

Martin Stark

**Modellgestützte Kostenprognose
für den Aufbau qualitätsgesicherter Geodatenbestände**

München 2005

**Verlag der Bayerischen Akademie der Wissenschaften
in Kommission beim Verlag C. H. Beck**

Modellgestützte Kostenprognose
für den Aufbau qualitätsgesicherter Geodatenbestände

Von der Fakultät für Bau- und Umweltwissenschaften
der Universität Stuttgart
zur Erlangung der Würde eines Doktor-Ingenieurs (Dr.-Ing.)
genehmigte Abhandlung

vorgelegt von

Dipl.-Ing. Martin Stark

Stuttgart, D93

München 2005

Verlag der Bayerischen Akademie der Wissenschaften
in Kommission beim Verlag C. H. Beck

Adresse der Deutschen Geodätischen Kommission:

Deutsche Geodätische Kommission

Marstallplatz 8 • D – 80 539 München

Telefon +49 - (0)89 - 23 031 -0 / -1113 • Telefax +49 - (0)89 - 23 031 -1283 / -1100

E-mail hornik@dgfi.badw.de • <http://dgk.badw.de>

Hauptberichter: Prof. Dr.-Ing. W. Möhlenbrink

Mitberichter: Prof. Dr.-Ing. M. Friedrich

Tag der Einreichung: 23.07.2003

Tag der mündlichen Prüfung: 03.11.2004

© 2005 Deutsche Geodätische Kommission, München

Alle Rechte vorbehalten. Ohne Genehmigung der Herausgeber ist es auch nicht gestattet,
die Veröffentlichung oder Teile daraus auf photomechanischem Wege (Photokopie, Mikrokopie) zu vervielfältigen

Inhalt

Inhalt	3
Zusammenfassung	5
Abstract	6
1 Einleitung	7
1.1 Motivation	7
1.2 Ziele der Arbeit	8
1.3 Vorgehensweise und Aufbau der Arbeit	8
2 Grundlagen und ökonomische Eigenschaften raumbezogener Daten	11
2.1 Grundlagen und Definitionen	11
2.1.1 Geodaten, Geobasis- und Geofachdaten	11
2.1.2 Geoinformationssysteme	12
2.1.3 Datenmodellierung	12
2.1.4 Signale, Daten, Nachricht und Informationen	13
2.1.5 Geodaten, Geoinformationen und Geoprodukte	14
2.2 Funktionalarchitektur von Informationsdienstleistungen mit Geodaten	15
2.3 Ökonomische Eigenschaften raumbezogener Daten	16
2.3.1 Anbieter und Nutzerstrukturen des Geodatenmarktes	16
2.3.2 Nutzen und wertbeeinflussende Eigenschaften	17
2.3.3 Kostenstrukturen amtlicher und privater Geodaten	18
3 Digitale Geodaten für die Intermodale Routensuche	19
3.1 Einleitung	19
3.2 Grundlagen einer intermodalen Routensuche	19
3.2.1 Informationsbedarf des Kunden	19
3.2.2 Betrachtete Umsteigebeziehungen	21
3.3 Aufbau der Geodatengrundlage	22
3.3.1 Vorgehensweise zur Realisierung der Geodatengrundlage	22
3.3.2 Nachbearbeitung der digitalen Straßenkarte	23
3.3.3 Erfassung des ÖV-Netzwerks	26
3.3.4 Modellierung der ÖPNV-Haltestellen	28
3.3.5 Georeferenzierung der Haltestellen und Linien	33
3.3.6 Überprüfung der Bearbeitung	36
3.4 Kostenanteile zur Realisierung des Intermodalen Routenplaners	37
3.5 Abschätzung des Kundennutzens	39
3.6 Reinvestition des Projektaufwands durch Nutzungsgebühren	42
3.6.1 Abschreibung der Investitionen	42
3.6.2 Lineare Abschreibung	43
3.6.3 Degressive Abschreibung	44
3.7 Abschätzung erzielbarer Erlöse durch Nutzungsentgelt	46
3.8 Berechnung des Return On Investment für die Nutzungsszenarien	49
3.9 Beurteilung des Ergebnisses	50
3.10 Zusammenfassende Bemerkung	51

4	Ein Konzept zur Ermittlung der Bereitstellungskosten von Geodaten	53
4.1	Einführung.....	53
4.2	Prozess der allgemeinen Geodatenbereitstellung	54
4.3	Der Kostenrahmen zur Bereitstellung von Geodaten	55
4.3.1	Grundlagen von Kostenmodellen.....	55
4.3.2	Einflussgrößen und deren Zusammenhang in Kostenmodellen	56
4.3.3	Kategorisierung kostenrelevanter Merkmale.....	56
4.3.4	Arten von Kostenmodellen.....	58
4.3.5	Prozessspezifisches Kostenmodell der Geodatenbereitstellung	59
4.4	Aufbaumöglichkeiten von Geodatenbeständen	60
4.4.1	Prinzipielle Aufbaumöglichkeiten.....	60
4.4.2	Entscheidungsdiagramm zur Herleitung der Aufbaumöglichkeiten.....	61
4.4.3	Varianten des Neuaufbaus von Geodatenbeständen.....	62
4.4.4	Varianten der Veredelung eigener Bestände	63
4.5	Die Kostenstufen der Aufbaumöglichkeiten	65
4.6	Das Berechnungsverfahren zur Kostenprognose.....	69
4.6.1	Grundlagen der HOAI und AHO	69
4.6.2	Die Kostenfunktionen und Ermittlungsmethodiken	71
4.7	Validierung des Prognosemodells	77
4.7.1	Datenbankbasierte Modellumsetzung	77
4.7.2	Kalibration und Validierung durch reale Kostenanalyse.....	78
4.8	Bemerkungen zum Ermittlungskonzept der Bereitstellungskosten von Geodaten.....	79
5	Qualitätssicherung von Geodaten	81
5.1	Qualitätsmerkmale als Beurteilungskriterien des Geodatenaufbaus.....	81
5.1.1	Allgemeine Betrachtung verwendeter Qualitätsmerkmale	82
5.1.2	Qualitätsmerkmale als Entscheidungskriterien	83
5.1.3	Qualitätsforderung an den Zieldatenbestand	90
5.2	Fehlerrate in Abhängigkeit inhaltlicher Qualitätsmerkmale.....	92
5.3	Der Begriff der Qualitätssicherung	93
5.4	Möglichkeiten zur Qualitätssicherung von Geodatenbeständen.....	95
5.4.1	Geeignete Verteilfunktion bei der Stichprobenprüfung von Geodaten	96
5.4.2	Ergebnisbewertung der Stichprobenuntersuchung	97
5.5	Die Operationscharakteristik	98
5.5.1	Einfluss der Fehlerrate auf den Stichprobenumfang	99
5.5.2	Kostenoptimale Stichprobenprüfpläne	101
5.5.3	Stichprobenprüfkosten unter Vollprüfungsrisiko	105
5.6	Die Kostenfunktionen der Qualitätssicherungsmaßnahmen.....	108
5.6.1	Kostenfunktion der Fehlerprüfung K_{FP}	108
5.6.2	Kostenfunktion der Fehlerkorrektur	109
5.6.3	Kostenfunktionen der Aufbaumöglichkeiten.....	110
5.6.4	Vorgehensweise zur Anwendung dieser Methodik.....	114
5.7	Beispielhafte Kostenrechnung der Aufbaumöglichkeiten	116
5.7.1	Der Modelldatenbestand zur einheitlichen Kostenbewertung	116
5.7.2	Die Kostenverläufe der Aufbaumöglichkeiten	118
5.8	Ergebnisse der dargestellten Methodik des Kostenvergleichs der Aufbaustrategien	121
5.9	Nutzen für künftige Informationsdienste.....	122
6	Ergebnisbewertung und Ausblick	123
7	Literatur	125

Zusammenfassung

Für moderne Verkehrs- und Telematikdienstleistungen spielen digitale, raumbezogene Daten eine entscheidende Rolle. Erst durch die Schaffung einer digitalen Straßenkarte erhielt z. B. die komplexe Aufgabe der Zielführung von Landfahrzeugen unter Verwendung von Ortungs- und Navigationsverfahren den zur Markteinführung notwendigen Reifegrad. Diese Geodaten sind nicht nur unverzichtbare Grundlage solcher Anwendungen, sie verzehren bei ihrer Erfassung und Pflege auch die meisten Kosten.

Mit der Datenalterung auf Grund mangelnden Aktualisierungen und fehlerhaften oder unvollständigen Inhalten geht ein hohes Maß an Akzeptanzverlust einher. Datenfehler können zudem zu erheblichen Mehraufwendungen und Folgekosten führen.

Es liegt deshalb im unternehmenseigenen Interesse, die Investitions- und Folgekosten, welche durch den Einsatz von digitalen Geodaten entstehen, möglichst zuverlässig prognostizieren zu können. Diese Kostenprognose ist für die Unternehmensentwicklung in Form eines Geschäftsmodells von zentraler Bedeutung. In dieser Arbeit wird deshalb ein Konzept zur modellgestützten Kostenprognose für den Aufbau qualitätsgesicherter Geodatenbestände vorgestellt.

Am Beispiel des im Rahmen des Forschungsprojekts MOBILIST in der Region Stuttgart geschaffenen Intermodalen Routenplaners wird gezeigt, mit welchen Arbeitsschritten und deren zeitlichen Umfängen es gelungen ist, von bereits vorhandenen digitalen raumbezogenen Daten zu einer digitalen Geodatengrundlage zu gelangen, die es gestattet, eine verkehrszübergreifende Routenplanung durchzuführen. Hintergrund dieser Entwicklung ist das Bestreben nach einem besseren Zusammenwirken der Verkehrsträger Straße und Schiene, um einen wichtigen Beitrag zu Sicherung der individuellen Mobilität bei stetig steigendem Verkehrsaufkommen zu leisten.

Hierbei wird aufgezeigt, dass die Gesamtkosten zur Realisierung der Geodatengrundlage je nach Umfang der Qualitätssicherungsmaßnahmen stark variieren. Der Aufwand zur Qualitätssicherung durch Fehleridentifikation und -korrektur kann die Herstellkosten übersteigen.

Hieraus motiviert sich die kostenmäßige Betrachtung von qualitätssichernden Maßnahmen sowie deren Einfluss auf die durchzuführende Strategie des Geodatenaufbaus. Mit der Integration dieser Betrachtung in das Kostenprognosemodell wird das Ziel verfolgt, nicht nur die reinen Aufbauposten des Datenbestandes abschätzen zu können, sondern auch die richtige, kostengünstigere Fehleridentifikations- und -korrekturmaßnahme auszuwählen. Darüber hinaus werden Aussagen getroffen, von welchen Einflussfaktoren eine Auswahl unterschiedlicher Aufbaustrategien abhängig ist.

Es wird anhand von Modellrechnungen nachgewiesen, dass es wirtschaftlich sinnvoll sein kann, statt der Nutzung des eigenen Datenbestandes einen kompletten fremden Datenbestand einzukaufen und nachträglich weitere Informationen hinzuzufügen. Ferner wird aufgezeigt, ab welchen Qualitätsforderungen eine vollständige Prüfung des gesamten Datenbestandes einer Stichprobenprüfung vorzuziehen ist und welchen Einfluss dies auf die Gesamtkosten hat.

Mit diesem praxistauglichen, anhand eines realen Geodatenaufbaus kalibrierten Kostenprognosemodells existiert eine Entscheidungsmethodik für den Aufbau von Geodatenbeständen. Zudem sind künftig die Qualitätssicherungsmaßnahmen systematisch planbar. Es ist weiterhin detailliert ausgearbeitet, welche Arbeitsschritte zur Projektion dieser modellgestützten Kostenprognose auf andere Anforderungen zum Aufbau von Geodatenbeständen notwendig sind.

Abstract

Spatial geodatabases are basic condition for many modern traffic management and telematic services. By realising the digital road map and using modern positioning techniques the complex task of target guidance finally took up speed and meets the required maturity for market introduction.

These geodatabases are not only indispensable for those services but their realisation and ongoing maintenance work is highly expensive. Data aging by less maintenance, data errors and incomplete databases will cause an increasing loss of acceptance. This errors can also cause additional work and the rising of the resulting costs.

Therefore it must be the own interest of an enterprise to get an precise forecast of investment and the resulting costs. Such a forecast plays an important roll to set up a new business model. Regarding aspects of quality a prediction model for the build-up costs of geodatabases is presented in this work.

By regarding the intermodal routing planner which had been realised within the official German research project MOBILIST the working tasks and their schedule are shown as well as the resulting costs to built up that intermodal database. With this database the interconnection of the street and railway network had been improved in order to ensure future individual mobility with respect to the increasing flood of traffic. It is shown that the total costs of realisation depends on the tasks of quality management. Costs of quality management can be higher than the total costs of realisation.

This motivates the research about the influence of building-up strategies of geodatabases on quality management costs. By integrating the quality aspects in the prediction model it will be able to forecast not only the build-up costs but also identify the cheapest way of error detection and correction. This leads to the possibility to say which build-up strategy with respect to quality requirements is cheaper and better.

It is proven by model calculations that it can be reasonable to reject own geodatabases and follow the new build-up strategy to buy a new geodataset and improve it by yourself. Furthermore the threshold value of quality requirements for the decision which way of quality inspection (full inspection or inspection by random sample) is given.

With this prediction model of the total costs regarding quality requirements a method was developed to decide which strategy of building up a geodatabase will be useful. It is shown how to transfer this method into other fields of applications on geodatabases.

1 Einleitung

1.1 Motivation

Digitale Geodaten sind für viele existierende und zukünftige Verkehrsinformationsdienstleistungen notwendige Grundvoraussetzung für deren erfolgreichen Einsatz (vgl. Abbildung 1.1). So konnte die Fahrzeugnavigation und Zielführung erst mit dem Aufbau einer digitalen, routingfähigen Straßenkarte in den Massenmarkt eingeführt werden. Flankiert durch die beiden Technologiepfade der Mobilkommunikation und satellitenbasierten Ortung wurden in der Vergangenheit eine Vielzahl von Verkehrstelematikanwendungen (z.B. TeleAID von DaimlerChrysler) aufgebaut. Sowohl fahrzeug- als auch personenbezogene Telematikanwendungen (Electronic Toll Collection, Location based services) der Zukunft werden sich nur dann am Markt durchsetzen können, wenn erstens deren Funktionalität auf die Kundenbedürfnisse zugeschnitten ist und zweitens durch deren Nutzung ein Kundenmehrwert entsteht für den auch eine Zahlungsbereitschaft vorhanden ist.

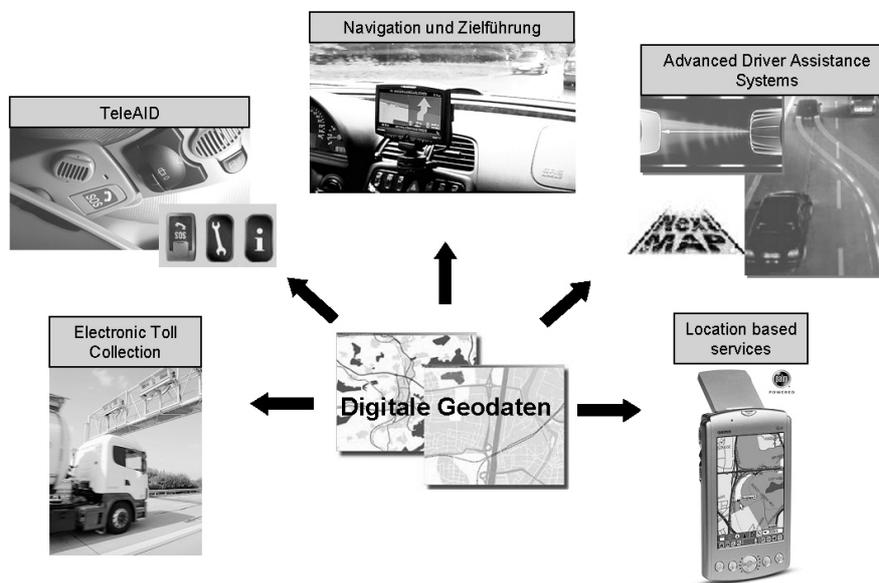


Abbildung 1.1:
Geodatenbedarf der
Informationsdienstleistungen

Zur Erfolgsabschätzung einer neuen Telematikanwendung gehört deshalb auch die Abschätzung der anfallenden Kosten für eine speziell dafür aufzubauende Geodatengrundlage. Die Richtigkeit und Genauigkeit sämtlicher Geschäftsmodelle dieser Informationsdienstleistungen hängt von dieser Kostenabschätzung ab. Eine solche Kostenschätzung geht einher mit der Art des Aufbaus der Geodatengrundlage (vgl. Abbildung 1.2).

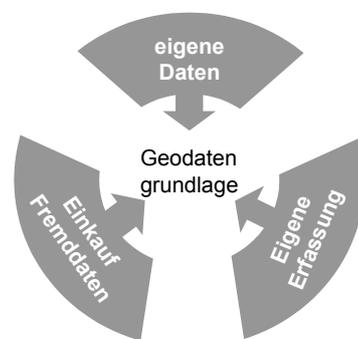


Abbildung 1.2:
Art des Geodatenaufbaus

Nach einer ausführlichen Anforderungsanalyse neuer Informationsdienstleistungen sowie einer notwendigen Betrachtung der augenblicklichen Situation in Form einer Bestandsaufnahme (Ist-Analyse) ergeben sich in der Regel mehrere Realisierungsmöglichkeiten des Geodatenbestandsaufbaus für eine Unternehmung. Dies trifft insbesondere dann zu, wenn entweder bereits umfangreiches eigenes Material zur Verfügung steht, vergleichbare Daten auf dem Markt einge-

kauf werden können, oder wenn die Anforderungsanalyse gezeigt hat, dass die Forderungen nicht durch den Einsatz eigener oder fremder Daten erfüllt werden können. Hier stellt sich den Unternehmen die Frage, ob eigene analoge Daten digitalisiert und mit fremden Daten fusioniert werden müssen, oder ob ein alleiniger Dateneinkauf sämtliche Nutzeranforderungen erfüllt. Im schlimmsten Fall wird, je nach Zieldefinition, die vollständige Neuerfassung der Daten notwendig.

Zur Beurteilung dieser unterschiedlichen Realisierungsmöglichkeiten sind neben den technischen und inhaltlichen Rahmenbedingungen auch die entsprechende Kostenbetrachtung relevant. Zwar werden in vielen Projekten bei Unternehmen Vollkostenrechnungen nachträglich durchgeführt, die den Umstellungs- bzw. Einführungsaufwand in Form eines Vergleichs der anfallenden Kosten zu rechtfertigen versuchen, jedoch eine vorausschauende Kostenschätzung basiert in den meisten Fällen auf Erfahrungswerten aus bereits durchgeführten Projekten. Oftmals können Kostenstrukturen nicht detailliert genug angegeben werden, um eine solche Kostenprognose mit den notwendigen Eingangsparametern zu versorgen.

1.2 Ziele der Arbeit

Das Ziel dieser Arbeit ist demzufolge die Erarbeitung einer Methodik zur modellbasierten Kostenschätzung, die sich für die Aufgabe des Aufbaus von Geodatenbeständen anwenden lässt. Insbesondere stehen hier die unterschiedlichen Möglichkeiten des Aufbaus im Vordergrund. Kernstück dieser Methodik ist ein dafür eigens aufgestelltes Kostenmodell. Die darin entwickelten Kostenfunktionen geben den funktionalen Zusammenhang zwischen Kostenparametern und Teilprozesskosten an. Diese Kostenfunktionen können je nach unterschiedlicher Möglichkeit des Geodatenaufbaus variiert werden. Sie stellen sozusagen einen transparenten Rahmen zur Kostenschätzung dar.

Neben der Erarbeitung eines Kostenmodells ist die transparente Darstellung der Vorgehensweise sowie die Bedeutung der universellen Einsetzbarkeit des Prognosemodells vorrangiges Ziel der Arbeit.

So wird ein Verfahren entwickelt, mit dessen Hilfe die Kostenverläufe der Aufbaumöglichkeiten in Abhängigkeit von Qualitätsforderungen angegeben werden können. Wesentlicher Einfluss auf diese Kostenverläufe haben Qualitätssicherungsmaßnahmen, die mit Hilfe der Stichprobenprüfung durchgeführt werden können. Es werden in dieser Arbeit die maßgeblich kostenverursachenden Faktoren des Geodatenaufbaus herausgearbeitet. Mit Hilfe eines Kostenmodells und darin enthaltenen Kostenfunktionen wird deren Einfluss auf die Gesamtkosten der Aufbaumöglichkeiten der Geodatenbestände nachgewiesen.

1.3 Vorgehensweise und Aufbau der Arbeit

Die Arbeit ist in insgesamt sechs Kapitel gegliedert. Neben dieser Einleitung und der Ergebnisbewertung in Kapitel 6 sind in **Kapitel 2** die theoretischen Grundlagen raumbezogener Daten erläutert, soweit sie das hier behandelte Thema berühren. Es werden in einer kurzen Übersicht die Fachbegriffe der Geodaten und Geofachdaten, Geoinformationssysteme und die Datenmodellierung sowie eine Unterscheidung zwischen Geoinformationen und Geodatenprodukten erläutert.

Neben diesen Grundlagen sind die ökonomischen Eigenschaften und wirtschaftliche Relevanz von Geodaten dargestellt. Zentraler Gegenstand der Betrachtung ist die Wertschöpfungskette bei der Herstellung von Geodatenprodukten. Die Geobasisdaten stehen am Anfang dieser Wertschöpfungskette. Deren Erfassung und Pflege ist mit sehr hohen Kosten bei gleichzeitig nur geringem Marktwert, bzw. Nutzwert für den Kunden verbunden, während die eigentliche Wertschöpfung erst durch die anforderungsspezifische Veredelung erfolgt und erst damit ein höherer Nutzwert entsteht. Durch die Darstellung der Anbieter- und Nutzerstrukturen auf dem Geodatenmarkt, deren wertbeeinflussende Faktoren sowie dem großen öffentlichen Interesse an Geodaten soll die Forderung nach einer nutzerorientierten Tarifierungsstrategie amtlicher Geobasisdatenproduzenten motiviert werden. Am Beispiel einer allgemein gültigen Funktionalarchitektur von Informationsdienstleistungen wird schließlich die Notwendigkeit einer prozessbegeleitenden Kostenbetrachtung motiviert.

Das **Kapitel 3** stellt den Realisierungsaufwand einer Geodatenbasis vor, welche zum Zwecke der intermodalen, also verkehrsträgerübergreifenden, Routenberechnung im Rahmen des BMBF-Forschungsprojekts MOBILIST für den Ballungsraum Stuttgart entstanden ist. Darin sind sämtliche Arbeitsschritte der Geodatenveredelung detailliert beschrieben und werden aufwandsbezogen bewertet und die Kostenstruktur herausgearbeitet. Auf der Grundlage einer Nutzungsabschätzung und mehrerer Entwicklungsszenarien sind verschiedene Zeitpunkte des Wiedergewinns der Investition (Return On Invest) abgeleitet. Aufgrund dieser Modellrechnung konnte die kritische Masse entgeltpflichtiger Nutzer eines solchen Informationsauskunftssystems ermittelt werden. Dieses Kapitel gibt ein Beispiel für die verbleibende Notwendigkeit der Integration und Homogenisierung von Geodaten unterschiedlicher Quellen und relativiert damit gewissermaßen die im Begriff der Interoperabilität implizit vorhandene pauschale Verzichtserklärung auf Homogenisierung von Geodaten.

Aufbauend auf diesem Beispiel wird in **Kapitel 4** ein Konzept zur Ermittlung der Bereitstellungskosten von Geodaten vorgestellt. Es basiert auf der Betrachtung des Prozesses der allgemeinen Geodatenbereitstellung, einer Analyse in Frage kommender Kostenmodelle sowie einem speziell für die Aufgabe der Geodatenbereitstellung aufgestellten Kostenmodell. Darin sind die funktionalen Abhängigkeiten zwischen den Kostenparametern in Form von Funktionen angegeben, die eine transparente Anwendung dieses Modells gestatten. Das Kernstück dieses Modells ist ein daraus abgeleitetes Berechnungsverfahren zur á priori Kostenschätzung. Unter besonderer Berücksichtigung von Qualitätsmerkmalen als Bewertungskriterien steht ein umfassendes Werkzeug zur Verfügung, die skizzierte Funktionalarchitektur einer Informationsdienstleistung in Form eines begleitenden Kostenmodells berechnen und bewerten zu können.

In **Kapitel 5** richtet sich die Kostenbetrachtung auf Qualitätssicherungsmaßnahmen bei der Geodatenproduktion. Insbesondere steht dabei die Stichprobenprüfung im Vordergrund. Es werden speziell angepasste Kostenfunktionen hergeleitet, mit denen es möglich ist, unter Variation der Eingangsparameter eine Aussage über auftretende Konkurrenzsituationen zweier verschiedener Aufbaumöglichkeiten von Geodatenbeständen zu treffen. So sind diejenigen Einflussparameter herausgearbeitet, deren Variation die größte Kostenveränderung nach sich ziehen. Damit können entsprechend einer existierenden Qualitätsforderung, die durch das Stichprobenverfahren garantiert werden soll, die Prüf- und Qualitätssicherungskosten vorhergesagt werden und so zu einem kostengünstigeren Geodatenaufbau beitragen.

2 Grundlagen und ökonomische Eigenschaften raumbezogener Daten

Raumbezogene Daten finden in vielen Einsatzgebieten ihre Anwendung. Die damit verbundenen Geoinformationssysteme (GIS) werden zur Lösung raumbezogener Aufgaben in zahlreichen Bereichen der Verwaltung, Wirtschaft und Industrie eingesetzt, um so beispielsweise bei Standortentscheidungen für Produktionsstätten oder bei Absatzmarktanalysen die Untersuchungen regional schärfer einzugrenzen und zuverlässigere Ergebnisse zu erhalten [BERNHARD, 2002].

2.1 Grundlagen und Definitionen

In diesem Abschnitt werden einige Begriffe aus dem Bereich der Verarbeitung raumbezogener Daten kurz definiert, bzw. erläutert werden. Auf eine ausführlichere Darstellung der Begriffe wurde bewusst verzichtet, da sie erstens innerhalb der Fachdisziplinen der Geoinformationsverarbeitung hinlänglich bekannt sind und zweitens eine ausführlicherer Darstellung für fachfremde Disziplinen hier keinen nennenswerten Informationsgewinn bezüglich des Verständnisses dieser Arbeit zur Folge hätte. Die weiteren Ausführungen beschränken sich im wesentlichen auf die nachfolgend aufgeführten Definitionen.

2.1.1 Geodaten, Geobasis- und Geofachdaten

Als Geodaten werden jegliche raumrelevanten Informationen mit einem geographischen Bezug bezeichnet. Geodaten lassen sich in die beiden Klassen der Referenzdaten, oder auch Basisdaten und die thematischen Daten unterscheiden. Eine umfassende, grundlegende Definition der Geodaten gibt [BILL, 1999] mit:

Definition 2-1:

Geodaten sind Daten über Gegenstände, Geländeformen und Infrastrukturen an der Erdoberfläche, wobei als wesentliches Element ein Raumbezug vorliegen muss. Sie beschreiben die einzelnen Objekte der Landschaft. Geodaten lassen sich über den Raumbezug miteinander verknüpfen, woraus insbesondere unter Nutzung von GIS-Funktionalitäten wiederum neue Informationen abgeleitet werden können. Auf und mit Ihnen lassen sich Abfragen, Analysen und Auswertungen für bestimmte Fragestellungen durchführen. Geodaten lassen sich in zwei große Teilkomplexe aufteilen, nämlich die Geobasisdaten und die Geofachdaten.

