

---

# Selektieren und Kombinieren von Modellen unter Berücksichtigung der Problematik fehlender Daten

Michael Schomaker

---



Dissertation  
an der Fakultät für Mathematik, Informatik und Statistik  
der Ludwig-Maximilians-Universität München

München, im Dezember 2009

## **Bibliografische Information der Deutschen Nationalbibliothek**

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

1. Aufl. - Göttingen: Cuvillier, 2010

Zugl.: München (Univ.), Diss., 2009

978-3-86955-330-6

Berichterstatter: PD Dr. Christian Heumann

Auswärtiger Gutachter: Prof. Dr. Susanne Rässler

Rigorosum: 23. Februar 2010

© CUVILLIER VERLAG, Göttingen 2010

Nonnenstieg 8, 37075 Göttingen

Telefon: 0551-54724-0

Telefax: 0551-54724-21

[www.cuvillier.de](http://www.cuvillier.de)

Alle Rechte vorbehalten. Ohne ausdrückliche Genehmigung des Verlages ist es nicht gestattet, das Buch oder Teile daraus auf fotomechanischem Weg (Fotokopie, Mikrokopie) zu vervielfältigen.

1. Auflage, 2010

Gedruckt auf säurefreiem Papier

978-3-86955-330-6

Do not fear to be eccentric in opinion,  
for every opinion now accepted was once eccentric.

Bertrand Russel



## Danksagung

---

An dieser Stelle möchte ich mich bei all denjenigen Menschen bedanken, die in unterschiedlicher Form zum Gelingen dieser Arbeit beigetragen haben. Mein besonderer Dank gilt Christian Heumann für die Betreuung der Dissertation, viele hilfreiche Vorschläge und Denkanstöße, seine Zeit und sein Interesse sowie Gespräche, die auch über den Tellerrand der Statistik hinausgingen. Helge Toutenburg danke ich für seine große Unterstützung während der letzten Jahre, eine immer produktive Zusammenarbeit und vor allem natürlich für seine humorvolle Art und eine abwechslungsreiche Zeit. Bei Alan Wan bedanke ich mich für sein außerordentliches Engagement, viel konstruktive und fachkundige Kritik und dass er mir einen in jeder Hinsicht lohnenden Aufenthalt in Hongkong ermöglicht hat. Mein Dank geht auch an Shalabh für viele wertvolle Hinweise in puncto konsequenten, effektiven wissenschaftlichen Arbeitens, viele unübertroffene Lebensweisheiten und eine äußerst herzliche Zusammenarbeit. Susanne Rässler danke ich für ihre Zeit und ihr Interesse und der Übernahme des auswärtigen Gutachtens.

Bedanken möchte ich mich auch bei Ingrid Kreuzmair für die schöne und ereignisreiche Zeit in und um Raum 340, Michael Obermeier für ausführliches Korrekturlesen und sachdienliche Hinweise zu einigen Absurditäten deutscher Rechtschreibung, Birgit Schrödle für ein detailliertes inhaltliches Feedback und wertvolle Anmerkungen zur Modellmitteilung, Benjamin Hofner für viele Vorschläge zur Präsentation meiner Inhalte, Stefan Pilz für hilfreiche Anmerkungen und Korrekturen zu dieser Arbeit, Jan Ulbricht für angeregte Diskussionen zu Akaike & Co. und Gero Walter und Andrea Wiencierz für Gespräche rund um die Statistik und das tägliche Dolce Vita.

Großer Dank gilt auch meinen Eltern, die mich in meinen Plänen immer unterstützt und ermutigt haben. Auch hätte ich ohne die zahlreichen Nachmittage in Iffezheim wohl nie Statistik studiert...



---

## Zusammenfassung

In den letzten Jahren haben sich Modellmittelungsverfahren als Alternative zur Modellselektion etabliert. Anstatt sich auf ein einziges Siegermodell zu beschränken, werden hierbei mehrere konkurrierende Modelle betrachtet und ihre Parameterschätzer gewichtet miteinander kombiniert. Das Hauptaugenmerk liegt dabei meist auf der Konstruktion der Gewichte, wie auch der Optimalität der daraus resultierenden gewichteten Parameterschätzung. In der vorliegenden Arbeit werden verschiedene Konzepte bayesianischer, vor allem aber auch frequentistischer Modellmittelung (Frequentist Model Averaging, FMA) erläutert und ihre Stärken und Schwächen gegenüber einer Vielzahl an traditionellen Modellselektionsmethoden herausgestellt.

Schwerpunkt ist dabei die Konstruktion und Diskussion verschiedener Strategien zur Verwendung von FMA-Methoden unter Berücksichtigung der Problematik fehlender Daten. Hierfür werden zwei Kernkonzepte vorgeschlagen: Der erste Ansatz konstruiert Gewichte für einen FMA-Schätzer auf Basis eines für fehlende Daten adjustierten Kriteriums, welches der aktuellen Literatur aus dem Bereich der Modellselektion entstammt und das das im Kontext fehlender Werte bekannte Prinzip des inverse probability weighting verwendet; der zweite Ansatz ersetzt die fehlenden Werte durch einfache und multiple Imputationen, um darauf aufbauend geeignete Punktschätzungen und deren Varianz mit Hilfe bekannter Modellmittelungsschätzer zu konstruieren. Zu diesem Zweck wird auch ein rekursiver Imputationsalgorithmus präsentiert, der die geläufige Idee einer Regressionsimputation unter Verwendung generalisierter additiver Modelle verallgemeinert.

Die Arbeit zeigt die Eigenheiten, Stärken und Schwächen der vorgestellten Ansätze im Kontext von linearen und logistischen Regressionsanalysen anhand weitreichender Monte-Carlo-Simulationen auf und diskutiert am Beispiel der Faktorenanalyse mögliche Erweiterungen und Verallgemeinerungen der angeführten Schätzer für weitere multivariate, statistische Analysemethoden. Alle Verfahren werden an realen Datensätzen illustriert.

Es zeigt sich, dass in vielen Situationen beide vorgestellten Konzepte einem Verwerfen der nicht-vollständigen Beobachtungen vorzuziehen sind, die Strategie einer Modellmittelung nach Imputation in der Regel bessere Resultate erzielt als die Verwendung eines FMA-Schätzers, der Gewichte auf Basis eines für fehlende Daten adjustierten Kriteriums verwendet, und insbesondere die technisch weniger aufwändigen Modellmittelungsverfahren zu besseren Schätzungen führen als diejenigen, die aus einer klassischen Modellselektion resultieren.

