vollfaktorieller Plan  $3^3$

Die Versuchspunkte des Plans  $3^3$  lassen sich schematisch durch einen Würfel im drei-dimensionalen Versuchsraum darstellen, bei dem neben den Würfecken des  $2^3$ -Plans die Flächenmitten, die Kantenmitten und der Zentralpunkt zu untersuchen sind.

In der Auswertematrix des Plans  $3^3$  treten neben den ganzen Zahlen -1, 0 und +1 die Brüche  $-2/3$  und  $+1/3$  auf. Die Koeffizienten des Regressionspolynoms werden also nicht alle in der gleichen Weise berechnet, vgl. [2], S.209. Für die Auswertung eines solchen Plans muss in jedem Fall ein Rechnerprogramm verwendet werden.

Ein Versuchsplan für dreistufige Faktoren wird meist dann in Frage kommen, wenn eine Untersuchung sich auf wenige Faktoren beschränken kann und ein Einzelversuch mit vergleichsweise geringem Aufwand durchführbar ist. In der Praxis ist es beispielsweise denkbar, dass in der "Umgebung" des Zentralpunkts, der einem Serienzustand entsprechen kann, nach besseren Einstellungen der Faktoren gesucht werden soll.

Manchmal handelt es sich bei einem oder mehreren der untersuchten Faktoren um diskrete Einflussgrößen wie z.B. Material oder Maschine. Die Betrachtung eines Regressionspolynoms und die Berechnung theoretischer Zwischenwerte ist in diesem Fall völlig sinnlos. Es genügt also, sich in diesem Fall auf eine varianzanalytische Auswertung zu beschränken und nach dem Prinzip "pick the winner" die Faktoreinstellungen mit dem besten Versuchsergebnis zu ermitteln.

### 8.2.2. Zentral zusammengesetzte Pläne

Wesentlicher Nachteil vollfaktorieller Pläne auf 3 Stufen ist, dass die Anzahl der Faktorstufenkombinationen sehr schnell mit der Anzahl der Faktoren wächst. Außerdem ergibt sich eine hohe Redundanz der Untersuchungen, denn es werden wesentlich mehr Versuche durchgeführt als unbekannte Koeffizienten in der Modellgleichung existieren, siehe Tabelle 8.4.

Anzahl Faktoren	Anzahl Koeffizienten im quadratischen Modell	Anzahl Versuche		Redundanzfaktor	
		$3^k$	$2^k+2k+1$	$3^k$	$2^k+2k+1$
2	6	9	9	1,5	1,5
3	10	27	15	2,7	1,5
4	15	81	25	5,4	1,7

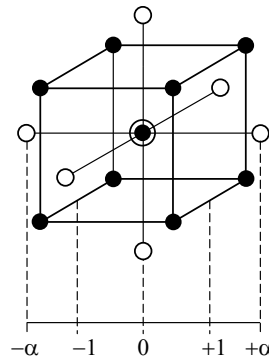
**Tabelle 8.4:** Anzahl Versuche bei zentral - zusammengesetzten Plänen

Eine geeignete Alternative bilden die sogenannten zentral-zusammengesetzten Plänen vom Typ  $2^k+2k+1$  dar. Der Aufbau eines zentral zusammengesetzten Versuchsplans für drei Faktoren ist aus der folgenden Darstellung ersichtlich. Zusätzlich zu den Versuchen in den "Würfelpunkten", wie sie ein vollständiger Plan  $2^3$  vorgibt, werden Versuche in den sogenannten Sternpunkten sowie dem





Mittelpunkt des Sterns, dem Zentralpunkt durchgeführt. Ein solcher Plan heißt zentral zusammengesetzt, wenn der Zentralpunkt des Sterns mit dem Zentralpunkt des Würfels übereinstimmt, so wie es Bild 8.2 zeigt.



**Bild 8.2:** Zentral - zusammengesetzter Plan mit 3 Faktoren

Offensichtlich wird jeder der drei Faktoren auf fünf Stufen variiert. Es wäre denkbar, für jeden Faktor die Stufen -2, -1, 0, +1, +2 zu wählen. Damit ginge allerdings die Orthogonalität des Plans verloren, d.h. die Koeffizienten des Regressionspolynoms könnten nicht mehr unkorreliert und mit der gleichen Varianz bestimmt werden (vgl. [2], S. 228). Damit diese Eigenschaft erhalten bleibt, müssen die Faktorstufen  $-\alpha, -1, 0, +1, +\alpha$  gewählt werden, wobei  $\alpha$  von der Zahl der Versuchsdurchführungen in den Würfelpunkten, in den Sternpunkten und im Zentralpunkt abhängt:

$$\alpha_{orth}^2 = 0.5 \cdot \sqrt{N \cdot 2^k - 2^k}, \quad (8.5)$$

dabei bedeutet  $N$  die Gesamtanzahl von Versuchen (hier  $2^k + 2k + n_0$ ),  $k$  die Anzahl der Faktoren und  $n_0$  die Anzahl der Wiederholungen des Zentralpunktes.

Eine weitere wichtige Eigenschaft eines Versuchsplans ist seine *Drehbarkeit*. Man nennt einen Plan drehbar, wenn die Breite des berechneten Vertrauensbereiches nur vom Abstand zum Zentrum des Würfels abhängt und nicht von der Richtung, vgl. [1], S.191. Man kann zeigen, dass diese Eigenschaft gegeben ist, wenn

$$\alpha_{dreh}^2 = \sqrt{2^k}. \quad (8.6)$$

Man kann durch Anpassung der Anzahl Wiederholungen im Zentralpunkt  $n_0$  erreichen, dass beide Eigenschaften verfügbar sind.

Manchmal ist aus technischen Gründen nur ein Plan auf 3 Stufen durchführbar, in diesem Fall kann  $\alpha=1$  verwendet werden. Dies muss allerdings die Ausnahme bleiben, denn ein solcher „flächenzentrierter“ Plan ist nicht orthogonal.

Als Basis für den Würfel können bei mehr als 3 Faktoren auch teilfaktorielle Pläne  $2^{k-p}$  mindestens der Auflösung  $V$  verwendet werden. In diesem Fall sind die Bedingungen für Orthogonalität und Drehbarkeit anzupassen (die Terme  $2^k$  sind zu ändern in  $2^{k-p}$ ).

In der Praxis bietet sich die Wahl eines zentral zusammengesetzten Plans an, wenn ein zunächst durchgeführter Plan des Typs  $2^k$  durch zusätzliche Versuche ergänzt werden soll. Es sind in diesem Fall lediglich  $2k + n_0$  weitere Versuchspunkte zu untersuchen.

### 8.2.3. D – optimale Pläne

D-optimale Pläne bieten einige große Vorteile:

- Sie bieten große Freiheit, was die Anpassung des Plans an eine (z.B. aus Kostengründen) vorgegebene Anzahl von Versuchen angeht.
- Sie können an ein vorgegebenes Modell angepasst werden und sind somit in der Lage, nichtlineares Verhalten zu erfassen.
- Beliebige Stufen können vorab festgelegt werden, welche in dem Plan vorhanden sein müssen. Dies erlaubt die Mitbenutzung von historischen Daten.



- Bereiche des Versuchsraums können ausgeschlossen werden, falls diese nicht von Interesse sind.
- Mehrstufige, quantitative und qualitative Faktoren können kombiniert werden

Ein gewisser Nachteil ist, dass die Pläne nicht orthogonal sind, wobei der Unterschied allerdings meistens nicht sehr groß ist. Das Vorgehen kann außerdem nur mit einer Software realisiert werden, allerdings bieten die meisten auf DoE spezialisierten Softwaretools mittlerweile auch D-optimale Pläne an.

Für die Bestimmung der Pläne werden iterativ jene Einzelversuche ausgewählt, welche es erlauben die Koeffizienten im Regressionsmodell möglichst genau (mit möglichst kleinen Vertrauensbereichen) zu bestimmen. Eine detaillierte Darstellung der Vorgehensweise würde den Rahmen dieses Heftes sprengen, somit wird darauf verzichtet. Weitere Details sind z.B. in [4] zu finden.

D-optimale Pläne bieten eine interessante Alternative zu den zentral-zusammengesetzten Plänen. Es wird empfohlen sie dann einzusetzen, wenn die zentral-zusammengesetzten nicht angewendet werden können. Typische Beispiele sind qualitative (kategorische) Faktoren mit mehr als 2 Niveaus, Begrenzungen im Versuchsraum, frei wählbare Blockgrößen, von der Voreinstellung abweichende Variantenzahl, Stufenwerte etc.

#### 8.2.4. Handlungsempfehlung

Pläne für nichtlineare Zusammenhänge sollten nur bei wenigen signifikanten Faktoren angewendet werden, um die Wirkzusammenhänge mit der Zielgröße im Detail zu bestimmen und ggf. eine Optimierung durchzuführen. Für die meisten Anwendungen dürften die Anzahl der signifikanten Faktoren zwischen 3 und 5 liegen. Falls diese nicht bekannt sind, müssen zunächst Screening-Untersuchungen mit linearen Ansätzen angewendet werden.

Es gelten folgende Empfehlungen, vgl. auch [1], S. 208:

- Im Regelfall ist ein zentral-zusammengesetzter Plan am besten geeignet. Bis zu 4 Faktoren kann ein vollfaktorieller Plan für den Würfel verwendet werden, bei 5 Faktoren genügt ein teilfaktorieller Plan der Auflösung V.
- Soweit wie möglich müssen die Stufenwerte des Plans so eingestellt werden, dass sich ein orthogonaler Plan ergibt. Durch geeignete Wahl der Anzahl Wiederholungen im Zentralpunkt muss auch eine Drehbarkeit des Plans erreicht werden.
- Sind in Ausnahmefällen aus technischen Gründen nur 3 Stufen durchführbar, kann ein flächenzentrierter Plan mit  $\alpha=1$  verwendet werden.
- Wenn kategorische Faktoren mit mehr als 2 Stufen, nicht realisierbare Kombinationen im Versuchsraum, von der Voreinstellung abweichende Stufenwerte o.ä. gegen den Einsatz von zentral-zusammengesetzten Plänen sprechen, müssen D-optimale Pläne verwendet werden.

#### 8.3. Alternative Modellansätze

Wenn keine physikalische Modellbildung möglich war, wurde in Kapitel 4.4 als allgemeiner Modellansatz meistens a-priori ein lineares oder quadratisches Polynom angenommen. Hintergrund hierzu bildet die Tatsache, dass jede stetige Funktion in der Umgebung eines Punktes durch eine Potenz-Reihe mit einem Fehler höherer Ordnung angenähert werden kann. Dieser Vorgang ist auch als Taylor-Reihenentwicklung bekannt.

Diese Näherung ist jedoch nur in der Umgebung des Entwicklungs-Punktes, also *lokal*, gültig. Probleme können auftreten, wenn in Teilbereichen des Versuchsraumes unterschiedliches Verhalten auftritt, so dass die *globalen* Verhältnisse nicht durch einfache Polynome ersten oder zweiten Grades erfasst werden können. Üblicherweise begegnet man solchen Schwierigkeiten, indem Potenzglieder höherer Ordnung in dem Modellansatz berücksichtigt werden. Dass diese Strategie nicht immer erfolgreich ist, soll mit dem folgenden Beispiel verdeutlicht werden.

*BEISPIEL 1:* Folgende Funktion soll einmal im Intervall  $I_1 = (-0.6; +0.6)$  und einmal im Intervall  $I_2 = (-3.0; +3.0)$  durch Polynome unterschiedlichen Grades angenähert werden. Die Funktion liegt tabellarisch in 7 gleichmäßig verteilten Punkten des jeweiligen Intervalls vor:



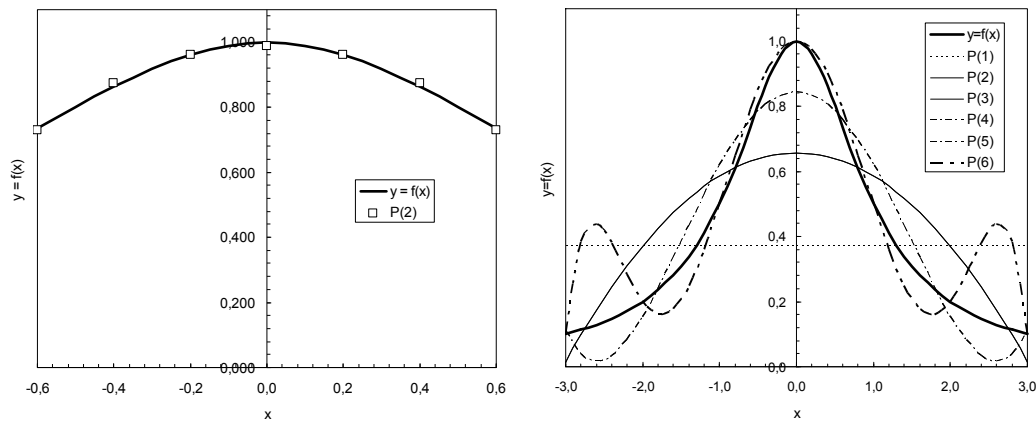
$$f(x) = \frac{1}{1+x^2}$$

x	y=f(x)
-0,6	0,735
-0,4	0,862
-0,2	0,962
0,0	1,000
0,2	0,962
0,4	0,862
0,6	0,735

x	y=f(x)
-3,0	0,1
-2,0	0,2
-1,0	0,5
0,0	1,0
1,0	0,5
2,0	0,2
3,0	0,1

**Tabelle 8.5:** Datenpunkte des Beispiels

Die berechneten Ausgleichskurven sind in Bild 8.3 dargestellt. Wie daraus ersichtlich ist, reicht für eine lokale Approximation im Intervall  $I_1$  ein quadratisches Polynom völlig aus. Für eine globale Betrachtung im Intervall  $I_2$  reicht der quadratische Ansatz jedoch nicht aus, die Erhöhung des verwendeten Polynomansatzes von Grad 1 (linear) bis Grad 6 verbessert zudem nur teilweise die Qualität der Interpolation. Insbesondere die höheren Polynomansätze neigen zum „Schwingen“ an den Intervallgrenzen und sind kaum geeignet das globale Verhalten der Funktion angemessen wiederzugeben. Dieses Phänomen wird auch „overfitting“ genannt.



**Bild 8.3:** Berechnete Ausgleichskurven

Das Problem besteht offensichtlich in dem prinzipiellen Unvermögen der Polynome, einen komplexen globalen Funktionsverlauf anzunähern. Dieser Umstand hat zur Entwicklung alternativer Modell-Ansätze geführt. Einige davon werden im Folgenden kurz zusammengefasst. Weitere Einzelheiten zu diesem Thema finden sich in [26,27].

- **Trigonometrische Funktionen:** Als Basisfunktionen können Sinus und Cosinus-Funktionen verwendet werden. Dies würde dem Vorgehen bei einer Fourier - Reihenentwicklung der Funktion entsprechen. Der Ansatz ist hier, wie beim klassischen Vorgehen, global, in dem Sinne, dass jede der Basisfunktionen im gesamten Versuchsraum einen Beitrag leistet.
- **Splines:** Ein Spline der Ordnung  $k$  ist eine glatte Kurve, welche stückweise aus unterschiedlichen Polynomen derselben Ordnung zusammengesetzt ist. Damit soll gewährleistet werden, dass ein lokal-unterschiedliches Verhalten angemessen abgebildet werden kann. Die Problematik dabei ist, den glatten verlauf des Splines an den „Nahtstellen“ zu gewährleisten, also in den Stützstellen, wo die unterschiedlichen Polynomen angrenzen. Häufig werden kubische Splines ( $k=3$ ) verwendet, welche durch Zusatzbedingungen sicherstellen, dass die Kurve in den Stützstellen 2-mal stetig differenzierbar ist, was bedeutet, dass keine Sprünge und „Knicke“ im Funktionsverlauf auftreten. Auf die explizite Darstellung der Spline-Basisfunktionen wird an dieser Stelle verzichtet, vgl. dazu [26]. Es soll nur erwähnt werden, dass es sich hierbei um Funktionen aus  $(k+1)$  Polynomstücken der Ordnung  $k$  handelt, die  $(k-1)$ -mal stetig differenzierbar zusammengesetzt werden und lediglich in einem über  $k+2$  benachbarte Knoten gebildeten Teilbereich des Versuchsraums von Null verschieden sind. Der Umstand, dass eine Funktion nur in einem begrenzten Teilbereich von Null verschieden ist, wird mit „Funktion mit kompakten Träger“ bezeichnet.
- **Radiale Funktionen:** Als Basis werden hier Funktionen verwendet, die zwar einen globalen Charakter haben können, welche aber mit wachsendem radialen Abstand (daher auch der Name)

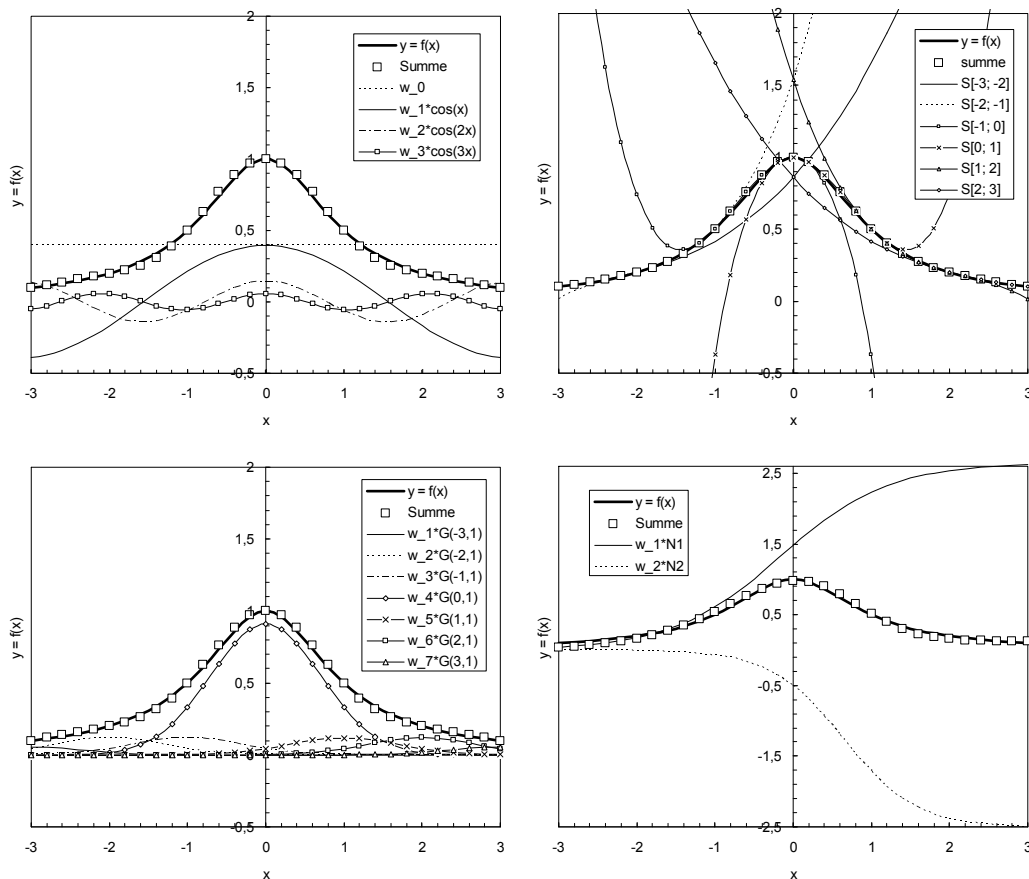


zu einem lokalen Punkt im Versuchsraum schnell abklingen und auf 1 normiert sind. Solche Funktionen werden auch „Kerne“ genannt. Typisches Beispiel sind Funktionen von der Form der Wahrscheinlichkeitsdichte der Gauss-Verteilung (Gauss'sche Glocken-Kurve), welche die wichtige Eigenschaft besitzen, dass sie sehr glatt auslaufen und somit auf natürliche Weise das Problem des stetigen Verlaufs der Modellgleichung lösen. Damit werden die Vorteile eines lokalen Modellansatzes sichergestellt, ohne die Nachteile eines nicht ausreichend glatten Verlaufs des Modells in Kauf nehmen zu müssen.

