
Entwicklung einer populations- basierten Validierungsmethodik für FE-Menschmodelle

Sonja Andrea Schneider



München 2022

Aus dem Institut für Rechtsmedizin

Institut der Universität München

Vorstand: Prof. Dr. med. Matthias Graw

Entwicklung einer populations- basierten Validierungsmethodik für FE-Menschmodelle

Dissertation

zum Erwerb des Doktorgrades der Humanbiologie
an der Medizinischen Fakultät der
Ludwig-Maximilians-Universität zu München

vorgelegt von

Sonja Andrea Schneider

aus Ulm

2022

Mit Genehmigung der Medizinischen Fakultät
der Universität München

Berichterstatter: Prof. Dr. rer. biol. hum. Dipl.-Ing.
Steffen Peldschus

Mitberichterstatter: PD Dr. Michael Hubig
PD Dr. Matthias Woiczinski

Mitbetreuung durch den
promovierten Mitarbeiter:

Dekan: Prof. Dr. med. Thomas Gudermann

Tag der mündlichen Prüfung:
04.07.2022

Inhaltsverzeichnis

Inhaltsverzeichnis	I
Zusammenfassung	V
Abstract	VII
Abbildungsverzeichnis	IX
Tabellenverzeichnis	XIII
Abkürzungsverzeichnis	XV
1 Einleitung	1
1.1 Einsatz von Menschmodellen in der Automobilindustrie	1
1.1.1 Vorteile der Menschmodelle gegenüber anthropomorphen Messsystemen	1
1.1.2 Existierende Menschmodelle und Teilkörpermodelle	3
1.1.3 Neue Entwicklungen – Parametrisierte Menschmodelle	5
1.2 Validierung von Menschmodellen	7
1.2.1 Begriffsdefinitionen	7
1.2.2 Validierung von Menschmodellen	8
1.2.3 Übersicht Validierung von THUMS Version 4 AM 50 und GHBMC AM 50	9
1.2.4 Konventioneller korridor-basierter Validierungsansatz.....	10
1.2.5 Validierungsansätze für parametrisierte Modelle	11
1.2.6 Limitationen aktueller Validierungsmethoden	12
2 Zielsetzung	15
2.1 Zielsetzung und wissenschaftliche Fragestellung	15
2.2 Struktur der Dissertation.....	16
3 Entwicklung des numerischen Modells eines Referenzexperiments	19
3.1 Einleitung.....	19
3.1.1 Zielsetzung	19
3.1.2 Auswahl des dynamischen 3-Punkt-Biegeversuchs des Femurs nach Forman et al.....	19
3.1.3 Besonderheit in den experimentellen Kurvenverläufen bei Forman et al.	21
3.2 Methoden.....	22
3.2.1 Numerische Modellerstellung des Ausgangsmodells für die Sensitivitätsanalysen	22
3.2.2 Methodik für die Definition der Fehler- und Sensitivitätsanalysen.....	23
3.2.3 Sensitivitätsanalysen bezüglich Modellierungsannahmen	25
3.2.4 Sensitivitätsanalysen zur Festlegung der Geometrie der Auflager.....	26
3.2.5 Sensitivitätsanalysen bezüglich Schaumdefinition	28

3.2.6	Sensitivitätsanalysen bezüglich Knochenpositionierung.....	28
3.2.7	Sensitivitätsanalysen bezüglich der Positionierung des Versuchsaufbaus	29
3.2.8	Einfluss unterschiedlicher Knochenmodelle auf den Simulationsaufbau	30
3.2.9	Vergleich des entwickelten Simulationsmodells mit der Realität.....	31
3.3	Ergebnisse.....	31
3.3.1	Sensitivitätsanalysen bezüglich Modellierungsannahmen	31
3.3.2	Sensitivitätsanalysen zur Festlegung der geometrischen Parameter	32
3.3.3	Sensitivitätsanalysen bezüglich der Schaumdefinition	33
3.3.4	Sensitivitätsanalysen bezüglich der Knochenpositionierung	33
3.3.5	Sensitivitätsanalysen bezüglich der Positionierung des Versuchsaufbaus	33
3.3.6	Einfluss unterschiedlicher Knochenmodelle auf die Modellvorhersage.....	34
3.3.7	Vergleich des entwickelten Simulationsmodells mit der Realität.....	34
3.4	Diskussion	35
3.4.1	Auswahlkriterien des Referenzexperiments	35
3.4.2	Definition und Durchführung der Sensitivitätsanalysen	36
3.4.3	Modellierung des Versuchsaufbaus	37
3.4.4	Sensitivitätsanalysen zur Festlegung der geometrischen Parameter	38
3.4.5	Sensitivitätsanalysen bezüglich der Schaumdefinition	39
3.4.6	Sensitivitätsanalysen bezüglich der Knochenpositionierung	40
3.4.7	Sensitivitätsanalysen bezüglich Positionierung des Versuchsaufbaus.....	40
3.4.8	Einfluss unterschiedlicher Knochenmodelle auf die Modellvorhersage.....	41
3.4.9	Problematik bei Literaturvalidierung	42
4	Prüfung der Eignung von Clusterverfahren für den Vergleich biomechanischer Datensätze	45
4.1	Mathematische Clusteranalysen	45
4.1.1	Einleitung.....	45
4.1.2	Klassifizierung von Clusterverfahren	46
4.1.3	Prozess eines Clusterverfahrens für Zeitreihen	48
4.1.4	Verwendete Metriken zur Distanzmessung	49
4.1.5	Verwendete Fusionierungsalgorithmen.....	50
4.1.6	Bestimmung der Clusteranzahl	51
4.1.7	Zielsetzung des Kapitels	51
4.2	Methoden	52
4.2.1	Übersicht der Vorgehensweise dieses Kapitels.....	52
4.2.2	Kriterien für die synthetische Datensatzgenerierung	53
4.2.3	Beschreibung der Testdatensätze	54

4.2.4	Durchführung der Clusteranalyse	57
4.2.5	Bewertungskriterien der Clusteranalyse für die Testdatensätze	58
4.3	Ergebnisse	60
4.3.1	Testdatensätze I–V	60
4.3.2	Testdatensätze VI–VIII	74
4.4	Diskussion	77
4.4.1	Methodik	77
4.4.2	Ergebnisse der Testdatensätze	84
4.4.3	Zusammenfassung und Ausblick	86
5	Entwicklung einer populationsbasierten Validierungsmethodik	89
5.1	Einleitung und Zielsetzung	89
5.2	Methoden	90
5.2.1	Vorgehensweise in diesem Kapitel	90
5.2.2	Entwicklung der populationsbasierten Validierungsmethode	91
5.2.3	Definition der stochastischen Population	99
5.2.4	Erzeugung der stochastischen Population	106
5.2.5	Vorstellung der Konzeptdatensätze zur Überprüfung des Validierungskonzeptes	108
5.3	Ergebnisse	110
5.3.1	Populationsbasierte Validierungsmethode mit Konzeptdatensatz I	110
5.3.2	Populationsbasierte Validierungsmethode mit Konzeptdatensatz II	114
5.4	Diskussion	117
5.4.1	Populationsbasierte Validierungsmethode	117
5.4.2	Populationsdefinition	122
5.4.3	Erzeugung der Population und Definition der Konzeptdatensätze	126
5.4.4	Validierungsergebnisse der Konzeptdatensätze	127
5.4.5	Zusammenfassung und Ausblick	131
6	Schlussfolgerungen und Ausblick	133
	Literaturverzeichnis	135
	Anhang	141
A1	Übersicht der Validierungslastfälle des THUMS-Modells [66]	141
A2	Ergebnisse Sensitivitätsanalysen	143
A2.1	Ergebnisse Sensitivitätsanalysen bezüglich Modellierungsannahmen	143
A2.2	Sensitivitätsanalysen zur Festlegung der geometrischen Parameter	144
A2.3	Sensitivitätsanalysen bezüglich der Schaumdefinition	145
A2.4	Sensitivitätsanalysen bezüglich der Knochenpositionierung	145

A2.5	Sensitivitätsanalysen bezüglich der Positionierung des Versuchsaufbaus	146
A2.6	Einfluss unterschiedlicher Knochenmodelle auf das Simulationssetup.....	147
A2.7	Vergleich des entwickelten Simulationsmodells mit der Realität.....	147
A3	Dendrogramme.....	148
A4	Interne Validierungsindizes	167
A5	Erzeugung LMU-Population	172
A6	LMU-Population – Übersicht Vernetzungsqualität.....	176
A7	Ergebnisse Konzeptpopulation I	177
A8	Ergebnisse Konzeptpopulation II	181
	Danksagung.....	185
	Affidavit	187

Zusammenfassung

Virtuelle FE-Menschmodelle finden heutzutage ihren Einsatz vor allem in der frühen Phase im Fahrzeugentwicklungsprozess zur verletzungsmechanischen Bewertung und Untersuchung von Sicherheitsfunktionen, um die existierenden numerischen Crashtest-Dummys zu ergänzen und zukünftig teilweise zu ersetzen. FE-Menschmodelle werden u.a. gegen Bewegungsmessungen mit Freiwilligen oder Versuche mit postmortalen Testobjekten validiert, um ein menschenähnliches Verhalten dieser Modelle, auch Biofidelität genannt, sicherzustellen. FE-Menschmodelle sind, im Gegensatz zu Crashtest-Dummys, grundsätzlich veränderbar. Mit Verwendung von mesh-morphing Algorithmen oder parametrisierten Modellen kann die Geometrie von FE-Menschmodellen variiert werden. Zudem können auch Beschreibungen der Gewebeeigenschaften verändert werden. Die Veränderung von FE-Menschmodellen ermöglicht beispielsweise die Untersuchung von Populationen mit bestimmten Eigenschaften. Dabei muss jedoch beachtet werden, dass die Modelle mit der Veränderung ihren validierten Bereich verlassen. Um zukünftig den Einsatz stochastischer Simulationen mit FE-Menschmodellen zu ermöglichen, muss Vertrauen in numerische Modellvorhersagen generiert werden, welche nicht direkt mit den experimentellen Testproben eines Validierungsexperimentes übereinstimmen.

Im Rahmen dieser Arbeit wird daher die wissenschaftliche Fragestellung bearbeitet, wie numerische Vorhersagen von Modellen, welche von experimentellen Testproben abweichende charakteristische Eigenschaften aufweisen, validiert werden können. Insbesondere liegt der Fokus auf der Erarbeitung eines Validierungskonzeptes, mit dem die Modellvorhersagen einer gesamten numerischen Population in einem Prozess mit den Ergebnissen eines Validierungsexperimentes verglichen werden können, ohne die einzelnen Kurven detailliert zu untersuchen.

Zu Beginn wird ein Validierungsexperiment aus der Literatur auf lokaler Ebene, ein dynamischer 3-Punkt-Biegeversuch des Femurs, als Referenzlastfall ausgewählt und ein numerisches Abbild des experimentellen Versuchsaufbaus entwickelt. Dieses numerische Abbild des Referenzlastfalls wird als Voraussetzung für die folgenden Schritte detailliert validiert und dient im weiteren Verlauf der Arbeit dazu, das populationsbasierte Validierungskonzept zu testen. Da die genaue numerische Abbildung experimenteller Lasten und Randbedingungen bei derartigen Versuchen bis dato eine große Herausforderung darstellt, wird eine umfassende Sensitivitätsanalyse zu den wichtigsten Parametern durchgeführt. Auf dieser Basis kann das numerische Abbild des Referenzlastfalls für die nachfolgenden Schritte qualifiziert werden. Die Sensitivitätsanalysen fokussieren sich auf die Parameter zur Geometrie- und Massendefinition der Auflager, Parameter zur Materialcharakterisierung des Schaumes um den Impaktor, der Knochenpositionierung, sowie der Positionierung der Auflager relativ zum Knochen und zu dem Impaktor. Die Masse der Auflager hat den größten Einfluss auf die Kurvenform des experimentellen Ergebnisses. Daneben kann im Rahmen der detaillierten Analyse des Experiments auch eine Ausgangsbasis für die Verwendung von Femora-Modellen geschaffen werden. Danach folgt eine Eignungsprüfung von mathematischen Clusteranalyseverfahren und internen Validierungsindizes für den vergleichenden Bewertungsprozess numerischer und experimenteller Ergebnisse. Untersucht werden dabei die Möglichkeiten der objektiven Bestimmung einer geeigneten Clusteranzahl nach der Clusteranalyse für die Anwendung an biomechanischen Datensätzen. Diese Eignungsprüfung wird an vereinfachten Testdatensätzen mit einer geringen Kurvenanzahl durchgeführt, welche charakteristische Eigenschaften verletzungsmechanischer Datensätze beinhalten. Es werden nur hierarchische Fusionierungsalgorithmen der Eignungsprüfung unterzogen.

Die geeigneten Kombinationen aus Clusteranalyseverfahren und internen Validierungsindizes werden in dem populationsbasierten Validierungskonzept verwendet, um die kombinierten Populationsergebnisse, bestehend aus der numerischen und der experimentellen Kurvenschar, zu vergleichen. Das populationsbasierte Validierungskonzept verwendet das Ergebnis der Clusteranalyse, um die Güte der numerischen Modellvorhersagen quantitativ zu bewerten.

Im letzten Schritt wird das populationsbasierte Validierungskonzept an dem numerischen Abbild des Referenzlastfalls getestet. Es wird eine numerische Testpopulation an Femora erstellt. Für diese Population werden die Modellvorhersagen mit dem validierten numerischen Versuchsaufbau des Referenzlastfalls generiert. Die Geometrie der numerischen Femora basiert auf postmortaler Bildgebung (CT) und die Gewebeeigenschaften basieren auf veröffentlichten Werten aus der Literatur. Die numerischen Modellvorhersagen und die publizierten experimentellen Ergebnisse des Referenzlastfalls werden in einen Datensatz zusammengeführt. Durch die Verringerung der Kurvenanzahl dieses Datensatzes wird ein zweiter Datensatz artifiziell generiert, welcher eine abweichende Struktur mit natürlichen Gruppierungen aufweist. An diesen beiden Testdatensätzen wird die populationsbasierte Validierungsmethode abschließend getestet.

Mit der Hilfe von Clusteranalyseverfahren können komplexe Datensätze in eine überschaubare Anzahl an Gruppen aufgeteilt werden. Kurven aus einer Gruppe weisen nur leichte Unterschiede in den Kurvenformen und weiterer knochenspezifischer Parameter, wie z.B. der Biegeresistenz des Knochens, auf. Die resultierenden Gruppen aus einer Clusteranalyse werden in der populationsbasierten Validierungsmethode einzeln analysiert und gegeneinander verglichen. Damit das Ergebnis einer Clusteranalyse an einem biomechanischen Datensatz erfolgreich in der populationsbasierten Validierungsmethode verwendet werden kann, muss sichergestellt werden, dass die numerischen Kurvenformen detailliert mit den experimentellen Kurvenformen übereinstimmen. Wenn signifikante Unterschiede zwischen den experimentellen und den numerischen Kurvenformen existieren würden, würde ein Clusteranalyseprozess die experimentellen und die numerischen Kurven voneinander trennen. In diesem Fall ist keine Bestimmung der Validierungsgüte mit dem populationsbasierten Validierungsprozess möglich. Größere Unterschiede in den Kurvenformen, wie z.B. auftretende Schwingungen, können daher zur Beeinträchtigung der Anwendbarkeit eines solchen Ansatzes führen. Das Validierungskonzept wird in dieser Arbeit mit Testdatensätzen, die eine geringe Kurvenanzahl enthalten, getestet. Diese geringe Kurvenanzahl in den Testdatensätzen ist bewusst gewählt, damit der Validierungsprozess visuell überprüft werden kann. Abschließend wird empfohlen, zukünftige Anstrengungen auf die Automatisierung des populationsbasierten Validierungsprozesses zu legen, um die Analyse von komplexeren und undurchsichtigen Datensätzen zu ermöglichen.

Abstract

Nowadays, numerical FE human body models are mainly used to biomechanically assess and investigate safety features during the early stage in a vehicle development process, complementing or in future partly replacing the existing numerical crash test dummies. FE human body models are validated against experimental test results from biomechanical experiments, performed with volunteers or postmortem human subjects, to create and ensure a human-like behavior of these models, which is also called biofidelity. In contrast to crash test dummies, FE human body models can in principle be modified. The geometry of FE human body models can be modified by using mesh-morphing algorithms or parametrized models. Additionally, the material characterization of tissues can be modified. The modification of FE human body models enables for instance the investigation of populations with certain characteristics. However, one has to keep in mind that the models leave with their modification the validated area where they were intended to be used. To enable the usage of stochastic numerical simulations with FE human body models in the future, one has to create trust in numerical model predictions, which are not directly comparable to the experimental test objects of a validation experiment.

In the scope of this thesis, the investigated scientific research question is therefore, how the predictions of numerical models, with differing characteristic properties to the experimental test objects can be validated. In particular, the focus lies on the elaboration of a validation concept by which the model predictions of a whole numerical population can be compared to the results of a validation experiment in one process, without investigating the single curves in detail.

First, a validation experiment on a local level, a dynamic 3-point bending test of a femur, is selected as the reference load case from literature. A numerical replicate is then created of the experimental test setup. This numerical replicate of the reference load case is the basis to test the population based validation concept in the further course of this thesis and is therefore thoroughly validated. Up to now, one major challenge in the numerical replication of such experimental tests is the exact numerical replication of the experimental loads and boundary conditions. Therefore, a comprehensive sensitivity analysis regarding the most important parameters is performed. This establishes the basis for qualifying the numerical replicate of the reference load case to be used in following steps. The sensitivity analyses focuses on parameters to characterize the geometry and mass definition of the roller supports, the parameters to characterize the foam material around the impactor, the positioning of the bone, as well as the positioning of the roller supports relative to the bone and the impactor. The mass of the roller support has the biggest influence on the curve form of the experimental test result. Besides, in the context of the detailed analysis of the experiments a starting point for the usage of femora models is created.

Then, a qualifying examination of mathematical cluster analysis methods and internal validation indices for the comparative evaluation process of numerical and experimental results is performed. The possibilities of objective determination of a suitable number of clusters after the cluster analysis for the application to biomechanical data sets are examined. This qualifying examination is performed on simplified test datasets containing a small amount of curves and representing characteristic properties of biomechanical datasets. Furthermore, only hierarchical fusion algorithms are subjected to the qualifying examination.

The suitable combinations of cluster analysis method and internal validation indices are used in the population based validation concept to compare the combined population results, consisting of the numerical and the experimental set of curves. The population based validation concept uses the result of the cluster analysis to examine the quality of the numerical model predictions quantitatively.

In the last step, the population based validation concept is tested on the numerical replication of the reference load case. A numerical test population of femora is created. The numerical model predictions are created for this test population with the validated numerical test setup of the reference load case. The geometry of the numerical femora is based on postmortem imaging (CT) and the tissue properties are based on published values from literature. The numerical model predictions and the published experimental test results of the reference load case are combined to one dataset. A second dataset is artificially created by reducing the amount of curves of the combined dataset, which shows a differing structure incorporating natural grouping of curves. Finally, the population based validation method is tested on both test datasets.

Complex datasets can be split up in a manageable number of groups by using cluster analysis methods. Curves within one group only differ slightly regarding their curve forms and further bone specific parameters, as for instance the bending resistance of the bone. In the population based validation method, the resulting groups from a cluster analysis are analyzed on their own and compared to each other. In order to successfully use the result of a cluster analysis of a biomechanical dataset in the population based validation method, it has to be ensured that the numerical curve forms match detailed to the experimental ones. If significant differences between the experimental and the numerical curve forms were present, the cluster analysis process would separate the experimental curves from the numerical ones. In this case, it would not be possible to determine the quality of the validation with the population based validation process. Significant differences in curve forms, as e.g. occurring vibrations, can impair the applicability of such an approach.

The validation concept is tested with test datasets incorporating a small amount of curves. This small amount of curves is chosen on purpose to enable a visual verification of the validation process. In conclusion, it is recommended to put future effort in the automation of the population based validation process to enable the analysis of complex and obscure datasets.

Abbildungsverzeichnis

Abbildung 1 – Von links nach rechts: Finite-Elemente (Abkürzung: FE) Modelle des ATD HYBRID III für frontale Belastung, des ATD SID II für seitliche Belastung und des THUMS™-Menschmodells [69]	1
Abbildung 2 – Neuartige Sitzpositionen in einem Konzeptfahrzeug [105]	2
Abbildung 3 – Überblick über aktuelle FE-Ganzkörper-Menschmodelle [113]	4
Abbildung 4 – Zunahme des Detaillierungsgrades der THUMS™-Modelle von Version 1 bis Version 5 [60]	4
Abbildung 5 – Entwickelte detaillierte und vereinfachte Modelle des GHBMC-Modells [22]	5
Abbildung 6 – Notwendige Arbeitsschritte für die Entwicklung eines parametrisierten Modells [113]	6
Abbildung 7 – Kategorisierung von Validierungslastfällen nach Fuchs mit Beispielen [40]	9
Abbildung 8 – Anzahl der Validierungslastfälle für die Modellentwicklung des GHBMC AM 50 [24]	9
Abbildung 9 – Korridor-basierter Validierungsansatz [93]	11
Abbildung 10 – Subjekt-Spezifischer Validierungs- und Kalibrierungsansatz [93].....	12
Abbildung 11 – Vorteil stochastischer Simulationen gegenüber den drei existierenden Perzentilen der ATDs	15
Abbildung 12 – Struktur der Dissertation und Verknüpfungen der einzelnen Kapitel	17
Abbildung 13 – Experimenteller Versuchsaufbau [38] überlagert mit numerischem Ausgangsmodell	22
Abbildung 14 – Knochenpositionierung.....	23
Abbildung 15 – Analysen für die kritische Beurteilung der Güte des numerischen Versuchsaufbaus	24
Abbildung 16 – Skizze starrer vs. deformierbarer Versuchsaufbau.....	25
Abbildung 17 – Variantendefinitionen für Schwerpunktsverlagerung	27
Abbildung 18 – Variation des Auflageradius	27
Abbildung 19 – Variation Auflagerblockbreite (Abkürzung: RBB).....	27
Abbildung 20 – Variation Drehpunkt	28
Abbildung 21 – Knochenschwerpunkt und Achsendefinition	28
Abbildung 22 – Drehung des Impaktors um 5° um die z-Achse (blau) im Vergleich zur Ausgangslage (weiß).....	29
Abbildung 23 – Schräge Positionierung der Auflager (vgl. blaue Darstellung)	29
Abbildung 24 – Auflagerversatz in y-Richtung um 1 cm (blauer Auflager).....	30
Abbildung 25 – Links: GHBMC-Femur-Modell [34]/ Rechts: THUMS Version 4 Femur-Modell [97]	30
Abbildung 26 – Modellvergleiche	31
Abbildung 27 – Beispiel einer Clusteranalyse einer Punktwolke [21].....	45
Abbildung 28 – Beispiel einer Clusteranalyse von einer Kurvenschar [80].....	45
Abbildung 29 – Klassifizierung der Clusteranalyseverfahren auf Basis der Zuordnung der Objekte [5]	47
Abbildung 30 – Klassifizierung von Clusteranalyseverfahren anhand Fusionierungsprozess [6]	47

Abbildung 31 – Links: Aufteilung einer Kurvenschar in 3 Cluster / rechts: Dendrogramm- Visualisierung [80]	49
Abbildung 32 – Darstellung des Mappings bei der euklidischen Distanzmessung (links) und der DTW Distanzmessung (rechts) [48].....	49
Abbildung 33 – Visualisierung der Distanzmessung zwischen den Clustern bei den Fusionierungsalgorithmen Single-Linkage, Complete-Linkage und Average- Linkage [47].....	50
Abbildung 34 – Methodik zur Ermittlung geeigneter Clusterverfahren und Clusteranzahl.....	52
Abbildung 35 – Vorgehen Bewertung der Clusteranalyse.....	53
Abbildung 36 – Visualisierung der Testdatensätze I–V	55
Abbildung 37 – Testdatensatz VI	56
Abbildung 38 – Testdatensatz VII	56
Abbildung 39 – Testdatensatz VIII	57
Abbildung 40 – Datenvorbereitung für die Clusteranalyse	58
Abbildung 41 – Vergleich der Dendrogramme zwischen den verwendeten Distanzmetriken	61
Abbildung 42 – Ergebnisse Testdatensatz I mit euklidischer Distanzmetrik.....	62
Abbildung 43 – Visualisierung der Ergebnisse der Clusteranalyse für Testdatensatz II.....	63
Abbildung 44 – Visualisierung der Ergebnisse der Clusteranalyse für Testdatensatz III bei Verwendung des euklidischen Distanzmaßes	64
Abbildung 45 – Visualisierung der Ergebnisse der Clusteranalyse für Testdatensatz III bei Verwendung der dynamischen Zeitnormierung als Distanzmaß.....	65
Abbildung 46 – Visualisierung der Ergebnisse der Clusteranalyse für Testdatensatz IV bei Verwendung des euklidischen Distanzmaßes	67
Abbildung 47 – Visualisierung der Ergebnisse der Clusteranalyse für Testdatensatz IV bei Verwendung der dynamischen Zeitnormierung als Distanzmaß.....	68
Abbildung 48 – Visualisierung der favorisierten und maximalen Clusteranzahl.....	69
Abbildung 49 – Visualisierung der Ergebnisse der Clusteranalyse für Testdatensatz V bei Verwendung des euklidischen Distanzmaßes	70
Abbildung 50 – Visualisierung der Ergebnisse der Clusteranalyse für Testdatensatz V bei Verwendung der dynamischen Zeitnormierung als Distanzmaß.....	71
Abbildung 51 – Vergleich der Dendrogramm-Struktur bei Verwendung beider Distanzmetriken für Testdatensatz VI	75
Abbildung 52 – Vergleich der Dendrogramm Struktur bei Verwendung beider Distanzmetriken für Testdatensatz VII.....	76
Abbildung 53 – Vergleich der Dendrogramme für Testdatensatz VIII bei Verwendung beider Distanzmetriken.....	77
Abbildung 54 – (a) Visualisierung der Datensätze, (b) form-basierte Clusteranalyse, (c) struktur- basierte Clusteranalyse [80]	81
Abbildung 55 – Testdatensatz VI mit euklidischer Distanzmetrik und Median Fusionierungsalgorithmus	84
Abbildung 56 – Vorgehensweise Kapitel 5 und Zusammenhang mit Kapitel 3.....	90
Abbildung 57 – Erkenntnisgewinn und Plausibilitätsbewertung nach Clusteranalyse	92
Abbildung 58 – Güte der Clusteranalyse in Relation zum Erkenntnisgewinn	93
Abbildung 59 – Links: Gleichmäßige Metadaten-Verteilung, Rechts: Deutliche Gruppierung der Metadaten in drei Cluster.....	93
Abbildung 60 – Clusteraufteilung des Dendrogramms	94
Abbildung 61 – Einschätzung der Plausibilität der Wertebereiche der Metadaten.....	94

Abbildung 62 – Veranschaulichung des Kompensationseffektes bezüglich Biegesteifigkeit innerhalb eines Clusters	95
Abbildung 63 – Bewertungsschema für die Clusteranalyse.....	98
Abbildung 64 – Links: Visualisierung der entnommenen Stichproben / Rechts: Zufallszahlen der entnommenen Stichproben	99
Abbildung 65 – Linearisiertes elastisch-plastisches Materialgesetz [101].....	100
Abbildung 66 – Verwendete Alterungsfunktionen von Dokko et al. [29].....	100
Abbildung 67 – Darstellung der Lage der untersuchten Querschnitte des Femurs und vereinfachte Berechnung [10].....	102
Abbildung 68 – Eigene Aufbereitung der NHANES-Datenbank [10] bez. Abhängigkeiten der Geometrie-Parameter	102
Abbildung 69 – Messung der Femurlänge in Kerrigan et al. [65].....	103
Abbildung 70 – Diskrete Wahrscheinlichkeitsverteilungen als Basis der Femur-Geometriedefinition	103
Abbildung 71 – Vergleich der Soll- und Ist-Population gegen die Referenzdatenbank NHANES [10] bez. kortikale Dicke und Innendurchmesser des Femurschafts und der Referenzdatenbank Park [85] bez. Femurlänge	105
Abbildung 72 – Beispiel der Femurmodellierung für Femur M22	106
Abbildung 73 – Modellierung Übergangsbereiche zwischen Femurschaft und den Endbereichen	106
Abbildung 74 – Modellierung Femur M22	107
Abbildung 75 – Schaubild der erzeugten numerischen Population an Femora	108
Abbildung 76 – Konzeptdatensatz I	109
Abbildung 77 – Konzeptdatensatz II	110
Abbildung 78 – Aufteilung der Kurvenschar in 5 Cluster nach Ward-Algorithmus.....	112
Abbildung 79 – Ergebnis der Clusteranalyse mit Aufteilung in 5 Cluster	112
Abbildung 80 – Vergleich der Wertebereiche der Cluster bezüglich Femurlänge mit der Referenzpopulation der Femurlängen aus Park [85]	113
Abbildung 81 – Vergleich der Wertebereiche der Cluster bezüglich Flächenträgheitsmoment mit einer Referenzpopulation bestehend aus den Flächenträgheitsmomenten von Ivarsson et al. [58], Funk et al. [42], Kennedy et al. [62] und Kerrigan et al. [65].....	113
Abbildung 82 – Ergebnis der Clusteranalyse mit Aufteilung in 4 Cluster	116
Abbildung 83 – Vergleich der Wertebereiche der Cluster bezüglich Femurlänge mit der Referenzpopulation der Femurlängen aus Park [85]	116
Abbildung 84 – Vergleich der Wertebereiche der Cluster bezüglich Flächenträgheitsmoment mit einer Referenzpopulation bestehend aus den Flächenträgheitsmomenten von Ivarsson et al. [58], Funk et al. [42], Kennedy et al. [62] und Kerrigan et al. [65].....	117
Abbildung 85 – Abgedeckter Wertebereich des E-Moduls bei 35 Jahren [29]	124
Abbildung 86 – Vergleich der Flächenträgheitsmomente zwischen HSA-Programm und anderen Quellen	128

Tabellenverzeichnis

Tabelle 1 – Kriterien für die Auswahl des dynamischen Biegeversuchs nach Forman et al. [38].....	20
Tabelle 2 – Vergleich der Experimente von Forman et al. und Funk et al.	21
Tabelle 3 – Variantendefinition und Gewicht der Auflager	26
Tabelle 4 – Materialeigenschaften Femur THUMS Version 4 [97].....	30
Tabelle 5 – Materialeigenschaften Femur GHBM [34]	30
Tabelle 6 – Ergebnisse der Sensitivitätsanalyse bezüglich der Modellierungsannahmen.....	32
Tabelle 7 – Ergebnisse der Sensitivitätsanalysen der geometrischen Parameter der Auflager	32
Tabelle 8 – Ergebnis Sensitivitätsanalysen bezüglich Schaumdefinition	33
Tabelle 9 – Ergebnis Sensitivitätsanalysen hinsichtlich Knochenpositionierung	33
Tabelle 10 – Ergebnisse der Sensitivitätsanalysen bezüglich der Positionierung des Versuchsaufbaus.....	34
Tabelle 11 – Reibwertkoeffizienten entnommen aus [39]	37
Tabelle 12 – Auflistung der verwendeten internen Validierungsindizes [20].....	51
Tabelle 13 – Eigenschaften der Testdatensätze.....	54
Tabelle 14 – Aufteilung des Testdatensatz III in vier Cluster	58
Tabelle 15 – Aufteilung des Testdatensatz IV in 3 Cluster	59
Tabelle 16 – Aufteilung des Testdatensatz IV in 4 Cluster	59
Tabelle 17 – Aufteilung des Testdatensatz IV in 7 Cluster	59
Tabelle 18 – Aufteilung von Testdatensatz V in 3 Cluster.....	59
Tabelle 19 – Aufteilung von Testdatensatz V in 4 Cluster.....	59
Tabelle 20 – Aufteilung von Testdatensatz V in 6 Cluster.....	59
Tabelle 21 – Übersicht der Ergebnisse der Testdatensätze I–V.....	61
Tabelle 22 – Zusammenfassung der Eignung der int. Clustervalidierungsindizes für die Testdatensätze I–V bei Verwendung der euklidischen Distanzmetrik	72
Tabelle 23 – Zusammenfassung der Eignung der int. Clustervalidierungsindizes für die Testdatensätze I–V bei Verwendung der dynamischen Zeitnormierung als Distanzmetrik	73
Tabelle 24 – Übersicht der Ergebnisse der Testdatensätze VI–VIII.....	74
Tabelle 25 – Zusammenfassung der geeigneten int. VIs für das euklidische Distanzmaß.....	84
Tabelle 26 – Fragebogen für die Datensatzanalyse	97
Tabelle 27 – Absolutwerte der Stichproben zur Materialdefinition	101
Tabelle 28 – Ziel-Geometrien der Femur-Population	104
Tabelle 29 – Ist-Werte der Femur Geometriedefinition der ausgewählten Femora des CT- Kollektivs.....	104
Tabelle 30 – Metadaten der numerischen Modelle	109
Tabelle 31 – Bewertungsfragebogen für Konzeptdatensatz I.....	111
Tabelle 32 – Bewertungsfragebogen für Konzeptdatensatz II.....	115

Abkürzungsverzeichnis

Abkürzung	Begriff
AM 50	50 th percentile american male: Englischer Begriff für das 50. Perzentil der maskulinen amerikanischen Bevölkerung
ASME	American Society of Mechanical Engineers: Vereinigung von Maschinenbauingenieuren in den Vereinigten Staaten
ATD	Anthropomorphic Test Device: englischer Begriff für Crashtest-Dummy
BMI	Body Mass Index: Kennzahl, die aus Körpergröße und -gewicht bestimmt wird
CT	Computertomographie: bildgebendes Verfahren in der Medizintechnik
DTW/DTWARP	Distanzmetrik der dynamischen Zeitnormierung
EUCL	Euklidische Distanzmetrik
FE	Finite-Elemente-Methode: Berechnungsmethode, in der das Gesamtsystem in eine hohe Zahl diskreter Elemente aufgeteilt wird.
GHBMC	Ein kommerzielles virtuelles FE-Ganzkörpermodell des Global Human Body Model Consortium
HSP	Hip Structural Program: Öffentlich verfügbare Datenbank als Teil der NHANES-Studie
int. VI	Interner Validierungsindex zur Bestimmung einer Clusteranzahl
MRT	Magnetresonanztomographie: bildgebendes Verfahren in der Medizintechnik
NASS-CDS	Englische Abkürzung für die Datenbank <i>Crashworthiness Data System of the National Automotive Sampling System</i>
NHANES	National Health and Nutrition Examination Survey: öffentlich verfügbare Studie der Vereinigten Staaten
NTHSA	National Highway Traffic Safety Administration: amerikanische Verkehrsbehörde
PMMA	Polymethylmethacrylat-Material wird für die Knocheneinbettung bei experimentellen Studien verwendet
RBB	Rollerblockbreite: die Breite der Auflager aus dem Referenzlastfall
THUMS™	Total Human Model for Safety: ein kommerzielles virtuelles FE-Ganzkörpermodell
VPS	Virtual Performance Solution: kommerzielle Finite-Elemente-Software

1 Einleitung

1.1 Einsatz von Menschmodellen in der Automobilindustrie

1.1.1 Vorteile der Menschmodelle gegenüber anthropomorphen Messsystemen

Bis heute werden in der Fahrzeugsicherheit so genannte Crashtest-Dummys verwendet, um die Auswirkungen von Sicherheitstechnologien auf Fahrzeuginsassen zu testen (siehe Abbildung 1). Der englische Begriff für Crashtest-Dummy lautet Anthropomorphic Test Device (Abkürzung: ATD). Im weiteren Verlauf wird die englische Abkürzung ATD verwendet, um eine bessere Lesbarkeit des Textes zu erreichen.



Abbildung 1 – Von links nach rechts: Finite-Elemente (Abkürzung: FE) Modelle des ATD HYBRID III für frontale Belastung, des ATD SID II für seitliche Belastung und des THUMS™-Menschmodells [69]

ATDs sind menschenähnliche Puppen, welche aus industriellen Werkstoffen, wie Stahl und Gummi, hergestellt werden. Im Inneren eines ATDs sind Sensoren für Beschleunigungs-, Kraft- und Wegmessungen verbaut, um die auf den Insassen einwirkenden Belastungen zu untersuchen. Diese Messgrößen werden während eines Versuches aufgenommen und können im Nachgang verwendet werden, um das Verletzungsrisiko auf den Insassen zu bestimmen.

ATDs werden nach Geschlecht, Alter, Körpergröße und Belastungsrichtung kategorisiert [115]. Die Bevölkerungsvariabilität der erwachsenen Bevölkerung über 18 Jahre wird mit ATDs meist durch die folgenden Perzentile bezüglich Körpergröße und Körpergewicht abgebildet: das 5. Perzentil der weiblichen Bevölkerung, das 50. Perzentil der männlichen Bevölkerung und das 95. Perzentil der männlichen Bevölkerung. Die Referenzbevölkerung für den ATD HYBRID III ist beispielsweise die amerikanische Bevölkerung [92].

Um ein annähernd menschenähnliches mechanisches Verhalten zu erlangen, genannt Biofidelität, werden die ATDs gegen Leichenversuche und Freiwilligentests validiert. Der technische Bericht ISO/TR 9790 spezifiziert beispielsweise die notwendigen Biofidelitätsbewertungen für einen ATD klassifiziert für laterale Lasteinwirkung [56]. ATDs sind heutzutage nicht mehr nur physisch, sondern auch als numerische Modelle verfügbar.

In der Anwendung weisen ATDs jedoch einige Limitationen auf, welche sich auf die Prognosefähigkeit und Prognosegüte der Insassenschutz-Bewertung auswirken. ATDs sind auf eine Belastungsrichtung validiert. Historisch wurden seit Ende der 40-er Jahre ATDs zuerst für frontale Belastungen (z.B. HYBRID III, THOR) und seit Ende der 70-er Jahre für seitliche Belastungen (z.B. Bio-SID, EuroSID) entwickelt [69].

Zudem stellen ATDs eine mechanische Repräsentation des Menschen dar. Mit Werkstoffen wie Stahl und Gummi kann nur eine begrenzte Biofidelität erreicht werden. Des Weiteren sind Sensoren für Belastungsmessungen nur an bestimmten Stellen verbaut. Yoganandan et al. [115] bieten beispielsweise einen guten Überblick über die Instrumentalisierung des ATDs Hybrid III für frontale Lasteinwirkung. Zuletzt muss noch berücksichtigt werden, dass aufgrund der überschaubaren Anzahl an existierenden Perzentilen die reale Bevölkerungsvariabilität nur sehr eingeschränkt mit ATDs untersucht werden kann.

Bei der Entwicklung von Menschmodellen wird das Ziel verfolgt, diese genannten Limitationen zu beseitigen. Menschmodelle verfügen über eine detaillierte Geometrie des menschlichen Körpers (siehe Abbildung 1). Neuere Modelle, wie z.B. die vierte Version der THUMS-Menschmodelle, besitzen neben dem Skelett auch Organe, sowie die Abbildung von Muskeln und Bändern [96]. Zudem besitzen Menschmodelle biologische Materialeigenschaften [113, 96]. Die Biofidelität von Menschmodellen wird analog zu ATDs ebenfalls mit Vergleichen zu biomechanischen Experimenten erreicht. Kitagawa et al. veröffentlichten z.B. eine Übersicht der verwendeten Experimente für die numerische Validierung der vierten Version des THUMS-Menschmodells [66].

Im Gegensatz zu ATDs sind Menschmodelle vom Modellierungsansatz her unabhängig hinsichtlich der Belastungsrichtung. Zudem können Verletzungen potenziell genauer prognostiziert werden. Zum Beispiel wird bei ATDs nur die axiale Kraft im Femur gemessen und in ein Verletzungsrisiko umgewandelt. Mit dem Menschmodell kann eine auf das Femur wirkende Belastung genauer untersucht werden. Es kann beispielsweise die Belastungsart untersucht werden und es kann grundsätzlich auch eine Vorhersage getroffen werden, an welcher Stelle des Knochens eine Fraktur auftritt.

Vor allem im Hinblick auf Zukunftstrends in der Automobilbranche, wie z.B. der Entwicklung neuer Sitzpositionen im Bereich des hochautomatisierten Fahrens (siehe Abbildung 2), ist der Einsatz von Menschmodellen und ihre omni-direktionale Abbildung der mechanischen Antwort unabdingbar. Die amerikanische Verkehrsbehörde National Highway Traffic Safety Administration (Abkürzung: NHTSA) fordert zudem explizit bei der Entwicklung neuer Sicherheitstechnologien einen Nachweis des Schutzpotenzials für Insassen jeder Größe und Altersgruppe [105]. Diese Forderung ist mit der eingeschränkten Bevölkerungsrepräsentation der ATDs anhand von drei Bevölkerungsperzentilen nicht mehr möglich.



Abbildung 2 – Neuartige Sitzpositionen in einem Konzeptfahrzeug [105]

Ein weiterer Vorteil von Menschmodellen gegenüber ATDs ist ihre erhöhte Adaptivität, da numerische Modelle leichter verändert werden können als physische Modelle. Beispielsweise werden die Perzentile für ATDs einmal zu Beginn der ATD Entwicklung festgelegt und bleiben über dessen Entwicklungs- und Lebenszyklus unverändert. Die Perzentile für den ATD Hybrid III wurden in den 70er Jahren definiert [87]. Reed et al. [87] überprüften beispielsweise die damals definierten Perzentile anhand der anthropometrischen Datenbank der National Health and Nutrition Examination Survey

(Abkürzung: NHANES) bezüglich der Jahre 2005 - 2008. Das Ergebnis ist, dass der ATD Hybrid-III des 50. Perzentils aktuell das 33. Perzentil bezüglich des Körpergewichts repräsentiert. Der ATD für das 95. Perzentil der Bevölkerung repräsentiert aktuell das 81. Perzentil bezüglich des Körpergewichts. Hinsichtlich der Körpergröße unterscheiden sich die aktuellen Perzentile um weniger als 5 Prozent. Im Gegensatz zu physischen ATDs ist eine derartige Geometrie- oder Massenaktualisierung für Menschmodelle möglich.

Aktuelle Entwicklungen im Bereich der Menschmodelle sind die Entwicklung von Algorithmen zur Veränderung der Geometrie eines FE-Netzes, sogenannte mesh-morphing Algorithmen [113]. Diese Algorithmen ermöglichen eine Veränderung der Geometrie der numerischen Modelle auf gewünschte Zielgrößen und die Entwicklung von parametrisierten Modellen (siehe Abschnitt 1.1.3). Diese parametrisierten Modelle bilden die Basis für stochastische Simulationen, d.h. die Simulation einer hohen Anzahl von Individuen unterschiedlicher Charakteristika. Diese stochastischen Simulationen ermöglichen beispielsweise die Untersuchung der Bevölkerungsvariabilität oder die Untersuchung gefährdeterer Bevölkerungsgruppen auf bestimmte Belastungen, wie beispielsweise ältere oder übergewichtige Personen.

1.1.2 Existierende Menschmodelle und Teilkörpermodelle

Yang et al. [112] veröffentlichten bereits 2006 eine systematische Übersicht über entwickelte numerische Modelle für die Anwendung in der biomechanischen Forschung. Diese systematische Übersichtsarbeit beschränkte sich auf numerische Teilkörpermodelle und Ganzkörpermodelle, die während der letzten 50 Jahre auf der Stapp Car Crash Konferenz veröffentlicht wurden. Es wird hierbei explizit erwähnt, dass diese Übersichtsarbeit keinen vollumfänglichen Anspruch besitzt, da sie lediglich Veröffentlichungen eines spezifischen Journals betrachtet. Für die einzelnen Modelle wurden jeweils Details zur Modellierung, den verwendeten Materialeigenschaften, der Validierungsgrundlage, sowie hinsichtlich des Anwendungsbereiches veröffentlicht. Die Veröffentlichung zeigt zudem die historische Entwicklung immer detaillierterer numerischer Modelle aufgrund der rasanten Entwicklung der Rechenleistung von Computern.

Heutzutage existieren zahlreiche numerische FE-Ganzkörpermodelle für die Untersuchung von Verletzungsmechanismen (siehe Abbildung 3). Die beiden kommerziell lizenzierten Modelle THUMS und GHBMC werden häufig in der Automobilbranche verwendet [52, 98].

Das Total Human Model for Safety (Abkürzung: THUMS) Modell wurde von Toyota Motor Corporation und von Toyota Central R&D Labs gemeinsam entwickelt [66]. Die erste Version des THUMS-Modells wurde im Jahr 2000 und die fünfte Version im Jahr 2015 veröffentlicht (siehe Abbildung 4). Mit jeder veröffentlichten Version stieg der Detaillierungsgrad und die Güte zur Verletzungsprädiktion der Modelle.

Das Global Human Body Model Consortium (Abkürzung: GHBMC) Modell wird von einem Konsortium bestehend aus mehreren Automobilherstellern, Zulieferern, Forschungseinrichtungen und der amerikanischen Verkehrsbehörde NHTSA entwickelt [22]. GHBMC-Modelle existieren als detaillierte Modelle und als vereinfachte Modelle (siehe Abbildung 5).

Beide Modelle, die THUMS-Modelle und die GHBMC-Modelle, wurden jeweils als Fußgänger-Modelle und als Insassen-Modelle entwickelt und existieren für dieselben Perzentile analog zu den ATDs. Aktuelle Entwicklungstendenzen im Bereich der Menschmodell-Entwicklung beinhalten u.a. die Entwicklung von Modellen höheren Alters (z.B. CHARM-70F in Abbildung 3) und von Modellen, welche übergewichtige Personen repräsentieren [23]. Zudem liegt vermehrt der Fokus auf der Entwicklung von Modellen mit der Möglichkeit der Muskelaktivierung für die Untersuchung von Szenarien kurz vor einem Unfallgeschehen (z.B. fünfte Version des THUMS-Modell in Abbildung 4) [23, 61].

In dieser Arbeit werden beide Modelle, das GHBMC-Modell, in der Version 4.3, und das THUMS-Modell, in der Version 4, verwendet.








Model	HUMOS	H-Model	Ford Model	WSU Model	THUMS	GHBMC	CHARM-70F
Reference	Robin (2001) Vezin and Verriest (2005)	Haug et al. (2004) Ito et al. (2012)	Ruan et al. (2003) El-Jawahri et al. (2010)	Shah et al. (2001) Kim et al. (2005)	Iwamoto et al. (2002) Hayashi et al. (2008)	Shin et al. (2012) Vavalle et al. (2013)	To be published
Software	Radioss PAM-CRASH	PAM-CRASH	LS-DYNA	LS-DYNA	LS-DYNA	LS-DYNA	LS-DYNA
Size	Midsize male Small female Large male Scale to others	Midsize male Small female	Midsize male	Midsize male	Midsize male Small female Large male	Midsize male Small female Large male	Midsize female
Age	No	35, 75	35, 55, 75	No	No	No (65)	70
Obesity	No	No	No	No	No	No	No
Posture	Seated	Seated	Seated	Seated	Seated and standing	Seated and standing	Seated and standing
Figure (most recent version)							
Geometry	External anthropometry, X-ray	Visible Human Project	Visible Human Project	Visible Human Project	CT scans	MRI, UMRI, CT, external anthropometry	CT scans
Elements	~70 k	Not Stated	~120 k	~120 k	~1.8 million	~2.2 million	~1.5 million

Abbildung 3 – Überblick über aktuelle FE-Ganzkörper-Menschmodelle [113]

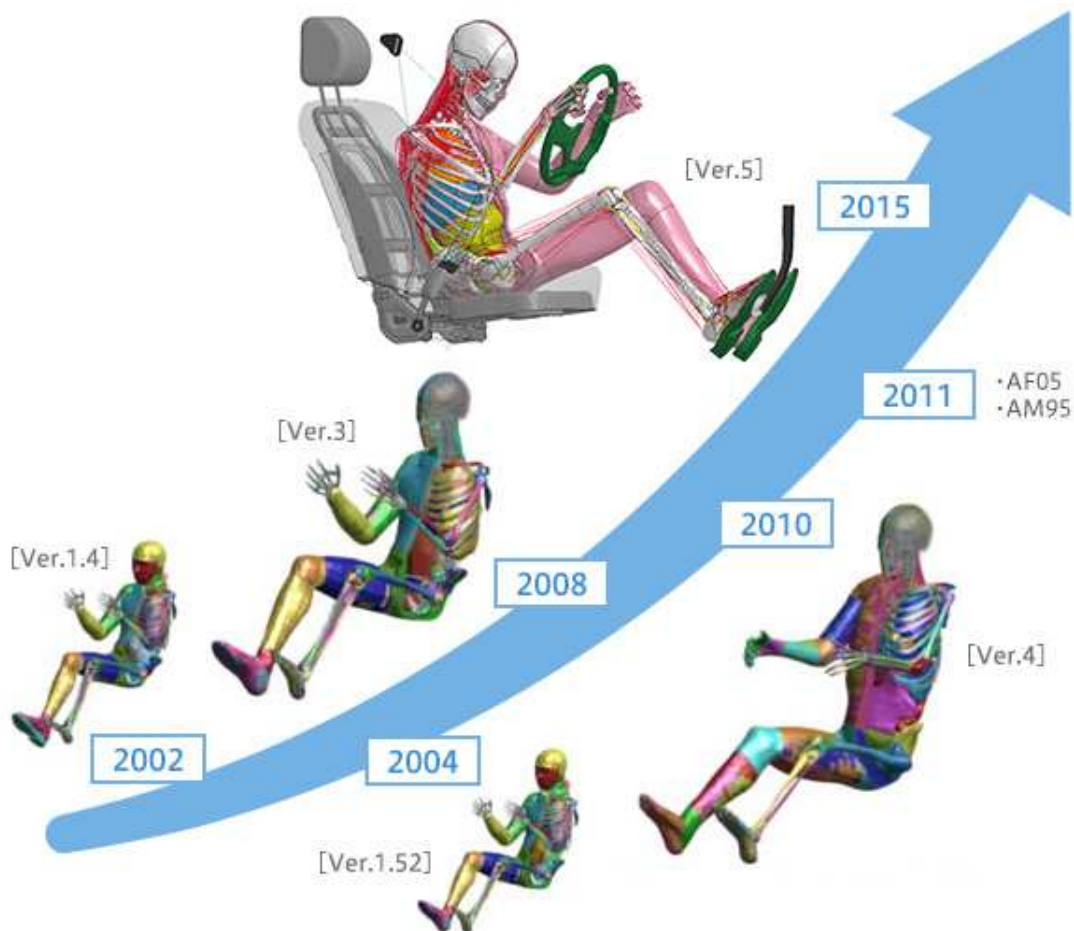


Abbildung 4 – Zunahme des Detaillierungsgrades der THUMS™-Modelle von Version 1 bis Version 5 [60]

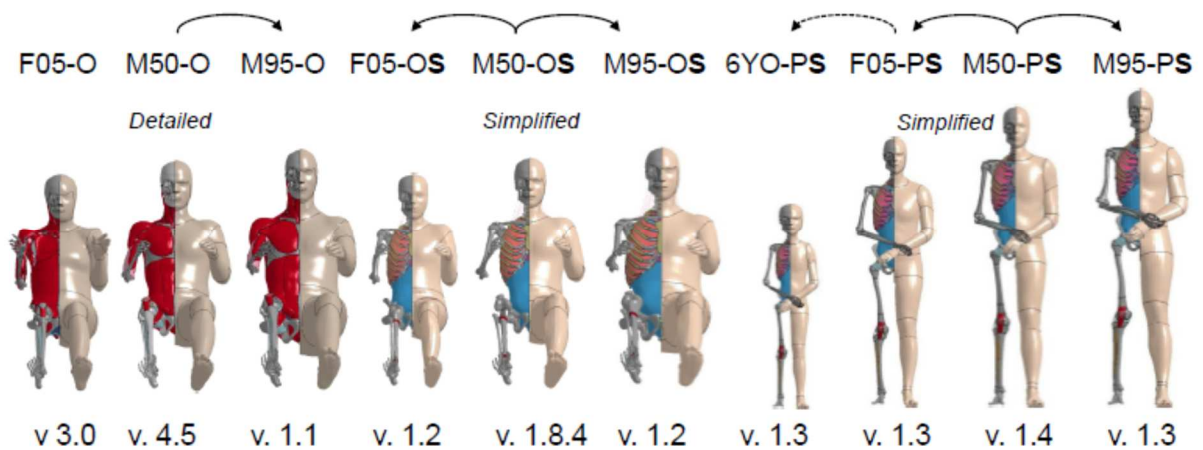


Abbildung 5 – Entwickelte detaillierte und vereinfachte Modelle des GHBMC-Modells [22]

1.1.3 Neue Entwicklungen – Parametrisierte Menschmodelle

Historisch wurden Menschmodelle analog zu den ATDs zuerst für die drei Perzentile entwickelt. Die Notwendigkeit von Simulationen mit Menschmodellen abweichend von den drei existierenden Perzentilen wird offensichtlich, wenn man die Faktoren betrachtet, welche das Verletzungsrisiko maßgeblich beeinflussen. Carter et al. verglichen die Faktoren Alter, Body Mass Index (Abkürzung: BMI) und Geschlecht hinsichtlich der Verletzungsschwere bei Unfällen [18]. Die Studie kam zu dem Ergebnis, dass alle drei Faktoren einen maßgeblichen Einfluss auf die Verletzungsschwere haben. Diese drei Faktoren wirken sich auf die Geometrie, die Materialeigenschaften und die konstitutionellen Eigenschaften von Knochen und Gewebe aus [113]. Unter dem Begriff Geometrie wird hierbei die Ganzkörper Geometrie, sowie die geometrischen Eigenschaften einzelner Knochen zusammengefasst [113]. Unter dem Begriff konstitutionelle Eigenschaften wird die Querschnittsgeometrie von Knochen und Gewebe zusammengefasst [113].

Dieser beschränkten Repräsentation der Bevölkerungsvariabilität anhand von Perzentilen kann mit parametrisierten Modellen entgegengewirkt werden. Ein parametrisiertes Modell kann als Basismodell angesehen werden, das durch die Veränderung bestimmter Parameter ein neues Modell mit veränderten Eigenschaften anhand eines automatisierten Prozesses generiert. Diese veränderbaren Parameter beziehen sich meist auf die Definition der Geometrie und Materialeigenschaften. In Abbildung 6 sind die notwendigen Arbeitsschritte für die Entwicklung parametrisierter Modelle beispielhaft anhand eines Brustkorbs und des Femurs dargestellt. Diese Arbeitsschritte werden im Folgenden kurz beschrieben. In Yang et al. [113] sind die Arbeitsschritte ausführlich dokumentiert.

Für die Entwicklung parametrisierter Modelle wird eine große Datenbasis an Computertomographie (Abkürzung: CT)-Aufnahmen oder Magnetresonanztomographie (Abkürzung: MRT)-Aufnahmen benötigt. Für jede Aufnahme wird eine bestimmte Anzahl an geometrischen Referenzpunkten extrahiert, welche die Geometrie des betrachteten Objektes beschreiben. Ziel dieser Daten-Extraktion ist es, für jedes Individuum der Datenbasis dieselben Referenzpunkte auszulesen, damit diese zwischen den einzelnen Individuen verglichen werden können. Diese Extraktion der Referenzpunkte kann auf unterschiedliche Weise geschehen. Die Referenzpunkte können für jedes Individuum individuell bestimmt werden. Eine andere Möglichkeit ist, ein Basismodell des betrachteten Objektes mit Referenzpunkten zu generieren und dieses Basismodell auf die jeweilige Geometrie jedes Individuums anzugleichen.

Nach der Datenextraktion ist das Ziel, ein statistisches Modell basierend auf einer geringen Anzahl charakteristischer Parameter zu generieren, welches die Geometrie eines Basismodells anhand definierter Referenzpunkte verändert. Parametrisierte Ganzkörpermodelle können z.B. anhand von globalen Parametern, wie Alter, BMI, Geschlecht oder Körpergröße, die Geometrie auf Basis der Referenzpunkte verändern. Als erster Schritt wird die Procrustes-Analyse durchgeführt, um die Referenzpunkte zwischen den einzelnen Individuen der Datenbasis anzugleichen, um die Vergleichbarkeit zwischen Objekten sicherzustellen und die weiteren statistischen Analysen zu ermöglichen. Danach wird die Hauptkomponentenanalyse durchgeführt, um die Anzahl der Referenzpunkte für die Geometriedefinition zu verringern. Bei der Hauptkomponentenanalyse wird die gesamte Anzahl an Geometrie-definierenden Referenzpunkten durch möglichst wenige Linearkombinationen, die Hauptkomponenten, angenähert. Ziel dabei ist, diejenigen Referenzpunkte auszusortieren, welche nicht maßgeblich für die geometrische Varianz verantwortlich sind. Danach wird eine multivariate Regressionsanalyse verwendet, um die Ergebnisse der Hauptkomponentenanalyse, der reduzierten Anzahl geometrischer Referenzpunkte, mit charakteristischen Parametern zu verknüpfen, um das statistische Modell für die Geometrieprädiktion zu erhalten. Auf Basis dieses statistischen Modells und eines Basismodells, kann anhand weniger charakteristischer Parameter ein neues Modell mit veränderten Eigenschaften generiert werden. Es existieren mehrere Möglichkeiten zur Veränderung des FE-Netzes des Basismodells. Dieser automatisierte Prozess der Veränderung des FE-Netzes kann auf verschiedenen Methoden, wie z.B. der Kriging-Methode oder radialer Basisfunktionen, beruhen.

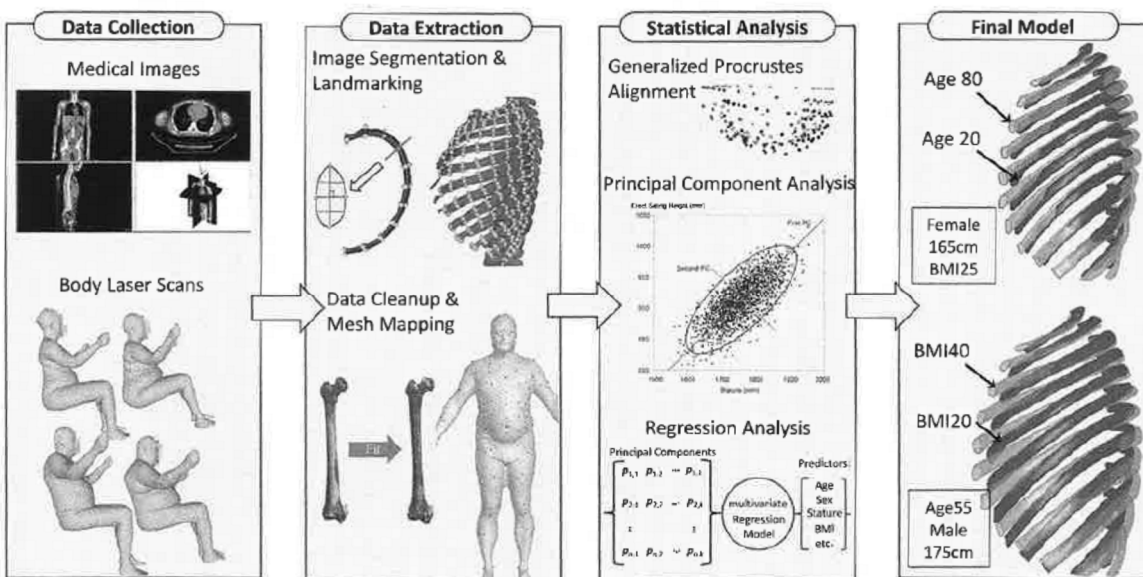


Abbildung 6 – Notwendige Arbeitsschritte für die Entwicklung eines parametrisierten Modells [113]

In der Literatur wurden bereits zahlreiche parametrisierte Modelle bezüglich einzelner Knochen oder Organe, bis hin zu parametrisierten Teil- und Ganzkörper Menschmodellen veröffentlicht [13, 15, 63, 67, 71, 84, 108, 114].

1.2 Validierung von Menschmodellen

1.2.1 Begriffsdefinitionen

Numerischen Modellen fehlt es oft an Glaubwürdigkeit aufgrund unzureichend durchgeführter Verifikations- und Validierungstätigkeiten [3]. In dem Leitfaden von Anderson et al. [3] werden Validierungs- und Verifikationskonzepte, sowie Konzepte für Sensitivitätsanalysen, speziell für die Anwendung in der numerischen Biomechanik, vorgestellt.

Die Basis für die Entwicklung der in Anderson et al. vorgestellten Konzepte waren Anstrengungen, welche in anderen Fachbereichen bezüglich Verifizierungs- und Validierungsstrategien bereits unternommen wurden, wie z.B. im Bereich der numerischen Strömungsmechanik oder im Bereich der numerischen Festkörpermechanik [94]. Bevor auf die Validierung von Menschmodellen und die dafür verwendeten Validierungsansätze näher eingegangen wird, werden in diesem Abschnitt folgende Fachbegriffe näher definiert: Verifizierung, Validierung, Kalibrierung, Sensitivitätsanalyse, Unsicherheit und Fehler.

Verifizierung und Validierung

Laut Anderson et al. [3] werden für ein numerisches Modell folgende Ziele mit einer Verifizierungs- und Validierungsstrategie angestrebt:

- Überprüfung ob die mathematischen Gleichungen, welche das System beschreiben, richtig implementiert wurden,
- Überprüfung ob das numerische Modell die Physik des realen experimentellen Versuchsaufbaus ausreichend genau widerspiegelt,
- Einschätzung von Fehler und Unsicherheiten ausgehend von dem numerischen Modell.

Laut dem Verifizierungs- und Validierungsleitfaden für die numerische Festkörpermechanik der American Society of Mechanical Engineers (Abkürzung: ASME) [94] umfasst der Begriff Verifizierung zwei Bereiche: die Verifizierung des Algorithmus und die Verifizierung der Simulation. Die Verifizierung des Algorithmus beinhaltet die Überprüfung, ob der Algorithmus richtig implementiert wurde und die mathematischen Gleichungen richtig gelöst werden [94]. Bei der Verifizierung der Simulation wird überprüft, ob die diskrete Lösung des Finite-Elemente-Problems ausreichend genau ist [94]. Verifizierung im Bereich der Biomechanik bezieht sich meist auf die Implementierung der konstitutiven Gleichungen und der Einschätzung von Diskretisierungsfehlern [3].

Bei der Validierung geht es um die quantitative Bewertung, mit welcher Güte das numerische Ergebnis mit dem experimentellen Ergebnis übereinstimmt. Es geht aus der Perspektive des Anwendungszwecks um die Bewertung, wie zufriedenstellend ein numerisches Modell die Realität abbildet [94].

Kalibrierung

Die Kalibrierung eines numerischen Modells beinhaltet die Veränderung der numerischen Parameter, um eine bessere Übereinstimmung zwischen der numerischen Modellvorhersage und den experimentellen Ergebnissen zu erlangen [25]. Die Tätigkeit verfolgt somit das Ziel, die Güte der Validierung zu erhöhen.

Sensitivitätsanalysen

Bei der Durchführung von Sensitivitätsanalysen werden Variablen des Modells gezielt verändert, um den Einfluss dieser Variablen auf die Modellvorhersage zu untersuchen. In der Literatur gibt es widersprüchliche Aussagen, wann Sensitivitätsanalysen durchgeführt werden sollen.

Laut Anderson et al. [3] sollten Sensitivitätsanalysen vor der Durchführung von Experimenten erfolgen, um wichtige Parameter zu identifizieren, aber auch während des gesamten Validierungsprozesses, nicht jedoch als separater Bestandteil nach dem Validierungsprozess.

Unsicherheit und Fehler

Laut dem Leitfaden nach Anderson et al. können Fehler in zwei Arten kategorisiert werden. Es gibt numerische Fehler und Modellierungsfehler. Zu den numerischen Fehlern gehören z.B. Diskretisierungsfehler, Fehler aus mangelnder Netzkonvergenz, sowie numerische Rundungsfehler. Modellierungsfehler entstehen aufgrund von Annahmen in der mathematischen Repräsentation des physikalischen Problems. Solche Fehler entstehen beispielsweise durch Inkonsistenzen zwischen Modell und Experiment bezüglich Geometrie, Randbedingungen, Materialeigenschaften und verwendeter konstitutiver Gleichungen [3].

Unsicherheiten können durch den Modellierungsprozess aufgrund von Mangel an Informationen, z.B. durch nicht definierte Materialeigenschaften, entstehen. Unsicherheiten und Fehler sind wie folgt zu unterscheiden: Fehler existieren permanent im System, während Unsicherheiten nur zeitweise auftreten [3].

1.2.2 Validierung von Menschmodellen

Verifizierungs- und Validierungstätigkeiten sind essentielle Schritte in der Entwicklung von Menschmodellen. Die Modelle werden während des Entwicklungsprozesses mit einer möglichst großen Anzahl an Validierungslastfällen unterschiedlicher Belastungsarten und Komplexität verglichen, um ihre Biofidelität zu bewerten. Mit der Biofidelitätsbewertung wird bestimmt, bis zu welchem Grad das numerische Menschmodell das physikalische Verhalten eines realen Menschen repräsentiert. Die Biofidelitätsbewertung verfolgt somit das Ziel, Vertrauen in das numerische Modell zu generieren. Die Validierung von Menschmodellen sollte schrittweise, mit zunehmender Komplexität der Validierungslastfälle, erfolgen [25].

Es existieren mehrere Möglichkeiten, Validierungslastfälle zu kategorisieren. In Abbildung 7 ist die Kategorisierung der Validierungslastfälle nach Fuchs [40] dargestellt. Sie verfeinerte 2018 die bereits existierende Kategorisierung von Cronin et al. [25] aus dem Jahr 2011. Cronin et al. kategorisierten die Validierungslastfälle anhand ihrer Komplexität in eine globale Ebene, regionale Ebene und lokale Ebene [25]. Fuchs splittete die lokale Ebene von Cronin et al. in eine lokale Ebene und eine Gewebeebe auf (siehe Abbildung 7). Materialtests, um Knochen und andere Gewebe zu charakterisieren, fallen laut Fuchs [40] in die Gewebeebe. In die lokale Ebene fallen Experimente, welche ganze Knochen und Gelenke beinhalten. In der regionalen Ebene werden alle Experimente zusammengefasst, in denen ganze Körperregionen betroffen sind. Die globale Ebene beinhaltet Ganzkörper Experimente. Von der Gewebeebe bis zur globalen Ebene nimmt die Komplexität der Experimente, hinsichtlich Randbedingungen und Freiheitsgrade, zu.

Diese Arbeit befasst sich mit einem Validierungsexperiment auf lokaler Ebene, welches in Kapitel 3 näher beschrieben wird.

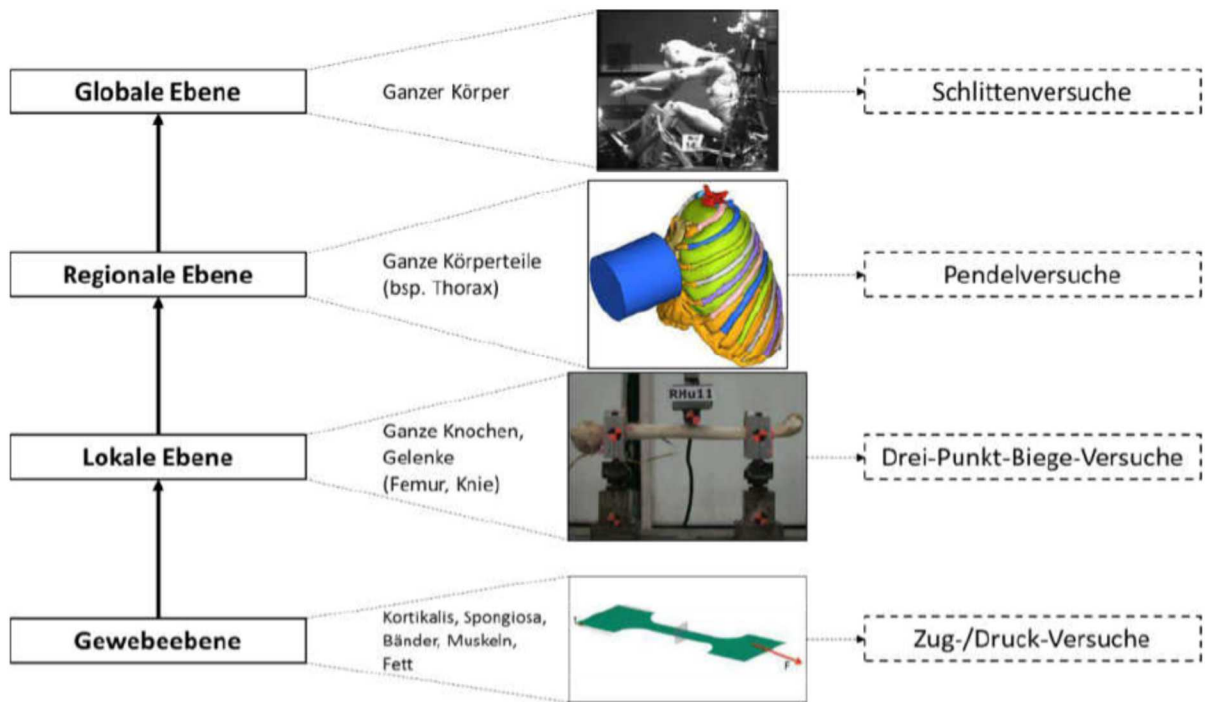


Abbildung 7 – Kategorisierung von Validierungslastfällen nach Fuchs mit Beispielen [40]

1.2.3 Übersicht Validierung von THUMS Version 4 AM 50 und GHBMC AM 50

Die beiden kommerziellen Modelle THUMS Version 4 AM 50 und GHBMC AM 50, die in dieser Arbeit verwendet werden, wurden gegen zahlreiche Freiwilligenversuche und Leichenversuche validiert. Die Abkürzung AM 50 in der Modellbezeichnung steht für die englische Bezeichnung 50th percentile american male. Die Modelle sollen das 50. Perzentil der männlichen amerikanischen Bevölkerung repräsentieren. In Abbildung 8 ist die Anzahl der verwendeten Validierungslastfälle (siehe blaue Ellipsen) pro Körperregion für das GHBMC-Modell dargestellt. Zudem wurde die Anzahl der zur Verfügung stehenden experimentellen Tests für die Validierung (siehe grüne Ellipsen) dargestellt.

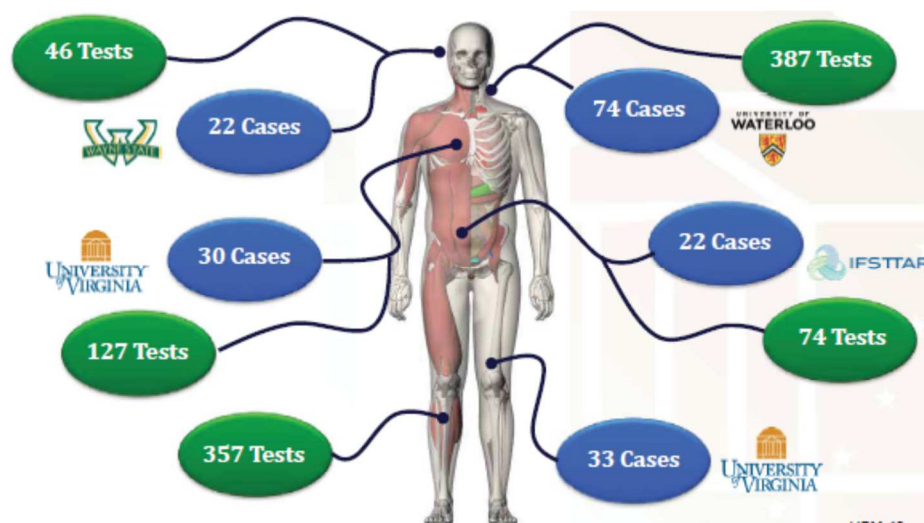


Abbildung 8 – Anzahl der Validierungslastfälle für die Modellentwicklung des GHBMC AM 50 [24]

Das THUMS Version 4 Modell wurde anhand von 27 Lastfällen validiert [66]. Die Abbildungen im Anhang A1 geben einen Überblick über die verwendeten Validierungslastfälle. Das GHBMC-Modell

wurde mit 181 Lastfällen (siehe Abbildung 8) validiert. Die beiden Modelle weisen eine sehr unterschiedliche Anzahl an Validierungslastfällen auf.

Zudem existiert eine Diskrepanz zwischen der in Abbildung 8 veröffentlichten Anzahl an Validierungslastfällen und den veröffentlichten 70 Validierungslastfällen im Handbuch des GHBMC-Modells [45]. Diese Anzahl reduziert sich erneut auf 37 Validierungslastfälle, wenn nur die Validierungslastfälle für die Entwicklung des GHBMC-Modells für den numerischen Solver Virtual Performance Solution betrachtet werden. Es sind keine genauen Angaben verfügbar, wie die Anzahl der Validierungslastfälle für das GHBMC-Modell in Abbildung 8 ermittelt wurde. Eine Ursache der starken Diskrepanz zwischen der im GHBMC-Handbuch [45] veröffentlichten Anzahl und der in Abbildung 8 veröffentlichten Anzahl könnte sein, dass seit der Veröffentlichung des GHBMC-Modells V4.3 und des Handbuchs im Jahr 2014 das Modell immer weiter entwickelt wurde und die Anzahl der Validierungslastfälle bis 2018 auf 181 Validierungslastfälle anstieg.

Die unterschiedliche Anzahl an Validierungslastfällen zwischen den Modellen kann mehrere Ursachen haben. Bei dem GHBMC-Modell werden z.B. experimentelle Studien, welche mit unterschiedlichen Belastungsgeschwindigkeiten durchgeführt wurden, als separate Validierungslastfälle gezählt. Zudem werden bei dem GHBMC-Modell Stabilitätslastfälle durchgeführt, welche ebenfalls als Validierungslastfällen gezählt werden. Diese Stabilitätslastfälle sind hochdynamische Lastfälle, die rein die numerische Stabilität eines Modells bewerten sollen. Derartige Lastfälle sind für das THUMS-Modell nicht gelistet. Die Anzahl der Validierungslastfälle lässt daher keine direkten Schlüsse auf die Qualität der Modellvorhersage eines Modells zu.

Die Referenzen für die Validierungslastfälle beider Modelle sind in den jeweiligen Benutzerhandbüchern veröffentlicht [45, 104].

1.2.4 Konventioneller korridor-basierter Validierungsansatz

Bei dem korridor-basierten Validierungsansatz (siehe Abbildung 9) wird mittels der veröffentlichten experimentellen Kurven ein experimenteller Korridor generiert. Das physikalische Experiment wird numerisch repliziert. Die numerische Modellvorhersage wird mit dem generierten experimentellen Korridor verglichen. Es findet eine Bewertung statt, wie gut die numerische Modellvorhersage in dem experimentellen Korridor liegt. Eine Methode für die Berechnung einer objektiven Bewertungskennzahl ist in der ISO/TS 18571 Norm verankert. Sie vereint mehrere Metriken, die CORA Corridor Rating Methode von Gehre et al. [43] und die EEARTH Rating Methode von Zhan et al. [116] zur Kurvenbewertung in einer Kennzahl. Die CORA-Methode berücksichtigt u.a. wie gut eine numerische Kurve in einen Korridor fällt und die EEARTH Rating Methode untersucht die Kurven hinsichtlich ihrer Charakteristika, wie beispielsweise Phasenverschiebung, Magnitude und Form. Wenn die Qualität der numerischen Modellvorhersage als unzureichend bewertet wird, d.h. die numerische Modellvorhersage nicht zufriedenstellend innerhalb des Korridors liegt, kann ein Optimierungsprozess der Materialeigenschaften des untersuchten Objektes durchgeführt werden, bis die numerische Modellvorhersage sich zufriedenstellend im experimentellen Korridor befindet.

Dieser Validierungsansatz liefert akzeptable Ergebnisse, wenn die Charakteristiken des numerischen Untersuchungsobjektes ähnlich den Charakteristiken der experimentellen Versuchsobjekte sind. Wenn jedoch z.B. numerisch ein Knochenmodell einer kleinen Frau untersucht wird und experimentell die Knochen von großen Männern untersucht wurden, dann müssen bei diesem Validierungsansatz Skalierungstechniken angewendet werden, um eine Vergleichbarkeit herzustellen. Hierbei können nach Davis et al. [28] entweder die numerischen Modelle geometrisch an die experimentellen Untersuchungsobjekte angeglichen werden, um eine Vergleichbarkeit der numerischen Modellvorhersage mit den experimentellen Ergebnissen zu erreichen, oder die numerische Modellvorhersage

wird basierend auf Lastfall spezifischen Skalierungsmethoden auf die Charakteristiken der experimentellen Testproben skaliert.

Die Verwendung von Skalierungsmethoden ist jedoch nach Park et al. [86] nicht geeignet, da die fundamentale Annahme für die Skalierbarkeit, die geometrische Ähnlichkeit zwischen Objekten, für komplexe Strukturen wie z.B. Knochen-Geometrien nicht gegeben ist.

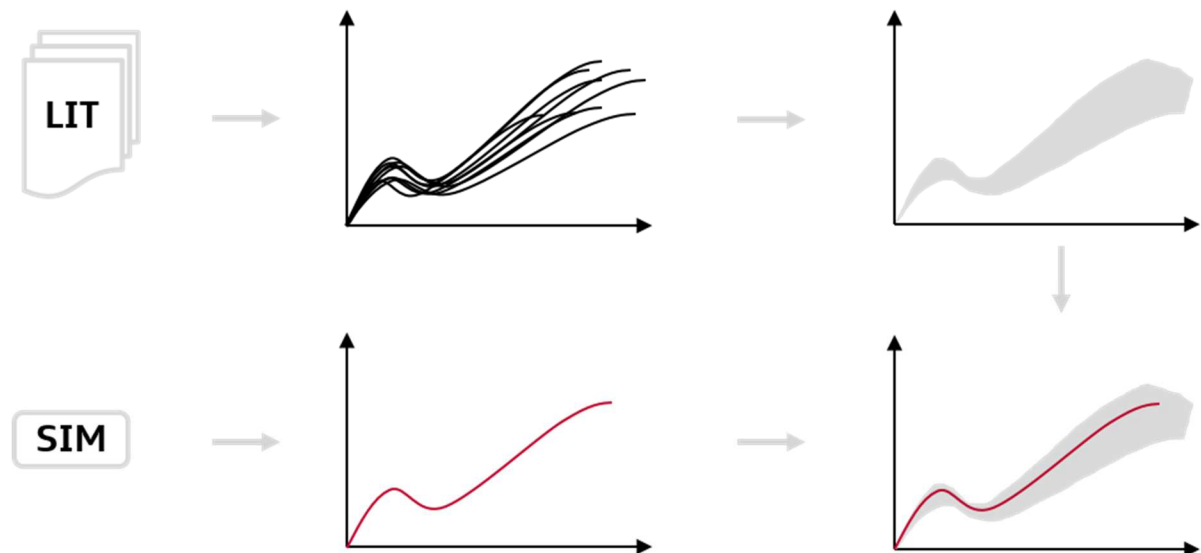


Abbildung 9 – Korridor-basierter Validierungsansatz [93]

1.2.5 Validierungsansätze für parametrisierte Modelle

Der Vorteil von parametrisierten Modellen ist, dass die Geometrie dieser Modelle verändert werden kann. Diese Eigenschaft kann bei der Validierung verwendet werden, um die Geometrie von experimentellen Testproben nachzubilden. In Abbildung 10 ist dieser Validierungsansatz beispielhaft anhand eines Knochenexperiments dargestellt.

Es wird ein numerisches Abbild des physikalischen Experiments erstellt, bei dem zusätzlich die zu untersuchende numerische Testprobe geometrisch an die jeweilige experimentelle Testprobe angeglichen wird. Es wird der Ansatz verfolgt, mögliche geometrische Einflussfaktoren auf die numerische Modellvorhersage anhand einer subjekt-spezifischen Modellierung zu eliminieren. Auf diese Weise wird eine verbesserte Validierung der gewählten Materialeigenschaften ermöglicht. Die numerische Modellvorhersage wird mit der jeweiligen experimentellen Kurve anhand eines objektiven Kriteriums verglichen. Wenn die Güte der Modellvorhersage nicht zufriedenstellend ist, kann ein Optimierungsprozess bezüglich der Materialeigenschaften durchgeführt werden (siehe Kalibrierungsprozess in Abschnitt 1.2.1), bis die Modellvorhersage die benötigte Güte erreicht. Diese geometrische Angleichung der numerischen Testprobe findet für jede experimentelle Testprobe statt, um einen Kurvenvergleich ohne geometrischen Einfluss zu ermöglichen.

Die Problematik in der Anwendbarkeit dieses Validierungsansatzes liegt darin, dass detaillierte geometrische Informationen bezüglich der experimentellen Testproben benötigt werden. Diese Informationen sind in wissenschaftlichen Veröffentlichungen nicht verfügbar. Daher kann dieser Validierungsansatz meist nur von Institutionen angewendet werden, welche die Experimente durchgeführt haben und daher über die notwendigen Informationen, wie z.B. CT-Aufnahmen, verfügen.

In stark vereinfachter Form kann dieser Ansatz auch für Validierungsaktivitäten basierend auf wissenschaftlichen Veröffentlichungen angewendet werden. Bei dem vereinfachten subjekt-

spezifischen Ansatz werden die charakteristischen Eigenschaften der experimentellen Testproben auf das parametrisierte Modell angewandt.

Dies bedeutet, es werden Eigenschaften wie BMI, Alter, Größe und Geschlecht der experimentellen Testprobe verwendet, um ein numerisches Modell basierend auf diesen Eigenschaften zu generieren. Da bei diesem Ansatz nur die Eigenschaften verwendet werden, jedoch nicht die exakte Geometrie repliziert wird, kann die numerische Modellvorhersage nicht 1:1 mit dem experimentellen Ergebnis verglichen werden. Es geht bei diesem vereinfachten subjekt-spezifischen Ansatz um die Untersuchung von gleichen Tendenzen zwischen den experimentellen Ergebnissen und den numerischen Modellvorhersagen [113].

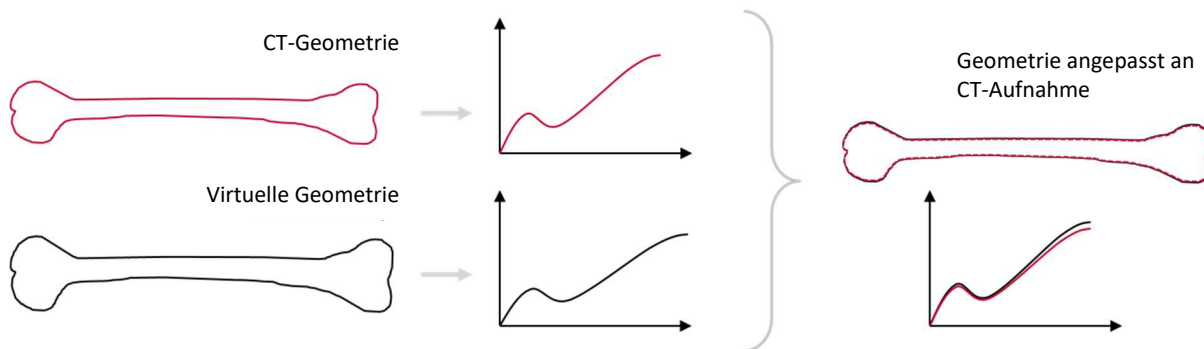


Abbildung 10 – Subjekt-Spezifischer Validierungs- und Kalibrierungsansatz [93]

Klein et al. [68] entwickelten einen Ansatz für den Vergleich von Tendenzen zwischen den numerischen Modellantworten und den experimentellen Ergebnissen. Sie entwickelte ein parametrisiertes Femur-Modell mit den Prädiktoren Alter, BMI und Femurlänge. Das parametrisierte Femur-Modell wurde verwendet, um numerische Femur-Modelle anhand dieser Prädiktoren zu erstellen, die mit den Charakteristiken der experimentellen Femora aus der Studie von Ivarsson et al. [58] übereinstimmten. Die generierten Femora wurden verwendet, um das dynamische Validierungsexperiment von Ivarsson et al. [58] numerisch zu reproduzieren. Klein et al. erzeugten zwei lineare gemischte statistische Modelle (engl. linear mixture models) mit den genannten Prädiktoren. Ein Modell wurde mit den numerischen Frakturkräften generiert, das andere mit den experimentellen Frakturkräften. Nach der statistischen Modellerzeugung wurden die Koeffizienten der Prädiktoren zwischen den statistischen Modellen verglichen. Es wurde für alle Koeffizienten des statistischen Modells, welches mit den experimentellen Frakturkräften generiert wurde, das 95% Konfidenzintervall berechnet. Danach wurde untersucht, ob die Koeffizienten des statistischen Modells, welches mit den numerischen Frakturkräften generiert wurde, in die berechneten Konfidenzintervalle fallen. Wenn dies der Fall ist, kann die Annahme erfolgen, dass die beiden statistischen Modelle denselben Tendenzen folgen. Das parametrisierte Modell kann in diesem Fall als validiert angesehen werden.

1.2.6 Limitationen aktueller Validierungsmethoden

In diesem Abschnitt werden einige Aspekte beschrieben, welche die Güte von Validierungsaktivitäten negativ beeinflussen können. Wissenschaftliche Veröffentlichungen von experimentellen Studien weisen oftmals einen unzureichenden Detaillierungsgrad in der Beschreibung von einem oder mehreren der folgenden Aspekte auf: des experimentellen Aufbaus, der Versuchsdurchführung, der experimentellen Objekte oder der experimentellen Ergebnisse. Jede fehlende oder unzureichend beschriebene Information bezüglich der genannten Aspekte beeinflusst die Qualität eines numerischen Abbilds und somit direkt die Güte der Validierung. Auf diese Problematik wird in Kapitel 3 näher eingegangen.

Zudem existieren zahlreiche experimentelle Studien, die eine sehr geringe Anzahl an Testproben aufweisen [1, 36]. Bei diesen Studien stellt sich die Frage der Sinnhaftigkeit, einen experimentellen Korridor mittels statistischer Analysen für Validierungszwecke zu entwickeln.

Überdies ist es nahezu unmöglich, bei einer geringen Anzahl an Testproben, Schlussfolgerungen hinsichtlich der realen biomechanischen Varianz der Bevölkerung zu ziehen.

Des Weiteren wurden zahlreiche experimentelle Studien mit Testproben durchgeführt, die ein sehr fortgeschrittenes Alter aufweisen [57, 59, 72]. Objekte fortgeschrittenem Alters sind fragiler, da sich der Knochenquerschnitt und die Knochensubstanz während des Alterungsprozesses verändert [38, 90]. Wenn Menschmodelle mit Daten aus derartigen Validierungslastfällen verglichen werden, beschränkt sich ihre Biofidelität auf die eine Bevölkerungsgruppe fortgeschrittenem Alters. NHTSA veröffentlichte eine Studie über die Haupttodesursachen der amerikanischen Bevölkerung für die Jahre 2010–2011 kategorisiert nach Altersgruppen [106]. Verkehrsunfälle zählen zu den fünf bedeutendsten Todesursachen bei der Altersgruppe 25–44 Jahren. Bei der Altersgruppe von 45–64 Jahre stehen Verkehrsunfälle auf Platz 9 und bei der Altersgruppe > 65 Jahre überwiegen krankheitsbedingte Ursachen. Es ist daher notwendig Methoden zu entwickeln, um die Biofidelität von Menschmodellen auch für die jüngere Bevölkerungsgruppe zu erreichen, mit dem Ziel, das Schutzpotenzial von Fahrzeugsicherheitssystemen möglichst für die gesamte Bevölkerung genau bewerten zu können.

2 Zielsetzung

2.1 Zielsetzung und wissenschaftliche Fragestellung

Menschmodelle werden gegen zahlreiche Validierungslastfälle verglichen, um ein Vertrauen in die numerischen Modellvorhersagen zu generieren und ein menschenähnliches, biofideles Verhalten zu erlangen. Dies wurde in Abschnitt 1.2 bereits detailliert erläutert. Neue Entwicklungstrends in der Menschmodellentwicklung umfassen u.a. die Entwicklung parametrisierter Modelle (vgl. Abschnitt 1.1.3). Diese parametrisierten Modelle bilden die Basis für den effizienten Einsatz stochastischer Simulationen. Bei Entwicklungsprozessen in der Fahrzeugsicherheit muss der gesamte Bereich vom 5. Perzentil bis zum 95. Perzentil der Bevölkerung berücksichtigt werden. Die Bewertung des Insassenschutzpotenzials für neue Sicherheitstechnologien könnte mit Hilfe stochastischer Simulationen mit Menschmodellen umsichtiger bewertet werden, als mit den existierenden drei Perzentilen der ATDs, da mit stochastischen Simulationen der gesamte Parameterraum umfassend untersucht werden kann (siehe Abbildung 11).

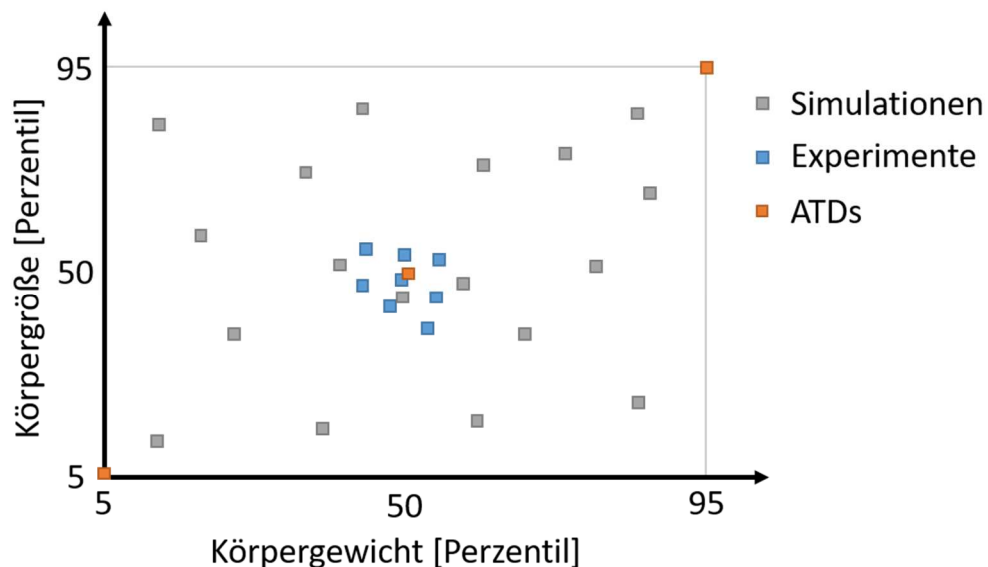


Abbildung 11 – Vorteil stochastischer Simulationen gegenüber den drei existierenden Perzentilen der ATDs

Die charakteristischen Eigenschaften von Testproben in experimentellen Studien, wie z.B. Körpergewicht und Körpergröße, decken meist nur einen kleinen Unterraum des gesamten Parameterraumes ab, welcher numerisch untersucht werden soll und für Entwicklungsprozesse relevant ist (vgl. Abbildung 11).

Bei der Validierung parametrisierter Modelle wird oftmals der subjekt-spezifische Validierungsansatz, welcher bereits in Abschnitt 1.2.5 beschrieben wurde, angewendet. Es erfolgt dabei die folgende vereinfachende Annahme: wenn das parametrisierte Modell gegen die experimentellen Testproben validiert wurde, gilt das Modell auch in anderen Zuständen als validiert, beispielsweise mit einer Geometrie abweichend von den Geometrien der experimentellen Testproben. In der Realität verlässt jedoch das parametrisierte Modell den validierten Bereich, wenn Modellvorhersagen mit charakteristischen Eigenschaften generiert werden, die von den charakteristischen Eigenschaften der experimentellen Testproben abweichen.

Es stellt sich daher die Frage, wie Modellvorhersagen mit charakteristischen Eigenschaften, welche von den charakteristischen Eigenschaften der experimentellen Testproben abweichen, hinsichtlich ihrer Biofidelität bewertet werden können. Zudem stellt sich die Fragestellung, wie stochastische Populationen, also eine hohe Anzahl unterschiedlicher Modellvorhersagen, effektiv validiert werden können.

Zielsetzung dieser Arbeit ist, eine populationsbasierte Validierungsmethode zu entwickeln, mit der eine Population von numerischen Modellvorhersagen gegen eine Population experimenteller Ergebnisse validiert werden kann. Der Begriff *Population* kann hierbei als Kurvenschar angesehen werden, bei der die individuellen Kurven charakteristische Eigenschaften repräsentieren.

2.2 Struktur der Dissertation

Die Struktur dieser Arbeit ist in Abbildung 12 visualisiert. In Kapitel 3 wird ein geeignetes Validierungsexperiment ausgewählt, welches in den darauffolgenden Kapiteln als Referenzlastfall verwendet wird. Zudem wird für dieses Validierungsexperiment, einem dynamischen 3-Punkt-Biegeversuch des Femurs, ein numerisches Abbild erstellt und dieser numerische Versuchsaufbau umfassend verifiziert. Ziel von Kapitel 4 ist die Untersuchung, ob Clusteranalyseverfahren und interne Validierungsindizes zur Bestimmung einer geeigneten Clusteranzahl (Abkürzung: int. VIs) für die Anwendung an biomechanischen Datensätzen geeignet sind. Diese Eignungsprüfung wird mit vereinfachten Testdatensätzen durchgeführt, welche charakteristische Eigenschaften realer biomechanischer Datensätze, wie z.B. variierende Frakturzeitpunkte oder -kraftniveaus, enthalten.

In Kapitel 5 wird mit den geeigneten Clusteranalyseverfahren aus Kapitel 4 eine populationsbasierte Validierungsmethode entwickelt, um den Grad der Biofidelität einer stochastischen Population beurteilen zu können. Die Validierungsmethode verwendet Clusteranalyseverfahren für den Vergleich einer Population an numerischen Modellvorhersagen mit einer Population an experimentellen Ergebnissen. Die Güte der numerischen Modellvorhersagen wird anhand eines entwickelten Bewertungskonzeptes bestimmt.

Zusätzlich wird in Kapitel 5 zur Überprüfung des entwickelten populationsbasierten Validierungskonzeptes eine numerische Population an Femora erstellt. Für diese numerische Population werden die Modellvorhersagen mit dem numerischen Versuchsaufbau des Referenzlastfalls aus Kapitel 3 erzeugt. Der kombinierte Datensatz aus numerischen Modellvorhersagen und experimentellen Ergebnissen wird einer Clusteranalyse unterzogen. Die Güte der numerischen Modellvorhersagen wird durch die populationsbasierte Validierungsmethode bestimmt und das entwickelte Validierungskonzept anhand dieses Referenzlastfalls überprüft.