---





---

## Abstract

Model averaging or combining is often considered as an alternative to model selection. Rather than attaching to a single „winning“ model, a model average estimator weights across the estimators of many potential models. The construction of appropriate weights and the properties of the resulting estimator are issues of high concern.

In this work, Frequentist Model Averaging (FMA) is considered extensively and strategies for the application of FMA methods in the presence of missing data based on two distinct approaches are presented and compared to concepts of traditional model selection: The first approach combines estimates from a set of appropriate models which are weighted by scores of a missing data adjusted criterion developed in the recent literature of model selection. The second approach averages over the estimates of a set of models with weights based on conventional approaches but with the missing data replaced by imputed values prior to estimating the models. For this purpose several imputation methods that have been programmed in currently available statistical software are considered, and a recursive algorithm is proposed to implement an extended version of regression imputation that relies on generalized additive models.

Focusing on the linear and binary logistic regression model, the properties of the FMA estimators resulting from these strategies are explored by means of Monte-Carlo studies and extensions of the presented methodologies to other areas of multivariate statistical modeling are discussed briefly in the context of factor analysis by way of example. As an illustration, the proposed methods are applied to real data.

The results show that in many situations both approaches are superior to a Complete Case Analysis, averaging after imputation is normally preferred to averaging using weights that adjust for the missing data and model average estimators, especially those of technically easy conception, often provide better estimates than those resulting from any single model.

---



# Inhaltsverzeichnis

<b>1. Einleitung</b>	1
<b>2. Modelle in Wissenschaft und Statistik</b>	9
<b>3. Modellselektion</b>	17
3.1 Modellselektion durch Betrachtung der Parameterschätzungen	19
3.1.1 Sukzessives Testen von Hypothesen	19
3.1.2 Shrinkage	23
3.2 Modellselektion auf Basis von Vorhersagefehlern	24
3.2.1 Mallows Kriterium ( $C_p$ )	24
3.2.2 Erwarteter Vorhersagefehler (EPE)	27
3.2.3 Kreuzvalidierungskriterium (CV)	28
3.2.4 Finaler Vorhersagefehler (FPE)	29
3.2.5 Weitere Ansätze	29
3.3 Informationstheoretische Selektionskriterien	30
3.3.1 Akaikes Informationskriterium (AIC)	31
3.3.2 Takeuchis Informationskriterium (TIC)	36
3.3.3 Regularisiertes Informationskriterium (RIC)	37
3.3.4 Korrigiertes Informationskriterium ( $AIC_c$ )	38
3.3.5 Informationskriterium bei Überdispersion (QAIC)	39
3.3.6 Weitere Ansätze	40

3.4	Bayesianische Modellselektion . . . . .	40
3.4.1	Schwarzsches Bayes-Kriterium (SBC) . . . . .	40
3.4.2	Weitere Ansätze . . . . .	42
3.5	Weitere Ansätze . . . . .	43
3.5.1	Minimum Description Length . . . . .	43
3.5.2	Dimensionskonsistente Kriterien . . . . .	45
3.5.3	Ad-hoc Ansätze . . . . .	46
3.5.4	Robuste Verfahren . . . . .	48
3.6	Asymptotische Optimalität . . . . .	50
<b>4.</b>	<b>Modellmittelung . . . . .</b>	<b>55</b>
4.1	Der bayesianische Ansatz . . . . .	57
4.2	Frequentistische Ansätze . . . . .	59
4.2.1	Kriteriums-basierte Schätzungen . . . . .	60
4.2.2	Der MMA-Schätzer . . . . .	61
4.2.3	Der OPT-Schätzer . . . . .	63
4.2.4	Schätzung der Varianz . . . . .	64
4.2.5	Modellmittelung in der Faktorenanalyse . . . . .	65
<b>5.</b>	<b>Berücksichtigung fehlender Werte . . . . .</b>	<b>69</b>
5.1	Modellselektion bei fehlenden Daten . . . . .	76
5.1.1	Gewichtetes Akaike Kriterium ( $AIC_W$ ) . . . . .	77
5.1.2	Selektion nach Imputation . . . . .	80
5.1.3	Weitere Ansätze . . . . .	88
5.2	Modellmittelung bei fehlenden Daten . . . . .	89
5.2.1	Mittelung mit adjustierten Kriterien . . . . .	90
5.2.2	Mittelung nach Imputation . . . . .	91

<b>6. Simulationsstudien</b>	95
6.1 Lineare Regression	95
6.2 Logistische Regression	109
6.3 Die Auswirkungen multipler Imputation	121
6.4 Zusammenfassung	131
<b>7. Anwendungsbeispiele</b>	135
7.1 Phasenangepasste Führung von Wachstumsunternehmen	135
7.1.1 Analyse der Zufriedenheit	137
7.1.2 Analyse der Effektivität	144
7.2 Muskeldystrophie vom Typ Duchenne	151
7.3 Olympischer Zehnkampf	158
<b>8. Résumé</b>	169
<b>Anhang</b>	175
<b>A. Symbolverzeichnis</b>	177
A.1 Lateinische Symbole	177
A.2 Griechische Symbole	178
A.3 Notation	179
A.4 Abkürzungen	180
<b>B. Detaillierte Simulationsergebnisse</b>	183
B.1 Lineare Regression	183
B.2 Logistische Regression	191
<b>C. Weitere Analysen</b>	199
<b>Literatur</b>	204



# 1. Einleitung

Die Konstruktion, die Wahl und das Verständnis von Modellen sind von zentraler Bedeutung innerhalb vieler wissenschaftlicher Erkenntnisprozesse: Modelle helfen, die wesentlichen Strukturen wahrnehmbarer und nicht wahrnehmbarer Gegenstandsbereiche aufzudecken und damit Phänomene darzustellen und zu erklären; um es mit den Worten von Frigg und Hartmann (2006) zu formulieren:

*„Models are vehicles for learning about the world“*

Um Phänomene erfassen und damit auch deren relevante Effekte beschreiben und verstehen zu können, werden innerhalb der Statistik eine Vielzahl an multivariaten Verfahren wie etwa parametrische und nichtparametrische Regressionsmodelle, autoregressive Prozesse, Bild-, Kontur-, Faktor- oder auch Clusteranalysen verwendet. Die konkrete Entscheidung zugunsten einer adäquaten Modellierung, also beispielsweise die Wahl geeigneter Kovariablen bei Regressionsmodellen oder die Bestimmung der Anzahl der Faktoren in der Faktorenanalyse, ist ein klassisches Modellselektionsproblem, für das zahlreiche Methoden unterschiedlicher Konzeption existieren. Typischerweise wird für die interessierende/n Größe/n eine im Kontext passende Verteilungsfamilie gewählt und es werden für die Wahl der zur Modellierung geeigneten Variablen sukzessive Hypothesentests, Shrinkageschätzungen, risikobasierte Entscheidungen auf Basis von Modellwahlkriterien, statistische Lernverfahren oder ad-hoc Vermutungen herangezogen. Einige aus diesem weitreichenden Feld repräsentative Ansätze werden im Verlauf der Arbeit noch näher diskutiert; eine ausführliche Übersicht findet sich auch bei Rao und Wu (2001).

## Modellmittelung

Ungeachtet der Schwierigkeit, ein passendes Verfahren für eine konkrete Fragestellung zu wählen, wird insbesondere seit Mitte der 1990er Jahre die Problematik der Modellselektionsunsicherheit in der Literatur diskutiert: Die Eigenschaften von Punktschätzungen statistischer Modelle, wie auch deren Varianz, hängen sowohl davon ab, auf welche Art und Weise das Modell gewählt wurde als auch von dessen stochastischen Begebenheiten.

In der Regel wird aber nur der letztgenannte Punkt beachtet und jede Form von Inferenz wird so durchgeführt, als wäre das betrachtete Modell *a priori* gewählt worden und die Schätzungen unabhängig von der Modellwahl. Tatsächlich ist dies nicht korrekt, da der Selektionsschritt datenbasiert ist und die Unsicherheit bezüglich der Modellselektion in den entsprechenden Schätzungen reflektiert werden sollte. Von Interesse sind damit also nicht die Eigenschaften der Schätzer bedingt auf *ein* gewähltes Modell, sondern die unbedingten Parameterschätzungen, die den durch die Modellselektion hervorgerufenen Inferenzschritt berücksichtigen, vergleiche auch Leeb und Pötscher (2003, 2005, 2006a, 2008b). Wird dieser Sachverhalt dennoch ignoriert, können als Konsequenz Parameterschätzungen verzerrt sein und die Varianz wird ob der nicht berücksichtigten Unsicherheit systematisch unterschätzt, was zu durchweg überoptimistischen Konfidenzintervallen führt.

Inzwischen ist weitgehend akzeptiert, dass als eine Möglichkeit, diesem Problem entgegenzutreten, Parameterschätzungen mehrerer Modelle kombiniert werden können. Kerngedanke ist dabei, dass nicht nur ein, sondern verschiedene Modellierungsansätze eine gute Beschreibung der Datenstruktur bieten können und eine gewichtete Mittelung vieler plausibler Schätzungen die Modellselektionsunsicherheit adäquat erfassen kann, also die Berechnung eines sinnvollen Schätzers *post model selection* ermöglicht. Dieser Gedanke wird in der Literatur erstmals sauber von Leamer (1978) formuliert, der einen bayesianischen Ansatz von *model averaging*, also Modellmittelung, propagiert: Hierbei werden für alle infrage kommenden Modelle posteriori-Wahrscheinlichkeiten berechnet und die entsprechenden Parameterschätzungen so kombiniert, dass die Schätzungen der Modelle mit höherer posteriori-Wahrscheinlichkeit auch ein insgesamt höheres Gewicht erhalten. Die zum damaligen Zeitpunkt noch sehr limitierten computergestützten Ressourcen ermöglichten jedoch keine konkrete Umsetzung der vorgeschlagenen Konzepte, so dass sich nur vereinzelte Veröffentlichungen in den darauffolgenden Jahren der Thematik annahmen. Erst die wegweisenden Artikel von Draper (1995) und Chatfield (1995) und eine mittlerweile deutlich erhöhte Computerpower führten zu neuen Denkanstößen und einer Vielzahl an kohärenten Ideen und Konzepten bayesianischer Modellmittelung innerhalb kürzester Zeit. Ein detaillierter Überblick findet sich unter anderem bei Hoeting et al. (1999).

Eine korrekte Umsetzung der Methodik verlangt jedoch weiterhin einen hohen computationalen Aufwand; insbesondere deswegen und auch aufgrund einiger strittiger philosophischer Positionen – so etwa die Frage was priori- und posteriori-Wahrscheinlichkeiten von Modellen überhaupt bedeuten – haben sich über die letzten 15 Jahre einige nicht-



bayesianische, in gewisser Weise frequentistische, Alternativen ergeben, die insgesamt einfacher und schneller umzusetzen sind, bisher jedoch nicht wirklich einem einheitlichen Konzept entsprechen. Aufgrund ihres großen, jedoch noch nicht vollständig erschlossenen Potentials bilden sie auch den Schwerpunkt der vorliegenden Arbeit. Einige Kerngedanken zur frequentistischen Methodik finden sich unter anderem bei Buckland, Burnham und Anderson (1997), Hjort und Claeskens (2003), Yuan und Yang (2005), Hansen (2007) und Schomaker, Wan und Heumann (2010).

Grundsätzlich stellen sich ungeachtet des zugrundeliegenden Paradigmas die Fragen: Welche Modelle sollen kombiniert werden? Wie sollen diese Modelle kombiniert werden? Was bedeutet dies für die kombinierten Parameterschätzungen? Prinzipiell lässt sich festhalten, dass die Punkt- und deren Varianzschätzungen für alle Modelle von Interesse gewichtet miteinander kombiniert werden sollten, wobei die Gewichte der einzelnen Modelle in der Regel so bestimmt werden, dass entweder über ein Selektionskriterium bzw. die posteriori-Wahrscheinlichkeit jedem Modell eine gewisse Plausibilität zugesprochen wird oder der finale, gemittelte Schätzer auf irgendeine Art und Weise optimal ist. Exemplarisch für diese beiden, grundsätzlich unterschiedlichen Vorgehensweisen werden in dieser Arbeit insbesondere dem Schätzer von Buckland, Burnham und Anderson (1997), der zur Gewichtung der Kandidatenmodelle exponentielle AIC-Gewichte verwendet sowie dem Schätzer von Hansen (2007), der in gewissem Sinne einen optimalen Modellmittelungsschätzer garantiert, besondere Beachtung geschenkt.