- **Sigmoidale Funktionen:** Solche Funktionen tauchen typischerweise im Rahmen des Konzeptes Künstlicher Neuronaler Netze (KNN) auf. Ein KNN ist ein Netzwerk aus künstlichen Strukturen, welche den natürlichen Neuronen (Nervenzellen) nachempfunden sind. Jedes Neuron kann mathematisch durch seine Aktivierungsfunktion beschrieben werden, welche seine Ausgabe in Abhängigkeit der Eingangsgrößen definiert. Als Aktivierungsfunktion wird häufig eine sigmoidale Funktion verwendet.

Die Eignung der erwähnten alternativen Modellansätze soll anhand eines Beispiels verdeutlicht werden:

*BEISPIEL 2:* Für die Funktion aus Beispiel 1 soll eine Regression im Bereich  $I_2$  anhand der dort aufgeführten 7 Stützstellen vorgenommen werden. Als Basisfunktionen für das Modell sollen nun nicht Polynome, sondern trigonometrische Funktionen, Spline-Basen, Radiale, sowie Sigmoidale Funktionen verwendet werden.



**Bild 8.4:** Ausgleichskurven mit alternativen Modellansätzen

Die Ergebnisse der 4 Ansätze sind in Bild 8.4 dargestellt. Die Kurve mit der höheren Stärke stellt die tatsächliche Funktion, die Kurve mit den großen Vierecken die jeweilige Approximation dar. Aus den Bildern ist ersichtlich, dass alle Ansätze sehr gut in der Lage sind, das globale Verhalten der Funktion wiederzugeben.

Desweiteren sind durch Kurven mit einer kleineren Stärke die verwendeten gewichteten Funktionen (bzw. die verwendeten Teilpolynome beim Spline) aufgeführt. Als trigonometrische Funktionen wurden 3 Cosinus-Terme und eine Konstante benutzt (Bild 8.4 oben links). Der kubische Spline im Bild 8.4 oben rechts wurde durch 6 Polynome 3. Grades  $S[a,b]$  definiert, jedes davon ist jeweils im Intervall  $[a,b]$  gültig. Bei den Radialen Basisfunktionen wurden 7 Gauss-Funktionen verwendet (Bild 8.4 unten links). Schließlich wurde ein KNN mit sigmoidalen Funktionen als Aktivierungsfunktion für die Approximation im Bild 8.4 unten rechts verwendet.



Die Berechnung der Gewichtskoeffizienten wurde (mit Ausnahme des Splines) durch Minimierung der Summe der Residuenquadrate mit dem Optimierer (Solver) von MS-Excel durchgeführt. Der Spline wurde mit dem Java-Applet von <http://www.arndt-bruenner.de/mathe/scripts/kubspline.htm> berechnet.

Obige univariate Beispiele haben alternative Wege zur Wahl des verwendeten Modellansatzes aufgezeigt. Der wesentliche Unterschied zu den klassischen Polynomansätzen besteht darin, lokal unterschiedliche Funktionen zu verwenden, anstatt ein globales, komplexes Polynom durch Parameteridentifikation an den Versuchsergebnissen anzupassen. In diesem Sinne spricht man bei solchen Ansätzen häufig auch von „nicht-parametrischen Modellen“. Bei Fragestellungen mit umfangreichen Versuchsdaten wird das Vorgehen kompliziert, so dass auf eine geeignete Software nicht verzichtet werden kann. Eine geeignete Lösung, welche den Ansatz der Gauss-Basisfunktionen implementiert, ist im Anhang aufgeführt.

#### 8.4. Stochastische Versuchspläne

In den vorangegangenen Kapiteln wurden unterschiedlichste Verfahren vorgestellt, mit denen der Zusammenhang zwischen Einfluss- und Zielgrößen ermittelt werden konnte. All diese Verfahren können als Samplingmethoden bezeichnet werden, da sie versuchen, durch die Generierung isolierter Stützstellen (die Faktorstufenkombinationen) einen kontinuierlichen Versuchsraum zu diskretisieren. Mit Hilfe dieser Stützstellen soll anschließend eine Antwortfläche (die Modellgleichung) berechnet werden, welche die wahren aber unbekanntem Wirkzusammenhänge zwischen Einfluss- und Zielgrößen annähernd beschreiben soll. Dabei handelt es sich durchweg um deterministische Ansätze, welche die Stützstellen nach einem bestimmten vorher festgelegten Schema generieren, bei einem vollfaktoriellen Plan sind dies beispielsweise die Ecken des Versuchsraums. Insofern können alle bisher vorgestellte Ansätze als *deterministische Samplingmethoden* bezeichnet werden.

Bei einer ganzen Gruppe von Anwendungen, wie bei der Analyse der Robustheit und Zuverlässigkeit, muss allerdings nicht nur der funktionale Zusammenhang zwischen Einfluss- und Zielgrößen beschrieben werden, sondern zusätzlich auch noch die Verteilung der streuenden Zielgröße in Abhängigkeit der streuenden Einflussgrößen. Die Verteilung der Zielgröße beantwortet die zentrale Frage von Robustheits- und Zuverlässigkeitsanalyse, welcher Anteil von allen möglichen Zuständen zulässig im Sinne der Designanforderungen ist.

Falls keine expliziten Modelle vorliegen, wird die Verteilung einer Zielgröße typischerweise durch eine Monte-Carlo-Simulation ermittelt, bei der für jede streuende Einflussgröße, je nach einer eingangs festgelegten Verteilung, eine zufällige Stichprobe von Werten generiert wird. Jede Kombination von Werten der unterschiedlichen Einflussgrößen bildet so ein mögliches „Design“, welches in der Stichprobe aller Designs bereits mit der „richtigen“ Häufigkeit auftritt. Somit kann die Verteilung der Zielgröße durch statistische Auswertung der Stichprobe aller Designs berechnet werden.

Eine zentrale Aufgabe bei der Monte-Carlo-Simulation ist die Generierung zulässiger Kombinationen aus Einflussgrößen im Designraum, wobei die Stichprobe einer jeden Einflussgröße der eingangs festgelegten Verteilung entstammen muss. Diese Aufgabe ähnelt sehr der Aufstellung eines Versuchsplans, andererseits unterscheidet sie sich auch von den klassischen DoE-Ansätzen darin, dass die Einflussfaktorkombinationen nicht nach einem festen Schema, sondern als zufällige Stichprobe einer Verteilung aufgestellt werden müssen. Insofern werden diese Verfahren als *stochastische Samplingmethoden* und die erzeugten Pläne als *stochastische Versuchspläne* bezeichnet. Im Folgenden werden die beiden wichtigsten Ansätze vorgestellt, weitere Details zu diesem Thema sind in [20,22] zu finden.

##### 8.4.1. Plain Monte Carlo (PMC)

Bei dem PMC-Ansatz handelt es sich um die naheliegendste Vorgehensweise, bei der zunächst mit einem Zufallszahlengenerator  $N$  pseudo-Zufallszahlen  $u_i$  generiert werden, welche im Intervall  $[0,1]$  gleichverteilt sind. Anschließend werden diese Zahlen als Wahrscheinlichkeiten der zugehörigen Merkmalswerte  $x_i$  einer bestimmten Verteilungsfunktion  $F(x)$  interpretiert. Die Merkmalswerte können so durch Anwendung der inversen kumulativen Verteilungsfunktion berechnet werden:

$$x_i = F_x^{-1}(u_i). \quad (8.7)$$

Mit diesem Ansatz können pseudo-Zufallszahlen für jede Verteilung berechnet werden, für die eine Inverse ihrer kumulativen Verteilungsfunktion definiert ist [22].

Der Hauptvorteil dieses Verfahrens ist in seiner Robustheit und Einfachheit zu suchen. Ein gewisser Nachteil liegt darin, dass bedingt durch die Zufälligkeit der Stichproben auch zufällige Abweichungen



von der Idealverteilung auftreten können, welche nach dem Gesetz der großen Zahlen erst für sehr große Stichprobenumfänge aufgehoben werden, so dass sich die Konvergenz einer Monte-Carlo-Simulation erst relativ langsam einstellt. Abhilfe schafft hier eine Weiterentwicklung des PMC-Verfahrens, welche im nächsten Kapitel vorgestellt wird.

#### 8.4.2. Latin Hypercube (LHC)

Beim PMC-Verfahren können, wie bereits erwähnt, zufällige Abweichungen von der Idealform der Verteilung auftreten, welche durch die Dichtefunktion  $F(x)$  vorgegeben ist. So werden z.B. bei der Stichprobe einer Normalverteilung die Werte  $F(x)$  in einem Normalverteilungsplot nicht auf einer idealen Gerade liegen, es werden zufällige kleine Abweichungen vorkommen.

Diese Abweichungen können wie folgt beschrieben werden: Falls der Umfang einer Stichprobe  $N$  ist, und  $N$  gleichwahrscheinliche (im Allgemeinen nicht unbedingt äquidistante!) Klassen gebildet werden, muss bei einer Klassierung im Idealfall jede Klasse je nur einmal besetzt werden. Durch zufällige Abweichungen vom Idealzustand kann es jedoch vorkommen, dass nicht jede dieser Klassen besetzt wird bzw. gewisse Klassen mehrfach besetzt werden. Hier setzt das LHC-Verfahren an. Falls der Umfang der Stichprobe  $N$  ist, wird das Wahrscheinlichkeitsintervall zunächst in  $N$  gleichwahrscheinliche Sub-Intervalle

$$D_i = (x_i, x_{i+1}); F(x_{i+1}) - F(x_i) = 1/N; i = 1 \dots N \quad (8.8)$$

unterteilt. Anschließend wird in *jedem* dieser Intervalle *eine* Zufallszahl wie bei dem PMC-Verfahren generiert, wobei dazu entweder die tatsächliche Verteilung berücksichtigt wird oder der Einfachheit halber (bei großen Stichprobenumfängen) auch die Mitte des Intervalls verwendet werden kann. Das LHC Verfahren versucht also, bildlich gesprochen, dem Zufall systematisch auf die Sprünge zu helfen.

Das LHC-Verfahren hat klare Vorteile in der Konvergenzgeschwindigkeit gegenüber dem PMC-Verfahren, so dass bei gleicher Genauigkeit hier kleinere Stichprobenumfänge verwendet werden können, vgl. auch [20]. Da der Stichprobenumfang auch die Anzahl der durchzuführenden Simulationen (externe Solver-Aufrufe) bei dem Monte-Carlo-Verfahren vorgibt, liegen die Vorteile klar auf der Hand.

#### 8.5. Spezielle Erprobungspläne

Für die Verifikation und Validierung von Qualitätsmerkmalen wie Funktionalität, Dauerhaltbarkeit, Robustheit oder Sicherheit werden Bauteile erprobt. Dabei kommen spezielle Versuchstechniken zum Einsatz, deren Behandlung den Rahmen des vorliegenden Heftes sprengen würde. Diese Verfahren werden deswegen hier nicht weiter behandelt, Details dazu sind u.a. in [17] zu finden.

Darüber hinaus muss allerdings erwähnt werden, dass generell nichts gegen den Einsatz von DoE-Ansätzen bei der Erprobung sprechen würde, um so das Maximum an Information über die Wirkzusammenhänge zu gewinnen und dabei Ressourcen nicht zu verschwenden. In manchen Fällen würde es schon genügen die Einflussgrößen, welche bei einem typischen Dauerlauf wirken, systematisch zu erfassen und bei der Auswertung zu berücksichtigen, um neue Erkenntnisse über das Produkt zu gewinnen. Oder es könnten Einflüsse auf die Lebensdauer von Bauteilen durch ein DoE-Schema untersucht werden, wobei der Lebensdauerversuch selbst eine komplette Serie von Einzelversuchen darstellt.

Damit entsteht ein hierarchisches Vorgehen, bei dem auf der Meta-Ebene die Struktur der Untersuchungen seitens eines DoE-Schema vorgegeben wird und auf der operativen Ebene beliebige Untersuchungen durchgeführt werden können, welche zur Ermittlung der betrachteten Zielgrößen führen. Diese müssen nicht einfache Versuche sein, sondern können selbst komplexe Untersuchungen mit langen Versuchsreihen ergänzt durch Simulationen beinhalten.



## 8.6. Versuchspläne nach Taguchi

Der hier vorgestellte Ansatz wurde vom japanischen Qualitätsingenieur und Statistiker G. Taguchi (geb. 1924) nach dem 2. Weltkrieg entwickelt, um die Produktion robuster gegenüber Störeinflüssen zu gestalten und Qualitätsverluste zu vermeiden, ein Kernpunkt japanischer Qualitätsphilosophie. Die Vorgehensweise fand in Verbindung mit der Qualitätsphilosophie Six Sigma schnell eine breite Anwendung und wird in manchen Kreisen fast schon als Synonym für Versuchsplanung verwendet.

Ohne die unbestrittenen Beiträge von Taguchi zur Entwicklung und Popularisierung der Versuchsmethodik und insbesondere sein Verständnis über die Beziehung von Qualität und Robustheit in Frage zu stellen, werden einige der von ihm vorgeschlagenen Lösungswege heute kritisch gesehen. Im Rahmen seiner Grundidee können allerdings ohne Weiteres die Erkenntnisse der klassischen Versuchsplanung verwendet werden, so wie sie in den vorangegangenen Kapiteln dargestellt wurden. Daraus ergibt sich eine Vorgehensweise, welche den Geist, aber nicht die umstrittenen Nachteile des Original-Ansatzes beinhaltet, so dass das Vorgehen auch heute noch seine Aktualität besitzt. Nachfolgend wird eine kurze Einführung gegeben, weitere Einzelheiten finden sich in [3,24], zahlreiche Anwendungen sind in [25] aufgeführt.

### 8.6.1. Motivation: robuste Produkte und Prozesse

Der Ursprung vieler experimenteller Untersuchungen liegt in einem Qualitätsproblem, bei dem aufgrund hoher Streuungen von Qualitätsmerkmalen ein hoher Anteil an Ausschuss produziert wird. Dabei wird nicht nur nach dem „warum“ gefragt, sondern auch „wie“ die auftretende Streuung verringert werden kann. Beispielsweise muss das Ergebnis eines Fertigungsprozesses möglichst insensitive (oder robust) gegenüber Schwankungen von Umgebungsbedingungen wie Temperatur und Luftfeuchtigkeit sein.

Ein robuster Prozess ist jener Prozess, dessen Ergebnis möglichst wenig von einem vorgegebenen Zielwert abweicht, d.h. dessen Streuung möglichst klein ist. Das Ziel der Prozessoptimierung ist demnach die Reduktion der Streuung.

Die grundsätzliche Frage, auf welche Weise Einflussgrößen die Robustheit beeinflussen können, wurde bereits in Kapitel 7.3 diskutiert. An dieser Stelle geht es nun darum, wie günstige Werte für jene Einflussparameter gefunden werden können, welche deutlichen Einfluss auf die Streuung haben, ohne dabei den Mittelwert des Prozesses zu verändern.

### 8.6.2. Versuchspläne nach Taguchi

Taguchi unterscheidet auf dem Weg zu robusten Prozessen zwei Arten von Größen, vgl. auch [1], S.149:

- Die **Steuerfaktoren** (control factors) sind solche Einflussfaktoren, welche sowohl im Laborversuch als auch in der Fertigung eingestellt werden können.
- Die **Rauschfaktoren** (noise factors) hingegen können zwar im Laborversuch eingestellt werden, nicht aber später in der Fertigung, z.B. aus Kostengründen. Beim robusten Prozess soll das Ergebnis möglichst insensitive gegen diese Faktoren sein, z.B. Position in der Vorrichtung, Umgebungstemperatur etc. Die Rauschfaktoren dürfen nicht mit den Störfaktoren verwechselt werden, welche prinzipiell unbekannt sind und daher auch nicht eingestellt werden können, weder im Versuch noch in der Produktion.

