

FORSCHUNGSBERICHT

# Methodik zur Prozessüberwachung, Fehleranalyse und Regelung durch wissensbasierte Modellierung von Produktionsprozessen mittels Bayes'scher Netze

ProBayes

Jonathan Lambers, Jakob Schüder, Jonas Krauß, Markus Schömig  
Benjamin Baudrit, Thomas Hochrein, Martin Bastian

# ProBayes

Methodik zur Prozessüberwachung, Fehleranalyse und Regelung durch wissensbasierte Modellierung von Produktionsprozessen mittels Bayes'scher Netze



Gefördert durch:

<b>Thema .....</b>	<b>1</b>
<b>Berichtszeitraum .....</b>	<b>1</b>
<b>Forschungsvereinigung .....</b>	<b>1</b>
<b>Forschungseinrichtung(en) .....</b>	<b>1</b>
<b>Abkürzungsverzeichnis .....</b>	<b>III</b>
<b>Projektsteckbrief .....</b>	<b>4</b>
<b>1 Einleitung .....</b>	<b>5</b>
1.1 Anlass für das Forschungsvorhaben .....	5
1.2 Problemstellung .....	5
1.3 Zielsetzung .....	6
<b>2 Stand der Technik .....</b>	<b>7</b>
2.1 Prozessregelung im Spritzgießen .....	7
2.1.1 Der Spritzgießprozess – Einstellparameter, Prozessgrößen und Qualitätsmerkmale .....	7
2.1.2 Prozessregelung beim Spritzgießen .....	8
2.2 Datenerfassung .....	9
2.3 Bayes'sche Netze .....	10
2.4 Simulation .....	11
<b>3 Lösungsweg zur Erreichung des Forschungsziels .....</b>	<b>13</b>
<b>4 Durchgeführte Arbeiten .....</b>	<b>14</b>
4.1 Aufbau der Spritzgießzelle .....	14
4.1.1 Spritzgießmaschine und Peripheriegeräte .....	14
4.1.2 Werkzeugauswahl .....	16
4.1.3 Qualitätsmesssysteme .....	17
4.1.4 Datenerfassung und Vernetzung .....	19
4.2 Entwicklung der topologischen Struktur des Bayes'schen Netzes .....	24
4.3 Vorversuche .....	27
4.4 Aufbau eines Simulationsmodells und Quantifizierung des Sim-to-Real Gap 29	
4.4.1 Modellierung der Randbedingungen in Moldex3D .....	29
4.4.2 Evaluation des digitalen Zwillings anhand der Vorversuche .....	34

---

---

4.4.3	Verknüpfung von Prozesssimulation und Bayes'schen Modellen .....	34
4.5	Hauptversuche .....	34
4.6	Datenanalyse .....	36
4.6.1	Datenaufbereitung und Erzeugung eines Datensatzes.....	36
4.6.2	Extraktion von Features aus den Thermografieaufnahmen.....	39
4.6.3	Extraktion von Features aus Zeitreihen .....	42
4.6.4	Features aus den Bauteilmaßen des Bildverarbeitungssystems.....	44
4.7	Feature Selection.....	45
4.7.1	Datenbasierte Feature Selection .....	45
4.7.2	Wissensbasierte Feature Selection .....	46
4.8	Kausalitätsanalyse.....	48
4.8.1	Allgemeines Vorgehen des Kausalitätstests nach DoWhy.....	48
4.8.2	Umsetzung im Projekt .....	49
4.9	Quantifizierung und Validierung des Bayes'schen Netzes.....	50
4.9.1	Validierung der Gewichtsvorhersage .....	50
4.9.2	Validierung der Parameterempfehlung.....	53
4.9.3	Validierung der Verzugsvorhersage .....	53
4.10	Transfer auf reale Anwendung mit Live-Demonstrator .....	54
4.11	Technologietransfer und Dokumentation .....	58
<b>5</b>	<b>Diskussion der Ergebnisse .....</b>	<b>59</b>
5.1	Fazit der Vorversuche.....	59
5.2	Vollvernetzte Spritzgießzelle.....	59
5.3	Sim-to-Real Gap .....	60
5.4	Validiertes Bayes'sches Netz für die Vorhersage des Bauteilgewichts und des Verzugs.....	64
5.5	Methodik für die Entwicklung Bayes'scher Netze für industrielle Anwendungsfälle .....	65
<b>6</b>	<b>Zusammenfassung und Ausblick.....</b>	<b>68</b>
6.1	Ergebnis 1 – Dateninfrastruktur und Prozessdaten.....	68
6.2	Ergebnis 2 – Sim-to-Real Gap .....	69
6.3	Ergebnis 3 – Bayes'sches Netz zur Prozessüberwachung und Methodik ..	69
6.4	Ausblick.....	69
<b>7</b>	<b>Literatur .....</b>	<b>71</b>
<b>8</b>	<b>Abbildungsverzeichnis .....</b>	<b>74</b>
<b>9</b>	<b>Anhang.....</b>	<b>77</b>



**Abkürzungsverzeichnis**

BLM	Boundary Layer Method
BN	Bayes'sches Netz
DoE	Design of Experiment / Versuchsplan
DSC	Dynamische Differenzkalorimetrie
kNN	Künstliche neuronale Netze
MQTT	Message Queuing Telemetry Transport
NTP	Network-Time-Protocol
OPC UA	Open Platform Communications Unified Architecture
PLS	Partial Least Squares Regression
SVR	Support Vector Regression

## Projektsteckbrief

Ziel des Projekts ProBayes ist die Entwicklung einer Methodik zur Prozessüberwachung, Fehleranalyse und Regelung durch wissensbasierte Modellierung von Produktionsprozessen mittels Bayes'scher Netze.

Dazu wurde eine Spritzgießzelle, bestehend aus einer Spritzgießmaschine und mehreren Peripheriegeräten, aufgebaut und mit Messsystemen zur Bestimmung der Bauteilqualität sowie mit einer Datenerfassung ausgestattet, welche eine zyklusbasierte Aufzeichnung der Prozessdaten ermöglicht. Basierend auf dem Prozesswissen der Spritzgießexperten am SKZ und auf den in umfassenden Versuchsreihen erzeugten Prozessdaten, wurde die topologische Struktur eines Bayes'schen Netzes erarbeitet, welches die Überwachung und Prognose des Qualitätsmerkmals „Bauteilgewicht“ ermöglicht. Zusätzlich wurde ein Spritzgießsimulationsmodell aufgesetzt und die Simulationsergebnisse mit denen realer Versuche verglichen, um die Lücke zwischen Simulation und Realität zu quantifizieren. Das auf realen Prozessdaten trainierte Bayes'sche Netz wurde an der aufgebauten Spritzgießzelle im Live-Betrieb getestet und validiert. Das Bayes'sche Netz war dabei in der Lage zu detektieren, wann das Bauteilgewicht außerhalb der Toleranz lag, die wahrscheinlichste Ursache dafür anzugeben und darüber hinaus Handlungsempfehlungen an den Maschinenbediener auszugeben.

AiF/IGF-Projekt Projektnummer: 21815 N

„ProBayes“

Dauer: 01.05.2021 – 30.04.2023

Unterstützt durch den projektbegleitenden Ausschuss:

KraussMaffei Technologies GmbH	Wago Kontakttechnik GmbH & Co. KG
Katulu GmbH	Evonik Industries AG
GEMÜ Gebr. Müller Apparatebau GmbH & Co. KG	Ziemann & Urban GmbH
Erium GmbH	Atlan-Tec Systems GmbH
Ewikon Heißkanalsysteme GmbH	Motan Holding GmbH
MOCOM Compounds GmbH & Co. KG	Pöppelmann GmbH
Rehau AG	Micro-Epsilon GmbH & Co. KG
BSH Hausgeräte GmbH	Codecentric AG
Wirthwein AG	Symate GmbH
Cavity Eye Deutschland GmbH	Dr. Schneider Holding GmbH i.Ins.
Plus10 GmbH	Enlyze GmbH

## **1 Einleitung**

### **1.1 Anlass für das Forschungsvorhaben**

Der globale Markt für Spritzgießmaschinen ist mit 15,33 Mrd. USD im Jahre 2022 und einer prognostizierten Wachstumsrate (CAGR) bis 2030 von jährlich 3,8 % ein sehr relevanter Markt der Kunststoffverarbeitung. Durch die Vielfalt der möglichen Anwendungen von Spritzgießprodukten und der eingesetzten Materialien, handelt es sich um eine sehr komplexe Prozesslandschaft. [1]

Die Spritzgießbranche ist dabei zunehmend vom Fachkräftemangel betroffen, was dazu führt, dass die Produktionsmitarbeiter für immer mehr Maschinen gleichzeitig verantwortlich sind. Zeitgleich findet eine Transformation des Kunststoffsektors zu mehr Nachhaltigkeit statt, sodass vermehrt Rezyklate zum Einsatz kommen. Rezyklate besitzen volatilere Materialkennwerte als virgine Materialien, was durch Verunreinigungen, Materialvermischung oder unterschiedliche Materialdegradation verursacht wird. Dies erschwert die Verarbeitung, denn die produzierten Kunststoffteile sollen eine möglichst einheitliche Qualität haben.

Diesen beiden Herausforderungen steht eine positive Entwicklung entgegen:

Die rasante Entwicklung im Bereich der Digitalisierung erreicht auch die Maschinenhersteller beziehungsweise die ganze Kunststoffbranche. Durch die Erweiterungen hinsichtlich der Datenerfassung, die Schaffung neuer übergeordneter Standard-Schnittstellen und der Einsatz von Data Analytics Methoden sowie künstlicher Intelligenz bieten sich neue Möglichkeiten, um die genannten Herausforderungen durch den Einsatz von hochmodernen Assistenzsystemen zu bewältigen.

Im Projekt wurden die neuen Entwicklungen im Bereich der Datenerfassung und der daten- und wissensbasierten Prozessmodellierung genutzt, um ein Assistenzsystem für Fachkräfte des Spritzgießprozesses zu entwickeln und zu validieren.

### **1.2 Problemstellung**

Eine Spritzgießmaschine hat eine große Zahl an Sensoren zur Messung von Prozessparametern, wodurch vielerlei Daten vorliegen. Aufgrund der komplexen Natur des Spritzgießprozesses beeinflussen sich viele dieser Prozessparameter gegenseitig und das wiederum abhängig von Maschinentyp, Werkzeuggeometrie und Materialkennwerten.

Aufgrund dessen ist eine hohe Expertise der Maschinenbediener gefragt, welche den Prozess mit möglichst geringer Zykluszeit und konstanter Qualität der Bauteile einstellen können. Da es tendenziell immer weniger dieser Experten gibt, ist die Entwicklung eines Assistenzsystems nötig, das folgende Eigenschaften mitbringt:

- Überwachung ausgewählter Qualitätsmerkmale der Bauteile und Erkennung von Abweichungen aus den geforderten Toleranzbereichen
- Unterstützung bei der Ursachenfindung für die Toleranzabweichungen
- Erteilung von Handlungsempfehlung an den Maschinenbediener
- Angabe von für den Maschinenbediener nachvollziehbaren Kriterien für die Handlungsempfehlung

- Prognose der Qualitätsmerkmale, falls der Maschinenbediener die Handlungsempfehlung übernimmt
- Einfache Adaption des Prozessmodells auf andere Kombinationen von Maschine, Werkzeug und Material

Für die Entwicklung eines solchen Assistenzsystems gibt es weitere Hürden zu überwinden, welche im Projekt adressiert wurden. Dazu gehört die ganzheitliche, zyklusbezogene Datenerfassung an Spritzgießzellen, wozu die Spritzgießmaschine selbst sowie weitere Peripheriegeräte wie Temperierung, Materialkonditionierung oder Qualitätssysteme zählen. Da die Prozessmodellierung daten- und wissensbasiert erfolgen soll, wurden ebenfalls verschiedene Möglichkeiten betrachtet, um den Aufwand für die Datengenerierung an der realen Spritzgießmaschine zu minimieren. Dies wurde durch Variation von Versuchsplänen sowie durch Spritzgießsimulationen untersucht, wobei letzteres aufgrund der Lücke von Simulation zu Realität (sog. Sim-to-Real Gap) zusätzliche Herausforderungen mit sich bringt.

### 1.3 Zielsetzung

Ziel des Forschungsvorhabens war es eine Methodik zu erarbeiten, anhand der wissensbasierte Prozessmodelle auf Basis Bayes'scher Netze (BN) im weiteren Sinne für die Überwachung, Fehleranalyse und Regelung industrieller Prozesse entwickelt werden können. Im Vorhaben wurden konkrete Bayes'sche Netze für den Spritzgießprozess erarbeitet und validiert, sodass aus dem Gelernten eine generalisierbare Methodik entwickelt werden konnte. Zudem war ein weiteres Ziel die Praxisrelevanz der Bayes'schen Netze zu demonstrieren, in dem eine vernetzte Spritzgießzelle im Technikum des SKZ aufgebaut und die BN in einer Live-Inferenz Handlungsempfehlungen an den Maschinenbediener geben konnten.

Konkret verfolgt das Projekt ProBayes folgende aufeinander aufbauende Ziele:

1. Aufbau einer vollvernetzten Spritzgießzelle die neben der Maschine auch Peripheriegeräte und Inline-Messsysteme beinhaltet einschließlich Integration von standardisierten Datenschnittstellen und Aufbau einer automatisierten zyklusbezogenen Datenerfassung aller Maschinen und Geräte
2. Erzeugung von Datensätzen zur Quantifizierung der Bayes'schen Netze und Erweiterung der in Realversuchen generierten Daten durch die Simulation des Prozesses in Moldex3D
3. Entwicklung einer topologischen Struktur für die Regelung mit Bayes'schen Netzen für ausgewählte Qualitätsmerkmale des Versuchswerkzeugs einschließlich Modellierung und Validierung der Prognosefähigkeit und Empfehlungsgüte des Netzes
4. Prototypischer Aufbau eines Live-Demonstrators in der vollvernetzten Spritzgießzelle zur Regelung des Prozesses

Die Möglichkeit zur Zielerreichung basiert dabei auf zwei Hypothesen, welche im Projekt verifiziert wurden:

Hypothese 1: „Die Zusammenhänge und gegenseitigen Abhängigkeiten von Einflussgrößen, Ausgangsgrößen sowie allen anderen prozessrelevanten Parametern eines Spritzgussprozesses lassen sich realitätsnah als hierarchisches Bayes'sches Netz abbilden.“

Hypothese 2: „Durch die Möglichkeit des bidirektionalen Schlussfolgerns in Bayes'schen Netzen sowie der Integration einer zeitlichen Komponente durch die dynamischen Bayes'schen Netzwerke sind die Voraussetzungen für das Erreichen des Forschungsziels geschaffen.“

## **2 Stand der Technik**

### **2.1 Prozessregelung im Spritzgießen**

#### **2.1.1 Der Spritzgießprozess – Einstellparameter, Prozessgrößen und Qualitätsmerkmale**

Aufgrund der großen wirtschaftlichen Bedeutung des Spritzgießprozesses wird an Methoden zur Prozess- und Qualitätsüberwachung sowie der Fehleranalyse bereits seit vielen Jahren geforscht. Die Optimierung der Qualität von im Spritzgießverfahren hergestellten Teilen ist ein technisches und wirtschaftliches Ziel, das entsprechend seit Jahrzehnten Gegenstand zahlreicher wissenschaftlicher Untersuchungen ist, wie in Abschnitt 2.1.2 näher beschrieben.

Bei der Analyse des Spritzgießprozesses, ergeben sich drei Kategorien von Größen: Einstellparameter, Prozessgrößen und Qualitätsmerkmale.

Die Einstellparameter sind solche Größen, die direkt an der Spritzgießmaschine oder deren unmittelbarer Peripherie (z. B. Temperiergeräte oder Heißkanalregler) vom Bediener eingestellt werden können. Dazu gehören Werte wie die Schneckendrehzahl, die Zylindertemperatur, die Wasservorlaufzeit, die Einspritzgeschwindigkeit oder die Nachdruckzeit. Diese dienen im Prozess meist als Sollwerte für die verschiedenen Regelkreisläufe, beispielsweise die Temperaturregelung der Heizbänder am Spritzzylinder, oder als Stellwerte für unregelte Größen.

Die Prozessgrößen sind die tatsächlichen physikalischen Größen, die während der Entstehung des Formteils im Spritzgießwerkzeug herrschen und damit seine Eigenschaften bestimmen. Dies sind zum Beispiel der Druck oder die Temperatur der Kunststoffschmelze sowie deren Veränderung über der Zeit. Dabei durchläuft der Kunststoff von der Schmelze bis zum Festkörper verschiedene Phasenübergänge, die auch zusätzliche Vorgänge wie die Entstehung von Kristallitstrukturen enthalten können. Diese Zustandsgrößen sind meist nicht direkt oder nur schwer zugänglich messbar. Eine gut messbare und häufig angewandte Prozessgröße ist der Werkzeuginnendruck, der mithilfe piezoelektrischer Werkzeuginnendrucksensoren gemessen wird. Diese messen den tatsächlichen Schmelzedruck und dessen zeitlichen Verlauf in der Kavität, wenn auch nur punktuell für eine bestimmte Stelle am Formteil. [2]

Die Merkmalswerte oder Qualitätsmerkmale sind schließlich die gemessenen Eigenschaften der hergestellten Teile. Diese werden durch den Einsatzzweck des Bauteils definiert

und können vielfältige physikalische Eigenschaften betreffen, z. B. Maße und Gewicht, Festigkeitskennwerte, Farbe oder Glanz. [2]

### 2.1.2 Prozessregelung beim Spritzgießen

Das Zusammenspiel vieler Einflussfaktoren, Nichtlinearitäten und komplexer Wechselwirkungen macht das Spritzgießen zu einem herausfordernden Anwendungsgebiet. Eine der größten Herausforderungen bei der Implementierung kontinuierlicher Produktüberwachung und Prozessoptimierung beim Spritzgießen ist die Erfassung der Prozessgrößen und der notwendigen Qualitätskennwerte für die Korrelation mit den Prozessparametern. Hier liegt auch einer der größten Defizite bekannter Veröffentlichungen.

Einige Arbeiten [3, 4] ziehen die Messung des Werkzeuginnendruckes für die kontinuierliche Prozessüberwachung heran. Der erfasste Druck der eingespritzten Kunststoffschmelze ermöglicht eine Bestimmung des idealen Verarbeitungspunkts. Die Bestimmung des Werkzeug-Innendruckes ermöglicht eine Kontrolle der Prozesskonstanz, jedoch keine Rückschlüsse auf die eigentliche Bauteilqualität.

Prozessvariablen beim Spritzguss unterliegen im Allgemeinen multivariaten Einflüssen und weisen Nichtlinearitäten auf. Um derartige Beziehungen abzubilden, eignen sich mathematische Näherungsmodelle (Approximationen), wie künstliche neuronale Netze (kNN), Support Vector Regression (SVR), Partial Least Squares Regression (PLS) und Gauß-Prozesse. Beispielsweise bildet Gao et al. den Zusammenhang zwischen Sensordaten und Produktqualität dagegen mittels SVR ab [5]. Ein Nachteil dieses Modells besteht in der Notwendigkeit großer Datenmengen als Grundlage des Modelltrainings. Das Modell ist sehr spezifisch an einen Anwendungsfall angepasst, weshalb die Wirkzusammenhänge kaum generalisierbar und auf veränderte Bedingungen, bspw. beim Umrüsten auf ein anderes Produkt, übertragbar sind [6]. Zudem sind die vorhandenen Modellstrukturen intransparent und erlauben daher keine Ursachenanalyse bei ungewünschten Veränderungen im Prozess. Der Maschinenbediener hat keine Möglichkeit die getroffenen Entscheidungen nachzuvollziehen.

Auch statistische Methoden werden zur Prozessüberwachung und Qualitätsprognose herangezogen. Beispielsweise leiten Zhang et al. statistische Variablen aus Prozessparametern ab, die signifikant hinsichtlich der Fertigungsqualität sind [7]. Zwar ermöglicht dies teilweise eine effektive Prozessüberwachung sowie die Detektion von Anomalien. Es kann jedoch nicht der quantitative Zusammenhang zwischen Produktqualität und Prozessparametern abgebildet werden, weshalb auch keine Maßnahmen zur Optimierung abgeleitet werden können.

Volle Marktreife haben bisher lediglich ein paar Systeme erreicht. Vor allem die Maschinenhersteller integrieren unternehmenseigene Prozessregelsysteme in deren Steuerung.

Ein bereits auf dem Markt verfügbares System zur adaptiven Prozessanpassung ist die patentierte APC plus Technologie der Firma KraussMaffei Technologies GmbH. Das System analysiert und regelt den Spritzgießprozess dynamisch, sodass bei jedem Schuss ein vollständiges Ausfüllen der Form gewährleistet ist und damit die Produktion von Ausschuss reduziert werden kann. Dazu werden insbesondere materialbedingte Schwankun-

gen der Kompressibilität und des Spritzdrucks sensorisch überwacht und mit einer Referenzkurve abgeglichen. Anschließend kann der Umschaltzeitpunkt und das Nachdruckprofil entsprechend des Massendrucks automatisch angepasst werden [8]. Einen ähnlichen Ansatz verfolgt das IQ-Weight Control Konzept des Maschinenherstellers Engel Austria GmbH. Das Druckprofil beim Einspritzen der Kunststoffschmelze in das Werkzeug wird kontinuierlich mit einer Referenzkurve verglichen und der Umschaltzeitpunkt sowie das Einspritzprofil angepasst. Die Füllmenge wird stabilisiert und somit Qualitätsschwankungen reduziert. Beide Systeme funktionieren jeweils in ihrem eigenen Ökosystem. In der Realität sind Anlagen und deren Peripheriegeräte jedoch häufig von unterschiedlichen Herstellern, sodass die beiden genannten Systeme bei herstellerübergreifenden Anwendungen an ihre Grenzen stoßen. Des Weiteren ist vom Anwender in der kunststoffverarbeitenden Industrie eine Prozesskonstanz gewünscht, welche teilweise durch diese Systeme unterbrochen wird, dadurch dass sie für den Maschinenbediener schwer nachvollziehbar bestimmte Einstellparameter automatisch nachregeln (Black-Box).

Neben den Maschinenhersteller gibt es Startups, wie das Fraunhofer Spin Off plus10 GmbH, welche sich mit der KI-gesteuerten Produktion beschäftigen. Der Qualitätsoptimierer *Hopper* von plus10 kombiniert verschiedene Modelle, um optimierte Einstellparameter vorzuschlagen. Auch hierbei ist die Optimierung für den Maschinenbediener nicht nachvollziehbar. [9]

## 2.2 Datenerfassung

Die Spritzgießbranche ist geprägt von kleinen und mittelständischen Unternehmen, die häufig einen heterogenen, organisch gewachsenen Maschinenpark unterhalten. In den allermeisten Spritzgießunternehmen wird man mindestens zwei unterschiedliche Hersteller von Spritzgießmaschinen finden, was teilweise eine bewusst gewählte Strategie ist, um die Abhängigkeit von einzelnen Maschinenherstellern zu umgehen oder die individuellen Vorteile einzelner Hersteller für spezifische Produkte zu nutzen. In Bezug auf die Datenerfassung bringt dies allerdings das Problem mit sich, dass unterschiedliche Schnittstellen bzw. Softwaresysteme zur Datenerfassung eingesetzt werden müssten. In Realität stellt dies häufig ein derart großes Hindernis dar, dass komplett auf die Erfassung von Prozessdaten verzichtet wird.

In den letzten Jahren hat sich die EUROMAP [10], als Schirmorganisation, die ca. 500 Unternehmen der Kunststoff- und Gummiindustrie vertritt und die relevanten Datenstandards definiert, mit der OPC Foundation [11] zusammengetan und die EUROMAP Standards mit den OPC UA Standards vereinheitlicht. Konkret spricht man von Companion Specifications, die definieren wie z. B. eine Spritzgießmaschine mit einem MES (Manufacturing Execution System) kommunizieren muss, und das herstellerunabhängig. Für die Kommunikation einiger Peripheriegeräte mit der Spritzgießmaschine sind bereits Companion Specifications veröffentlicht worden, z. B. EUROMAP 82.1 für Temperiergeräte (im Jahr 2020) oder EUROMAP 82.2 für Heißkanäle (im Jahr 2021), während andere sich aktuell (Stand Juli 2023) noch in der Spezifizierung befinden wie z. B. EUROMAP 86 für Materialversorgungssysteme.

Diese EUROMAP Schnittstellen erhalten langsam Einzug in die Produktion, entweder durch Nachrüstung oder Updates oder durch Erwerb gänzlich neuer Maschinen und Geräte, auch wenn die Schnittstellen in der Regel noch aufpreispflichtig sind. Das senkt die Hürde die Prozessdaten zu erfassen und mit diesen auch einen Mehrwert zu generieren.

Im Projekt wird, wo möglich, auf die bereits veröffentlichten Companion Specifications zurückgegriffen und somit der Reifegrad und der Mehrwert dieser Schnittstellen überprüft.

Da die Companion Specifications allerdings nur die Menge an Parametern beinhalten, auf die sich alle am Spezifikationsprozess beteiligten Maschinenhersteller einigen konnten, befinden sich aus Sicht der Prozessexperten einige Lücken darin. Beispielsweise sind hier Zeitreihendaten zu nennen, die innerhalb eines Spritzgießzyklus aufgezeichnet werden, wie z. B. der Verlauf einer Spritzdruckkurve. Um diese Daten, die für eine tiefere Prozessbewertung unerlässlich sind, zu erhalten, muss Stand heute auf proprietäre Softwaretools der jeweiligen Maschinenhersteller bzw. von Drittfirmen zurückgegriffen werden. Dies erhöht die Komplexität enorm, weshalb starker Aufwand zur Integration solcher Tools in die allgemeine Datenerfassungsstruktur betrieben werden muss, was in der Praxis häufig dazu führt, dass diese Daten außen vorgelassen werden.

### 2.3 Bayes'sche Netze

Die Modellstruktur eines BN besteht aus einer endlichen Menge von Variablen, die Prozessparameter darstellen. Jede dieser Variablen ist durch eine endliche Anzahl von Zuständen definiert, die sich gegenseitig ausschließen. In der Netzwerkstruktur werden die Variablen durch Knoten und die vorherrschenden Kausalitäten und Abhängigkeiten durch Kanten dargestellt. Jeder Zustand einer Variablen wird durch einen bedingten Wahrscheinlichkeitswert quantifiziert. Mathematisch formuliert, ist ein BN also ein Paar  $(G, \mathbf{P})$ , wobei  $G=(V, E)$  ein gerichteter, azyklischer Graph mit einer Menge von Zufallsvariablen  $V$  sowie einer Menge gerichteter Kanten  $E$  ist.  $\mathbf{P}$  ist eine Menge von bedingter Wahrscheinlichkeitsverteilungen, die die Stärke der Beziehungen in  $E$  angeben. Insbesondere enthält  $\mathbf{P}$  für jede  $V$  in  $V$  die bedingte Verteilung  $P(V|pa(V))$ , wobei  $pa(V)$  die Menge der Elternknoten von  $V$  in  $G$  ist. [12, 13]

Ein großer Vorteil von Bayes'schen Netzwerken ist ihre Fähigkeit zum bidirektionalen Schlussfolgern [14]. Im Produktionskontext ermöglichen sie sowohl die Vorhersage der Produktqualität auf der Grundlage der erfassten Prozessdaten als auch die Ausgabe von Prozessparameterempfehlungen bei identifizierten Qualitätsabweichungen. Ein weiterer Vorteil besteht darin, dass Expertenwissen in den Modellierungsprozess einbezogen werden kann, was bedeutet, dass weniger Daten für das Modelltraining benötigt werden. Es müssen jedoch Vorsichtsmaßnahmen getroffen werden, um die Netzwerktopologie auf physischer Ebene richtig zu gestalten und sicherzustellen, dass die Daten dem in der Topologie definierten Muster entsprechen. Im Gegensatz zu neuronalen Netzwerken, bei denen die Gewichte der Neuronen Zahlen zwischen 0 und 1 sind, repräsentieren die Knoten in einem Bayes'schen Netzwerk physikalische Parameter und können von Domänenexperten interpretiert werden, was die Interpretierbarkeit und Erklärbarkeit des Modells unterstützt.

Die Entwicklung Bayes'scher Modelle für einen Anwendungsbereich ist in der Regel ein iterativer Prozess, bei dem Wissens-elemente aus verschiedenen Quellen zusammengeführt werden müssen. Kjærulf stellt ein Vierphasenmodell zum Aufbau Bayes'scher Modelle vor (siehe Abbildung 1), das beliebig oft durchlaufen werden kann, bis das Modell den Forderungen genügt [14]. In der ersten Phase werden einerseits die notwendigen Variablen mit ihren Zuständen definiert und die topologische Struktur des Netzwerks festgelegt. Andererseits muss verifiziert werden, dass die durch die Struktur bedingten Abhängigkeits- bzw. Unabhängigkeitsannahmen mit dem Expertenwissen des Problem-bereichs konsistent sind.

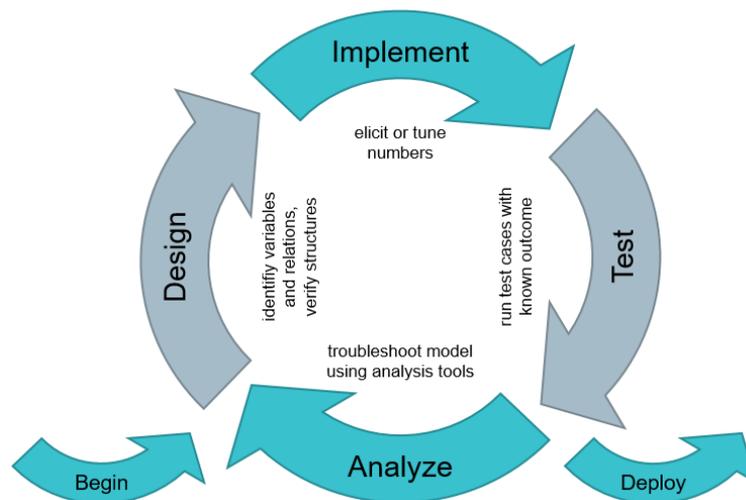


Abbildung 1: Die vier Phasen bei der Entwicklung Bayes'scher Modelle (nach [14]).

