

M.Eng. Daniel Kanzler

**Zuverlässigkeitsanalyse von
digitalen Radiographie-Systemen bei
der Prüfung von realen Materialdefekten**

Die vorliegende Arbeit entstand an der Bundesanstalt für Materialforschung und -prüfung (BAM).

Impressum

**Zuverlässigkeitsanalyse von
digitalen Radiographie-Systemen bei
der Prüfung von realen Materialdefekten**

2016

Herausgeber:

Bundesanstalt für Materialforschung und -prüfung (BAM)

Unter den Eichen 87

12205 Berlin

Telefon: +49 30 8104-0

Telefax: +49 30 8104-72222

E-Mail: info@bam.de

Internet: www.bam.de

Copyright© 2016 by

Bundesanstalt für Materialforschung und -prüfung (BAM)

Layout: BAM-Referat Z.8

ISSN 1613-4249

ISBN 978-3-9818270-1-9

Zuverlässigkeitsanalyse von digitalen Radiographie-Systemen bei der Prüfung von realen Materialdefekten

Dissertation
zur
Erlangung des akademischen Grades
Doktor-Ingenieur (Dr.-Ing.)
der Fakultät für Informatik und Elektrotechnik
der Universität Rostock

vorgelegt von

Daniel Kanzler Diplom-Wirtschaftsing. (FH) / M. Eng.
geb. am 26. Januar 1983 in Alzenau / Wasserlos

Gutachter:

Herr Prof. Dr. rer. nat. habil. Ewald - Universität Rostock
Herr Prof. Dr.-Ing. Gehlen - Technische Universität München
Frau Dr. rer. nat. Müller - Bundesanstalt für Materialforschung und -prüfung (BAM)

Abgabe der Dissertationsschrift: 2. Dezember 2015

Tag der wissenschaftlichen Aussprache: 2. Juni 2016

Kurzfassung

Die zerstörungsfreie Prüfung (ZfP) ist aus solchen Bereichen unseres Lebens nicht mehr wegzudenken, in denen Schäden mit hohen Folgekosten oder Gefährdungen von Menschenleben entstehen können (Beispiele: Transportwesen, Energieerzeugung, Chemieindustrie). In der Praxis kann ein Prüfsystem an seine Grenzen geraten, z.B. bei kleinen Defekten. Defekte mit kritischer Größe werden möglicherweise nicht detektiert. Daher müssen probabilistische Bewertungsverfahren das Prüfsystem beschreiben.

Es wird eine objektive Qualitätskennzahl gesucht, auf deren Basis die Anwendbarkeit der Prüfmethode definiert werden soll. Die Auffindwahrscheinlichkeit (engl. probability of detection - POD) erfüllt diese Anforderung. Die POD zeigt auf Basis des Zusammenhangs und der Streuung der Daten, ob das Verfahren für die Prüftätigkeit akzeptiert werden kann oder noch verbessert werden muss.

Das ursprüngliche POD-Verfahren wurde für quasi eindimensionale Defekte in dünnen Luftfahrtbauteilen entwickelt. In der industriellen Realität ist diese Bewertung ein Balanceakt zwischen Statistik und Durchführbarkeit: Die Prüfung soll mit realen Defektdaten für die spätere Produktion des Bauteils (bzw. wiederkehrende Wartungsprüfung) bewertet werden. Doch die notwendige Gegenüberstellung zwischen Schliifdaten, für die Erfassung der wahren Defektgröße von räumlich ausgeprägten Defekten und dem Signal eines ZfP-Systems stellt sich als herausfordernde und kostenintensive Aufgabe heraus. Sowohl die Aufstellung eines gemeinsamen Koordinatensystems als auch die Beschreibung und Angleichung der Daten stellen eine notwendige Vorarbeit dar. In dieser Arbeit wird ein mögliches Vorgehen entwickelt, dass im Weiteren eingesetzt werden kann. Während in der Literatur zum Thema POD häufig die Begrenzung des Einsatzes einer eindimensionalen POD (POD mit einem Defektparameter) für reale Defekte bereits erkannt wurde, soll außerdem in dieser Arbeit das Verfahren auf der Signalseite umfassender erweitert werden, um die Einbeziehung realer Defekte in die POD-Bewertung zu ermöglichen. Hierfür werden mit Hilfe dieser Arbeit zwei wesentliche Neuerungen in der POD-Bewertung eingeführt:

1. Die Anzeigenfläche wird als wichtiges Indiz zur Detektion in die Bewertung eingeführt. Dabei zeigt der Ansatz einer Observer-POD, bei dem der Detektierbarkeit eines Defekts beschrieben wird, eine Möglichkeit in die Bewertung zu erweitern. Jedoch wird die notwendige Datenanzahl die für eine Observer-POD selten mit Experimenten erreicht. Daher schlagen wir die Einführung eines Glättungsalgorithmus vor, um auch auf der Basis von wenigen Daten die Flächenabhängigkeit zu erfassen. Der Algorithmus wird hierbei durch simulierte

Daten auf seine Funktionsfähigkeit überprüft, bevor er auf reale Defekte angewendet wird. Gleichzeitig helfen die simulierten Daten einen Vergleich zu den vorhergegangenen Ansätzen zu ermöglichen.

2. Darüber hinaus reichen die Daten der realen Defekte häufig nicht aus, um die statistische Forderung zu gewährleisten, so dass es notwendig wird, künstliche Defekte mit einzubeziehen. Deshalb sollen die vorhandenen künstlichen Defekte in Form von Referenzdefekten mit einbezogen werden, um die statistische Grundlage zu erhöhen. Für die Prüfung von Referenzdefekten sind jedoch wichtige Einflussgrößen (z.B. Oberflächenrauigkeit) nicht vorhanden. Wegen der unterschiedlichen Aussagekraft der Daten und zur Vermeidung einer zu optimistischen Abschätzung, ist eine einfache Mischung der Daten ausgeschlossen. Um realen Defekten eine Möglichkeit dafür zu schaffen, dass die Eigenschaften der realen Defekte angemessen auf das Ergebnis der Bewertung des Verfahrens Einfluss nehmen können, wird eine gewichtete Kombination der Defektdaten für die Bewertung vorgestellt.

Das Vorgehen wird am Beispiel der radiographischen Prüfung einer elektronenstrahlgeschweißten Naht durchgeführt. Die Schweißnaht verbindet den Deckel zur Außenwand eines Kupferbehälters, der für die spätere Endlagerung von verbrauchten Brennstäben aus Kernkraftwerken entwickelt wurde. Die Messergebnisse stammen aus von der Firma Posiva Oy, dem zuständigen Unternehmen für die Endlagerung von verbrauchten Brennstäben aus Kernkraftwerken in Finnland. Hierbei stellt die POD-Bewertung ein wichtiges Element in der Gesamtrisikobewertung für das Endlagersystem dar.

Abstract

Nondestructive testing (ndt) systems are essential for areas in our lives, in which there is a high risk for failures that would induce high costs or even damage to people and the environment (i.e. transportation, energy production, chemical industry). It is necessary to find and to characterise every defect in the material which might jeopardise the functionality of the tested part. But in the praxis the testing system will be used at their limits, i.e. for detecting small defects. Thus, there is a probability that critical defects might be overseen, which must be quantified.

The evaluation is especially important for safety-relevant areas. The probability of detection (POD) characteristic is an objective number, which is widely used in these cases. It is used to provide a statement about the tested ndt system.

The POD can provide the statement whether the system is working well enough to be accepted to find the defects.

The original POD method was developed for one-dimensional defects in thin parts used in the aircraft industry. In reality, the evaluation is a compromise between statistics and costs. On the one hand, the real testing situation should be evaluated for the later use. On the other hand, the evaluation of real defects including the metallography and the comparison with the signals is a complex and expensive task. To find a coordinate system to compare the data is, therefore, an important prerequisite, before starting to evaluate. Therefore, this thesis will present a practical approach.

The research community, as well, sees the POD of the real defects as a challenge. It is necessary to extend the one-parametric POD approach by evaluating the whole NDT indication. The area of the NDT indication is one important fact which should be included.

The thesis will introduce two new aspects to the calculation of the POD:

1. The area of the indication will be introduced by using a smoothing algorithm, which is based on the known Observer-POD. The Observer-POD cannot be used due to the small amount of available data. After the smoothing, the data will be used for an advanced POD method. The method will be verified by simulated data, before applying it on real data.

2. Even after introducing a new method - to be able to include real defects - the amount of real defect data may not be sufficient for a statistical evaluation. Therefore, it is necessary to use also artificial defects. However, the artificial defects may lead to an overestimation of the ndt system. To combine the different types of defects a weighted combination will be

introduced. This approach helps to acknowledge the importance as well as the costs and the work done for the real defects, and leads to a useful estimation of the POD for the system.

The investigation was carried out on an example of radiographic testing system for an electron-beam weld. The weld is a critical part of the copper canister for the deposit of spent nuclear fuel. Data were measured and evaluated at Posiva Oy, the company responsible for the spent nuclear fuel disposal in Finland. The POD analysis is an important element of the risk assessment for the final deposit system.

Inhaltsverzeichnis

| | | |
|----------|--|----------|
| 1 | Einleitung | 1 |
| 2 | Zuverlässigkeitsbewertung eines zerstörungsfreien Prüfsystems am Beispiel der Radiographie | 5 |
| 2.1 | Zuverlässigkeit von zerstörungsfreien Prüfsystemen | 5 |
| 2.1.1 | Definitionen der Zuverlässigkeit | 5 |
| 2.1.2 | Beschreibung der Zuverlässigkeit von Prüfsystemen | 9 |
| 2.1.3 | Kennzahlen zur Zuverlässigkeit | 11 |
| 2.2 | Auffindwahrscheinlichkeit eines zerstörungsfreien Prüfsystems | 13 |
| 2.2.1 | Einführung des kritischen Defektparameters für die strukturelle Integrität | 14 |
| 2.2.2 | Bewertung der Auffindwahrscheinlichkeit mit Verwendung der quantitativen Signalhöhe | 16 |
| 2.2.3 | Likelihoodstatistik | 20 |
| 2.2.4 | Berechnung der Auffindwahrscheinlichkeit | 22 |
| 2.2.5 | Vertrauensbereich | 24 |
| 2.3 | Physikalische Grundlagen für die Zuverlässigkeitsbewertung von Radiographiesystemen | 30 |
| 2.3.1 | Radiographie als zerstörungsfreie Materialprüfung | 30 |
| 2.3.2 | Radiographie als Prozess mit zufälligen Einflüssen | 31 |
| 2.3.3 | Physikalische Grundlagen der digitalen radiographischen Prüfung mit Hochenergie-Röntgenstrahlung | 32 |
| 2.3.4 | Physikalischer Zusammenhang zwischen Signal und Defektgröße | 32 |
| 2.3.5 | Zusammenhangs zwischen durchstrahlte Länge und Kontrast für ein Strahlenspektrum | 34 |
| 2.3.6 | Beschreibung der variablen Einflüsse auf das Radiographiesystem | 35 |
| 2.3.7 | Bewertung der Auffindwahrscheinlichkeit in der Radiographie | 37 |
| 2.4 | Herausforderung bei dem Einsatz der POD in der Radiographie mit realen Defekten | 39 |

| | | |
|----------|---|------------|
| 2.4.1 | Arten zur Verfügung stehender Materialdefekte | 39 |
| 2.4.2 | Umgang mit realen Daten bei der Zuverlässigkeitsbewertung von ZfP-Systemen | 41 |
| 3 | Neue Methodik zur Bewertung von Radiographie-Systemen mit Hilfe von realen Defekten | 45 |
| 3.1 | Umgang mit Schlifflen realer Defekte bei der Bewertung durch die POD | 45 |
| 3.1.1 | Metallographische Untersuchung | 45 |
| 3.1.2 | Konventionelle POD mit realen Defekten unter Verwendung einzelner Schlifffbilder | 47 |
| 3.1.3 | POD-Bewertung anhand von mehrerer metallographischen Schlifflen | 49 |
| 3.1.4 | Unterschiede zwischen der radiographischen Anzeige und den Informationen aus den Schlifflen | 51 |
| 3.2 | Einführung der Anzeigenbreite in die Zuverlässigkeitsbewertung | 53 |
| 3.2.1 | Einsatz einer Glättung | 53 |
| 3.2.2 | Anzeigenflächen-bedingte Probability of Detection | 56 |
| 3.3 | Kombination von realen und künstlichen Defektdaten für die Zuverlässigkeitsbewertung | 64 |
| 3.3.1 | Bewertung der künstlichen und realen Defekte | 64 |
| 3.3.2 | Einfluss der Glättung auf die Ergebnisse | 65 |
| 3.3.3 | Beeinflussung der Gewichtung | 66 |
| 4 | Experimentelle POD-Bewertung des Radiographiesystems bei der Prüfung der Elektronenstrahlschweißnaht bei Endlagerkanistern | 69 |
| 4.1 | Beschreibung des zu bewertenden ZfP-Systems | 69 |
| 4.1.1 | Der Kupferkanister als Korrosionsbarriere | 69 |
| 4.1.2 | Zerstörungsfreie Prüfung des Kupferkanisters | 71 |
| 4.1.3 | Bewertungsgrundlage des Radiographie-Systems | 76 |
| 4.2 | Ergebnisse zur Bestimmung der POD gemäß der entwickelten Methodik | 80 |
| 4.2.1 | Modellierte Daten zur Überprüfung der Methodik | 80 |
| 4.2.2 | POD der Prüfung einer EB-Kupferschweißnaht | 88 |
| 4.2.3 | Kombination von realen und künstlichen Defekten | 99 |
| 5 | Zusammenfassung und Schlussfolgerung | 101 |

Abkürzungen

| | |
|--------------|---|
| CNR | Kontrast-Rausch-Verhältnis - engl. contrast noise ratio |
| DDA | digitaler Matrixdetektor - engl. digital detector array |
| DeP | Datenwert erzeugende Prozess |
| EBW | Elektronenstrahlschweißung - engl. electron beam welding |
| EDM | elektroerosivem Abtragungsverfahren - engl. electrical discharge machining |
| FBH | Flachbodensacklochbohrung - engl. flat-bottom-hole |
| FBSQ | Quadratische Flachbodenlöcher- engl. flat-bottom-squares |
| GW | Grauwerten |
| Linac | Linearbeschleuniger - engl. LINear ACcelerator |
| MaPOD | POD unterstützt durch Modellierungssoftware - engl. model-assisted POD |
| MLE | Maximum Likelihood Methode - engl. maximum likelihood estimation |
| POD | Auffindwahrscheinlichkeit - engl. probabability of detection |
| Pixel | Bildelemente in einer digitalen Anzeige - engl. PICTure ELEMent |
| ROC | Operationscharakteristik eines Empfängers - engl. receiver operating characteristic |
| RT | radiographischen Prüfung - engl. radiographic testing |
| SKB | Svensk Kärnbränslehantering AB - engl. Swedish Nuclear Fuel and Waste Management Co |
| SNR | Signalrauschverhältnis - engl. signal to noise ratio |
| UT | Ultraschallprüfung - engl. ultrasonic testing |
| WLSE | Schätzung mit gewichteten kleinsten Fehlerquadraten - engl. weighted least squares estimation |
| ZfP | zerstörungsfreie Prüfung - engl. nondestructive testing |

1 Einleitung

Tagtäglich vertrauen weltweit viele Menschen ihr Leben technischen Systemen an, zum Beispiel in Hochgeschwindigkeitszügen, in Flugzeugen, in Schiffen oder allgemein im Straßenverkehr. Ein Ausfall eines solchen Systems kann in vielen Fällen schwerwiegende Folgen haben. Aus diesem Grund werden für die Bereiche, bei denen durch einen Ausfall Menschen in Gefahr sind oder hohe materielle Schäden entstehen können, besonders sichere und zuverlässige Systeme gefordert. Ein wichtiges Mittel, ein System und somit seine Sicherheit zu überprüfen, stellt die zerstörungsfreie Prüfung - engl. nondestructive testing (ZfP) dar. Doch auch diese muss in ihrer Fähigkeit und Zuverlässigkeit abgeschätzt werden. Normen und Richtlinien gewährleisten eine Beschreibung der minimalen Fähigkeit des Prüfsystems. Auf der Suche nach dem tatsächlichen Können müssen jedoch probabilistische Beurteilungen unter möglichst realistischen Bedingungen durchgeführt werden. Die am weitesten verbreitete Beurteilung ist die Bestimmung der Auffindwahrscheinlichkeit - engl. probability of detection (POD). Eine POD ermöglicht beispielsweise mit Hilfe von Erfahrungswerten einen Vergleich mit alternativen Verfahren.

Während in vielen technischen Anwendungen teilweise über Jahrhunderte Erfahrung gesammelt wurden, ist die Gestaltung einer sicheren Endlagerung von verbrauchten Brennelementen aus Kernkraftwerken eine neue und komplexe Aufgabe. Es muss über viele Jahrtausende ein Schutz vor ionisierender Strahlung, chemischer Vergiftung und der Schutz des Grundwassers als Trinkwasserressource gewährleistet werden [1]. Dabei wurde beispielsweise an die Einlagerung in einem Bergwerk in tiefen geologischen Formationen gedacht [2]. Während eine schnelle Lösung für das bisher noch ungelöste Problem verlangt wird [3, 4], stellt das Thema Endlager eines der meist diskutierten Themen unserer Zeit dar [3], [5–8]. Die finnischen und schwedischen Ansätze, die verbrauchten Brennstäbe in ein Granitgestein einzulagern [9], besitzen dabei eine weltweite Vorreiterrolle, auf das besonderes Augenmerk gelegt wird [10], [7], [11]. Das Konzept zur Endlagerung enthält ein Sicherheitssystem mit mehreren Barrieren (KBS-3 - KBS: Kärn Bränsle Säkerhet [12]), in der ein Kupferbehälter mit einer 5,00 cm dicken Wand die Korrosionsbarriere darstellt (siehe Abbildung 1.1) [10], [13].

Die Anforderungen für die Kanisterkomponenten sind so hoch, dass der Kanister vollständig und mit verschiedenen Verfahren geprüft werden muss. Die einzelnen Kanisterbestandteile (siehe Abbildung 1.2) werden mit Schweißtechniken verbunden. Die Schweißnaht stellt eine Schwachstelle des Behälters dar.

1. Einleitung

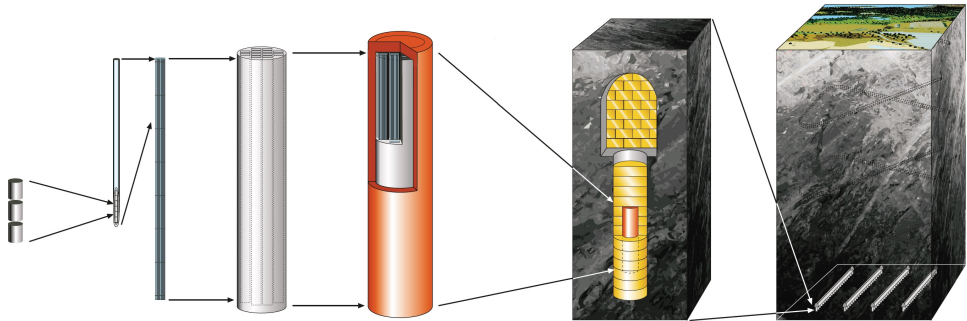


Abbildung 1.1: Mehrbarrierensystem für die Endlagerung von verbrauchten Brennstäben: Die schlecht wasserlöslichen Uranoxidpellets sind in Brennstäben aufgereiht und in Brennelementen angeordnet. Um den Scher- und Druckkräften standzuhalten, wird ein gusseiserner Stahlkocher verwendet, der von einem Kopperbehälter umschlossen ist. Der Bentonitlehmpuffer hat die Aufgabe Wärme abzuführen und Wasserkontakt mit dem Kopperbehälter zu vermeiden oder zu reduzieren. Das eigentliche Endlager liegt mehrere hundert Meter tief in Granitgestein (Quelle: [13])



Abbildung 1.2: Kanisterbauteile bestehend aus Kopperbehälter (Mantel mit Deckel), dem Stahlkocher und den Brennstäbe (Quelle: [13])

Die mit einem elektronenstrahlgeschweißte Naht wird daher mit vier verschiedenen ZfP-Verfahren geprüft: Ultraschallgruppenstrahlprüfung, Radiographie mit Linearbeschleuniger und Digitaldetektor, Wirbelstromprüfung mit einer hohen und einer niedrigeren Frequenz und eine visuelle Prüfung mit einem Kamerasystem [14].

Da es sich um ein Projekt mit Gefahr für Mensch und Umwelt handelt und auch in politischer Hinsicht großer Druck auf den Ausführenden lastet, müssen die einzelnen ZfP-Verfahren ihre Zuverlässigkeit unter Beweis stellen. Für die Bewertung von ZfP-Verfahren hinsichtlich der Defektgröße hat sich die POD-Einschätzung bewährt.

Jedoch treten bei der Bewertung der ZfP-Systeme für die Endlagerung neue Probleme auf: Meist bauen POD-Bewertungen auf einer großen Anzahl von Werten, beispielsweise durch wiederkehrende Prüfungen und somit einer Abschätzung von Defektgrößen auf (siehe Turbinenprüfung in der Luftfahrt [15]) oder es werden in einfachen Fällen vergleichbare Referenzdefekte bewertet, um einen Anhaltspunkt zu erhalten. Die wiederkehrende Prüfung ist jedoch nach Verschluss des Endlagers nicht mehr möglich. Auch ist die Produktionsqualität so hoch, dass kaum Materialdefekte auftreten [16]. Für die POD-Bewertung ist es daher wichtig zwischen künstlichen Defekten, die mit einfacher Geometrie in ein Bauteil eingebracht werden, realistischen Defekten, die mit Absicht hervorgerufen werden, und realen Defekten, die durch den Einsatz oder bei der Herstellung entstehen, zu unterscheiden [17]. Eine Einschätzung ist jedoch basierend auf Referenzdefekten und künstlichen Materialdefekten in einem so wichtigem Fall wie der Endlagerung nicht ausreichend, da sie die erwarteten Defekte teilweise nur unzureichend widerspiegelt.

Es wird daher eine Möglichkeit gesucht, die bewährte Bewertung von ZfP-Verfahren mit POD, mit Hilfe von realistischen Defekten zu ermöglichen. Dabei ist die Erweiterung des bekannten Bewertungsverfahrens hinsichtlich zusätzlicher Defektparameter von Nöten. Dadurch muss nicht nur das maximale Signal eines Defekts ausgewertet werden, sondern vielmehr ein Datenfeld von Kontrastverhältnissen zum Hintergrundmaterial. Ein ähnlicher Ansatz wurde bereits in der Ultraschallprüfung angewandt [18] und wird in dieser Arbeit an die Bewertung einer radiographischen Prüfung - engl. radiographic testing (RT) angepasst. Bei Überführung zur RT-Prüfung treten weitere Herausforderungen auf. Diese werden in dieser Arbeit ausgearbeitet und von uns entwickelten Vorschlägen zur Lösung unterbreitet.

Außerdem soll das Verfahren, das auf statistischen Grundsätzen beruht, trotz der geringen Menge an vorhandenen realen Materialdefekten sinnvoll durchgeführt werden. Die bewährten Methoden sollen daher in die Richtung weiterentwickelt werden, dass sie trotz der geringen Datenmengen ein realistisches Bild zur Bewertung der ZfP-Verfahren darstellen. Auch hierfür wird eine praktische Lösung entwickelt.

Daher wird in dieser Arbeit die Evaluation von ZfP-Systemen auf die genannten Herausforderungen weiterentwickelt: Nach einem Überblick über verschiedene Ansätze zur Bestimmung der Fähigkeitsanalyse wird der bisherige Ansatz zur Bestimmung einer POD für

1. Einleitung

die Radiographie erläutert (Kapitel 2). Hierbei werden die Anforderungen und die Einflüsse hinsichtlich einer konventionellen, ein-parametrischen Methode aufgelistet. Es wird außerdem diskutiert, weshalb der konventionelle Ansatz zur POD-Bewertung nur begrenzt für reale komplexe Defekte eingesetzt werden kann.

In Kapitel 3 wird die Gegenüberstellung der realen Defekte mit den zugehörigen Schlifffdaten aus Sicht der POD eingeführt. Basierend auf den Erkenntnissen hinsichtlich der Wahrnehmung zum Finden von Defekten, wird ein neuer erweiterter Ansatz für die POD mit multi-skalen Glättungen eingeführt. Die Glättung bezeichnet hierbei einen Kern mit dem die Daten im Ortsraum gefaltet werden, um in diesem Fall bestimmte physikalische Eigenschaften hervorzuheben oder zu unterdrücken. Die Einführung von verschiedenen Glättungsbandbreiten (multi-skalen) reflektiert die Größe der Anzeige und weitere Einflüsse, die bei der Gegenüberstellung der Anzeigen mit den Schlifffern ersichtlich werden. Die Einbeziehung weiterer Defektparameter ermöglicht nun auch die Einbeziehung von künstlichen Defekten in die Bewertung, um das Vertrauen in die Bewertung zu steigern. Die künstlichen Defekten sollen gewichtet und mit realen Defekten kombiniert werden, um auch die unterschiedliche Aussagekraft der Daten zu erfassen.

In Kapitel 4 wird das neue vorgestellte Vorgehen auf ihre Fähigkeit und ihre Methodik untersucht. Im ersten Schritt werden Ergebnissen aus einer Modellierungssoftware genutzt, bevor der Ansatz auf das zu bewertende RT-System der Firmen Svensk Kärnbränslehantering AB - engl. Swedish Nuclear Fuel and Waste Management Co (SKB) [9] und Posiva Oy [13] im Endlagerfall übertragen wird.

Kapitel 5 liefert eine Zusammenfassung der Arbeit und einen Ausblick für die weitere Entwicklung hinsichtlich der POD für die Bewertung von RT-Systemen.

Diese Arbeit bildet die Grundlage dafür, dass reale Defektdaten korrekt in die Bewertung von ZfP-Systemen einfließen und zeigt, welcher Aufwand für eine richtige Bewertung notwendig ist. Außerdem liefert die Arbeit einen Vorschlag, reale Daten in die POD einfließen zu lassen, selbst wenn die Anzahl nur gering ist.

2 Zuverlässigkeitsbewertung eines zerstörungsfreien Prüfsystems am Beispiel der Radiographie

2.1 Zuverlässigkeit von zerstörungsfreien Prüfsystemen

Ein System besteht im technischen Sinn aus einzelnen Einheiten [19]. Fällt eine Einheit aus, kann dies schwere Folgen nach sich ziehen [20]: So kann beispielsweise der Ausfall einer Welle dazu führen, dass ein Zug entgleist. Dabei entstehen beträchtliche Geldschäden und möglicherweise kommen Menschen zu Schaden (z.B.: das Zugunglück in Manila im Jahr 2014 [21] oder der Vorfall nahe Köln im Jahr 2008 [22]). Um solche Folgen zu vermeiden, soll gegen Ausfälle und Versagen von technischen Einheiten vorgegangen werden [20], [23]. Die technische Diagnostik hilft dabei, das System anhand verschiedener Konzepte, Methoden und Techniken zu überprüfen und zu beurteilen [24].

Ein wichtiges Werkzeug der technischen Diagnostik stellt die ZfP dar. Als ZfP werden nicht-invasive Methoden bezeichnet, die zum Prüfen, Evaluieren und Charakterisieren von Bauteilen eingesetzt werden. Dabei basieren die Methoden auf den physikalischen Prinzipien des Wahrnehmens und Beurteilens [24]. Es ist bei Anwendungen mit möglicherweise schwerwiegenden Folgen notwendig, das ZfP-System als ein technisches System anzusehen, das ebenfalls versagen kann. Damit wird verhindert, dass ein übersteigertes Vertrauen in das geprüfte Bauteil gesetzt wird. Im allgemeinen Sprachgebrauch wird die Zuverlässigkeit des Prüfsystems gefordert.

2.1.1 Definitionen der Zuverlässigkeit

Für den Begriff „Zuverlässigkeit“ existiert keine eindeutige Definition.¹ Aus diesem Grund sollen zunächst die Anforderungen an ein ZfP-System sowie die Unterscheidung zwischen einem zuverlässigen und einem unzuverlässigen System geklärt werden.

Bei dem Versuch die Zuverlässigkeit eines ZfP-Systems zu definieren, wird teilweise auf die allgemeine Bewertung eines Messsystems zurückgegriffen [26]. Hierbei steht die Genauigkeit eines Messsystems im Vordergrund. Die Genauigkeit teilt sich in Präzision und Richtigkeit auf [27], [28]. Die Richtigkeit beschreibt den systematischen Fehler bei der Messung. Die Präzision wird als die Nähe der Messpunkte zueinander bezeichnet. Die Vorteile bei dem

¹In den Normen [20] und [25] bezeichnet „reliability“ einerseits die Funktionsfähigkeit andererseits die Zuverlässigkeit.

2.1 Zuverlässigkeit von zerstörungsfreien Prüfsystemen

Ansatz zur Genauigkeitsbetrachtung sind die probabilistische Sichtweise und die Allgemeingültigkeit des Vorgehens [29]. Jedoch liefert ein Messsystem ein quantitatives Ergebnis, bei dem ein Messfehler angegeben werden kann [28], [30], [31]. Ein ZfP-System wiederum resultiert in einer qualitativen Aussage (“Anzeige” oder “keine Anzeige”). Die Bestimmung des Messfehlers schlägt daher häufig fehl.² Gleichzeitig hat sich gezeigt, dass die Bestimmung eines Messfehlers oder der Genauigkeit teilweise nicht den Ansprüchen und dem gewünschten Nutzen eines Kunden hinsichtlich der Fähigkeit eines Systems entspricht [32]. Dies gilt besonders für ein ZfP-System, da dessen Versagen nicht zu klaren Folgen führt. So ist beispielsweise ein nicht detektierter Materialfehler nicht gleichzusetzen mit dem Ausfall des getesteten Prüfstücks [33]. Der allgemeine Ansatz hinsichtlich der Genauigkeit wird vorerst als Bewertungsmethode abgelehnt, da die Komplexität des Systems nicht vollständig erfasst wird.

Im Weiteren wird in dieser Arbeit der Begriff Defekt für eine Materialinhomogenität genutzt, die im Prüfprozess nachgewiesen werden sollte.

Im zweiten amerikanisch-europäischen Workshop über ZfP-Zuverlässigkeit wurde Folgendes über die Zuverlässigkeit von ZfP-Systemen definiert [34]: Die Zuverlässigkeit von ZfP-Systemen ist der Fähigkeitsgrad, nach dem das ZfP-System seine Anforderung bezüglich Detektion, Charakterisierung und Fehlalarme erfüllt [17]. Die Anforderungen sind folgendermaßen definiert:

- Detektion beschreibt das Auffinden eines vorhandenen Defekts [35];
- Charakterisierung teilt sich in [35]:
 - Positionsbestimmung, Größenanzeige;
 - Klassifizierung; d.h. die Benennung und Bewertung der Anzeige hinsichtlich Defektart;
- Fehlalarm drückt die Fehlbewertung oder Missinterpretation eines defektfreien oder unkritischen Bauteils aus.

Das Detektieren von Defekten stellt einen essentiellen Schritt eines zuverlässigen ZfP-Systems dar, da sonst keine weitere Charakterisierung möglich ist. Daher liegen bei vielen Bewertungen der ZfP-Zuverlässigkeit die Defektdetektion im Fokus [36].

Ein damit verwandter Definitionsansatz für die Zuverlässigkeit von ZfP-Systemen betrachtet die Teilbereiche, die auf das Ergebnis der Prüfung wirken:

- innewohnende Faktoren (engl. *intrinsic capability*), die die physikalische Eigenschaft des Prüfsystems bezeichnen (Prüfsensoren, Hilfsmittel für die Prüfung, Manipulator, usw.) [37]

²Liefert das ZfP-System eine quantitative Größe (Materialfehlergröße), so ist die Berechnung eines Messfehler sinnvoll.

2.1 Zuverlässigkeit von zerstörungsfreien Prüfsystemen

- Anwendungsfaktoren (engl. application parameter), die auf die Prüfung wirken (Ankopplung, sonstige Rauschquellen, Einflüsse der Umwelt, etc.) [37]
- menschliche Faktoren (engl. human factors), die einen Einfluss auf die Prüfung haben (z.B. Faktoren aus der Umgebung des Menschen, der Organisation, der Arbeit wie auch menschliche oder individuelle Charakteristika, die einen sicherheitsrelevanten Einfluss auf das Arbeitsverhalten haben.) [38]
- inner- und außer-organisatorische Einflussfaktoren (engl. organisational context – organisatorische Prozesse der Prüfplanung und Prüfdurchführung, kulturelle Einflüsse, etc.) [39]

Aus Sicht der Fähigkeit Defekte zu detektieren, erfüllt ein ZfP-System seine Aufgabe zuverlässig, wenn nach der Abfrage, ob ein Defekt im Bauteil vorhanden ist, eine wahrheitsgemäße Antwort zurückgemeldet wird [35].

Für die Bewertung eines ZfP-Systems gibt es zwei unterschiedliche Ansätze. Der auf Normen basierte Ansatz, der überwiegend in Europa zum Einsatz kommt, und ein Leistungsnachweis (engl. performance demonstration), der im US-amerikanischen Gebiet überwiegt [40].

Im normbasierten Ansatz wird die Qualität einzelner Teilschritte im ZfP-Prozess sichergestellt. So regeln die Normen die theoretische Fähigkeit des ZfP-Equipments, beispielsweise die minimale Auflösbarkeit, das minimale Defektsignal für eine festgelegte Anzeige (Kontrast-Rausch-Verhältnis) und die notwendige Ausbildung des Prüfers [41]. Dabei entsteht eine konservative Bewertung des ZfP-Systems unter Idealbedingungen mit Hilfe eines Mindest-Qualitätsniveaus. Jedoch gibt die allgemeine Abschätzung nicht die tatsächliche Fähigkeit des Gesamtprozesses wieder. Dabei werden die probabilistischen Einflussgrößen in den Hintergrund gedrängt.

Der Leistungsnachweis beschreibt das System als Black-Box. Das Gesamtsystem wird nur hinsichtlich ihres Ergebnisses als ausreichend oder nicht ausreichend bewertet. Es wird dabei das ZfP-System, bestehend aus Gerät, Prüfanweisung und Prüfteam, überprüft [42]. Einsatz findet der Leistungsnachweis beispielsweise in der Bewertung von ZfP-Systemen im Kernkraftbereich in der USA.³ Bei diesem Ansatz liegt der Fokus auf den unvorhersehbaren und schwankenden Einflussgrößen. Wird trotz dieser Einflussgrößen ein ausreichendes Ergebnis erzielt, so kann das System im eigentlichen Prüfprozess eingesetzt werden. Mit anderen Worten handelt es sich hier um eine probabilistische Anforderung, die die Einflüsse der Einzelteile außer Acht lässt.

³Es sei jedoch angemerkt, dass in letzter Zeit sich in diesem Bereich der europäische Ansatz und der US Ansatz aufeinander bewegen. So werden rein probabilistische Ansätze mit deterministischen Untersuchungen ersetzt.

2.1 Zuverlässigkeit von zerstörungsfreien Prüfsystemen

Im optimalen ZfP-Managementkonzept müssen sowohl die minimalen Anforderung aus Sicht der Normen als auch die probabilistischen Anforderungen umfassend untersucht werden. Außerdem sollte der Prozess als Gesamtkontext in die reale Prüfumgebung eingebettet werden. Ein praktisches Beispiel für ein solches ganzheitliches Konzept, das alle Sichtweisen versucht zu vereinen, findet sich bei dem Managementprogramm von Posiva (siehe Abb. 2.1 nach [43], [44]). Dieses Konzept geht weit über die vorgestellten Ansätze hinaus.

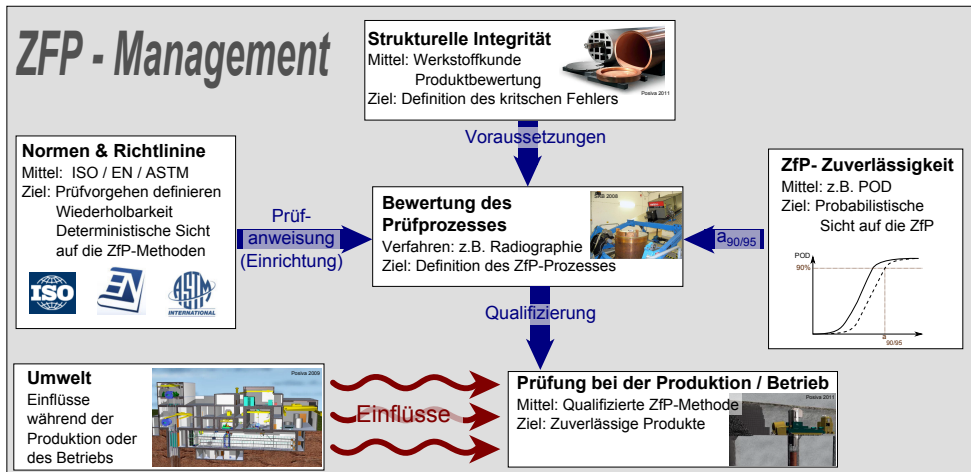


Abbildung 2.1: ZfP-Managementprogramm der Firma Posiva Oy: Ein Ansatz für die Zuverlässigkeit, der sowohl Normen wie auch probabilistische Ansätze in die Bewertung aufnimmt und zusätzlich auch die Gesamtsituation betrachtet.

Das Prüfziel der ZfP ist es, die strukturelle Integrität eines Bauteils oder einer Anlage nachzuweisen, die eine ungestörte Funktionalität gewährleistet. Ist die strukturelle Integrität durch Materialdefekte, die ein Versagen hervorrufen können, gestört, besteht das Ziel diese nachzuweisen und ggf. eine Reparatur oder einen Austausch vorzunehmen. Die Normen und Richtlinien schreiben die Mindestvoraussetzung des Verfahrens im geregelten Bereich vor.⁴ Die probabilistischen Anforderungen schätzen realistisch die tatsächliche Fähigkeit des Systems hinsichtlich möglicher Fehlentscheidungen ab. Hat ein System den Qualifizierungsprozess durchlaufen, müssen, für den tatsächlichen Einsatz, noch weitere Einflüsse betrachtet werden (z.B. in der Produktionsprüfung): Umwelteinflüsse [37], menschliche Faktoren [38] oder organisatorische Einflüsse [39]. In dieser Arbeit soll eine quantitative probabilistische Bewertung des ZfP-Systems im Fokus stehen. Der nächste Abschnitt stellt eine Auswahl von Verfahren zur Auswertung vor.

⁴Im unregulierten Bereich liegen oft Empfehlungen vor, die sich am geregelten Bereich orientieren.

2.1.2 Beschreibung der Zuverlässigkeit von Prüfsystemen

So sind die anfallenden Kosten für die Prüfung, die mögliche Fehlalarmrate oder die Anzahl nicht gefundener Defekte Beispiele für Bedingungen an die Fähigkeit des ZfP-Systems. Der Zusammenhang der einzelnen Bedingungen im System wird bei einer probabilistischen Betrachtung sichtbar, die die traditionellen Beurteilungen („Daumenregel“) abgelöst hat. Diese Gratwanderung führt zu neuen Sichtweisen, wodurch Unsicherheiten eine essentielle Rolle spielen [45]. Die ZfP-Verfahren werden meist an den Grenzen ihrer Detektionsfähigkeit eingesetzt, wodurch schwankende Ergebnisse wegen der Unsicherheiten im System und weiteren Einflüssen auf die Prüfung wichtig sein können [46]. Die Schwankungen müssen daher probabilistisch betrachtet werden.

Je eindeutiger eine richtige Entscheidung einer wahren Situation (Bauteil besitzt einen / keinen Defekt) zugeordnet werden kann, umso besser ist das System hinsichtlich seiner Fähigkeit [47]. Zur Beschreibung eines radiographischen Prüfprozesses [48] hilft das Modell in Abbildung 2.2, das an ein allgemeines Datenauswertungsprozessmodell angelehnt ist [49].

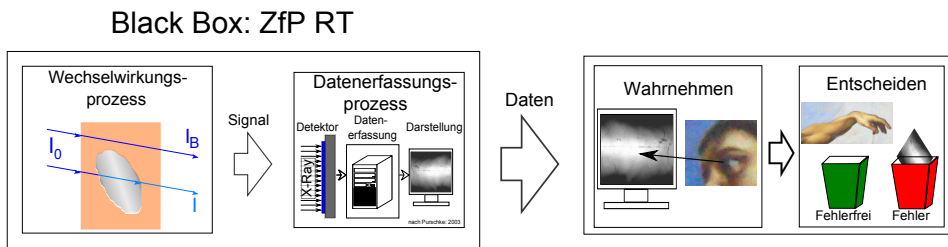


Abbildung 2.2: Die radiographische Defektprüfung aufgespalten in die einzelnen Prozesse von der Datenentstehung bis hin zur Entscheidung

Für die quantitative Bewertung eines komplexen ingenieurtechnischen Systems, wie dem ZfP-System, ist es nicht möglich, die Funktionsfähigkeit direkt zu bestimmen [49]. Der Datenwert erzeugende Prozess (DeP), der bewertet werden soll, kann jedoch als Black-Box dargestellt werden. Während Modellierungsprogramme einen großen Beitrag zum Verständnis des DeP leisten können, sind jedoch oft nicht alle Informationen gegeben. Mit dem System zu kommunizieren ist daher nur mit Hilfe von Stichproben und durch Charakterisierung der Fähigkeit des Systems möglich [49]. Der DeP kann eine beliebig große Zahl von Daten erzeugen. Die Anzahl der untersuchten Daten zur Beschreibung des DeP sind jedoch nur endlich groß. Die Beschreibung von Wahrscheinlichkeiten und Verteilungsfunktionen geben Rückschlüsse von inkompletten Informationen auf das wahrscheinliche Verhalten des Sys-

2.1 Zuverlässigkeit von zerstörungsfreien Prüfsystemen

tems (Inferenz). Die Inferenz wird dann besonders wichtig, wenn jeder Durchlauf des DeP sehr kostenintensiv ist, beispielsweise wegen hohen Prüfkosten oder teuren Prüfstücken.

Eine mögliche probabilistische Sichtweise stellt die ZfP als Detektionsprozess eines Defektsignals im verrauschten Hintergrund dar. Der verrauschte Hintergrund ist eine Verteilung von Signalen, die ein defektfreies Testobjekt liefert: die so genannte Rauschverteilung. In der vorliegenden Aufgabenstellung sind innerhalb dieser Verteilung keine Informationen hinsichtlich Materialdefekte vorhanden [50]. Gleichzeitig kann das Signal selbst stochastisches Verhalten aufweisen, wie in Abschnitt 2.3 gezeigt wird.

Wird ein Objekt geprüft, für das nicht bekannt ist, ob ein Materialdefekt vorliegt, wird das gemessene Signal mit der Rauschverteilung verglichen. In der allgemeine Signaldetektion kann diese Situation mit Hilfe eines Hypothesentests betrachtet werden [51]. Die Nullhypothese besagt, dass kein Defekt vorliegt; das Signal des unbekanntes Objekts ist vereinbar mit der Rauschverteilung [52]. Wird das gemessene Signal der Stichprobe so groß, dass die Nullhypothese abgelehnt werden muss, muss die Alternativhypothese angenommen werden. Die Alternativhypothese wiederum besagt, dass ein Defekt vorhanden ist. Die Entscheidungsschwelle legt fest, welches Signal als vereinbar mit der Nullhypothese oder der Alternativhypothese gilt. Aus ZfP-Sicht gibt diese das minimal geforderte Signalrauschverhältnis - engl. signal to noise ratio (SNR) an.

Sollen die möglichen Ergebnisse einer Prüfung (richtige Entscheidungen und Fehlentscheidungen) bewertet werden, werden diese häufig in einer Vier-Felder Tabelle (Tab. 2.1) dargestellt. Allgemein ergeben sich zwei alternative Situationen mit jeweils einer korrekten und einer fehlerhaften Bewertung [35]. Im Jargon der Signaldetektion werden die korrekten

Tabelle 2.1: Vier-Fehler Tabelle bei der Bewertung von Prüfsystemen: Die möglichen Fälle zwischen Response und Stimuli als bedingte Wahrscheinlichkeiten (P) ausgedrückt

| | | Response | |
|---------|--------------|-----------------|---------------|
| | | Defektanzeige A | Anzeigefrei N |
| Stimuli | defekthaft a | $P(A a)$ | $P(N a)$ |
| | defektfrei n | $P(A n)$ | $P(N n)$ |

Bewertungen $P(A|a)$ als „Treffer“ (engl. hit) und $P(N|n)$ als „korrekte Ablehnung“ (engl. correct rejection) bezeichnet. Die falschen Bewertungen heißen $P(A|n)$ „Fehlalarm“ (engl. false call) und $P(N|a)$ „Verpasser“ (engl. miss) [35].

Die reale Bewertung eines ZfP-Ergebnisses basiert häufig auf den möglichen Fehlentscheidungen. Der Fehler erster Art (Fehlalarm: $P(A|n)$) spielt eine wirtschaftliche Rolle, da funktionsfähige Bauteile verschrottet oder repariert werden müssen. Während ein Fehler zweiter Art (Verpasser: $P(N|a)$) ein Sicherheitsrisiko darstellt, da ein defektbehaftetes Bauteil weiter im Prozess bleibt [46], und es im schlimmsten Fall zu katastrophalen Kon-

sequenzen kommen kann [53]. Nach der ASTM E 1316 ergeben sich für die Bewertung von Anzeigen in der ZfP Falschanzeigen, relevante Anzeigen und nicht relevante Anzeigen [50].

Es ist somit eine wichtige Aufgabe, eine geeignete Entscheidungsschwelle zwischen relevanten Anzeigen und nicht relevanten Anzeigen festzulegen. Dabei sollen die Gesamtkosten sowie die Kosten, die aus dem „Fehler erster Art“ als auch aus dem „Fehler zweiter Art“ entstehen, zumutbar sein [35].

2.1.3 Kennzahlen zur Zuverlässigkeit

Aus Kostensicht zeigt sich, dass es sinnvoll ist, sich für eine nicht-optimale Prüfung statt für keinen ZfP-Einsatz zu entscheiden, da selbst eine nicht-optimale Prüfung die Anzahl der defekthaften Bauteile verringert, wodurch fatale Folgen vermieden werden [54]. Jedoch ist eine reine Entscheidungsfindung anhand von Kosten-Nutzen Analysen komplex, da die quantitative Messung von vielen Größen teils nicht möglich ist [35]. Statistische Kennzahlen bringen hierzu Abhilfe.

Risikofunktion

Die Risikofunktion ist eine Möglichkeit zur Bewertung von ZfP-Systemen, die sowohl Kosten als auch Nutzen bewertet [51]. Bei der Risikofunktion wird jedem Kostenfaktor eine Auftretenswahrscheinlichkeit zugewiesen. Das Gesamtrisiko R , bestehend aus der Summe der Einzelrisiken, soll minimiert werden:

$$R = C_{11}P_aP(A|a) + C_{21}P_aP(N|a) + C_{12}P_nP(A|n) + C_{22}P_nP(N|n), \quad (2.1)$$

wobei P_a die Auftretenswahrscheinlichkeit eines Defektes ist und P_n die Wahrscheinlichkeit, dass kein Defekt auftritt. Es gilt $P_a + P_n = 1$. Die Kosten C_{22} , die auftreten, wenn das defektfreie Bauteil nach der Prüfung im Betrieb eingesetzt wird, sind die Minimalkosten und bezeichnen nur die durch die Prüfung angefallenen Kosten. Die Kosten C_{11} bezeichnen die Kosten einer Verschrottung eines defekthaften Bauteils zuzüglich der Prüfkosten. Bei C_{21} und C_{12} handelt es sich um die Kosten der Fehlentscheidungen. Die Kosten C_{21} bezeichnen die Folgekosten eines in Betrieb genommenen Bauteils beim Ausfall zuzüglich die Prüfkosten. Die Abschätzung von C_{21} ist am schwierigsten und Kosten können die anderen Beiträge bei weitem überschreiten (siehe beispielsweise Folgen nach einem Flugzeugabsturz). C_{12} sind die Kosten für die Verschrottung eines defektfreien Bauteils und die Prüfkosten. Die Anwendungen der Risikofunktion für die Bewertung der ZfP in sicherheitsrelevanten Bereichen (z.B. in Kraftwerken, im Transportwesen, in der Luft- und Raumfahrt, im Militärwesen) ist jedoch meist nicht möglich, da für die Risikofunktion die konkreten Kosteninformationen essentiell sind [55], [56]. Im Fall von Schmerzempfinden, Verringerung der Lebensqualität,

2.1 Zuverlässigkeit von zerstörungsfreien Prüfsystemen

Umweltschäden (z.B. radioaktive Verseuchung von Lebensraum) oder Sterberaten können jedoch keine eindeutigen Kosteninformationen bestimmt werden [35].