Das Vorgehen besteht nun im Prinzip darin, jene Kombinationen der Steuergrößen zu finden, welche die Streuung der Zielgröße (infolge der Streuung der Rauschfaktoren) minimieren. Dabei werden zwei Arten von Prüfplänen verwendet:

- Für die Steuerfaktoren wird ein sogenanntes **inneres Feld** (Inner Array) verwendet, von Taguchi auch „orthogonales Feld“ bezeichnet. Es entspricht einem teilfaktoriellen Versuchsplan der Auflösung III in der klassischen statistischen Versuchsplanung, vgl. [1], S.159. Wie in Kapitel 8.1 diskutiert, sind Pläne dieser Auflösung mit Nachteilen verbunden und können durch Vermengungen der Haupteffekte und zweifach-Wechselwirkungen zu Problemen führen. An dieser Stelle empfiehlt es sich einfach einen Plan der Auflösung IV zu verwenden.
- Für die Rauschfaktoren wird ein **äußeres Feld** (Outer Array) verwendet, hierbei handelt es sich entweder um einen Faktor in mehreren Stufen oder einen teilfaktoriellen Versuchsplan der



Auflösung III für mehrere Faktoren. An dieser Stelle ist die niedrige Auflösung jedoch unkritisch, weil im weiteren Verlauf das äußere Feld als eine Art Versuchswiederholung gehandhabt wird.

**8.6.3. Auswertung und Ergebnis**

Die Auswertung erfolgt folgendermaßen:

- Zur Erfassung der Streuung kann das Signal-Rausch-Verhältnis

$$S/N = 10 \cdot \log_{10} \left( \frac{\bar{y}^2}{s^2} \right), \tag{8.9}$$

oder eine andere geeignete Größe (z.B. Standardabweichung  $\sigma$ , oder die Varianz  $s^2$ ) als separate Zielgröße neben dem Mittelwert  $y$  erfasst werden.

- Die Effekte von Mittelwert und S/N werden so berechnet, als sei das äußere Feld eine Wiederholung des Versuchs, obwohl dies nicht der Fall ist, da die Rauschfaktoren gezielt variiert werden. Damit ist eine konservative Signifikanzaussage möglich, denn die Signifikanz der Effekte wird geringfügig unterschätzt. Für die Zielgröße S/N gibt es nur eine Realisierung, deshalb wird diese Zielgröße wie bei Versuchen mit einfacher Realisierung  $n=1$  behandelt.
- Die Faktoren werden je nach Ergebnis folgendermaßen aufgeteilt:
  - Faktoren, die die Streuung (und ev. zusätzlich noch den Mittelwert) beeinflussen werden so gewählt, dass sie die Streuung minimieren.
  - Faktoren, die nur den Mittelwert beeinflussen werden dazu benutzt, die Zielgröße je nach Designvorgabe einzustellen.
  - Schließlich werden solche Faktoren, die weder Mittelwert noch Streuung beeinflussen zur Minimierung der Kosten eingesetzt.

*BEISPIEL: Ein Beispiel mit 3 Steuerfaktoren C1 bis C3 sowie 3 Rauschfaktoren N1 bis N3 ist in der folgenden Tabelle 8.6 dargestellt. Als Inner Array wird ein vollfaktorieller Plan auf 2 Stufen verwendet, als Outer Array der Plan  $2^{3-1}$ . Die Effekte von Mittelwert und S/N wurden nach deren Signifikanz mit „---“, „\*\*“, „\*\*\*“ (nicht signifikant, signifikant, hoch signifikant) bewertet. Mit dieser Auswertung würden sich die Faktoren C1 und C2 zur Erhöhung der Robustheit anbieten, denn sie wirken sich signifikant auf S/N aus. Deren Einstellung muss in einer Optimierung also auf jener Stufe erfolgen, welche die kleinere Streuung (hohes S/N) aufweist. Dies würde jedoch den Mittelwert der Zielgröße verändern, eine Korrektur kann mittels der Größe C3 vorgenommen werden, welche nur den Mittelwert der Zielgröße beeinflusst, nicht jedoch S/N. C3 wird also so eingestellt, dass die Zielgröße den erforderlichen Mittelwert aufweist.*

			N1	-	+	-	+			
			N2	-	-	+	+			
			N3	-	+	+	-			
Nr.	C1	C2	C3	1	2	3	4	y	s	S/N
1	+	+	+	$y_{11}$	$y_{12}$	$y_{13}$	$y_{14}$	$y_1$	$s_1$	$S/N_1$
...										
8	-	-	-	$y_{81}$	$y_{82}$	$y_{83}$	$y_{84}$	$y_8$	$s_8$	$S/N_8$
$\Delta_y$	*	---	*							
$\Delta_s$										
$\Delta_{S/N}$	**	*	---							

**Tabelle 8.6:** Auswertung eines Taguchi-Plans

Die Robustheit gegenüber Rauschfaktoren kann alternativ auch mittels klassischer Versuchspläne untersucht werden. Die Rauschfaktoren müssen darin als Faktor und die Streuung als Zielgröße berücksichtigt werden. Die Robustheit kann anhand von Wechselwirkungen zwischen Steuer- und Rauschfaktoren beeinflusst werden.





### 8.7. Shainin® Methode

Der amerikanische Unternehmensberater Dorian Shainin (1914 - 2000) propagiert eine Vorgehensweise, die sich von den anderen hier vorgestellten Ansätzen stark unterscheidet. Im erweiterten Sinne kann aber auch die Shainin®-Methode als eine Variante der Versuchsmethodik verstanden werden, welche alternative Ansätze zur Planung, Durchführung und Auswertung von Versuchen bietet. Bei den Begriffen in Kapitel 8.7, welche mit ® und ™ gekennzeichnet sind, handelt es sich um (eingetragene) Warenzeichen von Red® X Holdings LLC.

#### 8.7.1. Motivation: Die Suche nach dem Red X®

Ausgangspunkt der Shainin®-Methode ist die Annahme, dass bei vielen Fragestellungen das sogenannte Pareto-Prinzip gilt. Das Pareto-Prinzip besagt, dass ein Phänomen, das theoretisch sehr viele Ursachen haben kann, in Wirklichkeit nur sehr wenige Ursachen hat.

Shainin® empfiehlt fünf Aussiebeverfahren um diese relevanten Ursachen zu erkennen. Sie zeichnen sich zwar durch große Einfachheit aus, sind aber nur bei speziellen Fragestellungen anwendbar. Anstatt das relevante Vorwissen und die Erfahrungen abzufragen, empfiehlt Shainin "die Teile sprechen zu lassen". Diese Empfehlung setzt also voraus, dass das betrachtete Problem einen laufenden Fertigungsprozess betrifft. Die Zielgröße wird bei dieser Methode üblicherweise als Green Y® bezeichnet. Derjenige Faktor, der den größten Einfluss (50% des Effekts) auf das Ergebnis Y hat, wird Red X® genannt.

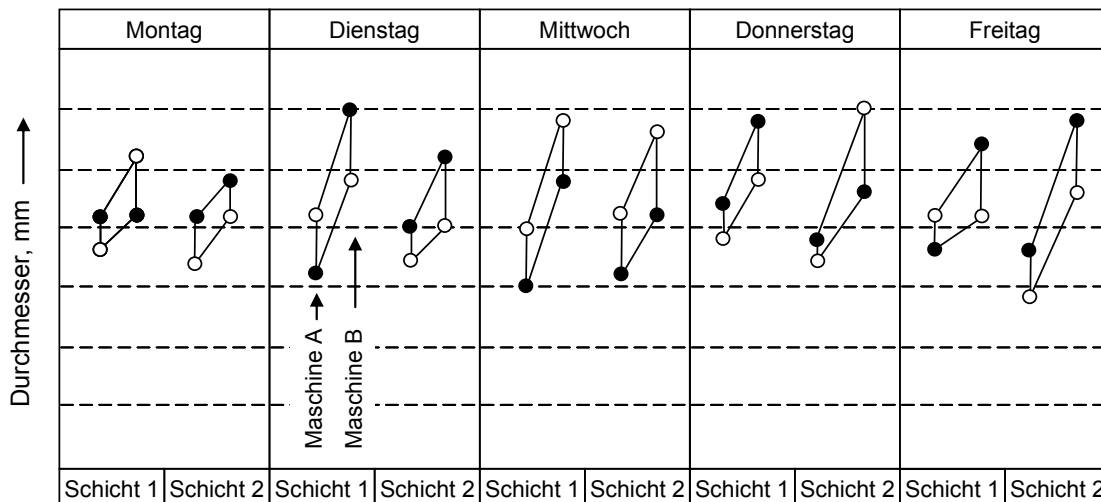
Das generelle Vorgehen des Ansatzes lässt sich wie folgt zusammenfassen:

1. Aus vielen Parametern durch systematische Beobachtung max. 20 als wesentlich identifizieren
2. Aus diesen 20 mittels einfacher Versuche die 4 wichtigsten ermitteln
3. Damit eventuell einen vollständigen Versuchsplan mit anschließender Optimierung durchführen

#### 8.7.2. Systematische Beobachtung

Zunächst geht es darum durch systematische Beobachtung „die Teile sprechen zu lassen“ und möglichst viel über den untersuchten Sachverhalt herauszufinden. Dieser Ansatz ist besonders in späten Phasen des PEP bzw. in der Fertigung sinnvoll, wenn Vorkenntnisse, Muster und Daten vorliegen.

##### 8.7.2.1. Multi-Vari-Charts



**Bild 8.5:** Multi-Vari-Chart

Eine wesentliche Technik der systematischen Beobachtung stellt die Multi-Vari-Chart-Methode dar. Diese wurde 1950 von L. Seder entwickelt. Schwankungen eines Prozesses werden graphisch dargestellt. Durch Urwerte oder vertikale Linien werden die Schwankungen ortsbedingt, chargenbedingt, zeitbedingt etc. voneinander unabhängig aufgezeichnet. Hierdurch wird unter Umständen deutlich, wo die Hauptursachen für die Schwankungen liegen. Shainin® unterscheidet:

2020-04-06 - SOCOS



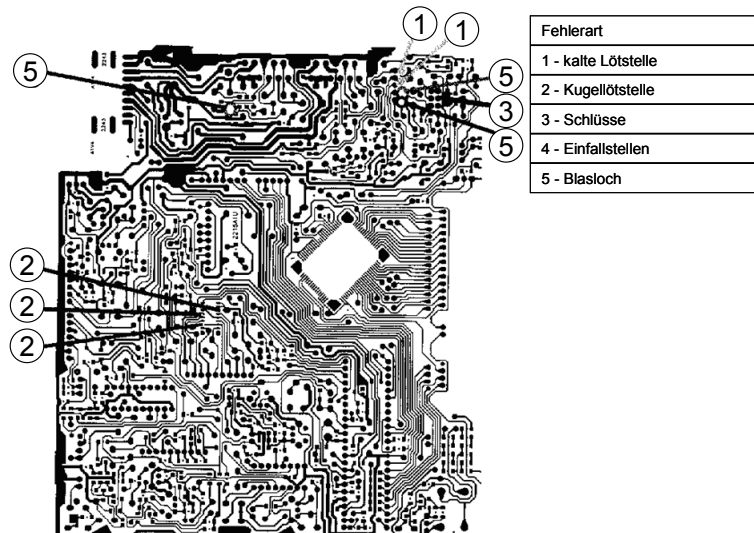
- a) Positionsabhängige Variationen, z.B. Position innerhalb einer Einheit, Maschine zu Maschine, Spannstelle zu Spannstelle, Bediener zu Bediener, Linie zu Linie, Werk zu Werk,
- b) Zyklische Variationen, z.B. Charge zu Charge oder Los zu Los,
- c) Zeitliche Variationen, z.B. Stunde zu Stunde, Schicht zu Schicht, Tag zu Tag, oder Woche zu Woche.

Ein Multi-Vari-Chart ist lediglich eine geschickte graphische Darstellung der nach Variationsfamilien geordneten Daten. Auf der Ordinate (y-Achse) wird jeweils der Messwert des untersuchten Merkmals (das Green Y®) aufgetragen. In Abbildung 8.5 sind die Mittelwerte von jeweils 3 Messwerten als Punkt eingezeichnet. Die schwarzen Punkte gehören zu Spannstelle 1, die weißen zu Spannstelle 3. Es ist ersichtlich, dass Maschine B systematisch größere Ergebnisse liefert als Maschine A.

### 8.7.2.2. Concentration Charts

Die Beobachtung der Ausprägung ist eventuell auch als „Konzentration“ der Auffälligkeit möglich, diese wird durch die Technik der concentration charts angewendet.

Sofern sich herausstellt, dass die Ursache für ein Problem innerhalb einer Einheit zu suchen ist, so bietet es sich an, den genauen Fehlerort und die Häufigkeit des bzw. der auftretenden Fehler in einem Fehlerortdiagramm aufzuzeichnen. Dazu wird eine einfache schematische Darstellung des Erzeugnisses verwendet, in welcher die Stelle markiert wird, an der ein Fehler oder eine Abweichung auftritt. Das Beispiel in Bild 8.6 stammt aus dem Heft Elementare Werkzeuge der Qualitätssicherung.



**Bild 8.6:** Concentration-Chart

### 8.7.3. Einfache Versuche

#### 8.7.3.1. Komponententausch™

Beim Komponententausch™ wird durch systematischen Austausch von Einzelteilen zwischen einer „guten“ Einheit und einer „schlechten“ Einheit das Einzelteil lokalisiert, das bei der schlechten Einheit zum Qualitätsmangel geführt hat. Bei der Reparatur defekter Geräte wird das Prinzip der Komponentensuche häufig verwendet.

Voraussetzungen:

- Anwendung hauptsächlich in der Montage
- Die Zielgröße muss mit einer Messunsicherheit messbar sein, die höchstens 20% der Toleranz oder der Prozessstreuung beträgt.
- Es muss möglich sein, gute und schlechte Teile zu zerlegen und wieder zusammenzubauen, ohne dass sich am Ergebnis "gut" bzw. "schlecht" etwas ändert.

Vorgehensweise (12 Schritte):

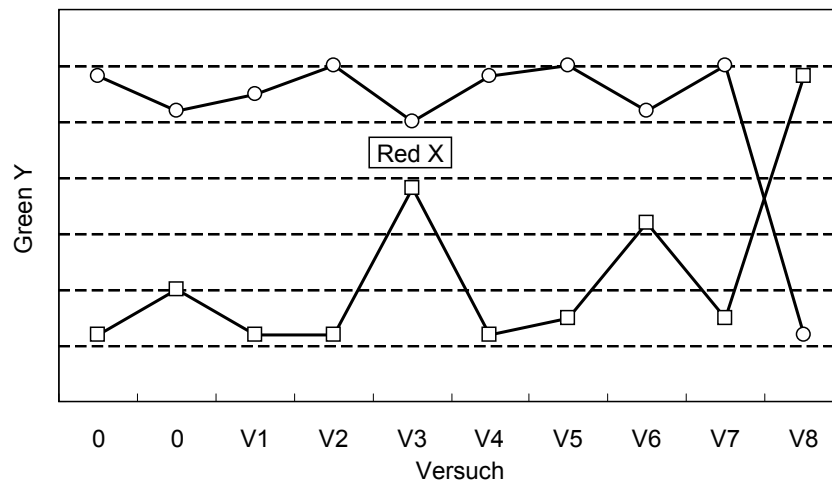


1. Auswahl von 2 Teilen aus einer Tagesproduktion, ein sehr gutes und ein sehr schlechtes (BOB = Best of the Best, WOW = Worst of the Worst). Je weiter die entsprechenden Messergebnisse auseinander liegen, um so leichter wird sich das Red X® finden lassen.
2. Zerlegen und Zusammenbauen beider Teile und zweimaliges Nachmessen der Zielgröße
3. Signifikanztest anhand der Messergebnisse. Ergebnisse des "Gut"-Teils müssen wesentlich besser sein als Ergebnisse des "Schlecht"-Teils. Kriterium:

$$D/d \geq 5, \text{ mit } D = \bar{x}_{BOB} - \bar{x}_{WOW} \text{ und } d = \frac{R_{BOB} + R_{WOW}}{2}, \text{ siehe Kap. 8.7.4.1.} \quad (8.10)$$

4. Falls  $D/d < 5$  liegt das Problem nicht bei den Komponenten sondern bei deren Montage. Vorschlag: Schrittweise die Montage untersuchen. Andernfalls müssen eine oder mehrere der Komponenten oder Teilkomponenten das Red X® sein. Dann Komponentensuche starten.
5. Komponenten nach Größe des vermuteten Einflusses ordnen
6. Tausch der Komponente A zwischen BOB und WOW (A = aussichtsreichster Kandidat)
7. Analyse des Ergebnisses
8. In jedem Fall Rückmontage, um zu sehen, ob Ausgangszustand wiederhergestellt wird
9. Wiederholung der Schritte 6, 7, 8 mit den nächsten Komponenten B, C usw.
10. Identifikation einer oder mehrerer Komponenten als das Red X®. Komponente ist Red X®, falls bei gleichzeitigem Austausch zwischen BOB und WOW eine vollständige Umkehrung erfolgt, d.h. aus einem BOB ein WOW wird und umgekehrt. Eine Wechselwirkung liegt vor, falls zwar aus einem BOB ein WOW wird oder aus einem WOW ein BOB aber nicht umgekehrt.
11. Bestätigungsversuch: Nimmt man ein BOB-Erzeugnis und ein WOW-Erzeugnis und tauscht die als unwesentlich erachteten Komponenten jeweils als Gruppe aus, so muss sich am Ergebnis BOB und Ergebnis WOW nichts ändern. Tauscht man dagegen alle als wesentlich erachteten Komponenten als Gruppe aus muss eine deutliche vollständige Umkehrung erfolgen.
12. Berechnung von Haupt- und Wechselwirkungseffekten

Versuchsergebnisse beim Komponententausch werden schematisch, nach [12] im Bild 8.7 dargestellt.



**Bild 8.7:** Komponententausch. Die ersten beiden Versuche sind lediglich eine Zerlegung mit anschließendem Zusammenbau. Beim Versuch V3 wurde das Red X® entdeckt. Der letzte Schritt ist ein Bestätigungsversuch.

### 8.7.3.2. Paarweise Vergleiche™

Die Methode der paarweisen Vergleiche™ ist dem Komponententausch™ sehr ähnlich. Sie wird angewandt, wenn Einheiten nicht in Einzelteile zerlegt werden können. Gute und schlechte Einheiten werden aus entsprechenden Grundgesamtheiten entnommen und messbare Qualitätsmerkmale werden miteinander verglichen. Qualitätsmerkmale, die sich in auffälliger Weise immer wieder



voneinander unterscheiden, sind möglicherweise für den Qualitätsunterschied der Einheiten verantwortlich. Diese Vermutung wird dann durch weitere Vergleiche überprüft.