Die zweite Phase dient zur Quantifizierung des Netzes, d. h. der Definition der bedingten Wahrscheinlichkeitswerte. In der dritten Phase wird das Verhalten des Modells getestet und validiert. Dies kann mit Hilfe geeigneter Testszenarien bewerkstelligt werden. Die vierte und letzte Phase dient der Analyse des Modells aus verschiedenen Blickwinkeln, wie z. B. mithilfe der einer Sensitivitätsanalyse, mit der bewertet werden kann, wie empfindlich das Modell auf kleine Änderungen der Parameter reagiert. Diese Phasen werden gegebenenfalls mehrmals durchlaufen.

## 2.4 Simulation

Für die Spritzgusssimulation gibt es diverse Softwarelösungen, wie z. B. Autodesk *Moldflow*, CoreTech *Moldex3D*, Simcon *Cadmould* und Sigmasoft *Virtual Molding*, die sich den Großteil des Marktes teilen. Damit haben Konstrukteure die Möglichkeit qualitative und quantitative Analysen und Diagnosen z. B. hinsichtlich Formteifüllung, Luftein-schlüssen, Bindenähten, Faserorientierung oder Schwindung durchzuführen. Im Folgenden sind die essenziellen Aspekte der Simulationsmodelle für den Spritzguss und der Aufbau einer Simulation exemplarisch für die Software *Moldex3D* näher erläutert:

Das digitale Abbild des Werkzeugs (CAD-Modell) ist essenziell, da mit ihm grundlegende Randbedingungen einhergehen. Im Verlauf des Produktentwicklungsprozesses wird das Modell des Werkzeugs dann noch um die Temperierung, die Kavität (Werkzeugs-einsatz), das Angusssystem und das eigentliche Stammwerkzeug ergänzt, um ein möglichst genaues Abbild darzustellen. Um den Rechenaufwand zu minimieren, findet

innerhalb dieses Schritts eine Vereinfachung nicht formgebender Baugruppen, z. B. Verschmelzung von Kontaktflächen, statt. Im Anschluss sind den einzelnen Komponenten entsprechende Attribute zuzuweisen, die Ihre Funktion im Werkzeug spezifizieren, bevor dann die eigentliche Vernetzung (Mesh) der Werkzeug-Geometrie erfolgt. Im Anschluss an die Mesh-Erzeugung kann den attribuierten Formteilen und Werkzeugelementen ein Material zugeordnet werden. Für einen digitalen Materialzwilling des Formteils sind diverse rheologische und thermodynamische Laboruntersuchungen notwendig, mit denen Materialeigenschaften wie z. B. pVT-Verhalten (Druck  $p$ , spezifisches Volumen  $V$ , Temperatur  $T$ ), Viskosität, Wärmeleitfähigkeit, spezifische Wärmekapazität oder Kristallinität zu bestimmen sind. Zu Teilen liegen diese Materialkennwerte bereits in ausreichender Genauigkeit in den Materialdatenbanken der Simulations-Software vor, aber häufig müssen sie durch besagte Laboruntersuchungen bestimmt werden. Zuletzt sind die Berechnungsparameter, in denen unter anderem Solver, Aufzeichnungsintervalle und Materialbesonderheiten wie z. B. die Kompressibilität der Schmelze einzustellen sind, einzustellen.

Teilweise besteht bei den Simulationstools außerdem die Möglichkeit das Ansprechverhalten der physischen Maschine sowie des Plastifizieraggregats auf Basis einer real gemessenen Maschinencharakteristik zu integrieren, um den Sim-to-Real Gap zu verkleinern.

### 3 Lösungsweg zur Erreichung des Forschungsziels

Zur Entwicklung eines Bayes'schen Netzes gehören die Topologie des Netzes sowie eine Quantifizierung mittels Wahrscheinlichkeitsverteilung der betrachteten Daten. Daher sieht der Lösungsweg, wie in Abbildung 2 dargestellt, vor, die topologische Struktur des Netzes mittels Expertenwissen zu ermitteln und gleichzeitig eine Spritzgießzelle für die Generierung realer Spritzgießdaten aufzubauen sowie das Aufsetzen eines Simulationsmodells für die Generierung simulierter Daten. Liegen die topologische Struktur des Netzes und dazu passende Prozessdaten vor, wird eine Quantifizierung des Netzes durchgeführt. Das nunmehr quantifizierte Netz wird dann in einer Live-Demonstration an der Spritzgießmaschine getestet und validiert.

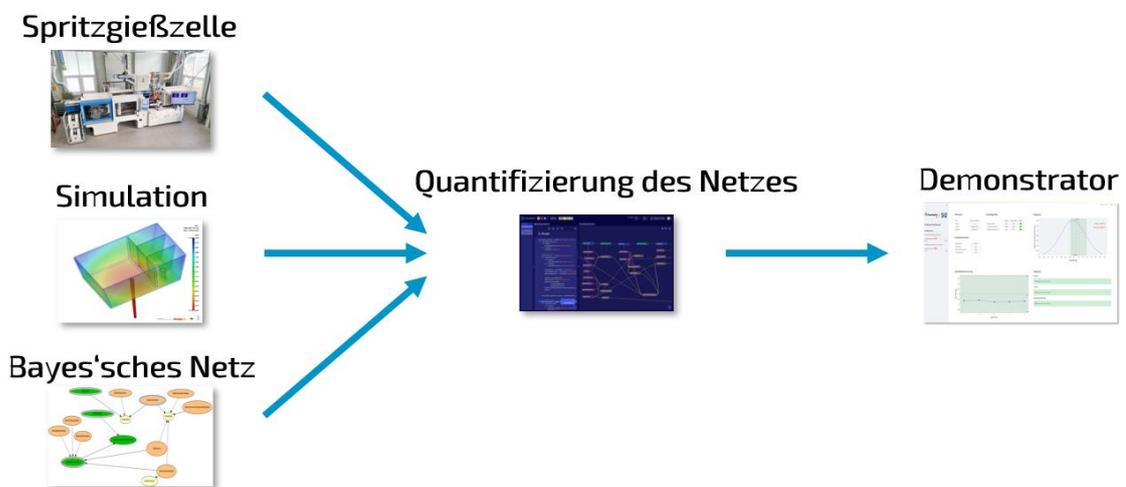


Abbildung 2: Lösungsweg von Datenerzeugung über Quantifizierung bis hin zum Demonstrator, wobei dieser Weg iterativ durchlaufen wurde

Der Projektansatz sah vor, dass die in Abbildung 2 gezeigte Vorgehensweise zunächst in der ersten Projekthälfte schnell durchlaufen wird und ein MVP (Minimum Viable Product) zur Verfügung steht, welches dann in weiteren Iterationsschleifen optimiert werden kann.

Der Grund warum der Aufbau der Spritzgießzelle inkl. Datengenerierung und die Entwicklung der Topologie parallel durchgeführt werden, ist darin begründet, dass ein hohes Maß an Abhängigkeit zwischen beiden Aktivitäten besteht. Da im Projekt erstmalig die gesamten Peripheriegeräte einer Spritzgießzelle datentechnisch angebunden wurden, konnte das Verständnis bzw. die Lehrmeinung der Spritzgießexperten durch reale Daten gegengeprüft werden. Die von den Experten formulierten Abhängigkeiten der einzelnen Einstell-, Prozess- und Qualitätsmerkmale voneinander konnten somit direkt in den Daten nachgewiesen bzw. auch widerlegt werden.

## 4 Durchgeführte Arbeiten

In den folgenden Unterkapiteln sind die durchgeführten Arbeiten dargestellt, geordnet nach den jeweiligen Domänen. Zusätzlich zu dem vorliegenden Bericht wurden im Projekt erhobene Datensätze [15–17] und erarbeiteter Quellcode [18] veröffentlicht, die an verschiedenen Stellen hier im Bericht referenziert sind.

### 4.1 Aufbau der Spritzgießzelle

#### 4.1.1 Spritzgießmaschine und Peripheriegeräte

##### 4.1.1.1 Spritzgießmaschine

Für die Spritzgießzelle wird eine Spritzgießmaschine der Firma KraussMaffei Technologies GmbH des Typs 160-750-PX und der in Tabelle 1 aufgelisteten Spezifikationen genutzt.

Tabelle 1: Spezifikation der Spritzgießmaschine

Hersteller	KraussMaffei
Name	160-750-PX
Steuerung	MC6
Antrieb	Elektrisch
Schließkraft	1600 kN
Holmabstand (horizontal*vertikal)	520 mm*470 mm
Schneckendurchmesser	35 mm

Um die digitale Datenerfassung zu ermöglichen, ist die Maschine mit der Standardschnittstelle EUROMAP77 und der KraussMaffei eigenen Datenerfassungshardware „DataXplorer“ ausgestattet. Das Handlingsystem an der Maschine ermöglicht die automatisierte Bauteilentnahme und die Einbindung von Inline-Qualitätsmesssystemen.



Abbildung 3: Genutzte Spritzgießmaschine im Technikum des SKZ

#### 4.1.1.2 Materialhandling und Temperierung

Das Materialhandling wurde vom Unternehmen MOTAN HOLDING bereitgestellt. Das System besteht aus der zentralen Trocknungsanlage LUXOR an welche die 2 Trockentrichter angeschlossen sind. In der Steuerung der Trocknungsanlage ist ein proprietärer OPC UA Server installiert, welcher zur Datenerfassung genutzt wird.

Um den Restfeuchtegehalt des Kunststoffgranulats zu überwachen, ist das Feuchtemessgerät „Moisture Minder“ am Auslass des Trockentrichters installiert. In Absprache mit Vertretern aus der Industrie wurde evaluiert, welchen ergänzenden Nutzen das verfügbare Feuchtemessgerät im Projekt hat. Dabei wurde erkannt, dass die momentane Genauigkeit und Zuverlässigkeit des Gerätes nicht zur Entwicklung der in diesem Projekt angestrebten Modelle ausreichen. Die Durchführung von manuellen Feuchtemessungen, d. h. einer regelmäßigen Probeentnahme während der Versuchsdurchführung, gewährleistete die Überprüfung des Restfeuchtegehaltes. Für die Feuchtigkeitsmessung des Granulats wurde deshalb das Messgerät der Sartorius AG Germany mit der Bezeichnung MA100C-000230V1 verwendet. Für die Messung werden 20 g Granulat eingewogen und 25 min bei einer Temperatur von 110 °C temperiert.

Zur Temperierung des Spritzguss-Werkzeugs wurden 2 Temperiergeräte des Unternehmens technotrans SE (ehemals gwK Gesellschaft Wärme Kältetechnik mbH) integriert. Die Temperiergeräte sind ebenfalls mit einem OPC UA Server ausgestattet, sodass aktuelle Vor- und Rücklauftemperaturen sowie der Durchfluss ausgelesen werden können.

### 4.1.2 Werkzeugauswahl

Für die Versuchsdurchführung wurde das bestehende Werkzeug „Verzugsschale“ in Abbildung 4 herangezogen. Dies enthält mehrere konstruktive Elemente, die sich zur Qualitätskontrolle eignen. Dazu gehören ein Dom, eine Rippenstruktur sowie ein quadratischer Ausschnitt. Zudem besteht die Möglichkeit über unterschiedliche Temperierkreise verschiedene Verzugsbilder zu generieren.

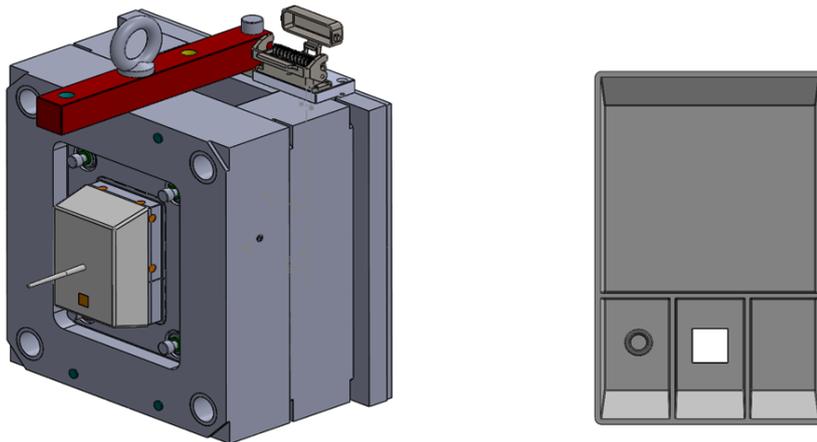


Abbildung 4: Spritzgieß-Werkzeug „Verzugsschale“ (links) und das entsprechende Bauteil (rechts)

Das Werkzeug ist jeweils mit einem kavitätsnahen Thermofühler auf der Düsen- und Auswerferseite ausgestattet. Die Auswerferseite enthält weiterhin einen angussnahen Drucksensor. Die präzise Messung des Werkzeuginnendruckes ist beim Spritzgießen von zentraler Bedeutung. Sein Verlauf gibt Auskunft über die Qualität jedes einzelnen Spritzgießteils und dient dazu, den Prozess zu überwachen und eine nachhaltige Qualitätssicherung zu erzielen. In diesem Fall bestand die Möglichkeit die Sensoren an die Spritzgießmaschine anzuschließen und die Daten über den DataXplorer von KraussMaffei Technologies GmbH auszulesen.

Um das Bauteil automatisiert zu entnehmen wurde ein bauteilspezifischer Greifer konstruiert und additiv gefertigt (vgl. Abbildung 5). Der Greifer entnimmt das Bauteil mittels Unterdruck und fährt dieses zu den entsprechenden Inline-Qualitätsmesssystemen.

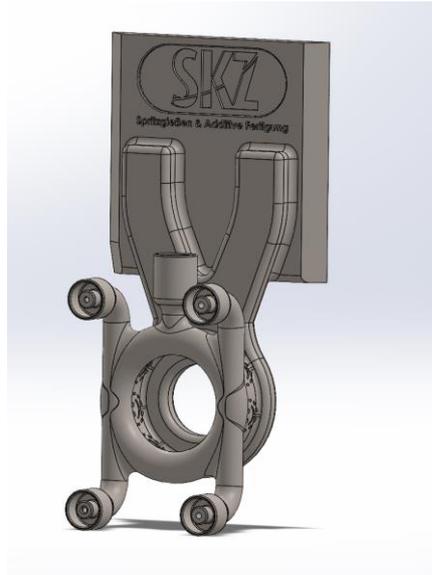


Abbildung 5: CAD-Modell des konstruierten Greifers zur automatisierten Bauelemententnahme

### 4.1.3 Qualitätsmesssysteme

Zur Entwicklung einer Live-Prozessregelung ist es notwendig, auch die Qualitätsmerkmale des Bauteils Inline zu erfassen, um darauf basierend Rückschlüsse auf die Einstellparameter zu ziehen. In die Spritzgießzelle am SKZ wurde deshalb ein Bildverarbeitungssystem, eine Waage sowie ein Thermographiemesssystem integriert. Das Handlingsystem der Spritzgussmaschine transportiert das Bauteil zu den einzelnen Systemen und triggert bei Ankunft an der Messposition die Messung. Anschließend erfolgt die Ablage auf dem Förderband und die Entnahme des darauffolgenden Bauteils.

#### 4.1.3.1 Thermographie

Für die Inline-Thermografiemessung wird eine Wärmebildkamera der Firma MICRO-EPSILON MESSTECHNIK GmbH & Co. KG vom Typ TIM 640 mit 33°-Optik verwendet. Zur Kommunikation mit der Kamera wird die Software TIM Connect verwendet, die Darstellung der Messergebnisse geschieht über die SKZ eigene Software SKZ TDI V3.6.0.3000323. Die Kamera ist in der Einhausung des Handlings positioniert. Um eine möglichst gute Aussage über die Entformungstemperatur zu erhalten, steuert das Handling die Thermographiekamera direkt nach der Bauelemententnahme an, wo mehrere Thermogramme im Abstand eines festgelegten Zeitintervalls erstellt werden.

Tabelle 2: Thermographiekamera von Micro-Epsilon Messtechnik GmbH & Co. KG

Eigenschaft	Name / Wert
Hersteller	Micro-Epsilon Messtechnik GmbH & Co. KG
Bezeichnung	TIM 640
Software	TIM Connect

Temperaturmessbereiche (umschaltbar)	-20°C - 100°C, 0°C - 250°C, 150°C - 900°C
Spektralbereich	7,5 -13 $\mu\text{m}$
Optische Auflösung	32 Hz@640x480 Pixel oder 125 Hz@640x120 Pixel
Thermische Auflösung	0,1 °C
Objektivbrennweite	f=18,4 mm / 33°-Objektiv

#### 4.1.3.2 Bildverarbeitungssystem

Das Bildverarbeitungssystem des Unternehmens ZIEMAN & URBAN GmbH ist in Abbildung 6 zu sehen und besteht aus einem Standgehäuse, den Kameras und der Belichtung sowie einem Bedienterminal.

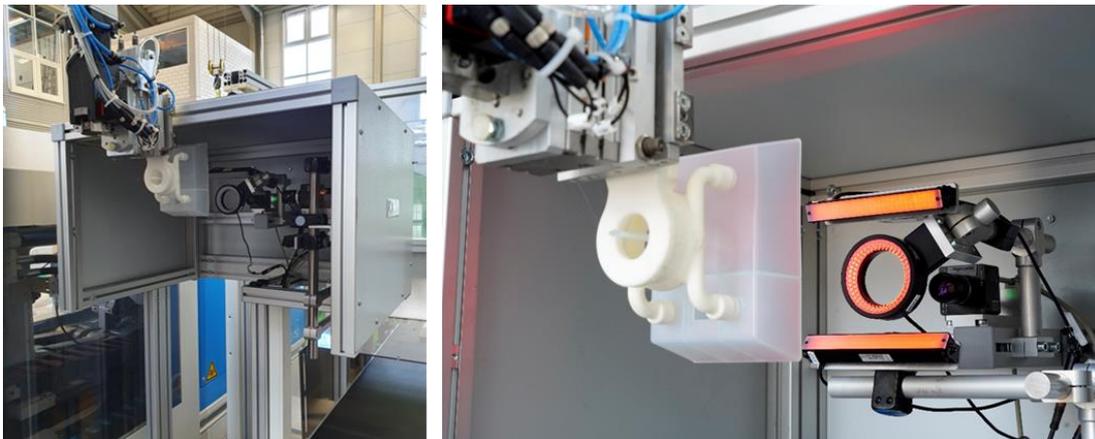


Abbildung 6: Bildverarbeitungssystem des Unternehmens Ziemann & Urban GmbH

Die Auswertung des aufgenommenen Bilds erfolgt über die Software von Ziemann & Urban. Das Programm enthält eine Vielzahl von Möglichkeiten das Bild zu analysieren und entsprechende Maße zu extrahieren. Über den proprietären OPC UA Server ist es möglich die Maße eindeutig einem Zyklus zuzuordnen. Die Vorgehensweise der Datenfusionierung von Spritzgussmaschine und Bildverarbeitungssystem ist in 4.1.4 beschrieben.

#### 4.1.3.3 Gewichtsmessung

Zur Erfassung des Qualitätsmerkmals Gewicht wurde in die Spritzgießzelle eine Waage des Typs „NewClassic MF“ der METTLER TOLEDO GMBH integriert (Abbildung 7). Das zusätzlich installierte edge device „IIoT Device ACI400 Edge“ ermöglichte die Übertragung des Gewichtswertes via OPC UA. Der Roboter legt das Bauteil auf der Waage ab, nimmt dieses nach Ermittlung des Gewichtswertes erneut auf und legt es anschließend auf dem Förderband ab.



Abbildung 7: Integrierte Industriewaage des Unternehmens Mettler-Toledo zur Aufzeichnung des Gewichtwertes

#### 4.1.3.4 Umweltsensor

Zur Überwachung der Umgebungstemperatur und -feuchtigkeit wurde ein Umgebungs-sensor entwickelt und in der Nähe der Thermographie Kamera platziert. Das dafür genutzte Developer Board ESP32-PoE-ISO-IND ist mit einem DHT22 Feuchte-Temperatur Sensor sowie einem Display (SH1107) ausgestattet. Dadurch ist es möglich unerwartete Bauteilfehler oder abweichende Entformungstemperaturen auf mögliche Veränderungen der Umgebungstemperaturen zurückzuführen.

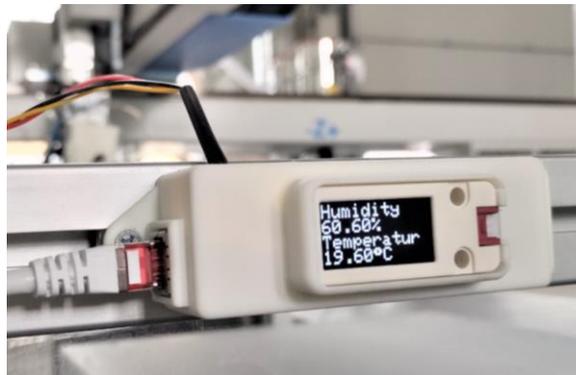


Abbildung 8: Im Projekt entwickelter Umweltsensor zur Aufzeichnung von Temperatur und Feuchtigkeit

#### 4.1.4 Datenerfassung und Vernetzung

Grundsätzlich wurde der Ansatz verfolgt, alle Parameter, die durch die verschiedenen Geräte zur Verfügung gestellt werden, in voller Auflösung und unkomprimiert aufzuzeichnen. Dies hat drei Gründe:

1. Zu Beginn des Projekts war unklar, welche Geräte welche Parameter aufzeichnen und dann auch nach außen zur Verfügung stellen. Dies ist abhängig vom Geräte-Typ, von der konkreten Konfiguration bzw. Ausstattung des Geräts und vom Hersteller selbst. Zudem werden nicht alle intern gemessenen oder berechneten Werte nach außen über Schnittstellen zugreifbar gemacht.
2. Für die Entwicklung eines robusten Prozessmodells muss auf vielerlei Parameter zugegriffen werden. Welche die relevanten, beeinflussenden Prozessgrößen für

die jeweiligen Qualitätsparameter sind, war zu Beginn noch unklar. Zudem ist es denkbar, dass einige Parameter nur aus anderen berechnet oder abgeschätzt werden können, z. B. die Schmelzeviskosität, oder Parameter aus Zeitreihen extrahiert werden müssen, wie z. B. das Integral des Werkzeuginnendrucks.

3. Die Spritzgießzelle soll als Versuchsplattform für ProBayes und auch darüber hinaus genutzt werden. Auch für Anwendungsfälle und Forschungsinhalte, die bis dato noch unbekannt sind. Daher ist es sinnvoll die Datenerfassung möglichst breit aufzustellen.

#### 4.1.4.1 Middleware

Herz der Datenerfassung ist die *Connectware* der Firma Cybus GmbH. Hierbei handelt es sich um eine Middleware, welche mittels eines Protocol-Mappers verschiedenste Kommunikationsprotokolle, wie OPC UA oder Modbus, auf MQTT Topics mappen kann. Diese Topics sind frei definierbar und können in mehreren Hierarchieebenen organisiert werden. Die Daten werden auf dem jeweiligen Topic gepublished und können von einem oder mehreren Clients abonniert werden. Die *Connectware* selbst besteht aus einzelnen Microservices, welche in separaten *Docker*-Containern laufen. Dabei können von Cybus GmbH zur Verfügung gestellte Docker-Images genutzt werden, wie beispielsweise einem OPC UA Client, oder öffentlich verfügbare (z. B. auf Dockerhub.com) Images wie *InfluxDB* oder *Grafana*. Dadurch kann die *Connectware* vielseitig eingesetzt und skaliert werden, was im Vorhaben ProBayes auf folgende Weise genutzt wurde.

Für jedes Gerät wurde ein eigener Service angelegt, der den zum Kommunikationsprotokoll passenden Client beinhaltet und das Mapping auf die MQTT Topics definiert. Des Weiteren wurde ein separater Service für die Daten eines Geräts zur Datenbank implementiert. Dieser vollmodulare Ansatz wurde bevorzugt, denn eine Separierung der Services ermöglicht eine schnelle Fehlerfindung, z. B. bei Verbindungsproblemen einzelner Geräte. Außerdem wird dadurch vermieden, dass Probleme einzelner Geräte die Datenübertragung anderer beeinträchtigen.

Durch die *Connectware* werden die ankommenden Daten noch um zusätzliche Attribute erweitert, bevor sie per MQTT veröffentlicht werden, sodass eine Datenstruktur wie in Tabelle 3 entsteht.

Tabelle 3: Datenstruktur umgesetzt in der Cybus *Connectware*

	Beispiel
Topic	TemperatureControl/Auswerferseite/MainTemperature
Value	50.1
Timestamp	1656590559227
Datatype	FLOAT

#### 4.1.4.2 Zykluszuordnung & Zeitsynchronisation

Die meisten der verwendeten Geräte, wie z. B. die Spritzgießmaschine oder das Bildverarbeitungssystem, erheben Daten, die direkten Bezug zu einem Spritzgießzyklus haben. Daher war es nötig, eine Zyklus-ID einzuführen, die es ermöglicht jeden Zyklus eindeutig zu identifizieren. Dazu wurde der über die Spritzgießmaschine (EUROMAP77) verfügbare MachineCycleCounter herangezogen. Hierbei handelt es sich um einen Zähler, welcher zu Beginn eines neuen Spritzgießzyklus, d. h. beim Schließen des Werkzeugs, inkrementiert wird. Diese im Weiteren als Zyklus-ID bezeichnete Zahl kann nicht zurückgesetzt werden, ist nicht manipulierbar und garantiert somit eine eindeutige Zuordnung.

Neben der Zyklus-ID ist auch die Zeit, wann ein Spritzgießzyklus begonnen hat, eindeutig. Dies setzt jedoch voraus, dass alle Geräte die gleiche, d. h. synchrone Uhrzeit haben. Dies wurde umgesetzt, indem ein NTP-Server (Network-Time-Protocol) im Netzwerk des SKZ installiert wurde, von dem die verwendeten Geräte ihre Uhrzeit beziehen können. Nicht alle der eingesetzten Geräte waren in der Lage ihre Uhrzeit auf den NTP-Server zu synchronisieren. Dies ist einer der Gründe, warum die in Tabelle 3 gezeigte Datenstruktur auch den Zeitstempel enthält. Dieser entspricht dem Zeitpunkt, wann der jeweilige Datenpunkt bei der *Connectware* angekommen ist. Da die Uhrzeit der *Connectware* synchron zur Uhrzeit des NTP-Servers ist, kann dieser Zeitstempel neben der Zyklus-ID ebenfalls als eindeutiger Identifikator für eine Zykluszuweisung dienen, sofern die Zeitverzögerung von Zykluswechsel bis zur Ankunft der Messdaten bekannt ist.

#### 4.1.4.3 Datenpersistierung

Die Datenspeicherung erfolgte nach dem Grundsatz: „Jedem Datentyp die passende Art von Datenbank / Datenspeicherung“. Daher entstanden mehrere Datenspeicherorte, welche im Nachgang über die Zyklus-ID bzw. den Zeitstempel verknüpfbar sind, wie in Tabelle 4 dargestellt.

Tabelle 4: Datenpersistierung in Abhängigkeit des Datentyps

Datentyp	Speicherort	Bemerkung
Thermogramme	Dateibasiert auf Netzlaufwerk	
Bilder aus Bildverarbeitung	Dateibasiert auf Ziemann-Urban System	Der Dateiname beinhaltet die Zyklus-ID, welche über OPC UA empfangen wird.
Aus Bilddaten extrahierte Maße des Bauteils	InfluxDB	Von Ziemann-Urban System per OPC UA gepublished
OPC UA Daten aus allen Geräten	InfluxDB	

Zeitreihendaten aus Spritzgießmaschine	Dateibasiert auf Spritzgießmaschine bzw. Netzlaufwerk	Automatisierte Extraktion via Softwaretool „ibaDatCoordinator“ und Ablage als „PARQUET-Datei“
Daten des Umweltse-nsors	InfluxDB	Per MQTT gepublished

Für alle Daten, welche via OPC UA empfangen wurden, wird die Zeitreihendatenbank InfluxDB ausgewählt, welche optimiert ist für die Aggregation von Zeitreihendaten. Da die Daten aus unterschiedlichen Quellen (also den unterschiedlichen Geräten) kommen, wird für jede Datenquelle ein separater sog. Bucket, bereitgestellt. Somit ist eine Zuordnung der Daten zum Quell-Gerät jederzeit möglich.

#### 4.1.4.4 Spritzgießmaschine

Die Maschine verfügt über mehrere Datenschnittstellen, welche alle gleichsam genutzt werden. Zunächst lässt sich die Qualitätstabelle (Istwerttabelle) über USB auslesen, welche einen guten Überblick über die durchgeführten Zyklen ermöglicht. Zusätzlich ist die Maschine mit der EUROMAP77 ausgestattet, welche die OPC UA Companion Specification OPC 40077 umsetzt. Diese Schnittstelle ist hauptsächlich für die Kommunikation von MES (Manufacturing Execution System) mit Spritzgießmaschinen gedacht, weshalb sie viele Meta-Informationen, beispielsweise die Art und Konfiguration der Maschine und auftragsbezogene Daten etc, enthält. Diese OPC UA Nodes können von einem OPC UA Client abonniert (Subscription) werden und liefern immer dann einen neuen Datenpunkt, wenn sich der Wert eines Parameters ändert. Des Weiteren können über die EUROMAP77 sogenannte OPC UA Events empfangen werden, welche durch ein Ereignis an der Maschine, z. B. einen Alarm oder das Ende des Spritzzyklus, ausgelöst werden. Als dritte Schnittstelle stellt die Spritzgießmaschine die Prozessgrößen der Maschine über den sog. *DataExplorer* zur Verfügung. Diese sind in 5 Millisekunden Zeitintervallen aufgelöst und stellen die Verläufe der einzelnen Prozessgrößen, wie z. B. den Einspritzdruck, dar. Diese Daten werden direkt auf einem Rechner, der Teil der SGM ist, gespeichert. Auch ist es möglich mit Hilfe der Software *ibaDatCoordinator* (und entsprechender dem SKZ verfügbarer Lizenz) eine Extraktionsaufgabe anzulegen, welche die Daten aus dem proprietären Dateiformat in das *Parquet*-Format umwandelt und die Daten auf einem Netzlaufwerk oder in eine Datenbank ablegt.