Wahrscheinlichkeitskennzahlen

Bleibt die Kosteninformation außen vor, so kann eine Bewertung des Prüfsystems nur anhand folgender Wahrscheinlichkeitsaussagen getroffen werden [46]:

- Auffindwahrscheinlichkeit (POD): Sensitivität der Detektion (Häufigkeitsbeschreibung: Quotient aus gefundenen Defekten und defektbehafteten Bauteilen)⁵
- Erkennungswahrscheinlichkeit (engl.: Probability of Recognition): Spezifität der fehlerfreien Bauteile (Häufigkeitsbeschreibung: Quotient aus korrekt akzeptierten Bauteilen und defektfreien Bauteilen)⁶
- Fehlalarmrate: Quotient aus fälschlicherweise abgelehnten Bauteilen und defektfreien Bauteilen, die getestet wurden
- Genauigkeit des Beobachters: Die Summe aus korrekt akzeptierten und gefundenen Defekten dividiert durch alle getesteten Bauteile

Welche Wahrscheinlichkeitsaussage zur Bewertung herangezogen wird, ist abhängig von der Prüfsituation, dem Materialverhalten, der Geometrie des Prüfstücks, der ZfP-Methode und von den möglichen Folgen eines nicht gefundenen Defekts [57].

Wird der Fehlalarmwahrscheinlichkeit, die hauptsächlich ökonomische Folgen hat, eine deutlich geringere Priorität als der Detektion der Defekte zugeordnet – beispielsweise in sicherheitsrelevanten Anwendungen – so steht die Bewertung anhand der Auffindwahrscheinlichkeit im Fokus [57]. Da jedoch die Fehlalarmwahrscheinlichkeit nicht konkret dargestellt wird, muss diese zusätzlich bewertet werden, da ansonsten die Bewertung irreführend sein kann [35]. Das schlechtere Detektionsvermögen des Prüfsystems unter realen Bedingungen liegt oft in nicht direkt erfassbaren Unsicherheiten. Für die Bewertung des ZfP-Systems sollte diese mit Hilfe von statistischen Methoden quantifiziert und bewertet werden, um ein aussagekräftiges Ergebnis zu erlangen [58]. Im weiteren Fokus dieser Arbeit wird die Auffindwahrscheinlichkeit der Defekte (POD) stehen, die im nächsten Abschnitt vertieft wird.

⁵Die Häufigkeitsbeschreibung ist der beste Schätzwert für die real vorliegende Wahrscheinlichkeit (hier gilt die Annahme, dass pro Bauteil nur ein oder kein Defekt vorhanden ist).

⁶siehe Kommentar Auffindwahrscheinlichkeit

2.2 Auffindwahrscheinlichkeit eines zerstörungsfreien Prüfsystems

Beim Auffinden eines Defekts in einem Bauteil besteht für jedes ZfP-System die Aufgabe, das vom Defekt herrührende Nutzsignal von dem Hintergrund des Rauschens zu trennen. Dabei kann auf Basis der Rauschverteilung und einer Entscheidungsschwelle auf den Fehler erster Art (Fehlalarmrate) und somit auf die Wirtschaftlichkeit eines ZfP-Systems geschlossen werden. Jedoch gibt die Rauschverteilung keine Information über die Auffindwahrscheinlichkeit an. Die Auffindwahrscheinlichkeit bezeichnet, die Tatsache ob ein bestimmter Defekt gefunden wird bzw. es Sinn ergibt das Verfahren überhaupt einzusetzen (Fehler zweiter Art - siehe Abbildung 2.3). Doch gerade diese Wahrscheinlichkeit ist der wichtigste Punkt im Entwicklungsprozess von ZfP-Verfahren, besonders für kleine Defekte oder sonstige schwierige Detektionsverhältnisse.

Für die Bestimmung der Auffindwahrscheinlichkeit muss die Verteilung der Defektsignale herangezogen werden, da die Auffindwahrscheinlichkeit den Bereich der Signalverteilung oberhalb der Entscheidungsschwelle beinhaltet (siehe hierzu Abbildung 2.3). Hinter der Defektsignalverteilung verbergen sich sämtliche mögliche Arten, Geometrien und Positionen der Defekte und deren Auftretswahrscheinlichkeiten. Aus diesem Grund muss zur Bestimmung der Signalverteilung eine repräsentative und ausreichend große Stichprobe aus der Grundgesamtheit der Defekte herangezogen werden [52]. Nicht vollständig bekannte oder

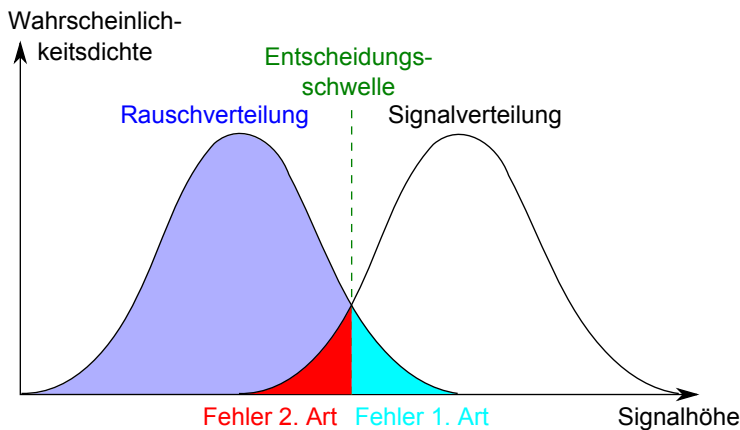


Abbildung 2.3: Rausch- und Signalverteilung mit Entscheidungsschwelle und zugehörigen Fehlentscheidungen

unerwartete Defektentstehungsprozesse und unbekannte Auftretswahrscheinlichkeiten machen es dabei schwierig, die Grundgesamtheit zu beschreiben [45]. Daher werden Modelle der

2.2 Auffindwahrscheinlichkeit eines zerstörungsfreien Prüfsystems

strukturellen Integrität genutzt, um, trotz der unbekanntenen Grundgesamtheit der Defekte, eine Aussage über die Lebensdauer des geprüften Objekts machen zu können.

2.2.1 Einführung des kritischen Defektparameters für die strukturelle Integrität

Der aus dem englischen Sprachraum stammende Begriff „strukturelle Integrität“ beschreibt zunächst die Bruchfestigkeit eines Bauteils unter den vorgesehenen Belastungsbedingungen. Der kritische Defekt wird so definiert, dass für ein Bauteil mit diesem Defekt die strukturelle Integrität nicht mehr gewährleistet werden kann [59]. Außerdem interessiert hierbei, welche Folgen durch diesen Defekt entstehen können. Dabei wird der Ausfallprozess genau unter die Lupe genommen und mögliche Defekte sowie Gründe für den Ausfall des Bauteils aufgeführt [60]. Bei dem vorzeitigen Ausfall handelt es sich um den Fall, dass die lokale Belastung größer ist als die lokale Kapazität, die das Bauteil aufnehmen kann [61]. Der tatsächliche Ausfallprozess ist selten deterministisch und oft schwierig nachzuweisen, z.B. beim Korrosionsprozess [45].

Gründe für den Ausfall sind entweder Überbelastung oder Aufsummierung von Einzel-schäden. Ob ein Ausfall stattfindet, ist hierbei abhängig von den Eigenschaften des verwendeten Materials, dem Bauteildesign und der Fehlbelastung des Bauteils [62]. Im Fall des Korrosionsprozesses entspricht die Wanddicke eines Bauteils der Kapazität, die das Bauteil aufnehmen kann. Wird die Bauteildicke durch einen oder mehrere Materialdefekte vermindert, kann dies zum Ausfall des Bauteils führen [63].

Dem Bauteildesign wird eine weitere wichtige Rolle zugemessen, da durch die schadenstolerante Auslegung trotz vorhandener Materialdefekte eine verlängerte Lebenszeit des Bauteils sichergestellt werden kann [46].

Aus Sicht der strukturellen Integrität wird daher eine erforderliche Restwanddicke als wesentliche Größe für den Einsatz eines Bauteils gewählt [44].

Defektparameter in der zerstörungsfreien Prüfung

Kann ein Defekttyp an einer bestimmten Position mit einer bestimmten Geometrie zum Ausfall führen [64], so wird dieser als kritisch definiert [65],

Jedoch werden keine festgelegten kritischen Defekte oder noch akzeptable Defektgrößen allgemeingültig angegeben [66]. Letztendlich ist die Bestimmung der kritischen Defektparameter und die Höhe notwendiger Sicherheitsfaktoren im Qualitätsbericht des Bauteilherstellers festgelegt (z.B. unter [67]).

Häufig wird die Defektgröße als der entscheidende Parameter für den Ausfall eines Bauteils herangezogen [68]. Dabei ermöglicht die ZfP die Lebenszeit eines Bauteils zu verlängern, da Bauteile mit großen Materialdefekten, die wahrscheinlich früher ausfallen, ausgesondert werden können [59]. Die von der ZfP nicht detektierten kritischen Materialdefekte müssen durch das Bauteildesign ausgeglichen werden.

Daher ist für die Gewährleistung der strukturellen Integrität die Bewertung der Auffindbarkeit eines Defekts für ein ZfP-System notwendig [65].

Aus Sicht der ZIP wird als passendes Gegenstück zum kritischen Defekt der äquivalente Defektgrößenparameter eingeführt, der sicher detektiert werden soll [69].

Zusammengefasst kann gesagt werden, dass die Auffindwahrscheinlichkeit von kritischen Defektgrößen und die Auslegung des Designs und die Verwendung des Materials für das Bauteil stark miteinander zusammenhängen [46]. Die Kommunikation zwischen Konstrukteur, Bruchmechaniker und Prüfer ist daher für sicherheitsrelevante Bauteile eine Notwendigkeit, die jedoch häufig zu kurz kommt.

Bestimmung der Auffindwahrscheinlichkeit abhängig von der kritischen Defektgröße

Im Allgemeinen wird ein kleiner Materialdefekt auch eine kleine Anzeigengröße hervorrufen. Unter idealen Prüfbedingungen können selbst kleine Anzeigen sicher detektiert werden: die Auffindwahrscheinlichkeit ist daher selbst für kleine Defekte hoch und das ZfP-System wird als zuverlässig erachtet. In der praktischen Anwendung jedoch steigt der Einfluss von unvorhersehbaren Quellen – die Messunsicherheit wird größer. Dadurch liegt eine größere Menge von nicht detektierten Anzeigen vor, dies gilt dann auch für größere Defekte (Abbildung 2.4) [53].

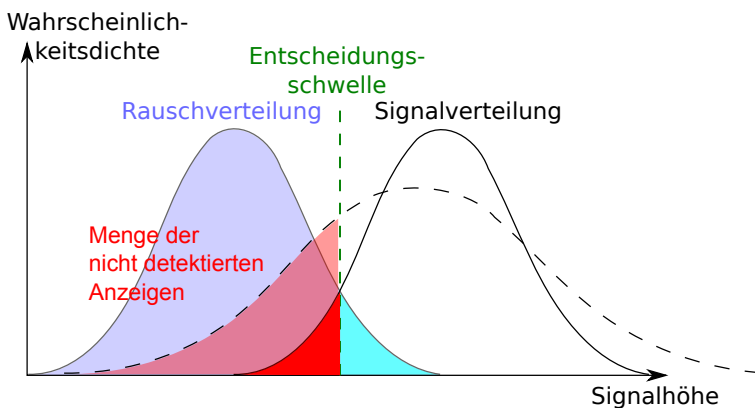


Abbildung 2.4: Einfluss der Messunsicherheit (Streuung der Signalverteilung) auf die Anzahl der nicht detektierten Defektanzeigen

Für eine Bewertung wird nun die Auffindwahrscheinlichkeit hinsichtlich der Defektgröße mit Hilfe von Stichproben geschätzt [45]. Die Schätzung kann parametrisch mit Verwendung von Verteilungsmodellen [68] oder nicht-parametrisch erfolgen [70]. Dabei können Defektgrößen danach bewertet werden, ob sie detektiert wurden (qualitative Signalantwort) [68] oder

2.2 Auffindwahrscheinlichkeit eines zerstörungsfreien Prüfsystems

es wird ein Zusammenhang zwischen der Defektgröße und der Detektierbarkeit des Defekts verwendet (quantitative Signalantwort) [15]. Die Auswahl des zu verwendenden Vorgehens ist abhängig von der Art der Daten, von der Anzahl der durchgeführten Experimente und den Informationen über die physikalischen Grundlagen des ZfP-Verfahrens.

Die quantitative Bewertung der Auffindwahrscheinlichkeit, basierend auf einem parametrischen Modell, wird im Allgemeinen nicht-parametrischen und qualitativen Verfahren vorgezogen. Der Grund liegt in aussagekräftigeren Ergebnissen - selbst bei geringer Datenmenge [15].

Außerdem können auch Abschätzungen von Auffindwahrscheinlichkeiten außerhalb der geprüften Defektgrößen abgeschätzt werden [45], [46]. Des Weiteren können zusätzliche Informationen über Sättigungsbereiche der Signale und Defektsignale, die im Rauschen verschwinden, in die Bewertung einbezogen werden [15].

Allerdings ist die Auswahl von passenden Modellen bei dem parametrischen Ansatz eine mögliche Schwachstelle. Daher muss das Modell mit Vorsicht ausgewählt und überprüft werden [71]. Eine Extrapolation ist zwar möglich, jedoch nicht immer zweckmäßig und kann zu falschen Schlussfolgerungen führen [52]. Trotz der notwendigen experimentellen Verifizierung ist der geringere Aufwand hinsichtlich der experimentellen Datenerhebung ein wesentlicher Vorteil für den Einsatz [64] und soll daher im Weiteren im Fokus stehen.

2.2.2 Bewertung der Auffindwahrscheinlichkeit mit Verwendung der quantitativen Signalhöhe

Bei vielen ZfP-Verfahren basiert die Auswertung und die Defektdetektion auf der Basis eines quantitativen Signals. Die mögliche Signalskala ist dabei so fein untergliedert, dass häufig von einem fein untergliederten diskreten Signal ausgegangen werden kann. Dies gilt z.B. für die elektrische Signalantwort eines Detektorpixels bei der digitalen RT. Reagiert ein kontinuierliches (oder fein untergliedertes diskretes) Signal auf einen Defektparameter, so kann ein Zusammenhang zwischen Signal und Parameter hergestellt werden [72]. Dieser kann mit mathematischen Variablen beschrieben werden [73].

Die einzelnen Variablen des Zusammenhangs können jedoch nicht rein aus der Theorie bestimmt werden, sondern müssen mit Hilfe von Experimenten ermittelt werden [74]. Für ein idealisiertes ZfP-System ruft ein großer Materialdefekt in der Regel ein großes Signal hervor. Dieser Zusammenhang zwischen Defektgröße und Signal bietet die Grundlage für die Signal-Antwort-POD (z.B. Reflektivität von Rissflanken bei der Ultraschallprüfung oder radiographischer Kontrast von Volumendefekten). Die Abschätzung des physikalischen Zusammenhangs (intrinsic Faktor) soll hinsichtlich der Bewertung von ZfP-Verfahren im Weiteren im Fokus stehen.

Hierbei wird mit Hilfe der Regressionsanalyse der physikalische Zusammenhang zwischen dem Signalparameter \hat{a} und dem Defektparameter a beschrieben [15].

Regressionsanalyse zwischen Signal und Defektparameter

Die physikalische Theorie führt zu der funktionellen Form des Modells (Abhängigkeit und Verteilungsmodell), während Experimente und beobachtete Daten zur Abschätzung der unbekannt Parameter benötigt werden [74].

Das Ziel ist es nun die Parameter aus Beobachtungen abzuschätzen und den Zusammenhang für die POD-Bewertung zu nutzen (konventionelle Signal-Antwort-POD). Dabei wird für jeden Defekt das maximale Signal und die maximale Defektgröße untersucht [15]. Die Beschränkung auf das maximale Signal vermeidet eine mehrfache Zählung von Anzeigen und daher eine Überbewertung der Fähigkeit des ZfP-Systems [75]. Neben dem rein theoretischen Zusammenhang bleiben stochastische Einflussgrößen (z.B. Messfehler) bestehen, deren Abschätzung das probabilistische Modell vervollständigt. Im Fall eines linearen Zusammenhangs zwischen dem Defektparameter a und dem dabei hervorgerufenen Signalparameter \hat{a} kann die Abhängigkeit wie folgt beschrieben werden:

$$\hat{a} = \beta_0 + \beta_1 \cdot a + \epsilon. \quad (2.2)$$

Während β_0 und β_1 die Parameter der Geradengleichung darstellen, steht das ϵ für die Reststreuung der Messwerte um die idealisierte Gerade (siehe Abbildung 2.5)⁷.

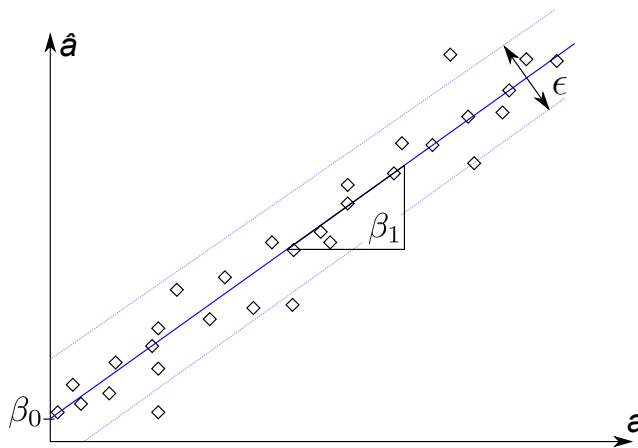


Abbildung 2.5: Abschätzung einer Gerade durch Datenwerte

⁷Im Allgemeinen sind \hat{a} , a und ϵ Vektoren. Für Abbildung 2.5 wird die Anforderung der homogenen Streuung aus Gleichung 2.8 vorgegriffen.

2.2 Auffindwahrscheinlichkeit eines zerstörungsfreien Prüfsystems

Der zufällige Fehlerterm, der durch die Unkenntnis des wahren Wertes entstanden ist, wird durch den Erwartungswert

$$E(\epsilon) = 0 \quad (2.3)$$

und die Streuungsmatrix

$$V(\epsilon) = \sigma^2 I \quad (2.4)$$

beschrieben. σ^2 ist hierbei die Varianz und I stellt die Informationsmatrix dar, auf die später noch genauer eingegangen wird.

Funktionaler Zusammenhang und reale Beobachtungsverteilungen

In Fällen in denen weder theoretische Überlegungen noch vorhergegangene Versuche ein vollständiges Modell für den Zusammenhang anbieten, beeinflussen meist Schwankungen und Beobachtungsfehler die Datenwerte. Es kann nicht der wahre Wert beobachtet werden, sondern vielmehr ein wahrer Wert mit einem zufälligen Element. Die Parameter des Zusammenhangs werden durch die Schätzung von Daten und die Stichprobenunsicherheit beeinflusst [73].

Bei der Schätzung der Parameter β_0 und β_1 können \hat{a} und a nicht direkt beobachtet werden, sondern nur die zwei zufälligen Variablen ξ und η . Diese sind definiert durch die wahren Werte \hat{a} , a und einen jeweils zufälligen Fehler δ und γ , die jeweils abhängig von der Messung i sind:

$$\xi_i = \hat{a}_i + \delta_i \quad (2.5)$$

mit

$$\eta_i = a_i + \gamma_i \quad (2.6)$$

wobei $i = 1, 2, \dots, n$.

Da die Parameter der Geradengleichung β_0 und β_1 nicht bekannt sind und durch $\hat{\beta}_0$ und $\hat{\beta}_1$ mit ξ und η geschätzt werden müssen [74], unterliegt die Schätzung einem Einfluss von beiden zufälligen Fehlern δ und γ .

Wird im Fall der ZfP davon ausgegangen, dass der Defektparameter a exakt bestimmt wurde, beispielsweise durch genaues Einbringen von künstlichen Defekten oder durch das Ausmessen durch die Metallographie, überwiegt δ_i , wodurch annähernd $\epsilon_i = \delta_i$ angenommen werden kann. Die Abschätzung der Parameter der Geradengleichung β_0 und β_1 unterliegt daher hauptsächlich dem Einfluss des Fehlers δ . Für eine korrekte Bewertung ist es daher

notwendig die gesamte Reststreuung ϵ möglichst gering zu halten (siehe Abbildung 2.5) [45].

Wird die abgeschätzte Gerade im „ \hat{a} vs. a “-Diagramm betrachtet, so verläuft sie durch die Punktwolke der Datenpunkte, die um die Gerade streuen. Im allgemeinen Fall haben die experimentellen Ergebnisse einen zufälligen Charakter. Die Ergebnisse werden von einer großen Anzahl von Größen beeinflusst und liegen teilweise nicht unter der Kontrolle des Experimentators. Daher ist es notwendig, ein stochastisches Modell für die Reststreuung zu wählen [76].

Für den Zusammenhang zwischen \hat{a} und a entsteht die Reststreuung durch die Abhängigkeit vom Defekt bzgl. Größe, Form, Orientierung und Position, die besonders bei realen Defekten schwer zu bestimmen sind [77]. Weitere Einflüsse sind zum Beispiel die Varianz des Defektsignals, geometrische Varianz oder Varianz im Entscheidungskriterium.

Die Reststreuung zwischen beobachteten und vorhergesagten Wert $\epsilon_i = \hat{a}_i - \bar{a}_i$ [47], wird meist durch ein theoretisches Verteilungsmodell angenähert [78].

Die Verteilungsannahme vereinfacht das weitere Vorgehen und stellt häufig eine überraschend gute Näherung dar [79]. Häufig wird bei der Streuung von Signalwerten um die Regressionsgerade von einer Normalverteilung N mit dem Mittelwert μ und der Standardabweichung σ ausgegangen [77] (Siehe Gleichung 2.7 für die Dichtefunktion).

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right), \quad (2.7)$$

mit $F(x) \sim N(\mu, \sigma)$ und $x \in \mathbb{R}$. Der Einsatz der Normalverteilung dient häufig als idealisiertes Modell für empirische Häufigkeitsverteilungen. Sie spielt als theoretische Verteilung eine wichtige Rolle, da viele andere Verteilungen gut approximiert werden können [80].

Eine weitere Annahme, die häufig getroffen wird und die Bewertung sehr vereinfacht, ist die Homogenität der Reststreuung ϵ (Homoskedastizität).

Dies bedeutet für das ZFP-Signal, dass an n -vielen Positionen der Defektparameter ϵ_i gleich sein soll [81]:

$$V(\epsilon_1) = V(\epsilon_2) = \dots = V(\epsilon_n) = \sigma_\epsilon^2. \quad (2.8)$$

Beobachtete Datenwerte sind dann paarweise unkorreliert, wenn die Kovarianzmatrix bis auf die Diagonale null ist [52]. Es kann von einer Unabhängigkeit der Experimente und Messwerte ausgegangen werden, falls eine korrekt durchgeführte Versuchsplanung stattgefunden hat. Liegt die konstante Varianz um den Mittelwert einer linearen Funktion mit unbekanntem Parameter vor und die beobachteten Werte sind paarweise unkorreliert, so kann selbst für kleine Stichprobengrößen ein lineares Modell genutzt werden [73].

2.2.3 Likelihoodstatistik

Wenn nun normal-verteilte Werte mit homogener Streuung angenommen werden, so kann die Maximum Likelihood Methode - engl. maximum likelihood estimation (MLE) eine Schätzung der Parameter für den Zusammenhang liefern [73]. Der Begriff der Mutmaßlichkeit (Likelihood) beschreibt hierbei die Inferenz von konkreten Datenwerten x_i auf den unbekannt Parameter $(\mu; \sigma) = \theta \in \Theta = \mathbb{R}, \mathbb{R}_+$ [45].

Die Idee hinter der Likelihoodstatistik ist, dass ein Modell an die vorhandenen Daten mit Hilfe von verschiedenen Modellparameterkombinationen angepasst wird. Kombinationen mit höheren Ergebnissen der Likelihoodfunktionen für die vorhandenen Daten sind plausibler und näher an den empirischen Daten als solche mit geringeren Likelihoodergebnissen [82].

Bei der Schätzung des allgemeinen Parameters (θ) werden verschiedene Likelihoodfunktionen $L(\theta)$ betrachtet. Die Entscheidung für die wahrscheinlichste Parameterkombination fällt für den größten Likelihoodwert [82], [45].

Die Unterscheidung zwischen zwei Parameterwerten θ' und θ'' aufgrund einer Stichprobe ist umso leichter, je stärker sich die Werte der Likelihood-Funktion unterscheiden [45].

Die Likelihoodfunktion für die Gesamtmenge wird allgemein durch die Dichte der Einzelversuche $(x_1 \dots x_n)$ gegeben, falls die Daten statistisch unabhängig sind [52]:

$$L(\theta, x) = f(x_1, \dots, x_n; \theta) = f(x_1; \theta) \cdot \dots \cdot f(x_n; \theta). \quad (2.9)$$

Die Likelihoodfunktion für eine Stichprobe aus $N(\mu, \sigma^2)$ mit zwei unbekannt Parametern ist:

$$L(\mu, \sigma^2, x) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right). \quad (2.10)$$

Eigenschaften und Forderungen der Likelihoodfunktion

Die Likelihoodstatistik bietet einige für die Statistik wünschenswerte Eigenschaften, so beispielsweise die Additivität. Sie beschreibt die Tatsache, dass unabhängige Daten kombinierbar sind [83].

Ein weiterer wichtiger Punkt ist die Konsistenz in der Likelihoodbewertung: Zwei Funktionen sind gleich bedeutend, wenn beide Funktionen die gleiche Menge an Wissen über die Daten liefern. Auf lange Sicht sollte die Funktion mit der größten zusätzlichen Information die richtige Funktion sein [83]. Die Konsistenz ist bei den meisten praktischen Problemen gegeben [74].

Eine direkte Interpretation der Likelihoodfunktion in Termen der Wahrscheinlichkeit ist hingegen nicht länger möglich [82]: Die Wahrscheinlichkeiten über alle Ergebnisdaten ergibt aufsummiert das sichere Ereignis. Während dies für Likelihoodfunktionen nicht gilt, da sie als Funktion von Parametern und Hypothesen bei fixen Ergebnisdaten angesehen werden [83]. Aus der Bayes'schen Sichtweise gilt: die Likelihood $L(H|Data)$ von der Hypothese H gegeben durch die Daten und einen bestimmten Modell sind proportional zu $P(Data|H)$, während die Proportionalitätskonstante beliebig ist [83].

Schätzung mit der Maximum Likelihood Methode

Wird die Likelihoodfunktion maximiert, so kann in häufigen Fällen ein akzeptabler Schätzwert erreicht werden [74]. Das Vorgehen wird MLE genannt. Die MLE kann auch für den Fall eingesetzt werden, dass ein Teil zensierter Daten vorliegt [84]. Bei zensierten Daten handelt es sich um Werte die außerhalb des Betrachtungsraums liegen und somit nicht die volle Aussagekraft besitzen [45]. Liegt ein annähernd normal-verteiltes Modell der Fehlerterme ϵ vor und handelt es sich um einen unzensierten Datensatz, so vereinfacht sich die Maximum-Likelihood Analyse zur Methode der kleinsten Fehlerquadrate [45].

Der wahre Parameter θ wird mit der Schätzung aus den Beobachtungen $\hat{\theta}$ bestimmt [85].

$$L(x|\hat{\theta}) = \operatorname{argmax}_{\theta' \in \mathbb{R}, \mathbb{R}_+} L(x|\theta'). \quad (2.11)$$

Wird eine Lösung für das Schätzproblem gesucht, so wird meist von der ersten Ableitung der Likelihoodfunktion ausgegangen. Werden zwei Parameter θ' und θ'' verglichen, so besitzt die Funktion an einer Stelle θ ein lokales Extremum:

$$\frac{\partial L}{\partial(\theta_i)} = 0, \quad (2.12)$$

wobei $\theta_i = (\theta_1, \theta_2, \dots, \theta_n)$. Für den Fall einer Normalverteilung siehe [46].

Verifizierung des angewendeten Modells

Wurden die Schätzwerte $\hat{\beta}_0$, $\hat{\beta}_1$ und $\hat{\epsilon}$ berechnet, so kann bewertet werden, ob der Zusammenhang richtig erkannt wurde [45]. Die Überprüfung der statistischen Analyse mit den Daten ist ein essentieller Schritt [86]. Es gibt mehrere Methoden zu Überprüfung der Verteilungen, beispielsweise den Residuengraph, das Bestimmtheitsmaß oder die Wahrscheinlichkeits-Plots (QQ-Graph) [45], [86]. Die Überprüfung der Verteilungsfunktion, auch hinsichtlich Ausreißer, ist wichtig. Im Fall einer Symmetrie ist jedoch die Normalverteilung eine gute Annahme [76].

2.2 Auffindwahrscheinlichkeit eines zerstörungsfreien Prüfsystems

Um das Verteilungsmodell zu schätzen, sollten die Histogramme der Residuen konstruiert werden, womit die Normalität und der Mittelwert der Residuen abgeschätzt werden können. Dabei sollte das Histogramm eine ungefähre Glockenform besitzen und außerdem symmetrisch zum Nullpunkt liegen. Im Allgemeinen ist die Form und die Symmetrie wichtiger als die strikte Normalität [47].

Die Anforderungen an die konstante Varianz der Residuen und die Linearität des Zusammenhangs kann durch einen Graph für die Residuen (Ordinate) und die durch das Modell beschriebenen Werte (Abszisse) bewertet werden: Liegen die Residuen innerhalb eines horizontalen Bands mit der Nullkoordinate im Zentrum und es kann kein systematischer Trend (positiv oder negativ) wahrgenommen werden, so kann eine konstante Varianz und der lineare Zusammenhang angenommen werden. Die konstante Varianz ist hierbei wiederum wichtiger als die Annahme, dass es sich um eine Normalverteilung handelt [47].

Rechnerisch kann das Ergebnis der Regressionsanalyse mit Hilfe des Korrelationskoeffizienten R^2 bewertet werden. Dieser beschreibt die Varianz, die mit Hilfe des Regressionsmodells beschrieben werden kann (siehe [52]).

Eine graphische Möglichkeit zur Bewertung des Verteilungsmodells ist der QQ-Plot (Wahrscheinlichkeitsplot) [86]. Dabei werden auf einem Graphen die vorhergesagten Werte gegenüber den wirklichen Datenwerten aufgetragen. Damit ist es möglich die Qualität der Regression zu evaluieren. Die aufgetragenen Punkte sollten dabei ungefähr auf der $y = x$ Geraden liegen [47]. Streuen die Punkte systematisch um Linie $y = x$, handelt es sich um ein Indiz für ein schlechtes Verteilungsmodell. Im Weiteren soll später die Normalverteilungsannahme in einem QQ-Plot geprüft werden, um zu sehen, wie die Form und die Textur der Verteilung ist [58].

2.2.4 Berechnung der Auffindwahrscheinlichkeit

Durch die Schätzung des Zusammenhangs zwischen Defektparameter a und Signalparameter \hat{a} konnten nun einerseits die Parameter für den Zusammenhang erfasst werden und andererseits das Verteilungsmodell für die Streuung der Daten beschrieben werden. Die POD definiert nun die Defekte a als gefunden, deren Signal oberhalb einer Entscheidungsschwelle \hat{a}_{dec} liegen (siehe Abbildung 2.6).

Dadurch dass die Signale \hat{a} für einen Defekt streuen, kann nur eine statistische Aussage getroffen werden, welcher Defekt detektiert werden kann. Hierfür wird die kumulative Verteilungsfunktion, die mit Hilfe der MLE abgeschätzt wurde, genutzt. Mathematisch kann die Auffindwahrscheinlichkeit für einen Defekt a bei der Annahme einer Normalverteilung wie folgt beschrieben werden:

$$POD = \int_{\hat{a}_{dec}}^{\infty} N(\mu, \sigma) d\hat{a}, \quad (2.13)$$

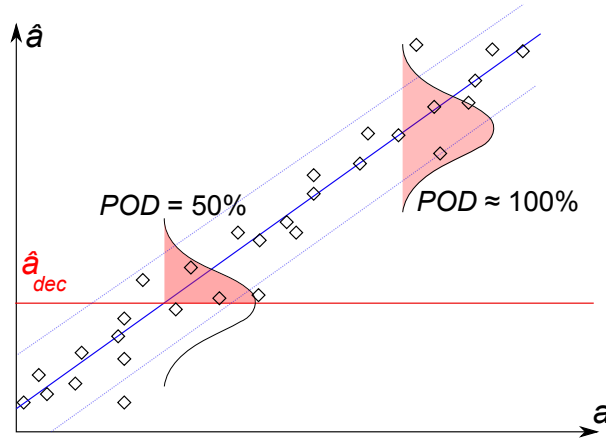


Abbildung 2.6: Bestimmung der Auffindwahrscheinlichkeit an der Stelle des Defektparameters a in einem eindimensionalen Modell

wobei für den oben eingeführten theoretischen Zusammenhang folgendes gilt [15]:

$$\mu = \frac{(\hat{a}_{dec}) - \beta_0}{\beta_1}, \quad (2.14)$$

$$\sigma = \frac{\sigma_\epsilon}{\beta_1} \quad (2.15)$$

Entscheidungskriterium

Während das Verteilungsmodell und die Schätzwerte schon erklärt wurden, stellt sich die Frage, wie die Entscheidungsschwelle \hat{a}_{dec} definiert und bestimmt wird. Dabei ist zu bedenken, dass die Entscheidungsschwelle einen großen Einfluss auf die POD hat [87].

Ein Defekt, der das Signal \hat{a} hervorruft, wird als gefunden bezeichnet, wenn \hat{a} größer als \hat{a}_{dec} ist. Das Entscheidungskriterium bestimmt nun, ob es sich um eine Anzeige bei einem Signal handelt, oder nicht. Ein Defekt, der durch ein Signal \hat{a} beschrieben wird, das genau auf dem Entscheidungskriterium \hat{a}_{dec} liegt, wird zu 50% gefunden, falls der Versuch oft genug durchgeführt wird [47]. Dies erklärt sich aus der Tatsache, dass durch die Streuung im Mittel 50% der Messwerte über und 50% unter der Schwelle liegen. Die Entscheidungsschwelle

2.2 Auffindwahrscheinlichkeit eines zerstörungsfreien Prüfsystems

verändert somit die mittlere Auffindwahrscheinlichkeit [88]. In dem praktischen Einsatz der ZfP kann jedoch eine Entscheidungsschwelle meist nicht frei gewählt werden: Prüfanweisung, Verfahren und zu findende Defekte legen diese fest. Entweder wird die Entscheidungsschwelle auf Basis der Rauschschwelle oder anhand eines Referenzdefekts bestimmt.

Im ersten Fall für die Abhängigkeit der Rauschschwelle wird die Anzahl der Fehlalarme auf eine akzeptable Menge reduziert [47]. Teilweise wird hier auch mit zusätzlichen Abschlägen gearbeitet (beispielsweise Ultraschallprüfung: dreifaches SNR [89], [90] oder mit Hilfe der Operationscharakteristik eines Empfängers - engl. receiver operating characteristic (ROC)-Analyse [91]). Dieses Vorgehen stützt sich auf den Fehler erster Art, der in der Anzeigenbewertung gemacht werden kann.

Im zweiten Fall wird die Signalthöhe eines Referenzdefekts als Entscheidungsschwelle festgelegt. Dabei sollte der Referenzdefekt so gewählt werden, dass er den gesuchten Defekt mit ähnlichem Signalverhalten abbildet. Es wird beispielsweise in der Ultraschallprüfung der Referenzdefekt auf 80% Bildschirmhöhe bei der Prüfung gelegt [92]. Im zweiten Verfahren steht der Fehler zweiter Art im Fokus.

In vielen Anwendungen wird der Schwellwert unabhängig von dem Defektparameter angesehen. Der Schwellwert bildet eine Parallele zur Abszisse. Im Fall der Radiographie werden nun auch nicht konstante Entscheidungsschwellen für die POD zum Einsatz gebracht. Hierbei spielt das tatsächliche Detektionsverhalten für eine Anzeige eine Rolle [93]. Auf diesen Fall wird später weiter eingegangen.⁸

Entstehung und Einsatz der POD-Kurve

Unter einer POD-Kurve wird die kumulierte Verteilung verstanden, wobei auf der Ordinate die kumulierte Wahrscheinlichkeit aufgetragen ist, während auf der Abszisse die Skala des zu bewertenden Defektparameters a aufgetragen ist (siehe Abbildung 2.7) [64].

Im Moment handelt es sich bei dieser POD-Kurve nur um eine Momentaufnahme. In ungefähr 50% der Fälle würde bei einer ähnlichen Durchführung von Experimenten und Auswertung ein schlechteres Ergebnis entstehen [47]. Die gewonnenen Schätzwerte müssen für den konkreten Einsatz für die strukturellen Integrität auf lange Sicht validiert werden [73]. Dennoch kann diese Kurve durchaus für die strukturelle Zuverlässigkeitsberechnung als Mittelwertkurve, beispielsweise bei einem Monte-Carlo-Ansatz nützlich sein [95].

2.2.5 Vertrauensbereich

Die Abschätzung der Auffindwahrscheinlichkeit nur basierend auf den MLE- Schätzwerten (aus Abschnitt 2.2) könnte zu einem zweifelhaften Ergebnis führen [45]. Der Grund hierfür

⁸Es sei angemerkt, dass sich im tatsächlichen Entscheidungsprozess neben den Schwellen für die zensierten Daten (Rausch-, Sättigungsschwelle) und der Entscheidungsschwelle, noch weitere Schwellen finden lassen. Beispielsweise die Registrierschwelle, die einen Defekt zwar ins Prüfprotokoll aufnimmt, ihn jedoch noch nicht als kritischen Defekt kennzeichnen lässt [94].

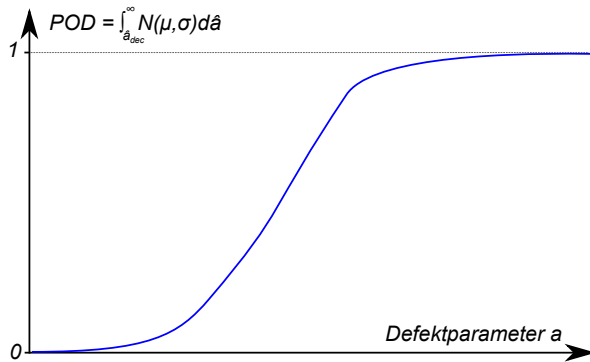


Abbildung 2.7: Typische POD-Kurve zwischen 0 und 1 aufgetragen für den Defektparameter a

ist, dass durch die Verwendung der reinen Schätzwerte nicht mehr die vollständige Dateninformation zur Verfügung steht [58].

Die Menge der Daten ist ein wichtiges Indiz für die Aussagekraft der POD [96]. In der Versuchsplanung wird die Stichprobenanzahl festgelegt, die im Experiment verwendet werden soll [88]. Während in vielen Fachgebieten des Ingenieurwesens die Stichprobenanzahl und somit auch die Folgerung auf die Grundgesamtheit keine Herausforderung darstellt, ist im Fall der ZfP-Prüfung innerhalb Werkstofftechnik jede Stichprobe meist mit einem fehlerbehafteten Bauteil verbunden. Daher wurde auch über Jahrzehnte immer weiter versucht, ein sinnvolles Ergebnis zu erlangen, selbst bei geringen Datenmengen, um die Bewertung wirtschaftlich zu halten.

Allerdings tragen nicht alle Datenpunkte gleich viel zur Beschreibung der POD-Kurve bei. Defektgrößen mit einer Auffindwahrscheinlichkeit nahe 0% oder 100% besitzen nur wenig Informationsgehalt [88]. Weiterhin ist zu betrachten, dass im Fall eines Regressionsmodells die Bewertung an den Punkten, an denen Daten vorhanden sind, die beste angepasste Gerade erreicht. Daher ist es ratsam, die POD für den wichtigen Bereich, d.h. Bereich der höchsten Steigung zu ermitteln [36].

Im Allgemeinen werden für die Bewertung mindestens 30-40 verschiedene Defekte mit der Auffindwahrscheinlichkeit zwischen 10% und 90% gefordert [36]. Werden diese Anforderungen erfüllt, so kann von einer repräsentativen Stichprobengröße für die POD ausgegangen werden. Die notwendigen Annahmen hinsichtlich des mathematischen Modells können getroffen werden [36].⁹

Die Aussage der POD-Kurve ist dennoch beschränkt auf den Fall der Experimente. Es handelt sich dabei sozusagen um eine Momentaufnahme des beobachteten Zusammenhangs

⁹Weiteres Vorgehen hinsichtlich der Versuchsplanung und korrekten Versuchsdurchführung findet sich in [47].

2.2 Auffindwahrscheinlichkeit eines zerstörungsfreien Prüfsystems

zwischen a und \hat{a} . Um eine Inferenz für die allgemeine Aussage über das Prüfsystems zu ermöglichen, reicht die reine Schätzung nicht aus [58].

Notwendigkeit eines Vertrauensbereichs

Der Vertrauensbereich verdeutlicht die Unsicherheit, die durch die begrenzte Datengrundlage entsteht [45].

Der Schätzwert ist der wahrscheinlichste wahre Wert für die Stichprobe. Wird ein Parameterwert mit großer Entfernung vom Schätzwert betrachtet, so handelt es sich mit großer Wahrscheinlichkeit nicht um den wahren Wert. Jedoch kann der weit entfernte Wert auch nicht als wahrer Wert ausgeschlossen werden. Der Bereich, in dem der wahre Wert liegen könnte, wird in der Praxis häufig prozentual festgelegt: So wird beispielsweise das 95% Vertrauensintervall für einen geschätzten Parameter gesucht. Dabei gesteht man der Schätzung eine 5% Chance zu, dass der Parameter außerhalb des Intervalls liegt, und somit die Entscheidung auf den falschen Parameter fällt.

Allgemein wird zwischen dem Vertrauensbereich hinsichtlich der Parameter und dem Toleranzbereich, der auch die Defektvariation umfasst, unterschieden [97]: Der Vertrauensbereich drückt die statistische Unsicherheit aus. Die statistische Unsicherheit wird sehr klein für sehr viele Daten. Der Toleranzbereich fällt jedoch mit steigender Anzahl der Daten nicht auf Null [97].¹⁰

Im Weiteren wird der Vertrauensbereich genutzt, da die Anzahl begrenzt ist. Der 95%-Vertrauensbereich für die POD, der erstmals bei der Triebwerksprüfung in der militärischen Luftfahrt eingesetzt wurde, hat sich auch allgemein als Standard für die Bewertung eines ZfP-Systems etabliert [88]. Als Kenngröße wurde hierbei $a_{90/95}$ benannt. Sie ist durch den Schnittpunkt zwischen der unteren Vertrauensgrenze von 95% und der 90%-Auffindwahrscheinlichkeit definiert. Die so berechnete Defektgröße bezeichnet eine Fehlergröße bei der 10% nicht durch das ZfP-Verfahren gefunden wurde. Diese Aussage wird mit 95% Sicherheit getroffen. In 5% aller Fälle wäre das Ergebnis falsch. Die Kenngröße $a_{90/95}$ wird für Lebenszeitmodelle und zur Bestimmung des nächsten Prüfzeitpunktes verwendet [57].

Ist die Anzahl der Daten gering, so ist der Vertrauensbereich breit. Durch die Breite des Vertrauensbereichs wird auch die Kenngröße $a_{90/95}$ groß. Der Vertrauensbereich und die Kenngröße $a_{90/95}$ bilden somit die Unsicherheit durch die begrenzte Anzahl von Datenpunkten ab (siehe Abbildung 2.8).

¹⁰Daher spielt der Toleranzbereich eine wichtige Rolle für die Bewertung der ZfP bei der strukturellen Integrität, für denn Fall dass viele Datenwerte vorhanden sind und nicht alle Einflüsse erfasst wurden [97].

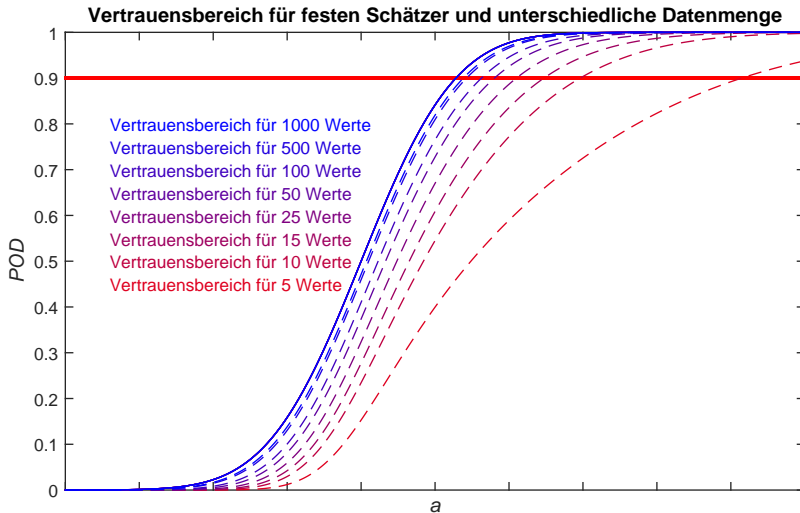


Abbildung 2.8: Einfluss der Datenanzahl auf den Vertrauensbereich der POD-Kurve (theoretische Daten)

Herangehensweise zur Berechnung eines Vertrauensbereichs

Die Art des Denkens und die Herangehensweise an Vertrauensbereiche führt zu verschiedenen Lösungsansätzen, die jedoch häufig zu gleichen Ergebnissen führt [73]. Jedoch ist die Auswahl des Ansatzes abhängig von der späteren Aussage und dem Einsatzgebiet [82], [73].

Für den Vertrauensbereich einer POD hat sich ein klassischer Häufigkeitsansatz durchgesetzt, den sowohl parametrische als auch nicht-parametrische Modelle nutzen können. Für die Bewertung mit Hilfe der Signal-Antwort-POD bietet sich ein Vertrauensbereich basierend auf einem parametrischen Modell an. Durch die Anforderungen an den linearen Zusammenhang von Signal und Defektparameter und der Annahme eines Verteilungsmodells liegt dieser Ansatz nahe.

Für die Berechnung eines Vertrauensbereichs für die POD werden die Parameter der Regression (β_0 , β_1 , ϵ) und die der Verteilungsfunktion (μ , σ) auf ihre Aussagekraft (bzw. Unsicherheit) bewertet.

Es werden somit gleichzeitig mehrere Parameter betrachtet. Dabei ist die getrennte Auswertung, für die jeder Parameter ein einzelnes Intervall erhalten würde, nicht lösbar [73].

Ein Ansatz zur Lösung ist die Einführung des simultanen Vertrauensbereichs. Hierfür wird ein Modell eingeführt, das die Verbindung zwischen Mittelwert μ und Streuung σ der Verteilungsfunktion abgeschätzt [45].