### 8.7.3.3. Variablensuche™

Die Variablensuche™ ist ein Verfahren, das auf möglichst effektive Art aus einer mittleren Anzahl (5 - 15) von Einflussfaktoren die wichtigsten aussiebt. In gewisser Hinsicht ist die Variablensuche eine Einfaktoruntersuchung. Die Variablensuche ist nur anwendbar, wenn im Prinzip bei der Festlegung zweier Stufen bei jedem Einflussfaktor bekannt ist, welche Stufe die „bessere“ und welche Stufe die „schlechtere“ ist. Diese unabdingbare Voraussetzung ist natürlich sehr einschneidend. Das heißt, bei der Variablensuche geht es nicht um die Suche nach einem Optimum, sondern um die Frage, welche Faktoren in entscheidender Weise zum Optimum beitragen.

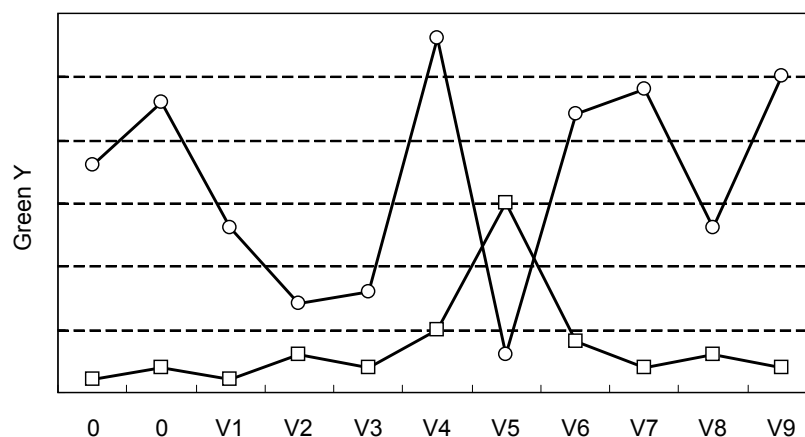
Vorgehensweise:

1. Durchführung von Vergleichsversuchen „alle Faktoren auf schlechter Stufe“  $V_S$  mit „alle Faktoren auf guter Stufe“  $V_G$ . Ist kein großer Unterschied erkennbar, dann lohnt sich eine weitere Untersuchung nicht.
2. Nun wird der Einfluss der einzelnen Faktoren getrennt untersucht. Ist  $A$  mutmaßlich der stärkste Faktor, so wird ein Versuch durchgeführt, bei dem  $A$  auf der guten und alle anderen Faktoren auf der schlechten Stufe festgelegt sind. Entspricht das Ergebnis etwa dem Ergebnis des Versuches  $V_G$  von 1., so kann angenommen werden, dass  $A$  einen dominanten Einfluss auf die Zielgröße hat.

Dies wird durch eine Gegenprobe, bei der  $A$  auf der schlechten und alle anderen auf der guten Stufe festgelegt sind, bestätigt. Im Falle der Dominanz von  $A$  müsste sich das Ergebnis von  $V_S$  wiederholen. Wenn  $A$  dominant ist, muss die Untersuchung nicht fortgesetzt werden. Wenn  $A$  nicht dominant ist, wird  $B$  in derselben Weise untersucht etc.

3. Ist keiner der Faktoren dominant, so werden die Faktoren, die einen Einfluss erkennen lassen, zusammengefasst und je einmal auf der guten bzw. schlechten Stufe und die anderen Faktoren auf der schlechten bzw. guten Stufe festgelegt.

Werden hierdurch die Ergebnisse von 1. wiederholt, so ist nachgewiesen, dass es ausreicht, die Faktoren, die bei den Einfaktorvergleichen einen Einfluss erkennen ließen, auf ihrer guten Stufe festzulegen und die übrigen Faktoren auf ihrer schlechten Stufe festgelegt werden können.



**Bild 8.8:** Variablensuche. Die ersten beiden Versuche sind lediglich eine Zerlegung mit anschließendem Zusammenbau. Im Versuch V5 wurde das Red X® entdeckt.

Ist nun die Zahl der Einflussfaktoren auf maximal 2 - 4 reduziert, empfiehlt Shainin® die bekannten voll-faktoriellen Pläne mit 2 bis 4 Faktoren auf 2 Stufen, vorzugsweise mit maximal 2 Wiederholungen pro Parameterkombination. Diese Pläne werden in der üblichen Weise ausgewertet, z.B. mit der Varianzanalyse. Auf diese Weise lässt sich feststellen, welche Einflussfaktoren bzw. welche Wechselwirkungen besonders wichtig sind. Da Shainin® relativ kleine Stichprobenumfänge empfiehlt, ist es klar, dass nur starke Effekte nachweisbar sind. Zu den empfohlenen Verfahren gehören auch Einfaktorvergleiche (B vs. C™).



Schließlich ist noch erwähnenswert, dass Shainin die Verwendung teil-faktorieller Pläne ablehnt, da bei vermengten Plänen Fehlinterpretationen möglich sind. Eine beispielhafte Variablensuche ist im Bild 8.8 dargestellt.

#### 8.7.3.4. Produkt-/Prozess-Suche

Das Ziel dieses Ansatzes ist die Trennung von wichtigen und unwichtigen Prozess-Parametern unter der Voraussetzung, dass es möglich ist, Teile während des Prozesses zu verfolgen und die während der Bearbeitung dieser Teile vorliegenden Prozessparameter zu messen.

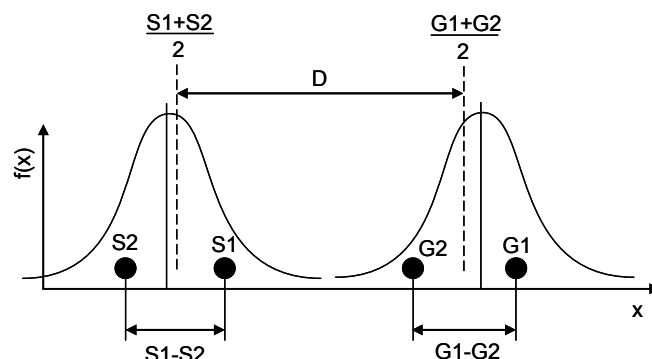
Vorgehensweise:

1. Auflistung der möglicherweise ursächlichen Prozessparameter, Reihenfolge entsprechend der vermuteten Wahrscheinlichkeit des Einflusses
2. Dokumentation welcher Prozessparameter von wem wo genau gemessen wird
3. Sicherstellen, dass Messunsicherheit < 20% der Toleranz
4. Sicherstellen, dass die tatsächlichen Parameterwerte, nicht nur die Einstellungen gemessen werden
5. Falls ein einzelner Prozessparameter bei der Überwachung keinerlei Variation zeigt, kann er von der folgenden Untersuchung ausgeschlossen werden.
6. 100%-Untersuchung der gefertigten Erzeugnisse so lange, bis mindestens 8 gute und 8 schlechte Teile gefunden wurden. Die Spannweite zwischen dem besten und dem schlechtesten Teil muss mindestens 80% der bisher beobachteten Prozessstreuung entsprechen.
7. Während der Herstellung eines Teils muss jeder der Prozessparameter aus der Liste gemessen und aufgezeichnet werden
8. (Paarweiser) Vergleich der Prozessparameterwerte, die zu "Gut"-Teilen geführt haben mit denjenigen, die zu den "schlecht"-Teilen geführt haben
9. Durchführung eines B vs. C™ –Tests.
10. Variablensuche™
11. Weitere Optimierung und ggf. Toleranzanpassung

#### 8.7.4. Weitere Tools

Als Alternative zu den klassischen Signifikanztests, wie z.B. dem t-Test, gibt Shainin® einige einfache Testverfahren an, die mit wenig Aufwand die gleichen Informationen liefern sollen. Diese und einige weitere Darstellungstools werden nachfolgend kurz erläutert.

##### 8.7.4.1. Regel D:d



**Bild 8.9:** Regel D/d.

Beim Komponententausch kann mit Hilfe der Regel D:d entschieden werden, ob die Differenz der Versuchsergebnisse mit den guten bzw. schlechten Einheiten signifikant ist oder nicht. Ein signifikanter Unterschied liegt vor, wenn  $D/d \geq 5$ . Die Regel lässt sich natürlich ebenso gut beim Variablenvergleich anwenden, um die Ergebnisse der Versuche auf der guten und der schlechten Stufe zu vergleichen.



Das folgende Schema soll die Bedeutung der Größen  $D$  und  $d$  sowie die Darstellung zur Regel  $D:d \geq 5:1$  veranschaulichen:

$$d = \frac{G_1 - G_2}{2} + \frac{S_1 - S_2}{2}, \quad D = \frac{G_1 + G_2}{2} - \frac{S_1 + S_2}{2}. \quad (8.11)$$

### 8.7.4.2. B vs. C™

B vs. C™ steht für "Better versus Current" und meint schlicht den Vergleich zwischen dem Aktuellen (current) und dem Besseren. Gegenstand des Vergleichs kann z.B. ein Produkt, ein Prozess, eine Methode oder ein Verfahren sein.

B vs. C™ ist ein einfacher (nichtparametrischer) statistischer Signifikanztest und lediglich eine Variante des Tuckey-Tests. Liegen jeweils 3 "B"-Werte und 3 "C"-Werte vor, die sich nicht überlappen, so ist deren Unterschied auf dem 95%-Niveau signifikant. Sofern eine Überlappung vorliegt, so muss der Tuckey-Test mit einem größeren Stichprobenumfang angewendet werden.

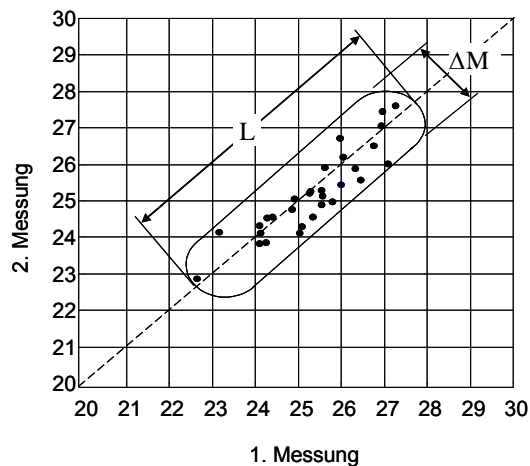
### 8.7.4.3. Isoplot®

Zur Untersuchung und Bewertung der Messunsicherheit (measurement uncertainty) im Verhältnis zum Streubereich des Teilemerkmals (product variation) schlägt Shainin® alternativ zum klassischen Vorgehen (siehe [15,16]) den sogenannten Isoplot® vor.

Dazu wird eine Stichprobe von 30 Teilen entnommen (Teile werden nummeriert). Das interessierende Merkmal jedes Teils wird gemessen (1. Messreihe). Es wird eine Wiederholmessung in zufälliger Reihenfolge durchgeführt (2. Messreihe). Die beiden an einem Teil ermittelten Ergebnisse bilden die  $x$ - und  $y$ -Koordinate des entsprechenden Punktes im Isoplot (Darstellung aller 30 Wertepaare).

Es ist darauf zu achten, dass Abszisse und Ordinate denselben Maßstab haben, die Graphik also quadratisch wird. Die Punkte werden durch eine Art "Zeppelin" eingeschlossen. Die Länge  $L$  dieses Bereichs (in mm) ist ein Maß für die Prozessstreuung (siehe Anmerkung), seine Breite  $\Delta M$  (ebenfalls in mm) ein Maß für die Messunsicherheit.

Nach Shainin® ist die Messunsicherheit genügend klein, wenn  $L/\Delta M \geq 8,5$ .



**Bild 8.10:** Isoplot®.



## 9. Anhang

### 9.1. Grundlegende Konzepte der Statistik

In diesem Kapitel werden in Kurzform einige grundlegende Konzepte vorgestellt, welche für das Verständnis des Themas Versuchsplanung erforderlich sind.

#### 9.1.1. Datenreihen und deren Kennwerte

Auch bei einer sorgfältigen Versuchsdurchführung werden die Ergebnisse eines Experiments unter gleichbleibenden Bedingungen bei einer mehrmaligen Ausführung unterschiedlich ausfallen, sie werden streuen. Eine Datenreihe

$$x_i, i \in 1..n, \tag{9.1}$$

mit  $n$ -Werten der betrachteten Zielgröße kann also allgemein als Ergebnis eines stochastischen Prozesses angesehen werden. Solche Datenreihen lassen sich im Rahmen der deskriptiven Statistik grundlegend durch folgende Kennwerte charakterisieren:

- Mittelwert

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \tag{9.2}$$

- Varianz

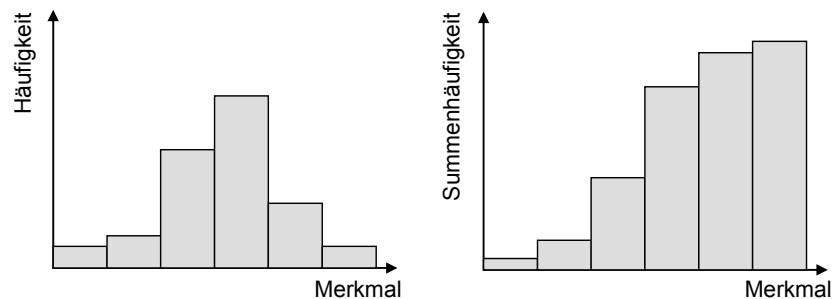
$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \tag{9.3}$$

oder deren Quadratwurzel, die Standardabweichung  $s$ , welche die Streuung der Werte um den Mittelwert charakterisieren.

Da die Einzelwerte der Datenreihe streuen, wird bei einer mehrmaligen Ermittlung der Datenreihe auch sie selbst unterschiedlich ausfallen. Damit werden auch deren Kennwerte unterschiedlich sein. Eine wichtige Erkenntnis ist also, dass der Mittelwert und die Standardabweichung ebenfalls einer Streuung unterliegen. Allerdings ist die Varianz des Mittelwertes deutlich kleiner als die Varianz der Einzelwerte und zwar umso kleiner, je größer der Stichprobenumfang ist. Man kann zeigen (vgl. auch [22], S.122), dass

$$s_{\bar{x}}^2 = \frac{s^2}{n}. \tag{9.4}$$

Datenreihen lassen sich z.B. in Histogrammen darstellen, welche bei der Klassierung der Werte der Datenreihe entstehen. In einem Histogramm werden also Häufigkeiten der auftretenden Merkmalsklassen dargestellt. Durch Aufsummierung der Klassen können Summenhäufigkeiten abgeleitet werden.



**Bild 9.1:** Histogramm und Summenhäufigkeitskurve (schematische Darstellung).



## 9.1.2. Verteilungen und deren Kennwerte

### 9.1.2.1. Verteilung und Verteilungsdichte, Momente

Eine etwas andere Sicht der Dinge entsteht, wenn die Daten als Stichprobe aus einer unendlich großen Grundgesamtheit angesehen werden. Wir deuten also

$$x_i, i \in 1 \dots n, \quad (9.5)$$

als Realisierungen einer Zufallsvariable  $X$ . Als theoretisches Modell für die Beschreibung der Grundgesamtheit aller möglichen Werte der Variable dient die Verteilungsdichtefunktion

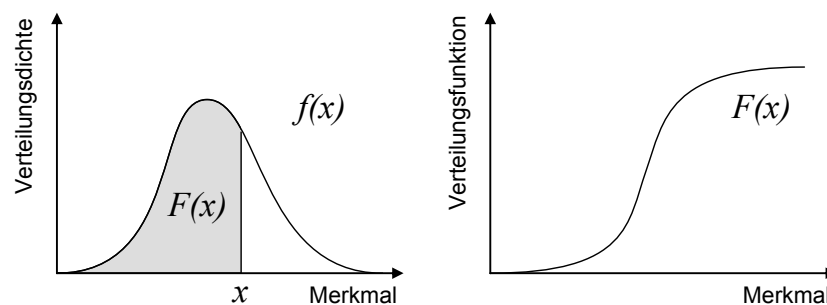
$$f(x). \quad (9.6)$$

Diese ordnet dem Wert  $x$  eine Zahl  $f(x)$  zu, ähnlich wie in einem Histogramm dem Wert  $x$  eine relative Häufigkeit zugeordnet wird. Man kann zeigen, dass falls der Umfang der betrachteten Stichprobe sehr groß und die Klassenbreiten beim Histogramm immer kleiner werden, aus dem Histogramm direkt die Verteilungsdichtefunktion abgeleitet werden kann, so dass das Histogramm als eine Näherung für die Dichtefunktion angesehen werden kann.

Aus der Summenhäufigkeitskurve kann durch Grenzübergang gegen unendlich ebenfalls eine Funktion abgeleitet werden, die Verteilungsfunktion  $F(x)$ . Sie gibt zu jedem Wert  $x$  die Wahrscheinlichkeit dafür an, dass eine Zufallsvariable  $X$  einen Wert kleiner als  $x$  annimmt. Man erhält allgemein die Verteilungsfunktion einer Verteilung, indem man über deren Dichtefunktion integriert:

$$P(X \leq x) = F(x) = \int_{-\infty}^x f(v) dv. \quad (9.7)$$

Die unabhängige Variable der Funktion  $f(v)$  wird dabei mit  $v$  bezeichnet, damit sie von der oberen Integrationsgrenze  $x$  unterschieden werden kann. Die Funktion  $F(x)$  ist also eine Funktion der oberen Integrationsgrenze,  $v$  ist hingegen eine Variable über welche integriert wird.



**Bild 9.2:** Verteilungsdichte und Verteilungsfunktion (schematische Darstellung).

Die Verteilungsfunktion ordnet dem Wert  $x$  der Variablen  $X$  eine Wahrscheinlichkeit zu, somit müssen einige Voraussetzungen dafür erfüllt werden, damit eine Funktion als Verteilungsfunktion verwendet werden kann:

- In den Grenzen muss gelten, dass

$$F(-\infty) = 0, F(+\infty) = 1, \text{ daraus folgt dass } \int_{-\infty}^{+\infty} f(v) dv = 1. \quad (9.8)$$

- Für alle  $-\infty < x < +\infty$  muss gelten, dass  $F(x) \geq 0$  sowie dass  $F(x)$  eine monoton steigende Funktion von  $x$  ist.

