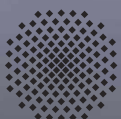


STUTTGARTER BEITRÄGE ZUR PRODUKTIONSFORSCHUNG

JOACHIM NEHER

Neuro-Fuzzy-Modellierung zur umfassenden Prozessüberwachung am Beispiel des Ultra- schallschweißens von Kunststoffteilen



Universität Stuttgart



Fraunhofer

Herausgeber:

Univ.-Prof. Dr.-Ing. Thomas Bauernhansl

Univ.-Prof. Dr.-Ing. Dr. h. c. Alexander Verl

Prof. Dr.-Ing. Prof. e. h. Dr.-Ing. e. h. Dr. h. c. mult. Engelbert Westkämper

Joachim Neher

**Neuro-Fuzzy-Modellierung zur umfassenden
Prozessüberwachung am Beispiel des Ultraschall-
schweißens von Kunststoffteilen**

Kontaktadresse:

Fraunhofer-Institut für Produktionstechnik und Automatisierung IPA, Stuttgart
Nobelstraße 12, 70569 Stuttgart
Telefon 07 11 9 70-00, Telefax 07 11 9 70-13 99
info@ipa.fraunhofer.de, www.ipa.fraunhofer.de

STUTTGARTER BEITRÄGE ZUR PRODUKTIONSFORSCHUNG**Herausgeber:**

Univ.-Prof. Dr.-Ing. Thomas Bauernhansl
Univ.-Prof. Dr.-Ing. Dr. h. c. Alexander Verl
Prof. Dr.-Ing. Prof. e. h. Dr.-Ing. e. h. Dr. h. c. mult. Engelbert Westkämper

Fraunhofer-Institut für Produktionstechnik und Automatisierung IPA, Stuttgart
Institut für Industrielle Fertigung und Fabrikbetrieb (IFF) der Universität Stuttgart
Institut für Steuerungstechnik der Werkzeugmaschinen und Fertigungseinrichtungen (ISW)
der Universität Stuttgart

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.
ISBN: 978-3-8396-0424-3

D 93

Zugl.: Stuttgart, Univ., Diss., 2011

Druck: Mediendienstleistungen des
Fraunhofer-Informationszentrum Raum und Bau IRB, Stuttgart

Für den Druck des Buches wurde chlor- und säurefreies Papier verwendet.

© by **FRAUNHOFER VERLAG**, 2012

Fraunhofer-Informationszentrum Raum und Bau IRB
Postfach 80 04 69, 70504 Stuttgart
Nobelstraße 12, 70569 Stuttgart
Telefon 07 11 9 70-25 00
Telefax 07 11 9 70-25 08
E-Mail verlag@fraunhofer.de
URL <http://verlag.fraunhofer.de>

Alle Rechte vorbehalten

Dieses Werk ist einschließlich aller seiner Teile urheberrechtlich geschützt. Jede Verwertung, die über die engen Grenzen des Urheberrechtsgesetzes hinausgeht, ist ohne schriftliche Zustimmung des Verlages unzulässig und strafbar. Dies gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen sowie die Speicherung in elektronischen Systemen.

Die Wiedergabe von Warenbezeichnungen und Handelsnamen in diesem Buch berechtigt nicht zu der Annahme, dass solche Bezeichnungen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und deshalb von jedermann benutzt werden dürften. Soweit in diesem Werk direkt oder indirekt auf Gesetze, Vorschriften oder Richtlinien (z.B. DIN, VDI) Bezug genommen oder aus ihnen zitiert worden ist, kann der Verlag keine Gewähr für Richtigkeit, Vollständigkeit oder Aktualität übernehmen.

GELEITWORT DER HERAUSGEBER

Produktionswissenschaftliche Forschungsfragen entstehen in der Regel im Anwendungszusammenhang, die Produktionsforschung ist also weitgehend erfahrungsbasiert. Der wissenschaftliche Anspruch der „Stuttgarter Beiträge zur Produktionsforschung“ liegt unter anderem darin, Dissertation für Dissertation ein übergreifendes ganzheitliches Theoriegebäude der Produktion zu erstellen.

Die Herausgeber dieser Dissertations-Reihe leiten gemeinsam das Fraunhofer-Institut für Produktionstechnik und Automatisierung IPA und jeweils ein Institut der Fakultät für Konstruktions-, Produktions- und Fahrzeugtechnik an der Universität Stuttgart.

Die von ihnen betreuten Dissertationen sind der marktorientierten Nachhaltigkeit verpflichtet, ihr Ansatz ist systemisch und interdisziplinär. Die Autoren bearbeiten anspruchsvolle Forschungsfragen im Spannungsfeld zwischen theoretischen Grundlagen und industrieller Anwendung.

Die „Stuttgarter Beiträge zur Produktionsforschung“ ersetzt die Reihen „IPA-IAO Forschung und Praxis“ (Hrsg. H.J. Warnecke / H.-J. Bullinger / E. Westkämper / D. Spath) bzw. ISW Forschung und Praxis (Hrsg. G. Stute / G. Pritschow / A. Verl). In den vergangenen Jahrzehnten sind darin über 800 Dissertationen erschienen.

Der Strukturwandel in den Industrien unseres Landes muss auch in der Forschung in einen globalen Zusammenhang gestellt werden. Der reine Fokus auf Erkenntnisgewinn ist zu eindimensional. Die „Stuttgarter Beiträge zur Produktionsforschung“ zielen also darauf ab, mittelfristig Lösungen für den Markt anzubieten. Daher konzentrieren sich die Stuttgarter produktionstechnischen Institute auf das Thema ganzheitliche Produktion in den Kernindustrien Deutschlands. Die leitende Forschungsfrage der Arbeiten ist: Wie können wir nachhaltig mit einem hohen Wertschöpfungsanteil in Deutschland für einen globalen Markt produzieren?

Wir wünschen den Autoren, dass ihre „Stuttgarter Beiträge zur Produktionsforschung“ in der breiten Fachwelt als substantiell wahrgenommen werden und so die Produktionsforschung weltweit auf ein neues Niveau heben.

Alexander Verl

Thomas Bauernhansl

Neuro-Fuzzy-Modellierung zur umfassenden
Prozessüberwachung am Beispiel des Ultraschallschweißens von
Kunststoffteilen

Von der Fakultät Konstruktions-, Produktions- und Fahrzeugtechnik
der Universität Stuttgart
zur Erlangung der Würde eines Doktor-Ingenieurs (Dr.-Ing.)
genehmigte Abhandlung

Vorgelegt von
Joachim Neher
aus Sindelfingen

Hauptberichter:	Prof. Dr.-Ing. Dr. h. c. Alexander Verl
Mitberichter:	Prof. Dr.-Ing. Arnold Kistner
Tag der Einreichung:	27. April 2011
Tag der mündlichen Prüfung:	15. Dezember 2011

Institut für Steuerungstechnik der Werkzeugmaschinen und Fertigungseinrichtungen (ISW)
der Universität Stuttgart

2011

Vorwort des Autors

Die vorliegende Arbeit entstand während meiner Tätigkeit am Fraunhofer Institut für Produktionstechnik und Automatisierung (IPA) in Stuttgart.

Herrn Prof. Dr.-Ing. Dr. h. c. Alexander Verl danke ich für seine wohlwollende Unterstützung und Förderung meiner Arbeit. Herrn Prof. Dr.-Ing. Arnold Kistner danke ich für die Übernahme des Mitberichts und die Impulse, die ich durch ihn bereits im Rahmen meines Studiums der Technischen Kybernetik an der Universität Stuttgart erhalten habe.

Ein herzliches Dankeschön gebührt auch Herrn Prof. Dr.-Ing. Günther Fischer von der Hochschule Esslingen. Von ihm habe ich sehr viel über Kunststofftechnik gelernt; durch die gemeinsamen Projekte hat er maßgeblich zum Gelingen der Arbeit beigetragen.

Ein ganz großer Dank gebührt Herrn Dipl.-Ing. Ernst Schmidberger. Er war für mich eine Art Mentor und hat durch seine Arbeit und die Entwicklung des NEPRES-Systems den Grundstein für diese Arbeit gelegt. Auch sonst habe ich viel von ihm gelernt und die gemeinsame Arbeit am Fraunhofer IPA hat meinen eingeschlagenen Weg nachhaltig geprägt.

Nicht unerwähnt lassen möchte ich auch meine anderen Kollegen der Abteilung Technische Informationsverarbeitung vom Fraunhofer IPA, die für eine offene und inspirierende Arbeitsatmosphäre verantwortlich zeichnen. Besonders zu erwähnen sind meine Kollegen Marius Pflüger und Markus Kaupp, mit denen ich mich zu jeder Zeit austauschen und über neue Ansätze oder über Probleme diskutieren konnte. Von den Studenten, die bei der Software-Entwicklung oder an Versuchsreihen mitgewirkt haben, gilt mein Dank insbesondere Christian Heigele für seine langjährige und stets bereichernde Arbeit.

Last but not least hat auch meine Freundin Karoline ihren erheblichen Anteil beigetragen, indem sie viele Abende und Wochenenden auf mich verzichten musste.

Sindelfingen, im Dezember 2011

Joachim Neher

Kurzinhalt

In Deutschland rechnet sich eine Fertigung oft nur noch für komplexe Produkte. Die benötigten Fertigungsprozesse sind zunehmend automatisiert und verkettet. In gleichem Maße steigt der Bedarf an Systemen zur Prozessüberwachung.

Diese Arbeit befasst sich mit dem Aufbau eines Prozessüberwachungssystems auf der Basis von erfassten Sensordaten aus dem Fertigungsprozess. Einerseits ermöglicht das System eine Qualitätsprognose. Andererseits können geänderte oder unbekannte Zustände, die ihre Ursache in Veränderungen im überwachten Prozess oder in Prozessen der vorausgehenden Prozesskette haben können, erkannt werden.

Kernelement dabei ist die Prozessmodellierung auf Basis von „general regression neuro-fuzzy networks“ (GRNFN), welche klassische Neuronale Netze mit Elementen von Fuzzy-Systemen verbinden. Die Eigenschaften der Modellstruktur werden detailliert untersucht. Anschließend werden die GRNFN-Modelle um einige neue Fähigkeiten erweitert. So ist es möglich, neben dem sonst auf Versuchsdaten basierenden Trainingsprozess auch vorhandenes Expertenwissen in Form von Fuzzy-Regeln direkt zu integrieren. Des Weiteren werden Indikatoren zum Erkennen unbekannter Zustände erarbeitet.

Durch systematische Versuchsreihen sowie die Analyse von Signalvarianzen können Kenngrößen abgeleitet werden, welche in Form der sogenannten Zustandsvektoren einen Fingerabdruck für jeden Fertigungszyklus bilden. Zudem werden verschiedene Methoden beschrieben, die der Optimierung der erzielbaren Modellgüte dienen.

Am Beispiel des Ultraschallschweißens wird das entwickelte System zur Prozessüberwachung sowohl an einem speziell entwickelten Probekörper als auch an Praxisbauteilen verifiziert und die Leistungsfähigkeit nachgewiesen. Neben der Prognose der Schweißnahtqualität können auch Änderungen des Werkstoffs, Variationen des vorgelagerten Spritzgießprozesses sowie Veränderungen des Feuchtegehalts, die auf eine veränderte Lagerung zurückzuführen sind, sicher erkannt werden.

Short summary

In Germany, manufacturing often is economic only for complex products. The required manufacturing processes are increasingly automated and chained up. Equally an increasing necessity for systems for process monitoring arises.

This work deals with the development of a process monitoring system based on sensor data collected from the manufacturing process. On the one hand, the system facilitates a quality prediction. On the other hand, changed or unknown conditions can be detected, that are caused either by changes in the monitored process or even in processes in the preceding process chain.

Core element of this framework are „general regression neuro-fuzzy networks” (GRNFN) that combine classical neural networks with elements of fuzzy systems. The properties of the model structure are examined in detail. Then the GRNFN models are expanded by several new capabilities. So it becomes possible to integrate existing expertise in the form of fuzzy rules directly into the model, which normally is trained by experimental data. Furthermore, indicators are developed for the detection of unknown states.

Through systematic experiments and the analysis of signal variance characteristics can be derived, which form the so-called state vectors and thereby form a kind of fingerprint for each production cycle. In addition, various methods are described that serve to optimize the achievable model accuracy.

Using the example of the ultrasonic welding, the developed system for process monitoring is verified and its performance is proven. This is done on a specially developed test specimens as well as on work pieces out of industrial production. In addition to the prediction of weld quality also changes of material, variations of the preceding injection molding as well as changes in moisture content, which can be induced by changes in storage, are detected reliably.

Inhaltsverzeichnis

Abkürzungsverzeichnis	xv
Abbildungsverzeichnis	xix
Tabellenverzeichnis	xxii
1 Einleitung	1
1.1 Ausgangssituation	1
1.1.1 Prozessablauf beim Ultraschallschweißen	3
1.1.2 Eigenschaften des Ultraschallschweißens	5
1.2 Problemstellung	8
1.3 Zielsetzung und Lösungsansatz	11
1.4 Aufgabenstellung und Vorgehensweise	14
2 Stand der Technik	17
2.1 Ultraschallschweißen: Maschinenteknik, Ablauf und Einflüsse	19
2.1.1 Maschinenteknik beim Ultraschallschweißen	20
2.1.2 Prozessablauf	23
2.1.3 Einflussgrößen	25
2.1.4 Prozessphasen im Schweißweg	36
2.2 Methoden zur Prozessüberwachung	37
2.2.1 Statistische Prozessüberwachung	39
2.2.2 Kontinuierliche Prozessüberwachung	41
2.3 Prozessüberwachung beim Ultraschallschweißen	45
2.3.1 Qualitätsmerkmale bei Schweißteilen	45
2.3.2 Bekannte Ansätze und Lösungen zur Überwachung	49
2.3.3 Verwendete Probekörpergeometrien	58
2.4 Modellierungsmethoden	61
2.4.1 Mathematisch-physikalische Modelle	61
2.4.2 Datenbasierte Modelle	62
2.4.3 Fuzzy-Systeme	75
2.4.4 Kombination von regel- und datenbasierten Methoden	78

2.5	Exemplarische Beispiele zur Prozessüberwachung	83
2.5.1	Polierprozesse in der Waferfertigung	83
2.5.2	Kunststoffspritzgießen	84
2.5.3	Reaktionsgießen von Polyurethanen	85
2.5.4	Klassifikation von Texturen	86
2.5.5	Gesenkschmieden und Stanzprozesse	87
2.5.6	Fügen von Blechen	88
2.5.7	Weitere Schweißverfahren	88
2.6	NEPRES-Systemtechnologie	90
2.6.1	Aufbau der Systemtechnik	90
2.6.2	Kenngößen	92
2.6.3	Modellierung	94
2.7	Fazit	94
3	Detaillierter Lösungsansatz	96
3.1	Signaldatenerfassung	96
3.1.1	Signalquellen	96
3.1.2	NEPRES-Mini-System	99
3.2	Versuchsanlage Probekörper	99
3.3	Definition eines Versuchsteils	101
3.4	Versuchsplan und Durchführung	103
3.5	Signalverarbeitung und Auswertung	106
3.5.1	Signalanalyse und Kenngößenbildung	106
3.5.2	Varianzanalysen und Kenngößenselektion	107
3.6	Modellbildung	108
3.7	Validierung und Umsetzung an Praxisbauteilen	109
3.8	Versuchsdaten zur Ergebnisdarstellung	109
4	Untersuchte Zustandsvektoren und Modellbildung	112
4.1	Projektion von Zustandsvektoren	112
4.1.1	PCA	113
4.1.2	Sammons-Mapping	114
4.2	GRNFN-Modelle: Aufbau und Training	117
4.2.1	Modellstruktur	117
4.2.2	Trainingsphase	119
4.3	Einfluss der Aktivierungsfunktion	123
4.3.1	Allgemeine Glockenkurve	124
4.3.2	Weitere Aktivierungsfunktionen	126

4.4	Eigenschaften von GRNFN-Modellen	131
4.4.1	Rechen- und Speicheraufwand	131
4.4.2	Robustheit und Fehlertoleranz	132
4.4.3	Interpretation als Fuzzy-Regelbasis	135
4.4.4	Gezielte Erweiterung der Fuzzy-Regelbasis	136
4.4.5	Zusammenfassen von Trainingsdaten	137
4.4.6	Dynamische Modellanpassung	138
4.5	Prognosegüte	139
4.6	Unbekannt-Klassifikation und Störungserkennung	143
4.6.1	Indikatoren und Kenngrößen für die Unbekannt-Klassifikation	145
4.6.2	Auswirkung des Nachtrainings	152
4.7	Kriterien für die getroffene Wahl der GRNFN-Modelle	154
5	Realisierte Signalverarbeitung	156
5.1	Visualisierung der Signale	156
5.2	Streuungsanalyse	159
5.2.1	Streuung der Prozesssignale	159
5.2.2	Ableitung von Kenngrößen	161
5.2.3	Untersuchung der Kenngrößenvarianz	165
5.3	Optimierung der Zustandsvektoren	169
5.3.1	Normierung und Skalierung	169
5.3.2	Korrelationsanalyse	170
5.3.3	Mittelwertvergleich und Signifikanzanalyse	176
5.3.4	Kenngrößenselektion durch Genetische Algorithmen	187
5.4	Fazit	193
6	Anwendung der realisierten Systemtechnik beim Ultraschallschweißen	195
6.1	Ergebnisse zum Körperschall	195
6.2	Ergebnisse mit Probekörpern	198
6.2.1	Überwachen der Prozesseinstellung	199
6.2.2	Prognose von Qualitätsmerkmalen durch Prozessmodelle	199
6.2.3	Variation des Spritzgießens	207
6.2.4	Veränderungen des Werkstoffs	209
6.2.5	Feuchtigkeitsaufnahme der Schweißteile durch Lagerung	211
6.3	Prognoseergebnisse an Praxisbauteilen	215
6.3.1	Einfluss eines veränderten Spritzgießprozesses	217
6.3.2	Erkennen von Materialänderungen bei einem Rückschlagventil	220
6.3.3	Prozessstörung und Arbeitspunktnachführung	223
6.3.4	Erkennen fehlender Einlegeteile	225

6.3.5	Qualitätsprognose unterschiedlicher Merkmale für Wasserverteiler . .	227
6.4	Zusammenfassung und Fazit der Anwendungen	229
7	Zusammenfassung und Ausblick	230
7.1	Zusammenfassung	230
7.2	Fazit und Ausblick	238
8	Summary	241
	Literaturverzeichnis	243

Abkürzungsverzeichnis

ABS	Acrylnitril-Butadien-Styrol
AiF	Arbeitsgemeinschaft industrieller Forschungsvereinigungen
ANFIS	adaptive neuro-fuzzy inference system
ANOVA	analysis of variance, Varianzanalyse
AP	Arbeitspunkt
ASI	American Supplier Institute
AWS	American Welding Society
CMP	Chemisch-mechanisches Polieren
CPC	continuous process control, kontinuierliche Prozessüberwachung
DoE	Design of Experiments, statistische Versuchsplanung
DSP	Digitaler Signalprozessor
DVS	Deutscher Verband für Schweißtechnik
ERG	Energierichtungsgeber
FEM	Finite-Elemente-Methode
FFT	Fast-Fourier-Transformation
FMEA	Fehlermöglichkeits- und Einflussanalyse
FT	Fourier-Transformation
GA	Genetischer Algorithmus
GF	Glasfaser(anteil)
GRNFN	general regression neural-fuzzy network
GRNN	general regression neural network

IBV	Industrielle Bildverarbeitung
IKP	Institut für Kunststoffprüfung und Kunststoffkunde
IKT	Institut für Kunststofftechnologie
iO	in Ordnung
IPA	Institut für Produktionstechnik und Automatisierung
IQR	Inter-Quartils-Range
IR	Infrarot
JE-Netz	Jordan-Elman-Netz
KI	Künstliche Intelligenz
kNN	k nächste Nachbarn
LSE	least square estimator
LVQ	Learning Vector Quantization
MLP	multi-layer perceptron, mehrschichtiges Perzeptron
MVR	melt volume rate
ND	Nachdruck
NEFCLASS	neuro-fuzzy classification
NEFCON	neuro-fuzzy control
NEPRES	Neuronales Prozess-Regelungs-System
NEPRES-Mini	mobile NEPRES-Plattform auf Laptop-Basis mit USB-Messkarte
niO	nicht in Ordnung
NN	neural network, Neuronales Netz
PA	Polyamid
PC	Polycarbonat
PCA	principal component analysis
PE	Polyethylen
PE-HD	Polyethylen hoher Dichte (high density)

PE-LD	Polyethylen niedriger Dichte (low density)
PK	Prozesskenngröße
PMMA	Polymethylmethacrylat (Plexiglas)
POM	Polyoxymethylen
PP	Polypropylen
PPA	Polyphthalamid
PS	Polystyrol
PTFE	Polytetrafluorethylen (Teflon)
PVC	Polyvinylchlorid
QFD	Quality Function Deployment
QK	Qualitätskenngröße
QN	Quetschnaht
QRK	Qualitätsregelkarte
QS	Qualitätssicherung
RBF	radial base function, radiale Basisfunktion
rF	relative Feuchte, Feuchtegehalt
RPN	Reference Point Numeric – Triggerpunkt beim Ultraschallschweißen
SAN	Styrol-Acrylnitril
SG	Spritzgießen
SNNS	Stuttgarter Neuronale Netze Simulator
SNR	signal-to-noise-ratio, Signal-Rausch-Verhältnis
SOM	Self-Organizing Maps
SPC	statistical process control, statistische Prozessüberwachung/-regelung
SVM	support vector machine
SVP	Statistische Versuchsplanung
TQM	Total Quality Management

TSK	Takagi-Sugeno-Kang
TTL	Transistor-Transistor-Logik
US	Ultraschall / Ultraschallschweißen
ZVEI	Zentralverband der Elektrotechnischen Industrie

Abbildungsverzeichnis

1.1	Aufbau und Komponenten einer Ultraschallschweißmaschine	2
1.2	Longitudinale Schwingungen des Schwinggebildes beim Ultraschallschweißen	4
1.3	Einzelteile und Schmelzebildung	5
1.4	Korrelation zwischen Prozesskenngrößen und dem Berstdruck	11
1.5	Modellbasierte Prozessüberwachung beim Ultraschallschweißen (Lösungsansatz)	13
2.1	Schwingungsamplituden der Sonotrode mit und ohne FEM-Optimierung . . .	21
2.2	Schubmodul und mechanischer Verlustfaktor für versch. Werkstoffe	27
2.3	Veränderung der Werkstoffeigenschaften von PA 66	28
2.4	Verschiedene Formen der Schweißzonengestaltung	30
2.5	Einflussparameter beim Ultraschallschweißen	35
2.6	Prozessphasen im Schweißweg	36
2.7	Verschiedene Probekörpergeometrien	57
2.8	Approximation einer eindimensionalen Funktion durch ein RBF-Netz	65
2.9	Voronoi-Diagramm	67
2.10	Entwicklung der Abdeckung eines rechteckigen Eingangsraums durch ein SOM	68
2.11	Verschiedene Aktivierungsfunktionen für Neuronen	70
2.12	Aufbau eines mehrschichtigen Perzeptrons mit Feed-Forward-Architektur . .	71
2.13	Aufbau eines Fuzzy-Systems	75
2.14	Vorgehen bei Fuzzy-Systemen (nach MAMDANI) zur Ableitung der Zielgröße	77
2.15	ANFIS-Modell (Typ 3 für TAKAGI-SUGENO-Fuzzysystem)	79
2.16	NEFCLASS-Modell mit zwei Eingängen, fünf Regeln und zwei Klassen . . .	80
2.17	Interpolierendes Neuro-Fuzzy-Modell	81
2.18	NEPRES-Prinzip	91
2.19	Originalsignal und Rekonstruktion mittels PCA-Kenngrößen	92
2.20	Adaptive Hüllkurven in NEPRES	93
3.1	Versuchsanlage mit Sensorik und System zur Messwerterfassung	100
3.2	Zyklusbezogene Signale eines Schweißzyklus	101
3.3	Realisierter Probekörper als Basis für Versuchsreihen	102
3.4	Prüfeinrichtungen für die Torsions-, Zug- und Berstdruckprüfung	104

4.1	Prinzip des Sammons-Mapping zur Projektion von N-dimensionalen Vektoren	114
4.2	Sammons-Mapping für den Einheitswürfel mit Zusatzpunkten	115
4.3	Verzerrungsbedingte Klassenüberlappung beim Sammons-Mapping	116
4.4	Architektur von general regression neural-fuzzy networks (GRNFN)	118
4.5	Einfluss der Parameter auf die Form der Glockenkurve	121
4.6	Verteilung der resultierenden Fuzzy-Zugehörigkeitsfunktionen	122
4.7	Modellgüte für die allg. Glockenkurve in Abhängigkeit der Parameter σ und b	125
4.8	Gaußglocke als Aktivierungsfunktion	127
4.9	Dreiecksfunktion als Aktivierungsfunktion	128
4.10	Trapezfunktion als Aktivierungsfunktion	130
4.11	Veränderung der Modellgüte bei ungültigen Eingängen	133
4.12	GRNFN-Modell mit manuell ergänzten (Fuzzy-)Regeln	136
4.13	Sammons-Mapping der Beispieldaten mit unterschiedlichen Zielgrößen	139
4.14	Vergleich zwischen Einstellung und GRNFN-Prognose (Lern- und Testdaten)	140
4.15	Prognoseergebnisse bei geänderter Codierung	142
4.16	Unbekannt-Klassifikation bei Training der Einstellungen A und C	144
4.17	Unbekannt-Klassifikation bei Training der Einstellungen B, D und E	146
4.18	Aktivierungswerte φ_i^k der einzelnen Komponenten für verschiedene Eingangs- vektoren	150
4.19	Distanzmaße für Training A und C mit zunehmendem Anteil an B	153
5.1	Abfolge verschiedener Darstellungsformen	157
5.2	Einstellungsabhängige Streuung der einzelnen Maschinensignale	160
5.3	Grafische Darstellung der Kenngrößenwerte über die Versuchszyklen	166
5.4	Verschiedene Verteilungskurven und zugeordnete Boxplots	167
5.5	Grafische Darstellung der Verteilung aller Kenngrößen (Farbhistogramme) .	168
5.6	Darstellung des Einflusses der Normierung	170
5.7	(Synthetische) Messreihen für die Korrelations- und Signifikanzberechnung .	171
5.8	(Schein)Korrelation bei Kenngrößen und Qualität für Probekörper aus SAN .	174
5.9	Verlauf exemplarischer Kenngrößen sowie variierte Einstellung	175
5.10	Verlauf des Berstdrucks und statistische Größen verschiedener Ausschnitte .	177
5.11	Signifikanzniveaus von Kenngrößen	183
5.12	Auswirkung der Kenngrößenauswahl und Potenzial der Signifikanzanalyse . .	184
5.13	Ergebnisse der Signifikanzanalyse (verschiedene Niveaus und Voting-Anteile)	188
6.1	Körperschallsignale ausgewählter Zyklen beim ERG-Probekörper aus PS . .	196
6.2	Zustandsvektor aus normiertem Amplitudenspektrum des Körperschalls . . .	198
6.3	Probekörper mit Quetschnahtgeometrie: Lerndaten für Festigkeitsprognose .	200
6.4	Probekörper mit Quetschnahtgeometrie: Qualitätsprognose für Berstdruck .	203

6.5	Qualitätsprognose für den Berstdruck mittels verschiedener MLP-Modelle . .	204
6.6	Auswirkung geänderter Spritzgießparameter (Probekörper PA-GF66, ERG) .	208
6.7	Schweißen von Probekörpern aus ABS und ABS-PS-Blend	209
6.8	Einfluss des Feuchtegehalts der Werkstücke	211
6.9	Prognose des Feuchtegehalts: Zustandsvektoren und Unbekannt-Indikatoren	213
6.10	Erfassung des Temperaturverlaufs mittels IR-Temperatursensoren	216
6.11	Schweißen von Automobilleuchten unterschiedlicher Gewichte	219
6.12	Verschweißen eines Rückschlagventils mit vertauschtem Material	220
6.13	Kenngrößenselektion für das Rückschlagventil zur Identifikation gelagerter Teile	222
6.14	Anschweißen eines Tasters mit Prozessänderung und Arbeitspunktanpassung	224
6.15	Verschweißen eines Wasserverteilers mit Filtereinsatz	226
6.16	Qualitätsprognose für einen Wasserverteiler (varierte Schweißparameter) . .	227

Tabellenverzeichnis

3.1	Versuchsreihe für Modellverifikation	109
3.2	Signifikanzwerte für den Vergleich des Berstdrucks bzgl. Prozesseinstellung	110
4.1	Min- und Maxwerte aller Datenzyklen pro Einstellung	147
4.2	Kenngößen zur Unbekannt-Klassifikation entsprechend Abbildung 4.18	149
5.1	Korrelations- und Signifikanzwerte zu Abbildung 5.7	173
5.2	Konfidenzintervalle und kritische Student-t-Werte zu Abbildung 5.10	180
6.1	Ergebnisse der Qualitätsprognose für GRNFN- und MLP-Modelle	205

1 Einleitung

1.1 Ausgangssituation

Im Zuge der Globalisierung haben sich während der letzten Jahre die Randbedingungen für produzierende Unternehmen deutlich verändert. Seit der Öffnung der Märkte in Osteuropa und dem wirtschaftlichen Wachstum in Indien und China gibt es dort viele Produzenten, die Produkte deutlich günstiger herstellen können als einheimische Betriebe. Insbesondere China hat sich durch die Verfügbarkeit preiswerter Arbeitskräfte zur verlängerten Werkbank für viele Firmen und Produkte entwickelt. Mit den hiesigen Kostenstrukturen – neben Lohn- und Lohnnebenkosten spielen auch andere Abgaben sowie hohe Energie- und Umweltkosten eine nicht zu unterschätzende Rolle – können Firmen ohne Änderungen gegen die neuen Märkte nicht bestehen, so dass die letzten Jahre eine starke Verlagerung der Produktion insbesondere im Bereich von Consumer-Produkten stattgefunden hat. Sichtbar ist dies unter anderem an den bis zum Jahr 2008 stark gestiegenen internationalen Frachtraten im Containerverkehr, wobei die Schifffrachtbranche über 90 % des internationalen Warenverkehrs abwickelt. Die Zukunftschancen für deutsche und europäische Firmen liegen darin, sich auf hochwertige Produkte zu konzentrieren, zu deren Herstellung entsprechend komplexe Maschinen und Prozesse mit angepasstem Prozess-Know-how erforderlich sind.

Im Vergleich zu Low-Tech-Produkten stellen solche Highvalue-Produkte auch durch die zunehmende Miniaturisierung und steigende Komplexität deutlich höhere Anforderungen an die Produzenten. Auch die zunehmende Funktionsintegration trägt ihren Teil dazu bei. Viele dieser Produktinnovationen sind erst durch neue Entwicklungen aus dem Bereich der Werkstoffe – insbesondere bei den Kunststoffen – ermöglicht worden, oft in Verbindung mit entsprechenden Fortschritten bei Verarbeitungsmethoden und -möglichkeiten. In vielen Bereichen werden „klassische Werkstoffe“ im Zuge einer Neugestaltung von Produkten durch Kunststoffpendants ersetzt oder bei Neuentwicklungen von Produkten konsequent aus Kunststoffen realisiert [Erh08, Fle91]. Neue oder verbesserte Kunststoffwerkstoffe erfüllen dabei dieselben Anforderungen bei deutlich geringerem Gewicht und bieten durch existierende Verarbeitungsmethoden – z. B. Spritzgießen (SG) und Ultraschallschweißen (US) – neue Gestaltungsmöglichkeiten. Beispiele hierfür sind der Automobil- und Kommunikationssektor (z. B. komplexe Beleuchtungsbaugruppen als Kompletmodul, Handys)

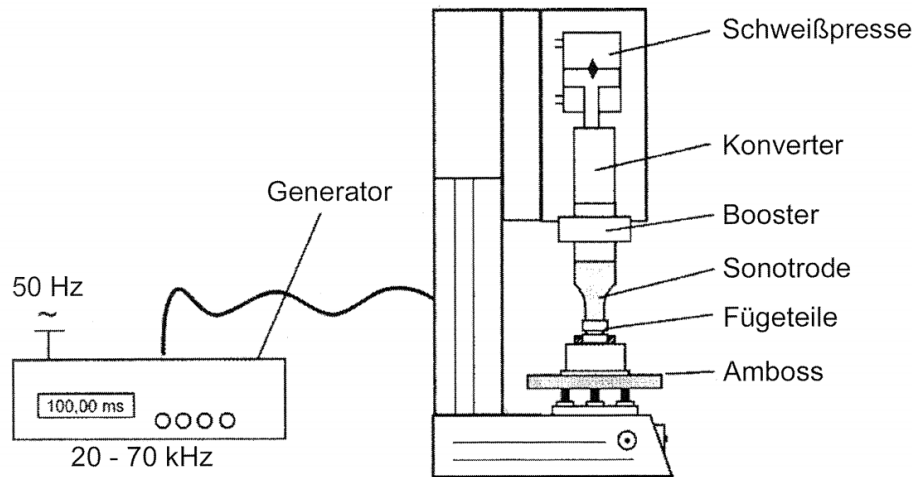


Abbildung 1.1: Aufbau und Komponenten einer Ultraschallschweißmaschine (aus [Ehr04])

oder Produkte aus dem Elektro- und Medizinbereich (z. B. mikrostrukturierte Sensorelemente mit optischen Kunststoffkomponenten, Dosiereinrichtungen). Durch die oft kürzeren Bearbeitungszeiten kann auch eine Steigerung der Effektivität und Wirtschaftlichkeit erzielt werden. Die vorwiegend mittelständisch geprägte kunststoffverarbeitende Industrie mit ca. 2900 Betrieben hatte 2009 in Deutschland einen Gesamtumsatz von knapp 46 Milliarden € [Pfe10].

Mit höheren Anforderungen an die Produkte geht eine Steigerung der Komplexität der Prozesse und des notwendigen Prozess-Know-hows einher. Gleichzeitig steigt der Aufwand, der notwendig ist, um die Prozesse abzusichern und die Produktqualität – als wesentlichen Wettbewerbsfaktor – zu gewährleisten. Insbesondere trifft dies bei sicherheitsrelevanten Teilen oder Baugruppen wie z. B. Hochdruckventilen, mechatronischen Sensor-Aktor-Elementen oder medizintechnischen Einheiten wie Dialysatoren zu. Durch die zunehmende Miniaturisierung und Funktionsintegration können auch geringe Prozessschwankungen, die bei jedem Prozess auftreten und unvermeidbar sind [Mas80, Pfe96], signifikante Auswirkungen haben und zum kompletten Ausfall oder einer Spezifikationsverletzung des Bauteils führen. Dieser Tatsache wird zum einen damit begegnet, dass für die Produktion leistungsfähigere und präzisere Maschinen (z. B. mit Linearantrieben oder geregelten Achsen) verwendet werden. Zum anderen wird auch die Prozessüberwachung in letzter Zeit deutlich ausgebaut. Als Teil der Prozessüberwachung wird meist auch die eingesetzte Qualitätssicherung und der notwendige Prüfaufwand verstanden, der mit zunehmender Komplexität der Teile deutlich ansteigt. Um die Qualität abzusichern, wird zunehmend auch der Aufbau einer automatisierten 100 %-Prüfung angestrebt. Dies ist unter anderem daran erkennbar, dass der Bereich der industriellen Bildverarbeitung (IBV) die letzten Jahre stetige Wachstumsraten meist im zweistelligen Prozentbereich aufweisen konnte; im Jahr 2008 stieg der Branchenumsatz um 5 % auf 1,2 Milliarden Euro [Sch09, Spa05]. Überwachungsmetho-

den, die direkt in den Fertigungsprozess integriert sind und nicht „nur“ anschließend prüfen, weisen größere Potenziale auf. Neben der Erkennung von fehlerhaften Teilen können teils auch langsame Prozessdriften identifiziert und diesen direkt entgegengewirkt werden, bevor Fehler entstehen [Wes97]. Solche prozessintegrierte Überwachungsmethoden basieren heutzutage oftmals auf der Grenzwertüberwachung einzelner Prozesskenngrößen oder einfachen Hüllkurvenmodellen. In einigen Bereichen wurden auch effektivere, aber aufwändigere und komplexere Methoden entwickelt, welche versuchen, die Prozesse mittels Modellen abzubilden und somit entsprechende genauere Aussagen treffen zu können [Bis06, HTF01]. Solche Lösungen haben sich jedoch – auch aufgrund des notwendigen Adaptionaufwands – noch nicht in der Breite durchgesetzt. Deshalb muss angestrebt werden, andere leistungsfähige und dennoch kostenmäßig darstellbare Methoden zur Prozessüberwachung einzusetzen. Dieses Bemühen sollte sich auf möglichst viele Anwendungsbereiche erstrecken, also auch auf die Kunststofftechnik und andere zyklische Produktionsprozesse.

1.1.1 Prozessablauf beim Ultraschallschweißen

Neben der Herstellung von Werkstücken und Halbzeugen durch Ur- und Umformprozesse gibt es eine Fülle von Verfahren zur Verbindung von Bauteilen. Neben den klassischen Verbindungsverfahren wie Schrauben oder Kleben existieren für Kunststoffteile weitere Möglichkeiten wie Montage- oder Mehrkomponentenspritzguss, Schnappverbindungen oder diverse Schweißverfahren [Ehr04]. Ein Schweißverfahren, welches sich durch kurze Schweißzeiten und eine gute Automatisierbarkeit und die damit verbundene hohe Wirtschaftlichkeit auszeichnet, ist das Ultraschallschweißen. Der prinzipielle Aufbau einer Ultraschallschweißmaschine mit ihren wesentlichen Komponenten ist in Abbildung 1.1 dargestellt. Mit Ultraschall sind formschlüssige und dichte Verbindungen auch ohne zusätzliche Dichtelemente wie z. B. O-Ringe möglich. Die Tatsache, dass neben den beiden zu verbindenden Werkstückteilen keine weiteren Elemente wie Schrauben, Klebstoff oder Dichtungen notwendig sind, macht das Verfahren zudem einfacher automatisierbar. Die Haupteinsatzgebiete des Ultraschallschweißens sind in der Automobil-, der Elektro- und Elektronik-, der Haushalts-, der Kommunikations-, der Medizin- und Labortechnik- sowie der Verpackungsindustrie. Das Verfahren kommt dabei vorwiegend bei technisch hochwertigen Anwendungen mit hohen Anforderungen an die Schweißqualität zum Einsatz [Ehr04]. Da mit dem Verfahren stoff- und gasdichte Verbindungen hergestellt werden können, wird es häufig zum „Abschließen“ von Baugruppen durch eine Verschweißung eines Deckels mit einem Gehäuse verwendet und findet seine Anwendung somit oftmals am Ende der Prozesskette. Ausschussteile, die durch eine fehlerhafte Schweißung entstehen, bedingen meist relativ hohe Fehlerkosten, da die Wertschöpfung aller vorangegangenen Prozessschritte sowie die Einzelbauteile verloren gehen.

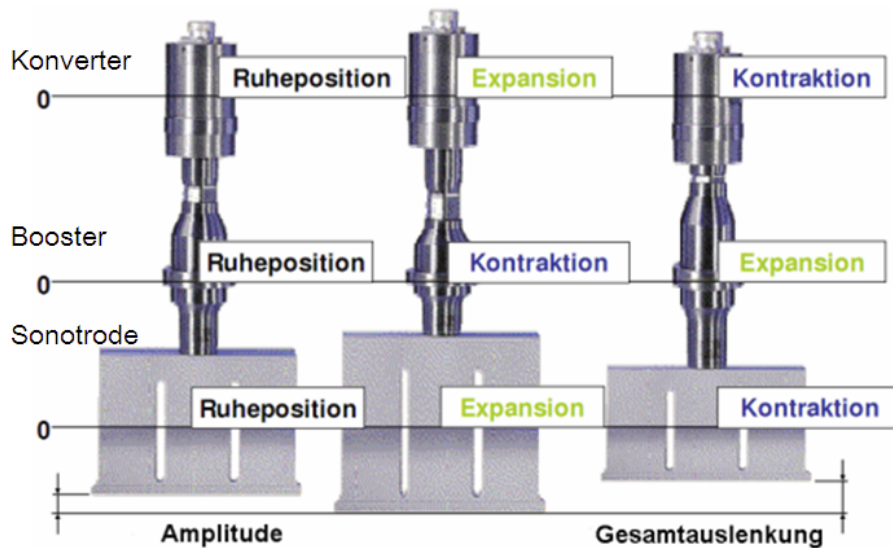


Abbildung 1.2: Longitudinale Schwingungen des Schwinggebildes beim Ultraschallschweißen (aus [Fis06])

Im Folgenden wird der Prozessablauf beim Ultraschallschweißen kurz beschrieben: Ein schwingungsfähiges System, bestehend aus Konverter, Booster und Sonotrode, wird durch Anregung mittels Piezoelementen im Konverter in longitudinale Schwingungen im Ultraschallbereich versetzt (Abbildung 1.2). Typische Schwingfrequenzen sind dabei 20, 30, 35 oder 40 kHz – je nach Maschinenhersteller und Größe der zu schweißenden Werkstücke. Die ursprüngliche Schwingungsamplitude im Bereich von ca. 6 bis 13 μm – abhängig von der Schwingfrequenz – wird durch den Booster (auch Amplitudentransformationsstück genannt) und die Sonotrode anwendungs- und werkstoffspezifisch um den Faktor 1-5 verstärkt. Durch die mechanische Ankopplung der Sonotrode wird die Schwingung in das Werkstück eingeleitet. Die Schwingungsenergie wird durch innere Reibung und Grenzflächenreibung absorbiert, wo die entstehende Wärme zu einem lokalem Aufschmelzen in der Fügezone und dadurch zur Verbindung der beiden Fügepartner führt. Für das Ultraschallschweißen existieren verfahrensspezifische Konstruktionsrichtlinien, um das Aufschmelzen in der Fügezone zu gewährleisten und somit eine gute Schweißverbindung erhalten zu können, die den gestellten Anforderungen genügt [CCCL00, Str95]. So dient der so genannte Energie richtungsgeber (ERG) zur Schall- und somit Energiefokussierung. Durch die kleine und definierte Kontaktfläche kommt es zudem zu einer lokalen Erhöhung der Druckwechselspannung und damit zu einem schnelleren Aufschmelzen in der Fügezone [Bec71, Pot71].

Der Ablauf beim Ultraschallschweißen gliedert sich in die drei Phasen [Ehr04, Fis06]:

- Das *Anfahren*, wobei die Sonotrode auf die zu verbindenden Bauteile fährt und die mechanische Ankopplung unter einer einstellbaren Kraft stattfindet.

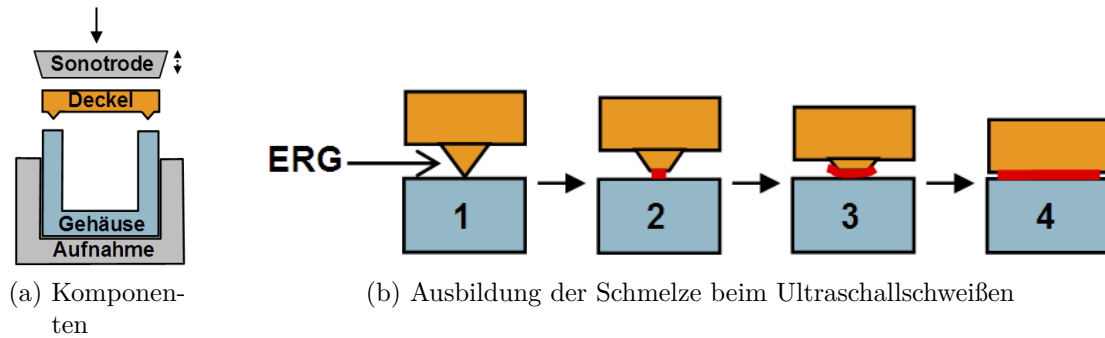


Abbildung 1.3: Einzelteile und Schmelzebildung

- Die eigentliche *Schweißphase*, während der die durch den Konverter erzeugte Schwingung unter einer anstehenden Kraft ins Bauteil eingeführt wird und zum Aufschmelzen und der Verbindung der Bauteile führt.
- Die *Haltephase* zum Abkühlen der beim Schweißen entstandenen Schmelze bis zur Ausbildung der festen Schweißnaht.

Die wesentlichen Stellgrößen zur Einstellung des Schweißprozesses sind dabei die aufgebrachten Fügekräfte während der drei Prozessphasen, die eingestellte Amplitude der Sonotrode, die Schweißzeit sowie der Schweißweg. Teilweise beeinflussen sich die Parameter gegenseitig, z. B. geht mit größeren Schweißwegen in der Regel auch eine höhere Schweißzeit einher. Zudem sind mehrere Schweißmodi zu unterscheiden, die das Abschaltkriterium, also das Ende der Schweißung, festlegen. Das Abschalten der Schwingung kann wählbar nach einer festgelegten Zeit, einem vorgegebenen Weg (relativ oder absolut) oder der vom Generator aufgebrachten Energie erfolgen. Neuere digitale Steuerungen der Schweißmaschinen ermöglichen teilweise auch eine Kombination aus verschiedenen Abschaltkriterien.

1.1.2 Eigenschaften des Ultraschallschweißens

Das Ultraschallschweißen als Fügeverfahren für Kunststoffteile weist im Vergleich zu anderen etablierten Schweißverfahren wie beispielsweise dem Vibrationsschweißen, dem Heizelementschweißen oder dem Laserschweißen eine Reihe an Vorteilen auf.

Vorteile Ein Hauptvorteil des Ultraschallschweißens sind die extrem kurzen Schweißzeiten von unter einer Sekunde [Ehr04]. Durch die hohen Schweißfrequenzen wird die Energie in kurzer Zeit eingeleitet und führt sehr schnell zum Erwärmen und Aufschmelzen der Fügeteile. Anhand eines Modells für die Erwärmung aufgrund des Verlustmoduls wurden Temperaturgradienten von 1000 °C pro Sekunde ermittelt und experimentell bestätigt [AWW72]. In [TDW83] wurden Temperaturen von 250 °C in der Schweißnaht bereits nach 10 ms erreicht (Gradient von > 10000 °C/s). Die damit verbundenen kurzen Zykluszeiten

erhöhen die Wirtschaftlichkeit des Verfahrens und machen es für den Einsatz in der Serienproduktion auch mit hohen Stückzahlen geeignet [N.N07b]. Durch Fügeprozesse wie das Ultraschallschweißen lassen sich auch häufig aufwändige und damit teure und störungsanfällige Spritzgießwerkzeuge einsparen, indem zunächst einfachere Bauteile spritzgegossen und diese anschließend gefügt werden [MKK97]. Ein weiterer Vorteil ist, dass die Maschinenteknik des Ultraschallschweißens eine hervorragende Integrierbarkeit in Fertigungsstraßen aufweist [BC89, HK04] und dabei relativ preiswert ist – insbesondere im Vergleich zu Laserschweißmaschinen. Bei technisch hochwertigen Teilen werden oft hohe Anforderungen an die Schweißnahtqualität gestellt und häufig hängt die Funktionstüchtigkeit des gesamten Bauteils von der Qualität der Schweißverbindung ab. Durch das breite Spektrum an zur Verfügung stehenden Schweißfrequenzen (20-70 kHz) können Ultraschallschweißmaschinen diese hohen Anforderungen bedienen [Ehr04].

Beim Ultraschallschweißen erfolgt die Erwärmung im Material selbst und nicht von außen. Die thermische Einwirkung auf das Bauteil ist somit relativ gering und lokal sehr begrenzt. Ein Oxidieren und ein Abkühlen der Fügeflächen wie z. B. beim Heizelementschweißen kann nicht auftreten [MP69]. Auch ist das Verfahren energetisch sehr günstig, da Energie nur während der eigentlichen Schweißung eingebracht wird und nicht wie beim Heizelementschweißen größere Maschinen- oder Werkzeuteile konstant auf einem hohen Temperaturniveau gehalten werden müssen.

Des Weiteren ist das Ultraschallschweißen für nahezu alle thermoplastischen Werkstoffe anwendbar. Spezielle Anforderungen an den Werkstoff oder die Werkstoffpaarung, wie sie z. B. beim Laserschweißen gestellt werden (Unterteil absorbierend, Oberteil transparent im Bereich der Wellenlänge des eingesetzten Lasers) [Ehr04], existieren beim Ultraschallschweißen nicht oder nur in viel kleinerem Maße. Da die Teile während des Schweißvorgangs keine Relativbewegungen wie beim Vibrations-, Orbital- oder Reibschweißen erfahren, sind genauere Lage- und Formtoleranzen erreichbar. Auch die Schweißnahtgeometrie unterliegt nicht so engen Einschränkungen wie beim Rotationsschweißen: Die Nahtgeometrie muss (in etwa) in einer Ebene liegen, kann darüber hinaus jedoch fast jeden Freiheitsgrad verwenden.

Durch die Tatsache, dass keine weiteren Hilfsstoffe wie z. B. Klebstoffe oder Lösungsmittel notwendig sind, ist das Ultraschallschweißen als Fügeverfahren auch umweltfreundlich, die mögliche Sortenreinheit der Schweißverbindungen wirkt sich positiv auf das Recycling aus.

Nachteile Aufgrund physikalisch begründeter Restriktionen bei der Größe der Schweißwerkzeuge ist das Verfahren des Ultraschallschweißens auf Werkstücke kleiner bis mittlerer Größe beschränkt – zumindest wenn durchgängige Nähte erzielt werden müssen [Ehr04]. Ein wesentlicher Nachteil des Ultraschallschweißens ist die Beeinflussung des Schweißprozesses durch Vorprozesse, Lagerung der Kunststoffteile sowie Umgebungsbedingungen. Die meisten Fügepartner beim Ultraschallschweißen werden durch Spritzgießen hergestellt. Dabei

haben die Parameter des Spritzgießens Einfluss auf die innere Gefügestruktur der Werkstücke und wirken sich auf das Schweißverhalten aus [Kor96]. Werkstoffabhängig spielen auch die Lagerbedingungen eine Rolle. Bei teilkristallinen Werkstoffen findet in den ersten Stunden bis Tagen nach der Herstellung noch eine Nachkristallisation statt, so dass sich die Lagerzeit auf die Gefügestruktur und somit wieder auf das Schweißverhalten auswirkt. Je mehr die Werkstoffe auskristallisiert sind, desto mehr Energie ist notwendig, die Kunststoffe in der Fügezone aufzuschmelzen und eine dauerhafte Verbindung herzustellen. Einen weiteren Einfluss stellt die Feuchtigkeitsaufnahme dar, die stark vom Werkstoff abhängig ist. Mit der Aufnahme von Feuchtigkeit verändern sich werkstoffspezifische Kenngrößen wie Schubmodul G und mechanischer Verlustfaktor $\tan \delta$ teilweise deutlich, was sich direkt auf die innere Reibung und die Schalleitung und somit auch auf das Schweißverhalten auswirkt [Lan78, NL96, TDW83]. Das Schweißen feuchter Fügeteile führt zu schlechten Ergebnissen mit längeren Schweißzeiten und porösen Schweißnähten [Ehr04]. Auch wenn die beeinflussenden Faktoren bekannt sind, können die tatsächlichen Zustände einzelner Werkstücke nicht ohne Weiteres ermittelt oder beurteilt werden. Dies gilt insbesondere unter Produktionsbedingungen.

Derzeit erfolgt die Prozessüberwachung beim Ultraschallschweißen nur mittels einfacher Grenzwertüberwachung verschiedener Maschinengrößen [Fis06]. Dabei werden verschiedene Kenngrößen wie z. B. Schweißzeit, absoluter Weg oder eingebrachte Energie mit vorgegebenen Grenzwerten verglichen. Die teilweisen Abhängigkeiten und Wechselwirkungen zwischen den Parametern in Verbindung mit den o. g. Einflüssen erschweren die Festlegung der Schwellwerte erheblich. Zudem weist der Ultraschallschweißprozess auch bei konstanter Prozesseinstellung relativ hohe Schwankungen auf. In der Praxis treten Fehler beim Ultraschallschweißen meist zufällig auf. Eine Ursachenfindung gestaltet sich meist als äusserst schwierig, da verschiedene Parameter wie Lager- oder Spritzgießbedingungen einfließen, die nur schwer oder teils gar nicht zu überprüfen sind.

Die Schweißnähte der bei der Prozesseinstellung verschweißten Versuchsteile müssen qualitativ bewertet werden. Die Nahtqualität ist dabei entweder nur visuell oder aber zerstörend prüfbar, was jedoch bei allen Schweißverfahren der Fall ist. Oft ist die Prüfung mit einem erheblichen apparativen und handhabungstechnischen Aufwand verbunden und kann bei zerstörender Prüfung nur an wenigen Stichproben durchgeführt werden.

Randbedingungen Aufgrund der Einleitung, der inneren Ausbreitung und der Umwandlung der longitudinalen Schwingungen muss die Schweißnaht in etwa in einer Ebene liegen. Beliebige dreidimensionale Konturen sind beim Ultraschallschweißen nicht möglich. Bei der Konstruktion der Bauteile sollten verfahrensspezifische Konstruktionsrichtlinien eingehalten werden, um eine ausreichende Qualität der Schweißnaht erzielen zu können.

Probleme bei Praxisteilen sind häufig darauf zurückzuführen, dass die Konstruktion nicht kunststoff- oder schweißoptimal ausgeführt ist. Auch gestaltet sich die Definition von Qualität und deren Prüfung, d. h. wie eine geforderte Eigenschaft (reproduzierbar) geprüft werden kann, als schwierig. Ein weiterer Punkt ist, dass die Verifikation der Vorprozesse in der Praxis nicht möglich ist. Es kann also in der Regel nicht nachvollzogen werden, welche chargenbedingten Materialschwankungen und welche Schwankungen bei der Herstellung der Fügepartner durch Spritzgießen aufgetreten sind. Auch sind die Lagerbedingungen – insbesondere wenn die Fügepartner durch einen Zulieferer hergestellt werden – schwer zu kontrollieren. Oftmals wird eine zerstörungsfreie Prüfung für jedes Teil gefordert, was nur durch aufwändige und teure Prüfverfahren möglich wird, z. B. Sichtprüfung oder Prüfung der Dichtigkeit. Die Festigkeit ist dagegen zerstörungsfrei nicht zu erfassen.

Da neben den zu verschweißenden Werkstücken keine weiteren Elemente wie O-Ringe, Schrauben oder Klebstoffe erforderlich sind, sind Ultraschallschweißprozesse vergleichsweise einfach zu automatisieren. Aufgrund der kurzen Schweißzeiten sind bei der automatisierten Fertigung hohe Taktraten erreichbar. Allerdings existiert an den Schweißmaschinen keine einheitliche und herstellerübergreifende Schnittstelle, über die Signale und Prozessparameter weiterführenden Systemen zur Prozessüberwachung verfügbar gemacht werden könnten.

1.2 Problemstellung

Das Ultraschallschweißen als Fügeverfahren für Kunststoffteile wird in vielen Branchen (u. a. im Automotive-, Elektro- und Health-Care-Bereich) eingesetzt, weist jedoch insgesamt starke Prozessschwankungen auf. Die Schwankungen im Prozess zeigen sich zum einen in den zeitlichen Verläufen der Prozessparameter, insbesondere im Leistungs- und Amplitudenverlauf. Auch die Eingangsimpedanz des Konverters, welche als Indikator für das dynamische Verhalten des gesamten Systems gesehen werden kann [LLLA06], weist deutliche Varianzen auf. Diese Schwankungen haben Auswirkungen auf die Nahtqualität, die bei konstanter Prozesseinstellung erzielt wird. Insbesondere die Festigkeit weist meist deutliche Streuungen auf – und dies bei unveränderten Maschinenparametern.

Der Schweißprozess selbst wird durch alle vorgelagerten Prozessschritte beeinflusst. Von der Herstellung des Kunststoffgranulats über die Trocknung, Prozessbedingungen und Prozesseinstellung beim Spritzgießen bis zur Lagerung und den Lagerbedingungen der Teile vor dem Schweißen beeinflusst jeder Schritt den Schweißprozess und wirkt sich auf die erzielbare Schweißqualität aus [Str95, MKK97]. Bereits der Prozess zur Herstellung des Kunststoffgranulats ist Schwankungen unterworfen, die sich in deutlichen Chargenschwankungen beim Spritzgießen bemerkbar machen können [Ujm00]. Der Spritzgießprozess bestimmt maßgeblich die mechanischen und geometrischen Eigenschaften der spritzgegossenen Werkstücke [Bic06]. Auch der Spritzgießprozess weist – je nach Güte des eingestellten Ar-

beitspunktes – mehr oder weniger deutliche Schwankungen auf, die sich auf die Qualität der Spritzgießteile auswirkt [BE93, NS07, Wor96]. Neben Auswirkungen auf die geometrischen Teileigenschaften (Füllgrad, Maßhaltigkeit, Gratbildung, Einfallstellen, ...) beeinflusst das Spritzgießen auch wichtige mechanische Eigenschaften der Werkstücke wie innere Spannungen, Kristallisationsgrad oder Faserorientierung [Win73]. Diese wirken sich deutlich auf das Ultraschallschweißen aus [Kor96], sind aber nur mit aufwändigen mikroskopischen Analysen exakt zu ermitteln. Ein weiterer Faktor ist die anschließende Lagerung inklusive der Lagerbedingungen, die sich insbesondere über die Feuchtigkeitsaufnahme der Werkstücke und die damit verbundene Änderung des E-Moduls sowie eine stattfindende Relaxation, d. h. den Abbau innerer Spannungen, auf den Schweißprozess auswirken [LE85]. Die jeweiligen Einflüsse sind bekannt und in der Literatur nachgewiesen, können in der Praxis jedoch meist nicht verifiziert werden. Dies gilt insbesondere dann, wenn die Fertigung der zu verschweißenden Spritzgießteile durch Zulieferer erfolgt. Bei Formteilen aus Thermoplasten ist die Auslagerung an spezialisierte Spritzgießbetriebe häufig der Fall.

Während der letzten Jahre gab es bei den Ultraschallschweißmaschinen insbesondere bei den Generatoren einige technologische Fortschritte [HK04]. Neben den bislang analogen Generatoren wurden digital gesteuerte Modelle entwickelt. Deren Steuerung basiert auf Mikroprozessoren; sie bieten im Vergleich zu den analogen Geräten eine Vielzahl neuer Parameter sowie Steuer- und Regelungsmöglichkeiten. Bislang wurden Ultraschallschweißprozesse vorwiegend anhand der Schweißzeit oder der eingebrachten Energie gesteuert, d. h. nachdem eine voreingestellte Schweißzeit abgelaufen oder eine zuvor festgelegte Menge an Energie vom Konverter ins System eingebracht war, wird der Schweißprozess beendet [BC89, CCCL00, LLLA06, LC01, vLvN96]. Die Schweißmodi *Zeit* und *Energie* führen jedoch zu relativ hohen Streuungen, so dass sie heutzutage nur noch bei einfacheren Anwendungen oder geringeren Anforderungen eingesetzt werden. Die neu entwickelten, digitalen Steuerungen bieten nunmehr auch die Möglichkeit, den Schweißvorgang zeit- oder geschwindigkeitsgesteuert zu führen [HK04]. Über die Auswirkungen dieser Schweißmodi und dabei relevante Prozessgrößen findet sich in der Literatur sehr wenig. Häufig liegen nur firmeninterne Erfahrungen zu spezifischen Produkten vor, die meist nicht strukturiert erfasst, sondern allein in Form von Mitarbeiter-Know-how vorhanden sind.

Das Ultraschallschweißen ist oft der letzte Schritt in der Prozesskette und Ausschuss durch die verlorene Wertschöpfung deshalb entsprechend teuer. Wenn Produkte mit unzureichender Schweißqualität als Baugruppe weiterverbaut werden, kann ein Ausfall des Teils zu weiteren Folgekosten für Nacharbeit oder gar zu einer Produkthaftung führen. Um dieses Risiko sowie die Ausschusskosten möglichst gering zu halten, sind die Anwender von Ultraschallschweißanlagen bestrebt, defekte Teile und Prozessdriften, die zu solchen führen könnten, zu erkennen und bei Bedarf schnellstmöglich Maßnahmen zur Stabilisierung des Prozesses einzuleiten. Aufgrund der teils großen Prozessschwankungen sowie der star-

ken Wechselwirkungen der einzelnen Prozessparameter und der vielfältigen Einflüsse der Vorprozesse gestalten sich die Erkennung und insbesondere die Ursachenfindung meist sehr schwierig.

Die Einhaltung der spezifizierten Qualität wird heutzutage meist mittels aufwändiger Prüfungen gewährleistet. Bei medizintechnischen oder sicherheitsrelevanten Bauteilen werden teils 100 %-Prüfungen durchgeführt. Dichtheitsprüfungen – ein durch die Schweißnaht begrenztes, abgeschlossenes Volumen vorausgesetzt – sind relativ aufwändig, da die Prüflinge mit einem System zur Druckaufbringung gekoppelt und entsprechend abgedichtet werden müssen. Dies gilt insbesondere bei automatisierten Prüfstationen innerhalb einer Fertigungslinie, die deshalb entsprechend teuer sind. Andere Qualitätseigenschaften, z. B. Austrieb, Dicke und Homogenität der Schweißnaht, werden mittels visueller Beurteilung durch speziell geschultes Prüfpersonal ermittelt, was eine subjektive Bewertung sowie einen möglichen Fehlerschlupf bedingt. Weitere Eigenschaften wie die Festigkeit der Schweißnaht, die eines der wichtigsten Qualitätskriterien darstellt¹, können nur zerstörend erfasst werden. Dies steht einem größeren Prüfungsumfang entgegen. Die Produzenten behelfen sich mit dem Einsatz von Stichprobenprüfungen und Statistischer Prozesskontrolle (SPC), die bei der Kunststoffverarbeitung jedoch als kritisch zu bewerten sind [PWHH93]. Durch den Stichprobencharakter können kurzfristige Störungen und singuläre Ausreißer nicht erfasst werden. Dies führt zu unentdeckten Schlechteilen. Basiert die SPC zudem nicht auf Prüfwerten, sondern – z. B. aufgrund des geringeren Aufwands oder um mehr Teile „prüfen“ zu können – auf ermittelten Prozesskenngrößen wie beispielsweise dem Maximum des Druckverlaufs, besteht die Gefahr, dass aufgrund mangelnden Zusammenhangs zwischen beobachteter Größe und der interessierenden Zielgröße Fehlalarme ausgelöst oder Qualitätsprobleme nicht erkannt werden [Häu95, Wor89].

Der zu erbringende Prüfaufwand führt insgesamt zu weiteren Kosten, die durch die Verfügbarkeit einer entsprechenden Prozessüberwachung effektiv reduziert werden können. Die Prozessüberwachung erfolgt heutzutage anhand der Grenzwertüberwachung einfacher Kenngrößen aus der Maschinensteuerung wie beispielsweise Schweißzeit, Schweißweg, Schweißenergie oder maximaler Leistung [Fis06]. Die Kenngrößen unterliegen jedoch den prozesstypischen Wechselwirkungen und einige Kenngrößen (z. B. die Schweißzeit im Schweißmodus Energie) weisen nachgewiesenermaßen keine Korrelation zur Festigkeit auf [vLvN96] – auch wenn sie in der Praxis dennoch (unzulässigerweise) verwendet werden. Eine wesentliche Voraussetzung, um eine Überwachung der Qualität anhand von Prozesskenngrößen zu ermöglichen, sind entsprechende Korrelationen und Zusammenhänge zwischen abgeleiteter Prozesskenngröße (PK) und Qualitätskenngröße (QK). In Abbildung 1.4 ist für den entwickelten Probekörper (siehe Abschnitt 3.3) der bei der Prüfung erreichte Berstdruck über den exemplarischen Kenngrößen Schweißzeit und Fügegeschwindigkeit aufgetra-

¹Jeder Prüfkörper (DVS oder AWS) ermöglicht die Ermittlung der Schweißnahtfestigkeit

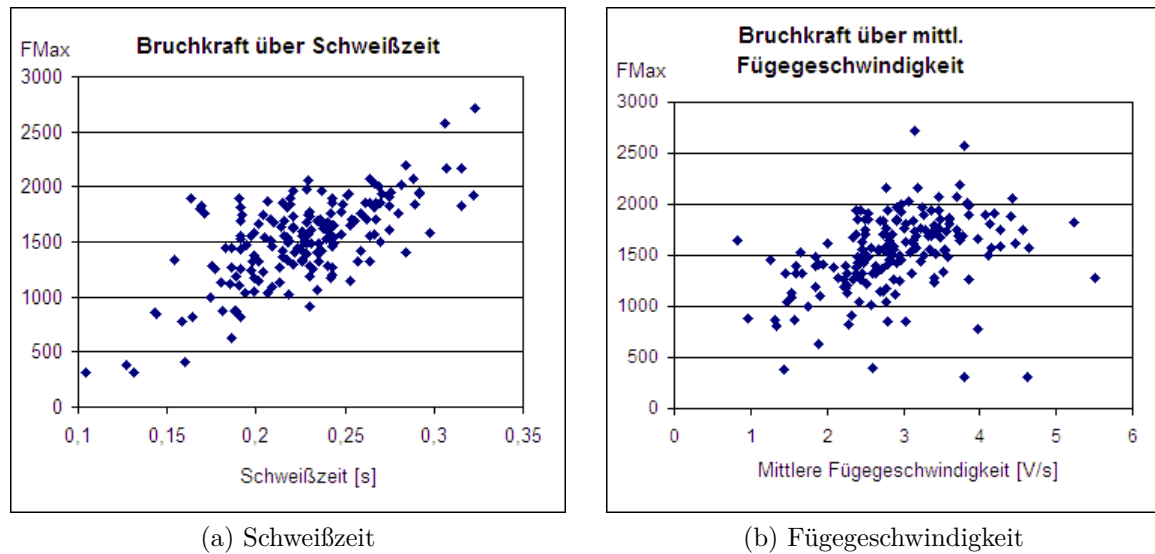


Abbildung 1.4: Korrelation zwischen Prozesskenngrößen und dem Berstdruck als Qualitätskenngröße

gen. Die Kenngrößen weisen zur Schweißnahtqualität eine Korrelation von 0,66 bzw. 0,39 auf, sind also zur Prozessüberwachung nicht geeignet. Klassische Korrelationsmethoden, wie sie z. B. auch Spannungs-Dehnungsdiagrammen im linearen Bereich zugrunde liegen, sind für das Ultraschallschweißen unzureichend.

Zusammenfassend kann somit bis dato gesagt werden: „[...] Zudem liefert das Ultraschallschweißen schwankende Schweißergebnisse und eine Prozesskontrolle fehlt.“ (Zitat aus [VH05], S. 33).

1.3 Zielsetzung und Lösungsansatz

Ausgehend von der Tatsache, dass für das Ultraschallschweißen keine leistungsfähige Prozessüberwachung existiert, soll die vorliegende Arbeit einen Beitrag zum Aufbau eines Überwachungssystems leisten. Ziel ist es, eine Methodik zu schaffen, für zyklische Produktionsprozesse eine zyklussynchrone Qualitätsbewertung zu ermöglichen, um Qualitätsprobleme und Störungen schnell erkennen und beheben zu können. Die Vorgehensweise soll am Beispiel des Ultraschallschweißens von Kunststoffteilen erarbeitet und verifiziert werden. Sie soll jedoch so allgemein bleiben, dass sie später auch auf andere zyklische Fertigungsprozesse übertragen werden kann. Wesentliches Merkmal der zu entwickelnden Methode zur Prozessüberwachung ist es, eine zuverlässige Gut-Schlecht-Erkennung für jedes produzierte Werkstück zu ermöglichen. Dies schafft eine zusätzliche Prozesssicherheit und kann auch zu einem entscheidenden Know-how-Zuwachs beim Anwender führen, indem der Prozess und die Zusammenhänge transparenter werden. Dies geht damit einher, dass Veränderungen im Schweißprozess viel schneller erkannt und angezeigt werden können, was den Ausschuss

reduziert und zudem die Zuordnung zu möglichen Ursachen deutlich vereinfachen kann. Des Weiteren kann durch eine funktionierende Prozessüberwachung eine existierende statistische Prozesskontrolle sinnvoll ergänzt und auf lange Sicht sogar ersetzt werden. Durch eine zuverlässige Prozessüberwachung wird es für Produzenten möglich, den derzeitigen Prüfaufwand, der teilweise eine 100 %-Prüfung umfasst, effektiv zu vermindern. Während heute einige wichtige Qualitätsmerkmale wie die erzielte Festigkeit der Schweißverbindung nur zerstörend geprüft werden können, soll die entwickelte Prozessüberwachung in der Lage sein, auch für solche Merkmale eine Qualitätsprognose zu ermöglichen.

Speziell für das Ultraschallschweißen von Kunststoffteilen sollen Möglichkeiten erarbeitet werden, Störungen oder Veränderungen in den Vorprozessen, die sich auf das Schweißen auswirken, zu erkennen und bei Problemen im Schweißprozess somit Hinweise auf mögliche Ursachen zu geben.

Ein Lösungsansatz ist der Aufbau einer modellbasierten Prozessüberwachung. Der schematische Aufbau eines solchen Überwachungssystems sowie die Zusammenhänge beim Ultraschallschweißen unter Einbezug der Vorprozesse und Störgrößen sind in Abbildung 1.5 dargestellt. Als Ausgangsbasis dient die während der letzten Jahre am Fraunhofer IPA erfolgreich entwickelte NEPRES²-Systemtechnik [RS01, SN05b].

Über die gesamte Prozesskette vom Spritzgießen über Transport und Lagerung bis zum Ultraschallschweißen wirken zufällige und nicht beeinflussbare Störgrößen auf die einzelnen Fertigungsstufen ein. Diese haben Auswirkungen auf die Werkstücke und können sich so aufgrund des durchgehenden Materialflusses auch in den Folgeprozessen auswirken, was letztlich zu schwankenden Qualitätsausprägungen des geschweißten Produkts führen kann. Häufige Qualitätsanforderungen an geschweißte Teile sind Dichtigkeit, Festigkeit der Schweißnaht oder Maße. Entsprechend der vereinbarten Sollwerte und Toleranzen ist ein geschweißtes Werkstück anhand seiner tatsächlichen Merkmalsausprägungen, die jedoch unbekannt sind, als Gut- oder Schlechtteil einzuordnen. Zum Aufbau einer modellbasierten Prozessüberwachung werden aus dem relevanten Fertigungsprozess – hier der Ultraschallschweißprozess – für jeden Fertigungszyklus Sensorsignale erfasst. Die aufgezeichneten Sensorsignale sollten dabei möglichst viele Informationen über den Prozess beinhalten, weshalb die Sensorik prozessspezifisch auszuwählen und zu erarbeiten ist. Um die qualitätsrelevanten Informationen zu verdichten und die Datenmenge handhabbar zu machen, werden aus Signalen im Fertigungstakt Prozesskenngrößen für jeden Fertigungszyklus berechnet. Diese dienen den verschiedenen Prozessmodellen als Eingangsgrößen. Die Prozessmodelle sind ein wichtiges Element der Prozessüberwachung und bilden den Zusammenhang zwischen Prozess- und somit Signalschwankungen, die wiederum in den Prozesskenngrößen abgebildet sind, und den Zielgrößen ab. Das Prozessmodell muss dabei spezifisch auf jeden Zielprozess adaptiert werden. Ein Prozessmodell soll Aussagen über die erzielte Teilequalität machen,

²Neuronales Prozess-Regelungssystem

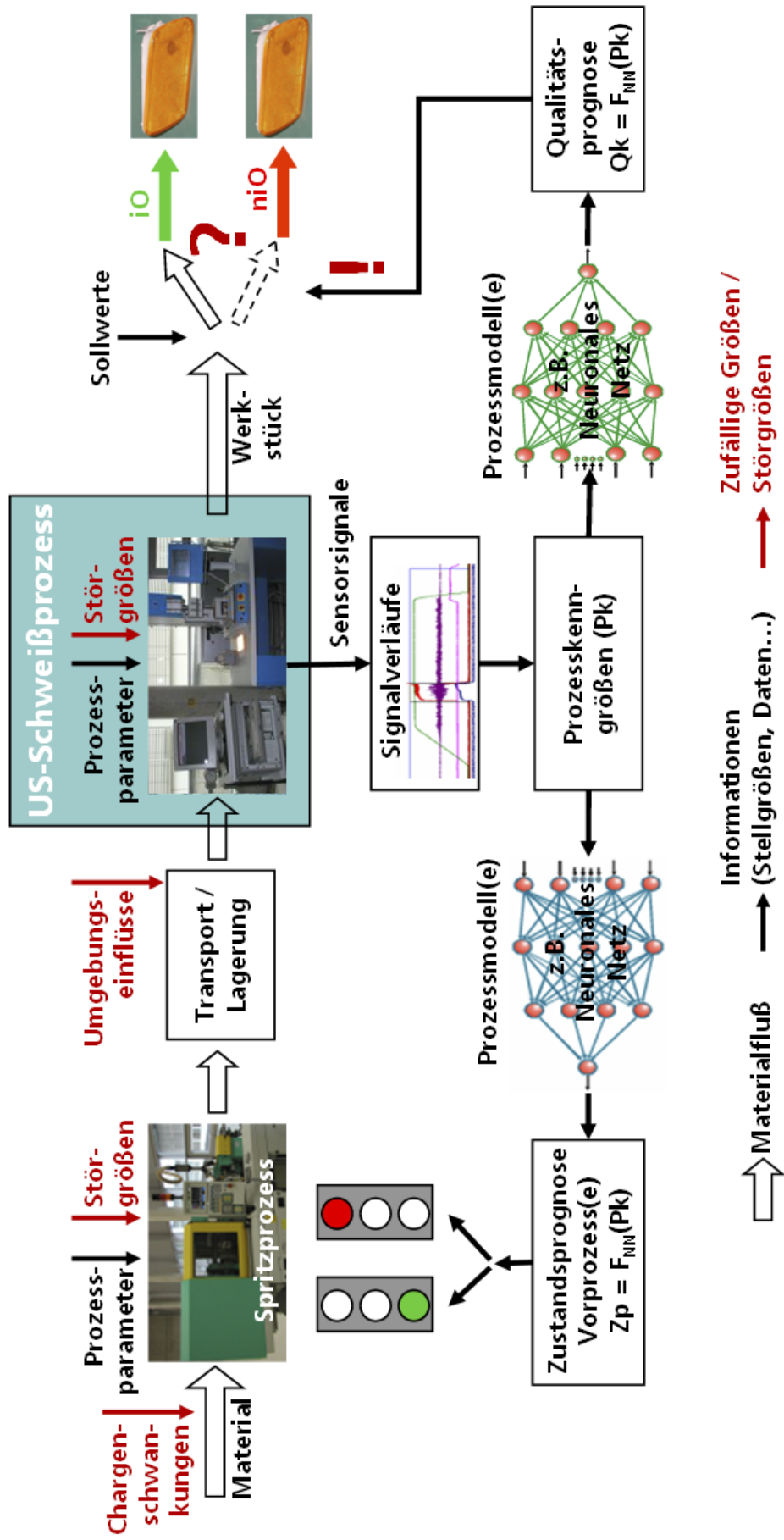


Abbildung 1.5: Lösungsansatz für die modellbasierte Prozessüberwachung beim Ultraschallschweißen

wodurch mittels Vergleich der Prognosewerte mit den Sollwerten Schlechteile erkannt und ausgeschleust werden können. Ein zweites Prozessmodell soll hier dazu dienen, die Zustände der Vorprozesse zu überwachen und bei entsprechenden Abweichungen oder Störungen dem Anwender eine Rückmeldung zu geben. Die Ursache von Fertigungsproblemen in Vorprozessen können damit schnell abgestellt und in der weiteren Produktion kann Ausschuss vermieden werden. Dies ist insbesondere für Prozesse mit kleinen Losgrößen interessant, da sich dort umfangreiche Versuchsreihen zur Modellbildung nicht lohnen [NK10a, NK10b].

1.4 Aufgabenstellung und Vorgehensweise

Damit der in Abbildung 1.5 dargestellte Lösungsansatz insgesamt umgesetzt werden kann, sind verschiedene Punkte zu bearbeiten und zu einem Gesamtverfahren zu kombinieren. Insgesamt ergeben sich die folgenden Teilaufgaben:

Signalerfassung: Um auftretende Prozessschwankungen – ursächlich für Qualitätsschwankungen – abbilden und bewerten zu können, müssen diese zuerst sensorisch erfasst werden. Dazu muss der Schweißprozess bezüglich relevanter Signale, die sich aufgrund der physikalischen Wirkungsweise ergeben, analysiert und die Signale müssen durch geeignete Sensoren erfasst werden. Es muss eine einheitliche Signalschnittstelle erarbeitet und definiert werden, damit die Signale auch an konventionellen Ultraschallschweißmaschinen, die im industriellen Umfeld eingesetzt werden, und nicht nur an speziell ausgerüsteten Labormaschinen erfasst werden können.

Definition Referenzteil: Um während der Analysephase reproduzierbare und verlässliche Ergebnisse zu erhalten, muss zuerst ein Referenzteil ausgewählt werden. Versuchstechnische Unzulänglichkeiten und Einschränkungen, die bei realen Bauteilen meist vorzufinden sind, können somit vermieden werden. Anhand von Schweißversuchen, die mit dem ausgewählten Referenzteil durchgeführt werden, können anschließend die wesentlichen Zusammenhänge zwischen Signalen, abgeleiteten Prozesskenngrößen sowie Qualitäten und Einflüssen der Vorprozesse untersucht und erarbeitet werden.

Versuchsplanung und -durchführung: Nach Auswahl des Referenzteils müssen die notwendigen Versuche systematisch geplant werden. Dabei müssen mögliche Einflüsse der verfügbaren Maschinenparameter berücksichtigt werden. Die Versuchsplanung muss auch im Hinblick auf die spätere Auswertung erfolgen und zur Identifikation von Wechselwirkungen und für eine Minimierung des Versuchsumfanges die Grundlagen der statistischen Versuchsplanung berücksichtigen. Anschließend müssen die aufgestellten Versuchspläne abgearbeitet werden, d. h. es müssen Schweißversuche mit den festgelegten Einstellungen durchgeführt werden. Für jedes Versuchsteil müssen die

Sensorsignale aufgezeichnet und die zuvor definierten Qualitätsparameter erfasst werden. Wesentlich ist, dass über alle Versuche eine eindeutige Zuordnung der Maschineneinstellung, der Signaldateien und der erfassten Qualitäten gegeben ist.

Signalanalyse: Nach der Versuchsdurchführung müssen die aufgezeichneten Signale hinsichtlich typischer Signalformen und auftretender Varianzen analysiert werden. Wesentlich ist die Aufteilung in Signalstreuungen im Arbeitspunkt (AP) des Ultraschallschweißprozesses und solche Streuungen, die mit geänderten Maschineneinstellungen oder Variationen in den Vorprozessen in Verbindung stehen. Für die angestrebte Qualitätsprognose und Zustandsüberwachung der Vorprozesse ist auch die Identifikation von Prozessstreuungen wichtig, die in Zusammenhang mit Qualitätsschwankungen oder Veränderungen der Vorprozesse gebracht werden können.

Kenngrößenbildung: Da die aufgezeichneten Signale in Rohform zu umfangreich und informationstheoretisch vielfach redundant sind, ist eine Vorverarbeitung und Informationskompression notwendig. Dies erfolgt, indem die in den Signalen enthaltenen, relevanten Informationen durch sogenannte Prozesskenngrößen auf skalare Werte reduziert werden, z. B. eine Reduktion des aufgezeichneten Schweißwegs auf Anfangs- und Endwert sowie auf eine mittlere Fügegeschwindigkeit. Aufbauend auf den durch die Signalanalyse identifizierten Streuungen und Varianzen müssen diese in entsprechende Prozesskenngrößen überführt werden. Für die so definierten Kenngrößen müssen Algorithmen für die automatische Berechnung erarbeitet werden, um die Werte bei der späteren Anwendung aus den Signalverläufen im Fertigungstakt automatisiert ableiten zu können.

Varianz- und Korrelationsanalyse: Für die bei den Versuchen generierten Daten müssen Varianzen und Korrelationen zwischen Prozesskenngrößen, die aus den aufgezeichneten Prozesssignalen gebildet werden, und interessierenden Zielgrößen, d. h. Qualitäten oder Zustände der Vorprozesse, untersucht werden. Ziel dabei ist es, die auftretenden Streuungen der Werte der einzelnen Prozesskenngrößen in Relation zu Qualitätsschwankungen oder zu detektierenden Ereignissen der Vorprozesse (z. B. hoher Feuchtegehalt im geschweißten Werkstück) zu bringen. Wesentlich ist die Berücksichtigung von Wechselwirkungen und Ursache-Wirk-Zusammenhängen, wobei die Arbeitsinhalte und Ergebnisse auch eng mit den Aufgaben der KenngrößenSelektion und Modellbildung verknüpft sind.

KenngrößenSelektion: Abhängig von der gefundenen Anzahl an Kenngrößen in Zusammenhang mit der Modellierungsmethode muss die Anzahl der verwendeten Prozesskenngrößen minimiert werden, um die Modellgüte zu steigern. Existierende Selektionsverfahren wie statistische Signifikanzanalysen oder Genetische Algorithmen (GA)

müssen hinsichtlich ihrer Eignung und Ergebnisse untersucht und bewertet werden. Nur durch eine geeignete Selektion der Prozesskenngrößen für den Modellinput können robuste und hinreichend genaue Modelle für die Prozessüberwachung aufgebaut werden.

Modellbildung: Der Zusammenhang zwischen Prozesskenngrößen und Zielgrößen muss mittels einer geeigneten Modellierungsmethode abgebildet werden. Während die Zusammenhänge innerhalb des betrachteten Anwendungsprozesses – hier das Ultraschallschweißen – mittels Versuchsplänen untersucht werden können, ist dies für Einflüsse und Störungen vorgelagerter Prozessschritte nicht möglich. Die gewählten Modelle müssen folglich in die Lage versetzt werden, neuartige Zustände als unbekannt zu detektieren. Daraufhin kann zielgerichtet nach der Ursache gesucht und das Modell angepasst werden. Da solche Störungen und Einflüsse oft nur deskriptiv beschrieben werden können, muss die Modellierungsmethode in der Lage sein, neben den durch die Versuchspläne generierten Daten auch vorhandenes Prozess- und Erfahrungswissen integrieren zu können. Ideal geeignet erscheinen dafür Ansätze aus dem Bereich der Neuro-Fuzzy-Modellierung. Wesentliche Kriterien zur Modellbildung sind neben den erzielten Korrelationen zwischen Ist- und Prognosewerten auch die Generalisierungs- und Anpassungsfähigkeit sowie der notwendige Aufwand zur Erstellung des Prozessmodells. Entscheidend ist, dass die verwendeten Modelleingangsparameter, d. h. die getroffene Auswahl an Kenngrößen, einen wesentlichen Einfluss auf die Modellgüte haben.

Validierung des Gesamtsystems: Abschließend müssen die entwickelte Vorgehensweise, die einzelnen Methoden und Systemtechniken kombiniert und für ein reales Bauteil angewendet werden. Nur so kann die Eignung verifiziert und die Übertragbarkeit auf andere Prozesse und Werkstücke nachgewiesen werden.

2 Stand der Technik

In nahezu allen wichtigen Produktbereichen haben technische Kunststoffprodukte mit unterschiedlicher Produktkomplexität und hohem Qualitätsstandard einen hohen und weiter steigenden Stellenwert erreicht [Erh08]. Auch ist der Trend ungebrochen, bei immer mehr Produkten andere Verbindungstechniken durch Schweißprozesse zu ersetzen [N.N07a]. Insbesondere die in dieser Arbeit betrachteten Ultraschallschweißprozesse haben dazu geführt, dass zunehmend komplexere Kunststoffprodukte gefertigt werden können. Anlagenseitig stehen dazu inzwischen digitale Schweißprozesssteuerungen mit Einzelparameterregelung der Einstellgrößen zur Verfügung [HK04]. Näheres zur Maschinenteknik und dem Verfahren des Ultraschallschweißens ist in Abschnitt 2.1 zu finden. Schwankungen der Schweißteileigenschaften (Werkstoff, Fügezone, Verzug), der Istwerte der Einstellparameter und der resultierenden Prozessverläufe vor allem in Grenzbereichen sowie Einflüsse der Umgebung werden steuerungsseitig derzeit weder erfasst noch berücksichtigt. Dies führt dazu, dass eine fehlerfreie Produktion trotz geregelter Schweißparameter nicht gewährleistet ist, insbesondere wenn die Vorprozesse und Lagerbedingungen unbekanntem Schwankungen unterliegen oder die Bauteile nicht unter Gesichtspunkten des Schweißprozesses konstruiert wurden [Str95]. Aus diesem Grund wird heute der Produktion häufig eine teilweise sehr aufwändige Qualitätsprüfung ultraschallgeschweißter Produkte nachgeschaltet. Diese erfolgt häufig (z. B. bei kritischen Produkten aus dem Bereich der Medizintechnik) durch eine vollständig automatisierte Dichtigkeitsprüfung, die jedoch entsprechend hohe Investitionskosten erfordert, sowie durch eine 100 %ige Sichtkontrolle. Bei konventioneller Sichtprüfung – im Vergleich zur automatisierten – ist jedoch die Fehlererkennungsrate bekanntermaßen begrenzt, z. B. aufgrund der notwendigen Daueraufmerksamkeit, des Entscheidungszwangs unter Zeitdruck und des ständigen Fixationswechsels der Augen [Mas80]. Dies trifft insbesondere auch für die visuelle Prüfung von Schweißnähten zu. Eine automatisierte Prüfung mittels Techniken der industriellen Bildverarbeitung ist meist jedoch nicht oder nur unzureichend möglich. Dies hat unterschiedliche Gründe, z. B. die oft schlechte Zugänglichkeit der Schweißnaht, der räumliche Verlauf der Schweißnaht oder unzureichende Kontrastverhältnisse. Transparente Kunststoffteile, das Aspektverhältnis zwischen Teilegröße und Schweißnahtgeometrie sowie die unzureichenden Möglichkeiten, Fehlermuster algorithmisch beschreiben zu können (z. B. Aufschäumen der Naht oder andere Inhomogenitäten), sind weitere Gründe.

Während optische Merkmale wie Austrieb oder Durchgängigkeit der Schweißnaht bei transparenten Kunststoffen mit entsprechendem Aufwand komplett geprüft werden können, ist dies bei Festigkeitswerten aufgrund des zerstörenden Charakters der Prüfung nicht möglich [LE85]. Schwankungen dieser Qualitätsmerkmale sind aus heutiger Sicht nur durch prozessintegrierte Überwachungssysteme erkennbar. Solche Systeme basieren stets auf der formalen Beschreibung oder Modellierung des Prozesses, womit anschließend durch die Überwachung von Prozessdaten auf Abweichungen der Produktqualität geschlossen wird. Somit kann dann auch kurzfristig auf unzulässige Prozess- und Qualitätsabweichungen reagiert werden.

In den vergangenen Jahren wurden verstärkt mittels statistischer Prozesskontrolle (SPC) (siehe Abschnitt 2.2.1) auch Festigkeitseigenschaften durch zerstörende Prüfung überwacht. Dabei wurden jedoch teils nur unzureichende Resultate erzielt, insbesondere bei statistisch ungesichert auftretenden Störeinflüssen [Nat92, VE00]. So können einzelne Ausreißer, wie sie bei der Verarbeitung von Kunststoffen auch aus dem Spritzgießen bekannt sind, nicht sicher erkannt werden. Während die SPC für messbare Qualitätsmerkmale der Produkte entwickelt wurde, ist sie insbesondere in der kunststoffverarbeitenden Industrie auf Prozesskenngrößen übertragen worden. Obwohl die SPC heutzutage sehr häufig eingesetzt wird, scheint sie hier nicht das ideale Mittel zu sein und führt oft aufgrund systembedingter Eigenschaften zu Problemen, besonders wenn versucht wird, eine Prozessüberwachung anhand einer SPC-Kontrolle von Prozesskenngrößen aufzubauen [PNH92, PWHH93]. Neben der Stichprobenprüfung mit einhergehender Mittelung liegt das auch an der Tatsache, dass die Zusammenhänge zwischen Prozesskenngrößen und Qualitätsmerkmalen nicht oder nur unzureichend bekannt sind. Werden signifikante Prozesskenngrößen bei der Überwachung nicht mit in Betracht gezogen, besteht die Gefahr, dass mögliche Qualitätsprobleme nicht erkannt werden. Werden für die SPC zu viele Größen verwendet, kann eine hohe Rate an Pseudoausschuss die Folge sein. Dies gilt vor allem dann, wenn auch Werte überwacht werden, die Streuungen unterworfen sind, welche in keiner Verbindung mit Schwankungen der Qualität der Schweißung stehen (siehe auch Abschnitt 2.2.1).

Derzeit werden bei Ultraschallschweißanlagen maschinenseitig bereits verschiedene Prozesskenngrößen zur Überwachung des Schweißvorgangs herangezogen. Dabei handelt es sich stets um Einzelgrößen wie Schweißzeit, Schweißweg, eingebrachte Energie oder maximal aufgetretene Leistung. Problematisch dabei ist, dass diese Größen meist keinen direkten Zusammenhang zur Qualität der Schweißverbindung aufweisen (siehe Abschnitt 2.3). Zu verschiedenen Schweißverfahren für Kunststoffe liegen Forschungsarbeiten vor, welche die genannte Problematik anhand ausgesuchter Signalwerte behandeln [Kor96, Nat92]. Aktuell sind jedoch keine Arbeiten bekannt, die – wie hier vorgesehen – eine Einbeziehung umfassender Prozesssignale mit Zusammenhangsbildung zur Qualität ermöglichen sowie weitere Sensorsignale wie Körperschall in größerem Umfang einbeziehen. Für andere Fertigungspro-

zesse wurden Entwicklungen vergleichbarer Art bereits an unterschiedlichen Stellen behandelt [FS96, Häu95, Lie05, SH00]. Sie basieren vorwiegend auf einer rechnergestützten Beschreibung der Zusammenhänge zwischen Prozess- und Qualitätsmerkmalen der Produkte in der Nähe des "Prozessarbeitspunktes". Eines der Hauptprobleme dabei ist die sensorische, direkte oder indirekte Erfassung der relevanten Prozessschwankungen, wodurch eine modellbasierte Adaption erst möglich wird. Heute sind überwiegend Prozessmodellierungen auf statistischer Basis in Form linearer und multivariater Regression, auf der Basis von Expertenwissen oder mittels selbstadaptierender Methoden forschungs- und praxisseitig in der Anwendung (siehe Abschnitt 2.4). Eines der Hauptprobleme bei der Abbildung der betrachteten Schweißprozesse durch ein Überwachungsmodell ist in den teils instationären Phasen des Prozessablaufs zu sehen, welche zu nichtlinearen Verknüpfungen zwischen den resultierenden Prozesssignalcharakteristika und den Qualitätsmerkmalen führen. Um bei vergleichbaren Randbedingungen eine Modellbildung auf einfache Art zu bewältigen, sind beispielsweise beim Kunststoffspritzgießen Systementwicklungen bekannt [LRS01, RS01, SN05b], mit denen eine Modellierung der Prozess- und Qualitätsvarianzen selbstadaptierend auf Lernbasis ermöglicht wird. Die Modellbildung erfolgt auf der Basis künstlicher Neuronaler Netze und unter Einbezug von Prozesskenngrößen. Für diese Systementwicklung wurde durch umfangreiche Praxisversuche der Nachweis erbracht, dass diese Technologie bereits heute anwendbar ist [Neu00].

2.1 Ultraschallschweißen: Maschinentechnik, Ablauf und Einflüsse

Thermoplastische Kunststoffe – gefüllt und ungefüllt – sowie auch einige Metalle (z. B. Aluminium oder Kupferlitze) können mittels Ultraschall gefügt werden [HK97, N.N03]. Zu den Fügeverfahren, die mittels Ultraschall funktionieren, zählen neben dem in dieser Arbeit betrachten Ultraschallschweißen auch das Nieten, Bördeln und Einbetten. Das Grundprinzip beim Schweißen von Thermoplasten ist, dass durch die eingebrachte Ultraschallenergie die Fügepartner an der Fügefläche erwärmt und somit erweicht und anschließend unter Aufbringung von Druck dauerhaft verschweißt werden.

Das Ultraschallschweißen besitzt im Vergleich zu anderen Fügeverfahren wie Kleben oder Schrauben verschiedene Vorteile. Ein wesentlicher Vorteil liegt in der sehr kurzen Zykluszeit sowie der guten Automatisierbarkeit. Dies ermöglicht insgesamt eine hohe Wirtschaftlichkeit beim Einsatz dieses Fügeverfahrens [HK04]. Mittels Ultraschallschweißen können auch gas- und mediumdichte Verbindungen ohne den Einsatz zusätzlicher Dichtelemente wie beispielsweise O-Ringe hergestellt werden. Ein weiterer Pluspunkt ist, dass keine weiteren Hilfsmittel oder -stoffe wie Kleber oder Schrauben benötigt werden. Im Hinblick auf

Umweltschutz und Recyclingfähigkeit sind die mögliche Sortenreinheit der Produkte sowie ein Verzicht auf zusätzliche Stoffe wie Kleber oder Lösungsmittel weitere positive Aspekte [Dev01]. Nachteile des Verfahrens liegen darin, dass spezielle Nahtkonturen erforderlich sind und bereits bei der Konstruktion der Teile berücksichtigt werden müssen. Auch sollten Eigenheiten und Anforderungen des Schweißverfahrens an die Schweißnahtgestaltung bereits bei der Konstruktion berücksichtigt werden, um qualitativ hochwertige und reproduzierbare Ergebnisse zu erhalten. Weiter ist zu beachten, dass bei einer Dichtschweißung die Größe der Teile durch die physikalischen Grundlagen des Verfahrens begrenzt ist. Zudem führt das Verfahren insbesondere bei Systemfrequenzen von 20 kHz zu größeren Schallemissionen, so dass in der Fertigung entsprechende Schallschutzeinrichtungen notwendig sind. Aufgrund der geringen Schmelzeschichtdicke sind mittels Ultraschallschweißen nur mittlere Nahtfestigkeiten zu erzielen [Ehr04].

2.1.1 Maschinenteknik beim Ultraschallschweißen

Der prinzipielle Aufbau einer Ultraschallschweißmaschine ist in Abbildung 1.1 zu sehen. Die genaue Konfiguration der einzelnen Maschinenelemente, wie z. B. Generator- und Konverterfrequenz, Amplitudenverstärkung oder Sonotrodengeometrie, ist dabei teile- und prozessspezifisch und hängt stark von den Anforderungen des zu schweißenden Produktes ab. Im Folgenden werden die Aufgaben und Funktionen der einzelnen Komponenten kurz beschrieben (siehe auch [Ehr04, Fis06, Pot04]).

Generator: Der Generator wandelt die versorgende Netzspannung ausgangsseitig in ein hochfrequentes Hochspannungssignal (HF-Signal). Die Frequenz liegt je nach Schweißanwendung zwischen 20 und 70 kHz. Alle weiteren Maschinenelemente müssen an diese Frequenz angepasst sein. Der verwendete Frequenzbereich ist dem Ultraschall zuzurechnen und liegt somit der Namensgebung „Ultraschallschweißen“ zugrunde. Der Generator ist zudem für die Regelung der Amplitude des hochfrequenten Spannungssignals zuständig. Die vom Generator erzeugte Spannung ist die elektrische Ausgangsbasis, welche die im Folgenden vorherrschende mechanische Schwingung definiert. Die Ausgangsleistung des Generators liegt in der Regel zwischen 1000 und 6000 Watt.

Konverter: Der Konverter ist aus piezostriktiven Elementen aufgebaut und wandelt durch den (inversen) Piezoeffekt die anliegende elektrische Spannung in eine mechanische Schwingung derselben Frequenz. Die elektrische Energie wird dabei mit einem Wirkungsgrad von bis zu 98 % in mechanische Energie umgewandelt [Str95]. Die Amplitude der mechanischen Schwingung ist proportional zur Amplitude des Spannungssignals und bewegt sich im Bereich von ca. 10 µm. Die entstehenden mechanischen Schwingungen sind stets longitudinal (siehe Abbildung 1.2).

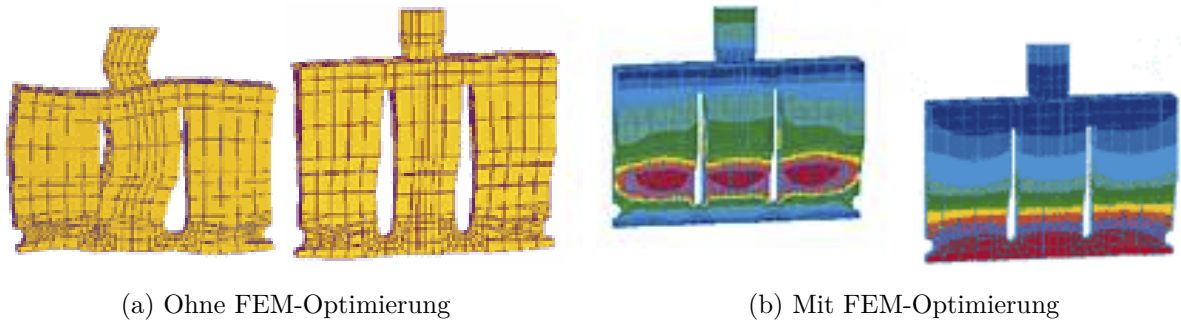


Abbildung 2.1: Schwingungsamplituden der Sonotrode mit und ohne FEM-Optimierung (Quelle: Fa. Herrmann Ultraschalltechnik)

Booster: Der Booster, auch Amplitudentransformationsstück genannt, überträgt die mechanische Schwingung vom Konverter zur Sonotrode und damit zum Füge­teil. Er ist das konstruktive Element, mit welchem das gesamte Schwinggebilde, bestehend aus Konverter, Booster und Sonotrode, an der Schweißmaschine befestigt ist. Die mechanische Kopplung muss sich dabei in einem Schwingungsknoten des Boosters befinden. Über den Booster wird damit auch die mechanische Kraft der Schweißpresse auf das Schweißteil übertragen.

Eine weitere Funktion des Boosters ist die (optionale) Verstärkung der Schwingung in einem festen Verhältnis, welches durch die Konstruktion bedingt ist. Das Transformationsverhältnis beträgt dabei normalerweise bis zu 1:3,5 und wird entsprechend der für die Schweißung erforderlichen Amplituden gewählt.

Sonotrode: Neben einer Verstärkung der Amplitude ist die Hauptaufgabe der Sonotrode die mechanische Übertragung der Schwingung sowie der Anpresskraft auf die Füge­teile. Dazu muss die Sonotrode geometrisch genau auf die Füge­teile abgestimmt sein. Der Generator sowie das Schwinggebilde arbeiten beim Ultraschallfügen in Resonanz. Um eine gleichmäßige und parallele Schwingung der Sonotrode an der Koppelfläche zu erreichen – eine notwendige Voraussetzung für reproduzierbare und hochwertige Schweißverbindungen –, werden diese bei der Konstruktion mittels FEM-Methoden ausgelegt [Haa89, Str95]. Nur so lassen sich unerwünschte Querschwingungen oder andere Unregelmäßigkeiten im Schwingverhalten der Sonotrode erkennen und beseitigen. Der Unterschied zwischen einer Sonotrode ohne und mit FEM-Optimierung ist in Abbildung 2.1 zu sehen.

Gefertigt sind Sonotroden meist aus speziellen Aluminiumverbindungen, gehärteten Stählen oder aus Titan. Aufgrund der auftretenden Schwingungsamplituden – diese bewegen sich zwischen 5 μm und 50 μm – und den in Verbindung mit der hohen Schwingfrequenz sehr hohen Beschleunigungswerten müssen höchste Anforderungen

an die Materialgüte und Verarbeitung gestellt werden. Eine mindere Gefügequalität beeinflusst die Schallgeschwindigkeit und damit das Schwingungsverhalten [Str95].¹

Amboss: Der Amboss dient zur Aufnahme des zu schweißenden Werkstücks und übernimmt dessen Fixierung. Die Werkstückaufnahme muss entsprechend der Sonotrode ausgerichtet und justiert werden und hat einen wesentlichen Einfluss auf die Reproduzierbarkeit einer hohen Schweißnahtgüte. Der Amboss sollte konstruktiv derart gestaltet sein, dass er das zu schweißende Werkstück so weit möglich unterhalb der Fügezone stützt, um die von der Schweißpresse aufgebracht und durch die Sonotrode übertragenen Schweißkräfte aufnehmen zu können, ohne dass es zu einer Aufbiegung oder Verformung des Schweißteils kommt. Auch muss der Amboss zusammen mit der Werkstückgeometrie und Sonotrode ein seitliches Abwandern der Füge Teile unterbinden. Zumeist sind die Aufnahmen aus Stahl, Aluminium oder Gießharzen gefertigt. Bei empfindlichen Bauteilen kann die Aufnahme auch mit Kork, Gummi, PTFE oder Leder ausgelegt werden. Es ist jedoch zu beachten, dass das Material der Aufnahme sich über die Dämpfung und die damit verbundene Reflektion der Schallwellen unterschiedlich auf die Schweißung auswirken kann. Ein nicht unerheblicher Anteil der eingebrachten Energie wandert über den Amboss in die Maschine ab und wird dort dissipiert – trägt also nicht zum Schweißvorgang bei [Dev01, NL96, Pot04].

Schweißpresse: Die Schweißpresse ist für den Vorschub der Sonotrode während der Schweißung und die Aufbringung der benötigten Schweiß- und Haltekräfte zuständig. Das gesamte Schwinggebilde aus Konverter, Booster und Sonotrode ist an der Schweißpresse montiert und wird durch diese bewegt. Es gibt verschiedene Ausführungen von Schweißpressen, die sich in der Art der Bewegungsgenerierung unterscheiden und damit verbunden auch unterschiedliche parametrierbare Größen aufweisen [HK04]. Die meisten Schweißpressen sind – auch aufgrund der verhältnismäßig günstigen Kosten – pneumatisch aufgebaut. Der Vorschub wird durch Beaufschlagung eines Kolbens mit Druck eingeleitet, welcher über ein schnelles Proportionalventil geregelt und somit die Schweißkraft eingestellt wird [Str95]. Der Schweißprozess kann steuerungsseitig nur über die Schweißkraft und die Schwingamplitude beeinflusst werden; der zeitliche Verlauf des Schweißwegs und damit der Fügegeschwindigkeit, welche einen wesentlichen Einfluss auf die Qualität der Schweißnaht hat (siehe Abschnitt 2.1.4), ergeben sich aus dem Prozess als resultierende Größen. Nachteilig wirkt die Kompressibilität der Luft, welche beim Schweißen zu Schwingungen und Nachgiebigkeiten führen kann und ein deutlich verzögertes Einwirken der eingestellten Schweißkraft auf das Füge teil bewirkt. Dies ist neben der Masseträgheit ursächlich für einen messbaren Abfall der

¹Aus $s(t) = \frac{\lambda}{2} \sin 2\pi ft$ ergibt sich mittels der zweifachen zeitlichen Ableitung für die Beschleunigung $a(t) = -4\frac{\lambda}{2}\pi^2 f^2 \sin 2\pi ft$. Bei einer Amplitude von $\frac{\lambda}{2} = 25\mu\text{m}$ und einer Frequenz von 20 kHz tritt demnach eine Maximalbeschleunigung von $\hat{a} = 4 * 25 * 10^{-6} * \pi^2 * (20 * 10^3)^2 \frac{\text{m}}{\text{s}^2} \approx 395000 \frac{\text{m}}{\text{s}^2}$ auf!

Schweißkraft im Moment des Aufschmelzens [HR88, Pot71].

Neben den pneumatischen existieren auch servoelektrische Schweißpressen. Neben einem Kraftprofil wie bei pneumatischen Pressen können auch Weg- und Geschwindigkeitsprofile vorgegeben und durch eine Regelung der Prozessparameter durch die Steuerung abgefahren werden. Die in der Schweißnaht wirkende Kraft ist nunmehr eine resultierende Größe und beinhaltet Informationen über den Schweißablauf. Aufgrund des Einsatzes von Servomotoren sowie der komplexen Ansteuereinheiten sind Servopressen mit deutlich höheren Kosten verbunden.

2.1.2 Prozessablauf

Nach dem Einlegen der beiden zu fügenden Werkstücke fährt die Schweißpresse die Sonotrode auf die Bauteile. Das Aufliegen der Sonotrode wird über die Steuerung entweder durch das Überschreiten einer einstellbaren Schwelle für die gemessene Kraft bestimmt oder dadurch, dass am gemessenen Wegsignal der Servopresse für eine längere Zeit keine Veränderung mehr eintritt. Ist dieser Triggerpunkt erreicht, wird der Generator angeschaltet und durch das Schwinggebilde die mechanische Schwingung erzeugt und ins Bauteil eingeleitet. Der Generator hat dabei die Aufgabe, während der Schweißung die in der Steuerung eingestellte Amplitude konstant zu halten oder das eingestellte Amplitudenprofil einzuhalten, was durch die Regelung der Anregungsfrequenz erreicht wird. Als Reaktion auf die sich ausbildenden Ankopplungsbedingungen zwischen Werkstück und Sonotrode kommt es zu einer Verschiebung der Resonanzfrequenz des Schwinggebildes [vLvN96]. Diese wird mittels der Regelung der Anregungsfrequenz durch den Konverter kompensiert, so dass die vorgegebene Amplitude eingehalten wird. Die mechanischen Schwingungen werden im Bauteil absorbiert und an der Grenzfläche reflektiert. Verschiedene physikalische Effekte wie die Molekular- und die Grenzflächenreibung der Füge Teile führen zu einer Erwärmung und dem Erweichen des Kunststoffes. Mit steigender Temperatur erhöht sich auch die Dämpfung des Kunststoffes und verstärkt die Absorption der mechanischen Schwingung. Die Reaktion beschleunigt sich somit selbst und es wird ein steigender Anteil der Schwingungsenergie in Wärme umgesetzt. Im Bereich der Fügezone bildet sich aufgrund der entstehenden flüssigen Schmelze eine Schallsperre, die zu einer weiteren Erhöhung der Reflektion und Absorption führt. Um ein definiertes Anschmelzen und somit einen reproduzierbaren Prozessablauf zu erhalten, sind besondere Nahtgeometrien an den Fügepartnern vorzusehen. Am häufigsten werden Energierichtungsgeber (ERG) oder die so genannte Quetschnaht verwendet. Der von der Schweißpresse aufgebraachte Druck bewirkt ein Fließen der entstehenden Schmelze und somit eine messbare Verfahrbewegung des Schwinggebildes: den Schweißweg. Nach dem Abschalten des Ultraschalls wird der Anpressdruck noch eine Weile aufrecht erhalten. In dieser Halte- bzw. Abkühlzeit erkaltet die Schmelze wieder und durch Erstarren wird die endgültige Schweißverbindung ausgebildet. Die Prozessparameter wie Amplitude, Kraft,

Schweißdauer und Haltezeit sind prozessspezifisch zu wählen und hängen von Material und Bauteilgeometrie ab.

Für das Abschalten des Ultraschalls gibt es mehrere Möglichkeiten, die im Folgenden zusammen mit einigen resultierenden Eigenschaften und Anwendungsgebieten aufgelistet sind (nach [Fis06] und [HK04]):

Konstante Schweißzeit: Das Schweißen wird nach dem Ablaufen einer fest eingestellten Zeit beendet. Alle anderen Schweißparameter (Schweißweg, Leistung, eingebrachte Energie, ...) schwanken mehr oder weniger. Die unterschiedlichen Fügewege führen zu unterschiedlichen Schmelzevolumina und dadurch zu Schwankungen der Schweißnahtfestigkeit und -dichtigkeit. Diese Schweißart wird meist nur für einfache Schweißungen mit geringen Anforderungen angewendet.

Konstanter Schweißweg: Das Schweißen wird nach Erreichen eines bestimmten Schweißwegs – bezogen auf den Startpunkt des Schweißvorgangs – beendet. Diese Art der Prozesssteuerung führt zu sehr stabilen Schweißprozessen und der konstante Fügeweg garantiert ein gleichmäßiges Schmelzevolumen und damit reproduzierbare Nahtfestigkeiten. Toleranzen der Fügepartner untereinander (z. B. aufgrund von Formnестtoleranzen beim Spritzgießen) sowie Toleranzen zwischen Teil und Werkzeug werden bei gleichbleibender Nahtqualität ausgeglichen. Schweißzeit, eingebrachte Leistung und Abmessungen nach dem Schweißen können variieren. Dieser Schweißmodus wird im Weiteren auch als RPN²-Modus bezeichnet.

Absoluter Wegendpunkt: Der Generator wird abgeschaltet, sobald das Wegmesssystem – bezogen auf die Ruhelage des Schwinggebildes – einen bestimmten Wert überschreitet. Dieser Modus dient zum Ausgleich von Höhentoleranzen der Fügepartner und zum Erzielen eines vorgegebenen Maßes in Schweißrichtung nach dem Verschweißen. Er sollte eingesetzt werden, wenn der Abmessung nach dem Schweißen ein höherer Stellenwert zukommt als reproduzierbaren Schweißnahtfestigkeiten. Schweißzeit und eingebrachte Energie sind variabel und der variierende Schweißweg hat schwankende Schmelzevolumina und somit schwankende Festigkeiten zur Folge. Bei großen Toleranzfeldern führt dies bei kleinsten Teilen zu niedrigen Festigkeiten.

Energie: Sobald eine vorgegebene Energie – definiert als Integral der vom Generator abgegebenen Leistung über der Zeit – in das Bauteil eingebracht ist, wird die Schweißung beendet. Für Teile ohne definierte Nahtgestaltung (z. B. Folien, Textilien, Membrane, ...) ist dies der Standardmodus. Dickenunterschiede, variierende Feuchtegehalte und Temperaturen sowie schwankende Schweißkräfte werden dadurch kompensiert. Dieser Schweißmodus liefert jedoch relativ hohe Prozessschwankungen, so dass er im

²Reference Point Numeric – Triggerpunkt

Allgemeinen nur für einfache Anwendungen oder bei geringen Anforderungen eingesetzt wird.

Leistung: Beim Erreichen einer bestimmten Generatorleistung (Leistungsspitze) wird der Schweißprozess beendet. Dieser Modus findet insbesondere beim Nieten Anwendung.

Kombination: Bei neueren Steuerungen ist es auch möglich, eine Kombination aus den anderen Schweißmodi zu verwenden. Die Schweißung wird beendet, sobald das erste der definierten Abschaltkriterien eintritt. Dadurch wird auch bei Prozessen mit größeren Toleranzen, z. B. bei Montageprozessen oder Spritzgußwerkzeugen mit vielen Kavitäten, ein sicheres Verschweißen möglich, auch wenn die Schweißung mit den anderen Modi nicht mehr beherrschbar ist. Die Kombination bietet beste Adaptionsmöglichkeiten an kritische Anwendungen mit sehr engem Prozessfenster. Beschädigungen an innen liegenden Strukturen bei hochpreisigen Bauteilen können mit diesem Modus am besten vermieden werden.

Insgesamt bleibt festzustellen, dass die erzielbare Nahtqualität stark von den Maschineneinstellparametern und den daraus resultierenden Prozessbedingungen abhängig ist.

2.1.3 Einflussgrößen

Im Prinzip sind alle thermoplastischen Kunststoffe mit Ultraschall schweißbar, wobei jedoch die Kombination von Werkstoffen einigen Beschränkungen unterliegt. So ist ein Verschweißen von amorphen mit teilkristallinen Werkstoffen nicht möglich. Thermoplaste aus der gleichen Produktklasse hingegen können meist miteinander verschweißt werden, wobei es problematisch sein kann, wenn insbesondere die Steifigkeit zu große Unterschiede aufweist [LE85]. Bei den zu schweißenden Fügeteilen wirken sich mehrere Faktoren auf den Schweißverlauf und damit das Schweißergebnis aus. Um qualitativ zufriedenstellende Ergebnisse zu erzielen, müssen Maschine, Schwinggebilde und Schweißparameter den entsprechenden Faktoren, die im Folgenden näher erläutert werden, Rechnung tragen [Ehr04, N.N07b].

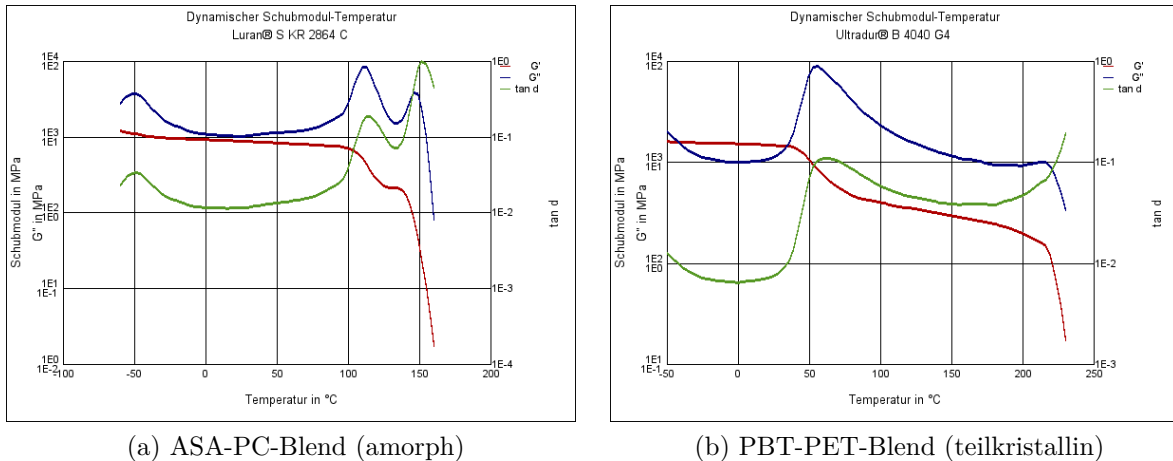
Werkstoffauswahl und -eigenschaften

Grundsätzlich gibt es zwei Arten von Kunststoffwerkstoffen, die sich in ihrem molekularen Aufbau grundsätzlich unterscheiden: Amorphe und teilkristalline Werkstoffe. Die Molekülketten amorpher Kunststoffe bilden im festen Zustand große Knäuel aus – es herrscht also keine spezielle Ordnung vor. Bekannte Werkstoffe dieser Gattung sind beispielsweise Acrylnitril-Butadien-Styrol (ABS), Polycarbonat (PC), Polymethylmethacrylat (PMMA, Plexiglas), Polystyrol (PS) und Polyvinylchlorid (PVC). Die Molekülketten teilkristalliner Kunststoffe richten sich dagegen aus und bilden – maßgeblich aufgrund polarer Anteile – eine Art regelmäßiges Kristallgitter. Bekannte Vertreter sind die Vielzahl an Poly-

amiden (PA), Polyethylen niedriger oder hoher Dichte (PE-LD, PE-HD) oder Polypropylen (PP). Dieser grundsätzlich unterschiedliche Aufbau wirkt sich auf die physikalischen Eigenschaften des Werkstoffs und somit auf die verschiedenen Materialkennwerte signifikant aus. Durch verschiedene Modifikationen und die Copolymerisation oder die Beimischung verschiedener Additive oder Füllstoffe lassen sich die Eigenschaften in weiten Grenzen den Bedürfnissen anpassen und teilweise maßschneidern, so dass zwischenzeitlich eine fast unerschöpfliche Anzahl an verschiedensten Kunststofftypen verfügbar ist [Mic06, HHH04]. In der frei zugänglichen Datenbank von MaterialDataCenter für Kunststoffe (Produktdatenblätter etc.) sind allein über 4600 Typen an Polyamiden von verschiedenen Herstellern gelistet [CAM, Mat].

Beim Ultraschallschweißen sind insbesondere der dynamische E-Modul, der Schubmodul G und der mechanische Verlustfaktor $\tan \delta$ für das Verhalten maßgeblich, wobei diese Werkstoffeigenschaften eine starke Temperaturabhängigkeit aufweisen und sich bei amorphen und teilkristallinen Werkstoffen signifikant unterscheiden. Die Temperaturabhängigkeit der genannten Größen bewirkt, dass sich diese Größen während des Schweißvorgangs aufgrund der auftretenden Erwärmung verändern [NL96, MP69]. Werkstoffe mit hohem Schubmodul G (und somit auch hohem E-Modul) zusammen mit einem niedrigen mechanischen Verlustfaktor $\tan \delta$ bis zur Glasübergangstemperatur bewirken eine gute Schallleitfähigkeit und damit gute Schweißigenschaften, da die Energie der Ultraschallschwingung verlustarm bis zur Schweißnaht geleitet wird, wo sie für die Schweißverbindung in Wärme umgewandelt werden soll. Diese Eigenschaften sind vorwiegend bei amorphen Kunststoffen zu finden [NL96]. Bei teilkristallinen Werkstoffen fällt die Schubmodulkurve bis zur Glasübergangstemperatur, welche dem Schmelzpunkt entspricht, stark ab, während gleichzeitig der Verlustfaktor ansteigt. Dies entspricht einer zunehmenden Dämpfung mit zunehmender Temperatur. Dies wiederum bedeutet, dass zum Schweißen deutlich höhere Energien notwendig sind und sich somit meist auch längere Schweißzeiten ergeben [LE85, NL96]. Exemplarische Verläufe für den Schubmodul G und den mechanischen Verlustfaktor $\tan \delta$ für je einen amorphen und teilkristallinen Werkstoff sind in Abbildung 2.2 dargestellt. Das Verhalten der Temperaturkurven lässt sich konsistent mit dem sich temperaturbedingt ändernden Verlustfaktor erklären (gemessenen in [TDW83]). Der mechanische Verlustfaktor $\tan \delta$, sprich die Dämpfung, gibt Auskunft über den Verlust der Schallenergie auf dem Weg von der Sonotrodenstirnfläche zur Fügezone [LE85]. Die verlorene Energie wird dabei in Wärme umgesetzt.