2.2 Auffindwahrscheinlichkeit eines zerstörungsfreien Prüfsystems

Der Vertrauensbereich wird auf $100(1 - \alpha)$ Prozent festgelegt. α stellt hierbei die Irrtumswahrscheinlichkeit dar, die noch akzeptiert wird (z.B. α gleich 5% für den 95%-Vertrauensbereich). Die Grenze zwischen dem Vertrauens- und dem Irrtumbereich wird als das Quantil x_p definiert.

Die Aufgabe besteht nun darin, das Quantil x_p so zu bestimmen, dass die Parameter der Verteilungsfunktion (μ, σ) für die gegebene Wahrscheinlichkeit $(1 - \alpha)$ im Vertrauensbereich liegen.

Berechnung eines Vertrauensbereichs für Verteilungsparameter

Eine Hauptschwierigkeit der Bestimmung des Vertrauensbereichs für die Parameter der Verteilungsfunktionen ist, dass nicht von einer konstanten Breite ausgegangen werden kann, da dies zu einer Überbewertung führen würde. Dabei würde ein zu breiter Vertrauensbereich entstehen, wodurch Auffindwahrscheinlichkeiten von über 1 oder unter 0 entstehen können [98].

Um dies zu vermeiden, wird eine Transformation in den Parameterraum durchgeführt. Im Parameterraum liegen alle Kombinationen der Parameter θ (die Werte μ und σ) [99]. Es soll gelten:

$$P[Q(\theta) \leq \gamma_{CL}] = 1 - \alpha. \quad (2.16)$$

Wobei Q den Bereich beschreibt, in dem der unbekannte transformierte Parameter Θ mit einer Wahrscheinlichkeit von $1 - \alpha$ liegt [98]. γ_{CL} ist hierbei die Grenze für den Bereich Q (Quantil). Auf der Abszisse des Parameterraums wird der unbekanntes Mittelwert μ der POD-Verteilung und auf der Ordinate die unbekanntes Standardabweichung σ aufgetragen. Auf der Applikate ist die Wahrscheinlichkeitsdichte der Werte μ und σ aufgezeigt.

Der Schätzwerts $\hat{\mu}$ und $\hat{\sigma}$ ist der Modus der Verteilung (für eine bivariate Normalverteilung auch der Mittelwert). Für die Berechnung des Vertrauensbereichs um den Schätzwert $\hat{\mu}$ und $\hat{\sigma}$ wird die Dichte über einen Bereich Q integriert [99].

Für die Annahme, dass der Schätzung eine ausreichend große Anzahl von Daten zugrunde liegt, wobei nur wenige oder keine Daten zensiert wurden, kann von einer bivariaten Normalverteilung als Dichte-Modell ausgegangen werden [100], [101].

Soll ein Vertrauensbereich für μ und σ mit einem festen Quantil beschrieben werden, so wird eine Ebene mit festem p -Wert gebildet. Diese schneidet die Oberfläche der Dichtefunktion in einer Ellipse Q (siehe Abbildung 2.9)[99]. Alle Werte außerhalb dieser Ellipse treten mit der Irrtumswahrscheinlichkeit α auf.

Mit Hilfe der Wald-Statistik Q , die annähernd Chi-Quadrat-verteilt ist [98], und Transformation von der drei-parametrischen Regressionsanalyse $(\beta_0, \beta_1, \epsilon)$ zur Verteilungsfunkti-

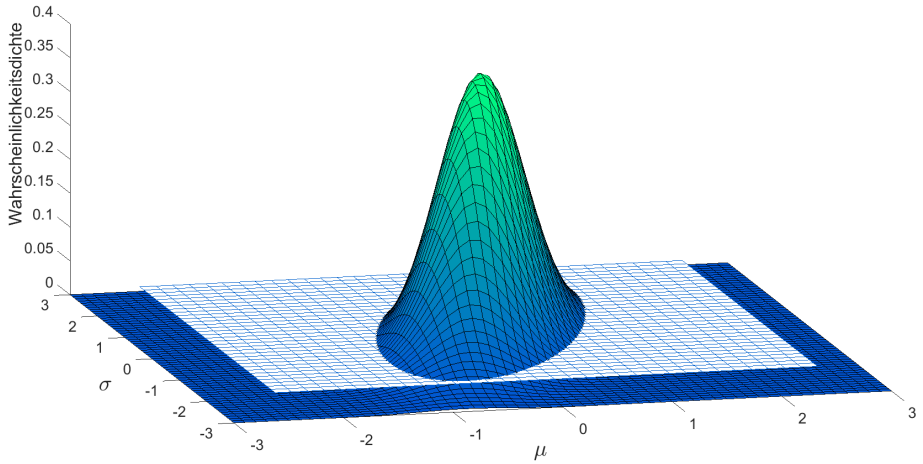


Abbildung 2.9: Bivariate Normalverteilung der Parameter des Verteilungsmodells mit Quantil-Ebene für einen festen Wert p

onsschätzung mit zwei Parametern (μ, σ) ergibt sich für die POD-Kurve folgender unterer Vertrauensbereich [15]:

$$POD_{\alpha}(a) = \Phi(\hat{z} - h). \quad (2.17)$$

Hierbei ist z der Parameter für die Standardnormalverteilung (2.18), der mit Hilfe der MLE in Kapitel 2.2 gewonnenen wurde [15], und h die untere Grenze des Vertrauensbereichs (2.19).

$$\hat{z} = \frac{a - \hat{\mu}}{\hat{\sigma}} \quad (2.18)$$

$$h = \sqrt{\frac{\gamma_{CL}}{n \cdot k_0} \left(1 + \frac{(k_0 \hat{z} + k_1)^2}{k_0 \cdot k_2 - k_1^2}\right)} \quad (2.19)$$

Die Parameter k_0 , k_1 und k_2 werden hierbei durch die Informationsmatrix $I(\mu, \sigma)$ geschätzt [15].

Das vorgestellte Vorgehen zur Berechnung des Vertrauensbereichs hat sich für den Einsatz in der POD als günstig erwiesen [15], [98].

2.3 Physikalische Grundlagen für die Zuverlässigkeitsbewertung von Radiographiesystemen

Im vorigen Kapitel wurden die statistischen Methoden eingeführt, die im Folgenden auf das ZfP-Verfahren überführt werden sollen. Denn es sei angemerkt, dass die Regressionsgerade nur einen statistischen Zusammenhang beschreibt [73]. Der funktionelle Zusammenhang benötigt die Bewertung des vorliegenden System hinsichtlich der physikalischen Grundlagen. Erst physikalische Grundlagen und mathematisches Vorgehen machen die POD zu einem zuverlässigen Werkzeug der ZfP.

Während die Regressionsanalyse die Kausalbeziehung des zu bewertenden Verfahren schätzt, wird der wahre Kausalzusammenhang zwischen den beiden Größen durch die physikalische Beziehung gegeben. Daher sollen für die korrekte Bewertung eines ZfP-Systems die physikalischen Grundlagen und die wichtigen Einflussparameter in diesem Kapitel dargestellt werden.

2.3.1 Radiographie als zerstörungsfreie Materialprüfung

Bei der Bewertung des anzuwendeten ZfP-Verfahrens ist neben den Kosten und dem Einsatzgebiet der zu erwartende Defekt von besonderem Interesse. Die physikalischen Grundlagen des Messprinzips sind ausschlaggebend dafür, welcher Materialdefekt detektiert werden kann [57]. Soll der zuverlässige Einsatz des geprüften Objekts gewährleistet werden, werden oft mehrere Verfahren parallel eingesetzt, um alle kritischen Defekte abzudecken (beispielsweise in sicherheitsrelevanten Bereichen)[102]. Der Einsatz von mehreren Verfahren ergibt insofern Sinn, da die einzelnen Verfahren unterschiedliche Materialdefekten besser detektieren können. Beispielsweise ist die konventionelle RT weniger gut für planare Defekte (z.B. Dopp-lungen in Blechen) geeignet, im Gegensatz zur konventionellen Ultraschallprüfung [54]. Für Porositäten und Schlackeeinschlüsse ist RT jedoch zuverlässiger als andere Verfahren [103]. Aus diesem Grund zählt die RT zu den bedeutendsten Verfahren der ZfP. Der häufigste Einsatz von RT ist die Durchstrahlung des Prüfobjekts. Dabei durchdringen elektroma-gnetische, kurzwellige und hochenergetische Strahlen das Prüfobjekt und treten mit dem Material in Wechselwirkung [104]. Dadurch kann die Durchstrahlungsprüfung nicht sichtbare Defekt detektieren, ohne das Bauteil langfristig in seiner Eigenschaft zu verändern. Der Aufbau der Durchstrahlungsprüfung ist gekennzeichnet durch die Strahlenquelle auf einer Seite des Prüfobjekts und dem Detektor auf der gegenüberliegenden Seite. Dabei geschieht die Auswertung der nicht sichtbaren elektromagnetischen Strahlung, meist mit einer pho-tographischen Schicht (Film) oder durch Anregungsdetektoren (digitaler Matrixdetektor - engl. digital detector array (DDA), Speicherfolien) [41]. In der vorliegenden Arbeit wird er DDA im Fokus stehen (siehe Schema in Abbildung 2.10).

2.3 Physikalische Grundlagen für die Zuverlässigkeitsbewertung von Radiographiesystemen

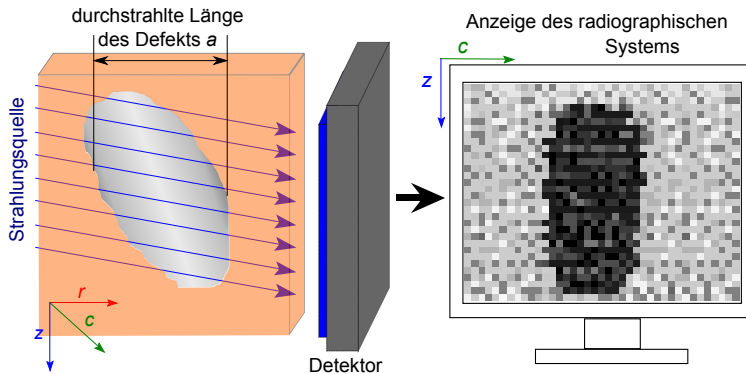


Abbildung 2.10: Schematischer Aufbau einer digitalen RT-Prüfung mit einer DDA: Von der Quelle bis zum Detektor dominiert hauptsächlich der physikalische Teil (innewohnende Faktor). Die Signalverarbeitung und die Bewertung am Monitor spielen ebenfalls eine wichtige Rolle bei der Gesamtbetrachtung des Systems.

2.3.2 Radiographie als Prozess mit zufälligen Einflüssen

Der Prozess der RT gliedert sich in drei Teilbereiche: die Signalentstehung, die Signalverarbeitung und die Auswertung. Auf jeden dieser Bereiche wirken verschiedene Einflüsse, die mit Hilfe des Datenerfassungsmodells (Siehe Abschnitt 2.2) bewertet werden können: Als Signal wird in der RT die auf den Detektor erfasste Strahlung genutzt. Materialdefekte führen im Gegensatz zu ihrem umgebenden Grundmaterial zu Intensitätsunterschieden, die verarbeitet und dargestellt werden und aus denen in der Datenauswertung auf Art, Form und Geometrie von Materialdefekten geschlossen werden kann.

Für die Signalentstehung spielen deterministische und stochastische Einflüsse eine Rolle. Während deterministische Einflüsse (Fokusgröße der Quelle, Prüfaufbau, usw.) während der Prüfung unverändert bleiben, können variable Einflüsse nur mit Hilfe der Wahrscheinlichkeitsbetrachtung beschrieben werden. Einflüsse sind beispielsweise stochastischer Natur und spielen für die Bewertung eines RT-Systems eine wichtige Rolle: z.B. die Detektorrespons, das Streuverhalten des Materials oder elektronisches Rauschen. Weitere unbekannte Einflüsse sind beispielsweise zeitliche Einflüsse oder Schwankungen im Energiespektrum.

Im folgenden Schritt der Daten- bzw. Signalverarbeitung finden sich ebenfalls Einflüsse, die direkt auf das Prüfergebnis wirken (beispielsweise: Kalibrierung des DDA, Darstellung und Auflösung des Ausgabegeräts, Software) [88].

Als letzten Schritt ist die Signalauswertung zu nennen. Es sei angemerkt, dass der Prüfer, der die Signale während der Prüfung beobachtet, und derjenige, der die Entscheidung zwischen Anzeige und Nicht-Anzeige trifft, oft in einer Person vereint sind [35]. Das bedeutet,

2.3 Physikalische Grundlagen für die Zuverlässigkeitsbewertung von Radiographiesystemen

dass für die Bewertung des Auswerteprozesses häufig diese Person im Mittelpunkt steht. Dabei erstreckt sich ein weites Feld von variablen Einflüssen auf die Prüfung. Diese umfassen die Faktoren in der Ausbildung [39], sowie Einflüsse auf den Menschen im ZIP-System (Human Factors) [38], und organisatorische Faktoren [39], [38]. In dieser Arbeit steht die Signalentstehung im Fokus.

2.3.3 Physikalische Grundlagen der digitalen radiographischen Prüfung mit Hochenergie-Röntgenstrahlung

Die Wechselwirkung der hochenergetischen Röntgenstrahlung mit dem Prüfmaterial bzw. dem Materialdefekt ist die Grundlage für die RT. Um einen Zusammenhang zwischen Signal und Defekt zu nutzen, ist es notwendig die Wechselwirkung zu kennen und einen mathematischen Zusammenhang für die Bewertung zu bestimmen.

Röntgenstrahlung und deren Schwächung im Material

Die physikalischen Wechselwirkungen entstehen durch die ausgesandte Strahlung, die teilweise absorbiert wird, im Material gestreut wird oder ungehindert das Material durchdringt. Die Menge der absorbierten und gestreuten Strahlung ist abhängig von der Beschaffenheit des durchstrahlten Materials (Kernladungszahl und Dichte), Dicke des durchstrahlten Prüfstücks und der Energie der Strahlung (weich, harte Strahlung) [46]. Für die Betrachtung der Abnahme der Intensität proportional zur Schichtdicke wird der Schwächungskoeffizient eingeführt.

Der Schwächungskoeffizient μ_{RT} berechnet sich wie folgt [104]:

$$\mu_{RT} = \tau + \sigma_{RT} + \pi, \quad (2.20)$$

wobei τ der Koeffizient für den photoelektrischen Effekt, σ_{RT} der Anteil der Streuung (Compton- und Rayleigh-Streuung) und π den Koeffizient für die Paarbildung darstellt.¹¹

2.3.4 Physikalischer Zusammenhang zwischen Signal und Defektgröße

Wird ein Werkstück mit einem Materialdefekt betrachtet, so kann das Werkstück gedanklich in einen defektfreien und einem defektbehafteten Abschnitt unterteilt werden (Siehe Abb. 2.11). Bei der Durchstrahlung trifft hinter dem defektfreien Teil eine Strahlungsintensität I_{P1} , bei dem defektbehafteten Teil eine Intensität I_{P2} auf den Detektor. Die Intensität, die ein Detektor für einen Bildpunkt (Pixel) aufnimmt und dem Beobachter im digitalen Bild

¹¹Im letzten Kapitel wurde μ und σ als Mittelwert und Standardabweichung eingeführt. Die Doppelnennung geht auf die allgemeine Nutzung der unterschiedlichen Literatur zurück.

2.3 Physikalische Grundlagen für die Zuverlässigkeitsbewertung von Radiographiesystemen

zur Verfügung stellt, wird Grauwert genannt. Der Grauwert kann als ein numerischer Wert im digitalen Bild erfasst werden [41].

Für den einfachen Fall der Prüfung mit monochromatischer Strahlung mit einer starken Bündelung kann die Schwächung mit Hilfe des Schwächungsgesetzes berechnet werden [105]. Dabei kann die Menge der durchdringenden Strahlung I_{P1} und I_{P2} hinsichtlich der Quellintensität I_0 , dem Material und der Wanddicke d und dem Defekt und dessen durchstrahlte Länge Δd beschrieben werden (Siehe Abb. 2.11).

So zum Beispiel für die defektfreien Intensität I_{P1} (vgl. Lambert-Beersches Gesetz):

$$\frac{I_{P1}}{I_0} = e^{-\mu_{RT}d_2}, \quad (2.21)$$

wobei I_0 die nicht geschwächte Strahlung, μ_{RT} der Schwächungskoeffizient und d_2 die Dicke des Materials, das durchstrahlt wurde [104]. Die Intensität I_{P2} kann bestimmt werden, in dem ein defektbehafteter Bereich in die Betrachtung mit einbezogen wird.¹²

Um einen Materialdefekt zu detektieren, spielt der Unterschied der Intensität zwischen defektfreiem und defektbehafteten Material (I_{P1} und I_{P2}) eine Rolle (siehe Abb. 2.11). Die Differenz zwischen den beiden Intensitäten wird auf dem Detektor als absoluter Kontrast C erfasst [106]. C wird praktisch zwischen dem Grundmaterial und der Diskontinuität oder dem Bildgüteprüfkörper in Grauwerten gemessen [41].

Für den Fall, dass $\mu_{RT} \cdot \Delta d$ viel kleiner ist als eins (z.B. bei kleiner Defektgröße bzgl. der Bauteildicke) kann ein linearer Zusammenhang angenommen werden.¹³

$$C := \Delta I = I_{P1} - I_{P2} \propto \mu_{RT} \cdot \Delta d. \quad (2.22)$$

Dabei ist der absolute Kontrast C proportional zur Wanddickenänderung durch einen Defekt Δd . Der Zusammenhang zwischen Defektparameter Δd und Kontrast C kann daher als annähernd linear vorausgesetzt werden, wodurch eine wichtige Voraussetzung für die POD gegeben ist.

Im Folgenden soll noch kurz auf die Intensität der kontrast-mindernden Streustrahlung I_S eingegangen werden (siehe Abb. 2.11)¹⁴. Eine einfache Möglichkeit die im Bauteil entstehende Streustrahlung zu betrachten, ist die Einführung des Streuverhältnisses k (Verhältnis von Streuintensität und Primärintensität $\frac{I_s}{I_p}$).

¹²Im Weiteren wird der Schwächungskoeffizient des defektbehafteten Bereich gleich dem Koeffizienten von Luft gesetzt, was bei einigen Materialdefekten angenommen werden kann (z.B. Gaseinschlüsse, Poren, Porositäten).

¹³Mit der ersten Näherung einer Taylorreihe (dem Abbruch nach dem ersten nicht konstanten Glied) wird der Zusammenhang angenähert [107].

¹⁴Es wird von einem konstanten Strahlungsbeitrag ausgegangen, der den relativen Kontrast mindert.

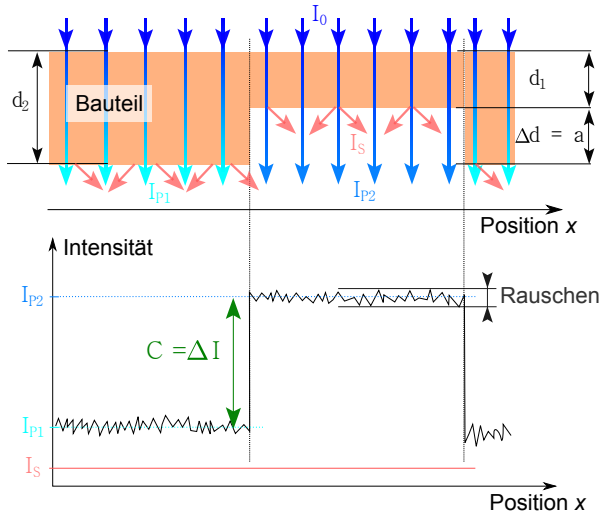


Abbildung 2.11: Beschreibung des Kontrast: Betrag der Primärstrahlung I_P und der Streuung I_S zum Kontrast C bei der Durchstrahlung eines Bauteils

Der relative Kontrast, der in Gleichung 2.23 definiert ist, geht hierbei auf die Streustrahlung ein

$$C_r = \frac{\Delta I}{I_0} = \frac{\mu}{1+k} \Delta d. \quad (2.23)$$

Da die Streustrahlung auch in Prüfrichtung erzeugt werden kann, wird hier der effektive Schwächungskoeffizient angegeben. Im weiteren Verlauf der Arbeit spielt die Streustrahlung wegen des Prüfaufbaus eine untergeordnete Rolle (siehe Kapitel 4.1).

2.3.5 Zusammenhang zwischen durchstrahlte Länge und Kontrast für ein Strahlenspektrum

Wird von Strahlung mit nicht monochromatischem Spektrum und einer realistischen Strahlbreite ausgegangen, gilt:

$$I(E) = I_0(E) \cdot e^{-\int \mu_{RT}(E) x dx}, \quad (2.24)$$

wobei E das Energiespektrum und x die durchstrahlte Länge darstellt. Das Intensitäts- und Schwächungsverhalten im Material können somit nicht mehr wie oben berechnet werden. In diesem Fall helfen Modellierungsprogramme der Radiographie (z.B. aRTist [108]), um die Beziehung zwischen Defektparameter und RT-Signal zu untersuchen. Auch mit Hilfe von einfach gehaltenen Modellierungsergebnissen konnte die annähernde Linearität (Linearität

der Kontrastabhängigkeit von der durchstrahlten Länge) erfasst werden (siehe Anhang 5). Die Grundlage für den Einsatz der Regressionsanalyse bei der Bewertung mit Hilfe der Signal-Antwort-POD ist daher als gegeben anzunehmen.

2.3.6 Beschreibung der variablen Einflüsse auf das Radiographiesystem

Für die POD-Bewertung wird des Weiteren ein adäquates Verteilungsmodell der Streuung um die Regressionsgerade und die Gleichheit der Varianz hinsichtlich der Defektgröße gefordert [36]. Auch spielen Schwellen für die Auswertung (Rausch-, Sättigungs- und Entscheidungsschwelle) während der ZfP-Prüfung eine wichtige Rolle.

Bewertung des Rauschens in der Radiographie

Wie schon erwähnt, kann bei einem niedrigen Kontrast-Rauschverhältnis ein Defekt nicht gefunden werden. Das Rauschen im Bild spielt daher eine wesentliche Rolle bei der Wahrnehmung einer Anzeige und gibt eine Untergrenze für die Prüfbarkeit vor (siehe Abbildung 2.11). In der normativen Literatur wird das Rauschen auch als „signifikantes Hindernis bei der Erfüllung der Anforderungen an die Bildgüte“ bezeichnet [41].

Wird die Bildgüte oder die Qualität der RT betrachtet, so hängt diese von vielen Faktoren ab:

1. Prüfobjekt (Material und Dicke)
2. Strahlenquelle (Qualität der Strahlung, Fokusgröße)
3. Prüfgeometrie (geometrische Unschärfe, Quelle-Objekt-Abstand, Objekt-Detektor-Abstand)
4. Abbildungssystem (Detektoreigenschaften, innere Unschärfe, Rauschverteilung des Detektors)

Dass das genaue Signal nicht vorhergesagt werden kann, liegt beispielsweise an Materialstrukturen (z.B. Inhomogenitäten) oder Oberflächenrauheiten, die dazu führen, dass die Verteilung der Streustrahlung nicht mehr als konstant angenommen werden kann. Gleichzeitig kann es dazu kommen, dass hierdurch die Bildqualität abnimmt [46].

Die Art der verwendeten Strahlung und die Prüfparameter spielen für die Prüfung eine wichtige Rolle. So beeinflussen Bestrahlungskonditionen die gemessenen Signale, da sie von der Dosis abhängen (siehe hierzu Photonenrauschen [109]).¹⁵ Mit einer höheren Dosis kann das Kontrast-Rausch-Verhältnis - engl. contrast noise ratio (CNR) gesteigert werden, was somit zu einer verbesserten Bildqualität führt.

¹⁵ Annahme: Benutzung eines dosisproportionalen Detektors

Bei dem Versuchsaufbau müssen hierbei beispielsweise die Seitenstreuung oder die Hintergrundstrahlung betrachtet werden, die das CNR senken können [41].

Im Fall von DDA bezeichnet das Strukturrauschen die Unterschiede zwischen den einzelnen Antworten der Detektor-Elemente. Auch thermische Effekte spielen hierbei eine große Rolle. Detektorpixel mit geringerer Sensitivität oder ohne Sensitivität (Bad Pixel) können außerdem die Rauschverteilung beeinflussen, falls diese nicht durch die Kalibrierung ausgeglichen werden. Eine korrekte Kalibrierung kann sowohl das Strukturrauschen deutlich senken [105] als auch Bad Pixeln entgegen wirken [110]. Der Prozess der Kalibrierung des DDA führt jedoch zusätzlich zu einer Veränderung der Rauschverteilung, wobei oftmals der genaue Prozess nicht bekannt ist.

Des Weiteren haben die Pixelgröße [111] und die Grauwerte einen weiteren Einfluss auf die Kontrast- bzw. räumliche Auflösung und somit auf das Prüfergebnis. Außerdem kann elektronisches Rauschen, beispielsweise in Steckverbindungen, Kabeln und Schaltungen, als Rauscheinfluss genannt werden.

Ist die durchstrahlte Länge Δd eines Materialdefekts klein, so ist der Kontrast C gering und geht im Rauschen unter. Der Defekt kann nicht detektiert werden. Diese Kontrastuntergrenze wird durch die Kontrastempfindlichkeit ausgedrückt und mit Hilfe der Draht-Typ-Bildgüteprüfkörper und Stufe-Lochprüfkörper gemessen [112]. Die Kontrastempfindlichkeit ist eine wesentliche Eigenschaft für die Effektivität des RT-Systems [57].

Es kann im Allgemeinen bei der mathematischen Beschreibung des Rauschens annähern von einer Normalverteilung ausgegangen werden. Dies sollte jedoch für jede Bewertung untersucht werden [113]. Dabei wird das Rauschen als Abweichung der Intensität über dem Hintergrund definiert. Mit der Annahme, dass sich die Streuung der Werte über dem gemessenen Defektbereich homogen gestaltet und die durchstrahlte Defektgröße gering ist, kann eine Signalstreuung mit der Standardabweichung des Rauschens angenähert werden. Dies ergibt Sinn, da viele Einflüsse des Rauschens ebenfalls das Defektsignal beeinflussen.

Unschärfe als wichtiger Einfluss

Ein weiterer Faktor bei der Bewertung eines RT-Systems ist ihre Unschärfe, die im Einzelnen folgende Bestandteile enthält:

- Bewegungsunschärfe
- Geometrische Unschärfe
- Innere Unschärfe
- Unschärfe, die am Bildschirm erscheint

Im Allgemeinen stehen die geometrische und innere Unschärfe im Fokus der Bewertung: Die Gesamtunschärfe u_t hängt einerseits von der Strahlenintensität u_t und dem Detektor (innere

2.3 Physikalische Grundlagen für die Zuverlässigkeitsbewertung von Radiographiesystemen

Unschärfe u_i) [106], andererseits von der Prüfgeometrie und dem Brennfleck der Quelle (geometrische Unschärfe u_g) [46] ab. Die Gesamtunschärfe u_t kann beispielsweise abgeschätzt werden durch:

$$u_t = \sqrt{u_i^2 + u_g^2}. \quad (2.25)$$

Die Unschärfe macht sich bemerkbar bei scharfen Kanten, die im Bild verwaschen angezeigt werden, und somit Einfluss auf die Detektierbarkeit eines Defekts haben. Bei kleinen Anzeigen oder großen Unschärfen kann es zu einer Abnahme des Kontrastes kommen [106]. Somit stellt die Unschärfe bei kleinen Anzeigen eine Detektionsuntergrenze dar.¹⁶

Aus Sicht der POD-Bewertung stellt die Unschärfe ein weiteres Hindernis dar, da selbst ein bekannter Defekt nicht direkt einer unscharfen Anzeige zugeordnet werden kann.

Wahrgenommener Kontrast

Die vorigen Abschnitte haben gezeigt, dass der Kontrast an sich nicht ausreicht, um aussagen zu können, ob ein Defekt detektiert wird. Vielmehr gilt für voluminöse Defekte, dass die Detektion vom Kontrast und der Größe seiner Anzeige abhängig ist. Des Weiteren spielen das Rauschverhalten des Detektors bzw. das Streuverhalten des Prüfobjekts und die minimale auflösbare Einheit des Detektors (SR_b) eine Rolle.

Zusammengefasst kann der normierte CNR (CNR_N) definiert werden als:

$$\frac{CNR_N}{\Delta d} = \frac{\mu_{RT}SNR}{SR_b}. \quad (2.26)$$

Bei der Detektion von kleinen Defekten ist der Quotient auf der rechten Seite wichtig [105]: das SNR wird durch die Einstellung und den Prüfaufbau, SR_b hauptsächlich durch die verwendeten Geräte und μ_{RT} durch das durchstrahlte Material und die verwendete Energie beeinflusst.

Zu SNR sollte erwähnt werden, dass dieses stark von der Qualität der Kalibrierung abhängt [41] und im Fall der Sättigung des DDA maximal wird. Bei maximalem SNR handelt es sich für die POD-Bewertungen um rechts-zensierte Datenpunkte. Somit ergibt sich bei dem maximalen SNR eine Sättigungsschwelle.

2.3.7 Bewertung der Auffindwahrscheinlichkeit in der Radiographie

In vielen Arbeiten hinsichtlich der Auffindwahrscheinlichkeit wurden hauptsächlich künstliche Defekte genutzt [36], wofür der lineare Zusammenhang zwischen Defektparameter und

¹⁶Aus der Bewertung mit Hilfe eines Doppeldrahtbildgüteprüfkörpers kann die Basisortsauflösung (SR_b) erfasst werden, die der effektiven Pixelgröße entspricht [41] (siehe auch [114], [115] oder [116]).

2.3 Physikalische Grundlagen für die Zuverlässigkeitsbewertung von Radiographiesystemen

Signalhöhe (aus Abschnitt 2.2) häufig ausreicht. In diesem Fall muss eine Korrelation zwischen realen und künstlichen Defektcharakteristika für den Einsatz der POD entweder theoretisch oder empirisch hergestellt werden [36]. Die Bildgüteprüfkörper sind ein wesentliches Hilfsmittel zur Beschreibung der Detektierbarkeit von bestimmten Materialdefekten, da eine empirische Relation besteht [104]: Bei dem Einsatz der RT für die Detektion von voluminösen Defekten (z.B. Poren und Löcher) konnte eine gute Übereinstimmung mit Ergebnissen von Stufe-Loch-Bildgüteprüfkörpern nachgewiesen werden [104]. Die zusätzliche Einführung von Unschärfe und die Fähigkeit des menschlichen Auges unterschiedliche Größen und Formen wahrzunehmen, muss daher das reine kontrastbasierende Detektionsmodell erweitern (siehe Detektierbarkeit von Stufe-Loch Bildprüfkörpern [104]).¹⁷

Sollen nun in die Bewertung zudem reale Defektdateien verwendet werden, so muss das Vorgehen zur Berechnung der POD und seine Voraussetzungen überprüft werden. Wird nur die maximale Materialdickenänderung Δd für reale und künstliche Defekte für die Bewertung genutzt, so ergibt sich hierbei kein absoluter oder relativer Kontrastunterschied zwischen realem und künstlichen Defekten. Sollen die oben genannten Faktoren zur Detektierbarkeit auch betrachtet werden, die durch ihren großen Einfluss in die Bewertung einfließen sollten [57], so muss das Modell erweitert werden. Durch ein erweitertes Modell können jedoch Unterschiede zwischen künstlichen Referenzdefekten und realen Materialdefekten bei der Nutzung in der POD-Analyse entstehen, die bisher noch nicht ausreichend behandelt wurden.

¹⁷Bei Defekten, für die die Orientierung oder weitere Parameter wichtig sind (z.B. reale Defekte) [117], spielt selbst der Winkel eine Rolle (siehe Risse [104]).

2.4 Herausforderung bei dem Einsatz der POD in der Radiographie mit realen Defekten

Bei dem Einsatz der POD in der RT treten einige Herausforderungen auf, falls reale Defekte verwendet werden. In diesem Abschnitt werden die Herausforderungen beschrieben und mögliche Lösungswege, die von uns gefunden wurden, vorgeschlagen. Für die Zuverlässigkeitsbewertung eines ZfP-Systems in der Praxis stehen vor allem die relevanten Materialdefekte im Mittelpunkt [17], [118]. Dabei sollte eine ausreichende Datenmenge, bestehend aus realen und repräsentativen Defekten, zur Verfügung stehen. Liegt eine geringere Menge an Defekten vor, kann dies zu unplausiblen und falschen Ergebnissen führen [77]. Es werden mindestens 40 unabhängigen Messungen für die Bewertung mit der POD gefordert [15]. Jedoch ist die Herstellung und Charakterisierung von repräsentativen Defekten schwierig [119] und die große Anzahl von Daten nur selten zu erreichen. Denn bei der POD-Bewertung müssen die repräsentative Aussagekraft und die Kosten im Blick behalten werden. Dadurch stellt die POD-Bewertung mit realen Defekten eine Herausforderung dar, die ein besonders Augenmerk auf die Eigenschaften der zu untersuchenden Defekte richten muss. Ein wichtiger Schritt ist die Unterscheidung der verschiedenen Materialdefekte.

2.4.1 Arten zur Verfügung stehender Materialdefekte

Ein ZfP-System findet seine Anwendung in der Herstellungsprüfung oder in der wiederkehrenden Prüfung. Es sollen dabei Materialdefekte detektiert werden, die das Funktionieren eines Bauteils oder Systems gefährden. Im Folgenden soll ein ZfP-System für die Herstellungsprüfung vor dem Einsatz bewertet werden. Dabei werden Referenzdefekte für die Kalibrierung genutzt, da für ein neues Produkt meist nur wenige bekannte reale Defekte vorliegen. Es wird allgemein zwischen realen, realistischen und künstlichen Materialdefekten unterschieden. Die richtige Wahl und die ausreichende Anzahl von Defekten ist ein wichtiger Punkt, der bei der Planung der notwendigen Experimente für die Bewertung des ZfP-Systems beachtet werden sollte [120].

Reale Materialdefekt

Reale Materialdefekte entstehen unbeabsichtigt bei der Herstellung und während des Einsatzes [17]. Jedoch sind die realen Defekte einzigartig hinsichtlich Bauteildesign, Herstellungsprozess, Material und Einsatzbedingung [121]. Die wenigen Möglichkeiten zum Untersuchen und Charakterisieren von realen Defekten sind meist kostenintensiv [121]. Für die Bewertung mit Hilfe der Signal-Antwort-POD wird teilweise sogar explizit die metallographische Untersuchung gefordert [88]. Die Metallographie ist ein aufwendiges Verfahren [14], das nur in wenigen Fällen für die Charakterisierung von Materialdefekten genutzt wird. Die hohen

2.4 Herausforderung bei dem Einsatz der POD in der Radiographie mit realen Defekten

Kosten und die nicht eindeutige Beziehung zwischen Struktur des Defekts und physikalischen Grundlagen des Verfahrens stellen hierbei die wesentlichen Kritikpunkte dar [121].

Realistische Materialdefekte

Um die Anzahl der Materialdefekte mit realen Eigenschaften zu erhöhen, wird auf realistische Defekte ausgewichen. Realistische Defekte werden in Prüfblöcken eingebracht, um metallurgische Erscheinungen von realen Defekten nachzuahmen [17].

Die Herstellung realistischer Defekte findet entweder durch direkte Implementierung oder durch Modifizierung der Herstellungsparameter des Bauteils (beispielsweise Prozessparameter bei der Herstellung einer Schweißnaht) statt [122].

Für die realistischen Defekte werden Bedingungen hinsichtlich des Detektionsverhaltens aufgestellt, wobei Vorwissen über die Belastungsart (z.B. Korrosion, Ermüdung) vorausgesetzt werden [103]. Durch ihre Relevanz in der Qualifikation von ZfP-Verfahren sind realistische Defekte von besonderer Bedeutung [17]. Jedoch können sie häufig nur in sehr kleinen Mengen in Laboren herbeigeführt werden [103]. Außerdem sollten selbst realistische Defekte letztendlich metallographisch charakterisiert werden, wodurch weitere Kosten anfallen [123].

Künstliche Materialdefekte

Der Einsatz von künstlichen Defekten bietet eine einfache und reproduzierbare Möglichkeit Materialdefekte zu erhalten [17]. So können künstliche Defekte für das konzeptionelle Detektionsverhalten herangezogen werden [34]. Beispiele für künstliche Materialdefekte sind Defekte, die durch elektroerosivem Abtragsverfahren - engl. electrical discharge machining (EDM) hergestellt werden, Bohrungen (z.B. Flachbodensacklochbohrung - engl. flat-bottom-hole (FBH)) oder Nuten [17].

Die Vorteile der künstlichen Defekte sind [121]:

- leichte und kostengünstige Herstellung
- einfache Charakterisierung durch bekannte Herstellungsparameter
- Nutzbarkeit für die Versuchsplanung durch einfache Parametrisierung
- Unterstützung bei dem Verständnis des Prüfers
- Geringe Verunreinigungen und Irregularitäten des Defekts
- Fokussierung auf wichtige Prüfmerkmale

Allerdings ist dieser Einsatz bei der Bewertung der Fähigkeit stark diskutiert [124]: Künstliche Defekte haben einfache geometrische Formen und spiegeln dadurch nur begrenzt die

2.4 Herausforderung bei dem Einsatz der POD in der Radiographie mit realen Defekten

Realität wider [103]. Dadurch können sich reale und künstliche Defekte in ihrem Signalverhalten stark unterscheiden [123] und die vereinfachten künstlichen Defekte wiederum zu einem zu optimistischen Ergebnis hinsichtlich der Detektionsfähigkeit führen [103].

Bewertung des ZfP-Systems mit den verschiedenen Defekten

Als Kompromiss zwischen Kosten und Repräsentativität werden für die Bewertung des ZfP-Systems Referenzdefekte gefordert, die reale Defekt abbilden sollen [81]. Häufig wird argumentiert, dass, falls ein kleiner künstlicher Defekt detektiert wird, ein deutlich größerer realer Defekt gefunden werden kann [103]. Das gilt jedoch nur bei einer guten Korrelation zwischen künstlichen, realistischen und realen Defekten, die für den Einsatz von künstlichen Defekten nachgewiesen werden muss [121]. In wie weit reale Materialdefekte mit künstlichen Defekten vergleichbar sind, ist jedoch in einigen Fällen stark umstritten [125], [126], [127], [128]. Gründe für die variierende Signalbildung sind die unterschiedliche geometrische Defektstruktur oder auch andere Signalentstehungsprozesse (z.B. Kantenreflexionsverhalten bei Ultraschallprüfung - engl. ultrasonic testing (UT)) [129], [130].

Das Signalverhalten von realen Defekten ist maßgeblich von Parametern beeinflusst, die möglicherweise nicht zu den geprüften Defektparametern gehören (z.B. Größe, Orientierung und Natur des Defekts) [124].

Aus diesen Gründen wird klar, dass eine Bewertung des ZfP-Systems nur auf der Basis von künstlichen Defekten, zu einer falschen Entscheidung führen kann [40]. Es sollen daher, falls möglich, reale Defekte miteinbezogen werden, um keine Fehleinschätzung des Systems zu begehen [121]. Jedoch ist die POD-Bewertung eines RT-Systems auf realen Defekten nur an sehr wenigen Stellen in der Literatur zu finden (Beispiele: [36], [70]) Im Weiteren kann somit nur begrenzt auf vorhandene Literatur zurückgegriffen werden.

2.4.2 Umgang mit realen Daten bei der Zuverlässigkeitsbewertung von ZfP-Systemen

Es ist grundsätzlich nicht möglich, die gesamte Population an möglichen Defekten auszuwerten. Einerseits ist die Population teilweise nicht bekannt, andererseits wäre die vollständige Charakterisierung zu kostenintensiv. Daher soll die Population auf eine praktische und konzeptionelle Weise durch eine beschränkte Zahl von Strichproben beschrieben werden [35]. Hierbei ist die klare Festlegung und Definition der wichtigen Defektparameter in der Versuchsplanung wichtig, um eine korrekte Bewertung durchzuführen [131].

Im Besonderen ist die Anzahl der Stichproben essentiell, um das Ergebnis auf die Population zurückführen zu können. Allgemein gilt, dass die Anzahl der Daten genügend groß sein muss, damit eine statistische Auswertung angewandt werden kann und die Aussage des Ergebnisses zuverlässig wird [35]. Jedoch bleibt letztendlich die Versuchsplanung dem individuellen Untersucher und dessen wissenschaftlichen Urteilsvermögens überlassen [35].

Erweiterte Detektionstheorie für die POD-Bewertung

Es kann gesagt werden, dass die Bewertung des RT-Systems mit Hilfe der POD sich auf den Kontrast stützen kann, da der physikalische Zusammenhang gegeben ist. Jedoch sind Lage, geometrische Form des Defekts und die gegebene Anzeigengröße und ihre Anzeigenform zusätzliche wichtige Punkte bei der Bestimmung der realen Detektionsfähigkeit. Hierfür sind erweiterte Konzepte für die Berechnung der POD notwendig. Beispiele sind die Einführung mehrerer Entscheidungsschwellen [132], die „Observer-POD“ [133] und die „Datenfeld-POD“ [18]. Während die verschiedenen Schwellen und die „Observer-POD“ eine große Zahl an Daten benötigt, um den Schwellwert für das System festzulegen, ist ein Kritikpunkt der „Datenfeld-POD“, dass es nicht eindeutig möglich ist, die reale Defektstruktur der Anzeige gegenüber zu stellen.

Die Forderung nach einem korrekten Vorgehen zur Berechnung der POD ist einerseits die Einbindung der Anzeigengröße (und somit Unschärfe und die Detektierbarkeit des Auges) und andererseits die möglichst realistische Gegenüberstellung von realen Defektparametern und der Anzeige.

Umgang mit wenigen realen Defekten bei der Bewertung

Steht ein Produkt vor der Einführung und es wird ein ZfP-System bewertet, so liegen häufig nur wenige relevante reale Defektdateien vor. Deswegen werden realistische Defekte oder künstliche Referenzdefekte zur Bewertung herangezogen.

Da für kleine Stichproben statistische Ansätze fehlschlagen können, kann der Einsatz von Verteilungsmodellen für die Signalstreuung zu einem falschen Ergebnis führen (vgl. [85], [45], [85], [53]).

Die Herausforderung liegt in der Suffizienz der Daten. Damit ist die Stichprobengröße beschrieben, bei der der Informationsverlust zu der Population gering ist und somit die Varianz nahe an dem minimal möglichen liegt [85], [82], [96].

Um die Aussage der Bewertung auf Basis von wenigen Daten zu unterstützen und zuverlässiger zu machen, wird die Erhöhung von Daten [96] oder der Einsatz von zusätzlichen Informationen vorgeschlagen [59]. Die Erhöhung der experimentellen Daten ist häufig aus ökonomischer Sicht keine Option.

Auch die wiederholte Prüfung von vorliegenden realen Defekten stellt kein Mittel dar, die mangelnde Anzahl auszubalancieren [88]. Ansätze für Bootstrap-Methoden oder Monte-Carlo-Analysen können in einigen Fällen von Nutzen sein [45], [96], [134]. Jedoch kann es auch hier zu einer Fehleinschätzung kommen, wenn die geringe Datenmenge Aspekte außer Acht lässt [96].

Eine weitere weit verbreitete Variante, Informationen über ein ZfP-Verfahren zu erlangen, sind Modellierungs- und Analyseprogramme [135]. In der POD unterstützt durch Modellierungssoftware - engl. model-assisted POD (MaPOD) wird das physikalische Prinzip mit

2.4 Herausforderung bei dem Einsatz der POD in der Radiographie mit realen Defekten

einem Modell erfasst, während die weiteren Einflüsse (probabilistisch) zusätzlich in der Bewertung generiert werden.¹⁸ Für den Einsatz der MaPOD müssen die Modellierungsergebnisse jedoch mit Experimenten überprüft werden [136]. Falsche Modellparameter oder unüberprüfbare Annahmen können daher zu einem unzulässigen Ergebnis der Bewertung führen [137]. Durch das subjektive Beeinflussen der auswertenden Person bezüglich der Annahmen kann es zu einer Voreingenommenheit hinsichtlich des Ergebnisses kommen [138]. Obwohl die Verifizierung von Modellierungsprogrammen weitreichend ist, sind auch unpräzise Modellierungsergebnisse nicht auszuschließen und können daher ebenfalls das Vorgehen gefährden [139]. Außerdem ist die Analogie zum Vertrauensbereich als Bewertungsmaß für modellierte Daten zu hinterfragen [140].

Ein weiterer Ansatz beschreibt die Nutzung von nicht-experimentellen Daten. Viele Information über das ZfP-System, die wichtig für die Bewertung des Systems sind [35], sind Meinungen und Erkenntnisse des Expertengremiums (siehe [67]). Diese liegen jedoch meist nicht quantifizierbar gemessen vor [141], [39]. Ein wichtiges Beispiel ist die subjektiv-technische Begründung der Äquivalenz zwischen einem hergestellten künstlichen Defekt und dem passenden realen Defekt (z.B. dem Urteil des Prüf- bzw. Qualitätsingenieurs und deren Erfahrungen mit Defektwachstum oder strukturellen Integritätsannahmen) [88]. Der subjektive Blickwinkel ist durchaus auch in der Praxis von Vorteil, da ein erfahrener Prüfer eine Anzeige an einer defektanfälligen Stelle intensiver in Augenschein nimmt [46].

Die Beachtung der subjektiven Einflussgrößen ist daher ein wichtiger Bestandteil der Bewertung [131], [79], [31]. Daher ist es sinnvoll, die subjektiven Meinungen und nicht-experimentelle Kenntnisse zu quantifizieren und diese für die Bewertung mit der POD zu nutzen [95].

Soll jedoch die Nutzung aller vorliegenden Ressourcen ermöglicht werden, so ist die Kombination der vorliegenden Daten ein Erfolg versprechender Ansatz. Obwohl für unterschiedliche Datenquellen eine getrennte Betrachtung bevorzugt werden sollte, kann teilweise die getrennte Datenbewertung nicht durchgeführt werden (z.B. bei zu geringer Datenmenge) [45].

Die Mischung von Daten aus unterschiedlichen Quellen zu einer Datengrundlage ist eine mögliche Methode, Daten erheblich zu erhöhen [35], [45]. Dies ist jedoch nur möglich, wenn die Unterschiede zwischen den beiden Datenmengen gering sind [45]. Bei der Mischung von sehr unterschiedlichen Population kann es jedoch dazukommen, dass eine Unterpoptulation bei der Auswertung überrepräsentiert ist, was zu einem falschen Schluss führen kann [45], [142], [35].

Unterschiedliche Prüfsituation können bei der Beurteilung von ZfP-Systemen mit Hilfe von Modellierungsergebnissen durch eine Transferfunktion miteinander bewertet werden [143]. Jedoch ist eine exakte Kenntnis des Zusammenhangs eine wesentliche Anforderung der

¹⁸Ein europäischer Einsatz der Modellierungsprogramme in der POD ist das Projekt PICASSO

2.4 Herausforderung bei dem Einsatz der POD in der Radiographie mit realen Defekten

Transferfunktion [60]. Liegt für den Transfer eine unbekannte Haupteinflussgröße vor oder es gibt eine wesentliche Veränderung in der übertragenen Situation (Einfluss von Parametern), so kann der Prozess zu Fehleinschätzungen führen. Die Annahmen (z.B. Streuverteilung, Materialhomogenität), die großen Einfluss auf das Ergebnis der Bewertung haben und nicht eindeutig abgesichert werden können, stellen daher ein weiteres wesentliches Hindernis dar.

Erweiterung des konventionellen POD-Ansatzes

Die Erweiterung des konventionellen POD-Ansatzes basiert hierbei im Weiteren auf der Nutzung von sowohl realen als auch künstlichen Defektdaten. Der Schlüssel für den Vergleich und somit auch die Kombination beider Datenpopulationen ist die Gegenüberstellung der RT-Anzeigen der realen Defekte mit den vorhandenen Schlifffdaten und der RT-Anzeige der künstlichen Defekte mit den durchstrahlten Defektanmessungen (siehe im zukünftigen Abschnitt 3.2). Hierbei spielt die oben erwähnte erweiterte Detektionstheorie eine bedeutende Rolle.