Die Dichte- und Verteilungsfunktion einer Verteilung haben charakteristische Kennwerte (manchmal auch Momente genannt), welche das Aussehen der Verteilung bestimmen. Die wichtigsten sind

- der Mittelwert (manchmal auch Erwartungswert genannt)

$$\mu = \int_{-\infty}^{+\infty} x \cdot f(x) dx \quad (9.9)$$





- und die Varianz

$$\sigma^2 = \int_{-\infty}^{+\infty} (x - \mu)^2 \cdot f(x) dx, \quad (9.10)$$

welche die Lage und Breite der Verteilung beschreiben.

### 9.1.2.2. Beispiele für Verteilungen

Obige Sachverhalte können am einfachsten am Beispiel der Normalverteilung verdeutlicht werden, welche die bekannte „glockenartige“ Verteilungsdichtefunktion besitzt:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}. \quad (9.11)$$

Durch Integration der Dichtefunktion erhält man die s-förmige Verteilungsfunktion der Normalverteilung

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}\left(\frac{v-\mu}{\sigma}\right)^2} dv. \quad (9.12)$$

Da das Integral nicht analytisch ausgewertet werden kann, lässt sich obige Gleichung auch nicht weiter vereinfachen.

Aus den Gleichungen ist ersichtlich, dass bei der Normalverteilung direkt die Kennwerte Mittelwert und Standardabweichung als Parameter in der Verteilungsfunktion auftreten, man bezeichnet sie oft auch mit  $N(\mu, \sigma)$ . Dies macht die Arbeit mit ihr besonders einfach. Dass dies aber nicht bei jeder Verteilung so sein muss, zeigt das Beispiel der Exponentialverteilung mit der Dichtefunktion

$$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0, \quad (9.13)$$

bei der ein Parameter  $\lambda$  auftritt. Die Beziehung zwischen Parameter und obigen Kennwerten ist gegeben durch folgende Beziehung:

$$\mu = \frac{1}{\lambda}; \quad \sigma^2 = \frac{1}{\lambda^2}. \quad (9.14)$$

Hier ist die Beziehung zwischen Verteilungsparameter und Kennwerten relativ einfach, so dass eine Umrechnung noch problemlos möglich ist. Komplizierter sieht es z.B. bei der Weibull-Verteilung aus, welche bei der Beschreibung von Lebensdauern eine wichtige Rolle spielt; dort ist eine Umrechnung nicht mehr ohne Weiteres möglich. Auf eine Darstellung wird allerdings an dieser Stelle verzichtet.

### 9.1.2.3. Kennwerte unabhängiger und abhängiger Variablen

Falls eine Variable  $Y$  eine lineare Abhängigkeit von  $n$  unabhängigen Zufallsvariablen  $X_i$  aufweist

$$Y = a_0 + \sum_{i=1}^n a_i X_i, \quad (9.15)$$

so wird  $Y$  ebenfalls eine Zufallsvariable sein. Die Verteilung von  $Y$  wird im allgemeinen nicht ohne Weiteres zu berechnen sein, insbesondere dann, wenn die Variablen  $X_i$  unterschiedliche Verteilungstypen haben. Trotzdem wird es möglich sein, Informationen über die Kennwerte der  $Y$ -Verteilung zu gewinnen, diese werden von den Kennwerten der Verteilung von  $X_i$  abhängen [22] S.150:

$$\begin{aligned} \mu_Y &= a_0 + \sum_{i=1}^n a_i \mu_{X_i} \\ \sigma_Y^2 &= \sum_{i=1}^n a_i^2 \sigma_{X_i}^2 \end{aligned} \quad (9.16)$$

Insbesondere gilt für den Mittelwert von  $n$  normalverteilten unabhängigen Variablen



$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \quad i = 1..n, \quad X_i \in N(\mu, \sigma), \quad (9.17)$$

dass dieser ebenfalls normalverteilt ist mit

$$\mu_{\bar{x}} = n \frac{\mu}{n} = \mu \quad \text{und} \quad \sigma_{\bar{x}}^2 = n \frac{\sigma^2}{n^2} = \frac{\sigma^2}{n}. \quad (9.18)$$

#### 9.1.2.4. Schätzung des Mittelwertes, Vertrauensbereich, t-Verteilung

Im Allgemeinen stellt sich die Frage, wie die Kennwerte der unbekanntes Grundgesamtheit ermittelt werden können. Es kann gezeigt werden, dass

- der Mittelwert einer Stichprobe  $\bar{y}$  als Schätzwert für den Mittelwert der Grundgesamtheit  $\mu$  verwendet werden kann sowie
- die Varianz einer Stichprobe  $s^2$  als Schätzwert für die Varianz der Grundgesamtheit  $\sigma^2$  verwendet werden kann.

Beide Schätzwerte sind zudem auch erwartungstreu, d.h. sie nähern sich bei vielen Wiederholungen den Werten der Grundgesamtheit.

Die Güte dieser Punkt-Schätzung ist zunächst allerdings offen, sie kann jedoch durch folgende Vorgehensweise bewertet werden. Falls eine Zufallsvariable  $X$  normalverteilt ist, so ist die Variable

$$u = \frac{X - \mu}{\sigma}, \quad X \in N(\mu, \sigma) \quad (9.19)$$

ebenfalls normalverteilt mit dem Mittelwert 0 und der Standardabweichung 1, man nennt diese Verteilung „Standardnormalverteilung“. Eine Abschätzung der Wahrscheinlichkeit  $P$ , mit welcher  $X$  in einem bestimmten Intervall  $\mu - u \cdot \sigma \leq X \leq \mu + u \cdot \sigma$  liegt, kann mit Hilfe des Parameters  $u$  erfolgen:

$$P(\mu - u \cdot \sigma \leq X \leq \mu + u \cdot \sigma) = 1 - \alpha. \quad (9.20)$$

Beispielsweise wird die Variable  $X$  mit einer Wahrscheinlichkeit von 68,3% im Intervall  $\mu \pm 1 \cdot \sigma$  liegen. Damit ist klar, dass  $u$  gleich dem Merkmalswert der Standardnormalverteilung bei Wahrscheinlichkeiten  $(1-\alpha/2)$  und  $\alpha/2$  ist.

Wir betrachten nun den Mittelwert von  $n$  unabhängigen normalverteilten Variablen

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \quad X_i \in N(\mu, \sigma), \quad i = 1..n, \quad \bar{X} \in N\left(\mu, \frac{\sigma}{\sqrt{n}}\right). \quad (9.21)$$

Die Größe

$$u = \sqrt{n} \frac{\bar{X} - \mu}{\sigma} \quad (9.22)$$

ist per Definition standardnormalverteilt. Allerdings wird in der Regel die Standardabweichung der Grundgesamtheit unbekannt sein, so dass sie aus einer Stichprobe geschätzt werden muss:  $\sigma \approx s$ . Man kann jedoch zeigen, dass in diesem Fall die Größe

$$t = \sqrt{n} \frac{\bar{X} - \mu}{s} \quad (9.23)$$

nicht mehr normalverteilt ist, deren Verteilung wird durch die sogenannte t-Verteilung (oder Student-Verteilung) beschrieben. Die t-Verteilung zeigt für kleine Werte des Parameters  $n$  eine größere Breite und Flankenbetonung als die Normalverteilung, für große Werte von  $n$  nähert sie sich der Normalverteilung an.

Eine Abschätzung der Wahrscheinlichkeit  $P$ , mit welcher  $\bar{X}$  in einem bestimmten Intervall liegt, kann mit Hilfe des Parameters  $t$  erfolgen. Umgekehrt, kann als sogenannter indirekter Schluss auch ein Intervall



$$\bar{X} - t \cdot \frac{s}{\sqrt{n}} \leq \mu \leq \bar{X} + t \cdot \frac{s}{\sqrt{n}} \quad (9.24)$$

angegeben werden, in dem der unbekannte Mittelwert der Grundgesamtheit liegt. Die Wahrscheinlichkeit

$$P\left[\bar{X} - t \cdot \frac{s}{\sqrt{n}} \leq \mu \leq \bar{X} + t \cdot \frac{s}{\sqrt{n}}\right] = 1 - \alpha \quad (9.25)$$

wird hier üblicherweise *Vertrauensniveau* (confidence level) und das Intervall als *Vertrauensintervall* (confidence interval) bezeichnet. Je sicherer man sein möchte, dass das Intervall den wahren Mittelwert enthält, desto höher also das Vertrauensniveau liegt, desto größer ist der Parameter  $t$  zu wählen. Tabellarische Werte von  $t$  sind im Anhang zu finden.

Analog kann auch ein Vertrauensintervall für die Varianz angegeben werden; darauf soll hier nicht weiter eingegangen werden. Details hierzu finden sich z.B. in [22].

### 9.1.3. t-Test, Vergleich von zwei Mittelwerten

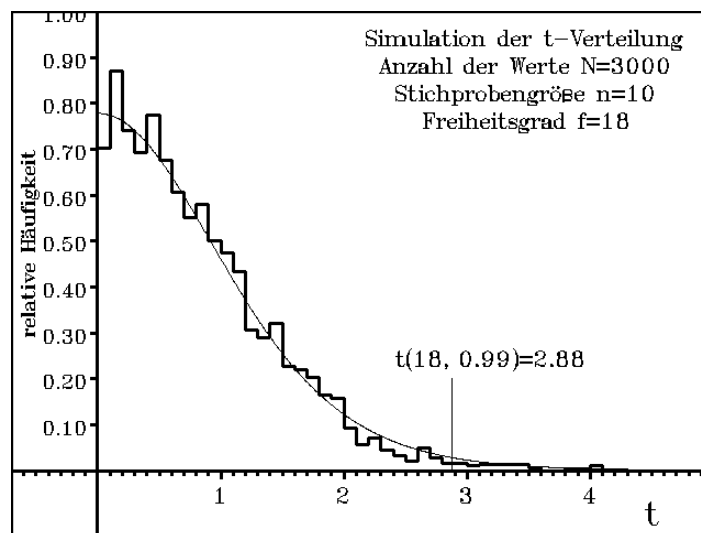
Die t-Verteilung kann auch dazu verwendet werden, um mittels eines sogenannten t-Tests zu entscheiden, ob zwei Stichproben sich signifikant unterscheiden oder nicht.

Um dies zu verdeutlichen, machen wir folgendes Gedankenexperiment: Wir entnehmen einer normalverteilten Grundgesamtheit  $N(\mu, \sigma)$  zwei Stichproben jeweils vom Umfang  $n$ , berechnen die Mittelwerte  $\bar{y}_1$  und  $\bar{y}_2$  sowie die Standardabweichungen  $s_1$  und  $s_2$  (bzw. die Varianzen  $s_1^2$  und  $s_2^2$ ) und damit den Wert

$$t = \sqrt{n} \frac{|\bar{y}_1 - \bar{y}_2|}{\sqrt{s_1^2 + s_2^2}}. \quad (9.26)$$

$t$  kann Werte zwischen 0 und  $+\infty$  annehmen. Wiederholen wir diesen Vorgang sehr oft, so erwarten wir, meist kleine Werte nahe null und nur selten sehr große Werte zu finden.

Dieses Gedankenexperiment wurde per Rechnersimulation durchgeführt. Für  $n=10$  und 3000 Stichprobenpaare ( $t$ -Werte) ergab sich das im Bild 9.3 dargestellte Histogramm.



**Bild 9.3:** Histogramm und Dichtefunktion der t-Verteilung.

Lässt man die Anzahl der Stichproben gegen unendlich und gleichzeitig die Klassenbreite gegen null streben, so nähert sich das Histogramm mehr und mehr der durchgezogenen Linie an, die die Dichtefunktion der t-Verteilung darstellt.

Die obere Grenze des 99%-Zufallsstrebereichs (Schwellenwert) ist in diesem Beispiel  $t(18; 99\%) = 2.88$ , d.h. nur in 1% aller Fälle können sich zufällig größere Werte als 2,88 ergeben. Die



Schwellenwerte der t-Verteilung sind für verschiedene Irrtumswahrscheinlichkeiten in Abhängigkeit vom Freiheitsgrad  $f=2(n-1)$  tabelliert.

Auf dem oben dargestellten Zusammenhang basiert die Vorgehensweise beim t-Test. Der t-Test ist ein statistisches Verfahren, mit dem entschieden werden kann, ob die arithmetischen Mittelwerte zweier vorliegender Messreihen jeweils vom Umfang  $n$  zu ein und derselben Grundgesamtheit gehören, so dass die Mittelwerte der Stichproben sich nur zufällig voneinander unterscheiden oder nicht, so dass der Unterschied in den Mittelwerten der Stichproben nicht zufällig ist. Als sogenannte Nullhypothese wird also angenommen, dass die Mittelwerte der jeweils zugehörigen Grundgesamtheit gleich sind.

Voraussetzungen für den Vergleich sind:

- Die Einzelwerte beider Stichproben sind repräsentativ
- Beide Stichproben stammen aus einer Normalverteilung
- Die Standardabweichung dieser Normalverteilung ist für beide Stichproben gleich. Diese Voraussetzung kann mit Hilfe eines F-Tests geprüft werden.

Das Vorgehen besteht nun darin, mit den Mittelwerten  $\bar{y}_1$  und  $\bar{y}_2$  sowie den Varianzen  $s_1^2$  und  $s_2^2$  die Testgröße

$$t = \sqrt{n} \frac{|\bar{y}_1 - \bar{y}_2|}{\sqrt{s_1^2 + s_2^2}} \quad (9.27)$$

zu berechnet.

Ergibt sich  $t > t_{2(n-1);0,99\%}$ , d.h.  $t$  liegt außerhalb des 99 %-Zufallsstreubereichs, so wird die Nullhypothese abgelehnt.

Bemerkungen:

- Der Ausdruck für die Prüfgröße  $t$  ist in dieser einfachen Form nur dann anwendbar, wenn die Stichprobenumfänge jeweils als gleich vorausgesetzt werden ( $n_1=n_2=n$ ). Eine analoge Berechnungsregel existiert allerdings auch für unterschiedliche Stichprobenumfänge.
- Der t-Test in der hier dargestellten Form testet die Nullhypothese  $\mu_1=\mu_2$  gegen die Alternative  $\mu_1 \neq \mu_2$ . Es liegt also eine zweiseitige Fragestellung vor. Aus diesem Grund steht im Ausdruck für  $t$  der Betrag der Mittelwertsdifferenz.  $t$  kann folglich nur Werte  $\geq 0$  annehmen, so dass sich die oben gezeigte Verteilung ergibt.
- Die Tabelle 2 in Kapitel 9.3 gibt entsprechend der zweiseitigen Fragestellung die Schwellenwerte 95%, 99%, und 99,9% der t-Verteilung an. Sie entsprechen den einseitigen Schwellenwerten 97,5%, 99,5% und 99,95%.

#### 9.1.4. Mindeststichprobenumfang

In der Vorbereitungsphase experimenteller Untersuchungen stellt sich dem Experimentator häufig die Frage, welcher minimale Mittelwertsunterschied  $\Delta = \bar{y}_1 - \bar{y}_2$  für ihn hinsichtlich seiner Zielsetzung (Systemoptimierung, Fertigungsvereinfachung, Kostenreduktion) überhaupt von Interesse ist, und welchen Mindeststichprobenumfang  $n$  er wählen muss, damit dieser minimale Mittelwertsabstand, sofern tatsächlich vorhanden, aufgrund der Versuchsauswertung erkennbar (signifikant) wird.

Aus dem Ausdruck für die Prüfgröße  $t$

$$t = \sqrt{n} \frac{|\bar{y}_1 - \bar{y}_2|}{\sqrt{s_1^2 + s_2^2}} \quad (9.28)$$

ist ersichtlich, dass für das "Ansprechen des t-Tests"  $n$  um so größer sein muss, je kleiner die Mittelwertsdifferenz  $\bar{y}_1 - \bar{y}_2$  ist und je größer die Varianzen  $s_1^2$  und  $s_2^2$  der beiden zu vergleichenden Messreihen sind. Dabei ist zu berücksichtigen, dass der Tabellenwert  $t_{\text{tabelle}}$  mit wachsendem Freiheitsgrad  $f = 2(n-1)$  kleiner wird.



Anschaulich bedeutet eine kleine Differenz der Mittelwerte bei gleichzeitig "großer" Varianz der Verteilungen, dass in einer graphischen Darstellung der beiden Messreihen die beiden Wertegruppen optisch nicht oder kaum noch unterscheidbar sind.

Aufgrund der vorstehenden Überlegungen wäre es also denkbar, den Mindeststichprobenumfang  $n$  grob abzuschätzen, indem man die Mittelwertsdifferenz als Vielfaches einer "mittleren Varianz"  $(s_1^2 + s_2^2)/2$  angibt und für verschiedene  $n$  die berechnete Prüfgröße  $t$  mit  $t_{\text{tabelle}}$  vergleicht (Freiheitsgrad und Signifikanzniveau beachten!). Neben diesem "Ausprobieren" gibt es jedoch eine aus statistischer Sicht exakte Herleitung des Mindeststichprobenumfangs, die wir an dieser Stelle nur grob skizzieren (Herleitung in [2]).

Beim Vergleich der Mittelwerte zweier Messreihen und der entsprechenden Testentscheidung sind zwei Arten von Irrtümern möglich.

- Im ersten Fall stammen beide Messreihen aus der gleichen Grundgesamtheit d.h., es gibt keinen signifikanten Unterschied. Entscheidet man hier aufgrund eines t-Tests, dass ein Unterschied der beiden Mittelwerte existiert, so begeht man einen Fehler 1. Art ( $\alpha$ ). Er entspricht dem Signifikanzniveau des t-Tests (zum Beispiel  $\alpha=1\%$ ).
- Liegt im zweiten Fall tatsächlich ein Unterschied der Mittelwerte vor, d.h. die Messreihen stammen aus zwei verschiedenen Grundgesamtheiten, so wird dies durch den Test nicht mit absoluter Sicherheit angezeigt. Rein zufällig kann das Testergebnis lauten, dass dieser Unterschied nicht existiert. Man spricht in diesem Fall von einem Fehler 2. Art ( $\beta$ ).