Weitere wichtige Werkstoffparameter sind Dichte, Reibungskoeffizienten und die spezifische Wärmekapazität c_p eines Werkstoffs [MP69]. Die spezifische Wärmekapazität beeinflusst dabei die Energie, die zum Schweißen notwendig ist. Je höher die Wärmekapazität eines Werkstoffs ist, desto mehr Energie muss aufgewendet werden, um eine Temperaturerhöhung und damit ein Aufschmelzen des Werkstoffs zu erreichen. Dies ist auch deshalb der



(a) ASA-PC-Blend (amorph)

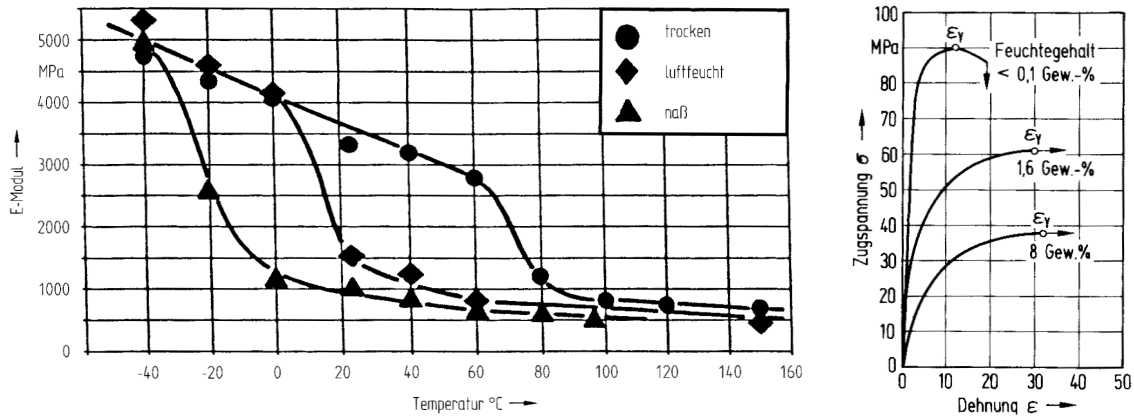
(b) PBT-PET-Blend (teilkristallin)

Abbildung 2.2: Schubmodul G und mech. Verlustfaktor $\tan \delta$ als Funktion der Temperatur bei verschiedenen Werkstoffen (Quelle: CAMPUS-Werkstoffdatenbank für Kunststoffe der BASF [CAM])

Fall, da ein Teil der eingebrachten Energie aufgrund der Dämpfung auch zur Erwärmung des Materials führt, das sich zwischen Einbringung der Schwingung durch die Sonotrode und der Schweißnaht befindet [NL96]. Nach dem Erreichen des Schmelzpunkts spielt der Schmelzetemperaturbereich eine Rolle. Er wird nach unten durch die Schmelztemperatur und nach oben durch die Zersetzungstemperatur begrenzt. Letztere ist dadurch gekennzeichnet, dass der Thermoplast ab dieser Temperatur beginnt, sich chemisch zu zersetzen, so dass beim Schweißprozess diese Grenze nicht erreicht werden sollte, um eine Materialschädigung zu vermeiden. Je enger der Schmelzetemperaturbereich ist, desto sorgfältiger ist der Schweißprozess einzustellen und zu führen.

Auch die Schmelzeviskosität hat einen Einfluss auf den Schweißprozess und die erzielte Schweißnahtqualität, da sie maßgeblich das Fließverhalten der Schmelze während der Schweißung beeinflusst. Hochmolekulare Kunststoffe sind zähfließender und benötigen mehr Energie, bis sie eine für die Verschweißung geeignete Viskosität erreicht haben. Werkstoffe mit niedrigerem Molekulargewicht schmelzen leichter und weisen eine niedrigere Viskosität auf. Dies führt dazu, dass die Schmelze unter Druck schneller aus der Fügezone austreten kann. Die Schweißparameter sowie die geometrische Gestaltung der Fügezone (siehe unten) müssen dementsprechend gewählt werden. Eine Abschätzung über die Schmelzeviskosität beim Schweißen liefert der MVR-Wert³ in $\text{cm}^3/10 \text{ min}$ nach ISO 1133 (auch Schmelze-Volumenrate genannt). Dieser Wert dient zur Charakterisierung des Fließverhaltens von Kunststoffen bei bestimmten Temperatur- und Druckbedingungen. Ein niedriger MVR-Wert entspricht einer hohen und ein hoher MVR-Wert einer niedrigen Schmelzeviskosität. Dabei ist zu beachten, dass die Viskosität von Kunststoffen nicht nur eine Temperaturabhängigkeit, sondern auch eine Frequenzabhängigkeit aufweist [NL96]. Auch die

³MVR = engl. Melt Volume Rate



(a) Elastizitätsmodul über der Temperatur bei verschiedener relativer Feuchte

(b) Spannungs-Dehnungs-Diagramm bei verschiedener relativer Feuchte

Abbildung 2.3: Veränderung der Werkstoffeigenschaften von PA 66 (Quelle [Erh08])

relative Molekularmasse der Kunststoffe, d. h. die Kettenlänge der Moleküle sowie deren Verteilung, wirkt sich auf verschiedene Eigenschaften wie Dichte, Kristallisation, Steifigkeit oder Viskosität und somit auf den Schweißprozess aus [Erh08]. Exemplarisch hat LAND für die Werkstoffe Polycarbonat (PC), Polyamid-6 (PA6) und Acrylnitril-Butadien-Styrol (ABS) deutliche Unterschiede beim Schweißverhalten nachgewiesen [Lan78]. Neben den unterschiedlichen Schmelzeviskositäten und damit dem Fließverhalten der Werkstoffe sind diese teils auch auf den variierten Feuchtegehalt zurückzuführen. Je nach chemischer Struktur – entscheidend sind dabei insbesondere polare Abschnitte – sind Kunststoffe teils sehr hygroskopisch, d. h. sie neigen zur Aufnahme von Wasser, was zu einem Ansteigen des relativen Feuchtegehalts führt. Der relative Feuchtegehalt – meist angegeben in Gewichtsprozent – kann sich sehr auf die mechanischen Eigenschaften wie Festigkeit und Elastizitätsmodul und somit das Schweißverhalten auswirken. LAND [LE85] führt Ergebnisse einer Zugprüfung des E-Moduls nach DIN 53457 an, wonach bei hochmolekularem PA der E-Modul von konditionierten Teilen (d. h. mit hoher Feuchte) von 3000 N/mm² auf 1200 N/mm² absinkt. ZIEGLTRUM [Zie00] untersucht u. a. den Einfluss des Feuchtegehalts auf die Schweißbarkeit und erzielbare Nahtfestigkeit für verschiedene Werkstoffe. Während bei den untersuchten amorphen Werkstoffen die Festigkeit um maximal 15% absinkt, wirkt sich ein höherer Feuchtegehalt bei teilkristallinen Werkstoffen und insbesondere bei PA teils deutlich aus. Bei feuchtem PA geht der Einfluss teilweise so weit, dass der Energierichtungsgeber teilweise nur noch verpresst und nicht mehr durchgängig verschweißt wird. In Abbildung 2.3 sind beispielhaft die Auswirkung des Feuchtegehalts sowie der Temperatur auf den E-Modul für PA 66 dargestellt.

Bei Thermoplasten kommen verschiedenste Zusatzstoffe zum Einsatz. Die oft eingesetzten Füll- und Verstärkungsmaterialien wie Glasfasern, Kohlenstofffasern oder Talkum erhöhen die Dichte und damit die Schalleitfähigkeit. Andere Zusatzstoffe wie Farbstoffe, Brandschutzmittel, Stabilisatoren oder Weichmacher sorgen für spezielle geforderte oder angepasste Eigenschaften des Werkstoffs. Die Art und Menge der Zusätze kann zwischen verschiedenen Werkstoffen stark variieren und sich auf das Schweißverhalten insgesamt auswirken [Dev01, Net92]. Modifizierte Kunststoffe haben gegenüber den Ausgangsmaterialien ein geändertes Schweißverhalten – sichtbar am veränderten Verlauf von Schubmodul und mechanischem Verlustfaktor [LE85]. LAND stellt fest, dass glasfaserverstärktes Polycarbonat (PC) zum Aufschmelzen ca. 50% länger benötigt als unverstärktes PC [Lan78]. LIU und CHANG haben dargestellt, dass glasfaserverstärktes Polypropylen (PP) beim Ultraschallschweißen höhere Festigkeiten liefert als normales PP [LC01]. Dabei wird weniger Energie benötigt, was damit erklärt wird, dass Glasmaterialien einen höheren E-Modul und geringeren Verlustfaktor haben, was wiederum dazu führt, dass mehr Energie zum Energierichtungsgeber gelangt und dort zum Verschweißen beiträgt.

Neben der schwankenden Verteilung der relativen Molekularmasse sind auch die schwankende Zusammensetzung der Zusatzstoffe sowie Feuchteschwankungen für die oftmals auftretenden Unterschiede zwischen verschiedenen Materialchargen im Verarbeitungsprozess von Thermoplasten ursächlich. Dies gilt neben dem Schweißen beispielsweise auch für das Spritzgießen als Urformprozess [JM04, Mic06].

Konstruktive Gestaltung

Bereits während der Planungs- und Konstruktionsphase eines Produkts wird die erreichbare Schweißnahtgüte maßgeblich mitbestimmt [MKK97]. Die Konstruktion sollte schweißgerecht ausgeführt sein – sowohl im Hinblick auf die Herstellung der Fügepartner (z. B. Bindnähte, Einfallstellen durch Materialanhäufung) als auch den Schweißprozess (Schalleinleitung, Schalleitung, Aufbiegung etc.). Die Konstruktion sollte sich nach den an die Schweißnaht gestellten Anforderungen richten im Hinblick auf Belastbarkeit, Dichtigkeit oder Optik unter Berücksichtigung von Schmelzeaustritt und Partikeln [Dev01, Ehr04, Vow80]. Neben der Auswahl des Werkstoffs, welche die mechanischen Grundeigenschaften festlegt, spielen unter anderem folgende, konstruktiv beeinflussbare Faktoren eine Rolle:

- Lage und Ausbildung der Fügeflächen (Planizität, Spielfreiheit, ...),
- Positionierung und Führung der Fügepartner zueinander,
- Gestaltung der Nahtgeometrie: ERG, Quetschnaht, ... (Abbildung 2.4),
- Freiliegen des Abschweißwegs,

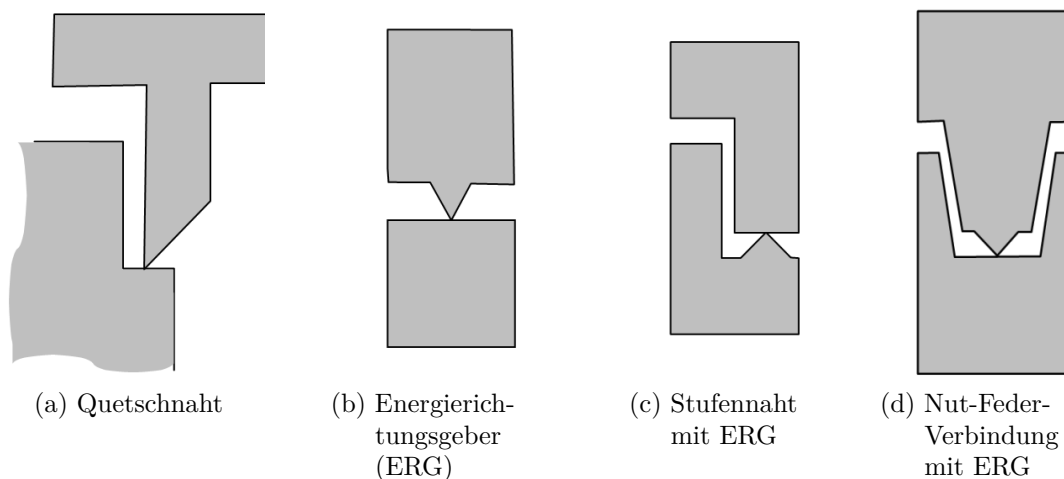


Abbildung 2.4: Verschiedene Formen der Schweißzonen-gestaltung

- Anlagefläche und Ankopplung der Sonotrode am Oberteil,
- Auflage und Fixierung des Unterteils sowie Abstützung der Schweißfläche durch Am-boss.

Je nach Abstand zwischen Schwingungseinleitung durch die Sonotrode und Schweißzone wird zwischen Nahfeld und Fernfeld unterschieden, wobei die Grenze bei ca. 6 mm liegt. Fernfeldschweißungen lassen sich am besten mit steifen Thermoplasten durchführen, da eine durch die Schweißkraft bedingte Aufbiegung der Teile möglichst gering gehalten oder ganz vermieden werden sollte. Aufgrund des Schubmoduls und mechanischen Verlustfaktors werden meist nur amorphe Kunststoffe im Fernfeld verschweißt. Teilkristalline Kunststoffe, die meist schon deutlich unterhalb der Schmelztemperatur einen deutlichen Dämpfungsanstieg aufweisen, sollten nur im Nahfeld verschweißt werden.

Um einen guten Schweißablauf zu erreichen, sollte die Fügefläche in einer Ebene und möglichst parallel zur Sonotrodenstirnfläche liegen, an welcher die Schwingungseinleitung erfolgt. Die günstigste und effektivste Schalleinleitung ergibt sich, wenn sich die Sonotrode direkt über der Schweißnaht ankopple.

Der Gestaltung der Schweißzone kommt insbesondere für ein reproduzierbares und gleichmäßiges Aufschmelzen eine große Bedeutung zu. Eine Reduzierung der Fügefläche führt in Verbindung mit der Fügekraft zu einer starken Erhöhung der Druckwechselbeanspruchung infolge der mechanischen Schwingungen. Aufgrund der Dämpfungseigenschaften der Kunststoffe führt dies zu einer schnellen Erwärmung und dem Aufschmelzen des Materials in der Fügezone. Die Reduzierung der Fügefläche wird konstruktiv durch eine anfangs linienförmige Kontaktfläche erzielt, welche sich durch das Aufschmelzen verbreitert, um die geforderte Schweißnahtdicke und das notwendige Schmelzevolumen zu erreichen. In der Praxis haben sich dazu hauptsächlich zwei Konstruktionsvarianten etabliert: Die Nahtge-

staltung mittels Energierichtungsgeber (ERG) sowie die Quetschnaht (QN). Über Höhe und Spitzenwinkel des ERG können das entstehende Schmelzevolumen und der mögliche Schweißweg gestaltet werden. Mittels Spalte und Spielpassungen bei der Nahtgeometrie können Volumina für die entstehende Schmelze definiert werden, um den Schmelzeffluss zu lenken, den Schmelzeaustrieb für eine Mindestnahtbreite zu steuern oder zu unterbinden sowie bei optischen Anforderungen die Schweißnaht zu verbergen (siehe Abbildung 2.4). Es ist dabei wichtig, den ERG konstruktiv im richtigen Abstand zum Ort der Schalleinleitung zu platzieren. Er sollte in einem Schwingungsbauch liegen, welche sich bei $\lambda/4$, $3\lambda/4$, ... in Bezug auf die Wellenlänge befinden [BC89]. Der ERG wird vorzugsweise verwendet und führt zu einer Erhöhung von Schweißgeschwindigkeit und Qualität [CCCL00]. Auch die geometrische Form des ERG beeinflusst die erzielbare Nahtfestigkeit. Während ein ERG mit 90° Spitzenwinkel für die meisten amorphen Werkstoffe geeignet ist, wird für Polycarbonat, Acrylate und teilkristalline Werkstoffe ein ERG mit 60° bevorzugt [Dev01]. Neben dreieckigen ERG mit verschiedenen Winkeln sind auch rechteckige oder halbrunde ERG untersucht worden. Nach CHUAH [CCCL00] sollte aufgrund theoretischer Überlegungen eine Halbkreisgeometrie immer höhere Festigkeiten liefern, wobei dies bei Versuchen mit verschiedenen Werkstoffen nur teilweise zutraf [LC01, LLC⁺99]. Mit einer Quetschnahtgeometrie können um 30 % bis 50 % höhere Festigkeiten erzielt werden als bei Einsatz eines ERG. Eine Quetschnaht ist insbesondere dann anzuraten, wenn eine Dichtigkeit gegenüber Gas, Dampf oder Flüssigkeit erreicht werden muss, wobei sie bei dünnen Wanddicken jedoch erhebliche konstruktive Schwierigkeiten mit sich bringt [LE85]. Die Quetschnahtgeometrie kommt vorwiegend bei Feuerzeugtanks und Kleingehäusen zum Einsatz [Vow80]. Auch bei Kunststoffen, die beim Erreichen des Schmelzpunkts in einem engen Temperaturbereich vom festen in den flüssigen Zustand übergehen – also bei teilkristallinen Thermoplasten wie POM und PA –, ist die Quetschnaht oft notwendig, da bei einer ERG-Geometrie das Material unkontrolliert wegspritzt [Mei00].

Herstellungbedingungen

Die meisten zu verschweißenden Teile werden mittels Kunststoffspritzguss hergestellt, wobei jedoch auch durch Blasformen oder Extrudieren gefertigte Teile mit Ultraschall gefügt werden. Die Werkstoffeigenschaften und Bauteilgeometrie sind beim Spritzgießen keine unveränderlichen Größen, sondern stark vom Herstellungsprozess abhängig [Rie88]. Erfahrungen aus der Praxis zeigen, dass Fehlschweißungen ihre Ursache oft in Unzulänglichkeiten oder Schwankungen der Urformprozesse der Fügepartner haben [Ehr04, MKK97, MK94a, Str95]. Zu den Ursachen zählen neben chargenbedingten Änderungen der Materialeigenschaften unter anderem Maßabweichungen (verursacht durch Schwindung, verschiedene Formnester, Werkzeugverschleiß), Verzug, Gewichtsschwankungen oder Oberflächenfehler (Einfallstellen, ungeeignete Fügeflächengeometrie, unvollständig ausgeformte Werkstücke). Auch In-

homogenitäten in der Schweißnaht, z. B. durch Lunker oder Bindenähte, sowie Zersetzungen oder Entmischungen durch Verarbeitungsfehler oder zu hohe innere Spannungen können zu Problemen im Schweißprozess führen. KORTE [Kor96] hat mit zwei teilkristallinen, unverstärkten Werkstoffen (PP und POM) den Einfluss des Spritzgießens auf den Ultraschallschweißprozess untersucht. Durch eine gezielte Variation von Einspritzgeschwindigkeit sowie Masse- und Werkzeugtemperatur beim Spritzgießen werden Werkstücke mit unterschiedlichen Sphärolithdurchmessern und Orientierungen hergestellt. Die resultierenden Werkstoffeigenschaften erweisen sich bei Zugversuchen an CAMPUS-Schulterproben als signifikant unterschiedlich. Für das Schweißen dieser Werkstücke mit unterschiedlicher Morphologie wird ein signifikanter Einfluss auf den Prozessverlauf beim Schweißen, z. B. auf Schweißgeschwindigkeiten und erzielte Schweißwege, sowie die erzielbare Festigkeit nachgewiesen.

Der Einfluss der Spritzgießparameter auf den Schweißverlauf und somit auf die erzielte Schweißverbindung ist auch für andere Schweißverfahren nachgewiesen. HANNING und NATROP [Han94, Nat92] haben einen sehr starken Einfluss der Spritzgießparameter auf die Teilemaße und die Festigkeit der Schweißverbindung für das Heizelementschweißen aufgezeigt. Auch für das Vibrationsschweißen sind entsprechende Zusammenhänge belegt [Ueb95]. Bei HOFFMANN [HHL07] wird die Teilmorphologie durch unterschiedliche Masse- und Werkzeugtemperaturen beeinflusst und der Nachweis für die Abhängigkeit der Schweißnahtfestigkeit von Veränderungen im Spritzgießprozess für das Laserdurchstrahlschweißen erbracht.

Eine weitere wichtige Rolle spielen die Lagerbedingungen und die Historie der Fügepartner zwischen der Teileherstellung und der Schweißung. Je nach Material haben z. B. Lagerzeit und -klima zwischen Herstellung und Schweißprozess einen entscheidenden Einfluss auf die erzielbare Schweißnahtqualität [MKK97]. Bei teilkristallinen Werkstoffen spielt der Kristallisationsgrad eine wesentliche Rolle für die Schweißung. Spritzfrische Teile besitzen gegenüber gelagerten Teilen eine geänderte Gefügestruktur, was sich wiederum – wie bei geänderten Spritzparametern – deutlich im Schweißablauf zeigt. Die für das Aufschmelzen benötigte Energie hängt von der Temperatur der Fügepartner ab, so dass sich unterschiedliche Teiletemperaturen beim Schweißen (z. B. im Sommer oder Winter) direkt in schwankenden Schweißparametern (Schweißzeit, Schweißenergie) abbilden. Alle diese Einflüsse hängen direkt mit Werkstoffeigenschaften zusammen, die sich auf den Schweißprozess auswirken (siehe S. 25). Die Lagerung der Teile zwischen Herstellung und Schweißung kann sich – je nach Material – zudem stark auf den Feuchtegehalt der Teile auswirken. LAND [Lan78] war es nicht möglich, Teile aus PA erfolgreich zu verschweißen, wenn diese einen hohen Feuchtegehalt aufweisen. Insbesondere bei PA bewirkt ein höherer Feuchtegehalt u. a. eine Abnahme der Steifigkeit und kann in der Folge auch das Auftreten poröser Fügenähte verursachen [LE85].

Schweißparameter

Einen großen Einfluss auf die Schweißung haben die Maschinenparameter, wobei die wesentlichen Größen die Amplitude sowie die auf das Werkstück einwirkende Schweißkraft sind. Schweißkraft und Amplitude bestimmen zusammen maßgeblich den Ankoppelungszustand zwischen Sonotrode und Füge teil und somit den Energieeintrag, der zum Aufschmelzen in der Fügezone führt [Kor96, Net92]. Die Bildung der ersten Schmelze erhöht die Dämpfung und somit die Energieabsorption, was wiederum den Aufschmelzvorgang beschleunigt und somit insgesamt eine Rückkopplung und Kettenreaktion in Gang setzt [HR88]. Ist der Fügedruck zu hoch, wird die Schmelze aus der Fügezone ausgedrückt, was die Kettenreaktion abdämpft. Während die statische Fügekraft zu einem Quetschfließen führt, ist die Schmelze durch die hochfrequenten dynamischen Schwingungen auch dynamisch belastet, weswegen bei einer Modellierung der physikalischen Vorgänge Trägheits- und Elastizitätseffekte berücksichtigt werden müssen. Da hierfür jedoch keine ausreichenden Modellansätze existieren, wird das Verhalten durch dimensionslose Kennzahlen – genauer Reynolds-, Deborah-, Mach- und Weissenbergzahl – abgeschätzt. Das Ergebnis zeigt, dass die Schmelze mit zunehmendem Schweißprozessfortschritt ein immer elastischeres Verhalten aufweist und somit wie eine harte Feder reagiert [Ben87, Kor96, Net92].

Die Fließvorgänge in der Schmelze führen zu Schmelzedeformationen in Form von Schub und Dehnung und sind für die Ausbildung einer spezifischen Morphologie verantwortlich. Diese weist einen direkten Zusammenhang zur erzielbaren Schweißnahtfestigkeit auf [CCCL00, MK93, MKDE94, Str95]. Zwischen der maximalen Dehngeschwindigkeit und der Schweißnahtfestigkeit besteht ein Zusammenhang derart, dass die Festigkeit mit zunehmender Dehngeschwindigkeit abnimmt, was teils auch auf Eigenspannungen und fließbedingte Molekülorientierungen zurückzuführen ist [KM94]. Ein Versagen der Schweißnaht erfolgt stets in der Mitte, wo die höchsten Dehngeschwindigkeiten auftreten [Net92]. Die Fließvorgänge sind neben der Amplitude und Schweißkraft aufgrund der temperaturabhängigen Viskosität der Werkstoffe auch stark von der Temperatur abhängig. Die Temperatur in der Fügezone wird dabei hauptsächlich durch ein Energiegleichgewicht zwischen Dissipation, Wärmefluss ins feste Material und die Kühlung durch Strömung aufgrund des Schmelzeaustriebs beeinflusst [NL96]. Da die Erwärmung der Schweißnaht proportional zum Quadrat der eingeleiteten Energie ist, welche wiederum proportional zum Produkt aus Amplitude und Sonotrodenfläche ist, kann die Schmelztemperatur durch die Amplitude beeinflusst werden [Str95]. Das Aufschmelzen und die anschließenden Fließvorgänge führen zu einem Schweißwegverlauf, der sich aus dem Verfahrensweg der Sonotrode ergibt. In Prozessanalysen hat sich der zeitliche Schweißwegverlauf als eine sehr wesentliche Prozessgröße herausgestellt [Net92], wobei sich verschiedene Phasen unterscheiden lassen. Diese sind in Abschnitt 2.1.4 näher erläutert. Die Abschmelzgeschwindigkeit ist dabei nach HABENICHT u. a. [HR88] ein direktes Maß für die vom Material absorbierte und in Wärme umgewandelte Energie.

Da sich bei Beginn des Abschmelzens der ERG verkürzt und das Schwingsystem durch die Masseträgheit nicht gleich nachkommt, kann anhand eines kurzzeitigen Einbruchs der statischen Fügekraft der Beginn des Aufschmelzens ermittelt werden [Pot71].

Da sich Amplitude und Schweißkraft in vielfältiger Weise beeinflussen und sich auf den Schweißverlauf einwirken, ist nachvollziehbar, dass die Einstellung und Optimierung des Schweißprozesses auch heute noch stark erfahrungsbasiert zumeist nach dem Trial-and-Error-Prinzip erfolgt [Kor96, LLC⁺99, NL96]. Für eine erste Starteinstellung werden manchmal auch tabellierte Erfahrungswerte für den Fügedruck verwendet, die vom eingesetzten Material abhängen. Zusammen mit der teilespezifischen Größe der Fügefläche kann so über $F = p \cdot A$ ein erster Wert für die Schweißkraft ermittelt werden. In der Praxis werden dabei meist zu hohe Schweißkräfte verwendet. Allgemein sollte die niedrigste Kraft gewählt werden, die zu guten Schweißnähten führt, während die Amplitude, sofern Sonotrodenform und Bauteil – insbesondere innenliegende Elektronik – es erlauben, möglichst hoch gewählt werden sollte [LE85]. Nach NETZE kann dabei meist ein amplituden- und schweißkraftspezifisches Festigkeitsoptimum erreicht werden [Net92].

Andere Untersuchungen haben ergeben, dass sich durch eine Reduzierung der Schweißkraft gegen Ende des Schweißprozesses höhere Festigkeiten erzielen lassen [MK93, MKW07]. Solche Schweißkraftprofile werden durch den Einsatz schnellschaltender Servoventile ermöglicht. Teilweise konnte durch diese Prozessführung die Reißkraft um ca. 30 % erhöht werden [MK93]. Dies ist konsistent damit, dass die maximale Dehngeschwindigkeit in der Schmelze durch eine Kraftreduzierung verringert wird [MKDE94]. Ein ähnlicher Zusammenhang zwischen niedriger Schweißkraft und höherer Festigkeit aufgrund größerer Schmelzeschichtdicken ist auch für das Heizelementschweißen beschrieben [TDW83].

Eine andere Möglichkeit zur Prozesssteuerung ist die Verwendung von Amplitudenprofilen. Wird dabei mit einer hohen Amplitude gestartet und diese im Prozessverlauf abgesenkt, führt dies zu ca. 50 % höheren Schmelzeschichtdicken. Damit geht direkt eine deutliche Erhöhung der Festigkeit einher – teils bis zu 80 % bei gleichzeitiger Reduktion der Standardabweichung [Str95].

Die prinzipiellen Einflussfaktoren auf den Schweißprozess und ihr Wirkungsort in der Prozesskette sind nochmals in Abbildung 2.5 zusammenfassend dargestellt. Ihre große Anzahl sowie die gegenseitige Beeinflussung erschweren die Prozessführung und Qualitätssicherung immens. Außerdem ist es aufgrund der Vielzahl von Effekten, die wie dargestellt zumeist nur grob abgeschätzt werden können, nicht möglich, ein allgemein gültiges Modell zu erstellen. Auch das Prozessführungskonzept, d. h. ob die Abschaltung nach relativem Weg oder eingestellter Schweißzeit erfolgt, führt bei manchen Werkstoffen zu unterschiedlich hohen Festigkeiten; bei anderen Werkstoffen jedoch ist keine Abhängigkeit von der Strategie zu erkennen (z. B. für PC bzw. ABS in [HK04]). Für eine Prozessüberwachung muss somit für jedes Werkstück aufgrund der eigenen Geometrie sowie der Werkstoff- und Werkstück-

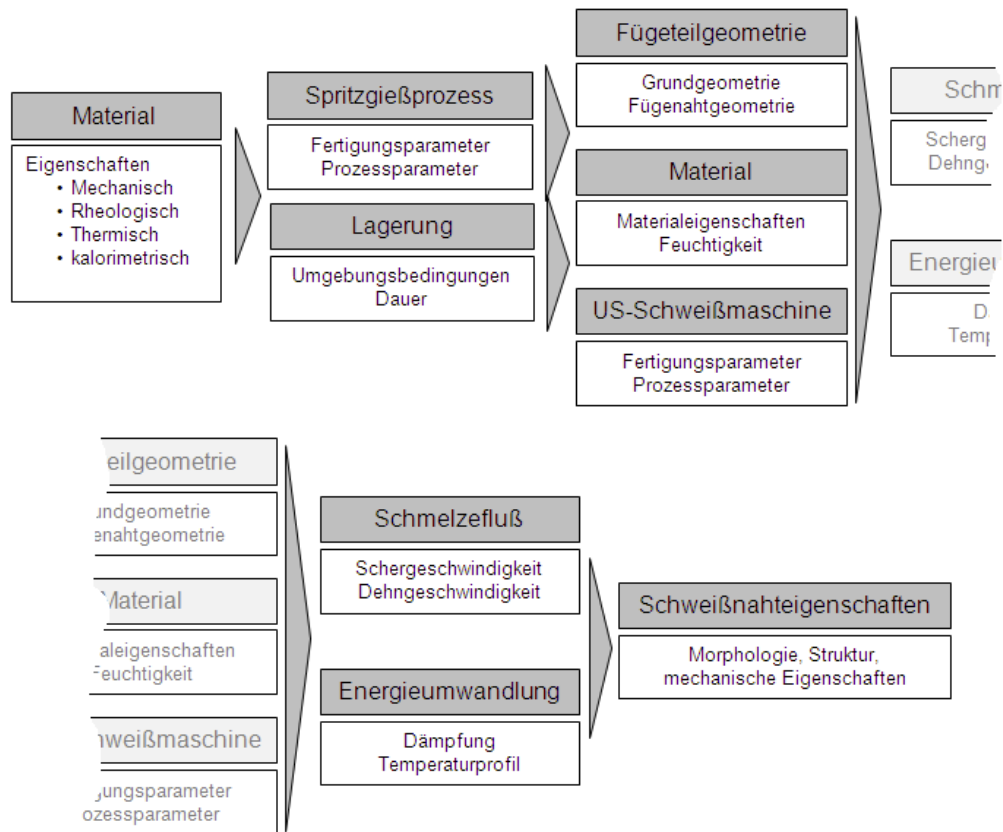


Abbildung 2.5: Einflussparameter beim Ultraschallschweißen

eigenschaften ein eigenes Prozessmodell erstellt oder angepasst werden. Hierfür eignen sich insbesondere statistische Prozessmodelle (siehe Abschnitte 2.2.2 und 2.4). Wichtig für eine hohe Modellgüte ist dabei auch die geeignete Auswahl von signifikanten Kenngrößen, die den zugrundeliegenden Prozess hinreichend genau abbilden. Bei WORTBERG [Wor96] sind Zitate zur prozessspezifischen Kenngrößendefinition beim Kunststoffspritzgießen aufgeführt:

- „Nach den heute vorliegenden Erfahrungen läßt sich nicht mehr ausschließen, daß die Auswahl der zu überwachenden Prozessparameter vom Werkzeug und vom Formteil abhängig sind“ [Wil90] .
- „Welche Parameter von Bedeutung sind, hängt in starkem Maße von den Qualitätsanforderungen und den verwendeten Komponenten ab“ [Fis90].

Aufgrund der beschriebenen, komplexen Zusammenhänge wird Entsprechendes auch für das Ultraschallschweißen erwartet. Auch die sehr große Vielfalt der Bauteile sowie die Verarbeitung des Werkstoffs Kunststoff tragen ihren Teil dazu bei.

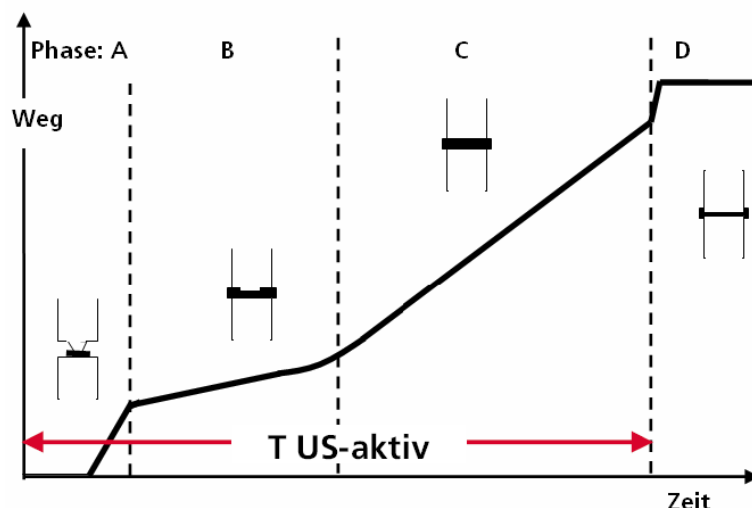


Abbildung 2.6: Prozessphasen im Schweißweg

2.1.4 Prozessphasen im Schweißweg

NETZE [Net92] kommt bei einer eingehenden Prozessanalyse des Ultraschallschweißens zu dem Ergebnis, dass der Fügewegverlauf die wichtigste Prozessgröße darstellt. Die Untersuchungen erfolgten anhand des DVS-Probekörpers (siehe Abschnitt 2.3.3) mit verschiedenen Nahtgeometrien. Bei einem kraftgesteuerten Schweißprozess ist der Fügewegverlauf eine resultierende Größe und lässt sich – vorausgesetzt, die Schweißung wird nicht vorzeitig abgebrochen – in vier Phasen unterteilen. Diese Phasen sind in Abbildung 2.6 zu sehen, wobei der Verlauf des Fügewegsignals über der Zeit aufgetragen und die prinzipielle Schweißnahtausprägung während der Phasen bei einer Schweißnaht mit ERG dargestellt sind. Die erste Phase wird durch das beginnende Abschmelzen des ERG definiert. Dieser weist aufgrund des dreieckigen Querschnitts an der Spitze wenig Material auf, was zu einer relativ hohen Fügegeschwindigkeit führt. Die Fügegeschwindigkeit lässt sich einfach anhand der ersten zeitlichen Ableitung des Schweißwegs gewinnen. In der zweiten Phase findet die Ankopplung zwischen Ober- und Unterteil statt. Dabei sinkt die Fügegeschwindigkeit wieder ab und ist während der gesamten zweiten Phase relativ gering. Nach erfolgter kompletter Ankopplung steigt die Fügegeschwindigkeit wieder an. Der Prozess geht in die dritte Phase über, in der ein stationärer Abschmelzprozess stattfindet. Aufgrund der weiterhin durch die Sonotrode eingebrachten Energie wird immer mehr Material aufgeschmolzen. Da sich der Querschnitt der Fügefläche, über den Schwingungsenergie in die Naht eingeleitet wird, nach der kompletten Ankopplung nicht mehr vergrößert, stellt sich eine konstante Fügegeschwindigkeit und somit eine lineare Zunahme des Schweißwegs ein. Neben der Werkstückgeometrie hängen der Weg und die Dauer bis zum Erreichen der dritten Phase auch von Amplitude und Kraft ab. Mit steigender Energieeinleitung und Abschmelzgeschwindigkeit werden erst bei größeren Wegen die notwendigen stationären thermischen

Zustände erreicht, die für diese Phase kennzeichnend sind. Nachdem der Generator abgeschaltet wurde, folgt die Haltephase als vierte Prozessphase. Dabei führt der Wegfall der dynamischen Belastung zu Veränderungen der mechanischen Eigenschaften des Fügeverbindes sowie der Ankopplungsbedingungen der Sonotrode [MNK92]. Der Wegfall der Schwingungen führt dazu, dass die Schmelze ihre elastischen Schmelzeschichteigenschaften verliert und sich weitgehend viskos verhält. Dadurch führt die Sonotrode eine Nachsetzbewegung in Richtung Werkstück aus [Kor96], und das vorhandene Schmelzevolumen, dessen Größe u. a. von den Werten der Amplitude und der Schweißkraft abhängig ist, wird dabei ausgedrückt. Insgesamt führt dies am Anfang der Haltephase zu einem Wegsprung. Dieser Effekt ist auch vom Vibrationsschweißen als so genannter Weißenberg-Effekt bekannt [Kai92, PK88], dessen Größe durch richtungsabhängige Normalspannungsdifferenzen in der Schmelzeschicht beeinflusst wird [TG06]. Im Folgenden ist während der Haltephase noch ein geringes Ansteigen des Schweißwegs zu erkennen. Dies hat seine Ursachen sowohl im letzten Ausfließen der Schmelze aufgrund des nunmehr statisch einwirkenden Schweißdrucks als auch in der Schwindung des Kunststoffes bei der Abkühlung. Die Schwindung ist dabei stark vom Material abhängig, in der Regel bei Kunststoffen jedoch $> 1\%$. Die in der Haltephase auftretenden rheologischen Vorgänge sind wiederum instationär.

Insgesamt ist der Fügewegverlauf als die qualitätsbestimmende Prozessgröße zu sehen [MK93, MNK92, Net92]. Für das Erreichen bestimmter Prozessphasen wird eine Korrelation mit bestimmten Schweißnahtfestigkeitsniveaus nachgewiesen. Nach MICHAELI und KORTE [MK93] sind hohe Festigkeitswerte zu erwarten, sobald die dritte Prozessphase erreicht wurde. Dementsprechend geht ein konstantes Geschwindigkeitsniveau während des Schweißvorgangs mit guten bis sehr guten Schweißnahtfestigkeiten einher [HK04]. Eine Kraftumschaltung während der Schweißung macht nur innerhalb der stationären Schweißphase Sinn, da deren Prozessbedingungen die Schweißnahtfestigkeit bestimmen [MK93]. Die starke Materialabhängigkeit des Erwärmungs- und Aufschmelzprozesses führt bei den kurzen Prozesszeiten und relativ geringen Schweißnahtdicken dazu, dass der Schweißprozess auf unterschiedliche innere Strukturen sehr sensibel reagiert [Str95].

2.2 Methoden zur Prozessüberwachung

Jeder Fertigungsprozess hat die Herstellung von Produkten zum Ziel, wobei die Hauptkriterien die Herstellkosten, die Zykluszeit und die Qualität der Produkte sind. Je nach Einsatzgebiet und Anwendungsbereich sind die Anforderungen an die Qualität sehr unterschiedlich. Die Gewährleistung der Qualität ist die Aufgabe der Qualitätssicherung und des übergeordneten Qualitätsmanagements, das weit mehr Aspekte umfasst. Aufbau und Anforderungen an ein Qualitätswesen sind in Normen DIN EN ISO 9000ff festgesetzt. Der weitreichende Ansatz, die Teilequalität nicht losgelöst, sondern im Kontext mit der Prozess-

qualität aller beteiligten Prozesse sowie unter Einbezug der Mitarbeiter und aller Phasen des Produktlebenszyklus zu sehen, führte zum Konzept des **Total Quality Management (TQM)** [CH07]. Die erste wesentliche Methode des TQM, angesiedelt in der Entwicklungsphase eines Produkts, ist das **Quality Function Deployment (QFD)**. Dabei handelt es sich um eine Methode zur systematischen Planung der Qualität eines Zielproduktes – ausgehend von kunden- und marktseitigen Qualitätsanforderungen. Jedoch existiert bis heute keine umfassende und einheitliche Definition dieser Methode, sondern es gibt nur unterschiedliche Varianten und Entwicklungstendenzen [Pfe96]. Gegenwärtig orientiert sich die Anwendungspraxis meist an der durch das Institut der amerikanischen Zulieferindustrie (American Supplier Institute, ASI) formalisierten vierstufigen Vorgehensweise. Nach der Produktplanung als erste Stufe folgen die Prozess- und Produktionsplanung. Dabei werden die Produktionsprozesse inklusive Parametern ausgewählt und daraus Qualitätssicherungsmaßnahmen abgeleitet. Eine weitere wichtige Methode des TQM ist die **FMEA (Fehlermöglichkeit- und Einflussanalyse, Failure Mode and Effects Analysis)**, die sich in eine Konstruktions- und Prozess-FMEA aufteilen lässt. Durch konsequenten Einsatz dieser Methoden wird versucht, mögliche Fehler im Herstellungsprozess bereits konstruktiv zu vermeiden und die Prozesssicherheit bei reduziertem Prüfaufwand zu erhöhen. Ein wichtiges Instrument während der Herstellungsphase eines Produkts ist der Einsatz der **Qualitätsregelkartentechnik**. Damit kann bei der Prozessanalyse und bei der Prozesslenkung in der Serienfertigung das Prozessverhalten bezüglich seiner Lage und Streuung visualisiert werden. Regelkarten werden auch bei der **Statistischen Prozessüberwachung und -regelung (Statistical Process Control, SPC)** verwendet (siehe Abschnitt 2.2.1). Eine wichtige Kenngröße ist dabei die Streuung Sigma σ eines Prozesses. Ist eine Größe, z. B. ein Qualitätsmerkmal, normalverteilt, so kann die Verteilung durch Angabe des Mittelwerts μ und der Standardabweichung σ vollständig beschrieben werden. In das 3σ -Intervall $\mu - 3\sigma \leq x \leq \mu + 3\sigma$ fallen dabei 99,73 % der Werte [HEK09, Kle03]. Auch wenn der Begriff **Six Sigma**, entsprechend 3,4 von 1 Million Ereignissen außerhalb des 6σ -Intervalls, auf die statistische Prozessregelung zurückgeht, wurde **Six Sigma** zu einer eigenen Qualitätsmanagementphilosophie entwickelt, die sich am gesamten Lebenszyklus eines Produktes orientiert [HS00, PNC00].

Unabhängig von der Art und Ausbaustufe des eingesetzten Qualitätsmanagements eines Betriebs ist die Qualitätssicherung eine Kernaufgabe während der Herstellung eines Produktes [Mas80, Pfe96]. Zu Beginn einer Fertigung muss ein stabiler Arbeitspunkt gefunden werden, welcher Teile mit hinreichender Qualität liefert. Der Nachweis der Stabilität eines Prozesses erfolgt über Regelkarten [DS05, Pfe90]. Die Stabilität ist notwendige Voraussetzung für die Beherrschbarkeit eines Prozesses und kann mittels der Fähigkeitsindizes c_p und c_{pk} beschrieben werden. Während der Index c_p die Breite der Prozessstreuung zur zulässigen

Toleranzbreite, d. h. den vorgegebenen Toleranzwerten, ins Verhältnis setzt, berücksichtigt der c_{pk} -Index zusätzlich die Lage der Verteilung [Pfe90].

Neben der Qualität können auch weitere Merkmale wie die Robustheit oder Ressourceneffizienz zur Charakterisierung eines Arbeitspunkts verwendet werden. Robustheit bedeutet dabei, dass sich auftretende Prozessschwankungen nicht oder nur geringfügig auf die Produktqualität auswirken. Die **statistische Versuchsplanung** (*SVP* oder *Design of Experiments, DoE*) bietet verschiedene Verfahren, einen Arbeitspunkt zu finden oder zu optimieren sowie relevante Prozessparameter und Steuergrößen zu identifizieren [Kle03, May97, Sch97]. Hauptpunkte der SVP sind dabei systematisch geplante Versuche und deren Auswertung. Ihren Ursprung hat die Methodik in den 1920er Jahren durch die Arbeiten von Sir Ronald FISHER [Fis35a, Fis35b]. Um den Versuchsaufwand, der exponentiell mit der Anzahl der Parameter steigt, reduzieren zu können, werden vielfach auch teilfaktorielle Versuchspläne eingesetzt, die je nach Ausprägung auch mehrfache Wechselwirkungen aufzeigen können [Mon09]. Am bekanntesten dürften dabei die Verfahren von TAGUCHI sein [Tag87]. Ein wichtiger Aspekt bei ihm ist die Signal-Rausch-Analyse, mit welcher die Robustheit einer Prozesseinstellung beurteilt werden kann. Einige einfache Verfahren, die wesentlichen Einflussgrößen zu identifizieren, gehen auf D. SHAININ zurück [SS88a]. Eine Gegenüberstellung der verschiedenen Methoden zur statistischen Versuchsplanung ist bei KLEPPMANN zu finden [Kle92].

Eine Qualitätssicherung muss auftretende Störungen im Prozessablauf schnell erkennen und bewerten können. Nur so kann die Ausschussrate möglichst gering gehalten werden. Einen wesentlichen Beitrag hierzu liefert der Einsatz von Überwachungsmethoden direkt am Prozess, welche sich sowohl der Mess- und Prüftechnik als auch Methoden der Statistik bedienen. Für die Prozessüberwachung von Fertigungsprozessen existiert heutzutage eine Vielzahl von Methoden und Verfahren. Diese können einerseits in statistische und kontinuierliche Verfahren, andererseits in Prozess- und Qualitätsüberwachungen aufgeteilt werden, abhängig davon, ob Prozessgrößen oder direkte Qualitätsausprägungen der Produkte überwacht werden. Damit eng verknüpft ist auch die Unterscheidung von indirekter zu direkter Qualitätsüberwachung. Die einzelnen Methoden werden im Folgenden näher erläutert.

2.2.1 Statistische Prozessüberwachung

Für eine Überwachung von Fertigungsprozessen wird heutzutage häufig die **Statistische Prozessüberwachung** oder **-regelung** (*Statistical Process Control, SPC*) eingesetzt, welche als wichtiges Element *Qualitätsregelkarten (QRK)* einsetzt. Diese dienen dazu, die Qualität der Übereinstimmung zwischen dem Istzustand der Fertigung und dem Sollzustand zu überwachen [Mas80]. Die wissenschaftlichen Grundlagen wurden von SHEWHART hergeleitet und beschrieben [She31]. Umfangreiche statistische Untersuchungen und theoretische Ableitungen führten schließlich zur Entwicklung der Regelkarten, um die gewon-

nenen Erkenntnisse in die praktische Anwendung zu überführen. Die Basis für die SPC bilden Stichproben, die in regelmäßigen Abständen der Produktion entnommen werden. Anschließend wird die Qualität jeder Stichproben durch eine Prüfung bestimmt. Die Stichprobe wird durch einige statistische Größen charakterisiert, z. B. Mittelwert, Streuung oder Spannweite, und diese werden in ein Diagramm eingetragen. Die Werte hängen vom eingesetzten Regelkartentyp ab, wobei die am häufigsten verwendeten Typen die Mittelwert- und Standardabweichungskarte sind [Pfe96, RM95]. Anhand des zeitlichen Verlaufs der aufgetragenen Stichprobengrößen können mit den darin enthaltenen Warn- und Eingriffsgrenzen Tendenzen im Prozess frühzeitig erkannt und Ausschussteile durch rechtzeitige Gegenmaßnahmen vermieden werden. Wenn das Qualitätsmerkmal erst Stunden nach der Fertigstellung ermittelt werden kann, stößt die klassische SPC schnell an ihre Grenzen [WH91]. Dies ist zum Beispiel bei Festigkeitswerten oder bei Geometriemaßen von Produkten aus teilkristallinen Kunststoffen der Fall, die aufgrund der Nachkristallisation noch über mehrere Stunden schwinden, bevor sie ihre endgültigen Abmessungen erreichen. Auch bei kurzfristigen Prozessabweichungen zwischen zwei Stichproben werden diese durch die SPC nicht erkannt [HW95].

Dieselben Daten, die in der QRK zur Überwachung der laufenden Produktion eingetragen werden, dienen auch als Basis für Prozessfähigkeitsuntersuchungen [DS05]. Solche QRK finden heute vielfach Anwendung, um industrielle Fertigungsprozesse zu überwachen. Die Robustheit einer Produktion kann über die auf Stichproben basierenden Verfahren der statistischen Prozesskontrolle nachgewiesen werden (siehe z. B. [Mes93, Rab93, Seu92]). Mit gewissen Modifikationen findet die „klassische“ SPC auch Anwendung bei chemischen Prozessen und Produkten, z. B. in der Kunststoffherstellung [Zab95].

SPC für Prozessparameter

Motiviert durch die überwiegend positiven Erfahrungen beim Einsatz der klassischen SPC, die zur Prozessüberwachung anhand von Messwerten der Qualität von Produkten entwickelt wurde, konnte in den letzten Jahren verstärkt der Trend beobachtet werden, die Verfahren auch ohne Modifikationen auf Prozessgrößen der Fertigung wie Drücke, Zeiten, Maxima etc. zu übertragen [PNH92]. Verstärkt wurde dies durch die immer größere Verbreitung von Rechnersystemen und digitalen Steuerungen, die ohne großen Mehraufwand die nötigen Parameter zur Verfügung stellen konnten. Aufgrund der nachfolgend aufgelisteten Punkte muss dieses Vorgehen jedoch zumindest als fragwürdig bezeichnet werden [MBVW94, PWHH93, Wor96, WH91]:

- Die SPC beruht auf Stichproben mit $n \geq 2$ und erlaubt es aufgrund der gebildeten Kenngrößen wie Mittelwert und Standardabweichung nicht, jedes produzierte Teil einzeln zu betrachten. Einzelne Ausreißer können im Mittelwert verschwinden und werden nicht erkannt.

- Aufgrund des Stichprobencharakters der SPC werden nicht alle Informationen genutzt, die durch eine kontinuierliche Datenerfassung gesammelt werden. Ein großer Teil der erfassten Daten bleibt somit unberücksichtigt.
- Die Eingriffsgrenzen von Prozessparametern, die in Regelkarten eingetragen werden, stehen in keinem direkten Zusammenhang mit Qualitätsmerkmalen des Produktes, sondern werden einzig aus dem Streuverhalten der Größe im bisherigen Prozessverlauf abgeleitet.
- Ein spezifiziertes Qualitätsmerkmal hängt meist nicht nur von einem, sondern von mehreren Prozesskenngrößen ab. Feste Grenzen für die Einzelparameter können damit nicht mehr angegeben werden, da sie aufgrund von Wechselwirkungen von den Werten der anderen relevanten Kenngrößen abhängen. Somit sind Einzelparameterüberwachungen unzureichend.

Dieser letzte Punkt bedeutet für Parameter, die nur einer geringen Streuung unterliegen und folglich mit engen Toleranzgrenzen überwacht werden, dass bei einem Verlassen dieser Grenzen Einflüsse auf die Produktqualität und systematische Einflüsse diagnostiziert werden, obwohl diese unter Umständen gar nicht vorliegen und die Produktqualität völlig unbeeinflusst ist. Werden umgekehrt Prozessparameter, die einen signifikanten Zusammenhang zur Produktqualität aufweisen, nicht überwacht, besteht die Gefahr, dass unerlaubte Prozessabweichungen von der SPC nicht erkannt werden – zumindest so lange sie sich nicht auch in anderen Prozessgrößen abbilden.

2.2.2 Kontinuierliche Prozessüberwachung

Im Gegensatz zur SPC, bei welcher die Überwachung rein anhand von Stichproben und anschließender Datenaggregation durchgeführt wird, wird bei der **kontinuierlichen Prozessüberwachung** (*Continuous Process Control, CPC*) jeder Fertigungszyklus betrachtet und überwacht. Eine CPC stellt somit eine 100%-Online-Prozesskontrolle dar. Außerdem ermöglicht die CPC eine Trendüberwachung [Han94, Wor96]. Die CPC führt zu einer effektiven Reduzierung der stichprobenartigen Endproduktprüfung, kann diese jedoch nicht vollständig ersetzen [PNH92]. Bei der Überwachung innerhalb einer mehrstufigen Prozesskette kann zudem bei frühzeitiger Ausschusserkennung das Teilfertigprodukt angesteuert werden, womit der wirtschaftliche Ausfall des fertigen Endprodukts vermieden wird. Auch können Fehler in den einzelnen Prozessschritten identifiziert und an Ort und Stelle behoben werden. Gegenüber der SPC haben die auf Prozessdaten basierenden Methoden den Vorteil, dass jeder Fertigungszyklus separat betrachtet wird und einzelne Ausreißer nicht durch Mittelung einer Stichprobe quasi verloren gehen können. Die Prozessüberwachung im Rahmen der CPC kann dabei mit verschiedenen Daten erfolgen: Anhand geprüfter Qualitätsdaten

oder auf der Basis messtechnisch erfasster Prozessdaten. Die wesentlichen Punkte werden im Folgenden kurz dargestellt.

... auf Basis der Qualitätsprüfung

Eine Möglichkeit der CPC besteht in einer 100 %-Prüfung aller gefertigten Teile. Ist die Prüfung als Endprüfung ausgeführt, werden zwar alle Ausschussteile entdeckt – aber verbunden mit hohem Aufwand und Kosten. Eine Prüfung am Endprodukt kann je nach Organisation der Fertigung einen größeren zeitlichen Abstand zum Produktionsprozess aufweisen, wodurch eine zeitnahe Reaktion bei auftretenden Fertigungsproblemen erschwert oder unmöglich wird. Die Prüfung kann dabei entweder durch Prüfpersonal durchgeführt werden, was in Deutschland unter anderem aus wirtschaftlichen Gründen immer weniger angewendet wird (Dreischichtbetrieb, Nachtzulage, ...), oder durch entsprechende Prüfautomaten oder prozessintegrierte Prüftechniken. Zur letzteren Gruppe zählt beispielsweise der in den letzten Jahren stark gewachsene Bereich der industriellen Bildverarbeitung. Voraussetzung für den Einsatz solcher Prüfverfahren ist, dass die relevanten Qualitätsmerkmale messtechnisch erfassbar sind. Beim Einsatz der Bildverarbeitung sind die Sichtbarkeit, Zugänglichkeit und Beleuchtung wichtig. Merkmale wie Dichtigkeit sind über die Bildverarbeitung nicht direkt erkennbar und bedürfen entsprechend teurer spezieller Prüfeinrichtungen. Dagegen können Qualitätsmerkmale, wie z. B. die Festigkeit, nur zerstörend geprüft werden. Sie entziehen sich so einer kontinuierlichen Prozessüberwachung durch Prüfverfahren vollständig und können nur stichprobenhaft erfasst werden.

... auf Basis von Prozessdaten

Als Grundlage dienen Daten, die direkt aus dem Fertigungsprozess gewonnen werden. Zum einen können direkt Werte aus der Maschinensteuerung des Prozesses (z. B. Zeiten, Maximalwerte etc.) oder erfasste Sensorsignale verwendet werden, die verschieden weiterverarbeitet oder aggregiert werden können. Für die Möglichkeit der Hüllkurven- oder Grenzwertüberwachung gilt ebenso wie für die SPC für Prozessparameter (siehe Abschnitt 2.2.1), dass sie meist keinen unmittelbaren Zusammenhang mit Qualitätsmerkmalen des Produktes aufweisen. Sie zählen zu den indirekten Verfahren, da anhand von Signal- oder Kenngrößen- ausprägungen nur indirekt auf die Qualität des Teils geschlossen wird. Durch den fehlenden direkten Zusammenhang ist das Setzen von geeigneten Grenzwerten und Toleranzen meist nur erfahrungsbasiert anhand von Versuchen möglich und mit entsprechender Unsicherheit behaftet. Auch die darin implizit enthaltene Annahme, dass alle Teile innerhalb der Toleranzen iO-Teile sind – also Teile, die innerhalb der Spezifikationen liegen – und bei Fertigungszyklen, die diese Grenzen verletzen, unzulässige Qualitätsmerkmale (niO-Teile) vorherrschen, ist in den allermeisten Fällen zumindest mit Vorsicht zu genießen. Dies grün-

det auch auf der Tatsache, dass mit diesen Verfahren Wechselwirkungen und kombinierte Einflüsse mehrerer Größen in der Regel nicht erfasst werden können. Im Gegensatz dazu bilden modell- oder wissensbasierte Systeme den Zusammenhang zwischen erfassten Prozessgrößen und den Zielgrößen, meist der Qualität, direkt ab. Diese Methoden können auch Wechselwirkungen berücksichtigen und zählen zu den direkten Überwachungsverfahren.

Hüllkurvenüberwachung: Maßgebliches Kriterium bei der Hüllkurvenüberwachung ist, ob sich die einzelnen erfassten Signale des Prozesses oder einzelne Teilbereiche innerhalb zuvor definierter Grenzverläufe befinden. Zur Definition der Grenzkurven wird dabei zu meist ein einzelner Zyklus als so genannter Referenzzyklus direkt verwendet oder dieser aus mehreren Fertigungszyklen als Mittelwertkurve berechnet. Dieser Referenzverlauf wird anschließend mit einem gewissen Toleranzband überlagert. Definitionsgemäß sind Fertigungszyklen, deren Signalverläufe innerhalb des Bandes liegen, gute Teile und solche ausserhalb schlechte Teile. Meist werden diese Toleranzen durch Angabe einer maximal zulässigen, relativen Abweichung in Prozent vorgegeben oder aus der Streuung der Signale abgeleitet. Die Prozessüberwachung anhand von Hüllkurven kommt u. a. bei verschiedensten Verfahren zur Metallbearbeitung zum Einsatz, wie z. B. beim Gesenkschmieden, Stanzen, Feinschneiden oder Clinchen [N.N98, Ros99, Ter94].

Die NEPRES⁴-Technologie (siehe Abschnitt 2.6) verfügt über adaptive Hüllkurven, die sich aus der Streubreite der Signale an jedem einzelnen Datenpunkt ableiten. Beim Einsatz von leistungsfähigen Methoden zur Prozessmodellierung, z. B. mittels Neuronaler Netze (siehe Abschnitt 2.4), können die Hüllkurven zur Absicherung der Modellgültigkeit verwendet werden. Auf diese Weise wird sichergestellt, dass sich alle aufgezeichneten Signale innerhalb des Dynamikbereichs befinden, der auch bei den Trainingsdaten vorgeherrscht hat. Auch der Ausfall eines Sensors kann so sicher erkannt werden.

Grenzwertüberwachung: Die einfachste Stufe ist eine Überprüfung auf die Einhaltung von Grenzwerten einzelner diskreter Größen nach unten oder oben, beispielsweise minimal und maximal zulässige Schweißzeit beim Ultraschallschweißen. Eine Grenzwertüberwachung bestimmter Parameter ist heutzutage für verschiedene Signale und Messgrößen in vielen Steuerungen vorhanden. Oft wird jedoch diese Prozessüberwachung durch zusätzliche Sensorsignale und eine Auswertung mittels externer Systeme erweitert, wie beispielsweise beim Spritzgießen. Je mehr Parameter eines Fertigungsprozesses überwacht werden sollen, desto größer wird der dazu nötige Aufwand. Auch das Setzen der „korrekten“ Grenzwerte ist schwierig und oft kompromissbehaftet. Bei zu engen Toleranzgrenzen werden viele Teile als Ausschuss deklariert, obwohl sie sich bei einer Prüfung als iO-Teile erweisen. Werden die Grenzen zu weit gesetzt, werden schadhafte Teile nicht erkannt und an den Kunden

⁴Neuronales Prozess-Regelungs-System

ausgeliefert. Ein weiteres Problem sind die Wechselwirkungen zwischen verschiedenen Prozessparametern, die durch Kenngrößen aus einzelnen Signalen nicht berücksichtigt werden [NS07].

Prozessmodelle: Prozessmodelle bilden einen direkten Zusammenhang zwischen Prozessgrößen und interessierenden Zielgrößen ab. Dazu werden zumeist durch Versuchsreihen entsprechende Datensätze gewonnen, anhand derer die Modelle anschließend trainiert, also an den Prozess adaptiert werden. Beim Einsatz von Prozessmodellen werden zwei Phasen unterschieden: Die Adaption der Modelle inklusive deren Verifikation sowie die Anwendung. Während der Anwendung werden aus dem Fertigungsprozess die für das Modell notwendigen Prozessgrößen erfasst und durch das Modell verarbeitet, welches als Ergebnis eine Prognose für den Zielwert liefert. Je nach Modellgüte, die während der Verifikation zu beurteilen ist, ist diese Prognose mit einem mehr oder weniger großen Fehler behaftet. Um hochwertige Modelle zu erhalten, ist eine exakte Ermittlung der Zielgrößen – hier: der Qualitätsmerkmale – sehr wichtig [Ueb95]. Da die Bestimmung der Modellparameter durch eine Fehlerminimierung zwischen Modellausgabe und Messwerten erfolgt, kann ein Prozessmodell nie genauer sein als die zur Modellbildung eingesetzten Messmittel. Zudem ist für eine gute Modellierungsleistung oft eine geeignete und problemangepasste Vorverarbeitung der Signaldaten notwendig, bevor diese als Eingangsdaten in die Modelle eingegeben werden können.

Es gibt die verschiedensten Methoden zum Aufbau von Prozessmodellen. Je nach Art werden sie eher der Mathematik, Statistik, Informatik, der Künstlichen Intelligenz oder dem Machine Learning zugeordnet, wobei die Grenzen und Übergänge fließend sind. Ein grober Überblick über häufiger eingesetzte Verfahren – wie Regressionsmodelle, Neuronale Netze oder Support Vector Machines – ist in Abschnitt 2.4 zu finden. Prozessmodelle sind geeignet, sowohl Wechselwirkungen zwischen Parametern abzubilden als auch Merkmale durch Prognosen zu überwachen, welche nicht direkt oder gar nicht geprüft werden können. Beispiele hierfür sind die Festigkeit, die Abmessungen eines warmen Werkstücks aus einem teilkristallinen Kunststoff nach erfolgter Schwindung oder Eigenschaften von Bauteilen nach (künstlicher) Alterung. Prozessmodelle werden in den verschiedensten Bereichen eingesetzt, wie beispielsweise dem Kunststoffspritzen [FS96, Gie91, LRS01, SH00], der Halbleiterindustrie [Kum05], der Robotik [VP96] oder auch dem Ultraschallschweißen von Kunststoffen oder Metallen [Kor96, N.N03]. Dieselben Modellierungsverfahren bilden auch die Basis für den Aufbau von Regelungen [FSK98, Pei96, PL95] und werden auch bei der Mustererkennung verwendet, z. B. der Texturanalyse [Pan05], der Spracherkennung [PKM05] oder der Identifikation von Personen in biometrischen Systemen [AA07]. Einige Beispiele sind in Abschnitt 2.5 ausführlicher behandelt.

2.3 Prozessüberwachung beim Ultraschallschweißen

Bevor eine Prozessüberwachung aufgebaut werden kann, müssen die Anforderungen und Ziele sowie die Randbedingungen ermittelt und festgelegt werden. Für das Ultraschallschweißen sind in den folgenden Abschnitten die wichtigsten Punkte und der Stand der Technik zusammengefasst.

2.3.1 Qualitätsmerkmale bei Schweißteilen

Ein wichtiger Schritt ist die Festlegung der relevanten Qualitätsmerkmale. Diese geschieht während der Planungs- und Konstruktionsphase. Wichtig ist auch, die Möglichkeiten zur Prüfung der einzelnen Qualitätsanforderungen zu berücksichtigen. Merkmale, die sich nicht oder nur sehr schwer quantifizieren lassen, sind dabei ungeeignet. Die Qualität von Produkten, die mittels Ultraschall geschweißt werden, wird meist durch eines oder mehrere der Qualitätsmerkmale Dichtigkeit, Schweißnahtfestigkeit oder optische Eigenschaften der Schweißnaht bestimmt. Auf die einzelnen Merkmale mit ihren wichtigsten Eigenschaften sowie deren Prüfung wird im Folgenden kurz eingegangen.

Dichtigkeit

Ein relativ häufig gefordertes Qualitätsmerkmal ist die Dichtigkeit der Schweißnaht gegenüber Gas oder anderen Medien. Diese Forderung besteht z. B. beim Aufschweißen eines Deckels auf ein Gehäuse. Eine Undichtigkeit tritt meist sehr lokal begrenzt auf. Ursachen können zum einen Verschmutzungen der Schweißfläche sein (z. B. durch Öltropfen oder Fingerabdrücke), was eine durchgehend dichte Verbindung von Ober- und Unterteil auf molekularer Ebene verhindert. Eine andere Ursache können kleine Risse in den Schweißteilen sein, die ihren Grund beispielsweise im Spritzgießen haben. Auch können Eigenspannungen, die sich während des Abkühlens nach der Schweißung bilden können, zu lokalen Spannungsspitzen und damit Mikrorissen und somit zu Undichtigkeiten führen. Über entsprechende konstruktive Maßnahmen bei der Schweißnahtgestaltung sollten mögliche Probleme mit der Dichtigkeit minimiert werden. Wenn Undichtigkeiten nur lokal auftreten, muss deren Detektion mittels einer signalbasierten Prozessüberwachung über integrale Messgrößen, wie sie beim Ultraschallschweißen erfasst werden können, als wenig erfolversprechend angesehen werden.

Für die manuelle Prüfung der Dichtigkeit bietet sich eine Blasenprüfung unter Wasser an. Während der Fertigung kann sie über entsprechende Prüfstationen ermittelt werden. Voraussetzung dafür ist, dass eventuell vorhandene Öffnungen in der Prüfvorrichtung abgedichtet werden können. Die meisten Prüfverfahren beruhen auf einer Differenzdruckmessung, nachdem das zu prüfende Bauteil angeschlossen, abgedichtet und entweder mit Über-

oder Unterdruck beaufschlagt wurde – je nach späterem Einsatzgebiet. Wird ein entsprechender Druckabfall festgestellt, bedeutet dies jedoch nicht zwangsläufig eine Undichtigkeit der Schweißnaht, sondern kann auch auf eine Leckage am Anschluss der Prüfanlage oder einer Abdichtung zurückzuführen sein.

Geometriemaße

Bei manchen Bauteilen spielen die Abmessungen eine große Rolle. Die meisten Geometriemaße werden dabei bereits durch die Herstellprozesse der Fügepartner festgelegt. Beim Schweißen kann zum einen ein seitlicher Versatz auftreten, dem durch konstruktive Maßnahmen wie Führungen am Teil oder durch die Werkstückaufnahme begegnet werden muss. Zum anderen erfolgt die Einhaltung eines Maßes in Schweißrichtung entweder durch konstruktive Maßnahmen wie Anschläge oder durch Schweißen auf absoluten Weg mittels des entsprechenden Schweißmodus. Die Schweißmodi nach Zeit, Energie oder relativem Weg sind hierbei eher ungeeignet. Die Überwachung im Schweißprozess erfolgt nach einmaliger Justage anhand des Wegsignals der Schweißpresse. Die Prüfung des Maßes kann entweder über Lehren, Messschieber oder über entsprechende Prüfeinrichtungen wie taktile oder optische Koordinatenmessmaschinen erfolgen.

Schweißaustrieb und optische Anforderungen

Bei Schweißnähten im Sichtbereich kann es die Forderung nach einem geringen oder keinem Schweißaustrieb geben. Der Schweißaustrieb entsteht durch Ausfließen der Schmelze während des Schweißvorgangs und hängt maßgeblich von der Konstruktion, dem Material sowie den eingestellten Schweißparametern ab. Bei manchen Bauteilen existiert auch die Anforderung, dass die Sonotrode keine Abdrücke auf dem Bauteil hinterlässt. Dies ist maßgeblich von der Konstruktion und Fertigung von Sonotrode und Werkstück sowie den von Material und Schweißparametern bedingten Ankopplungsbedingungen und der Ausrichtung von Werkzeug, Werkstück und Sonotrode abhängig. Durch konstruktive Maßnahmen sowie eine Optimierung des Arbeitspunktes (z. B. Reduzierung der Amplitude) können Abdrücke vermieden werden. In kritischen Fällen kann auch das Einlegen einer Folie zwischen Sonotrode und Werkstück während der Schweißung Abhilfe schaffen. Die Prüfung des Austriebs im Innenbereich kann – sofern entsprechende Öffnungen vorhanden sind – durch Endoskope erfolgen, im Aussenbereich über eine Sichtprüfung. Eine automatische Prüfung ist nur im Aussenbereich möglich, wobei sich diese je nach Schweißnahtgeometrie aufgrund von Sichtbarkeit und Teilehandling oder wegen Beleuchtung und Reflektionen als äusserst schwierig und meist unpraktikabel darstellt. Für die Prüfung von Abdrücken gilt bezüglich Beleuchtung und Reflektion Entsprechendes.

Die Gleichmäßigkeit der Schweißung und das Auftreten von Austrieb kann bei entsprechender Sichtbarkeit der Naht während der Einrichtphase auch mittels Thermografie sichtbar gemacht und zur Prozessbeurteilung herangezogen werden. Für eine Serienprüfung ist dieses Verfahren in der Regel zu teuer und aufwändig. Grundlegende Veränderungen des Austriebs aufgrund von materialbedingten Viskositätsschwankungen, z. B. aufgrund von Feuchtigkeit, dürften sich auf das Schweißverhalten in Form der Leistungsaufnahme und des Schweißwegverlaufs auswirken und sollten über Prozessmodelle erkennbar sein.

Schweißnahtfestigkeit

Ein häufig geforderes Qualitätskriterium ist die Mindestfestigkeit einer Schweißnaht. Die Ermittlung der Festigkeit erfolgt mittels werkstoffmechanischer Prüfungen. Mechanisch tritt ein Versagen ein, sobald aufgrund einer einwirkenden Kraft eine werkstoffabhängige, maximale Spannung überschritten wird. Dabei gilt $\sigma_{Bruch} = F_N/A_{proj}$ bzw. $F_N = \sigma_{Bruch} * A_{proj}$ mit F_N als Kraft senkrecht zur Fläche und A_{proj} der Querschnittsfläche der Schweißnaht. Festigkeitsuntersuchungen erfolgen dabei anhand genormter Prüfgeometrien und Prüfbedingungen (z. B. DIN EN ISO 527-1 für Zugversuche [CH07]). Bei der Zugprüfung ist auf einen einachsigen Spannungszustand zu achten. Beim Auftreten von mehrachsigen Spannungszuständen sowie Kerben können die Messwerte gegenüber der realen Festigkeit signifikant beeinflusst werden. Bei Festigkeitsprüfungen ist stets mit einer gewissen Unsicherheit und Streubreite der Messwerte zu rechnen. Dies zeigte sich z. B. auch bei der Durchführung von Eignungsprüfungen von Prüflaboratorien, wobei bei der Zugprüfung für eine Festigkeit von etwa 320 MPa Standardabweichungen von ca. 32 MPa, d. h. $\approx 10\%$, festgestellt wurden [Vog03]. Bei ZIEGLTRUM ergeben sich für die Ergebnisse der Zugprüfung teils Standardabweichungen von ca. 20 % der mittleren Festigkeit, was den Erfahrungen aus früheren Versuchen und der Praxis entspricht [Zie00]. Aufgrund ihres zerstörenden Charakters kann eine derartige Prüfung nicht wiederholt werden. Diese Prüfung muss deshalb stets sehr sorgfältig erfolgen. Bei der Zugprüfung von Kunststoffen sollte die Prüfgeschwindigkeit entsprechend gering sein (ca. 10 bis 20 mm/min [LE85]). Bei der Auswertung sollte, falls möglich, mit Mittelwerten mehrerer gleichwertiger Proben anstatt mit Einzelwerten gearbeitet und auf konstante Umgebungs- und Lagerbedingungen geachtet werden. Die teilweise großen Standardabweichungen insbesondere bei der Prüfung von Kunststoffteilen resultieren teils aus sehr lokal initiierten Anrissen der Schweißverbindung und dem darauffolgenden Versagen der kompletten Naht. Exemplarisch hierfür weist die Reißspannung nach KORTE bei ultraschallgeschweißten Teilen aus teilkristallinem Polypropylen relativ große Schwankungen auf, obwohl ähnliche Dehngeschwindigkeiten berechnet wurden. Die Ursache liegt laut KORTE darin, dass ein Bruch der Schweißnaht durch einen sehr lokal initiierten Riss eingeleitet wird und damit die ermittelte Festigkeit stark von der lokalen Spannungsverteilung abhängt, welche sehr unterschiedlich sein kann [Kor96]. Allgemein weisen polymere Werkstoffe ein

je nach Typ und Beanspruchung sehr unterschiedliches Bruchverhalten auf, welches unter anderem von Kristallisationsgrad und Sphärolithgrößen, der Temperatur, der Verformungsgeschwindigkeit oder den Umgebungsbedingungen abhängen kann [Mic92, Hei09].

Bei der Prüfung von geschweißten Produkten treten im Vergleich zu genormten Prüfkörpern einige weitere Punkte auf, die sich nachteilig auf die Prüfung und deren Ergebnisse auswirken können. Die erste Hürde ist oft schon die Einspannung des Bauteils in die entsprechende Prüfvorrichtung. Ist dies nicht möglich, können gewisse Abschnitte aus dem Bauteil ausgeschnitten und geprüft werden. Dabei ist jedoch mit Ungenauigkeiten zu rechnen. Zum einen weisen Proben von unterschiedlichen Stellen desselben Werkstücks teils deutlich andere Eigenschaften auf (z. B. in [Ueb95]), zum anderen bereitet auch die Übertragbarkeit von Werkstoffkennwerten, die an kleineren Proben genommen wurden, auf die gesamten Bauteilabmessungen und Beanspruchungen oft Schwierigkeiten [CH07]. Auch das Auftreten von mehrachsigen Spannungszuständen lässt sich nicht immer vermeiden. So treten z. B. bei der Berstdruckprüfung eines Gehäuses durch Aufbiegung mehrachsige Spannungszustände und eine Kerbwirkung an den Kanten auf, die sich beide auf den gemessenen Festigkeitswert auswirken. Ein weiterer Punkt sind die Schwankungen der Nahtquerschnittsfläche. Während bei genormten Prüfkörpern die Querschnittsfläche exakt definiert ist, ist der durch die Schweißung erzielte Nahtquerschnitt teils deutlichen Schwankungen unterworfen. Dies wiederum wirkt sich in einer Streuung der gemessenen Zugkräfte aus, welche die Festigkeit charakterisieren, da die zum Erreichen der Bruchspannung notwendige Kraft proportional der Querschnittsfläche ist. UEBBING berichtet von hohen Streuungen bei der Ermittlung der Festigkeit beim Vibrationsschweißen, was zu Problemen bei der Modellbildung führt. Eine Modellbildung für ein Längenmaß, welches messtechnisch sehr exakt erfassbar ist, ist hingegen problemlos möglich [Ueb95]. LAND schlägt für eine zerstörungsfreie Prüfung in der Fertigung vor, 50 - 70 % der (praxisgerechten) Last aufzubringen, die bei einer zerstörenden Prüfung zum Bruch oder Versagen führt [LE85].

Eingangszustand der Werkstücke

Wie in Abschnitt 2.1.3 dargestellt, hat der Zustand der Schweißpartner hinsichtlich ihrer Vorgeschichte (d. h. Herstellung, Transport und Lagerung) einen nicht unerheblichen Einfluss auf den Schweißprozess. Bei den meisten Anwendungen werden die einzelnen Schweißpartner als Schüttgut bereitgestellt. Eine Zuordnung der einzelnen Bauteile zu ihrer Vorgeschichte ist nur bedingt – d. h. wenn überhaupt, dann nur chargenweise – möglich. Erschwert wird dies durch die Tatsache, dass bei vielen Anwendungen mindestens ein Schweißpartner durch einen (externen) Zulieferer hergestellt wird. Nur bei relativ wenigen, meist sicherheitsrelevanten Produkten (z. B. in der Luftfahrt oder Medizintechnik), ist eine komplette Zuordnung der Prozessgrößen von Spritzgießen und Schweißen gegeben, also gut zu überwachen.

Relevante Größen wie Feuchtegehalt, Kristallisationsgrad, Molekülorientierungen oder Eigenspannungen können in der Praxis im Rahmen von Wareneingangskontrollen nicht erfasst werden. Die Schwankungen können nur gering gehalten werden, indem alle Vorprozesse gut beherrscht und überwacht sind. Dies funktioniert insbesondere bei Zulieferteilen nur mit organisatorischen Maßnahmen, wie z. B. detaillierten Lasten- und Pflichtenheften sowie Zertifizierungen.

2.3.2 Bekannte Ansätze und Lösungen zur Überwachung

Aus der Literatur sind einige Untersuchungen bekannt über die Zusammenhänge zwischen variierten Prozessparametern und ihren Auswirkungen auf die Qualität der Schweißteile, insbesondere der Schweißnahtfestigkeit. Aufbauend darauf werden vereinzelt Methoden zur Prozessüberwachung angewendet. Als Abschaltkriterium des Schweißprozesses dienen meist eine eingestellte Schweißzeit und manchmal die eingebrachte Schweißenergie. Im Folgenden wird ein Überblick über wesentliche Untersuchungen und Zusammenhänge gegeben.

RITTER [Rit86] untersucht an plattenförmigen Proben detailliert die Energieumwandlung und das Schwingungsverhalten beim Ultraschallschweißen bei verschiedenen Amplituden und Schweißkräften. Messtechnisch werden die vom Generator abgegebene Leistung, die resultierende Kraft unter dem Amboss sowie die Schwingung verschiedener Komponenten in verschiedenen Raumrichtungen untersucht. Ebenfalls werden verschiedene Ansätze zur Temperaturmessung in der Schweißzone – sowohl mit Thermoelementen als auch per Thermographiekamera – untersucht, die jedoch alle keine plausiblen und zufriedenstellenden Ergebnisse liefern. Eine einfache Messung der Temperatur ist unter Praxisgesichtspunkten somit nicht möglich.

Als wesentliche Mechanismen der Energieumwandlung weist RITTER die Kombination aus innerer Reibung und Grenzflächenreibung nach. Bei allen Versuchen tritt eine kritische Schweißkraft auf, die mit der maximal abgegebenen Schweißenergie und der maximalen Festigkeit – ermittelt durch Zugprüfversuche – zusammenfällt. Die vom Generator abgegebene Leistung weist also eine hohe Korrelation mit der Schweißnahtfestigkeit auf, so dass diese als wichtige Größe zur Qualitätskontrolle verwendet werden kann.

BENATAR [Ben87] ermittelt einen Einfluss der Schweißzeit auf die Festigkeit derart, dass bei zu langen Schweißzeiten die Festigkeit absinkt. Dies wird auf den großen Schmelzefluss bei langen Schweißzeiten zurückgeführt, was eine Orientierung der Molekülketten in Fließrichtung zur Folge hat [TDW83]. Derselbe Effekt wird auch von CHUAH ET AL. beobachtet [CCCL00].

Nach MICHAELI ET AL. [MNK92] ist der Schweißwegverlauf die wichtigste Prozesskenngröße beim Ultraschallschweißen und lässt sich unabhängig davon, ob der Werkstoff amorph, teilkristallin oder kurzglasfasergefüllt ist, in vier Phasen aufteilen. Die Leistungsaufnahme

des Konverters erweist sich unabhängig von der Schweißkraft als nahezu konstant, ist also für die Identifikation der verschiedenen Prozessphasen ungeeignet. Wirkleistung und Energie sind demnach nicht geeignet, qualitätsrelevante Schmelzestandsänderungen während der Schweißphasen zu detektieren. Veränderungen sind nur bei starken Schwingungszustandsvariationen sichtbar. Diese treten für höhere Fügekräfte auf, da sich bessere Ankopplungsbedingungen zwischen Sonotrode und Werkstück einstellen. Auch andere Veröffentlichungen [MEBN88, MKK97, MK94a, MK94c, MK95, MKDE94, MWK94, NM91] haben im Wesentlichen dieselben Inhalte und gehen zumeist auf die Dissertationen von NETZE [Net92] und KORTE [Kor96] zurück.

NETZE [Net92] untersucht den Ultraschallschweißprozess anhand des DVS-Probekörpers (siehe Abschnitt 2.3.3) und identifiziert den Schweißwegverlauf und seine verschiedenen Phasen (siehe Abschnitt 2.1.4) als die wichtigste Prozessgröße. Für die ersten beiden Wegphasen des Ankoppelns und Anschweißens wird eine Modellierung durch gekoppelte Feder-Masse-Schwinger untersucht. Das Fließverhalten in der dritten Phase wird über ein analytisches Modell inklusive Scher- und Dehngeschwindigkeiten abgebildet, wobei jedoch eine Reihe von Vereinfachungen und Annahmen getroffen wird. Für die Versuche wird die Schweißmaschine mit Zusatzsensorik für Pneumatikdruck, statische Schweißkraft, Sonotrodenamplitude, Generatorwirkleistung und Fügeweg ausgestattet; bei einigen Versuchen wird zusätzlich ein Ni/CrNi-Thermoelement in die Fügeebene eingebracht. Um Veränderungen anhand der Signale beschreiben zu können, bildet NETZE für die Auswertung aus den Signalen ein knappes Dutzend diskreter Kenngrößen: Amplitudenmittelwert, minimale und maximale Kraft sowie deren Auftrittszeitpunkt, Werte des Wegsignals zu festen Zeitpunkten (Ende Schweißzeit, Ende Haltezeit) und die Fügegeschwindigkeit in der dritten Prozessphase. Da das Erreichen der dritten Prozessphase für das Erreichen einer Mindestnahtfestigkeit erforderlich ist, sind die Phasenübergänge wichtige Kennzeichen. NETZE kommt jedoch zu dem Ergebnis, dass Wirkleistung und eingebrachte Energie nicht oder nur sehr wenig auf die Phasenübergänge reagieren und sich nur bei starken Schwingungszustandsvariationen, z. B. durch höhere Kräfte, deutlich verändern. Eine Qualitätssicherung kann dadurch erfolgen, dass der Fügewegverlauf visuell beurteilt wird, was jedoch nur zu einer Information über die prinzipielle Eignung der aktuellen Maschineneinstellung führt. Als weitere sinnvolle Messgröße wird die statische Schweißkraft empfohlen, anhand derer der Zeitpunkt des Aufschmelzens genauer erkannt werden kann.

STROHFUSS [Str95] bezieht sich in den wesentlichen Prozessverläufen und -analysen vorwiegend auf Arbeiten von NETZE [Net92] und KORTE/MICHAELI [MK93, MK94a, MKDE94, MWK94]. STROHFUSS unterstreicht den Einfluss der Spritzgießbedingungen auf den Schweißvorgang und beschreibt die Möglichkeiten, die Festigkeit durch Verwendung von Kraft- oder Amplitudenprofilen zu optimieren. Auch ist nach STROHFUSS die zur Überwachung als Qualitätskriterium herangezogene Maschinenenergie nicht sensibel genug, um auf

die qualitätsrelevanten Aufschmelzprozesse in der Fügeebene zu reagieren. Trotz gleicher abgegebener Maschinenenergie ergeben sich in der Praxis unterschiedliche Festigkeiten. Für Untersuchungen an Probekörpern unter Laborbedingungen sind für statistische Prozessmodelle mittels Regressionsgleichungen Bestimmtheitsmaße von 95 bis 98 % erzielt worden. In der Praxis zeigte sich jedoch, dass die Toleranzen von Werkstoff, Bauteil und Prüfmethode zum Teil deutlich größer sind als die Variation der Prozesskennzahlen, so dass die Modellgüte entsprechend abnimmt. Wichtig ist demnach, die Bauteile bereits in der Konstruktionsphase schweiß- und prüfgerecht auszulegen. Für verlässliche Schweißergebnisse müssen auch die Spritzgießparameter sorgfältig ermittelt und während der Serienfertigung kontrolliert werden.

KORTE [Kor96] weist detailliert den Einfluss von veränderten Spritzgießbedingungen auf den Schweißprozess anhand eines modifizierten Prüfkörpers nach. Folglich ist eine Prozessoptimierung beim Schweißen nur sinnvoll, wenn Einflüsse durch nicht reproduzierbar hergestellte Fügeteile ausgeschlossen sind (siehe auch [Str95]). Auch eine effektive Qualitätssicherung ist demnach nur prozessübergreifend erfolgsversprechend. KORTE erfasst sensorisch fast dieselben Daten wie NETZE [Net92], lediglich auf die Aufzeichnung der Sonotrodenamplitude verzichtet er zugunsten der Erfassung der Ambossbeschleunigung mittels eines Quarzkristallbeschleunigungsaufnehmers als Maß für die Schwingungsübertragung. Aus den erfassten Signalen werden offline im Wesentlichen dieselben Kenngrößen ermittelt wie bei NETZE, ergänzt um zwei Fügegeschwindigkeiten zu Prozessbeginn und einige Kenngrößen der Ambossbeschleunigung. Ziel ist die Identifikation von Schweißprozessparametern, welche die Einflüsse der Vorprozesse signifikant abbilden. Prozessdaten aus dem Herstellprozess der Probekörper werden nicht erfasst, da das Spritzgießwerkzeug nicht über Sensoren zur Erfassung der relevanten Druck- und Temperatursignale verfügt. Die Auswertung aller Daten erfolgt mittels Varianzanalyse [Mon09, Sch97], um trotz der stark streuenden Ergebnisse der Festigkeitsuntersuchungen belastbare Aussagen zu erhalten. Die Fügegeschwindigkeit reagiert sensibel auf geänderte Gefügestrukturen, die durch Variationen im Spritzgießprozess gezielt herbeigeführt werden. Allerdings wirken sich Variationen von Amplitude und Schweißkraft im Schweißprozess noch stärker auf die Größen aus, so dass ein gekoppelter Einfluss vorherrscht. Auch der absolute Fügeweg verändert sich bei geänderten Herstellbedingungen aufgrund des höheren Schmelzeenergiebedarfs bei größerer Gefügestruktur, jedoch ist dies auch auf das Schweißen mit fest eingestellter Schweißzeit zurückzuführen. Die aufgezeichnete Ambossbeschleunigung weist im Zeitbereich sehr unterschiedliche Verläufe auf – sowohl innerhalb als auch zwischen den verschiedenen Parametersätzen. Es treten zwar häufig Übereinstimmungen beim Auftreten von Peaks in Ambossbeschleunigung und vom Konverter abgegebener Wirkleistung auf, eine allgemein gültige Korrelation kann jedoch nicht nachgewiesen werden. Die ermittelten Temperaturanstiegsgradienten von fast 4000 °C/s passen laut KORTE zu anderen Untersuchungen wie z. B.

RAMARATHNAM [RNW91]; die Möglichkeiten der Fehlmessung aufgrund schwingungsinduzierter Reibungswärme zwischen Thermoelement und Werkstück, wie von POTENTE und RITTER [Pot71, Rit86] dargestellt, werden nicht bewertet.

Anhand statistischer Versuchspläne modelliert KORTE den Schweißprozess mittels multipler linearer Regression, die Verwendung nichtlinearer Terme brachte keine Verbesserung. Die Modellierung erfolgt sowohl für amorphe wie für teilkristalline Materialien. Dabei führt die Modellierung mittels Prozesskenngrößen aus den aufgezeichneten Signalen stets zu besseren Ergebnissen, als wenn allein die von der Schweißmaschine ausgegebenen Prozessdaten verwendet werden. Insgesamt werden für die Prognose der Festigkeit Bestimmtheitsmaße im Bereich von 74 bis 80 % erreicht. Die erzielbare Modellgüte bei einem Längenmaß stieg auf 88 %, wenn auch Qualitätsmerkmale der Fügepartner – hier Geometriemaße, deren Ausprägung vom Spritzgießprozess abhängt – in die Modellierung mit einbezogen werden. Die Folgerung daraus ist, dass bessere Prozessmodelle möglich sind, wenn prozesskettenübergreifende Daten mit einfließen. Plausibel wird dies auch dadurch, dass die beim Schweißen aufgetretene Prozessvarianz Anteile enthält, die auf geänderte Maschinenparameter, äußere Einflüsse wie die variierten Herstellbedingungen und statistische Streuungen zurückzuführen sind.

KORTE kommt allgemein zu dem Schluss, dass die Vorhersage von Qualitätsmerkmalen beim Ultraschallschweißen durch Prozessmodelle prinzipiell möglich ist, wenn das Merkmal reproduzierbar ermittelt werden kann und zudem der Prozess Größenverläufe aufweist, die sich abhängig vom Niveau des Qualitätsmerkmals signifikant ändern. Diese Aussage trifft jedoch für alle Prozesse zu und ist eine Grundlage statistischer Prozessmodelle.

WIJK ET AL. [vLvN96] vergleichen anhand des DVS-Probekörpers (siehe Abschnitt 2.3.3) die Leistungsfähigkeit verschiedener Maschinen unterschiedlicher Hersteller. Zudem werden die Einflüsse verschiedener Schweißparameter mittels mehrerer Versuchspläne (vollfaktoriell und nach TAGUCHI) untersucht. Als Parameter werden dabei Amplitude, Schweißkraft und eingebrachte Energie variiert; Abschaltkriterium ist die Schweißenergie. Die Schweißnähte werden zum einen durch eine Zugprüfung, zum anderen durch eine optische Begutachtung des Austriebs und der Bruchzone qualifiziert. Zur Ermittlung eines Zusammenhangs zwischen den Einstellparametern und der Festigkeit wurden Regressionsanalysen (siehe Abschnitt 2.4.2) mit den Einstellgrößen Schweißamplitude, Schweißenergie, Schweißkraft sowie deren Quadrate und Kreuzterme als Eingangsgrößen mit dem Ziel durchgeführt, einen optimalen Arbeitspunkt zu finden. Die Optimierung des Arbeitspunkts erfolgt neben einer hohen Festigkeit alternativ auch nach einem robusten Arbeitspunkt. Als Größe für die Robustheit wird die Streuung der Festigkeit bei den einzelnen Maschineneinstellungen anhand des Signal-Rausch-Verhältnisses $S/N = 10 \log_{10} \left(\frac{y^2}{s^2} - \frac{1}{n} \right)$ nach TAGUCHI [Tag87] ermittelt. Dabei stellt y den Mittelwert und s die Standardabweichung der Zugfestigkeit der n Wiederholungen im betrachteten Betriebspunkt dar. Je nach Optimierungsziel ergeben

sich verschiedene optimale Maschineneinstellungen. WIJK ET AL. stellen zum einen fest, dass Amplitude und Energie einen großen Einfluss auf die Festigkeit haben, die Schweißkraft jedoch vernachlässigbar ist. Allerdings ist der Einfluss der Parameter je nach Material deutlich unterschiedlich ausgeprägt. Für eine Prozessüberwachung ist folglich eine Adaption auf jeden Prozess mit einer Selektion der wichtigen Kenngrößen erforderlich. Die durchgeführten Impedanzmessungen im angekoppelten Zustand des Schwinggebildes erlauben keine Aussage über die erzielten Schweißnahtfestigkeiten. Ebenso ist die Schweißzeit als Überwachungskriterium ungenügend, da sie bei konstanter Energie keine Korrelation zur Festigkeit aufweist.

NONHOF [NL96] detailliert die Vorgänge und Größenverhältnisse beim Ultraschallschweißen anhand theoretischer Überlegungen und Abschätzungen. Insgesamt kommt er zu einer großen Übereinstimmung mit existierenden Praxisrichtlinien der DVS für das Ultraschallschweißen [N.N07b]. Ausgangsbasis ist die Tatsache, dass der temperaturabhängige mechanische Verlustfaktor für Kunststoffe im Bereich 20-40 kHz, dem Frequenzbereich des Ultraschallschweißens, nicht bekannt ist, sondern nur durch Extrapolation von Messungen bei geringerer (3-110 Hz) und höherer Frequenz (500-3500 kHz) errechnet wurde. Anhand von Betrachtungen der Energieumsetzung, des Materialflusses in der Fügezone sowie des Wärmeflusses kommt er u. a. zu dem Ergebnis, dass das Ultraschallschweißen aus energetischer Sicht eher ineffizient ist, da nur einige Prozent der eingebrachten Energie tatsächlich für das Aufschmelzen der Fügezone aufgewendet werden. Eine weitere Schlussfolgerung ist, dass es aufgrund der Dämpfungseigenschaften der Kunststoffe am günstigsten ist, die Sonotrode möglichst nahe an der Schweißnaht zu platzieren.

Nach MICHAELI ET AL. [Mic97] hat die gesamte Prozesskette von der Produktentwicklung (d. h. Materialauswahl und Bauteilkonstruktion) über die Prozessplanung bis zur Bauteilherstellung – vom Materialeingang beim Spritzgießer über das Spritzgießen und die Lagerung bis zum Ultraschallschweißen – einen entscheidenden Einfluss auf die erzielbare Schweißnahtqualität. In der Praxis erfolgt heute jedoch häufig erst am Ende der Prozesskette eine Überwachung der Schweißnahtqualität, wobei die dabei verwendeten Prozessdaten der Schweißmaschine im Allgemeinen keine oder nur eine geringe Korrelation zur tatsächlich erzielten Schweißnahtqualität aufweisen. Am Beispiel eines Ventilgehäuses, welches aufgrund der hohen Ausschussrate von ca. 40% und nicht reproduzierbarer Schweißnahtqualitäten nicht wirtschaftlich gefertigt werden konnte, wird eine umfangreiche Prozessanalyse – d. h. grundlegende Untersuchungen der Zusammenhänge zwischen aufgezeichneten Verläufen von Prozessparametern und erreichter Teilequalität – durchgeführt. Dazu werden anhand einer mit zusätzlicher Messtechnik ausgerüsteten Labormaschine Zusammenhänge zwischen Prozesssignalen und der Dichtigkeit der Ventile untersucht und in qualitätsrelevanten Prozesskennzahlen abgebildet. Dabei erweisen sich einige Kenngrößen als besonders signifikant und ein aus ihnen erstelltes Prozessmodell in Form einer Regressionsgleichung

kann undichte Teile mit hoher Trefferquote erkennen. Ein anderes Prozessmodell mit weiteren Kenngrößen sowie insbesondere ein Modell mit Prozessdaten, die standardmäßig von der Maschine erfasst werden, führt dagegen nicht zum Erfolg. Durch eine anschließende Modifikation der Bauteilgeometrie kann der Prozess dahingehend stabilisiert werden, dass der Ausschuss von 40 % auf ca. 40 ppm gesenkt werden konnte [MKK97].

CHUAH ET AL. [CCCL00] untersuchen den Einfluss verschiedener Parameter wie Schweißzeit, Schweißkraft und Form des Energierichtungsgebers auf die erzielbare Festigkeit der Schweißnaht. Die Schweißzeit, welche als Abschaltkriterium des Schweißprozesses dient, wird als wichtiger Parameter identifiziert. Wie BENATAR [Ben87] stellen CHUAH ET AL. das Vorhandensein einer kritischen Schweißzeit fest, deren Überschreiten zu einer Reduktion der Festigkeit aufgrund zunehmender Molekülkettenorientierung führt. Die Schweißkraft weist werkstoffabhängig teils einen Effekt auf; für den Werkstoff ABS verändert sich die erzielbare Festigkeit durch eine Variation der Schweißkraft jedoch nicht. Durch die Definition der Schweißeffektivität als Quotient aus erzielter Festigkeit zur Festigkeit des Grundmaterials wird ein Vergleich verschiedener Einstellungen in einem Diagramm möglich. Weiter wird die Temperaturverteilung im Werkstück oberhalb des ERG mittels mehrerer Thermoelemente mit definiertem Abstand erfasst, um die Energieübertragungsmechanismen im Werkstück zu verstehen. Dabei zeigt sich kurz nach dem Start des Schweißprozesses ein Anstieg der Temperatur an allen Messpunkten, mit dem höchsten Wert direkt an der Schweißnaht und stark abfallend Richtung Sonotrode. Aus den Temperaturen wird für PE und ABS die jeweilige Energieabsorption berechnet. Das führt zu dem Ergebnis, dass der ERG aus ABS mehr Energie absorbiert, was darauf zurückgeführt wird, dass ABS als amorpher Thermoplast den Ultraschall besser leitet und somit die Energie einfacher zur Schweißzone gelangen kann. Dies deckt sich mit den in Abschnitt 2.1.3 dargestellten mechanischen Verlustfaktoren bei Thermoplasten.

ZIEGLTRUM [Zie00] untersucht das Ultraschallschweißen zum Fügen von Werkstücken aus verschiedenen Werkstoffen. Die umfangreichen Untersuchungen umfassen Materialkombinationen aus 23 Werkstoffen, die insbesondere in der Automobil- und Elektroindustrie zum Einsatz kommen. Ein wesentliches Ergebnis ist ein Überblick, wie gut sich verschiedene Materialpaarungen schweißen lassen und welche Festigkeiten erzielt werden können. Für die Untersuchungen verwendet ZIEGLTRUM einen radialsymmetrischen Probekörper mit ERG im Nahfeld (siehe Abschnitt 2.3.3). Das Schweißen aller Probekörper erfolgt im RPN-Modus, um einen konstanten Schweißquerschnitt zu erhalten. Die Festigkeit als relevante Zielgröße wird mittels einer Zugprüfung ermittelt, wobei als Festigkeit der Mittelwert von zehn geprüften Probekörpern verwendet wird. Die Prüfergebnisse der einzelnen Versuchszyklen weisen teils eine Standardabweichung von bis zu 20 % des mittleren Festigkeitsniveaus auf. Diese Größenordnung entspricht laut ZIEGLTRUM auch den Erfahrungen aus anderen Versuchen und der Praxis. Anhand von Materialkenngrößen der verschiedenen verwendeten

Werkstoffe erfolgt eine Abschätzung der notwendigen Schweißenergie. Neben der Verifikation der aus der Literatur bekannten Abhängigkeit der erzielbaren Festigkeit von der eingestellten Fügekraft untersucht ZIEGLTRUM auch den Einfluss des Feuchtegehalts sowie der Anbringung des ERG an Ober- oder Unterteil auf Schweißbarkeit und Nahtfestigkeit. Abschließend prüft er noch die Möglichkeit, Werkstücke mit mehr als einer Fügeebene zu verschweißen, was neue Freiheitsgrade in der Gestaltung von Bauteilen erlaubt.

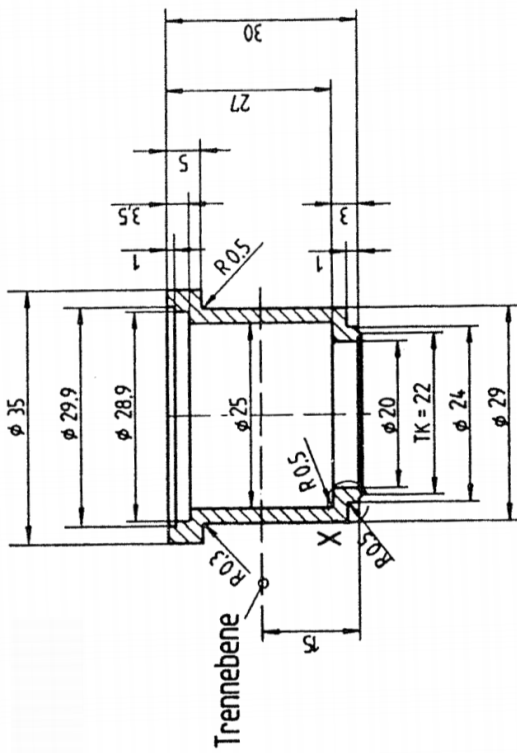
LIU ET AL. [LC01, LLC⁺99] untersuchen den Einfluss von sechs Einstellgrößen auf die Festigkeit anhand eines Probekörpers für verschiedene Materialien (PS, PP, PP+10%GF, PP+30%GF). Neben Schweißzeit (zugleich Abschaltkriterium), Schweißkraft und Amplitude werden auch Haltezeit, Haltedruck und die Geometrie des Energierichtungsgebers in je drei Einstellungen variiert. Da dies bei einem vollfaktoriellen Versuchsplan 3^6 (also 729) Einstellungen bedeuten würde, erfolgt über einen L18-Versuchsplan nach TAGUCHI eine Reduktion auf 18 Einstellungen. Das Optimierungsziel war die Maximierung der Festigkeit. Analog dem Signal-Rausch-Verhältnis S/N (signal-to-noise-ratio, SNR), welches von einer Verlustfunktion abgeleitet ist [Mon09], wird als Optimierungsfunktion $S/N = -10 \log_{10}(\frac{1}{n}) \sum_{i=1}^n (\frac{1}{y_i^2})$ verwendet. Für jede der Einstellungen wird die SNR und anschließend der sich daraus ergebende Effekt jeder Faktorstufe berechnet. Das Optimum setzt sich dann aus genau jenen Stufen aller Faktoren zusammen, welche die höchste SNR aufweisen. Da die so ermittelte optimale Einstellung nicht im Versuchsplan enthalten war, wird für diese über $SNR_{opt} = SNR_{MW} + \sum(SNR_{opt.Faktor} - SNR_{MW})$ eine Vorhersage getroffen. In Verifikationsversuchen zeigt sich, dass diese höhere Festigkeitswerte aufweist als im ursprünglichen Versuch. Die wichtigsten Einflüsse auf den Prozess wurden mittels ANOVA, einem statistischen Verfahren zur Beurteilung der Varianz, identifiziert. Bei allen untersuchten Materialtypen sind dies neben der Amplitude die Schweißzeit sowie die Form des ERG. Die anderen drei Parameter spielen jeweils keine signifikante Rolle. Der halbkreisförmige ERG führte insgesamt zu den höchsten Festigkeitswerten.

Im Rahmen eines AIF-PROJEKTS [N.N03] wurde der Aufbau eines auf Neuronalen Netzen basierenden Qualitätssicherungssystems für das Ultraschallschweißen von Metallen untersucht, da auch dort die Verbindungsgüte Streuungen aufweist und eine zerstörungsfreie, 100%-ige Qualitätssicherung nicht zufriedenstellend gelöst ist. Die Maschinensignale Weg und Leistung sowie externe Messsysteme für Beschleunigung des Ambosses und Temperaturmessung werden verwendet, um aus Versuchszyklen Eingangsdaten für ein Neuronales Netz (NN) abzuleiten. Als Eingangsdaten des NN dienen die Signalwerte zu festen Zeitpunkten. Die Ergebnisse anhand der Maschinen zeigen, dass ein QS-System auf der Basis von NN in der Lage ist, innerhalb entsprechender Vertrauensgrenzen für die Schweißnaht eine Prognose über die Schälzugfestigkeit der gefertigten Schweißverbindung zu treffen. Auch einige Fehlerarten werden hinsichtlich des Erreichens einer Mindestfestigkeit vom NN richtig eingeschätzt. Bei einer Variation des Litzentyps konnte ein anhand von Produktionsdaten

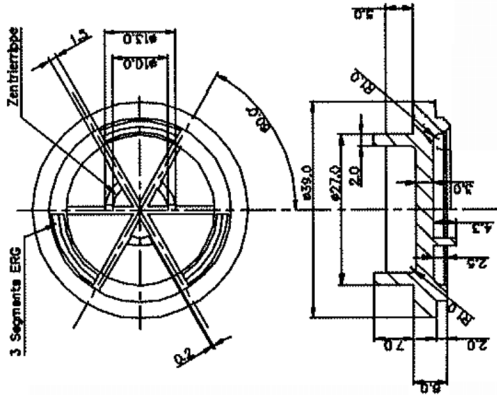
trainiertes Netz die Festigkeit nicht ausreichend genau vorhersagen. Das Netztraining muss also für jeden Anwendungsfall separat erfolgen, ein „allgemein gültiges“ Netz existiert nicht. Die Autoren ziehen als Fazit, dass bei richtigem Einsatz NN einen Beitrag zur Erhöhung der Fertigungssicherheit leisten können. Es ist zu erwarten, dass mittels NN mittelfristig die Möglichkeiten zur Qualitätssicherung in der Ultraschall-Schweißtechnik erweitert werden.

HABERSTROH [HK04] beschreibt verschiedene Prozessführungsstrategien beim Ultraschallschweißen sowie die Möglichkeiten der Geschwindigkeitsregelung durch eine servoelektrisch angetriebene Schweißmaschine. Untersuchungen mit dem DVS-Probekörper (siehe Abschnitt 2.3.3) für verschiedene Werkstoffe ergeben, dass der Einfluss der verschiedenen Prozessführungsstrategien auf die erzielbare Schweißnaht materialabhängig ist und keine allgemein gültigen Aussagen getroffen werden können. Untersucht wird zudem, ob mit den am Markt relativ neu verfügbaren digitalen Generatoren und den damit zusätzlich verfügbaren Prozesssignalen wie Frequenz, Strom und Phase eine bessere Prozessüberwachung als bei analogen Generatoren möglich ist. Die ermittelten Regressionsgleichungen ergeben eine Zunahme des Bestimmtheitsmaßes bei der Qualitätsvorhersage von 78,1 % auf 79 %. Die Verwendung digitaler Generatoren allein bringt somit keine wesentliche Verbesserung bei der Prozessüberwachung. Beim Schweißen von Folien konnten hingegen für die Reißfestigkeit Bestimmtheitsmaße von über 90 % erzielt werden.

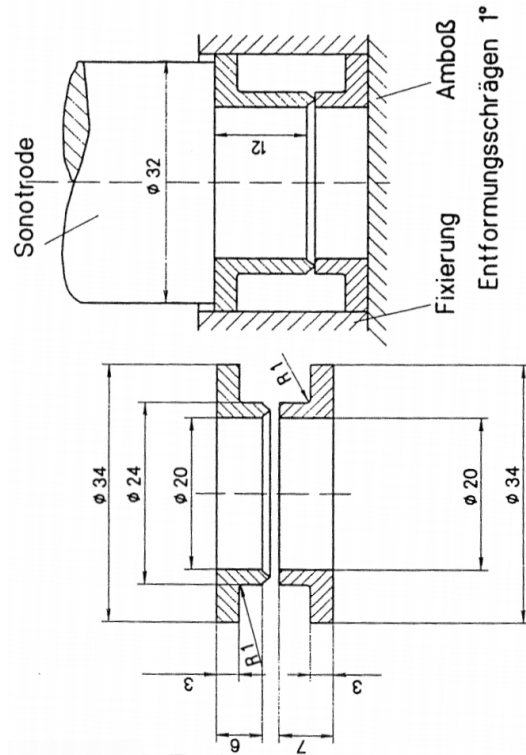
LING [LLLA06] verwendet NN zur Prognose der Festigkeit bei Versuchen mit konstanter Schweißzeit anhand des AWS-Probekörpers (siehe Abschnitt 2.3.3). Er zeigt, dass die Impedanzmessung einer Piezokeramik bei Anregung mit hochfrequenter Spannung dazu verwendet werden kann, Veränderungen an einem gekoppelten mechanischen Bauteil zu erkennen [LX01]. Analog wird beim Ultraschallschweißen das angeregte System aus dem Schwinggebilde mit gekoppeltem Werkstück gebildet, während der Konverter Aktor und Sensor zugleich ist. Da sich Konverter und Schwinggebilde im Laufe einer Schweißung mechanisch nicht verändern, ist eine zeitliche Variation der Impedanz des Konverters ein Maß für Veränderungen im Werkstück während der Schweißung, d. h. sehr hoch korreliert mit dem thermisch-mechanischen Verhalten des Werkstücks in der Fügezone. Da eine direkte Verwendung des Real- und Imaginärteils des Impedanzverlaufs aufgrund der Datengröße zu sehr großen Netzen mit langen Trainingszeiten führen würde, werden mittels Principal Component Analysis (PCA) die Daten auf eine handhabbare Größe reduziert, ohne wesentliche Information zu verlieren. Die mit den PCA-Daten trainierten NN sind geeignet, die Festigkeit für dem Netz unbekannte Testdaten innerhalb eines gewissen Vertrauensintervalls zu prognostizieren.



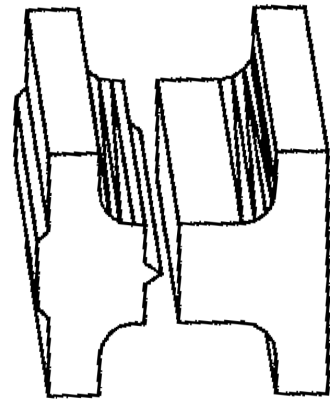
(a) DVS-Probekörper (aus [Net92])



(b) Probekörper von ZIEGLTRUM (aus [Zie00])



(c) Probekörper von KORTE (aus [Kor96])



(d) AWS-Probekörper (aus [Zie00])

Abbildung 2.7: Verschiedene Probekörpergeometrien

2.3.3 Verwendete Probekörpergeometrien

In der Literatur sind verschiedene Geometrien für Probekörper aufgeführt (siehe Abbildung 2.7), die eine Durchführung größerer Versuchsreihen zur Untersuchung des Ultraschallschweißens ermöglichen sollen. Je nach eingesetztem Probekörper sind verschiedene Prüfmethode und somit Analysen unterschiedlicher Qualitätsmerkmale möglich. Im Folgenden wird ein kurzer Überblick über Werkstücke gegeben, die zur Charakterisierung und Untersuchung des Schweißprozesses angewendet werden (größtenteils nach [Net92]):

- Das Verschweißen bzw. Aneinanderschweißen von Norm-Zugstäben (z. B. DIN 53 455, Nr. 3) ermöglicht Zug- oder Druckscherprüfungen [FW80]. Die Zugstäbe verfügen nicht über ERG, so dass allein aufgrund der Konstruktion keine idealen und verfahrensgerechten Schweißungen mit reproduzierbaren Ergebnissen möglich sind.
- BECKERT ET AL. schlagen einen zylindrischen Torsionsprüfkörper mit sehr großen Wanddicken (max. 15 mm) und ohne zusätzliche ERG oder Quetschnähte vor, mit dem Nahfeld- und Fernfeldschweißungen möglich sind [BMB76].
- Anhand von Streifenproben mit einer oder zwei Energierichtungsgeber-Reihen ([Rit86] bzw. [NM91]), die aus spritzgegossenen Platten geschnitten werden, können einfach überlappende Zugscherproben geschweißt werden. Vorteil dabei ist das günstige Spritzgießwerkzeug sowie die Sichtbarkeit der Naht von allen Seiten.
- Möglichkeiten zur Zug- und Torsionsprüfung sowie einer Innendruckbelastung in Fern- und Nahfeldschweißung bieten die spritzgegossenen Zylinderpaare [Mül88], wobei keinerlei Maße angegeben sind.
- Ein plattenförmiger Prüfkörper mit Doppel-T-Profil ermöglicht Fern- und Nahfeldschweißung mit verschiedenen Energierichtungsgeber-Geometrien und einer Quetschnahtform (siehe Abbildung 2.7(d)). Dieser Prüfkörper wird 1990 von der American Welding Society (AWS) vorgestellt [Som90]. Die Eigenschaften des unverschweißten Werkstoffes können mit einem separaten Prüfkörper mit derselben Gesamtgeometrie untersucht und mit dem geschweißten Teil verglichen werden.
- Der Zentralverband der Elektrotechnischen Industrie (ZVEI) entwickelte 1981 eine zylinderförmige Prüfkörpergeometrie, die in zwei Ausführungen alle grundsätzlichen Verfahrensvarianten des Ultraschallschweißens zulässt [Lan81]. Später wurde diese Geometrie auch vom Deutschen Verband für Schweißtechnik (DVS) übernommen und ist bis heute die am häufigsten verwendete Probekörper-Geometrie (siehe Abbildung 2.7(a)). Durch den treppenartigen Aufbau ermöglicht der Probekörper sowohl Nahfeld- als auch Fernfeldschweißungen mit nur einer Probekörpergeometrie. Durch

einen wechselbaren Einsatz im Spritzgießwerkzeug ist es zudem möglich, die Geometrie der Fügeebene zu verändern und Werkstücke mit Energierichtungsgeber- und mit Quetschnaht-Geometrie zu fertigen. Bei der Entwicklung dieses Probekörpers war neben der Fertigung in nur einem Spritzgießwerkzeug mit möglichst wenigen Einsätzen ein weiteres Ziel, die Geometrie auch für die anderen Schweißverfahren (Reibschweißen, Vibrationsschweißen und Heizelementstumpfschweißen) verwenden zu können. Verschiedene Untersuchungen, die anhand dieses Probekörpers durchgeführt worden sind (z. B. [EB88, MEBN88, NM91]), zeigen einige wesentliche Mängel im Hinblick auf das Schweißverhalten, die möglichen Prüfverfahren zur Quantifizierung der Schweißnaht und die Möglichkeiten zur Modellierung für FEM-Simulationen. Wesentliche Schwachpunkte sind:

- Schallumlenkung beim Schweißen von Quetschnähten im Nahfeld sowie ERG im Fernfeld,
 - der geringe Abschmelzweg bei einer ERG-Schweißung,
 - zu geringe Sonotrodenkontaktfläche beim Fernfeldschweißen,
 - starke Beeinflussung des Schwingungsverhaltens des Prüfkörpers durch überstehenden Zylinder beim Nahfeldschweißen,
 - Zug- und Torsionsprüfung sowie Innendruckprüfung als einzig mögliche Prüfverfahren,
 - biegemomentbedingtes Auftreten von mehrachsigen Spannungszuständen bei der Zugprüfung mit der Folge stark schwankender und unsicherer Prüfergebnisse,
 - Notwendigkeit der Anpassung der für den Schweißprozess verwendeten Sonotroden an die Schwindmaße, die jeweils vom verwendeten Werkstoff abhängen,
 - beim Schweißen notwendige Führung der Schweißpartner durch eine innenliegende Zentrierung, die an unterschiedliche Schwindmaße der verschiedenen Werkstoffe angepasst werden muss [Zie00].
- NETZE entwickelt im Rahmen seiner Arbeiten einen Prüfkörper ausschließlich für das Ultraschallschweißen, der verschiedene der aufgeführten Schwachstellen der DVS-Geometrie verbessern soll [Net92]. Um die Vergleichbarkeit zu bereits bestehenden Festigkeitsergebnissen für verschiedene Werkstoffe zu ermöglichen, ist auf eine möglichst identische Fügeflächengeometrie und gleiche Wanddicken des Prüfkörpers wie bei der DVS-Geometrie geachtet worden. Optimierungen sind insbesondere in Bezug auf Schalleinleitung, Schallleitung im Bauteil bis zur Fügeebene sowie die Energieumsetzung vorgenommen worden. Dabei wurde auf einfache Geometrien geachtet, um Wärmeleitprozesse, Schwingungsverhalten und Fließprozesse in der Fügeebene mittels FEM und analytischer Methoden modellierbar zu machen. Durch verschiedene

Einsätze im Spritzgießwerkzeug sind verschiedene ERG- und Quetschnahtformen sowie verschiedene Längen des Probekörpers (6, 12 und 30 mm) zwischen Fügeebene und vergrößerter Sonotrodenkontaktfläche möglich. Aufgrund der gewählten Zylindergeometrie sind ausschließlich dieselben Prüfmethode wie beim DVS-Probekörper möglich.

- KORTE [Kor96] verwendet großteils den von NETZE entwickelten Probekörper, der eine Variation des Abstands zwischen Sonotrode und Fügeebene sowie des ERG (Höhe und Spitzenwinkel) ermöglicht. Bei weiteren Untersuchungen wird eine Quetschnaht als Nahtgeometrie eingesetzt [MK94b]. Da bei Schweißverbindungen mit höherer Festigkeit, insbesondere bei amorphen Thermoplasten, häufig ein Bruch im Grundmaterial auftritt und somit keine Festigkeit der Schweißnaht ermittelt werden kann, modifiziert KORTE den Probekörper für weitere Versuche (siehe Abbildung 2.7(c)). Der Zylinder wird konisch ausgebildet und der Übergang zur Deckel- und Bodenfläche mit größeren Radien versehen. Dadurch wird gewährleistet, dass die maximalen Spannungen in der Fügeebene auftreten und dort der Bruch erfolgt.
- Der Probekörper des DVS wird auch von WIJK zu Untersuchungen herangezogen [vLvN96]. Auch dort wird dem Probekörper bei der Zugprüfung der Nachteil bescheinigt, die Schweißnaht aufgrund einer starken Biegekomponente einer kombinierten Zug- und Schälprüfung auszusetzen.
- ZIEGELTRUM [Zie00] verwendet einen eigens konstruierten zylindrischen Probekörper mit ERG für Nahfeldschweißungen (siehe Abbildung 2.7(b)). Die Schweißnaht ist nicht durchgängig, sondern in Segmente aufgeteilt. Dadurch können Ober- und Unterteil des Probekörpers symmetrisch aufgebaut werden und es ist nur ein Spritzgießwerkzeug notwendig. Zur Minimierung von auftretenden Biegungen bei der Zugprüfung, die zu mehrachsigen Spannungszuständen führen, ist der Probekörper mittels Rippen verstärkt. Das Auftreten von entsprechenden Biegekräften ist beim DVS-Probekörper ein bekanntes Problem. Die Zentrierung beim Schweißen erfolgt durch die Probekörper selbst, eine äußere, auf die Schwindung abzustimmende Aufnahme und innere Führung wie beim DVS-Probekörper ist nicht notwendig. Geometrieänderungen der Schweißnahtkontur können durch Verwendung anderer Werkzeugeinsätze realisiert werden.
- Die DVS-Arbeitsgruppe AG W 4.1d hat sich unter besonderem Engagement der BASF AG, Ludwigshafen, um einen einheitlichen Prüfkörper bemüht [MKW07]. Dieser ist zusammen mit den Prüfbedingungen in einer DVS-Richtlinie beschrieben [N.N05]. Die Geometrie basiert im Wesentlichen auf der von NETZE entwickelten Geometrie, jedoch

wurden ähnlich wie bei KORTE [Kor96] Modifikationen vorgenommen, insbesondere bei Radien, der Verjüngung des Zylinders sowie den Entformungsschrägen.