## **Fehlende Daten**

Die vorliegende Arbeit beschäftigt sich mit Modellselektion und Modellmittelung unter Berücksichtigung der Problematik fehlender Daten. In der Praxis sind fehlende Werte ein häufig auftretendes Problem, eine weitere Komponente der Unsicherheit, die verschiedenste Ursachen besitzen kann: etwa wenn in Meinungsumfragen die befragten Personen Antworten verweigern, bei klinischen Studien Patienten nicht über die volle Zeitspanne beobachtet werden können oder in naturwissenschaftlichen Experimenten Ergebnisse aufgrund fehlerhafter Messungen verworfen werden müssen. Um diese Problematik im Rahmen statistischer Inferenz explizit zu berücksichtigen, werden typischerweise zwei konzeptionell unterschiedliche Ansätze verfolgt:

- Zum einen können die fehlenden Werte durch andere, „plausible“ Werte ersetzt werden; so etwa indem konkrete Ausprägungen ähnlicher Beobachtungen impu-

tiert werden (Chen und Shao (2000)) oder die Abhängigkeitsstruktur der Variablen über ein passendes Regressionsmodell innerhalb der vollständigen Fälle erfasst wird und dessen Vorhersagen als Imputationen verwendet werden (Little und Rubin (2002)). Bei einem solchen oder ähnlichen Vorgehen wird jedoch vernachlässigt, dass die neuen, imputierten Werte nur datenbasierte Schätzungen für die wahren, fehlenden Daten sind und somit eine Unsicherheit bezüglich der Imputation vorliegt, die nicht vernachlässigt werden sollte. Dieser Sachverhalt kann durch die Verwendung multipler Imputationen (Rubin (1978), Rubin (1996)) berücksichtigt werden: Hierbei wird jeder fehlende Wert einer Datenmatrix durch  $M > 1$  zufällig gezogene Werte aus der prädiktiven a-posteriori-Verteilung der fehlenden Daten gegeben die beobachteten Daten (oder einer Approximation davon) ersetzt, wodurch  $M$  neue Datensätze zur Verfügung stehen, deren Schätzungen kombiniert werden können, um so die Unsicherheitskomponente bezüglich der Imputation zu erfassen.

- Zum anderen kann der Umstand fehlender Daten auch direkt im Inferenzprozess mitberücksichtigt werden, etwa über den EM-Algorithmus, bei dem die Likelihood des interessierenden Parameters gegeben die beobachteten Daten approximiert wird (Dempster, Laird und Rubin (1977)) oder auch unter Verwendung von Gewichtungsansätzen, bei denen typischerweise die Inferenz auf den gewichteten, vollständigen Fällen durchgeführt wird, um konsistente Schätzungen zu ermöglichen (vergleiche auch Molenberghs und Kenward (2007, Kapitel 10)).

Es stellt sich die Frage, wie die Konzepte von Modellselektion, Modellmittelung und fehlenden Daten zusammenpassen und miteinander kombiniert werden können.

## Übergreifende Ansätze

Derzeit existieren im Wesentlichen vier wissenschaftliche Arbeiten, die die Auswirkungen fehlender Werte auf die Modellselektion untersuchen, nämlich die Artikel von Shimodaira (1994), Cavanaugh und Shumway (1998), Hens, Aerts und Molenberghs (2006) und Claeskens und Consentino (2008). Alle angeführten Arbeiten unterstreichen, dass die Verwendung gängiger Verfahren der Modellselektion für den Fall fehlender Daten zur Wahl unpassender Modelle führen kann und schlagen allesamt eine Modifizierung von Akaikes Informationskriterium (AIC, Akaike (1973)) vor. Der in dieser Arbeit zentrale Artikel von Hens, Aerts und Molenberghs (2006) greift dabei die in vielen Teilgebieten

der Statistik bekannte und bereits angedeutete Idee des *inverse probability weighting* (Horvitz und Thompson (1952)) auf, bei der nur die vollständigen Beobachtungen eines Datensatzes verwendet werden – gewichtet mit ihrer inversen geschätzten Auswahlwahrscheinlichkeit. Die Autoren schlagen darauf aufbauend ein adjustiertes Selektionskriterium ( $AIC_W$ ) vor und konstatieren diesem in einer Auswahl an Simulationsstudien ein durchweg gutes Verhalten im Kontext fehlender Werte.

Es existiert bis dato jedoch keine einzige Veröffentlichung, die die Modellselektionsunsicherheit und die Problematik fehlender Werte gemeinsam diskutiert. Es stellt sich die Frage, ob und wie sich die Konzepte von Modellmittelung und fehlenden Daten in Einklang bringen lassen und inwiefern sich eine solche Kombination auch praktisch umsetzen lässt. Die vorliegende Arbeit zeigt die Vielfältigkeit beider Themengebiete auf und versucht sowohl für unterschiedlichste Ansätze aus dem Bereich fehlender Werte als auch aus dem Bereich der Modellmittelung einige repräsentative Ideen aufzugreifen und entsprechend zu erweitern.

Hierfür werden exemplarisch für ein mögliches Vorgehen drei konträre Ansätze erläutert: Es werden verschiedene Modellselektions- und Modellmittelungsschätzer unterschiedlicher Konzeption betrachtet, so etwa die oben angeführten Schätzer von Buckland, Burnham und Anderson (1997) bzw. Hansen (2007), wobei für deren Berechnung

- (i) die fehlenden Werte verworfen und nur die vollständigen Fälle eines Datensatzes verwendet werden,
- (ii) die fehlenden Werte durch einfache oder multiple Imputationen ersetzt werden und die Inferenz auf den entsprechend aufgefüllten Datensätzen durchgeführt wird und
- (iii) das für fehlende Werte adjustierte Kriterium von Hens, Aerts und Molenberghs (2006) zur Konstruktion von Gewichten für einen neuen, kriteriumsasierten Modellmittelungsschätzer verwendet wird.