#### 4.1.4.5 Temperiergerät

Die beiden Temperiergeräte verfügen ebenfalls über OPC UA Server, welche die EUROMAP82.1 umsetzen. Diese gibt Parameter wie Durchflussrate, Vorlauf- und Rücklauftemperaturen uvm. in einem sekundlichen Takt aus.

#### 4.1.4.6 Trocknungssystem

Teil des Trocknungssystems sind die Trockentrichter, der Trockenlufterzeuger und der Dosierer sowie das Inline-Feuchtemessgerät wie in Abbildung 9 zu sehen. Diese sind

intern, u.A. über Modbus, miteinander in Kommunikation und bieten nach außen hin einen proprietären OPC UA Server, der alle Daten zur Verfügung stellt.



Abbildung 9: Bestandteile des Trocknungssystems und deren Vernetzung

#### 4.1.4.7 Bildverarbeitungssystem

Das Bildverarbeitungssystem der Firma Ziemann-Urban verfügt über einen OPC UA Server mit beschreibbaren Variable Nodes. Daher wird die aus der Spritzgießmaschine verfügbare Zyklus-ID an einen Node des OPC UA Servers des Bildverarbeitungssystems gesendet. Das Bildverarbeitungssystem weiß somit, welchen Zyklus die Spritzgießmaschine aktuell fährt, und kann entsprechend die Zyklus-ID dann in die Benennung der einzelnen Bilddateien sowie der Log-Datei, welche die von den Bildern extrahierten Maße enthält, direkt integrieren. Zusätzlich werden diese extrahierten Maße per OPC UA gepublished und via *Connectware* in einem eigenen InfluxDB Bucket gespeichert. Die Bilddateien selbst werden nicht übertragen und müssen manuell nach abgeschlossenen Versuchsreihen kopiert werden.

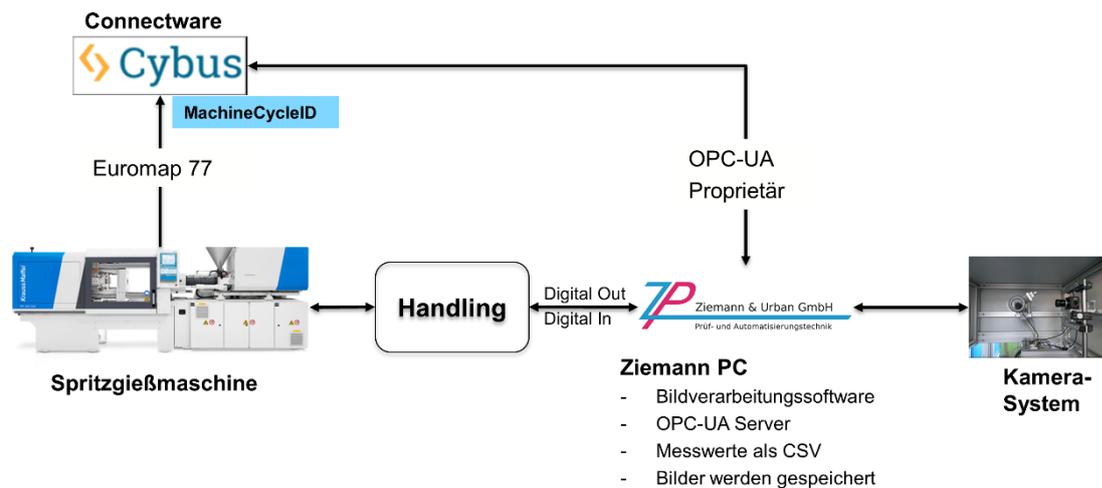


Abbildung 10: Kommunikation des Bildverarbeitungssystems mit *Connectware* und Spritzgießmaschine

#### 4.1.4.8 Thermographiemesssystem

Die von der Thermographiekamera aufgenommenen Thermogramme werden in der SKZ eigenen Software SKZ TDI V3.6.0.3000323 verarbeitet und als csv-Datei gespeichert, sodass für jeden Pixel einen Temperaturwert in °C vorhanden ist. Im Forschungsvorhaben wurde ein Python-Skript zur automatisierten Extraktion der relevanten Features des Thermogramms, wie Gradienten oder Min-Max Werte entwickelt (vgl. Kapitel 4.6.2). Für den Live-Betrieb können die extrahierten Features in einem lokal auf dem Messrechner laufenden MQTT Client an die *Connectware* gepublished werden.

#### 4.1.4.9 Umweltsensor

Mithilfe der Entwicklungsumgebung Arduino IDE wurde ein ESP32 programmiert, so dass er alle 30 s die Luftfeuchtigkeits- und Temperaturwerte des Sensors abfragt. Auf dem Display (vgl. Abbildung 8) werden die aktuellen Werte für den Maschinenbediener angezeigt und gleichzeitig per MQTT gepublished. Die *Connectware* fungiert hier als MQTT Broker, der die Daten des ESP32 subscribed und wiederum an eine InfluxDB zur Persistierung weiterleitet.

## 4.2 Entwicklung der topologischen Struktur des Bayes'schen Netzes

Der erste Schritt in der Entwicklung einer Netztopologie ist die Analyse der vorherrschenden Ursache-Wirkungs-Beziehungen. Ein klassisches Ursache-Wirkungs-Diagramm, in dem die Einflussgrößen mit Pfeilen mit ihrer Zielgröße verbunden werden, erwies sich aufgrund der hohen Prozesskomplexität als zu unübersichtlich. Es wurde daher ein Excel-Template erarbeitet, in dem zum einen der Prozess in seine einzelnen Phasen untergliedert und zum anderen die Wirkmechanismen von der Maschineneinstellung bis zum Qualitätsmerkmal des fertigen Formteils abgebildet werden können. Dafür wurden die jeweiligen Prozessparameter eingeteilt in Einstell-, Prozess- und Qualitätsgrößen. Abbildung 11 zeigt exemplarisch einen Ausschnitt aus der erarbeiteten Einflussmatrix. Die Einstell-

größen in der ersten Spalte entsprechen den Maschineneinstellungen, die von Anlagenbedienern manuell an der Spritzgussanlage vorgenommen werden. Durch sie werden Prozessgrößen beeinflusst, die wiederum das Qualitätsmerkmal des Formteils beeinflussen. Diese Wirkzusammenhänge werden durch entsprechende Kreuze in der Matrixstruktur dokumentiert. Ihr Aufbau ermöglicht auch die Abbildung direkter Abhängigkeiten zwischen Qualitätsmerkmalen und Einstellgrößen oder Abhängigkeiten von Prozessgrößen untereinander, die gestrichelt dargestellt sind. Von den Qualitätsmerkmalen des Formteils (blau) werden die Zwischenqualitätsmerkmale (grün), die unter Experten als Indikatoren für die Prozessqualität gelten, unterschieden. Sie werden in der Matrix neben den Qualitätsmerkmalen erfasst, da sie häufig teilprozessübergreifend von Einstell- und Prozessgrößen beeinflusst werden. Der Werkzeuginnendruck wird beispielsweise von Einstell- und Prozessgrößen der Phasen Einspritzen, Nachdrücken und Kühlen beeinflusst. Die Massetemperatur wird von Maschineneinstellungen im Teilprozess Dosieren beeinflusst. Da sie während des Abkühlens wiederum Auswirkungen auf das Qualitätsmerkmal Gewicht des Formteils hat, wird sie dort in der ersten Spalte mitaufgenommen, farblich unterschieden von den Einstellgrößen.

Einflussgrößen				Q-Merkmal		
Einstellgrößen	Prozessgrößen			Werkzeuginnendruck	Masstemperatur	Gewicht
Teilprozess Kühlen						
Masstemperatur	x	x			<del>x</del>	x
Soll-Temperatur (DS)	x					
Soll-Temperatur (WS)		x				
Restkühlzeit	----->					
	Werkzeugtemperatur (DS)	----->		x		x
		Werkzeugtemperatur (AS)		x		x
			Durchflussrate			

Abbildung 11: Exemplarischer Ausschnitt aus der erarbeiteten Einflussmatrix

Parallel zum zuvor beschriebenen wissensbasierten Ansatz wurden die Ursache-Wirkungs-Beziehungen auch datenbasiert untersucht. Hierzu wurden verschiedene Feature Selection Algorithmen genutzt, die in der Lage sind, die für eine definierte Zielgröße wichtigsten Einflussfaktoren (sog. „Feature Importance“) zu quantifizieren. Mit der Durchführung einer datenbasierten Feature-Auswahl wurden mehrere Ziele verfolgt. Zum einen sollte eine Bestätigung und Nachschärfung der in der Einflussmatrix festgehaltenen Ergebnisse aus den Expertenworkshops erfolgen. Zum anderen konnten so auch nachträglich berechnete Kenngrößen in die Auswahl einbezogen werden, über die eine wissensbasierte Aussage nur schwer möglich ist. Außerdem ermöglichten die verschiedenen Feature Selection Verfahren eine objektive Bewertung der Relevanz von Einflussgrößen, z. B. durch statistische Kenngrößen und sollten in der Einflussmatrix möglicherweise nicht beachtete Beziehungen aufdecken. Das detaillierte Vorgehen bei der Feature Selection ist in Kapitel 4.7 beschrieben.

Die Ergebnisse der datenbasierten Analyse wurden anschließend mit den im Excel-Template festgehaltenen Ursache-Wirkungs-Beziehungen abgeglichen und mit den Prozessexperten des SKZ diskutiert. Besonders im Fokus standen dabei die Plausibilität der datenbasierten Ergebnisse sowie Diskrepanzen in den Ergebnissen aus den unterschiedlichen Ansätzen. Durch die Kombination von wissens- und datenbasiertem Ansatz entstanden schließlich die in Abbildung 12 gezeigte Struktur für das Teilnetz Gewicht und die Abbildung 13 gezeigte Struktur für das Teilnetz Verzug. Durch die Kombination der beiden Teilnetze ergibt sich das in Abbildung 14 dargestellte Gesamtnetz.

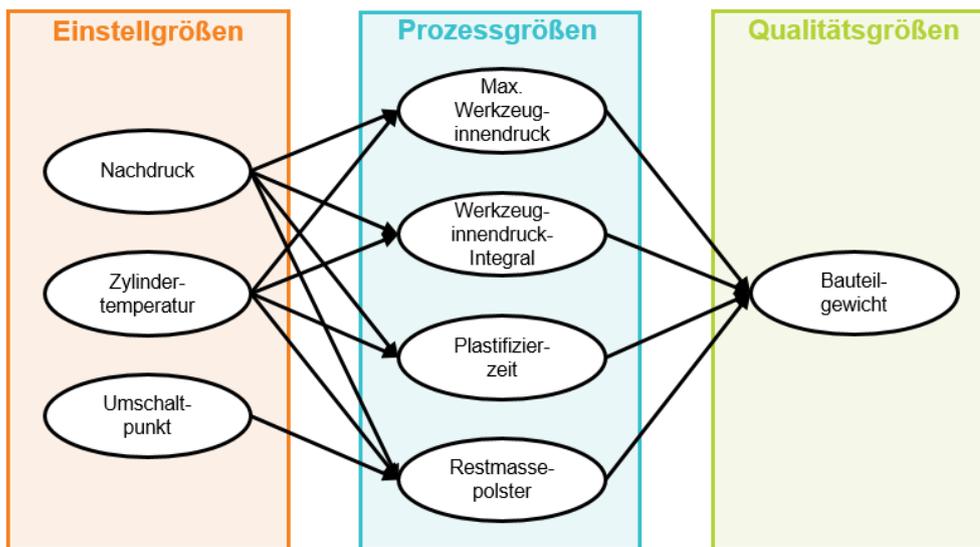


Abbildung 12: Teilnetz für den Qualitätsindikator „Bauteilgewicht“

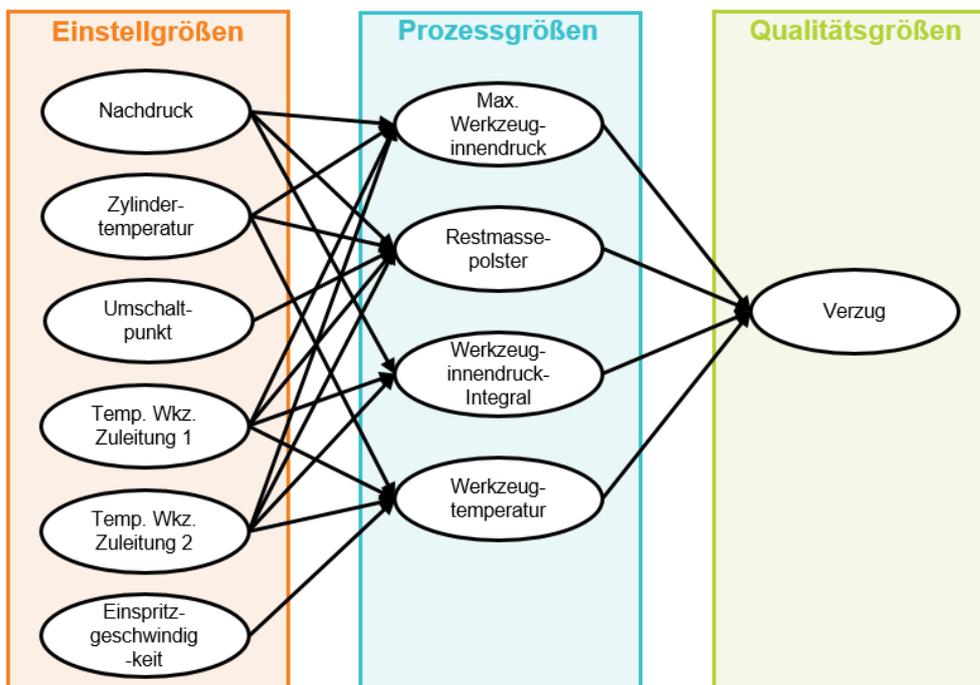


Abbildung 13: Teilnetz für den Qualitätsindikator „Verzug“

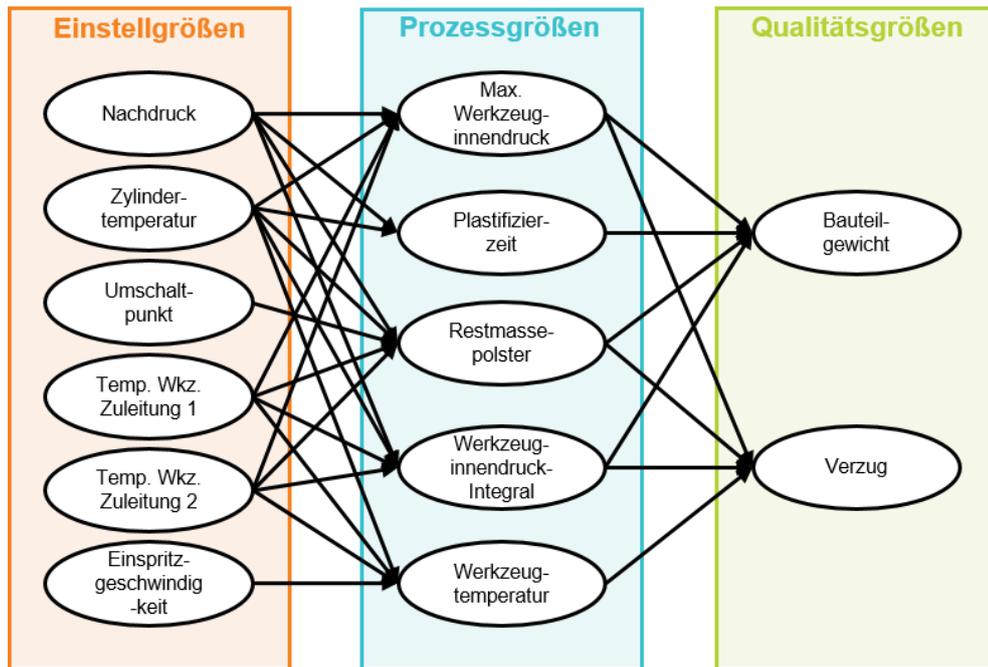


Abbildung 14: Gesamtnetz für die Qualitätskriterien „Bauteilgewicht“ und „Verzug“

### 4.3 Vorversuche

Für die Versuchsdurchführung wurde das in 4.1.2 beschriebene Werkzeug „Verzugschale“ ausgewählt. Das Ziel der Vorversuche bestand im Identifizieren der Parametergrenzen der ausgewählten Kombination aus Werkzeug und Material, dem Testen der integrierten Datenerfassung, dem Festlegen der Zielgrößen bzw. Qualitätsparameter, sowie der Generierung eines ersten Datensatzes für das Trainieren des Bayes'schen Netzes.

#### Materialauswahl

Bei der Werkstoffauswahl lag der Fokus auf einer optimalen Verarbeitbarkeit basierend auf den technischen Eigenschaften sowie einer hohen industriellen Relevanz. Aus diesem Grund wurden unverstärktes Polypropylen (PP) und Arcylnitril-Butadien-Styrol (ABS) ausgewählt.

Die Bereitstellung des teilkristallinen Standardkunststoffs PP mit dem Handelsnamen HE125MO erfolgte durch das Unternehmen Borealis AG. Das PP ist sehr leichtfließend. Dies ermöglicht ein breites Parameterfenster und einen erweiterten Handlungsspielraum für die Entwicklung des Bayes'schen Netzes bzw. des Demonstrators.

Der amorphe Kunststoff ABS, mit dem Handelsnamen Terluran® ECO GP-22 MR70, repräsentiert einem technischen Thermoplast, der in verschiedenen Branchen wie der Haushalts-, Automobil- und Elektroindustrie weit verbreitet ist. Das verwendete ABS weist einen recycelten Anteil von 70% auf.

#### Versuchsplan

Die DOE-Software *MESOS* unterstützte beim Aufsetzen eines vollfaktoriellen Versuchsplans ( $2^4+1$ ), dargestellt in Tabelle 5 und Tabelle 6 mit insgesamt je 17 Versuchsreihen

für beide Materialien. Der Versuchsplan variiert die 4 Einstellparameter, welche die Qualität des Bauteils am stärksten beeinflussen. Dazu zählt die Zylindertemperatur, die Werkzeugtemperatur, die Einspritzgeschwindigkeit sowie der Nachdruck.

Tabelle 5: Parameterbereich der Vorversuche - Material PP

	Minimaler Wert	Zentralpunkt	Maximaler Wert
Zylindertemperatur [°C]	200	225	250
Werkzeugtemperatur [°C]	20	25	50
Einspritzgeschwindigkeit [ccm/s]	15	58	100
Nachdruck [bar]	60	130	200

Tabelle 6: Parameterbereich der Vorversuche - Material ABS

	Minimaler Wert	Zentralpunkt	Maximaler Wert
Zylindertemperatur [°C]	220	240	260
Werkzeugtemperatur [°C]	30	55	80
Einspritzgeschwindigkeit [ccm/s]	20	50	80
Nachdruck [bar]	200	275	350

In den Versuchsreihen traten unterschiedliche Fehlerbilder auf, die teilweise durch die gewählten Parametereinstellungen absichtlich provoziert wurden. Die in Abbildung 15 dargestellten Fehlerbilder stellen die Klassifikationslabel dar, womit jeder Spritzgießzyklus im Datensatz annotiert ist. Es wurde das Label OK/NOK definiert, wobei ein Bauteil das Label OK erhält, wenn keiner der Fehlerfälle vorliegt.



Abbildung 15: Weitere erzeugte optische Fehlerbilder am Bauteil „Verzugsschale“

#### 4.4 Aufbau eines Simulationsmodells und Quantifizierung des Sim-to-Real Gap

Auf Basis der Ergebnisse aus den rheologischen Untersuchungen sollen die Zusammenhänge der einzelnen Parameter untersucht und die Verteilungen für die Quantifizierung der Bayes'schen Netze genutzt werden. Die Simulationen wurden für diesen Zweck in Moldex3D modelliert und berechnet, die Ergebnisse anschließend exportiert, um sie dann in das Bayes'sche Modell einzupflegen.

Im Folgenden soll auf die Arbeitsschwerpunkte hinsichtlich Modellierung der Randbedingungen, Evaluation des digitalen Zwillings und den Optimierungsiterationen näher eingegangen werden.

##### 4.4.1 Modellierung der Randbedingungen in Moldex3D

###### Geometrisches Modell:

Für die Modellierung des geometrischen Zwillings wurde für das Simulationsmodells der *Verzugsschale* das Formteil inkl. der Anbindung, den Temperierkanälen sowie Position des Werkzeuginnendruckensors aus dem CAD-Modell entnommen.

Die Netzgröße wurde hierfür nach der *nodes with enough space* Methode ausgelegt, bei der das Formteil so auf den Bildschirm zu positionieren ist, dass die längste Seite horizontal auf dem Bildschirm zu sehen ist. Die zweitlängste Seite ist in dessen vertikal dazu auszurichten. Abschließend ist das ausgerichtete Formteil im Bedienfenster mittels heran oder hinweg zoomen so zu positionieren, dass es so viel Raum wie möglich einnimmt, ohne bereits Teilbereiche abzuschneiden. Abschließend ist der Abstand zwischen zwei Netzpunkten so zu wählen, dass zwischen diese noch zwei bis drei weitere Netzpunkte passen würden. Abbildung 16 zeigt für die *Verzugsschale* die gedanklich zu positionierenden Punkte bei einem Netzabstand von 2 mm.

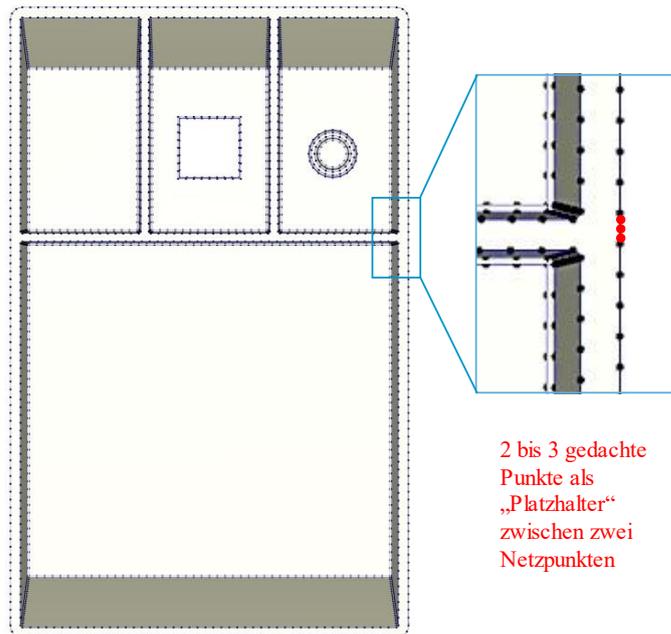


Abbildung 16: Festlegung des Netzabstands für das Formteil *Verzugsschale* nach der *nodes with enough space* Methode

Die folgende Abbildung 17 zeigt das vernetzte Formteil nach Anwendung des beschriebenen Vorgehens. Aufgrund der gleichmäßigen und flächigen Geometrie ist keine weitere, lokale Anpassung des Netzabstands notwendig.

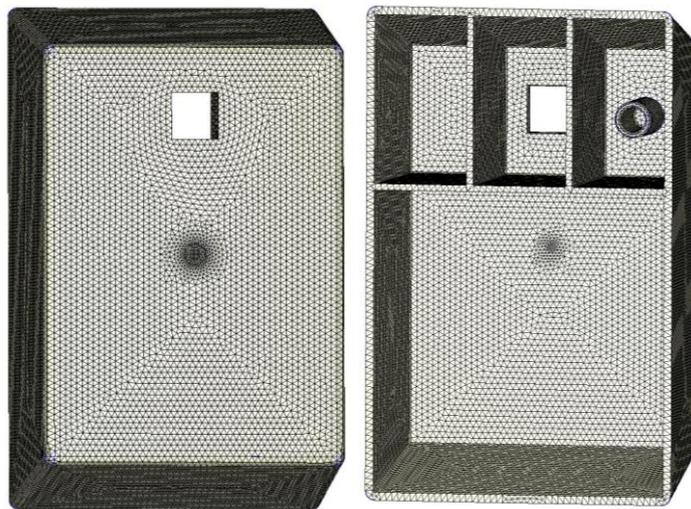


Abbildung 17: Darstellung des vernetzten Formteils (Netzabstand = 2 mm)

Zur Vermeidung von Rechenfehlern und zur Steigerung der Recheneffizienz wurden aus Anbindung und Temperierkanälen die Mittellinien extrahiert, und in Moldex3D parametrisch wieder modelliert. Hierbei spricht man vom sogenannten kurvenbasierten Vernetzen, bei dem die Elementzahl reduziert und damit die Rechenzeit optimiert werden kann, ohne qualitative Einschränkungen in der Ergebnisvorhersage hinnehmen zu müssen. Die nachfolgende Abbildung 18 stellt die unterschiedlichen Vernetzungsarten dar, unter anderem auch die eben erwähnte kurvenbasierte (Vgl. Erstarrender Anguss, Düsen-Dummy, Temperierkanal oder Verbindungsbrücke). Für das Formteil kommt das BLM-Mesh (Boundary Layer Method) zum Einsatz, welches dafür optimiert ist, die Interaktionen

zwischen Kavität und Schmelze an der Grenzschicht bzw. in der näheren Umgebung, um diese zu untersuchen. Verbunden werden diese BLM-Schichten durch eine Tetraeder-schicht, die als Koppelschicht dient. Der Werkzeug-Dummy selbst wird rein durch die zuletzt beschriebenen Tetraeder abgebildet. Diese sind für die notwendigen thermische Berechnungen ohne Fluidkontakt, wie sie im Werkzeug vorliegen ausreichend. Im Werkzeug verbaute Sensorik wird mittels im dreidimensionalen Raum definierten Sensorpunkten abgebildet, die zu einem variablen Zeitpunkt an der definierten Position den Wert für beliebige Mess- oder Qualitätsgrößen ermitteln.

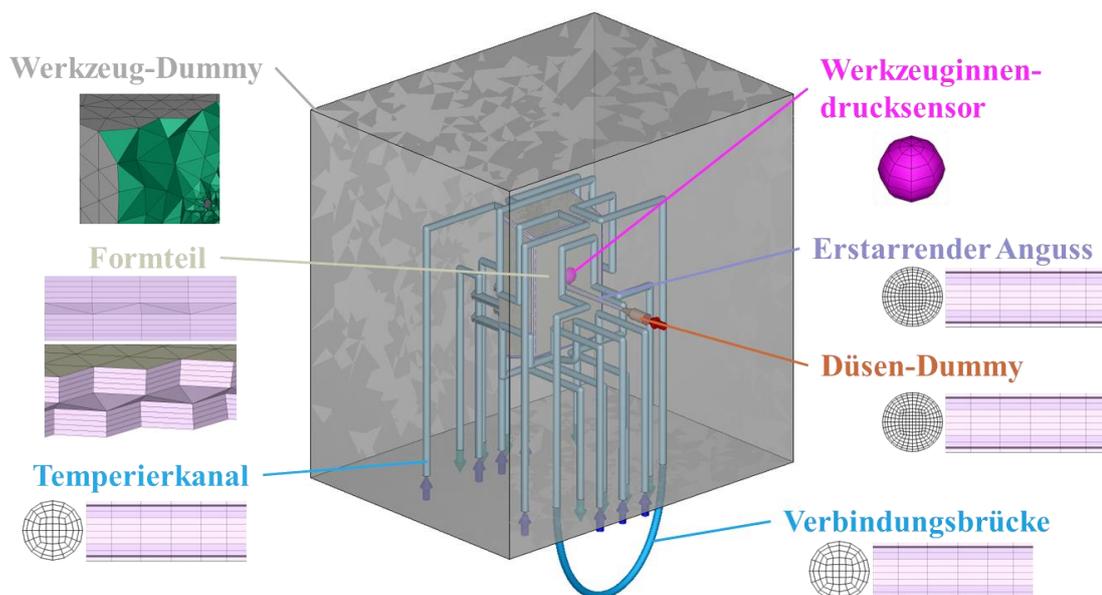


Abbildung 18: Übersicht der attributspezifischen Vernetzungsarten im Simulationsmodell

#### Thermografie:

Zum Abgleich der Auswerttemperaturen wird die Temperaturverteilung zu den Einstellparametern und den Qualitätsmerkmalen verglichen. Die innenliegende Bauteiloberfläche des Simulationsmodells ist für diesen Zweck mit rund 28.000 Messpunkten überzogen, an denen zu einem definierten Zeitpunkt die lokal vorliegende Temperatur ausgewertet werden kann.