Gemäß dieser Definition bildet der Begriff der Geodaten die Gesamtmenge aller Geobasis- und Geofachdaten. Für die Begriffe der Geobasisdaten und Geofachdaten gelten nach [BILL, 1999] folgende Definitionen:

Definition 2-2:

Geobasisdaten sind eine Teilmenge der Geodaten. Zu Ihnen zählen insbesondere die Daten der Vermessungsverwaltungen, die als Grundlage für viele Anwendungen geeignet sind. Speziell umfasst der Geobasisdatensatz die vorhandenen Daten aus der ALK, dem ALB und aus ATKIS sowie die bisher separat geführte DGM und die gescannten topographischen Kartenwerke. Zukünftig zählen hierzu auch die Bilddaten wie Orthophotos, Luft- und Satellitenbilder.

Definition 2-3:

Geofachdaten (oder Fachdaten) sind die in den jeweiligen Fachdisziplinen erhobenen Daten. Durch den Zusatz ‚Geo‘ soll konkretisiert werden, dass auch diese Daten einen Raumbezug besitzen. Zumeist wird dieser Zusatz aber weggelassen.

Die Definition von Geobasisdaten muss vor dem Hintergrund der Weiterentwicklung des Geodatenmarktes durch mindestens zwei private Anbieter (TeleATLAS, NavigationTechnologies) um die routingfähige, europaweit verfügbare digitale Straßenkarte auf der Grundlage der GDF3.0-Spezifikation [CEN, 1995] erweitert werden. Diese Straßendaten bilden mittlerweile die Grundlage vielzähliger Web-GIS-Applikationen [BERNHARD, 2001]. Durch diese Entwicklung wurde der jahrelang staatlich dominierte Markt der Geobasisdaten erstmalig um privatwirtschaftliche Anbieter erweitert.

2.1.2 Geoinformationssysteme

Eine allgemeingültige Definition von Geoinformationssystemen (GIS) kann nicht gegeben werden. Dazu sind die Anforderungen, aber auch die bereits entstandenen Systeme zu unterschiedlich. Aus diesem Grund werden hier einige Definitionen aus der Literatur zitiert:

Ein Geoinformationssystem dient der Erfassung, Speicherung, Analyse und Darstellung aller Daten, die einen Teil der Erdoberfläche und die darauf befindlichen technischen und administrativen Einrichtungen sowie geowissenschaftliche, ökonomische und ökologische Gegebenheiten beschreiben. [BARTHELME, 1989]

...ein Informationssystem, dessen Datenbank aus Beobachtungen räumlich verteilter Objekte, Aktivitäten oder Ereignisse besteht, die durch Punkte, Linien oder Flächen definierbar sind. [DUEKER, 1979]

...eine umfassende Sammlung von Werkzeugen für die Erfassung, Speicherung, Bereitstellung im Bedarfsfall, Transformation und Darstellung raumbezogener Daten der realen Welt im Rahmen spezieller Anwendungen. [BURROUGH, 1979]

...ein Informationssystem, das alle raumbezogenen Daten der Atmosphäre, der Erdoberfläche und der Lithosphäre enthält und die systematische Erfassung, Aktualisierung, Verarbeitung und Umsetzung dieser Daten auf der Grundlage eines einheitlichen räumlichen Bezugssystems gestattet. [GÖPFERT, 1991]

...ein System zur Unterstützung der Entscheidungsfindung, das raumbezogene Daten in einer Problemlösungsumgebung integriert. [COWEN, 1988]

Bei diesen Definitionen stehen je nach unterschiedlicher Betrachtungsweise die Geodaten, die Lösungswerkzeuge für raumbezogene Analysen, die Hard- und Softwarekomponenten oder die Technologie im Vordergrund. Auf eine weiterführende Betrachtung der unterschiedlichen Definitionen und Betrachtungsweisen von Geoinformationssystemen kann hier verzichtet werden, da durch sie kein relevanter Informationsgewinn im Hinblick auf das Thema dieser Arbeit zu erwarten ist. Es wird deshalb an dieser Stelle auf die gängige Literatur [BARTHELME, 2000], [BILL, 1999], [LONGLEY, 1999], [MAGUIRE, 1991], etc. verwiesen.

2.1.3 Datenmodellierung

Um in einem Geoinformationssystem die reale Welt anhand von Daten repräsentieren zu können und daraus Informationen über Phänomene und Eigenschaften der realen Welt abzuleiten muss die reale Welt durch die spezielle fachgebundene Sicht der Wirklichkeit abstrakt dargestellt werden. Dieser Abstraktionsvorgang wird als **Modellierung** bezeichnet.

Das Datenmodell kann folglich als eine Vorschrift betrachtet werden, um die Objekte der realen Welt zu selektieren, sie zu benennen und festzulegen welche geometrischen und beschreibenden Merkmale die Objekte für die konzipierte Anwendung charakterisieren. Die Projektion einer durch das Datenmodell vorgegebenen Teilmenge der realen Welt wird als abstraktes Abbild der realen Welt bezeichnet. Die europäische Norm „Geoinformation - Datenbeschreibung – Qualität“ [prEVN12656, 1998] verwendet dazu den Begriff der „konzeptuellen Wirklichkeit.“

Das abstrakte Abbild der realen Welt muss unabhängig davon, wie komplex die Möglichkeiten zur Modellierung sind, immer ein eindeutiges, widerspruchsfreies Modell der Realität darstellen. Dadurch wird letztendlich festgelegt, was bei der Erfassung der Daten berücksichtigt werden soll. Erst wenn die Frage nach dem Informationsprodukt hinreichend geklärt ist, kann überlegt werden, welche Maßnahmen zur Gewinnung dieser Produkte benötigt werden und welche Daten dazu erforderlich sind.

Man kann nach der Art der Modellierung in *räumliche, konzeptionelle, logische und physikalische* Modellierung unterscheiden. In [BILL, 1994] ist diese Unterteilung als das *Vier-Schalen-Modell* der Modellierung bezeichnet. Es finden sich in der Literatur noch weitere Einteilungsmöglichkeiten wie z.B. in das externe, anwendungsspezifische Schema, das konzeptionelle Schema und das interne Schema [BARTHELME, 1995].

Eine ältere Einteilungsmöglichkeit geht auf [PEUQUET, 1984] zurück. Sie bezeichnet die hierarchische Ordnung als Datenmodell, Datenstruktur und Dateienstruktur. [SCHILCHER, 1996] gibt einen sequentiellen Ablauf der Modellierung in Form von anwendungs- und systemabhängigen Phasen an. In diesem Ablauf lässt sich die Auswirkung einer veränderten Sichtweise auf die Reale Welt zeigen. Durch sie wird eine neue Ist-Erhebung notwendig, welche in das konzeptuelle Datenmodell überführt werden muss.

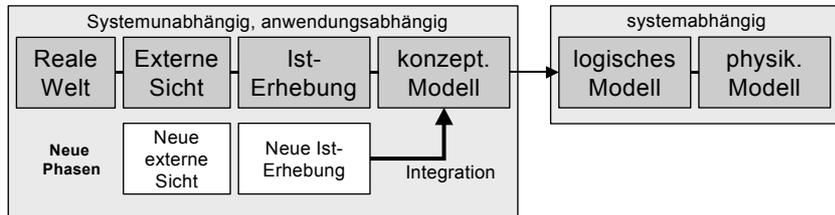


Abbildung 2.1: Phasen der Modellierung (verändert nach [Schilcher, 1996])

Diese ganzen Unterscheidungen haben alle die Gemeinsamkeiten, dass sie die reale Welt mit Hilfe eines anwendungsspezifischen Modells abstrahieren, um so die eindeutige Interpretation der Phänomene der realen Welt zu ermöglichen, die Interpretation in Form von Daten gemäß ihrer Bedeutung (Semantik) und wechselseitigen Beziehungen (Relationen) in eine Datenstruktur zu überführen und diese dann in einem physikalischen Speicher abzulegen. Zur Bewahrung der begrifflichen Klarheit wird fortan die Unterscheidung nach [BILL, 1994] verwendet, da man diese Unterteilung mittlerweile weitverbreitet in der Literatur zitiert findet.

Da das Datenmodell nicht die Daten modelliert, wie der Name vermuten lässt, sondern die reale Welt oder auch die Informationen, die in der realen Welt enthalten sind, wird in der Informatik oft auch der Begriff des Informationsmodells verwendet. Das logische und physikalische Modell modellieren tatsächlich die Daten, weshalb der Begriff *Datenmodell* hierfür angebracht erscheint. In der Literatur wird der Begriff *Datenmodell* allerdings oft für alle Stufen der Abstraktion von der realen Welt bis zur Speicherbelegung verwendet.

Da die Kosten der Datenerfassung und die Pflege von Geodatenbeständen wesentlich stärker von der Beschreibung von Phänomenen der realen Welt und deren Abstraktion in Form eines konzeptuellen Modells als von der Verwaltung der Daten und Informationen im logischen und physikalischen Modell abhängen, wird im Rahmen dieser Arbeit der Begriff des **Informationsmodells** weiter verwendet. Die Kostenauswirkungen unterschiedlicher logischer und physikalischer Modellierung gehen über das hier betrachtete Thema und die Zielsetzung der Arbeit hinaus und wird deshalb nicht berücksichtigt.

2.1.4 Signale, Daten, Nachricht und Informationen

Da im Zusammenhang mit raumbezogenen Informationen die Begriffe der Signale (Meß-), Daten (Meß-), Nachricht und Information Verwendung finden, und zumindest die Daten und Informationen oftmals synonym verstanden werden können, soll der Informationsbegriff nach [WIESER, 1989] hier eingeführt werden. Diese Begriffe finden maßgeblich in Informationssystemen ihre Verwendung, welche ihrerseits als Summe formalisierter, systembeteiligter Kommunikationsprozesse beschrieben werden können.

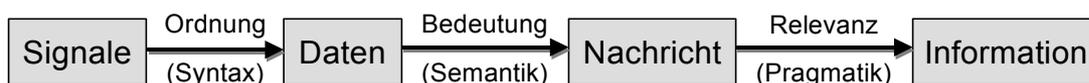


Abbildung 2.2: Zusammenhang zwischen Signalen, Daten, Nachrichten und Informationen

Originäre Messsignale werden, einer festgelegten Ordnung (Syntax) folgend, zu Daten aggregiert (ausgewertet, berechnet). Diese Daten können erst durch ihre Bedeutung, bzw. ihre Beziehung untereinander und zu ihrer Umwelt (Semantik) zu Nachrichten zusammengefasst werden. Der Begriff der Semantik versteht hierbei die Beziehungen zwischen Zeichen und Signalen und der sich daraus ergebenden Bedeutung. Die Informationen am Ende der Kette besitzen einen anwendungsbezogenen Charakter. Um aus Nachrichten Informationen zu generieren, bedarf es selektiver Maßnahmen wie z.B. der Relevanz gewisser Nachrichten bezüglich ihres kontextuellen Bezuges. So ist beispielsweise eine über Rundfunk ausgestrahlte Verkehrsstörung nur dann eine Information gemäß dieser Definition, wenn sie den Nutzer be-

trifft, also zeitlich aktuell und räumlich zutreffend ist. Sie ist relevant für den Nutzer. Für die Daten ergeben sich im Vergleich zu Informationen nach [BARTHELME, 2000] folgende typische Eigenschaften:

- Daten sind weniger strukturiert als Informationen,
- Semantische Aspekte werden codiert, wobei diese Codes nach Konventionen vergeben und interpretiert werden,
- anwendungsrelevante Aspekte fehlen, oder sie sind nur implizit im Kontext vorhanden

Aus den genannten Gründen wird in dieser Arbeit von Daten immer dann gesprochen, wenn keine direkt damit verbundene Anwendung im Bezug zur Argumentation steht und es sich um eine allgemeingültige Aussage handelt. Der Informationsbegriff findet folglich nur in Bereichen eine Verwendung, bei denen aus Daten nutzerspezifische Informationen erzeugt werden, welche einen direkten zeitlichen und räumlichen Bezug zum Informationsbedarf des Nutzers von Daten aufweisen. Er dient zur Verdeutlichung der strukturellen, syntaktischen, semantischen und pragmatischen Relevanz von Daten für deren Nutzung.

2.1.5 Geodaten, Geoinformationen und Geoprodukte

In enger Anlehnung an die Definition der Abhängigkeit zwischen Daten und Informationen kann auch der Zusammenhang zwischen Geodaten und Geoinformationen angegeben werden. Während es sich, vereinfacht ausgedrückt bei den **Geodaten** um den Inhalt einer Datenbank eines Geoinformationssystems handelt (inklusive seiner komplexen Strukturierung), sind zur Generierung von Geoinformationen Nutzerinteraktionen notwendig.

Ein Beispiel für **Geoinformationen** sind dynamisch generierte Abschnittskarten, die im Rahmen vieler Routingdienstleistungen im Internet angeboten werden. Raumbezogene Auskünfte auf mobilen Endgeräten zählen ebenso zu den Geoinformationen.

Die Palette der **Geodatenprodukte** ist vielfältig und lässt sich anhand mehrerer Alleinstellungsmerkmale charakterisieren. So werden Geodatenprodukte im Hinblick auf eine definierte Anwendung mit eingeschränktem Funktionsumfang einer festgelegten Zielgruppe zur kommerziellen Vermarktung angeboten. Die dem Produkt zu Grunde liegenden Geodaten werden bezüglich eines definierten Abgabeformat aufbereitet, welches sich entweder auf marktübliche Standards bezieht oder vor dem direkten Zugriff des Nutzers geschützt ist und die Nutzung nur durch eine, dem Produkt als wesentliche Komponente, hinzugefügte Softwarekomponente ermöglicht wird. Die Anbieter von Geodatenprodukten stellen Anwendungen z.B. für die Navigation oder das Geomarketing her. Dabei bilden nicht die Geodaten die eigentliche Information für den Kunden, sondern der spezifische Informationsgehalt entsteht erst durch die lösungsorientierte Einbindung der Geodaten in das Produkt. Beispiele hierfür sind die digitalen Rasterdaten der Vermessungsverwaltungen vieler Bundesländer (CD-ROM TOP50-Baden-Württemberg) oder im Handel frei verfügbare Routenplaner auf CD-ROM (z.B. Falk Routenplaner auf CD-ROM).

2.2 Funktionalarchitektur von Informationsdienstleistungen mit Geodaten

Zur Entwicklung von Informationsdienstleistungen auf der Basis von Geodaten bzw. deren Weiterentwicklung zu Geodatenprodukten kann folgende, in Abbildung 2.3 dargestellte Funktionalarchitektur der Informationsdienstleistungen angegeben werden [MÖHLENBRINK, 1998]. Sie bildet die notwendige Schnittstelle im Sinne einer Wertschöpfungskette von den Geobasisdaten hin zu Geoinformationen bzw. -produkten. Im Wesentlichen besteht sie aus den drei Institutionen der Informationsanbieter, den Informationsdiensten und den Systembetreibern.

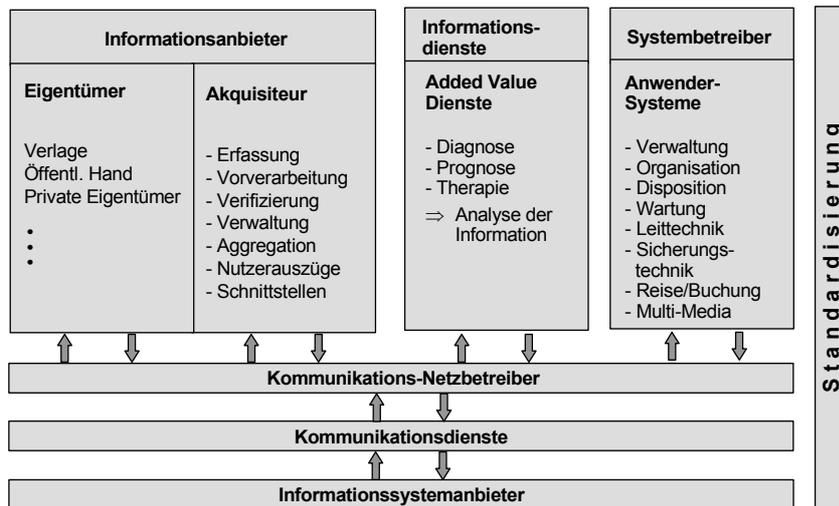


Abbildung 2.3:
Funktionalarchitektur der
Informationsdienstleister
[MÖHLENBRINK, 1998]

Die Informationsanbieter können sowohl Eigentümer als auch Akquisiteure von Geodaten sein. Sie haben einerseits den urheberrechtlichen Anspruch an den Daten und sind andererseits auch für deren Erfassung bzw. Aktualisierung verantwortlich. Diese Verantwortlichkeit ist entweder durch den gesetzlichen Auftrag zur Bereitstellung von Geobasisdaten festgelegt oder liegt im eigenen wirtschaftlichen Interesse privater Anbieter begründet.

Die Aufgaben der Datenakquisition umfasst die Aufgaben der Erfassung, der Vorverarbeitung, der Verifizierung, der Verwaltung, der Aggregation sowie die Erstellung von Nutzerausügen über definierte Schnittstellen.

Erst wenn diese Funktionalitäten ausreichend zur Verfügung stehen können Informationsdienste, also sogenannte Mehrwertdienste (Added value) daraus Diagnosen, Prognosen bzw. Lösungswege (Therapien) ableiten und so einem Nutzer situations- bzw. problemspezifische Auskünfte erteilen. Ob nun diese Informationsdienste Teil der Systembetreiber sind, oder ob sie lediglich durch dies hinzugekauft und integriert werden hängt vom Konzept des Systembetreibers ab. Dieser Systembetreiber wird jedenfalls durch die spezifische Informationszulieferung durch den Dienst in die Lage versetzt, eine Vielzahl von Anwendungssysteme mit Informationen zu versorgen und so dem Kunden diejenige Information zukommen zu lassen, nach der er nachgefragt hat. Wie bereits angesprochen verschwimmen die Grenzen in der Praxis häufig systembedingt zwischen den drei maßgeblichen Institutionen. Deren Funktionalitäten bleiben allerdings von den Systemgrenzen unbeeinflusst.

Diese Funktionalarchitektur, die den strukturellen Ablauf des Wertschöpfungsprozesses beschreibt, ist ohne eine begleitende Kostenbetrachtung zur Planung einer geodatenbasierten Informationsdienstleistung unvollständig. So müssen die, bei der Veredelung (added value), durchgeführten kostenverursachenden Arbeitsschritte durch ein speziell dafür geeignetes Verfahren modelliert und berechnet werden können. Erst durch die Erarbeitung eines Kostenmodells und dazugehörigen Berechnungsverfahren wird es möglich sein, den Prozess der Wertschöpfung auch kostenspezifisch begleiten zu können und damit das Niveau des Wertschöpfungsvorgangs anzugeben. Es ist demzufolge die zentrale Aufgabe dieser Arbeit ein solches Kostenmodell aufzustellen und dessen Wirkungsweise zu beurteilen. In folgendem Kapitel 3 wird deshalb zum besseren Verständnis eines Wertschöpfungsvorgangs die Realisierung einer digitalen Geodatenbasis zur intermodalen Routenberechnung beispielhaft erläutert und deren Kostenstrukturen aufgezeigt.

2.3 Ökonomische Eigenschaften raumbezogener Daten

Die Nutzung raumbezogener Daten zur Optimierung der Betriebsabläufe und zur Effizienzsteigerung von Unternehmen drang in den letzten Jahren immer stärker aus dem klassischen Geoinformationsmarkt in z.B. Immobilienmärkte, Freizeit- und Consumermärkte und nicht zuletzt auch in den logistischen Marktsektor (Auslieferungs- und Flottenlogistik, Direktmailing - Unternehmen) vor. Setzt man die richtigen Geodaten richtig ein, dann können damit die unterschiedlichsten Aufgaben der Navigation, Zielführung, Standortoptimierungen, Verbesserung der Transport- und Abfuhrlogistik etc. schneller, effizienter und vor allem kostensparender durchgeführt werden. Dadurch kommt den Geodaten eine besondere wirtschaftliche Bedeutung zu. Deswegen widmet sich dieser Abschnitt den ökonomischen Grundlagen und Eigenschaften, die mit Geodaten verbunden sind.

Ziel dieses Abschnitts ist es, die wirtschaftliche Relevanz raumbezogener Daten darzustellen und ihre Bedeutung für existierende und zukünftige Anwendungen in den unterschiedlichen Markt Bereichen aufzuzeigen. Dabei stehen in erster Linie die heterogenen Kundenanforderungen, die sich aus dem zahlreichen Anwendungsspektrum von Geodaten ergeben im Vordergrund. Ebenso soll der Frage nach den Hemmnissen eines bisher noch unzureichend erschlossenen Geodatenmarktes nachgegangen werden. Die Frage soll im Wesentlichen anhand zweier Marktstudien begründet werden.

Die Motivation für diese Betrachtungsweise liegt im kostendeckungsorientierten Gebührensystem staatlicher Organisationen begründet. So sind amtliche Geobasisdaten für privaten Nutzer derzeit abschreckend teuer. Dieses Gebührensystem wurde mit dem Ziel geschaffen, die hohen Kosten der Datenersterfassung und Aktualisierung nach dem Kostendeckungsprinzip an die Kunden weiterzugeben. Bei der Konzeption dieses Gebührensystems wurde jedoch übersehen, dass sich die Produktion von Geobasisdaten am Anfang der Wertschöpfungskette befindet und sich diese Daten erst durch die nutzungsorientierte Veredelung zu einer echten Wertschöpfung für den Kunden entwickelt. In der Marktstudie zur Aktivierung des Geodatenmarktes in Nordrhein-Westfalen [FORNEFELD, 2001] wird dabei vom „Wertschöpfungs-Paradoxon“ gesprochen (Abbildung 2.4).

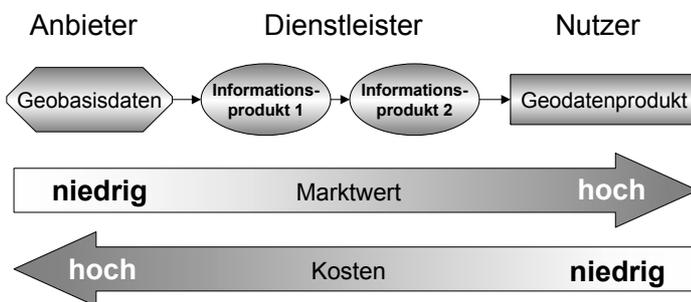


Abbildung 2.4: Das Wertschöpfungs-Paradoxon

Die staatlichen Anbieter stehen am Anfang der Wertschöpfungskette. Ihre Daten bilden zwar einen wichtigen Rohstoff für den Markt, sind aber für sich genommen nur von geringem Gebrauchswert. Zwischen Erhebung der Geobasisdaten und kundengerechtem Angebot von Geodatenprodukten liegt diese Wertschöpfungskette, in der viele Unternehmen Geodaten z.B. in andere Formate konvertieren, mit anderen Geodaten verschneiden, durch eigene Geofachdaten ergänzen und /oder für den Kunden in ein GIS integrieren. Nur für den enge Sektor der GIS - Anwendungen für Raum- und Städteplanung, der Vermessung, der Eigentumssicherung sowie der Besteuerung von Grundeigentum genügen diese Geobasisdaten i.d.R. den Anforderungen. Der eigentliche Mehrwert für den Nutzer entsteht erst mit der Datenveredelung durch entsprechende Dienstleister, obwohl diese i.d.R. mit deutlich geringeren Kosten verbunden sind als die Datengenerierung.

2.3.1 Anbieter und Nutzerstrukturen des Geodatenmarktes

Die Zahl der Anbieter und Nutzer von Geodaten ist viel zu umfangreich und komplex als dass sie im Rahmen dieser Arbeit ausführlich behandelt und diskutiert werden könnte. Zumal es bereits eine Vielzahl von Publikationen gibt, welche sich ausführlich mit der Entwicklung des Geodatenmarktes beschäftigen z.B. [BUHMANN, 2003], werden in diesem

Abschnitt lediglich die theoretischen Grundstrukturen aufgezeigt. Auch für eine detaillierte Übersicht der historischen Marktentwicklung sei hier ausdrücklich auf die Literatur z.B. [SCHILCHER ET AL, 1996] verwiesen.

Eine historische Entwicklung des Geodatenmarktes bzw. der damit einhergehenden Entwicklung der neuen Fachdisziplin Geoinformatik gibt [BARTHELME, 2000]. Er unterteilt die Entwicklung der noch relativ jungen Fachdisziplin in insgesamt 5 Phasen:

1. **Zeit der Pioniere** (1955-1975): Einzelne, individuelle Lösungen entstehen. Unternehmen gehen getrennte Wege. Es existieren kaum Geodaten.
2. **Zeit der Behörden** (1970-1985): Umstellung der analogen auf die digitalen Verwaltung von Geodaten.
3. **Zeit der Firmen** (1982-1990): Hard- und Softwarehersteller entdecken den Markt, erste Geoinformationssysteme entstehen.
4. **Zeit der Nutzer** (ab ca. 1988): Nutzerspezifische Lösungen werden zunehmend wichtiger, Datenmodellierungen und Applikationen werden weiterentwickelt.
5. **Zeit des offenen Marktes** (sein ca. 1995): Durch die Verbreitung des Internets ergab sich ein enormer Druck auf die Firmen, auf internetfähige Produkte umzusteigen. Dies beeinflusste auch die Bestrebungen nach Normung und Standardisierung von Geodaten.

Bedingt durch diese Entwicklung, entstand die bis heute noch existierende Monopolstellung der behördlichen Geobasisdatenproduzenten. Lediglich im Bereich der Fahrzeugnavigation konnte mit dem europäischen Forschungsprogramm EUREKA, das im Jahre 1985 unter Beteiligung von 25 europäischen Staaten gegründet wurde, der Grundstein für die europaweit verfügbare, navigations- und routingfähige digitale Straßenkarte gelegt werden, das mit Hilfe der beiden darin enthaltenen Projekte EDRM und PANDORA eine Spezifikation eines Objekt- und Attributwertkatalogs zur Beschreibung der Straßendaten erarbeitete. Daraus entstanden ist die heutige, auf internationaler Ebene standardisierte Spezifikation des Daten- und Austauschformates des Geographic Data File (GDF).

Diese Spezifikation bildete sozusagen die Grundlage für das Entstehen eines privaten Marktsektors zur Produktion von Geobasisdaten, in dem sich mittlerweile die beiden Firmen TeleATLAS und NavTech gleichermaßen etabliert haben.

Neben diesem, aus privatwirtschaftlichen Interesse heraus entstandenen, eigenständigen Marktsektor fungieren inzwischen mittlerweile auch viele Unternehmen als private Vertriebspartner öffentlicher Geodatenproduzenten. Die zum Teil unbefriedigende Situation der öffentlichen Geobasisdaten, führte dazu, dass einige Unternehmen (z.B. Telekom, Energieversorgungsunternehmen) unabhängig vom öffentlichen Vermessungswesen bei Bedarf eigene Geobasisdaten erhoben haben [GABRIEL, 2001] und diese auch vertreiben und damit in Konkurrenz zur öffentlichen Hand treten [HERDEG, 1996].