Für den zweiten der genannten Punkte werden dabei drei unterschiedliche Imputationsmethoden betrachtet. Zum einen die  $k$ -Nächste-Nachbarn-Methode (Chen und Shao (2000), Gottardo (2008)), bei der jeder fehlende Wert durch das arithmetische Mittel seiner  $k$  nächsten, vollständig beobachteten, Nachbarn ersetzt wird; ferner wird die Idee der Regressionsimputation (Little und Rubin (2002)) aufgegriffen, bei der die fehlenden Werte durch Vorhersagen einer linearen Regression ersetzt werden. Dieses Konzept wird

verallgemeinert, indem zur Vorhersage flexiblere, generalisierte additive Modelle verwendet und in einen rekursiven Algorithmus integriert werden. Darüber hinaus werden einfache und multiple Imputationen über das Paket „Amelia II“ (Honaker, King und Blackwell (2008)) der statistischen Software *R* generiert, das einen schnellen und robusten Bootstrap-basierten Ansatz zur Modellierung der prädiktiven a-posteriori-Verteilung der fehlenden Daten gegeben die beobachteten Daten verwendet.

Es zeigt sich im Verlauf der Arbeit, dass das Verhalten der vorgestellten Modellselektions- und Modellmittlungsstrategien höchst unterschiedlich ist und sich für die entsprechenden Schätzer abhängig von der konkreten Behandlung fehlender Werte interessante und vielfältige Eigenschaften nachweisen lassen.

Die Ansatzpunkte und Konzepte im Bereich der Modellwahl und Modellmittelung sind außerordentlich vielfältig: Je nach Zweck, Ziel und Fragestellung ergeben sich unterschiedliche Ausgangspositionen, die zu Methoden, Kriterien und Verfahren führen, die einerseits unterschiedlicher nicht sein könnten, andererseits am Ende dann doch zu ähnlichen oder sogar identischen Resultaten führen können. Die vorliegende Arbeit versucht, im Gegensatz zu vielen anderen Veröffentlichungen, die Kernkonzepte zur Wahl und Kombination von Modellen relativ weitläufig zu diskutieren, um so die Chancen und Risiken für Modellselektions- und Modellmittelungsschätzer im Kontext fehlender Daten unter den verschiedensten Gesichtspunkten für viele Themenbereiche zu öffnen. Dabei sollen ausgewählte Verfahren und Ideen in ein einheitliches Spektrum eingeordnet und ihre Eigenheiten, Stärken und Schwächen herausgestellt werden.

Ziel der Arbeit ist es, interessante und repräsentative Verfahren der Modellselektion und Modellkombination in den Kontext fehlender Werte zu übertragen und die Eigenschaften und Erfolgsaussichten der vorgestellten Ideen mit Hilfe weitreichender Überlegungen, Monte-Carlo-Simulationen und Anwendungsbeispielen zu diskutieren, bewerten, analysieren und illustrieren.

## Ausblick

Kapitel 2 versucht zunächst zu klären, was ein Modell überhaupt ist, was man darunter generell in der Wissenschaft verstehen kann, inwiefern sich dies im Rahmen statistischer Modellierung wiederfindet, welche Ziele mit Modellselektion verfolgt werden und auf welche Art und Weise die vielfältigen Konzepte zu rechtfertigen sind. Einen Überblick ausgewählter Modellwahlverfahren, insbesondere aus dem Anwendungsbereich linearer und

generalisierter linearer Regressionsmodelle, werden in Kapitel 3 vorgestellt. Wie Modelle kombiniert werden können und in welche Richtung aktuelle Entwicklungen führen, wird in Kapitel 4 diskutiert. Der Schwerpunkt liegt dabei in Anbetracht der darauffolgenden Analysen vor allem auf nicht-bayesianischen Konzepten. Kapitel 5 vermittelt grundlegende Ideen aus dem Bereich fehlender Daten und versucht repräsentative Methoden frequentistischer Modellselektion und Modellmittelung in diesem Kontext zu erweitern, ergänzen und verbessern. Die vorgeschlagenen Schätzer werden anschließend mit Hilfe ausführlicher Simulationsstudien (Kapitel 6) und mehrerer Anwendungsbeispiele (Kapitel 7) bewertet und illustriert. Schwerpunkt sind dabei das lineare und das logistische Regressionsmodell mit fehlenden Werten in den Kovariablen sowie mit Einschränkungen auch ausgewählte Aspekte der Faktorenanalyse. Kapitel 8 fasst die wichtigsten Resultate noch einmal zusammen und diskutiert die Chancen, Erfolgsaussichten und Probleme der betrachteten Verfahren unter Beachtung aller gewonnen Erkenntnisse. Ein Ausblick auf mögliche Erweiterungen der vorgestellten Strategien wird ebenfalls gegeben.



## 2. Modelle in Wissenschaft und Statistik

Das Ziel der Wissenschaft wird in der allgemeinen Wissenschaftstheorie mit Erkenntnisgewinn, dem Erwerb von neuem Wissen<sup>1</sup>, gleichgesetzt. Insbesondere will die Wissenschaft nicht nur Tatsachen feststellen, sondern auch Ursachen und Erklärungen für diese finden. Die Suche allgemeiner Strukturen und Beziehungen beschränkt sich dabei nicht nur auf wahrnehmbare, sondern auch auf nicht-wahrnehmbare Gegenstandsbereiche. Die Systematisierung dieser Bereiche, also die Reduzierung ihrer Vielfältigkeit auf einige wenige elementare Faktoren, ist dabei von größtem Interesse. Diese Form der Abstraktion wird in der Wissenschaft unter anderem durch die Konstruktion von *Modellen* erreicht.

Seien  $G_1$  und  $G_2$  Gegenstände<sup>2</sup>,  $S_1$ ,  $S_2$  Sätze<sup>3</sup> und bedeute  $M(A, B)$ , dass  $A$  Modell für  $B$  ist; dann lassen sich folgende Formen von Modellen unterscheiden (vgl. auch Detel (2007, Seite 94)):

- (i)  $M(G_1, G_2)$   $G_1$  ist ein *strukturelles Modell* für  $G_2$ ,
- (ii)  $M(S_1, G_1)$   $S_1$  ist ein *abstraktes bzw. idealisiertes Modell* für  $G_1$ ,
- (iii)  $M(G_1, S_1)$   $G_1$  ist ein *semantisches Modell* für  $S_1$ ,
- (iv)  $M(S_1, S_2)$   $S_1$  ist ein *theoretisches Modell* für  $S_2$ .