Bei der, von uns vorgeschlagenen, Kombination der Daten soll eine bewusste Beeinflussung der Parameterschätzung abhängig von der subjektiven Expertenmeinung (ähnlich der nicht-experimentellen Datenbewertung) ermöglicht werden. Dadurch kann erreicht werden, dass für eine geringe Datengrundlage an realen Defekten, die künstlichen eine größere Rolle spielen, während bei einer ausreichenden Menge an realen Defekten, der Einfluss der künstlichen Defekte auf die Bewertung verschwindet [144].

Neben dem Einsatz der Detektierbarkeit von künstlichen Defekten, als a priori Wissen im Bayes'schen Ansatz bei sehr geringer Datengrundlage [144], soll hier als Alternative auf die gewichtete Schätzung von Verteilungsparametern für die POD-Funktion eingegangen werden (siehe Abschnitt 3.3).

Das generelle Anliegen dieser Arbeit ist der korrekte Einsatz von realen Defekten für die POD-Bewertung, wobei auf die Herausforderungen beim Vergleich mit Referenzverfahren, den verschiedenen Defektparametern und der geringen Datenmenge Rücksicht genommen wird.

3 Neue Methodik zur Bewertung von Radiographie-Systemen mit Hilfe von realen Defekten

Im folgenden Kapitel werden die wesentlichen Neuerungen bei der Bewertung von RT-Systemen durch die POD eingeführt, die von uns vorgeschlagen werden für die Nutzung von realen Defekten. Im ersten Teil wird der Umgang mit realen Daten und die Gegenüberstellung von Schliffen und RT-Anzeige ausgearbeitet. Im nächsten Teil folgt die Einführung einer POD-Methodik, die abhängig von der Anzeigenfläche ist. Außerdem wird eine Möglichkeit zur Bewertung geringer realer Datenmengen vorgestellt.

3.1 Umgang mit Schliffen realer Defekte bei der Bewertung durch die POD

Die Bewertung eines ZfP-Systems mit Hilfe der POD basiert auf der Annahme, dass die Defektparameter bekannt sind. Die Abmessungen von künstlich eingebrachten Defekten werden häufig als gegeben angenommen, wodurch weitere Untersuchungen umgangen werden. Die wahren Abmessungen von realen Defekten sind jedoch schwieriger zu bestimmen. Dafür sind kostenintensive und komplexe Untersuchungen notwendig. Die Metallographie ist eine Möglichkeit, die häufig eingesetzt wird, um reale Defektgrößen zu bestimmen. Die Gegenüberstellung der Daten (Defekte zu ZfP-Signale) muss anschließend in einem gemeinsam gewählten Koordinatensystem geschehen.

In diesem Kapitel werden Ansätze aufgezeigt, die Metallographie mit den RT-Signalen zu vergleichen. Es wird ein Überblick über die übliche Vorgehensweise bei der RT-Bewertung durch Nutzung eines Schliffbildes gegeben und ein räumliches Koordinatensystem aufgestellt. Im Folgenden wird der Ansatz mit den Informationen aus mehreren Schliffbildern für reale Defekte erweitert. Dabei stellt die Abkehr vom Ansatz der höchsten Signal vs. größten Defektparameters in der RT eine wesentliche Erweiterung dar.

3.1.1 Metallographische Untersuchung

Die Metallographie stellt ein genaues Verfahren zur Charakterisierung und Bemaßung von Defekten dar, die im Volumen liegen.

3.1 Umgang mit Schlifflen realer Defekte bei der Bewertung durch die POD

Dabei werden an der Stelle des erwarteten Materialdefekts Schichten von Material abgetragen, bis die Oberfläche aufreißt. Schritt für Schritt können weitere Schichten entfernt werden, um den Defekt zu untersuchen. Außerdem kann hierbei die Oberfläche behandelt (geschliffen, geätzt und poliert) und hochauflösend digitalisiert werden [16]. Eine exakte Bemaßung ist dann für die jeweilige Schicht möglich. Jedoch besteht auch hier die Möglichkeit, dass durch das Schleifen und das Polieren der Defekt während des Metallographieprozesses verkleinert wird oder aber durch das Ätzen vergrößert wird. Die Bemaßung ist daher mit einem Fehler behaftet (vgl. Abschnitt 2.2 Gleichung 2.6) Werden diese Untersuchungen in bestimmten räumlichen Abständen durchgeführt, ist eine ungefähre Abschätzung in der dritten Ebene möglich.

Für die POD werden die Informationen aus der Metallographie und die Anzeige der RT gegenübergestellt. Dazu ist es notwendig, ein gemeinsames Koordinatensystem zu erstellen, wobei auf bestimmte Informationen Rücksicht genommen werden muss. Dabei stehen folgende Punkte im Fokus:

- das geprüfte Objekt (z.B. Geometrie und Bemaßung des Prüfstücks) im Folgenden: Zylindrisches Prüfobjekt [145]
- der genutzte Prüfaufbau für die RT (z.B. Aufnahmeposition, Durchstrahlungsrichtung) im Folgenden: Zeilendetektor [43]
- das metallographische Vorgehen (z.B. untersuchte Position, Abtragerichtung, Bemaßung der digitalisierten Aufnahme) [16]

Aufgrund dieser Überlegungen und für den weiteren Einsatz wird ein zylindrisches Koordinatensystem eingeführt (siehe Abb. 3.1 und Abschnitt 4.2). Die Axialachse z wird in Schweißnahttiefen-Richtung festgelegt. Entlang der Radialachse r kann die Dicke des Prüfstücks abgelesen werden. Im Schweißnahtbereich enthält diese auch die zu prüfende Schweißnaht. Im Weiteren wird die Radialachse r auch als Durchstrahlachse gewählt. Die Umfangsachse c beschreibt in Umlaufrichtung die Koordinate, in der auch die Schweißnaht verläuft. Im Allgemeinen wird im zylindrischen Koordinatensystem hier der Winkel ϕ angegeben, um eine unabhängige Größe zu bilden. Im vorliegenden Fall wird nur der Schweißnahtbereich verwendet, wodurch die Radialachse hinsichtlich r festgelegt ist. Dadurch dass die Schlibfbilder nur in einer rechteckigen Form gemacht werden konnten wird im Folgenden eine Umfangskoordiante in Millimeter statt die Grad angegeben.¹

¹Die Umfangskoordiante kann berechnet werden aus $c = 2 \cdot r \cdot \sin(\frac{1}{2}\phi)$. Außerdem wird der Nullpunkt auf eine gekennzeichnete Stelle am Kanisterrand gelegt.

3.1 Umgang mit Schliffen realer Defekte bei der Bewertung durch die POD

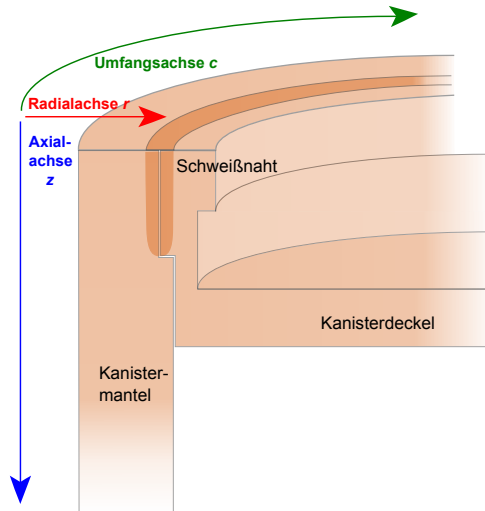


Abbildung 3.1: Zylindrisches Koordinatensystem für eine umlaufende Schweißnaht

3.1.2 Konventionelle POD mit realen Defekten unter Verwendung einzelner Schliffbilder

Der Zusammenhang zwischen der größten durchstrahlten Defektlänge und dem maximalen Kontrast wurde in Kapitel 2.3 beschrieben. Um den Ansatz für die Signal-Antwort-POD zu nutzen, wird das Schliffbild bei der größten durchstrahlten Defektlänge bewertet. Bei räumlicher Lage des Schliffbildes in der r - c Ebene bildet beispielsweise das Schliffbild die Information in Umfangsachse (genähert)² und Radialachse (Durchstrahlungsrichtung) ab (siehe Abb. 3.2).

Das radiographische Bild liefert die Informationen in der Ebene der Umfangs- und Axialachse (z - c) wobei der RT-Strahl in radialer Richtung verläuft und den Verlauf des Schwächungskoeffizienten aufnimmt und integriert. Deshalb kann auf die Tiefe des Defekts in Radialachse nur indirekt aus dem Kontrast geschlossen werden. Außerdem ist durch die Projektion keine Aussage über die genaue Tiefenposition möglich. Da das RT-Bild (Abb. 3.3) auch eine gröbere Auflösung (Bildelemente in einer digitalen Anzeige - engl. PICTURE ELEMENT (Pixel)) als die digitalisierte Metallographie besitzt, müssen die metallographischen Pixel für die Gegenüberstellung zusammengefasst werden.³

²Die Schliffe besitzen eine quadratische Fläche, während die Umfangskoordinate im zylindrischen System rund ist. Durch den großen Durchmesser der Umfangskoordinate und der geringen Breite des Schliffs kann der Winkel vernachlässigt werden.

³Im Fall eines Schliffs bleiben weitere Informationen außerhalb des für den Defekt angesprochenen Pixelbereichs unberücksichtigt. Denn mehrfache Zählung für einen Defekt führt im Allgemeinen zu einer Überbewertung bei der POD [75].

3.1 Umgang mit Schliffen realer Defekte bei der Bewertung durch die POD

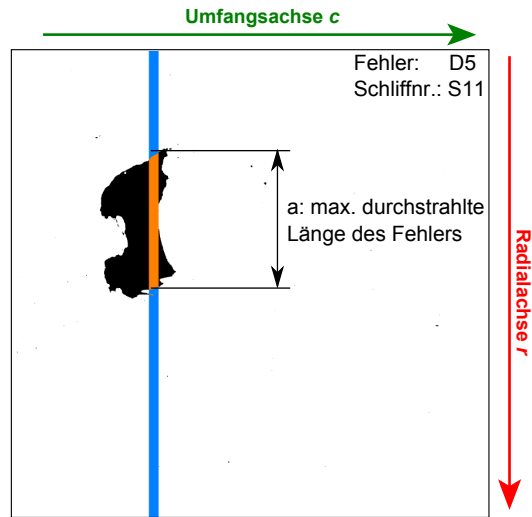


Abbildung 3.2: Die in der einfachen POD genutzte durchstrahlte Länge hinsichtlich eines Detektorpixels im Fall eines bestimmten Defekts (Schliffnummer 11 des Defekts Nr. 5)

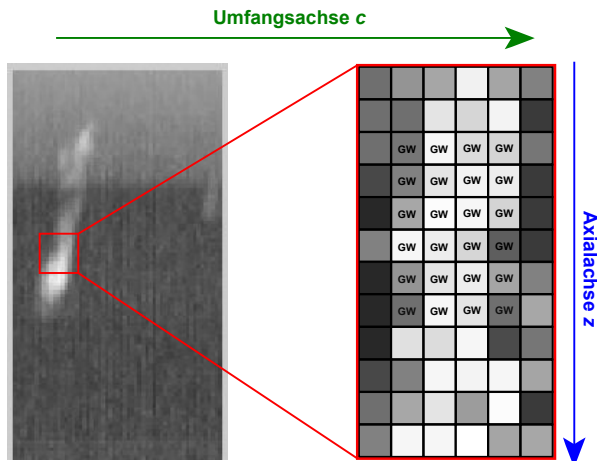


Abbildung 3.3: Aufbau eines radiographisches Bildes bestehend aus Grauwertpixel

3.1 Umgang mit Schlifflen realer Defekte bei der Bewertung durch die POD

Bei der Zuordnung der maximalen Kontraste zu den größten durchstrahlen Defektlängen liegen einige Unsicherheiten vor, die sich in der Streuung der Datenwerte widerspiegeln: Bei der metallographischen Untersuchung kann nicht davon ausgegangen werden, dass der Schliff an der größten Fehlerlänge erfolgt ist. Die Bearbeitungsschritte des Schneidens und Schleifens können die maximale Ausdehnung zerstört haben. Außerdem ist eine direkte Zuordnung nicht möglich, da die wirkliche Position des Schliffls unsicher ist. Auch spielt die deutlich gröbere Pixelquantifizierung durch den RT-Detektor eine Rolle, so dass die Zuordnung nicht exakt dem metallographischen Pixel entspricht.

Der größte Nachteil des Verfahrens ist die Beschränkung auf den maximalen Kontrast. Bei Beschränkung auf voluminöse Defekte gilt, dass besonders auch die Fläche der Anzeige auf dem RT-Bild eine wesentliche Rolle bei der Detektion von Defekten spielt ([115], [146], [147], [105], [148]). Schon vom empirischen Standpunkt ist klar, dass eine größere runde Fläche besser zu detektieren ist als nur ein Punkt (Pixel). Diese Erkenntnis muss sich ebenfalls in der POD widerspiegeln. Ein Ansatz dazu ist die Einführung eines Entscheidungskriteriums, das abhängig von der Anzeigenfläche ist. Dies sollte die mathematischen Voraussetzungen grundsätzlich unverändert lassen (vgl. Observer-POD) [149]. Für die Observerschwelle sind jedoch viele Datenwerte notwendig. Es ist nur selten möglich, die notwendige Datenmenge nur mit Experimenten zu erreichen. Modellierungsergebnisse können teilweise nicht die reale Detektierbarkeit oder die wahre Defektcharakteristika widerspiegeln. Aus diesem Grund soll ein alternatives Vorgehen zur Observer-POD vorgeschlagen werden, indem die Anzeigenfläche eine Rolle spielt, aber der Fokus weiter auf den Defektparameter liegt. Dabei soll auch auf die Unterschiede zwischen realen und künstlichen Defekten eingegangen werden.

3.1.3 POD-Bewertung anhand von mehrerer metallographischen Schlifflen

Ein idealer Fall wäre die exakte Bestimmung der räumlichen Ausdehnung eines Defektes. Jedoch gibt die metallographische Untersuchung nicht Aufschluss über den vollständigen Informationsgehalt. Ein Grund hierfür ist der Bearbeitungsprozess, bei dem Informationen verloren gehen (durch das Sägen und Polieren). Außerdem sind die Untersuchungen pro Schliff kostenintensiv, wodurch auf einen Kompromiss zwischen Kosten und Informationsgehalt geachtet werden muss.

Aber selbst mit wenigen Schliffbildern ist die räumliche Bestimmung des Defekts möglich (siehe Abb. 3.4) [16]. Dabei müssen die Abmessungen des Defekts zwischen den Schlifflen geschätzt werden. Je größer die Schrittweite zwischen den Schlifflen ist, um so größer ist dabei auch der Informationsverlust.

Ein visueller Vergleich zwischen der RT-Anzeige und der Projektion einer ersten Rekonstruktion von Schliffdaten zeigt bereits eine starke Ähnlichkeit (Abb. 3.5).

3.1 Umgang mit Schliffen realer Defekte bei der Bewertung durch die POD

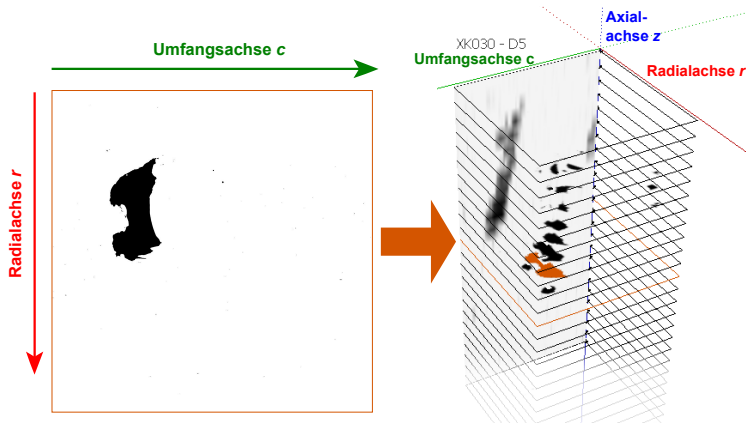


Abbildung 3.4: Auf der linken Seite ist ein digitalisierter Schliff für eine Position dargestellt. Die rechte Seite zeigt eine räumliche Darstellung (Stapel) der einzelnen Schliffe. Die Projektion der groben Rekonstruktion des Stapels ergibt das Schattenbild.

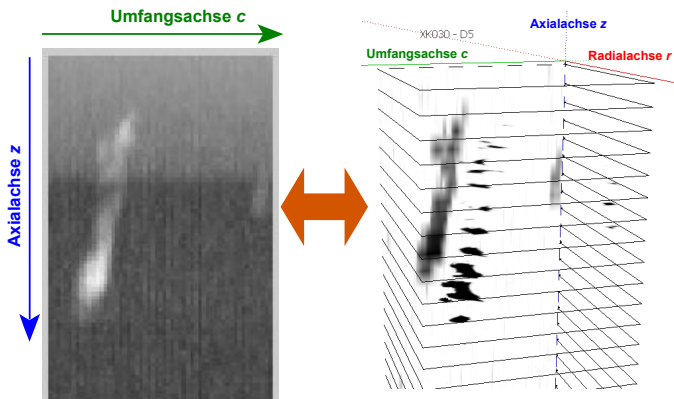


Abbildung 3.5: Vergleich RT-Bild und Schliff-Rekonstruktion

Tabelle 3.1: Vergleich zwischen der RT-Anzeige und dem Schliffbild

| Radiographisches Ergebnis | Metallographisches Ergebnis |
|--------------------------------------|---|
| Anzeigenrauschen | kein oder kaum Rauschelemente |
| geringe Pixelauflösung des Detektors | hohe Pixelauflösung bei der Digitalisierung |
| geometrische Unschärfe | wahre Größe des Defekts in Schliffebene |
| axiale Anzeigengröße ist darstellbar | unsichere axiale Größe durch Schätzung |

3.1.4 Unterschiede zwischen der radiographischen Anzeige und den Informationen aus den Schliffen

Für die Beschreibung des Zusammenhangs zwischen Defekt und Anzeige ist die rein visuelle Auswertung zu ungenau. Daher ist eine detaillierte Gegenüberstellung notwendig. Jedoch werden hierbei folgende Herausforderungen sichtbar (Tab. 3.1):

Während das RT-Bild sowohl Rauschen als auch ein Offset (fester Wert über Null) aufweist, besteht das digitalisierte Schliffbild nur aus binären Informationen (1 für Defektpixel, 0 kein Defekt). Dabei treten nur selten Fehlzuordnungen (Rauschen) auf. Während die Auflösung des RT-Systems durch die Pixelgröße des Detektors gegeben ist, kann der metallographische Schliff mit hoher Auflösung digitalisiert werden.

Die unterschiedliche Pixelauflösung stellt für den Vergleich ein Problem dar. Eine Möglichkeit ist, die metallographischen Pixel in „Bins“ zusammenzufassen [150] (siehe auch Histogramm bei Dichteschätzung), so dass einem RT-Pixel einem Bin gegenübersteht. Bins sind daher Bereiche mit einer Anzahl von Pixel und einer Positionsangabe. Dabei ist nicht klar, wie ein Bin zusammengefasst werden muss, um genau einem RT-Pixel gegenübergestellt zu werden.

Außerdem existiert die Ungenauigkeit des RT-Systems (Unschärfe, Rotationsungenauigkeit, usw.), das sich zwar im RT-Ergebnis widerspiegelt, jedoch nicht in den Schliffinformationen zu finden ist. Wodurch eine weitere Abweichung zwischen den Ergebnissen von RT-Signal und Schliffbild ersichtlich wird.

Weitere Unterschiede zwischen RT-Anzeige und Schliffinformationen in axialer Richtung können durch folgende Umstände entstehen:

1. Winkeleinstrahlung (falls nicht in dem gleichen Winkel geschliffen, wie durchstrahlt wurde)
2. Ungenauigkeit in der Positionierung (dem Schliff kann nicht direkt eine Pixelzeile in der RT-Anzeige zu geordnet werden)
3. Rotationsungenauigkeit (in verschiedenen Abschnitten werden unterschiedliche Pixelzeilen dem Schliff zugeordnet)

Daher ergibt es Sinn einen Mittelwert aus mehreren Pixelzeilen der RT-Anzeige in Axialrichtung zu bilden, um diese Unsicherheit abzuschwächen. Der Mittelwert reagiert hierbei

3.1 Umgang mit Schliffen realer Defekte bei der Bewertung durch die POD

wie ein Filter. Ungenauigkeiten werden ausgeglichen, jedoch sehr kleine Anzeigen, die nur über eine Pixelzeile zu sehen sind, in ihrem Kontrast geschwächt.

3.2 Einführung der Anzeigenbreite in die Zuverlässigkeitsbewertung

Trotz der obengenannten Nachteile besitzt die Metallographie den Vorteil, die nicht sichtbare Form des Defekts zu erfassen. Dadurch ist es möglich, nicht nur die maximal durchstrahlte Länge für einen RT-Pixel, sondern auch die Form und wahre Geometrie des Defekts in der Bewertung als Information zu nutzen. Die Relevanz der Gesamtanzeige der RT, in die beispielsweise Form, Fläche und Signalhöhe eingeht, wurde bereits angedeutet [147], [148], [105]. Ein wichtiges Ziel dieser Arbeit und eine wesentliche Neuerung für die Bewertung mit Hilfe der POD ist es, die Rolle der Anzeigenfläche für die Detektierbarkeit festzustellen und in die Bewertung miteinzubeziehen. In dem vorliegenden Fall wird primär die Breite, die eindeutig der metallographischen Informationen zugeordnet werden kann, genutzt. Die zusätzliche Nutzung der axialen Ausdehnung der RT-Anzeige führt zu einem räumlichen Modell. Ein räumliches Modell ist jedoch durch die grobe Unterteilung in axialer Richtung durch den Schlißprozess schwierig zu erstellen. Daher wird in dem vorliegenden Fall nur die Ausdehnung der Defekte in Umfangsrichtung betrachtet. Dennoch lässt sich durch Einbeziehung der Umfangsrichtung die prinzipielle Möglichkeit der flächigen Bewertung zeigen und die POD-Bewertung erweitern. Während die Umfangsrichtung in die Bewertung einfließt, wird über die axiale Richtung der Mittelwert gebildet, der gleichzeitig die Unsicherheit bei der Auswertung abfedern kann. Ein weiterer Vorteil der reinen Umfangsbewertung ist die einzelne Bewertung der Schlitze. Während für ein räumliches Modell aus mehreren Schlitzen ein Defekt erstellt wird, bleibt bei einer Schliß-basierten Bewertung die Datenanzahl hoch, da jeder Schliß einzeln in die Bewertung einfließt.

Für die erweiterte Auswertung wird nun zwischen den metallographischen Daten und den RT-Daten ein Profil gebildet. Dabei handelt es sich um die Häufigkeit der Defekt-Pixel im metallographischen Bild und des in axiale Richtung gemittelten Kontrastwerts im RT-Bild. Diese werden jeweils auf die Umfangskoordinate aufgetragen [151] (siehe als Beispiel Abb. 3.6). Durch die Mittlung der Kontrastwerte und dem Vergleich mit den einzelnen Schliß-bilder wird die Anzeigenfläche auf die Anzeigenbreite reduziert. Ein Einsatz der Glättung innerhalb der POD-Bewertung wurde teilweise unter [152] an Beispielen gezeigt.

3.2.1 Einsatz einer Glättung

Die Unterschiede zwischen der Metallographie und den RT-Daten verhindern einen direkten Vergleich der Daten. Als Alternative wurde daher ein Glättungsalgorithmus für beide Profildaten eingesetzt. Die Glättung (siehe auch Tiefpassfilterung) ist eine Klasse von Faltungsoperatoren, die eingesetzt werden, um Rauschen oder feine Details zu unterdrücken [153]. Jedoch ist häufig die Ausprägung von Signalhöhe und Anzeigenfläche hinsichtlich dem Unterscheidungskriterium entscheidend, ob es sich um Rauschen oder eine Defektanzeige handelt.

3.2 Einführung der Anzeigenbreite in die Zuverlässigkeitsbewertung

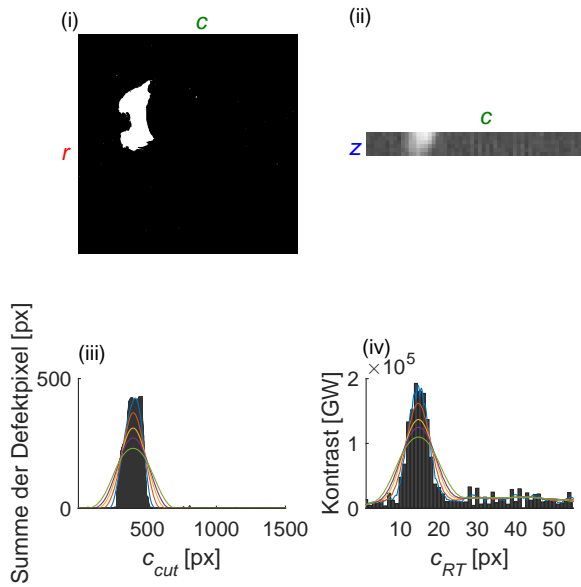


Abbildung 3.6: Gegenüberstellung von Schliff (i) und Schliffprofil (ii) zu RT-Anzeige (iii) und Profil für die Anzeige (iv)

Die Glättung eines Signals entspricht bildlich einem “Abschleifen” der Kanten. Dabei wird die Signalvariabilität unterdrückt, die aus unbestimmten Gründen auftritt und die durch ihre hochfrequente Natur gekennzeichnet ist. Damit werden systematische Merkmale hervorgehoben [154]⁴.

In Folgenden wurde die Glättung mit Hilfe eines Kerns durchgeführt, der im Weiteren eingehender beschrieben wird. Die Glättung mit einem Kern ist eine weitverbreitete Methode, da das Ergebnis einfach und interpretierbar ist [150]. Des Weiteren wird durch geeignete Wahl der Glättungskerns die gewonnene Funktion stetig und differenzierbar, wodurch die Zusammenfassung in diskrete Bins nicht mehr von Nöten ist [155].

Glättungskern

Als Muster wurde hier ein Epanechnikov-Kern genutzt [156]. Weitere Kerne sind der Boxplot, die Dreiecksfunktion, die Gleichverteilung und der in der RT häufig eingesetzte Gauss-Kern [157]. Dabei ist die Wahl des Glättungsparameters relevanter als die Wahl des Kerns [154]. Der Epanechnikov-Kern G gehört zu der Familie Gasser-Müller [154] und wird in der Gleichung 3.1 beschrieben. Der Glättungskern G wird für t definiert:

$$G(t) = \begin{cases} \frac{3}{4}(1 - t^2), & t \leq 1 \\ 0, & t > 1 \end{cases}. \quad (3.1)$$

Für die konkrete Glättung wird dem Kern G die Bandbreite h zugeordnet, bezogen auf die Koordianten der angewandten Daten $(t - x)$:

$$G_h(t - x) = \frac{1}{h} G\left(\frac{t - x}{h}\right) \quad (3.2)$$

und die Daten $y(t)$ mit dem Kern gefaltet, wobei $x \in \mathbb{N}_+$ und $t \in \mathbb{R}$:

$$T_h(x, t) = \sum_t y(t) \cdot G_h(t - x). \quad (3.3)$$

Der Wert t wird hierauf für jedes x so gewählt, dass $T_h(x)$ maximal wird.

So ergibt sich eine detektierte Anzeige, wenn das gefaltete RT-Signal, das vom Positionparameter x und vom Glättungsparameter h abhängt, oberhalb der Schwelle z_h liegt:

$$T_h(x) > z_h. \quad (3.4)$$

⁴Das Hervorheben der Kanten wird nicht betrachtet, da die intrinsische Fähigkeit sich hauptsächlich in den Intensitäten widerspiegelt.

Glättungsparameter

Der Glättungsparameter h ist als Bandbreite des Kerns definiert [151], dessen Größe essentiell für die Glättung ist. Je größer der Parameter h gewählt wird, um so kleiner wird die Varianz der geglätteten Ergebnisse. Dies ist intuitiv plausibel, da für größere h mehrere Pixel in die Glättung miteinbezogen werden [154].

Im Fall von einer großen Bandbreite ist die Kehrseite der Glättung, dass die Signal-Täler und -Berge über- bzw. unterbewertet werden, wodurch schmale Signale ins Rauschen fallen können. Wird die Bandbreite zu gering gewählt, so findet keine Unter- bzw. Überbewertung statt, jedoch bleibt die Streuung groß [154]. Hinsichtlich der ZFP kann die Situation wie folgt zusammengefasst werden:

Besitzen die Daten ein hohes Rauschen, so ist es schwierig kleine Signale zu detektieren. Um den verrauschten Hintergrund zu filtern, ist die optimale Bandbreite möglicherweise so groß, dass auch kleinere Signale noch detektiert werden können [154]. Die Bandbreite sollte sich daher nach der Anzeige des kleinsten noch zu findenden Defekts und dem größtem zu unterdrückenden Rauschsignals richten.

Häufig ist zur Auswahl der Bandbreite die "per Auge"-Methode ausreichend [151]. Zusätzlich können mehrere Werte für h verwendet werden, um verschiedene wichtige Merkmale der Daten hervorzuheben [154]. In der Abbildung 3.6 sind Glättungen mit verschiedenen Bandbreiten h in unterschiedlichen Farben dargestellt.

3.2.2 Anzeigenflächen-bedingte Probability of Detection

Die Wahrscheinlichkeit einen Defekt mit RT zu finden, wird von den Informationen, die dargestellt werden, und von der Wahrnehmung des Prüfers beeinflusst. Während die Wahrnehmung von der Erfahrung des Prüfers abhängen kann, sind die Informationen, die dem Prüfer zur Verfügung stehen, rein technischer Natur. In den bisherigen POD-Bewertungen wurde nur das zugrundeliegende physikalische Prinzip basierend auf dem maximale Kontrast (z.B. Schwächungsgesetz in RT) verwendet. Diese beschreiben aber nur die technischen Sichtweisen. Dadurch ist es weder möglich, reale und künstliche Defekte miteinander zu kombinieren (siehe Abschnitt 3.2), noch damit eine Basis für die Bewertung hinsichtlich der Wahrnehmung von Anzeigen zu erstellen. Außerdem wird die entscheidende Rolle der Anzeigenbreite und der Defektgeometrie bei der Detektion außer Acht gelassen.

Die Einbeziehung der Anzeigenfläche und -Form in die POD wurde bereits in der Bewertung der UT erforscht, um die Detektierbarkeit in C-Scans (Anzeigen-Bewertung als Draufsicht) zu erfassen [158]. Ein weitere Ansatz der die Einflüsse der Anzeigenfläche für die Detektion erkennt, ist die Observer-POD [133]. Die Glättung der RT-Anzeige bietet jedoch ein weitaus fortgeschrittenes Hilfsmittel, die Anzeigenfläche in die Bewertung mit-einzubeziehen.

Einführung der flächenabhängigen Detektionsschwelle

Ein Ansatz, der die Abhängigkeit der Anzeigenflächen für die POD beschreibt, stellt der Observerschwellwert dar. Da es wichtige Parallelen zwischen dem Glättungsansatz und der Observerschwelle gibt, soll nun detaillierter auf die flächenabhängige Detektionsschwelle eingegangen werden. Die detaillierte Beschreibung der Methodik findet sich unter [133] wider. Dieses Vorgehen basiert teilweise auf dem POD-Modell für die Detektierbarkeit von Anzeigen im Einsatz für Modellierungsprogrammen [159], welches wiederum auf das Detektionsmodell von Anzeigenflächen zurückgeht [160].

Dass die Detektierbarkeit von Defekten neben dem Kontrast auch von der Anzeigenfläche abhängt, wurde bereits angesprochen. Dies in der Berechnung der POD miteinzubeziehen stellt eine große Herausforderung dar. Durch den zusätzlichen Parameter kann das einparametrische POD-Konzepte nicht ohne Weiteres übernommen werden. Eine Möglichkeit das statistische Konzept und die Bestimmung der funktionellen Abhängigkeit von Kontrast und durchstrahlter Länge unverändert zu lassen, ist die Einführung von mehreren Schwellen. Diese sind dann abhängig von der Anzeigenflächen. So muss für jede Anzeigenfläche eine eigene Schwelle und somit eine neue POD berechnet werden. Handelt es sich bei den Defekten hauptsächlich um räumliche Defekte (z.B. Poren), kann davon ausgegangen werden, dass die durchstrahlte Länge des Defekts gleich der Ausdehnung des Defekts in axialer oder Umfangsrichtung ist.⁵ In diesem Fall ist es sinnvoll, Umfangsausdehnung und durchstrahlte Länge als gleich anzunehmen. Da es gleichzeitig durch die Vergrößerung der Ausdehnung in Axial- und Umfangsrichtung auch zu einer Vergrößerung der RT-Anzeigenfläche kommt, kann der Fokus auf die Bestimmung der Schwellenabhängigkeit gerichtet werden. Hierzu konnte mit Modellierungsergebnissen eine Detektionsschwelle abhängig von der Anzeigenfläche erfasst werden [133] (siehe Abb. 3.7).

Dabei wird ersichtlich, dass der detektierbare Kontrast, der gerade noch wahrgenommen werden kann, kleiner wird, wenn die Anzeigenfläche sich vergrößert. Dies ist eine intuitiv verständliche Tatsache bei der Detektion von Anzeigen und hat somit einen Einfluss auf die POD-Bewertung. Bei der Annahme von festen Detektionsschwellen werden entweder kleinere Defektflächen überbewertet oder größere Defekte unterbewertet (siehe Abb. 3.8).

Die Bestimmung der Observerschwelle erfordert eine große Datenmenge von Defekten, die gerade noch oder gerade nicht mehr zu detektieren sind. Während experimentelle Daten zu teuer sind, kann die Datenmenge durch Modellierungsergebnisse erreicht werden. Ein weiteres Problem liegt in der objektiven Bestimmung der Detektierbarkeit der Anzeigen, da hierbei menschliche Faktoren eine wichtige Rollen spielen können. Die Anforderung, dass die durchstrahlte Länge gleich der Defektausdehnung in axialer und Umfangsrichtung sein soll, ist teilweise vereinfacht, da reale Schweißdefekte häufig eine favorisierte Ausdehnungs-

⁵Ein ähnliches Vorgehen wird auch in der ASTM E 476 für die Referenzkörper mit den T 1 Bohrungen angenommen.

3.2 Einführung der Anzeigenbreite in die Zuverlässigkeitsbewertung

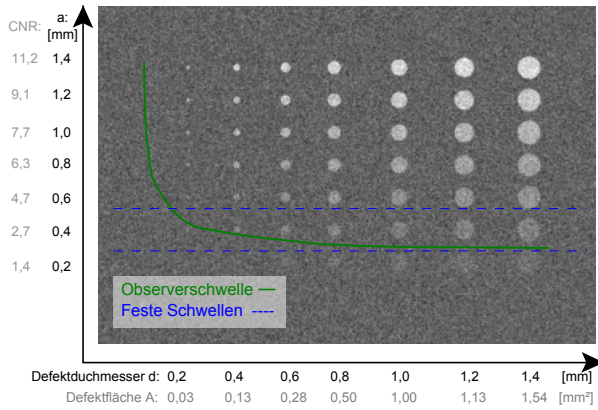


Abbildung 3.7: Observerschwelle für modellierte FBH

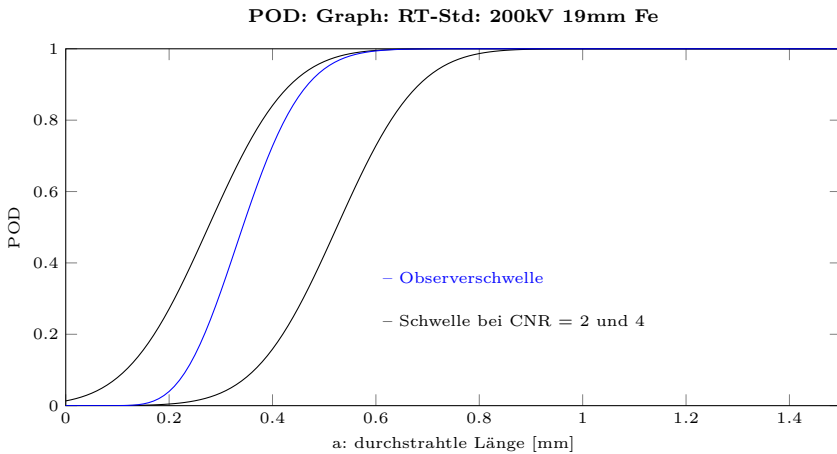


Abbildung 3.8: Observer-POD im Vergleich zu zwei konstanten Schwellen

richtung besitzen [16]. Aus diesen Gründen ist es sinnvoll ein neues Vorgehen mit Hilfe der Glättung einzuführen, das auf den Erkenntnissen der Observer-POD basiert.

Beziehung zwischen Signal und Defektparametern unter Einbeziehung der Anzeigenbreite

Basierend auf den physikalischen Gesetzmäßigkeiten soll nun das POD-Konzept mit Hilfe verschiedener Glättungsparameter erweitert werden. Die Gegenüberstellung von \hat{a} und a lässt sich auf Basis des Schwächungsgesetzes erklären (siehe Abschnitt 2.2). Dies kann nach der Glättung der RT-Anzeigenprofile und Schliffprofile im Allgemeinen beibehalten werden. Dies gilt besonders für großflächige Anzeigen, da hier selbst breite Glättungsbandbreiten h nur einen sehr geringen Einfluss auf die Amplitude haben (siehe Abb. 3.9).

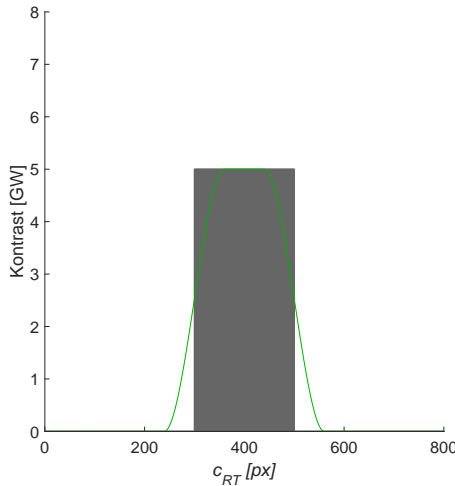


Abbildung 3.9: Glättung einer großen Anzeige ohne Rauschen

Des Weiteren ist die Breite der Anzeige abhängig von der Umfangsbreite des Defekts. Wird ein großer Defekt durchstrahlt, so wird auch eine große Anzeige erwartet. Durch die Glättung wird sowohl das RT-Profil als auch das Schliffprofil in der Breite beeinflusst. Während bei großen Anzeigen und bei adäquat gewählten Parametern die Beziehung beibehalten werden sollte, muss bei großen Bandbreiten damit gerechnet werden, dass kleine Anzeigen im Rauschen untergehen (siehe Abb. 3.10).

Ein weiterer Vorteil für die Detektierbarkeit ist das Filtern des hochfrequenten Teils des Rauschens. Dies hat den großen Vorteil, dass der Schwellwert niedriger gelegt werden kann, da bei geringerem Rauschen Anzeigen besser erkannt werden [161]. Die Filterung des Rauschens und die Senkung der Amplitude bei kleinen Defektgrößen sind daher zwei Konsequenzen der Glättung, die es gegeneinander abzuwägen gilt. Inwieweit kleine Defekte nach

3.2 Einführung der Anzeigenbreite in die Zuverlässigkeitsbewertung

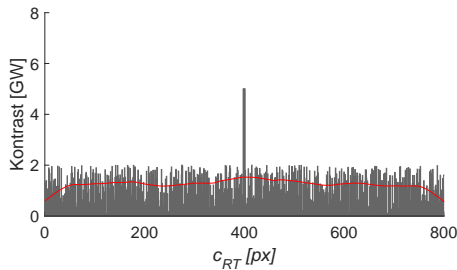


Abbildung 3.10: Kleinere Anzeigen könnten nach der Glättung im Rauschen untergehen

der Glättung gefunden werden können, wird maßgeblich durch die Detektionsschwelle beschrieben.

Einfluss der Glättung auf die Detektionsschwelle

Auf die Detektionsschwelle haben der Kontrast des Defektes zu seiner Umgebung und die Anzeigenbreite Einfluss. Dieses Verhalten wird auch im Trend wiedergegeben, wenn Glättungsverfahren eingesetzt werden. Die Form des Defekts kann vorerst vernachlässigt werden [133].

An dieser Stelle soll die Verbindung zwischen Glättung und Detektionsschwelle gezeigt werden: So zeigt sich, dass es für jede Anzeigenbreite je nach Rauschen einen optimalen Glättungsparameter gibt. Wird die Bandbreite zu klein gewählt, bleibt das Rauschen des Hintergrunds bestehen und die Detektionsschwelle liegt relativ hoch. Wird die Bandbreite zu groß gewählt, so sinkt das Signal des Defekts bezogen auf das Rauschen, wodurch die Aussage verfälscht wird. Gleichzeitig verändert sich auch die Schwelle durch die Glättung, da bei der Unterdrückung von Rauschen auch die Detektionsschwelle niedriger gesetzt werden kann. Um eine aussagekräftige Bewertung zu erhalten, wird jeder einzelnen Anzeige ein Glättungsparameter zugeordnet und individuelle Schwellen abhängig von ihrer Breite zugewiesen. Im der vorliegenden Arbeit wird hierfür ein multi-skalen Glättungsansatz verwendet. Dabei wird mit verschiedenen Bandbreiten die vorliegenden Daten untersucht. Es gilt hierbei, dass je größer eine Defektfläche ist, um so größer auch der Glättungsparameter gewählt werden kann. Intuitiv scheint dieses Vorgehen auch dem Detektionsverhalten eines menschlichen Auges nahe zukommen, da bei einer größeren Fläche mehr Informationen integriert werden, wodurch das Rauschen gefiltert wird.

3.2 Einführung der Anzeigenbreite in die Zuverlässigkeitsbewertung

Eine ähnliche Erkenntnis zeigt auch die Schwelle bei der Observer-POD [133]. Je größer die Anzeigenfläche ist, um so geringer ist das notwendige CNR, um den Defekt zu detektieren. Im Fall der Glättung werden mit größer werdender Fläche die größeren Bandbreiten eingesetzt (siehe Abb. 3.11, wobei $h_1 < h_2 < \dots < h_5$ ist).

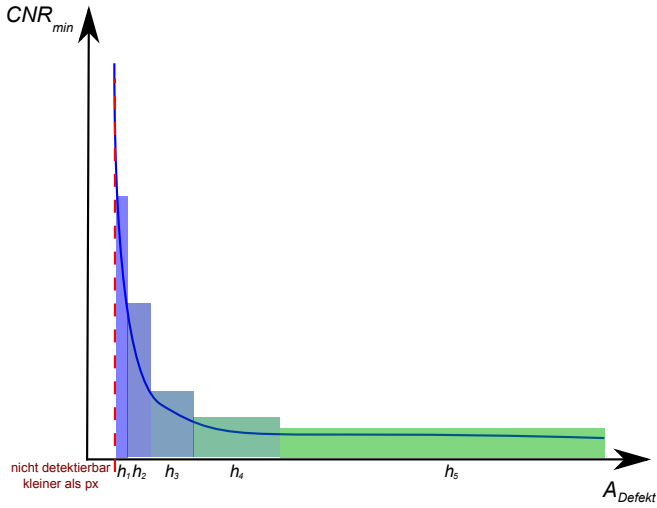


Abbildung 3.11: Bedeutung der Glättungsparameter in Beziehung zur Observer-POD-Schwelle

Der Einsatz der multi-skalen Glättung auf diese Weise in der POD ist eine Neuerung in dem Portfolio der POD-Methodiken.

Berechnung der Auffindwahrscheinlichkeit mit Berücksichtigung der Anzeigenbreite

Die Abhängigkeit von der Anzeigenbreite b , kann innerhalb der POD-Bewertung mit Hilfe des Glättungsparameters h beschrieben werden. Gleichzeitig bleibt jedoch der Hauptparameter für die POD-Bewertung a (bei RT die durchstrahlte Defektlänge) unverändert.

$$POD(b, a) = POD_{gl}(h, a) \quad (3.5)$$

Bei der Berechnung wird die Glättung sowohl der RT-Anzeige als auch des Schliffprofils mit mehreren Bandbreiten durchgeführt. Durch die höhere Auflösung des Schliffprofils im Vergleich zum RT-Profil und durch die Unschärfe des Prüfprozesses ist die Bandbreite der Glättung des Schliffprofils größer als die der RT-Anzeige. Für den Einsatz in der POD ist es notwendig die Bandbreiten aufeinander anzupassen. Es spielen sowohl die Form des Profils

3.2 Einführung der Anzeigenbreite in die Zuverlässigkeitsbewertung

als auch die Breite eine Rolle. Die verschiedenen Glättungsergebnisse werden dabei Skalen genannt. Dabei sei angemerkt, dass für jede Skala, sowohl die lineare Abhängigkeit zwischen Signal und Defektparameter als auch das Modell für die Streuung unverändert bleiben. Dadurch kann für jede Skala eine eigene POD-Kurve berechnet werden. Des Weiteren ist die optimale Wahl der Bandbreite h vom Defekt abhängig. In dem vorliegenden Fall ist dies hauptsächlich von der Anzeigenbreite beeinflusst. Soll nun eine POD bestimmt werden, so wird auf Basis der Defektgeometrie die passende Bandbreite und somit die dazugehörige POD ausgewählt. Um den Einfluss der Glättung und somit der Anzeigenbreite auf die POD nachvollziehen zu können, sollen hierbei theoretisch zwei Bandbreiten betrachtet werden. Während h_1 eine geringe Bandbreite besitzt und schmale Anzeigenbreiten repräsentiert, steht die große Bandbreite h_5 für große Anzeigenbreiten. Den Einfluss der Bandbreitengröße auf Profilanzeigen findet sich in Anhang III.