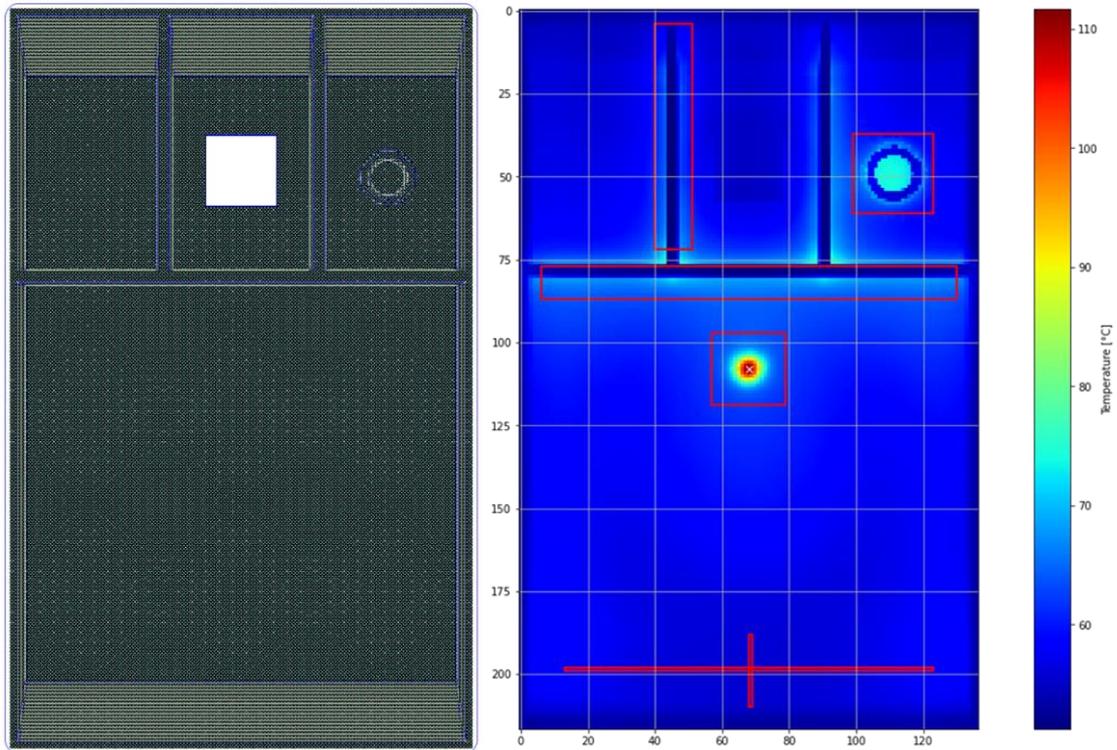
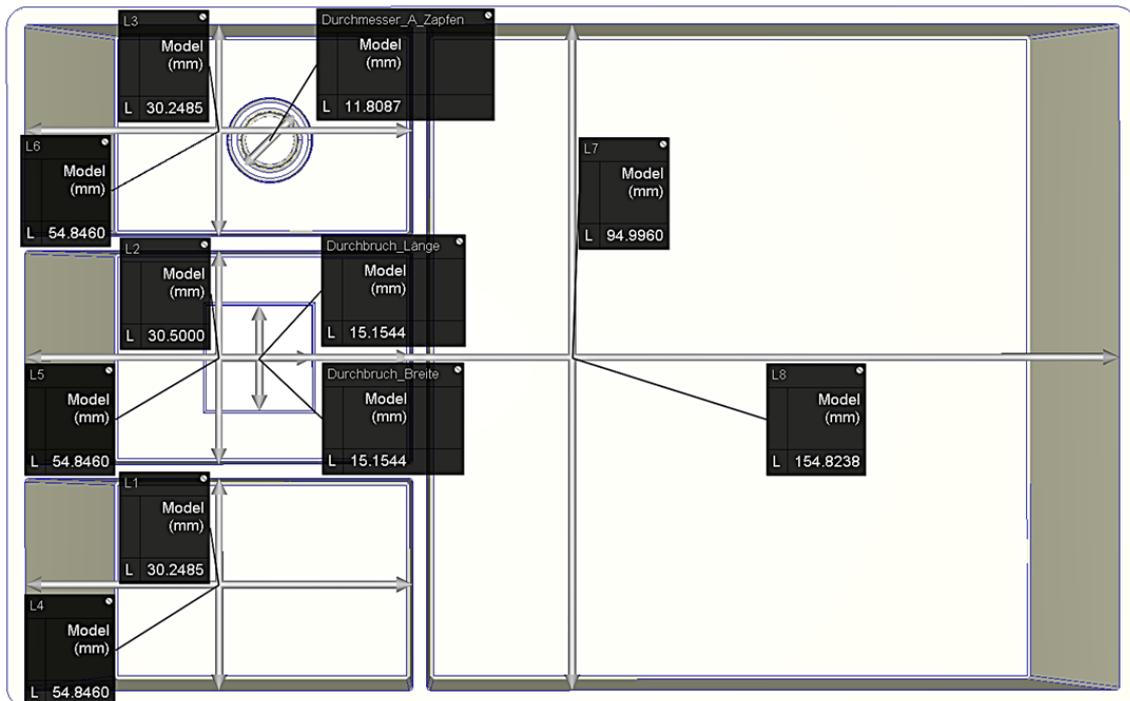


Abbildung 19: Verteilung der Messpunkte für die thermografische Auswertung (ca. 28.000 Messpunkte) links und thermografische Aufnahme einer exemplarischen Stichprobe aus dem realen Prozess (rechts)

### Bildverarbeitung:

Zur Evaluation des Verzugs sind analog zu den realen Versuchen Längenmaße auszuwerten. Das exportierte Maß entspricht den Bauteilabmaßen nach erfolgter Abkühlung auf Raumtemperatur. Die definierten Längenmaße werden mittels CSV-Datei aus Moldex3D exportiert und können anschließend weiterverarbeitet werden.



### Gewichtsmessung:

Die Auswertung des Gewichts erfolgt nach Beendigung der Nachdruckphase. Die benötigten Daten werden mittels CSV-Datei aus Moldex3D exportiert und können anschließend weiterverarbeitet werden.

Im Zeitraum des Projekts konnte festgestellt werden, dass die Softwareversion Moldex3D 2021R2 einen Versionsbug enthielt, der das Bauteilgewicht in der Nachdruckphase nicht adäquat vorhergesagt hat. Die Software wurde diesbezüglich auf die Version 2022R1 aktualisiert und die bisher durchgeführten Versuche wiederholt.

### Materialmodell:

Die Materialkarten werden durch Moldex3D bzw. den Distributor SimpaTec zur Verfügung gestellt. Für das Material Terluran ECO GP-22 MR70 sind aufgrund des Rezyklat-Anteils keine genauen Materialeigenschaften bekannt. Aufgrund dessen wird die Simulation mit dem Virgin Material Terluran GP-22 durchgeführt. Dies birgt das Risiko eines höheren Sim-to-Real-Gaps.

Die folgende Abbildung zeigt eine Bewertung der beiden verwendeten Materialkarten. Beide weisen eine nur mäßig zuverlässige Prognosewahrscheinlichkeit auf, wobei der Datensatz des verwendeten PPs zum Großteil aus vom Hersteller bereitgestellten Daten besteht, die in der Regel ausreichend valide Simulationsergebnisse liefern. Kritisch zu hinterfragen sind jedoch die generischen Kristallisationsdaten und nicht weiter spezifizierte Herkunft der Struktur-viskoelastischen Daten.

Das ABS verweist zu dem für spezifische Wärmekapazität, Wärmeleitfähigkeit und den mechanischen Eigenschaften auf Literaturdaten.

PP HE125MO BOREALIS	Data Source	ABS Terluran GP-22 Styrolution	Data Source
Viscosity	Manufacturer 	Viscosity	Manufacturer 
PVT	Manufacturer 	PVT	Manufacturer 
Crystallinity	Generic 	Specific Heat	Campus 5.0 
Specific Heat	Manufacturer 	Thermal Conductivity	Campus 5.0 
Thermal Conductivity	Manufacturer 	Viscoelasticity	Others 
Viscoelasticity	Others 	Mechanical Properties	Campus 5.0 
Mechanical Properties	Manufacturer 	Structure VE	Generic (2148) 
Structure VE	Generic 	Process condition	Manufacturer 
Optics	Others 		
Process condition	Manufacturer 		

				
<b>optimale</b> Datenqualität	<b>gute</b> Datenqualität	<b>nicht</b> quantifizierbare Datenqualität	<b>schlechte</b> Datenqualität	<b>unzureichende</b> Datenqualität

Abbildung 20: Bewertungsmatrix der Materialdatenqualität für die rheologischen Untersuchungen für PP-HE125MO (links) und ABS-TerluranGP-22 (rechts)

#### Prozessmodell:

Das Prozessmodell wurde analog zu den praktischen Versuchen modelliert.

#### **4.4.2 Evaluation des digitalen Zwillings anhand der Vorversuche**

Mit Hilfe des im Dezember 2021 bestellten DoE-Moduls für *Moldex3D* kann der gesamte  $2^4 (+1)$  Versuchsplan, wie er für die Vorversuche entwickelt und schon abgefahren wurde, in einem Stück simuliert werden. Zusätzlich sollen auch Versuchsreihen mit Zwischenschritten für die vier zu variierenden Parameter Einspritzgeschwindigkeit, Massetemperatur, Werkzeugtemperatur und Nachdruck simuliert werden. Als Ausgabe der Simulation wurde das Bauteilgewicht, sowie die Temperaturen der Bauteiloberfläche an verschiedenen Punkten, sowie der Verzug untersucht. Das entspricht den Daten, welche an der realen Spritzgießanlage mittels der verbauten Messtechnik erhoben wurden.

#### **4.4.3 Verknüpfung von Prozesssimulation und Bayes'schen Modellen**

Wie oben beschrieben zeigte der Vergleich der Simulationsdaten mit den real gemessenen Daten auch nach den Optimierungsmaßnahmen nur mäßige Übereinstimmung, weshalb für die Definition und Quantifizierung des Bayes'schen Netzes die realen Versuchsdaten herangezogen wurden.

#### **4.5 Hauptversuche**

Ziel der Hauptversuche ist die Generierung einer breiten Datenbasis zur Quantifizierung des entwickelten Bayes'schen Netzes. Um ausreichend Trainingsdaten zu generieren, wurde sich für die Durchführung eines D-optimalen Versuchsplans mit PP entschieden.

Dies bietet die Möglichkeit auch nicht lineare Zusammenhänge und 2-fach Wechselwirkungen, wie sie beim Spritzgießen auftreten, zu erkennen und abzubilden. Dies ist möglich da neben dem minimalen und maximalen Einstellwert auch eine Variation von Zwischenwerten erfolgt (vgl. Abbildung 21). Im Versuchsplanassistenten MESOS wird über die Eingabe von Levelstufen die Anzahl an Variationen eines Einstellparameters angegeben.

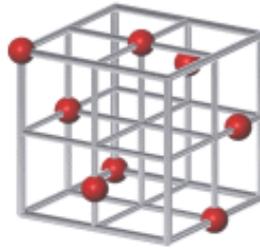


Abbildung 21: Beispielhafte Darstellung der Versuchspunkte bei einem D-optimalen Versuchsplan

Insgesamt enthält der gewählte Versuchsplan 28 Versuchsreihen. Wie in den Vorversuchen werden die Parameter Zylindertemperatur, Werkzeugtemperatur, Einspritzgeschwindigkeit und Nachdruck variiert. Der Parameterbereich und die Schrittweite zwischen den Levelstufen sind in Tabelle 7 dargestellt.

Tabelle 7: D-Optimaler Versuchsplan mit 28 Versuchsreihen

	Parameterbereich	Schrittweite	Stufen
Zylindertemperatur [°C]	220 - 260	20	3
Werkzeugtemperatur [°C]	30 - 60	15	3
Einspritzgeschwindigkeit [ccm/s]	40 - 105	22,5	4
Nachdruck [bar]	150 - 300	50	4

Um den Einfluss der variierten Einstellparameter unverfälscht zu erfassen, ist es notwendig sonstige Einflüsse auf das Qualitätsmerkmal identisch zu halten, Maschinenabläufe analog zu vorherigen Versuchsreihen einzustellen und prozesstechnische Charakteristiken zu beachten. Die getroffenen Entscheidungen sind im Folgenden erläutert.

- **Temperierung des Werkzeugs:** Die Temperierung des Werkzeugs erfolgt über mehrere Temperierkreisläufe, welche stets identisch angeschlossen sind.
- **Handlingentnahme:** Das zu Beginn programmierte Handlingprogramm wird bei jeder Versuchsreihe bzw. Versuchstag eingelesen und nicht verändert. Zu beachten ist, dass trotz der verschiedenen Inline-Qualitätsmesssysteme nicht die Zykluszeit durch das Handling beeinflusst wird. Die Fahrgeschwindigkeiten und das Zeitintervall zwischen Bauteilentnahme und Thermographieaufnahme müssen stets identisch sein.

- Maschinenablaufprogramm: Das Ablaufprogramm wurde abgespeichert und bis auf die variierenden Parameter nicht abgeändert.
- Werkzeugatmung: Beim Spritzgießen kann die Schmelze beim Einspritzvorgang und während der Nachdruckphase die Werkzeughälften minimal auseinander drücken. Dies geschieht, wenn die auftreibende Kraft des Werkzeuginnendrucks die Schließkraft übersteigt. Die kann die Ursache für verschiedene Spritzgießfehler wie Gratbildung sein und auch das Bauteilgewicht beeinflussen. Um die Werkzeugatmung gleich zu halten, wird diese zu Beginn mit einer Messuhr überprüft und über die Schließkraft optimal eingestellt.
- Einschwingverhalten des Prozesses: Das Einschwingverhalten beschreibt die Zeitspanne, die notwendig ist, bis sich Prozessgrößen stabilisiert haben. Je nachdem welche Einstellparameter verändert werden, sind unterschiedliche Zeitspannen notwendig. Insbesondere bei Veränderung der Zylinder- oder Werkzeugtemperatur ist mehr Zeit bis zur Einstellung eines thermischen Gleichgewichts notwendig, bis ein thermisches Gleichgewicht erreicht wird.

Bei jeder der 28 Versuchsreihen erfolgte die Entnahme von 12 Bauteilen, um die Prozesskonstanz innerhalb einer Versuchsreihe sicherstellen zu können.

## 4.6 Datenanalyse

Im folgenden Kapitel werden die Schritte zur Aufbereitung der Rohdaten und zur Erzeugung eines Datensatzes erläutert. Außerdem wird die Extraktion von Features aus Zeitreihen- und Bilddaten beschrieben.

### 4.6.1 Datenaufbereitung und Erzeugung eines Datensatzes

Die Messdaten standen direkt nach der Erfassung in den Versuchen noch nicht einheitlich konfiguriert und fusioniert, d.h. eindeutig und korrekt einander zugeordnet, zur Verfügung. Ziel der Vorverarbeitung war daher zunächst die Erzeugung einer weitestgehend automatisierbaren Routine zur Datenaufbereitung, die alle Messdaten eindeutig dem entsprechenden Zyklus zuordnet und zudem auf die zugehörigen Referenzdaten (Thermogramme, Bilder und Zeitreihen) verweist. Diese Arbeiten konnten direkt für die Umsetzung des Demonstrators verwertet werden. Für diese Vorverarbeitung und Zusammenführung der Daten aus unterschiedlichen Quellen wurde ein Skript in der Programmiersprache Python erarbeitet, das sich verschiedener Teilfunktionen aus bestehenden Programmbibliotheken (NumPy, pandas, u. a.) bedient, um automatisiert einen vollständigen Datensatz zu erstellen.

Die erfassten Daten können in vier verschiedene Typen eingeteilt werden:

- Zeitreihendaten je Zyklus, wie z. B. der Nachdruckverlauf, die die Änderung eines Werts über der Zeit innerhalb eines Spritzgießzyklus darstellen.
- Zyklenübergreifende Zeitreihendaten von Peripheriegeräten, wie z. B. Temperiergeräten oder dem Trocknungs- und Dosiersystem.
- Einzelwerte, wie z. B. der maximale Werkzeuginnendruck, die einen einzelnen Wert pro Zyklus liefern.

- Zyklen- bzw. bauteilbezogene Bilddaten, die durch die Infrarot-Kamera sowie das Bildverarbeitungssystem generiert werden.

Abbildung 22 zeigt schematisch das Vorgehen von den Rohdaten aus verschiedenen Quellen hin zu einem einheitlichen und vollständigen Datensatz.

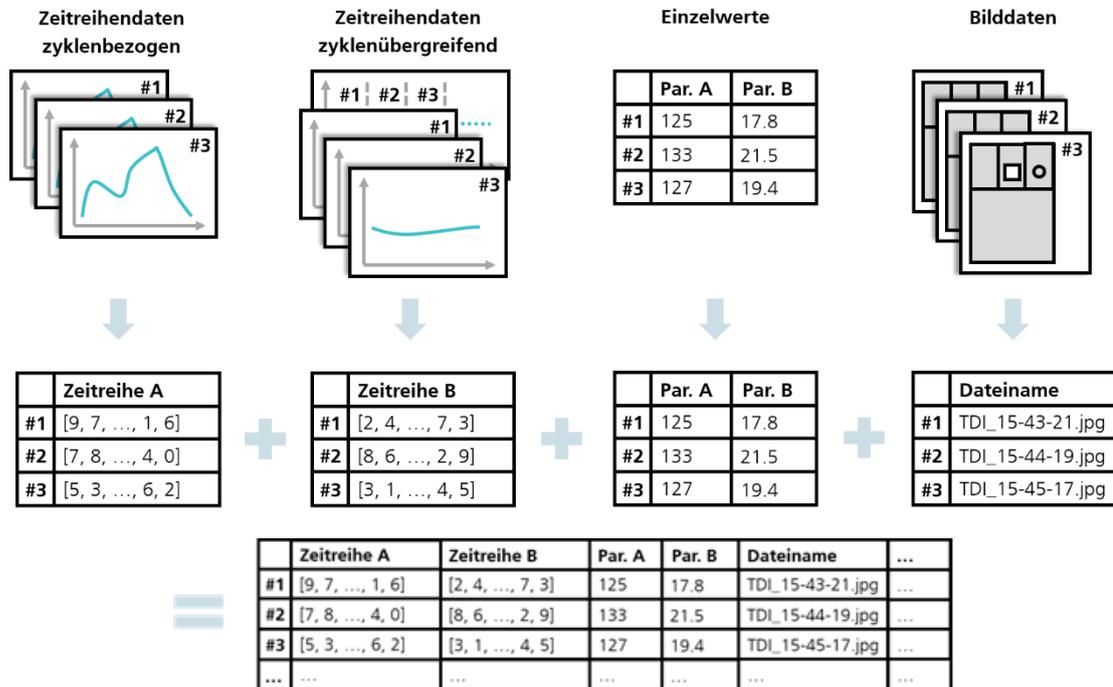


Abbildung 22: Schematische Darstellung der Aufbereitung und Zusammenführung unterschiedlicher Datentypen

Ziel der Datenaufbereitung war es einen Datensatz zu erzeugen, der die Daten aus allen Quellen enthält und in dem jede Zeile einem Spritzgießzyklus entspricht. In den Spalten sind die Werte der verschiedenen Parameter enthalten. Im ersten Schritt mussten die einzelnen Dateien eingelesen und je nach Bedarf verschiedene Vorverarbeitungsschritte durchgeführt werden, um eine weitere Analyse zu ermöglichen. Dies war unter anderem notwendig, um die unterschiedlichen Formate (*csv*, *xlsx*, *parquet*, *txt*, *jpg*) und die daraus resultierenden Dateistrukturen verarbeiten zu können. Erforderliche Vorverarbeitungsschritte umfassten beispielsweise das Entfernen irrelevanter Informationen aus den Dateien, die Anpassung von Parameter-Bezeichnungen, die Umstrukturierung von Daten sowie das Auffüllen oder Entfernen leerer Zellen. Im nächsten Schritt wurden den verschiedenen Daten die zugehörigen Spritzgießzyklen zugeordnet. Einzelne Daten (z. B. aus dem *DataXplorer* oder dem Bildverarbeitungssystem) verfügten bereits durch die Parameter *JobCycleID* oder *MachineCycleID* über die notwendigen Informationen. Bei anderen Daten (z. B. Thermogramme) musste eine Zuordnung über den Zeitstempel umgesetzt werden. Die Daten wurden nacheinander vorverarbeitet, einem Zyklus zugeordnet und zusammengeführt. So entstand schrittweise und automatisiert der vollständige Datensatz. Zyklenbezogene Zeitreihendaten wurden in Arrays überführt und konnten dadurch ebenfalls zeilenweise hinzugefügt werden. Einzelwerte konnten zum jeweiligen

Zyklus im Datensatz hinzugefügt werden. Bei Bilddaten wurde der Dateiname im Datensatz abgespeichert. Abbildung 23 zeigt einen Ausschnitt aus dem finalen Datensatz, der in den ersten Spalten Meta-Informationen wie die Zyklusnummer, die Versuchsreihe und den Zeitstempel sowie in den folgenden Spalten alle Parameter-Werte enthält.

Meta-Informationen				Parameter-Werte			
MET_JobCycleID	MET_MachineCycleID	MET_ExperimentNumber	MET_Timestamp	SCA_PartWeight	ENV_AirTemperature	ENV_AirHumidity	...
1	18535	A16	2021-12-07 08:27:34	91.39	17.1	32.0	
2	18536	A16	2021-12-07 08:28:31	91.37	17.1	32.0	
3	18537	A16	2021-12-07 08:29:28	91.42	17.1	32.0	
4	18538	A16	2021-12-07 08:30:25	91.23	17.1	32.0	
...	...	...	...	...	...	...	...

Abbildung 23: Ausschnitt aus dem finalen Datensatz

Die Bezeichnung der einzelnen Parameter erfolgt dabei nach einem systematischen Prinzip, indem jedem Parameter jeweils die zugrundeliegende Datenquelle in Form eines Präfixes vorangestellt ist. Dies trägt zur Übersichtlichkeit bei und erlaubt beispielsweise das einfache Auswählen aller zu einer bestimmten Datenquelle zugehörigen Parameter (siehe Tabelle 8).

Tabelle 8: Bezeichnung und Abkürzung der Datenquellen

Datenquelle (DE)	Datenquelle (EN)	Präfix	Beispiel-Parameter
<b>Qualitätstabelle</b>	Quality Table	QUA	QUA_CushionVolume
<b>Euromap77</b>	Euromap77	E77	E77_AverageCycleTime
<b>DataXplorer</b>	DataXplorer	DXP	DXP_Inj1PrsAct
<b>Temperiereinheit Auswerferseite</b>	Temperature Control Unit Ejector Side	TCE	TCE_TemperatureReturnLine
<b>Temperiereinheit Düsenseite</b>	Temperature Control Unit Nozzle Side	TCN	TCN_TemperatureReturnLine
<b>Trocknungssystem</b>	Drying system	DRY	DRY_HT101_acTempReturnAir
<b>Dosiersystem</b>	Dosing system	DOS	DOS_acComp1DosRate
<b>Bildverarbeitungssystem</b>	Computer Vision	CV	CV_Width1
<b>Thermo- bzw. IR-Kamera</b>	Thermal / IR camera	IR	IR_Image1Name
<b>Meta-Informationen</b>	Meta data	MET	MET_JobCycleID
<b>Waage</b>	Scale	SCA	SCA_PartWeight
<b>Umweltsensor</b>	Environmental sensor	ENV	ENV_AirTemperature
<b>Manuelle visuelle Inspektion</b>	Manual visual defect inspection	LBL	LBL_SinkMarks

Dieser aufbereitete Datensatz stellt ein wichtiges Zwischenergebnis dar. Die eindeutige und konsistente Zuordnung der Daten aus den verschiedenen Datenquellen zu den jeweiligen Spritzgießzyklen und die standardisierten Parameter-Bezeichnungen ermöglichen eine hohe Datenqualität. Erst durch diesen Datensatz werden weitergehende Analysen möglich. Insgesamt umfasst die Datenbasis zum Ende des Projekts ca. 870 Spritzgießvorgänge mit jeweils über 300 Parametern. Es wurden datenbasierte Analysen von Einflussfaktoren auf Qualitätsmerkmale durch Korrelationsanalysen durchgeführt und der Datenraum systematisch durch die Berechnung zusätzlicher Features erweitert. Der Datensatz bildete die Grundlage für die weiteren Analysen der Wirkzusammenhänge und die Merkmalsselektion („Feature Selection“) sowie die Berechnung von statistischen Verteilungen für die Quantifizierung der Bayes’schen Netze.

#### 4.6.2 Extraktion von Features aus den Thermografieaufnahmen

Ein wichtiger Teil dieser Analysen war die Extraktion weiterer Parameter bzw. aussagekräftiger Datenmerkmale aus den Rohdaten (sog. „Feature Extraction“), die am Beispiel der Thermokamera-Daten näher erläutert wird. Ziel war es hierbei, relevante Bereiche auf der Bauteiloberfläche zu identifizieren und für diese Bereiche signifikante Merkmale zu berechnen, wie zum Beispiel die mittlere Temperatur oder Temperaturverläufe. Relevant bedeutet in diesem Fall, dass die extrahierten Parameter einen Zusammenhang mit dem Qualitätsmerkmal Bauteilverzug herstellen sollen. Diese extrahierten Parameter wurden ebenfalls zum Datensatz hinzugefügt (in Abbildung 24 schematisch durch die Parameter T1, T2 und T3 dargestellt).

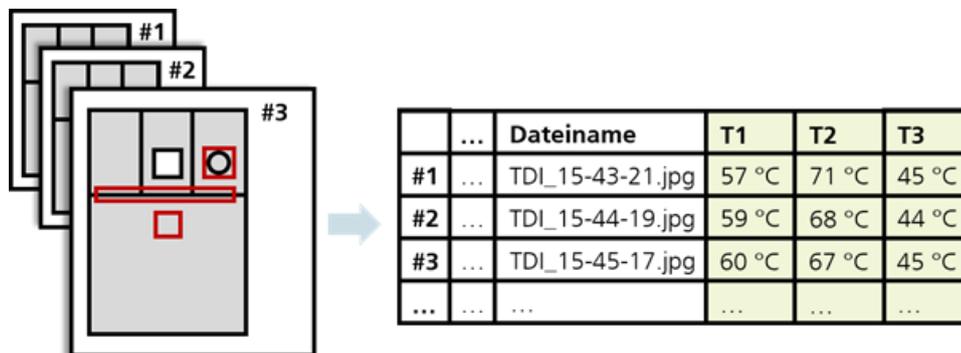


Abbildung 24: Schematische Darstellung der Erweiterung des Datensatzes um extrahierte Features aus den Thermokamera-Daten

Die Thermokamera-Aufnahmen liegen als *csv*-Dateien mit 640 Spalten und 480 Zeilen vor, was der Kamera-Auflösung entspricht. Jede Zelle enthält einen Temperaturwert. Zur Verarbeitung und Parameter-Extraktion wurde ein Python-Skript entwickelt, das die Bilddaten automatisiert einliest, die Bauteilposition bzw. -kontur im Bild erkennt und die gewünschten Temperaturwerte von definierten Bereichen auf der Bauteiloberfläche berechnet. Zur Identifizierung der interessanten Bereiche (sog. „Regions of Interest“, kurz „ROI“) wurden für jeden Pixel die Standardabweichungen der Temperaturwerte berech-

net. Dies wurde für die einzelnen Versuchsreihen, als auch über alle Versuchsreihen hinweg und jeweils nur mit der ersten Aufnahme pro Teil sowie mit allen drei Aufnahmen jedes Teils durchgeführt. Abbildung 25 zeigt das Ergebnis anhand von zwei Beispielen. Rote Bereiche zeigen eine hohe Standardabweichung und blaue Bereiche eine niedrige. Besonders beim Anguss, beim Dom sowie an den Kanten der drei Taschen sind gegenüber dem restlichen Bauteil hohe Standardabweichungen zu erkennen. Bei einigen Versuchsreihen konnten zudem deutliche Einfallstellen erkannt werden (siehe Abbildung 25, rechts). Dies zeigt sich durch die stark schwankenden Standardabweichungen an den Wänden im oberen und unteren Bereich des Teils.

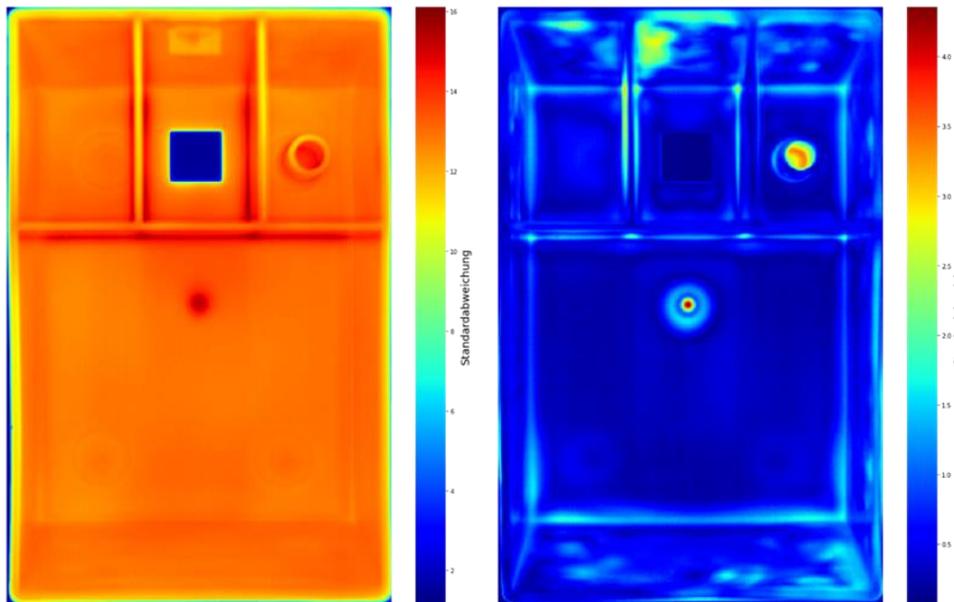


Abbildung 25: Standardabweichung der Temperaturwerte über alle Versuche und Aufnahmen mit PP hinweg (links) und für die Versuchsreihe A06 mit PP über alle Aufnahmen hinweg (rechts)

Basierend auf den Ergebnissen der vorherigen Analyse wurden die in Abbildung 26 dargestellten Bereiche („Full“, „EdgeVer“, „Dome“, „Sprue“ und „EdgeHor“) und Verläufe („GradVer“ und „GradHor“) definiert. Von den rechteckigen Bereichen wurden jeweils die Mittelwerte berechnet. Für die Verläufe wurden die Standardabweichungen sowie die Differenzen zwischen maximalen und minimalen Temperaturwerten berechnet. Außerdem wurden jeweils die Gradienten ermittelt und hierfür ebenfalls Standardabweichungen und Differenzen zwischen maximalen und minimalen Temperaturen.