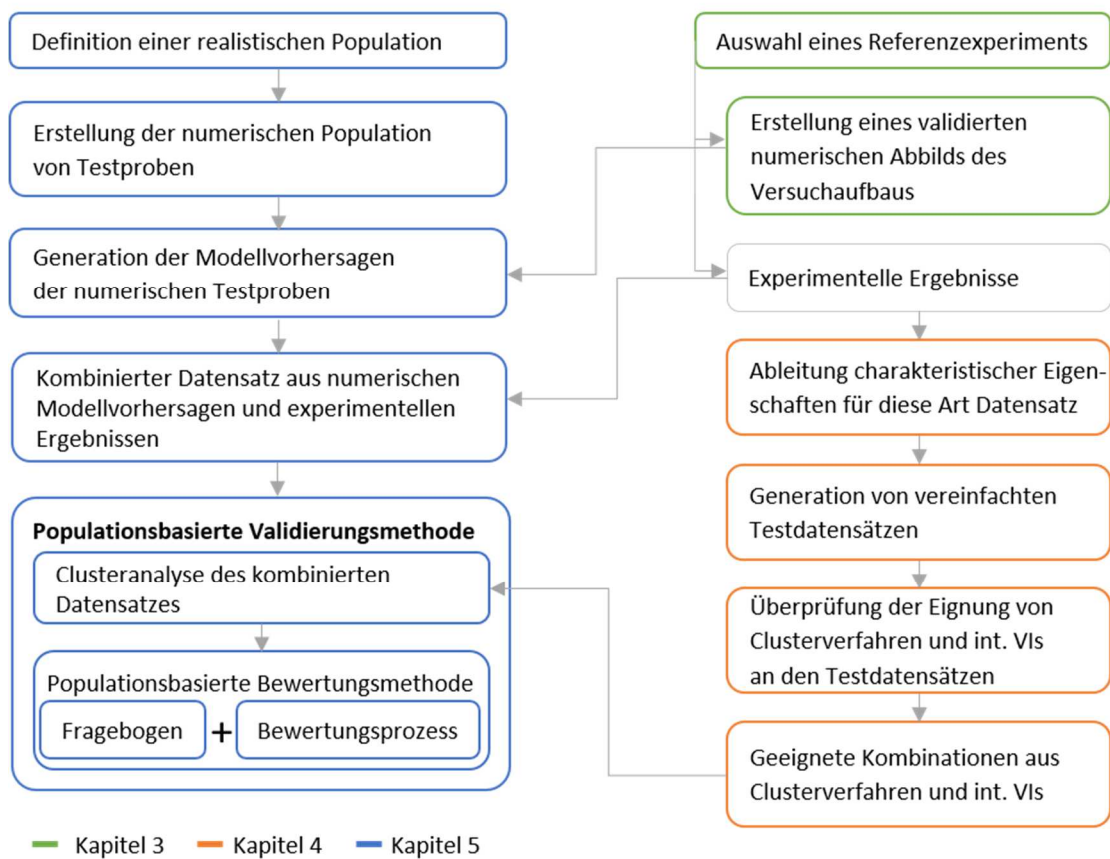


Abbildung 12 – Struktur der Dissertation und Verknüpfungen der einzelnen Kapitel

3 Entwicklung des numerischen Modells eines Referenzexperiments

3.1 Einleitung

3.1.1 Zielsetzung

Das Ziel dieses Kapitels ist die Erstellung eines numerischen Modells als Abbild eines dynamischen Referenzexperiments auf lokaler Ebene. Das Referenzexperiment und sein numerisches Abbild werden für die Anwendung der populationsbasierten Validierungsmethodik in Kapitel 5 verwendet.

3.1.2 Auswahl des dynamischen 3-Punkt-Biegeversuchs des Femurs nach Forman et al.

Für das Femur existiert eine hohe Datenverfügbarkeit hinsichtlich Materialcharakterisierung und anthropometrischer Varianz. Takahashi et al. [101] bieten einen Überblick über vorhandene Literatur zur Materialcharakterisierung des Femurs. In dieser Veröffentlichung wurden Daten zur Femur-Materialcharakterisierung aus unterschiedlichen Quellen zusammengeführt. Ziel war die Untersuchung der Varianz der einzelnen Materialparameter in Bezug auf das Alter. Es wurden Alterungsfunktionen für diejenigen Materialparameter erstellt, welche für die Definition eines elastisch-plastischen Materialgesetzes zur Charakterisierung des Femurs benötigt werden.

Zudem existiert für die anthropometrische Charakterisierung des Femurschaft-Querschnitts eine öffentlich verfügbare US-Datenbank der NHANES-Studie [10], welche für die Populationsdefinition in Kapitel 5 verwendet werden kann.

Einen Überblick über existierende Femur-Biegeversuche bis zum Jahr 2004 geben Kennedy et al. [62] und Kerrigan et al. [65]. In nachfolgenden Jahren wurden zusätzliche Versuche von Forman et al. [38] und Ivarsson et al. [58] veröffentlicht.

Für diese Dissertation wurde der dynamische 3-Punkt-Biegeversuch von Forman et al. [38] als Referenzexperiment ausgewählt. Kriterien für die Auswahl dieses Versuches sind in Tabelle 1 dargestellt. Ein offensichtlicher Nachteil bei der Auswahl dieses Experiments sind die geometrischen Unsicherheiten aufgrund der nicht definierten Auflagergeometrie. Diese Unsicherheiten werden in Abschnitt 3.2.4 speziell untersucht, um die wahrscheinlichste Parameterkonfiguration für die Auflagerdefinition zu finden.

Eine größere biomechanische Relevanz hat das Referenzexperiment für den Fußgängerschutz als für den Insassenschutz. Forman et al. [37] veröffentlichten 2019 eine Untersuchung über Verletzungstrends in gegurteten Frontalkollisionen mittels einer Analyse der Datenbank NASS-CDS bezogen auf die Jahre 1998 bis 2015. Die Analyse zeigt, dass vor allem für die unteren Extremitäten und Schädelfrakturen das Verletzungsrisiko in neueren Fahrzeugmodellen im Vergleich zu Älteren deutlich gesunken ist. Auftretende Femurschaft-Frakturen bei Fahrern wurden bei neueren Fahrzeugmodellen ab 2009 im Vergleich zu älteren Fahrzeugmodellen (1989–2008) um mindestens 50% reduziert.

Bei einer Fußgänger-Kollision treten Frakturen des Femurs häufig auf. Die amerikanische Verkehrsbehörde NHTSA veröffentlichte 2001 eine Untersuchung von Fußgängerunfällen [102]. Die Datenerhebung begann 1994 und endete 1998. Während dieser Zeit wurden 511 Fußgängerunfälle an sechs

verschiedenen geographischen Positionen in den Vereinigten Staaten untersucht. Laut dieser Studie erlitten diese Fußgänger 4184 Verletzungen. Verletzungen der unteren Extremitäten waren mit 33% die am häufigsten verletzte Körperregion, gefolgt von den oberen Extremitäten mit 20%, dem Kopf mit 17% und dem Gesicht mit 16%. Unter den 10 häufigsten Verletzungen befinden sich Frakturen des Femurs an 7. Stelle mit 43 Verletzungsfällen. Tibia-Frakturen stehen mit 130 Fällen an 2. Stelle, Fibula-Frakturen mit 113 Fällen an 3. Stelle und Beckenfrakturen mit 59 Fällen an 5. Stelle.

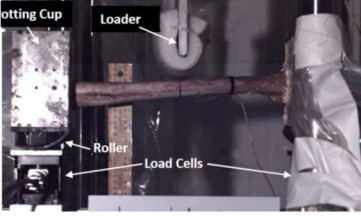
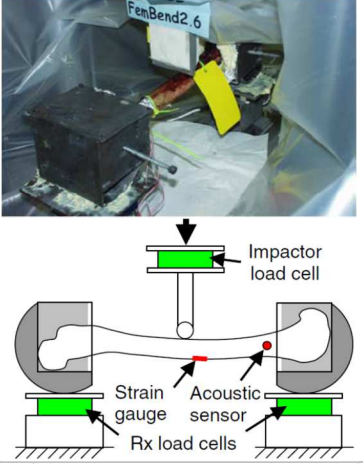
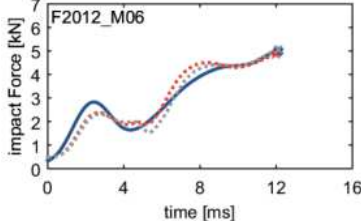
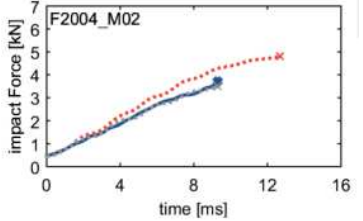
Tabelle 1 – Kriterien für die Auswahl des dynamischen Biegeversuchs nach Forman et al. [38]

Isolierter Knochen/Gewebe/Gelenk		Femur
Dynamischer Lastfall		Ja
Geometrie- definition	Auflager	Nicht definiert
	Schaum	Definiert
	Impaktor	Definiert
Testproben	Anzahl	15
	Alter	19–57
	Mehrfachverwendung	Nein
	Anthropometrische Angaben	Flächenträgheitsmoment in der Mitte des Femurschafts Femurlänge
	Behandlung der Testproben	Spezifiziert
Rand- bedingungen	Impaktorlage	Mitte des Femurschafts
	Impaktorgeschwindigkeit	1,5 m/s
	Knocheneinbettung	Definiert
Ergebnisse	Alle Testergebnisse vorhanden	Ja; Impaktkraft-Zeit-Kurve
	Filterung	Channel Frequency Class (CFC) 180
	Simulationsnachbereitung	Nullpunkt der Impaktkraft-Zeit Kurve definiert, wenn Impaktkraft 0,4 kN über- schreitet

3.1.3 Besonderheit in den experimentellen Kurvenverläufen bei Forman et al.

Die Unterschiede zwischen den beiden dynamischen 3-Punkt-Biegeversuchen des Femurs von Forman et al. [38] und Funk et al. [42] sind in Tabelle 2 gelistet.

Tabelle 2 – Vergleich der Experimente von Forman et al. und Funk et al.

	Forman et al. [38]	Funk et al. [42]
Impaktorgeschwindigkeit (m/s)	1,5	1,2
Schaum um Impaktor	Ja	Nein
Auflagergeometrie		
Exemplarische Versuchskurve (blaue Kurve)		

Bezüglich der Auflagergeometrie des in Funk et al. [42] verwendeten Versuchsaufbaus existiert eine sichtbare Diskrepanz zwischen der veröffentlichten Skizze in Funk et al. [42] und einer veröffentlichten Abbildung des von Funk et al. [42] verwendeten Versuchsaufbaus in der Veröffentlichung von Park [85].

Charakteristisch für die experimentellen Kurven von Forman et al. [38] ist ein schwingungsbehafteter Kurvenverlauf mit einem ausgeprägten ersten lokalen Maximum in Form einer niederfrequenten Schwingung mit markanter Amplitude. Im weiteren Kurvenverlauf ist die Amplitude signifikant geringer. Die Kurvenverläufe von Funk et al. [42] verlaufen nahezu linear.

Für die Ursache der Schwingungen in den Kurvenverläufen bei Forman et al. [38] sind zwei Hypothesen zu klären, um die belastbare Verwendung der Daten zu ermöglichen:

- Die Ursache der Schwingungen bei Forman et al. [38] liegt in einer zu geringen Steifigkeit der Auflager. Die Auflager werden während der Versuchsdurchführung vor dem Start des Abrollvorgangs deformiert. Diese Deformation erzeugt das signifikante erste lokale Maximum.
- Die Ursache der Schwingungen bei Forman et al. [38] liegt in der Massenträgheit der Auflager. Die Auflager besitzen eine zu große Masse. Die Massenträgheit die vor dem Start der Abrollbewegung überwunden werden muss, verursacht das erste signifikante lokale Maximum.

Diese Hypothesen werden in den Abschnitten 3.2.3–3.2.4 näher untersucht.

3.2 Methoden

3.2.1 Numerische Modellerstellung des Ausgangsmodells für die Sensitivitätsanalysen

Geometrische Definition des Versuchsaufbaus

Der Impaktor aus Aluminium besitzt einen Durchmesser von 13 mm und ist mit einem Schaum von ca. 25 mm Dicke umwickelt [38]. Der Impaktor ist als Starrkörper definiert. Die Dichte des Schaummaterials ist nicht veröffentlicht und ist mit einer Dichte von 50 kg/m^3 modelliert. Das Kompressionsverhalten des Schaumes ist nach Park [85] definiert. Außer dem veröffentlichten Foto des experimentellen Versuchsaufbaus in Abbildung 13, ist die Geometrie und das Material der Auflager nicht näher definiert. Die englische Bezeichnung für Auflager ist Roller (vgl. Abbildung 13). In den Darstellungen der Ergebnisse im Anhang wird anlehnend an Abbildung 13 der Begriff Roller verwendet. Die Geometrie der Auflager ist mit Schalenelemente modelliert. Die Auflagertiefe wird auf 80 mm festgelegt. Den Schalenelementen wird eine Dicke von 2 mm zugewiesen. Die Auflager sind als Stahl charakterisiert.

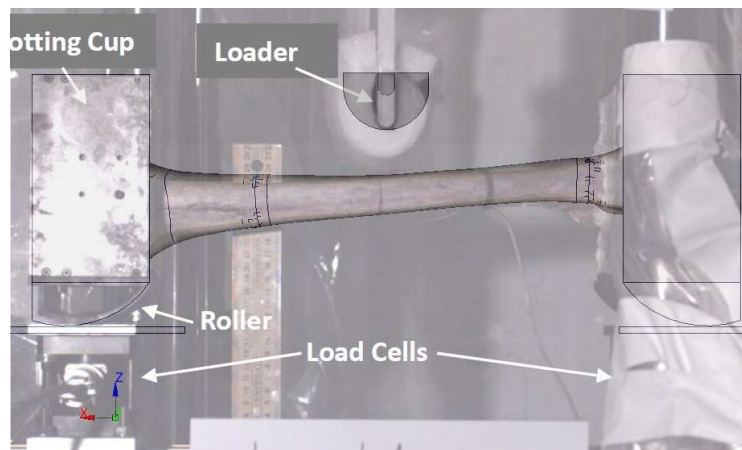


Abbildung 13 – Experimenteller Versuchsaufbau [38] überlagert mit numerischem Ausgangsmodell

Ein Urethanschaum wurde in den Experimenten verwendet, um den Knochen an seinem distalen und proximalen Ende in die Auflager einzubetten [38]. Dieser Einbettungsblock ist mit 3D-Elementen und einer Dichte von 1000 kg/m^3 modelliert.

An der Schnittstelle zwischen Einbettungsblock und Auflagergehäuse besitzen die Bauteile gemeinsame Knoten. Diese Modellierung führt zu einem Auflagergewicht von 2,0 kg.

Die Bodenplatte der Auflager ist als Starrkörper definiert. Die Kontaktfläche zwischen Bodenplatte und Auflager wurde in den Experimenten geschmiert, um den Rollwiderstand so weit wie möglich zu reduzieren [38]. Der numerische Reibwertkoeffizient ist mit $\mu = 0,2$ definiert.

Positionierung des Knochens

Der Knochen ist so positioniert, dass der Mittelpunkt des Femurkopfes (lat.: caput ossis femoris), welches dem Rotationspunkt des Femurkopfes in der Beckenpfanne entspricht, und die distale femorale Kerbe (lat.: fossa intercondylaris) durch eine horizontale Linie zu verbinden sind (siehe Abbildung 14a) [38]. Der Knochen ist in x-Richtung so eingebettet, dass jeweils die äußersten proximalen und distalen Punkte des Knochens an das Auflagergehäuse anstoßen. In der Draufsicht des Versuchsaufbaus sollte der Knochen parallel zur x-Achse positioniert sein, damit der Impaktor die Knochenpositionierungslinie im rechten Winkel trifft (siehe Abbildung 14b).

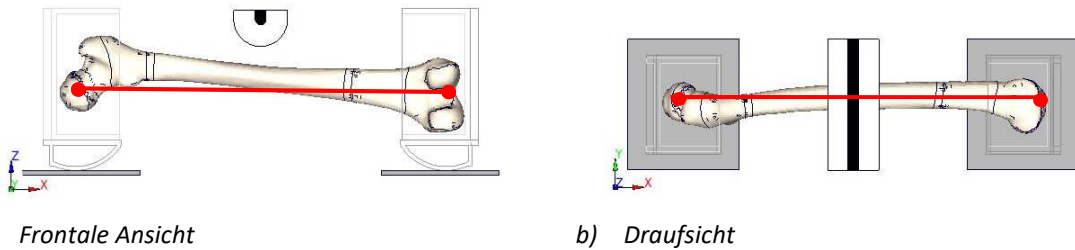


Abbildung 14 – Knochenpositionierung

Definition der Randbedingungen

Die Freiheitsgrade des Impaktors sind beschränkt, damit eine reine Abwärtsbewegung ermöglicht wird. Die Impaktorgeschwindigkeit von 1,5 m/s greift am Schwerpunkt des definierten Starrkörpers des Impaktors an. Der Impaktor ist so ausgerichtet, dass er das Femur in der Mitte des Femurschaftes trifft. Die horizontale Knochenpositionierungslinie und die vertikale Linie der Impaktorbewegung treffen sich in einem rechten Winkel [38].

Das Modell wird für die Sensitivitätsanalysen in zwei Konfigurationen verwendet. Bei dem numerischen Modell mit starrer Konfiguration erfolgt die Einbettung des Knochenendes in die Auflager durch die Definition eines Starrkörpers, welcher das Auflagergehäuse, den Einbettungsblock und das Knochenende beinhaltet. Für die deformierbare Konfiguration erfolgt die Starrkörperdefinition nur mit dem oberen rechteckigen Teil des Auflagergehäuses, dem Einbettungsblock und dem Knochenende. Der runde Teil des Auflagergehäuses bleibt deformierbar (siehe Abbildung 16).

Simulationdurchführung und Nachbereitung

Alle numerischen Simulationen werden mit Virtual Performance Solution (Abkürzung: VPS) Version 2015 mit doppelter Präzision des Löser durchgeföhrt, um den Einfluss numerischer Rundungsfehler so weit wie möglich zu begrenzen. Bei allen numerischen Simulationen wird die Energiebilanz, die Hourglass-Energie und die dynamische Massenzunahme kontrolliert. Zudem wird visuell die Kinematik überprüft. Die Nachbereitung der Simulationsergebnisse erfolgt mit der Software Animator Version 2.3.4.

Als Simulationsergebnis wird der Kontaktkraft-Zeit-Verlauf zwischen den Bodenplatten und den Auflagern extrahiert und mit dem Filter CFC180 der Channel Frequency Class analog zu den experimentellen Ergebnissen gefiltert.

Wenn die Simulationsergebnisse mit den experimentellen Ergebnissen verglichen werden, z.B. in Abbildung A2.7.1c, wird die Zeitachse bei der Modellvorhersage analog zu den experimentellen Kurven neu definiert. Der Nullpunkt der Zeitachse wird an dem Punkt der Kontaktkraft-Zeit-Kurve definiert, wenn die Kontaktkraft 0,4 kN überschreitet [85].

Die experimentellen Ergebnisse sind in [85, 86] veröffentlicht. Wenn die Kontaktkraft 0,4 kN erreicht, ist der Schaum zwischen Knochen und Impaktor vollständig komprimiert.

3.2.2 Methodik für die Definition der Fehler- und Sensitivitätsanalysen

Bei der Validierungsstrategie für den in Abschnitt 3.2.1 beschriebenen numerischen Versuchsaufbau liegt der Fokus auf der Überprüfung der Physik, sowie der Einschätzung von Fehlern und Unsicherheiten (siehe Abbildung 15). Ein weiterer Fokus liegt auf der Ergründung der in Abschnitt 3.1.3 beschriebenen auftretenden Schwingungen bei den experimentellen Kurven. Sensitivitätsstudien, bei denen die Masse und die Struktursteifigkeit der Auflager variiert werden, sollen dieses Phänomen erklären.

Ziel der Fehleranalysen ist eine Beurteilung der Ungenauigkeit, welche durch mangelnde Informationen und vereinfachende Modellannahmen entsteht. Mittels der Sensitivitätsanalysen wird das Ziel verfolgt, wichtige Parameter zu identifizieren, welche die Modellvorhersagen signifikant beeinflussen.

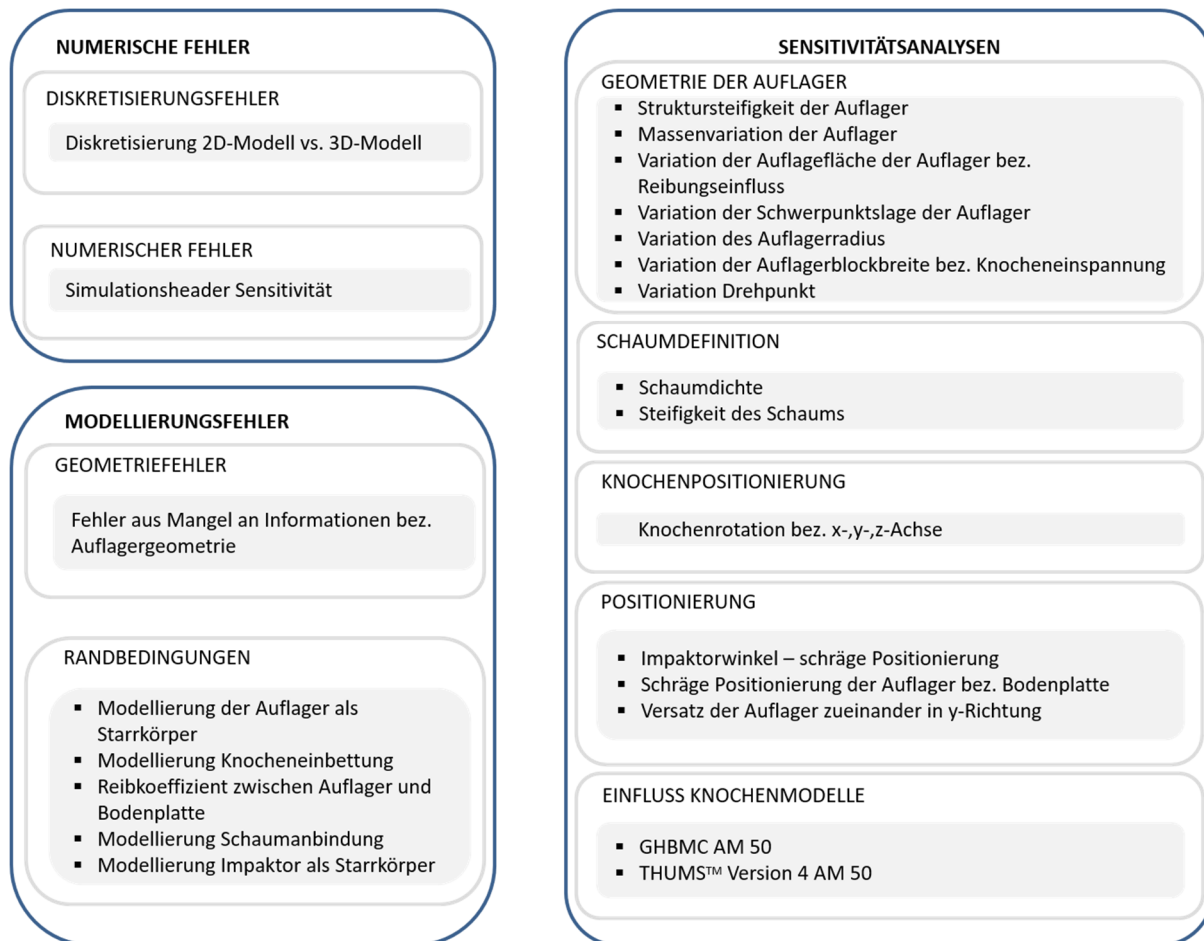


Abbildung 15 – Analysen für die kritische Beurteilung der Güte des numerischen Versuchsaufbaus

Bei den Sensitivitätsanalysen liegt der Fokus auf zwei wesentlichen Aspekten. Zum einen werden Parameter variiert, wie z.B. die Auflagertiefe, welche nicht publiziert wurden. Dies ist wichtig, um die gewählte Parameterkombination für das in Abschnitt 3.2.1 entwickelte numerische Modell zu überprüfen, welche den experimentellen Aufbau repräsentiert. Zum anderen werden Sensitivitätsanalysen von publizierten Parametern, wie z.B. der Knochenlage, durchgeführt. Bei dieser Art von Parameter ist es wahrscheinlich, dass während der experimentellen Durchführung eine Abweichung vom definierten Sollwert auftritt.

Um den geometrischen Modellierungsfehler aufgrund mangelnder Informationen bezüglich der Auflagergeometrie zu ermitteln, wird das 2D-Ausgangsmodell mit dem realitätsnahen numerischen 3D-Modell verglichen, welches in Abschnitt 3.2.9 beschrieben wird. Eine quantitative Validierung der Modellvorhersagen, wie z.B. mit der CORA-Bewertungsmethodik [44], findet nicht statt. Es wird qualitativ die Güte der Modellvorhersage, durch einen visuellen Kurvenvergleich der numerischen Modellvorhersage mit den experimentellen Ergebnissen, in Abschnitt 3.2.9 betrachtet. Ziel ist eine Überprüfung, ob die Modellvorhersage hinsichtlich der Kurvenform und des Versagenszeitpunktes eine plausible Repräsentation der Experimente darstellt. Eine Überprüfung der Implementierung des Algorithmus erfolgt nicht.

3.2.3 Sensitivitätsanalysen bezüglich Modellierungsannahmen

Starrer vs. deformierbarer Versuchsaufbau

Die Untersuchung umfasst zwei Möglichkeiten für die Modellierung der Auflager (vgl. Abbildung 16). Bei dem starren Versuchsaufbau umfasst die Starrkörperdefinition das Auflagergehäuse, den Einbettungsblock, in dem das Knochenende eingebunden ist, sowie das Knochenende. Bei dem deformierbaren Versuchsaufbau wird der untere runde Teil der Auflager nicht in die Starrkörperdefinition mit einbezogen.



Abbildung 16 – Skizze starrer vs. deformierbarer Versuchsaufbau

Variation Reibkoeffizient zwischen Auflager und Bodenplatte

In dieser Studie wird der Reibkoeffizient zwischen Auflager und Bodenplatte mit folgenden Werten variiert: keine Reibung, 0,05, 0,1, 0,2, 0,5, 1,0 und 3,0. Die Studie verwendet den starren 2D-Versuchsaufbau.

Variation Schaumanbindung an Impaktor

Die Anbindung zwischen Schaum und Impaktor wird auf zwei Arten realisiert. Bei einer Modellierung teilen sich Schaum und Impaktor an der Schnittstelle die Knoten. Bei der zweiten Modellierung werden die Bauteile Impaktor und Schaum unabhängig voneinander erzeugt. Die Schnittstelle wird mit einem Kontakt modelliert. Zudem werden die äußersten Knoten des Schaums an der Schnittstelle zum Impaktor in die Starrkörperdefinition des Impaktors mit aufgenommen. Die Studie wird mit starrem 2D-Versuchsaufbau durchgeführt.

Vereinfachte vs. realistische Knocheneinbettung

Diese Studie beinhaltet den Vergleich einer vereinfachten Knocheneinbettung des starren Versuchsaufbaus mit einer realitätsnäheren Knocheneinbettung. Bei der realitätsnäheren Knocheneinbettung wird mit einer booleschen Operation das Volumen des Knochenendes aus dem Einbettungsblock herausgeschnitten. Die Schnittstelle zwischen Knochenende und Einbettungsblock wird als idealisierte Anbindung mit gemeinsamen Knoten modelliert. Das PMMA-Material wird mit einem elastischen Materialgesetz charakterisiert (E-Modul: 2,0 GPa, Querdehnungskennzahl: 0,3, Dichte: 1,5 g/cm³). Das Modell mit realitätsnäherer Knocheneinbettung wird in zwei Varianten simuliert. Eine Variante wird mit deformierbaren Einbettungsblöcken und starrem Auflagergehäuse aus Stahl definiert. Die andere Variante wird als kompletter Starrkörper simuliert und mit dem starren Ausgangsmodell ohne boolesche Operation verglichen. Für diese Studie wird der 3D-Versuchsaufbau (siehe Abschnitt 3.2.9) verwendet.

Impaktor als Starrkörper oder als deformierbares Aluminiumteil

Es wird die vereinfachende Annahme untersucht, den Impaktor als Starrkörper zu modellieren, anstatt als deformierbares Aluminium-Bauteil. Das Aluminium ist mit einem elastisch-plastischen Materialgesetz und einer Dichte von $2,75 \text{ g/cm}^3$ charakterisiert. Die Studie verwendet den starren 3D-Versuchsaufbau (siehe Abschnitt 3.2.9).

3.2.4 Sensitivitätsanalysen zur Festlegung der Geometrie der Auflager

Variation der Struktursteifigkeit der Auflager

Bei dem Ausgangsmodell wird die Materialstärke des Auflagergehäuses variiert. Es werden folgende Materialstärken untersucht: 2 mm (festgelegte Stärke des Ausgangsmodells), 3 mm, 5 mm und 7 mm. Die Materialdichte des Auflagergehäuses wird für jede Variante angepasst, damit das Auflagergewicht konstant bleibt und dem Gewicht des Ausgangsmodells mit 2 mm Dicke entspricht. Die Studie wird in deformierbarer Konfiguration des 2D-Ausgangsmodells durchgeführt.

Variation der Auflagertiefe

Es wird die Auflagertiefe des Ausgangsmodells in y-Richtung variiert. Die Auflagertiefe des Ausgangsmodells wird auf das 1,5-fache und 2,0-fache erhöht. Die Studie wird in zwei Ausführungen durchgeführt. Bei einer Studie wird die Auflagertiefe variiert ohne das Auflagergewicht zu verändern. Für jede Variante dieser Studie wird die Dichte des Einbettungsblockes und des Auflagergehäuses angepasst, damit das Gesamtgewicht des Auflagers konstant bleibt und dem Gewicht der Ausgangskonfiguration (1,0-fache Auflagertiefe) entspricht.

Die zweite Studie variiert die Auflagertiefe, um das Auflagergewicht zu verändern (siehe Tabelle 3). Diese Studie wird ohne Reibung zwischen Bodenplatte und Auflager durchgeführt. Beide Studien verwenden das 2D-Modell in deformierbarer Konfiguration.

Tabelle 3 – Variantendefinition und Gewicht der Auflager

Auflagertiefe	Auflagermasse [kg]
1,0-fache Auflagertiefe (Basis)	2,0
1,5-fache Auflagertiefe	2,82
2,0-fache Auflagertiefe	3,63

Variation der Schwerpunktslage der Auflager

Diese Studie variiert die Lage des Schwerpunktes der Auflager (siehe Abbildung 17). Die homogene Gewichtsverteilung des Einbettungsblockes wird dahingehend verändert, dass das gesamte Gewicht des Einbettungsblockes nur auf die roten Bereiche homogen verteilt ist. Das restliche Volumen des Einbettungsblockes besitzt kein Gewicht. Auf diese Weise wird der Schwerpunkt je um ca. 12 mm vertikal in z-Richtung (siehe Abbildung 17c und d) und je um ca. 6 mm horizontal in x-Richtung (siehe Abbildung 17a und b) verschoben. Die Studie verwendet die realitätsnahe 3D-Modellierung des Auflagergehäuses in starrer Konfiguration.

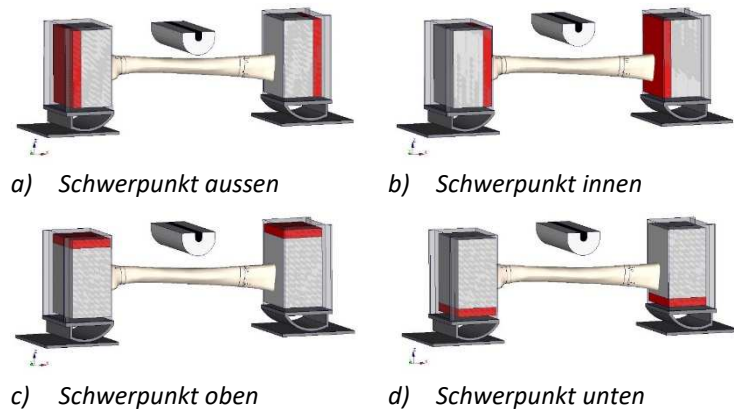


Abbildung 17 – Variantendefinitionen für Schwerpunktsverlagerung

Variation des Auflagerradius

Bei dieser Studie wird nur der runde Teil des Auflagergehäuses variiert. Die Dichte des Auflagergehäuses wird für jede Variante dementsprechend angepasst, damit das Gewicht des Auflagergehäuses immer dem Gewicht des Ausgangsmodells entspricht. Der Drehpunkt der Modelle befindet sich immer an derselben Stelle. Der runde Teil der Auflager des Ausgangsmodells wird für die Definition der Varianten jeweils um das 0,75-fache gestaucht (siehe Abbildung 18b) oder um das 1,25-fache gestreckt (siehe Abbildung 18c). Die Studie verwendet das 2D-Ausgangsmodell in der starren Konfiguration.

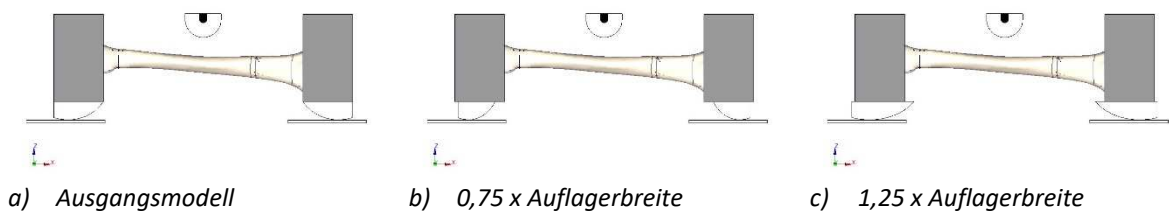


Abbildung 18 – Variation des Auflagerradius

Variation Knocheneinspannung über Variation der Auflagerblockbreite

Es wird der obere Teil der Auflager in der Breite variiert. Für die Variantendefinition wird die Breite des Auflagerblocks des Ausgangsmodells (siehe gestrichelte Linie in Abbildung 19) um das 0,75-fache gestaucht (siehe oranger Block in Abbildung 19) oder um das 1,25-fache gestreckt (siehe blauer Block in Abbildung 19). Es erfolgt eine Anpassung der Dichte, damit das Gewicht der Auflager bei jeder Variante dem Gewicht des Ausgangsmodells entspricht. Die Studie verwendet das 2D-Ausgangsmodell in starrer Konfiguration.

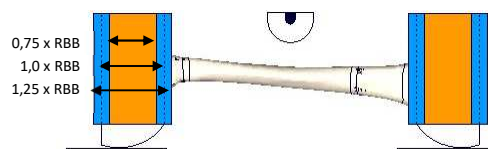


Abbildung 19 – Variation Auflagerblockbreite (Abkürzung: RBB)

Variation Drehpunkt

Die Breite des Auflagergehäuses des Ausgangsmodells (siehe grauer Block in Abbildung 20) wird für die Variantendefinition um das 0,75-fache gestaucht oder um das 1,25-fache gestreckt (siehe blauer Block in Abbildung 20).

Es erfolgt eine Dichteanpassung, damit das Auflagergewicht jeder Variante dem Auflagergewicht des Ausgangsmodells entspricht. Auf diese Weise werden die Abstände der Drehpunkte variiert. Die Studie verwendet das 2D-Ausgangsmodell in starrer Konfiguration.

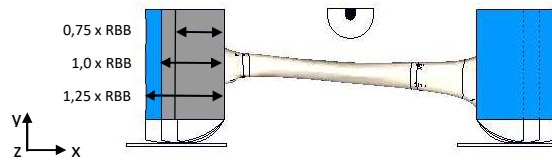


Abbildung 20 – Variation Drehpunkt

3.2.5 Sensitivitätsanalysen bezüglich Schaumdefinition

Schaumdichte

Die Schaumdichte wird in einem Bereich von $\rho = 25 \text{ kg/m}^3$ bis $\rho = 62 \text{ kg/m}^3$ variiert. Die Schaumdichte des numerischen Ausgangsmodells ist mit $\rho = 50 \text{ kg/m}^3$ definiert. Zudem wird die Wirkung des Schaums untersucht, indem eine Simulation ohne Schaum durchgeführt wird. Die Studie wird mit dem 2D-Ausgangsmodell in starrer Konfiguration durchgeführt.

Kompressionskurve

Die Kurve für die Definition der Kompressionscharakteristik des Schaums wurde von Park [85] veröffentlicht. Die Kompressionskurve wird in folgendem Bereich variiert: 0,9-fache Steifigkeit bis 1,2-fache Steifigkeit. Die Studie wird mit dem 2D-Ausgangsmodell in starrer Konfiguration gerechnet.

3.2.6 Sensitivitätsanalysen bezüglich Knochenpositionierung

Bei dieser Studie wird der Knochen in seiner Ausgangslage um den Knochenschwerpunkt in x-, y- und z-Richtung um jeweils 2° , -2° und -4° rotiert (siehe Abbildung 21). Die Studie verwendet das 2D-Ausgangsmodell in starrer Konfiguration.

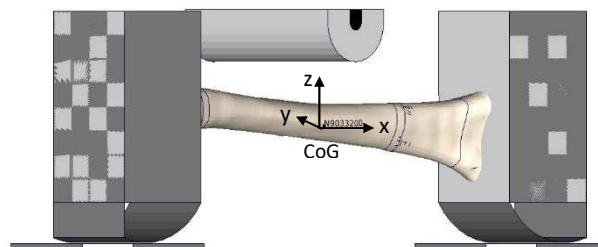


Abbildung 21 – Knochenschwerpunkt und Achsendefinition

3.2.7 Sensitivitätsanalysen bezüglich der Positionierung des Versuchsaufbaus

Variation der Impaktorpositionierung

Der Impaktor mit Schaum wird um den Schwerpunkt des Impaktors gedreht. Die Varianten sind wie folgt definiert: Drehung um 1° , 3° und 5° um die z-Achse. Diese Varianten werden mit der rechtwinkligen Ausgangslage verglichen (siehe Abbildung 22). Die Studie wird mit dem 2D-Ausgangsmodell in deformierbarer Konfiguration durchgeführt.

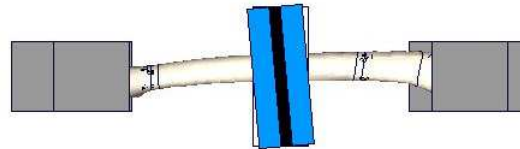


Abbildung 22 – Drehung des Impaktors um 5° um die z-Achse (blau) im Vergleich zur Ausgangslage (weiß)

Variation der Auflagerpositionierung

In diesem Schritt werden die Auflager in eine schräge Anfangsposition gebracht (siehe Abbildung 23). Die Auflager werden um ihren Schwerpunkt in jeweils y- (siehe Abbildung 23a und b) und z-Richtung (siehe Abbildung 23c und d) um 3° und 5° gedreht. Pro Konfiguration werden zwei Varianten untersucht. Bei einer Variante wird nur ein Auflager gedreht (siehe Abbildung 23a und c) und bei der anderen Variante werden beide Auflager gedreht (siehe Abbildung 23b und d). Die Studie verwendet das 2D-Ausgangsmodell in deformierbarer Konfiguration.

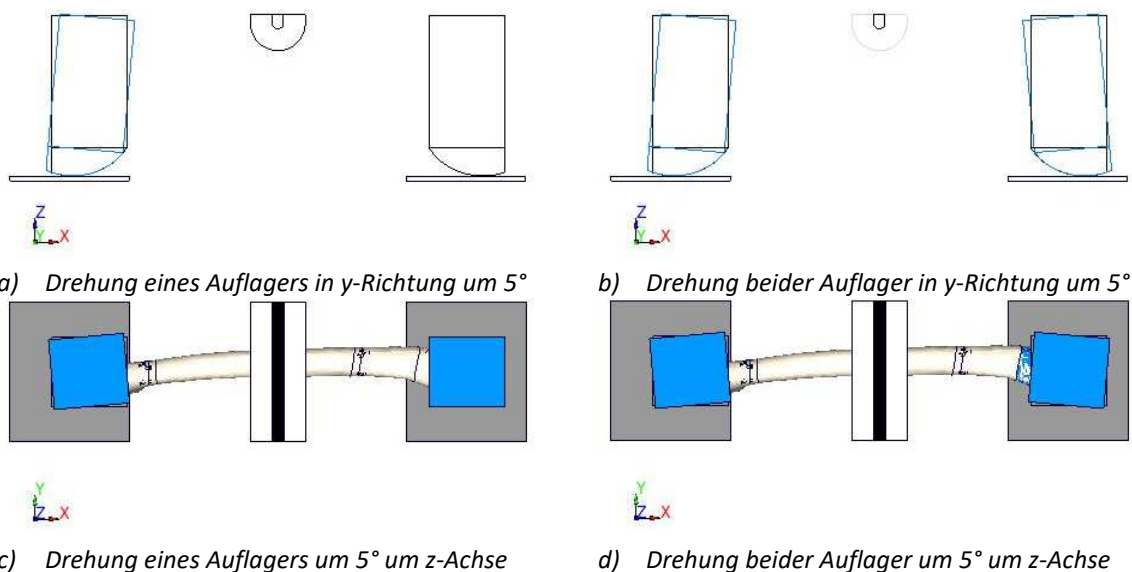


Abbildung 23 – Schräge Positionierung der Auflager (vgl. blaue Darstellung)

Variation Auflagerversatz

Bei dieser Studie wird ein Auflagerversatz durchgeführt. Ein Auflager wird um 1 cm und 3 cm (siehe Abbildung 24) in y-Richtung verschoben. Auf diese Weise werden die Schwerpunkte der Auflager in y-Richtung versetzt, damit die beiden Schwerpunkte in y-Richtung nicht mehr auf gleicher Höhe liegen. Die Studie wird mit dem 2D-Ausgangsmodell in deformierbarer Konfiguration durchgeführt.

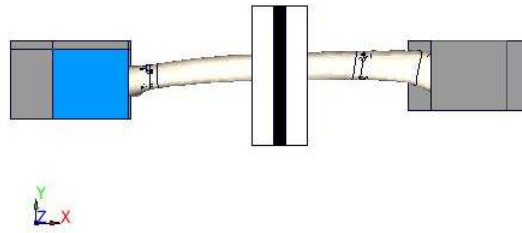


Abbildung 24 – Auflagerversatz in y-Richtung um 1 cm (blauer Auflager)

3.2.8 Einfluss unterschiedlicher Knochenmodelle auf den Simulationsaufbau

Das Femur-Modell des THUMS Version 4 besitzt eine durchgängige Kortikalis-Schicht mit 3D-Elementen variabler Dicke (siehe Abbildung 25). Das Innere des Femur-Modells ist mit 3D-Elementen gefüllt, dass die Spongiosa des Knochens darstellt. Die Geometrie des THUMS-Femurschaftes ist in der Frontalebene konvex von der Medianebene weg gewölbt (siehe Abbildung 25). Die Materialeigenschaften des THUMS-Femur-Modells sind in Tabelle 4 charakterisiert.

Tabelle 4 – Materialeigenschaften Femur THUMS Version 4 [97]

	E-Modul [MPa]	Querkontraktionszahl	Versagensspannung (MPa)
Kortikale Schicht	17 000	0,3	150
Spongiosa	145,6	0,45	30,6

Das GHBMC-Femur-Modell besitzt eine mit 3D-Elementen modellierte Kortikalis-Schicht im Bereich des Schaftes (siehe Abbildung 25). Das Innere des Femurschaftes ist nicht mit Spongiosa ausgefüllt. Das distale und proximale Ende des Femurs ist mit einer dünnen Kortikalis-Schicht mit 2D-Elementen modelliert. Jedes Element besitzt hierbei eine eigene Dickenzuweisung. Zudem ist dieser Bereich mit 3D-Elementen ausgefüllt, welche die Spongiosa modellieren. Die Materialeigenschaften der Kortikalis-Schicht sind in Tabelle 5 definiert. Im Vergleich zum Femur-Modell des THUMS besitzt die Geometrie des GHBMC-Femurschaftes eine konkave Wölbung in der Frontalebene hin zur Medianebene.

Tabelle 5 – Materialeigenschaften Femur GHBMC [34]

	E-Modul [MPa]	Querkontraktionszahl	Versagensspannung (MPa)	Versagensdehnung (-)
Kortikale Schicht	13 500	0,3	134	0,0218

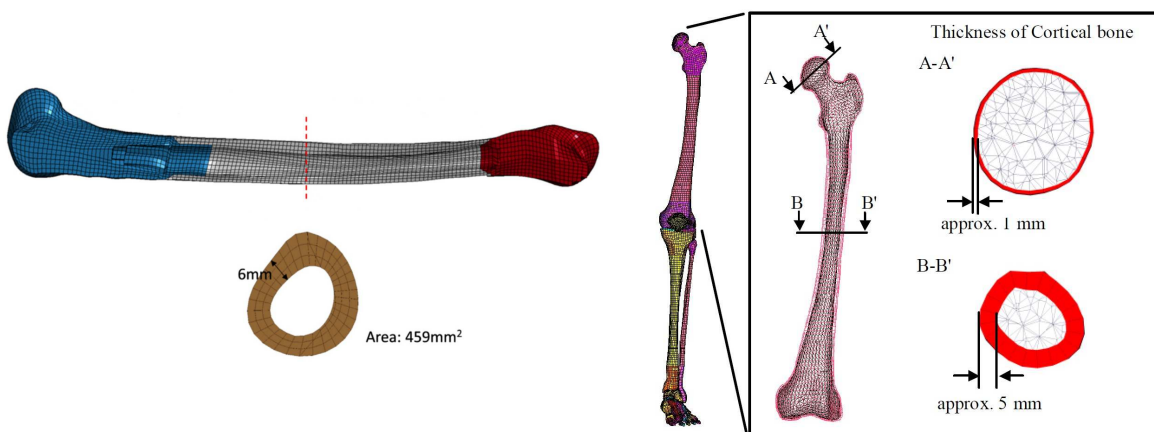


Abbildung 25 – Links: GHBMC-Femur-Modell [34]/ Rechts: THUMS Version 4 Femur-Modell [97]

Bei dieser Studie wird das 2D-Ausgangsmodell in deformierbarer Konfiguration mit dem THUMS Version 4 AM 50 Femur-Modell und mit dem GHBMCM Version 4 AM 50 Femur-Modell simuliert. Zudem wird ein modifiziertes Modell generiert, welches die Materialeigenschaften des GHBMCM-Femur-Modells (siehe Tabelle 5) und die Geometrie des THUMS-Femur-Modells besitzt. Die Modellvorhersage des modifizierten Modells ermöglicht einen Vergleich mit der Modellvorhersage des GHBMCM-Modells bezüglich der unterschiedlichen Femur Geometrien und Modellierungsarten, unabhängig von den Gewebeeigenschaften.

3.2.9 Vergleich des entwickelten Simulationsmodells mit der Realität

Die reale Geometrie der Auflager wurde von Wissenschaftlern der University of Virginia zur Verfügung gestellt. Die Modellierung der Auflager erfolgt in einer Ausführung mit 3D-Elementen (siehe Abbildung 26c) und in der Anderen mit 2D-Elementen (siehe Abbildung 26b). Die drei Modellierungsvarianten der Auflager, welche in Abbildung 26 dargestellt sind, werden jeweils in der starren und deformierbaren Konfiguration simuliert.

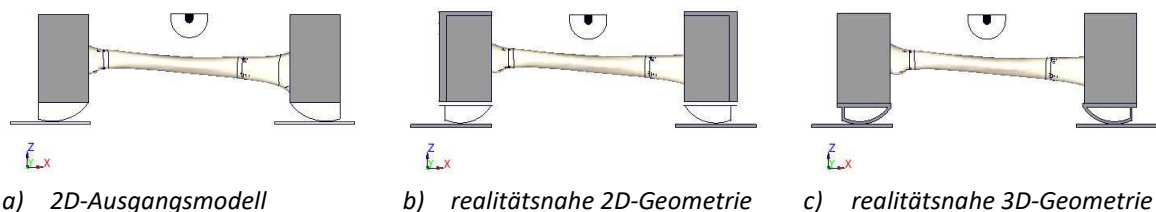


Abbildung 26 – Modellvergleiche

3.3 Ergebnisse

Die Ergebnisse der durchgeführten Studien werden mittels visueller Kurvenanalyse individuell auf ihre mechanische Relevanz und Plausibilität hin untersucht. Die Ergebnisse der Studien werden in zwei Kategorien eingestuft: signifikanter Einfluss des untersuchten Parameters oder vernachlässigbarer Einfluss. Mit dieser Kategorisierung wird das Ziel verfolgt, wichtige Parameter zu identifizieren, welche die Modellvorhersage beeinflussen. Als signifikante Parameter werden diejenigen Parameter eingestuft, welche den Kurvenverlauf der Modellvorhersage in einem realistischen Parameterbereich beeinflussen.

3.3.1 Sensitivitätsanalysen bezüglich Modellierungsannahmen

Die Ergebnisse der Studien bezüglich Modellierungsannahmen sind in Tabelle 6 und Anhang A2.1 zusammengefasst.

Es gibt einen signifikanten Unterschied im Kurvenverlauf der Modellvorhersage zwischen der starren und der deformierbaren Konfiguration des Versuchsaufbaus. Die Modellvorhersage des Versuchsaufbaus in der deformierbaren Konfiguration besitzt das in Abschnitt 3.1.3 beschriebene Schwingungsphänomen mit ausgeprägtem erstem lokalem Maximum. Im weiteren Kurvenverlauf treten fortlaufend niederfrequente Schwingungen auf. Die Modellvorhersage des Versuchsaufbaus in der starren Konfiguration ist weniger schwingungsbehaftet. Das erste signifikante lokale Maximum ist nicht vorhanden. Es tritt lediglich eine Änderung des Kurvengradienten zu dem Zeitpunkt ein (nach ca. 22 ms), bei dem das Auflager anfängt sich abzurollen.

Tabelle 6 – Ergebnisse der Sensitivitätsanalyse bezüglich der Modellierungsannahmen

Studie	Signifikanter Einfluss	Vernachlässigbarer Einfluss	Ergebnis
Starrer vs. deformierbarer Versuchsaufbau	x		Abbildung A2.1.1a
Variation Reibungskoeffizient Auflager – Bodenplatte		x	Abbildung A2.1.1b
Schnittstellenmodellierung Impaktor – Schaum		x	Abbildung A2.1.1c
Modellierung der Knocheneinbettung		x	Abbildung A2.1.1d
Modellierung des Impaktors		x	Abbildung A2.1.1e

Der Einfluss des Reibungskoeffizienten zwischen Auflager und Bodenplatte, welcher bei dem Ausgangsmodell mit $\mu = 0,2$ festgelegt ist, ist zwischen $\mu = 0,0$ und $\mu = 0,2$ gering. Ab einem größeren Reibwert von $\mu = 0,5$ ist der Unterschied signifikant, was eine Änderung des Kurvenverlaufs hervorruft. Die unterschiedlichen Modellierungsarten der Schaumanbindung und die unterschiedlichen Modellierungsarten der Knocheneinbettung haben einen geringen Einfluss auf die Modellvorhersage. Bei diesen Studien divergieren die Kurvenverläufe ab einem Kontaktkraftniveau von ca. 4,5 kN geringfügig.

Die Annahme der Starrkörperdefinition des Impaktors hat keinen Einfluss auf die Modellvorhersage. Die Modellvorhersagen können als identisch betrachtet werden.

3.3.2 Sensitivitätsanalysen zur Festlegung der geometrischen Parameter

Die Ergebnisse der Sensitivitätsanalysen zur Festlegung der geometrischen Parameter des Versuchsaufbaus sind zusammenfassend in Tabelle 7 und Anhang A2.2 dargestellt. Die Variation der Breite der Auflager und damit der freien Schaftlänge des Knochens, sowie die Verschiebung des Schwerpunktes in z-Richtung sind Parameter, welche einen vernachlässigbaren Einfluss auf die Modellvorhersage ausüben.

Tabelle 7 – Ergebnisse der Sensitivitätsanalysen der geometrischen Parameter der Auflager

Studie	Signifikanter Einfluss	Vernachlässigbarer Einfluss	Ergebnis
Variation der Struktursteifigkeit	x		Abbildung A2.2.1a
Variation des Auflagergewichts	x		Abbildung A2.2.1b
Variation der Auflagerfläche	x		Abbildung A2.2.1c
Verschiebung des Schwerpunktes in x-Richtung	x		Abbildung A2.2.1d
Verschiebung des Schwerpunktes in z-Richtung		x	Abbildung A2.2.1e
Variation des Radius der Auflager	x		Abbildung A2.2.1f
Variation der Breite der Auflager		x	Abbildung A2.2.1g
Variation des Drehpunktes in x-Richtung	x		Abbildung A2.2.1h

Bei der Variation der Struktursteifigkeit der Auflager nähert sich mit zunehmender Materialstärke der Kurvenverlauf der Modellvorhersage dem Kurvenverlauf der Modellvorhersage mit starrer Konfiguration an. Eine Zunahme des Auflagergewichts induziert verstärkt Schwingungen mit großer

Amplitude. Ein ähnliches Verhalten wird durch eine Zunahme der Kontaktfläche zwischen Bodenplatte und Auflager bei gleichem Auflagergewicht erzeugt. Durch die vergrößerte Kontaktfläche tritt der Effekt der Reibkraft verstärkt auf. Dies bewirkt Schwingungen mit niedrigerer Amplitude als bei der Variation des Auflagergewichts. Die Frequenz der Schwingungen nimmt mit zunehmender Kontaktfläche zu.

Eine Verlagerung des Auflagerschwerpunktes in x-Richtung hat einen Einfluss auf die Modellvorhersage. Die Kurvenformen der Modellvorhersagen der Varianten sind ähnlich der Kurvenform der Modellvorhersage des Ausgangsmodells.

Variationen des Auflagerradius und des Auflagerpunktes der Auflager haben maßgeblich Einfluss auf das Abrollverhalten zu Beginn der Auflagerbewegung nach ca. 21 ms. Zudem bewirkt ein größerer Auflagerradius oder eine Verringerung der freien Stützweite eine Reduktion von Schwingungen im Modell.

3.3.3 Sensitivitätsanalysen bezüglich der Schaumdefinition

Die Ergebnisse bezüglich der Schaumcharakterisierung sind in Tabelle 8 und Anhang A2.3 dargestellt. Die untersuchten Parameter haben keinen signifikanten Einfluss auf die Modellvorhersage.

Tabelle 8 – Ergebnis Sensitivitätsanalysen bezüglich Schaumdefinition

Studie	Signifikanter Einfluss	Vernachlässigbarer Einfluss	Ergebnis
Variation Schaumdichte		x	Abbildung A2.3.1a
Variation Schaumsteifigkeit		x	Abbildung A2.3.1b

Die Modellvorhersage ändert sich signifikant, wenn der Schaum weggelassen wird und der Impaktor direkt auf den Knochen trifft (siehe Abbildung A2.3.1a). Die Modellvorhersage ohne Schaum weist mehr Schwingungen im Kurvenverlauf auf, vor allem nach ca. 22 ms beim Anfang der Abrollbewegung.

3.3.4 Sensitivitätsanalysen bezüglich der Knochenpositionierung

Die Ergebnisse der Studien zur Untersuchung des Einflusses der Knochenpositionierung auf die Modellvorhersage sind in Tabelle 9 und Anhang A2.4 dargestellt.

Tabelle 9 – Ergebnis Sensitivitätsanalysen hinsichtlich Knochenpositionierung

Studie	Signifikanter Einfluss	Vernachlässigbarer Einfluss	Ergebnis
Rotation des Knochens um x-Achse		x	Abbildung A2.4.1a
Rotation des Knochens um y-Achse		x	Abbildung A2.4.1b
Rotation des Knochens um z-Achse		x	Abbildung A2.4.1c

Eine Knochenrotation um die z-Achse wirkt sich nicht auf die Modellvorhersage aus. Knochenrotationen bezüglich der x- und y-Achse wirken sich leicht auf die Modellvorhersage aus. Hierbei wird der Kurvenverlauf geringfügig bezüglich der Zeitachse verschoben. Diese Einwirkung ist jedoch vernachlässigbar.

3.3.5 Sensitivitätsanalysen bezüglich der Positionierung des Versuchsaufbaus

Die Ergebnisse der Sensitivitätsanalysen hinsichtlich der Positionierung des Versuchsaufbaus sind in Tabelle 10 und Anhang A2.5 zusammenfassend dargestellt.

Eine z-Rotation des Impaktors oder ein Versatz in y-Richtung während der Auflagerpositionierung wirken sich nicht auf die Modellvorhersage aus. Wenn jedoch ein bzw. beide Auflager in eine schräge Ausgangslage positioniert werden, d.h. um die x- oder z-Achse gedreht werden, ist ein signifikanter Einfluss auf die Modellvorhersage vorhanden. Der Effekt auf die Modellvorhersage bei beiden Studien ist verstärkt, wenn beide Auflager gedreht werden. Eine Drehung um die y-Achse wirkt sich stärker auf die Kurvenform der Modellvorhersage aus, als eine Drehung um die z-Achse.

Tabelle 10 – Ergebnisse der Sensitivitätsanalysen bezüglich der Positionierung des Versuchsaufbaus

Studie	Signifikanter Einfluss	Vernachlässigbarer Einfluss	Ergebnis
Z-Rotation des Impaktors		x	Abbildung A2.5.1a
Z-Rotation eines Auflagers um seinen Schwerpunkt	x		Abbildung A2.5.1b
Z-Rotation beider Auflager um jeweils ihren Schwerpunkt	x		Abbildung A2.5.1c
Y-Rotation eines Auflagers um seinen Schwerpunkt	x		Abbildung A2.5.1d
Y-Rotation beider Auflager um jeweils ihren Schwerpunkt	x		Abbildung A2.5.1e
Auflagerversatz in y-Richtung		x	Abbildung A2.5.1f

3.3.6 Einfluss unterschiedlicher Knochenmodelle auf die Modellvorhersage

In Abbildung A2.6.1a (siehe Anhang A2.6) sind die Modellvorhersagen unterschiedlicher Knochenmodelle mit dem Versuchsaufbau in deformierbarer Konfiguration dargestellt. Der Kurvenverlauf des THUMS-Femur-Modells hat einen geringeren Kontaktkraft-Anstieg als das GHBMC-Femur-Modell. Charakteristisch für die GHBMC-Modellvorhersage sind niederfrequente Schwingungen mit ungleichmäßiger Amplitude. Die Modellvorhersage des THUMS-Femur-Modells zeichnet sich durch gleichmäßige Schwingungen höherer Frequenz und niedriger Amplitude aus.

In Abbildung A2.6.1b sind die Ergebnisse der unterschiedlichen Geometrien der Knochen mit einer Materialanalogie und dem Versuchsaufbau in starrer Konfiguration dargestellt. Aufgrund der Verwendung des starren Modellaufbaus weisen die Ergebnisse geringere Schwingungen auf. Auch mit Materialanalogie zwischen den beiden Knochenmodellen weist das THUMS-Knochenmodell einen geringeren Kontaktkraft-Anstieg als das GHBMC-Femur-Modell auf.

3.3.7 Vergleich des entwickelten Simulationsmodells mit der Realität

In Abbildung A2.7.1a (siehe Anhang A2.7) sind die Modellvorhersagen mit verschiedenen Modellierungsarten für die Auflager in der starren Konfiguration dargestellt. In Abbildung A2.7.1b sind die Ergebnisse unterschiedlicher Modellierungsarten für die Auflager in der deformierbaren Konfiguration visualisiert. In Abbildung A2.7.1c werden die Modellvorhersagen der 3D-Modellierung der Auflager in starrer Konfiguration und des 2D-Ausgangsmodells (siehe Abschnitt 3.2.1) in deformierbarer Konfiguration mit den experimentellen Kurven verglichen.

Bei einer starren Konfiguration der Auflager weist das 2D-Ausgangsmodell (siehe Abbildung A2.7.1a) nicht das charakteristische lokale Maximum der experimentellen Kurven auf (siehe Abbildung A2.7.1c). Mit deformierbarer Konfiguration der Auflager stimmt der Kurvenverlauf der Modellvorhersage des 2D-Ausgangsmodells (siehe Abbildung A2.7.1b) gut mit den experimentellen Kurvenverläufen überein (siehe Abbildung A2.7.1c). Auch die Modellvorhersage der realitätsnahen 3D-Modellierung der Auflager in starrer Konfiguration (siehe Abbildung A2.7.1a) stimmt im Kurvenverlauf sehr gut mit den

experimentellen Ergebnissen überein (siehe Abbildung A2.7.1c), wohingegen diese Modellierung in deformierbarer Konfiguration gleichmäßige Schwingungen mit größerer Amplitude besitzt (siehe Abbildung A2.7.1b) und nicht mehr mit den Kurvenverläufen der experimentellen Ergebnisse übereinstimmt. Ein Vergleich der Modellvorhersagen der realitätsnahen 2D- und 3D-Modellierung der Auflager in starrer Konfiguration weist Unterschiede im Kurvenverlauf auf. Der Kurvenverlauf mit der realitätsnahen 2D-Modellierung weist deutlich mehr Schwingungen mit unterschiedlichen Amplituden auf. Die Unterschiede in den Kurvenverläufen der realitätsnahen 2D-Modellierung der Auflager zwischen der starren (siehe Abbildung A2.7.1a) und deformierbaren (siehe Abbildung A2.7.1b) Konfiguration sind vernachlässigbar.

3.4 Diskussion

3.4.1 Auswahlkriterien des Referenzexperiments

In diesem Abschnitt werden die Auswahlkriterien für das Referenzexperiment, dem dynamischen 3-Punkt-Biegeversuch des Femurs von Forman et al. [38], diskutiert.

Es wird ein Experiment auf lokaler Ebene gesucht, da eine möglichst isolierte Betrachtung von Effekten auf lokaler Ebene möglich ist. Aufgrund der experimentellen Randbedingungen sind die Freiheitsgrade meist eingeschränkt (z.B. Biegeversuch auf lokaler Ebene von Forman et al. [38] vs. Ganzkörper Schlittentests auf globaler Ebene von Shaw et al. [95]) und es tritt keine Überlagerung von Effekten, resultierend durch Interaktion einzelner Komponenten, auf (z.B. Forman et al. [38] vs. dynamische Beanspruchung des Knie zur Untersuchung der Hüftbelastbarkeit mit einem ganzen Bein auf regionaler Ebene von Rupp et al. [91]). Außerdem wird eine isolierte Probe bevorzugt (z.B. Forman et al. [38]) gegenüber einer Probe mit Weichteilmantel (z.B. dynamischer Biegeversuch des Femurs mit Weichteilmantel von Kerrigan et al. [65]) zur Vermeidung der potentiellen Überlagerung von Effekten.

Ein weiteres Kriterium ist die Auswahl eines dynamischen Lastfalls. Der Fokus liegt auf der Auswahl eines Fahrzeugcrash-relevanten Lastfalls, welcher eine ähnliche Belastung der Testproben während des Experiments bewirkt. Alle quasi-statischen Tests werden deshalb ausgeschlossen.

Bei der Auswahl wird die Qualität der wissenschaftlichen Veröffentlichung hinsichtlich ihres Informationsgehaltes bezüglich der Testproben, sowie der detaillierten Beschreibung des experimentellen Versuchsaufbaus, der experimentellen Randbedingungen, der Versuchsdurchführung und der Datenaufbereitung beurteilt. Der experimentelle Versuchsaufbau sollte möglichst geringe Unsicherheiten beinhalten, wie z.B. keinen nicht-charakterisierten Schaum oder fehlende Angaben bezüglich der Geometrie des Versuchsaufbaus, um die Unsicherheiten in der numerischen Replikation so gering wie möglich zu halten.

Zudem muss die wissenschaftliche Veröffentlichung die Ergebniskurven aller Testproben enthalten. Nur die Veröffentlichung des Versagenspunktes einer Testprobe ist nicht ausreichend, da die Ergebniskurven in Kapitel 5 für die Entwicklung und Anwendung der populationsbasierten Validierungsmethodik benötigt werden.

Das Referenzexperiment soll eine möglichst hohe Anzahl an Testproben aufweisen, da die Anwendung einer populationsbasierten Validierungsmethodik erst ab einer gewissen Anzahl an Testergebnissen sinnvoll ist.

Experimente mit einer Mehrfachverwendung von Testproben, wie z.B. in der Studie zur Untersuchung der Belastbarkeit des Thorax durch verschiedene Sitzgurtarten von Kent et al. [64], werden nicht berücksichtigt. Bei dieser Art von Studien kann nicht ausgeschlossen werden, dass unerkannte Schäden

der Testproben durch vorangegangene Versuche das Testergebnis bei späteren Versuchen beeinflussen.

Das Alter der Testproben soll im Erwachsenen-Alter zwischen 19 und 64 Jahren liegen, welches von der Weltgesundheitsorganisation [83] definiert wurde. Dieses Erwachsenen-Alter wurde bewusst gewählt, da Validierungsexperimente meist Testproben sehr hohen Alters aufweisen. Der Fokus dieser Untersuchung wird bewusst auf diese Altersgruppe gesetzt, da das Alter vieler Autofahrer in diese Altersgruppe fällt.

Ein weiteres wichtiges Kriterium für die Auswahl des Referenzexperiments ist eine hohe Datenverfügbarkeit bezüglich Geometrievarianz und Materialcharakterisierung der verwendeten Probe. Dieses Kriterium ist essentiell, da für dieses Referenzexperiment eine virtuelle Population bezüglich der verwendeten Probe in Kapitel 5 generiert wird. Diese virtuelle Population ist die Voraussetzung für die Entwicklung und Anwendung der populationsbasierten Validierungsmethodik in Kapitel 5.

3.4.2 Definition und Durchführung der Sensitivitätsanalysen

Ziel der Sensitivitätsanalysen ist die Überprüfung, ob das entwickelte numerische 2D-Ausgangsmodell die Mechanik der experimentellen Versuchsdurchführung korrekt abbildet. Bei der Definition der Sensitivitätsanalysen (siehe Abbildung 15) wird daher kein Anspruch auf eine umfassende Modellvalidierung erhoben. Für den Anspruch einer umfassenden Modellvalidierung und Verifizierung fehlen Sensitivitätsanalysen folgender Parameter:

- Vergleich unterschiedlicher Simulationssoftware: z.B. VPS vs. LS-Dyna,
- Unterschiedliche Versionen einer Software: z.B. VPS Version 2015 vs. VPS Version 2019,
- Numerische Parameter,
- Netzkonvergenzstudien.

Es erfolgt keine Überprüfung der Implementierung des Algorithmus, da die numerischen Simulationen mit VPS Version 2015, einer kommerziellen Software, durchgeführt werden. Es wird bei der Nutzung einer kommerziellen Software davon ausgegangen, dass die Lösungsalgorithmen adäquat implementiert und überprüft wurden.

Zudem findet keine quantitative Validierung der Modellvorhersagen statt, da das virtuelle Knochenmodell ein Individuum repräsentiert, welches eine andere Geometrie und ein anderes Materialverhalten als die experimentellen Knochen besitzt. Daher kann keine subjekt-spezifische Validierung erfolgen (siehe Abschnitt 1.2.5).

Die meisten Sensitivitätsanalysen werden mit dem 2D-Ausgangsmodell durchgeführt. Wenige Studien werden mit der realitätsnahen 3D-Modellierung der Auflager simuliert. Beide Modelle weisen eine ausreichende geometrische Ähnlichkeit zu dem experimentellen Versuchsaufbau und zueinander auf (siehe Abbildung 26). Zudem sind die Kurvenverläufe der Modellvorhersagen beider Modelle ähnlich zu den experimentellen Modellvorhersagen. Daher ist die Verwendung beider Modelle für die Sensitivitätsanalysen gerechtfertigt.

Des Weiteren wird pro Sensitivitätsanalyse nur ein Parameter isoliert variiert und immer mit der Modellvorhersage der Ausgangskonfiguration verglichen. Diese isolierte Effektbetrachtung ermöglicht die Bewertung in signifikante und vernachlässigbare Parameter bei dieser Art von Versuchsaufbau. Die Bewertung der Sensitivitätsstudien in signifikante Parameter und vernachlässigbare Parameter ist für den Validierungszweck angemessen.

Für eine detaillierte Modellvalidierung müsste jedoch eine genauere Bewertung der Unterschiede zwischen Modellvorhersagen erfolgen. Vor allem müssten Kurvenvergleiche bei einer detaillierten Modellvalidierung auf Basis möglichst objektiver Kriterien erfolgen.

3.4.3 Modellierung des Versuchsaufbaus

In diesem Abschnitt werden die durchgeführten Analysen bezüglich der Modellierungsannahmen diskutiert, welche bei der Erstellung des numerischen Versuchsaufbaus getroffen werden. Diese Analysen dienen zur Abschätzung des Modellierungsfehlers in Bezug auf die Randbedingungen des entwickelten Versuchsaufbaus (siehe Abbildung 15).

Der signifikante Unterschied im Kurvenverlauf zwischen dem starren- und dem deformierbaren 2D-Ausgangsmodell kommt aufgrund einer Deformation der Auflager während der Abrollbewegung zustande. Dies deutet darauf hin, dass die gewählte Materialstärke des Stahls mit 2 mm bei dem Ausgangsmodell zu einer zu geringen Auflagersteifigkeit führt. Die experimentellen Auflager werden für die Versuchsdurchführung aller Experimente verwendet. Eine ausreichend hohe Auflagersteifigkeit ist deshalb wichtig, damit sich die Auflager während der Versuchsdurchführung nicht deformieren, so dass ungewollte Einflüsse auf den Versuch vermieden werden und die Auflager mehrfach verwendet werden können. Dieser Aspekt der Auflagersteifigkeit wird daher separat in Abschnitt 3.2.4 untersucht und in Abschnitt 3.4.4 und Abschnitt 3.4.9 erneut diskutiert.

Forman et al. [38] beschrieben, dass die Kontaktflächen zwischen den Auflagern und der Bodenplatte geschmiert wurde. Diese Schmierung wird durchgeführt, um die Reibung so weit wie möglich zu reduzieren. In VPS Version 2015 gibt es nur die Möglichkeit, einen einzigen Reibungskoeffizienten zu definieren. Eine separate Definition für Haft- und Gleitreibung ist nicht möglich. Der Reibungskoeffizient des Ausgangsmodells wird mit $\mu = 0,2$ definiert. Bis zu diesem Wert sind die Kurvenverläufe sehr ähnlich. Bei einem größeren Reibungskoeffizienten ändert sich der Kurvenverlauf signifikant. Der Grund dafür ist, dass sich bei größerem Reibungskoeffizienten die Rollbewegung zu einer ruckartigen Rutschbewegung entwickelt. In [39] sind verschiedene Reibungskoeffizienten der betrachteten Oberflächen gelistet (siehe Tabelle 11). Der festgelegte Reibungskoeffizient von $\mu = 0,2$ liegt im Bereich der Haftreibung ohne Schmierung. Für eine geschmierte Oberfläche und ein realistischeres Abbild des Versuchsaufbaus wäre ein geringerer Reibungskoeffizient zwischen $\mu = 0,04$ bis $\mu = 0,12$ empfehlenswerter. Der Modellierungsfehler des Ausgangsmodells mit einem Reibungskoeffizient von $\mu = 0,2$ ist aufgrund des vernachlässigbaren Unterschiedes im Kurvenverlauf (siehe Abbildung A2.1.1b) vertretbar. Der Parameter Reibungskoeffizient wird in Abschnitt 3.3.1 als vernachlässigbar eingestuft, da laut Tabelle 11 nur eine Parametervariation bis zu einem Reibungskoeffizienten $\mu = 0,3$ realistisch ist. Deshalb werden die Varianten mit $\mu > 0,2$ aus der Ergebnisbewertung ausgeschlossen.

Tabelle 11 – Reibwertkoeffizienten entnommen aus [39]

Werkstoff	Reibungsart	Schmierung	Reibwert
Stahl – Stahl	Haftreibung	Trocken	0,15–0,30
Stahl – Stahl	Haftreibung	Geschmiert	0,10–0,12
Stahl – Stahl	Gleitreibung	Trocken	0,10–0,12
Stahl – Stahl	Gleitreibung	Geschmiert	0,04–0,07

Bei dem experimentellen Versuchsaufbau von Forman et al. [38] wird der Schaum mit einer Schnur an den Impaktor angebunden (siehe Tabelle 2). Zudem sind die Bauteile Schaum und Impaktor hinsichtlich ihres Materialverhaltens und ihrer Steifigkeit unterschiedlich. Die Schnittstelle zwischen Schaum und Impaktor wird deshalb bei dem 2D-Ausgangsmodell mit Kontakten modelliert. Diese Modellierungsart stellt ein besseres Abbild der Realität dar, als eine idealisierte Anbindung mit geteilten Knoten. Aufgrund des vernachlässigbaren Unterschiedes zwischen beiden Modellierungsarten (siehe Abbildung A2.1.1c) sind beide Modellierungsarten gerechtfertigt.

Hinsichtlich der Einbettung des Knochens in die PMMA-Einbettungsblöcke werden zwei Aspekte untersucht. Die Starrkörperdefinition der Auflagergeometrie wird bei dem starren 2D-Ausgangsmodell verwendet, um das Knochenende in die Einbettungsblöcke einzubinden. Diese vereinfachte Annahme beherbergt zwei mögliche Fehlerquellen. Zum einen besitzt das Auflager eine höhere Masse und damit Massenträgheit als in der Realität. Denn das Volumen des PMMA-Einbettungsblocks, welches normalerweise von dem Knochenende verdrängt wird, ist mit dem PMMA-Material ausgefüllt. Zum anderen kann durch die Starrkörperdefinition das PMMA-Material keine Energie aufnehmen. Deshalb werden zwei Simulationen durchgeführt, um eine realistischere Knocheneinbettung mit dem starren 3D-Ausgangsmodell zu vergleichen. Bei einer Simulation werden die Knochenenden realistischer, mit verdrängendem Volumen, in den Einbettungsblöcken gelagert. Der Einbettungsblock ist hierbei deformierbar. Diese Simulation dient zur Untersuchung der Energieaufnahme der deformierbaren Einbettungsblöcke. Bei der anderen Simulation werden die Knochenenden realistischer, mit verdrängendem Volumen, in den Einbettungsblöcken gelagert. Der Einbettungsblock wird jedoch als Starrkörper simuliert, um den Effekt des Massenunterschiedes und der daraus resultierenden Massenträgheit zwischen dem starren 3D-Ausgangsmodell und der realistischeren 3D-Starrkörperdefinition zu untersuchen. Der verursachte Fehler durch die vereinfachende Annahme des 2D-Ausgangsmodells, der Starrkörpermodellierung ohne boolesche Operation, kann als vernachlässigbar eingestuft werden (siehe Abbildung A2.1.1d).

Der Impaktor aus Aluminium besitzt eine wesentlich höhere Bauteilfestigkeit als Knochengewebe. Aus diesem Grund kommt es während der Versuchsdurchführung zu keiner merklichen Deformation des Aluminiums. Die Starrkörperdefinition für den Impaktor ist daher gerechtfertigt (siehe Abbildung A2.1.1e).

Zusammenfassend muss das Ausgangsmodell hinsichtlich seiner Struktursteifigkeit noch genauer untersucht werden. Alle anderen vereinfachenden Annahmen zur Modellbildung sind gerechtfertigt, da der resultierende Modellierungsfehler hinsichtlich der Randbedingungen gering ist.

3.4.4 Sensitivitätsanalysen zur Festlegung der geometrischen Parameter

Die geometrische Definition der Auflager ist in den wissenschaftlichen Veröffentlichungen zur Studie nicht enthalten. Es werden daher Sensitivitätsanalysen hinsichtlich geometrischer Parameter durchgeführt, um die Parameterauswahl des 2D-Ausgangsmodells zu überprüfen.

Bei der Variation der Struktursteifigkeit der Auflager wird die Dichte des Auflagergehäuses für jede Variante angepasst, um eine konstante Massenträgheit über alle Varianten zu gewährleisten. Auf diese Weise ist es möglich, den Effekt der Struktursteifigkeit isoliert zu betrachten.

Die Modellvorhersagen mit einer Materialstärke von 5 mm und 7 mm sind nahezu identisch (siehe Abbildung A2.2.1a). Es kann die Schlussfolgerung gezogen werden, dass ab einer Materialstärke von 5 mm die Auflager steif genug ausgelegt sind, damit sie während des Abrollvorgangs nicht mehr deformiert werden. Das realitätsnahe 3D-Modell besitzt eine Materialstärke von ca. 4,6 mm was die Schlussfolgerung unterstützt. Die Materialstärke des 2D-Ausgangsmodells wird auf 2 mm festgelegt, da der Kurvenverlauf dieser Modellvorhersage visuell gut mit den experimentellen Kurvenverläufen übereinstimmt (siehe Abbildung A2.7.1c). Dem entgegengesetzt ist der Kurvenverlauf der Modellvorhersage des 2D-Ausgangsmodells mit starrer Konfiguration sehr unterschiedlich zu den experimentellen Kurven. Diese Beobachtung legt die Vermutung nahe, dass sich die erste Hypothese in Abschnitt 3.1.3, dass die zu geringe Materialsteifigkeit für das erste signifikante lokale Maximum des experimentellen Ergebnisses verantwortlich ist, als wahr erweist.

Bei der Variation der Auflagertiefe werden zwei Aspekte untersucht. Bei einer Studie wird durch die Auflagertiefenvariation eine Variation des Auflagergewichts angestrebt (siehe Abbildung A2.2.1b). Bei

der anderen Studie wird die Auflagertiefe, ohne Gewichtsveränderung der Auflager, variiert. Diese Studie hat das Ziel, die Auflagerfläche der Auflager zu variieren und den Einfluss des Reibungseffektes zu untersuchen (siehe Abbildung A2.2.1c). Beide Parameter beeinflussen den Kurvenverlauf signifikant. Eine Erhöhung beider Parameter wirkt sich auf das Schwingungsverhalten im Kurvenverlauf aus. Für beide Parameter, das Auflagergewicht und die Auflagertiefe bzw. Auflagerfläche, erzeugen die festgelegten Werte für das 2D-Ausgangsmodell einen Kurvenverlauf der Modellvorhersage ähnlich zu den experimentellen Kurvenverläufen (siehe Abbildung A2.7.1c). Die Kurvenverläufe der Modellvorhersagen der einzelnen Varianten dieser Studien unterscheiden sich stark von den experimentellen Kurvenverläufen.

Mit der 3D-Modellierung der Auflager in starrer Konfiguration wird eine Studie zur Verlagerung des Schwerpunktes durchgeführt. Es wird untersucht, wie sich die Massenverteilung der Auflager auf die Modellvorhersage auswirkt. Die Ergebnisse zeigen, dass vor allem die Lage des Schwerpunktes in x-Richtung Einfluss auf die Modellvorhersage hat (siehe Abbildung A2.2.1d). Die Art, wie die Verlagerung des Schwerpunktes numerisch realisiert wird (siehe Abbildung 17) ist nicht realistisch, der Bereich der Verlagerung jedoch schon. Eine derartige Verlagerung des Schwerpunktes könnte in Realität erfolgen, indem diejenige Seite des Auflagergehäuses eine höhere Materialstärke zugewiesen bekommt, in dessen Richtung der Schwerpunkt verschoben werden soll. Vorausgesetzt das Auflagergehäuse ist mit einer konstanten Materialstärke modelliert, wird die Materialverteilung des 2D-Ausgangsmodells als realistisch angenommen und bestätigt.

Die Variation des Auflagerradius und die Variation des Drehpunktes, d.h. der Distanz zwischen den beiden Drehpunkten der Auflager, beeinflussen die Kurvenform der Modellvorhersage. Da jedoch diese zwei Parameter in dem 2D-Ausgangsmodell nur sehr leicht von dem experimentellen Versuchsaufbau abweichen (siehe Abbildung 13), kann die Parameterwahl des Ausgangsmodells als akzeptabel eingestuft werden.

Die Variation der Auflagerblockbreite, und damit der freien Schaftlänge des Femur Knochens, kann als vernachlässigbar angesehen werden.

Zusammenfassend wird bei der Planung eines experimentellen 3-Punkt-Biegeversuches empfohlen, die Auflager steif genug auszulegen und die Massenträgheit nicht zu hoch zu wählen, damit die experimentellen Ergebnisse nicht durch den Versuchsaufbau zusätzlich beeinflusst werden. Zudem muss die Gewichtsverteilung innerhalb eines Auflagers, genauer gesagt die Lage des Auflagerschwerpunktes in x-Richtung, und der gewählte Radius der Auflager bei der Planung eines Experimentes mitberücksichtigt werden. Zusätzlich können die Variation der Auflagerfläche, aufgrund der Reibkraftveränderung, und der Abstand der Drehpunkte der Auflager in x-Richtung das experimentelle Ergebnis signifikant beeinflussen.

3.4.5 Sensitivitätsanalysen bezüglich der Schaumdefinition

Die Dichte des Schaums wurde nicht veröffentlicht und daher in einem Bereich zwischen $\rho = 25 \text{ kg/m}^3$ bis $\rho = 62 \text{ kg/m}^3$ variiert. Weichschäume besitzen eine Dichte zwischen $\rho = 20 \text{ kg/m}^3$ bis $\rho = 40 \text{ kg/m}^3$ und Kalt-Schäume besitzen eine Dichte zwischen $\rho = 30 \text{ kg/m}^3$ bis $\rho = 50 \text{ kg/m}^3$ [54]. Die Dichte von Schäumen in Autositzen variiert zwischen $\rho = 55 \text{ kg/m}^3$ bis $\rho = 75 \text{ kg/m}^3$.

Der Variationsbereich dieses Parameters kann aufgrund der Referenzwerte als realistisch eingestuft werden. Der Parameter wirkt sich nicht auf die Modellvorhersage aus. Daher ist eine Definition der Schaumdichte mit $\rho = 50 \text{ kg/m}^3$ gerechtfertigt.

Die Kompressionskurve des Schaums wurde in Park [85] veröffentlicht. Sie wird leicht in ihrer Charakteristik im Bereich der 0,9-fachen bis 1,1-fachen Steifigkeit variiert, um die Sensitivität der Modellvorhersage bezüglich der Kompressionscharakteristik abzuschätzen. Die Ergebnisse zeigen,

dass eine leichte Abweichung der Schaumsteifigkeit keinen Einfluss auf die Modellvorhersage hat. Dies bedeutet, dass für diese Art von Versuchsaufbau auch ein anderer Schaum mit geringfügig abweichender Charakteristik verwendet werden kann, ohne die Modellvorhersage signifikant zu beeinflussen. Es ist jedoch davon auszugehen, dass bei einer stärkeren Änderung der Steifigkeit oder der Kurvenform der Kompressionscharakteristik die Modellvorhersage beeinflusst wird. Ein Indiz dafür ist die signifikante Beeinflussung der Modellvorhersage, wenn der Schaum nicht verwendet wird (siehe Abbildung A2.3.1a).

Bei der Planung einer experimentellen Versuchsdurchführung kann die Verwendung einer Schaumschicht um den Impaktor empfehlenswert sein, dass der Knochen beim Auftreffen des Impaktors an der Kontaktstelle nicht splittert. Zusammenfassend wird empfohlen, die Kompressionscharakteristik bei Verwendung eines Schaumes in einem Experiment zu veröffentlichen.

3.4.6 Sensitivitätsanalysen bezüglich der Knochenpositionierung

Die Knochenlage wurde bei der experimentellen Versuchsdurchführung, wie in Abschnitt 3.2.1 beschrieben, veröffentlicht. Jede experimentelle Versuchsvorbereitung wird in manueller Arbeit von einer Person durchgeführt. Dies bedeutet, dass Abweichungen von dieser angestrebten Knochenlage entstehen können. Mittels dieser isolierten Knochenrotationen um die x-, y- und z-Achse soll festgestellt werden, in welchem Maß der Parameter Knochenlage die Modellvorhersage beeinflusst. Eine Rotation um 4° ist eine sehr offensichtliche Veränderung. Diese signifikante Variation müsste bei der Versuchsvorbereitung zu erkennen sein. Eine Veränderung von $\pm 2^\circ$ ist eine leichte Abweichung, bei der es als realistisch eingestuft werden kann, dass eine solche Abweichung unerkannt bleibt.

Die Ergebnisse zeigen, dass der Parameter der Knochenlage keinen signifikanten Einfluss auf die Modellvorhersage ausübt.

3.4.7 Sensitivitätsanalysen bezüglich Positionierung des Versuchsaufbaus

Die Motivation für die Sensitivitätsanalysen hinsichtlich der Positionierung des Versuchsaufbaus ist dieselbe als in Abschnitt 3.4.6. Aufgrund der manuellen Versuchsvorbereitung können Abweichungen von der idealen Positionierung des Versuchsaufbaus auftreten.

Wenn der Impaktor um die z-Achse gedreht wird, damit nicht mehr parallel zur y-Achse auf den Knochen auftrifft, wird die Modellvorhersage nicht beeinflusst. Dieses Ergebnis ist nachvollziehbar, da sich die impaktierte Fläche, wo der Impaktor auf den Knochen auftrifft, nur sehr geringfügig ändert. Zudem ändert sich der Knochenquerschnitt im Bereich der Mitte des Femurschaftes nicht schlagartig. Es ist zu erwarten, dass erst ab einem signifikanten Drehwinkel von ca. 45° die Modellvorhersage beeinflusst wird, da sich dann auch die impaktierte Fläche des Knochens deutlich vergrößert. Es wird darauf verzichtet, eine solche Fehlpositionierung zu modellieren, da es nicht wahrscheinlich ist, dass eine derartige Fehlpositionierung in der Realität eintritt.

Bei einer versetzten Positionierung der Auflager in y-Richtung tritt ebenfalls keine Änderung der Modellvorhersage auf. Da die Auflager trotz y-Versatzes immer noch parallel in x-Richtung positioniert sind, wird demnach die ursprüngliche Kinematik der Abrollbewegung nicht verändert.

Eine Fehlpositionierung eines oder beider Auflager, aufgrund einer Drehung um die y- oder z-Achse, hat einen signifikanten Einfluss auf die Modellvorhersage. Eine Drehung der Auflager um 5° ist signifikant (siehe Abbildung 23) und sollte vor der Versuchsdurchführung visuell erkennbar sein. Es ist jedoch als realistisch zu bewerten, dass eine Drehung bis 3° unerkannt bleibt. Eine Drehung um die z-Achse bewirkt eine Änderung des Abrollwinkels der Auflager zueinander. Die Abrollbewegung der

Auflager erfolgt nicht mehr parallel zur x-Achse. Diese Veränderung der Auflagerkinematik bewirkt eine Änderung des Kurvenverlaufs der Modellvorhersage.

Bei einer Drehung um die y-Achse wird der Knochen schräg in die Auflager eingebettet. Bei dieser Art der Fehlpositionierung wird der Schwerpunkt des gedrehten Auflagers in x-Richtung näher zum eingebetteten Knochen verschoben. Diese Verschiebung des Schwerpunkts bewirkt, dass die Massenträgheit einfacher überwunden wird, da die neue Lage des Schwerpunkts sich in x-Richtung weiter entfernt vom Drehpunkt befindet.

Bei der Versuchsdurchführung eines Experimentes, welches zwei Stützlager für die Einbettung der Probe verwendet, sollte auf die korrekte Positionierung beider Stützlager geachtet werden, da eine Fehlpositionierung signifikant das experimentelle Ergebnis beeinflussen kann. Es ist darauf zu achten, dass die Auflager parallel in x-Richtung zueinander positioniert sind und eine y-Rotation der Auflager vermieden wird.

3.4.8 Einfluss unterschiedlicher Knochenmodelle auf die Modellvorhersage

Alle Sensitivitätsanalysen zur Untersuchung des numerischen Versuchsaufbaus werden mit dem Femur des GHBMCM AM 50 Modells durchgeführt (siehe Abbildung 25). Der Kurvenverlauf der Modellvorhersage mit dem 2D-Ausgangsmodell in deformierbarer Konfiguration und dem GHBMCM AM 50 Femur ist ähnlich zu den experimentellen Kurvenverläufen (siehe Abbildung A2.7.1c).

In der Bevölkerung existieren zahlreiche Individuen, die bezüglich ihrer Körpergröße und des Körpergewichts dem 50. Perzentil der männlichen Bevölkerung entsprechen. Diese Individuen können zu einer Population zusammengefasst werden. Bezogen auf das 50. Perzentil der männlichen Bevölkerung gibt es daher eine Population individueller Femora.

Das Ganzkörper GHBMCM AM 50 Modell wurde auf Basis eines Individuums entwickelt, welches bezüglich Körperhöhe und Körpergewicht dem 50. Perzentil der männlichen Bevölkerung entspricht. Das bedeutet, dass das Femur-Modell des GHBMCM AM 50 eine Stichprobe unter allen Femora oben genannter Population ist. Das Femur des THUMS AM 50 Modells ist ebenfalls eine solche Stichprobe. Das bedeutet, dass beide Modelle das Femur eines 50. Perzentils der männlichen Bevölkerung darstellen.

Wenn beide Modelle nun mit demselben numerischen Versuchsaufbau in derselben Konfiguration simuliert werden, sind die Modellvorhersagen bezüglich des Kurvenverlaufes sehr unterschiedlich (siehe Abbildung A2.6.1a). Das Femur des GHBMCM-Modells besitzt ein signifikant steiferes Verhalten als das Femur des THUMS-Modells. Dieser Unterschied in der Modellvorhersage ist legitim, da diese Femora-Individuen innerhalb der Population des 50. Perzentils sind und durchaus sehr unterschiedliche Geometrien und Gewebeeigenschaften aufweisen können.

Das Modellverhalten eines numerischen Femur-Modells wird vor allem durch die Geometrie, die Materialeigenschaften und die Modellierung beeinflusst. Unter Modellierung fallen Aspekte wie Netzauflösung oder Elementwahl. Das THUMS-Modell verwendet beispielsweise zur Charakterisierung der Kortikalis-Schicht ein elastisch-plastisches Materialmodell, welches eine Fließspannung von 34,5 MPa [68] aufweist. Dies ist außerhalb des in der Literatur üblichen Bereiches von 100–150 MPa [101]. Wenn nun für beide Modelle die Materialmodellierung des GHBMCM (siehe Tabelle 5) verwendet wird, entsteht eine Materialanalogie zwischen beiden Modellen. D.h. die Modelle unterscheiden sich nur noch bezüglich ihrer Modellierung und Geometrie (siehe Abbildung A2.6.1b). Mit dieser Materialanalogie ist der Unterschied zwischen den Modellvorhersagen bezüglich des Kurvengradienten deutlich geringer. Die Kurvenformen bleiben jedoch unterschiedlich. Gründe für diesen bleibenden Unterschied können die Modellierung oder die Knochengeometrie sein. Wie bereits in Abschnitt 3.2.8 beschrieben, unterscheiden sich die Modelle in beiden Aspekten. Es kann daher nicht näher definiert

werden, welcher der beiden Faktoren den bleibenden Unterschied in der Modellvorhersage trotz Materialanalogie verursacht, oder ob beide Faktoren gleichermaßen dafür verantwortlich sind.

Wie bereits erwähnt sind die Unterschiede in der Modellvorhersage zwischen dem Femur des THUMS-Modells und dem Femur des GHBMC-Modells aus dem Blickwinkel zweier unterschiedlicher Individuen legitim.

Trends in der Automobilindustrie v.a. in Bezug auf das automatisierte Fahren zeigen, dass in Zukunft vermehrt Menschmodelle bei der Fahrzeugentwicklung zur Anwendung kommen werden. Die amerikanische Verkehrsbehörde NHTSA veröffentlichte diesbezüglich ein Dokument Vision For Safety [105], indem ausdrücklich zu virtuellen Absicherungsmethoden mit Menschmodellen ermutigt wurde. Wenn nun Modelle des 50. Perzentils sich stark in ihren Modellvorhersagen unterscheiden, könnte das Modell für den Entwicklungsprozess ausgesucht werden, welches für ein bestimmtes Entwicklungsziel die besten Ergebnisse produziert. Dieser nichtkonformen Anwendung könnte vorgebeugt werden, indem das zu verwendende Menschmodell für ein bestimmtes Entwicklungsziel vorgeschrieben wird. Würde dieser Weg verfolgt werden, dann würde höchstwahrscheinlich auf Dauer nur ein Menschmodell überleben und angewendet werden. Es wäre daher empfehlenswerter, für zukünftige Entwicklungsziele mit Menschmodellen, die Simulation von Populationen vorzuschreiben, welche die Bevölkerungsvariabilität widerspiegeln, anstatt die Verwendung eines spezifischen Modells. Bei der Simulation von Populationen würde eine ganze Kurvenschar als Populationsantwort entstehen. Dies würde bedeuten, dass ein solcher Unterschied zweier Kurven nicht mehr eine bedeutende Rolle besitzt, da die Plausibilität der gesamten Populationsantwort gegen experimentelle Daten im Vordergrund steht.

3.4.9 Problematik bei Literaturvalidierung

Vergleich des entwickelten Simulationsmodells mit der Realität

Das in Abschnitt 3.2.1 beschriebene 2D-Ausgangsmodell ist auf Basis verfügbarer Informationen aus der Literatur entwickelt. Ein Vergleich des Kurvenverlaufes der Modellvorhersage mit den experimentellen Kurvenverläufen (siehe Abbildung A2.7.1c) zeigt eine akzeptable Übereinstimmung. Die in Abbildung 15 definierten und in den Abschnitten 3.2.3 bis 3.2.9 durchgeführten Sensitivitätsanalysen bestätigen weitestgehend die Definition der Parameter des Ausgangsmodells.

Der Reibungskoeffizient zwischen den Bodenplatten und den Auflagern ist der einzige Parameter, bei dem eine Änderung angestrebt werden muss. Ein niedrigerer Reibungskoeffizient ist empfehlenswert. Die Sensitivitätsanalysen in Abschnitt 3.2.4 scheinen die Hypothese zunächst zu bestätigen, dass die Struktursteifigkeit der Auflager bei einer Dicke von 2 mm zu gering ist. Diese zu geringe Struktursteifigkeit bewirkt eine Deformation der Auflager, bevor der Abrollvorgang startet. Diese Deformation ist die Ursache des ersten lokalen Maximums des Kurvenverlaufs der Modellvorhersage. Wenn jedoch die Auflager realitätsnah als 3D-Modell simuliert werden, tritt dieser Effekt des ersten lokalen Maximums ebenfalls auf (siehe Abbildung A2.7.1c). Bei dem realitätsnahen 3D-Modell besitzen die Auflager eine Materialstärke von 4,6 mm. Bei der Sensitivitätsanalyse bezüglich der Struktursteifigkeit der Auflager des 2D-Ausgangsmodells (siehe Abbildung A2.2.1a) zeigt sich, dass ab ca. 5 mm die Auflager steif genug ausgelegt sind, da die Modellvorhersage der deformierbaren Konfiguration sich der Modellvorhersage der starren Konfiguration angleicht. Da die Auflager des 3D-Modells steif genug ausgelegt sind, muss die erste Hypothese der zu geringen Struktursteifigkeit (siehe Abschnitt 3.1.3) verworfen werden. Es ist daher anzunehmen, dass die zweite Hypothese (siehe Abschnitt 3.1.3) zutrifft, dass das lokale erste Maximum der Modellvorhersage durch eine zu große Masse der Auflager und daher durch eine zu große Massenträgheit der Auflager verursacht wird. Wenn

die beiden Auflagervarianten, im 2D-Ausgangsmodell und im realitätsnahen 3D-Modell, miteinander verglichen werden, besitzt das Auflagergewicht des 3D-Modells mehr als die doppelte Masse. Es ist daher für zukünftige Studien empfehlenswert, ein geringeres Auflagergewicht zu wählen, um ein derartiges signifikantes erstes lokales Maximum in der Modellvorhersage zu vermeiden. Daher ist es empfehlenswert, vor der experimentellen Versuchsdurchführung numerische Sensitivitätsanalysen durchzuführen, um signifikante Parameter im Vorfeld zu identifizieren und um derartige artifiziell generierte Effekte im Probenverhalten und dadurch auch in der Modellvorhersage zu vermeiden.