2.4 Modellierungsmethoden

Verschiedenste Modellierungsmethoden ermöglichen den Aufbau von Modellen, die einen Zusammenhang zwischen Eingangs- und Zielgröße ermöglichen. Bei der Prozessüberwachung sind die Eingangsgrößen Messgrößen aus dem Prozess; ein Qualitätsmerkmal des gefertigten Teils stellt die Zielgröße dar. Die wichtigsten und gebräuchlichsten Methoden zur Modellierung von Systemen und Prozessen sind im Folgenden aufgeführt. Daneben gibt es noch eine Vielzahl von anderen Beschreibungsformen wie beispielsweise Bayes-Netze oder Hidden-Markov-Modelle, auf die jedoch insbesondere aufgrund unterschiedlicher Prozessvoraussetzungen oder Modellierungsziele nicht näher eingegangen wird. Für einen erweiterten Überblick über die verschiedenen Methoden wird auf die entsprechende Literatur verwiesen (z. B. [BH99, Bis06, DHS00, HTF01, Mit97, Vap95]).

2.4.1 Mathematisch-physikalische Modelle

Die exaktesten und genauesten Modelle liefert eine mathematisch-physikalische Prozessbeschreibung. Dazu wird der zu modellierende Prozess durch mechanische, chemische und elektrische Gesetzmäßigkeiten beschrieben. Dabei kommen maßgeblich auch die Erhaltungssätze von Masse und Energie zum Einsatz. Eine solche Modellierung ist zwar genau, liefert aber meist schon für einfachste Systeme gekoppelte Differenzialgleichungen mit entsprechenden prozessabhängigen Konstanten (z. B. Federsteifigkeiten, Dämpfungskoeffizienten, Materialkonstanten etc.). Für reale Prozesse sind diese Konstanten nur schwer oder gar nicht quantifizierbar. Oft sind sie zudem von anderen Zustandsvariablen abhängig, z. B. temperaturabhängige Dämpfung. Aufgrund ihrer Komplexität sowie unbekannter oder nicht quantifizierbarer Einflüsse sind mathematisch-physikalische Modelle für die allermeisten Prozesse aus dem Bereich des Ur- und Umformens nicht verfügbar (siehe z. B. [MBH⁺92]). Für das Spritzgießen, das sehr intensiv erforscht ist und wofür verschiedene Simulationswerkzeuge wie Mouldflow z. B. für Füllvorgänge verfügbar sind, gibt es unter bestimmten Randbedingungen gültige Modelle für Teilbereiche, wie z. B. die pvT- oder pmT-Regelung [MLBB89]. Es existiert jedoch kein umfassendes Modell, das den Gegebenheiten in der Praxis gerecht wird. Insbesondere treten bei mathematisch-physikalischen Modellen Probleme auf, wenn in der Praxis Randbedingungen oder Störeinflüsse vorliegen, die im Modell nicht erfasst sind [PNH92].

2.4.2 Datenbasierte Modelle

Aufgrund des immensen Aufwands zur Erstellung von mathematisch-physikalischen Modellen werden für die Beschreibung von Fertigungsprozessen unter realen Bedingungen zu meist empirische Modelle eingesetzt. Diese werden auf der Basis von „Beobachtungen“ des Prozesses – auch Trainingsdaten genannt – erstellt, d. h. in der Regel anhand von aufgezeichneten Messdaten. Nachteilig bei der datenbasierten Modellierung ist die Tatsache, dass nur Zusammenhänge erkannt und modelliert werden können, die sich in den Trainingsdaten abbilden. Um die Modelle entsprechend absichern zu können, ist ein relativ hoher Versuchsaufwand notwendig. Idealerweise wird für Versuche ein statistischer Versuchsplan erstellt, der bei entsprechender Planung auch die Abbildung von Wechselwirkungen erlaubt [Kle03, Mon09, Sch97, Tag87]. Ein Nachteil der statistischen Versuchsplanung ist der notwendige zeitliche, personelle und finanzielle Aufwand. Dieser lohnt sich für kleine bis mittlere Losgrößen meist nicht, so dass der ideale Einsatzbereich in der Produktionsplanung von Großserienteilen zu sehen ist [PNH92].

Die am häufigsten verwendeten Methoden sind die Modellierung mittels Regressionsmodellen und mit Neuronalen Netzen in Form des mehrschichtigen Perzeptrons (MLP). Auf die wichtigsten der datenbasierten Modelle wird im Folgenden kurz eingegangen.

Regressionsmodelle

Während die Korrelationsanalyse, d. h. die Berechnung des Korrelationskoeffizienten, ein qualitatives Maß für einen (linearen) Zusammenhang zwischen zwei Größen liefert, wird durch die Regressionsanalyse ein quantitativer, funktionaler Zusammenhang hergestellt [HEK09, Mas80]. Besteht zwischen den beiden Merkmalen X (Regressor) und Y (Regressand) ein linearer Zusammenhang oder wird ein solcher vermutet, kann anhand der n Datentupel (x_i, y_i) mit den Ausprägungen der beiden Merkmale eine Regressionsgerade berechnet werden. Anhand der n Gleichungen $y_i = \alpha + \beta x_i + e_i$, wobei e_i ein zufälliger Fehler (mit Mittelwert 0) ist, können die Werte α und β bestimmt werden, so dass die Quadratsumme der Fehler $S^2 = \sum_{i=1}^n (y_i - \alpha - \beta x_i)^2$ minimal ist. Die Güte des erreichten Regressionsmodells wird durch das Bestimmtheitsmaß $R^2 = \frac{Cov(X,Y)^2}{Var(X)Var(Y)}$ beschrieben, welches sich aus der Kovarianz und den Einzelvarianzen berechnet. Das Bestimmtheitsmaß ist stets im Wertebereich $0 \leq R^2 \leq 1$, wobei für einen perfekten linearen Zusammenhang der Wert $R^2 = 1$ erreicht wird.

Bei der multiplen Regression wird anhand verschiedener Regressoren X auf den interessierenden Regressanden Y geschlossen. Dabei wird ein funktionaler Zusammenhang der Form $y(x_1, \dots, x_k) = \alpha + \sum_{j=1}^k \beta_j x_j$ bestimmt. Voraussetzung für gute Modelle ist, dass ein linearer Zusammenhang vorliegt, was für praktische Prozesse meist nur sehr bedingt erfüllt ist. Nichtlinearitäten oder Wechselwirkungen zwischen den Einflussgrößen können auf diese

Weise nicht erfasst werden. Durch eine entsprechende Substitution, z. B. $\hat{x} = \log(x)$ oder $\hat{x} = x^2$, ist es möglich, auch nichtlineare Zusammenhänge in bestimmtem Umfang zu modellieren. Es bleibt jedoch das Problem, dass die Art des Zusammenhangs bekannt sein muss, um ein hochwertiges Modell zu erhalten.

Die multiple Regressionsrechnung bietet die Möglichkeit, das Streuverhalten am Arbeitspunkt zu beschreiben [HE84, WH91]. Auch wenn zwischen Prozessgrößen und Qualitätsmerkmalen einzeln keine Korrelationen vorherrschen, kann durch die Verknüpfung von mehreren Kenngrößen eine hohe Korrelation und ein hohes Bestimmtheitsmaß erreicht werden [PNH92]. Die Regressionsanalyse für die Prozessmodellierung wird für die Berechnung eines Qualitätsmerkmals (Regressand) in Abhängigkeit von mehreren Prozessparametern (Regressoren) in der Literatur häufiger und für verschiedene Bereiche verwendet (z. B. [Kor96, Nat92, Ueb95]).

Beim Heizelement-Stumpfschweißen ist der Aufbau eines Prozessmodells zur Qualitätssicherung durch eine Regressionsanalyse oft nicht möglich, da ein breites Band möglicher Einstellungen beim Schweißen von Polypropylen (PP) zum gleichen Qualitätsniveau führt [Kre87]. Beim Heizelementschweißen ist eine Prozessmodellierung im Arbeitspunkt allein anhand der Prozesssteuerung nicht möglich und für den Aufbau einer kontinuierlichen Prozessüberwachung (CPC) ist die Durchführung eines Versuchsplans zwingend notwendig. Die Gründe liegen darin, dass die äußeren Einflüsse auf die Schweißnahtqualität gravierender sind als die geringen Streuungen der Prozessparameter im Arbeitspunkt [PNH92].

Neuronale Netze

Ein anderer Ansatz zur datenbasierten Modellierung, der häufig Anwendung findet, wird durch Neuronale Netze (NN) bereitgestellt. In einer Auswertung von „intelligenten Maschinenmodellen“ über den Zeitraum 2002 - 2007 basierten knapp 60 % auf NN [ANS10]. Der Begriff (künstliche) NN steht dabei für eine ganze Klasse von so genannten konnektionistischen Modellen und wird oft mit „künstlicher Intelligenz“ (KI) in Verbindung gebracht. Ein wesentliches Merkmal bei NN ist, dass die Signalverarbeitung bzw. Modellierungsfähigkeit durch eine größere Anzahl an gleichartigen Elementen – Neuronen genannt – mit relativ einfachem Aufbau bewerkstelligt wird. Diese stehen über Verknüpfungen oder gewichtete Verbindungen untereinander in Beziehung und beeinflussen sich gegenseitig durch anregende oder hemmende Signale, die vom Aktivierungszustand der einzelnen Neuronen abhängig sind. Die Motivation dieser Modelle ist der Biologie entlehnt, da auch die leistungsfähigste „Signalverarbeitungsmaschine“ – das Gehirn – aus vielen einzelnen Neuronen besteht, die mittels der so genannten Synapsen sehr stark miteinander vernetzt sind. Auch wenn es eine Vielzahl verschiedener Ansätze zur Abbildung und Verknüpfung der Neuronen gibt, wird unter dem Begriff „NN“ meist das so genannte mehrschichtige Perzeptron (multi-layer perceptron, MLP) verstanden, das mittels eines Backpropagation-Algorithmus trainiert wird.

Die wesentliche Eigenschaft von NN ist ihre Fähigkeit, existierende Zusammenhänge – auch nichtlinearer Art – anhand von vorliegenden Daten zu „lernen“. Der Lernvorgang erfolgt dabei meist iterativ, d. h. es werden die Muster dem Netz so lange präsentiert und anschließend Modellparameter verändert, bis ein bestimmtes Abbruchkriterium, wie z. B. maximale Größe des Fehlers oder die Anzahl der Trainingsdurchgänge, erfüllt ist. Dabei sind überwachte und unüberwachte Lernverfahren zu unterscheiden. Bei Ersteren ist für jedes Muster des Datensatzes der Wert der Zielgröße bekannt. Aus der Differenz zwischen Sollwert und tatsächlicher Netzprognose, die dem Prognosefehler entspricht, kann das Netzmodell so angepasst werden, dass der Fehler über alle Muster minimal wird. Bei unüberwachten Verfahren findet keine Vorgabe der Zielgröße statt, so dass das Netz im Wesentlichen die Ähnlichkeit der Mustersätze untereinander interpretieren und abbilden muss. Der Lernalgorithmus versucht somit, in den Eingangsdaten vorhandene Cluster zu finden. Das am weitesten verbreitete unüberwachte Lernverfahren findet sich in den Self-Organizing Maps (SOM) wieder (siehe Seite 68), nach ihrem Entwickler auch Kohonen-Karten genannt.

Sehr eng mit dem Lernvorgang ist dabei die so genannte Abstraktions- oder Generalisierungsfähigkeit verbunden, d. h. es können auch Zustände, die nicht explizit in den Trainingsdaten vorhanden sind, sondern sich zwischen zwei oder mehreren Trainingsmustern befinden, richtig interpretiert werden. Es ist jedoch auf die mögliche Gefahr des Overfittings – d. h. einer Überanpassung – des Netzes an die Daten zu achten, was mit einem „Auswendiglernen“ der Muster verglichen werden kann. Dabei werden zwar die Trainingsdaten sehr gut abgebildet, die Generalisierungsfähigkeit geht dabei jedoch verloren und dem Netz unbekannte Muster werden nicht mehr sinnvoll interpretiert. Bei der Adaption von NN werden deshalb die verfügbaren Datensätze in Lern-, Test- und Validierungsdaten aufgeteilt. Die Lerndatenmenge dient dazu, die freien Parameter des Netzes zu bestimmen, während die Testmenge zur Ableitung eines Abbruchkriteriums und zur Abschätzung des Kompromisses zwischen Generalisierungsfähigkeit und Auswendiglernen dient. Die Validierungsdaten dienen als unabhängiger Datensatz dazu, das fertig trainierte bzw. adaptierte Modell zu validieren, wobei dies jedoch – meist aufgrund einer stark begrenzten Menge an Datensätzen – oft anhand der Testdaten erfolgt. Die so durch das Training entstehenden Modelle sind in der Regel robust gegenüber Störungen und einem Rauschen in den Eingangssignalen.

Die Neuronalen Netze unterscheiden sich sowohl in der Anordnung und den „Aktivierungseigenschaften“ der einzelnen Neuronen, der Anzahl und Richtung ihrer Verknüpfungen als auch in der Anzahl und Art ihrer wählbaren Parameter sowie den eigenschaftsabhängigen Algorithmen zum Training und der „Wissensabbildung“. Mathematisch können NN als Abbildungsvorschriften $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ betrachtet werden, die den n -dimensionalen Raum der Eingangsdaten auf einen m -dimensionalen Raum abbilden, der bei der Prognose eines

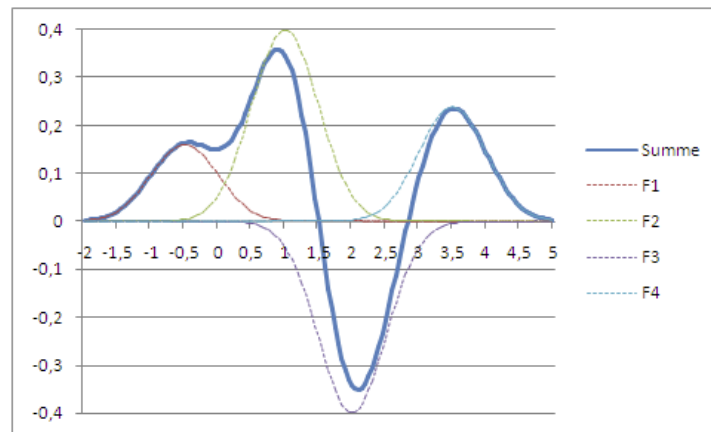
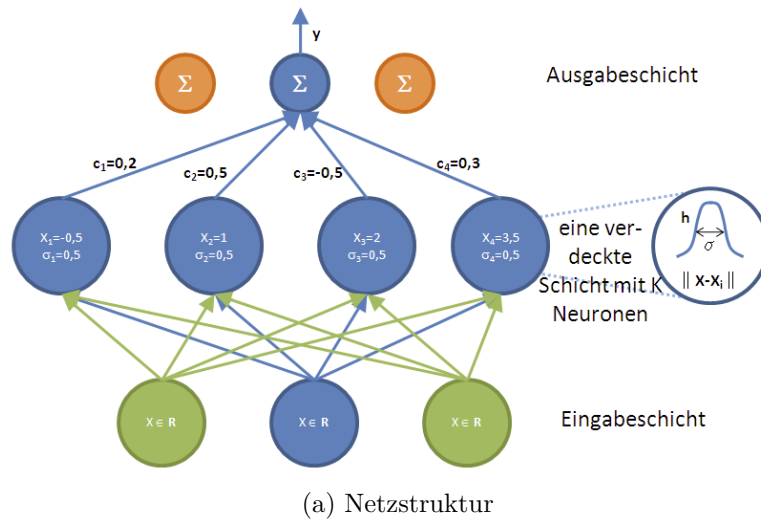


Abbildung 2.8: Approximation einer eindimensionalen Funktion durch ein RBF-Netz mit vier Neuronen an den Stützpunkten $x = \{0, 5; 1; 2, 5; 3\}$ und mit den Gewichtungsfaktoren $c = \{0, 3; 0, 5; -0, 3; 0, 4\}$

Qualitätsmerkmals meist eindimensional ist. Meist können die Abbildungsvorschriften und Algorithmen über entsprechende Matrizenoperationen beschrieben werden. Einen detaillierten Einblick in die Thematik der NN sowie deren Beschreibung und Simulation geben beispielsweise BISHOP [Bis95] oder ZELL [Zel94]. ZELL hat auch maßgeblich an der Entwicklung des *Stuttgarter Neuronale Netze Simulators*, kurz SNNS, mitgearbeitet – einem Softwaresystem zur rechnerbasierten Simulation der verschiedensten Netztypen. Im Folgenden wird ein Überblick über einige wichtige Netzwerkmodelle mit ihren wesentlichen Eigenschaften gegeben.

Radial-Base-Function-Networks (RBF-Netze): RBF-Netze (siehe Abbildung 2.8) sind spezielle, vorwärtsgerichtete neuronale Netze mit nur einer verdeckten Schicht. Ihren Namen haben sie von den radialsymmetrischen Aktivierungsfunktionen der Neuronen der

versteckten Schicht. Dies macht auch die Besonderheit dieses Netztyps und den Unterschied zu MLP-Modellen (siehe Seite 69) aus, die mehrere versteckte oder Hidden-Schichten haben können. Aus mathematischer Sicht bilden diese Aktivierungsfunktionen die n -dimensionalen Basisfunktionen eines Funktionensystems zur Approximation von mehrdimensionalen Funktionen [Zel94]. Die K Neuronen bilden dabei die K Stützstellen $X_i \in \mathbb{R}^n$ für die Approximation und die jeweilige Aktivierungsfunktion $h_j(d) \geq 0$ mit $1 \leq i \leq K$ ordnet dem Neuron für den Eingabevektor X einen skalaren Wert zu. Der Parameter $d = \|X - X_j\|$ ist dabei der Abstand des Eingabevektors X zum Stützpunkt X_j . Der Wert der Aktivierungsfunktion $h_j(d)$ nimmt mit steigendem Abstand des n -dimensionalen Eingabemusters zur Stützstelle des Neurons monoton ab. Für die Berechnung des Abstands wird dabei meist die von der L_2 -Norm induzierte Metrik, d. h. der euklidische Abstand, verwendet. Anders ausgedrückt, sind die Neuronen die Zentren der verwendeten radialen Basisfunktionen. Da die Aktivität eines Neurons mit steigendem Abstand des Eingangs zum Zentrum der Basisfunktion abnimmt, ergibt sich für RBF-Netze insgesamt die Eigenschaft, dass sie für unbekannte Muster, d. h. solche, die der gesamten Trainingsmenge unähnlich sind, eine sehr niedrige Aktivität aller Neuronen aufweisen. Dies steht im Gegensatz zum mehrschichtigen Perzeptron (siehe Seite 69), bei dem die Ausgabe für Muster außerhalb des trainierten Bereichs aufgrund der verwendeten Sigmoidfunktionen zu unvorhersagbaren Ausgaben führen kann.

Die Netzausgabe bei RBF-Netzen ergibt sich aus der gewichteten Summe der Aktivität der einzelnen Neuronen, d. h. $y = f(X) = \sum_{i=1}^K c_i \cdot h_i(\|X - X_i\|)$ mit c_i als Wichtigkeitsfaktoren. Diese können, ausgehend von den N Trainingsmustern sowie der gewählten Stützstellen, über definierte Matrizenoperationen – genauer: die Bestimmung einer (Pseudo)Inversen – berechnet werden. Die Parameter von RBF-Netzen können also aufgrund ihres einfachen Aufbaus komplett berechnet werden und müssen nicht über einen langwierigen iterativen Trainingsprozess bestimmt werden. Werden alle N Trainingsmuster als Stützstellen verwendet (d. h. $K = N$), entspricht dies einer Interpolation durch die so definierten N Funktionswerte, wobei zwischen den Stützstellen starkes Überschwingen auftreten kann. Bei $K \neq N$ bilden RBF-Netze eine Approximation an die gegebenen K Stützstellen mit der Minimierung des Approximationsfehlers. Durch eine weitere Verallgemeinerung kann auch die „Glattheit“ der approximierten Funktion als Kriterium vorgegeben und eingehalten werden. Für Details wird auf die einschlägige Literatur verwiesen (z. B. [Bis95, Bis06, Zel94]).

Entscheidend für die Güte der Abbildung ist bei RBF-Netzen die Wahl der K Zentren der Basisfunktionen. Die Wahl der Stützstellen kann dabei durch eine zufällige Auswahl aus den Trainingsdaten oder durch ein vorgeschaltetes Clustering-Verfahren (z. B. k-Means oder ein LVQ-Modell, siehe nachstehend) erfolgen. Eine iterative Anpassung von nicht-optimal gewählten Zentren sowie der Gewichtungsfaktoren c_i kann im Bedarfsfall anschließend über gradientenbasierte Lernverfahren, z. B. Backpropagation, erfolgen, wie sie auch beim mehrschichtigen Perzeptron zum Einsatz kommen (siehe Seite 69). SCHNERR-HÄSELBARTH

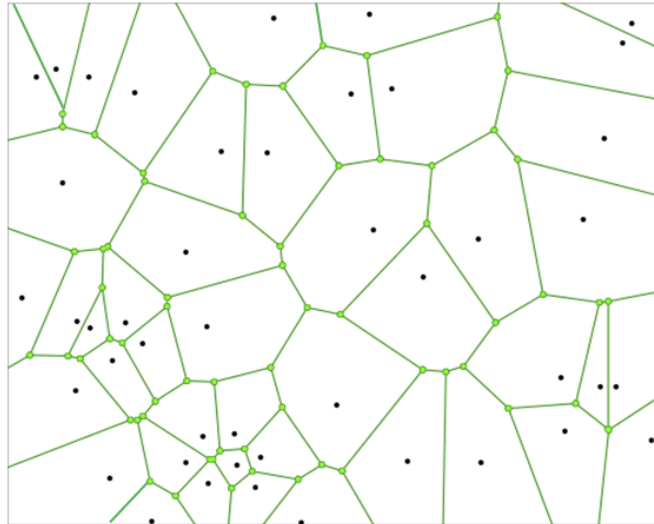


Abbildung 2.9: Voronoi-Diagramm zur Darstellung der Zuordnung bei Nächster-Nachbar-Klassifikation (z. B. LVQ)

verwendet RBF-Netze zur Prognose von attributiven Qualitätsmerkmalen beim Spritzgießen [SH00].

Learning-Vector-Quantization (LVQ): Die LVQ ist ein überwachtes Lernverfahren für attributive Größen, d. h. die Daten werden entsprechenden Merkmalsklassen zugewiesen. Sie wurde von KOHONEN in der Zeit um 1985 entwickelt [Koh01]. Dabei handelt es sich um ein einschichtiges Neuronales Netz, dessen K Neuronen die so genannten Codebook-Vektoren W_i bilden. Die Codebook-Vektoren sind n -dimensionale Vektoren entsprechend der Dimension des Eingangsraums, denen jeweils ein Klassenlabel zugeordnet ist. Die Trainingsmenge für LVQ-Modelle besteht aus N Musterdatensätzen, deren n -dimensionalen Eingangsdaten jeweils ein Label, d. h. eine Zielklasse, zugeordnet ist. Durch entsprechende Trainingsalgorithmen werden die in einer Initialisierungsphase gewählten K Codebook-Vektoren so im Raum angeordnet, dass ihre Verteilung der Dichteverteilung der Eingangsdaten möglichst gut entspricht. Häufungsregionen in den Eingangsdaten, so genannte Cluster, können somit gut ermittelt werden. Während des iterativen Trainingsvorgangs werden anhand eines definierten Abstandsmaßes die dem Trainingsmuster ähnlichsten Codebook-Vektoren gesucht und diese entsprechend verschoben. Abhängig davon, ob die Labels von Codebook-Vektor und Trainingsmuster identisch oder unterschiedlich sind, wird der Codebook-Vektor zum Trainingsmuster hingezogen oder abgestoßen. Mathematisch geschieht dies dadurch, dass ein durch die so genannte Lernrate definierter Anteil α des Differenzvektors zwischen Codebook-Vektor und Muster zum definierenden Vektor W_i des Codebook-Vektors hinzuaddiert oder abgezogen wird. Die Lernrate sinkt dabei in der Regel mit steigendem Lernfortschritt, um eine bessere Konvergenz zu erreichen. Bei LVQ-Modellen ist darauf zu achten, dass das Modell nicht zu lange trainiert wird, da ansonsten

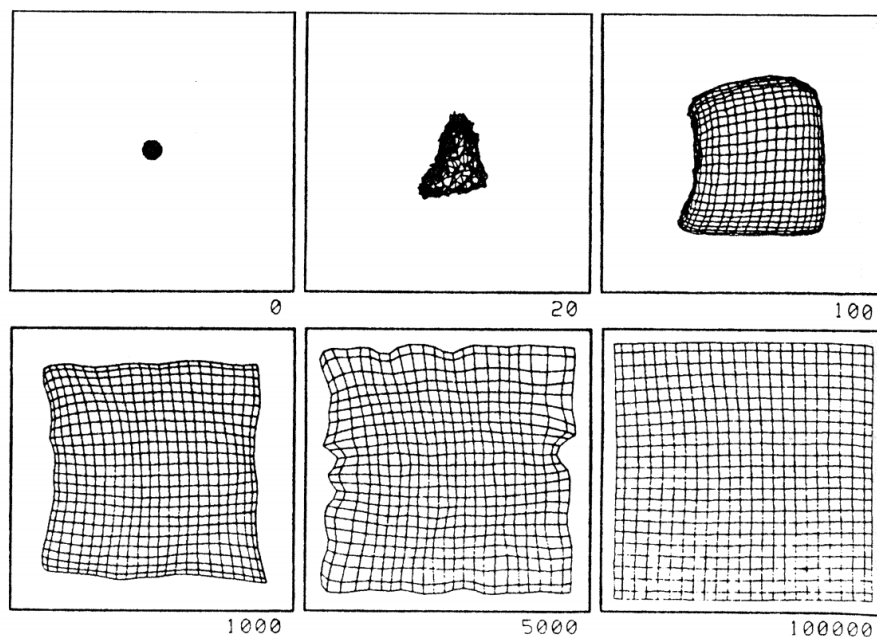


Abbildung 2.10: Entwicklung der Abdeckung eines rechteckigen Eingangsraums durch ein SOM (aus [Koh01])

die Gefahr einer Überanpassung an die Trainingsdaten existiert und dabei die Generalisierungsfähigkeit stark eingeschränkt wird [Zel94].

In der Anwendung wird für ein zu klassifizierendes Testmuster der ähnlichste Codebook-Vektor gesucht, was dem Codebook-Vektor mit dem geringsten Abstand bezüglich des gewählten Abstandsmaßes entspricht. Das Label des so ermittelten Gewinners bestimmt die Ausgabe des Netzes. Es handelt sich bei LVQ somit um einen Nächste-Nachbar-Klassifikator, der den Eingangsraum in entsprechende Klassenregionen unterteilt. Bei der grafischen Darstellung der Zuordnung des Eingangsraums zu den einzelnen Codebook-Vektoren entstehen so genannte Voronoi-Diagramme, die sich stückweise aus den Mittelsenkrechten zwischen zwei Codebook-Vektoren zusammensetzen.

Entsprechende Unstetigkeiten und Fehlklassifikationen insbesondere bei nicht optimal getrennten Klassen im Eingangsraum können wie bei allen Nächste-Nachbar-Klassifikatoren abgemildert werden. Dazu wird nicht nur der nächste Nachbar betrachtet, sondern es werden die nächsten k Nachbarn betrachtet und über entsprechende Regeln aus der „Homogenität“ ihrer Klassenzugehörigkeit Entscheidungen über die Klassifikation des Testmusters abgeleitet.

Self-Organizing Maps (SOM, Kohonen Feature Maps): SOM sind eng verwandt mit den LVQ-Verfahren und werden nach ihrem Erfinder auch oft als Kohonen-Karten oder Kohonen-Feature-Maps bezeichnet [Koh01]. Auch bei SOM wird eine zuvor festgelegte Zahl an Codebook-Neuronen verwendet, die jedoch im Gegensatz zu LVQ durch eine lokale und invariante Nachbarschaftsbeziehung der Neuronen untereinander ergänzt wird. Die Neu-

ronen werden meist in einem rechteckigen oder hexagonalen Gitter angeordnet und es wird eine Nachbarschaftsfunktion definiert, die für jedes Neuron den Grad der Nachbarschaft zu jedem anderen Neuron festlegt. Jedem dieser Neuronen ist ein n -dimensionaler Vektor im Eingangsraum zugeordnet. Beim Trainingvorgang wird analog zu LVQ über eine Nächste-Nachbar-Ermittlung das Neuron gesucht, das dem präsentierten Testmuster am ähnlichsten ist. Im Gegensatz zu LVQ existiert jedoch keine Zuordnung zu Merkmalsklassen, anhand deren eine Anziehungs- oder Abstoßungsreaktion durchgeführt werden kann. Vielmehr wird das Gewinnerneuron sowie dessen Nachbarn entsprechend der durch die Nachbarschaftsfunktion definierten Stärke der Nachbarschaft anteilig in Richtung des angelegten Musters verschoben. Man kann sich eine Kohonenkarte als eine Art elastisches Gitternetz vorstellen, dessen Gewinnerneuron verschoben wird und bei dem die durch „Gummibänder“ mit ihm verbundenen Nachbarn entsprechend mitgezogen werden. Die Konvergenz des entsprechenden Lernverfahrens ist mathematisch nachgewiesen [Koh01]. Die erzielte Abdeckung des Eingangsraums durch die Neuronen der Kohonenkarte ist exemplarisch in Abbildung 2.10 dargestellt. Bei SOM handelt es sich um ein unüberwachtes Lernverfahren. Die Adaption erfolgt rein anhand von Ähnlichkeiten zwischen Neuronen und Eingangsmustern. Die Neuronenschicht wirkt als topologische Merkmalskarte. Ähnlichkeiten und Dichteverteilung der Trainingsdaten werden auf eine SOM abgebildet. Regionen im Eingangsraum, denen verhältnismäßig viele Trainingsmuster zugeordnet sind, haben entsprechend viele Neuronen mit großer Ähnlichkeit, d. h. kleinem Abstand im n -dimensionalen Eingangsraum, zur Folge. Biologisch sind unüberwachte Lernverfahren wie SOM am plausibelsten. Biologische Untersuchungen zeigen, dass sich im Gehirn ähnliche Eingangsreize in der Aktivität von Regionen widerspiegeln, die nahe beieinander liegen, der Musterraum der Eingangsreize also quasi kartiert wird. Solche topologieerhaltenden Karten sind auch im neuronalen Kortex von Säugetieren nachgewiesen [Zel94].

Multi-Layer Perceptron (MLP): Das Modell des Perzeptrons geht auf Arbeiten von ROSENBLATT Anfang der 60er Jahre zurück [Ros58]. Da es nur eine Schicht trainierbarer Gewichte enthält, handelt es sich um ein einstufiges Neuronales Netz. Es kann deshalb nur sehr wenige Funktionen repräsentieren. Die Erweiterung des Perzeptrons um weitere Schichten variabler, d. h. trainierbarer Gewichte führt zum mehrschichtigen Perzeptron (MLP), das (fast) beliebige funktionale, auch nichtlineare Zusammenhänge abbilden kann. Das mehrschichtige Perzeptron ist wohl das am meisten verbreitete Modell von NN bei der Modellierung technischer Prozesse. In sehr vielen Veröffentlichungen ist von NN die Rede, wobei jedoch genau genommen das MLP gemeint ist. Ein MLP besteht aus einer Eingabe- oder Inputschicht von Neuronen, einer oder mehreren versteckten Schichten – so genannte Hidden-Schichten – sowie einer Ausgabe- oder Output-Schicht. Die Anzahl der Eingabe- und Ausgabe-Neuronen bestimmt die Dimension des Eingangs- und Ausgangsraums in der ma-

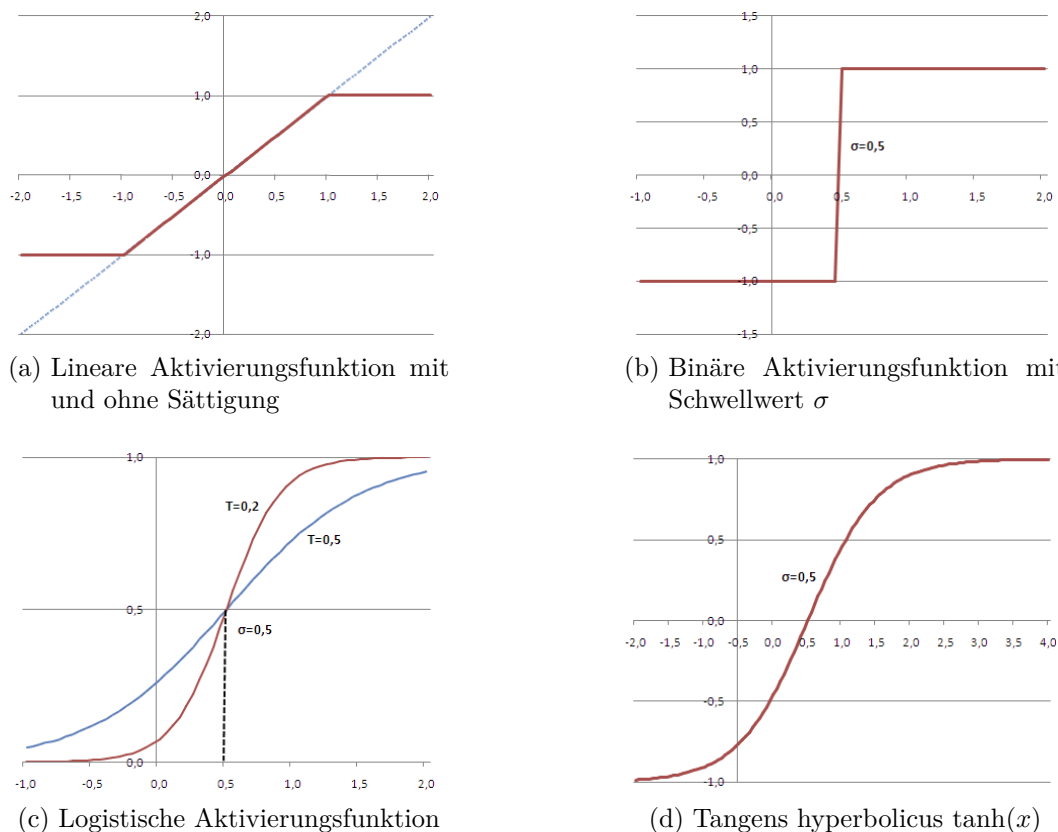


Abbildung 2.11: Verschiedene Aktivierungsfunktionen für Neuronen

thematischen Beschreibung. Die wesentlichen Größen jedes Neurons sind die Netzeingabe, d. h. die Eingangsimpulse in ein Neuron, die Aktivierungsfunktion sowie die Ausgabefunktion. Für Inputneuronen entspricht die Netzeingabe dem – meist normierten – Wert des n -dimensionalen Eingabemusters p an der entsprechenden Komponente (das erste Neuron erhält die erste Komponente des n -dimensionalen Eingabervektors, das zweite Neuron die zweite Komponente usw.), d. h. $net_{pj} = in_{pj}$. Für Neuronen der Hidden- und Ausgabeschicht bildet die Summe der durch die jeweiligen Verbindungsgewichte bewerteten Ausgabe der Vorgängerneuronen die Netzeingabe, d. h. $net_{pj} = \sum_i o_{pi} w_{ij}$ mit o_{pi} als Ausgabe des Neurons i bei dem Muster p und w_{ij} als Verbindungsgewicht von Neuron i zu Neuron j . Die Vorschrift zur Bildung der Netzeingabe wird auch als Propagierungsfunktion bezeichnet. Aus der Netzeingabe eines Neurons wird über die Aktivierungsfunktion die Aktivität des Neurons bestimmt. Ein Auswahl möglicher Aktivierungsfunktionen ist in Abbildung 2.11 dargestellt. Wesentliche Merkmale einer Aktivierungsfunktion sind ihr Schwellwert sowie der (meist) begrenzte Wertebereich. Die wohl am häufigsten verwendete Aktivierungsfunktion ist die logistische mit $f_{log}(x) = \frac{1}{1+e^{-(x-\sigma)/T}}$ mit Schwellwert σ und Steigungsparameter T . Bei einer Netzeingabe im Bereich des Schwellwerts ändert sich die Ausgabe der Aktivierungsfunktion am stärksten, so dass das Neuron auf Änderungen in diesem Bereich am sensibelsten reagiert. Als Ausgabefunktion, welche den an Folgeneuronen weitergeleite-

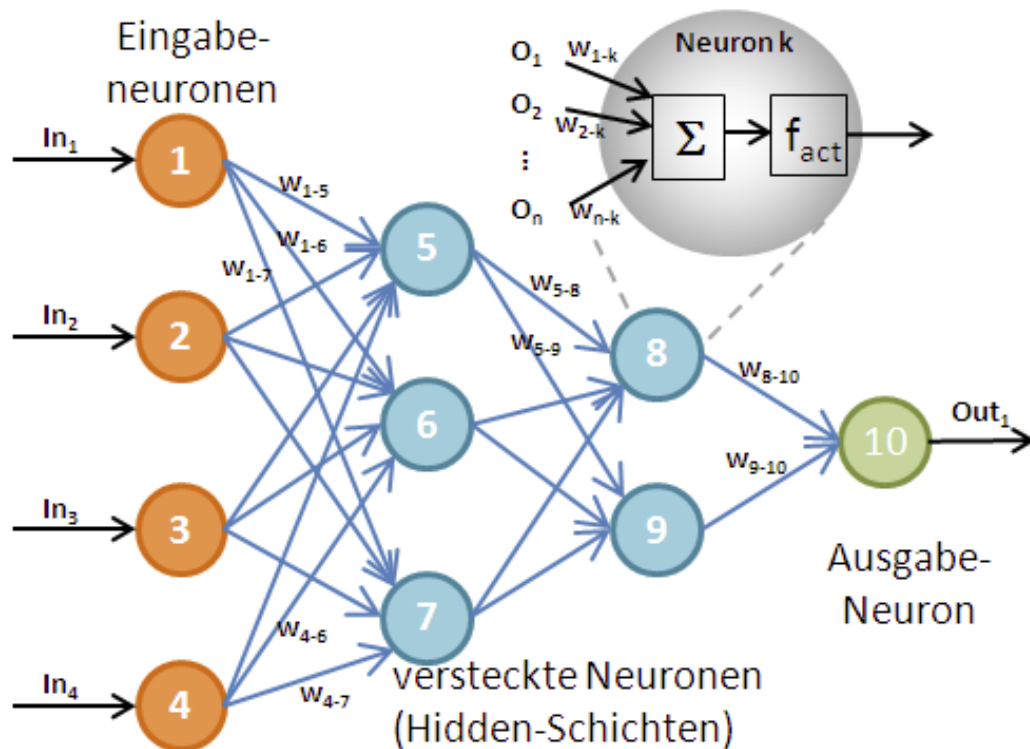


Abbildung 2.12: Aufbau eines mehrschichtigen Perzeptrons mit Feed-Forward-Architektur

ten Ausgabewert eines Neurons in Abhängigkeit der Aktivierung bestimmt, wird meist die Identität benutzt. Auch wenn die beiden letztgenannten Funktionen oft zusammengefasst werden, ist für manche Netzsimulationen mit synchronem Berechnungsmodus die genannte Aufteilung sinnvoll [Zel94]. Im Gegensatz zu LVQ-Netzen, die nur Klassenmerkmale verarbeiten können, ist beim MLP die Prognose kontinuierlicher Merkmale möglich. Der schematische Aufbau eines mehrschichtigen Perzeptrons mit Feed-Forward-Architektur sowie den Signalverarbeitungsbestandteilen in einem Neuron ist in Abbildung 2.12 zu sehen. Während das Netztraining (nachfolgend genauer beschrieben) relativ aufwändig ist, wird in der Anwendung zur Netzprognose nur ein konstanter Zeitaufwand benötigt, der sich insbesondere aus der Anzahl der für die Vorwärtsrechnung benötigten Multiplikationen, Additionen und Funktionsauswertungen der Aktivierungsfunktion zusammensetzt. Auf PCs bedeutet dies einen Aufwand im Millisekundenbereich.

Die „Gedächtnisleistung“ des MLP und sein Wissen ist auf die Verbindungsgewichte verteilt. Eine Zuordnung der „Information“ zu ihrem genauen Speicherort kann deshalb nicht vorgenommen werden, was im Zusammenhang mit NN oftmals auch zur Verwendung des Ausdrucks „Black-Box“ führt. Das Training des MLP erfolgt anhand von Datensätzen, für die zu den Netzeingängen auch die zugehörige Zielgröße bekannt ist – es handelt sich also um ein überwachtes Lernverfahren. Ausgehend von einer Startinitialisierung für die Gewichte und die Schwellwerte der einzelnen Aktivierungsfunktionen erfolgt das Training iterativ. Für einen Datensatz wird durch Vorwärtsrechnen die Ausgabe des Netzes beim

Anlegen der Werte an die Netzeingänge bestimmt. Anhand des berechneten Prognosefehlers, der die Differenz zwischen Netzausgabe und Sollausgabe ist, werden anschließend die Gewichte und Schwellwerte angepasst, um den Fehler zu verringern. Diese Schritte werden für alle Trainingsmuster so lange wiederholt, bis ein zuvor definiertes Abbruchkriterium erfüllt ist. Beim (einschichtigen) Perzeptron erfolgt die Änderung der Gewichte nach der so genannte Deltaregel, die den anteiligen Beitrag eines Gewichts zum Prognosefehler verwendet. Da bei mehrschichtigen Netzen für alle Hidden-Neuronen keine Soll-Ausgabewerte vorgegeben sind, ist die Deltaregel auf Netze mit mehreren Ebenen und semilinearen, d. h. monotonen und differenzierbaren Aktivierungsfunktionen erweitert worden und unter dem Namen Backpropagation bekannt [RM86]. Dabei handelt es sich um ein gradientenbasiertes Lernverfahren, das auf der Rückrechnung des Fehlers basiert und für die Gewichtsänderung die Ableitung der Aktivierungsfunktion des Neurons verwendet. Da für die Sigmoidfunktion die Ableitung als Funktion der Ausgabe des Neurons dargestellt werden kann, vereinfacht und beschleunigt sich die Berechnung, weswegen die Sigmoidfunktion – oder die logistische Funktion als ein Spezialfall – sehr häufig verwendet wird.

Vor dem Durchführen des Netztrainings muss die Netzstruktur – d. h. Anzahl der Schichten, Anzahl der Neuronen in der letzten Schicht, Aktivierungsfunktionen – festgelegt werden. Zudem müssen verschiedene Parameter wie Abbruchkriterium, Gütemaß, Anzahl maximaler Lernepochen, Lernfaktor – und bei Abwandlungen von Backpropagation noch weitere – gewählt werden. Jeder Parameter kann sich dabei je nach Lerndaten auf die Geschwindigkeit und Güte des Netztrainings auswirken, so dass das Training insgesamt viel an Erfahrung und Aufwand bedarf. Neben den vielen Parametern ist ein weiteres Problem beim MLP – wie bei allen NN – das schon in Abschnitt 2.4.2 erwähnte Auswendiglernen. Eine Möglichkeit, einen möglichst optimalen Kompromiss zwischen Lerngüte, d. h. möglichst gutem Abbilden der Lerndaten, und Generalisierungsfähigkeit zu erhalten, ist die Verwendung von Testmustern beim Netztraining. Diese unabhängige Datenmenge wird nicht zum Training des Netzes, d. h. der Anpassung der Gewichte, verwendet. Vielmehr wird in gewissen Abständen der Prognosefehler des bis dahin trainierten Netzes für die Testmuster berechnet. Ein Abnehmen des Fehlers zeugt von einem zunehmenden Lernerfolg. Sobald jedoch der Testfehler wieder zunimmt – bei weiterhin abnehmendem Prognosefehler für die Lernmuster –, ist dies ein Zeichen für ein Auswendiglernen und eine abnehmende Generalisierungsfähigkeit. Da der Backpropagation-Algorithmus ein Gradientenabstiegsverfahren ist, bleibt das Verfahren in der Regel in einem lokalen Fehlerminimum hängen und erreicht das globale Maximum nicht. Um diese und andere Schwächen von Backpropagation (z. B. langsame Konvergenz auf flachen Plateaus, Oszillationen in steilen Tälern, etc.) zumindest abzuschwächen, existiert eine Vielzahl an Modifikationen oder Erweiterungen des Algorithmus wie z. B. Quickprop, Momentum-Term, Flatspot-Elimination, Backpercolation etc. Für eine genauere Beschreibung wird auf einschlägige Literatur (z. B. [Bis95, Zel94]) verwiesen.

Der Backpropagation-Algorithmus wurde für Feed-Forward-Netze entwickelt, d. h. für Netze, deren gerichtete Gewichtsverbindungen immer von Eingangsschicht in Richtung Ausgabeschicht gehen und nie rückwärts. Dies ist der mit Abstand verbreitetste Typ eines mehrschichtigen Perzeptrons. Das Konzept des MLP ermöglicht jedoch auch Kopplungen zwischen Neuronen einer Schicht oder gar Rückkopplungen zu Neuronen einer vorherigen Schicht. Dabei stellen sich jedoch Fragen zur Adaption der Rückkopplungen sowie der Berechnung der Netzprognose durch mehrfache Iterationen und der synchronen Propagierung aller Neuronen. Eng damit ist auch die Stabilität der Netzausgabe und des Netzzustands verbunden. Eine Art von Neuronen mit Rückkopplungen sind die Jordan-Elman-Netze, die im Folgenden kurz erläutert werden.

Jordan-Elman-Netze: Jordan-Elman-Netze sind ein mehrschichtiges Perzeptron mit Rückkopplungen. Als Erweiterung zum „normalen“ MLP werden so genannte Kontextneuronen eingefügt. Diese verfügen nicht über Eingänge von aussen, sondern werden nur über eine Rückkopplung mit den Ausgangsneuronen (Jordan-Netze) oder mit Neuronen der versteckten Schichten (Elman-Netze) sowie eventuell über direkte Rückkopplungen mit sich selbst mit Eingangsreizen versorgt. Diese rückgekoppelten Verbindungen besitzen feste, unveränderliche Werte γ (meist $\gamma = 1$). Dadurch und durch die Tatsache, dass die Aktivität der Kontextneuronen erst nach erfolgter Ausgabe des Netzes aktualisiert wird, werden Probleme mit der Herleitung eines Trainingsalgorithmus sowie der Netzstabilität und Oszillation vermieden. Die Kontextneuronen bilden somit quasi den Zustand und Verlauf der Eingangswerte und des Netzes ab. Das Abklingverhalten und damit das „Vergessen“ kann über die Stärke der Rückkopplungen beeinflusst werden.

SCHNERR-HÄSELBARTH verwendet diese Netzform, um aus dem in mehreren Etappen angelegten Eingangssignal einen Zustand der Kontextneuronen zu generieren, der wiederum zu einer Netzausgabe führt, die den zeitlichen Verlauf inklusive Schwankungen der Signale abbildet. Eine Definition und Extraktion von relevanten Kenngrößen wird somit vermieden. Nachteilig dabei ist insbesondere der hohe Rechenaufwand beim Netztraining von bis zu 30.000 Lernzyklen. Allein für 10.000 Lernzyklen bei einer Abtastrate von nur 10 Hz werden fast 2,5 Stunden benötigt [SH00].

Counterpropagation: Auch wenn Counterpropagation nach einer Abwandlung des Trainingsalgorithmus Backpropagation klingt, handelt es sich um einen eigenen Typ von NN. Ein Counterpropagation-Netz ist eine Kombination von zwei Netzarten (Kohonen und MLP) und hat neben einer Kohonen-Eingabeschicht eine so genannte Grossberg-Schicht, die einem MLP mit einer Schicht entspricht und Neuronen mit sigmoider Aktivierungsfunktion besitzt. Counterpropagation funktioniert als eine Art „look-up table“, die über eine Generalisierungsfähigkeit verfügt. Einsatzgebiet sind vorwiegend Probleme der Musterklas-

sifikation und -erkennung sowie der Mustervervollständigung. Eine interessante Eigenschaft eines vollständigen Counterpropagation-Netzwerks, das als Ein- und Ausgabeschicht jeweils den Summenvektor aus Ein- und Ausgabedaten eines Musters verfügt, ist die Tatsache, dass es neben der Abbildung $f(X) = Y$ zugleich auch die inverse Funktion $f^{-1}(Y) = X$ abbilden kann. Für Details wird auf ZELL verwiesen [Zel94].

Weitere Netzformen: Neben den genannten Arten existiert noch eine Vielzahl von anderen Netzformen wie z. B. Hopfield-Netze, der bidirektionale Assoziativspeicher, Time-Delay-Netze (TDNN), ART-Netze (Adaptive Resonance Theory) oder ALN (Adaptive Logische Netze). Da diese weniger verbreitet und teils auf spezielle Anforderungen oder Anwendungen zugeschnitten sind, wird hier nicht weiter auf sie eingegangen (vgl. hierzu z. B. ZELL [Zel94]).

Support Vector Machines

Support Vector Machines (SVM) sind eine weitere Klasse von statistischen, relativ jungen Klassifikatoren und gehen großteils auf VAPNIC zurück [BGV92, CV95]. Hauptmerkmal der SVM ist, dass sie eine Trennung von Mustern ermöglichen, die zu zwei verschiedenen Klassen gehören, indem sie Muster mittels geeigneter Funktionen – so genannten Kernels – in einen höherdimensionalen Raum transformieren, in dem sie linear separierbar sind, d. h. es existiert eine Hyperebene, welche die beiden Klassen optimal trennt. Diese optimale Trennebene wird mittels der Support-Vektoren als Untermenge der Trainingsdaten beschrieben [CST00]. Der minimale Abstand aus allen Datenvektoren zur Trennebene ist Maß für die Trenngüte, was für andere statistische Modellierungsarten derart nicht existiert. SVM sind die letzten Jahren sehr erfolgreich für Klassifikationsaufgaben auch im Bereich der Bildverarbeitung eingesetzt worden [CHV99, OFG97, PV98]. Das Hauptproblem der SVM ist, geeignete Kernel-Funktionen zur Transformation der Daten in einen höherdimensionalen Raum zu finden, die eine lineare Separation ermöglichen. Des Weiteren sind SVM für die Lösung von Zwei-Klassen-Problemen entwickelt worden. Für die Klassifikation von Mustern mit mehr als zwei Zielklassen wird für jede Klasse ein eigenes Modell generiert, das dieser Klasse als Komplementärmenge die Vereinigung der Daten aller anderen Klassen entgegengesetzt und eine Trennung dieser beiden Mengen ermöglicht. Problem dabei ist die mögliche Mehrdeutigkeit, dass einzelne Testdaten bei mehr als einem Modell die (scheinbare) Zugehörigkeit zu der Ein-Klassen-Menge aufweisen und so Mehrdeutigkeiten auftreten [CCL11, HV10, KU02, RK04, WW99].

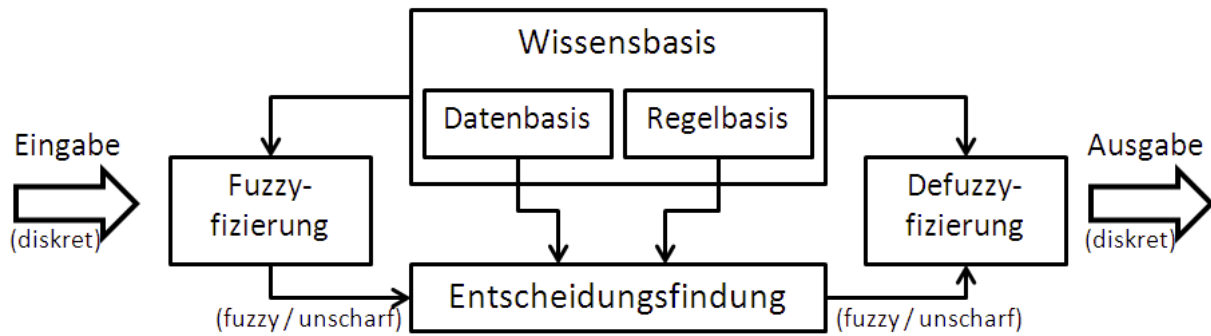


Abbildung 2.13: Aufbau eines Fuzzy-Systems

2.4.3 Fuzzy-Systeme

Fuzzy-Systeme wurden in den siebziger und achtziger Jahren des vorigen Jahrhunderts vorwiegend im akademischen Bereich verwendet, ein größeres öffentliches Interesse kam erst in den neunziger Jahren auf. Mit Fuzzy-Systemen war es möglich, die Erfahrungen und das Know-how eines erfahrenen Maschinenführers mathematisch abzubilden und für Regelungssysteme zugänglich zu machen. Einer aktuellen Studie zufolge basieren etwa 15 % der „intelligenten“ Überwachungssysteme für Maschinen oder Prozesse auf Fuzzy-Systemen [ANS10].

Im Gegensatz zu den datenbasierten Modellierungsmethoden, die ihr Wissen anhand exemplarischer Trainingsdaten generieren, beruhen Fuzzy-Systeme auf vorhandenem Experten-Wissen, das in Form von Regeln abgebildet wird. Der prinzipielle Aufbau eines Fuzzy-Systems ist in Abbildung 2.13 dargestellt. Die Fuzzy-Logik als systemtheoretische Grundlage aller Fuzzy-Systeme geht auf ZADEH zurück [Zad65]. Hauptmerkmal ist der Übergang von scharfen, diskreten Werten – wie sie der booleschen Logik eigen sind – zu so genannten unscharfen Werten (fuzzy: unscharf, undeutlich). Für eine Basisvariable werden verschiedene unscharfe Mengen (fuzzy sets) und entsprechende Zugehörigkeitsfunktionen μ_k definiert, die den Grad der Zugehörigkeit eines Wertes zu den einzelnen Mengen durch einen Wert $[0..1]$ beschreiben. Damit wird es möglich, die linguistische Variable genannten Eingangsgrößen mit sprachlichen Einteilungen zu beschreiben und somit unscharfes, sprachlich formuliertes Wissen und Zusammenhänge abzubilden. Über entsprechende Operatoren ist auch eine Abbildung von Attributen wie „sehr“, „wenig“ oder „stark“ möglich. Die vorhandenen Eingangsvariablen werden über Wenn-Dann-Regeln mit den Ausgangsvariablen verknüpft, wobei in diesen Regeln das sprachlich formulierte und vorhandene Wissen über Zusammenhänge abgelegt ist. Diese Regeln formen die so genannte Regelbasis, die zusammen mit der Datenbasis, welche insbesondere die Operatoren und Zugehörigkeitsfunktionen beinhaltet, den „Wissenspeicher“ bei Fuzzy-Systemen bildet. Mittels der Fuzzy-Inferenz werden dann Eingabewerte auf Ausgabewerte abgebildet. Es existieren verschiedene Fuzzy-Typen, die sich insbesondere in der Konklusion, d. h. der Abbildung

des Dann-Teils der Regeln, unterscheiden. Ein einfaches Beispiel für ein Fuzzy-System ist in Abbildung 2.14 dargestellt, wobei für die Konklusion ein Fuzzy-Inferenz-System nach MAMDANI zum Einsatz kommt.

Zur Bestimmung der Ausgangs- und Zielgröße bei Fuzzy-Systemen sind die folgenden Schritte notwendig:

1. Fuzzifizierung: Zum Übergang der scharfen Eingangswerte zu unscharfen Werten wird für alle linguistischen Variablen x_i anhand der Zugehörigkeitsfunktionen μ_k^i aller k Fuzzy-Sets die Zugehörigkeit des Eingangswerts zu jedem Set bestimmt.
2. Inferenz: Mit den fuzzifizierten Eingangswerten werden die einzelnen Regeln der Regelbasis ausgewertet und ihr Erfülltheitsgrad bestimmt. Die Auswertung erfolgt dabei wie folgt:
 - a) Aggregation: Anhand der Einzelwerte wird für jede Regel der Erfülltheitsgrad der Prämisse, d. h. des WENN-Teils, bestimmt.
 - b) Implikation: Basierend auf dem Erfülltheitsgrad der Prämisse wird der Erfülltheitsgrad der zugehörigen Konklusion, d. h. des DANN-Teils der Regel, ermittelt.
3. Akkumulation: Um die Zielgröße zu ermitteln, wird das Inferenzergebnis der einzelnen Regeln der Wissensbasis zusammengefasst, d. h. die Einzelergebnisse werden vereinigt.
4. Defuzzifizierung: Das unscharfe Inferenzergebnis wird durch die Defuzzifizierung auf einen dezidierten Wert zurückgeführt. Häufige Methoden hierfür sind 'Center of Area' oder 'Mean of Maxima'.

In den einzelnen Schritten finden jeweils Operatoren Anwendung, die je nach Fuzzy-Typ (siehe unten) verschiedentlich ausgeführt sein können. Als Aggregationsoperatoren werden meist das Minimum für „UND“ und das Maximum für „ODER“ verwendet, die Implikation kann über Minimum oder das Produkt und die Akkumulation über das Maximum oder den gewichteten Durchschnitt erfolgen [Alt95, JS95, Zim93]. Die meisten der in der Literatur vorgeschlagenen Fuzzy-Inferenz-Systeme können, je nach Ausprägung der Operatoren und insbesondere des Konklusions-Teils, einem der drei folgenden Typen zugeordnet werden:

- Fuzzy-Typ 1 (TSUKAMOTO [Tsu79]): Die Konklusion verwendet Fuzzy-Sets mit monotoner Zugehörigkeitsfunktion, also beispielsweise auch Sigmoid-Funktionen wie bei MLP. Die Aggregation erfolgt über eine mit den Erfülltheitsgraden gewichtete Mittelwertbildung der einzelnen Aktivierungswerte der ausgangsseitigen Zugehörigkeitsfunktionen. Das Ergebnis ist ein diskreter Wert, so dass eine zeit- und rechenaufwändige Defuzzifizierung entfällt.

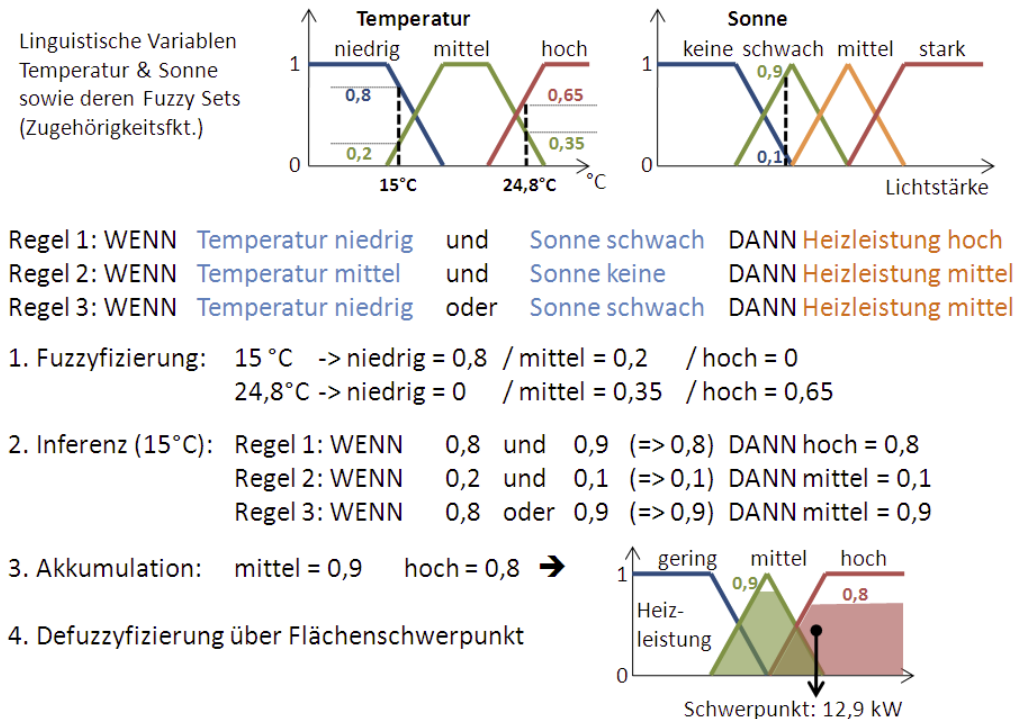


Abbildung 2.14: Vorgehen bei Fuzzy-Systemen (nach MAMDANI) zur Ableitung der Zielgröße

- Fuzzy-Typ 2 (MAMDANI [BH94, JS95]): Die Konklusion einer Regel besteht ebenso aus unscharfen Beschreibungen wie die Prämisse, d. h. die Ausgangsvariablen sind auch Fuzzy-Sets (siehe Beispiel in Abbildung 2.14). Die Akkumulation erfolgt über eine Vereinigung der entsprechend des Erfülltheitsgrades gekappten Ausgangsmengen – meist über den MAX-Operator. Die Ableitung des diskreten Zielwertes erfolgt anschließend über eine Defuzzifizierung, indem z. B. der Schwerpunkt der Vereinigungsfläche ermittelt wird (Center-of-Area-Methode). Aufgrund der notwendigen Defuzzifizierung ist der Rechenaufwand im Vergleich zu den Fuzzy-Typen 1 und 3 deutlich höher.
- Fuzzy-Typ 3 (SUGENO-KANG [SK88, TS85]): Der Ausgabewert einer Regel ist als funktionaler Zusammenhang $z = f(x, y, \dots)$ in Abhängigkeit der linguistischen Eingangsvariablen x, y, \dots definiert, wobei sich die „Stärke“ oder Gewichtung der Regel aus dem Erfülltheitsgrad der Prämisse ergibt. Wie bei TSUKAMOTO-Systemen erfolgt die Aggregation über die Bildung des gewichteten Mittelwerts. Dieser Typ wird manchmal auch als TAKAGI-SUGENO-KANG-Modell (TSK) bezeichnet.

Einen wesentlichen Einfluss auf ein Fuzzy-System haben die Ausprägung und der Typ der einzelnen Fuzzy-Sets der linguistischen Input-Variablen. Wesentliche Merkmale von Fuzzy-Sets sind die Lage des Maximalwertes, ihre Breite sowie ihre Form. Häufige Formen sind

Dreiecke oder Trapeze, aber auch Gaußfunktionen können verwendet werden. Zusammen mit dem verwendeten Fuzzy-Typ und der Komplexität seiner Operatoren bestimmt die Form der Fuzzy-Sets maßgeblich den Aufwand für die Implementierung sowie den notwendigen Rechenzeitbedarf. Werden Fuzzy-Systeme zur Steuerung oder Regelung von Prozessen eingesetzt, kann es für stabile Ergebnisse notwendig sein, die einzelnen Regeln der Wissensbasis mit geeigneten Wichtungsfaktoren zu versehen. Diese werden bei der Akkumulation der Regeln entsprechend berücksichtigt. Für größere Systeme oder Systeme mit vielen Einflussgrößen kann es mitunter sehr aufwändig und schwierig sein, eine entsprechende vollständige Regelbasis zu erstellen.

2.4.4 Kombination von regel- und datenbasierten Methoden

Während datenbasierte Modellierungsmethoden ihr Wissen einzig anhand von Datensätzen erlernen, fließt in Fuzzy-Systeme ausschließlich vorhandenes Wissen in Form von Regeln ein, die bekannte Zusammenhänge beschreiben. Beide Varianten haben Vor- und Nachteile bezüglich der Lern- und Generalisierungsfähigkeit, Extrapolation, Anzahl der Parameter oder dem Einfluss von Rauschen. Mittels der Kombination der Verfahren wird versucht, die Stärken der jeweiligen Methoden zu vereinen und gleichzeitig die Schwächen zu reduzieren. Ergebnis sind die so genannten Neuro-Fuzzy-Modelle, wobei der Unterschied zu „normalen“ Neuronalen Netzen in den Verbindungsgewichten und den Propagierungs- und Aktivierungsfunktionen liegen [NK97]. Viele Neuro-Fuzzy-Modelle basieren auf mehrschichtigen Feed-Forward-Netzen [BK92, BH94, HG94, NK96, TGm95], aber auch andere Netzstrukturen wie SOM als Basis sind möglich [BTP92, Vuo94]. Neuro-Fuzzy-Modelle können demnach verschiedene Ausprägungen aufweisen, wobei im Folgenden auf einige näher eingegangen wird.

Bei JANG und SUN [JS95] wird das Modell eines „adaptiven Netzwerks“ (adaptive network) vorgestellt, welches eine Verallgemeinerung aller Arten von Neuronalen Netzen mit überwachten Lernfähigkeiten darstellt. Sowohl Netzwerke mit Backpropagation-Lernalgorithmus als auch RBF-Netze und Fuzzy-Systeme können als spezielle Formen eines adaptiven Netzwerks betrachtet werden. Das Modell des adaptiven Netzwerks umfasst Knoten sowie gerichtete Verbindungen zwischen einzelnen Knoten. Die Verbindungen geben dabei die Richtung des Datenflusses an, wobei mit den Verbindungen im Allgemeinen jedoch keine Gewichte verknüpft sind. Das Ein-Ausgabeverhalten der Knoten wird durch die Werte aller veränderbaren Parameter, die Teil der Knotenfunktionalität sein können, bestimmt. Insgesamt gibt es Knoten mit und ohne Parameter. Die übliche Darstellung eines „adaptiven Netzwerks“ erfolgt mittels Schichten. Bei der Repräsentation eines Feed-Forward-Netzes gibt es keine Verbindungen zwischen Knoten derselben Schicht. Ausgänge von Knoten sind nur mit Knoten von darauffolgenden Schichten verbunden. Sowohl MLP-Netze als auch RBF-Netze sind Sonderfälle dieser allgemeinen Netzform. Die Parameter werden in lineare

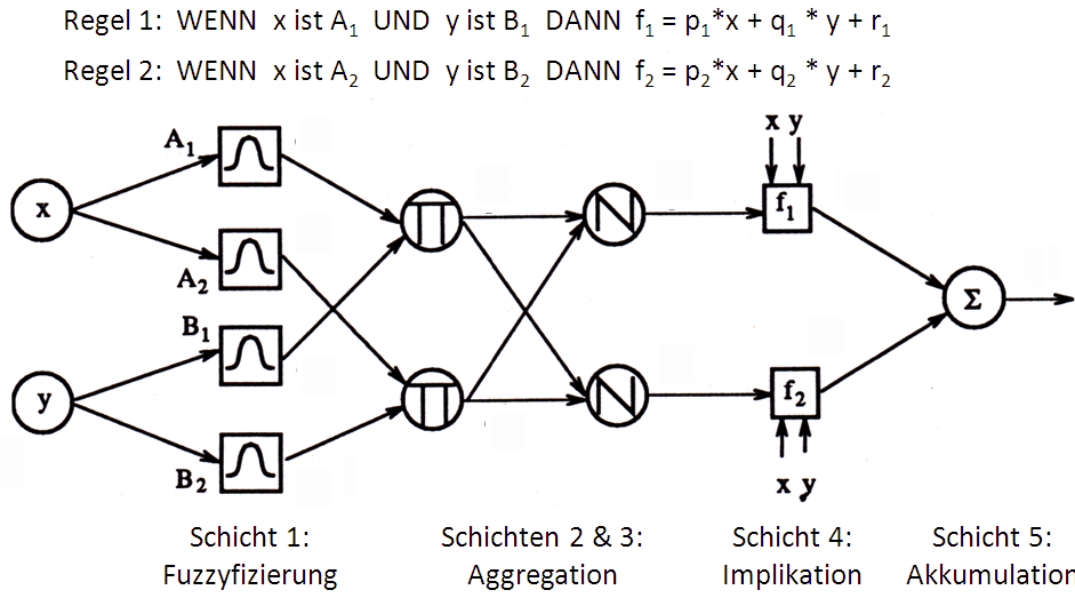


Abbildung 2.15: ANFIS-Modell (Typ 3 für TAKAGI-SUGENO-Fuzzysystem), aus [Zel94]

und nichtlineare aufgeteilt. Die hybriden Lernverfahren sind eine Kombination aus linearen Methoden der Kleinste-Fehler-Quadrate und gradientenbasierten Methoden. Da durch die kleinsten Fehlerquadrate die linearen Parameter auf ein Optimum gesetzt werden, wird der Lösungsraum effektiv reduziert und somit insgesamt eine schnellere Konvergenz beim Training erreicht. Eine solche Kombination kommt auch bei ANFIS-Modellen zum Einsatz.

Adaptive Neuro-Fuzzy Inference System (ANFIS)

Bei ANFIS handelt es sich um die wohl verbreitetste Kombination aus Neuronalen Netzen und Fuzzy-Logik. Die Entwicklung von ANFIS geht auf JANG zurück [JS95, Jan93]. Dabei werden die Vorteile von Fuzzy-Systemen – die einfache Erstellung der Regelbasis – mit der Trainierbarkeit von NN verbunden. Die vorhandene Wissensbasis wird dabei auf eine speziell angepasste vorwärtsgerichtete Netzwerkstruktur mit mehreren Schichten übertragen. Die Fuzzifizierung sowie Prämissen- und Konklusionsteile werden durch verschiedene Schichten abgebildet, wobei es Knoten mit und ohne Parameter gibt. Den verschiedenen Fuzzy-Inferenztypen wird durch verschiedene ANFIS-Typen entsprochen (siehe z. B. [Zel94]). In Abbildung 2.15 ist ein ANFIS-Modell vom Typ 3 als Pendant für ein TSK-Fuzzysystem dargestellt. Die Parameter der Regelbasis des Fuzzysystems, wie z. B. die definierenden Parameter der Zugehörigkeitsfunktionen, werden durch ein hybrides Lernverfahren anhand von gewonnenen Versuchsdaten optimiert. Da sich die Prämissenparameter als Linearkombination der Parameter des Konklusionsteils darstellen lassen, können ihre optimalen Werte über die Methode der kleinsten Fehlerquadrate (least square estimator, LSE) ermittelt werden. Anschließend werden die Konklusionsparameter über ein Gradientenverfahren wie z. B.

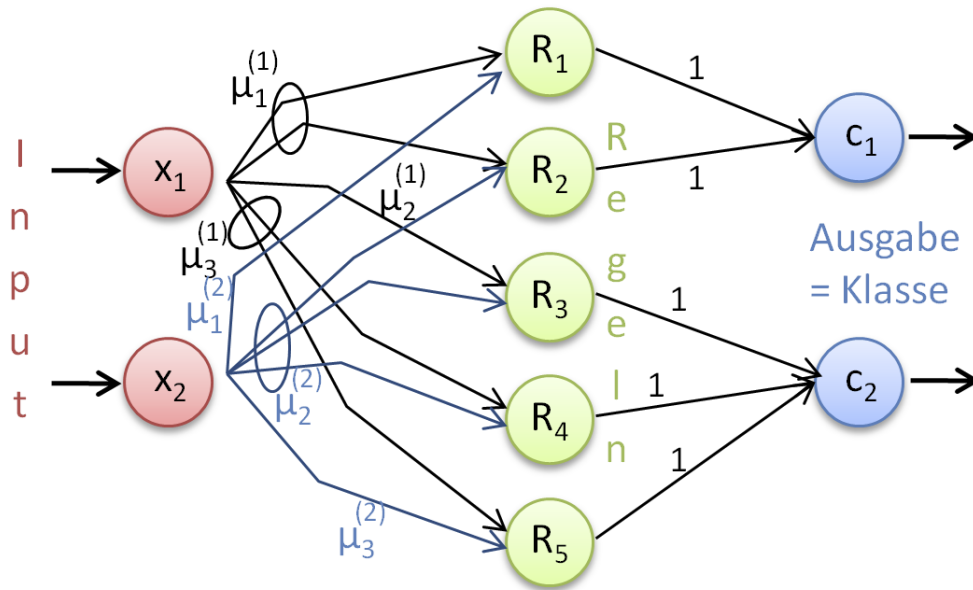


Abbildung 2.16: NEFCLASS-Modell mit zwei Eingängen, fünf Regeln und zwei Klassen

Backpropagation angepasst. Die beiden Schritte des so definierten hybriden Lernverfahrens werden iterativ so lange wiederholt, bis eine hinreichende Modellgüte erreicht ist.

SCHNERR-HÄSELBARTH verwendet ANFIS-Modelle, um die Versuchsplanung bei Spritzgießprozessen zu optimieren. Vorhandenes Expertenwissen wird anhand von Versuchen verfeinert; damit werden die wichtigsten Einstellparameter in Abhängigkeit von der Gewichtung verschiedener Qualitätsmerkmale wie z. B. Einfallstellen abgeleitet [SH00].

In der Literatur sind weitere ähnliche Ansätze bekannt, die eine Verbindung von Fuzzy-Komponenten mit Neuronalen Netzen ermöglichen (z. B. [KKWG97, KK99, MH09, MDP97, MH00, PH00, SYO98]).

Neuro-Fuzzy Classification (NEFCLASS)

Bei dieser Modellart – auch Fuzzy-Perzeptron genannt – handelt es sich um ein gewöhnliches mehrschichtiges Perzeptron, wobei die Gewichte als Fuzzy-Sets modelliert sind. Des Weiteren sind Aktivierungs-, Ausgabe- und Propagierungsfunktionen entsprechend angepasst, um die üblichen Schritte bei der Fuzzy-Inferenz abbilden zu können. Die Absicht dieses Modelltyps ist es, ein Framework für Lernalgorithmen zu bieten, deren Ergebnis in Form linguistischer Variablen interpretierbar ist und das in der Lage ist, a priori vorhandenes Wissen in Form von Regeln als Ausgangspunkt zu verwenden. Insgesamt handelt es sich bei NEFCLASS um eine heuristische Technik zum Ableiten von Fuzzy-Regeln aus Trainingsdaten. Das Anpassen der Regeln erfolgt dabei durch Modifikation der Parameter der jeweiligen Zugehörigkeitsfunktionen [Nau94, NK93, NK96, NK97].

Ein spezielles NEFCLASS-Modell, vorgestellt durch NAUCK und KRUSE [NK95], folgt diesem Schema; es ist ein 3-schichtiges Feed-Forward-Netz, das zur Klassifikation verwen-

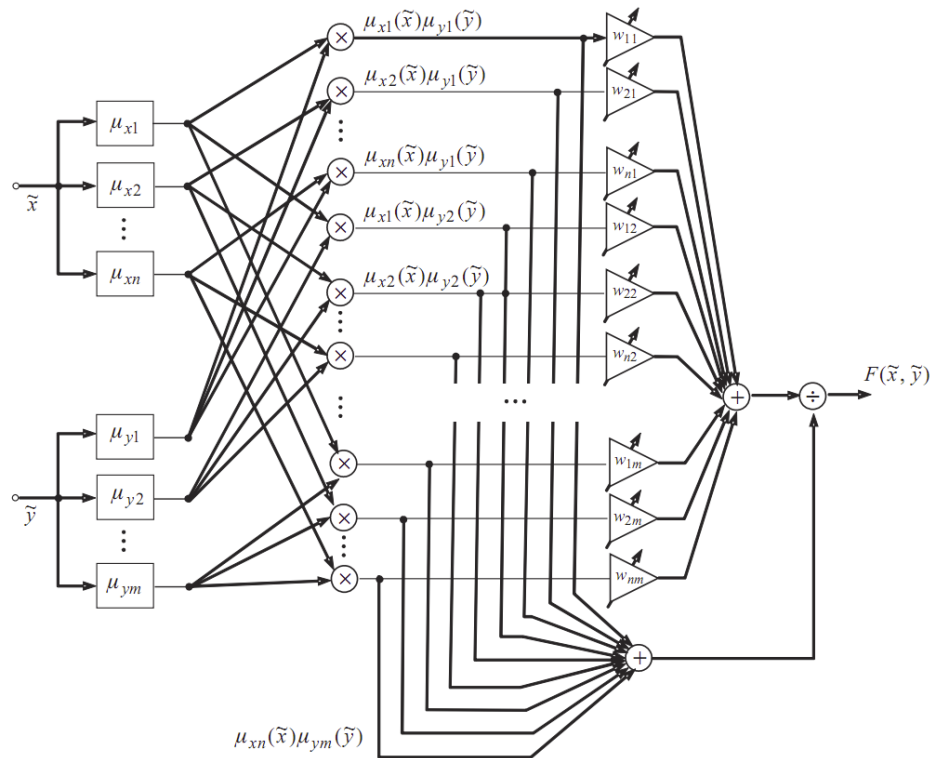


Abbildung 2.17: Interpolierendes Neuro-Fuzzy-Modell (aus [BK08])

det wird. Das heißt, jedem Eingangsmuster wird genau eine von m möglichen Zielklassen zugeordnet. Die Gewichte werden durch Zugehörigkeitsfunktionen gebildet, jedes Eingangsneuron steht für eine linguistische Variable, jedes Neuron der mittleren Schicht für eine Regel der Regelbasis und jedes Ausgangsneuron für eine der Zielklassen. Wesentliches Merkmal sind die so genannten *shared links*, d. h. bei Verwendung desselben Fuzzy-Sets einer Eingangsvariablen in mehreren Regeln ist diesen immer dieselbe Zugehörigkeitsfunktion zugeordnet, wodurch die Konsistenz der Regelbasis gewährleistet wird. Ein NEFCLASS-Modell für zwei linguistische Variablen mit je drei Fuzzy-Sets, mit fünf Regeln sowie zwei Zielklassen ist in Abbildung 2.16 dargestellt.

Das Training des NEFCLASS-Modells erfolgt durch Startwerte für die Zugehörigkeitsfunktionen sowie eine Anzahl k der zu definierenden Regeln. Aus den Trainingsmustern werden dann anhand entsprechender Algorithmen die Regeln abgeleitet, wobei es verschiedene Strategien wie „*best*“ *rule learning* oder „*best per class*“ *rule learning* gibt. Anschließend werden die Parameter der Zugehörigkeitsfunktionen über Lernschritte angepasst – Details sind z. B. bei NAUCK und KRUSE [NK97] zu finden. Auch das NEFCON-Modell – Abkürzung für Neuro-Fuzzy Controller – zur Neuro-Fuzzy-Regelung basiert auf diesem Ansatz.

Interpolierende Neuro-Fuzzy-Modelle

Eine effektive Alternative zu Netzwerken, deren Training anhand der Optimierung von Netzparametern hinsichtlich einer Zielgröße erfolgt (*optimization-based networks*), sind so genannte *memory-based networks*, d. h. Netzwerkmodelle, deren Verhalten sich durch die direkte Verknüpfung von N „Stützpunkten“ ergibt. RBF-Netze können auf diese Weise interpretiert werden. Ein weiterer wichtiger Vertreter dieser Klasse sind *General Regression Neural Networks (GRNN)*, wie sie von SPECHT beschrieben sind [Spe91]. Diesen Modellen liegen als Grundidee so genannte Parzen-Windows und Nadaraya-Watson-Kernel-Schätzer zugrunde [Bis95, HTF01, Par62]. Die Modellanpassung erfolgt in einem einzigen Durchgang – es ist also kein iteratives Verfahren – und besteht aus einer eindeutigen Festlegung von mehrdimensionalen radialsymmetrischen Basisfunktionen, deren Zentralpunkte durch die Trainingsdaten gegeben sind. Sie folgen also dem Prinzip „neurons at the data points“, wobei die Scheitelhöhe der RBF-Funktionen durch die Ausgabewerte des zu beschreibenden Zielsystems an den durch die Trainingsdaten definierten Stützstellen festgelegt sind.

Auf der Basis dieser Architektur wurde von BODYANSKIY und KULISHOVA ein interpolierendes Neuro-Fuzzy-System entwickelt, das für die Interpolation der Reflektionskoeffizienten von Druckertinten verwendet wurde [BK08]. Wesentliche Anforderungen waren dabei die exakte Interpolation an den Stützpunkten sowie eine glatte Interpolationsfunktion. Die Trainingsdaten sind durch die Funktionswerte $f(x_i, y_j)$ eines zweidimensionalen, unregelmäßigen $n \times m$ -Gitters mit den Stützpunkten $\{x_i, y_j\}$ mit $i = 1, 2, \dots, n; j = 1, 2, \dots, m$ gegeben. Das Modell soll diese vorgegebenen Stützpunkte genau wiedergeben. Zudem darf bei der Interpolation zwischen den einzelnen Stützpunkten kein Vorzeichenwechsel in der Ableitung stattfinden, die Funktion muss also glatt sein. RBF-Netze erwiesen sich dafür als ungeeignet, da sie teilweise zu starkem Überschwingen neigen, wenn ihre Ausgabe den Sollwerten der Stützstellen genau entsprechen muss [Kul06]. Die vorgeschlagene Netzwerkstruktur für die beiden reellen Eingangsgrößen x und y ist in Abbildung 2.17 zu sehen. Die Netzeingänge werden entsprechend des den Trainingsdaten zugrunde liegenden Rasters auf $[0..1]$ normiert. Die zweite Netzschicht besteht aus $n + m$ Knoten, welche die Fuzzy-Basis-Funktionen beinhalten und die Fuzzifizierung durchführen. Als Basisfunktionen werden Dreiecksfunktionen mit dem Zentralpunkt an den Rasterpunkten sowie alternativ kubische Splines verwendet, wobei jeweils die Ruspini-Partition erfüllt ist, d. h. für alle Eingangswerte x oder y die Summe aller Erfülltheitsgrade eins ergibt. Die nächste Schicht aus nm -Multiplikationsknoten führt die Aggregation der Erfülltheitsgrade der einzelnen Zugehörigkeitsfunktionen aus. Die nächste Schicht beinhaltet die synaptischen Gewichte, welche während des Trainings adaptiert werden. Sowohl die tatsächlichen Erfülltheitsgrade als auch ihre gewichteten Werte werden aufsummiert und der letzten Schicht zugeführt, die durch eine Normalisierung das Ausgabesignal des Neuro-Fuzzy-Systems berechnet. Diese Architektur deckt sich mit der Architektur eines WANG-MENDEL Neuro-Fuzzy-Systems [WM92] mit einer Rasterung

des Eingangsraums, was wiederum eine Modifikation des TAKAGI-SUGENO Fuzzy-Inferenz-Systems nullter Ordnung darstellt [TS85]. Werden die synaptischen Gewichte den Werten der Zielgröße in den Trainingspunkten gleichgesetzt, kann unter der Voraussetzung der Ruspini-Partition gezeigt werden, dass es sich um ein interpolierendes Modell handelt, das die Trainingspunkte exakt abbildet.

Eine Erweiterung dieses Modellansatzes auf mehr als zwei Eingangsdimensionen und das Fallenlassen der Anforderung, dass die Stütz- und Trainingspunkte auf einem definierten Raster liegen müssen, führt zu den so genannten *General Regression Neuro-Fuzzy Networks (GRNFN)* [BOPT07, BT08]. Als Zugehörigkeitsfunktionen bei der Fuzzifizierung wird dabei die verallgemeinerte Glockenkurve $\mu_{a,b,c}(x) = \left(1 + \left|\frac{x-c}{a}\right|^{2b}\right)^{-1}$ verwendet. Der Parameter c gibt dabei den Zentralpunkt an, während der Parameter a die Breite der Glocke sowie der Quotient $\frac{b}{a}$ die Steigung in den Wendepunkten definiert. Eine detaillierte Analyse der Eigenschaften dieses Modellansatzes, der durch leichte Anpassungen die im Rahmen dieser Arbeit identifizierten Anforderungen sehr gut erfüllt, ist in Kapitel 4 zu finden.

2.5 Exemplarische Beispiele zur Prozessüberwachung

In der Literatur finden sich viele Beispiele für eine Prozessüberwachung, die eine oder mehrere der genannten Methoden in Kombination verwenden. Wesentlich ist dabei jeweils, die Eigenheiten des Prozesses genau zu untersuchen und eine anwendungsspezifische Sensorik, Signalverarbeitung und Modellierung aufzubauen. Im Folgenden werden als kleiner Auszug und Beleg der breiten Palette einige Anwendungen in Kürze dargestellt, wobei sich die Liste beliebig fortführen ließe.

2.5.1 Polierprozesse in der Waferfertigung

Am Fraunhofer IPA wurde von KUMAR ein agentenbasiertes Diagnosesystem entwickelt, um beim chemisch-mechanischen Polieren (CMP-Prozess) von Wafern Defekte frühzeitig erkennen zu können [Kum05]. Der erste wichtige Schritt dabei ist eine genaue Analyse der Abläufe und Zusammenhänge bei den komplexen Polierprozessen sowie den auftretenden Fehlern und ihren Ursachen. Neben vorgelagerten Prozessen wie dem Aufkleben der Wafer auf die beim Polieren benötigten Träger, die jedoch nicht näher betrachtet werden, spielen insbesondere Prozess- und Maschinenparameter der Polierstation eine entscheidende Rolle. Beispielhaft sind dies die Verläufe der Rotationsgeschwindigkeiten von Polier- und Trägerplatten, Anpresskräften und Temperaturen während der mehrminütigen Poliervorgänge sowie chemische und physikalische Eigenschaften des Schleifmittels und der Polierscheiben, die sich mit der Zeit und zunehmenden Anzahl an Polierprozessen verändern. Bei der Waferfertigung ist eine sehr hohe Informationsverkettung vorhanden, was eine lückenlose Pro-

duktverfolgung und damit die Verknüpfung auch mit Prozessparametern von Vorprozessen ermöglicht. Somit können neben den messtechnisch im Sekundenraster erfassten Zustandsdaten (Kraft, Drehzahl, Durchfluss, ...) auch Prüfergebnisse und Qualitätsdaten der vorigen Polierzyklen, die in dem CMP-Prozess nachgeschalteten Prüfschritten für jeden Wafer erfasst werden, in die Bewertung eines Polierdurchgangs mit einfließen. Damit wird auch der Einfluss der fortschreitenden Abnutzung und Veränderung der Polierscheiben und -medien implizit mit erfasst. Um bei den anfallenden großen Datenmengen eine Defektanalyse und Zuordnung der so genannten „Verhaltensmuster“ („behaviorial patterns“) zu Defektklassen zu ermöglichen, setzt KUMAR Genetische Algorithmen in Verbindung mit Agentensystemen ein. Der Aufbau einer geeigneten Regel- und Wissensbasis für ein im Rahmen dieser Arbeit betrachtetes Produkt dauerte ca. ein Jahr. Zeitliche Veränderungen von Prozessparametern während des Prozesses können dabei nur über Regelerweiterungen abgebildet werden, wie z. B. „ODER Erhöhung der Drehzahl um 5 rpm während der letzten 20 Sekunden“. Erkennt das entwickelte System dabei Probleme oder auffällige Trends, kann aufgrund der existierenden Wissensbasis eine Abhilfe vorgeschlagen werden. Ein Eingriff in den Prozess oder die Anpassung der Prozessparameter – Recipe genannt – kann aber ausschließlich durch hochspezialisierte Prozessingenieure und nicht automatisiert erfolgen.

2.5.2 Kunststoffspritzgießen

Ein bedeutsamer Fertigungsbereich, bei welchem modellbasierte Systeme zur Prozessüberwachung und Qualitätssicherung in größerem Umfang entwickelt worden sind, ist das Kunststoffspritzgießen – insbesondere von Thermoplasten. Das prinzipielle Vorgehen bei allen bekannten Arbeiten ist stets ähnlich. Zuerst müssen aus dem Prozess Sensorsignale gewonnen werden, die auftretende Prozessschwankungen und somit qualitätsrelevante Änderungen möglichst gut abbilden. Nachdem sich die ausschließliche Verwendung von Maschinensignalen wie Hydraulikdruck und Schneckenweg bei höheren Anforderungen an die Prognosegüte als zu unsensibel erwiesen hat, werden für leistungsfähige Systeme zusätzlich Werkzeugsignale wie Druck- und Temperaturverlauf im Formnest verwendet. Da diese am Ort der Teileentstehung erfasst werden, besitzen sie einen entsprechend hohen Informationsgehalt über die Vorgänge im formgebenden Werkzeug. Aus den erfassten Signalen werden anschließend Kenngrößen abgeleitet, welche die wesentlichen Prozessabläufe in verdichteter Form wiedergeben und einen Zusammenhang zur Zielgröße – meist Teilgewicht, Geometriemerkmale oder Einfall- und Glanzstellen – aufweisen. Anhand von Trainingsdaten, die durch entsprechende Versuche mit anschließender Ermittlung der resultierenden Teilequalität gewonnen werden, wird ein Modell des Spritzgießprozesses zur Qualitätsüberwachung aufgestellt. Dabei kommen verschiedene Modellierungsmethoden wie Regressionsanalysen, Neuronale Netze oder Support Vector Machines zum Einsatz. SCHNERR-HÄSELBARTH untersucht die Anwendung von Regressionsmodellen, RBF-Netzen, Neuronalen Netzen und

Jordan-Elman-Netzwerken (JE-Netze) für die Qualitätsmodellierung [SH00]. Durch den Einsatz der JE-Netze kann auf eine Kenngrößenextraktion aus den Signalen verzichtet werden, da durch die rekurrenten, d. h. rückgekoppelten Verbindungen der Netze der zeitliche Verlauf der Signale direkt abgebildet wird. Allerdings ist der Trainingsaufwand für die Netze sehr hoch und es werden maximal zwei Signale mit relativ niedriger Anzahl an Datenpunkten (< 200) verwendet. Durch eine Matrizendarstellung wird die Abbildung von prinzipiellem Prozesswissen beim Spritzgießen über den Zusammenhang zwischen Einstellparametern und verschiedenen Qualitätsmerkmalen ermöglicht. Um entsprechend der Wichtigkeit einzelner Merkmale eine automatische Versuchsplanung mit den signifikantesten Prozessparametern zu ermöglichen, wird daraus die Regelbasis für ein Fuzzy-System abgeleitet. Damit die feste Regelbasis anhand von Beispielen und Versuchen prozessspezifisch angepasst und erweitert werden kann, setzt SCHNERR-HÄSELBARTH Neuro-Fuzzy-Systeme in Form von ANFIS ein.

Am Fraunhofer IPA ist die Thematik der Prozessüberwachung beim Kunststoffspritzgießen seit ca. 1992 zuerst in einem Industrieprojekt mit regionalen Firmen aus dem Kunststoffbereich und anschließend in zwei Verbundforschungsprojekten des Landes Baden-Württemberg in enger Kooperation mit den beiden Kunststoffinstituten IKP und IKT der Universität Stuttgart bearbeitet worden [FS96, GHOS95]. Nachdem erfolgreich Prozessmodelle zur Vorhersage der Teilequalität aufgestellt wurden, ist im nächsten Schritt die Nachregelung des Prozesses bearbeitet worden [LRS01, LRSF01, NS07, SN05a]. Bei Abweichungen von der Sollqualität wird über weitere Modellfunktionen die Prozesseinstellung dahingehend abgeändert, dass den Ursachen der Qualitätsabweichung entgegengewirkt wird. Die entwickelte Technologie war beim Einsatz in der Praxis sehr erfolgreich – der Ausschuss in der Serienfertigung konnte auf Null reduziert werden [Neu00]. Großteils basierend auf Ergebnissen aus den Forschungsprojekten, aus denen die NEPRES-Systemtechnik (siehe Abschnitt 2.6) hervorgegangen ist, ist auch die Arbeit von TABATABAI entstanden [Tab01]. Als Teil der gesamten Qualitätssicherung werden Neuronale Netze zur Prozessmodellierung eingesetzt, die thermo-rheologische Kennzahlen aus Düsen- und Werkzeugsensoren als Eingangsgrößen verwenden. Auch an anderen Hochschulen und in anderen Ländern der EU sind Arbeiten zur Prozessüberwachung und Regelung beim Spritzgießen durchgeführt worden [Gie91, Han94, Häu95, Mus00, Rib05, Ujm00, Vac96].

2.5.3 Reaktionsgießen von Polyurethanen

Das Reaktionsgießen von Polyurethanen bietet aufgrund der weitreichenden Möglichkeiten, die Produkteigenschaften durch Variation der Chemierohstoffe gezielt einzustellen, ein nahezu unbegrenztes Einsatzspektrum. Obwohl moderne Verarbeitungsanlagen eine Überwachung sämtlicher Maschineneinstellgrößen erlauben, gibt es keine Möglichkeit, die für die Werkstückqualität entscheidenden chemischen Reaktionen im Werkzeug zu kontrollieren.

Aufgrund der Kopplung der chemischen Reaktion mit dem Formgebungsprozess sind die Einflüsse auf die resultierende Formteilqualität sehr vielschichtig und analytisch kaum zu beschreiben. Aufgrund der Abhängigkeit der Werkstückmerkmale von einer Vielzahl von Einstellparametern können Qualitätsschwankungen nur sehr bedingt auf Schwankungen einzelner Einstellgrößen zurückgeführt werden. Dies macht eine manuelle und somit kostenintensive Qualitätskontrolle der Produkte häufig unumgänglich. HENZE und HABERSTROH haben eine Werkzeugsensorik vorgestellt, die – basierend auf der dielektrischen Spektroskopie – eine Erfassung des Leitwertverlaufs im Werkzeug erlaubt [HH02]. Dieser korreliert gut mit dem Verlauf der chemischen Reaktion. Zusammen mit der Druck- und Temperatursensorik im Werkzeug können so qualitätsrelevante Signalverläufe erfasst und Prozesskenngrößen extrahiert werden, welche die Erstellung eines statistischen Prozessmodells erlauben. Bei der Überwachung des erreichten E-Moduls konnte beim Einsatz der multiplen Regression ein Bestimmtheitsmaß von 95,8 %, mittels Neuronalen Netzen sogar von 98,2 % erzielt werden.

2.5.4 Klassifikation von Texturen

Zur Qualitätssicherung wird in den letzten Jahren verstärkt industrielle Bildverarbeitung eingesetzt. Dabei ist bei der Interpretation der Bilder neben der Bestimmung von Maßen auch die automatisierte Bewertung von Oberflächen bezüglich ihrer Struktur, wie sie sonst durch Prüfpersonal bewerkstelligt wird, eine immer häufiger gestellte Anforderung. PANNEKAMP untersucht Möglichkeiten zur Bewertung von texturierten Oberflächen mittels Bildverarbeitung [Pan05]. Die Lösung der Aufgabe weist dabei dieselben Komponenten auf, wie sie im Wesentlichen bei allen Systemen zur Signalverarbeitung, Klassifikation und Machine Learning zu finden sind: Merkmalsbestimmung und -auswahl sowie anschließende Klassifikation. Die Merkmale müssen dabei einen Zusammenhang zur interessierenden Zielgröße aufweisen, d. h. sie müssen geeignet sein, eine Textur informationstechnisch zu beschreiben und in ihrer Gesamtheit unterschiedliche Texturen unterscheiden zu können. Die bei der Texturanalyse verwendeten Merkmale stammen dabei zum einen aus der Zweipunkt-Statistik, andererseits kommen so genannte lokale binäre Muster zum Einsatz. Bei den Merkmalen sind dabei auch Punkte wie Rotations- und Spiegelinvarianz, Multiskalen und Mehrkanalbilder zu berücksichtigen. Bei der Merkmalsselektion wird aus der Vielzahl an möglichen Merkmalen eine Untermenge ausgewählt, um das für die konkrete Aufgabe gewählte Gütekriterium – z. B. die Unterscheidung zweier Texturen – möglichst optimal zu lösen [Lie05, PSK98, SS88b]. Ein für die Selektion geeignetes Gütekriterium hängt dabei stark mit dem eingesetzten Klassifikator – der den abschließenden Schritt bestimmt – zusammen. Für einen Normalverteilungsklassifikator ist ein Maß, welches die den Daten innewohnende Normalverteilung beschreibt, gut geeignet, während für eine Nächste-Nachbar-Klassifikation (wie z. B. LVQ) ein Überlappingsmaß sinnvoll ist [Kön95, Lam98]. Motiviert wird die Merkmalsselektion zum einen durch die Beschleunigung der Merkmals-

berechnung bei der Anwendung, zum anderen durch eine daraus resultierende Reduktion der Fehlerrate bei der Klassifikation. Dieser scheinbare Widerspruch, dass mehr Merkmale nicht zwangsläufig zu einer niedrigeren Fehlerrate führen, wird dabei in einer Vielzahl von Untersuchungen beschrieben (z. B. [JZ97, OD92, PS98]). Ursächlich für dieses Phänomen sind die Schätzfehler des gebildeten Modells, die bei endlicher Lernstichprobe mit der Anzahl der Merkmale zunehmen [JZ97, PMW00, Pan05, Tru79]. PANNEKAMP verwendet zur Merkmalsselektion hauptsächlich verschiedene Suchstrategien, die sukzessive Merkmale aus der Gesamtmenge selektieren, bis eine zuvor festgelegte Anzahl an Merkmalen oder ein maximal zulässiger Modellfehler erreicht ist. Eine Alternative zur Merkmalsselektion stellt die Merkmalsoptimierung dar. Dabei werden Parameter von Funktionen optimiert, die bei der Bildung von Merkmalen eine Rolle spielen. Bei der Bildverarbeitung sind dies beispielsweise die Parameter von eingesetzten Filtern oder Filterbänken, d. h. die Anzahl der verwendeten Filter oder bei Gaborfiltern die Zentralfrequenzen und Analyserichtungen. Das entwickelte System wird u. a. bei der Herstellung von medizinischen Kollagenschwämmen zur Bewertung der Porenverteilung erfolgreich eingesetzt.

2.5.5 Gesenkschmieden und Stanzprozesse

Möglichkeiten zur Prozessüberwachung beim Gesenkschmieden werden von TERZYK beschrieben [Ter94]. Es werden Auswirkung verschiedener Prozessparameterveränderungen (z. B. Schmiedetemperatur, Rohteilmasse, Werkstücklage, ...) auf Körperschall und mehrdimensionale Kraftmessungen untersucht und bewertet. Basierend darauf wird ein Überwachungssystem vorgeschlagen, das verschiedene Funktionen erfüllen soll: Eine Sofortüberwachung zur Klassifikation nach jedem Fertigungszyklus bezüglich eines „normalen“ Schmiedevorgangs oder einer Störung, eine Trendüberwachung zur längerfristigen Prozesskontrolle sowie eine Hinweisfunktion auf Fehlerursachen. Wesentlich dabei ist die Generierung und Auswahl geeigneter Merkmale aus den Kraftsignalen, die sich auch nach Prozessphasen und Winkelsegmenten einteilen lassen. Abschließend steht die Zusammenführung zu einer Klassifikationsausgabe. SOLTYSIAK entwickelt für den Gesenkschmiedeprozess ein Gesamtsystem, das neben der Visualisierung und der sich daraus ergebenden Interpretierbarkeit des Prozesses auch mehrere Überwachungsmethoden umfasst [Sol95]. Aus erfassten Prozesssignalen, die einen Zusammenhang zur Teilequalität aufweisen müssen, werden Kenngrößen abgeleitet und mittels Regelkartentechnik überwacht. Über wissensbasierte Modelle werden bekannte Zusammenhänge zwischen Qualität, Einstellparametern und Prozesssignalen abgebildet, um bei Störungen Hinweise auf deren Behebung ableiten zu können.

STRACHE betrachtet die Überwachung von Stanzprozessen anhand von Kraft- und Körperschallsignalen [Str00]. Um möglichst viele Informationen über den Prozess erhalten zu können, müssen die Sensoren möglichst nahe am Umformprozess platziert sein. Da eine Integration direkt in Werkzeuge aus Kosten- und Aufwandsgründen meist nicht möglich

ist, wird zuerst eine Multi-Sensorplatte entwickelt, die relativ einfach zwischen Werkzeug und Presse eingebaut werden kann und über Kraft- und Körperschallsensoren verfügt. Anhand der so gewonnenen Signale können verschiedene Störungen des Prozesses – wie beispielsweise verringerte Blechdicken – erkannt werden. Die Überwachung erfolgt nach der Signalauswertung durch Vergleich signifikanter Größen mit zuvor ermittelten Sollwerten. Eine Modellbildung zur Online-Überwachung erfolgt nicht. STRACHE beschreibt auch eine Veränderung des Zeit-Frequenz-Spektrums bei zunehmendem Werkzeugverschleiß. Um aus den Signalen Rückschlüsse auf den Verschleiß durchführen zu können, ist jedoch noch Forschungsbedarf vorhanden.

2.5.6 Fügen von Blechen

RICHTER überwacht den Druckfügeprozess von Blechen anhand der Sensorsignale für Fügekraft, Weg und Körperschall [Ric97]. Durch eine geeignete Verarbeitung der Signale ist es möglich, bei einem ansonsten stabilen Prozess alle wesentlichen Prozessfehler erkennen zu können. Die Modellbildung beschränkt sich im Wesentlichen darauf, zu bestimmten Zeitpunkten die Ist-Werte mit Vorgabewerten zu prüfen, d. h. es wird das aktuelle Sensorsignal anhand von zuvor festgelegten Hüllkurven beurteilt.

Für das Radialclinchen entwickelt BRECKWEG ein auf Neuronalen Netzen basierendes Prozessüberwachungssystem zur Erkennung und Klassifikation von Prozessfehlern [Bre07]. Die Signale von Stempelweg, Fügekraft, Drehmoment und Drehzahl sind dabei wesentliche Prozessgrößen. Aus ihnen werden zuerst eine Reihe von Kenngrößen ermittelt, anschließend signifikante Größen ausgewählt und diese als Eingabewerte für die Neuronalen Netze verwendet. Eine experimentelle Verifikation bestätigt dabei die Eignung dieser Methode zur Prozessüberwachung und zur Fehlerklassifikation.

2.5.7 Weitere Schweißverfahren

Für das Heizelementschweißen erarbeitet NATROP ein Verfahren, um basierend auf mathematisch-physikalischen Prozessbeschreibungen der einzelnen Phasen die Einstellparameter im Hinblick auf eine optimale Festigkeit rechnergestützt zu optimieren [Nat92]. Anhand von gemessenen Größen des Prozesses erstellt er mittels multipler Regression ein Prozessmodell zur Vorhersage des Qualitätsmerkmals Festigkeit. Für die Regression verwendet NATROP dabei auch logarithmische Werte von ermittelten Prozessgrößen, was aufgrund der logarithmischen Zusammenhänge einzelner Größen – wie sie sich aus der mathematisch-physikalischen Beschreibung ergeben – plausibel ist und insgesamt zu guten Prognoseergebnissen führt.

UEBBING untersucht das Vibrationsschweißen anhand einer sechseckigen Halbschale als Probekörper mit verschiedenen Werkstoffen [Ueb95]. Er analysiert die verschiedenen Pro-

zessphasen beim Schweißen und führt nach Versuchen mit Parametervariationen eine Varianzanalyse zur Identifikation der wichtigsten Einflussgrößen durch. Als Qualitätsmerkmale dienen ihm neben den Ergebnissen der Innendruckprüfung auch die Abmessung der geschweißten Teile in Schweißrichtung. Neben der Prozessüberwachung mittels Regelkarten anhand der Schweißzeit untersucht er insbesondere die Möglichkeit der Prozessüberwachung mittels Regressionsmodellen. Dazu verwendet er insgesamt 18 Werte pro Zyklus, die wichtige Größen aus dem Prozess abbilden. Die Güte der Prozessmodelle ist stark schwankend. UEBBING belegt die Wichtigkeit einer reproduzierbaren und exakten Qualitätsprüfung. Festigkeitsaussagen sind anhand von Prozessparametern nur mäßig bestimmt möglich. UEBBING erklärt dies damit, dass sich viele Einflussgrößen, die mit den Maschinenparametern nicht erfasst werden, auf die Prüfung der Kurzzugfestigkeit auswirken. Ergänzend untersucht er die Steigerung der Modellgüte, wenn Daten aus dem Spritzgießen der Halbschalen in die Regressionsmodelle mit einfließen. Für das Geometriemaß bestimmen Kenngrößen aus dem Spritzgießen über 60 % der Modellgüte, wodurch der Einfluss des Vorprozesses belegt ist.

Auch beim Vibrationsschweißen existieren verschiedene Prozessphasen, die einen Zusammenhang zur erreichbaren Schweißnahtfestigkeit haben. GIESE betrachtet die Auswirkungen von Parameteränderungen auf die Prozessphasen beim Vibrationsschweißen, indem er aus dem Prozess Sensorsignale für Schweißweg, Schwingamplitude, Schweißkraft und Reibkraft aufzeichnet und auswertet [Gie95]. Für die Detektion, wann sich der Prozess in der wichtigen dritten, stationären Schweißphase befindet, verwendet er ein Neuronales Netz, das als Eingangssignal den aktuellen Weg sowie den Wegverlauf und dessen Veränderung in jüngster Vergangenheit erhält. Anhand dieses Prozessmodells ist GIESE in der Lage, die Prozessphasen während der Schweißung online zu erkennen und daraus relevante Kenngrößen wie den Eintrittszeitpunkt abzuleiten. Ein Vergleich dieser Kenngrößen mit vorgegebenen Toleranzen erlaubt anschließend eine Beurteilung der Schweißgüte. Des Weiteren weist er nach, dass einige der mittels dem Neuronalen Netz ermittelten Kenngrößen auf Veränderungen im vorgelagerten Spritzprozess, wie eine Absenkung der Werkzeugtemperatur, sensitiv reagieren. Veränderungen im Vorprozess wirken sich im Schweißprozess aus und können teilweise anhand des Schweißverhaltens erkannt werden.

AHREND untersucht die Grenzen und Möglichkeiten von Neuronalen Netzen beim Widerstandspunktschweißen [Ahr04]. Aus dem erfasstem Verlauf von Schweißstrom und -spannung werden einige charakteristische Größen (z. B. Min- und Max-Wert) abgeleitet und als Netzinput verwendet. Untersucht wird die Eignung dieser Modellierungsmethode bei einer Variation der Einstellparameter, geänderten Werkstückeigenschaften (Beschichtung, Dicke, Material) und dem Verschleiß der Schweißzangen. Einige Variationen bei den geschweißten Stahlblechen können dabei sehr gut abgebildet werden, andere gestalten sich in der Erkennung als schwieriger. Die Extrapolationsleistung ist nicht zufriedenstellend. Probleme

ergeben sich auch bei geänderten Materialien, da dem Netz hierüber mangels geeigneter Sensorik Informationen fehlen. Bei Aluminium ist dies aufgrund vieler signifikanter Materialeinflüsse (Alterung, Oxidation u. a.) noch gravierender. Prinzipiell sind Neuronale Netze dazu geeignet, das Widerstandspunktschweißen zu überwachen. Allerdings müssen die Randbedingungen sorgfältig in einem Bereich gehalten werden, der dem Netz bekannt ist. Alternativ können für die verschiedenen Stufen der Einflüsse eigene Prozessmodelle trainiert werden. Ein Allgemeinmodell für jeden Prozess ist nicht möglich.

2.6 NEPRES-Systemtechnologie

Am Fraunhofer IPA wurde unter der Federführung von SCHMIDBERGER im Rahmen mehrerer Verbundforschungsprojekte NEPRES⁵ zur qualitätsbasierten Prozessüberwachung und -regelung entwickelt [GHOS95, LRSF01, SN05b, SNG07, SSP97]. Das Prinzip der Systemtechnik ist in Abbildung 2.18 dargestellt. Während anfangs die Qualitätssicherung von Kunststoffspritzgießprozessen im Vordergrund stand, wurde im Laufe der letzten Jahre diese Systemtechnik auch auf andere Fertigungsbereiche wie das Clinchen von Blechen [BSNS04] und das Ultraschallschweißen ausgeweitet [SF05, NF10].

2.6.1 Aufbau der Systemtechnik

Kerngedanke von NEPRES ist die Modellierung des zu überwachenden Prozesses mittels Neuronaler Netze auf der Basis von Sensorsignalen aus dem Fertigungsprozess. Sehr wichtig ist dabei, dass die Sensoren ihre Informationen möglichst nahe am „Fertigungsort,“ gewinnen, um auch kleine Veränderungen im Prozess erfassen zu können. Hauptaufgabe des eingesetzten Modells ist es, einen Zusammenhang zwischen Schwankungen der Signale und Qualitätsschwankungen des Produktes herzustellen. Die Systemtechnik von NEPRES besteht aus einer Kombination von Hard- und Software-Komponenten.

Die **Hardware** umfasst neben dem Rechnersystem für die Signalverarbeitung auf PC-Basis die gesamte Messkette von den Sensoren über die Messverstärker bis zur Messwertfassungskarte sowie weitere notwendige elektrische und mechanische Kopplungen zu weiteren Komponenten wie z. B. die Ansteuerung einer Sortierweiche oder Schnittstelle zur Maschinensteuerung. Die notwendige Hardware ist immer abhängig vom spezifischen Anwendungsfall und muss entsprechend ausgewählt und aufeinander abgestimmt werden.

Die für NEPRES entwickelte **Software** beinhaltet alle notwendigen Verarbeitungsschritte und -routinen. Basis des NEPRES-Systems bildet die kommerziell erhältliche Messwertverarbeitungsplattform DASYLab [Das]. Die Programmierung und Definition aller notwendigen Verarbeitungsschritte erfolgt grafisch. DASYLab ist modular strukturiert und verfügt

⁵Neuronales Prozess-Regelungs-System

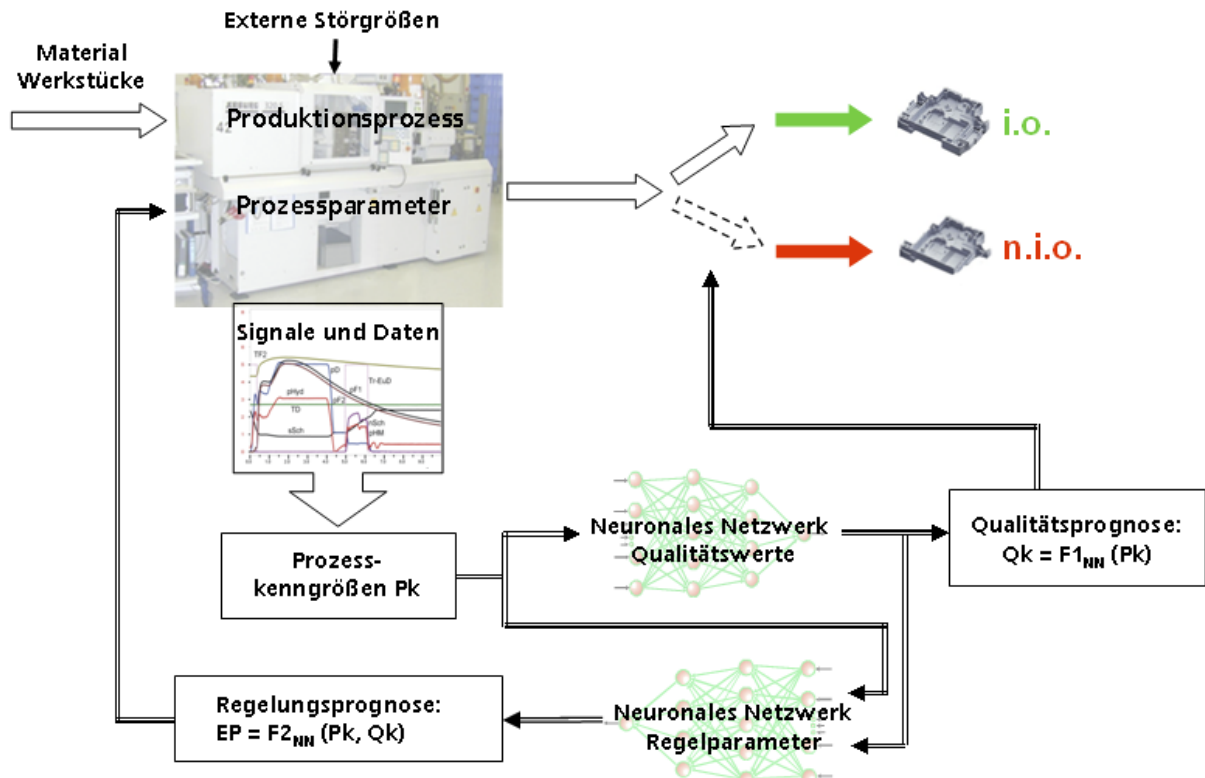


Abbildung 2.18: NEPRES-Prinzip

über einen Satz von Standardmodulen, deren Zusammenspiel über entsprechende Signalverbindungen gesteuert wird. Neben dem Zugriff auf eine Vielzahl von Messwerterfassungskarten sind verschiedene, nach Kategorien sortierte Module u. a. auch für die Visualisierung und den Dateizugriff verfügbar. Die gesamten, am IPA entwickelten NEPRES-Module sind über eine Plug-In-Schnittstelle direkt in DASyLab integriert.

Diese Module lassen sich in verschiedene Kategorien einteilen, die im Folgenden kurz beschrieben sind:

Vorverarbeitung: Zu dieser Gruppe gehören verschiedene Filter (z. B. zur Rauschunterdrückung), die Offsetkompensation und die Skalierung. Damit werden aufgezeichnete Signale vor der Weiterverarbeitung entsprechend aufbereitet, um in den nachfolgenden Modulen belastbare und robuste Größen erhalten zu können. Weitere Module dienen dazu, Signale zu normieren oder nur interessierende Signalbereiche durchzuschleusen.

Ableitung von Kenngrößen: Diese Kategorie beinhaltet Module (siehe Abschnitt 2.6.2), die zur Datenaggregation oder -komprimierung verwendet werden. Die so erhaltenen Kenngrößen dienen insbesondere den nachfolgenden Modellierungswerkzeugen als Eingangsgrößen. Zudem werden durch die Aggregation der Information in einzelnen Kenngrößen zielgerichtete Untersuchungen und Analysen konkreter Fertigungsprozesse deutlich erleichtert.

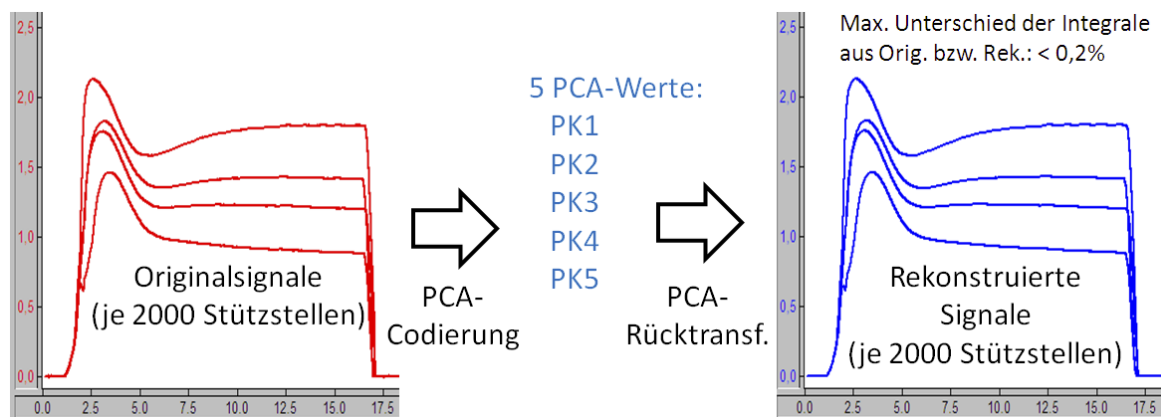


Abbildung 2.19: Originalsignal und Rekonstruktion von Werkzeugdrucksignalen mittels 5 PCA-Kenngrößen

Modellierung: Diese Module beinhalten die Prozessmodelle, welche die Verknüpfung von Prozessdaten und interessierenden Zielgrößen herstellen (siehe Abschnitte 2.4 und 2.6.3). Obwohl in den allermeisten Fällen die Modellierung durch Neuronale Netze erfolgt, die durch ihre Prognosefähigkeiten eine direkte Qualitätsüberwachung ermöglichen, stehen in NEPRES auch andere Methoden wie Grenzwertüberwachung oder adaptive Hüllkurven für eine indirekte Qualitätsüberwachung zur Verfügung.

Generatoren: Dies sind Programmmodule, die zur Erzeugung notwendiger Strukturen verwendet werden. Die wichtigsten beiden dienen zur Erstellung und Verwaltung einer zentralen Konfiguration, welche das Zusammenspiel der verschiedenen Module festlegt. Der Netzgenerator wird zur Erstellung der Prozessmodelle verwendet. Des Weiteren gibt es einen Versuchsplan-Generator, der unter Zugriff auf die Konfiguration einen statistischen, anwendungsspezifisch optimierten Versuchsplan generiert. Dieser kann mit einem entsprechenden Anwendungsprogramm online am Prozess gefahren werden.

Sonstiges: Eine weitere, heterogene Gruppe von Modulen stellt verschiedenste Hilfsfunktionen bereit, z. B. für die Dokumentation, Visualisierung, Datenorganisation, Versuchsdurchführung, Ansteuerung einer Sortierweiche oder zur Datenanalyse.