Wenn für das Schliffprofil und das RT-Profil jeweils mit einer geringen und einer großen Bandbreite geglättet werden und für die POD-Bewertung im \hat{a} vs. a -Graph gegenübergestellt werden, so kann Folgendes beobachtet werden (siehe Abb. 3.12):

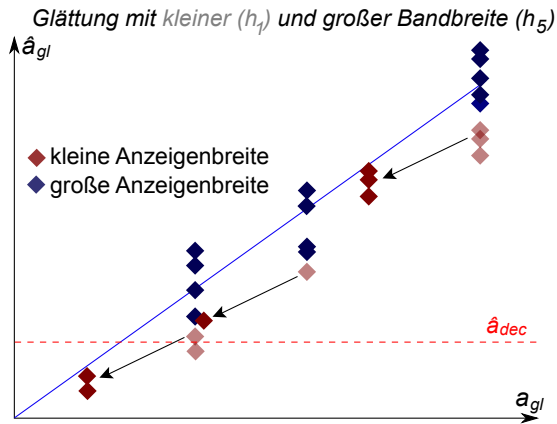


Abbildung 3.12: Einfluss der Bandbreite auf den \hat{a} vs. a Graph

Bei großen Anzeigen besitzt die Glättung weder mit einer großen Bandbreite noch mit einer kleinen Bandbreite Einfluss auf die Signalhöhe und auch nicht auf das Schliffprofil. Wird eine große Bandbreite gewählt, so bleiben die großen Anzeigen unverändert, jedoch sinken die kleinen Anzeigen ab. Dies gilt sowohl für das Schliffprofil als auch für das RT-Profil. Durch den unterschiedlichen Einfluss der Glättung auf Schliff- und RT-Profil, wird ein stärkerer Drift auf der Abszisse erwartet. Dadurch zeigt sich jedoch eine geringere Streuung der Datenwerte um die Regressionsgerade. Es wurde beobachtet, dass die Signalwerte

3.2 Einführung der Anzeigenbreite in die Zuverlässigkeitsbewertung

von kleineren Anzeigenflächen geringe Signalwerte besitzen. Diese werden durch die Glättung geringeren a_s zugeordnet, wobei die Streuung des Gesamtmodells sinkt. Gleichzeitig wird das Verteilungsmodell und der theoretische Zusammenhang besser geschätzt. Dadurch verringert sich auch die Breite des Vertrauensbands. Ein weiterer Vorteil ist die Möglichkeit einer sinkenden Entscheidungsschwelle bei größeren Bandbreiten. Die erwartenden Auswir-

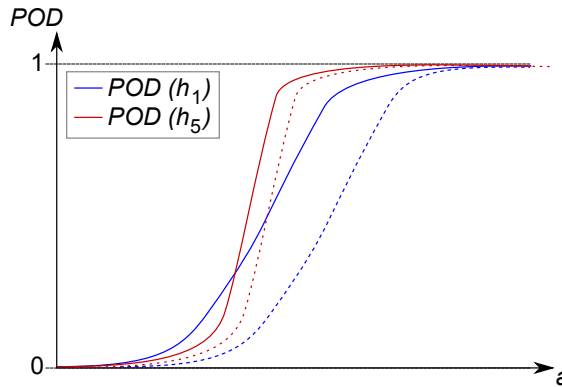


Abbildung 3.13: Einfluss der Bandbreite auf die POD-Kurve

kungen auf die POD-Kurve sind ein schmaleres Vertrauensband und eine steilere Kurve durch die Reduktion der Datenstreuung (siehe Abb. 3.13). Ob jedoch kleine voluminöse Defekte besser gefunden werden, hängt einerseits an der sinkenden Schwelle gleichzeitig jedoch auch am sinkenden Signal. Durch diese beiden Effekte kommt es vor, dass sich die POD-Kurven von geringen und großen Bandbreiten schneiden.

Die Abszisse des neu erstellten POD-Graphen muss nun in eine durchstrahlte Länge übergeführt werden: Dabei wird eine Umrechnung von der Auflösung des Schliffprofils in durchstrahlte Länge (in Millimetern) in abhängig von der Defektgröße (bzw. Glättungsbandbreite) notwendig. Somit kann die POD in Abhängigkeit von der durchstrahlten Defektlänge in Millimeter für die gewählte Defektbreite angegeben werden.

3.3 Kombination von realen und künstlichen Defektdaten für die Zuverlässigkeitsbewertung

Die Einführung der statistischen Methodik für die Evaluation von ZfP-Verfahren war ein wichtiger Schritt, dem probabilistischen Charakter der ZfP Rechnung zu tragen. Um ein aussagekräftiges Ergebnis mit Hilfe der Statistik zu erreichen, sollte jedoch die Datengrundlage repräsentativ und ausreichend groß sein. Beide Anforderung führen zu hohen Kosten (siehe [16]). Aus diesem Grund wurden immer wieder Ansätze verfolgt die notwendige Anzahl von Daten intelligent zu reduzieren und trotzdem ein aussagekräftiges Ergebnis zu erhalten. Dies wurde einerseits durch die Einführung physikalischer Zusammenhänge (Signal-Antwort-POD) als auch durch Hinzunahme von Modellierungsergebnissen (MaPOD) versucht. Gerade dann, wenn die Experimente sehr kostenintensiv sind, ist der korrekte Einsatz der Statistik unter Einbeziehung sämtlicher wichtiger Informationen unumgänglich [80]. Daher ist es in vielen Anwendungen notwendig, Wissen über die Detektion von Defekten mit einer limitierten Anzahl realer Daten zu verbinden [45].

Im Fall der Signal-Antwort-POD sollte jede Situation einzeln betrachtet werden (z.B. Defektarten, Prüfsituation, Geräteeinstellung, usw.). Die Vertreter der Transferfunktion behaupten, diese Brücke mit Hilfe von Modellierungssoftware schlagen zu können [143]. Jedoch sind die Anforderungen für den Transfer durchaus schwierig zu erfüllen. Die aus Sicht der Transferfunktion geforderte Annahme um Daten auf eine andere Situation zu übertragen, benötigt eine sehr detaillierte Einschätzung der Prüfsituation.

In dieser Arbeit wird eine alternative Methodik vorgeschlagen: Dabei sollen die vorhandenen Daten genutzt werden, die bereits die Fähigkeit des ZfP-Systems abschätzen. Während die allgemeine Signal-Antwort-POD nicht zwischen einem Referenzdefekt, einem künstlichen oder einem realen Defekt unterscheidet, soll es hier berücksichtigt werden. Die Glättung ist ein adäquates Hilfsmittel um diese verschiedenen Situationen miteinander zu verbinden. Dabei rückt die Form des Fehlers in den Hintergrund und die Anzeigenfläche kann zusätzlich verwendet werden.

3.3.1 Bewertung der künstlichen und realen Defekte

Die Beschreibung der künstlichen und realen Defekte ist ein essentieller Punkt bei der Bewertung eines ZfP-Systems, da beide Arten von Defekten berücksichtigt werden. Da sich diese Defekte grundsätzlich unterscheiden, sollen hier die Differenzen aufgezeigt werden:

Die Herstellung von realistischen Defekten und die Größenbestimmungen von realen Defekten ist ein teures Unterfangen [125]. Daher stehen reale Defektdaten meist nicht in ausreichender Form zur Verfügung. Während reale Defekte selten für die Bewertung zur Verfügung stehen, sind künstliche Defekte relativ einfach herzustellen. Aus diesem Grund sollten die

3.3 Kombination von realen und künstlichen Defektdaten für die Zuverlässigkeitsbewertung

wertvollen realen Defekte in die Bewertung einfließen. Dies kann die Richtigkeit und die und Aussagekraft des späteren Ergebnisses erhöhen.

Hingegen liegt die Bedeutung für künstliche Defekte in der prinzipiellen Abschätzung der Fähigkeit des ZfP-Systems während des Kalibrierungsprozesses und im Zusammenhang mit den Normen [162]. Bei der Bewertung eines Systems im Einsatz (z.B. Herstellungs-, In-Service-Prüfung) sind reale Defekte die ideale und richtige Wahl [133]. Die Repräsentativität muss für künstliche Defekte hinterfragt werden. Selbst wenn die künstlichen Defekte nach Einfluss auf Produktzuverlässigkeit, Auftrittswahrscheinlichkeit und Defektart bzw. -Form ausgewählt werden müssen, ist ihr Einfluss fraglich, da die Auswahl der Defekte subjektiv vom Experimentator bestimmt wird [133].

Im Vergleich zwischen den RT-Anzeigen und dem wahren Defektparameter, können sich Unsicherheiten sowohl bei dem Defektparameter als auch bei der RT-Anzeige ergeben:

Für künstlich eingebrachte (gebohrt oder geschnittene) Defekte hängt der Parameter vom mechanischen Prozess ab. Häufig wird die wahre Defektgeometrie basierend auf der technischen Auftragszeichnung angenommen, da die metallographische Untersuchung vermieden wird. Die einzigen Möglichkeiten die Geometrie von realen Defekten zu verifizieren, sind hochauflösenden ZfP-Methoden (z.B. Computertomographie, Synthetischen Apertur Fokus Technik (SAFT), hochauflösende Ultraschalltauchttechnikprüfung [163]) oder metallographische Untersuchungen [16] (siehe auch Abschnitt 3.2).

Bei künstlichen Defekten liegt meist eine klar definierte Form vor. Sowohl die Oberfläche als auch die Geometrie verändert sich innerhalb des Defekts nur geringfügig, was wiederum nur geringfügige Variationen des Kontrastes innerhalb der Defektgrenzen zur Folge hat. Bei realen Defekten können sich Oberfläche, Geometrie und Form der Defektgrenze signifikant ändern. Den realen Defekt zeichnen unerwartete Änderungen des Kontrastes aus, was zu einer deutlich größeren Streuung von Datenwerten, um die zu erwartende Signal-Antwort-Regressionsgeraden führt.

Durch diese Tatsache ist eine einfache Mischung der Daten nicht mehr möglich. Eine Möglichkeit diese Daten trotzdem miteinander zu kombinieren, ist eine gewichtete Regressionsanalyse [133].

3.3.2 Einfluss der Glättung auf die Ergebnisse

Während bei der bisherigen gewichteten Regressionsanalyse der Zusammenhang zwischen Defektparameter und Anzeigensignal im Fokus stand [133], verschiebt sich dieser durch die Einführung der Glättung zu der unterschiedlichen Streuung der Anzeigenwerte von realen und künstlichen Defekten.

Zur Einschätzung des Regressionszusammenhangs wird nun im Weiteren die Maximum-Likelihood- Schätzung (siehe Abschnitt 2.2) genutzt. Bei der Streuung wird jedoch auf die Unterschiede zwischen realen und künstlichen Defekten eingegangen. Der gewichtete

3.3 Kombination von realen und künstlichen Defektdaten für die Zuverlässigkeitsbewertung

Anpassungsfehler (Residuen) wird daher auf folgendes neues Modell für die geglätteten Daten erweitert:

$$r_{gew} = \left(\sum_i \frac{(y_i - f(x_i))^2}{\sigma_{real}^2} \cdot w_{real} + \sum_j \frac{(y_j - f(x_j))^2}{\sigma_{artificial}^2} \cdot w_{artificial} \right) / w_{total} \quad (3.6)$$

i beschreibt hierbei die Anzahl der realen Datenwerte, die eine Anzeigenstreuung von σ_{real}^2 aufweisen. y steht für den Kontrast der Anzeige, x für die Defektgröße. j definiert die Anzahl der künstlichen Datenwerte mit Streuung $\sigma_{artificial}^2$.

Bei der Abschätzung der Streuung wird eine subjektive Gewichtung w_{total} eingeführt, auf das das Expertenkonsortium direkt Einfluss nehmen kann. Dieses Expertenkonsortium sollte dabei beachten, dass künstliche Defekte die realen Defekte nicht optimal darstellen oder die Anzahl der realen Defekte zu gering im Vergleich zu den künstlich Eingebachten ist.

$$w_{total} = w_{artificial} + w_{real} \quad (3.7)$$

w_{real} kann beispielsweise auf 1 festgelegt werden, da die Gewichtung sich relativ zu den Einzelbeiträgen berechnet. Ein datenbasierender Einsatz, der sich auf die gewichtete Regressionsanalyse stützt, stellt folgendes Beispiel dar:

$$w_{artificial} = \frac{\sigma_{real}^2}{\sigma_{artificial}^2} \quad (3.8)$$

Diese Gewichtung hat Einfluss auf den Vertrauensbereich, wie auch auf die Schätzung mit gewichteten kleinsten Fehlerquadraten - engl. weighted least squares estimation (WLSE) [133]:

$$n_{total} = n_{artificial} \cdot w_{artificial} + n_{real} \quad (3.9)$$

3.3.3 Beeinflussung der Gewichtung

Die Gewichtung muss primär durch die Signifikanz der Daten hinsichtlich der POD bestimmt werden (vgl. WLSE) [133]. Eine niedrige Gewichtung kann Ausreißer deutlich abschwächen, wodurch einerseits eine Möglichkeit geschaffen wird, bestimmte Effekte zu unterdrücken, andererseits können auch essentielle Eigenschaften und Parametereinflüsse in der Bewertung abgeschwächt werden. Daher ist ein Fokus auf die Defektarten und Versuchsplanung notwendig. Aus diesen Gründen ist das Wissen eines unabhängigen Expertenkreises unabdingbar [45].

3.3 Kombination von realen und künstlichen Defektdaten für die Zuverlässigkeitsbewertung

An dieser Stelle sei auf die kritischen Punkte der zu starken Beeinflussung der Gewichtung hingewiesen: Es soll nicht die Möglichkeit geschaffen werden, anhand von Daten ein Ergebnis nach eigenen Vorstellungen zu erschaffen, sondern vielmehr vorhandene Daten zu nutzen, um eine möglichst realistische Beschreibung zu erhalten. Hierbei kann eine Parallele zu Bayesschen Ansätzen und deren Kritik gezogen werden: Wie immer, wenn ein subjektiver Charakter mit Daten verbunden wird, ist dies der stark diskutierte und kritisierte Aspekt des Vorgehens [164]: Zwei Bewerter können mit den gleichen Daten verschiedene Ergebnisse erhalten [80]. Als Gegenargument kann genannt werden, dass schon allein die Auswahl der Daten im Fall der Fähigkeitsbestimmung subjektiv ist [165]. Um diese Punkte jedoch zu entschärfen, sollen hierbei einige Richtlinien aufgezählt werden: Die Gewichtung sollte aus umfassender Erfahrung aus der Vergangenheit und physikalischer Theorie gebildet werden [45]. Klar ist auch, dass mit einer zunehmenden Anzahl realer Daten die künstlichen Daten an Einfluss verlieren müssen [80], [79]. Des Weiteren wurde schon an mehreren Stellen der Versuch unternommen, Wissen und Erfahrung in quantitative Zahlen zu fassen. Auch diese können hierbei zu einem Gewinn beitragen [31], [95], [166]. In wie weit unterschiedliche POD-Bewertung jedoch miteinander verglichen werden können, ist ein Diskussionspunkt der momentanen Forschung [167]. Gezeigt wurde, dass der einfache Einsatz von Bildgüteprüfkörper [112] nicht den Einsatz von realen Defekte für die POD-Bewertung oder im Einsatz von Ringversuchen ersetzen kann. Durch die gleichzeitige statistisch unzureichende Anzahl der realen Defekte, ist das vorgestellte Verfahren eine Lösungsmöglichkeit.

4 Experimentelle POD-Bewertung des Radiographiesystems bei der Prüfung der Elektronenstrahlschweißnaht bei Endlagerkanistern

4.1 Beschreibung des zu bewertenden ZfP-Systems

Der vorgestellte Ansatz zur Abschätzung der Fähigkeit eines ZfP-Systems soll bei der Implementierung der ZfP für die Endlagerung von hochradioaktiven Abfällen aus finnischen Kernkraftwerken Einsatz finden. Sowohl Finnland als auch Schweden fassen die Lagerung im fenno-skandinavischen Granitgestein für die Endlagerung ins Auge. Die Auswahl dieses Gesteins fordert unter anderem eine zusätzliche Sicherheitsbarriere des Behälters, der das radioaktive Material umfasst, gegen Korrosion. Eine adäquate Sicherheitsbarriere besteht darin, den Behälter mit einer Außenhülle aus Kupfer mit einer Wanddicke von 5 cm auszustatten [168].

Die Kupferummantlung muss gemäß seiner strukturellen Integrität bestimmte Forderungen hinsichtlich Zeitraum und Korrosionsrate standhalten. Eine dieser Forderungen ist die Detektion von Herstellungsdefekten ab einer bestimmten Größe, die die Funktionalität des Kanisters gefährden können. Diese sollen dann ausgesondert werden. Hierbei soll die ZfP zum Einsatz kommen. Eines der eingesetzten ZfP-Verfahren ist die RT, die mit Hilfe von verschiedenen Defekten evaluiert werden soll.

4.1.1 Der Kupferkanister als Korrosionsbarriere

Korrosionsvorgänge ziehen kostenintensive Folgen nach sich, da durch ihre Einflüsse beispielsweise Produktionsprozesse ausfallen oder angehalten werden müssen [24]. Bei der Korrosion handelt es sich um die Degradation von Material (Materialverschlechterung). Dabei wird durch einen chemischen Angriff die Bauteildicke einer Komponente reduziert. Es wird hierbei zwischen der allgemeinen (großflächigen Abtrag der Wanddicke) und der lokalen Korrosion (Angriff an einer kleinen Fläche) unterschieden [46]. Der Korrosionsprozess hängt hierbei vom Bauteil und den Umgebungseinflüssen ab [46].

Obwohl die Korrosion ein komplexer Schädigungsmechanismus ist, so stellt sich dennoch der Kupferbehälter als adäquate Korrosionsbarriere für den notwendigen Zeitraum einer Endlagerung dar [168].

4.1 Beschreibung des zu bewertenden ZfP-Systems

Ist die Kupferummantlung fehlerfrei, so ist ein Leck, das durch Korrosion hervorgerufen wurde, unwahrscheinlich. Liegt jedoch ein Materialdefekt vor, so könnte auch bei sehr geringer Korrosionsrate durch den großen Zeitraum die Ummantlung durchlässig werden. Dabei werden Oberflächen- und oberflächennahe Defekte [168] und Defekte, die während des Herstellungsprozesses im Material entstehen können [54], getrennt behandelt.

Als Anforderung an den hergestellten Kupferkanister gilt, dass in 100% des Umfangs eine intakte Wanddicke von 3,50 cm vorhanden sein muss und 4,00 cm in 99% des Umfangs [14].

Elektronenstrahlschweißung der Kupferummantelung

Obwohl die Schweißnaht kein besonderes Verhalten hinsichtlich der Korrosion zeigt [168], bildet sie dennoch durch die Auftrittswahrscheinlichkeit der Defekte eine Schwachstelle für den Kanister [33].

Die Schweißnaht der Kupferummantlung wird beispielsweise mit Hilfe einer Elektronenstrahlschweißung - engl. electron beam welding (EBW) hergestellt.¹ Dabei handelt es sich bei der EBW um ein Schmelzschweißverfahren, wobei ein fokussierter Elektronenstrahl eingesetzt wird [169], um die Temperatur auf das Schweißgut zu übertragen.

Die EBW bietet sich für die Schweißung des Kupferdeckels an die Mantelfläche an, da sie eine große Eindringtiefe (im vorliegenden Fall: 4,20-7,00 cm [14]) erreichen kann [170]. Doch gerade die Schweißung dicker Schichten ist eine Herausforderung, was bedeutet, dass mit Materialdefekten während des Schweißprozesses zu rechnen ist [170].

Die Untersuchung des EBW-Prozesses gibt spezifische und zu erwartende Defekte an, die während des Prozesses in der Schweißnaht entstehen können [171], [172]. Spezifische Defekte konnten im Entwicklungsprozess nachgewiesen werden, während zu erwartende Defekte bisher bei der Untersuchung noch nicht gefunden wurden [44]. Beide Arten von Defekten könnten jedoch die Funktionsfähigkeit des Kanisters gefährden und daher sollten sie durch die ZfP entdeckt werden. Ein Gesamtüberblick über die Defekte findet sich unter [44].

Die Planung der ZfP spielt dabei eine essentielle Rolle. Wesentliche Gesichtspunkte, die durch ein Expertenkonsortium festgelegt werden, sind die Auswahl der ZfP-Methoden, die Beurteilung der Materialdefekte und der Entscheidungsprozess bei der Prüfung [33], [44].

Beurteilung der Materialdefekte hinsichtlich der Endlagerung

Bei der Beurteilung der Materialdefekte stehen die möglichen Folgen eines Defekts im Vordergrund. Könnte ein Defekt zu einem Leck im Kanister und damit zum Austritt von radioaktivem Material führen, muss dieser Defekt als kritisch gekennzeichnet werden. Dabei sind das Materialverhalten (z.B. Korrosionsrate, Eigenspannung, Kriechverhalten), das Bau-

¹Sowohl die schwedischen als auch die finnischen Endlagerfirmen haben sich seit Kurzem für die Rührschweißung entschieden.

teildesign (z.B. erlaubte Geometrietoleranz) und der kritische Defekt (Defektgeometrie und Defektart) essentiell zur Beschreibung des Ausfallrisikos des Kanisters [44].

Aus probabilistischer Sicht besteht das Ausfallrisiko aus drei wesentliche Faktoren [33]:

1. Auftrittswahrscheinlichkeit: es tritt ein Defekt in der Schweißnaht auf
2. Übersehenswahrscheinlichkeit (1-POD): ein Defekt wird bei der Prüfung übersehen. (Im Weiteren wird das Gegenereignis, die Auffindwahrscheinlichkeit, genutzt: POD)
3. Ausfallwahrscheinlichkeit: innerhalb eines bestimmten Zeitraums führt ein nicht detektierter Defekt zu einem Leck des Kanisters

Die Auftrittswahrscheinlichkeit des Defekts ist abhängig vom Herstellungsprozess und dessen Parameter. Die Defekte, die während der EBW entstehen können und die nachgewiesen werden sollen, sind u.a. Krater, Einbrandkerben, Hohlräumen sowie Risse [172], [173].

Die Auffindwahrscheinlichkeit ist durch den Einsatz der ZfP-Verfahren bedingt und besitzt einen wesentlichen Einfluss auf die Gesamtbewertung [33]. Die Bestimmung der POD ist dabei abhängig von der klaren Spezifikation von Defekten, die auftreten und die detektiert werden müssen [53].

Auf die Ausfallwahrscheinlichkeit wirkt die Umgebung und die sonstigen Bestandteile des Kanister (siehe hierzu [168]).

4.1.2 Zerstörungsfreie Prüfung des Kupferkanisters

Eine erneute Prüfung nach einem bestimmten Zeitraum im Einsatz (Wartungsprüfung) ist im Fall der Endlagerung nicht möglich. Daher müssen relativ kleine Defekte sicher detektiert werden. Ein einzelnes ZfP-Verfahren kann jedoch nicht alle unterschiedlichen Defektarten sicher finden. Während es sich bei Kratern oder Einbrandkerben um Oberflächendefekte handelt [173], liegen Hohlräume und Risse teilweise vollständig im Material, ohne ein Anzeichen an der Oberfläche zu hinterlassen. Hohlräume zählen zu den räumlichen Defekten [174], während Risse als planare oder flächige Defekte beschrieben werden [174]. Die Kombination von mehreren ZfP-Verfahren bietet in diesem Fall entscheidende Vorteile [46].

Die visuelle Prüfung mit einem Kamerasystem und die hochfrequente Wirbelstromprüfung (30kHz) detektieren Oberflächendefekte. Die niederfrequente Wirbelstromprüfung (200 Hz) erfasst oberflächen- und oberflächennahe Defekte, während für Volumendefekte die Ultraschallprüfung und die RT eingesetzt werden [14] (siehe Abb. 4.1).

Der entscheidende physikalische Parameter bei der RT-Prüfung ist die durchstrahlte Länge. Damit ist die RT-Prüfung eher für voluminöse Defekte geeignet, die eine Ausdehnung in Durchstrahlungsrichtung besitzen [46]. Im Gegensatz dazu ist für die Ultraschallprüfung die Fläche lotrecht zur Schallausbreitungsrichtung wichtig. Die Ultraschallprüfung kann daher planare Defekttypen mit einer großen Fläche sicher detektieren [33], [175], [176].

4.1 Beschreibung des zu bewertenden ZfP-Systems

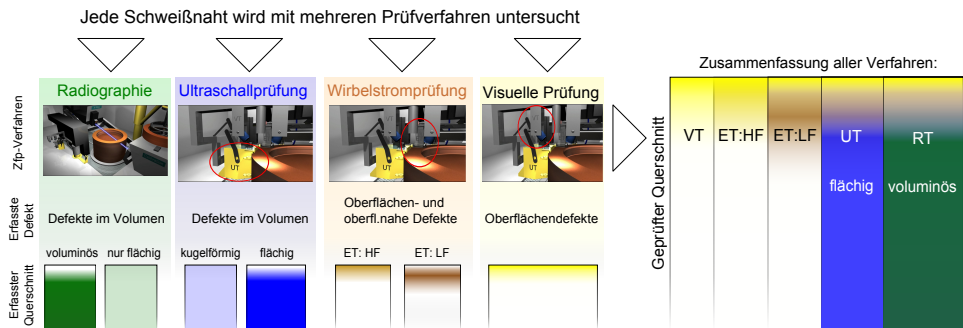


Abbildung 4.1: Die vier verschiedenen Prüfverfahren zur Sicherstellung einer defektfreien Schweißnaht

Voluminöse Defekte für die radiographische Prüfung

Im Weiteren wird der Fokus dieser Arbeit auf die Defekte gelegt, die eine räumliche Ausdehnung und damit eine Ausdehnung in Durchstrahlungsrichtung besitzen und mit der RT nachgewiesen werden sollen. In Hinsicht auf die strukturelle Integrität der Schweißnaht stehen folgende Defekte im Fokus [173]:

- Gasporositäten
- Hohlräume durch Spiking
- Ungenügende Durchschweißung
- Unterwölbung

Bei den Gasporositäten handelt es sich um spezielle Hohlräume [174]. Diese können die Form und Größe von kleinen Poren besitzen oder sich als Schlauchpore über die gesamte Länge der Schweißnaht ziehen. In wie weit der Defekt bei der Korrosion zur Gefährdung des Kanisters führt, ist noch nicht bekannt. Daher wird die Gasporosität vorerst als kritisch eingestuft [172]. Liegen sie jedoch tiefer als 1,50 cm im Material, so sind sie nach Bowyer unkritisch. Mit Lunkern, einer weiteren Art von Hohlräumen, ist in ähnlicher Weise zu verfahren, wobei die Tiefeneinschränkung nicht gilt [172]. Hohlräume sind der Haupteinsatzgrund für die Radiographie, da sich bei kleinen realen Defekten mit irregulärer Form Detektionsschwierigkeiten für die UT ergeben [171].

Bei den Hohlräumen durch Spiking (Durchlöchern - engl. „Gun Discharge Defects“) handelt es sich um eine Defektkategorie, die speziell bei der EBW auftritt. Bei einem elektrischen

Überschlag können Defekte in der Schweißnaht bis zur kompletten Tiefe entstehen [172]. Diese Defekte können durch eine gute Elektronikontrolle des Schweißequipments und durch einen adäquaten Strahldurchmesser vermieden werden [170]. Außerdem kann ihr Auftreten während des Schweißprozesses anhand von Blitzen beobachtet werden [177].

Eine weitere Gruppe von Defekten entsteht durch eine nicht komplette Schweißung der geplanten Tiefe [174], beispielsweise wegen ungenauer Führung des Elektronenstrahls. Dieser Defekt kann durch die RT (aber auch die UT [172]) nachgewiesen werden [173].

Unterwölbungen entstehen, wenn geschmolzenes Metall aus der Schweißnaht läuft und nicht mehr für die Schweißung zur Verfügung steht (engl. „Run-out“). Gleichzeitig kann sich eine Oberflächenüberhöhung ausbilden [174]. Die RT ist hierbei ein Verfahren zur Detektion dieses Defekts [173].

Beschreibung des zu bewertenden RT-Systems

Im Weiteren werden die technischen Daten des Prüfsystems vorgestellt. Das zu bewertende RT-System wird als Durchstrahlungsprüfung betrieben. Als Strahlenquelle dient ein Linearbeschleuniger, der in einem festen Winkel (10°) auf das Prüfobjekt gerichtet ist. Der Kanister als Prüfobjekt, in der Mitte der Anordnung, ist drehbar gelagert. Auf der Gegenseite dient ein kollimierter Zeilendetektor zum Erfassen der durchgehenden Strahlung (Siehe Abb. 4.2).

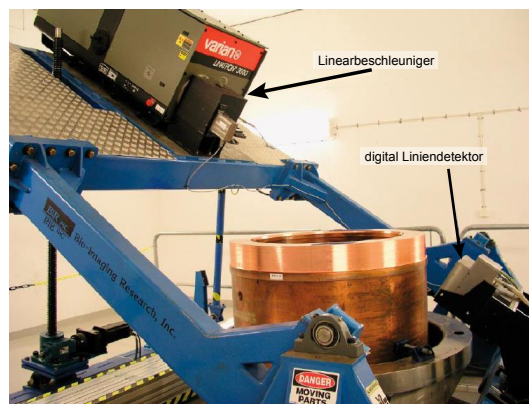


Abbildung 4.2: Der Prüfaufbau mit Linearbeschleuniger, Detektor und Manipulatorsystem (Quelle: [178] hier mit einem Einstrahlwinkel von 30°)

Der Linearbeschleuniger - engl. LINear ACcelerator (Linac) Linatron 3000 der Firma Varian dient als Strahlenquelle. Der Linac wurde für die Pilotprüfungen der Schweißnaht des Kupferkanisters von der schwedischen Firma SKB angeschafft [43]. Der Linac erzeugt Röntgenstrahlung durch linear beschleunigte Elektronen, die auf ein Transmissionstarget treffen [178]. Das für die EBW-Prüfung genutzte Strahlenspektrum besitzt eine Grenzenergie von

4.1 Beschreibung des zu bewertenden ZfP-Systems

9 MeV. Das Maximum des Spektrums wurde auf 3 MeV abgeschätzt [179], [180], [181]. Die Brennfleckgröße des Linacs wurde mit 2,40 mm vertikal und 2,60 mm horizontal bestimmt [181]. Hinter der Strahlenquelle wird die Strahlform mit Hilfe eines Schlitzkollimators (Tiefe: 7,00 cm, Material: Wolfram, Schlitzbreite 0,40 mm [181]), 3,00 cm hoch und 3 mm breit geformt (der Strahlkegel besitzt hinter dem Kollimator eine Breite von 3°) [43]. Durch diesen Aufbau wird für eine Aufnahme nur ein kleiner Bereich des Materials der Strahlung ausgesetzt.² Durch das hohe Strahlenspektrum der ausgesandten Energie, die große durchstrahlte Wanddicke und dem Winkel zwischen Strahlenquelle und Prüfobjekt ist die Anordnung nicht geeignet für eine tomographische Aufnahme, wie in der Computertomographie. Daher soll dieses Thema nicht weiter betrachtet werden.

Als Detektor für die Prüfung der EBW-Naht wurde ein DDA ausgewählt [178]. Für den DDA-Einsatz sprechen die geringe notwendige Strahlung für eine Aufnahme und die direkte elektronische Speicherung und Weiterverarbeitung am Computer. Jedoch ist die minimale Auflösung relativ gering im Vergleich zu einem Film (typischerweise 0,10 mm) [178].

Der hier eingesetzte Detektor ist ein Liniendetektor der Firma Bio-Imaging Research [43]. Der Einsatz des Liniendetektors mit einem Kollimator führt zu einer starken Reduktion der eintreffenden Streustrahlung. Ein Schlitzkollimator vor dem Detektor, mit der Breite von 0,40 mm in horizontaler Richtung, führt zusätzlich zu einer Reduktion der Streustrahlung. Daher ist die Streustrahlung kleiner als 3% der Primärstrahlung und kann in der weiteren Bewertung außer Acht gelassen werden [180].

Die Szintillationselemente des Detektors besitzen eine Länge von 100 mm und liegen parallel zum Schlitz. Der Detektor besitzt 2048 Linien mit einer Breite von 0,05 mm. Acht Linien bilden einen Pixel im Bild. Dies entspricht in etwa einer Pixelbreite von 0,40 mm für das RT-Bild [43]. Das Szintillationsmaterial ist $CdWO_4$ mit einer Dicke von 5 mm. Dies wandelt die Röntgenstrahlung in Licht um. Das Licht trifft dabei die Photodioden und lässt ein elektrisches, messbares Signal entstehen [43]. Die Abbildung 4.3 zeigt eine schematische Skizze des Aufbaus und ein Bild des eingesetzten Detektor.

Der Winkel zwischen Strahlenquelle und Prüfobjekt ist ein wichtiger Prüfparameter. Abhängig von der Art und Position der Schweißnaht wird der Einstrahlwinkel für das Prüfproblem optimiert. Für die hier vorliegende Prüfung der EBW ist ein Prüfwinkel von 10° festgelegt [43].

Weitere wichtige Einstellungen sind die Abstände zwischen Quelle, Prüfobjekt und Detektor. Für die Prüfaufgabe wurde ein Abstand zwischen dem Target der Quelle bis zum Detektor von 2180 mm eingestellt [182]. Zwischen Quelle und durchstrahltem Bereich der Kupferummantlung liegen 1800 mm. Dabei sei angemerkt, dass im Prüfobjekt durch den schrägen Einstrahlwinkel und die Geometrie des Deckels die durchstrahlte Wanddicke in-

²Der bestrahlte Bereich wird meist so gewählt, dass der Detektor vollständig genutzt werden kann.

4.1 Beschreibung des zu bewertenden ZfP-Systems

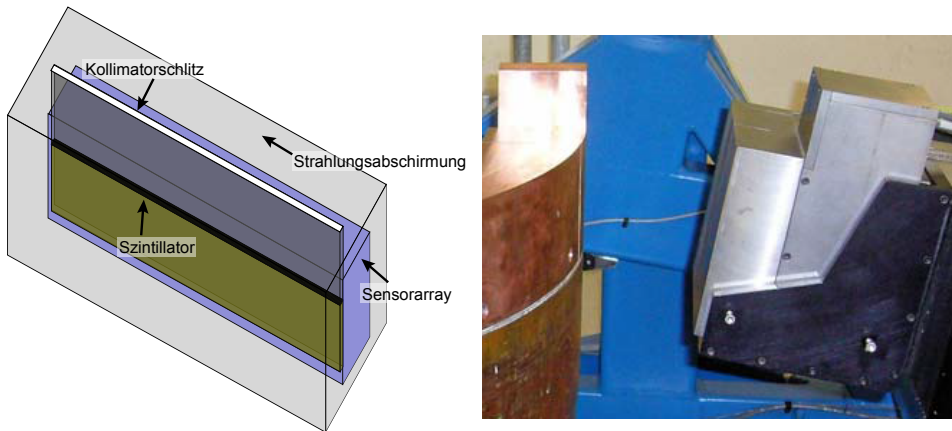


Abbildung 4.3: Schematischer Aufbau und Bild des genutzten Detektors (Quelle: [145])

nerhalb einer Aufnahme variiert. Die Abbildung 4.4 zeigt das leicht vereinfachte Modell für Modellierungssoftware bei der beschriebenen Prüfanordnung.

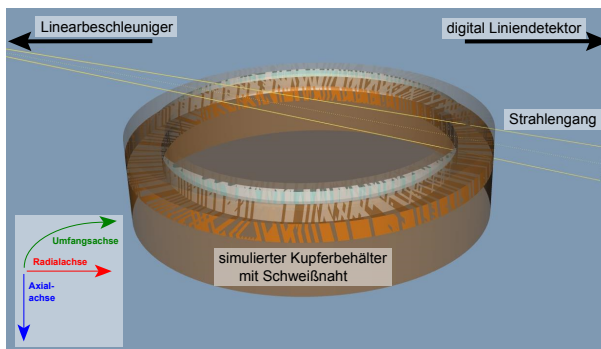


Abbildung 4.4: Die Modellierung für die RT, bei der Prüfung der Schweißnaht

Um die gesamte Schweißnaht mechanisiert zu prüfen, wird mit Hilfe eines Manipulatorsystems der Kanister mit einer konstanten Geschwindigkeit rotiert [43]. Die Aneinanderreihung der Bilder des Liniendetektors bilden das zu analysierende RT-Bild. Bei der gesamten Prüfung muss auf einen gleichbleibenden Quelle-Objekt-Abstand geachtet werden, um die Schweißnaht zuverlässig und wiederholbar zu prüfen.

4.1.3 Bewertungsgrundlage des Radiographie-Systems

Zur Bewertung der Zuverlässigkeit eines RT-Systems helfen die normativen Anforderungen, die schon in Abschnitt 2.1 erwähnt wurden. Dabei sind die Beurteilung von Aufnahmen mit Hilfe von Einzel- und Doppeldraht- und Stufe-Lochbildgüteprüfkörper gemeint. Für den Fall einer einwandigen Durchstrahlung eines 100-125 mm dicken Kupferbauteils wurden bereits während der Entwicklung des Prüfverfahrens die Bildgüteklasse A fast für den vollständigen Umfang erreicht (mit Hilfe des Kompensationsprinzips II) [43], [41]. Dabei konnte eine Schwankung in der Bildqualität abhängig von der gemessenen Position wahrgenommen werden [43]. Teilweise wurden Basis-Ortsauflösungen unterhalb der Detektorpixelgröße erreicht, da während der Prüfung eine Projektionsvergrößerung stattfindet [43]. Für das weitere Vorgehen wird jedoch der niedrigste gemessene Wert der maximalen Basis-Ortsauflösung SR_b^{image} von 0,50 mm genutzt [43], um das Ergebnis der Zuverlässigkeitsbewertung möglichst konservativ zu gestalten.³ Neben den Betrachtungen der Normen soll die Fähigkeit unter Verwendung von künstlichen und realen Defekten beurteilt werden.

Künstliche Defekte

Für die Untersuchungen wurden folgende künstlichen Defekte in ein geschweißtes Segment eines Kanisters eingebracht [43]: verschiedene Nuten, quadratische Flachbodenbohrungen und kreisförmige Flachbodenlöcher (siehe Abb. 4.5). Diese spiegeln Aspekte der zu erwartenden realen Defekte wider. Folgende durchstrahlte Tiefen (in Richtung der durchstrahlten Länge) und Bohrungsdurchmesser (bzw. Nutenlängen) wurden für die Bewertung ausgewählt.

- Quadratische Flachbodenlöcher- engl. flat-bottom-squares (FBSQ) für die Kantenlänge 8 mm, 4 mm, 2 mm und 1 mm mit jeweils den Tiefen von 4 mm, 2 mm und 1 mm
- Kreisförmige FBH für den Durchmesser 8 mm, 4 mm, 2 mm und 1 mm mit jeweils den Tiefen von 4 mm, 2 mm und 1 mm
- Nuten mit der Länge 12 mm, 6 mm und 3 mm mit jeweils den Tiefen von 4 mm, 2 mm und 1 mm mit einer maximalen Breite von 1 mm

Die Abbildung 4.6 zeigt das Ergebnis einer RT des Referenzprüfkörpers mit dem oben beschriebenen System. Hier sind die künstlichen Materialdefekte (und Bildgüteprüfkörper (engl. IQI: Image Quality Indicator)) zu erkennen. Der Einsatz von Drahtbildgüteprüfkörpern macht hier nur begrenzt Sinn, da diese häufig im Vergleich mit rissähnlichen Anzeigen genutzt werden. Die oben aufgeführten erwarteten Defekte jedoch eine voluminöse Struktur besitzen.

³Es sei zu bemerken, dass bei Posiva ein weiterer Detektor mit geringer Pixelgröße und somit kleiner Basis-Ortsauflösung in Planung ist [183].

4.1 Beschreibung des zu bewertenden ZfP-Systems



Abbildung 4.5: Referenzprüfkörper für die RT-Prüfung mit Nuten und Bohrungen (Quelle: [178])

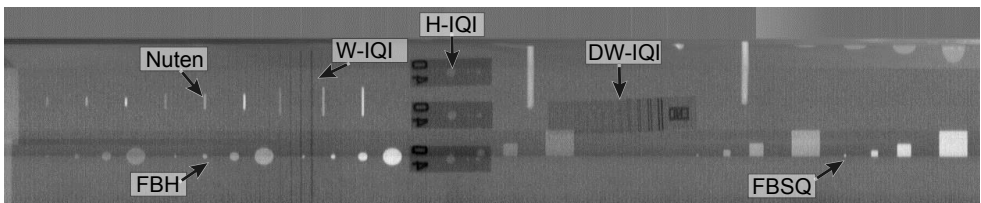


Abbildung 4.6: Radiographische Anzeige des Referenzprüfkörpers mit Flachbodenbohrungen (FBH), Flachbodenvierecke (FBSQ), Nuten und Bildgüteprüfkörper (Draht- (W-IQI), Doppeldraht- (DW-IQI) und Stufe-Loch-Bildgüteprüfkörper (H-IQI))

Realistische Defekte

Das Referenzprüfstück ist eine preiswerte Möglichkeit, Defekte für die Beurteilung herzustellen. Hierzu müssen allerdings die genauen Defektparameter bekannt sein. Jedoch stellt sich die Herstellung bestimmter Parameter als kompliziert dar (wie Rauheit [122]). Die realistischen Defekte, die mit Hilfe von Schweißparametermodifizierungen erreicht werden, sind deutlich schlechter zu kontrollieren [122], jedoch ähnlicher dem zu erwartenden Defekt im Schweißprozess als die künstlichen Defekte.

Für den Einsatz der Signal-Antwort-POD ist eine exakte Bestimmung der Defektgröße notwendig. Die Größen dieser realistischen Defekte müssen mit Hilfe von tomographischen (z.B. Computerröntgentomographie) oder metallographischen Verfahren bestimmt werden. Beide Verfahren besitzen ihre Nachteile. Tomographische Verfahren sind hochauflösender und somit genauer als die eingesetzten ZfP-Verfahren. Jedoch kann von einer sicheren Defektgröße oder der Existenz eines Defektes nicht immer ausgegangen werden. Die Metallographie wiederum liefert ebenfalls keinen exakten Wert für die maximale Größe des Defekts [84], da die exakte Abmessung des Defekts nur an der Stelle des Schliffs vorliegt. Bearbeitungsschritte und die genaue Bestimmung der Schlifffosition sind jedoch die Herausforderung in diesem Prozess.

Für den Fall der vorliegenden Bewertung wurden die realistischen Defekte mit mehreren Schliffbildern untersucht [14]. Dadurch ergibt sich die Möglichkeit einer räumlichen Rekonstruktion des Defekts. Für den Vergleich der unterschiedlichen Daten wurde ein zylindrisches Koordinatensystem gemäß des Kupferbehälters festgelegt (siehe hierzu Abschnitt 3.1):

Die Durchstrahlungsrichtung der RT ist dabei in Radialachse. Die Tiefe bei künstlichen Defekten bzw. die durchstrahlte Länge für reale Defekte sind auf dieser Achse angeordnet. Jedoch sollte man beachten, dass die Einstrahlung in einem Winkel von 10° vorgenommen wurde. Die geringe Abweichung bei der Bemaßung der realen Defekte kann in den Schliffbildern im Vergleich zur RT-Anzeige durch den Einstrahlwinkel vernachlässigt werden. Im RT-Bild geben die verschiedenen Grauwerte die Informationen in Radialrichtung wider. Im Fall der ein-parametrischen Signal-Antwort-POD wird daher der Kontrast der durchstrahlten Länge gegenüber gestellt (siehe Abschnitt 2.2).

Die Breite der künstlichen Defekte verläuft in Umfangsrichtung. Für reale Defekte wurde diese Koordiante auf 20 mm beschränkt. Die Umfangskoordiante beschreibt dabei gleichzeitig die Lage an der die Schriffe aus der Schweißnaht entnommen wurden. Im radiographischen Bild hat dieser Bereich eine Breite von ca. 58 px. Da die Auffindbarkeit von Defekten von der Anzeigenfläche abhängt, ist dies eine wichtige Information für das weitere Vorgehen (siehe Abschnitt 3.2).

Für künstliche Defekte, die hinreichend bekannt sind, verläuft die Länge der Nuten in axialer Richtung. Durch Schriffe in 2 mm Abständen konnte für diese Defekte nur eine Abschätzung durchgeführt werden. Im radiographischen Bild werden die 2 mm dicken Schriffe

4.1 Beschreibung des zu bewertenden ZfP-Systems

als 6 px dargestellt. Die RT ist in dieser Richtung genauer als die Metallographie. Um die Daten in der Axialachse miteinander vergleichen zu können, wurde die RT-Anzeige über diese 6 px gemittelt (siehe Abschnitt 3.2).

4.2 Ergebnisse zur Bestimmung der POD gemäß der entwickelten Methodik

Der folgende Abschnitt zeigt die Resultate für den Einsatz der POD, die die Anzeigenfläche mit in die Bewertung einbezieht. Dabei ist die Bestimmung der Detektionsschwelle ein wichtiger Bestandteil der Auswertung. Außerdem ermöglicht die Erweiterung der POD eine Kombination aus Daten von realen und künstlichen Defekten. Eine Modellierungssoftware ist ein passendes Werkzeug zur Erstellung der Detektionsschwelle und ihre Ergebnisse können mit den künstlichen und realen Defektdaten verglichen werden und sie unterstützen. Künstliche Defekte sind die am häufigsten gebrauchte Datengrundlage für POD-Bewertungen zur Herstellungsprüfung [36]. Im Fall der realen Defekte steht der Vergleich zwischen metallographischen Schliffen und RT-Anzeige im Fokus der Bewertung. Bei der Einführung der Anzeigenfläche ist zu bedenken, dass sich ab einer bestimmten Anzeigenfläche eine Sättigung ergibt. In dieser Arbeit wird von einer maximalen Anzeigenbreite und -höhe von 10 mm ausgegangen, da bei deutlich größeren Defekten der angenommene Zusammenhang gestört werden könnte.

4.2.1 Modellerte Daten zur Überprüfung der Methodik

Die Modellierungsergebnisse tragen maßgeblich dazu bei, die Detektionsgrenze zu beschreiben, da eine Vielzahl von Daten generiert werden kann. Dies wäre im Allgemeinen wegen des enormen Kostenaufwands bei künstlichen und realen Defektdaten nicht erschwinglich. Die Kosten für Prüfstücke und Defekteinbringung sowie auch die Prüfzeit sind hierfür verantwortlich. Allerdings können Modellierungsergebnisse nur unter gewissen Voraussetzungen für die Zuverlässigkeitsbewertung benutzt werden. Dabei ist der Vergleich zwischen modellierten Ergebnissen und Experiment unerlässlich. Grundsätzlich ist die Nutzung von Modellierungsergebnissen für die POD-Berechnung kritisch zu sehen, da möglicherweise wichtige Defekt-, Prüf-, oder Umgebungsparameter nicht in die Modell-Bildung mit einbezogen wurden. Jedoch können fundierte Modellierungsprogramme als ein erster Schritt helfen die Methodik zu verifizieren. Für den späteren Einsatz sind weitere Annahmen und der Vergleich mit Daten zwingend notwendig.

Im ersten Schritt wird eine Verbindung zur Observer-POD hergestellt. Dabei wird die RT-Prüfung einer Schweißnaht in einem Stahlblech betrachtet. Hierbei kann leicht auf vorhandene Erfahrungen zurückgegriffen werden, da es sich dabei um eine häufig auftretende Prüfsituation handelt. Im zweiten Schritt steht die Prüfung der EBW-Naht im Kupferkanister im Vordergrund. Die modellierten Daten ermöglichen einen Einblick in die Fähigkeit des zu bewerteten Systems. Dadurch ist eine Abschätzung der Flächenabhängigkeit in der POD für die erfassten Prüfergebnisse a priori möglich.

Modellierung der Schweißnahtprüfung eines Stahlblechs

Ein mit Defekten versehenes Stahlblech ist eine leicht zu modellierende Prüfsituation und daher gut geeignet für eine allgemeine Überprüfung der Methodik. Für die Verifizierung der Detektionsschwelle und wegen der Verbindung zur Observer-POD wurde die RT-Prüfung eines 19 mm Stahlblechs mit verschiedenen Löchern modelliert. Als Quelle dient eine Röntgenröhre mit einer Beschleunigungsspannung von 200 kV. Ein D4-Film wird als Detektor angenommen (Pixelgröße 0,05 mm und einer Basisortsauflösung von 0,08 mm). Als Verhältnis zwischen Primär- und Streustrahlung wurde ein kollimiertes System mit einem Build-up Faktor ($1 + k$) von 2 angenommen (siehe Abschnitt 2.3). Die Schwärzung im Hintergrund soll eine optische Dichte von 2 betragen. Die durchstrahlte Länge der Defekte wird im Bereich der Detektionsgrenze zwischen 1,00 und 0,10 mm variiert. Die Höhe der Anzeige wird auf 6 px beschränkt, um einen späteren Vergleich mit den Bewertungen von realen Defekten zu ermöglichen. Daher kann auch im Weiteren eine vereinfachte Abhängigkeit von der Anzeigenbreite statt von der Fläche ausgenommen werden. Als Modellierungssoftware wurde hierbei „aRTist“ verwendet [108].

Dabei entstehen Anzeigen in dem Bereich “klar zu detektieren” bis “im Rauschen untergehen” (siehe Abb. 4.7). Für die Modellierungsergebnisse liegen keine wirklichen Schlitze vor, es wurde daher basierend auf den angegebenen Defektdaten ein theoretischer Schliff kreiert (Pseudoschliff). Die angenommene Auflösung für das Schliffbild wurde auf 0,04 mm festgelegt.