Die realitätsnahe Auflagergeometrie wird einmal als 3D-Modell und einmal als 2D-Modell realisiert (siehe Abschnitt 3.2.9). Der Vergleich beider Modellvorhersagen in der starren Konfiguration zeigt einen deutlichen Unterschied im Kurvenverlauf zwischen beiden Modellierungsarten (siehe Abbildung A2.7.1a). Auch wenn die Geometrie bei der 2D-Modellierung so genau wie möglich abgebildet wird, gibt es geringe geometrische Unterschiede aufgrund der 2D-Modellierung. Diese Unterschiede verändern geringfügig die Lage des Drehpunktes der Auflager und das Auflagergewicht. Dies wirkt sich auf die Kurvenform der Modellvorhersage aus.

Wenn das 3D-Modell in deformierbarer Konfiguration (siehe Abbildung A2.7.1b) simuliert wird, weist es deutlich signifikantere Schwingungen auf als in starrer Konfiguration (siehe Abbildung A2.7.1a). Diese Unterschiede in der Modellvorhersage werden durch die Modellsensitivität bezüglich der Simulationssteuerungsparameter verursacht. Die Modellvorhersage ist sensitiv bezüglich des Zeitschrittes, da die Simulation ohne dynamische Massenskalierung durchgeführt wird. Bereits bei einem Zeitschritt von $0,1 \mu\text{s}$ entsteht diese signifikante Veränderung in der Modellvorhersage. Da Menschmodelle für den Einsatz in Fahrzeugumgebungen entwickelt werden, ist es zumindest in der Anwendung nicht möglich, den Zeitschritt beliebig zu reduzieren. Es ist daher für die Nutzung standardisierter Simulationsumgebungen empfehlenswert, die Auflager in starrer Konfiguration zu verwenden, da bewiesen ist, dass sie steif genug ausgelegt sind.

Auswahl des Validierungslastfalls und Verwendung von Literaturdaten

Bei einer Validierung auf Basis von Literaturdaten ist die Güte eines jeden numerischen Simulationsmodells abhängig von der Genauigkeit der Angaben in der Veröffentlichung. Es ist empfehlenswert, bei der Auswahl von Validierungslastfällen darauf zu achten, dass alle in Tabelle 1 gelisteten Kriterien erfüllt sind. Der Referenzlastfall des dynamischen 3-Punkt-Biegeversuchs des Femurs von Forman et al. [38] wird ausgewählt, obwohl die Auflagergeometrie nicht spezifiziert ist, da kein geeigneterer Lastfall bezüglich der definierten Kriterien verfügbar ist.

Die Diskussion des letzten Abschnitts zeigt, dass zwei Modelle mit unterschiedlicher Mechanik ähnliche Kurvenverläufe erzeugen können. Beide Modellvorhersagen sind ähnlich zu den experimentellen Kurvenverläufen (siehe Abbildung A2.7.1c). Der Unterschied der Modelle liegt darin, dass bei einem Modell der charakteristische Kurvenverlauf durch eine zu hohe Massenträgheit verursacht wird und bei dem anderen Modell durch eine zu geringe Steifigkeit der Auflager. Eine solche falsche Abbildung der Mechanik des Experiments entsteht durch eine nicht charakterisierte Auflagergeometrie. Bei dieser Art von Lastfall beeinflusst die Geometrie, die Lage des Schwerpunkts, sowie das Gewicht der Auflager wesentlich die Modellvorhersage. Um die Mechanik dieser Art von Experimente abbildbar zu machen, müssen diese Parameter definiert sein. Zudem ist es empfehlenswert, vor der Versuchsdurchführung in einem experimentellen Protokoll die Positionierung des Versuchsaufbaus zu überprüfen, da eine schräge oder nicht parallele Positionierung der Auflager die Modellvorhersage signifikant beeinflussen kann (siehe Anhang A2.5).

4 Prüfung der Eignung von Clusterverfahren für den Vergleich biomechanischer Datensätze

4.1 Mathematische Clusteranalysen

4.1.1 Einleitung

Mathematische Clusteranalysen werden angewendet, wenn die Frage zu klären ist, ob verschiedene Objekte Ähnlichkeiten zueinander aufweisen. Clusteranalysen finden in zahlreichen Disziplinen, wie beispielsweise der Medizin, der Archäologie, der Soziologie, der Linguistik, der Biologie oder den Wirtschaftswissenschaften ihre Anwendung [6]. Jede Clusteranalyse unterliegt dem Grundprinzip der Homogenität [5]. Ziel einer Clusteranalyse ist es, Objekte zu Gruppen zusammenzufassen. Objekte innerhalb einer Gruppe sollten untereinander ähnlich sein, wohingegen Objekte aus verschiedenen Gruppen sich möglichst stark unterscheiden sollten [5]. In Abbildung 27 ist beispielhaft das Ergebnis einer Clusteranalyse dargestellt. Eine Punktwolke (links) wurde in drei Gruppen (rechts) aufgeteilt.

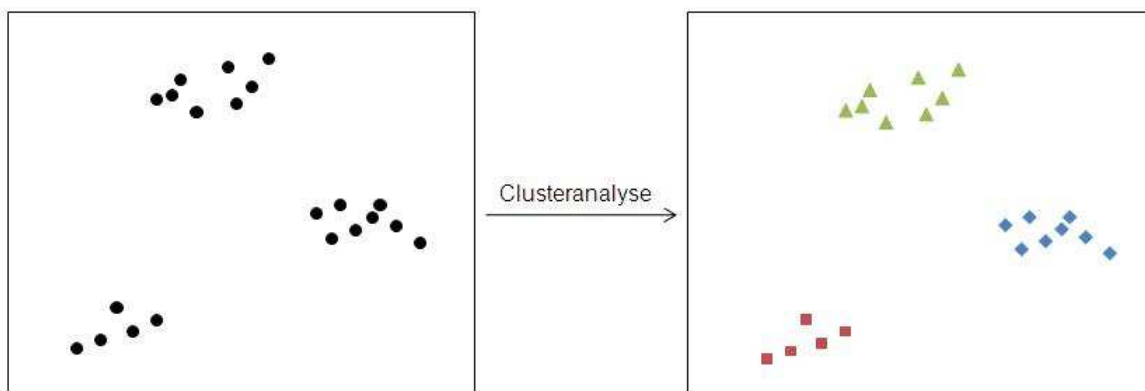


Abbildung 27 – Beispiel einer Clusteranalyse einer Punktwolke [21]

Neben Punktwolken können auch Kurven einer Clusteranalyse unterzogen werden. Die Clusteranalyse von Kurvenscharen wird ebenfalls in den unterschiedlichsten Disziplinen, wie beispielsweise der Astronomie, der Biologie, der Klimaforschung, des Finanzsektors, sowie der Medizin, eingesetzt [2]. Aghabozorgi et al. [2] veröffentlichten einen Überblick über die Anwendungsgebiete der Clusteranalyse von Kurvenscharen mit expliziten Beispielen. In Abbildung 28 ist beispielhaft das Ergebnis einer Clusteranalyse einer Kurvenschar in drei Gruppen dargestellt.

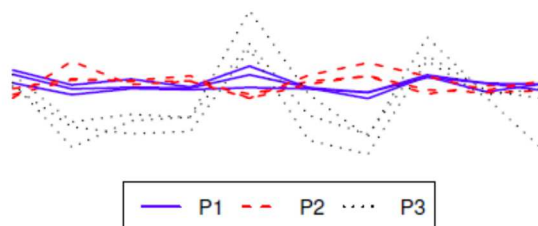


Abbildung 28 – Beispiel einer Clusteranalyse von einer Kurvenschar [80]

In der Biologie wird die Clusteranalyse von Zeitreihen beispielsweise in der Genforschung verwendet [33, 41]. In der Medizin findet die Methode unter anderem Anwendung in der Gehirnforschung zum Messen der Gehirnaktivität [110], oder um pathologische Fälle in klinischen Studien zu identifizieren [48].

In der Biomechanik wurde unter anderem eine Clusteranalyse von Zeitreihen von Symeonidis et al. [99, 100] durchgeführt. Symeonidis et al. unternahmen eine Studie, um den Einfluss von bewussten und unbewussten Verzögerungen von Freiwilligen auf einer Schlittenanlage zu untersuchen. Bei unbewussten Verzögerungen wurden vor dem Hintergrund potenziell automatisiert durchgeführter Bremsvorgänge motorisierter Zweiräder, die Probanden vorab nicht von der Bremsung in Kenntnis gesetzt. Mittels einer Clusteranalyse wurden die Trajektorien der Kopfbewegung der Motorradfahrer untersucht. Die individuell unterschiedlichen Reaktionen führten zu unterschiedlicher Stabilisierung des Kopfes. Dies äußerte sich in einer separaten Clusterbildung.

Neben biomechanischer Trajektorien, können auch biomechanische Datensätze mit anderen Charakteristika einer Clusteranalyse unterzogen werden. Es können beispielsweise biomechanische Datensätze, bestehend aus Kraft-Weg-Verläufen oder Kraft-Zeit-Verläufen einzelner Individuen, mittels einer Clusteranalyse untersucht werden. Zum Beispiel kann bei einem Datensatz, welcher die Knochenbelastungen einer definierten Belastungsart bis zur Fraktur für verschiedene Individuen enthält, eine Clusteranalyse durchgeführt werden. Werden Frakturzeit und Frakturkraft für jedes einzelne Individuum dieses Datensatzes entnommen, dann könnte eine Clusteranalyse dieser Punktwolke durchgeführt werden, um das Knochenversagen im Speziellen, anstatt des vollständigen Belastungsverlaufes, zu untersuchen. Im weiteren Verlauf der Arbeit wird die Clusteranalyse von Punktwolken nicht näher betrachtet.

Es wird im Folgenden explizit die Clusteranalyse von Kurvenscharen betrachtet, da bei dieser Analyse das gesamte Knochenverhalten berücksichtigt wird. Das Integral der mathematischen Kurve eines Knochens ab Belastungsbeginn bis zur Fraktur ist eine Kenngröße bezüglich der Frage, wie viel Energie-Eintrag ein Knochen bis zur Fraktur aufnehmen kann.

4.1.2 Klassifizierung von Clusterverfahren

Clusteranalyseverfahren gehören mathematisch betrachtet zu den Struktur entdeckenden, multivariaten Analyseverfahren [6]. Es existieren zahlreiche Clusteranalyseverfahren, die sich im Wesentlichen in zwei Aspekten unterscheiden: dem Maß mit dem die Distanz zwischen Objekten berechnet wird und der Wahl des Gruppierungsverfahrens, d.h. nach welchem Vorgehen Objekte gruppiert werden [6]. Die Klassifizierung von Clusteranalyseverfahren kann auf verschiedene Arten erfolgen. Eine Möglichkeit ist die Klassifizierung anhand der Zuordnung der Objekte zu den Clustern (siehe Abbildung 29). Eine weitere Möglichkeit ist die Klassifizierung anhand der Vorgehensweise des Fusionierungsalgorithmus (siehe Abbildung 30).

Bei der Klassifizierung anhand der Objektzuordnung wird zwischen heuristischen und modellbasierten Verfahren unterschieden [5]. Bei modellbasierten Verfahren liegt ein Wahrscheinlichkeitsmodell zugrunde [5].

Die unvollständigen Clusteranalyseverfahren ergeben eine Visualisierung der Objekte. Die Zuordnung der Objekte zu bestimmten Gruppen wird vom Anwender übernommen [5]. Diese Art von Verfahren wird in Abbildung 30 als graphentheoretische Verfahren bezeichnet.

Bei den deterministischen Clusteranalyseverfahren werden alle Objekte einem oder mehreren Clustern mit einer Wahrscheinlichkeit von 1 zugeteilt [5]. Zu dieser Gruppe von Clusteranalyseverfahren gehören die in Abbildung 30 dargestellten hierarchischen und partitionierenden Verfahren.

Bei der Anwendung von probabilistischen Clusteranalyseverfahren werden die Objekte einem Cluster nicht mit der Wahrscheinlichkeit 1 oder 0 zugeordnet, sondern mit einem dazwischen liegenden Wahrscheinlichkeitswert [5].

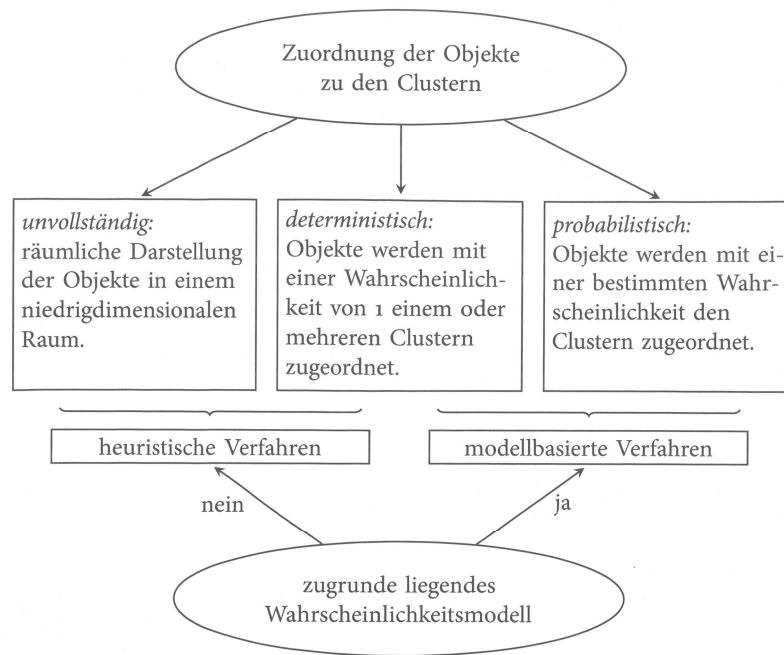


Abbildung 29 – Klassifizierung der Clusteranalyseverfahren auf Basis der Zuordnung der Objekte [5]

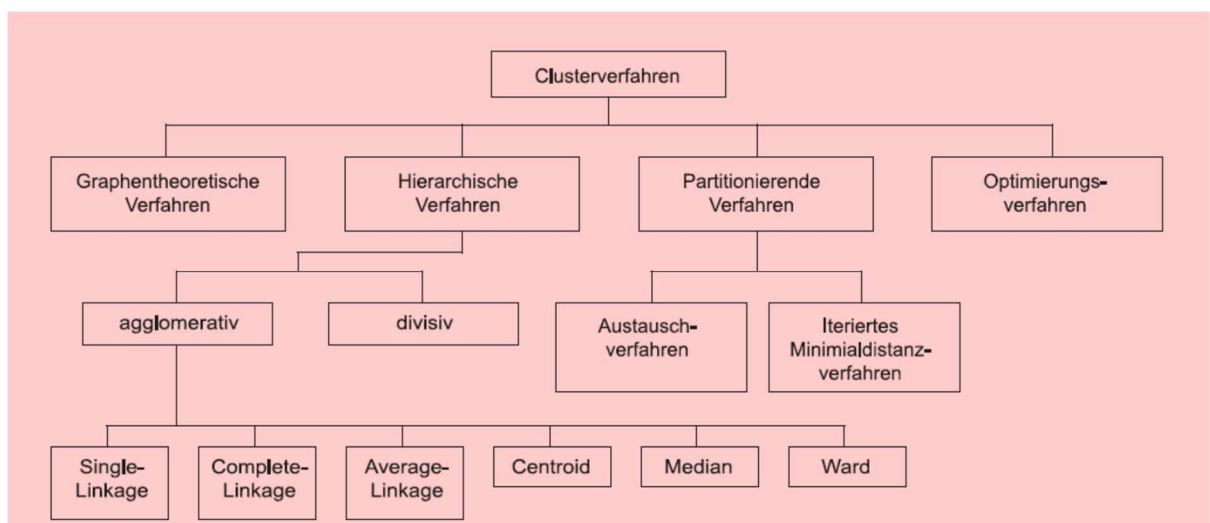


Abbildung 30 – Klassifizierung von Clusteranalyseverfahren anhand Fusionierungsprozess [6]

Im Folgenden werden die hierarchischen und partitionierenden Verfahren kurz erläutert. Bei den partitionierenden Verfahren wird die Clusteranzahl im Voraus festgelegt. Die einzelnen Objekte werden während der Clusteranalyse den festgelegten Clustern zugeteilt, so dass ein bestimmtes Kriterium maximiert oder minimiert wird [5]. Bei den hierarchischen Verfahren wird zwischen agglomerativen und divisiven Verfahren unterschieden. Bei den divisiven Verfahren befinden sich zu Beginn der Clusteranalyse alle Objekte in einem Cluster. Dieses Cluster wird während der Clusteranalyse immer weiter aufgeteilt. Am Ende der Clusteranalyse befinden sich alle Objekte in ihrem eigenen Cluster.

Das agglomerative Verfahren verfolgt das entgegengesetzte Prinzip. Am Anfang der Clusteranalyse bildet jedes Objekt sein eigenes Cluster. Während der Clusteranalyse werden bei jedem Fusionierungsschritt die beiden zueinander ähnlichsten Cluster zusammengefügt [5]. Am Ende der Clusteranalyse befinden sich alle Objekte in demselben Cluster.

In dieser Arbeit werden hierarchisch agglomerative Clusteranalyseverfahren verwendet, welche in Abschnitt 4.1.5 kurz erläutert werden.

4.1.3 Prozess eines Clusterverfahrens für Zeitreihen

Jedes Clusteranalyseverfahren lässt sich in vier Arbeitsschritte unterteilen [6]:

- Bestimmung der Distanzmetrik,
- Auswahl des Fusionierungsalgorithmus,
- Bestimmung der Clusteranzahl,
- Interpretation der Ergebnisse.

Zu Beginn wird eine Distanzmetrik bestimmt. Mit dieser Distanzmetrik wird eine Distanz-Matrix berechnet, welche die berechneten Distanzen zwischen allen möglichen Kurvenpaaren der betrachteten Kurvenschar beinhaltet. Diese Distanz-Matrix dient im weiteren Verlauf als Entscheidungsbasis für die Fusionierungsschritte. In Abschnitt 4.1.4 werden die Distanzmetriken kurz erläutert, die in dieser Arbeit verwendet werden.

Als zweiter Schritt erfolgt die Auswahl eines Fusionierungsalgorithmus. Bei den hierarchisch agglomerativen Clusterverfahren, welche in dieser Arbeit ihre Anwendung finden und in Abschnitt 4.1.5 kurz erläutert sind, werden bei jedem Fusionierungsschritt die zwei zueinander ähnlichsten Cluster zu einem gemeinsamen Cluster zusammengefügt. Die Auswahl der beiden ähnlichsten Cluster wird auf Basis der Distanz-Matrix ermittelt. Nach jedem Fusionierungsschritt muss diese Distanz-Matrix aktualisiert und reduziert werden. Das Ergebnis einer jeden Clusteranalyse kann graphisch als Dendrogramm dargestellt werden (siehe Abbildung 31). Die Äste des Dendrogramms symbolisieren die Distanz zwischen zwei Objekten. Desto geringer die Distanz, umso ähnlicher sind die Objekte zueinander.

Als dritter Schritt wird die Clusteranzahl bestimmt. In diesem Schritt wird untersucht, wie viele Gruppen gebildet werden und wie sich die Ergebnisse bei verschiedenen Gruppenzahlen verändern [6]. Es existieren verschiedene Methoden, eine Clusteranzahl zu bestimmen. Die bekanntesten Methoden sind graphische Verfahren und die Berechnung von Validierungsindizes [2]. Die Validierungsindizes werden in externe und interne Indizes aufgeteilt. Bei den externen Validierungsindizes ist die *wahre* Clusteraufteilung (engl. Begriff: ground truth) bekannt und das Ergebnis der Clusteranalyse kann gegen die *wahre* Clusteraufteilung verglichen werden [2]. Bei den internen Validierungsindizes (Abkürzung: int. VIs) ist die *wahre* Clusteraufteilung nicht bekannt. Int. VIs verwenden nur Informationen der Clusteranalyse und Eigenschaften des Datensatzes, um die Clusteranzahl zu bestimmen [2, 20]. Graphische Verfahren sind beispielsweise die Elbow-Methode, die Gap-Statistik oder die Silhouette-Methode [26]. Bei den graphischen Verfahren wird dem Anwender meist ein Diagramm erzeugt, welches als Basis für den Anwender dient, die Clusterzahl festzulegen. Diese Methoden besitzen den Nachteil, dass sie dem Anwender bei der Festlegung der Clusterzahl einen gewissen subjektiven Spielraum lassen. Deshalb werden in dieser Arbeit ausschließlich int. VIs verwendet, die in Abschnitt 4.1.6 kurz vorgestellt werden.

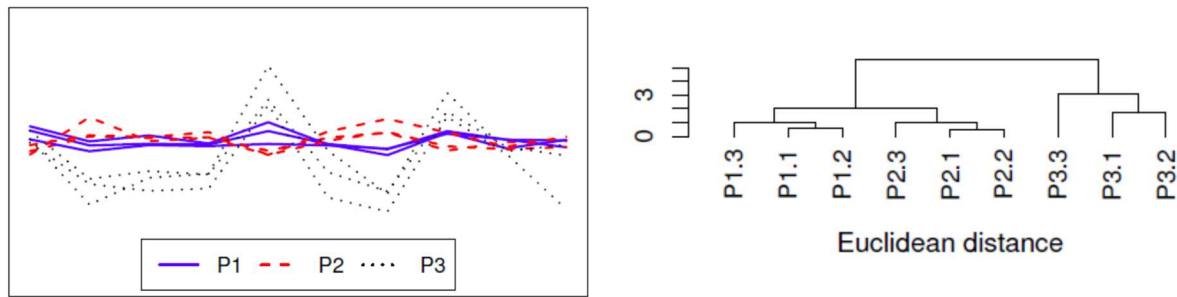


Abbildung 31 – Links: Aufteilung einer Kurvenschar in 3 Cluster / rechts: Dendrogramm-Visualisierung [80]

Als letzter Schritt erfolgen die Analyse und die Interpretation der Ergebnisse der Clusteranalyse. Dabei geht es unter anderem um die Frage, wie sich die einzelnen Cluster voneinander unterscheiden und ob sich die Ergebnisse sinnvoll interpretieren lassen. Auf dieses Thema wird in Abschnitt 4.2.5 und in Kapitel 5 näher eingegangen.

4.1.4 Verwendete Metriken zur Distanzmessung

Die euklidische Distanz und die dynamische Zeitnormierung (engl. Begriff: Dynamic Time Warping (Abkürzung: DTW)) werden in dieser Arbeit als Metriken zur Distanzmessung verwendet. Die euklidische Distanz wird auch als L_2 -Norm bezeichnet. Die euklidische Distanz zwischen zwei Zeitreihen \mathbf{X}_T und \mathbf{Y}_T kann wie folgt berechnet werden [80]:

$$d(\mathbf{X}_T, \mathbf{Y}_T) = \sqrt{\sum_{t=1}^T (X_t - Y_t)^2}.$$

Die DTW-Distanzmessung wurde von Berndt et al. [12] beispielsweise verwendet, um Muster in Zeitreihen zu erkennen. Die DTW-Distanzmessung verfolgt das Ziel, ein Mapping zwischen zwei Zeitreihen zu finden, bei dem die Distanz zwischen gekoppelten Beobachtungen ($\mathbf{X}_{a_i}, \mathbf{X}_{b_i}$) minimiert wird [80]. Sie wird wie folgt berechnet [80]:

$$d_{DTW}(\mathbf{X}_T, \mathbf{Y}_T) = \min_{r \in M} (\sum_{i=1, \dots, m} |X_{a_i} - Y_{b_i}|).$$

In Abbildung 32 wird das verwendete Mapping zwischen zwei Zeitreihen zwischen beiden Arten der Distanzmessung graphisch dargestellt. Die DTW-Distanzmessung eignet sich auch zur Anwendung bei Zeitreihen unterschiedlicher Länge. Zudem erkennt sie ähnliche Kurvenverläufe, auch wenn sie einer Transformation unterliegen, wie beispielsweise einer zeitlichen Verschiebung oder Skalierung bezüglich der Zeitachse [80].

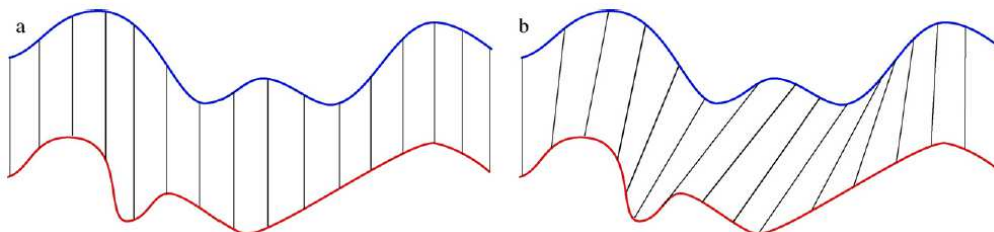


Abbildung 32 – Darstellung des Mappings bei der euklidischen Distanzmessung (links) und der DTW Distanzmessung (rechts) [48]

4.1.5 Verwendete Fusionierungsalgorithmen

In dieser Arbeit werden folgende Fusionierungsalgorithmen verwendet:

- Ward-Algorithmus in zwei Varianten,
- Single-Linkage-Algorithmus,
- Complete-Linkage-Algorithmus,
- Average-Linkage-Algorithmus,
- Mc Quitty-Algorithmus,
- Median-Algorithmus,
- Centroid-Algorithmus.

Die Fusionierungsalgorithmen unterscheiden sich im Wesentlichen darin, wie die Distanz zwischen zwei Clustern berechnet wird. Auf Basis dieser Distanzmessung zwischen den Clustern wird entschieden, welche Cluster während des Fusionierungsschrittes zusammengruppiert werden. Jeder dieser Fusionierungsalgorithmen wird nachfolgend kurz erläutert. In Charrad et al. [20] sind zu jedem der Fusionierungsalgorithmen detaillierte Informationen, wie eine Erklärung der Arbeitsweise des Algorithmus, der Berechnungsformel, sowie weitere Referenzen, veröffentlicht.

Der Ward-Algorithmus fusioniert zwei Cluster mit dem Ziel, dass die Varianz innerhalb des neu fusionierten Clusters so wenig wie möglich durch die Fusionierung zunimmt [20]. Der Ward-Algorithmus existiert in zwei Ausführungen, welche im folgenden *Ward.D* und *Ward.D2* genannt werden. Die Varianten unterscheiden sich dahingehend, dass nur bei der Variante *Ward.D2* das Ward-Cluster-Kriterium implementiert wurde, bei dem die Distanzen vor dem Fusionierungsschritt quadriert werden [20]. Dieser Fusionierungsalgorithmus wurde 1963 von Ward et al. publiziert [109].

Der Single-Linkage-Algorithmus fusioniert zwei Cluster, wenn zwei Elemente aus den einzelnen Clustern die geringste Distanz zueinander aufweisen (siehe Abbildung 33).

Der Complete-Linkage-Algorithmus berechnet jeweils die größtmögliche Distanz zwischen den Elementen von zwei unterschiedlichen Clustern. Es werden diejenigen zwei Cluster fusioniert, welche die geringste größtmögliche Distanz zueinander aufweisen (siehe Abbildung 33).

Bei dem Average-Linkage-Algorithmus ist die Distanz zwischen zwei Cluster der Mittelwert der Distanzen zwischen allen Element-Paaren zwischen den Clustern [20].

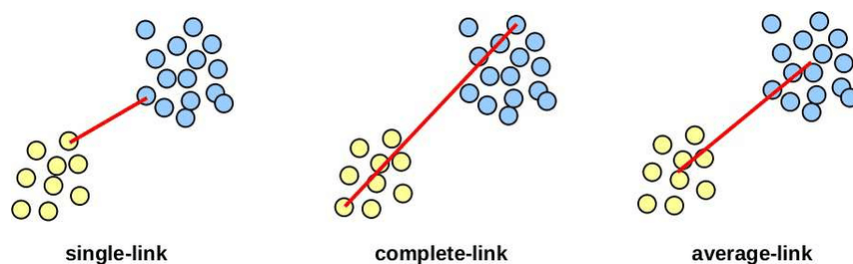


Abbildung 33 – Visualisierung der Distanzmessung zwischen den Clustern bei den Fusionierungsalgorithmen Single-Linkage, Complete-Linkage und Average-Linkage [47]

Die Distanzmessung des Mc Quitty-Algorithmus [75] erfolgt auf Basis des gewichteten Mittelwerts der Distanzen zwischen den Clustern [20]. Die Gewichtung erfolgt auf Basis der Anzahl der Objekte in den jeweiligen Clustern [107].

Der Median-Algorithmus berechnet die Distanz zwischen zwei Clustern als gewichtete Distanz zwischen den Clusterzentren [107]. Die Gewichtung erfolgt auf Basis der Anzahl der Objekte in den jeweiligen Clustern [107].

Der Centroid-Algorithmus berechnet die Distanz zwischen zwei Clustern als quadrierte euklidische Distanz bezüglich der beiden geometrischen Schwerpunkte der beiden Cluster, dies bedeutet zwischen den beiden geometrischen Clusterzentren [20].

4.1.6 Bestimmung der Clusteranzahl

In dieser Arbeit werden int. VIs zur Bestimmung der Clusteranzahl verwendet. Tabelle 12 gibt einen Überblick über die verwendeten Indizes. In Charrad et al. [20] wurden alle verwendeten Indizes ausführlich beschrieben, die Formel für die Berechnung der Indizes veröffentlicht und die dazugehörigen Referenzen angegeben. Aus diesem Grund findet in dieser Arbeit nur eine Auflistung der verwendeten Indizes mit Referenzen statt.

Tabelle 12 – Auflistung der verwendeten internen Validierungsindizes [20]

Index Abkürzungen	Name des Index	Referenz
kl	KL Index	[70]
ch	Calinski und Harabasz (CH) Index	[17]
hartigan	Hartigan Index	[51]
cindex	C-Index	[55]
db	Davies und Bouldin Index	[27]
silhouette	Silhouette Index	[89]
duda	Duda Index	[30]
pseudot2	Pseudo t^2 Index	[31]
beale	Beale Index	[9]
ball	Ball Index	[8]
ptbiserial	Punkt-biserialer Korrelationskoeffizient	[76, 77]
gap	Gap Index	[103]
mcclain	Mc Clain und Rao Index	[74]
gamma	Gamma Index	[7]
gplus	Gplus Index	[77, 88]
tau	Tau Index	[77, 88]
dunn	Dunn Index	[32]
sdindex	SD Validierungsindex	[50]
sdbw	SDbw Index	[49]

Int. VIs verwenden Informationen über die Kompaktheit von Clustern und die Abgrenzung zwischen den Clustern für die Bestimmung der Clusteranzahl [20]. Zudem werden weitere Faktoren für die Berechnung der Clusteranzahl verwendet, wie beispielsweise geometrische oder statistische Eigenschaften des Datensatzes oder der Distanzmatrix [20].

Die in Tabelle 12 enthaltenen Index-Abkürzungen werden in dieser Arbeit in den Ergebnis-Tabellen der Clusteranalysen bezüglich der Clusteranzahlen verwendet. Mit der Verwendung der Index-Abkürzungen, anstatt einer einfachen Nummerierung der Indizes, wird eine verbesserte Lesbarkeit angestrebt.

4.1.7 Zielsetzung des Kapitels

Ziel dieses Kapitels ist es, eine Strategie zur Überprüfung der Eignung eines Clusteranalyseverfahrens für die Analyse von biomechanischen Datensätzen zu entwickeln. Es stellt sich die Frage, wie bei einer solchen Überprüfung vorgegangen wird, damit ein transparentes Vorgehen sichergestellt und eine möglichst objektive Evaluierung ermöglicht wird. Dieses Vorgehen umfasst zudem die Eignungsprüfung von int. VIs zur Bestimmung einer geeigneten Clusteranzahl nach der Clusteranalyse. Es stellt sich dabei vor allem die Frage, ob es eine oder mehrere Kombinationen aus Clusterverfahren und int. VIs gibt, welche für alle Testdatensätze gleich gut geeignet sind. Wenn dies nicht der Fall ist, muss die Frage geklärt werden, ob Kombinationen von Clusterverfahren und int. VIs existieren, welche für bestimmte Eigenschaften der Datensätze geeignet sind.

In diesem Rahmen stellt sich auch die Frage, wie geeignete Testdatensätze aussehen müssten, anhand derer eine solche Evaluierung stattfinden könnte.

4.2 Methoden

4.2.1 Übersicht der Vorgehensweise dieses Kapitels

In Abbildung 34 ist die Vorgehensweise dieses Kapitels visualisiert. Als erster Schritt werden in Abschnitt 4.2.2 Kriterien festgelegt, auf deren Basis die synthetischen Testdatensätze in Abschnitt 4.2.3 generiert werden. Diese Kriterien umfassen Eigenschaften, die in realen biomechanischen Datensätzen vorhanden sein können. Nach der Erstellung der Testdatensätze werden die Clusteranalysen wie in Abschnitt 4.2.4 beschrieben, mit den in Abschnitt 4.1.4 vorgestellten Distanzmetriken und den in Abschnitt 4.1.5 vorgestellten Fusionierungsalgorithmen, durchgeführt.

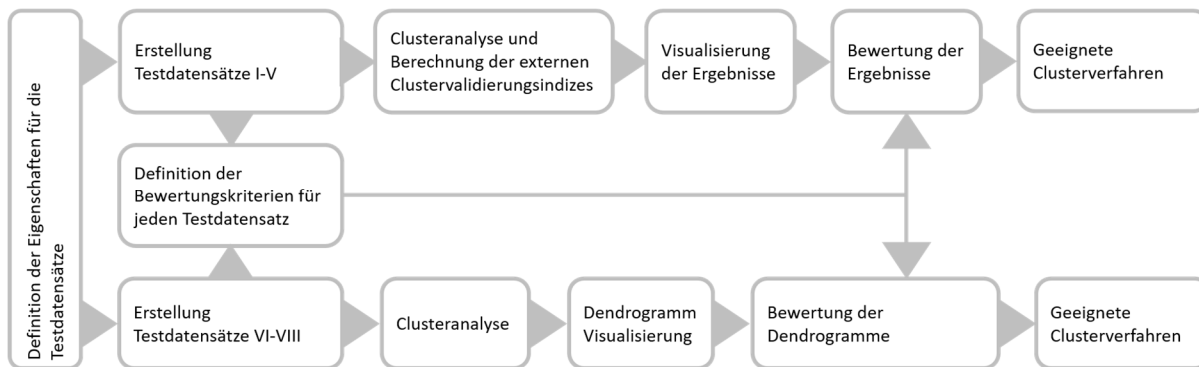


Abbildung 34 – Methodik zur Ermittlung geeigneter Clusterverfahren und Clusteranzahl

Für die Testdatensätze I–V werden nach der Clusteranalyse die in Abschnitt 4.1.6 vorgestellten int. VIs berechnet, um die möglichen Clusteranzahlen zu bestimmen. Jeder int. VI berechnet die Clusteranzahl nach einer unterschiedlichen Formel. Dies bedeutet, dass unterschiedliche Validierungsindizes unterschiedliche Clusteranzahlen für denselben Testdatensatz und dieselbe Kombination von Fusionierungsalgorithmus und Distanzmetrik berechnen können. Es werden für jeden Testdatensatz zwei Tabellen erzeugt, eine Tabelle für jede Distanzmetrik. Die Tabellen enthalten die berechneten Clusteranzahlen für alle Kombinationen von Fusionierungsalgorithmen und int. VIs (siehe Abbildung 35). Alle diese Kombinationen werden visualisiert und anhand der in Abschnitt 4.2.5 beschriebenen Bewertungskriterien bewertet. Diese manuelle Bewertung basiert auf einer visuellen Überprüfung der Dendrogramm-Struktur und der Visualisierung der Clusteraufteilung der Kurvenschar. Rot eingefärbte Tabellenfelder bedeuten eine nicht geeignete Kombination von Fusionierungsalgorithmus und int. VI. Orange eingefärbte Tabellenfelder signalisieren eine geeignete, jedoch nicht favorisierte Kombination. Grün eingefärbte Tabellenfelder bedeuten eine geeignete Kombination.

Es wird für jeden Datensatz überprüft, welche Kombinationen aus Fusionierungsalgorithmen und int. VIs sich als geeignet erweisen. Diese geeigneten Kombinationen der einzelnen Testdatensätze werden in einer Tabelle zusammengefasst (siehe Abbildung 35). Die römische Zahl kennzeichnet die Nummer des Testdatensatzes.

Die beiden Ergebnistabellen werden erneut bewertet, um die geeignetsten Kombinationen aus int. VIs und Fusionierungsalgorithmen zu bestimmen. Wenn eine Kombination nur für einen Testdatensatz oder für zwei Testdatensätze geeignet ist, dies bedeutet für die geringste Anzahl an Testdatensätzen geeignet ist, wird die Kombination als *ungeeignet* bewertet und das dazugehörige Tabellenfeld rot eingefärbt. Erweist sich eine Kombination für die drei Testdatensätze II–IV als geeignet, dann wird die Kombination als *geeignet* bewertet und das Tabellenfeld orange eingefärbt. Wenn eine Kombination sich für mehr als die Testdatensätze II–IV als geeignet erweist, dann wird die Kombination als *sehr gut geeignet* bewertet und das dazugehörige Tabellenfeld grün eingefärbt.

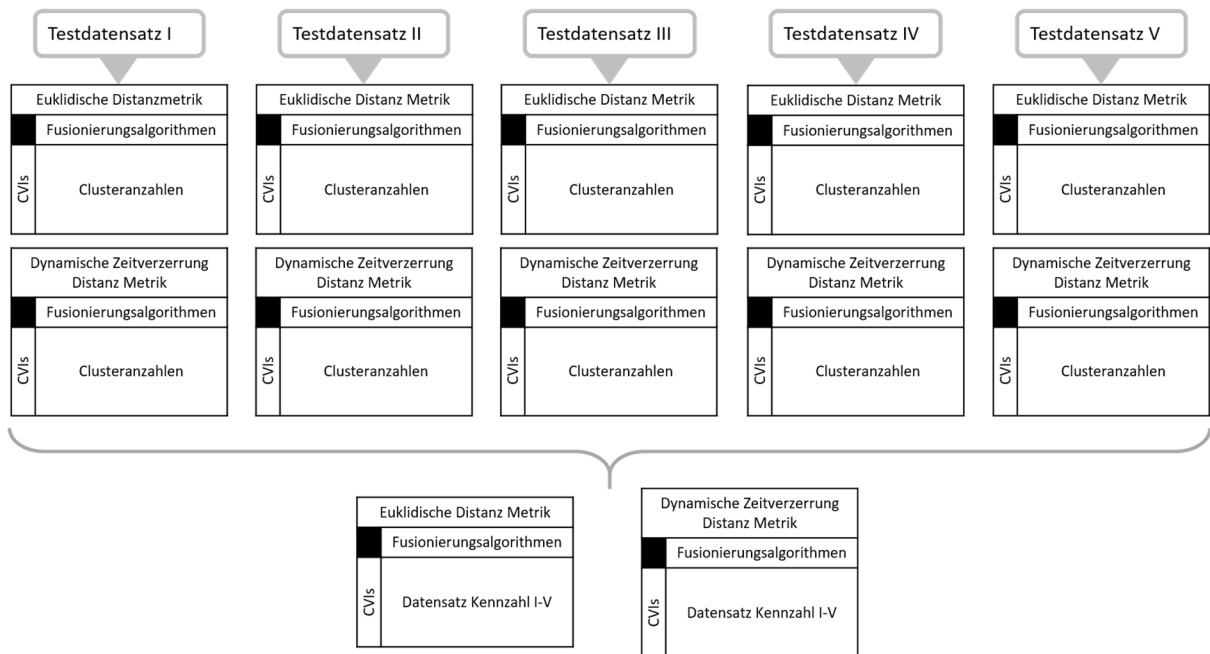


Abbildung 35 – Vorgehen Bewertung der Clusteranalyse

Zudem werden einige Kombinationen für weitere Betrachtungen ausgeschlossen und die dazugehörigen Tabellenfelder grau eingefärbt. Dies bedeutet, dass sich diese Kombinationen nicht zur Verwendung an biomechanischen Datensätzen eignen. Es werden Fusionierungsalgorithmen ausgeschlossen, welche inverse Äste in den Dendrogrammen produzieren. Zudem werden int. VIs ausgeschlossen, welche über alle Datensätze invariant dieselbe Clusteranzahl berechneten. Ebenfalls ausgeschlossen werden Kombinationen, welche für alle Datensätze nicht geeignet waren.

Die Testdatensätze VI–VIII stellen zu betrachtende Sonderfälle dar und fließen daher nicht in die bereits beschriebene Bewertung mit ein. Bei der Analyse der Testdatensätze VI–VIII werden die int. VIs nicht berücksichtigt. Die Analyse dieser Testdatensätze basiert ausschließlich auf der Untersuchung der Dendrogramm-Struktur auf Basis der in Abschnitt 4.2.5 definierten Bewertungskriterien (siehe Abbildung 34). Bei der Bewertung der Dendrogramme wird nur in *geeignete* Dendrogramm-Strukturen und *ungeeignete* Dendrogramm-Strukturen unterschieden.

4.2.2 Kriterien für die synthetische Datensatzgenerierung

Biomechanische Datensätze, wie sie für diese Arbeit relevant sind, besitzen unter anderem folgende Eigenschaften:

- unterschiedliche Frakturzeitpunkte, d.h. Kurvenlänge,
- unterschiedliche Frakturkräfte,
- Unterschiede in der Kurvendichte,
- vorhandene natürliche Gruppierungen.

Diese vier Eigenschaften müssen von einem Clusteranalyseverfahren korrekt interpretiert werden. Zudem sollte bei der Erstellung der Testdatensätze darauf geachtet werden, dass die verwendeten Kurvenformen eine biomechanische Relevanz aufweisen, d.h. eine Ähnlichkeit zu realen experimentellen Kurvenverläufen besitzen.

Des Weiteren werden die im Folgenden beschriebenen Sonderfälle untersucht. Es wird untersucht, ob ein Clusteranalyseverfahren die Kurvenform in der Clusteranalyse mitberücksichtigt und ob es die Fähigkeit besitzt, systematische Fehler, wie beispielsweise eine Stauchung oder Streckung der Kurven, zu erkennen.

4.2.3 Beschreibung der Testdatensätze

In diesem Abschnitt werden die synthetisch generierten Testdatensätze vorgestellt. Sie werden manuell mit Matlab-Version R2016a generiert und sind in Abbildung 36 bis Abbildung 39 visualisiert. Tabelle 13 gibt einen Überblick über die Eigenschaften der Testdatensätze gemäß der Festlegung in Abschnitt 4.2.2. Die Testdatensätze I–V (siehe Abbildung 36a–e) sind bezüglich der Eigenschaften: Kurvenform, Kurvenlänge und Kurvendichte, vereinfachte Testdatensätze. Die Testdatensätze VI–VIII (siehe Abbildung 37 bis Abbildung 39) weisen bezüglich dieser Eigenschaften einen größeren Realitätsbezug auf. Die Eigenschaften der Unterschiede in der Kurvendichte und der Existenz natürlicher Gruppierungen können aufgrund der Undurchsichtigkeit der realitätsnahen Testdatensätze VI–VIII nicht bewertet werden. Diese Felder werden in Tabelle 13 mit einem Fragezeichen markiert.

Tabelle 13 – Eigenschaften der Testdatensätze

Untersuchte Aspekte	Nummer des Testdatensatzes							
	I	II	III	IV	V	VI	VII	VIII
Variation Kurvenlänge		x	x	x		x	x	x
Variation Endpunkte in y-Richtung	x	x	x	x	x	x	x	x
Unterschiede in Kurvendichte			x		x	?	?	?
Existenz natürlicher Gruppierungen		x				?	?	?
Komplett homogener Datensatz	x							
Variation Kurvenform							x	x
Stauchung / Streckung der Kurven						x		

Die Testdatensätze I–V (siehe Abbildung 36a–e) besitzen eine ähnliche Kurvenform wie die experimentellen Daten des dynamischen 3-Punkt-Biegeversuches von Forman et al. [38]. Die Datensätze bestehen immer aus zwei Kurvenscharen: Einer Kurvenschar symbolischer numerischer Kurven, welche in der Kurvenbezeichnung mit dem Buchstaben *S* gekennzeichnet sind, und einer Kurvenschar symbolischer experimenteller Kurven, welche in der Kurvenbezeichnung mit dem Buchstaben *E* gekennzeichnet sind.

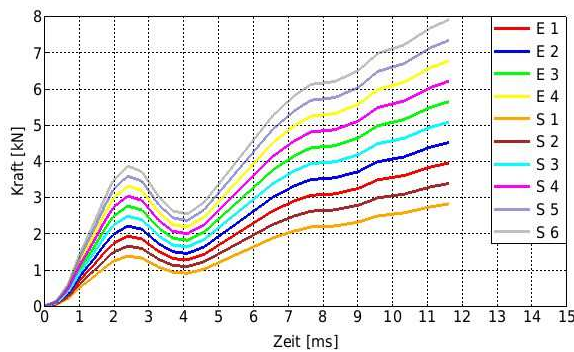
Der Testdatensatz I (siehe Abbildung 36a) stellt einen vollständig homogenen Datensatzes dar. Die Kurvenform, die Kurvenlänge und die Abstände zwischen zwei benachbarten Kurven sind für alle Kurven des Datensatzes gleich.

Der Testdatensatz II (siehe Abbildung 36b) weist eine doppelte natürliche Gruppierung auf. Er besteht aus zwei Kurvenscharen unterschiedlicher Länge. Zwischen den Kurvenscharen, das bedeutet zwischen den Kurven *E5* und *S1*, existiert zudem ein doppelt so großer Kurvenabstand als bei den restlichen benachbarten Kurven.

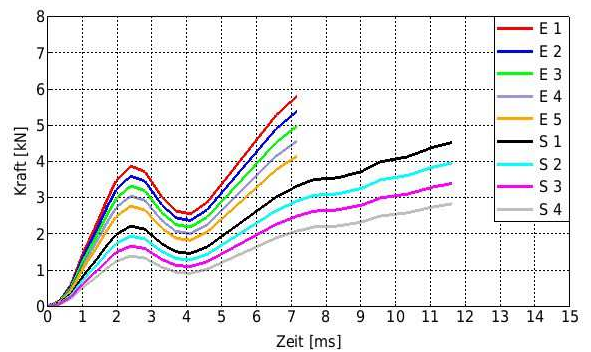
Testdatensatz III (siehe Abbildung 36c) besteht aus zwei sich überlappenden Kurvenscharen unterschiedlicher Länge. Innerhalb jeder Kurvenschar besitzen zwei benachbarte Kurven denselben Abstand zueinander.

Testdatensatz IV (siehe Abbildung 36d) weist dieselben Kurvenabstände zwischen zwei benachbarten Kurven auf. Die Kurvenlänge variiert zwischen zwei Längen. Die Definition der Kurvenlänge für eine Kurve ist zufällig gewählt.

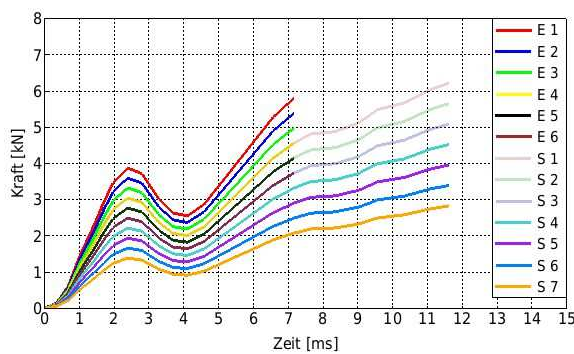
Testdatensatz V (siehe Abbildung 36e) weist bei gleichbleibender Kurvenlänge eine Variation der Kurvendichte auf, d.h. der Abstände zwischen zwei benachbarten Kurven.



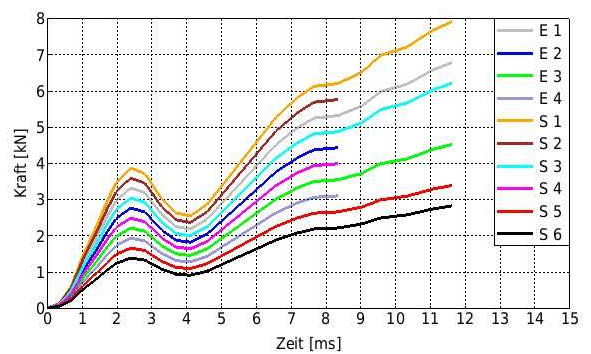
a) Testdatensatz I



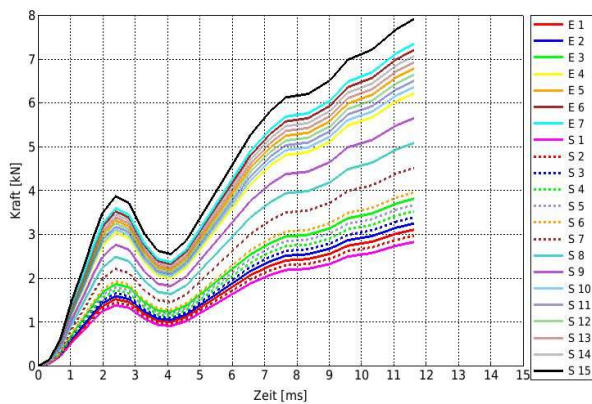
b) Testdatensatz II



c) Testdatensatz III



d) Testdatensatz IV



e) Testdatensatz V

Abbildung 36 – Visualisierung der Testdatensätze I–V

Testdatensatz VI (siehe Abbildung 37) besteht aus zwei Kurvenscharen. Eine Kurvenschar besteht aus den experimentellen Kurven von Forman et al. [38]. Diese Kurven sind mit *M01–M15* gekennzeichnet. Bei der zweiten Kurvenschar wird die Zeitachse um den Faktor 0,5 gestaucht.

In Testdatensatz VII (siehe Abbildung 38) existiert zu jeder experimentellen Kurve, z.B. *M01*, eine lineare Kurve, z.B. *M 01 Lin..* Die jeweils zusammengehörenden Kurvenpaare besitzen denselben Start- und Endpunkt.

Testdatensatz VIII (siehe Abbildung 39) stellt eine realitätsnähere Variante von Testdatensatz VII dar. Bei diesem Testdatensatz wurde von Park et al. [85] für jede experimentelle Kurve eine subjekt-spezifische numerische Kurve generiert (siehe subjekt-spezifischer Validierungsansatz in Abschnitt 1.2.5).

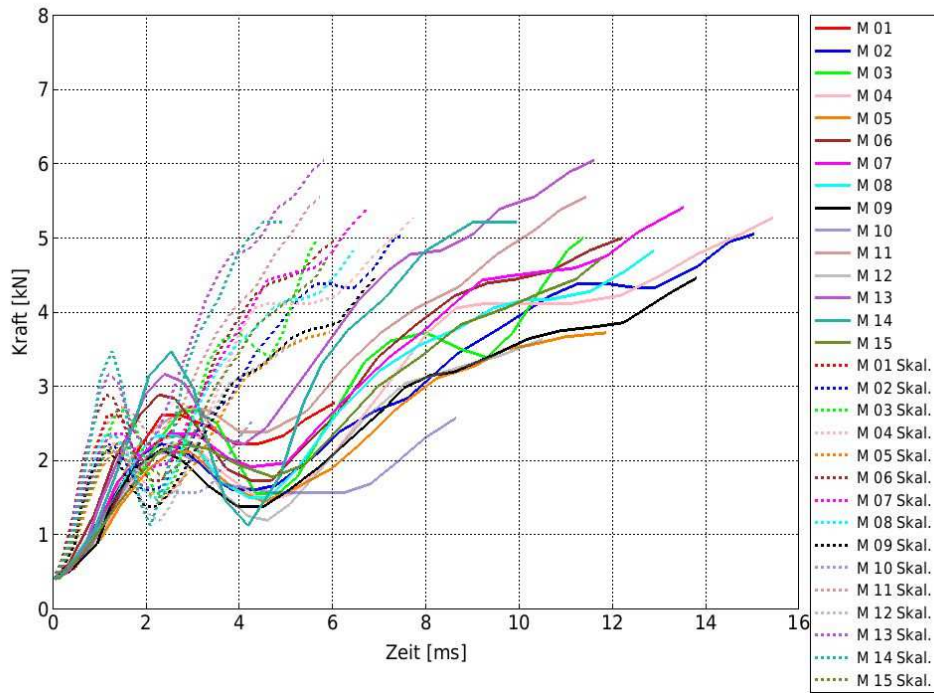


Abbildung 37 – Testdatensatz VI

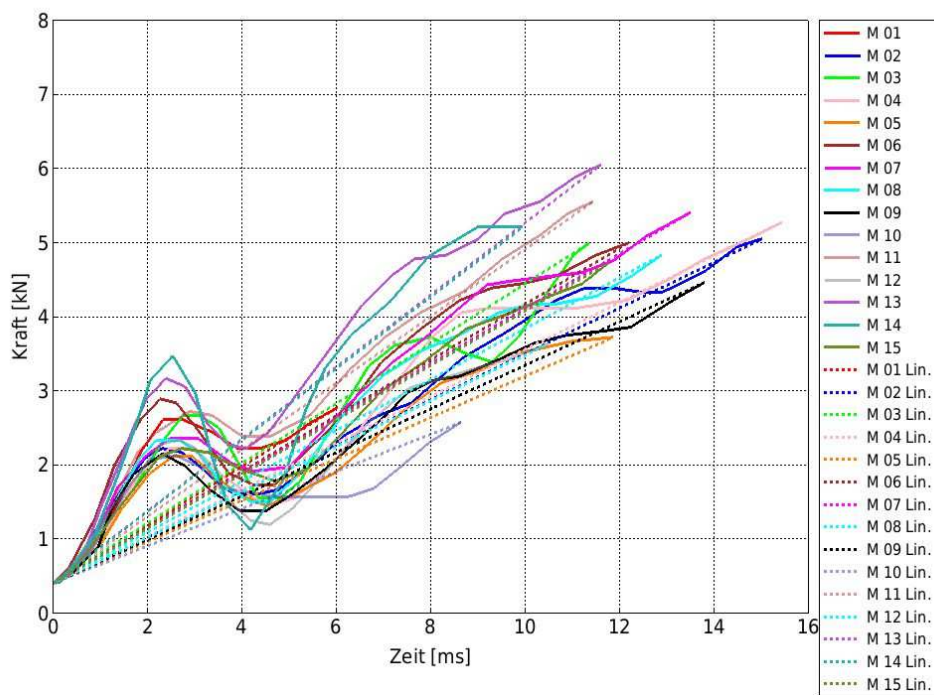


Abbildung 38 – Testdatensatz VII

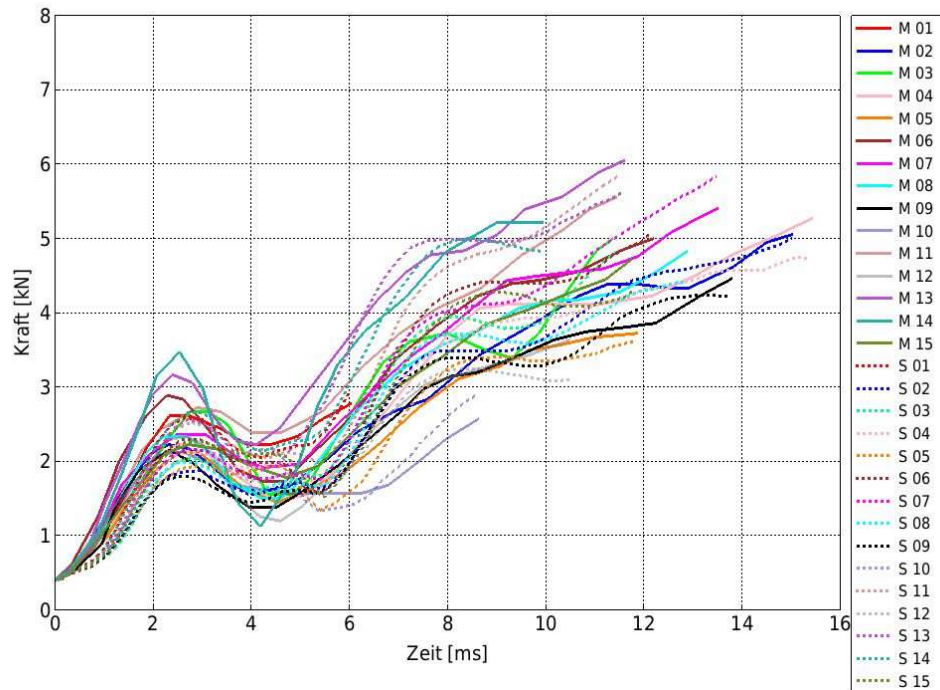


Abbildung 39 – Testdatensatz VIII

4.2.4 Durchführung der Clusteranalyse

Die Clusteranalysen werden mit der öffentlich verfügbaren Statistik-Software RStudio-Version 3.3.4 durchgeführt. Für die Clusteranalysen wird das TScust-Paket verwendet und für die Berechnung der Clusteranzahl wird das NbClust-Paket verwendet. Bei der Berechnung der int. VIs mit dem NbClust-Paket existiert die Limitation, dass die maximal mögliche Clusteranzahl, mit der Kurvenanzahl n , $n-1$ beträgt.

Bevor eine Kurvenschar einer Clusteranalyse unterzogen werden kann, müssen die einzelnen Kurven vorbereitet werden. Für die Clusteranalyse müssen alle Kurven dieselbe Länge aufweisen. Dies bedeutet, dass zuerst die maximale Kurvenlänge aus der gesamten Kurvenschar ermittelt wird, um den maximalen x-Wert der Kurvenschar zu erhalten. Alle Kurven der Kurvenschar, welche einen kleineren x-Wert als Kurvenendpunkt besitzen, müssen mit Null-Werten bis zu dem maximalen x-Wert aufgefüllt werden. Nach der Anpassung der Kurvenlänge, werden die Kurvenpunkte der einzelnen Kurven bezüglich der x-Werte vereinheitlicht. In Abbildung 40 ist dieser Prozess der Datenvorbereitung visualisiert.

Danach wird die Clusteranalyse, wie in folgendem Pseudo-Code veranschaulicht, durchgeführt:

```

For i = 1:1:length(Test_Data_Sets)
  For j = 1:1:length(Distance_Metrics)
    For k = 1:1:length(Fusion_Algorithms)
      Calculate Distance Matrix with Distance Metric j
      Apply Fusion Algorithm k to Distance Matrix
      Visualize Dendrogram of Clustering Result
      (Determine Internal CVIs)
    End
  End
End
End.

```

Für jeden der in Abschnitt 4.2.3 definierten Testdatensätze wird mit allen möglichen Kombinationen aus in Abschnitt 4.1.4 vorgestellten Distanzmaßen und in Abschnitt 4.1.5 vorgestellten Fusionierungsalgorithmen eine Clusteranalyse durchgeführt. Der Clusteranalyseprozess wurde bereits in Abschnitt 4.1.3 ausführlich erläutert. Für die Testdatensätze I–V werden zudem die int. VIs berechnet, welche bereits in Abschnitt 4.1.6 vorgestellt wurden.

Nach der Clusteranalyse kann die Bewertung der Ergebnisse wie in Abschnitt 4.2.1 beschrieben, mit den in Abschnitt 4.2.5 definierten Bewertungskriterien erfolgen.

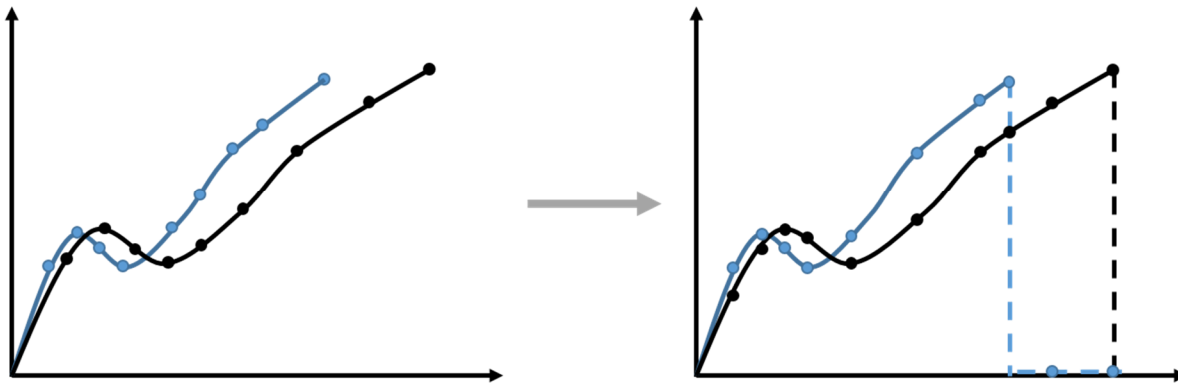


Abbildung 40 – Datenvorbereitung für die Clusteranalyse

4.2.5 Bewertungskriterien der Clusteranalyse für die Testdatensätze

In diesem Abschnitt werden die Bewertungskriterien für jeden Testdatensatz kurz beschrieben. Anhand dieser Kriterien wird festgestellt, ob eine Kombination von Fusionierungsalgorithmus, Distanzmaß und int. VI für den jeweiligen biomechanischen Datensatz geeignet ist.

Das Bewertungskriterium für Testdatensatz I hinsichtlich der vollständigen Homogenität ist, dass entweder jede Kurve ihr eigenes Cluster bildet oder alle Kurven in einem Cluster zusammengefasst sind.

Testdatensatz II weist eine doppelte natürliche Gruppierung auf (siehe Abschnitt 4.2.3). Als Bewertungskriterium wird festgelegt, dass ein zufriedenstellendes Ergebnis der Clusteranalyse die Kurven in exakt zwei Cluster aufteilt, basierend auf der Kurvenlänge.

Die Aufteilung der einzelnen Kurven auf die jeweiligen Cluster basiert bei Datensatz III ebenfalls auf der Kurvenlänge. Kurven unterschiedlicher Länge können nicht in einem Cluster zusammengefasst werden. Die Aufteilung der Kurven kann auf zwei Cluster, mit den Elementen E1–E6 in einem Cluster und den Elementen S1–S7 im zweiten Cluster, oder auf vier Cluster (siehe Tabelle 14) erfolgen.

Tabelle 14 – Aufteilung des Testdatensatz III in vier Cluster

Name des Clusters	Elemente
Cluster 1	E1–E3
Cluster 2	E4–E6
Cluster 3	S1–S3
Cluster 4	S4–S7

Für den Testdatensatz IV ist ebenfalls ein Bewertungskriterium, dass Kurven unterschiedlicher Länge nicht in einem Cluster zusammengefasst werden sollen. Die minimale Clusteranzahl wird auf drei festgelegt (siehe Tabelle 15) und die maximale Clusteranzahl beträgt sieben Cluster (siehe Tabelle 17). Die erwarteten Clusteraufteilungen sind die Aufteilung in drei oder vier Cluster (siehe Tabelle 16).

Tabelle 15 – Aufteilung des Testdatensatz IV in 3 Cluster

Name des Clusters	Elemente
Cluster 1	S1, E1, S3
Cluster 2	E3, S5, S6
Cluster 3	S2, E2, S4, E4

Tabelle 16 – Aufteilung des Testdatensatz IV in 4 Cluster

Name des Clusters	Elemente
Cluster 1	S1, E1, S3
Cluster 2	E3, S5, S6
Cluster 3	E2, S4, E4
Cluster 4	S2

Tabelle 17 – Aufteilung des Testdatensatz IV in 7 Cluster

Name des Clusters	Elemente
Cluster 1	S1
Cluster 2	E1, S3
Cluster 3	E3
Cluster 4	S5, S6
Cluster 5	S2
Cluster 6	E2, S4
Cluster 7	E4

Das Bewertungskriterium für Testdatensatz V besteht darin, dass die Clusteraufteilung die Variation der Kurvendichte des Datensatzes berücksichtigt. Es sind folgende drei Clusteraufteilungen geeignet: drei Cluster, vier Cluster und sechs Cluster. Die minimale Clusteranzahl beträgt drei Cluster (siehe Tabelle 18). Eine Clusteraufteilung in drei Cluster oder vier Cluster (siehe Tabelle 19) wird favorisiert. Die maximale Clusteranzahl beträgt sechs Cluster (siehe Tabelle 20).

Tabelle 18 – Aufteilung von Testdatensatz V in 3 Cluster

Name des Clusters	Elemente
Cluster 1	S15, E7, E6, S14, S13, E5, S12, S11, S10, E4
Cluster 2	S9, S8, S7
Cluster 3	S6, E3, S5, S4, S3, E2, E1, S2, S1

Tabelle 19 – Aufteilung von Testdatensatz V in 4 Cluster

Name des Clusters	Elemente
Cluster 1	S15
Cluster 2	E7, E6, S14, S13, E5, S12, S11, S10, E4
Cluster 3	S9, S8, S7
Cluster 4	S6, E3, S5, S4, S3, E2, E1, S2, S1

Tabelle 20 – Aufteilung von Testdatensatz V in 6 Cluster

Name des Clusters	Elemente
Cluster 1	S15
Cluster 2	E7, E6, S14, S13, E5, S12, S11, S10, E4
Cluster 3	S9
Cluster 4	S8
Cluster 5	S7
Cluster 6	S6, E3, S5, S4, S3, E2, E1, S2, S1

Das Bewertungskriterium für Testdatensatz VI ist, dass eine geeignete Dendrogramm-Struktur den systematischen Fehler bezüglich der Zeitachse erkennt. Ähnliche Kurvenformen, auch wenn eine davon einer zeitlichen Transformation unterliegt, sollten zusammengruppiert werden. Eine geeignete Dendrogramm-Struktur sollte jeweils in der detailliertesten Ebene des Dendrogramms, die jeweiligen korrespondierenden Kurvenpaare, wie beispielsweise *M01 Skal.* mit *M01*, verbinden.

Testdatensatz VII enthält zwei Kurvenscharen mit deutlich unterschiedlichen Kurvenformen. Das Bewertungskriterium für diesen Testdatensatz ist, dass die Dendrogramm-Struktur die beiden Kurvenscharen mit den Kurven *M01 Lin.* bis *M15 Lin.* und *M01* bis *M15* deutlich voneinander trennt.

Der Testdatensatz VIII enthält subjekt-spezifische Kurvenpaare. Eine geeignete Dendrogramm-Struktur sollte jeweils in der detailliertesten Ebene des Dendrogramms, die jeweiligen subjekt-spezifischen Kurvenpaare miteinander verbinden.

4.3 Ergebnisse

4.3.1 Testdatensätze I–V

Übersicht der Ergebnisse der Testdatensätze I–V

Im Anhang A3 werden für jeden Testdatensatz die Dendrogramme der Clusteranalyse für alle möglichen Kombinationen der in Abschnitt 4.1.4 beschriebenen Distanzmaße und in Abschnitt 4.1.5 enthaltenen Fusionierungsalgorithmen visualisiert. Im Anhang A4 befinden sich für jeden Testdatensatz zwei Tabellen. Eine Tabelle bezieht sich auf das euklidische Distanzmaß und die Andere auf die dynamische Zeitnormierung. Die Tabellen enthalten die berechneten Clusteranzahlen für alle Kombinationen von Fusionierungsalgorithmen und int. VIs. Tabellenfelder, die mit rot markiert werden, bedeuten eine nicht geeignete Kombination von Fusionierungsalgorithmus und Clustervalidierungsindex. Tabellenfelder, die mit grün markiert werden, bedeuten eine geeignete Kombination.

Tabelle 21 enthält eine Übersicht der Ergebnisse der Clusteranalyse für die Testdatensätze I–V. Die Abkürzung *EUCL* wird für die euklidische Distanzmetrik und die Abkürzung *DTW* für die dynamische Zeitnormierung verwendet. Die Tabelle enthält Referenzen zu allen Ergebnissen der Clusteranalyse, welche in den bereits genannten Anhängen dargestellt sind. In diesem Abschnitt werden vereinzelt Dendrogramme zur Veranschaulichung der Bewertung und Visualisierung der Clusterergebnisse, basierend auf den in Abschnitt 4.2.5 definierten Bewertungskriterien, verwendet.

Für die Testdatensätze II–V werden alle Kombinationen aus Fusionierungsalgorithmen und int. VIs aus der Bewertung kategorisch ausgeschlossen, bei denen jede Kurve ihr eigenes Cluster bildet oder alle Kurven zu einem Cluster zusammengefasst sind.

Tabelle 21 – Übersicht der Ergebnisse der Testdatensätze I–V

Testdatensatz	Distanzmetrik	Spezifikation	Ergebnisse
Testdatensatz I	EUCL	Dendrogramm	Abbildung A3.1
Testdatensatz I	EUCL	Int. Vls	Tabelle A4.1
Testdatensatz I	DTW	Dendrogramm	Abbildung A3.2
Testdatensatz I	DTW	Int. Vls	Tabelle A4.2
Testdatensatz II	EUCL	Dendrogramm	Abbildung A3.3
Testdatensatz II	EUCL	Int. Vls	Tabelle A4.3
Testdatensatz II	DTW	Dendrogramm	Abbildung A3.4
Testdatensatz II	DTW	Int. Vls	Tabelle A4.4
Testdatensatz III	EUCL	Dendrogramm	Abbildung A3.5
Testdatensatz III	EUCL	Int. Vls	Tabelle A4.5
Testdatensatz III	DTW	Dendrogramm	Abbildung A3.6
Testdatensatz III	DTW	Int. Vls	Tabelle A4.6
Testdatensatz IV	EUCL	Dendrogramm	Abbildung A3.7
Testdatensatz IV	EUCL	Int. Vls	Tabelle A4.7
Testdatensatz IV	DTW	Dendrogramm	Abbildung A3.8
Testdatensatz IV	DTW	Int. Vls	Tabelle A4.8
Testdatensatz V	EUCL	Dendrogramm	Abbildung A3.9
Testdatensatz V	EUCL	Int. Vls	Tabelle A4.9
Testdatensatz V	DTW	Dendrogramm	Abbildung A3. 10
Testdatensatz V	DTW	Int. Vls	Tabelle A4.10

Testdatensatz I

Bei ausschließlicher Betrachtung der Dendrogramm-Struktur wird das in Abschnitt 4.2.5 festgelegte Bewertungskriterium für Testdatensatz I nur bei Verwendung des Single-Linkage-Algorithmus in Kombination mit dem euklidischen Distanzmaß erfüllt. Das in Abbildung 41a dargestellte Dendrogramm visualisiert deutlich, dass entweder alle Kurven in einem Cluster zusammengefasst sind, oder alle Kurven ihr eigenes Cluster bilden (siehe Abbildung 42a). Bei der Verwendung des Single-Linkage-Algorithmus in Kombination mit der dynamischen Zeitnormierung ist dies nicht der Fall (siehe Abbildung 41b).

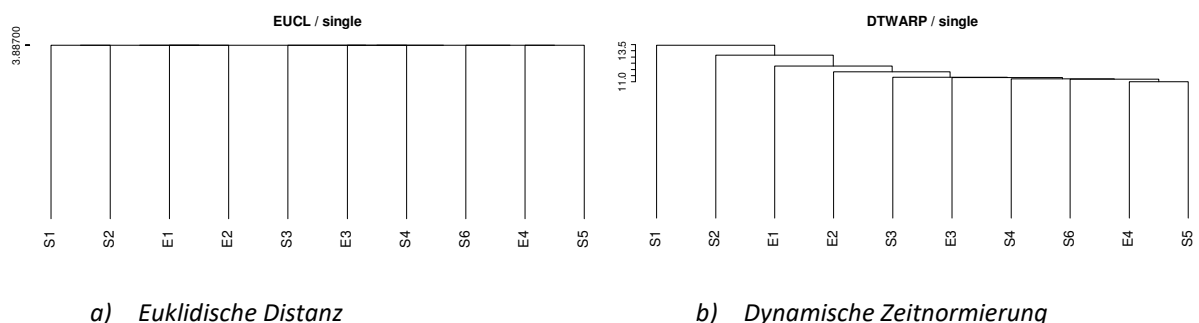
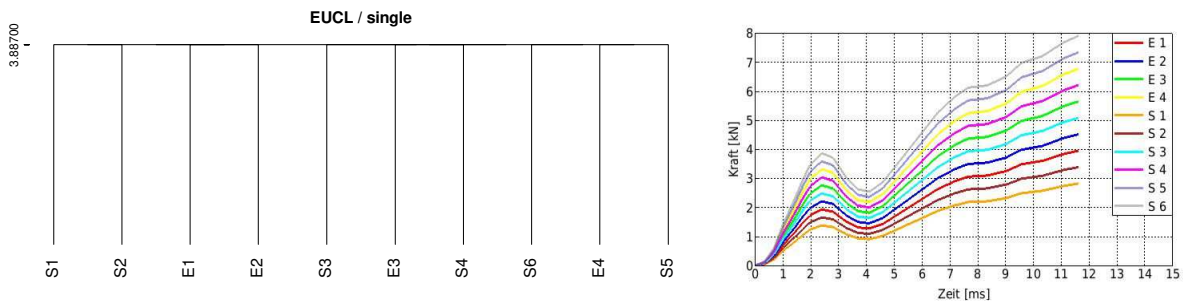
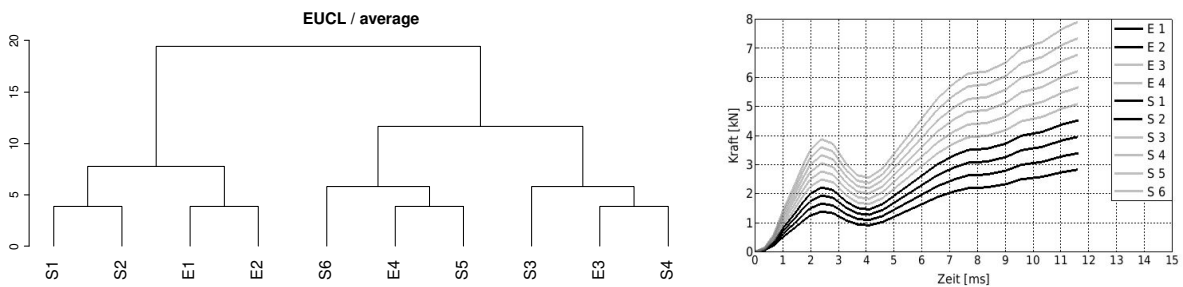


Abbildung 41 – Vergleich der Dendrogramme zwischen den verwendeten Distanzmetriken

Die restlichen Fusionierungsalgorithmen generieren für beide Distanzmaße ein ähnlich strukturiertes Dendrogramm mit ausgeprägten Ästen (siehe Abbildung 42b). Das Vorhandensein von ausgeprägten Ästen in einem Dendrogramm suggeriert die Existenz von Clustern, was bei einem vollständig homogenen Datensatz nicht gegeben ist.



a) Dendrogramm und Visualisierung der Clusteranalyse bei Verwendung des Single-Linkage-Algorithmus



b) Beispiel eines Dendrogramms und einer Visualisierung der Clusteranalyse für die restlichen Fusionierungsalgorithmen

Abbildung 42 – Ergebnisse Testdatensatz I mit euklidischer Distanzmetrik

Einige Kombinationen aus Fusionierungsalgorithmen und int. VIs berechnen beispielsweise eine Aufteilung in zwei Cluster (siehe Abbildung 42b). Bei der Visualisierung dieser Clusteraufteilung wird deutlich, dass eine solche Dendrogramm-Struktur für einen vollständig homogenen Datensatz nicht geeignet ist (siehe Abbildung 42b).

Es werden daher alle Kombinationen aus Fusionierungsalgorithmen und int. VIs, bei denen die berechnete Clusteranzahl von eins oder neun abweicht, für beide Distanzmaße als nicht geeignet bewertet. Eine Clusteranzahl von 9 kann in diesem Fall derart gedeutet werden, dass alle Kurven ihr eigenes Cluster bilden (vgl. Abschnitt 4.1.6).

Testdatensatz II

Es existieren einige Kombinationen aus Fusionierungsalgorithmen und int. VIs, welche das Bewertungskriterium, wie in Abbildung 43a visualisiert, erfüllen (vgl. Tabelle A4.3 und Tabelle A4.4). Bei allen Kombinationen aus Fusionierungsalgorithmen, int. VIs und Distanzmetriken, welche dieselbe Clusteranzahl berechnen, wird die Aufteilung der Kurven zu den jeweiligen Clustern zwischen den Kombinationen verglichen. Alle Kombinationen für vier und sechs Cluster weisen dieselbe Clusteraufteilung auf. Die Ergebnisse sind in Abbildung 43b–d als Negativ-Beispiele visualisiert.

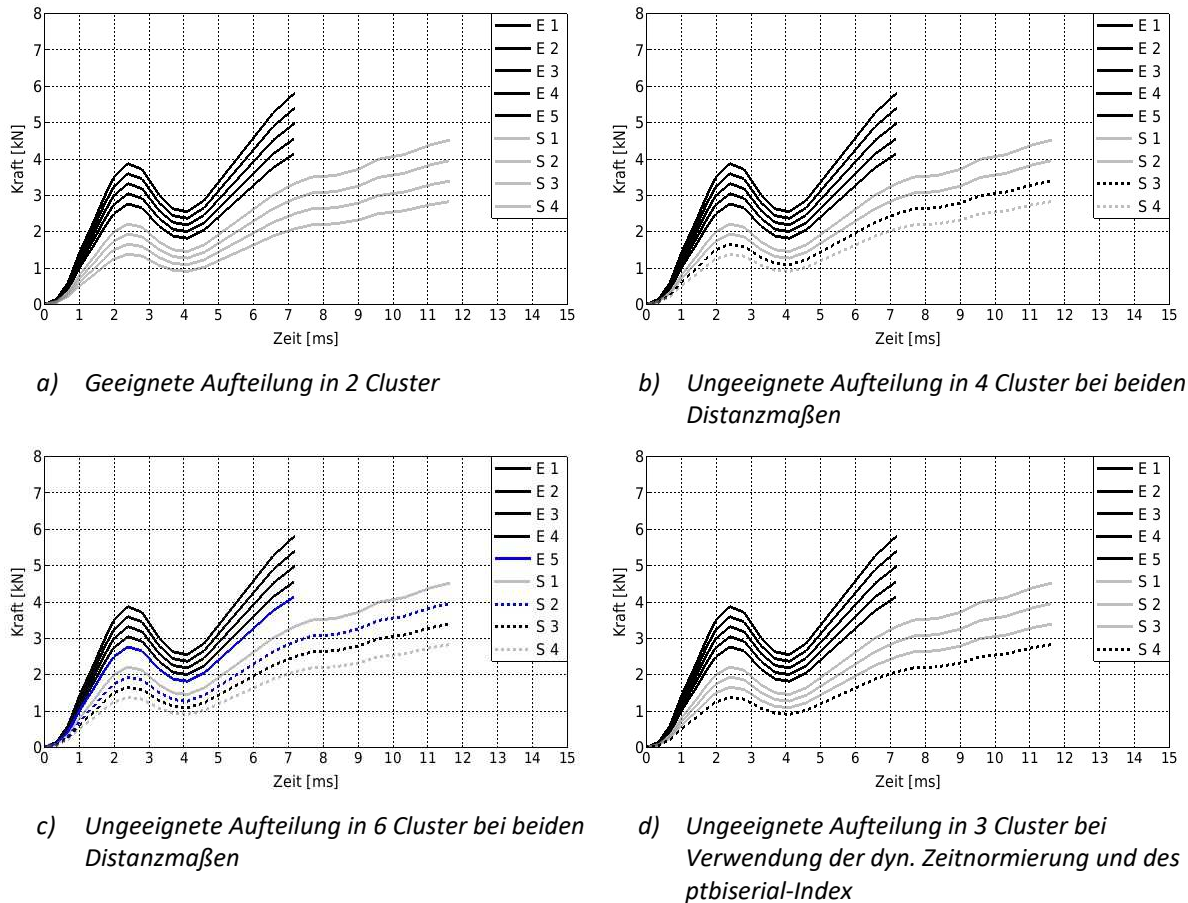
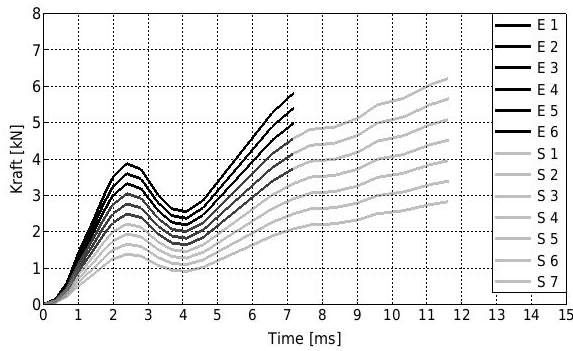


Abbildung 43 – Visualisierung der Ergebnisse der Clusteranalyse für Testdatensatz II

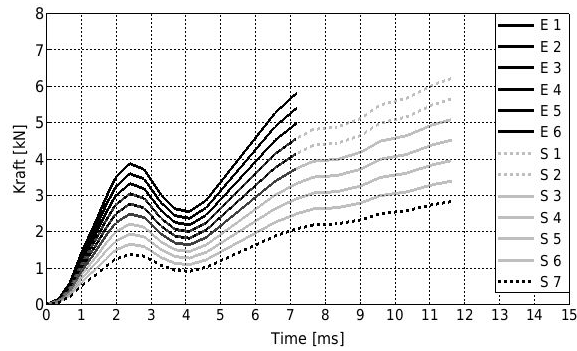
Testdatensatz III

Abbildung 44 enthält die Visualisierungen der berechneten Clusteranzahlen bei Verwendung des euklidischen Distanzmaßes und Abbildung 45 enthält die Visualisierungen der berechneten Clusteranzahlen bei Verwendung der dynamischen Zeitverzerrung. Für jedes Distanzmaß werden die Kombinationen aus Fusionierungsalgorithmen und int. VIs, welche dieselbe Clusteranzahl berechneten, hinsichtlich Gleichheit bei der Clusteraufteilung verglichen. Weisen Kombinationen unterschiedliche Clusteraufteilungen auf, werden sie getrennt visualisiert und untersucht (vgl. Abbildung 45a–c).

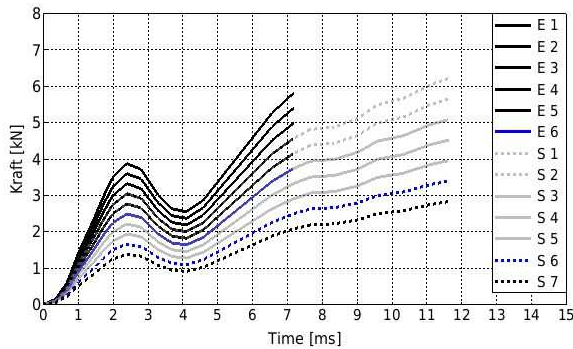
Kategorisch ausgeschlossen werden alle Kombinationen aus Fusionierungsalgorithmen und int. VIs, bei denen jede Kurve ihr eigenes Cluster bildet oder alle Kurven zu einem Cluster zusammengefasst sind. Basierend auf den Bewertungskriterien aus Abschnitt 4.2.5 ist nur die Clusteraufteilung in 2 Cluster (siehe Abbildung 44a) für das euklidische Distanzmaß geeignet. Alle anderen Clusteraufteilungen, welche in Abbildung 44b–e visualisiert sind, werden als nicht geeignet bewertet.



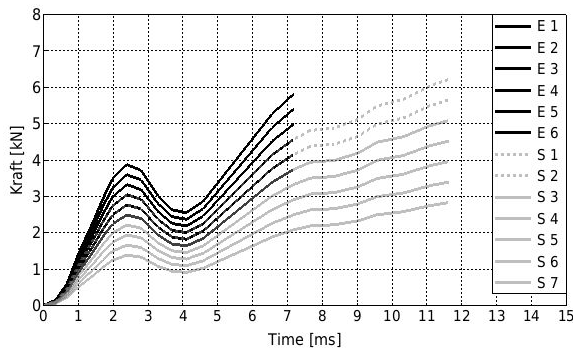
a) Geeignete Aufteilung in 2 Cluster



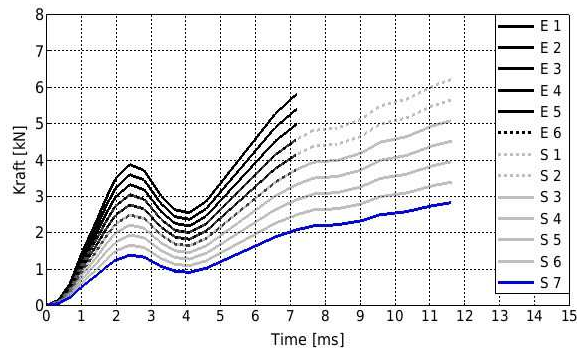
b) Ungeeignete Aufteilung in 4 Cluster



c) Ungeeignete Aufteilung in 6 Cluster



d) Ungeeignete Aufteilung in 3 Cluster mit Ward-Algorithmus

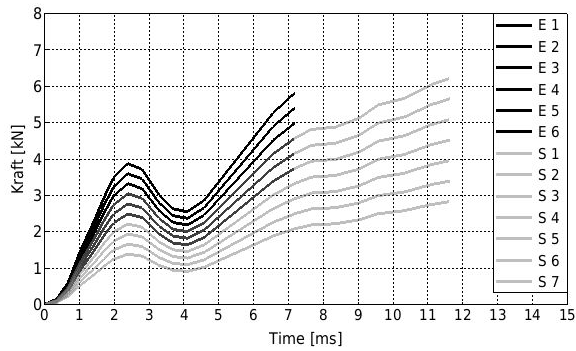


e) Ungeeignete Aufteilung in 5 Cluster mit Ward-Algorithmus

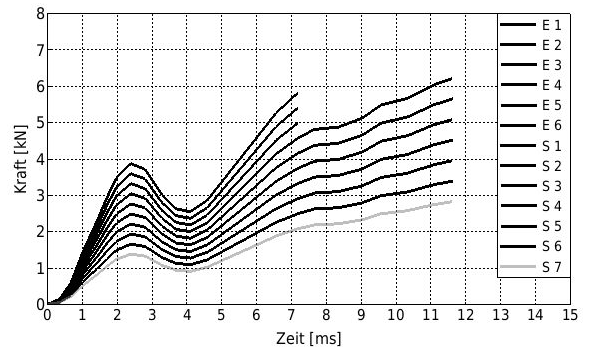
Abbildung 44 – Visualisierung der Ergebnisse der Clusteranalyse für Testdatensatz III bei Verwendung des **euklidischen Distanzmaßes**

Bei Verwendung der dynamischen Zeitnormierung als Distanzmaß, wird nur die Clusteraufteilung in zwei Cluster (siehe Abbildung 45a), jedoch nicht bei Verwendung des Centroid- und Average-Linkage-Algorithmus (siehe Abbildung 45b–c), als geeignet bewertet. Alle anderen Clusteraufteilungen mit einer abweichenden Clusteranzahl werden als nicht geeignet bewertet (siehe Abbildung 45d–g).

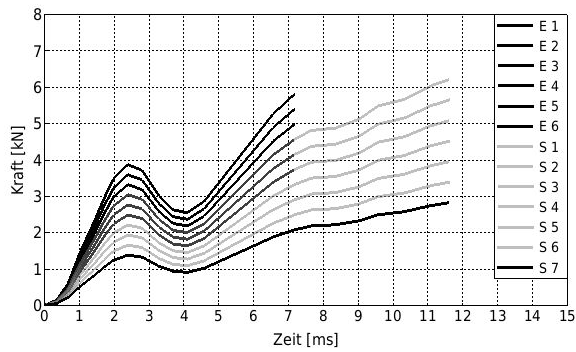
Die Clusteraufteilungen für zwei Cluster (vgl. Abbildung 44a mit Abbildung 45a), fünf Cluster (vgl. Abbildung 44e mit Abbildung 45e) und sechs Cluster (vgl. Abbildung 44c mit Abbildung 45f) sind für beide Distanzmaße identisch.



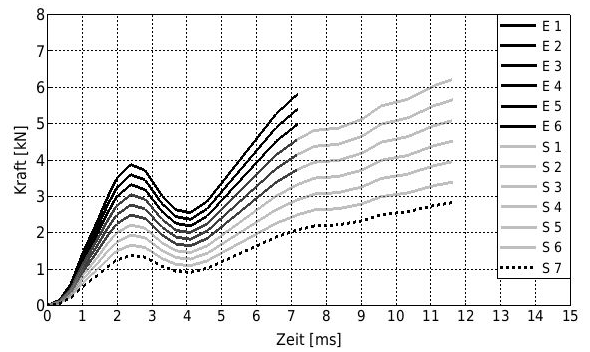
a) Geeignete Aufteilung in 2 Cluster (außer Centroid- und Average-Linkage-Algorithmus)



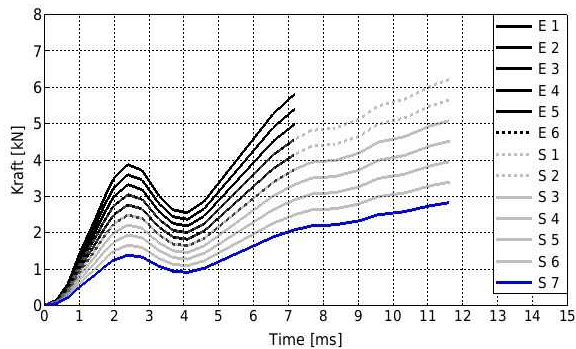
b) Ungeeignete Aufteilung in 2 Cluster bei Verwendung des Centroid-Algorithmus



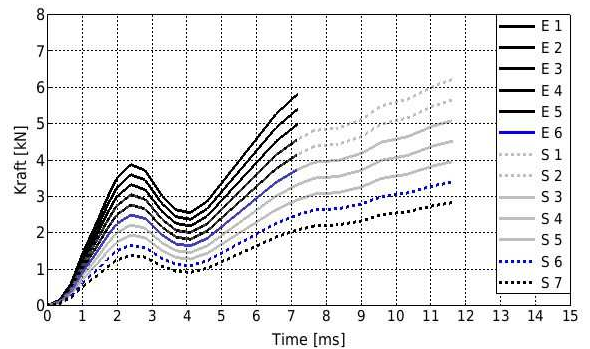
c) Ungeeignete Aufteilung in 2 Cluster bei Verwendung des Average-Linkage-Algorithmus



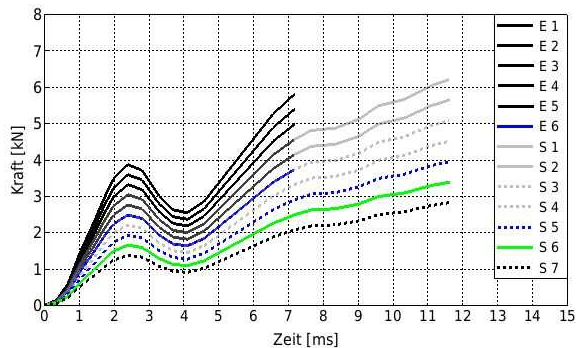
d) Ungeeignete Aufteilung in 3 Cluster



e) Ungeeignete Aufteilung in 5 Cluster



f) Ungeeignete Aufteilung in 6 Cluster



g) Ungeeignete Aufteilung in 7 Cluster bei Verwendung des Average-Linkage-Algorithmus

Abbildung 45 – Visualisierung der Ergebnisse der Clusteranalyse für Testdatensatz III bei Verwendung der **dynamischen Zeitnormierung** als Distanzmaß

Testdatensatz IV

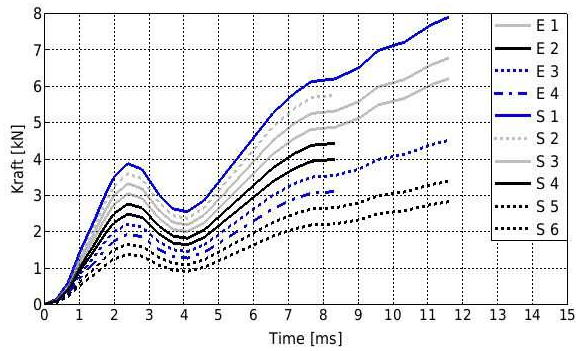
Bei Verwendung des euklidischen Distanzmaßes sind die berechneten Clusteraufteilungen in Abbildung 46 visualisiert und bei Verwendung des Distanzmaßes der dynamischen Zeitverzerrung sind die berechneten Clusteraufteilungen in Abbildung 47 visualisiert.

Für jedes Distanzmaß werden die Kombinationen aus Fusionierungsalgorithmen und int. VIs, welche dieselbe Clusteranzahl berechneten, hinsichtlich Gleichheit bei der Clusteraufteilung verglichen. Alle Kombinationen weisen dieselbe Clusteraufteilung bezüglich einer bestimmten Clusteranzahl auf.

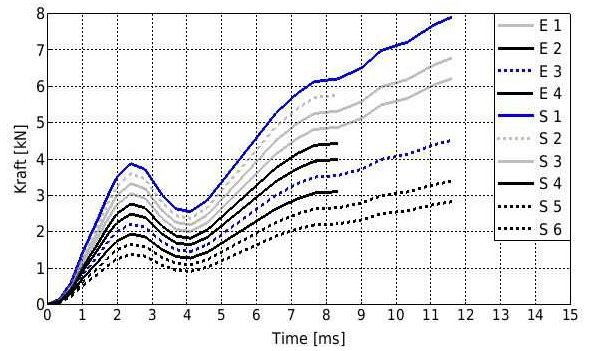
Es werden die Clusteranzahlen, welche die maximale Clusteranzahl überschreiten (vgl. definierte Bewertungskriterien in Abschnitt 4.2.5) als ungeeignet bewertet.

Basierend auf den Bewertungskriterien in Abschnitt 4.2.5 werden bei Verwendung des euklidischen Distanzmaßes die Clusteraufteilungen in sieben Cluster und sechs Cluster als geeignet bewertet (siehe Abbildung 46a–b), die Clusteraufteilungen in drei Cluster und 4 Cluster als favorisiert bewertet (siehe Abbildung 46c–d) und die Aufteilungen in zwei Cluster oder 5 Cluster als nicht geeignet bewertet (siehe Abbildung 46e–f).

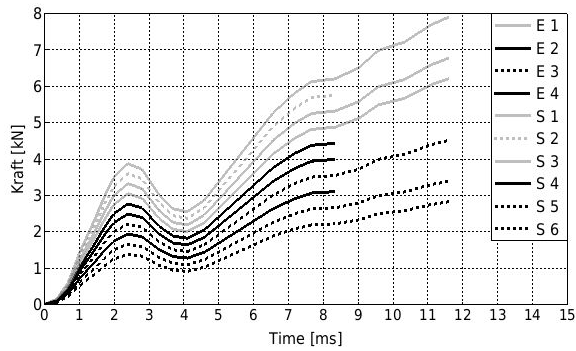
Bei Verwendung des Distanzmaßes der dynamischen Zeitverzerrung werden alle Kombinationen aus Fusionierungsalgorithmen und int. VIs als nicht geeignet bewertet, da bei jeder Kombination Kurven mit unterschiedlicher Kurvenlänge in einem Cluster zusammengefasst werden (siehe Abbildung 47a–e).