2.6.2 Kenngrößen

Wesentliche Voraussetzung für die erfolgreiche Anwendung adaptiver Prozessmodelle ist eine entsprechende Komprimierung der relevanten Signalinformation auf eine handhabbare Anzahl von Daten. Dazu werden nach eventuell notwendiger Vorverarbeitung, z. B. Filterung, Kenngrößen abgeleitet. Ziel dieser Prozesskenngrößen (PK) ist es, die Schwankungen der Signale mit wenigen – möglichst aussagekräftigen – Größen zu beschreiben. Diese Kenngrößen sind meist sehr prozessspezifisch und entscheiden maßgeblich über die

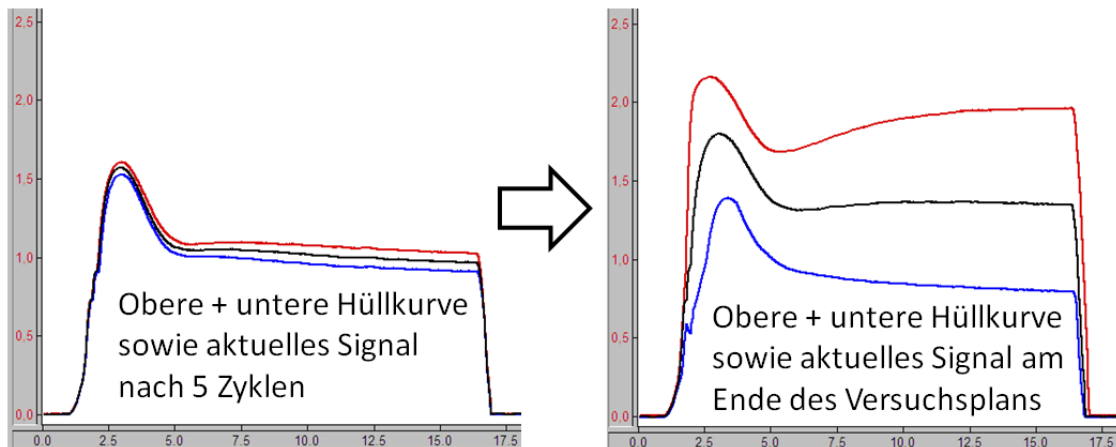


Abbildung 2.20: Adaptive Hüllkurven in NEPRES (z. B. für Werkzeuginnendruck)

Leistungsfähigkeit aller nachfolgend angewandter Methoden zur Prozessmodellierung und -überwachung (siehe Abschnitt 2.4).

Die Kenngrößen lassen sich in verschiedene Arten unterteilen. Diskrete Kenngrößen sind „klassische“ Werte der Signalverarbeitung, wie beispielsweise Integral, Maximum, maximaler Gradient oder zugehörige Zeitpunkte. Neben diesen allgemeinen, auf alle Signale anwendbaren Größen umfasst NEPRES auch viele signalspezifische Kenngrößen. Diese versuchen, die fertigungsprozessspezifischen Gegebenheiten möglichst optimal abzubilden und beinhalten somit einen wesentlichen Teil des vorhandenen Prozess-Know-hows. Die wichtigsten Signale beim Spritzgießen sind neben dem Hydraulikdruck und dem Schneckenweg insbesondere der Druck- und Temperaturverlauf direkt aus der Kavität. In NEPRES existieren entsprechende Kenngrößenmodule für jedes dieser Signale; sie berechnen relevante Größen wie das Einspritzintegral oder die Entformungstemperatur. Die Festlegung, welche Kenngrößen aus welchem Signal berechnet werden, bildet das vorhandene Prozesswissen ab und stellt bei der Anwendung den Schlüssel zum Erfolg dar. Für das Spritzgießen wurden aussagekräftige Größen im Rahmen mehrerer Forschungsprojekte und Dissertationen erarbeitet (u. a. [Gie91, Han94, Mus00, SSP97, SH00, Vac96]).

Eine andere Art von Kenngrößen ist nicht prozessbezogen. Durch diese aus der Mathematik hergeleiteten Kenngrößen wird versucht, Schwankungen und Unterschiede in Signalen möglichst optimal zu beschreiben, ohne einen direkten Bezug zu physikalischen Zusammenhängen herzustellen. Als sehr leistungsfähig haben sich Kenngrößen erwiesen, welche mittels einer PCA⁶-Transformation gewonnen werden [PSK98, PS98]. Dabei wird durch eine Hauptachsentransformation die dominante Signaldynamik ermittelt, welche bei hinreichend ähnlichen und glatten Signalen eine Datenreduktion auf wenige Kenngrößen ermöglicht, ohne wesentliche Prozessinformationen zu verlieren. Beim Spritzgießen können Druck- und Temperatursignale aus dem Werkzeug meist mit fünf PCA-Kenngrößen hinlänglich ge-

⁶Principal Component Analysis

nau beschrieben werden, ohne dass bei der Rekonstruktion des Signals aus diesen fünf PCA-Kenngrößen nennenswerte Verfälschungen auftreten (siehe Abbildung 2.19). Auf den Einsatz der PCA als Projektionsmethode wird in Abschnitt 4.1.1 näher eingegangen.

2.6.3 Modellierung

Für eine einfache signalbasierte Überwachung beinhaltet NEPRES adaptive Hüllkurven, die sich automatisch an die Signaldynamik anpassen (siehe Abbildung 2.20). Diese bilden somit die Streuung des Prozesses ab, haben jedoch keinen direkten Bezug zur produzierten Teilequalität. Bei der direkten Qualitätsüberwachung mittels NEPRES durch den Einsatz von Neuronalen Netzen werden die adaptiven Hüllkurven verwendet, um die Gültigkeit der Modelle abzusichern und stark abweichende Prozesszustände sowie Fehler wie Sensor- oder Kabelbruch zu erkennen.

Zur Prozessmodellierung verwendet NEPRES Neuronale Netze mit Feed-Forward-Architektur, die mittels Backpropagation-Algorithmen trainiert werden. Dabei ist ein wesentlicher Gestaltungsfaktor, dass jedes zu überwachende Merkmal mit einem eigenen Netz modelliert wird, um bessere Modellierungsergebnisse zu erzielen. Neben den Überwachungsnetzen für die Prognose von Qualitätsmerkmalen werden Neuronale Netze auch eingesetzt, um eine Regelung aufzubauen, die qualitätsbasiert die Einstellparameter nachregelt (siehe Abbildung 2.18). Die Daten für die Adaption der Modelle werden mittels Versuchen generiert, die sich aus einer statistischen Versuchsplanung ableiten.

Die mit den Modellen erzielten Prognosen dienen bei Qualitätsmerkmalen direkt der Prozessüberwachung und der Detektion und Ausschleusung von Produkten, welche die geforderte Spezifikation oder Qualität nicht erfüllen. Bei größeren Abweichungen der Qualitätswerte bietet NEPRES somit die Möglichkeit, regelnd in den Prozess einzugreifen – entsprechende Schnittstellen vorausgesetzt [RS01, SN05a, SN05b]. Die Prognose des Maschinenzustands wird dazu verwendet, eine geänderte Einstellung für die Maschinenparameter zu ermitteln, die minimale Abweichungen aller Qualitätsmerkmale von ihrem Sollwert zum Ziel hat. Somit wird durch rechtzeitige Regelungseingriffe Ausschuss effektiv vermieden.

2.7 Fazit

Im heutigen Fertigungsumfeld kommt eine Vielzahl von Prozessen zum Einsatz. Die große Anzahl an Faktoren – wie Funktionsintegration, steigende Produktkomplexität, neue Werkstoffe, billigere Fertigung in Asien – führt dazu, dass die Anforderungen an die Qualität steigen. Dies macht zunehmend den Einsatz von Systemen zur Prozessüberwachung erforderlich. Die komplexen Produkte stellen zudem hohe Werte dar, so dass Ausschuss insbesondere wegen Prüfsystemen erst am Ende der Prozesskette sehr teuer ist. Erfolgt die

Prozessüberwachung dagegen anhand von Sensorsignalen direkt im Fertigungsprozess, können Fertigungsprobleme sehr zeitnah erkannt und entsprechende Maßnahmen eingeleitet werden. Für den Aufbau solcher Prozessüberwachungssysteme stehen heutzutage viele Modellierungsmethoden zur Verfügung, die jeweils ihre Vor- und Nachteile haben. Unterschiedlichste Verfahren werden bereits zur Überwachung verschiedenster Prozesse genutzt.

Das Ultraschallschweißen ist ein Verfahren, das meist am Ende der Prozesskette eingesetzt wird und das hohe Qualitätsanforderungen insbesondere an Dichtigkeit und Festigkeit erfüllen muss. Die Prüfung dieser Eigenschaften ist sehr aufwändig und zudem die Festigkeit nur zerstörend prüfbar. Einer funktionierenden Prozessüberwachung mit der Möglichkeit, Aussagen über die Produktqualität abzuleiten, kommt dementsprechend eine hohe Bedeutung zu. Bei der Verarbeitung von Kunststoffen spielen viele Faktoren, die sich über die gesamte Prozesskette von der Rohstoffherstellung über die Verarbeitung und Lagerung erstrecken, eine wesentliche Rolle. Sie können sich entsprechend auf den Schweißprozess auswirken, was auch häufiger zu Problemen führt, ohne dass deren Ursache klar ersichtlich ist. Der Einfluss der Bedingungen beim Spritzgießen für große Prozessvariationen sowie die prinzipielle Modellbildung zur Prognose der Festigkeit sind für das Ultraschallschweißen nachgewiesen. Die bisher in der Praxis verfügbaren Überwachungssysteme beschränken sich jedoch auf die Grenzwertüberwachung weniger Parameter, deren Zusammenhang mit der Qualität oft fraglich ist. Eine Überwachung der Vorprozesse findet nicht statt, so dass sich Störungen oder geänderte Bedingungen negativ auf den Schweißprozess auswirken, ohne dass die Ursachen bekannt oder ersichtlich sind. Eine umfassende Prozessüberwachung – wie in Abbildung 1.5 dargestellt – mit einer Prognose der Schweißqualität sowie einer Komponente zur Identifikation von unbekanntem Veränderungen, die ihre Ursache im Schweißprozess oder den Vorprozessen haben können, kann hierbei einen wesentlichen Beitrag zur Erhöhung der Prozesssicherheit, der Qualität und der Wirtschaftlichkeit leisten.

3 Detaillierter Lösungsansatz

Für die Prozessüberwachung beim Ultraschallschweißen wird eine Vorgehensweise gewählt, die auch bei anderen Prozessarten zum Erfolg geführt hat (siehe Abschnitt 2.5). Als Basisplattform bietet sich die NEPRES-Systemtechnik (siehe Abschnitt 2.6) an, die für das Kunststoffspritzgießen über alle notwendigen Bausteine verfügt. Im Rahmen dieser Arbeit wird diese entsprechend erweitert, um den in Abbildung 1.5 dargestellten Lösungsansatz umzusetzen. Dies umfasst neben einer Weiterentwicklung der eingesetzten Hardware insbesondere Module zum Berechnen prozessspezifischer Kenngrößen, Module für die Analyse der Kenngrößensignifikanz sowie als zentralen Punkt Werkzeuge zum Training und zur Anwendung der gewählten GRNFN¹-Modelle (siehe Abschnitt 4.2). Das so erarbeitete Gesamtsystem wird im Folgenden als NEPRES-Mini bezeichnet.

3.1 Signaldatenerfassung

Der erste Schritt für den Aufbau einer Prozessüberwachung stellt die Erfassung relevanter Prozessdaten dar.

3.1.1 Signalquellen

Aus dem Prozessablauf und den bekannten Zusammenhängen lassen sich für das Ultraschallschweißen nachfolgend aufgeführte Größen ableiten, die mit unterschiedlichem Aufwand und individueller Genauigkeit messtechnisch erfassbar sind:

- Der **Schweißweg** als Zeitsignal ist eine wichtige Prozessgröße und beinhaltet auch die Fügegeschwindigkeit als wesentliche Kenngröße (siehe Abschnitt 2.1.4 und 2.3.2). Die Erfassung des Wegsignals erfolgt anhand des in der Maschine bereits vorhandenen potentiometrischen Wegmesssystems. Dieses wird als analoges Messsignal auf die Messkarte aufgelegt; dabei sind bei genügend hohem Eingangswiderstand keine Rückwirkungen auf die Maschinensteuerung zu befürchten.
- Die **Leistungsaufnahme des Generators** ist ein Maß für die eingebrachte Energie. Sie kann anhand von Strom- und Spannungsverlauf unter Berücksichtigung des Phasenwinkels ermittelt werden. Bei der eingesetzten Versuchsanlage (siehe Abschnitt 3.2)

¹general regression neuro-fuzzy networks

wird die Leistung steuerungsintern anhand vorliegender Messgrößen digital berechnet und über einen Digital-Analog-Ausgang als ein zur Wirkleistung proportionales Spannungssignal zur Verfügung gestellt.

- In Kombination mit der Leistungsaufnahme ist die **Schwingungsamplitude** der Sonotrode ein wichtiger Indikator für die Wechselwirkungen zwischen Anregung der Maschine und Reaktion des Werkstücks. Die Amplitude kann über den zu ihr proportionalen Strom des Konverters einfach erfasst werden. Generatorintern dient dieser auch als Führungsgröße zur Regelung der in der Steuerung eingestellten Amplitude.
- Signifikanten Einfluss auf den Schweißablauf hat die aufgebrachte **Fügekraft**. Die Schweißkraft beeinflusst maßgeblich die Ankopplungsbedingungen und Einbringung der Schwingungen sowie die Fließvorgänge in der Schweißnaht während des Schweißvorgangs (siehe Abschnitt 2.1.3). Die Kraft ergibt sich aus dem Druck im Pneumatikzylinder der Schweißpresse sowie der wirksamen Zylinderfläche. Die Erfassung als Spannungssignal erfolgt anhand des Stellsignals des eingesetzten Proportionalventils.

Alle diese Größen sind heutzutage bei Standardmaschinen zum Ultraschallschweißen ohne größere Umbauten oder besondere Sensorausrüstung verfügbar. Wenn für die Prozessmodellierung nur aus diesen Signalen abgeleitete Kenngrößen verwendet werden, ist auch ein Einsatz in der heutigen Fertigung ohne hohe Zusatzkosten möglich. Darüber hinaus existieren weitere Größen, bei denen ein wesentlicher Informationsgehalt vermutet wird. Deren Erfassung bedingt jedoch zusätzliche Ausrüstungen und Umbauten an Maschine oder Werkzeug, die teils recht aufwändig und teuer sind. Bei einem regelmäßigen Wechsel der gefertigten Produkte, wie es in der Praxis häufig vorkommt, sind diese Signale als weitere Inputgrößen nicht tragfähig. Für einige Versuchsreihen am Probekörper wurden sie jedoch mit erfasst und ausgewertet:

- Die direkte Erfassung der Schweißkraft kann mittels **Kraftmesselementen** in der Aufspannung der Werkzeugaufnahme erfolgen. Dort wird die tatsächlich wirkende Kraft erfasst, wobei sich auch Schwingungen sowie implizit das Dämpfungsverhalten des Werkstücks abbilden. Die Probleme der starken Dämpfung aufgrund der Kompressibilität der Luft zwischen Vorschubzylinder und Druckregelventil entfallen. Nachteilig ist die aufwändige Integration ins Werkzeug sowie die hohen Kosten für ausreichend steife und genaue Sensoren und Messverstärker.
- Mechanische Schwingungen sind das kennzeichnende Element des gesamten Schweißprozesses. Als Maß für die Schwingungen kann der Körperschall an der Werkzeugaufnahme mit entsprechenden **Schwingungssensoren** erfasst werden. Die Körperschallsignale werden auch durch Änderungen des mechanischen Verlustfaktors beeinflusst, der die Dämpfung der Schwingungen im Werkstoff bedingt. Die Änderung

der Dämpfung ist sowohl durch die Temperaturänderung des Werkstoffs als auch durch die Schmelzebildung, die eine Dämpfungsschicht darstellt, bedingt (siehe Abschnitt 2.1.3). Langjährige Maschineneinrichter sind in der Lage, Veränderungen am Schweißprozess anhand des Klangs der Schweißung einzuschätzen. Diese akustische Informationsquelle nutzen sie dann auch zum Einrichten des Prozesses. Offen ist, inwiefern der Körperschall mit dem Luftschall übereinstimmt, da letzterer auch Schall-schwingungen von Werkstück- und Sonotrodenschwingungen umfasst. Eine Erfassung des Luftschalls mittels Messmikrofonen ist ungeeignet, da sich im Maschinenumfeld zu viele andere Schallquellen befinden, die im Signal nicht separiert werden können. Bei der sensorteknischen Erfassung des Körperschalls sind die Schwingfrequenz von 20 kHz sowie die damit verbundenen sehr hohen Beschleunigungswerte zu beachten. Bei der Auswahl der Sensoren ist deren Frequenzgang inklusive Grenz- und Resonanzfrequenz zu berücksichtigen. Auf dem Markt sind nur wenige Sensoren verfügbar, deren Grenzfrequenz über der Schwingfrequenz des Konverters der Schweißmaschine liegt. Somit ist der Einsatz entsprechender analoger Anti-Aliasing-Tiefpässe im Messverstärker notwendig, um aussagefähige Signale ohne Scheinschwingungen erfassen zu können.

- Eine weitere wichtige Größe ist die Temperatur in der Schweißnaht. Sie kann durch den Einsatz einer **Thermographiekamera** erfasst und zur Prozessanalyse verwendet werden. Aufgrund der sehr schnellen Schweißvorgänge ist jedoch eine entsprechend hochwertige und damit teure Kamera notwendig. Die Kosten sowie die aufwändige Auswertung der aufgezeichneten Bildsequenzen verbieten einen Einsatz in der Serienfertigung. Eine weitere Möglichkeit zur Erfassung der Oberflächentemperatur in der Nahtzone ist der Einsatz von optischen **Infrarot-Temperatursensoren** bei größeren Versuchsreihen. Die Messung ist dabei nur punktuell am fokussierten Messfleck möglich. Außerdem spielen bei der optischen Temperaturmessung materialabhängige Emissionsgrade, Reflexionen sowie optische Abbildungsfehler wie die sphärische und chromatische Abberation eine große Rolle und können sich zu erheblichen Messunsicherheiten addieren. Für den Einsatz solcher Systeme sind auch entsprechende Fachkräfte notwendig, was einem Einsatz in der Serienfertigung entgegensteht. Beide Verfahren zur optischen Temperaturmessung bedingen eine Sichtbarkeit der Schweißnaht. Dieser Punkt ist bei Praxisteilen oft nicht gegeben, da durch entsprechende Gestaltung versucht wird, die Schweißnaht zu verstecken.

Die aufgeführten Signale sind durchweg Analogsignale – meist in Form einer Spannung mit 0-10 Volt –, die mit entsprechend hoher Abtastrate erfasst werden müssen. Zudem sind noch einige digitale Triggersignale erforderlich, die zur Steuerung der Signalaufzeichnung sowie der Zuordnung der Analogsignale zu den verschiedenen Prozessphasen wie Anfahren,

Schweißen oder Halten verwendet werden. Die Bereitstellung von Digitalsignalen ist in der Regel kein Problem, da diese meist durch vorhandene Schaltelemente wie z. B. Taster definiert sind. Allerdings muss bei Digitalsignalen auf den Spannungspegel geachtet werden. Während Messkarten und Laborgeräte in der Regel mit TTL-Pegeln von 5 Volt arbeiten, ist im Maschinen- und Anlagenbau eine Spannung von 24 Volt üblich, so dass entsprechende Spannungsbrücken oder Optokoppler zur Wandlung (und galvanischen Trennung) verwendet werden müssen.

3.1.2 NEPRES-Mini-System

Die bisherige Hardwareplattform von NEPRES basiert auf einem Industrie-PC mit integrierter Messkarte sowie Vorschalt elektronik mit Überspannungsschutz und erwähnter Pegelanpassung der Digitalsignale. Mit diesem System sind Abtastraten von maximal 1 kHz zuverlässig zu erreichen. Das hier neu entwickelte Hardwarekonzept umfasst eine USB-Messkarte mit integriertem digitalen Signalprozessor (DSP) der Fa. Goldammer und sowie eine am Fraunhofer IPA entwickelte Elektronikschaltung mit Optokopplern, galvanischer Trennung und Überspannungsschutz. Diese sind zusammen mit einem Netzteil zur Spannungsversorgung in ein gemeinsames Gehäuse eingebaut. Dieses ist unter einen Laptop montiert, welcher über eine DasyLab-Installation verfügt und der somit als Messrechner dient. Mit diesem System kann eine Summenabtastrate von über 300 kSamples erreicht werden. Bis zu acht Analogsignale können somit mit einer Abtastfrequenz von 45 kHz aufgezeichnet werden. Das komplette Messsystem – im Weiteren auch als **NEPRES-Mini** bezeichnet – ist in Abbildung 3.1 am linken Rand zu sehen. Durch die kompakte Baugröße ist das ganze System mobil und kann ohne großen Aufwand an industrielle Fertigungsprozesse gekoppelt werden.

3.2 Versuchsanlage Probekörper

Als Versuchsanlage zur Durchführung von grundlegenden Versuchsreihen wurde die in Abbildung 3.1 dargestellte Ultraschallschweißanlage „DIALOG digital control“ der Fa. Herrmann Ultraschalltechnik GmbH & Co. KG, Karlsbad, verwendet. Die Maschine verfügt über einen Generator mit 2 kW Leistung und 20 kHz Schweißfrequenz. Die Sonotrode und die Werkstückaufnahme sind auf den in Abschnitt 3.3 dargestellten Probekörper abgestimmt. Die Erfassung der Prozesssignale (siehe Abschnitt 3.1) erfolgt über eine speziell definierte Schnittstelle. Maschinenseitig werden dabei die Signale *Schweißweg*, *Schweißamplitude*, *Leistungsaufnahme* des Generators sowie der *Ist-Druck* am Proportionalventil des Vorschubzylinders als Spannungssignale 0-10 V zur Verfügung gestellt. Zusätzlich liefert die Schnittstelle die Digitalsignale „*Prozess aktiv*“ sowie „*Ultraschall aktiv*“, welche zur

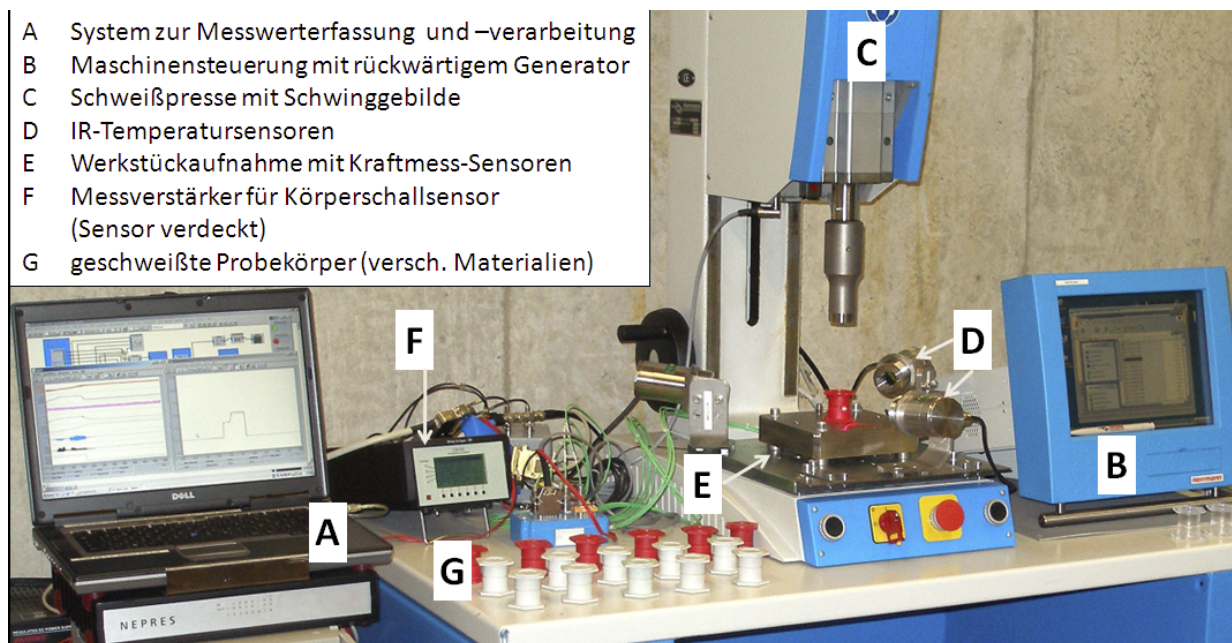
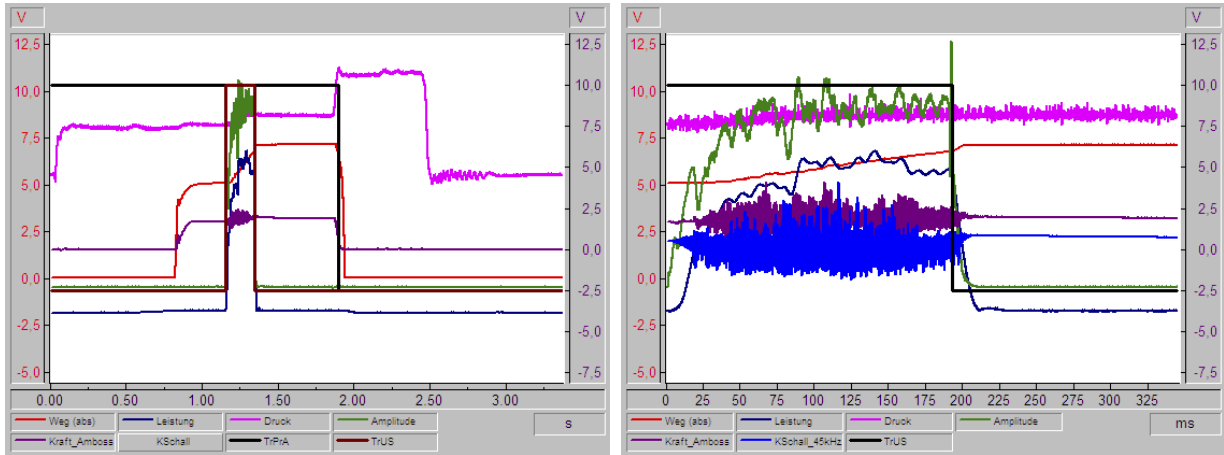


Abbildung 3.1: Versuchsanlage mit erweiterter Sensorik sowie NEPRES-System zur Messwerterfassung

getriggerten Aufzeichnung der Signale verwendet werden, als 0/24 Volt-Signal. Als weiteres Signal wird der entstehende *Körperschall* aufgezeichnet, welcher über einen piezoelektrischen Schwingungsaufnehmer (Hersteller Brüel & Kjær, Deltatron Typ 4394, Grenzfrequenz 25 kHz (Amplitude $\pm 10\%$), Resonanzfrequenz 52 kHz) erfasst und einen Messverstärker mit entsprechendem Anti-Aliasing-Tiefpass auf 0-10 V verstärkt wird. Zur reproduzierbaren und sicheren Positionierung und um ein durch die Schwingungen verursachtes Abheben zu vermeiden, ist der Körperschallsensor (im Bild durch das eingelegte Werkstück verdeckt) mittels Stiftschraube an der Werkstückaufnahme fixiert. Als weitere Sensorsignale stehen noch die real wirkende *Schweißkraft* – erfasst über piezoelektrisch messende Unterlagsmessscheiben der Fa. Kistler zwischen Werkstückaufnahme und Maschinen – sowie bei einigen Versuchen die über Lasersensoren punktuell erfasste Oberflächentemperatur zur Verfügung. Die Versuchsanlage mit zusätzlich installierter Messtechnik sowie dem überarbeiteten NEPRES-System zur Messwerterfassung und -verarbeitung ist in Abbildung 3.1 zu sehen.

Aufgrund der anregenden Schweißfrequenz von 20 kHz muss die Abtastfrequenz des Körperschallsignals nach dem Abtasttheorem mindestens 40 kHz betragen, so dass die Frequenz für die Datenerfassung auf 45 kHz festgelegt wurde. Diese hohen Frequenzen treten nur während des eigentlichen Schweißvorganges auf, wenn die Sonotrode zu Schwingungen angeregt wird. Um die Datenmenge zu reduzieren und die im NEPRES-System verwendete Verarbeitung eines Zyklus als Gesamtblock zu ermöglichen, wurden die Daten mit 45 kHz nur während des aktiven Ultraschalls abgespeichert. Die Signale des Gesamtzyklus inklusive



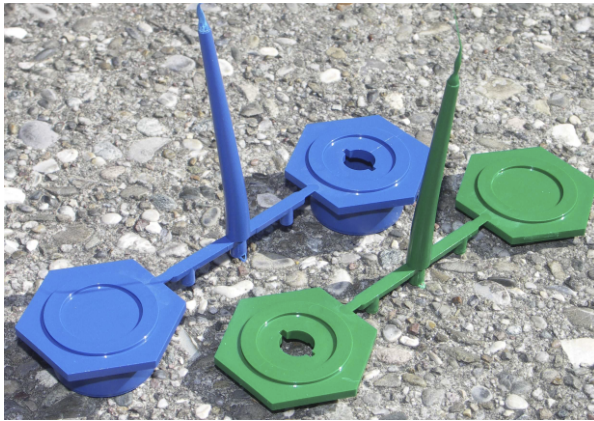
(a) Niederfrequenter Gesamtzyklus (NF) mit 2 kHz (b) Hochfrequente Schweißphase (HF) mit 45 kHz

Abbildung 3.2: Zyklusbezogene Signale eines Schweißzyklus (Probekörper ERG, Material PA66 GF30)

An- und Abfahren der Sonotrode sowie der Haltephase wurden in einem zweiten Verarbeitungszweig auf 2 kHz umgerechnet und separat abgespeichert. Aufgrund der mechanischen Trägheit des Gesamtsystems, der Dämpfungseigenschaften der Luftsäule im Pneumatikzylinder sowie der Tatsache, dass die an der Maschine verfügbaren Signale Leistung und Amplitude über einen D/A-Wandler mit 1 kHz bereitgestellt werden, ist dies ausreichend. Insgesamt ergeben sich für einen Schweißzyklus somit zwei Signaldateien: Eine niederfrequente mit 2 kHz, welche den gesamten Schweißzyklus inklusive Anfahren und Haltephase ab Trigger „Prozess aktiv“ umfasst, sowie eine hochfrequente mit 45 kHz, die den durch den Trigger „Ultraschall aktiv“ definierten eigentlichen Schweißvorgang umfasst. Exemplarisch sind diese beiden Signalblöcke für einen Schweißzyklus in Abbildung 3.2 dargestellt.

3.3 Definition eines Versuchsteils

Für reproduzierbare und belastbare Versuchsdaten, die Grundlage jeder Untersuchung sind, ist ein stabiler Prozess notwendig, bei dem möglichst alle Einflussgrößen kontrolliert werden können. Der Auswahl eines geeigneten Versuchsteils kommt somit grundlegende Bedeutung zu. Aus der Literatur sind verschiedene Geometrien bekannt, die für Untersuchungen des Ultraschallschweißprozesses als Probekörper eingesetzt wurden (siehe Abschnitt 2.3.3). Jedoch haben die verschiedenen Probekörper, wie erwähnt, unterschiedliche Stärken und Schwächen. Im Rahmen der beiden von der Landesstiftung geförderten Verbundprojekte QP-UVS [SF05] und SC-QUPUS [NF10] wurde eine eigene **Probekörpergeometrie** entwickelt, um systematisch Untersuchungen am Schweißprozess sowie zu den Zusammenhängen zwischen Schwankungen im Prozess, der Prozesskette und der Schweißnahtqualität durchführen zu können. Dem Probekörper liegt – ähnlich den Probekörpern von NETZE und des



(a) Gespritzter Probekörper (Fern- und Nahfeld)



(b) Probekörper, verschweißt

Abbildung 3.3: Realisierter Probekörper als Basis für Versuchsreihen

DVS – eine Zylindergeometrie zugrunde, wobei durch die Nahtgestaltung eine Fixierung und Zentrierung erreicht und einem Versatz beim Schweißen entgegengewirkt wird. Weiter ermöglicht der Sechskant eine reproduzierbare Positionierung der Teile. Neben Anforderungen des Schweißprozesses wurden spritzgießtechnische Besonderheiten beachtet. So sind am Bauteil entsprechende Entformungsschrägen vorhanden und zur Reduktion von Einfallstellen wurden Materialanhäufungen vermieden. Die Wanddicken von ca. 2 mm entsprechen praxisüblichen Maßen von Kunststoffteilen. Die Konstruktionen von Werkstück und Spritzgießwerkzeug ermöglichen es, Teile aus verschiedenen thermoplastischen Werkstoffen – auch mit stark unterschiedlichen Verarbeitungsparametern und Schwindungseigenschaften – zu fertigen und zu untersuchen. Die ringförmigen Sonotrodenkontaktflächen sowie die Wanddicken ohne Sprünge gewährleisten beim Schweißen eine optimale Schalleinleitung sowie eine gleichmäßige und ungestörte Schalleitung im Inneren des Werkstückes. Aufgrund der Symmetrie der Zylinderform sind über den gesamten Umfang gleiche Amplituden zu erwarten, was zu reproduzierbaren Schweißungen führt. Der Probekörper existiert in zwei Varianten für Nah- und Fernfeldschweißung (siehe Abbildung 3.3), die sich lediglich in der Länge des Zylinderkörpers unterscheiden. Durch Verwendung von Wechseleinsätzen können zudem verschiedene Nahtgeometrien gefertigt und untersucht werden.

Gegenüber bisher verfügbaren Probekörpern sind beim **Spritzgießwerkzeug** des dargestellten Probekörpers Druck- und Temperatursensoren integriert. Somit können die Bedingungen beim Spritzgießen sowie Schwankungen während der Fertigung der Probekörper mit dem am IPA entwickelten NEPRES-System (siehe Abschnitt 2.6) dokumentiert und bei Bedarf in die Auswertung der Schweißungen einbezogen werden. Da dieselbe Software auch als Basis für die Aufzeichnung beim Ultraschallschweißen dient, sind die Daten von beiden Prozessen im selben Datenformat gespeichert und können leicht zusammengeführt werden. Damit kann die gesamte Prozesskette inklusive der Einflüsse des Spritzgießens auf

die Schweißung untersucht werden. Notwendige Voraussetzung hierfür ist die sichere und eindeutige Zuordnung der Daten vom Spritzgießen und Ultraschallschweißen sowie der Ergebnisse der Qualitätsprüfung zu den einzelnen Teilen.

Die Möglichkeiten zur **Qualitätsprüfung** wurden im Vergleich zu existierenden Probekörper deutlich erweitert und vereinfacht. Die bei der ERG-Geometrie rundum freiliegende Schweißnaht ermöglicht eine einfache optische Beurteilung auf Gleichmäßigkeit und Schweißaustrieb. Auch die Erfassung der Temperatur während des Schweißvorgangs mittels Thermographiekamera oder optischen IR-Sensoren wird dadurch ermöglicht. Der auf der Unterseite des Probekörpers vorhandene Bajonettanschluss ermöglicht den einfachen und schnellen Anschluss eines Adapters zur Dichtigkeits- und Berstprüfung. Die Prüfeinrichtung zur Berstdruckprüfung besteht im Wesentlichen aus einem wassergefüllten Tank, in welchem sich am Boden der Druckanschluss inklusive Druckmesszelle befindet. In den eingesetzten und zuvor mit Wasser gefüllten Prüfkörper wird mittels eines Kolbens Wasser eingepresst. Der Kolben ist durch einen Elektromotor angetrieben, der zur Erzeugung des notwendigen Moments über ein Getriebe angeschlossen ist. Durch die Prüfung mit und unter Wasser sind zum Erreichen der Berstdrücke von teils über 30 bar nur geringe Volumina notwendig. Durch die geringe Kompressibilität wird zudem die Gefahr eines „Explodierens“ des Teils im Gegensatz zu Luft als Druckmedium minimiert. Der Berstdruck wird durch einen Druckverstärker mit integriertem Maximumspeicher sicher ermittelt. Als weiteres Konstruktionsmerkmal verfügen Ober- und Unterteil des Probekörpers über eine Sechskantgeometrie. Diese ermöglicht das Einspannen der Teile während der Berstprüfung und bei der Zug- oder Torsionsprüfung. Für die Zugprüfung wird der Sechskant in eine entsprechende Aufnahme eingeschoben, die mittels eines Kardangelenks in der Zugprüfmaschine eingebracht ist – ein aufwändiges Einspannen entfällt. Das Kardangelenk minimiert die Einbringung von Querkräften und die daraus resultierenden mehrachsigen Spannungszustände. Weitere Vorteile des Sechskants sind die einfach zu realisierende Führung der Probekörper sowie eine Sicherung gegen das Verdrehen von Ober- zu Unterteil beim Schweißen. Die Prüfeinrichtungen für die verschiedenen durchführbaren Festigkeitsprüfungen sind in Abbildung 3.4 dargestellt.

3.4 Versuchsplan und Durchführung

Für die Untersuchung verschiedener Einflussfaktoren sowie die Erarbeitung von Prozessmodellen ist die Durchführung von Versuchen und deren anschließende Auswertung unerlässlich. Für belastbare Aussagen und die klare Zuordnung von Auswirkungen zu Einflussgrößen ist eine sorgfältige Planung der Versuche notwendig. Erster Schritt bei der Versuchsplanung ist dabei die Festlegung, welche Einflussgrößen untersucht werden sollen. Um den Umfang einer Versuchsreihe in vertretbarem Rahmen zu halten, sollten nicht zu

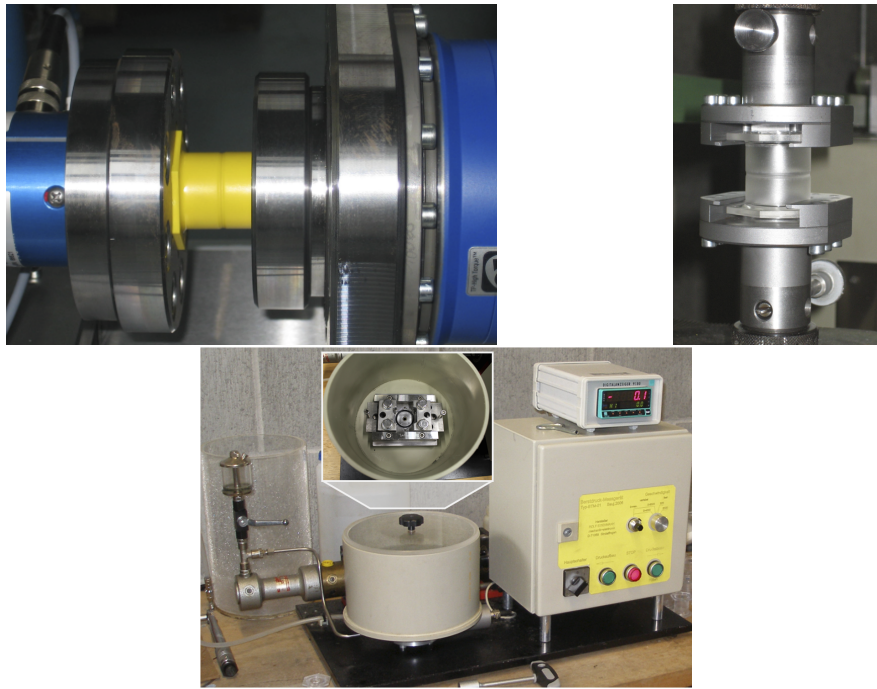


Abbildung 3.4: Prüfeinrichtungen für die Torsions-, Zug- und Berstdruckprüfung

viele Größen gleichzeitig untersucht und variiert werden. Auch die Anzahl der Stufen pro Einstellung spielt eine Rolle. Wenn neben den Haupteffekten, d. h. den unmittelbaren Einflüssen einzelner Parameter auf die Zielgröße, auch Nebeneffekte und Wechselwirkungen, also Einflüsse eines Parameters in Abhängigkeit anderer Einstellgrößen, zu analysieren sind, sollten die Versuche mittels statistischer Versuchsplanung geplant und ausgewertet werden [Kle03, May97, Mon09, Sch97]. Ihr Aufbau ermöglicht es, auch Aussagen über Effekte und Faktorstufenkombinationen abzuleiten, die nicht explizit im Versuchsplan enthalten sind. Ohne entsprechende statistische Versuchspläne explodiert die Anzahl der notwendigen Versuche schon bei geringer Anzahl von Einflussgrößen und Stufen, da diese exponentiell in die Anzahl der Versuchsschritte eingehen. Für die Wahl der Einflussgrößen ist es sehr hilfreich, vorhandenes Prozesswissen und -verständnis heranzuziehen und damit die Auswahl der Größen stark einzuschränken. So ist beim Ultraschallschweißen die Variation der Haltezeit nicht zielführend. Für einen stabilen Prozess ist eine bestimmte Haltedauer bis zur Erstarrung der Schmelze notwendig. Eine Verlängerung der Haltezeit hat keinen weiteren Einfluss auf die Naht, so dass eine Variation nicht sinnvoll und die Aufnahme dieser Größe in den Versuchsplan unnötig ist. In den allermeisten Fällen sind solche Parameter wichtige Größen, die in der Maschinensteuerung eingestellt und von erfahrenem Maschinenpersonal als „Stellglieder“ zum Einrichten des Prozesses oder bei der Nachjustage bei Problemen verwendet werden. Diese sind hier insbesondere der Schweißweg, die Schweißkraft sowie die Höhe der Amplitude. Für das Spritzgießen sind die Temperaturführung und die Nachdruckhöhe sowie für die Lagerung die Umgebungsbedingungen und Dauer maßgeblich. Wichtig

ist auch, dass die gewählten Größen gezielt und reproduzierbar beeinflusst werden können. Theroetisch können auch Größen wie Verschleißzustand von Werkzeugen oder die Molmen-
genverteilung von Kunststoffen in den Versuchsplan mit aufgenommen werden. Da diese jedoch nicht oder nur sehr aufwändig beeinflusst werden können, ist eine dementsprechende Durchführung der Versuche nur eingeschränkt möglich.

Ebenso wichtig wie die Wahl der Einflussgrößen ist die Festlegung der Zielgröße, d. h. des zu untersuchenden Qualitätsmerkmals. Dieses sollte mess- oder prüftechnisch möglichst genau und reproduzierbar zu erfassen sein und bei variiertem Prozess auch gewisse Streubreiten aufweisen. Sind in den Merkmalsausprägungen der Zielgröße keine Variationen zu erfassen, sind eine sinnvolle Auswertung und Aussagen zu diesem Merkmal nicht möglich. So wurden bei Versuchsreihen, deren Arbeitspunkt hinsichtlich der erzielbaren Festigkeitswerte optimiert war, in der Regel alle Teile als dicht geprüft. Undichte Teile waren entweder auf Leckagen der Anlagendichtung des Prüfsystems oder auf sonstige Prozessfehler wie schief eingelegte Teile zurückzuführen. Eine Aussage zum Zusammenhang zwischen Prozessschwankungen und Dichtigkeit ist damit nicht möglich, weshalb auf die Dichtigkeit im Folgenden nicht weiter eingegangen wird. Die Aussagen wurden durch Erfahrungen bei Versuchen an Praxisbauteilen bestätigt. Auch die Angemessenheit des Aufwands für die Prüfung der Versuchsteile sollte bei der Versuchsplanung berücksichtigt werden. Da umfangreiche Versuchsreihen schnell deutlich mehr als 100 Teile umfassen, sind die erforderlichen Zeiträume andernfalls nicht mehr darstellbar. Zudem muss darauf geachtet werden, dass insbesondere bei Kunststoffen allein die Lagerung zu Veränderungen im Werkstoff führen kann. Bei hygroskopischen – d. h. Feuchtigkeit aufnehmenden – Werkstoffen ist dies besonders zu berücksichtigen, da sich der Feuchtegehalt direkt auf Werkstoffeigenschaften auswirkt (siehe auch Abbildung 2.3). Auch bei teilkristallinen Werkstoffen kann es bei längerer Lagerung zu Veränderungen in Form des Abbaus von inneren Spannungen kommen. Somit besteht die Gefahr, dass dadurch beeinflusste Prüfergebnisse den ursprünglich im Versuch variierten Einflussgrößen zugeschrieben werden, obwohl sie ihre Ursache in Werkstoffveränderungen während der Lagerzeit zwischen Schweißen und Prüfen haben.

Für verlässliche und belastbare Aussagen ist eine gewissenhafte und reproduzierbare Durchführung der Versuche notwendig. Mögliche Störgrößen müssen so weit wie möglich ausgeschaltet oder zumindest deren Auswirkung minimiert werden. Für die Auswertung ist eine detaillierte und gewissenhafte Dokumentation notwendig und eine gesicherte Zuordnung der Versuchsteile zu den aufgezeichneten Signalen und Prüfergebnissen wichtig. Um die Ergebnisse statistisch absichern zu können und Aussagen zum Streuverhalten und damit zur Robustheit treffen zu können, sind pro Einstellung des Versuchsplans mehrere Teile zu fertigen. Sofern es möglich und mit vertretbarem Aufwand realisierbar ist, sollte auch die Reihenfolge der Versuchszyklen zufällig ausgewählt werden, um systematische Fehler ausschließen oder erkennen zu können.

Im Rahmen der vorliegenden Arbeit sind für das Ultraschallschweißen je nach Zielsetzung der Analysen unterschiedliche Parameter variiert worden. Für die Qualitätsprognose beim Schweißen sind insbesondere Schweißkraft, Schweißweg und Amplitude einflussreiche Parameter (siehe Abschnitt 2.1.3), so dass diese bei den Versuchen gezielt verändert wurden. Zur Entwicklung von Prozessmodellen zur Störungserkennung in den Vorprozessen müssen die Versuchspläne Parameter aufweisen, die für auftretende Störungen ursächlich sind. Dies beinhaltet die Lagerung mit unterschiedlicher Dauer, Temperatur und Feuchtigkeit, was sich in geänderten Werkstoffeigenschaften auswirkt. Darüber hinaus wurden beim Spritzgießen Parameter variiert sowie Materialverwechslungen untersucht, indem andere Materialchargen oder andere Materialtypen verwendet wurden. Solche Eigenschaften sind anhand gespritzter Werkstücke ohne Laboruntersuchungen meist nicht zu erkennen. Als Zielgröße wurde insbesondere die erzielte Festigkeit betrachtet, da diese – im Gegensatz zur Dichtigkeit – deutliche Streuungen aufweist und entsprechende Prüfmethode verfügbar sind. Für die Bestimmung der Festigkeit kommen sowohl die Berstdruck- als auch die Zugprüfung zum Einsatz. Da die Prüfmethode hierfür zerstörend sind, ist eine Wiederholung der Messung oder der Vergleich beider Werte für ein und dasselbe Teil nicht möglich. Ein wichtiger Punkt ist die Streuung der Prüfmethode, welche die Streuung der Festigkeit der Schweißnähte überlagert.

Alle Variationen führten insgesamt zur Durchführung einiger Dutzend Versuchsreihen im Rahmen dieser Arbeit. Aus Gründen der Übersichtlichkeit werden diese jedoch im Weiteren nicht einzeln dargestellt, sondern die wesentlichen Erkenntnisse an einer ausgewählten Versuchsreihe (siehe Abschnitt 3.8) dargestellt. Dies dient zudem der Verifikation der Ergebnisse.

3.5 Signalverarbeitung und Auswertung

Um ein funktionierendes Prozessmodell aufstellen zu können, müssen die aufgezeichneten Versuchsdaten detailliert untersucht und ausgewertet werden.

3.5.1 Signalanalyse und Kenngrößenbildung

Nach Durchführung der Versuche müssen die aufgezeichneten Signale zuerst hinsichtlich ihrer Streuung und charakteristischer Signalausprägungen analysiert werden. Da die Ursache für Qualitätsschwankungen in Prozessschwankungen liegt, müssen diese möglichst gut beschrieben und interpretiert werden, um in die Prozessüberwachung und Modellbildung einzufließen. Hierbei ist es hilfreich, vorhandenes Prozesswissen und Kenntnisse über Wirkzusammenhänge einfließen zu lassen. Bei der Signalanalyse muss neben der Beurteilung auch der durchgeführte Versuchsplan mit seinen Veränderungen der Prozesseinstellung be-

rücksichtigt werden. Damit sind Aussagen darüber möglich, wie sich Veränderungen im Prozess sowie Variationen der Vorprozesse auf Signalform, Streubreite u. ä. auswirken.

Im nächsten Schritt müssen die identifizierten Signalcharakteristika, von denen ein hoher Informationsgehalt oder ein Zusammenhang mit der Qualität vermutet wird, in signifikante Kenngrößen überführt werden. Je besser diese Kenngrößen die Schwankungen beschreiben, desto mehr Informationen stehen den Prozessmodellen zur Verfügung. Die aus den Kenngrößen gebildeten Zustandsvektoren dienen den Modellen als Eingangsdaten. Kenngrößen, die häufig verwendet werden, sind beispielsweise Maximalwert oder Integrale der einzelnen Signale (siehe z. B. [Ahr04, BSNS04, Häu95, Kor96, N.N03, SH00, Ter94]). Diese Größen sind jedoch häufig nicht sehr selektiv bezüglich wesentlicher Qualitätseinflüsse. Deswegen müssen möglichst signifikante Kenngrößen erarbeitet werden, um die verschiedenen Ursachen und Einflüsse unterscheidbar zu machen und eine Modellierung mit hoher Güte zu ermöglichen. Die Methoden zur Signalanalyse, Kenngrößenbildung und deren Signifikanz sind in Kapitel 5 näher beschrieben.

3.5.2 Varianzanalysen und Kenngrößenselektion

Basierend auf der Signalanalyse und den daraus abgeleiteten Kenngrößen müssen diese hinsichtlich ihrer Streuungen und auf Zusammenhänge mit relevanten Prozesseigenschaften untersucht werden. Hierzu können verschiedene Verfahren aus der Statistik wie Korrelations- und Varianzanalysen verwendet werden. Damit können Zusammenhänge der verschiedenen Größen mit diversen Zielgrößen, z. B. Einstellparameter oder Qualitätsmerkmale, statistisch nachgewiesen werden. Aufbauend darauf erfolgt eine Selektion, um die wichtigsten Kenngrößen zu ermitteln. Die Dimension des beschreibenden Kenngrößenvektors, der als Zustandsvektor die Eingangsinformation der Modelle bildet, wird somit bei maximalem Informationsgehalt minimiert. Eine Optimierung der Kenngrößenauswahl führt meist zu kleineren und robusteren Prozessmodellen bei höherer Genauigkeit. Dies hängt damit zusammen, dass einerseits weniger freie Modellparameter ermittelt werden müssen und andererseits durch Eliminierung nicht signifikanter Kenngrößen die störenden „Rauschanteile“ beim Modelltraining reduziert werden. Dabei ist es wichtig, die gegenseitige Beeinflussung der Modellbildung und der Kenngrößenselektion zu berücksichtigen. Für die Selektion können Verfahren wie statistische Mittelwertvergleiche oder auch heuristische Verfahren wie Genetische Algorithmen (GA) zum Finden einer optimierten Lösung sowie Kombinationen eingesetzt werden. Die untersuchten Methoden hierzu werden in Abschnitt 5.3 näher behandelt.

3.6 Modellbildung

Kernelement einer modellbasierten Prozessüberwachung ist ein geeignetes Modellierungsverfahren, welches mit seinen Eigenschaften und Fähigkeiten möglichst gut zu den gestellten Anforderungen passt. An die Modelle des in Abbildung 1.5 dargestellten Lösungsansatzes werden mehrere Anforderungen gestellt: Die Prognose der Schweißqualität und die Erkennung von geänderten Vorprozessen, die sich in dem Modell „unbekannten Zuständen“ auswirken. Für die Qualitätsüberwachung der geschweißten Werkstücke, z. B. bei der Festigkeit, ist die **Prognose eines kontinuierlichen Merkmalswertes** durch das Modell notwendig. Modelle, die nur mit Klassendaten – d. h. die Zielgröße nimmt genau einen Wert aus einer endlichen Wertemenge an – arbeiten, sind dafür relativ schlecht geeignet. Dies gilt insbesondere für Modellierungsansätze, die auf Zwei-Klassen-Probleme ausgelegt sind, wie die Support Vector Machines [CV95]. Es existieren zwar entsprechende Erweiterungen auf Mehr-Klassen-Probleme, die das Problem auf mehrere Zwei-Klassen-Probleme zurückführen. Diese werden jedoch recht schnell unpraktikabel, da für jede existierende Klasse eine Klassifikation erstellt wird, die jeweils die Daten der Klasse entweder von jeder anderen Klasse separat oder von den verbleibenden Restdaten als Gesamtklasse unterscheidet. Bei k Klassen sind somit insgesamt k (one-versus-rest) oder $\frac{k*(k-1)}{2}$ (one-versus-one) Klassifikatoren sowie eine verknüpfende Operation zur Ableitung einer Gesamtaussage notwendig [HV10, RK04, WW99].

Während für die Prognose der Schweißqualität die für die Modelladaption notwendigen Trainingsdaten in der Regel über entsprechend geplante Versuche generiert werden können, trifft dies in der Praxis für die Störungserkennung in Vorprozessen nicht zu. Veränderungen in den Vorprozessen, die sich auf die Schweißbarkeit und das Schweißergebnis auswirken, müssen sich somit auch auf den Prozess auswirken und entsprechende Variationen in den aufgezeichneten Signalen bedingen. Solche Abweichungen können durch ein entsprechendes Modell erkannt werden, welches eine **Unbekannt-Detektion** – *novelty detection* – ermöglicht. Bei neueren Ansätzen des *active learning* werden solche Daten auch hochinformativ oder *highly informative* genannt [Bis06]. Hierfür ist es notwendig, aus den Trainingsdaten bestimmte Eigenheiten oder Ähnlichkeiten zu extrahieren, um eine Art Fingerabdruck des Prozesses zu erstellen und diesen in einem Modell abzubilden. Daraus ergeben sich wiederum Anforderungen an das Prozessmodell und die Modellbildung. Für eine verfügbare Trainingsmenge muss es möglich sein, ein Maß für die „Bekanntheit“ abzuleiten und veränderte Prozesssignale gegenüber diesen bekannten Zuständen zu unterscheiden. Es ist also ein **Maß für die Unbekanntheit** oder Neuheit erforderlich, welches mit dem entsprechenden Modell korrespondiert [NK10a, NK10b].

Zur Bewertung der Leistungsfähigkeit ist die Festlegung einer Modellgüte notwendig. Diese muss verschiedene Kriterien abbilden, die den Erfülltheitsgrad bestimmter Anforder-

Einstellung	A	B	C	D	E
Anz. Teile	20	20	19	20	19
Diagramm-Marker	■	◆	▲	●	✘
Amplitude (%)	90	100	90	90	90
Weg RPN (mm)	0,32	0,32	0,35	0,32	0,32
Triggerkraft (N)	300	300	300	300	100
Schweißkraft (N)	300	300	300	350	100
mittlerer Berstdruck (N)	20,9	26,0	29,9	29,6	27,1
minimaler Berstdruck (N)	15,5	19,9	24,8	26,6	24,5
maximaler Berstdruck (N)	26,8	30,5	33,2	35,1	32,4
Standardabw. Berstdruck (N)	3,87	2,91	2,07	2,54	2,30

Tabelle 3.1: Versuchsreihe für Modellverifikation

rungen darstellen. Die Modellgüte ist eng mit der Lernstrategie, der Anpassung der Modellparameter sowie der Optimierung des verwendeten Kenngrößenvektors durch Selektion und Skalierung verknüpft (siehe auch Abschnitte 4.5 und 5.3).

Die genannten Anforderungen werden durch GRNFN-Modelle in weitem Umfang unterstützt. Deshalb wird diese Modellart hier favorisiert und ihre Eigenschaften in Kapitel 4 näher untersucht.

3.7 Validierung und Umsetzung an Praxisbauteilen

Abschließend müssen die erarbeiteten Methoden und Verfahren, die anhand eines möglichst idealen Probekörpers erarbeitet wurden, verifiziert werden. Für einen Nachweis der Gültigkeit müssen die Verfahren anhand von realen Bauteilen, wie sie auch in der industriellen Praxis gefertigt werden, untersucht werden. Um eine Umsetzung direkt an einem Schweißprozess zu ermöglichen, müssen die Verfahren so weit implementiert werden, dass insbesondere die Kenngrößenermittlung automatisch mittels geeigneter Algorithmen erfolgt und keine manuelle Ableitung von Kenngrößen mehr notwendig ist. Ergebnisse aus Versuchen mit industriellen Produktionsprozessen sind in Abschnitt 6.3 aufgeführt.

3.8 Versuchsdaten zur Ergebnisdarstellung

Die exemplarische Darstellung der Einflüsse von Prozessvariationen erfolgt nachstehend immer anhand derselben Versuchsdaten. Bei der Versuchsreihe wurden Probekörper aus ABS mit ERG-Nahtgeometrie verschweißt, wobei ein definierter Weg als Abschaltkriterium vorgegeben war. Variiert wurden bei der Versuchsreihe die Schweißamplitude, die Trigger- und Schweißkraft sowie der eingestellte Schweißweg. Auf die einzelnen Einstellungen, die mit den Buchstaben A bis E bezeichnet sind, wird nachfolgend öfters Bezug genommen

Einstellung	A	B	C	D	E
MW Berstdruck (N)	20,9	26,0	29,9	29,6	27,1
Standardabw. (N)	3,87	2,91	2,07	2,54	2,30
Signifikanz (%)	Unterschiede des Berstdrucks				
A	— X —	99,996 %	100 %	100 %	100 %
B	78,03 %	— X —	99,998 %	99,98 %	79,43 %
C	98,95 %	84,32 %	— X —	32,87 %	99,97 %
D	92,68 %	43,88 %	60,74 %	— X —	99,78 %
E	96,86 %	67,43 %	33,82 %	32,16 %	— X —
	Unterschiede der Standardabw.				

Tabelle 3.2: Signifikanzwerte für den Vergleich von Mittelwert und Standardabweichung (obere bzw. untere Dreiecksmatrix) des Berstdrucks bezüglich der Einstellungen

und in Projektionen der angegebene Diagramm-Marker verwendet. Die exakten Werte der einzelnen Parameter sind in Tabelle 3.1 aufgeführt. Mit jeder Einstellung wurden 20 Probekörper geschweißt, um statistische Aussagen ableiten und eine Trennung in separate Lern- und Testmuster für Analysen vornehmen zu können. Von den insgesamt 100 Schweißungen waren zwei wegen Fehlern bei der Signalaufzeichnung oder Überschreitens der zulässigen Schweißzeit, die auf eine zu geringe gewählte Schweißkraft zurückzuführen war, unbrauchbar. Insgesamt ergeben sich somit 98 Zyklen als Datenbasis. Für jeden der Datenzyklen wurden aus den Signalen entsprechende Kenngrößen abgeleitet, die nach den in Kapitel 5 dargestellten Methoden erarbeitet wurden. Aus dieser Gesamtmenge von deutlich über 100 Größen wurden mit den in Abschnitt 5.3 beschriebenen Methoden insgesamt 30 Kenngrößen ausgewählt, die für die Analysen den Zustandsvektor für jeden Zyklus bilden.

Tabelle 3.2 zeigt eine statistische Auswertung dahingehend, ob sich die Einstellungen bezüglich der erzielbaren Schweißnahtfestigkeit unterscheiden. Dazu wurden die erzielten Berstdrücke der Einstellungen paarweise einem statistischen Vergleich bezüglich dem Mittelwert unterzogen. Dies erlaubt eine statistisch abgesicherte Aussage, ob sich die Mittelwerte unterscheiden und bis zu welchem Vertrauensbereich dies der Fall ist. Der Mittelwertvergleich erfolgt auf der Basis der Student-t-Verteilung. Das Konfidenzintervall entspricht dem Anteil, der zu 100 % fehlt. Neben den Messwerten an sich spielt auch die Anzahl der Messwerte – dies entspricht den Freiheitsgraden – und deren Streuung eine Rolle. Je größer ein Stichprobenumfang, desto eher können Unterschiede im Mittelwert auch bei größeren Streuungen als signifikant ermittelt werden. Die Wahrscheinlichkeiten, zu denen die Mittelwerte des Berstdrucks zwischen zwei Einstellungen unterschiedlich sind, enthält Tabelle 3.2 in der oberen Dreiecksmatrix. In den meisten Fällen unterscheiden sich die erzielbaren Schweißnahtfestigkeiten mit einer Wahrscheinlichkeit $> 99,9\%$. Dies ist direkt auf die variierten Schweißparameter zurückzuführen. In der unteren Dreiecksmatrix sind zusätzlich die Wahrscheinlichkeiten (ermittelt mit dem sogenannten F-Test) dafür angegeben, dass

sich die Varianzen unterscheiden. Die Varianz ist identisch dem Quadrat der Standardabweichung und somit ein Maß für die Streuung. Damit ist eine Abschätzung über die Robustheit einer Einstellung möglich.

Insgesamt ist zu erkennen, dass eine Erhöhung des eingestellten Schweißwegs für diesen Prozess zu einer Erhöhung der Festigkeit bei gleichzeitig reduzierter Streuung führt. Die Statistik ist das wesentliche Werkzeug, um Versuchsreihen auszuwerten. Eine Reduktion des Versuchsaufwands sowie die strukturierte und gezielte Lösung gegebener Aufgabenstellungen kann mittels einer statistischen Versuchsplanung erreicht werden. Bei entsprechender Gestaltung der Versuchspläne können mit minimalem Aufwand neben Haupt- und Nebenefekten auch Wechselwirkungen ermittelt und andere Einflüsse abgeschätzt werden. Bei allen Auswertungen und Folgerungen ist jedoch zu berücksichtigen, dass die Messwerte durch die Streuung der eingesetzten Messmittel beeinflusst sind. Für weitere Details zu statistischen Größen, Versuchsplanung und Auswertungen wird auf die einschlägige Literatur verwiesen (z. B. [HEK09, Mon09, Sch97]).

4 Untersuchte Zustandsvektoren und Modellbildung

Die Anforderungen an ein Prozessmodell, das den in Abbildung 1.5 dargestellten Lösungsansatz umzusetzen vermag, sind bereits im Abschnitt 3.6 näher erläutert worden. Im Wesentlichen sind dies zwei Fähigkeiten: Die Prognose der Ausprägung eines trainierten Qualitätsmerkmals und die Identifikation von Prozesszuständen, die sich von der Trainingsmenge signifikant unterscheiden, als unbekannt. Die Zusammenfassung von N Kenngrößen x_1, x_2, \dots, x_N eines Fertigungszyklus in einem Vektor ergibt den Zustandsvektor $x = [x_1, x_2, \dots, x_N]^T$, der den Modellen als Eingangsvektor dient. Die Zusammenstellung der Komponenten des Kenngrößenvektors wird als Kenngrößenselektion (*feature selection*) bezeichnet und hat maßgeblichen Einfluss auf die erzielbare Modellgüte. Kenngrößen und Möglichkeiten zur Auswahl werden in Kapitel 5 näher untersucht. Im Folgenden wird zuerst eine Möglichkeit vorgestellt, hochdimensionale Kenngrößenvektoren durch Projektionen grafisch darzustellen. Anschließend werden die hier favorisierten Neuro-Fuzzy-Modelle in Form von *general regression neuro-fuzzy networks (GRNFN)* (siehe Abschnitt 3.6) im Detail beschrieben. Zur Analyse ihrer Eigenschaften und Randbedingungen ist insbesondere der Einsatz der in Abschnitt 4.1.2 vorgestellten Projektionsmethode des Sammons-Mappings hilfreich.

4.1 Projektion von Zustandsvektoren

Jede Menge D aus N -dimensionalen Vektoren kann als Punktwolke in \mathbb{R}^N interpretiert werden. Die Punkte besitzen Abstände zueinander, die sich bezüglich einer definierten Metrik berechnen lassen. Zumeist wird als Metrik der euklidische Abstand verwendet, der als euklidische Norm des Differenzvektors zweier Punkte definiert ist. D. h. der Abstand zweier Punkte $x, y \in \mathbb{R}^N$ berechnet sich durch

$$d(x, y) = \|x - y\| = \sqrt{(x_1 - y_1)^2 + \dots + (x_N - y_N)^2}. \quad (4.1)$$

Der Abstand ist ein Maß dafür, wie ähnlich sich zwei Vektoren oder Punkte sind. Entsprechend ihrer Abstände ergeben sich Streuungen und Cluster innerhalb der Punktwolke.

Ein Cluster ist dabei eine Untermenge an Punkten, die jeweils einen vergleichsweise geringen Abstand zueinander haben. Vektoren, die zu verschiedenen Clustern gehören, weisen dabei einen großen Abstand zueinander auf. Jedem Vektor können Zielgrößen oder Eigenschaften zugeordnet werden, so dass die Verteilung der Vektoren auch die Anordnung der Zielgrößen bestimmt. Projektionen ermöglichen es dabei, diese Punkteverteilungen im hochdimensionalen Raum darstellbar zu machen, um so Zusammenhänge und Abhängigkeiten erkennen zu können. Es erfolgt also eine Abbildung des N -dimensionalen Ursprungsraums in den zweidimensionalen Bildraum. Eine Modellierung versucht, unabhängig vom Verfahren, die Zusammenhänge im Ursprungsraum zu beschreiben. Das Verständnis von Modellen ist somit eng mit der Verteilung von Punktemengen verknüpft und kann durch geeignete Projektionen gefördert werden. Zwei Projektionsmethoden werden im Folgenden kurz vorgestellt.

4.1.1 PCA

Eine häufige Methode zur Berechnung von Bildpunkten hochdimensionaler Vektoren ist die Karhunen-Loève-Transformation oder Hauptkomponentenanalyse PCA (*principal component analysis*) [Jol86]. Dabei handelt es sich um eine Hauptachsentransformation der durch die Vektoren gebildeten Punktwolke. Das Verfahren kommt auch sehr häufig in der Bildverarbeitung zum Einsatz. Durch die Bildung einer Reihenfolge entsprechend des Informationsgehalts der transformierten Kenngrößen kann durch Weglassen von weniger wichtigen Kenngrößen eine Dimensionsreduktion erzielt werden, ohne wesentliche Informationen zu verlieren. Auch Rauschanteile können so effizient ausgefiltert werden. Auch die in Abschnitt 2.6.2 vorgestellte Kenngrößenbildung bei NEPRES basiert auf der PCA.

Die Komponenten der ersten beiden Hauptachsen dienen als Koordinaten des zweidimensionalen Bildraums. Für die Matrix, welche aus den zeilenweise angeordneten Vektoren gebildet wird, werden die Eigenwerte und Eigenvektoren berechnet. Die beiden Eigenvektoren mit den größten Eigenwerten werden zu den beiden Koordinatenachsen. Für die Darstellung und Erkennung von Clustern und Ähnlichkeiten, die durch gegenseitig geringe Abstände der zugeordneten Vektoren gekennzeichnet sind, ist die PCA jedoch ungeeignet, da die Abstandsinformation verloren geht. Die Nähe zweier Punkte in der Bildebene bedeutet nicht zwangsweise, dass die zugeordneten Vektoren im Originalraum auch nahe beieinander liegen. Beispielsweise wird bei einem Ellipsoid bei Anwendung der PCA die Bildebene durch die Ellipse in der Äquatorebene gebildet, deren Länge a bzw. b ist und den beiden größten Hauptachsen des Ellipsoids entspricht. Beiden Pole des Ellipsoids in Richtung der dritten Hauptachse werden durch die PCA auf denselben Punkt $O(0/0)$ – den Mittelpunkt der Bildebene – projiziert, unabhängig von ihrem Abstand $2c$, dem Doppelten der Länge der dritten Hauptachse.

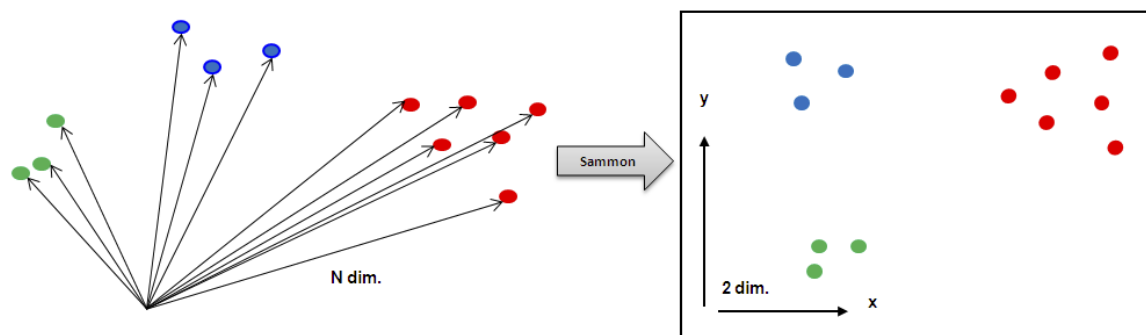


Abbildung 4.1: Prinzip des Sammons-Mapping zur Abbildung von N -dimensionalen Ursprungsvektoren in eine 2-dimensionale Bildebene

4.1.2 Sammons-Mapping

Eine andere nichtlineare Methode zur Projektion hochdimensionaler Vektorräume ist das so genannte *Sammons-Mapping* [Sam69]. Kerngedanke des Sammons-Mapping ist, dass die paarweisen Abstände der Vektoren in der Bildebene den paarweisen Abständen im Originalraum möglichst gut entsprechen. Folglich liegen Vektoren, die sich ähnlich sind – d. h. im N -dimensionalen Raum einen geringen Abstand aufweisen – auch in der Bildebene nahe beieinander, Vektoren mit großem Abstand liegen auch in der Bildebene weit auseinander. Der Grundgedanke ist in Abbildung 4.1 dargestellt. Cluster im N -dimensionalen Ursprungsraum bleiben somit auch nach der Projektion erhalten, so dass die Verteilung der Daten visuell abgeschätzt werden kann. Weiter kann den entsprechenden Vektoren eine Zielgröße zugeordnet und diese mittels einer Farbcodierung dargestellt werden. Damit kann eine Abschätzung der Homogenität der Daten hinsichtlich eines Zielmerkmals erfolgen.

Berechnung

Die Berechnung der Projektion für einen Satz an Datenvektoren erfolgt iterativ. Für die Abbildung wird folgende Zielfunktion definiert:

$$E_s = \frac{1}{\sum_i \sum_{j>i} d_{ij}} \sum_i \sum_{j>i} \frac{(d_{ij} - \|r_i - r_j\|)^2}{d_{ij}}. \quad (4.2)$$

Dabei ist d_{ij} der Abstand des i ten zum j ten Vektor im N -dimensionalen Ursprungsraum bezüglich der definierten Metrik und r_i bzw. r_j entspricht den zugeordneten Vektoren in der Bildebene. Für den euklidischen Abstand werden diese Werte nach der Gleichung (4.1) für den Ursprungs- und Bildraum berechnet. Die Differenz zwischen den paarweisen Abständen im Original- und Bildraum wird quadriert und für alle Vektoren aufsummiert. Während der Berechnung der Projektion wird diese Zielfunktion minimiert, um insgesamt eine gute Abbildung der Ähnlichkeitsverteilung zu erhalten. Ein besonderes Merkmal dabei

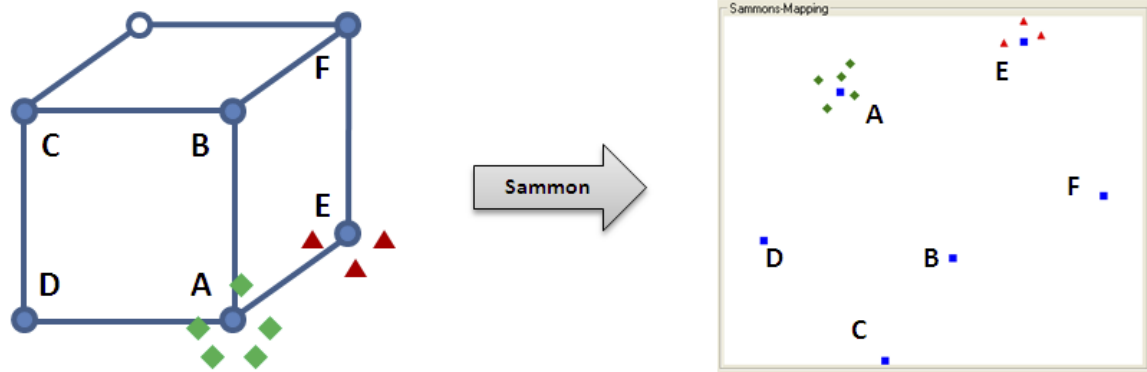


Abbildung 4.2: Sammons-Mapping für sechs Eckpunkte des Einheitswürfels inklusive ecknaher Zusatzpunkte

ist die Division der Abstandsdifferenz durch den Abstand der Vektoren in \mathbb{R}^N . Dadurch wird erreicht, dass kleine Abstände insgesamt stärker gewichtet werden und keine Dominanz der großen Abstände auftritt. Die Minimierung dieses Gütemaßes erfolgt meist durch ein iteratives Gradientenverfahren. Ähnlich der Backpropagation-Lernregel bei Neuronalen Netzen kann durch eine partielle Ableitung der Zielfunktion nach den Komponenten der Vektoren r_i im Bildraum, also den Bildkoordinaten, eine Regel zur iterativen Veränderung der Bildvektoren der Form $r_i(m+1) = r_i(m) - \alpha \cdot \Delta r$ abgeleitet werden [Sam69]. $r_i(m)$ bezeichnet den Vektor im (zweidimensionalen) Bildraum im m -ten Iterationsschritt und α entspricht der Lernrate, von SAMMON auch „magic factor“ genannt, mit einem empirisch ermittelten Wert zwischen 0,3 und 0,4. Ausgehend von einer Startinitialisierung der Bildvektoren wird der Iterationsprozess nach einer vorgegebenen Anzahl an Schritten abgebrochen. Das Ergebnis des hier realisierten Sammons-Mapping für sechs Eckpunkte des Einheitswürfels in \mathbb{R}^3 ist beispielhaft in Abbildung 4.2 dargestellt. Durch die niedrige Dimension der Ausgangsvektoren ist gut nachvollziehbar, dass das Verhältnis der Abstände der Eckpunkte in der Projektion erhalten bleibt. Auch die zufällig um die Eckpunkte A und E gewählten Punkte liegen nach der Abbildung in entsprechender Nähe.

Eigenschaften des Sammons-Mapping

Der iterative Prozess zur Berechnung der Projektion bedingt einige Eigenschaften des Sammons-Mapping. Die genaue Lage der Bildpunkte ist stark von deren Startinitialisierung abhängig, die meist zufällig erfolgt, und eine Veränderung der Datenbasis, z. B. durch Hinzufügen einzelner Vektoren, kann sich teils deutlich auf das Ergebnis der Projektion auswirken. Außerdem existiert keine mathematisch darstellbare Abbildungsvorschrift, mit der die Projektion weiterer Daten berechnet werden kann. Eine näherungsweise Abbildung kann jedoch durch ein adaptiertes Schätzverfahren erfolgen, sofern die Originaldaten entsprechend verfügbar sind. Für den neuen Vektor werden die drei nächsten Nachbarn der

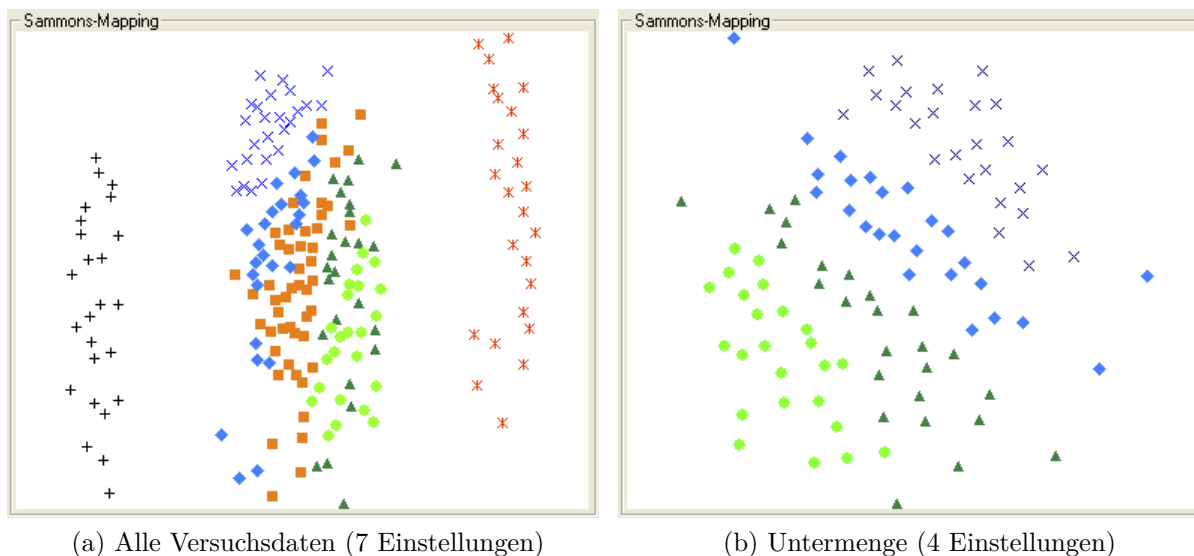


Abbildung 4.3: Verzerrungsbedingte Klassenüberlappung beim Sammons-Mapping

Originaldaten sowie ihre Bildpunkte ermittelt und die jeweiligen Abstände in \mathbb{R}^N sowie in \mathbb{R}^2 für die erste Schätzung des Bildpunkts berechnet. Über ein Minimierungsverfahren, z. B. ein Downhill-Simplex-Verfahren [PTVF07], kann der Bildpunkt verschoben werden, bis das Verhältnis der Abstände zwischen Original- und Bildraum hinreichend genau ist – in der Regel, bis der Fehler kleiner als ein Pixel des Darstellungsmediums ist.

Die Projektion kann – wie alle Transformationen in einen Raum mit niedrigerer Dimension – gewissen Verzerrungen unterliegen, die umso größer ausfallen können, je stärker die Dimensionsreduktion ist. Dies ist dadurch bedingt, dass einzig die Summe aller Abstandsfehler minimiert wird. Die grobe Aufteilung der Cluster ist immer zu erkennen, jedoch sollte die Berechnung für mehrere unterschiedliche Startinitialisierungen erfolgen, um die Struktur der Daten durch Vergleiche besser abschätzen zu können. So ist es anhand der Projektionen auch nicht möglich, für einen Datenpunkt den nächsten Nachbarn im N -dimensionalen Ursprungsraum sicher anzugeben. Im Bildraum können durch die Verzerrungen reale Punkte mit größerer Entfernung in der Projektion näher erscheinen. So ist es oft sinnvoll, nach einer Grobsichtung der Gesamtdaten für einzelne Cluster oder Untermengen eine separate Projektion zu berechnen. In Abbildung 4.3 ist exemplarisch die Verteilung von Kenngrößenvektoren für eine Versuchsreihe mit der Ultraschallschweißanlage dargestellt, bei der u. a. die Schweißparameter variiert wurden. Die unterschiedlichen Einstellungen sind mit unterschiedlichen Markern und Farben gekennzeichnet. In der linken Projektion, die für alle sieben durchgeführten Prozesseinstellungen berechnet ist, ist deutlich die Ausbildung von Clustern zu erkennen, die aus den veränderten Prozesseinstellungen herrühren. Dabei scheint es, dass sich einige der Cluster in der Mitte durchdringen. Wird jedoch nur für die vier relevanten Einstellungen aus diesem Bereich ein separates Sammons-Mapping berechnet (Abbildung 4.3(b)) ist zu erkennen, dass sich die Einstellungen komplett trennen lassen.

Auch die mit Dreiecken und Kreisen zuvor durchmischten Klassen sind nun komplett separierbar. Die Ursache für die scheinbare Durchdringung ist in Verzerrungen der Projektion begründet.

4.2 GRNFN-Modelle: Aufbau und Training

Im Folgenden wird der Aufbau von GRNFN-Modellen erläutert, um einen Bezug zu den daraus abgeleiteten Eigenschaften und die Funktionweise bei der Qualitätsprognose sowie Unbekannt-Klassifikation zu ermöglichen.

4.2.1 Modellstruktur

Bei GRNFN handelt es sich um interpolierende Neuro-Fuzzy-Modelle, wie sie schon in Abschnitt 2.4.4 eingeführt wurden. Die Architektur von GRNFN ist in Abbildung 4.4 dargestellt. Bei GRNFN handelt es sich um speicherbasierte (memory-based) Modelle. Sie verfügen im Wesentlichen über L N -dimensionale Datenvektoren, die als Stützpunkte für die Interpolation dienen. Diesbezüglich sind GRNFN den RBF-Netzen (siehe Abschnitt 2.4.2) sehr ähnlich, jedoch werden die Eingabewerte nicht direkt verarbeitet, sondern zuerst fuzzifiziert. Die allgemeine Aufgabe bei der Modellierung einer Strecke – z. B. einem Fertigungsprozess – anhand von L ermittelten Messpunkten $(x^*, y^*)^k$, $k \in [1; L]$ besteht darin, den durch die Strecke vorgegebenen, nichtlinearen Zusammenhang $y = F(x)$ zu identifizieren. Bei speicherbasierten Modellen sind die Stützstellen durch die Messpunkte definiert. Des Weiteren spielen für die Approximation eine (stützpunktspezifische) Distanzfunktion $D_k(s)$ entsprechend einer festgelegten Metrik und eine Kernelfunktion φ – in der Regel eine Gaußfunktion – eine Rolle. Die Distanzfunktion $D_k(s)$ definiert den Abstand des Eingabevektors s zum der Funktion zugeordneten Stützvektor x_k^* . Eine Schätzung $y_{prog} = \hat{y}(t)$ als Ausgabegröße der Strecke für einen beliebigen Punkt x berechnet sich dann als

$$y_{prog} = \hat{y}(x) = \frac{\sum_{k=1}^L y^*(k) \varphi(D_k(x))}{\sum_{k=1}^L \varphi(D_k(x))} \quad (4.3)$$

(siehe [BT08, HTF01, Spe91]). Ein GRNFN-Modell bildet durch den im Folgenden beschriebenen Aufbau einen entsprechenden Schätzer. Das Modell verfügt über N Eingänge $x_1 \dots x_N$ und kann somit Eingangsvektoren $x = [x_1, x_2, \dots, x_N]^T$ mit der Dimension N der modelleigenen Stützvektoren verarbeiten. Jeder Eingangsvektor x wird zuerst durch eine Normierung in den Vektor \tilde{x} transformiert, so dass für alle Komponenten \tilde{x}_i gilt: $\tilde{x}_i \in [-0,5; 0,5]$, $i \in [1..N]$. Der so normierte Eingangsvektor \tilde{x} wird an jeden der L Blöcke angelegt, die jeweils durch einen Stützvektor sowie N Zugehörigkeitsfunktionen μ_i^k , $i \in [1..N]$, $k \in [1..L]$ definiert sind. Die N Zugehörigkeitsfunktionen μ_i^k eines Blocks k transformieren jeweils eine Komponente des Eingangs in einen entsprechenden Aktivierungswert $\varphi_i^k \in [0; 1]$.

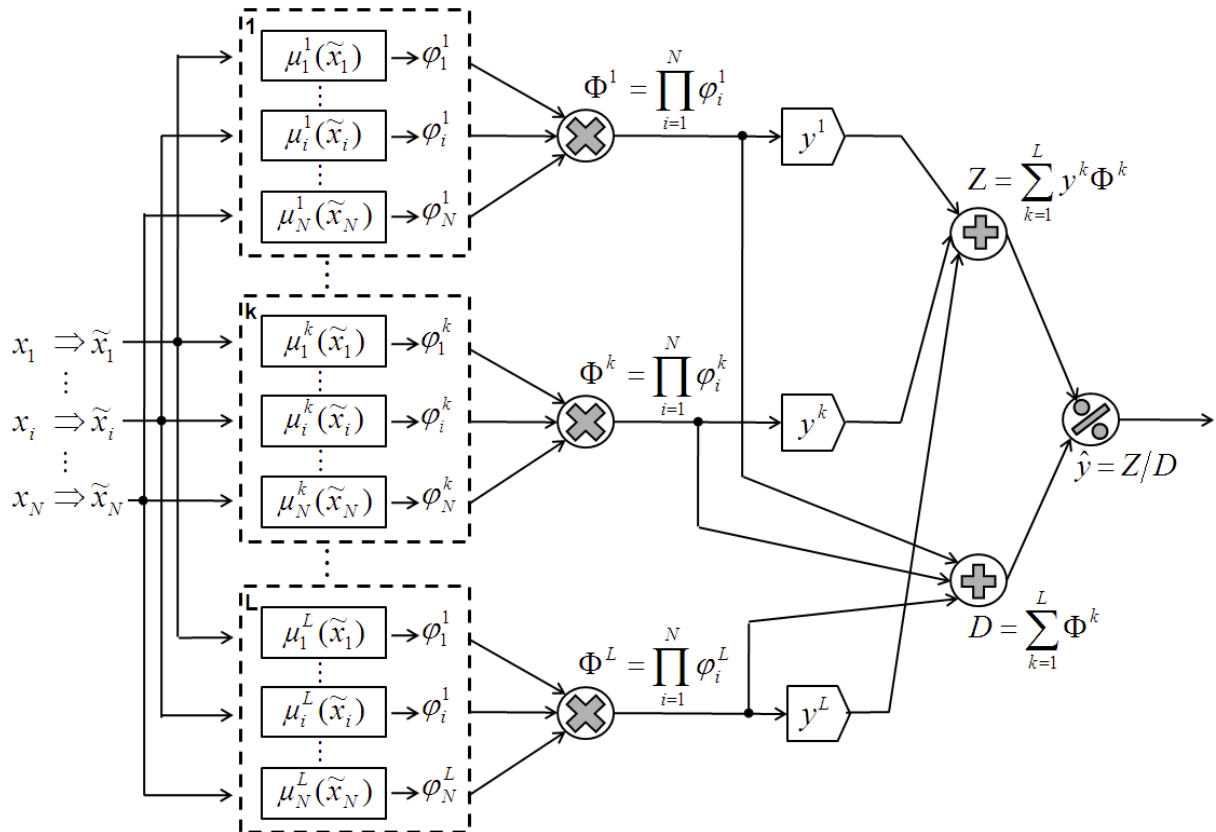


Abbildung 4.4: Architektur von general regression neural-fuzzy networks (GRNFN) (nach [BOPT07])

Durch Multiplikation der Aktivierungswerte φ_i^k der einzelnen N Zugehörigkeitsfunktionen ergibt sich für jeden der L Blöcke eine vom Eingangsvektor abhängige Gesamtaktivierung Φ^k , $k \in [1..L]$. Jeder dieser Gesamtaktivierungswerte Φ^k wird mit einem spezifischen Faktor y^k multipliziert und anschließend zu Z aufsummiert. Parallel dazu wird aus den unskalierten Gesamtaktivierungswerten Φ^k eine Gesamtsumme D gebildet. Abschließend wird der dem anliegenden Eingangsvektor x zugeordnete Ausgabewert $\hat{y} = Z/D$ des Netzes als Quotient dieser beiden Summen gebildet.

Ein GRNFN-Modell lässt sich damit als sechsschichtiges Netzmodell darstellen, wobei einige Schichten über Knoten mit Parametern verfügen und andere aus reinen „funktionalen Knoten“ bestehen:

Schicht 1 – Normierung: In der ersten Schicht werden alle Komponenten des Eingangsvektors normiert, so dass sie Werte im Bereich $[-0,5; 0,5]$ annehmen. Die Ausgabe der ersten Schicht besteht somit aus N Werten im Bereich $[-0,5; 0,5]$.

Schicht 2 – Fuzzifizierung: Die zweite Schicht verfügt über L Blöcke mit jeweils N Zugehörigkeitsfunktionen. Die N Komponenten des Netzeingangs werden mittels der Zugehörigkeitsfunktionen der einzelnen Blöcke fuzzifiziert und in die Aktivierungswerte φ_i^k überführt. Insgesamt verfügt die Schicht 2 somit über $N \cdot L$ Ausgabewerte.

Schicht 3 – Aggregation der Zugehörigkeitswerte: Aus den N Zugehörigkeits- oder Aktivierungswerten eines Blocks wird durch Multiplikation eine Gesamtaktivierung Φ^k gebildet. Die dritte Schicht verfügt somit über L Ausgabewerte. Diese haben einen Wertebereich von $[0; 1]$, da auch die multiplikativ verknüpften Zugehörigkeitsfunktionen diesen Wertebereich haben.

Schicht 4 – synaptische Gewichte: Die vierte Schicht verfügt über die synaptischen Gewichte, welche maßgeblich das Ausgabeverhalten des Netzes bestimmen. Die Gesamtaktivierungswerte Φ^k werden durch Multiplikation mit den entsprechenden Gewichten y^k skaliert. Dieser Schritt entspricht der Interferenz bei Fuzzy-Systemen.

Schicht 5 – Akkumulation: In der fünften Schicht findet die Akkumulation der Regelbasis statt, die im Wesentlichen aus den L Blöcken der ersten Schichten gebildet wird. Ausgabewerte sind die beiden Summen der skalierten und unskalierten Gesamtaktivierungswerte Φ^k .

Schicht 6 – Netzausgabe: In der letzten Schicht wird die Netzausgabe \hat{y} als Gewichtung der Summe der skalierten Gesamtaktivierungswerte mit der Summe der Gesamtaktivierungswerte berechnet.

4.2.2 Trainingsphase

Während die Operationen der Schichten 3, 5 und 6 definiert sind und keinen Freiheitsgrad bieten, sind die Parameter der ersten, zweiten und vierten Schicht im Laufe der Trainingsphase eines GRNFN-Modells festzulegen. Das Training erfolgt anhand einer Trainingsdatenmenge Θ , bestehend aus M Datensätzen. Jeder dieser Datensätze ist durch einen N -dimensionalen Vektor als Netzzinput sowie die zugehörigen Ausgabe- oder Zielgröße(n) definiert. Damit handelt es sich bei GRNFN um Modelle mit überwachtem Lernverfahren, d. h. für die Trainingsdaten sind die zugeordneten Zielgrößen bekannt.

Normierungsgrenzen

Die Normierungsgrenzen für die Netzeingänge werden komponentenweise anhand der Trainingsdaten bestimmt. Für jede der N Komponenten werden die minimalen und maximalen Werte \hat{x}_i^{min} und \hat{x}_i^{max} der M Trainingsdaten bestimmt. Die Normierung der Netzeingänge erfolgt dann nach der Formel

$$\tilde{x}_i = \frac{x_i - \hat{x}_i^{min}}{\hat{x}_i^{max} - \hat{x}_i^{min}} - 0,5 ; i \in [1..N]. \quad (4.4)$$

Für die Anwendung in der Praxis ist es auch möglich, die Normierungswerte etwas aufzuweiten (z. B. um 5%). Dadurch kann gewährleistet werden, dass alle Eingabewerte auf

den Bereich $[-0,5; 0,5]$ abgebildet werden, auch wenn einzelne Eingangswerte außerhalb des Trainingsbereichs liegen. Noch größere Abweichungen können durch eine vorgelagerte Vorverarbeitung im Rahmen einer Vergleichsoperation mit den Intervallgrenzen erkannt und die Daten als *Unbekannt* gekennzeichnet werden. Im Rahmen der in Abschnitt 4.6 vorgestellten Verwendung des GRNFN-Modells zur Unbekannt-Klassifikation ist dieser Schritt jedoch impliziert und kann somit als separate Funktion entfallen.

Festlegung der Blockanzahl L

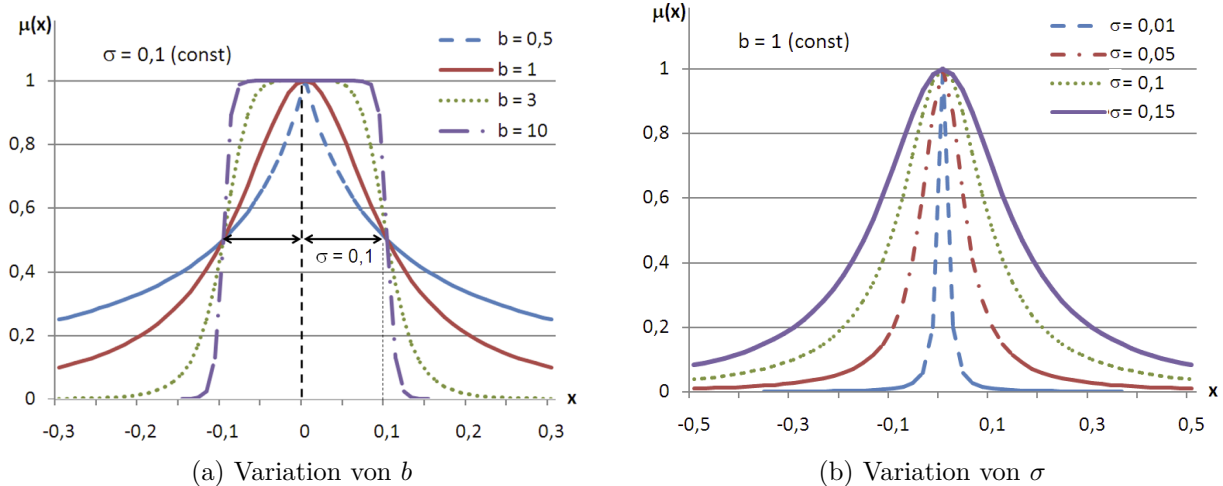
Die Anzahl L der Blöcke in Schicht 2 entspricht im Allgemeinen der Anzahl M der Trainingsdaten. Deren N -dimensionale Vektoren dienen jeweils als Stützvektoren. Darin spiegelt sich der Grundgedanke der speicherbasierten Modellstruktur (memory based) wider. Jeder Block ist somit mit einem Neuron bei RBF-Netzen vergleichbar (siehe Abschnitt 2.4.2). Im Unterschied zu RBF-Netzen besitzt jeder Block jedoch N eindimensionale Zugehörigkeits- bzw. Aktivierungsfunktionen zur Fuzzifizierung der Eingangsgrößen, während RBF-Netze über eine Aktivierungsfunktion mit N -dimensionalem Argument verfügen.

Eine Abweichung der Blockanzahl von der Anzahl der Trainingsdaten kann dann auftreten, wenn entweder einige Trainingsdaten identisch sind oder für einige Daten keine zugeordneten Ausgänge verfügbar sind. Dann kann jedoch durch Weglassen der entsprechenden Vektoren die Trainingsmenge Θ in eine Menge Θ^* überführt werden, die komplett in das Modelltraining einfließt. Ein weiterer Sonderfall, der zu einer größeren Anzahl der Blöcke führt, tritt ein, wenn zusätzlich zu den Lerndaten in Form von Vektoren manuell festgelegte Fuzzy-Regeln, wie sie bei regulären Fuzzy-Systemen Anwendung finden, mit berücksichtigt werden sollen (siehe Abschnitt 4.4.4).

Aktivierungsfunktionen

Die Aktivierungsfunktionen μ_i^k mit $i \in [1..N]$, $k \in [1..L]$ in den L Blöcken der zweiten Schicht, auch Zugehörigkeitsfunktionen oder Fuzzy-Memberfunktionen genannt, sind die wesentlichen Elemente für die Fuzzifizierung der Eingangsgrößen. Wesentliches Element jeder Aktivierungsfunktion μ_i^k ist dabei ihr „Zentral-“ oder Stützpunkt c_i^k . Die Festlegung der Funktionen sowie deren Parameter erfolgt anhand der Trainingsdaten Θ und bestimmt maßgeblich das Modellverhalten. Die wesentlichen Eigenschaften jeder Funktion μ_i^k sind:

- eine Abbildung der Art $\mu : \mathbb{R} \rightarrow \mathbb{R}_0^+$,
- die Definitionsmenge entspricht (mindestens) den reellen Zahlen im Bereich $[-0,5; 0,5]$, der Wertebereich den reellen Zahlen im Bereich $[0; 1]$,
- die Funktion besitzt ein Maximum an der Stelle $c_{i,k}$ und es gilt $\mu_i^k(c_{i,k}) = 1$.

Abbildung 4.5: Einfluss der Parameter auf die Form der Glockenkurve ($c = 0$)

Die Aktivierungsfunktion entspricht in der allgemeinen Approximationsformel (4.3) dem verwendeten Kernel φ inklusive der Distanzfunktion $D(s)$. Je größer der Abstand zum Zentrum ist, desto geringer ist die Aktivität.

Eine häufig verwendete Aktivierungsfunktion ist die verallgemeinerte Gaußkurve – auch allgemeine Glockenkurve genannt – mit den drei Parametern c , σ und b und der Berechnungsformel

$$\mu_c(\tilde{x}_i) = \left(1 + \left| \frac{\tilde{x}_i - c}{\sigma} \right|^{2b} \right)^{-1}, \quad b \geq 0,5. \quad (4.5)$$

Das Zentrum der Funktion ist durch c gegeben. Bei $\tilde{x}_i = c$ hat die Glockenkurve ihr Maximum mit dem Funktionswert $\mu_c(c) = 1$. Die Breite der Glocke ist durch den skalaren Wert σ und die Form der Glocke durch den skalaren Parameter b bestimmt. Die resultierende Glockenkurve für verschiedene Werte von σ und b mit Zentrum $c = 0$ ist in Abbildung 4.5 dargestellt.

Auch BODYANSKIY verwendet die allgemeine Glockenkurve als Aktivierungs- bzw. Fuzzy-Memberfunktion für die Fuzzifizierung in der zweiten Netzwerkschicht [BT08]. Die Zentren der einzelnen Funktionen μ_i^k werden durch die Werte der i -ten Komponente des k -ten Trainingsmusters bestimmt. Die Festlegung des Breitenparameters σ , der für alle verwendeten Funktionen identisch ist, erfolgt anhand der allgemeinen Empfehlung, dass sich die Graphen benachbarter Fuzzy-Zugehörigkeitsfunktionen – d. h. die (nach c_i^k aufsteigend sortierten) L Funktionen für jede der N Komponenten – leicht überlappen sollten [HTF01]. Für die allgemeine Gaußkurve führt dies zur Abschätzung $\sigma < \frac{L-1}{2 \div 3}$. Die aus diesen Festlegungen resultierenden Zugehörigkeitsfunktionen aus $L = 8$ Trainingsmusters für eine Komponente i mit $\sigma = \frac{1}{L-1}$ und $b = 1,5$ sind exemplarisch in Abbildung 4.6 dargestellt. Der Einfluss der Aktivierungsfunktion sowie ihrer Parameter wird in Abschnitt 4.3 näher untersucht.

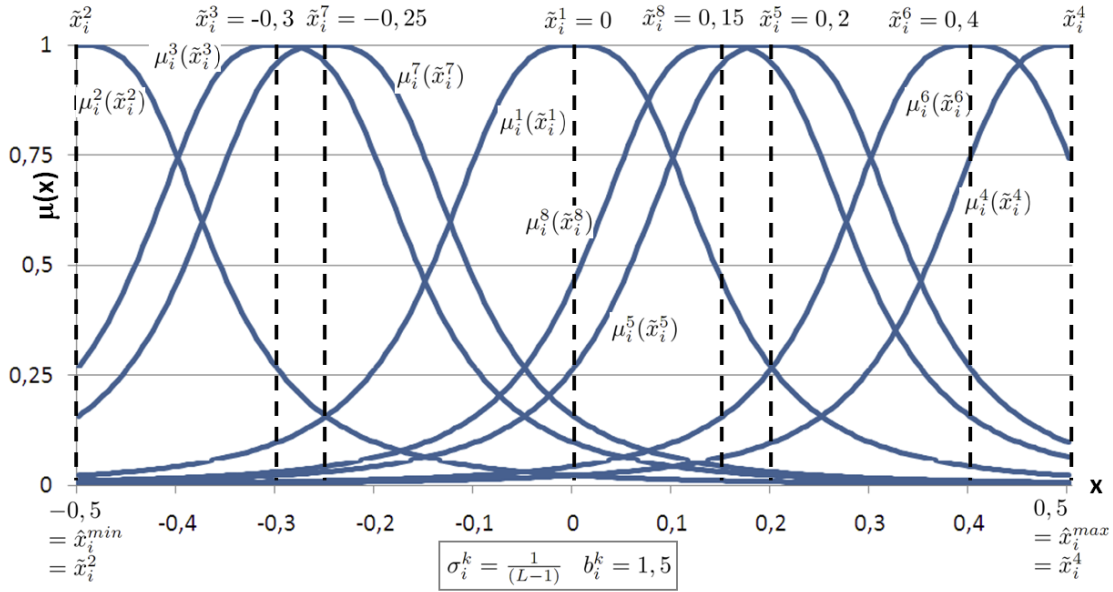


Abbildung 4.6: Verteilung der resultierenden Fuzzy-Zugehörigkeitsfunktionen einer Komponente i für $L = 8$

Synaptische Gewichte

Ein weiteres Element der GRNFN-Modelle, welche über die Approximationseigenschaften entscheiden, sind die Faktoren in der vierten Schicht. Diese Faktoren stellen die synaptischen Gewichte des Modells dar. Für die Netzausgabe eines GRNFN-Modells ergibt sich aus dem Aufbau folgender Zusammenhang, wobei der Einfachheit halber direkt der normierte Eingabevektor \tilde{x} verwendet wird:

$$\hat{y}(\tilde{x}) = \frac{\sum_{k=1}^L y^k \Phi^k(\tilde{x})}{\sum_{k=1}^L \Phi^k(\tilde{x})} = \frac{\sum_{k=1}^L y^k \prod_{i=1}^N \mu_i^k(\tilde{x}_i)}{\sum_{k=1}^L \prod_{i=1}^N \mu_i^k(\tilde{x}_i)}. \quad (4.6)$$

Durch Vergleich mit der allgemeinen Approximationsformel (4.3) ergibt sich neben der Kongruenz der Kernel- und Distanzfunktion mit der Aktivierungsfunktion, dass die synaptischen Gewichte y^k des GRNFN-Modells direkt den Ausgabewerten der einzelnen Trainingsdaten entsprechen.

Trainingsablauf

Das Training erfolgt in der Regel anhand eines zuvor definierten Satzes an Trainingsdaten. Der Trainingsvorgang an sich erfolgt – wie auch bei RBF-Netzen – in einem Durchgang. Anhand der Trainingsdaten – wobei zuvor Daten mit ungültigen oder unvollständigen Einträgen ausgefiltert werden müssen – stehen direkt die Parameter N und L fest. Zuerst werden die Normierungsgrenzen für alle N Komponenten bestimmt und anschließend die Trainingsdaten damit normiert. Dann erfolgt die Festlegung der Parameter für die $N * L$

Aktivierungsfunktionen der zweiten Schicht. Deren Zentren sind direkt durch die Trainingsdaten gegeben, die weiteren Parameter können nach festen Regeln oder Abschätzungen bestimmt werden. Die synaptischen Gewichte der vierten Schicht sind, wie zuvor erläutert, direkt durch die jeweiligen Ausgangsgrößen der Trainingsdaten definiert.

Nach den beschriebenen Schritten, deren Berechnungsdauer proportional zur Anzahl der Trainingsdaten und der Input-Dimension steigt, ist das GRNFN-Modell vollständig definiert. Es ist also im Gegensatz zu Neuronalen Netzen mit Backpropagation kein langdauerndes iteratives Training notwendig. Zudem ist – bei festgelegten Regeln für die Parameterfindung – das Ergebnis der Trainingsphase bei GRNFN deterministisch und nicht wie bei Neuronalen Netzen von der (zufälligen) Anfangsinitialisierung abhängig.

Prinzipiell ist es auch möglich, ein GRNFN-Modell parallel zum Betrieb einer Anlage aufzubauen. Dazu werden in festen Zeitabständen oder von laufenden Fertigungszyklen Messwerte und zugeordnete Ausgabegrößen erfasst und als zusätzliche Blöcke einem bereits bestehenden GRNFN-Modell hinzugefügt. Auf diese Weise ist eine Art Beobachter zur Online-Identifikation möglich. Gleichzeitig können mit der Einbindung neuer Trainingsdaten auch „ältere“ entfernt werden. Somit können auch langsame Veränderungen der Strecke abgebildet werden, ohne das Modell immer weiter anwachsen zu lassen, was zu einer stetig zunehmenden Berechnungsdauer führen würde. Durch das gezielte Hinzufügen neuer Trainingsdaten können auch neue Eigenheiten der Strecke, insbesondere nach vorheriger Klassifikation als Unbekannt, in das Modell integriert werden. Dies ist auch als *active learning* bekannt [HTF01]. Somit kann ein GRNFN-Modell zur Prozessüberwachung auch prozessbegleitend erstellt werden – die Kenntnis entsprechender Zielgrößen vorausgesetzt [NK10a, NK10b].

4.3 Einfluss der Aktivierungsfunktion

Die Aktivierungsfunktion ist ein wesentliches Element von GRNFN-Modellen. Im Folgenden wird ihr Einfluss näher untersucht. Es sind verschiedene Aktivierungsfunktionen möglich, die jeweils über einen Satz an P definierenden Parametern verfügen. Prinzipiell können diese Parameter – bis auf den Lageparameter c – global festgelegt werden, d. h. alle Funktionen μ_i^k verwenden dieselben Werte. Alternativ können die Parameter für jede Funktion an jedem Stützpunkt separat festgelegt werden. Dies bedingt zwar einen erheblichen Mehraufwand bei der Festlegung der Parameter, kann sich aber – insbesondere für die Unbekannt-Klassifikation (siehe Abschnitt 4.6) – als lohnenswert erweisen.

4.3.1 Allgemeine Glockenkurve

Die verwendete Aktivierungsfunktion hat einen wesentlichen Einfluss auf die Prognose- oder Approximationsfunktion, wie auch aus den Gleichungen (4.3) und (4.6) ersichtlich ist. BODYANSKIY verwendet die allgemeine Glockenkurve (siehe Gleichung (4.5)) als Fuzzy-Memberfunktion [BT08]. Die allgemeine Glockenfunktion verfügt über drei definierende Parameter. Neben der durch den Parameter c definierten Lage wird ihre Form durch die Parameter σ und b bestimmt. Die für ein Eingabemuster ermittelte Aktivität wird somit durch die Parameter σ und b beeinflusst, welche wiederum durch den multiplikativen Einfluss der Aktivierung in die Prognose und damit in die Modellgüte eingehen. Der Wahl geeigneter Parameter beim Training kommt somit eine entscheidende Bedeutung zu.

Kenngrößen für die Modellgüte sind – jeweils bezogen auf eine festgelegte Datenmenge – die Korrelationen zwischen Netzprognose und realem Messwert sowie mittlerem oder maximalem Fehler. Der Verlauf dieser drei Größen in Abhängigkeit der Variation der bestimmenden Parameter σ und b ist exemplarisch in Abbildung 4.7 dargestellt. Als Daten wurden die Festigkeitswerte einer Versuchsreihe beim Ultraschallschweißen mit Probekörpern aus ABS (siehe Abschnitt 3.3) verwendet. Während der Versuche wurden mehrere Schweißparameter variiert und alle Versuche für das Erstellen des Prozessmodells verwendet. Detaillierte Angaben zur Prozessüberwachung beim Ultraschallschweißen anhand des Probekörpers sind in Abschnitt 6 zu finden.

In Abbildung 4.7 ist deutlich zu erkennen, dass die Modellgüte stark von den Parametern σ und b der glockenförmigen Aktivierungsfunktion abhängt. Prinzipiell sind anhand des Verlaufs von Korrelation und Fehler folgende Zusammenhänge zu erkennen:

- die Korrelation zwischen Netzprognose und tatsächlichem Wert sinkt mit zunehmendem σ und zunehmendem b ,
- der Fehler – mittlerer und maximaler – erhöht sich sowohl mit steigendem σ als auch steigendem b .

Die daraus mögliche Folgerung, die beiden Parameter σ und b möglichst gering zu wählen, ist jedoch keine optimale Lösung. Denn dabei ist das Prognoseverhalten für Datensätze, die nicht für das Modelltraining verwendet wurden, nicht berücksichtigt. Jedes Modell sollte einerseits die Trainingsdaten möglichst gut wiedergeben können, jedoch auch einen Interpolationscharakter für „Zustände dazwischen“ aufweisen, was der Generalisierungsfähigkeit entspricht. Obwohl es sich bei GRNFN um ein stützpunktbasiertes Modellierungsverfahren handelt, und alle dem Diagramm in Abbildung 4.7 zugrunde liegenden Daten auch für das Modelltraining verwendet wurden, entspricht die Netzprognose nicht immer dem realen Wert. Ursächlich ist der Wirkzusammenhang zwischen Parameter, Aktivierungsfunktion, daraus ermittelter Aktivität aller L Trainingsmuster innerhalb der entsprechenden Blöcke

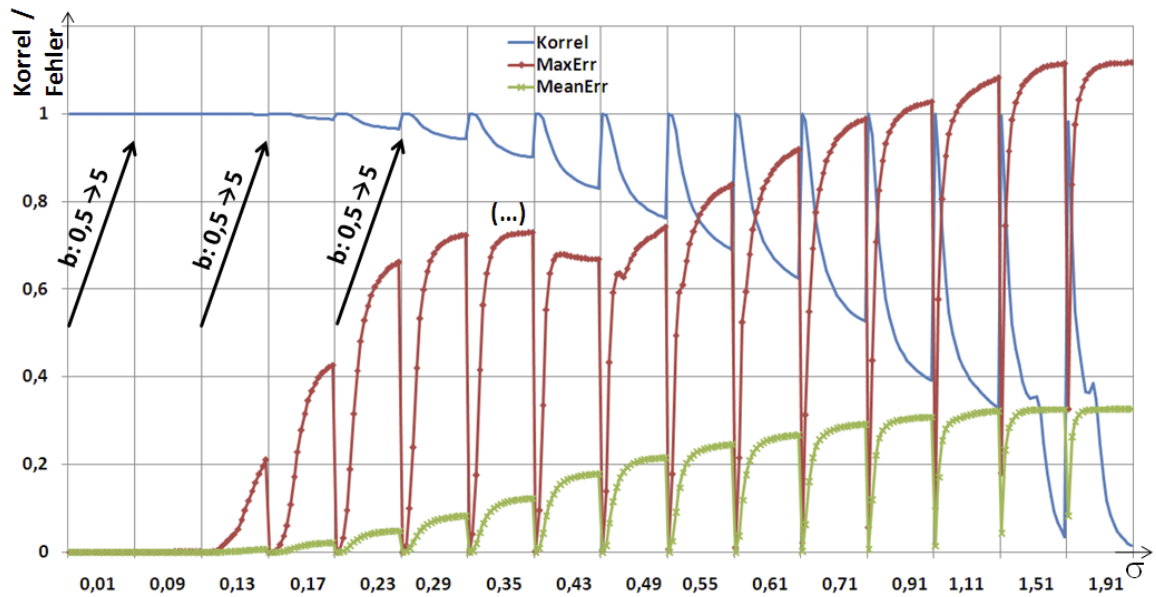


Abbildung 4.7: Modellgüte der Trainingsdaten für die allgemeine Glockenkurve in Abhängigkeit der Parameter σ und b

sowie der Berechnung des Modellausgangs (vergleiche Abbildungen 4.4, 4.5 und 4.6). Je kleiner die Werte für σ , desto schmaler ist die Form der Glockenkurven. Für Eingangswerte entsprechend der Trainingsdaten werden im jeweiligen Block hohe Aktivierungswerte erzielt. Weichen die Werte nur gering von den Trainingswerten ab, fallen die Aktivierungswerte stark ab. Die anschließende Multiplikation in Schicht 3 führt zu Werten sehr nahe Null, wodurch keines der synaptischen Gewichte, die durch die Lernmuster bestimmt sind, aktiviert wird. Dadurch wird auch keine sinnvolle Interpolation zwischen mehreren Stützvektoren der Trainingsdaten – bzw. eine gewichtete Überlagerung – erreicht. Die Netzausgabe ergibt weitestgehend sinnlose Werte, so dass solche Eingangsdaten als unbekannt zu verwerfen sind (siehe auch Abschnitt 4.6). Ein solches Modell hat quasi auswendig gelernt. Es kennt nur noch die Lernmuster, hat darüber hinaus aber seine Generalisierungsfähigkeit verloren. Umgekehrt werden für steigende Werte von σ die Glockenkurven breiter, so dass eine Aktivierungsfunktion μ_i^k auch für die Stützwerte „benachbarter“ Funktionen μ_i^m , $m \neq k$ hohe Aktivierungswerte liefert. Dadurch werden mehrere Lernmuster aktiviert und es erfolgt eine Interpolation entsprechend dem Verhältnis der Aktivierungswerte. Das GRNFN-Modell kann so generalisieren und auch Werte zwischen den Lernmustern sinnvoll interpretieren. Der Parameter b bestimmt maßgeblich darüber, wie schnell die Aktivierung abfällt. Für große Werte von b fällt die Aktivierungskurve steil ab und weist zunehmend ein binäres Schaltverhalten auf, was zu un stetigen Übergängen und sprunghaftem Verhalten des Modells führen kann¹. Kleinere Werte von b sorgen für einen sanfteren Übergang zwischen den

¹Vom mehrschichtigen Perzeptron (MLP) ist Entsprechendes bekannt – dies ist einer der Gründe, die zur Ablösung von binären Schwellen z. B. durch die Sigmoidfunktion führten (siehe Abschnitt 2.4.2 oder [Bis95, Zel94]) .

verschiedenen Stützpunkten. Damit beeinflussen σ und b gemeinsam, wieviele verschiedene Lernmuster für die Interpolation verwendet werden und mit welcher Gewichtung. Im Extremfall – sehr großen Werten für σ und relativ unabhängig von b – ist die Modellausgabe für alle Datensätze quasi konstant und entspricht weitestgehend einer Mittelung der Zielgrößen aller Trainingsdaten. Dies führt zwangsläufig zu einer Korrelation von 0. Auf diesen Zusammenhängen basiert auch die bereits in Abschnitt 4.2.2 erwähnte Forderung, dass sich die einzelnen Aktivierungsfunktionen teils überlappen sollten.

Die dargestellten Überlegungen zu den Parametern gelten für jedes GRNFN-Modell. Da jedes Modell in Schicht 1 eine Normierung der Eingänge auf den Bereich $[-0,5; 0,5]$ durchführt, können beim Training anhand von Richtlinien sinnvolle Parameterwerte ermittelt werden. Dies gilt allgemein auch für andere Aktivierungsfunktionen wie die Glockenkurve. Für die Glockenkurve haben sich die Parameterwerte $\sigma = 0,05$ und $b = 0,5$ für die meisten untersuchten Modelle als sehr geeignet erwiesen. Die Anzahl der für das Training verwendeten Versuchszyklen war dabei Bereich von 40–200 Zyklen. Bei noch mehr Trainingsdaten sollte der Wert für σ entsprechend reduziert werden.

4.3.2 Weitere Aktivierungsfunktionen

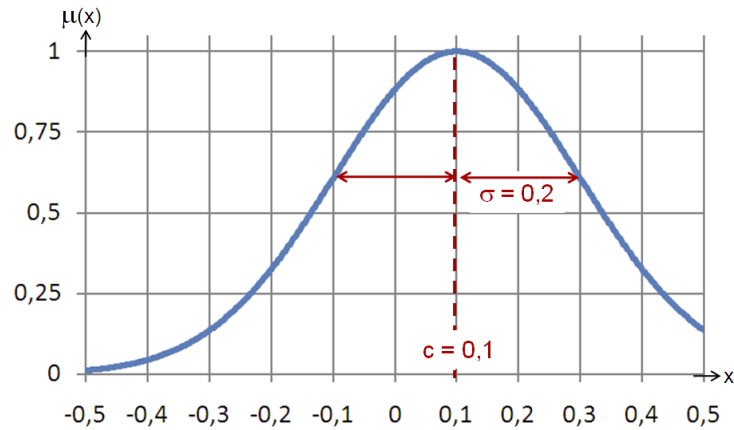
Neben der allgemeinen Glockenkurve mit $\mu_c(\tilde{x}_i) = \left(1 + \left|\frac{\tilde{x}_i - c}{\sigma}\right|^{2b}\right)^{-1}$ (Formel (4.5)) gibt es weitere Funktionen, welche die in Abschnitt 4.2.2 genannten Anforderungen erfüllen und sich demnach als Aktivierungsfunktion für GRNFN-Modelle eignen.

Gaußglocke

Eine naheliegende Möglichkeit ist die Verwendung einer Normalverteilung – auch Gaußsche Glockenkurve oder oft nur Glockenkurve genannt. Diese Funktion verfügt über zwei Parameter. Neben der durch c definierten Mittellage bestimmt der Parameter σ die Breite der Gaußglocke. Eine Eigenschaft der Normalverteilung ist, dass ihr Gesamtintegral den Wert 1 ergibt. Dies wird durch die Multiplikation mit dem Vorfaktor $\frac{1}{\sigma\sqrt{2\pi}}$ erreicht. Dies führt in der Standardform jedoch dazu, dass der Maximalwert mit steigendem σ stetig abnimmt. Um die Forderung $\mu_c(c) = 1$ der Aktivierungsfunktion, d. h. der Aktivierungswert am Zentralpunkt ist eins, zu erfüllen, wird der normierende Vorfaktor weggelassen. Dies führt zu einer Aktivierungsfunktion in Form der Gaußkurve mit

$$\mu_{c,\sigma}(x) = \exp\left(-\frac{1}{2}\left(\frac{x-c}{\sigma}\right)^2\right) \quad (4.7)$$

als Berechnungsvorschrift. Eine solche Gaußglocke sowie die sie definierenden Parameter sind in Abbildung 4.8(a) dargestellt.



