



Patrick Rene Kröpfl, BSc.

**Beitrag zur Beschreibung des Bewegungsverhaltens
von Kleinsendungen und Polybags mit Hilfe von
Regressionsmodellen**

MASTERARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieur (Dipl. Ing.)

Masterstudium Wirtschaftsingenieurwesen-Maschinenbau

eingereicht an der

Technischen Universität Graz

Betreuer

Assoc.Prof. Dipl.-Ing. Dr.techn. Landschützer Christian

Dipl.-Ing. BSc Schedler Michael

Graz, April 2022

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Das in TUGRAZonline hochgeladene Textdokument ist mit der vorliegenden Masterarbeit identisch.

Graz,.....

Datum

.....

Unterschrift

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master's thesis.

Graz,.....

Date

.....

Signature

KURZFASSUNG

In der vorliegenden Arbeit wird die Entwicklung einer Methodik für die Ermittlung der Durchlaufzeit von Stückgütern in Sortieranlagen auf Basis von Regressionsmodellen behandelt. Die Aufgabenstellung leitet sich aus der Durchsatzberechnung des Free-Fall-Sorters (FFS) ab. Der FFS ist ein neuartiges Sortierkonzept, bei dem Sendungen schwerkraftgetrieben die Sortieranlage durchlaufen. Da das Stückgut nicht wie bei herkömmlichen Sortieranlagen durch ein Fördermittel angetrieben wird, müssen für eine analytische Berechnung der Durchlaufzeit (wodurch auf den Durchsatz geschlossen werden kann) eine Vielzahl an Vereinfachungen angenommen werden, durch welche es zu großen Fehlern in der Berechnung kommt.

Abgeleitet aus den Problemen bei der analytischen Berechnung hat sich die Fragestellung für eine alternative Methode zur Bestimmung der Durchlaufzeit ergeben. Dabei ist die Wahl auf einen statistischen Ansatz gefallen, in dem Regressionsmodelle als Werkzeug verwendet werden. Für die Erzeugung des Regressionsmodells ist im Zuge dieser Arbeit eine Methodik entwickelt worden, welche in neun Schritten den Ablauf von der Vorbereitung bis zur Validierung beinhaltet. Die ersten Schritte befassen sich dabei mit der Analyse der Sortieranlage und des dazugehörigen Stückguts. Mit Hilfe von Versuchsplänen werden die Einflüsse der Eigenschaften des Stückgutes sowie der Sortieranlage auf die Durchlaufzeit untersucht.

Anschließend folgt die Erzeugung der Regressionsmodelle. Dabei wird eine Logik implementiert, welche es erlaubt gesamte Anlagenlayouts aus einzelnen Submodellen aufzubauen. Diese ermöglicht es beliebige Anlagenlayouts modular nachzubilden.

Am Ende der Methodik wird eine Validierung und Optimierung des vorliegenden Predictionmodells durch Realversuche durchgeführt.

Neben der allgemeinen Darstellung der Methodik werden in dieser Arbeit die einzelnen Schritte anhand des Free-Fall-Sorters angewandt und in weiterer Folge validiert. Dadurch lässt sich eine allgemeine Gültigkeit der Methodik für Sortieranlagen von Stückgütern nachweisen. Als abschließender Punkt werden noch Erweiterungsmöglichkeiten und Denkanstöße für weitere Anwendungen angeführt.

ABSTRACT

This thesis describes the development of a method to determine the lead time of general cargo in a sorting system by using regression models. The problem is defined by the need of a throughput calculation of the Free-Fall-Sorter (FFS). This sorter is a new approach, where the shipments go through the sorting system only by influence of gravitation. Because of that reason an analytical calculation of the lead time (which leads to the throughput) needs a lot of simplifications which is the reason for a big deviation.

Derived from the problems with the analytical calculation, the question for an alternative method for the determination of the lead time has arisen. Which leads to a statistical approach, where regression models are used. This thesis describes a method (with nine steps) which can be used as a guide from the system definition up to the validation.

The first steps are including an overview about the considered system and the general cargo. After that, DoE (design of experiment) is used to evaluate the influence on the lead time by the property from the general cargo and the sorting system.

The next steps are including the creation of the regression model which makes it possible to simulate whole system layouts by merging submodels.

At last, there is a validation and optimization of the prediction model by the use of real experiments.

Besides the general description, the method is also applied to the Free-Fall-Sorter. With the use of the validation data of the FFS it becomes possible that the method could be used for general sorting systems and general cargo. In the last chapters of the thesis will also be a list with extension possibilities and examples for other applications.

Inhaltsverzeichnis

1	EINLEITUNG	1
1.1	Ausgangslage	1
1.2	Zielsetzung und Motivation	2
1.3	Stand der Technik	2
2	GRUNDLAGEN	5
2.1	FreeFallSorter	5
2.2	Statistische Versuchsplanung	7
2.2.1	Aufbau von Versuchsplänen	7
2.2.2	Faktorielle Versuchspläne	9
2.2.3	Effektbestimmung	11
2.2.4	Korrelationsanalyse	12
2.3	Regressionsmodelle	13
2.3.1	Stichprobengröße	15
2.3.2	t-Verteilung	16
2.3.3	Kreuzvalidierung	18
3	AUFBAU DER METHODIK	20
3.1	Schritt 1 – System definieren	22
3.2	Schritt 2 – Parameterscreening-Versuchsplan erzeugen	22
3.2.1	Versuchsplan erstellen	24
3.2.2	Anlage vorbereiten	24
3.2.3	Proben vorbereiten	24
3.2.4	Messgrößenerfassung standardisieren	24
3.2.5	Versuchsvorschrift erstellen	25
3.2.6	Stichprobenumfang ermitteln	25
3.3	Schritt 3 – Versuche für das Parameterscreening durchführen	26
3.4	Schritt 4 – Einflüsse der Parameter analysieren	26
3.5	Schritt 5 – Teachingset-Versuchsplan erzeugen	26
3.6	Schritt 6 – Versuche für das Teachingset durchführen	27
3.7	Schritt 7 – Regressionsmodelle erzeugen	28
3.8	Schritt 8 – Regressionsmodelle validieren	29

3.9	Schritt 9 – Userinterface erstellen	30
4	EXEMPLARISCHE ANWENDUNG DER METHODIK.....	31
4.1	Anwendung Schritt 1 - System definieren.....	31
4.1.1	Stückgutparameter	34
4.1.2	Anlagenparameter.....	34
4.2	Anwendung Schritt 2 - Parameterscreening-Versuchsplan erzeugen	35
4.2.1	Versuchsplan erstellen	35
4.2.2	Anlage vorbereiten	37
4.2.3	Proben vorbereiten.....	39
4.2.4	Messgrößenerfassung standardisieren	43
4.2.5	Versuchsvorschrift erstellen	45
4.2.6	Stichprobenumfang ermitteln	46
4.3	Anwendung Schritt 3 – Versuche für das Parameterscreening durchführen	48
4.4	Anwendung Schritt 4 - Einflüsse der Parameter analysieren	49
4.5	Anwendung Schritt 5 Teachingset -Versuchsplan erzeugen	53
4.6	Anwendung Schritt 6 – Versuche für das Teachingset durchführen	54
4.7	Anwendung Schritt 7 – Regressionsmodelle erzeugen.....	54
4.7.1	Regressionsmodelle mit bekannten Anfangsbedingungen	54
4.7.2	Regressionsmodelle mit unbekanntem Anfangsbedingungen	56
4.7.3	Zusammensetzung der Regressionsmodelle	59
4.8	Anwendung Schritt 8 – Regressionsmodelle validieren	59
4.8.1	Validierung der Rutsche oben (RO)	60
4.8.2	Validierung der Umlenkung	60
4.8.3	Validierung der Rutsche unten (RU)	61
4.8.4	Validierung der Gesamtzeit	62
4.8.5	Anpassung an die optimale Stützstellenanzahl.....	62
4.9	Anwendung Schritt 9 – Userinterface erstellen	63
5	ERGEBNISSE UND INTERPRETATION.....	65
5.1	Zusammenfassung der Ergebnisse der exemplarischen Anwendung.....	65
5.2	Beurteilung der Gültigkeit der Methodik	66
5.2.1	Möglichkeiten für die Beurteilung der Gültigkeit der Methodik	66
5.2.2	Gültigkeit der Submodelle.....	66
5.2.3	Gültigkeit der Zusammensetzung der Submodelle zu einem Predictionmodell.....	67

5.2.4	Einschränkungen der Methodik.....	68
5.2.5	Verallgemeinerung der Methodik für weitere Anwendungsfälle.....	68
5.3	Ausblick.....	69
5.3.1	Prozessoptimierung und Anlagensteuerung.....	69
5.3.2	Anlagenplanung und Optimierung.....	69
5.3.3	Verbesserung durch Machine Learning.....	69
5.4	Fazit.....	70
6	VERZEICHNISSE.....	71
6.1	Literaturverzeichnis.....	71
6.2	Abbildungsverzeichnis.....	73
6.3	Tabellenverzeichnis.....	75
	ANHANG.....	76
	Anhang A1 – Quantile der Kolmogorov-Smirnov-Teststatistik.....	76
	Anhang A2 – Kolmogorov-Smirnov-Test der Submodelle.....	77
	Anhang B1 – Parameterscreening.....	80
	Anhang B2 – Korrelationsanalyse.....	82
	Anhang C – Predictionmodell erzeugen.....	85

1 EINLEITUNG

In diesem Kapitel folgt eine kurze Beschreibung der Ausgangslage, der Zielsetzung und Motivation der Arbeit sowie ein Einblick in den aktuellen Stand der Technik.

1.1 Ausgangslage

Durch die starke Globalisierung in den letzten Jahren, dem veränderten Einkaufsverhalten von Kunden und neuen Businessmodellen im Bereich E-Commerce, erreichte die Kurier-Express-Paket (KEP) Branche einen sehr starken Zuwachs. Dieses Wachstum wird sich laut Prognosen auch in den nächsten Jahren noch fortsetzen. Als Folge daraus kam es auch zu einer ständigen Weiterentwicklung in Bezug auf die Verpackungen von Sendungen. Klassische Verpackungen aus Papier oder Kartonage werden häufig durch Kunststoffbeutel, welche auch Polybags genannt werden ersetzt [EK21]. Jene Polybags weisen viele physikalische Eigenschaften auf, für welche herkömmliche Sortieranlagen nicht ausgelegt sind. Hierbei kommt es häufig zu Problemen, wie „Verlieren“ der Sendung innerhalb der Verarbeitung, Problemen bei der maschinellen Lesbarkeit des Labels und bei der Kontrolle/Regelung der Dynamik der Polybags [SSK+21].

Aus diesem Grund wurde am Institut für Technische Logistik (ITL) ein neuartiges Konzept für das Sortieren von Kleinsendungen entwickelt. Daraus entstand das Konzept des Free-Fall-Sorters (FFS), dessen Aufbau auf einem schwerkraftgetriebenen Sortierprozess beruht. Dabei gleiten Sendungen ohne externen Antrieb durch die Sortieranlage. Der Aufbau ähnelt einem entscheidungsbaumartigen Layout, dessen Äste durch Rutschen ersetzt werden. Für die Umlenkungen zwischen den Rutschen werden automatisch verstellbare Klappen verbaut, welche die Sendung in die geforderte Richtung für die jeweilige Endstelle leiten. Eine detailliertere Beschreibung des FFS folgt in Kapitel 2.1.

Neben dem Ziel der hohen Energieeffizienz, bei gleichzeitiger Leistungssteigerung wird ein Hauptaugenmerk auf die physikalischen Eigenschaften der Sendung gelegt. Da durch die hohe Variabilität der Sendung das Systemverhalten stark beeinflusst werden kann. Die Betrachtung dieser Einflüsse gilt als ein bis dato kaum erforschter Bereich und wird häufig nicht berücksichtigt oder nur stark vereinfacht abgebildet [SL21].

Um dieses vielversprechende Konzept mit herkömmlichen Sortieranlagen vergleichen zu können, ist neben ökonomischen und ökologischen Faktoren der Durchsatz der Anlage eine maßgebliche Kenngröße. Es liegen bereits Ansätze für eine Durchsatzberechnung des FFS vor, welche sich jedoch als analytisch nicht generell gültig erweisen und nur eine Näherungslösung liefern.

1.2 Zielsetzung und Motivation

Abgeleitet aus der Durchsatzberechnung des Free-Fall-Sorters, ergibt sich als Zielsetzung die Entwicklung einer Methodik zur Erzeugung eines Predictionmodells für die Ermittlung der Durchlaufzeit von Stückgütern in Sortieranlagen.

Als Eingabeparameter für das Predictionmodell sollen physikalische Eigenschaften des Stückgutes und Anlagenparameter der Sortieranlage dienen. Um unterschiedliche Layouts nachzubilden, muss das Predictionmodell modular¹ erweiterbar sein und für einen breiten Parameterbereich verwendbar sein.

Wie in nahezu allen Bereichen liegt auch hier ein Hauptaugenmerk darauf, den Aufwand für die Erzeugung des Predictionmodells möglichst gering zu halten. Um das Modell nur mit den ausschlaggebenden Informationen bezüglich des Stückgutes und der Anlage zu versorgen, sollte im Vorfeld ein Auswahlverfahren für die Ermittlung jener Eigenschaften definiert werden.

Da sich in der Versuchshalle des Technischen Logistik Institutes ein Prototyp des FFS befindet, wird die entwickelte Methodik auf diesen angewendet. Als Stückgut werden Polybags verwendet, welche durch ihre Variabilität das breite Spektrum der Stückgüter sehr gut abbilden. Um die Gültigkeit der Methodik und des daraus abgeleiteten Predictionmodells zu überprüfen, werden die vom Modell berechneten Durchlaufzeiten mit Realdaten aus Versuchen am FFS validiert.

1.3 Stand der Technik

Aktuelle Durchsatzermittlungen von Sortieranlagen beziehen sich auf die Fördergeschwindigkeit des Verteilförderers, dem minimalen Abstand der Sortiergüter und der Anzahl der parallelen Tragmittel. Der hierbei ermittelte Durchsatz ist jedoch lediglich ein technischer Grenzdurchsatz. Es wird dabei nicht berücksichtigt, ob Güter zur Einschleusung vorliegen oder ob freie oder passende Tragmittel zur Verfügung stehen. Die physikalischen Eigenschaften des Stückgutes werden hierbei nur bedingt berücksichtigt, indem zwischen verschiedenen Gutklassen differenziert wird [JH12].

Hinter den Kombinationen der Stückguteigenschaften können sich jedoch Einflussfaktoren für den Durchsatz verbergen, welche bei einer verallgemeinerten Betrachtung nicht berücksichtigt werden. Wodurch die Ergebnisse der Durchsatzberechnung stark verfälscht werden können. Wie stark der Einfluss der jeweiligen Faktoren ist, hängt jedoch wiederum stark von dem verwendeten Fördermittel ab.

¹ Die Sortieranlage wird in einzelne Submodelle aufgeteilt, welche beliebig zu anderen Layouts zusammengefügt werden können.

Für die Untersuchung des Reibungsverhaltens bei gleitenden Sendungen beschäftigte sich bereits Böcker in seiner Dissertation „Auslegung von Rutschen an Sorterendstellen“. Dabei untersuchte er den Einfluss der Gutgeschwindigkeit und des Gutgewichtes auf den Reibungsbeiwert. Dabei ergab sich aus Versuchen eine meist quadratische Abhängigkeit des Reibwertes von der Geschwindigkeit. Wobei dieser Zusammenhang nicht nur auf den Materialeigenschaften der Reibpartner beruht, sondern auch auf Strömungseinflüssen. Diese wiederum lassen sich in den Strömungswiderstand des Körpers und in den Auf- und Abtrieb unterteilen. Mit diesen Erkenntnissen, wurde eine näherungsweise Berechnung des Gleitreibungskoeffizienten erstellt [Tho96].

Dafür wurden folgende Annahmen getroffen:

- Masse des Körpers bleibt konstant
- Kontur der Gleitfläche wird vorgegeben
- Definierte Ein- und Auslaufkanten
- Möglichkeit, dass Luft zwischen Körper und Gleitfläche in Gleitrichtung strömt besteht
- Berührungsflächen der Reibpartner sind plan
- Form des Körpers ist quadratisch
- Ein- und Auslaufkante biegt sich durch die Anströmung auf, wodurch es zu einem Düseneffekt kommt

Die, aus diesen Annahmen entstandene Analytik, wurde durch einen entsprechenden Versuchsaufbau untersucht und erwies sich als teilweise gültig. Die dabei untersuchten Parameter sind jedoch nur bedingt auf den FFS anzuwenden. Da bei den Untersuchungen die Untergrenze der Masse der Körper mit 100g festgelegt wurde und die untersuchten Reibpaarungen nicht den gesamten Anwendungsbereich² des FFS abdecken. Weiteres ist auch die Annahme von quadratischen Körpern und einer planen Kontaktfläche eine starke Vereinfachung für den Anwendungsfall Polybag als Stückgut. Die Ergebnisse dieser Arbeit liefern somit keinen Mehrwert für den Aufbau der Methodik.

² Bereiche der betrachteten Masse entspricht 20-220g Gesamtgewicht. Bereich des Gleitreibungskoeffizienten der Reibpaarung entspricht 0,16-0,31 [SSK+21].

Auch für den FFS liegt bereits eine Analytik für die Durchsatzberechnung³ vor. Abgeleitet aus dem Drallsatz ergibt sich hierbei eine Durchlaufzeit, welche von der Anfangsgeschwindigkeit, Rutschenlänge, Rutschenwinkel, Reibwert und der Erdbeschleunigung abhängig ist. Dieser Berechnung liegen folgende Annahmen zugrunde:

- Anfangsgeschwindigkeit bei der Einschleusung wird gemessen
- Beschleunigung der Sendung ist konstant
- Strömungswiderstände haben keinen Einfluss
- Sendung gleitet, ohne zu rollen und/oder ohne sich zu verdrehen
- keine Anfangsgeschwindigkeit nach der Umlenkung
- Reibung nach Coulomb
- Gestalt der Sendung unverändert
- Auftrieb nicht berücksichtigt

Aus physikalischer Sicht ist es nicht möglich all diese getroffenen Annahmen analytisch nachzuweisen. Viele davon beruhen auf Systemkenntnissen und Beobachtungen, aus bereits zuvor erfolgten Versuchen. Die vorhandene Analytik liefert somit nur eine näherungsweise Durchsatzberechnung, welche Spielraum für Erweiterungen zulässt.

Abgeleitet aus den vorliegenden Erkenntnissen zeigt sich eine analytische Berechnung der Durchlaufzeit nicht als zielführend. Um den gesamten Parameterbereich der Sendung und der Anlage beherrschen zu können, muss somit eine alternative Beschreibungsmöglichkeit gefunden werden.

³ Grundlage für die Analytik, wird in einer zu diesem Zeitpunkt noch nicht veröffentlichten Dissertation beschrieben. Die Ergebnisse der Analytik sind für die vorliegende Problemstellung nicht relevant. Sie kann jedoch als Orientierungshilfe gesehen werden, da in ihr die Problemstellen der Berechnung erkennbar sind.

2 GRUNDLAGEN

Um dem Inhalt der Arbeit besser folgen zu können, werden in diesem Kapitel alle wichtigen theoretischen Grundlagen zusammengefasst, auf welche im Zuge der Arbeit zurückgegriffen wurde. Des Weiteren wird auch das Prinzip des FFS etwas näher erläutert, da die Methodik daran angewendet und validiert wird.

2.1 FreeFallSorter

Der FFS ist die Entwicklung eines neuartigen Sortiersystems für den Sendungstyp Polybag und Kleinsendungen. Den Anstoß für die Konzipierung des Sorters gaben die häufig auftretenden Probleme, bei der Verarbeitung von Sendungen, deren Verpackung von herkömmlicher Kartonage abweicht. Solche Verpackungen werden unter dem Begriff Polybag zusammengefasst. Ein besonderes Merkmal dieser Verpackungen ist die hohe Variabilität der physikalischen Eigenschaften. State-Of-The-Art Sortieranlagen stoßen bei der Verarbeitung von Polybags auf Probleme, wie „Verlieren“ der Sendung innerhalb der Verarbeitung, Fehler bei der maschinellen Lesbarkeit der Label sowie bei der Kontrolle/Regelung ihrer Dynamik.

Das Grundprinzip wie in Abbildung 2.1 dargestellt, beruht auf einem Layout, welches auf Basis eines Entscheidungsbaumes aufgebaut ist. Die Äste (Pfad) des Entscheidungsbaumes werden in der Anlage durch Rutschen ersetzt. Entlang dieser gleiten die Sendungen ohne externen Antrieb, nur durch die Schwerkraft getrieben, bis zur nächsten Knotenstelle. Für Layouts, mit sehr flachen Neigungswinkeln (Pfadneigung), können die Rutschen auch durch Fördermittel ersetzt werden. In den Knotenstellen befindet sich eine Umlenkklappe, welcher den weiteren Weg der Sendung durch den Sorter vorgibt. Um die Klappe steuern zu können, werden durch einen Code-Leser bei der Einschleusung (Infeed) Daten übermittelt, wodurch alle Stellungen der Umlenkklappe, bis zur Endstelle vorgegeben werden[SL21]. Die Zuförderung für die Einschleusung kann beliebig an die Anforderungen angepasst werden und unterliegt nur der Bedingung, dass die Geschwindigkeit mit welcher der Paketstrom zugeführt wird, nicht den Durchsatz der Anlage limitiert. Ebenso liegt für die Ausführung der Endstellen eine hohe Flexibilität⁴ vor.

⁴ Die Ausführung der Endstellen kann durch eine Vielzahl an Möglichkeiten realisiert werden. Beispielhaft dafür wäre ein Automatischer Abtransport durch Förderbänder, AGV's (automated guided vehicle) oder auch ein manueller Abtransport von Auffangbehälter [Hof21].

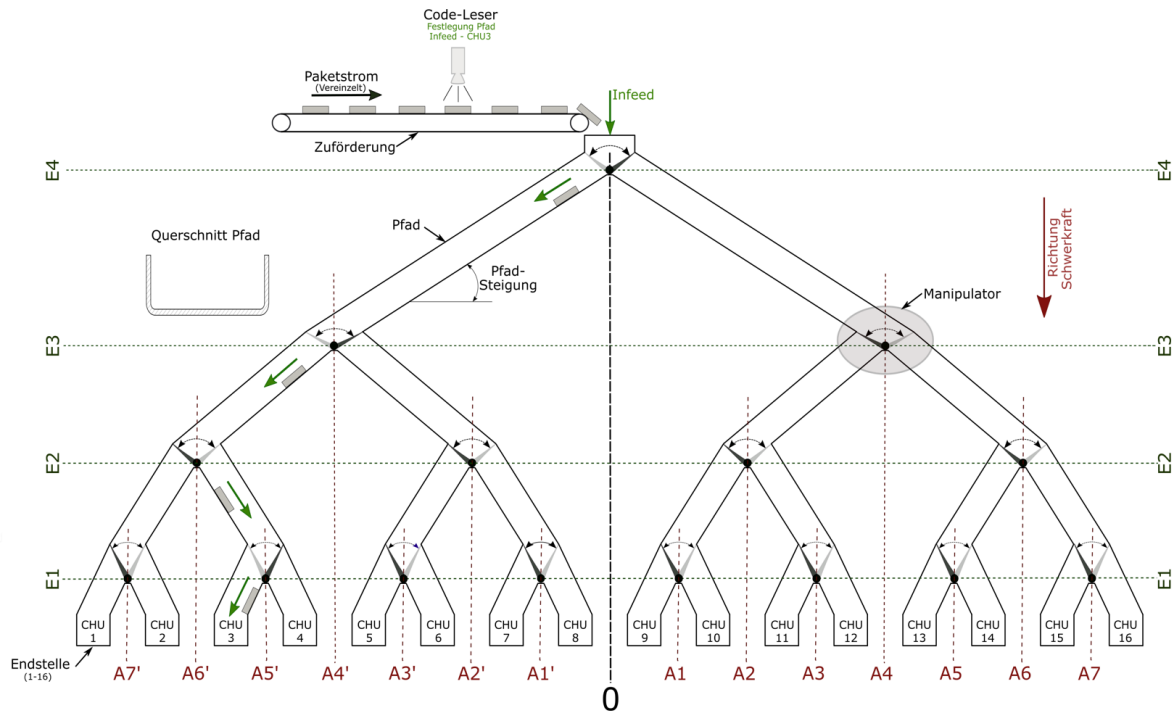


Abbildung 2.1: Funktionsprinzip des FFS [SL21]

Das Konzept des FFS ist somit eine Sortieranlage, welche neben herkömmlichen Sendungen, auch Sendungstypen mit sehr variablen Verhalten verarbeiten kann. Neben der Anpassung an die Anforderungen von neuen Sendungstypen, werden laut dem Erfinder auch die Anforderungen, welche für State-of-the-Art Systeme bestehen erfüllt.

Als grundlegende Forderung gelten laut Schedler [SL21]:

- geringer regelungstechnischer Aufwand
- geringer Flächenbedarf
- hohe Energieeffizienz bei niedrigem Gesamtenergieverbrauch
- breitestmögliche funktionale Absicherung der Entwurfsqualität durch virtuelle Produktentwicklung
- Bereitstellung von Konstruktions-Know-how, durch methodisches Konstruktion für anpassbare Variantenkonstruktionen und Ableitung einer anwenderfreundlichen Durchsatzberechnung

2.2 Statistische Versuchsplanung

Statistische Versuchsplanung oder Design of Experiment (DoE) ist eine Methode für die Planung und Auswertung von Versuchsreihen. Sie ist interdisziplinär einsetzbar und kann auf nahezu alle Problemstellungen angewandt werden. Einer der häufigsten Gründe für die Verwendung von Statistischen Versuchsplänen ist die Untersuchung von realen und auch scheinbaren Effekten, welche bei Variationen der Versuchsparameter entstehen. Diese Untersuchungen können bereits in den ersten Auslegungsschritten Erkenntnisse liefern, auf welche Systemeigenschaften ein besonderes Augenmerk gelegt werden muss. Die benötigten Versuchsergebnisse für die Auswertung der Versuchspläne können aus realen Versuchen oder aus computergestützten Simulationsergebnissen gewonnen werden. Dadurch ist es möglich, auch Versuchspläne mit einer hohen Anzahl von Versuchen mit geringem Aufwand zu erzeugen. Speziell Versuche aus Computer Aided Engineering (CAE) gestützten Anwendungen gewinnen durch die stetige Leistungssteigerung von Computern an Bedeutung, was den Trend zur Einsparung von Prototypen unterstützt. Worin sich ein weiterer Vorteil von CAE gestützten Versuchen verbirgt, da reale Versuche in der Regel kostenintensiv sind und sich die Kontrolle der gesteuerten Parameter als sehr komplex erweist [SvH17].

Durch die Erkenntnisse, aus den Versuchsplänen, können beispielhaft aufwendige Rechenmodelle für Auslegungen stark vereinfacht werden. Auch Effekte aus bestehenden Systemen können durch Versuchspläne nachgewiesen werden. Somit können auch Effekte dessen Ursprung physikalisch nicht zuordenbar sind, ihrer Herkunft zugewiesen werden. Dadurch ist die Versuchsplanung ein statistisches Werkzeug, welches für Vorauslegungen bis hin zu Produkttestungen eingesetzt werden kann.

2.2.1 Aufbau von Versuchsplänen

Da eine der größten Stärken von Versuchsplänen in der weltweit standardisierten Vorgehensweise und Ergebnisdarstellung liegt, werden auch in dieser Arbeit die üblichen Begrifflichkeiten verwendet. Das zu untersuchende System wird wie eine Blackbox betrachtet. Dazu müssen im ersten Schritt Systemgrenzen festgelegt werden. Das betrachtete System zeichnet sich aus, indem es eine oder mehrere Funktionen erfüllt. Die erfüllte Funktion wird als Qualitätsmerkmal oder Zielgröße definiert und entsteht durch die Interaktion der Einflussgrößen innerhalb der Systemgrenze. Bei den Einflussgrößen unterscheidet man zwischen Faktoren und Parametern. Als Parameter werden alle Größen bezeichnet, welche auf das System wirken können. Parameter welche steuerbar sind, werden als Faktoren bezeichnet. Die Betrachtung als Blackbox mit den dazugehörigen Eingangsgrößen und der Zielgröße wird in Abbildung 2.2 dargestellt [SvH17].

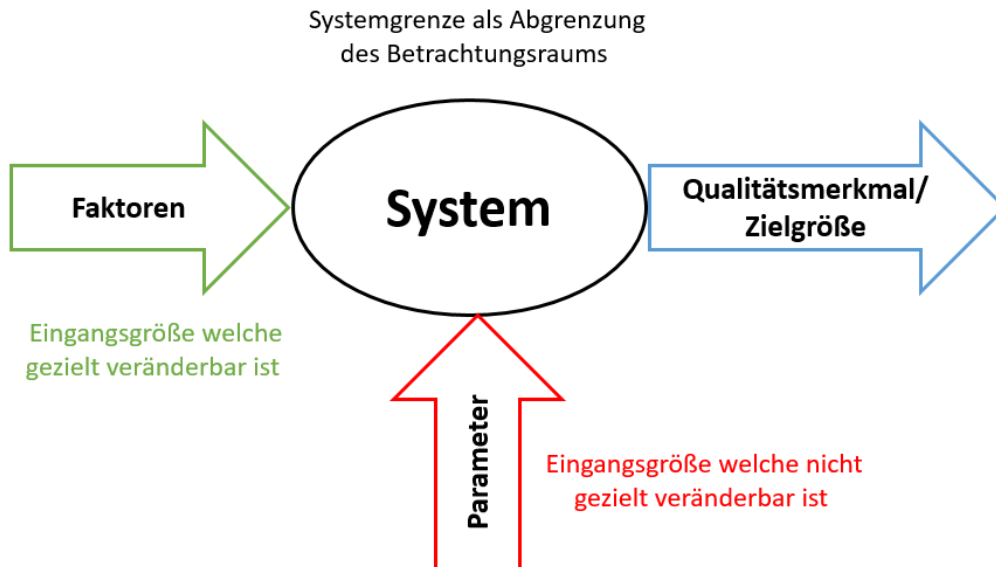


Abbildung 2.2: Allgemeine Darstellung eines Modells als Blackbox

Damit Effekte der Faktoren auf die Zielgröße nachweisbar sind, ist es notwendig, dass Qualitätsmerkmale kontinuierliche und messbare Größen sind. Faktoren sind jene Größen, welche in weiterer Folge in den Versuchsplan aufgenommen und untersucht werden. Somit ist bei der Definition von Faktoren darauf zu achten, dass diese Parameter reproduzierbar einstellbar sind. Dabei kann zwischen qualitativen und quantitativen Faktoren unterschieden werden [Kle13]. Bei qualitativen Faktoren behilft man sich, indem man einer Faktoreinstellung zum Beispiel „Strom ein“ den Wert + und „Strom aus“ den Wert - zuweist. Für quantitative Faktoren kann direkt der vorliegende Zahlenwert der Größe im Versuchsplan verwendet werden.

Bei der Untersuchung der Faktoren werden die nicht einstellbaren Parameter möglichst konstant gehalten, damit deren statistische Schwankung die Ergebnisse nicht zu stark beeinflusst. Die Variation in denen die Faktoren geprüft werden, nennt man Stufen oder Level, wobei eine Stufe je eine Einstellung des Faktors bedeutet [SvH17]. Bevor die Stufenanzahl festgelegt wird, muss der zu untersuchende Bereich durch eine Ober- und Untergrenze festgelegt werden. Ist die verwendete Anzahl der Stufen bekannt, müssen diese über dem Intervall zwischen den Grenzen in definierten Abständen verteilt werden. Hierbei ist noch zu berücksichtigen, ob die gewählten Faktoreinstellung am verwendeten Versuchsaufbau eingehalten werden können, oder ob die Faktoren aus physikalischen oder fertigungstechnischen Gründen nur eingeschränkt aufteilbar sind. Für die Aufteilung der Stufen gibt es in der Literatur eine Vielzahl von Methoden, welche an den jeweiligen Anwendungsfall angepasst sind.