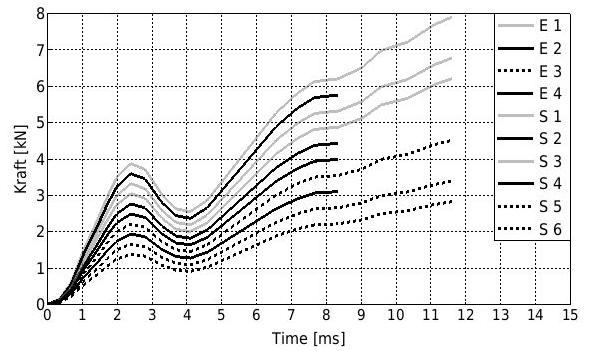
a) Geeignete Aufteilung in 7 Cluster



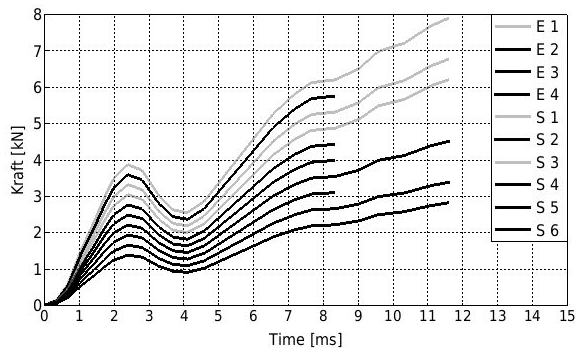
b) Geeignete Aufteilung in 6 Cluster



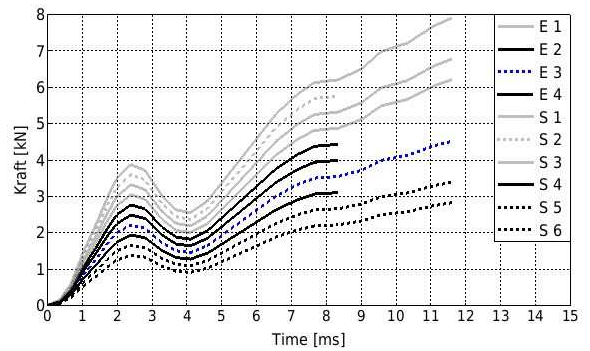
c) Favorisierte Aufteilung in 4 Cluster



d) Favorisierte Aufteilung in 3 Cluster

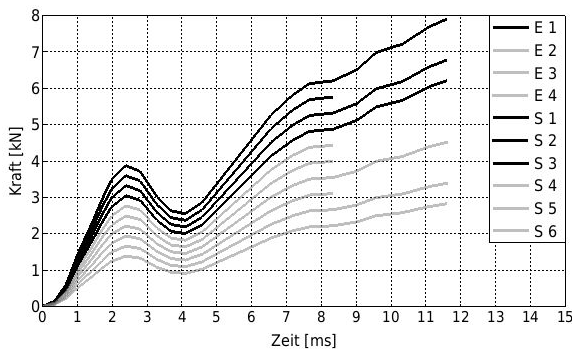


e) Ungeeignete Aufteilung in 2 Cluster

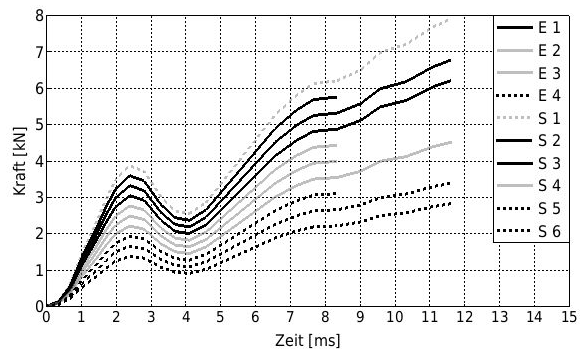


f) Ungeeignete Aufteilung in 5 Cluster

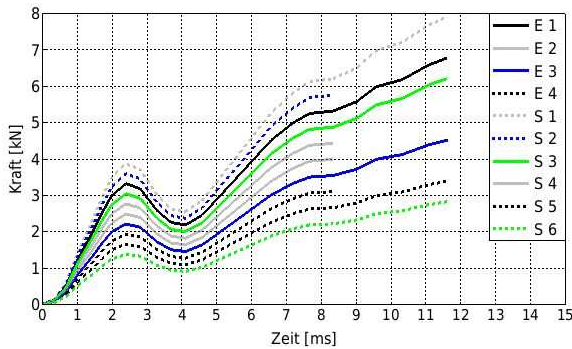
Abbildung 46 – Visualisierung der Ergebnisse der Clusteranalyse für Testdatensatz IV bei Verwendung des **euklidischen Distanzmaßes**



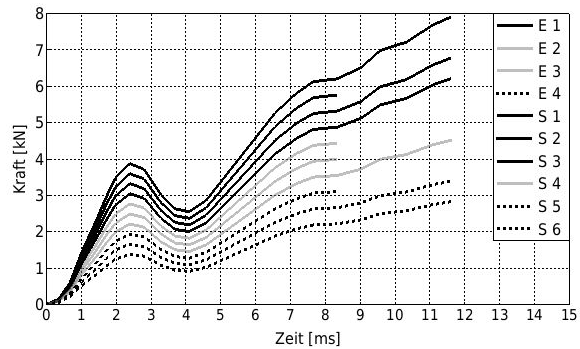
a) Ungeeignete Aufteilung in 2 Cluster



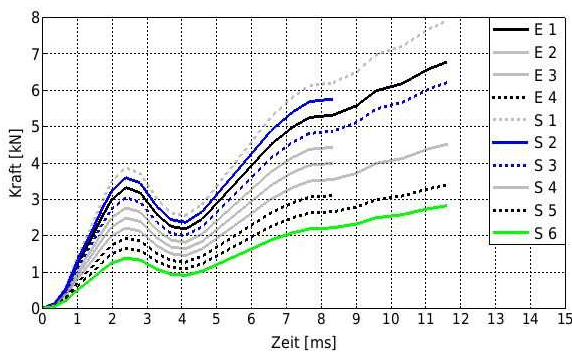
b) Ungeeignete Aufteilung in 4 Cluster



c) Ungeeignete Aufteilung in 8 Cluster



d) Ungeeignete Aufteilung in 3 Cluster bei Verwendung des Complete-Linkage- und des Ward-Algorithmus

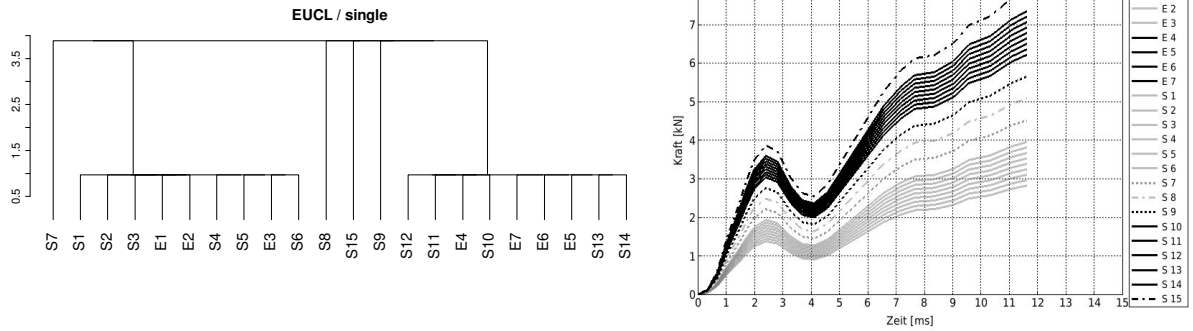


e) Ungeeignete Aufteilung in 7 Cluster bei Verwendung des Centroid- und Median-Algorithmus

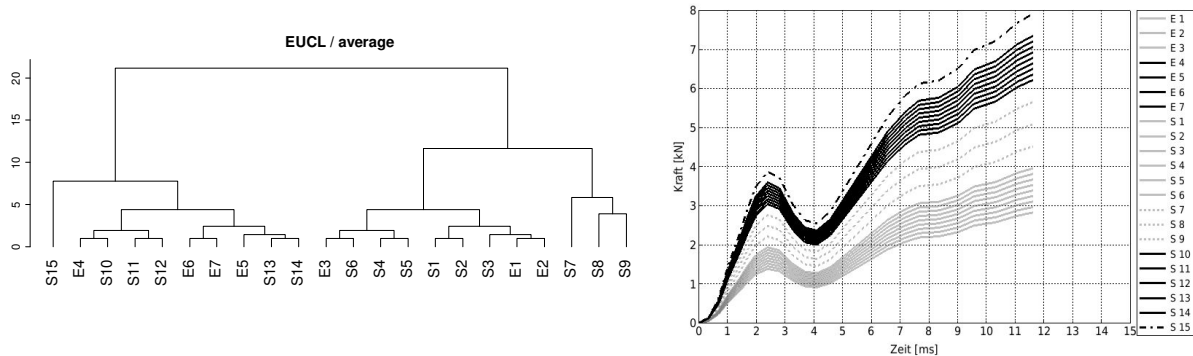
Abbildung 47 – Visualisierung der Ergebnisse der Clusteranalyse für Testdatensatz IV bei Verwendung der **dynamischen Zeitnormierung** als Distanzmaß

Testdatensatz V

Bei Verwendung des Single-Linkage-Fusionierungsalgorithmus in Kombination mit der euklidischen Distanzmetrik existiert kein int. VI, der eine geeignete Clusterlösung vorschlägt. Eine graphische Bewertung des Dendrogramms zeigt eine eindeutige Clusterlösung von sechs Clustern (Abbildung 48a). Dies bedeutet, dass der Single-Linkage-Algorithmus bei diesem Datensatz ausschließlich mit einer Dendrogramm-Bewertung verwendet werden kann, ohne die Verwendung von int. VIs. Die favorisierte Clusterlösung ist in Abbildung 48b dargestellt.



a) Visuelle Aufteilung des Dendrogramms in 6 Cluster bei Verwendung des Single-Linkage Algorithmus

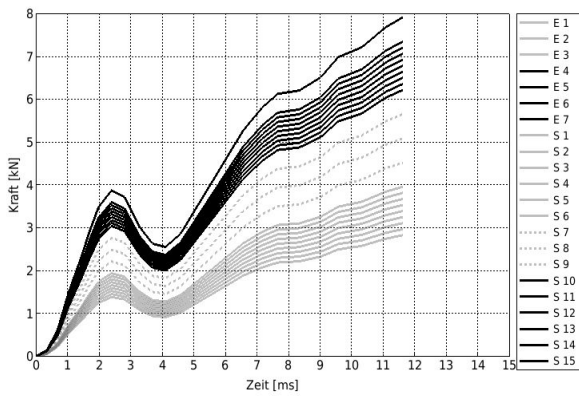


b) Favorisierte Aufteilung in 4 Cluster

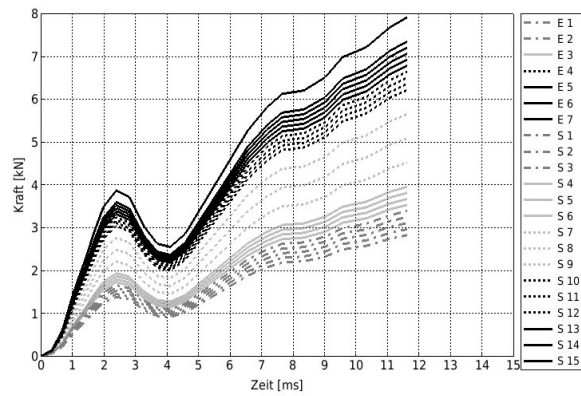
Abbildung 48 – Visualisierung der favorisierten und maximalen Clusteranzahl

Abbildung 49 enthält die Visualisierungen der berechneten Clusteranzahlen bei Verwendung des euklidischen Distanzmaßes und Abbildung 50 enthält die Visualisierungen der berechneten Clusteranzahlen bei Verwendung der dynamischen Zeitverzerrung. Für jedes Distanzmaß werden die Kombinationen aus Fusionierungsalgorithmen und int. VIs, welche dieselbe Clusteranzahl berechnen, hinsichtlich Gleichheit bei der Clusteraufteilung verglichen. Weisen Kombinationen unterschiedliche Clusteraufteilungen auf, werden sie getrennt visualisiert und untersucht (vgl. Abbildung 45a–c). Kategorisch ausgeschlossen werden alle Kombinationen aus Fusionierungsalgorithmen und int. VIs, bei denen die definierte maximale Clusteranzahl überschritten oder die definierte minimale Clusteranzahl unterschritten wird.

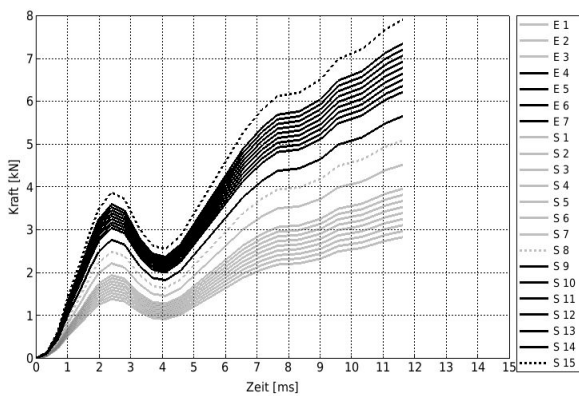
Basierend auf den Bewertungskriterien, welche in Abschnitt 4.2.5 definiert sind, ist nur die Clusteraufteilung in 3 Cluster (siehe Abbildung 49a) für das euklidische Distanzmaß geeignet. Alle anderen Clusteraufteilungen, welche in Abbildung 49b–e visualisiert sind, werden als nicht geeignet bewertet. Die Clusteraufteilung in 3 Cluster (siehe Abbildung 50c) und in 4 Cluster (siehe Abbildung 50g) bei Verwendung des Centroid-Fusionierungsalgorithmus in Kombination mit der dynamischen Distanzmetrik werden ebenfalls als geeignet bewertet. Alle anderen Clusteraufteilungen bei Verwendung der dynamischen Distanzmetrik werden als ungeeignet bewertet (vgl. Abbildung 50a–b und Abbildung 50d–f).



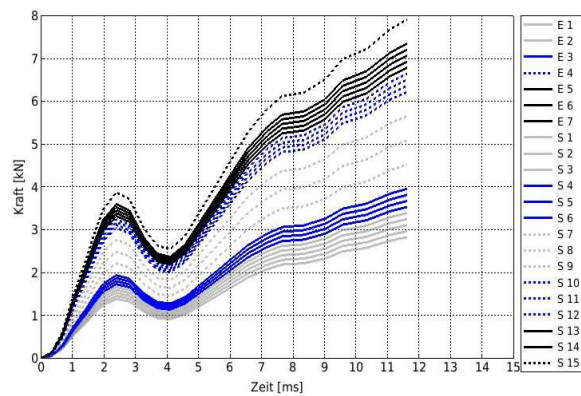
a) Geeignete Aufteilung in 3 Cluster



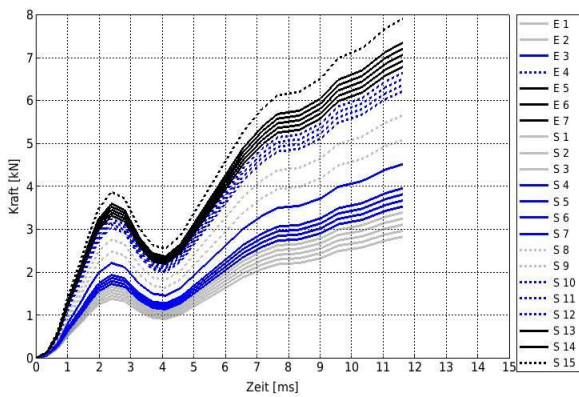
b) Ungeeignete Aufteilung in 5 Cluster bei Verwendung des Ward-Algorithmus



c) Ungeeignete Aufteilung in 4 Cluster bei Verwendung des Single-Linkage-Algorithmus

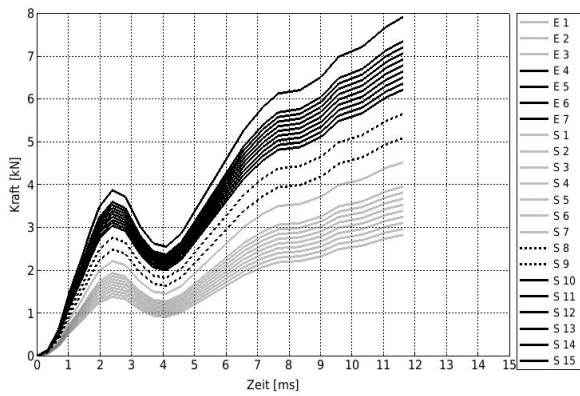


d) Ungeeignete Aufteilung in 6 Cluster bei Verwendung des Ward-Algorithmus

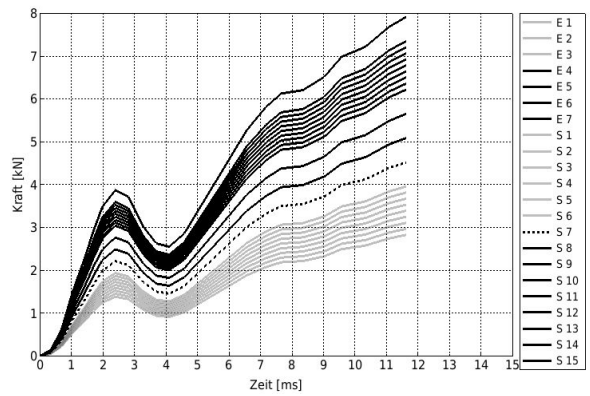


e) Ungeeignete Aufteilung in 6 Cluster bei Verwendung des Complete-Linkage-Algorithmus

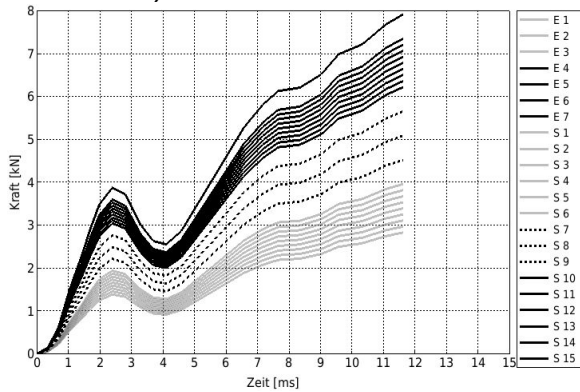
Abbildung 49 – Visualisierung der Ergebnisse der Clusteranalyse für Testdatensatz V bei Verwendung des **euklidischen Distanzmaßes**



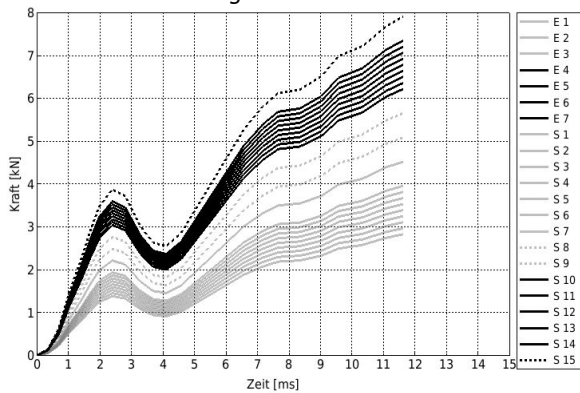
a) *Ungeeignete Aufteilung in 3 Cluster bei Verwendung der Algorithmen Ward, Median und Mc Quitty*



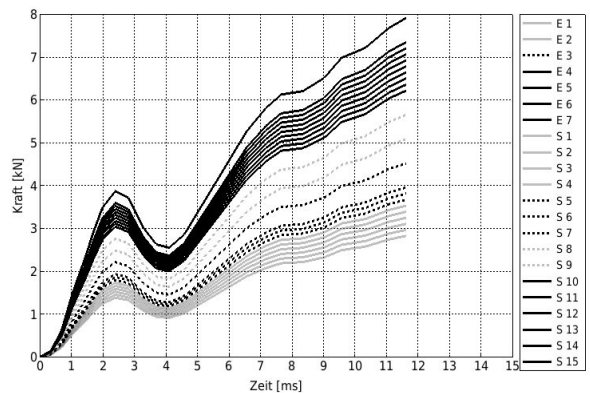
b) *Ungeeignete Aufteilung in 3 Cluster bei Verwendung des Single-Linkage-Algorithmus*



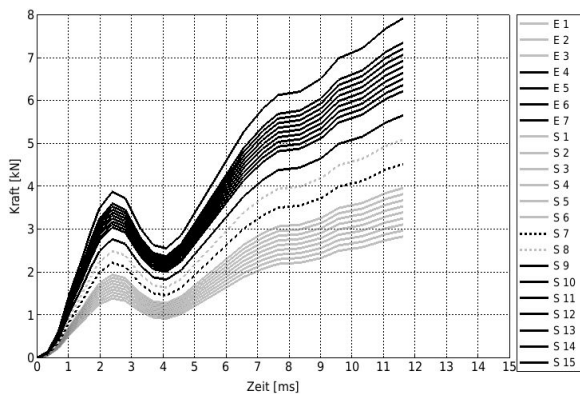
c) *Geeignete Aufteilung in 3 Cluster bei Verwendung des Centroid-Algorithmus*



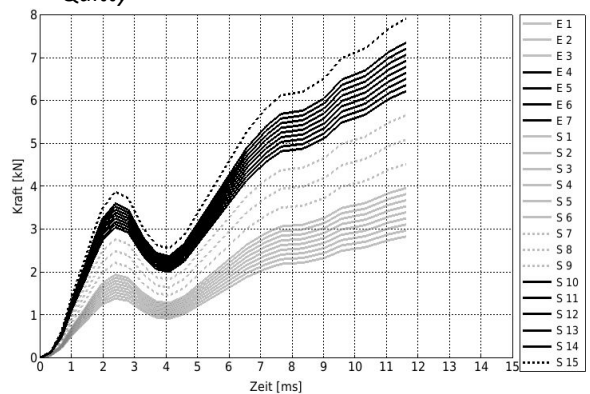
d) *Ungeeignete Aufteilung in 4 Cluster bei Verwendung des Average-Linkage-Algorithmus*



e) *Ungeeignete Aufteilung in 4 Cluster bei Verwendung der Algorithmen Median und Mc Quitty*



f) *Ungeeignete Aufteilung in 4 Cluster bei Verwendung des Single-Linkage-Algorithmus*



g) *Geeignete Aufteilung in 4 Cluster bei Verwendung des Centroid-Algorithmus*

Abbildung 50 – Visualisierung der Ergebnisse der Clusteranalyse für Testdatensatz V bei Verwendung der **dynamischen Zeitnormierung** als Distanzmaß

Zusammenfassung int. VIs der Testdatensätze I–V

Es wird für jeden Datensatz überprüft, welche Kombinationen aus Fusionierungsalgorithmen und int. VIs sich als geeignet erweisen. Diese geeigneten Kombinationen werden in Tabelle 22 zusammengefasst. Die römische Zahl kennzeichnet die Nummer des Testdatensatzes. Tabelle 22 enthält nur die Ergebnisse bei Verwendung der euklidischen Distanzmetrik.

Tabelle 22 – Zusammenfassung der Eignung der int. Clustervalidierungsindizes für die Testdatensätze I–V bei Verwendung der **euklidischen Distanzmetrik**

Zusammenfassung der Ergebnisse für die Testdatensätze I–V								
Int. VIs	Average-Linkage	Centroid	Complete-Linkage	Mc Quitty	Median	Single-Linkage	Ward.D	Ward.D2
kl	II / III	II / III	II / III	II / III	II / III	II / III / IV	II / III	II / III
ch	II / III	II / III	II / III	II / III	II / III	II / III / IV	II / III	II / III
hartigan	I	I	I	I	I	I	I	I
cindex	IV	IV	IV	IV	IV	IV	IV	IV
db	I	I	I	I	I	I	I	I
silhouette	I	I	I	I	I	I	I	I
duda	II / III	II / III	I / II / III	II / III	II / III	II / III	II	II
pseudot2	II / III	II / III	I / II / III	II / III	II / III	II / III	II	II
beale	II / III / IV / V	II / III / IV / V	II / III / IV	II / III / IV / V	II / III / IV / V	II / III / IV	II / IV	II / IV
ball	II / III	II / III	II / III	II / III	II / III	II / III	II / III	II / III
ptbiserial	II / III / IV	II / III / IV	II / III / IV	II / III / IV	II / III / IV	II / III / IV	II / III / IV	II / III / IV
gap	I	I	I	I	I	I	I	I
mcclain								
gamma	II / III	II / III	II / III	II / III	II / III	II / III	II / III	II / III
gplus	II / III	II / III	II / III	II / III	II / III	II / III	II / III	II / III
tau	III	III	III	III	III	III / IV	III	III
dunn								
sdindex			IV					
sdbw	I	I	I	I	I	I	I	I

Erklärung der Einfärbung der Tabellenfelder:
 Grau: Ausgeschlossene Kombinationen aus int. VI und Fusionierungsalgorithmus
 Grün: Geeignete Kombinationen aus int. VI und Fusionierungsalgorithmus für mehr als drei Testdatensätze
 Orange: Geeignete Kombinationen aus int. VI und Fusionierungsalgorithmus für die drei Testdatensätze II–IV
 Rot: Ungeeignete Kombinationen aus int. VI und Fusionierungsalgorithmus

Die Fusionierungsalgorithmen *Centroid* und *Median* werden für weitere Betrachtungen bezüglich der Eignung der Kombinationen ausgeschlossen. Der Grund für den Ausschluss der beiden Fusionierungsalgorithmen ist, dass inverse Äste in der Dendrogramm-Struktur auftreten können, welche schwer zu interpretieren sind (vgl. bspw. Abbildung A3.4). Die int. VIs *hartigan*, *db*, *gap*, *sdbw* werden ausgeschlossen, da sie invariant bei allen Testdatensätzen dasselbe Ergebnis berechnen. Die int. VIs *dunn* und *mcclain* werden ebenfalls von weiteren Betrachtungen ausgeschlossen, da sie für keinen der fünf Testdatensätze eine geeignete Clusteranzahl berechnen. Zudem wird der *silhouette*-Index von weiteren Betrachtungen ausgeschlossen, da er nur für Testdatensatz I die richtige Clusteranzahl berechnet. Die Validierung der Homogenität eines Datensatzes kann anhand einer visuellen Analyse der Dendrogramm-Struktur erfolgen. Für alle ausgeschlossenen Fusionierungsalgorithmen und int. VIs werden die Tabellenfelder in Tabelle 22 grau markiert. Kombinationen die für die drei Testdatensätze II–IV geeignet sind, werden orange eingefärbt. Kombinationen, welche für mehr als die drei genannten Testdatensätze geeignet sind, werden grün eingefärbt. Die restlichen Tabellenfelder werden als ungeeignet bewertet und rot eingefärbt.

Für die dynamische Zeitnormierung werden die Ergebnisse in Tabelle 23 dargestellt.

Tabelle 23 – Zusammenfassung der Eignung der int. Clustervalidierungsindizes für die Testdatensätze I–V bei Verwendung der **dynamischen Zeitnormierung als Distanzmetrik**

Zusammenfassung der Ergebnisse für die Testdatensätze I–V								
Int. VIs	Average-Linkage	Centroid	Complete-Linkage	Mc Quitty	Median	Single-Linkage	Ward.D	Ward.D2
kl	II	II	II/III	II/III	II/III	II/III	II/III	II/III
ch	II	II	II/III	II/III	II/III	II/III	II/III	II/III
hartigan	I	I	I	I	I	I	I	I
cindex								
db	I	I	I	I	I	I	I	I
silhouette	I	I	I	I	I	I	I	I
duda	I/II	I/II	I/II/III	I/II/III	I/II/III	I/II/III	I/II/III	I/II/III
pseudot2	I/II	I/II	I/II/III	I/II/III	I/II/III	I/II/III	I/II/III	I/II/III
beale	II	II	II/III	II/III	II/III	I/II/III	II/III	II/III
ball	II	II	II/III	II/III	II/III	II/III	II/III	II/III
ptbiserial								
gap	I	I	I	I	I	I	I	I
mcclain		V		V	V	V		
gamma	I/II	I/II	I/II	I/II	I/II	I/II	I/II	I/II
gplus	I/II	I/II	I/II	I/II	I/II	I/II	I/II	I/II
tau				V	V			
dunn	I	I	I	I	I	I	I	I
sdindex	I/V	I/V	I	I/V	I/V	I	I/V	I/V
sdbw	I	I	I	I	I	I	I	I

Erklärung der Einfärbung der Tabellenfelder:
 Grau: Ausgeschlossene Kombinationen aus int. VI und Fusionierungsalgorithmus
 Grün: Geeignete Kombinationen aus int. VI und Fusionierungsalgorithmus für mehr als drei Testdatensätze
 Orange: Geeignete Kombinationen aus int. VI und Fusionierungsalgorithmus für die drei Testdatensätze II–IV
 Rot: Ungeeignete Kombinationen aus int. VI und Fusionierungsalgorithmus

Die int. VIs *hartigan*, *gap* und *sdbw* werden ausgeschlossen, da sie invariant bei allen Testdatensätzen dasselbe Ergebnis berechnen. Ebenfalls werden die int. VIs *cindex* und *ptbiserial* ausgeschlossen, da sie für keinen der Testdatensätze geeignet sind. Des Weiteren werden die Fusionierungsalgorithmen der Median-Methode und der Centroid-Methode, wegen auftretenden inversen Ästen bei den Dendrogrammen wie bei der euklidischen Distanzmetrik, für weitere Betrachtungen ausgeschlossen. Es existiert keine geeignete Kombination aus Fusionierungsalgorithmus und int. VI für Testdatensatz IV. Es werden alle Kombinationen aus Fusionierungsalgorithmen und int. VIs daher als nicht geeignet bewertet und die Tabellenfelder rot eingefärbt.

4.3.2 Testdatensätze VI–VIII

Übersicht der Ergebnisse der Testdatensätze VI–VIII

Tabelle 24 enthält eine Übersicht der Ergebnisse der Clusteranalyse für die Testdatensätze VI–VIII. Die Abkürzung *EUCL* wird für die euklidische Distanzmetrik verwendet und die Abkürzungen *DTW* und im Anhang *DTWARP* werden für die dynamische Zeitnormierung verwendet. Die Tabelle enthält Referenzen zu allen Ergebnissen der Clusteranalyse. In diesem Abschnitt werden vereinzelt Dendrogramme zur Veranschaulichung verwendet.

Tabelle 24 – Übersicht der Ergebnisse der Testdatensätze VI–VIII

Testdatensatz	Distanzmetrik	Spezifikation	Ergebnisse
Testdatensatz VI	EUCL	Dendrogramm	Abbildung A3.11
Testdatensatz VI	DTW	Dendrogramm	Abbildung A3.12
Testdatensatz VII	EUCL	Dendrogramm	Abbildung A3.13
Testdatensatz VII	DTW	Dendrogramm	Abbildung A3.14
Testdatensatz VIII	EUCL	Dendrogramm	Abbildung A3.15
Testdatensatz VIII	DTW	Dendrogramm	Abbildung A3.16

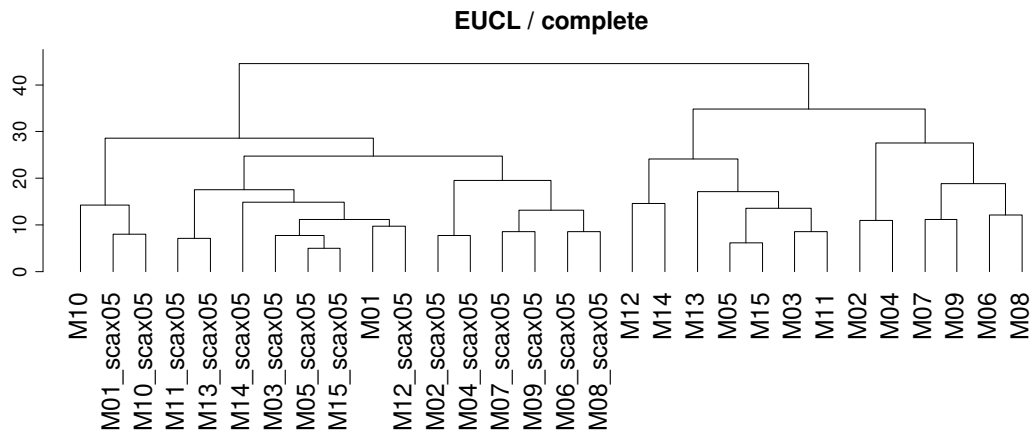
Testdatensatz VI

In einem Vergleich der Dendrogramme bei Verwendung des euklidischen Distanzmaßes wird ersichtlich, dass alle Dendrogramme hinsichtlich der Kurvenlänge strukturiert sind. Wenn z.B. die Dendrogramm-Struktur bei Verwendung des Complete-Linkage-Fusionierungsalgorithmus und der euklidischen Distanzmetrik betrachtet wird, weist die Dendrogramm-Struktur zwei Haupt-Äste auf (siehe Abbildung 51a). Ein Haupt-Ast des Dendrogramms enthält überwiegend die skalierten, transformierten Kurven und der zweite Haupt-Ast enthält vorwiegend die Kurven mit der originalen Länge (vgl. Abbildung 37). Bei jedem Fusionierungsalgorithmus werden mindestens zwei Kurven mit originaler Länge zu den skalierten Kurven gruppiert, da sie bezüglich der Länge besser zu den skalierten Kurven passen.

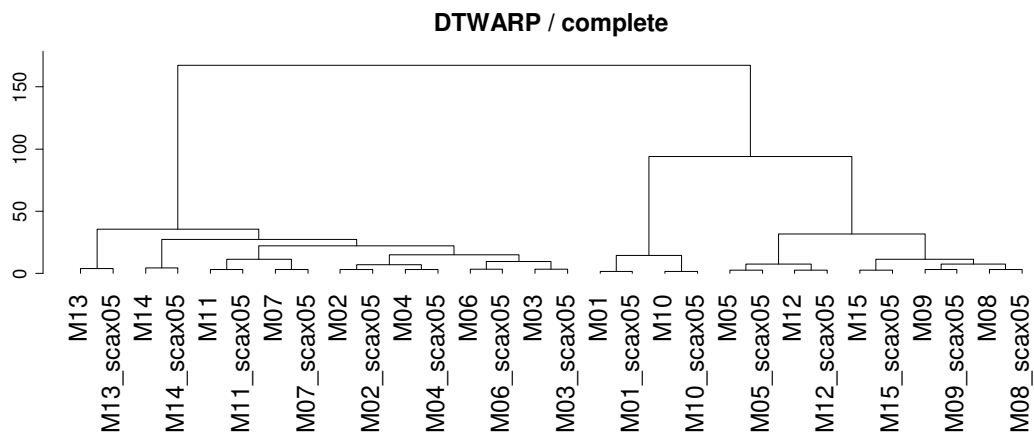
In einer vergleichenden Betrachtung der Dendrogramme bei Verwendung der dynamischen Zeitverzerrung als Distanzmetrik fällt auf, dass jeweils die skalierte Kurve und ihre korrespondierende originale Kurve in der detailliertesten Ebene des Dendrogramms zusammen gruppiert werden. Dies gilt für alle Fusionierungsalgorithmen gleichermaßen (siehe Abbildung 51b). Bei Verwendung der dynamischen Zeitnormierung wird die skalierte Kurvenform erkannt und mit der originalen Kurvenform zusammengruppiert.

Für beide Distanzmaße werden die Dendrogramme, die mit dem Median-Fusionierungsalgorithmus oder dem Centroid-Fusionierungsalgorithmus erzeugt werden, wegen auftretender inverser Äste von der Betrachtung ausgeschlossen.

Jeder Fusionierungsalgorithmus in Kombination mit der dynamischen Zeitverzerrung ist in der Lage, die ähnliche Kurvenform und somit den systematischen Fehler zu erkennen.



a) Dendrogramm bei Verwendung des Complete-Linkage-Algorithmus mit dem euklidischen Distanzmaß



b) Dendrogramm bei Verwendung des Complete-Linkage-Algorithmus mit der dynamischen Zeitverzerrung als Distanzmetrik

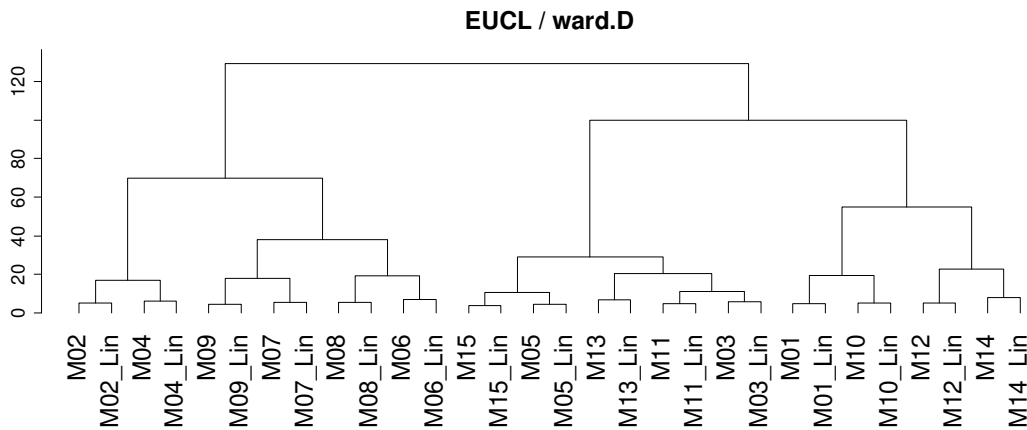
Abbildung 51 – Vergleich der Dendrogramm-Struktur bei Verwendung beider Distanzmetriken für Testdatensatz VI

Testdatensatz VII

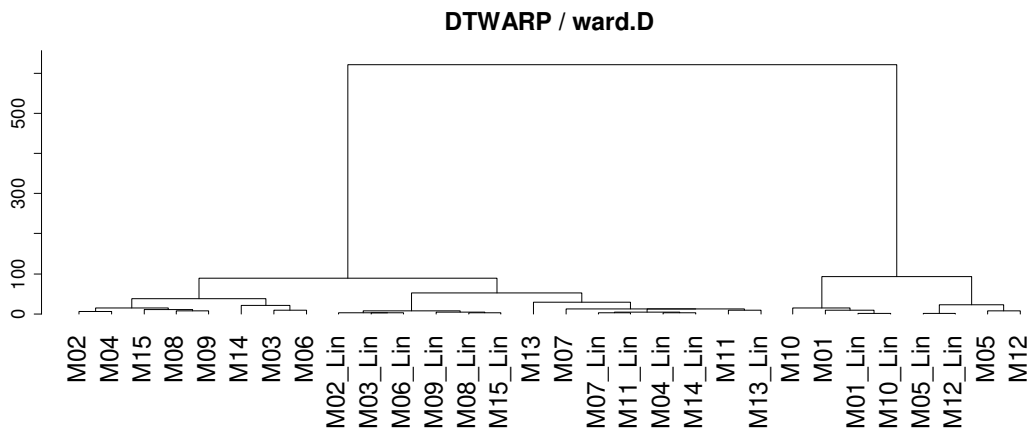
Bei Verwendung der euklidischen Distanzmetrik weisen alle Dendrogramme dieselbe Struktur auf, welche in Abbildung 52a beispielhaft bei Verwendung des Ward-Fusionierungsalgorithmus dargestellt ist. In der detailliertesten Ebene des Dendrogramms werden jeweils die experimentelle Kurve und die lineare Kurve zusammengruppiert, welche denselben Start- und Endpunkt aufweisen (vgl. Abbildung 38).

Bei Verwendung der dynamischen Zeitverzerrung als Distanzmetrik weisen alle Dendrogramme für die einzelnen Fusionierungsalgorithmen eine blockartige Struktur auf, wie sie beispielhaft für den Ward-Fusionierungsalgorithmus in Abbildung 52b dargestellt wird. Es existieren überwiegend Gruppen, welche jeweils nur lineare Kurven oder nur experimentelle Kurven enthalten. Es existiert jedoch auch eine falsche Zuordnung in der untersten Ebene des Dendrogramms, bei der die Kurve *M11* mit der linearisierten Kurve *M13 Lin.* zusammengruppiert wird.

Bei Verwendung der dynamischen Zeitnormierung als Distanzmetrik wird überwiegend nach der Kurvenform gruppiert. Eine eindeutige Trennung zwischen den linearisierten Kurven und den originalen Kurven, z.B. in zwei Haupt-Äste, ähnlich dem Ergebnis für Testdatensatz VI bei Verwendung des euklidischen Distanzmaßes, findet nicht statt.



a) Dendrogramm bei Verwendung des Ward-Algorithmus mit der euklidischen Distanzmetrik



b) Dendrogramm bei Verwendung des Ward-Algorithmus mit der dynamischen Zeitverzerrung als Distanzmetrik

Abbildung 52 – Vergleich der Dendrogramm Struktur bei Verwendung beider Distanzmetriken für Testdatensatz VII

Testdatensatz VIII

Die Strukturen der Dendrogramme für die einzelnen Fusionierungsalgorithmen bei Verwendung der euklidischen Distanzmetrik sind vergleichbar. In Abbildung 53a wird die Dendrogramm-Struktur beispielhaft bei Verwendung des Ward-Fusionierungsalgorithmus dargestellt. In der detailliertesten Ebene des Dendrogramms wird die jeweilige subjekt-spezifische Kurve mit der korrespondierenden experimentellen Kurve zusammengruppiert.

Wenn die dynamische Zeitnormierung als Distanzmetrik verwendet wird, werden nicht alle korrespondierenden Kurvenpaare zusammengruppiert (siehe Abbildung 53b). Bei Verwendung der dynamischen Zeitnormierung wird auch bei diesem Datensatz, wie erwartet, nach der Kurvenform gruppiert. Die subjekt-spezifischen Kurvenpaare, bei denen die numerische Modellvorhersage eine sehr gute Qualität erreicht, werden in der untersten Ebene, wie bei der euklidischen Distanzmetrik, zusammengruppiert. Beispiele hierfür sind *S07* und *E07*, *S01* und *E01*. Es existieren jedoch einige nicht subjekt-spezifisch zusammengruppierte Paare. Dies bedeutet, dass die subjekt-spezifische numerische Modellqualität nicht detailliert genug die experimentelle Kurve reproduzieren kann.

Die Dendrogramme des Median-Fusionierungsalgorithmus und des Centroid-Fusionierungsalgorithmus werden aufgrund inverser Äste von der Betrachtung für beide Distanzmaße ausgeschlossen.

Es eignen sich daher für diesen Datensatz die restlichen Fusionierungsalgorithmen in Kombination mit beiden Distanzmaßen. Bei Verwendung der euklidischen Distanzmetrik wird vor allem nach der

Kurvenlänge gruppiert, wohingegen die dynamische Zeitverzerrung die Güte der numerischen Modellvorhersagen hinsichtlich ihrer Kurvenformen berücksichtigt.

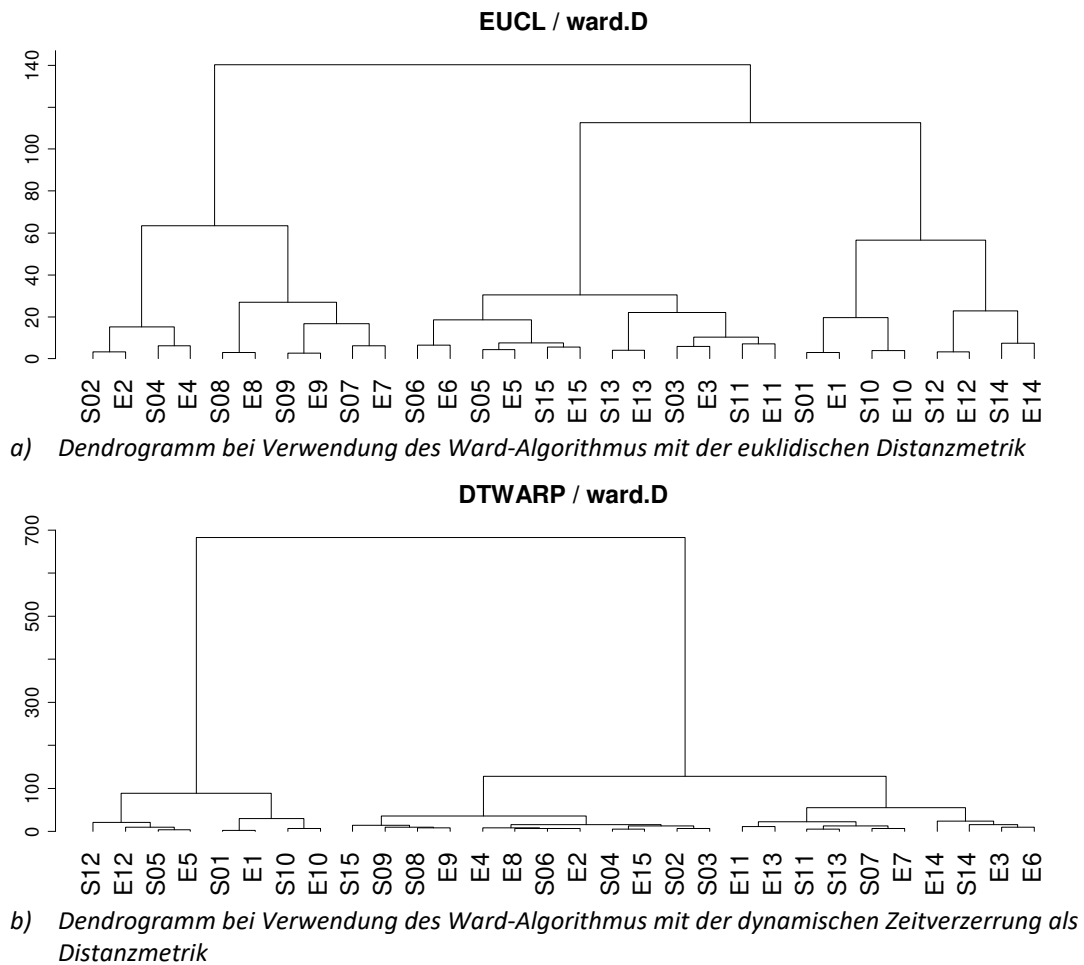


Abbildung 53 – Vergleich der Dendrogramme für Testdatensatz VIII bei Verwendung beider Distanzmetriken

4.4 Diskussion

4.4.1 Methodik

Notwendigkeit der Eignungsprüfung für biomechanische Datensätze

In Abschnitt 4.1 wurde bereits erläutert, dass Clusterverfahren in den unterschiedlichsten wissenschaftlichen Disziplinen angewendet werden. Auch in der Biomechanik wurden Clusterverfahren bereits von Simeonidis et al. [99, 100] auf kinematische Trajektorien angewendet. Bei der Entwicklung des TSclust-Pakets [80], welches in diesem Kapitel für die Clusteranalysen verwendet wird, wurden die Algorithmen anhand mehreren Testdatensätzen mit unterschiedlichen Kurvenformen verifiziert (vgl. Abbildung 28). Nachfolgend wird erläutert, weshalb es dennoch notwendig ist, eine Strategie zur Eignungsprüfung von Clusterverfahren für die Analyse von biomechanischen Datensätzen zu entwickeln, obwohl die Clusterverfahren des TSclust-Pakets bereits erfolgreich an Testdatensätzen getestet wurden.

Es existieren keine allgemein gültigen Empfehlungen, nach welchen Kriterien ein geeignetes Clusterverfahren für einen bestimmten Datensatz ausgewählt werden kann. Ebenfalls existieren keine

allgemein gültigen Empfehlungen in Bezug auf die Auswahl einer geeigneten Möglichkeit zur Bestimmung der Clusteranzahl.

In Simeonidis et al. [99, 100] wurde z.B. die Ward-Methode erfolgreich an einem biomechanischen Datensatz bestehend aus Kinematik-Trajektorien angewandt. Simeonidis et al. begründeten jedoch nicht die Auswahl des Fusionierungsalgorithmus. Zudem wich die bearbeitete Fragestellung in Simeonidis et al. ab von der in dieser Arbeit behandelten Fragestellung. Simeonidis et al. führten Clusteranalysen an einem Datensatz durch, welcher die Kinematik-Trajektorien von *bewussten* und *unbewussten* Bremsmanövern beinhaltete. Diese unterscheiden sich in der Kurvenform der Trajektorie. Dieser Unterschied wurde durch die Clusteranalyse bestätigt.

In dieser Arbeit wird das Ziel verfolgt, in Kapitel 5 einen Konzeptvorschlag für eine populationsbasierte Validierungsmethodik auf Basis von Clusteranalyseverfahren zu entwickeln. Dabei wird ein Datensatz, welcher die Modellvorhersagen einer numerischen Population und die experimentellen Kurven enthält, für den Zweck einer numerischen Modellvalidierung mit Clusterverfahren analysiert. Da die Clusteranalyse für Validierungszwecke verwendet werden soll, ist zu erwarten, dass die Kurvenformen der Individuen des Datensatzes ähnlich sind. Die Verifizierung der Algorithmen des *TSclust*-Pakets erfolgte jedoch mit Testdatensätzen, in denen die Kurvenformen sich deutlich unterscheiden [80]. Zudem enthalten diese Testdatensätze immer Kurven gleicher Länge. Eine charakteristische Eigenschaft einer experimentellen Kurvenschar eines dynamischen Validierungsexperimentes ist jedoch der unterschiedliche Versagenszeitpunkt der jeweiligen Individuen. Bei derartigen biomechanischen Datensätzen ist die Kurvenlänge bis zum Versagenszeitpunkt ein charakteristisches Maß für die Belastbarkeit einer Komponente. Bei der Anwendung von Clusterverfahren muss daher die Kurvenlänge korrekt interpretiert werden.

Es existieren Studien von Milligan [78] et al. und Arbelaitz et al. [4], welche die Eignung von verschiedenen Kombinationen von Fusionierungsalgorithmen und int. VIs an Testdatensätzen bestehend aus Punktwolken untersuchten. Diese Untersuchungen erfolgten jedoch nicht für mathematische Kurvenscharen mit biomechanischen Charakteristiken.

Aus diesen Gründen liegt der Fokus dieses Kapitels auf der Entwicklung eines Konzeptes für die Eignungsprüfung von Clusterverfahren für die Analyse von (verletzungs-)biomechanischen Datensätzen. Dieses Konzept basiert auf der Verwendung von idealisierten Testdatensätzen, um die möglichen Eigenschaften von biomechanischen Datensätzen möglichst isoliert voneinander analysieren zu können. Es können keine realen biomechanischen Datensätze für die Eignungsprüfung verwendet werden, da reale biomechanische Datensätze zu komplex sind, um einzelne Charakteristiken isoliert bewerten zu können (vgl. Kurven *M01* bis *M15* in Abbildung 37).

Biomechanische Relevanz der Testdatensätze

In Abschnitt 4.2.2 werden charakteristische Eigenschaften definiert, welche in biomechanischen Datensätzen vorhanden sein können. Auf Basis dieser charakteristischen Eigenschaften werden die Testdatensätze I–V definiert, welche in Abschnitt 4.2.3 vorgestellt werden. Die Testdatensätze VI–VIII werden definiert und erstellt, um Sonderfälle zu untersuchen. Die Erkenntnisse dieser Testdatensätze fließen in die Entwicklung des Validierungskonzeptes in Kapitel 5 mit ein.

Im Folgenden werden die Testdatensätze hinsichtlich ihrer biomechanischen Relevanz diskutiert. Die Testdatensätze I–V müssen als stark idealisierte Datensätze bewertet werden. Diese Datensätze repräsentieren aufgrund der starken Vereinfachung nicht die Realität. Die Verwendung der Datensätze wird jedoch als gerechtfertigt angesehen, um isoliert bestimmte Eigenschaften zu untersuchen, welche in realen biomechanischen Datensätzen vorkommen können. Die Testdatensätze I–V sind vor allem

hinsichtlich folgender Eigenschaften als idealisiert zu bewerten: Kurvenform, Kurvenlänge und Kurvendichte.

Die Testdatensätze I–V werden auf Basis der Kurvenform einer experimentellen Kurve von Forman et al. [38, 85] erstellt, um eine biomechanisch relevante Kurvenform zu untersuchen. Wenn jedoch eine Kurvenschar einer experimentellen Studie betrachtet wird, wie z.B. die Kurvenschar *M01* bis *M15* der Studie von Forman et al. in Abbildung 37, weisen die individuellen Kurven eines Datensatzes in der Realität eine ähnliche Grundform auf, jedoch stets mit einer überlagerten individuellen Varianz. Diese individuelle Varianz existiert nicht nur bezüglich der Kurvenform, sondern auch der Kurvenlänge. Die Existenz von nur ein oder zwei verschiedenen Kurvenlängen in den Testdatensätzen I–V ist daher idealisiert. In der Realität ist es sehr unwahrscheinlich, dass in einem Validierungsexperiment zwei Testproben getestet werden, welche exakt dieselbe Kurvenlänge aufweisen (vgl. Abbildung 37). Aus biomechanischer Sicht symbolisieren unterschiedliche Kurvenlängen eine unterschiedliche Energieaufnahmekapazität bis zur Fraktur. Diese Energieaufnahmekapazität ist eine charakteristische Eigenschaft jedes Knochens. Daher muss bei einer Clusteranalyse vermieden werden, dass Kurven mit stark unterschiedlicher Kurvenlänge in einem Cluster zusammengruppiert werden. Dieser Aspekt wird mit den Testdatensätzen II–IV betrachtet.

Ein weiteres Kriterium für die Testdatensatzgeneration ist, dass eine Clusteranalyse neben den Frakturzeitpunkten, oder Kurvenlängen, auch die Höhe der Frakturkräfte berücksichtigt. Bei experimentellen Kurven, welche nahezu dieselbe Kurvenlänge, jedoch Unterschiede in den Frakturkräften aufweisen, existieren ebenfalls Unterschiede in der Energieaufnahmekapazität, welche in der Clusteranalyse berücksichtigt werden muss. Dieser Aspekt wird vor allem in Testdatensatz IV betrachtet.

Die Kurvendichte wird in Testdatensatz V gezielt variiert. Es kann als sehr wahrscheinlich angenommen werden, dass reale biomechanische Datensätze entweder natürliche Gruppierungen oder Dichteunterschiede enthalten. Diese Dichteunterschiede können vorhanden sein, wenn z.B. eine numerische Kurvenschar mittels ihrer Charakteristik einen Konzeptraum absteckt und in bestimmten Bereichen dieses Konzeptraumes zusätzlich experimentelle Kurven fallen. Diese Überlagerung kann zu Dichteunterschieden im gesamten Datensatz führen.

In realen biomechanischen Datensätzen kann bei der visuellen Betrachtung der Kurven oftmals keine Aussage über vorhandene Dichteunterschiede oder natürliche Gruppierungen, aufgrund der Undurchsichtigkeit des Datensatzes resultierend aus der individuellen Varianz, getroffen werden (vgl. Abbildung 37). Deshalb werden in Tabelle 13 diese beiden Eigenschaften bei den realitätsnäheren Testdatensätzen VI–VIII mit einem Fragezeichen markiert.

Die Präsenz von natürlichen Gruppierungen wird in den Testdatensätzen II–III ebenfalls betrachtet. In Testdatensatz II ist eine natürliche Gruppierung aufgrund der unterschiedlichen Länge jeder Kurvenschar und einem erweiterten Kurvenabstand zwischen den Kurvenscharen realisiert. In Testdatensatz III überlappen sich die Kurvenscharen und provozieren eine erhöhte Kurvendichte im Überlappungsbereich. Dennoch existiert aufgrund der unterschiedlichen Kurvenlänge jeder Kurvenschar eine natürliche Gruppierung.

Der Testdatensatz I ist ein vollständig homogener Datensatz, welcher eine gleichmäßige Durchmischung zwischen einer experimentellen und einer numerischen Kurvenschar darstellt (vgl. Abbildung 36a). Diesem Datensatz kommen zwei Bedeutungen zu. Zum einen stellt dieser Datensatz den extremsten Anwendungsfall für eine Clusteranalyse dar. Eine Clusteranalyse verfolgt das Ziel, Unterschiede in einem Datensatz zu finden. Dieser Datensatz ist jedoch perfekt gleichmäßig. Mit diesem Datensatz wird untersucht, welches Ergebnis eine Clusteranalyse produziert, wenn sie außerhalb ihres vorhergesehenen Anwendungsbereichs angewendet wird, womit bei numerischen Populationssimulationen zumindest näherungsweise gerechnet werden muss.

Bei der Definition der möglichen charakteristischen Eigenschaften, welche in diesem Kapitel untersucht werden, wird kein Anspruch auf Vollständigkeit erhoben. Der Fokus dieses Kapitels liegt auf der Entwicklung eines Konzeptes für die Überprüfung der Eignung von Clusteranalyseverfahren zur Anwendung an biomechanischen Datensätzen. Eine starke Limitation dieser Studie besteht darin, dass nur Datensätze einer Kurvenform untersucht werden. Für die Prüfung des entwickelten Konzeptes ist dies ausreichend. Das Konzept wird insgesamt mit acht Datensätzen überprüft. Es wäre jedoch eine zukünftige Studie empfehlenswert, welche die existierenden Kurvenform-Typen und charakteristischen Eigenschaften aus experimentellen Studien, welche für die Menschmodell-Validierung verwendet werden, bestimmt. Eine Ausgangsbasis für eine solche Studie sind z.B. die dargestellten Validierungslastfälle in Anhang A1.

Mit diesen Erkenntnissen könnte eine systematische Studie durchgeführt werden, in der Testdatensätze für die verschiedenen Kurvenform-Typen und charakteristischen Eigenschaften generiert werden und für die Eignungsprüfung der Clusteranalyseverfahren verwendet werden. Eine solche Studie ist notwendig, um Aussagen über den Gültigkeitsbereich der Erkenntnisse, welche in dieser Studie gewonnen werden, auch für weitere Kurvenformen, wie z.B. Kinematik-Trajektorien, zu ermöglichen.

Die Testdatensätze VI–VIII stellen Sonderfälle dar, um systematische Fehler zu untersuchen. Diese werden im Folgenden kurz erläutert. Diese Testdatensätze dienen vor allem zur Analyse der Vor- und Nachteile bei der Verwendung der jeweiligen Distanzmetrik.

Testdatensatz VI enthält zwei ursprünglich identische Kurvenscharen, wovon eine Kurvenschar bezüglich ihrer zeitlichen Achse gestaucht wird. Testdatensatz VI beschreibt einen systematischen Fehler, der beispielsweise durch einen Fehler im numerischen Modell oder einem Fehler in der Kurvenvorbereitung der Clusteranalyse entstehen kann. Bei diesem Datensatz wird überprüft, ob ein derartiger systematischer Fehler erkannt werden kann.

In Testdatensatz VII werden zwei grundsätzlich unterschiedliche Kurvenform-Typen, wellig vs. linear, untersucht. Der Kurvenform-Typ ist v.a. abhängig vom gewählten experimentellen Versuchsaufbau und den Randbedingungen. Wenn z.B. numerische Modellvorhersagen in ihrer Kurvenform sich deutlich von den experimentellen Kurvenformen unterscheiden, kann davon ausgegangen werden, dass ein Fehler im numerischen Modell existiert und die Mechanik des Experiments unzureichend abgebildet wird. Dieser Testdatensatz dient daher der Überprüfung, ob Unterschiede in der Kurvenform von den Clusterverfahren erkannt werden.

Testdatensatz VIII enthält gegensätzlich zu Testdatensatz VII zu jeder experimentellen Kurve eine subjekt-spezifische numerische Modellvorhersage. Subjekt-spezifische numerische Modellvorhersagen sollten immer die ähnlichste Repräsentation eines experimentellen Ergebnisses sein. In diesem Fall sollten Clusterverfahren die korrespondierenden Kurvenpaare zusammengruppieren.

Verwendete Distanzmaße

In dieser Arbeit werden das euklidische Distanzmaß und das Distanzmaß der dynamischen Zeitverzerrung verwendet. Beide Distanzmaße werden in Abschnitt 4.1.4 detailliert beschrieben.

Es werden zwei Distanzmaße untersucht, um das entwickelte Konzept zur Eignungsprüfung in größerem Umfang zu testen. Es werden bewusst zwei möglichst unterschiedliche Distanzmaße verwendet, um die Vorzüge und Nachteile der Charakteristiken jedes Distanzmaßes zu beleuchten [99]. Das euklidische Distanzmaß wird verwendet, da es von Simeonidis et al. [99, 100] bereits an biomechanischen Datensätzen erfolgreich angewendet wurde. Diese Distanzmetrik weist besondere Sensitivität bezüglich Signaltransformationen, wie Streckung oder Stauchung der x-Achse auf [80]. Die

dynamische Zeitverzerrung ermöglicht das Erkennen ähnlicher Kurvenformen, auch wenn diese Signaltransformationen, wie Streckung oder Stauchung der x-Achse unterliegen [80].

In der Literatur wurden zahlreiche weitere Distanzmaße veröffentlicht. Aghabozorgi et al. [2] veröffentlichten eine Zusammenfassung der in der Literatur verwendeten Distanzmaße und ihre Eigenschaften mit literarischen Referenzen. Das TSclust-Paket, welches in dieser Arbeit verwendet wird, beinhaltet bereits über 30 Distanzmetriken [80].

Laut Montero et al. [80] wird nicht empfohlen, eine Studie mit allen Distanzmaßen durchzuführen, um das Distanzmaß auszuwählen, welches das beste Clusterergebnis liefert. Es wird vielmehr empfohlen, ein geeignetes Distanzmaß nach dem spezifischen Ziel der Clusteranalyse auszuwählen [80]. Es wird in der Literatur jedoch keine eindeutige Empfehlung gegeben, für welches Problem welches Distanzmaß geeignet ist. Daher wird die Vorgehensweise der Untersuchung verschiedener Distanzmaße anhand von spezifischen Testdatensätzen als geeignet angesehen.

Generell kann zwischen form-basierten und struktur-basierten Distanzmaßen unterschieden werden [80]. Form-basierte Distanzmaße arbeiten überwiegend mit lokalen Vergleichen, wohingegen struktur-basierte Distanzmaße den globalen Verlauf vergleichen [80]. Abbildung 54 visualisiert die unterschiedliche Vorgehensweise beider Kategorien. Bei einer form-basierten Clusteranalyse werden die beiden Datensätze *P1* und *P2* aufgrund ihrer geringeren Distanz zueinander zusammen gruppiert, wohingegen bei einer struktur-basierten Clusteranalyse die Datensätze *P1* und *P3* aufgrund ihrer ähnlichen Kurvenform zusammen gruppiert werden.

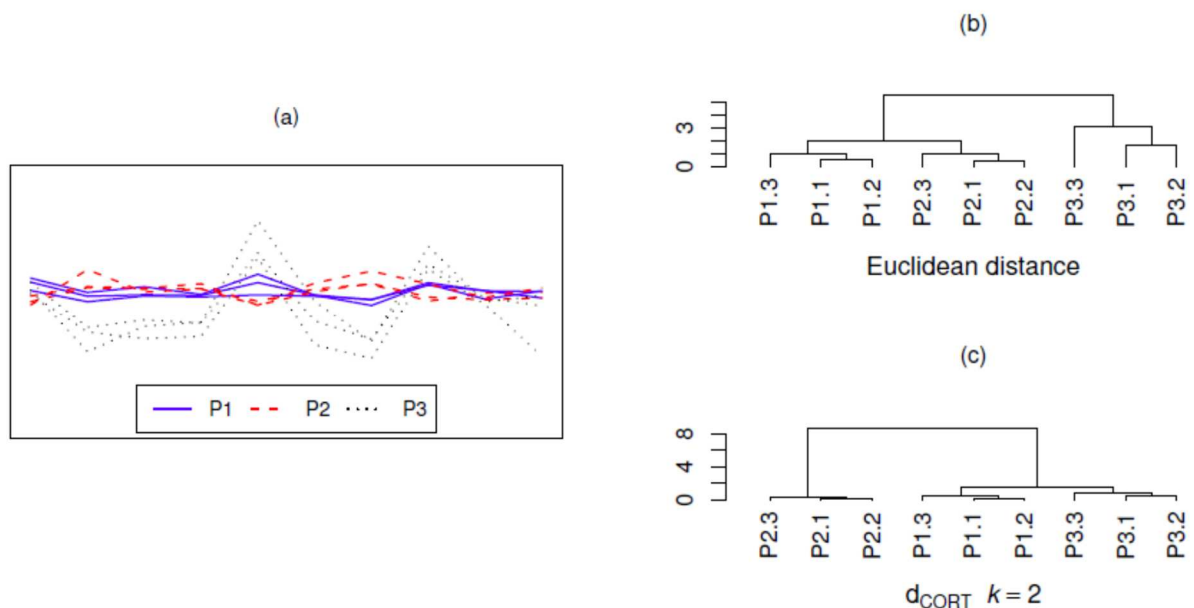


Abbildung 54 – (a) Visualisierung der Datensätze, (b) form-basierte Clusteranalyse, (c) struktur-basierte Clusteranalyse [80]

Die beiden verwendeten Distanzmaße in dieser Arbeit gehören zu den form-basierten Distanzmaßen. In Zukunft ist es empfehlenswert, weitere Distanzmaße auf ihre Eignung zu überprüfen. Zukünftig könnte z.B. die Fréchet-Distanz untersucht werden, da sie die Möglichkeit bietet Kurven unterschiedlicher Länge zu untersuchen [80]. Des Weiteren könnten Distanzmaße, welche der Kategorie der struktur-basierten Distanzmaße angehören, untersucht werden.

Verwendete Fusionierungsalgorithmen

Alle verwendeten Fusionierungsalgorithmen, welche in Abschnitt 4.1.5 kurz vorgestellt werden, sind in dem *NbClust*-Paket von Charrad et al. [20] in der R-Studio-Software implementiert. In dieser Arbeit

werden Fusionierungsalgorithmen der hierarchisch agglomerativen Kategorie untersucht, da eines dieser Verfahren, der Ward-Fusionierungsalgorithmus, bereits in Simeonidis et al. [99, 100] erfolgreich an biomechanischen Datensätzen angewendet wurde. Da in Simeonidis et al. die Auswahl des Ward-Fusionierungsalgorithmus nicht begründet wurde, werden alle implementierten hierarchisch agglomerativen Fusionierungsalgorithmen in diesem Kapitel systematisch der Eignungsprüfung unterzogen. Die Eignung der einzelnen Fusionierungsalgorithmen für die Anwendung an biomechanischen Datensätzen wird in Abschnitt 4.4.2 analysiert.

Die untersuchte Anzahl von acht Fusionierungsalgorithmen ist höher im Vergleich zu anderen Studien. Milligan et al. [78] führten eine ähnliche Studie durch, bei der an verschiedenen Testdatensätzen von Punktwolken die Eignung verschiedener int. VIs überprüft wurde. Milligan et al. [78] verwendeten für diese Studie vier hierarchische Clusterverfahren, die Ward-Methode, die Single-Linkage-Methode, die Complete-Linkage-Methode und die Average-Linkage-Methode.

Arbelaitz et al. [4] führten eine ähnliche Studie wie Milligan et al. durch, welche anhand von Testdatensätzen aus Punktwolken die Eignung von int. VIs überprüfte. Arbelaitz et al. [4] verwendeten hierzu die Fusionierungsalgorithmen der K-Means-Methode, Ward-Methode und Average-Linkage-Methode.

In zukünftigen Studien wäre es ratsam, weitere Clusterverfahren einer Eignungsprüfung zu unterziehen. Es wäre empfehlenswert, Fusionierungsalgorithmen der folgenden Kategorien zu untersuchen: partitionierende Verfahren (vgl. Abbildung 30), probabilistische Verfahren, oder den speziellen Fall der überlappenden Clusteranalyse der deterministischen Verfahren (vgl. Abbildung 29) [5].

Bestimmung der Clusteranzahl

In Tabelle 12 in Abschnitt 4.1.6 werden alle verwendeten int. VIs aufgelistet. Alle verwendeten int. VIs werden im Rahmen des *NbClust*-Pakets von Charrad et al. [20] in R-Studio zu Verfügung gestellt. Dieses Paket stellt insgesamt 30 verschiedene int. VIs zu Verfügung. Im Folgenden werden die Gründe dargelegt, weshalb nur 19 int. VIs der 30 zur Verfügung gestellten int. VIs getestet werden. Zwei der zu Verfügung gestellten VIs gehören zu der Kategorie der graphischen Indizes. Das Ergebnis dieser VIs ist ein Diagramm, anhand dessen die Clusteranzahl bestimmt wird. Bei dem *DIndex*-Validierungsindex wird eine Kurve erzeugt. Der Knickpunkt bestimmt dabei die Clusteranzahl. Bei der Verwendung dieser Art von Indizes wird dem Anwender eine gewisse subjektive Entscheidungsfreiheit bei der Graphen-Interpretation zugestanden. In dieser Arbeit wird auf ein möglichst transparentes und langfristig objektives Vorgehen Wert gelegt. Deshalb wird diese Art int. VIs nicht betrachtet.

Die restlichen int. VIs des *NbClust*-Pakets werden nicht betrachtet, da sie bei Verwendung des Testdatensatzes I nicht berechnet werden können. Um einen möglichst transparenten Clusteranalyse Prozess zu erhalten, wird der Prozess automatisiert durchgeführt. Diese int. VIs werden ausgeschlossen, um den Prozess möglichst robust zu halten. Die Ergebnisse des Testdatensatzes I zeigen jedoch, dass dieser Datensatz am besten anhand eines Dendrogrammes analysiert wird. Es war daher nicht der geeignetste Datensatz, um über einen solchen Ausschluss zu entscheiden. Es wird daher empfohlen, in einer zukünftigen Studie diese int. VIs erneut zu testen.

Der Ausschluss dieser int. VIs hat jedoch weder Einfluss auf die Qualität dieser Studie noch auf die Prüfung des entwickelten Konzeptes der Eignungsprüfung. Die getestete Anzahl von 19 int. VIs stellt eine umfassende Studie dar, welche jedoch keinen Anspruch auf Vollständigkeit erhebt. Selbst bei Einbeziehen aller im Software-Paket zur Verfügung stehenden int. VIs, könnte kein Anspruch auf Vollständigkeit erhoben werden, da darüber hinaus eine sehr große Anzahl an veröffentlichten int. VIs existiert. Die Anzahl von 19 int. VIs wird jedoch als ausreichend bewertet, um die entwickelte Methodik umfassend zu testen. Die bereits in vorigem Paragraphen beschriebenen vergleichbaren Studien von

Milligan et al. [78] und Arbelaitz et al. [4] untersuchten jeweils 30 verschiedene int. VIs. Welche int. VIs sich besonders für den Einsatz an biomechanischen Datensätzen eignen, wird in Abschnitt 4.4.2 diskutiert.

Durchführung der Clusteranalyse

Es wird bewusst die Verwendung der Statistik-Software R-Studio mit den öffentlich verfügbaren Paketen für die Clusteranalyse ausgewählt, anstatt Algorithmen selbst zu implementieren. Diese verfügbaren Pakete enthalten implementierte Algorithmen, die an verschiedenen Testdatensätzen bereits verifiziert wurden. Aus diesem Grund kann in dieser Arbeit auf die Tätigkeit der Verifizierung verzichtet werden. Der Fokus kann somit auf die Eignungsprüfung der Algorithmen für die Anwendung an biomechanischen Datensätzen gelegt werden.

Die Durchführung der Clusteranalyse und die Erstellung der Dendrogramme, sowie die Berechnung der int. VIs für die Testdatensätze I–V, wird als automatisierter Prozess ausgeführt, um manuelle Fehler so weit wie möglich zu reduzieren.

Die Vorbereitung der Datensätze für die Clusteranalyse wird aus demselben Grund ebenfalls als automatisierter Prozess ausgeführt. Der gewählte Zeitschritt wird bei der Datenvorbereitung sehr klein gewählt (0,1 ms), um den Informationsverlust aufgrund der Zeitschrittgröße so weit wie möglich zu limitieren.

Bewertungsprozess der Ergebnisse der Clusteranalyse

Die Bewertung der Eignung der einzelnen Kombinationen aus Fusionierungsalgorithmen, Distanzmaßen und int. VIs für jeden der Testdatensätze I–V erfolgt auf Basis manueller Sichtung anhand der in Abschnitt 4.2.5 definierten Bewertungskriterien.

Die Bewertungskriterien für die Testdatensätze I–V können detailliert definiert werden, so dass eine automatisierte Bewertung in Zukunft möglich und empfehlenswert wäre. Eine solche automatisierte Bewertung könnte anhand externer Validierungsindizes erfolgen. Externe Validierungsindizes vergleichen ein Ergebnis der Clusteranalyse mit einer vorgegebenen, wahren Clusteraufteilung (engl. ground truth) [2]. Aghabozorgi et al. [2] geben einen Überblick über mögliche externe Validierungsindizes.

Die Bewertung der Ergebnistabellen, um die geeignetsten Kombinationen aus int. VIs und Fusionierungsalgorithmen auf Basis der Testdatensätze I–V zu bestimmen, werden im Folgenden kurz erläutert. Um eine geeignete Clusterlösung zu erhalten, muss eine Kombination aus Distanzmaß, Fusionierungsalgorithmus und int. VI zumindest für die Testdatensätze II–IV geeignet sein. Diese Anforderung resultiert daraus, dass die Testdatensätze II–IV u.a. die Eigenschaft unterschiedlicher Kurvenlängen aufweisen, welche von einem Clusterverfahren korrekt interpretiert werden muss. Kurven unterschiedlicher Länge sollten nicht zusammen in einem Cluster vereint werden, da die Kurvenlänge in Zusammenhang mit der Belastbarkeit des Knochens steht.

Zudem werden diejenigen Clusterverfahren ausgeschlossen, welche Dendrogramme mit inversen Ästen produzieren, wie beispielsweise in Abbildung 55 dargestellt. Inverse Äste sind schwer interpretierbar und verhindern die Möglichkeit zur visuellen Analyse des Dendrogramms. Eine visuelle Analyse des Dendrogramms ist jedoch hilfreich, wenn beispielsweise unterschiedliche int. VIs verschiedene Clusteranzahlen für eine Kombination aus Testdatensatz, Distanzmetrik und Fusionierungsalgorithmus vorschlagen. In einem solchen Fall können die vorgeschlagenen Clusteranzahlen anhand einer visuellen Analyse des Dendrogramms bezüglich ihrer Element-Zuordnung nachvollzogen werden.

Zudem werden diejenigen int. VIs für weitere Betrachtungen ausgeschlossen, welche über alle Testdatensätze invariant dieselbe Clusteranzahl berechnen. Dies wird als Indiz dafür bewertet, dass diese int. VIs für diese Art von Datensätzen nicht geeignet sind. Diese Bewertung erfolgt auf der

Grundlage, dass jeder der Testdatensätze unterschiedliche charakteristische Eigenschaften besitzt. Es ist daher unwahrscheinlich, dass für alle Testdatensätze dieselbe Clusteranzahl berechnet wird.

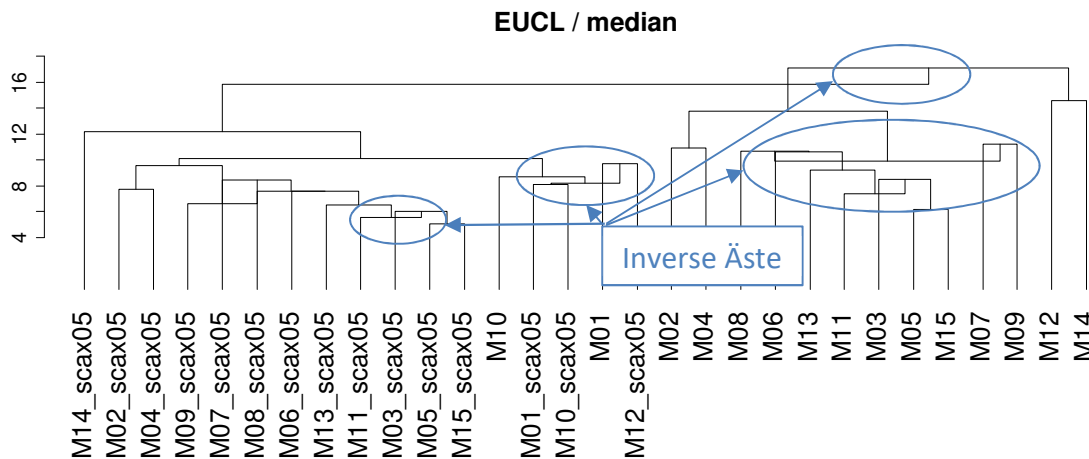


Abbildung 55 – Testdatensatz VI mit euklidischer Distanzmetrik und Median Fusionierungsalgorithmus

Die Testdatensätze VI–VIII dienen vorwiegend dem Erkenntnisgewinn bez. der Vor- und Nachteile bei Verwendung der beiden Distanzmetriken. Bei diesen Testdatensätzen ist es nicht möglich, Bewertungskriterien in dem Detaillierungsgrad festzulegen, der für eine Automatisierung des Bewertungsprozesses nötig wäre. Daher erfolgt die Bewertung auf manueller Basis. Zukünftig könnten Überlegungen dahingehend getroffen werden, die Bewertungskriterien für diese Art von Datensätzen so zu definieren, dass sie auch automatisiert bewertet werden könnten.

4.4.2 Ergebnisse der Testdatensätze

Testdatensätze I–V

Die Zusammenfassung der Ergebnisse für die Testdatensätze I–V in Tabelle 22 für die euklidische Distanzmetrik zeigt, dass sich einige Kombinationen aus Fusionierungsalgorithmen und int. VIs als geeignet oder sehr gut geeignet erweisen. Diese Kombinationen werden in Tabelle 25 nochmals zusammengefasst.

Tabelle 25 – Zusammenfassung der geeigneten int. VIs für das euklidische Distanzmaß

Bewertungsgüte	Fusionierungsalgorithmus	Abkürzung int. VI
Sehr gut geeignet	Average-Linkage	beale
Sehr gut geeignet	Mc Quitty	beale
Geeignet	Average-Linkage	ptbiserial
Geeignet	Complete-Linkage	beale
Geeignet	Complete-Linkage	ptbiserial
Geeignet	Mc Quitty	ptbiserial
Geeignet	Ward.D	ptbiserial
Geeignet	Ward.D2	Ptbiserial
Geeignet	Single-Linkage	kl
Geeignet	Single-Linkage	ch
Geeignet	Single-Linkage	beale
Geeignet	Single-Linkage	ptbiserial

Es existiert kein int. VI, welches für alle Fusionierungsalgorithmen und Datensätze geeignet ist. Diese Erkenntnis stimmt mit den Erkenntnissen der Studie von Milligan et al. [78] überein. Bei der Studie von Milligan et al. erwies sich der Index *ch* als geeignetster Index, der Index *beale* erreichte Rang 5 und der

Index *ptbiseria* erreichte Rang 7 von 30 überprüften Indizes. Die Indizes *beale* und *ptbiseria* erweisen sich auch in dieser Studie als besonders geeignet (vgl. Tabelle 25). Es wird empfohlen zukünftig weitere Untersuchungen dahingehend zu unternehmen, auch außerhalb der Kategorie der int. VIs weitere Möglichkeiten zu identifizieren, wie geeignete Clusteranzahlen nach einer Clusteranalyse bestimmt werden können, welche alle definierten charakteristischen Eigenschaften von Datensätzen erkennen. Innerhalb einer Kategorie der Bewertungsgüte *geeignet* oder *sehr gut geeignet* kann keine Aussage darüber getroffen werden, welche Kombinationen vorzugsweise verwendet werden sollten. Alle diese Kombinationen innerhalb einer jeweiligen Kategorie erweisen sich nach dieser Studie als gleich gut geeignet. Um zukünftig detailliertere Einschätzungen diesbezüglich zu erhalten, sind weiterführende Studien mit sensitiveren Eigenschaften der Testdatensätze notwendig, als in dieser Studie getestet werden. Einige in Tabelle 25 aufgeführten geeigneten Kombinationen beinhalten den Single-Linkage-Fusionierungsalgorithmus. Dieser Fusionierungsalgorithmus eignet sich zum Erkennen von *Ausreißern* in einem Datensatz [6]. Kontrahierende Fusionierungsalgorithmen, wie beispielsweise der Single-Linkage-Algorithmus, neigen dazu wenige große Cluster zu bilden, denen einige kleine Cluster oder *Ausreißer* gegenüberstehen [6]. Zudem neigt er zur Bildung von Kettenstrukturen innerhalb der großen Gruppen, weshalb *schlecht* getrennte Gruppen unerkannt bleiben (vgl. Dendrogramm des Single-Linkage-Algorithmus in Abbildung A3.11 oder Abbildung A3.13).

Aus diesen Gründen werden diese Kombinationen in Tabelle 25 grau eingefärbt. Es wird empfohlen, die anderen Kombinationen bevorzugt zu verwenden. Diese Fusionierungsalgorithmen gehören entweder der Kategorie der konservativen- oder der dilatierenden Fusionierungsalgorithmen an [6]. Dilatierende Fusionierungsalgorithmen neigen dazu Objekte zu gleich großen Clustern zusammenzufassen. Die Kategorie der konservativen Fusionierungsalgorithmen neigt zu keinerlei Tendenzen. Der Complete-Linkage-Algorithmus gehört zu der Kategorie der dilatierenden Algorithmen, wohingegen die restlichen Algorithmen der konservativen Kategorie angehören.

Alle Kombinationen mit der Bewertungsgüte *geeignet*, erweisen sich für den Testdatensatz I und den Testdatensatz V als nicht geeignet. Diese beiden Testdatensätze sind in ihrer Struktur sehr homogen und weisen dieselbe Kurvenlänge auf. Bei sehr homogenen Datensätzen, oder wenn die Eigenschaft der Homogenität bei einem Datensatz zu überprüfen ist, ist es empfehlenswert, das Dendrogramm mit dem Single-Linkage-Fusionierungsalgorithmus und der euklidischen Distanzmetrik für die Analyse des Datensatzes zu verwenden. In diesem speziellen Fall sind die Kombinationen aus Fusionierungsalgorithmen und int. VIs aus der Bewertungskategorie *geeignet* nicht anwendbar.

Im Vergleich zu Testdatensatz I besitzt Testdatensatz V Dichteunterschiede in den Kurvenabständen. Diese werden von den Kombinationen mit der Bewertungsgüte *sehr gut geeignet* richtig interpretiert. Die Zusammenfassung der Ergebnisse für die Testdatensätze I–V in Tabelle 23 für die Distanzmetrik dynamische Zeitverzerrung zeigt, dass sich keine Kombination aus Fusionierungsalgorithmen und int. VIs als geeignet erweist. Diese Bewertung erfolgt, da für Testdatensatz IV keine geeignete Kombination aus Fusionierungsalgorithmen und int. VIs existiert. Bei diesem Testdatensatz werden immer Kurven unterschiedlicher Länge zusammengruppiert. Bei biomechanischen Datensätzen ist die Kurvenlänge jedoch wichtig, da sie maßgeblich das Verhalten unter grenzwertiger Belastung widerspiegelt und ein Maß für den Energieeintrag in die Testprobe darstellt.

Testdatensätze VI–VIII

Die Analyse der Clusterergebnisse der Testdatensätze VI–VIII ergibt, dass sich das Distanzmaß der dynamischen Zeitverzerrung vor allem zum Erkennen systematischer Fehler, wie z.B. einer Streckung oder Stauchung der Kurven in x-Richtung (Zeitachse) bei Testdatensatz VI oder dem Erkennen unterschiedlicher Kurvenformen bei den Testdatensätzen VII–VIII, eignet.

Zudem eignet sich dieses Distanzmaß nicht nur zum Erkennen signifikant unterschiedlicher Kurvenformen, sondern auch zur Analyse der Güte von subjekt-spezifischen numerischen Modellvorhersagen, wie in Testdatensatz VIII analysiert wird. Die Analyse des Testdatensatzes VIII betont eine Schwachstelle des subjekt-spezifischen Validierungsansatzes, dass auch bei einer genauen numerischen Abbildung der Geometrie der Testprobe mit anschließendem Materialoptimierungsprozess die subjekt-spezifische numerische Modellvorhersage oftmals nicht mit hoher Güte reproduziert werden kann. Die Ungenauigkeit der numerischen Modellvorhersagen kann durch die Materialcharakterisierung zustande kommen, da für eine genaue numerische Materialcharakterisierung die Testprobe nach der Durchführung eines dynamischen Experiments zusätzlich Materialtests zur Materialcharakterisierung durchlaufen müsste. Diese Materialtests werden jedoch meist nicht durchgeführt.

Eine wichtige Erkenntnis der Verwendung dieser Testdatensätze ist ebenfalls, die Limitationen der euklidischen Distanzmetrik zu kennen. Das euklidische Distanzmaß ist nicht in der Lage, Kurvenformen zu erkennen. Wenn z.B. zwei unterschiedliche Kurvenformen denselben Start- und Endpunkt besitzen und die geringste euklidische Distanz zueinander aufweisen, werden diese zusammengruppiert, obwohl die Kurvenformen unterschiedlich sind. Diese Limitation dieser Distanzmetrik muss bei der Anwendung an biomechanischen Datensätzen berücksichtigt werden.

4.4.3 Zusammenfassung und Ausblick

Dieses Kapitel bearbeitet die Fragestellung, wie ein Konzept zur Eignungsprüfung von mathematischen Clusteranalyseverfahren zur Analyse und zum Vergleich von (verletzungs-)biomechanischen Datensätzen aussehen kann. Biomechanische Datensätze, wie z.B. die Ergebniskurven experimenteller Studien, besitzen oftmals die Eigenschaft, für den Nutzer sehr undurchsichtig zu sein. Sie können daher nicht auf rein visueller Basis analysiert oder verglichen werden. Aufgrund dieser Eigenschaft können Clusteranalyseverfahren nicht direkt an biomechanischen Datensätzen einer Eignungsprüfung unterzogen werden.

In diesem Kapitel wird ein Konzept erarbeitet, bei dem charakteristische Eigenschaften aus biomechanischen Datensätzen definiert werden, welche in Zukunft untersucht werden sollen. Mit diesen charakteristischen Eigenschaften werden vereinfachte Testdatensätze generiert, anhand dieser die Clusteranalyseverfahren einer Eignungsprüfung unterzogen werden. Die Eignungsprüfung überprüft, ob die definierten charakteristischen Eigenschaften von den Clusteranalyseverfahren erkannt werden. Die Kombinationen aus Distanzmetriken, Fusionierungsalgorithmen und int. VIs, welche sich für die meisten Datensätze als geeignet erweisen, werden als geeignete Kombinationen für die Analyse und den Vergleich von biomechanischen Datensätzen definiert (siehe Tabelle 25).

In zukünftigen Studien sollte überprüft werden, ob noch weitere Clusteranalyseverfahren oder Möglichkeiten zur Bestimmung einer geeigneten Clusteranzahl sich als geeignet erweisen. Zudem sollte das Konzept an weiteren Datensätzen mit unterschiedlichen Kurvenformen, wie z.B. Kinematik-Trajektorien, getestet werden.

Weiterführende Untersuchungen können auch dahingehend unternommen werden, ob andere strukturentdeckende Verfahren im Bereich der Datenanalyse für eine derartige Fragestellung in Betracht kommen, wie beispielsweise Gaussian Mixture Models oder Machine Learning.

In Kapitel 5 werden die geeigneten Kombinationen aus Distanzmetriken, Fusionierungsalgorithmen und int. VIs in einem populationsbasierten Validierungskonzept verwendet, um die Modellvorhersagen einer numerischen Population gegen experimentelle Kurven zu validieren.

5 Entwicklung einer populationsbasierten Validierungsmethodik

5.1 Einleitung und Zielsetzung

Aktuelle Entwicklungstrends in der Menschmodellentwicklung umfassen u.a. die Entwicklung von mesh-morphing Algorithmen und die Entwicklung parametrisierter Modelle. Beide Entwicklungen wurden bereits in Kapitel 1 erläutert. Diese Entwicklungstrends bilden die Basis für stochastische Simulationen, da sie eine Geometrieänderung existierender Modelle ermöglichen. Mit stochastischen Simulationen kann z.B. die Varianz einer Population mit einer hohen Anzahl an numerischen Simulationen mit unterschiedlichen Charakteristiken der Testproben untersucht werden. Eine Herausforderung bei der Anwendung stochastischer Simulationen ist es, Vertrauen in die numerischen Modellvorhersagen zu generieren.

Bei Menschmodellen findet der Vertrauensaufbau durch einen Nachweis der Biofidelität mittels Abgleich zu biomechanischen Experimenten statt. Die Menschmodelle werden mit dem Ziel kalibriert und validiert, dass sie nach dem Validierungsprozess dieselben charakteristischen Eigenschaften wie die experimentellen Testproben aufweisen (vgl. Kapitel 1 Abschnitt 1.2).

Die Anwendung von mesh-morphing Algorithmen oder parametrisierten Modellen ermöglicht es, dass neue Modelle mit Eigenschaften generiert werden können, welche von den charakteristischen Eigenschaften der experimentellen Testproben abweichen. Derartige Modelle besitzen charakteristische Eigenschaften, welche außerhalb des validierten Bereichs liegen. Dies bedeutet, dass ein erneuter Nachweis der Biofidelität stattfinden muss.

Der Prozess der Vertrauensgewinnung sollte möglichst objektiv sein und in zwei Schritten erfolgen. Im ersten Schritt muss nachgewiesen werden, dass die durchgeführten Veränderungen des numerischen Ausgangsmodells, wie z.B. Geometrieänderungen oder Veränderungen der Materialeigenschaften, plausibel sind. In einem zweiten Schritt muss nachgewiesen werden, dass das veränderte numerische Modell sich weiterhin biofidel, d.h. mechanisch genügend ähnlich zum menschlichen Körper, und glaubwürdig im Bereich des vorhergesehenen Einsatzgebietes verhält.

Im Folgenden wird ein Beispiel zur Veranschaulichung beschrieben. Die Varianz einer Femora-Population soll anhand eines dynamischen 3-Punkt-Biegeversuches untersucht werden. Bei der Erzeugung einer stochastischen Population an Femora muss zum einen nachgewiesen werden, dass die erzeugten Femora bezüglich der Geometrie- und Materialcharakterisierung plausibel sind, d.h. innerhalb der Varianzen gemessener Vergleichswerte liegen. Zu diesem Aspekt gehört auch der Nachweis, dass die Modelle eine ausreichende Güte bezüglich Geometrie- und Materialmodellierung besitzen. Zum anderen muss die Frage geklärt werden, ob die numerischen Modelle aussagekräftig bezüglich ihres Einsatzbereiches sind. In diesem Fall muss ein Nachweis erbracht werden, dass die numerischen Modelle hinsichtlich einer dynamischen Biegebelastung ein realistisches Verhalten aufweisen, d.h. sich biofidel verhalten.

Dieses Kapitel widmet sich daher der Fragestellung, wie numerische Modellvorhersagen validiert d.h. hinsichtlich ihrer Biofidelität bewertet werden können, auch wenn sie andere charakteristische Eigenschaften besitzen als die experimentellen Testproben. Es stellt sich im Hinblick auf die

Anwendbarkeit stochastischer Simulationen zudem die Frage, ob in einem Validierungsprozess eine gesamte Population validiert werden kann.

Ziel dieses Kapitels ist es, eine Validierungsmethodik zu entwickeln, welche die numerischen Modellvorhersagen einer stochastischen Population in einem Prozess validiert. Zudem muss der Aspekt berücksichtigt werden, dass diese stochastische Population auch numerische Testproben enthalten kann, die andere charakteristische Eigenschaften als die experimentellen Testproben aufweisen.

5.2 Methoden

5.2.1 Vorgehensweise in diesem Kapitel

Der Inhalt und die Vorgehensweise in diesem Kapitel ist in Abbildung 56 visualisiert. Als erstes wird in Abschnitt 5.2.2 das Konzept der populationsbasierten Validierungsmethode beschrieben. Das Ziel der nachfolgenden Abschnitte besteht darin, das entwickelte Validierungskonzept an Konzeptdatensätzen zu testen. Der Referenzlastfall, mit dem die populationsbasierte Validierungsmethode getestet wird, ist der in Kapitel 3 ausgewählte dynamische 3-Punkt-Biegeversuch von Forman et al. [38, 85].

In Abschnitt 5.2.3 werden die charakteristischen Eigenschaften für eine Population, bestehend aus Femora, definiert. Diese Definition erfolgt auf Basis abgeleiteter Varianzen aus existierenden Datenbanken. Der darauffolgende Abschnitt 5.2.4 beschreibt die Erzeugung der numerischen Femora-Population, welche in Kombination mit dem validierten numerischen Versuchsaufbau des Referenzlastfalls verwendet wird, um die numerischen Modellvorhersagen dieser Population zu generieren. Der Abschnitt 5.2.5 beschreibt die verwendeten Konzeptdatensätze für die Überprüfung des populationsbasierten Validierungskonzeptes. Die numerischen Modellvorhersagen in Kombination mit den experimentellen Ergebnissen bilden die Basis für die Definition der Konzeptdatensätze.

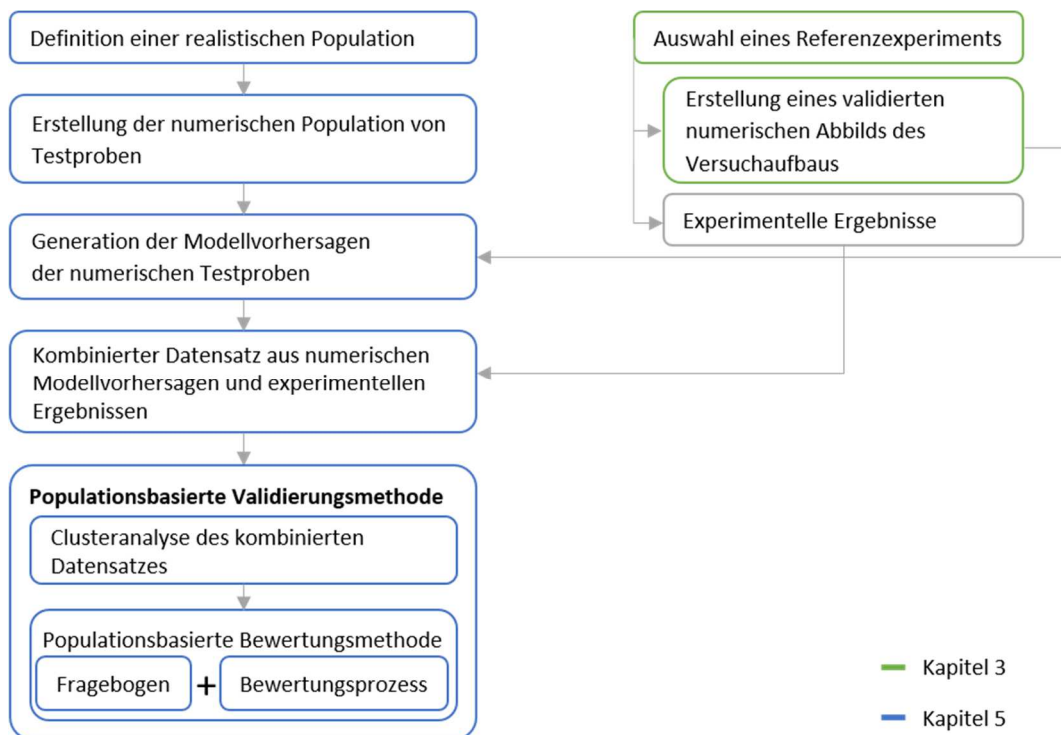


Abbildung 56 – Vorgehensweise Kapitel 5 und Zusammenhang mit Kapitel 3

5.2.2 Entwicklung der populationsbasierten Validierungsmethode

Der Begriff Validierung wurde bereits in Abschnitt 1.2.1 erläutert. Die Besonderheit bei der Validierung einer stochastischen Population ist, dass eine gesamte Kurvenschar an numerischen Modellvorhersagen gegen eine Kurvenschar an experimentellen Kurven zu vergleichen ist. Für diesen Zweck wird im Rahmen dieser Arbeit ein Konzept für eine populationsbasierte Validierungsmethode entwickelt, um die Güte und die Plausibilität der Ergebnisse stochastischer Simulationen quantitativ zu bewerten.

Der Validierungsansatz wird populationsbasierte Validierungsmethode genannt, da eine Kurvenschar an numerischen Modellvorhersagen, welche auch als numerische Population bezeichnet werden kann, gegen eine Kurvenschar an experimentellen Ergebnissen, welche auch experimentelle Population genannt werden kann, verglichen und validiert wird.