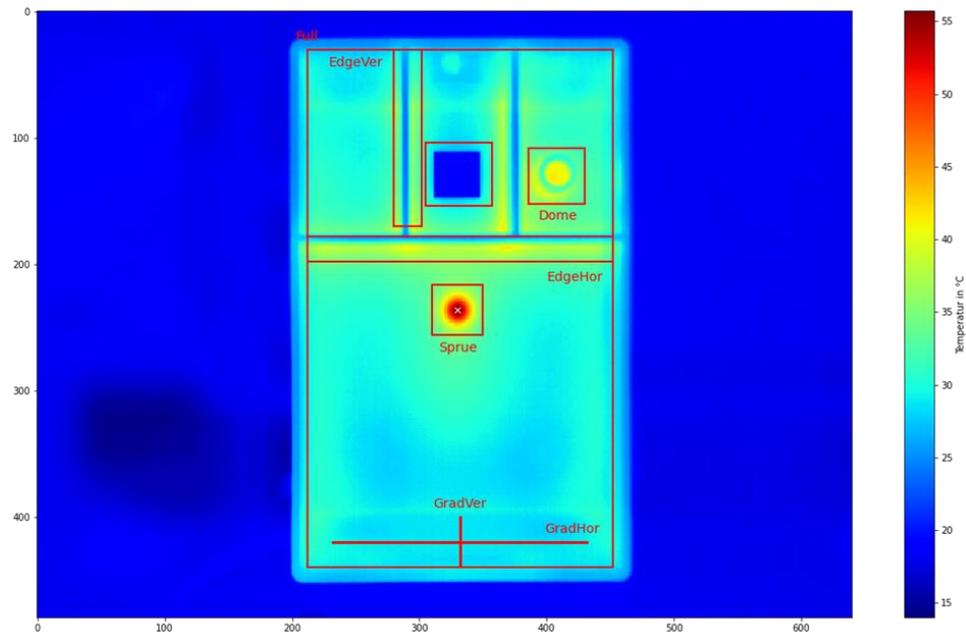


Abbildung 26: Thermogramm mit festgelegten „Regions of Interest“ (ROIs)

Um die Zusammenhänge zwischen extrahierten Features aus den Thermogrammen und bestimmten Fehlerkategorien bzw. Fehlerbildern zu bewerten, wurden Pearson-Korrelationskoeffizienten berechnet. Abbildung 27 zeigt einen ausgewählten Teil der Ergebnisse in einer Korrelationsmatrix. Eine besonders hohe Korrelation weisen die abgebildeten zum vertikalen Gradienten bzw. Verlauf gehörenden Features mit dem Fehlerbild Einfallstellen („LBL\_SinkMarks“) auf.

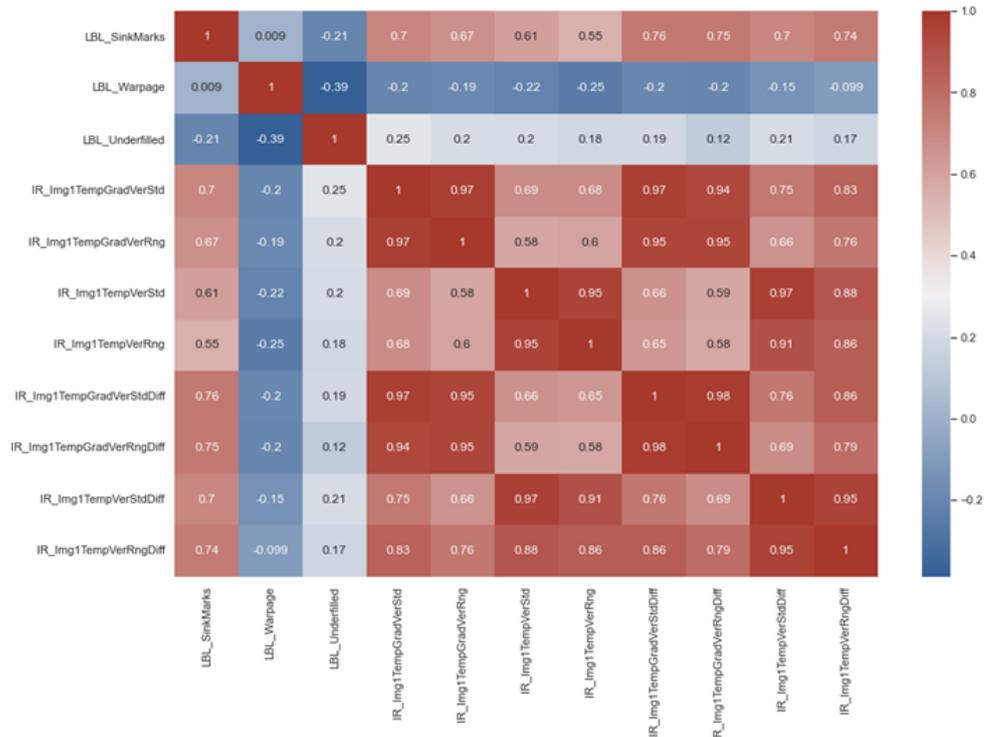


Abbildung 27: Pearson-Korrelationskoeffizienten zwischen ausgewählten Fehlerbildern und extrahierten Features aus den Thermogrammen

Die extrahierten Features stellen eigene Prozessparameter bzw. Qualitätsmerkmale der produzierten Bauteile dar. Da diese jedoch für den Maschinenbediener schwer greifbar und sehr komplex in einer konsistente Qualitätsbewertung sind, sind diese Features zwar im Datensatz enthalten, werden jedoch nicht für die Quantifizierung der Bayes'schen Netze herangezogen.

#### 4.6.3 Extraktion von Features aus Zeitreihen

Auch aus den Zeitreihendaten (z. B. Druckverläufe aus dem *DataXplorer*) wurden Features extrahiert. Hierzu wurde unter anderem das Python-Package *tsfresh* [19] verwendet, mit dem automatisch eine große Anzahl von Zeitreihenmerkmalen bestimmt werden kann. Dies wird beispielhaft anhand des Einspritzdruckes in Abbildung 28 dargestellt. Hier wurden der maximale und minimale Einspritzdruck sowie der Mittelwert bestimmt und im Diagramm visualisiert. Insgesamt können mit *tsfresh* über 700 einzelne statistische Merkmale aus einer Zeitreihe extrahiert werden. Zu beachten ist jedoch, dass viele dieser Größen, im Gegensatz zu Merkmalen wie Minimal- oder Maximal-Werten, für Menschen nur schwer interpretierbar und nachvollziehbar sind. Auf diese schlecht interpretierbaren Features wurde daher verzichtet, da sie dem erklärbaren und nachvollziehbaren Ansatz der im Projekt zu entwickelnden Methodik widersprechen.

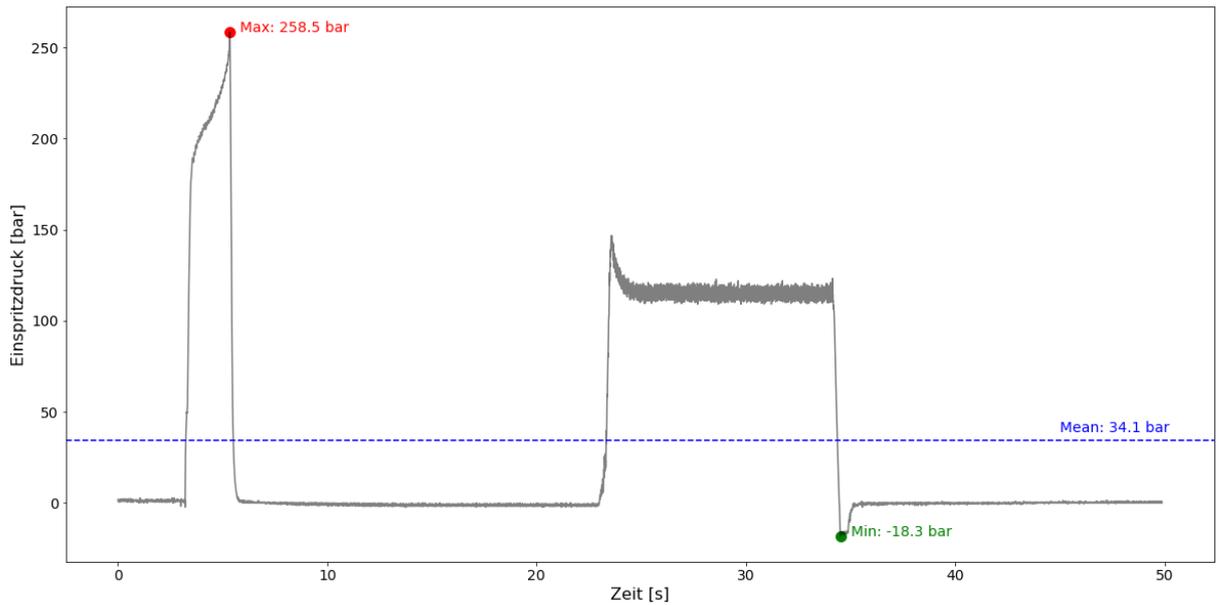


Abbildung 28: Exemplarische Extraktion von Features aus der Einspritzdruckkurve

Der Spritzgießprozess lässt sich in verschiedene Phasen unterteilen. Diese sind in Abbildung 29 anhand der Triggerwerte aus dem *DataXplorer* dargestellt. Die im Projekt erzeugten Zeitreihendaten entstanden, wie in Kapitel 4.6.1 erläutert, entweder kontinuierlich über mehrere Zyklen hinweg (beispielsweise bei den Temperiergeräten) oder erstreckten sich, wie z. B. beim *DataXplorer*, jeweils nur über einen einzelnen Zyklus. Um auch phasenspezifisch Features extrahieren zu können, wurde ein Skript geschrieben, das basierend auf den Triggerwerten aus dem *DataXplorer* Merkmale für einzelne Phasen des Spritzgießzyklus extrahieren kann. Dies ist beispielhaft in Abbildung 30 dargestellt, in der speziell für die Einspritzphase Minimal- und Maximal-Wert sowie Mittelwert des Einspritzdruckes berechnet wurde.

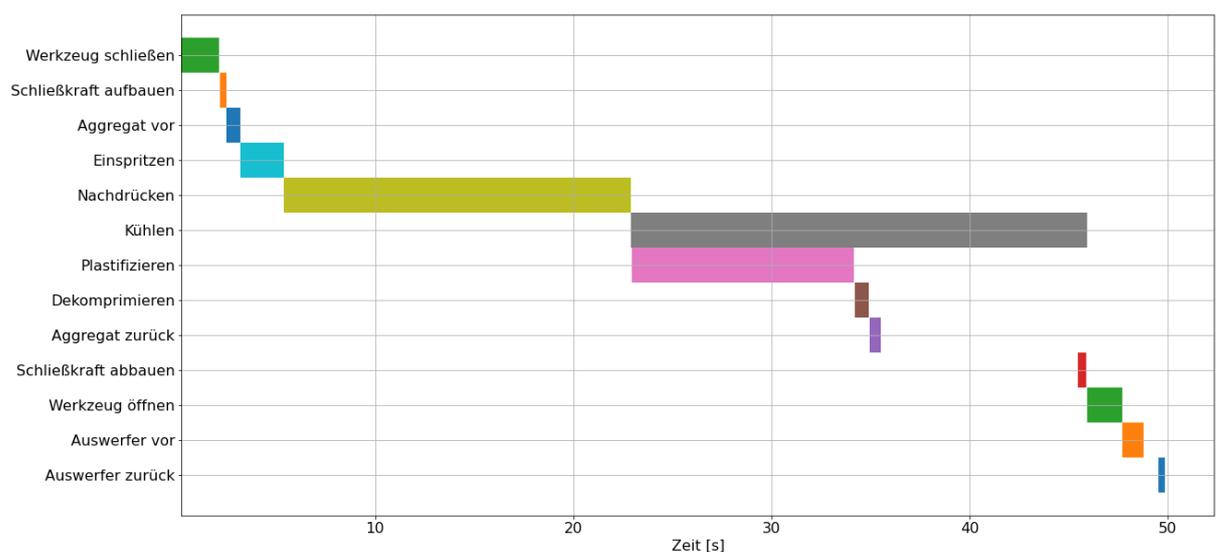


Abbildung 29: Phasen des Spritzgusszyklus basierend auf den Triggerwerten des DataXplorers

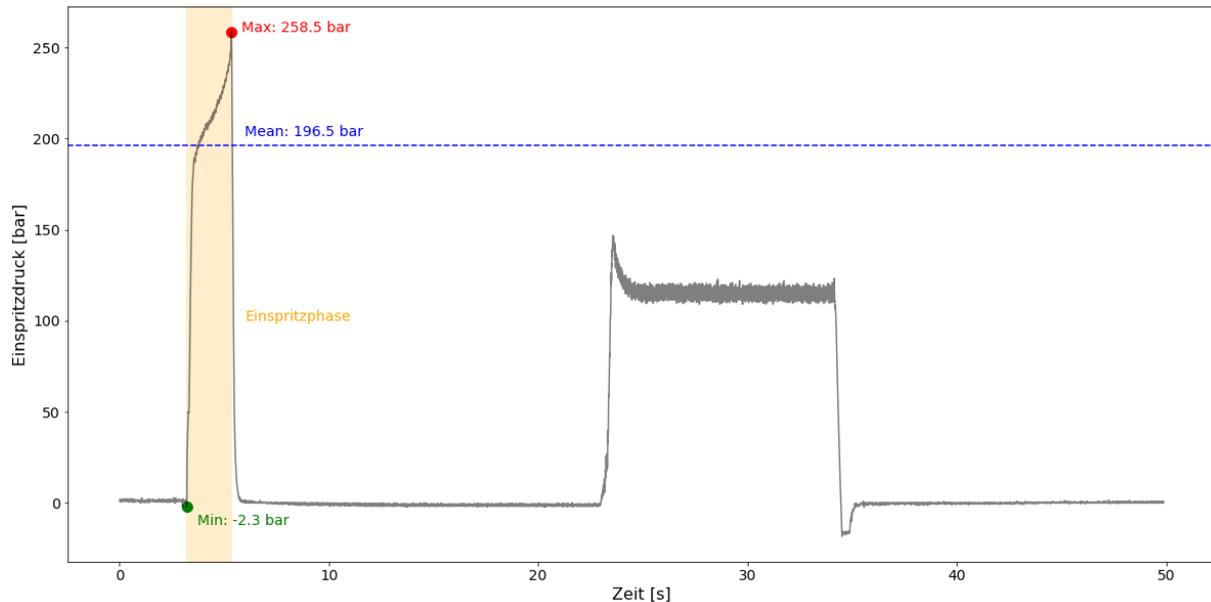


Abbildung 30: Phasenspezifische Extraktion von Features aus der Einspritzdruckkurve

#### 4.6.4 Features aus den Bauteilmaßen des Bildverarbeitungssystems

Um den Verzug anhand einer Kennzahl bewerten zu können, wurde eine Formel entwickelt, die sich aus der Summe der Beträge der Abweichungen zwischen Soll- und Ist-Maßen der drei Taschen der Verzugsschale zusammensetzt (vgl. Abbildung 31 und Formel 1). Sollten Soll- und Ist-Maße für alle drei Kammern übereinstimmen, ist die Verzugskennzahl 0. Mit steigender Abweichung, also stärkerem Verzug, steigt auch der Wert der Kennzahl an. Es handelt sich dabei um eine arbiträre, für dieses Spritzgießwerkzeug passende Kennzahl und nicht um eine normierte Kennzahl, deren Maximalwert 1 beträgt. Die Kennzahl wurde für alle Bauteile berechnet und dem Datensatz unter der Bezeichnung CV\_Warpage hinzugefügt.

$$\text{Verzug} = |l_{1s} - l_{1i}| + |l_{2s} - l_{2i}| + |l_{3s} - l_{3i}| \quad (1)$$

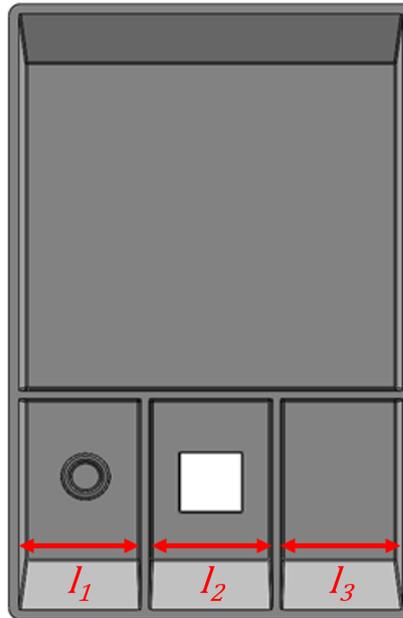


Abbildung 31: Verzugschale mit den zur Berechnung der Verzugskennzahl verwendeten Maßen

## 4.7 Feature Selection

Die Merkmalsauswahl („Feature Selection“) ist ein Prozess, bei dem die relevantesten Merkmale oder Variablen in einem Datensatz ausgewählt werden. Das Ziel ist es, eine Untermenge der ursprünglichen Merkmale zu finden, die die Vorhersagegenauigkeit eines Modells verbessert und gleichzeitig die Komplexität reduziert. Durch die Auswahl der besten Merkmale können unnötige Redundanzen und Rauschen vermieden werden, was zu effizienteren und interpretierbaren Modellen führt. Die Merkmalsauswahl fand im Projekt sowohl daten- als auch wissensbasiert statt, wobei die Ergebnisse aus beiden Herangehensweisen zusammengeführt wurden.

### 4.7.1 Datenbasierte Feature Selection

Für die datenbasierte Merkmalsauswahl wurde ein Skript programmiert („feature\_selection.ipynb“) [18], das für eine ausgewählte Qualitätsgröße automatisiert die fünf wichtigsten Prozessparameter ermittelt und für diese anschließend jeweils die wichtigsten fünf Einstellparameter. Hierfür wurden zunächst alle relevanten Merkmale einer der drei Kategorien Einstellparameter, Prozessparameter und Qualitätsgrößen zugeordnet. Im Skript wird die Qualitätsgröße (z. B. das Bauteilgewicht) definiert, für die die Feature Selection durchgeführt werden soll. Die Daten werden in Trainings- und Testdatensatz aufgeteilt. Zunächst wird ein Random Forest basierend auf allen verfügbaren Prozessgrößen mit dem angegebenen Qualitätsmerkmal als Zielgröße trainiert und der  $R^2$ -Wert ausgegeben, um eine Ausgangs- und Vergleichsbasis zu haben.

Anschließend wird die sogenannte „Permutation Feature Importance“ berechnet. Dabei handelt es sich um ein Modellprüfungsverfahren, das mit verschiedenen Modelltypen verwendet werden kann, solange die zugrundeliegenden Daten tabellarisch sind. Besonders nützlich ist das Verfahren für nichtlineare oder intransparente Modelle. Die „Permutation

Feature Importance“ ist definiert als die Verringerung eines Modell-Scores, wenn ein einzelner Merkmalswert zufällig ausgetauscht („permutiert“) wird. Durch dieses Verfahren wird die Beziehung zwischen dem Merkmal und der Zielgröße unterbrochen, so dass der Rückgang der Modell-Performance einen Hinweis darauf gibt, wie sehr das Modell von dem Merkmal abhängt. Diese Technik hat den Vorteil, dass sie modellunabhängig ist und viele Male mit verschiedenen Permutationen des Merkmals berechnet werden kann.

Das Skript gibt anschließend die ermittelten fünf besten Prozessparameter für die Vorhersage des Qualitätsmerkmals in Listenform aus. Zusätzlich werden die berechneten „Permutation Feature Importances“ für die einzelnen Parameter sowie Streudiagramme für alle Kombinationen aus Zielgröße und den ausgewählten Prozessparametern geplottet, um eine Beurteilung durch die Prozessexperten zu ermöglichen. Außerdem wird ein R2-Score angegeben für einen Random Forest, der nur auf den ausgewählten Parametern basiert. Für jeden der fünf Prozessparameter werden nun nach derselben, zuvor beschriebenen, Vorgehensweise die besten fünf Einstellparameter ermittelt und auch wieder Plots generiert.

Zu beachten ist, dass die Gestaltung des Versuchsplans, mit dem die Datenbasis generiert wurde, einen starken Einfluss auf die datenbasierte Feature Selection haben kann. Ein Versuchsplan, der eine breite Palette von Einstellparametervariationen abdeckt, kann dazu beitragen, dass eine größere Variabilität der Daten erzeugt wird. Dadurch können verschiedene Muster und Zusammenhänge zwischen den Merkmalen und der Zielvariable identifiziert werden, die sonst möglicherweise übersehen worden wären. Eine erhöhte Variabilität der Daten kann die Chancen erhöhen, relevante Features zu entdecken. Ein gut gestalteter Versuchsplan kann auch dazu beitragen, Interaktionen zwischen den Merkmalen zu identifizieren. Interaktionen treten auf, wenn der Effekt einer Variablen von einer anderen Variablen abhängt. Durch systematische Variation der Einstellungen in verschiedenen Kombinationen können solche Interaktionen erkannt und als relevante Features in das Modell aufgenommen werden. Ein Versuchsplan, der systematisch verschiedene Einstellungen abdeckt, kann auch dabei helfen, irrelevante Features zu identifizieren. Wenn bestimmte Einstellungen keine Auswirkungen auf die Zielvariable haben, können diese Merkmale als irrelevant betrachtet und bei der Feature Selection ausgeschlossen werden. Dadurch kann die Modellkomplexität reduziert und Overfitting vermieden werden. Es ist jedoch zu beachten, dass die konkreten Auswirkungen des Versuchsplans auf die Feature Selection stark von der spezifischen Problemstellung, den verfügbaren Daten und dem Machine Learning Algorithmus abhängen. Es gibt keine allgemeingültige Regel, welche Auswirkungen ein Versuchsplan haben wird, aber eine sorgfältige Gestaltung kann die Chancen erhöhen, relevante Features zu identifizieren.

#### **4.7.2 Wissensbasierte Feature Selection**

Die Ergebnisse aus der datenbasierten Feature Selection wurden nicht direkt übernommen, sondern zunächst gemeinsam mit den Prozessexperten interpretiert und diskutiert. Unter Zuhilfenahme der in Kapitel 4.2 vorgestellten Einflussmatrix, wurden die aufgedeckten Beziehungen auf Plausibilität hin überprüft. Als Basis für Diskussion wurden unter anderem die in der datenbasierten Feature Selection erzeugten Plots verwendet. Abbildung 32 zeigt beispielhaft die ermittelten „Permutation Feature Importances“ für alle

Prozessparameter in Bezug auf das Bauteilgewicht. Hier wird auf den ersten Blick deutlich, dass der maximale Werkzeuginnendruck („E77\_CavityPressureMaximum“) der mit Abstand wichtigste Parameter ist. Die beiden darauffolgenden Parameter weisen noch einen signifikanten, wenn auch geringen, Permutation Feature Importance Score auf, während die übrigen nicht mehr von Relevanz sind. In Abbildung 33 sind Streudiagramme für das Bauteilgewicht und die fünf besten Prozessparameter dargestellt. Deutlich erkennbar ist, dass besonders bei den Parametern „E77\_CavityPressureMaximum“, „E77\_CushionVolume“ und „DXP\_AreaCavityPressure“ starke lineare Zusammenhänge zum Bauteilgewicht bestehen. Basierend auf diesen verschiedenen Informationen und dem Wissen der Experten wurde schließlich die finale Auswahl der relevanten Parameter getroffen.

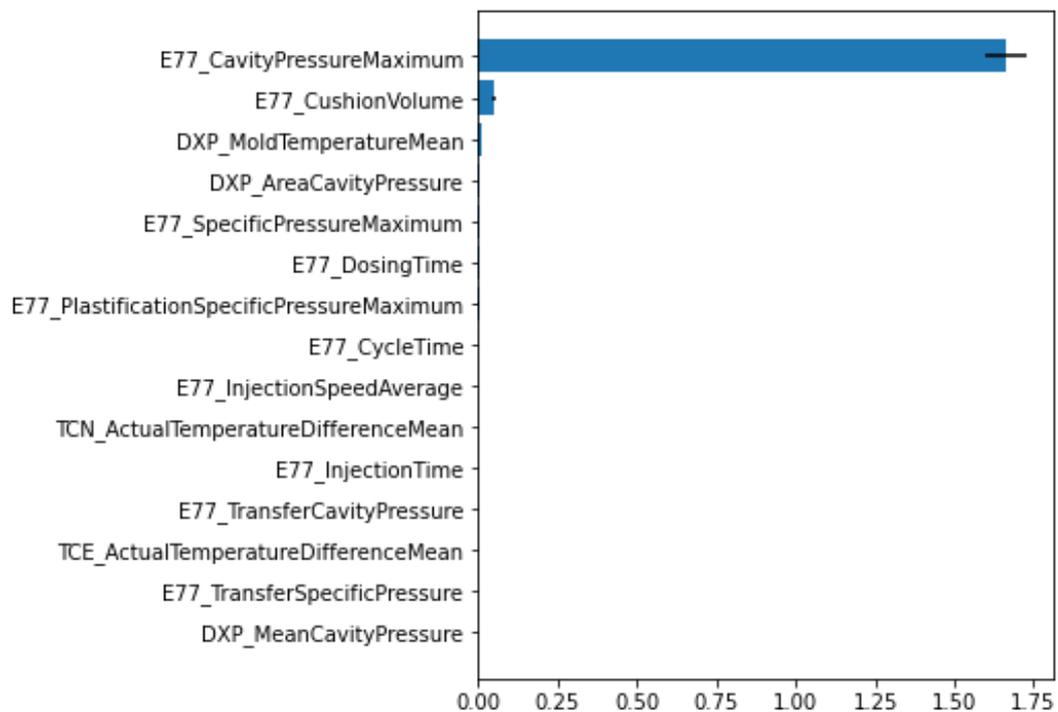


Abbildung 32: Ermittelte „Permutation Feature Importances“ für alle Prozessparameter in Bezug auf das Bauteilgewicht

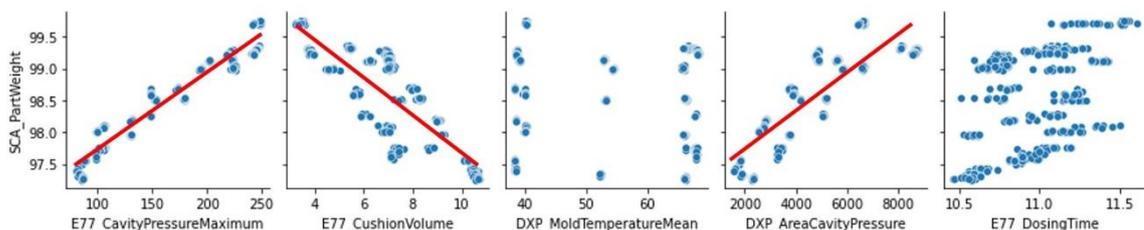


Abbildung 33: Streudiagramme für das Bauteilgewicht und die fünf besten Prozessparameter

## 4.8 Kausalitätsanalyse

In Bayes'schen Netzen wird grundsätzlich ein kausaler Zusammenhang zwischen den einzelnen Knoten angenommen. Es gibt verschiedene Definitionen von Kausalität wie z. B. die Granger-Kausalität [20], die auch verschieden getestet werden. Die im Projekt eingesetzt Python Library „DoWhy“ [21] basiert auf den Vorarbeiten von Pearl [22] und definiert einen Zusammenhang von Treatment T auf Outcome Y als kausal, wenn eine Änderung von T eine Änderung von Y bewirkt, während alles andere konstant gehalten wird.

### 4.8.1 Allgemeines Vorgehen des Kausalitätstests nach DoWhy

Der Test auf Kausalität besteht aus folgenden vier Schritten:

1. Model - Erstellen des kausalen Graphs
2. Identify - Identifizierung der Zusammenhänge
3. Estimate - Abschätzen des kausalen Effekts
4. Refute - Test und Ablehnung der Gegenhypothese

In Schritt 1 werden die Knoten und Kanten des kausalen Graphs definiert, was im Grunde der topologischen Struktur des Bayes'schen Netzes entspricht. Im zweiten Schritt wird festgelegt, welche der Größen der Outcome (Y), welche die Treatment-Größen (T) und welche die Confounder (W) sind. Wie in Abbildung 34 schematisch dargestellt wird im Folgenden der kausale Effekt der Treatment-Größen auf den Outcome untersucht.

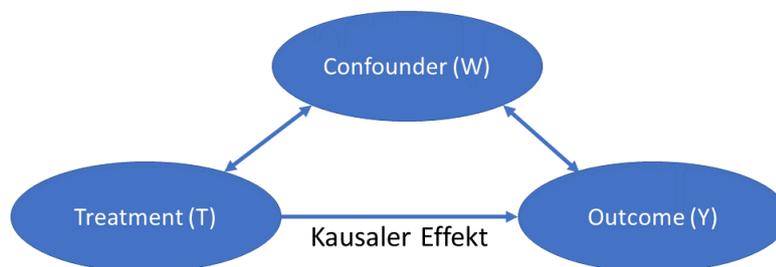


Abbildung 34: Schematische Darstellung des kausalen Effekts (angelehnt an [21])

Dies erfolgt in dem die Werte der Treatment-Größen um eine arbiträre Einheit erhöht oder verringert werden. Ändert sich der Outcome mit einem Signifikanzwert (p-value) von  $< 0,05$  wird die Kausalhypothese angenommen.

Im letzten Schritt „Refute“ erfolgt der Nachweis, dass der untersuchte Zusammenhang nicht nicht-kausal ist, ergo der Gegenbeweis zur Kausalhypothese. Dies wurde auf drei verschiedene Weisen durchgeführt:

- Placebo: Ersetzen einer Treatment-Größe durch eine Zufallsgröße. Der erwartete Estimate geht gegen 0
- Random Common Cause: Hinzufügen einer zufälligen Confounder Größe. Der erwartete Estimate bleibt gleich wie zuvor
- Data Subset: Es wird nur ein Anteil des gesamten Datensatzes verwendet für die Analyse. Der erwartete Estimate bleibt gleich.

### 4.8.2 Umsetzung im Projekt

Im mit diesem Bericht veröffentlichten Jupyter-Notebook „CausalityTest\_BayesianNetwork\_Partweight.ipynb“ [18] wurde der Test auf Kausalität für die topologische Struktur des entwickelten Bayes'schen Netzes „Bauteilgewicht“ (vgl. Abbildung 12) durchgeführt. Dazu wurden fünf kausale Graphen erstellt und entsprechend Tabelle 9 die Einstell- und Prozessgrößen den Confounders, Treatments und dem Outcome zugeordnet:

Tabelle 9: Kausale Graphen abgeleitet aus dem entwickelten Bayes'schen Netz "Bauteilgewicht"

Nr.	Name des kausalen Graphs im Quellcode	Confounders	Treatment	Outcome
1	Gesamtes Modell	Alle Einstellgrößen	Alle Prozessgrößen	Bauteilgewicht
2	Cushion Modell	-	Alle Einstellgrößen	Restmassepolster
3	Dosing Modell	-	Nachdruck, Zylindertemperatur	Plastifizierzeit
4	Area_Cavity Modell	-	Nachdruck, Zylindertemperatur	Werkzeuginnen-druck-Integral
5	Cavity_Max Modell	-	Nachdruck, Zylindertemperatur	Max. Werkzeuginnendruck

Somit sind alle Abhängigkeiten, namentlich die des Outcomes von den Prozessgrößen und die der Prozessgrößen von den Einstellgrößen, Gegenstand der Prüfung auf Kausalität. Die oben erwähnten Schritte 3 und 4 wurden für jeden dieser kausalen Graphen durchgeführt. Die berechneten Causal Estimates mitsamt der Signifikanzwerte sind im Anhang in Tabelle 13 aufgelistet.