2.2.2 Faktorielle Versuchspläne

Als Beispiel für einen vollfaktoriellen Versuchsplan dient die Untersuchung von drei Faktoren, welche in zwei Stufen getestet werden. Vollfaktoriell bedeutet hierbei, dass alle Kombinationen der Faktoren unter Variation der Stufen miteinander getestet werden. Dafür ergibt sich ein Versuchsplan, wie in Tabelle 2.1 dargestellt. Einer Stufenanzahl von zwei liegt die Annahme eines linearen Verlaufs des Effektes zu Grunde. Werden nicht lineare Verläufe erwartet, so findet man speziell angepasste Versuchspläne für nichtlineare Zusammenhänge. Dabei wird die Stufenanzahl meist erhöht und die Verteilung der Stufenwerte normiert [Kle13] .

Wie in Tabelle 2.1 ersichtlich, werden Versuche für die Kombination aller Faktoren durchgeführt. Durch diese Auswertung können die einzelnen Effekte sowie auch die Effekte, welche durch die Kombination der Faktoren (AB, AC, BC und ABC) auftreten, untersucht werden.

Tabelle 2.1: Vollfaktorieller Versuchsplan

Versuchs Nr.	Faktoren/Effekte						
	A	B	C	AB	AC	BC	ABC
Z1	+	-	-	+	+	+	-
Z2	-	-	-	-	-	+	+
Z3	+	+	-	-	+	-	+
Z4	-	+	-	+	-	-	-
Z5	+	-	+	+	-	-	+
Z6	-	-	+	-	+	-	-
Z7	+	+	+	-	-	+	-
Z8	-	+	+	+	+	+	+

Die Anzahl der benötigten Versuche kann laut Formel (2.1) berechnet werden. Diese besagt, dass m Versuche, bei einer Betrachtung von s Stufen und k Faktoren benötigt werden.

$$m = s^k \quad (2.1)$$

Da die Versuchszahl, bei einer steigenden Anzahl von Faktoren, stark zunimmt, ist es wichtig den Einfluss des jeweiligen Faktors auf die Zielgröße zu kennen. Um diesen Einfluss der Faktoren ermitteln zu können, gibt es spezielle Screening-Versuchspläne. Diese dienen dazu Faktoren mit geringen Einflüssen, vor der weiteren Untersuchung, auszusortieren. Die Faktoren, welche sich aus den Screening Versuchsplänen als nicht relevant erweisen, können im weiteren Verlauf der Versuche konstant gehalten werden, wodurch sich die Versuchsanzahl für den weiteren Ablauf reduziert.

Um die Versuchsanzahl zu senken, gibt es neben Screening Versuchsplänen auch die Möglichkeit mit teilfaktoriellen Versuchsplänen zu arbeiten. Hierbei werden Annahmen getroffen, dass ausgewählte Einzeleffekte zu einem Effekt summiert werden und/oder Wechselwirkungen höherer Ordnung vernachlässigt werden. Bei der Summierung spricht man von einem Vermengen oder Falten von Effekten. Wodurch sich die benötigte Versuchsanzahl laut der Formel (2.2) berechnen lässt. Der Unterschied zu Formel (2.1) besteht einzig durch den Subtraktionsterm p in der Hochzahl, welcher für die Anzahl der Faltungen steht.

$$m = s^{k-p} \quad (2.2)$$

In Tabelle 2.2 wird ein Beispiel angeführt, wie aus dem vollfaktoriellen Versuchsplan aus Tabelle 2.1, ein teilfaktorieller Versuchsplan erzeugt wird. Hierbei werden Kombinationen der Effekte A, B und C zu einem neuen Faktor D vermengt. Durch die Vermengung kann bei der Auswertung zwischen den einzelnen Effekten nicht mehr unterschieden werden [Kle13]. Der Vorteil von teilfaktoriellen Versuchsplänen wird hierbei jedoch schnell ersichtlich, da sich die Anzahl der benötigten Versuche um nahezu die Hälfte verringert.

Tabelle 2.2: Teilfaktorieller Versuchsplan

Versuchs Nr.	Faktoren/Effekte			
	A	B	C	D
Z1	+	-	-	-
Z2	-	-	-	+
Z3	+	+	-	+
Z4	-	+	-	-
Z5	+	-	+	+
Z6	-	-	+	-
Z7	+	+	+	-
Z8	-	+	+	+

Da in vielen Betrachtungen die Effekte aus der Kombination von mehr als zwei Faktoren nicht berücksichtigt werden und häufig ein Vermengen der Effekte von zwei Faktoren zulässig ist, finden teilfaktorielle Versuchspläne sehr häufig ihre Anwendung. Sollten die Ergebnisse für die Beurteilung zu ungenau sein, so können teilfaktorielle Versuchspläne nachträglich wiederum zu einem vollfaktoriellen Versuchsplan ergänzt werden.

Vollfaktorielle und teilfaktorielle Versuchsplänen sind Grundbausteine für die Untersuchung von Effekten. In der Literatur gibt es noch unzählige Varianten, wie Versuchspläne an spezifische Anwendungsfälle angepasst werden können [Kle13].

2.2.3 Effektbestimmung

Die Hauptaufgabe von Versuchsplänen ist die Effektbestimmung eines Faktors auf eine Zielgröße. Als Haupteffekt wird in der Literatur der alleinige Effekt eines Faktors bezeichnet. Alle anderen Effekte werden hingegen als Wechselwirkungen bezeichnet. Die Kombinationswirkung von mehreren Faktoren wird entsprechend als zweifach- oder dreifach-Wechselwirkung bezeichnet [BM13].

Wie groß der Einfluss eines Faktors auf die Zielgröße ist, lässt sich durch Mittelwertbildung als quantitative Größe kennzeichnen. Dazu wird der Mittelwert des Zielwertes bei einer Faktoreinstellung „+“ von dem Mittelwert der Faktoreinstellung „-“ subtrahiert.

Als Beispiel für die Berechnung des Effektes A wird der vollfaktorielle Versuchsplan aus Tabelle 2.1 herangezogen. In Formel(2.3) werden wie zuvor angeführt die Mittelwerte der Zielgröße y bei der Faktoreinstellung „-“ von den Mittelwerten der Zielgröße bei der Faktoreinstellung „+“ subtrahiert.

$$E_A = \frac{y_{Z1} + y_{Z3} + y_5 + y_{Z7}}{4} - \frac{y_{Z2} + y_{Z4} + y_6 + y_{Z8}}{4} \quad (2.3)$$

Der Effekt des jeweiligen Faktors wird dann wie in Abbildung 2.3 in einem Effektdiagramm abgebildet. Auf der vertikalen Achse wird die Zielgröße aufgetragen und auf der horizontalen Achse sind die Faktoren in seinen unterschiedlichen Stufen angeführt. Die Kodierung der Stufen dient einer dimensionslosen Darstellung und ermöglicht es somit, Faktoren mit unterschiedlichen Einheiten direkt miteinander vergleichen zu können. Wenn die Stufeneinstellungen normiert sind, lässt sich die Stärke des Effekts über die Steigung der Geraden ermitteln. Die Gerade wird erzeugt, indem man die Punkte der jeweiligen Mittelwerte der gleichen Stufeneinstellungen miteinander verbindet [SvH10].

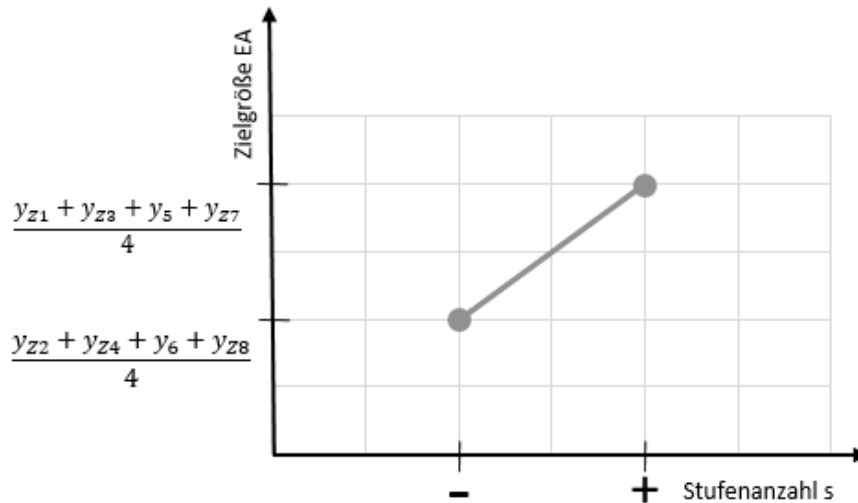


Abbildung 2.3.: Effektogramm des Faktors A

2.2.4 Korrelationsanalyse

In der Korrelationsanalyse wird der Zusammenhang zwischen einer Eingangsgröße und einer Ausgangsgröße ermittelt. Dazu dient der Korrelationskoeffizient als ein statistisches Mittel, für die Definition der Stärke des linearen Zusammenhangs. Somit lässt sich nicht nur bestimmen, ob eine Korrelation besteht, sondern auch ob die Korrelation einen entscheidenden Einfluss aufweist oder ob sie vernachlässigt werden kann. Zum Ausdruck gebracht wird dies über den empirischen Korrelationskoeffizienten R , der auch als Bravais-Pearson-Korrelationskoeffizient bezeichnet wird [FHK+16]. Die Berechnung erfolgt laut Formel (2.4)⁵

$$R = r_{XY} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{\tilde{s}_{XY}}{\tilde{s}_X \cdot \tilde{s}_Y} \quad (2.4)$$

Wobei der Zähler der Formel (2.4) die empirische Kovarianz, durch die Formel (2.5) ausdrückt.

$$\tilde{s}_{XY} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \quad (2.5)$$

Der Nenner der Formel (2.4) besteht aus einer Multiplikation, der Standardabweichung, der Merkmale X und Y , welche in den Formeln (2.6) und (2.7) abgebildet werden.

⁵ Alle angeführten Formeln für die Korrelationsanalyse stammen aus dem Werk [FHK+16].

$$\tilde{s}_X = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (2.6)$$

$$\tilde{s}_Y = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.7)$$

Der Zweck des Terms \tilde{s}_{XY} liegt in der Normierung des Korrelationskoeffizienten. Die beiden Terme \tilde{s}_X und \tilde{s}_Y ergeben die jeweilige Summe der Abweichungsprodukte. Da die Streuung positiv wie auch negativ sein kann, befindet sich der Korrelationskoeffizient R in einem Intervall von $[-1,1]$. Die Stärke der Korrelation wird laut Tabelle 2.3 definiert [FHK+16].

Tabelle 2.3: Definition der Stärke eines Korrelationskoeffizienten r

Schwache Korrelation	$ R < 0,5$
Mittlere Korrelation	$0,5 \leq R < 0,8$
Starke Korrelation	$0,8 \leq R $

Um den Korrelationskoeffizienten genauer zu beschreiben, können zusätzlich zur Größe, auch die Vorzeichen des Korrelationskoeffizienten gedeutet werden. Dazu werden die Begriffe gleichsinniger- und gegensinniger Zusammenhang verwendet. Bei gleichsinnigen Zusammenhängen treten bei großen x -Werten auch große y -Werte auf, beziehungsweise bei kleinen x -Werten liegen kleine y -Werte vor. Dies zeichnet sich mit einem positiven Vorzeichen im Korrelationskoeffizient ($R > 0$) ab, was tendenziell auf eine Gerade mit einer positiven Steigung hinweist. Gegensinnig wird der Zusammenhang bezeichnet, wenn kleine x -Werte auf große y -Werte treffen und vice versa. Kennzeichen dafür ist ein negativer Korrelationskoeffizient ($R < 0$), was einer negativen Steigung einer Geraden entspricht. Bei dem Fall, dass der Koeffizient ($R = 0$) entspricht, tritt entweder keine Korrelation auf oder der Zusammenhang ist nicht linear, da durch den Korrelationskoeffizienten nur lineare Korrelationen nachgewiesen werden können [FHK+16].

2.3 Regressionsmodelle

Durch die Erstellung eines Regressionsmodells wird der Zusammenhang aus mindestens zwei erklärenden Variablen und einer Reaktionsvariablen modelliert [Esr21]. Das Grundprinzip besteht darin, dass eine Schätzgerade durch vorliegende Funktionswerte gelegt wird. Durch die dazugehörige mathematische Funktion ist es dann möglich, Werte für unbekannte Variablen zu ermitteln. Einsatz finden sie in naturwissenschaftlichen bis hin zu sozialwissenschaftlichen Anwendungen [Fro18].

Grundlage für alle Regressionsmodelle bildet ein Datensatz von Variablen deren Reaktionsgröße durch eine mathematische Funktion angenähert wird. Als Beispiel für das einfachste Regressionsmodell wird in Abbildung 2.4 ein lineares Modell abgebildet. Die mathematische Beschreibung entspricht hierbei einer Geradengleichung, welche durch die Messpunkte angenähert wird. In Abhängigkeit von der betrachteten Problemstellung können die Daten einem Trend folgen, saisonale Schwankungen aufweisen oder einer Kombination der beiden Verläufe unterliegen. Variablen, dessen Verlauf sich nicht durch die zuvor genannten Effekte beschreiben lassen, müssen aus der Betrachtung ausgeschlossen werden, wobei diese zu einer Streuung in den Daten führt [FKL09].

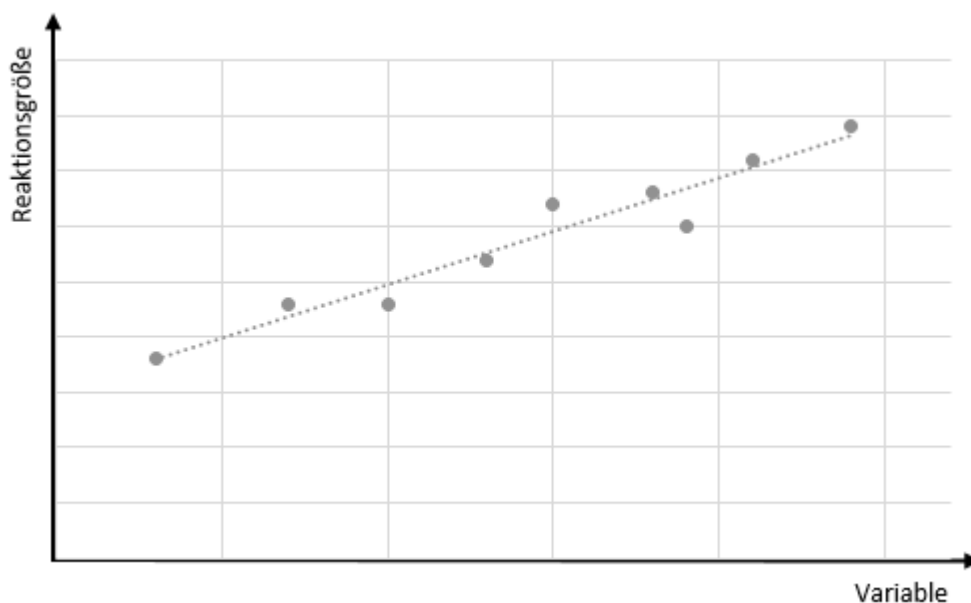


Abbildung 2.4: Beispiel für die Darstellung einer linearen Regression

Durch die Anwendung auf immer umfangreichere Problemstellungen werden immer komplexere mathematische Beschreibungsmodelle benötigt. Für die Ermittlung des passenden Modells dienen für erste Betrachtungen meist grafische Darstellungsmethoden. Wobei in den meisten Anwendungen Streudiagramme eingesetzt werden, um Muster in den Datensätzen zu erkennen. Diese Methode funktioniert jedoch nur für Modelle, welche sich durch einfache mathematische Funktionen beschreiben lassen. Für komplexere Datensätze werden mathematischen Berechnungsprogramme zu Hilfe gezogen, worin eine Vielzahl von Beschreibungsmöglichkeiten hinterlegt sind.

Für die Findung des optimalen Modells durch Berechnungsprogramme wird für den vorgegebenen Datensatz jedes zur Verfügung stehende Beschreibungsmodell angewandt und dessen

Brauchbarkeit beurteilt. Ziel hierbei ist es ein Modell zu finden, welches die geringste Abweichung gegenüber dem realen Modell aufweist.

Als Kriterium für die Brauchbarkeit des Modells wird der Wert des Root-Mean-Square-Errors (RMSE) herangezogen. Dieser gibt den quadrierten Wert der Abweichung zwischen dem realen Wert und dem vom Modell prognostizierten Wert aus. Dafür wird der zur Verfügung stehende Datensatz in ein Trainingsset und ein Validierungsset unterteilt. Das jeweilige erzeugte Regressionsmodell wird in weiterer Folge mit den Daten aus dem Validierungsset überprüft. Wichtig hierbei ist es, dass das Modell unabhängig von den Validierungsdaten erzeugt wird. Um diese Methode anzuwenden, muss der Datensatz ausreichend groß sein, da ansonsten durch die Aufteilung der Daten eine zu hohe Ungenauigkeit entsteht. Für die Aufteilung des Datensatzes gibt es verschiedene Möglichkeiten der Kreuzvalidierung, welche in Kapitel 2.3.3 näher erläutert werden [FKL09]. Als Beispiel für ein mathematisches Berechnungsprogramm dient hierbei Matlab. Als Übersicht werden die darin verfügbaren Beschreibungsmodelle in Tabelle 2.4 angeführt.

Tabelle 2.4.: Verfahren für Regressionsmodelle aus dem Machine learning Tool von Matlab [The22]

Linear Regression	Support Vector Machine	Gaussian Process Regression	Regression Trees
<ul style="list-style-type: none"> • Linear • Interactions • Linear • Robust Linear • Stepwise Linear 	<ul style="list-style-type: none"> • Linear SVM • Quadratic SVM • Cubic SVM • Gaussian SVM 	<ul style="list-style-type: none"> • Squared Exponential GPR • Matern 5/2 GPR • Exponential GPR • Rational Quadratic GPR 	<ul style="list-style-type: none"> • Fine Tree • Medium Tree • Coarse Tree

2.3.1 Stichprobengröße

Stichproben werden verwendet, um etwas über die Grundgesamtheit der vorliegenden Versuche zu lernen. Die Größe des Stichprobenumfangs hängt von der gewünschten Genauigkeit ab, mit der das Modell die Wirklichkeit abbilden sollte. Ein weiterer Faktor für die Wahl der Stichprobengröße ist die Größe des nachzuweisenden Effektes. Für kleine Effekte wird daher eine hohe Anzahl an Stichproben benötigt, da die Effekte ansonsten nicht erkennbar sind [Kle13].

Nach Kenntnis der Autoren aus dem Werk [FKL09] kann für die Größe des Stichprobenumfanges keine allgemeine Aussage getroffen werden. Grund dafür ist die starke Abhängigkeit von der Problemstellung und dem verwendeten Beschreibungsmodell. Daher kann für die optimale Stichprobengröße nur eine näherungsweise Lösung gefunden werden. Als Hilfe dazu dienen statistische Berechnungen und Systemkenntnisse [FKL09].

2.3.2 t-Verteilung

Die t-Verteilung, auch häufig als student'sche Verteilung bekannt, kann für normalverteilte Stichproben deren Standardabweichung nicht bekannt ist, mittels Schätzung der Standardabweichung, ein Vertrauensintervall über einen unbekanntem Mittelwert festlegen. Diese Methode findet häufig Anwendung, da die t-Verteilung schon mit einer geringen Versuchsanzahl die Verteilungsfunktion sehr gut approximiert [Eck19]. Grundvoraussetzung für die Anwendung der t-Verteilung ist es, dass die Stichproben normalverteilt sind und zufällig aus der Grundgesamtheit entnommen werden.

Die Anzahl der benötigten Stichproben ergibt sich aus der gewünschten Überdeckungswahrscheinlichkeit, welche den Vertrauensbereich (Konfidenzintervall) für die jeweilige Stichprobenanzahl festlegt [Wal15]. Der Verlauf der Dichtefunktion der t-Verteilung ändert sich mit der Stichprobenanzahl n . Für eine geringe Anzahl besitzt die Funktion im Vergleich zur Standardnormalverteilung einen Verlauf, der an den Enden nicht so stark abfällt. Somit ist die Funktion breiter verteilt, wodurch auch die Wahrscheinlichkeit steigt, dass der Wert im Zentrum von Null abweicht. Die Fläche unter der Dichtefunktion bleibt jedoch für jede Stichprobenanzahl gleich [FHK+16]. Für eine Stichprobenanzahl $n > 30$ geht die t-Verteilung in eine Standardnormalverteilung über [BBK17].

Der Vertrauensbereich einer t-Verteilung unterliegt, wie auch die Standardnormalverteilung einer Signifikanz, welche die Wahrscheinlichkeit angibt, mit welcher der Mittelwert innerhalb des Konfidenzintervalls liegt. Abhängig vom erforderlichen Signifikanzniveau muss der entsprechende $t_{n,G}$ -Wert, welcher für die Berechnung der oberen und unteren Grenze benötigt wird, gewählt werden. In Abbildung 2.5 wird ein Teil einer Tabelle für die Ermittlung des $t_{n,G}$ -Wertes dargestellt. Dabei wird in der Spalte mit der geforderten Wahrscheinlichkeit, der $t_{n,G}$ -Werte bei der benötigten Stichprobenanzahl ermittelt.

Anzahl der Einzelwerte	Werte für t und t/\sqrt{n}							
	für $P = 95,0\%$		für $P = 99,0\%$		für $P = 99,73\%$		für $P = 68,3\%$	
n	t	t/\sqrt{n}	t	t/\sqrt{n}	t	t/\sqrt{n}	t	t/\sqrt{n}
(3)	(4,30)	(2,48)	(9,92)	(5,72)	(19,21)	(11,00)	1,32	0,762
(4)	(3,18)	(1,59)	(5,84)	(2,92)	(9,22)	(4,61)	1,20	0,600
5	2,78	1,243	4,60	2,057	6,62	2,961	1,15	0,514
6	2,57	1,049	4,03	1,645	5,51	2,249	1,11	0,453
7	2,45	0,926	3,71	1,402	4,90	1,852	1,09	0,412
8	2,37	0,838	3,50	1,237	4,53	1,602	1,09	0,385
9	2,31	0,770	3,36	1,120	4,27	1,423	1,07	0,357
10	2,26	0,715	3,25	1,028	4,09	1,293	1,06	0,335

Abbildung 2.5.: Tabelle für die Ermittlung des t-Wertes [Goß20]

Die Berechnung der Grenzwerte für die ausgewählte Stichprobenzahl erfolgt nach Formel (2.8) und Formel (2.9). Der Mittelwert der Stichproben wird hierbei mit \bar{x} gekennzeichnet. Da die Berechnung des $t_{n,G}$ -Wertes sich als sehr komplex erweist, wird dieser meist aus Tabellen für die jeweilige Stichprobenanzahl bei dem gewünschten Signifikanzniveau abgelesen.

Die Standardabweichung für Stichproben wird mit s gekennzeichnet [Goß20]. Daraus erfolgt die Berechnung der oberen und unteren Grenze wie folgt:

$$\text{obere Grenze} = \bar{x} + t_{n,G} \frac{s}{\sqrt{n}} \quad (2.8)$$

$$\text{untere Grenze} = \bar{x} - t_{n,G} \frac{s}{\sqrt{n}} \quad (2.9)$$

Um eine aussagekräftige Stichprobengröße zu ermitteln, kann somit festgelegt werden, wie weit die Grenzen des Vertrauenswertes von dem gemessenen Wert abweichen dürfen. Durch den dafür benötigten $t_{n,G}$ -Wert ergibt sich in weiterer Folge die benötigte Anzahl an Versuchen, welche benötigt werden, um die Vorgaben der Grenzwerte einzuhalten. Aus Abbildung 2.6 wird ersichtlich, wie sich die Grenzen für das Konfidenzintervall festlegen lassen.

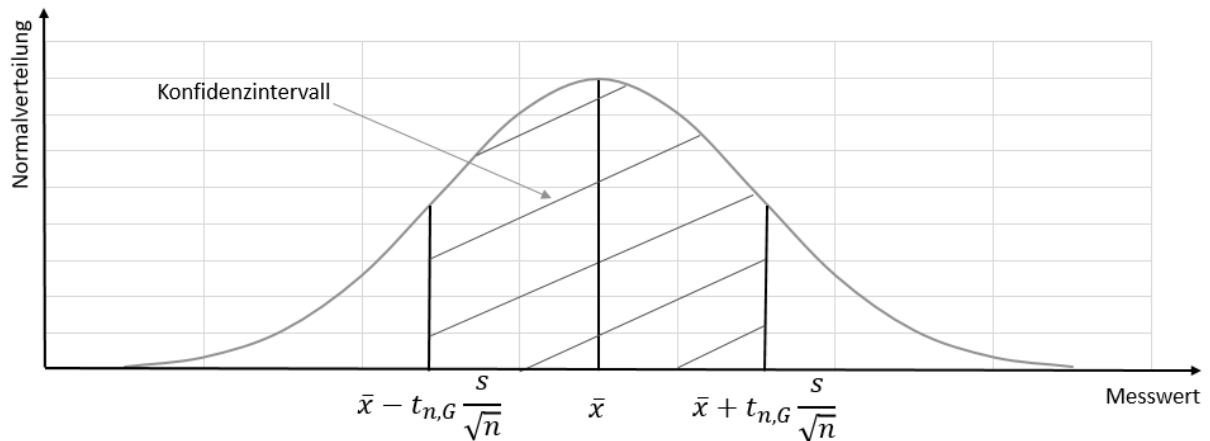


Abbildung 2.6.: Darstellung des Konfidenzintervalls inklusive der dazugehörigen Grenzen

2.3.3 Kreuzvalidierung

Im Bereich des maschinellen Lernens ist die Kreuzvalidierung eine statistische Möglichkeit, um in den ersten Schritten ein geeignetes Regressionsmodell zu finden oder auch um die Genauigkeit der prognostizierten Daten eines Modells zu überprüfen. Als Grundlage der Kreuzvalidierung dient die Aufteilung der Daten in Trainings- und Testdaten. Die Trainingsdaten (Teachingset) werden verwendet, um das Vorhersagemodell aufzubauen. Mit den Testdaten (Validierungsset), welche nicht zur Erzeugung des Modells verwendet wurden, wird die Abweichung der prognostizierten Werte von den Testwerten ermittelt. Als Maß für die Güte des Modells kann die Summe der quadratischen mittleren Abweichungen herangezogen werden. Dieses Vorgehen dient als allgemeines Grundgerüst der Kreuzvalidierung [FKL09]. Um die Aussagekraft zu erhöhen, gibt es verschiedene Möglichkeiten der Kreuzvalidierung. Diese unterscheiden sich in der Art, wie Trainings- und Testdaten aufgeteilt werden oder wie viele Tests durchgeführt werden. Im Weiteren werden die Methoden der zufälligen Untergruppen-Validierung und der k-fach (k-fold) Kreuzvalidierung ausführlicher betrachtet.

Zufällige Untergruppen Kreuzvalidierung

Hierbei wird der Datensatz zufällig in Trainings- und Testdaten unterteilt. Nachdem von allen Testdaten der Fehler ermittelt wurde, wird erneut zufällig in Training- und Testdaten unterteilt und dieser Vorgang anschließend mehrmals wiederholt. Der Mittelwert aller mittleren quadratischen Abweichungen wird dann als Gütekriterium verwendet. Das Problem dieser Betrachtung liegt darin, dass es vorkommen kann, dass Daten durch die zufällige Auswahl mehrmals und andere nie getestet werden, was zu einer Verfälschung des Ergebnisses führt [FKL09].

k-fach Kreuzvalidierung

Bei der k-fach Validierung handelt es sich um die am häufigsten angewandte Form der Kreuzvalidierung. Der gesamte Datensatz wird in k zufällig ausgewählte Teilmengen unterteilt. Jeder dieser Datensätze wird einmal für die Validierung verwendet und dient in den anderen Durchgängen als Trainingsatz. In Abbildung 2.7 wird das Vorgehen grafisch für eine vierfache Faltung dargestellt. Da die nicht getesteten Daten für die Generierung des Modells verwendet werden, muss das Modell mehrmals mit neuen Daten trainiert werden, was zu einer längeren Rechenzeit führt [SvH17].

FOLD1	Testdaten	Trainingsdaten	
FOLD2	Trainingsdaten	Testdaten	Trainingsdaten
FOLD3	Trainingsdaten		Testdaten
FOLD4	Trainingsdaten		Testdaten
Gesamter Datensatz			

Abbildung 2.7: Vorgehensweise für eine 4-fach Kreuzvalidierung

3 AUFBAU DER METHODIK

In diesem Kapitel wird der Aufbau der Methodik erläutert. Abbildung 3.1 bildet diesen in einem Flussdiagramm ab. Im Anschluss an die Abbildung des Prozesses folgt eine Erklärung der einzelnen Prozessschritte, um das Vorgehen allgemein zu beschreiben. In Kapitel 4 folgt danach die exemplarische Anwendung der Methodik am Free-Fall-Sorter.

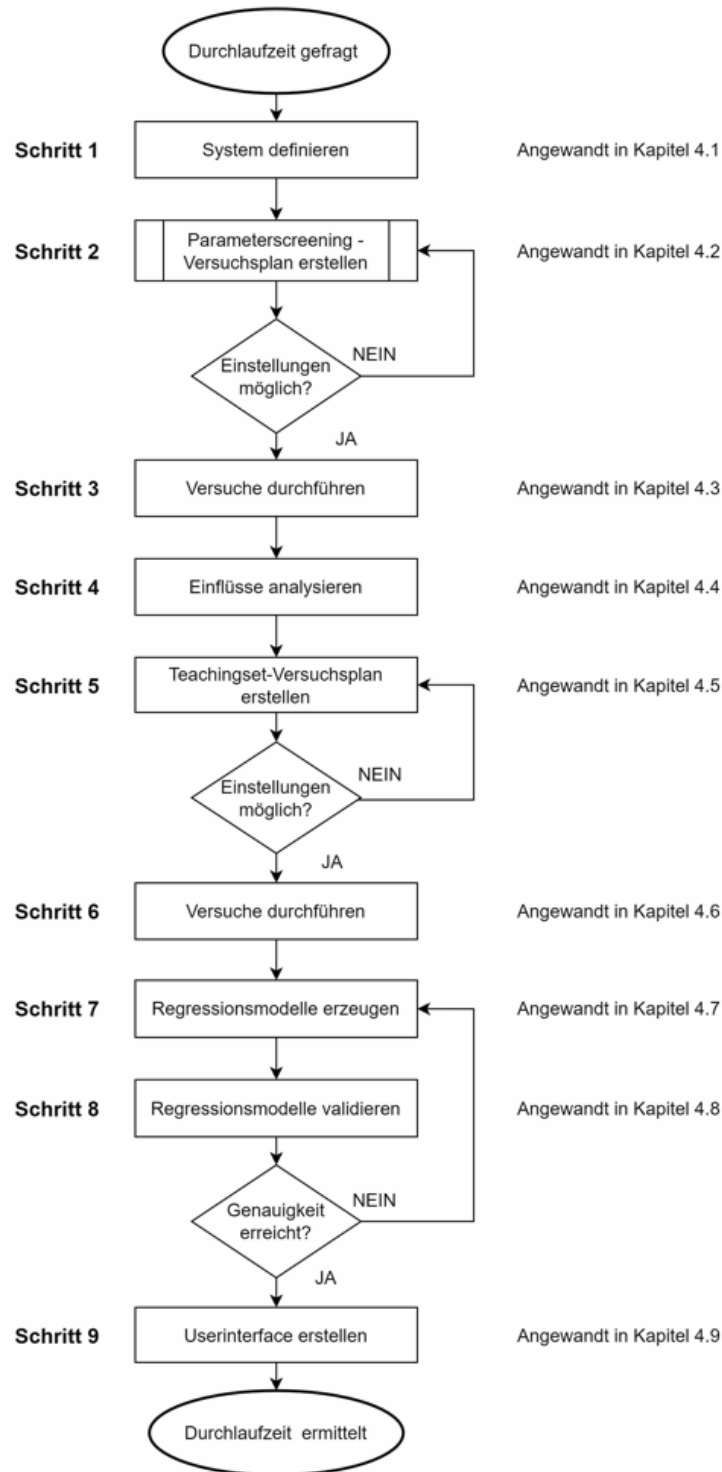


Abbildung 3.1.: Aufbau der Methodik für die Durchlaufzeitermittlung

In Tabelle 3.1 wird ein Überblick über den Inhalt der einzelnen Schritte gegeben. Darin wird die Kernaufgabe jedes Schrittes beschrieben.