Bei der Validierung einer numerischen Kurvenschar gegen eine experimentelle Kurvenschar können die Kurven nicht, wie bei der subjekt-spezifischen Validierung, 1:1 verglichen werden. Jede Kurve der experimentellen und numerischen Kurvenschar kann unterschiedliche Werte für die Variablen der Metadaten besitzen. Mit dem Begriff Metadaten werden alle Informationen des Datensatzes, abgesehen von den Ergebnis-Kurven, zusammengefasst. Dies sind meist charakteristische Kenngrößen zur Geometrie-Definition oder Kenngrößen bezüglich der Materialcharakterisierung der betrachteten Proben.

Da ein konkreter Vergleich zwischen einzelnen Kurven nicht möglich ist, wird im Rahmen dieser Arbeit ein Ansatz basierend auf der Anwendung von Clusteranalyseverfahren verwendet. In Kapitel 4 wurden bereits die Grundlagen für diesen Ansatz erarbeitet. Es wurden Kombinationen aus hierarchisch agglomerativen Fusionierungsalgorithmen, Distanzmaßen und internen Validierungsindizes (Abkürzung: int. VIs) einer Eignungsprüfung bezüglich der Anwendbarkeit an biomechanischen Datensätzen unterzogen. Die Eignungsprüfung lieferte einige geeignete Kombinationen (vgl. Tabelle 25) und Erkenntnisse (vgl. Abschnitt 4.4.2), welche in diesem Kapitel bei der Entwicklung eines populationsbasierten Validierungsansatzes verwendet werden.

Bei dem Validierungskonzept wird zu Beginn eine Clusteranalyse an einem kombinierten Datensatz, welcher die numerischen Modellvorhersagen und die experimentellen Ergebnisse enthält, durchgeführt (vgl. Abbildung 56). Es kann eine beliebige Kombination aus Fusionierungsalgorithmus und int. VI für die Clusteranalyse und die Bestimmung der Clusteranzahl aus Tabelle 25 ausgewählt werden. Die Datensatzvorbereitung und die Durchführung der Clusteranalyse wird, wie in Abschnitt 4.2.4 beschrieben, durchgeführt. Bei der Clusteranalyse wird der kombinierte Datensatz in Gruppen, sogenannte Cluster, aufgeteilt (siehe Abbildung 57). Innerhalb eines Clusters sind diejenigen Kurven vereint, welche anhand eines gewählten Distanzmaßes die geringste Distanz zueinander aufweisen, d.h. mathematisch am ähnlichsten sind. In Abbildung 57 wird durch eine Clusteranalyse ein Datensatz beispielhaft in die drei Cluster x, y und z aufgeteilt. Jedes Cluster enthält die ihm zugeteilten Kurven und die dazugehörigen Metadaten.

Nach der Clusteranalyse erfolgt eine Plausibilitätsprüfung der numerischen Modellvorhersagen mittels eines Clustervergleichs. Es werden die Cluster untereinander hinsichtlich der Kurven und der Metadaten verglichen. Anhand der Metadaten und der zugrundeliegenden Physik des Experiments, kann eine Aussage getroffen werden, aus welchen Gründen die jeweiligen Kurven in einem Cluster zusammengefasst sind. Auf diese Weise können auch rein numerische Cluster auf ihre Plausibilität untersucht werden.

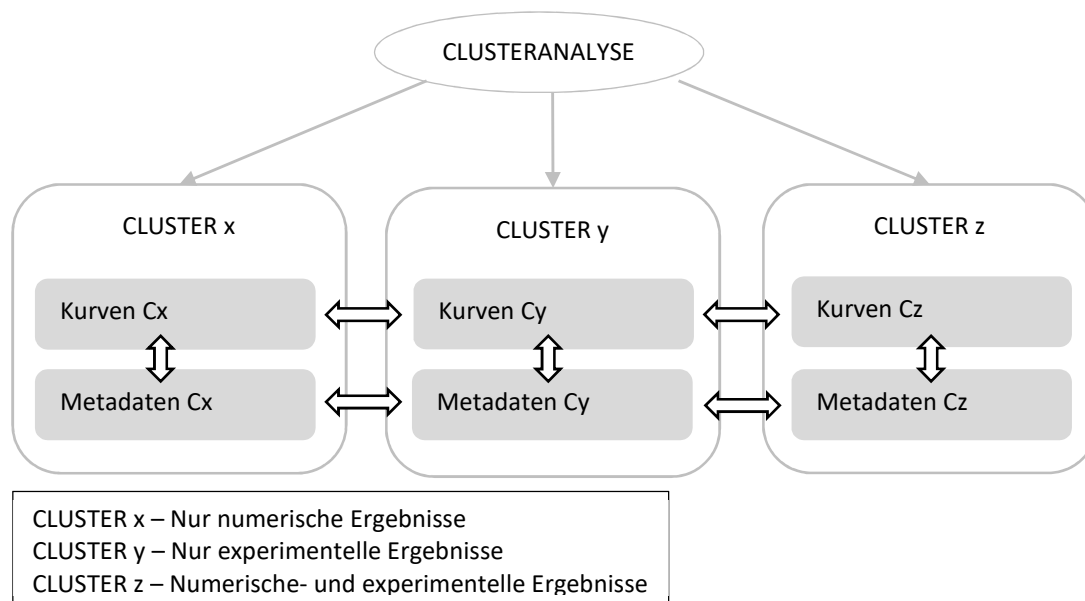


Abbildung 57 – Erkenntnisgewinn und Plausibilitätsbewertung nach Clusteranalyse

Nachstehend wird das entwickelte populationsbasierte Validierungskonzept beschrieben, welches in Abbildung 63 dargestellt ist. Aus Gründen der Übersichtlichkeit enthält der Ablaufplan in Abbildung 63 nur Fragennummern. Tabelle 26 enthält die ausführlichen Fragen, mit denen relevante Aspekte zur Beurteilung der Güte und Plausibilität von numerischen Modellvorhersagen geprüft werden. Das Validierungskonzept basiert auf der Untersuchung der Kurvenschar und der Metadaten des kombinierten Datensatzes vor und nach der Clusteranalyse. Im Folgenden werden alle Aspekte gelistet, welche zur quantitativen Beurteilung der Güte und Plausibilität der numerischen Modellvorhersagen in dem populationsbasierten Validierungskonzept berücksichtigt werden:

- Überprüfung der Charakteristik der Clusteraufteilung bezüglich der Kompaktheit der Cluster,
- Existenz von gemischten Clustern, welche numerische und experimentelle Kurven enthalten (vgl. Abbildung 57),
- Überprüfung der Homogenität eines kombinierten Datensatzes,
- Überprüfung der Plausibilität der verwendeten Metadaten, die zur Erzeugung der numerischen Testproben verwendet werden,
- Überprüfung der Plausibilität der Kurvenformen der numerischen Modellvorhersagen.

Es existieren meist mehrere Möglichkeiten zur Überprüfung dieser genannten Aspekte. Das entwickelte populationsbasierte Validierungskonzept verfolgt dabei einen ganzheitlichen Ansatz, bei dem alle verfügbaren Analysemöglichkeiten zur Untersuchung eines Aspektes kombiniert verwendet werden. Es werden nun die Analysen zur Untersuchung der einzelnen Aspekte des Validierungskonzeptes vorgestellt.

Überprüfung der Charakteristik der Clusteraufteilung

Die Charakteristik der Cluster hinsichtlich der Kompaktheit und Streuung der Kurven und Metadaten, wird mit den Fragen 16 und 17 untersucht. Die Untersuchung erfolgt auf visueller Basis. In Abbildung 58 ist beispielhaft eine Clusteranalyse mit geringer Streuung, d.h. mit kompakter Clusteraufteilung, und mit hoher Streuung dargestellt.

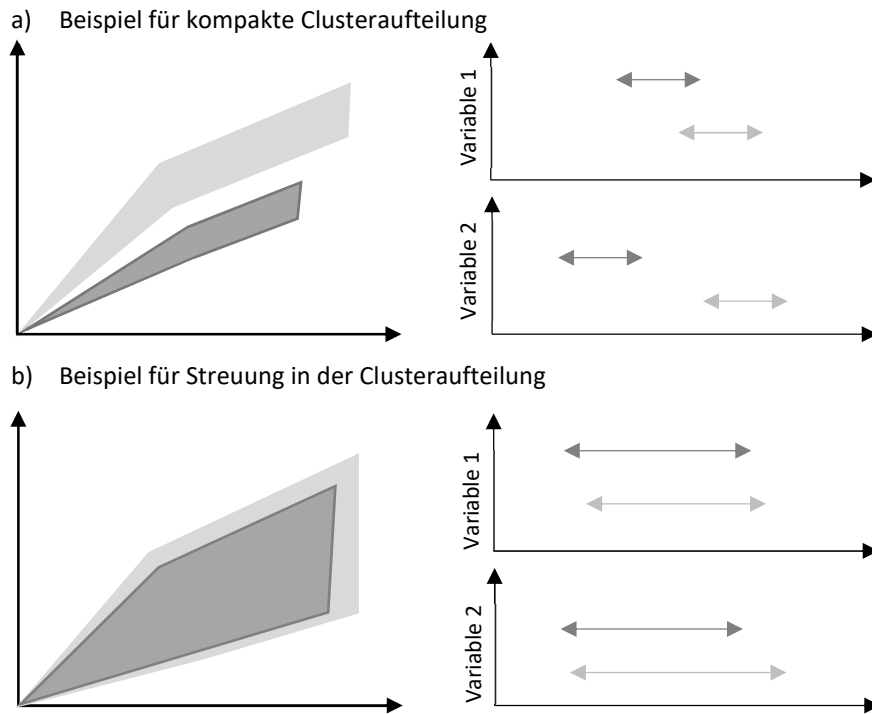


Abbildung 58 – Güte der Clusteranalyse in Relation zum Erkenntnisgewinn

Als weiterer Aspekt, welcher ebenfalls auf Überprüfung der Güte des Clusteranalyse-Ergebnisses abzielt, ist die Plausibilitätsüberprüfung der gewählten Clusteranzahl, die anhand eines int. Vis berechnet wird. Dieser Aspekt wird mit den Fragen 11 bis 15, sowie Frage 1 und Frage 7, aus verschiedenen Blickwinkeln betrachtet. Zuerst wird mit Frage 1 visuell die Kurvenschar des kombinierten Datensatzes vor der Clusteranalyse daraufhin untersucht, ob natürliche Gruppierungen vorhanden sind (vgl. Abbildung 58a). Diese Analyse wird ebenfalls in Frage 7 hinsichtlich der Metadaten durchgeführt. Es wird für jede Metadaten-Variable des kombinierten Datensatzes vor der Clusteranalyse die Gruppenanzahl visuell bestimmt. Die maximale Clusteranzahl der untersuchten Metadaten wird für weitere Betrachtungen verwendet (vgl. Abbildung 59).

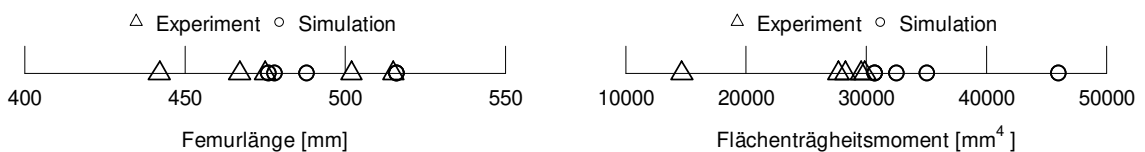


Abbildung 59 – Links: Gleichmäßige Metadaten-Verteilung, Rechts: Deutliche Gruppierung der Metadaten in drei Cluster

In Abbildung 59 ist ein Beispiel für eine nicht eindeutige Gruppierung der Metadaten (links) und ein Beispiel für eine deutliche Gruppierung der Metadaten (rechts) gegeben. Es wird empfohlen, bei dieser Untersuchung nur deutliche Gruppierungen zu verwenden. Mit Frage 14 wird basierend auf den Antworten der Fragen 1 und 7 überprüft, ob in der Kurvenschar des kombinierten Datensatzes oder in den dazugehörigen Metadaten natürliche Gruppierungen vorhanden sind. Diese ermittelten Clusteranzahlen werden mit der index-basierten Clusteranzahl in Frage 15 verglichen.

Zudem wird mit Frage 11 die Dendrogramm-Struktur des Clusteranalyse-Ergebnisses untersucht, ob die gewählte Clusteranzahl das Dendrogramm an längeren Ästen schneidet. Abbildung 60 stellt ein Positiv- und Negativbeispiel diesbezüglich dar. Bei einem akzeptablen Dendrogramm-Schnitt an längeren Ästen besitzen die einzelnen Cluster möglichst viele Objekte, wohingegen bei einer unvorteilhaften Clusteraufteilung oftmals viele Cluster mit nur sehr wenigen Objekten entstehen.

Als weitere Untersuchung wird die Clusteranzahl für den rein experimentellen Datensatz ermittelt. Für diesen Zweck werden die experimentellen Kurven einer Clusteranalyse mit derselben Kombination aus Fusionierungsalgorithmus und int. VI unterzogen, welche auch für die Clusteranalyse des kombinierten Datensatzes verwendet wird. Diese Clusteranzahl wird in Frage 12 ebenfalls mit der index-basierten Clusteranzahl des kombinierten Datensatzes verglichen. Wenn die index-basierte Clusteranzahl geringer ist als die Clusteranzahl des rein experimentellen Datensatzes, wird in Frage 13 visuell untersucht, ob experimentelle Cluster mit Hilfe von numerischen Kurven zu zusammenhängende Cluster verbunden werden.

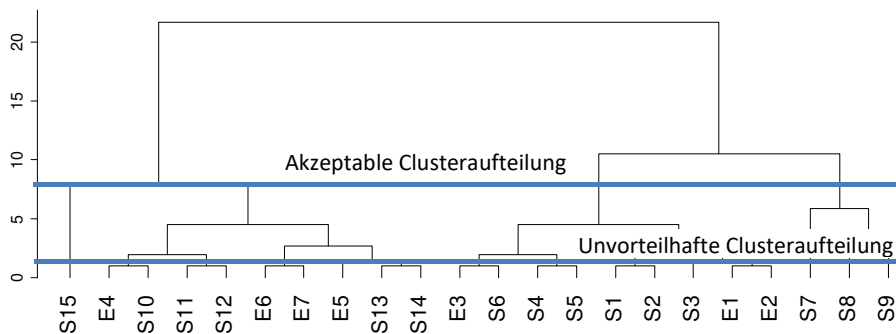


Abbildung 60 – Clusteraufteilung des Dendrogramms

Zusätzlich wird mit Frage 8 die unterste Ebene des Dendrogramms untersucht. Es werden die Verbindungen, die eine experimentelle Kurve mit einer numerischen Kurve kombinieren, visualisiert. Es wird untersucht, ob diese Kurvenpaare sich visuell signifikant unterscheiden.

Existenz gemischter Cluster

Mit den Fragen 4, 5 und 20 wird untersucht, ob gemischte Cluster existieren, welche numerische und experimentelle Ergebnisse enthalten. Mit Frage 4 wird visuell untersucht, ob die Kurvenschar eine Trennung zwischen numerischen und experimentellen Ergebnissen aufweist. Diese Trennung wird hinsichtlich der Metadaten in Frage 5 untersucht. In Frage 20 wird überprüft, ob mindestens ein Cluster existiert, welches numerische und experimentelle Kurven enthält.

Bewertung der Plausibilität der Metadaten

Die plausible Auswahl der Metadaten für die Erzeugung der numerischen Testproben wird mit Frage 6 überprüft. Die Metadaten werden gegen externe Datenbanken auf Plausibilität geprüft. Wenn die Werte der Metadaten in den Wertebereichen der Datenbanken liegen, wird die Metadaten-Auswahl als plausibel bewertet (siehe Abbildung 61).

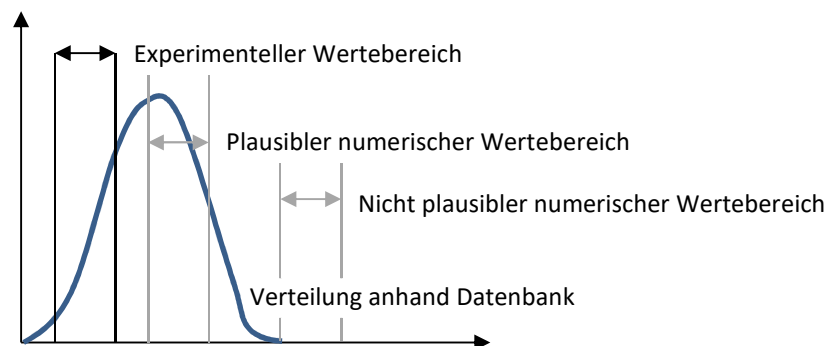


Abbildung 61 – Einschätzung der Plausibilität der Wertebereiche der Metadaten

Überprüfung der Plausibilität der Kurvenformen der Modellvorhersagen

Die Plausibilität der Kurvenformen der numerischen Modellvorhersagen wird mit den Fragen 3, 10 und 18 bis 19 überprüft. Mit Frage 3 wird untersucht, ob bei visueller Betrachtung der Kurvenschar, die Kurvenformen der numerischen Modellvorhersagen sich signifikant von den Kurvenformen der experimentellen Ergebnisse unterscheiden. In Frage 10 wird derselbe Aspekt mit einer Analyse der Dendrogramm-Struktur untersucht. Das Dendrogramm wird mit einer Clusteranalyse erzeugt, bei der die dynamische Zeitverzerrung als Distanzmaß verwendet wird. Es wird untersucht, ob die Dendrogramm-Struktur blockartige Gruppierungen zwischen experimentellen Ergebnissen und numerischen Modellvorhersagen aufweist (vgl. Testdatensatz VII in Kapitel 4, Abschnitt 4.4.2).

Mit Frage 18 und Frage 19 wird der Kompensationseffekt betrachtet. Es wird mit Frage 18 visuell untersucht, ob in einzelnen Clustern bei den Variablen der Metadaten signifikante Trennungen der Wertebereiche existieren. In diesem Fall wird mit Frage 19 überprüft, ob diese Trennungen durch einen Kompensationseffekt zustande kommen. Der Kompensationseffekt wird nachstehend kurz erläutert. Wenn beispielhaft ein Knochenquerschnitt unter Biegebelastung betrachtet wird, ist eine charakteristische Eigenschaft das Flächenträgheitsmoment, welches den Widerstand gegen Biegebelastung ausdrückt. Das Flächenträgheitsmoment wird bei einem Knochenquerschnitt maßgeblich beeinflusst durch den Innendurchmesser der kortikalen Schicht des Knochens und der kortikalen Dicke. Es gibt mehrere Möglichkeiten, wie mit einer unterschiedlichen Wahl der Metadaten Innendurchmesser und kortikale Dicke dasselbe Flächenträgheitsmoment berechnet wird. Dieser ausgleichende Effekt, der durch das Zusammenspiel mehrerer einflussreicher Variablen zustande kommt, wird im weiteren Verlauf als Kompensations-Effekt bezeichnet (vgl. Abbildung 62).

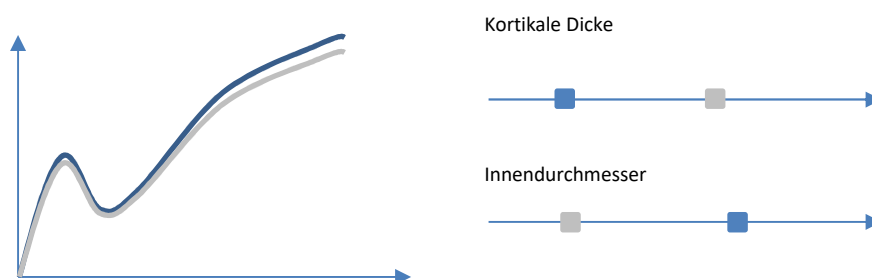


Abbildung 62 – Veranschaulichung des Kompensationseffektes bezüglich Biegesteifigkeit innerhalb eines Clusters

Analyse der Homogenität des Datensatzes

Die Homogenität eines Datensatzes wird mit Frage 2 und Frage 9 überprüft. In Frage 2 wird die Homogenität eines Datensatzes visuell anhand der Kurvenschar des kombinierten Datensatzes vor der Clusteranalyse überprüft (vgl. Testdatensatz I Kapitel 4, Abschnitt 4.4.2). In Frage 9 wird die Dendrogramm-Struktur, die durch eine Clusteranalyse mit Verwendung des Single-Linkage-Algorithmus in Kombination mit der euklidischen Distanzmetrik erstellt wurde, auf Homogenität überprüft. Abbildung 42a in Kapitel 4 visualisiert ein diesbezügliches Positivbeispiel und Abbildung 42b ein Negativbeispiel.

Quantitative Bewertung des populationsbasierten Validierungskonzeptes

Wie die einzelnen Untersuchungen bewertet werden, visualisiert der Ablaufplan in Abbildung 63. Es wird nachfolgend der Prozess beschrieben, wie die Validierungsgüte berechnet wird. Der Ablaufplan wird von Anfang bis Ende durchlaufen. Wird während des Durchlaufens des Ablaufplans auf einen roten Punkt getroffen, muss der Validierungsprozess abgebrochen werden. Wird während des Durchlaufens des Ablaufplans auf orangene Punkte getroffen, kann der Prozess weiter durchlaufen werden. Die orangenen Punkte werden während des Durchlaufens gesammelt. Jeder gesammelte orangene Punkt verringert die maximal erreichbare Güte um einen Punkt. Die maximal erreichbare Güte der Clusteranalyse beträgt 7 Punkte.

Die resultierende Punkteanzahl repräsentiert ein Maß für die Güte der numerischen Modellvorhersagen und des möglichen Erkenntnisgewinns durch den Vergleich der Cluster, welcher am Anfang dieses Abschnitts bereits beschrieben wurde. Sie kann mit der maximal möglichen Punkteanzahl als Referenz verglichen werden.

Tabelle 26 – Fragebogen für die Datensatzanalyse

#	FRAGE	Antwort	JA	NEIN
Visuelle Analyse				
1	Existieren bei visueller Betrachtung Gruppierungen in der Kurvenschar? Wenn ja, Clusteranzahl?			
2	Ist bei visueller Betrachtung die Kurvenschar homogen?			
3	Unterscheiden sich die Kurvenformen der numerischen Modellvorhersagen signifikant zu den Kurvenformen der experimentellen Ergebnisse?			
4	Existiert eine räumliche Trennung zwischen den numerischen Modellvorhersagen und den experimentellen Ergebnissen?			
Metadaten Analyse				
5	Fallen für mindestens eine Variable der numerischen Metadaten Werte in den Wertebereich der experimentellen Metadaten?			
6	Für die numerischen Werte der Metadaten außerhalb des experimentellen Wertebereichs: Wahl der Werte anhand Datenbank-Vergleich plausibel?			
7	Existieren natürliche Gruppierungen für die einzelnen Variablen der Metadaten? Wenn ja, Ermittlung der Gruppenanzahl pro Variable. Was ist die maximale Anzahl?			
Clusteranalyse				
8	Analyse der untersten Ebene des Dendrogramms. Visualisierung der Verbindungen, die eine experimentelle Kurve mit einer numerischen Kurve verbinden. Unterscheiden sich die zusammengehörigen Kurvenpaare signifikant?			
9	Überprüfung der Homogenität eines Datensatzes anhand des Dendrogramms bei Verwendung des Single-Linkage Algorithmus mit euklidischer Distanzmetrik.			
10	Überprüfung anhand eines Dendrogramms, welches mit dem Distanzmaß der dynamischen Zeitverzerrung erstellt wurde, ob sich die Kurvenformen der numerischen Modellvorhersagen signifikant von den Kurvenformen der experimentellen Ergebnisse unterscheiden.			
11	Schneidet die index-basierte Clusteranzahl das Dendrogramm an längeren Ästen?			
12	Vergleich der index-basierten Clusteranzahl mit der Clusteranzahl aus rein experimentellem Datensatz. Ist die index-basierte Clusteranzahl geringer?			
13	Wenn F12 bejaht wurde: Visuelle Überprüfung, ob die numerischen Kurven zwei oder mehr experimentelle Cluster miteinander verbinden?			
14	Wurde F1 oder F7 oder wurden beide Fragen bejaht?			
15	Wenn F14 bejaht wurde: Ist die index-basierte Clusteranzahl kleiner als die Clusteranzahl aus F1 oder F7?			
16	Visuelle Betrachtung der Kurvenscharen der einzelnen Cluster. Sind die Kurvenscharen kompakt definiert?			
17	Existiert bezüglich des Wertebereichs eine visuell starke Streuung in den Clustern für mehr als eine Variable der Metadaten?			
18	Existieren innerhalb eines Cluster bei mehreren Variablen der Metadaten signifikante Trennungen zwischen Gruppen von Wertebereichen?			
19	Wenn F18 bejaht wurde: Ist dies physikalisch durch Kompensationseffekte erklärbar?			
20	Existiert mindestens ein Cluster, welches numerische- und experimentelle Kurven enthält?			

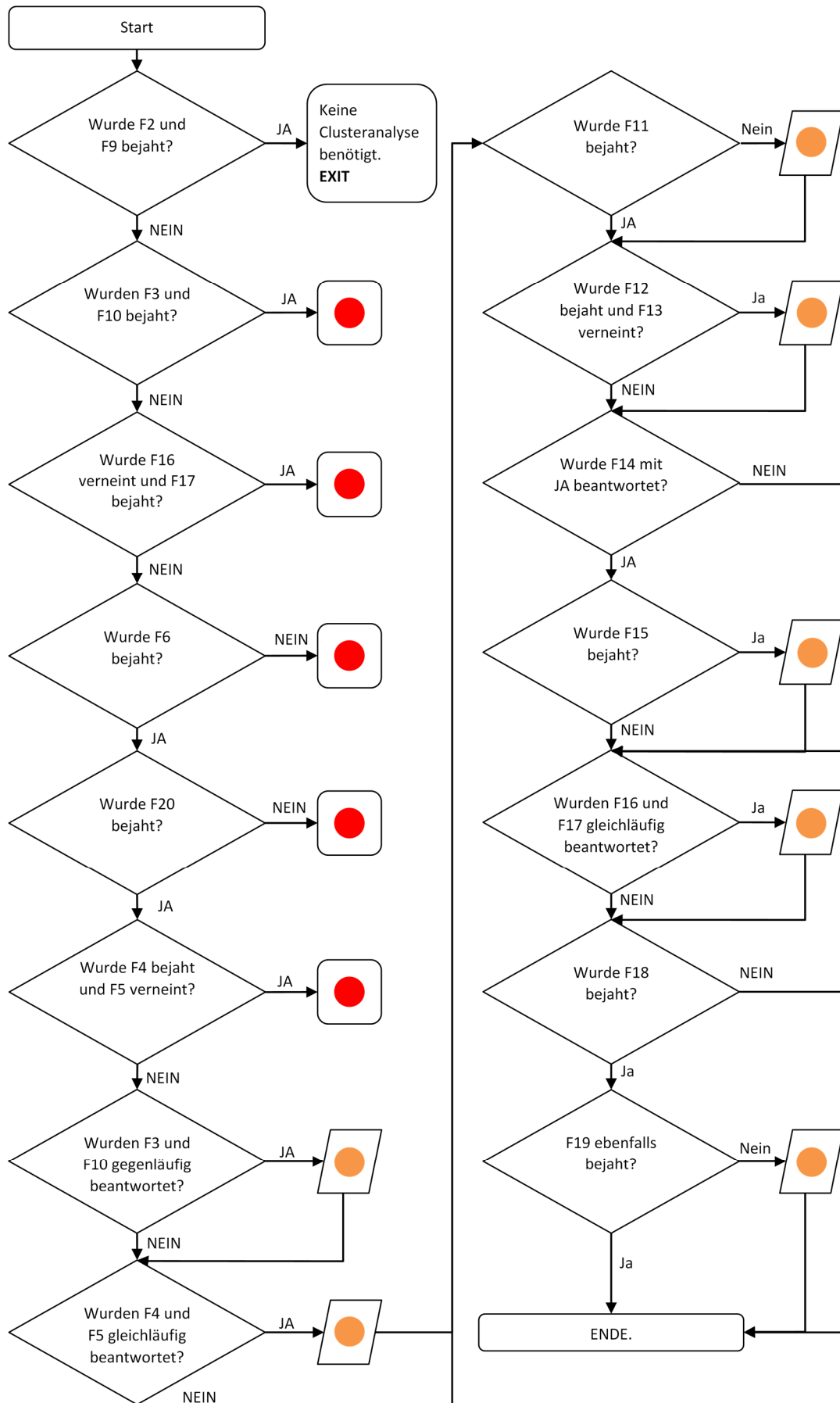


Abbildung 63 – Bewertungsschema für die Clusteranalyse

5.2.3 Definition der stochastischen Population

Latin-Hypercube-Stichprobenverfahren

Die stochastische Population wird mit dem Latin-Hypercube-Stichprobenverfahren mit einer Stichprobenanzahl von $n = 8$ erzeugt. Es werden je drei Variablen benötigt, um die Knochengometrie des Femurs oder die Materialeigenschaften für die kortikale Schicht des Femurs zu definieren.

Die Knochengometrie wird mit folgenden Variablen definiert: Femurlänge, Innendurchmesser des Femurschaftes und kortikale Dicke des Femurschaftes. Zur Definition der kortikalen Materialeigenschaften werden die Variablen E-Modul, Versagensspannung und Versagensdehnung verwendet.

Es erfolgt die vereinfachende Annahme, dass alle Variablen linear unabhängig zueinander sind. Die jeweiligen drei Variablen kreieren einen dreidimensionalen Raum. Für die Entnahme von Stichproben wird eine einheitliche Verteilungsfunktion für alle drei Variablen angenommen. Die Verteilungsfunktion jeder Variablen wird in n gleichmäßige Intervalle aufgeteilt, mit $n = \text{Stichprobenanzahl}$. Bei der Entnahme der Stichproben wird darauf geachtet, dass nur eine Stichprobe pro Reihe, Spalte und Tiefe des aufgeteilten Würfels existiert. Jede entnommene Stichprobe repräsentiert einen Punkt in diesem 3-dimensionalen Raum (siehe Abbildung 64). Nach der Entnahme der Stichproben können die Werte der Variablen für die jeweiligen Stichproben anhand der diskreten Verteilungsfunktion der Variablen berechnet werden.

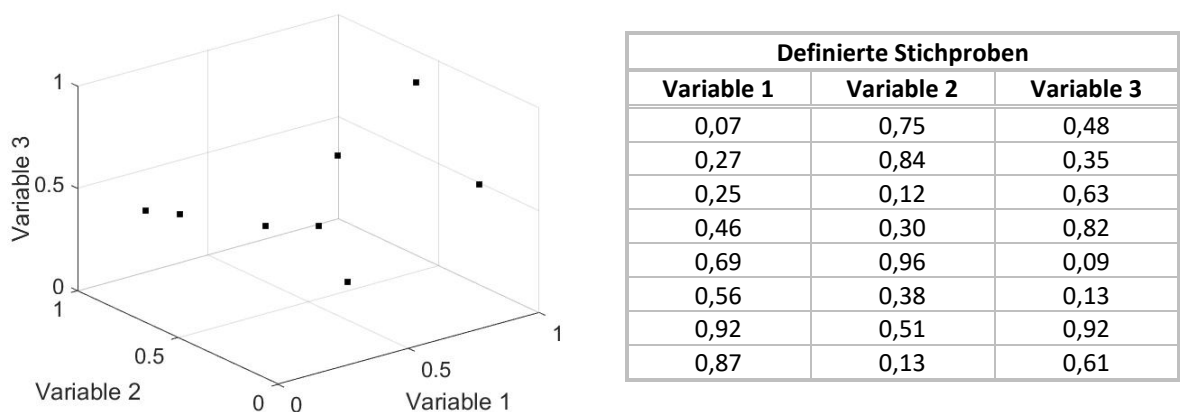


Abbildung 64 – Links: Visualisierung der entnommenen Stichproben / Rechts: Zufallszahlen der entnommenen Stichproben

Definition der kortikalen Materialeigenschaften

Für die Definition der kortikalen Materialeigenschaften wird ein linearisiertes elastisch-plastisches Materialgesetz analog zu Takahashi et al. [101] verwendet (siehe Abbildung 65). Zur Definition des Materialgesetzes werden folgende Variablen verwendet: Querdehnungskennzahl, E-Modul, Fließspannung, Versagensspannung und Versagensdehnung. Die Querdehnungskennzahl wird mit $\nu = 0,3$ analog zu Takahashi et al. [101] festgelegt.

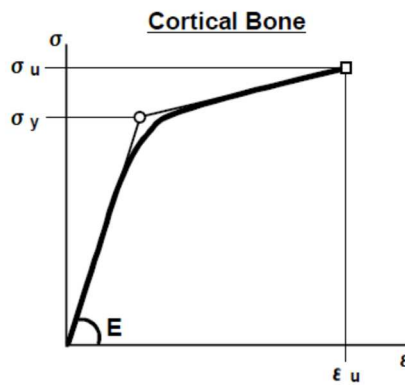


Abbildung 65 – Linearisiertes elastisch-plastisches Materialgesetz [101]

In dieser Arbeit werden die veröffentlichten Alterungsfunktionen von Dokko et al. [29] für die Charakterisierung der kortikalen Materialeigenschaften verwendet (siehe Abbildung 66). Die Variable 1 aus Abbildung 64 entspricht den Stichproben für den E-Modul, Variable 2 entspricht den Stichproben der Versagensspannung und Variable 3 entspricht den Stichproben der Versagensdehnung. Die Basis für die Berechnung der Absolutwerte der Stichproben für die einzelnen Variablen anhand Normalverteilungstabellen bilden der Mittelwert, sowie die +/- 1,64 Standardabweichung, welche für ein Alter von 35 Jahren aus den Diagrammen in Abbildung 66 ausgelesen werden.

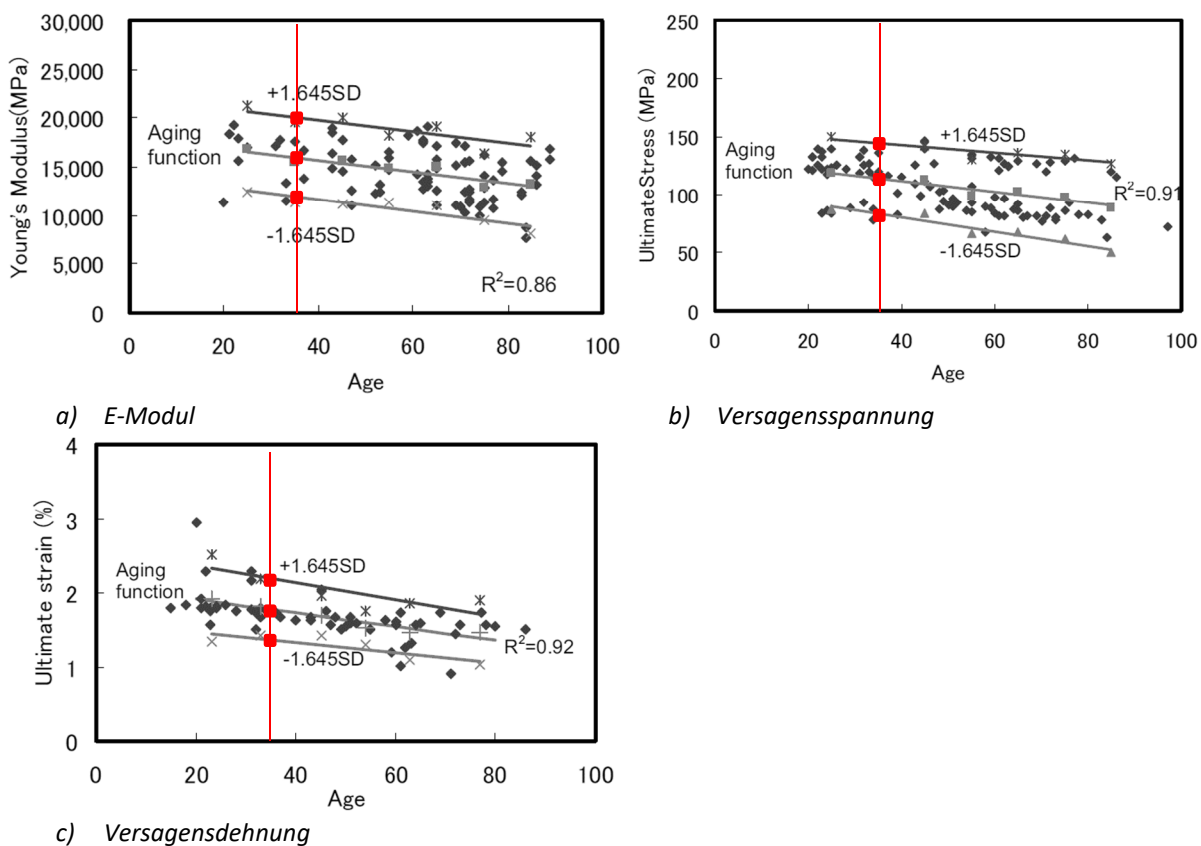


Abbildung 66 – Verwendete Alterungsfunktionen von Dokko et al. [29]

Tabelle 27 enthält die berechneten Werte für die Materialcharakterisierung der Stichproben. Die Fließspannung wird in Bezug zur Versagensspannung analog zu Dokko et al. [29] wie folgt definiert:

$$\sigma_F = 0,87 * \sigma_V.$$

Die Dehnratenabhängigkeit wird analog dem GHBMC-Modell in der Materialkarten-Definition berücksichtigt.

Tabelle 27 – Absolutwerte der Stichproben zur Materialdefinition

E-Modul [GPa]	Versagensspannung [MPa]	Versagensdehnung [-]
12,5	126	1,77
14,5	132	1,69
14,4	92	1,87
15,8	104	2,01
17,3	147	1,45
16,4	108	1,50
19,6	114	2,14
18,8	93	1,85

Definition der Knochengometrie

Für die Definition der Knochengometrie werden folgende drei Variablen verwendet: die Femurlänge, der Innendurchmesser des Knochens und die kortikale Dicke des Knochens für die Querschnittscharakterisierung an einer bestimmten Stelle in Längsrichtung des Femurs. Es werden zwei verschiedene Quellen verwendet, um die Geometrie der Femur-Population festzulegen.

Der Innendurchmesser des Knochens und die kortikale Dicke werden anhand der NHANES-Datenbank definiert. Diese beiden Variablen unterliegen der Annahme, dass die Varianzen dieser Referenzpopulationen den realen Varianzen in der US-Bevölkerung entsprechen.

Die amerikanische Gesundheits- und Ernährungsstudie NHANES (engl.: National Health and Nutrition Examination Survey) wird seit den 1960er Jahren durchgeführt, um den Gesundheits- und Ernährungsstatus der amerikanischen Bevölkerung zu untersuchen [82]. Die Studie wird vom Zentrum für Gesundheitsstatistik (engl.: National Center for Health Statistics) durchgeführt. Die NHANES-Studie wird seit den 1960er Jahren bis heute in Episoden durchgeführt. Jede dieser Episoden fokussiert sich auf bestimmte Bevölkerungsgruppen oder bestimmte Gesundheitsaspekte [82]. Die Studie untersucht jedes Jahr eine repräsentative Population von ca. 5000 Personen [82]. Die NHANES-Studie untersucht demographische, sozioökonomische, dietätische und gesundheitsbezogene Fragen [82]. Die Untersuchung umfasst medizinische, zahnärztliche, physiologische Messungen, sowie Laboruntersuchungen [82]. Die Episode mit dem Namen NHANES III wurde im Zeitraum 1988–1994 durchgeführt [19]. Innerhalb dieser Episode wurde auch eine Studie zur Untersuchung der Hüfte (engl. Hip Structural Analysis Program (Abkürzung: HSA)) durchgeführt. In diesem Programm wurden außer Knochendichtemessungen des Hüftknochens, auch die Struktur des Knochenquerschnitts am proximalen Ende des Femurs bis zu Beginn des Femurschaftes untersucht [10]. Abbildung 67 zeigt die Lage der untersuchten Querschnitte des Femurs. Zudem wird die vereinfachende Betrachtung dargestellt, den Querschnitt am Femurschaft als Kreisring anzunehmen, um weitere Kennzahlen wie z.B. das Flächenträgheitsmoment abzuleiten. Der in Abbildung 67 dargestellte Querschnitt am proximalen Ende des Femurschaftes, mit der vereinfachenden Annahme des Kreisringquerschnitts, wird für die Auswahl der Femora für die Population verwendet.

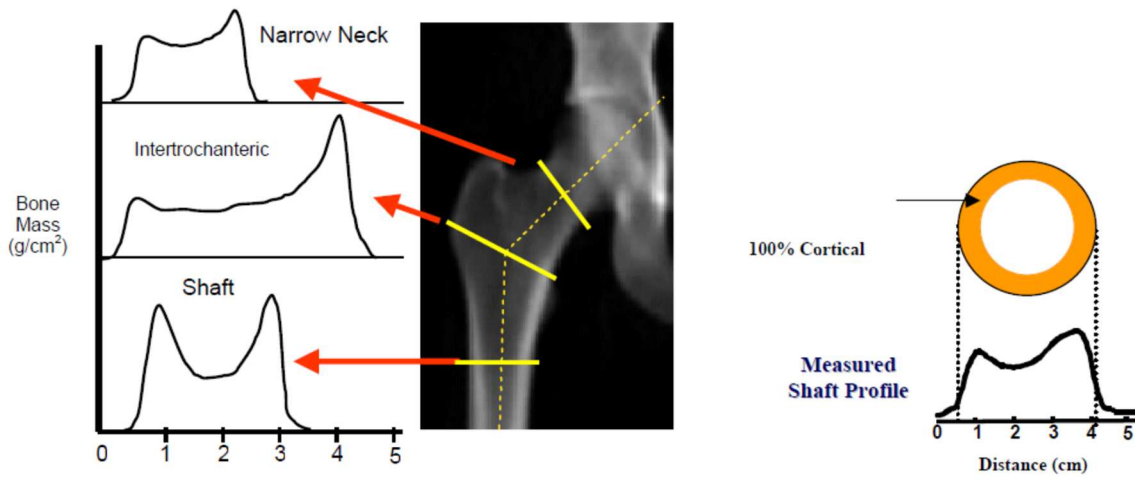


Abbildung 67 – Darstellung der Lage der untersuchten Querschnitte des Femurs und vereinfachte Berechnung [10]

Das HSA-Programm stellt ca. 13 600 Daten zu Verfügung. Abbildung 68 stellt eine eigene Aufbereitung der Daten des HSA-Programms dar und zeigt die Abhängigkeiten zwischen den einzelnen Geometrieparametern auf. Die Parameter des Innendurchmessers und der kortikalen Dicke sind die Daten des Schaftquerschnitts (vgl. Abbildung 67). Zudem ist in Abbildung 68 ersichtlich, dass die Daten des HSA-Programms keine Femurlängen enthalten, lediglich die Oberschenkellänge der untersuchten Personen.

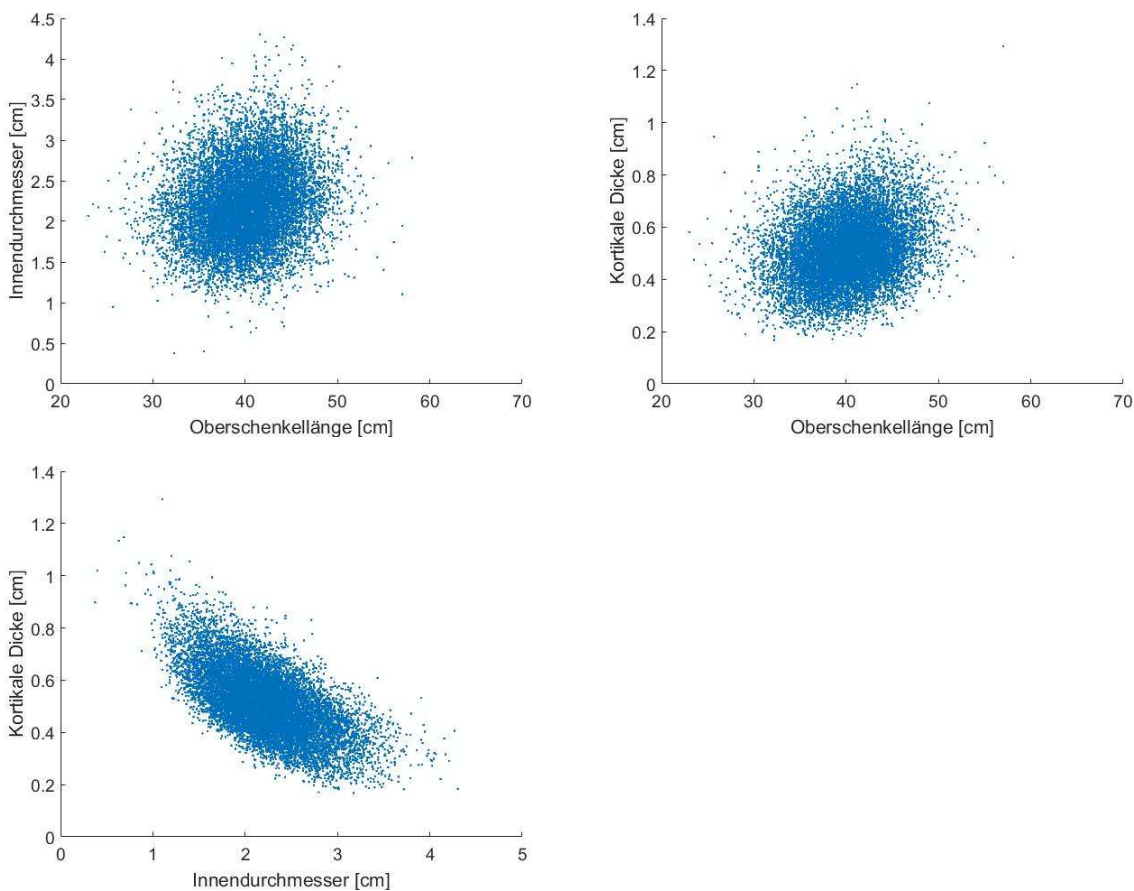


Abbildung 68 – Eigene Aufbereitung der NHANES-Datenbank [10] bez. Abhängigkeiten der Geometrieparameter

In der Dissertation von Park [85] wurden die Femurlängen 121 männlicher Femora veröffentlicht. Es erfolgt ebenfalls die Annahme, dass diese Referenzpopulation dieselbe Varianz bezüglich der Femurlänge aufweist als die US-Bevölkerung. In Park [85] sind keine Angaben veröffentlicht, wie die Femurlänge gemessen wurde. Beim Vermessen der Femora des Rechtsmedizin-Archivs wurde daher die Komplettlänge des Femurs verwendet, analog dem Vorgehen in Kerrigan et al. [65] (Abbildung 69).

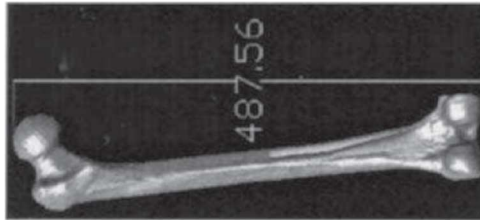
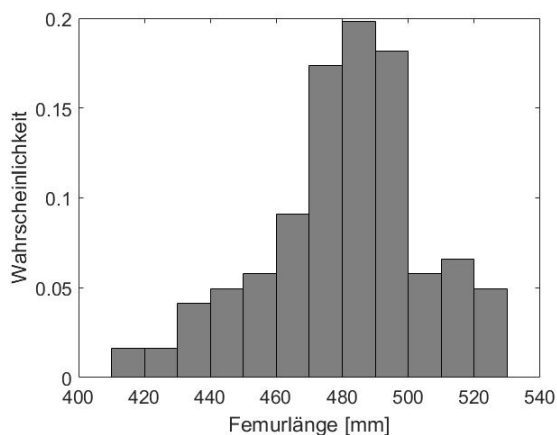


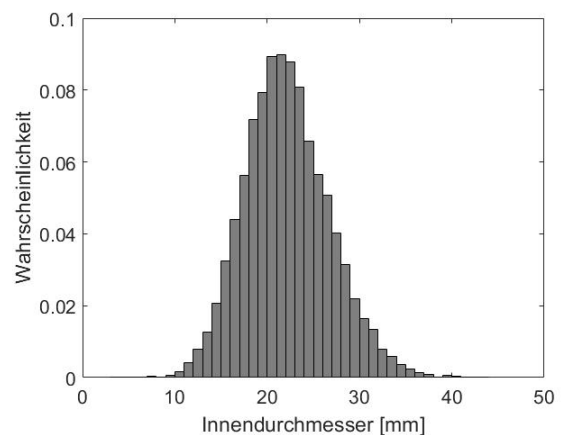
Abbildung 69 – Messung der Femurlänge in Kerrigan et al. [65]

Mit den definierten Zufallszahlen für die einzelnen Stichproben in Abbildung 64 und den diskreten Verteilungen in Abbildung 70 werden die Absolutwerte für die einzelnen Stichproben jeder Variablen berechnet.

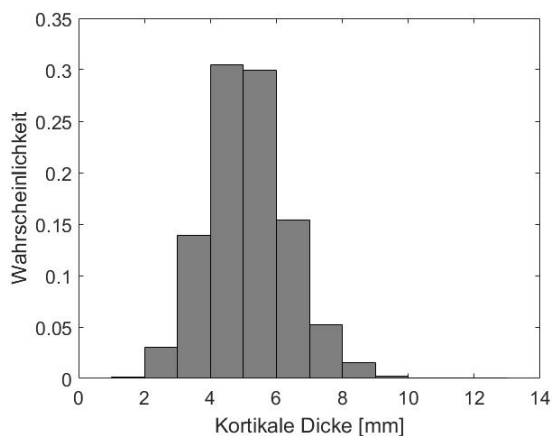
Die Variable 1 aus Abbildung 64 entspricht den Stichproben für die Femurlänge, Variable 2 entspricht den Stichproben der kortikalen Dicke und Variable 3 entspricht den Stichproben des Innendurchmessers. In Tabelle 28 sind die Ziel-Werte der Geometriedefinitionen der acht Femora dargestellt.



a) Femurlängen von Park [85]



b) Innendurchmesser aus NHANES-Datenbank [10]



c) Kortikale Dicke aus NHANES-Datenbank [10]

Abbildung 70 – Diskrete Wahrscheinlichkeitsverteilungen als Basis der Femur-Geometriedefinition

Tabelle 28 – Ziel-Geometrien der Femur-Population

Femurlänge [mm]	Kortikale Dicke [mm]	Innendurchmesser [mm]
438	5,8	22
469	6,3	20
466	3,7	23
481	4,5	26
491	7,5	16
484	4,7	17
515	5,1	29
504	3,8	23

Auswahl der Femur-Testproben aus CT-Kollektiv der Rechtsmedizin

Es werden aus dem CT-Kollektiv des Instituts für Rechtsmedizin diejenigen Femora ausgewählt, die am besten mit den Zielvorgaben in Tabelle 28 übereinstimmen, um realitätsnahe FE-Modelle der Femora zu erzeugen. Das in der Abteilung Biomechanik und Unfallforschung erstellte Kollektiv verfügt über 36 auf Basis von CT-Aufnahmen vermessene Femora, welche in 18 männliche und 18 weibliche Femora aufgeteilt sind. Die weiblichen Femora werden nicht in die mögliche Auswahl für die Population genommen, da die Femurlängen aus Park [85] nur von männlichen Femora stammten. Für die definierten Ziel-Geometrien aus Tabelle 28, welche mittels dem Latin-Hypercube-Stichprobenverfahren definiert sind, werden diejenigen Femora aus dem Archiv ausgewählt, welche am ähnlichsten bezüglich der Ziel-Geometrien sind, d.h. die geringste normierte euklidische Distanz zueinander aufweisen.

Aufgrund der begrenzten Auswahlmöglichkeit von 18 Femora des verwendeten Archivs verringert sich die Population auf 5 Femora. Die Charakteristiken der ausgewählten 5 Femora sind in Tabelle 29 dargestellt.

Tabelle 29 – Ist-Werte der Femur Geometriedefinition der ausgewählten Femora des CT-Kollektivs

Bezeichnung	Femurlänge [mm]	Kortikale Dicke [mm]	Innendurchmesser [mm]
M22	516	7	17
M30	476	9	14
M31	483	6	19
M46	488	5	26
M100	478	6	18

In Abbildung 71 werden die geometrischen Parameter der Ist-Population aus dem CT-Kollektiv (vgl. Tabelle 29) und der Ziel-Population aus dem Latin-Hypercube-Stichprobenverfahren (vgl. Tabelle 28) gegen Referenz-Datenbanken verglichen, um die Auswahl auf Plausibilität zu prüfen. Die Parameter beider Populationen, der Ist-Population des CT-Kollektivs und der Ziel-Population, befinden sich innerhalb der Varianzen der Referenzdatenbanken.

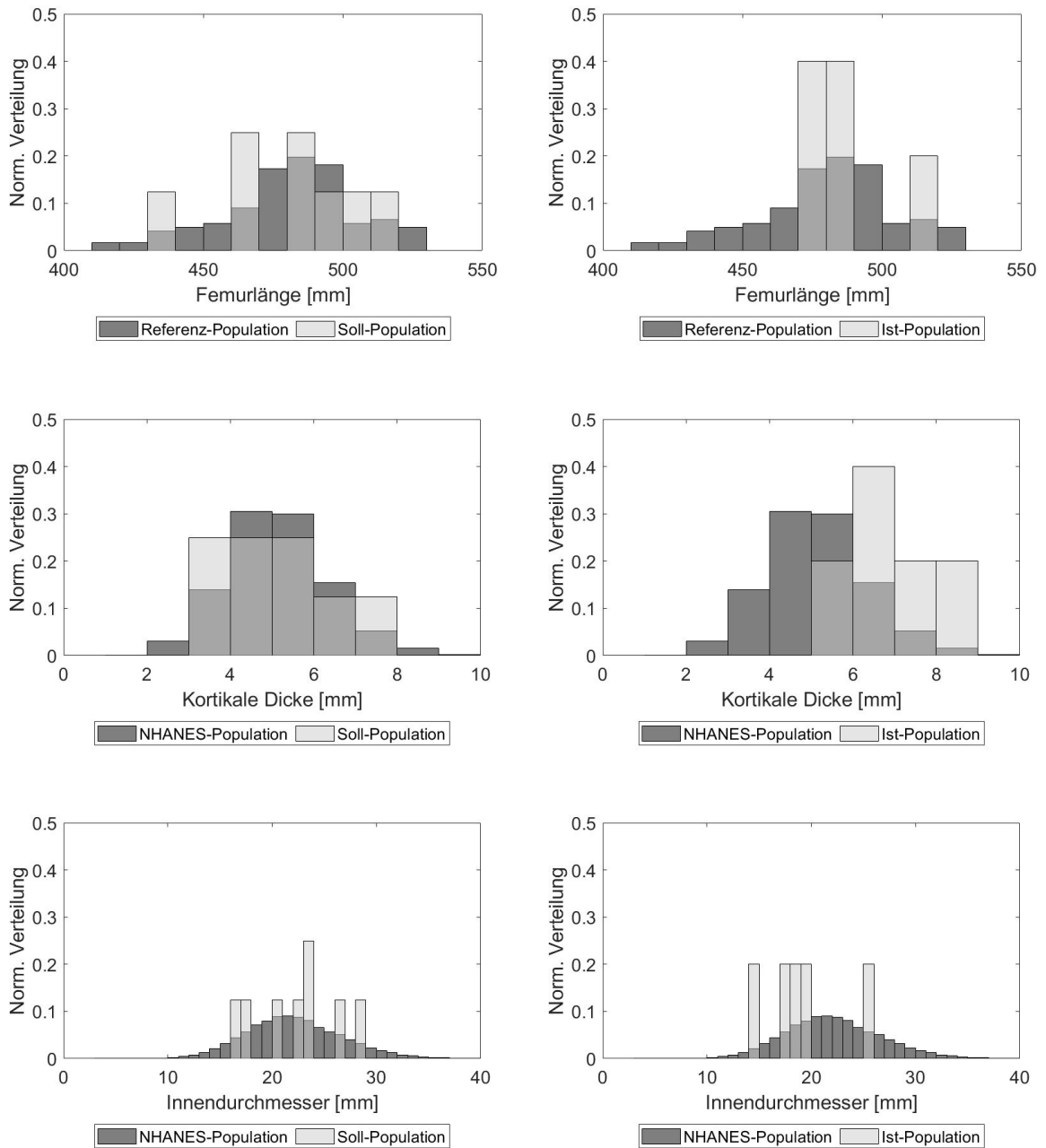


Abbildung 71 – Vergleich der Soll- und Ist-Population gegen die Referenzdatenbank NHANES [10] bez. kortikale Dicke und Innendurchmesser des Femurschafts und der Referenzdatenbank Park [85] bez. Femurlänge

5.2.4 Erzeugung der stochastischen Population

Geometrie Erzeugung

Für die fünf ausgewählten Individuen des CT-Kollektivs wurde aus dem Scan ein 3D Volumenmodell des Knochens von der Abteilung Biomechanik und Unfallforschung des Instituts für Rechtsmedizin zu Verfügung gestellt. Die kortikale Schicht des Femurs konnte aufgrund der groben CT-Auflösung nur im Schaftbereich extrahiert werden, da sie am distalen und proximalen Ende des Femurs sehr dünn wird und dort nicht mehr extrahiert werden konnte. Für jedes Femur wurden zwei Volumenmodelle bereitgestellt: Ein Volumenmodell der äußeren Hülle des kompletten Knochens und ein Volumenmodell des Schaftbereiches mit der extrahierten kortikalen Schicht.

Der Schaft-Bereich des Femurs wird mit zwei Elementen über die Dicke vernetzt. Es werden Hexaeder-Elemente verwendet und die Ziel-Elementgröße wird auf 3 mm definiert.

Der proximale- und distale Bereich eines Femurs, bei dem die Extraktion der kortikalen Schicht nicht möglich ist, wird mit einer Elementschicht von 2 mm Dicke vernetzt (siehe Abbildung 72).

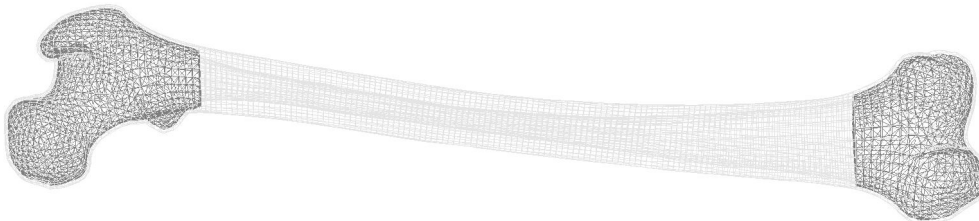


Abbildung 72 – Beispiel der Femurmodellierung für Femur M22

Der Übergang zwischen dem Schaftbereich mit zwei Elementen über die kortikale Dicke zu den Endbereichen mit nur einem Element über die kortikale Dicke, erfolgt mittels gemeinsamer Knoten (siehe Abbildung 73). Die Epiphysen des Femurmodells werden mit Tetraeder-Elementen ausgefüllt. Dieses aufgefüllte Volumen stellt die Spongiosa dar.

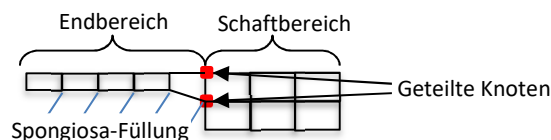


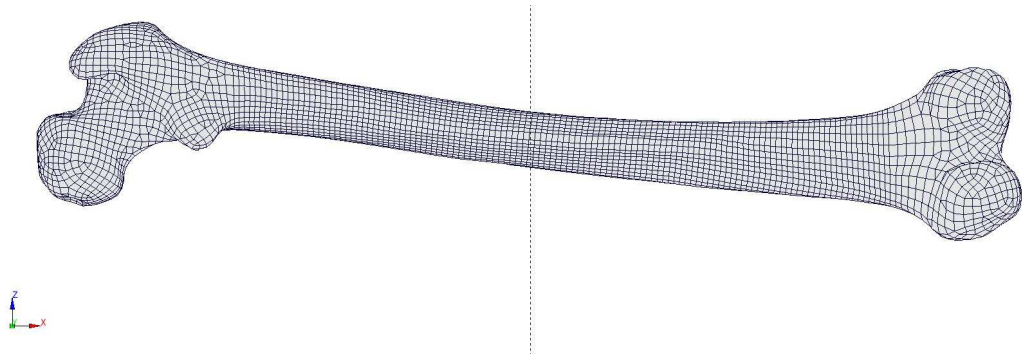
Abbildung 73 – Modellierung Übergangsbereiche zwischen Femurschaft und den Endbereichen

Die Vernetzung der Femurmodelle wird nach den Qualitätsstandards von Park et al. [85] bezüglich Jacobi-Verhältnis, Verwerfung (engl. warping), Kantenlängenverhältnis (engl. aspect ratio) und Schiefe (engl. skewness) überprüft. Die Ergebnisse sind in Anhang A6 gelistet.

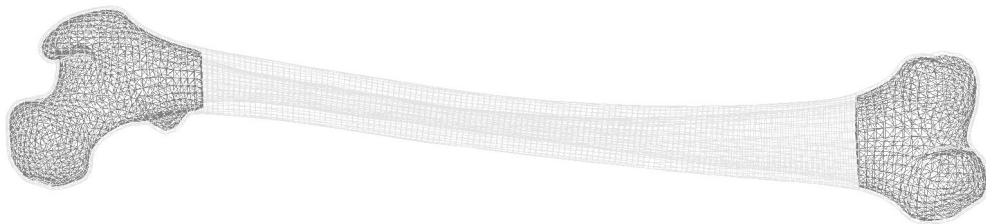
Für jedes Femurmodell wird folgendes visualisiert:

- die äußere Vernetzung des Femurmodells,
- eine transparente Darstellung zum Erkennen der Schaftgeometrie und des Schaft-übergangs zu den Endbereichen,
- eine Querschnittsdarstellung der Mitte des Femurschaftes,
- das Femurmodell eingebettet in der Versuchsumgebung.

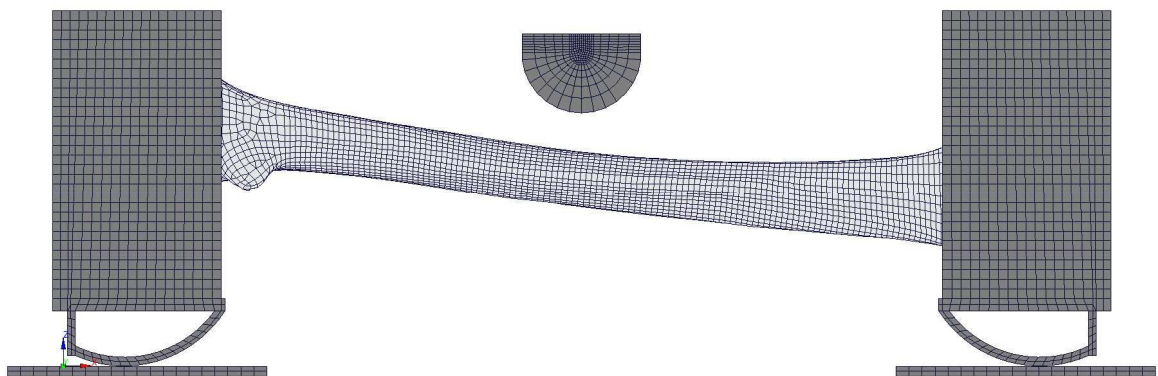
In Abbildung 74 werden die Ansichten beispielhaft für das Femurmodell M22 dargestellt. Anhang A5 enthält die Visualisierung aller weiteren erstellten Femurmodelle.



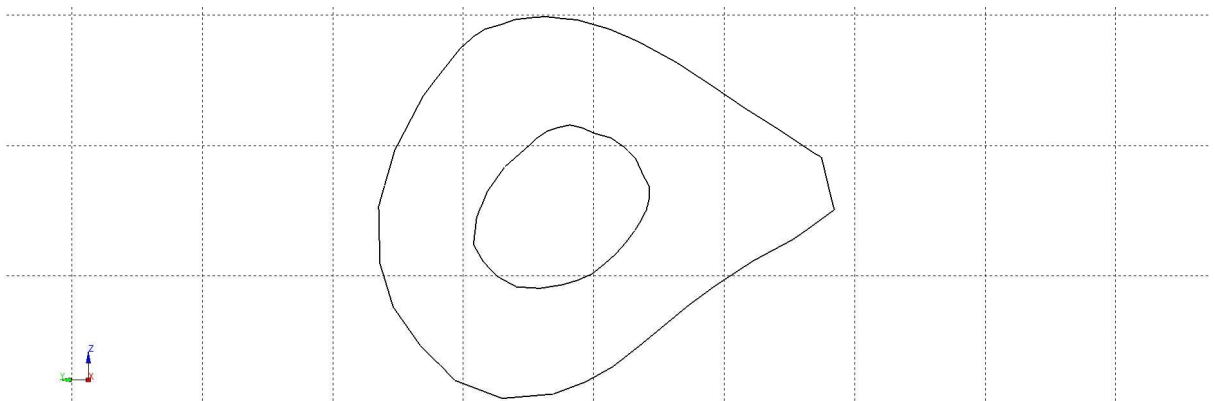
a) Darstellung der Femurvernetzung und Schnittlage des Femur M22



b) Darstellung der Vernetzung des Femurinneren (M22)



c) Femur eingebettet in den Simulationsaufbau (M22)



d) Femurquerschnitt resultierend aus dem Schnitt dargestellt in Graphik a)

Abbildung 74 – Modellierung Femur M22

Zusammensetzung der Population

Es wird analog zur Definition der Material-Stichproben und der Geometrie-Stichproben ebenfalls die Annahme getroffen, dass zwischen den Femurmodellen und den definierten Materialgesetzen eine lineare Unabhängigkeit existiert. Zudem wird die vereinfachende Annahme getroffen, dass die kortikalen Materialeigenschaften und die Materialeigenschaften der Spongiosa linear unabhängig voneinander sind. Abbildung 75 visualisiert die Zusammensetzung der Femurmodelle. Für jedes Femurmodell werden acht FE-Simulationsmodelle, eins für jedes der acht definierten Materialgesetze, erstellt. Insgesamt wird eine numerische Population von 40 Femora generiert. Die Materialdefinition der Spongiosa in den Endbereichen entspricht der Formulierung des GHBM-Modells und wird bei allen erstellten Femurmodellen konstant gehalten.

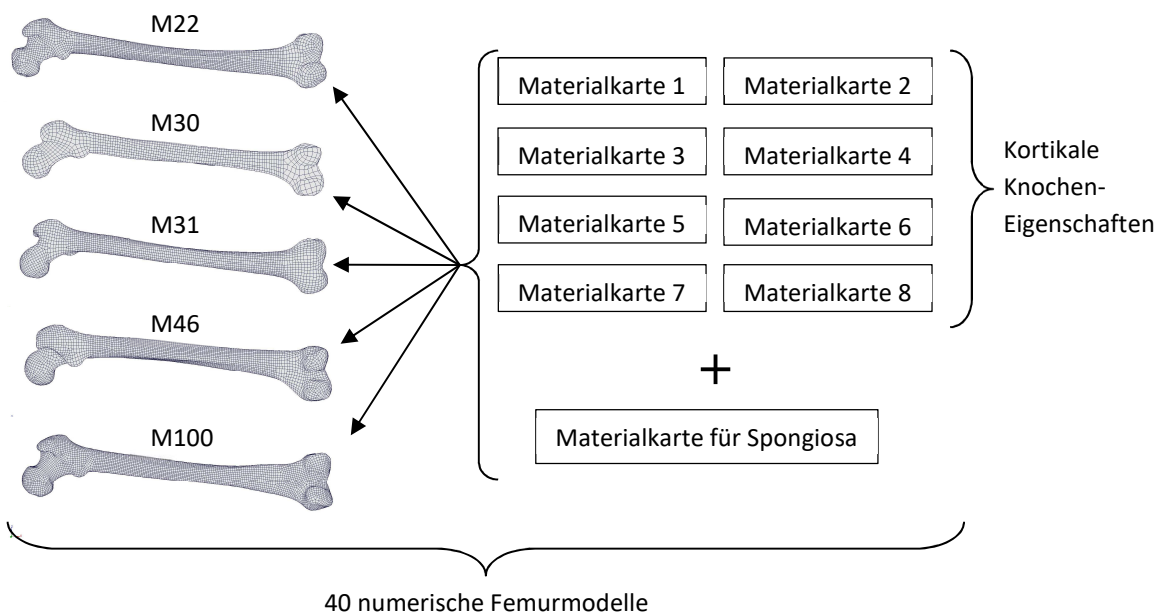


Abbildung 75 – Schaubild der erzeugten numerischen Population an Femora

Erzeugung der Modellvorhersagen

Die numerischen Modellvorhersagen für alle 40 numerischen Knochenmodelle werden mit dem Modell des Referenzexperiments, das in Kapitel 3 entwickelt und validiert wurde, erzeugt. Für die numerische Versuchsdurchführung wird der 3D-Versuchsaufbau in starrer Konfiguration verwendet (vgl. Abbildung 26). Die Simulationsdurchführung und Nachbereitung der Ergebnisse erfolgt entsprechend der Beschreibung in Abschnitt 3.2.1.

5.2.5 Vorstellung der Konzeptdatensätze zur Überprüfung des Validierungskonzeptes

Das populationsbasierte Validierungskonzept wird mit zwei Konzeptdatensätze getestet, welche im Folgenden kurz vorgestellt werden. Konzeptdatensatz I besteht aus dem kombinierten Datensatz, welcher die 40 numerischen Modellvorhersagen und die 15 experimentellen Ergebnisse beinhaltet (siehe Abbildung 76). Konzeptdatensatz II ist ein vereinfachter Datensatz des Konzeptdatensatzes I (siehe Abbildung 77). Für Konzeptdatensatz II werden nur wenige Kurven des Konzeptdatensatzes I verwendet. Konzeptdatensatz II wird artifiziell generiert, damit er natürliche Gruppierungen aufweist.

Als experimentelle Metadaten sind die Femurlängen in Park et al. [86] und die Flächenträgheitsmomente in der Mitte des Femurschaftes in Forman et al. [38] veröffentlicht. Damit die Metadaten der numerischen Modelle mit den experimentellen Metadaten vergleichbar sind, wird ebenfalls für die numerischen Modelle das Flächenträgheitsmoment in der Mitte des Femurschaftes ermittelt. Zudem wird als weitere Metadaten-Variable eine Materialkenngröße für die numerischen Modelle berechnet, welche im Folgenden Energie-Aufnahmekapazität des Knochens genannt wird. Diese Kenngröße entspricht dem Integral unter der definierten elastisch-plastischen Materialkurve der kortikalen Knochenschicht. Die Metadaten der numerischen Modelle sind in Tabelle 30 dargestellt.

Tabelle 30 – Metadaten der numerischen Modelle

Modellbezeichnung	Femurlänge [mm]	I_{xx} [mm ⁴]	Energieaufnahmekapazität des Knochens [-]							
			S1	S2	S3	S4	S5	S6	S7	S8
M22	516	30729	0,77	0,78	0,67	0,83	0,73	0,60	1,0	0,71
M30	476	32567	0,77	0,78	0,67	0,83	0,73	0,60	1,0	0,71
M31	486	24047	0,77	0,78	0,67	0,83	0,73	0,60	1,0	0,71
M46	488	45992	0,77	0,78	0,67	0,83	0,73	0,60	1,0	0,71
M100	478	35066	0,77	0,78	0,67	0,83	0,73	0,60	1,0	0,71

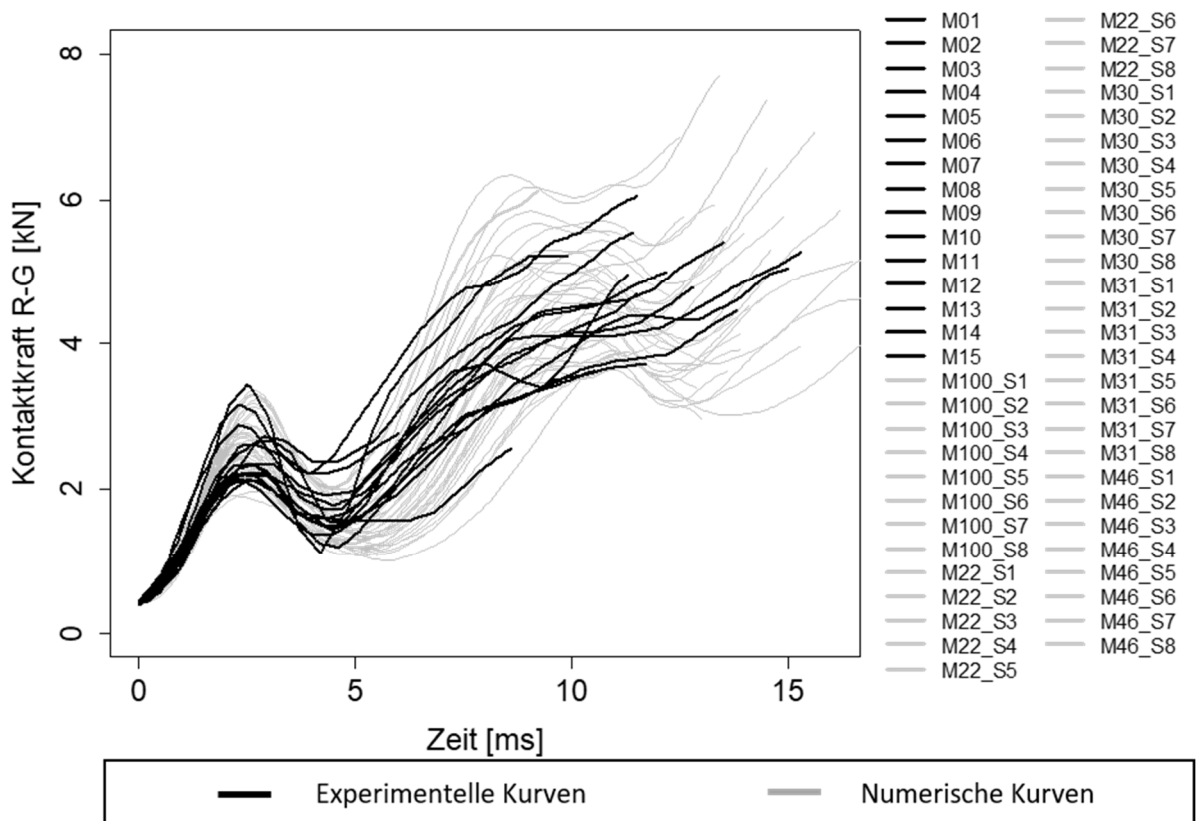


Abbildung 76 – Konzeptdatensatz I

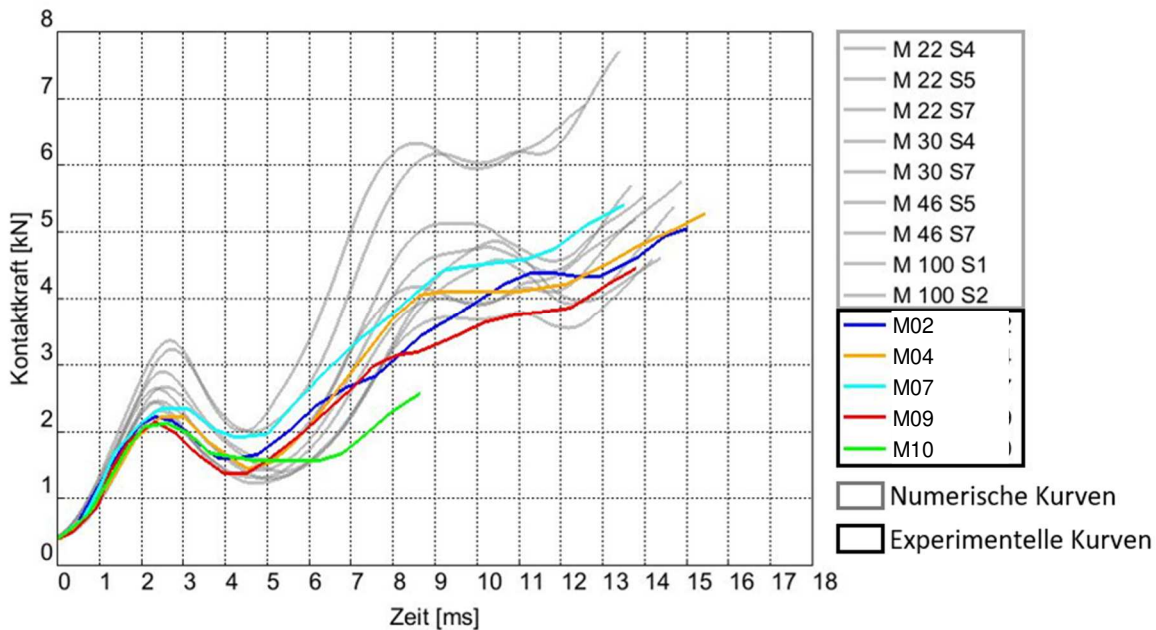


Abbildung 77 – Konzeptdatensatz II

5.3 Ergebnisse

5.3.1 Populationsbasierte Validierungsmethode mit Konzeptdatensatz I

In Tabelle 31 sind alle Informationen zusammengefasst, die für den populationsbasierten Validierungsprozess aus Abbildung 63 benötigt werden. Um die Lesbarkeit des Textes zu gewährleisten, befinden sich einige Informationen für die Beantwortung des Fragebogens in Tabelle 31 in Anhang A7. Die Informationen in Anhang A7 umfassen u.a. abweichende visuelle Darstellungen mit ähnlichem Inhalt wie Abbildung 78 bis Abbildung 81 (vgl. Abbildung A7.1 und Abbildung A7.2), Dendrogramme (vgl. Abbildung A7.4 bis Abbildung A7.6), mögliche Clusteraufteilungen (vgl. Tabelle A7.1), sowie die Visualisierung der Dendrogramme auf unterster Ebene (vgl. Abbildung A7.3) zur Beantwortung der Frage der Ähnlichkeit der Kurvenformen zwischen experimentellen und numerischen Ergebnissen in Frage 8. Zudem ist in Abbildung A7.7 in Anhang A7 rot markiert, an welcher Stelle der Validierungsprozess abgebrochen ist.

Die Fragen 1 bis 4 bez. der visuellen Analyse der Charakteristik des kombinierten Datensatzes werden mit Abbildung 76 beantwortet. Für die Clusteranalyse wird der Ward-Fusionierungsalgorithmus in Kombination mit dem *ptbserial*-Index zur Ermittlung der Clusteranzahl ausgewählt (vgl. Tabelle A7.1). Der int. VI berechnet eine Clusteraufteilung in 5 Cluster. Um den Vergleich zwischen der Clusteranzahl des rein experimentellen Datensatzes und der Clusteranzahl des kombinierten Datensatzes in Frage 12 durchführen zu können, werden die experimentellen Kurven ebenfalls mit dem Ward-Fusionierungsalgorithmus und dem int. VI *ptbserial*-Index einer Clusteranalyse unterzogen. Die berechnete experimentelle Clusteranzahl beträgt vier Cluster.

Tabelle 31 – Bewertungsfragebogen für Konzeptdatensatz I

#	FRAGE	Antwort	JA	NEIN
Visuelle Analyse				
1	Existieren bei visueller Betrachtung Gruppierungen in der Kurvenschar? Wenn ja, Clusteranzahl?	Abbildung 76		x
2	Ist bei visueller Betrachtung die Kurvenschar homogen?	Abbildung 76		x
3	Unterscheiden sich die Kurvenformen der numerischen Modellvorhersagen signifikant zu den Kurvenformen der experimentellen Ergebnisse?	Abbildung 76	x	
4	Existiert eine räumliche Trennung zwischen den numerischen Modellvorhersagen und den experimentellen Ergebnissen?	Abbildung 76		x
Metadaten Analyse				
5	Fallen für mindestens eine Variable der numerischen Metadaten Werte in den Wertebereich der experimentellen Metadaten?	Abbildung A7.1	x	
6	Für die numerischen Werte der Metadaten außerhalb des experimentellen Wertebereichs: Wahl der Werte anhand Datenbank-Vergleich plausibel?	Abbildung A7.2	x	
7	Existieren natürliche Gruppierungen für die einzelnen Variablen der Metadaten? Wenn ja, Ermittlung der Gruppenanzahl pro Variable. Was ist die maximale Anzahl?	3 Abbildung A7.1	x	
Clusteranalyse				
8	Analyse der untersten Ebene des Dendrogramms. Visualisierung der Verbindungen, die eine experimentelle Kurve mit einer numerischen Kurve verbinden. Unterscheiden sich die zusammengehörigen Kurvenpaare signifikant?	Abbildung A7.3	x	
9	Überprüfung der Homogenität eines Datensatzes anhand des Dendrogramms bei Verwendung des Single-Linkage Algorithmus mit euklidischer Distanzmetrik.	Nicht homogen. Abbildung A7.4		x
10	Überprüfung anhand eines Dendrogramms, welches mit dem Distanzmaß der dynamischen Zeitverzerrung erstellt wurde, ob sich die Kurvenformen der numerischen Modellvorhersagen signifikant von den Kurvenformen der experimentellen Ergebnisse unterscheiden.	Abbildung A7.5		x
11	Schneidet die index-basierte Clusteranzahl das Dendrogramm an längeren Ästen?	Abbildung A7.6	x	
12	Vergleich der index-basierten Clusteranzahl mit der Clusteranzahl aus rein experimentellem Datensatz. Ist die index-basierte Clusteranzahl geringer?			x
13	Wenn F12 bejaht wurde: Visuelle Überprüfung, ob die numerischen Kurven zwei oder mehr experimentelle Cluster miteinander verbinden?		-	-
14	Wurde F1 oder F7 oder wurden beide Fragen bejaht?		x	
15	Wenn F14 bejaht wurde: Ist die index-basierte Clusteranzahl kleiner als die Clusteranzahl aus F1 oder F7?			x
16	Visuelle Betrachtung der Kurvenscharen der einzelnen Cluster. Sind die Kurvenscharen kompakt definiert?	Abbildung 78		x
17	Existiert bezüglich des Wertebereichs eine visuell starke Streuung in den Clustern für mehr als eine Variable der Metadaten?	Abbildung 79	x	
18	Existieren innerhalb eines Cluster bei mehreren Variablen der Metadaten signifikante Trennungen zwischen Gruppen von Wertebereichen?	Abbildung 79		x
19	Wenn F18 bejaht wurde: Ist dies physikalisch durch Kompensationseffekte erklärbar?		-	-
20	Existiert mindestens ein Cluster, welches numerische- und experimentelle Kurven enthält?	Abbildung 78	x	

In Abbildung 78 ist die Kurvenschar des kombinierten Datensatzes wie folgt farblich in die einzelnen Cluster aufgeteilt: das Cluster 1 ist blau markiert, das Cluster 2 ist in rot dargestellt, das Cluster 3 ist in grün dargestellt, das Cluster 4 ist in rosa gefärbt und das Cluster 5 ist der Farbe grau zugeordnet. Mit dieser Abbildung können die Fragen 16 und 20 bez. der Kompaktheit der Clusterdefinition (vgl. Abbildung 58) und der Existenz gemischter Cluster (vgl. Abbildung 57) beantwortet werden. Zudem ist visuell erkennbar, dass die Kurvenendpunkte der einzelnen Cluster bez. der Frakturkraft signifikant streuen.

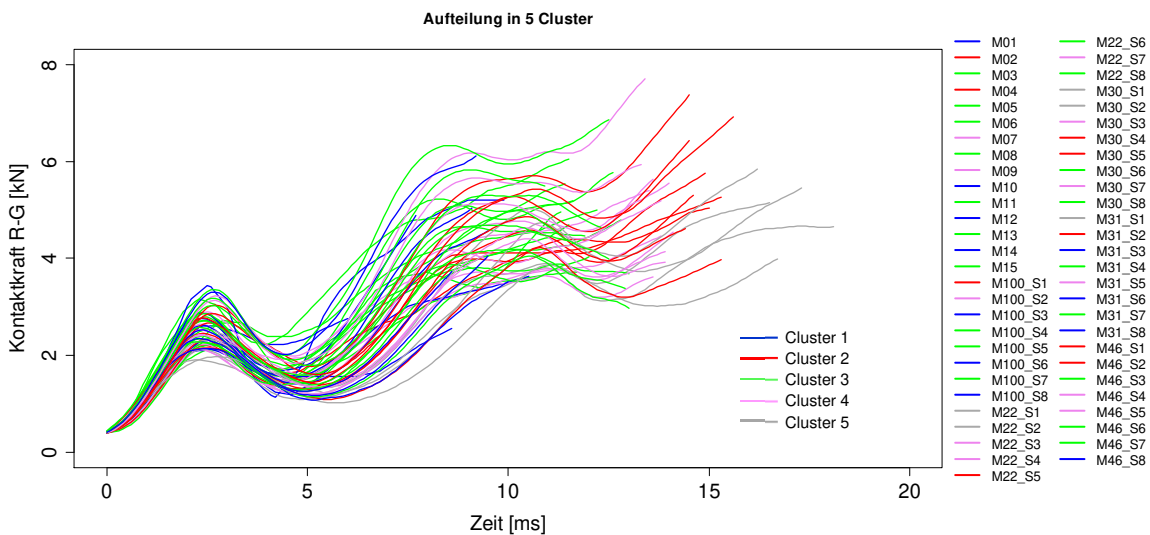


Abbildung 78 – Aufteilung der Kurvenschar in 5 Cluster nach Ward-Algorithmus

Abbildung 79 enthält analog zu Abbildung 78 dieselbe farbliche Zuordnung der Metadaten zu den jeweiligen Clustern. Anhand Abbildung 79 können die Fragen 17 bis 19 beantwortet werden. Mit diesen Fragen werden visuell die Metadaten bez. Streuung und Kompensationseffekte (vgl. Abbildung 62) untersucht. Es ist analog zu Abbildung 78 visuell deutlich erkennbar, dass die Wertebereiche der einzelnen Cluster bei allen Metadaten eine Streuung aufweisen. Lediglich die Metadaten-Variable der Energieaufnahme des Knochens besitzt bez. der Cluster 1 (blau), 2 (rot) und 5 (grau) eine geringe Streuung der Werte. Diese Cluster können bez. dieser Metadaten-Variable als kompakt definiert angesehen werden.

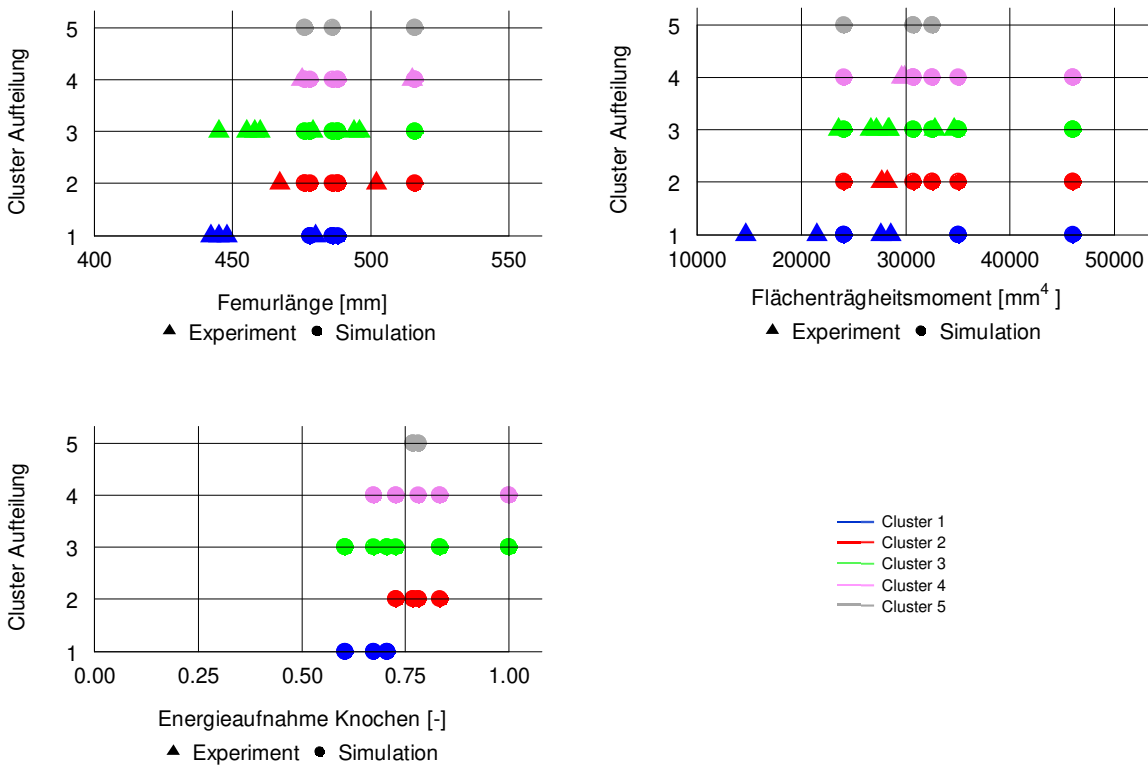


Abbildung 79 – Ergebnis der Clusteranalyse mit Aufteilung in 5 Cluster

In Abbildung 80 und Abbildung 81 sind die Wertebereiche der Metadaten der einzelnen Cluster im Vergleich zu Histogrammen von Referenz-Datenbanken dargestellt, um die Plausibilität der Metadaten

zu überprüfen. Die Wertebereiche der Population für beide Metadaten-Variablen liegen innerhalb der Histogramme der Referenz-Datenbanken. Die Wertebereiche der Population werden daher als plausibel bewertet. Für beide Metadaten-Variablen ist erkennbar, dass die Wertebereiche der einzelnen Cluster sehr groß sind. Zudem überschneiden sich oftmals bei beiden Metadaten-Variablen die Wertebereiche der einzelnen Cluster. Es ist in diesem Fall kein Erkenntnisgewinn dahingehend möglich, welche Wertebereiche der Referenz-Datenbanken die einzelnen Cluster repräsentieren.

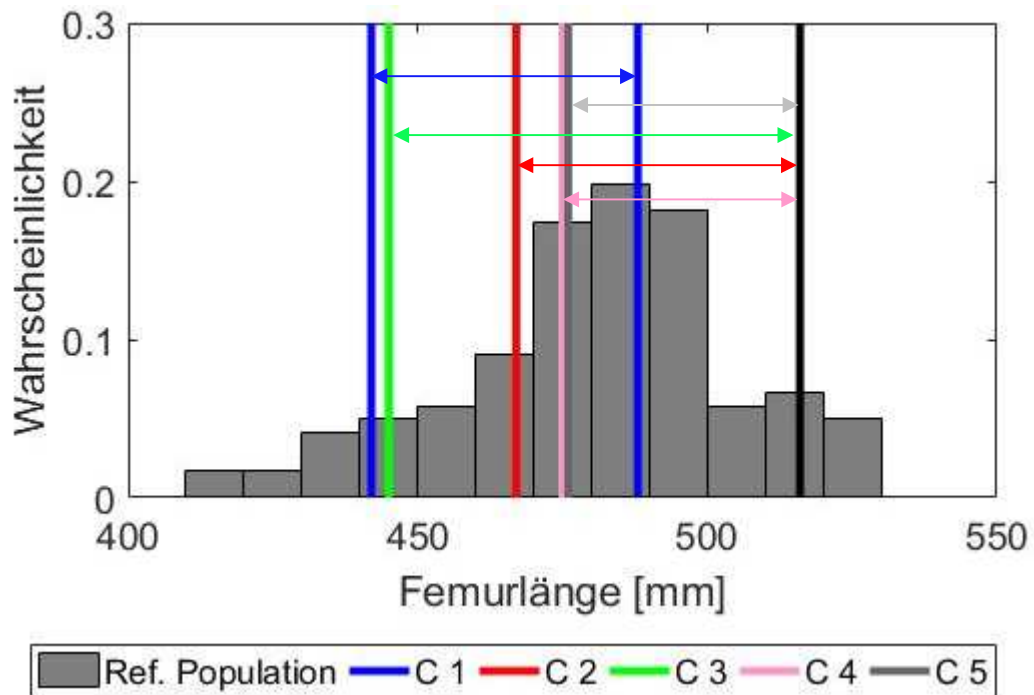


Abbildung 80 – Vergleich der Wertebereiche der Cluster bezüglich Femurlänge mit der Referenzpopulation der Femurlängen aus Park [85]

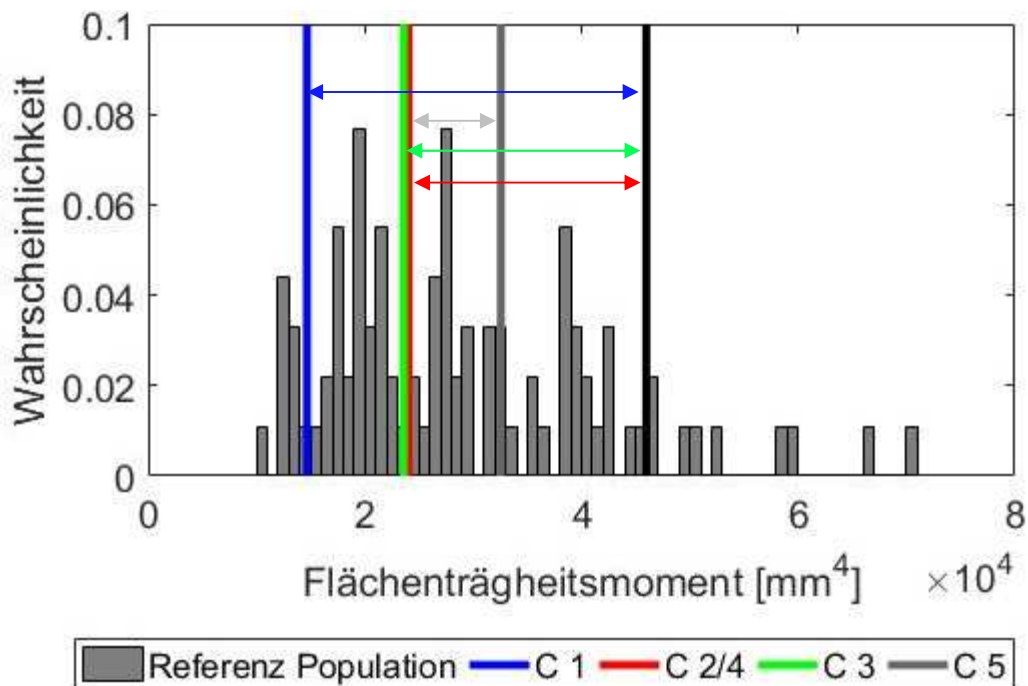


Abbildung 81 – Vergleich der Wertebereiche der Cluster bezüglich Flächenträgheitsmoment mit einer Referenzpopulation bestehend aus den Flächenträgheitsmomenten von Ivarsson et al. [58], Funk et al. [42], Kennedy et al. [62] und Kerrigan et al. [65]

Die Metadaten werden mit den Fragen 5 bis 7 hinsichtlich der Durchmischung der experimentellen und numerischen Metadaten, existierender Gruppierungen (vgl. Abbildung 59) und Plausibilität der Metadaten mit Abbildung A7.1 und Abbildung A7.2 untersucht.

Abbildung A7.3 bis Abbildung A7.6 analysieren den kombinierten Datensatz auf Basis von Dendrogramm-Strukturen. Diese Abbildungen helfen die Fragen 8 bis 11 bez. der Ähnlichkeit der Kurvenformen zwischen numerischen- und experimentellen Kurven, der Homogenität des Datensatzes und der Plausibilitätsüberprüfung der Clusteranzahl, zu beantworten. Frage 10 untersucht anhand eines Dendrogramms, welches mit der Distanzmetrik der dynamischen Zeitverzerrung und dem Ward-Fusionierungsalgorithmus erstellt wird, ob sich die experimentellen Kurven signifikant von den numerischen Modellvorhersagen unterscheiden. Etwa ein Drittel der experimentellen Kurven werden blockartig zusammengruppiert. Dies wird als Indiz dafür gewertet, dass sich die Kurvenformen zwischen numerischen Modellvorhersagen und experimentellen Ergebnissen nicht signifikant unterscheiden.