(a) Gaußglocke mit definierenden Parametern

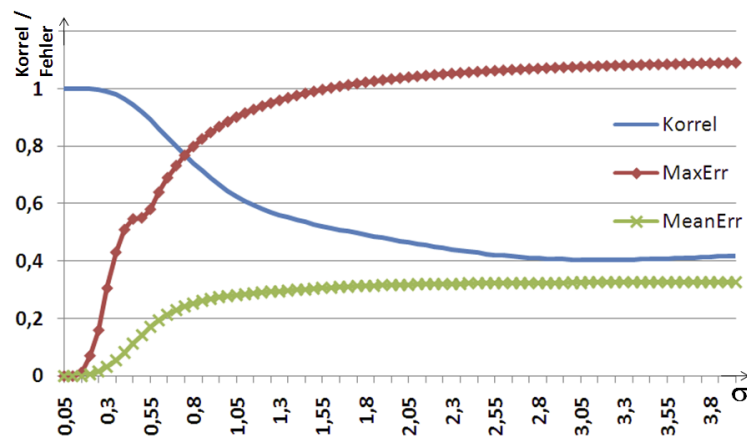
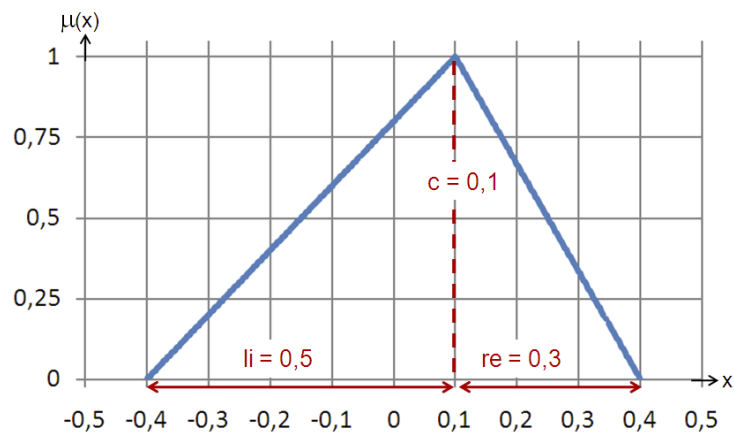
(b) Modellgüte in Abhängigkeit des Parameters σ

Abbildung 4.8: Gaußglocke als Aktivierungsfunktion

Für den Einfluss der Parameter auf die Aktivierungshöhe bei anliegendem Eingangsmuster sowie die Auswirkungen auf die Modellausgabe gilt Ähnliches wie für die allgemeine Glockenfunktion. Der Parameter c definiert die Zentrallage der Glocke und ist durch die Trainingsmuster bestimmt. Der Parameter σ bestimmt – wie bei der allgemeinen Glockenkurve – die Breite der Verteilung und somit die Höhe der Aktivierungsfunktion von benachbarten Stützvektoren. Die Auswirkungen der Variation von σ auf die Modellgüte ist in Abbildung 4.8(b) zu sehen. Entsprechend den Ergebnissen für die allgemeine Glockenkurve fällt mit steigenden σ die Korrelation, während gleichzeitig die Prognosefehler steigen – obwohl die zur Berechnung der Modellgüte verwendete Datenmenge den Trainingsdaten entspricht. Die in Abschnitt 4.3.1 gegebene Begründung gilt für die Gaußkurve entsprechend. Um die Vergleichbarkeit zu ermöglichen, wurden dabei jeweils dieselben Daten verwendet. Als Daten für das GRNFN-Training mit variierten Parametern und die anschließende Ermittlung der dargestellten Modellgüte dient immer die in Abschnitt 3.8 vorgestellte Versuchsreihe mit 98 Zyklen. Bei der Gaußglocke sollte, wie aus Abbildung 4.8(b) ersichtlich,



(a) Dreiecksfunktion mit definierenden Parametern

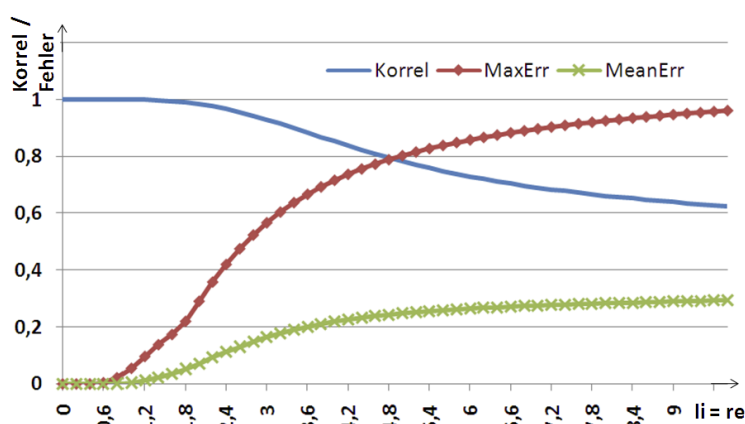

 (b) Modellgüte in Abhängigkeit der Parameter li und re (hier $li = re$)

Abbildung 4.9: Dreiecksfunktion als Aktivierungsfunktion

der Parameter σ für diesen Datenumfang maximal 0,5 betragen. Für größere Versuchsreihen sollte der Wert sukzessive weiter reduziert werden, um für einen Wert nicht zu viele Zugehörigkeitsfunktionen und damit die Mittelung zu vieler Trainingsdaten zu aktivieren (vergleiche Abbildung 4.6).

Bei GRNFN-Modellen wird die Aktivierungsfunktion für die Fuzzifizierung der Eingänge verwendet. Somit liegt es nahe, auch solche Funktionen als Aktivierungsfunktionen zu untersuchen, die bei Fuzzy-Systemen verwendet werden.

Dreiecksfunktion

Eine häufige Funktion, die bei konventionellen Fuzzy-Systemen zur Definition der Fuzzy-Sets verwendet wird, ist die sogenannte „Dreiecksfunktion“. Sie hat insgesamt drei Parameter, welche die Lage (Parameter c) sowie die Breite nach links und rechts bestimmen,

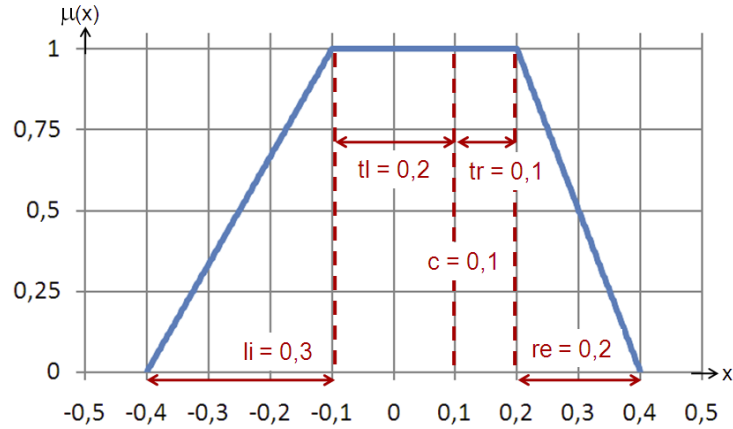
bis die Aktivität 0 erreicht ist. Die Berechnungsformel für eine Dreiecksfunktion mit den genannten Parametern lautet:

$$\begin{aligned}\mu_{li,re}(x) &= \max \left(0; \min \left(1; \frac{1}{li}x + \frac{li-c}{li}; \frac{-1}{re}x + \frac{c+re}{re} \right) \right) \\ &= \max \left(0; \min \left(1; \frac{li+x-c}{li}; \frac{re+c-x}{re} \right) \right).\end{aligned}\quad (4.8)$$

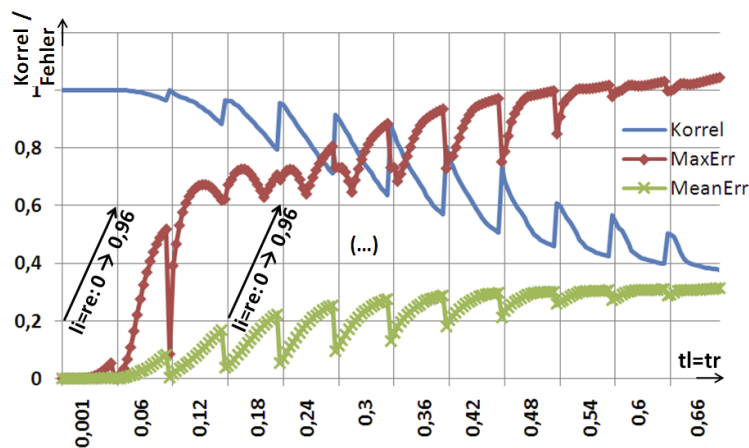
Die Dreiecksfunktion inklusive der definierenden Parameter ist in Abbildung 4.9(a) dargestellt, ebenso wie die Auswirkungen einer Parametervariation auf die Modellgüte, wie es schon bei der Gauß- und allgemeinen Glockenkurve erfolgt ist. Die Parameter li und re wurden dabei auf identische Werte gesetzt, um die Symmetrie der Dreiecksfunktion beizubehalten. Über die Parameter wird dabei wieder die Breite der Aktivierungsfunktion bestimmt, was wiederum dem Parameter σ bei der Gauß- oder allgemeinen Glockenkurve entspricht. Mit steigender Breite sinkt wiederum die Modellgüte.

Gegenüber der Gauß- und der Glockenkurve hat die Dreiecksfunktion einige Besonderheiten. Durch die getrennten Parameter für die linke und rechte Breite lassen sich auch asymmetrische Zugehörigkeitsfunktionen realisieren. Dies ermöglicht insbesondere eine einseitig scharfe Abgrenzung zu Nachbarwerten bei gleichzeitiger Beibehaltung der Interpolationseigenschaft auf der anderen Seite durch Aktivierungswerte deutlich größer als Null. Eine derartige Parametrierung ist jedoch stark von den zu beschreibenden Daten abhängig. Für die Auswertung in Abbildung 4.9(b) wurden aufgrund der mangelnden allgemeinen Aussagekraft keine asymmetrischen Funktionen berücksichtigt. Des Weiteren nimmt die Dreiecksfunktion für Werte, die weiter als die durch die Parameter li bzw. re definierte Breite vom Zentralpunkt entfernt sind, den Wert 0 an. Trifft dieser Fall für eine Kenngröße pro Block in der zweiten Netzwerkschicht (siehe Abbildung 4.4) zu, ergibt sich durch die nachfolgende Multiplikation eine Null und folglich aufgrund der nachfolgenden Division durch Null ein undefiniertes Verhalten in der Modellausgabe. Die zuvor vorgestellten Glockenfunktionen nähern sich zwar asymptotisch der Null, sind aber stets größer, so dass ein solch undefiniertes Verhalten mathematisch nicht auftreten kann. Bei der Realisierung auf einem Rechnersystem liefern jedoch auch die Gauß- und die allgemeine Glockenfunktion aufgrund der beschränkten Genauigkeit der Datentypen float und double den Wert Null, was zum selben Effekt führt.

Eine weitere Eigenschaft der Dreiecksfunktion ist ihre einfache Berechnung: Es sind nur drei Additionen, eine Division und zwei Vergleiche notwendig. Bei der Gauß- oder der allgemeinen Glockenkurve ist jeweils mindestens eine komplexere Funktion für die Berechnung der e-Funktion oder der Potenz notwendig. Da die Aktivierungsfunktion für die Berechnung der Modellausgabe zu einem einzelnen Muster $N \cdot L$ Mal aufgerufen wird, macht sich hier jede Optimierung hinsichtlich Rechenzeit deutlich bemerkbar.



(a) Trapezfunktion mit definierenden Parametern



(b) Modellgüte in Abhängigkeit der Parameter, mit $li = re$ und $tl = tr$

Abbildung 4.10: Trapezfunktion als Aktivierungsfunktion

Die Dreiecksfunktion – ebenso wie die nachfolgende Trapezfunktion – ermöglicht es, die Regelbasis „normaler“ Fuzzy-Systeme direkt in GRNFN-Modellen abzubilden (siehe Abschnitt 4.4.4). Ein wesentlicher Punkt von Fuzzy-Systemen ist die Definition der verwendeten Fuzzy-Sets. Sehr häufig haben diese eine dreieckige Form, so dass diese Art von Zugehörigkeitsfunktion sehr häufig ist. Des Weiteren kann durch die Verwendung solcher Funktionen, die stückweise aus linearen Abschnitten zusammengesetzt sind, die Einhaltung der Ruspini-Partitionierung einfach gewährleistet werden. Eine Ruspini-Partitionierung sagt aus, dass für jeden Wert x die Summe aller Zugehörigkeitsfunktionen an dieser Stelle 1 ergibt, d. h. $\sum_{k=1}^L \mu^k(x) = 1$ gilt. Diese Forderung wird häufiger an Fuzzy-Sets gestellt, da sich daraus weitere Eigenschaften für das gesamte Fuzzy-System ableiten lassen [CDM09]. Bei Glockenkurven ist diese Bedingung in der Regel nicht erfüllt.

Trapezfunktion

Eine weitere Funktion, die häufig bei Fuzzy-Systemen zum Einsatz kommt, ist die Trapezfunktion. Wesentliche Parameter legen dabei die Lage der Funktion, die Breite des Plateaus sowie die Steigung der Flanken fest. Mittels dieser Parameter ist die Definition jedoch mehrdeutig, da die Lage des Stützpunktes und Zentrums nicht eindeutig festgelegt ist. Deswegen wird die Parametrierung der Plateaubreite in einen linken und rechten Anteil gesplittet. Der Wert der Zugehörigkeitsfunktion für jeden Punkt x kann über folgende Formel berechnet werden:

$$\begin{aligned} \mu_{li,tl,tr,re}(x) &= \max \left(0; \min \left(1; \frac{1}{li}x + \frac{li + tl - c}{li}; \frac{-1}{re}x + \frac{c + tr + re}{re} \right) \right) \\ &= \max \left(0; \min \left(1; \frac{(li + tl) + x - c}{li}; \frac{(tr + re) + c - x}{re} \right) \right). \end{aligned} \quad (4.9)$$

Die Trapezfunktion inklusive ihrer so definierenden Parameter ist in Abbildung 4.10(a) dargestellt, die Modellgüte in Abhängigkeit der Parameter in Abbildung 4.10(b). Dabei sind – wie schon bei der Dreiecksfunktion – die Parameter symmetrisch gehalten. Somit wird ein Vergleich mit der Gauß- und allgemeinen Glockenfunktion möglich. Zudem ist die über die Parameter darstellbare Unsymmetrie in ihrer Anwendung stark von den jeweiligen Trainingsdaten abhängig. Bezüglich der Parameter und deren Auswirkungen ergeben sich wiederum Analogien zur Glockenfunktion: Die Plateaubreite, bestimmt durch tl und tr , entspricht σ und bestimmt die Breite. Die Parameter li und re bestimmen ähnlich b die Steilheit und damit die Grenzscharfe der Zugehörigkeit.

Wie die Dreiecksfunktion ermöglicht die Trapezfunktion die Abbildung von Fuzzy-Sets, von regulären Fuzzy-Systemen sowie die Erfüllung der Ruspini-Partitionierung. Auch bezüglich der notwendigen Rechenoperationen bei der Implementierung ist die Trapezfunktion mit drei Vergleichen, zwei Additionen und einer Division günstig.

4.4 Eigenschaften von GRNFN-Modellen

Aus dem in Abschnitt 4.2 beschriebenen Aufbau von GRNFN-Modellen ergeben sich einige interessante Eigenschaften, auf die im Folgenden kurz eingegangen wird.

4.4.1 Rechen- und Speicheraufwand

Bei vielen Modellierungsverfahren spielt der mit zunehmender Eingangsdimension überproportional – oft sogar exponentiell – steigende Aufwand bezüglich Rechenzeit und Speicherplatz eine Rolle. Dies wird allgemein als *curse of dimensionality* bezeichnet [Bis06, HTF01].

Beispielsweise steigt bei Feed-Forward-Netzen und MLP-Modellen (siehe Abschnitt 2.4) die Anzahl der Gewichte und damit auch der Speicher- und Rechenaufwand bei Training und Anwendung in der Regel quadratisch². So ergibt sich für ein Modell mit N Eingangsnuroneuronen und $\frac{N}{2}$ Neuronen in der (ersten) versteckten Schicht die Anzahl der Gewichte proportional zu N^2 .

GRNFN-Modelle besitzen in der zweiten Schicht nur eindimensionale Aktivierungsfunktionen. Je nach gewählter Art sind diese auch mit sehr wenigen Rechenoperationen verbunden (siehe Abschnitt 4.3). Die Anzahl der Aktivierungsfunktionen ist proportional zur Eingangsdimension N , so dass auch bei der Anwendung der notwendige Aufwand der Ordnung $\mathcal{O}(N)$ entspricht und nicht $\mathcal{O}(N^2)$ wie bei MLP-Modellen. Für das Training, welches in einem Durchlauf erfolgen kann, ist daher auch nur ein Aufwand proportional zu N notwendig.

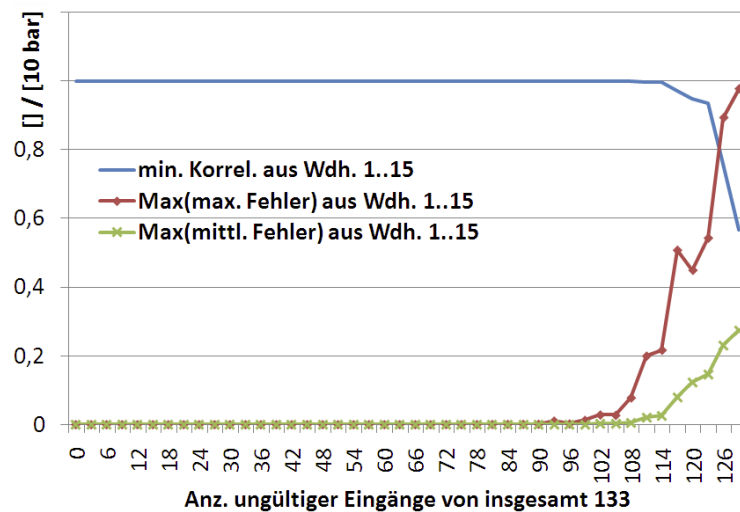
Der notwendige Speicherbedarf für GRNFN-Modelle hängt von drei Faktoren ab: Der Eingangsdimension N , der Anzahl der Lernmuster L sowie der Art, in der die Parameter für die Aktivierungsfunktionen festgelegt werden. Der Speicher wächst jeweils proportional zu N und L , da hierdurch die Anzahl der Blöcke in der zweiten Modellschicht und die Lageparameter c_i^k , die für jede Aktivierungsfunktion spezifisch sind, festgelegt ist. Alle weiteren $(P-1)$ Parameter der Aktivierungsfunktionen können entweder global oder für jede Aktivierungsfunktion separat festgelegt werden. Dementsprechend kommt ein konstanter (und vernachlässigbarer) Speicherbedarf von einem (bei Gauß) bis drei Werten (Trapez) hinzu, bei separater Festlegung wächst der Speicherbedarf insgesamt statt mit $L \cdot N$ mit $(P-1) \cdot L \cdot N$, was aber immer noch $\mathcal{O}(N)$ entspricht.

Bezüglich der Approximation ist festzustellen, dass Neuro-Fuzzy-Systeme mit GRNFN-Modellen die Approximationseigenschaften von RBF-Netzen aufweisen. Jedoch leiden sie deutlich weniger unter dem *curse of dimensionality*, da nur eindimensionale Aktivierungsfunktionen anstelle N -dimensionaler RBF-Funktionen verwendet werden [BOPT07].

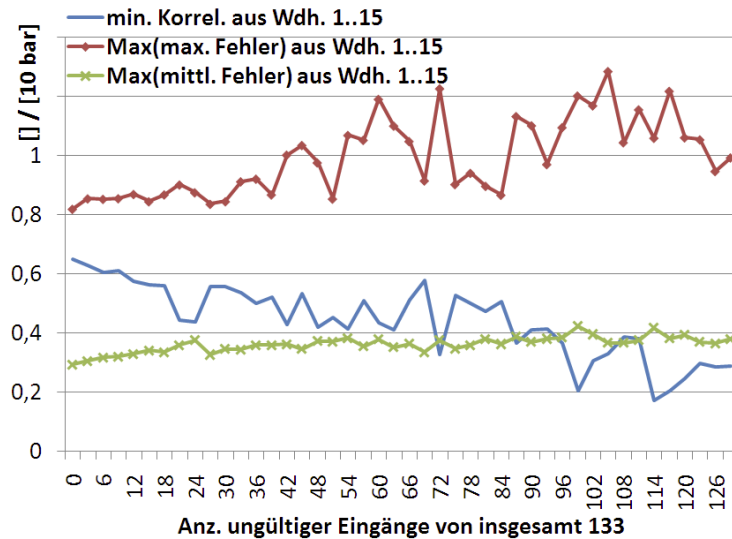
4.4.2 Robustheit und Fehlertoleranz

Bei GRNFN-Modellen geht jede Komponente des Eingangsvektors nach der Fuzzifizierung multiplikativ in die Gesamtaktivierung Φ^k jedes Blocks k , $k \in [1..L]$ ein. Für jedes der L Trainingsmuster, das jeweils einen Block definiert, ist der Aktivitätswert der einzelnen Komponenten $\varphi_i^k = 1$ und damit ergibt sich auch die Gesamtaktivität zu $\Phi^k = 1$. Der Wegfall einzelner Komponenten bei der Produktbildung ändert daran nichts. Bei Eingangsvektoren, die keinem der Trainingsvektoren hinreichend ähnlich sind, und die somit als unbekannt gelten, ist davon auszugehen, dass sich Unterschiede nicht nur in einzelnen, sondern in mehreren Komponenten widerspiegeln. Die Gesamtaktivität eines Blocks, die sich

²Ausnahmen sind Modelle, deren Anzahl an Neuronen in den versteckten Schichten unabhängig von der Eingangsdimension festgelegt ist.



(a) Trainingsdaten



(b) Verifikationsdaten

Abbildung 4.11: Veränderung der Modellgüte bei zunehmender Anzahl von ungültigen Eingängen

als Produkt der Einzelaktivitäten ergibt, ist demnach nahe Null, da mehrere niedrige Aktivitäten miteinander multipliziert werden. Auch hier ändert der Wegfall einzelner Faktoren nichts Entscheidendes.

Folglich ist es bei trainierten GRNFN-Modellen möglich, in der Anwendung einen gewissen Anteil der Eingangskomponenten (zeitweise) wegfallen zu lassen. An der ermittelten Gesamtaktivität ändert dies wenig. Voraussetzung ist jedoch, dass in den Eingängen eine gewisse Redundanz und Kopplung vorhanden ist. Beim Einsatz der Modelle zur Prozessüberwachung ist dies in der Regel jedoch gegeben, da die einzelnen Messgrößen physikalisch über den Prozess gekoppelt sind. Meist ist es nicht möglich, dass die einzelnen aus Prozesssignalen abgeleiteten Kenngrößen, die dem Netz als Eingangswerte dienen, unabhängig

voneinander sind. Beispielsweise bedingen beim Ultraschallschweißen hohe Wegsignale am Prozessanfang zwangsläufig, dass auch die Endwerte deutlich höher sind, und veränderte Kräfte wirken sich im Wegverlauf, der Schweißzeit und der Leistungsaufnahme aus (siehe Abschnitt 2.1).

Diese Eigenschaft lässt sich zum Aufbau von robusten und fehlertoleranten Prozessmodellen ausnutzen. Die einzelnen Eingangskomponenten werden dabei aus verschiedenen Prozesssignalen abgeleitet, welche in der Praxis von kurzzeitigen Störungen oder gar Ausfällen betroffen sein können. Fallen einzelne Signale und damit Kenngrößen temporär aus, können diese bei der Anwendung des GRNFN-Modells ausgeblendet werden und dennoch ist die Prognose anhand der verbliebenen Kenngrößen möglich. Dies ist ein deutlicher Vorteil gegenüber MLP-Modellen. Diese benötigen immer den kompletten Eingangsdatensatz, da über die gewichtete Summenbildung in den Neuronen der versteckten Schichten alle Eingänge direkt miteinander verknüpft werden. Einzelne Eingangsgrößen können folglich nicht einfach entfallen oder ausgeblendet werden.

Auch gegenüber gestörten Eingangssignalen sind GRNFN-Modelle robust und weisen Vorteile gegenüber MLP-Modellen auf: Wenn eine Störung zu deutlichen Veränderungen einzelner Kenngrößen führt, können diese bei einem GRNFN ausgeblendet und die Netzprognose mit den restlichen Größen erstellt werden. Werden solche Kenngrößen nicht ausgeblendet, führt ein deutliches Überschreiten der Normierungsgrenzen zu einer Einzelaktivierung von $\varphi_i \approx 0$, was sich durch die Multiplikation in einer Gesamtaktivierung $\Phi \approx 0$ auswirkt. Der ungültige oder unbekannte Zustand kann damit sicher erkannt werden und es führt nicht zu unvorhersagbaren Netzausgaben. Anders verhält sich dies jedoch bei Neuronalen Netzen in Form von MLP. Zum einen liefert bei einem MLP jedes Neuron aufgrund der meist sigmoiden Aktivierungsfunktion einen gültigen Wert, der bis zur Netzausgabe weiter verrechnet wird – unabhängig davon, wie weit der Wert von den Trainingsdaten entfernt ist. Die Netzausgabe ist somit ungültig, ohne dass dies vom Modell erkannt werden kann.

Die Auswirkung einer zunehmenden Anzahl an ungültigen Modellinputs auf die Modellgüte von GRNFN-Modellen ist in Abbildung 4.11 dargestellt. Der Analyse liegt dieselbe Versuchsreihe wie für die Darstellung der Aktivierungsfunktionen in Abschnitt 4.3 zugrunde. Für das Modelltraining – Glockenkurve als Aktivierungsfunktion, $\sigma = 0,2$ und $b = 0,5$ – wurden 80 % der verfügbaren Daten verwendet, so dass 20 % der Daten als unabhängige Verifikationsdaten verbleiben. Die Daten bestehen aus 133 Kenngrößen aus den Schweißsignalen (siehe Abschnitt 5.2.2) sowie dem ermittelten Berstdruck als Zielgröße. Die Anzahl der ungültigen Kenngrößen wurde schrittweise von 0 auf 126 – d. h. fast alle Kenngrößen – erhöht und dann die Modellgüte anhand von 15 Durchläufen berechnet. In jedem Durchlauf wurden dabei die ungültigen Inputs zufällig ausgewählt. Die resultierende Modellgüte für die Trainings- sowie Verifikationsdaten (79 bzw. 19 Zyklen) ist in Abbildung 4.11 dar-

gestellt. Über die 15 Durchläufe mit derselben Anzahl an (unterschiedlichen) ungültigen Inputs wurde für die Korrelation der minimale und für die Fehler der jeweils maximal auftretende Wert verwendet. Es ist deutlich zu erkennen, dass die Modellgüte erst abnimmt, wenn ein signifikanter Anteil an Modelleingängen ungültig ist. Der mittlere Fehler bei den Verifikationsdaten liegt bei ca. 11 % des Festigkeitsniveaus. Dies liegt innerhalb der üblichen Schwankungen bei der Festigkeitsprüfung von Schweißnähten an Kunststoffteilen, die häufig Standardabweichungen von ca. 20 % der mittleren Festigkeit aufweisen (siehe Abschnitt 2.3.1).

Die Eigenschaft, Eingänge ohne ein neues Netztraining abschalten zu können, kann auch für die Kenngrößenselektion verwendet werden. Dabei wird ein GRNFN-Modell mit den Gesamtdaten trainiert. Anschließend werden nach einem Algorithmus iterativ definierte Eingänge abgeschaltet und die Modellgüte für die verbliebenen Größen ermittelt, bis ein bestimmtes Abbruchkriterium (Fehler bei Verifikationsdaten, vorgegebene Anzahl an Eingängen, ...) erreicht ist. Ein mögliches Vorgehen, dies zu bewerkstelligen, ist der Einsatz von Genetischen Algorithmen (GA, siehe Abschnitt 5.3.4).

4.4.3 Interpretation als Fuzzy-Regelbasis

Die Architektur von GRNFN-Modellen (siehe Abbildung 4.4) stimmt mit der Architektur eines Fuzzy-Systems nach WANG/MENDEL überein, welche aus den Elementen Fuzzifizierer, Fuzzy-Regelbasis, Inferenz und Defuzzifizierer besteht [WM92]. Die Regelbasis besteht dabei aus M Regeln folgender Form:

$$R_j : IF x_1 is A_1^j and x_2 is A_2^j and \dots and x_n is A_n^j, THEN z is B^j \quad (4.10)$$

mit $j = 1, 2, \dots, M$. Weiter sind x_i ($i = 1, 2, \dots, n$) die Eingangsgrößen des Fuzzy-Systems, z die Ausgangsvariable sowie A_i^j und B^j die durch Zugehörigkeitsfunktionen charakterisierten Fuzzy-Sets der linguistischen Variablen. Diese Form der Regeln entspricht einem Takagi-Sugeno-Fuzzy-Inferenz-System nullter Ordnung (siehe Abschnitt 2.4.3 und [TS85]). Mit der Multiplikation als Verknüpfungsoperator ergibt sich damit für GRNFN-Modelle folgende Analogie:

- Die normierten Eingangsgrößen \tilde{x}_i mit $i \in [1..N]$ entsprechen den x_i des Fuzzy-Systems,
- jeder der L Blöcke, gebildet durch die Trainingsdaten bzw. Stützvektoren, entspricht einer Regel der Regelbasis,
- der Aktivitätswert φ_i^k entspricht dem Grad der Zugehörigkeit der i -ten Eingangsvariable zum Fuzzy-Set A_i^j , wobei die Aktivierungsfunktion μ_i^k die definierende Zugehörigkeitsfunktion des Fuzzy-Sets bildet,

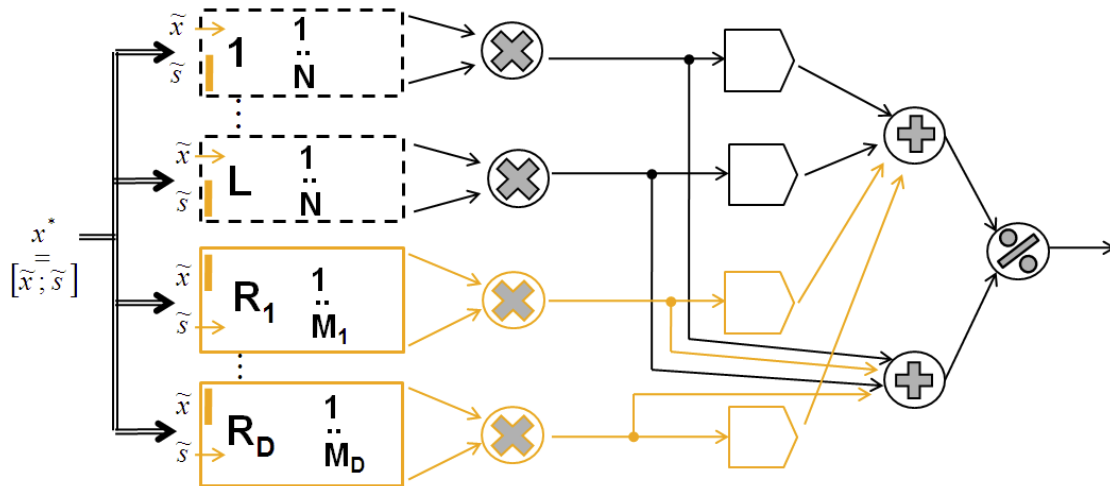


Abbildung 4.12: GRNFN-Modell mit manuell ergänzten (Fuzzy-)Regeln

- die zugeordnete Ausgangsgröße y^k mit $k \in [1..L]$ entspricht der B^j ,
- die Defuzzifizierung erfolgt durch Summation und anschließende Division der Gesamaktivierungen Φ^k (analog zu WANG/MENDEL [WM92]).

Das anhand von Lerndaten erstellte GRNFN-Modell kann somit direkt als Fuzzy-System mit Regelbasis interpretiert werden. Dies kann auch das Verständnis von Zusammenhängen durch Interpretation der Regeln der enthaltenen Regelbasis erleichtern.

4.4.4 Gezielte Erweiterung der Fuzzy-Regelbasis

Der wesentliche Vorteil von Fuzzy-Systemen liegt in der Fähigkeit, Erfahrungen und vorhandenes Wissen in Form von Regeln zu beschreiben und in ein formales System und damit auch in Computerprogramme überführen zu können. Modelle, die anhand von Trainingsdaten erstellt werden, können solche Informationen nicht integrieren und erschweren die Interpretierbarkeit der durch sie abgebildeten Zusammenhänge.

Auch GRNFN-Modelle werden anhand von Trainingsdaten erstellt (siehe Abschnitt 4.2), können aber, wie in Abschnitt 4.4.3 beschrieben, als Fuzzy-System mit Regelbasis interpretiert werden. Wenn für einen zu modellierenden Prozess neben den Trainingsdaten, denen Messdaten zugrunde liegen, weitere Erfahrungen (z. B. der Prozessbediener) bezüglich Zusammenhängen und Einflüssen vorhanden sind, können diese in das GRNFN-Modell integriert werden. Dazu wird das vorhandene Wissen mit Regeln beschrieben, die wie Gleichung (4.10) aufgebaut sind. Die Regeln können aber komplett aus eigenen Variablen bestehen, die im bisherigen Datenvektor x nicht vorhanden sein müssen. Dazu sind zuerst geeignete linguistische Variablen $s = [s_1, \dots, s_G]$ sowie deren Aufteilung in Fuzzy-Sets zu bilden. Die Beschreibung der Fuzzy-Sets geschieht über Zugehörigkeitsfunktionen μ_g^d , welche dann im

GRNFN-Modell als Aktivierungsfunktionen dienen. Vor diesem Hintergrund sind die in Abschnitt 4.3 beschriebenen Dreiecks- und Trapezfunktionen hilfreich. Die auf diese Weise definierten D Fuzzy-Regeln werden als zusätzliche Blöcke in Schicht 2 des GRNFN-Modells eingefügt. Ein so erweitertes Modell ist in Abbildung 4.12 dargestellt. Die farbigen Elemente im unteren Bereich entsprechen den vom Benutzer definierten D Fuzzy-Regeln. Die Anzahl der verknüpften Variablen kann dabei in den einzelnen Regeln verschieden sein – angedeutet durch die verschiedene Anzahl an Zugehörigkeitsfunktionen $M_1 \dots M_D$. Die Auswahl der in einer Regel verwendeten Variablen erfolgt über einen geeigneten „Filter“. Der neue Eingabevektor x^* besteht aus dem um die (normierten) linguistischen Variablen \tilde{s} ergänzten Eingangsvektor \tilde{x} . Die Verarbeitung dieses vergrößerten Vektors in den ersten L Blöcken stellt kein Problem dar, da die neuen linguistischen Variablen ausgefiltert und ausgeblendet werden können, indem sie als ungültig markiert werden (siehe Abschnitt 4.4.2). Bei den benutzerdefinierten Regeln gilt Entsprechendes für die Ausblendung der Komponenten von \tilde{x} sowie der unbenutzten Variablen aus \tilde{s} .

4.4.5 Zusammenfassen von Trainingsdaten

Größe, Speicherbedarf und Berechnungsdauer eines GRNFN-Modells steigen proportional zur Anzahl der Trainingsmuster. Ab einer gewissen Größe kann eine Reduktion der Trainingsdaten und damit der Modellgröße sinnvoll sein. Dies kann dadurch erreicht werden, dass die Trainingsdaten geeignet vorverarbeitet und zu Clustern zusammengefasst werden. Dies kann durch den Einsatz entsprechender Clustering-Verfahren wie kMeans oder LVQ erreicht werden (siehe [HTF01]). Die einem Cluster zugeordneten Datenvektoren können dann durch den cluster-definierenden Vektor ersetzt werden – die Verfügbarkeit eines geeigneten zugeordneten Zielwertes vorausgesetzt. Während die zugeordneten Datenvektoren einen gewissen Bereich im Originalraum abbilden, geht dieses Raumvolumen beim Ersetzen durch einen einzelnen Datenpunkt verloren. Durch eine geeignete Anpassung der Parameter der Zugehörigkeitsfunktionen kann die ursprüngliche Streubreite der Ausgangsdaten jedoch im GRNFN-Modell – zumindest teilweise – wieder abgebildet werden.

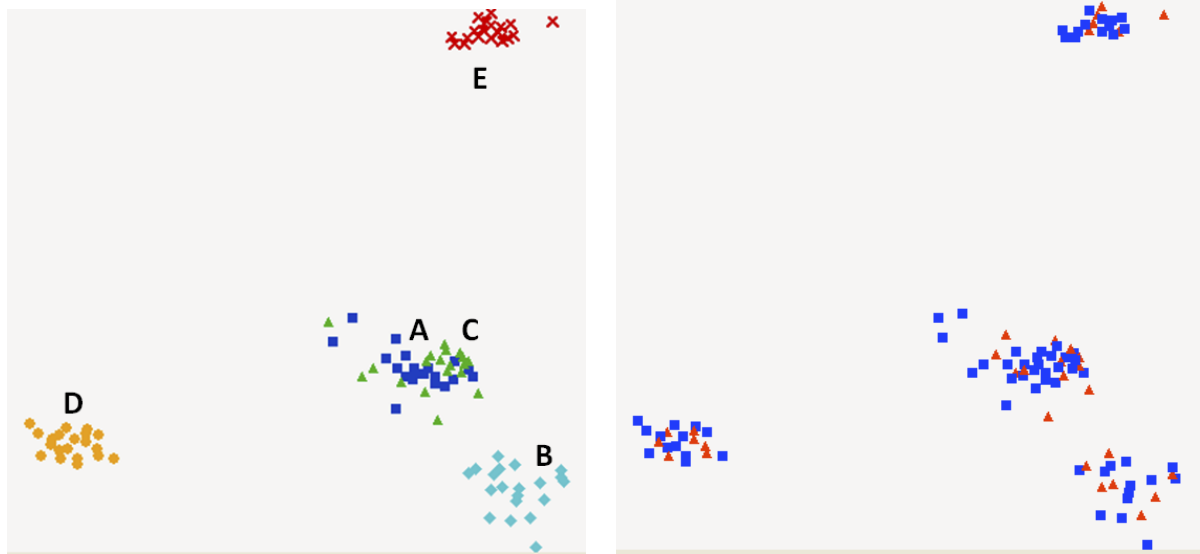
Für die Interpolation verwendet das GRNFN-Modell den Stützvektoren zugeordnete Ausgangsgrößen in der vierten Netzschicht. Der Zielwert für die ermittelten Cluster kann entweder durch Mittelung oder Medianbildung der zugeordneten Daten abgeschätzt werden, was jedoch je nach Streubreite der Daten zu entsprechenden Fehlern führt. Alternativ kann die Zielgröße festgelegt werden, indem die Forderung nach „klassenreinen“ Clustern als Randbedingung in die Clusterverfahren einfließt. Dies kann zwar die Überführung von kontinuierlichen Merkmalswerten in entsprechende Klassen erfordern, stellt jedoch kein Problem dar, da dies bei Fuzzy-Systemen durch die notwendige Definition von Fuzzy-Sets bereits für andere Größen impliziert ist.

Eine Clustering der Trainingsdaten erfolgt anhand einer Metrik bezüglich der Eingangsvektoren x in \mathbb{R}^N . Um für zusammengefasste Cluster sinnvolle Zielgrößen ableiten zu können, ist Voraussetzung, dass sich die Cluster auch in Clustern bezüglich gleicher oder ähnlicher Zielgrößen abbilden. Anders ausgedrückt: Die Zielgrößen sollten eine gewisse Homogenität aufweisen, was bedeutet, dass ähnliche Eingangsvektoren auch ähnliche Zielgrößen haben. Die Ähnlichkeit wird dabei über eine Metrik oder ein Distanzmaß definiert, wobei häufig der euklidische Abstand zweier Vektoren verwendet wird. Dabei gilt, dass die Ähnlichkeit umso größer ist, je kleiner der ermittelte Abstand ist. Zur Visualisierung der Ähnlichkeit von N -dimensionalen Vektoren kann beispielsweise das Sammons-Mapping verwendet werden (siehe Abschnitt 4.1). Eine Homogenität und Stetigkeit der Trainingsdaten ist auch Voraussetzung für ein stabiles Prozessmodell und wird maßgeblich durch die Zusammensetzung des Eingangsvektors und die Wahl seiner Komponenten bestimmt. Eine Optimierung erfolgt mittels einer geeigneten Kenngrößenselektion (*feature selection*). Diese ist stark prozessabhängig; einige mögliche Verfahren werden in Abschnitt 5.3 näher erläutert.

4.4.6 Dynamische Modellanpassung

Die Struktur von GRNFN-Modellen ermöglicht es, dynamische Modellanpassungen vorzunehmen und die modeldefinierenden Daten zu verändern. Wie schon in Abschnitt 4.4.4 beschrieben, können neben den im einmaligen Trainingsdurchgang integrierten Daten weitere Blöcke in der zweiten Netzschicht hinzugefügt werden. Durch nachträgliches Hinzufügen können dem Modell neue oder erweiterte Zusammenhänge beigebracht werden. Entweder kann das Hinzufügen neuer Muster zu festgelegten Zeitpunkten oder ereignisbasiert erfolgen. Ist ein angelegtes Eingangssignals durch das GRNFN-Modell als „unbekannt“ klassifiziert worden, kann dieses Eingangsmuster und damit der Zustand, der bis dato unbekannt Informationen enthält, in das Modell integriert werden. Damit wird eine prozessbegleitende Adaption und ein sogenanntes *active learning* möglich. Bei jedem Hinzufügen von neuen Daten müssen jedoch die entsprechenden Zielgrößen verfügbar sein. So ist es möglich, dass ein Modell an Trends und längerfristige Prozessdriften, wie sie beispielsweise durch Verschleiß oder Erwärmung auftreten können, angepasst werden kann. Entsprechend können auch definierte Zustände „vergessen“ und aus dem Modell entfernt werden. Bei Prozessen, die langsamen Trends oder Schwebungen unterliegen, ist es damit auch möglich, den aktuellen Zustand im Kontext des aktuellen Betriebspunkts zu überwachen, indem regelmäßig neue Datensätze hinzugefügt und die ältesten gelöscht werden.

Beim Hinzufügen neuer Daten kann es vorkommen, dass einige Komponenten außerhalb der bisherigen Normierungsgrenzen aus der ersten Netzschicht liegen. Entweder können die Daten dann über die bisherigen Grenzen extrapoliert – vorausgesetzt, dass dadurch nicht der Definitionsbereich von in Schicht 2 verwendeten Aktivierungsfunktionen überschritten



(a) Darstellung der verschiedenen Prozesseinstellungen
 (b) Aufteilung der Daten in Lern- und Testmuster
 (blau bzw. rot)

Abbildung 4.13: Darstellung des Sammons-Mapping der bisherigen Beispieldaten mit unterschiedlichen zugeordneten Zielgrößen

wird – oder die Normierungsgrenzen neu bestimmt werden. Dies zieht dann jedoch eine entsprechende Anpassung der Lageparameter c aller Aktivierungsfunktionen μ_n^k nach sich, die den mit den aktuellen Grenzen normierten Werten der ursprünglichen Trainingsdaten entsprechen. Sind die Aktivierungsfunktionen nicht global definiert, kann es nach größeren Modelländerungen notwendig sein, neben c die weiteren Parameter der Aktivierungsfunktionen neu einzustellen oder gar ein Neutraining durchzuführen. Allerdings ist es dazu nicht notwendig, die bisherigen Trainingsdaten separat zur Verfügung zu haben, da diese implizit durch die Lageparameter c sowie die Werte in der Netzschicht 4 im GRNFN-Modell verfügbar sind.

4.5 Prognosegüte

In Abbildung 4.13 ist die Projektion der in Abschnitt 3.8 beschriebenen Beispieldaten dargestellt. Es ist klar zu erkennen, dass sich die Vektoren in mehrere Cluster aufteilen. In Abbildung 4.13(a) ist den Vektoren die jeweilige Prozesseinstellung zugeordnet, wobei gleiche Farben und Marker für dieselbe Einstellung stehen. Es ist deutlich zu sehen, dass die einzelnen Cluster jeweils eine Prozesseinstellung widerspiegeln. Dies ist physikalisch auch plausibel, da sich eine Änderung der Prozesseinstellung in veränderten Signalen und folglich in daraus abgeleiteten Zustandsvektoren auswirkt. Die Zustandsvektoren einer Einstellung sollten sich wiederum ähnlich sein. Durch eine solche Darstellung kann auch die Stabilität

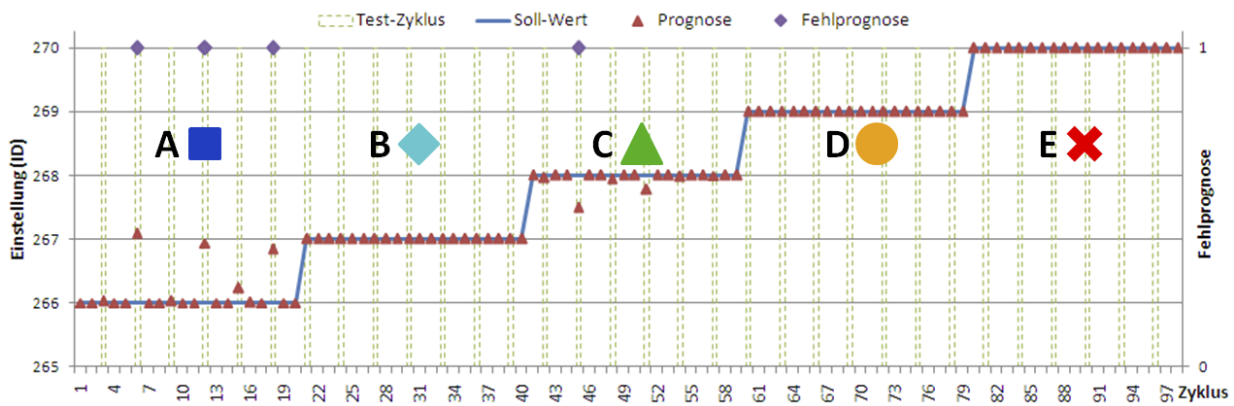


Abbildung 4.14: Vergleich zwischen realer Einstellung und GRNFN-Prognose für Lern- und Testdaten

einer Prozesseinstellung dargestellt und durch entsprechende Farbcodierung eines Qualitätsmerkmals visuell beurteilt werden.

In Abbildung 4.13(b) ist die Aufteilung in Lern- und Testmuster dargestellt, wie sie im folgenden Abschnitt 4.6 für die Analyse der GRNFN-Fähigkeiten zur Prognose von Merkmalswerten und der Unbekannt-Klassifikation verwendet werden. Die Verteilung der Bildpunkte ist identisch mit der Verteilung links, da sie sich einzig nach den Zustandsvektoren richtet und unabhängig vom Zielmerkmal ist, welches durch die Farbgebung codiert wird. Die einbezogenen 98 Zyklen der Versuchsdaten wurden aufgeteilt, indem jeweils zwei Zyklen für das Training und jeder dritte Zyklus ausschließlich für die Verifikation eines GRNFN-Modells verwendet wurden. Die Trainingsmenge umfasst somit 66 Zyklen, die Testmenge 32 Zyklen. Die Aufteilung ist notwendig, um die Leistungsfähigkeit an einem unabhängigen Datensatz ermitteln zu können. Da GRNFN-Modelle datenbasiert arbeiten, ist es durch eine geeignete Wahl der Parameter der Aktivierungsfunktionen möglich, dass die Trainingsdaten exakt abgebildet werden. Über die Generalisierungsfähigkeit und somit die Eignung als Prognosemodell kann so keine Aussage getroffen werden.

Zur Darstellung der Prognose- und Generalisierungsfähigkeit wurde anhand der 66 genannten Zyklen ein GRNFN-Modell erstellt, wobei als Zielgröße die als Zahl codierte Maschineneinstellung dient. Als Aktivierungsfunktion wurde die Glockenfunktion mit den Parametern $\sigma = 0,05$ und $b = 0,5$ verwendet. Mit diesem Netz wurde anschließend für alle 98 verfügbaren Datenzyklen eine Prognose – also eine vom GRNFN-Modell getroffene Zuordnung der Prozesseinstellung anhand der aus den Kenngrößen gebildeten Zustandsvektoren – berechnet. Das Ergebnis ist in Abbildung 4.14 dargestellt. Die tatsächliche Einstellung der Zyklen ist als durchgehende Treppenfunktion dargestellt. Eine Zuordnung zu Abbildung 4.13(a) und Tabelle 3.1 ist durch die Buchstaben sowie farbigen Marker gegeben. Die Prognose der Einstellung für jeden Zyklus ist als rotes Dreieck dargestellt; die gestrichelten Balken bei jedem dritten Zyklus kennzeichnen die Testdaten. Anhand des Diagramms lassen

sich mehrere Aussagen und Eigenschaften bezüglich der Prognose mit GRNFN-Modellen ableiten:

- Während der zugeordnete Zahlenwert der Einstellungen ganzzahlig ist, ergeben sich als Prognosewerte reelle Zahlen und somit beliebige Zwischenwerte. Ursächlich ist die zugrunde liegende Berechnungsvorschrift nach Formel (4.3). Eine Zuordnung des Prognosewertes zu einer Einstellung kann als einfache Rundung zur nächstgelegenen Ganzzahl erfolgen.
- Die Einstellungen B, D und E werden zu 100 % richtig zugeordnet – also auch die in den 59 Zyklen enthaltenen 20 Testzyklen.
- Die Prognose weicht nur in den Einstellungen A und C bei insgesamt 6 der 39 Zyklen ab. Daraus ergeben sich nach oben beschriebener Rundung noch vier Prognosefehler, die in Abbildung 4.14 durch Rauten am oberen Rand separat gekennzeichnet sind. Beim Vergleich mit Abbildung 4.13(a) ist zu erkennen, dass sich genau die beiden Einstellungen A und C durch die Projektion nicht trennen lassen, sie sich also teils sehr ähnlich sind.
- Die Ähnlichkeit zweier Vektoren ist gleichbedeutend mit einem geringen Abstand der Zustandsvektoren. Dies ist wiederum nur möglich, wenn sich die einzelnen Vektorkomponenten kaum unterscheiden. Bei der Berechnung des GRNFN-Modells für Zustandsvektoren der Einstellung A liefern folglich Blöcke, denen Trainingsdaten der Einstellung A oder C zugeordnet sind, nennenswerte Aktivierungswerte Φ^k . Alle diese Aktivierungswerte fließen in die Berechnung nach Gleichung (4.3) ein. Sie werden mit den Zielwerten der Trainingsdaten verknüpft und gemittelt.
- Die vier falsch zugeordneten Zyklen der Einstellungen A und C werden nicht der jeweils anderen, nahe liegenden Einstellung C bzw. A zugeordnet, sondern der Einstellung B, die einen deutlichen Abstand zu A und C aufweist. Ursache hierfür ist die Codierung der Einstellungen mit Zahlen, wobei Einstellung B mit einer Zahl zwischen der für A und C codiert wurde. Die zuvor erläuterte Mittelung bei der Interpolation durch das GRNFN-Modell führt dann zu der im ersten Augenblick unplausiblen Zuordnung. Die Ursache liegt demnach nicht im GRNFN-Modell, sondern in der ungeeigneten Codierung der definierten Zielklassen. Bei diskreten Merkmalen mit festgelegten Klassen sollte die Codierung der Klassen entsprechend ihrem Abstand oder einer geordneten Reihenfolge erfolgen. Zur Analyse dieses Problems kann die Projektion durch Sammons-Mapping gute Dienste leisten.
- Die Werte, mit welchen die einzelnen Einstellungen codiert werden, können frei gewählt werden. Somit können auch die Zahlenwerte der Einstellungen A und B vertauscht werden. D. h. im Beispiel wird allen Zyklen von Einstellung A jetzt 267 statt

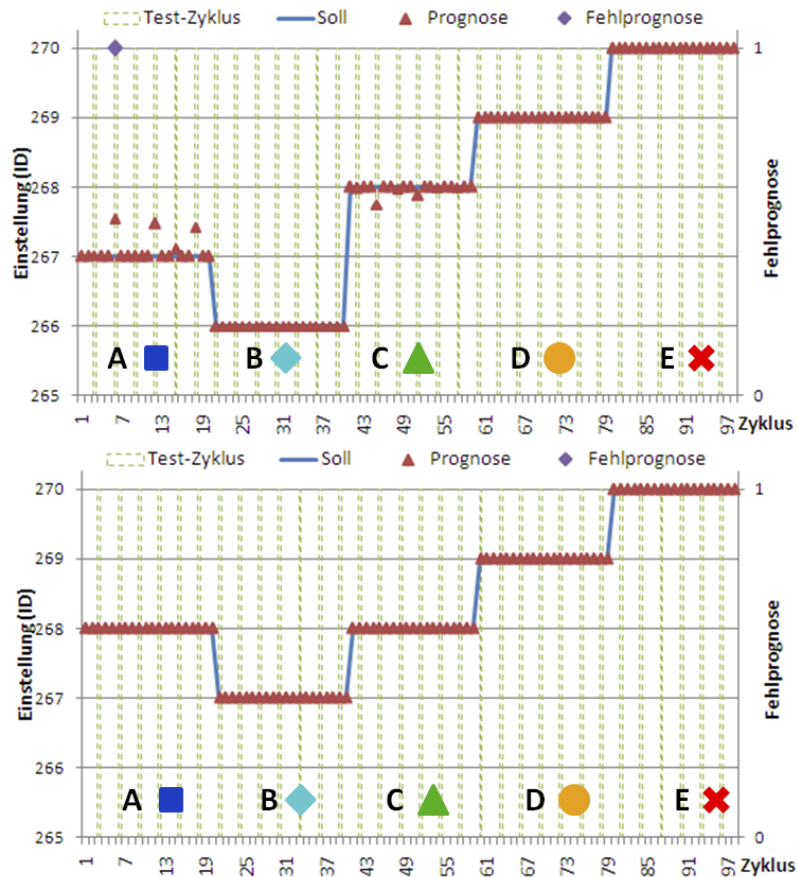


Abbildung 4.15: Prognoseergebnisse bei geänderter Codierung

wie bisher 266 und allen Zyklen von B 266 statt bisher 267 als Zielwert zugeordnet. In der Folge ergibt sich bei der Modellerstellung und Verifikation, dass nur noch ein Zyklus der Einstellungen A und C falsch klassifiziert wird: Es wurde der zu A gehörende Zyklus jedoch C zugeordnet – und nicht B. Wenn die beiden Klassen A und C zusammengefasst werden, ergibt sich für alle 98 Zyklen eine richtige Prognose. Diese beiden Fälle sind in Abbildung 4.15 analog zu Abbildung 4.14 dargestellt.

- Als weitergehende Konsequenz bedeutet dies, dass eine gute Prognosegüte nur erreicht wird, wenn sich die interessierende Zielgröße im Sammons-Mapping als hinreichend „glatt“ erweist, d. h. nahe beieinander liegende Zustandsvektoren auch ähnliche oder gleiche Zielgrößen aufweisen. Die Verteilung der Zustandsvektoren wird maßgeblich durch die enthaltenen Kenngrößen sowie ihre Normierung bestimmt und kann durch geeignete Methoden entsprechend beeinflusst werden (siehe Abschnitt 5.3).

Die Beschreibung der Prognosegüte kann über verschiedene Größen erfolgen. Mögliche Kenngrößen sind dabei:

- Anzahl der falsch zugeordneten Zyklen (bei Klassenmerkmalen),
- maximaler und/oder mittlerer Fehler (bei kontinuierlichen Merkmalen),

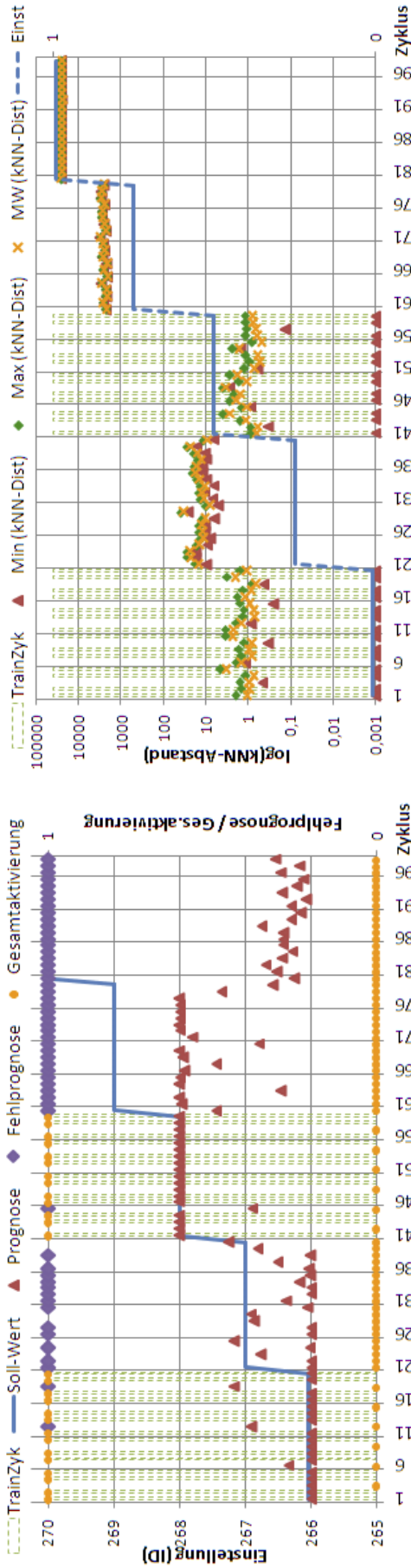
- Absolutwert des maximalen und/oder mittleren Fehlers,
- prozentualer Anteil von falsch oder unzureichend prognostizierten Zyklen,
- Korrelationskoeffizient zwischen Sollwert und Prognose über alle Zyklen, wahlweise für Gesamtdaten, Lerndaten oder nur Testdaten.

Je nach Anwendungsfall ist eine geeignete Größe auszuwählen. Zudem können diese Werte auch verknüpft werden, z. B. weniger als 2% Fehlklassifikation und maximale Abweichung < 5 .

Insgesamt ist festzustellen, dass GRNFN-Modelle über eine gute Generalisierungsfähigkeit verfügen. Bei der Codierung von Merkmalen ist jedoch auf eine Wertgebung zu achten, die benachbarte Cluster entsprechend berücksichtigt. Dabei kann das Sammons-Mapping eine wichtige Hilfestellung geben. Die Prognosegüte kann mit verschiedenen Werten quantifiziert werden. Die Güte eines Modells hängt stark davon ab, wie homogen sich die Zustandsvektoren entsprechend der relevanten Zielgröße verteilen. Dieser Punkt spielt in Abschnitt 5.3 bei der Kenngrößenselektion und -optimierung noch eine wichtige Rolle.

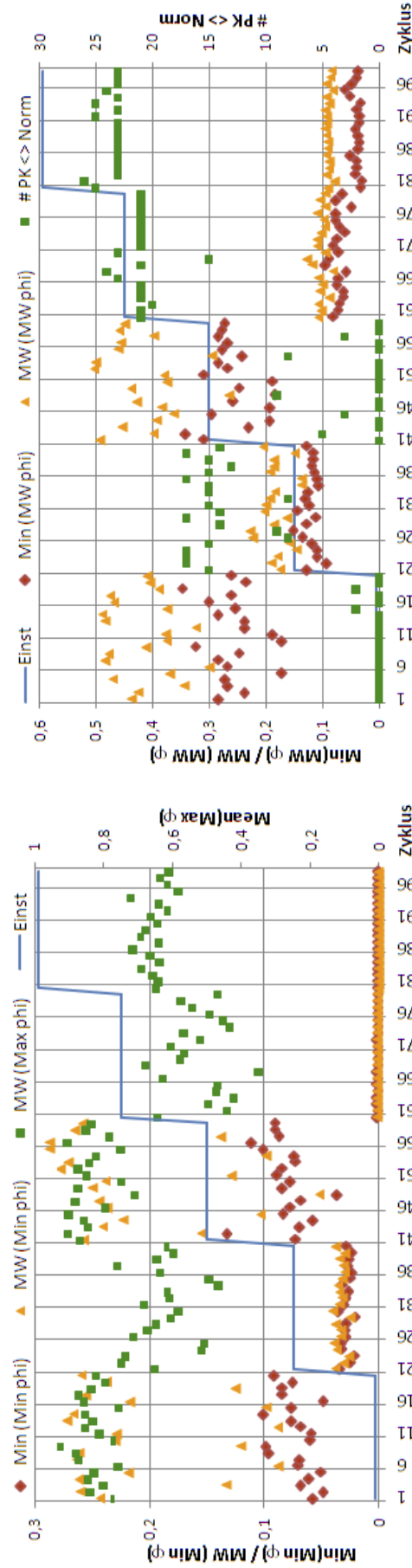
4.6 Unbekannt-Klassifikation und Störungserkennung

Primäre Aufgabe eines Prozessmodells ist hier, eine Aussage über den aktuellen Zustand eines Prozesses und eines gefertigten Produktes ableiten zu können. Für trainierte Zustände ist mit GRNFN-Modellen eine Aussage zu den Zielgrößen möglich, wobei die in Abschnitt 4.5 erläuterten Zusammenhänge maßgeblich sind. Aufgrund der Generalisierungsfähigkeiten der Modellstrukturen sind für hinreichend ähnliche Zustandsvektoren, die nicht in der Trainingsmenge enthalten waren, Prognosen mit hoher Güte möglich. Wird die Prozesseinstellung verändert oder durch Einflüsse im Fertigungsprozess oder den Vorprozessen wesentlich beeinflusst, verändern sich auch die Prozesssignale; die daraus abgeleiteten Zustandsvektoren und die prognostizierten Zielgrößen können Werte in bisher nicht abgedeckten Bereichen annehmen. Die Berechnungsvorschriften der Modelle erlauben es jedoch, auch für solch gänzlich neue – d. h. dem Modell unbekannte – Zustandsvektoren Zielgrößen zu berechnen. In der Regel stehen diese allerdings in keinem Zusammenhang mehr mit den tatsächlichen Größen. Ursache ist die Anwendung des Modells auf unbekannte Zustände. Es ist somit ebenfalls Aufgabe einer umfassenden Prozessüberwachung, solche unbekannt Zustände erkennen zu können. Dazu ist es notwendig, den am Modelleingang anliegenden Zustandsvektor mit dem aktuellen Modellwissen in Beziehung zu setzen.



(a) Prognose der Einstellung sowie max. Aktivierung Φ_{max}^k

(b) Distanzmaße bzgl. kNN mit $k=6$



(c) Aktivitätskenngrößen für Min und Max

(d) Aktivitätskenngrößen für MW und Normierung

Abbildung 4.16: Kenngrößen zur Unbekannt-Klassifikation bei Training der Einstellungen A und C

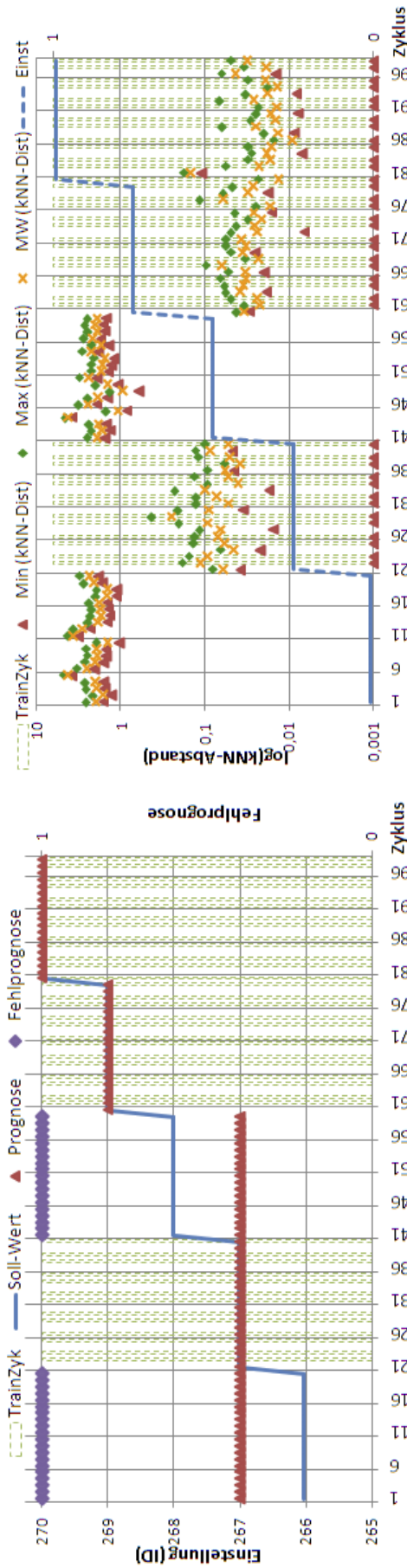
4.6.1 Indikatoren und Kenngrößen für die Unbekannt-Klassifikation

Die Entwicklung einer adäquaten Unbekannt-Klassifikation erfolgt ebenso anhand des schon mehrfach verwendeten Beispieldatensatzes aus Abschnitt 3.8. Für das Modelltraining werden jedoch – im Gegensatz zu Abschnitt 4.5 – nicht $\frac{2}{3}$ der Zyklen aller Einstellungen, sondern nur von den Einstellungen A und C verwendet. Das andere Drittel von A und C sowie die kompletten Einstellungen B, D und E werden für die Verifikation verwendet. Dabei sollten die Verifikationszyklen von A und C aufgrund der Ähnlichkeit zu den Trainingsdaten als bekannt und die Einstellungen B, D und E als unbekannt klassifiziert werden. Das Ergebnis der Prognose der Einstellung des so erstellten Modells für alle Datenzyklen des Versuchs ist in Abbildung 4.16(a) analog zu Abschnitt 4.5 dargestellt. Im Gegensatz zum vorigen Abschnitt sind diesmal jedoch die Trainingsdaten – und nicht die Testdaten – gekennzeichnet. Wie zu erwarten, werden die Einstellungen A und C weitestgehend richtig prognostiziert – die Ursachen für die drei enthaltenen Fehlzyklen sind in Abschnitt 4.5 erläutert³. Für alle Zyklen der Einstellungen B, D und E werden falsche Werte prognostiziert. Diese Aussage der Fehlprognose ist jedoch nur bei Kenntnis des tatsächlichen Wertes möglich, was in der Praxis per se nicht der Fall ist. Aus Abbildung 4.13(a) wird ersichtlich, dass die Einstellungen B, D und E als unbekannt anzusehen sind, da sie insbesondere einen großen Abstand zu den Einstellungen A und C aufweisen. Der Abstand zu den Trainingsdaten ist somit ein mögliches Maß für den Aufbau einer Unbekannt-Klassifikation. Geringe Abstände zwischen Vektoren bedingen sehr ähnliche Kenngrößen, was sich wiederum in den Aktivierungswerten der Zugehörigkeitsfunktionen in den Schichten 2 und 3 des GRNFN-Modells abbildet. Für den Aufbau einer Unbekannt-Klassifikation ergeben sich dementsprechend folgende Möglichkeiten zur Bildung von Indikatoren und Kenngrößen:

1. Relation jeder Komponente eines Zustandsvektors zu den jeweiligen Normierungsgrenzen aus Netzschicht 1,
2. Abstand des Zustandsvektors am Modelleingang zu den Vektoren des GRNFN-Modells in Schicht 2,
3. Betrachtung der (maximalen) Gesamtaktivierung der GRNFN-Blöcke in Schicht 3,
4. statistische Größen der Aktivierungswerte in Schicht 2,
5. ungültige Prognosewerte.

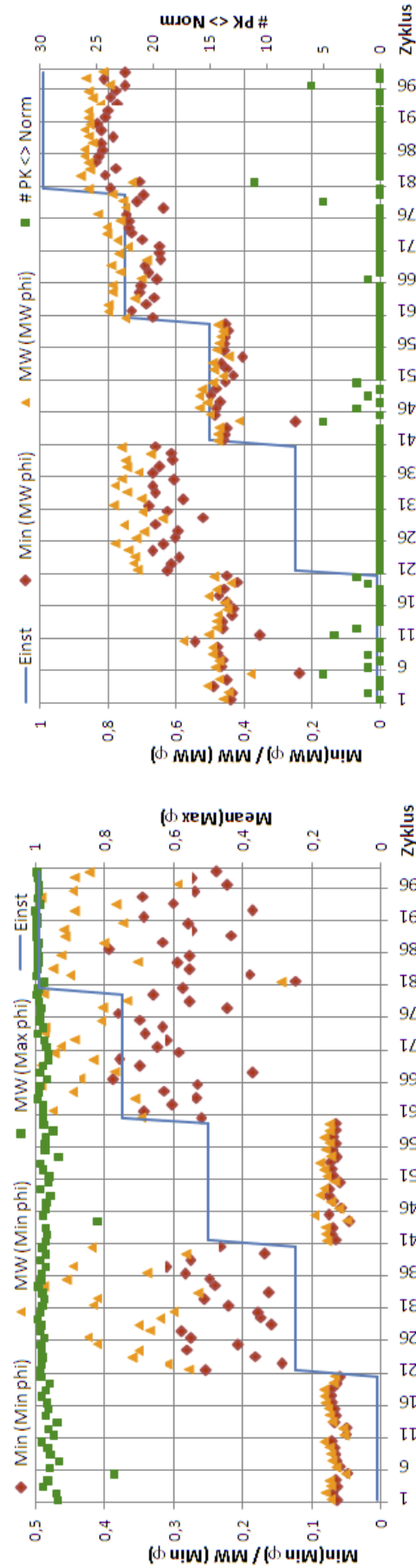
Auf die einzelnen Punkte wird in den folgenden Abschnitten näher eingegangen.

³Die Ursache für die Reduktion von 4 auf 3 Fehlzyklen liegt darin, dass für die Modellausgabe statt aller Modellstützvektoren nur noch die k nächsten Nachbarn verwendet werden (siehe unten).

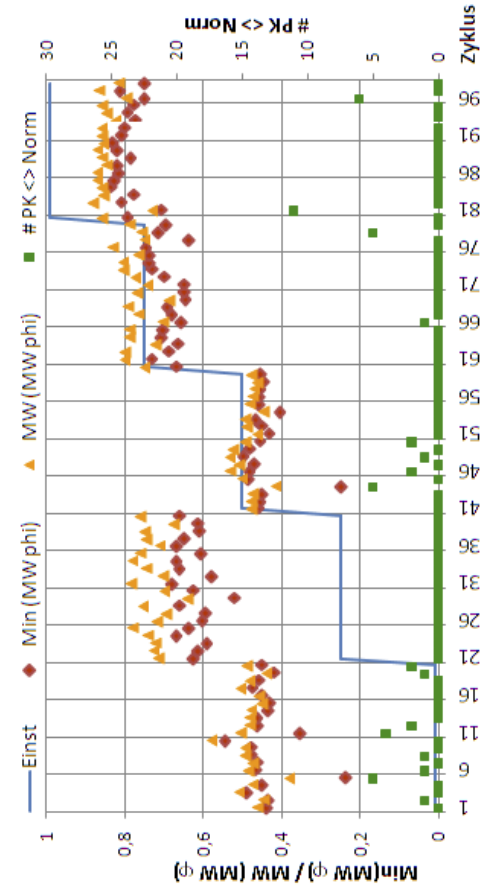


(a) Prognose der Einstellung

(b) Distanzmaße bzgl. kNN mit $k=6$



(c) Aktivitätskenngrößen für Min und Max



(d) Aktivitätskenngrößen für MW und Normierung

Abbildung 4.17: Unbekannt-Klassifikation bei Training der Einstellungen B, D und E

Einstellung	A	B	C	D	E	gesamt
Anz. PK mit globalem Minimum	4	0	3	1	22	30
Anz. PK mit globalem Maximum	0	16	1	11	2	30

Tabelle 4.1: Min- und Maxwerte aller Datenzyklen pro Einstellung

Zur Verifikation der allgemeineren Anwendbarkeit wurden in einem zweiten Modell mit entsprechender Auswertung die Rollen der Einstellungen A/C sowie B/D/E vertauscht. Das Modell wurde anhand der Einstellungen B, D und E erstellt (mit $\frac{2}{3}$ der Zyklen) und alle verbliebenen Zyklen dienten nur zur Verifikation. Die Ergebnisse sind in Abbildung 4.17 dargestellt.

Verletzung der Normierungsgrenzen

Der erste Indikator zur Identifikation unbekannter Eingangszustände ist der Vergleich jeder Komponente des Zustandsvektors mit den jeweiligen Normierungsgrenzen. Diese wurden beim Modelltraining ermittelt und werden für die Normierung innerhalb der ersten Netzwerkschicht verwendet – sind also als Modellparameter vorhanden. Liegen vermehrt Komponenten eines Zustandsvektors außerhalb der Normierungsgrenzen, ist dies ein Hinweis auf einen veränderten Prozesszustand. In Abbildung 4.16(d) ist für jeden der insgesamt 98 Datenzyklen der Versuchsreihe die Anzahl der Komponenten dargestellt, die für das mit den Einstellungen A und C trainierte Modell außerhalb der Normierungsgrenzen liegen. Auch wenn für die Verifikationszyklen der Einstellungen A und C einzelne Kenngrößen betroffen sind, verletzt für alle drei unbekanntesten Einstellungen B, D und E eine signifikant höhere Anzahl an Komponenten die Normierungsgrenzen. Beim Vergleich mit der relativen Lage der Zustandsvektoren anhand der Sammons-Projektion in Abbildung 4.13(a) ist dies auch plausibel, da die trainierten Einstellungen A und C im „Inneren“ des gesamten Zustandsraumes liegen. Im Umkehrschluss muss sich nicht jede Veränderung im Prozess in einer Verletzung der Normierungsgrenzen auswirken. Dies ist insbesondere dann der Fall, wenn der Modellerstellung ein Versuchsplan mit größerer Variation der Einstellparameter und damit Randgebiete des Zustandsraums zugrunde liegen. Dann können auch unbekannte Abweichungen vom Arbeitspunkt auftreten, die sich dennoch für alle Kenngrößen innerhalb der Normierungsgrenzen befinden. Entsprechendes ist anhand des Modells, welches für die Einstellungen B, D und E erstellt wurde, in Abbildung 4.17(d) ersichtlich. Obwohl die Einstellungen A und C dem Modell unbekannt sind, ist kein signifikanter Anteil an Kenngrößen außerhalb der Normierungsgrenzen. Dies stimmt mit Werten der Tabelle 4.1 überein, die für jede Einstellung die Anzahl der Kenngrößen angibt, für welche die minimalen und maximalen Werte über alle 98 Datenzyklen vorhanden sind. Während durch die Einstellungen A und C nur insgesamt sieben Minimal- und ein Maximalwert definiert werden, sind es bei den restlichen drei Einstellungen 23 Minimal- und 29 Maximalwerte.

Abstandsmaß zu GRNFN-Stützvektoren

Der Abstand eines Zustandsvektors zu den Stützvektoren eines GRNFN-Modells ist ein wichtiger Indikator für unbekannte Zustände. Hierbei ist insbesondere die Abstandsinformation zu den k nächsten Nachbarn (kNN) von Relevanz. Da sich die Zentrumsparameter der Aktivierungsfunktionen in der zweiten GRNFN-Modellschicht durch die Trainingsdaten definieren, sind die Trainingsvektoren implizit im Modell enthalten. Ein Abstandsmaß entsprechend der gewählten Metrik – meist wird der euklidische Abstand verwendet – kann somit allein mit vorhandenen Modelldaten ermittelt werden. Für die Auswertungen in den Abbildungen 4.16, 4.17 und 4.18 sowie in Tabelle 4.2 wurden zu jedem Eingangsvektor die sechs nächstgelegenen Stützvektoren ermittelt und verwendet. Auch in die Prognose von Zielgrößen fließen aufgrund der Eigenschaft der Aktivierungsfunktionen, mit steigendem Abstand zum Zentrum geringere Werte zu liefern, insbesondere die nächsten Nachbarn verstärkt ein. Stützvektoren mit zunehmendem Abstand hingegen steuern aufgrund der geringeren Aktivierungswerte ihrer Zugehörigkeitsfunktionen einen zu vernachlässigenden Anteil zur Ermittlung der Netzausgaben nach Formel (4.6) bei. Deshalb wurde der GRNFN-Modellansatz dahingehend verändert, dass auch in die Berechnung des Ausgabewertes nur die k nächsten Nachbarn – mit k im Bereich von $[3..10]$, je nach Umfang der Trainingsmenge – berücksichtigt werden. Dies erfordert in einem ersten Schritt zusätzlichen Rechenaufwand zur Ermittlung der nächsten Nachbarn; anschließend ist jedoch die Anzahl der zu berechnenden Aktivierungswerte entsprechend reduziert. Die Änderung der Anzahl an Fehlprognosen von 4 auf 3 im Vergleich zwischen den Abbildungen 4.14 und 4.16(a) ist auf die Tatsache zurückzuführen, dass für letztere Auswertung nur die sechs nächsten Nachbarn anstelle aller Stützvektoren verwendet wurden. Anhand der Abbildungen 4.16(a) und 4.17(a) ist zu erkennen, dass die Prognose der Einstellung einzig anhand der $k = 6$ nächsten GRNFN-Stützvektoren (für die trainierten Zustände) gleich gute Resultate liefert, wie sie in Abschnitt 4.5 mit allen GRNFN-Stützvektoren erzielt wurden. In Tabelle 4.2 sind der minimale Abstand sowie der mittlere Abstand zu den sechs nächstgelegenen GRNFN-Vektoren für je einen beliebig gewählten Vektor pro Einstellung angegeben. Der auch beim Training verwendete Zyklus 1 weist folglich einen minimalen Abstand von 0 auf. Dabei weisen die dem Modell unbekanntesten Einstellungen B, D und E deutlich höhere Abstände auf als die Verifikationszyklen aus den Einstellungen A und C. Die minimalen, mittleren und maximalen quadratischen euklidischen Abstände⁴ zu den jeweiligen nächsten Nachbarn sind für alle Datenzyklen in den Abbildungen 4.16(b) und 4.17(b) dargestellt. Die Achsskalierung ist logarithmisch und für die Trainingszyklen wurde der Abstand, der eigentlich – wie aus Tabelle 4.2 ersichtlich – Null beträgt, aus Gründen der Darstellung auf 0,001 gesetzt. Die

⁴Aus Gründen der Laufzeitoptimierung wird bei der Berechnung des euklidischen Abstands auf die Berechnung der Quadratwurzel verzichtet, so dass alle angegebenen Abstandsmaße dem quadratischen euklidischen Abstand entsprechen.

Einstellung	A	A	B	C	D	E
Zyklusnummer	1	3	24	42	69	87
Verifikationszyk. / Einst. trainiert	- / X	X / X	X / -	X / X	X / -	X / -
minimaler quadratischer Abstand	0	0,46	9,42	0,34	2506	26240
mittl. quad. Abstand (kNN mit k=6)	1,09	0,99	14,71	0,62	2532	26326
Φ_{max}^k	1	1,6E-13	1,5E-27	2,9E-14	3,4E-44	2,0E-53

Tabelle 4.2: Kenngrößen zur Unbekannt-Klassifikation entsprechend Abbildung 4.18

Bandbreite der Abstände – die Differenz zwischen minimalem und maximalem Abstand – variiert teils deutlich. Aber für beide Modelle weisen alle drei in den Abbildungen 4.16(b) und 4.17(b) dargestellten Größen jeweils für die nicht trainierten Einstellungen signifikant höhere Werte auf, so dass sie sich als Kenngrößen zur Unbekannt-Klassifikation eignen.

Maximaler Aktivierungswert

Eine Ähnlichkeit von Zustandsvektoren zu Stützvektoren – gleichbedeutend mit einem geringen Abstand – bedeutet auch nur kleine Unterschiede in den Einzelkomponenten. Daher liegt der Schluss nahe, dass geringe Abstände auch zu hohen Aktivierungswerten Φ^k der entsprechenden Blöcke in Schicht 2 des GRNFN-Modells führen und die maximale Gesamtaktivierung Φ_{max}^k ein Maß für die Unbekanntheit eines Eingangsvektors darstellt. Wie aus Abbildung 4.16(a) ersichtlich, liefern die Trainingsmuster für Φ_{max}^k Werte von Eins. Alle anderen – inklusive der Verifikationsdaten der Einstellungen A und C – führen jedoch augenscheinlich zu einer maximalen Gesamtaktivierung von Null. Bei Aktivierungswerten im Bereich von $[0; 1]$ scheint dies vordergründig widersprüchlich. Aus Tabelle 4.2 wird jedoch ersichtlich, dass die Werte zwar nicht Null sind, sich aber im Bereich von 10^{-13} bis 10^{-55} bewegen. Die Ursache hierfür ist, dass die Einzelaktivitäten miteinander multipliziert werden. Eine größere Anzahl an Faktoren zusammen mit der in Abbildung 4.18 ersichtlichen Tatsache, dass die einzelnen Kenngrößen häufig Aktivitäten $\ll 1$ liefern, führt zu einem schnellen Abfallen der Gesamtaktivität. Das Aufsummieren für die GRNFN-Prognose nach Gleichung (4.6) führt zu einem Zähler nahe Null. Dies führt jedoch nicht auch zu Ausgabegrößen von Null, da der Nenner eine gleiche Größenordnung aufweist. Jedoch können bei der Umsetzung auf einem Rechner die notwendigen Normalisierungen mit zunehmender Anzahl an Multiplikationen zu nicht mehr zu vernachlässigenden Rundungsfehlern führen, so dass auch aus praktischen Gründen eine Reduktion der Dimension des Eingangsvektors erfolgen sollte (siehe Abschnitt 5.3).

Die Werte der Einzel- und Gesamtaktivitäten sind stark von der Wahl der jeweiligen Aktivierungsfunktion und ihren Parametern abhängig. So führen hohe Werte für σ bei Glocken- und Gaußfunktionen dazu, dass die Aktivitätswerte φ auch für Kenngrößen mit größerem Abstand zum Zentrum der jeweiligen Aktivierungsfunktion im Bereich von 0,5 liegen. Da-

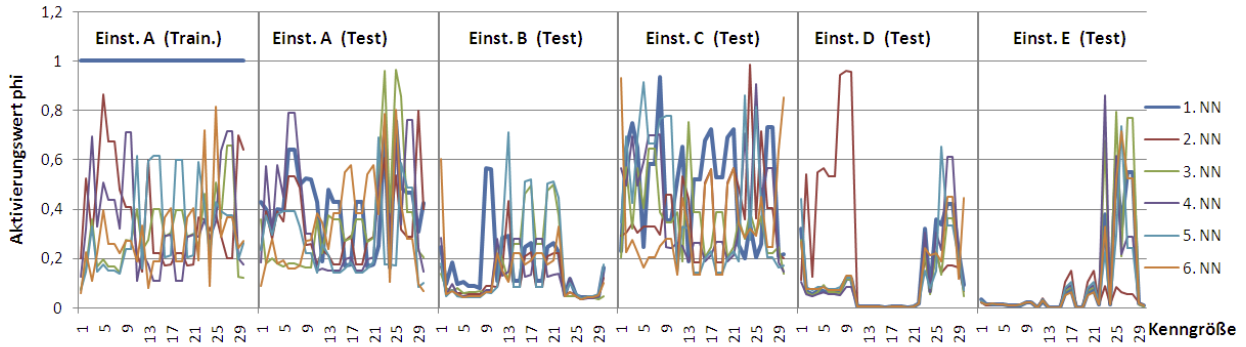


Abbildung 4.18: Aktivierungswerte φ_i^k der einzelnen Komponenten der nächstgelegenen GRNFN-Vektoren für Eingangsvektoren der verschiedenen Einstellungen

mit wird einerseits ein allzu schnelles Abfallen der Gesamtaktivität vermieden. Andererseits führt es aber durch die damit verbundene „Mittelung“ zu sinkenden Prognosegüten (siehe Abschnitt 4.5) und mit den nachfolgend beschriebenen statistischen Aktivitätskenngrößen zu einer Verschlechterung der Erkennung von unbekanntem Zuständen.

Statistische Kenngrößen der Einzelaktivitäten

Wie aus Abbildung 4.18 ersichtlich ist, können die zu einem Eingangsvektor nächstgelegenen Stützvektoren zu stark unterschiedlichen Verläufen der Aktivierungswerte φ_i^k , $i \in \{1 \dots N\}$ über alle Komponenten führen. Die Klassifikationsentscheidung nach bekannten und unbekanntem Zuständen anhand einer einzigen daraus abgeleiteten Größe, wie beispielsweise der Gesamtaktivierung Φ^k , gestaltet sich dementsprechend als schwierig und fehleranfällig. Eine Beschreibung der Aktivität mit statistischen Größen führt zu robusteren Größen und verlässlicheren Aussagen. Dies ist auch im Hinblick auf die in Abschnitt 4.4 angesprochene Robustheit der Modelle in Bezug auf Störungen oder den Ausfall einzelner Eingangsgrößen wichtig. Auch die zuvor untersuchte Abstandsberechnung weist durch die vielen Summanden eine gewisse Filterfunktion gegenüber Ausreißern in einzelnen Kenngrößen auf. Aus den Einzelaktivitäten φ_i^k eines GRNFN-Blocks der zweiten Modellschicht lassen sich Werte wie minimale, mittlere oder maximale Aktivität berechnen. Diese können jedoch für die k GRNFN-Stützvektoren, die sich durch die Nächste-Nachbar-Bestimmung ergeben haben, sehr unterschiedlich ausfallen. Dies ist insbesondere an der maximalen Aktivität für den Verifikationszyklus aus Einstellung D in Abbildung 4.18 zu sehen. Es ist also naheliegend, diese Größen über alle k nächsten Nachbarn zusammenzufassen. Mit den statistischen Operatoren Minimum, Maximum und Mittelwert ergibt sich somit eine Verkettung von jeweils zwei aufeinanderfolgenden Operationen. Die so erhaltenen Größen

1. Minimum der minimalen Aktivitätswerte: $Min_{min} = \min_k \left(\min_i \left(\varphi_i^k \right) \right)$,

2. mittlere minimale Aktivität: $MW_{min} = \frac{1}{M} \sum_k \min_i \left(\varphi_i^k \right)$,

3. mittlere maximale Aktivität: $MW_{max} = \frac{1}{M} \sum_k \max_i (\varphi_i^k)$,
4. minimale mittlere Aktivität: $Min_{MW} = \min_k \left(\frac{1}{N} \sum_i \varphi_i^k \right)$,
5. Mittelwert der mittleren Aktivität: $MW_{MW} = \frac{1}{M} \sum_k \left(\frac{1}{N} \sum_i \varphi_i^k \right) = \frac{1}{MN} \sum_k \sum_i \varphi_i^k$,

mit jeweils $i \in [1..N]$, $k \in [1..M]$ sind für das mit A und C trainierte Modell in den Abbildungen 4.16(d) und 4.16(c) sowie für das Modell für B, D und E in den Abbildungen 4.17(d) und 4.17(c) dargestellt. N ist dabei die Dimension des Zustandsvektors und M die Anzahl der berücksichtigten nächsten Nachbarn, wobei in den Beispielen $M = 6$ gilt. Alle diese Größen sind dabei für die nicht trainierten Einstellungen größer als für die trainierten und können somit als Indikatoren für unbekannte Zustände des Eingangsvektors dienen. Einzig für die Größe MW_{max} gilt dies nur für das Prognosemodell in den Einstellungen A und C. Eine Lösung dafür ist der Einsatz eines sogenannten *Votings* (siehe nachstehend).

Ungültige Prognosewerte

Während bei Glockenkurven als Aktivierungsfunktion aufgrund des asymptotischen Verhaltens keine Aktivitäten φ_i^k identisch Null auftreten können und damit auch die Gesamtaktivität Φ^k ungleich Null ist, ist dies bei Dreiecks- und Trapezfunktionen nicht der Fall. Bereits wenn ein Aktivitätswert φ_i^k Null wird, ist die Gesamtaktivität Φ^k identisch Null. Weicht nur eine einzige Komponente des Eingangsvektors stark von allen Trainingsdaten ab oder weisen die entsprechenden Aktivierungsfunktionen einen zu steilen Abfall auf, sind alle Gesamtaktivitäten Φ^k Null. Dies führt in Gleichung (4.6) zur Ermittlung des Prognosewerts dazu, dass Zähler und Nenner Null sind und somit kein gültiger Wert berechnet werden kann. Ungültige Prognosewerte sind somit – bei Einsatz von Dreiecks- und Trapezfunktionen – ein Indikator für unbekannte Zustände.

In den vorigen Abschnitten wurden mehrere mögliche Indikatoren und Kenngrößen zur Identifikation unbekannter Zustände erarbeitet. Dabei ist für die meisten Indikatoren die Festlegung eines Schwellwerts notwendig. Dieser kann anhand der Lage und Streuung der Trainingsdaten abgeschätzt werden. Solche Schwellwerte können zwar optimiert werden, wenn die zu erkennenden, unbekannt Zustände bekannt sind und dann entsprechend berücksichtigt werden. In der Praxis jedoch sind Daten „unbekannter Zustände“ in der Regel nicht vorhanden und insgesamt steht dies der eigentlichen Grundidee der Unbekannt-Klassifikation entgegen. Aus den Abbildungen 4.16 und 4.17 ist ersichtlich, dass auch optimierte Schwellwerte nicht für alle Fälle zu einer 100 %-Trennung bekannt/unbekannt führen. Zudem ist am Beispiel der mittleren Maximalaktivierung MW_{max} erkenntlich, dass nicht für jedes Modell alle Größen geeignet sind. Es bietet sich somit der Einsatz von sogenannten *Voting-Verfahren* an. Dabei werden für eine Klassifikation verschiedene Indikatoren oder

„Unterklassifikatoren“ eingesetzt. Anstatt die Entscheidung anhand eines Wertes zu treffen, werden dabei mehrere Werte berechnet und zu einer Gesamtaussage verknüpft. Für die Unbekannt-Klassifikation bedeutet dies, dass ein Zustandsvektor am Eingang dann als unbekannt klassifiziert wird, wenn beispielsweise mindestens drei der vorgestellten Indikatoren größer als der geschätzte Schwellwert sind. Voting-Verfahren (*committee based methods*) wurden in den letzten Jahren insbesondere im Bereich der Unbekannt-Klassifikation sowie dem damit verbundenen *active learning* erfolgreich eingesetzt [CAL94, DE95, TSH03].

Wichtig für eine gute Trennschärfe ist auch die Reduktion von „Rauschanteilen“ in den Zustandsvektoren. Solche Rauschanteile werden durch Kenngrößen gebildet, die aber keinen Zusammenhang mit der Zielgröße oder der Trennung von Einstellungen haben, sondern gleichverteilt streuen. Alle derartigen Kenngrößen liefern unabhängig vom tatsächlichen Zustand Aktivitätswerte $\gg 0$, beeinflussen damit insbesondere die aus der mittleren Aktivität abgeleiteten Kenngrößen und reduzieren die Aussagekraft dieser Größen. Solche Kenngrößen ohne Trennschärfe oder Informationsgehalt wirken sich auch auf die Abstandsberechnung aus. Sie führen in der Folge zu einer schlechteren Trennung – im originalen Zustandsraum und in der mittels Sammons-Mapping gewonnenen Projektion – und insgesamt zu einer geringeren Modellgüte. Solche Größen beeinflussen somit alle Bereiche der Modellbildung. Die Aufteilung in relevante und irrelevante Kenngrößen ist wesentliche Aufgabe der Merkmalsselektion, die Teil der Optimierung der Zustandsvektoren ist und in Abschnitt 5.3 behandelt wird.

4.6.2 Auswirkung des Nachtrainings

Bei einer intelligenten Prozessüberwachung ist ein Ziel, unbekannte Prozesszustände zuerst zu erkennen und anschließend in das Prozessmodell zu übernehmen. Auf diese Weise wird das Modell fortwährend angepasst und erweitert. Der anfängliche Versuchsaufwand kann also minimiert werden, da das (Nach)Training ereignisbezogen erfolgt. Wie in Abschnitt 4.4 erläutert, sind GRNFN-Modell in der Lage, neue Zustände und Zusammenhänge zusätzlich zu integrieren, ohne ein komplettes Neutraining zu erfordern. Eine Prozesseinstellung, die beim Modelltraining nicht enthalten war, kann mit den im vorigen Abschnitt beschriebenen Kenngrößen erkannt werden. Das Hinzufügen – also Trainieren – eines Zustandsvektors einer veränderten Prozesseinstellung führt dann entsprechend zu geänderten Modellzuständen, wenn anschließend der nächste Vektor der neuen Einstellung bewertet werden soll.

Das Nachtraining mit als unbekannt identifizierten Daten wurde anhand der Beispieldaten durchgeführt und näher untersucht. Dabei wurde wie schon in Abschnitt 4.6.1 ein Modell nur anhand der Daten der Einstellungen A und C trainiert. Über die erarbeiteten Kenngrößen ist es möglich, alle Zyklen der Einstellung B, die relativ nahe an den Trainingsdaten A und C liegen, wie angestrebt als unbekannt zu identifizieren. Wie sich das sukzessive Hinzufügen von Datenvektoren der Einstellung B zum GRNFN-Modell auf die

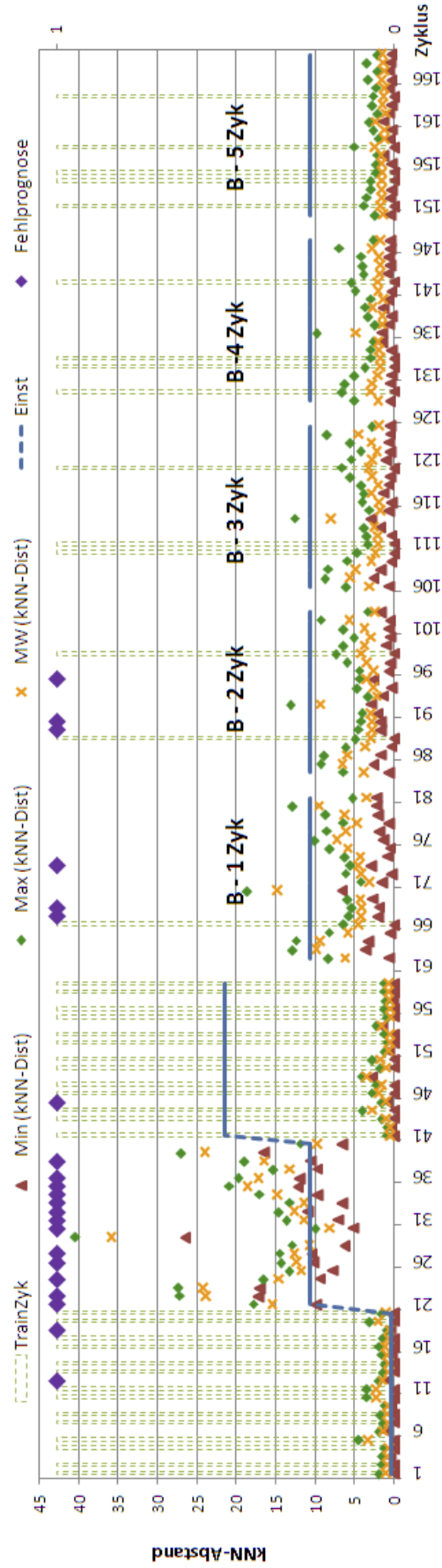


Abbildung 4.19: Distanzmaße für Training A und C mit zunehmendem Anteil an B (kNN mit $k=6$)

Kenngrößen zur Unbekannt-Klassifikation und das Prognoseergebnis auswirkt, ist in Abbildung 4.19 dargestellt. Für jeden weiteren Vektor von B sind die bereits im Modell integrierten die nächstgelegenen GRNFN-Vektoren, so dass sich auch die Abstandsmaße entsprechend verändern. In Abbildung 4.19 ist zu sehen, dass bereits mit einem Trainingsvektor aus Einstellung B 17 der 20 Datenzyklen von B richtig klassifiziert werden. Zudem führt bereits das Training des ersten Vektors von B zu einem Abfall der minimalen kNN-Distanz in den Bereich der Trainingsdaten. Jeder zusätzliche Trainingsvektor der Einstellung von B führt zu einer Annäherung der Aktivitäts- und Distanzmaße an Bereiche, wie sie sich auch für die ursprüngliche Trainingsmenge aus A und C ergeben. In der Folge sind weitere Vektoren von B nur noch schwer als unbekannt zu identifizieren. Bereits ab drei Vektoren von B – was der Hälfte der für die Klassifikation verwendeten sechs nächsten Nachbarn entspricht – werden alle Vektoren der richtigen Einstellung zugeordnet und es tritt keine Fehlklassifikation mehr auf. Die Einstellung B wird folglich sehr schnell als bekannt klassifiziert. Um stets eine sichere Erkennung zu erzielen und die schrittweise Annäherung der Unbekannt-Identifikatoren zu vermeiden, ist es ratsam, neue Vektoren nicht einzeln dem GRNFN-Modell hinzuzufügen. Statt dessen ist es besser, als unbekannt klassifizierte Vektoren separat zu sammeln. In der Folge wird jeder weitere, ebenfalls neue Zustandsvektor mit dem unveränderten Modell mit derselben Sicherheit wie der erste Vektor als Unbekannt klassifiziert. Sobald mindestens $M/2$ bis M Vektoren, die untereinander ähnlich sind, vorhanden sind, können diese en bloc dem Modell hinzugefügt werden. Die neue Einstellung ist dem Modell dann komplett bekannt und der unsichere Übergang wie bei der sukzessiven Aufnahme einzelner Vektoren wird vermieden.

4.7 Kriterien für die getroffene Wahl der GRNFN-Modelle

Um den in Abbildung 1.5 dargestellten Lösungsansatz umsetzen zu können, muss die gewählte Modellierungsmethode anhand eines Zustandsvektors in der Lage sein, einerseits einen **Prognosewert für die Zielgröße** – im gewählten Anwendungsfall die Qualität eines mit Ultraschall gefügten Produktes – abzuleiten. Andererseits muss das Modell geeignet sein, im Vergleich zu den Trainingsdaten geänderte Zustände zu erkennen. Nur so sind Veränderungen in den Vorprozessen, die sich auf die Schweißung auswirken, auch zu detektieren. GRNFN-Modelle erfüllen sämtliche Anforderungen, die sich durch diese Aufgabenstellung ergeben. Über die Approximationsformel (4.3) liefern GRNFN-Modelle beliebige **kontinuierliche Ausgangswerte** und sind damit RBF-Netzen oder MLP-Modellen ebenbürtig. Auch **diskrete Merkmale**, deren Ausprägung sich aus einem begrenzten Satz an Werten speist, können über entsprechende Schwellwerte oder Rundungsoperationen ab-

gebildet werden. Eine **Unbekannt-Klassifikation** bei GRNFN-Modellen ist mittels der in Abschnitt 4.6 erarbeiteten Indikatoren möglich. MLP-Modelle, wie sie in der NEPRES-Systemtechnik (siehe Abschnitt 2.6) bisher zum Einsatz kommen, bieten diese Möglichkeiten nicht.

Zudem verfügen GRNFN-Modelle über eine Reihe von Eigenschaften, die sie anderen Modellierungsverfahren überlegen erscheinen lassen: Das **Training ist sehr schnell**, da es in einem Durchgang erfolgt und keine größere Anzahl an Iterationen benötigt, wie es bei MLP-Modellen oder auch bei ANFIS der Fall ist. Auch ist das **Training deterministisch**, d. h. aus einem wiederholten Training mit denselben Daten resultiert ein identisches Modell. Bei MLP-Modellen ist dies aufgrund der meist zufälligen Startinitialisierung nicht der Fall. Ferner ist bei GRNFN-Modellen die **Anzahl der Modellparameter gering**, die während des Trainings festgelegt werden müssen. Im Wesentlichen sind dies nur die (globalen) Parameterwerte für die Aktivierungsfunktionen, wobei die in Abschnitt 4.3 dargestellten Diagramme über die Modellgüte in Abhängigkeit der Parameterwerte als Anhaltspunkte dienen können. Die geringe Anzahl an Modellparametern, die sich nicht direkt aus den Trainingsdaten ableiten, machen GRNFN-Modelle auch **für hohe Dimensionen des Zustandsvektors geeignet**. Bei MLP-Modellen hingegen steigt die Anzahl der Gewichte, die durch das Training festgelegt werden müssen, in der Regel quadratisch; analog steigt auch der Umfang der erforderlichen Versuchsdaten. Ein weiterer Vorteil von GRNFN-Modellen ist die **Toleranz und Robustheit gegenüber dem Ausfall von Eingangsdaten** oder Rauschen, wie sie z. B. bei Sensordefekten auftreten können. Diese Eigenschaft ist auch bei der Merkmalsselektion über Suchstrategien wie Genetische Algorithmen hilfreich, wenn die Modellgüte als Auswahlkriterium fungiert. Die Prognosewerte und damit die Modellgüte können für verschiedene selektierte Kenngrößen des Zustandsvektors ohne ein Neutraining berechnet werden. GRNFN-Modelle unterstützen die **Optimierung des Gesamtsystems** in Kombination mit anderen Verfahren, wie sie in Abschnitt 5.3 beschrieben werden. Durch die **Interpretierbarkeit** der Netzstruktur als Fuzzy-Regeln wird das Verständnis der Zusammenhänge gefördert und sogenannte *black boxes* wie beispielsweise bei MLP-Modellen können reduziert werden. Über die Möglichkeit, manuell festgelegte Fuzzy-Regeln in das Modell zu übernehmen, kann auch eine existierende Fuzzy-Regelbasis und damit **vorhandenes Prozesswissen in das Modell integriert** werden. Abschließend ist anzuführen, dass GRNFN-Modelle auch **nach dem Training anpassbar** sind, indem gezielt Daten aus dem Modell entfernt und damit „vergessen“ oder neue Zustände hinzugefügt und damit in das Modell integriert werden. Letzteres ist insbesondere im Hinblick auf *active learning* sehr hilfreich, da ein Nachtraining und somit die **Erweiterung des Modellwissens ohne ein komplettes Neutraining** erfolgen kann.

5 Realisierte Signalverarbeitung

Ein weiteres wichtiges Kernelement der Prozessüberwachung ist neben der Modellbildung eine leistungsfähige und problemangepasste Signalverarbeitung. Ein Ergebnis der Signalverarbeitung ist die Bereitstellung von signifikanten Kenngrößen in Form eines Zustandsvektors, der dem Prozessmodell als Eingangsgröße dient und den aktuellen Prozesszustand abbildet. Die Güte eines Prozessmodells hängt neben der Modellierungsmethode maßgeblich von geeigneten Prozesskenngrößen (PK) ab. Die notwendigen Schritte von der Signalvisualisierung zur Kenngrößenfindung bis zur Optimierung des beschreibenden Kenngrößenvektors sind im Folgenden erläutert.

5.1 Visualisierung der Signale

Ausgangspunkt der Signalanalyse und Kenngrößendefinition sind zyklusbezogene Signalverläufe der verschiedenen Sensoren. Als Startbedingung für die Aufzeichnung werden entsprechende Triggersignale verwendet. Diese dienen anschließend auch als Referenzsignale für den Vergleich und die Zuordnung verschiedener Zyklen. Die Signale für einen Schweißzyklus der Versuchsanlage sind in Abbildung 3.2 dargestellt. Wie bereits in Abschnitt 3.2 beschrieben, werden aufgrund der unterschiedlich schnellen Abläufe im Gesamtprozess und während der eigentlichen Schweißphase zwei Signalscharen mit unterschiedlichen Zeitskalen und Triggerereignissen erfasst. Die nachfolgend beschriebenen Verfahren können unabhängig von der Aufzeichnungsfrequenz und -dauer angewendet werden.

Abbildung 5.1 zeigt Darstellungsformen und Verarbeitungsschritte, die insbesondere für die Analyse der Streuungen und die Ableitung von Kenngrößen notwendig sind. Die einzelnen Diagramme und ihre Bedeutung werden im Folgenden kurz erläutert.

Signale (Abb. 5.1(a)): Ausgangspunkt für alle Analysen sind die Signalverläufe eines Schweißzyklus, deren Aufzeichnung durch ein reproduzierbares Triggersignal gesteuert wird. Anhand der einzelnen Signale kann der Prozessablauf klar nachvollzogen werden. Der Trigger „*Prozess aktiv*“ – blau dargestellt – startet die Aufzeichnung und wird am Ende des Schweißvorgangs wieder zurückgesetzt. Der Pneumatikdruck in der Schweißpresse – anhand der Stufenform gut erkennbar – wird vom Start des Schweißprozesses an mit kleiner Verzögerung aufgebaut. Während des Schweißvorgangs bleibt der Druck konstant, da Triggerkraft und Schweißkraft auf denselben Wert eingestellt

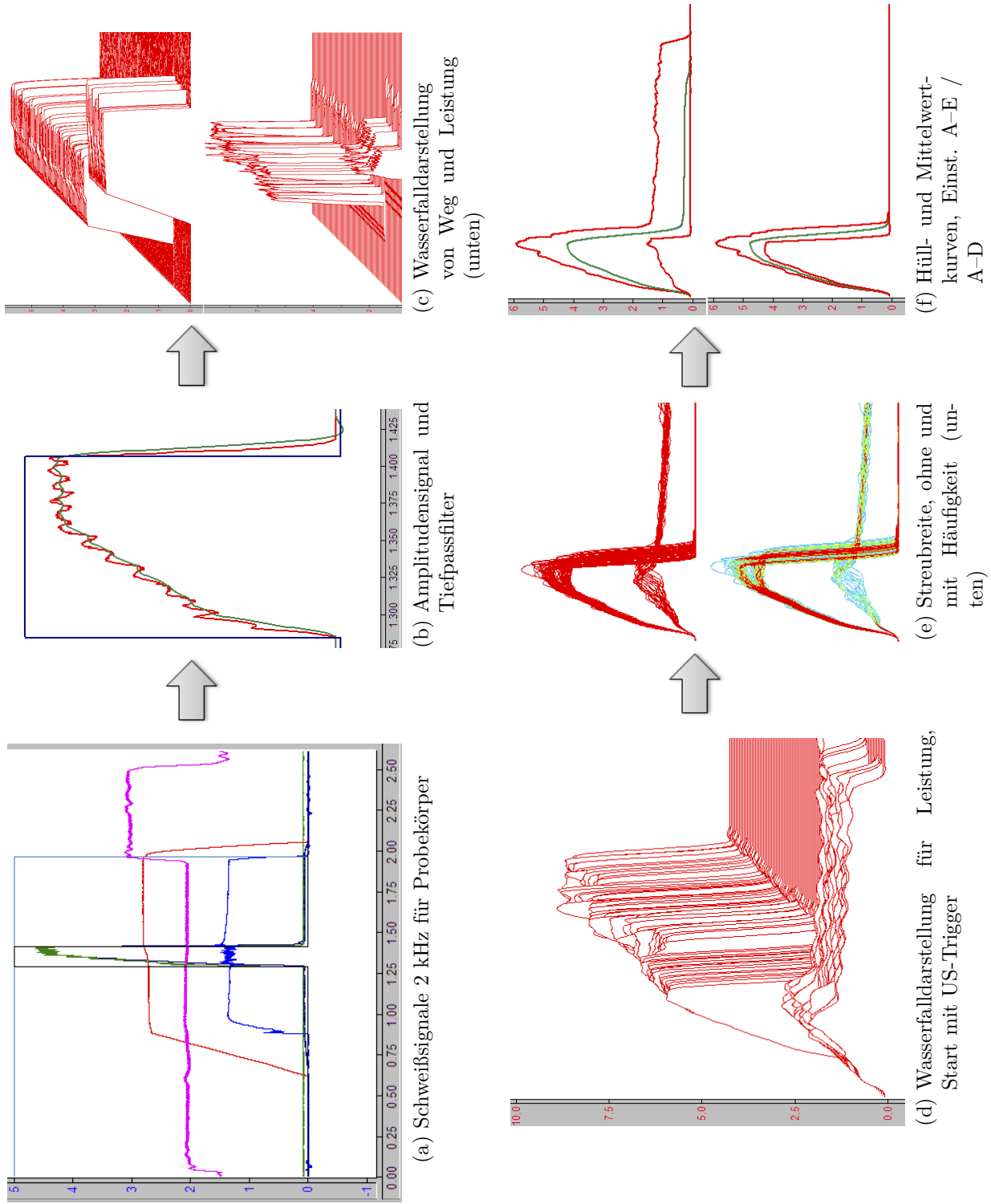


Abbildung 5.1: Abfolge verschiedener Darstellungsformen

sind. Der Anstieg am Ende ist durch das Rückfahren der Sonotrode begründet. Das Absenken der Sonotrode auf das Werkstück führt zu einem linearen Anstieg des rot dargestellten Wegsignals. Mit dem Aufsetzen der Sonotrode bleibt der Weg konstant und die Kraftmessung unter der Werkstückaufnahme registriert einen Kraftanstieg. Anschließend setzt die Steuerung den Trigger „*Ultraschall aktiv*“, der als Steuersignal für den Generator dient. Während der Schweißung steigen Amplitude und Leistung an, bis die Sonotrode mit der eingestellten Amplitude schwingt. Gleichzeitig ist die Schwingung auch am Kraftsignal unter dem Amboss zu erkennen. In der anschließenden Haltephase schwingt die Amplitude aus. Bis zum Ende der Haltezeit verbleibt die Sonotrode unter Einwirkung der eingestellten Haltekraft auf dem Werkstück, bis die Schmelze erstarrt ist. Abschließend wird die Sonotrode zurückgefahren.

Signalfilter (Abb. 5.1(b)): Je nach Signalgüte ist als erstes eine Signalfilterung notwendig. Dargestellt ist eine Tiefpassfilterung des Amplitudensignals zur Reduktion der überlagerten Schwingung, die vermutlich auf das Regelverhalten des Generators zurückzuführen ist. Auch das Kraftsignal unter dem Amboss muss gefiltert und auf den statischen Anteil reduziert werden. Der dynamische Anteil der Kraft ist auf die Sonotrodenschwingung zurückzuführen. Die darin enthaltenen Frequenzanteile im Bereich der Erregerfrequenz des Generators dürften jedoch aufgrund der nach dem Datenblatt aufgeführten Dämpfung der Sensoren bei > 5 kHz nicht vorkommen. Die Ursache ist vermutlich in der Messkette zu suchen und eine Interpretation bezüglich der Sonotrodenschwingung wäre unzulässig.

Wasserfalldarstellung (Abb. 5.1(c)): Sie vermittelt einen ersten Eindruck über die auftretenden Schwankungen über mehrere Zyklen hinweg. Dabei wird der Verlauf einzelner Signale über mehrere Zyklen gestaffelt dargestellt, der letzte Zyklus zuvorderst. In Abbildung 5.1(c) ist dies anhand des Weg- (oben) und Leistungssignals (unten) für alle fünf Einstellungen der Versuchsreihe aus Abschnitt 3.8 gezeigt. Auffällig ist der deutliche Zeitversatz für die letzte Einstellung sowohl im Weg- als auch im Leistungssignal. Ursächlich dafür ist die geringer eingestellte Kraft und die deshalb langsamere Sonotrodenbewegung, was zu einer zeitlichen Verschiebung des Ultraschall-Triggers führt. Insbesondere im Leistungssignal sind deutliche Unterschiede zwischen den Zyklen zu erkennen und die letzte Einstellung weist eine gänzlich unterschiedliche Signalform über alle 19 Zyklen auf. Grund ist das deutlich geänderte Prozessverhalten, bedingt durch die stark verringerte Schweißkraft (siehe auch Abschnitt 2.1). Dies ist auch bei der Optimierung des Zustandsvektors (Abschnitt 5.3) mehrfach auffällig.

Getriggerte Wasserfalldarstellung (Abb. 5.1(d)): Die zeitlichen Schwankungen einer Wasserfalldarstellung aus den Schweißsignalen (Abbildung 5.1(a)) erschweren eine Interpretation der Signalform immens. Abhilfe kann dadurch geschaffen werden, dass

die Signale anhand eines weiteren – evtl. auch künstlichen – Triggersignals synchronisiert werden. Da Leistungs- und Amplitudensignal nur während der eigentlichen Schweißphase relevant sind, bietet sich der Trigger *Ultraschall aktiv* an. Das Resultat ist in Abbildung 5.1(d) zu sehen.

Signalstreuung (Abb. 5.1(e)): Die Streuung der Signale ist in der Wasserfalldarstellung prinzipiell erkennbar, durch den Versatz jedoch nur schwer quantifizierbar. Wenn die Signale ohne Versatz übereinander gelegt werden, sind die relativen Schwankungen deutlich besser zu sehen. Dies führt zu der oberen Darstellung in Abbildung 5.1(e), in der die gesamte Streubreite des Signals deutlich wird. Nachteilig dabei ist, dass Häufungen nicht erkennbar sind. In der unteren Darstellung ist die „Häufigkeit der Trefler“ farblich codiert. Innerhalb der Streubreiten können so Häufungen oder Cluster, die beispielsweise durch eine Variation von Einstellparametern bedingt sind, erkannt werden.

Hüll- und Mittelwertkurven (Abb. 5.1(f)): Mittels Hüllkurven, wie sie auch schon in Abbildung 2.20 eingeführt wurden, kann die gesamte Streubreite eines Signals beschrieben werden. Jedoch können auf die Hüllkurven schon einzelne Ausreißer einen signifikanten Einfluss haben. Eine andere Möglichkeit für Analysezwecke ist die dargestellte Mittelwertkurve, die bei genügend großer Trainingsmenge durch einzelne Ausreißer nicht wesentlich beeinflusst wird. Die Kombination beider Einstellungen entspricht den statistischen Rechengrößen Mittelwert und Streubreite für jeden einzelnen Zeitpunkt der Messreihe. Insgesamt spielen die verwendeten Daten eine wesentliche Rolle, wie an den beiden Darstellungen in Abbildung 5.1(f) zu erkennen ist. Während in Abbildung 5.1(f) oben alle Zyklen – somit auch die der stark variierte Einstellung E – zur Definition von Hüll- und Mittelwertkurven verwendet wurden, sind die entsprechenden Kurven der unteren Darstellung nur anhand der Einstellungen A bis D entsprechend der beschriebenen Versuchsreihe (Abschnitt 3.8) erfolgt.

5.2 Streuungsanalyse

Die Analyse der Streuungen erfolgt in mehreren Stufen. Ausgehend von der Streuung der Signale auch in Abhängigkeit von Prozessvariationen erfolgt die Ableitung von Kenngrößen. Diese können anschließend wieder auf Streuverhalten und Zusammenhänge bezüglich Qualitäts- und Versuchsdaten untersucht werden.