Als erste Bewertungsvariante wird die ein-parametrische POD errechnet (nach Abschnitt 2.2). Hierbei wurde die durchstrahlte Länge als Defektparameter ausgewählt. Die Abbildung 4.8 zeigt den linearen Zusammenhang zwischen Defektparameter und Signal (maximaler Kontrast in Grauwerten) mit einer starken Streuung der Datenwerte um die Regressionsgerade. Die Homoskedastizität und die symmetrische Verteilung der Streuung können angenommen werden. Ein Wechsel des Regressionsmodells oder des Verteilungsmodells scheint daher nicht notwendig zu sein. Jedoch ist die Spannweite der Streuung groß. Es muss bei dem konventionellen Ansatz mit einer hohen Detektionsschwelle gearbeitet werden, um die zu erwartende Falschalarmrate für dieses Modell gering zu halten. Dadurch wird auch die Auffindwahrscheinlichkeit geringer geschätzt und das System möglicherweise schlechter bewertet, als es in Wirklichkeit ist. Die unsichere Schätzung der Spannweite zeigt sich auch im breiten Vertrauensbereich. Es ist nachvollziehbar, dass das Vertrauensintervall hier keine große Aussagekraft besitzt, da durch eine Modellierungssoftware beliebig viele Daten erstellt werden können. Aus diesem Grund läuft der Vertrauensbereich gegen den Mittelwert.⁴ Dennoch kann der Vertrauensbereich hier als Qualitätsmerkmal angesehen werden. Dabei ermöglicht die Breite des Vertrauensbereichs eine Abschätzung, in wie weit das ma-

⁴Ein adäquater Vertrauensbereich für Modellierungsergebnisse ist momentan Gegenstand der aktuellen Diskussionen auf internationalem Niveau [184].

4.2 Ergebnisse zur Bestimmung der POD gemäß der entwickelten Methodik

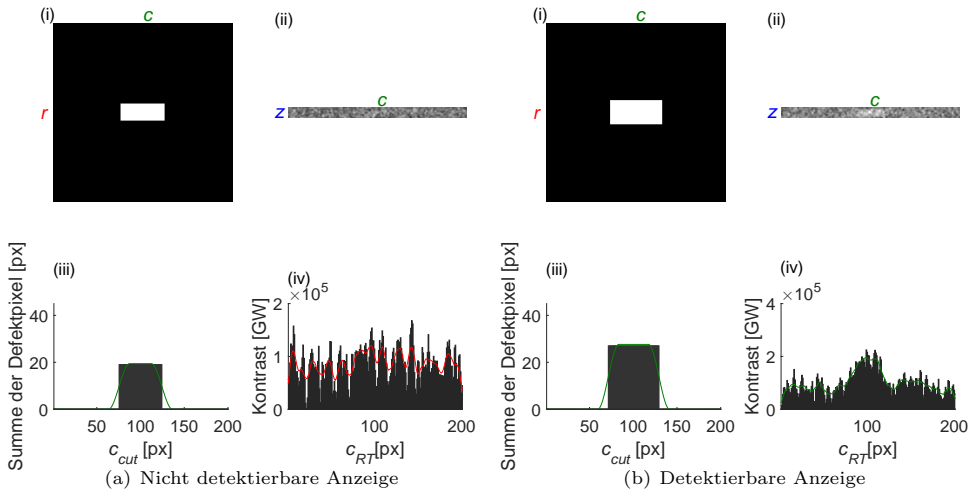


Abbildung 4.7: Pseudo-Schliff (i), RT-Anzeige (ii) und Profile mit Glättungskurve (iii, iv) von modellierten Daten für ein 19 mm FE-Blech

thematische Modell adäquat geschätzt wurde. Der große Streubereich legt nahe, dass die ein-parametrische Analyse nicht ausreichend ist.

Für den Einsatz der geglätteten Daten wird das multi-skalen Glättungsmodell (siehe Abschnitt 3.2) eingesetzt. Die Daten werden mit mehreren Bandbreiten geglättet. Beispielhaft wurde dies für zwei verschiedene Bandbreiten ($h = 2 \text{ px}$ und $h = 8,50 \text{ px}$) gezeigt. Die Bandbreite bezieht sich auf die Breite des Glättungskerns hinsichtlich der RT-Anzeige. Der Glättungskern des Schliffprofils wird analog dazu abhängig von der unterschiedlichen Auflösung und der Abschätzung der Unschärfe des RT-Systems bestimmt. Für die POD ist es notwendig einen Defektparameter, im vorliegenden Fall, die durchstrahlten Länge des Defekts als Basis auszuwählen. Die linearen Abhängigkeiten im \hat{a} vs. a Graphen (jeweils linke Seite) und die daraus entstehende POD Kurve (rechte Seite) sind in den Abb. 4.9 und Abb. 4.10 zu sehen. Die Detektionsschwelle wurde in diesem Fall visuell durch ein Experiment mit wenigen Prüfern bestimmt (siehe auch [185]). Die Entscheidungsschwelle wurde auf $CNR = 2,50$ festgelegt (zwischen $1,60 \cdot 10^5$ und $2,00 \cdot 10^5 \text{ GW}$ vgl. hierzu [186]).

Die Auswahl der passenden Bandbreite soll auf Basis der Anzeigenfläche geschehen. Ein Zusammenhang zwischen Bandbreite und Defektbreite kann mit Hilfe von zusätzlichen Experimenten mit Prüfern geschätzt werden, ergibt jedoch in diesem Fall keinen Sinn, da die Anzahl der Prüfer groß sein sollte. Ein Zusammenhang ist gleichzeitig an der korrekten Bestimmung der POD gebunden. Für den vorliegenden Fall soll die durch die Software mo-

4.2 Ergebnisse zur Bestimmung der POD gemäß der entwickelten Methodik

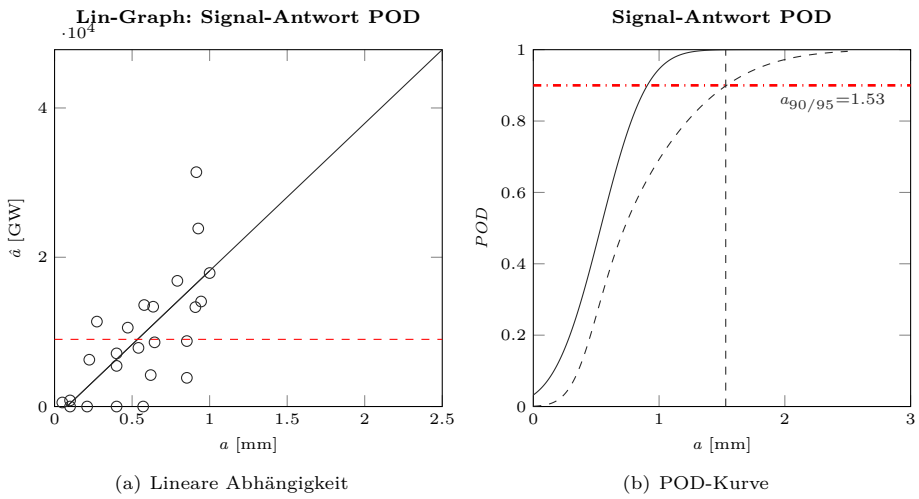


Abbildung 4.8: Conventionele POD-Bewertung der Modellierung einer Schweißnaht eines Stahlblechs

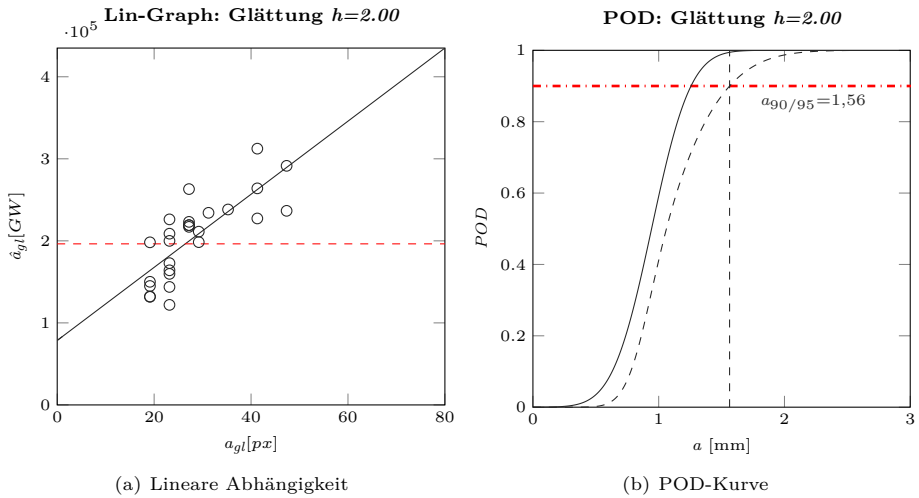


Abbildung 4.9: POD-Analyse mit geglätteten Daten (Glättungsparameter $h = 2$ px)

4.2 Ergebnisse zur Bestimmung der POD gemäß der entwickelten Methodik

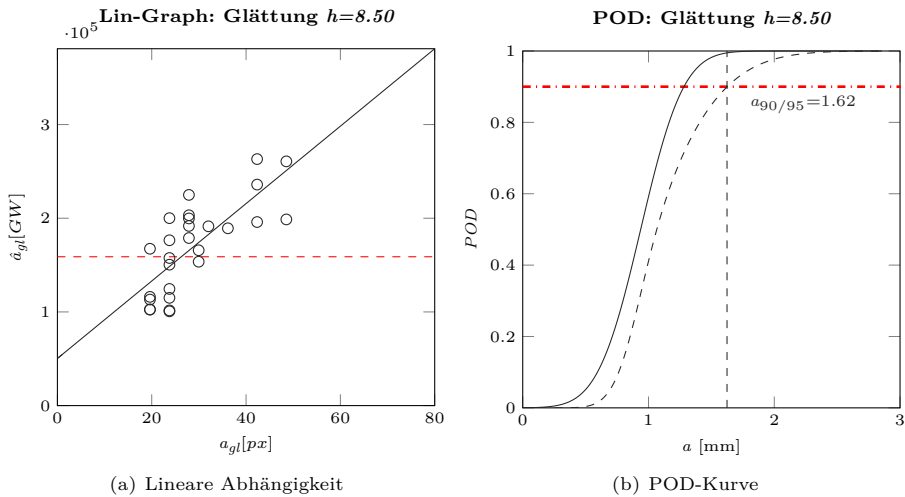


Abbildung 4.10: POD-Analyse mit geglätteten Daten (Glättungsparameter $h = 8,50$ px)

dellierte Prüfung der Stahlblech jedoch nur die Durchführbarkeit der Methodik bewerten. Über die Methodik kann folgendes gesagt werden: Der Einsatz von verschiedenen Glättungsparametern erhöht den Komplexitätsgrad der Berechnung für die erweiterte Analyse. Zudem hängt die Auswahl der Bandbreite h und die Zuordnung zu den Defektbreiten stark vom Rauschen und den zu detektierenden Defektgrößen ab.

Die geglätteten Daten reduzieren die Streuung stark. Dies führt für die POD zu einem schmalen Vertrauensbereich. Dies unterstützt die Auswahl der Modellparameter. Gleichzeitig ermöglicht die Einführung der multi-skalen Glättung eine Abhängigkeit der POD von der Anzeigenfläche. Werden die verschiedenen Glättungsparameter gegenübergestellt, so zeigt sich im vorliegenden Fall, keine signifikante Änderung der Detektierbarkeit durch die größere Anzeigenflächen (vgl. die $a_{90/95}$ -Werte in Abb. 4.9 und Abb. 4.10). Dies kann unter anderem daran liegen, dass die Reduzierung des Rauschens und die Unterbewertung von schmalen Maxima zwei gegensätzliche Eigenschaften sind, die im vorliegenden Fall vorerst keinen klaren Trend zu erkennen geben.

Als Fazit aus den Modellierungsexperimenten bei der Stahlblechprüfung kann abgeleitet werden, dass das Vorgehen der flächenabhängigen POD mit geglätteten Daten genutzt werden kann und einige wesentliche Vorteile im Gegensatz zu dem ein-parametrischen Ansatz besitzt, besonders wenn die bruchmechanisch kritische Defektfläche (-Breite) bekannt ist.

Modellierung der Schweißnahtprüfung eines Kupferkanisters

Analog zur vorhergegangenen Bewertung wurde für die Modellierung der Schweißnahtprüfung des Kupferbehälters das Programm "aRTist" verwendet. Dabei ergeben sich drei Punkte, die bei dem Vergleich zwischen Modellierungsergebnis und wahrer Prüfung beachtet werden müssen:

1. Prüfaufbau (z.B. Vergrößerung, Drehung)
2. Strahlenspektrum (z.B. verwendete Energien)
3. Kanistergeometrie und Prüfbereich (z.B. durchstrahlte Wanddicken)

Die Drehung des Kanisters im Strahlengang und die Aufnahme mit einem Liniendetektor wird innerhalb der Modellierungssoftware in zusätzlichen Arbeitsschritten verwirklicht (wie die Aneinanderreihung der einzelnen Linienaufnahmen). Diese sind notwendig, da durch eine Vereinfachung von Geometrien oder Prüfaufbauten die Ergebnisse (z.B. Streustrahlung, Entfernung zwischen Prüfobjekt und Detektor) verfälscht werden können. Für die Modellierung wurde hierbei zusätzlich das aRTist-Modul Tomosynth genutzt.

Der Einsatz einer Modellierungssoftware für den hochenergetischen Bereich birgt einige Hindernisse: So ist die Anwendbarkeit in höheren Energiebereichen ($>2\text{ MeV}$) seltener in Gebrauch. Dadurch wäre eine zusätzliche Verifizierung mit Experimenten sinnvoll. Ein Grund hierfür liegt in den Wechselwirkungsprozessen im hohen Energiebereich für die eine korrekte Erfassung aufwendig ist (z.B. wegen des Paarbildungseffekt). Außerdem ist die Bestimmung des Strahlenspektrums für einen 9 MeV -Linearbeschleuniger eine herausfordernde Aufgabe [180], [113]. Das Spektrum kann nicht gemessen werden. Letztendlich kann das Strahlenspektrum nur abgeschätzt werden, wodurch es zu möglichen Abweichungen zu den wahren Werten kommen kann. Des Weiteren handelt es sich bei der Schweißnaht des Kupferbehälters um eine komplexe Form, so dass sich die durchstrahlte Materialdicke in mehreren Bereichen ändert. Aus diesem Grund wird die Modellierungssoftware vorerst nur in dem Bereich eingesetzt, für den die Materialdicke konstant ist (siehe Abb. 4.11 a). Als Bewertungsgrundlage dienen ebenfalls Pseudo-Schliffe (Auflösung $0,01\text{ mm}$), die modellierten Anzeigen, die jeweiligen Profile und die Glättungsergebnisse (z.B. siehe Abb. 4.11 b).⁵ Es wurden hierbei 108 Werte mit Defektkantenlängen von $0,38\text{ mm}$ bis 5 mm betrachtet.

Trotz der aufgezählten Unterschiede zwischen Modellierungsergebnissen und wahren Messwerten, kann hier das Vorgehen an Hand des Falls einer experimentnahen Modellierung überprüft und eine grobe Abschätzung ermittelt werden. Ähnlich wie im vorherigen Beispiel wird der Glättungsparameter in Abhängigkeit von der Breite des Defekts ausgewählt. Im Folgenden wurden zwei Glättungsparameter beispielhaft ausgewählt ($h = 2\text{ px}$ und $h = 8,50\text{ px}$) (siehe Abb. 4.12 und 4.13).

⁵Vorerst wurde die Entwicklung der POD im konstanten Wanddickenbereich durchgeführt. Die durchstrahlte Wanddicke hat einen Einfluss auf die POD, diese wird jedoch im Weiteren nicht vertieft, da im Späteren eine Änderung des Prüfaufbaus die hier beschriebene Abhängigkeit hinfällig werden lässt.

4.2 Ergebnisse zur Bestimmung der POD gemäß der entwickelten Methodik

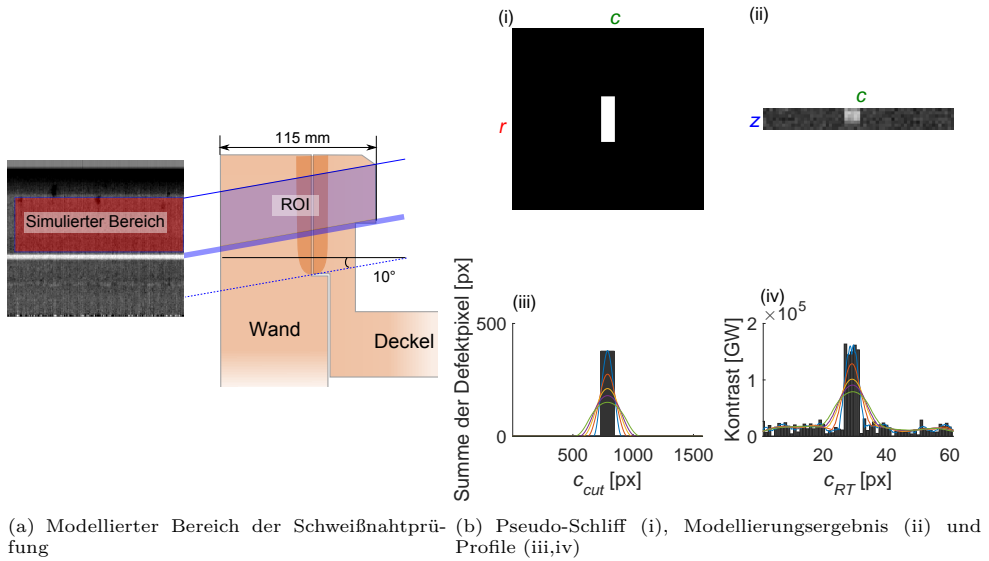


Abbildung 4.11: Modellierungsergebnisse für die Radiographie des Kupferbehälters zum Einsatz in der Endlagerung (Region of Interest im Kanisterquerschnitt - ROI)

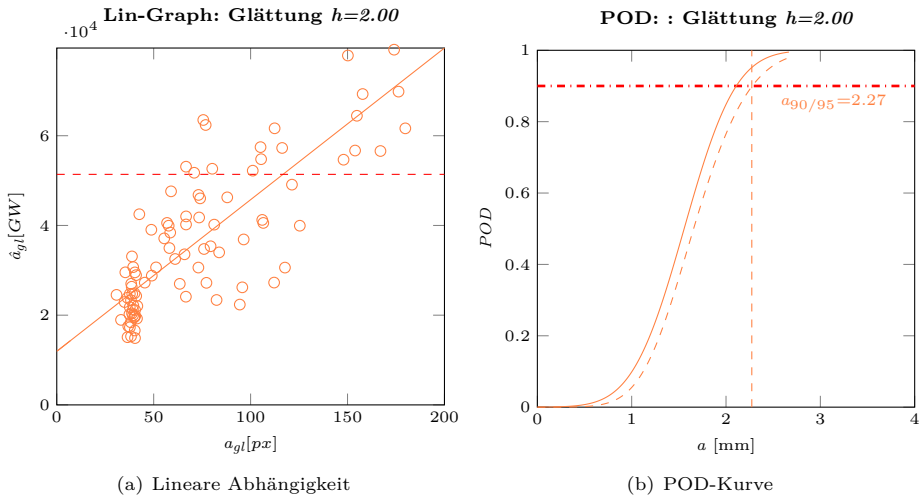


Abbildung 4.12: POD-Analyse mit geglätteten Daten der modellierten Kupfernahtprüfung (Glättungsparameter $h = 2$ px)

4.2 Ergebnisse zur Bestimmung der POD gemäß der entwickelten Methodik

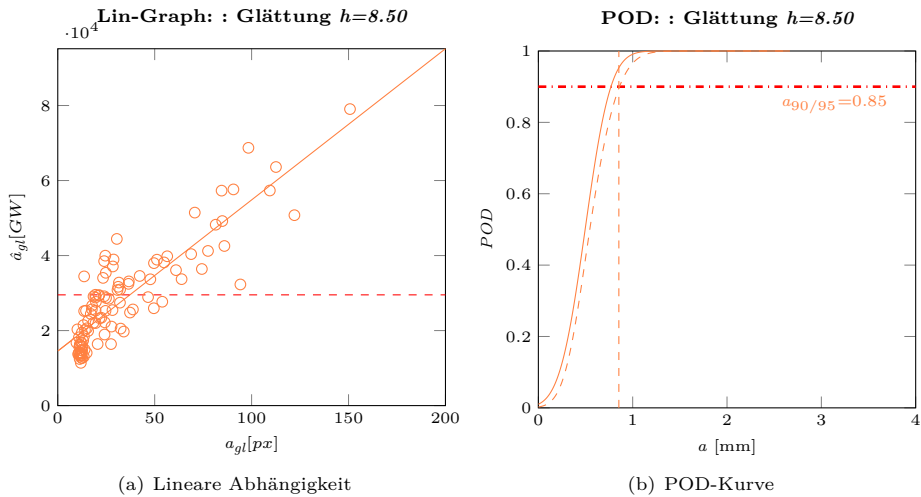


Abbildung 4.13: POD-Analyse mit geglätteten Daten der modellierten Kupfernahtprüfung (Glättungsparameter $h = 8,50$ px)

Besonderes Augenmerk soll hier auf folgende Tatsache gelegt werden: Die Streuung der Daten und die Linearität scheint aus visueller Sicht bei sehr kleinen Daten nicht mehr zu stimmen. Die Werte liegen eigentlich unter der Rauschgrenze, können daher nicht genau bestimmt werden. Es kann für diese Werte somit nur das Rauschen gemessen werden und nicht das tatsächliche Signal (zensierte Daten) [45]. Die Linearität scheint an diesem Punkt gestört. Daher ist der Einsatz der Maximum-Likelihood-Schätzung notwendig, da die Signalwerte auf ihren wahrscheinlichsten Wert geschätzt werden. Unabhängig von diesem Fall streuen die Werte aber symmetrisch um die Gerade. Die Entscheidungsschwelle wird jeweils auf $CNR = 2.5$ festgelegt (vgl. [186]). Im Weiteren wird der Kontrast in Grauwerten (GW) ausgedrückt und somit liegt die Schwelle zwischen $\hat{a}_{thr}(h_1) = 51\,000$ GW und $\hat{a}_{thr}(h_5) = 30\,000$ GW. Sie ist abhängig von den Werten die gerade noch detektiert werden sollen. Außerdem wird sie auch von der Fehlalarmrate, die minimiert werden soll, und somit von der Streuung um das Modell beeinflusst. Auch hier zeigen die Ergebnisse, dass größere Anzeigenflächen leichter zu detektieren waren als kleinere (siehe $a_{90/95}$ -Werten in Abbildungen 4.12 und 4.13), jedoch geht mit fallendem Kontrast die Anzeige auch schnell im Rauschen unter. Die Kurve ist steiler. Dies wurde auch schon bei der Observer-POD beobachtet. Auch andere Glättungskurven für die modellierten Werte unterstützen den oben beschriebenen Zusammenhänge (siehe Tab. 4.1).

Die Ergebnisse der Modellierungsprogramme spiegeln nur teilweise die experimentellen Daten wider, wodurch eine Überführung nur beschränkt möglich ist. Dies liegt in den Ein-

4.2 Ergebnisse zur Bestimmung der POD gemäß der entwickelten Methodik

Tabelle 4.1: Zusammenfassung der Glättungskurven für modellierte Werte: Im Vergleich Glättungsparameter h , \hat{a}_{dec} und $a_{90/95}$

| h | \hat{a}_{dec} | $a_{90/95}$ |
|---------|-----------------|-------------|
| 2,00 px | 51 390 GW | 2,27 mm |
| 4,20 px | 35 917 GW | 1,37 mm |
| 6,00 px | 33 036 GW | 1,14 mm |
| 7,00 px | 31 652 GW | 1,00 mm |
| 8,50 px | 29 522 GW | 0,85 mm |

schränkungen, die für die Modellierung bzgl. der Schweißnaht und dem Prüfaufbau getroffen werden müssen. Trotz der Einschränkungen zeigt sich eine klare Linearität zwischen den Signalwerten und den Defektparametern für die geglätteten Profile.

Bei der Untersuchung der Detektionsschwelle tritt eine weitere Einschränkung auf. Während bei Experimenten für deren Detektionsschwelle die Defekte an einer bestimmten Stelle erwartet wurden, ist eine Überbewertung der Detektierbarkeit möglich. Im realen Prüfalltag wäre die Detektierbarkeit somit etwas niedriger. Daher sollte die Schwelle später durch eine erhöhte Anzahl von Prüfern in Blindversuchen objektiver bestimmt werden, um die Daten über eine Bewertung der Durchführbarkeit hinaus zu nutzen. Mit Hilfe eines Sicherheitsaufschlags kann die vorliegende Detektionsschwelle der Praxis angenähert werden.

4.2.2 POD der Prüfung einer EB-Kupferschweißnaht

Für die Bewertung des tatsächlichen RT-Systems stehen künstliche und realistische Defekte zur Verfügung, die mit Absicht hervorgerufen wurden. Zuerst werden die künstlichen Defekte behandelt: Um den Nutzen des erweiterten Verfahrens zu veranschaulichen, wird wiederum eine ein-parametrische Bewertung als Vergleich durchgeführt. Im weiteren Schritt wird die Bewertung mit dem Konzept der Glättung erweitert. Dadurch kann die Durchführbarkeit der Methodik bewertet werden. Für den Einsatz bei realen Defekten werden die Schlibilder den RT-Anzeigen gegenübergestellt. Da es sich um reale Schriffe handelt, entstehen weitere Herausforderung. Die realen Defekte spielen eine wichtige Rolle, da sie bzgl. des Detektionsverhaltens den zukünftigen Schweißdefekten ähneln. Die Durchführung der erweiterten POD stellt daher das eigentliche Einsatzgebiet der neuen Methodik dar. Die experimentellen Daten sind ein wichtiger Bestandteil der POD, unabhängig ob die Defekte künstlich eingebracht wurden oder während des Schweißprozesses zufällig entstanden sind. Bei der Modellierung können nicht alle relevanten Prüfparameter in das Modell eingetragen werden, da es zum Einen nicht möglich ist und zum Anderen teilweise wichtige Parameter unbekannt sein können. Die experimentellen Daten der Referenzdefekte und der realen Defekte können ihrerseits nicht einfach gemischt werden, da Referenzdefekte und reale De-

fekte für die POD unterschiedlich gewichtet werden. Deshalb werden sie vorerst getrennt diskutiert.

Künstliche Referenzdefekte in der Schweißnaht

Künstliche Defekte zeichnen sich durch ihre bekannten Abmessungen und ihre deutliche Anzeige im RT-Bild aus (siehe Abb. 4.6). Dadurch ist es nicht unbedingt notwendig, die Defekte metallographisch oder mit anderen Referenzmethoden zu überprüfen und zu bemaßen, wenn angenommen werden kann, dass der Prozess, mit dem die Defekte erstellt wurden, unter Kontrolle ist. Die Auflösung des Pseudoschliffbilds beträgt 0,01 mm (analog zur Auflösung der wahren Schliffbilder von realistischen Defekten). Im Gegensatz dazu ist die Auflösung der RT-Anzeige 0,35 mm. Ein Beispiel für die Gegenüberstellung eines Pseudoschliffbilds mit der RT-Anzeige ist in Abbildung 4.14 zu sehen.

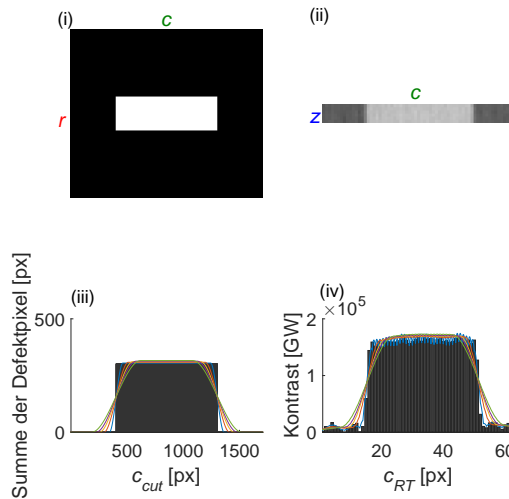


Abbildung 4.14: Pseudo-Schliff (i), modellierte RT-Anzeige (ii), Profile mit Glättungskurven (iii,iv) für einen künstlichen Defekt im Referenzsegment

Während die Bandbreite für h bei RT-Anzeigen sich im Bereich von einigen Pixelgrößen (2 - 8,50 px) lag, wurde die Bandbreite h für das metallographische Profil auf 60 - 220 px festgelegt.

Es wird zuerst die konventionelle Signal-Antwort POD als Bewertungsverfahren herangezogen. Bei der Gegenüberstellung des höchsten Kontrastes mit der größten durchstrahlten Länge kann folgender Trend erkannt werden: größere Defekte weisen einen höheren Kontrast

4.2 Ergebnisse zur Bestimmung der POD gemäß der entwickelten Methodik

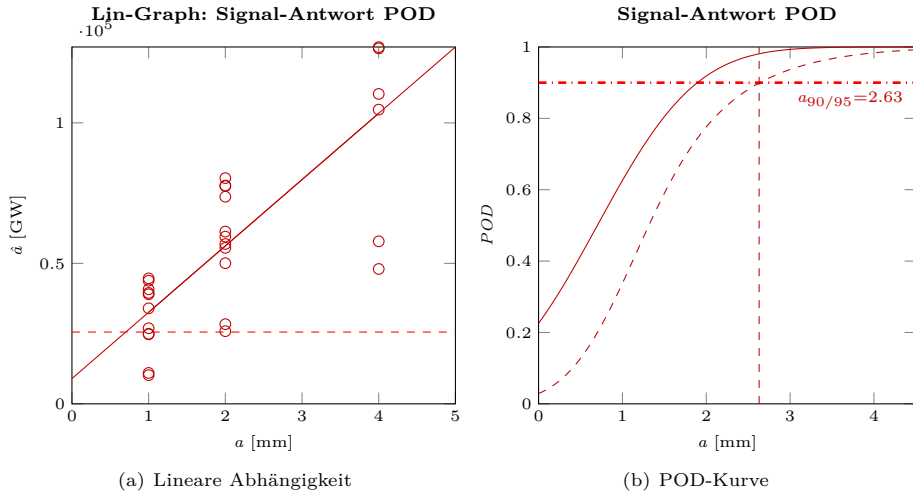


Abbildung 4.15: Konventionelle POD-Bewertung mit Hilfe von künstlichen Defektdaten

auf (siehe Abb. 4.15 a). Bei einer einfachen linearen Regression zeigt sich jedoch, dass die Streuung sowohl groß (R^2 liegt bei 0,71), als auch inhomogen ist. Außerdem ist die Streuung nicht symmetrisch. Daher sollte die Annahme der normal-verteilten Streuung in Frage gestellt werden.

Sowohl bei inhomogenem und unsymmetrischem Streuungsverhalten als auch bei einer breiten Datenstreuung sollte der konventionelle Ansatz zur Berechnung einer POD abgelehnt werden. Es muss dabei eigentlich von einer deutlich höheren Detektionsschwelle ausgegangen werden, da ansonsten aus Sicht der POD eine hohe Fehlalarmrate zu erwarten ist (siehe Schnittpunkt der POD-Kurve mit der Ordinate) [187]. Des Weiteren ist die Kenngröße $a_{90/95}$, trotz einer hohen zu erwarteten Fehlalarmrate, über 2,00 mm. Bei der visuellen Betrachtung heben sich 2,00 mm Defekte jedoch deutlich von der Umgebung ab. Dies führt zu einer Unterbewertung des Verfahrens. Zudem kann die nicht symmetrische und große Streuung auch als Forderung nach mehr-parametrischen Ansätzen gesehen werden.

Obwohl die Daten die Forderungen (siehe [81]) nicht ausreichend erfüllen, soll dennoch eine konventionelle POD-Kurve für den späteren Vergleich berechnet werden. Die Entscheidungsschwelle bei 26 000 GW wurde hierbei auf den Signalwert gelegt, der gut erkannt werden konnte. Was hierbei auffällt ist der Schnittpunkt mit der Ordinate der nicht durch den Nullpunkt geht. Aus Sicht der Praxis müsste eine POD-Kurve für keinen vorhandenen Defekt

4.2 Ergebnisse zur Bestimmung der POD gemäß der entwickelten Methodik

keine Auffindwahrscheinlichkeit besitzen, somit eine POD von 0 haben. Dabei gelten vor allem folgende Beobachten aus der Realität:

$$\lim_{a \rightarrow 0} \hat{a} = 0 \quad (4.1)$$

und

$$\lim_{a \rightarrow 0} \sigma_{\hat{a}} = 0. \quad (4.2)$$

Allerdings wird in dem mathematischen Modell die Annahme der Homoskedastizität eingeführt, wodurch gilt:

$$\sigma_{\hat{a}}(a) = \textit{konstant}, \forall a \quad (4.3)$$

(vgl. auch Gleichung 2.8). Es wird daher ersichtlich, dass durch die Annahme einer Normalverteilung mit einer konstanten Varianz, die $POD(0)$ einen Wert besitzt. Dies verstärkt sich weiter, wenn der \hat{a} vs. a Graph nicht durch den Ursprung verläuft (siehe Abb. 4.16). Hierbei gibt es die Möglichkeit die Gerade durch den Ursprung zu zwingen und somit gleichzeitig von der am Besten geschätzten Kurve abzuweichen. Dass die Gerade nicht durch den Ursprung verläuft, ist ein praktisches Problem: Bei den Experimenten kann der Kontrast nicht optimal berechnet werden kann, da sich der Hintergrund teilweise auch innerhalb eines Bildes verändert (vgl. Abb. 4.6). Es kann somit gesagt werden, dass der Verlauf der POD -Kurve an der Stelle $a = 0$ eine Schwäche des mathematischen Modells darstellt, das bei der Bewertung von Prüfungssystemen teilweise in Kauf genommen oder sogar mit Absicht herbeigeführt wird (vgl. [187]).

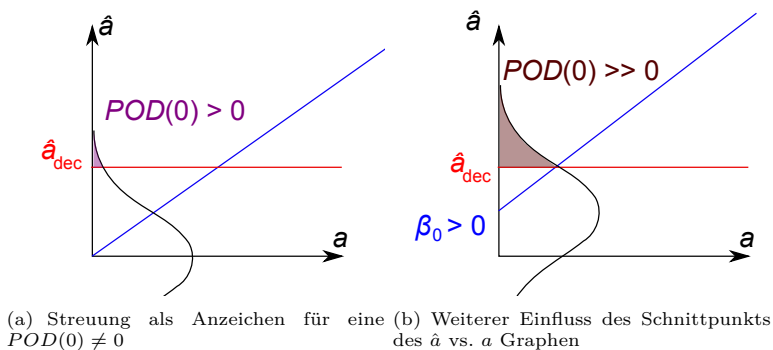


Abbildung 4.16: Modellannahmen, die einen Einfluss auf den Schnittpunkt von POD an der Stelle 0 aufweisen

4.2 Ergebnisse zur Bestimmung der POD gemäß der entwickelten Methodik

Für die erweiterte POD wird nun die Defektbreite hinzugezogen, wobei die Glättung analog zu den Modellierungsergebnissen durchgeführt wird. Zur Bestimmung der POD stehen 33 künstliche Defekte zur Verfügung. Die lineare Abhängigkeit ist im \hat{a} vs. a -Graphen und die daraus entstehende POD-Kurve für zwei Bandbreiten ($h = 8,50$ px und $h = 4,20$ px) in Abb. 4.18 und Abb. 4.17 zu sehen. Alle künstlichen Defekte wurden detektiert. Deshalb könnte die Detektionsschwelle auf das Signal eines kleinen Defekts festgelegt werden. Jedoch spielt die Streuung auch weiterhin eine Rolle insbesondere bei der Wahl der Entscheidungsschwelle (wegen der Fehlalarmrate). Die Detektionsschwelle wird ebenfalls auf $CNR = 2.5$ entsprechend zwischen 23 000 GW ($h = 2$ px) und 16 000 GW ($h = 8,50$ px) festgelegt (vgl. [186]).

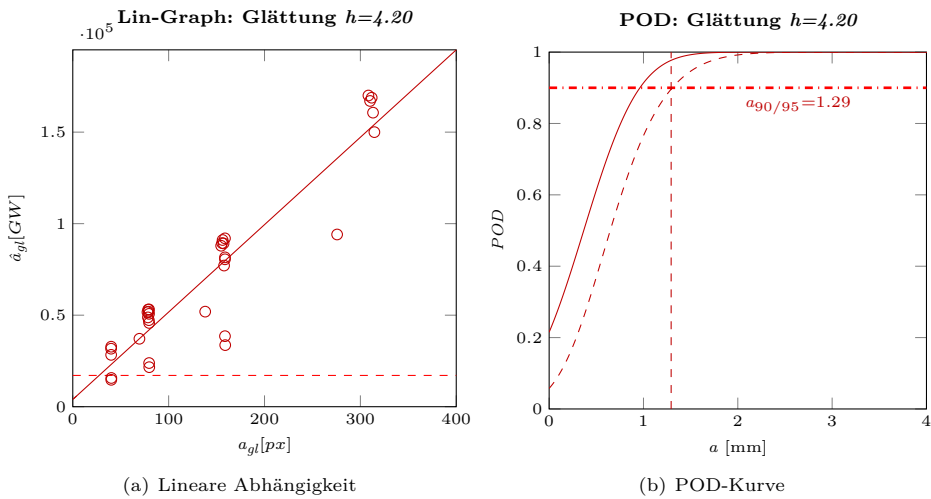


Abbildung 4.17: POD-Analyse mit geglätteten Daten (Glättungsparameter $h = 4,20$ px)

Bei dem Vergleich zwischen dem konventionellen und dem erweiterten Ansatz wird ersichtlich, dass die Glättung Einfluss auf das Signal und somit auf den linearen Zusammenhang hat. Es standen nur drei verschiedene maximal durchstrahlte Längen von Defekten zur Verfügung. Daher ist die Streuung für das konventionelle Modell an diesen drei Stellen von a für \hat{a} sehr ausgeprägt. Im Fall der Glättung fließen auch die Defektbreiten in die Bewertung ein, wodurch die Werte stärker auf der Abszisse verteilt werden. Die zusätzliche Bewertung eines weiteren Defektparameters führt gleichzeitig zu der gewünschten Reduktion der Streuung.

Kleine Glättungsparameter ($h < 2$ px) oder große Anzeigenflächen zeigen dieses Verhalten nicht, da die Anzeigenfläche keine Einfluss hat. Hieraus lässt sich schließen, dass die konventionelle POD dann verwendet werden kann, wenn die Defekte eine sehr große Anzeigenfläche haben (im Vergleich zum Glättungsparameter). In diesem Fall wäre dann die

4.2 Ergebnisse zur Bestimmung der POD gemäß der entwickelten Methodik

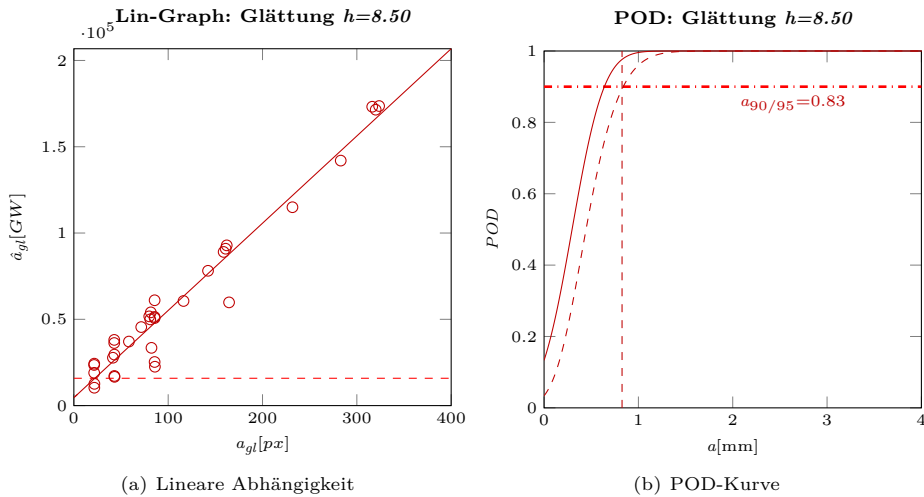


Abbildung 4.18: POD-Analyse mit geglätteten Daten (Glättungsparameter $h = 8,50$ px)

Amplitude ausschlaggebend. In der Observer-POD läuft die Schwelle in diesem Bereich auf einen konstanten Wert zu.

Der Graph für eine geringere Bandbreite zeigt, dass der Schnitt mit der Ordinate immer noch relativ hoch liegt, jedoch die POD für höhere Werte sich verändert hat. Weshalb die Glättung vorerst die großen Datenwerte beeinflusst, ist dahingegen zu erklären, dass das Rauschen reduziert wird; das heißt, dass auch die Schwelle niedriger angenommen werden kann. Gleichzeitig ist die Streuung der Daten noch nicht genug geschwächt, um die POD gegen den Ursprung zu zwingen. Diese Kenntnisse spiegeln sich auch in der Zusammenfassung für die Glättung für künstliche Defekte wider (siehe Tab. 4.2).

Tabelle 4.2: Zusammenfassung der Glättungskurven für künstliche Defekte: Im Vergleich Glättungsparameter h , \hat{a}_{dec} und $a_{90/95}$

| h | \hat{a}_{dec} | $a_{90/95}$ |
|---------|-----------------|-------------|
| 0,00 px | 25 550 GW | 2,63 mm |
| 2,00 px | 22 703 GW | 1,80 mm |
| 4,20 px | 17 087 GW | 1,29 mm |
| 6,00 px | 16 202 GW | 1,09 mm |
| 7,00 px | 16 035 GW | 0,95 mm |
| 8,50 px | 15 841 GW | 0,83 mm |

Mit Hilfe des erweiterten Ansatzes wird die Streuung deutlich reduziert und besitzt ein homogeneres Verhalten über die durchstrahlte Länge. Gleichzeitig übt die Glättung in Abhängigkeit von der Anzeigengröße einen Einfluss auf das Signal aus. Dennoch fallen einige

4.2 Ergebnisse zur Bestimmung der POD gemäß der entwickelten Methodik

Datenwerte außergewöhnlich niedrig für das Modell aus. Untersuchungen zeigten, dass eine bestimmte Form des Defekts dafür ausschlaggebend ist. Schmale Nuten werden nicht optimal von dem aufgestellten Modell erfasst. Ein Grund dafür ist die eingeschränkte Höhe der RT-Anzeige, die die Anzahl der gegebenen Pixel beeinflusst. In der Praxis sind schmale Nuten die Referenzdefekte für rissähnliche Anzeigen. Risse werden jedoch aus Sicht der Detektionierbarkeit anders als voluminöse Defekte behandelt [104], weshalb das vorgestellte Vorgehen vorerst nur im Fall von voluminösen Defekten eingesetzt werden soll.

Die Reduzierung der Streuung zeigt sich auch am Vertrauensbereich. Da die Streuung durch den neuen Ansatz deutlich besser mit Hilfe des mathematischen Modells (homoskedastische Normalverteilung) beschrieben werden kann, wird den Modellparametern auch ein höheres Vertrauen entgegengebracht. Wird die Bandbreite der Glättung erhöht, so bewirkt dies, dass die POD steiler wird. Zudem spiegelt es die Tatsache wider, dass Defekte mit einer größeren Anzeigenfläche auch noch bei geringerem Kontrast gefunden werden können, kleinere Defekte jedoch nicht. Die geringere Bandbreite bezieht sich jedoch auf ein Modell, bei dem mit einer hohen Fehlalarmrate zu rechnen ist (siehe Schnittpunkt mit Ordinate).

Die Untersuchung der Auswirkung der Bandbreite auf die Anzeige stellt sich bei der geringen Anzahl von Defekten als schwierig dar. Daher soll die Methodik auf die Ergebnisse der Modellierungsprogramme angepasst werden. Die Abhängigkeit zwischen Anzeigenfläche und Bandbreite (siehe Abschnitt 3.2) kann als linear angenommen werden, wenn nur voluminöse Defekte betrachtet und Formaspekte vernachlässigt werden. Dadurch reduziert sich die funktionelle Abhängigkeit der Detektionsschwelle von der Anzeigenfläche auf die Abhängigkeit von ihrer Breite, die in dieser Bewertung Ausschlag gebend ist. Unter der Annahme, dass die Modellierungsexperimente, teilweise auf die realen Messungen übertragen werden können und unter der zweiten Annahme, dass die hier berechnete erweiterte POD korrekt ist, lassen beide Modelle einen Zusammenhang zwischen Bandbreite und detektierter Anzeigenbreite zu (siehe Anhang VI). Für $h = 2$ px ergibt sich eine $POD = 50\%$ bei einem $a = 0,44$ mm. Bei der Observerschwelle entspricht dies einer Anzeigenfläche von $0,50$ mm² ($l(\text{FBSQ}) = 0,70$ mm) [133]. Im Gegensatz dazu liegt $h = 8,50$ px bei einer Anzeigenfläche von $2,40$ mm² ($l(\text{FBSQ}) = 1,50$ mm). Damit können nun unter den genannten Voraussetzungen die angegebenen Graphen unter Berücksichtigung der Bandbreite der Anzeigenfläche bzw. -breite zugeordnet werden.

Als Zusammenfassung kann der Einsatz einer Bewertung des RT-Systems, die von der Anzeigenfläche abhängt, mit experimentellen Daten als sinnvoll und notwendig für die korrekte Bewertung betrachtet werden.

Reale Defekte in der EB-Naht

Der Herstellungsprozess von realistischen Defekten ist kompliziert. So ist es fast unmöglich, bestimmte Defekte mit vorgegebenen Geometrien herzustellen. Außerdem werden nur

bestimmte Defekte untersucht, da die Kosten hoch sind. Es liegen auch nur diese als Bewertungsgrundlage vor, während nicht gefundene Defekte nicht bewertet werden. Durch die verschiedenen Prüfverfahren ist jedoch die Wahrscheinlichkeit, einen kritischen Defekt zu übersehen, gering (vgl. Abschnitt 4.1). Für die Bewertung der realen Defekte werden die wahren Abmessungen des Defekts durch die Metallographie bestimmt. Diese Schlifflinformationen werden mit der RT-Anzeige verglichen.

Die Abstände zwischen den einzelnen Schliffl ermöglichen es nicht die durchstrahlten Defektlängen genau den RT-Grauwerte zu zuordnen. Außerdem wird dies durch die Pixelauflösung der Radiographie erschwert. Die wirkliche Position des Defekts ist zudem nicht exakt bestimmbar, da schon geringe Schwankungen in der Rotationsgeschwindigkeit oder eine leichte Exzentrizität des Kanisters zu ungenauen Positionen für die Anzeigen führen können. Gleichzeitig kann nicht von einem defektfreien Bereich, wie für Modellierungsergebnisse oder künstliche Defektbewertungen, ausgegangen werden. Der defektfreie Bereich muss vielmehr mit Schliffl verifiziert werden. Es werden daher jedem Schliff 6 px der RT-Anzeige in Axialrichtung zugeordnet (RT-Pixelgröße: 0,40 mm) (Abb. 4.19). Zudem zeigt das RT-Signal einen Offset, da auch der Hintergrund ein Signal aufnimmt. Um die Profile des Schliffls und der RT vergleichbar zu machen, wird ein durchschnittliches Rauschen für das RT-Bild bestimmt und auf den 0 GW-Wert gesetzt. Es wird somit der Kontrast zum durchschnittlichen Hintergrund betrachtet. Bei der Schätzung des Kontrasts und des Offsets sind Schätzfehler möglich, sodass der \hat{a} vs. a Graph nicht exakt den Ursprung trifft.

Für die Breite des Profils liegt dem Schliff eine Auflösung von 0,01 mm vor. Zur Gegenüberstellung ist daher Binning oder ein Glättungsalgorithmus nötig, um die Daten vergleichbar zu machen.