Die **Nutzer von Geodaten** können hinsichtlich Ihres Nachfrageverhaltens in elastisch und unelastisch eingeteilt werden. Nutzer mit unelastischer Nachfrage sind solche, die auf Geobasisdaten notwendigerweise angewiesen sind (Energieversorgungsunternehmen, Raum- und Umweltplaner, Architekten, etc.). Ihr Nachfrageverhalten bleibt in der Regel konstant und wird sich entsprechend langsamer der Preisentwicklung von Geodaten anpassen. Elastische Nachfrage nach Geodaten existiert hingegen im derzeit stark wachsenden Consumer- und Freizeitmarkt. Werden entsprechende Geodaten bzw. Geoinformationen zu einem nutzerorientierten Preis angeboten, steigt die Nachfrage an. Diese Situation trifft insbesondere für die Kunden digitaler Freizeitkarten die z.B. auf CD-ROM angeboten werden zu.

Hier existiert für die amtlichen Geobasisdaten ein enormer Wachstumsmarkt und entsprechend Chancen sich darin zu positionieren.

2.3.2 Nutzen und wertbeeinflussende Eigenschaften

Die wirtschaftliche Relevanz von Geodaten kann durch ein Satz von Merkmalen beschrieben werden, die ihrerseits von den unterschiedlichen Kundengruppen und Anbietern abhängig sind. Dazu ist insbesondere die Frage zu stellen, inwiefern Geodaten wirtschaftliche Güter darstellen, welche produktionsbeeinflussende Eigenschaften sie in welchen Markt-

sektoren besitzen und welche Chancen und Risiken sich durch den Einsatz von Geodaten in einem Unternehmen ergeben.

Allgemein betrachtet bestimmen die Eigenschaften von Waren bzw. Gütern maßgeblich deren Absatzchancen auf dem Markt. Für Geodaten als Ware, lassen sich dabei folgenden allgemeingültige Eigenschaften definieren:

- Geodaten unterliegen derzeit besonderen **Kostenstrukturen**.
- Das **öffentliche Interesse** an Geodaten ist **sehr groß**.
- Geodaten verfügen über **externe Effekte**
- Das **Marktpotential** aber auch die **Hindernisse** liegen im **breiten Anwendungsspektrum**. Ihre **Konfektionierung** für die Absatzmärkte gelingt **nur in bestimmten Segmenten**.
- Geodaten müssen in **rein und in nicht reine öffentliche Güter** unterschieden werden können. Die Geodaten-dienstleister müssen ihr Angebot inhaltlicher (Dateninhalt) vielfältiger gestalten und die Tarifierungsstrategien entsprechend detaillierter ausgestalten.

Zur Erarbeitung der charakterisierenden Merkmale für Geodaten als wirtschaftliche Güter sind die unterschiedlichen Marktbereiche in denen Geodaten aktuell und künftig zum Einsatz kommen sollen sowie die existierenden heterogenen Kundenanforderungen zu analysieren, damit aus dieser Analyse die entsprechenden Strategien zur Bereitstellung der Geodaten entwickelt werden können.

2.3.3 Kostenstrukturen amtlicher und privater Geodaten

Die Kostenstrukturen von Geodaten können allgemein entsprechend den Bearbeitungsphasen in Erhebungs-, Aktualisierungs-, und Diffusionskosten unterschieden werden.

Zu den **Erhebungskosten** zählen die Kosten für die erstmalige Erfassung. Sie werden repräsentiert durch die Kosten für Fahrten, Material, Geräte, Lagerung, Aufbereitung und Personal, sowie durch die Kapitalkosten für einmalig getätigte Investitionen. Für bestimmte Geodaten fallen die Erhebungskosten nur einmalig an (z. B. Grundstücks- und Katasterdaten). Liegen diese Daten flächendeckend vor, zählen die Kosten der Datenfortführung und Verbesserung zu den Aktualisierungskosten. Werden unter Umständen in einem bestimmten Zyklus Totalrevisionen notwendig, so entstehen diese Erhebungskosten erneut (z.B. Bestandessituation der Waldflächen). Dagegen existieren Daten, bei denen die Notwendigkeit einer Aktualisierung einer Neuerfassung gleichkommt (z.B. Volkszählung). Diese Situation ist von der zeitlichen und mengenbezogenen Stabilität der Daten abhängig.

Die Kosten der **Aktualisierung** sind Kosten für den Werterhalt einer einmalig getätigten Investition. Auch sie können durch die Parameter Fahrtkosten, Material., Geräte, Lagerung, Aufbereitung und Personal charakterisiert werden, da oftmals die Prozesse der Aktualisierung entsprechend gleichzusetzen sind mit denen der Erhebung.

Die **Diffusionskosten** fallen beim Vertrieb der Geodaten an und beinhalten die Kosten für die Bearbeitung der Anfrage, die Bereitstellung der Daten, das Porto sowie die Fakturierungskosten (Kosten der Rechnungsschreibung). Zu den Diffusions- oder Vertriebskosten zählen außerdem die zum Vertrieb notwendigen Kosten für Personal und Infrastruktur (Informatikmittel, Mobiliar, Räume).

3 Digitale Geodaten für die Intermodale Routensuche

3.1 Einleitung

Durch ein besseres Zusammenwirken der Verkehrsträger (Straße, Schiene, Wasser, Luft) kann ein wichtiger Beitrag zur Sicherung der individuellen Mobilität bei stetig steigendem Verkehrsaufkommen geleistet werden. Eine Möglichkeit des besseren Zusammenwirkens besteht in einer verkehrnetzübergreifenden, intermodalen Routenplanung in Ballungszentren. Dort sind Verkehrsbelastungen der Straßen sehr hoch und gleichzeitig das Verkehrsnetz des öffentlichen Personennahverkehrs (ÖPNV) gut ausgebaut. Zahlreiche Projekte der Öffentlichen Hand, wie z. B. die durch das Bundesministerium für Bildung und Forschung (BMBF) geförderten Leitprojekte (1998-2003) zum Thema „Mobilität in Ballungsräumen“ (WayFlow, MOBILIST, Intermobil, StadtInfoKöln, MobiNet, CashCar) haben eine Steigerung der Attraktivität des Öffentlichen Verkehrs (ÖV) im Rahmen von Forschungsvorhaben (Mobilitätsmanagementzentralen, ÖV-Anschlussinformationssysteme, System zur IV/ÖV-Verkehrsprognose) zum Ziel.

In diesem Abschnitt soll anhand des im Leitprojekt MOBILIST entstandenen *Intermodalen Dienst* [MOBILIST, 2001] der Aufwand dokumentiert werden, der notwendig gewesen war, eine digitale Geodatenbasis für die intermodale Routenberechnung zwischen dem Individualverkehr (Pkw, Fahrrad, Fuß) und den Verkehrsmitteln des öffentlichen Personennahverkehrs aufzubauen. Dieser Intermodale Dienst wurde in Zusammenarbeit zwischen der Forschung der DaimlerChrysler AG, dem Verkehrs- und Tarifverbund Stuttgart (VVS), der MentzDV GmbH und den Stadtwerken Ulm realisiert. Durch die Projektbeteiligung des Instituts für Anwendungen der Geodäsie im Bauwesen der Universität Stuttgart im Rahmen der MOBILIST - Begleitforschung ergab sich hierbei meinerseits die Möglichkeit zur detaillierten Analyse der Arbeitsprozesse sowie der zugehörigen Kostenstrukturen.

Dieses Kapitel zeigt am praktischen Beispiel die notwendigen Arbeitsschritte des Geodatenaufbaus vor dem Hintergrund einer verkehrstelematischen Anwendung und analysiert die dabei entstehende Kostenstruktur. Anhand dieses Beispiels soll die Forderung nach einer modellgestützten Kostenprognose zusätzlich motiviert werden.

3.2 Grundlagen einer intermodalen Routensuche

Die beiden grundlegenden technischen Voraussetzungen eines Intermodalen Routenplaners sind einerseits die Routen- und Tourenplanungssysteme des Individualverkehrs (IV) sowie andererseits die Fahrplan - Auskunftssysteme internationaler und regionaler Unternehmen des öffentlichen Personenverkehrs.

Das Problem der Berechnung der kürzesten Wege in gerichteten bzw. ungerichteten Graphen für den IV ist durch die Algorithmen von z. B. [MOORE, 1957], [BELLMANN, 1958], [DIJKSTRA, 1959] schon seit Jahren gelöst und wird inzwischen in Kombination mit der europaweit verfügbaren digitalen Straßenkarte, basierend auf dem Datenstandard des *Geographic Data File* (GDF), auf zahlreichen Internetplattformen (Map24, map&guide, etc.) als kommerzielle oder frei zugängliche Softwareprodukte bereitgestellt. Verkehrsmittelübergreifende und –interne Fahrtauskünfte des ÖV basieren hingegen auf elektronischen Kursbüchern mit statisch hinterlegten Fahrplandaten. In Deutschland dominieren im ÖPNV die Fahrplan-Auskunftssysteme EFA (Fa.MentzDV) und HAFAS (Deutsche Bahn AG). Durch jahrelange Entwicklungsarbeit haben sie eine hohe Auskunftqualität erreicht. Die darin angebotenen Dienstleistungen reichen von einer Haltestelle-zu-Haltestelle-Auskunft bis hin zur Möglichkeit der Adresseingabe für Start- und Zielpunkt [FIEHN, 2002].

3.2.1 Informationsbedarf des Kunden

Bei einer intermodalen Routenberechnung werden diese beiden Systeme des IVs und des ÖVs so miteinander gekoppelt, dass einem Reisenden neben der reinen Fahrtauskunft mit dem Pkw bzw. einem öffentlichen Verkehrsmittel eine Reiseauskunft unter Nutzung beider Verkehrsmittel inklusive der Beschreibung des Umsteigevorgangs zur Auswahl gestellt werden kann. Notwendige Voraussetzung dieser Kopplung ist eine gemeinsame Geodatenbasis, welche die Umsteigepunkte (Haltestellen des ÖV) zwischen dem Straßennetz und dem Verkehrsnetz des öffentlichen Personenverkehrs beinhaltet und umsteigerelevante Informationen in Form von Attributen (Fußwegzeit, P&R-Platz, B&R-Platz etc.) enthält. Konkret bedeutet die gemeinsame Geodatenbasis eine Verknüpfung dieser Verkehrsnetze (Routingfähiges Stra-

Bennetz mit ÖV-Netz inklusive Schienennetz, Abbildung 3.1) in Form einer Zuweisung der Haltestellen des ÖVs zu den entsprechenden Straßenkanten.

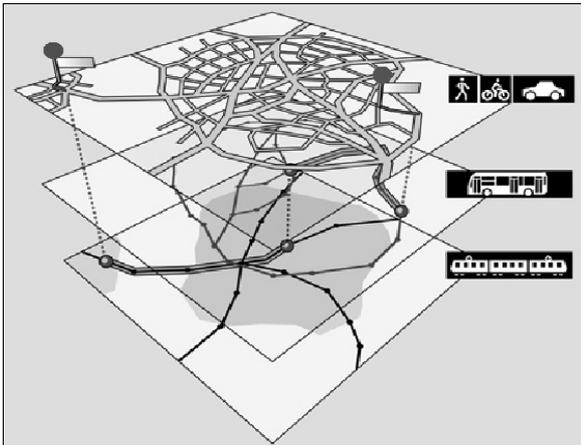


Abbildung 3.1: Verkehrsnetze [BRACHT, 2002]

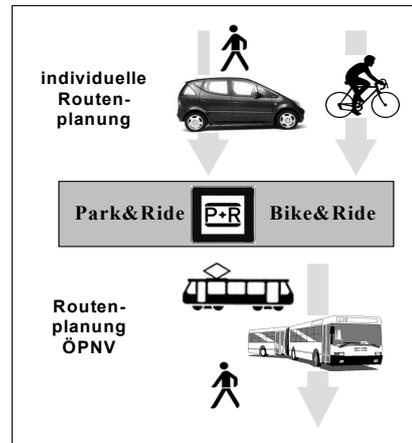


Abbildung 3.2: Intermodale Wegekette

Die Start- und Zielpunkte einer Intermodalen Routenberechnung sind folglich die georeferenzierten Umsteigestellen (ÖV-Haltestellen) zwischen den Verkehrsnetzen. Der Ablauf der intermodalen Routenberechnung erfolgt, entsprechend nach logischen Ebenen getrennt, durch die Teilphasen:

- der Suche nach geeigneten Übergangspunkten (ÖV-Haltestellen),
- der Suche von Teilrouten in den verschiedenen Verkehrsnetzen,
- der Kombination der gefundenen Teilrouten sowie durch
- die Auswahl der besten Gesamtrouten unter Berücksichtigung von Präferenzen.

Eine intermodale Route wird vom Kunden allerdings nur dann als Alternative zu einer reinen IV- bzw. ÖV-Route gewählt werden, wenn die Übergangswiderstände an den Umsteigestellen minimal sind. Die Übergangswiderstände können sowohl durch persönliche Präferenzen, z. B. Gewohnheit, Anforderungen an eine bestimmte Art der Gepäckbeförderung, aber auch durch Informationsdefizite über den Umsteigeweg, die benötigte Zeit sowie die angeschlossenen Verkehrsmittel bestimmt sein. Aus diesem Grund ist eine alleinige Objektrelation zwischen Haltestellenobjekten und Straßenkanten unzureichend. Vielmehr müssen sämtliche umsteigerelevanten Informationen (vgl. Abbildung 3.3) modelliert und einer gemeinsamen Datenbank zugeführt werden.

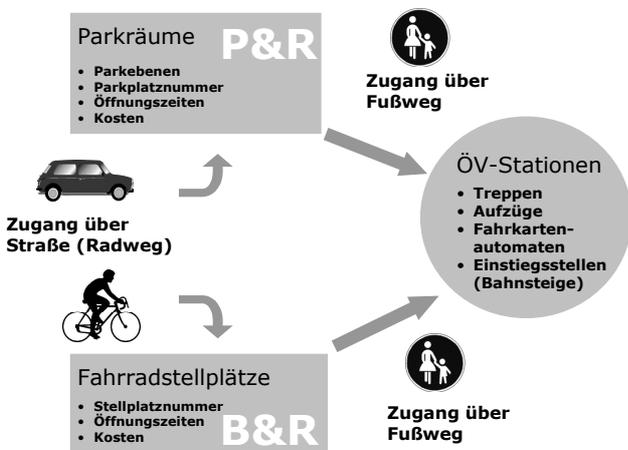


Abbildung 3.3: Informationen an den Netzübergängen [STARK, 2001]

Die innerhalb des Forschungsprojektes MOBILIST hauptsächlich durch den Verkehrs- und Tarifverbund Stuttgart (VVS) realisierte Geodatengrundlage beinhaltet die Umsteigeinformationen in Form von Zugängen zu ÖV-Haltestellen, angebundene Park&Ride (P&R) sowie Bike&Ride (B&R) – Plätze und berücksichtigt die Umsteigezeit in Form einer

Fußwegematrix zwischen den Teilbereichen der ÖV-Haltestellen. Darüber hinausgehende Informationen wie z. B. Parkebenen, Parkplatznummern, freie Parkräume oder Buchungsfunktionalitäten sind nicht enthalten.

3.2.2 Betrachtete Umsteigebeziehungen

Das Umsteigen zwischen den Verkehrsmitteln (Pkw, Fahrrad, Fuß, Bahn, Bus) der Verkehrsträger (Straße, Schiene) ist prinzipiell in allen Kombinationen möglich. Jedoch unterliegen die Umsteigebeziehungen gewissen Wahrscheinlichkeiten oder gar Einschränkungen. So kann die Wahrscheinlichkeit eines Wechsels zwischen Pkw und Schiene als hoch angenommen werden, wogegen der Umstieg von Fahrrad zu Pkw als sehr gering einzuschätzen ist. Aus Sicht einer Verkehrsinformationsdienstleistung (z. B. EFA) betrachtet, die solche Umsteigevorgänge berechnet und den Kunden aufbereitet weitergibt, werden auf Grund diesen Wahrscheinlichkeiten und Einschränkungen bestimmte Verkehrsmittelkombinationen aus der intermodalen Routenberechnung ausgeschlossen.

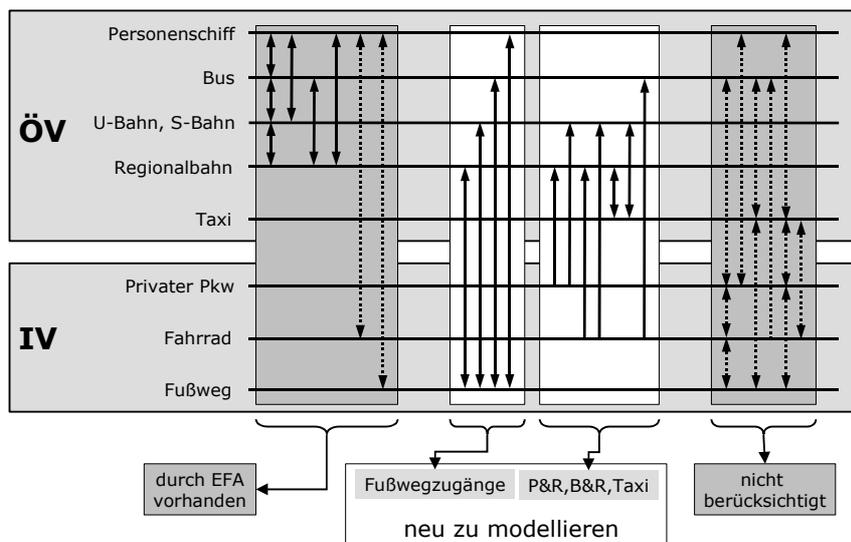


Abbildung 3.4:
vorhandene, hinzugefügte und
unberücksichtigte Umsteige-
beziehungen

Abbildung 3.4 zeigt die durch das Fahrplanauskunftssystem EFA bereits vorhandenen, die neu zu modellierenden sowie die in diesem ersten Schritt des Geodatenaufbaus nicht berücksichtigten Umsteigebeziehungen zwischen IV und ÖV. Derzeit noch nicht berücksichtigt wurden IV-interne Umsteigebeziehungen und solche, die den ÖV-Anteil an die erste Stelle vor den IV-Anteil der Reisekette stellen. Zur Realisierung einer solchen Reisekette wären weit größere Investitionen in die Infrastruktur (Pkw/Fahrrad-Leihstellen an den Haltestellen) sowie Buchungs-, Abrechnungs- und Kontrollfunktionen notwendig. Sie sind deshalb aus der hier beschriebenen Betrachtung ausgeschlossen.

Im Rahmen dieser zusätzlichen Modellierung werden die Fußwegverbindungen zwischen den Haltestellen und den Straßenelementen flächendeckend realisiert sowie die Umsteigebeziehungen zwischen IV und ÖV als bidirektionale Verbindungen vom IV-Verkehrsmittel (Pkw, Fahrrad) zum ÖV-Verkehrsmittel (Regionalbahn, S- bzw. U-Bahn) eingeführt. Da keine der Bushaltestellen des VVS zum Zeitpunkt des Geodatenaufbaus über P&R-Plätze verfügt, wurde die Umsteigebeziehung IV-Bus lediglich für das Fahrrad in Form der Fahrradmitnahme berücksichtigt.

3.3 Aufbau der Geodatengrundlage

In den Ballungszentren entstanden in den letzten Jahren verstärkt infrastrukturgestützte Parkleitsysteme sowohl im Zentrums-kern als auch an peripheren ÖPNV-Haltestellen in Form spezieller Park&Ride - Informationssysteme. Trotz dieser baulichen Infrastrukturmaßnahmen und Informationstafeln verbleibt für den Nutzer eines IV-Verkehrsmittels das Problem der Anschlusssicherung sowie die Transparenz bezüglich seines individuellen Vorteils bei einem Verkehrsmittelwechsel. Es ist ihm nur schwer möglich, innerhalb kürzester Zeit bei der Vorbeifahrt an einer P&R-Informationstafel einschätzen zu können, ob er die darauf angebotene ÖV-Alternative zum IV auch tatsächlich erreicht und welche zeitlichen oder monetären Vorteile ihm daraus entsteht. Aus diesem Grund trägt die Gewährleistung der Anschlusssicherung und die Schaffung einer transparenten Kosten-Nutzendarstellung maßgeblich zur Minimierung der Umsteigewiderstände bei. Ein Beitrag zur Minimierung dieser Widerstände kann geleistet werden, wenn der Umsteigeprozess durch eine solche, in Abbildung 3.5 dargestellte, prozessbegleitende, lückenlose Informationskette beschrieben werden kann.



Nicht in dieser Informationsprozesskette dargestellt sind z. B. die Park- und Fahrscheinkosten, Informationen über die Parkraumsicherheit und –verfügbarkeit. Eine solche umsteigeprozessbegleitende Informationskette kann nur dann individuell, dynamisch und aktuell berechnet werden, wenn ein Datenbestand existiert, welcher die Park&Ride (P&R) bzw. Bike&Ride (B&R) Plätze, Zugangswege und -zeiten enthält und mit dem die berechneten Umsteigeverbindungen visualisiert werden können.

3.3.1 Vorgehensweise zur Realisierung der Geodatengrundlage

Aufbauend auf der existierenden Datengrundlage des VVS, den digitalen Kursbüchern, die der Elektronischen Fahrplanauskunft (EFA) bisher zur ÖV - Routenberechnung genügten, wurde die digitale Straßenkarte (NavTECH) für die IV – Routenberechnung hinzugekauft. Diese beiden Komponenten bildeten die Grundlage für die weiter notwendigen Arbeitsschritte, der Erweiterung um Zusatzdaten und die damit verbundene Verknüpfung der Verkehrsnetze.



Das Umsetzungswerkzeug war hierfür das Geoinformationssystem DIVAGeo der Fa. MentzDV. Es integrierte die digitale Straßenkarte zusammen mit den digitalen Kursbüchern der Elektronischen Fahrplanauskunft und einer georeferenzierten, gescannten analogen topographischen Übersichtskarte (TÜK 200) der Region Stuttgart im Maßstab 1:200 000.

Zu Beginn der Verknüpfung der Verkehrsnetze konnten die Haltestellen der EFA nur sehr ungenau geographisch dargestellt werden, da jedes Haltestellenobjekt nur über eine so genannte Haltestellenhauptkoordinate verfügte, die ausschließlich dem Zweck diente, sie einigermaßen lagerichtig vor dem Hintergrund der TÜK200 darzustellen. Eine Verbindung zwischen Haltestelle und Straßenelement bestand zu diesem Zeitpunkt noch nicht.

Da diese Datengrundlage (digitale Kurspläne und TÜK200) von Beginn an bereits vorlag und sie ohne Zusatzaufwand vollständig übernommen werden konnte, werden die Kostenfaktoren dieses Grundlagenaufbaus und der Datenübernahme in dieser Betrachtung nicht berücksichtigt. Die in diesem Abschnitt herausgearbeitete Kostenstruktur bezieht sich

damit auf diejenigen Kostenanteile, mit denen gerechnet werden muss, wenn eine elektronische Fahrplanauskunft in der Vergangenheit bereits realisiert wurde und sie nun um die Aufgaben der intermodalen Routenberechnung erweitert werden soll.

Zur Garantie einer geschlossenen Reisekette über die Umsteigestellen hinweg, sind die vorhandenen Übergänge zwischen den verschiedenen Verkehrsmitteln zu modellieren, in die entsprechende Datenbankstruktur zu integrieren und die notwendigen Informationen zu erfassen und hinzuzufügen. Diese Übergänge werden dann durch Zuordnung der ÖV-Haltestellen zum digitalen Straßennetz mittels Georeferenzierung, also der Zuweisung einer geographischen Position (Koordinate) mit entsprechender Objektverknüpfung, realisiert.

Der Arbeitsablauf der Modellierung und Erfassung komplexer Haltestellen, ihrer Zuordnung zum korrespondierenden Straßenelement sowie deren Attributierung wird nachfolgend unter dem Begriff der **IV/ÖV-Modellierung** verwendet. Der Arbeitsablauf besteht dabei aus den Einzelschritten der

- Nachbearbeitung der digitalen Straßenkarte, der
- Erfassung des ÖV-Netzwerks, der
- Erweiterung des ÖV-Haltestellenmodells und Datenerfassung, der
- Georeferenzierung der Haltestellen und Linien und der
- Überprüfung der Bearbeitung.

Da in den folgenden Abschnitten die technische Umsetzung der Einzelschritte anhand von Aufwandsabschätzungen quantifiziert wird, ist es hier vorab nötig, die Vorgehensweise zur Ermittlung des Arbeitsaufwands zu erläutern. Die Aufwandsabschätzung greift dazu auf die Arbeitszeitprotokolle des für die IV/ÖV-Modellierung verantwortlichen Bearbeiters der VVS zurück. Die Protokolle standen in Form einer Datenbank für diese Untersuchung zur Verfügung, in der die einzelnen Bearbeitungsschritte mit Datums- und Zeitmarkierungen versehen waren. Die daraus ableitbare Arbeitsdauer ist jedoch als reine Bearbeitungszeit ohne Maßnahmen zur Bereitstellung von Erfassungsgrundlagen, Einkauf und Verwaltung zu betrachten und wird nachfolgend als Netto-Arbeitszeit bezeichnet. Eine Betrachtung der Netto-Zeit alleine reicht zur Aufwandsschätzung jedoch nicht aus. Vielmehr sind die Maßnahmen zur Arbeitsvorbereitung mit zu berücksichtigen.

Durch die Analyse der im Rahmen des Projekts MOBILIST zur Verfügung stehenden Kostenstrukturen ergab sich ein Verhältnis zwischen den Arbeitsanteilen der Arbeitsvorbereitung und den Arbeitsprozessen von 56,5% zu 43,5%. Der Nachweis, dass der Aufwand der Arbeitsvorbereitung den der eigentlichen Arbeitsprozesse beim Geodatenaufbau übersteigen kann, konnte bereits im Rahmen des Benchmark Tests zur prototypischen Realisierung der digitalen Straßenkarte [KRONJÄGER, 1992] nachgewiesen werden.

Dieses prozentuale Verhältnis zwischen der eigentlichen Erfassungsleistung und den notwendigen Vorarbeiten wird fortan zur Abschätzung des zeitlichen „Overheads“ veranschlagt. Dabei ist besonders darauf hinzuweisen, dass eine allgemeine Gültigkeit dieses prozentualen Verhältnisses nicht pauschal für alle Erfassungsvorgänge von Geodaten vorausgesetzt werden kann. So reduzieren z. B. Routinetätigkeiten, die auf die einmalig hergestellte arbeitsvorbereitende Grundlage zurückgreifen, dieses prozentuale Verhältnis nachhaltig. Es behält folglich nur für den erstmaligen Geodatenaufbau seine Gültigkeit und muss z. B. für den Vorgang der Fortführung und Aktualisierung entsprechend prozessbezogen verändert werden.

3.3.2 Nachbearbeitung der digitalen Straßenkarte

Da die digitale Straßenkarte zum Zwecke der Fahrzeugnavigation und Routenberechnung aufgebaut wurde [CEN,1995], enthält diese keinerlei flächendeckende Informationen über das Fußgängerrouting. Diese Informationen sind hinsichtlich der zu erreichenden Auskunftqualität des intermodalen Routenplaners notwendig, um die Wegstrecke von Start- bzw. Zielkante zur Start- bzw. Zielhaltestelle entsprechend den Anforderungen an ein Fußgängerouting richtig berechnen zu können. Darüber hinaus stehen darin derzeit zur adressscharfen Routenberechnung nur unzureichend Straßenkanten zur Verfügung, die mit Hausnummernbereichen versehen sind. Die Nachbearbeitung der digitalen Straßenkarte umfasst deshalb

- die Erfassung von Fußwegen an Haltestellen,
- die Änderung der Attributierung der Straßenelemente sowie
- das Nacherfassen fehlender Hausnummernbereiche.