5.2.1 Streuung der Prozesssignale

In Abbildung 5.2 sind die Streuungen der vier Maschinensignale Weg, Leistung, Druck und Amplitude für verschiedene Versuchsumfänge der Beispieldaten aus Abschnitt 3.8 anhand

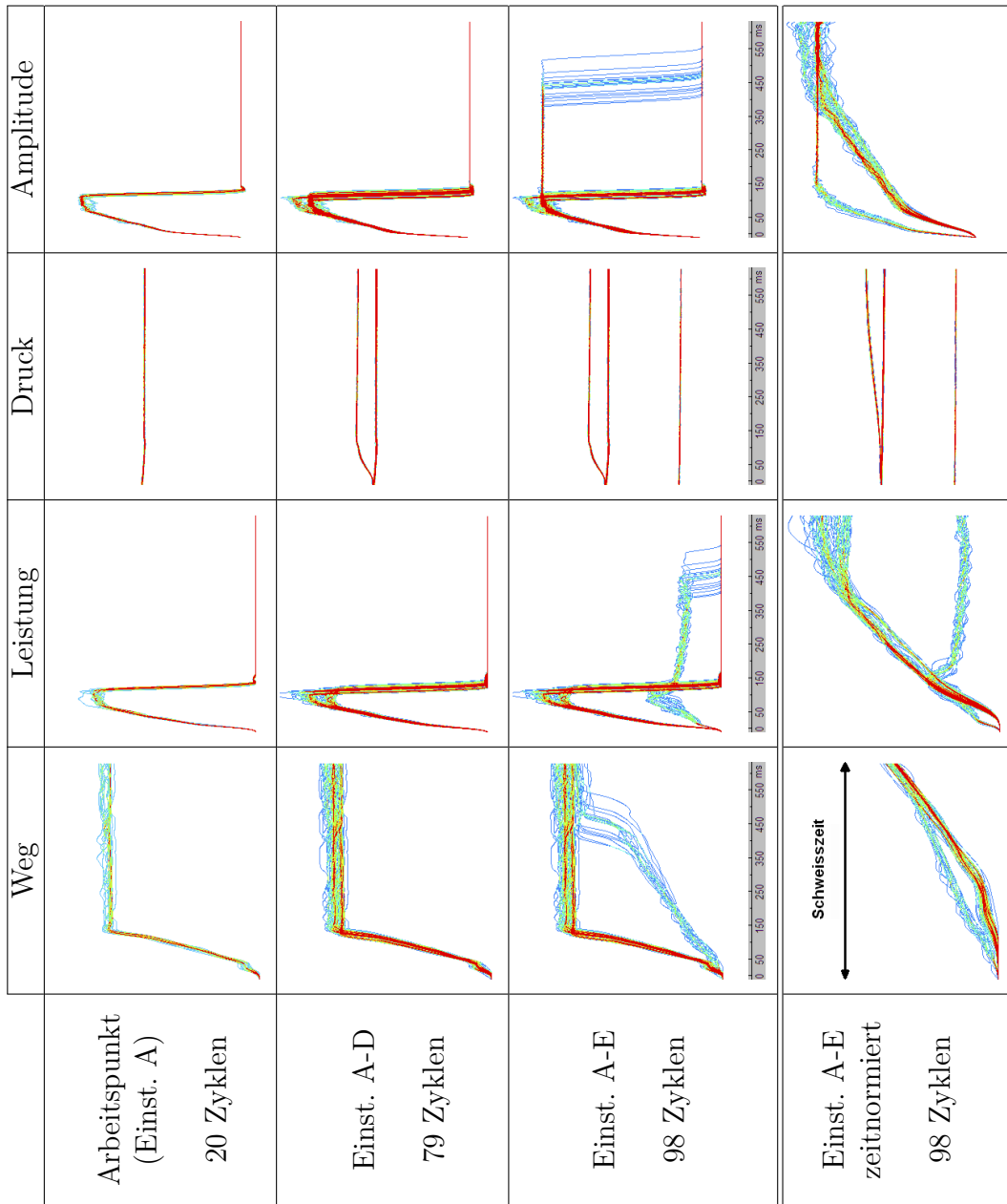


Abbildung 5.2: Einstellungsabhängige Streuung der einzelnen Maschinensignale

der in Abbildung 5.1(e) eingeführten Visualisierung dargestellt. Die dargestellte Zeitspanne von 600 ms richtet sich nach der längsten Schweißzeit der Versuchsreihe mit 515 ms plus Haltezeit. Nach Abschalten des Generators vergehen nochmals in etwa 25 ms, bis die Amplitudenschwingung abgeklungen ist. Die Diagramme der ersten Zeile umfassen nur die 20 Zyklen der Einstellung A, dem Arbeitspunkt des Prozesses. Im Wegsignal sind deutliche Schwankungen im Verlauf der ersten 40 ms sowie dem Endweg während der Haltephase zu erkennen. Leistungs- und Amplitudensignal weisen ein reproduzierbares Anlaufverhalten auf, schwanken aber im weiteren Verlauf. Die Amplitudenschwankungen sind geringer, was auf die Regelung der Amplitude durch die Maschinensteuerung zurückzuführen ist. Umgekehrt führen Regeleingriffe des Generators zu einer geänderten Leistungsaufnahme, die sich teilweise im Leistungssignal abbildet. Das Drucksignal ist über alle Zyklen sehr konstant. Dies ist darauf zurückzuführen, dass die Schweißkraft eine geregelte Stellgröße des Prozesses ist und somit den Prozessablauf bestimmt. Alle anderen Signale hingegen registrieren die Reaktion des Prozesses und weisen dementsprechend Schwankungen auf. Für die nächsten Zeilen in Abbildung 5.2 sind die Zyklen weiterer Einstellungen hinzugekommen. Die geänderten Prozesseinstellungen wirken sich als Prozessvariation aus und sind als Schwankungen zu erkennen. Im Drucksignal sind klar die zwei oder drei Stufen der Versuchsvariation zu erkennen. Ebenso ist die Verzögerung von ca. 100 ms zu erkennen, die das Gesamtsystem aufgrund von Trägheit benötigt, um ausgehend von den 300 N Triggerkraft die Schweißkraft von 350 N zu erreichen. Auch in den anderen Signalen sind Häufungstendenzen zu erkennen, die mit den geänderten Parametern korrespondieren. Das Abschalten des Ultraschalls nach Erreichen des Abschaltkriteriums – hier das Erreichen des Schweißwegs von 0,32 mm – hat einen Rückgang von Leistungs- und Amplitudensignal auf Null zur Folge. Diese Signalfanken sind gut zu erkennen und bilden die schwankenden Schweißzeiten sichtbar ab. Durch eine Normierung der Signale anhand der resultierenden Schweißzeit werden konstante Signallängen erreicht und die Signalschwankungen entsprechend transformiert. Dies ist in der letzten Zeile von Abbildung 5.2 dargestellt. Eine einheitliche Zeitskala kann natürlich nicht mehr angegeben werden. Dafür können durch die konstante Signallänge Methoden der Signalverarbeitung wie die PCA (siehe Abschnitt 2.6.2 und 4.1) verwendet werden, die konstante Vektorlängen voraussetzen. Dadurch können die Schwankungen neben prozessspezifischen Kenngrößen mit existierenden Methoden der NEPRES-Technik beschrieben werden (siehe Abschnitt 5.2.2).

5.2.2 Ableitung von Kenngrößen

Auf der Basis der identifizierten Prozessstreuung der aufgezeichneten Signale werden Kenngrößen definiert, um die Signalvarianzen beschreiben zu können. Ziel der Kenngrößenbildung ist die Reduktion der Datenmenge, ohne wesentliche Informationen zu verlieren. Zudem wird durch die Definition eines Kenngrößensatzes eine konstante Dimension des durch sie

gebildeten Zustandsvektors erzielt. Dies ist für eine Vielzahl von Modellierungsmethoden eine zwingende Voraussetzung, da die Modelle eine feste Eingangsdimension aufweisen.

Zuerst wird ein Satz an Kenngrößen abgeleitet, der möglichst viele Eigenheiten des Prozesses und Auffälligkeiten der Signale beschreibt. Eine anschließende Optimierung des Zustandsvektors (siehe Abschnitt 5.3) dient dazu, die Daten weiter zu verdichten und nur solche Kenngrößen zu selektieren, die einen Zusammenhang zur betrachteten Zielgröße aufweisen. Die Dimensionsreduktion muss so weit erfolgen, dass die resultierenden Zustandsvektoren als Eingangsdaten für Prozessmodelle dienen können, ohne deren Leistungsfähigkeit zu begrenzen. Bei zu wenigen Informationen fehlen wesentliche Zusammenhänge, so dass diese vom Modell nicht abgebildet werden können. Daraus resultiert direkt eine schlechtere Modellgüte. Ist der Umfang der Eingangsdaten zu groß, führt dies je nach Modellierungsart zu einer hohen Anzahl freier Modellparameter. Deren Festlegung steht in enger Verbindung mit den notwendigen Versuchsumfängen, die mit wachsender Größe schnell unpraktikabel werden. Dies führt letztendlich zu schlechten oder auch instabilen Prozessmodellen. Eine weitere Folge einer zu großen Eingangsdimension liegt in steigendem Trainingsaufwand, erhöhtem Speicherbedarf oder höherem Aufwand und Zeitbedarf für die Berechnung der Modellausgabe.

Prozessunabhängige Kenngrößen

Die Signalstreuungen können zu einem gewissen Anteil durch eine Reihe an „Standardkenngrößen“ beschrieben werden. Im Abschlussbericht [N.N03] werden als Kenngrößen für jedes Signal die Werte an $N = 10$ festgelegten Zeitpunkten verwendet. Dieses Vorgehen hat den Vorteil, dass es einfach implementiert werden kann. Nachteilig ist jedoch, dass weder auf prozess- und zyklusspezifische Eigenheiten eingegangen wird, noch schwankende Prozesszeiten berücksichtigt werden können. Andere Kenngrößen werden durch Mittelwerte und Integrale verschiedener Signale – auch orientiert an Prozessphasen – sowie durch Start- und Endwerte zu definierten Zeitpunkten gebildet. Weitere wichtige Größen sind Minimal- und Maximalwerte sowie der Zeitpunkt ihres Auftretens. Zur weiteren Beschreibung der Schwankungen dienen erweiterte Kenngrößen wie die Steigung von Regressionsgeraden oder Sekanten. Start- und Endpunkte werden dabei datenabhängig festgelegt. Interessante Intervalle können beispielsweise vom Start bis zum Maximum oder vom Maximum eines Signals bis zum Ende der Prozessphase sein. Dies ist beim Leistungssignal der Einstellung E interessant (siehe Abbildung 5.2). Für dieselben Bereiche können Mittelwerte oder (Teil)Integrale gebildet werden. Alle diese Größen können sowohl für die Originalsignale als auch für die zeitnormierten Signale (siehe Abbildung 5.2) berechnet werden. Die Implementierung der bisher genannten Größen ist mit zunehmendem Aufwand verbunden. Sind die Routinen zur Berechnung aber einmal umgesetzt, sind sie allgemein gültig und können für alle Arten von Prozessen verwendet werden.

Es ist auch möglich, die Schwankung der Signalformen mittels einer PCA-Codierung zu beschreiben (siehe Abschnitte 2.6.2 und 4.1.1). Allerdings muss hierfür erst in einem separaten Trainingsdurchlauf die zugrunde liegende Hauptachsentransformation mittels aller Versuchsdaten bestimmt werden. Alle anderen Kenngrößen können ohne Vorarbeiten direkt vom ersten Zyklus an berechnet werden. Zu beachten ist, dass für die PCA konstante Signallängen Voraussetzung sind, die über eine Zeitnormierung erreicht werden können.

Einige weitere Kenngrößen sind der deskriptiven Statistik entlehnt. Zur Beschreibung und Charakterisierung von Verteilungen werden dort Momente k -ter Ordnung verwendet. Das erste Moment einer Zufallsvariablen X ist dabei der Mittelwert $\mu = E(X)$ mit E als Erwartungswert. Die zentralen Momente k -ter Ordnung – zentral bedeutet, dass sie mittelwertbereinigt sind – berechnen sich allgemein zu $\mu_k = E((X - \mu)^k)$. Die Varianz ist identisch mit dem zweiten zentralen Moment, das dritte wird als Schiefe und das vierte als Exzess oder Wölbung bezeichnet. Werden die aufeinander folgenden Werte einer Zeitreihe als Werte einer Verteilungsfunktion betrachtet, können die Momente entsprechend berechnet werden. Die so berechneten Werte können als eine Art „Formparameter“ der Kurve interpretiert und so zur Unterscheidung herangezogen werden. Der zeitnormierte Verlauf des Schweißwegs (Abbildung 5.2, links unten) weist zwei deutliche Cluster und Verlaufsformen auf. Für die erste Verlaufsform, bestehend aus den 79 Zyklen der Einstellungen A–D, ergeben sich für die Schiefe ein Wert von 2 und für die Wölbung ein Wert von 4,7 bei einer Varianz nahe Null. Für den Wegverlauf der Einstellung E ergeben sich für Schiefe und Wölbung die Werte 1,6 und 2,9 – sind also deutlich unterschiedlich. Als weitere Kenngrößen mit ähnlichem Hintergrund können das Bogenmaß und der Schwerpunkt der Fläche unter einer Kurve berechnet werden.

Kenngrößen für Schwingungssignale

Sowohl der Körperschall als auch die resultierende Schweißkraft in der Werkstückaufnahme sind Schwingungssignale. Diese Signale oszillieren mit einer bestimmten Amplitude und Frequenz um einen Mittelwert, wobei alle drei Größen zeitvariabel sind. Die bisher vorgestellten Größen sind für Schwingungssignale nur bedingt geeignet. Die erste Möglichkeit besteht darin, die Schwingung über einen entsprechenden Tiefpass auszufiltern und den resultierenden Mittelwertverlauf mit den bisher genannten Kenngrößen zu beschreiben. Des Weiteren kann aus den Schwingungssignalen die Einhüllende ermittelt werden. Dieses Signal ist wieder sehr niederfrequent und kann entsprechend der anderen Signale behandelt werden. Ein anderer Ansatz für Schwingungs- und Schallsignale besteht in der Frequenzanalyse, deren Kernelement eine Transformation der Signale in den Frequenzraum ist. Dies geschieht mittels der Fouriertransformation (FT), wobei in der diskreten Signalverarbeitung fast ausschließlich die Fast-Fourier-Transformation (FFT) eingesetzt wird. Diese kommt mit einem Bruchteil der Rechenoperationen aus, setzt dafür aber eine Zweierpotenz als Block-

größe voraus. Das Ergebnis der (F)FT ist das Amplituden- und Phasenspektrum über der Frequenz. Die maximale Frequenz entspricht der halben Abtastfrequenz des Ursprungssignals. Zumeist wird nur das Amplitudenspektrum ausgewertet, da es im Gegensatz zum Phasenspektrum anschaulich ist und einfach interpretiert werden kann. Aus dem Amplitudenspektrum werden dann Kenngrößen wie Maxima, deren Frequenzen sowie mittlere Amplituden oder Integrale über definierte Frequenzbereiche ermittelt. Sind die Grenzen der Frequenzintervalle durch ein festes Verhältnis definiert, spricht man allgemein auch von Ordnungsanalyse – bei einem Frequenzverhältnis von 2:1 speziell von einer Oktavanalyse.

Bei Anwendung der FFT ist es wichtig, sich über die so genannte Unschärferelation im Klaren zu sein. Vereinfacht ausgedrückt bedeutet dies, dass für kurze Zeitsignale die Ermittlung der darin enthaltenen Frequenzen nur in einem relativ groben Raster möglich ist. Um die Auflösung im Frequenzbereich zu erhöhen, müssen längere Zeitabschnitte betrachtet werden. Verändern sich währenddessen die enthaltenen Frequenzen, so liefert die FFT entsprechende Mittelwerte. Weitere Details sind z. B. bei BRIGHAM [Bri88] (sehr mathematisch) oder KARREBERG [Kar09] (sehr anschaulich) zu finden. Für das Ultraschallschweißen ist dies insofern von Bedeutung, als aufgrund der kurzen Schweißzeiten weder eine leichte Verschiebung der Schweißfrequenz noch deren Amplitudenvariation während des Schweißvorgangs ermittelt werden können. Wie in Abschnitt 2.1 dargestellt, lassen beide Größen theoretisch Rückschlüsse auf Änderungen der Dämpfung im Werkstoff zu und sind zur Beschreibung von Prozessveränderungen geeignet. In der Praxis können diese Daten jedoch aufgrund der Unschärferelation nicht hinreichend genau aus den Signalen bestimmt werden.

Prozessspezifische Kenngrößen

Die Kenntnis über Zusammenhänge sowie vorhandenes Prozesswissen können dazu genutzt werden, prozessspezifische Größen zu definieren. Nachteilig dabei ist, dass sie für jede Prozessart separat erarbeitet werden müssen und anschließend eine spezifische Kenngrößenberechnung implementiert werden muss. Andererseits verfügen prozessspezifische Kenngrößen aufgrund ihrer individuellen Herleitung über einen entsprechend hohen Informationsgehalt bezüglich Prozess- und Produktschwankungen. Für das Ultraschallschweißen ist erwiesen, dass der Wegverlauf verschiedene Phasen charakterisiert und die Fügegeschwindigkeit in der stationären Phase eine sehr wichtige Größe darstellt (siehe Abschnitt 2.1.4). Aus dem Wegsignal kann durch numerische Differenziation die Fügegeschwindigkeit über den gesamten Schweißvorgang ermittelt werden. Wird diese nur für die zweite Hälfte der Schweißzeit berechnet, kann dieser Wert als Abschätzung der Fügegeschwindigkeit in der stationären Schweißphase verwendet werden. Weiter bilden sich Geometrieschwankungen der Werkstücke sowie Teilemaße in Schweißrichtung nach dem Schweißen in der Höhe des Wegsignals zu Beginn der Schweißphase und am Ende der Haltephase ab. Die Schweißkräfte inklusive

eventuell vorhandener Kraftanpassungen während des Schweißens können durch eine Mittelwertbildung über entsprechende Zeitfenster sehr genau bestimmt werden. Oftmals lassen sich solche spezifischen Größen aus den zuvor genannten allgemeinen Größen ableiten, wenn die Parameter wie Zeitpunkte oder Ereignisse entsprechend gesetzt werden. So ist auch die Differenz zwischen dem Wert des Wegsignals am Anfang und am Ende der Haltezeit eine sehr aussagekräftige Größe, da sie ein Maß für das Ausquetschen der Schmelze in Kombination mit der Schwindung während des Abkühlens darstellt. Eine solche Verknüpfung von Werten ist zwar theoretisch auch mit allen prozessunabhängigen Größen möglich, macht jedoch nur für entsprechend physikalisch motivierte Größen Sinn, wie sie aus Prozesswissen abgeleitet werden können.

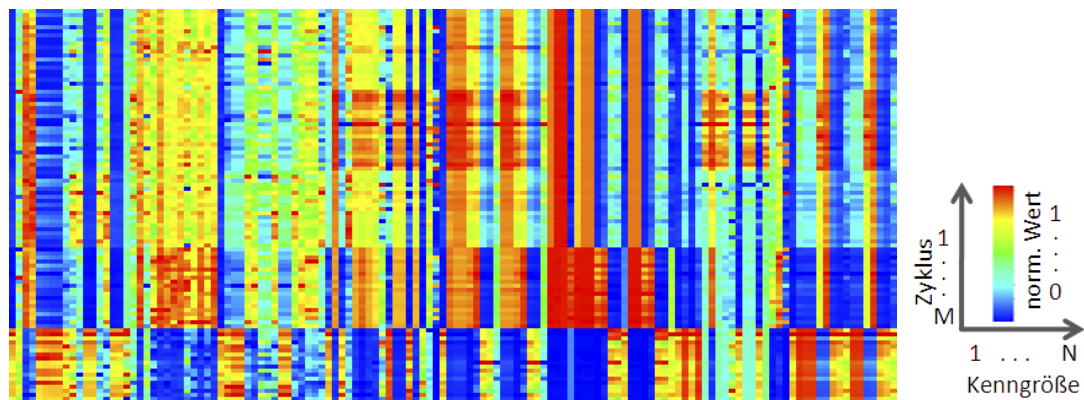
5.2.3 Untersuchung der Kenngrößenvarianz

Die Varianz der Kenngrößen kann mit verschiedenen Methoden untersucht werden. Im Folgenden werden einige kurz vorgestellt.

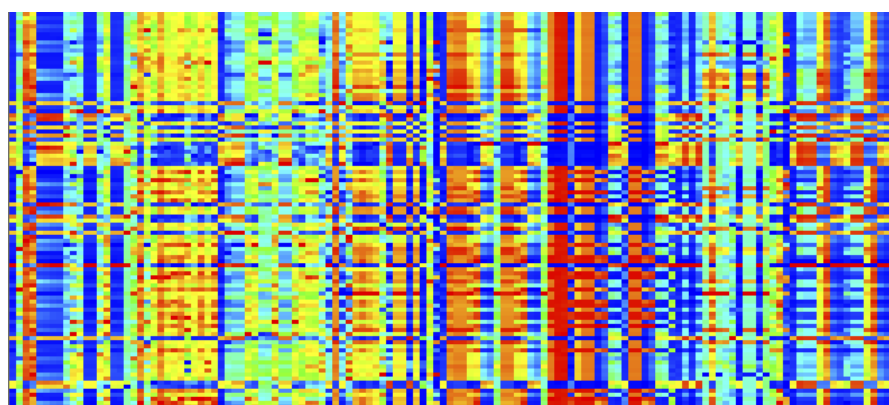
Visualisierung der Kenngrößenwerte

Nachdem die Kenngrößen definiert sind, können ihre Werte für jeden Zyklus berechnet werden. Bei M Zyklen und N Kenngrößen ergeben sich so insgesamt $G = M \cdot N$ Werte, welche die Prozessschwankungen entsprechend abbilden. Einen ersten Überblick über die Kenngrößenvarianzen ermöglicht die Visualisierungsform in Abbildung 5.3, wobei $N = 132$ Kenngrößen entsprechend Abschnitt 5.2.2 für die $M = 98$ Zyklen der Versuchsreihe aus Abschnitt 3.8 dargestellt sind. Jeder Datenzyklus wird in einer Zeile abgebildet, jede Spalte gibt eine Kenngröße wieder und die Farbe repräsentiert den Wert der Kenngröße. Blau steht für den geringsten Wert einer Kenngröße, rot für den höchsten. Um eine einheitliche Farbskalierung zu ermöglichen, wird in einem ersten Schritt jede Kenngröße unter Berücksichtigung ihrer Extremwerte über alle Zyklen auf den Bereich $[0; 1]$ normiert. Um erste Zusammenhänge der Kenngrößen zu einer interessierenden Zielgröße erkennen zu können, werden die Zyklen entsprechend der Zielgröße sortiert. Ohne eine separate Sortierung sind die Zyklen entsprechend des Versuchsplans angeordnet, so dass vorhandene Unterschiede zwischen den einzelnen Einstellungen sichtbar werden¹. In Abbildung 5.3(a) sind die fünf Einstellungen deutlich als horizontale Balken zu erkennen. Anhand der komplementären Farben ist für viele Kenngrößen ersichtlich, dass sie für Einstellung E – abgebildet am unteren Rand – ganz andere Wertebereiche als für die anderen Einstellungen aufweisen. Unter Berücksichtigung der teils sehr verschiedenen Signalverläufe für Einstellung E, zu sehen in Abbildung 5.2, ist dies nahe liegend. Auch für die anderen Einstellungen sind Kenngrößen

¹Vorausgesetzt, dass den Daten ein Versuchsplan zugrunde liegt, und dieser nicht randomisiert durchgeführt wurde.



(a) Zyklen sortiert nach Versuchsplan, d. h. Einstellung



(b) Zyklen sortiert nach steigendem Berstdruck

Abbildung 5.3: Grafische Darstellung der Kenngrößenwerte über die Versuchszyklen

vorhanden, die eine Parameteränderung entsprechend abbilden. Die Tatsache, dass sich teilweise auch breitere vertikale Balken auszubilden scheinen, liegt darin begründet, dass die Kenngrößen blockweise für die einzelnen Signale vorliegen. Es ist klar zu sehen, dass einige Kenngrößen innerhalb der Einstellungen deutliche Varianzen zeigen. Weist eine Kenngröße innerhalb der einzelnen Einstellungen relativ homogene Werte auf – zu erkennen an ähnlichen Farben – und im Vergleich zwischen zwei Einstellungen sehr unterschiedliche, bildet diese Kenngröße ein wesentliches Merkmal zur Unterscheidung der beiden Einstellungen.

In Abbildung 5.3(b) sind die Zyklen nach aufsteigendem Berstdruck sortiert. Insgesamt ist keine große Homogenität erkennbar, d. h. es ist kein direkter (linearer) Zusammenhang zwischen den Kenngrößen und dem Berstdruck als Qualitätsmerkmal vorhanden. Dabei wirkt sich natürlich aus, dass die Berstdrucke der verschiedenen Einstellungen teils sehr nahe beieinander liegen und deutliche Schwankungen aufweisen. Ausschnitte des Berstdruckverlaufs zu gewählten Zyklen sind in Abbildung 5.10 zu sehen. Die Auswahl einer Untermenge an Kenngrößen zur Modellbildung erfolgt mittels der in Abschnitt 5.3 beschriebenen Methoden.

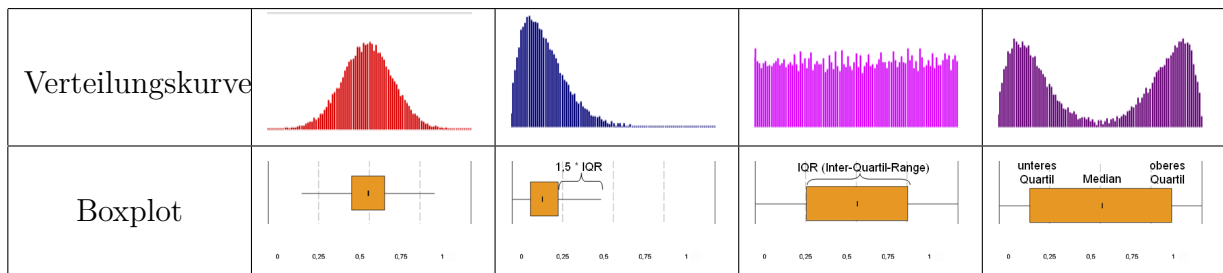


Abbildung 5.4: Verschiedene Verteilungskurven und zugeordnete Boxplots

Boxplots

Eine verbreitete Methode, einen Überblick über die Lage und Streuung einer Größe zu erhalten, sind Boxplots oder Box-Whisker-Plots. Diese Diagramme fassen wesentliche statistische Größen in einem Diagramm zusammen. In Abbildung 5.4 sind die Histogramme für eine große Anzahl zufälliger Werte, die durch einen Zufallszahlengenerator mit einer vorgegebenen Verteilung generiert wurden, nebst den daraus abgeleiteten Boxplots dargestellt. Entsprechend dem Gesetz großer Zahlen ist die jeweilige Verteilungsfunktion am Histogramm erkennbar. Die Messreihen aus Abbildung 5.4 folgen einer Normal-, einer Gamma- und einer Gleichverteilung. Die letzte Verteilung ist die Summe einer Gamma-Verteilung sowie einer an der Mittelachse gespiegelten Gammaverteilung, um eine zweigipfelige Verteilung zu erzielen. Für einen Boxplot sind neben dem Median als Zentralwert auch das obere und untere Quartil notwendig. Für einen Datensatz ermittelt sich das obere Quartil als der Schwellwert, für den 25 % der Daten darüber und 75 % darunter liegen. Entsprechendes gilt für den Median mit 50 % zu 50 % und das untere Quartil mit 75 % darüber und 25 % darunter. Diese drei Werte bilden das mittlere Element eines Boxplots. Durch ihre Lage zueinander sind sowohl der Schwerpunkt als auch die Breite und die Schiefe der Daten schnell abzuschätzen. Der Quartilsabstand wird auch mit IQR bezeichnet. Als weiteres Element besteht ein Boxplot noch aus dem oberen und unteren Whisker, die quasi das obere und untere Limit der Verteilung angeben. Zumeist wird die Länge des Whiskers durch das anderthalbfache des Quartilsabstands definiert. Liegen der minimale oder der maximale Wert aller Messungen zwischen Whiskerende und Quartil, wird der Whisker durch den Extremwert begrenzt. Alle Messwerte, die außerhalb der Whisker liegen, werden als Ausreißer deklariert und entsprechend markiert. Manchmal werden die Whisker unabhängig vom IQR einfach durch Minimal- und Maximalwerte der Kenngrößen festgelegt – Ausreißer gibt es dann innerhalb der Messreihe keine. Boxplots werden häufig für Qualitätswerte eingesetzt, um die Streuung und Robustheit eines Prozesses beurteilen zu können. Zudem ist über Boxplots ein schneller Vergleich mehrerer Einstellungen bezüglich der Lage und Streuung der Zielgröße möglich. Mit zunehmender Anzahl an Einstellungen oder Kenngrößen wird die Darstellung durch Boxplots jedoch unpraktikabel und unübersichtlich, so dass sie sich nicht zur Analyse großer Kenngrößensätze eignen.

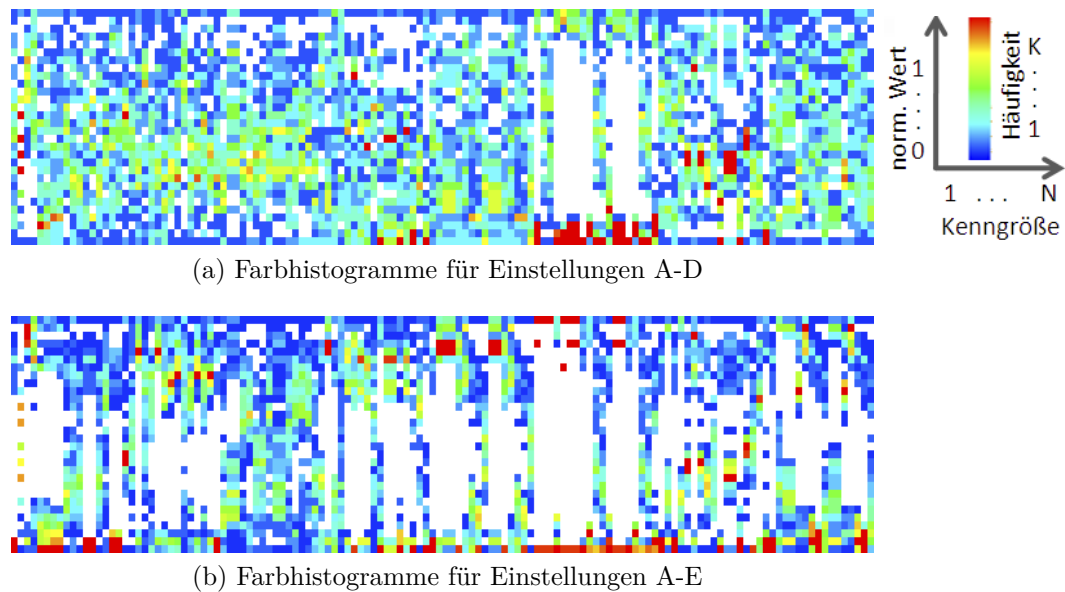


Abbildung 5.5: Grafische Darstellung der Verteilung aller Kenngrößen über Farbhistogramme

Visualisierung der Kenngrößenstreuung

Eine andere Art, die Streuung der N Kenngrößen über alle vorhandenen M Zyklen darzustellen, ist in Abbildung 5.5 zu sehen. Wie in Abbildung 5.3 sind Ausgangspunkt die über alle Zyklen auf $[0; 1]$ normierten Kenngrößen. Auf der Abszisse sind wieder die Kenngrößen angeordnet und die Ordinate bildet den Wertebereich $[0; 1]$ mit Z Zellen gleicher Breite ab. Für jede Kenngröße wird über alle Zyklen ermittelt, wie oft die Kenngröße Werte innerhalb der einzelnen Zellen aufweist, und die Anzahl der „Treffer“ farbcodiert dargestellt. Jede Spalte der Darstellung entspricht somit einem Histogramm der Werteverteilung und die Farbe stellt die Häufigkeit jeder Klasse dar. In Abbildung 5.5(a) ist dies für die 79 Zyklen der Einstellungen A-D der Versuchsreihe aus Abschnitt 3.8 dargestellt, in Abbildung 5.5(b) für die 98 Zyklen aller fünf Einstellungen. Beide Verteilungen sind für $Z = 30$ Histogrammklassen dargestellt. Weiße „Löcher“ stehen für Klassen, denen keine Werte zugeordnet sind. Liegen bei der Anwendung von GRNFN-Modellen einzelne Werte der Eingangsdaten innerhalb solcher „Löcher“ oder Nullklassen, forcieren diese eine Klassifikation als Unbekannt (siehe Abschnitt 4.6). In Abbildung 5.5(a) ist zu erkennen, dass für die Einstellungen A–D nur wenige Kenngrößen größere Löcher aufweisen und die meisten relativ gleichmäßig verteilt sind. Wird wie in Abbildung 5.5(b) jedoch noch die Einstellung E zur Auswertung hinzugezogen, werden für einen Großteil der Kenngrößen weiße Löcher sichtbar. Die mit der geringen Schweißkraft von Einstellung E verbundenen, deutlich geänderten Prozessverläufe führen zu anderen Wertebereichen der Kenngrößen, was sich direkt in deren Verteilung abbildet. Für Analysen ist es folglich wichtig, entsprechend relevante Daten zu verwenden.

5.3 Optimierung der Zustandsvektoren

Wie insbesondere in Kapitel 4 mehrfach dargestellt, ist für gute Modellierungs- und Überwachungsergebnisse eine Optimierung desjenigen Zustandsvektors notwendig, der einen Fertigungszyklus anhand der daraus ermittelten Kenngrößen, dem daraus gebildeten Zustandsvektor und damit einem Datenpunkt in \mathbb{R}^N beschreibt. Die wichtigsten Schritte umfassen dabei die komponentenweise Normierung und Skalierung des Zustandsvektors sowie die Selektion relevanter Kenngrößen. Die einzelnen Punkte sind im Folgenden näher erläutert.

5.3.1 Normierung und Skalierung

Die einzelnen Komponenten eines Zustandsvektors werden durch die verschiedenen berechneten (und durch eine Optimierung ausgewählten) Kenngrößen gebildet. Der Wertebereich der einzelnen Komponenten kann dabei sehr unterschiedlich sein und richtet sich nach der Signalart, der erfolgten Vorverarbeitung (z. B. Verstärkungsfaktoren und Umrechnung) sowie der Art der Kenngrößenbildung. Den einzelnen Komponenten ist zwar eine Einheit zugeordnet, welche bei der Modellbildung jedoch nicht berücksichtigt wird. Dort werden die Komponenten insbesondere bei der Abstandsberechnung miteinander verrechnet, auch wenn sie physikalisch unterschiedliche Einheiten haben. Beispielsweise wird das Leistungssignal der in Abschnitt 3.2 beschriebenen Versuchsanlage als Spannungssignal im Bereich von 0-10 Volt erfasst. Die Kenngröße der maximalen Leistung eines Zyklus liegt also im Wertebereich 0-10. Erfolgt eine Umrechnung auf die physikalische Größe, liegt die maximale Leistung bei einem 4 kW-Generator im Bereich von 0-4 kW oder auch 0-4000 W. Zwei Schweißzyklen mit 2 kW oder 4 kW Maximalleistung, die einzig durch die Kenngröße „maximale Leistung“ beschrieben werden, weisen je nach Beschreibung einen euklidischen Abstand von 5 (V), 2 (kW) oder 2000 (W) auf. Entsprechendes gilt für die Schweißdauer (Angabe in Sekunden oder Millisekunden) sowie für jede weitere Kenngröße. Der Wertebereich der einzelnen Kenngrößen bestimmt damit maßgeblich den Abstand zweier Zustandsvektoren. So hat die Beschreibung der Zustandsvektoren direkten Einfluss auf die Ähnlichkeit von Vektoren, die Clusterbildung und damit auch auf die Ergebnisse der Projektion durch das Sammons-Mapping sowie die Separierbarkeit verschiedener Ereignisse und Zielgrößen. In Abbildung 5.6 ist jeweils für die 98 Zyklen der Beispielversuchsreihe (siehe Abschnitt 3.8) die Projektion mittels Sammons-Mapping (siehe Abschnitt 4.1.2) dargestellt. Die verwendeten Zustandsvektoren bestehen aus einer Untermenge von 30 Kenngrößen², die in Abbildung 5.6(a) entsprechend ihrem ursprünglichen Wertebereich und in Abbildung 5.6(b) in normierter Form als Ausgangsbasis für die Projektionsberechnung verwendet wurden. Die Normierung bildet für jede Kenngröße den in den Daten auftretenden Wertebereich auf das Intervall $[0; 1]$ ab. Jede Kenngröße fließt damit gleichgewichtet in die Abstandsberechnung

²andere wie in Kapitel 4

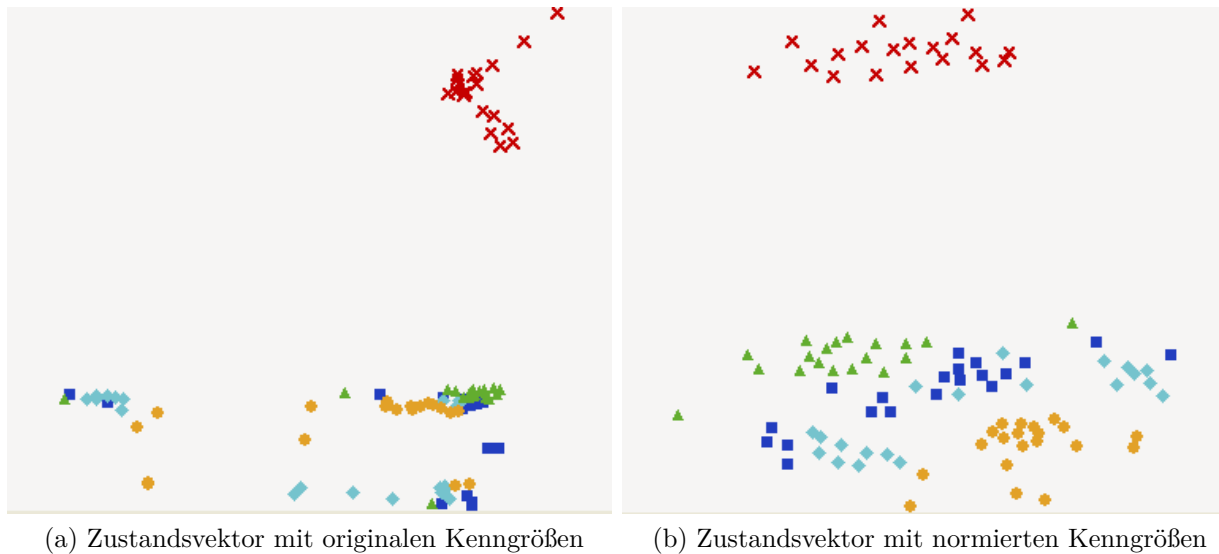


Abbildung 5.6: Darstellung des Einflusses der Normierung

ein. Während sich mit den unveränderten Zustandsvektoren einzig die Einstellung E separieren lässt, wogegen die anderen Einstellungen komplett durchmischt sind, ist allein durch die Normierung der Einzelkomponenten der Zustandsvektoren (größtenteils) eine Trennung der Einstellungen möglich. Wie in Abschnitt 4.5 beschrieben, hat die Separierbarkeit direkten Einfluss auf die Prognosegüte. Wenn für einzelne Kenngrößen eine größere oder geringere Wichtigkeit bekannt ist, können diese Werte durch Einführung von Skalierungsfaktoren entsprechend gewichtet werden.

Für die Prognose von Werten mittels GRNFN-Modellen hat eine Normierung keinen Einfluss, da alle Werte für die Aktivierungsfunktionen in Netzwerkschicht 1 auf das Intervall $[-0,5; 0,5]$ abgebildet werden. Die meisten Indikatoren für die Unbekannt-Klassifikation basieren jedoch auf den nächsten Nachbarn, so dass die Wertebereiche in die Abstandsberechnung eingehen und damit die nächsten Nachbarn bestimmen. Wenn auch für die Prognose nur – wie in Abschnitt 4.6 beschrieben – die nächsten Nachbarn verwendet werden, ist auch die Prognose abhängig von der Normierung der einzelnen Kenngrößen.

5.3.2 Korrelationsanalyse

Für eine erfolgreiche Modellbildung ist es wichtig, Größen als Modelleingangsdaten zu verwenden, die einen Zusammenhang zur Zielgröße aufweisen. Eine häufig verwendete Maßzahl zur Überprüfung eines Zusammenhangs zwischen zwei Messgrößen ist deren Korrelation, genauer gesagt ihr Korrelationskoeffizient. Der Korrelationskoeffizient r nach Bravais-Pearson ist eine dimensionslose Maßzahl der deskriptiven Statistik und beschreibt die Stärke des linearen Zusammenhangs zweier Größen. Dabei wird die Kovarianz der beiden Messreihen X und Y mit dem Produkt ihrer jeweiligen Standardabweichung normiert. Der Korrelati-

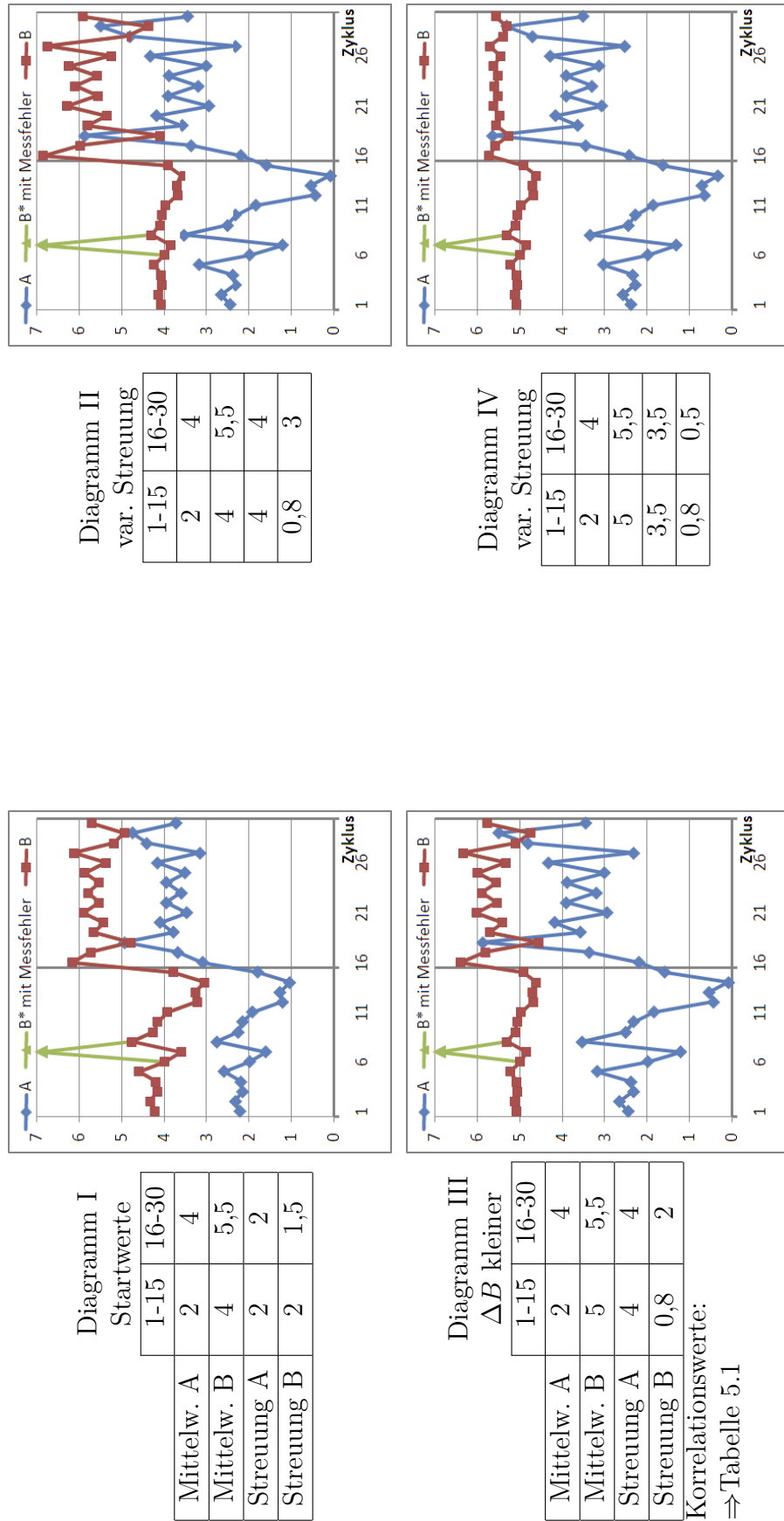


Abbildung 5.7: (Synthetische) Messreihen für die Korrelations- und Signifikanzberechnung

onskoeffizient r errechnet sich anhand einer Reihe von n Wertepaaren $(x_i; y_i)$ entsprechend der Formel

$$r_{XY} = \frac{Cov(X, Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left(\sum_{i=1}^n (x_i - \bar{x})^2\right) \left(\sum_{i=1}^n (y_i - \bar{y})^2\right)}} \quad (5.1)$$

mit $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ und $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ als Mittelwert der beiden Messreihen X und Y . Der Korrelationskoeffizient liefert Werte im Bereich $-1 \leq r_{XY} \leq 1$. Ein Korrelationskoeffizient mit Betrag eins bedeutet dabei einen idealen linearen Zusammenhang, d. h. es gilt $y = ax + b$ (für alle einbezogenen Werte). Für ein negatives Vorzeichen von r_{XY} ist die Geradensteigung negativ und große Werte von y gehen mit kleinen Werten von x einher und umgekehrt. Bei positiven Werten von r_{XY} treten für niedrige Werte von x auch niedrige Werte von y und für hohe x -Werte treten auch hohe y -Werte auf. Ab Werten von ca. 0,8 spricht man von einer hohen Korrelation. Für die Beschreibung des durch ein Modell erklärten Anteils einer auftretenden Varianz wird häufig auch das Bestimmtheitsmaß B mit $0 \leq B \leq 1$ verwendet, welches für die einfache lineare Regression mit dem quadrierten Korrelationskoeffizienten $B \equiv r_{XY}^2$ identisch ist [HE84].

Mit den synthetischen Messreihen in Abbildung 5.7 werden einige Eigenschaften des Korrelationskoeffizienten erläutert. Die Werte wurden erzeugt, indem gleichverteilte Zufallszahlen, die für A und B identisch waren, mit der angegebenen Streubreite symmetrisch um einen Mittelwert verteilt wurden. Für beide Messreihen A und B umfasst jeder der Bereiche 15 Zyklen, wobei in allen Abbildungen die jeweils angegebenen Parameter verwendet wurden. Während bei den ersten 15 Zyklen – als Einstellung 1 bezeichnet – die Zahlen für A und B jeweils zum Mittelwert addiert wurden, wurden sie in der zweiten Hälfte – Einstellung 2 – zu A addiert und von B subtrahiert. Zusätzlich ist von Messreihe B eine Kopie B^* vorhanden, in der ein Wert deutlich erhöht wurde, um einen Ausreißer oder Messfehler zu simulieren. Die sich ergebenden Korrelationswerte und Signifikanzniveaus (siehe Abschnitt 5.3.3) sind in Tabelle 5.1 aufgeführt. Gebildet wurden die Korrelationen zwischen A und der Einstellungs-codierung (15 Elemente mit Wert 1 und 15 mit Wert 2), zwischen A und B sowie zwischen A und B^* – jeweils für die 30 Wertepaare. Da die Messreihen gleichgerichtet sind, sind alle Korrelationskoeffizienten positiv. Obwohl visuell in allen vier Fällen ein deutlicher Zusammenhang erkennbar ist, schwanken die Korrelationskoeffizienten für den Zusammenhang zwischen A und B stark. Einzig im ersten Fall kann mit $r_{AB} = 0,80$ noch von einer hohen Korrelation gesprochen werden. Die Variation der Streubreite oder der Differenz der Mittelwerte in einer Messreihe wirken sich teils stark auf den Korrelationskoeffizienten aus. Im dritten Fall kann bei einem Wert $r_{AB} = 0,17$ nicht mehr von einer Korrelation gesprochen werden. Die Messreihe B könnte dabei die Ergebnisse eines Versuchsplans eines verfahrenstechnischen Prozesses darstellen: Mit einer Erhöhung der Stellgröße A wurde die Ausbeute B von 5,0 auf 5,5 um 10% mit dem Nebeneffekt einer

Diagramm	Korrelation			Signifikanz (%)		
	A – Einst	A – B	A – B*	A	B	B*
I	0,89	0,80	0,61	100	100	100
II	0,68	0,39	0,21	100	100	100
III	0,68	0,17	0,00	100	99,96	98,32
IV	0,73	0,68	0,31	100	100	98,88

Tabelle 5.1: Korrelations- und Signifikanzwerte zu Abbildung 5.7

höheren Streuung gesteigert, ohne dass dies am Korrelationsfaktor zu erkennen wäre. Folglich beeinflussen überlagerte Streuungen die Korrelationswerte teilweise sehr stark, so dass die Korrelationsanalyse vorhandene (lineare) Zusammenhänge nicht immer aufdeckt, insbesondere wenn – wie bei Festigkeitswerten – mit nicht unerheblichen Streuungen zu rechnen ist. Eng damit verbunden ist der Einfluss von Messfehlern, wie sie auch in Abbildung 5.7 durch die Messreihe B^* abgebildet sind. Ein einziger deutlicher Ausreißer aufgrund eines Messfehlers reduziert die Korrelationswerte bereits erheblich. Für die Korrelationsanalyse ist es wichtig, dass Ausreißer zuvor identifiziert und – sofern es sich um Messfehler handelt – geglättet oder gestrichen werden. Umgekehrt dürfen Werte nicht deshalb gestrichen werden, weil sie nicht in die Reihe passen und die Korrelation reduzieren. Eine Streichung ist nur zulässig, wenn es sich tatsächlich um Messfehler handelt.

Bei Anwendung des Pearsonschen Korrelationskoeffizienten r_{XY} sind zudem folgende Punkte zu beachten, um falsche Schlussfolgerungen zu vermeiden:

1. Der Korrelationskoeffizient berücksichtigt nur einen *linearen* Zusammenhang. Ein Wert von Null bedeutet demnach nicht, dass *kein* Zusammenhang zwischen x und y besteht, sondern nur, dass kein *linearer* Zusammenhang existiert. So ist für Messwerte der Normalparabel $y = x^2$ der Korrelationskoeffizient $r_{XY} = 0$, wenn ausschließlich korrespondierende x -Werte, d. h. $\pm x$, sowie zugehörige y -Werte für die Berechnung verwendet werden.
2. Von *Scheinkorrelation* ist die Rede, wenn die Korrelation von zwei Größen x und y berechnet wird, die jeweils linear mit einer dritten, nicht berücksichtigten Größe z gekoppelt sind. Insbesondere tritt dies ein, wenn physikalisch nicht berücksichtigte Wirkzusammenhänge existieren. Dies ist anhand einer Versuchsreihe mit Probekörpern aus dem Werkstoff SAN in Abbildung 5.8 dargestellt. Die aus Prozesssignalen berechneten Größen Schweißzeit und Leistungsintegral weisen eine (scheinbar) hohe Korrelation zum Qualitätsmerkmal Bruchkraft auf und scheinen deswegen ideal zur Modellbildung. Wenn jedoch zusätzlich die Größe des Schweißwegs betrachtet wird, ist zu erkennen, dass dieser während der Versuche variiert wurde. Dies ist der eigentliche Grund für die Varianzen in Leistung, Schweißzeit und Bruchkraft. Höhere

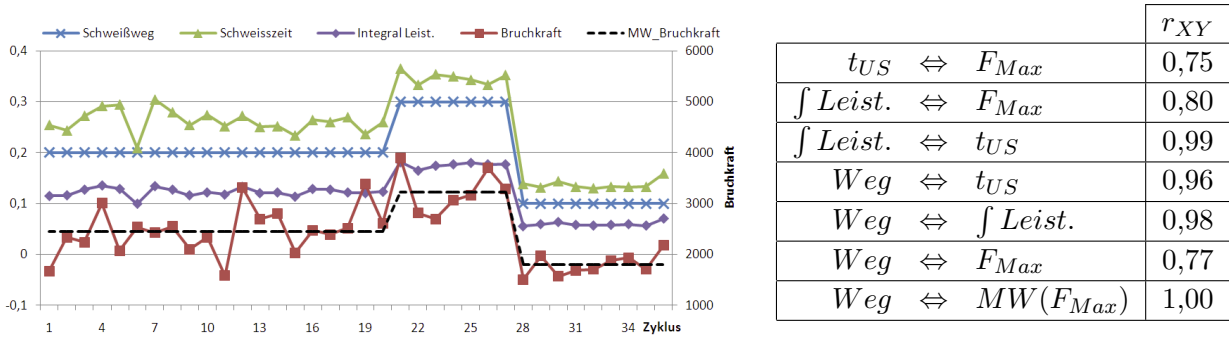


Abbildung 5.8: (Schein)Korrelation bei Kenngrößen und Qualität für Probekörper aus SAN

Schweißwege führen durch das größere Schmelzevolumen zu höherer Festigkeit, aber gleichzeitig sind längere Schweißzeiten sowie mehr Energieeintrag notwendig.

3. Wenn für willkürliche Größen die Korrelation berechnet wird, von denen aufgrund vorhandenen Erfahrungs- oder Prozesswissens kein kausaler Zusammenhang plausibel ist, besteht die Gefahr der *Nonsenskorrelation*. So weist beim Vergleich verschiedener Länder die Geburtenrate rechnerisch eine hohe Korrelation mit der Anzahl der im Land lebenden Störche auf – ein kausaler Zusammenhang besteht jedoch nicht [HEK09].
4. Bei der Korrelationsanalyse sollte beim Einfluss *mehrerer Größen jede separat* analysiert werden, da ansonsten existierende Effekte durch die Überlagerung nicht erkannt werden und verloren gehen. In Abbildung 5.9 ist für die in Abschnitt 3.8 beschriebene Versuchsreihe der Verlauf dreier Kenngrößen dargestellt, die aus dem Leistungs- oder Amplitudensignal abgeleitet wurden. Ferner ist markiert, bei welchen Zyklen welche Versuchsvariation durchgeführt wurde. In der Tabelle in Abbildung 5.9 rechts sind die Korrelationskoeffizienten angegeben, die sich für die Kenngrößen in Relation zu den verschiedenen Einstellgrößen ergeben. Die Wertepaare für jeden Zyklus bestehen also aus dem Wert der Kenngröße sowie dem eingestellten Wert der Maschineneinstellung. Während in der linken Spalte die Korrelation über alle 98 Zyklen ermittelt wurde, sind für die Werte der rechten Seite nur die Zyklen der Arbeitspunkt-Einstellung sowie diejenigen verwendet, in denen der korrelierte Einstellparameter verändert wurde. Die Zyklen, in denen die anderen Parameter variiert wurden, sind also nicht enthalten. Wie in Abbildung 5.9 ergeben sich mit beiden Kenngrößen für beide Datenmengen sehr deutliche Unterschiede. Einzig für die Korrelation der Kenngrößen mit der Veränderung der Schweißkraft ergeben sich immer hohe Werte. Dies hängt insbesondere damit zusammen, dass die Schweißkraft stark verändert wurde. Dies führte zu einem deutlich geänderten Schweißverhalten und signifikant anderen Kenngrößen. Die Änderung des Kenngrößenniveaus bei niedriger Kraft ist um ein Vielfaches höher als die Streubreiten bei den anderen Einstellungen. Der Unterschied der Mittelwerte zwei-

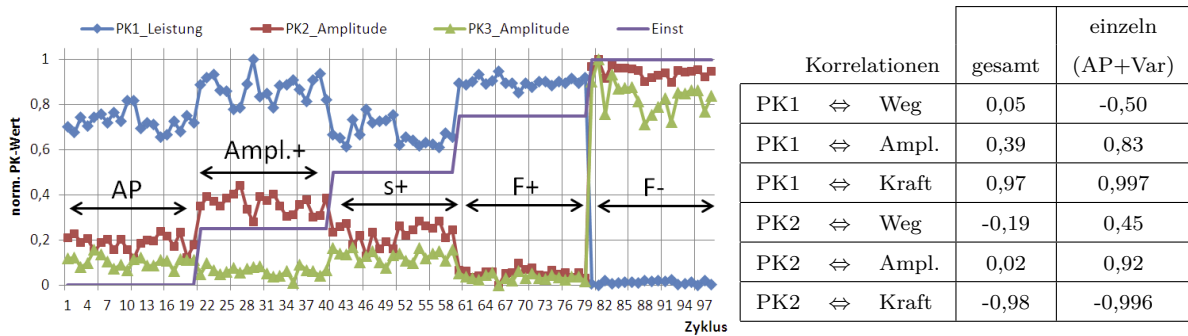


Abbildung 5.9: Verlauf exemplarischer Kenngrößen sowie variierte Einstellung

er Einstellungen im Verhältnis zur Streuung hat einen deutlichen Einfluss auf die erzielbaren Korrelationsfaktoren (vergleiche Abbildung 5.7).

Für die Untersuchung auf Korrelationen wurden im Rahmen dieser Arbeit mehrere Varianten umgesetzt, um die paarweisen Korrelationskoeffizienten zwischen Kenngrößen und Zielgrößen weitestgehend automatisiert bestimmen zu können. Die erste Variante besteht darin, die Korrelationskoeffizienten aller Merkmalspaarungen über alle verfügbaren Zyklen zu berechnen, um einen ersten Überblick erhalten zu können. Um den im vorigen Abschnitt beschriebenen Problemen bei der Korrelationsanalyse Rechnung zu tragen, können die Werte für r_{XY} auch nur über Teilbereiche mit frei wählbarer Länge berechnet werden. Diese Teilbereiche sind durch ihren Startzyklus sowie ihre Länge definiert und können durch sukzessive Erhöhung des Offsets über die gesamte verfügbare Datenmenge verschoben werden. Pro Paarung aus Kenn- und Zielgröße ergeben sich so mehrere Korrelationskoeffizienten, wobei insbesondere das Maximum sowie der Mittelwert über alle Bereiche interessant sind. So können Zusammenhänge teils auch bei existierenden Wechselwirkungen zwischen den Parametern identifiziert werden, d. h. dann, wenn sich die Auswirkungen eines Parameters auf die Zielgröße nur zeigen, wenn ein weiterer Parameter auf einer bestimmten Stufe steht. Auch den zuvor unter Punkt 4 genannten Einschränkungen kann begegnet werden, indem die Daten umsortiert und die Teilbereiche entsprechend den durchgeführten Variationen gewählt werden. Mittels einer weiteren realisierten Variante kann der in Abbildung 5.7 beschriebenen negativen Auswirkung von Streuungen innerhalb der einzelnen Einstellung entgegen gewirkt werden. Für alle Größen werden automatisiert die Mittelwerte pro Versuchseinstellung, die durch ein entsprechendes Merkmal eindeutig gekennzeichnet ist, ermittelt und die Korrelationswerte anhand der Mittelwerte über alle Einstellungen berechnet. Dies ist allerdings nicht möglich, wenn nur zwei Einstellungen vorhanden sind. Abgesehen vom Sonderfall des gleichen Mittelwerts für eine Größe ergibt sich für diesen Fall immer ein Korrelationskoeffizient mit Wert eins. Wenn $x_1 < x_2$ ist, gibt es für die Werte y des zweiten Korrelationspartners nur zwei Möglichkeiten: Wenn $y_1 < y_2$ ist, sind x und y gleichgerichtet und es ergibt sich $r_{XY} = +1$, für $y_1 > y_2$ ergibt sich dann entsprechend $r_{XY} = -1$.

Durch die Außerachtlassung der Streuung ist auch keine Aussage darüber möglich, mit welcher Wahrscheinlichkeit der berechnete Mittelwert dem tatsächlichen entspricht. Somit sind auf diese Weise – zumindest wenn die Mittelwerte relativ nahe beieinander liegen – leicht falsche Folgerungen möglich. Für den Vergleich zweier Mittelwerte ist es deshalb besser, entsprechende statistische Verfahren zu verwenden. Dies führt zu den Mittelwert- und Signifikanzanalysen, die im Folgenden beschrieben werden.

5.3.3 Mittelwertvergleich und Signifikanzanalyse

Bei der Variation oder Optimierung eines Prozesses ist eine häufige Fragestellung, ob sich eine Zielgröße im Mittel verändert hat. Beim Ultraschallschweißen ist beispielsweise interessant, ob eine Veränderung der Schweißkraft zu einer höheren Festigkeit der Schweißnaht führt. Bei der Auswahl von Kenngrößen stellt sich die Frage, ob sich der (Mittel)Wert einer Kenngröße in einer Einstellung von dem einer geänderten Einstellung unterscheidet und somit als Indikator für eine entsprechende Veränderung dienen kann. Die Differenz zwischen zwei Mittelwerten wird auch Effekt genannt. Die Abbildung eines Zusammenhangs mittels Modellen kann nur erfolgen, wenn Informationen zu den vorhandenen Effekten in den Daten vorhanden sind, die dem Netz als Eingang zur Verfügung gestellt werden. Eine Aussage zu solchen Fragestellungen liefern Mittelwertvergleiche.

Berechnung von Mittelwerten

Ein Vergleich von Mittelwerten – gleichbedeutend mit der Ermittlung eines Effekts – macht jedoch nur Sinn, wenn die Werte hinlänglich genau ermittelt werden können. Allein das arithmetische Mittel aus mehreren Größen ist nicht aussagefähig, da die Streuung vernachlässigt wird. Bereits in Abschnitt 5.3.2 wurde deutlich, dass sich die Höhe des Unterschieds sowie die Streubreite deutlich in statistischen Kenngrößen bemerkbar machen. Bei jeder Messreihe treten in den einzelnen Größen mehr oder weniger große Streuungen auf. Dies gilt immer und ist unabhängig davon, ob die Messreihe aus den ermittelten Festigkeiten von zehn Werkstücken besteht oder ob an einem Werkstück dasselbe Maß mehrfach gemessen wurde. Je kleiner die Streuung ist, desto näher liegt der aus der Messreihe ermittelte Mittelwert \bar{x} am realen Mittelwert μ des Merkmals, der als Erwartungswert bezeichnet wird. In der Statistik wird der Bereich, in dem der Erwartungswert μ mit einer vorgegebenen Wahrscheinlichkeit liegt, Konfidenzintervall genannt. Das Konfidenzintervall ist für jede Messreihe spezifisch, d. h. bei einer Wiederholung der Messreihe ergeben sich leicht unterschiedliche Werte. Die Grenzen des Konfidenzintervalls hängen von mehreren Faktoren ab: Mittelwert und Streuung der Messwerte, Anzahl der Messwerte und der geforderten Vertrauenswahrscheinlichkeit, auch als Signifikanzniveau bezeichnet. Keine Statistik kann als Ergebnis eine absolute Gewissheit liefern, sondern alle Aussagen beziehen sich auf eine

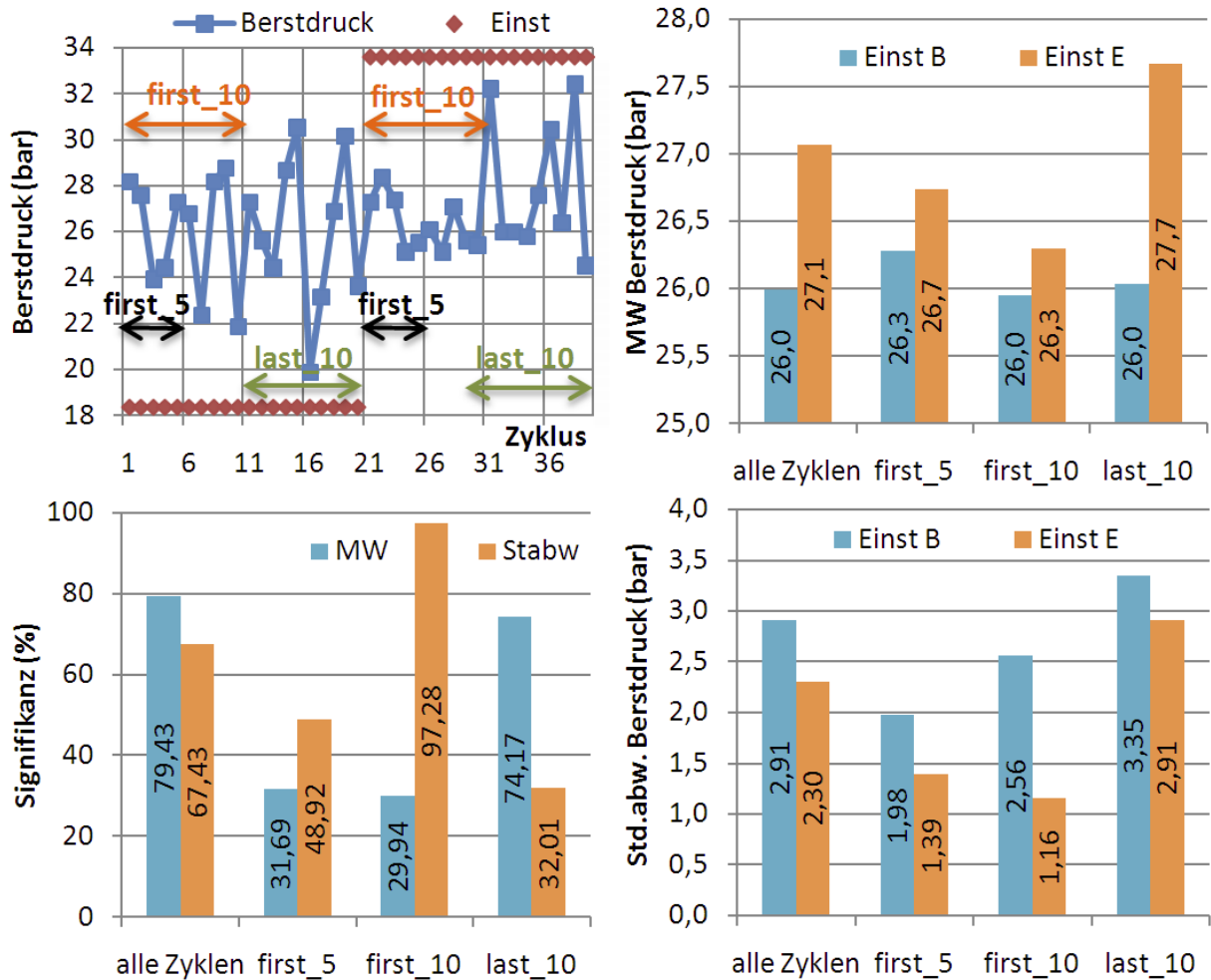


Abbildung 5.10: Verlauf des Berstdrucks und statistische Größen verschiedener Ausschnitte

(maximal zulässige) Irrtumswahrscheinlichkeit α , die direkt mit dem Signifikanzniveau P über $P = 1 - \alpha$ verknüpft ist. Grundsätzlich gelten folgende Zusammenhänge:

- Je kleiner die Streuung, desto besser kann der Erwartungswert abgeschätzt werden und desto enger ist das Konfidenzintervall,
- je größer die Anzahl an Messwerten ist, desto besser kann der Erwartungswert abgeschätzt werden – auch bei größeren Streuungen,
- je größer das Signifikanzniveau, je kleiner also die Irrtumswahrscheinlichkeit ist, desto breiter ist das Konfidenzintervall.

In Abbildung 5.10 sind der Verlauf des Berstdrucks für die beiden Einstellungen B und E der in Abschnitt 3.8 eingeführten Versuchsreihe dargestellt. Zudem sind als statistische Größen der Mittelwert pro Einstellung, die Standardabweichung sowie die Signifikanz für Unterschiede im Mittelwert und der Standardabweichung zwischen den Einstellungen B und E angegeben. Um die angesprochene Abhängigkeit dieser Größen von der zugrunde liegenden Stichprobe zu verdeutlichen, wurden alle Größen für mehrere Bereiche berechnet.

Basierend auf diesen Daten wurden die Konfidenzintervalle für die verschiedenen Bereiche der Einstellungen B und E für mehrere Signifikanzniveaus berechnet und sind in Tabelle 5.2 dargestellt. Die Grenzen der Konfidenzintervalle ergeben sich aus den n Messwerten anhand der Schätzungen für den Mittelwert

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (5.2)$$

und der berechneten Standardabweichung

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (5.3)$$

entsprechend der Formel

$$\text{Konfidenzintervall}_{Min} = \bar{x} - \frac{t_{krit} \cdot s}{\sqrt{n}} < \mu < \bar{x} + \frac{t_{krit} \cdot s}{\sqrt{n}} = \text{Konfidenzintervall}_{Max}. \quad (5.4)$$

Die darin verwendeten Werte t_{krit} ergeben sich aus der Studentschen t-Verteilung mit Mittelwert 0 und Streuung 1. In die Verteilung fließt die Anzahl der Messpunkte in Form des Freiheitsgrads $\nu = n - 1$ ein, so dass sich bei wenigen Messpunkten – gleichbedeutend mit größerer Unsicherheit – höhere t-Werte ergeben. Für $\nu = 1$ entspricht die Student-Verteilung der Cauchy-Verteilung. Andererseits erhält man für $\nu \rightarrow \infty$ die Standard-Normalverteilung. In den kritischen Wert t_{krit} , der die Grenze zur Annahme oder Ablehnung der zugrunde liegenden Nullhypothese H_0 – wie sie für jede statistische Aussage zuerst aufgestellt werden muss – bestimmt, fließt die gewünschte Irrtumswahrscheinlichkeit $\alpha \in (0; 1)$ und das gewünschte Signifikanzniveau $1 - \alpha$ ein. Meist werden die Werte t_{krit} entsprechenden Tabellen entnommen, deren Spalten und Zeilen durch das gewünschte Signifikanzniveau und den Freiheitsgrad definiert werden. Beim Ermitteln der t-Werte muss berücksichtigt werden, ob es sich um eine ein- oder zweiseitige Nullhypothese handelt und ob dementsprechend die Werte $t(\nu, 1 - \alpha)$ oder $t(\nu, 1 - \frac{\alpha}{2})$ verwendet werden müssen. Die Berechnung von Werten der t-Verteilung ist nicht trivial, da sie die Gamma-Funktion mit uneigentlichem Integral enthält. Die Wahrscheinlichkeitsdichte der Student-t-Verteilung mit ν Freiheitsgraden, Mittelwert μ und Standardabweichung σ berechnet sich nach

$$p_{\nu, \mu, \sigma}(x) = \frac{\Gamma(\frac{1}{2}(\nu + 1))}{\Gamma(\frac{1}{2}\nu) \sqrt{\nu\pi}\sigma} \left(1 + \frac{1}{\nu} \left(\frac{x - \mu}{\sigma}\right)^2\right)^{-\frac{1}{2}(\nu+1)}; \quad -\infty < x < +\infty \quad (5.5)$$

mit der Gamma-Funktion $\Gamma: \mathbb{R} \rightarrow \mathbb{R}$ definiert durch

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt \quad \text{für alle } x \in \mathbb{R}.$$

Der t -Wert zu einem gegebenen Signifikanzniveau $P \in [0; 1]$ – auch als P -Quantil ξ_P bezeichnet – ist der Wert, für den das Integral $\int_{-\infty}^t p(x)dx$ mit $p(x)$ aus Gleichung 5.5 genau P ergibt. Für weitere Details und Hintergründe zur Statistik wird auf die Fachliteratur verwiesen (z. B. [HEK09]). Für die Berechnung von t -Werten sowie entsprechenden Wahrscheinlichkeiten in Verbindung mit Mittelwertvergleichen wurden im Rahmen dieser Arbeit numerische Algorithmen aus PRESS ET AL. [PTVF07] verwendet.

Für die nachfolgende Beschreibung des Mittelwertvergleichs und die darauf aufbauende Signifikanzanalyse zur Kenngrößenauswahl sind folgende Punkte relevant, die anhand der Werte in Tabelle 5.2 nachvollzogen werden können:

- Eine höhere Anzahl an Versuchspunkten hat kleinere Werte von t_{krit} und somit engere Grenzen des Konfidenzintervalls zur Folge (bei vergleichbarer Streuung).
- Höhere Werte des Signifikanzniveaus P , d. h. die Forderung nach einer geringeren Irrtumswahrscheinlichkeit α , führen zu höheren Werten von t_{krit} und damit zu breiteren Konfidenzintervallen.
- Die Streuung der Stichprobe hat maßgeblichen Einfluss auf die Intervallbreite. Für ein Signifikanzniveau von 99 % ist der Bereich, in dem der Erwartungswert für den Berstdruck liegt, im Mittel über beide Einstellungen über 5 bar breit. Ursächlich für die Breite sind die hohen Ausreißer, die zumindest teilweise in der Unsicherheit der Festigkeitsprüfung begründet liegen.
- Jede Stichprobe führt aufgrund verschiedener Schätzwerte für Mittelwert und Standardabweichung zu anderen Konfidenzintervallen, die teilweise relativ stark streuen. Bei 99 % Signifikanz schwankt für Einstellung B die Breite zwischen 3,72 und 8,15 bar – bei einem geschätzten Mittelwert des Berstdrucks von 26 bar –, was bis zu 31,3 % des Endwerts entspricht.

Vergleich von Mittelwerten

Aufgrund der Tatsache, dass allein die Bestimmung eines Mittelwerts nicht trivial ist und statistisch zu einem Konfidenzintervall in Verbindung mit einem Signifikanzniveau führt, erscheint ein Vergleich der Mittelwerte \bar{x} und \bar{y} zweier Messreihen (berechnet nach Gleichung 5.2) noch schwieriger. Der direkte Vergleich der empirischen berechneten Mittelwerte \bar{x} und \bar{y} über einen einfachen Relationsoperator wie größer „>“ oder kleiner „<“ ist – für eine statistisch gesicherte Aussage – unzulässig. Für den Vergleich zweier Werte gibt es entsprechende, statistisch gesicherte Verfahren, die wiederum auf der Student- t -Verteilung aufsetzen. Dabei ist zu unterscheiden, ob ein Vergleich mit ein- oder zweiseitiger Hypothese vorliegt. Die Fragestellung – auch Nullhypothese H_0 – bei einem einseitigen Test lautet: „Ist $\bar{x} \leq \bar{y}$ “, während der zweiseitige Test die Frage „Ist $\bar{x} = \bar{y}$ “ zu beantworten sucht. Zu

Daten	alle Zyklen			first_5			first_10			last_10		
	95 %	99 %	99,9 %	95 %	99 %	99,9 %	95 %	99 %	99,9 %	95 %	99 %	99,9 %
Sig.niv.												
B: t_{krit}	2,093	2,861	3,883	2,776	4,604	8,610	2,262	3,250	4,781	2,262	3,250	4,781
B: Min	24,63	24,13	23,47	23,82	22,20	18,66	24,12	23,32	22,07	23,63	22,58	20,96
B: Max	27,35	27,85	28,51	28,74	30,36	33,90	27,78	28,58	29,83	28,43	29,48	31,10
B: Breite	2,72	3,72	5,05	4,91	8,15	15,24	3,67	5,27	7,75	4,80	6,89	10,14
E: t_{krit}	2,101	2,878	3,922	2,776	4,604	8,610	2,262	3,250	4,781	2,262	3,250	4,781
E: Min	25,96	25,55	25,00	25,01	23,88	21,39	25,47	25,11	24,55	25,59	24,68	23,27
E: Max	28,18	28,59	29,14	28,47	29,60	32,09	27,13	27,49	28,05	29,75	30,66	32,07
E: Breite	2,22	3,04	4,14	3,45	5,73	10,71	1,66	2,39	3,51	4,16	5,98	8,80

Tabelle 5.2: Berechnete Konfidenzintervalle und kritische Student-t-Werte zu Abbildung 5.10

jeder Nullhypothese gibt es eine komplementäre Alternativhypothese H_A , die entsprechend „ $\bar{x} > \bar{y}$ “ für den einseitigen und „ $\bar{x} \neq \bar{y}$ “ bzw. „ $\bar{x} < \bar{y} \vee \bar{x} > \bar{y}$ “ für den zweiseitigen Test lautet. Unabhängig von der Formulierung liegt jeder Annahme oder Ablehnung einer Hypothese eine Irrtumswahrscheinlichkeit α zugrunde, die dem Signifikanzniveau $P = 1 - \alpha$ entspricht. Folglich impliziert die Aussage „ \bar{x} unterscheidet sich signifikant von \bar{y} “ immer eine Wahrscheinlichkeit. Häufig werden 95 %, 99 % oder 99,9 % verwendet.

Die Grundidee beim Mittelwertvergleich ist, die Differenz der Mittelwerte – als Effekt bezeichnet – zu berechnen und zu prüfen, ob der Wert 0 im Konfidenzintervall des Effekts enthalten ist. Der Effekt $\bar{d} = \bar{x} - \bar{y}$ ist ein Schätzer für den wahren Unterschied δ der Mittelwerte und für die Grenzen des Konfidenzintervalls wird – vereinfacht ausgedrückt – entsprechend das im vorigen Abschnitt beschriebene Verfahren verwendet. Enthält das Konfidenzintervall den Wert 0, bedeutet dies, dass zum gewählten Signifikanzniveau kein Effekt vorliegt, und somit die Mittelwerte nicht verschieden sind.

Für das Konfidenzintervall nach Gleichung 5.4 werden die Standardabweichung s der Stichprobe sowie der Umfang n benötigt. Beim Mittelwertvergleich liegen im Grunde zwei Stichproben vor – für jede Gruppe oder Einstellung jeweils eine. Die Schätzung der Standardabweichung s kann nur bei gleicher Varianz und gleichem Umfang der Stichproben direkt nach Gleichung 5.3 berechnet werden. Bei unterschiedlichen Stichprobenumfängen oder unterschiedlicher Varianz sind entsprechende Anpassungen notwendig. Die Prüfung der Varianz auf Gleichheit erfolgt dabei mit dem so genannten F-Test. Für Details wird auf die Literatur verwiesen, z. B. HARTUNG/ELPELT oder SCHEFFLER [HE84, Sch97].

Anstatt die Konfidenzintervalle zum Effekt anhand einer geschätzten Standardabweichung zu verwenden, kann auch auf der Basis der Stichproben für beide Gruppen eine Schätzung des resultierenden t -Werts der Student-Verteilung als Prüfgröße vorgenommen werden. Ein Vergleich mit dem kritischen Wert t_{krit} , der auch bei der Berechnung der Konfidenzintervalle verwendet wird, führt zur Akzeptanz oder Ablehnung der gestellten Hypothese. Die von PRESS ET AL. [PTVF07] vorgestellten Algorithmen folgen diesem Ansatz, wobei es unterschiedliche Varianten gibt – je nachdem, ob davon ausgegangen werden kann, dass beiden Gruppen eine identische oder eine unterschiedliche Varianz zugrunde liegt. Die Verfahren für eine unterschiedliche Varianz liefern dabei konservativer geschätzte Werte, so dass diese immer angewendet werden können und zu robusten und statistisch abgesicherten Aussagen führen.

Der Vollständigkeit halber sei noch erwähnt, dass ein Mittelwertvergleich voraussetzt, dass die Werte beider zu vergleichenden Gruppen jeweils normalverteilt sind. Dies kann wiederum mit Tests wie beispielsweise dem Kolmogorow-Smirnow-Test³ – auch KS-Test genannt – oder dem Shapiro-Wilk-Test erfolgen. Während der Nachweis, dass den Daten eine Normalverteilung zugrunde liegt, sehr schwierig ist, reagieren die Tests jedoch sehr

³andere Schreibweisen sind Kolmogoroff-Smirnoff oder Kolmogorov-Smirnov

empfindlich, wenn die Daten *nicht* normalverteilt sind, so dass für robuste Analysen diese Hypothese daraufhin geprüft werden sollte.

Signifikanzanalyse zur Kenngrößenselektion

Auf der Basis statistischer Mittelwertvergleiche, wie im vorigen Abschnitt beschrieben, erfolgt der Aufbau einer Signifikanzanalyse mit dem Ziel der Kenngrößenselektion. Das Ergebnis des Mittelwertvergleichs liefert eine boolesche Aussage dahingehend, ob zwei Messreihen oder Gruppen an Werten einen unterschiedlichen Mittelwert bezüglich eines zuvor festgelegten Signifikanzniveaus (z. B. 99 %) aufweisen oder nicht. Dies ist auch dem Umstand geschuldet, dass die t-Verteilung nicht einfach berechnet werden kann und die kritischen Werte t_{krit} , welche die Grenze für die Signifikanz eines Effekts bedeuten, in der Regel aus Tabellen mit einem definierten Satz an Signifikanzniveaus entnommen wurden. Numerische Algorithmen ermöglichen es zwischenzeitlich, die boolesche Ja/Nein-Aussage dahingehend zu erweitern, dass als Ergebnis eines statistischen Vergleichs die Wahrscheinlichkeit für einen unterschiedlichen Mittelwert direkt berechnet wird. Ein Vergleich der so berechneten Werte mit vorgegebenen Signifikanzniveaus ergibt dieselben Resultate wie der „klassische“ Mittelwertvergleich.

Mittels dieser berechneten Wahrscheinlichkeit kann für eine Kenngröße – unabhängig davon, ob es sich um einen Qualitätswert oder eine aus Signalen abgeleitete Kenngröße handelt – ihre Fähigkeit als Indikator für geänderte Zustände beurteilt werden. Hierfür sind zuerst Bedingungen festzulegen, anhand derer aus den zu untersuchenden Daten zwei Gruppen definiert werden. Anschließend werden die Zyklen entsprechend der Bedingungen in die zwei Gruppen einsortiert. Dabei ist es nicht notwendig, dass alle Zyklen verwendet werden oder jeder Zyklus zwingend nur einer Gruppe zugeordnet wird – obwohl die zuletzt genannte Forderung meist sinnvoll ist. Dann wird für alle Kenngrößen mittels eines statistischen Mittelwertvergleichs ermittelt, mit welcher Wahrscheinlichkeit sich die Mittelwerte der so gebildeten beiden Gruppen für diese Kenngröße unterscheiden.

In Abbildung 5.11 sind die Ergebnisse dieses Vorgehens für 40 Kenngrößen der Beispielsversuchsreihe (siehe Abschnitt 3.8) dargestellt, die aus Prozesssignalen berechnet wurden. Die Gruppen in Abbildung 5.11(a) wurden entsprechend der Einstellungen gebildet. Im ersten Fall bilden alle Zyklen der Einstellung A die erste und alle Zyklen der Einstellung C die zweite Gruppe. Wie schon aus Abbildung 4.13(a) ersichtlich, sind sich diese beiden Einstellungen sehr ähnlich⁴ und können demnach durch ein Modell schlecht getrennt werden (siehe auch Abschnitt 4.5). Durch die Signifikanzanalyse in dieser Gruppeneinteilung können diejenigen Kenngrößen ermittelt werden, die am besten geeignet sind, diese beiden Einstellungen zu separieren. Der zweite Fall in Abbildung 5.11(a) dient zur Identifikation der Kenngrößen, die es ermöglichen, die Einstellung A von der „Restmenge“ – bestehend

⁴immer in Bezug auf die für die Abbildung verwendeten Zustandsvektoren

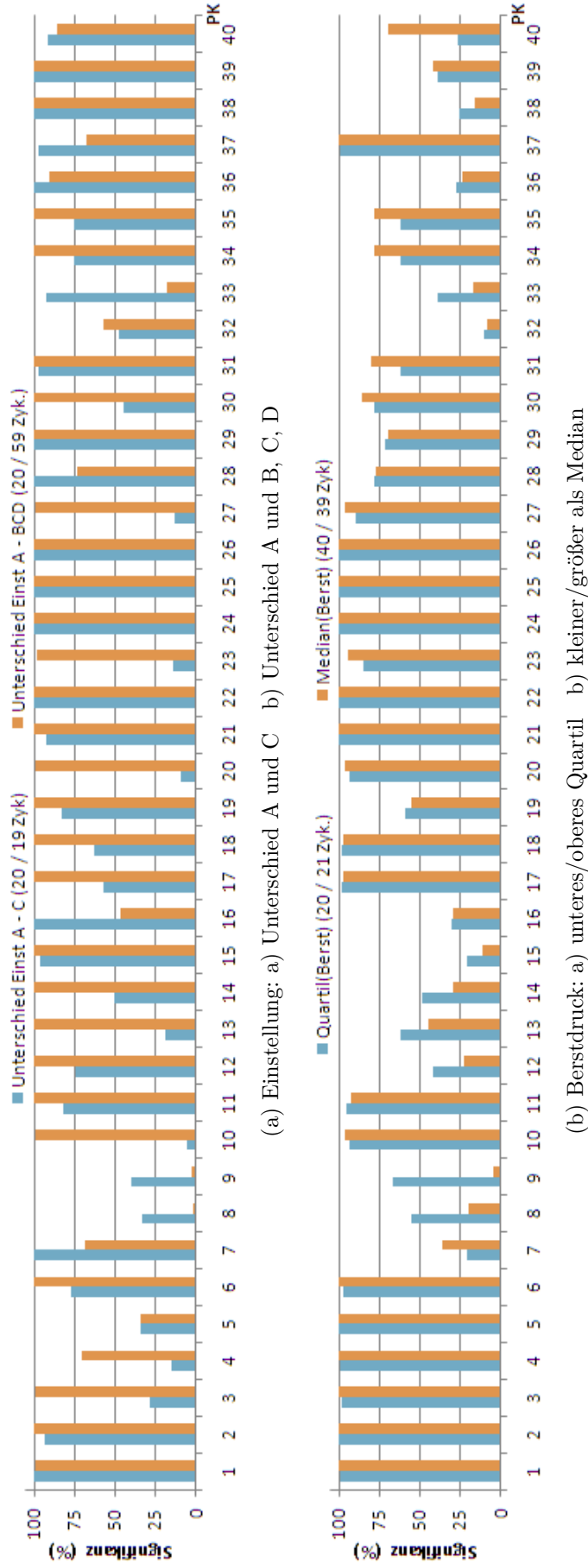


Abbildung 5.11: Signifikanzniveaus von Kenngrößen bezüglich Unterscheidungsfähigkeit verschiedener Zielgrößen

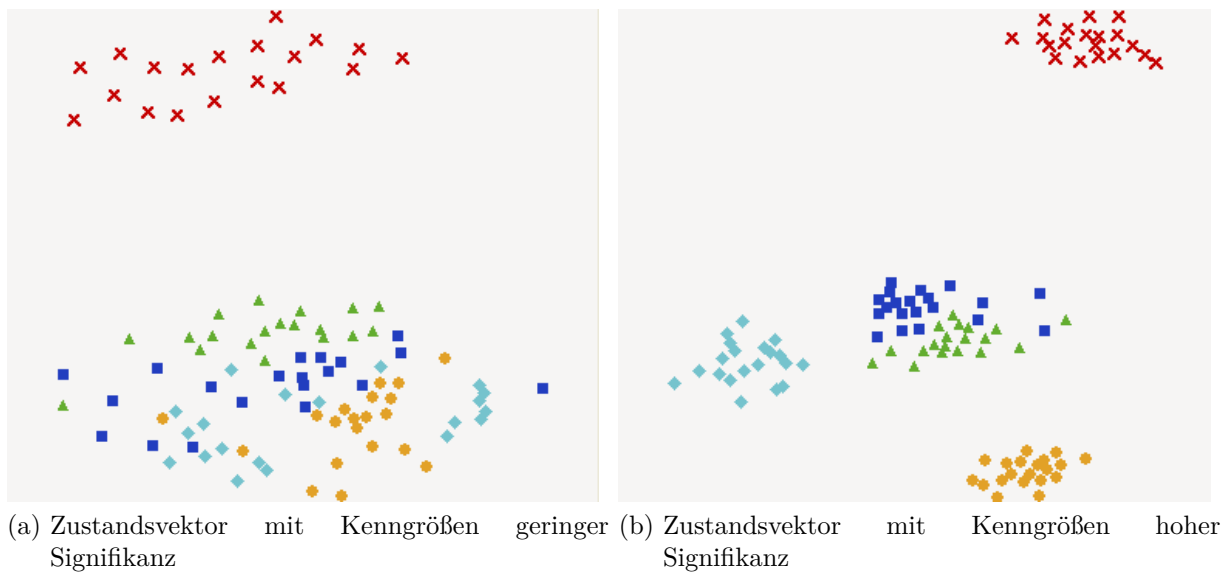


Abbildung 5.12: Auswirkung der Kenngrößenauswahl und Potenzial der Signifikanzanalyse

aus den Einstellungen B, C und D – zu unterscheiden. Die Einstellung E wurde für diese Untersuchungen komplett ausgeblendet. Aufgrund der stark variierten Schweißkraft hat sie nämlich deutlich geänderte Prozessbedingungen zur Folge und ist für fast alle Kenngrößen signifikant verschieden. Im Vergleich der Einstellungen A und E haben über 85 % aller Kenngrößen eine Signifikanz von 100 %, der Mittelwert liegt bei 97,96 %. Bei der Signifikanzanalyse zur Unterscheidung der Einstellung E zu den restlichen 79 Zyklen weisen über 93 % der Kenngrößen eine Signifikanz von 100 % und im Mittel eine Signifikanz von 99,2 % auf. Eine Gruppenbildung entsprechend der verschiedenen Einstellungen kann auch dazu genutzt werden, die Effekte auf Qualitätswerte zu beurteilen und damit den Prozess zu optimieren. Dies entspricht direkt dem Vorgehen, wie es auch im Rahmen von ANOVA zur Auswertung und Interpretation von Versuchsplänen durchgeführt wird (vgl. MONTGOMERY [Mon09]).

Für die Signifikanzanalyse in Abbildung 5.11(b) ist mit dem Berstdruck ein Qualitätsmerkmal gruppenbestimmend. Die erste Gruppe wird durch Zyklen gebildet, die einen geringeren Berstdruck als eine definierte Schwelle l_1 aufweisen; der zweiten Gruppe sind die Zyklen zugewiesen, deren Probekörper einen Berstdruck größer einer Schwelle l_2 haben. Im einen Fall ist $l_1 = l_2$ auf den Median der Berstdrücke gesetzt, d. h. jeder der Zyklen ist einer der beiden Gruppen zugewiesen, denen letztendlich gleich viele Zyklen zugeordnet sind⁵. Im zweiten Fall sind die Grenzen l_1 und l_2 identisch mit dem oberen und unteren Quartil. Folglich ist der einen Gruppe das Viertel der Zyklen zugewiesen, die den höchsten Berstdruck lieferten. Die 25 % der Versuchsdaten mit den geringsten Berstdrücken gehören zur anderen Gruppe. Die Signifikanzanalyse mit einem Zielmerkmal zur Gruppenzuweisung

⁵eine ungerade Anzahl an Gesamtzyklen führt zu einem Unterschied von einem Zyklus

wird dazu verwendet, diejenigen Kenngrößen zu identifizieren, die einen maximalen Beitrag zur Prognose des Zielmerkmals liefern können. Ein Zustandsvektor, der einen deutlichen Anteil an Kenngrößen beinhaltet, die für unterschiedliche Niveaus der Zielgröße signifikant andere Werte aufweisen, begünstigt eine Clusterbildung entsprechend der Zielgröße. Dies wiederum ermöglicht den Aufbau von robusten Modellen, welche die Zielgröße hinreichend genau abbilden können. Im Gegensatz dazu führen deutlich unterschiedliche Werte der Zielgröße wegen der damit verbundenen statistisch abgesicherten unterschiedlichen Werte der Kenngrößen zu großen Distanzen zwischen den zugeordneten Zustandsvektoren. Diese Abstände fließen maßgeblich sowohl in die Projektion der Daten durch das Sammons-Mapping als auch in die Prognosegüte der GRNFN-Modelle ein (entsprechend der Abschnitte 4.1.2 und 4.5). Die Auswirkung der Zusammensetzung des Zustandsvektors in Bezug auf Kenngrößen, die vorwiegend eine hohe oder niedrige Signifikanz zur Zielgröße aufweisen, ist in Abbildung 5.12 dargestellt. Im einen Fall ist der beschreibende Zustandsvektor aus 30 Kenngrößen zusammengesetzt, die mehrheitlich keine große Signifikanz bezüglich des Unterschieds der Zielgröße – in diesem Fall der Einstellung – aufweisen. Im anderen Fall beinhaltet der 30-dimensionale Zustandsvektor vorwiegend Größen mit hoher Signifikanz. Der Zusammenhang zwischen den ermittelten Signifikanzniveaus der Kenngrößen sowie den gruppenbestimmenden Werten der Zielgröße ist deutlich zu erkennen.

Die Signifikanzanalyse zur Auswahl von Kenngrößen kann bei der Modellbildung durch die erzielte bessere Trennung bezüglich der Zielgröße zu deutlich verbesserten Modellen beitragen. Dabei bietet dieses Vorgehen mehrere Vorteile:

- Eine Normierung oder (lineare) Umskalierung der einzelnen Kenngrößen (siehe Abschnitt 5.3.1) wirkt sich nicht auf den Wert des Signifikanzniveaus aus. Bei der Normierung verschiebt sich neben dem Mittelwert auch die Streuung entsprechend, so dass die Signifikanzanalyse für originale und normierte Kenngrößen identische Wahrscheinlichkeiten liefert.
- Der Einfluss der Streuung wird durch die statistischen Verfahren berücksichtigt und gefundene Zusammenhänge sind statistisch abgesichert.
- Im Vergleich zur Korrelationsanalyse wird der Einfluss der Streuung deutlich reduziert, was zu einer deutlich erhöhten Trennschärfe führt. In Tabelle 5.1 sind neben den Korrelationskoeffizienten auch die Signifikanzniveaus und Wahrscheinlichkeiten angegeben. Diese geben an, wie deutlich sich die Mittelwerte der beiden Einstellungen mit je 15 Zyklen für die einzelnen Größen unterscheiden. Für alle Messreihen – insbesondere auch B^* inklusive des enthaltenen Messfehlers – werden sehr hohe Signifikanzlevel erreicht. Selbst im dritten Fall, in dem die Korrelation keine Aussagen mehr zulässt, liegen die Wahrscheinlichkeiten für einen Unterschied zwischen 98,32 % und 100 %.

- Einzelne Ausreißer, die in Messfehlern begründet sein können, haben kaum Einfluss, und das Verfahren liefert entsprechend robuste Ergebnisse.
- Die Analyse kann auch für Prozesse angewendet werden, für die entweder kein Versuchsplan oder ein solcher nur mit sehr großem Aufwand durchgeführt werden kann. Liegen allein aus dem Arbeitspunkt größere Datenmengen vor, können die beeinhalteten Schwankungen mit entsprechend ermittelten Qualitätswerten verknüpft werden. Hierbei ist entscheidend, dass aufgrund der statistischen Grundlage mit zunehmendem Datenumfang auch kleinere Varianzen sicher bewertet werden können. Eine hinreichend genaue und abgesicherte Ermittlung der Zielgröße ist dabei jedoch zwingende Voraussetzung, da ansonsten die Schwankung der Prüfmittel dominiert und zu Scheinzusammenhängen führen kann.

Die Auswahl von Kenngrößen mittels der Signifikanzanalyse kann für dasselbe Merkmal allein aufgrund geänderter Grenzen des Zielmerkmals bei der Gruppenbildung merklich variieren (siehe Abbildung 5.11(b)). Wenn die Gruppen aus je einer Einstellung bestehen, werden einerseits Kenngrößen zur Unterscheidung identifiziert; andererseits sind jedoch die anderen Einstellungen nicht berücksichtigt, da ein Mittelwertvergleich nur für zwei Gruppen funktioniert. Insgesamt ist es somit nötig, die Kenngrößenselektion auf der Basis mehrerer Signifikanzanalysen mit unterschiedlichen Gruppenparametern zu kombinieren. Zur Berücksichtigung aller Einstellungen können alle Einstellungen paarweise, jede Einstellung separat gegen die verbliebenen (siehe Abbildung 5.11(a)) oder beides betrachtet werden. Bei fünf Einstellungen ergeben sich so 10 Signifikanzen pro Merkmal bei 1:1 oder 5 Signifikanzen für 1:Rest. Insgesamt führt dies bei K Signifikanzanalysen zu K Signifikanzniveaus pro Kenngröße. Neben unterschiedlichen Mittelwerten kann auch die Größe der Varianz bewertet und als Auswahlkriterium aufgenommen werden. Eine kleinere Varianz bedeutet dabei eine geringere Streuung und ein robusteres Verhalten.

Die endgültige Selektion von Kenngrößen erfolgt wiederum über ein Votingverfahren – ähnlich wie bei der Unbekannt-Klassifikation (siehe Abschnitt 4.6). Zur Auswahl von Kenngrößen ist anschließend ein Mindestwert für die Signifikanz vorzugeben – alle Kenngrößen, die eine höhere Signifikanz aufweisen, werden selektiert. Entweder kann dies pro Kenngröße anhand des Mittelwerts oder anderer Größen über alle K Signifikanzwerte erfolgen. Oder der Mindestwert wird auf jede der K Auswertungen angewandt, wobei er jeweils einzeln festgelegt werden kann. Auch ist es möglich, das Niveau so zu wählen, dass mindestens M Kenngrößen selektiert sind. Für die abschließende Auswahl können dann z. B. nur die Kenngrößen verwendet werden, die in mindestens 50% der Signifikanzanalysen über dem festgelegten Signifikanzniveau liegen. Das genaue Auswahlkriterium ist flexibel und könnte auch entsprechend der Gewichtung der Zielgrößen oder anderer Randbedingungen gewählt werden.

Insgesamt haben viele Auswertungen ergeben, dass sich eine bessere Modellgüte erzielen lässt, wenn auch einige weniger signifikante Kenngrößen in den Zustandsvektor aufgenommen werden. Wenn ausschließlich Kenngrößen mit Signifikanzen von $\geq 99,9\%$ in allen Einzelanalysen verwendet werden, wird das Modell zu stark auf genau diesen einen Zusammenhang getrimmt. Es treten dann ähnliche Effekte auf wie beim Auswendiglernen bei Neuronalen Netzen, welches mit einem Verlust an Generalisierungsfähigkeit einhergeht. Außerdem fehlen einem solchen Modell auch Kenngrößen, die Änderungen im Zustand anzeigen, welche nicht direkt mit den Größen der verwendeten Signifikanzanalyse zusammenhängen. Damit geht auch sukzessive die Fähigkeit verloren, bisher unbekannte Zustände zu erkennen. Für die Beispielversuchsreihe aus Abschnitt 3.8 ist die erzielbare Trennung bezüglich der Einstellungen für verschiedene Signifikanzniveaus und Voting-Anteile in Abbildung 5.13 dargestellt. Unter dem Begriff „Voting-Anteile“ wird dabei der prozentuale Anteil verstanden, der von den einzelnen Wahrscheinlichkeitswerten jeder Kenngröße überschritten werden muss, damit diese in den Zustandsvektor übernommen wird. Für die dargestellten Daten wurde jeweils die Diskriminierungsfähigkeit der Kenngrößen von einer Einstellung gegenüber der Restmenge untersucht. Bei den vier verwendeten Einstellungen A, B, C und D⁶ ergibt das insgesamt vier Signifikanzwerte pro Kenngröße. Bei einem Voting-Anteil von 75% muss eine Kenngröße somit mindestens drei Signifikanzwerte größer als den Schwellwert aufweisen, um selektiert zu werden. Die Signifikanzanalyse zur Kenngrößenselektion wurde wiederum nur anhand der Trainingsdaten – also $\frac{2}{3}$ der Gesamtdaten – durchgeführt, für die anschließende Beurteilung der Selektion wurden alle Datenzyklen verwendet. Anhand Abbildung 5.13 ist zu erkennen, dass hohe Voting-Anteile in Verbindung mit hohen Signifikanzniveaus zu schlechteren Separierungsergebnissen führen. Die Bedeutung der mit J_5 bezeichneten Werte wird im folgenden Abschnitt 5.3.4 beschrieben. Zudem ist die Anzahl der Kenngrößen aufgeführt, die jeweils die angegebenen Signifikanzniveaus übertreffen. Es ist einleuchtend, dass mit steigender Signifikanz- oder Voting-Schwelle die Anzahl der verbleibenden Kenngrößen abnimmt.

5.3.4 Kenngrößenselektion durch Genetische Algorithmen

Eine andere Möglichkeit, die Kenngrößen für einen optimierten Zustandsvektor auszuwählen, ist der Einsatz von Genetischen Algorithmen (GA). GA sind ein heuristisches Optimierungsverfahren und gehören zur Klasse der Evolutionären Algorithmen. Diese Verfahren adaptieren sehr erfolgreich die Strategien der (biologischen) Evolution, wonach sich über Generationen hinweg die Individuen durchsetzen, deren durch ihre Gene bestimmten Eigenschaften die besten Anpassungen an die Umgebung ermöglichen. Dies wird oftmals auch als „survival of the fittest“ bezeichnet. GA ermöglichen eine heuristische und parallele Su-

⁶Einstellung E wurde aus den oben genannten Gründen nicht für die Signifikanzanalyse verwendet, wohl jedoch für die Verifikation

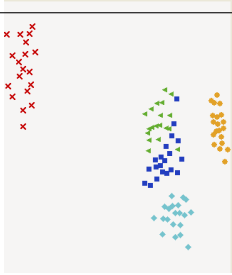
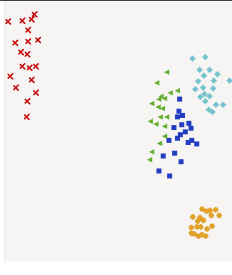
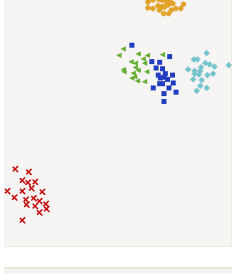
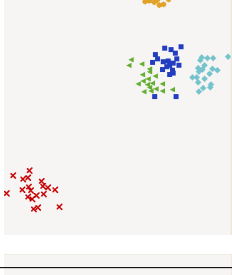
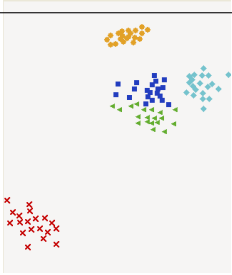
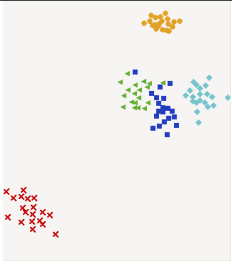
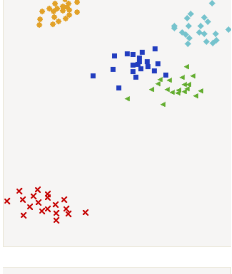
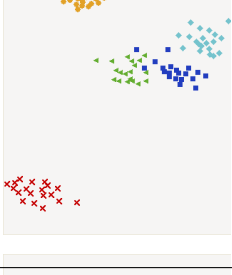
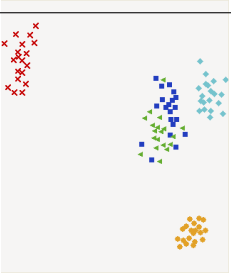
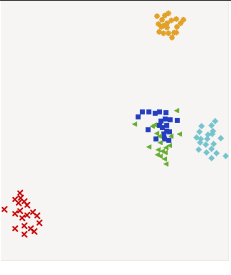
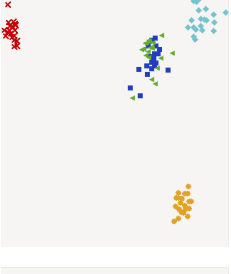
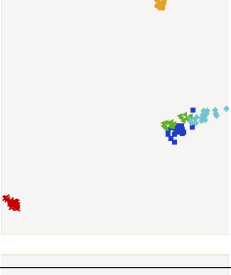
Voting 50 %	Signifikanzniveau $P = 90 \%$		Signifikanzniveau $P = 95 \%$		Signifikanzniveau $P = 99,5 \%$		Signifikanzniveau $P = 99,9 \%$	
	Anz. PK: 120 Overlap $J_5 = 0,9818$	Anz. PK: 115 Overlap $J_5 = 0,9829$	Anz. PK: 101 Overlap $J_5 = 0,9819$	Anz. PK: 93 Overlap $J_5 = 0,9804$	Anz. PK: 84 Overlap $J_5 = 0,9740$	Anz. PK: 53 Overlap $J_5 = 0,9676$	Anz. PK: 47 Overlap $J_5 = 0,9689$	
								
Voting 75 %	Anz. PK: 94 Overlap $J_5 = 0,9820$	Anz. PK: 52 Overlap $J_5 = 0,9109$	Anz. PK: 30 Overlap $J_5 = 0,9327$	Anz. PK: 30 Overlap $J_5 = 0,9093$				
Voting 100 %								

Abbildung 5.13: Ergebnisse der Signifikanzanalyse für verschiedene Niveaus und Voting-Anteile

che auch in sehr großen Lösungsräumen, die kombinatorisch keinesfalls zu erfassen wären. Evolutionäre Algorithmen bieten sich insbesondere für Probleme an, für die keine speziellen Lösungsansätze existieren. Nachteilig bei GA ist der hohe Rechenaufwand sowie die Tatsache, dass es trotz der Gewinnung sehr guter Ergebnisse keine Garantie für das Auffinden des Optimums gibt.

Ausgangspunkt bei Genetischen Algorithmen ist eine Population aus L Individuen. Jedes der Individuen verfügt über ein Genom aus M einzelnen Genen, welche die Eigenschaften des Individuums bestimmen. Die Anzahl und Art der Gene sind entsprechend der zu optimierenden Aufgabe festzulegen. Anhand seiner Gene kann für jedes Individuum ein Fitnesswert berechnet werden, der ein Maß dafür ist, wie gut ein Individuum die gestellte Aufgabe löst. Ziel der durch GA simulierten Evolution ist die Maximierung des Fitnesswertes⁷. Zu Anfang wird jedes Individuum der Population initialisiert, zumeist mit zufälligen Werten für jedes Gen. Jedes Individuum dient als Ausgangspunkt einer Suche im Lösungsraum, wodurch die oben genannte parallele Suche erreicht wird. Anschließend durchläuft die Population eine größere Anzahl an Evolutionsschritten. Dabei kommen insbesondere drei Operationen zum Einsatz, die Kernbestandteil eines jeden Evolutionären Algorithmus sind und jeweils zur nächsten Generation führen:

Selektion: Bei ihr werden aus der aktuellen Generation zufällig Individuen ausgewählt, die entweder direkt oder durch Rekombination in die nächste Generation übernommen werden. Individuen mit einer höheren Fitness, die durch eine als Fitness-Funktion bezeichnete Zielfunktion anhand des Genoms ermittelt wird, werden dabei bevorzugt gewählt. Dies begünstigt die Weitergabe von Genen, die zu hoher Fitness führen und damit eine relativ gute Lösung des gestellten Problems bilden, verhindert aber nicht die Selektion schlechter Gene.

Rekombination: Sie bezeichnet die Vermischung des Erbguts zweier als Eltern bezeichneter Individuen auf die aus ihnen generierten Kind-Individuen. Die Bildung der Kind-Genome kann mit verschiedenen Crossover-Operatoren erfolgen. Häufig ist der 1- (oder n -)Punkt-Crossover als kombinierender Operator ausgeführt, wobei die Eltern-Genome jeweils an einer (oder n) zufällig gewählten Position(en) aufgetrennt und die Teilabschnitte der Eltern bei den Kindern vertauscht werden. Es existieren auch interpolierende Operatoren wie der arithmetische Crossover, der durch eine gewichtete Verknüpfung die Gene der Eltern miteinander verrechnet. Die Rekombination ist bei GA der Hauptoperator.

⁷Es kann auch Ziel der Optimierung sein, den Fitnesswert F – beispielsweise durch einen Fehlerwert repräsentiert – zu minimieren. Durch Verwendung des reziproken Werts $F^* = 1/F$ kann die Optimierung wieder in ein Maximierungsproblem überführt werden. Im Weiteren wird deshalb nur noch die Maximierung als Ziel betrachtet.

Mutation: Sie verändert einzelne Gene zufällig, wobei die Mutationsrate meist auf sehr geringe Werte gesetzt wird. Die Mutation erfolgt zumeist als letzte Operation und dient zur Ausdehnung der Suche auf die relative Nachbarschaft des Individuums sowie der Aufrechterhaltung der Diversität. Bei GA spielt die Mutation eine untergeordnete Rolle.

Grundlegende Arbeiten und die Verwendung von GA für Such- und Optimierungsprobleme gehen auf HOLLAND zurück [Hol92]. Er führt auch die kanonische Form des Genoms als Binärstring ein, der nur aus Nullen und Einsen besteht. Im Folgenden werden nur noch die Punkte näher besprochen, die für die hier gestellte Aufgabe der Kenngrößen Selektion relevant sind. Tiefergehende Details zu Genetischen Algorithmen sind der weiterführenden Literatur zu entnehmen, beispielsweise GOLDBERG oder WEICKER [Gol89, Wei07].

Die Umsetzung der GA im Rahmen dieser Arbeit erfolgte unter Verwendung der C++-Bibliothek *GAlib* (Version 2.4.7, März 2007) als Framework. *GAlib* wurde von WALL im Rahmen seiner Promotion am Massachusetts Institute of Technology (MIT) entwickelt und ist dort verfügbar und zur Verwendung freigegeben [Wal07].

Festlegung des Genoms

Hauptaufgabe bei der Optimierung der Zustandsvektoren ist die Auswahl geeigneter Kenngrößen. Die Codierung dieser Aufgabenstellung kann sehr einfach mit einem binärcodierten Genom erfolgen. Jedes Gen kann dabei nur die Werte 1 oder 0 annehmen – für die Beschreibung „Kenngröße selektiert“ oder „Kenngröße nicht selektiert“. Die Genomlänge ist durch die Anzahl der insgesamt verfügbaren Kenngrößen festgelegt, die Anzahl der selektierten Kenngrößen eines Individuums entspricht der Anzahl der Gene mit Wert 1. Eine Erweiterung der reinen Selektion ist zusätzlich die unterschiedliche Gewichtung der einzelnen Kenngrößen. Dies entspricht der in Abschnitt 5.3.1 angesprochenen Skalierung. Die Skalierung kann mittels binärer Genome nicht dargestellt werden, so dass die Gene durch Gleitkomma-Zahlen repräsentiert werden. Der Wert eines Genoms entspricht direkt dem Skalierungsfaktor der zugeordneten Kenngröße. Um die absolute Dominanz einzelner Kenngrößen durch sehr hohe Skalierungsfaktoren auszuschließen, ist es sinnvoll, den zulässigen Wertebereich für die Gene zu begrenzen, z. B. auf das Intervall $[0, 01; 10]$. Bei der Skalierung können zwei Versionen unterschieden werden: Die reine Skalierung, bei der für jede Kenngröße ein Faktor definiert wird, oder eine Kombination aus Skalierung und Selektion. Die Codierung erfolgt dann derart, dass positive Werte eines Gens den Skalierungsfaktor wiedergeben und ein negativer Faktor für eine deselektierte Kenngröße steht. Dieses Vorgehen hat gegenüber der Alternative, deselektierte Kenngrößen durch den Wert Null zu repräsentieren, den Vorteil, bereits gefundene Skalierungsfaktoren, die sich im Laufe der Generationen bewährt haben, bei der Deselektion nicht zu verlieren. Eine Umschaltung des Selektionszustands erfolgt einfach durch Multiplikation des Gens mit -1 .

Aus Gründen des Speicherbedarfs der Modelle, der Laufzeit bei der Anwendung oder Reduzierung der Modellgröße und der einhergehenden Reduktion von festzulegenden Modellparametern kann es erforderlich sein, Randbedingungen wie eine maximale (oder auch genaue) Anzahl an zu selektierenden Kenngrößen zu definieren. Dies führt zu Bedingungen, welchen den Individuen bei der Optimierung genügen, und die im Rahmen der Rekombinations- und Mutationsoperatoren des GA berücksichtigt werden müssen. Hierfür gibt es prinzipiell zwei Ansätze. Zum einen können Individuen, die Randbedingungen verletzt haben, durch Modifikation des Fitnesswerts – indem beispielsweise ein Faktor < 1 als Strafterm eingeführt wird – bestraft werden, so dass sie aufgrund der geringeren Fitness mit deutlich kleinerer Wahrscheinlichkeit im Rahmen der Selektion in die nächste Generation übernommen werden. Dieses Vorgehen hat den Nachteil, dass sehr eng gesetzte Randbedingungen, wie beispielsweise die Auswahl von exakt 13 Kenngrößen, zu vielen ungültigen Genomen und damit vielen schlechten Fitnesswerten führt, was das Auffinden einer guten Lösung erschwert. Die andere Möglichkeit besteht in der Reparatur von Genomen, welche die gestellten Bedingungen verletzen. Sind zu viele Kenngrößen selektiert, können zufällig einzelne deselektiert werden, bis die gestellten Bedingungen wieder erfüllt werden.

Fitnessfunktion

Ein Kernelement bei GA ist die Festlegung einer geeigneten Fitnessfunktion. Da die Fitnessfunktion pro Generation für jedes Individuum einmal ausgewertet werden muss, stellt sie das laufzeitkritische Element beim Einsatz von GA dar. Für die Kenngrößenselektion gibt es verschiedene Möglichkeiten, zu aussagekräftigen Fitnesswerten zu kommen.

1. Da die Kenngrößenselektion zur Optimierung der Prozessmodellierung dient, ist es nahe liegend, mit der durch das Genom eines Individuums definierten Kenngrößenselektion oder Skalierung ein Modell zu erstellen. Für dieses Modell kann anschließend ein Gütemaß ermittelt (siehe Abschnitt 4.5) und als Fitnesswert für die zugrunde liegende Kenngrößenauswahl verwendet werden. Nachteilig bei dieser Variante ist, dass der Trainingsaufwand für ein Modell relativ hoch ist. Bei der Modellbildung mit Neuronalen Netzen auf MLP-Basis kommt noch hinzu, dass ihr Training durch die Initialisierung mit Zufallszahlen nicht deterministisch ist und folglich die Modellgüte stark schwanken kann. Dem kann durch mehrmaliges Training und eine anschließende Mittelwertbildung begegnet werden, was jedoch den Aufwand zur Ermittlung des Fitnesswerts entsprechend vervielfacht.

Bei GRNFN-Modellen spielt der letztgenannte Punkt aufgrund spezifischer Eigenschaften der Modellstruktur, die in Abschnitt 4.4 näher erläutert sind, keine so tragende Rolle. Das Netztraining erfordert kein iteratives und dadurch zeitaufwändiges Lernverfahren. Zudem ist es deterministisch bezüglich der verwendeten Trainingsdaten, womit ein mehrfaches Training mit anschließender Mittelung unnötig wird.

In Verbindung mit der in Abschnitt 4.4.2 erläuterten Fehlertoleranz, die eine Netzprognose trotz fehlender Eingangsdaten ermöglicht, kann der Aufwand noch weiter reduziert werden. Es ist mithin nur ein einmaliges Training mit allen Kenngrößen während der Initialisierungsphase des Genetischen Algorithmus notwendig. Für jede beliebige Kenngrößenselektion kann anschließend die Modellgüte ohne Neutraining ermittelt werden, indem bei der Berechnung der Modellprognose die deselektierten Kenngrößen für alle Datenzyklen als ungültig markiert werden.

- Wie schon in Kapitel 4 und in den vorigen Abschnitten deutlich wurde, hängt die Clusterbildung der Datenvektoren eng mit der erreichbaren Modellgüte zusammen. Da die Ausbildung und Verteilung der Cluster einzig von den Kenngrößen abhängt, kann eine Maßzahl, welche die Verteilung und Durchdringung der Cluster beschreibt, direkt als Fitnesswert für die GA verwendet werden. In [Kön95] schlägt KÖNIG ein nicht-parametrisches Überlappungsmaß vor, das die Durchdringung verschiedener Cluster zu quantifizieren sucht. Die Grundidee beruht darauf, die Klassenzugehörigkeit jedes Datenvektors mit denen seiner k nächsten Nachbarn zu vergleichen und die Übereinstimmung zu bewerten. Die Berechnung erfolgt nach folgender Berechnungsvorschrift:

$$\begin{aligned}
 J_k &= \frac{1}{N} \sum_{i=1}^N q_i & q_i &= \frac{\sum_{j=1}^k q_{ij} + \sum_{j=1}^k n_{ij}}{2 \sum_{j=1}^k n_{ij}} & (5.6) \\
 n_{ij} &= 1 - \frac{d_{ij}}{d_{ik}} & q_{ij} &= \begin{cases} n_{ij} & , w_i = w_j \\ -n_{ij} & , w_i \neq w_j \end{cases}
 \end{aligned}$$

- mit
- N: Anzahl der Datenpunkte
 - k: Anzahl der nächsten Nachbarn, die berücksichtigt werden
 - d_{ij} : Abstand zw. Datenpunkt x_i und seinem j-ten nächsten Nachbarn x_{ij}
 - d_{ik} : Abstand zw. Datenpunkt x_i und dem entferntesten der k Nachbarn x_{ik}
 - w_i : Klasse des Datenpunktes x_i
 - w_j : Klasse des j-ten Nachbarn x_{ij} .

Für jeden Datenvektor wird ein lokaler Überlappungsgrad durch Vergleich der zugeordneten Klasse mit den nächstgelegenen Datenvektoren berechnet. Dabei fließen über den Wert n_{ij} Nachbarn mit geringerem Abstand stärker ein als weiter entfernte. Durch q_{ij} und die Berechnung für q_i liefern übereinstimmende Klassen einen Beitrag, während unterschiedliche Klassen sich aufheben. Das Überlappungsmaß ergibt sich abschließend als arithmetischer Mittelwert aus den Überlappungsgraden aller Datenvektoren. Durch die Normierung in q_i ergibt sich für das Überlappungsmaß ein Wertebereich von $J_k \in [0; 1]$. Je größer der Wert von J_k ist, desto geringer ist die Überlappung zwischen den Clustern. Das Abstandsmaß wird auch bei PANNEKAMP [Pan05] erfolgreich zur Identifikation von Merkmalen zur

Trennung verschiedener Texturklassen eingesetzt. Aufgrund der Betrachtung der Klassenzugehörigkeit ist das Überlappungsmaß nur für diskrete Merkmale geeignet, deren Werte sich dezidierten Klassen zuordnen lassen. Für kontinuierliche Merkmale ist es notwendig, diese zuvor anhand mindestens eines Schwellwerts in Klassenmerkmale zu überführen.

Zuvor wurde schon auf die in Abbildung 5.13 angegebenen und mit „Overlap J_5 ” bezeichneten Werte hingewiesen. Dabei handelt es sich um das von KÖNIG definierte Überlappungsmaß in der vorgestellten Form für $k=5$ nächste Nachbarn. Das Überlappungsmaß für die in Abbildung 5.12(a) dargestellten Zustandsvektoren ergibt sich für fünf nächste Nachbarn zu $J_5 = 0,7943$ und für $k = 10$ zu $J_{10} = 0,7637$, für Abbildung 5.12(b) ergeben sich die Werte $J_5 = 0,9688$ und $J_{10} = 0,9643$. Der Zusammenhang zwischen höheren Werten von J_k und der besseren Trennschärfe der Cluster ist klar zu erkennen, so dass vor dem Hintergrund der vielfach nachgewiesenen Optimierungserfolge von GA und den in den vorigen Abschnitten hergeleiteten Zusammenhängen auf weitere analytische Untersuchungen verzichtet wird.