Im Falle struktureller Modelle sind also Gegenstände Modelle anderer Gegenstände; so zum Beispiel das maßstabsgetreue Modell einer Brücke oder das Doppelhelix-Modell der DNA. Bei abstrakten und idealisierten Modellen werden hingegen Sätze als Modelle für Gegenstände verwendet. Sie versuchen die Komplexität von Phänomenen einzuschränken, nur ihre Kernelemente zu erfassen und sie unter dem Ziel der Verständlichkeit

---

<sup>1</sup> Die traditionelle Definition nach Plato bezeichnet „Wissen“ als *wahre, gerechtfertigte Meinung*. Auch wenn diese Definition nicht unumstritten ist (vgl. Gettier (1963)), so ist sie im Kontext von Modellbildung und Modellselektion völlig ausreichend.

<sup>2</sup> In der Philosophie bezeichnet ein Gegenstand eine Sache oder eine Entität, die Eigenschaften besitzen und Beziehungen zu anderen Gegenständen haben kann.

<sup>3</sup> Ein Satz ist eine im Sinne der Logik widerspruchsfreie Aussage (z.B. über einen oder mehrere Gegenstände), die mittels eines Beweises, das heißt aus Axiomen und bereits vorhandenen Sätzen, hergeleitet werden kann.

vereinfachend darzustellen. Damit bilden sie das Rückgrat aller empirischen Wissenschaften. Ein Beispiel dafür ist etwa das Modell idealer Gase, mit dem sich unter gewöhnlichen Bedingungen viele thermodynamische Prozesse von Gasen verstehen und beschreiben lassen, das für tiefe Temperaturen und hohen Druck jedoch keine adäquate Modellierung mehr bietet; auch die vereinfachende Annahme unabhängiger und identisch verteilter Beobachtungen einer Stichprobe, um Verfahren der statistischen Inferenz besser verwenden zu können, kann als idealisiertes Modell verstanden werden; oder das Gesetz von Henry Darcy, das die Wasserströmung in porösen Flüssigkeiten modelliert und ursprünglich durch Versuche in einem Sandbett entstanden ist, heutzutage jedoch vor allem als spezielle Lösung der Navier-Stokes-Gleichungen motiviert werden kann. Bei semantischen Modellen sind es Gegenstände, die Sätze wahr machen; beispielsweise die rationalen Zahlen  $\mathbb{Q}$  ohne Null, mit der Multiplikation als Verknüpfung und der Eins als neutralem Element, für die abelschen Gruppen im Bereich der Gruppentheorie der Mathematik. Sind Sätze Modelle für andere Sätze, so spricht man von theoretischen Modellen - insbesondere dann, wenn wissenschaftliche Theorien vorteilhaft für andere Theorien verwendet werden können. In diesem Sinne ist die klassische Mechanik ein theoretisches Modell für die Quantenmechanik.

Auch in der Statistik spielen Modelle eine herausragende Rolle. Das Sammeln, das Aufbereiten, die Analyse und die Interpretation von Daten eröffnet eine Fülle an Möglichkeiten statistischer Modellierung. Als weitgehend empirische Wissenschaft sind es insbesondere abstrakte und idealisierte Modelle, denen eine große Bedeutung zuzuschreiben ist. Aufgrund von Beobachtungen in der Form von Daten, sollen statistische Modelle Phänomene abstrahieren und beschreiben. Ein statistisches Modell in seiner allgemeinsten Form bezeichnet dabei eine parametrisierte Familie von Wahrscheinlichkeitsverteilungen

$$\mathcal{F} = \{f(y; \theta), \theta \in \Theta\}, \quad (2.1)$$

bei der die Beziehung einzelner Elemente eines Phänomens anhand einer Parametrisierung der Dichte von  $y$  durch  $\theta$  in  $\mathcal{F}$  beschrieben wird. Der Parameterraum  $\Theta$  muss dabei nicht endlich-dimensional sein. Ein Beispiel ist die Menge aller Normalverteilungen

$$\mathcal{F} = \left\{ \frac{1}{\sqrt{2\pi}\sigma} \exp \left( -\frac{1}{2} \left[ \frac{y - \mu}{\sigma} \right]^2 \right); \theta = (\mu, \sigma) \in \mathbb{R}^p \times (0, \infty) \right\},$$

wobei  $\mu$  den Erwartungswert und  $\sigma^2$  die Varianz von  $y$  beschreibt. Häufig meint ein statistisches Modell auch die funktionale Beziehung zwischen einer Zielgröße  $y$  und potentiellen Einflussgrößen  $X_1, \dots, X_p$ ,



$$y = f(X_1, \dots, X_p; \theta) + \epsilon, \quad (2.2)$$

wobei  $f(\cdot)$  eine noch unbestimmte Funktion und  $\epsilon$  eine Zufallskomponente bezeichnet. Ein typisches Beispiel ist das lineare Regressionsmodell

$$y = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p + \epsilon, \quad \epsilon \sim N(0, \sigma^2 I),$$

bei dem die Zielgröße  $y$  über eine Linearkombination der  $X_1, \dots, X_p$  modelliert wird. In der Regel betrachtet man zur Modellierung dieser Abhängigkeitsbeziehung die bedingte, parametrisierte Dichte  $f(y|X_1, \dots, X_p; \theta)$ , weswegen (2.2) als Spezialfall von (2.1) aufgefasst werden kann.

Ziel der statistischen Modellselektion ist es, aus einer Menge von Kandidatenmodellen  $\mathcal{M} = \{M_1, \dots, M_k\} \subset \mathcal{F}$  ein Modell  $M_{\kappa^*}$  auszuwählen, das die Daten – unter noch zu definierenden Gesichtspunkten – gut beschreibt. Dies betrifft insbesondere die Wahl geeigneter Regressoren bei Regressionsmodellen, die Anzahl von Kontrollpunkten in der Kontur- und Bildanalyse, die Ordnung autoregressiver Prozesse, die Anzahl von Faktoren in der Faktorenanalyse, die Wahl eines geeigneten Kerndichteschätzers und andere Problemstellungen der Statistik und verwandter Gebiete, vergleiche auch Linhart und Zucchini (1986) sowie Rao und Wu (2001).