3.3.2.1 Erfassung von Fußwegen

Die Erfassung von Fußwegen dient zur Erhöhung der Erreichbarkeit der Haltestellen (vgl. [JERMANN, 2002] und [PRINZ, 2002]) von den umgebenden Straßenkanten. Diese Fußwege sind als neue Vektorelemente hinzuzufügen, zu attributieren (Befahrbarkeit, zugelassenes Verkehrsmittel) und mit den Straßenkanten zu verknüpfen. Da eine flächendeckende Verfügbarkeit von Fußwegen nicht durch die digitalen Straßendaten bereitgestellt werden kann, werden die Fußwege durch das Verfahren der Bildschirmdigitalisierung hinzugefügt. Als Erfassungsgrundlage dient eine überlagerte Darstellung am Erfassungsbildschirm mit der topographische Karte (TK) im Maßstab 1:25.000 bzw. in kleinmaßstäbliche Bereichen die topographische Übersichtskarte (TÜK-1:200.000) und den vorhandenen digitalen Straßenkanten (Abbildung 3.7).



Abbildung 3.7: Erfassung von Fußwegen

Bei einer Gesamtanzahl von 3440 Haltestellen im Verkehrsgebiet des Verkehrs- und Tarifverbundes Stuttgart (VVS) mussten für 1376 Haltestellen neue Fußwegverbindungen geschaffen werden. Da im Durchschnitt pro Haltestelle ca. zwei Fußwege hinzugekommen sind, kann von einer Gesamtanzahl von 2750 neuen Fußwegen in Form von Kanten-elementen ausgegangen werden.

3.3.2.2 Änderung der Attributierung der Straßenelemente

Die Änderung der Attributierung der Straßenelemente im Umkreis der Haltestellen ist notwendig, da die digitalen Straßendaten im originalen NavTech-Bestand Attribute enthalten, die die Nutzung des Straßenelements in Form von zugelassenen Verkehrsmitteln (Motorisierter Verkehr, Bevorrechtigter Verkehr, Eisenbahn, Straßenbahn) beschränken. Eine Änderung dieser Attributierung ist für die Berechnung der Fußwege von der Haltestelle zur Zielkante notwendig, da sonst ein Routingalgorithmus wichtige Kanten außer Acht lässt, was unter Umständen zu einem längeren und aus Sicht eines Kunden falsch berechneten Fußweg führt.

Diese Änderung wird durch die Erweiterung der Attributwertliste der zugelassenen Verkehrsmittel um „Fußweg“ und „Radweg“ realisiert. Speziell bei innerstädtischen Bundesstraßen in der Nähe von ÖPNV-Haltestellen muss deren Attributierung spezifisch angepasst werden, da nicht jede Bundesstraße über Geh- oder Radwege verfügt. Die Ermittlung der Netto-Arbeitszeit für die Fußwegerefassung und Änderung der Straßenkantenattributierung stützt sich auf die Auswertung der VVS-internen Haltestellendatenbank, welche bei jeder Bearbeitung dem neuen Element eine Datums- und Zeitmarkierung hinzufügt. Mit Hilfe dieser Datenbank konnte die Netto-Arbeitszeit zur Nachbearbeitung der Straßenkanten in der Umgebung von Schienenhaltestellen mit 28 h angegeben werden.

Die Ermittlung der Netto-Arbeitszeit für die Straßenkantenbearbeitung im Umkreis von Bushaltestellen stützt sich auf eine Hochrechnung von empirisch gemessener Bearbeitungsdauer (Netto-Zeit). Es wurden mit Hilfe von 10 repräsentativen Kartenausschnitten und der darin enthaltenen Anzahl von 5 Haltestellen eine durchschnittliche Bearbeitungsdauer von 4 min pro Kartenausschnitt ermittelt.

Für die 3100 Bushaltestellen ergeben sich damit rund 620 zu bearbeitende Kartenausschnitte, woraus eine Netto-Arbeitsdauer von 41 h resultiert. Für die Nachbearbeitung der Straßenkanten im Umkreis von Schienen und Bushaltestellen entsteht damit eine Netto-Arbeitsdauer von hochgerechneten 69 h (siehe Zusammenfassung in Tabelle 3-1).

Tabelle 3-1: Netto-Arbeitszeiten für die Erfassung der Fußwege und deren Attributierung

Arbeitsschritte	Netto-Arbeitszeit [h]
Erfassung von Fußwegen und Attributänderung	
a) an den Schienenhaltestellen	28,0
b) an den Bushaltestellen	41,0
<i>Gesamtsumme:</i>	<i>69,0</i>

Für die Berechnung der Bruttoarbeitszeit wird zusätzlich der Overheadfaktor für die notwendigen Vorarbeiten angesetzt (vgl. Tabelle 3-2).

Tabelle 3-2: Arbeitsschritte mit Zeitanteil der Nachbearbeitung der Straßenkanten

Arbeitsschritte		Zeitanteil	Zeit (h)
Vorarbeiten	• Beschaffung und Georeferenzierung TK+TÜK im GIS als Hintergrund	56,5%	90
	• Einteilung Kartenausschnitte für Bushaltestellen		
Eigentliche Arbeit	• Erfassung von Fußwegen und Attributierungsänderung an Schienenhaltestellen	43,5%	28
	• Erfassung von Fußwegen und Attributierungsänderung an Bushaltestellen		41
<i>Brutto-Arbeitszeit:</i>		<i>100%</i>	<i>159</i>

Die **Hausnummernbereiche**, welche eine Grundvoraussetzung für das Fußgängerouting von den Haltestellen zu den Start- und Zielkanten darstellen (Haustür-zu-Haustür), werden derzeit nur unvollständig im Datenbestand der digitalen Straßenkarte vorgehalten. So lagen z. B. im März 2001 für das Gebiet der Bundesrepublik Deutschland im TeleAtlas-Datenbestand (TELEATLAS, 2001) nur für ca. 48% der Bevölkerung tatsächlich erfasste Hausnummernbereiche für die Straßenkanten vor. Die folgende Abbildung 3.8 visualisiert diejenigen Gebiete der Region Stuttgart, die im TeleAtlas Datenbestand (Release 0.07) über Hausnummernbereiche der Straßenkanten verfügen. Nebenstehende Tabelle 3-3 zeigt den Attributierungsgrad der Straßenkanten sowie den prozentualen Anteil von Hausnummernbereichen an der Bevölkerung.

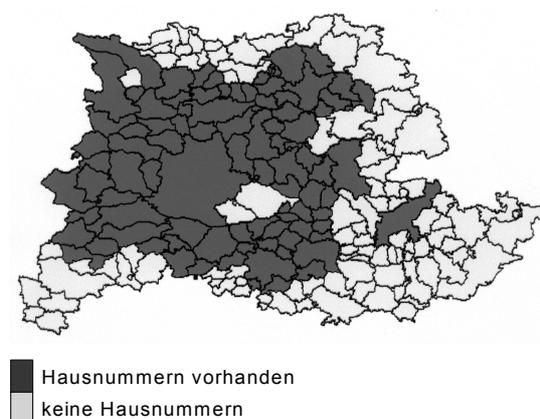


Abbildung 3.8: Hausnummernbereiche bei TeleAtlas (Rel.0.07, 12/00)

Kreis	Digitalisierungsgrad	vollständig attribuiert	Hausnummern vorhanden
		in % Anteil an Bevölkerung	
SK Stuttgart	100%	100%	100%
LK Böblingen	100%	100%	75,60%
LK Esslingen	100%	89,10%	63,30%
LK Göppingen	100%	77,70%	21,60%
LK Ludwigsburg	100%	97,40%	83,40%
LK Waiblingen	100%	97,60%	80,30%

Tabelle 3-3: Attributierungsgrad der digitalen Karte von TeleAtlas (Stand Dez.2000)

Neben den in den digitalen Straßendaten teilweise vorhandenen Adressdatenbeständen in Form von Hausnummernbereichen, musste die alte VVS-interne statische Straßen-Haltestellen-Zuordnung in das Auskunftssystem mit integriert werden. Bei dieser statischen Zuordnung wurden Straßenkanten den Haltestellen in ihrer Umgebung in tabellarischer Form zugeordnet. Diese statische Zuordnung ist bezüglich der Verkehrsmittel- und Routenwahl zu unflexibel gegenüber der hausnummernscharfen Adresszuordnung zum Straßenelement. Da jedoch das vollständige Nacherfassen fehlender Hausnummernbereiche der Region ein enormer Arbeits- und Kostenaufwand bedeuten würde, entschied man sich, bis zur Verfügbarkeit der kompletten Hausnummernbereiche im Bestand der digitalen Straßenkarte bzw. der vollständigen Verfügbarkeit der Hausnummernkoordinaten aus Katasterkarten an beiden Möglichkeiten zur Adresslokalisierung und Haltestellenzuordnung festzuhalten. Folgende Möglichkeiten ergeben sich nach der beschriebenen Fallunterscheidung zu:

- 1.) Sind Hausnummernbereiche im digitalen Straßendatenbestand verfügbar, werden diese zur Haltestellenbestimmung herangezogen.
- 2.) In Gebieten ohne Hausnummernbereiche werden die statischen „Straße-zu-Haltestelle“ Zuweisungen weiter verwendet. Darin enthaltene lange Straßenabschnitte, die nicht nur einer Haltestelle zuzuordnen sind, mussten entsprechend der vorhandenen Straßensegmentierung aufgeteilt und die Hausnummernbereiche nacherfasst werden.

Insgesamt wurden durch die Firma MentzDV 22.314 Adressdatensätze in die GIS- Daten integriert, was ungefähr 4000 Straßen entspricht. Durch die leistungsbezogene Auftragsvergabe des VVS fielen rund 800 Brutto-Arbeitsstunden an, die durch eine angeleitete Hilfskraft durchgeführt werden konnten. Daraus ergibt sich folgende Kostenabschätzung:

- | | |
|--|----------|
| • Kosten für eigene Bearbeitung
(VVS – Fußwege + Attributierung – 159h – 70€/h) | € 11.130 |
| • Kosten für Leistungseinkauf
(MentzDV – Adressdatenbestände – 800h – 50€/h) | € 40.000 |

Gesamtkosten für die Nachbearbeitung der digitalen Straßenkarte	€ 51.130
---	-----------------

3.3.3 Erfassung des ÖV-Netzwerks

Die geographisch lagerichtige Erfassung des ÖV-Netzwerks (Schienen- und Buslinien) ist nicht notwendige Grundvoraussetzung der Intermodalen Routenberechnung, da die ÖV-Fahrtauskünfte durch digitale Kursbücher auch ohne geographischen Bezug ermittelt werden können.

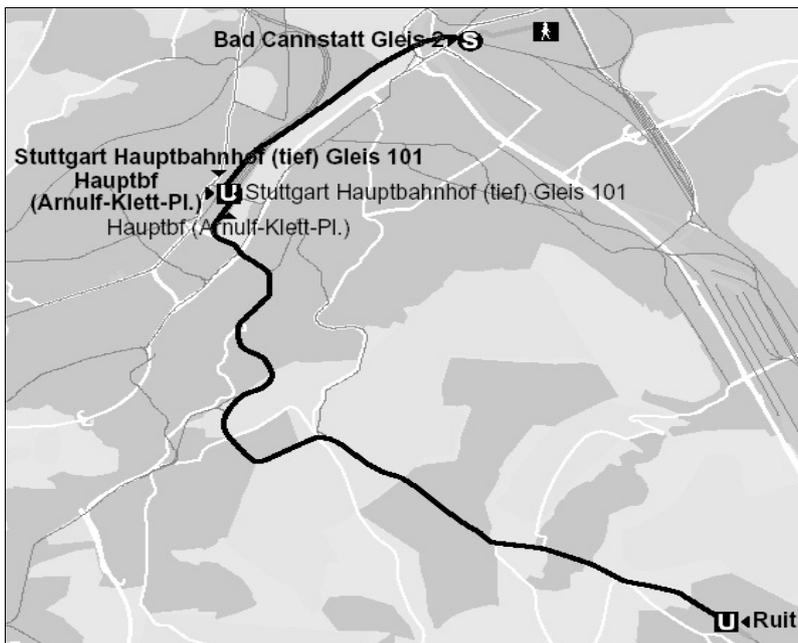


Abbildung 3.9: dynamisch generierte, individuelle Verkehrslinienpläne

Ohne eine solche Erfassung würden allerdings die Verbindungen zweier ÖV – Haltestellen als Geradenverbindung im Rahmen der kartographischen Visualisierung erscheinen. Die Erfassung der Netzwerkgeometrie wurde deshalb durch den VVS zusätzlich durchgeführt und dient zur verbesserten Auskunftqualität in Form von dynamisch generierbaren Verbindungsplänen, die die ermittelten ÖV-Verbindungen lagerichtig darstellen können. Die lagerichtige Fahrstreckendarstellung war jedoch nicht Hauptmotivation des VVS, das ÖV-Netzwerk digital zu erfassen, sondern der dadurch entstehende digitale Verkehrslinienplan, der künftige Arbeitsprozesse der kostengünstigen Pflege und Aktualisierung erst ermöglichen sollte.

Außer den VVS-internen analogen Verkehrslinienplänen, die mit enormem kartographischem Aufwand bei jeder Fahrplanumstellung angepasst werden müssen, lagen keine digitalen Geometriedaten über die ÖV-Linien des VVS vor. Da allerdings die Schienenwege der Deutschen Bahn AG teilweise im GDF-basierten digitalen Datenbestand der Herstellerfirmen (TeleAtlas bzw. NavTech) zum Zweck der kartographischen Präsentation und besseren Orientierung in der Kartenanzeige von Kfz-Navigationssystemen enthalten sind, konnte diese Geometrie zum Aufbau des ÖV-Netzwerks genutzt werden.

Da sowohl S-Bahnen als auch Regionalbahnen die Schienen der Deutschen Bahn AG nutzen, musste nur in ländlichen Gebieten und dort nur für die Regionalbahnen ca. 10% der Streckengeometrie (ca. 56 km) nacherfasst werden. Die Grundlage zur Referenzierung der S-Bahnstrecken konnte vollständig durch die Geometrie in der digitalen Straßenkarte bereitgestellt werden. Die Liniengeometrie der Stadtbahnen im VVS-Gebiet waren zu Beginn der IV/ÖV-Modellierung allerdings nicht digital verfügbar. Sie wurde mit Hilfe von analogen Verkehrslinienplänen und Orthophotos durch Digitalisieren erfasst. Für die Buslinien war keine eigene Netzwerkgeometrie nachzuerfassen, da die Linienführung der Busse ausschließlich auf den vorhandenen Straßenelementen stattfindet.

Die Aufgabe der Erfassung des ÖV-Netzes bestand demnach darin, die Linienverläufe der Stadtbahnen vollständig und die Verläufe der fehlenden DB-Strecken (~10% Regionalbahnen) nachzudigitalisieren. Aus den Arbeitszeitprotokollen ergibt sich entsprechend Tabelle 3-4.

Tabelle 3-4: Netto-Arbeitszeit der Digitalisierung des Schienennetzes nach Zeitprotokoll

Arbeitsschritte der Digitalisierung	Netto-Arbeitszeit in [h]
Schienenstrecken der Stadtbahnen mit ca. 203 km	32
Schienenstrecken der Regionalbahnen mit ca. 56 km	

Berücksichtigt man erneut des zeitlichen Overhead zwischen eigentlicher Erfassung und den Vorarbeiten, so ergibt sich nachfolgende Abschätzung:

Tabelle 3-5: Arbeitsschritte mit Zeitanteil der Digitalisierung des Schienennetzes

Arbeitsschritte:		Zeitanteil	Zeit (h)
Vorarbeiten	• Bestandsaufnahme und Bearbeitung in der digitalen Straßenkarte enthaltender Schienengeometrie	56,5%	42
	• Orientierung und Identifikation fehlender Geometrie		
	• Individuelle Zusammenstellung der Erfassungsgrundlagen		
Eigentliche Arbeit	• Neuerfassung der fehlenden Schienengeometrie	43,5%	32
Brutto-Arbeitszeit:		100%	74

Da diese Tätigkeit durch den VVS ohne eine Leistungsvergabe an Dritte durchgeführt wurde, resultieren die Kosten der ausführenden Arbeit mit einem, in dieser Arbeit angenommenen, Stundesatz von €70/h zu € 5.180.

3.3.4 Modellierung der ÖPNV-Haltestellen

Das Ziel der IV/ÖV-Modellierung ist die Zuordnung einzelner Haltestellenobjekte zu den Straßen- oder Schienenelementen. Die dafür notwendige Vorschrift ergibt sich aus dem Datenmodell der Firma MentzDV [MENTZDV, 2001].

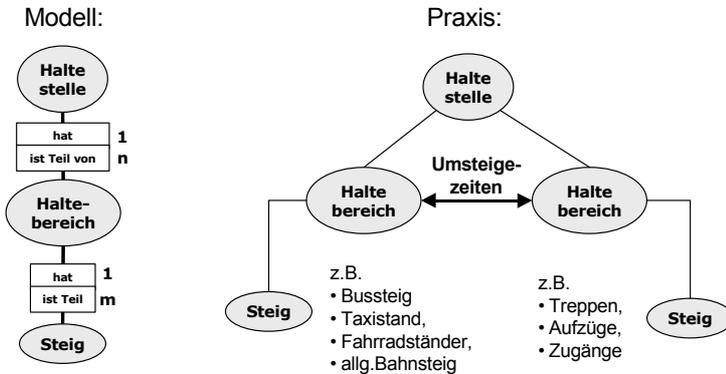


Abbildung 3.10: Datenmodell von ÖPNV-Haltestellen

Darin wird eine *Haltestelle* durch ein Objekt repräsentiert, welches in die Objektteile *Haltestellenbereiche* und dieses wiederum in *Steige* aufgegliedert sind. Eine *Haltestelle* wird immer dann in Bereiche unterteilt, wenn mehr als ein ÖV- oder IV- Verkehrsmittel (S-Bahn, U-Bahn, Auto, Taxi, Fahrrad, Bus) an diese *Haltestelle* angeschlossen ist. Die *Haltestellenbereiche* können ihrerseits in die *Steige* mit den zugehörigen Attributen der Steignummer, des Steignamens, der geographischen Referenz (Koordinate) unterteilt werden. Bei diesen *Steigen* kann es sich sowohl um Bahnsteige als auch um Treppen, Aufzüge und ggf. Zugangswege für Fußgänger zu Haltestellen handeln. Sie sind die geographisch genauen Punkte des Ein- oder Ausstiegs bzw. des Wechsels des Verkehrsmittels. Zur Festlegung der Umsteigezeiten dienen die Haltebereiche. Sie sind gemäß Datenmodell die Start- und Zielpunkte der Fußwege. Dieses Datenmodell ist Grundlage der Integration der Haltestellen mit den daran angeschlossenen intermodalen Bereichen sowie den umsteige-relevanten Informationen. Zur übersichtlicheren Darstellung kann der Arbeitsprozess der Modellierung der ÖPNV-Haltestellen in die Teilprozesse der

- Beschaffung von Erfassungsgrundlagen (Neuerfassung, interne Erhebung), in die
- Modellierung und Erfassung intermodaler Haltestellenbereiche (inkl. Kontrolle Erfassungsgrundlagen) sowie in die
- Erfassung und Integration der Umsteigezeiten zwischen alten und neuen Bereichen

unterteilt werden. Der Teilprozess der Modellierung intermodaler Haltestellenbereiche ist zweckmäßigerweise zusätzlich in den *Modellierungsvorgang bezüglich P&R, B&R, Taxi –Bereiche* und in den *Modellierungsvorgang der Fußgängerzugangssteige* zu unterteilen.

3.3.4.1 Beschaffung von Erfassungsgrundlagen

Als Erfassungsgrundlage dieser Haltestellenmodellierung dienten beim VVS vorhandene Haltestellenumgebungspläne, die zur Orientierung an den Haltestellen ausgehängt und als zusätzliche Serviceleistungen als Karten im Internet angeboten werden. Diese Umgebungspläne wurden in der Vergangenheit durch Fremdaufträge digital erstellt und lagen zu Beginn der IV/ÖV-Modellierung für die Gesamtzahl von 3440 Haltestellen nahezu vollständig vor. Lediglich für 72 Regionalbahnhaltestellen in ländlichen Regionen des VVS-Gebiets mussten diese Umgebungspläne durch eine „Vor-Ort“-Erfassung neu erstellt werden. Neben diesen Umgebungsplänen dienten weitere analoge Kartenwerke als sekundäre Erfassungsgrundlagen. Ihr Zweck war die Überprüfung und ggf. Aktualisierung der hausinternen Umgebungspläne. Durch diese sekundären Erfassungsgrundlagen konnte gewährleistet werden, dass die in der realen Welt neu entstandene P&R-Plätze auch in die Umgebungspläne integriert wurden. Für die Herstellung und Aktualisierung der 3440 Haltestellenumgebungspläne waren die folgenden beiden Arbeitstätigkeiten durchzuführen:

- Erfassung von 72 neu zu erhebenden Plänen (Handskizzen und deren digitale Ausarbeitung)
- Überprüfung von 3368 Umgebungsplänen (vorhandene intermodale Bereiche).

3.3.4.2 Erfassung intermodaler Haltestellenbereiche

Da im Rahmen der früheren Realisierung des Fahrplan-Auskunftssystems EFA sämtliche Haltestellenobjekte (337 Schienenthaltestellen und 3103 Bushaltestellen) bereits in die entsprechenden Haltestellenbereiche und Steige im internen Datenbestand des VVS vorliegen (Regionalbahn, S-Bahn, Stadtbahn, Bushaltestellen), mussten die zur intermodalen Routenberechnung notwendigen Bereiche und Steige sowie Fußwegzugänge bei insgesamt 122 Schienenthaltestellen hinzugefügt werden. Eine Erweiterung der existierenden Bushaltestellen um **Intermodale Haltestellenbereiche** (P&R, B&R, Taxi) fand nicht statt, da die damit verbundenen Umsteigebeziehungen *Bus-Pkw*, *Bus-Fahrrad*, *Bus-Taxi* in beiden Richtungen nicht Gegenstand der IV/ÖV-Modellierung sind (vgl. Abbildung 3.4).

Tabelle 3-6: Neue Bereiche der Schienenthaltestellen des VVS¹

Existierende ÖV Kombination	Anzahl	Neue P&R	Neue B&R	Neue Taxi
Mit 1 Bereich (default)	23	21	12	3
Mit 2 Bereichen	82	62	53	29
Mit 3 Bereichen	16	13	14	14
Mit 4 Bereichen	1	1	0	1
Gesamt	122	97	79	47

In Tabelle 3-6 ist die Anzahl neu anzulegender Haltestellenbereiche in Abhängigkeit zu den bereits existierenden Haltestellenkombinationen dargestellt. So existierten beispielsweise vor der IV/ÖV-Modellierung 82 Schienenthaltestellen mit zwei Bereichsunterteilungen, denen im Zuge der Modellierung insgesamt 62 neue P&R-Plätze, 53 neue B&R-Plätze sowie 29 Taxistände hinzugefügt worden sind. Insgesamt wurden im Zuge der IV/ÖV-Modellierung 122 Schienenthaltestellen des VVS um 223 Bereiche erweitert.

3.3.4.3 Modellierung und Erfassung der Fußgängerzugänge

Neben der Modellierung und Erfassung der Bereiche für den intermodalen Verkehrsmittelwechsel kommt der Modellierung und Erfassung der Fußgängerzugangswege (kurz: Zugänge) besondere Bedeutung zu. Für diese Zugänge können entweder extra Steige einem bereits existierenden Haltestellenbereich hinzugefügt oder eigene Haltestellenbereiche (Zugangsbereiche) geschaffen werden. Die Einführung eines Haltestellenbereichs „Zugang“ dient hauptsächlich bei großen verteilten Umsteigehaltestellen zur richtigen Berechnung der Gehzeiten zwischen den Ausgängen (Treppe, Rolltreppe, Aufzug) und den vorhandenen Haltestellenbereichen.

Die Abbildung 3.11 zeigt die unterschiedliche Modellierung in Fußgängersteige und –bereiche am Beispiel der Haltestelle „Stuttgart Hauptbahnhof (S) tief“. Darin wird die Unterscheidung in die vier Zugangsbereiche Zugang-West, Zugang-Mitte, Zugang-Ost und Zugang-Süd mit den dazu gehörenden Eingängen, die als Steige den Zugangsbereichen zugeordnet sind, deutlich. Wegen seiner Komplexität ist der Stuttgarter Hauptbahnhof insgesamt in 3 Objekte von Typ Haltestelle unterteilt. Es existiert für die S-Bahn, die Stadt-Bahn und die Regionalbahn eine eigene Haltestelle mit den entsprechenden Zugängen. Am Haltestellenbereich Nr.6-Taxi-Platz sieht man die Problematik der richtigen Gehzeitschätzung, da hier der Fußweg von den beiden seitlich von Bahnhof gelegenen Taxiständen zu den S-Bahngleisen wesentlich länger dauert als vom Taxisteig direkt vor dem Hauptausgang. Da sich die Angabe der Gehzeiten auf den Fußweg zwischen den jeweiligen Bereichen bezieht, wird sie sich entsprechend am zeitintensivsten bzw. am weitesten Fußweg orientieren. Im Rahmen der IV/ÖV-Modellierung waren für insgesamt 12 Haltestellen neue Zugangsbereiche dieser Art zu schaffen.

¹ Der Grund für die fehlende Unterteilung von Haltestellen mit einem Bereich liegt in der praktischen Durchführung der Modellierung begründet. Grundsätzlich werden alle Haltestellenobjekte mit einem Standardbereich angelegt, da ein einheitliches Datenmodell keine direkte Zuordnung von Steigen zu Haltestellen erlaubt. Eine Haltestelle ohne Bereich entspricht sozusagen einer Haltestelle mit nur diesem Standardbereich.

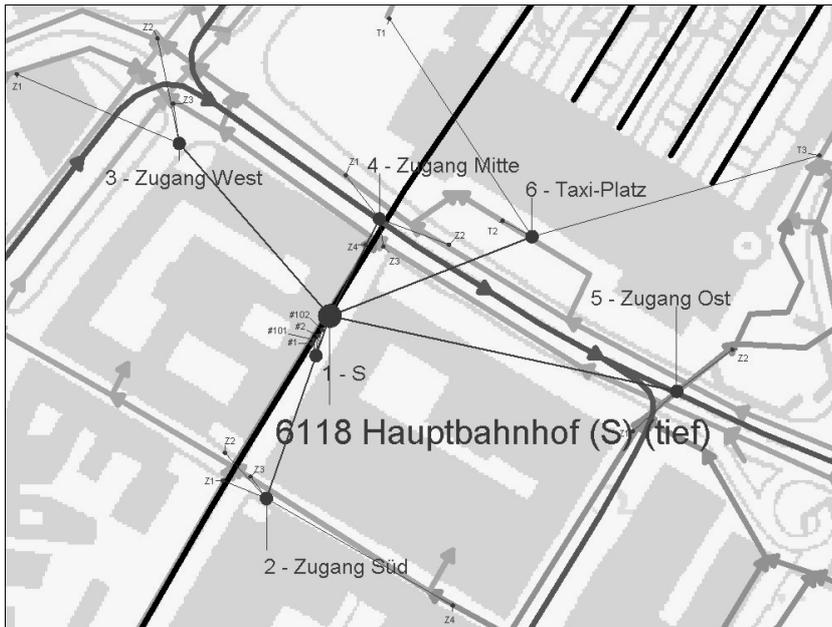


Abbildung 3.11: Haltestelle Stuttgart Hauptbahnhof (S-Bahn) tief

Die Anzahl neu zu modellierender Zugangssteige hängt direkt von der Anzahl an Haltebereichen einer Schienenhaltestelle ab, da pro Schienentyp mindestens ein Zugangssteig vom Straßennetz zum entsprechenden Haltebereich hinzuzufügen war. Es ergeben sich mindestens 359 neu anzulegende Zugangssteige (Tabelle 3-7). Berücksichtigt man ferner, dass für jeden Schienentyp im Durchschnitt 1,5 Steige angelegt wurden, dann sind 540 Zugangssteige, welche direkt an einen existierenden ÖV-Bereich angefügt werden, als gute Näherung für die weitere Abschätzung anzunehmen. Der durchschnittliche Netto-Zeitbedarf zur Modellierung eines Zugangssteiges von 1 min führt zum Netto-Zeitaufwand der vollständigen Bearbeitung von 540 min (~9 h).