Für alle kausalen Graphen ergibt die Abschätzung des Causal Estimate (vgl. Schritt 3) einen sehr geringen Signifikanzwert (p-value), sodass die Kausalhypothese mit großer Sicherheit angenommen werden kann. Bei Schritt 4, dem Ablehnen der Gegenhypothese, ist zu beobachten, dass der Placebo Test, also das Hinzufügen einer zufälligen Treatmentgröße den Estimate (in Tabelle 13 bezeichnet als „new causal estimate“) gegen 0 gehen lässt. Eine zufällige Treatmentgröße hat also keinen messbaren Effekt auf den Outcome, d. h. es besteht kein zufälliger, sondern ein kausaler Zusammenhang. Der zweite Gegenbeweis „Random Common Cause“ wird ebenfalls erfolgreich durchgeführt, denn der neue Estimate ändert sich auch nicht auf die vierte Nachkommastelle. Auch beim dritten Gegenbeweis „Data Subset“ sind der alte (d.h. der gesamte Datensatz wurde verwendet) und der neue (d.h. nur ein Teil des Datensatz wurde verwendet) Estimate identisch.

In Summe ist somit der Nachweis erbracht, dass alle Abhängigkeiten von Einstell- auf Prozessgrößen und von Prozessgrößen auf das Qualitätsmerkmal im Bayes'schen Netz „Bauteilgewicht“ kausaler Natur sind gemäß der oben genannten Definition für Kausalität.

## 4.9 Quantifizierung und Validierung des Bayes'schen Netzes

Die topologische Struktur des Bayes'schen Netzes „Bauteilgewicht“, wie sie in 4.2 beschrieben wurde, bildete die Grundlage für die Quantifizierung des Netzes. Für die Implementierung der Struktur und die Durchführung der Quantifizierung standen zwei Softwarelösungen zur Verfügung: *Hugin-Expert* der Firma HUGIN EXPERT A/S und *Halerium* der Firma Erium GmbH. Beide Softwarelösungen wurden getestet und gegeneinander evaluiert. Letzten Endes wurde die Implementierung der im Projekt entwickelten Netze mit *Halerium* durchgeführt, aus folgenden drei Gründen:

- *Halerium* ist rein Python-basiert und passt somit zum sonstigen Software-Stack
- *Halerium* bietet die Möglichkeit kontinuierliche Verteilungen abzubilden, während bei *Hugin-Expert* eine Diskretisierung nötig ist.
- Support durch die Firma Erium durch Bereitstellung der Software und engem Austausch bzgl. fachlicher und implementierungstechnischer Fragestellungen

Die konkrete Umsetzung in *Halerium* ist dem mitveröffentlichten Quellcode zu entnehmen und erfolgte mit dem Datensatz der Hauptversuche, der mit einem d-optimalen Versuchsplandesign erstellt wurde.

### 4.9.1 Validierung der Gewichtsvorhersage

Die Bewertung der Güte des entwickelten Bayes'schen Modells für das Bauteilgewicht erfolgte anhand verschiedener Metriken wie dem Bestimmtheitsmaß (R<sup>2</sup>), dem mittleren absoluten Fehler (MAE) und anhand von Konfusionsmatrizen. Das optimierte Teilnetz für das Bauteilgewicht erzielte im Endzustand sehr gute Ergebnisse. Die erzielten R<sup>2</sup>- und MAE-Werte sind in Tabelle 10 dargestellt. Sowohl der Trainings- als auch der Testdatensatz enthielten Experimente aus den Versuchsreihen A01 bis A28 der Hauptversuche. Ein R<sup>2</sup>-Wert von 99,4 % für den Testdatensatz deutet auf eine sehr hohe Vorhersagegenauigkeit hin, ebenso wie der geringe mittlere absolute Fehler von 0,044 g, der deutlich unterhalb der Gewichtstoleranz von ca.  $\pm 0,9$  g (entspricht  $\pm 1$  % des Bauteilgewichts) liegt.

Tabelle 10: Erzielte R<sup>2</sup>- und MAE-Werte für die Gewichtsvorhersage des Bayes'schen Netzes

	R <sup>2</sup>	MAE
<b>Train Set</b>	99,5 %	0,043 g
<b>Test Set</b>	99,4 %	0,044 g

Um das Netz mit neuen und unbekanntem Daten zu validieren, wurden weitere Versuchsreihen mit neuen Einstellparameterkombinationen durchgeführt. Das Netz wurde dann verwendet, um die Gewichtswerte für diese unbekanntem Daten vorherzusagen und wieder der MAE berechnet. In Tabelle 11 sind die Einstellparameter aus drei Validierungsversuchsreihen aufgeführt und die dafür vom Bayes'schen Netz erzielten MAE. Da es sich hier um komplett unbekanntem Daten handelt, sind die MAE hier deutlich höher als in Tabelle 10. Bei zwei der drei Versuchsreihen lag der MAE jedoch trotzdem noch innerhalb der Gewichtstoleranz. Zu beachten ist auch, dass zwischen den Versuchen zur Erzeugung der Trainingsdaten und den Validierungsversuchen mehrere Monate lagen, in

denen die verwendete Spritzgießmaschine in eine andere Halle umgezogen wurde. Um mögliche Auswirkungen zu ermitteln, wurde zunächst am neuen Standort eine bereits bekannte Einstellparameterkombination abgefahren. Die dabei ermittelten Gewichte zeigten im Mittel eine Abweichung von 0,3 g zu den identischen Versuchen mehrere Monate zuvor am alten Standort der Maschine.

Tabelle 11: Einstellparameter und erzielte MAE des Bayes'schen Netzes für drei Validierungsversuchsreihen

Versuchsreihe	VAL-05	VAL-07	VAL-09
<b>Zylindertemperatur [°C]</b>	220	220	235
<b>Werkzeugtemperatur [°C]</b>	45	55	45
<b>Einspritzgeschwindigkeit [bar*s]</b>	90	90	72,5
<b>Nachdruck [bar]</b>	200	200	225
<b>MAE [g]</b>	0,74	0,56	1,06

Abbildung 35 zeigt beispielhaft die Vorhersage des Bayes'schen Netzes in Form eines Wahrscheinlichkeitsdichteplots. Dieser kann aus der Vorhersage des Bayes'schen Netzes, also einem Mittelwert inklusive Standardabweichung, erzeugt werden. Durch diese Informationen ist es möglich, eine Aussage darüber zu treffen, mit welcher Wahrscheinlichkeit das Bauteilgewicht im i.O.-Bereich liegen wird. Im gezeigten Beispiel beträgt das Soll-Gewicht 92,0 g und die Toleranz  $\pm 1\%$ . Deshalb erstreckt sich der i.O.-Bereich von 91,1 g bis 92,9 g (in der Abbildung grün hinterlegt). Liefert das Bayes'sche Netz eine Gewichtsvorhersage von  $92,7 \pm 0,05$  g kann über die Wahrscheinlichkeitsdichtefunktion errechnet werden, dass das vorhergesagte Gewicht mit einer Wahrscheinlichkeit von 97,7 % im grünen i.O.-Bereich liegen wird.

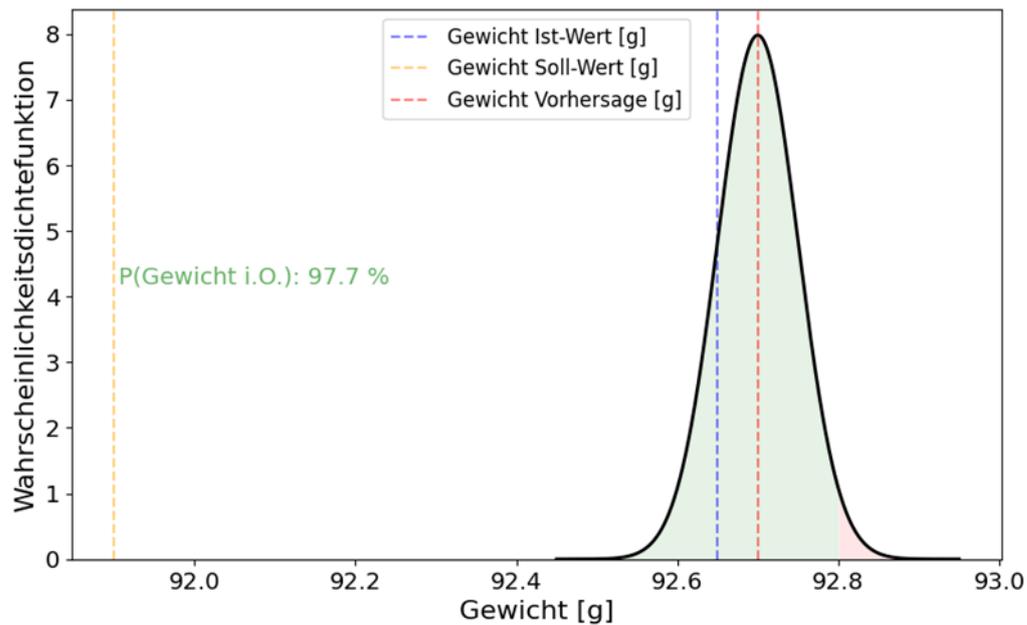


Abbildung 35: Vorhersage des Formteilgewichts durch das Bayes'sche Netz in Form eines Wahrscheinlichkeitsdichteplots

Eine weitere Modellevaluierung wurde durch eine Einteilung der gefertigten Bauteile in i.O. und n.i.O. anhand des Qualitätskriteriums Bauteilgewicht umgesetzt. Hierfür wurden sowohl alle Gewichtswerte aus dem Testdatensatz als auch die Gewichtsvorhersagen des Bayes'schen Netzes einer der beiden Kategorien i.O. oder n.i.O. zugeordnet. Durch diese Klassifikation konnten die vorhergesagten und tatsächlichen Kategorien miteinander verglichen werden. In Abbildung 36 sind die Ergebnisse einer solchen Bewertung in einer Konfusionsmatrix dargestellt. Der Testdatensatz umfasste in diesem Beispiel 65 Bauteile. Bei 51 Bauteilen wurde korrekt prognostiziert, dass das Gewicht außerhalb des i.O.-Bereichs liegen wird. Für 8 Bauteile wurde korrekterweise ein Gewicht im i.O.-Bereich vorhergesagt. Lediglich bei 6 Bauteilen, wurde fälschlicherweise ein Gewicht im i.O.-Bereich vorhergesagt, obwohl es tatsächlich außerhalb der Toleranz lag.

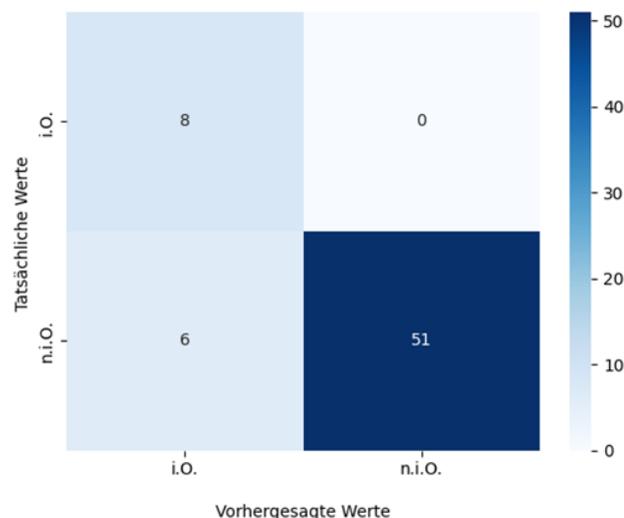


Abbildung 36: Konfusionsmatrix für die Gewichtsvorhersage des Bayes'schen Netzes

#### 4.9.2 Validierung der Parameterempfehlung

Neben der Güte der Vorhersagequalität wurde auch die Güte der Parameterempfehlungen validiert. Hierfür wurden ebenfalls verschiedene Validierungsversuchsreihen (VAL-01 bis VAL-11) gefahren (vgl. Abbildung 37). Bei den türkis hinterlegten Reihen wurden komplett neue Einstellparameterkombinationen verwendet. In der Spalte „Ergebnis“ befindet sich das gemessene mittlere Ist-Gewicht aus diesen Versuchsreihen. Durch die Pfeile wird angezeigt, ob das Ist-Gewicht zu hoch oder zu niedrig ist. In der Spalte „Ziel“ findet sich das Soll-Gewicht mit der erlaubten Toleranz. Dieses wurde in allen Fällen so gewählt, dass das Ist-Gewicht nicht innerhalb der Toleranz liegt, um eine Empfehlung auszulösen. In der Spalte „Empfehlung“ ist der vom Bayes'schen Netz ausgegebene Vorschlag für den einzustellenden Nachdruck angegeben. Die Pfeile symbolisieren hier, ob der Nachdruck erhöht oder gesenkt werden soll. Es wurde zunächst nur isoliert die Empfehlung für einen einzelnen Parameter, in diesem Fall der Nachdruck, betrachtet. Orange hinterlegte Parameter befinden sich außerhalb der Grenzen, die für den Versuchsplan zur Erzeugung der Trainingsdaten genutzt wurden. In der Ergebnis-Spalte ist das Gewicht aufgeführt, das nach Übernehmen der Nachdruckempfehlung erreicht wurde. Außerdem ist das Delta zwischen erreichtem und Ziel-Gewicht angegeben.

MAE = 0.13 g

Neue Einstellkombinationen
  Einstellungen außerhalb des Versuchsplans

Reihe	Eingestellte Parameter					Ergebnis	Ziel	Empfehlung	Ergebnis	
	Nr.	Zylinder-temp. [°C]	Werkzeug-temp. [°C]	Einspritzgeschw. [bar*s]	Nachdruck [bar]	Umschalt-punkt [mm]	Gewicht [g]	Gewicht [g]	Nachdruck [bar]	Gewicht [g]
VAL-01	220	45	62	300	24,6	99,47	98,5 ± 0.5%	225	98,61	+ 0,11
VAL-02	220	45	62	225	24,6	98,61	99,5 ± 0.5%	300	99,47	- 0,03
VAL-03	220	45	62	300	24,6	99,47	98,5 ± 0.5%	225	98,57	+ 0,07
VAL-04	220	45	62	225	24,6	98,57	97,5 ± 0.5%	135	97,42	- 0,08
VAL-05	220	45	90	200	25,3	98,28	100 ± 0.5%	362	100,4	+ 0,4
VAL-06	220	45	60	362	25,3	100,4	97,5 ± 0.5%	144	97,5	0
VAL-07	220	55	90	200	24,6	98,15	99,5 ± 0.5%	316	99,53	+ 0,03
VAL-08	220	55	90	316	24,6	99,53	97,7 ± 0.5%	154	97,6	- 0,1
VAL-09	235	45	72,5	225	25,2	98,74	97,7 ± 0.5%	153	97,85	+ 0,15
VAL-10	235	45	72,5	153	25,2	97,85	99,4 ± 0.5%	307	99,75	+ 0,35
VAL-11	235	40	72,5	225	25,2	99,85	97,6 ± 0.5%	144	97,74	+ 0,14

Abbildung 37: Validierung der Parameterempfehlungen

Bei allen Validierungsversuchsreihen wurde das Ziel-Gewicht durch Anpassung des Nachdrucks gemäß der Empfehlung des Bayes'schen Netzes erreicht. Auch bei unbekanntem Einstellkombinationen oder Zielgewichten außerhalb des bekannten Bereichs wurde das erwünschte Gewicht durch die Nachdruckanpassung erfolgreich erreicht. Außerdem war das Bayes'sche Netz auch in der Lage, falls nötig, Nachdruckempfehlungen außerhalb des bekannten Bereichs zu generieren.

#### 4.9.3 Validierung der Verzugsvorhersage

Als Parameter zur Beurteilung des Verzugs wurde die Gesamtbauteilbreite mit einem Sollwert von 100 mm verwendet. Datenbasis für die Validierung waren die Hauptversu-

che. Sowohl im Trainings- als auch im Testdatensatz sind Versuche aus allen Versuchsreihen (A01 bis A28) enthalten. Mit dem Testdatensatz wurde ein R2-Wert von 92,5 % und ein MAE von 0,11 mm erreicht (vgl. Tabelle 12). Damit ist die Vorhersage zwar etwas schlechter als beim Gewicht, aber immer noch auf einem guten Niveau. Zu beachten ist, dass der Versuchsplan nicht für das Qualitätsmerkmal Verzug, sondern für das Bauteilgewicht optimiert wurde.

Tabelle 12: Erzielte R2- und MAE-Werte für die Verzugsvorhersage des Bayes'schen Netzes

	R2	MAE
<b>Train Set</b>	94,8 %	0,11 mm
<b>Test Set</b>	92,5 %	0,11 mm

#### 4.10 Transfer auf reale Anwendung mit Live-Demonstrator

Zur Validierung des entwickelten Bayes'schen Netzes wurde ein Live-Demonstrator konzipiert und implementiert. Die zugrundeliegende Architektur ist in Abbildung 38 dargestellt. Die für das Modell benötigten Daten aus der Spritzgießmaschine und den Peripheriegeräten werden teilweise direkt über OPC UA oder die EUROMAP77-Schnittstelle an die *Connectware* übertragen. Bestimmte Daten aus dem *DataXplorer* werden zunächst automatisch in einem Python-Skript vorverarbeitet z. B., um Merkmale aus Zeitreihendaten zu extrahieren und anschließend über MQTT an die *Connectware* gesendet. Alle Daten werden in einer Datenbank (*InfluxDB*) gespeichert. Die graphische Benutzeroberfläche des Demonstrators wurde mit Hilfe von *Streamlit* [23] realisiert, einer Open Source Python Bibliothek zur Erstellung interaktiver Webanwendungen. Im Backend werden die Daten zyklusbasiert über MQTT aus der *Connectware* abgefragt und an das mit der Python-Bibliothek *Halerium* implementierte Bayes'sche Netz übergeben. Dieses führt eine Inferenz durch und gibt anschließend Vorhersagen für die Qualitätsgrößen sowie Handlungsempfehlungen an das Backend zurück. Auf Basis dieser Werte wird die Darstellung im Frontend aktualisiert. Der Benutzer hat die Möglichkeit in der graphischen Oberfläche ein Zielgewicht auszuwählen. Wird dieses geändert, wird die Benutzereingabe an das Backend und von dort an das Bayes'sche Netz weitergeleitet, welches die Empfehlungen entsprechend anpasst.

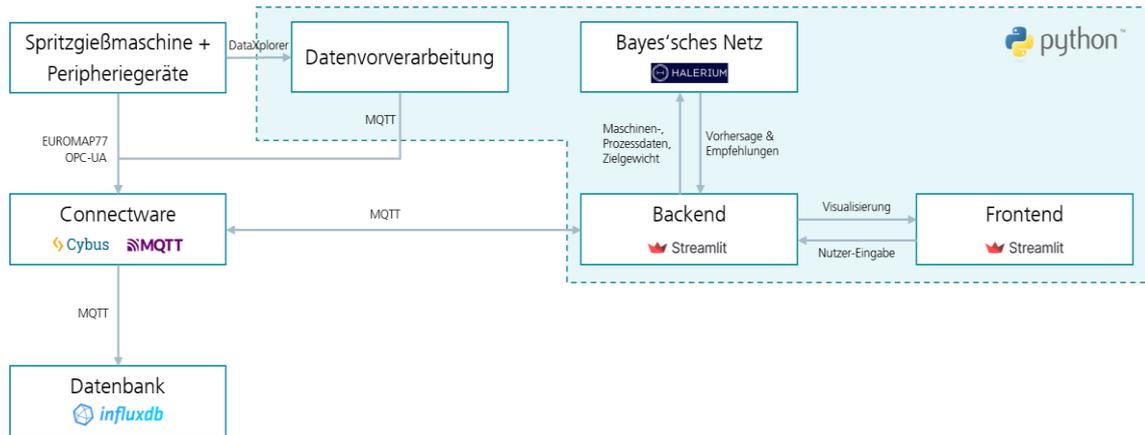


Abbildung 38: Schematische Darstellung der Architektur des Live-Demonstrators

Abbildung 39 zeigt die graphische Benutzeroberfläche des Demonstrators, die in fünf Bereiche unterteilt ist. In der linken Seitenleiste können das Zielgewicht und die zulässige Gewichtstoleranz eingestellt werden. Oben links werden in der Übersicht wichtige Informationen wie aktuelle Uhrzeit, Zyklusnummer, Bauteil und Material angezeigt. Außerdem werden die jeweils aktuellen Prozessparameter und Einstellgrößen mit aktuellem und empfohlenem Wert angezeigt. Eine Ampel symbolisiert, ob Ist- und Sollwert übereinstimmen. Im unteren linken Bereich wird der Gewichtsverlauf über die Zyklen dargestellt. Das Sollgewicht sowie die obere und untere Toleranzgrenze werden durch die Linien und das grüne Band visualisiert. Oben rechts befindet sich die Qualitätsprognose, darunter der Diagnosebereich.

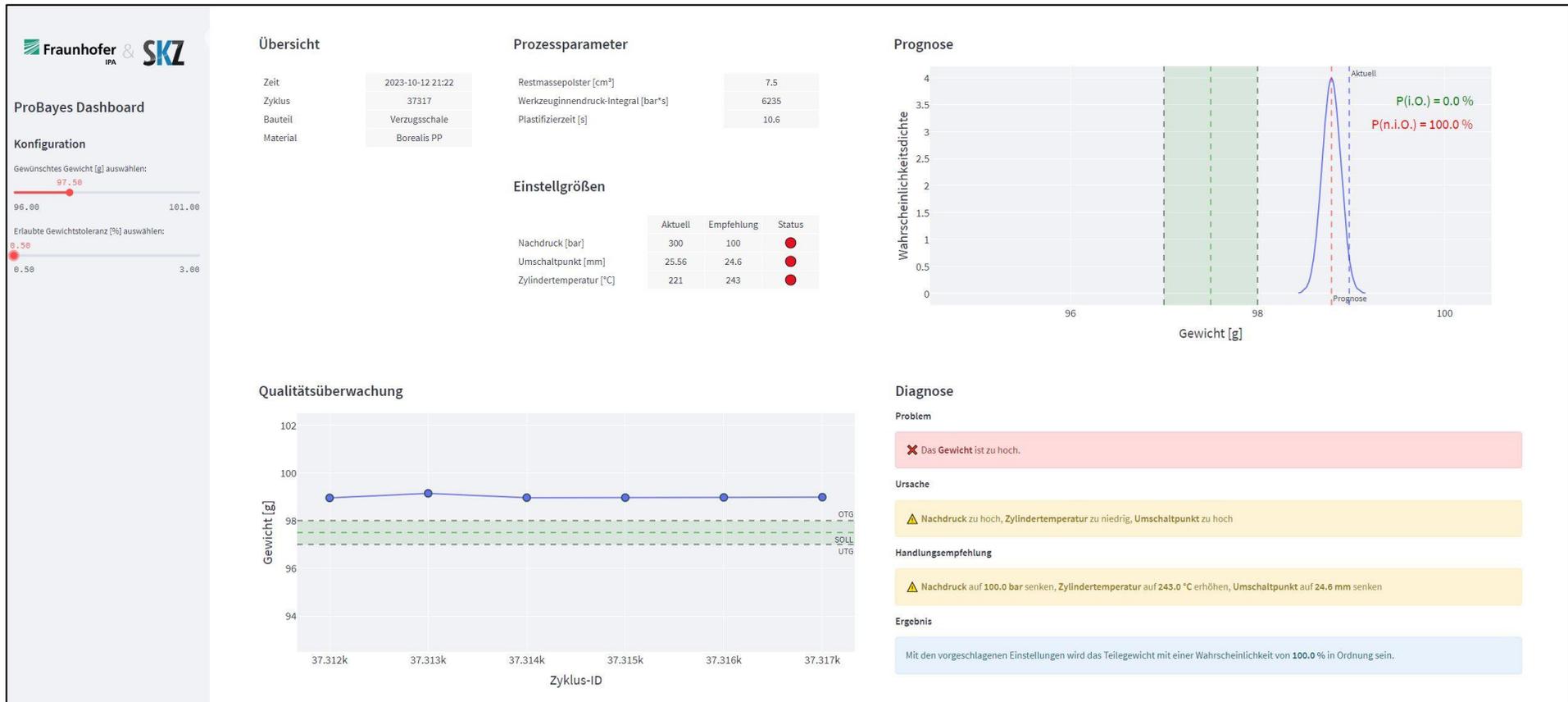


Abbildung 39: Graphisches User Interface des Demonstrators, umgesetzt mit *Streamlit* [23]

Der Prognosebereich enthält ein Wahrscheinlichkeitsdichtediagramm und ist vergrößert in Abbildung 40 dargestellt. Die grüne gestrichelte Linie zeigt das Sollgewicht an, die grüne Fläche den Toleranzbereich. Die blaue gestrichelte Linie zeigt den zuletzt von der Waage gemessenen Gewichtswert an und die rote gestrichelte Linie die Vorhersage des Bayes'schen Netzes für das Gewicht auf der Grundlage der aktuellen Einstellparameter. Die blaue Wahrscheinlichkeitsdichtekurve gibt an mit welcher Wahrscheinlichkeit das Gewicht innerhalb bzw. außerhalb der Toleranz liegen wird. Im dargestellten Fall liegt die Kurve vollständig außerhalb des grünen Toleranzbereichs, so dass eine Sicherheit von 100% angegeben wird, dass das Gewicht nicht in Ordnung ist.

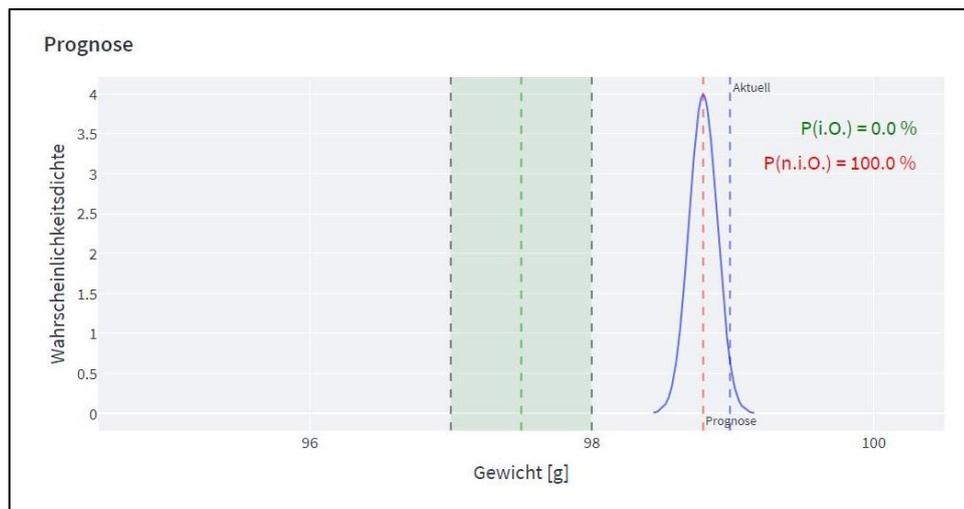


Abbildung 40: Prognosebereich des Dashboards

In Abbildung 41 ist der Diagnosebereich des Demonstrator-Dashboards (entspricht unten rechts in Abbildung 39) vergrößert dargestellt. Er ist in die vier Bereiche „Problem“, „Ursache“, „Handlungsempfehlung“ und „Ergebnis“ unterteilt.



Abbildung 41: Diagnosebereich des Dashboards

Das Problem wird durch einen einfachen Soll-Ist-Vergleich der Zielgröße identifiziert. Die Ursache wird ermittelt, indem der Sollwert der Zielgröße in das Bayes'sche Netz eingegeben wird, dieses Empfehlungen für Einstellparameter ausgibt und diese Empfehlungen mit den aktuellen Einstellungen verglichen werden. Die Handlungsempfehlung

entspricht den Sollwerten der Stellgrößen, die zuvor bereits für die Ursachenermittlung berechnet wurden. Als Ergebnis wird eine Wahrscheinlichkeit angegeben, mit der die Zielgröße im Toleranzbereich liegt, wenn die empfohlenen Einstellungen übernommen werden. Um die aktuelle Situation auf den ersten Blick beurteilen zu können, werden die Angaben zu Problem, Ursache und Handlungsempfehlung in verschiedenen Farben (grün, wenn alles in Ordnung ist bzw. gelb oder rot, falls nicht) und mit verschiedenen Symbolen (grüner Haken, rotes Kreuz, gelbes Warnzeichen) dargestellt.

#### **4.11 Technologietransfer und Dokumentation**

Die Prozessdaten der im Projekt durchgeführten Versuchsreihen (geteilt in Vorversuche und Hauptversuche) sind in aufbereiteter Form auf der europäischen Datenplattform EU-DAT unter der Lizenz CC-BY veröffentlicht worden, d.h. die Daten dürfen in jeglicher Hinsicht genutzt werden, unterliegen aber einer Zitationspflicht [15–17]. Dort sind neben den reinen Daten auch eine ausführliche Datenbeschreibung, sowie eine Liste der Parameternamen samt Kurzerläuterung und ein Pandal-Profiling Report, der einen schnellen Überblick über die Daten ermöglicht, hinterlegt. Die Daten werden bereits von einigen Firmen und Forschungsinstituten, darunter Katulu GmbH, HBF Dr. Bauer KG, Strem.ai GmbH, Kunststoff-Zentrum Leipzig, Jheronimus Academy of Data Science und weiteren, für verschiedene Zwecke verwendet.

Im Rahmen des Projekts wurden zwei Paper erstellt, welche beide auf der 16th CIRP ICME 2022 veröffentlicht wurden. Das Paper von Krauß et al [24] ist ein Literature Review, welches den aktuellen Stand von KI-Systemen zur Prädiktion von Bauteilgewicht im Spritzguss zusammenfasst. Das Paper von Borchardt et al [25] beschreibt die in Pro-Bayes entwickelte Methodik zur Erstellung eines Bayes'schen Netzes zur Überwachung des Spritzgussprozesses. Des Weiteren wurde das Projekt und die Ergebnisse beim Berlin Workshop on Artificial Intelligence for Engineering Applications 2022 im Rahmen der SEFM 2022 (Internationale Konferenz Software Engineering and Formal Methods) mit Poster [26] und Vortrag [27] vorgestellt.