Für den Experimentator sind beide Fehlerarten unangenehm, weil er beispielsweise aufgrund des vermutlichen Effekts eines Einflussfaktors möglicherweise kostspielige weitere Untersuchungen oder sogar Änderungen in einem Produktionsprozess vorschlägt (Fehler 1. Art), bzw. weil er einen tatsächlich vorhandenen Effekt nicht erkennt und ihm die Chance für mögliche Prozessverbesserungen entgeht (Fehler 2. Art).

Der Mindeststichprobenumfang  $n$ , der notwendig ist, um einen echten Mittelwertsunterschied anhand eines Vergleichs von 2 Stichproben des gleichen Umfangs  $n_1=n_2=n$  zu erkennen, hängt entsprechend obiger Plausibilitätsbetrachtung von dem in Einheiten der Standardabweichung  $\sigma$  angegebenen Abstand der beiden Mittelwerte

$$D = \frac{\Delta}{\sigma} \tag{9.29}$$

und den Irrtumswahrscheinlichkeiten  $\alpha$  und  $\beta$  ab:

$$n = 2 \frac{(u_\alpha + u_\beta)^2}{D^2} . \tag{9.30}$$

Im konkreten Fall des Vergleichs zweier Messreihen sind die Mittelwerte  $\mu_1$  und  $\mu_2$  sowie die Standardabweichung  $\sigma$  der Grundgesamtheit (folglich auch  $D$ ) nicht bekannt. Sie werden durch die empirischen Werte  $\bar{y}_1$ ,  $\bar{y}_2$  und  $s$  geschätzt. Aus diesem Grund muss man bei der Berechnung von  $n$  nach der angegebenen Formel die t-Verteilung zugrunde legen. Demzufolge sind  $u_\alpha$  und  $u_\beta$  die Abszissenwerte  $u$ , bei denen die t-Verteilung die Werte  $\alpha$  (zweiseitig) bzw.  $\beta$  (einseitig) annimmt.

Kleinere Irrtumswahrscheinlichkeiten, d.h. kleinere Fehler 1. Art ( $\alpha$ ) und 2. Art ( $\beta$ ) bedeuten, dass sich die beiden zu vergleichenden Verteilungen und damit auch die zugehörigen Verteilungen der Mittelwerte nur wenig überlappen dürfen. Dazu muss bei gegebenem Mittelwertsabstand  $D$  der Stichprobenumfang  $n$  ausreichend groß gewählt werden.

Beabsichtigt man die Durchführung eines Versuchs anhand eines vollfaktoriellen Versuchsplans mit Faktoren auf 2 Stufen, so muss die Anzahl der Versuchsrealisierungen jeder Faktorstufenkombination so gewählt werden, dass die Gesamtanzahl Versuche

$$N_{\text{ges}} \geq 4 \frac{(u_\alpha + u_\beta)^2}{D^2} , \tag{9.31}$$

da der Effekt eines jeden Faktors in diesem Fall als Differenz der Mittelwerte zweier Stichproben vom Umfang  $n = N_{\text{tot}} / 2$  definiert wird.



### 9.1.5. Varianzanalyse

#### 9.1.5.1. F-Test, F-Verteilung

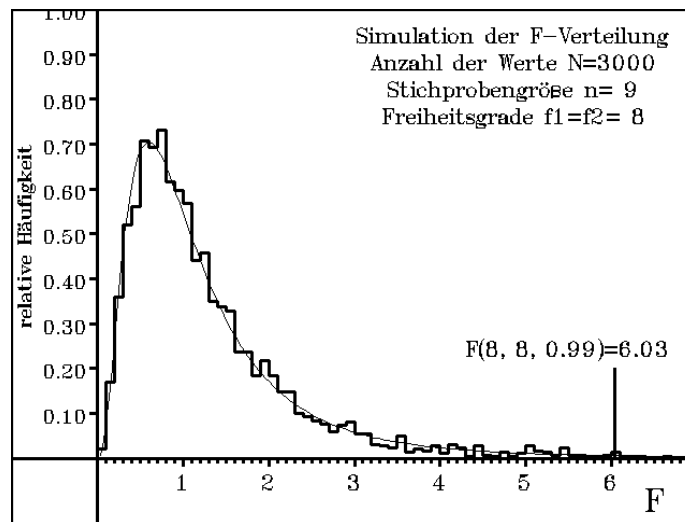
Der F-Test ist ein statistisches Verfahren, mit dem entschieden werden kann, ob die Varianzen zweier Stichproben signifikant verschieden sind. Die Funktionsweise des Tests lässt sich wie im Falle des t-Tests anhand des Ergebnisses einer Rechnersimulation erläutern.

Wir entnehmen einer normalverteilten Grundgesamtheit  $N(\mu, \sigma)$  zwei Stichproben der Umfänge  $n_1$  bzw.  $n_2$ , berechnen die Stichproben-Varianzen  $s_1^2$  und  $s_2^2$  und daraus schließlich die Größe

$$F = \frac{s_1^2}{s_2^2} \quad (9.32)$$

$F$  kann Werte zwischen 0 und  $+\infty$  annehmen, wird aber nur zufällig von 1 abweichen. Es ist plausibel, dass sich bei häufiger Wiederholung dieses Vorgangs nur selten kleine Werte nahe null bzw. sehr große Werte ergeben.

Das Ergebnis einer Rechnersimulation, bei der  $F$ -Werte für  $N=3000$  Stichprobenpaare mit Stichprobenumfang  $n_1=n_2=9$  bestimmt wurden, ist im Bild 9.4 als Histogramm dargestellt.



**Bild 9.4:** Histogramm und Dichtefunktion der F-Verteilung.

Lässt man die Anzahl der Stichproben gegen unendlich und gleichzeitig die Klassenbreite gegen null streben, so nähert sich das Histogramm der durchgezogenen Linie (Dichtefunktion der F-Verteilung) an. Die Form des Histogramms hängt von den Stichprobengrößen  $n_1$  und  $n_2$  der untersuchten Stichprobenpaare ab; dementsprechend ist der Kurvenverlauf der Dichtefunktion der F-Verteilung von den Freiheitsgraden  $f_1 = n_1 - 1$  und  $f_2 = n_2 - 1$  abhängig.

Die obere Grenze des 99%-Zufallsstrebereichs (Schwellenwert) ist im berechneten Beispiel  $F_{8,8,0,99}=6,03$ , d.h. nur in 1% aller Fälle (Irrtumswahrscheinlichkeit) ist zufällig  $s_1^2 \geq 6,03s_2^2$ .

Die Schwellenwerte der F-Verteilung sind für verschiedene Irrtumswahrscheinlichkeiten in Abhängigkeit von den Freiheitsgraden  $f_1$  und  $f_2$  im Anhang tabelliert.

Der oben dargestellte Zusammenhang macht die Vorgehensweise beim F-Test verständlich. Es soll entschieden werden, ob zwei vorliegende Messreihen vom Umfang  $n_1$  und  $n_2$  aus zwei normalverteilten Grundgesamtheiten mit gleicher Varianz (die Mittelwerte brauchen nicht bekannt zu sein) stammen oder nicht. Als Nullhypothese wird also angenommen, dass die Varianzen der zugehörigen Grundgesamtheiten gleich sind:  $\sigma_1^2 = \sigma_2^2$ .

Anschließend wird aus den Varianzen  $s_1^2$  und  $s_2^2$  der beiden Messreihen die Testgröße

$$F = \frac{s_1^2}{s_2^2} \quad (9.33)$$



berechnet und mit dem Schwellenwert der F-Verteilung verglichen. Ergibt sich  $F > F_{n_1-1; m_2-1; 0,99}$ , d.h.,  $F$  liegt außerhalb des 99%-Zufallsstrebereichs, so wird die Nullhypothese abgelehnt.

ANMERKUNG:

- Die Alternativhypothese ist  $\sigma_1^2 > \sigma_2^2$ , es handelt sich um eine einseitige Fragestellung.
- Wenn man grundsätzlich die größere der beiden Varianzen  $s_1^2$  und  $s_2^2$  über den Bruchstrich schreibt, so kann  $F$  nur Werte größer 1 annehmen; es liegt dann eine zweiseitige Fragestellung vor. Bei einer Irrtumswahrscheinlichkeit  $\alpha=1\%$  muss dann aus Tabelle 2, Kapitel 9.3, der Schwellenwert zu 99,5% ermittelt werden.

### 9.1.5.2. Varianzanalyse

Ob die Mittelwerte zweier Messreihen sich signifikant unterscheiden kann mit Hilfe des t-Tests geprüft werden. Die zu vergleichenden Messreihen können formal jeweils als Versuchsergebnisse für die beiden Stufen 1 (z.B. Material  $A$ ) und 2 (Material  $B$ ) eines einzelnen Einflussfaktors (Material) betrachtet werden. Erweitert man den Einfaktorversuch auf mehr als zwei Stufen (allgemein:  $m$  Stufen), so ist ein Vergleich der Mittelwerte mit Hilfe des t-Tests nicht mehr möglich. In diesem Fall kann eine Auswertung mittels der Varianzanalyse erfolgen.

Betrachten wir folgende Versuchsergebnisse aus  $m$ -Stufen eines Faktors mit  $n$ -Realisierungen pro Stufe:

Stufe	Ergebnisse	Mittelwert	Varianz
1	$y_{11}, y_{12}, \dots, y_{1n}$	$\bar{y}_1$	$s_1^2$
2	$y_{21}, y_{22}, \dots, y_{2n}$	$\bar{y}_2$	$s_2^2$
$m$	$y_{m1}, y_{m2}, \dots, y_{mn}$	$\bar{y}_m$	$s_m^2$

**Tabelle 9.1:** Versuchsergebnisse aus  $m$ -Stufen eines Faktors mit  $n$ -Realisierungen pro Stufe

Falls der Faktor keinen Einfluss auf das Messergebnis hat, so können alle Einzelergebnisse als aus der gleichen Grundgesamtheit stammend angesehen werden. Diese und damit auch die Mittelwerte unterliegen dann nur zufälligen Abweichungen ("Versuchsräuschen") vom gemeinsamen Mittelwert.

Im anderen Fall — der Faktor hat einen signifikanten Einfluss auf das Messergebnis — werden sich die Mittelwerte der zu den Stufen des Faktors gehörenden Verteilungen signifikant unterscheiden.

Im Rahmen der Varianzanalyse setzt man  $m$  unabhängige, normalverteilte Grundgesamtheiten mit gleicher Varianz voraus und stellt die Nullhypothese auf, dass alle Messwerte aus Grundgesamtheiten mit dem gleichen Mittelwert  $\mu_1 = \mu_2 = \dots = \mu_m = \mu$  stammen. Da gleiche Varianzen vorausgesetzt wurden, bedeutet die Nullhypothese, dass alle Messwerte aus ein und derselben Grundgesamtheit stammen.

Die Grundidee der Varianzanalyse ist die Schätzung der Varianz der Mittelwerte auf zwei unterschiedlichen Wegen. Zunächst kann die Varianz der Mittelwerte direkt aus den Versuchsdaten berechnet werden:

$$s_{\bar{y}}^2 = s_{\bar{y}}^2 = \frac{1}{m-1} \sum_{i=1}^m (\bar{y}_i - \bar{\bar{y}})^2, \text{ mit } \bar{\bar{y}} = \frac{1}{m} \sum_{i=1}^m \bar{y}_i \quad (\text{Gesamtmittelwert}) \quad (9.34)$$

Auf der anderen Seite kann die Varianz der Mittelwerte aus der Varianz der Einzelwerte geschätzt werden. Letztere ergibt sich als Mittelwert über die Varianzen von allen  $m$  Stufen zu

$$\bar{s}_y^2 = \frac{1}{m} \sum_{i=1}^m s_i^2 \quad (9.35)$$

und stellt ein Maß für die Versuchsstreuung dar. Falls es keine wahren Unterschiede zwischen den einzelnen Stufen gibt, beträgt die Varianz der Mittelwerte somit



$$s_{\bar{y}}^2 = \frac{s_y^2}{n} \quad (9.36)$$

Schließlich führt man mit der Testgröße

$$F = \frac{s_{\bar{y}}^2}{s_y^2} = \frac{ns_{\bar{y}}^2}{s_y^2} \quad (9.37)$$

einen F-Test durch (Vergleich der beiden Schätzungen) und lehnt die oben formulierte Nullhypothese ab, falls  $F > F(m-1; m(n-1); 99\%)$  (Weitere Schwellenwerte für  $F$  im Anhang).

Je größer die Prüfgröße ist, desto unwahrscheinlicher ist es, dass es sich hierbei um ein zufälligen (nicht-signifikanten) Effekt handelt, umso stärker ist also der Hinweis auf einen wahren (signifikanten) Unterschied zwischen den Stufen. Ablehnung der Nullhypothese bedeutet: es liegt ein signifikanter Unterschied der Mittelwerte der Messergebnisse für die Stufen des Faktors vor oder: der Faktor hat einen signifikanten Einfluss auf das Messergebnis.

### 9.1.5.3. Herleitung der Testgröße

Der Begriff Varianzanalyse oder Streuungszerlegung beruht auf der mit dem oben dargestellten Formalismus verbundenen Zerlegung der Varianz aller Messwerte in die beiden Anteile zufällige Streuung ("Versuchsräuschen") und systematische Abweichung der Mittelwerte. Im Folgenden wird diese Zerlegung beschrieben. Wenn  $m$  die Anzahl der Zeilen und  $n$  die Anzahl der Messwerte (Versuche) pro Zeile bezeichnen, so ist die Gesamtvarianz aller  $m \cdot n$  Messwerte gegeben durch

$$s^2 = \frac{1}{n \cdot m - 1} \cdot \sum_{i=1}^n \sum_{j=1}^m (y_{ij} - \bar{\bar{y}})^2 \quad (9.38)$$

Die Größe  $Q = (n \cdot m - 1)s^2$  heißt Summe der quadratischen Abweichungen (S.d.q.A.).

$$Q = \sum_{i=1}^n \sum_{j=1}^m (y_{ij} - \bar{\bar{y}})^2 \quad (9.39)$$

$$Q = \sum_{i=1}^n \sum_{j=1}^m (y_{ij} - \bar{y}_j + \bar{y}_j - \bar{\bar{y}})^2 \quad (\text{Erweiterung mit Null}) \quad (9.40)$$

$$Q = \sum_{i=1}^n \sum_{j=1}^m ((y_{ij} - \bar{y}_j)^2 + 2(y_{ij} - \bar{y}_j)(\bar{y}_j - \bar{\bar{y}}) + (\bar{y}_j - \bar{\bar{y}})^2) \quad (9.41)$$

Man kann zeigen, dass der mittlere Term Null ist. Daher:

$$Q = \sum_{i=1}^n \sum_{j=1}^m (y_{ij} - \bar{y}_j)^2 + \sum_{i=1}^n \sum_{j=1}^m (\bar{y}_j - \bar{\bar{y}})^2 \quad (9.42)$$

$$Q = \sum_{j=1}^m (n-1)s_j^2 + \sum_{i=1}^n (m-1)s_{\bar{y}}^2 \quad (9.43)$$

$$(n \cdot m - 1)s^2 = m(n-1)\overline{s_j^2} + n(m-1)s_{\bar{y}}^2 \quad (9.44)$$

$$Q = Q_R + Q_A \quad (9.45)$$

"Gesamtstreuung = Versuchsräuschen + Streuung der Mittelwerte"

$$\text{Freiheitsgrad von } Q_R: f_R = m(n-1) \quad (9.46)$$

$$\text{Freiheitsgrad von } Q_A: f_A = m-1 \quad (9.47)$$

$$\text{Freiheitsgrad von } Q: f = n \cdot m - 1 \quad (9.48)$$

Bilanz der Freiheitsgrade:

$$f = f_A + f_R \quad (9.49)$$





$$n \cdot m - 1 = m - 1 + m(n - 1) \quad (9.50)$$

$$n \cdot m - 1 = n \cdot m - 1 \quad (9.51)$$

Testgröße:

$$F = \frac{\frac{Q_A}{f_A}}{\frac{Q_R}{f_R}} = \frac{\frac{n(m-1)}{m-1} \frac{s_y^2}{s_y^2}}{\frac{m(n-1)}{m(n-1)} \frac{s_y^2}{s_y^2}} = \frac{ns_y^2}{s_y^2} \quad (9.52)$$

### 9.1.6. Regressionsanalyse

Bei quantitativen Faktoren auf  $n$  Stufen geht es darum, mittels Regression ein vorab definierten Modellansatz optimal an die Versuchsergebnisse anzupassen. Nehmen wir beispielhaft an, dass  $N$  Versuche und somit  $N$  Werte der Zielgröße  $y$  bei  $N$  Werten eines Faktors  $x$  vorliegen:

$$(y_i, x_i); i = 1 \dots N. \quad (9.53)$$

Für die Abhängigkeit  $y = f(x)$  vermuten wir einen linearen Zusammenhang, von dem die streuende Zielgröße nur geringfügig und zufällig um einen Betrag  $\varepsilon$  abweicht:

$$y_i = b_0 + b_1 x_i + \varepsilon_i. \quad (9.54)$$

Bei der Regression geht es nun darum, die Koeffizienten  $b_0$  und  $b_1$  so zu berechnen, dass eine möglichst gute Prognose der Zielgröße möglich wird. Mit

$$\hat{y}_i = b_0 + b_1 x_i \quad (9.55)$$

wird dabei der vom Modell prognostizierte Wert der Zielgröße an der Stelle  $x_i$  bezeichnet. Die „Güte“ der Prognose wird beim üblichen Verfahren der Kleinsten Quadrate sichergestellt, indem die Summe der quadratischen Abweichungen zwischen Versuchswert und Prognose an allen  $N$  Stellen  $x_i$  minimiert wird:

$$\sum_{i=1}^N (y_i - \hat{y}_i)^2 = \sum_{i=1}^N (y_i - (b_0 + b_1 x_i))^2 \rightarrow \min. \quad (9.56)$$

Die notwendige Bedingung für die Minimierung obiger Funktion ist, dass die ersten partiellen Ableitungen bezüglich  $b_0$  und  $b_1$  null werden. Daraus können die Koeffizienten wie folgt berechnet werden (vgl. auch [22], S.158):

$$b_1 = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^N (x_i - \bar{x})^2}, \quad b_0 = \bar{y} - b_1 \bar{x}, \quad (9.57)$$

mit den Bezeichnungen

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i, \quad \bar{y} = \frac{1}{N} \sum_{i=1}^N y_i. \quad (9.58)$$

Unter der Voraussetzung, dass

- für jede Einstellung der Einflussgrößen, die Zielgröße normalverteilt ist,
- ein linearer Zusammenhang zwischen den Einflussgrößen und dem Mittelwert der Zielgrößen besteht,  $\mu_y(x) = a_0 + a_1 x$ ,
- die Abweichungen zwischen dieser linearen Abhängigkeit nur zufälliger Natur sind, und
- die Standardabweichung welche diese zufällige Streuung charakterisiert konstant über dem Versuchsraum ist,

kann gezeigt werden, dass die mittels einer linearen Regression berechneten Schätzwerte für die Koeffizienten der Modellgleichung erwartungstreu sind, d.h. deren Mittelwerte über viele Versuche



stimmen mit den tatsächlichen Werten überein und es können Vertrauensbereiche angegeben und damit auch Signifikanzaussagen über die Koeffizienten gemacht werden, vgl. [1], S.171ff.