Tabelle 3-7: Zugangssteige an Schienenhaltestellen

Schienenkombination	Anzahl Haltestellen	Zugangssteige angelegt (mindestens)
S+R+U	1	3 (für je 3 Schientypen)
S+R	13	26 (für je 2 Schientypen)
S+U	6	12 (für je 2 Schientypen)
R+U	0	0 (für je 2 Schientypen)
S	52	52 (für je 1 Schientyp)
R	75	75 (für je 1 Schientyp)
U	191	191 (für je 1 Schientyp)
Gesamt:		359 Zugangssteige

Erfassung und Integration der Umsteigezeitwerte: Sind die intermodalen Bereiche sowie die Fußgängerbereiche und –steige einer Haltestelle modelliert, dann können die Umsteigezeiten zwischen den neuen und alten Haltestellenbereichen erfasst und eingetragen werden. Dafür steht eine Umsteigezeitmatrix, die Bestandteil des GIS-Systems DIVAGeo (MentzDV) ist, zur Verfügung. In Abbildung 3.12 ist die Fußwegematrix der Haltestelle Stuttgart-Vaihingen als Beispiel abgebildet. Der Wertebereich der Zeiten liegt darin zwischen 0 und 5 min mit einer Auflösung von einer Minute.

	06002 6	06002 7	06002 1	06002 4	06002 3	06002 2	06002 5	06002 10	06002 11	06002 12	
	BuBüJ / 0	BuRoh / 0	S / 0	StAnk / 0	StFeA / 0	StMöh / 0	ZOB / 0	P+R / 0	B+R / 0	Taxi / 0	
6	BuBüJ / 0	00.0	01.0	03.0	01.0	02.0	02.0	02.0	03.0	03.0	02.0
7	BuRoh / 0	01.0	00.0	04.0	02.0	03.0	03.0	03.0	03.0	03.0	02.0
1	S / 0	03.0	04.0	02.0	03.0	02.0	02.0	03.0	03.0	03.0	03.0
4	StAnk / 0	01.0	02.0	03.0	00.0	01.0	01.0	01.0	02.0	03.0	02.0
3	StFeA / 0	02.0	03.0	02.0	01.0	00.0	01.0	02.0	02.0	03.0	02.0
2	StMöh / 0	02.0	03.0	02.0	01.0	01.0	00.0	02.0	02.0	03.0	02.0
5	ZOB / 0	02.0	03.0	03.0	01.0	02.0	02.0	01.0	04.0	03.0	01.0
10	+R / 0	03.0	03.0	03.0	02.0	02.0	02.0	04.0	xxx	xxx	xxx
11	+R / 0	03.0	03.0	03.0	03.0	03.0	03.0	03.0	xxx	xxx	xxx
12	axi / 0	02.0	02.0	03.0	02.0	02.0	02.0	01.0	xxx	xxx	xxx
1	ulnd / 0	04.0	05.0	04.0	04.0	04.0	04.0	05.0	04.0	04.0	05.0

Abbildung 3.12: Fußwegematrix der Haltestelle Stuttgart-Vaihingen

Der umrandete Teil stellt die im Rahmen der IV/ÖV-Modellierung neu hinzugefügten Umsteigezeitwerte für die Haltestelle Stuttgart-Vaihingen dar. Am Beispiel dieses umrandeten Teils wird die Vorgehensweise der Ermittlung des Arbeitsaufwandes zur Modellierung dieser Relationen deutlich. Der gesamte Zeitbedarf ist abhängig von der Gesamtzahl der neu zu schaffenden Umsteigezeitwerte. Diese wiederum sind ihrerseits von der Anzahl der Haltebereiche einer Haltestelle abhängig. Demzufolge berechnet sich die Gesamtanzahl der neuen Umsteigezeitwerte N_G^{Um} aus der Anzahl der Haltebereiche n vor der IV/ÖV-Modellierung und den neu geschaffenen Haltebereichen m allgemein mit:

$$N_G^{Um} = 2(m \cdot n) + m^2 - m$$

Die Umsteigezeitmatrix ist in der Regel eine symmetrische Matrix, da die Fußwegzeiten zwischen Hin- und Rückweg normalerweise gleich groß sind. Hiervon ausgenommen können z. B. Zeiten für lange Treppen (Rolltreppen) sein. Dort müssen unter Umständen für das Emporsteigen längere Zeiten als für das Hinabsteigen angesetzt werden. Deshalb kann auf eine getrennte Speicherung der Fußwegzeiten im Hin- und Rückweg nicht generell verzichtet werden.

Tabelle 3-8 zeigt die Ermittlung der Anzahl neuer Umsteigezeiten getrennt in diejenigen Haltestellen, denen ein, zwei oder drei neue Bereiche hinzugefügt wurden. So existierten z. B. 3 Haltestellen, die bereits in 7 Bereiche unterteilt waren und denen drei neue Bereiche hinzugefügt wurden (erste Tabellenzeile, links-3 neue Bereiche). Insgesamt waren 1035 neue Zeitwerte im Rahmen der Ermittlung neuer Umsteigezeitwerte zu erfassen und in die Umsteigezeitmatrix zu integrieren. Dieser Aufwand verteilt sich entsprechend den Prozentangaben auf die 3 Gruppen (neu hinzuzufügende Bereiche).

Tabelle 3-8: Ermittlung der Anzahl neuer Umsteigezeiten

3 neue Bereiche			2 neue Bereiche			1 neuer Bereich		
exist. Bereich	Anzahl Haltestellen	Neue Umsteigezeiten	exist. Bereich	Anzahl Haltestellen	Neue Umsteigezeiten	exist. Bereich	Anzahl Haltestellen	Neue Umsteigezeiten
7	3	72	7	1	16	9	1	10
6	5	105	6	1	14	8	3	27
5	4	72	4	5	50	7	1	8
4	8	120	3	10	80	6	4	28
3	11	132	2	20	120	5	3	18
2	5	45	1	3	12	4	2	10
1	1	6				3	7	28
						2	14	42
						1	10	20
Anzahl Haltestellen	37	552	Anzahl Haltestellen	40	292	Anzahl Haltestellen	45	191
Gesamtzahl Bereiche	111	53,30%	Gesamtzahl Bereiche	80	28,21%	Gesamtzahl Bereiche	45	18,45%

Erleichterter Erfassungs- und Integrationsaufwand: Der Arbeitsaufwand zur Schaffung dieser Fußwegzeiten lässt sich mit Hilfe einer Eingabefunktion, die die Zuweisung der Matrixelemente $a_{i,j} = a_{j,i}$ vornimmt, erleichtern. Dieser Vorgang entspricht der Spiegelung einer Matrix an Ihrer Hauptdiagonalen. Ferner werden die Umsteigezeitwerte mit einer Auflösung von einer Minute durch das GIS streckenabhängig ermittelt. Bei der Ermittlung des jeweiligen Zahlenwerts

der Umsteigezeit wird oftmals mit einer pauschalen Zeit von einigen Minuten gearbeitet. Diese pauschale Zeit kann, je nach räumlicher Entfernung zwischen den Haltestellenbereichen, einen distanzabhängigen Zuschlag bekommen. Die dafür verwendete Strecke ist die euklidische Distanz zwischen den Koordinaten der entsprechenden Haltebereiche.

Arbeitsaufwand zur Erfassung der Grundlagen: Aus einer Analyse der Arbeitszeitprotokolle des Bearbeiters beim VVS resultiert eine Netto-Erfassungsdauer für die 72 neu zu erhebenden Pläne von 38 Arbeitsstunden. Diese Zeitdauer umfasst ausschließlich die An- und Abfahrtszeiten zu den Haltestellen sowie die Dauer der Erstellung der Handskizzen der Haltestellen vor Ort (*Eigentliche Arbeit*). Berücksichtigt man darüber hinaus die Zeiten

- der Tourenplanung (Ausarbeitung der Haltestellenreihenfolge),
- der Überprüfung vorhandener Umgebungspläne sowie
- die Ausarbeitung der Handskizzen zu digital gezeichneten Umgebungsplänen

mit dem Faktor von 130% („Overhead“ für Arbeitsvorbereitung), so resultiert daraus die Brutto-Erfassungsdauer von 87h für die 72 neu zu zeichnenden Umgebungspläne.

Arbeitsaufwand der Erfassung: Folgende Tabelle gibt die Netto-Arbeitszeiten in Stunden für die Teilprozesse der Haltestellenerfassung an. Die Zeitwerte stammen wie bei der „Vor-Ort“-Erfassung aus der Analyse der Arbeitszeitprotokolle des Bearbeiters.

Tabelle 3-9: Netto-Arbeitszeiten nach Zeitprotokoll

Arbeitsschritte	Netto-Arbeitszeit in [h]
Erfassung intermodaler Bereiche	7,4
Erfassung Zugangsbereiche (Fußgänger)	1,0
Erfassung Zugangssteige (Fußgänger)	9,0
Erfassung und Integration Umsteigezeiten	8,6
<i>Gesamtsumme Nettoarbeitszeit:</i>	<i>26,0</i>

Die Korrektheitsprüfung der existierenden Haltestellenumgebungspläne (neue Intermodale Bereiche in der Realen Welt) findet nicht innerhalb der Erfassung von Umgebungsplänen statt, sondern ist Teil des Modellierungs- und Erfassungsvorgangs der Haltestelle. Ihr Aufwand wird bei der Erfassung und nicht bei der Bereitstellung der Erfassungsgrundlagen berücksichtigt. Es ergibt sich die Gesamtsumme der Nettoarbeitszeit der Haltestellenbearbeitung zu 26 Stunden. Berücksichtigt man denselben Zeitfaktor von 130% („Overhead“ für Arbeitsvorbereitung) zwischen Netto- und Brutto-Erfassungszeit so ergibt sich die Einteilung gemäß Tabelle 3-10.

Tabelle 3-10: Arbeitsschritte mit Zeitanteil der Haltestellenerfassung

Arbeitsschritte:		Zeitanteil	Zeit (h)
Arbeitsvorbereitung	• Orientierung des Bearbeiters	56,5%	34
	• Auswahl der entsprechenden Erfassungsgrundlage (Suche und Auswahl der Umgebungspläne + Stadtkarten)		
	• Vergleich der Erfassungsgrundlagen (gegenseitige Konsistenzprüfung - visuelle Prüfung)		
Eigentliche Arbeit	• Erfassung der Bereiche der Haltestellen (Intermodale Bereiche P&R,B&R,Taxi, Fußgängerbereiche und –steige, Umsteigezeiten zwischen neuen und alten Bereichen)	43,5%	26
<i>Brutto-Erfassungsdauer der Haltestellenbearbeitung</i>		<i>100%</i>	<i>60</i>

Aus den beiden Brutto-Zeiten der Erfassung der Modellierungsgrundlagen mit 87h und der Modellierung mit 60h folgt die hochgerechnete Gesamtzeit zur vollständigen Modellierung der ÖV-Haltestellen zu **147h**. Unter Verwendung des Stundensatzes von €70 resultieren daraus die Kosten der ausführenden Arbeiten mit **€10.290**.

3.3.5 Georeferenzierung der Haltestellen und Linien

Die Georeferenzierung der Haltestellen und Linien stellt den eigentlichen Kern der IV/ÖV-Modellierung dar. In diesem Arbeitsschritt werden die vollständig erfassten Haltestellen (unterteilt in Bereiche und Steige) mit dem Straßennetz und dem nacherfassten ÖV-Liniennetz verknüpft. Diese Verknüpfung wird über eine semantische Relation (sachliche Zuordnung) realisiert. Im einfachsten Fall (z. B. den meisten Bushaltestellen) genügt eine Verknüpfung zwischen der ID des Straßenelements und der ID der Haltestelle in Form einer Kanten-Knoten-Relation. Bei komplexen Haltestellen mit mehreren Treppen, Aufzügen, Ebenen und Gleisen kann diese Relation nicht mehr nur zwischen dem Haltestellenobjekt und der Straße gebildet werden. Je nach Anzahl der Zugangsmöglichkeiten zur komplexen Haltestelle sind die entsprechenden Verknüpfungen über die Haltestellenbereiche oder die Steige zu schaffen.

Die im Rahmen der Georeferenzierung von Haltestellen und Linien durchzuführenden Arbeiten lassen sich in die einzelnen Arbeitsschritte aufteilen:

- Georeferenzierung der Schienen- und Bushaltestellen,
- Linienreferenzierung der Schiene- und Busverbindungen

3.3.5.1 Georeferenzierung der Schienen- und Bushaltestellen

Zum Zeitpunkt der Überlagerung modellierter Haltestellendaten (Schiene) mit den Verkehrsnetzen existiert außer der Hauptkoordinate, die lediglich zur Platzierung der Haltestelle in der Karte dient, noch keine Georeferenz für die einzelnen Bereiche und Steige. Sie „hängen“ in Form einer aufgereihten Perlenkette an der Haltestelle. Durch manuelle Auswahl und Verschieben der Bereiche und Steige wird ihnen jeweils eine Koordinate zugeordnet. Durch z. B. Verschieben eines Steiges auf ein Straßenelement wird die Koordinate des gerechneten Lotfußpunktes dem Steig zugeordnet und als semantische Relation abgelegt. Die Graphiken in Abbildung 3.13 und Abbildung 3.14 wurden aus dem Geoinformationssystem DIVA-Geo der Fa.MentzDV entnommen, welches beim VVS zur umfassenden Pflege der Haltestellendaten als Bearbeitungssoftware der internetbasierten elektronischen Fahrplanauskunft (EFA) eingesetzt wird.

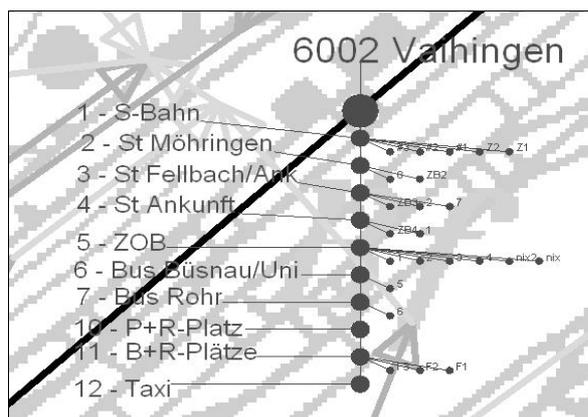


Abbildung 3.13: Haltestellenbereich mit Steigen vor der Georeferenzierung

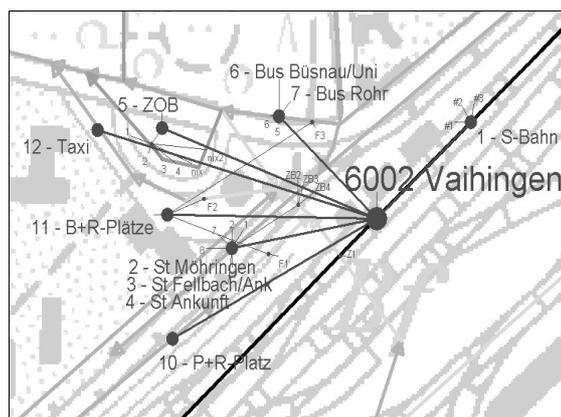


Abbildung 3.14: Haltestellenbereich mit Steigen nach der Georeferenzierung

3.3.5.2 Georeferenzierung der Bushaltestellen

Die Georeferenzierung der Bushaltestellen kann im Gegensatz zu den Schienenhaltestellen mit Hilfe der Haltestellenhauptkoordinate weitestgehend automatisiert ablaufen. Dabei wird von der Koordinate der Bushaltestelle aus, das nächstgelegene Straßenelement berechnet und dort die Bushaltestelle positioniert. Verbleibende fehlerhafte Positionierungen auf Grund von Mehrdeutigkeiten der Zuweisung werden anschließend, im Rahmen der Linienreferenzierung der Bushaltestellen überprüft und berichtigt.

Eine Besonderheit bei der Georeferenzierung der Bushaltestellen existiert bei speziellen Kreuzungssituationen (z. B. Kreisverkehr). Abbildung 3.15 zeigt einen Kreisverkehr und die vier existierenden Haltepunkte. Diese Haltepunkte sind als Steige modelliert und zählen zusammen zu einer Haltestelle (ohne Bereich). Neben der automatisierten Positionierung der Bushaltestelle mittels Hauptkoordinaten muß der Bearbeiter den Steigen die geographisch richtige Position der Haltepunkte zuweisen.

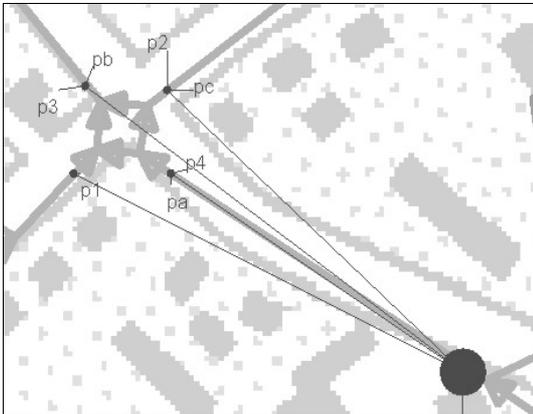


Abbildung 3.15: Ausnahmefall Bushaltestelle im Kreuzungsbereich

Die Häufigkeit dieses Ausnahmefalls ist sehr gering und wird gegenüber der verbleibenden Arbeit der Haltestellenarbeit allerdings vernachlässigt.

In Tabelle 3-11 sind die modellierten Schienen- und Bushaltestellen entsprechend ihrer Häufigkeit und Zugehörigkeit zur Anzahl der existierenden Bereiche unterteilt. Es existieren demzufolge Schienenhaltestellen mit bis zu 10 und Bushaltestellen mit bis zu 4 Haltestellenbereichen. Während sich ein fast gleiches Verhältnis zwischen dem Haltestellenobjekt *Bus* und der Gesamtzahl der Bushaltestellenbereiche ergibt, verhalten sich die Haltestellenobjekte *Schiene* zu ihrer Gesamtanzahl an Haltestellenbereiche im Verhältnis von 1:3.

Tabelle 3-11: Klassifikation der Schienen und Bushaltestellen

Klassifikation Schienenhaltestellen			Anzahl neuer Bereiche
Bereiche	Häufigkeit	Verteilung %	
0	108	32,0	108
1	25	7,4	25
2	50	14,8	100
3	39	11,6	117
4	42	12,5	168
5	21	6,2	105
6	20	5,9	120
7	12	3,6	84
8	6	1,8	48
9	10	3,0	90
10	4	1,2	40
Summe	337	100	1005

Klassifikation Bushaltestellen			Anzahl neuer Bereiche
Bereiche	Häufigkeit	Verteilung %	
0	2752	88,7	2752
1	304	9,8	304
2	38	1,2	76
3	6	0,2	18
4	3	0,1	12
Summe	3103	100	3162

Daraus lässt sich der Arbeitsaufwand zur Georeferenzierung der Haltestellen für die intermodale Routensuche quantifizieren. Es müssen insgesamt 4167 Haltestellenbereiche (1005 Schiene, 3162 Bus) geographisch positioniert und Relationen zu Straßen- und Schienelementen erstellt werden. Darüber hinaus ist die geographische Lage aller 3440 Haltestellenobjekte zu überprüfen und ggf. zu verändern. Die im Rahmen mehrerer Experteninterviews durchgeführten Befragungen des zuständigen Bearbeiters und die Auswertung seiner Arbeitszeitprotokolle ergab einen Netto-Zeitaufwand zur Georeferenzierung aller Bus- und Schienenhaltestellen des VVS-Gebiets von ca. 70 Arbeitsstunden. Anders als die Georeferenzierung der Haltestellen kann die Linienreferenzierung der Schienenverbindungen weitestgehend automatisiert ablaufen. Die Schienenhaltestellen werden, ähnlich wie die Bushaltestellen, über ihre Haltestellenhauptkoordinate auf die vorhandene Schienengeometrie (siehe Erfassung der ÖV-Linien) positioniert. Anschließend kann mit Hilfe des digitalen Streckenfahrplanes die Attributwertzuweisung zu den zwischen den Haltestellen liegenden Teilstrecken stattfinden. Dabei wird den Teilstrecken die entsprechende Liniennummer des ÖV-Verkehrsmittels fahrtrichtungsspezifisch

zugeordnet (z. B. U2 Hin- und Rückweg). Für die Bushaltestellen kann diese Linienreferenzierung allerdings nicht automatisiert erfolgen. Sie findet in drei aufeinander folgenden halbautomatischen Arbeitsschritten statt. Nach der Georeferenzierung der einzelnen Haltestellenobjekte (Bus) und Überprüfung der geometrisch richtigen Position auf den Straßenkanten (Schritt 1; Abbildung 3.16) erfolgt im zweiten Schritt die automatische Verbindung der Haltestellen, entsprechend ihrer sequentiellen Reihenfolge, über die Informationen aus den Streckenfahrplänen (vgl. Abbildung 3.16). Diese Verbindung geschieht automatisch (Schritt 2; Abbildung 3.16) vgl. [STARK, 2003].



Abbildung 3.16: Ablauf der Linienreferenzierung von Bushaltestellen

Im dritten Schritt werden die Teilstücke zwischen den einzelnen Haltestellen durch die Interaktion des Bearbeiters dem Straßennetz zugeordnet. Dabei wird der Bearbeiter weitestgehend durch eine Objektfangfunktion des GIS unterstützt, in dem er die Teilstrecken zwischen den Haltestellen näherungsweise in Richtung der Straßenführung verschiebt. Der Arbeitsaufwand dieser Linienreferenzierung ergibt sich näherungsweise aus der Gesamtstrecke von 3.700 km und der Anzahl 3103 Bushaltestellen. Es existieren folglich rund 3110 Verbindungen zwischen zwei Bushaltestellen mit der durchschnittlichen Länge von ca. 1,2 km. Die Netto-Arbeitszeit der Linienreferenzierung betrug 185 Arbeitsstunden. Damit ergibt sich folgende Zusammenstellung der Netto-Arbeitszeiten:

• Georeferenzierung aller Haltestellen, Bereiche und Steige:	70h
• Linienreferenzierung von Buslinien:	185h
Gesamte Netto-Arbeitszeit der vollständigen Georeferenzierung:	255h

Nach Übergang von der Netto- auf die Bruttoarbeitszeit mit dem Faktor 130% ergeben sich die in dargestellten Zeitan- teile und Bearbeitungszeiten der notwendigen Arbeitsvorbereitung.

Tabelle 3-12: Arbeitsschritte mit Zeitanteil der vollständigen Georeferenzierung

Arbeitsschritte:		Zeitanteil	Zeit (h)
Arbeits- vorbe- reitung	• Orientierung und Auswahl zu bearbeitender Haltestellen	56,5%	332
	• Starten und Kontrolle der automatischen Linienreferenzierung		
	• Starten und Kontrolle automatischer Verbindungen der Bushaltestellen		
Eigent- liche Arbeit	• Georeferenzierung der Haltestellen	43,5%	255
	• Linienreferenzierung von Buslinien		
Brutto-Arbeitszeit der gesamten Georeferenzierung		100%	587

Die Kosten der ausgeführten Arbeiten der gesamten Georeferenzierung ergeben sich daraufhin mit dem Stundensatz von €70 zu €41.055.

3.3.6 Überprüfung der Bearbeitung

Eine Kontrolle der IV/ÖV-Modellierung ist zur Reduzierung der Fehlerhäufigkeit von Routenberechnungen zwingend notwendig, da im Falle einer fehlerhaften Systemauskunft die Reisekette nicht realisiert werden kann, was die Akzeptanz der Nutzer stark beeinträchtigt und im schlimmsten Fall zu keiner erneuten Inanspruchnahme des Dienstes führt. Eine fehlerhafte Systemauskunft liegt z. B. dann vor, wenn durch eine fehlende Verknüpfung von Fußwegen mit dem Straßennetz die Routenberechnung von Startkante zur Umsteigehaltestelle nicht den direkten Fußweg berechnet, sondern einen anderen Weg ausgibt, welcher tatsächlich oder nach dem subjektiven Empfinden des Kunden länger als eine mögliche Alternative ist. Eine ebenso falsche Auskunft wird generiert, wenn Haltestellenzugänge nicht dem richtigen Straßenelement zugewiesen wurden. Daraus resultieren falsch berechnete Umsteigewege, die dazu führen können, dass die gewünschten Anschlussverbindungen nicht erreicht werden.

Die Überprüfung der korrekten Nachbearbeitung der Straßenkarte und korrekten Erfassung der Haltestellenbereiche geschah mittels Testläufe des Routenplaners. In der Umgebung jeder Schienenhaltestellen sind 3 Startkanten (Adressen) zufällig ausgewählt worden, um die daraufhin ermittelten Verbindungen des Routenplaners manuell auf Richtigkeit (korrekte Fußwege, Umsteigewege) und Plausibilität (Umsteigezeiten) zu überprüfen und bei auftretenden Fehlern zu korrigieren. Diese Überprüfung fand für alle 337 Schienenhaltestellen des VVS-Gebiets statt. Von den insgesamt 3103 vorhandenen Bushaltestellen wurden als Stichprobe 30% der Haltestellen verwendet. Dabei wurde darauf geachtet, dass die Stichprobenauswahl alle diejenigen Bushaltestelle, die im Einzugsbereich von größeren Wohngebieten liegen, beinhaltet.

Bei diesen Überprüfungsläufen des Routenplaners waren individuelle, der Situation angepasste, Einstellungen der Verkehrsmittelwahl zu treffen. So mussten z. B. für die Überprüfung der richtigen Wegeberechnung von Startkante zu den U-Bahnstationen alle Umsteigemöglichkeiten in andere Verkehrsmittel unterbunden werden. Gerade bei komplexen Haltestellen mit mehreren Haltestellenbereiche, waren diejenigen Verkehrsmittel auszublenden, die die zu überprüfende Umsteigebeziehung möglicherweise verfälschen könnten. Der zeitliche Überprüfungsaufwand war bei dieser Arbeitstätigkeit nicht aus den Zeitprotokollen zu entnehmen. Die durchschnittliche Brutto-Arbeitszeit von 30 min pro zu überprüfender Haltestelle entstammt aus einer Experteneinschätzung des Bearbeiters des VVS. Rechnet man den gesamten Bruttozeitaufwand dieser Überprüfung auf der Grundlage dieser Einschätzung aus, dann erhält man eine Arbeitszeit zur Kontrolle der richtigen Nachbearbeitung der Straßenkarte und Modellierung der Haltestellen von ca. **634 Stunden**. Daraus abgeleitet ergeben sich die Überprüfungskosten der ausgeführten Arbeiten (30%ige Stichprobe) mit dem Stundensatz von € 70,00 zu ca. **€ 44.380**. Für eine vollständige Überprüfung aller Haltestelle ergeben sich die Kosten entsprechend zu ca. **€ 120.400**.