Tabelle 3.1.: Ablauf der Methodik

Schritt 1	Standardisierten Ablauf, welcher eine vollständige Beschreibung des zu untersuchenden Systems liefert, definieren. Dabei werden alle Parameter ermittelt, welche einen Einfluss auf die Zielgröße – Durchlaufzeit haben können.
Schritt 2	Screening-Versuchsplan für die Analyse der Einflussfaktoren auf die Zielgröße, mit den ermittelten Parametern aus Schritt 1 erstellen. Dabei werden auch Vorkehrungen getroffen, welche einen standardisierten Ablauf bei der Versuchsdurchführung in Schritt 3 gewährleisten.
Schritt 3	Versuche des Screening-Versuchsplans unter Einhaltung der Durchführungsvorschrift durchführen. Die entstehenden Messergebnisse dienen als Grundlage für die Analysen in Schritt 4.
Schritt 4	Parameterscreening und Korrelationsanalyse für die ermittelten Daten aus Schritt 3 durchführen. Alle Parameter, mit geringem Einfluss, müssen nicht mehr in den Versuchsplan für die Trainingsdaten in Schritt 5 aufgenommen werden.
Schritt 5	Teachingset-Versuchsplan mit den Faktoren aus Schritt 4 erstellen. Auf Basis dieses Versuchsplans werden in Schritt 6 Versuchsdaten für die Regressionsmodelle erhoben.
Schritt 6	Versuche des Teachingset-Versuchsplans aus Schritt 5, unter Einhaltung der Versuchsvorschrift, durchführen. Die daraus entstehenden Daten dienen in Schritt 7 als Trainingsdaten für die Regressionsmodelle.
Schritt 7	Mit den Trainingsdaten aus Schritt 6 Regressionsmodelle für einzelne Abschnitte (Submodelle) der Anlage erzeugen und zu einem Gesamtmodell (Predictionmodell) zusammenführen.
Schritt 8	Überprüfen mit welcher Abweichung das Predictionmodell Ergebnisse liefert. Dazu werden die Ergebnisse von Realversuchen mit jenen des Predictionmodells validiert.
Schritt 9	Userinterface erstellen, um die Anwendung des Predictionmodells anwenderfreundlich zu gestalten.

3.1 Schritt 1 – System definieren

Der Grundstein, auf der die gesamte Methodik aufbaut, sind Versuchsdaten, welche für ein Parameterscreening, eine Korrelationsanalyse und die Erzeugung der Regressionsmodelle benötigt werden. Aus diesem Grund wird in Schritt 1 ein standardisierter Ablauf definiert, welcher eine vollständige Beschreibung des zu untersuchenden Systems (Sortieranlage) liefert. Dabei werden mit Hilfe verschiedener Methoden⁶ alle Parameter ermittelt, welche einen Einfluss auf die Zielgröße (Durchlaufzeit) haben können. Die Parameter werden dazu in Stückgut- und Anlagenparameter unterteilt und in weiterer Folge auf ihren Einfluss auf die Durchlaufzeit untersucht.

Als Diskussionsgrundlage für das weitere Vorgehen wird eine Skizze der Anlage erstellt. Darin werden Aufbau sowie alle Abmessungen, Kenngrößen und Bauteile inklusive ihrer Bezeichnungen eingezeichnet. Die, in der Skizze definierten, Bezeichnungen dienen als Kommunikationsgrundlage für das gesamte Vorgehen und müssen daher durchgängig eingehalten werden.

3.2 Schritt 2 – Parameterscreening-Versuchsplan erzeugen

Um festzustellen, ob der Effekt eines Parameters aus Schritt 1 zu berücksichtigen ist, wird in Schritt 4 ein Parameterscreening und eine Korrelationsanalyse durchgeführt. Dazu wird in Schritt 2 ein Screening-Versuchsplan erstellt, welcher die zuvor ermittelten Parameter (Stückgutparameter und Anlagenparameter) beinhaltet.

Damit gewährleistet wird, dass alle Stufen des Versuchsplans eingestellt werden können, müssen Vorbereitungen an der Anlage⁷ sowie den Proben⁸ getroffen werden. Um diesen Vorgang zu standardisieren, wird im Zuge der Versuchsplanerzeugung eine Vorschrift für die Versuchsdurchführung erstellt. Diese beinhaltet eine Dokumentation für die Vorbereitung der Anlage und der Proben. Weiters wird eine standardisierte Datenerfassung für die Messdaten eingeführt. Als letzter Schritt wird der benötigte Stichprobenumfang ermittelt. Dadurch wird festgelegt, wie viele Versuche für jede Einstellung durchgeführt werden müssen.

Lässt sich eine Stufe des Versuchsplans nicht einstellen oder mit anderen Stufen kombinieren, muss der Screening-Versuchsplan iterativ angepasst werden. Nach Beendigung von Schritt 2, kann

⁶ Möglichkeiten für die Untersuchung bieten Methoden wie z.B. Ishikawa Diagramme, Brainstorming und Mind-Mapping.

⁷ Parameter der Anlage, welche variiert werden

⁸ Eigenschaften der Proben, welche variiert werden

der dadurch entstandene Versuchsplan als Grundlage für die Versuchsdurchführung in Schritt 3 verwendet werden. In Abbildung 3.2 wird der Schritt 2 noch einmal als Unterprozess dargestellt.

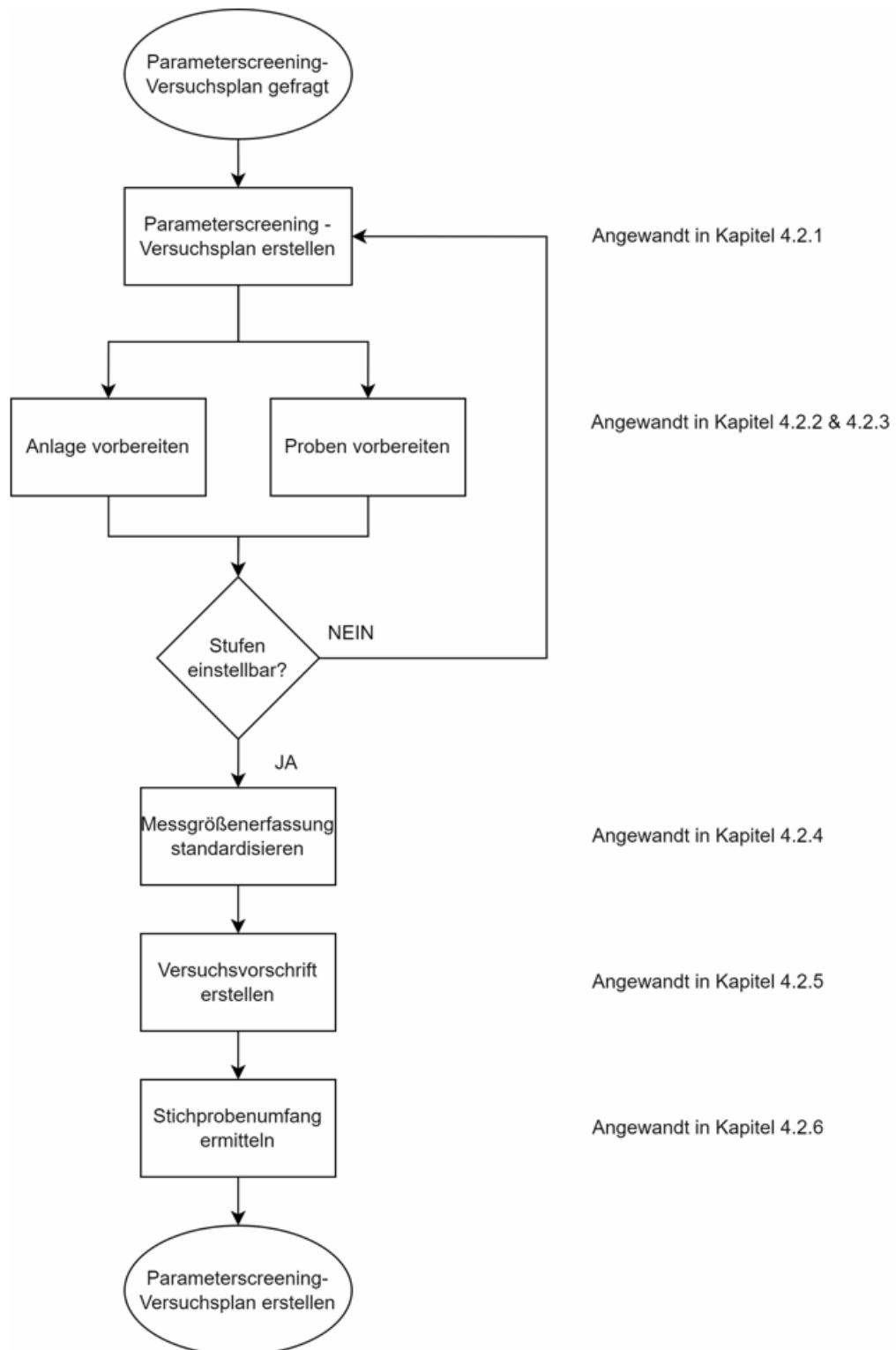


Abbildung 3.2.: Unterprozess Schritt 2 – Parameterscreening-Versuchsplan erstellen

In den folgenden Unterkapiteln wird die Erzeugung des Versuchsplans, die Vorbereitung der Proben und der Anlage, die Durchführungsvorschrift sowie die Wahl des Stichprobenumfangs erläutert.

3.2.1 Versuchsplan erstellen

Um einen vollfaktoriellen Versuchsplan wie in Tabelle 2.1 erstellen zu können, müssen allen Faktoren, welche getestet werden, Zahlenwerte für die Ober- und Untergrenze zugewiesen werden. Bei Faktoren, welche sich durch physikalische Größen beschreiben lassen, wird nur der Zahlenwert der Grenzen angegeben. Für Faktoren, welche durch subjektive Merkmale beschrieben werden, wird den Merkmalen ein Zahlenwert (siehe Tabelle 4.1.) zugewiesen.

3.2.2 Anlage vorbereiten

Vor den Versuchen muss sichergestellt werden, dass alle Stufeneinstellungen mit einer hohen Wiederholgenauigkeit reproduziert werden können. Daher muss im Vorfeld überprüft werden, ob alle benötigten Faktorstufen der Anlage einstellbar sind und ob sie sich miteinander kombinieren lassen.

3.2.3 Proben vorbereiten

Für die Einstellung der Stückgutparameter werden für jede Stufeneinstellung Proben präpariert, welche für das gesamte Screening die jeweilige Stufeneinstellung vertreten. Für jeden zu testenden Faktor müssen je nach Stichprobengröße (siehe Kapitel 3.2.6), entsprechend viele Proben vorbereitet werden. Hierbei ist wiederum zu beachten, dass alle Einstellungen miteinander kombinierbar sein müssen. Ansonsten muss der Versuchsplan iterativ angepasst werden.

3.2.4 Messgrößenerfassung standardisieren

Eine standardisierte Messgrößenerfassung wird für die Dokumentation und Auswertung der Messdaten benötigt. Als Messdaten werden Daten bezeichnet, welche durch die Durchlaufzeitmessung entstehen. Dazu muss ein entsprechender messtechnischer Aufbau für die Ermittlung der Durchlaufzeit der Proben am Sorter verbaut werden. Die daraus entstehenden Messdaten müssen danach durch ein geeignetes Computerprogramm aufgezeichnet und ausgewertet werden.

3.2.5 Versuchsvorschrift erstellen

Um die Versuche mit einer möglichst hohen Wiederholgenauigkeit durchführen zu können, muss eine Versuchsvorschrift erstellt werden. Diese beinhaltet eine Dokumentation aller Vorbereitungsschritte. Dabei wird neben der Vorbereitung auch der darauffolgende Ablauf der Durchführung der Versuche definiert.

3.2.6 Stichprobenumfang ermitteln

Die Ermittlung des Stichprobenumfangs wird benötigt um die statistische Schwankung, welcher die Versuche unterliegen, zu minimieren. Die Schwankungen lassen sich in Schwankungen aus Umgebungseinflüssen und in Schwankungen aus den Messungen unterteilen. Schwankungen aus Umgebungseinflüssen sind nicht steuerbar, können jedoch durch konstant halten der Umgebungsparameter als ein Grundrauschen aus den Messdaten gefiltert werden. Die Schwankungen, welche durch die Messung entstehen, können wiederum durch die Stichprobengröße beeinflusst werden. Durch die Berechnung der optimalen Stichprobenanzahl wird diese minimiert.

Für die Anzahl der benötigten Stichproben wird die zufällige Schwankung der Messwerte der Zielgröße herangezogen. Für die Ermittlung der Schwankung werden Versuche mit den gleichen Stückguteigenschaften sowie Anlageneinstellungen durchgeführt. Aus diesen Versuchen werden zufällige Stichproben gezogen, deren Schwankung mittels der t-Verteilung ermittelt wird. Über die jeweilig verwendete Anzahl und der geforderten Wahrscheinlichkeit, mit der die Zielgröße im Vertrauensbereich liegt, wird über die Standardabweichung der t-Verteilung die Ober- und die Untergrenze für die Zielgröße festgelegt. Aus dieser Standardabweichung der Stichproben lässt sich in weiterer Folge ermitteln, wie viele Versuche benötigt werden, damit der Messwert innerhalb des vorgegebenen Intervalls liegt.

Kolmogorov-Smirnov Test

Als Voraussetzung für die t-Verteilung gilt, dass die Stichproben normalverteilt sein müssen und zufällig gezogen werden. Daher muss vor der Berechnung der benötigten Stichprobenanzahl überprüft werden, ob eine Normalverteilung vorliegt.

Dafür wird mit den Versuchsdaten der Durchlaufzeit ein Kolmogorov-Smirnov Test durchgeführt. Dabei wird überprüft, ob die Nullhypothese für die Messungen verworfen werden kann. Dies ist der Fall, wenn die maximale Abweichung den kritischen Wert übersteigt. Der kritische Wert wird bei einem Signifikanzniveau von 0,05 aus der Tabelle A1.1 im Anhang A1 für die dazugehörige Stichprobengröße entnommen. Liegt der kritische Wert unter der ermittelten maximalen

Abweichung, so kann die Hypothese nicht verworfen werden, wodurch die Aussage getroffen werden kann, dass eine Normalverteilung der Daten vorliegt.

t-Verteilung

Nachdem nachgewiesen wurde, dass die Versuche normalverteilt sind, erfolgt die Ermittlung der benötigten Stichprobenanzahl. Diese wird über die Vorgabe der zulässigen Abweichung der Messgrößen bestimmt. Dabei werden über die t-Verteilung die Abweichungen der Messwerte für verschiedene Stichprobengrößen ermittelt. Beginnend bei zwei Stichproben wird die Anzahl gesteigert, bis die Abweichung der Messergebnisse sich in der gewünschten Toleranz befindet.

3.3 Schritt 3 – Versuche für das Parameterscreening durchführen

In Schritt 3 werden alle Versuche des im Schritt 2 definierten Versuchsplans durchgeführt. Dabei muss die in Schritt 2 definierte Vorschrift bei der Versuchsdurchführung eingehalten werden.

Die Versuche sollten dabei nicht willkürlich abgearbeitet werden, sondern es sollte zuvor die Versuchsdurchführung so geplant werden, dass der Aufwand für den Umbau der Anlage möglichst gering bleibt. Dadurch wird nicht nur der Aufwand minimiert, sondern auch Fehlerquellen, welche durch das Umrüsten entstehen können. Die Messdaten, welche sich durch die Versuche ergeben, dienen als Grundlage für das Parameterscreening und die Korrelationsanalyse in Schritt 4.

3.4 Schritt 4 – Einflüsse der Parameter analysieren

Die Analyse der Einflüsse der Parameter auf die Durchlaufzeit wird benötigt, um die Versuchsanzahl bei der Erzeugung der Trainingsdaten möglichst gering zu halten. Durch ein Parameterscreening und eine Korrelationsanalyse wird ermittelt, ob ein nennenswerter Einfluss des Parameters auf die Zielgröße besteht. Alle Parameter, welche nur einen geringen Einfluss aufweisen, müssen im Teachingset-Versuchsplan in Schritt 5 nicht berücksichtigt werden.

3.5 Schritt 5 – Teachingset-Versuchsplan erzeugen

Der Teachingset-Versuchsplan⁹ dient als Basis für die Erzeugung der Regressionsmodelle. Dazu wird ein Versuchsplan mit den entscheidenden Parametern aus Schritt 4 erzeugt. Im Gegensatz zum Screening-Versuchsplan werden nicht nur die obere und untere Grenze benötigt, sondern weitere Stützstellen, welche über den Parameterbereich aufgeteilt werden. In Schritt 5 wird dafür die Anzahl und die Verteilung der Stützstellen ermittelt. Als Ergebnis ergibt sich ein Versuchsplan,

⁹ Trainingsdaten, an welche das Regressionsmodell angepasst wird.

welcher die benötigten Versuche für das Teachingset der Regressionsmodelle beinhaltet. Dieser gilt als Grundlage für die Versuchsdurchführung in Schritt 6.

Die Wahl der Stützstellenanzahl wird iterativ durchgeführt. Die minimale Stützstellenanzahl ist mit zwei (obere und untere Grenze) je Parameter festgelegt. Ein Versuchsplan mit zwei Stützstellen je Parameter liegt durch das Parameterscreening bereits vor und kann somit auch als erstes Teachingset herangezogen werden. Um Vergleichswerte zu generieren und somit festzustellen, ob eine höhere Stützstellenanzahl benötigt wird, werden zwei weitere Versuchspläne erstellt.

Der erste Versuchsplan wird mit drei Stützstellen und der zweite mit vier Stützstellen je Parameter erzeugt. Die Aufteilung über dem Parameterbereich¹⁰ erfolgt hierbei möglichst¹¹ symmetrisch.

Bei der Wahl der Stützstellenanzahl gibt es zwei entscheidende Aspekte. Der erste Aspekt ist das Bestreben den Versuchsaufwand möglichst gering zu halten, um möglichst wenig Versuche zu benötigen. Als zweiter Aspekt gilt die Gefahr des Overfittings¹².

Aus diesen beiden Gründen fällt der erste Iterationsschritt der Stützstellenanzahl auf drei Stützstellen je Parameter. Um die Ergebnisse der Modelle mit drei Stützstellen je Parameter auf ein Overfitting untersuchen zu können, wird noch ein weiteres Modell mit vier Stützstellen je Parameter benötigt. Daher werden drei Modelle¹³ für die Wahl der optimalen Stützstellenanzahl benötigt.

3.6 Schritt 6 – Versuche für das Teachingset durchführen

Im Schritt 6 werden die Versuche aus dem Teachingset-Versuchsplan durchgeführt. Dabei muss wiederum die Durchführungsvorschrift aus Schritt 2 berücksichtigt werden. Die daraus entstehenden Daten dienen in Schritt 7 als Trainingsdaten für die Erzeugung der Regressionsmodelle.

¹⁰ Der Parameterbereich wurde im Kapitel 3.1 definiert.

¹¹ Bei der Verteilung können Einschränkungen durch z.B. fertigungstechnische oder physikalische Grenzen auftreten.

¹² Overfitting entsteht, wenn dem Regressionsmodell zu viele Werte für die Erzeugung vorgegeben werden. Das führt dazu, dass das Model durch die vielen Stützpunkte nicht mehr dynamisch genug ist, um sich an neue Werte anzupassen. Daher resultieren aus einem Overfitting fehlerhafte Vorhersagen.

¹³ Modell 1 – zwei Stützstellen je Parameter; Modell 2 – drei Stützstellen je Parameter; Modell 3 – vier Stützstellen je Parameter

3.7 Schritt 7 – Regressionsmodelle erzeugen

Mit den Trainingsdaten aus Schritt 6 werden für die einzelnen Submodelle der Anlage Regressionsmodelle erzeugt, welche zu einem Gesamtmodell zusammengeführt werden (siehe Beispiel Abbildung 4.25). Dazu werden die Regressionsmodelle durch ein computergestütztes Berechnungsprogramm erzeugt. Aus der Zusammenführung der Submodelle entsteht ein Predictionmodell für die Durchlaufzeit des Stückgutes.

Der Ablauf bei der Erzeugung der Regressionsmodelle wird in Abbildung 3.3 dargestellt. Dabei werden drei Regressionsmodelle mit den unterschiedlichen Stützstellen je Parameter aus Schritt 5 erzeugt.

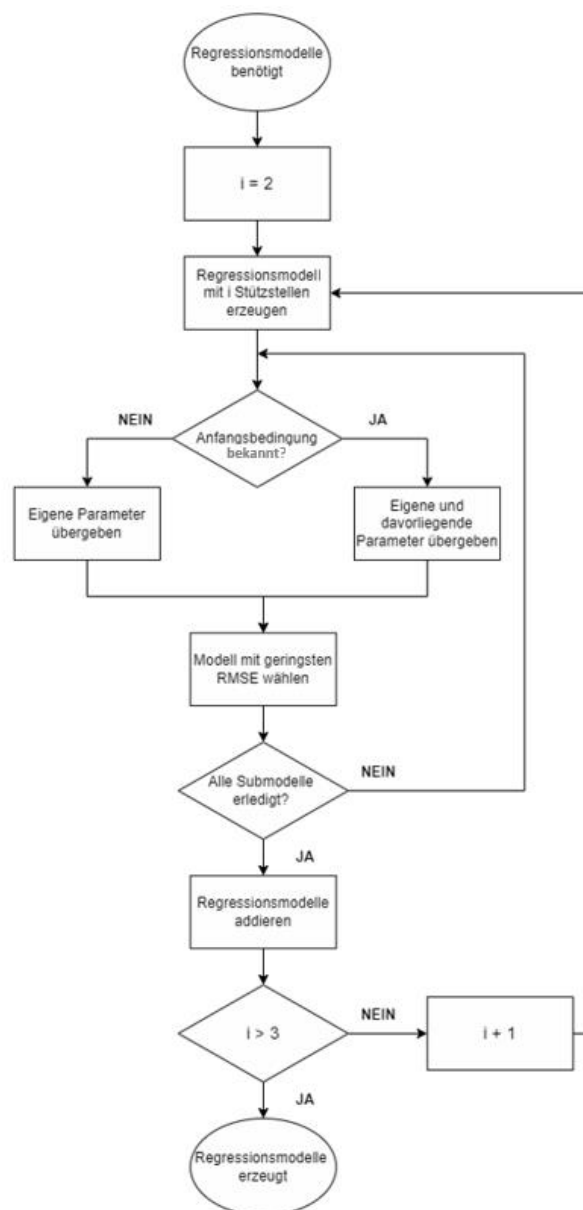


Abbildung 3.3.: Prozess für die Erzeugung von Regressionsmodellen

3.8 Schritt 8 – Regressionsmodelle validieren

Dieser Schritt dient der Überprüfung, mit welcher Genauigkeit das Predictionmodell Ergebnisse liefert. Weiters wird aus den Ergebnissen der Validierung der Modelle die optimale Stützstellenanzahl für jedes Submodell ersichtlich. Dazu werden die Ergebnisse von realen Versuchen mit jenen des Predictionmodells verglichen. Durch die Erkenntnisse der Validierung wird festgestellt, ob eine weitere Iteration¹⁴ benötigt wird oder ob das Predictionmodell die Anforderungen an die Genauigkeit erfüllt.

Um die Abweichung zwischen realen Werten und jenen aus dem Predictionmodell zu bestimmen, werden zusätzlich zu den Versuchen des Teachingsets auch Validierungsversuche durchgeführt. Dazu werden Proben zufällige Parameter zugewiesen, welche bei verschiedenen Anlageneinstellungen getestet werden¹⁵. Die Größe des Validierungssets hängt hierbei von der Größe des jeweiligen Teachingsets ab. Als Richtwert sollte der Umfang des Validierungssets ein Drittel der Versuchsanzahl aus dem Teachingset entsprechen [SvH17]. Wichtig hierbei ist es, dass die Validierungsdaten unabhängig vom Teachingset sind und nicht für die Erzeugung eines Regressionsmodells verwendet werden.

Für die zufälligen Parameter der Validierungsversuche werden in weiterer Folge die Durchlaufzeiten mit Hilfe des Predictionmodells ermittelt. Durch den Vergleich der realen und der vorhergesagten Werte lässt sich die Abweichung zu den realen Werten festlegen. Der Ablauf der Validierung ist in Abbildung 3.4 dargestellt.

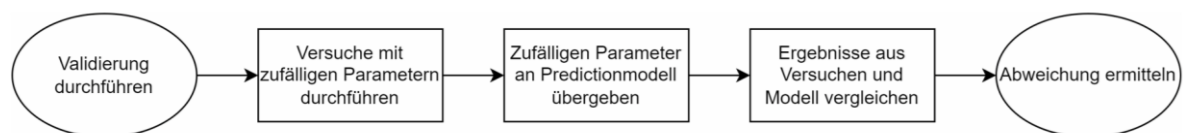


Abbildung 3.4.: Validierungsprozess des Predictionmodells

¹⁴ Anpassen der Stützstellenanzahl oder in weiterer Folge eine Änderung des Stichprobenumfangs

¹⁵ Proben- und Anlagenparameter werden innerhalb der oberen und unteren Grenze des Parameterbereichs aus Kapitel 3.2 variiert und getestet.

3.9 Schritt 9 – Userinterface erstellen

Um das Programm nicht über die Benutzeroberfläche des mathematischen Computerprogramms bedienen zu müssen, wird im letzten Schritt ein bedienerfreundliches Userinterface mittels eines GUI (Graphical-User-Interface) erzeugt.

Hierbei gibt es abhängig vom Anwendungsfall verschiedene Möglichkeiten des Aufbaus. Dabei kann über das Interface der gesamte Ablauf von Parameterscreening bis hin zur Validierung gesteuert werden. Auch Einstellungen wie die Stichprobengröße, Stützstellenanzahl oder Faltungszahl der Validierung können dazu eingebettet werden.

4 EXEMPLARISCHE ANWENDUNG DER METHODIK

In diesem Kapitel wird die Methodik aus Kapitel 3 am Beispiel des Free-Fall-Sorter angewandt. Als Anlage dient dabei ein Versuchstand des FFS, welcher sich in der Versuchshalle des Institutes für Technische Logistik befindet. Entsprechend dem FFS werden Polybags als Stückgut verwendet. Durch die Anwendung der Methodik wird neben der Veranschaulichung der einzelnen Schritte, auch eine Validierung der Gültigkeit der Methodik durchgeführt.

4.1 Anwendung Schritt 1 - System definieren

Für die Definition des Systems (FFS) ergibt sich eine Skizze der Anlage wie in Abbildung 4.1 dargestellt. Darin eingezeichnet sind alle relevanten Bauteile und Kenngrößen inklusive der dazugehörigen Bezeichnung.

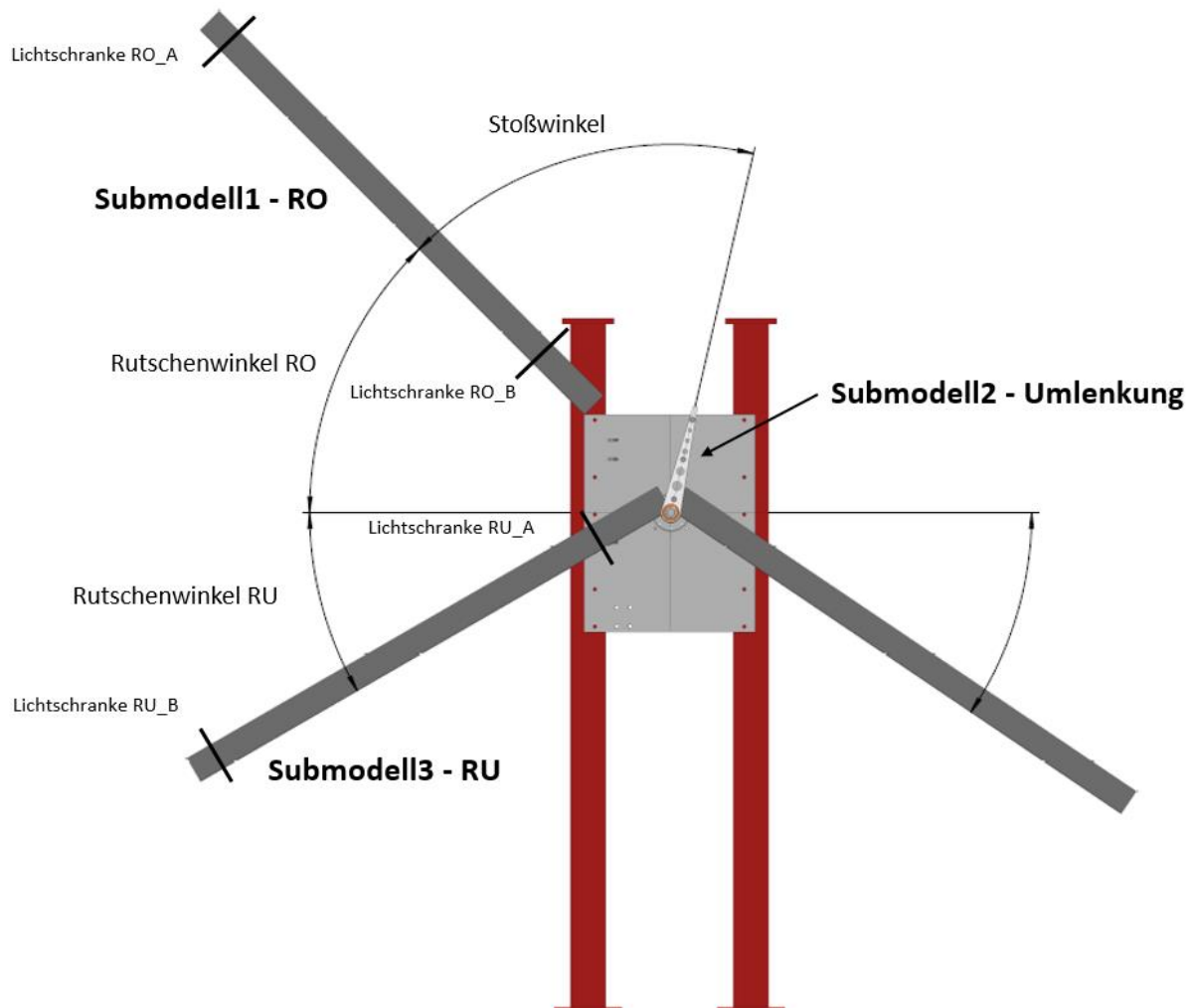


Abbildung 4.1: Aufbau des FFS Versuchstands

Abgeleitet aus der Anlagenskizze und den Systemkenntnissen der durchführenden Personen¹⁶, wird für den FFS ein Ishikawa Diagramm erstellt (siehe Abbildung 4.2). Dabei werden die Einflüsse auf die Durchlaufzeit in den sechs Bereichen: Mensch, Material, Maschine, Messung, Methode und Milieu untersucht. Besonderes Augenmerk wird hierbei auf die Vollständigkeit gelegt, da alle zu untersuchenden Parameter für den Versuchsplan im Schritt 2 aus diesem Diagramm abgeleitet werden.