Bei der metallographischen Untersuchung werden teilweise auch Defekte hervorrufen, die zum Zeitpunkt der Prüfung nicht vorhanden waren. Beispiele hierfür sind Kratzer an der Oberfläche durch den Bearbeitungsprozess oder Markierungen, die zu einem Signal im Profil führen (z.B. D8-S1 und D2-S1).⁶ Diese Anzeigen sollen aus der POD-Bewertung des ZfP-System ausgeschlossen werden, da sie sonst das Ergebnis verfälschen.

Ein weiteres Problem bei der Gegenüberstellung sind die Einflüsse durch geometrische Ungenauigkeiten und die Rotationsungenauigkeit des Prüfobjekts. Hierbei kann ein Defekt zwar eine Anzeige liefern, jedoch auf dem zugeordneten Schliff nicht zu sehen sein: Liegt das Ende (bzw. der Anfang) eines Defekt genau zwischen den 6 px, so ist dieser deutlich im Schliff zu sehen, besitzt jedoch durch den gemittelten Grauwerte nur eine kleine RT-Anzeige. Diese besitzen ein kleines \hat{a} oder können im Rauschen untergehen. Aber auch eine Erhöhung des Grauwerts kann die Folge sein, wenn die maximale Länge nicht durch den Schliff getroffen wurde. Dies ist ein Einfluss der Auswertung und gibt nicht die Fähigkeit des

⁶D2 bezeichnen Schliffl, die an defektfreien Stelle erstellt wurden (Gespräch mit Pitkänen).

4.2 Ergebnisse zur Bestimmung der POD gemäß der entwickelten Methodik

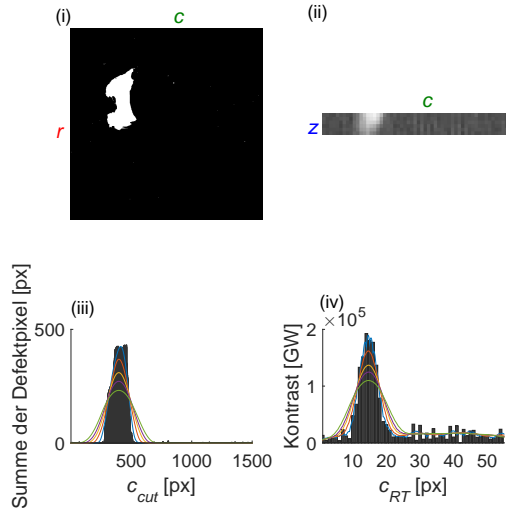


Abbildung 4.19: Beispiel eines realen Defektes mit Schlifffdaten (i), RT-Anzeige (ii) und jeweiligen Profilen (iii,iv)

Systems wieder. Daher werden die aufgelisteten Daten im Weiteren als Ausreißer behandelt. Sie fließen nicht in die Bewertung mit ein.

Die Einflüsse der Metallographie auf die Auswertung, der Unterschied zwischen Schlift und radiographischer Anzeige und weitere unbekannte Einflüsse geben eine Untergrenze für die Genauigkeit der Auswertung. Diese Auswertungsgrenze in radialer Ebene befindet sich für die metallographische Anzeige bei 100 px (ca. 1,33 mm). Außerdem zeigt sich hier das probabilistische Verhalten eines ZfP-Systems für reale Defekte. Während manche Defekte in diesem Bereich deutlich zu erkennen sind, gehen andere Defekte mit vergleichbarer Größe im Rauschen unter. In einer zukünftigen räumlichen Bewertung werden diese vermieden. Dies ist die Motivation für die Entwicklung eines räumlichen Ansatzes.

Ähnlich den vorhergegangenen Analysen, zeigt sich auch hier der Vorteil der Glättung. Es wurden 108 Defekte in den Schliffen erfasst, die als Bewertungsgrundlage dienen.⁷ Es ist klar ersichtlich, dass es sich bei den realen Defekten um sehr kleine Anzeigen handelt, da viele im Rauschen untergehen. Als Entscheidungsschwelle wurde basierend auf dem Rauschen ($CNR = 2.5$) [186] ein Schwellwert zwischen 65 000 GW und 43 000 GW festgelegt.

⁷Eine Vorauswahl und Beschränkung auf relevante Defektgrößen ergab 43 Defekte.

4.2 Ergebnisse zur Bestimmung der POD gemäß der entwickelten Methodik

Als Beispiel werden für die geglätteten Defektdaten POD-Kurven mit der durchstrahlten Länge als a für zwei Bandbreiten ($h = 2 \text{ px}$, $h = 8,50 \text{ px}$) in Abbildung 4.20 gezeigt.

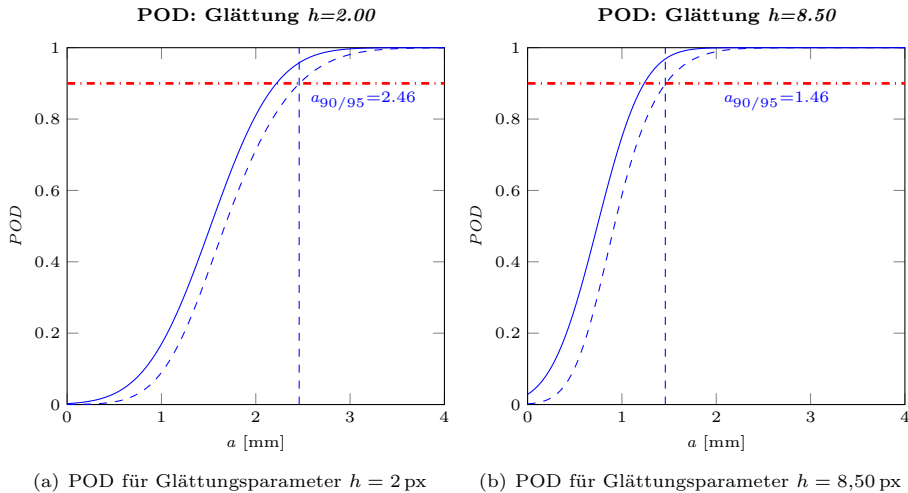


Abbildung 4.20: POD-Kurve bezogen auf den jeweiligen Glättungsparameter

Wenn die Linearität und die Streuung der Daten effektiv geschätzt werden und eine hohe Anzahl von Daten verwendet wird, bildet sich ein schmaler Vertrauensbereich aus. Gleichzeitig zeigt sich bei unterschiedlichen Glättungsparametern, dass der $a_{90/95}$ -Wert von 1,46 mm auf 2,46 mm steigen kann, falls die Anzeigenfläche geringer wird.

Bei den realen Defekten wird der eigentliche Nutzen des erweiterten Vorgehens ersichtlich: Die Abhängigkeit der Anzeigenfläche besitzt den größten Einfluss auf kleine Defekte. Diese kleinen Defekte stehen bei den realen Defekten im Vordergrund. Viele der Defekte sind klein und teilweise vielleicht unkritisch für den späteren Einsatz. Aus Sicht der Versuchsplanung jedoch bilden sie den Bereich, auf dem die POD-Berechnung aufgebaut werden soll, weil die Steigung der POD-Kurve in diesen Bereich fällt. Der Einsatz der Glättungsparameter zeigt, dass die Detektierbarkeit innerhalb einiger Millimeter (durchstrahlten Länge) schwanken kann (Tab. 4.3).⁸

Die Umrechnung vom Glättungsparameter zur Breite stellt sich für die realen Defekte als schwierig heraus, da sich durch die Rotation und durch die weiteren Prüfeinflüsse der Hintergrund stark ändert. Daher ist in diesem Fall nur eine ungefähre Abschätzung möglich. Bei einer Bandbreite von $h = 2 \text{ px}$ liegen viele Werte knapp unter der Schwelle, die eine Breite von 1,00 mm aufweist. Daher kann eine Parallele zu dem Zusammenhang mit

⁸Wird die Detektionsbereich mit der Wanddicke (ca. 100 mm) verglichen, wird die Fähigkeit des ZfP-System ersichtlich.

4.2 Ergebnisse zur Bestimmung der POD gemäß der entwickelten Methodik

Tabelle 4.3: Zusammenfassung der Glättungskurven für reale Defekte: Im Vergleich Glättungsparameter h , \hat{a}_{dec} und $a_{90/95}$

| h | \hat{a}_{dec} | $a_{90/95}$ |
|---------|-----------------|-------------|
| 2,00 px | 64 834 GW | 2,46 mm |
| 4,20 px | 46 399 GW | 1,83 mm |
| 6,00 px | 45 953 GW | 1,73 mm |
| 7,00 px | 46 414 GW | 1,67 mm |
| 8,50 px | 42 632 GW | 1,46 mm |

künstlichen Defekten gezogen werden. Analog zu den künstlichen Defekten wird der Bandbreite $h = 2$ px einer Anzeigenfläche von $0,50 \text{ mm}^2$ und $h = 8,50$ px bei einer Anzeigenfläche von $2,40 \text{ mm}^2$ zugeordnet. Dies gilt ausdrücklich nur für ein ähnliches Detektionsverhalten von realen und künstlichen Defekten, dem korrekten Modell für die erweiterte POD und fundierten Modellierungsergebnissen.

Es zeigt sich, dass die Methodik für reale Defektdaten einsetzbar ist und mehr Informationen liefert als der konventionelle Ansatz. Dies gilt besonders für den Fall, wenn der $a_{90/95}$ -Wert nahe an der Akzeptanzschwelle der Bruchmechanik liegt. Im Fall des vorliegenden Projekts ist die Schwelle für die Akzeptanz eines Kanisters noch weit entfernt von dem hier bestimmten $a_{90/95}$. Das RT-Verfahren ist somit auf Basis der POD einsetzbar. Die Verbindung der Bandbreite zur Defektbreite konnte mit Hilfe von Modellierungsergebnissen abgeschätzt werden, sollte jedoch für den weiteren Einsatz verifiziert werden.

Abschließend soll hier noch einmal die Ergebnisse der Bewertungen in der Tabelle 4.4 zusammengestellt werden.

Tabelle 4.4: Zusammenfassung der Ergebnisse der Modellierungen, der künstliche Defekte und der realen Defekte im Vergleich Anzahl der Datenwert n und Resultat $a_{90/95}$

| Art der Defekte | n | $a_{90/95}$ |
|---------------------------|-----|-----------------|
| Modell. Stahlschweißnaht | 28 | 1,53 mm-1,62 mm |
| Modell. Kupferschweißnaht | 108 | 0,83 mm-2,27 mm |
| künstliche Defekte | 33 | 0,83 mm-2,63 mm |
| reale Defekt | 43 | 1,46 mm-2,46 mm |

Wie schon in der Übersicht zu sehen ist, ist die Anzahl der zur Verfügung stehenden Daten essentiell ist, für die Auswerten, soll im Weiteren der Fall von sehr wenigen Daten vertieft werden, der besonders bei realen Defekte auftreten kann.

4.2.3 Kombination von realen und künstlichen Defekten

Ein typisches Problem für die POD-Berechnung ist die geringe Anzahl von realen Defekten. Eine geringe Anzahl von Datenwerten bewirkt ein großes Vertrauensband. Gleichzeitig bilden die künstlichen Defekte nur bestimmte Informationen ab.

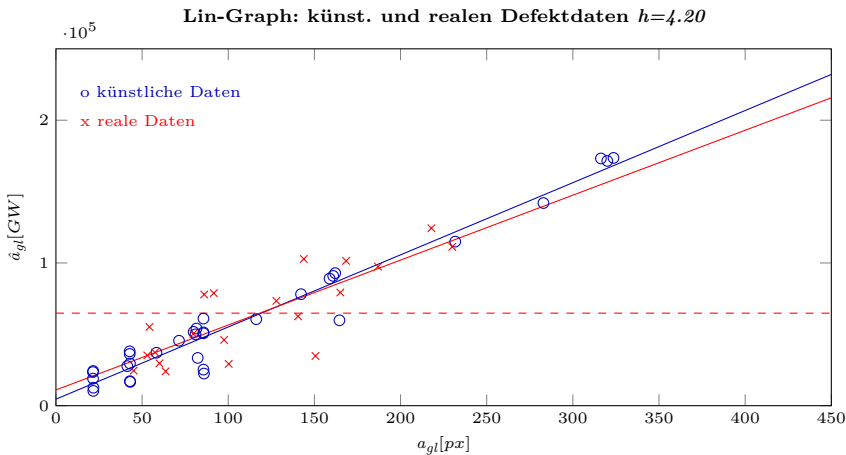


Abbildung 4.21: Signal-Antwort Graph für einige reale und künstliche Datenwerte aus der Bewertung mit einem Glättungsparameter

Für die Kombination von verschiedenen Datenwerten wurde eine Untergruppe von realen Defekten als Datenbasis ($n = 20$) herangezogen. Die realen Defekte liegen teils nahe an der Rauschschwelle, wodurch eine beliebige Reduzierung der Daten vorerst nicht möglich ist. Als Beispiel wurden die geglätteten Daten mit einer Bandbreite von $h = 4,20$ px herangezogen (Abb. 4.21). Die Gegenüberstellung der adäquaten Bandbreiten mit ihren zugehörigen Anzeigenflächen stellt im Moment noch eine Herausforderung dar, da sich die Prüfsituationen und die vorhandenen Datenmengen unterscheiden. Dennoch war es möglich, für den vorliegenden Fall eine Ähnlichkeit der Pixelgröße für künstliche und reale Defekte und daher auch der Bandbreite festzustellen, die eine mögliche Kombination nahelegt. In der weiteren Verwendung sollten diese Ähnlichkeiten mit Untersuchungen verifiziert werden.

Im Folgenden werden Kurve unterschiedlicher Anzahl von realen Defektdaten dargestellt (Abb. 4.22). Die Kurve für reale Defekte verläuft hier fast durch den Ursprung, was jedoch zufällig ist, wie an einer größeren Datenmenge zu sehen ist.

Bei dem Beispiel ist das größere Vertrauensband bei realen Defekten zu finden. Für die Kombination wurden die künstlichen Defekte mit ca. 60% auf Grund des geringeren Informationsgehalts gegenüber den realen Defekten gewichtet.⁹ Somit werden die künstliche

⁹Der Wert 60% wurden während einer Expertenrunde abgeschätzt und diskutiert.

4.2 Ergebnisse zur Bestimmung der POD gemäß der entwickelten Methodik

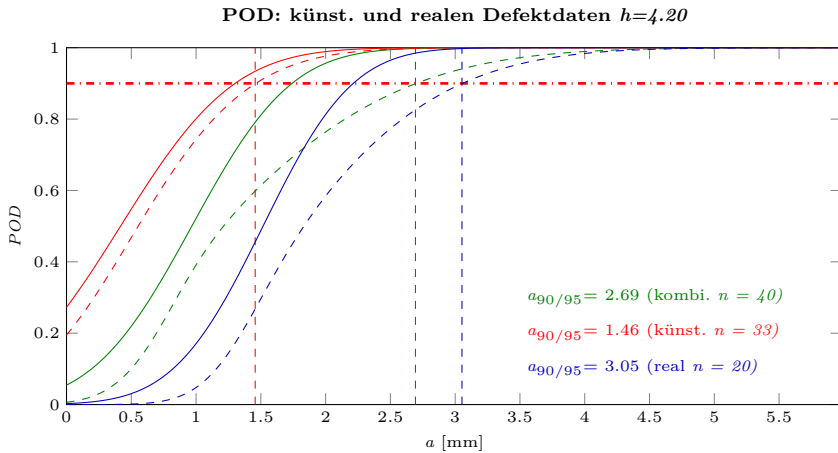


Abbildung 4.22: POD-Kurven für wenige reale, für künstliche und für die kombinierte Datenanzahl

Defektanzahl von 33 Datenwerte auf 20 Datenwerte verringert. Somit ergibt sich als Summe 40 Datenwerte. Die künstlichen Defekte dominieren nicht mehr das Modell, wie es bei einer einfachen Mischung passiert.

Sowohl die $a_{90/95}$ als auch das gesamte Vertrauensband zeigt eine Verringerung. Dies resultiert aus der Zunahme der Datenmenge durch die künstlichen Defektdate. Bei der Kombination nimmt die Anzahl zu und somit wird das Vertrauensband kleiner. Auch der Einfluss auf den $a_{90/95}$ -Wert zeigt eine bessere Einschätzung mit Hilfe der Kombination. Während die wenigen Datenwerte den Wert zu hoch einschätzen, liegt die Kombination näher an dem Wert der für eine größere Anzahl von Daten geschätzt wurde ($a_{90/95}(h = 4,20 \text{ px}) = 1,83 \text{ mm}$). Eine Beeinflussung der Gewichtung mit subjektiven Bewertungen wäre hier möglich, wurde jedoch vorerst nicht durchgeführt.

Die vorgestellte Kombination ermöglicht daher, zwei unterschiedliche Datenmengen so zu verbinden und damit eine ausgeglichene Datenmischung durchzuführen.

5 Zusammenfassung und Schlussfolgerung

Die POD stellt ein wichtiges Mittel zur Bewertung eines RT-Systems dar. Dabei spielen die realen Defekte eine wichtige Rolle in der Analyse, da sie am besten einen Einblick in die wirklichen Herstellungsbedingungen und in die Zuverlässigkeit des Bauteils erlauben. Doch für die Bewertung des Systems auf Basis von realen Defekten sind die konventionellen Ansätze (vgl. [15] und [81]) nicht mehr zweifelsfrei anwendbar. Es wurde die Entwicklung von einem ein-parametrischen Ansatz zu einem erweiterten Ansatz ausgeführt. Dieser ermöglichte den Einsatz der POD mit realen Defekten und außerdem eine Kombinationsmöglichkeit von Daten verschiedener Defekttypen.

Diese Arbeit stellte ein Vorgehen für den möglichen Umgang mit realen Defekten vor und zeigte die auftretenden Herausforderungen auf. Dabei waren die wesentlichen Punkte der neuen Methodik:

1. die Gegenüberstellung der metallographischen Schlitze zu den RT-Anzeigen
2. die Formulierung physikalischer Eigenschaften hinsichtlich des Vergleichs
3. die Einführung der Anzeigenfläche als einen weiteren Parameter für die Detektion von realen Defekten

Bei der RT-Prüfung der EBW-Schweißnaht des Kupferbehälters für die Endlagerung von Brennstäben aus Kernkraftwerken lag ein Prüfanordnung vor, für die ein zylindrisches Koordinatensystem sinnvoll für die Gegenüberstellung von RT-Anzeigen und Schlitfbildern war. Da im vorliegenden Fall die Entfernung zwischen den Schlitzen ca. 2 mm betrug und eine Rekonstruktion eines räumlichen Defektmodells nicht möglich war, musste die Methodik vorerst auf Schlitze beschränkt werden. Die Gegenüberstellung der Schlitze mit den Anzeigen war in wesentlicher Schritt bei dem Umgang mit realen Defekten und bildete die Grundlage für die weitere Bewertung.

Die Gegenüberstellung zeigte auch, dass die Einführung eines Parameters für die Anzeigenfläche im Fall der Bewertung eines RT-Systems notwendig war. Eine reine Beschränkung auf die konventionellen Verfahren hätte voraussichtlich in einem verfälschten Ergebnis resultiert. Die hieraus entstehende Überbewertung des Systems könnte für den späteren Einsatz schwerwiegende Folgen mit sich führen. Daher wurden die Anzeigengröße und die physikalischen Eigenschaften durch adäquate Glättungsparameter bei der Bewertung in das Modell

5. Zusammenfassung und Schlussfolgerung

eingeführt. Die Glättung wurde sowohl für die Schlifffdaten als auch für die Anzeigen angewandt. Der Einsatz verschiedener Glättungsparameter resultierte in einem multi-skalen Ansatz. Die multi-skale Glättung wurde zur Bewertung von verschiedenen Informationen genutzt und ermöglicht somit die Überführung von Schliffinformationen und Anzeigendaten in ein gemeinsames System der POD. Dies stellte ein neues Einsatzgebiet für die Glättung dar. Auch wenn diese zusätzlich für das Entfernen von Rauschen (Wiederherstellen von Strukturen) eingesetzt wurde, so könnten auch weitere physikalische Eigenschaften des Systems mit Hilfe der Glättung eingeführt werden: Dabei ist die Modellierung der Unschärfe des Systems oder die Einführung der Detektierbarkeit von Anzeigen zu nennen. Während die Ergebnisse der Modellierungssoftware primär für die Überprüfung der Methodik eingesetzt wurde, zeigte sich schon im Fall der modellierten Kupferschweißnaht, dass mit steigender Bandbreite der Glättung, der Schwellwert \hat{a}_{dec} von $5,14 \cdot 10^4$ GW auf $2,95 \cdot 10^4$ sank, gleichzeitig wurden geringere $a_{90/95}$ Werte sicher erfasst (2,27 mm bis 0,85 mm). Ähnliches galt auch für die Experimente mit künstlichen (\hat{a}_{dec} : $2,27 \cdot 10^4$ GW auf $1,58 \cdot 10^4$ und für $a_{90/95}$: 1,80 mm auf 0,83 mm) und auch realen Defekten (\hat{a}_{dec} : $6,48 \cdot 10^4$ GW auf $4,26 \cdot 10^4$ und für $a_{90/95}$: 2,46 mm auf 1,46 mm). Diese Ergebnisse spiegelten das erwartete Verhalten wider, dass größere Anzeigenflächen bei geringerem Kontrast erkannt werden können. Das Nicht-einbeziehen der Anzeigenfläche und somit der Einsatz der konventionellen POD resultierte in einer Unterbewertung des Systems, was besonders gut für künstliche Defekte zu sehen war (vgl. konservatives $a_{90/95} = 2,63$ mm).

Die Einführung realer Defektdaten stellte außerdem noch eine Herausforderung dar, da nur im Allgemeinen eine geringere Anzahl von Datenwerten vorhanden war und die Kosten für die detaillierte Untersuchung der realen Defekte im Allgemeinen hoch sind (beispielsweise durch die metallographische Untersuchung). Daher wurde eine gewichtete Kombination von verschiedenen Daten eingeführt, die es ermöglichte, auf die wichtigen Informationen der realen Defekte zugreifen zu können und dabei die statistische Sicherheit durch andere Datenquellen zu erhöhen. Im vorliegenden Fall wurden dabei künstliche Defekte als zusätzliche Informationsquelle ausgewählt. Die Kombination war jedoch nur für eine gemeinsame Datenstruktur möglich, daher spielten hierbei sowohl die Gegenüberstellung als auch die multi-skale Glättung (durch die Einbeziehung weiterer Defekteigenschaften) eine essentielle Rolle. Während das Ergebnis für die künstlichen Defekte das System zu gut bewertet hatte ($\hat{a}_{90/95} = 1,46$ mm), lag die Kombination beider Daten bei $\hat{a}_{90/95} = 2,69$ mm. Einerseits konnte das Vertrauen in die Daten gestärkt werden, da das Vertrauensintervall für die wenigen realen Daten durch die Erhöhung der Daten (von 20 realen Daten auf 40 repräsentative Daten) verringert werden konnte. Gleichzeitig wurde die Bewertung weniger von den künstlichen Defekten beeinflusst als beispielsweise bei einer Mischung der Daten (33 künstliche Daten wurden als 20 repräsentative Daten abgeschätzt). Der Einsatz von wenigen realen Defekten im Kombinationsmodus soll dem Einsatz von rein künstlichen De-

fektdaten vorgezogen werden, da die realen Defekte die eigentliche Prüfaufgabe am Besten umschreiben.

Während die Glättung bei der POD-Bewertung sowohl für simulierte als auch für experimentelle Daten den erwarteten Erfolg gebracht hat, ist die Frage hinsichtlich der korrekten Entscheidungsschwelle noch weitgehend ungeklärt. Im Allgemeinen wird die Frage nach den Schwellen häufig in der Prüfanweisung beantwortet. Im vorliegenden Fall wurde eine Mischung aus Ringversuch und CNR-Kriterium genutzt. Bei dem Ringversuch innerhalb dieser Arbeit konnten nur wenige Personen herangezogen werden, daher sollte die Detektionsschwelle im Weiteren überprüft werden.

Außerdem hat sich gezeigt, dass die statistischen Methoden in der Praxis fehlschlagen. POD für das nicht Vorhandensein eines Defektes ($POD(0) \neq 0$) genauso wie der beste Fit der Gerade, der nicht durch den Ursprung verläuft sind hierbei zu nennen. Ein Ansatz wäre die Annahme der homogenen Streuung abzulehnen und für jede Defektposition eine individuelle Streuung zu erstellen. Die gebräuchlichen Modelle hierfür resultieren wiederum in einer deutlich größeren Anzahl von notwendigen Datenwerten. Die neue Gestaltung der POD-Bewertung ist daher notwendig, um gleichzeitig praktische wie auch mathematische Anforderung zu beachten. Dies ist auch ein wichtiger Punkt mit dem sich die POD-Community im Moment intensiv beschäftigt [184].

Diese Arbeit ermöglicht durch die Einführung des erweiterten Verfahrens einen Weg zu öffnen, um noch näher zur realen Auffindwahrscheinlichkeit im Prüfprozess in der Praxis zu gelangen. Hierfür sind die mathematischen Anforderungen eine wichtige Voraussetzung, deren Umgang und Lösung der Widersprüche zwischen praktischen Daten und statistischer Theorie eine wichtiges Ziel in diesem Thema darstellt. Ein weiteres Ziel der zukünftigen Forschung stellt die räumliche Rekonstruktion der Schlibfbilder für die Bewertung mit Hilfe der POD dar. Dabei werden einige beobachtete Ausreißer von künstlichen Daten und einige Ungereimtheiten bei der Datenbeurteilung verhindert. Gleichzeitig wird die korrekte Rekonstruktion und die zusätzliche Absicherung mit weiterem Wissen neue Herausforderungen schaffen, die die Kommunikation zwischen ZfP-Prüfer, Statistiker und Materialwissenschaftler unumgänglich machen.

Außerdem ist die Abhängigkeit der Glättung vom realen Defekttyp wichtig zu beurteilen. Während sich diese Arbeit hauptsächlich auf voluminöse Defekte stützt, ist die Einbeziehung weiterer Defekttypen (z.B. Rissen) und der Einsatz in anderen ZfP-Verfahren eine sinnvolle Erweiterung des Verfahrens. Mit dem Fokus auf voluminöse Defekte stellt die Erweiterung hinsichtlich computertomographischer Prüfsysteme ein weiterer Anwendungsbereich dar. Eine Möglichkeit die anzeigeflächenabhängige Detektionsschwelle objektiver zu untersuchen, ist ein Ringversuch zur Bestimmung der Detektierbarkeit. Der Einsatz von Augenbewegungsverfolgung anhand des Eye-Trackers zur Untersuchung der Detektionsfähigkeit verschiedener Prüfer, Fragebögen und psychologischen Untersuchungen ermöglicht

5. Zusammenfassung und Schlussfolgerung

die Einbeziehung von subjektiven Faktoren auf die objektive Kennzahl der POD. Gleichzeitig kann in den Experimenten der Zusammenhang zwischen Glättungsbandbreite und Anzeigengröße untersucht werden.

Die POD-Bewertung hat gezeigt, dass nicht nur die Forderung nach der Kennzahl POD immer lauter werden, sondern auch die richtige Methodik für bestimmte Anwendungen immer dringender benötigt wird.

Anhang

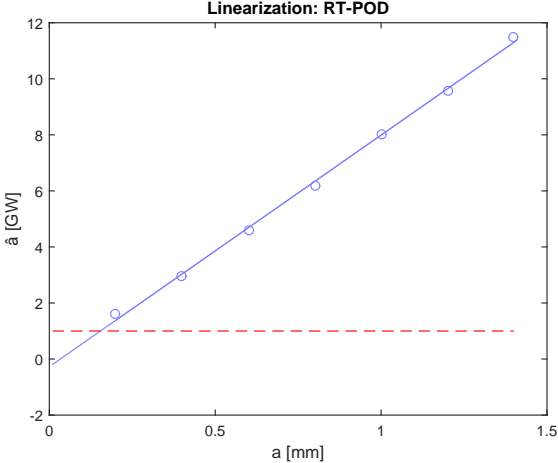
Anhang I: Modellierung der Linearität des Zusammenhangs zwischen Defektparameter und Signal

Zur Verifizierung des Zusammenhangs zwischen durchstrahlter Länge des Defekts und Kontrast soll beispielhaft ein einfacher Prüfaufbau simuliert werden (mit Hilfe von aRTist): Ein Stahlblech mit 19 mm Dicke soll durchstrahlt werden. Dafür werden eingebrachten runden Flachbodenbohrungen mit gleichem Durchmesser und verschiedener Länge simuliert. Dabei wurden folgende Werte gemessen:

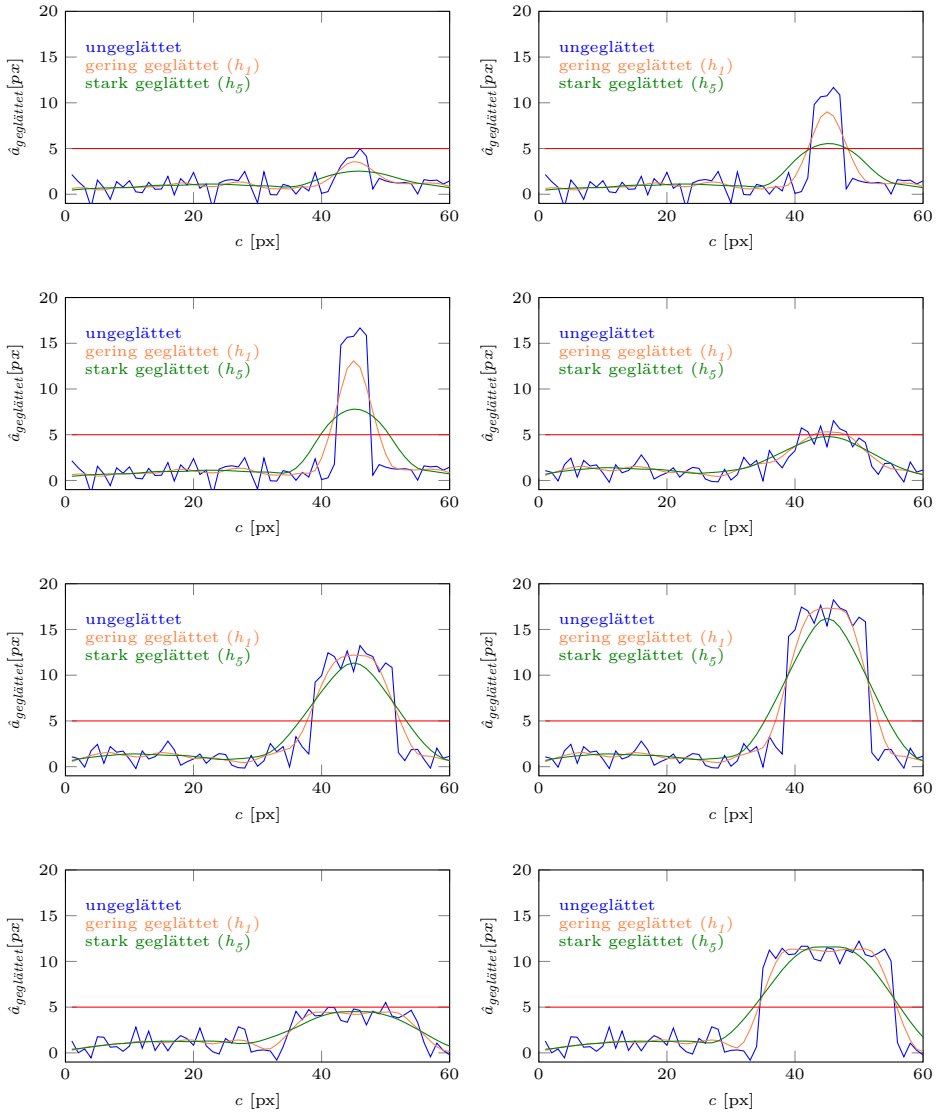
- Mittelwert der Signalintensität im Hintergrund: $\mu_I(B) = 25135$
- Standardabweichung (Rauschen): $\sigma_I(B) = 295.8$

| Durchstrahlte Länge in mm | Signal (I) in GW | Kontrast (C) in GW | CNR |
|---------------------------|------------------|--------------------|------|
| 1,4 | 28437 | 3302 | 11,2 |
| 1,2 | 27931 | 2796 | 9,5 |
| 1,0 | 27343 | 2299 | 7,8 |
| 0,8 | 26854 | 1719 | 5,8 |
| 0,6 | 26490 | 1355 | 4,6 |
| 0,4 | 26043 | 908 | 3,1 |
| 0,2 | 25610 | 475 | 1,6 |

Das CNR ist definiert als der Quotient aus der Differenz der mittleren Signalpegel von zwei Bildbereichen und der gemittelten Standardabweichung der Signalpegel [41]. Mit der Annahme des gleichen Hintergrundmaterials und bei gleicher Materialdicke besteht annähernd konstantes Rauschen, wodurch sich ein linearer Zusammenhang zwischen CNR und durchstrahlter Länge des Defekts ergibt.

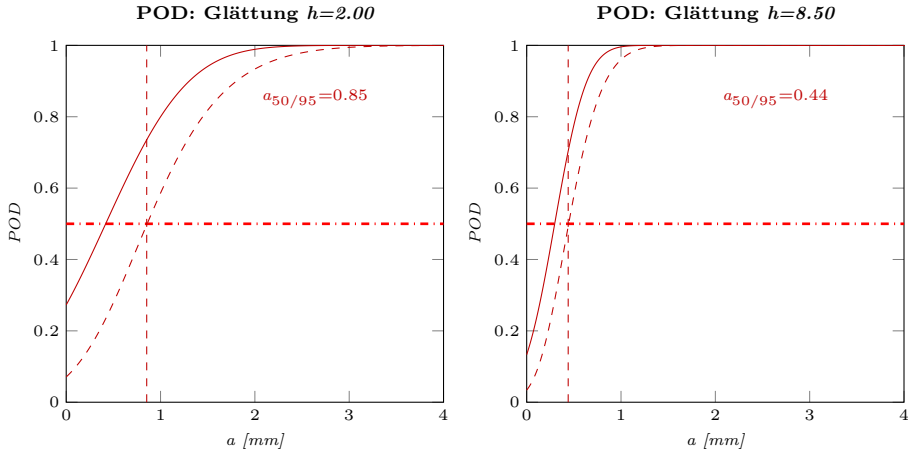


Anhang II: Einfluss der Glättung auf verschiedene Anzeigen

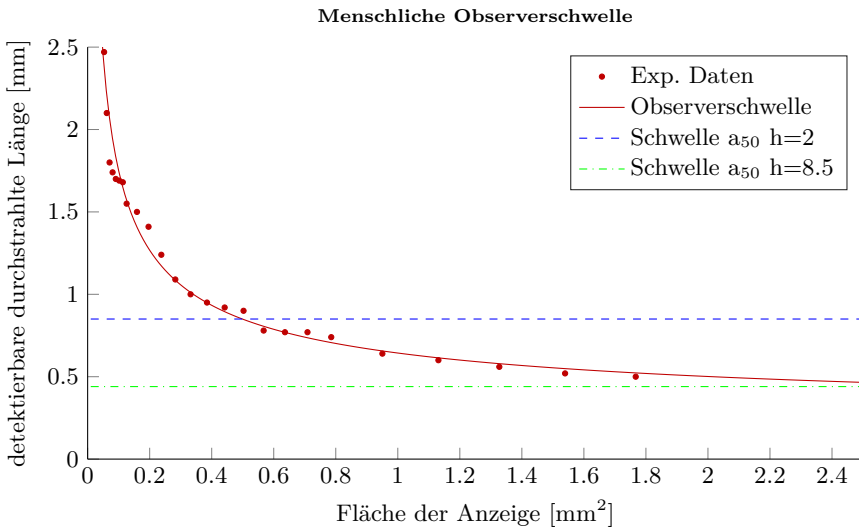


Anhang III: Abschätzung des Zusammenhangs zwischen Glättungsbandbreite und Fläche

Die POD bei 50% gibt einen Wert für die durchstrahlte Länge an, der gerade noch gefunden werden kann. Wegen der Streuung und den geringen Daten für die POD wurde auch hier der 95% Vertrauensbereich gewählt.



Die Observerschwelle basiert auf mehr Werten von Experimenten mit Prüfern, die subjektiv entschieden haben, ob eine Anzeige noch sichtbar war:



Der Schnittpunkt zwischen dem a_{50} und der Observerschwelle, schätzt die detektierbare Fläche. Jedoch ist dies abhängig vom Experiment und der Entscheidungsschwelle, wann ein Defekt gefunden wurde, und von den Einflussgrößen der POD. Eine weitere Untersuchung dieser Zusammenhänge ist wünschenswert.

Literatur

- [1] W. Weiss, C.-M. Larsson, Chr. McKenney, J.-P. Minon, S. Mobbs, T. Schneider, H. Umeki, W. Hilden, C. Pescatore und M. Vesterind. *Radiological Protection in Geological Disposal of Long-Lived Solid Radioactive Waste*. Annals of the ICRP 122. ICRP Publication. International Commission on Radiological Protection, 2011.
- [2] BMWi / IA8. *Endlagerung hochradioaktiver Abfälle in Deutschland - Das Endlagerprojekt Gorleben*. Hrsg. von Bundesministerium für Wirtschaft und Technologie. Peschke Druck GmbH, München. 2008.
- [3] D. Louvar und P. Metcalf. „Closing the Cycle: Disposal options for low level waste have been developed and good prospects for geological disposal of radioactive waste appear to be on the horizon in several member states“. In: *IAEA Bulletin* 51.2 (2010), S. 20–23. URL: <http://www.iaea.org/Publications/Magazines/Bulletin/Bull512/51202642023.pdf>.
- [4] C. Peachey. „Disposal Plans“. In: *Nuclear Engineering International* June (2012), S. 38–48.
- [5] C. O. Richter. *Finnland: Atommüll - Ja gerne!* NDR Fernsehen. 15.01.2013 23:30 Uhr. 2013. URL: <http://www.ndr.de/fernsehen/sendungen/weltbilder/videos/weltbilder2747.html>.
- [6] C. Vider. *Quarks und CO: Atommüll - Endlager verzweifelt gesucht*. WDR Fernsehen. Köln, 9. Nov. 2010. URL: podcast.wdr.de.
- [7] L. Potterton. „Living with Nuclear: In the Swedish city of Oskarshamn, public support for nuclear power and acceptance of nuclear waste is strong“. In: *IAEA Bulletin* 51.2 (2010), S. 28–31.
- [8] E. Clarke. „The truth about... nuclear waste“. In: *ClimateChangeCorp.com Technology*. Technology (2009). URL: www.climatechangecorp.com.
- [9] N.N. *Encapsulation When, Where, How and Why*. Hrsg. von Svensk Kärnbränslehantering. Booklet. ISBN 91-975606-4-2. BOX 250, SE-10134 Stockholm, Sweden, 2008. URL: <http://www.skb.se/>.
- [10] N.N. *Final disposal of spent nuclear fuel in Olkiluoto*. Hrsg. von Posiva Oy. Olkiluoto, 27160 Eurajoki, Finland, 2011.

Literatur

- [11] N.N. *Geological Disposal of Radioactive Waste: Technological Implication of Retrievability*. NW-T- 1.19. Technical Report. International Atomic Energy Agency, 2009.
- [12] N.N. *Design and production of the KBS-3 repository*. TR- 10-12. Technical Report. Svensk Kärnbränslehantering AB, 2010.
- [13] N.N. *Safe Final Disposal of Spent Nuclear Fuel*. Hrsg. von Posiva Oy. Olkiluoto, 27160 Eurajoki, Finland, 2010. URL: www.posiva.fi.
- [14] J. Pitkänen. *Inspection of Bottom and Lid Welds for Disposal Canisters*. Posiva 2010-04. Report. Posiva Oy, 2010.
- [15] A.P. Berens. *NDE Reliability Data Analysis: Nondestructive Evaluation and Quality Control*. In: *Metals Handbook*. 9. Aufl. Bd. 17. Materials Park, Ohio: ASM International, 1989, S. 689–701.
- [16] J. Pitkänen, R. Paussu, P. Pohjanne, I. Virkkunen, M. Kemppainen, A. Lipponen, M. Sarkimo, K. Simola und K.-M. Reddy. „Metallographic study of Detected Indications in EB-Copper welds for Verifying the NDT Reliability of Inspection“. In: *8th International Conference on NDE in Relation to Structural Integrity for Nuclear and Pressurised Components*. Hrsg. von Deutsche Gesellschaft für zerstörungsfreie Prüfung e.V. Berlin, 29. Sep. 2010. ISBN: 9-789279-203213.
- [17] Directorates General Environment and Joint Research Centre. *ENIQ Glossary*. ENIQ Report 12. EUR18102EN. European Commission, 1999.
- [18] K. Takahashi, M. Pavlovic, M. Bertovic, U. Ewert, C. Müller und U. Ronneteg. „Application of POD to complex industrial problems: New approach“. In: *Review of Progress in Quantitative Nondestructive Evaluation: Proceedings of the 35th Annual Review of Progress in Quantitative Nondestructive Evaluation*. AIP Conference Proceedings 1096.1 (2009), S. 1824–1831.
- [19] DIN. *Internationales Elektrotechnisches Wörterbuch*. Deutsch / Englisch. Norm-Entwurf: DIN IEC 50(191). Deutsches Elektrotechnische Kommission, Nov. 1994.
- [20] DIN. *Zuverlässigkeit Begriffe*. Norm: DIN 40041. Deutsche Elektrotechnische Kommission Ausschuss Qualitätssicherung und angewandte Statistik, Dez. 1990.
- [21] FAZ-Redaktion, AP und Reuters. „Personenzug entgleist in Manila“. In: *FAZ*. Unglücke www.faz.net (2014). URL: www.faz.net/aktuell/gesellschaft/ungluecke/dutzende-menschen-verletzt-personenzug-entgleist-in-manila-13097558.html.
- [22] FAZ-Redaktion. „Ermittlungen wegen Kölner ICE-Unfall eingestellt“. In: *FAZ*. Gesellschaft www.faz.net (2009). URL: <http://www.faz.net/-gum-134fv>.
- [23] ISO. *Risk management*. Richtlinie: ISO-Guide 73. GUIDE 73. ISO, Nov. 2009.

- [24] H. Czichos. *Scope of Technical Diagnostic*. In: *Handbook of Technical Diagnostics. Fundamentals and Application to Structures and Systems*. Hrsg. von Horst Czichos. 1. Aufl. Industrial Radiology. Berlin Heidelberg: Springer Verlag, 2013, S. 3–9. ISBN: 978-3-642-25849-7.
- [25] ISO. *Condition monitoring and diagnostics of machines*. Norm: ISO 13372. ISO. ISO, Mai 2004.
- [26] M. Turnbow, S. Daisy und J. Thompson. „Reliability of NDE - The Criticality of Procedure and Personnel“. In: *4th European-American Workshop on Reliability*. Hrsg. von Deutsche Gesellschaft für zerstörungsfreie Prüfung e.V. BB 116-CD. Berlin, 24. Juni 2009. ISBN: 978-3-940283-17-7.
- [27] DIN. *Genauigkeit (Richtigkeit und Präzision) von Meßverfahren und Meßergebnissen TEIL 1: Allgemeine Grundlagen und Begriffe*. Norm: ISO 5725-1. Normenausschuß Qualitätsmanagement, Statistik und Zertifizierungsgrundlagen (NQSZ), Nov. 1994.
- [28] JCGM. *Evaluation of measurement data - Guide to the expression of uncertainty in measurement*. Richtlinie: GUM:JCGM 100:2008. Working Group 1 - Joint Committee for Guides in Metrology (JCGM/WG 1), Sep. 2008.
- [29] W. Rummel. „Challenges to Nondestructive Testing Reliability“. English. In: *Materials Evaluation* 73.1 (2015), S. 38–42.
- [30] W. Hässelbarth. *BAM-Leitfaden zur Ermittlung von Messunsicherheiten bei quantitativen Prüfergebnissen*. Forschungsbericht. BAM, 2004.
- [31] B. N. Taylor und C. E. Kuyatt. *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*. NIST 1297. Technical Note. United States Department of Commerce: Technology Administration: National Institute of Standards und Technology, 1994.
- [32] D. Placko. *Metrology in Industry. The Key of Quality*. Wiley-ISTE. French College of Metrology, 2006. ISBN: 978-1-905209-51-4.
- [33] J.-E. Holmberg und P. Kuusela. *Analysis of probability of defects in the disposal canisters*. Working Report: WR 2011-36. VTT Research Report. VTT, 2011.
- [34] C. Müller, T. Fritz, G.-R. Tillack, C. Bellon und M. Scharmach. „Theory and Application of the Modelar Approach of NDE Reliability Bridging the Gab Between Safety Requirements adn Economy“. In: *NDT.net* 06.09 (2001).
- [35] J. A. Swets und R. M. Pickett. *Evaluation of diagnostic systems*. Hrsg. von Academic Press. Academic Press, 1982. ISBN: 0-12-679080-9.
- [36] R. Singh. *Three Decades of Reliability Assessment*. Karta- 3510-99-01. Report. San Antonio: Karta, 2000.

- [37] C. Müller und M. Golis. „Basic Ideas of the American-European Workshops 1997 in Berlin and 1999 in Boulder“. In: *15th World conference on NDT*. Hrsg. von ICNDT. Bd. 15th. Roma, 15. Okt. 2000.
- [38] M. Bertovic. „Human Factors in Non-Destructive Testing (NDT): Risks and Challenges of Mechanised NDT“. Diss. Berlin: Technische Universität Berlin, 2015.
- [39] R. Holstein, M. Bertovic, D. Kanzler und C. Müller. „NDT Reliability in the Organizational Context of Service Inspection Companies“. In: *Materials Testing* 56.7-8 (2014), S. 607–610.
- [40] *General Principles of Reliability Assessment of Nondestructive Diagnostic Systems and Its Applicability to the Demining Problem*. 8th ECNDT. 2003. (Besucht am 15.08.2015).
- [41] ISO. *Zerstörungsfreie Prüfung von Schweißverbindungen Durchstrahlungsprüfung Teil 2: Röntgen- und Gammastrahlungstechniken unter Anwendung digitaler Detektoren*. Norm-Entwurf: FprEN ISO 17636-2. Technischen Komitee ISO/TC 44 Technischen Komitee CEN/TC 121, Mai 2013.
- [42] H.-F. Shyu und W.-J. Shong. „Statistical analysis results of performance demonstration for piping welds“. In: *Insight* 47.11 (2005). PERFORMANCE DEMONSTRATION, S. 688–690.
- [43] S. Sandlin. *Defect Detectability in the Disposal Canister Lid-Weld Using the 9 MeV Linear Accelerator*. Posiva WR- 2009-84. Working Report. Posiva Oy, 2010.
- [44] J. Pitkänen. *Inspection of Disposal Canisters Components*. English. Posiva 2012-35. Technical Report. Posiva, 2013.
- [45] W. Q. Meeker und L. A. Escobar. *Statistical Methods for Reliability Data*. New York: Wiley Series in Probability und Mathematical Statistics, 1998. ISBN: 978-0-471-14328-6.
- [46] A. F. Grandt. *Fundamentals of Structural Integrity: Damage tolerant design and nondestructive evaluation*. 1. Aufl. Hoboken, New Jersey: John Wiley & Sons, Inc, 2004. ISBN: 0-471-21459-0.
- [47] J. Brown und S. James. *ASTM Standard Practice for Probability of Detection Analysis for a-hat versus a Data*. WK 40772. Work Item. ASTM, 2013.
- [48] M. Purschke. *The X-Ray Inspection (RT/RS)*. Bd. 7. NDT compact and understandable - Informative booklets for non-destructive testing. Wuppertal: Castell-Verlag GmbH, 2004. ISBN: 3-934 255-22-1.
- [49] K. V. Bury. *Statistical Models in Applied Science*. Hrsg. von John Wiley & Sons. London: Wiley Series in Probability und Mathematical Statistics, 1975. ISBN: ISBN 0-471-12590-3.