3.4 Kostenanteile zur Realisierung des Intermodalen Routenplaners

Zu den Realisierungskosten des Intermodalen Routenplaners zählen neben den Kosten der ausgeführten Arbeiten der IV/ÖV-Modellierung auch die Einkaufs- und Lizenzierungskosten der digitalen Straßenkarte sowie die Anschaffungskosten für die benötigten Hard- und Softwarekomponenten.

Kosten der digitalen Straßenkarte: Um im Rahmen dieser Aufwandsabschätzung zu belastbaren Zahlen zu gelangen, wurden Angebote der beiden Hersteller von digitalen Straßendaten (NavTech und TeleAtlas) eingeholt [STARK, 2001a]. Die Datenbereitstellung geschieht in der Regel bundeslandbezogen. Eine kleinräumigere Gebieteinteilung z. B. für die Region Stuttgart kann nicht direkt über die Hersteller bezogen werden, sondern ist über Partnerfirmen abzuwickeln (z. B. PTV für NavTech). Da bei einer Entscheidung, den gesamten Datenbestand eines Bundeslandes direkt beim Hersteller einzukaufen oder ihn gebietsbezogen durch Partnerfirmen zu beziehen, auch die Größe und inhaltliche Dichte (Anteil Objekte am Gesamtbestand) des zu selektierenden räumlichen Gebietes sowie die Selektionskosten der Partnerfirma eine maßgebliche Rolle spielt, wurden hier die Kosten von € 38.000 des Herstellers als Kostenfaktor weiterverwendet. Ein signifikanter Preisunterschied zwischen beiden Möglichkeiten war laut Auskunft der am Aufbau des intermodalen Routenplaners beteiligten Partner nicht vorhanden.

Beide Hersteller machen ihre Preise und Lizenzgebühren von dem zugrunde gelegten Geschäftsmodell der Dienstleistung abhängig. Da der hier betrachtete Intermodale Routenplaner im Sinne eines Demonstrators der technischen Machbarkeit und verkehrlichen Wirksamkeit im Rahmen des Forschungsprojekts MOBILIST aufgebaut wurde, lässt sich das künftige Geschäftsmodell nur sehr schwer definieren. Grundsätzlich kann aber bei beiden Herstellern zwischen zwei Abgabevarianten der digitalen Straßenkarte unterschieden werden.

1. **Variante: Daten zur ausschließlichen internen Nutzung:** Bei dieser Variante darf der Datenbestand an maximal 5 Arbeitsplätzen des Unternehmens genutzt werden. Es sind Analysen und Berechnungen (z. B. zur Routenplanung, Tourenplanung und zur Flottenlogistik) sowie die Weitergabe der Analyse- und Berechnungsergebnisse an Dritte erlaubt. Ein Weiterverkauf der Daten im Original oder als veredelter Datenbestand ist bei dieser Variante nicht gestattet. Die Kosten sind als einmalige Bereitstellungs- und Lizenzgebühren in einer Höhe von ca. € 38.000,00 für das Bundesland Baden-Württemberg zu entrichten.
2. **Variante: Daten zur Online-Nutzung:** Bei dieser Variante kann der Datenbestand zentral auf einem Server vorgehalten werden, auf den die unterschiedlichen Anwendungen zugreifen können. So können z. B. auch Internet-Routenplaner mit diesen Daten versorgt werden oder Teilausschnitte von Daten als aufbereitete Karten weitergegeben werden. Zusätzlich zu einer Mindestgebühr von € 30.000 werden die Kosten transaktionsgebunden abgerechnet. Je nach Anzahl an Zugriffen werden zusätzliche Kosten von ca. € 1.670 bis zu € 104.000 in Rechnung gestellt (vgl. Tabelle 3-13).

Tabelle 3-13: zugriffsabhängige Lizenzgebühr nach Gespräch mit Kartenhersteller, 1999.

Zugriffe	Kartenpräsentationen		Karte+Route		
	manuell	automatisch	manuell	automatisch	Dynamisch
50.000	1.667 €	2.167 €	3.333 €	4.333 €	6.667 €
250.000	3.067 €	3.978 €	6.183 €	7.973 €	12.267 €
100.000.000	104.067 €	135.287 €	208.133 €	270.573 €	416.267 €

Da im Rahmen dieses Forschungsprojektes weder das Geschäftsmodell klar zu definieren ist noch die Anzahl der Zugriffe über potentielle Nutzer bekannt ist, werden zur Abschätzung aller Kostenanteile die Kosten der digitalen Straßenkarte für Baden-Württemberg mit € 38.000 angesetzt. Dabei handelt es sich um die Kostenzusammensetzung gemäß Variante 2 unter Annahme dynamischer Kartenpräsentation und angenommenen 500.000 jährlichen Zugriffen.

Kosten der Hard- und Software: Da zur Einrichtung des Intermodalen Routenplaners insgesamt 4 vernetzte, handelsübliche Standard-Desktop-PCs verwendet wurden, betragen die Anschaffungskosten inkl. Betriebssystem, Netzwerkkarten und –software nach marktüblichen Preise ca. € 15.000.

Berücksichtigt man darüber hinaus die Einrichtung und Pflege der Steuerungssoftware² mit ca. € 10.000, dann resultieren daraus die Hard- und Softwarekosten zur Einrichtung des intermodalen Routenplaners rund € 25.000. Es ergeben sich die folgenden Kostenanteile (auf volle 100 € gerundet) und der überprüfungsabhängige Gesamtaufwand:

Tabelle 3-14: Kostenanteile der Realisierung des Intermodalen Geodatenbestandes

Prozess	Kosten	ohne Prüfung	30% Prüfung	100% Prüfung
1.) Einkauf digitale Straßenkarte	€ 38.000,00	X	X	X
2.) Hard- und Software	€ 25.000,00	X	X	X
3.) Nachbearbeitung der Straßenkarte	€ 51.100,00	X	X	X
4.) Erfassung des ÖV-Netzes	€ 5.100,00	X	X	X
5.) Modellierung und Erfassung der Haltestellen	€ 10.300,00	X	X	X
6.) Georeferenzierung	€ 41.100,00	X	X	X
7a.) bei 30% iger Überprüfung der Bearbeitung	€ 44.400,00	-	X	-
7b.) bei 100%iger Überprüfung der Bearbeitung	€ 120.400,00	-	-	X
<i>Gesamtaufwand</i>		<i>€ 170.600,00</i>	<i>€ 215.000,00</i>	<i>€ 291.000,00</i>

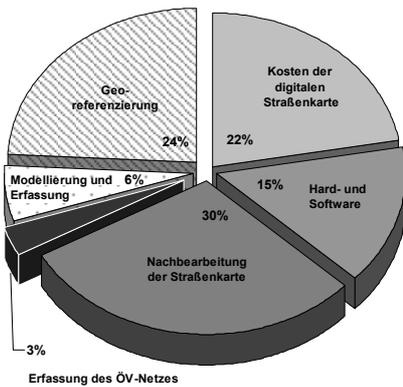


Abbildung 3.17 Kostenanteile ohne Prüfung

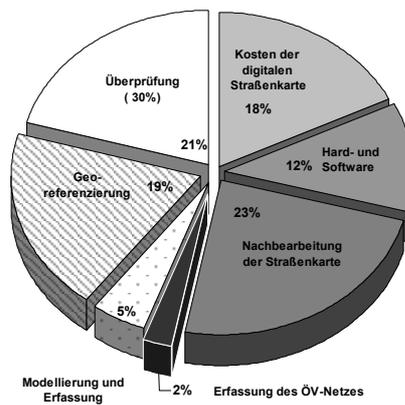


Abbildung 3.18: Kostenanteile mit 30% iger Prüfung

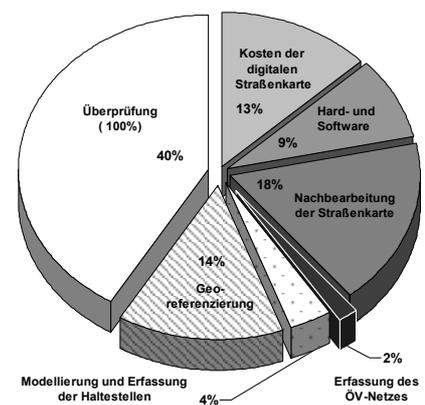


Abbildung 3.19: Kostenanteile mit 100% iger Prüfung

Vergleicht man die Kostenanteile der Realisierung des Intermodalen Geodatenbestands entsprechend der Art der Überprüfung so kann folgendes festgehalten werden:

- Ein vollständig überprüfter Datenbestand ist hierbei nahezu um 2/3 teurer in seiner Realisierung als ein ungeprüfter Datenbestand.
- Selbst bei nur 30% iger Prüfung ist dies ein erheblicher Kostenfaktor
- Eine 100% Prüfung kann die reinen Realisierungskosten erreichen und wie hier dargestellt sogar übersteigen.

Damit kann festgehalten werden, dass der Kostenanteil von Qualitätssicherungsmaßnahmen einen erheblichen Einfluss auf die Gesamtkosten der Realisierung ausübt und es deshalb für Informationsdienstleister von Vorteil ist diesen Einfluss möglichst exakt prognostizieren zu können. Bevor jedoch auf diese Prognosemöglichkeit eingegangen wird, betrachten die folgenden Abschnitte den Kundennutzen der Intermodalen Routenberechnung sowie die Möglichkeiten eines damit verbundenen „Return On Investments“.

² Software im Sinne von kleineren modularen Steuerprogrammen, PHP/CGI-Skripten

3.5 Abschätzung des Kundennutzens

Der wirtschaftliche Erfolg einer Informationsdienstleistung ist neben der qualitativ hochwertigen technischen Realisierung maßgeblich vom nachweisbaren Nutzen für den Kunden und dessen Zahlungsbereitschaft abhängig. Der Kunde muss den für ihn entstehenden Mehrwert der Informationsbereitstellung durch eine solche Dienstleistung klar erkennen und auch bereit sein, eigene Gewohnheiten der Verkehrsmittelwahl zu ändern. Das hierfür erfolgreichste Instrument einer Dienstleistung ist die transparente Darstellung resultierender individueller Vorteile wie z. B. Kosten- und Zeiterparnis oder auch den Zugewinn an Beförderungskomfort. Daraus entsteht die Verantwortung des Geschäftsplans, neben der reinen Bereitstellung von richtigen Umsteigeinformationen, auch die auf den Nutzer individuell zugeschnittene und transparente Darstellung seines Kosten- und Zeitvorteils aufzubereiten.

Während der Zeitvorteil über die Berechnung durchschnittlicher IV-Reisezeiten auf dem Straßennetz und den Fahrplaninformationen der ÖV-Linien vergleichend ermittelt werden kann, müssen verschiedene Kostenfaktoren, die dem Nutzer entlang der Intermodalen Route entstehen, aufbereitet werden. Zu den maßgeblichen Kostenfaktoren, welche die Kundenentscheidung beeinflussen, zählen:

Tabelle 3-15: Entscheidungsbeeinflussende Kostenfaktoren des Kunden

Kostensteigernd:	Kostensenkend:
<ul style="list-style-type: none"> • P&R-Kosten ($K_{P\&R}$) • Fahrscheinkosten (K_{Fk}) 	<ul style="list-style-type: none"> • Transaktionskostensparnis (K_{taK}) • Betriebsmittlersparnis (Kraftstoffe) (K_{Bm})

Die kostensteigernden Faktoren lassen sich durch entsprechende administrative Maßnahmen senken, um eine Kostentransparenz herzustellen. So existiert bereits heute in mehreren Ballungszentren die Vereinbarung, dass ein gelöster Parkschein gleichzeitig als Fahrschein genutzt werden kann. Denkbar wäre es auch, den Kunden der intermodalen Dienstleistung Preisnachlässe beim Erwerb von Zeitfahrberechtigungen (Wochen-, Monats- oder Jahreskarten) einzuräumen. Die Betriebsmittlersparnis des Kunden kann über Verbrauchswerte und Fahrleistungen transparent dargestellt werden. Aus diesen kostensteigernden und kostensenkenden Einflussfaktoren lassen sich zwei Grundbedingungen ableiten, die erfüllt sein müssen, damit der Kunde den Dienst in Anspruch nimmt:

$$K_{taK} + K_{Bm} - (K_{P\&R} + K_{Fk}) > 0$$

Die Ersparnisse durch die Dienstleistung müssen größer sein als die fixen Kosten

$$K_{taK} + K_{Bm} - (K_{P\&R} + K_{Fk}) - K_N \geq 0$$

Die Einsparungen durch die Dienstleistung müssen größer oder mindestens genauso groß sein wie die Summe der fixen Kosten und der Nutzungsgebühr

Bei den bisher erläuterten kostensteigernden und –senkenden Faktoren bedarf es zur transparenten Darstellung lediglich einer entsprechend aufbereiteten und nachvollziehbaren Darstellung. Für den Kunden wesentlich weniger nachvollziehbar sind die Ersparnisse, die ihm durch die Senkung der Transaktionskosten entstehen. Aus diesem Grund wird in diesem Abschnitt der Effekt der Transaktionskostensenkung durch den Intermodalen Routenplaner dargestellt. Die Theorie der Transaktionskosten geht dabei auf die Arbeiten von [COARSE, 1991] zurück.

Definition 3-1: Transaktion nach Coarse [JOST, 2001]

Als Transaktion wird der Austausch von Gütern zwischen zwei Akteuren (Transaktionspartner) des wirtschaftlichen Lebens (Wirtschaftssubjekte) bezeichnet. Die Partner sind nicht notwendigerweise Individuen, sondern können auch andere ökonomische Akteure sein. Unter Gütern wird alles das verstanden, was in irgendeiner Weise dafür geeignet ist, für einen Transaktionspartner einen Nutzen zu stiften.

Daraus abgeleitet ergeben sich die Transaktionskosten mit folgender Definition:

Definition 3-2: Transaktionskosten [JOST, 2001]

Die Transaktionskosten werden allgemein als Kosten zur Durchführung einer Transaktion definiert. Für eine Unternehmung umfassen Transaktionskosten also die Kosten für die Einrichtung, die Erhaltung und die Veränderung der Organisation sowie die laufenden Kosten, die mit dem Betrieb der Unternehmung verbunden sind. Sie resultieren unmittelbar aus der Art und Weise, wie Arbeitsteilung und Kooperation in einer Unternehmung gestaltet sind.

Die Transaktionskosten lassen sich vor allem als Such- und Informationskosten sowie Kosten zur Entscheidungsfindung bezeichnen. Je geringer die Transaktionskosten innerhalb eines Unternehmens sind, desto schneller kann es wachsen. Die Senkung der Transaktionskosten ist im Kern nichts anderes als die Reduzierung von Eigentumsprämien bei der Erstellung von Produkten, Dienstleistungen oder dem Verkauf von Wissen. Die Grenze für eine Unternehmensgröße ist an dem Punkt erreicht, an dem die Transaktionskosten gleich den Kosten im Markt werden. Da in vielen Firmen die Transaktionskosten innerhalb der Firma höher liegen als auf den Märkten, werden Aktivitäten ausgelagert [SCHMIDT, 1999].

Der intermodale Routenplaner reduziert die Widerstände an den Übergangsstellen der Verkehrsnetze des IV und ÖV durch Umsteigeinformationen. Einer dieser Widerstände, der seither den Verkehrsmittelwechsel beeinflusst hat, ist der extrem hohe Aufwand, eine kombinierte, intermodale Reisekette mit herkömmlichen Mitteln, wie z. B. Stadtplan, Verkehrslinienplan, Fahrplan zusammen zu stellen. Es scheiterte oftmals bereits an der Nicht-Verfügbarkeit einer dieser Hilfsmittel.

Daraus lässt sich schlussfolgern, dass der intermodale Routenplaner neben der Bereitstellung der IV/ÖV-Anschlussinformationen die Such- und Informationszeit reduziert, die nötig wäre, eine intermodale Route mit herkömmlichen Mitteln zusammen zu stellen. Durch einen Vergleich der benötigten Zeit zur Ermittlung der intermodalen Route mit und ohne Verwendung des intermodalen Routenplaners kann die Zeitersparnis, also die Reduktion der Transaktionskosten, angegeben werden.

Dieser Vergleich wurde mit Testpersonen durchgeführt, deren Aufgabe es war, einmal mit den angegebenen Hilfsmitteln eine intermodale Route von Ihrem Wohnort aus zu einem Zielort manuell zusammenzustellen (manueller Fall) und anschließend dieselbe Verbindung mit dem Intermodalen Routenplaner berechnen zu lassen (automatischer Fall). Für beide Verfahren wurde der zeitliche Aufwand ermittelt. Folgende Tabelle 3-16 zeigt die Rahmenbedingungen des Testszenarios.

Tabelle 3-16: Rahmenbedingungen des Testszenarios

<i>Testpersonen</i>	10
<i>Start der Route</i>	Wohnort der Testperson
<i>Ziel der Route</i>	Esslingen a.N. - Berkheim, Rainweg
<i>Hilfsmittel:</i>	Stadtplan, Verkehrslinienplan, Fahrplan
<i>Bemerkung:</i>	Die Verwendung der Elektronischen Fahrplanauskunft konnte zur reinen Fahrzeitermittlung verwendet werden. Ein automatische Berechnung des Verkehrsmittelwechsel im ÖV hingegen wurde ausgeschlossen.

Bei der Auswahl der Testpersonen wurde darauf geachtet, dass deren Wohnort entweder im Norden oder im Zentrum der Region Stuttgart lag. Diese Festlegung war nötig, da nur so gewährleistet werden konnte, dass eine Strecke zurückgelegt werden musste, die lange genug für einen effektiven Verkehrsmittelwechsel war. Je kürzer die Strecke, um so benachteiligter ist der ÖV gegenüber dem IV bezüglich der ermittelten Reisezeit.

Als weitere Vorgabe des manuellen Falls wurde festgelegt, dass der Umsteigepunkt bei ca. 50% der Wegstrecke liegen sollte. Als Orientierung hierfür diente die Luftlinienverbindung zwischen Start- und Zielpunkt. Zur Auswertung dieser Tests wurde ein Zeitprotokoll von jeder Testperson geführt, welches die Schritte der Zusammenstellung der intermodalen Route enthält. Diese Schritte sind:

- Identifikation des Startortes in der Karte,
- Identifikation des Zielortes in der Karte,
- Auswahl der ÖV-Haltestelle in der Nähe des Zielortes mit Fußwegzeitschätzung,
- Auswahl des P&R-Platzes und Identifikation der ÖV-Linien zum Ziel,
- Zusammenstellung der Einzelzeiten der Reisekette.

Folgende Tabelle zeigt den Zeitbedarf beider Fälle der 10 Testpersonen:

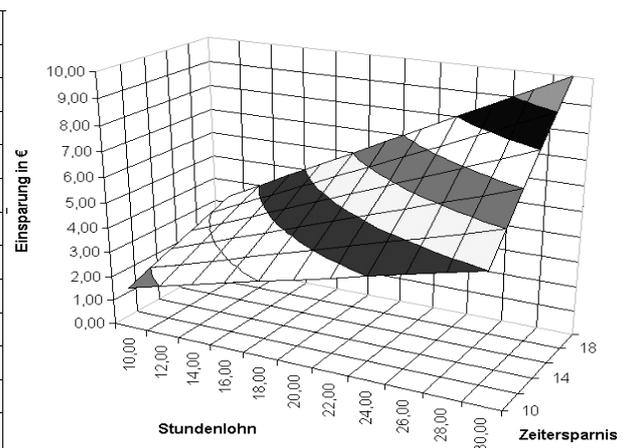
Tabelle 3-17: Zeitbedarf der Testpersonen für den manuellen und den automatischen Fall

Aktion	Zeitbedarf der Testpersonen in [min]									
	A	J	C	H	G	I	B	D	F	E
Ident. Start	2,0	1,0	1,0	1,0	0,5	0,5	0,5	0,5	0,5	1,0
Ident. Ziel	1,5	2,5	2,0	2,0	1,0	1,0	2,0	2,5	0,5	0,5
ÖV-Halt am Ziel	1,0	1,5	1,0	1,0	1,0	2,0	1,0	0,5	1,0	1,0
P&R und ÖV-Linie	2,0	2,0	1,5	2,0	1,5	1,0	1,0	2,0	1,5	1,5
Reisekette	10,0	9,5	10,0	8,0	7,5	7,5	6,5	5,0	6,5	5,5
Addieren Gesamtzeit	1,0	0,5	1,0	1,0	1,0	0,5	1,0	1,0	0,5	0,5
Zeitbedarf (manueller Fall)	17,5	17,0	16,5	15,0	12,5	12,5	12,0	11,5	10,5	10,0
Zeitbedarf (automatischer Fall)	1	1,5	1	2	1,5	1	2,5	1,5	1	1

Der Zeitbedarf der Testpersonen beim manuellen Fall schwankt zwischen 10 und 17,5 min, beim automatischen Fall zwischen 1 und 2,5 min. Da die Zeitschwankungen beim automatischen Fall stark von den Zugriffszeiten der Internetverbindung beeinflusst sind und bei den Testpersonen ein überdurchschnittliches Bildungsniveau vorausgesetzt werden kann, welches sich durch den schnelleren weil gewohnteren Umgang mit dem Internet widerspiegelt, kann für den Bevölkerungsdurchschnitt der Region Stuttgart eine höhere Aufwandsreduktion angenommen werden. Ohne entsprechende empirische Umfrage und Analyse wird von einer Aufwandsreduktion bei der automatischen Berechnung einer intermodalen Route von 10 bis 20 min ausgegangen.

Tabelle 3-18: lohnabhängige Monetarisierung der Zeitersparnis

Stundenlohn €	Ersparnisse pro Minute					
	10	12	14	16	18	20
10,00	1,67	2,00	2,33	2,67	3,00	3,33
12,00	2,00	2,40	2,80	3,20	3,60	4,00
14,00	2,33	2,80	3,27	3,73	4,20	4,67
16,00	2,67	3,20	3,73	4,27	4,80	5,33
18,00	3,00	3,60	4,20	4,80	5,40	6,00
20,00	3,33	4,00	4,67	5,33	6,00	6,67
22,00	3,67	4,40	5,13	5,87	6,60	7,33
24,00	4,00	4,80	5,60	6,40	7,20	8,00
26,00	4,33	5,20	6,07	6,93	7,80	8,67
28,00	4,67	5,60	6,53	7,47	8,40	9,33
30,00	5,00	6,00	7,00	8,00	9,00	10,00



Die Ableitung der Transaktionskostensenkung aus den Werten der Zeitersparnis kann nur über einen geldwerten Bemessungsfaktor der Zeit monetarisiert werden. Das bedeutet, dass die Monetarisierung der Zeitersparnis nicht für alle Situationen des täglichen Lebens und für alle Nutzergruppen einheitlich anzugeben ist. Die Nutzerprofile in verschiedenen Situationen lassen eine unterschiedliche Interpretation des individuellen Nutzens von Zeitersparnissen zu.

Tabelle 3-18 zeigt die monetarisierte Zeitersparnis durch den intermodalen Routenplaner in Abhängigkeit des Stundenlohns von 10 bis 30€. Aus der nebenstehenden Visualisierung des linearen Kostenverhältnisses kann schnell der Kostenvorteil in Abhängigkeit des Stundenlohns und der Zeitersparnis abgelesen werden.

3.6 Reinvestition des Projektaufwands durch Nutzungsgebühren

Für den wirtschaftlichen Betrieb eines intermodalen Routenplaners als Internetdienstleistung ist neben der transparenten Darstellung des Kundennutzens auch die Wiedergewinnung der Investition, der **Return on Investment (ROI)** entscheidend. Die Refinanzierungsmöglichkeiten von Informationsdienstleistungen lassen sich nach folgenden Gesichtspunkten unterscheiden:

1. Die Dienstleistung liegt zum überwiegenden Teil im öffentlichen Interesse. Ihre (Re-)finanzierung erfolgt über die öffentliche Hand. (z. B. Stadtplan als interaktive Karte im Internet).
2. Die Informationen werden als Serviceleistungen eines Unternehmens an die Kunden weitergegeben. Ihre Refinanzierung kann über eine anderweitige Erhöhung des Geldmittelverkehrs zwischen Unternehmen und Kunden kompensiert werden (z. B. tagesaktuelle kundenspezifische Börsenanalysen).
3. Die Informationen werden als Serviceleistungen mit eingeschränktem Funktionsumfang zur Neukundengewinnung kostenlos angeboten. Die Refinanzierung erfolgt über gewonnene Neukunden (z. B. Routenplaner mit eingeschränkter Nutzung – nur straßenscharfes, nicht adressscharfes Routen möglich).
4. Die Informationen werden als Serviceleistungen kostenlos zur Verfügung gestellt. Ihre Refinanzierung geschieht über Werbemaßnahmen.
5. Die Informationsdienstleistungen werden dem Kunden in Rechnung gestellt (z. B. Nutzungsgebühren des Deutschen Wetterdienstes).

Neben diesen dargestellten Möglichkeiten zur Finanzierung einer Informationsdienstleistung sind darüber hinaus auch Kombinationen dieser Möglichkeiten denkbar. Grundsätzlich ist der ROI von mehreren Faktoren abhängig. Diese können sein:

- Höhe des Investitionsvolumens der Systemeinführung
- Geplanter Nutzungszeitraum in Jahren
- Abzinsungsfaktor zur Ermittlung des Barwerts am Ende der Nutzungsperiode (wenn vorhanden).
- Nutzungshäufigkeit pro Periode
- Zeitliche Entwicklung von Nutzungsgebühren
- Abschreibungen der Hard- und Software, Wartungskosten sowie Personalkosten der DV-Betreuung.

In diesem Abschnitt soll deshalb abgeschätzt werden, nach wie viel Jahren mit welcher Nutzerzahl und Nutzungsgebühren der ROI für den Intermodalen Routenplaner zu erreichen ist. Dazu wird ein geplanter Nutzungszeitraum von 10 Jahren und ein Barwert am Ende des 10-ten Nutzungsjahres von Null angenommen. Darüber hinaus muss mit einem jährlichen Pflegeaufwand der digitalen Geodatengrundlage von ca. 5-10% (vgl. [BEHR, 1998]) des Investitionsvolumens aufgrund infrastruktureller Änderungen gerechnet werden. Speziell bei amtlichen Geodaten kann sich der Aufwand auch zwischen 8% und 14% bewegen [STARK, 2001A].

3.6.1 Abschreibung der Investitionen

Den Abschreibungen kommen im Rahmen des betrieblichen Rechnungswesen grundsätzlich die Aufgabe zu, alle Wertminderungen am ruhenden Vermögen einer Unternehmung buchhalterisch zu erfassen. Während die **bilanzielle** (planmäßige) **Abschreibung**, die bestimmten handels- und steuerrechtlichen Bewertungsvorschriften zu entsprechen hat, im Interesse einer periodenadäquaten Aufwandsverrechnung die Aufwands- und Herstellkosten der Anlagegüter auf die Jahre der Nutzung verteilt, stellt die **kalkulatorische Abschreibung** eine Methode zur Erfassung des bewerteten Güterverzehr derjenigen abnutzbaren Wirtschaftsgüter dar, die laufend dem Sachziel der Unternehmung dienen. Sowohl bei der bilanziellen als auch der kalkulatorischen Abschreibung können den Wertminderungen verschiedene Ursachen zugrunde liegen, die in aller Regel nicht exakt feststellbar sind, sondern auf Hypothesen beruhen.

Die kalkulatorische Abschreibung erfassen folgende Ursachen des Verzehr abnutzbarer Anlagegüter [FREIDANK, 1997], [WARNECKE ET AL., 1996]:

- Abnutzungsbedingter Verschleiß
- Substantielle Wertminderung
- Natürlicher Verschleiß (z. B. Verwitterung)
- Technische oder wirtschaftliche Überholung
- Ablauf von Fristen (z. B. Lizenzvereinbarungen, Patentrechte)

Folgt man dieser Argumentation so kann grundsätzlich festgehalten werden, dass Abschreibungen nur bei abnutzbaren Anlagegütern (Produktionsfaktoren, Kapitalinvestitionen, etc.), die einem bestimmten Werteverzehr durch Gebrauch oder veränderten Marktpreisen unterliegen, durchgeführt werden.