In Abbildung 4.2 sind alle Parameter, von denen ein nennenswerter Einfluss auf die Durchlaufzeit zu erwarten ist, rot markiert. Aus dem Ishikawa Diagramm lassen sich für den Versuchsplan in Schritt 2 somit folgende Parameter ableiten:

- Anlagenparameter
 - Rutschenwinkel
 - Stoßwinkel
- Stückgutparameter
 - Form
 - Gewicht (Gesamtgewicht des Polybags = Verpackung + Inhalt)
 - Material
 - Reibwert
 - Biegeschlaffheit

¹⁶ Um eine ganzheitliche Betrachtung zu ermöglichen, sollten für die Untersuchungen in Schritt 1 alle Personen mitwirken, welche über Kenntnisse des Systems verfügen. Für die Einflüsse komplexer Parameter können weiters auch Spezialisten/Spezialistinnen des jeweiligen Fachgebietes zu Rat gezogen werden.

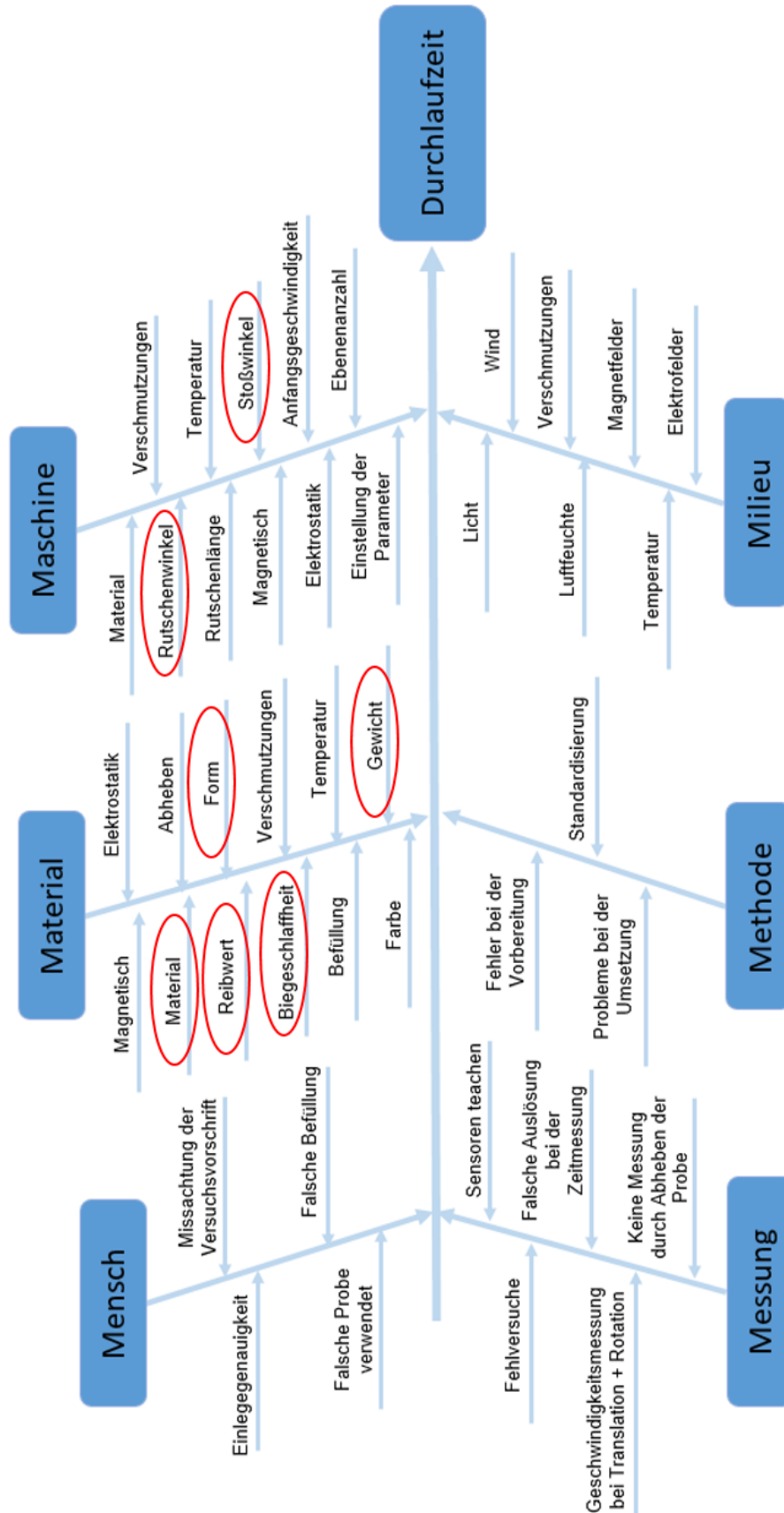


Abbildung 4.2: Ishikawa Diagramm mit den möglichen Einflussfaktoren auf die Durchlaufzeit

Nachdem alle zu untersuchenden Parameter durch das Ishikawa-Diagramm ermittelt worden sind, wird für jeden Parameter der zu untersuchende Bereich durch eine obere und untere Grenze definiert. Die daraus folgenden Einteilungen werden in Kapitel 4.1.1 und 4.1.2 erläutert.

4.1.1 Stückgutparameter

Die Festlegung der Grenzen der Parameter lassen sich aus dem Anwendungsfall, für welchen die Anlage bestimmt ist, ableiten. Im Falle des FFS handelt es sich hierbei um Kleinsendungen, dessen Anforderungen laut [SL21] definiert sind. Die Grenzen können dabei allgemein physikalische Größen, Verhältnisgrößen, dimensionslose Kennzahlen oder auch subjektive Qualitätsmerkmale sein.

Für den Polybag ergeben sich somit für die weitere Untersuchung folgende Grenzen:

- Gesamtgewicht [20g-220g]
- Biegeschlaffheit [schlaff – steif]
- Verpackung abhängig vom Reibwert [$\mu=0,18$ - $\mu=0,31$]
- Verpackung abhängig vom Material [Kartonage – Kunststoff]
- Form der Sendung [flach – ausgefüllt]

4.1.2 Anlagenparameter

Die entscheidenden Faktoren der Anlage ergeben sich, wie auch jene des Stückgutes aus der Betrachtung des Ishikawa Diagrammes. Auch hierbei werden den Parametern Grenzen¹⁷ zugewiesen, welche sich aus Versuchen im Vorfeld ergeben.

Für den FFS werden die folgenden Grenzen definiert:

- Stoßwinkel [29°-124°]
- Rutschenwinkel [31°- 47°]

Ist bei der Zuweisung eines Faktors unklar, ob er zum Stückgut oder der Anlage gehört, so ist die Zuordnung frei wählbar, da die Unterscheidung keinen direkten Einfluss auf das weitere Vorgehen hat, sondern nur der Übersichtlichkeit dient.

¹⁷Definition der Grenzen für die Anlage wird im Kapitel 3.2.2 – Anlage vorbereiten, ausführlicher erklärt.

4.2 Anwendung Schritt 2 - Parameterscreening-Versuchsplan erzeugen

In diesem Schritt wird die Erzeugung des Parameterscreening Versuchsplans des FFS durchgeführt. Dazu wird als erstes ein Versuchsplan mit den zuvor definierten Grenzen aus Schritt 1 erstellt. Danach wird die Vorbereitung der Anlage sowie der verwendeten Proben demonstriert. Weiters beinhaltet der Schritt 2 die Vorbereitung der Messgrößenerfassung, die Erstellung einer Versuchsvorschrift und die Ermittlung des Stichprobenumfangs.

4.2.1 Versuchsplan erstellen

Um den Versuchsplan für das Parameterscreening erstellen zu können, müssen den in Schritt 1 definierten Faktoren Werte für die Ober- und Untergrenze zugewiesen werden.

Bei den Faktoren Gesamtgewicht, Reibwert, Stoßwinkel und Rutschenwinkel können die jeweiligen Zahlenwerte der Grenzen für den Versuchsplan verwendet werden. Für den Polybag ergeben sich jedoch durch die Faktoren Biegeschlaffheit, Form und Material weitere Stufen, welche nicht direkt durch eine physikalische Größe beschrieben werden können.

Daher wird wie in Tabelle 4.1 ersichtlich, eine Codierung eingeführt. Den Eigenschaften werden dabei die Zahl 1 (untere Grenze) oder 2 (obere Grenze) zugewiesen. Dadurch ist es auch möglich subjektive Merkmale in den Versuchsplan aufzunehmen. Hierbei bedarf es jedoch einer klaren Vorschrift wie das Merkmal definiert ist, da es ansonsten zu falschen Interpretationen führen kann. Die Bezeichnung biegeschlaff definiert hierbei das Verhalten des Polybags, durch welches ein plastischer Stoß entsteht. Steif hingegen wird für die Bezeichnung von Materialien mit einem teilweise elastischen bis vollelastischen Stoßverhalten verwendet. Das Verhalten von biegeschlaffen und steifen Proben wird in Abbildung 4.8 dargestellt.

Die Kennzeichnung einer flachen Form wird durch die projizierte Fläche des Polybags definiert. Dazu wird ein Verhältnis der Oberfläche durch die projizierte Anströmfläche eingeführt. Hierbei wird die Form als flach bezeichnet, wenn dieses Verhältnis einen Wert kleiner als 30 aufweist. Werte größer 30 ergeben wiederum eine Form, welche als ausgefüllt definiert wird. Dargestellt wird dies in Abbildung 4.10.

Die Zuweisung der Materialien erfolgt nach ihrer Zusammensetzung, wobei zwischen Kartonage und Kunststoff unterschieden wird (siehe Abbildung 4.7).

Tabelle 4.1.: Codierung der Faktoren

Parameter	Codierung	
	1	2
Biegeschlaffheit	schlaff	steif
Form	flach	ausgefüllt
Material	Kartonage	Kunststoff

In Tabelle 4.2 wird der Aufbau des daraus entstehenden Versuchsplans für das Parameterscreening dargestellt. Hierbei werden die Zeilen nach dem Gray-Code Schema¹⁸ so lange vervollständigt bis alle Faktoreinstellungen miteinander kombiniert worden sind.

Für die zu untersuchenden Faktoren wird angenommen, dass die Effekte sich linear auf die Zielgröße auswirken. Daher genügt es das Screening mit einem zweistufigen Versuchsplan durchzuführen. Wobei die Stufeneinstellungen die Ober- und Untergrenze des jeweiligen Faktors sind.

Tabelle 4.2.: Vollfaktorieller Versuchsplan für das Parameterscreening

Versuchs Nr.	Faktoren ¹⁹						
	Gewicht	Reibwert	Material	Form	Rutschenwinkel	Stoßwinkel	Biegeschlaffheit
1	20	0,18	1	1	31	29	1
2	220	0,18	1	1	31	29	1
3	20	0,31	1	1	31	29	1
4	220	0,31	1	1	31	29	1
5	20	0,18	2	1	31	29	1
6	220	0,18	2	1	31	29	1
7	20	0,31	2	1	31	29	1
8	220	0,31	2	1	31	29	1
9	20	0,18	1	2	31	29	1
10	220	0,18	1	2	31	29	1
11	20	0,31	1	2	31	29	.
12	220	0,31	1	2	31	.	.
13	20	0,18	2	2	.	.	.
14	220	0,18	2
15	20	0,31
16	220

¹⁸ Das Ausfüllen des Versuchsplans nach dem Gray-Code Schema ist in Tabelle 2.1 ersichtlich.

¹⁹ Die Faktorstufen werden für den Polybag im Anwendungsfall des FFS nach [SL21] definiert. Die Herkunft der Anlagenparameter liegt in Versuchen, welche in Kapitel 3.2.2 erläutert werden.

4.2.2 Anlage vorbereiten

Bei der Vorbereitung der Anlage werden die beiden Faktoren Rutschenwinkel und Stoßwinkel variiert. Die Rutschenlänge und das Material der Rutsche sind zwei Parameter, welche mit Sicherheit einen Einfluss auf die Durchlaufzeit haben. Jedoch werden diese als konstant angenommen, da von einer genormten Rutschenlänge und einem gleichbleibenden Material ausgegangen wird.

In den folgenden Absätzen wird dargestellt, wie die Variation der Anlagenparameter am Versuchsaufbau definiert sind.

Rutschenwinkel

Die Grenzen für den Rutschenwinkel ergeben sich aus Versuchen im Vorfeld. Die Untergrenze ergibt sich mit 28° als Losreißwinke, welcher durch Versuche ermittelt worden ist. Um ein Haften der Polybags auszuschließen, fällt daher die Wahl der Untergrenze auf 30° . Die Obergrenze liegt bei 50° Steigung. Steilere Stellungen führen zu einem Abheben des Polybags, wodurch er aus der Anlage fällt. Bei der Stufeneinstellung treten Einschränkungen durch die Bauform des Prüfstandes auf. Somit können die Grenzen nicht exakt, sondern nur mit 31° - Untergrenze (siehe Abbildung 4.3 links) und 47° - Obergrenze (siehe Abbildung 4.3 rechts) angenähert werden.

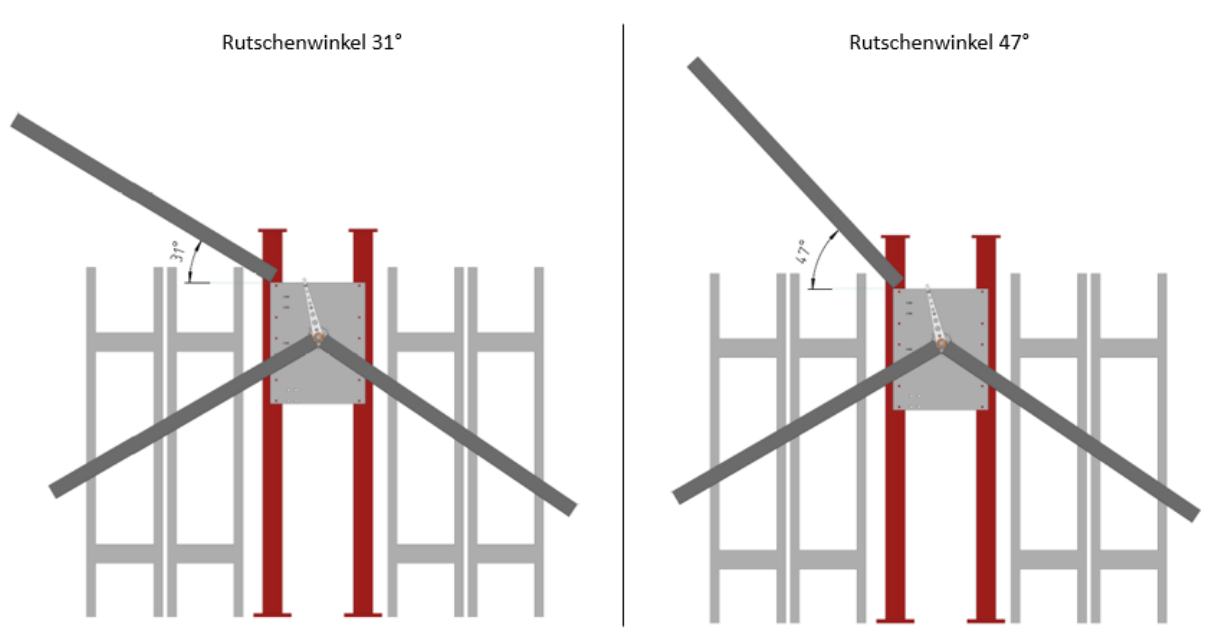


Abbildung 4.3: Grenzen für den Rutschenwinkel

Stoßwinkel

Der Winkel zwischen der oberen Rutsche und der Umlenkklappe wird als Stoßwinkel bezeichnet. Hierbei variiert der Winkel wie in Abbildung 4.4 dargestellt, zwischen 29° und 124° , wobei der Stoß mit 29° als steiler Stoß (siehe links Abbildung 4.4) und der Fall mit 124° als flacher Stoß (siehe rechts Abbildung 4.4) bezeichnet wird. Um den Winkel verändern zu können, befindet sich auf der Rückseite der Anlage ein Teilungsapparat, mit dem der Winkel der Klappe eingestellt werden kann.

Da der Stoßwinkel zwischen der Rutsche und der Klappe gemessen wird, muss die Klappenstellung an die Rutschenneigung angepasst werden. Daraus ergibt sich der flache Stoß aus einer Kombination der Rutsche mit 31° und der steile Stoß durch eine Rutschenstellung von 47° . Der flache Stoß wird, wie auch der flache Rutschenwinkel, durch ein Haften des Polybags begrenzt. Hierbei kommt als weiterer Gesichtspunkt hinzu, dass der Polybag bei noch flacheren Klappeneinstellungen nicht mehr umgelenkt wird, sondern über die Klappe hinweg in die falsche Richtung gleitet. Der steile Stoß wird durch die Polybag Größe limitiert. Durch die steile Klappenstellung wird der Platz zwischen Rutsche und Klappe so stark verringert, dass der benötigte Platz für den Polybag nicht mehr gegeben ist.

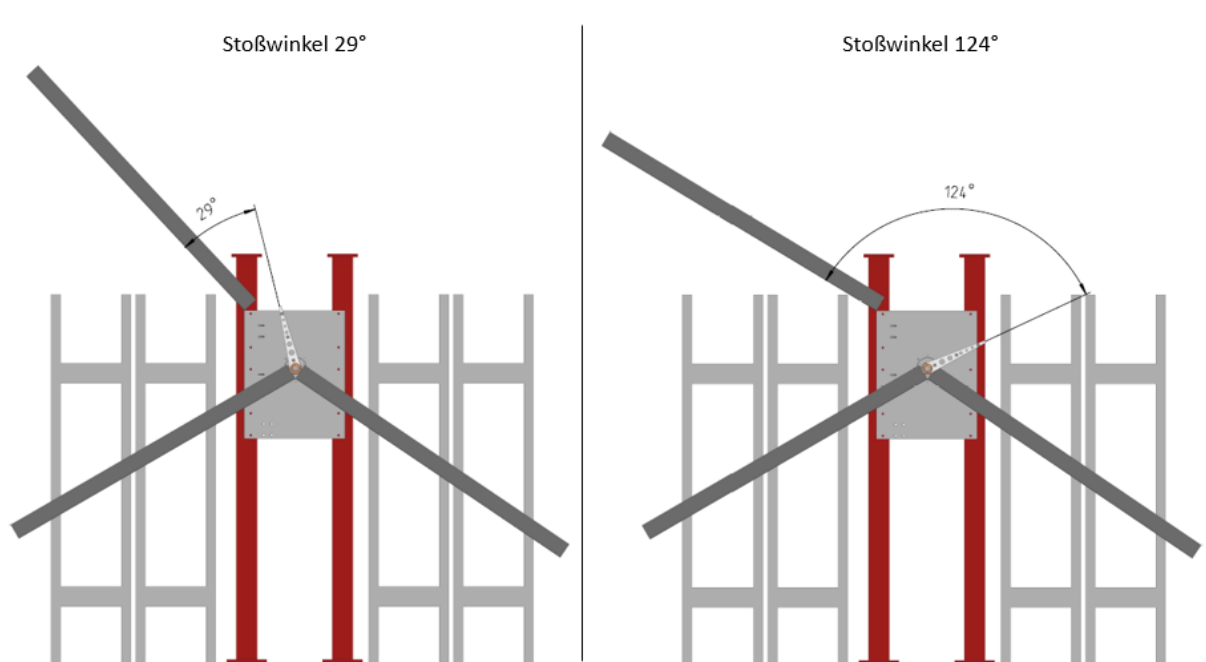


Abbildung 4.4.: Grenzen für den Stoßwinkel

4.2.3 Proben vorbereiten

Für den Polybag als Stückgut wird nachfolgend gezeigt, wie die Vorbereitung der Proben durchgeführt wird. Dabei wird jeder Faktor getrennt betrachtet. Durch die Vorbereitung wird gezeigt, wie die Proben an die jeweiligen Stufeneinstellungen angepasst werden. Angelehnt an die Ermittlung des Stichprobenumfangs aus Kapitel 4.2.6 müssen je Stufeneinstellung sechs Proben vorbereitet werden.

Gewicht

Bei der Befüllung der Polybags wird darauf geachtet, dass das Gewicht als Beladung inklusive Verpackung definiert ist. Dies ist erforderlich, da der Verwendungsbereich von Polybags zwischen 20g und 220g liegt, wobei das Gewicht der Verpackung alleine schon zwischen 10g und 80g schwanken kann [SSK+21]. Damit der Einfluss des Verpackungsgewichtes für alle Versuche konstant bleibt, werden die Proben wie in Abbildung 4.5 so modifiziert, dass die leeren Proben ein Gewicht von 20g besitzen. Somit können die leeren Proben für die Testung der unteren Stufe verwendet werden. Um die Obergrenze zu testen, werden die Proben zusätzlich mit 200g Kieselsteinen befüllt, wodurch die Probe ein Gesamtgewicht von 220g aufweist. Damit die Verteilung des Gewichtes keinen Einfluss auf die Ergebnisse hat, wird die Beladung nicht lose in den Polybag geleert, sondern mit einem zweiten Beutel im Zentrum des Polybags fixiert. Ansonsten könnten Bewegungen des Inhalts die Versuche verfälschen.



Abbildung 4.5.: Standardisieren des Verpackungsgewichtes auf 20g

Material – Gleitreibungskoeffizient μ

Für die Unterscheidung des Materials werden zwei Betrachtungen²⁰ herangezogen. Als erstes Betrachtungsmerkmal wird der Gleitreibungskoeffizient verwendet. Hierbei liegen die Grenzen für den Reibwert bei $\mu=0,18$ bis $\mu=0,31$, wodurch der Einsatzbereich, der am häufigsten verwendeten Verpackungen, laut [SSK+21] abgedeckt wird. Dazu werden zwei Kunststoffproben (siehe Abbildung 4.6) verwendet.

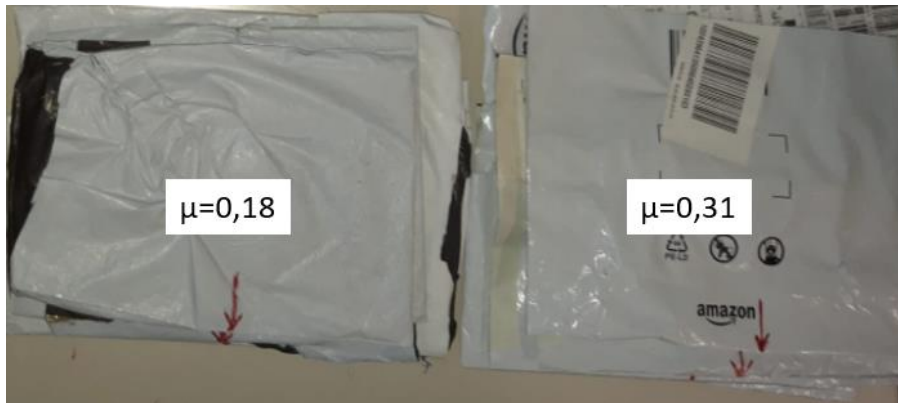


Abbildung 4.6: Kunststoffproben mit verschiedenen Reibwerten

Material – Reibwert unabhängige Betrachtung

Um das Material nicht nur auf den Reibwert zu reduzieren, bietet die zweite Betrachtung einen Vergleich von zwei Polybags, welche beide denselben Reibwert besitzen, aber aus unterschiedlichen Materialien bestehen. In Abbildung 4.6 sind beide Proben abgebildet. Links die Probe aus Kartonage und auf der rechten Seite eine Probe aus Kunststoff.



Abbildung 4.7: Proben mit selbem Reibwerten, aber aus verschiedenen Materialien

²⁰ Betrachtung 1 – Material wird durch den Gleitreibungskoeffizienten ($\mu=0,18$ - $\mu=0,31$) definiert.

Betrachtung 2 – Material wird durch seine Zusammensetzung (Kartonage – Plastik) definiert.

Biegeschlaffheit

Da das Stoßverhalten bei der Umlenkung abhängig von der Art des Stoßes ist, werden dazu Sendungen mit denselben Eigenschaften so modifiziert, dass sie einmal ein elastisches und einmal ein plastisches Stoßverhalten aufweisen. Ohne Bearbeitung sind die verwendeten Materialien biegeschlaff. Durch Versteifungen auf der Innenseite der Proben durch Kartonage wird das Verhalten eines teilweise elastisches Stoßverhalten angenähert. Diese Modifikation wird benötigt, um eine Kombination aus demselben Materialverhalten mit verschiedenen Stoßeigenschaften zu ermöglichen. In Abbildung 4.8 sieht man wie sich die Probe biegeschlaff (vorne) und im versteiften Zustand (hinten) verhält.



Abbildung 4.8: Darstellung wie aus einer biegeschlaffen Probe eine steife Probe entsteht

Form

Durch die Eigenschaft des Polybags sich an die Form des Inhaltes anzupassen, kann er bei selbem Gewicht eine sehr geringe oder auch große Oberfläche aufweisen. Diese Oberfläche resultiert in einer vergrößerten angeströmten Querschnittsfläche, welche einen Einfluss auf das strömungsdynamische Verhalten haben kann. Alle Effekte²¹, welche durch die Anströmung entstehen, werden in dem Faktor „Form“ gebündelt. Damit dieser Einfluss sichtbar gemacht werden kann, werden Polybags mit denselben physikalischen Eigenschaften durch die Befüllung so angepasst, dass alle Parameter mit kleinem und großem projiziertem Querschnitt getestet

²¹ Als Effekte der Anströmung zählen der Auftrieb, der Luftwiderstand und der Anpressdruck durch die Strömung über dem Polybag.

werden können. Um das Gesamtgewicht des Polybags nicht zu beeinflussen, geschieht die Anpassung durch Befüllung mit aufgeblasenen Luftballons, deren Gewicht geringer als ein Gramm ist (siehe Abbildung 4.9).



Abbildung 4.9.: Befüllung der Proben für die Faktorstufe mit großem projizierten Querschnitt

Als Vergleich sieht man in Abbildung 4.10 zwei Proben, welche beide mit 200g Steinen befüllt sind. Zusätzlich zu den Steinen wurde die linke Probe noch mit Luftballons beladen. Damit vergrößert sich der projizierte Querschnitt der Probe, ohne das Gewicht signifikant zu verändern. Bei solchen Modifizierungen ist darauf zu achten, dass die Beladung nicht den Versuch verfälscht. Daher müssen auch die Luftballons in der Mitte der Probe fixiert werden. Würden diese an der Vorderkante der Probe platziert sein, so würde sich wiederum das Stoßverhalten verändern, wodurch die Aussagen über den Einfluss der Biegeschlaffheit verfälscht werden würden.



Abbildung 4.10: Veränderung des projizierten Querschnittes bei selbem Gewicht

4.2.4 Messgrößenerfassung standardisieren

Der entsprechende messtechnische Aufbau für die Durchlaufzeitmessung am FFS wird in Abbildung 4.9 dargestellt. Darin werden die Positionen der Lichtschranken gekennzeichnet und die Durchlaufzeiten der Submodelle (exemplarisch für RO) definiert.

Die Messung der Durchlaufzeit erfolgt durch Lichtschranken, welche jeweils am Anfang und Ende jeder Rutsche montiert sind. Die Durchlaufzeit der Umlenkung ergibt sich entsprechend aus der Differenz der Gesamtzeit und den Zeiten der Rutschen.

Die Messdaten der Lichtschranken werden an ein für den FFS speziell erstelltes B&R Programm übergeben, welches aus den Sensordaten die jeweiligen Durchlaufzeiten für die Messpunkte auswertet.

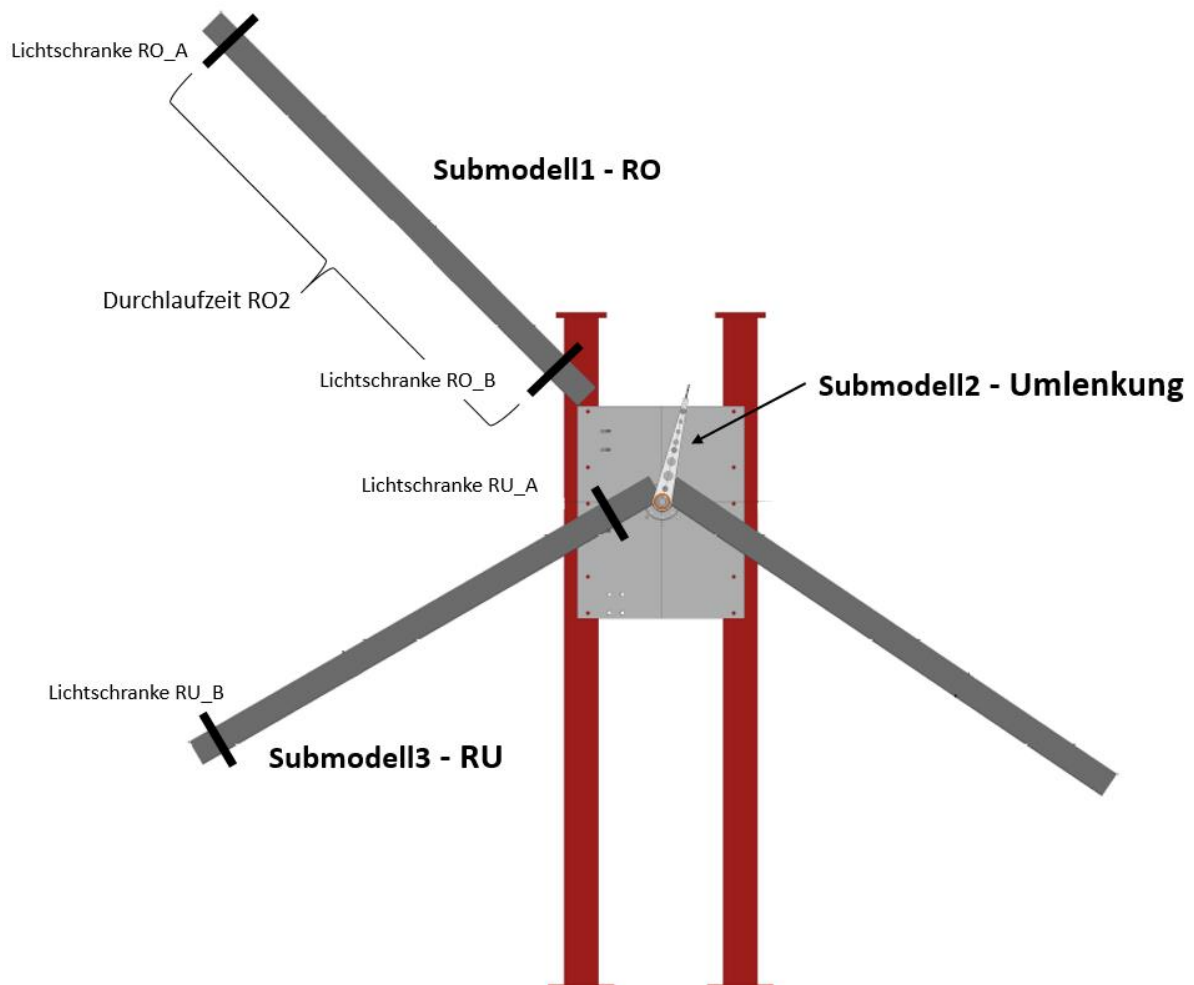


Abbildung 4.11.: Kennzeichnung der Messvorrichtungen am FFS

4.Exemplarische Anwendung der Methodik

Abbildung 4.12 zeigt die Benutzeroberfläche dieses Programmes. Darin sieht man eine Darstellung des Prüfstandes, wobei alle verbauten Sensoren mit einem Ampelsymbol angedeutet sind. Die grüne Farbe symbolisiert, dass der jeweilige Sensor ausgelöst wurde.