Neben der Selektion der Kenngrößen ist es auch möglich, GA dazu zu verwenden, die Parameter der Aktivierungsfunktionen individuell zu bestimmen. Untersuchungen hierzu haben jedoch ergeben, dass die Modellgüte bei der Verwendung von Standardparametern, welche die Anzahl der Trainingsdaten berücksichtigen, hinreichend groß ist. Ein Aufwand der individuellen Parameterermittlung mittels GA ist aufgrund des sehr hochdimensionalen Lösungsraums $d = N * L * m$ mit m als Anzahl der Parameter der Aktivierungsfunktion nicht zielführend und birgt zudem die Gefahr der Überanpassung ausschließlich an die bekannten Datenvektoren. Wenn die Parameter nur global – d. h. gleiche Werte für alle Aktivierungsfunktionen – optimiert werden sollen, ist eine kombinatorische Suche mit entsprechender Schrittweite, wie sie auch für die Abbildungen in Abschnitt 4.3 verwendet wurden, die deutlich schnellere Lösung.

5.4 Fazit

Aufbauend auf den durchgeführten Versuchsreihen mit dem Probekörper wurden unterschiedliche Möglichkeiten der Signalverarbeitung vorgestellt. Als Ausgangspunkt dienen verschiedene Visualisierungsarten, die es ermöglichen, relevante Einflüsse und Streuungen erkennen und Zusammenhänge ableiten zu können. Aufbauend darauf werden verschiedene Kenngrößen abgeleitet mit dem Ziel, die wesentlichen Informationen in komprimierter Form abzubilden. Die Kenngrößen werden zum Zustandsvektor zusammengeführt, der für die Prozessmodelle als Eingangssignal dient.

Für die Modellierung mit GRNFN-Modellen sind die Abstände zwischen Zustandsvektoren im \mathbb{R}^N -Raum in mehrerlei Hinblick wesentlich, so dass die Skalierung der einzelnen Kenngrößen einen wesentlichen Einfluss hat. Die komponentenbezogene Normierung

der Zustandsvektoren anhand einer gegebenen Trainingsmenge ist eine daraus abgeleitete, zwingende Standardoperation bei der Modellbildung. Zusätzlich kann eine Skalierung einzelner Kenngrößen erfolgen, um einzelne Größen stärker zu gewichten oder vorhandenes Prozesswissen einfließen zu lassen. Ist die Modellgüte, die mit den normierten Zustandsvektoren erreicht werden kann, nicht ausreichend – insbesondere aufgrund einer in Relation zur Nachbarschaft inhomogenen Verteilung der Zielklassen – muss der Zustandsvektor, wie in Abschnitt 5.3 gezeigt, optimiert werden. Bei der Optimierung werden die hinsichtlich einer Zielgröße weniger relevanten Kenngrößen identifiziert und ausgeblendet. Ihre Varianzen, die hinsichtlich der zielgrößenbezogenen Clusterbildung wie Rauschfaktoren wirken, beeinflussen somit die Abstandsberechnung nicht mehr, so dass eine bessere Trennung bezüglich der Zielgrößen und somit bessere GRNFN-Modellgüten erreicht werden. Eine Möglichkeit zur Auswahl von Kenngrößen stellt die Signifikanzanalyse dar, die auf statistischen Mittelwertvergleichen aufsetzt. Abhängig vom Wert der Zielgröße werden die Trainingsdaten in zwei Gruppen aufgeteilt und anschließend die Signifikanz jeder Kenngröße auf die Verschiedenheit der Mittelwerte bezüglich der beiden Gruppen untersucht. Eine andere Möglichkeit zur Optimierung des Zustandsvektors besteht im Einsatz von Suchstrategien wie GA. Zur notwendigen Beschreibung der Güte einer Kenngrößenauswahl können verschiedene Kriterien verwendet werden. Da die Trenngüte bezüglich eines Qualitätsmerkmals wesentlich für leistungsfähige Modelle ist, kann diese mittels eines so genannten Überlappungsmaßes quantifiziert werden, in das für jeden Trainingszyklus der Vergleich des Zielwerts mit seinen Nachbarn einfließt. Ein weiteres Gütemaß ist die mit der aktuellen Kenngrößenauswahl erzielbare Modellgüte für unabhängige, d. h. dem Modell unbekannte Testzyklen. Suchstrategien sind im Gegensatz zur Signifikanzanalyse beim Finden eines optimierten Zustandsvektors nicht auf den Vergleich von zwei Klassen beschränkt und neben der Kenngrößenauswahl ist auch eine gleichzeitige Skalierung der Kenngrößen möglich. Dafür ist der Rechenaufwand bei Suchstrategien wie GA sehr viel höher als bei der Signifikanzanalyse und trotz sehr guter Erfahrungen in vielen Bereichen kann das Auffinden der optimalen Lösung bei ersteren nicht garantiert werden.

6 Anwendung der realisierten Systemtechnik beim Ultraschallschweißen

In den folgenden Abschnitten sind die Ergebnisse unter Anwendung der erarbeiteten und vorgestellten Methoden und Systemtechnik beim Ultraschallschweißen dargestellt. Im Wesentlichen erfolgten die Arbeiten mit der in Abschnitt 3.1.2 beschriebenen NEPRES-Mini-Plattform, in welche die in den Kapiteln 4 und 5 softwareseitig realisierten Methoden integriert wurden. Entsprechend der ursprünglichen Zielsetzung (siehe Abschnitt 1.3 und insbesondere Abbildung 1.5) wird einerseits die Leistungsfähigkeit zur Prognose der erzielten Werkstückqualität beschrieben. Andererseits wird die Detektion von Veränderungen im Schweißprozess, die auf Einflüsse der Vorprozesse zurückzuführen sind, untersucht. Die analytischen Arbeiten erfolgten vorwiegend anhand des Probekörpers (siehe Abschnitt 3.3), da dort die gesamte Prozesskette überwacht und beeinflusst werden kann. Darüber hinaus werden Ergebnisse von Versuchsreihen an Praxisbauteilen vorgestellt.

6.1 Ergebnisse zum Körperschall

Bei den Versuchsreihen mit dem Probekörper sowie einigen Versuchen an Praxisbauteilen wurde der Körperschall beim Schweißvorgang als weitere Signalquelle erfasst. Die Erfassung des Körperschalls ist mit Randbedingungen verknüpft, welche die Anwendung beeinflussen. Im Gegensatz zu den Maschinensignalen Weg, Leistung, Amplitude und (aufgebrachte) Schweißkraft muss der Körperschall am Werkzeug erfasst werden. Der Körperschallsensor muss aufgrund der hohen Beschleunigungen und zur reproduzierbaren Positionierung mittels einer Stiftschraube fixiert werden. Dies verhindert einen schnellen Werkzeug- und Produktwechsel, da die Sensorik inklusive Messkabel beim Umbau in der Regel entfernt werden muss.

Der Körperschallsensor erfasst die beim Schweißvorgang resultierenden Schwingungen der Werkzeugaufnahme. Diese Schwingungen werden durch die Sonotrode eingebracht, aber bei der weiteren Fortpflanzung durch das Produkt, welches sich durch die Schweißung zudem verändert, und durch die Reflektionen und Überlagerungen in der Werkstückaufnahme un-

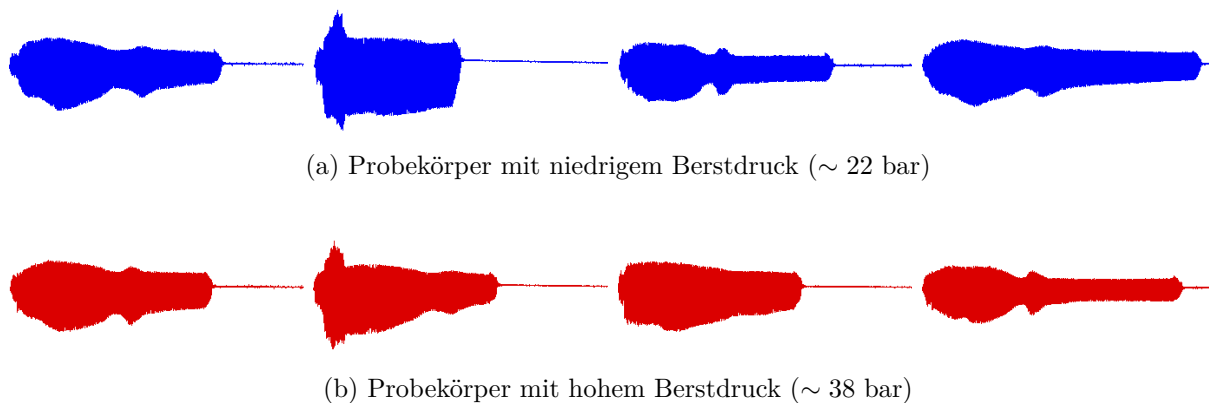


Abbildung 6.1: Körperschallsignale ausgewählter Zyklen beim ERG-Probekörper aus PS

vorhersehbar moduliert. Da die Werkzeuge meist an mehreren Punkten mit Schrauben auf dem Maschinentisch der Schweißmaschine fixiert sind, kann sich ein verändertes Anzugsmoment direkt auf die Schwingungsmoden der Werkstückaufnahme inklusive Grundplatte und somit das Körperschallsignal auswirken. Ebenso kann sich das erfasste Körperschallsignal bereits bei einem leichten Versatz der Sensorposition signifikant verändern. Die beiden letztgenannten Einflüsse wurden in begleitenden Versuchen festgestellt. Ferner bedingen die hohen Frequenzen des Körperschallsignals und die damit anfallenden großen Datenmengen eine leistungsfähige Hardware zur Signalerfassung. Dies begründet unter anderem auch die Weiterentwicklung der PC-basierten Hardwareplattform des NEPRES-Systems (siehe Abschnitt 2.6) zum neuen, auf einem Laptop mit USB-Messkarte basierenden NEPRES-Mini-System (siehe Abschnitt 3.1.2). Auch die Verarbeitung von hochfrequenten Schwingungssignalen ist deutlich aufwändiger, da die Zeitsignale nicht direkt verwendet werden können. Entweder muss zuerst die Einhüllende ermittelt oder eine (Fast-)Fourier-Transformation durchgeführt werden, bevor Kenngrößen abgeleitet werden können (siehe Abschnitt 5.2.2).

Insgesamt ist festzustellen, dass die Körperschallsignale auch bei nicht verändertem Prozess sehr stark streuen. In Abbildung 6.1 sind für acht Zyklen die Körperschallsignale dargestellt, wobei jedes der dargestellten Signale aus etwa 18.000 Messwerten besteht. Geschweißt wurden Probekörper mit Energierichtungsgeber aus Polystyrol mit Schweißzeiten von ca. 200 – 400 ms. Das durch das Abschalten des Ultraschalls gekennzeichnete Ende der Schweißphase ist klar zu erkennen. Während die Probekörper in Abbildung 6.1(a) sehr niedrige Berstdrücke im Bereich von 22 bar ergaben, hatten die Abbildung 6.1(b) zugeordneten Probekörper durchweg hohe Berstdrücke von ca. 38 bar. In beiden Gruppen der Signale sind einerseits die sehr starken Schwankungen der Signalform zu erkennen, andererseits kommen in beiden Gruppen ähnliche Formen vor. Es besteht somit keinerlei Zusammenhang zwischen Signalform und erzielbarer Schweißnahtgüte. Auch für Kenngrößen, die sowohl aus der Einhüllenden als auch dem Amplitudenspektrum der Signale gewonnen wurden, konnte

kein eindeutiger Zusammenhang zur Schweißnahtfestigkeit gefunden werden. Eine andere Analyse des Zusammenhangs zwischen Körperschallsignal und visuell beurteilter Nahtgüte ergab ein ähnliches Ergebnis. Es ist keine Systematik zwischen Schwankungen des Amplitudenspektrums und dem resultierenden Nahtbild zu erkennen. Diese Analyse wurde jedoch nicht am Probekörper durchgeführt, sondern im Rahmen von Versuchen an Praxisteilen.

Der Grund zur Untersuchung des Körperschallsignals war die Erkenntnis aus der Praxis, dass erfahrene Einrichter die beim Schweißen entstehenden Geräusche teils zur Unterstützung beim Einrichten des Prozesses verwenden. Für eine Versuchsreihe bei einem Industriepartner – gefertigt wurde ein Werkstück, welches in der medizinischen Laboranalyse eingesetzt wird – hat ein erfahrener Prozessverantwortlicher für jeden Schweißzyklus eine (subjektive) Beschreibung und Bewertung der Schweißgeräusche sowie der optischen Nahtgüte vorgenommen. Die Ergebnisse wurden gruppiert und in eine Rangfolge gebracht. Dem abgeleiteten Kriterium *Klanghomogenität* wurden anschließend ganzzahlige Werte im Bereich $[-2; 2]$ zugeordnet. Für die 250 Versuchszyklen, in denen neben der Schweißkraft auch die Amplitude leicht variiert wurde, wurden Zustandsvektoren gebildet, die ausschließlich aus dem Amplitudenspektrum des Körperschallsignals abgeleitet wurden. Die Verteilung dieser (normierten) Zustandsvektoren ist in Abbildung 6.2(a) dargestellt. Blaue Quadrate stehen für -2 und rote Kreuze für $+2$, weisen also den größten Abstand zueinander auf. Insgesamt ist relativ wenigen Zyklen eine der „Extremklassen“ zugeordnet. Diese lassen jedoch keinerlei Strukturierung oder Clusterbildung erkennen, sondern verteilen sich über den gesamten Darstellungsbereich. Negative und positive Extremwerte sind teils sehr nahe beieinander. Dies bedeutet wiederum, dass sich die zugrunde liegenden Amplitudenspektren sehr ähnlich sind. Bereits durch leichte Frequenzverschiebungen zueinander können Dissonanzen auftreten, die sich im Klang auswirken können. Geringe Verschiebungen, die sich während der Schweißung zusätzlich ändern, sind jedoch bei den kurzen Prozesszeiten durch die FFT nicht aufzulösen (siehe Abschnitt 5.2.2). Insgesamt ist ein Zusammenhang zwischen dem aufgezeichneten Körperschallsignal und dem akustisch wahrnehmbaren Schweißgeräusch also fraglich. Eine nahe liegende Begründung liegt darin, dass auch Schwingungen von Sonotrode und Werkstück den Luftschall beeinflussen und somit als „Gesamtschall“ wahrgenommen werden. Im Körperschall bilden sich diese jedoch nicht ab.

Bei dem untersuchten Prozess werden die zu verschweißenden Werkstücke durch Spritzgießen mit je einem Zwei-Kavitäten-Werkzeug für Ober- und Unterteile hergestellt, d. h. pro Fertigungszyklus entstehen zwei Produkte. Diese sind aufgrund von Fertigungstoleranzen nicht hundertprozentig identisch und bilden somit zwei „Varianten“. Da es sich bei dem Werkstück um ein sehr diffiziles Teil handelt, werden die Varianten mit leicht unterschiedlichen Parametern verschweißt. Die beiden Schweißprogramme unterscheiden sich in der Schweißamplitude um 2 %, in der Schweißkraft um etwa 1 %, sowie in der Triggerkraft um etwa 5 %. In Abbildung 6.2(b) sind die beiden Varianten als Rechteck und Dreieck

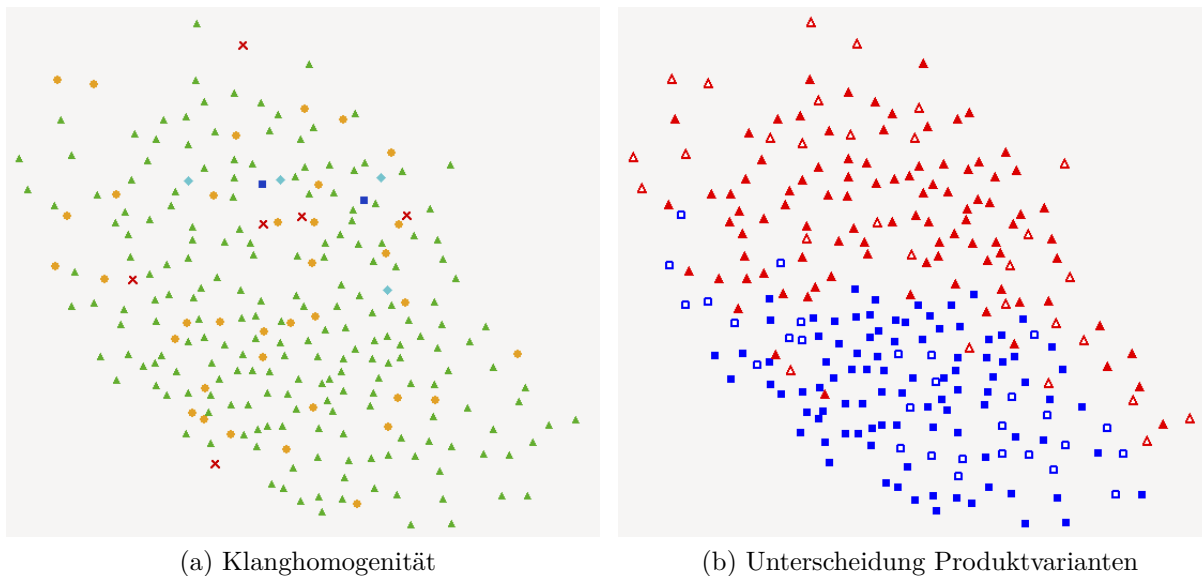


Abbildung 6.2: Zustandsvektor einzig aus normiertem Amplitudenspektrum des Körperschalls

dargestellt. Während der Versuchsreihe wurde für je 30 Zyklen die Schweißkraft für beide Varianten ausgehend vom Arbeitspunkt um fast 10 % erhöht – gekennzeichnet durch die nicht ausgefüllten Marker. Insgesamt ist zu erkennen, dass sich die Zustandsvektoren, deren Komponenten ausschließlich aus Kenngrößen des Amplitudenspektrums bestehen, für beide Varianten deutlich unterscheiden. Dies ist entweder den spritzgießbedingten, geringen Unterschieden zwischen den beiden Varianten oder den aufgrund der unterschiedlichen Triggerkraft geänderten Ankopplungsbedingungen geschuldet. Gegen die Triggerkraft als Ursache spricht jedoch die Tatsache, dass sich eine Variation der Schweißkraft für beide Varianten ohne erkennbare Systematik auf das Amplitudenspektrum auswirkt. Es erscheint wenig plausibel, dass sich die Triggerkraft deutlich auf das Schwingungsverhalten auswirken soll, die Schweißkraft aber nicht.

6.2 Ergebnisse mit Probekörpern

Der Probekörper hat durch die Verfügbarkeit aller Prozesse den Vorteil, dass die Versuchsreihen bei entsprechender Planung über die gesamte Prozesskette – vom Spritzgießen über die Lagerzeiten und -bedingungen bis zum Schweißen – genau kontrolliert und beeinflusst werden können. Des Weiteren können mehrere Prüfverfahren durch die ebenfalls verfügbaren Prüfeinrichtungen durchgeführt werden, um die Qualität reproduzierbar und genau beschreiben zu können. Auf der Basis verschiedener Versuchsreihen werden die Ergebnisse der Modellierung mit GRNFN-Modellen dargestellt und bei der Qualitätsprognose mit den Ergebnissen von MLP-Modellen verglichen. Da MLP-Modelle nicht für die Störungserkennung

in Vorprozessen geeignet sind (siehe Abschnitt 4.6), ist hierfür ein Vergleich mit Ergebnissen bei der GRNFN-Modellierung nicht möglich.

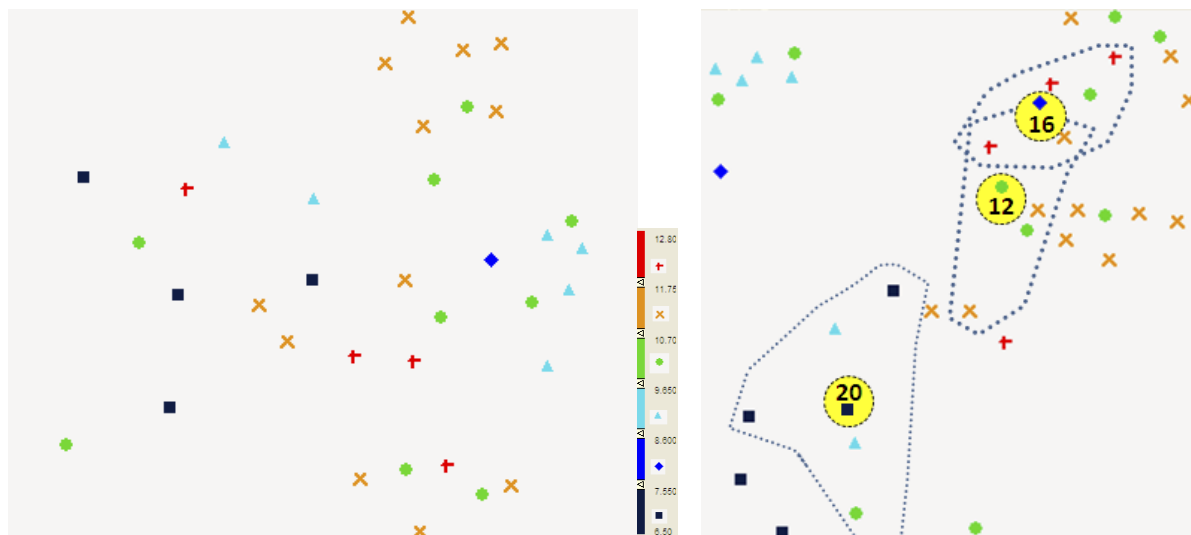
6.2.1 Überwachen der Prozesseinstellung

Wie beschrieben, wirken sich Veränderungen in der Prozesseinstellung der Schweißmaschine direkt auf den Schweißverlauf und damit auf viele abgeleitete Kenngrößen aus. Dadurch verändert sich auch der beschreibende Zustandsvektor signifikant. Die auftretenden Prozessschwankungen haben den geänderten Zustandsvektor als Schwerpunkt, so dass sich eine Änderung der Einstellung als eigener Cluster im Zustandsraum und auch bei der Projektion durch das Sammons-Mapping bildet. Dies wurde bereits im Zuge der Modellbildung und Signalverarbeitung (Kapitel 4 und 5) mehrfach dargestellt – z. B. in den Abbildungen 4.13, 5.6 und 5.12. Die angeführten Untersuchungen erfolgten jeweils mit dem Probekörper aus dem Werkstoff ABS mit ERG-Nahtgeometrie; die zugrunde liegende Versuchsreihe inklusive der variierten Parameter ist in Abschnitt 3.8 erläutert. Auch bei dem Probekörper aus ABS mit Quetschnaht-Geometrie, der in Abschnitt 6.2.2 zur Darstellung der Ergebnisse der Qualitätsprognose verwendet wird, ist die Auswirkung einer Parametervariation auf den Zustandsvektor und eine entsprechende Clusterbildung anhand Abbildung 6.3(c) zu erkennen.

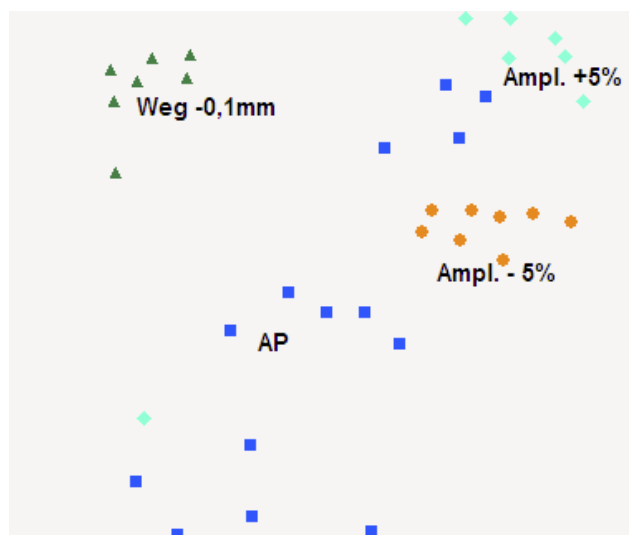
Durch entsprechende GRNFN-Modelle ist es möglich, die Schweißzyklen den einzelnen Einstellungen zuzuordnen oder neue, unbekannte Einstellungen zu erkennen, die durch eine (versehentliche oder gewollte) Änderung von Maschinenparametern entstehen können. Dies kann natürlich prinzipiell durch die Maschinensteuerung abgefangen und durch Zugangsbeschränkungen verhindert werden. Interessanter ist die Fähigkeit der Modelle, Maschinen- oder Prozessstörungen zu registrieren, die von der Steuerung unter Umständen nicht erkannt werden. So wirkt sich ein Booster mit geändertem Verstärkungsfaktor auf die notwendige Leistungsaufnahme aus oder ein zunehmend verstopfter Druckluftfilter in der Vorschubeinheit führt durch die damit verbundene Druckminderung zu anderen Schweißkraftverhältnissen. Beides wirkt sich auf den Prozessverlauf und den berechneten Zustandsvektor eines Schweißzyklus aus und kann ab einer gewissen Größe durch die Modelle erkannt werden.

6.2.2 Prognose von Qualitätsmerkmalen durch Prozessmodelle

Für die Darstellung der Fähigkeiten eines Prozessmodells zur Prognose von Qualitätsmerkmalen werden Probekörper mit Quetschnahtgeometrie aus ABS verwendet. Die zugrunde liegende Versuchsreihe besteht aus 47 Zyklen, die alle für die Bestimmung der Normierungsgrenzen herangezogen werden. Neben dem Schweißen im Arbeitspunkt sind auch die Schweißparameter Amplitude und Abschaltweg variiert worden, als zu prognostizierendes Qualitätsmerkmal wird der Berstdruck der geschweißten Teile verwendet. Die Aufteilung



(a) Verteilung der normierten Zustandsvektoren mit allen Kenngrößen (b) Verteilung nach Optimierung der Zustandsvektoren (Kreise: vgl. Abb. 6.4)



(c) Zuordnung der Prozessvariation zur Verteilung aus Abb. 6.3(b)

Abbildung 6.3: Probekörper mit Quetschnahtgeometrie: Lerndaten für Festigkeitsprognose

der vorhandenen Zyklen in Lern- und Verifikationsmuster für die Modellgenerierung erfolgt im Verhältnis 3:1; d. h. jeder vierte Zyklus wird nur als Testzyklus verwendet, so dass sich 36 Lern- und 11 Testzyklen ergeben. Die weiteren Operationen zur Modellbildung – Kenngrößenselektion und Training des Modells – werden einzig anhand der 36 Lerndaten durchgeführt. Die Verteilung der normierten Zustandsvektoren, wie sie aus den Schweißsignalen abgeleitet werden, sind in Abbildung 6.3(a) zu sehen. Der Berstdruck als Qualitätsmerkmal liegt für alle Zyklen im Bereich von 6,5 - 12,8 bar. Auch wenn das Merkmal einen kontinuierlichen Wertebereich umfasst, wurde der Berstdruck in Abbildung 6.3 aus Darstellungsgründen in 6 gleichverteilte Qualitätsklassen mit je 1,05 bar Breite aufgeteilt. Deren

Wertebereiche sowie die Zuordnung zu den Markern der Sammons-Mapping-Darstellung sind der Farbskala in Abbildung 6.3(a) zu entnehmen.

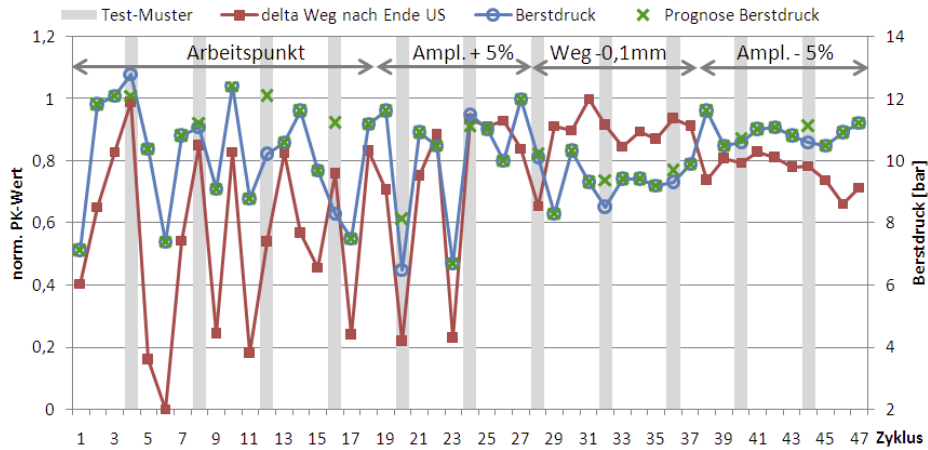
Die Optimierung der Zustandsvektoren erfolgte mittels einer Signifikanzanalyse (siehe Abschnitt 5.3). Für die Zustandsvektoren wurden nur diejenigen Kenngrößen verwendet, die eine Signifikanz von über 95 % gegenüber den Quartilen des Berstdrucks über alle Lernzyklen hinweg aufwiesen. Aus den 36 Lernzyklen wurden also zuerst die 9 Zyklen (ein Viertel) mit dem jeweils höchsten oder niedrigsten Berstdruck ausgewählt. Für jede Kenngröße wurde anschließend die Signifikanz dafür berechnet, dass ihre Werte für die Zyklen mit niedriger Festigkeit einen unterschiedlichen Mittelwert gegenüber denen mit hoher Festigkeit aufweisen. Nur Kenngrößen, deren Signifikanzniveau über 95 % lag, wurden in den optimierten Zustandsvektor übernommen, so dass sich insgesamt eine Reduktion von 133 auf 16 Kenngrößen ergab. Die Verteilung der so erhaltenen Vektoren ist in Abbildung 6.3(b) dargestellt. Zusätzlich sind noch die drei Testzyklen 12, 16 und 20 inklusive der ihnen zugeordneten fünf nächsten Nachbarn dargestellt, auf die im Abschnitt 6.2.2 noch näher eingegangen wird. Die optimierten Zustandsvektoren weisen eine deutlich bessere Verteilung als die Vektoren mit allen Kenngrößen auf. Insbesondere liegen Zyklen mit höherer Festigkeit in Abbildung 6.3(b) deutlich kompakter zusammen und sind nicht wie in Abbildung 6.3(a) über den gesamten Zustandsraum verteilt.

Qualitätsprognose mit GRNFN

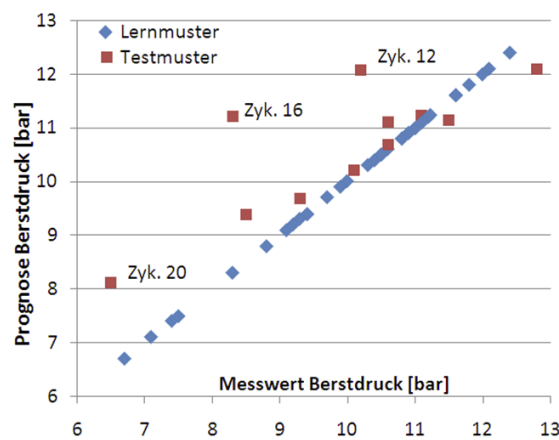
Ausgehend von den optimierten Zustandsvektoren in Abbildung 6.3(b) erfolgt der Aufbau eines GRNFN-Modells zur Prognose des Berstdrucks. Wie in Kapitel 4 mehrfach ausgeführt, bedingen ähnliche Prozessabläufe ähnliche Zustandsvektoren, die wiederum zu geringen Abständen in der Projektion mittels Sammons-Mapping führen. Aus Abbildung 6.3(a) ist anhand der Verteilung der Zyklen ersichtlich, dass benachbarte Zyklen und somit ähnliche Prozessabläufe teils zu sehr unterschiedlichen Festigkeiten (zu erkennen am Markertyp) führen. Für die optimierten Zustandsvektoren aus Abbildung 6.3(b) ist eine Zuordnung der Zyklen zur Prozesseinstellung des Schweißprozesses durch einen direkten Vergleich mit Abbildung 6.3(c) möglich. Die Marker repräsentieren dabei die unterschiedlichen Einstellungen. Bei diesem Vergleich ist auch zu erkennen, dass insbesondere der Arbeitspunkt – zumindest hinsichtlich des Berstdrucks – relativ instabil ist. Von den 14 Zyklen im Arbeitspunkt (Abbildungen 6.3(b) und 6.3(c)) weisen 5 Zyklen eine der drei niedrigsten und 6 eine der beiden höchsten Berstdruckklassen auf. Die drei variierten Einstellungen (2 x Amplitude, 1 x Weg) mit je 7 bzw. 8 Zyklen sind deutlich homogener. Einzig die Einstellung mit 5 % erhöhter Amplitude weist einen Festigkeits”ausreißer” in der niedrigsten Berstdruckklasse auf – in den Abbildungen 6.3(b) und 6.3(c) links unten. Dessen Zustandsvektor unterscheidet sich jedoch deutlich von den anderen Zyklen.

Eine Analyse des Zusammenhangs der einzelnen Prozesskenngrößen, die aus den Schweißsignalen abgeleitet werden, mit dem Berstdruck lieferte einige Kenngrößen, die insbesondere im Arbeitspunkt eine hohe Korrelation aufweisen. Die Streuung des Berstdrucks ist somit kein Problem der Berstdruckprüfung, sondern hat ihre Ursachen tatsächlich im Schweißprozess. Insbesondere der Weg, den die Sonotrode während der Haltezeit aufgrund von Schmelzeaustrieb und abkühlungsbedingter Schwindung zurücklegt, korreliert stark mit den großen Unterschieden im Berstdruck. Der Verlauf dieser Wegkenngröße, des gemessenen Berstdrucks sowie der Prognosewert des Berstdrucks mittels eines GRNFN-Modells sind in Abbildung 6.4(a) zu sehen. Das GRNFN-Modell wurde anhand der Daten aus Abbildung 6.3(b) erstellt. Dargestellt sind in Abbildung 6.4(a) die genannten Größen für alle 47 Zyklen – d. h. die Lerndaten sowie die dem GRNFN unbekannt, also unabhängigen Testmuster. Eine direkte Gegenüberstellung von Messwert und Prognose ist in Abbildung 6.4(b) zu sehen. Eine stets absolut richtige Prognose (Prognosegüte von 100 %) würde in Abbildung 6.4(b) dazu führen, dass alle Wertepaare auf der ersten Winkelhalbierenden abgebildet würden; alle Prognosefehler sind dagegen durch entsprechende Abweichungen zu erkennen. Die Lernmuster weisen auch bei starken Berstdruckschwankungen eine 100 %ige Übereinstimmung auf. Anhand der Testmuster ist zu erkennen, dass auch für diese der Berstdruck relativ genau prognostiziert werden kann.

Durch größere Abweichungen der Prognose vom gemessenen Berstdruck fallen insbesondere die Testzyklen 12, 16 und 20 auf, weswegen diese in Abbildung 6.3(b) mit ihrer gemessenen Festigkeit zusätzlich dargestellt sind. Zudem sind jeweils die fünf Trainingsmuster – und damit Stützvektoren des GRNFN-Modells – hervorgehoben, die den geringsten Abstand zu den Testzyklen aufweisen und somit vom GRNFN-Modell zur Ableitung der Prognose herangezogen werden. Anhand der Festigkeitswerte dieser Nachbarzyklen ermittelt das GRNFN-Modell auch seine Prognose, so dass das Zustandekommen der Prognose nachvollziehbar ist. Dies steht im Gegensatz zu MLP-Modellen, die bezüglich ihrer Wissensrepräsentation als „Black-Box“ funktionieren und bei (Fehl)Prognosen keinerlei Verknüpfung mit den Trainingsdaten ermöglichen. In Abbildung 6.4(a) können die Änderungen des gemessenen und prognostizierten Berstdrucks sowie die Wegkenngröße, welche stark mit dem Berstdruck korreliert, für die Zyklen 12, 16 und 20 abgelesen werden. Für die Zyklen 12 und 20 ist jeweils die richtige Tendenz der Prognose zu erkennen. Die Veränderung des Berstdrucks sowie der Prognose im Vergleich mit den jeweils vorigen Zyklen weist in dieselbe Richtung wie die der Wegkenngröße, lediglich das Ausmaß der Änderung ist unterschiedlich. Anders sieht es hingegen bei Zyklus 16 aus: Im Vergleich zu Zyklus 15 weist dieser eine deutlich geringere Festigkeit auf, obwohl die relevante Wegkenngröße viel größer ist. Diese Wegkenngröße beeinflusst maßgeblich den Prognosewert des Berstdrucks, der folglich deutlich über dem Messwert liegt. Aufgrund der beschriebenen Zusammenhänge ist



(a) Verlauf von Wegkenngroße, Berstdruck und Prognose für Lern- und Testmuster



(b) Berstdruck-Diagramm: Messwert und Prognose

Abbildung 6.4: Probekörper mit Quetschnahtgeometrie: Qualitätsprognose für Berstdruck mittels GRNFN-Modell

es insbesondere für den Zyklus 16 fraglich, ob nicht ein Messfehler, eine Signalstörung oder eine sonstige Vorschädigung für den geschweißten Probekörper vorliegt.

Qualitätsprognose mit MLP

Auch die NEPRES-Systemtechnologie, auf welcher diese Arbeit aufsetzt, ermöglicht die Prognose von Qualitätswerten mittels Neuronalen Netzen. Als Modellstruktur werden MLP mit Feed-Forward-Architektur eingesetzt (siehe Abschnitt 2.4.2). Das Training erfolgt mit einem eigens entwickelten Netzgenerator, der in verschiedenen Punkten bezüglich der Bedürfnisse der Modellierung von Fertigungsprozessen optimiert ist. Dies betrifft beispielsweise die Auswahl der Lern- und Testmuster oder den automatischen Lernabbruch anhand des Testfehlers, um einem Verlust der Generalisierungsfähigkeiten des Modells entgegenzuwirken. Anhand der in Abschnitt 6.2.2 beschriebenen Daten wurden mehrere MLP-Modelle zur Prognose des Berstdrucks erstellt, wobei das Training jeweils über 500 Epochen ging. Das

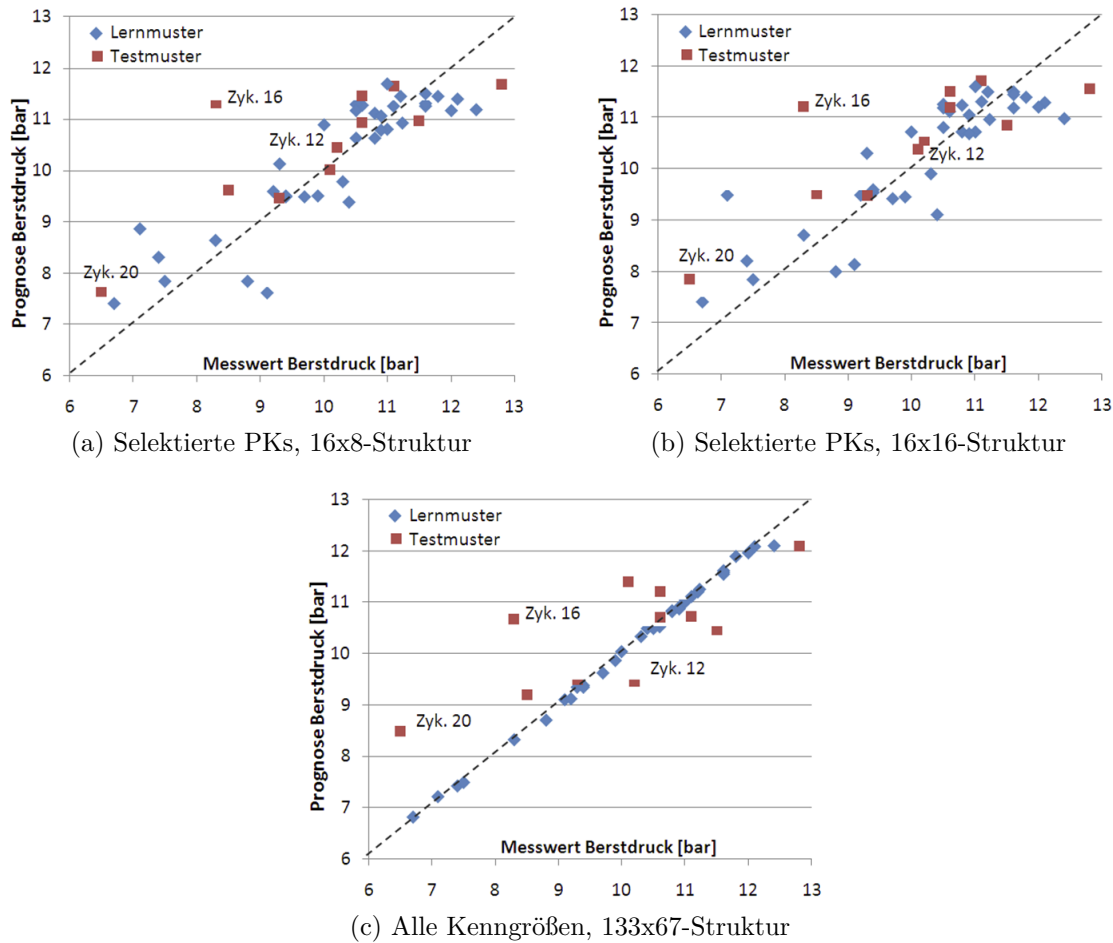


Abbildung 6.5: Qualitätsprognose für den Berstdruck mittels verschiedener MLP-Modelle

Training erfolgte mit denselben Lern- und Testdaten wie für das in Abschnitt 6.2.2 beschriebene GRNFN-Modell, um die Ergebnisse vergleichen zu können. Alle MLP-Modelle wurden mehrfach mit unterschiedlichen Initialisierungswerten der Gewichte trainiert. Sie führten jeweils zu vergleichbaren Resultaten. Dabei gab es auch einige Ausreißer nach unten, die durch das frühe Erreichen eines lokalen Minimums der Fehlerfunktion während des Trainings begründet sind.

In Abbildung 6.5 sind die Prognoseergebnisse für drei verschiedene MLP-Modelle dargestellt. Das erste Neuronale Netz entspricht der Struktur, wie sie bei NEPRES standardmäßig angelegt wird, und besteht aus 16 Eingangsneuronen – entsprechend der Dimension des optimierten Zustandsvektors –, acht Neuronen in der Hidden-Schicht sowie einem Ausgabeneuron. Insgesamt ergeben sich somit $N = 16 \cdot 8 + 8 \cdot 1 = 136$ Gewichtswerte sowie die neun Schwellwerte der internen Aktivierungsfunktionen der Hidden- und Ausgabeneuronen, die durch das Training festgelegt werden müssen. Die Ergebnisse der Prognose des Berstdrucks mit diesem MLP sind in Abbildung 6.5(a) dargestellt. Es ist zu erkennen, dass die Modellierung des Zusammenhangs zwischen Zustandsvektor und dem Berstdruck als Qualitätsgröße

	GRNFN	MLP (16x8)	MLP (16x16)	MLP (133x67)
Abbildung	6.4(b)	6.5(a)	6.5(b)	6.5(c)
Trainingsdauer	0,26 Sek.	0,86 Sek.	0,92 Sek.	2,9 Sek.
Anz. zu bestimmender Parameter	1-4	145	289	9046
Korrelation	0,92	0,84	0,83	0,92
mittlerer Prognosefehler [bar]	0,20	0,61	0,64	0,24
maximaler Prognosefehler [bar]	2,9	3,0	2,9	2,3
	ohne Zyklus 16 (wegen vermutetem Messfehler)			
Korrelation	0,96	0,88	0,87	0,95
mittlerer Prognosefehler [bar]	0,14	0,55	0,59	0,20
maximaler Prognosefehler [bar]	1,9	1,8	2,4	2,0

Tabelle 6.1: Ergebnisse der Qualitätsprognose für GRNFN- und MLP-Modelle

tendenziell klappt, das Modell jedoch Prognosefehler für Lern- und Testmuster aufweist. Als ein Problem wurde die relativ geringe Anzahl von $P = 136 + 9 = 145$ freien Netzparametern vermutet, so dass das zweite Modell mit 16 Neuronen in der Hidden-Schicht über $N = 16 \cdot 16 + 16 \cdot 1 = 272$ Gewichtswerte und 17 Schwellwerte als freie Parameter verfügt. Das Ergebnis ist in Abbildung 6.5(b) dargestellt. Auch ein weiteres Modell mit einer zweiten versteckten Schicht wies vergleichbare Ergebnisse auf. Abschließend wurde ein weiteres MLP-Modell angelegt, welches nicht mit dem optimierten, sondern mit dem ursprünglichen Zustandsvektor mit 133 Kenngrößen trainiert wurde. Das Modell verfügt über mehr als 9000 freie Netzparameter¹, die für die Modellierung der Zusammenhänge der 36 Lernmuster à 133 Kenngrößen zur Verfügung stehen. Aufgrund der im Verhältnis zu den verfügbaren Lerndaten hohen Anzahl an freien Parametern können diese die Lernmuster sehr gut abbilden, wie aus Abbildung 6.5(c) ersichtlich ist. Es stellt sich jedoch die Frage, wie es um die Generalisierungsfähigkeit eines solchen Netzes und damit die Verlässlichkeit der Prognose bei unbekanntem Daten bestellt ist.

Bei allen MLP-Modellen fällt auf, dass die Zyklen 16 und 20, wie schon beim GRNFN-Modell, nur unzureichend prognostiziert werden. Erfahrungen aus der praktischen Anwendung von NEPRES zeigen, dass für den Fall, dass ein Wert vom MLP-Modell trotz wiederholten Trainings auch mit geänderten Lern- und Testmustern nicht gut wiedergegeben werden kann, zumeist ein Messfehler vorliegt. Für Geometriemaße oder das Teilegewicht, die im Gegensatz zur Festigkeit wiederholt gemessen werden können, hat sich dies mehrfach bestätigt.

Gegenüberstellung der Prognoseergebnisse von GRNFN und MLP

Um die Leistungsfähigkeit von GRNFN-Modellen bei der Prognose von Qualitätswerten beurteilen zu können, wurde für den Probekörper eine Prognose des Berstdrucks erstellt.

¹ $N = 133 \cdot 67 + 67 \cdot 1 = 8978$ Gewichte sowie 68 Schwellwerte

Anschließend erfolgte ein Vergleich zu MLP-Modellen, wie sie in NEPRES seit Jahren erfolgreich für die Qualitätsprognose verwendet werden. Um einen direkten Vergleich der Ergebnisse zu ermöglichen, wurde das Training der GRNFN- und MLP-Modelle mit denselben Lern- und Testdaten durchgeführt. Aus den Ergebnissen wurden einige Kennwerte zur Beschreibung der Modellgüte für das GRNFN-Modell sowie die drei MLP-Varianten abgeleitet. Diese sind in Tabelle 6.1 aufgeführt. Da beim Zyklus 16 höchstwahrscheinlich ein Messfehler vorliegt (siehe vorstehend „*Qualitätsprognose mit GRNFN*“ und „... *mit MLP*“), sind die Güte-Kennwerte auch für die Versuchsdaten ohne Zyklus 16 aufgeführt, was in allen Fällen zu einer deutlichen Verbesserung der Werte führt. Insgesamt ist festzustellen, dass das GRNFN-Modell in fast allen Größen gleich gut oder sogar deutlich besser abschneidet als die MLP-Modelle.

Insgesamt bleibt festzustellen, dass die **Qualitätsprognose mit GRNFN-Modellen mindestens so gut** möglich ist wie mit MLP-Modellen. GRNFN bieten jedoch noch weitere Vorteile. Der wichtigste ist, dass GRNFN im Gegensatz zu MLP-Modellen eine **Erkennung von unbekanntem Zuständen** ermöglichen. Bei GRNFN-Modellen sind nur **wenige Parameter**² festzulegen und ein wiederholtes Training ergibt immer ein identisches Modell, d. h. das **Training ist deterministisch**. Beim MLP müssen zuerst die Netzstruktur festgelegt und anschließend durch das Training eine große Anzahl an Gewichten eingestellt werden. Da das resultierende MLP-Modell und damit die erzielbare Modellgüte von der Startinitialisierung abhängen, sollte das Training mehrfach wiederholt werden. Nur so kann ein schlechtes Modelltraining – bedingt durch ein frühes lokales Minimum der Fehlerfunktion – ausgeschlossen werden. Der dadurch resultierende Trainingsaufwand der MLP-Modelle führt zu einer deutlich **kürzeren Trainingsdauer** für GRNFN. Ergeben sich Prognosefehler, können bei GRNFN-Modellen die Ursachen nachvollzogen werden. In Kombination mit der Projektion durch das Sammons-Mapping machen **GRNFN-Modelle die Zusammenhänge transparent** und ermöglichen eine Interpretation. Dies wurde zuvor anhand der analytischen Betrachtung der Zyklen 12, 16 und 20 deutlich. MLP-Modelle hingegen sind eine Black-Box. Bei ihnen ist nicht nachvollziehbar, in welchen Gewichten eines MLP welche Information wie abgespeichert wird. Die Gründe für schlechte Prognosewerte können somit nicht näher eingegrenzt werden. Einen nicht zu vernachlässigenden Faktor für die Praxis stellt der deutlich geringere Versuchsaufwand zur Realisierung einer GRNFN-Adaption gegenüber der von MLP-Modellen dar. GRNFN-Modelle können im Wesentlichen für den Arbeitspunkt trainiert werden und eine Absicherung der Modellgültigkeit ist über die Unbekannt-Klassifikation gegeben. Für MLP-Modelle muss ein Versuchsplan, der um den Arbeitspunkt herum angesiedelt ist, durchgeführt werden. Nur so kann eine Gültigkeit des Modells gewährleistet werden, da für MLP eine Extrapolation über den Trainingsbe-

²Die freien Parameter entsprechen den Parametern der (global verwendeten) Aktivierungsfunktion. Geeignete Standardwerte können in Abhängigkeit vom Umfang der Trainingsdaten abgeleitet werden (siehe Abschnitt 4.3)

reich hinaus unzulässig ist. Für MLP-Modelle kann so schnell ein Vielfaches an Aufwand erforderlich werden, der insbesondere durch die notwendige Anzahl an Versuchsschritten am Zielprozess verursacht wird. Die Anzahl der benötigten Trainingsdaten hängt zudem stark von der abgeleiteten Dimension der optimierten Zustandsvektoren ab. Je höher die Eingangsdimension des Netzes und damit die Anzahl der freien Parameter ist, desto mehr Daten werden für ein zuverlässiges Training benötigt. Zu den weiteren Vorteilen, welche auch maßgeblich zur Auswahl der GRNFN-Modelle als Modellierungsverfahren beigetragen haben, sei auf Abschnitt 4.7 verwiesen.

Neben dem Berstdruck können andere Qualitätsmerkmale mittels GRNFN-Modellen analog überwacht werden. Für aussagekräftige Modelle ist jedoch wichtig, dass die Qualität der Werkstücke, die für das Training verwendet werden, möglichst exakt und reproduzierbar ermittelt werden kann und dass Schwankungen der Qualitätswerte mit Schwankungen im Prozess einhergehen. Für die Dichtigkeit als Qualitätsmerkmal waren – abhängig von der Prozesseinstellung – alle Teile entweder dicht oder undicht. Eine Überwachung ist damit indirekt über eine Kontrolle der Einstellung (siehe Abschnitt 6.2.1) möglich. Für die sehr wenigen Ausreißer, d. h. als undicht geprüfte Probekörper einer sonst dichten Einstellung, konnten als Ursachen ein nicht vollständig abschließendes Dichtelement bei der Ankopplung an das Dichtheitsprüfgerät, lokale Verschmutzungen auf der Nahtgeometrie beim Schweißen oder aber bereits durch das Spritzgießen induzierte Mikrorisse ermittelt werden. Letztere sind nicht oder nur schwer zu prognostizieren.

6.2.3 Variation des Spritzgießens

Um die Auswirkungen eines veränderten Spritzgießprozesses auf das Schweißverhalten untersuchen zu können, wurden zwei Varianten des Probekörpers mit variierten Spritzgießparametern hergestellt. Die erste Variante V_1 wurde im ermittelten Arbeitspunkt für den Probekörper aus PA-GF66 mit ERG-Geometrie gespritzt, bei der Variante V_2 wurde der Nachdruck (ND) deutlich reduziert. Auf den ersten Blick waren die beiden gespritzten Varianten optisch nicht zu unterscheiden. Nur in stark vergrößerten Aufnahmen war zu erkennen, dass der Energierichtungsgeber mit Höhe 0,3 mm bei V_2 – den ohne Nachdruck gefertigten Probekörpern – nicht immer komplett ausgeprägt war, sondern teilweise eine abgerundete Spitze aufwies. Zudem waren die Gewichtswerte leicht unterschiedlich.

Mit jeder der beiden Probekörpervarianten wurden anschließend Schweißversuche durchgeführt. Dabei wurden auch die Schweißparameter nach identischem Versuchsplan mit einer Veränderung von Schweißkraft und Amplitude variiert. Für jede SchweißEinstellung und jede Spritzgießvariante V_1 und V_2 wurden jeweils 20 Probekörper geschweißt. Ausgehend von der Versuchsreihe mit den Probekörpern V_1 wurden die Normierungsgrenzen aller 133 Kenngrößen für die Modellierung bestimmt; eine weitere spezifische Optimierung

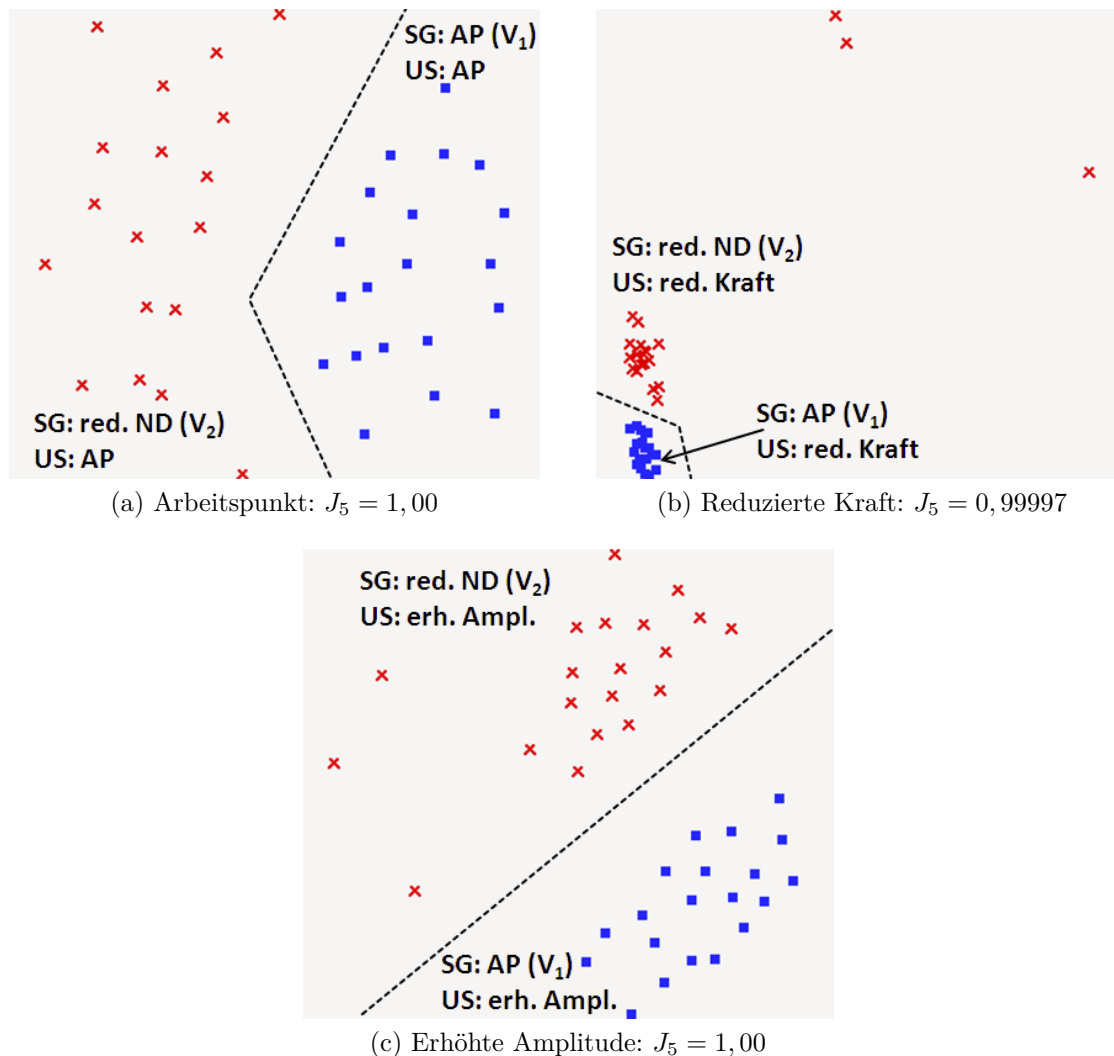


Abbildung 6.6: Auswirkung geänderter Spritzgießparameter (Probekörper PA-GF66, ERG)

des Zustandsvektors ist nicht erfolgt, d. h. es wurden alle 133 Kenngrößen verwendet. Anschließend wurden die Zustandsvektoren für beide Versuchsreihen normiert – jeweils mit den identischen Normierungsgrenzen, die anhand der Reihe V_1 ermittelt wurden. In den drei Projektionen in Abbildung 6.6 sind die so gewonnenen Zustandsvektoren für jeweils eine Einstellung des Schweißprozesses dargestellt. Neben dem ermittelten Arbeitspunkt der Schweißmaschine (Abbildung 6.6(a)) wurde die Schweißkraft reduziert (dargestellt in Abbildung 6.6(b)) und die Amplitude um 5% erhöht (Abbildung 6.6(c)). In allen drei Fällen kommen im Sammons-Mapping die Zustandsvektoren der geschweißten Probekörper V_2 – dargestellt durch rote Kreuze – deutlich getrennt von den Probekörpern V_1 zu liegen – gefertigt im Arbeitspunkt des Spritzgießprozesses. Die sich ergebenden Überlappungsmaße J_5 (siehe Formel (5.6)) sind ebenfalls angegeben. Neben der geänderten Lage im Zustandsraum ist auch die Streuung für den Probekörper V_2 deutlich höher. Dies entspricht einem weniger robusten Schweißprozess, was insbesondere auf den nicht optimalen Energierichtungsgeber

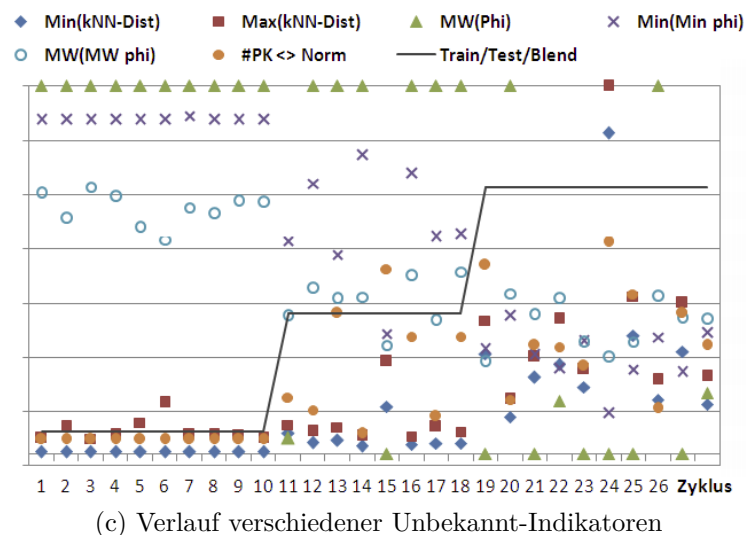
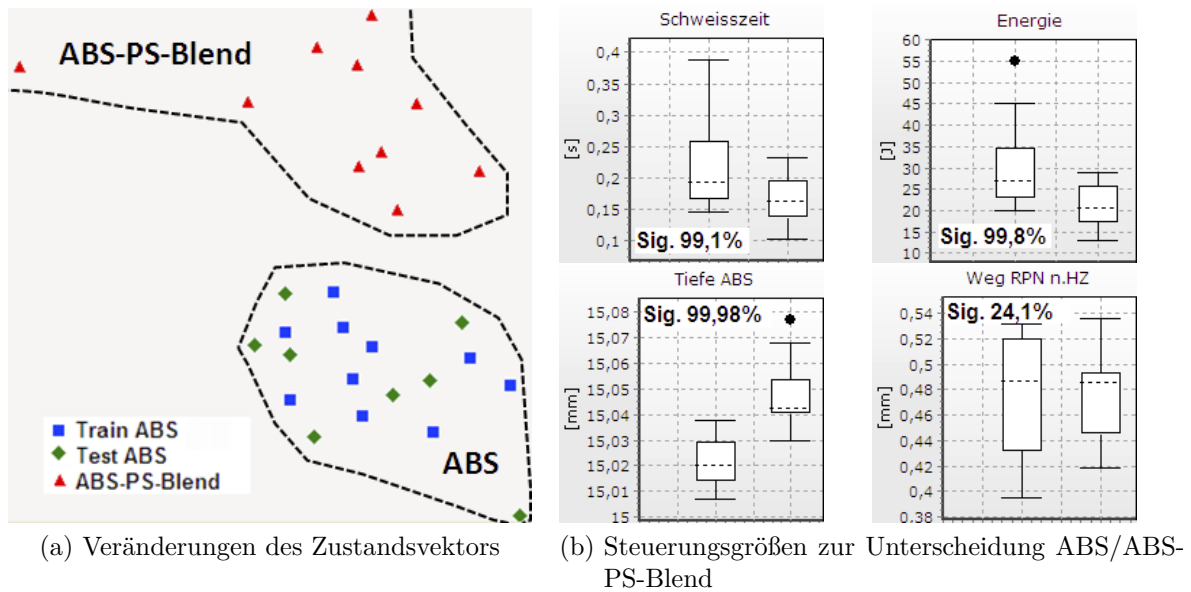


Abbildung 6.7: Schweißen von Probekörpern aus ABS und ABS-PS-Blend

zurückzuführen sein dürfte. Des Weiteren weisen die Probekörper V_2 gegenüber den korrekt gefertigten eine teils um 40 % verringerte Schweißnahtfestigkeit auf – ermittelt durch eine Zugprüfung. Ein veränderter Spritzgießprozess für die Probekörper wirkt sich folglich auf den Schweißprozess aus – im Prozessablauf und in der Qualität – und eine Unterscheidung ist einzig anhand von Größen aus dem Schweißprozess möglich.

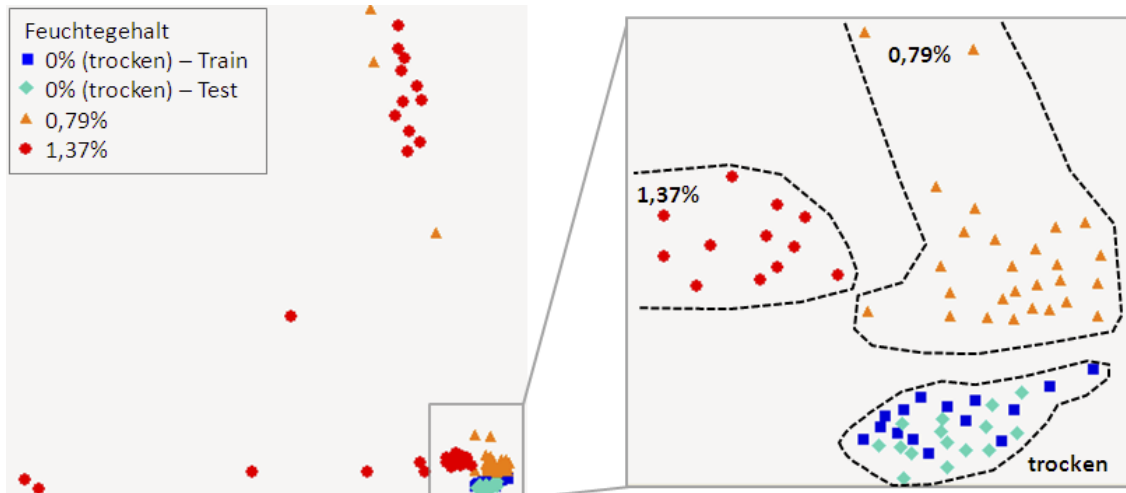
6.2.4 Veränderungen des Werkstoffs

Einen wesentlichen Einfluss auf den Schweißprozess hat der Werkstoff der zu schweißenden Werkstücke. Aufbauend auf der Versuchsreihe aus Abschnitt 6.2.2 wurde für die Herstellung der hier verwendeten Probekörper das ABS-Material durch einen ABS-PS-Blend mit

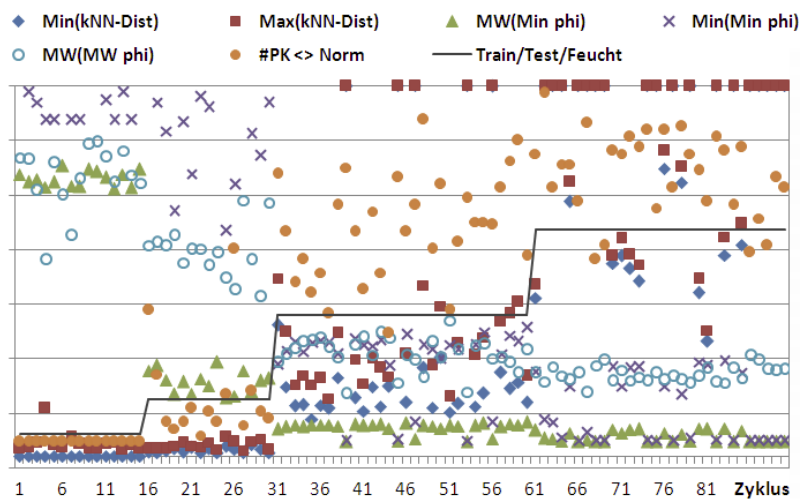
25 % Polystyrol-Anteil ersetzt. Mit den so modifizierten Probekörpern wurden 10 Schweißungen mit den Schweißparametern des Arbeitspunktes aus Abschnitt 6.2.2 durchgeführt. Abbildung 6.7(a) zeigt die Darstellung der normierten Zustandsvektoren der insgesamt 28 Schweißungen – alle mit Arbeitspunkteinstellung, davon 18 Probekörper aus ABS und 10 aus ABS-PS-Blend. Der geänderte Werkstoff wirkt sich deutlich auf das Schweißverhalten aus, so dass eine Überwachung mittels Prozessmodellen möglich ist. Für die Unbekannt-Klassifikation wurde ein GRNFN-Modell erstellt und dessen Unbekannt-Indikatoren berechnet. Für das Modelltraining wurde wiederum nur ein Teil der ABS-Daten – die ersten 10 Zyklen – verwendet, der Rest sowie die ABS-PS-Probekörper dienten als Verifikationsdaten. Der Verlauf verschiedener Unbekannt-Indikatoren des GRNFN-Modells für die drei Datentypen Lerndaten ABS, Testdaten ABS und ABS-PS-Blend ist in Abbildung 6.7(c) dargestellt. Für die Probekörper aus ABS-PS-Blend sind deutlich andere Werte der meisten Indikatoren festzustellen, während die Veränderung für die Verifikationsdaten aus ABS (Zyklen 11–18) deutlich geringer ausfällt. Eine Ausnahme bildet dabei Zyklus 15, der trotz des ursprünglichen Werkstoffs eine hohe Signifikanz für einen unbekanntem Zustand aufweist. Da dieser Zyklus in Abbildung 6.7(a) dem abseits gelegenen Datenpunkt rechts unten entspricht, der einen hohen Abstand zu den sonstigen Schweißungen im Arbeitspunkt aufweist, ist eine Klassifikation als unbekannt gerechtfertigt.

Separation anhand verfügbarer Maschinendaten

Die Maschinensteuerung der verwendeten Versuchsanlage stellt verschiedene Größen zur Verfügung. Diese können mit Toleranzen und Grenzwerten versehen werden, um eine Prozessüberwachung zu realisieren. Die Streuung verschiedener solcher Größen für die Probekörper aus den verschiedenen verwendeten Materialien ABS und ABS-PS-Blend sind in Abbildung 6.7(b) als Boxplot dargestellt. Die Länge des Whiskers beträgt dabei das Anderthalbfache des Interquartilabstands und Ausreißer sind separat gekennzeichnet. Die sich durch die Schweißung ergebende Wegdifferenz – bezeichnet als Weg RPN nach Haltezeit (Abbildung 6.7(b) rechts unten) – weist für beide Werkstoffe keinen signifikanten Unterschied auf. Angesichts des eingestellten RPN-Schweißmodus, d. h. der Beendigung der Schweißung nach einem eingestellten relativen Schweißweg, ist dies plausibel. Die bereits in Abbildung 6.4(a) dargestellte starke Schwankung dieser entsprechenden Prozessgröße wirkt sich direkt auf den Boxplot aus und ist anhand des großen Interquartilsabstands IQR zu erkennen. Die Größe des absoluten Schweißwegs (Tiefe ABS, Abbildung 6.7(b) links unten) weist mit 99,98 % eine sehr hohe Signifikanz auf. In dieser Größe bildet sich direkt das unterschiedliche Schwindungsverhalten der unterschiedlichen Werkstoffe ab, welches sich in einer geringfügig unterschiedlichen Länge des Probekörpers auswirkt (Delta von $\sim 0,02\text{ mm}$ bei 38 mm Gesamtlänge). Auch andere Größen wie die Schweißzeit und die insgesamt durch den Konverter aufgenommene Energie haben signifikant unterschiedliche Mittelwerte. Eine



(a) Veränderungen der Zustandsvektoren bei Zunahme des Feuchtegehalts – Normierungsgrenzen aus trockenem Zustand



(b) Verlauf verschiedener Unbekannt-Indikatoren für Training im trockenem Zustand

Abbildung 6.8: Einfluss des Feuchtegehalts der Werkstücke

eindeutige Zuordnung einer einzelnen Schweißung zur Gruppe der Werkstücke aus ABS oder ABS-PS einzig anhand der Steuerungsdaten ist jedoch nicht möglich, da sich für alle aufgeführten Kenngrößen die Streuungen deutlich überlappen. Dies ist anhand der Boxplots in Abbildung 6.7(b) schnell zu erkennen. Mit GRNFN-Modellen hingegen ist eine eindeutige Zuordnung möglich, wie in der Abbildung 6.7(a) durch die lineare Separierbarkeit der beiden Gebiete leicht ersichtlich ist.

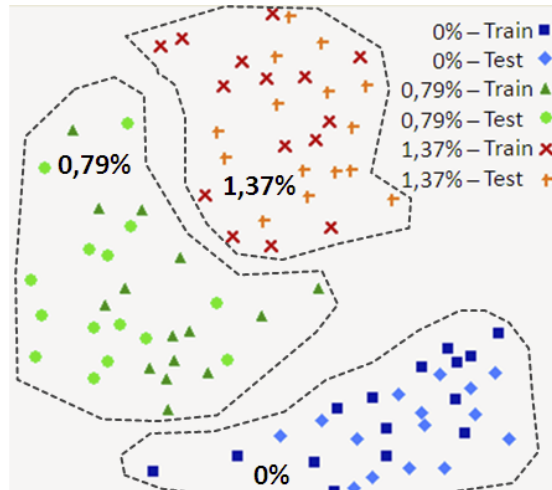
6.2.5 Feuchtigkeitsaufnahme der Schweißteile durch Lagerung

Die verschiedenen Kunststoffe unterscheiden sich teils deutlich in ihrer Hygroskopie, d. h. ihrem Bestreben, Wasser aufzunehmen. Der Feuchtegehalt hat einen wesentlichen Einfluss auf die Schweißbarkeit, da er sich auch stark auf das Dämpfungsverhalten der Werkstoffe

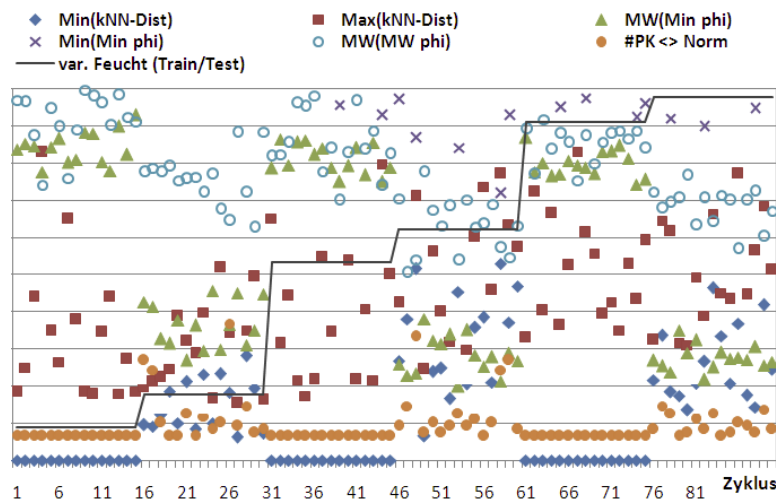
auswirkt (siehe Abschnitt 2.1.3). Bei zu hohem Feuchtegehalt ist teilweise kein Schweißen mehr möglich. Der Feuchtegehalt der Werkstücke wird maßgeblich durch die Dauer der Lagerung sowie die dabei herrschenden Umgebungsbedingungen beeinflusst. Da beim Ultraschallschweißen die Schweißpartner oftmals Zulieferteile sind, entziehen sich diese Größen einer effektiven Kontrolle und sind somit unbekannt.

Zur Untersuchung der Auswirkungen eines variierten Feuchtegehalts der Werkstücke beim Schweißen wurde ein Probekörper aus PA66-GF30 mit ERG-Geometrie im Nahfeld verwendet. Das Kunststoffgranulat wurde vor dem Spritzgießen der Probekörper getrocknet, so dass nur noch ein zu vernachlässigender Anteil an Feuchte enthalten war. Ein Teil der so gefertigten „trockenen“ Probekörper wurde im Klimaschrank bei 40 °C und 90 % Luftfeuchtigkeit kontrolliert konditioniert. Die nach 25 und 76 Stunden entnommenen Probekörper wiesen einen Feuchtegehalt rF von 0,79 % und 1,37 % auf. Ermittelt wurden diese Werte durch die Gewichtszunahme während der Lagerung. Anschließend wurden je 30 Probekörper mit identischen Schweißparametern und den drei Feuchtigkeitsniveaus trocken, $rF=0,79\%$ und $rF=1,37\%$ geschweißt. Insbesondere die Werkstücke mit hohem Feuchtegehalt führten zu Problemen beim Schweißen. Eine zunehmende Anzahl an Teilen wies nach dem Schweißen einen Versatz auf, dessen Ursache vermutlich in einem ungleichmäßigen Aufschmelzen zu suchen ist. Gleichzeitig waren vermehrt Probekörper undicht, so dass für diese auch keine Berstprüfung durchgeführt werden konnte. Wiederum andere Teile konnten bei der Berstprüfung nicht zum Bersten gebracht werden. Lokale Schwachstellen, vermutlich zurückzuführen auf eine unvollständige Verschweißung an dieser Stelle, führten zu Undichtigkeiten und verhinderten einen weiteren Druckaufbau. Für die 30 Schweißungen der Probekörper mit 1,37 % Feuchtegehalt liegen nur für 6 verwertbare Prüfwerte des Berstdrucks vor, während bei den anderen Feuchteniveaus (trocken und $rF=0,79\%$) für nur zwei der 60 Teile kein Wert ermittelt werden konnte. Die mittlere Höhe des Berstdrucks zwischen den einzelnen Gruppen wies keinen statistisch signifikanten Unterschied auf.

Wie in Abschnitt 5.3.1 beschrieben, ist die Normierung des Zustandsvektors der erste, wichtige Schritt bei der Modellbildung. Die Bestimmung der Normierungsgrenzen erfolgte bei diesem Teilversuch anhand der ersten 15 Schweißzyklen mit „trockenen“ Probekörpern. Die Zustandsvektoren für die insgesamt 90 Schweißungen wurden anhand der so gebildeten Grenzen normiert. Abbildung 6.8(a) zeigt die Projektion der so normierten Zustandsvektoren für alle 90 Schweißzyklen. Während die Schweißungen von Probekörpern im trockenen Zustand sehr eng beieinander liegen (siehe Abbildung 6.8(a)), weisen die Schweißungen im befeuchteten Zustand deutlich höhere Schwankungen auf. Diese Schwankungen bilden auch die instabilen Prozesszustände ab, die sich durch den steigenden Feuchtegehalt einstellen. Die Schweißungen mit verschiedenen Feuchtigkeitsniveaus sind in der Sammons-Projektion deutlich getrennt. Es ist somit leicht nachvollziehbar, dass die in Abbildung 6.8(b) dargestellten Unbekannt-Indikatoren eines GRNFN-Modells, das nur anhand



(a) Zustandsvektoren bei Zunahme des Feuchtegehalts



(b) Verlauf verschiedener Unbekannt-Indikatoren

Abbildung 6.9: Prognose des Feuchtegehalts: Zustandsvektoren und Unbekannt-Indikatoren

der ersten 15 Schweißzyklen mit trockenen Teilen trainiert wurde, für die befeuchteten Werkstückzustände deutliche Unterschiede aufweist. Beim Schweißen ist es somit möglich, über GRNFN-Modelle einen geänderten Feuchtegehalt der Werkstücke zu erkennen.

Überwachung des Feuchtegehalts

Wenn Prozessmodelle – wie im vorigen Abschnitt – mittels Werkstücken eines einzigen definierten Feuchtegehalts erstellt werden, können geänderte Zustände zwar erkannt, aber nicht quantifiziert werden. Für die Überwachung des Feuchtegehalts anhand der Signale, die während der Schweißung aufgezeichnet werden, wurde ein weiteres GRNFN-Modell erstellt. Diesmal wurden von allen drei Stufen des Feuchtegehalts jeweils die ersten 15 Zyklen für die Bestimmung von neuen Normierungsgrenzen sowie für das Training des neuen

GRNFN-Modells verwendet. Die verbliebenen 50 % der Versuchsdaten dienen als unabhängige Verifikationsdaten. Die Sammons-Projektion aller 90 Zyklen, deren Zustandsvektoren anhand der neuen Normierungsgrenzen ermittelt wurden, ist in Abbildung 6.9(a) dargestellt. Dabei sind anhand der verschiedenen Marker neben dem Feuchtegehalt auch die Aufteilung in Lern- und Testmuster zu erkennen. Im Vergleich von Abbildung 6.9(a) mit 6.8(a) ergibt sich eine deutlich geänderte Verteilung – zurückzuführen auf die geänderten Normierungsgrenzen. Die Schweißungen von Probekörpern mit verschiedenem Feuchtegrad können, wie Abbildung 6.9(a) zeigt, klar getrennt werden. Das GRNFN-Modell wurde anhand der Hälfte der verfügbaren Versuchsdaten trainiert und anschließend alle 90 Zyklen zur Verifikation verwendet. Der sich ergebende Verlauf der Unbekannt-Indikatoren ist in Abbildung 6.9(b) dargestellt. Die meisten Indikatoren – insbesondere die sehr aussagekräftigen Werte der kNN-Abstände – sind für die Testdaten nicht signifikant unterschiedlich. Dies ist insofern plausibel, als alle Feuchtigkeitsniveaus auch in der Trainingsdatenmenge vorhanden waren. Umgekehrt belegt dies die Wirksamkeit der Unbekannt-Klassifikation mit GRNFN-Modellen. Das erstellte GRNFN-Modell hatte die Prognose des Feuchtegehalts als Zielgröße. Für die gesamten Versuchsdaten wurde eine korrekte Klassifikation für alle Zyklen erzielt – was einer Rate von 100 % entspricht. Dies wiederum bedeutet, dass eine Überwachung des Feuchtegehalts von Werkstücken mit GRNFN-Modellen einzig anhand der Schweißsignale möglich ist. Für eine Wareneingangskontrolle bei Zulieferteilen eröffnet dies ganz neue Möglichkeiten. Voraussetzung ist jedoch, dass für das Systemtraining gezielt Versuchsdaten feuchter Teile zur Verfügung gestellt werden.

Veränderung des Temperaturverlaufs beim Schweißen

Um den Verlauf und die Schwankungen der Temperatur in der Schweißnaht besser beurteilen zu können, wurde bei der zuvor genannten Versuchsreihe mit Probekörpern aus PA66-GF30 zusätzlich der Temperaturverlauf der Naht punktuell durch optische IR-Temperatursensoren erfasst. Dabei wurden drei Sensoren verwendet, die um 120° versetzt um den Probekörper angeordnet sind. Jeder Sensor verfügt über einen Messfleck, der auf die Schweißnaht fokussiert wurde. Der Durchmesser des Messflecks ist abhängig vom Abstand zum Sensor und beträgt etwa 0,5 bis 5 mm. Der Messaufbau ist in Abbildung 6.10(a) zu sehen. Dieses Vorgehen ist jedoch nur bei Probekörpern mit Energierichtungsgeber sinnvoll, da nur dort die Schweißnaht sichtbar ist. Bei der ebenfalls verfügbaren Quetschnahtgeometrie ist die Schweißnaht verdeckt. Dann spielt die Wärmeleitung des Kunststoffes eine dominierende Rolle, was zu stark gedämpften und verzögerten Signalverläufen der Temperatur führt. Der Temperaturverlauf der drei IR-Temperatursensoren für einen Schweißzyklus ist in Abbildung 6.10(b) dargestellt. Der kurze Impuls am Anfang ist der Ultraschall-Aktiv-Trigger. Insgesamt ist ein sehr schnelles Ansteigen der Temperatursignale innerhalb von ca. 0,3 Sekunden auf den Maximalwert festzustellen. Die anschließende Abkühlung erfolgt

deutlich langsamer und lässt den typischen Verlauf einer abfallenden e-Funktion erkennen. Zwischen den einzelnen Zyklen sowie den Signalen der einzelnen Temperatursignale treten teils deutliche Schwankungen auf. Diese sind insbesondere darauf zurückzuführen, dass das Messsignal als Integral über dem Messfleck von ca. 2 mm Durchmesser ermittelt wird und somit das Auftreten und die Menge des Schmelzeaustriebs einen maßgeblichen Einfluss auf das Messsignal hat. Um diese Schwankungen auszugleichen, wird für die Auswertung das Mittelwertsignal aus allen drei Sensoren bestimmt. Aus diesem Signal werden dann am Ende der Messdauer von 30 Sekunden Kenngrößen wie Maximalwert, Mittelwert oder die Endtemperatur ermittelt. Der Verlauf dieser drei Größen ist in Abbildung 6.10(c) zu sehen. Während das Maximum starke Schwankungen aufweist, sind die Mittel- und Endwerte für die einzelnen „Einstellungen“ des Feuchtegehalts sehr stabil. Für Probekörper, die einen nicht zu vernachlässigenden Feuchtegehalt aufweisen, sind die Mittel- und Endwerte deutlich höher. Eine mögliche Erklärung hierfür ist, dass das aufgenommene Wasser eine höhere Wärmekapazität als der (trockene) Kunststoff besitzt, so dass auch mehr Wärme abgegeben werden kann. Zudem ist möglich, dass durch die feuchtigkeitsbedingt höhere Dämpfung auch die sonotrodenseitige Probekörperhälfte mehr Schwingungen absorbiert und sich dadurch leicht erwärmt. Die leicht höhere Grundtemperatur des Probekörpers führt dann aufgrund der Integration über den Messfleck der Sensoren zu einem merklichen Anstieg. Eine Erhöhung des Feuchtegehalts über 0,8 % hat keinen weiteren Einfluss auf die Temperaturerhöhung. Über eine Temperaturmessung ist es somit möglich, für Werkstücke aus PA66 einen trockenen oder unzulässigen Feuchtegehalt anhand der Temperatursignale zu erkennen. Über die Höhe des Feuchtegehalts kann allerdings keine Aussage getroffen werden.

Es ist jedoch festzuhalten, dass diese Aussage nicht allgemein gültig ist, sondern nur für PA66 gilt. Dieselbe Versuchsreihe wurde für ein PPA-Material – ein Polyamid mit aromatischem Anteil – durchgeführt. Für dieses Material war mit ansteigendem Feuchtegehalt ein Abfall der maximalen und mittleren Temperatur verbunden. Ursache für das andere Verhalten ist die werkstoffabhängige Glasübergangstemperatur, die sich bei Befeuchtung deutlich verschiebt. Die Übertragung von Aussagen zu einem Werkstoff auf andere ist somit meist nicht ohne weiteres möglich, da zudem Werkstoffeigenschaften, die insbesondere bei Kunststoffen stark unterschiedlich und von weiteren Einflüssen abhängig sein können, eine große Rolle spielen.

6.3 Prognoseergebnisse an Praxisbauteilen

Im Rahmen der beiden Verbundforschungsprojekte QP-UVS und SC-QUPUS [SF05, NF10] wurden gemeinsam mit den Industriepartnern Versuche an verschiedenen Schweißprozessen durchgeführt. Die untersuchten Prozesse wurden in den Firmen betrieben, um firmeneigene Produkte zu verschweißen. Aus Gründen der geforderten Vertraulichkeit und der Wahrung

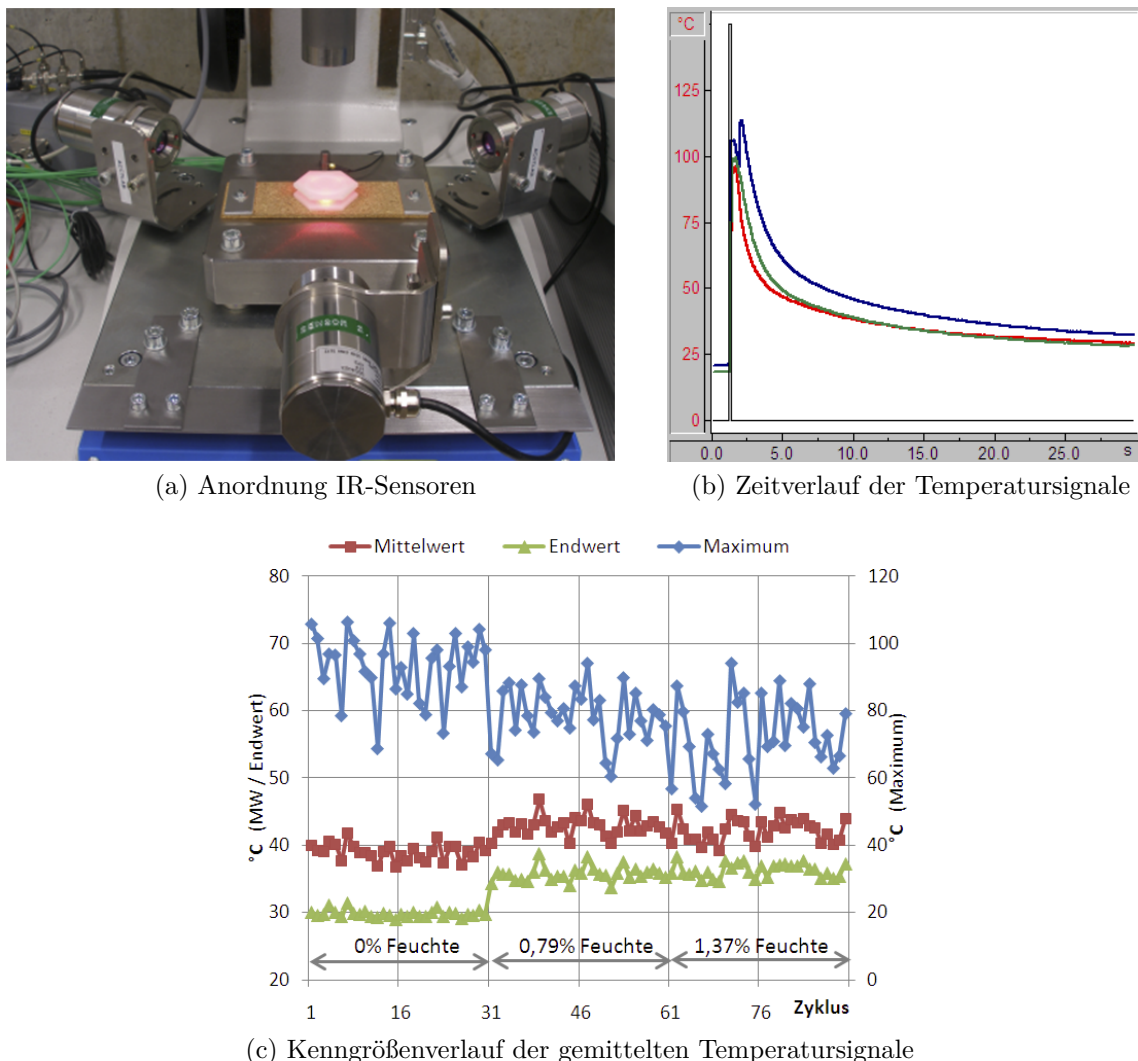


Abbildung 6.10: Erfassung des Temperaturverlaufs mittels IR-Temperatursensoren (für PA66-GF30)

von Betriebsgeheimnissen können im Folgenden weder Firmen- oder Produktnamen genannt noch Bilder aller Produkte dargestellt werden.

Bei der Auswahl und Durchführung der Versuchsreihen waren einige Randbedingungen zu beachten, die bei Probekörperversuchen so nicht auftraten, und sich deshalb insbesondere auf die Gestaltung der Versuchsreihen auswirken:

- Die zu verschweißenden Produkte sind oftmals Zulieferteile. Die Spritzgießbedingungen können somit weder überwacht noch beeinflusst werden.
- In vielen Fällen sind die Lagerzeiten sowie die Umgebungsbedingungen nicht bekannt, bei Zulieferteilen kommen noch die Einflüsse des Transports hinzu. Diese Größen können zwischen einzelnen Werkstücken deutlich schwanken und den Prozess beeinflussen, ohne verifizierbar zu sein.

- Die Spritzgießprozesse verfügen in der Regel über keine (Werkzeug)Sensorik und können somit nicht wie beschrieben überwacht werden. Eine Überprüfung von Schwankungen kann also nicht erfolgen.
- Bei teils verketteter Fertigung sind die Spritzgießprozesse direkt eingebunden, so dass sie nicht variiert werden können. Versuchsumfang und -variationen sind damit eingeschränkt.
- Bei vielen Produkten existieren keine quantifizierbaren Merkmale für die Qualität oder es gibt keine entsprechenden Prüfeinrichtungen und -vorschriften. Beim Anschweißen eines Druckknopfes besteht z. B. „nur“ die Anforderung, dass die Schweißung die beiden Bauteile verbindet – schwankende „Qualitätswerte“ können damit nicht erfasst werden. Auch bei größeren Prozessvariationen ist dies gegeben, so dass Prozessschwankungen einem konstanten, jedoch weit interpretierbaren „Qualitätskriterium“ gegenüberstehen und somit keine Auswertungen und Modellbildungen möglich sind.
- Sind quantifizierbare Qualitätsmerkmale festgelegt, fehlen oftmals geeignete Prüfverfahren. Ansonsten ist der Prüfaufwand oft so groß, dass er nicht wirtschaftlich realisiert werden kann, ohne die laufende Fertigung zu beeinträchtigen.
- Insbesondere bei der Dichtheitsprüfung sind die Prüfwerte teilweise mit prüfspezifischen Streuungen oder Fehlern überlagert. Eine derartige Dichtheitsprüfung lässt keinen sicheren Schluss auf die Ursache der Undichtigkeit zu – entweder eine mangelhafte Schweißnaht oder eine mangelhafte Ankopplung und Abdichtung der Prüfeinrichtung z. B. durch Fussel.

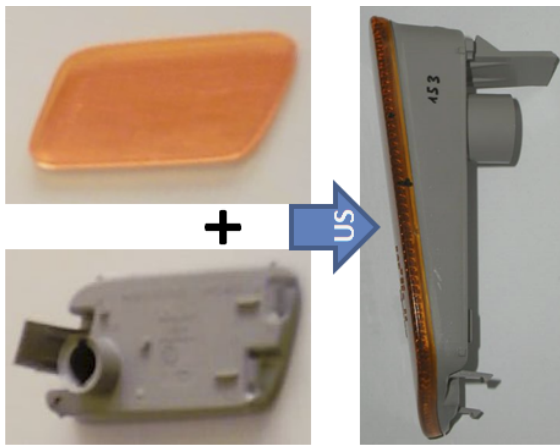
Diese Fakten führen dazu, dass sich die Versuche mit Praxisteilen an den Gegebenheiten orientieren müssen und oftmals Abstriche beim Versuchsplan hinzunehmen sind. Eine Verifikation der bisher beschriebenen, verschiedenen Fähigkeiten der GRNFN-Modelle anhand eines einzigen Produktes ist somit nicht durchführbar, da die einsetzbaren Überwachungsmöglichkeiten stark vom Produkt abhängen.

6.3.1 Einfluss eines veränderten Spritzgießprozesses

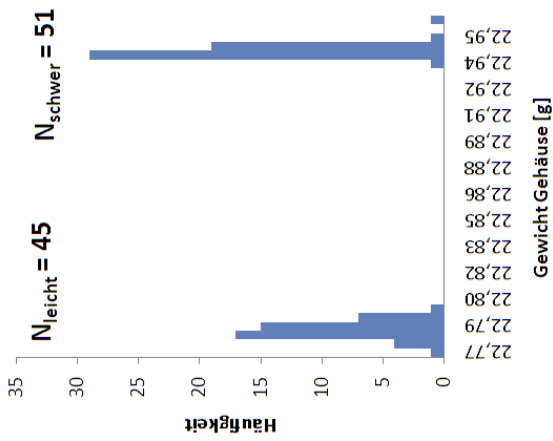
Ein Projektpartner verschweißt ein Gehäuse aus einem ABS-Blend mit einer gefärbten Scheibe aus transparentem PMMA (Plexiglas) zu einer Automobilleuchte (siehe Abbildung 6.11(a)). Die Nahtkonstruktion umfasst einen Energierichtungsgeber, der am Gehäuse angebracht ist. Scheiben- und Gehäuseteile für den Schweißprozess stammen aus der Serienfertigung des Industriepartners. Die Gehäuseteile lassen sich in zwei getrennte Gewichtsklassen aufteilen (siehe Abbildung 6.11(b)). Der Gewichtsunterschied ist dabei ca.

0,2 g, was $< 1\%$ des Sollgewichts entspricht. Die Teile stammen offensichtlich aus verschiedenen Fertigungschargen. Über die Ursachen der Gewichtsschwankung – andere Charge des Spritzgießmaterials oder veränderte Spritzgießparameter – ist nichts Näheres bekannt. Alle 96 Versuchsteile wurden mit identischen Schweißparametern verschweißt. Die Verteilung der normierten Zustandsvektoren ist in Abbildung 6.11(c) dargestellt, Zyklen mit leichtem oder schwerem Gehäuse wurden entsprechend markiert. Die beiden Gewichtsklassen wirken sich in einer deutlichen Trennung der Zustandsvektoren aus. Wenn bekannt ist, dass leichte und schwere Gehäuse auftreten können, so kann ein entsprechendes GRNFN-Modell zur Unterscheidung einzig anhand von Daten aus dem Schweißprozess verwendet werden. Anhand der Zustände in Abbildung 6.11(c) ist – unter Einbezug der Erkenntnisse aus Kapitel 4 – eine Unterscheidung der Teilegewichte und damit des veränderten Spritzgießprozesses mit hoher Güte möglich.

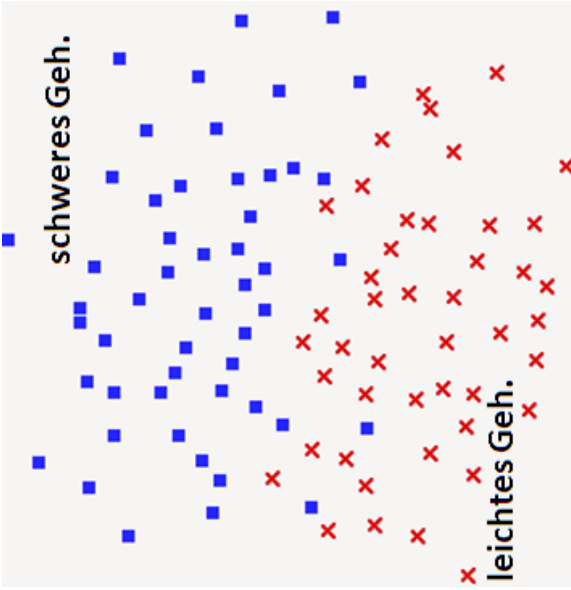
In der Regel ist jedoch nur eine Charge zum Einrichten des Prozesses und zum Aufbau der Prozessüberwachung vorhanden. Das Auftreten einer Charge mit anderem Teilegewicht muss dann als neuer Zustand – und meist als Störung mit Ursachen in den Vorprozessen – identifiziert werden. Um die Möglichkeit des Einsatzes von GRNFN-Modellen beurteilen zu können, wurden die vorhandenen Versuchszyklen aufgeteilt. Die ersten 35 Zyklen mit leichtem Gehäuse dienen zum Training des GRNFN-Modells. Die restlichen 10 Zyklen mit leichtem und die 51 Zyklen mit schwerem Gehäuse dienen der Verifikation des Modells und einer Aussage zu den Fähigkeiten einer Unbekannt-Klassifikation. Die in Abschnitt 4.6 vorgestellten Indikatoren für Zustände, die einem GRNFN-Modell unbekannt sind, werden für alle 96 Versuchszyklen berechnet. Eine Auswahl der Indikatoren ist in Abbildung 6.11(d) dargestellt. Aus Darstellungsgründen wurden die einzelnen Größen teilweise umskaliert, weshalb auf die Angabe einer Achsskalierung verzichtet wird. Die Indikatoren, die sich aus den Abständen zu den k Nächsten Nachbarn ergeben, sind wiederum logarithmisch skaliert; die anderen, auf Aktivierungswerten und Anzahl der Überschreitung der Normierungsgrenzen basierenden Indikatoren sind linear aufgetragen. Die Lern- und Testmuster mit leichtem Gehäuse sowie die mit schwerem Gehäuse sind – in dieser Reihenfolge – anhand der eingezeichneten Treppenfunktion in Abbildung 6.11(d) zu erkennen. Insgesamt ist festzuhalten, dass die Indikatoren für alle dem Modell nicht trainierten Zyklen – also den 10 leichten und 51 schweren Gehäusen – teils starke Schwankungen aufzeigen. Diese sind jedoch für die leichten Gehäuse im Mittel von der Veränderung anders als bei den Zyklen mit schwerem Gehäuse. Auch wenn es keinen einzelnen Wert gibt, der für alle Zyklen eine gesicherte Aussage ermöglicht, so kann doch für jede Größe ein Schwellwert angegeben werden, der für die meisten Zyklen eine korrekte Trennung ermöglicht. Diese Schwellwerte können am besten anhand der Werte, welche sich für untrainierte Zyklen ergeben, abgeleitet werden. Abbildung 6.11(c) zeigt, dass sich die beiden Klassen nicht komplett trennen lassen, sondern eine gewisse Überlappung oder Durchmischung existiert. Dies ist unter anderem darauf



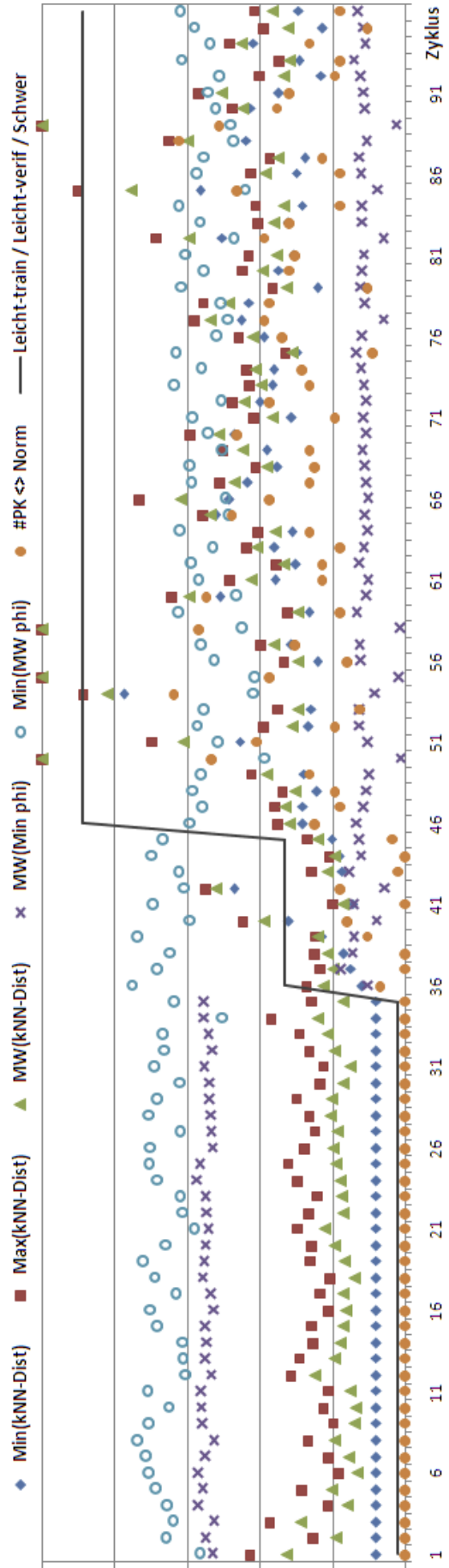
(a) Produkt Automobilleuchte



(b) Gewichtsverteilung der Gehäuse aus Serienfertigung



(c) Verteilung der Zustandsvektoren beim Schweißen der beiden Gehäuseklassen



(d) Verlauf verschiedener Unbekannt-Indikatoren bei unterschiedlichem Gehäusegewicht

Abbildung 6.11: Schweißen von Automobilleuchten unterschiedlicher Gewichte

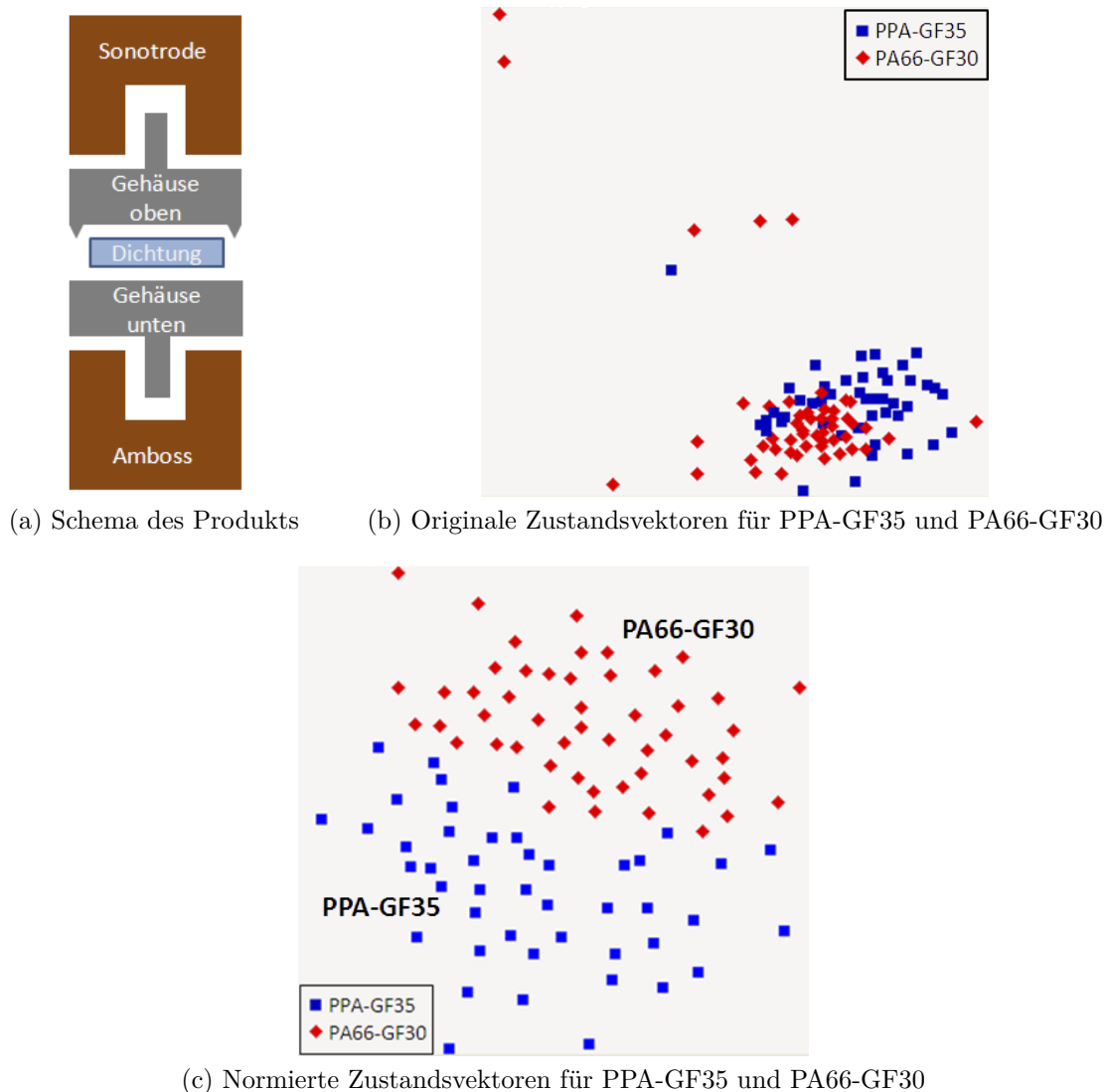


Abbildung 6.12: Verschweißen eines Rückschlagventils mit vertauschtem Material

zurückzuführen, dass neben dem hier betrachteten Gewicht der Gehäuse auch die Scheiben sowie insbesondere der Schweißprozess selbst Schwankungen unterliegen, die sich zusätzlich bemerkbar machen können. Prinzipiell ist jedoch der Einfluss des Spritzgießens auf den Schweißprozess sowie die Möglichkeit nachgewiesen, dies einzig anhand von Kenngrößen aus den Schweißsignalen mittels GRNFN-Modellen und daraus abgeleiteten Indikatoren erkennen zu können.

6.3.2 Erkennen von Materialänderungen bei einem Rückschlagventil

Im Rahmen der Versuche mit den Partnern wurde auch ein Rückschlagventil als Produkt ausgewählt. Im Wesentlichen besteht dieses aus einem Ober- und einem Unterteil, zwischen

denen vor dem Schweißen eine Kunststoffplatte als Ventilelement für die Sperrfunktion in einer Richtung eingelegt wird. Die beiden Gehäuseteile sind zylinderförmig und verfügen jeweils über einen Anschlussstutzen für die spätere Medieneinleitung. Am oberen Teil ist konstruktiv der Energierichtungsgeber angebracht. Der schematische Aufbau des Rückschlagventils ist in Abbildung 6.12(a) zu sehen. Die wesentlichen Qualitätskriterien sind Dichtigkeit und Sperrfunktion des Rückschlagventils. Die Dichtigkeit ist konstruktionsbedingt unkritisch und die Sperrfunktion wird maßgeblich durch die Kunststoffplatte erreicht, die als reines Einlegeteil nicht an der Schweißung beteiligt ist. Es existieren somit keine prüf- baren Qualitätskriterien der Schweißung. Der Fokus der Versuchsreihe wurde aus Gründen der Praxisrelevanz auf die Variation der Vorprozesse gelegt.

Geänderter Grundwerkstoff

Die Partnerfirma, die das Rückschlagventil fertigt, hat für ihre Produkte einige Standardwerkstoffe und bevorzugt insbesondere PA66-GF30. Das Rückschlagventil ist werkstoff- technisch jedoch ein Sonderfall und als Material wird PPA-GF35 – und nicht PA66-GF30 – verwendet. Ein Problem ist die Verwechslung des Materials beim Spritzgießen. Es ist schon vorgekommen, dass die Gehäuseteile versehentlich aus einem PA66-GF30 als Standardma- terial statt des PPA-GF35 gefertigt wurden. Da beide Materialien schwarz eingefärbt sind und eine ähnliche Dichte haben, ist eine schnelle visuelle Unterscheidung nicht möglich. Für die durchgeführten Schweißversuche wurden neben Gehäuseteilen aus PPA-GF35 auch sol- che aus PA66-GF30 verschweißt. Die direkt aus den Signalen abgeleiteten Zustandsvektoren der 95 Versuchszyklen sind in Abbildung 6.12(b) dargestellt. Dabei ist ein Hauptcluster zu erkennen, in dem die Mehrzahl aller Zyklen – unabhängig vom Material – liegt. Aufgrund der lokalen Begrenzung und homogenen Verteilung weisen die Zustandsvektoren ähnliche Abstände zueinander auf, was auf einen relativ stabilen Prozess hinweist. Allerdings gibt es einige deutliche Ausreißer. Eine Trennung der Schweißzyklen nach den verwendeten Ma- terialien anhand der originalen Zustandsvektoren erscheint nicht möglich.

Anschließend wurden anhand der Schweißteile mit dem für die Serienfertigung vorgese- henen Material PPA-GF35 Normierungsgrenzen für die einzelnen Komponenten des Zu- standsvektors bestimmt und auf die Schweißzyklen für beide Materialien angewendet. Die Projektion der so erhaltenen, normierten Zustandsvektoren ist in Abbildung 6.12(c) dar- gestellt. Für jeden der beiden Werkstoffe ergibt sich ein kompakter, nicht überlappender Cluster. Somit lassen sich die Cluster prinzipiell trennen und eine Unterscheidung der Werk- stoffe anhand der Zustandsvektoren ist möglich. Aufgrund der relativen Nähe der Cluster kann es nur in der Nähe der Trennebene zu einer Falschklassifikation kommen. Da in der Praxis die Bauteile jedoch chargenweise verschweißt werden, ist die prinzipielle Trennung der Chargen wichtig. Eine Charge kann bei Berücksichtigung einer Stichprobe mit mehreren Schweißungen sicher zugeordnet werden. Damit ist der Nachweis erbracht, dass auch in der

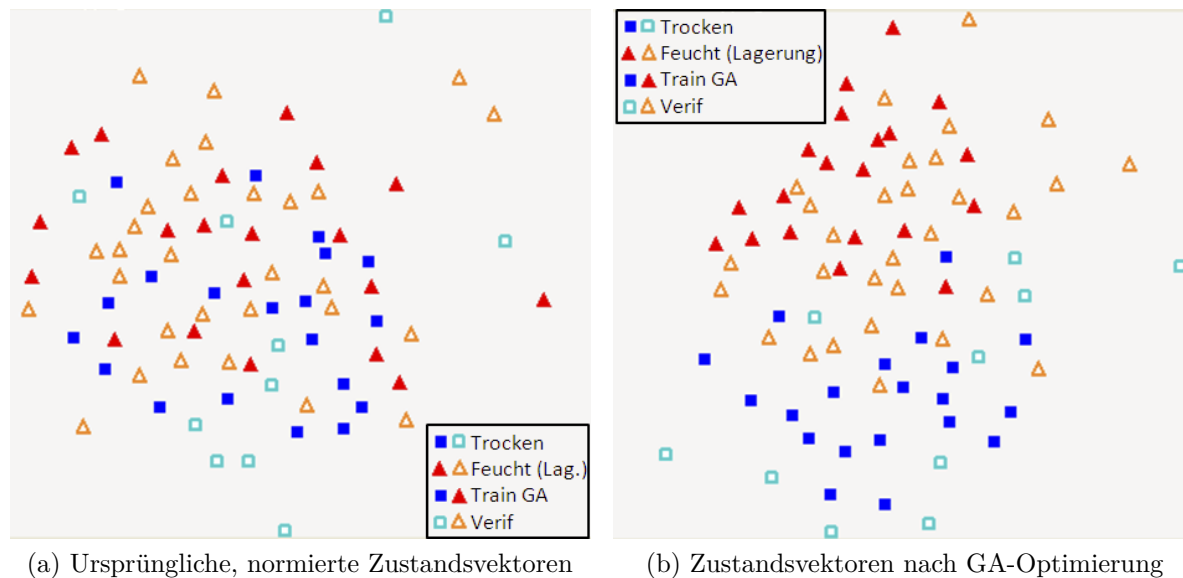


Abbildung 6.13: Kenngrößenselektion für das Rückschlagventil zur Identifikation gelagerter Teile

Praxis eine Vertauschung des Materials anhand der Schweißsignale prinzipiell erkannt werden kann. Dabei ist es nicht zwingend notwendig, für die Modellerstellung Werkstücke aus allen denkbaren „falschen“ Materialien zu verschweißen. Die in Abbildung 6.12(c) erzielte Trennung ergab sich einzig anhand der für das Serienmaterial ermittelten Normierungsgrenzen. Gebiete außerhalb des Clusters der Zustandsvektoren mit Serienmaterial können also standardmäßig als Fehler angesehen werden, die teils auf ein geändertes Material zurückgeführt werden können.

Feuchtigkeitsaufnahme durch Lagerung

Der für das Rückschlagventil verwendete Werkstoff gehört zu den Polyamiden, so dass die Aufnahme von Feuchtigkeit eine Rolle spielt. Die Feuchtigkeitsaufnahme erfolgt dabei während der Lagerung aus der Umgebungsluft. Der Feuchtegehalt von PPA-Materialien ist dabei mit ca. 0,4 % deutlich geringer als es beim Probekörper aus PA66-GF30 (Abschnitt 6.2.5) der Fall war. Um eine Feuchtigkeitsaufnahme aus der Umgebung zu verhindern, wurden für eine Versuchsreihe einige Gehäuseteile direkt nach dem Spritzgießen in spezielle Kunststoffbeutel gepackt, die zusätzlich mit einem Schutzgas gefüllt und anschließend versiegelt wurden. Andere Gehäuse wurden für einige Wochen bei Raumklima gelagert. Die beiden Gruppen wurden anschließend mit identischer Maschineneinstellung verschweißt. Für die Zustandsvektoren wurden aus den ersten 20 Zyklen der trockenen gelagerten Gehäuseteile Normierungsgrenzen ermittelt und für die Normierung der Zustandsvektoren aller Versuchszyklen verwendet. Die sich so ergebenden Zustandsvektoren sind in Abbildung 6.13(a) dargestellt. Eine klare Trennung der beiden Gehäusegruppen ist nicht möglich.

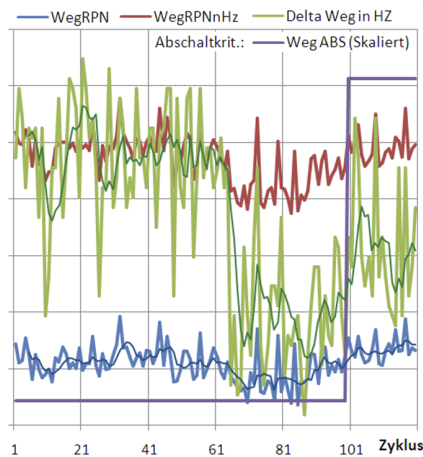
Anschließend wurde versucht, den Zustandsvektor durch GA hinsichtlich der Kenngrößen zu optimieren. Ziel war, 15 bis 30 Kenngrößen aus den ursprünglich 133 verfügbaren Kenngrößen des Zustandsvektors auszuwählen und ihnen zusätzlich einen Skalierungsfaktor zuzuordnen, so dass die beiden Gruppen (geschützt und ungeschützt gelagert) möglichst gut zu trennen sind. Als Gütekriterium wurde das Überlappungsmaß verwendet (Formel (5.6)). Die Trainingsdaten für die Optimierung mittels GA bestanden aus den jeweils ersten zwanzig Zyklen beider Gruppen. Die verbleibenden 40 Schweißzyklen (10 für trockene, 30 für gelagerte Gehäuse) dienten anschließend der Verifikation. Die Projektion der optimierten Zustandsvektoren für alle 80 verfügbaren Zyklen ist in Abbildung 6.13(b) zu sehen. Die Umrissform der Marker dient zur Unterscheidung zwischen geschützt und ungeschützt gelagerten Teilen, die gefüllten Marker zeigen die zum Training verwendeten Vektoren an. Für die Trainingsmenge ist eine gute Trennung der beiden Bereiche zu erkennen. Auch die Verifikationsdaten ordnen sich tendenziell wie die Trainingsdaten an. In der Mitte ist jedoch ein gewisser Überlappungsbereich erkennbar, was hier zu Fehlklassifikationen führen kann. Durch eine Vergrößerung der Trainingsmenge oder ein gezieltes Hinzufügen kritischer Zyklen ist indes eine weitere Optimierung zur Separierung denkbar.

Nachteilig gegenüber der Erkennung des vertauschten Grundwerkstoffs ist, dass für das Training beide zu unterscheidenden Zustände verfügbar sein müssen. Dies ist aber durch eine entsprechende Lagerung verhältnismäßig einfach zu erreichen. Prinzipiell ist jedoch nachgewiesen, dass auch ein geänderter Feuchtegehalt der Schweißteile erkannt werden kann.

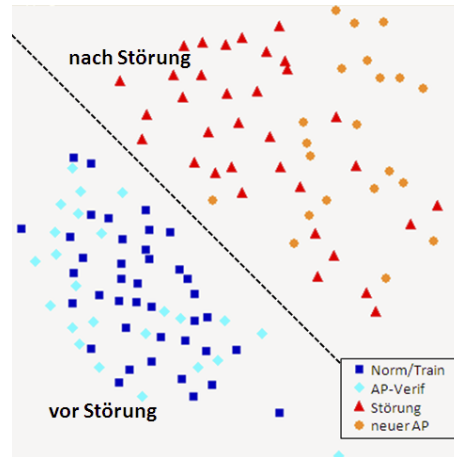
6.3.3 Prozessstörung und Arbeitspunktnachführung

In einer weiteren Versuchsreihe bei einem Industriepartner ging es um das Anschweißen eines Tasters an eine Baugruppe. Eine wesentliche Qualitätsanforderung an das Produkt ist die Einhaltung von Geometriemaßen in Schweißrichtung. Als Abschaltkriterium des Schweißprozesses ist somit ein absoluter Wegpunkt eingestellt. Da die Maßprüfung parallel zur Fertigung zu aufwändig gewesen wäre und ansonsten kein weiteres prüfbares Qualitätskriterium existiert, wurden die Schweißdaten und -signale während der Fertigung aufgezeichnet und später ausgewertet. Im Folgenden ist ein Ausschnitt der Daten näher untersucht worden.