Übertragen auf die Herstellung und Pflege von Geodaten lässt sich folgendes festhalten: Geodaten unterliegen bei mangelhafter oder fehlender Aktualisierung bzw. Fortführung einem zeitlichen Alterungsprozess, welcher die Ursache des Werteverzehrs ist. Ein Werteverzehr aufgrund von Abnutzung durch Gebrauch ist bei Geodaten nicht vorhanden. Da jedoch dem Alterungsprozess von Geodaten eben durch Aktualisierungs- und Fortführungsmaßnahmen entgegenge wirkt wird, kann, bei unterstellter vollständiger, aktueller und richtiger Fortführung, kein Werteverzehr festgestellt werden. Damit entspricht im umgekehrten Sinne der monetäre Fortführungsaufwand einer jährlichen Abschreibung mit dem Unterschied, dass dieser Betrag den Wert der Geodaten nicht mindert, sondern zu dessen Werterhaltung beiträgt. Diese Betrachtungsweise trifft für Geodatenbestände mit exklusivem Charakter (z. B. amtliche Datenbestände wie ATKIS, ALK) zu, welche z. B. ganz individuell für eingeschränkte Anwendungsfelder vorgehalten werden. Dieser Werteverzehr kann grob dargestellt werden, in dem die Anzahl der Veränderungsnachweise des Vermessungskatasters (für das Bundesland Baden-Württemberg im Jahr 2000) von 1,1 Mio VN der Anzahl der Flurstücke von 8,9 Mio gegenübergestellt werden. Daraus wird deutlich, dass ohne stetige Fortführung und Aktualisierung der digitale Bestand der ALK in ca. 8 Jahren vollständig veraltet wäre [STARK, 2001a]. Für Geodatenprodukte (z. B. Routenplaner auf CD-ROM, Navigations- - CDs für die Kfz-Navigation), die über einen bestimmten Zeitraum genutzt werden können und die keine inkrementelle Fortführung gestatten, sind Abschreibungen jedoch durchaus zweckmäßig.

Grundsätzlich kann zwischen **linearer**, **degressiver** und **progressiver** sowie der Abschreibung nach **Leistung** und **Inanspruchnahme** unterschieden werden. Da die progressive Abschreibung steuerlich nicht zulässig ist und die Abschreibung nach Leistung und Inanspruchnahme eher im Rahmen der innerbetrieblichen Nachkalkulation verwendet wird, ist nachfolgend nur die lineare und degressive Methode der Abschreibung Gegenstand der Betrachtung.

3.6.2 Lineare Abschreibung

Der **linearen Abschreibung** liegt die Annahme zugrunde, dass der Verzehr abnutzbarer Anlagegüter sich in Abhängigkeit vom Zeitablauf vollzieht und die Abschreibungsbeträge pro Periode (q_t) für die gesamte Nutzungsdauer (T) konstant sind. Bei der Festlegung der planmäßigen Abschreibungsbasis muss, wie auch bei den anderen Verfahren, ein eventuell am Ende der Nutzungsdauer verbleibender Rest- oder Schrottwert (R_T) berücksichtigt werden, da nur der planmäßige Verzehr in die Rechnung mit einfließen soll. Von den Anschaffungs- bzw. Herstellungskosten (AK) der abzuschreibenden Betriebsmittel müssen deshalb die entstehenden Rest- bzw. Schrottwerte abgezogen werden. Für die lineare Abschreibung gelten somit folgende Beziehungen (w_t = Abschreibungsprozentsatz pro Periode) [FREIDANK, 1997]:

$$AK - R_T = \sum_{t=1}^T q_t$$

$$q_t = \frac{AK - R_T}{T}$$

$$w_t = \frac{q_t}{AK - R_T} \cdot 100$$

- Die Summe der periodischen Abschreibungsbeträge entspricht der Differenz von Anschaffungskosten und Restwert.
- Die periodischen Abschreibungsbeträge sind konstant.
- Der periodische Abschreibungsprozentsatz ist konstant.

Die lineare Abschreibung kann Anwendung finden zur Berücksichtigung des natürlichen Verschleißes, der technischen und wirtschaftlichen Überholung, des Fristablaufs und zur Erfassung des Gebrauchsverschleißes.

3.6.3 Degressive Abschreibung

Im Gegensatz zur linearen Abschreibung geht die **degressive Methode** von sinkenden Abschreibungsbeträgen aus, unterstellt aber auch einen nicht linearen Zusammenhang zwischen dem Betriebsmittelverzehr und dem Zeitablauf. Diese Abschreibung findet statt, wenn:

- die Zeitwerte (monetären Beträge) der Wirtschaftsgüter nicht linear, sondern aufgrund technischer oder wirtschaftlicher Überholung in den ersten Jahren schneller als in späteren Jahren fallen.
- die Gebrauchsfähigkeit der Anlagen in den ersten Nutzungsperioden höher ist als in den folgenden und /oder
- in späteren Jahren mit ansteigenden Reparatur- und Instandhaltungsarbeiten gerechnet werden muss.

Die degressive Abschreibung kann ihrerseits in die arithmetisch-degressive und geometrische - degressive Abschreibung unterschieden werden. Während bei der arithmetische - degressiven Methode die jährlichen Abschreibungsbeträge um einen gleichbleibenden Betrag fallen, so sinken bei der geometrisch-degressiven Methode die jährlichen Abschreibungsbeträge um einen kleiner werdenden Betrag (Buchwertverfahren). Da beim Buchwertverfahren immer vom jeweiligen Restbuchwert ausgegangen wird, kann sich für geometrisch-degressiv abgeschriebene Anlagegüter am Ende der Nutzungsdauer niemals ein Restwert von Null ergeben. Diese Möglichkeit wird zur Kostenabschätzung des hier betrachteten intermodalen Routenplaners nicht verwendet, da nicht á priori der Restwert der Hard- und Software bzw. der Restwert des investierten Kapitals zum Geodatenaufbau abgeschätzt werden kann. Für die arithmetisch-degressive Abschreibung kann folgender Zusammenhang angegeben werden:

$$AK - R_T = \sum_{t=1}^T q_t$$

- Die Summe der periodischen Abschreibungsbeträge entspricht der Differenz von Anschaffungskosten und Restwert.

$$q_t = (T - t + 1) \cdot \frac{AK - R_T}{\sum_{t=1}^T t}$$

- Die periodischen Abschreibungsbeträge verringern sich um einen konstanten Betrag pro Periode.

$$w_t = \frac{q_t}{AK - R_T} \cdot 100$$

- Die periodischen Abschreibungsprozente sind nicht konstant.

Entsprechend diesen Methoden der Abschreibungen können die Gesamtinvestitionen des Aufbaus des intermodalen Routenplaners wie folgt durchgeführt werden:

<p>1. Hard- und Softwareinvestitionen sind arithmetisch-degressiv abzuschreiben.</p>	<p>Hier wird der Tatsache Rechnung getragen, dass die Gebrauchsfähigkeit der Hard- und Software in der ersten Nutzungsperiode größer ist als in den folgenden und dass in späteren Jahren mit ansteigenden Reparatur- und Wartungsaufwand zu rechnen ist. Ferner ist der Gesamtzeitraum von 10 Jahren für die Hardware eindeutig zu groß. Deshalb erfolgt nach 5 Jahren eine Neubeschaffung der Hardware, was den Nutzungszeitraum entsprechend halbiert.</p>
<p>2. Kapitalinvestitionen des Geodatenaufbaus werden linear abgeschrieben.</p>	<p>Dies wird unter der Vereinfachung durchgeführt, dass sich die Zinsentwicklung konstant verhält, da ansonsten der Kapitalwert der Investition über die Abschreibung gemäß Leistungs- bzw. Inanspruchnahme geschehen müsste. Bei dieser linearen Abschreibung wird man auch möglichen Zeitabläufen von Lizenz- und Patentrechten gerecht.</p>

Der Werteverfall der Geodaten wird, wie bereits begründet, nicht abgeschrieben. Ihm wird durch einen jährlichen Pflegeaufwand von 5-10% der Anfangsinvestition Rechnung getragen.

Folgende Tabelle zeigt die kumulierten jährlichen Kosten bezüglich des Betrachtungszeitraums von 10 Jahren und unterscheidet hinsichtlich des Pflegeaufwandes, der sich nach der Art des Datenüberprüfungsverfahrens richtet.

Tabelle 3-19: kumulierte Kosten mit Abschreibung der Investitionen (Auszug)

Abschreibungsgegenstand	Abschreibungszeitraum / Nutzungszeitraum des Intermodalen Routenplaners									
	2. Jahr		4. Jahr		6. Jahr		8. Jahr		10. Jahr	
	wt %	qt	wt %	qt	wt %	qt	wt %	qt	wt%	qt
digitale Straßenkarte	27	10.133,33 €	13	5.066,67 €	33	12.666,67 €	20	7.600,00 €	7	2.533,33 €
Hard- und Software	27	6.666,67 €	13	3.333,33 €	33	8.333,33 €	20	5.000,00 €	7	1.666,67 €
Geodatengrundlage 100% Prüfung	10	25.300,00 €	10	25.300,00 €	10	25.300,00 €	10	25.300,00 €	10	25.300,00 €
Geodatengrundlage Stichprobe	10	17.700,00 €	10	17.700,00 €	10	17.700,00 €	10	17.700,00 €	10	17.700,00 €
Nachbearbeitung	30% mit 10%	511,00 €		511,00 €		511,00 €		511,00 €		511,00 €
	100% mit 5%	255,50 €		255,50 €		255,50 €		255,50 €		255,50 €
Modellierung	30% mit 10%	103,00 €		103,00 €		103,00 €		103,00 €		103,00 €
	100% mit 5%	51,30 €		51,30 €		51,30 €		51,30 €		51,30 €
Georeferenzierung	30% mit 10%	411,00 €		411,00 €		411,00 €		411,00 €		411,00 €
	100% mit 5%	205,50 €		205,50 €		205,50 €		205,50 €		205,50 €
1/2 Personalstelle Pflege / Support		30.000,00 €		30.000,00 €		30.000,00 €		30.000,00 €		30.000,00 €
Summe Kosten 30%ige Prüfung		65.525,00 €		57.125,00 €		69.725,00 €		61.325,00 €		52.925,00 €
kumulierte Kosten 30%		135.250,00 €		253.700,00 €		376.350,00 €		503.200,00 €		613.250,00 €
Summe Kosten 100%ige Prüfung		72.612,30 €		64.212,30 €		76.812,30 €		68.412,30 €		60.012,30 €
kumulierte Kosten 100%		149.424,60 €		282.049,20 €		418.873,80 €		559.898,40 €		684.123,00 €

Es wird zur Abschätzung des Pflegeaufwandes davon ausgegangen, dass sich bei einer am Anfang durchgeführten vollständigen Überprüfung ein jährlicher Pflegeaufwand von 5% und für die Stichprobenprüfung von 10% ergibt. Ferner werden die Investitionen in die digitale Straßenkarte und in die Hard- und Software degressiv auf zweimal fünf Jahre abgeschrieben, da mit einer Alterung des Geodatenproduktes und einem Wertverlust auf Grund technischer Überholung der Hard- und Software zu rechnen ist. Für die Personalkosten sind Lohnkosten einer halben Stelle mit konstanten € 30.000 zur Systempflege angesetzt. Daraus resultieren, abhängig von dem Pflege- und Überprüfungsaufwand die kumulierten Kosten am Ende des Nutzungszeitraums von 613 T€ bzw. 684 T€. Auf eine vollständige Darstellung der kumulierten Kosten ist hier wegen der Übersichtlichkeit verzichtet worden.

3.7 Abschätzung erzielbarer Erlöse durch Nutzungsentgeld

Die Refinanzierung von Informationsdienstleistungen durch Nutzungsgebühren ist direkt von der Akzeptanz des Kunden in Form eines individuellen Informationsgewinns, seiner Nutzungshäufigkeit und dessen Zahlungsbereitschaft abhängig. Unterstellt man dem Informationsdienst eine kundenorientierte und transparente Aufbereitung des individuellen Nutzungsvorteils, so stellt sich dem Informationsdienstleister die Frage nach der Höhe der Entgelte, die er vom Nutzer erheben kann und nach der mindestens notwendigen jährlichen Nutzungshäufigkeit, um das Dienstleistungsangebot wirtschaftlich betreiben zu können.

In diesem Abschnitt soll die Vorgehensweise zur Ermittlung von zu erwartender Nutzungshäufigkeit am Beispiel des intermodalen Routenplaners aufgezeigt und prognostiziert werden. Die Prognose der Nutzungshäufigkeit ist allgemein von folgenden Faktoren abhängig:

- Mögliche, potentielle Nutzeranzahl (inkl. Bevölkerungsentwicklung)
- Internetverfügbarkeit für potentielle Nutzer,
- Bekanntheitsgrad des Dienstes (Penetrationsrate)
- Grad der Inanspruchnahme des Dienstes.

Die Vorgehensweise zur Abschätzung der Nutzungshäufigkeit kann durch folgende Abbildung 3.20 dargestellt werden. Sie geschieht nach derselben, von der Arbeitsgruppe Wissenschaft von MOBILIST entwickelten Bewertungsmethodik zur Abschätzung der Nutzungshäufigkeit künftiger Informationsdienstleistungen [HAUTZINGER ET AL., 2001].

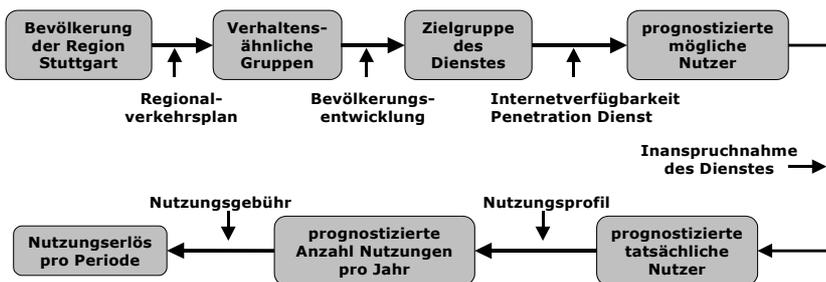


Abbildung 3.20: Vorgehensweise der Abschätzung erzielbarer Nutzungserlöse

Um die möglichen, potentiellen Nutzer des intermodalen Dienstes zu ermitteln, wird hier auf den Regionalverkehrsplan der Region Stuttgart zurückgegriffen [RVP, 2001]. Darin ist die Mächtigkeit von so genannten verhaltensähnlichen Gruppen, die ein ähnliches Verkehrsverhalten besitzen bzgl. des Bezugsjahres 2010 enthalten. Die Gesamtsumme von Personen dieser verhaltensähnlichen Gruppen (Erwerbstätiger mit Pkw - EmP, Teilerwerbstätiger mit Pkw - TemP, Nicht-Erwerbstätiger mit Pkw - NemP und Rentner mit Pkw - RmP) dient als Ausgangsbasis zur Abschätzung und wird im Weiteren als potentielle Nutzeranzahl bezeichnet. Diese Mächtigkeit ist entsprechend der jährlichen Bevölkerungsentwicklung von ca. +0,5% [RVP, 2001] auf den Betrachtungszeitraum (10 Jahre) verteilt. Dieser Personengruppen sind damit als Zielgruppe des intermodalen Routenplaners identifiziert.

Tabelle 3-20: Anteil verhaltensähnlicher Gruppen entsprechend Nutzungszeitraum

Jahr	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Anzahl Jahre	1	2	3	4	5	6	7	8	9	10
EmP	797301	801307	805334	809381	813448	817536	821644	825773	829902	834051
TemP	188393	189340	190292	191248	192209	193175	194145	195121	196097	197077
NemP	103636	104157	104680	105206	105735	106266	106800	107337	107874	108413
RmP	282670	284090	285518	286952	288394	289844	291300	292764	294228	295699
Summe:	1372000	1378894	1385824	1392788	1399786	1406821	1413890	1420995	1428100	1435240

Tabelle 3-20 zeigt die Anteile der verhaltensähnlichen Gruppen mit dem Bezugsjahr 2010 des Regionalverkehrsplans und einer unterstellten, konstanten Bevölkerungsentwicklung von +0,5% über den betrachteten Nutzungszeitraum von

10 Jahren. Eine personengruppenspezifische Entwicklung wurde hier vernachlässigt. Um von der potentiellen auf die prognostizierte, mögliche Nutzerzahl zu kommen, muss einerseits die Internetverfügbarkeit bzw. –nutzung und andererseits der Bekanntheitsgrad des Dienstes (Penetration) mit in die Rechnung einfließen. Aus einer Pressemitteilung des Statistischen Landesamtes Baden-Württemberg [STAT.LABW,2003] heißt es dazu:

Der Nutzungsgrad von Personalcomputer und Internet ist vor allem bei jungen Leuten in sehr hohem Maße verbreitet. So nutzten 94% der 16-bis24-jährigen im ersten Quartal 2002 das Internet; in der Altersgruppe der 25- bis 34-jährigen waren es 88 Prozent. Auffällig ist, dass der Anteil der Computernutzer mit zunehmendem Alter erkennbar sinkt. Es nutzten im gleichen Quartal 2002 nur etwa die Hälfte der 55- bis 64-Jährigen einen PC. Darüber hinaus ist eine hohe Abhängigkeit zwischen Internetnutzung und sozialer Stellung zu verzeichnen. Es setzen nahezu alle Schüler und Studierende mittlerweile einen PC ein (99%), während Personen ohne Erwerb und Ausbildung nur zu ca. 50% und Personen im Ruhestand nur zu ca. 25% einen PC nutzen. Entsprechend niedriger war auch die Internetnutzung. Nur etwa jeder 5-te. Ruheständler, etwa ein Drittel der Hausfrauen und- männer sowie ca. 50% der Arbeitslosen nutzten in dieser Periode das Internet.

Berücksichtigt man darüber hinaus die Dienstpenetration, als eine Entwicklung des Bekanntheitsgrades von anfänglich 5% bis zu 20% im Jahre 2012 und die entsprechende prozentuale Steigerung der Internetnutzung (N_I), bezogen auf die verhaltensähnlichen Gruppen mit jährlichen Steigerungsraten von 0,1%, so ergibt sich die Anzahl prognostizierter, mögliche Nutzer pro Jahr.

Tabelle 3-21: prognostizierte, mögliche Nutzer (Internetverfügbarkeit, Dienstpenetration)

Jahr	2003		2004		2005		2006		2007		2008		2009		2010		2011		2012	
Anzahl Jahre	1		2		3		4		5		6		7		8		9		10	
Penetration Dienst	5%		7%		10%		12%		14%		16%		18%		19%		20%		20%	
	NI	Nutzer	NI	Nutzer	NI	Nutzer	NI	Nutzer	NI	Nutzer	NI	Nutzer	NI	Nutzer	NI	Nutzer	NI	Nutzer	NI	Nutzer
EmP (Nutzung)	70%	27906	72%	40386	73%	58789	74%	71873	75%	85412	76%	99412	77%	113880	78%	122380	79%	131124	80%	133448
TemP (Nutzung)	65%	6123	66%	8748	67%	12750	68%	15606	69%	18567	70%	21636	71%	24812	72%	26693	73%	28630	74%	29167
NemP (Nutzung)	50%	2591	51%	3718	52%	5443	53%	6691	54%	7994	55%	9351	56%	10765	57%	11625	58%	12513	59%	12793
RmP (Nutzung)	25%	3533	25%	4972	26%	7423	27%	9297	28%	11305	29%	13449	30%	15730	31%	17244	32%	18831	33%	19516
Summe:	40153		57823		84406		103467		123278		143848		165187		177941		191099		194925	

Die mit dieser Prognose ermittelten Nutzerzahlen dienen im folgenden als Bezugsgröße der möglichen Nutzer. Sie stellen sozusagen diejenigen Personen dar, welche aufgrund gleicher oder ähnlicher Rahmenbedingungen Zugang zum Intermodalen Dienst haben, über dessen Existenz informiert sind und die den entsprechenden verhaltensähnlichen Gruppen angehören.

Tabelle 3-22: Szenarien der Nutzungsentwicklung

Optimistische Entwicklung	Pragmatische Entwicklung	Pessimistische Entwicklung
<ul style="list-style-type: none"> Inanspruchnahme des Dienstes nach 10 Jahren bei ca. 30% Sättigung Exponentiell-degressive Zunahme Stammkundschaft hat sich etabliert. Mobile PDA unterstützen die Online-Auskunft 	<ul style="list-style-type: none"> Inanspruchnahme des Dienstes nach 10 Jahren bei ca. 20% Sättigung Langsame exp.- degressive Zunahme Kleine Stammkundschaft Verbesserte Umsteigeinformationen an Haltestellen 	<ul style="list-style-type: none"> Inanspruchnahme des Dienstes nach 10 Jahren bei ca. 10% Sättigung Stetige, fast lineare Zunahme Wechselnde Kunden, sporadische Teilnahme Keine mobile Auskunft über PDA.
<p>Nutzungen</p> <p>30% der progn. mögl. Nutzer</p> <p>Zeit</p>	<p>Nutzungen</p> <p>20% der progn. mögl. Nutzer</p> <p>Zeit</p>	<p>Nutzungen</p> <p>10% der progn. mögl. Nutzer</p> <p>Zeit</p>

Um von den prognostizierten, möglichen Nutzern auf die prognostizierten tatsächlichen Nutzer schließen zu können, muss angenommen werden, dass auch nur ein Teil der möglichen Nutzer den Dienst tatsächlich in Anspruch nimmt und einem eigenen Nutzungsprofil unterliegt. Diese Abschätzung der grundsätzlichen Inanspruchnahme geschieht über die in Tabelle 3-22 dargestellten drei Szenarien der potentiellen Nutzungsentwicklung. Diese drei Szenarien unterscheiden sich hauptsächlich in der nach Ablauf des Betrachtungszeitraums von 10 Jahren erreichte Penetrationsrate von 10- bis 30-ig Prozent sowie einer exponentiell-degressiven Entwicklung der Nutzerzahlen. Es werden der Dienstleistung damit bei der Hälfte ihrer Sättigungskurve die größten Zuwächse an Nutzern unterstellt. Neben der Definition von Szenarien der Dienstentwicklung spielt das **Nutzungsprofil** der möglichen Nutzer eine wesentliche Rolle. Es muss zur wirksamen Abschätzung der potentiellen Nutzungen pro Periode zwischen verschiedenen Nutzungshäufigkeiten unterschieden werden. Nicht alle möglichen Nutzer werden dasselbe Nutzungsverhalten besitzen. Nachfolgend sind die Nutzungen entsprechend den drei Ausgangsszenarien und einer Auswahl von Nutzungsprofilen hochgerechnet. Diese Nutzungsprofile der Dienstleistung sind wie folgt zu verstehen:

Tabelle 3-23: Tabelle Nutzungsprofile

Profil-Nr.	Nutzungshäufigkeit	Abkürzung	Faktor	Berücksichtigung
1	1 malige Nutzung im Halbjahr	1/Halbjahr	2	-
2	1-malige Nutzung im Vierteljahr	1/Vierteljahr	4	-
3	1-malig Nutzung im Monat	1/Monat	12	Szenario opt.;
4	5-malige Nutzung im Monat	5/Monat	60	Szenario opt.,-prag.,pessim.
5	10-malige Nutzung im Monat	10/Monat	120	Szenario opt.,-prag.,pessim.
6	20-malige Nutzung im Monat	20/Monat	240	Szenario pessim.

Von den in Tabelle 3-23 definierten Nutzungsprofilen werden auf Grund der Übersichtlichkeit der weiteren Betrachtung nur die Profile 3 bis 6 bei den Szenarien berücksichtigt.

Tabelle 3-24: Hochrechnung der tatsächlichen Nutzungsanzahl für die Szenarien

Jahr	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
	1	2	3	4	5	6	7	8	9	10
Summe mögl. Nutzer	40153	57823	84406	103467	123278	143848	165187	177941	191099	194925
Inanspruchnahme opt.	5%	6%	8%	10%	15%	20%	25%	27%	29%	30%
Nutzer opt.	2008	3469	6752	10347	18492	28770	41297	48044	55419	58477
Inanspruchnahme prag.	3%	4%	5%	8%	10%	13%	15%	17%	19%	20%
Nutzer prag.	1205	2313	4220	8277	12328	18700	24778	30250	36309	38985
Inanspruchnahme pess.	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%
Nutzer pess.	402	1156	2532	4139	6164	8631	11563	14235	17199	19492
Nutzungen pro Jahr für optimistisches Szenario										
1mal / Monat	24092	41633	81030	124161	221901	345235	495562	576527	665023	701728
5mal / Monat	120458	208164	405148	620803	1109503	1726177	2477810	2882636	3325114	3508641
10mal / Monat	240916	416328	810295	1241607	2219005	3452355	4955620	5765272	6650229	7017282
Nutzungen pro Jahr für pragmatischen Szenario										
5mal / Monat	72275	138776	253217	496643	739668	1122015	1486686	1814993	2178523	2339094
10mal / Monat	144549	277552	506435	993285	1479337	2244031	2973372	3629986	4357046	4678188
Nutzungen pro Jahr für pessimistisches Szenario										
5mal / Monat	24092	69388	151930	248321	369834	517853	693787	854114	1031932	1169547
10mal / Monat	48183	138776	303861	496643	739668	1035706	1387574	1708229	2063864	2339094
20mal / Monat	96366	277552	607721	993285	1479337	2071413	2775147	3416458	4127728	4678188

Während die Penetrationsrate des Dienstes, die über den Dienst informierten Nutzer von den nicht informierten trennt, werden über den prozentualen Anteil der Inanspruchnahme diejenigen Nutzer selektiert, die den Dienst kennen und auch nutzen. Legt man diesen prognostizierten, tatsächlichen Nutzern das spezifisches Nutzerprofil (Tabelle 3-23) zu Grunde, dann erhält man die Nutzungen pro Jahr für die entsprechenden Szenarien (vgl. Tabelle 3-24). Neben den szenarien-abhängigen, prognostizierten tatsächlichen Nutzungen sind die Nutzungsgebühren, welche der Dienst beim Kunden erhebt, zur Berechnung des Zeitpunkts des *Return On Investment* entscheidend. Zu dessen Berechnungen im folgenden Abschnitt wird von Nutzungsgebühren von 5 bzw. 10 Eurocent ausgegangen.

3.8 Berechnung des Return On Investment für die Nutzungsszenarien

Vergleicht man die kumulierten Kosten mit den Erlösen durch Nutzungsgebühren der drei Szenarien, dann ergeben sich erwartungsgemäß unterschiedliche Zeitpunkte des *Return On Investment*.