## Sparsamkeit

Um Phänomene beschreiben und verstehen zu können, sollte ein gewähltes Modell einerseits ein möglichst genaues Abbild der Realität liefern, andererseits nur die Komplexität in Anspruch nehmen, die nötig ist, um die wichtigsten Kausalitäten und Merkmale der Daten abzubilden. Da zudem insbesondere im statistischen Kontext die Varianz, etwa der Parameterschätzung von  $\theta$ , mit der Komplexität steigt, der Bias in der Regel dagegen fällt, stellt sich die Frage nach einem geeigneten Kompromiss. Ist ein statistisches Modell zu komplex, so nennt man es überangepasst, ist es zu simpel, so nennt man es unterangepasst. In der Literatur wird häufig die Auffassung vertreten, dass bei gleicher Erklärungskraft das weniger komplexe Modell gewählt werden sollte. Dieses Prinzip ist

auch als *Prinzip der Sparsamkeit* bzw. als *Occam's Razor*<sup>4</sup> bekannt. Die Wahl eines solchen Modells ist jedoch keineswegs trivial, Burnham und Anderson (2002) bemerken hierzu:

*„Parsimony lies between the evil of under- and overfitting“*

und Forster (1998) fragt:

*„How much better must the complex model fit before we say that the extra parameter is necessary? Or, when should the better fit of complex models be ‘explained away’ as arising from the greater tendency of complex models to fit noise? How do we trade off fit with simplicity?“*

Um sich in der Statistik, auch unter Beachtung der Sparsamkeit, zwischen mehreren konkurrierenden Modellen für ein bestes entscheiden zu können, sind Verfahren und Kriterien notwendig. Statistische Modellselektion umfasst dabei häufig

- (i) eine risikobasierte Entscheidung durch Optimierung eines Selektionskriteriums,
- (ii) das sukzessive Testen von Hypothesen,
- (iii) oder ein ad-hoc Vorgehen.

Punkt (i) beinhaltet insbesondere Selektionskriterien auf Basis von Vorhersagefehlern, im Rahmen der Informationstheorie und bayesianischer Natur. Ausgewählte Verfahren und Methoden zu diesen drei, wie auch einigen anderen Punkten sollen in den Abschnitten 3.1–3.5 vorgestellt und motiviert werden. Dies geschieht, wenn möglich, in allgemeinsten Form; stets jedoch im Hinblick auf die Wahl geeigneter Regressoren in linearen und generalisierten linearen Regressionsmodellen.

---

<sup>4</sup> William Ockham (1285–1349) formulierte als erster ein Prinzip der Sparsamkeit (häufig wiedergegeben als *„entia non sunt multiplicanda praeter necessitatem“*), das weit über die Statistik hinaus in Biologie, Medizin und Philosophie bekannt ist. Heutzutage existieren im Detail viele verschiedene Fassungen und Versionen dieses Prinzips; prinzipiell lässt es sich jedoch so verstehen, dass für zwei wissenschaftliche Theorien bzw. Erklärungen unter festen Bedingungen diejenige zu bevorzugen ist, die einfacher ist. Die Rechtfertigung dieses Prinzips wird in der Wissenschaftstheorie diskutiert, vergleiche etwa Sober (1981) und Forster und Sober (1994). Die Begründungen sind dabei sowohl pragmatisch motiviert, wie etwa der oben angedeutete Punkt, dass Phänomene auf diese Weise besser verstanden werden können, als auch wissenschaftstheoretisch; so das Argument, dass das Ziel der Wissenschaft in der Approximation der Wahrheit besteht (vgl. auch die untenstehende Diskussion) und dieses Ziel nicht ohne die Berücksichtigung von Sparsamkeit erreicht werden kann. Kritische Stimmen sprechen dem Prinzip keine allgemeine Gültigkeit zu. So meint selbst Sober (2002): *„It may turn out, that simplicity has no global justification – that its justification varies from problem to problem“*.

Die meisten der vorgestellten Methoden beinhalten dabei das Sparsamkeitsprinzip mehr oder weniger explizit. Es stellen sich in diesem Zusammenhang jedoch die grundlegenden Fragen: Was ist Sparsamkeit? Wie lässt sich Sparsamkeit messen und konstruieren? Ist Sparsamkeit eindeutig? Diesen Fragen entspringt ein natürlicher Diskurs, ob Sparsamkeit Teil eines empirischen Vorgehens sein kann oder ob es ein künstliches, insofern rationalistisches, Konzept ist. Prinzipiell setzt Empirismus voraus, dass jede Form von Wissen über Erfahrung, beispielsweise über Daten, gewonnen wird, während unter einer rationalistischen Denkweise Erkenntnis in erster Linie unabhängig von unseren Sinnesindrücken entsteht. Deswegen wird in der Literatur teilweise argumentiert, dass das Sparsamkeitsprinzip ein extraempirisches Element ist, das primär in der Statistik, aus pragmatischen Gründen, herangezogen wird und somit als rationalistisch angesehen werden muss; so schreiben Forster und Sober (1994):

*„Giving weight to simplicity thus seems to embody a kind of rationalism“*

Die folgenden Kapitel werden zeigen, dass dieses Argument nicht korrekt ist. Obgleich Bestandteil fast jedes statistischen Verfahrens bzw. Kriteriums, kann die Interpretation der Aufsplittung in einen Anpassungs- und einen Sparsamkeitsterm als Kompromiss zwischen Bias und Varianz in der Regel nur a posteriori erfolgen; a priori liegt den Methoden meist ein grundlegend anderes Prinzip zugrunde, etwa die Approximation von Wahrheit, die Verringerung von Vorhersagefehlern oder die Maximierung von posteriori-Wahrscheinlichkeiten – die Sparsamkeitsterme entstehen dabei gewissermaßen als „Abfallprodukt“ bei der Evaluierung der eigentlichen Zielsetzung. Es zeigt sich daher, dass die Konzepte statistischer Modellselektion fast ausschließlich datengestützt und empirisch motiviert sind und keine rationalistische Rechtfertigung benötigen.

Eine Ausnahme bildet die von Jorma Rissanen begründete Theorie der Minimum Description Length. Hierbei bildet der Kompromiss zwischen Anpassung und Sparsamkeit das Fundament aller von Rissanen (1978) erarbeiteten Verfahren. Konzepte aus der Informationstheorie helfen dabei, die Länge wissenschaftlicher Theorien, im Speziellen statistischer Modelle, zu beschreiben und dadurch Modelle zu wählen, die Wissen am besten „verschlüsseln“ können und dennoch anschließend dieses Wissen am besten zu reproduzieren vermögen. Dadurch wird nicht nur das Prinzip der Sparsamkeit direkt bei der Konstruktion von Modellwahlkriterien verwendet, sondern auch erstmals eine konkrete Ausarbeitung davon präsentiert, wie Sparsamkeit zu messen ist. Eine überschaubare Einführung bietet Abschnitt 3.5.1; die dort angegebenen Referenzen erlauben darüber hinaus einen tieferen Einblick in die Thematik, die in den folgenden Kapiteln keine zentrale Rolle einnehmen wird.