Der Kunststofffachwelt wurden die Projektfortschritte und -erkenntnisse in Beiträgen für die Fachzeitschriften *Plastverarbeiter* [28] und *SKZ-Aktuell* [29] bekannt gemacht. Des Weiteren werden die Ergebnisse regelmäßig in den „Arbeitskreis Digitalisierung im Spritzgießen – DigInject“ (Gremium bestehend aus über 20 Firmen, organisiert vom SKZ) hineingetragen sowie dem SKZ-Fachbeirat präsentiert. Zusätzlich wurde das Projekt im „The Erium Podcast“ [30] vorgestellt, dessen Hörerschaft aus angehenden Data Scientists und Young Professionals besteht, um selbigen einen Data Science Anwendungsfall aus der Praxis näher zu bringen.

## **5 Diskussion der Ergebnisse**

### **5.1 Fazit der Vorversuche**

Als Zielgröße bzw. Qualitätsmerkmal des Netzes wurde das Bauteilgewicht ausgewählt. Das Gewicht ist eine einfach und exakt messbare Größe. Im Gegensatz zum Qualitätsmerkmal Verzug sind die Abhängigkeiten zu den Einstellparametern gut abbildbar. Zudem bleibt das Gewicht nach der Entformung konstant, während sich der Verzug infolge der Nachkristallisation bei teilkristallinen Kunststoffen noch über Stunden ändern kann.

Des Weiteren sind vereinzelt Messfehler am Bildverarbeitungssystem aufgrund von stark unterschiedlichen Lichtverhältnissen im Technikum aufgetreten. Trotz zusätzlicher Abdunkelung im Bereich der Bildaufnahme konnte dieser Fehler nie vollständig eliminiert werden. Dies lag zum Teil auch an der unterschiedlichen Ausprägung des Verzugs, welcher verschiedene Schatten erzeugte, sodass die optische Kantenerfassung nicht mit den ursprünglichen Einstellungen bzw. Toleranzen funktionierte. Die Bilddateien der fehlerhaft gemessenen Bauteile wurden erneut von der Software mit angepasster Kantenerfassung vermessen, um einen vollständigen und fehlerfreien Trainingsdatensatz zu garantieren. Für eine Live-Regelung wurde sich angesichts der genannten Gründe für das zuverlässiger messbare Qualitätsmerkmal Gewicht entschieden.

Nach Auswertung der Vorversuche wurde erkannt, dass für die Quantifizierung der Bayes'schen Netze umfassendere Versuchspläne mit mehr Zwischenstufen notwendig sind. Um den Versuchsaufwand zu stemmen, wurde der Fokus auf ein Material (PP) gelegt. Dies ist auch mit der Tatsache begründet, dass für das rABS keine detaillierte Materialkarte zur Simulation des Spritzgießprozesses bereitsteht und die Simulation bei diesem Material nicht zur Erweiterung des Datensatzes beitragen kann.

### **5.2 Vollvernetzte Spritzgießzelle**

Es wurde eine vollvernetzte Spritzgießzelle aufgebaut und mit einer Datenerfassung ausgestattet. Die Spritzgießzelle vernetzt heterogene Maschinen, Peripheriegeräte und Messsysteme mit Hilfe aktueller Kommunikationsstandards wie OPC UA und MQTT. Durch den Einsatz der Inline-Messtechnik wird ein direktes Feedback der Bauteilqualität ermöglicht. Alle Daten werden zyklusbezogen in eine Influx-Datenbank geschrieben und ein umfassendes Datenabbild des Prozesses aufgezeichnet. Abbildung 42 zeigt die Architektur der vernetzten Spritzgießzelle.

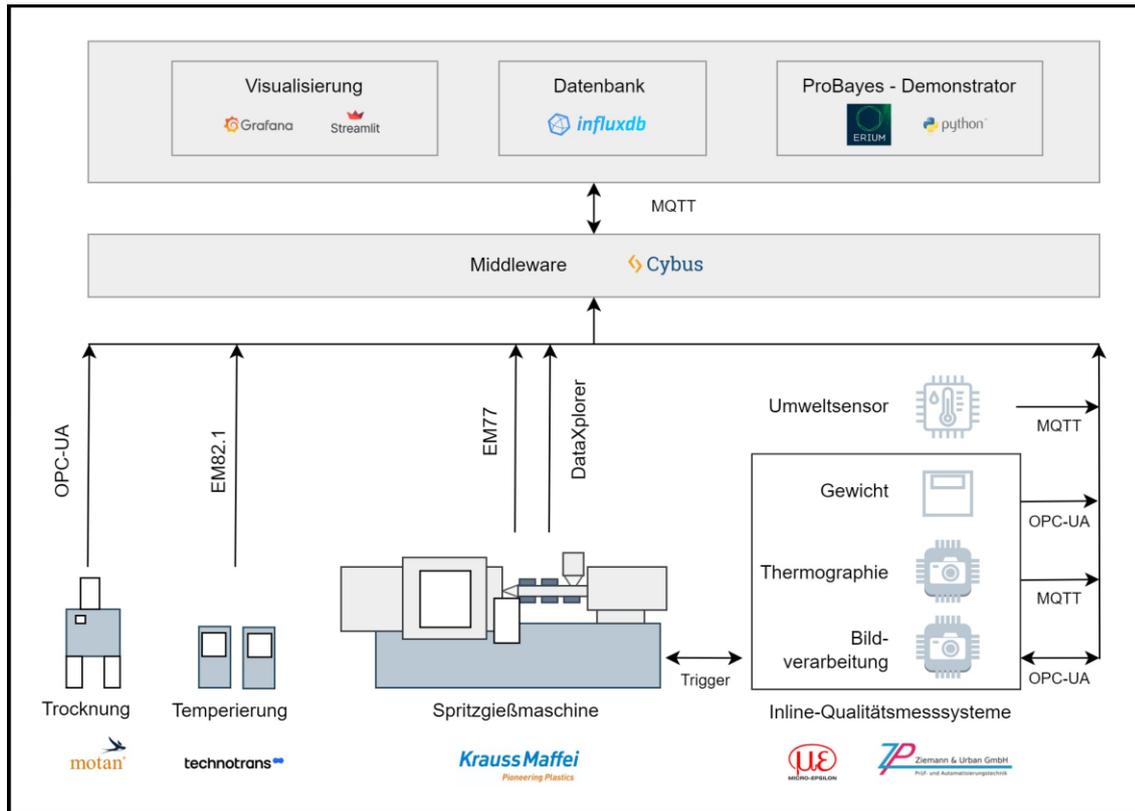


Abbildung 42: Schematische Darstellung aller vernetzten Geräte und der Spritzgießmaschine

Der Versuchsdatensatz, bestehend aus Einstellparametern, Prozessgrößen und Qualitätsmerkmalen, ist auf der öffentlich zugänglichen Seite EUDAT bereitgestellt. Data Scientists können diesen als Trainingsdatensatz für sonstige Entwicklungsprojekte nutzen.

Gleichermaßen kann die Spritzgießzelle für eine Live-Demonstration der entwickelten Bayes'schen Netze eingesetzt werden. Die Daten-Pipeline ist derart modular gestaltet, dass neue Prozessparameter mit geringem Aufwand in die Pipeline neu hinzugefügt werden können. Dies wurde aufgrund der iterativen Vorgehensweise bei der Netz-Entwicklung bereits mehrfach genutzt, ist aber auch für Folgeprojekte von großem Vorteil. Gleichermaßen ist die Visualisierung sowie auch die Implementierung des Bayes'schen Netzes jeweils als separates Modul gestaltet, sodass auch andere Modellierungsmethoden und Darstellungsweisen auf dieser neuen Forschungsplattform eingesetzt werden können.

### 5.3 Sim-to-Real Gap

Die Simulation des Spritzgießprozesses war dazu gedacht, den Datensatz für die Quantifizierung des Bayes'schen Netzes anzureichern und somit Versuchsaufwand an der realen Maschine einzusparen. Der Vergleich der Simulationsergebnisse mit den real durchgeführten Versuchen ergab allerdings, je nach Bauteilmerkmal, eine signifikante Abweichung von Simulation zu Realität, weshalb Maßnahmen zur Reduktion dieser Sim-to-Real Gap ergriffen wurden.

Dies umfasst eine Materialcharakterisierung mit Hilfe einer DSC-Messung, um das Kristallisationsverhalten in der Simulation besser abbilden zu können. Außerdem wurden die

Eingangsparameter der Simulation an die reale Düsen- und Schneckengeometrie angepasst, um die Sim-to-Real Gap zu minimieren.

Exemplarisch ist ein Vergleich anhand des Bauteilgewichts in den folgenden beiden Abbildungen dargestellt, der die Signifikanz des Sim-to-Real Gap zeigt.

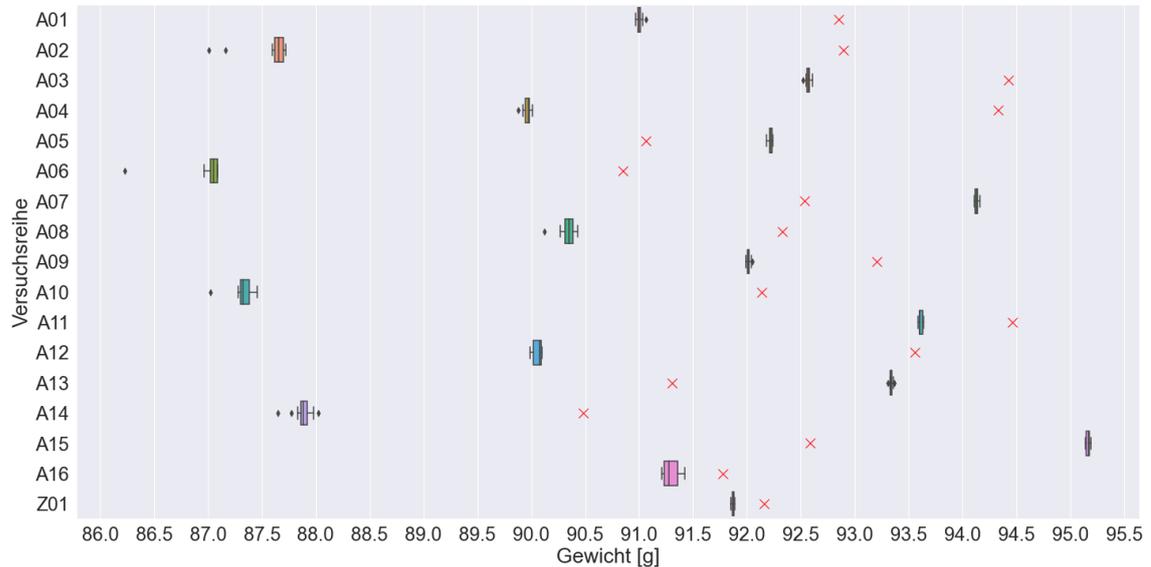


Abbildung 43: Auszug aus den Vorversuchen Versuchsplan für PP-HE125MO des Vergleichs von gemessenen und simulierten Formteilgewichten vor dem Bug-Fix (rotes X = Standardsimulation, Boxplot = 12 real gespritzte Teile)

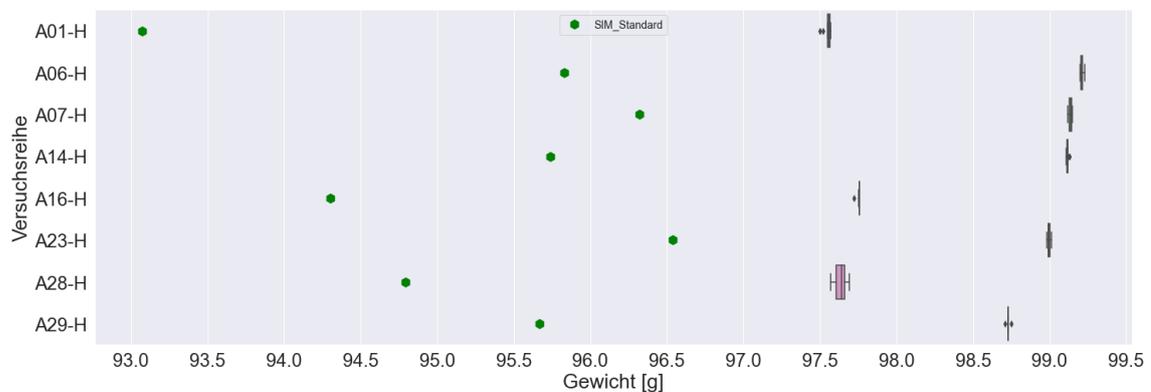


Abbildung 44: Auszug aus dem D-Optimalen Versuchsplan für PP-HE125MO des Vergleichs von gemessenen und simulierten Formteilgewichten nach dem Bug-Fix (grünes Sechseck = Standardsimulation, Boxplot = 12 real gespritzte Teile)

Um Anwenderfehler auszuschließen, wurden die Simulationen gemeinsam mit den Experten von SimpaTec begutachtet und als mögliche Fehlerquellen für den Sim-to-Real Gap der nicht berücksichtigte Maschinenzwilling genannt und die verwendeten Materialdaten hinterfragt.

Um dem Sim-to-Real Gap entgegenzuwirken, sollte wie bereits erwähnt die Berücksichtigung der Maschine in Form eines digitalen Maschinenzwillings in Verbindung mit dem Plasticizing Modul von Moldex3D – hiermit ist das Plastifizieraggregat und die darin stattfindenden Aufschmelzvorgänge detailliert abzubilden – und eine ganzheitliche rheo-

logische Materialquantifizierung erfolgen. Da dies jeweils mit enormen Kosten verbunden ist, war es nicht möglich die genannten Maßnahmen zusätzlich im Rahmen des Projekts durchzuführen.

Um den generellen Einfluss der potenziellen Optimierungsmaßnahmen zu evaluieren, wurde zunächst für das PP vom Typ HE125MO eine Teil-Materialcharakterisierung der Kristallinität durchgeführt. Im Rahmen dessen sind die generischen Kristallisationsdaten durch analytisch vermessene Daten zu ersetzen. Die Kristallisation kann bei Betrachtung des Bauteilgewichts einen signifikanten Einfluss auf das Nachdruck- und Abkühlverhalten darstellen. Zudem bestehen am SKZ die notwendigen analytischen Laborgeräte, um diese Messung intern durchzuführen.

Abbildung 45 zeigt den Vergleich der generischen (orange) zu den vermessenen (blau) Datensätzen. Die vermessenen Datensätze zeigen abgesehen von einem absoluten Versatz der Kurven hin zu niedrigeren Temperaturen eine deutlich flachere Steigung mit Annäherung an die maximale relative Kristallinität bzw. steiler bei minimaler relativer Kristallinität. Ersteres trifft vor allem auf die Kurven mit höhere Abkühlraten zu.

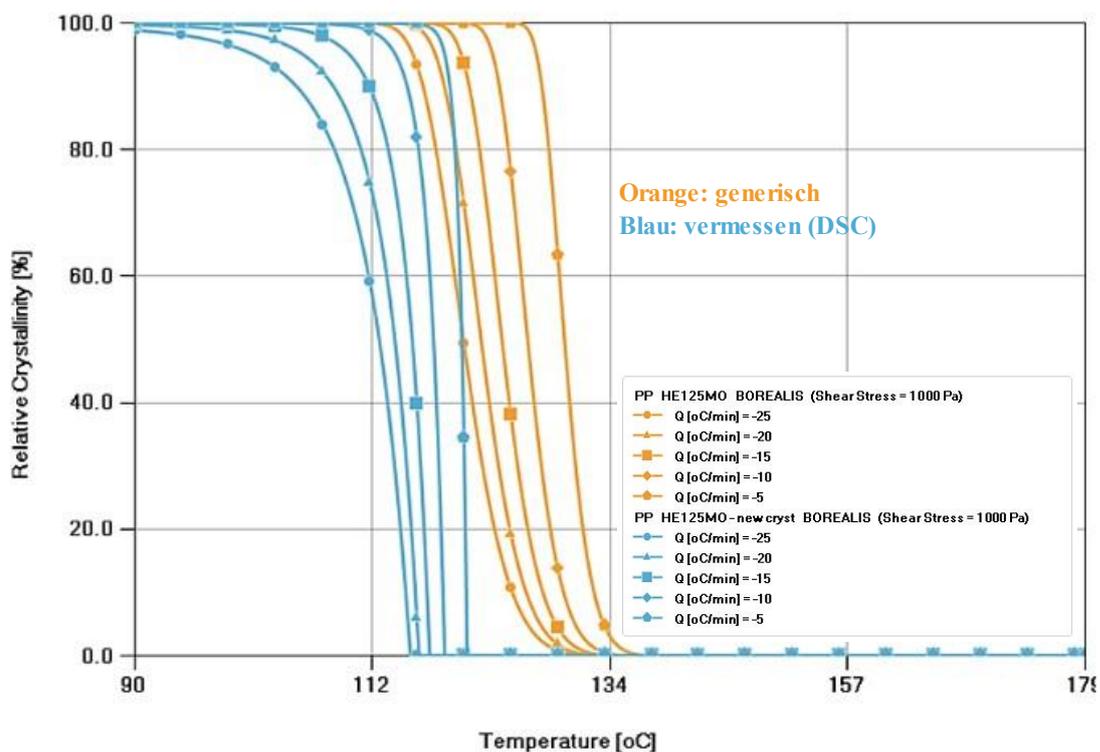


Abbildung 45: Vergleich der generischen (orange) und vermessenen (blau) Daten zur Kristallinität des Materials PP-HE125MO

Im Ergebnis (Abbildung 46) zeigte sich, dass die Gewichte des Bauteils über alle Versuchsreihen hinweg geringer ausfallen. Die relative Gewichtsveränderung von Standardsimulation zu der Simulation mit modifizierten Materialdaten streut zudem unterschiedlich stark.

Die Tatsache, dass das Bauteilgewicht weiter sinkt, erscheint in Anbetracht von Abbildung 45 schlüssig. Ein Kristall stellt molekular betrachtet die dichteste gepackte Struktur

dar und weist damit die vergleichsweise höchste Dichte auf. Nimmt die relative Kristallinität nun ab bedeutet das im Umkehrschluss, dass der amorphe Anteil im Material zunimmt, der wiederum weniger dicht gepackt ist. Daraus resultiert ein geringeres Bauteilgewicht. Dieser Effekt wird bei höheren Abkühlraten (Faktor 10), wie sie beim Spritzgießen oftmals auftreten können nochmals verstärkt.

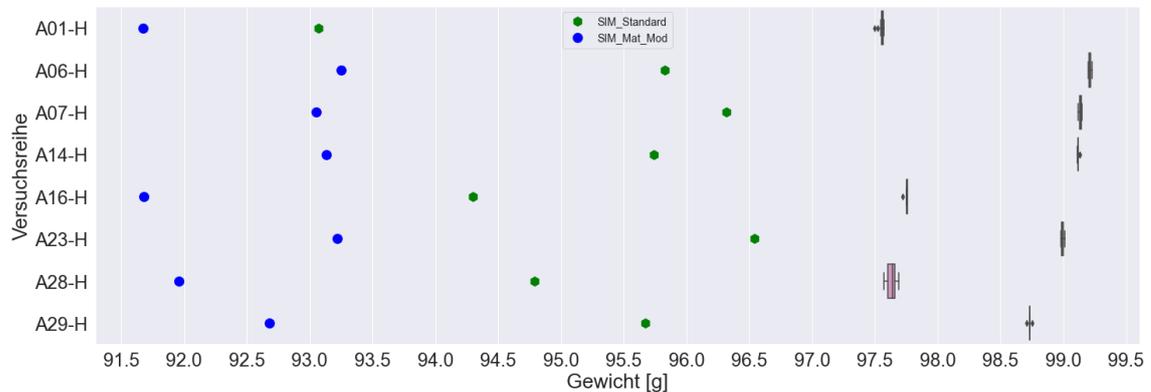


Abbildung 46: Auszug aus dem D-Optimalen Versuchsplan für PP-HE125MO des Vergleichs von gemessenen und simulierten Formteilgewichten (grünes Sechseck = Standardsimulation, blauer Kreis = Simulation mit modifizierten Materialdaten, Boxplot = 12 real gespritzte Teile)

Im nachfolgenden Schritt wurde eine Maschinencharakterisierung „light“ durchgeführt. Hierbei kam ersatzweise eine Spritzgießmaschine vom Typ KM160PX-SP750-50 zum Einsatz, die bereits als Maschinenzwilling in der Software Moldex3D hinterlegt war. Für die Maschinencharakterisierung „light“ wurde die Schneckengeometrie vom hinterlegten 50 mm Aggregat durch die Geometriedaten des am SKZ eingesetzte 35 mm Aggregat ersetzt. Versuche zur Maschinenbewegungsreaktion fanden nicht statt.

Die Ergebnisse in Abbildung 47 zeigen die angepasste Schneckengeometrie bzw. die Verwendung der damit einhergehenden Spritzgießmaschine (rote Dreiecke) in Relation zu den vorangegangenen Ergebnissen. Die Maßnahmen der Maschinencharakterisierung „light“ zeigen keinen signifikanten Einfluss.

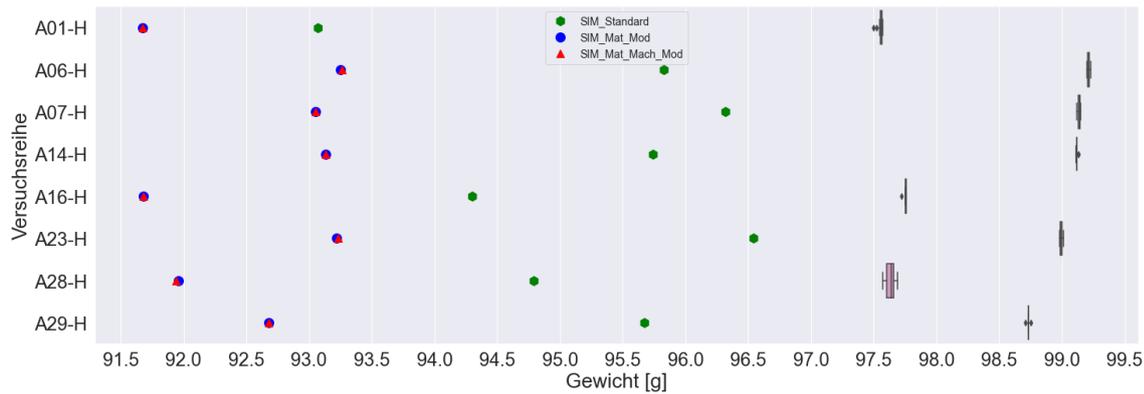


Abbildung 47: Auszug aus dem D-Optimalen Versuchsplan für PP-HE125MO des Vergleichs von gemessenen und simulierten FormteilmGewichten (grünes Sechseck = Standardsimulation, blauer Kreis = Simulation mit modifizierten Materialdaten, rotes Dreieck = Simulation mit modifizierten Material- und Maschinendaten, Boxplot = 12 real gespritzte Teile)

Zusammengefasst ist keine signifikante Verbesserung mit den vereinfachten bzw. reduzierten Maßnahmen zu erzielen. Da weitere Maßnahmen nicht mehr im Projekt umgesetzt werden können, kann der Datensatz für die Quantifizierung des Bayes'sches Netzes nicht mit Simulationsdaten angereichert werden. Das Bayes'sche Netz wird daher rein mit den real gemessenen Prozessdaten quantifiziert.

#### 5.4 Validiertes Bayes'sches Netz für die Vorhersage des Bauteilgewichts und des Verzugs

Die entwickelten Bayes'schen Netze wurden durch verschiedene Metriken ( $R^2$ , MAE) sowie Visualisierungen (Wahrscheinlichkeitsdichteplot, Konfusionsmatrix) bewertet und verglichen. Die finale Version des Netzes für das Bauteilgewicht ist in Abbildung 48 dargestellt. Das Bayes'sche Netz für die Gewichtsvorhersage ist in der Lage das Bauteilgewicht mit hoher Genauigkeit vorherzusagen. Die hohe Sicherheit der Vorhersage spiegelt sich auch im niedrigen mittleren absoluten Fehler (MAE) wider. Die Generalisierungsfähigkeit der Netze auf unbekannte Versuchsreihen wurde ebenfalls überprüft. Hierbei wurden zwar etwas schlechtere Werte ermittelt, dennoch war eine gute Vorhersage auch für komplett unbekannte Versuchsreihen möglich. Auch bei der Validierung der ausgegebenen Parameterempfehlungen wurden sehr gute Ergebnisse erzielt.

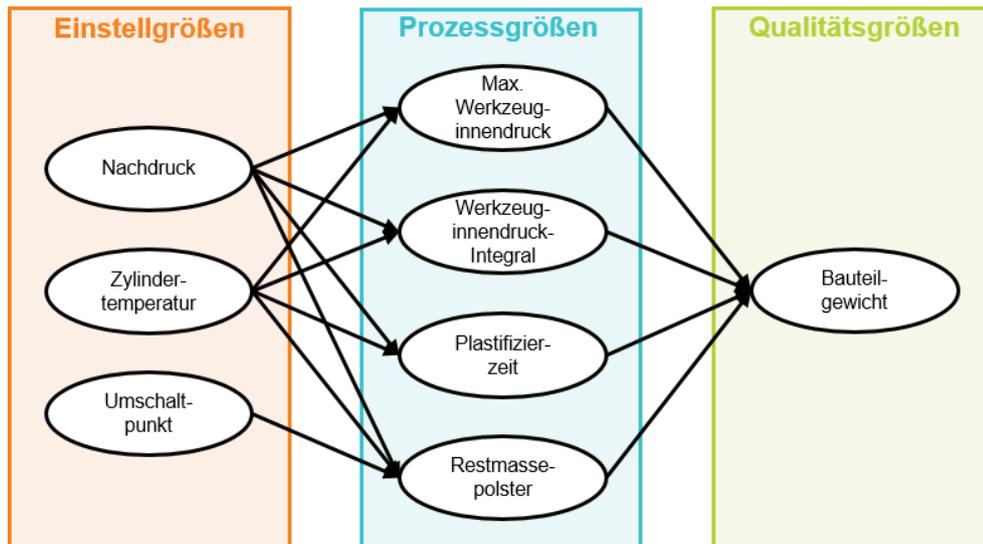


Abbildung 48: Finales Teilnetz für den Qualitätsindikator „Bauteilgewicht“

Das finale Teilnetz für den Verzug ist in Abbildung 49 dargestellt. Auch hier wurden gute Ergebnisse erzielt, die sich durch einen hohen R2-Wert und einen niedrigen MAE auszeichnen.

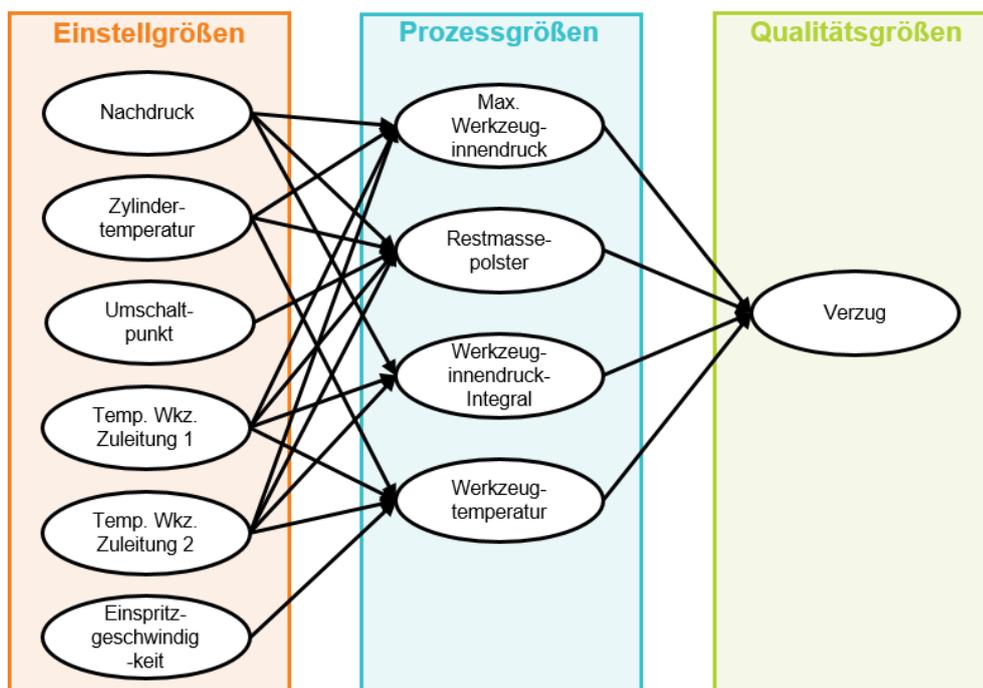


Abbildung 49: Finales Teilnetz für den Qualitätsindikator „Verzug“

## 5.5 Methodik für die Entwicklung Bayes'scher Netze für industrielle Anwendungsfälle

In der Literatur ist keine standardisierte und übertragbare Methodik für die Entwicklung von Bayes'schen Netzen für industrielle Anwendungsfälle zu finden. Daher wurde im Rahmen des Projektes eine Methodik entwickelt, die in Abbildung 50 dargestellt ist. Sie

orientiert sich an der gängigen Vorgehensweise zur Entwicklung von maschinellen Lernmodellen, enthält jedoch zusätzlich spezifische Schritte, die nur für Bayes'sche Netze relevant sind (markiert durch die roten gestrichelten Linien). Die einzelnen Schritte sind in grün dargestellt. Darunter sind in blau die Ergebnisse der einzelnen Schritte dargestellt. Im ersten Schritt werden Daten generiert, z. B. durch das Abfahren von Versuchsreihen an einer Maschine. Die gewonnenen Daten werden aufbereitet und analysiert. Dazu gehören Korrelationsanalysen und die Auswahl relevanter Merkmale („Feature Selection“). Zusammen mit dem Expertenwissen ergeben sich daraus die Zusammenhänge zwischen den Parametern. Dieses Wissen fließt auch in den vierten Schritt, die Graph-Entwicklung, ein. Zuvor werden die Daten in Trainings- und Testdaten aufgeteilt. Die Topologie des Graphen basiert auf dem Expertenwissen und den datenbasierten Erkenntnissen, wobei widersprüchliche Aussagen diskutiert und abgewogen werden müssen, da sowohl die individuelle Erfahrung der Experten als auch die Wahl des Versuchsplans bei der Datenerzeugung eine Rolle spielen. Im Graphen werden für die einzelnen Parameter die statistischen Verteilungen hinterlegt, die aus dem Trainingsdatensatz ermittelt werden. Dieser Graph wird nun an die Daten angepasst, d.h. ein Modell wird trainiert, wobei unbekannte Parameter gelernt werden. Das trainierte Modell wird als Posterior-Graph bezeichnet. Mit diesem können Vorhersagen, z. B. von Qualitätsgrößen, generiert werden. Das Modell wird evaluiert und ggf. optimiert.