Auf der oberen Seite des Interfaces sind alle vom Programm ermittelten Messdaten angeführt. Wie in der dargestellten Tabelle werden die Messdaten in einer CSV-Datei abgespeichert. Um Fehlversuche nicht in die Datenerfassung aufzunehmen, werden vom Programm nur Daten gespeichert, welche den gesamten Versuchsaufbau durchlaufen. Nachdem alle Sensoren ausgelöst wurden, erscheint auf der rechten Seite im unteren Drittel der Skizze ein Feld, welches die Anzahl der gültigen Versuche angibt. Auf der linken Seite der Benutzeroberfläche sind weitere Felder für Kommentare und für die Laufnummer des Versuches.

System Neustart Reset **USB Stick gefunden** Username: Passwort: Login

PB-NR	Gesamtzeit RO_2 RU_1 [ms]	Gesamtzeit RO_2 RU_2 [ms]	Zeit RO_2 [ms]	Zeit RU_1 [ms]
1	0	1452	416	0

Zeit RU_2 [ms]	Zeit Umlenkung [ms]	Geschw. RO_2_O [m/s]	Geschw. RO_2_U [m/s]
939	97	3.000000	4.583333

-	Geschw. RU_1_U [m/s]	Geschw. RU_2_O [m/s]	Geschw. RU_2_U [m/s]
-	0.000000	45.000000	2.750000

Distanz LR RO_2_O [m]:

Distanz LR RO_2_U [m]:

Distanz LR RU_1_U [m]:

Distanz LR RU_2_O [m]:

Distanz LR RU_2_U [m]:

Rutsche frei?

1

Abbildung 4.12.: Benutzeroberfläche nach einem gültigen Versuch

4.2.5 Versuchsvorschrift erstellen

Die Versuchsvorschrift für die Versuchsdurchführung am FFS wird wie folgt definiert:

Vor der Versuchsdurchführung müssen alle zuvor erläuterten Vorbereitungen²² eingehalten werden. Wenn alle Proben und die Anlage wie gefordert eingestellt sind, müssen die Sensoren noch einmal durch ein Teach-In neu kalibriert werden. Vor dem Einlegen der Probe, wird die Laufnummer (PB-Nummer) in das Kommentarfeld in der Benutzeroberfläche eingegeben. Dadurch können die Ergebnisse bei der Auswertung der dazugehörigen Stufenvariation im Versuchsplan zugewiesen werden. Die Laufnummer für das Parameterscreening des FFS baut sich aus den Zahlen 1 und 2 wie in Tabelle 4.3 ersichtlich auf. Daraus entsteht die PB-Nummer: Gewicht-Biegeschlaffheit-Form-Material-Rutschenwinkel-Stoßwinkel²³.

Tabelle 4.3.: Codierung der Parameter

Parameter	Codierung	
	1	2
Gewicht	20g	220g
Biegeschlaffheit	schlaff	steif
Form	flach	ausgefüllt
Material	Kartonage	Kunststoff
Rutschenwinkel	31°	47°
Stoßwinkel	flach	steil

Diese Codierung wird auch für die Versuche des Teachingsets verwendet. Da es nach dem Parameterscreening Faktoren geben kann, welche nicht weiter in den Versuchsplan für das Teachingset aufgenommen werden müssen, können diese Faktoren aus der Codierung gelöscht werden. Der Aufbau der verbleibenden Nummer bleibt danach jedoch gleich angeordnet.

Nachdem alle Sensoren bereit sind und die PB-Nummer in das Benutzerinterface eingetragen ist, können die Versuche durchgeführt werden. Die Proben dürfen dabei nicht willkürlich in die Anlage eingelegt werden. Da die Einschleusung per Hand erfolgt, werden alle Proben an der markierten Stelle (siehe Abbildung 4.13) aus der Ruhe losgelassen²⁴.

Neben der Vorgabe des Startpunktes muss die Probe so platziert werden, dass sie vor dem Loslassen plan auf der Rutschenoberfläche aufliegt. Da die Proben modifiziert wurden, müssen die Proben mit dem roten Pfeil (siehe Abbildung 4.13) voraus losgelassen werden, damit der Effekt

²² Anlage vorbereiten, Proben vorbereiten und Messgrößenerfassung standardisieren

²³ Als Beispiel für ein schlaffes, mit 220g ausgefülltes Paket, welches aus Kartonage besteht und bei einem Rutschenwinkel von 31° bei einem steilen Stoß getestet wird, ergibt sich die PB-Nummer: 212112.

²⁴ Damit ist als Anfangsbedingung mit einer Anfangsgeschwindigkeit=0 klar definiert.

des Stoßverhaltens nachgebildet werden kann. Nachdem die Probe alle Sensoren ausgelöst hat, kann die nächste Probe nach demselben Vorgehen geprüft werden.

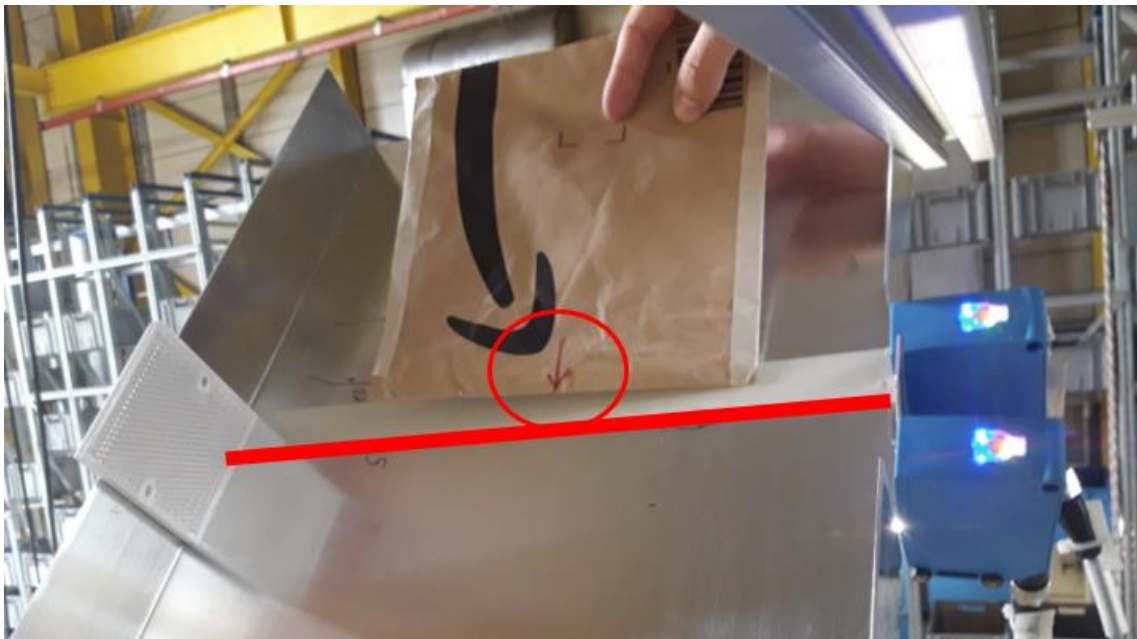


Abbildung 4.13.: Einschleusung der Proben per Hand

4.2.6 Stichprobenumfang ermitteln

Um die statistische Schwankung der Messungen zu minimieren, wird für die Durchlaufzeitmessung der Polybags der benötigte Stichprobenumfang berechnet. Um die t-Verteilung anzuwenden, muss zuerst sichergestellt werden, dass die Versuche einer Normalverteilung unterliegen. Dies wird durch einen Kolmogorov-Smirnov Test überprüft. Dazu werden exemplarisch die Durchlaufzeiten der Rutsche oben (RO) überprüft. Die Berechnung für die übrigen Submodelle (Umlenkung und Rutsche unten (RU)) wird im Anhang A2 dargestellt.

Kolmogorov-Smirnov Test

In Abbildung 4.14 wird exemplarisch dargestellt, wie die Auswertung für die Durchlaufzeit der Rutsche oben (RO) aufgebaut ist. Hierbei werden nur die Ergebnisse der Auswertung abgebildet. Die dazugehörige Tabelle mit den Messdaten, für welche der Test durchgeführt wurde, ist im Anhang A2 zu finden.

Bei den untersuchten Messdaten ergibt sich wie in Abbildung 4.14 ersichtlich, dass die Nullhypothese des Tests nicht verworfen werden kann. Daraus lässt sich eine Normalverteilung der Messdaten ableiten.

Anzahl [#]
35
Mittelwert [ms]
1079,771429
Standardabweichung [ms]
27,11097641
Max. Abweichung = Teststatistik
0,061487356
Kritischer Wert für $\alpha=0,05$
0,22424
Auswertung
Teststatistik < kritischer Wert
Nullhypothese kann nicht verworfen werden - Normalverteilung liegt vor

Abbildung 4.14.: Ergebnisse des Kolmogorov-Smirnov Tests für die Durchlaufzeit Rutsche oben (RO)

t-Verteilung

Nachdem nachgewiesen ist, dass die Versuche normalverteilt sind, erfolgt die Ermittlung der benötigten Stichprobenanzahl. Die maximal zulässige Abweichung der gemessenen Durchlaufzeit für den FFS liegt bei 5%. Um die benötigte Stichprobengröße zu finden, bei der die Schwankung des Messwertes innerhalb dieser Grenzen liegt, wird die Berechnung für einen Bereich von 2 – 10 Stichproben durchgeführt. Die Ergebnisse der Abweichungen für die jeweilige Stichprobengröße werden in Tabelle 4.4 dargestellt. Darin wird für die Mittelwerte der Durchlaufzeiten der Rutsche oben (RO2) und unten (RU2) die Ober- und Untergrenze für die jeweilige Stichprobenanzahl ermittelt. Über die Grenzen und den Mittelwert kann auf die relative Abweichung zurückgeschlossen werden. Die Berechnung startet bei einer Stichprobengröße von zwei und wird wiederholt, bis die geforderte Abweichung vorliegt.

Tabelle 4.4.: Berechnung der Abweichung der Durchlaufzeiten je Stichprobengröße

2-Stichproben		4-Stichproben		6-Stichproben		8-Stichproben		10-Stichproben	
RO2	RU2	RO2	RU2	RO2	RU2	RO2	RU2	RO2	RU2
Mittelwert [ms]		Mittelwert [ms]		Mittelwert [ms]		Mittelwert [ms]		Mittelwert [ms]	
1074,00	912,00	1080,00	913,00	1078,67	927,33	1075,00	910,00	1071,6	910,8
Standardabw.[ms]		Standardabw.[ms]		Standardabw.[ms]		Standardabw.[ms]		Standardabw.[ms]	
42,43	96,17	29,39	57,45	25,00	49,73	22,60	53,15	22,19	52,50
Obergrenze [ms]		Obergrenze [ms]		Obergrenze [ms]		Obergrenze [ms]		Obergrenze [ms]	
1129,20	1037,12	1097,64	947,47	1090,00	949,87	1083,63	930,30	1080,07	928,40
Untergrenze [ms]		Untergrenze [ms]		Untergrenze [ms]		Untergrenze [ms]		Untergrenze [ms]	
1018,80	786,88	1057,60	878,53	1067,34	904,80	1066,37	889,70	1060,19	883,80
Abweichung		Abweichung		Abweichung		Abweichung		Abweichung	
3,95%	10,54%	2,72%	6,29%	2,32%	5,36%	2,10%	5,84%	2,07%	5,76%

In Abbildung 4.15 ist der Verlauf der relativen Abweichung über die Stichprobenanzahl abgebildet. Hierbei lässt sich erkennen, dass ab einer Anzahl von sechs Stichproben die Abweichung nichtmehr nennenswert reduziert wird. Durch Gegenüberstellung des Aufwandes, welcher ein höherer Stichprobenumfang mit sich bringt, fällt die Wahl auf sechs Stichproben für die Versuche des Parameterscreenings sowie für die Generierung des Teachingsets. Daraus ergeben sich 768 Versuche²⁵ für das Parameterscreening. Eine Erhöhung der Stichprobenanzahl auf acht würde 1024 Versuche²⁶ benötigen, wobei der Aufwand um rund ein Viertel steigen würde, die Genauigkeit hingegen sich nur noch im Promille Bereich verbessert.

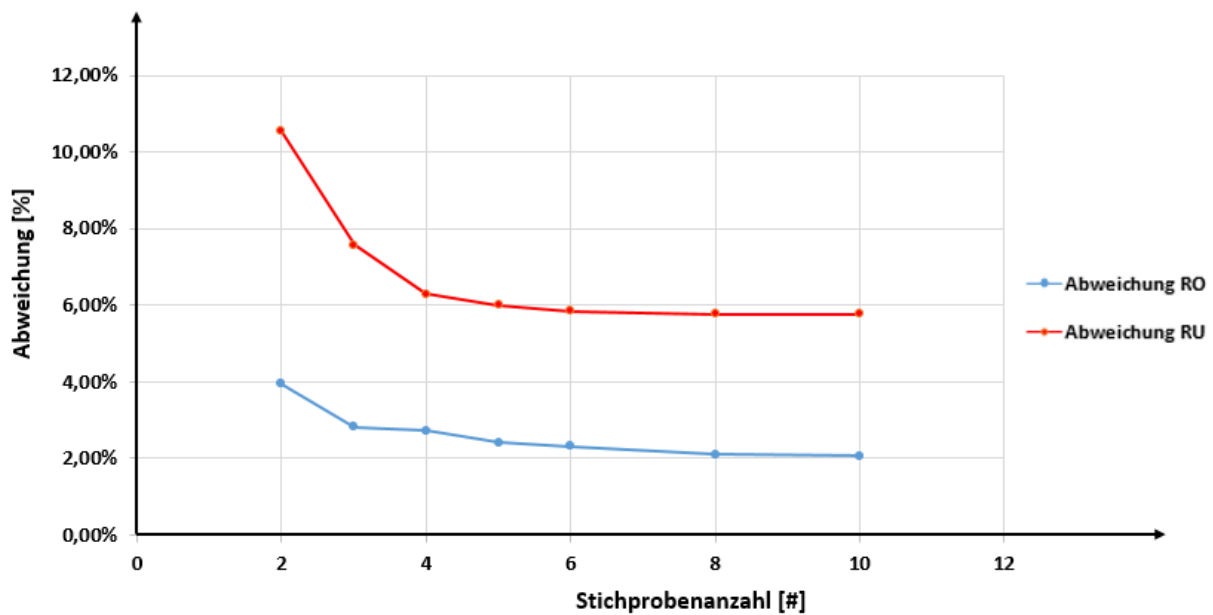


Abbildung 4.15.: Darstellung der Abweichung der Durchlaufzeit je Stichprobenanzahl

4.3 Anwendung Schritt 3 – Versuche für das Parameterscreening durchführen

In Schritt 3 werden alle Versuche des im Schritt 2 definierten Versuchsplans durchgeführt. Dabei muss die in Schritt 2 definierte Vorschrift bei der Versuchsdurchführung eingehalten werden. Die Versuche sollten dabei nicht willkürlich abgearbeitet werden, sondern es sollte zuvor die Versuchsdurchführung so geplant werden, dass der Aufwand für den Umbau der Anlage möglichst gering bleibt. Dadurch wird nicht nur der Aufwand minimiert, sondern auch Fehlerquellen, welche durch das Umrüsten entstehen können. Die Messdaten, welche sich durch die Versuche ergeben, dienen als Grundlage für das Parameterscreening und die Korrelationsanalyse in Schritt 4.

²⁵ Laut Formel (2.1) Stichprobenanzahl=6; Stufenanzahl s=2; Faktoren k=7 → Versuchsanzahl=6 * 7² = 768

²⁶ Laut Formel (2.1) Stichprobenanzahl=8; Stufenanzahl s=2; Faktoren k=7 → Versuchsanzahl=8 * 7² = 1024

4.4 Anwendung Schritt 4 - Einflüsse der Parameter analysieren

Für die Untersuchung der Ergebnisse der Versuche am FFS wird das mathematische Computerprogramm Matlab verwendet. Dieses eignet sich besonders für diese Aufgabenstellung, da es für numerische Berechnungen auf Matrizenbasis ausgelegt ist. Um den Programmcode übersichtlich zu gestalten, werden für das Parameterscreening sowie für die Korrelationsanalyse vordefinierte Funktionen²⁷ verwendet. Ein weiterer Vorteil des Programmes liegt in der Erweiterung für Datenanalyse, womit im Schritt 7 die Regressionsmodelle erzeugt werden.

Bevor die Versuchsdaten des Parameterscreenings an das Programm übergeben werden können, muss der Versuchsplan aus Schritt 4.2 um die Ergebnisse (Durchlaufzeiten) der einzelnen Submodelle (siehe Abbildung 4.11) und der gesamten Durchlaufzeit ergänzt werden. Wichtig ist es hierbei, dass alle Versuche exakt nach dem erstellten Versuchsplan durchgeführt werden, da das Auslassen eines Versuches eine Verschiebung der Ergebnisse für alle weiteren Faktoreinstellungen mit sich bringt.

Im folgenden Unterpunkt wird das Parameterscreening und die Korrelationsanalyse der Subsysteme Rutsche oben (RO), Umlenkung, Rutsche unten (RU) und der gesamten Anlage durchgeführt. Als zu untersuchende Größe gilt dabei die jeweilige Durchlaufzeit. Der Matlab Programmcode dazu befindet sich im Anhang B.

Wie in Tabelle 2.3 ersichtlich, werden Parameter mit einem Korrelationskoeffizienten $R > 0,8$ als stark korrelierend bezeichnet. Parameter mit einem Wert von $R 0,5 - 0,8$ weisen eine mittlere Korrelation auf. Und alle Parameter mit einem Korrelationskoeffizienten $R < 0,5$ werden für die weiteren Untersuchungen vernachlässigt, da sie nur eine sehr geringe Korrelation aufweisen.

Parameterscreening und Korrelationsanalyse

Als erstes wird der Vergleich zwischen einer Kunststoffprobe und einer Probe aus Kartonage mit demselben Gleitreibungskoeffizienten durchgeführt. Dabei wird ermittelt, ob es genügt die Unterscheidung der Materialien aufgrund des Reibwertes zu treffen. Dazu wird der Effekt auf die Durchlaufzeit Rutsche oben (RO) betrachtet. Zusätzlich zur graphischen Darstellung des Effektes wird über jedem Effektdiagramm der dazugehörige Korrelationskoeffizient R angegeben.

²⁷ Durch die Funktion „maineffectsplot“ wird das Parameterscreening durchgeführt und als Ergebnis ein Effektdiagramm (siehe Abbildung 4.17) ausgegeben. Der Korrelationskoeffizient R zweier Vektoren wird durch die „corr“ Funktion aufgerufen.

4.Exemplarische Anwendung der Methodik

In Abbildung 4.16 wird ersichtlich, dass sich kein nennenswerter Unterschied der Durchlaufzeit der Rutsche oben (RO) zwischen Kunststoff und Kartonage ergibt. Daher genügt für die weiteren Betrachtungen die Untersuchung des Reibwertes, stellvertretend für die Materialeigenschaften.

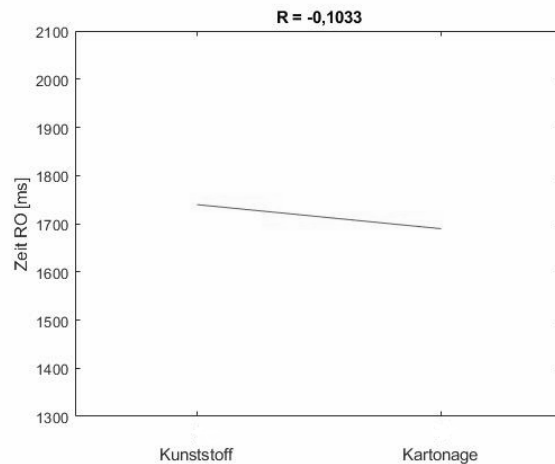


Abbildung 4.16.: Unterscheidung zweier Proben mit demselben Reibungskoeffizienten, aber unterschiedlicher Materialzusammensetzung

Das Parameter Screening angewandt für die Gesamtzeit, mit der Umlenkung steiler Stoß ergibt ein Effektdiagramm wie in Abbildung 4.17 dargestellt.

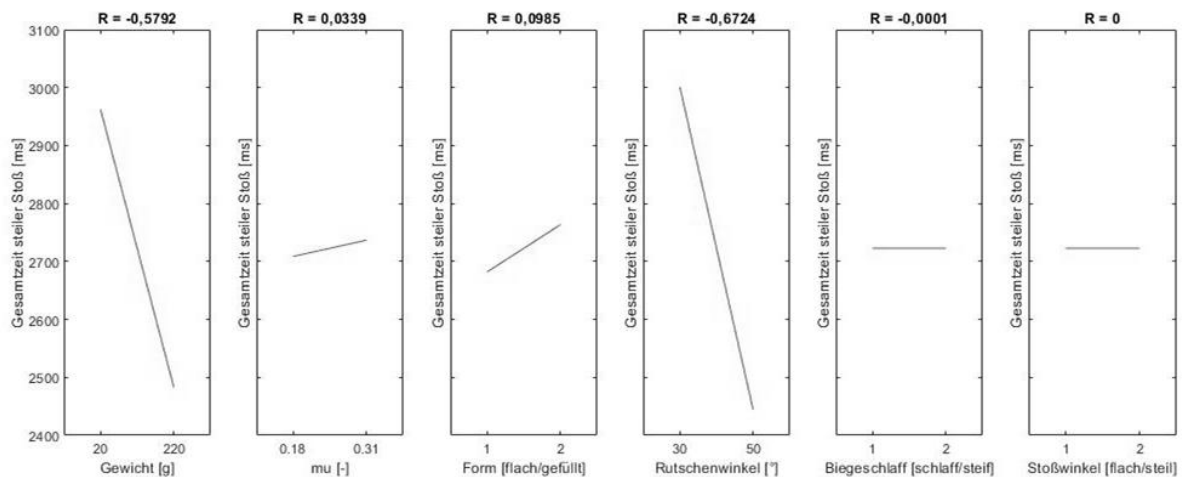


Abbildung 4.17.: Effekte auf die Gesamtzeit bei einem steilen Stoß

Abbildung 4.18 zeigt wiederum den Einfluss der Faktoren auf die Gesamtzeit für den Fall eines flachen Stoßwinkels.

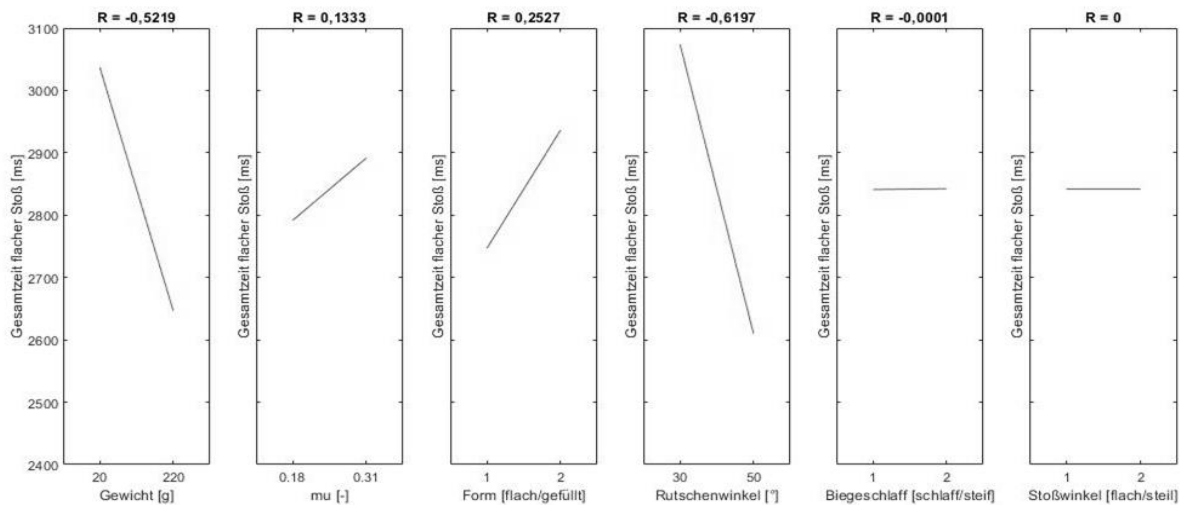


Abbildung 4.18.: Effekte auf die Gesamtzeit bei einem flachen Stoß

Bei der Betrachtung der Gesamtzeiten für den steilen Stoß (siehe Abbildung 4.17) und flachen Stoß (siehe Abbildung 4.18) ergibt sich wie zu erwarten ein starker Einfluss des Rutschenwinkels und des Probengewichts. Hingegen ist der Einfluss des Gleitreibungskoeffizienten gering ausgefallen. Unter der Annahme des Reibgesetzes nach Coulomb wurde im Vorfeld ein größerer Effekt erwartet. Die Effekte des Stoßwinkels sowie der Biegeschlaffheit weisen in der Betrachtung der Gesamtzeit keinerlei Effekte auf. Um diese Effekte sichtbar zu machen, wird in Abbildung 4.19 gesondert der Umlenkvorgang betrachtet. Diese Betrachtung wird benötigt, da die Gesamtzeit rund zehnmal²⁸ größer ist als die Dauer des Umlenkvorganges. Daher werden die Auswirkungen der Parameter in der Betrachtung der Gesamtzeit nicht ersichtlich.

In Abbildung 4.20 werden dazu die Einflüsse der Parameter bezogen auf die Dauer des Umlenkvorganges dargestellt. Dabei ergibt sich eine starke Abhängigkeit des Umlenkvorganges vom Stoßwinkel. Wie bereits aus der Betrachtung der Gesamtzeit ergibt sich wiederum eine starke Abhängigkeit vom Sendungsgewicht und dem Rutschenwinkel. Parameter wie Form, Reibwert und Biegeschlaffheit weisen hingegen keine nennenswerten Effekte auf den Umlenkvorgang auf.

²⁸ Die Gesamtzeit beträgt rund 1700 – 2400[ms]. Der Umlenkvorgang hingegen benötigt im Schnitt 90-240[ms].

4.Exemplarische Anwendung der Methodik

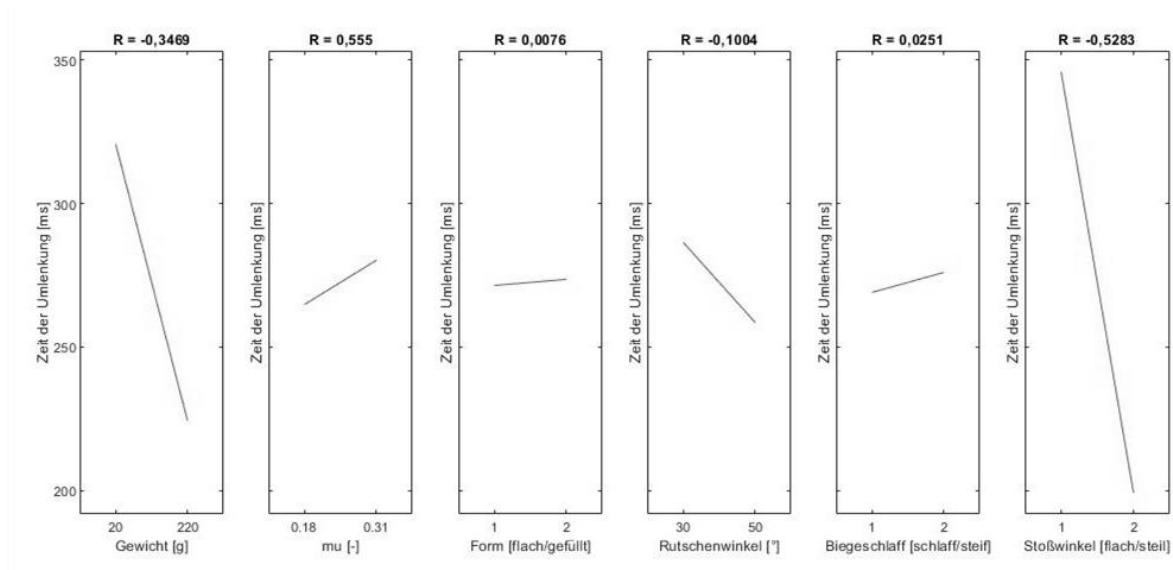


Abbildung 4.19.: Effekte auf den Umlenkvorgang

In Abbildung 4.20 wird der Effekt der Parameter auf die Einschleusung (Rutsche oben (RO)) untersucht. Daraus lässt sich der Einfluss der Parameter gesondert auf die Rutschen abbilden. Dabei stehen wiederum die Effekte des Sendungsgewichtes und des Rutschenwinkels im Vordergrund.

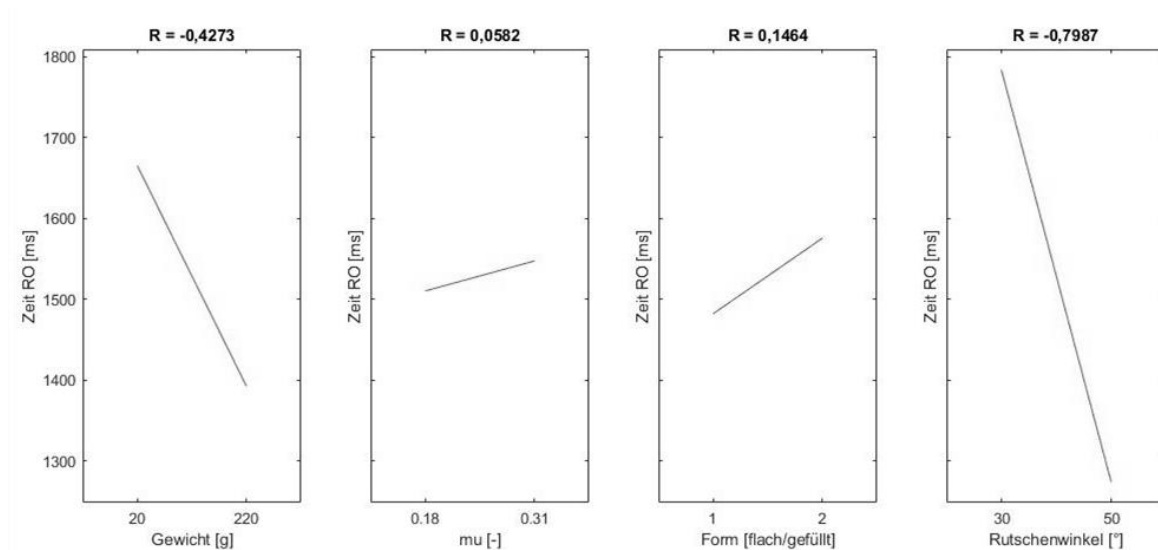


Abbildung 4.20.: Effekte auf die Einschleusung RO

Alle genannten Effekte, welche aus dem Parameterscreening ersichtlich sind, werden auch durch den dazugehörigen Korrelationskoeffizienten aus Tabelle 2.3 untermauert. Somit werden für weitere Untersuchungen alle Parameter mit einem Korrelationskoeffizienten $R < 0,5$ nicht berücksichtigt.

Aus den Ergebnissen der Analysen lässt sich für die weitere Untersuchung festhalten, dass die Parameter Gewicht, Rutschenwinkel und Stoßwinkel jene Größen sind, welche einen Einfluss auf die Zielgröße haben. Für die Erzeugung der Regressionsmodelle (Schritt 7) und dem dazugehörigen Teachingset (Schritt 6) werden daher nur jene drei Parameter berücksichtigt. In Tabelle 4.5 werden noch einmal die Ergebnisse des Parameterscreenings zusammengefasst.