## Wahrheit

Ein Vergleich und eine Beurteilung der einzelnen Verfahren, insbesondere der Selektionskriterien, erfolgt meist durch die Betrachtung asymptotischer Güteigenschaften, wie der Konsistenz und der Effizienz. Abschnitt 3.6 beschäftigt sich hiermit ausführlich. Es zeigt sich, dass eine der entscheidenden Voraussetzungen zur Optimalität eines Modellwahlkriteriums in der Tatsache liegt, ob in der Menge der Kandidatenmodelle  $\mathcal{M} = \{M_1, \dots, M_k\}$  das wahre, datengenerierende Modell  $M_\kappa^*$  enthalten ist oder nicht. Es zeigt sich ferner, dass das Konzept eines wahren Modells als solches problematisch ist. Obgleich Voraussetzung in der Konstruktion vieler populärer Modellwahlkriterien, ist seine Existenz nicht unumstritten. Burnham und Anderson (2002) bemerken

*„The words ‘true model’ represent an oxymoron“*

und de Leuw (1988) meint in diesem Zusammenhang lapidar

*„Truth is elusive“*

Die Diskussion dieses Aspekts verlagert sich in der statistischen Literatur jedoch meist in Richtung der nahezu gleichwertigen Frage der Dimensionalität eines wahren Modells. Ist es von unendlicher Dimension, so ist es implizit nicht in der Menge der Kandidatenmodelle enthalten; ist es von endlicher Dimension, so kann es durchaus Bestandteil dieser Menge sein. Abschnitt 3.6 präsentiert wichtige Resultate zur Optimalität von Modellwahlkriterien und diskutiert ihre Nützlichkeit anhand ausgewählter Aspekte.

In gewisser Weise ist diese Diskussion auch Bestandteil einer alten Realismus-Debatte: Die Annahme einer denkunabhängigen Realität, einer Realität die sich erkennen und erfassen lässt und damit auch letztlich zu Wissen unabhängig von menschlichen Theorien und Konventionen führt, kann als Position für einen erkenntnistheoretischen, wissenschaftlichen Realismus verstanden werden.<sup>5</sup> Eine solche Sichtweise impliziert, dass eine Wirklichkeit, eine Wahrheit existiert und wir diese erfahren und beschreiben können und dass die Annäherung an diese Wahrheit insofern auch Ziel der Wissenschaft ist. Im Gegensatz dazu existieren viele nicht-realistische Positionen, die sich in philosophischen Denkweisen, wie etwa dem Relativismus, dem Instrumentalismus oder dem konstruktiven Empirismus äußern. Letzterer geht auf van Fraassen (1980) zurück und verneint das

<sup>5</sup> Tatsächlich umfasst der Begriff des Realismus eine Vielzahl philosophischer Positionen, die sich auf unterschiedliche Gegenstandsbereiche beziehen. Diese sind im Kontext dieser Arbeit jedoch nicht weiter relevant; der Kerngedanke und das Stichwort einer „denkunabhängigen Realität“ genügt für die folgende Diskussion.

Ziel der Wissenschaft als Approximation von Wahrheit. Die Doktrin eines konstruktiven Empirismus sieht vor, die Wissenschaft als reinen Beobachter zu betrachten, der wahre Aussagen über beobachtbare Phänomene und Experimente machen kann, ausdrücklich aber nicht über unbeobachtbare Phänomene und damit über eine den Beobachtungen zugrundeliegende Wahrheit. Es ist fragwürdig, ob solche oder andere (beispielsweise relativistische) Standpunkt hilfreich sind. Sober (2000) quittiert die Diskussion mit den Worten:

*„Realism says that the goal of science is to discover which theories are true; [constructive] empiricism maintains that the goal is to discover theories that are empirically adequate [...] In both cases, truth is the property that matters“*

Tatsächlich ist die Realismus-Debatte im Kontext statistischer Modellselektion nicht entscheidend. Auch wenn aufgrund ihrer Konzeption viele der in Kapitel 3 vorgestellten Methoden eine zugrundeliegende Wahrheit, zumindest in Form eines datengenerierenden Prozesses, benötigen und damit auch eine prinzipiell realistische Sichtweise angenommen wird, so steht vor allem die Identifizierung relevanter Effekte eines Phänomens im Vordergrund. Ob Aussagen über eine Wahrheit und ihre Existenz getroffen werden müssen, ist fraglich. Dies macht auch Abschnitt 3.6 deutlich: Ob die oben erwähnten Qualitätsmerkmale von Konsistenz und Effizienz im Kontext statistischer Modellwahl sinnvoll sind, wird dort diskutiert.

## Grenzen des Wissens

Die Suche geeigneter Modelle zur Charakterisierung von Phänomenen unterliegt häufig gewissen Beschränkungen, insbesondere solchen, die sich aus den Grenzen empirischen Wissens ergeben. Dies betrifft vor allem die mangelnde Verfügbarkeit und das Fehlen von Daten: Die Herausforderung, Modelle zu bilden, zu wählen und zu schätzen, auch unter Beachtung der oben diskutierten Sachverhalte der Sparsamkeit und Optimalität, erfährt im Kontext fehlender Werte eine zusätzliche Dimension. Die Diskussion und Evaluierung von statistischen Methoden zur Modellselektion unter Berücksichtigung dieser Problematik ist Schwerpunkt dieser Arbeit und wird weitgehend in den Kapiteln 5-7 erörtert. Zusätzlich berücksichtigt werden dabei auch die Grenzen statistischer Modellselektionsverfahren. Die Unsicherheit bezüglich der Wahl eines geeigneten Modells ist ein weiterer Faktor, dem besondere Beachtung geschenkt wird. Die Kombination verschiedener Konzepte aus verschiedenen Teilgebieten der Statistik sollen helfen, ein relativ allgemein angelegtes Sammelsurium an Methoden zur Bewältigung dieser Probleme zu liefern. Die Illustration dieser Methoden beschränkt sich aus Gründen der Übersichtlichkeit dabei auf