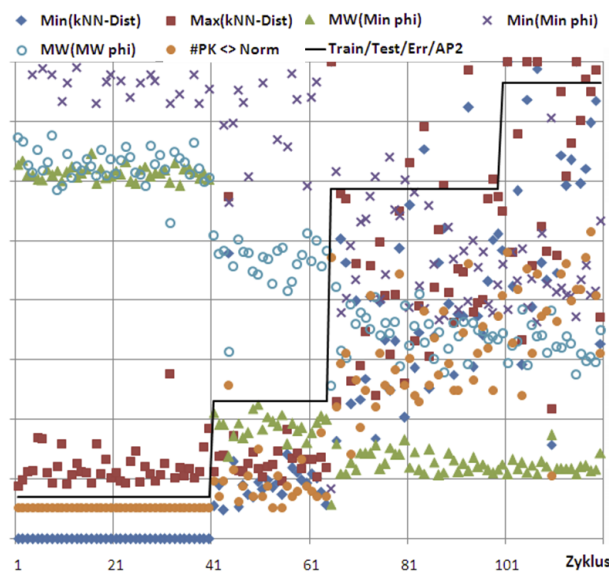
In Abbildung 6.14(a) ist der Verlauf verschiedener Wegkenngrößen inklusive deren gleitender Mittelwerte zu sehen. Wegen der relativ starken Schwankungen der Werte gestaltet sich die Festlegung von Toleranzen zur Prozessüberwachung als schwierig. Dargestellt ist auch die Größe *Weg RPN*, also der während des aktiven Ultraschalls zurückgelegte Schweißweg. Aufgrund des Schweißmodus *Weg absolut* unterliegt diese Größe Schwankungen. Bei den bisher vorgestellten Versuchen war *Weg RPN* meist das Abschaltkriterium für den Ultraschall und damit konstant. Ab Zyklus 65 ist eine deutliche Veränderung der Wegkenngrößen zu erkennen, die auf eine nicht näher spezifizierte Störung im Prozess zurückgeht.



(a) Verlauf von Weg-Kenngrößen



(b) Normierte Zustandsvektoren vor und nach der Störung



(c) Verlauf verschiedener Unbekannt-Indikatoren

Abbildung 6.14: Anschweißen eines Tasters mit Prozessänderung und Arbeitspunktanpassung

Diese Störung ist nicht singular, sondern dauert an. Knappe 40 Zyklen später hat ein Prozessverantwortlicher auf die Veränderung im Prozess reagiert und das Abschaltkriterium angepasst. Es ist zu vermuten, dass im Rahmen einer Stichprobenkontrolle das erforderliche Teilemaß nach der Schweißung nicht mehr in der Toleranz lag und infolgedessen der Schweißweg, bei dem der Schweißprozess beendet wird, erhöht wurde. Über die Ursachen ist nichts Näheres bekannt, aber eine neue Charge an Werkstücken wäre eine Möglichkeit.

Interessant ist die Frage, ob die aufgetretene Prozessveränderung mit einem Überwachungsmodell früher hätte erkannt werden können. Anhand der ersten 40 Zyklen wurden die Normierungsgrenzen für die Komponenten des Zustandsvektors ermittelt. Die Projektion der damit gebildeten normierten Zustandsvektoren für alle Zyklen inklusive der Störung sowie die nachfolgende Prozessanpassung sind in Abbildung 6.14(b) zu sehen. Es ist eine

deutliche Ausbildung von Clustern zu erkennen. Die Zyklen 1 - 64 (linke untere Bildhälfte) entsprechen den ungestörten Zyklen im ursprünglichen Arbeitspunkt. Die zur Normierung verwendeten Zyklen sind separat markiert. Bei den Zyklen ab Nr. 65 (Cluster rechts oben) trat die Prozessstörung auf. Die Zyklen mit anschließend geändertem Abschaltkriterium sind wiederum mit einem anderen Marker versehen. Aufgrund der deutlichen Trennung der Cluster scheint eine Erkennung über die Unbekannt-Klassifikation mittels GRNFN-Modellen möglich. Mit den ersten 40 Zyklen, die auch zur Berechnung der Normierungsgrenzen verwendet wurden, wurde ein GRNFN-Modell erstellt. Die anschließend mit dem Modell ermittelten Unbekannt-Indikatoren für alle Zyklen sind in Abbildung 6.14(c) zu sehen. Die eingezeichnete Treppenfunktion gibt die verschiedenen Zyklusbereiche an: Die für die Normierung und das Modelltraining verwendeten Zyklen, die „Verifikationszyklen“ im ungestörten Arbeitspunkt, die Zyklen des gestörten Prozesses sowie zuletzt bei geändertem Arbeitspunkt. Mit Auftreten der Störung ist eine deutliche Veränderung in den Indikatoren erkennbar. Die Störung wäre demnach in der Tat durch das GRNFN-Modell zeitnah erkannt worden, so dass auch die Reaktion des Prozessverantwortlichen auf die geänderten Prozesszustände hätte früher erfolgen können.

6.3.4 Erkennen fehlender Einlegeteile

Im Rahmen einer anderen Versuchsreihe wurde das Ultraschallschweißen an einem Wasserverteiler für niedrige Durchflussraten untersucht. Das Produkt besteht aus zwei Gehäuseteilen mit einer ERG-Nahtgeometrie, in die vor dem Schweißen eine weiche Filtermatte eingelegt wird. Der schematische Produktaufbau ist in Abbildung 6.15(a) skizziert. Das Schweißen der Wasserverteiler im Arbeitspunkt erfolgte teils mit und teils ohne eingelegten Filter, um den denkbaren Fehler eines vergessenen Einlegeteils während der Produktion untersuchen zu können. Die normierten Zustandsvektoren der insgesamt 50 Schweißzyklen (je 25 mit und ohne eingelegten Filter) sind in Abbildung 6.15(b) dargestellt. Die Normierung erfolgte anhand der ersten 15 Zyklen des vorschriftsmäßigen Arbeitspunktes mit eingelegtem Filter. Dabei ist keine Trennung der Schweißungen mit und ohne Filter erkennbar. Das Überlappungsmaß als Maß für die Trennbarkeit ergibt sich zu $J_5 = 0,67$.

Ausgehend davon wurde – wie bei den Versuchen mit Rückschlagventilen (Abschnitt 6.3.2) – versucht, den Zustandsvektor dahingehend zu optimieren, dass eine bessere Trennbarkeit erzielt wird – auch für unabhängige Verifikationsmuster. Als Daten für die Optimierung wurden von jeder Einstellung, d. h. mit und ohne eingelegten Filter, die ersten 15 Zyklen verwendet, so dass noch jeweils 10 Zyklen für die Verifikation verblieben. Das Ergebnis der Optimierung durch Signifikanzanalyse mit einem Signifikanzniveau von 85 % ist in Abbildung 6.15(c) zu sehen. Die Zyklen, die für Normierung, Optimierung durch Kenngrößen-selektion und Verifikation verwendet wurden, sind anhand der Marker zu erkennen. Auch wenn „nichtrelevante“ Komponenten durch die Signifikanzanalyse entfernt wurden und ein

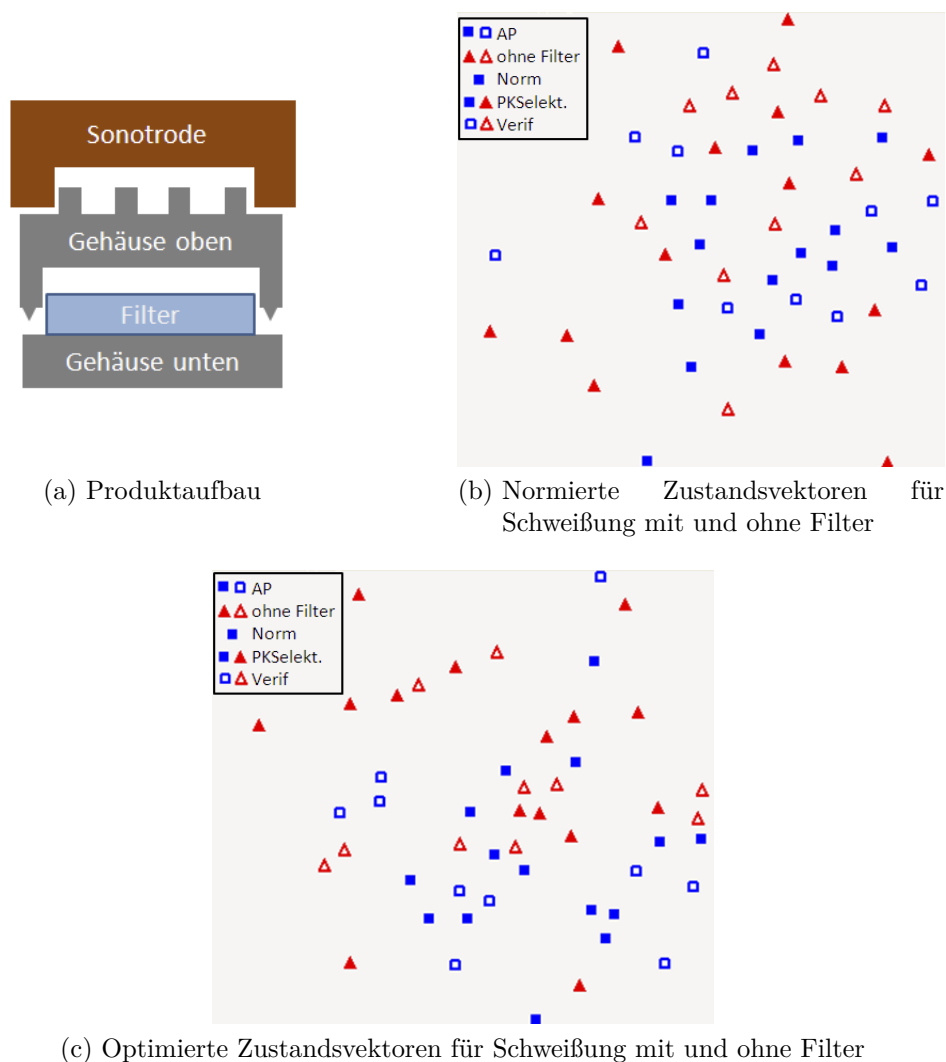
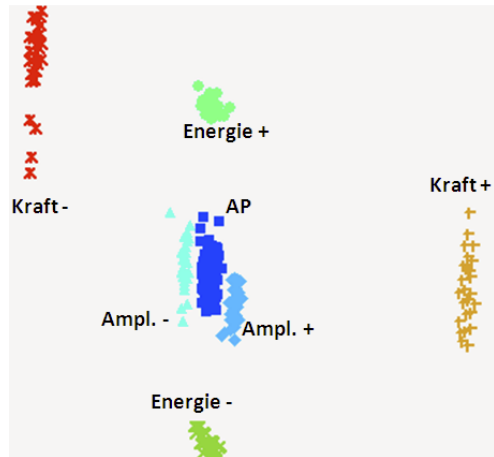


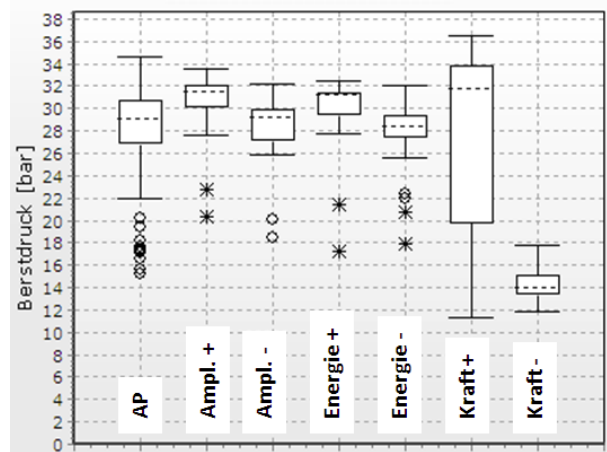
Abbildung 6.15: Verschweißen eines Wasserverteilers mit Filtereinsatz

Zustandsvektor jetzt nur noch aus 13 Komponenten besteht, ist keine bessere Trennung zu erreichen (Überlappungsmaß $J_5 = 0,64$). Ähnliche Ergebnisse ergeben sich auch bei der Kenngrößenselektion und -skalierung durch GA.

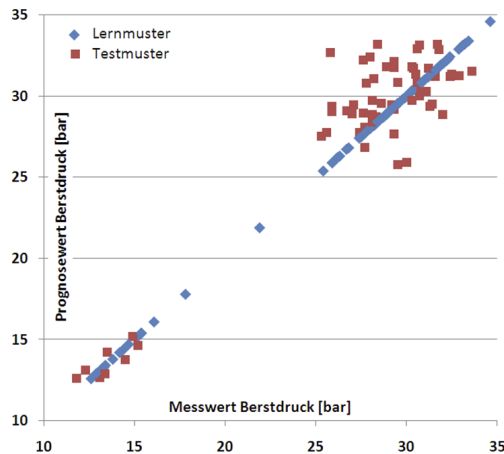
Dass eine Trennung von Schweißzyklen mit und ohne Filter nicht möglich ist, erscheint physikalisch plausibel, da der Filter keinen signifikanten Einfluss auf die Schweißung hat: Er liegt komplett im Inneren des Gehäuses und weist keine nennenswerten Klemm- oder Widerstandskräfte auf, die auf den Schweißprozess rückwirken könnten. Dass bei nicht existierenden Zusammenhängen auch keine Verbesserung und keine hohe Trennschärfe der Zustandsvektoren erzielt wird, heißt im Umkehrschluss, dass das vorgestellte Verfahren robust ist. Eine Kenngrößenselektion der realisierten Art führt also nicht nur zu scheinbar hohen Modellgüten – die Verfügbarkeit einer ausreichenden Menge an Daten und unabhängigen Verifikationsdaten vorausgesetzt –, sondern gibt reale Unterschiede im Prozess wieder, sofern diese existieren. Sollten sich Einlegeteile etwa in Form federnder Rückwirkung (z. B.



(a) Verteilung der normierten Zustandsvektoren



(b) Boxplot des Berstdrucks für variierte Prozesseinstellungen



(c) Berstdruck-Diagramm: Prognose und Messwert

Abbildung 6.16: Qualitätsprognose mit variierten Schweißparametern für einen Wasserverteiler

O-Ringe) auf den Schweißablauf auswirken, müssten fehlende Teile mit diesem Überwachungsansatz eindeutig zu erkennen sein.

6.3.5 Qualitätsprognose unterschiedlicher Merkmale für Wasserverteiler

Für den Wasserverteiler (Abbildung 6.15(a)) gibt es Qualitätsmerkmale, für die Prüfeinrichtungen existieren und die somit auch quantifizierbar sind. Deshalb wurden weitere Versuche mit einer Variation verschiedener Prozessparameter durchgeführt, um die Möglichkeiten einer Qualitätsprognose zu untersuchen. Für den Schweißprozess ist dabei Abschaltkriterium die eingebrachte Energie und nicht – wie bei den meisten bisher betrachteten Prozessen – der Schweißweg. Der Schweißweg ist also eine resultierende Größe jeder Schweißung und

deshalb nicht als Stellgröße variiert worden. Das Sammons-Mapping aller normierten Zustandsvektoren ist in Abbildung 6.16(a) dargestellt; zur Bestimmung der Normierungsgrenzen wurden 45 der 75 Schweißungen im Arbeitspunkt verwendet. Eine deutliche Trennung der einzelnen Einstellungen ist erkennbar, so dass eine Überwachung der Einstellung – wie für den Probekörper in Abschnitt 6.2.1 beschrieben – problemlos möglich ist.

Die Variation der Schweißparameter hat Einfluss auf die erzielbare Berstdruckfestigkeit der Schweißnaht als wesentliches Qualitätsmerkmal. Für die Versuchsteile wurde der Berstdruck in einem Druckprüfstand mit Wasser als Medium ermittelt. Die Streuung des Berstdrucks für die verschiedenen Maschineneinstellungen sind in Abbildung 6.16(b) anhand von Boxplots dargestellt. Neben der Lage ist auch die Streubreite gut zu erkennen. Auffällig ist, dass es einige deutliche Ausreißer nach unten gibt, die sich in den Signalen und ebenso in den Zustandsvektoren nicht wiederfinden. Die Ausreißer liegen demnach höchstwahrscheinlich in der Prüfung begründet. Der Wasserverteiler verfügt über mehrere Anschlüsse, die vor der Prüfung abgedichtet werden müssen. Eine mögliche Ursache für einen zu niedrigen Berstdruck liegt darin, dass diese Abdichtungen dem Prüfdruck nachgeben und zu Leckagen führen. Die Prüfung wurde mit Erreichen eines eingebrachten Flüssigkeitsvolumens beendet, ohne dass die Schweißnaht geborsten wäre. Als Ergebnis lieferte die Prüfanlage den insgesamt erzielten Maximaldruck, der durch Leckagen entsprechend reduziert sein kann und nicht zum Versagen der Schweißnaht geführt hat.

In Abbildung 6.16(b) ist weiter auffällig, dass insbesondere bei einer Erhöhung der Schweißkraft die Streuung stark ansteigt. Die Schweißkraft ist der wesentliche Parameter des Prozesses, der einerseits maßgeblich über die Höhe der erzielbaren Schweißnahtfestigkeit entscheidet, andererseits aber auch kritisch hinsichtlich der Prozessstabilität und -robustheit ist. Anhand der Versuchsdaten wurden GRNFN-Modelle für die Qualitätsgrößen Berstdruck, Spaltmaß und Sonotrodenabdruck erstellt, wobei jeweils nur zwei Drittel der verfügbaren Schweißzyklen für das Training verwendet wurden. Das verbleibende Drittel dient zur Verifikation der erzielten Modellgüte. Die Ergebnisse der Berstdruckprognose durch das GRNFN-Modell sind in Abbildung 6.16(c) dargestellt. Dabei wurden die erläuterten Ausreißer sowie die Einstellung mit erhöhter Schweißkraft aufgrund der Prozessinstabilität nicht berücksichtigt, so dass von den 225 Versuchszyklen insgesamt 179 in diese Untersuchung einbezogen wurden. Die Korrelation zwischen dem geprüften Berstdruck und dem durch das GRNFN-Modell prognostizierten Berstdruck über die 179 Versuchszyklen beträgt $r = 0,976$ mit einem maximalen Fehler von 6,85 bar und einem mittleren Fehler von 0,58 bar.

Für das Spaltmaß, welchem im Rahmen einer visuellen Kontrolle eine von drei möglichen Spaltklassen zugeordnet wurde, ergibt sich ein Korrelationskoeffizient von $r = 0,956$ mit maximaler Abweichung der Prognose von einer Klasse. Auch für den visuell klassifizierten Sonotrodenabdruck mit Wertigkeiten von 0 bis 3 ergibt sich bei einem maximalen

Prognosefehler von einer Klasse eine Korrelation von $r = 0,967$. Das Auftreten von Spalten und Sonotrodenabdrücken ist stark von der Prozesseinstellung abhängig, so dass die GRNFN-Modelle auch für diese subjektiv beurteilten Klassenmerkmale aufgrund der guten Trennung der verschiedenen Einstellungen sehr hohe Prognosegüten erreichen.

6.4 Zusammenfassung und Fazit der Anwendungen

In diesem Hauptkapitel wurden unterschiedliche Anwendungen der vorgestellten Methoden und Verfahren am Probekörper sowie an Praxisteilen beschrieben. Dabei wurde eine breite Palette möglicher Praxisprobleme und auftretender Störungen über die ganze Prozesskette hinweg versuchstechnisch abgebildet. Zudem wurde eine Zusatzsensorik zur Erfassung von Körperschall und Temperatur auf ihre Eignung als zusätzliche Datenquelle untersucht. Ausgehend von „Erfahrungen“ aus der Praxis wurde vom Körperschall ein hoher Informationsgehalt in Bezug auf die Schweißnahtqualität erwartet. Diese Erwartungen haben sich jedoch nicht erfüllt. Durch die optische Erfassung des Temperaturverlaufs in der Schweißnaht war es indes möglich, Rückschlüsse auf einen geänderten Feuchtegehalt in der Schweißzone zu ziehen.

Im Rahmen mehrerer Versuchsreihen mit dem Probekörper wurden die Fähigkeiten von GRNFN-Modellen aufgezeigt, neben parameterbedingten Veränderungen im Schweißprozess auch die Schweißnahtqualität überwachen zu können. Auch war es möglich, geänderte Herstellbedingungen der zu verschweißenden Werkstücke im Spritzgießen, eine Vertauschung des Werkstoffs sowie variierende Lagerbedingungen, die sich in der Aufnahme von Feuchtigkeit auswirken, mittels GRNFN-Modellen zu überwachen. Dabei wurden Probekörper sowohl mit unterschiedlicher Nahtgeometrie (ERG- und Quetschnaht) als auch aus verschiedenen Materialien verwendet. Auch beim Training der GRNFN-Modelle einzig anhand von Daten aus dem Arbeitspunkt des Schweißprozesses – d. h. Störungen der Vorprozesse wurden nicht berücksichtigt – war es dennoch für alle genannten Fälle möglich, die Störungen in der vorgelagerten Prozesskette durch GRNFN-Modelle zu erkennen. Anschließend wurden alle vorgestellten Überwachungsmöglichkeiten anhand verschiedener Praxisbauteile untersucht. Auch dabei konnte die Eignung von GRNFN-Modellen für die Erfüllung aller Anforderungen nachgewiesen werden.

Mit GRNFN-Modellen können somit zuverlässige Prognosen von Qualitäts- und Prozessmerkmalen erzielt werden – sowohl für kontinuierliche Prüfmerkmale als auch für Sichtprüfmerkmale mit Klassenzuordnung. Die Qualitätsprognose von GRNFN-Modellen wurde mit MLP-Modellen, wie sie in NEPRES und anderen Systemen seit Jahren erfolgreich zum Einsatz kommen, verglichen. Der exemplarische Vergleich des Systemansatzes mit MLP hat deutlich die Vorteile des hier favorisierten GRNFN-Einsatzes aufgezeigt.

7 Zusammenfassung und Ausblick

Um vor dem Hintergrund zunehmender Globalisierung mit hochwertigen Produkten weiter am Markt bestehen zu können, steigt der Bedarf an leistungsfähigen Systemen zur Prozess- und Qualitätsüberwachung stetig. Die vorliegende Arbeit soll hierzu einen Beitrag liefern.

7.1 Zusammenfassung

Die Kunststoffverarbeitung ist ein sehr komplexer Produktionszweig und die Anzahl der aus Kunststoff gefertigten Produkte nimmt ständig zu. Ein häufiges Fügeverfahren für Kunststoffe ist das Ultraschallschweißen. Die kurzen Zykluszeiten, die preiswerte Maschinenteknik sowie die gute Integrierbarkeit machen das Verfahren auch für automatisierte Fertigungslinien sehr interessant. Da das Ultraschallschweißen vorwiegend am Ende der Prozesskette stattfindet, sind mit Ausschussteilen hohe Wertschöpfungsverluste und Kosten verbunden. Die Ursachen für Schweißfehler können dabei auch in anderen, vorgelagerten Prozessschritten liegen, da sich diese über Änderungen der Werkstoffeigenschaften teils signifikant auf den Schweißprozess auswirken.

Moderne Produktionsanlagen verfügen zwar häufig über Steuerungen, die viele Einzelgrößen auf die Verletzung von Bereichsgrenzen überwachen können. Dabei gestaltet sich jedoch das Festlegen der Toleranzgrenzen als schwierig, insbesondere da meist ein direkter Zusammenhang zu den relevanten Qualitätsmerkmalen fehlt. Auch optische Inspektionssysteme (industrielle Bildverarbeitung) können nicht alle Qualitätskriterien überwachen – dies gilt insbesondere für Festigkeitswerte. Folglich werden andere Systeme, z. B. modellbasierte Prozess- und Qualitätsüberwachungssysteme, benötigt, die diese Defizite kompensieren können.

Im Rahmen dieser Arbeit wird als Lösung eine modellbasierte Prozessüberwachung am Beispiel des Ultraschallschweißens von Kunststoffteilen unter Berücksichtigung der Einflüsse der Vorprozesse dargestellt. Die Prozessüberwachung erfolgt in heutigen Ultraschallschweißanlagen bestenfalls über Toleranzen für einfache Maschinengrößen wie Schweißzeit oder Schweißenergie. Diese weisen jedoch meist keine hinreichende Korrelation mit der Schweißnahtqualität auf. Eine effektive Prozess- und Qualitätsüberwachung fehlt somit bis dato. Die Zielsetzung dieser Arbeit war der Aufbau einer Prozessüberwachung, die sowohl eine Gut-Schlecht-Erkennung bis hin zur Prognose von Qualitätswerten für jedes geschweißte

Werkstück als auch die Erkennung von Störungen in den Vorprozessen ermöglicht. Der Lösungsansatz basiert im Wesentlichen auf der am IPA für das Spritzgießen entwickelten NEPRES-Systemtechnologie (siehe Kapitel 1). Diese wurde auf das Ultraschallschweißen übertragen und so erweitert, dass die Prozessmodelle Informationen aus dem Schweißprozess als Eingangsdaten erhalten und daraus Aussagen über die Schweißnahtqualität sowie den Zustand der Vorprozesse möglich sind. Der Lösungsansatz ist in Abbildung 1.5 dargestellt. Das Vorgehen stützt sich maßgeblich auf systematisch geplante und durchgeführte Versuchsreihen. Zuerst wurde ein spezieller Probekörper definiert, der verschiedene Prüfungen zur Ermittlungen der relevanten Qualitätsmerkmale ermöglicht (Abschnitt 3.3). Die Probekörper wurden durch einen überwachten Spritzgießprozess gefertigt und anschließend auf einer Ultraschallschweißmaschine geschweißt. Diese war durch die Integration der notwendigen Messtechnik zur Erfassung der Prozesssignale zur Versuchsanlage aufgerüstet worden (Abschnitte 3.1 und 3.2). Zur Analyse der verschiedenen Einflüsse wurden schließlich in den Versuchen gezielt relevante Parameter entlang der gesamten Prozesskette vom Spritzgießen, über die Lagerung bis zum Schweißen variiert. Die Prozesssignale beim Schweißen wurden aufgezeichnet und anschließend analysiert und ausgewertet.

In Kapitel 2 wurde der Stand der Technik umfassend untersucht. Dabei standen der Prozess des Ultraschallschweißens sowie die Prozessüberwachung im Fokus. Neben der Maschinenteknik wurden insbesondere die wesentlichen Einflussgrößen dargestellt (Werkstoff, Konstruktion, Herstellbedingungen, Schweißparameter). Vor allem die Materialkennwerte, z. B. Wärmekapazität oder Schmelzpunkt, unterscheiden sich bei den verschiedenen Kunststoffarten und sind zudem oft abhängig von der Temperatur und dem Feuchtegehalt. Insbesondere der mechanische Verlustfaktor $\tan \delta$ beeinflusst die Umsetzung der Schwingungen in Wärme und damit den Schweißablauf. Auch die konstruktive Gestaltung der Werkstücke, die Schweißnahtgeometrie sowie die Herstell- und Lagerbedingungen beeinflussen maßgeblich die Stabilität eines Schweißprozesses. Erfolgt ihre Herstellung durch Spritzgießen, können sich Produktionsschwankungen in Maßabweichungen oder veränderten inneren Eigenschaften (Kristallisationsgrad, Vakuolen, Spannungen, Feuchtegehalt, ...) abbilden und sich so auf den Schweißvorgang auswirken. Für diesen sind die eingestellten Schweißparameter von entscheidender Wichtigkeit, vor allem die Schweißkraft sowie die Amplitude. Die Einstellung des Schweißprozesses erfolgt vorwiegend durch erfahrene Einsteller im Trial-and-Error-Verfahren. Die Vielzahl der Einflussgrößen auf den Schweißprozess sind in Abbildung 2.5 zusammengefasst.

Die Prozessüberwachung ist eng mit dem Qualitätsmanagement verknüpft. Im Wesentlichen ist zwischen einer statistischen und einer kontinuierlichen Kontrolle zu unterscheiden. Bei der statistischen Prozessüberwachung (SPC) werden Stichproben entnommen und ausgewertet, wobei insbesondere Regelkarten eine wichtige Methode darstellen. Singuläre Ereignisse und kurzfristige Störungen können mit der SPC jedoch nicht sicher erkannt werden.

Die Anwendung der SPC auf Prozessparameter – so häufig in der Kunststoffverarbeitung – ist als fraglich anzusehen, da meist kein direkter Zusammenhang zur Qualität vorliegt. Bei der kontinuierlichen Prozessüberwachung (CPC) wird jeder Fertigungszyklus anhand von Sensorsignalen (z. B. mittels Hüllkurven) oder mittels Prozessgrößen (Schweißweg, maximale Drücke, Zeiten, ...) und zugeordneten Grenzwerten bewertet. Wie bei der SPC ist dabei jedoch der Zusammenhang mit der Teilequalität oft fragwürdig. Deshalb ist für eine kontinuierliche Qualitätsüberwachung entweder eine explizite Qualitätsprüfung jedes Teils – manuell oder durch integrierte Prüfsysteme wie beispielsweise Bildverarbeitung – oder der Einsatz von Systemen mit entsprechenden Prozessmodellen notwendig (Abschnitt 2.2).

Die bisherige Prozessüberwachung beim Ultraschallschweißen beruht vor allem auf den Arbeiten von NETZE und KORTE, die beide beim DVS-Probekörper ansetzen. NETZE identifiziert den Verlauf des Schweißwegs als wesentliche Prozessgröße und teilt ihn in verschiedene Phasen ein. Der Eintritt in die dritte Phase mit stationärem Fließverhalten in der Schweißnaht geht laut NETZE mit dem Erreichen einer Mindestfestigkeit der Schweißnaht einher. Für die Überwachung wird eine visuelle Beurteilung des Schweißwegverlaufs vorgeschlagen. KORTE untersucht die Auswirkungen eines veränderten Spritzgießprozesses auf den Schweißprozess und weist signifikante Einflüsse nach. Um Zusammenhänge zwischen den Größen und den stark schwankenden Festigkeitswerten zu belegen, wendet er statistische Auswertungen an. Eine Modellierung der Festigkeit erfolgte bei KORTE mit linearen Regressionsmodellen anhand von Werten, die aus Signalen des Ultraschallschweißprozesses abgeleitet werden. Durch zusätzliche Informationen aus dem Spritzgießprozess konnte ein Bestimmtheitsmaß von knapp 90 % für die Modellgüte erreicht werden. Andere Autoren wenden mehrfach die Methode der statistischen Versuchsplanung erfolgreich an, um die Einstellung des Schweißprozesses hinsichtlich der Festigkeit zu optimieren. Die Kenngrößen der Wirkleistung und der eingebrachten Energie erweisen sich entgegen den Aussagen z. B. von Arbeiten der 80er Jahre als nicht sensibel genug für eine Qualitätsüberwachung. Im Rahmen eines AiF-Projektes wurde die Anwendung von Neuronalen Netzen zur Prognose der Festigkeit beim Ultraschallschweißen von Metallteilen untersucht. Demnach ist eine Prognose prinzipiell möglich und die Neuronalen Netze sind in der Lage, einige Fehlerarten sowie das Erreichen einer Mindestfestigkeit zu erkennen. Die Modelle müssen dabei für jeden Prozess separat angepasst werden (Abschnitt 2.3).

Im Rahmen eines Überblicks über verschiedene Modellierungsverfahren (Abschnitt 2.4) wurde der Fokus auf datenbasierte Modelle gelegt, d. h. solche Modelle, die anhand von Beispieldaten erstellt werden. Eine große Klasse bilden dabei die Neuronalen Netze (NN), von denen es verschiedenste Ausprägungen gibt: RBF-Netze basieren auf radialsymmetrischen Funktionen, deren Zentren an den Trainingsvektoren liegen. Das mehrschichtige Perzeptron (MLP) ist für die Prozessmodellierung der wohl am häufigsten verwendete Modelltyp. Problematisch sind bei ihm vor allem die Festlegung der richtigen Struktur und Lernpa-

parameter, der vom Initialisierungszustand abhängige Trainingsverlauf sowie die Gefahr des „Auswendiglernens“. Auch LVQ und SOM sind wichtige Modellierungsverfahren, bei denen Stützvektoren und Abstandsinformationen zwischen Vektoren eine wichtige Rolle spielen. Fuzzy-Modelle ermöglichen die Abbildung von Expertenwissen in Form einer Regelbasis und werden – im Gegensatz zu Neuronalen Netzen – nicht anhand von Beispieldaten trainiert. Abschließend erfolgte im Überblick zur Modellierung die Darstellung einiger Ansätze, die eine Verknüpfung der Vorteile von Neuronalen Netzen und Fuzzy-Methoden versuchen (Abschnitt 2.4.4): ANFIS, das u. a. auch beim Spritzgießen eingesetzt worden ist, und insbesondere GRNFN, die im Wesentlichen Elemente von RBF-Netzen und Fuzzy-Systemen kombinieren. Aufgrund ihrer Eigenschaften, die in Kapitel 4 näher untersucht wurden, sind GRNFN ein sehr geeignetes Modellierungsverfahren zur Umsetzung des gewählten Lösungsansatzes.

Die prinzipielle Eignung des Ansatzes belegen die in Abschnitt 2.5 genannten Beispiele der Prozessüberwachung in verschiedensten Bereichen. Allen Anwendungsfällen ist gemeinsam, dass Prozessspezifika in Kenngrößen übertragen und über verschiedene Modellierungsmethoden ein Zusammenhang zwischen Prozessdaten und Qualitäts- oder Fehlergrößen abgebildet wird.

Abschließend wurden die Elemente der NEPRES-Systemtechnik, die als Basis für die im Rahmen dieser Arbeit entwickelten Methoden und Verfahren dient, kurz vorgestellt. NEPRES umfasst neben Möglichkeiten der Messdatenakquisition umfangreiche Verarbeitungs- und Modellierungsmethoden und ist über eigene Module beliebig erweiterbar (Abschnitt 2.6).

Basierend auf dem Stand der Technik wurde in Kapitel 3 der Lösungsansatz detailliert dargestellt. Neben den erfassten Maschinensignalen wie Schweißweg, Leistung, Schwingungsamplitude und Fügekraft wurden weitere Sensoren als mögliche Informationsquellen betrachtet. Anschließend wurden die verwendete Versuchsanlage mit gekoppelter Messwert-erfassung und -verarbeitung sowie der definierte Probekörper für die systematischen Versuchsreihen beschrieben. Der Probekörper ermöglichte eine detaillierte Untersuchung über die gesamte Prozesskette hinweg – vom Spritzgießen bis zum Schweißen. Seine Konstruktion reduziert einige versuchstechnische Nachteile existierender Probekörpergeometrien und ermöglicht vielfältige Prüfverfahren zur Qualifizierung der Schweißung. Es folgte ein Überblick über die geplanten Versuche zur Analyse der verschiedenen Einflüsse entlang der Prozesskette. Auch wurden das Vorgehen zur Signalanalyse und Kenngrößenbildung skizziert und wesentliche Anforderungen an die Modellbildung herausgearbeitet. Abschließend wurde eine exemplarische Versuchsreihe vorgestellt, die in den anschließenden Kapiteln mehrfach zur Darstellung, Erläuterung und Verifikation der entwickelten Methoden diente (Abschnitt 3.8).

Wesentlich für jede Modellierung ist – wie in Kapitel 4 eingangs dargelegt – der Zustandsvektor, der den Prozessmodellen als Eingangsvektor dient. Seine Komponenten werden aus den Prozesssignalen abgeleitet und sollten einen Fertigungszyklus sowie die Veränderungen im Prozess abbilden. Zyklen, die mit gleicher Prozesseinstellung gefertigt wurden, liefern sehr ähnliche Signalverläufe und damit ähnlichen Kenngrößen. Dies ist äquivalent mit ähnlichen Zustandsvektoren, d. h. Vektoren mit geringem Abstand entsprechend der gewählten Metrik, die in der Folge so genannte Cluster bilden. Eine Visualisierung und Beurteilung der Cluster wird durch das Sammons-Mapping ermöglicht (Abschnitt 4.1.2). Dabei handelt es sich um eine Methode zur Abbildung des hochdimensionalen Zustandsraums in die Ebene. Dies erfolgt derart, dass das Verhältnis der Abstände zwischen den Vektoren möglichst gut beibehalten wird und somit Cluster im Zustandsraum auch im Bildraum wieder Cluster formen. Die Zusammenhänge der abgebildeten Zustandsvektoren mit den Zielgrößen können dann mittels Farbcodierung der Zielwerte dargestellt und analysiert werden.

Die verwendeten Modelle mussten im Wesentlichen zwei Anforderungen erfüllen: Zum einen die Schweißnahtqualität prognostizieren und zum anderen „unbekannte Zustände“ erkennen, die aus Störungen oder Veränderungen der Vorprozesse resultieren. Beide Anforderungen konnten mittels der systemseitig realisierten GRNFN erfüllt werden. In Kapitel 4 wurden die Eigenschaften von GRNFN detailliert untersucht und notwendige Weiterentwicklungen zur Erfüllung der Zielsetzung erarbeitet. GRNFN basieren wie RBF-Netze auf Stützvektoren, die aus den Trainingsdaten abgeleitet werden, und haben ähnliche Interpolationseigenschaften. Die Interpolation ist jedoch unabhängig von der Dimension des Zustandsvektors und basiert auf eindimensionalen Zugehörigkeits- oder Aktivierungsfunktionen. Diese bilden den Fuzzy-Anteil von GRNFN-Modellen und dienen zur Abbildung aller Komponenten des Zustandsvektors am Modelleingang. Daraus werden Aktivitätswerte abgeleitet und es wird durch Verknüpfung mit synaptischen Gewichten, die während des Trainings festgelegt wurden, der Prognosewert als Netzausgabe ermittelt. Da GRNFN-Modelle als Fuzzy-Regelbasis interpretiert werden können, ermöglichen sie eine Analyse der Zusammenhänge sowie die Identifikation wesentlicher Einflussfaktoren. Des Weiteren können so auftretende Prognosefehler auf ihre Ursachen hin untersucht werden (vgl. Abschnitt 4.4ff).

In Abhängigkeit der Parameter der Aktivierungsfunktion wurde die Modellgüte analysiert, so dass daraus sinnvolle Werte für die Parameter abgeleitet werden können. GRNFN erwiesen sich als robust gegenüber dem Weglassen von Eingangskomponenten. Auch für unvollständige Zustandsvektoren wurden hohe Prognosegüten erzielt. Daraus ergibt sich eine gute Fehlertoleranz der GRNFN und der Ausfall von Sensoren kann kompensiert werden, was das Verfahren für die Praxis sehr interessant macht. Die Verwendung von Dreiecks- und Trapezfunktion als Aktivierungsfunktion ermöglicht es, vorhandene Fuzzy-Regeln abzubilden und in das Modell zu integrieren. Ein anhand von Trainingsdaten generiertes

GRNFN-Modell kann so gezielt um vorhandenes Wissen erweitert werden. Somit ist eine effektive Kombination von daten- und wissensbasierter Modellbildung möglich. Ein weiteres Vorteil von GRNFN-Modellen ist die dynamische Modellanpassung. Es ist nachträglich möglich, alte Trainingsdaten zu entfernen und neue hinzuzufügen. Damit können Prozesse, die langsamen Trends oder Driften unterliegen, abgebildet werden.

Durch erarbeitete Indikatoren konnten anhand von GRNFN-Modellen auch Zustände, die sich von den Trainingsdaten deutlich unterscheiden, als unbekannt klassifiziert werden. Durch nachträgliches Hinzufügen neuer Trainingsdaten kann auch ein so genanntes *active learning* vergleichsweise einfach realisiert werden (Abschnitt 4.6).

Vor der Aufstellung eines Prozessmodells waren die Zustandsvektoren zu erarbeiten. Die notwendigen Schritte der Signalverarbeitung zur Ableitung von Kenngrößen sind in Kapitel 5 am Beispiel des Ultraschallschweißens beschrieben. In Abschnitt 5.3 wurde dargestellt, wie Zustandsvektoren optimiert und so signifikant bessere Modelle erzielt werden können.

Aus den visualisierten Signalen eines Schweißzyklus können notwendige Vorverarbeitungsfunktionen wie Filter abgeleitet werden. Anschließend können über verschiedene Arten der Darstellung die aufgetretenen Signalschwankungen analysiert und signifikante Kenngrößen zu ihrer Beschreibung erarbeitet werden. Allerdings müssen diese Kenngrößen eigens für jeden Fertigungsprozess (z. B. Spritzgießen, Ultraschallschweißen, ...) erarbeitet und implementiert werden. Gleichwohl gibt es prozessunabhängige Kenngrößen, die relativ allgemein angewendet werden können, z. B. Integrale und Maxima. Eine weitere Möglichkeit besteht in der Anwendung der PCA zur Codierung der Signalform und Abbildung der Schwankungen – dabei ist jedoch eine feste Signallänge Voraussetzung. Auch in der Statistik existieren Kenngrößen, welche die Lage und Form einer Verteilung (Schiefe, Wölbung, ...) grob beschreiben können. Ziel war, die Prozessschwankungen durch all diese Prozesskenngrößen möglichst umfassend zu beschreiben und anschließend durch ein Modell mit den für die Trainingsdaten ermittelten Qualitätsschwankungen in Beziehung zu setzen. Es wurde insgesamt ein Satz von knapp 130 Kenngrößen erarbeitet, in Algorithmen zur Berechnung überführt und in NEPRES integriert. Die so für die Versuche berechneten Kenngrößen bildeten die Basis für alle weiteren Analysen und Anwendungen.

Mit einer Streuungsanalyse (Abschnitt 5.2) können relevante Prozesskenngrößen identifiziert werden. Auch sind verschiedene Möglichkeiten der Visualisierung erarbeitet und umgesetzt worden, um Zusammenhänge einfach erkennen zu können. Über so genannte Boxplots kann ein Überblick über die Lage und Streuung von Kenngrößen gewonnen werden. Bei Anwendung auf Qualitätsmerkmale führt dies schnell zu Aussagen über die Prozessstabilität.

Mittels des Sammons-Mapping (Abschnitt 4.1.2) kann die Clusterbildung der Zustandsvektoren analysiert werden. Da die Wertebereiche der einzelnen Komponenten jedoch stark unterschiedlich sind und über die implizite Abstandsberechnung die Abbildung beeinflusst

sen, ist die Normierung aller Kenngrößen wesentlich für jede Optimierung. Die normierten Kenngrößenvektoren führten für das Ultraschallschweißen in einigen Fällen schon zu einer ausreichenden Auftrennung der Cluster und zu guten Prognosemodellen. Eine weitere Optimierung kann durch eine gezielte Skalierung der einzelnen Kenngrößen erfolgen. Jede geänderte Zusammenstellung der Zustandsvektoren führt zu anderen Abstandsmaßen und somit einer geänderten Clusterbildung und Modellgüte. Eine Merkmalsselektion stellt eine weitere Optimierungsmöglichkeit dar. Mittels einer einfachen Korrelationsanalyse können lineare Zusammenhänge zwischen zwei Größen erkannt werden – für die Modellbildung ist der Zusammenhang Prozesskenngröße zur Qualität relevant. Da jedoch bereits wenige Ausreißer einen starken Einfluss haben und die Korrelation einer schwankenden Prozessgröße mit einer konstanten Zielgröße (z. B. eingestellte Schweißkraft) nicht aussagekräftig ist, wurde eine Methode zur Kenngrößenselektion erarbeitet, die eine Signifikanzanalyse auf der Basis statistischer Mittelwertvergleiche ermöglicht (Abschnitt 5.3.3). Die so erarbeitete Signifikanzanalyse ist robust gegenüber Ausreißern und verringert den Einfluss von Rauschanteilen.

Als weitere Optimierung wurden Genetische Algorithmen (GA) als Suchstrategie untersucht (Abschnitt 5.3.4). Inspiriert durch die Evolution bilden GA den Selektionsprozess nach dem Motto „survival of the fittest“ ab. Wesentlich beim Einsatz von GA ist die Definition eines Gütekriteriums. Im Rahmen dieser Arbeit wurde dafür ein Überlappungsmaß verwendet, welches die Trennung und Homogenität der durch die Zustandsvektoren gebildeten Cluster beschreibt. Im Gegensatz zur Signifikanzanalyse funktioniert das Überlappungsmaß auch für Merkmale mit mehr als zwei Klassen ohne weitere Voting-Operationen.

In Kapitel 6 wurde die Leistungsfähigkeit der entwickelten Methoden im Hinblick auf die Zielsetzung der Arbeit an Probekörpern und an Praxisbauteilen untersucht. Insgesamt wurde eine breite Palette möglicher Praxisprobleme und auftretender Störungen über die ganze Prozesskette hinweg versuchstechnisch abgebildet. Von Schwingungssignalen, die mittels Körperschallsensoren erfasst werden, wurde ein hoher Informationsgehalt vermutet. Schon erste Versuche zeigten starke Varianzen im aufgezeichneten Körperschallsignal – unabhängig von der erzielten Schweißnahtqualität. Trotz umfangreicher Untersuchungen konnte kein Zusammenhang zwischen Körperschall (Zeitsignal oder Amplitudenspektrum) und Schweißnahtfestigkeit gefunden werden.

Anhand des Probekörpers wurde aufgezeigt, dass es mit einer Modellbildung möglich ist, eine Schweißung verschiedenen Prozesseinstellungen der Schweißmaschine zuzuordnen und diese somit zu unterscheiden. Dies kann insbesondere der Erkennung von Maschinenstörungen oder -defekten dienen, die sich auf den Schweißablauf auswirken, durch die Steuerung aber nicht erkannt werden (z. B. ein verstopfter Druckluftfilter). Ein wesentliches Element zur Überwachung des Schweißprozesses ist die Prognose der Schweißnahtqualität. Für das Training von Modellen zur Prognose eines Qualitätsmerkmals ist es wichtig, dass zuvor

die Zielgröße in den Trainingsdaten sicher und mit hoher Güte ermittelt wurde – dies gilt insbesondere für die Festigkeitsprüfung. Am Probekörper wurde die Prognose der Schweißnahtfestigkeit mittels GRNFN untersucht. Als Ergebnis ist festzuhalten, dass mit GRNFN-Modellen eine Prognose der Festigkeit in hoher Güte möglich war (Abschnitt 6.2.2). Insgesamt wurde ein Korrelationskoeffizient von 0,96 erzielt. Die Prognoseergebnisse wurden anschließend mit den Ergebnissen mehrerer MLP-Modelle, wie sie bei der Prozessüberwachung in vielen Arbeiten und auch in NEPRES eingesetzt werden, verglichen. Das GRNFN-Modell lieferte in allen Bereichen gleich gute und meist sogar bessere Resultate als die MLP-Modelle.

Für den Probekörper konnten entlang der gesamten Prozesskette weitere Einflüsse auf den Schweißprozess nachgewiesen und durch GRNFN-Modelle als Veränderung im Schweißprozess erkannt werden. Veränderungen im Spritzgießen konnten durch die erarbeiteten Methoden für verschiedene Einstellungen der Schweißmaschine zuverlässig erkannt und separiert werden (Abschnitt 6.2.3). Eine Änderung des Werkstoffs wirkt sich insbesondere durch die Einflüsse wichtiger Werkstoffkenngrößen aus. Schweißzyklen mit geändertem Material konnten mittels GRNFN-Modellen zuverlässig als Unbekannt und somit als Störung der Vorprozesse erkannt werden (Abschnitt 6.2.4). In der Maschinensteuerung vorhandene Größen wie Schweißzeit oder Schweißweg wiesen bei einer statistischen Auswertung zwar Unterschiede zwischen den Materialien auf. Die Streubereiche überlappten jedoch so stark, dass für einzelne Zyklen eine sichere Zuordnung nicht möglich war. Eine weitere wichtige Einflussgröße ist der Feuchtegehalt der Werkstücke. Mittels GRNFN-Modellen war es möglich, allein anhand der Schweißsignale Änderungen des Feuchtegehalts zu erkennen (Abschnitt 6.2.5).

Während mittels des Probekörpers die gesamte Prozesskette an einem Werkstück durchgehend untersucht werden konnte, sind Praxisbauteile den herrschenden Randbedingungen unterworfen. So können zugelieferte Spritzgießteile nicht in der Herstellung variiert werden, bei manchen Bauteilen sind aussagefähige Qualitätsmerkmale nicht definiert oder können nicht geprüft werden. Die Untersuchungen in Abschnitt 6.3 mussten sich deshalb an den Möglichkeiten der verfügbaren Bauteile orientieren. Losgelöst davon bestätigte sich die Unabhängigkeit der Methoden vom konkreten Anwendungsfall und somit ihre Übertragbarkeit.

Bei einer Automobilleuchte wurden leichte Gewichtsunterschiede zwischen den Gehäuseteilen festgestellt, die nur auf eine Veränderung im Spritzgießprozess zurückzuführen waren. Ein GRNFN-Modell, das ausschließlich anhand der Schweißzyklen einer Gewichtsklasse erstellt wurde, war in der Lage, die Schweißungen mit Teilen der anderen Gewichtsklasse größtenteils als unbekannt zu klassifizieren (Abschnitt 6.3.1). Für ein Rückschlagventil wurde der Fall eines veränderten Grundmaterials untersucht. Auch hier war es möglich, die beiden Materialtypen nur auf der Basis von Zustandsvektoren aus dem Schweißprozess großteils zu separieren. Eine vorherige Kenntnis möglicher Störungen und die Abbildung

der Störung für das Training waren dabei nicht notwendig. Für das Rückschlagventil wurde auch der Einfluss eines Feuchtegehalts durch Lagerung untersucht und konnte nach einer Optimierung der Zustandsvektoren mit GA separiert werden (Abschnitt 6.3.2).

Für ein weiteres Bauteil, ein Taster, waren keine quantifizierbaren Qualitätsmerkmale verfügbar. Im Rahmen einer Signalaufzeichnung parallel zur Fertigung wurde festgestellt, dass der Prozess einer Veränderung unterlag, die ein Nachführen des Prozesses seitens des Maschinenführers notwendig machte. Dies erfolgte jedoch mit einer deutlichen Verzögerung, da die Auswirkungen durch eine nachgelagerte Qualitätssicherungsmaßnahme der SPC erst später erkannt wurden. Eine anschließende Analyse brachte den Nachweis, dass ein anhand des Arbeitspunkts erstelltes Prozessmodell die Prozessveränderung unmittelbar und früh erkannt hätte (Abschnitt 6.3.3).

Abschließend wurde der Schweißprozess sowie die Prognose der erzielten Nahtqualität an einem Wasserverteiler untersucht. Ein möglicher Montagefehler lag dabei im Nicht-Einlegen eines Filterelements. Da dieses jedoch physikalisch keine Rückwirkung auf den Schweißprozess hat, konnten (richtigerweise) fehlende Einlegeteile trotz gezielter Optimierung der Zustandsvektoren nicht erkannt werden. Die Optimierungsmethoden erwiesen sich demnach als entsprechend robust und tendierten nicht dazu, Scheinzusammenhänge zu generieren. Verschiedene Parametereinstellungen des Schweißprozesses konnten durch das Sammons-Mapping gut getrennt und durch die GRNFN-Modelle sicher zugeordnet werden. Auch die Prognose des Berstdrucks war mit relativ hoher Genauigkeit möglich (Abschnitte 6.3.4 und 6.3.5). Wie an den Streuungen der Messwerte zu erkennen ist, war der Prozess jedoch in einigen Einstellungen instabil; auch die Prüfung lieferte einige deutliche Ausreißer. In diesen Bereichen waren folglich auch entsprechende Prognosefehler zu finden. Insgesamt ist jedoch festzustellen, dass die Prognose der Nahtqualität anhand der Schweißsignale auch für Praxisbauteile funktioniert.

7.2 Fazit und Ausblick

Das in der Arbeit vorgestellte modellbasierte Überwachungssystem ermöglicht es, sowohl Qualitätsprognosen aus Prozesssignalen abzuleiten als auch Störungen der Vorprozesse zu erkennen. Die Konzepte der Zustandsvektoren und der Modellbildung sind allgemein gültig und können in verschiedensten Anwendungsfeldern eingesetzt werden. Kernelement ist die auf *general regression neuro-fuzzy networks* (GRNFN) basierende Modellbildung. Eine prozessspezifische Signalanalyse und Kenngrößenbildung sind für die Entwicklung von Zustandsvektoren sehr wichtig. Die diskutierte Optimierung der Zustandsvektoren führt zu einer signifikanten Erhöhung der erzielbaren Modellgüte.

Für das exemplarisch untersuchte Ultraschallschweißen von Kunststoffteilen konnte für den Probekörper zum einen die Nahtfestigkeit mit hoher Genauigkeit prognostiziert werden.

Weiter wurde der Nachweis erbracht, dass sich sowohl Änderungen im vorgelagerten Spritzgießprozess als auch bei der Lagerung signifikant auf den Schweißprozess auswirken und durch die erarbeiteten GRNFN-Modelle erkannt werden können. Eine Überwachung der Vorprozesse auf Störungen ist somit einzig anhand der Signale des Schweißprozesses möglich. Im direkten Vergleich mit MLP-Modellen (mehrschichtiges Perzeptron), wie sie bisher meist zur Prozessmodellierung verwendet werden, sind mit GRNFN-Modellen bessere Modellgüten erzielt worden. Neben einem einfacheren und deterministischen Training sowie den Möglichkeiten der Unbekannt-Klassifikation und dem nachträglichen Hinzufügen von Lerndaten weisen GRNFN-Modelle gegenüber MLP noch weitere Vorteile wie die Fähigkeit zur Integration von Erfahrungswissen und die Toleranz gegenüber unvollständigen Eingangsdaten auf. Insgesamt bieten GRNFN-Modelle deutlich mehr Möglichkeiten als MLP-Modelle bei gleichzeitig höherer Prognosegüte. Bestätigt wurden die Ergebnisse am Probekörper durch die erfolgreiche Verifizierung der Methoden auch an mehreren Schweißprozessen mit Praxisbauteilen. Somit wurde insgesamt das angestrebte Ziel der Arbeit erreicht.

Mit den realisierten Methoden steht ein Werkzeug zur effizienten qualitätsbasierten Überwachung von Ultraschallschweißprozessen zur Verfügung. Die umfassenden Versuche und Anwendungen zeigten jedoch, dass unterschiedliche Weiterentwicklungen zukünftige Systemeinsätze positiv beeinflussen könnten, auf die im Folgenden kurz eingegangen wird. Hinsichtlich der Methoden ist eine Weiterentwicklung für die Unbekannt-Klassifikation sowie die Reduktion von Prognosefehlern denkbar. Für die Indikatoren zur Unbekannt-Klassifikation sollten Strategien und Kenngrößen zur automatischen Ableitung von Schwellwerten erarbeitet werden. Auch fehlen effektive Möglichkeiten wie Voting-Ansätze zur Verknüpfung der Unbekannt-Indikatoren. Bei der Modellbildung ist die automatische Festlegung optimaler Parameter der Aktivierungsfunktion in Abhängigkeit von den Trainingsdaten ein möglicher Ansatz. Der Einsatz von verschiedenen Aktivierungsfunktionen für einzelne Komponenten oder Trainingsdaten sowie die individuelle Festlegung ihrer Parameter könnte näher untersucht werden. Durch eine Analyse der Kenngrößenverteilung sowie der zugeordneten Qualitätswerte könnte bei ähnlichen Kenngrößenwerten und unterschiedlichen Zielgrößen mittels geeigneter Parameter eine niedrigere Aktivierung forciert und damit die Beeinflussung der Prognose sowie der resultierende Prognosefehler reduziert werden.

Im Hinblick auf das hier untersuchte Ultraschallschweißen gibt es noch weitere Handlungsfelder. Im Rahmen dieser Arbeit wurde der Einfluss der Vorprozesse nachgewiesen. Es wurde jedoch nicht untersucht, ob und wie sich die verschiedenen Einflüsse unterscheiden lassen. Die Unbekannt-Klassifikation sollte um eine Möglichkeit erweitert werden, anhand der Indikatoren eine Aussage über die Ursache ableiten zu können. Da sich insbesondere Veränderungen von wesentlichen Materialkenngrößen auswirken, ist ein möglicher Ansatzpunkt die Verknüpfung mit existierenden Materialdatenbanken. Für die Kenngrößen beim Ultra-

schallschweißen könnten Zusammenhänge zu Werkstoffkennwerten erarbeitet und daraus bei Kenngrößenänderungen Rückschlüsse auf mögliche Ursachen abgeleitet werden.

Die positiven Ergebnisse beim Ultraschallschweißen legen es nahe, die realisierte Systemtechnik auf andere Fertigungsbereiche zu übertragen (z. B. Spritzgießen, Umformprozesse oder Schweißprozesse für Metall). Wesentliche Aufgabe ist dabei jeweils die Identifikation von relevanten Einflussgrößen sowie die Erarbeitung von aussagekräftigen Prozesskenngrößen. Dabei wird auch interessant sein, inwieweit sich auch in anderen Anwendungsbereichen die Einflüsse von Prozessen, die in der Prozesskette weiter vorne angesiedelt sind, auswirken und wie diese erkannt werden können.

Im Zuge eines umfassendes Qualitätsmanagements sowie einer Neugestaltung der Qualitätsüberwachung könnten zudem jede Stufe in einer Prozesskette mittels eigener Modelle überwacht und Informationen daraus in nachgelagerten Prozessen verwendet werden. Damit wäre langfristig auch eine Nachführung von Folgeprozessen auf der Basis von Informationen aus den Vorprozessen denkbar.

8 Summary

Since requirements to pieces and their complexity constantly increase, for production often a process chain with several stages is necessary. The achievable level of quality of every process depends also on the earlier stages of the process chain. As a result, the necessary process know-how increases and the process windows are getting closer. Due to material substitution, lightweight design and miniaturization as well as the increasing integration of functions this applies in particular for the production of technical plastic products. In addition, thermoplastics have very complex material behavior and properties that also depend on the processing parameters and environmental conditions. The high demands on the processes and products require a comprehensive quality monitoring. This is usual very expensive and can often only be done with control samples. They are analyzed by a quality assurance lab and results therefore are only available with a clear time lag. Since such a sample-based SPC monitoring for plastic products has proven inadequacies, procedures are necessary to allow a 100% monitoring directly in the production cycle.

In this thesis a model-based monitoring procedure was developed. It depicts the process-specific relationships between process variations and the quality of components and thus allows for a quality forecast synchronous to production cycle. For modeling "general regression neuro fuzzy networks" (GRNFN) are used. This model structure combines the advantages of data based and knowledge based methods. First, the properties of GRNFN and the influence of the available parameters are examined in detail. Following the GRNFN models are expanded to be combined fuzzy rules. So a GRNFN model which has been trained by the use of training samples can be extended with existing empirical knowledge in terms of a set of fuzzy rules. In addition, indicators are developed by which input data can be detected, which is not or only poorly mapped by the model. These indicators can be used, to identify new states of the process. Thus the process model is capable to detect changes or disturbances that may be caused by the material, the process adjustment or preliminary processes. Hence committee can be avoided effectively. By performing a simple retrain of the model new relationships can be introduced to the model. Also it's possible to remove older knowledge. GRNFN models are therefore suitable for an active learning procedure and can adapt dynamically to a process and its drift.

To enable process models to give reliable statements, meaningful process data must be provided. Process characteristics are built from process-integrated sensor signals, for ex-

ample, by deriving significant parameters. These process characteristics form the so-called state vector that reflects process variations and therefore represents some kind of fingerprint of the process. The composition of the state vector is crucial for the achievable model accuracy. In this work the derivation of characteristics, possibilities of visualization and ways to improve the state vectors are closely examined.

The performance of GRNFN models and their suitability are examined on the basis of ultrasonic welding of plastics, also including custom signal processing. All attempts are first made by using a specimen specially designed for ultrasonic welding and afterwards on thermoplastic parts out of everyday production. On the one hand quality characteristics as the weld strength are predicted by GRNFN models. If the measured quality values, which have been used to train the model, were determined reliably and reproducibly, the achievable performance of quality prediction is very good. The forecast results were also compared with neural networks with feedforward structure and backpropagation training as they are used in many systems for process modeling. The created GRNFN models yield the better results. Furthermore, the influence of parameters is investigated, that are subject to variations in practice and thus may lead to poor welding results. Besides changes in material these factors include changes in molding conditions and a modified storage (temperature, humidity, ...) of the parts. All these factors could be mapped by the GRNFN models or if they were not included in the training patterns are identified as unknown states. Thus a model-based monitoring of ultrasonic welding in terms of weld quality and changed process conditions due to preliminary processes is possible by using the presented GRNFN models.

Literaturverzeichnis

- [AA07] ABIYEV, R. H. ; ALTUNKAYA, K.: Iris recognition for biometric personal identification using neural networks. In: *Proceedings of the 17. international conference on artificial neural networks (ICANN'07)*, Berlin, Heidelberg: Springer, 2007, S. 554–563
- [Ahr04] AHREND, M.T.: *Grenzen und Möglichkeiten Neuronaler Netze zur Qualitätssicherung beim Widerstandspunktschweißen unter fertigungsgerechten Bedingungen*. Aachen: Shaker, 2005, RWTH Aachen, Diss., 2004
- [Alt95] VON ALTROCK, C.: *Fuzzy Logic - Band 1: Technologie*. München: Oldenbourg, 1995
- [ANS10] ABELLAN-NEBOT, J.V. ; SUBIRÓN, F.R.: A review of machining monitoring systems based on artificial intelligence process models. In: *International Journal of Advanced Manufacturing Technology* 47 (2010), Nr. 1-4, S. 237–257
- [AWW72] ALOISIO, C.J. ; WAHL, D.G. ; WHETSEL, E.E.: A simplified thermoviscoelastic analysis of ultrasonic bonding. In: *Annual Technical Conference, Society of Plastics Engineers* Bd. 18, 1972, S. 445–451
- [BC89] BENATAR, A. ; CHENG, Z.: Ultrasonic welding of thermoplastics in the far-field. In: *Polymer Engineering and Science* 29 (1989), Nr. 23, S. 1699–1703
- [BE93] BOURDON, R. ; EHRENSTEIN, G.W.: Prozessrobustheit beim Spritzgießen optimieren. In: *Qualität und Zuverlässigkeit* 38 (1993), Nr. 2, S. 29
- [Bec71] BECKER, K.: *Untersuchung der Vorgänge bei der Ultraschallschweißung thermoplastischer Kunststoffe mit Hilfe elektronischer Meßverfahren*, TU Braunschweig, Diss., 1971
- [Ben87] BENATAR, A.: *Ultrasonic welding of advanced thermoplastic composites*, Massachusetts Institute of Technology, Diss., 1987
- [BGV92] BOSER, B. ; GUYON, I. ; VAPNIK, V.: A training algorithm for optimal margin classifiers. In: *Proceedings of the Fifth Annual Workshop on Computational Learning Theory* New York: ACM Press, 1992, S. 144–152

- [BH94] BUCKLEY, J.J. ; HAYASHI, Y.: Fuzzy Neural Networks: A survey. In: *Fuzzy Sets and Systems* 66 (1994), Nr. 1, S. 1–13
- [BH99] BERTHOLD, M. ; HAND, D.J.: *Intelligent Data Analysis - An Introduction*. Berlin, Heidelberg: Springer, 1999
- [Bic06] BICHLER, M.: *Kunststoffteile fehlerfrei spritzgießen*. 2., überarb. und erw. Auflage. Heidelberg: Hüthig, 2006
- [Bis95] BISHOP, C.M.: *Neural Networks for Pattern Recognition*. New York: Oxford University Press, 1995
- [Bis06] BISHOP, C.M.: *Pattern Recognition and Machine Learning*. New York: Springer, 2006
- [BK92] BERENJI, H.R. ; KHEDKAR, P.: Learning and Tuning Fuzzy Logic Controllers Through Reinforcements. In: *IEEE Trans. Neural Networks* 3 (1992), Nr. 5, S. 724–740
- [BK08] BODYANSKIY, Y. ; KULISHOVA, N.: Memory-based neuro-fuzzy system for interpolation of reflection coefficients of printing inks. In: *Cybernetics and Systems Analysis* 44 (2008), Nr. 5, S. 625–632
- [BMB76] BECKERT, M. ; MARTINEK, I. ; BRINKE, G.: Verfahren und Einrichtungen zur Prüfung der Ultraschallschweißignung von Thermoplasten. In: *ZIS-Mitteilungen* 22 (1976), Nr. 7, S. 671–685
- [BOPT07] BODYANSKIY, Y. ; OTTO, P. ; PLISS, I. ; TESLENKO, N.: Nonlinear process identification and modeling using general regression neuro-fuzzy network. In: *52. Internationales Wissenschaftliches Kolloquium, TU Ilmenau, 10-13. September 2007*, 2007
- [Bre07] BRECKWEG, Achim: *Automatisiertes und prozessüberwachtes Radialclinchen höherfester Blechwerkstoffe*, Institut für Industrielle Fertigung und Fabrikbetrieb, Universität Stuttgart, Diss., 2007
- [Bri88] BRIGHAM, E.O.: *The Fast Fourier Transform and its applications*. New Jersey: Prentice-Hall, 1988
- [BSNS04] BRECKWEG, A. ; SCHMIDBERGER, E. ; NEHER, J. ; SCHRAFT, R.D.: Inline-Qualitätskontrolle beim Radialclinchen. In: *Automatisierungstechnische Praxis atp* 46 (2004), Nr. 2, S. 68–71

- [BT08] BODYANSKIY, Y. ; TESLENKO, N.: General regression neuro-fuzzy network for identification of nonstationary plants. In: *International Journal 'Information Technologies and Knowledge' 2* (2008), Nr. 2, S. 136–142
- [BTP92] BEZDEK, J.C. ; TSAO, E.C.-K. ; PAL, N.R.: Fuzzy Kohonen Clustering Networks. In: *Proc. IEEE Int. Conf. on Fuzzy Systems 1992*. San Diego, 1992, S. 1035–1043
- [CAL94] COHN, D. ; ATLAS, L. ; LADNER, R.: Improving Generalization with Active Learning. In: *Machine Learning* 15 (1994), Nr. 2, S. 201–221
- [CAM] BASF CAMPUS® Werkstoffdatenbank für Kunststoffe <http://campus-i.plasticsportal.net/matdb/matdb.php> (6. September 2010)
- [CCCL00] CHUAH, Y.K. ; CHIEN, L.-H. ; CHANG, B.C. ; LIU, S.-J.: Effects of the shape of the energy director on far-field ultrasonic welding of thermoplastics. In: *Polymer Engineering and Science* 40 (2000), Nr. 1, S. 157–167
- [CCL11] CHANG, C.-C. ; CHIEN, L.-J. ; LEE, Y.-J.: A novel framework for multi-class classification via ternary smooth support vector machines. In: *Pattern Recognition* 44 (2011), Nr. 6, S. 1235–1244
- [CDM09] CODARA, P. ; D'ANTONA, O.M. ; MARRA, V.: A characterisation of bases of triangular fuzzy sets. In: *IEEE International Conference on Fuzzy Systems 2009 - FUZZ-IEEE 2009*, 2009, S. 604–609
- [CH07] CZICHOS, H. (Hrsg.) ; HENNECKE, M. (Hrsg.): *Hütte - Das Ingenieurwissen*. 33. Auflage. Berlin, Heidelberg: Springer, 2007
- [CHV99] CHAPELLE, O. ; HAFFNER, P. ; VAPNIK, V.N.: Support vector machines for histogram-based image classification. In: *IEEE Transactions on Neural Networks* 10 (1999), Nr. 5, S. 1055–1064
- [CST00] CRISTIANINI, N. ; SHAWE-TAYLOR, J.: *An Introduction to Support Vector Machines*. Cambridge: Cambridge University Press, 2000
- [CV95] CORTES, C. ; VAPNIK, V.: Support-vector network. In: *Machine Learning* 20 (1995), Nr. 3, S. 273–297
- [Das] DASYLab - Data Acquisition System Laboratory <http://www.dasylab.com> (6. September 2010)
- [DE95] DAGAN, I. ; ENGELSON, S.P.: Committee-Based Sampling For Training Probabilistic Classifiers. In: PRIEDITIS, A. (Hrsg.): *Machine learning: Proceedings*

- of the 12th International Conference on Machine Learning, 9 - 12. July 1995, Tahoe City, California, 1995, S. 150–157*
- [Dev01] DEVINE, J.: Ultrasonic plastics welding basics. In: *Welding Journal* 80 (2001), Nr. 1, S. 29–33
- [DHS00] DUDA, R.O. ; HART, P.E. ; STORK, D.G.: *Pattern Classification*. 2. Auflage. Oxford: John Wiley & Sons, 2000
- [DS05] DIETRICH, E. ; SCHULZE, A.: *Statistische Verfahren zur Maschinen- und Prozessqualifikation*. München, Wien: Hanser, 2005
- [EB88] EL BARBARI, N.: *Ultraschallschweißen von Thermoplasten - Möglichkeiten der Einsatzoptimierung*, RWTH Aachen, Diss., 1988
- [Ehr04] EHRENSTEIN, G.W. (Hrsg.): *Handbuch Kunststoff-Verbindungstechnik*. München, Wien: Hanser, 2004
- [Erh08] ERHARD, G.: *Konstruieren mit Kunststoffen*. 4. Auflage. München, Wien: Hanser, 2008
- [Fis35a] FISHER, R.: *The Design of Experiments*. Edinburgh: Oliver and Boyd, 1935
- [Fis35b] FISHER, R.: *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd, 1935
- [Fis90] FISCHBACH, G.: Rechnergestützte Einflußgrößenüberwachung aus der Sicht des Maschinenherstellers. In: *Süddeutsches Kunststoff-Zentrum: SKZ-Fachtagung "SPC im Spritzgießbetrieb", 27. November 1990, Würzburg, 1990*
- [Fis06] FISCHER, T. *Schulungsunterlagen 'Ultraschall Fügetechnologien für thermoplastische Kunststoffe', Fa. Herrmann Ultraschalltechnik, Karlsbad. 2006*
- [Fle91] FLEISCHER, D.: Entwicklungstendenzen polymerer Werkstoffe. In: *Süddeutsches Kunststoff-Zentrum: SKZ-Fachtagung "Hochleistungswerkstoffe in der technischen Anwendung", 4.-5. Juni 1991, Würzburg, 1991, S. 7–23*
- [FS96] FISCHER, G. ; SCHMIDBERGER, E.: Einsatz Neuronaler Netze zur Prozess- und Qualitätsüberwachung beim Kunststoffspritzgießen (Industrie-Verbundprojekt NEPRES). In: *Tagungsband IPA-Technologie Forum F21, Fraunhofer IPA, 4. Dezember 1996, Stuttgart, 1996, S. 49–68*
- [FSK98] FREY, C. W. ; SAJIDMAN, M. ; KUNTZE, H.-B.: Ein neuro-adaptives Regelungskonzept mit online-fähigem Kalman-Filter-Lernverfahren für stochastisch

- gestörte nichtlineare Prozesse. In: *VDI/VDE-GMA-Tagung "Computational Intelligence", Berlin, 3./4. März 1998, VDI-Berichte 1381*, Düsseldorf: VDI-Verlag, 1998, S. 381–392
- [FW80] FRANKEL, E.J. ; WANG, K.K.: Energy transfer and bond strength in ultrasonic welding of thermoplastics. In: *Polymer Engineering and Science* 20 (1980), Nr. 6, S. 396–401
- [GHOS95] GUTBROD, O. ; HEINZELMANN, F. ; OHNGEMACH, D. ; SCHMIDBERGER, E.: Neuronale Netze beim Spritzgießen - Ein neuer Weg zur Qualitätssicherung. In: *Kunststoffe* 85 (1995), Nr. 5, S. 620–623, 626
- [Gie91] GIERTH, M.: *Methoden und Hilfsmittel zur prozessnahen Qualitätssicherung beim Spritzgießen von Thermoplasten*. Aachen : Verl. der Augustinus-Buchh., 1992, RWTH Aachen, Diss., 1991
- [Gie95] GIESE, M.: *Fertigungs- und werkstofftechnische Betrachtungen zum Vibrations-schweißen von Polymerwerkstoffen*, Universität Erlangen-Nürnberg, Diss., 1995
- [Gol89] GOLDBERG, D.E.: *Genetic Algorithms in Search, Optimization, and Machine Learning*. Boston: Addison-Wesley Professional, 1989
- [Haa89] HAAS, R.: Finite-Elemente-Methode für die Auslegung von Ultraschallsonotroden. In: *Industrie-Anzeiger* 111 (1989), Nr. 87, S. 22–23
- [Han94] HANNING, D.: *Continuous Process Control - Qualitätssicherung im Kunststoffverarbeitungsprozess auf Basis Statistischer Prozessmodelle*. Aachen: Shaker, 1995, Universität-Gesamthochschule Paderborn, Diss., 1994
- [HE84] HARTUNG, J. ; ELPELT, B.: *Multivariate Statistik*. München: Oldenbourg, 1984
- [Hei09] HEINRICH, G.: Bruchmechanik und Statistische Mechanik von verstärkten Elastomeren. In: *12. Problemseminar "Deformation und Bruchverhalten von Kunststoffen", 24.-26. Juni 2009* Kunststoff-Kompetenzzentrum Halle-Merseburg, 2009
- [HEK09] HARTUNG, J. ; ELPELT, B. ; KLÖSENER, K.-H.: *Statistik: Lehr- und Handbuch der angewandten Statistik*. 15. Auflage. München: Oldenbourg, 2009
- [HG94] HALGAMUGE, S.K. ; GLESNER, M.: Neural Networks in Designing Fuzzy Systems for Real World Applications. In: *Fuzzy Sets and Systems* 65 (1994), Nr. 1, S. 1–12

- [HH02] HENZE, E. ; HABERSTROH, E.: Prozessüberwachung PUR - Statistische Prozessmodelle und neuronale Netze kontrollieren das Reaktionsgießen von Polyurethanen. In: *Elastomers and Plastics* 55 (2002), Nr. 5, S. 232–235
- [HHH04] HELLERICH, W. ; HARSCH, G. ; HAENLE, S.: *Werkstoff-Führer Kunststoffe: Eigenschaften, Prüfungen, Kennwerte*. 9. Auflage. München, Wien: Hanser, 2004
- [HHL07] HOFFMANN, W.-M. ; HABERSTROH, E. ; LÜTZELER, R.: Einfluss der Spritzgießparameter auf den Laserdurchstrahlschweißprozess. In: *Joining Plastics - Fügen von Kunststoffen* 1 (2007), Nr. 1, S. 67–74
- [HK97] HEROLD, H. ; KUSSMAUL, K.: Ultraschallnahtschweißen von Aluminiumwerkstoffen / Universität Magdeburg, MPA Stuttgart. 1997. – Abschlussbericht AIF-Projekt Nr. 10.111B
- [HK04] HABERSTROH, E. ; KUHLMANN, K.: Prozessführung beim Ultraschallschweißen. In: *Kunststoffe* 84 (2004), Nr. 7, S. 35–38
- [Hol92] HOLLAND, J.H.: *Adaptation in Natural and Artificial Systems*. Cambridge: MIT Press, 1992
- [HR88] HABENICHT, G. ; RITTER, J.: Energieumwandlung beim Ultraschallschweißen von Thermoplasten. In: *Kunststoffe* 78 (1988), Nr. 6, S. 546–551
- [HS00] HARRY, M. ; SCHROEDER, R.: *Six Sigma - Prozesse optimieren, Null-Fehler-Qualität schaffen, Rendite radikal steigern*. Frankfurt/Main: Campus, 2000
- [HTF01] HASTIE, R. ; TIBSHIRANI, R. ; FRIEDMAN, J.: *The Elements of Statistical Learning - Data Mining, Inference and Prediction*. New York: Springer, 2001
- [Häu95] HÄUSSLER, J.: *Eine Qualitätssicherungsstrategie für die Kunststoffverarbeitung auf der Basis Neuronaler Netze*. Düsseldorf: VDI-Verlag, 1995 (Fortschritts-Bericht VDI Reihe 20, Nr. 149)
- [HV10] HÜLLERMEIER, E. ; VANDERLOOY, S.: Combining predictions in pairwise classification: an optimal adaptive voting strategy and its relation to weighted voting. In: *Pattern Recognition* 43 (2010), Nr. 1, S. 128–142
- [HW95] HÄUSSLER, J. ; WORTBERG, J.: Anpassungsfähige Modelle gesucht. In: *Qualität und Zuverlässigkeit* 40 (1995), Nr. 10, S. 1160–1164
- [Jan93] JANG, J.-S.R.: ANFIS: Adaptive-network-based fuzzy inference system. In: *IEEE Transactions on Systems, Man and Cybernetics* 23 (1993), Nr. 3, S. 665–685

- [JM04] JOHANNABER, F. ; MICHAELI, W.: *Handbuch Spritzgießen*. 2. Auflage. München, Wien: Hanser, 2004
- [Jol86] JOLLIFFE, I.T.: *Principal Component Analysis*. New York: Springer, 1986
- [JS95] JANG, J.-S.R. ; SUN, C.-T.: Neuro-Fuzzy Modeling and Control. In: *Proceedings of the IEEE* 83 (1995), Nr. 3, S. 378–406
- [JZ97] JAIN, A. ; ZONGKER, D.: Feature selection: evaluation, application, and small sample performance. In: *IEEE Trans. Pattern Analysis and Machine Intelligence* 19 (1997), Nr. 2, S. 153–158
- [Kai92] KAISER, H.: *Prozessanalyse und Prozessführung beim linearen Vibrations-schweißen von Kunststoffen*. Düsseldorf : Dt. Verl. für Schweißtechnik, 1992, Universität-Gesamthochschule Paderborn, Diss., 1992
- [Kar09] KARRENBURG, U.: *Signale, Prozesse, Systeme - eine multimediale und interaktive Einführung in die Signalverarbeitung*. 5., neu bearb. u. erw. Auflage. Berlin, Heidelberg: Springer, 2009
- [KK99] KIM, J. ; KASABOV, N.: HyFIS: adaptive neuro-fuzzy inference systems and their application to nonlinear dynamical systems. In: *Neural Networks* 12 (1999), Nr. 9, S. 1301–1319
- [KKWG97] KASABOV, N.K. ; KIM, J. ; WATTS, M.J. ; GRAY, A.R.: FuNN/2 - a fuzzy neural network architecture for adaptive learning and knowledge acquisition. In: *Information Sciences* 101 (1997), Nr. 3-4, S. 155–175
- [Kle92] KLEPPMANN, W.G.: Statistische Versuchsplanung - Klassisch, Taguchi oder Shainin? In: *Qualität und Zuverlässigkeit* 37 (1992), Nr. 2, S. 89–92
- [Kle03] KLEPPMANN, W. ; BRUNNER, Franz J. (Hrsg.): *Taschenbuch Versuchsplanung - Produkte und Prozesse optimieren*. 3. überarb. Auflage. München, Wien: Hanser, 2003
- [KM94] KORTE, W. ; MICHAELI, W.: Untersuchung neuer Prozessführungsstrategien zur Steigerung der Nahtfestigkeit beim Ultraschallschweißen / IKV, RWTH Aachen. 1994. – Abschlussbericht zum AiF-Forschungsvorhaben Nr. 8626
- [Kön95] KÖNIG, A.: *Neuronale Strukturen zur sichtgestützten Oberflächeninspektion von Objekten in industrieller Umgebung*, Technische Hochschule Darmstadt, Diss., 1995

- [Koh01] KOHONEN, T.: *Self-Organizing Maps*. 3. Auflage. Berlin, Heidelberg: Springer, 2001
- [Kor96] KORTE, W.H.: *Einfluß der Spritzgießherstellbedingungen und Möglichkeiten der prozessnahen Qualitätssicherung mittels statistischer Methoden beim Ultraschallschweißen*, RWTH Aachen, Diss., 1996
- [Kre87] KREITER, J.: *Optimierung der Schweißnahtfestigkeit von Heizelementstumpfschweißungen von Formteilen durch verbesserte Prozessführung und Selbsteinstellung*. Düsseldorf : Dt. Verl. für Schweisstechnik, 1988, Universität-Gesamthochschule Paderborn, Diss., 1987
- [KU02] KIJSIRIKUL, B. ; USSIVAKUL, N.: Multiclass support vector machines using adaptive directed acyclic graph. In: *Proceedings of International Joint Conference on Neural Networks (IJCNN 2002)*. Hawaii, 12.-27. Mai 2002, S. 980–985
- [Kul06] KULISHOVA, N.Y.: Interpolation of color reflection coefficients with the help of an artificial radial basis neural network. In: *Bionika Intellekta* 64 (2006), Nr. 1, S. 105–109
- [Kum05] KUMAR, A.P.: *Agent Based Diagnostic System for the Defect Analysis during Chemical Mechanical Polishing (CMP)*. Heimsheim: Jost-Jetter, 2005, Universität Stuttgart, Diss., 2005
- [Lam98] LAMBERT, G.: *Echtzeit-Bildverarbeitung zur Qualitätssicherung bei texturierten Oberflächen unter besonderer Berücksichtigung strukturierter Texturen*. Aachen: Shaker, 1998, TU Darmstadt, Diss., 1998
- [Lan78] LAND, W.: Untersuchungen zum Ablauf des Ultraschallschweißens. In: *Kunststoffe* 68 (1978), Nr. 4, S. 233–237
- [Lan81] LAND, W.: Kunststoff-Formteile Schweißen + Prüfen. In: *Verbindungstechnik* 13 (1981), Nr. 5, S. 26–27
- [LC01] LIU, S.-J. ; CHANG, I.-T.: Factors affecting the joint strength of ultrasonically welded polypropylene composites. In: *Polymer Composites* 22 (2001), Nr. 1, S. 132–141
- [LE85] LAND, W. ; ERLenkÄMPER, E.: Erfahrungen beim Schweißen von Formteilen aus Thermoplasten. In: *Kunststoffe* 75 (1985), Nr. 6, S. 315–320
- [Lie05] LIEDL, P.: *Automatische Online-Früherkennung von Qualitätsschwankungen und Störungen bei zyklischen und kontinuierlichen Prozessen*. Berlin: dissertation.de, 2005, Universität Stuttgart, Diss., 2005

- [LLC⁺99] LIU, S.-J. ; LIN, W.-F. ; CHANG, B.-C. ; WU, G.-M. ; HUNG, S.-W.: Optimizing the joint strength of ultrasonically welded thermoplastics. In: *Advances in Polymer Technology* 18 (1999), Nr. 2, S. 125–135
- [LLLA06] LING, S.-F. ; LUAN, J. ; LI, X. ; ANG, W.L.Y.: Input electrical impedance as signature for nondestructive evaluation of weld quality during ultrasonic welding of plastics. In: *Independent Nondestructive Testing and Evaluation* 39 (2006), Nr. 1, S. 13–18
- [LRS01] LEHTONEN, P. ; RÜB, G. ; SCHMIDBERGER, E.: *Abschlussbericht zum Verbundprojekt: Sicherung der Qualität technischer Spritzgießprodukte durch moderne Methoden der Informationsverarbeitung in der betrieblichen Praxis (SGQtech)*. 2001. – Universität Stuttgart, Fraunhofer IPA. 2001
- [LRSF01] LEHTONEN, P. ; RÜB, G. ; SCHMIDBERGER, E. ; FRITZ, H.-G.: NEPRES-System regelt Spritzgießprozesse. In: FRITZ, C.D. (Hrsg.): *17. Stuttgarter Kunststoff-Kolloquium* Universität Stuttgart, Institut für Kunststofftechnologie, 2001
- [LX01] LING, S.-F. ; XIE, Y.: Monitoring structural integrity using a piezoelectric inertial actuator cum sensor. In: *Journal of Sound and Vibration* 247 (2001), Nr. 4, S. 731–737
- [Mas80] MASING, W.: *Handbuch Qualitätsmanagement*. München, Wien: Hanser, 1980
- [Mat] *Material Data Center* <http://www.materialdatacenter.de/mb/index.php> (6. September 2010)
- [May97] MAYERS, B.: *Prozess- und Produktoptimierung mit Hilfe der Statistischen Versuchsmethodik*. Aachen: Shaker, 1997, RWTH Aachen, Diss., 1997
- [MBH⁺92] MICHAELI, W. ; BREUER, P. ; HOHENAUER, K. ; VON OEPEN, R. ; PHILIPPS, M. ; PÖTSCH, G. ; RECKER, H. ; VACULIK, R.: Qualitätsgesichertes Spritzgießen - Qualitätsregelung hält Formteileigenschaften konstant. In: *Kunststoffe* 82 (1992), Nr. 12, S. 1167–1171
- [MBVW94] MICHAELI, W. ; BLUHM, R. ; VACULIK, R. ; WYBITUL, K.: Formteilfehler sicher erkennen. In: *Kunststoffe* 84 (1994), Nr. 8, S. 979–982
- [MDP97] MITRA, S. ; DE, R.K. ; PAL, S.K.: Knowledge-Based Fuzzy MLP for Classification and Rule Generation. In: *IEEE Transactions on Neural Networks* 8 (1997), Nr. 6, S. 1338–1350

- [MEBN88] MICHAELI, W. ; EL BARBARI, N. ; NETZE, C.: Ermittlung der Zusammenhänge praxisrelevanter Prozessparameter zur Optimierung des Einsatzes von Ultraschall bei der Verbindung von Kunststoffen / IKV, RWTH Aachen. 1988. – Abschlussbericht zum AiF-Forschungsvorhaben Nr. 6972
- [Mei00] MEISSNER, H. *Übersicht von Schweißnahtgestaltungen beim Ultraschallschweißen*. Heusenstamm: Ultratech GmbH. 2000
- [Mes93] MESCH, F.: Statistische Qualitätsregelung bei Stückgut und Fließprozessen, Teil 1. In: *Automatisierungstechnische Praxis - atp 35* (1993), Nr. 10, S. 553 – 562
- [MH00] MITRA, S. ; HAYASHI, Y.: Neuro-Fuzzy Rule Generation: Survey in Soft Computing Framework. In: *IEEE Transactions on Neural Networks* 11 (2000), Nr. 3, S. 748–768
- [MH09] MASTOROCOSTAS, P.A. ; HILAS, C.S.: A block-diagonal recurrent fuzzy neural network for system identification. In: *Neural Computing and Applications* 18 (2009), Nr. 7, S. 707–717
- [Mic92] MICHLER, G.H.: *Kunststoff-Mikromechanik - Morphologie, Deformation und Bruchmechanismen von polymeren Werkstoffen*. München, Wien: Hanser, 1992
- [Mic97] MICHAELI, W.: Entwicklung neuer Qualitätssicherungskonzepte für das Ultraschallschweißen / IKV, RWTH Aachen. 1997. – Abschlussbericht zum AiF-Forschungsvorhaben Nr. 9572
- [Mic06] MICHAELI, W.: *Einführung in die Kunststoffverarbeitung*. 5. Auflage. München, Wien: Hanser, 2006
- [Mit97] MITCHELL, T.M.: *Machine Learning*. New York: McGraw-Hill, 1997
- [MK93] MICHAELI, W. ; KORTE, W.: Bessere Nahtqualität beim Ultraschallschweißen. In: *Kunststoffe* 83 (1993), Nr. 9, S. 667–670
- [MK94a] MICHAELI, W. ; KORTE, W.: Einfluß der Spritzgießbedingungen auf Schweißprozess und Nahtfestigkeit beim Ultraschallschweißen. In: *Plaste und Kautschuk* 41 (1994), Nr. 6, S. 293–297
- [MK94b] MICHAELI, W. ; KORTE, W.: Einfluß der Spritzgußherstellbedingungen auf die Qualität ultraschallgeschweißter Formteile / IKV, RWTH Aachen. 1994. – Abschlussbericht zum AiF-Forschungsvorhaben Nr. 8880

- [MK94c] MICHAELI, W. ; KORTE, W.: Untersuchung neuer Prozessführungsstrategien zur Steigerung der Nahtfestigkeit beim Ultraschallschweißen / IKV, RWTH Aachen. 1994. – Abschlussbericht zum AiF-Forschungsvorhaben Nr. 8626
- [MK95] MICHAELI, W. ; KORTE, W.: Quality assurance in ultrasonic welding on the basis of statistical process models - prediction of weld strength. In: *Welding in the World* 35 (1995), Nr. 5, S. 348–351
- [MKDE94] MICHAELI, W. ; KORTE, W. ; DHAWAN, K. ; EXNER, S.: Festigkeit von Ultraschallschweißnähten vorausberechnen. In: *Plastverarbeiter* 45 (1994), Nr. 10, S. 40–44
- [MKK97] MICHAELI, W. ; KLEIN, H. ; KORTE, W.: Qualitätssicherung - Ausschuss verringern. In: *Plastverarbeiter* 48 (1997), Nr. 3, S. 66–72
- [MKW07] MOLL, H. ; KOSTRZEWA, M. ; WILKE, L.: Ultraschallschweißen von DVS-Prüfkörpern mit Energierichtungsgebern aus ABS. In: *Joining Plastics - Fügen von Kunststoffen* 1 (2007), Nr. 1, S. 84–89
- [Mül88] MÜLLER, W.: Prüfkörper für das NF-Schweißen und US-Schweißen. In: *ZIS-Mitteilungen* 30 (1988), Nr. 7, S. 734–740
- [MLBB89] MICHAELI, W. ; LAUTERBACH, M. ; BREUER, P. ; BUCKES, G.: Spritzgießen von Thermoplasten - Maschineneinstellung der Nachdruckphase mit der pmT-Steuerung. In: *Plastverarbeiter* 40 (1989), Nr. 11, S. 204–210
- [MNK92] MICHAELI, W. ; NETZE, C. ; KORTE, W.: Prozessanalyse beim Ultraschallschweißen am Beispiel hochtemperaturbeständiger Thermoplaste - Strategien zur Erzielung hoher Schweißnahtqualitäten. In: *DVS-Berichte Band 147 'Kunststoffschweißen und -kleben'* (1992), S. 56–59
- [Mon09] MONTGOMERY, D.C.: *Design and Analysis of Experiments*. 7. Auflage. New York: John Wiley & Sons, 2009
- [MP69] MENGES, G. ; POTENTE, H.: Schallfelder und Energieumsetzung beim Ultraschallschweißen von Kunststoffen. In: *Kunststoffe* 59 (1969), Nr. 6, S. 369–374
- [Mus00] MUSTAFA, M.A.-H.: *Modellbasierte Ansätze zur Qualitätsregelung beim Kunststoffspritzgießen*. Aachen: Shaker, 2000, Universität-Gesamthochschule Essen, Diss., 2000
- [MWK94] MICHAELI, W. ; WEGENER, M. ; KORTE, W.: Qualitätssicherheit beim Ultraschallschweißen - Prozessmodell verdrängt Empirie. In: *Kunststoffe Synthetics* 41 (1994), Nr. 7, S. 12–14

- [Nat92] NATROP, J.: *Rechnerunterstützte Parameteroptimierung und Qualitätssicherung in der Serienproduktion am Beispiel des Heizelementschweißens*. Düsseldorf : Dt. Verl. für Schweißtechnik, 1992, Universität-Gesamthochschule Paderborn, Diss., 1992
- [Nau94] NAUCK, Detlef: A Fuzzy Perceptron as a Generic Model for Neuro-Fuzzy Approaches. In: *Proc. Fuzzy-Systeme '94*. 20. Oktober 1994, München, 1994
- [Net92] NETZE, C.: *Ultraschallschweißen von hochtemperaturbeständigen, unverstärkten, kurglasfaser- und endlosfaserverstärkten Thermoplasten*. Aachen: Verl. der Augustinus-Buchh., 1992, RWTH Aachen, Diss., 1992
- [Neu00] NEUNER, M.: Auf dem Weg zur Null-Fehler-Produktion - Neuronale Netze beim Kunststoffspritzgießen. In: *Industrie-Anzeiger* 122 (2000), Nr. 36, S. 66–67
- [NF10] NEHER, J. ; FISCHER, G.: *Abschlussbericht zum Verbundforschungsprojekt "SC-QUPUS: Signalbasiertes Clustering zur qualitätsorientierten Prozessüberwachung beim Ultraschallschweißen thermoplastischer Kunststoffprodukte"*. 2010. – Fraunhofer IPA, Hochschule Esslingen. 2010
- [NK93] NAUCK, D. ; KRUSE, R.: A Fuzzy Neural Network Learning Fuzzy Control Rules and Membership Functions by Fuzzy Error Backpropagation. In: *Proc. IEEE Int. Conf. on Neural Networks 1993*. San Francisco, März 1993, S. 1022–1027
- [NK95] NAUCK, D. ; KRUSE, R.: NEFCLASS – A Neuro-Fuzzy Approach for the Classification of Data. In: GEORGE, K.M. et a. (Hrsg.): *Applied Computing 1995. Proc. of the 1995 ACM Symposium on Applied Computing, Nashville, 26.-28. Februar 1995*. New York: ACM Press, 1995, S. 461–465
- [NK96] NAUCK, D. ; KRUSE, R.: Designing Neuro-Fuzzy Systems Through Backpropagation. In: PEDRYCZ, W. (Hrsg.): *Fuzzy Modelling: Paradigms and Practice*. Dordrecht: Kluwer, 1996
- [NK97] NAUCK, D. ; KRUSE, R.: A neuro-fuzzy method to learn fuzzy classification rules from data. In: *Fuzzy Sets and Systems* 89 (1997), Nr. 3, S. 277–288
- [NK10a] NEHER, J. ; KAUPP, M.: Active Learning-Based Adaptive Control in Low-Runner Processes. In: TETI, R. (Chairman) u.a.; CIRP u. (Hrsg.): *Intelligent Computation in Manufacturing Engineering - CIRP ICME '10: Innovative and Cognitive Production Technology and Systems. 7th CIRP International Conference*. Capri, Italy, 23-25. Juni 2010

- [NK10b] NEHER, J. ; KAUPP, M.: A new approach for quality monitoring in small batch processes. In: LIEN, T.K. (Hrsg.): *Responsive, customer demand driven, adaptive assembly. Proceedings of the 3rd CIRP Conference on Assembly Technologies and Systems*, Trondheim: Tapir Academic Press, 2010, S. 205–208
- [NL96] NONHOF, C.J. ; LUITEN, G.A.: Estimates for process conditions during the ultrasonic welding of thermoplastics. In: *Polymer Engineering and Science* 36 (1996), Nr. 9, S. 1177–1183
- [NM91] NETZE, C. ; MICHAELI, W.: Correlation of welding parameters, energy conversion and mechanical weld seam properties for ultrasonic welding. In: *49th ANTEC, 5.-9.5.1991, Montreal (Canada), Society of Plastic Engineers (SPE)* Bd. 37, 1991, S. 2490–2493
- [N.N98] N.N.: D2H - Doppelt-dynamische Hüllkurve: Meilenstein der BRANKAMP Überwachungstechnologie. In: *Brankamp International Journal* 15 (1998), Nr. A12/1
- [N.N03] N.N.: Entwicklung eines Qualitätssicherungs(QS)-Systems für das Ultraschallschweißen auf Basis Neuronale Netze unter Nutzung der von der Maschine zur Verfügung gestellten Messwerte / Institut für Materialprüfung, Werkstoffkunde und Festigkeitslehre (IMWF), Universität Stuttgart. 2003. – Abschlussbericht AIF-Projekt Nr. 12.937 N / DVS-Nr. 5.022 / BA-Nr. 1163
- [N.N05] N.N.: *DVS-Prüfkörper mit Energierichtungsgeber zum Ultraschallschweißen thermoplastischer Kunststoffe und Prüfbedingungen; Richtlinie 2216-1 Beiblatt 1 (Ausgabe 09/2005)*. Düsseldorf: DVS-Verlag, 09 2005
- [N.N07a] N.N.: Die andere Art des Schweißens. In: *Automation* 25 (2007), Nr. 6, S. 74–77
- [N.N07b] N.N.: *DVS 2216-1, Ultraschallschweißen von Kunststoffserienteilen - Prozessbeschreibung, Maschinen und Geräte, Einflussgrößen, Konstruktion, Qualitätssicherung*. Düsseldorf: DVS-Verlag, 04 2007
- [NS07] NEHER, J. ; SCHMIDBERGER, E.: Qualitätsbasierte Überwachung und Regelung von Spritzgießprozessen mit NEPRES. In: *Spritzgießen von Präzisionsteilen - Zukunftsorientierte Qualitätssicherung für die Serienfertigung*, Stuttgart: FpF, 25. September 2007. – Fraunhofer IPA Technologieforum Bd. F150, S. 114–137
- [OD92] OHANIAN, P.P. ; DUBES, R.C.: Performance evaluation for four classes of textural features. In: *Pattern Recognition* 25 (1992), Nr. 8, S. 819–833

- [OFG97] OSUNA, E. ; FREUND, R. ; GIROSI, F.: Training support vector machines: An application to face detection. In: *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR'97), Puerto Rico, 17-19. Juni, 1997*
- [Pan05] PANNEKAMP, J.: *Adaptive Verfahren zur Bewertung texturierter Oberflächen*. Heimsheim: Jost-Jetter, 2005, Universität Stuttgart, Diss., 2005
- [Par62] PARZEN, E.: On the estimation of a probability density function and the mode. In: *Annals of Mathematic Statistics* 33 (1962), Nr. 3, S. 1065–1076
- [Pei96] PEINEMANN, F.W.: *Neuronale Regelungskonzepte zur sensomotorischen Koordination von Industrierobotern*. Düsseldorf: VDI-Verl., 1996, Universität Bremen, Diss., 1996
- [Pfe90] PFEIFER, T.: *Untersuchungen zur Qualitätssicherung, Stand und Bewertung, Empfehlungen für Maßnahmen*. Karlsruhe: Kernforschungszentrum Karlsruhe, 1990 (KfK-PFT-Bericht 155)
- [Pfe96] PFEIFER, T.: *Qualitätsmanagement*. München, Wien: Hanser, 1996
- [Pfe10] PFENNING, W.: Nehmerqualitäten - GKV Jahrespressekonferenz 2010. In: *Plastverarbeiter* 61 (2010), Nr. 3, S. 12–15
- [PH00] PARK, S. ; HAN, T.: Iterative Inversion of Fuzzified Neural Networks. In: *IEEE Transactions on Fuzzy Systems* 8 (2000), Nr. 3, S. 266–280
- [PK88] POTENTE, H. ; KAISER, H.: Prozessanalyse des Vibrationsschweißens - Abschlussbericht zum AiF-Forschungsvorhaben Nr. 6740 / Universität-Gesamthochschule Paderborn. 1988. – Forschungsbericht
- [PKM05] PAWAR, R. V. ; KAJAVE, P.P. ; MALI, S.N.: Speaker identification using neural networks. In: *World Academy of Science, Engineering and Technology* 12 (2005), S. 31–35
- [PL95] PHAM, D.T. ; LIU, X.: *Neural Networks for Identification, Prediction and Control*. New York: Springer, 1995
- [PMW00] PANNEKAMP, J. ; MELCHIOR, K. ; WESTKÄMPER, E.: Adaptex: Adaptive Bewertung texturierter Oberflächen. In: *Bildverarbeitung im industriellen Einsatz, Stuttgart*. Düsseldorf: VDI/VDE-GMA, 2000. – (VDI-Bericht, Bd. 1572), S. 199–206
- [PNC00] PANDE, P.S. ; NEUMAN, R.P. ; CAVANAGH, R.R. ; PANDE, Peter S. (Hrsg.): *The Six Sigma Way: how GE, Motorola, and other top companies are honing their performance*. New York: McGraw-Hill, 2000

- [PNH92] POTENTE, H. ; NATROP, J. ; HANNING, D.;; Konzepte zur Qualitätssicherung. In: *Kunststoffe* 82 (1992), Nr. 11, S. 1078–1083
- [Pot71] POTENTE, H.: *Untersuchung der Schweißbarkeit thermoplastischer Kunststoffe mit Ultraschall*, RWTH Aachen, Diss., 1971
- [Pot04] POTENTE, H.: *Fügen von Kunststoffen: Grundlagen, Verfahren, Anwendung*. München, Wien: Hanser, 2004
- [PS98] PANNEKAMP, J. ; SCHMUTZ, M.: Methodology for evaluation and selection of features. In: *Workshop on Texture Analysis 1998 (WTA '98)*, Albrecht Ludwigs-Universität Freiburg. 1998, S. 151–164
- [PSK98] PANNEKAMP, J. ; SCHMIDBERGER, E. ; KISCHKAT, R.: Using the Karhunen-Loeve Expansion for Feature Extraction on Small Sample Sets. In: *IECON '98. Proceedings of the 24th Annual Conference of the IEEE Industrial Electronics Society* Bd. 3, 1998, S. 1582–1586
- [PTVF07] PRESS, W.H. ; TEUKOLSKY, S.A. ; VETTERLING, W.T. ; FLANNERY, B.P.: *Numerical Recipes - The Art of Scientific Computing*. 3. Auflage. New York: Cambridge University Press, 2007
- [PV98] PONTIL, M. ; VERRI, A.: Support vector machines for 3D object recognition. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20 (1998), Nr. 6, S. 637–646
- [PWHH93] POTENTE, H. ; WORTBERG, J. ; HANNING, D. ; HÄUSSLER, J.: Kritische Betrachtungen zur Prozessüberwachung in der Kunststoffverarbeitung, Teil 1. In: *Plastverarbeiter* 44 (1993), Nr. 8, S. 44–49
- [Rab93] RABE, J.: SPC als Grundlage und Voraussetzung beim Spritzgießen technischer Präzisionsformteile. In: *Süddeutsches Kunststoff-Zentrum: SKZ-Fachtagung 'Statistische Prozesskontrolle (SPC) im Spritzgießbetrieb' am 11./12. Dezember 1993*. Würzburg, 1993
- [Rib05] RIBEIRO, B.: Support Vector Machines for Quality Monitoring in a Plastic Injection Molding Process. In: *IEEE Transactions on Systems, Man, and Cybernetics* 35 (2005), Nr. 3, S. 401–410
- [Ric97] RICHTER, A.: *Möglichkeiten und Grenzen der Prozessüberwachung zur Qualitätssicherung beim Fügen durch Umformen*, TU Hamburg-Harburg, Diss., 1997
- [Rie88] RIES, H.: *Veränderung von Werkstoff- und Formteilstruktur beim Spritzgießen von Thermoplasten*, RWTH Aachen, Diss., 1988

- [Rit86] RITTER, J.: *Untersuchungen zur Energieumwandlung und zum Schwingungsverhalten des Systems Sonotrode, Füge­teile und Amboß beim Ultraschallschweißen ausgewählter Thermoplaste*. Düsseldorf: Dt. Verl. für Schweißtechnik, 1987, TU München, Diss., 1986
- [RK04] RIFKIN, R. ; KLAUTAU, A.: In defense of one-vs-all classification. In: *Journal of Machine Learning Research* 5 (2004), S. 101–141
- [RM86] RUMELHART, D.E. ; MCCLELLAND, J.L.: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 1*. Cambridge: MIT Press, 1986
- [RM95] RINNE, H. ; MITTAG, H.-J.: *Statistische Methoden der Qualitätssicherung*. 3. Auflage. München, Wien: Hanser, 1995
- [RNW91] RAMARATHNAM, G. ; NORTH, T.H. ; WOODHAMS, R.T.: Effect of Tie-Layer Crystallinity on Heat Generation and High Speed Filming of Ultrasonic Welding. In: *Berichte Soc. of Plastics Engineers (SPE) ANTEC'91, 5.-9. May 1991*. Montreal, Kanada, 1991, S. 2486–2489
- [Ros58] ROSENBLATT, F.: The perceptron: a probabilistic model for information storage and organization in the brain. In: *Psychological review* 65 (1958), Nr. 6, S. 386–408
- [Ros99] ROSKAM, R.: *In-Prozess-Überwachung von Pressen der Blechverarbeitung*, Universität Hannover, Diss., 1999
- [RS01] RAUH, W. ; SCHMIDBERGER, E.: Die Messlatte liegt bei Null - Fehler beim Spritzgießen von Kunststoffteilen erzeugen teuren Ausschuss. In: *Spektrum der Wissenschaft* (2001), Nr. 5, S. 88–89
- [Sam69] SAMMON, J.W.: A Nonlinear Mapping for Data Structure Analysis. In: *IEEE Transactions on Computers* C-18 (1969), Nr. 5, S. 401–409
- [Sch97] SCHEFFLER, E.: *Statistische Versuchsplanung und -auswertung: Eine Einführung für Praktiker*. 3. Auflage. Stuttgart: Deutscher Verlag für Grundstoffindustrie, 1997
- [Sch09] SCHWARZKOPF, P.: Industrielle Bildverarbeitung: Nach Rekordjahr 2008 derzeit in schwierigem Fahrwasser / Frankfurt/Main: VDMA. 2009. – Forschungsbericht
- [Seu92] SEUTHE, A.: Qualität im Spritzgießbetrieb. In: *Kunststoffberater* 37 (1992), Nr. 11, S. 34 – 42

- [SF05] SCHMIDBERGER, E. ; FISCHER, G.: *Abschlussbericht zum Verbundforschungsprojekt "Qualitätsprognose mittels Prozessinformation beim Ultraschall- und Vibrationsschweißen (QP-UVS)".* 2005. – Fraunhofer IPA, Hochschule Esslingen. 2005
- [SH00] SCHNERR-HÄSELBARTH, O.: *Automatisierung der Online-Qualitätsüberwachung beim Kunststoffspritzgießen.* Aachen: Mainz, 2000, RWTH Aachen, Diss., 2000
- [She31] SHEWHART, W.: *Economic control of quality of manufactured product.* New York: Van Nostrand, 1931
- [SK88] SUGENO, M. ; KANG, G.T.: Structure identification of fuzzy model. In: *Fuzzy Sets and Systems* 28 (1988), Oktober, Nr. 1, S. 15–33
- [SN05a] SCHMIDBERGER, E. ; NEHER, J.: Höchste Aufmerksamkeit - Zyklische Produktionsprozesse qualitätsorientiert regeln. In: *Qualität und Zuverlässigkeit* 50 (2005), Nr. 2, S. 58–59
- [SN05b] SCHMIDBERGER, E. ; NEHER, J.: *FQS-DGQ. Bd. 85-05: Leitfaden für die qualitätsorientierte Prozessüberwachung und -regelung zyklischer Produktionsprozesse (NEPRES II). Schlussbericht. AiF-Projekt-Nr. 13383 N/1.* Frankfurt/Main: Forschungsgemeinschaft Qualitätssicherung e.V.; Deutsche Gesellschaft für Qualität e.V., 2005
- [SNG07] SCHMIDBERGER, E. ; NEHER, J. ; GEBBERT, T.: Brücke zwischen Theorie und Praxis. In: *Plastverarbeiter* 58 (2007), Nr. 1, S. 26–27
- [Sol95] SOLTYSIAK, J.T.: *Überwachung und Steuerung von Gesenkschmiedeprozessen auf Basis messtechnisch erfassbarer Prozessgrößen.* Aachen: Shaker, 1995, Universität-Gesamthochschule Paderborn, Diss., 1995
- [Som90] SOMLUK, G.R.: Why it's time to take an other look at ultrasonic welding. In: *Modern Plastics International* 30 (1990), Nr. 12, S. 47–50
- [Spa05] SPAETH, B.: Industrielle Bildverarbeitung: enormes Wachstum. In: *VDI-Z* 147 (2005), Nr. 7/8, S. 14
- [Spe91] SPECHT, D.F.: A General Regression Neural Network. In: *IEEE Transactions on Neural Networks* 2 (1991), Nr. 6, S. 568–576
- [SS88a] SHAININ, D. ; SHAININ, P.: Better than Taguchi Orthogonal Tables. In: *Quality and Reliability Engineering International* 4 (1988), Nr. 2, S. 143–149

- [SS88b] SIEDLECKI, W. ; SKLANSKY, J.: On Automatic Feature Selection. In: *International Journal on Pattern Recognition and Artificial Intelligence* 2 (1988), Nr. 2, S. 197–220
- [SSP97] SCHMIDBERGER, E. ; SCHMUTZ, M. ; PANNEKAMP, J. *Verfahren zur Überwachung eines zyklischen Produktionsprozesses*. Patent DE 19743600 A1: 19990415. 10 1997
- [Str95] STROHFUSS, W.: “Schwarze Magie“? In: *Kunststoffe* 85 (1995), Nr. 6, S. 827–833
- [Str00] STRACHE, W.: *Multisensorielle Überwachung des Stanzprozesses*. Düsseldorf : VDI-Verl., 2000, Universität Hannover, Diss., 2000
- [SYO98] SENG, M. ; YUSOF, R. ; OMATU, S.: Adaptive Neuro-Fuzzy Control Systems by RBF and GRNN Neural Networks. In: *Journal of Intelligent and Robotic Systems* 23 (1998), Nr. 2-4, S. 267–289
- [Tab01] TABATABAI, S.S.: *Entwicklung und Analyse von Strategien zur prozesszeitkonformen Qualitätssicherung beim Thermoplast-Spritzgießen*, Universität Stuttgart, Diss., 2001
- [Tag87] TAGUCHI, G.: *Systems of Experimental Design. Vol. 1 and Vol. 2*. Dearborn/Michigan: American Supplier Institute, 1987
- [TDW83] TOLUNAY, M.N. ; DAWSON, P.R. ; WANG, K.K.: Heating and bonding mechanisms in ultrasonic welding of thermoplastics. In: *Polymer Engineering and Science* 23 (1983), Nr. 13, S. 726–733
- [Ter94] TERZYK, T.: *Prozessüberwachung beim Gesenkschmieden mit Grat durch Auswertung prozessbedingter Signale*. Düsseldorf : VDI-Verl., 1995, Universität Hannover, Diss., 1994
- [TG06] TADMOR, Z. ; GOGOS, C.G.: *Principles of Polymer Processing*. 2. Auflage. New York: John Wiley & Sons, 2006
- [TGm95] TSCHICHOLD-GÜRMAN, N.: Generation and Improvement of Fuzzy Classifiers with Incremental Learning using Fuzzy RuleNet. In: GEORGE, K.M. et al. (Hrsg.): *Applied Computing 1995. Proc. of the 1995 ACM Symposium on Applied Computing, Nashville, Feb. 26–28*. New York: ACM Press, Februar 1995, S. 466–470

- [Tru79] TRUNK, G.V.: A Problem of Dimensionality: A Simple Example. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1 (1979), Nr. 3, S. 306–307
- [TS85] TAKAGI, T. ; SUGENO, M.: Fuzzy identification of systems and its applications to modeling and control. In: *IEEE Transactions on Systems, Man and Cybernetics* 15 (1985), Nr. 1, S. 116–132
- [TSH03] TUR, G. ; SCHAPIRE, R.E. ; HAKKANI-TÜR, D.: Active Learning For Spoken Language Understanding. In: *Proceedings of International Conference on Acoustics, Speech and Signal Processing (ICASSP'03): April 6- 10, 2003, Hong Kong Convention & Exhibition Center, Hong Kong Bd. 1*, IEEE Operations Center, 2003, S. 276–279
- [Tsu79] TSUKAMOTO, Y.: An approach to fuzzy reasoning method. In: GUPTA, M.M. (Hrsg.) ; RAGADE, R.K. (Hrsg.) ; YAGER, R.R. (Hrsg.): *Advances in Fuzzy Set Theory and Applications*, Amsterdam: Elsevier North-Holland, 1979, S. 137–149
- [Ueb95] UEBBING, M.: *Berechnungsmöglichkeiten und Qualitätssicherung beim Vibrationsschweißen*. Düsseldorf: Dt. Verl. für Schweisstechnik, 1995, Universität-Gesamthochschule Paderborn, Diss., 1995
- [Ujm00] UJMA, A.: *Auswirkungen von Rohstoffchargenschwankungen auf den Spritzgießprozess und regelungstechnische Kompensationsstrategien*, Universität-Gesamthochschule Paderborn, Diss., 2000
- [Vac96] VACULIK, R.: *Regelung der Formteilqualität beim Spritzgießen auf der Basis statistischer Prozessmodelle*. Aachen: Verl. der Augustinus-Buchh., 1996, RWTH Aachen, Diss., 1996
- [Vap95] VAPNIK, V.: *The Nature of Statistical Learning Theory*. New York: Springer, 1995
- [VE00] VETTER, J. ; EHRENSTEIN, G.W.: Online-Qualitätserkennung beim biaxialen Vibrationsschweißen von Thermoplasten. In: *DVS-Berichte* 209 (2000), S. 238–244
- [VH05] VATTERODT, T. ; HÄNSCH, D.: Durchdringende Technik: Kunststoffe fügen - Qualität per Laser. In: *Plastverarbeiter* 56 (2005), Nr. 5, S. 32–33
- [vLvN96] VAN WIJK, H. ; LUITEN, G.A. ; VAN ENGEN, P.G. ; NONHOF, C.J.: Process Optimization of Ultrasonic Welding. In: *Polymer Engineering and Science* 36 (1996), Nr. 9, S. 1165–1176

- [Vog03] VOGT, C.-C.: *Die Eignungsprüfung als Instrument der Qualitätssicherung und Kompetenzbewertung von Prüflaboratorien für mechanisch-technologische Werkstoffprüfungen*, Bergische Universität Wuppertal, Diss., 2003
- [Vow80] VOWINKEL, H.: Ultraschallschweißen von Formteilen aus Acetalcopolymerisat. In: *Industrie-Anzeiger* 102 (1980), Nr. 93, S. 35–39
- [VP96] VISSER, A. ; PEINEMANN, F.W.: Einsatzmöglichkeiten neuronaler Netze in der Industrierobotik. In: *Forschung im Ingenieurwesen* 62 (1996), Nr. 6, S. 149–160
- [Vuo94] VUORIMAA, P.: Fuzzy Self-Organizing Map. In: *Fuzzy Sets and Systems* 66 (1994), Nr. 2, S. 223–231
- [Wal07] WALL, M.: *GAlib: A C++ Library of Genetic Algorithm Components*, 2007
- [Wei07] WEICKER, K.: *Evolutionäre Algorithmen*. 2. Auflage. Teubner, 2007
- [Wes97] WESTKÄMPER, E.: *Null-Fehler-Produktion in Prozessketten; Maßnahmen zur Fehlervermeidung und -kompensation*. Berlin, Heidelberg: Springer, 1997
- [WH91] WORTBERG, J. ; HANNING, D.: SPC und Prozessüberwachung am Beispiel des Spritzgießprozesses. In: *Kunststoffe* 81 (1991), Nr. 10, S. 965–971
- [Wil90] WILLE, H.: SPC - angewendet bei Prozessdaten. In: *Süddeutsches Kunststoff-Zentrum: SKZ-Fachtagung "SPC im Spritzgießbetrieb", 27. November 1990, Würzburg*, 1990
- [Win73] WINTERGERST, S.: Orientierungen und Spannungen in Spritzgußteilen. In: *Kunststoffe* 63 (1973), Nr. 5, S. 636–642
- [WM92] WANG, L.-X. ; MENDEL, J.M.: Fuzzy basis functions, universal approximation and orthogonal least squares. In: *IEEE Transactions on Neural Networks* 3 (1992), Nr. 5, S. 807–814
- [Wor89] WORTBERG, J.: Qualitätssicherung beim Spritzgießen von Thermoplasten. In: *Kunststoffe* 79 (1989), Nr. 2, S. 121–132
- [Wor96] WORTBERG, J.: *Qualitätssicherung in der Kunststoffverarbeitung*. München, Wien: Hanser, 1996
- [WW99] WESTON, J. ; WATKINS, C.: Support Vector Machines for Multi-Class Pattern Recognition. In: *Proceedings of European Symposium on Artificial Neural Networks (ESANN'99), 21.-23. April, Brüssel, Belgien* (1999), S. 219–224

- [Zab95] ZABROCKI, K.: SPC in der Kunststoffherzeugung. In: *Qualität und Zuverlässigkeit* 40 (1995), Nr. 10, S. 1172–1176
- [Zad65] ZADEH, L.A.: Fuzzy Sets. In: *Information and Control* 8 (1965), Nr. 3, S. 338–353
- [Zel94] ZELL, A.: *Simulation Neuronale Netze*. München: Oldenbourg, 1994
- [Zie00] ZIEGLTRUM, M.: *Untersuchung von Ultraschallschweißverbindungen in der Kombination unterschiedlicher Thermoplaste mit einer und mehreren Fügeebenen*. München: Utz, 2001, Universität-Gesamthochschule Paderborn, Diss., 2000
- [Zim93] ZIMMERMANN, H.-J. (Hrsg.): *Fuzzy Technologien - Prinzipien, Werkzeuge, Potentiale*. Düsseldorf: VDI-Verlag, 1993

Die Arbeit untersucht den Einsatz von »general regression neuro-fuzzy networks« (GRNFN) für die Modellbildung bei der Prozessüberwachung. Anhand von Zustandsvektoren können Qualitätsprognosen abgeleitet, aber auch neue, unbekannte Prozesszustände erkannt werden. Eigenschaften von GRNFN-Modellen werden im Detail untersucht und diese erfolgreich zur Prozessüberwachung der Prozesskette beim Ultraschallschweißen von Kunststoffteilen eingesetzt.

ISBN 978-3-8396-0424-3



FRAUNHOFER VERLAG