Der Validierungsprozess aus Abbildung 63 muss bei dem dritten Prozessschritt (siehe Abbildung A7.7) abgebrochen werden, da die Frage 16 verneint und die Frage 17 bejaht wird. Dies bedeutet, dass eine signifikante Streuung in den Kurvenscharen der einzelnen Cluster (siehe Abbildung 78) und den dazugehörigen Metadaten auftritt (siehe Abbildung 79).

5.3.2 Populationsbasierte Validierungsmethode mit Konzeptdatensatz II

Die populationsbasierte Validierungsmethode wird in diesem Abschnitt an Konzeptdatensatz II, wie in Abschnitt 5.3.1 bereits ausführlich beschrieben, angewendet. In Tabelle 32 sind alle Antworten des Fragebogens übersichtlich zusammengefasst. Zusätzliche Informationen befinden sich analog zu Abschnitt 5.3.1 in Anhang A8.

Die Frage F7, die zur Überprüfung der Existenz natürlicher Gruppierungen in den Metadaten dient, wird mit drei natürlichen Gruppierungen durch das Flächenträgheitsmoment beantwortet. Alle Kombinationen aus Fusionierungsalgorithmus und int. VI in Tabelle A8.1, welche eine Aufteilung in vier Cluster berechnen, besitzen dieselbe Clusteraufteilung. Es wird der Average-Linkage-Fusionierungsalgorithmus mit dem *ptbiserial*-Index für die Clusteranalyse verwendet.

Frage F10 bez. des Vergleichs der Kurvenformen zwischen numerisch- und experimentellen Ergebnissen wird anhand des Dendrogramms, welches mit der Distanzmetrik der dynamischen Zeitverzerrung und dem Average-Linkage-Fusionierungsalgorithmus erstellt wird, beantwortet. Die Mehrheit der experimentellen Kurven wird blockartig zusammengruppiert. Dies wird als Indiz dafür gewertet, dass sich die Kurvenformen zwischen numerischen Modellvorhersagen und experimentellen Ergebnissen unterscheiden. Um Frage F12 beantworten zu können, mit der die indexbasierte Clusteranzahl des kombinierten Datensatzes mit der Clusteranzahl des rein experimentellen Datensatzes verglichen wird, werden die experimentellen Kurven ebenfalls mit dem Average-Linkage-Fusionierungsalgorithmus und dem int. VI *ptbiserial*-Index einer Clusteranalyse unterzogen. Die berechnete experimentelle Clusteranzahl beträgt zwei Cluster.

Aus dem Bewertungsschema in Abbildung 63 und den Antworten in Tabelle 32 ergibt sich eine Validierungsgüte von 6 Punkten. Verglichen mit der maximal möglichen Anzahl von 7 Punkten ist die Validierungsgüte hoch.

Tabelle 32 – Bewertungsfragebogen für Konzeptdatensatz II

#	FRAGE	Antwort	JA	NEIN
Visuelle Analyse				
1	Existieren bei visueller Betrachtung Gruppierungen in der Kurvenschar? Wenn ja, Clusteranzahl?	3 Cluster Abbildung 77	x	
2	Ist bei visueller Betrachtung die Kurvenschar homogen?	Abbildung 77		x
3	Unterscheiden sich die Kurvenformen der numerischen Modellvorhersagen signifikant zu den Kurvenformen der experimentellen Ergebnisse?	Abbildung 77		x
4	Existiert eine räumliche Trennung zwischen den numerischen Modellvorhersagen und den experimentellen Ergebnissen?	Abbildung 77		x
Metadaten Analyse				
5	Fallen für mindestens eine Variable der numerischen Metadaten Werte in den Wertebereich der experimentellen Metadaten?	Abbildung A8.2	x	
6	Für die numerischen Werte der Metadaten außerhalb des experimentellen Wertebereichs: Wahl der Werte anhand Datenbank-Vergleich plausibel?	Abbildung A8.3	x	
7	Existieren natürliche Gruppierungen für die einzelnen Variablen der Metadaten? Wenn ja, Ermittlung der Gruppenanzahl pro Variable. Was ist die maximale Anzahl?	3 Abbildung A8.2	x	
Clusteranalyse				
8	Analyse der untersten Ebene des Dendrogramms. Visualisierung der Verbindungen, die eine experimentelle Kurve mit einer numerischen Kurve verbinden. Unterscheiden sich die zusammengehörigen Kurvenpaare signifikant?	Abbildung A8.4		x
9	Überprüfung der Homogenität eines Datensatzes anhand des Dendrogramms bei Verwendung des Single-Linkage Algorithmus mit euklidischer Distanzmetrik.	Nicht homogen. Abbildung A8.5		x
10	Überprüfung anhand eines Dendrogramms, welches mit dem Distanzmaß der dynamischen Zeitverzerrung erstellt wurde, ob sich die Kurvenformen der numerischen Modellvorhersagen signifikant von den Kurvenformen der experimentellen Ergebnisse unterscheiden.	Abbildung A8.6	x	
11	Schneidet die index-basierte Clusteranzahl das Dendrogramm an längeren Ästen?	Abbildung A8.7	x	
12	Vergleich der index-basierten Clusteranzahl mit der Clusteranzahl aus rein experimentellem Datensatz. Ist die index-basierte Clusteranzahl geringer?			x
13	Wenn F12 bejaht wurde: Visuelle Überprüfung, ob die numerischen Kurven zwei oder mehr experimentelle Cluster miteinander verbinden?		-	-
14	Wurde F1 oder F7 oder wurden beide Fragen bejaht?		x	
15	Wenn F14 bejaht wurde: Ist die index-basierte Clusteranzahl kleiner als die Clusteranzahl aus F1 oder F7?			x
16	Visuelle Betrachtung der Kurvenscharen der einzelnen Cluster. Sind die Kurvenscharen kompakt definiert?	Abbildung 82	x	
17	Existiert bezüglich des Wertebereichs eine visuell starke Streuung in den Clustern für mehr als eine Variable der Metadaten?	Abbildung 82		x
18	Existieren innerhalb eines Cluster bei mehreren Variablen der Metadaten signifikante Trennungen zwischen Gruppen von Wertebereichen?	Abbildung 82		x
19	Wenn F18 bejaht wurde: Ist dies physikalisch durch Kompensationseffekte erklärbar?	Abbildung 82	-	-
20	Existiert mindestens ein Cluster, welches numerische- und experimentelle Kurven enthält?	Abbildung 82	x	

Im Folgenden wird Abbildung 82 bis Abbildung 84 betrachtet. In diesen Abbildungen sind die einzelnen Cluster wie folgt farblich gekennzeichnet: Cluster 1 in blau, Cluster 2 in rot, Cluster 3 in grün und Cluster 4 in rosa. Die Kurven mit den höchsten Frakturkräften sind in die Cluster 3 und 4 aufgeteilt. Die zu diesen Kurven gehörenden Femora besitzen ein höheres Flächenträgheitsmoment als die restlichen Femora, wohingegen die Femurlänge und die Werte für die Energieaufnahmekapazität des Knochens im Bereich derjenigen der anderen Kurven liegen. Der einzige Unterschied zwischen Cluster 3 und Cluster 4 liegt in der Materialcharakterisierung. Cluster 2 besitzt im Vergleich zu den anderen Clustern eine geringe Femurlänge und ein geringes Flächenträgheitsmoment. Es ist zudem ein rein experimentelles Cluster.

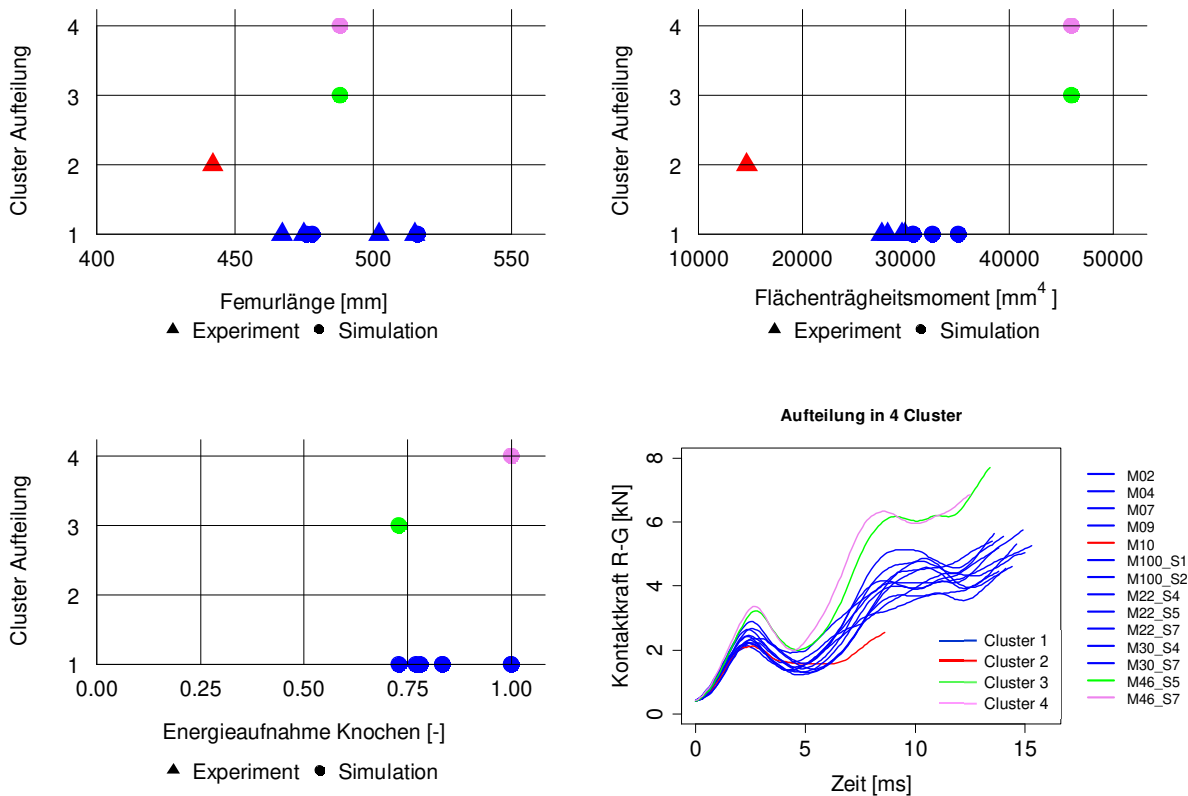


Abbildung 82 – Ergebnis der Clusteranalyse mit Aufteilung in 4 Cluster

Werden die Wertebereiche der einzelnen Cluster mit Referenz-Datenbanken verglichen, wie beispielsweise in Abbildung 83 und Abbildung 84 dargestellt, repräsentieren die Cluster 3 und 4 bezüglich der Femurlänge die Mitte der Bevölkerung, jedoch den rechten Randbereich der Bevölkerungsverteilung in Bezug auf das Flächenträgheitsmoment. Das Cluster 2 repräsentiert bezüglich der Femurlänge und des Flächenträgheitsmomentes den linken Randbereich der Bevölkerung. Das Cluster 1 repräsentiert für beide Variablen die Mitte der Bevölkerung.

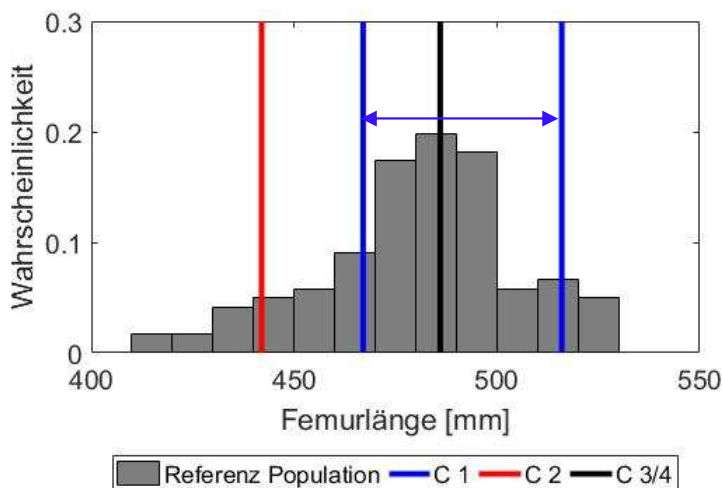


Abbildung 83 – Vergleich der Wertebereiche der Cluster bezüglich Femurlänge mit der Referenzpopulation der Femurlängen aus Park [85]

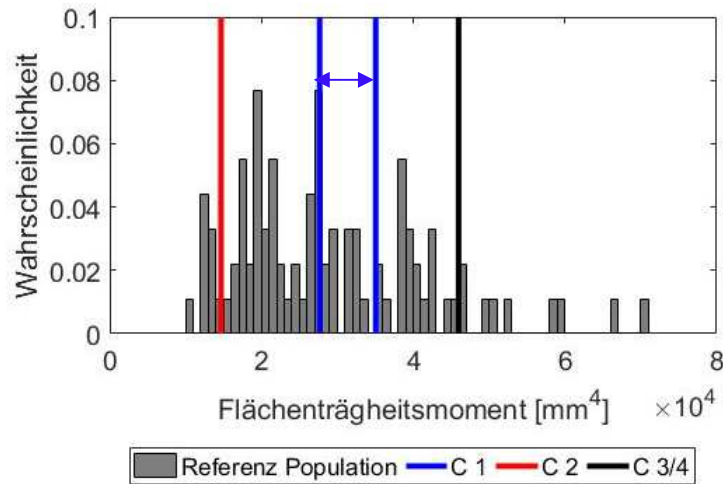


Abbildung 84 – Vergleich der Wertebereiche der Cluster bezüglich Flächenträgheitsmoment mit einer Referenzpopulation bestehend aus den Flächenträgheitsmomenten von Ivarsson et al. [58], Funk et al. [42], Kennedy et al. [62] und Kerrigan et al. [65]

5.4 Diskussion

5.4.1 Populationsbasierte Validierungsmethode

Im Rahmen dieser Arbeit wird ein populationsbasiertes Validierungskonzept entwickelt, welches auf der Anwendung von Clusteranalyseverfahren basiert. Das Validierungskonzept untersucht die Kurvenscharen und die Metadaten des kombinierten Datensatzes vor und nach der Clusteranalyse. Im Folgenden werden Aspekte diskutiert, welche bei der Entwicklung des Validierungskonzeptes berücksichtigt werden.

Wenn die Homogenität eines Datensatzes mit zwei unabhängigen Analysen bestätigt wird, wird kein Validierungsprozess benötigt. Testdatensatz I aus Abbildung 36 ist ein Beispiel für einen homogenen Datensatz. Diese Struktur eines Datensatzes sollte als Sonderfall betrachtet werden. Es ist keine Aufteilung des Datensatzes durch eine Clusteranalyse notwendig, da die Plausibilität der numerischen Modellvorhersagen durch die Gleichmäßigkeit der Kurvenschar visuell überprüft werden kann. Es ist keine weitere Vertrauensgewinnung bezüglich der numerischen Modellvorhersagen, auch außerhalb des experimentellen Bereiches, notwendig.

Experimentelle Kurven und numerische Modellvorhersagen ergänzen sich und bilden einen homogenen Datensatz. Die Wahrscheinlichkeit, dass dieser Sonderfall auftritt ist gering. Er wird dennoch vollständigshalber berücksichtigt.

Die Clusteranalyse wird verwendet, um einen großen Datensatz in kleinere Teildatensätze aufzuteilen, um diese vereinfacht analysieren zu können. Das Ergebnis der Clusteranalyse kann für eine Plausibilitätsprüfung der numerischen Modellvorhersagen verwendet werden. Ein Erkenntnisgewinn durch einen Vergleich der Cluster hinsichtlich der Metadaten und der Kurvenscharen ist jedoch nur möglich, wenn die Clusteranalyse den Datensatz in kompakt definierte Cluster aufteilt (vgl. Abbildung 58a). Es kann anhand der Tendenzen in den Metadaten und der zugrundeliegenden Physik des Experiments eine Aussage getroffen werden, aus welchen Gründen die jeweiligen Kurven in einem Cluster zusammengefasst sind.

Das Validierungsverfahren wird abgebrochen, wenn kein Erkenntnisgewinn aufgrund starker Streuung in den Metadaten und den Kurvenscharen der Cluster möglich ist. Insbesondere bei sehr großen Datensätzen ist diese Beobachtung als wahrscheinlich einzuordnen. Für einen Erkenntnisgewinn wäre

in einer derartigen Situation eine Weiterentwicklung der Methodik nötig. Hier wäre v.a. der Einsatz automatisierter Analysen auf der Grundlage der in dieser Arbeit vorgestellten Ansätze vorstellbar.

Im Folgenden werden die Fälle betrachtet, welche zu einer Verringerung der Validierungsgüte führen. Die Validierungsgüte verringert sich, wenn die Kurvenscharen der einzelnen Cluster eine Streuung aufweisen, die korrespondierenden Metadaten jedoch nicht. Zudem gilt der entgegengesetzte Fall, wenn die Metadaten der Cluster eine signifikante Streuung aufweisen, die zugehörigen Kurvenscharen jedoch nicht. Diese Widersprüche deuten auf Fehler in den numerischen Simulationen hin. Aus dieser Analyse kann jedoch keine Schlussfolgerung bezüglich der Fehlerquelle, d.h. ob der Fehler in den numerischen Modellen der Testobjekte liegt oder ob der Fehler im numerischen Versuchsaufbau liegt, gezogen werden.

Die Wahl der Clusteranzahl beeinflusst maßgeblich die Kompaktheit und Streuung der Cluster durch die Clusteraufteilung. Daher ist die Plausibilitätsprüfung der gewählten Clusteranzahl mit einem int. VI ein weiterer wichtiger Schritt. Ein Indiz für eine geeignete Clusteranzahl ist, wenn das Dendrogramm an langen Ästen geschnitten wird. Längere Äste bedeuten einen größeren Abstand zwischen den Objekten. Wenn das Dendrogramm an sehr kurzen Ästen geschnitten wird, bedeutet dies, dass Objekte getrennt werden, welche zueinander sehr ähnlich sind. Die Folge wäre die Generierung von unscharf getrennten Clustern. In diesem Fall verringert sich die Validierungsgüte, da der Erkenntnisgewinn erschwert wird. Eine weitere Möglichkeit die Clusteranzahl des int. VI zu bewerten, ist der Vergleich mit der experimentellen Clusteranzahl, welche aus der Clusteranalyse des rein experimentellen Datensatzes resultiert. Ist die Clusteranzahl des kombinierten Datensatzes geringer als die Clusteranzahl des rein experimentellen Datensatzes, kann der einzige nachvollziehbare Grund hierfür sein, dass die numerischen Modellvorhersagen mindestens zwei experimentelle Cluster zu einem größeren Cluster zusammenführen. Ist dies der Fall, dann kann die Clusteranzahl des kombinierten Datensatzes als plausibel angesehen werden. Wenn der entgegengesetzte Fall eintritt, verringert sich die Güte der Clusteranalyse. Des Weiteren kann überprüft werden, ob die Variablen der Metadaten oder die Kurven des kombinierten Datensatzes natürliche Gruppierungen aufweisen. Die bestimmten Clusteranzahlen aus den natürlichen Gruppierungen sollten ebenfalls nicht geringer sein als die Clusteranzahl, welche mit einem int. VI berechnet ist. Wenn die Clusteranzahl des int. VI geringer ist, ist es ein Indiz für eine nicht geeignete Clusteranzahl, da der int. VI die natürliche Datenstruktur nicht korrekt interpretiert. In diesem Fall verringert sich ebenfalls die Validierungsgüte. Des Weiteren wird der Kompensationseffekt betrachtet, der bereits in Abschnitt 5.2.2 beschrieben wurde. Wenn anhand des Kompensationseffektes physikalisch erklärt werden kann, dass ähnliche Kurven signifikant unterschiedliche Wertebereiche in den Metadaten besitzen, ist es gerechtfertigt, dass sie in einem Cluster zusammengefasst sind. Wenn dies nicht der Fall ist, ist es ein Indiz dafür, dass die Clusteraufteilung nicht geeignet ist, oder dass Fehler in den numerischen Simulationen vorhanden sind. In diesem Fall wird die Validierungsgüte verringert.

Es ist in Zukunft kritisch zu hinterfragen, ob dieses harte Kriterium der Kompaktheit für alle Cluster notwendig ist, oder ob ein weiches Kriterium, dass beispielsweise die Mehrheit der Cluster kompakt definiert sein sollte, ausreichend wäre, um eine gewisse Merkmalsvarianz innerhalb eines Clusters zuzulassen. Zudem sollte in Zukunft untersucht werden, ob ein objektiveres Kriterium zur Beurteilung der Streuung gefunden werden kann, als eine rein visuelle Betrachtung. Es könnte beispielsweise die Streubreite eines Clusters in Relation zur Streubreite des kombinierten Datensatzes vor der Clusteranalyse gesetzt werden. Bei diesem Ansatz müsste untersucht werden, bis zu welchem Streubreiten-Verhältnis ein Cluster als *kompakt definiert* gilt und ab welchem kritischen Maß es als *streuungsbehaftet* bewertet werden muss.

Des Weiteren sollte untersucht werden, ob bei der Dendrogramm-Analyse ebenfalls ein objektiveres Bewertungskriterium eingeführt werden kann. Es ist nicht empfehlenswert, das Dendrogramm bei

einer sehr geringen Clusteranzahl zu schneiden, da in diesem Fall zu erwarten ist, dass die Cluster sehr groß gewählt sind und eine signifikante Streuung in den Clustern existiert. Es ist jedoch ebenfalls nicht empfehlenswert, das Dendrogramm bei einer sehr großen Clusteranzahl zu schneiden, da in diesem Fall zu erwarten ist, dass die Anzahl der Elemente pro Cluster sehr klein ist und dies ebenfalls die Analyse und den Erkenntnisgewinn erschwert. Es ist zu untersuchen, ob ein Bereich definiert werden kann, z.B. zwischen 30% und 70% der max. Dendrogramm-Höhe, in dem eine geeignete Clusteranzahl liegen sollte. Diese Analyse sollte dann komplementär zu der bereits existierenden Überprüfung erfolgen, ob die Clusteraufteilung an längeren Dendrogramm-Ästen erfolgt.

Darüber hinaus ist es zukünftig empfehlenswert, ein objektiveres Kriterium für die Untersuchung natürlicher Gruppierungen in den Metadaten-Variablen zu entwickeln. Es wird empfohlen, für jede Metadaten Variable eine eindimensionale Clusteranalyse mit derselben Kombination aus Fusionierungsalgorithmus und int. VI durchzuführen, die auch für die Clusteranalyse des kombinierten Datensatzes verwendet wird. Auf diese Weise kann eine bessere Objektivität bezüglich der Bestimmung der Clusteranzahl hinsichtlich der Metadaten-Variablen erreicht werden.

Durch eine Clusteranalyse können folgende drei Cluster-Arten definiert werden (vgl. Abbildung 57):

- Cluster enthält nur experimentelle Kurven,
- Cluster enthält nur numerische Modellvorhersagen,
- Cluster enthält experimentelle und numerische Kurven.

Der Begriff Validierung wurde bereits in Abschnitt 1.2.1 spezifiziert und beinhaltet einen quantitativen Vergleich zwischen numerischen Modellvorhersagen und Experimenten. Wenn nach einer Clusteranalyse kein Cluster gegeben ist, welches experimentelle und numerische Kurven enthält, wird der Validierungsprozess abgebrochen. In diesem Fall könnte nur eine Trendanalyse hinsichtlich der Kurven und Metadaten zwischen Clustern erfolgen, welche entweder nur experimentelle Kurven enthalten oder nur numerische Modellvorhersagen. In diesem Fall ist keine direkte Vergleichbarkeit zwischen experimentellen und numerischen Modellvorhersagen gegeben. Bei einem Cluster, das experimentelle und numerische Kurven enthält, kann untersucht werden, ob eine ähnliche Existenz der Metadaten hinsichtlich experimentellem und numerischem Ergebnis zu ähnlichen Kurven führt. Dieser direkte Vergleich zwischen numerischen- und experimentellen Ergebnissen ist ein essentieller Schritt bei jedem Validierungsprozess und stellt auch bei der entwickelten populationsbasierten Validierungsmethodik einen bedeutenden Schritt in der Plausibilitätsprüfung der numerischen Modellvorhersagen dar.

Zudem wird visuell untersucht, ob in der Kurvenschar des kombinierten Datensatzes eine Trennung zwischen experimentellen und numerischen Kurven existiert. Ebenfalls wird untersucht, ob für mindestens eine Variable der Metadaten, die numerischen Werte in den experimentellen Wertebereich fallen. Existiert eine Trennung hinsichtlich der Kurvenschar und der Metadaten-Variablen zwischen experimentellen und numerischen Daten, obwohl ein gemeinsames Cluster existiert, wird der Validierungsprozess abgebrochen. Dies ist ein Indiz dafür, dass die Clusteraufteilung nicht geeignet ist, da sie eine offensichtliche Trennung berücksichtigen müsste. In diesem Fall wäre kein gemeinsames Cluster mit experimentellen und numerischen Kurven vorhanden. Wenn der Fall eintritt, dass bei visueller Betrachtung der Kurvenschar eine Trennung zwischen numerischen und experimentellen Ergebnissen existiert, die Wertebereiche der Metadaten-Variablen zwischen numerischen und experimentellen Werten sich jedoch überlappen, wird die Validierungsgüte verringert. Dies wird auch im entgegengesetzten Fall umgesetzt. In diesen beiden Fällen ist es ein Indiz für die Existenz von Fehlern im numerischen Modell oder in der Modellierung der numerischen Testproben.

Ein weiterer Aspekt, welcher bei der Entwicklung des populationsbasierten Validierungskonzepts berücksichtigt wird, ist die Plausibilitätsprüfung der Metadaten der numerischen Testproben. Wenn die Werte der Metadaten-Variablen außerhalb real existierender Varianzen liegen, was durch einen Datenbank-Vergleich überprüft werden kann (vgl. Abbildung 61), sind die entwickelten numerischen Testproben als nicht realistisch zu bewerten. In diesem Fall wird der Validierungsprozess abgebrochen.

Die Überprüfung der Plausibilität der Kurvenformen der numerischen Modellvorhersagen ist ein weiterer Aspekt, der bei der Entwicklung des Validierungskonzeptes berücksichtigt wird. Die vorgeschlagenen Kombinationen aus Fusionierungsalgorithmen und int. VIs aus Tabelle 25 verwenden die euklidische Distanzmetrik zur Durchführung der Clusteranalyse. Dieses Distanzmaß berücksichtigt nicht die Kurvenform während der Clusteranalyse (vgl. Testdatensatz VII in Abschnitt 4.3.2). Die Kurvenform repräsentiert jedoch die Physik des Experiments. Zudem spiegelt die Fläche unter der Kurve die Energieaufnahmekapazität des Knochens während der gesamten Belastung bis zur Fraktur wider. Da das euklidische Distanzmaß die Kurvenform bei der Clusteranalyse nicht berücksichtigt, muss dieser Aspekt separat überprüft werden. Wenn sich die Kurvenformen der numerischen Modellvorhersagen visuell signifikant von den Kurvenformen der experimentellen Ergebnisse unterscheiden, ist es ein Indiz dafür, dass der numerische Versuchsaufbau die Physik des realen Experiments nicht korrekt widerspiegelt. Wird dieses Indiz zusätzlich durch eine blockartige Dendrogramm-Struktur, die mit der dynamischen Zeitverzerrung als Distanzmetrik erzeugt wird, gestärkt, wird die Clusteranalyse abgebrochen. Wenn nur eine Analyse einen signifikanten Unterschied in den Kurvenformen zwischen den numerischen Modellvorhersagen und den experimentellen Ergebnissen bestätigt, wird dies als Indiz gewertet, dass der numerische Versuchsaufbau fehlerbehaftet ist. Die Güte des Validierungsergebnisses wird in diesem Fall nur verringert, der Validierungsprozess jedoch nicht abgebrochen. Zukünftig werden Untersuchungen dahingehend empfohlen, ob noch objektivere Alternativen existieren, als die Untersuchung einer blockartigen Dendrogramm-Struktur unter Verwendung des Distanzmaßes der dynamischen Zeitverzerrung, um unterschiedliche Kurvenformen zu identifizieren.

In Kapitel 4 wurden einige Kombinationen aus Distanzmetrik, Fusionierungsalgorithmus und int. VI für den Einsatz an biomechanischen Datensätzen als geeignet bewertet (vgl. Tabelle 25). Es kann bis zum jetzigen Zeitpunkt keine Aussage dahingehend getroffen werden, welche Kombinationen sich bevorzugt eignen. Erweist sich während des Validierungsprozesses eine Kombination aus Fusionierungsalgorithmus und int. VI als nicht geeignet, äußert sich dies in einer deutlich verringerten Validierungsgüte. In diesem Fall wird empfohlen, den Validierungsprozess mit einer anderen Kombination zu wiederholen. Zukünftig ist es empfehlenswert, diesen Aspekt weiter zu untersuchen. Es ist davon auszugehen, dass sich für unterschiedliche Datensätze mit unterschiedlichen Charakteristiken unterschiedliche Kombinationen eignen.

Die maximale Güte des Validierungsprozesses wird mit sieben Punkten definiert. Sie resultiert aus der maximalen Punkteanzahl, welche bei der Verringerung der Güte während des Validierungsprozesses möglich sind.

Die Bewertungsskala wird so definiert, dass jeder rote Punkt zu einem Abbruch des Bewertungsprozesses führt und jeder orangene Punkt die maximal erreichbare Güte um eine Stufe verringert. Bei der Prozessentwicklung werden die betrachteten Aspekte nicht hinsichtlich der Schwere ihres Einflusses bewertet, um die Bewertung durch Gewichtungsfaktoren zu beeinflussen. Dennoch existiert in dem entwickelten Bewertungsschema (vgl. Abbildung 63) eine indirekte Gewichtung, da einzelne Aspekte mit einer unterschiedlichen Anzahl an Fragen in die Bewertung miteinfließen. Dieses Bewertungsschema sollte daher anhand weiterer Testdatensätze und

Validierungslastfälle untersucht werden und könnte gegebenenfalls mit Wichtungsfaktoren angepasst werden.

Bereits Klein et al. [68] beschäftigten sich mit der Entwicklung einer Strategie, um Modellvorhersagen von parametrisierten Modellen auch außerhalb des experimentell abgedeckten Validierungsbereiches zu rechtfertigen. Das entwickelte Validierungskonzept von Klein et al. [68] wurde bereits in Abschnitt 1.2.5 vorgestellt. Ein Nachteil des entwickelten Validierungskonzeptes von Klein et al. ist, dass nur die Versagenszeitpunkte und die linearisierten Kurvensteigungen am Anfang der Belastungskurve der Testproben für die Trendanalyse zwischen den numerischen- und experimentellen Ergebnissen betrachtet wurden. Es wurde jedoch nicht die gesamte Belastungskurve der Testobjekte betrachtet. Ein weiteres Konzept, welches oftmals für die Validierung von parametrisierten Modellen verwendet wird, ist der subjekt-spezifische Validierungsansatz, welcher bereits in Abschnitt 1.2.5 detailliert beschrieben wurde. Es wird die Schlussfolgerung gezogen, dass das parametrisierte Modell auch außerhalb des experimentell abgedeckten Bereiches als validiert gilt, wenn die subjekt-spezifischen Replikationen der experimentellen Ergebnisse eine ausreichende Güte aufweisen. Der subjekt-spezifische Validierungsprozess ist sehr arbeitsaufwändig und oftmals nur in begrenztem Ausmaß möglich. Dieser Ansatz kann meist nur von der Institution durchgeführt werden, welche auch die experimentelle Studie durchgeführt hat, da für diesen Ansatz die CT-Aufnahmen der Testproben sowie sehr detaillierte Beschreibungen der Gewebe benötigt werden, welche meist nicht veröffentlicht werden.

Bei dem vereinfachten subjekt-spezifischen Validierungsprozess wird anhand globaler Parameter, wie z.B. Alter oder Körpergröße, die *subjekt-spezifische* Replikation erzeugt (vgl. Abschnitt 1.2.5). In diesem Fall ist der 1:1 Vergleich zwischen der *subjekt-spezifischen* Replikation und dem korrespondierenden Experiment streng genommen nicht gerechtfertigt, da aufgrund der individuellen Varianz anhand globaler Parameter keine subjekt-spezifische Replikation einer Testprobe erfolgen kann.

Das im Rahmen dieser Arbeit entwickelte populationsbasierte Validierungskonzept besitzt den wesentlichen Vorteil, dass aufwändige Replikationen der Testproben, wie es der subjekt-spezifische Ansatz erfordert, entfallen. Ein weiterer Vorteil des populationsbasierten Validierungsansatzes ist, dass die gesamte Belastungskurve der Testproben, sowie die Metadaten, berücksichtigt werden. Es wird aufgrund des ganzheitlichen Bewertungsansatzes der Kurvenscharen und Metadaten des kombinierten Datensatzes, vor und nach der Clusteranalyse, eine Plausibilitätsbewertung der numerischen Modellvorhersagen, auch außerhalb des experimentell abgedeckten Raumes, ermöglicht. Ein weiterer Vorteil dieses Validierungskonzeptes besteht darin, dass eine große Anzahl an stochastischen Simulationen anhand eines Prozesses bewertet werden kann.

Es muss jedoch an dieser Stelle nochmals deutlich darauf hingewiesen werden, dass diese vorgestellte populationsbasierte Validierungsmethode einen ersten Konzeptentwurf darstellt. Der Fokus dieser Arbeit ist die Entwicklung eines ganzheitlichen Bewertungskonzeptes. Hinsichtlich der berücksichtigten Aspekte zur Entwicklung des Validierungskonzeptes besteht kein Anspruch auf Vollständigkeit. In Zukunft ist zu prüfen, ob noch weitere relevante Aspekte mitberücksichtigt werden sollten. Diese können an das bereits existierende Bewertungsschema (vgl. Abbildung 63) angehängt werden. Die größte Limitation dieses Validierungskonzeptes sind die subjektiven Fragestellungen, welche in diesem Abschnitt ausführlich diskutiert werden. Zukünftige Anstrengungen sollten dahingehend erfolgen, objektivere Bewertungskriterien zu erarbeiten, damit der Bewertungsprozess zukünftig automatisiert erfolgen kann.

Diese populationsbasierte Validierungsmethode eignet sich vor allem als Ergänzung der Validierung mit bestehenden Vorgehensweisen, wie z.B. der subjekt-spezifischen Validierung, bei der meistens nur

wenige subjekt-spezifische numerische Modellvorhersagen mit den korrespondierenden experimentellen Ergebnissen verglichen werden (siehe Abschnitt 1.2.5).

5.4.2 Populationsdefinition

In Abschnitt 5.2.3 wird eine Population an Femora definiert, die der Belastung des Referenz-experiments unterzogen wird. Die Modellvorhersagen dieser Femora-Population werden in Abschnitt 5.2.5 verwendet, um die Konzeptdatensätze zu definieren. Mit diesen Konzeptdatensätzen wird das entwickelte populationsbasierte Validierungskonzept getestet.

Im Folgenden wird das Vorgehen bei der Populationsdefinition in Abschnitt 5.4.2 diskutiert. Das Latin-Hypercube-Stichprobenverfahren wird für die Bestimmung der Stichproben ausgewählt, da es bei der Entnahme der Stichproben die zugrundeliegenden Verteilungsfunktionen der Variablen berücksichtigt und für eine geringere Anzahl an Stichproben geeignet ist [73]. Bei einer geringen Stichprobenanzahl wäre eine Stichprobenentnahme auf Basis von Perzentilen ebenfalls eine Alternative gewesen. Bei dieser Methodik wäre bei jedem X.-Perzentil eine Stichprobenentnahme erfolgt. Das Latin-Hypercube-Stichprobenverfahren stellt einen speziellen Fall des Monte-Carlo-Stichprobenverfahrens dar [73]. Bei dem Monte-Carlo-Stichprobenverfahren wird jedoch eine relativ hohe Anzahl an Stichproben benötigt, um die Verteilung der einzelnen Variablen abzubilden. Die benötigte Stichprobenanzahl liegt nach Laz et al. [73] in der Größenordnung von mehreren tausend Stichproben. Um einen adäquaten Ansatz für eine geringere und größere Stichprobenanzahl vorzuschlagen, wird das Latin-Hypercube-Stichprobenverfahren ausgewählt, da es bei der Stichprobenentnahme die zugrundeliegenden Verteilungsfunktionen der Variablen bereits bei einer geringen Stichprobenanzahl berücksichtigt.

Bei der Populationsdefinition mit dem Latin-Hypercube-Stichprobenverfahren erfolgt die Annahme, dass alle Variablen linear unabhängig voneinander sind. Diese Annahme der linearen Unabhängigkeit wird nicht nur innerhalb der geometrie- oder materialdefinierenden Variablen, sondern auch zwischen den geometrie- und materialdefinierenden Variablen getroffen. Im Folgenden wird diese vereinfachende Annahme der linearen Unabhängigkeit der beteiligten Variablen diskutiert.

Die HSA-Datenbank der NHANES-Studie enthält keine Messdaten bezüglich der Femurlänge. Sie enthält lediglich Daten bezüglich der kortikalen Dicke und des Innendurchmessers am proximalen Ende des Femurschafts, sowie die Messdaten der Oberschenkelhöhe (vgl. Abbildung 68). Die Oberschenkelhöhe entspricht natürlich nicht der Femurlänge. Es kann jedoch die Annahme erfolgen, dass die beiden Maße korrelieren. Unter dieser Annahme und bei Verwendung des Maßes der Oberschenkelhöhe, anstatt der Femurlänge, kann die Annahme der linearen Unabhängigkeit bezüglich der Variablen: Oberschenkelhöhe, kortikale Dicke und Innendurchmesser überprüft werden (siehe Abbildung 68). Zwischen dem Innendurchmesser und der Oberschenkelhöhe, sowie der kortikalen Dicke und der Oberschenkelhöhe, existiert bei visueller Betrachtung keine signifikante Abhängigkeit. Zwischen diesen beiden Variablen-Paaren ist die Annahme der linearen Unabhängigkeit gerechtfertigt. Zwischen der kortikalen Dicke und dem Innendurchmesser existiert jedoch eine signifikante Abhängigkeit. Die Abhängigkeit dieser beiden Variablen sollte in zukünftigen Studien berücksichtigt werden. Die Auswirkung der vernachlässigten Abhängigkeit der beiden Variablen ist in dieser Arbeit als gering einzuschätzen, da die numerischen Femurmodelle auf der Basis realer Femur-Geometrien entwickelt werden. Bei der Erstellung numerischer Populationen mit parametrisierten Modellen, sollte dieser Aspekt der Abhängigkeit zwischen Innendurchmesser und kortikaler Dicke berücksichtigt werden. Bei Vernachlässigung dieser Abhängigkeit könnten Modelle erzeugt werden, welche unrealistische Geometrien aufweisen. Es wird für zukünftige Populationsdefinitionen empfohlen, ein zweidimensionales Latin-Hypercube-Stichprobenverfahren für die Populationsdefinition der Femurlänge und der kortikalen Dicke, oder der

Femurlänge und des Innendurchmessers, zu verwenden und danach die abhängige Variable zu berechnen.

Zwischen den geometriedefinierenden und den materialdefinierenden Variablen wird ebenfalls eine lineare Unabhängigkeit angenommen.

In der Literatur wurde eine Abhängigkeit zwischen der Knochendichte und des E-Moduls der kortikalen Knochenschicht und der Spongiosa berichtet. Morgan et al. [81] berichteten über die Abhängigkeit zwischen der Knochendichte und des E-Moduls der Spongiosa. Helgason et al. [53] veröffentlichten eine umfassende Literaturrecherche über die Abhängigkeit zwischen der Knochendichte und des E-Moduls der kortikalen Knochenschicht und der Spongiosa bezüglich verschiedener Knochen. Nach Helgason et al. [53] existieren signifikante Unterschiede zwischen Studien bezüglich der Formulierung der Abhängigkeit, von z.B. einer linearen bis zu einer exponentiellen Beziehung. In bereits existierenden parametrisierten Femurmodellen wurden diese Abhängigkeiten teilweise berücksichtigt. In dem entwickelten Femurmodell von Bryan et al. [14] wurde die Abhängigkeit zwischen Knochendichte und E-Modul der Spongiosa nach Morgan et al. [81] berücksichtigt. In dem parametrisierten Femurmodell von Park et al. [86] wurde die Abhängigkeit zwischen Knochendichte und E-Modul der kortikalen Knochenschicht berücksichtigt. In Park et al. [86] wurde ebenfalls auf die Unterschiede in der Abhängigkeitsbeschreibung zwischen Knochendichte und des E-Moduls der kortikalen Knochenschicht zwischen Studien hingewiesen.

Es wird empfohlen, den Einfluss der Abhängigkeit zwischen Knochendichte und E-Modul auf eine numerische Modellvorhersage in zukünftigen Studien zu untersuchen. Hierbei sollte speziell auch der Einfluss verschiedener Formulierungen der Abhängigkeit auf eine numerische Modellvorhersage untersucht werden. Wenn der Einfluss dieser Variablen-Abhängigkeit in diesen Studien signifikant ist, sollte diese in zukünftigen Populationsdefinitionen mitberücksichtigt werden. Da der Unterschied in der Formulierung der Abhängigkeit zwischen einzelnen Studien signifikant ist, wird davon ausgegangen, dass der Fehler in dieser Arbeit, durch die Vernachlässigung dieser Variablen-Abhängigkeit, innerhalb der Varianz zwischen betrachteten Studien liegt.

Takahashi et al. [101] veröffentlichten eine Literaturrecherche hinsichtlich der Datenverfügbarkeit bezüglich materialdefinierender Parameter und entwickelten Alterungsfunktionen für die materialdefinierenden Parameter E-Modul, Versagensspannung und Versagensdehnung, die für die Definition eines linear elastisch-plastischen Materialgesetzes benötigt werden. Takahashi et al. [101] veröffentlichten keine Informationen bezüglich der Qualitätskriterien der Literaturrecherche und der verwendeten Studien. Die Alterungsfunktionen von Takahashi et al. [101] wiesen ein Maximum in den Materialparametern bei einem Alter von 40 bis 50 Jahren auf. Dies wird als nicht realistisch gewertet, da laut Forman et al. [38] die maximalen Werte für die Materialcharakterisierung zwischen einem Alter von 25 bis 45 Jahren auftreten. Dokko et al. [29] überarbeiteten die verwendeten Literaturquellen von Takahashi et al. [101] und erzeugten realitätsnähere lineare Alterungsfunktionen für dieselben materialdefinierenden Parameter (vgl. Abbildung 66). Diese Alterungsfunktionen werden in dieser Arbeit verwendet. Dokko et al. [29] schlossen einzelne Studien mit einer sehr hohen Varianz in den experimentellen Ergebnissen und Daten von Testproben unter 15 Jahren aus. Beide Veröffentlichungen enthalten keine Informationen bezüglich des Zustands der Testproben in den experimentellen Studien. Es ist jedoch erwiesen, dass sich die physikalischen Eigenschaften der Knochensubstanz abhängig von ihrem Zustand ändert. Die Eigenschaften von z.B. formalinfixierten Knochen oder ausgetrockneten Knochen unterscheiden sich beispielsweise zu hydrierten Knochen oder bereits gefroren gewesenen Knochen [46, 111]. Diese Aspekte sollten zukünftig bei der Auswahl von experimentellen Studien bezüglich Materialeigenschaften mitberücksichtigt werden. Zudem sollten in Zukunft die Daten neuerer Studien, wie z.B. von Mirzaali et al. [79], ebenfalls in die Betrachtung miteinbezogen werden.

Die Definition der materialdefinierenden Parameter erfolgt auf Basis der Verteilungen bei einem Alter von 35 Jahren (vgl. Abbildung 66). Der Altersdurchschnitt der getesteten Testobjekte von Forman et al. [38] lag bei 31,5 Jahre. Da eine Überlappung der numerischen- und experimentellen Population erreicht werden soll, wird ein Alter von 35 Jahren ausgewählt, da es nah am experimentellen Durchschnittsalter liegt. Zudem ist der experimentell abgedeckte Bereich sehr groß (vgl. Abbildung 85). Eine andere Möglichkeit zur Materialdefinition wäre gewesen, alle verfügbaren Messpunkte in den Diagrammen von Abbildung 66 in Histogramme zu überführen und analog dem Vorgehen bei der Definition der geometriedefinierenden Variablen zu verfahren. Dieser Ansatz wurde jedoch nicht verfolgt, da jeder materialdefinierende Parameter eine unterschiedliche Anzahl an Messpunkten und eine unterschiedliche Anzahl an Studien enthält. Es kann kein Schluss dahingehend gezogen werden, dass diese Histogramme die reale Verteilung der Bevölkerungsvariabilität abbilden.

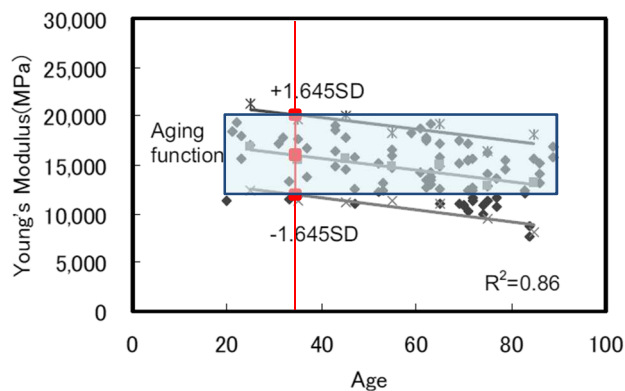


Abbildung 85 – Abgedeckter Wertebereich des E-Moduls bei 35 Jahren [29]

Eine dritte Möglichkeit der Materialdefinition, welche u.a. in Park et al. [85] verwendet wurde, ist die gemeinschaftliche Betrachtung von Messwerten aller Röhrenknochen der unteren Extremitäten, um die Datenbasis zu erweitern und eine realitätsnähere Verteilung zu erhalten. Ob dieses Vorgehen gerechtfertigt ist, bleibt fragwürdig. Es wird keine wissenschaftliche Veröffentlichung gefunden, welche dieses Vorgehen rechtfertigt. Yamada et al. [111] veröffentlichten Werte für die Versagensspannung für Femur, Tibia, Fibula, Humerus, Radius und Ulna. Wenn diese Werte gegen den Streuungskorridor von Dokko et al. [29] in Abbildung 66 verglichen werden, sind sie innerhalb des Korridors. Dies könnte als Indiz verwendet werden, das für eine gemeinschaftliche Betrachtung der Knochen spricht. Andererseits beschreiben Benninghoff et al. [11], „dass das Knochengewebe selbst in der Substantia compacta und in derselben Querschnittshöhe eines Röhrenknochens keineswegs überall die gleiche Festigkeit besitzt. Die Knochenfestigkeit geht weitgehend parallel zum Mineralsalzgehalt, und dieser drückt sich wiederum in der Röntgendichte eines Knochens aus“ [11]. Diese Aussage widerspricht deutlich dem Vorgehen, dass Werte unterschiedlicher Knochen gemeinschaftlich als homogenes Kontinuum betrachtet werden können. Vielmehr wird die Aussage gemacht, dass selbst eine homogene Materialformulierung innerhalb eines Knochens, wie sie auch in dieser Arbeit verwendet wird, bereits eine deutlich vereinfachte Modellierung darstellt.

In dieser Arbeit steht die Entwicklung der populationsbasierten Validierungsmethodik im Fokus. Es wird auf die Entwicklung eines parametrisierten Femur Modells verzichtet, da bereits viele Ansätze diesbezüglich u.a. von Park et al. [86], Bryan et al. [14], Klein et al. [68] existieren. Es wird eine geringe Stichprobenanzahl von acht Stichproben gewählt, um den Arbeitsaufwand zur manuellen Entwicklung der numerischen Femora Modelle überschaubar zu halten. Diese Stichprobenanzahl von acht Stichproben reduziert sich auf fünf Stichproben, aufgrund der geringen Anzahl von 18 Femora der verwendeten Datenbasis, welche für die Stichprobenauswahl genutzt werden kann. Für jede definierte

Stichprobe wird das ähnlichste zur Verfügung stehende Femur der CT-Datenbasis ausgewählt. Wenn ein Femur bereits einer anderen Stichprobe zugeteilt ist, wird das zweitähnlichste Femur ausgewählt. Wenn dieses ebenfalls bereits einer anderen Stichprobe zugeteilt ist, dann wird diese Stichprobe nicht mehr berücksichtigt. Auf diese Weise verringerte sich die Stichprobenanzahl auf fünf Stichproben. Es kann daher in diesem Rahmen keine konkrete Empfehlung gegeben werden, wie groß eine Population zukünftig sein sollte, um die Bevölkerungsvarianz detailliert abzubilden. Dieser Aspekt sollte in zukünftigen Studien separat betrachtet werden.

Ziel der Populationsdefinition ist es, eine Population von Femora zu erzeugen, die dieselbe Varianz einer Referenz-Bevölkerungsgruppe aufweist. Wenn nun die Soll-Populationen der einzelnen Parameter mit den Referenzdatenbanken verglichen werden (vgl. Diagramme der linken Spalte in Abbildung 71) wird deutlich, dass eine Stichprobenanzahl von acht Stichproben nur im Ansatz eine Verteilung widerspiegeln kann. Eine höhere Anzahl an Stichproben wäre vorteilhaft, um die reale Varianz detaillierter abbilden zu können, wengleich dies die manuelle Prüfung der Vorgehensweise der populationsbasierten Validierung erschweren würde. Es wird daher die Empfehlung gegeben, in zukünftigen Studien eine Stichprobenanzahl von $n > 8$ zu wählen. Da diejenigen Femora der verwendeten Datenbasis von 18 Femora ausgewählt werden, welche der Soll-Population am ähnlichsten sind, werden die Ist-Verteilungen der einzelnen Parameter in der rechten Spalte in Abbildung 71 ebenfalls mit den Referenzdatenbanken verglichen. Bezüglich der beiden Parameter Femurlänge und Innendurchmesser kann die zugrundeliegende Referenzverteilung grob abgebildet werden, wohingegen bei dem Parameter der kortikalen Dicke eine signifikante Abweichung zwischen den Verteilungen der Referenz-Population und der Ist-Population ersichtlich ist. Die erzeugte Population bildet bez. der kortikalen Dicke nicht dieselbe Varianz der Referenz-Bevölkerungsgruppe ab. Sie bleibt jedoch legitim, da sie auf existierenden Individuen basiert.

Für zukünftige Studien wird empfohlen, numerische Populationen anhand von parametrisierten Modellen zu generieren, um eine bessere Übereinstimmung zwischen Soll- und Ist-Population zu erreichen.

Es bleibt noch die Erklärung offen, weshalb im Rahmen dieser Arbeit verschiedene Quellen für die Populationsdefinition verwendet werden. Die Gründe hierfür liegen in einer anwenderorientierten Sichtweise. Im universitären Bereich ist es möglich, auf unkomplizierte Weise Zugang zu Datenbanken zu erhalten, welche beispielsweise eine große Anzahl an vermessenen Femora beinhalten. Die Definition der geometriedefinierenden Parameter anhand einer solchen Datenbank ist genauer, da alle Parameter und die abgeleiteten Verteilungen von denselben Objekten stammen. Wenn der Zugang zu dieser Art von Datenbanken gegeben ist, wird empfohlen, diese in zukünftigen Studien zu verwenden. FE-Menschmodelle werden entwickelt, um sie bei der Bewertung neuer Sicherheitstechnologien in Entwicklungsprozessen der Fahrzeugsicherheit v.a. in der Industrie einzusetzen. Daher finden auch im industriellen Bereich vermehrt Validierungsaktivitäten hinsichtlich FE-Menschmodellen statt. Deshalb wird im Rahmen dieser Arbeit der Ansatz verfolgt, dass die verwendeten Datenbanken öffentlich verfügbar sind oder die verwendeten Daten aus veröffentlichter Literatur stammen. Aus industrieller Sichtweise heraus ist es jedoch aktuell nicht möglich, auf unkomplizierte Weise Zugang zu dieser Art von anthropometrischen Datenbanken zu erhalten. In diesem Fall wird für zukünftige Studien die Verwendung von parametrisierten FE-Menschmodellen empfohlen.

5.4.3 Erzeugung der Population und Definition der Konzeptdatensätze

In diesem Abschnitt werden die Modellierung der Population und die Definition der Konzeptdatensätze diskutiert.

Die Struktur der entwickelten Knochenmodelle, den Femurschaft als Hohlkörper zu definieren und das proximale und distale Ende des Femurs mit Spongiosa zu modellieren, kann als gängige Modellierungsvariante bewertet werden. Das Femur des GHBM-Modells, sowie die entwickelten Femora in Takahashi et al. [101] und Park et al. [85, 86] wurden ebenfalls auf dieselbe Weise modelliert.

Am proximalen und distalen Ende des Femurs wird die kortikale Schicht mit 2mm Dicke modelliert, da aufgrund der groben Auflösung der CT-Aufnahmen die kortikale Schicht in diesen Bereichen nicht extrahiert werden konnte. Nach Shigeta et al. [96] besitzt das THUMS-Modell in der Mitte des Femurschaftes eine kortikale Dicke von 5 mm und am Femurkopf eine Dicke von 1 mm. Es ist davon auszugehen, dass mit einer kortikalen Schicht von 2 mm der Knochen an den Epiphysen zu steif modelliert ist. Dieser Modellierungsfehler ist in diesem Fall als vernachlässigbar anzunehmen, da die Epiphysen in den Auflagern fest gelagert sind und nicht betrachtet werden (vgl. Anhang A5). In diesen Übergangsbereichen, zwischen Femurschaft und den Epiphysen, ist eine Dicke von 2 mm in Einklang zu den existierenden FE-Menschmodellen, wie z.B. dem GHBM- und dem THUMS-Modell.

An den Übergängen des Femurschaftes zu den Epiphysen entsteht, aufgrund der abrupten Dickenänderung der kortikalen Dicke auf 2 mm, ein Steifigkeitssprung (vgl. Abbildung 73). Während der Erzeugung der numerischen Modellvorhersagen für die generierten Modelle, wird dieser Aspekt genauer untersucht. Die numerischen Modelle werden dem dynamischen 3-Punkt-Biegeversuch nach Forman et al. [38] bis zur Fraktur unterzogen. Die Fraktur wird während des Simulationsprozesses mit Element-Eliminierung erzeugt. Der Steifigkeitssprung wird als nicht relevant bewertet, wenn das Löschen der Elemente zuerst in der Mitte des Femurschaftes oder gleichzeitig im Bereich des Steifigkeitssprungs und der Mitte des Femurschaftes auftritt. Dieses Kriterium wird von allen fünf Femurmodellen in allen 40 Simulationen erfüllt. Diese Untersuchung zeigt jedoch nur, dass der Steifigkeitssprung keine offensichtlichen Fehler während der Simulation erzeugt. Es ist jedoch davon auszugehen, dass der Steifigkeitssprung einen Einfluss auf das Simulationsergebnis besitzt. Es ist in diesem Rahmen jedoch keine Quantifizierung des Effektes möglich. Es wird daher ausdrücklich empfohlen, in zukünftigen Studien auf derartige Steifigkeitssprünge zu verzichten.

Nach der Erstellung des FE-Netzes der einzelnen Knochen erfolgt eine Überprüfung der Elementqualität, da Elemente schlechter Qualität eine Quelle für Ungenauigkeiten oder Fehler sein können [113]. Burkhart et al. [16] veröffentlichten Empfehlungen zur Überprüfung der FE-Netzqualität hinsichtlich der Modellierung von Knochengewebe. Eine weitere Quelle, welche Empfehlungen zur Überprüfung der FE-Netzqualität in Bezug auf die Entwicklung von FE-Menschmodellen enthält, ist das Buch von Yang [113]. Die Vernetzungsqualität der erstellten Knochenmodelle wird mit denselben Kriterien aus Park [85] überprüft (vgl. Anhang A6). Alle Modelle entsprechen den geprüften Kriterien. Die Kriterien zur Überprüfung der Vernetzungsqualität von Park [85] sind jedoch weniger streng als in Burkhart et al. [16] oder Yang [113]. Burkhart et al. [16] stuften auf Basis von Felippa [35] z.B. das Aspekt-Verhältnis eines Elements zwischen 1 bis 3 als *gut* und zwischen 3 bis 10 als *mit Vorsicht zu genießen* ein. Yang [113] empfiehlt ein Aspekt-Verhältnis unter 5. Vier der fünf Femurmodelle besitzen ein Aspekt-Verhältnis im Bereich 3 bis 5 und ein Femurmodell liegt im Bereich 5 bis 8. Hinsichtlich der Jacobi-Determinante empfiehlt Yang [113] einen Wert größer 0,4, wohingegen Park [85] gegen den Wert größer 0,3 prüft. Die FE-Femurmodelle besitzen alle eine min. Jacobi-Determinante im Bereich 0,3 bis 0,6. Sie würden der strengeren Bewertung von Yang [113] nicht standhalten. Da keine allgemeingültigen Empfehlungen für die Elementqualität eines FE-Netzes existieren und die individuellen Empfehlungen stark abweichen, kann zusammengefasst werden, dass die Netzqualität

der Femurmodelle akzeptabel ist. Zukünftig ist es jedoch empfehlenswert, eine bessere Netzqualität anzustreben.

Die erzeugten Knochenmodelle enthalten Elemente stark unterschiedlicher Größe. Vor allem in den Epiphysen des Femurs existieren große Elemente. Der verursachte Fehler kann in diesem Fall vernachlässigt werden, da diese Bereiche des Femurs in den Auflagern eingebettet sind und in der Deformation nicht betrachtet werden. In Zukunft ist eine gleichmäßige Elementgröße innerhalb eines FE-Modells anzustreben.

Die relativ große Ziel-Elementgröße von 3mm wird gewählt, da existierende FE-Menschmodelle, wie das THUMS- und das GHBMC-Modell, ebenfalls in diesem Bereich modelliert sind. FE-Menschmodelle, welche zur Anwendung in Entwicklungsprozessen in der Fahrzeugsicherheit konzipiert wurden, können nicht beliebig verfeinert werden. Diese Modelle wurden entwickelt, um in kompletten Fahrzeugumgebungen angewendet zu werden. Eine zu geringe Elementgröße führt zu einer starken Reduktion des Lösungszeitschritts, was wiederum zu einem unververtretbaren Simulationsaufwand führen würde. Es wird im Rahmen dieser Arbeit auf eine Netzkonvergenzstudie zur Fehlereinschätzung bewusst verzichtet, da in Park [85] eine Netzkonvergenzstudie für denselben Referenzlastfall und einer sehr ähnlichen Femur-Modellierung durchgeführt wurde. Um eine Netzkonvergenz zu erreichen, musste in Park [85] das FE-Netz in der Mitte des Femurschaftes mit sieben Elementen über die kortikale Dicke diskretisiert werden. Bei einer Diskretisierung von zwei Elementen über die kortikale Dicke, wie sie in dieser Arbeit verwendet wird, berichtete Park [85] eine 1%-ige Abweichung der realen Frakturkraft. Dieser Fehler wird in dieser Arbeit akzeptiert.

Konzeptdatensatz I (vgl. Abbildung 76) stellt einen kombinierten Datensatz dar, der die gesamte numerisch erstellte Population und alle experimentellen Kurven enthält. Von der Struktur ist Konzeptdatensatz I ein eher undurchsichtiger Datensatz, der bewusst ohne natürliche Gruppierungen in den Kurven erzeugt wird. Konzeptdatensatz II (vgl. Abbildung 77) ist ein kombinierter Datensatz mit wenigen Kurven, welcher künstlich von Konzeptdatensatz I abgeleitet wird und drei natürliche Gruppierungen aufweist. Auf diese Weise kann die entwickelte populationsbasierte Validierungsmethode an zwei Konzeptdatensätzen angewendet werden, welche gegensätzliche Charakteristiken aufweisen.

5.4.4 Validierungsergebnisse der Konzeptdatensätze

In diesem Abschnitt werden die Validierungsergebnisse der beiden Konzeptdatensätze diskutiert. Angenommen, ein Stahlrohr wird in Auflagern gelagert und erfährt mittig einen Anstoß, ähnlich dem Referenzexperiment aus Kapitel 3, dann wird das Stahlrohr einer Biegebelastung unterzogen. Die Resistenz eines Stahlrohrs gegen diese Art von Belastung unter Biegung ist dabei abhängig von der freien Länge des Stahlrohrs zwischen den Auflagern, des Flächenträgheitsmomentes des Stahlrohrquerschnitts in der Belastungsebene und der Materialfestigkeit des Stahlrohrs. Die Resistenz eines Stahlrohrs gegen Biegebelastung wird verringert, wenn die freie Länge zwischen den Auflagern zunimmt, das Flächenträgheitsmoment verringert wird, oder die Materialsteifigkeit reduziert wird.

Ein Röhrenknochen, wie beispielsweise das Femur, besitzt eine deutlich komplexere Struktur als ein simples Stahlrohr. Die Grundgesetze der Mechanik der Biegebelastung und Biegeresistenz sind jedoch übertragbar für eine Plausibilitätsprüfung der numerischen Modellvorhersagen nach einer Clusteranalyse. Ein hohes Flächenträgheitsmoment in der Mitte des Femurschaftes erhöht beispielsweise die Resistenz gegen Biegung und führt zu höheren Frakturkräften als ein Knochen mit sehr geringem Flächenträgheitsmoment.

Die Wertebereiche der einzelnen Cluster werden bei beiden Konzeptdatensätzen gegen Referenz-Datenbanken verglichen. Für die Femurlänge werden im Sinne einer Referenzdatenbank die publizierten Femurlängen von Park et al. [86] verwendet. Die Wahl der Referenzdatenbank ist stark abhängig vom jeweiligen Analyseziel. In dieser Arbeit wird dieselbe Referenzdatenbank für den Vergleich der Wertebereiche der einzelnen Cluster verwendet, welche auch zur Populationsdefinition herangezogen wird. Es wird überprüft, welche Bereiche des Histogramms von welchem Cluster abgedeckt werden. Es könnte jedoch ebenfalls eine Populationsdefinition auf Basis einer Verteilung einer speziellen Bevölkerungsgruppe erfolgen, welche später gegen die Verteilung der gesamten Bevölkerung verglichen wird. In diesem Fall würden für die Populationsdefinition und für den Referenz-Vergleich verschiedene Datenbanken verwendet werden.

Das HSA-Programm der NHANES-Studie veröffentlichte Flächenträgheitsmomente im proximalen Bereich des Femurschaftes (vgl. Abbildung 67). Wenn diese Werte mit den Flächenträgheitsmomenten in der Mitte des Femurschaftes der experimentellen Testproben von Forman et al. [38] verglichen werden, unterscheiden sich die Werte signifikant (vgl. linkes Diagramm in Abbildung 86). Ein Grund dafür könnte sein, dass die beiden Datenquellen den Knochenquerschnitt an einer unterschiedlichen Position des Femurs untersuchen. Ein weiterer Grund könnte die vereinfachte Berechnungsmethode für das Flächenträgheitsmoment auf Basis von Kreisringquerschnitten des HSA-Programms sein [10]. Aufgrund der starken Diskrepanz von experimentell in [38] ermittelten Werten und den Werten des HSA-Programms, wird für das Flächenträgheitsmoment in der Mitte des Femurschaftes eine Datenbank aus den folgenden Veröffentlichungen erzeugt: Ivarsson et al. [58], Funk et al. [42], Kennedy et al. [62] und Kerrigan et al. [65]. Die erzeugte Referenz-Datenbank enthält 91 Datenpunkte und stimmt zufriedenstellend mit den experimentellen Ergebnissen überein (vgl. rechtes Diagramm in Abbildung 86).

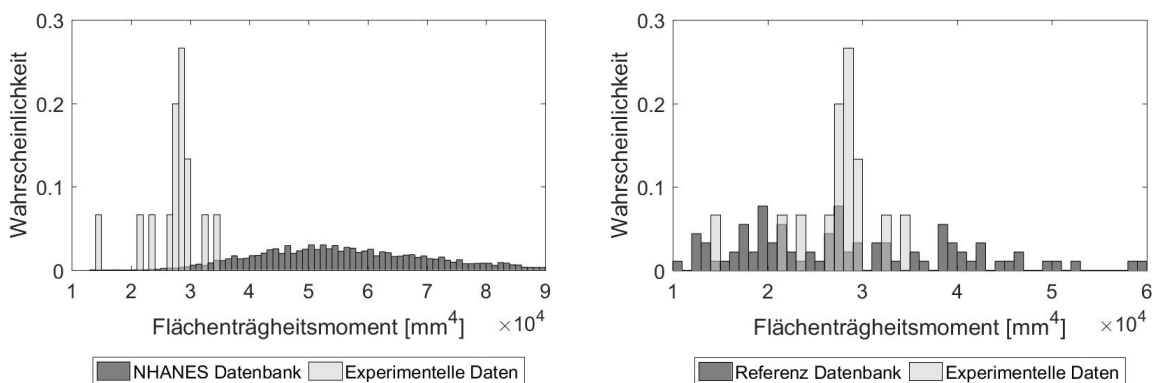


Abbildung 86 – Vergleich der Flächenträgheitsmomente zwischen HSA-Programm und anderen Quellen

Der Konzeptdatensatz II weist mit einem Validierungsergebnis von 7 der 8 möglichen Punkte eine hohe Validierungsgüte auf. Die folgenden Plausibilitätsbetrachtungen erfolgen auf Basis von Abbildung 82 bis Abbildung 84. Ein sehr niedriges Flächenträgheitsmoment (vgl. Cluster 2) führt zu einer sehr geringen Frakturkraft. Ein sehr hohes Flächenträgheitsmoment (vgl. Cluster 3 und Cluster 4) führt dagegen zu einer sehr hohen Frakturkraft. Die Femurlängen und die Werte für die Energieaufnahmekapazität des Knochens von Cluster 3 und Cluster 4 fallen in den Wertebereich des Cluster 1. Daher kann die Vermutung geäußert werden, dass das Flächenträgheitsmoment der einflussreichste Parameter gegen Biegeresistenz ist und die beiden anderen Parameter eine untergeordnete Rolle spielen. Diese Erkenntnis stimmt mit den Erkenntnissen der Studie in Funk et al. [42] überein, dass die geometrischen Aspekte einen signifikanten Einfluss haben, die materialspezifischen Aspekte jedoch nicht. Das Cluster 3 und das Cluster 4 unterscheiden sich in der Frakturkraft und -Zeit. Der einzige Unterschied ist die Energieaufnahmekapazität des Knochens. Obwohl das Cluster 3 die höhere Frakturkraft und den späteren Frakturzeitpunkt aufweist, weist es die

geringere Energieaufnahmekapazität des Knochens auf. Dieser Aspekt kann nicht physikalisch erklärt werden. Es wird vermutet, dass dieser Aspekt durch die Dynamik des Experiments zustande kommt. Durch die Dynamik des Experiments werden lokale Schwingungen produziert, welche die Modellvorhersage mit beeinflussen. Zusammengefasst kann gesagt werden, dass dieser Konzeptdatensatz II eine sehr gute Validierungsgüte besitzt, die anhand der Überschaubarkeit des Datensatzes visuell gut nachvollziehbar ist. Dieser Datensatz enthält deutliche Gruppierungen, die durch das Clusterverfahren ebenfalls richtig erkannt werden. Zudem existiert weder in den Metadaten noch in den Kurven selbst eine hohe Streuung in den einzelnen Clustern.

Bei dieser Güte eines Datensatzes können durch die Anwendung von Clusterverfahren auch weitere Erkenntnisse bez. der einzelnen Cluster mittels Vergleich gegen Referenz-Datenbanken gewonnen werden. Abbildung 84 zeigt sehr deutlich bez. des Flächenträgheitsmomentes, dass beispielweise Cluster 2 eher den linken Rand der Verteilung der Referenzdatenbank abbildet, wohingegen das Cluster 3 und 4 eher den rechten Rand der Verteilung der Referenzdatenbank abbildet. Cluster 3 befindet sich weitestgehend in der Mitte der Verteilung.

Die folgenden Ausführungen bezüglich des Validierungsprozesses von Konzeptdatensatz I beziehen sich auf Abbildung 78 bis Abbildung 81. Der Validierungsprozess von Konzeptdatensatz I muss aufgrund starker Streuung in den einzelnen Clustern und in den Metadaten abgebrochen werden. Dies bedeutet, dass die Qualität des Clusteranalyseergebnisses keinen Erkenntnisgewinn, d.h. keine Plausibilitätsprüfung der numerischen Modellvorhersagen, zulässt. Es kann keine Aussage getroffen werden, weshalb bestimmte Kurven in einem Cluster zusammengefasst sind, da alle Cluster ähnliche Wertebereiche in den Metadaten-Variablen aufweisen. In diesem Fall erweist sich die Clusteraufteilung als nicht geeignet für einen Erkenntnisgewinn.

Es ist ersichtlich, dass die Cluster hinsichtlich der Kurvenlänge gebildet werden, hinsichtlich der Frakturkraft streuen die Kurven stark. Die signifikanten Streuungen bez. der Frakturkraft in den Clustern werden hauptsächlich durch das verwendete Distanzmaß und den Clusteranalyseprozess erzeugt. Der Clusteranalyseprozess verlangt, dass vor Beginn des Prozesses alle Kurven auf dieselbe Kurvenlänge gebracht werden müssen. Dies bedeutet, dass alle Kurven außer der längsten Kurve des kombinierten Datensatzes auf die Kurvenlänge der längsten Kurve extrapoliert werden müssen. Diese kürzeren Kurven werden mit Null-Werte aufgefüllt, bis sie die Kurvenlänge der längsten Kurve erreichen. Durch die zusätzliche Verwendung der euklidischen Distanzmetrik fallen nun die Enden der kürzeren Kurven stärker ins Gewicht als Kurven mit ähnlichem Frakturzeitpunkt. Dieses Vorgehen ist maßgeblich dafür verantwortlich, dass die Kurven hauptsächlich bez. des Frakturzeitpunktes gruppiert werden, jedoch weniger hinsichtlich der Frakturkraft. Dies ist ein Hauptgrund für die signifikanten Streuungen in den Clustern von Konzeptdatensatz I (vgl. Abbildung 78). In Zukunft sollten Untersuchungen dahingehend erfolgen, wie für die spezielle Situation in verletzungsmechanischen Experimenten mit variierenden Versagenszeitpunkten die Distanzberechnung zwischen den Kurven geeigneter erfolgen kann, um die folgenden Faktoren mit gleicher Gewichtung in der Clusteranalyse zu berücksichtigen: Kurvenform, Frakturzeitpunkt und Frakturkraft.

Wenn der Validierungsprozess aufgrund starker Streuung in den Clustern abgebrochen wird, gibt es zwei Möglichkeiten zur weiteren Verfahrensweise. Es können die möglichen Kombinationen aus Fusionierungsalgorithmen und int. VIs aus Tabelle 25 erneut betrachtet werden (siehe Tabelle A7.1), ob eine Kombination existiert, welche eine höhere Clusteranzahl vorschlägt. Wenn dies der Fall ist, wird empfohlen, den Validierungsprozess erneut mit der höheren Clusteranzahl durchzuführen. Bei Konzeptdatensatz I werden bei der Kombination des Complete-Linkage-Verfahrens und des Beale-Index 12 Cluster berechnet. Bei dieser hohen Clusteranzahl empfiehlt es sich jedoch nicht, den

Validierungsprozess erneut zu durchlaufen, da das Dendrogramm an sehr kurzen Ästen geschnitten wird und die einzelnen Cluster nur noch sehr wenige Objekte enthalten würden.

Wenn es keine weiteren geeigneten Vorschläge für Clusteranzahlen gibt, oder es erneut zum Abbruch des Validierungsprozesses kommt, kann bei einer sehr ungleichmäßigen Dendrogramm-Struktur zwischen dem linken und rechten Dendrogramm-Ast wie folgt vorgegangen werden: jedes existierende Cluster aus der bereits durchgeführten Clusteranalyse kann erneut einer Clusteranalyse unterzogen werden. Es ist zu empfehlen, für diesen zweiten Clusteranalyse-Zyklus dieselbe Kombination aus Fusionierungsalgorithmus und int. VI wie bei der ersten Clusteranalyse zu verwenden. Danach wird der Validierungsprozess mit den neu definierten Clustern erneut durchgeführt. Dieser Schritt der erneuten Clusteranalyse sollte jedoch nur mit wenigen Zyklen und sehr großen Datensätzen durchgeführt werden, da bei jeder weiteren Clusteranalyse das Dendrogramm feiner aufgeteilt wird, die Anzahl der Cluster sich vervielfachen und sich der Erkenntnisgewinn somit ebenfalls erschwert. Es sollte bei jedem weiteren Clusteranalyse-Zyklus überprüft werden, ob die Streuung zum vorherigen Clusteranalyse-Zyklus abgenommen hat. Wenn die Streuung nicht abnimmt, könnte dies folgende zwei Ursachen haben: die Kurven sind innerhalb eines Clusters sehr homogen verteilt oder der Datensatz ist stark heterogen. Bei einem stark heterogenen Datensatz unterscheiden sich selbst die zusammengehörigen Kurvenpaare auf unterster Ebene des Dendrogramms signifikant voneinander. In diesem Fall wird der Erkenntnisgewinn erschwert, da die Clusteraufteilung immer eine Streuung enthält. Für den Konzeptdatensatz I kommt eine erneute Clusteranalyse nicht in Betracht, da das Dendrogramm eine sehr gleichmäßige Struktur aufweist und der Datensatz relativ klein ist (vgl. Abbildung A7.6).

Ein Grund für den Abbruch des Validierungsprozesses ist, dass der Datensatz eine sehr heterogene Struktur aufweist, bei der selbst die Kurvenpaare auf unterster Ebene des Dendrogramms einen signifikanten Unterschied zueinander aufweisen (vgl. Abbildung A7.3). Dies deutet auf eine unzureichende Güte der numerischen Modellvorhersagen hin, was wiederum als Indiz dafür gewertet werden könnte, dass die Modellierung der numerischen Femora, wie beispielsweise der existierende Steifigkeitssprung oder die variierende Elementgröße, einen negativen Einfluss auf die Güte der numerischen Modellvorhersagen hat. In Abbildung A7.3 ist ersichtlich, dass die numerischen Modellvorhersagen stärkere Schwingungen aufweisen als die experimentellen Ergebnisse. Der numerische Versuchsaufbau wird in Kapitel 3 umfassend validiert. Es kann jedoch nicht abschließend geklärt werden, ob die Schwingungen auf die Modellierung des Versuchsaufbaus oder der Testproben zurückzuführen sind. Zukünftig sollte daher auf eine hohe Modellgüte der numerischen Population geachtet werden, um das Validierungskonzept zufriedenstellend testen zu können. Außerdem würde eine Gestaltung des experimentellen Aufbaus, die unnötige Schwingungsphänomene in den mechanischen Antworten verhindert, die Anwendbarkeit des Validierungskonzeptes auf Basis von Clusteranalyseverfahren erleichtern.

Ein wesentlicher Unterschied zwischen den beiden Konzeptdatensätzen ist, dass Konzeptdatensatz II natürliche Gruppierungen enthält, wohingegen Konzeptdatensatz I einen zusammenhängenden und undurchsichtigen Datensatz darstellt. Es ist in weiteren Studien zu untersuchen, ob die Clusteranalyse sich besser zur Anwendung an Datensätzen eignet, welche wie Konzeptdatensatz II, bereits natürliche Gruppierungen aufweisen. Dies wäre z.B. der Fall, wenn in einem Datensatz gezielte Sub-Populationen enthalten sind.

5.4.5 Zusammenfassung und Ausblick

Im Rahmen dieser Arbeit wird ein Konzept eines populationsbasierten Validierungsansatzes auf Basis von Clusteranalyseverfahren für die Validierung stochastischer Simulationen entwickelt. Dieses Konzept wird mit zwei Konzeptdatensätzen anhand eines lokalen Lastfalls getestet. Ein Hauptansatzpunkt für weitere Untersuchungen ist die Art der Ermittlung der Distanz zwischen den Kurven eines Datensatzes, damit bei dem anschließenden Clusteranalyseverfahren der Frakturzeitpunkt, die Frakturkraft und die Kurvenform gleichmäßig in den Clusteranalyseprozess mit einfließen, um Cluster mit weniger Streuungen zu erzeugen. Zudem sollten zukünftige Untersuchungen dahingehend erfolgen, den gesamten Bewertungsprozess aus Abbildung 63 zu objektivieren, damit der gesamte Validierungsprozess automatisiert durchgeführt werden kann. Dieser Aspekt ist besonders wichtig, um größere stochastische Datensätze untersuchen zu können. Bereits in Konzeptdatensatz I (vgl. Abbildung 76) bei einer Kurvenschar von 40 Kurven ist erkennbar, dass der Datensatz undurchsichtig wirkt und sich bereits nur erschwert visuell untersuchen lässt.

In Zukunft sollte dieses Konzept an weiteren Lastfällen mit möglichst unterschiedlichen Kurvenformen getestet werden. Hierbei ist es empfehlenswert, das populationsbasierte Validierungskonzept zunächst an weiteren Lastfällen auf lokaler Ebene zu testen, bevor Lastfälle auf regionaler oder globaler Ebene mit mehreren Einflussfaktoren betrachtet werden (vgl. Abbildung 7).

Zudem wird empfohlen, für die weitere Überprüfung des Validierungskonzeptes Datensätze mit möglichst unterschiedlichen Datensatz-Strukturen zu verwenden, wie beispielsweise zusammenhängende Datensätze und Datensätze, welche bereits Sub-Populationen enthalten.

Abschließend wird empfohlen, zukünftig die numerischen Populationen mit parametrisierten Modellen zu erstellen und weitere Anstrengungen hinsichtlich der Objektivierung und Verfeinerung der Validierungsmethodik zu unternehmen.

6 Schlussfolgerungen und Ausblick

In dieser Arbeit wird ein Konzept für eine populationsbasierte Validierungsmethode auf Basis von Clusteranalyseverfahren entwickelt, um die Validierung stochastischer Simulationen mit FE-Menschmodellen gegen experimentelle Ergebnisse zu ermöglichen. Es wird damit eine Vorgehensweise vorgestellt, die gängige Validierungsansätze für Menschmodelle ergänzt. Die Methode verfolgt das Ziel, Vertrauen in numerische Modellvorhersagen zu generieren, bei denen die verwendeten numerischen Testproben abweichende Charakteristiken zu den experimentellen Testproben aufweisen. Das entwickelte Validierungskonzept wird an einem Referenzlastfall mit einer numerischen Testpopulation erfolgreich getestet.

Nennenswerte Vorteile dieses Validierungsansatzes gegenüber bereits existierenden Ansätzen sind, dass bei dieser Methode der komplette Kurvenverlauf der Modellvorhersagen berücksichtigt wird und dass eine Population an numerischen Modellvorhersagen mit unterschiedlichen Charakteristiken anhand eines Prozesses analysiert und validiert werden kann. Zudem eignet sich diese Methode dort, wo der subjekt-spezifische Validierungsansatz nicht verfolgt werden kann, weil keine detaillierten Informationen zur subjekt-spezifischen Replikation der Testproben vorhanden sind.

In Zukunft sollte das entwickelte Validierungskonzept an weiteren Lastfällen und auch Testpopulationen mit unterschiedlichen Datensatzstrukturen verifiziert werden. Es wird empfohlen, das Validierungskonzept zunächst an weiteren Lastfällen mit unterschiedlichen charakteristischen Kurvenverläufen auf lokaler Ebene zu testen. Danach kann das Validierungskonzept an Lastfällen aus anderen Validierungsebenen bis hin zu Ganzkörpervalidierungslastfällen überprüft werden (vgl. Abschnitt 1.2.2).

Mit dieser Validierungsmethode können nicht nur monoton ansteigende Kurven untersucht werden, wie beispielsweise Belastungskurven bis zu einem Versagenspunkt, sondern auch schleifenartige Kurven, wie beispielsweise Kinematik-Trajektorien.

Das entwickelte Validierungskonzept besitzt somit das Potenzial, parametrisierte Modelle zu validieren, um das Vertrauen in die Modellvorhersagen parametrisierter Modelle zu stärken und ihre Akzeptanz zu steigern. Für eine vollumfängliche Bewertung des Schutzpotenzials von Fahrzeugsicherheitssystemen im Fahrzeugentwicklungsprozess ist es wichtig, die Varianzen in der Bevölkerung zuverlässig numerisch abbilden und untersuchen zu können (vgl. Abschnitt 2.1). Dies ist nur mit validierten parametrisierten Modellen in Kombination mit stochastischen Methoden möglich.

Hinsichtlich der Weiterentwicklung des Clusteranalyse-Prozesses innerhalb des Validierungskonzeptes sind Untersuchungen bezüglich der Eignung von weiteren Clusteranalyseverfahren für die Anwendung an biomechanischen Datensätzen und weitere Möglichkeiten zur zuverlässigen Bestimmung einer geeigneten Clusteranzahl notwendig. Daneben ist ein wichtiger Aspekt die Automatisierung des gesamten Validierungsprozesses, damit zukünftig größere stochastische Datensätze ohne manuelles Eingreifen ausgewertet werden könnten.

Neben der Entwicklung eines Konzeptes für eine populationsbasierte Validierungsmethode und eines Konzeptes für die Überprüfung der Eignung von Clusteranalyseverfahren für die Anwendung an biomechanischen Datensätzen, beschäftigt sich diese Arbeit aus größerem Blickwinkel mit der Menschmodellvalidierung und mit der Erstellung eines numerischen Abbilds eines Referenzlastfalls als Validierungslastfall. In Abschnitt 1.2.3 wurde bereits die Problematik geschildert, dass unterschiedliche Menschmodelle mit einer unterschiedlichen Anzahl an Validierungslastfällen validiert sind und dass zudem dasselbe Modell für unterschiedliche numerische Solver eine unterschiedliche Anzahl an

Validierungslastfällen aufweist. Zudem wurde in Abschnitt 1.2.6 die Problematik bei der Erstellung von Validierungslastfällen auf Basis wissenschaftlicher Veröffentlichungen erläutert. Bei dieser Art von Validierungsaktivität sind oftmals unzureichend detaillierte oder mangelhafte Informationen für die Erstellung des numerischen Abbilds eines Validierungslastfalls verfügbar. Um eine hohe Aussagegüte numerischer Menschmodellvorhersagen für die Bewertung von Sicherheitssystemen in der Fahrzeugentwicklung zu erhalten, müssen verschiedene Menschmodelle dieselbe Aussagegüte, d.h. Validierungsqualität besitzen. Daher sind in Zukunft weitere Anstrengungen im Bereich der Harmonisierung von Validierungsaktivitäten hinsichtlich unterschiedlicher Menschmodelle, unterschiedlicher numerischer Solver und der Vereinheitlichung der numerischen Abbildungen von experimentellen Versuchen notwendig. Zudem sollten zukünftig durchgeführte experimentelle Studien detailliert veröffentlicht werden, mit dem Gedanken, dass ein akkurates numerisches Abbild der experimentellen Studie auf Basis der Veröffentlichung möglich ist.

Das Sicherstellen einer hohen Biofidelität der numerischen Menschmodelle ist eine notwendige Voraussetzung für die Bewertung von zukünftigen Insassenschutzkonzepten in der Fahrzeugsicherheit, wie beispielsweise im Bereich des automatisierten Fahrens, welche mit ATDs aufgrund der neuartigen Belastungsarten nicht mehr vollumfänglich bewertet werden können (vgl. Abschnitt 1.1.1).

Literaturverzeichnis

- [1] Acosta, S. M., Ash, J. H., Lessley, D. J., Shaw, C. G., Heltzel, S. B., and Crandall, J. R. 2016. Comparison of whole body response in oblique and full frontal sled tests. In *Proceedings of IRCOBI Conference*.
- [2] Aghabozorgi, S., Shirkhorshidi, A. S., and Wah, T. Y. 2015. Time-series clustering-A decade review. *Information Systems* 53, 16–38.
- [3] Anderson, A. E., Ellis, B. J., and Weiss, J. A. 2007. Verification, validation and sensitivity studies in computational biomechanics. *Computer methods in biomechanics and biomedical engineering* 10, 3, 171–184.
- [4] Arbelaitz, O., Gurrutxaga, I., Muguerza, J., Pérez, J. M., and Perona, I. 2013. An extensive comparative study of cluster validity indices. *Pattern Recognition* 46, 1, 243–256.
- [5] Bacher, J., Pöge, A., and Wenzig, K. 2011. *Clusteranalyse. Anwendungsorientierte Einführung in Klassifikationsverfahren*. Oldenbourg Verlag.
- [6] Backhaus, K., Erichson, B., Plinke, W., and Weiber, R. 2016. *Multivariate analysemethoden*. Springer.
- [7] Baker, F. B. and Hubert, L. J. 1975. Measuring the power of hierarchical cluster analysis. *Journal of the American statistical association* 70, 349, 31–38.
- [8] Ball, G. H. and Hall, D. J. 1965. *ISODATA, a novel method of data analysis and pattern classification*. Stanford research inst Menlo Park CA.
- [9] Beale, E. M. 1969. *Euclidean cluster analysis*. Scientific Control Systems Limited.
- [10] Beck, T. J. 2002. Hip structural analysis (HSA) program. *Baltimore: Johns Hopkins University*.
- [11] Benninghoff, A., Goerttler, K., and Staubesand, J. 1980. *Lehrbuch der Anatomie des Menschen. Makroskopische und mikroskopische Anatomie unter funktionellen Gesichtspunkten. Allgemeine Anatomie, Cytologie und Bewegungsapparat/von Jochen Staubesand*. Urban & Schwarzenberg.
- [12] Berndt, D. J. and Clifford, J. 1994. Using dynamic time warping to find patterns in time series. In *KDD workshop*, 359–370.
- [13] Besnault, B., Lavaste, F., Guillemot, H., Robin, S., and Le Coz, J.-Y. 1998. A parametric finite element model of the human pelvis. *SAE transactions*, 2680–2694.
- [14] Bryan, R., Mohan, P. S., Hopkins, A., Galloway, F., Taylor, M., and Nair, P. B. 2010. Statistical modelling of the whole human femur incorporating geometric and material properties. *Medical engineering & physics* 32, 1, 57–65.
- [15] Bryan, R., Nair, P. B., and Taylor, M. 2009. Use of a statistical model of the whole femur in a large scale, multi-model study of femoral neck fracture risk. *Journal of biomechanics* 42, 13, 2171–2176.
- [16] Burkhart, T. A., Andrews, D. M., and Dunning, C. E. 2013. Finite element modeling mesh quality, energy balance and validation methods. A review with recommendations associated with the modeling of bone tissue. *Journal of biomechanics* 46, 9, 1477–1488.
- [17] Caliński, T. and Harabasz, J. 1974. A dendrite method for cluster analysis. *Communications in Statistics-theory and Methods* 3, 1, 1–27.
- [18] Carter, P. M., Flannagan, C. A. C., Reed, M. P., Cunningham, R. M., and Rupp, J. D. 2014. Comparing the effects of age, BMI and gender on severe injury (AIS 3+) in motor-vehicle crashes. *Accident Analysis & Prevention* 72, 146–160.
- [19] CDC/National Center for Health Statistics. 2018. *NHANES Questionnaires, Datasets, and Related Documentation*. <https://wwwn.cdc.gov/nchs/nhanes/Default.aspx>. Accessed 8 December 2019.
- [20] Charrad, M., Ghazzali, N., Boiteau, V., and Niknafs, A. 2014. NbClust. An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software*.
- [21] 2009. *Clusteranalyse. Technische Universität München*. https://www-m9.ma.tum.de/material/felix-klein/clustering/Allgemeines/Was_ist_Clustering.php. Accessed 10 December 2019.
- [22] Combest, J. 2016. *Current status and future plans of the ghbmc (Global Human Body Models Consortium)*. http://www.ghbmc.com/wp-content/uploads/2016/10/HuMo16_15_Combest_GHBMC.pdf. Accessed 23 November 2019.

- [23] Combest, J. 2018. *CURRENT STATUS AND FUTURE PLANS OF THE GHBMCM. (Global Human Body Models Consortium)*. <http://www.ghbmc.com/status-of-the-ghbmc-project/>. Accessed 19 December 2019.
- [24] Cronin, D., Panzer, M., Vezin, P., Yoganandan, N., Broolin, K., Reed, M., Beillas, P., and Gayzik, S. 2018. *Human Body Modeling and Validation with Biomechanics Experiments. IRCOBI 2018 Workshop Series*. <http://www.ircobi.org/wordpress/downloads/ircobi-hbm-2018.pdf>. Accessed 19 December 2019.
- [25] Cronin, D. S. 2011. Explicit finite element method applied to impact biomechanics problems. *IRCOBI, Krakow*.
- [26] DATANOVIAM. *Determining The Optimal Number Of Clusters: 3 Must Know Methods*. <https://www.datanovia.com/en/lessons/determining-the-optimal-number-of-clusters-3-must-know-methods/>. Accessed 12 December 2019.
- [27] Davies, D. L. and Bouldin, D. W. 1979. A cluster separation measure. *IEEE transactions on pattern analysis and machine intelligence*, 2, 224–227.
- [28] Davis, M. L., Koya, B., Schap, J. M., and Gayzik, F. S. 2016. *Development and full body validation of a 5th percentile female finite element model*. SAE Technical Paper.
- [29] Dokko, Y., Ito, O., and Ohashi, K. 2009. *Development of human lower limb and pelvis fe models for adult and the elderly*. SAE Technical Paper.
- [30] Duda, R. O. and Hart, P. E. 1973. *Pattern recognition and scene analysis*. Wiley, New York.
- [31] Duda, R. O., Hart, P. E., and Stork, D. G. 1973. *Pattern classification and scene analysis* 3. Wiley New York.
- [32] Dunn, J. C. 1974. Well-separated clusters and optimal fuzzy partitions. *Journal of cybernetics* 4, 1, 95–104.
- [33] Ernst, J., Nau, G. J., and Bar-Joseph, Z. 2005. Clustering short time series gene expression data. *Bioinformatics* 21, suppl_1, i159-i168.
- [34] F. Khor, D.S. Cronin, B. Watson, D. Gierczycka, and S. Malcolm. Importance of asymmetry and anisotropy in predicting cortical bone response and fracture using human body model femur in three-point bending and axial rotation.
- [35] Felippa, C. 2016. Introduction to FEM, FEM Modeling. Mesh, Loads and BCs. *Lecture Notes, University of Colorado Boulder*.
- [36] Forman, J., Lopez-Valdes, F., Lessley, D., Kindig, M., Kent, R., Ridella, S., and Bostrom, O. 2009. *Rear seat occupant safety. An investigation of a progressive force-limiting, pretensioning 3-point belt system using adult PMHS in frontal sled tests*. SAE Technical Paper.
- [37] Forman, J., Poplin, G. S., Shaw, C. G., McMurry, T. L., Schmidt, K., Ash, J., and Sunnevang, C. 2019. Automobile injury trends in the contemporary fleet. Belted occupants in frontal collisions. *Traffic injury prevention* 20, 6, 607–612.
- [38] Forman, J. L., Dios, E. de, Symeonidis, I., Duart, J., Kerrigan, J. R., Salzar, R. S., Balasubramanian, S., Segui-Gomez, M., and Kent, R. W. 2012. Fracture tolerance related to skeletal development and aging throughout life:3-point bending of human femurs. In *IRCOBI Conference Proceedings, Dublin, Ireland*, 529–539.
- [39] 2019. *Formelsammlung und Berechnungsprogramme Anlagenbau. Reibwerte von verschiedenen Materialien*. <https://www.schweizer-fn.de/stoff/reibwerte/reibwerte.php>. Accessed 23 September 2019.
- [40] Fuchs, T. 2018. *Objektivierung der Modellbildung von verletzungsmechanischen Experimenten für die Validierung von Finite-Elemente Menschmodellen*, Imu.
- [41] Fujita, A., Severino, P., Kojima, K., Sato, J. R., Patriota, A. G., and Miyano, S. 2012. Functional clustering of time series gene expression data by Granger causality. *BMC systems biology* 6, 1, 137.
- [42] Funk, J. R., Kerrigan, J. R., and Crandall, J. R. 2004. Dynamic bending tolerance and elastic-plastic material properties of the human femur. In *Annual Proceedings/Association for the Advancement of Automotive Medicine*, 215.
- [43] Gehre, C., Gades, H., and Wernicke, P. 2009. Objective rating of signals using test and simulation responses. In *Proceedings: International Technical Conference on the Enhanced Safety of Vehicles*.
- [44] Gehre, Christian, Heinrich Gades, and Philipp Wernicke, Ed. 2009. *Objective rating of signals using test and simulation responses*. National Highway Traffic Safety Administration.
- [45] Global Human Body Models Consortium, LLC. *User Manual: M50 Occupant Version 4.3 for PAM-CRASH. Distributed by Elemance, LLC* 2014.

- [46] Greenberg, S. W., Gonzalez, D., Gurdjian, E. S., and Thomas, L. M. 1968. *Changes in physical properties of bone between the in vivo, freshly dead, and embalmed conditions*. SAE Technical Paper.
- [47] Guevara Alvez, P. B. 2011. *Inference of a human brain fiber bundle atlas from high angular resolution diffusion imaging*, Paris 11.
- [48] Gullo, F., Ponti, G., Tagarelli, A., Tradigo, G., and Veltri, P. 2012. A time series approach for clustering mass spectrometry data. *Journal of Computational Science* 3, 5, 344–355.
- [49] Halkidi, M. and Vazirgiannis, M. 2001. Clustering validity assessment. Finding the optimal partitioning of a data set. In *Proceedings 2001 IEEE International Conference on Data Mining*, 187–194.
- [50] Halkidi, M., Vazirgiannis, M., and Batistakis, Y. 2000. Quality scheme assessment in the clustering process. In *European Conference on Principles of Data Mining and Knowledge Discovery*, 265–276.
- [51] Hartigan, J. A. 1975. *Clustering algorithms*. John Wiley & Sons, Inc.
- [52] Hayashi, S., Yasuki, T., and Kitagawa, Y. 2008. *Occupant kinematics and estimated effectiveness of side airbags in pole side impacts using a human FE model with internal organs*. SAE Technical Paper.
- [53] Helgason, B., Perilli, E., Schileo, E., Taddei, F., Brynjólfsson, S., and Viceconti, M. 2008. Mathematical relationships between bone density and mechanical properties. A literature review. *Clinical Biomechanics* 23, 2, 135–146.
- [54] <https://www.schaumstoffbedarf.de/content/7-schaumstoff-infothek>. <https://www.schaumstoffbedarf.de/content/7-schaumstoff-infothek>. Accessed 3 October 2019Uhr.
- [55] Hubert, L. J. and Levin, J. R. 1976. A general statistical framework for assessing categorical clustering in free recall. *Psychological bulletin* 83, 6, 1072.
- [56] ISO, T. R. 1999. 9790, Road vehicles-anthropomorphic side impact dummy-Lateral impact response requirements to assess the biofidelity of the dummy. *IO f. Standardization*, Ed.
- [57] Ivancic, P. C., Panjabi, M. M., Tominaga, Y., and Malcolmson, G. F. 2006. Predicting multiplanar cervical spine injury due to head-turned rear impacts using IV-NIC. *Traffic injury prevention* 7, 3, 264–275.
- [58] Ivarsson, B. J., Genovese, D., Crandall, J. R., Bolton, J. R., Untaroiu, C. D., and Bose, D. 2009. *The tolerance of the femoral shaft in combined axial compression and bending loading*. SAE Technical Paper.
- [59] Jaffredo, A. S., Potier, P., Robin, SR, Le Coz, J.-Y., and Lassau, J.-P. 2000. Cadaver lower limb dynamic response in inversion-eversion. In *IRCOBI Conference on the Biomechanics of Impact*, 183–194.
- [60] JSOL - Computer Aided Engineering Solution. *Total Human Model for Safety THUMS TM. THUMS Features*. <https://www.jsol-cae.com/en/product/structure/thums/feature/>. Accessed 20 December 2019.
- [61] Kato, D., Nakahira, Y., Atsumi, N., and Iwamoto, M. 2018. Development of human-body model THUMS version 6 containing muscle controllers and application to injury analysis in frontal collision after brake deceleration. In *2018 IRCOBI Conference Proceedings. IRCOBI*.
- [62] Kennedy, E. A. 2004. *Lateral and posterior dynamic bending of the mid-shaft femur. Fracture risk curves for the adult population*, Virginia Tech.
- [63] Kent, R., Lee, S.-H., Darvish, K., Wang, S., Poster, C. S., Lange, A. W., Brede, C., Lange, D., and Matsuoka, F. 2005. *Structural and material changes in the aging thorax and their role in crash protection for older occupants*. SAE Technical Paper.
- [64] Kent, R., Lessley, D., and Sherwood, C. 2004. *Thoracic response to dynamic, non-impact loading from a hub, distributed belt, diagonal belt, and double diagonal belts*. SAE Technical Paper.
- [65] Kerrigan, J. R., Drinkwater, D. C., Kam, C. Y., Murphy, D. B., Ivarsson, B. J., Crandall, J. R., and Patrie, J. 2004. Tolerance of the human leg and thigh in dynamic latero-medial bending. *International Journal of Crashworthiness* 9, 6, 607–623.
- [66] Kitagawa, Y. and Yasuki, T. 2014. Development and Application of THUMS Version 4. In *proceedings of 5th International Symposium of Human Modeling and Simulation in Automotive Engineering (Munich, Germany, Oct. 16-17), CARHS*.
- [67] Klein, K. F., Hu, J., Reed, M. P., Hoff, C. N., and Rupp, J. D. 2015. Development and validation of statistical models of femur geometry for use with parametric finite element models. *Annals of biomedical engineering* 43, 10, 2503–2514.
- [68] Klein, K. F., Hu, J., Reed, M. P., Schneider, L. W., and Rupp, J. D. 2017. Validation of a parametric finite element human femur model. *Traffic injury prevention* 18, 4, 420–426.