Tabelle 4.5.: Ergebnisse des Parameterscreenings

Parameter	Beurteilung
Gewicht	berücksichtigen
Gleitreibungskoeffizient	verwerfen
Form	verwerfen
Rutschenwinkel	berücksichtigen
Biegeschlaffheit	verwerfen
Stoßwinkel	berücksichtigen

4.5 Anwendung Schritt 5 Teachingset -Versuchsplan erzeugen

Die minimale Stützstellenzahl ist laut Definition der Methodik mit zwei²⁹ (obere und untere Grenze) je Parameter festgelegt. Weiters folgen zwei weitere Versuchspläne mit jeweils drei und vier Stützstellen je Parameter³⁰. Die Stützstellen für die Parameter des FFS sind in Tabelle 4.6 angeführt. Das Gewicht der Proben kann stufenlos angepasst werden. Daher kann dieser Faktor symmetrisch über den Wertebereich verteilt werden. Für den Stoßwinkel und den Rutschenwinkel müssen bauliche Beschränkungen berücksichtigt werden. Daher ergibt sich hierbei eine unsymmetrische Verteilung. Mit diesen Stützstellen (siehe Tabelle 4.6) wird ein vollfaktorieller Versuchsplan erstellt, welcher als Grundlage für die Versuche im Schritt 6 dient.

Tabelle 4.6.: Stützstellen der Faktoren

Stützstellen je Parameter	Faktoren		
	Gewicht	Stoßwinkel	Rutschenwinkel
1	20	29°	31
2	86	45°	36
3	154	108°	44
4	220	124°	47

³⁰ Da die Stützstellenanzahl für jeden Parameter gelten, ergeben sich für zwei Stützstellen $2^3 = 8$, für drei Stützstellen $3^3=27$ und für vier Stützstellen $4^3=64$ Stützpunkte durch die Kombinationen der einzelnen Stützstellen.

4.6 Anwendung Schritt 6 – Versuche für das Teachingset durchführen

Im Schritt 6 werden die Versuche aus dem Teachingset-Versuchsplan durchgeführt. Dabei muss wiederum die Durchführungsvorschrift aus Schritt 2 berücksichtigt werden. Die daraus entstehenden Daten dienen in Schritt 7 als Trainingsdaten für die Erzeugung der Regressionsmodelle.

4.7 Anwendung Schritt 7 – Regressionsmodelle erzeugen

Mit den Trainingsdaten aus Schritt 6 werden für einzelne Submodelle der Anlage Regressionsmodelle erzeugt und in weiter Folge zu einem Gesamtmodell des FFS zusammengeführt (siehe Abbildung 4.25). Dazu werden die Regressionsmodelle mit Hilfe von Matlab erzeugt.

Als erster Schritt gilt bei der Erzeugung der Regressionsmodelle die Unterscheidung zwischen Modellen mit bekannten und unbekanntem Anfangsbedingungen. Als System mit bekannten Anfangsbedingungen gilt die Rutsche oben (RO) (Einschleusung), da hier die Geschwindigkeit mit null als definierter Zustand gilt. Die Submodelle Umlenkung und Rutsche unten (RU) dienen als Beispiel für Regressionsmodelle ohne bekannte Anfangsbedingungen.

In den folgenden Unterkapiteln werden die Regressionsmodelle der Submodelle erzeugt und in weiter Folge zu einem Gesamtmodell zusammengefügt. Der dazugehörige Matlab-Code befindet sich im Anhang C.

4.7.1 Regressionsmodelle mit bekannten Anfangsbedingungen

Wie bereits erwähnt, kann für den Versuchsstand des FFS die Einschleusung als ein Submodell angesehen werden, dessen Anfangsbedingungen bekannt sind³¹. Als Datengrundlage für das Regressionsmodell wird der Teachingset-Versuchsplan inklusive der ermittelten Durchlaufzeiten an Matlab übergeben. Wie in Abbildung 4.21 ersichtlich, werden die übermittelten Daten im Regression Learner in Predictors³² und in Response unterteilt. Als Predictors werden für ein Regressionsmodell mit bekannten Anfangsbedingung nur die eigenen Parameter (Probengewicht und Rutschenwinkel) übergeben.

³¹Durch das Loslassen aus der Ruhe (Geschwindigkeit=0) ist der Anfangszustand klar definiert.

³² Parameter, welche für die Ermittlung der Zielgröße eingegeben werden müssen

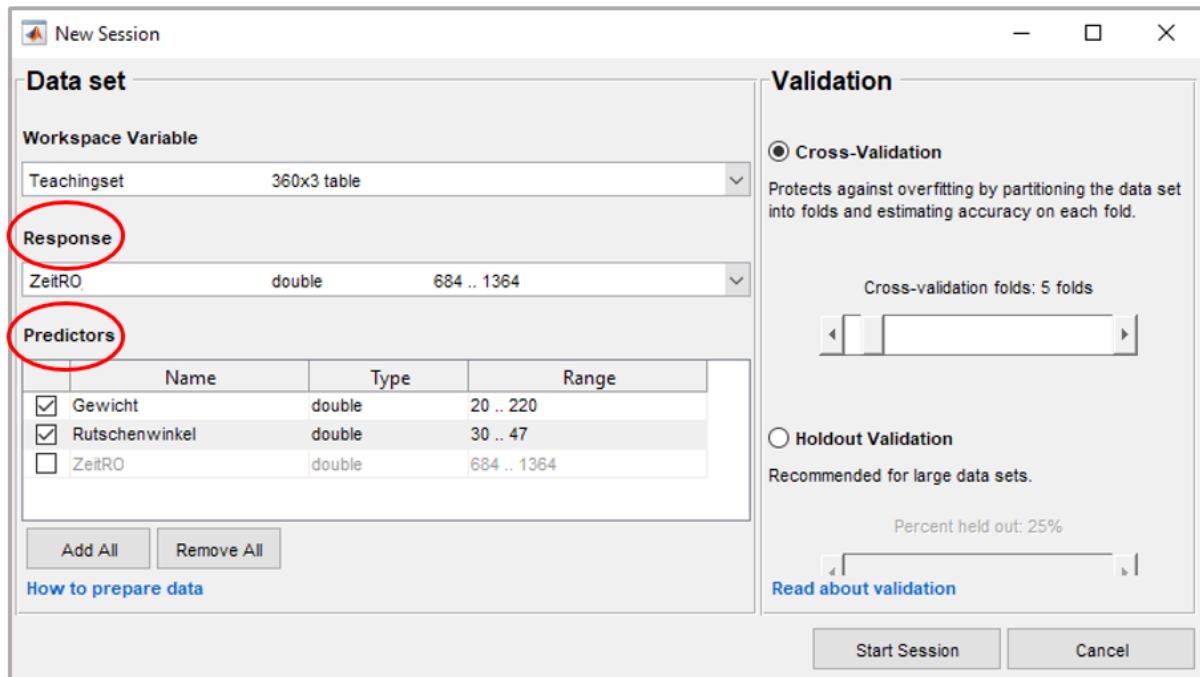


Abbildung 4.21.: Unterteilung des Teachingsets der Rutsche oben (RO) in Response und Predictors

Um für die übermittelten Trainingsdaten das am besten geeignete Regressionsmodell zu finden, werden alle im Programm hinterlegten Beschreibungsmodelle³³ getestet. Dabei wird die Abweichung, der vom Regressionsmodell erzeugten Daten zu den Validierungsdaten, ermittelt. Als Beurteilungsgröße wird hierbei der Root Mean Square Error (RMSE) ausgegeben. Das Modell mit dem geringsten RMSE und somit der besten Übereinstimmung wird in weiterer Folge für die Beschreibung des Submodells verwendet.

In Abbildung 4.22 sind alle getesteten Regressionsmodelle inklusive des jeweiligen RMSE angeführt. Das zu bevorzugende Regressionsmodell mit der geringsten Abweichung ist darin markiert. Für die Durchlaufzeit der Rutsche oben (RO) hat sich das Gaussian Process Regression -Modell Matern 5/2 mit einem RMSE von 51,579ms³⁴ als das beste Modell herausgestellt. Das Diagramm in Abbildung 4.22 zeigt die dafür verwendeten Trainingsdaten (blau) und die daraus vorhergesagten Ergebnisse (gelb) an.

³³ Alle Regressionsmodelle, welche in Matlab hinterlegt sind, wurden im Kapitel 2.3 in Tabelle 2.4 dargestellt.

³⁴ Das bedeutet eine Abweichung von durchschnittlich 51,579 ms auf die Durchlaufzeit von 700-1300ms. Dies entspricht einer ungefähren Abweichung von 5-10% der Gesamtzeit.

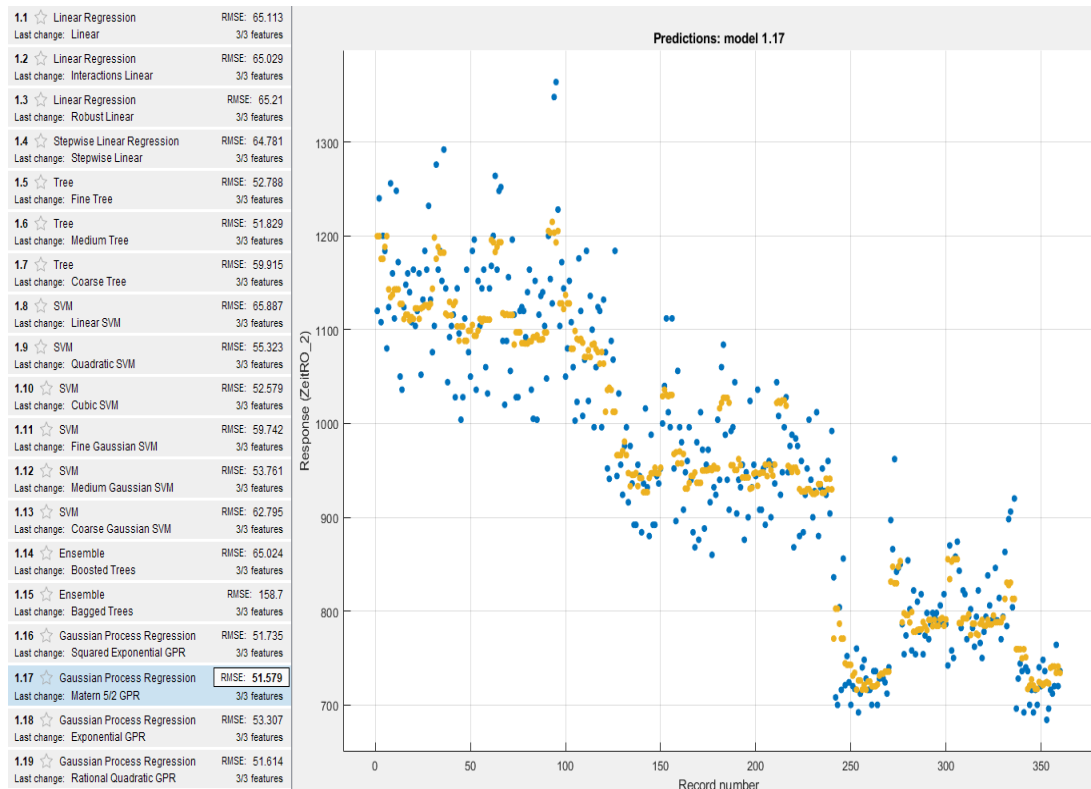


Abbildung 4.22.: Wahl des Regressionsmodells mit dem geringsten RMSE

4.7.2 Regressionsmodelle mit unbekanntem Anfangszustand

Für die Submodelle deren Anfangszustand nicht klar definiert ist, müssen bei der Erzeugung des Regressionsmodells die Anfangsbedingungen des Stückgutes berücksichtigt werden.

Probleme bei der Bestimmung des Anfangszustandes

Ein Ansatz für die Bestimmung des Anfangszustandes ist die Geschwindigkeitsmessung am Eintritt jedes Submodells. Als Geschwindigkeit wird hierbei die translatorische Geschwindigkeit des Schwerpunktes parallel zur Rutschenoberfläche benötigt. Dabei tritt jedoch das Problem auf, dass das Stückgut nicht in jedem Fall eine rein translatorische Bewegung aufweist. Besonders nach dem Umlenkvorgang kommt es häufig zur Überlagerung von translatorischen und rotatorischen Bewegungen.

Die Geschwindigkeitsmessung durch zwei Lichtschranken (siehe Abbildung 4.23) berechnet die Geschwindigkeit, indem der Abstand der beiden Lichtschranken durch die verstrichene Zeit, zwischen dem Auslösen der Lichtschranken, dividiert wird. Die zuvor beschriebene Überlagerung der Bewegungen verfälscht in weiterer Folge die Geschwindigkeitsmessung des Prüfstandes und führt zu falschen Ergebnissen.

In Abbildung 4.23 wird dargestellt, wie das beschriebene Problem anhand eines Realversuches aussieht. Hierbei wird im Bild links ersichtlich, wie die erste Lichtschranke ausgelöst wird. Danach

führt das Paket gleichzeitig eine Rotation und Translation aus. Durch die Rotation wird die zweite Lichtschranke (Abbildung in der Mitte) innerhalb kürzester Zeit ausgelöst, wodurch die Geschwindigkeitsmessung eine zu hohe Geschwindigkeit ausgibt. Zusätzlich zu der Verfälschung des Ergebnisses wird hierbei nicht die Geschwindigkeit des Schwerpunktes, sondern jene der unteren Kante des Paketes gemessen, wodurch diese Messung nicht zielführend ist.



Abbildung 4.23.: Herausforderung bei der Messung der Geschwindigkeit

Ein weiterer Ansatz für die Geschwindigkeitsmessung wäre eine optische Messung, mittels einer Highspeed Kamera, welche über eine Bilderkennung die Geschwindigkeit ermittelt. Diese Lösung ist mit einem hohen Investitionsaufwand verbunden und garantiert nicht für jedes Stückgut eine brauchbare Lösung. Wenn das Stückgut zum Beispiel einer einfarbigen Kugel entspricht, kann eine Bilderkennung die translatorische und rotatorische Bewegung nichtmehr unterscheiden und somit ist die Geschwindigkeitsmessung für solche Anwendungen wiederum nicht brauchbar.

Anfangswertbestimmung durch statistische Grundlagen

Wegen den zuvor genannten Problemen der Anfangswertbestimmung durch messtechnische Einrichtungen wird als Lösung ein statistischer Ansatz gewählt. Dabei werden bei der Erzeugung des Regressionsmodells neben den Parametern des vorherzusagenden Submodells auch die Parameter der davorliegenden Submodelle übergeben.

Als Beispiel dazu dient die Abbildung 4.24. Hierbei wird das Regressionsmodell für die untere Rutsche des Prüfstands erzeugt. Im Vergleich zu Abbildung 4.21 wird hierbei ersichtlich, dass neben den Faktoren der Rutsche unten (RU), auch die Faktoren der Umlenkung und der Rutsche oben (RO) in das Regressionsmodell einfließen. Dadurch entsteht eine Verknüpfung mit den vorherliegenden Submodellen, wodurch automatisch die Anfangsbedingung des Stückgutes berücksichtigt wird.

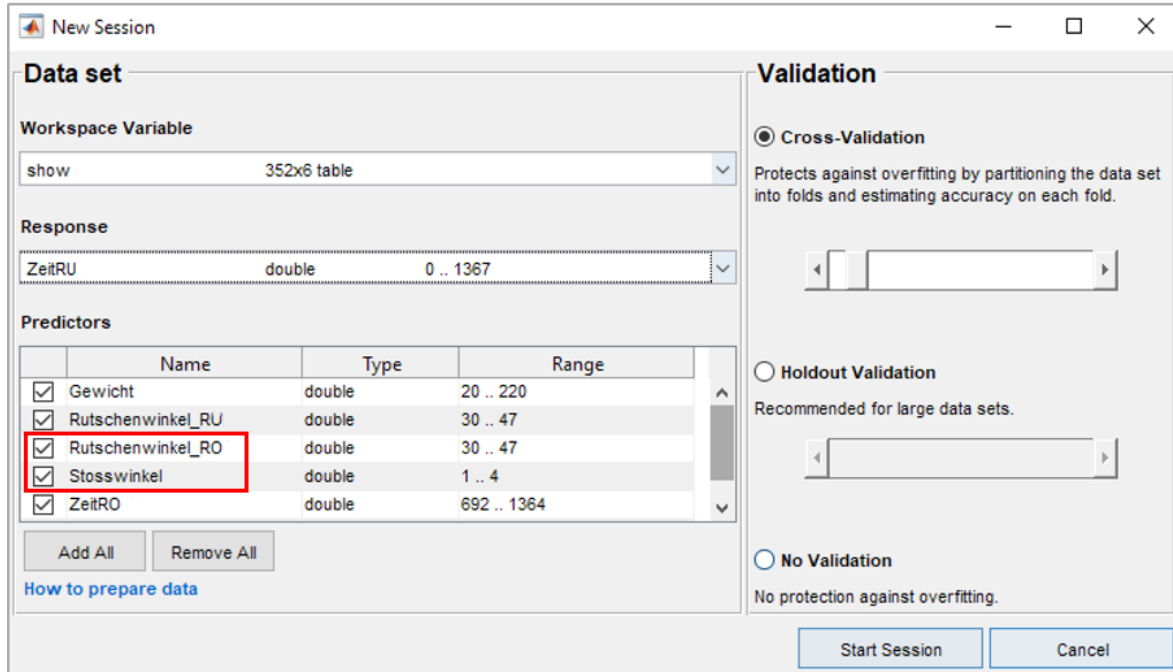


Abbildung 4.24.: Erzeugung des Regressionsmodells RU abhängig von RO und der Umlenkung

Dieses Netzwerk, lässt sich entlang des gesamten Layouts des FFS erweitern. Als Beispiel für das Regressionsmodell der Rutsche unten (RU), wird in Abbildung 4.25 noch einmal anhand einer Skizze des FFS der Ablauf dargestellt. Dabei berücksichtigt das Regressionsmodell mit welcher Einstellung die Rutsche oben (RO) verbaut wurde und welcher Umlenkvorgang stattgefunden hat. Aus der Summe der Submodelle lässt sich weiters ein Predictionmodell der gesamten Anlage erstellen.

Durch den statistischen Ansatz für die Berücksichtigung der Anfangsbedingung muss kein zusätzlicher messtechnischer Aufwand betrieben werden. Die benötigten Parameter für die Durchlaufzeitermittlung sind dabei schon bekannt³⁵.

³⁵ Der Rutschenwinkel und Stoßwinkel werden bereits durch das Anlagenlayout vorgegeben. Das Gewicht des Polybags kann entsprechend über eine Bandwaage vor der Einschleusung ermittelt werden oder durch das Scannen des Labels bekannt sein.

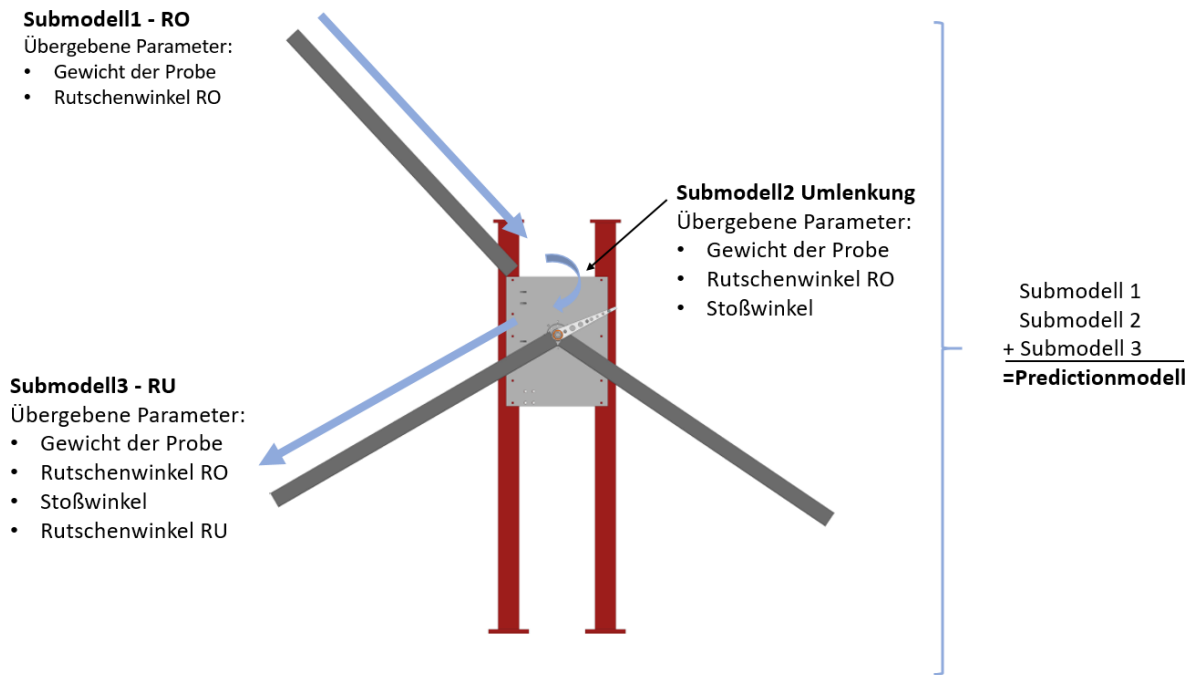


Abbildung 4.25.: Darstellung der Zusammensetzung der Submodelle

4.7.3 Zusammensetzung der Regressionsmodelle

Nachdem die Regressionsmodelle der Submodelle entsprechend erzeugt worden sind, ist es möglich die Durchlaufzeiten der einzelnen Submodelle zu addieren. Dadurch erhält man in weiterer Folge die Durchlaufzeit der Stückgüter, über welche man auf den Durchsatz schließen kann. Ein weiterer Vorteil, der durch die Betrachtung als einzelne Submodelle entsteht, liegt in der Anpassungsmöglichkeit des verwendeten Regressionsmodells. Dabei besteht die Möglichkeit für jedes Submodell das optimale mathematische Beschreibungsmodell zu finden.

Dadurch kann beispielhaft die Umlenkung durch ein anderes Regressionsmodell beschrieben werden als jenes der Rutschen. Nachdem für jede Stützstellenanzahl ein Predictionmodell nach diesem Vorgehen erzeugt worden ist, kann in Schritt 8 die Validierung der Modelle erfolgen.

4.8 Anwendung Schritt 8 – Regressionsmodelle validieren

Die Validierung der Methodik am FFS wird mit den Daten aus 270 realen Versuchen durchgeführt. Diese Versuche werden bei zufälligen Parametereinstellungen durchgeführt. Für die zufälligen Parameter wird gleichzeitig mit Hilfe des Predictionmodells eine Vorhersage der Durchlaufzeiten ermittelt. Die vorhergesagten Werte werden dann durch die realen Werte validiert.

Betrachtet wird dabei für jede Stützstellenkonstellation die Abweichung der Durchlaufzeit der Rutsche oben (RO), Rutsche unten (RU), Umlenkung und der daraus abgeleiteten Gesamtzeit. Die Abweichung für jedes Modell wird mit Hilfe eines Histogramms dargestellt.

4.8.1 Validierung der Rutsche oben (RO)

Für die Untersuchung der Abweichung der Rutsche oben (RO) ergibt sich wie in Abbildung 4.26 dargestellt, dass die Wahl einer vierten Stützstelle je Parameter ein besseres Ergebnis liefert. Aus dem Histogramm lässt sich ableiten, dass bei einem Teachingset mit drei Stützstellen je Parameter 202 vorhergesagte Durchlaufzeiten um +/- 5% vom Realwert abweichen. Bei vier Stützstellen je Parameter hingegen sind es 214 Werte, welche sich in diesem Bereich befinden. Da die Genauigkeit von drei auf vier Stützstellen nur geringfügig steigt, ist hierbei keine weitere Iteration für eine fünfte Stützstelle vorgesehen. Durch diese Erkenntnis wird die Stützstellenanzahl für das Teachingset in Kapitel 4.5 auf vier Stützstellen angepasst.

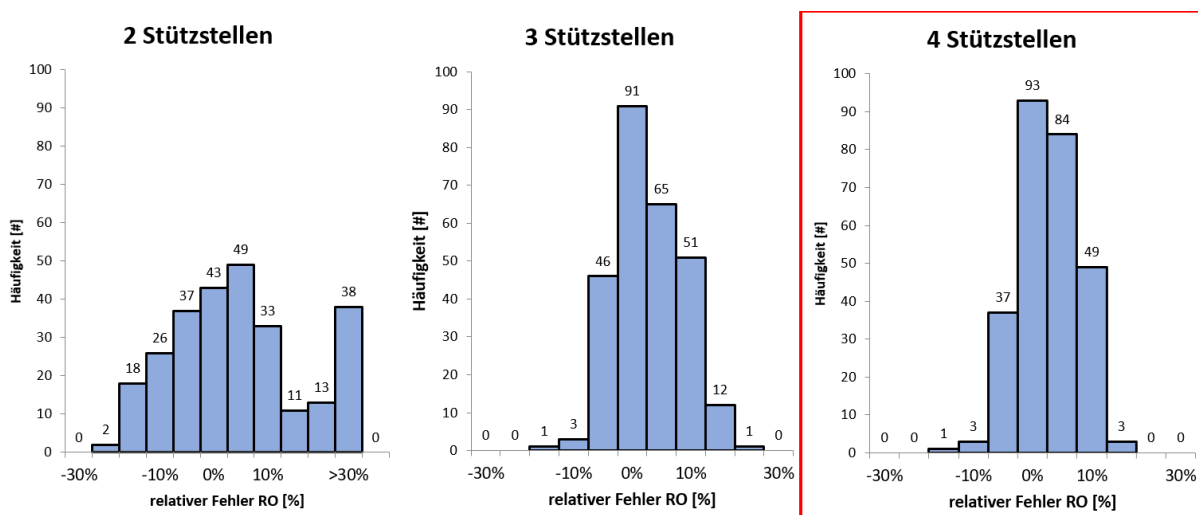


Abbildung 4.26.: Relativer Fehler der Durchlaufzeit der Rutsche oben (RO)

4.8.2 Validierung der Umlenkung

Aus den Ergebnissen der Validierung in Abbildung 4.27 lässt sich ableiten, dass der Umlenkungsprozess, das am komplexesten zu beschreibende Modell ist. Aufgrund der vielen Einflussfaktoren im Umlenkprozess war dies bereits im Vorfeld zu erwarten. Da die Dauer der Umlenkung jedoch nur rund einem Zehntel der Durchlaufzeit der Rutschen entspricht, wirkt sich die Ungenauigkeit auf das Gesamtsystem, wie später in Kapitel 4.8.4 ersichtlich, nur geringfügig aus. Neben der großen Streuung, welcher der Umlenkprozess unterliegt, lässt sich aus Abbildung 4.27 auch die optimale Stützstellenanzahl von drei Stützstellen je Parameter feststellen. Bei zwei Stützstellen je Parameter hingegen liegen noch 128 der vorhergesagten Durchlaufzeiten bei einer Abweichung von rund 50%. Eine Stützstellenanzahl von vier bewirkt gegenüber drei Stützstellen eine Verschlechterung. Dahinter lässt sich ein Overfitting vermuten.

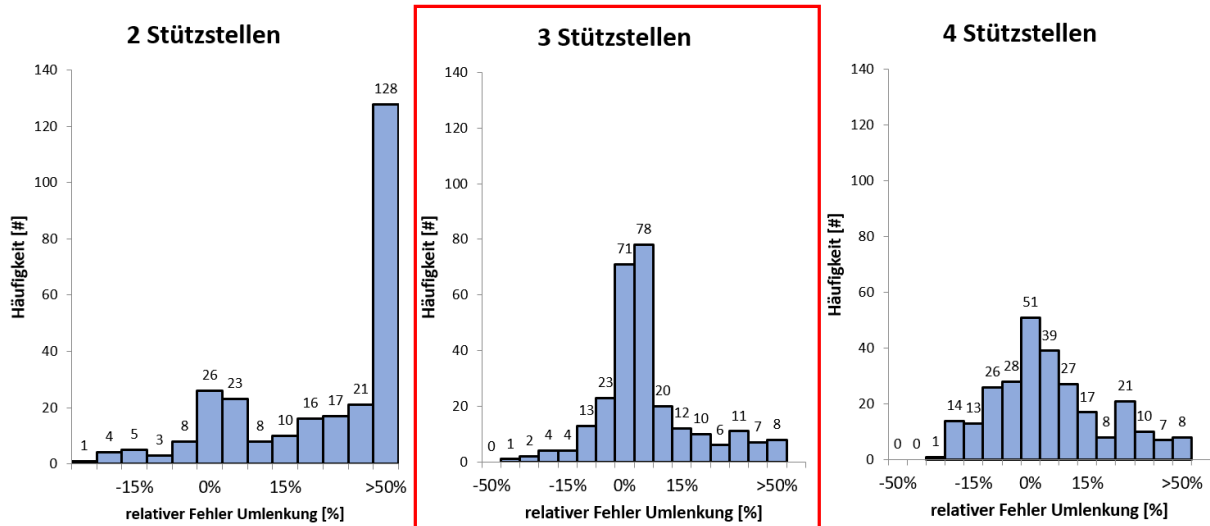


Abbildung 4.27.: Relativer Fehler der Durchlaufzeit der Umlenkung

4.8.3 Validierung der Rutsche unten (RU)

Abbildung 4.28 zeigt, dass sich bei einer Stützstellenanzahl von drei, 257 von 270 Durchlaufzeiten innerhalb einer Abweichung von +/-5% befinden. Eine Erhöhung der Stützstellen führt auch in diesem Fall zu keiner Verbesserung, sondern würde auch hier wieder zu einem Overfitting führen. Somit entspricht die gewählte Stützstellenanzahl von drei dem Optimum, wodurch auch für dieses Modell keine Iteration notwendig ist.

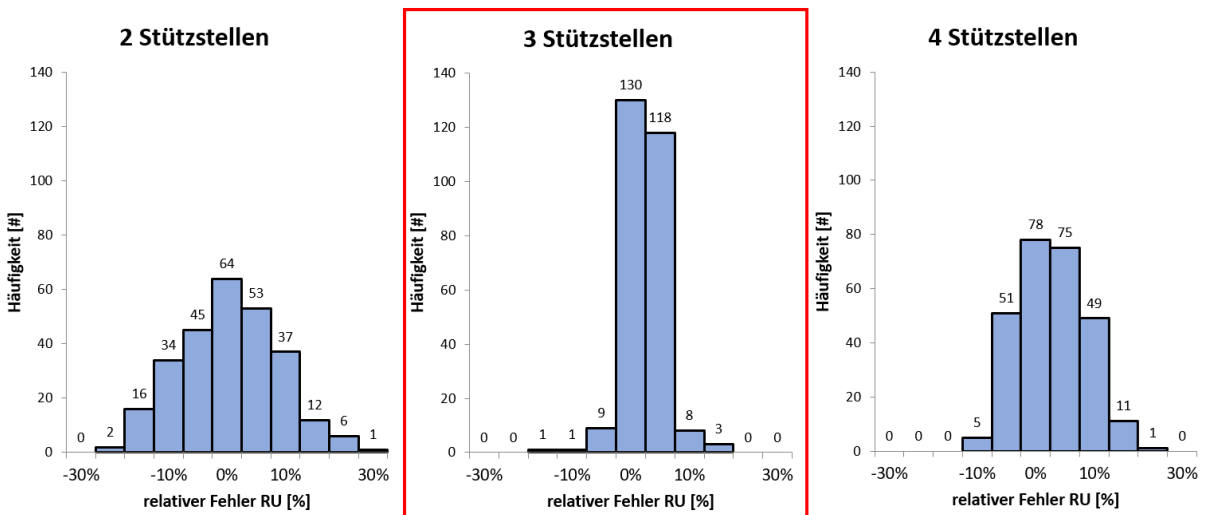


Abbildung 4.28.: Relativer Fehler der Durchlaufzeit der Rutsche unten (RU)

4.8.4 Validierung der Gesamtzeit

Für die Validierung der zusammengesetzten Durchlaufzeit wird die Durchlaufzeit durch Addition der einzelnen Submodelle mit der realen Durchlaufzeit ins Verhältnis gesetzt. In Abbildung 4.29 wird ersichtlich, dass bei einer Wahl von drei Stützstellen je Parameter (für jedes Submodell), 248 von 270 Durchlaufzeiten innerhalb der Abweichung von $\pm 5\%$ liegen.