Um die Güte der Anpassung zu bewerten, kann die Summe der quadratischen Abweichungen (S.d.q.A.) der  $y$ -Werte wie folgt zerlegt werden:

$$\sum_{i=1}^N (y_i - \bar{y})^2 = \sum_{i=1}^N (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (9.59)$$

$$Q = Q_{\text{Regr}} + Q_{\text{Resid}}$$

Der erste Term kann durch die Regressionsgerade erklärt werden, der zweite Term ist der Reststreuung (Residuen) zugeordnet. Die Methode der kleinsten Quadrate besteht wie erwähnt darin, den zweiten Term zu minimieren. Der sogenannte RMS-Error (root mean square error)

$$s_R = \sqrt{\frac{Q_{\text{Resid}}}{f_{\text{Resid}}}} \quad (9.60)$$

auch Standardfehler bzw. Standardabweichung der Regression genannt, eignet sich somit gut, das Ergebnis dieser Minimierung zu charakterisieren. Der Freiheitsgrad  $f_{\text{Resid}} = N - p$  hängt von der Anzahl  $p$  der freien Koeffizienten ab (2 im linearen Fall).

Die Güte der Anpassung ist umso höher, je größer der Anteil ist, welcher durch die Regression erklärt wird. Somit ist es auch naheliegend, folgenden Quotienten als Gütemaß zu verwenden:

$$R^2 = \frac{Q_{\text{Regr}}}{Q} \quad (9.61)$$

Diese Größe heißt Bestimmtheitsmaß und nimmt Werte immer zwischen 0 und 1 an. Werte nahe an 1 sprechen für einen hohen Anteil der Gesamtstreuung, den die Regression erklärt. Alternativ kann auch ein weiteres Maß, der sogenannte Korrelationskoeffizient verwendet werden:

$$R = \text{sgn}(b_1) \cdot \sqrt{R^2} \quad (9.62)$$

Der Korrelationskoeffizient nimmt Werte zwischen -1 und +1 an.

Ob der Anteil der Gesamtstreuung, der durch die Regression erklärt wird, signifikant größer ist, als der Anteil, welcher der Reststreuung zugeordnet wird, kann anhand eines F-Testes mit der Testgröße

$$F_{\text{RegrT}} = \frac{Q_{\text{Regr}} / f_{\text{Regr}}}{Q_{\text{Resid}} / f_{\text{Resid}}} = \frac{Q_{\text{Regr}} / f_{\text{Regr}}}{s_R^2} \quad (9.63)$$

analysiert werden. Mit  $f_{\text{Regr}} = p - 1$  und  $f_{\text{Resid}} = N - p$  wurden die entsprechenden Freiheitsgrade bezeichnet,  $p$  ist die Anzahl der freien Modellparameter. Die Bewertung erfolgt wie bei jedem F-Test: falls  $F_{\text{RegrT}} > F(f_{\text{Regr}}; f_{\text{Resid}}; 99\%)$  wird die Nullhypothese „Der Streuungsanteil, welcher durch Regression erklärt wird ist gleich hoch dem Streuungsanteil der Residuen“ abgelehnt, so dass die Regression einen Teil der Streuung signifikant erklärt. Schwellenwerte für  $F$  sind im Anhang zu finden.

Der Term  $Q_{\text{Resid}}$  ist, wie erwähnt, den Residuen (Reststreuung) zugeordnet, also jenem Teil in der S.d.q.A. der  $y$ -Werte, der nicht durch Regression erklärt wird. Es stellt sich prinzipiell die Frage, ob dieser Anteil auf Unzulänglichkeiten des Modells oder der reinen Versuchstreuung zurückzuführen ist. Im ersten Fall kann eine weitere Reduzierung der Reststreuung durch hinzufügen weiterer (höherer) Terme im Modell erreicht werden. Im zweiten Fall wäre eine Verbesserung des Modells nur durch Berücksichtigung bisher unbekannter Störgrößen möglich. Um diese Fälle unterscheiden zu können, ist an mindestens einer Stelle im Versuchsraum eine mehrfach-Realisierung (Versuchswiederholung) erforderlich, welche die Abschätzung des reinen Versuchsfehlers erlaubt. Angenommen, die  $N$  Versuche wurden so ausgeführt, dass der  $i$ -te von  $m$  Versuchspunkte  $n_i$ -mal wiederholt wurde. Folgende Zerlegung von  $Q_{\text{Resid}}$  kann nun durchgeführt werden:

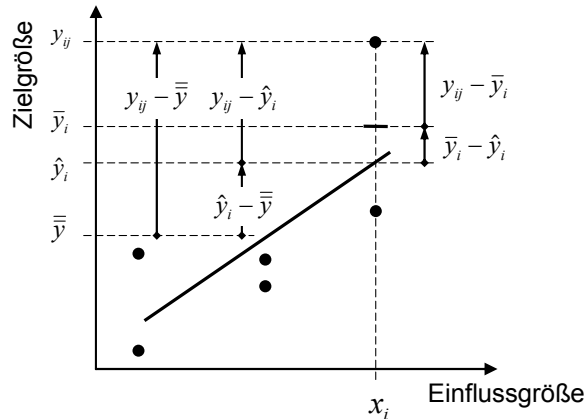
$$\sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \hat{y}_i)^2 = \sum_{i=1}^m n_i (\bar{y}_i - \hat{y}_i)^2 + \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 \quad (9.64)$$

$$Q_{\text{Resid}} = Q_{\text{LoF}} + Q_{\text{PE}}$$

Der erste Anteil  $Q_{\text{LoF}}$  entspricht der gewichteten S.d.q.A. des Gruppenmittelwertes aller Realisierungen im  $i$ -ten Versuchspunkt vom prognostizierten Wert. Er wird „Modellschwäche“ (Lack of Fit) genannt.



Der zweite Anteil  $Q_{PE}$  entspricht der S.d.q.A. der Einzelwerte vom jeweiligen Gruppenmittelwert aller Realisierungen im  $i$ -ten Versuchspunkt. Dieser Term hängt nicht vom anhand des vorliegenden Modells prognostizierten Wert ab und wird deswegen „reiner Versuchsfehler“ (Pure Error) genannt. Die Beziehungen werden im Bild 9.5 grafisch veranschaulicht:



**Bild 9.5:** Komponenten in einer Regressions-Zerlegung der S.d.q.A. (schematisch)

Die Güte des Modells hinsichtlich einer ausreichend hohen Ordnung des verwendeten Ansatzes (nicht jedoch hinsichtlich der Berücksichtigung aller relevanten Faktoren!) kann anhand eines F-Testes mit der Testgröße

$$F_{LoFT} = \frac{Q_{LoF} / f_{LoF}}{Q_{PE} / f_{PE}} \tag{9.65}$$

analysiert werden. Dieser Test wird auch „Modellschwäche-Test“ (Lack-of-Fit-Test) genannt. Mit  $f_{LoF} = m - p$  und  $f_{PE} = N - m$  wurden die entsprechenden Freiheitsgrade bezeichnet,  $p$  ist die Anzahl der freien Modellparameter. Falls  $F_{LoFT} > F(f_{LoF}; f_{PE}; 99\%)$  wird die Nullhypothese „Streuung aufgrund der Modellschwäche und reine Versuchsstreuung sind gleich“ abgelehnt, so dass die Modellschwäche signifikant größer als die reine Versuchsstreuung ist (Schwellenwerte für  $F$  im Anhang).

	Freiheitsgrad	S.d.q.A.	Varianz	F-Wert	p-Wert
Regression	$f_{Regr}$	$Q_{Regr}$	$s_{Regr}^2 = \frac{Q_{Regr}}{f_{Regr}}$	$F_{RegrT} = \frac{s_{Regr}^2}{s_R^2}$	$p_{Regr}$
Residuen	$f_{Resid} = f_{LoF} + f_{PE}$	$Q_{Resid} = Q_{LoF} + Q_{PE}$	$s_R^2 = \frac{Q_{Resid}}{f_{Resid}}$		
Modellschwäche (Lack-of-Fit)	$f_{LoF}$	$Q_{LoF}$	$s_{LoF}^2 = \frac{Q_{LoF}}{f_{LoF}}$	$F_{LoFT} = \frac{s_{LoF}^2}{s_{PE}^2}$	$p_{LoFT}$
reiner Versuchsfehler (Pure Error)	$f_{PE}$	$Q_{PE}$	$s_{PE}^2 = \frac{Q_{PE}}{f_{PE}}$		
Summe	$f = N - 1 = f_{Regr} + f_{Resid}$	$Q = Q_{Regr} + Q_{Resid}$	$s^2 = s_{Regr}^2 + s_R^2$		

**Tabelle 9.2:** ANOVA-Tabelle der Regression

Bei manchem statistischen Tool wird zusätzlich zur Testgröße  $F_{LoFT}$  bzw.  $F_{RegrT}$  eine entsprechende Größe  $p_{Regr}$  bzw.  $p_{LoFT}$  zur Beurteilung der Signifikanz angegeben, bei der gilt

$$F_{RegrT} = F(1 - p_{RegrT}; f_1; f_2) \text{ bzw. } F_{LoFT} = F(1 - p_{LoFT}; f_1; f_2). \tag{9.66}$$

Diese Größe kann als „Irrtumswahrscheinlichkeit“ (Fehler 1. Art) interpretiert werden: je kleiner diese Wahrscheinlichkeit ist, desto signifikanter wird die Regression bzw. die Modellschwäche sein. Die Bewertung erfolgt analog zur Testgröße: ein signifikantes Ergebnis liegt vor, wenn z.B.  $p_{Regr} \leq 1\%$ , ein hoch-signifikantes, wenn  $p_{Regr} \leq 0,1\%$ .



Die beschriebenen Größen werden typischerweise in einer Tabelle zusammengefasst. Diese wird analog zur Varianzanalyse-Tabelle als (globale) ANOVA-Tabelle des Regressionsmodells bezeichnet, siehe Tabelle 9.2.

## 9.2. Softwaretools

Falls anhand der Versuchsdaten eine Modellbildung mit mehreren Faktoren durchgeführt werden soll, sind sowohl das Aufstellen der Versuchspläne als auch die meisten Auswertungen umständlich und von vielen Anwendern ohne Softwarehilfe nicht zu bewerkstelligen. Insbesondere bei Anwendern ohne viel Hintergrundwissen wird hier auch ein Tabellenkalkulationsprogramm wie MS-Excel nicht unbedingt weiterhelfen. Mancher Anwender wird auch bei einer linearen Regression oder einer Varianzanalyse mit einem Faktor Softwareunterstützung brauchen. Für die Versuchsplanung und -auswertung ist deswegen generell die Anwendung geeigneter Software zu empfehlen. Im Allgemeinen können folgende Tools empfohlen werden:

- **Minitab®**

Hierbei handelt es sich um eine Standard-Lösung für statistische Auswertungen im Rahmen von Six-Sigma-Initiativen, welche auch DoE-Funktionalitäten besitzt und für die Aufstellung und Auswertung von DoE-Versuchsplänen verwendet werden kann. Diese Lösung ist vor allem für gelegentliche DoE-Aufgaben und Anwender mit bereits vorliegender breiter Erfahrung mit diesem Tool sinnvoll. Für Minitab liegt eine Konzernlizenz vor.

Weitere Info im Internet unter <http://www.additive-net.de/software/minitab>

- **Cornerstone®**

Hier handelt es sich um ein auf DoE spezialisiertes Softwaretool, mit dem auch umfangreichere DoE-Untersuchungen geplant und ausgewertet werden können. Vor allem fortgeschrittene Anwender können von einigen auf DoE optimierten Bearbeitungs- und Darstellungsmöglichkeiten profitieren, welche eine enorme Hilfe bei der Bearbeitung reeller experimenteller Daten leisten. Eine Konzernlizenz liegt nicht vor.

Weitere Info sowie eine Testlizenz für einen Monat ist im Internet zu finden unter <http://www.versuchsplanung.de/Versuchsplanung/Cornerstone>

- **ASCMO®**

Hierbei handelt es sich um ein DoE-Tool, das bei Bosch seit 2004 entwickelt wird. Eine Besonderheit der Software ist, dass die verwendete Modellgleichung nicht auf Polynomen, sondern auf Funktionen mit lokalen Trägern basiert, was sehr hilfreich für die response-surface Modellierung und der detaillierten Optimierung eines komplexen Verhaltens sein kann. Die Software ist als Matlab- und Standalone-Variante verfügbar. Ein Betrieb ist interaktiv über eine Benutzeroberfläche (welche auch hilfreiche Bearbeitungs- und Darstellungsmöglichkeiten wie bei Cornerstone beinhaltet) oder im Batch-Modus über Skripte möglich. Für Bosch-interne Anwender ist die Software derzeit kostenlos.

Weitere Infos über [PJ-ASCTeam-Mailbox@de.bosch.com](mailto:PJ-ASCTeam-Mailbox@de.bosch.com)

- **Optislang®**

Dies ist ein Softwaretool für Sensitivitätsstudien, Optimierung und Robustheitsuntersuchungen im CAE-Bereich, welches auch DoE-Funktionalitäten vor allem im Bereich Computerexperimente besitzt. Das Tool weist Schnittstellen zu vielen numerischen Simulationstools auf, welche als Solver automatisch für deterministische oder stochastische Optimierungen angesteuert werden können. Für Optislang liegt eine Konzernlizenz vor.

Weitere Info im Internet unter <http://www.dynardo.de/software/optislang/>



**9.3. Tabellen**

Tabelle 1:

Schwellenwerte der t-Verteilung (zweiseitig)

<i>f</i>	Aussagesicherheit		
	95%	99%	99,9%
1	12,7	63,7	636,6
2	4,3	9,93	31,6
3	3,18	5,84	12,9
4	2,78	4,60	8,61
5	2,57	4,03	6,87
6	2,45	3,71	5,96
7	2,37	3,50	5,41
8	2,31	3,36	5,04
9	2,26	3,25	4,78
10	2,23	3,17	4,59
11	2,20	3,11	4,44
12	2,18	3,06	4,32
13	2,16	3,01	4,22
14	2,15	2,98	4,14
15	2,13	2,95	4,07
16	2,12	2,92	4,02
17	2,11	2,90	3,97
18	2,10	2,88	3,92
19	2,09	2,86	3,88
20	2,09	2,85	3,85
25	2,06	2,79	3,73
30	2,04	2,75	3,65
35	2,03	2,72	3,59
40	2,02	2,70	3,55
45	2,01	2,69	3,52
50	2,01	2,68	3,50
100	1,98	2,63	3,39
200	1,97	2,60	3,34
300	1,97	2,59	3,32
400	1,97	2,59	3,32
500	1,97	2,59	3,31
$\infty$	1,96	2,58	3,30

2020-04-06 - SOCOS



Tabelle 2:

Schwellenwerte der F-Verteilung (einseitig zum Wert 95%)

$f_2$	$f_1=1$	$f_1=2$	$f_1=3$	$f_1=4$	$f_1=5$	$f_1=6$	$f_1=7$	$f_1=8$	$f_1=9$
1	161	200	216	225	230	234	237	239	241
2	18,5	19,0	19,2	19,2	19,3	19,3	19,4	19,4	19,4
3	10,1	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21
32	4,15	3,30	2,90	2,67	2,51	2,40	2,31	2,24	2,19
34	4,13	3,28	2,88	2,65	2,49	2,38	2,29	2,23	2,17
36	4,11	3,26	2,87	2,63	2,48	2,36	2,28	2,21	2,15
38	4,10	3,24	2,85	2,62	2,46	2,35	2,26	2,19	2,14
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04
70	3,98	3,13	2,74	2,50	2,35	2,23	2,14	2,07	2,02
80	3,96	3,11	2,72	2,49	2,33	2,21	2,13	2,06	2,00
90	3,95	3,10	2,71	2,47	2,32	2,20	2,11	2,04	1,99
100	3,94	3,09	2,70	2,46	2,31	2,19	2,10	2,03	1,97
150	3,90	3,06	2,66	2,43	2,27	2,16	2,07	2,00	1,94
200	3,89	3,04	2,65	2,42	2,26	2,14	2,06	1,98	1,93
1000	3,85	3,00	2,61	2,38	2,22	2,11	2,02	1,95	1,89

2020-04-06 - SOCCOS



Tabelle 2 (Fortsetzung)

Schwellenwerte der F-Verteilung (einseitig zum Wert 95%)