- [69] Kramer, F. 1998. *Passive Sicherheit von Kraftfahrzeugen*. Springer.
- [70] Krzanowski, W. J. and Lai, Y. T. 1988. A criterion for determining the number of groups in a data set using sum-of-squares clustering. *Biometrics*, 23–34.
- [71] Lamecker, H., Lange, T., and Seebass, M. 2004. Segmentation of the liver using a 3D statistical shape model.
- [72] Lamielle, S., Vezin, P., Verriest, J.-P., Petit, P., Trosseille, X., and Vallancien, G. 2008. *3D deformation and dynamics of the human cadaver abdomen under seatbelt loading*. SAE Technical Paper.
- [73] Laz, P. J. and Browne, M. 2010. A review of probabilistic analysis in orthopaedic biomechanics. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine* 224, 8, 927–943.
- [74] McClain, J. O. and Rao, V. R. 1975. Clustisz. A program to test for the quality of clustering of a set of objects. *Journal of Marketing Research*, 456–460.
- [75] McQuitty, L. L. 1966. Similarity analysis by reciprocal pairs for discrete and continuous data. *Educational and Psychological measurement* 26, 4, 825–831.
- [76] Milligan, G. W. 1980. An examination of the effect of six types of error perturbation on fifteen clustering algorithms. *Psychometrika* 45, 3, 325–342.
- [77] Milligan, G. W. 1981. A Monte Carlo study of thirty internal criterion measures for cluster analysis. *Psychometrika* 46, 2, 187–199.
- [78] Milligan, G. W. and Cooper, M. C. 1985. An examination of procedures for determining the number of clusters in a data set. *Psychometrika* 50, 2, 159–179.
- [79] Mirzaali, M. J., Schwiedrzik, J. J., Thaiwichai, S., Best, J. P., Michler, J., Zysset, P. K., and Wolfram, U. 2016. Mechanical properties of cortical bone and their relationships with age, gender, composition and microindentation properties in the elderly. *Bone* 93, 196–211.
- [80] Montero, P., Vilar, J. A., and others. 2014. TSclust. An R package for time series clustering. *Journal of Statistical Software* 62, 1, 1–43.
- [81] Morgan, E. F., Bayraktar, H. H., and Keaveny, T. M. 2003. Trabecular bone modulus-density relationships depend on anatomic site. *Journal of biomechanics* 36, 7, 897–904.
- [82] National Center for Health Statistics. *National Health and Nutrition Examination Survey. Overview*. https://www.cdc.gov/nchs/data/nhanes/nhanes_13_14/NHANES_Overview_Brochure.pdf. Accessed 8 December 2019.
- [83] Organization, W. H. 2001. *Men, ageing and health. Achieving health across the life span*. World Health Organization.
- [84] Park, B.-K. and Reed, M. P. 2015. Parametric body shape model of standing children aged 3-11 years. *Ergonomics* 58, 10, 1714–1725.
- [85] Park, G. May / 2017. *Injury Risk Functions Based on Responses of Population-Based Finite Element Models: Application to Femurs under Dynamic Loading*, University of Virginia.
- [86] Park, G., Kim, T., Forman, J., Panzer, M. B., and Crandall, J. R. 2017. Prediction of the structural response of the femoral shaft under dynamic loading using subject-specific finite element models. *Computer methods in biomechanics and biomedical engineering* 20, 11, 1151–1166.
- [87] Reed, M. P. and Rupp, J. D. 2013. An anthropometric comparison of current ATDs with the US adult population. *Traffic injury prevention* 14, 7, 703–705.
- [88] Rohlf, F. J. 1974. Methods of comparing classifications. *Annual Review of Ecology and Systematics* 5, 1, 101–113.
- [89] Rousseeuw, P. J. 1987. Silhouettes. A graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics* 20, 53–65.
- [90] Ruff, C. B. and Hayes, W. C. 1988. Sex differences in age-related remodeling of the femur and tibia. *Journal of Orthopaedic Research* 6, 6, 886–896.
- [91] Rupp, J. D., Reed, M. P., van Ee, C. A., Kuppa, S., Wang, S. C., Goulet, J. A., and Schneider, L. W. 2002. *The tolerance of the human hip to dynamic knee loading*. SAE Technical Paper.
- [92] Schneider, L. W. 1983. Development of anthropometrically based design specifications for an advanced adult anthropomorphic dummy family, volume 1. final report.

- [93] Schneider, S. 2018. *Introduction of a Population-Based Validation Method for Human Body Models*. CAHRS, Berlin.
- [94] Schwer, L. E. 2007. An overview of the PTC 60/V&V 10. Guide for verification and validation in computational solid mechanics. *Engineering with Computers* 23, 4, 245–252.
- [95] Shaw, G., Parent, D., Purtsezov, S., Lessley, D., Crandall, J., Kent, R., Guillemot, H., Ridella, S. A., Takounts, E., and Martin, P. 2009. *Impact response of restrained PMHS in frontal sled tests. Skeletal deformation patterns under seat belt loading*. SAE Technical Paper.
- [96] Shigeta, K., Kitagawa, Y., and Yasuki, T. 2009. Development of next generation human FE model capable of organ injury prediction. *Proceedings of the 21st Annual Enhanced Safety of Vehicles*, 15–18.
- [97] Shigeta, K., Kitagawa, Y., and Yasuki, T. 2009. Development of next generation human FE model capable of organ injury prediction. *Proceedings of the 21st Annual Enhanced Safety of Vehicles*, 15–18.
- [98] Shin, J., Yue, N., and Untaroiu, C. D. 2012. A finite element model of the foot and ankle for automotive impact applications. *Annals of biomedical engineering* 40, 12, 2519–2531.
- [99] Symeonidis, I., Kavadarli, G., Erich, S., Graw, M., and Peldschus, S. 2012. Analysis of the stability of PTW riders in autonomous braking scenarios. *Accident Analysis & Prevention* 49, 212–222.
- [100] Symeonidis, I., Kavadarli, G., Fraga, F., Lex, v. R., Schuller, E., and Peldschus, S. 2010. Characterisation of motorcyclist's upper body motion during braking manoeuvre. In *2010 IRCOBI Conference Proceedings Toyota Motor Corporation Britax Thatcham*.
- [101] Takahashi, Y., Kikuchi, Y., Konosu, A., and Ishikawa, H. 2000. *Development and validation of the finite element model for the human lower limb of pedestrians*. SAE Technical Paper.
- [102] 2001. *the pedestrian crash data study*. National Highway Traffic Safety Administration.
- [103] Tibshirani, R., Walther, G., and Hastie, T. 2001. Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 63, 2, 411–423.
- [104] Toyota Motor Corporation. *Documentation Total Human Model for Safety (THUMS) AM50 Pedestrian/Occupant Model Version 4.0*, 2011.
- [105] U.S. Department of Transportation. *Automated Driving Systems 2.0 - A Vision for Safety*. https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/13069a-ads2.0_090617_v9a_tag.pdf. Accessed 7 October 2019 Uhr.
- [106] U.S. Department of Transportation - NHTSA. 2015. *Motor Vehicle Traffic Crashes as a Leading Cause of Death in the United States, 2010 and 2011*. <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812203>. Accessed 29 December 2019.
- [107] Université Laval. 2008. *Analyse multivariable. Cluster analyses*. http://biol09.biol.umontreal.ca/ULaval08/Chapitre_3.pdf. Accessed 12 January 2020.
- [108] Wang, Y., Cao, L., Bai, Z., Reed, M. P., Rupp, J. D., Hoff, C. N., and Hu, J. 2016. A parametric ribcage geometry model accounting for variations among the adult population. *Journal of biomechanics* 49, 13, 2791–2798.
- [109] Ward Jr, J. H. 1963. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association* 58, 301, 236–244.
- [110] Wismüller, A., Lange, O., Dersch, D. R., Leinsinger, G. L., Hahn, K., Pütz, B., and Auer, D. 2002. Cluster analysis of biomedical image time-series. *International Journal of Computer Vision* 46, 2, 103–128.
- [111] Yamada, H., Evans, F. G., and others. 1970. Strength of biological materials.
- [112] Yang, K. H., Hu, J., White, N. A., King, A. I., Chou, C. C., and Prasad, P. 2006. *Development of numerical models for injury biomechanics research. A review of 50 years of publications in the Stapp Car Crash Conference*. SAE Technical Paper.
- [113] Yang, K.-H. 2017. *Basic finite element method as applied to injury biomechanics*. Academic Press.
- [114] Yates, K. M., Lu, Y.-C., and Untaroiu, C. D. 2016. Statistical shape analysis of the human spleen geometry for probabilistic occupant models. *Journal of biomechanics* 49, 9, 1540–1546.
- [115] Yoganandan, N., Nahum, A. M., and Melvin, J. W. 2014. *Accidental injury. Biomechanics and prevention*. Springer.
- [116] Zhan, Z., Fu, Y., and Yang, R.-J. 2011. *Enhanced error assessment of response time histories (EARTH) metric and calibration process*. SAE Technical Paper.

Anhang

A1 Übersicht der Validierungslastfälle des THUMS-Modells [66]

<p>1. Head Impact (Nahum et al. 1977)</p> <p>Cylindrical Impactor Impactor Weight: 5.6 kg Impact Velocity: 6.3 m/s Frankfurt Anatomic Plane</p>	<p>6. Thorax Belt (Cesari et al. 1990)</p> <p>Belt Loading Impactor Weight: 22.4 kg Impactor Velocity: 7.78 m/s</p>	<p>2. Head Rotation (Gemarelli et al. 1982)</p> <p>Angular acceleration Brain CoG Rotational center 91mm</p>	<p>7. Thorax Lateral Impact (Shaw et al. 2006)</p> <p>Cylindrical Impactor Impactor Diameter: 152 mm Impactor Weight: 23 kg Impact Velocity: 2.76 m/s</p>	<p>3. Head Lateral Drop (Vogannadan et al. 2004)</p> <p>Free fall Pad Load cell</p>	<p>8. Abdomen Impact (Cavanaugh et al. 1989)</p> <p>Bar Impactor Impactor Diameter: 25 mm Impactor Weight: 32 kg Impact Velocity: 6.1 m/s</p>	<p>4. Head-Neck Drop (Nightingale et al. 1997)</p> <p>Base (18kg) Impact plane (Angle 15(deg)) Load cell 3.18m/s</p>	<p>9. Abdomen Belt (Foster et al. 2006)</p> <p>Belt Loading Device Width: 50 mm Velocity: 6.9 m/s</p>	<p>5. Thorax Anterior Impact (Krcel et al. 1974)</p> <p>Cylindrical Impactor Impactor Diameter: 152 mm Impactor Weight: 23 kg Impact Velocity: 7.2 m/s</p>	<p>10. Knee-Thigh-Hip Impact (Rupp et al. 2008)</p> <p>Load cell PAD Table Impactor</p>	<p>11. Pelvis Impact (Guillemot et al. 1997)</p> <p>Impactor PAD Ball Fracture Fix</p>	<p>12. Humerus Bending (Kemper et al. 2005)</p> <p>Bar Loading Device Diameter: 20 mm Impact Velocity: 10 mm/s Loading Head Humerus</p>	<p>13. Humerus Compression (Kemper et al. 2005)</p> <p>Circular Impactor Diameter: 152 mm Compression: 4 m/s Flesh Humerus Circular Table Diameter: 152 mm</p>	<p>14. Femur Bending (Yamada et al. 1970)</p> <p>Bar Loading Device Diameter: 20 mm Impact Velocity: 10 mm/s Impactor Femur</p>	<p>15. Knee Lateral Impact (Kajzer et al. 1997)</p> <p>Impactor Loading Impactor Diameter: 100 mm Impactor Mass: 6.25 kg Impact Velocity: 40 km/h Fixed 0.470 N</p>	<p>16. Knee 4p Bending (Bose et al. 2004)</p> <p>Instron Actuator 100mm 150mm 300mm 450mm</p>	<p>17. Tibia Bending (Schreiber et al. 1997)</p> <p>Impactor Loading Impactor Diameter: 25 mm Impactor Mass: 9.48 kg Impact Velocity: 5.55 m/s</p>	<p>18. Foot-Ankle Impact (Kitagawa et al. 1998)</p> <p>Impactor (18kg) Sole plate Load cell #1 Load cell #2 Tension Force (1.8kN) offset (50mm) Stopper Fix</p>	<p>19. Head-Neck Bending (Stemper et al. 2004)</p> <p>Frankfort Plane T1 Sled AV = 2.6 m/s</p>	<p>20. Lumbar Spine Tension/Compression/Bending (Demetropoulos et al. 1998)</p> <p>Loading Rate 100 mm/s Fix Fix</p>
--	--	---	--	--	--	---	--	---	--	---	--	---	--	--	--	---	--	---	---

Abbildung A1.1 – Übersicht der lokalen bis regionalen Validierungslastfälle des THUMS-Modells

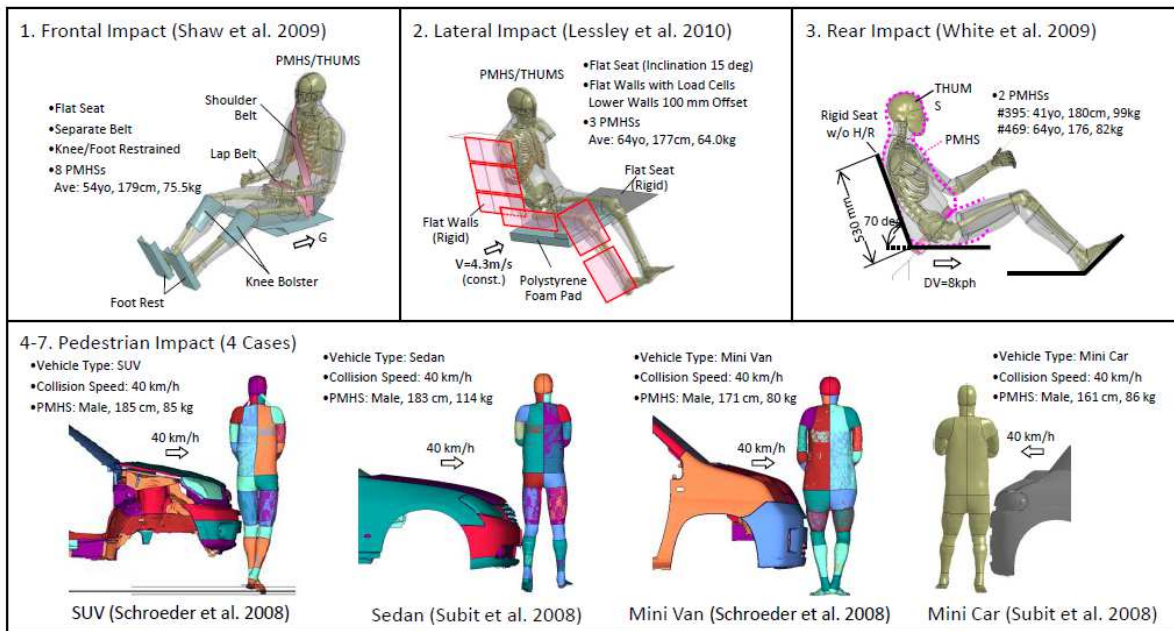
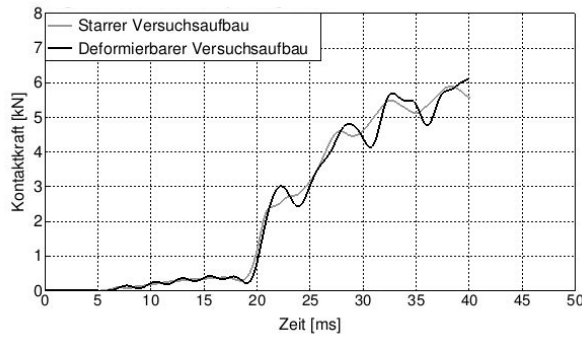


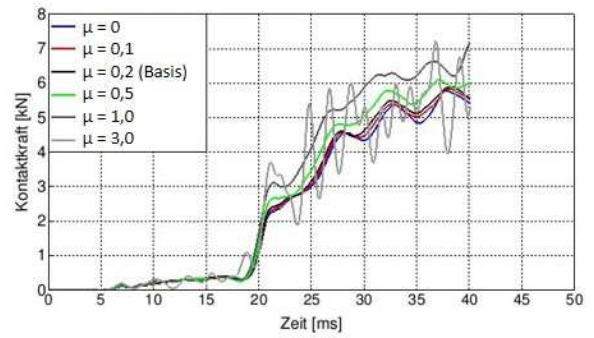
Abbildung A1.2 – Übersicht der Ganzkörper-Validierungslastfälle des THUMS-Modells [66]

A2 Ergebnisse Sensitivitätsanalysen

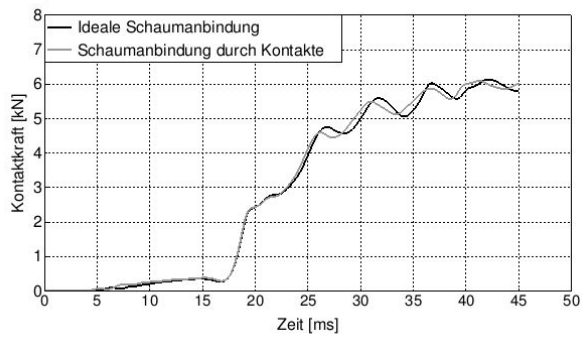
A2.1 Ergebnisse Sensitivitätsanalysen bezüglich Modellierungsannahmen



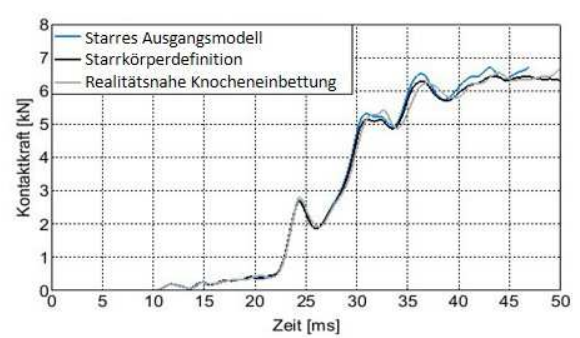
a) Starrer vs. deformierbarer Versuchsaufbau



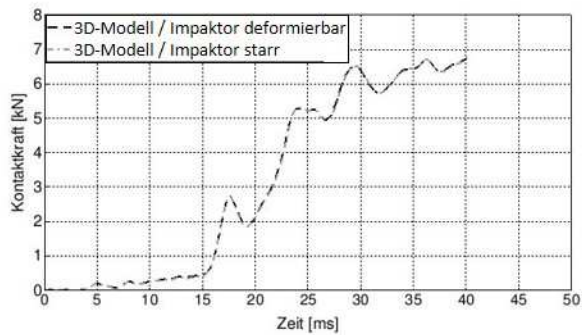
b) Variation Reibungskoeffizient zwischen Auflager und Bodenplatte



c) Modellierung der Schnittstelle Schaum – Impaktor



d) Vereinfachte Modellierung der Knocheneinbettung



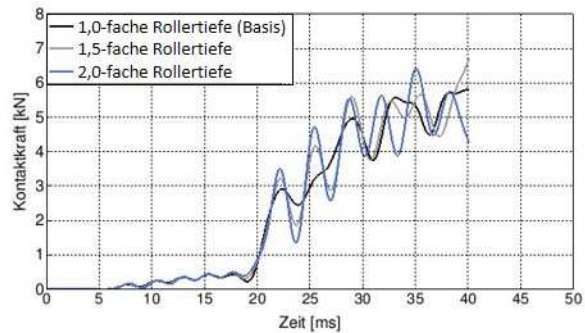
e) Vereinfachte Modellierung des Impaktors

Abbildung A2.1.1 – Überprüfung der Modellierungsfehler hinsichtlich der Randbedingungen

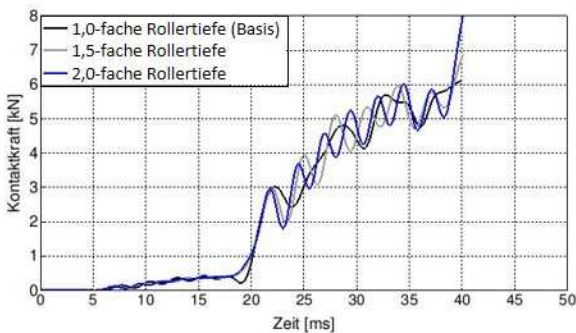
A2.2 Sensitivitätsanalysen zur Festlegung der geometrischen Parameter



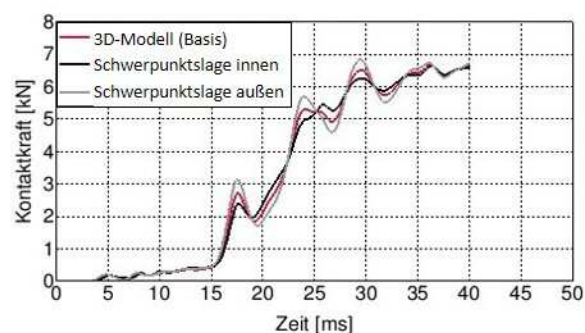
a) Variation der Struktursteifigkeit der Auflager



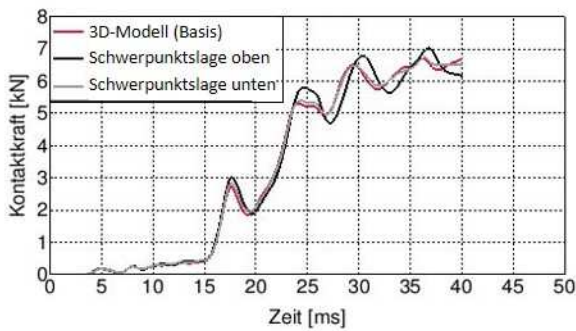
b) Variation Auflagertiefe mit Massenvariation



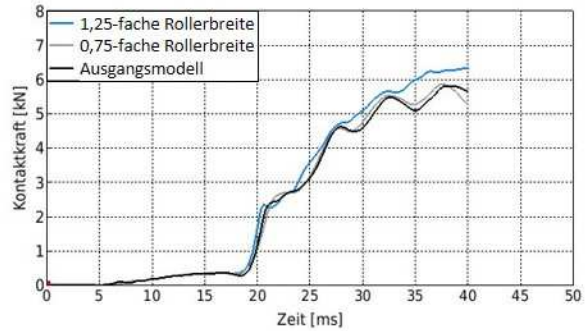
c) Variation der Auflagefläche der Auflager



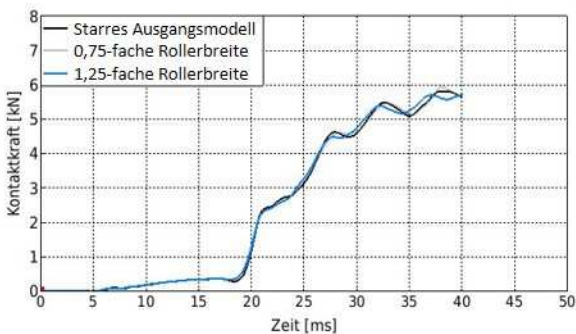
d) Variation der Schwerpunktslage der Auflager in x-Richtung



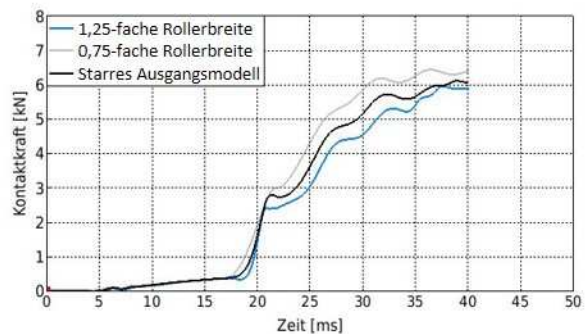
e) Variation der Schwerpunktslage der Auflager in z-Richtung



f) Variation des Radius der Auflager



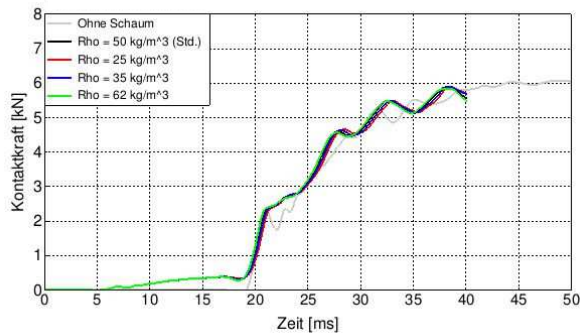
g) Variation Auflagerblockbreite



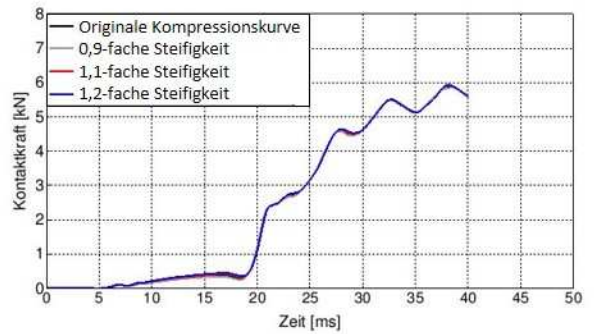
h) Variation Drehpunkt

Abbildung A2.2.1 – Sensitivitätsanalysen bezüglich geometrischer Parameter

A2.3 Sensitivitätsanalysen bezüglich der Schaumdefinition



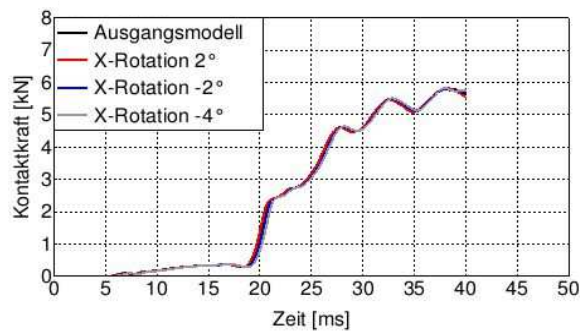
a) Variation der Dichte des Schaums



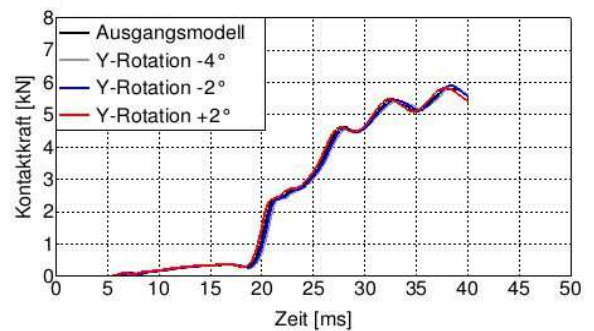
b) Variation der Steifigkeit des Schaums

Abbildung A2.3.1 – Sensitivitätsanalysen hinsichtlich Schaumdichte und Steifigkeit

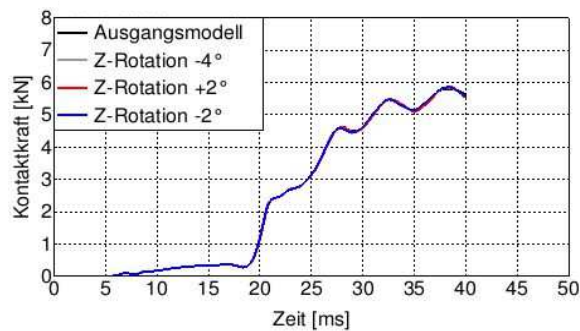
A2.4 Sensitivitätsanalysen bezüglich der Knochenpositionierung



a) Rotation des Knochens um x-Achse



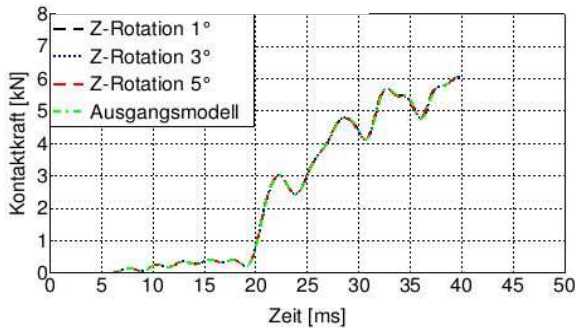
b) Rotation des Knochens um y-Achse



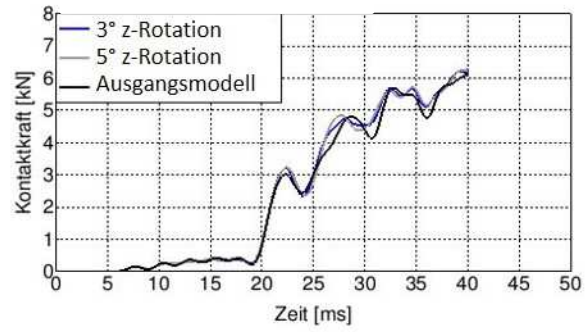
c) Rotation des Knochens um z-Achse

Abbildung A2.4.1 – Sensitivitätsanalysen hinsichtlich Positionierung des Knochens

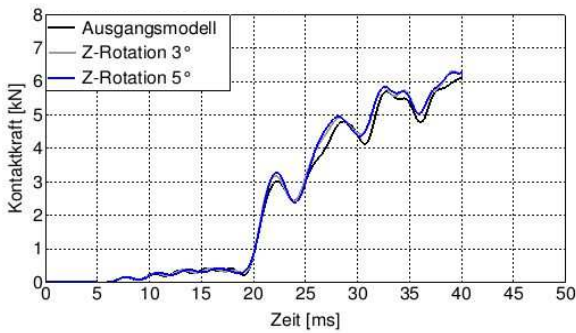
A2.5 Sensitivitätsanalysen bezüglich der Positionierung des Versuchsaufbaus



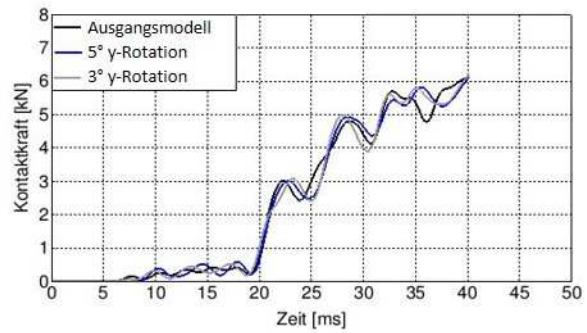
a) Schräge Positionierung des Impaktors



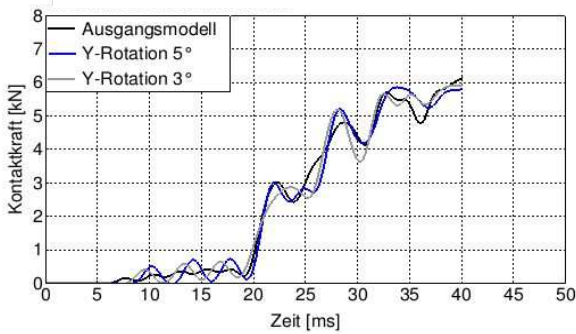
b) Z-Rotation eines Auflagers – keine parallele Lage der Auflager zueinander



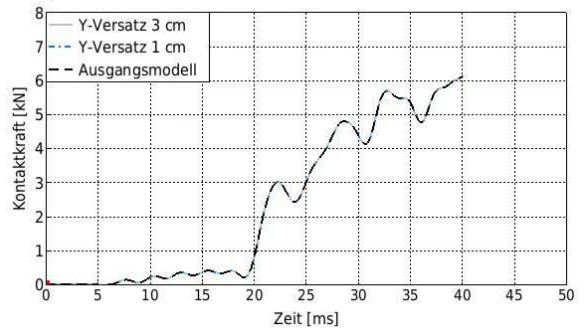
c) Z-Rotation beide Auflager



d) Y-Rotation eines Auflagers – keine parallele Lage der Auflager zueinander



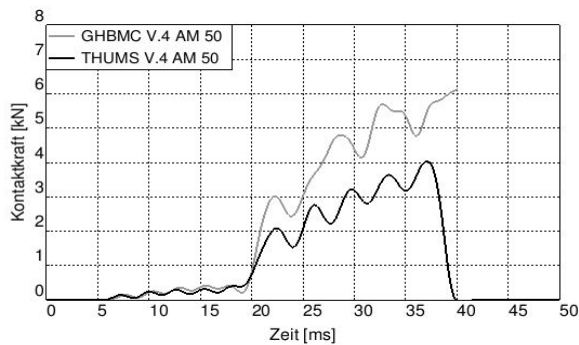
e) Y-Rotation beider Auflager



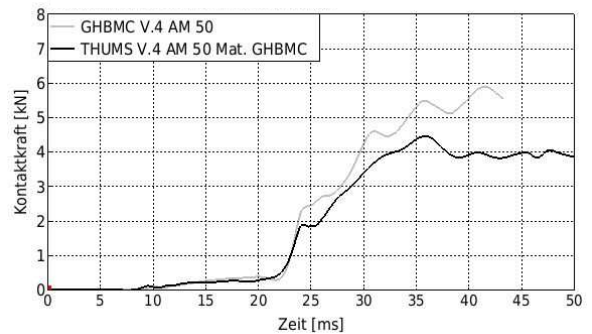
f) Auflagerversatz in y-Richtung eines Auflagers

Abbildung A2.5.1 – Sensitivitätsanalysen bezüglich der Positionierung des Versuchsaufbaus

A2.6 Einfluss unterschiedlicher Knochenmodelle auf das Simulationssetup



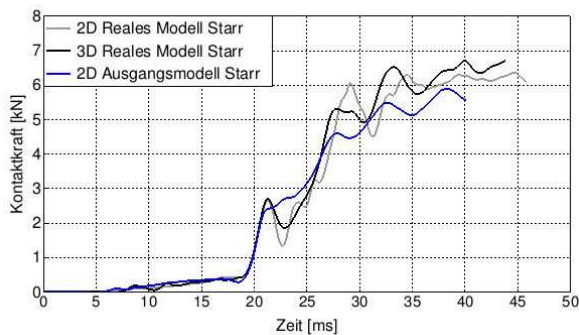
a) Modellvorhersage mit unterschiedlichen Knochenmodellen



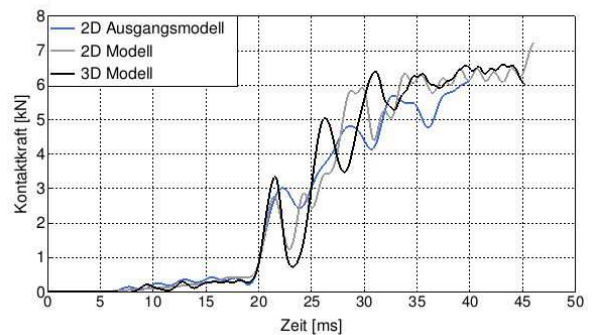
b) Modellvorhersage unterschiedlicher Knochengemetrien

Abbildung A2.6.1 – Einfluss unterschiedlicher Knochenmodelle auf die Modellvorhersage

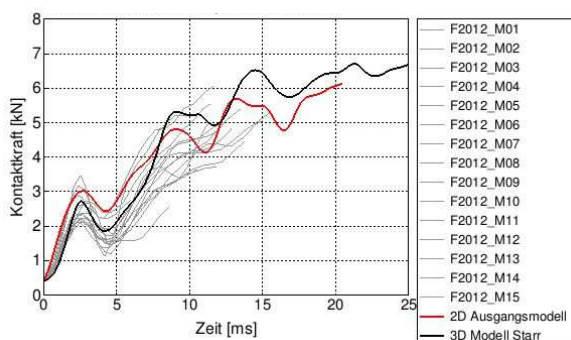
A2.7 Vergleich des entwickelten Simulationsmodells mit der Realität



a) Vergleich der unterschiedlichen starren Modellierungsarten des Versuch Aufbaus



b) Vergleich der unterschiedlichen deformierbaren Modellierungsarten des Versuch Aufbaus



c) Vergleich Ausgangsmodell mit dem realen Modell

Abbildung A2.7.1 – Vergleich des Ausgangsmodells mit realen Modell und Vergleich der unterschiedlichen Modellierungsarten: 2D- vs. 3D-Modellierung

A3 Dendrogramme

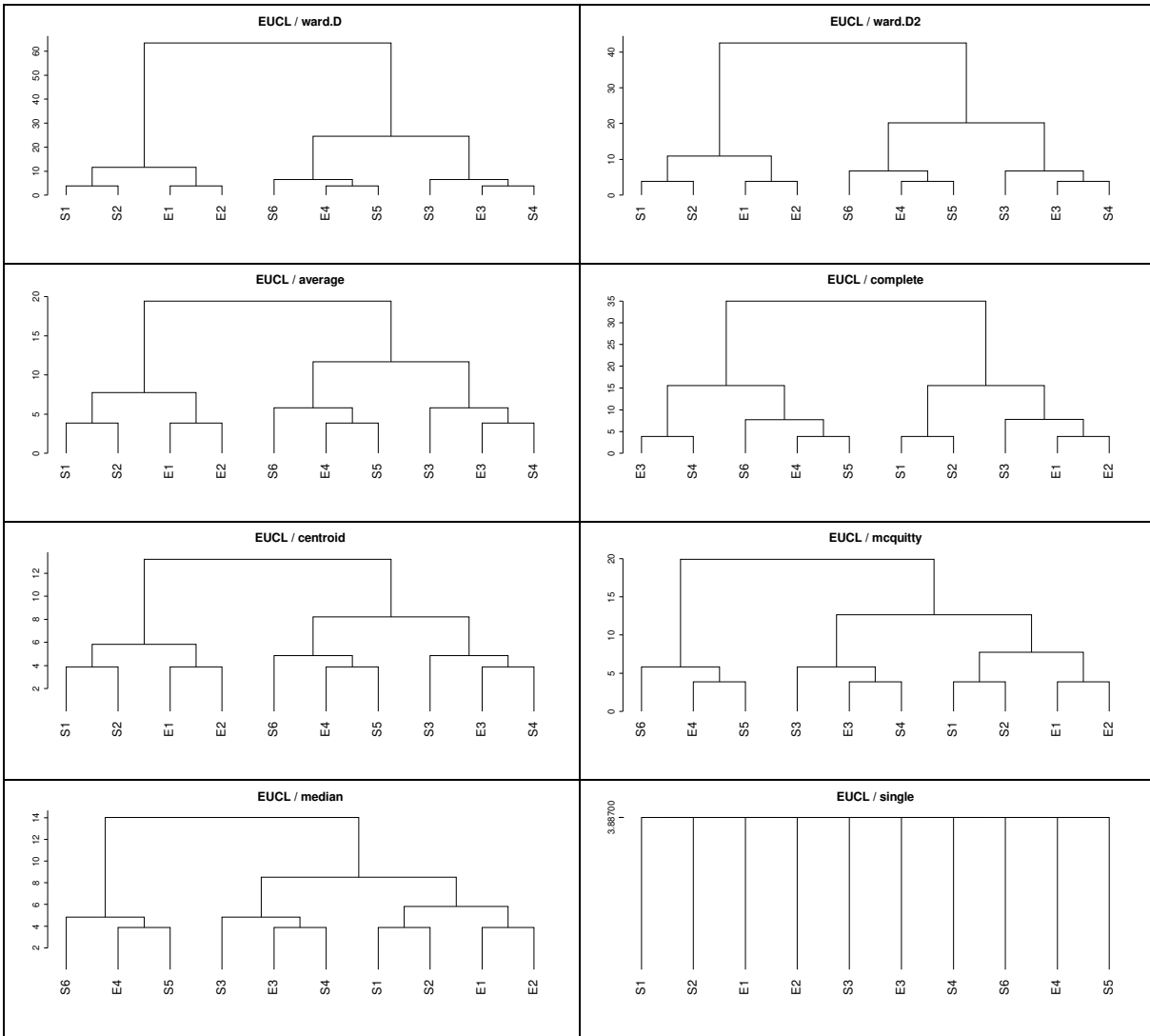
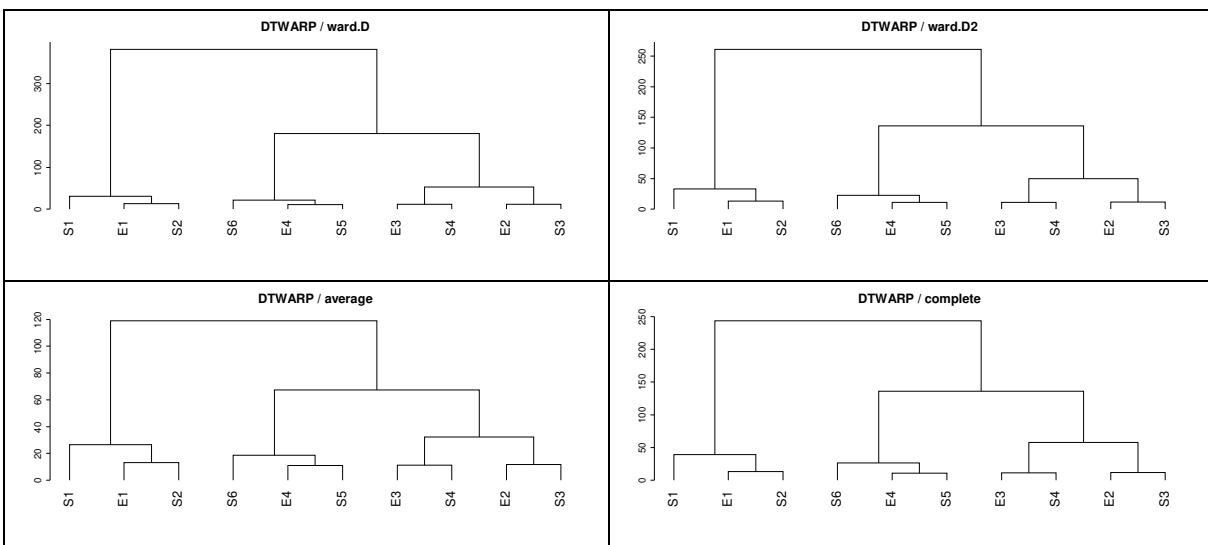


Abbildung A3.1 – Dendrogramme Test Datensatz I mit euklidischer Distanz



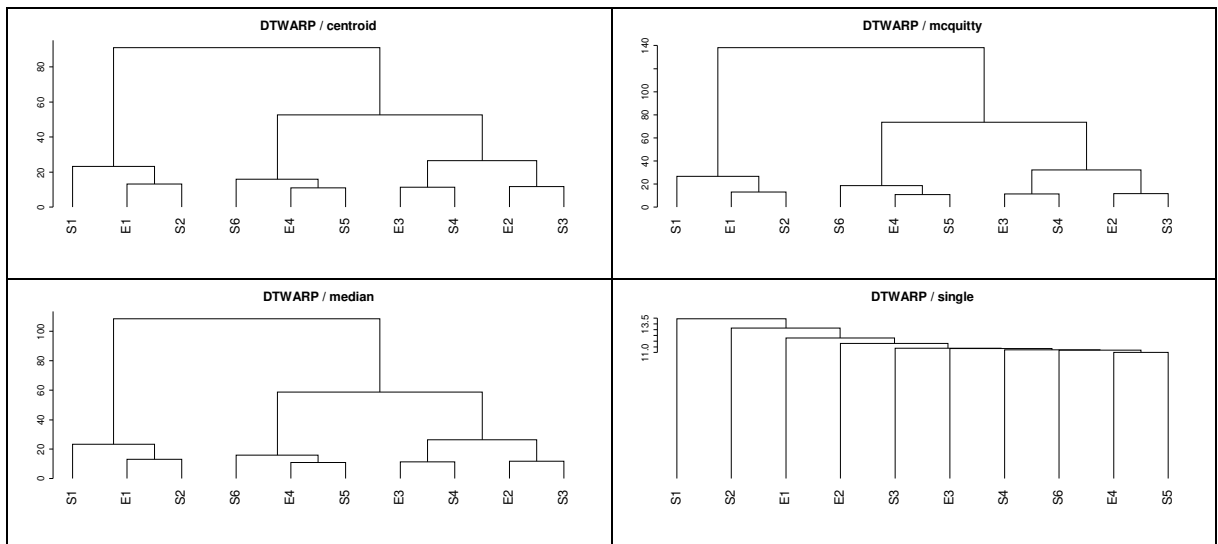


Abbildung A3.2 – Dendrogramme Test Datensatz I mit **dynamischer Zeitverzerrung**

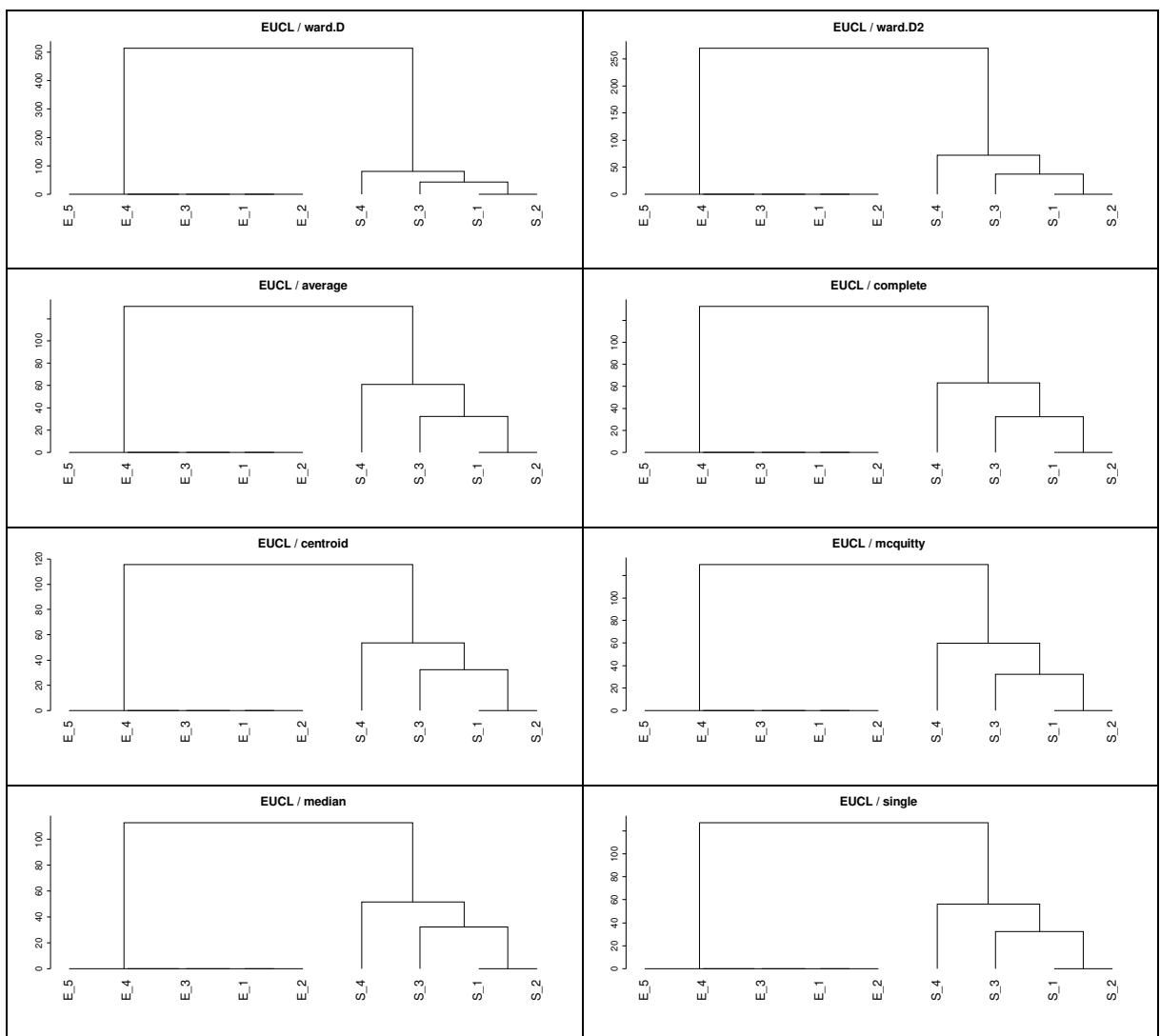


Abbildung A3.3 – Dendrogramme Testdatensatz II mit **euklidischer Distanz**

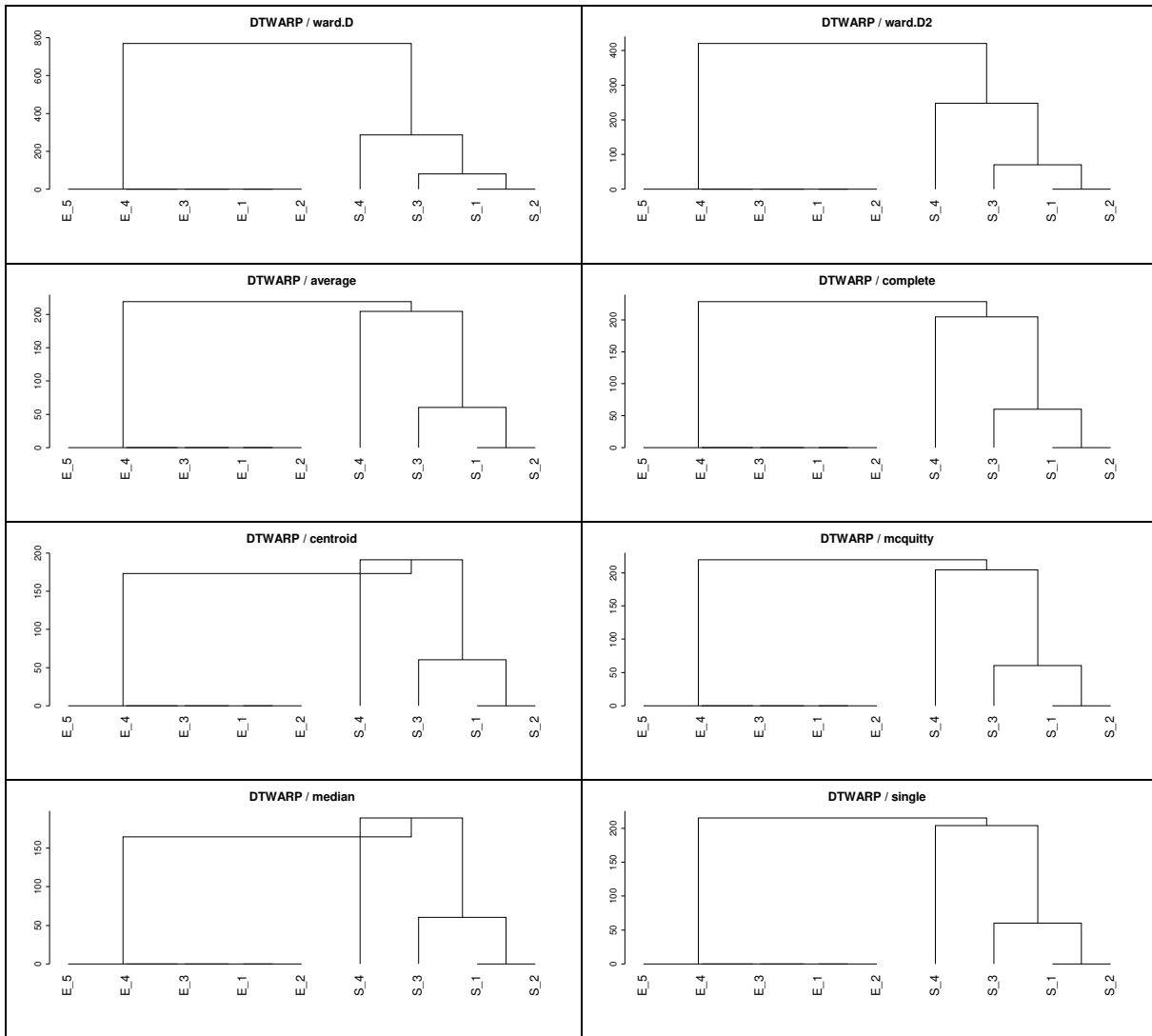
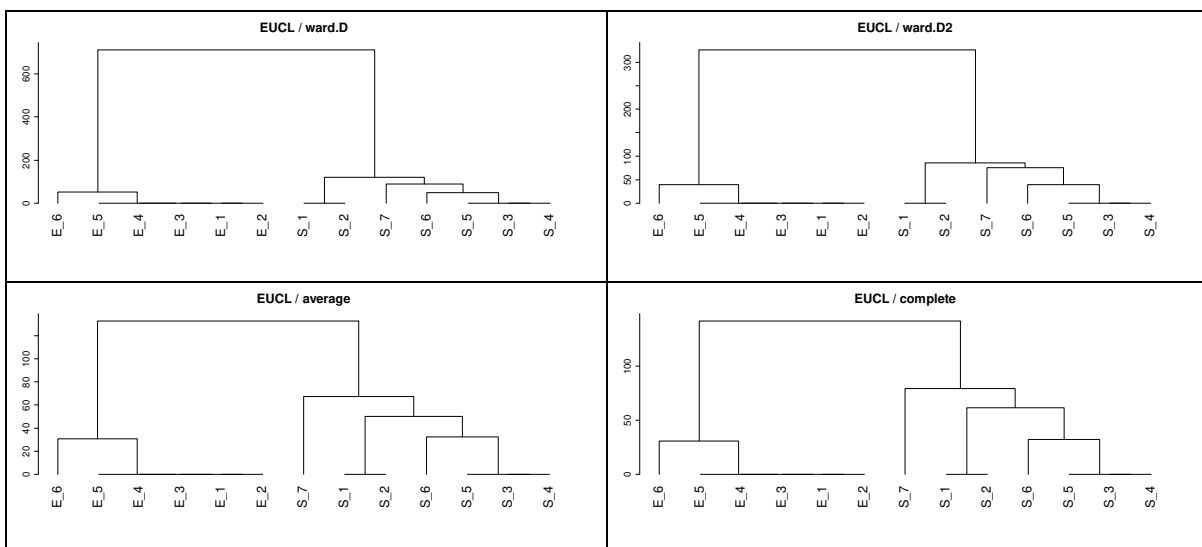


Abbildung A3.4 – Dendrogramme Testdatensatz II mit **dynamischer Zeitverzerrung**



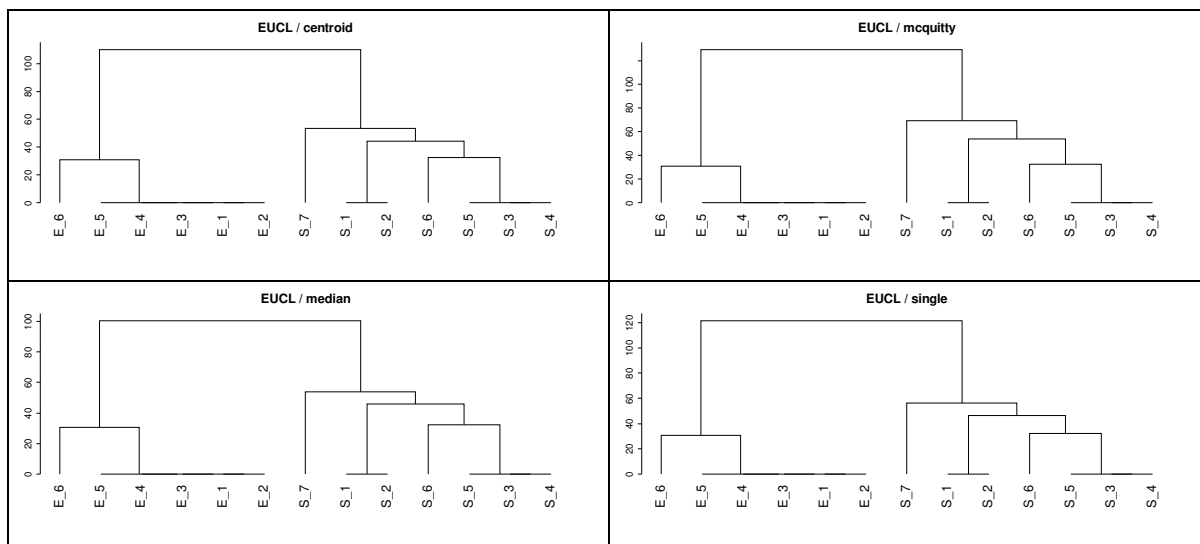


Abbildung A3.5 – Dendrogramme Test Datensatz III mit **euklidischer Distanz**

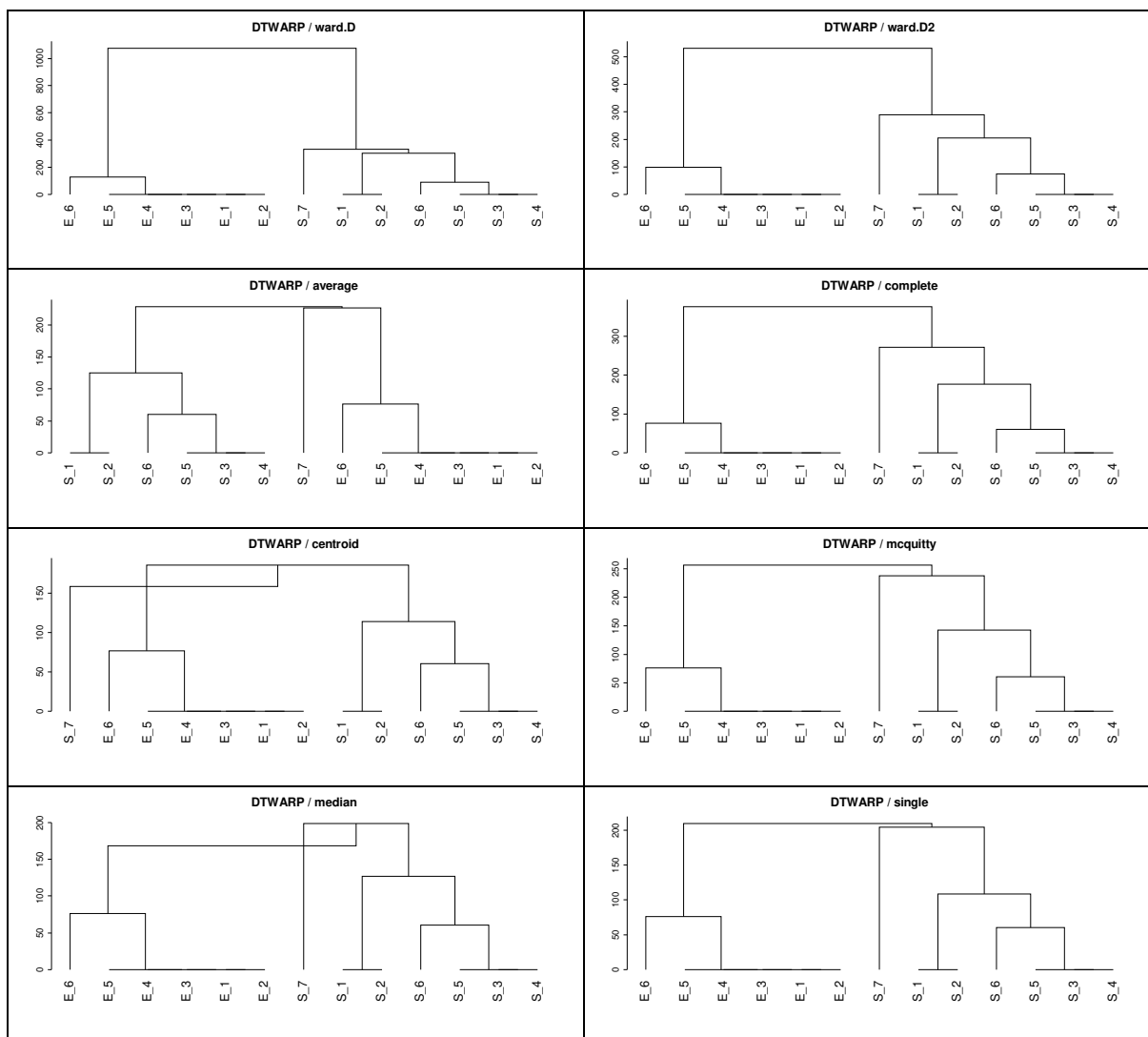


Abbildung A3.6 – Dendrogramme Testdatensatz III mit **dynamischer Zeitverzerrung**

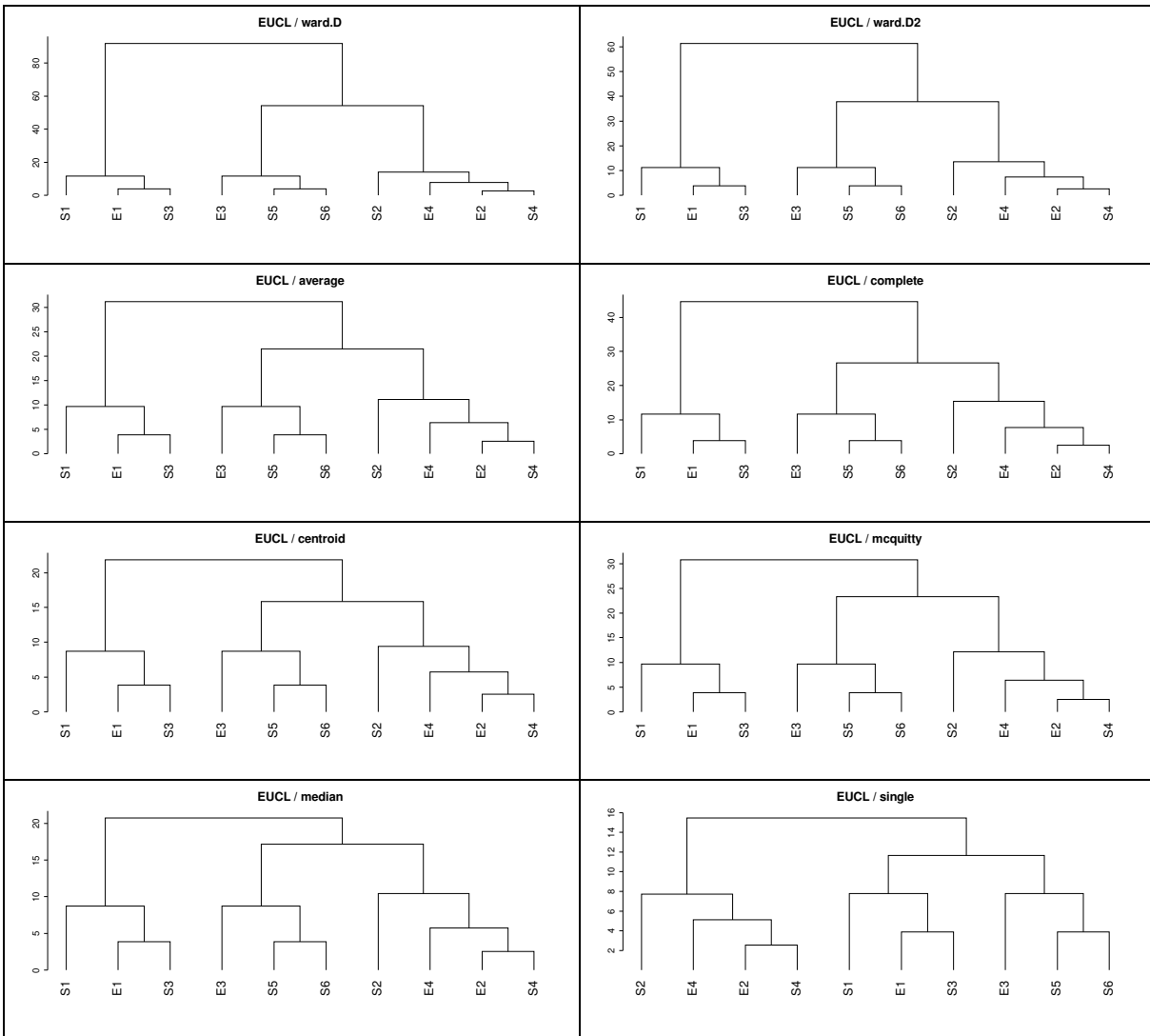
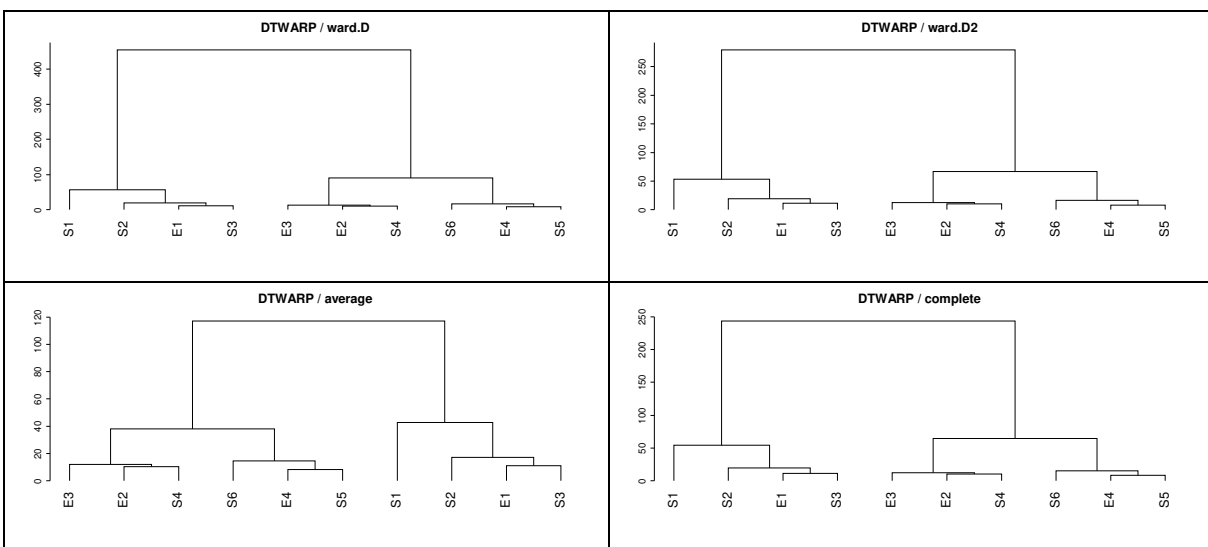


Abbildung A3.7 – Dendrogramme Test Datensatz IV mit **euklidischer Distanz**



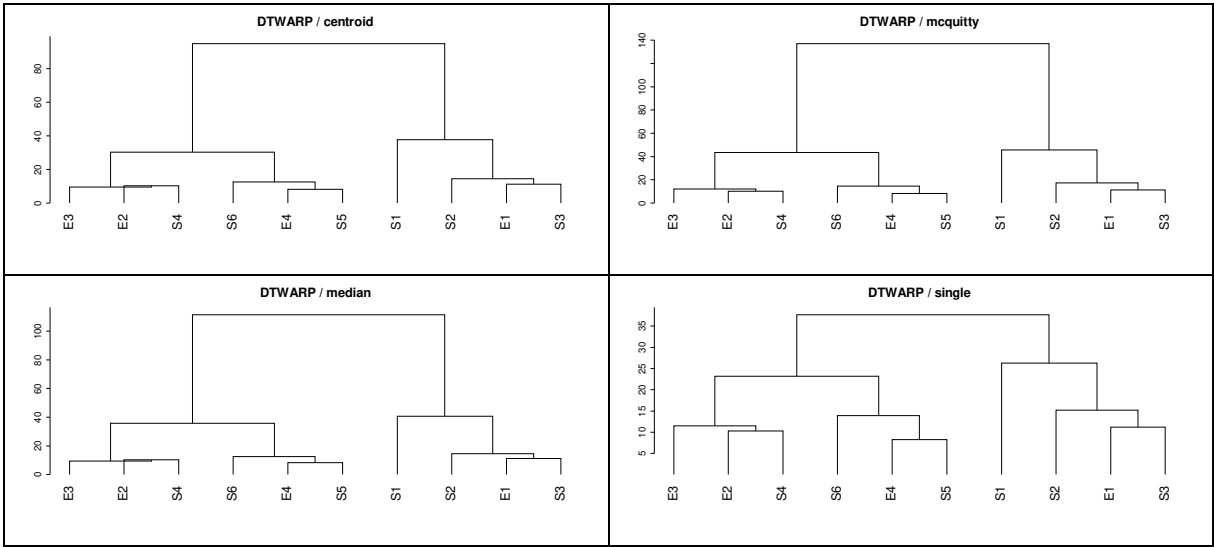


Abbildung A3.8 – Dendrogramme Testdatensatz IV mit dynamischer Zeitverzerrung

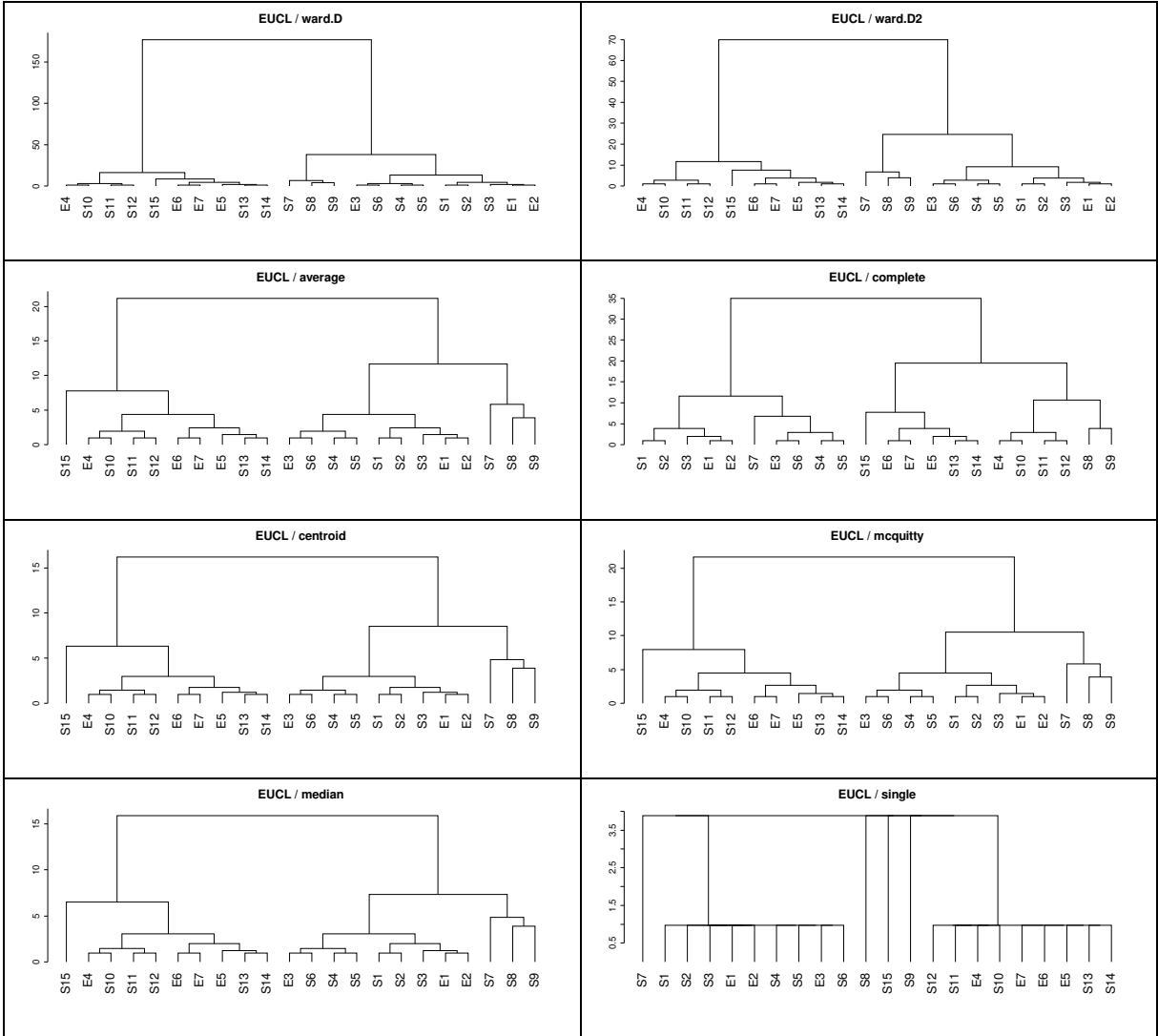


Abbildung A3.9 – Dendrogramme Testdatensatz V mit euklidischer Distanz

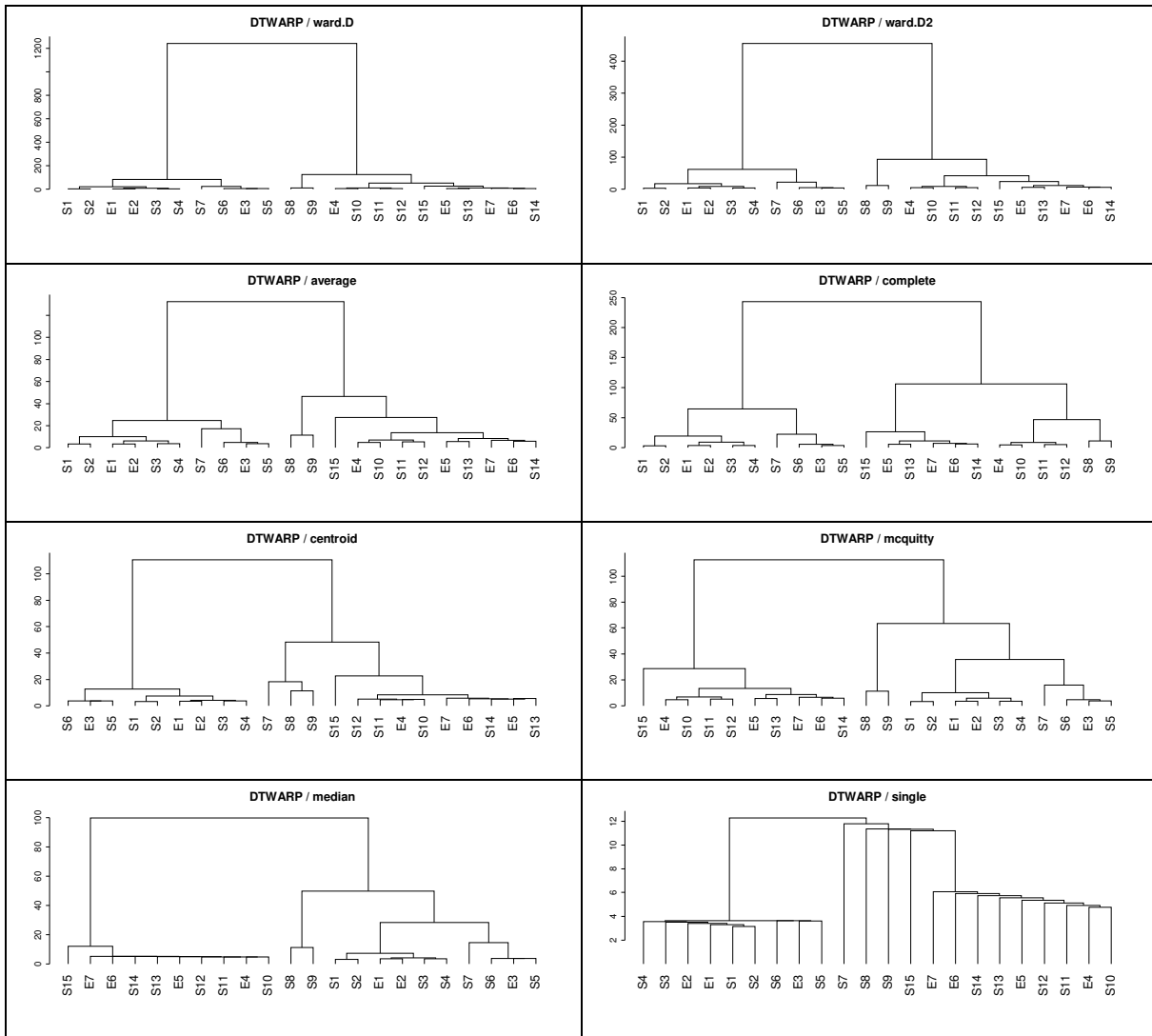
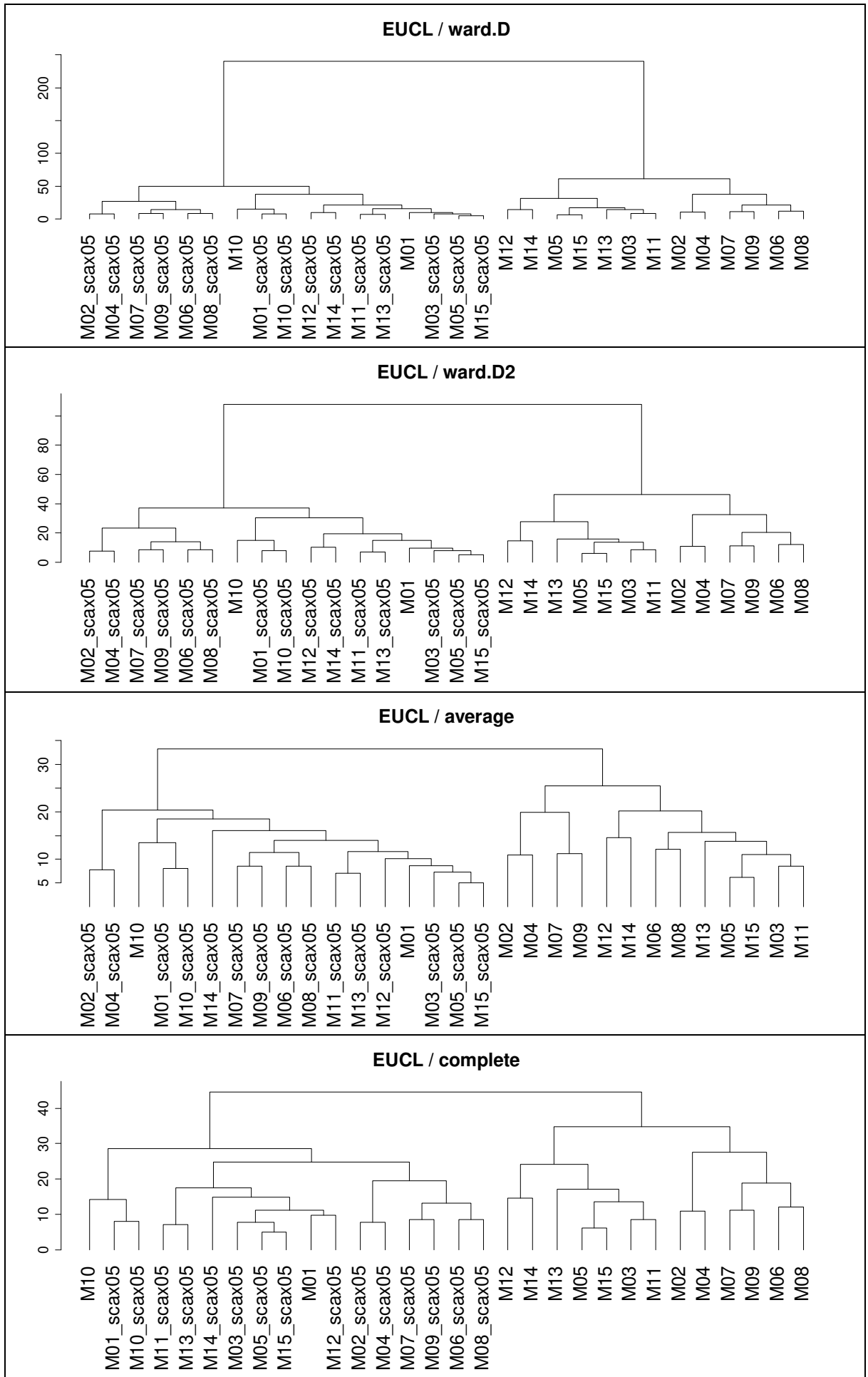


Abbildung A3. 10 – Dendrogramme Testdatensatz V mit **dynamischer Zeitverzerrung**



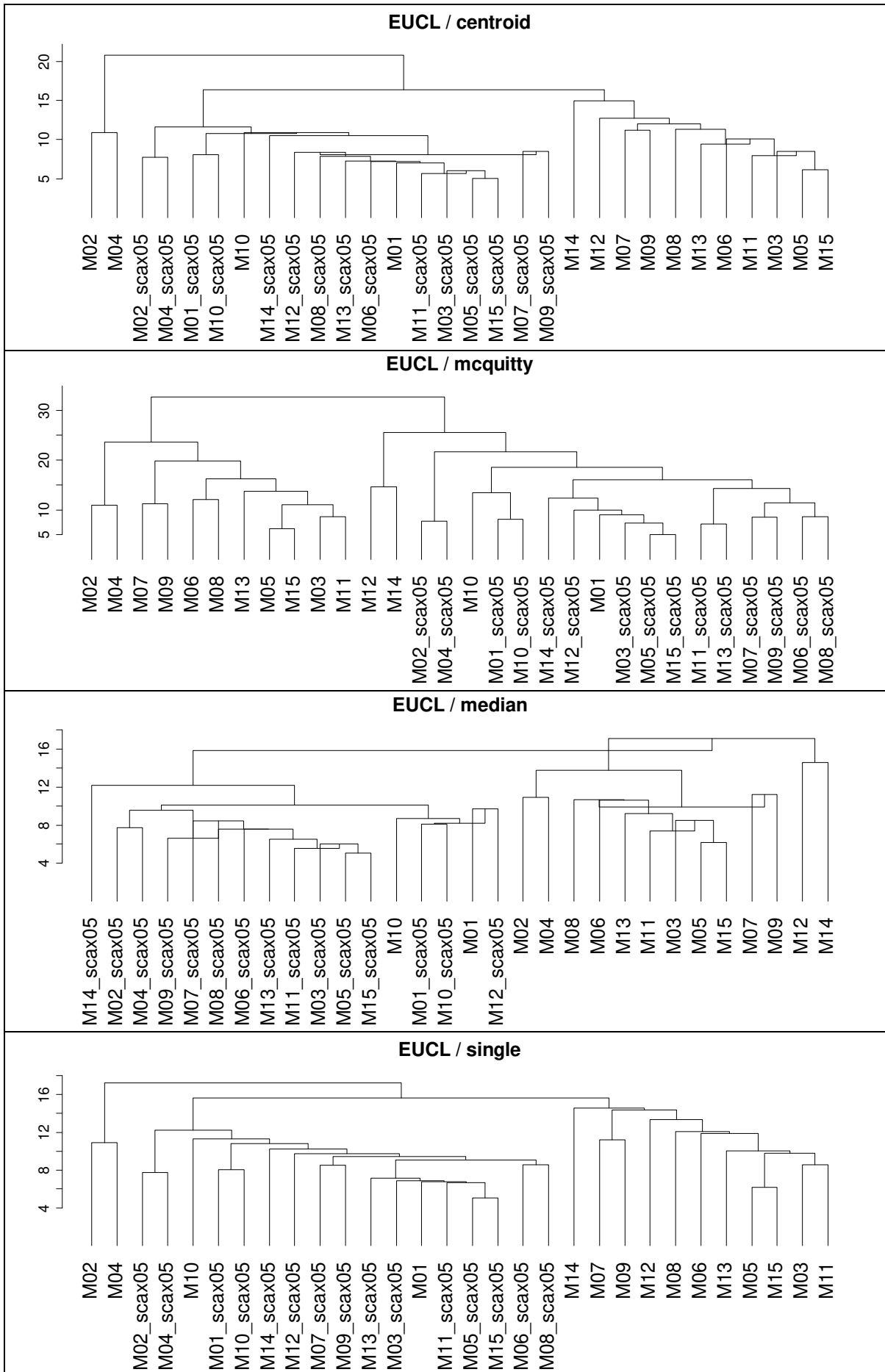
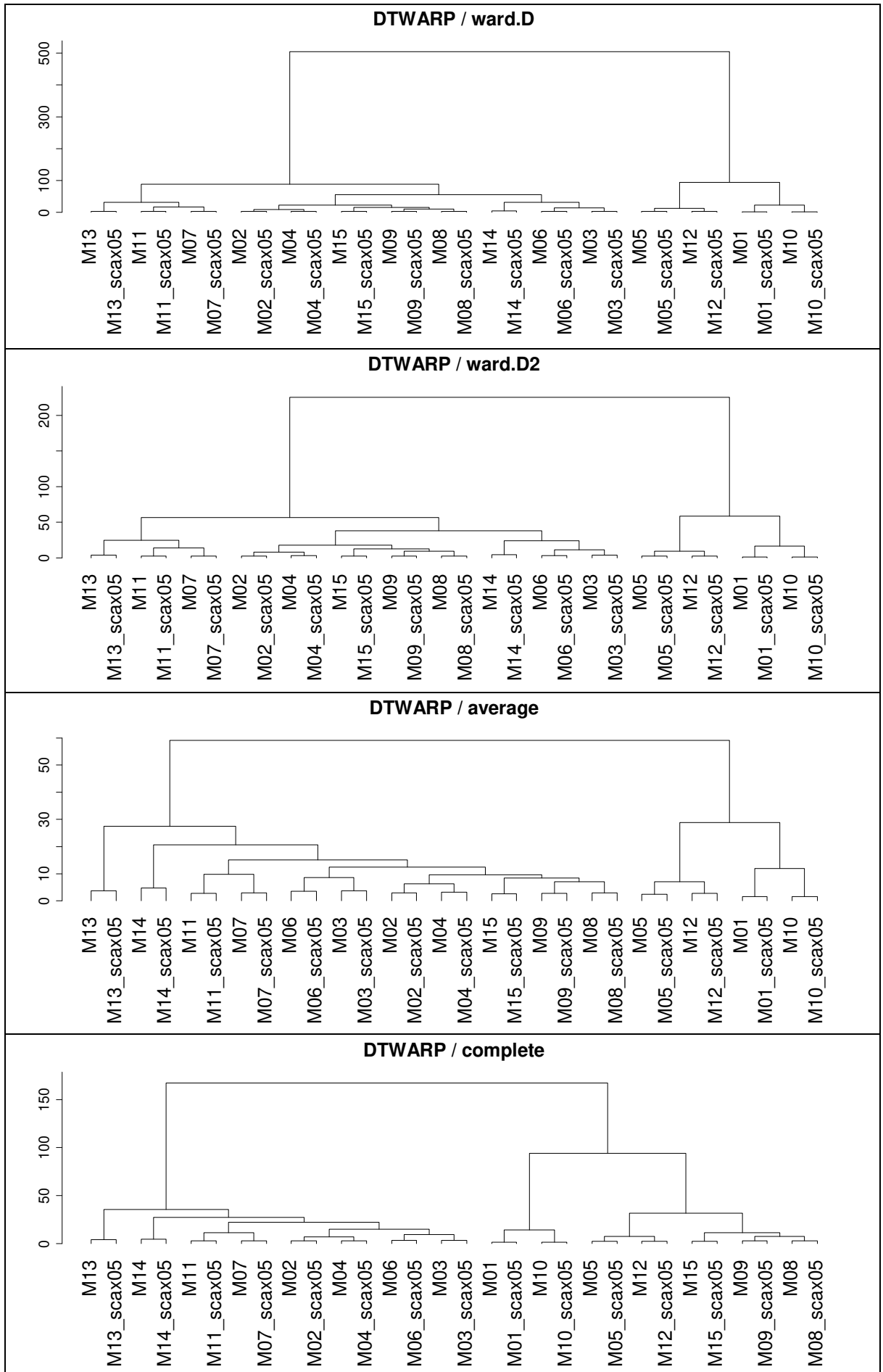


Abbildung A3.11 – Dendrogramme Testdatensatz VI mit euklidischer Distanz



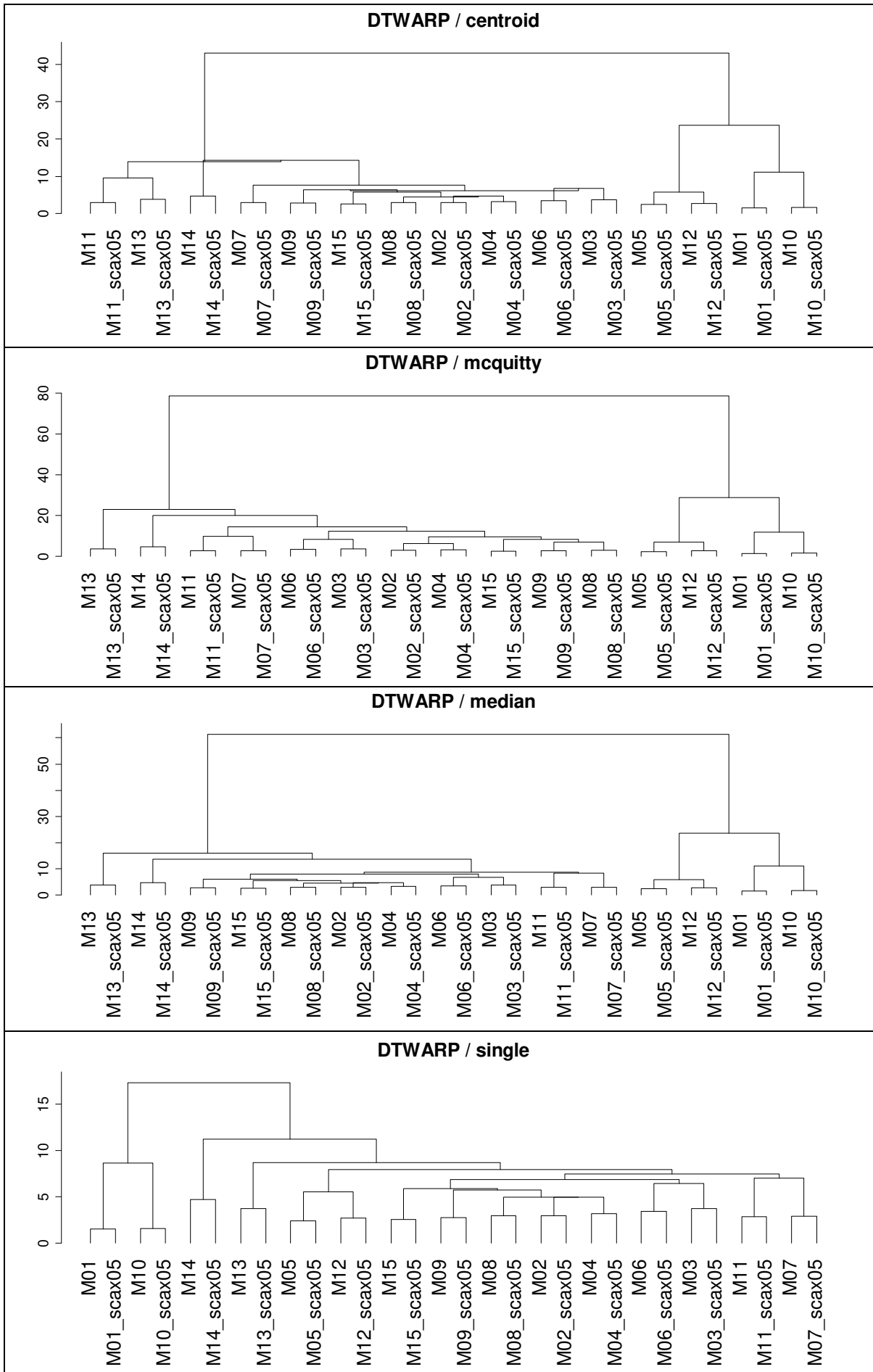
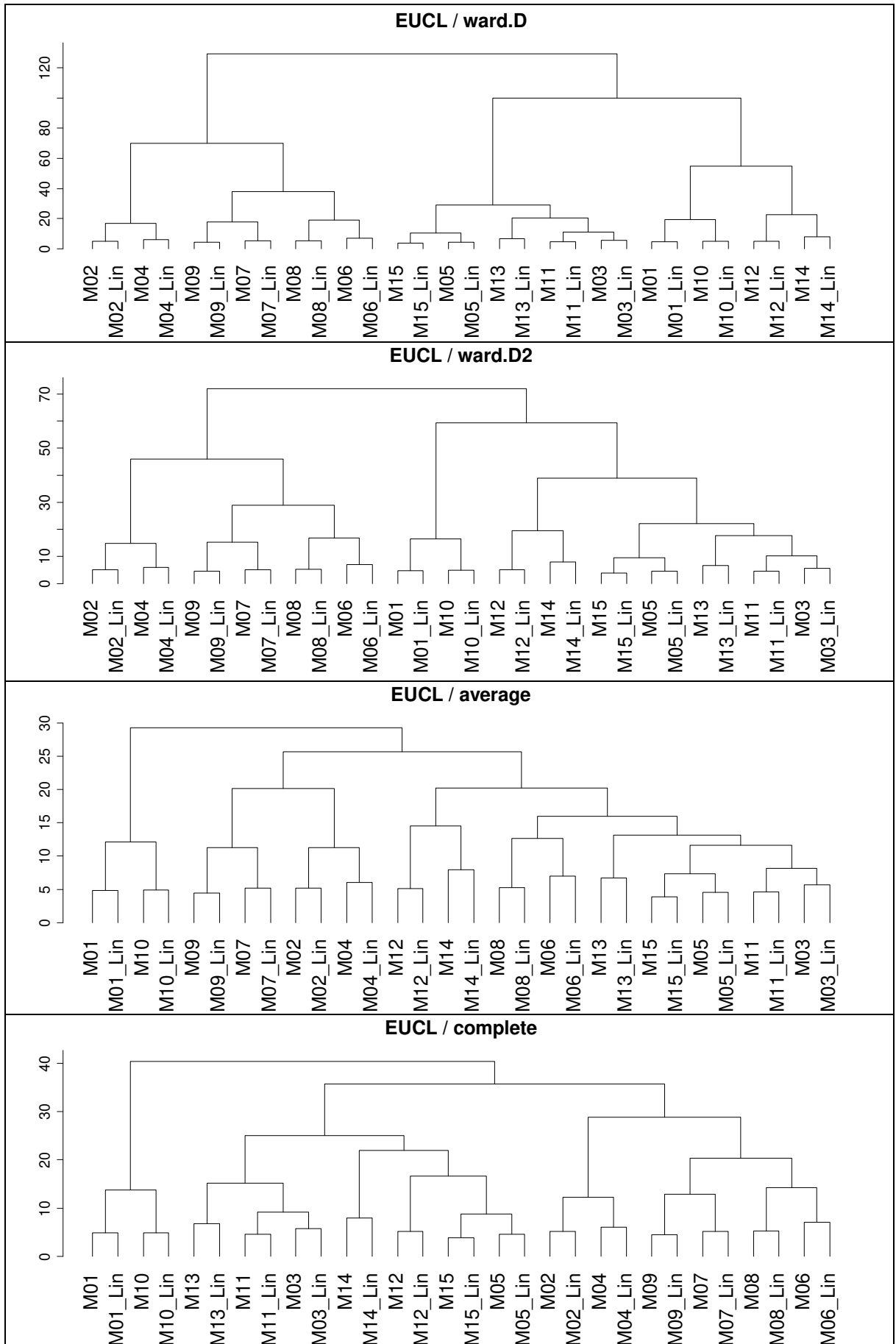