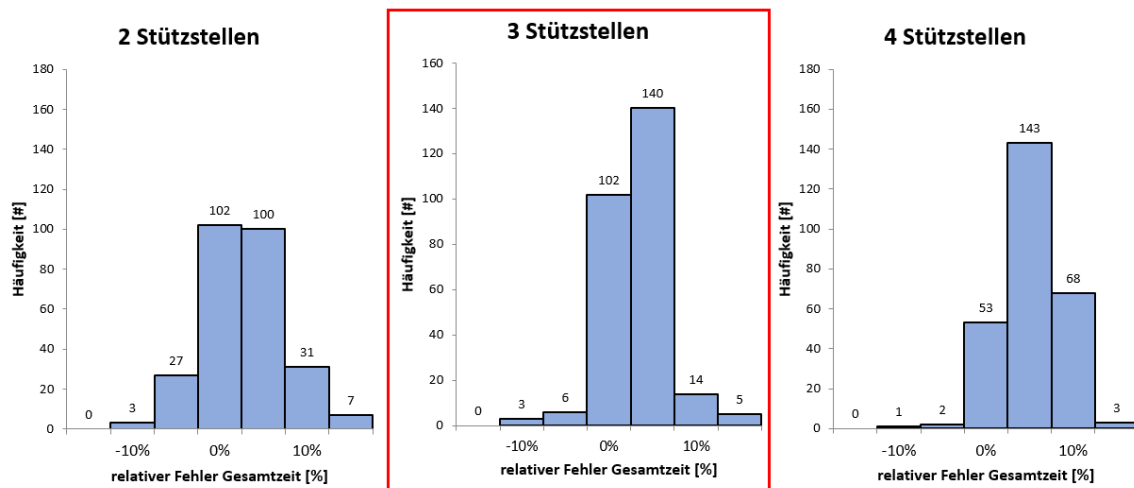


Abbildung 4.29.: Relativer Fehler der gesamte Durchlaufzeit

4.8.5 Anpassung an die optimale Stützstellenanzahl

Anhand der zuvor ermittelten optimalen Stützstellenanzahl jedes Submodells wird der Versuchsplan für das Teachingset angepasst. Daraus ergibt sich ein neues Predictionmodell für die gesamte Durchlaufzeit, worin jedes Submodell mit der optimalen Stützstellenanzahl erzeugt wird (siehe Abbildung 5.1). Die Ergebnisse der Validierung dieses Modells werden in Abbildung 4.30 dargestellt. Dabei wird ersichtlich, dass 166 der 170 vorhergesagten Durchlaufzeiten um 5% von der realen Zeit abweichen. Somit ergibt sich für 98% der durchgeführten Versuche eine maximale Abweichung von 5%. Angesichts dieser hohen Genauigkeit wird keine weitere Anpassung³⁶ des Modells benötigt. Auf Basis dieser Ergebnisse erfolgt in Kapitel 5 eine Beurteilung über die Gültigkeit der Methodik.

³⁶ Wenn die geforderte Genauigkeit durch die Anpassung der Stützstellen nicht erreicht werden kann, dann dient die Stichprobengröße als weitere Stellschraube.

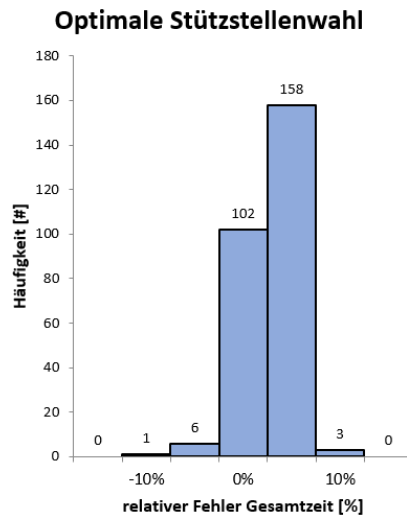


Abbildung 4.30.: Anpassung an die optimale Stützstellenanzahl für die gesamte Durchlaufzeit

4.9 Anwendung Schritt 9 – Userinterface erstellen

Das für den FFS erstellte Interface teilt sich in die Bereiche Anwendung, Voreinstellungen und einer grafischen Darstellung der Ergebnisse auf. Bei diesem Anwendungsfall liegt das Hauptaugenmerk in der Vorhersage der benötigten Durchlaufzeiten für ein vorgegebenes Sendungsspektrum. Daher findet sich im oberen Bereich des GUI der Button „Prüfdaten Laden“, durch welchen die zu untersuchenden Sendungsdaten aus der Datei des Eingabefeldes darunter an das Programm übergeben werden. Durch das Laden dieser Datei wird das Predictionmodell auf diese Daten angewandt und liefert eine weitere Datei mit den Ergebnissen der Durchlaufzeit. Der Speicherort dieser Datei wird im Fenster auf der rechten Seite unter „Ergebnisausgabe“ angezeigt. Der untere Bereich des GUI dient für die Voreinstellungen des Programmes. Darin befinden sich die Buttons „Versuchsplan Parameterscreening Laden“, „Versuchsplan Teachingset Laden“ und „Versuchsplan Validierungsset Laden“. Die jeweiligen Dateien werden durch das Eingabefeld unter dem jeweiligen Button ausgewählt und anschließend geladen.

Auf der rechten Seite des Interfaces befindet sich eine Zusammenfassung der Dateien, welche für die Ermittlung der Durchlaufzeit verwendet wurden. Dabei wird auch die Ergebnisausgabe inklusive dem dazugehörigen Ordnerpfad aufgelistet. Darunter befindet sich noch eine grafische Darstellung der Ergebnisse in Form eines Histogramms. In Abbildung 4.31 wird das für den FFS erzeugte Userinterface dargestellt.

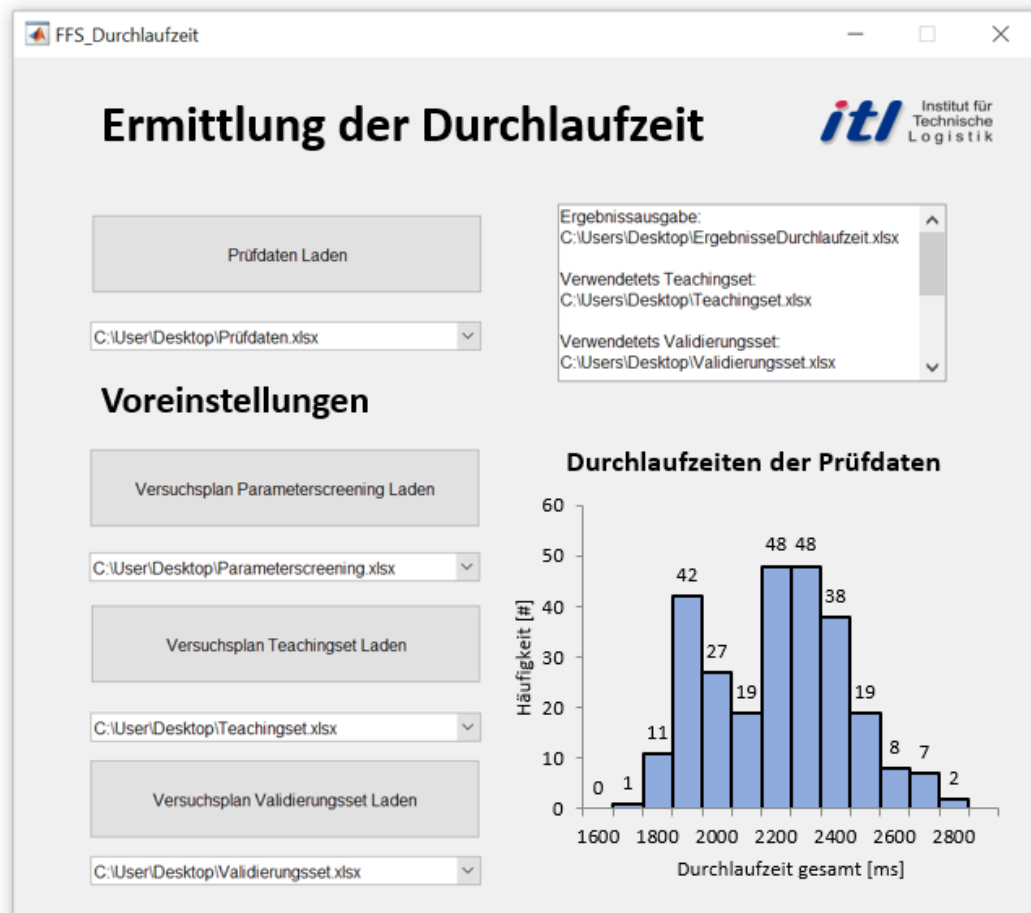


Abbildung 4.31.: Userinterface für den Free-Fall-Sorter

5 ERGEBNISSE UND INTERPRETATION

In diesem Kapitel werden die Ergebnisse aus Kapitel 4, welche durch die Anwendung der Methodik auf den FFS entstanden sind, interpretiert. Des Weiteren wird die allgemeine Gültigkeit der Methodik für weitere Stückgüter betrachtet und in einem Ausblick festgehalten, was durch diese Untersuchungen in Zukunft möglich ist.

5.1 Zusammenfassung der Ergebnisse der exemplarischen Anwendung

In diesem Unterkapitel wird das Ergebnis aus der Anwendung der Methodik aus Kapitel 4 zusammengefasst. Als Ergebnis der Methodik wird wie in Abbildung 5.1 dargestellt ein Predictionmodell mit der optimalen Stützstellenanzahl erstellt. Dieses entsteht aus der Kombination einzelner Submodelle, wobei für jedes Submodell gesondert die ideale Stützstellenanzahl (je Parameter) gewählt wird.

Daraus ergibt sich ein Predictionmodell, welches für 98% der vorgegebenen Proben eine Durchlaufzeit mit einer Voraussagegenauigkeit $> 95\%$ liefert.

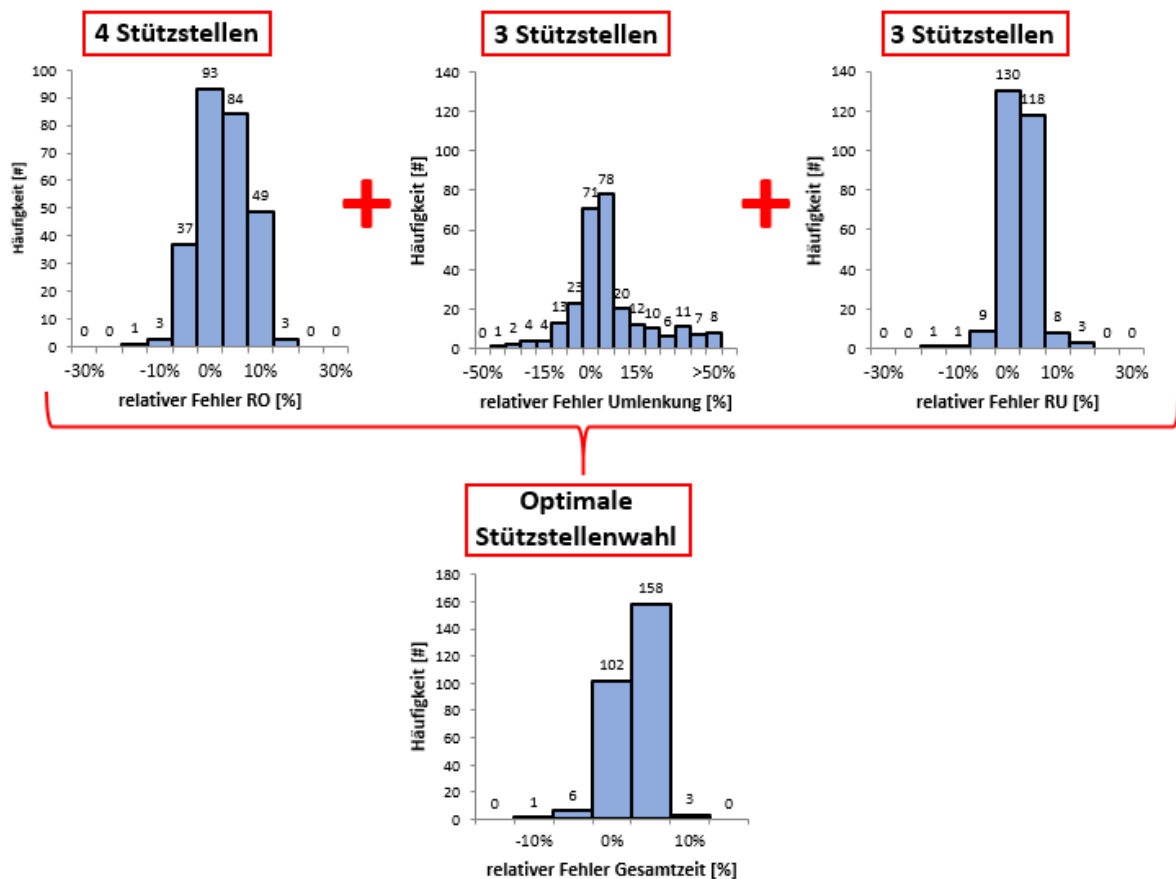


Abbildung 5.1.: Ergebnisse der exemplarischen Anwendung am FFS

5.2 Beurteilung der Gültigkeit der Methodik

Unter Betrachtung der Submodelle und dem daraus entstehenden Predictionmodells des FFS wird eine Aussage über die Gültigkeit der Methodik getroffen.

5.2.1 Möglichkeiten für die Beurteilung der Gültigkeit der Methodik

Als Grundlage, um eine Aussage über die Gültigkeit der Methode treffen zu können, liegen zwei Möglichkeiten der Validierung der Ergebnisse vor. Die erste Möglichkeit baut auf einer Validierung mittels Realversuchen auf, wohingegen die zweite Möglichkeit Validierungsdaten über ein LHS (Latin Hypercupe Sampling) erzeugt. Wie bereits in der Anwendung der Methode ersichtlich, fällt die Wahl auf die Überprüfung mit Versuchsdaten aus Realversuchen³⁷.

5.2.2 Gültigkeit der Submodelle

In diesem Unterkapitel erfolgt die Beurteilung für die Durchlaufzeitermittlung von Submodellen durch Regressionsmodelle. Für die Überprüfung der Gültigkeit wird die Validierung der Submodelle aus dem Kapitel 4.8 herangezogen. Dabei hat sich ergeben, dass für die Betrachtung der Rutschen 95% der vorhergesagten Werte eine Abweichung liefern, welche nur +/-5% von den Realversuchen abweicht. Eine solche Schwankung bedeutet in diesem Anwendungsfall eine Abweichung von rund 50ms³⁸. Daraus lässt sich eine sehr hohe Genauigkeit der erzeugten Submodelle ableiten.

Die Betrachtung des Umlenkvorganges ergibt eine Schwankung von +/-10% für 76% der 270 Versuche. Angesichts der kurzen Dauer der Umlenkung ergibt sich hierbei eine durchschnittliche Abweichung von 20ms³⁹. Daraus ist ersichtlich, dass das Submodell, welches vom Stoßverhalten abhängt, das am komplexesten zu beschreibende Modell ist. Hierbei sollte jedoch nicht unerwähnt bleiben, dass der Polybag durch seine hohe Variabilität als ein besonders schwer zu beschreibendes Stückgut gilt. Daher sind Abweichungen von 10% durchaus zufriedenstellend. Aus diesen Ergebnissen lässt sich ableiten, dass die Methodik für weniger komplexe Stückgüter eine höhere Genauigkeit aufweisen wird.

³⁷ Die Wahl von Realversuchen wird durch zwei Umstände begründet. Als erster Grund dient, dass bereits für das Parameterscreening Versuche benötigt werden. Daher muss der Aufwand der Probenvorbereitung ohnehin betrieben werden. Als zweiter Gesichtspunkt gilt, dass bei einer Validierung mittels LHS die Ergebnisse des LSH zuvor wiederum mit Realversuchen validiert werden müssen.

³⁸ Die Durchlaufzeit für ein Rutschenelement beträgt 700-1400ms.

³⁹ Der Dauer eines Umlenkvorgangs beträgt 90-200ms.

Berücksichtigt man neben der Abweichung, welche sich aus der Validierung ergibt, die statistische Schwankung, welcher jeder Versuch unterliegt, kann die Annahme getroffen werden, dass die Schwankung nicht nur dem Predictionmodell zuzuordnen ist, sondern auch einem statistischen Grundrauschen⁴⁰, welchem die Versuche unterliegen.

Hierbei darf auch nicht vernachlässigt werden, dass die Versuche unter idealen Bedingungen durchgeführt werden, was in der Praxisanwendung sicher zu höheren Abweichungen führt. Aufgrund der Größenordnungen der zufälligen Schwankungen (negative Auswirkung) und der Effekte der Idealisierung (positive Auswirkung) wird erwartet, dass sich die Effekte in Summe ausgleichen.

Aus diesen Erkenntnissen kann abgeleitet werden, dass sich die einzelnen Submodelle, in welche die Sortieranlage unterteilt wurde, sehr gut durch Regressionsmodelle der Submodelle annähern lassen. Durch die hohe Genauigkeit der Ergebnisse lässt sich daraus eine Gültigkeit der Methodik für die Submodelle bestätigen.

5.2.3 Gültigkeit der Zusammensetzung der Submodelle zu einem Predictionmodell

In diesem Absatz wird die Frage beantwortet, ob es möglich ist Submodelle zu einem Gesamtmodell modular zusammen zu setzen. Dies galt zu Beginn der Arbeit als die Frage mit der größten Ungewissheit, da ein solches Vorgehen aus der Literatur nicht bekannt ist.

Die Validierung des Gesamtmodells ergab, dass sich 98 % der vorhergesagten Durchlaufzeiten innerhalb einer Schwankung von +/-5% befinden. Darin widerspiegelt sich die Schwankung, welche bereits in den Submodellen auftrat. Da die Dauer der Umlenkung nur rund 10% der gesamten Durchlaufzeit beinhaltet, verschwindet⁴¹ die erhöhte Schwankung dieses Submodells bei der Betrachtung der Gesamtzeit nahezu.

Für die Validierung der Gesamtzeit gilt auch hierbei die Annahme der zufälligen Schwankungen, welche dem Predictionmodell zugeschrieben werden und die Verringerung der zufälligen Schwankungen, welche der Idealisierung zugrunde liegen.

⁴⁰ Dieses Grundrauschen wurde in der Betrachtung der Normalverteilung der Versuche untersucht und beträgt für diesen Anwendungsfall 2%.

⁴¹ Da die Dauer der Umlenkung nur 10% der Gesamtzeit beträgt, wirkt sich eine Schwankung von 10% des Umlenkvorgangs nur als 1% der Gesamtzeit aus.

Aufgrund der geringen Abweichung der Gesamtzeit des Predictionmodells gegenüber den Realversuchen von nur 5% wird die Zusammensetzung der Submodelle nicht nur als gültig, sondern auch als sehr genau angesehen.

5.2.4 Einschränkungen der Methodik

Bevor eine Aussage über eine allgemeine Gültigkeit für Sortiervorgänge von Stückgütern getroffen wird, sollten auch die Einschränkungen der Methodik nicht unerwähnt bleiben.

Als eine Einschränkung gilt die Idealisierung der Massenverteilung. Dabei wird angenommen, dass die Beladung in der Mitte des Polybags fixiert ist und sich nicht bewegt. Durch die Bewegung des Inhaltes kann jedoch das Stoßverhalten und/oder das Bewegungsverhalten verändert werden. Ein weiteres einschränkendes Kriterium ist das Bewegungsverhalten des Polybags bzw. des Stückgutes allgemein. Das gesamte Spektrum der verwendeten Stückgüter muss für die Anwendung der Methodik dasselbe Bewegungsverhalten aufweisen. Dabei ist unerheblich, ob das Stückgut rollt, gleitet oder eine überlagerte Bewegung ausführt. Wichtig dabei ist nur, dass diese Bewegung für alle Stückgüter gleich ist. Wenn hingegen dasselbe Stückgut einmal eine reine Translation und einmal eine reine Rotation aufweist, treten starke Abweichungen der vorhergesagten Durchlaufzeit auf.

Als letzter Punkt darf nicht unerwähnt bleiben, dass die Methodik bisher nur für normalverteilte Abläufe untersucht worden ist. Daher ist dies eine weitere Einschränkung, welche bei der Anwendung berücksichtigt werden muss.

5.2.5 Verallgemeinerung der Methodik für weitere Anwendungsfälle

Die allgemeine Gültigkeit der Methodik besagt, dass sich die Durchlaufzeit für Stückgüter in Sortieranlagen durch zusammengesetzte Regressionsmodelle bestimmen lässt. Ermöglicht wird dies durch das strukturierte Vorgehen von der Betrachtung der Sortieranlage bis hin zur Validierung des Modells. Als Besonderheit hierbei gilt der versuchsbasierte Aufbau, wodurch auch Effekte, welche nicht physikalisch abbildbar sind, berücksichtigt werden können. Somit ist man an keine analytischen Berechnungen gebunden. Durch die Gültigkeit für einen breiten Parameterbereich wird das Anwendungsfeld zusätzlich stark ausgeweitet.

Vergleicht man die bestehende Analytik für die Durchlaufzeitberechnung des FFS mit den Ergebnissen des Predictionmodells, so lassen sich die Abweichungen von 10% auf 5% verringern. Bei der Anwendung für den Polybag wurde die Methodik für ein Stückgut getestet, welches an Variabilität nahezu nicht zu übertreffen ist. Die hohe Genauigkeit trotz des komplexen Verhaltens unterstreichen dabei ein weiteres Mal die Gültigkeit und Güte der Methode. Daraus lässt sich

ableiten, dass die Methode allgemein angewandt auf Sortieranlagen von Stückgütern Ergebnisse mit einer sehr hohen Genauigkeit liefert.

5.3 Ausblick

Das Hauptaugenmerk der Entwicklung der Methodik lag auf der Ermittlung der Durchlaufzeit für einen breiten Parameterbereich von Stückgütern. In der Entwicklungsphase zeigten sich jedoch mehrere Optionen auf, für welche die Methode verwendet werden kann. Dazu zählen die in den weiteren Unterpunkten erwähnten Anwendungen.

5.3.1 Prozessoptimierung und Anlagensteuerung

Unter die Prozessoptimierung fällt die Anpassung des Einschleusvorgangs des Stückgutes an die benötigte Durchlaufzeit des davorliegenden Stückgutes. Dabei kann diese maximiert werden, wodurch der Durchsatz und daraus die Anlagenauslastung optimiert werden. Des Weiteren kann die Steuerung der Klappe für die Umlenkung über die Daten aus dem Predictionmodell gesteuert werden. Es ist auch möglich Sendungen, bei denen Probleme bei der Verarbeitung auftreten, bereits im Vorfeld auszusortieren. Der Ansatz der Klappensteuerung und der Überwachung des Einschleusvorganges werden laut [SL21] unter dem Begriff „predictiv sorting“ zusammengefasst. Anhand der Ergebnisse dieser Arbeit kann der Gedanke von Schedler als umsetzbar betrachtet werden. Wodurch sich ein weiterer Vorteil aus der Entwicklung des Free-Fall-Sorters erschließt.

5.3.2 Anlagenplanung und Optimierung

Für ein vorliegendes Stückgüterspektrum können in der Planungsphase der Sortieranlage Optimierungen des Layouts getroffen werden. Aktuell wird das Layout des FFS durch die geometrischen Gegebenheiten der vorliegenden Hallenstruktur und der geforderten Endstellen erzeugt. Diese Auslegung kann erweitert werden, indem sie um eine Durchsatzberechnung durch das Predictionmodell ergänzt wird. Dabei kann für ein vorliegendes Sendungsspektrum der Durchsatz vorab abgeschätzt werden und dadurch die Auslastung des jeweiligen Layouts beurteilt werden.

5.3.3 Verbesserung durch Machine Learning

Durch den Verbau von zusätzlichen Sensoren in der Anlage können im Laufe des Betriebs Messdaten aufgenommen werden, durch welche das Predictionmodell sich laufend selbst optimiert. Somit kann das Predictionmodell direkt aus den Messdaten lernen, welche im realen Einsatz der Anlage entstehen. Daraus ergibt sich auch die Möglichkeit das Modell laufend an neue Stückgüter anzupassen, ohne ein separates Teaching und Parameterscreening durchzuführen.

5.4 Fazit

Unter Betrachtung der hohen Genauigkeit und mit allen Erweiterungsmöglichkeiten erweist sich die entwickelte Methode als sehr praktikabel. Besonders die Anpassungsfähigkeit der Methodik an weitere Anwendungsfälle zählt als großer Pluspunkt. Somit kann die Methodik als Grundstamm und Inspiration für weitere Anwendungen in ähnlichen Bereichen angesehen werden.

6 VERZEICHNISSE

6.1 Literaturverzeichnis

- [BBK17] Brell, C.; Brell, J.; Kirsch, S.: Statistik von Null auf Hundert. Springer Berlin Heidelberg, Berlin, Heidelberg, 2017.
- [BM13] Brunner, E.; Munzel, U.: Nichtparametrische Datenanalyse. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.
- [Dul18] Duller, C.: Tabellen. In (Duller, C. Hrsg.): Einführung in die nichtparametrische Statistik mit SAS, R und SPSS. Springer Berlin Heidelberg, Berlin, Heidelberg, S. 323–374, 2018.
- [Eck19] Eckstein, P. P.: Statistik für Wirtschaftswissenschaftler. Springer Fachmedien Wiesbaden, Wiesbaden, 2019.
- [EK21] Esser, K.; Kurte, J.: KEP-Studie 2021 – Analyse des Marktes in Deutschland. Eine Untersuchung im Auftrag des Bundesverbandes Paket und Expresslogistik e. V. (BIEK), Cologne, 2021.
- [Esr21] esri.com: Regressionsmodell erstellen.
<https://doc.arcgis.com/de/insights/latest/analyze/create-regression-model.htm>, Stand: 01.12.2021.
- [FHK+16] Fahrmeir, L. et al.: Statistik. Springer Berlin Heidelberg, Berlin, Heidelberg, 2016.
- [FKL09] Fahrmeir, L.; Kneib, T.; Lang, S.: Regression. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.
- [Fro18] Frost, I.: Einfache lineare Regression. Springer Fachmedien Wiesbaden, Wiesbaden, 2018.
- [Goß20] Goßlau, D.: Fahrzeugmesstechnik. Springer Fachmedien Wiesbaden, Wiesbaden, 2020.
- [Hof21] Hofmann: BA-1180255-Hofmann, 2021.
- [JH12] Jodin, D.; Hompel, M. ten: Sortier- und Verteilsysteme. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.
- [Kle13] Kleppmann Wilhelm: Versuchsplanunug. Produkte und Prozesse optimieren, 2013.

- [SL21] Schedler, M.; Landschützer, C.: Methodische Entwicklung eines neuartigen Sortiersystems für Polybags. Logistic Journal 2021, 2021.
- [SSK+21] Schadler, M. et al.: Characteristics of ‘polybags’ used for low value consignments (small packets) in the mail, courier, express and parcel industry. Logistic Journal, 2021.
- [SvH10] Siebertz, K.; van Bebber, D.; Hochkirchen, T.: Statistische Versuchsplanung. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.
- [SvH17] Siebertz, K.; van Bebber, D.; Hochkirchen, T.: Statistische Versuchsplanung. Springer Berlin Heidelberg, Berlin, Heidelberg, 2017.
- [The22] The MathWorks, Inc.: Train Regression Models in Regression Learner App. <https://de.mathworks.com/help/stats/train-regression-models-in-regression-learner-app.html>, Stand: 09.03.2022.
- [Tho96] Thorsten Böcker: Auslegung von Rutschen an Sortierendstellen. unter besonderer Berücksichtigung der Gleitreibungsbedingungen von Stückgütern, Dortmund, 1996.
- [Wal15] Waldi, R.: Statistische Datenanalyse. Springer Berlin Heidelberg, Berlin, Heidelberg, 2015.

6.2 Abbildungsverzeichnis

Abbildung 2.1: Funktionsprinzip des FFS [SL21]	6
Abbildung 2.2: Allgemeine Darstellung eines Modells als Blackbox	8
Abbildung 2.3.: Effektdiagramm des Faktors A	12
Abbildung 2.4: Beispiel für die Darstellung einer linearen Regression	14
Abbildung 2.5.: Tabelle für die Ermittlung des t-Wertes [Goß20]	17
Abbildung 2.6.: Darstellung des Konfidenzintervalls inklusive der dazugehörigen Grenzen	18
Abbildung 2.7: Vorgehensweise für eine 4-fach Kreuzvalidierung	19
Abbildung 3.1.: Aufbau der Methodik für die Durchlaufzeitermittlung	20
Abbildung 3.2.: Unterprozess Schritt 2 – Parameterscreening-Versuchsplan erstellen	23
Abbildung 3.3.: Prozess für die Erzeugung von Regressionsmodellen	28
Abbildung 3.4.: Validierungsprozess des Predictionmodells	29
Abbildung 4.1: Aufbau des FFS Versuchsstands	31
Abbildung 4.2: Ishikawa Diagramm mit den möglichen Einflussfaktoren auf die Durchlaufzeit	33
Abbildung 4.3: Grenzen für den Rutschenwinkel	37
Abbildung 4.4.: Grenzen für den Stoßwinkel	38
Abbildung 4.5.: Standardisieren des Verpackungsgewichtes auf 20g	39
Abbildung 4.6: Kunststoffproben mit verschiedenen Reibwerten	40
Abbildung 4.7: Proben mit selben Reibwerten, aber aus verschiedenen Materialien	40
Abbildung 4.8: Darstellung wie aus einer biegeschlaffen Probe eine steife Probe entsteht	41
Abbildung 4.9.: Befüllung der Proben für die Faktorstufe mit großem projizierten Querschnitt	42
Abbildung 4.10: Veränderung des projizierten Querschnittes bei selbem Gewicht	42
Abbildung 4.11.: Kennzeichnung der Messvorrichtungen am FFS	43
Abbildung 4.12.: Benutzeroberfläche nach einem gültigen Versuch	44
Abbildung 4.13.: Einschleusung der Proben per Hand	46
Abbildung 4.14.: Ergebnisse des Kolmogorov-Smirnov Tests für die Durchlaufzeit Rutsche oben (RO)	47
Abbildung 4.15.: Darstellung der Abweichung der Durchlaufzeit je Stichprobenanzahl	48
Abbildung 4.16.: Unterscheidung zweier Proben mit demselben Reibungskoeffizienten, aber unterschiedlicher Materialzusammensetzung	50
Abbildung 4.17.: Effekte auf die Gesamtzeit bei einem steilen Stoß	50
Abbildung 4.18.: Effekte auf die Gesamtzeit bei einem flachen Stoß	51
Abbildung 4.19.: Effekte auf den Umlenkvorgang	52
Abbildung 4.20.: Effekte auf die Einschleusung RO	52

Abbildung 4.21.: Unterteilung des Teachingsets der Rutsche oben (RO) in Response und Predictors	55
Abbildung 4.22.: Wahl des Regressionsmodells mit dem geringsten RMSE	56
Abbildung 4.23.: Herausforderung bei der Messung der Geschwindigkeit	57
Abbildung 4.24.: Erzeugung des Regressionsmodells RU abhängig von RO und der Umlenkung	58
Abbildung 4.25.: Darstellung der Zusammensetzung der Submodelle	59
Abbildung 4.26.: Relativer Fehler der Durchlaufzeit der Rutsche oben (RO)	60
Abbildung 4.27.: Relativer Fehler der Durchlaufzeit der Umlenkung	61
Abbildung 4.28.: Relativer Fehler der Durchlaufzeit der Rutsche unten (RU)	61
Abbildung 4.29.: Relativer Fehler der gesamte Durchlaufzeit	62
Abbildung 4.30.: Anpassung an die optimale Stützstellenanzahl für die gesamte Durchlaufzeit	63
Abbildung 4.31.: Userinterface für den Free-Fall-Sorter	64
Abbildung 5.1.: Ergebnisse der exemplarischen Anwendung am FFS	65

6.3 Tabellenverzeichnis

Tabelle 2.1: Vollfaktorieller Versuchsplan	9
Tabelle 2.2: Teilfaktorieller Versuchsplan.....	10
Tabelle 2.3: Definition der Stärke eines Korrelationskoeffizienten r	13
Tabelle 2.4.: Verfahren für Regressionsmodelle aus dem Machine learning Tool von Matlab [The22].....	15
Tabelle 3.1.: Ablauf der Methodik	21
Tabelle 4.1.: Codierung der Faktoren.....	36
Tabelle 4.2.: Vollfaktorieller Versuchsplan für das Parameterscreening.....	36
Tabelle 4.3.: Codierung der Parameter.....	45
Tabelle 4.4.: Berechnung der Abweichung der Durchlaufzeiten je Stichprobengröße.....	47
Tabelle 4.5.: Ergebnisse des Parameterscreenings	53
Tabelle 4.6.: Stützstellen der Faktoren.....	53
Tabelle A1.1.: Tafel mit den Quantilen der Kolmogorov-Smirnov-Teststatistik [Dul18].....	76
Tabelle A2.2.: Kolmogorov-Smirnov-Test für die Rutsche unten (RU).....	78
Tabelle A2.3.: Kolmogorov-Smirnov-Test für die Umlenkung.....	79

ANHANG

Anhang A1 – Quantile der Kolmogorov-Smirnov-Teststatistik

In Tabelle A1.1 sind die kritischen Werte für der Vergleich der Teststatistik des Kolmogorov-Smirnov-Tests angeführt.