Die Übertragbarkeit der entwickelten Modelle auf andere Maschinen, Werkzeuge oder Materialien wurde bisher nicht untersucht. Die Validierung erfolgte ausschließlich für die einzigartige Kombination aus eingesetzter Maschine, Werkzeug (Verzugsschale) und eingesetztem Material. Im Rahmen des Projekts konnte nicht explizit nachgewiesen werden, dass eine Übertragbarkeit der topologischen Struktur möglich ist. Dies ist Gegenstand weiterer Forschung. Selbst wenn eine Übertragung der topologischen Struktur möglich wäre, bestünde die Herausforderung darin, dass für jede neue Maschine-Werkzeug-Material-Kombination neue Prozessdaten generiert werden müssten, um ein angepasstes Prozessmodell trainieren zu können. Da dies für die auf Kosten- und Zeiteffizienz bedachte industrielle Praxis eine große Hürde darstellt, besteht eine Lösungsmöglichkeit darin, die Generierung der notwendigen Prozessdaten in einen Standardprozess wie die Werkzeugabmusterung zu integrieren. Dies ist Gegenstand des bereits beantragten Folgeprojektes „ProBayes ReLOAD“.

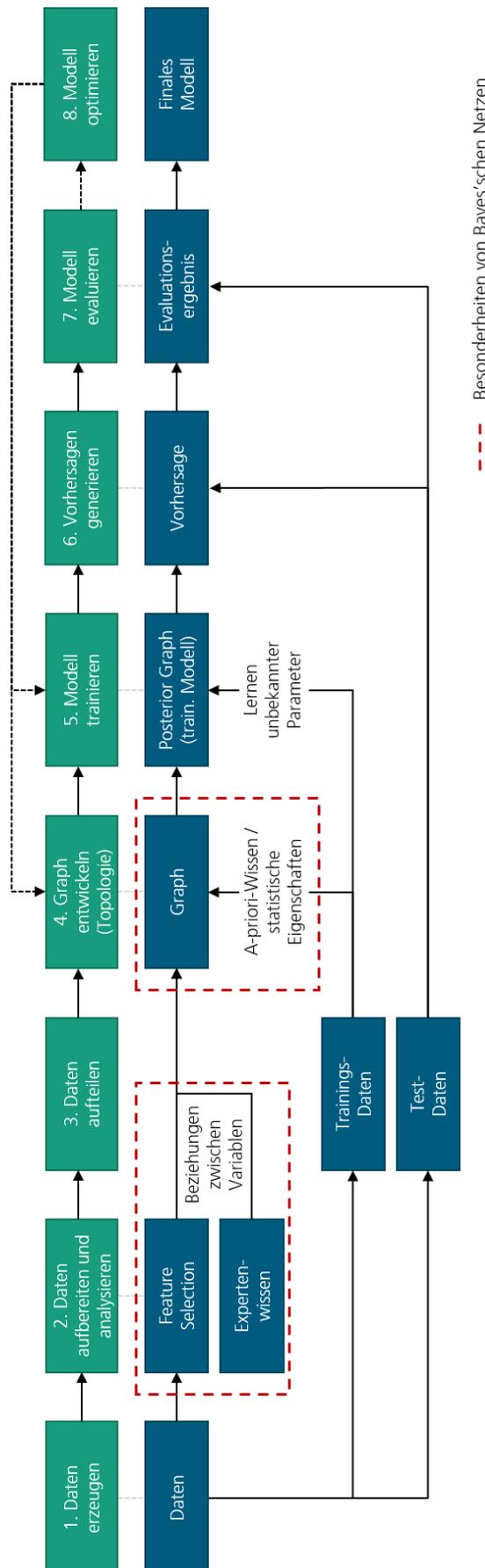


Abbildung 50: Erarbeitete Methodik für die Entwicklung Bayes'scher Netze

## 6 Zusammenfassung und Ausblick

Im Projekt konnten vielversprechende Ergebnisse erarbeitet werden und die aufgestellten Hypothesen teilweise verifiziert werden. So ist die Modellierung des Spritzgießprozesses in Bezug auf ausgewählte, relevante Qualitätsmerkmale mit Hilfe von Bayes'schen Netzen grundsätzlich demonstriert worden, wobei die dynamische Komponente aus Hypothese 2 nicht notwendig war. Der Ansatz hierarchischer Bayes'scher Netze (vgl. Hypothese 1) wurde durch die Zerlegung des Gesamt-Spritzgießprozesses in einzelne Teilbereiche abhängig vom betrachteten Qualitätsmerkmal ersetzt.

Zentrale Ergebnisse des Projekts sind neben der aufgebauten Spritzgießzelle und den veröffentlichten Datensätzen, das entwickelte und validierte Bayes'sche Netz für das Bauteilgewicht sowie die dem zu Grunde liegende Methodik zur Entwicklung Bayes'scher Netze für industrielle Prozesse. Der entwickelte und umgesetzte Live-Demonstrator beweist die Funktionsfähigkeit der Bayes'schen Netze für die Prozessüberwachung und -optimierung und stellt den Whitebox Charakter, mit dessen Hilfe nicht nur eine Vorhersage der Zielgröße, sondern auch Ursachenanalysen und die Ableitung von Handlungsempfehlungen möglich ist, als Vorteil der Bayes'schen Netz eindrücklich heraus.

### 6.1 Ergebnis 1 – Dateninfrastruktur und Prozessdaten

Durch den Aufbau der Spritzgießzelle mit allen Peripheriegeräten und den eingesetzten Schnittstellen (maßgeblich OPC UA) wurde Pionierarbeit geleistet, indem neueste Companion Specifications bzw. noch unveröffentlichte Spezifikationen einem Praxistest unterzogen wurden. Die Mitglieder des projektbegleitenden Ausschusses sowie des „Arbeitskreises Digitalisierung im Spritzguss - DigInject“ profitieren direkt von dieser Arbeit, indem gefundene Bugs der OPC UA Server auf den Geräten beseitigt werden können und die Gerätehersteller lernen, was Kunden wirklich brauchen. Insbesondere KMU mit begrenztem IT-Know-How und geringerer Fachkräfteverfügbarkeit wird so der Weg zur Digitalisierung erleichtert.

Durch die intensiven Untersuchungen der verfügbaren Prozessdaten im Projekt, konnte einer Fragestellung, die viele Firmen umtreibt, auf den Grund gegangen werden: „Welche Prozessdaten sind überhaupt von Relevanz für eine Prozessüberwachung oder -regelung und welche Messsysteme sind zu deren Erhebung von Nöten?“

Dadurch dass die Prozessdaten, welche in Versuchsreihen auf der Spritzgießzelle erhoben wurden, in für Data Scientists gut nutzbarer Form veröffentlicht wurden, können sie sowohl für das laufende als auch für zukünftige Forschungsprojekte herangezogen werden. Das im Projekt erzeugte Wissen im Umgang mit diesen Prozessdaten stellt ein weiteres Ergebnis dar, das für zukünftige Forschung hilfreich sein wird. So wurden bereits viele Korrelationen und Zusammenhänge algorithmisch untersucht und von Prozessexperten bewertet.

Zur geschaffenen Infrastruktur gehört des Weiteren auch die Middleware, die Anbindung an Datenbanken und der ProBayes-Demonstrator. Alles ist modular und wiederverwertbar konzipiert, sodass es für andere Prozessregelungen oder KI-Modelle mit geringem Aufwand angepasst werden kann. Die in *Streamlit* entwickelte Visualisierung ist für

Bayes'sche Netze ausgelegt, jedoch kann der zugrundeliegende Quellcode (inkl. Vorverarbeitung, *MQTT-Client* etc.) auf andere Visualisierungsarten angepasst werden.

## 6.2 Ergebnis 2 – Sim-to-Real Gap

Das im Projekt erstellte Simulationsmodell in Moldex3D ist hochgradig individuell. Zum Vergleich von Simulation und Realität wurden die gleichen Versuchsreihen sowohl abgefahren als auch simuliert. Dabei zeigt sich eine Lücke, die zu groß ist, als dass Simulationsdaten auch für das Training des an der realen Maschine einzusetzenden Prozessmodells genutzt werden könnten. Erste Maßnahmen zur Minimierung der Sim-to-Real Gap, wie eine Materialcharakterisierung über DSC-Messung, wurden im Projekt umgesetzt, die allerdings keine Verbesserung brachten. Weitere Ansätze wurden erdacht, wie z. B. die Charakterisierung der Spritzgießmaschine und das Einbeziehen der Schnecke in die Simulation, konnten allerdings nicht vollumfänglich im Projektrahmen umgesetzt werden. Jedoch wurden weiterer Forschungsbedarf abgeleitet und weitere Forschungsprojekte geplant. Die bisher getätigten Untersuchungen zur Vergleichbarkeit können als Benchmark herangezogen werden und können so die Verbesserungen im Angleichen von Simulation und Realität quantifizieren.

## 6.3 Ergebnis 3 – Bayes'sches Netz zur Prozessüberwachung und Methodik

Die Methodik zur Entwicklung Bayes'scher Netze wurde im laufenden Projekt erarbeitet, erprobt, dokumentiert und veröffentlicht [25]. Anhand dieser Methodik wurde ein Bayes'sches Netz für die Vorhersage und Überwachung des Formteilgewichts entwickelt, welches in einem Live-Demonstrator validiert wurde. Die grundlegende topologische Struktur der entwickelten Bayes'schen Netze (für die Qualitätsmerkmale Bauteilgewicht bzw. Verzug) wurde erarbeitet, wobei eine Übertragung auf andere Maschinen-Werkzeug-Material Kombinationen denkbar ist, aber noch nicht im Rahmen des Projekts nachgewiesen wurde.

## 6.4 Ausblick

Zusammenfassend kann festgehalten werden, dass die Projektergebnisse die Grundlagen gelegt haben für einen weiteren Entwicklungsschritt in der Spritzgießfertigung. Dieser erstreckt sich über die Betrachtung der reinen Maschinenverfügbarkeit und -stabilität hinaus hin zu einer ganzheitlichen Betrachtung von Maschine und Peripherie sowie der aktiven Nutzung der verfügbaren Prozessparameter zur Optimierung der Produktion.

„ProBayes“ wird als eine Art „Keim-Projekt“ gesehen, woraus sich viele weitere Forschungsthemen eröffnen:

- *Tiefgehende Untersuchung zur Quantifizierung und Minimierung der Sim-to-Real Gap*  
Bei hinreichender Minimierung der Sim-to-Real Gap können sowohl aussagekräftigere Simulationen durchgeführt werden als auch Simulationen als Mittel zur Erzeugung von Trainingsdaten für real eingesetzte Prozessmodelle nutzbar gemacht werden.
- *Übertragbarkeit von Prozessmodellen auf andere Maschinen, Werkzeuge und Materialien*

In weiteren Projekten (Forschung und Industrie) kann die schnelle Adaption der Modelle auf spezifische Werkzeuge untersucht werden. Dabei wird die Erzeugung und praxisnahe Integration der nötigen Trainingsdaten in Standardabläufen, wie eine Werkzeugbemusterung, angestrebt.

- *Niveau der Datenerfassung und -haltung*

In ProBayes wurde durch den Aufbau der Spritzgießzelle der industrielle Stand der Technik hinsichtlich Messtechnik und Datenerfassung übertroffen. Dieser steinige Weg ist im Folgenden nun leichter zu begehen, denn die Vorgehensweise zum Aufbau einer funktionstüchtigen Architektur ist in diesem Bericht dokumentiert und wird in u. a. durch Vorträge, Beratungs-Workshops und Kurse in die Industrie transferiert.

- *OPC UA Companion Specifications*

Die OPC UA Companion Specifications (z. B. EUROMAP77) wurden im Projekt intensiv genutzt und herausgearbeitet, welche dort verfügbaren Parameter für die Prozessbewertung relevant sind. Darüber hinaus wurden ebenfalls wichtige Werte aus den Zeitreihendaten extrahiert, welche bisher nicht über die EUROMAP77 verfügbar waren. Dieses Wissen kann einerseits genutzt werden, um in Zukunft KI-Modelle basierend auf den existierenden Standards mit möglichst viel Aussagekraft zu entwickeln und andererseits, um eine mehrwertbringende Erweiterung der Standards zu erwirken.

Das Projekt hat gezeigt, dass die Entwicklung eines robusten Prozessmodells, das teilweise über die trainierten Prozessfenstergrenzen hinaus sinnvolle Vorhersagen und Handlungsempfehlungen gibt, möglich ist. Der im Projekt gewählte Ansatz des Human-in-the-loop wurde aus sicherheitstechnischen Gründen bevorzugt, jedoch zeigt sich das Potential, dass KI-Modelle Maschinen auch direkt regeln können. Dafür bedarf es weiterer Robustheitsuntersuchungen, erweiterter Sicherheitsmechanismen und einer Betrachtung von vielen, auch juristischen, Randbedingungen, die Gegenstand zukünftiger, interdisziplinärer Forschung sein werden. Nichtsdestotrotz bietet die erarbeitete Methodik zur Entwicklung Bayes'scher Netze eine solide Grundlage, um auch für andere industrielle Prozesse hochwertige, robuste Modelle zur Überwachung, Regelung und Fehleranalyse zu entwickeln.

## 7 Literatur

- [1] „Injection Molding Machine Market Size, Share, & Trends Analysis Report.” <https://www.grandviewresearch.com/industry-analysis/injection-molding-machine-market#>
- [2] W. M. Friedrich Johannaber, *Handbuch Spritzgießen*. München, 2001.
- [3] Meaghan Charest, Ryan Finn und Rickey Dubay. „Integration of artificial intelligence in an injection molding process for on-line process parameter adjustment.”
- [4] Dae Seong Baek, Chengjun Li und Jungsoo Nam et al. „A Study on Condition Monitoring and Diagnosis of Injection Molding Process Using Probabilistic Neural Network Method: Conference: ASME 2014 International Manufacturing Science and Engineering Conference collocated with the JSME 2014 International Conference on Materials and Processing and the 42nd North American Manufacturing Research Conference.”
- [5] Robert X. Gao, Xinyao Tang, Guthrie Gordon und David Kazmer. „Online product quality monitoring through in-process measurement.”
- [6] Xundao Zhou, Yun Zhang, Ting Mao und Huamin Zhou. „Monitoring and Dynamic Control of Quality Stability for Injection Molding Process.”
- [7] Yun Zhang und Z. H. Ting Mao. „A statistical quality monitoring method for plastic injection molding using machine built-in sensors.”
- [8] KraussMaffei. „PC Plus – Stabilität, Präzision und Kosteneffizienz in der Fertigung.” [https://km.kraussmaffei.com/media/files/imm\\_broschueren/IMM\\_FL\\_APCPlus\\_de.pdf](https://km.kraussmaffei.com/media/files/imm_broschueren/IMM_FL_APCPlus_de.pdf) (Zugriff am: 22. April 2020).
- [9] plus10 GmbH. „HOPPER - der Online-Qualitätsoptimierer.” <https://www.plus10.de/hopper>
- [10] „EUROMAP - Overview OPC UA Specifications.” <https://www.euromap.org/i40/overview> (Zugriff am: 5. Oktober 2023).
- [11] OPC Foundation. „Home Page - OPC Foundation.” <https://opcfoundation.org/> (Zugriff am: 5. Oktober 2023).
- [12] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, 1. Aufl. s.l.: Elsevier Reference Monographs, 2014. [Online]. Verfügbar unter: <http://gbv.ebib.com/patron/FullRecord.aspx?p=1876690>
- [13] A. Darwiche, *Modeling and reasoning with Bayesian networks*. Cambridge: Cambridge University Press, 2009.
- [14] U. B. Kjærulff, *Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis*, 2. Aufl. (Information Science and Statistics Ser v.22). Dordrecht: Springer, 2011.
- [15] J. Lambers, J. Schüder und J. Krauß, "Injection-Molding Production Data with Quality Labels (Version 1)", 2022, doi: 10.23728/b2share.03133fb279294389a15baefd55e4257a. [Online]. Verfügbar unter: <https://b2share.eudat.eu/records/03133fb279294389a15baefd55e4257a>

- [16] J. Lambers, J. Schüder und J. Krauß, "Injection-Molding Production Data with Quality Labels (Version 2)", 2022, doi: 10.23728/B2SHARE.4C5692B886DB419180F716ACF895BF06. [Online]. Verfügbar unter: <https://b2share.eudat.eu/records/4c5692b886db419180f716acf895bf06>
- [17] J. Lambers, J. Schüder, J. Krauß und M. Schömig, "Injection-Molding Production Data with Quality Labels (d-optimal Design of Experiment pattern)", 2023, doi: 10.23728/b2share.3f80952ce5ff4be88ae4cf6a3bdfe732. [Online]. Verfügbar unter: <https://b2share.eudat.eu/records/3f80952ce5ff4be88ae4cf6a3bdfe732>
- [18] J. Lambers und J. Krauß. „ProBayes Repository.” [https://dev.azure.com/SKZ-public/ProBayes\\_public](https://dev.azure.com/SKZ-public/ProBayes_public) (Zugriff am: 13. Oktober 2023).
- [19] M. Christ, N. Braun, J. Neuffer und A. W. Kempa-Liehr. „tsfresh: feature extraction of time series data.” <https://tsfresh.readthedocs.io/en/latest/index.html> (Zugriff am: 9. Oktober 2023).
- [20] C. W. J. Granger, "Investigating Causal Relations by Econometric Models and Cross-spectral Methods," *Econometrica*, Jg. 37, Nr. 3, S. 424, 1969, doi: 10.2307/1912791.
- [21] A. Sharma und E. Kiciman, "DoWhy: An End-to-End Library for Causal Inference", Nov. 2020. [Online]. Verfügbar unter: <https://arxiv.org/pdf/2011.04216>
- [22] J. Pearl, *Causality: Models, reasoning and inference*, 2021. Aufl. Cambridge: Cambridge University Press, 2009.
- [23] „Streamlit documentation.” <https://docs.streamlit.io/> (Zugriff am: 9. Oktober 2023).
- [24] J. Krauß und I. Borchardt, "Prediction and control of injection molded part weight using machine learning – A literature review," in *16th CIRP Conference on Intelligent Computation in Manufacturing Engineering - CIRP ICME '22 Virtual Conference: Procedia CIRP*, Bd. 118, Elsevier, 2022, S. 867–872.
- [25] I. Borchardt, J. Krauß, J. Lambers und J. Schüder, "Approach for inline monitoring and optimization of a thermoplastic injection molding process with Bayesian networks taking the example of the quality feature weight," in *16th CIRP Conference on Intelligent*, Bd. 118, S. 851–856.
- [26] J. Lambers, J. Schüder und J. Krauß, "Poster - ProBayes,".
- [27] J. Lambers, "Qualitätsmodell zur Optimierung und Online Überwachung des Spritzgussprozesses mit Hilfe Bayes'scher Netze,".
- [28] J. Schüder und J. Lambers, "IIoT Spritzgießzelle – Versuchsplattform für Industrie und Forschung," *Plastverarbeiter*, 22-10, 2022. [Online]. Verfügbar unter: <https://www.plastverarbeiter.de/automation/versuchsplattform-fuer-industrie-und-forschung-204.html>
- [29] J. Lambers und J. Schüder, "Expertenwissen & Maschinelles Lernen im Spritzguss: SKZ und Fraunhofer IPA erforschen die Überwachung und Steuerung des Spritzgussprozesses," , Nr. 22.1, S. 10, 2022. [Online]. Verfügbar unter: [https://www.skz.de/fileadmin/user\\_upload/News/SKZ\\_aktuell/PDF/SKZ-Aktuell-22-1-web.pdf#page=10](https://www.skz.de/fileadmin/user_upload/News/SKZ_aktuell/PDF/SKZ-Aktuell-22-1-web.pdf#page=10)

- [30] The Erium Podcast - Data Science & Machine Learning. „The Erium Podcast.” <https://theeriumpodcast.de/> (Zugriff am: 20. Januar 2023).
- [31] „ProBayes: Qualitätsmodell zur Optimierung und Online-Überwachung von Spritzgussprozessen mit Hilfe Bayes'scher Netze.” <https://www.skz.de/forschung/projekt/probayes> (Zugriff am: 15. November 2022).
- [32] Fraunhofer-Institut für Produktionstechnik und Automatisierung IPA. „Digitale Qualitätsmodelle zur Online-Überwachung und Optimierung von Spritzgussprozessen – ProBayes - Fraunhofer IPA.” <https://www.ipa.fraunhofer.de/de/referenzprojekte/ProBayes.html> (Zugriff am: 15. November 2022).
- [33] „Posten | LinkedIn.” [https://www.linkedin.com/posts/skz-%E2%80%93-das-kunststoff-zentrum\\_skz-opendata-digitalstrategie-activity-6939899899190636545-4K4q/?utm\\_source=share&utm\\_medium=member\\_desktop](https://www.linkedin.com/posts/skz-%E2%80%93-das-kunststoff-zentrum_skz-opendata-digitalstrategie-activity-6939899899190636545-4K4q/?utm_source=share&utm_medium=member_desktop) (Zugriff am: 15. November 2022).
- [34] „LinkedIn-Post Spritzgießzelle.” [https://www.linkedin.com/posts/skz-%E2%80%93-das-kunststoff-zentrum\\_digitale-spritzgie%C3%9Fzelle-des-skz-activity-7000449219576614912-1NJm?utm\\_source=share&utm\\_medium=member\\_desktop](https://www.linkedin.com/posts/skz-%E2%80%93-das-kunststoff-zentrum_digitale-spritzgie%C3%9Fzelle-des-skz-activity-7000449219576614912-1NJm?utm_source=share&utm_medium=member_desktop) (Zugriff am: 21. November 2022).

## 8 Abbildungsverzeichnis

Abbildung 1: Die vier Phasen bei der Entwicklung Bayes'scher Modelle (nach [14]).	11
Abbildung 2: Lösungsweg von Datenerzeugung über Quantifizierung bis hin zum Demonstrator, wobei dieser Weg iterativ durchlaufen wurde.....	13
Abbildung 3: Genutzte Spritzgießmaschine im Technikum des SKZ .....	15
Abbildung 4: Spritzgieß-Werkzeug „Verzugsschale“ (links) und das entsprechende Bauteil (rechts) .....	16
Abbildung 5: CAD-Modell des konstruierten Greifers zur automatisierten Bauteilentnahme.....	17
Abbildung 6: Bildverarbeitungssystem des Unternehmens Ziemann & Urban GmbH..	18
Abbildung 7: Integrierte Industriewaage des Unternehmens Mettler-Toledo zur Aufzeichnung des Gewichtwertes .....	19
Abbildung 8: Im Projekt entwickelter Umweltsensor zur Aufzeichnung von Temperatur und Feuchtigkeit.....	19
Abbildung 9: Bestandteile des Trocknungssystems und deren Vernetzung .....	23
Abbildung 10: Kommunikation des Bildverarbeitungssystems mit <i>Connectware</i> und Spritzgießmaschine .....	24
Abbildung 11: Exemplarischer Ausschnitt aus der erarbeiteten Einflussmatrix.....	25
Abbildung 12: Teilnetz für den Qualitätsindikator „Bauteilgewicht“ .....	26
Abbildung 13: Teilnetz für den Qualitätsindikator „Verzug“ .....	26
Abbildung 14: Gesamtnetz für die Qualitätskriterien „Bauteilgewicht“ und „Verzug“ ..	27
Abbildung 15: Weitere erzeugte optische Fehlerbilder am Bauteil „Verzugsschale“ ....	29
Abbildung 16: Festlegung des Netzabstands für das Formteil <i>Verzugsschale</i> nach der <i>nodes with enough space</i> Methode.....	30
Abbildung 17: Darstellung des vernetzten Formteils (Netzabstand = 2 mm).....	30
Abbildung 18: Übersicht der attributspezifischen Vernetzungsarten im Simulationsmodell.....	31
Abbildung 19: Verteilung der Messpunkte für die thermografische Auswertung (ca. 28.000 Messpunkte) links und thermografische Aufnahme einer exemplarischen Stichprobe aus dem realen Prozess (rechts) .....	32
Abbildung 20: Bewertungsmatrix der Materialdatenqualität für die rheologischen Untersuchungen für PP-HE125MO (links) und ABS-TerluranGP-22 (rechts) .....	34
Abbildung 21: Beispielhafte Darstellung der Versuchspunkte bei einem D-optimalen Versuchsplan .....	35
Abbildung 22: Schematische Darstellung der Aufbereitung und Zusammenführung unterschiedlicher Datentypen .....	37
Abbildung 23: Ausschnitt aus dem finalen Datensatz .....	38
Abbildung 24: Schematische Darstellung der Erweiterung des Datensatzes um extrahierte Features aus den Thermokamera-Daten .....	39

Abbildung 25: Standardabweichung der Temperaturwerte über alle Versuche und Aufnahmen mit PP hinweg (links) und für die Versuchsreihe A06 mit PP über alle Aufnahmen hinweg (rechts).....	40
Abbildung 26: Thermogramm mit festgelegten „Regions of Interest“ (ROIs) .....	41
Abbildung 27: Pearson-Korrelationskoeffizienten zwischen ausgewählten Fehlerbildern und extrahierten Features aus den Thermogrammen .....	42
Abbildung 28: Exemplarische Extraktion von Features aus der Einspritzdruckkurve ...	43
Abbildung 29: Phasen des Spritzgusszyklus basierend auf den Triggerwerten des DataXplorers .....	43
Abbildung 30: Phasenspezifische Extraktion von Features aus der Einspritzdruckkurve .....	44
Abbildung 31: Verzugsschale mit den zur Berechnung der Verzugskennzahl verwendeten Maßen .....	45
Abbildung 32: Ermittelte „Permutation Feature Importances“ für alle Prozessparameter in Bezug auf das Bauteilgewicht .....	47
Abbildung 33: Streudiagramme für das Bauteilgewicht und die fünf besten Prozessparameter .....	47
Abbildung 34: Schematische Darstellung des kausalen Effekts (angelehnt an [21]) .....	48
Abbildung 35: Vorhersage des Formteilgewichts durch das Bayes'sche Netz in Form eines Wahrscheinlichkeitsdichteplots .....	52
Abbildung 36: Konfusionsmatrix für die Gewichtsvorhersage des Bayes'schen Netzes .....	52
Abbildung 37: Validierung der Parameterempfehlungen .....	53
Abbildung 38: Schematische Darstellung der Architektur des Live-Demonstrators .....	55
Abbildung 39: Graphisches User Interface des Demonstrators, umgesetzt mit <i>Streamlit</i> [23].....	56
Abbildung 40: Prognosebereich des Dashboards .....	57
Abbildung 41: Diagnosebereich des Dashboards .....	57
Abbildung 42: Schematische Darstellung aller vernetzten Geräte und der Spritzgießmaschine .....	60
Abbildung 43: Auszug aus den Vorversuchen Versuchsplan für PP-HE125MO des Vergleichs von gemessenen und simulierten Formteilgewichten vor dem Bug-Fix (rotes X = Standardsimulation, Boxplot = 12 real gespritzte Teile) .....	61
Abbildung 44: Auszug aus dem D-Optimalen Versuchsplan für PP-HE125MO des Vergleichs von gemessenen und simulierten Formteilgewichten nach dem Bug-Fix (grünes Sechseck = Standardsimulation, Boxplot = 12 real gespritzte Teile) .....	61
Abbildung 45: Vergleich der generischen (orange) und vermessenen (blau) Daten zur Kristallinität des Materials PP-HE125MO .....	62
Abbildung 46: Auszug aus dem D-Optimalen Versuchsplan für PP-HE125MO des Vergleichs von gemessenen und simulierten Formteilgewichten (grünes Sechseck =	

---

Standardsimulation, blauer Kreis = Simulation mit modifizierten Materialdaten, Boxplot = 12 real gespritzte Teile).....	63
Abbildung 47: Auszug aus dem D-Optimalen Versuchsplan für PP-HE125MO des Vergleichs von gemessenen und simulierten Formteilgewichten (grünes Sechseck = Standardsimulation, blauer Kreis = Simulation mit modifizierten Materialdaten, rotes Dreieck = Simulation mit modifizierten Material- und Maschinendaten, Boxplot = 12 real gespritzte Teile).....	64
Abbildung 48: Finales Teilnetz für den Qualitätsindikator „Bauteilgewicht“ .....	65
Abbildung 49: Finales Teilnetz für den Qualitätsindikator „Verzug“ .....	65
Abbildung 50: Erarbeitete Methodik für die Entwicklung Bayes'scher Netze.....	67

## 9 Anhang

Tabelle 13: Quantifizierung des kausalen Effekts für alle Abhängigkeiten des Bayes'schen Netzes "Bauteilgewicht"

Modell	Causal estimator		Refuter								
	Causal estimate	p_value	Refuter Placebo			Refuter common cause			Refuter data subset		
old causal estimate			new causal estimate	p_value	old causal estimate	new causal estimate	p_value	old causal estimate	new causal estimate	p_value	
<b>Gesamtes Modell</b>	0.0642	<10 <sup>-10</sup>	0.0642	0.0	2.0	0.0642	0.0642	0.9	0.0642	0.0642	0.98
<b>Cushion Modell</b>	-0.0941	<10 <sup>-10</sup>	-0.0941	0.0	2.0	-0.0941	-0.0941	0.98	-0.0941	-0.0941	0.86
<b>Dosing Modell</b>	0.0534	<10 <sup>-10</sup>	0.0534	0.0	2.0	0.0534	0.0534	0.90	0.0534	0.0533	0.94
<b>Area_Cavity Modell</b>	0.0905	<10 <sup>-10</sup>	0.0905	0.0	2.0	0.0905	0.0905	0.88	0.0905	0.0905	0.9
<b>Cavity_max Modell</b>	0.0962	<10 <sup>-10</sup>	0.0962	0.0	2.0	0.0962	0.0962	0.92	0.0962	0.0962	0.96