$f_2$	$f_1=10$	$f_1=15$	$f_1=20$	$f_1=30$	$f_1=40$	$f_1=50$	$f_1=100$	$f_1 \rightarrow \infty$
1	242	246	248	250	251	252	253	254
2	19,4	19,4	19,4	19,5	19,5	19,5	19,5	19,5
3	8,79	8,70	8,66	8,62	8,59	8,58	8,55	8,53
4	5,96	5,86	5,80	5,75	5,72	5,70	5,66	5,63
5	4,74	4,62	4,56	4,50	4,46	4,44	4,41	4,37
6	4,06	3,94	3,87	3,81	3,77	3,75	3,71	3,67
7	3,64	3,51	3,44	3,38	3,34	3,32	3,27	3,23
8	3,35	3,22	3,15	3,08	3,04	3,02	2,97	2,93
9	3,14	3,01	2,94	2,86	2,83	2,80	2,76	2,71
10	2,98	2,85	2,77	2,70	2,66	2,64	2,59	2,54
11	2,85	2,72	2,65	2,57	2,53	2,51	2,46	2,40
12	2,75	2,62	2,54	2,47	2,43	2,40	2,35	2,30
13	2,67	2,53	2,46	2,38	2,34	2,31	2,26	2,21
14	2,60	2,46	2,39	2,31	2,27	2,24	2,19	2,13
15	2,54	2,40	2,33	2,25	2,20	2,18	2,12	2,07
16	2,49	2,35	2,28	2,19	2,15	2,12	2,07	2,01
17	2,45	2,31	2,23	2,15	2,10	2,08	2,02	1,96
18	2,41	2,27	2,19	2,11	2,06	2,04	1,98	1,92
19	2,38	2,23	2,16	2,07	2,03	2,00	1,94	1,88
20	2,35	2,20	2,12	2,04	1,99	1,97	1,91	1,84
22	2,30	2,15	2,07	1,98	1,94	1,91	1,85	1,78
24	2,25	2,11	2,03	1,94	1,89	1,86	1,80	1,73
26	2,22	2,07	1,99	1,90	1,85	1,82	1,76	1,69
28	2,19	2,04	1,96	1,87	1,82	1,79	1,73	1,65
30	2,16	2,01	1,93	1,84	1,79	1,76	1,70	1,62
32	2,14	1,99	1,91	1,82	1,77	1,74	1,67	1,59
34	2,12	1,97	1,89	1,80	1,75	1,71	1,65	1,57
36	2,11	1,95	1,87	1,78	1,73	1,69	1,62	1,55
38	2,09	1,94	1,85	1,76	1,71	1,68	1,61	1,53
40	2,08	1,92	1,84	1,74	1,69	1,66	1,59	1,51
50	2,03	1,87	1,78	1,69	1,63	1,60	1,52	1,44
60	1,99	1,84	1,75	1,65	1,59	1,56	1,48	1,39
70	1,97	1,81	1,72	1,62	1,57	1,53	1,45	1,35
80	1,95	1,79	1,70	1,60	1,54	1,51	1,43	1,32
90	1,94	1,78	1,69	1,59	1,53	1,49	1,41	1,30
100	1,93	1,77	1,68	1,57	1,52	1,48	1,39	1,28
150	1,89	1,73	1,64	1,53	1,48	1,44	1,34	1,22
200	1,88	1,72	1,62	1,52	1,46	1,41	1,32	1,19
1000	1,84	1,68	1,58	1,47	1,41	1,36	1,26	1,08

2020-04-06 - SOCOS



Tabelle 2 (Fortsetzung)

Schwellenwerte der F-Verteilung (einseitig zum Wert 99%)

$f_2$	$f_1=1$	$f_1=2$	$f_1=3$	$f_1=4$	$f_1=5$	$f_1=6$	$f_1=7$	$f_1=8$	$f_1=9$
1	4052	4999	5403	5625	5764	5859	5928	5982	6022
2	98,5	99,0	99,2	99,3	99,3	99,3	99,4	99,4	99,4
3	34,1	30,8	29,5	28,7	28,2	27,9	27,7	27,5	27,3
4	21,2	18,0	16,7	16,0	15,5	15,2	15,0	14,8	14,7
5	16,3	13,3	12,1	11,4	11,0	10,7	10,5	10,3	10,2
6	13,7	10,9	9,78	9,15	8,75	8,47	8,26	8,10	7,98
7	12,2	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72
8	11,3	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91
9	10,6	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35
10	10,0	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19
14	8,86	6,51	5,56	5,04	4,70	4,46	4,28	4,14	4,03
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18
28	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07
32	7,50	5,34	4,46	3,97	3,65	3,43	3,26	3,13	3,02
34	7,44	5,29	4,42	3,93	3,61	3,39	3,22	3,09	2,98
36	7,40	5,25	4,38	3,89	3,57	3,35	3,18	3,05	2,95
38	7,35	5,21	4,34	3,86	3,54	3,32	3,15	3,02	2,92
40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89
50	7,17	5,06	4,20	3,72	3,41	3,19	3,02	2,89	2,79
60	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72
70	7,01	4,92	4,08	3,60	3,29	3,07	2,91	2,78	2,67
80	6,96	4,88	4,04	3,56	3,26	3,04	2,87	2,74	2,64
90	6,93	4,85	4,01	3,54	3,23	3,01	2,84	2,72	2,61
100	6,90	4,82	3,98	3,51	3,21	2,99	2,82	2,69	2,59
150	6,81	4,75	3,92	3,45	3,14	2,92	2,76	2,63	2,53
200	6,76	4,71	3,88	3,41	3,11	2,89	2,73	2,60	2,50
1000	6,66	4,63	3,80	3,34	3,04	2,82	2,66	2,53	2,43

2020-04-06 - SOCOS





Tabelle 2 (Fortsetzung)

Schwellenwerte der F-Verteilung (einseitig zum Wert 99%)

$f_2$	$f_1=10$	$f_1=15$	$f_1=20$	$f_1=30$	$f_1=40$	$f_1=50$	$f_1=100$	$f_1 \rightarrow \infty$
1	6056	6157	6209	6261	6287	6300	6330	6366
2	99,4	99,4	99,4	99,5	99,5	99,5	99,5	99,5
3	27,2	26,9	26,7	26,5	26,4	26,4	26,2	26,1
4	14,5	14,2	14,0	13,8	13,7	13,7	13,6	13,5
5	10,1	9,72	9,55	9,38	9,29	9,24	9,13	9,02
6	7,87	7,56	7,40	7,23	7,14	7,09	6,99	6,88
7	6,62	6,31	6,16	5,99	5,91	5,86	5,75	5,65
8	5,81	5,52	5,36	5,20	5,12	5,07	4,96	4,86
9	5,26	4,96	4,81	4,65	4,57	4,52	4,42	4,31
10	4,85	4,56	4,41	4,25	4,17	4,12	4,01	3,91
11	4,54	4,25	4,10	3,94	3,86	3,81	3,71	3,60
12	4,30	4,01	3,86	3,70	3,62	3,57	3,47	3,36
13	4,10	3,82	3,66	3,51	3,43	3,38	3,27	3,17
14	3,94	3,66	3,51	3,35	3,27	3,22	3,11	3,00
15	3,80	3,52	3,37	3,21	3,13	3,08	2,98	2,87
16	3,69	3,41	3,26	3,10	3,02	2,97	2,86	2,75
17	3,59	3,31	3,16	3,00	2,92	2,87	2,76	2,65
18	3,51	3,23	3,08	2,92	2,84	2,78	2,68	2,57
19	3,43	3,15	3,00	2,84	2,76	2,71	2,60	2,49
20	3,37	3,09	2,94	2,78	2,69	2,64	2,54	2,42
22	3,26	2,98	2,83	2,67	2,58	2,53	2,42	2,31
24	3,17	2,89	2,74	2,58	2,49	2,44	2,33	2,21
26	3,09	2,82	2,66	2,50	2,42	2,36	2,25	2,13
28	3,03	2,75	2,60	2,44	2,35	2,30	2,19	2,06
30	2,98	2,70	2,55	2,39	2,30	2,25	2,13	2,01
32	2,93	2,66	2,50	2,34	2,25	2,20	2,08	1,96
34	2,89	2,62	2,46	2,30	2,21	2,16	2,04	1,91
36	2,86	2,58	2,43	2,26	2,17	2,12	2,00	1,87
38	2,83	2,55	2,40	2,23	2,14	2,09	1,97	1,84
40	2,80	2,52	2,37	2,20	2,11	2,06	1,94	1,80
50	2,70	2,42	2,27	2,10	2,01	1,95	1,82	1,68
60	2,63	2,35	2,20	2,03	1,94	1,88	1,75	1,60
70	2,59	2,31	2,15	1,98	1,89	1,83	1,70	1,54
80	2,55	2,27	2,12	1,94	1,85	1,79	1,66	1,49
90	2,52	2,24	2,09	1,92	1,82	1,76	1,62	1,46
100	2,50	2,22	2,07	1,89	1,80	1,73	1,60	1,43
150	2,44	2,16	2,00	1,83	1,73	1,66	1,52	1,33
200	2,41	2,13	1,97	1,79	1,69	1,63	1,48	1,28
1000	2,34	2,06	1,90	1,72	1,61	1,54	1,38	1,11

2020-04-06 - SOCCOS



## 10. Literatur

- [1] W. Kleppmann: Taschenbuch Versuchsplanung. Produkte und Prozesse Optimieren, Hanser Verlag, München, 1998
- [2] E. Scheffler: Einführung in die Praxis der statistischen Versuchsplanung, Deutscher Verlag für Grundstoffindustrie, Leipzig, 1986
- [3] J. Krottmaier: Versuchsplanung. Der Weg zur Qualität des Jahres 2000, Verlag Industrielle Organisation Zürich, Verlag TÜV Rheinland, Köln, 1990
- [4] H. Petersen: Grundlagen der statistischen Versuchsplanung, Band 2, ecomed Verlag Landsberg/Lech, 1991
- [5] W.G. Cochran, G.M. Cox: Experimental Designs, Wiley, New York, 1957
- [6] VDA Reihe Qualitätsmanagement in der Automobilindustrie, Band 4 Sicherung der Qualität während der Produktrealisierung – Methoden und Verfahren, Kapitel 5: Versuchsmethodik (DoE), 4. Auflage 2003
- [7] VDA Reihe Qualitätsmanagement in der Automobilindustrie, Band 3 „Zuverlässigkeitssicherung bei Automobilherstellern und Lieferanten“, Kapitel 2 „Zuverlässigkeitsmethoden und Hilfsmittel“
- [8] VDA Reihe Qualitätsmanagement in der Automobilindustrie, Band 4 Sicherung der Qualität während der Produktrealisierung – Methoden und Verfahren“, Kapitel 2 „Entwicklungsabläufe“
- [9] ISO/TS 16949 „Qualitätsmanagementsysteme - Besondere Anforderungen bei Anwendung von ISO 9001:2000 für die Serien- und Ersatzteilproduktion in der Automobilindustrie“
- [10] Chrysler, Ford, GM: QS-9000
- [11] R. A. Fisher: The Design of Experiments, Oliver&Boyd, London 1935
- [12] K. R. Bhole: Qualität – Der Weg zur Weltspitze, IQM, Großbottwar, 1990
- [13] H. Bossel: Systeme, Dynamik, Simulation. Modellbildung, Analyse und Simulation komplexer Systeme, Books on Demand, Norderstedt, 2004
- [14] G. Ropohl: Eine Systemtheorie der Technik, Hanser Verlag, München, 1979
- [15] Qualitätsmanagement in der Bosch-Gruppe, Heft 8: Messunsicherheit
- [16] Qualitätsmanagement in der Bosch-Gruppe, Heft 10: Fähigkeit von Mess- und Prüfprozessen
- [17] E. Haibach: Betriebsfestigkeit, VDI-Verlag, Düsseldorf, 1989
- [18] J. Riedel: Gewichtsoptimierung eines Kreuzfahrtschiffes unter Spannungsrestriktionen, Institutskolloquium Bauhaus-Universität Weimar, 2000, (siehe [www.optislang.de/Projekte/Anlagenbau.htm](http://www.optislang.de/Projekte/Anlagenbau.htm))
- [19] T. Wember: Skript Technische Statistik und statistische Versuchsplanung, Eigenverlag
- [20] OptiSlang Version 3.0, Theory Manual und User's Guide, dynardo Dynamic Software and Engineering GmbH, 2009
- [21] J. Gamweger, O.Jöbstl, M. Strohmam, W. Suchowerskyj: Design for Six Sigma, Hanser, München, 2009
- [22] A. Haldar, S. Mahadevan: Probability, Reliability and Statistical Methods in Engineering Design, J. Wiley, 2000
- [23] G. Taguchi: Introduction to Quality Engineering; Designing Quality into Products and Processes, Unipub Kraus International Publications and American Supplier Institute, 1986
- [24] G. Taguchi: System of experimental design, Vol. 1 and 2, Unipub Kraus International Publications and American Supplier Institute, 1987
- [25] T. Bendell (Editor): Taguchi Methods; Proceedings of the 1988 European Conference, Elsevier, 1989
- [26] L. Fahrmeir, T. Kneib, S. Lang: Regression. Modelle, Methoden und Anwendungen. 2. Auflage, Springer, 2009
- [27] H. Pruscha: Statistisches Methodenbuch, Springer, 2006



**11. Index**

Adjusted Response .....	51	-verfahren .....	24
Alternativhypothese .....	83, 86	Methode der kleinsten Quadrate .....	41
Anforderungen .....	8	Mindeststichprobenumfang .....	83
ANOVA-Tabelle		Modell .....	4, 62
des Regressionsmodells .....	91	-ansatz .....	25
Aussiebeplan .....	61	-kalibrierung .....	7
Auswertematrix .....	40	Meta- .....	7
Basisfunktion		nicht parametrisch .....	68
radiale .....	66	strukturtreues .....	4
sigmoidal .....	67	verhaltensgültiges .....	4
trigonometrisch .....	66	Modellbildung	
Bestimmtheitsmaß .....	44, 89	empirische .....	6
Black-Box-Modell s. Modellbildung, empirische		grey-box .....	7, 29
Blockbildung .....	33	Gründe .....	4
C&E-Matrix .....	21, 23, 24	physikalische .....	6, 21, 25
Computerexperiment .....	7, 35	Vorgehen .....	4
Concentration Chart .....	73	Modellschwäche .....	89
DFSS .....	9, 70	Modellschwäche-Test .....	90
Dokumentation .....	33	Multilinearform .....	40
Effekt .....	40	Multi-Vari-Chart .....	72
Einflussgröße .....	3, 21	Nullhypothese .....	83, 85, 86
qualitative .....	24	Optimierung .....	20, 23
End-of-Life-Test .....	9	Orthogonalität .....	64
Erkenntnisgewinn .....	19	overfitting .....	66
Erprobung .....	8	Paarweise Vergleich .....	21, 74
Experiment .....	3	Parameteridentifikation .....	7
Faktor		Pareto	
-stufen .....	27, 28	Optimierung .....	23
Fehler		-Prinzip .....	72
1. Art .....	32, 50, 84, 90	P-Diagramm .....	22
2. Art .....	84	Plackett-Burman-Plan .....	61
modellbedingter .....	25	Polynomansatz .....	26
systematischer .....	11	Predicted Response .....	52
zufälliger .....	11	Produkt-/Prozess-Suche .....	76
Feldversuch .....	3	Prognose .....	20
fit-for-standard .....	9	Prüfling .....	8, 13
fit-for-use .....	9	Prüflings	
Freiheitsgrad .....	85, 87	-archivierung .....	33
F-Test .....	85	-bezeichnung .....	33
Glass-Box-Modell s. Modellbildung,		-entwurf .....	34
physikalische		-herstellung .....	35
Histogramm .....	85	-streuung .....	10, 11, 12, 23, 33
historische Daten .....	23	Pure Error .....	42, 90
Interaction Graph .....	52	p-Wert .....	50, 90
Irrtumswahrscheinlichkeit .....	84, 86	Randomisieren .....	33
Ishikawa-Diagramm .....	21	Regression .....	41
Kerne .....	67	Standardabweichung .....	44
kompakter Träger .....	66	Standardfehler .....	44
Komponententausch .....	73	reiner Versuchsfehler .....	s. Pure Error
Koordinator .....	19	Residuum .....	44
Korrelationskoeffizient .....	44, 89	Ressourcen .....	19, 34
Künstliche Neuronale Netze (KNN) .....	67	RMS-Error .....	44, 89
Laborversuch .....	3	Robustheits	
Lack of Fit .....	89	-analyse .....	20
Lack-of-Fit-Test .....	90	-optimierung .....	20
Mess		S.d.q.A. .....	89
-auflösung .....	24	Samplingmethoden	
-fehler .....	22	deterministisch .....	68
-mittel .....	24	stochastisch .....	68



Scheineffekt .....	61	Varianz .....	85
Schwellenwert .....	83, 85	Varianzanalyse .....	86
Screening.....	20	einfache .....	49
Shainin-Methode .....	72	faktorielle .....	50
Signifikanz .....	10, 11, 12, 33, 34	multiple .....	50
Signifikanzbewertung		Tabelle.....	50
bei Regression .....	89	Verifizierung .....	8
einfache.....	42	Vermengung .....	61
Six-Sigma .....	9	Versuch .....	3, 35
Spline.....	66	Versuchs	
Standardfehler		-durchlauf .....	33
des Regressionskoeffizienten.....	46	-einrichtung.....	34, 35
Sternpunkt .....	64	-fähigkeit.....	34
Störfallplan.....	33	-methodik.....	3
Störgröße.....	22, 33	-plan, D-optimaler.....	25
Summe der quadratischen Abweichungen		-plan, einfaktorierter.....	28
(S.d.q.A.).....	87	-plan, stochastischer .....	68
System.....	4	-plan, vollfaktorierter .....	31, 40
-analyse.....	4, 21	-plan, zentral-zusammengesetzter.....	64
-ausgang .....	5	-planung .....	11
-eingang .....	5	-planung, statistische .....	3, 11
-elemente .....	4	-raum .....	25
-grenze .....	4	-rauschen .....	87
-wechselwirkungen .....	5	-realisierung.....	3
Systemidentifikation s. Modellbildung,		-realisierung, einfache.....	32
empirische		-strategie .....	19
Systemtheorie.....	4	-streuung .....	42
Taguchi-Methode.....	70	Vertrauens	
Taylor-Entwicklung .....	26	-intervall.....	82
test to failure .....	9	-niveau.....	82, 84
test to pass .....	9	Wechselwirkung.....	30, 53
Transferfunktion.....	5, 6	White-Box-Modell s. Modellbildung,	
Trial and Error.....	28	physikalische	
Validierung.....	8	Zentralpunkt .....	64
Variablensuche.....	75	Zielgröße .....	3, 21
Variablentransformation .....	27, 41	qualitative .....	23



**Robert Bosch GmbH**

C/QMM

Postfach 30 02 20

D-70442 Stuttgart

Germany

Phone +49 711 811-0

**[www.bosch.com](http://www.bosch.com)**