- [50] ASTM International. *Standard Terminology for Nondestructive Examinations*. Norm: ASTM E 1316 - 09a. ASTM International, Juli 2009.
- [51] K. Kroschel. *Statistische Informationstechnik*. 4. Aufl. Springer-Verlag, 2004. ISBN: 3-540-40237-3.
- [52] S. M. Ross. *Statistik für Ingenieure und Naturwissenschaftler*. Hrsg. von Elsevier. 3. Aufl. Elsevier München: Spektrum akademischer Verlag, 2006. ISBN: 978-3-8274-1621-6.
- [53] R. B. Thompson. „Early work on the use of models in the determination of POD / Inspection reliability in the US and UK“. In: *Review of Progress in Quantitative NDE* 30.1335 (2011), S. 1581–1588.
- [54] N. S. Goujon. *Safety Implications of TOFD for In-manufacture Inspections*. RR 433. Research Report. Mitsui Babcock Energy Limited for Health und Safety Executive, 2006.
- [55] International Electrotechnical Commission. *Functional Safety and IEC 61508*. Webseite. 10. Nov. 2014.
- [56] United State Nuclear Regulatory Commission. *Safety-related*. Hrsg. von U.S. NRC. Webseite. United State Nuclear Regulatory Commission, 28. Okt. 2014. URL: <http://www.nrc.gov/reading-rm/basic-ref/glossary/safety-related.html>.
- [57] B. D. Olin und W. Q. Meeker. „Applications of Statistical Methods to Nondestructive Evaluation“. In: *Technometrics* 38.2 (1996), S. 95–130.
- [58] Y. Pawitan. *In all Likelihood: Statistical Modelling and Inference Using Likelihood*. Hrsg. von Clarendon Press. Oxford Science Publications, 2001. ISBN: 0-19-850765-8.
- [59] R. B. Thompson. „POD of Aircraft Engines“. In: *WFNDEC 2010*. Hrsg. von World Federation of NDE Centers. San Diego, 17. Juli 2010.
- [60] H. Czichos. *Application of Technical Diagnostics*. In: *Handbook of Technical Diagnostics. Fundamentals and Application to Structures and Systems*. Hrsg. von Horst Czichos. 1. Aufl. Industrial Radiology. Berlin Heidelberg: Springer Verlag, 2013, S. 11–22. ISBN: 978-3-642-25849-7.
- [61] H. Czichos. *Physics of Failure*. In: *Handbook of Technical Diagnostics. Fundamentals and Application to Structures and Systems*. Hrsg. von Horst Czichos. 1. Aufl. Industrial Radiology. Berlin Heidelberg: Springer Verlag, 2013, S. 23–40. ISBN: 978-3-642-25849-7.
- [62] T. Boellinghaus. *Principles and Concepts of Technical Failure Analysis*. In: *Handbook of Technical Diagnostics. Fundamentals and Application to Structures and Systems*. Hrsg. von Horst Czichos. 1. Aufl. Industrial Radiology. Berlin Heidelberg: Springer Verlag, 2013, S. 341–353. ISBN: 978-3-642-25849-7.

- [63] A. Erhard. *Non-destructive Evaluation*. In: *Handbook of Technical Diagnostics. Fundamentals and Application to Structures and Systems*. Hrsg. von Horst Czichos. 1. Aufl. Industrial Radiology. Berlin Heidelberg: Springer Verlag, 2013, S. 161–174. ISBN: 978-3-642-25849-7.
- [64] O. Førli. *Guidelines for NDE reliability determination and description*. NT TR 394. Nordtest Technical Report. Det Norske Veritas AS, 1998.
- [65] G. Redmond. „From ‘Safe Life’ to Fracture Mechanics - F111 Aircraft Cold Temperature Proof Testing at RAAF Amberley“. In: *10th Asia-Pacific Conference on Non-Destructive Testing*. Hrsg. von AINDT. Brisbane. Brisbane, 17. Sep. 2001.
- [66] ASTM International. *Standard Practice for Ultrasonic Inspection of Aluminum - Alloy Wrought Products for Aerospace Applications*. Norm: ASTM B 594-02. ASTM International, Apr. 2002.
- [67] J. Pitkänen, M. Bertovic, D. Brackrock, G. Brekow, U. Ewert, D. Kanzler und C. Müller. „Reliable Evaluation of Acceptability of Weld for Final Disposal Based on the Canister Copper Weld Inspection Using Different NDT Methods“. In: *Materials Testing* 56.9 (2014), S. 748–757.
- [68] A.P. Berens und P.W. Hovey. *Evaluation of NDE Reliability Characterization*. AFWAL-TR 81-4160. Technical Report. Materials Laboratory - Air Force Wright Aeronautical Laboratories, 1981.
- [69] J. P. Gallagher, C. A. Babish und J. C. Malas. „Damage Tolerant Risk Analysis Techniques for Evaluating the Structural Integrity of Aircraft Structures“. In: *Proceedings of 11th International Conference on Fracture 2005*. Hrsg. von Inc. Curran Associates. Turin, Italy, 20. Mai 2005. ISBN: 9781617820632.
- [70] W. Rummel. „Recommended Practice for a Demonstration of Nondestructive Evaluation (NDE) Reliability on Aircraft Production Parts“. In: *Material Evaluation* 40.9 (1982), S. 922–932.
- [71] L. Gandossi und C. Annis. *Probability of Detection Curves: Statistical Best-Practices*. English. Statistical best-practices 41. scientific report. ENIQ, 2010.
- [72] C.R.A Schneider und G.A. Georgiou. „Radiography of thin-section welds, Part 2: Modelling“. In: *Insight* 45.2 (2003), S. 119–121.
- [73] M. G. Kendall. *Kendall’s advanced theory of statistics. Classical inference and relationship. Advanced theory of statistics*. 5. Aufl. Bd. 2. Kendall, Maurice G.: The advanced theory of statistics. London: Alan Stuart, John K. Ord und Anthony O’Hagan, 1991. ISBN: 0-340-52923-7.

- [74] N. J.D. Nagelkerke. *Maximum Likelihood Estimation of Functional Relationships. Lecture Notes in Statistics*. New York: Springer Verlag, 1992. ISBN: 0-387-97721-X 3-540-97721-X.
- [75] A. Walker, W. Daniels und S. Wedge. „Use of Monte-Carlo methods to Derive Quantitative Probability of Defect Detection Data“. In: *4th European-American Workshop on Reliability*. Hrsg. von Deutsche Gesellschaft für zerstörungsfreie Prüfung e.V. BB 116-CD. Berlin, 24. Juni 2009. ISBN: 978-3-940283-17-7.
- [76] J. Pilz. *Bayesian Estimation and Experimental Design in Linear Regression Models*. New York: Wiley Series in Probability und Mathematical Statistics, 1991. ISBN: 0-471-91732-X.
- [77] R. H. Burkel, D. J. Sturges, W.T. Tucker und R.S. Gilmore. „Probability of Detection for Applied Ultrasonic inspection“. In: *Review of Progress in Quantitative NDE 15* (1996), S. 1991–1998.
- [78] A. Wald. *Sequential Analysis*. 1. Aufl. New York: Dover Publications & John Wiley & Sons, 1973, S. 212.
- [79] S. J. Press. *Bayesian statistics: Principles, Models, and Applications*. Hrsg. von S. James Press. Wiley Series in Probability und Mathematical Statistics, 1989. ISBN: 0 471 63729 7.
- [80] G. D. Kleiter. *Bayes-Statistik: Grundlagen und Anwendungen. Grundlagen und Anwendungen*. 1. Aufl. Berlin: W. de Gruyter, 1980. ISBN: 978-3110082739.
- [81] Department of Defense. *Nondestructive Evaluation System Reliability Assessment*. Richtlinie: MIL-HDBK-1823A. MIL. SUPERSEDING MIL-HDBK-1823 14 April 2004. Departments and Agencies of the Department of Defense, 2009.
- [82] T. A. Severini. *Likelihood Methods in Statistics*. A.C. Atkinson and J.B. Copas and D.A. Pierce and M.J. Schervish and D.m. Titterington. Statistical science. Oxford: Oxford Statistical Science Series, 2000. ISBN: 0 19 850650 3.
- [83] A. W. F. Edwards. *Likelihood - An account of the statistical concept of likelihood and its application to scientific inference*. Bentley House, 200 Euston Road, London: Cambridge at the University Press, 1972. ISBN: 978-0801844430.
- [84] Y. Wang und W. Q. Meeker. „A Statistical Model to adjust for flaw-size bias in the computation of Probability of Detection“. In: *Review of Quantitative Nondestructive Evaluation 25*.Cp820 (2006), S. 1854–1861.
- [85] L.R. Shenton und K.O. Bowman. *Maximum Likelihood estimation in small samples*. Professor Alan Stuart. Bd. 38. Oxford: Griffin’s Statistical Monographs und Coruses No. 38, 1977. ISBN: 0 85264 238 5.

- [86] C. Annis, J. C. Aldrin und H. A. Sabbath. „What is Missing in Nondestructive Testing Capability Evaluation?“ In: *Materials Evaluation* 73.1 (2015), S. 44–54.
- [87] A. Prosch und B. Larson. *Real-time Radiography - An Introductory Course Module for NDT Students*. Hrsg. von North Central Collaboration For Education in NDE/NDT. Iowa, 2000.
- [88] A. P. Berens und P. W. Hovey. *Flaw Detection Reliability Criteria*. AD-A 141 001. Final technical rept. May 82-Sep 83. University of Dayton, 1984.
- [89] M. Pavlovic, K. Takahashi und C. Müller. „Probability of detection as a function of multiple influencing parameters“. In: *Insight* 54.11 (2012), S. 606–611.
- [90] D. Kanzler, C. Müller und J. Pitkänen. *Probability of defect detection of Posiva’s electron beam weld*. WR 2013-70. Working Report. Posiva Oy, 2013.
- [91] S. Feistkorn. „Gütebewertung qualitativer Prüfaufgaben in der zerstörungsfreien Prüfung im Bauwesen am Beispiel des Impulsradarverfahrens“. Diss. Berlin: Technischen Universität Berlin, 2012.
- [92] V. Deutsch, M. Platte, W. A. Karl Deutsch, V. Schuster und M. Vogt. *Die Ultraschallprüfung (UT)*. 2. Aufl. Bd. 2. ZfP kompakt und verständlich. Castell-Verlag GmbH Wuppertal, 2010. ISBN: 978-3-934255-52-4.
- [93] D. Kanzler, C. Müller und J. Pitkänen. „Probability of detection for surface-breaking holes with low-frequency eddy current testing - a non-linear multi-parametric approach“. In: *Insight* 56.12 (2014), S. 664–668.
- [94] DIN. *Zerstörungsfreie Prüfung von Schweißverbindungen - Allgemeine Regeln für metallische Werkstoffe*. Norm: DIN EN ISO 17635. ISO, Aug. 2010.
- [95] L. Gandossi, K. Simola und B. Shepherd. „Application of a Bayesian model for the quantification of the European methodology for qualification of non-destructive testing“. In: *International Journal of Pressure Vessels and Piping* 87 (2010), S. 111–116.
- [96] C. Annis und L. Gandossi. *Influence of Sample Size and Other Factors on Hit/Miss Probability of Detection Curves*. ENIQ TGR Technical Document 47. 24429EN. European Network for Inspection und Qualification, 2011.
- [97] M. Li, F.W. Spencer und W.Q. Meeker. „Quantile Probability of Detection: Distinguishing Between Uncertainty and Variability in Nondestructive Testing“. In: *Materials Evaluation* 73.1 (2015), S. 89–95.
- [98] R. C. H. Cheng und T. C. Iles. „Confidence Bands for Cumulative Distribution Functions of Continuous Random Variables“. In: *Technometrics* 25.1 (1983), S. 77–86.

- [99] A. M. Mood, F. A. Graybill und D. C. Boes. *Introduction to the Theory of Statistics*. Hrsg. von David Blackwell und Herbert Solomon. 3. Aufl. Series in Probability and Statistics. New York: McGraw-Hill, 1974.
- [100] F. W. Spencer. „The Calculation and Use of Confidence Bounds in Pod Models“. In: *Review of Quantitative Nondestructive Evaluation* 26 (2007), S. 1791–1798.
- [101] Y. Hong, L. A. Escobar und W. Q. Meeker. „Coverage Probabilities of Simultaneous Confidence Bands and Regions for Log-Location-Scale Distributions“. In: *Journal of Statistical Planning and Inference* 139.9 (2009), S. 3231–3245.
- [102] B. Raj und B. Venkatraman. *Overview of Diagnostics and Monitoring Methods and Techniques*. In: *Handbook of Technical Diagnostics. Fundamentals and Application to Structures and Systems*. Hrsg. von Horst Czichos. 1. Aufl. Industrial Radiology. Berlin Heidelberg: Springer Verlag, 2013, S. 43–68. ISBN: 978-3-642-25849-7.
- [103] R. W. Nichols. *An Introductory Overview*. In: *improving the Effectiveness and Reliability of Non-Destructive Testing*. Hrsg. von W.E. Gardner. 1. Aufl. England: Pergamon Press, 1992, S. 1–24. ISBN: 0080369812.
- [104] R. Halmshaw. *Industrial Radiography - Theory and Practice*. Hrsg. von Applied Science Publishers. British Library Cataloguing. Englewood, New Jersey: Applied Science Publishers, 1982. ISBN: 0-85334-105-2.
- [105] U. Ewert, G.-R. Jaenisch, U. Zscherpel und K. Osterloh. *Industrial Radiology*. In: *Handbook of Technical Diagnostics. Fundamentals and Application to Structures and Systems*. Hrsg. von Horst Czichos. 1. Aufl. Industrial Radiology. Berlin Heidelberg: Springer Verlag, 2013, S. 221–247. ISBN: 978-3-642-25849-7.
- [106] R. Halmshaw. „The Factors Involved in an Assessment of Radiographic Definition“. In: *The Journal of Photographic Science* 3.3 (1955), S. 161–168.
- [107] L. Papula. *Mathematik für Ingenieure und Naturwissenschaftler. Fachbücher der Technik - Ein Lehr- und Arbeitsbuch für das Grundstudium*. 11. Aufl. Bd. 1. Wiesbaden: Vieweg Teubner Verlag, 2007.
- [108] C. Bellon, A. Deresch und G.-R. Jaenisch. „Radiography Simulation with aRTist - Combining Analytical and Monte Carlo Methods“. In: *Digital Industrial Radiology and Computed Tomography (DIR 2015)*. Hrsg. von Deutsche Gesellschaft für zerstörungsfreie Prüfung e.V. Digital Industrial Radiology and Computed Tomography. Deutsche Gesellschaft für zerstörungsfreie Prüfung e.V. Ghent Belgium, 22. Juni 2015.
- [109] W. Stolz. *Effekte der Physik und ihre Anwendung*. Hrsg. von Manfred von Ardenne. Hrsg. von Gehard Musiol. Hrsg. von Uwe Klemradt. Bd. 3. Frankfurt am Main: Harri Deutsch, 2005. ISBN: 978-3817116829.

Literatur

- [110] R. Halmshaw. *Development in Radiographic Inspection - Film versus Real-time Radiography (Radioscopy)*. In: *Improving the Effectiveness and Reliability of Non-Destructive Testing*. Hrsg. von W.E. Gardner. 1. Aufl. England: Pergamon Press, 1992, S. 137–164. ISBN: 0080369812.
- [111] N. Brierley, T. Tippetts und P. Crawley. „Data fusion for automated non-destructive inspection“. In: *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 470.2167 (2014).
- [112] DIN. *Non-destructive testing - Image quality of radiographs - Part 1: Determination of the image quality value using wire-type image quality indicators*. Norm: DIN EN ISO 19232-1. DIN, 2005.
- [113] G.-R. Jaenisch, C. Bellon und U. Ewert. „Reliability Assessment by Simulation for High Energy Radiography“. In: *Review of Quantitative Nondestructive Evaluation* 26.Cp894 (2007), S. 1831–1838.
- [114] CEN. *Bildgüte von Durchstrahlungsaufnahmen*. Norm: EN 462-5. CEN, März 1996.
- [115] ISO. *Zerstörungsfreie Prüfung - Bildgüte von Durchstrahlungsaufnahmen - Teil 5: Bestimmung der Bildunschärfezahl mit Doppeldraht-Typ-Bildgüteprüfkörpern*. Norm: DIN EN ISO 19232-5. DIN, 2005.
- [116] ASTM International. *Standard Practice for Determining Total Image Unsharpness in Radiology*. Norm: ASTM E2002-98. ASTM International, 2009.
- [117] M. P. Souza. „Detection of Lack of Fusion weld defects by radiography“. In: *Review of Quantitative Nondestructive Evaluation* 28.CP1096 (2009), S. 1174–1181.
- [118] D. Firth. *Demonstrating Effect Inspection*. In: *improving the Effectiveness and Reliability of Non-Destructive Testing*. Hrsg. von W.E. Gardner. 1. Aufl. England: Pergamon Press, 1992, S. 165–172. ISBN: 0080369812.
- [119] ENIQ. *European Methodology for Qualification of Non-destructive Testing*. LDNA 22906. ENC-002. European Network for Inspection und Qualification, 2007.
- [120] W. Q. Meeker. *Trends in the Statistical Assessment of Reliability*. In: *Advances in Degradation Modeling. Applications to Reliability, Survival Analysis, and Finance*. Bd. *Advances in Degradation Modeling. Statistics for Industry and Technology*. Boston: Birkhäuser Boston, 2010, S. 3–16. ISBN: 978-0-8176-4923-4.
- [121] S. Crutzen, P. Lemaitre und I. Iacono. „Realistic Defects Suitable for ISI Capability Evaluation and Qualification“. In: *14th International Conference on NDE in the Nuclear and Pressure Vessel Industries*. Hrsg. von European Commission Joint Research Centre. ASM international. Stockholm, 24. Sep. 1996.

- [122] I. Virkkunen, M. Kempainen, H. Ostermeyer, R. Paussu und T. Dunhill. „Grown cracks for NDT development and qualification: Test Block“. In: *Insight* 51.5 (2009), S. 271–278.
- [123] J. Caldwell, M. Moles, L. Robertson, A. Hull und A. Sinclair. „Reference Discontinuities for Small Surface-breaking Cracks“. In: *Materials Evaluation* 08 (2013), S. 928–935.
- [124] D. P. Johnson. „Inspection Uncertainty: The Key Element in Nondestructive Inspection“. In: *Nuclear Engineering and Design* 43.43 (1977), S. 219–226.
- [125] M. Kemppainen und I. Virkkunen. „Crack Characteristics and Their Importance to NDE“. In: *Journal of Nondestructive Evaluation* 30 (2011), S. 143–157.
- [126] H. Wüstenberg und A. Erhard. „Problems with artificial test reflectors at the performance demonstration of Ultrasonic Inspection“. In: *European Conference on Non-Destructive Testing*. Nice, 24. Okt. 1994.
- [127] B. Neundorf, G. Csapo und A. Erhard. „Optimizing the NDT of Boiling Water Reactors by Using Realistic Flaws in the Cladding“. In: *7th European Conference on Non-Destructive Testing*. Hrsg. von ICNDT. Copenhagen, 26. Mai 1998.
- [128] D. Gülle, M. Gaal, M. Bertovic, C. Müller, M. Scharmach und M. Pavlovic. *South-east Europe Interim Report Field Trial Croatia*. Interim Report Field Trial Croatia -. Final Report. Bundesanstalt für Materialforschung und -Prüfung, 2006.
- [129] S. Dugan und S. Zickler. „Comparison of Different Test Reflectors for Ultrasonic Testing“. In: *35th MPA-Seminar*. Hrsg. von MPA. Stuttgart, 9. Okt. 2009.
- [130] S. Majidnia, C. Schneider, J. Rudlin und R. Nilavalan. „A method of determining the eddy current instrument response from models“. In: *Insight* 56.2 (2014), S. 82–86.
- [131] J. P.A. Ioannidis. „Why most published research findings are false“. In: *PLoS Medicine* 2.8 (2005), S. 696–701.
- [132] R. B. Thompson und W. Q. Meeker. „Assessing the POD of Hard-alpha Inclusions from Field Data“. In: *Review of Quantitative Nondestructive Evaluation* 26.Co894 (2007), S. 1759–1766.
- [133] D. Kanzler, U. Ewert, C. Müller und J. Pitkänen. „Observer POD for radiographic testing“. English. In: *AIP Conference Proceedings* 1650 (2015), S. 562–570.
- [134] J. S. Knopp, F. Ciarallo und R. V. Grandhi. „Development in Probability of Detection Modeling and Simulation Studies“. English. In: *Materials Evaluation* 73.1 (2015), S. 55–61.

Literatur

- [135] J. N. Gray, T. A. Gray, N. Nakagawa und R.B. Thompson. *Models for Predicting NDE Reliability*. Hrsg. von ASM International. Bd. ASM Handbook 17. ASM International, 1989. Kap. Models for Predicting NDE Reliability, S. 702–715.
- [136] F. Jenson und P. Benoist. „Overview of the MAPOD (SISTAE) General Approach“. In: *WFNDEC 2010*. Hrsg. von World Federation of NDE Centers. San Diego, 17. Juli 2010.
- [137] W. Q. Meeker. „Probability of Detection“. In: *WFNDEC 2010*. Hrsg. von World Federation of NDE Centers. San Diego, 17. Juli 2010.
- [138] N. Dominguez und F. Jenson. „Simulation assisted POD of a high frequency eddy currents inspection procedure“. In: *10th European conference and exhibition on non-destructive testing*. Hrsg. von RSNTTD. Moskow, 7. Juli 2010.
- [139] K. Beddek, R. H. De Staelen, S. Clenet, O. Moreau, Y. Le Menach und R. Van Keer. „Direct Method and Bayesian Inference for Uncertainty Quantification on Stochastic NDT Problem“. In: *Fifth International Conference on Advanced Computational Methods in Engineering*. Hrsg. von University of Liege. ACOMEN 2011. Liege, Belgium, 14. Nov. 2011. ISBN: 9782960114317.
- [140] R. B. Thompson. „Model-assisted Probability of Detection (Mapod) Overview“. In: *WFNDEC 2010*. Hrsg. von World Federation of NDE Centers. San Diego, 17. Juli 2010.
- [141] C. G. Drury und J. Watson. *Good Practices in Visual Inspection*. FAA - Forschungsbericht. Applied Ergonomics Group Inc, 2002.
- [142] Y. Hong, W. Q. Meeker und J. D. McCalley. „Prediction of Remaining Life of Power Transformers Based on Left Truncated and Right Censored Lifetime Data“. In: *The Annals of Applied Statistics* (2009), S. 857–879.
- [143] R. B. Thompson, L. J. Brasche, D. Forsyth, E Lindgren, P. Swindell und W. Winfree. „Recent advances in model-assisted Probability of Detecion“. In: *4th European-American Workshop on Reliability*. Hrsg. von Deutsche Gesellschaft für zerstörungsfreie Prüfung e.V. BB 116-CD. Berlin, 24. Juni 2009. ISBN: 978-3-940283-17-7.
- [144] D. Kanzler, C. Müller, U. Ewert. und J. Pitkänen. „Einsatz der Bayesschen Statistik in der Zuverlässigkeitsbestimmung von zerstörungsfreien Prüfsystemen“. In: *Materials Testing* 55.4 (2013), S. 254–260.
- [145] J. Pitkänen, T. Salonen, S. Sandlin und U. Ronneteg. „Defect detectability in EB-welded copper disposal cansiter with 9 MeV accelerator“. In: *6th International Conference on NDE in Relation to Structural Integrity for Nuclear and Pressurized Components*. Hrsg. von European Commission’s Joint Research Centre. Budapest, Hungary, 8. Okt. 2007. ISBN: 978-92-79-08897-1.

- [146] ASTM-International. *Standard Practice for Determining Relative Image Quality Response of Industrial Radiographic Imaging Systems*. Norm: ASTM E 746-07. ASTM-International, Dez. 2007.
- [147] A. E. Burgess. „The Rose model, revisited: Brigham and Women’s Hospital and Harvard Medical School, Boston Massachusetts 02115“. In: *Optical Society of America* 16.3 (1999), S. 633–646.
- [148] K. Osterloh und U. Zscherpel. „Ways of Digital Image Processing to Obtain Fast and Unbiased Results“. In: *18th World Conference on Nondestructive Testing*. Hrsg. von SAINT. 18th World Conference on Nondestructive Testing. Durban, South Africa, 16. Apr. 2012. ISBN: 978-0-620-52872-6.
- [149] D. Kanzler, C. Müller und J. Pitkänen. „Combining of different data pools for calculating a reliable POD for real defects“. English. In: *AIP Conference Proceedings* 1650 (2015), S. 1924–1932.
- [150] J. Fan und J. S. Marron. „Fast implementations of nonparametric curve estimators“. In: *Journal of Computational and Graphical Statistics* 3.1 (1993), S. 35–56.
- [151] A. W. Bowman und A. Azzalini. *Applied Smoothing Techniques for Data Analysis. The Kernel Approach with S-Plus Illustrations*. Oxford Statistical Science Series 18. Oxford: Oxford Science Publications, 1997. ISBN: 0198523963.
- [152] D. Kanzler und C. Müller. „How much Information do we need? A reflection of the correct use of real defects in POD-evaluations“. In: *AIP Conference Proceedings* (eingereicht) (2016).
- [153] B. Jähne. *Digitale Bildverarbeitung*. 3. Aufl. Berlin: Springer Verlag, 1993. ISBN: 3-540-53768-6 2.
- [154] J. D. Hart. *Nonparametric Smoothing and Lack-of-Fit Tests*. Springer Series in Statistics. Texas: Springer Verlag, 1997. ISBN: 0-387-94980-1.
- [155] R. Schlittgen und B. H.J. Streitberg. *Zeitreihenanalyse*. 9. Aufl. Lehr- und Handbücher der Statistik. Oldenbourg, 2001. ISBN: 978-3486257250.
- [156] T. Zuva, K. Zuva, S. O. Ojo und S. M. Ngwira. „Image Representation Using Epanechnikov Density Feature Points Estimator“. In: *Signal & Image Processing : An International Journal* 4.1 (2013), S. 75–82.
- [157] E. Parzen. „On estimation of a probability density function and mode“. In: *The Annals of Mathematical Statistics* (1962), S. 1065–1076.
- [158] M. Pavlovic, K. Takahashi, C. Müller, R. Böhm und U. Ronneteg. *NDT reliability - Final report Reliability in non-destructive testing (NDT) of the canister components*. R- 08-129. Technical Report. Sweden: Svensk kärnbränslehantering AB, 2008.

- [159] C. Gollwitzer, C. Bellon, A. Deresch, U. Ewert und G.-R. Jaenisch. „On POD Estimations with Radiographic Simulator aRTist“. In: *International Symposium on Digital Industrial Radiology and Computed Tomography*. Hrsg. von Deutsche Gesellschaft für zerstörungsfreie Prüfung e.V. International Symposium on Digital Industrial Radiology and Computed Tomography. Berlin, 20. Juni 2011.
- [160] A. Rose. „Television Pickup Tubes and the Problem of Vision“. In: *Advances in Electronics and electron Physics* 1 (1948), S. 131–166.
- [161] M. Zontak und I. Cohen. „Defect Detection in Patterned Wafers Using Multichannel Scanning Electron Microscope“. In: *Signal Processing* 89 (2009), 1511–152.
- [162] DIN. *Luft- und Raumfahrt Prüfverfahren für metallische Werkstoffe Teil 3: Vergleichskörper*. Norm: DIN EN 4050-3. DIN, Aug. 2011.
- [163] D. Gohlke, D. Kotschate und T. Heckel. „Visualisierung der Volumenmessdaten von hochauflösenden Ultraschalltauchtechnikprüfungen“. In: *DGZfP-Jahrestagung*. Hrsg. von Deutsche Gesellschaft für zerstörungsfreie Prüfung e.V. DGZfP-Jahrestagung BB 127-CD. Bremen, 30. Mai 2011.
- [164] L. Piccinato, E. Jaynes, B. Hill, D. Lane, J. Kadane, M Schervish, T. Seidenfeld, R Litterman, H van Dijk und T Kloek etc. *Bayesian Inference and Decision Techniques*. Hrsg. von Prem K. Goel, Arnold Zellner und Joseph B. Kadane. Bd. 6. Essays in Honor of Bruno de Finetti. Elsevier Science Publishers, 1991, S. 496. ISBN: 0-444-87712-6.
- [165] B. G. W. Yee, F. H. Chang, J. C. Couchman und G. H Lemon. *Assessment of NDE Reliability Data*. NASA CR- 134991. Working Report. NASA, 1975.
- [166] H. Raiffa. *Decision Analysis - Introduction Lectures on Choices under Uncertainty*. Hrsg. von Harvard University. Reading, Massachusetts: Addison-Wesley, 1970. ISBN: 978-0075548669.
- [167] C. Müller, M. Bertovic, D. Kanzler und U. Ronneteg. „Conclusions of the 6th European American Workshop on Reliability of NDE“. In: *AIP Conference Proceedings* (eingereicht) (2016).
- [168] F. King, C. Lilja, K. Pedersen und P. Pitkänen. *An update of the state-of-the-art report on the corrosion of copper under expected conditions in a deep geologic repository*. TR- 10-67. SKB Technical Report. SKB, 2010.
- [169] DIN. *Schweißen und verwandte Prozesse - Begriffe für Metallschweißprozesse*. Norm: DIN EN 14610:2004. DIN, 2004.
- [170] H. Aalto. *EB-welding of the copper canister for the nuclear waste disposal*. Posiva 98-03. Forschungsbericht. Posiva Oy, 1998.

- [171] W. H. Bowyer. *A Study of Defects which Might Arise in the Copper Steel Canister*. SKI Report 00:19. Forschungsreport. SKI, 1999.
- [172] W. H. Bowyer. *Defects Which Might Occur in the Copper- Iron Canister Classified According to their Likely Effect on Canister Integrity*. SKI Report 00:21. Forschungsreport. SKI, 2000.
- [173] J. Pitkänen, R. Paussu, P. Seppälä, T. Saastamoinen, D. Kanzler und C. Müller. „Remote Visual Inspection of Disposal Canister Eb Weld“. In: *9th International Conference on NDE in Relation to Structural Integrity for Nuclear and Pressurised Components*. Hrsg. von European Commission's Joint Research Centre. BB 103-CD. Seattle, Washington, 22. Mai 2012. ISBN: 978-92-79-28212-6.
- [174] ISO. *Schweißen und verwandte Prozesse - Einteilung von geometrischen Unregelmäßigkeiten an metallischen Werkstoffen - Teil 1: Schmelzschweißen*. Norm: EN ISO 6520-1. Technischen Komitee ISO/TC 44/SC 7 "Welding und allied processes", Nov. 2007.
- [175] R. DeNale und C. A. Lebowitz. „A Comparison of Ultrasonics and Radiography for Weld Inspection“. In: *Review of Progress in Quantitative NDE* Vol. 8b (1989), S. 2003–2010.
- [176] A. M. Amaro, P. N. B. Reis, M. F. S. F. de Moura und J. B. Santos. „Damage detection on laminated composite materials using several NDT techniques“. In: *Insight* 54.1 (2012), S. 14–20.
- [177] H. Raiko, I. Meuronen, J. Pitkänen, T. Salonen, N.-C. Wikström und V.-M. Ämmälä. *EB-Demo - Canister Sealing Demonstration*. Posiva 2009-126. Working Report. Posiva Oy, 2009.
- [178] S. Sandlin. *High Energy Radiography for Inspection of the Lid Weld in Disposal Canister*. WR- 2009-82. Working Report 2009-82. Posiva Oy, 2010.
- [179] U. Zscherpel, U. Ewert und B. Redmer. *Experiments for Verification of Radiography at SKB*. DR-SKB -. Working Report. Federal Institute for Materials Research und Testing, Okt. 2004.
- [180] C. Bellon, G.-R. Jaenisch und B. Redmer. *BAM-Experiments at SKB 2005*. SKB 2005. Forschungsbericht. Federal Institute for Materials Research und Testing, 2005.
- [181] C. Müller u. a. *Reliability of Nondestructive Testing (NDT) of the Copper Canister Seal Weld*. R- 06-08. Scientific Report. Sweden: Svensk kärnbränslehantering AB, 2006.
- [182] S. Sandlin. *X-Ray Inspection Setups for the Disposal Canister Lid Weld*. Posiva 2009-98. Working Report. Eurajoki: Posiva Oy, 2010.

Literatur

- [183] J. Pitkänen, S. Sandlin, S. Nenonen und K. Banzuzi. „A new linear detector array concept for accelerator based digital radiography of the final disposal canisters for spent nuclear fuel“. In: *9th International Conference on NDE in Relation to Structural Integrity for Nuclear and Pressurised Components*. Hrsg. von European Commission's Joint Research Centre. Seattle, Washington, 22. Mai 2012. ISBN: 978-92-79-28212-6.
- [184] M. Bertovic, P. Calmon, L. Carter, J. Fischer, D. Forsyth, R. Holstein, B. McGrath, C. Müller, M. Pavlovic, U. Ronneteg, W. Rummel, F. Schubert und G. Selby. „Summary of the Open Space Technology Discussions on Reliability of NDE“. In: *Materials Testing* 56.7-8 (2014), S. 602–606.
- [185] D. Kanzler, C. Müller, M. Bertovic und J. Pitkänen. „Evaluation of RT-Systems with an Observer-POD“. In: *5th European-American Workshop on Reliability of NDE*. Hrsg. von Deutsche Gesellschaft für zerstörungsfreie Prüfung e.V. BB 144-CD. Berlin, 7. Okt. 2013.
- [186] ASTM International. *Standard Practice for Radiological Examination Using Digital Detector Arrays*. Norm: ASTM E 2698-10. ASTM International, Juli 2010.
- [187] F. W. Spencer. „Curve Fitting for Probability of Detection Data: A 4-Parameter Generalization“. In: *40th Annual Review of Progress in Quantitative Nondestructive Evaluation: Incorporating the 10th International Conference on Barkhausen Noise and Micromagnetic Testing* 1581.1 (2014), S. 2055–2062.

Abbildungsverzeichnis

| | | |
|------------|--|----|
| Abb. 1.1: | Mehrbarrierensystem für die Endlagerung von verbrauchten Brennstäben | 2 |
| Abb. 1.2: | Kanisterbauteile bestehend aus Kupferbehälter, dem Stahlköcher und den Brennstäbe | 2 |
| Abb. 2.1: | ZfP-Managementprogramm der Firma Posiva Oy | 8 |
| Abb. 2.2: | Die radiographische Defektprüfung aufgespalten in die einzelnen Prozesse von der Datenentstehung bis hin zur Entscheidung | 9 |
| Abb. 2.3: | Rausch- und Signalverteilung mit Entscheidungsschwelle und zugehörigen Fehlentscheidungen | 13 |
| Abb. 2.4: | Einfluss der Messunsicherheit auf die Anzahl der nicht detektierten Defektanzeigen | 15 |
| Abb. 2.5: | Abschätzung einer Gerade durch Datenwerte | 17 |
| Abb. 2.6: | Bestimmung der Auffindwahrscheinlichkeit an der Stelle des Defektparameters a in einem eindimensionalen Modell | 23 |
| Abb. 2.7: | Typische POD-Kurve zwischen 0 und 1 aufgetragen für den Defektparameter a | 25 |
| Abb. 2.8: | Einfluss der Datenanzahl auf den Vertrauensbereich der POD-Kurve | 27 |
| Abb. 2.9: | Bivariate Normalverteilung der Parameter des Verteilungsmodells mit Quantil-Ebene für einen festen Wert p | 29 |
| Abb. 2.10: | Schematischer Aufbau einer digitalen RT-Prüfung mit einer DDA | 31 |
| Abb. 2.11: | Beschreibung des Kontrast: Betrag der Primärstrahlung I_P und der Streuung I_S zum Kontrast C bei der Durchstrahlung eine Bauteils | 34 |
| Abb. 3.1: | Zylindrisches Koordinatensystem für eine umlaufende Schweißnaht | 47 |
| Abb. 3.2: | Die in der einfachen POD genutzte durchstrahle Länge hinsichtlich eines Detektorpixels | 48 |
| Abb. 3.3: | Aufbau eines radiographisches Bildes bestehend aus Grauwertpixel | 48 |
| Abb. 3.4: | Gegenüberstellung eines digitalisierten Schliffs und der räumlichen Darstellung einzelner Schliffe | 50 |
| Abb. 3.5: | Vergleich RT-Bild und Schliff-Rekonstruktion | 50 |

Abbildungsverzeichnis

| | | |
|------------|--|----|
| Abb. 3.6: | Gegenüberstellung von Schliff und RT-Anzeige | 54 |
| Abb. 3.7: | Observer Schwelle für modellierte FBH | 58 |
| Abb. 3.8: | Observer-POD im Vergleich zu zwei konstanten Schwellen | 58 |
| Abb. 3.9: | Glättung einer großen Anzeige ohne Rauschen | 59 |
| Abb. 3.10: | Kleinere Anzeigen könnten nach der Glättung im Rauschen untergehen . | 60 |
| Abb. 3.11: | Bedeutung der Glättungsparameter in Beziehung zur Observer- POD- Schwelle | 61 |
| Abb. 3.12: | Einfluss der Bandbreite auf den \hat{a} vs. a Graph | 62 |
| Abb. 3.13: | Einfluss der Bandbreite auf die POD-Kurve | 63 |
| Abb. 4.1: | Die vier verschiedenen Prüfverfahren zur Sicherstellung einer defektfrei- en Schweißnaht | 72 |
| Abb. 4.2: | Der Prüfaufbau mit Linearbeschleuniger, Detektor und Manipulatorsys- tem | 73 |
| Abb. 4.3: | Schematischer Aufbau und Bild des genutzten Detektors | 75 |
| Abb. 4.4: | Die Modellierung für die RT, bei der Prüfung der Schweißnaht | 75 |
| Abb. 4.5: | Referenzprüfkörper für die RT-Prüfung mit Nuten und Bohrungen | 77 |
| Abb. 4.6: | Radiographische Anzeige des Referenzprüfkörpers | 77 |
| Abb. 4.7: | Pseudo-Schliff (i), RT-Anzeige (ii) und Profile mit Glättungskurve (iii, iv) von modellierten Daten für ein 19 mm FE-Blech | 82 |
| Abb. 4.8: | Konventionelle POD-Bewertung der Modellierung einer Schweißnaht ei- nes Stahlblechs | 83 |
| Abb. 4.9: | POD-Analyse mit geglätteten Daten (Glättungsparameter $h = 2$ px) . . . | 83 |
| Abb. 4.10: | POD-Analyse mit geglätteten Daten (Glättungsparameter $h = 8,50$ px) . | 84 |
| Abb. 4.11: | Modellierungsergebnisse für die Radiographie des Kupferbehälters zum Einsatz in der Endlagerung | 86 |
| Abb. 4.12: | POD-Analyse mit geglätteten Daten der modellierten Kupfernahtprü- fung (Glättungsparameter $h = 2$ px) | 86 |
| Abb. 4.13: | POD-Analyse mit geglätteten Daten der modellierten Kupfernahtprü- fung (Glättungsparameter $h = 8,50$ px) | 87 |
| Abb. 4.14: | Pseudo-Schliff (i), modellierte RT-Anzeige (ii), Profile mit Glättungs- kurven (iii,iv) für einen künstlichen Defekt im Referenzsegment | 89 |
| Abb. 4.15: | Konventionelle POD-Bewertung mit Hilfe von künstlichen Defektdaten . | 90 |
| Abb. 4.16: | Modellannahmen, die einen Einfluss auf den Schnittpunkt von POD an der Stelle 0 aufweisen | 91 |
| Abb. 4.17: | POD-Analyse mit geglätteten Daten (Glättungsparameter $h = 4,20$ px) . | 92 |
| Abb. 4.18: | POD-Analyse mit geglätteten Daten (Glättungsparameter $h = 8,50$ px) . | 93 |
| Abb. 4.19: | Beispiel eines realen Defektes mit Schliffdaten (i), RT-Anzeige (ii) und jeweiligen Profilen (iii,iv) | 96 |

Abb. 4.20: POD-Kurve bezogen auf den jeweiligen Glättungsparameter 97

Abb. 4.21: Signal-Antwort Graph für einige reale und künstliche Datenwerte aus
der Bewertung mit einem Glättungsparameter 99

Abb. 4.22: POD-Kurven für wenige reale, für künstliche und für die kombinierte
Datenanzahl 100

Tabellenverzeichnis

| | | |
|-----------|--|----|
| Tab. 2.1: | Vier-Fehler Tabelle bei der Bewertung von Prüfsystemen | 10 |
| Tab. 3.1: | Vergleich zwischen der RT-Anzeige und dem Schliffbild | 51 |
| Tab. 4.1: | Zusammenfassung der Glättungskurven für modellierte Werte | 88 |
| Tab. 4.2: | Zusammenfassung der Glättungskurven für künstliche Defekte | 93 |
| Tab. 4.3: | Zusammenfassung der Glättungskurven für reale Defekte | 98 |
| Tab. 4.4: | Zusammenfassung der Ergebnisse | 98 |

Formelverzeichnis

| | | |
|-----------|--|----|
| Gl. 2.1: | Formel für das Gesamtrisiko zur Bewertung eines ZfP-Systemen | 11 |
| Gl. 2.2: | Formel für einen linearen Zusammenhang zwischen Defektparameter und hervorgerufenen Signalparameter | 17 |
| Gl. 2.3: | Erwartungswert für den zufälligen Fehlerterm | 18 |
| Gl. 2.4: | Streuungsmatrix für den zufälligen Fehlerterm | 18 |
| Gl. 2.5: | Wahrer Wert für die Signalantwort mit zufälligen Fehler | 18 |
| Gl. 2.6: | Wahrer Wert für den Defektparameter mit zufälligen Fehler | 18 |
| Gl. 2.7: | Gleichung für die Normalverteilung | 19 |
| Gl. 2.8: | Definition der Homoskedastizität | 19 |
| Gl. 2.9: | Definition der Likelihoodfunktion | 20 |
| Gl. 2.10: | Likelihoodfunktion einer normalverteilten Stichprobe | 20 |
| Gl. 2.11: | Bester Schätzwert bei Likelihoodfunktionen | 21 |
| Gl. 2.12: | Lokales Extremum der Likelihoodfunktionen | 21 |
| Gl. 2.13: | Definition der POD | 23 |
| Gl. 2.14: | Mittelwert der POD-Verteilung | 23 |
| Gl. 2.15: | Standardabweichung der POD-Verteilung | 23 |
| Gl. 2.16: | Definition des Vertrauensbereichs für die geschätzten Verteilungspara- meter | 28 |
| Gl. 2.17: | Unterer Vertrauensbereich für die POD-Kurve | 29 |
| Gl. 2.18: | Parameter für die Standardnormalverteilung | 29 |
| Gl. 2.19: | Parameter für den unteren Vertrauensbereich | 29 |
| Gl. 2.20: | Schwächungskoeffizient in der radiographischen Prüfung | 32 |
| Gl. 2.21: | Intensität für eine defektfreie Stelle | 33 |
| Gl. 2.22: | Definition des Kontrasts | 33 |
| Gl. 2.23: | Definition des relativen Kontrats | 34 |
| Gl. 2.24: | Allgemeine Definition des Schwächungskoeffizients | 34 |
| Gl. 2.25: | Eine Definition der Gesamtunschärfe | 37 |
| Gl. 2.26: | Definition des normierten Kontrast-Rausch-Verhältnisses | 37 |
| Gl. 3.1: | Definition des Epanechnikov-Glättungskerns | 55 |
| Gl. 3.2: | Erweiterung des Kerns mit der Bandbreite | 55 |

Formelverzeichnis

| | | |
|----------|---|----|
| Gl. 3.3: | Beschreibung der Glättung mit Hilfe des Kerns | 55 |
| Gl. 3.4: | Definition einer detektierten Anzeige nach der Glättung | 55 |
| Gl. 3.5: | Definition der POD mit zusätzlicher Abhängigkeit von der Anzeigenbreite | 61 |
| Gl. 3.6: | Formel für die gewichtete Regressionsanalyse | 66 |
| Gl. 3.7: | Definition des Gesamtgewichts | 66 |
| Gl. 3.8: | Abschätzung des Gewichts für künstliche Datenwerte | 66 |
| Gl. 3.9: | Definition des normierten Kontrast-Rausch-Verhältnisses | 66 |
| Gl. 4.1: | Erwartetes Verhalten der Signalantwort aus Sicht der Praxis | 91 |
| Gl. 4.2: | Erwartetes Verhalten der Streuung der Signalantworten aus Sicht der Praxis | 91 |
| Gl. 4.3: | Annahme der Homoskedastizität als Modellvoraussetzung | 91 |

Danksagung

“Wer einem Ertrinkenden noch so anschaulich einen Rettungsring beschreibt, wird trotzdem nicht sein Leben retten.” Walter Ludin

Dass diese Arbeit tatsächlich in der Praxis Einsatz findet, ist einer Anzahl von Personen zu verdanken:

Erst einmal möchte ich meinem Doktorvater Professor Dr. Hartmut Ewald danken, der mir die Möglichkeit gegeben hat, diese Arbeit zu erstellen und mich auch während der Arbeit mit hilfreichen Anmerkungen und Ideen begleitet hat. Außerdem danke ich der Fakultät für Informatik und Elektrotechnik in Rostock, deren Namen diese Dissertationsschrift schmücken darf. Auch Ihnen vielen Dank für die organisatorische Hilfe im Verlauf der Arbeit. Für das zweite Gutachten bin ich Professor Dr. Gehlen zu Dank verpflichtet.

Des Weiteren geht spezieller Dank an meine Projektleiterin und Mentorin Dr. Christina Müller. Ohne sie hätte ich nicht die Möglichkeit zum Schreiben dieser Arbeit bekommen. Danke, dass Du mir mit Rat und Tat zur Seite gestanden und durch viele schwierige Phasen und arbeitsintensive Stunden begleitet hast. Danke auch für die Möglichkeit mit nationalen und internationalen Wissenschaftler in Kontakt zu treten.

Ganz besonderen Dank geht an meinen leider viel zu früh verstorbenen Freund und Kollegen, dem diese Arbeit gewidmet ist: Lieber Jorma Pitkänen, noch weit bevor das erste Wort geschrieben wurde, hatten wir unzählige Diskussion und konstruktive Gespräche zu meinem Thema. Wenn eine Person stolz auf die Fortschritte innerhalb dieser Arbeit gewesen wäre, dann Du.

Auch möchte ich mich bei meinen Kollegen bedanken, die mich in der gesamten Zeit tatkräftig unterstützt haben. Im Besonderen möchte ich mich bei Dr. Gerd-Rüdiger Jaenisch bedanken, für sein offenes Ohr bei Fragen hinsichtlich Radiographie und der Erstellung von wissenschaftlichen Texten. Für seine Expertise in der Statistik möchte ich auch Prof. Dr. Spokoiny vom Weierstraß-Institut danken. Meinen Dank gilt meiner Kollegin Julia Lakämper für ihre Arbeit mit der Modellierungssoftware und Martina Rosenthal bei der Zusammenstellung der räumlichen Schliffansichten. Mein Dank geht auch an Prof. Dr. Ewert für seine Unterstützung meiner Tätigkeiten, die zur Fertigstellung dieser Arbeit geführt haben.

Vielen Dank an meine Kollegin, Freundin und Leidensgenossin Dr. Marija Bertovic für die unzähligen englischen Reviews, für die Gespräche und die Feierabendbiere, die einem wieder den Überblick erlangen ließen.

Danksagung

Außerdem möchte ich meiner Partnerin Kristin danken: Du hast mir Kraft und Ausdauer gegeben. Für ihre Unterstützung und ihren Glauben an mich möchte ich auch meiner Familie danken, genau wie vielen Freunde und Kollegen, die mich auf diesem Weg unterstützt haben.