Abbildung A3.12 – Dendrogramme Testdatensatz VI mit dynamischer Zeitverzerrung



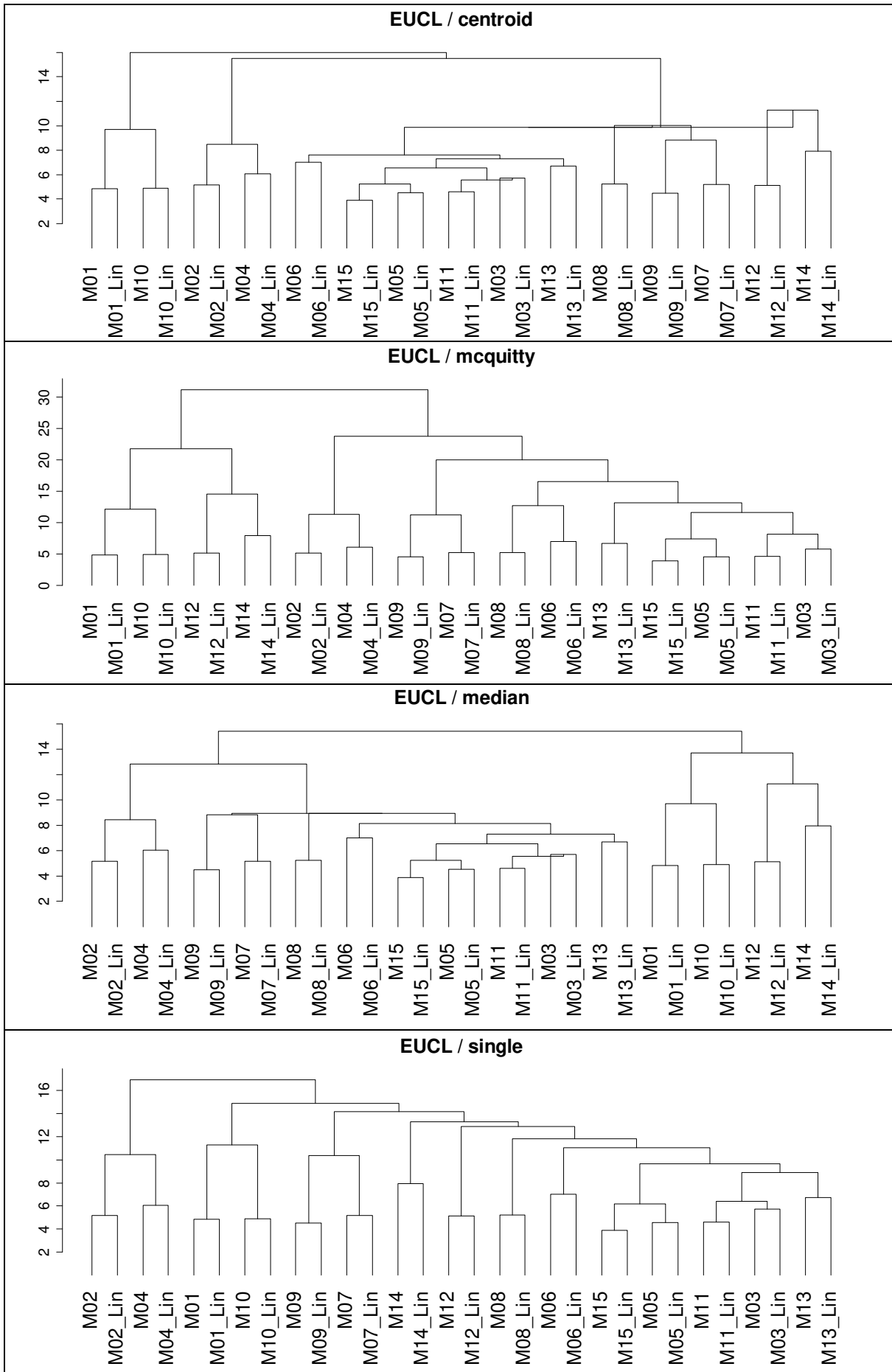
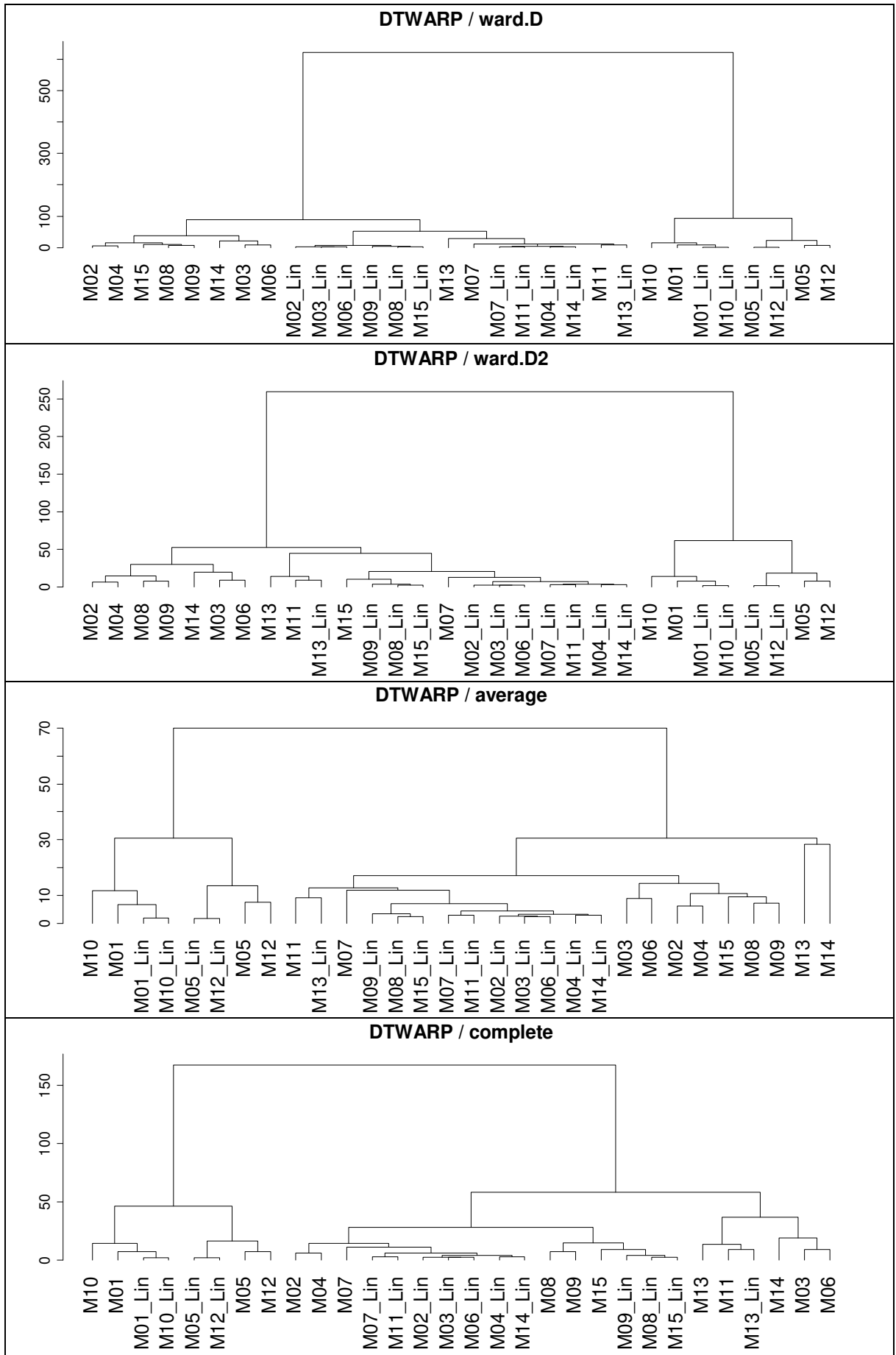


Abbildung A3.13 – Dendrogramme Testdatensatz VII mit **euklidischer Distanz**



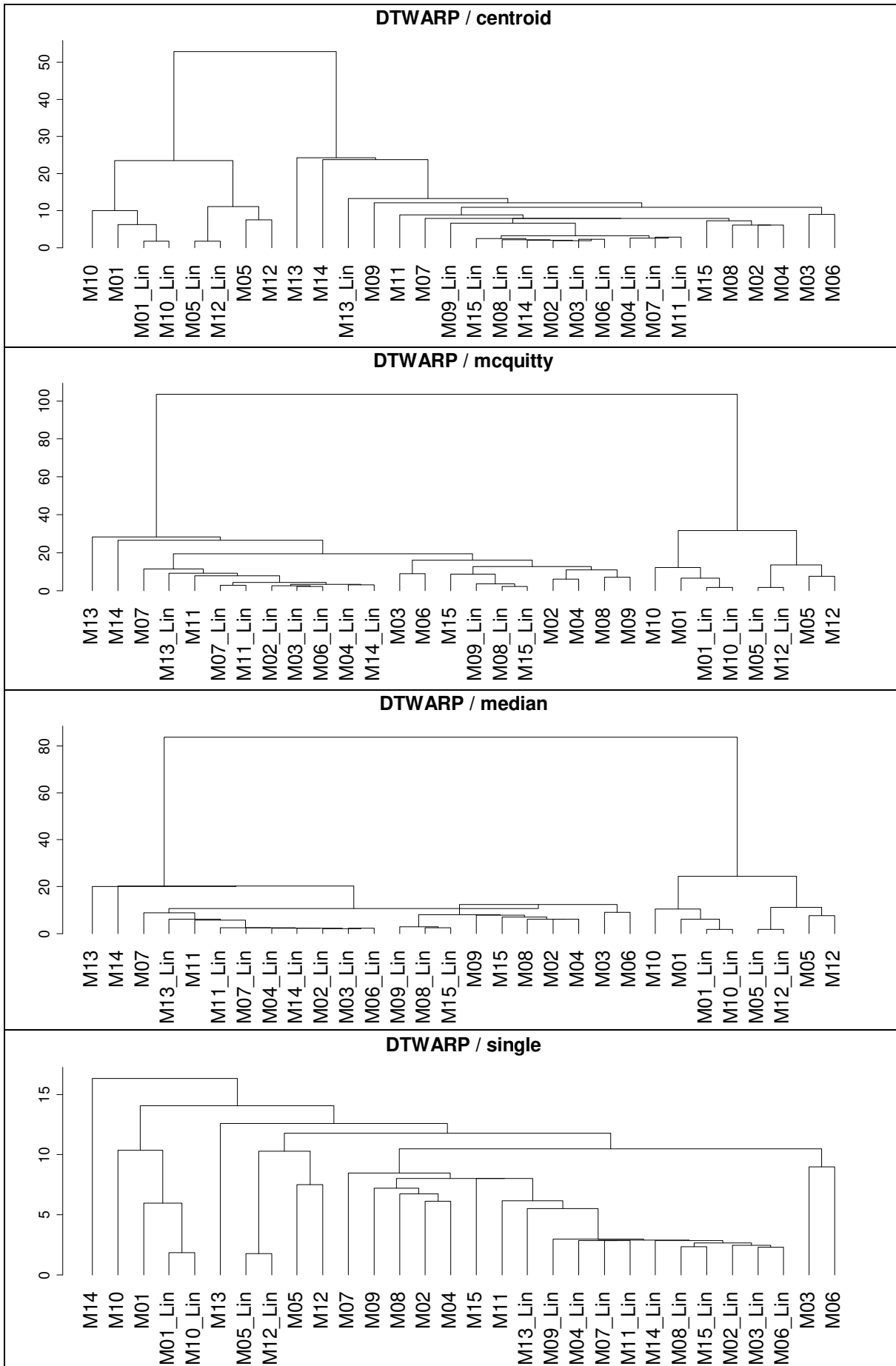
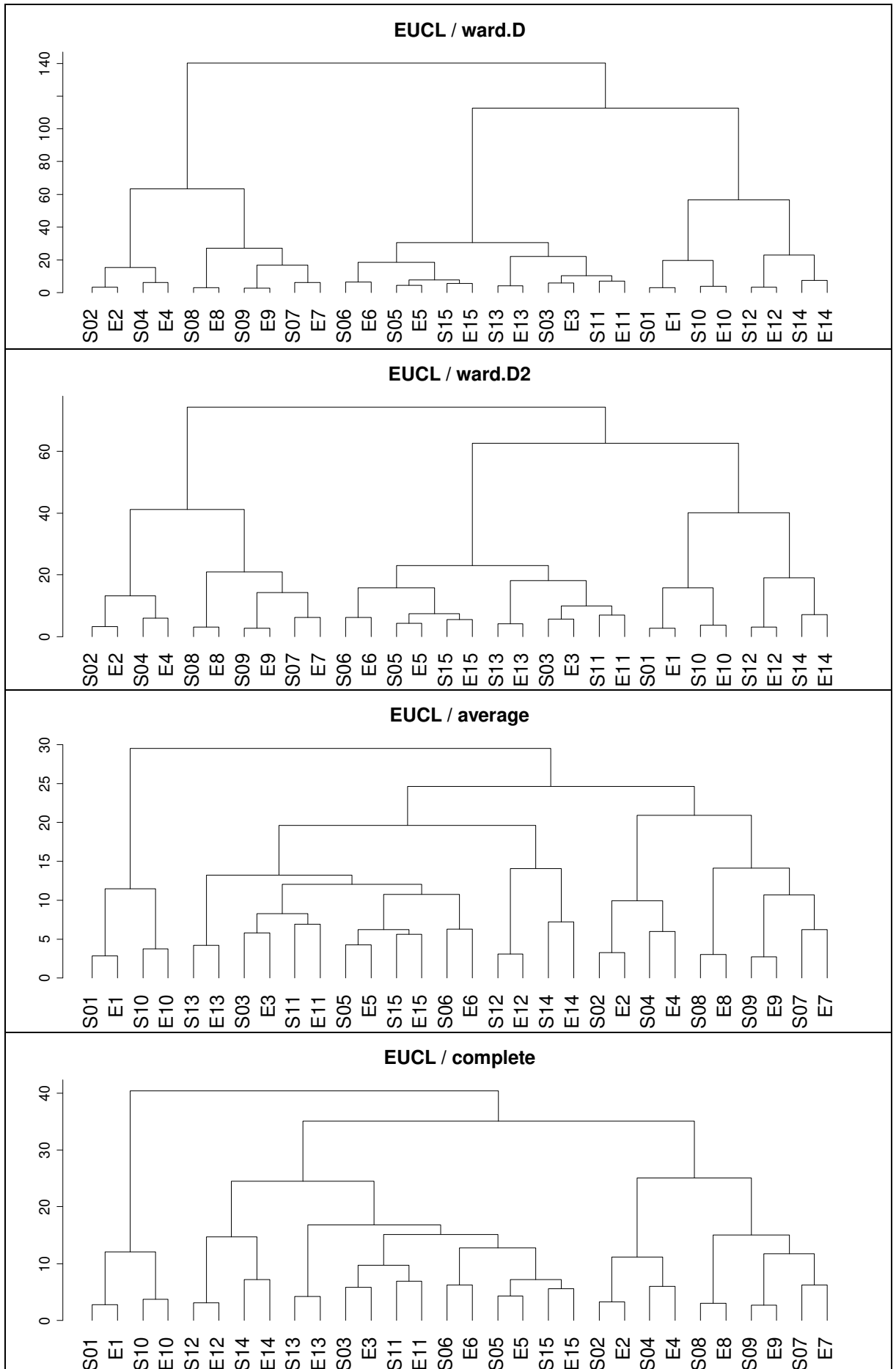


Abbildung A3.14 – Dendrogramme Testdatensatz VII mit **dynamischer Zeitverzerrung**



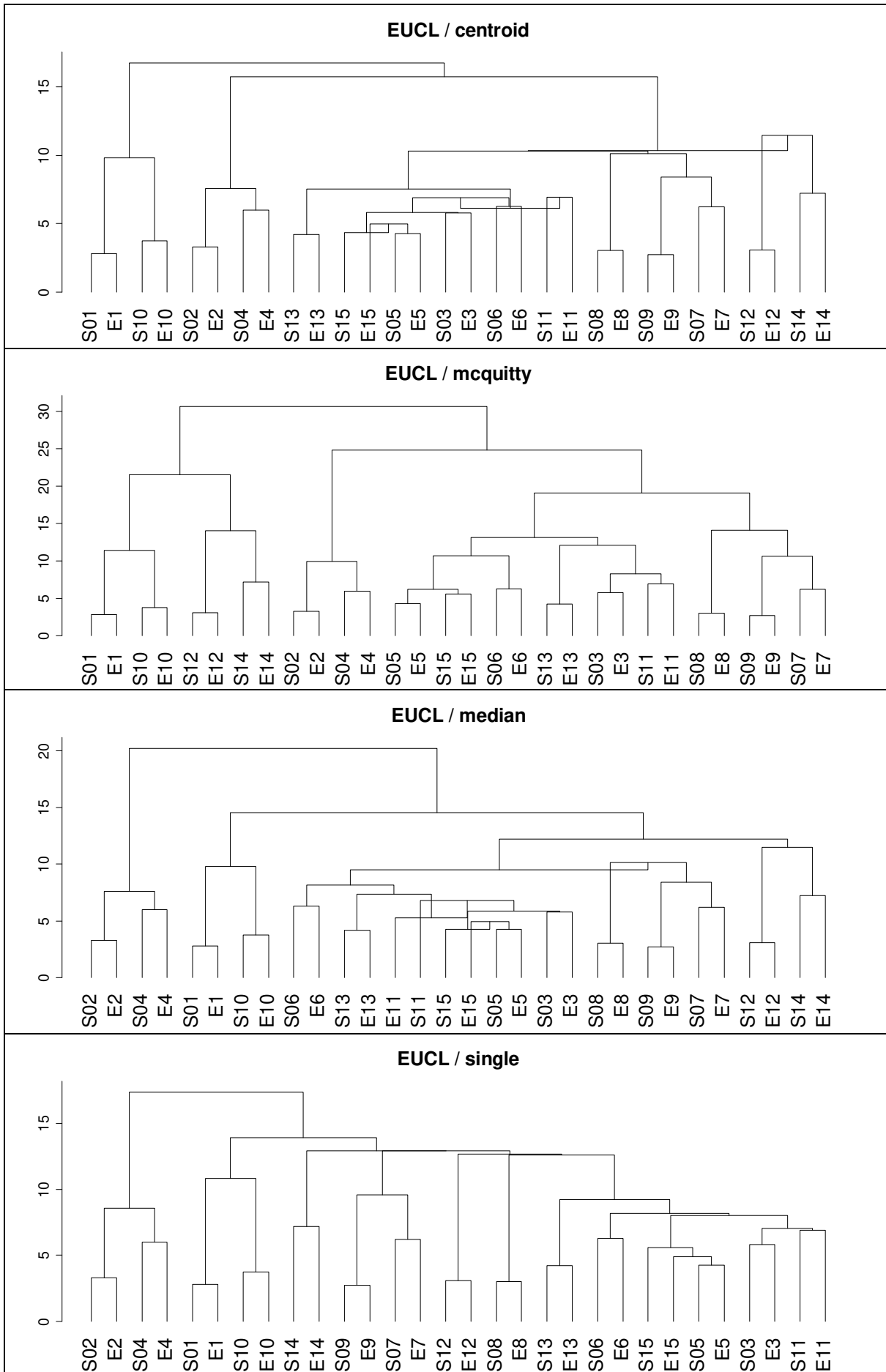
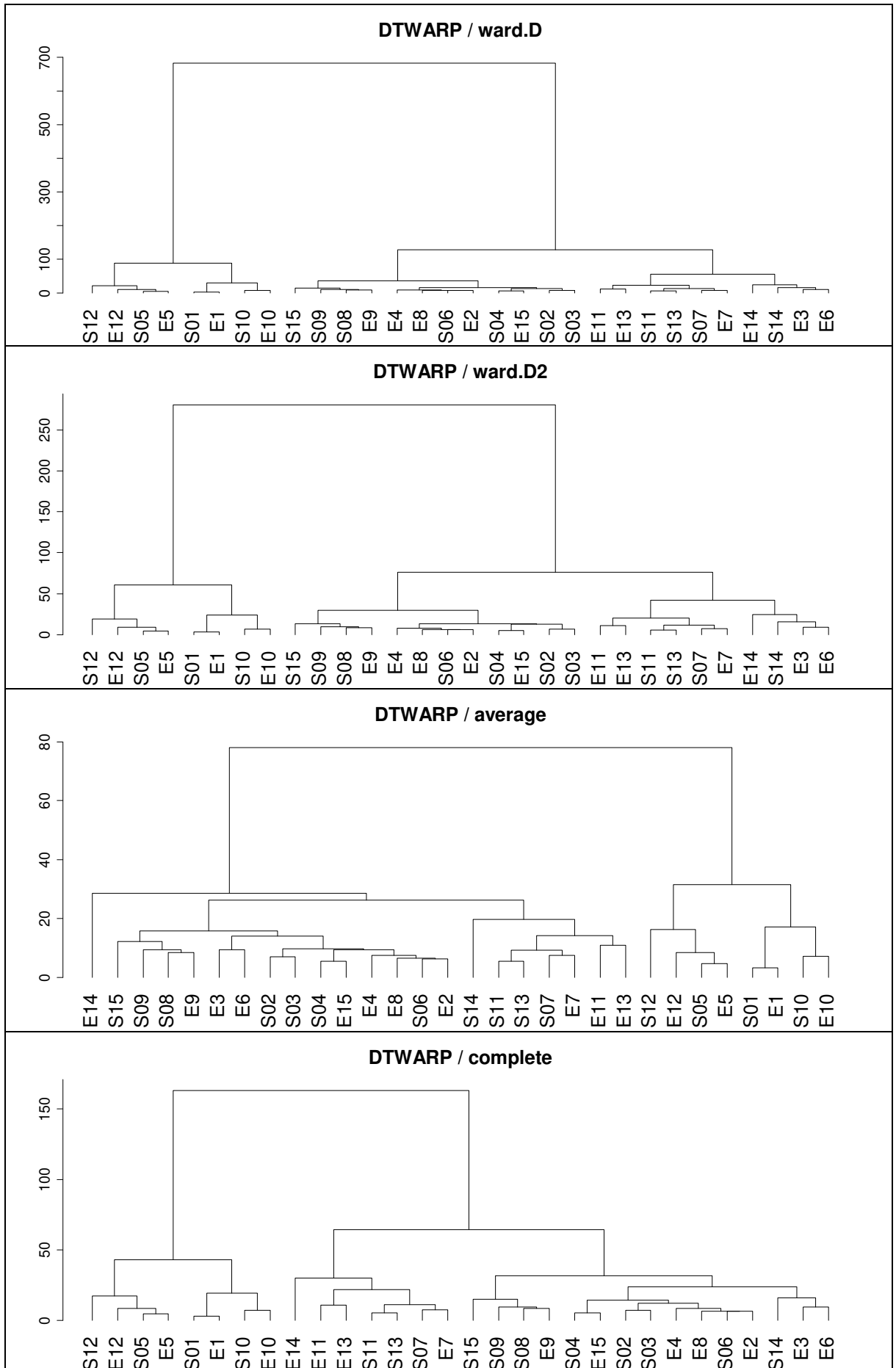


Abbildung A3.15 – Dendrogramme Testdatensatz VIII mit **euklidischer Distanz**



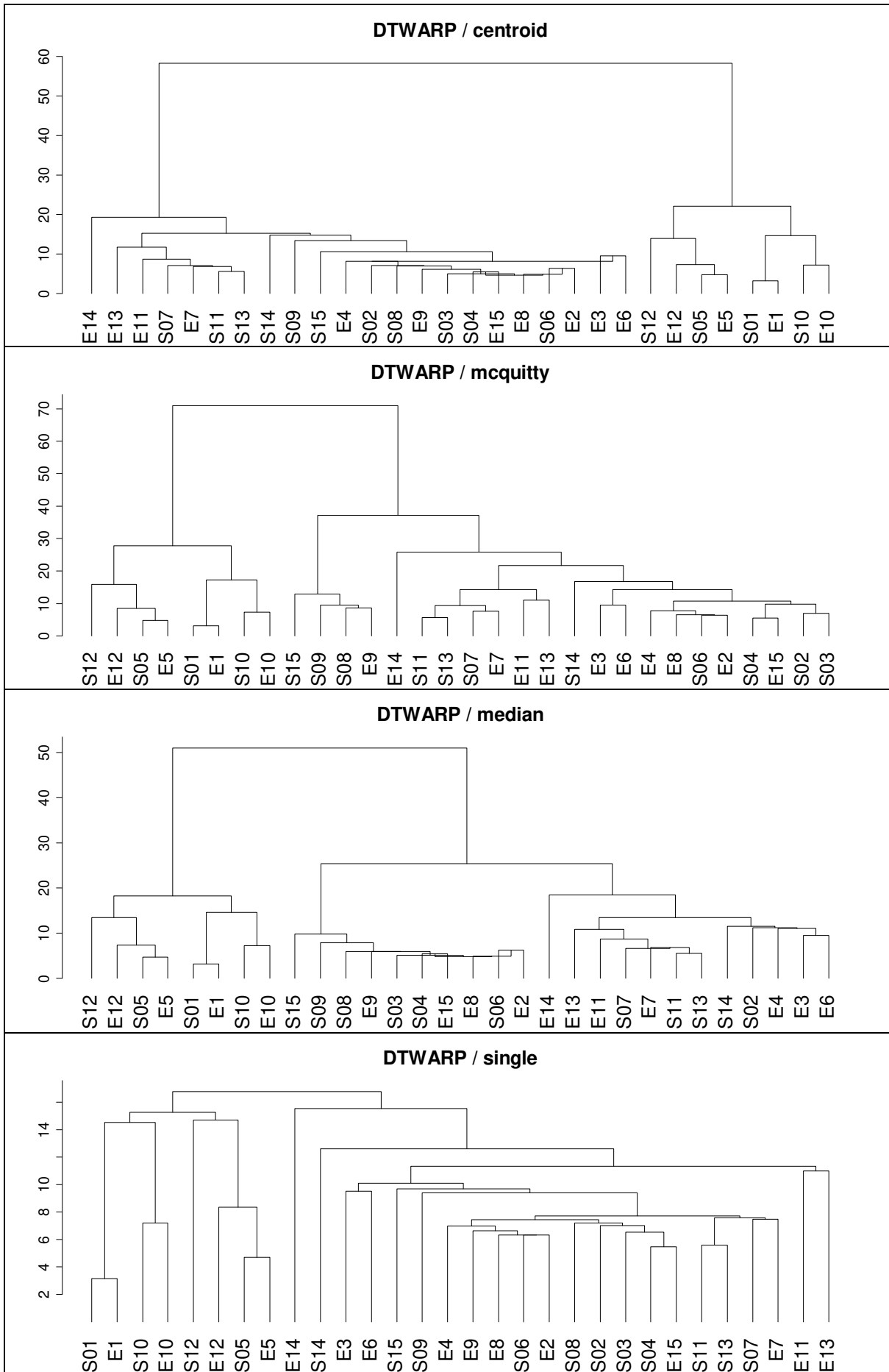


Abbildung A3.16 – Dendrogramme Testdatensatz VIII mit **dynamischer Zeitverzerrung**

A4 Interne Validierungsindizes

Tabelle A4.1 – Interne Validierungsindizes für Testdatensatz I mit **euklidischer** Distanz

Testdatensatz I mit euklidischer Distanz								
Int. Vls	Average-Linkage	Centroid	Complete-Linkage	Mc Quitty	Median	Single-Linkage	Ward.D	Ward.D2
kl	7	7	7	7	7	7	7	7
ch	2	2	2	7	7	2	2	2
hartigan	9	9	9	9	9	9	9	9
cindex	4	4	4	4	4	2	4	4
db	9	9	9	9	9	9	9	9
silhouette	9	9	9	9	9	9	9	9
duda	2	2	1	4	4	2	2	2
pseudot2	2	2	1	4	4	2	2	2
beale	4	4	4	4	4	2	4	4
ball	2	2	2	2	2	2	2	2
ptbiserial	2	2	2	3	3	2	2	2
gap	1	1	1	1	1	1	1	1
mcclain	2	2	2	2	2	2	2	2
gamma	6	6	6	6	6	6	6	6
gplus	6	6	6	6	6	6	6	6
tau	2	2	2	3	3	2	2	2
dunn	6	6	6	6	6	6	6	6
sdindex	7	7	7	7	7	7	7	7
sdbw	9	9	9	9	9	9	9	9

Tabelle A4.2 – Interne Validierungsindizes für Testdatensatz I mit **dynamischer Zeitverzerrung**

Testdatensatz I mit dynamischer Zeitverzerrung								
Int. Vls	Average-Linkage	Centroid	Complete-Linkage	Mc Quitty	Median	Single-Linkage	Ward.D	Ward.D2
kl	3	3	3	3	3	7	3	3
ch	2	2	2	2	2	7	2	2
hartigan	9	9	9	9	9	9	9	9
cindex	4	4	4	4	4	2	4	4
db	9	9	9	9	9	7	9	9
silhouette	9	9	9	9	9	9	9	9
duda	1	1	1	1	1	1	1	1
pseudot2	1	1	1	1	1	1	1	1
beale	4	4	4	4	4	1	4	4
ball	2	2	2	2	2	2	2	2
ptbiserial	2	2	2	2	2	4	2	2
gap	1	1	1	1	1	1	1	1
mcclain	2	2	2	2	2	2	2	2
gamma	9	9	9	9	9	9	9	9
gplus	9	9	9	9	9	9	9	9
tau	2	2	2	2	2	4	2	2
dunn	9	9	9	9	9	9	9	9
sdindex	9	9	9	9	9	4	9	9
sdbw	9	9	9	9	9	9	9	9

Tabelle A4.3 – Interne Validierungsindizes für Testdatensatz II mit **euklidischer Distanz**

Testdatensatz II mit euklidischer Distanz								
Int. Vls	Average-Linkage	Centroid	Complete-Linkage	Mc Quitty	Median	Single-Linkage	Ward.D	Ward.D2
kl	2	2	2	2	2	2	2	2
ch	2	2	2	2	2	2	2	2
hartigan	8	8	8	8	8	8	8	8
cindex	4	4	4	4	4	4	4	4
db	8	8	8	8	8	8	8	8
silhouette	4	4	4	4	4	4	4	4
duda	2	2	2	2	2	2	2	2
pseudot2	2	2	2	2	2	2	2	2
beale	2	2	2	2	2	2	2	2
ball	2	2	2	2	2	2	2	2
ptbiserial	2	2	2	2	2	2	2	2
gap	1	1	1	1	1	1	1	1
mcclain	4	4	4	4	4	4	4	4
gamma	2	2	2	2	2	2	2	2
gplus	2	2	2	2	2	2	2	2
tau	6	6	6	6	6	6	6	6
dunn	4	4	4	4	4	4	4	4
sdindex	8	8	8	8	8	8	8	8
sdbw	8	8	8	8	8	8	8	8

Tabelle A4.4 – Interne Validierungsindizes für Testdatensatz II mit **dynamischer Zeitverzerrung**

Testdatensatz II mit dynamischer Zeitverzerrung								
Int. Vls	Average-Linkage	Centroid	Complete-Linkage	Mc Quitty	Median	Single-Linkage	Ward.D	Ward.D2
kl	2	2	2	2	2	2	2	2
ch	2	2	2	2	2	2	2	2
hartigan	8	8	8	8	8	8	8	8
cindex	4	4	4	4	4	4	4	4
db	8	8	8	8	8	8	8	8
silhouette	4	4	4	4	4	4	4	4
duda	2	2	2	2	2	2	2	2
pseudot2	2	2	2	2	2	2	2	2
beale	2	2	2	2	2	2	2	2
ball	2	2	2	2	2	2	2	2
ptbiserial	3	3	3	3	3	3	3	3
gap	1	1	1	1	1	1	1	1
mcclain	4	4	4	4	4	4	4	4
gamma	2	2	2	2	2	2	2	2
gplus	2	2	2	2	2	2	2	2
tau	6	6	6	6	6	6	6	6
dunn	4	4	4	4	4	4	4	4
sdindex	8	8	8	8	8	8	8	8
sdbw	8	8	8	8	8	8	8	8

Tabelle A4.5 – Interne Validierungsindizes für Testdatensatz III mit **euklidischer Distanz**

Testdatensatz III mit euklidischer Distanz								
Int. Vls	Average-Linkage	Centroid	Complete-Linkage	Mc Quitty	Median	Single-Linkage	Ward.D	Ward.D2
kl	2	2	2	2	2	2	2	2
ch	2	2	2	2	2	2	2	2
hartigan	12	12	12	12	12	12	12	12
cindex	6	6	6	6	6	6	6	6
db	12	12	12	12	12	12	12	12
silhouette	6	6	6	6	6	6	6	6
duda	2	2	2	2	2	2	3	3
pseudot2	2	2	2	2	2	2	3	3
beale	2	2	2	2	2	2	3	3
ball	2	2	2	2	2	2	2	2
ptbiserial	2	2	2	2	2	2	2	2
gap	1	1	1	1	1	1	1	1
mcclain	6	6	6	6	6	6	6	6
gamma	2	2	2	2	2	2	2	2
gplus	2	2	2	2	2	2	2	2
tau	2	2	2	2	2	2	2	2
dunn	6	6	6	6	6	6	6	6
sdindex	4	4	4	4	4	4	5	5
sdbw	12	12	12	12	12	12	12	12

Tabelle A4.6 – Interne Validierungsindizes für Testdatensatz III mit **dynamischer Zeitverzerrung**

Testdatensatz III mit dynamischer Zeitverzerrung								
Int. Vls	Average-Linkage	Centroid	Complete-Linkage	Mc Quitty	Median	Single-Linkage	Ward.D	Ward.D2
kl	7	3	2	2	2	2	2	2
ch	3	3	2	2	2	2	2	2
hartigan	12	12	12	12	12	12	12	12
cindex	6	6	6	6	6	6	6	6
db	12	12	12	12	12	12	12	12
silhouette	6	6	6	6	6	6	6	6
duda	2	1	2	2	2	2	2	2
pseudot2	2	1	2	2	2	2	2	2
beale	2	1	2	2	2	2	2	2
ball	2	2	2	2	2	2	2	2
ptbiserial	3	3	3	3	3	3	3	3
gap	1	1	1	1	1	1	1	1
mcclain	6	6	6	6	6	6	6	6
gamma	3	3	3	3	3	3	3	3
gplus	3	3	3	3	3	3	3	3
tau	3	3	3	3	3	3	3	3
dunn	6	6	6	6	6	6	6	6
sdindex	5	5	5	5	5	5	5	5
sdbw	12	12	12	12	12	12	12	12

Tabelle A4.7 – Interne Validierungsindizes für Testdatensatz IV mit **euklidischer Distanz**

Testdatensatz IV mit euklidischer Distanz								
Int. Vls	Average-Linkage	Centroid	Complete-Linkage	Mc Quitty	Median	Single-Linkage	Ward.D	Ward.D2
kl	5	5	6	5	5	4	5	5
ch	5	5	8	5	5	4	5	5
hartigan	9	9	9	9	9	9	9	9
cindex	3	3	3	3	3	2	3	3
db	9	9	9	9	9	9	9	9
silhouette	9	9	9	9	9	9	9	9
duda	1	1	1	1	1	1	1	1
pseudot2	1	1	1	1	1	1	1	1
beale	3	3	3	3	3	3	3	3
ball	2	2	2	2	2	2	2	2
ptbiserial	3	3	3	3	3	3	3	3
gap	1	1	1	1	1	1	1	1
mcclain	2	2	2	2	2	2	2	2
gamma	7	7	7	7	7	7	7	7
gplus	7	7	7	7	7	7	7	7
tau	2	2	2	2	2	3	2	2
dunn	9	9	9	9	9	9	9	9
sdindex	6	6	6	6	6	6	6	6
sdbw	9	9	9	9	9	9	9	9

Tabelle A4.8 – Interne Validierungsindizes für Testdatensatz IV mit **dynamischer Zeitverzerrung**

Testdatensatz IV mit dynamischer Zeitverzerrung								
Int. Vls	Average-Linkage	Centroid	Complete-Linkage	Mc Quitty	Median	Single-Linkage	Ward.D	Ward.D2
kl	2	7	3	2	7	2	3	3
ch	2	2	3	2	2	2	3	3
hartigan	9	9	9	9	9	9	9	9
cindex	2	2	3	2	2	2	3	3
db	9	9	9	9	9	9	9	9
silhouette	9	9	9	9	9	9	9	9
duda	1	1	1	1	1	1	1	1
pseudot2	1	1	1	1	1	1	1	1
beale	2	2	3	2	2	2	3	3
ball	2	2	2	2	2	2	2	2
ptbiserial	2	2	2	2	2	2	2	2
gap	1	1	1	1	1	1	1	1
mcclain	2	2	2	2	2	2	2	2
gamma	4	4	4	4	4	4	4	4
gplus	4	4	4	4	4	4	4	4
tau	2	2	2	2	2	2	2	2
dunn	9	9	9	9	9	9	9	9
sdindex	8	8	8	8	8	8	8	8
sdbw	9	9	9	9	9	9	9	9

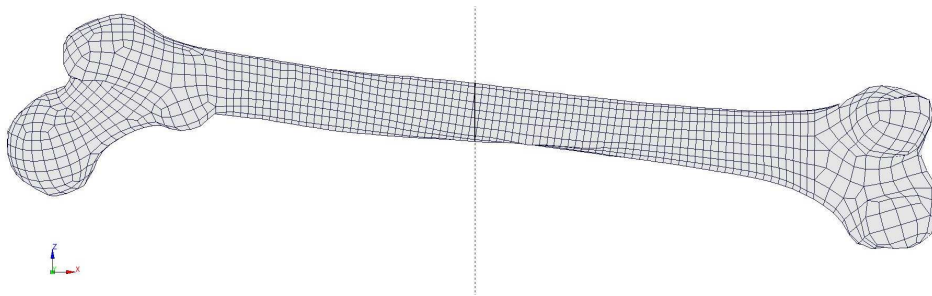
Tabelle A4.9 – Interne Validierungsindizes für Testdatensatz V mit **euklidischer Distanz**

Testdatensatz V mit euklidischer Distanz								
Int. Vls	Average-Linkage	Centroid	Complete-Linkage	Mc Quitty	Median	Single-Linkage	Ward.D	Ward.D2
kl	14	14	14	14	14	20	14	14
ch	20	20	20	20	20	20	20	20
hartigan	21	21	21	21	21	21	21	21
cindex	21	21	21	21	21	21	21	21
db	21	21	21	21	21	21	21	21
silhouette	21	21	21	21	21	21	21	21
duda	1	1	1	1	1	1	1	1
pseudot2	1	1	1	1	1	1	1	1
beale	3	3	1	3	3	1	5	5
ball	2	2	2	2	2	2	2	2
ptbiserial	2	2	2	2	2	2	2	2
gap	1	1	1	1	1	1	1	1
mcclain	2	2	2	2	2	2	2	2
gamma	19	19	19	19	19	19	19	19
gplus	19	19	19	19	19	19	19	19
tau	2	2	2	2	2	2	2	2
dunn	20	20	20	20	20	20	20	20
sdindex	4	4	6	4	4	4	6	6
sdbw	21	21	21	21	21	21	21	21

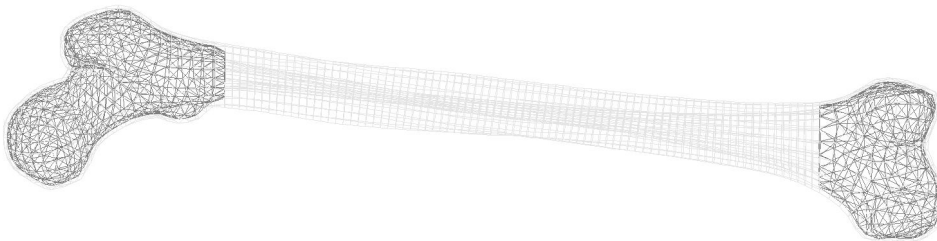
Tabelle A4.10 – Interne Validierungsindizes für Testdatensatz V mit **dynamischer Zeitverzerrung**

Testdatensatz V mit dynamischer Zeitverzerrung								
Int. Vls	Average-Linkage	Centroid	Complete-Linkage	Mc Quitty	Median	Single-Linkage	Ward.D	Ward.D2
kl	19	19	16	19	19	21	19	19
ch	19	19	19	19	19	21	19	19
hartigan	21	21	21	21	21	21	21	21
cindex	2	2	2	2	2	2	2	2
db	21	21	21	21	21	19	21	21
silhouette	21	21	21	21	21	21	21	21
duda	1	1	1	1	1	1	1	1
pseudot2	1	1	1	1	1	1	1	1
beale	1	1	1	4	4	1	1	1
ball	2	2	2	2	2	2	2	2
ptbiserial	2	2	2	2	2	3	2	2
gap	1	1	1	1	1	1	1	1
mcclain	2	3	2	3	3	4	2	2
gamma	21	21	21	21	21	21	21	21
gplus	21	21	21	21	21	21	21	21
tau	2	2	2	3	3	3	2	2
dunn	21	21	21	21	21	21	21	21
sdindex	4	4	16	3	3	17	3	3
sdbw	21	21	21	21	21	21	21	21

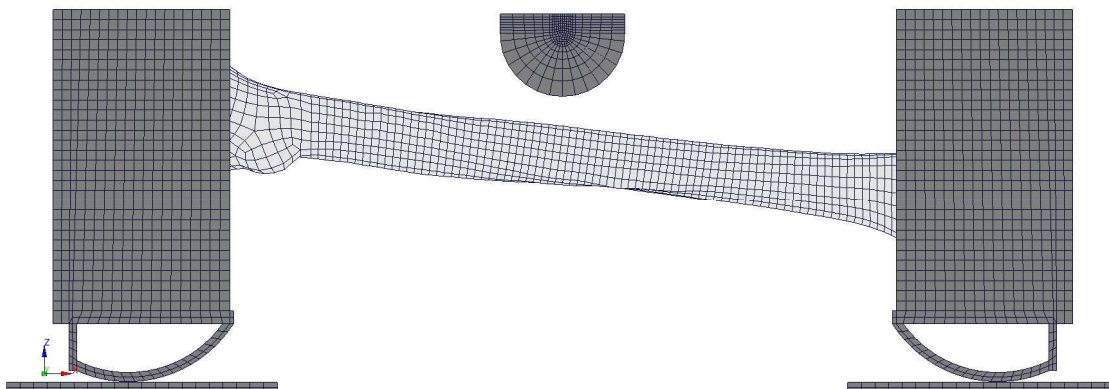
A5 Erzeugung LMU-Population



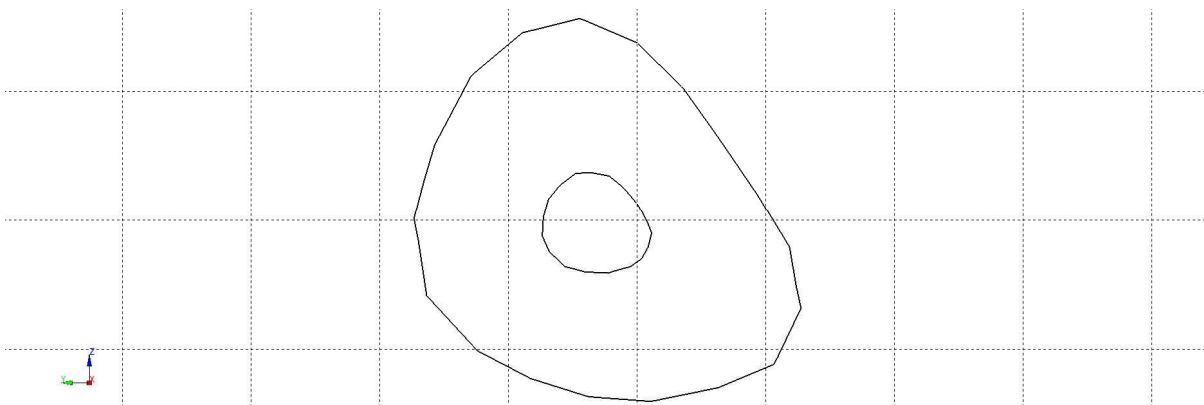
a) Darstellung der Knochenvernetzung und Schnittlage des Knochens M30



b) Darstellung der Vernetzung des Knocheninneren (M30)

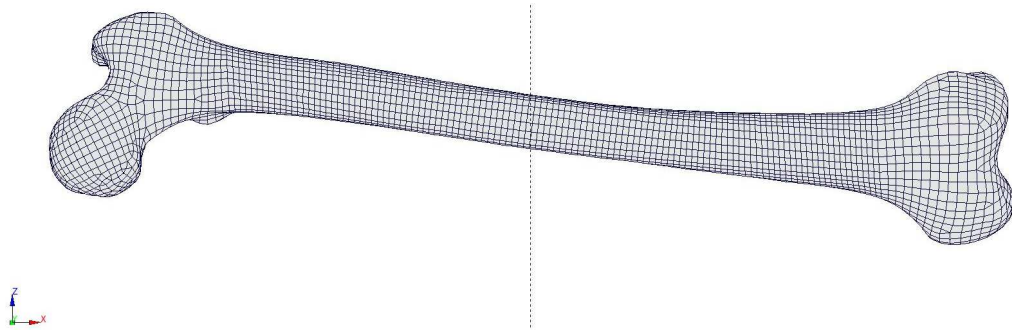


c) Knochen eingebettet in den Simulationsaufbau (M30)

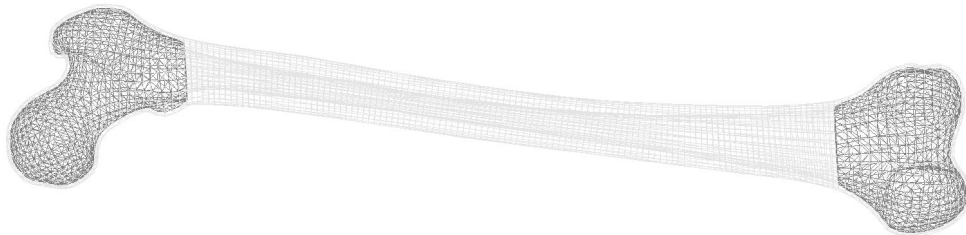


d) Knochenquerschnitt resultierend aus dem Schnitt dargestellt in Graphik a)

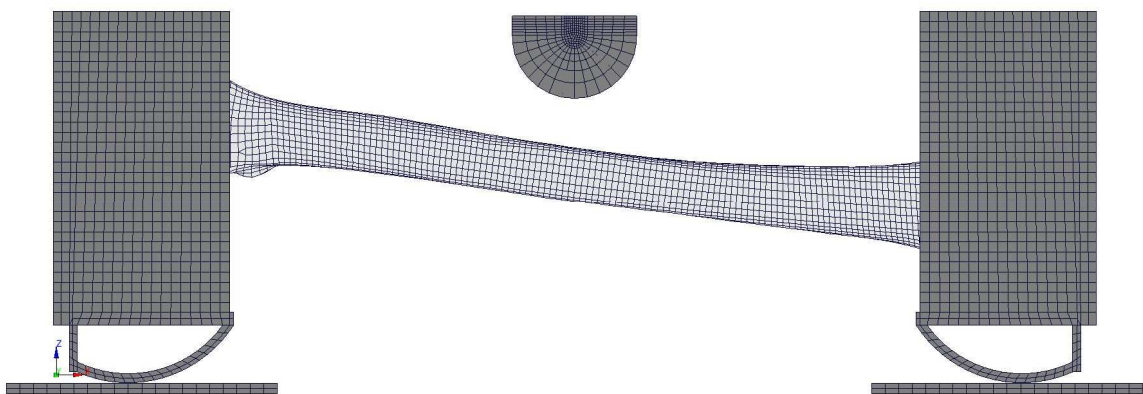
Abbildung A5.1 – Modellierung Knochen M30



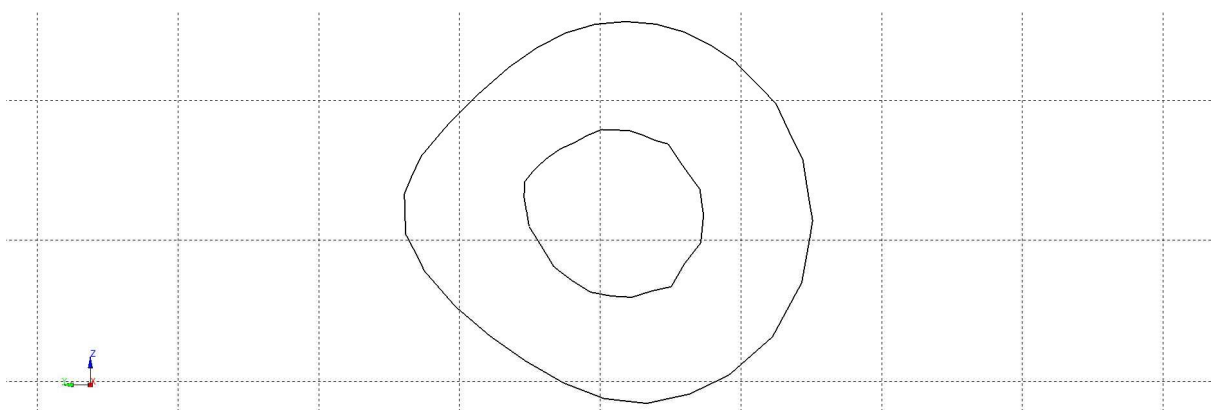
a) Darstellung der Knochenvernetzung und Schnittlage des Knochens M31



b) Darstellung der Vernetzung des Knocheninneren (M31)

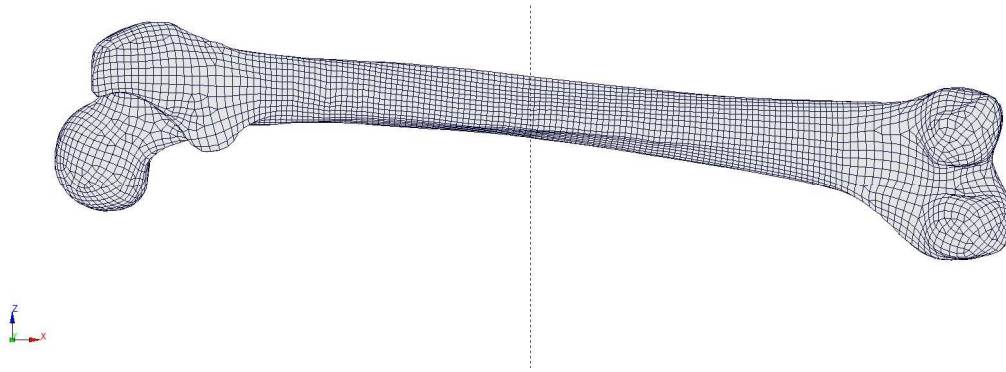


c) Knochen eingebettet in den Simulationsaufbau (M31)

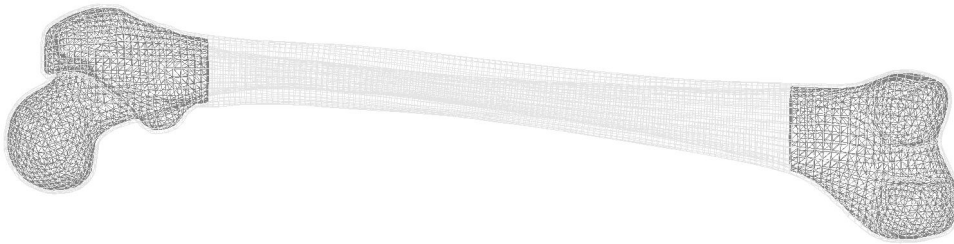


d) Knochenquerschnitt resultierend aus dem Schnitt dargestellt in Graphik a)

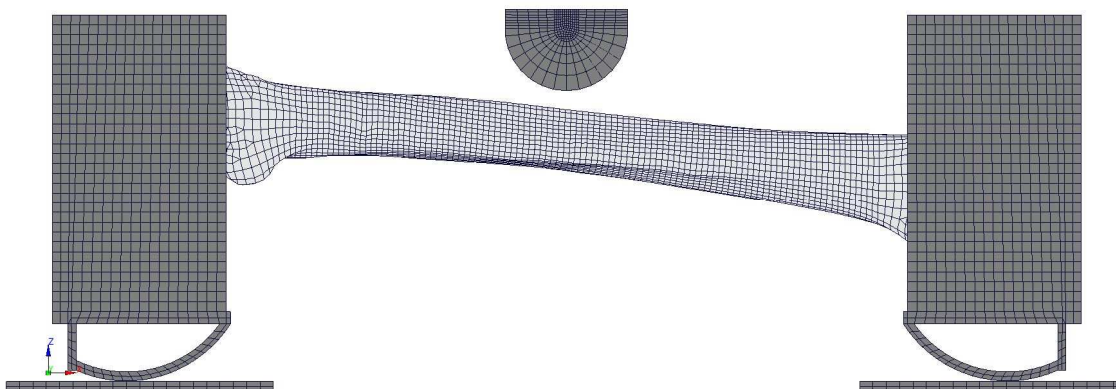
Abbildung A5.2 – Modellierung Knochen M31



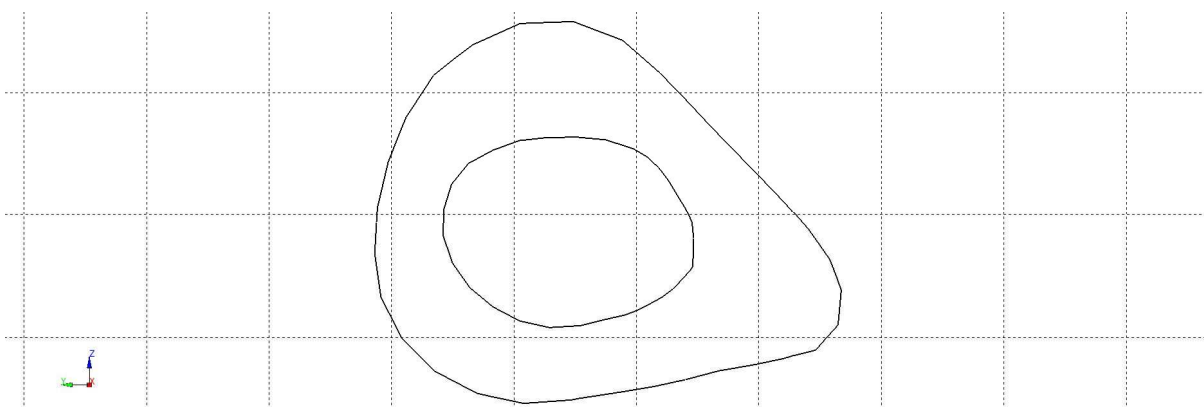
a) Darstellung der Knochenvernetzung und Schnittlage des Knochens M46



b) Darstellung der Vernetzung des Knocheninneren (M46)

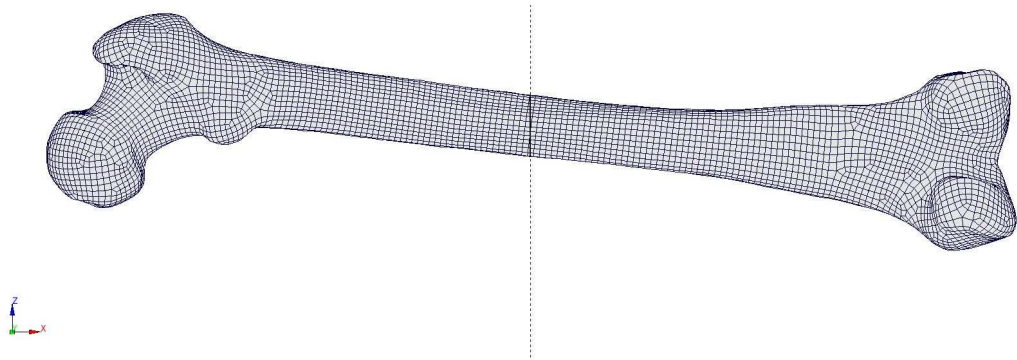


c) Knochen eingebettet in den Simulationsaufbau (M46)

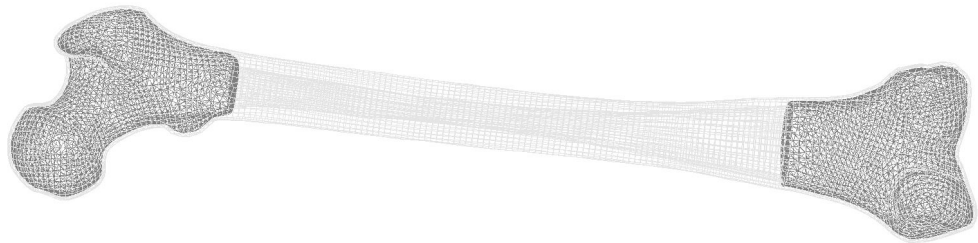


d) Knochenquerschnitt resultierend aus dem Schnitt dargestellt in Graphik a)

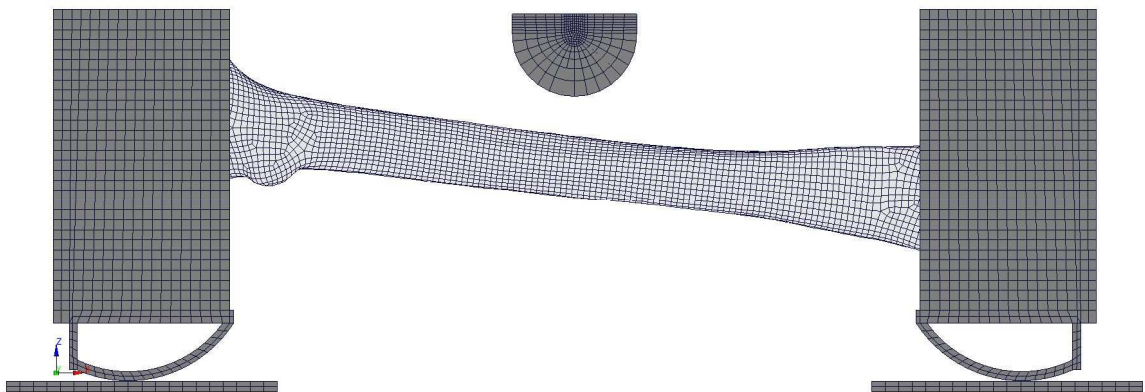
Abbildung A5.3 – Modellierung Knochen M46



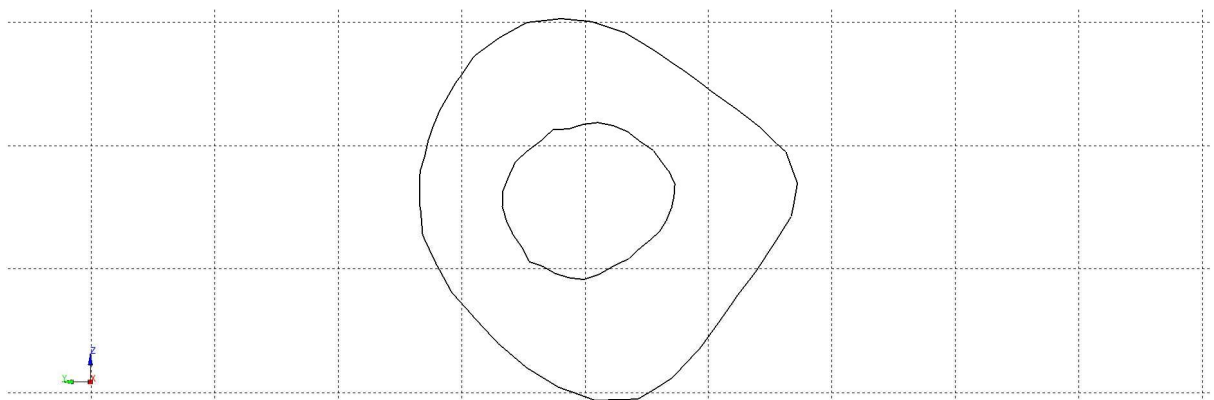
a) Darstellung der Knochenvernetzung und Schnittlage des Knochens M100



b) Darstellung der Vernetzung des Knocheninneren (M100)



c) Knochen eingebettet in den Simulationsaufbau (M100)



d) Knochenquerschnitt resultierend aus dem Schnitt dargestellt in Graphik a)

Abbildung A5.4 – Modellierung Knochen M100

A6 LMU-Population – Übersicht Vernetzungsqualität

	Target Value	Criteria	M 22	M30	M31	M46	M100
Min. Edge Length	1	No. Violated (%)	1.17	0.04	0.46	0.15	0.0
		Min. Value	0.728	0.956	0.9024	0.3837	1.011
		Max. Value	8.956	7.46	8.195	7.719	6.404
Max. Edge Length	5	No. Violated (%)	64.46	90.75	63.1	67.20	32.15
		Min. Value	2.409	3.15	2.738	2.751	1.805
		Max. Value	12.64	12.54	12.59	10.805	9.525
Max. Aspect Ratio	8	No. Violated (%)	0.0	0.0	0.0	0.0	0.0
		Min. Value	0.9975	0.94	0.988	0.9419	0.9057
		Max. Value	4.601	4.93	4.129	6.472	3.579
Max. Face Warping	50	No. Violated (%)	0.0	0.0	0.0	0.0	0.0
		Min. Value	0.0635	0.083	0.03232	0.04651	0.02272
		Max. Value	30.263	32.55	25.51	31.24	19.879
Max. Face Skewness	70	No. Violated (%)	0.0	0.0	0.0	0.0	0.0
		Min. Value	0.349	0.6974	0.3683	0.3483	0.202
		Max. Value	57.175	59.842	66.529	57.111	56.351
Min. Jacobian	0.3	No. Violated (%)	0.0	0.0	0.0	0.0	0.0
		Min. Value	0.5119	0.4076	0.4336	0.3169	0.4369
		Max. Value	1.0	1	1	1	1
Tetra Collapse	0.2	No. Violated (%)	0.0	0.0	0.0	0.0	0.0
		Min. Value	0.3303	0.332	0.2232	0.3468	0.3468
		Max. Value	0.8065	0.8065	0.8065	0.8065	0.8065
Tetra Aspect Ratio	8	No. Violated (%)	0.0	0.0	0.0	0.0	0.0
		Min. Value	1.03	1.047	1.063	1.045	1.038
		Max. Value	2.717	3.432	3.546	2.944	3.243

A7 Ergebnisse Konzeptpopulation I

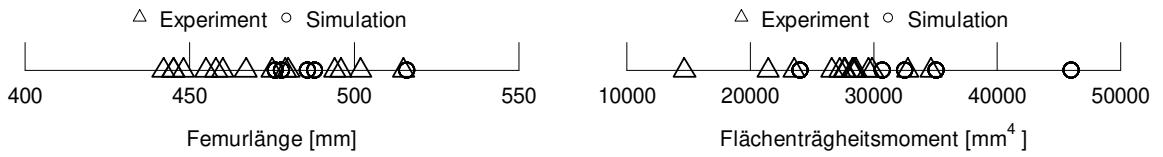


Abbildung A7.1 – Vergleich der experimentellen und numerischen Metadaten

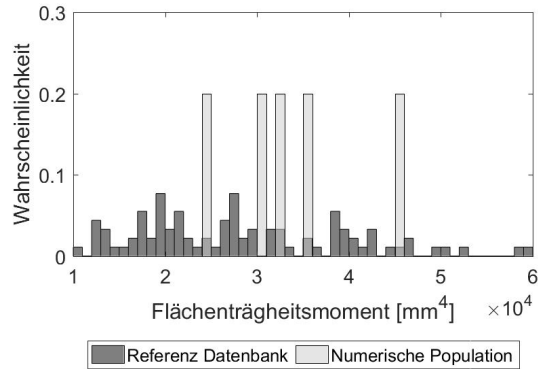
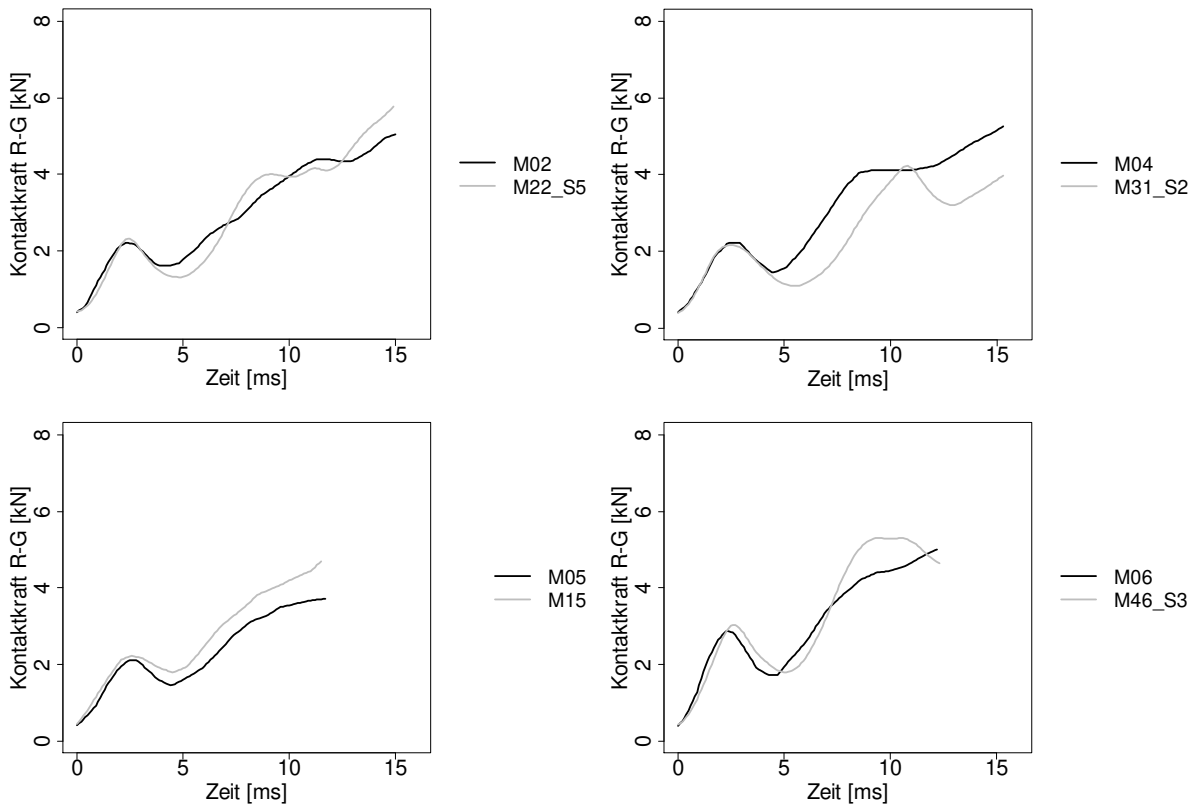


Abbildung A7.2 – Vergleich der Flächenträgheitsmomente der numerischen Population gegen Referenz-Datenbank



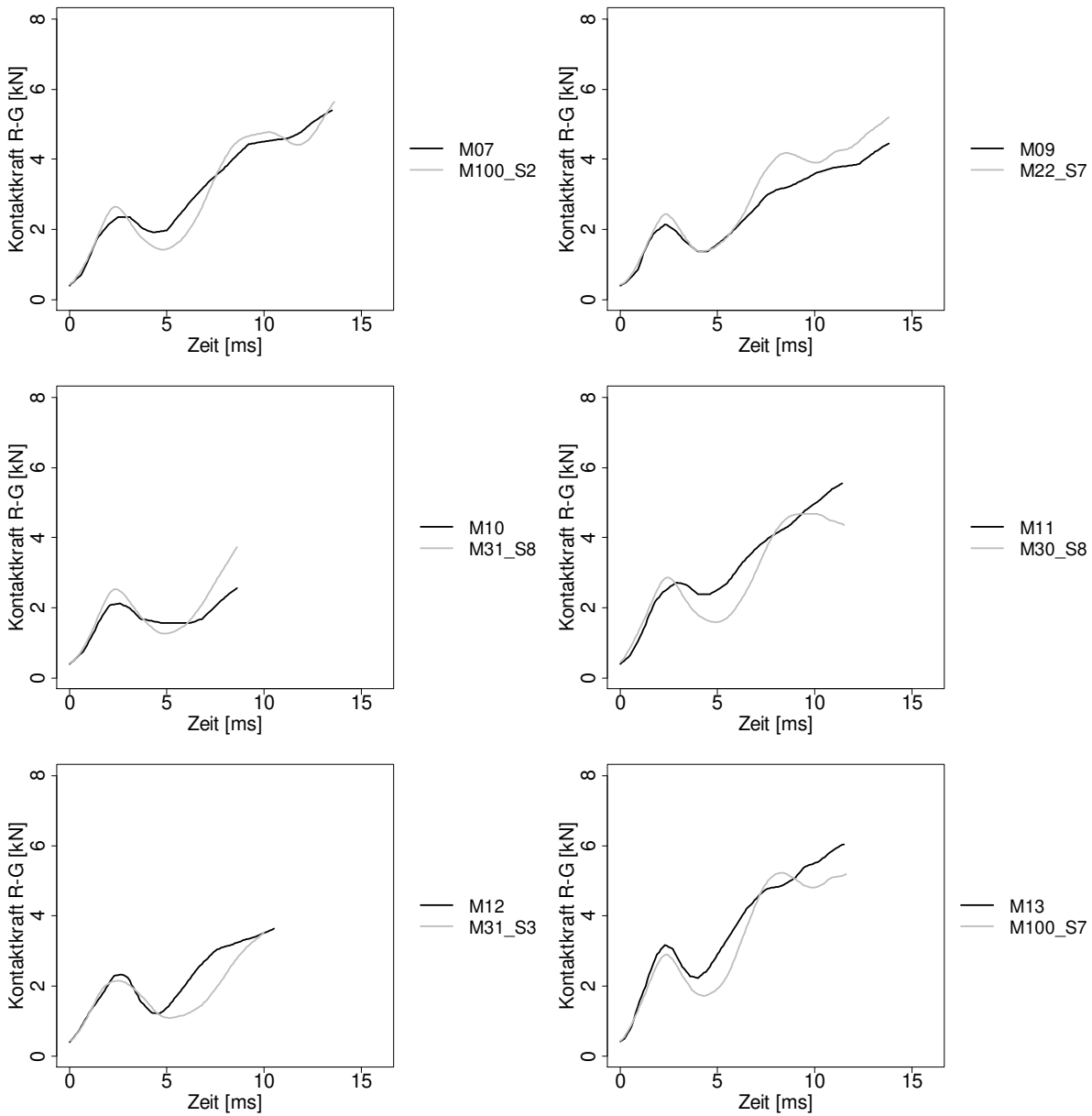


Abbildung A7.3 – Kombinierte Kurvenpaare der untersten Ebene des Dendrogramms

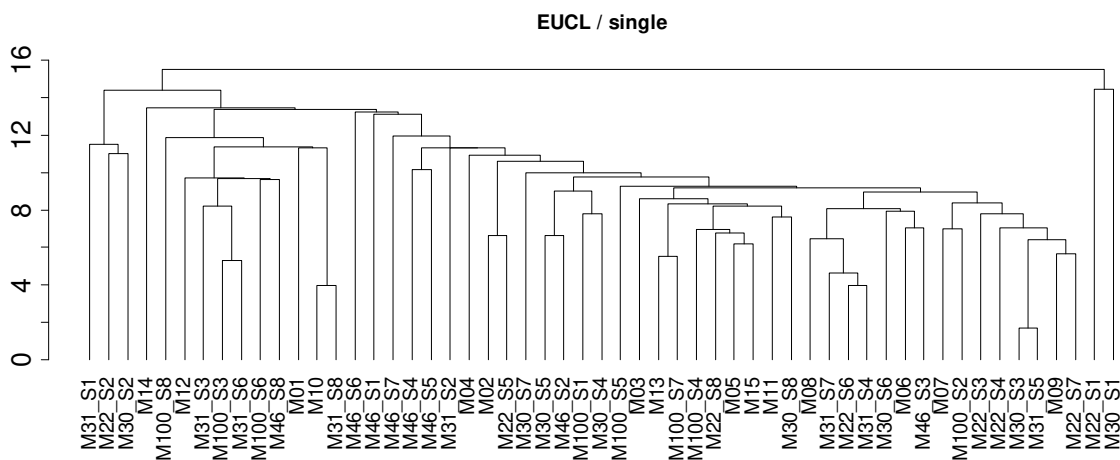


Abbildung A7.4 – Dendrogramm des Single-Linkage-Algorithmus mit euklidischer Distanz-Metrik

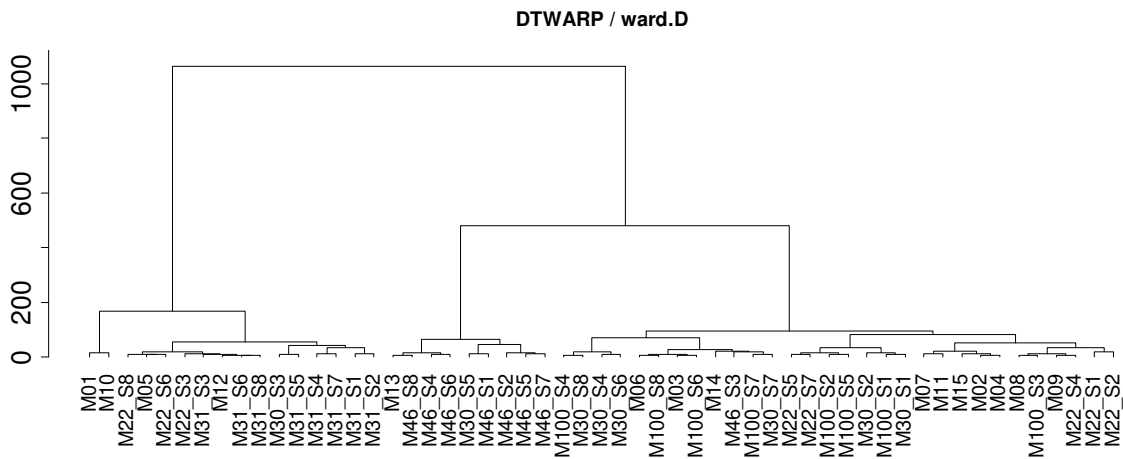


Abbildung A7.5 – Dendrogramm Ward-Algorithmus mit dynamischer Zeitverzerrung

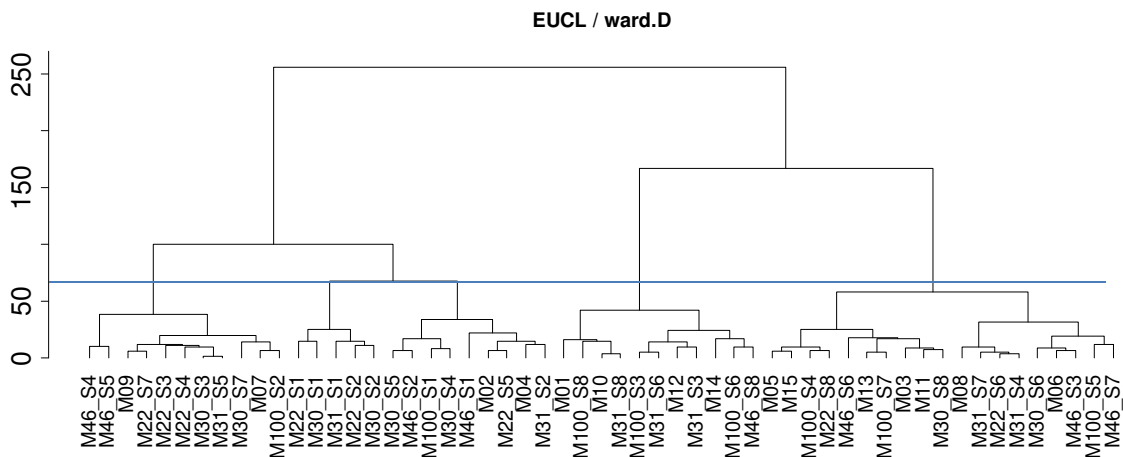


Abbildung A7.6 – Dendrogramm Ward-Algorithmus mit euklidischer Distanz-Metrik

Tabelle A7.1 – Zusammenfassung der geeigneten int. VIs für das euklidische Distanzmaß

Bewertungsgüte	Fusionierungsalgorithmus	Abkürzung int. VI	Clusteranzahl
Sehr gut geeignet	Average-Linkage	beale	4
Sehr gut geeignet	Mc Quitty	beale	2
Geeignet	Average-Linkage	ptbiserial	3
Geeignet	Complete-Linkage	beale	12
Geeignet	Complete-Linkage	ptbiserial	4
Geeignet	Mc Quitty	ptbiserial	3
Geeignet	Ward.D	ptbiserial	5
Geeignet	Ward.D2	ptbiserial	5
Geeignet	Single-Linkage	kl	18
Geeignet	Single-Linkage	ch	54
Geeignet	Single-Linkage	beale	-
Geeignet	Single-Linkage	ptbiserial	9

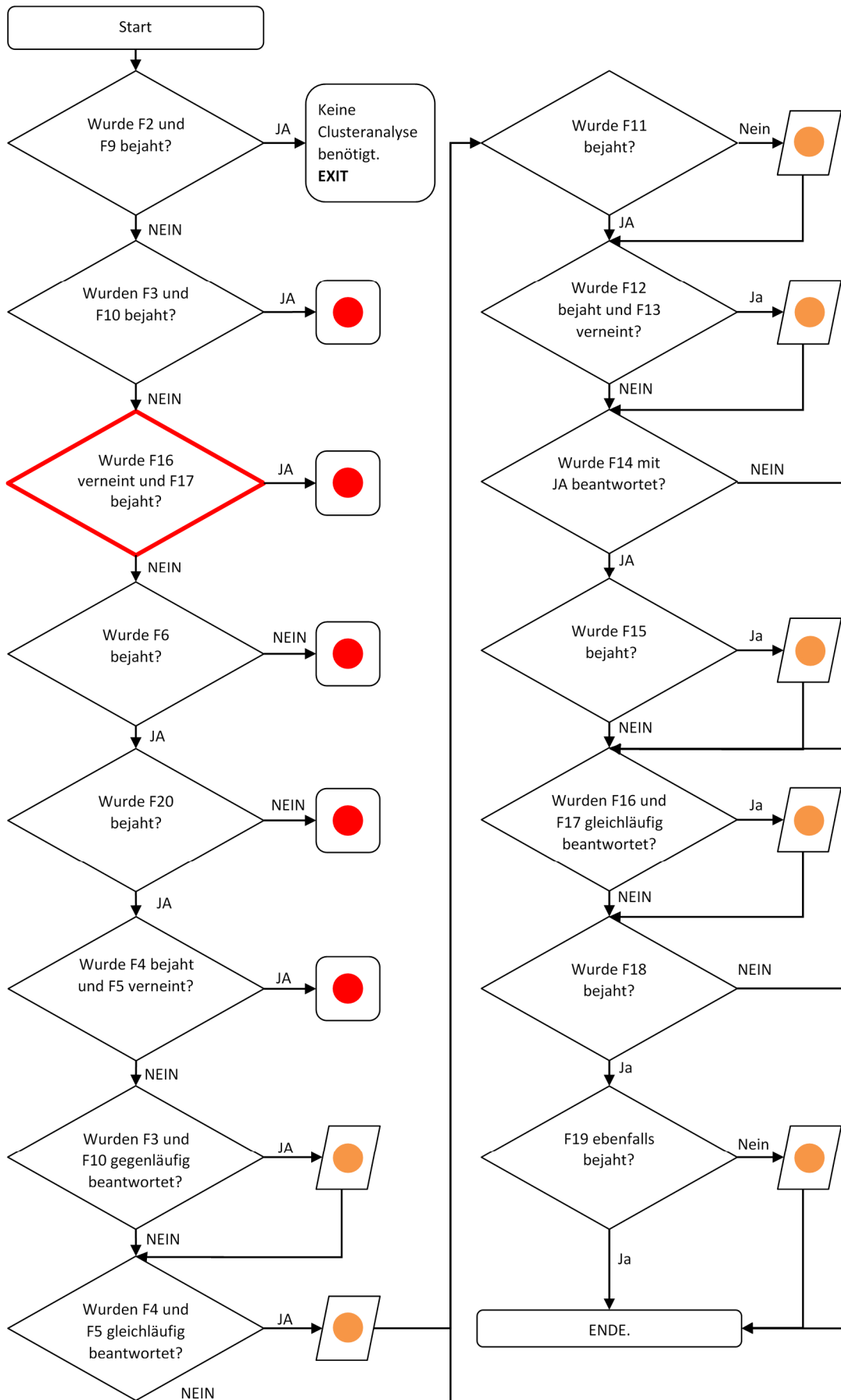


Abbildung A7.7 – Farbliche Markierung des Abbruchs des Validierungsprozesses

A8 Ergebnisse Konzeptpopulation II

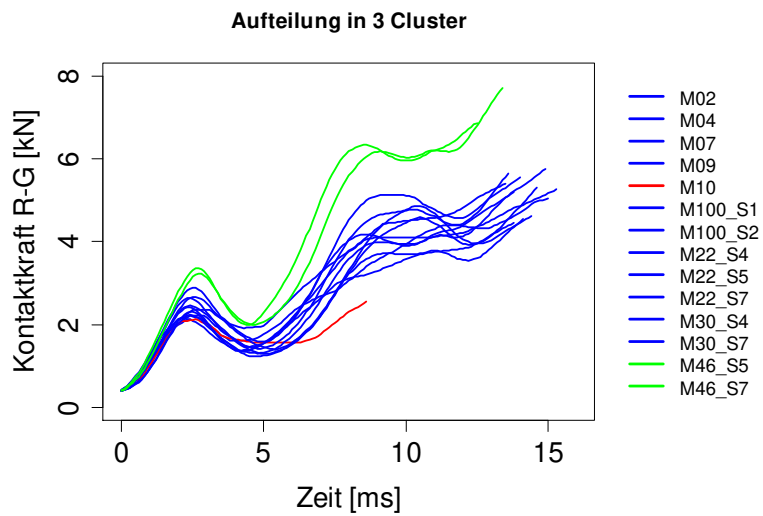


Abbildung A8.1 – Visuelle Aufteilung in 3 Cluster

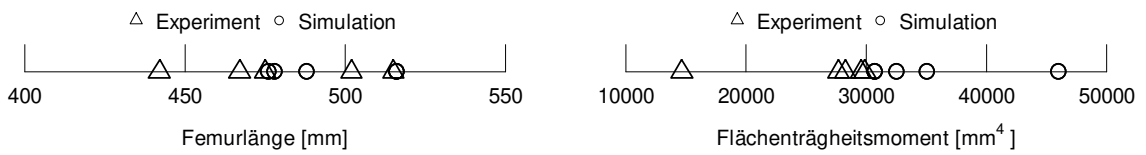


Abbildung A8.2 – Vergleich der Metadaten

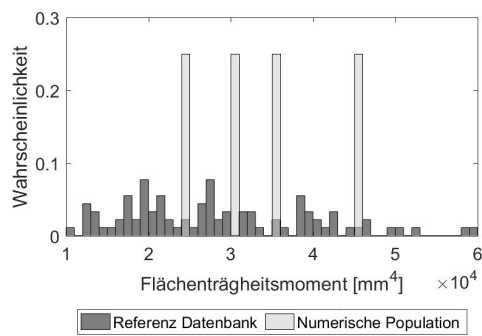


Abbildung A8.3 – Vergleich der Flächenträgheitsmomente anhand veröffentlichter Werte

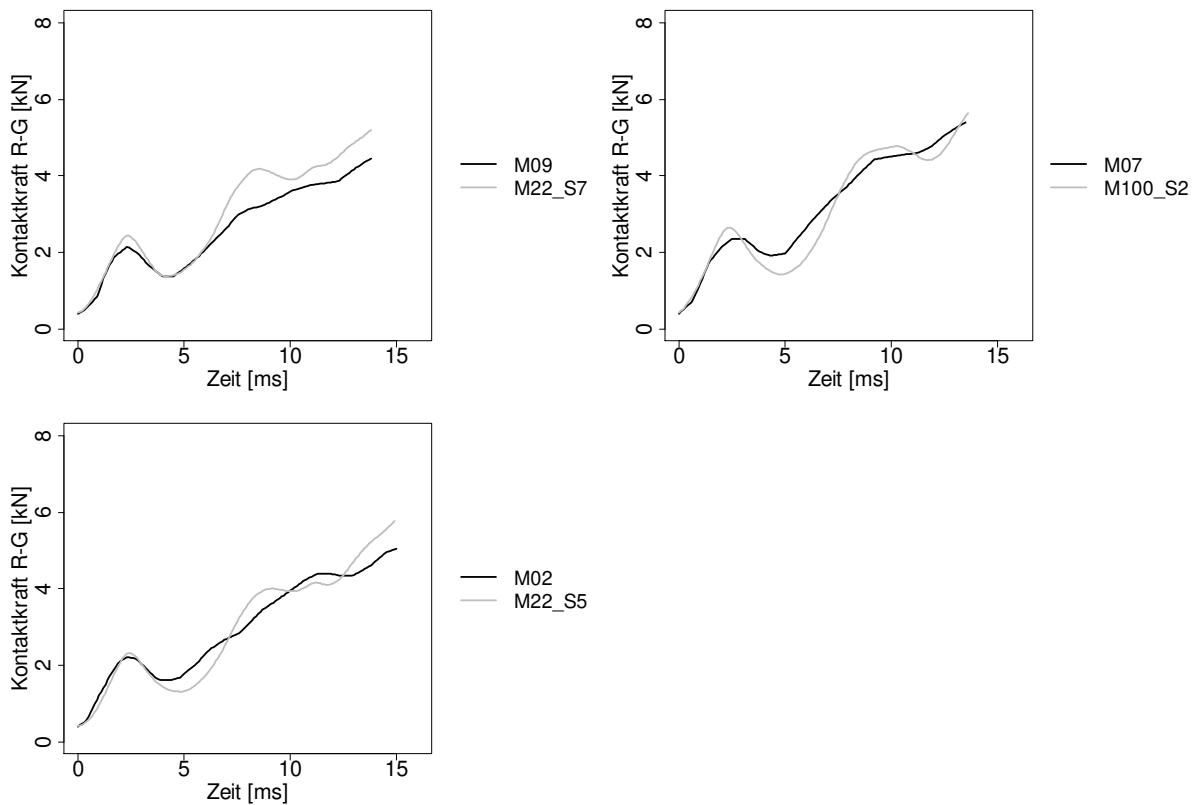


Abbildung A8.4 – Kombinierte Kurvenpaare der experimentellen Kurve mit der zugehörigen numerischen Modellvorhersage

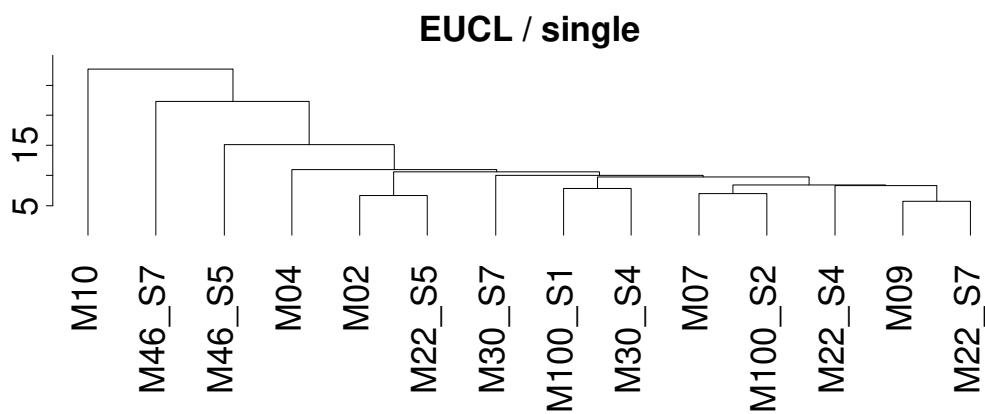


Abbildung A8.5 – Dendrogramm des Single-Linkage-Algorithmus mit euklidischer Distanz-Metrik

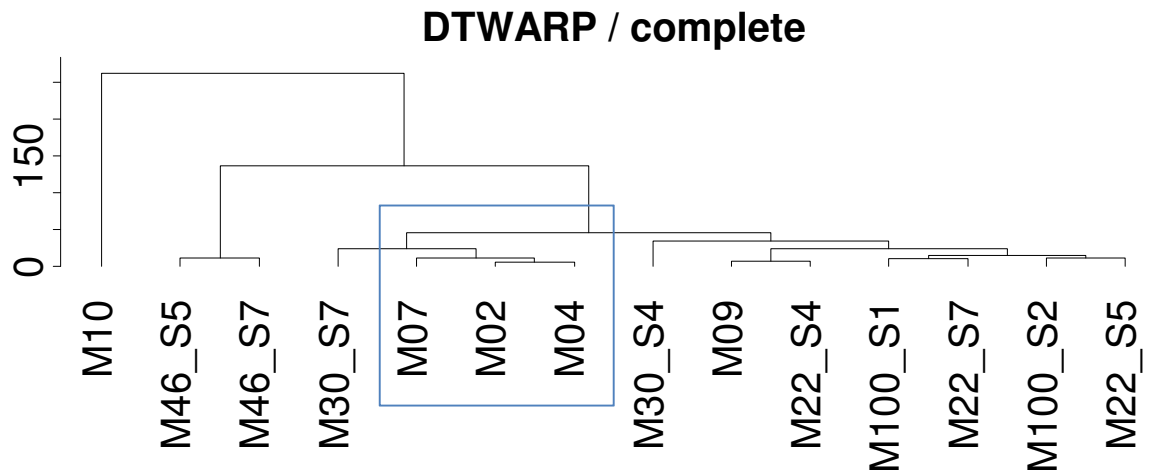


Abbildung A8.6 – Dendrogramm Complete-Linkage-Algorithmus mit dynamischer Zeitverzerrung

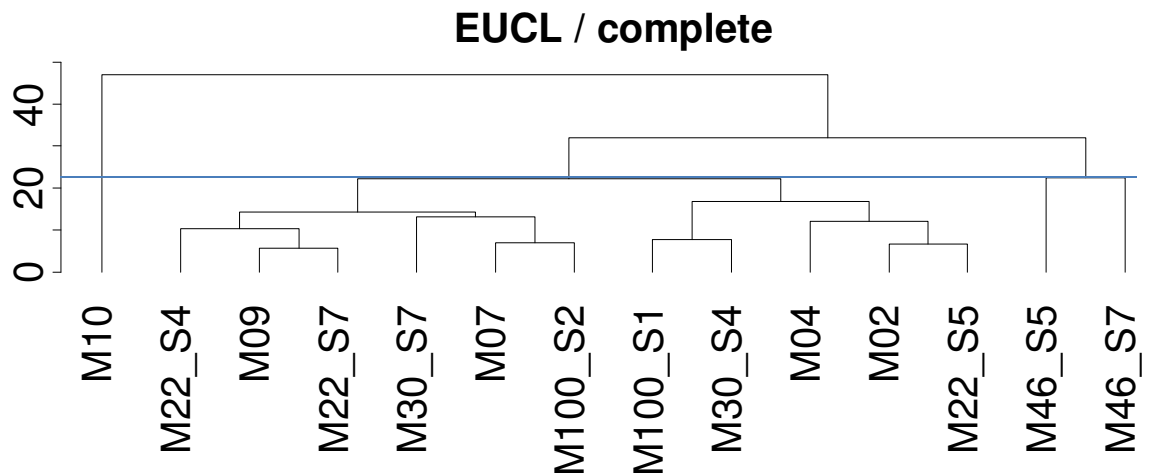


Abbildung A8.7 – Dendrogramm Complete-Linkage-Algorithmus mit euklidischer Distanz-Metrik

Tabelle A8.1 – Zusammenfassung der geeigneten int. VIs für das euklidische Distanzmaß

Bewertungsgüte	Fusionierungsalgorithmus	Abkürzung int. VI	Clusteranzahl
Sehr gut geeignet	Average-Linkage	beale	-
Sehr gut geeignet	Mc Quitty	beale	-
Geeignet	Average-Linkage	ptbiseri	4
Geeignet	Complete-Linkage	beale	-
Geeignet	Complete-Linkage	ptbiseri	4
Geeignet	Mc Quitty	ptbiseri	4
Geeignet	Ward.D	ptbiseri	2
Geeignet	Ward.D2	ptbiseri	2
Geeignet	Single-Linkage	kl	8
Geeignet	Single-Linkage	ch	4
Geeignet	Single-Linkage	beale	2
Geeignet	Single-Linkage	ptbiseri	4

Danksagung

Die vorliegende Arbeit entstand am Institut für Rechtsmedizin der Ludwig-Maximilians-Universität in München in Zusammenarbeit mit dem Fachbereich für Fahrzeugsicherheit der Audi AG.

An erster Stelle möchte ich mich ganz herzlich bei meinem Doktorvater Prof. Dr. Steffen Peldschus für die Möglichkeit bedanken, diese Dissertation in der Abteilung für Biomechanik und Unfallforschung anfertigen zu können. Ich möchte mich bei ihm besonders für die konstruktiven Diskussionen und zahlreichen Ideen während meiner gesamten Promotionszeit, sowie für die angenehme Betreuung bedanken.

Zudem gilt mein Dank Herrn Christian Schramm, meinem Leiter während der Promotionszeit bei der Audi AG, der meine Promotion seitens der Audi AG unterstützte und förderte. Ich konnte in seinem Team wertvolle Einblicke in Entwicklungsprozesse im Bereich der Fahrzeugsicherheit erlangen und dennoch stets fokussiert an meiner Promotion arbeiten. Besonderer Dank gilt meinen Kollegen Herrn Franz Fürst und Herrn Dr. Giacomo Marini für die unermüdliche Unterstützung während der gesamten Promotionszeit.

Zuletzt möchte ich mich bei meiner Familie, meinen Freunden und Bekannten bedanken, dass sie für die notwendige Abwechslung zu meiner Arbeit gesorgt haben.

Affidavit



Eidesstattliche Versicherung

Schneider, Sonja Andrea

Name, Vorname

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel:

Entwicklung einer populationsbasierten Validierungsmethodik für FE-Menschmodelle

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

Ingolstadt, 04.07.2022

Ort, Datum

Sonja Andrea Schneider

Unterschrift Doktorandin bzw. Doktorand