Tabelle A1.1.: Tafel mit den Quantilen der Kolmogorov-Smirnov-Teststatistik [Dul18]

p	0,8	0,9	0,92	0,95	0,96	0,98	0,99
$n = 1$	0,900	0,950	0,960	0,975	0,980	0,990	0,995
2	0,684	0,776	0,800	0,842	0,859	0,900	0,929
3	0,565	0,636	0,658	0,708	0,729	0,785	0,829
4	0,493	0,565	0,585	0,624	0,641	0,689	0,734
5	0,447	0,509	0,527	0,563	0,580	0,627	0,669
6	0,410	0,468	0,485	0,519	0,534	0,577	0,617
7	0,381	0,436	0,452	0,483	0,497	0,538	0,576
8	0,358	0,410	0,425	0,454	0,468	0,507	0,542
9	0,339	0,387	0,402	0,430	0,443	0,480	0,513
10	0,323	0,369	0,382	0,409	0,421	0,457	0,489
11	0,308	0,352	0,365	0,391	0,403	0,437	0,468
12	0,296	0,338	0,351	0,375	0,387	0,419	0,449
13	0,285	0,325	0,338	0,361	0,372	0,404	0,432
14	0,275	0,314	0,326	0,349	0,359	0,390	0,418
15	0,266	0,304	0,315	0,338	0,348	0,377	0,404
16	0,258	0,295	0,306	0,327	0,337	0,366	0,392
17	0,250	0,286	0,297	0,318	0,327	0,355	0,381
18	0,244	0,279	0,289	0,309	0,319	0,346	0,371
19	0,237	0,271	0,281	0,301	0,310	0,337	0,361
20	0,232	0,265	0,275	0,294	0,303	0,329	0,352
21	0,226	0,259	0,268	0,287	0,296	0,321	0,344
22	0,221	0,253	0,262	0,281	0,289	0,314	0,337
23	0,216	0,247	0,257	0,275	0,283	0,307	0,330
24	0,212	0,242	0,251	0,269	0,277	0,301	0,323
25	0,208	0,238	0,246	0,264	0,272	0,295	0,317
26	0,204	0,233	0,242	0,259	0,267	0,290	0,311
27	0,200	0,229	0,237	0,254	0,262	0,284	0,305
28	0,197	0,225	0,233	0,250	0,257	0,279	0,300
29	0,193	0,221	0,229	0,246	0,253	0,275	0,295
30	0,190	0,218	0,226	0,242	0,249	0,270	0,290
31	0,187	0,214	0,222	0,238	0,245	0,266	0,285
32	0,184	0,211	0,219	0,234	0,241	0,262	0,281
33	0,182	0,208	0,215	0,231	0,238	0,258	0,277
34	0,179	0,205	0,212	0,227	0,234	0,254	0,273
35	0,177	0,202	0,209	0,224	0,231	0,251	0,269
36	0,174	0,199	0,206	0,221	0,228	0,247	0,265
37	0,172	0,196	0,204	0,218	0,225	0,244	0,262
38	0,170	0,194	0,201	0,215	0,222	0,241	0,258
39	0,168	0,191	0,199	0,213	0,219	0,238	0,255
40	0,165	0,189	0,196	0,210	0,216	0,235	0,252
$n > 40$	$\frac{1,07}{\sqrt{n}}$	$\frac{1,22}{\sqrt{n}}$	$\frac{1,27}{\sqrt{n}}$	$\frac{1,36}{\sqrt{n}}$	$\frac{1,40}{\sqrt{n}}$	$\frac{1,52}{\sqrt{n}}$	$\frac{1,63}{\sqrt{n}}$

Anhang A2 – Kolmogorov-Smirnov-Test der Submodelle

In Tabelle A2. wird der Kolmogorov-Smirnov-Test für die Rutsche oben (RO) durchgeführt. Dabei befinden sich im oberen Abschnitt die dazugehörigen Messwerte und darunter die Auswertung des Tests.

Tabelle A2.1.: Kolmogorov-Smirnov-Test für die Rutsche oben (RO)

Zeit RO [ms]	Rang	(Rang-1)/n	tats.kum.Anteil [ms]	Differenz [ms]
1008	1	0,000	0,004	0,004
1040	2	0,029	0,071	0,043
1044	3	0,057	0,094	0,036
1048	4	0,086	0,121	0,035
1052	5	0,114	0,153	0,039
1052	6	0,143	0,153	0,010
1060	7	0,171	0,233	0,061
1060	8	0,200	0,233	0,033
1064	9	0,229	0,280	0,052
1064	10	0,257	0,280	0,023
1068	11	0,286	0,332	0,046
1068	12	0,314	0,332	0,018
1068	13	0,343	0,332	0,011
1072	14	0,371	0,387	0,016
1072	15	0,400	0,387	0,013
1072	16	0,429	0,387	0,041
1076	17	0,457	0,445	0,012
1076	18	0,486	0,445	0,041
1080	19	0,514	0,503	0,011
1084	20	0,543	0,562	0,019
1084	21	0,571	0,562	0,009
1088	22	0,600	0,619	0,019
1088	23	0,629	0,619	0,009
1092	24	0,657	0,674	0,017
1092	25	0,686	0,674	0,012
1096	26	0,714	0,725	0,011
1100	27	0,743	0,772	0,029
1100	28	0,771	0,772	0,001
1100	29	0,800	0,772	0,028
1108	30	0,829	0,851	0,023
1112	31	0,857	0,883	0,026
1116	32	0,886	0,909	0,024
1124	33	0,914	0,949	0,034
1132	34	0,943	0,973	0,030
1132	35	0,971	0,973	0,002
Anzahl [#]				
35				
Mittelwert [ms]				
1079,771429				
Standardabweichung [ms]				
27,11097641				
Max. Abweichung = Teststatistik				
0,061487356				
Kritischer Wert für $\alpha=0,05$				
0,22424				
Auswertung				
Teststatistik < kritischer Wert				
Nullhypothese kann nicht verworfen werden - Normalverteilung liegt vor				

7.Anhang

In Tabelle A2.2 wird der Kolmogorov-Smirnov-Test für die Rutsche unten (RU) durchgeführt. Dabei befinden sich im oberen Abschnitt die dazugehörigen Messwerte und darunter die Auswertung des Tests.

Tabelle A2.2.: Kolmogorov-Smirnov-Test für die Rutsche unten (RU)

Zeit RU [ms]	Rang	(Rang-1)/n	tats.kum.Anteil [ms]	Differenz [ms]
820	1	0,000	0,027	0,027
840	2	0,029	0,063	0,034
844	3	0,057	0,073	0,016
848	4	0,086	0,085	0,001
848	5	0,114	0,085	0,029
852	6	0,143	0,098	0,045
864	7	0,171	0,146	0,025
868	8	0,200	0,165	0,035
876	9	0,229	0,207	0,021
880	10	0,257	0,231	0,026
888	11	0,286	0,282	0,004
896	12	0,314	0,338	0,024
896	13	0,343	0,338	0,005
896	14	0,371	0,338	0,033
904	15	0,400	0,398	0,002
908	16	0,429	0,429	0,000
908	17	0,457	0,429	0,028
920	18	0,486	0,524	0,038
932	19	0,514	0,617	0,103
936	20	0,543	0,647	0,104
936	21	0,571	0,647	0,075
940	22	0,600	0,676	0,076
940	23	0,629	0,676	0,047
956	24	0,657	0,781	0,124
956	25	0,686	0,781	0,095
956	26	0,714	0,781	0,066
960	27	0,743	0,803	0,061
960	28	0,771	0,803	0,032
964	29	0,800	0,825	0,025
964	30	0,829	0,825	0,004
976	31	0,857	0,879	0,022
980	32	0,886	0,895	0,009
984	33	0,914	0,908	0,006
984	34	0,943	0,908	0,034
1016	35	0,971	0,975	0,004
Anzahl [#]				
35				
Mittelwert [ms]				
917,0285714				
Standardabweichung [ms]				
50,3198509				
Max. Abweichung = Teststatistik				
0,124				
Kritischer Wert für $\alpha=0,05$				
0,22424				
Auswertung				
Teststatistik < kritischer Wert				
Nullhypothese kann nicht verworfen werden - Normalverteilung liegt vor				

7.Anhang

In Tabelle A2.3 wird der Kolmogorov-Smirnov-Test für die Umlenkung durchgeführt. Dabei befinden sich im oberen Abschnitt die dazugehörigen Messwerte und darunter die Auswertung des Tests.

Tabelle A2.3.: Kolmogorov-Smirnov-Test für die Umlenkung

Zeit Umlenkung[ms]	Rang	(Rang-1)/n	tats.kum.Anteil [ms]	Differenz [ms]
180	1	0,000	0,027	0,027
184	2	0,029	0,032	0,004
212	3	0,057	0,092	0,035
212	4	0,086	0,092	0,006
232	5	0,114	0,169	0,055
236	6	0,143	0,189	0,046
240	7	0,171	0,210	0,038
244	8	0,200	0,232	0,032
244	9	0,229	0,232	0,003
248	10	0,257	0,255	0,002
252	11	0,286	0,279	0,006
252	12	0,314	0,279	0,035
256	13	0,343	0,305	0,038
260	14	0,371	0,331	0,040
264	15	0,400	0,359	0,041
268	16	0,429	0,387	0,042
272	17	0,457	0,416	0,041
284	18	0,486	0,504	0,019
288	19	0,514	0,534	0,020
292	20	0,543	0,563	0,020
304	21	0,571	0,649	0,078
308	22	0,600	0,676	0,076
308	23	0,629	0,676	0,048
312	24	0,657	0,702	0,045
312	25	0,686	0,702	0,017
312	26	0,714	0,702	0,012
316	27	0,743	0,728	0,015
316	28	0,771	0,728	0,044
328	29	0,800	0,797	0,003
332	30	0,829	0,817	0,012
332	31	0,857	0,817	0,040
356	32	0,886	0,912	0,026
372	33	0,914	0,950	0,036
376	34	0,943	0,958	0,015
416	35	0,971	0,993	0,022
Anzahl [#]				
35				
Mittelwert [ms]				
283				
Standardabweichung [ms]				
53,74574441				
Max. Abweichung = Teststatistik				
0,078				
Kritischer Wert für $\alpha=0,05$				
0,22424				
Auswertung				
Teststatistik < kritischer Wert				
Nullhypothese kann nicht verworfen werden - Normalverteilung liegt vor				

Anhang B1 – Parameterscreening

In diesem Abschnitt wird der Matlab Programmcode des Parameterscreenings dargestellt.

```

clear all; close all;
%% Import data from spreadsheet
% Script for importing data from the following spreadsheet:
%
%   Workbook: C:\Users\prkl994\Desktop\FFS_Versuche_Masterarbeit\Screening.xlsx
%   Worksheet: Tabelle1
%
% Auto-generated by MATLAB on 26-Oct-2021 08:42:23

%% Setup the Import Options
opts = spreadsheetImportOptions("NumVariables", 17);

% Specify sheet and range
opts.Sheet = "Tabelle1";
opts.DataRange = "B2:R385";

% Specify column names and types
opts.VariableNames = ["RO", "RORUf", "RORUs", "ROs", "RUs", "StozeitS", "RORUg",
"Rof", "RUF", "Stozeitf", "Gewicht", "mul", "Form", "Rutschenwinkel", "Biegeschlaff",
"Stozeitfs", "Stowinkel"];
opts.SelectedVariableNames = ["RO", "RORUf", "RORUs", "ROs", "RUs", "StozeitS",
"RORUg", "Rof", "RUF", "Stozeitf", "Gewicht", "mul", "Form", "Rutschenwinkel",
"Biegeschlaff", "Stozeitfs", "Stowinkel"];
opts.VariableTypes = ["double", "double", "double", "double", "double", "double",
"double", "double", "double", "double", "double", "double", "double", "double",
"double", "double", "double"];

% Import the data
tbl = readtable("C:\Users\prkl994\Desktop\Masterarbeit\Screening.xlsx", opts,
"UseExcel", false);

%% Convert to output type
RO = tbl.RO;
RORUf = tbl.RORUf;
RORUs = tbl.RORUs;
ROs = tbl.ROs;
RUs = tbl.RUs;
StozeitS = tbl.StozeitS;
RORUg = tbl.RORUg;
Rof = tbl.Rof;
RUF = tbl.RUF;
Stozeitf = tbl.Stozeitf;
Gewicht = tbl.Gewicht;
mul = tbl.mul;
Form = tbl.Form;
Rutschenwinkel = tbl.Rutschenwinkel;
Biegeschlaff = tbl.Biegeschlaff;
Stozeitfs = tbl.Stozeitfs;
Stowinkel = tbl.Stowinkel;

%Darstellung der Effektdiagramme
figure(1);
maineffectsplot(RO, {Gewicht,mul,Form,Rutschenwinkel}, 'varnames', {'Gewicht [g]', 'mu [-
]', 'Form [flach/gefüllt]', 'Rutschenwinkel [°]'})

```

7.Anhang

```
figure(2);
maineffectsplot(RORUf, {Gewicht, mul, Form, Rutschenwinkel, Biegeschlaff, ✓
Stowinkel}, 'varnames', {'Gewicht [g]', 'mu [-]', 'Form [flach/gefüllt]', 'Rutschenwinkel ✓
[°]', 'Biegeschlaff [schlaff/steif]', 'Stoßwinkel [flach/steil]'})

figure(3);
maineffectsplot(RORUs, {Gewicht, mul, Form, Rutschenwinkel, Biegeschlaff, ✓
Stowinkel}, 'varnames', {'Gewicht [g]', 'mu [-]', 'Form [flach/gefüllt]', 'Rutschenwinkel ✓
[°]', 'Biegeschlaff [schlaff/steif]', 'Stoßwinkel [flach/steil]'})

figure(4);
maineffectsplot(Stozeitfs, {Gewicht, mul, Form, Rutschenwinkel, Biegeschlaff, ✓
Stowinkel}, 'varnames', {'Gewicht [g]', 'mu [-]', 'Form [flach/gefüllt]', 'Rutschenwinkel ✓
[°]', 'Biegeschlaff [schlaff/steif]', 'Stoßwinkel [flach/steil]'})
```


Anhang B2 – Korrelationsanalyse

In diesem Abschnitt wird der Matlab Programmcode Korrelationsanalyse dargestellt.

```

clear all; close all;
%% Import data from spreadsheet
% Script for importing data from the following spreadsheet:
%
%   Workbook: C:\Users\prk1994\Desktop\FFS_Versuche_MaSterarbeit\Screening.xlsx
%   Worksheet: Tabelle1
%
% Auto-generated by MATLAB on 26-Oct-2021 08:42:23

%% Setup the Import Options
opts = spreadsheetImportOptions("NumVariables", 17);

% Specify sheet and range
opts.Sheet = "Tabelle1";
opts.DataRange = "B2:R385";

% Specify column names and types
opts.VariableNames = ["RO", "RORUf", "RORUs", "ROs", "RUs", "StozeitS", "RORUg",
"Rof", "RUF", "Stozeitf", "Gewicht", "mul", "Form", "Rutschenwinkel", "Biegeschlaff",
"Stozeitfs", "Stowinkel"];
opts.SelectedVariableNames = ["RO", "RORUf", "RORUs", "ROs", "RUs", "StozeitS",
"RORUg", "Rof", "RUF", "Stozeitf", "Gewicht", "mul", "Form", "Rutschenwinkel",
"Biegeschlaff", "Stozeitfs", "Stowinkel"];
opts.VariableTypes = ["double", "double", "double", "double", "double", "double",
"double", "double", "double", "double", "double", "double", "double", "double",
"double", "double", "double"];

% Import the data
tbl = readtable("C:\Users\prk1994\Desktop\FFS_Versuche_MaSterarbeit\Screening.xlsx",
opts, "UseExcel", false);

%% Convert to output type
RO = tbl.RO;
RORUf = tbl.RORUf;
RORUs = tbl.RORUs;
ROs = tbl.ROs;
RUs = tbl.RUs;
StozeitS = tbl.StozeitS;
RORUg = tbl.RORUg;
Rof = tbl.Rof;
RUF = tbl.RUF;
Stozeitf = tbl.Stozeitf;
Gewicht = tbl.Gewicht;
mul = tbl.mul;
Form = tbl.Form;
Rutschenwinkel = tbl.Rutschenwinkel;
Biegeschlaff = tbl.Biegeschlaff;
Stozeitfs = tbl.Stozeitfs;
Stowinkel = tbl.Stowinkel;

%Korrealtion der Parameter für die Rutsche RO

corr_RO_Gewicht=corrcoef(RO,Gewicht);
corr_RO_mul=corrcoef(RO,mul);
corr_RO_Form=corrcoef(RO,Form);

```

7.Anhang

```
corr_RO_Rutschenwinkel=corrcoef(RO,Rutschenwinkel);
corr_RO_Biegeschlafheit=corrcoef(RO,Biegeschlaff);
corr_RO_Stowinkel=corrcoef(RO,Stowinkel);

%Ausgabe des Korrelationskoeffizienten für die Rutsche RO
corr_RO_Gewicht_c=corr_RO_Gewicht(1,2);

% plot(eins, corr_RO_Gewicht_c);

corr_RO_mul_c=corr_RO_mul(1,2);
corr_RO_Form_c=corr_RO_Form(1,2);
corr_RO_Rutschenwinkel_c=corr_RO_Rutschenwinkel(1,2);
corr_RO_Biegeschlafheit_c=corr_RO_Biegeschlafheit(1,2);
corr_RO_Stowinkel_c=corr_RO_Stowinkel(1,2);

%Korrealtion der Parameter für den gesamten Versuch mit flachen Stoß
corr_RORUf_Gewicht=corrcoef(RORUf,Gewicht);
corr_RORUf_mul=corrcoef(RORUf,mul);
corr_RORUf_Form=corrcoef(RORUf,Form);
corr_RORUf_Rutschenwinkel=corrcoef(RORUf,Rutschenwinkel);
corr_RORUf_Biegeschlafheit=corrcoef(RORUf,Biegeschlaff);
corr_RORUf_Stowinkel=corrcoef(RORUf,Stowinkel);

%Ausgabe des Korrelationskoeffizienten mit flachen Stoß
corr_RORUf_Gewicht_c=corr_RORUf_Gewicht(1,2);
corr_RORUf_mul_c=corr_RORUf_mul(1,2);
corr_RORUf_Form_c=corr_RORUf_Form(1,2);
corr_RORUf_Rutschenwinkel_c=corr_RORUf_Rutschenwinkel(1,2);
corr_RORUf_Biegeschlafheit_c=corr_RORUf_Biegeschlafheit(1,2);
corr_RORUf_Stowinkel_c=corr_RORUf_Stowinkel(1,2);

%Korrealtion der Parameter für den gesamten Versuch mit steilen Stoß
corr_RORUs_Gewicht=corrcoef(RORUs,Gewicht);
corr_RORUs_mul=corrcoef(RORUs,mul);
corr_RORUs_Form=corrcoef(RORUs,Form);
corr_RORUs_Rutschenwinkel=corrcoef(RORUs,Rutschenwinkel);
corr_RORUs_Biegeschlafheit=corrcoef(RORUs,Biegeschlaff);
corr_RORUs_Stowinkel=corrcoef(RORUs,Stowinkel);

%Ausgabe des Korrelationskoeffizienten mit steilen Stoß
corr_RORUs_Gewicht_c=corr_RORUs_Gewicht(1,2);
corr_RORUs_mul_c=corr_RORUs_mul(1,2);
corr_RORUs_Form_c=corr_RORUs_Form(1,2);
corr_RORUs_Rutschenwinkel_c=corr_RORUs_Rutschenwinkel(1,2);
corr_RORUs_Biegeschlafheit_c=corr_RORUs_Biegeschlafheit(1,2);
corr_RORUs_Stowinkel_c=corr_RORUs_Stowinkel(1,2);

%Korrealtion der Parameter nur für den Stoßvorgang
corr_Stozeitfs_Gewicht=corrcoef(Stozeitfs,Gewicht);
corr_Stozeitfs_mul=corrcoef(Stozeitfs,mul);
corr_Stozeitfs_Form=corrcoef(Stozeitfs,Form);
corr_Stozeitfs_Rutschenwinkel=corrcoef(Stozeitfs,Rutschenwinkel);
corr_Stozeitfs_Biegeschlafheit=corrcoef(Stozeitfs,Biegeschlaff);
corr_Stozeitfs_Stowinkel=corrcoef(Stozeitfs,Stowinkel);
```

7.Anhang

```
%Ausgabe des Korrelationskoeffizienten für den Stoßvorgang
corr_Stozeitfs_Gewicht_c=corr_Stozeitfs_Gewicht(1,2);
corr_Stozeitfs_mul_c=corr_Stozeitfs_mul(1,2);
corr_Stozeitfs_Form_c=corr_Stozeitfs_Form(1,2);
corr_Stozeitfs_Rutschenwinkel_c=corr_Stozeitfs_Rutschenwinkel(1,2);
corr_Stozeitfs_Biegeschlafheit_c=corr_Stozeitfs_Biegeschlafheit(1,2);
corr_Stozeitfs_Stowinkel_c=corr_Stozeitfs_Stowinkel(1,2);

%Matrix der Korrelationskoeffizienten
Korrelationskoeffizienten=[corr_RO_Gewicht_c corr_RO_mul_c corr_RO_Form_c ✓
corr_RO_Rutschenwinkel_c corr_RO_Biegeschlafheit_c corr_RO_Stowinkel_c]
```

Anhang C – Predictionmodell erzeugen

In Anhang C wird exemplarisch die Anwendung des Predictionmodells für ein Modell mit zwei Stützstellen je Parameter angeführt. Die Anwendung für drei und vier Stützstellen unterscheidet sich nur durch die Verwendung eines anderen Teachingsets. Der Aufbau hingegen bleibt ident.

```

clc; clear all;
%% Import data from spreadsheet
% Script for importing data from the following spreadsheet:
%
%   Workbook: C:\Users\prkl1994\Desktop\Masterarbeit\Matlab\Teachingset2STS.xlsx
%   Worksheet: RO2RU2_Teachingset2STS
%
% Auto-generated by MATLAB on 26-Jan-2022 11:45:51

%% Setup the Import Options
opts = spreadsheetImportOptions("NumVariables", 9);

% Specify sheet and range
opts.Sheet = "RO2RU2_Teachingset2STS";
opts.DataRange = "A2:I33";
% Specify column names and types
opts.VariableNames = ["Gewicht", "Rutschenwinkel", "Stosswinkel", "Var4",
"GesamtzeitR22", "ZeitRO_2", "Var7", "ZeitRU_2", "ZeitUmlenkung"];
opts.SelectedVariableNames = ["Gewicht", "Rutschenwinkel", "Stosswinkel",
"GesamtzeitR22", "ZeitRO_2", "ZeitRU_2", "ZeitUmlenkung"];
opts.VariableTypes = ["double", "double", "double", "char", "double", "double",
"char", "double", "double"];
opts = setvaropts(opts, [4, 7], "WhitespaceRule", "preserve");
opts = setvaropts(opts, [4, 7], "EmptyFieldRule", "auto");

% Setup rules for import
opts.ImportErrorRule = "omitrow";
opts.MissingRule = "omitrow";
opts = setvaropts(opts, [1, 2, 3, 5, 6, 8, 9], "TreatAsMissing", '');

% Import the data
RO2RU2_Teachingset = readtable("C:\
\Users\prkl1994\Desktop\Masterarbeit\Matlab\Teachingset2STS.xlsx", opts, "UseExcel",
false);

%% Import data from spreadsheet
% Script for importing data from the following spreadsheet:
%
%   Workbook: C:\Users\prkl1994\Desktop\Masterarbeit\Matlab\Teachingset2STS.xlsx
%   Worksheet: RO2RU2_Validierungsset2STS
%
% Auto-generated by MATLAB on 26-Jan-2022 11:51:34

%% Setup the Import Options
opts = spreadsheetImportOptions("NumVariables", 9);

% Specify sheet and range
opts.Sheet = "RO2RU2_Validierungsset2STS";
opts.DataRange = "A2:I91";

% Specify column names and types
opts.VariableNames = ["Gewicht", "Rutschenwinkel", "Stosswinkel", "Var4",
"GesamtzeitR22", "ZeitRO_2", "Var7", "ZeitRU_2", "ZeitUmlenkung"];
opts.SelectedVariableNames = ["Gewicht", "Rutschenwinkel", "Stosswinkel",
"GesamtzeitR22", "ZeitRO_2", "ZeitRU_2", "ZeitUmlenkung"];

```

7.Anhang

```
opts.VariableTypes = ["double", "double", "double", "char", "double", "double", "char", "double", "double"];
opts = setvaropts(opts, [4, 7], "WhitespaceRule", "preserve");
opts = setvaropts(opts, [4, 7], "EmptyFieldRule", "auto");

% Setup rules for import
opts.ImportErrorRule = "omitrow";
opts.MissingRule = "omitrow";
opts = setvaropts(opts, [1, 2, 3, 5, 6, 8, 9], "TreatAsMissing", '');

% Importieren der Validierungsdaten
RO2RU2_Validierungsset = readtable("C:\Users\prk1994\Desktop\Masterarbeit\Matlab\Teachingset2STS.xlsx", opts, "UseExcel", false);

%% Import data from spreadsheet
% Script for importing data from the following spreadsheet:
%
%   Workbook: C:\Users\prk1994\Desktop\Masterarbeit\Matlab\Teachingset2STS.xlsx
%   Worksheet: RO2RU2_2STS
%
% Auto-generated by MATLAB on 26-Jan-2022 11:26:46

%% Import data from spreadsheet
% Script for importing data from the following spreadsheet:
%
%   Workbook: C:\Users\prk1994\Desktop\Masterarbeit\Matlab\Teachingset2STS.xlsx
%   Worksheet: RO2RU2_2STS
%
% Auto-generated by MATLAB on 26-Jan-2022 11:52:30

%% Setup the Import Options
opts = spreadsheetImportOptions("NumVariables", 9);

% Specify sheet and range
opts.Sheet = "RO2RU2_2STS";
opts.DataRange = "A2:I271";

% Specify column names and types
opts.VariableNames = ["Gewicht", "Rutschenwinkel", "Stosswinkel", "Var4", "GesamtzeitR22", "ZeitRO_2", "Var7", "ZeitRU_2", "ZeitUmlenkung"];
opts.SelectedVariableNames = ["Gewicht", "Rutschenwinkel", "Stosswinkel", "GesamtzeitR22", "ZeitRO_2", "ZeitRU_2", "ZeitUmlenkung"];
opts.VariableTypes = ["double", "double", "double", "char", "double", "double", "char", "double", "double"];
opts = setvaropts(opts, [4, 7], "WhitespaceRule", "preserve");
opts = setvaropts(opts, [4, 7], "EmptyFieldRule", "auto");

% Setup rules for import
opts.ImportErrorRule = "omitrow";
opts.MissingRule = "omitrow";
opts = setvaropts(opts, [1, 2, 3, 5, 6, 8, 9], "TreatAsMissing", '');

% Import the data
```

7.Anhang

```
RO2RU2_2STS = readtable("C:\Users\prkl994\Desktop\Masterarbeit\Matlab\Teachingset2STS.
xlsx", opts, "UseExcel", false);
load("RO2_predict_2STS.mat", "RO2_predict_2STS")
load("RU2_predict_2STS.mat", "RU2_predict_2STS")
load("Umlenkung_predict_2STS.mat", "Umlenkung_predict_2STS")
load("RO2RU2_predict_2STS.mat", "RO2RU2_predict_2STS")

% %Tabelle mit Validierungsdaten an das Predictionmodell übergeben
testset=RO2RU2_2STS;

% %Vorhergestagten Durchlaufzeiten für die Division als Tabelle definieren
RO2_time=table(RO2_predict_2STS.predictFcn(testset));
RU2_time=table(RU2_predict_2STS.predictFcn(testset));
Umlenkung_time=table(Umlenkung_predict_2STS.predictFcn(testset));
RO2RU2_time=table(RO2RU2_predict_2STS.predictFcn(testset));

% Prediction Ergebnisse in Tabelle laden
RO2_val=RO2RU2_2STS(:,5);
RU2_val=RO2RU2_2STS(:,6);
Umlenkung_val=RO2RU2_2STS(:,7);
RO2RU2_val=RO2RU2_2STS(:,4);

%Berechnung der Abweichung von Validierungsdaten zu Predictionsergebnissen
Deviation_RO2 = RO2_time(:,1)./RO2_val(:,1);
Deviation_RU2 = RU2_time(:,1)./RU2_val(:,1);
Deviation_Umlenkung = Umlenkung_time(:,1)./Umlenkung_val(:,1);
Deviation_RO2RU2=RO2RU2_time(:,1)./RO2RU2_val(:,1);
Deviation=table(Deviation_RO2,Deviation_RU2,Deviation_Umlenkung,Deviation_RO2RU2)
RO=table(RO2_val,RO2_time);

%Erstellen der Histogramme mit den Abweichungen
figure(1);
edges=linspace(0.8, 1.2, 30);
histogram(Deviation_RO2, 'BinEdges', edges);
grid on;
figure(2);
edges=linspace(0.8, 1.2, 30);
histogram(Deviation_RU2, 'BinEdges', edges)
edges=linspace(0.8, 1.2, 30);
figure(3);
histogram(Deviation_Umlenkung, 'BinEdges', edges)
figure(4);
histogram(Deviation_RO2RU2, 'BinEdges', edges)

%Abweichungen in ein Excelsheet laden
filename = 'Deviation2STS.xlsx';
writetable(Deviation,filename,'Sheet',1,'Range','D1')
```