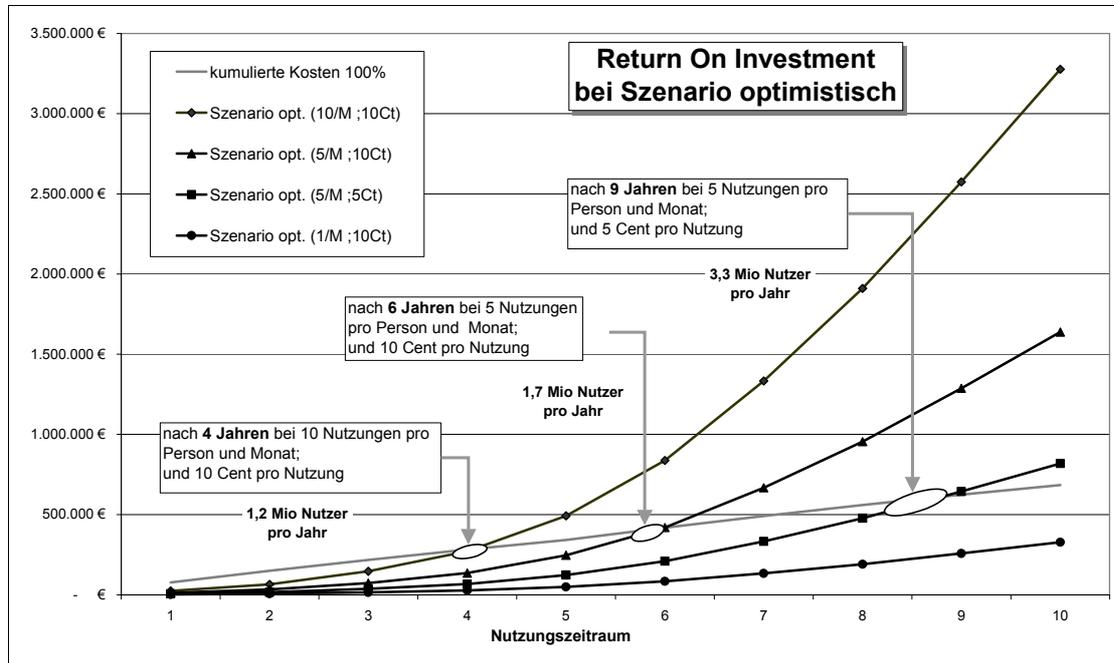


Abbildung 3.21: Zeitpunkte des Return On Invests bei Szenario optimistisch

Im Falle einer optimalen Dienstentwicklung (Szenario opt.) ergeben sich je nach Nutzungshäufigkeit und Nutzungspreise nach vier, sechs und neun Jahren die Reinvestition der Investitionssumme.

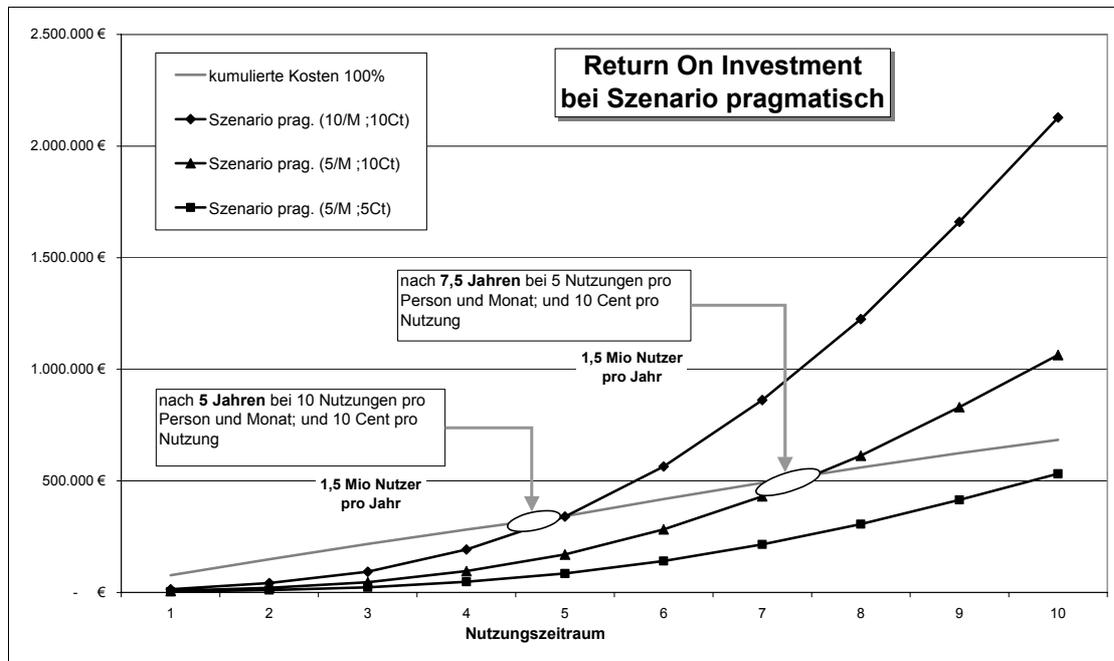


Abbildung 3.22: Zeitpunkte des Return On Invests bei Szenario pragmatisch

Unterstellt man der Dienstentwicklung ein eher pragmatisches Entwicklungsverhalten, dann stellt sich der ROI bei gleichem Nutzungsverhalten und Nutzungsentgelte nach 5 bzw. 7,5 Jahren, ein.

Zwar stellt sich der ROI beim pessimistischen Szenario ebenfalls wie beim pragmatischen Szenario nach 5 bzw. 7,5 Jahren ein, jedoch ist das Nutzungsverhalten von 10 auf 20 bzw. 5 auf 10 Nutzungen pro Person und Monat verdoppelt

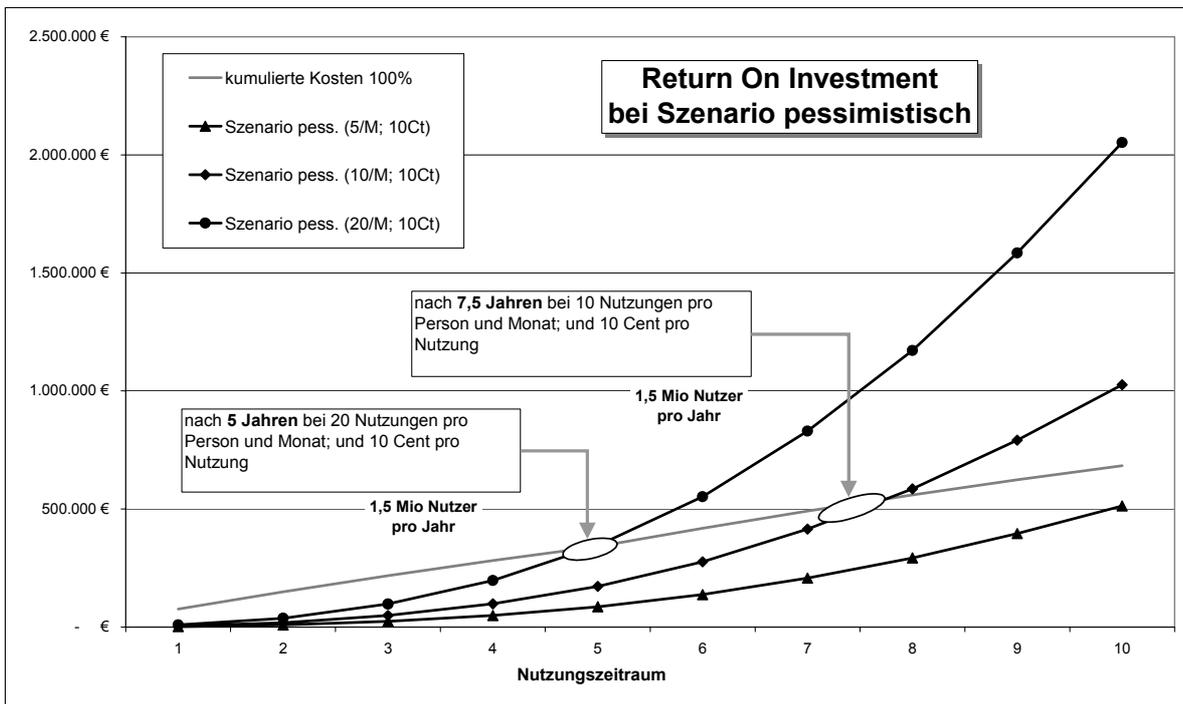


Abbildung 3.23: Zeitpunkte des Return On Invests bei Szenario pessimistisch

3.9 Beurteilung des Ergebnisses

Als Ergebnis dieses Vergleichs der kumulierten Kosten mit den prognostizierten, szenarienabhängigen kumulierten Erlösen kann festgehalten werden, dass grundsätzlich der *Return On Investment* innerhalb des betrachteten zehnjährigen Nutzungszeitraums möglich ist und dass es zu keiner nennenswerten zeitlichen Verschiebung des ROI - Zeitpunktes durch die Kostenunterschiede, die sich durch die 30%ige Prüfung oder die vollständige Überprüfung ergeben, kommt. Sämtliche Szenarien zeigen einen diesbezüglichen zeitlichen Unterschied zwischen 3 bis 6 Monaten. Es kann deshalb hier die Empfehlung zugunsten der vollständigen Überprüfung ausgesprochen werden, da mit ihr eine Steigerung der Auskunftqualität und einer damit verbundenen steigenden Nutzerzufriedenheit erreichbar ist.

Betrachtet man zuerst die Nutzerzahlen bei den einzelnen ROI - Zeitpunkten, so stellt man fest, dass unabhängig vom zugrunde gelegten Nutzungsszenario mindestens 1,5 Millionen jährliche Nutzer notwendig sind, um nach der entsprechenden Anzahl von Jahren den Return On Invest erzielen zu können. Legt man das optimistische Nutzungsszenario mit einer Nutzungshäufigkeit von 5 Nutzungen pro Person und Monat sowie dem geringen Entgelt von 5 Cent pro Nutzung zu Grunde, so kommt es sogar noch nach Ablauf von neun Jahren mit allerdings 3,3 Millionen notwendiger Nutzer zu einem ROI (vgl. Abbildung 3.21 – Szenario opt. – (5/M;5Ct)).

Für eine Nutzungshäufigkeit von nur einer Nutzung pro Person und Monat bei 10 Cent Nutzungsgebühren existiert kein ROI innerhalb des Betrachtungszeitraums. Der erste ROI - Zeitpunkt ergibt sich für das optimistische Nutzungsszenario nach Ablauf des vierten Jahres bei angenommenen 10 Nutzungen pro Person und Monat sowie der Entgeltgebühr von 10 Cent pro Nutzung. Dieser Zeitpunkt ist als kritisch zu beurteilen, da er von einem sehr hohen Nutzungsprofil ausgeht. Dies würde bedeuten, dass sich die Nutzer fast jeden zweiten bis dritten Tag eines Monats eine für sie entgeltspflichtige Route berechnen lassen. Dies widerspricht allerdings dem Profil der überwiegenden Zahl der Nutzer, die sich als Berufspendler bezeichnen lassen. Deshalb ist dieser ROI - Zeitpunkt als unwahrscheinlich zu bezeichnen.

Ein realistischerer ROI - Zeitpunkt ist der nach Ablauf von 6 Jahren bei 5 Nutzungen pro Person und Monat und einer Gebühr von 10 Cent. Unter einem solchen Nutzungsprofil könnte man sich das Bewegungsverhalten eines Berufspend-

lers vorstellen, welcher ca. 5 mal im Monat an Orte geführt werden will, die nicht seiner Arbeitstätte entsprechen und bei denen er über keine Ortskenntnis verfügt.

Betrachtet man hingegen das pragmatische Nutzungsszenario (Abbildung 3.22) so ergeben sich die Zeitpunkte des ROI nach Ablauf von fünf Jahren bei einer Nutzung pro Person von 10 mal im Monat mit einer Gebühr von 10 Cent pro Nutzung und nach 7,5 Jahren bei der Hälfte der Nutzungen pro Person (5 mal). Auch hier gilt, dass eine 10 – malige Nutzung pro Person und Monat eher unwahrscheinlich ist. Für eine fünfmalige Nutzung pro Person und Monat bei 5 Cent Nutzungsgebühren existiert bei diesem pragmatischen Szenario kein ROI-Zeitpunkt innerhalb des Betrachtungszeitraums.

Um bei dem pessimistischen Szenario (Abbildung 3.23) überhaupt eine Chance auf Reinvestition zu bekommen, muss von einer Nutzungshäufigkeit pro Person und Monat von 10 bzw. 20 Nutzungen ausgegangen werden. Auch dies ist als sehr unwahrscheinlich zu bezeichnen. Die ROI – Zeitpunkte ergeben sich bei gleichem Nutzungsentgelt von 10 Cent hier auch nach Ablauf von 5 bzw. 7,5 Jahren. Während beim pragmatischen Szenario die fünfmalige Nutzung pro Person und Monat bei 10 Cent Nutzungsgebühren noch zum ROI innerhalb des Betrachtungszeitraums führt, erhält man für das pessimistische Szenario keinen ROI innerhalb dieser 10 Jahre.

3.10 Zusammenfassende Bemerkung

Zusammenfassend kann festgehalten werden, dass eine Chance auf Wiedergewinnung der Investition nur bei den ersten beiden Szenarien (optimistisch und pragmatisch) besteht. Dies gilt aber auch nur bei gleichzeitiger, realistischer Einschätzung der Nutzungshäufigkeit pro Person und Monat. So kann eigentlich nur von den beiden realistischen ROI Zeitpunkten nach Ablauf von **9 Jahren** bei optimistischer und nach Ablauf von **7,5 Jahren** bei pragmatischer Gesamtnutzerzahl ausgegangen werden. Die anderen, prinzipiell möglichen **ROI – Zeitpunkte** scheiden wegen zu hoch angenommenen Nutzungshäufigkeiten aus.

- Betrachtet man mit Hilfe dieser Kenntnis die **kritische Nutzerzahl**, welche überhaupt zu erreichen ist um eine Chance auf Wiedergewinnung der Investition zu erhalten, dann kann ein Spektrum **zwischen 1,5 Mio und 3,3 Mio** Nutzer pro Jahr festgestellt werden. Dieses Spektrum entspricht einem ca. 10-20%igem Anteil an den jährlichen entgeltfreien Nutzungen des elektronischen Fahrplanauskunftssystems EFA der Region des Mittleren Neckarraums (Auskunft des Verkehrs- und Tarifverbunds Stuttgart).
- Es muss folglich das Ziel des Dienstebereibers sein, mindestens **10 bis 20% der Kundschaft** (bzw. Neukunden) **des Fahrplanauskunftssystems EFA** für einer entgeltpflichtige Nutzung des Intermodalen Routenplaners zu gewinnen um eine Chance auf Refinanzierung der Investition und der laufenden Kosten zu erreichen. Damit ist jedoch das Problem alleine noch nicht gelöst, denn es verbleibt immerhin noch die fragliche Zahlungsbereitschaft der Nutzer, bzw. das ebenfalls mit Kosten verbundene Abrechnungsverfahren.
- Die Akzeptanz und Inanspruchnahme des Dienstes sowie die Bereitschaft des Kunden für die Informationsbereitstellung ein Entgelt zu entrichten wächst mit steigendem Verkehrsaufkommen und damit verbundenem individuellen Zeit- und Geldverlust. Stellt man sich die Dynamisierung des Intermodalen Routenplaners, unterstützt durch mobile Organzier (PTA - Personal Travel Assistant, PDA - Personal Digital Assistant) und durch leistungsverbesserte Mobilkommunikation (UMTS) als Zukunftsszenario vor, dann wäre nicht nur eine bereits heutzutage in Ansätzen vorhandene Stauumfahrung mit Hilfe von Verkehrsmeldungen über RDS-TMC in Kombination mit einem Kfz-Navigationssystem sondern auch die „onboard“-Berechnung einer intermodalen Alternativroute denkbar.

In einem solchen Fall könnten Fahrzeuge an der letzten Möglichkeit von dem Stauanfang die Strecke verlassen und per Navigationssystem auf einen Park&Ride-Platz geführt werden und deren Fahrer in öffentliche Verkehrsmittel umsteigen. Wenn diese, heutzutage bereits existierenden, technischen Einzelkomponenten zu einem wirtschaftlichen Gesamtkonzept zusammengeführt werden, ist auch von einer steigenden Nutzernachfrage des intermodalen Routenplaners auszugehen.

4 Ein Konzept zur Ermittlung der Bereitstellungskosten von Geodaten

4.1 Einführung

Die Aufbau- und Bereitstellungskosten nutzbarer Geodaten für verkehrliche Anwendungen werden derzeit über grobe Vorkalkulationsverfahren unter Einsatz von Expertenwissen prognostiziert oder mit Hilfe von Erfahrungswerten [BEHR, 1998] extrapoliert. Eine solche Schätzung beinhaltet eine Menge pauschalisierte Kostenfaktoren die ohne Fachwissen nur schwer durchschaubar und nur unzureichend auf den eigenen Sachverhalt übertragbar sind. Deshalb wird in diesem Kapitel ein Konzept vorgestellt, dessen Aufgabe es ist, die Bereitstellungskosten eines Geodatenbestandes für verkehrliche Anwendungen zu prognostizieren. Dieses Konzept soll dabei folgenden Anforderungen genügen:

- Das Konzept zur Ermittlung der Bereitstellungskosten von Geodaten muss sich bei allen Anforderungen von Verkehrsinformationsdienstleistungen anwenden lassen und es hat sämtliche Möglichkeiten des Aufbaus von Geodaten zu berücksichtigen.
- Das Konzept muss bereits in der Definitionsphase des Aufbaus von Geodatenbeständen anwendbar sein und hat notwendige Grundlagen in Form von Kostenstrukturen für strategische Unternehmensentscheidungen zu liefern.
- Das Konzept hat die Aufgabe, transparente Beurteilungskriterien zu schaffen, die zur Entscheidung über die Vorgehensweise des Aufbaus von Geodatenbeständen herangezogen werden.

Durch diese Beurteilungskriterien, die mit Hilfe von Qualitätsforderungen ausgedrückt werden, soll erstens die Komplexität und die Kostenintensivität, die mit der Nutzung von Geodaten allgemein verbunden ist, nachvollziehbar angegeben werden können sowie zweitens ein Beitrag dazu geleistet werden, die Unsicherheiten, die bei der Definition der richtigen Vorgehensweise der Geodatenbereitstellung existieren, zu minimieren. Diese Minimierung trägt entscheidend dazu bei, die richtige Strategie im Spannungsfeld zwischen kostensenkendem unternehmerischen Denken und technisch Machbarem zu finden.

Die Kernelemente des hier vorgestellten Bereitstellungskonzepts sind, in Analogie zu Prognosemodellen des industriellen Fertigungsprozesses, ein **Kostenmodell** sowie ein **Berechnungsverfahren**, welche sich für verschiedene Möglichkeiten der Bereitstellung von Geodatenbeständen verwenden lassen (Abschnitt 4.3). Da ein Kostenmodell grundsätzlich von der Struktur des Prozesses abhängig ist, auf das es angewendet werden soll, ist in Abschnitt 4.2 der Prozess der **allgemeinen Geodatenbereitstellung** in seinen wesentlichen Hauptprozessen zerlegt dargestellt.

Die verschiedenen Möglichkeiten, durch die Geodatenbestände aufgebaut und bereitgestellt werden können, hängen einerseits von der Ausgangssituation, nämlich der Existenz eigener Datenbestände und andererseits von den Aktionen des Zukaufs von Fremddaten, der Integration und Homogenisierung zweier Datenbestände sowie von eventuell notwendig werdender eigener Erfassung von Daten ab. Wie nun aus der jeweiligen Ausgangssituation und aus den Einzelaktionen die richtige Vorgehensweise ausgewählt wird und welche Abhängigkeiten sich daraus ergeben, ist Gegenstand der Betrachtung prinzipiell in Frage kommender **Aufbaumöglichkeiten** in Abschnitt 4.4.

Das Kostenmodell, der Prozess der allgemeinen Geodatenbereitstellung und die verschiedenen Aufbaumöglichkeiten bilden die Grundlage für das in Abschnitt 4.6 beschriebene Rechenverfahren zur Kostenprognose. Dieser Zusammenhang ist in Abbildung 4.1 verdeutlicht.

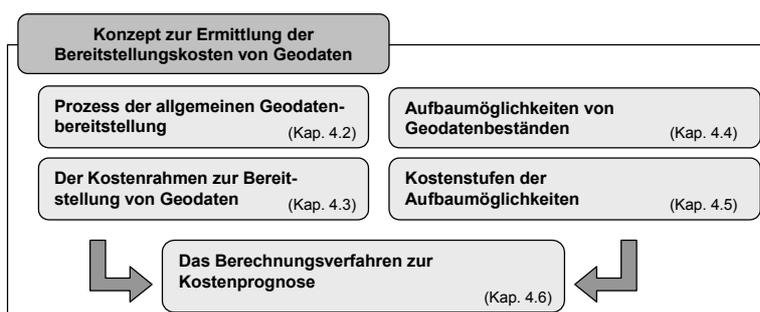


Abbildung 4.1: Konzept zur Ermittlung der Bereitstellungskosten von Geodaten

4.2 Prozess der allgemeinen Geodatenbereitstellung

Der Vorgang der Geodatenbereitstellung lässt sich allgemein in die Hauptprozesse der *Anforderungsanalyse*, der *Datenerhebung*, der *Integration*, sowie des *Homogenisierens* mit den zugehörigen Kostenfaktoren (vgl. Abbildung 4.2) unterteilen. Unter dem Begriff der **Geodatenbereitstellung** werden im Rahmen dieser Arbeit die Aufgaben der Datenveredelung, der vollständige Neuaufbau, die Fusion zweier Datenbestände sowie der Vorgang des Aktualisierens zusammengefasst. Welche dieser Aufgaben letztendlich im Rahmen der Geodatenbereitstellung durchzuführen sind, hängt von der Ausgangssituation bzw. dem Ausgangsdatenbestand ab und wird in Form von Aufbaumöglichkeiten in Abschnitt 4.4 weiter behandelt.

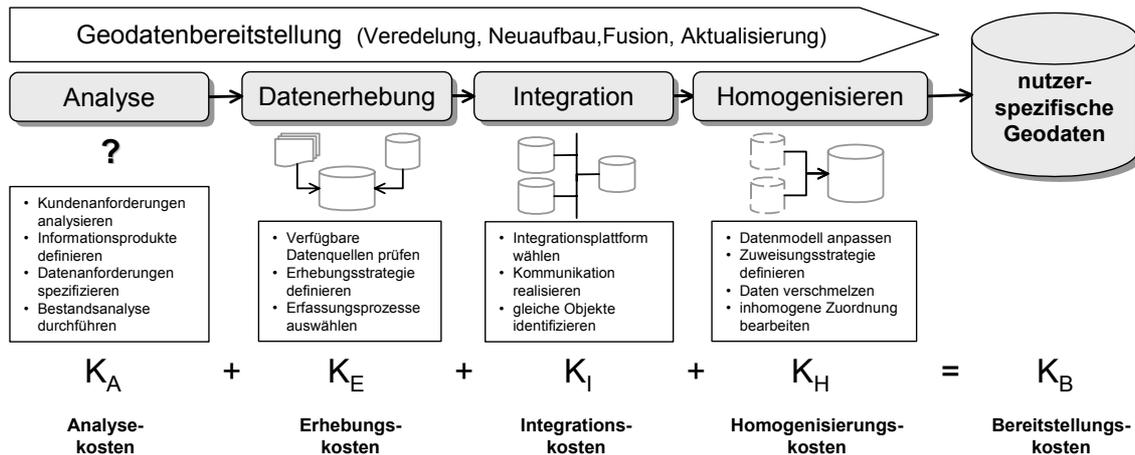


Abbildung 4.2: Prinzipieller Ablauf der Geodatenbereitstellung mit Prozesskosten

Der hier dargestellte Prozessablauf soll einen einheitlichen Rahmen schaffen, in den sich die erwähnten Aufgaben und Strategien der Geodatenbereitstellung einordnen lassen. Durch die weiterführende Unterteilung der Hauptprozesse *Analyse*, *Datenerhebung*, *Integration* und *Homogenisieren* in weitere, so genannte Teilprozesse (Kap. 4.4 und Kap 4.5) soll ein Diskretisierungsgrad erreicht werden, mit dessen Hilfe die Erarbeitung **spezifischer Kostenstrukturen** für die genannten Aufgaben und die damit verbundenen Aufbaumöglichkeiten der Geodatenbereitstellung möglich wird. Dies gelingt nur, wenn auch die den Hauptprozessen zugeordneten Kostenfaktoren mit demselben Diskretisierungsgrad angegeben werden können, wie die der Teilprozesse.

Eine spezifische Kostenstruktur orientiert sich an dem jeweils betrachteten Prozessablauf einer Datenbereitstellung. Durch ein noch zu definierendes **Kostenmodell** können die Kostenanteile der Hauptprozesse angegeben werden. Der für die Kostenstruktur verantwortliche Prozessablauf ergibt sich aus der Vorgehensweise des Geodaten verarbeitenden Unternehmens. Trotz der unterschiedlichen Aufgaben, die zur Geodatenbereitstellung zusammengefasst sind, gelten folgende allgemeingültige Hauptprozesse.

Der Hauptprozess **Analyse** umfasst im wesentlichen vorbereitende Arbeiten. Zu ihnen zählt die Analyse der Kundenanforderungen sowie die daraus unmittelbar resultierende Definition des Informationsproduktes, welches die Kundenanforderung erfüllt. Durch die Kenntnis dieses Produktes und die dahinter liegenden Funktionalitäten sind die Anforderungen an die darin zu verarbeitenden Daten bekannt und spezifizierbar. Am Ende des Hauptprozesses *Analyse* steht die Untersuchung des eigenen Datenbestandes. Sie untersucht, inwieweit eigene Bestände zur Erfüllung der Anforderungen herangezogen werden können. Entsprechend der spezifizierten Datenanforderung und der durchgeführten Analyse eigener Datenbestände erfolgt die vorläufige Festlegung darauf, ob die Bereitstellung der Daten durch einen Neuaufbau oder die Veredelung des eigenen Bestandes stattfindet. Eine endgültige Festlegung kann allerdings erst nach Bearbeitung aller Hauptprozesse stattfinden.

Der zweite Hauptprozess **Datenerhebung** beinhaltet als ersten Schritt die Auswahl und Prüfung verfügbarer Datenquellen. Auf der Grundlage der Erkenntnisse dieses ersten Schritts wird zwischen *vollständigem Neuaufbau* und *Veredelung des eigenen Bestandes* entschieden und die Vorgehensweise endgültig festgelegt. Damit kommen den Hauptprozessen

Analyse und Datenerhebung besondere Bedeutung zu. Konkretisiert wird sie schließlich durch die Auswahl und das Zusammenwirken der Datenerfassungsverfahren sowie die Vorgehensweise zur Integration und Homogenisierung. Diejenigen Aufbaumöglichkeiten, die im Rahmen dieser Arbeit zur Aufstellung von Kostenmodellen herangezogen werden, sind in Abschnitt 4.4 hergeleitet und dargestellt.

Der dritte Hauptprozess **Integration** hat die Realisierung des einheitlichen Zugriffs auf die im vorangegangenen Prozess spezifizierten, erhobenen und erfassten Daten zum Ziel. Innerhalb der Integration unterschiedlicher Geodaten muss erstens die Integrationsplattform (Rechner, Betriebssystem, DBMS, etc.) gewählt, zweitens die Kommunikation (Schnittstellen, Protokolle) realisiert sowie drittens das Problem der Identifikation gleicher Objekte (Geodaten) gelöst werden. Es ist vom definierten Informationsprodukt abhängig, ob eine Integration der Geodaten ausreichend ist oder ob die weiterführende Homogenisierung stattfinden muss.

Der vierte und abschließende Hauptprozess **Homogenisierung** baut unmittelbar auf der mittlerweile zur Verfügung stehenden Integrationslösung auf. Von einer Homogenisierung spricht man, wenn die Inhalte mehrerer Datenbestände vollständig zu einem gemeinsamen Datenbestand verschmolzen werden bzw. die Datenveredelung so durchgeführt wird, dass die Zusatzdaten vollständig dem Grunddatenbestand zuweisbar sind. Im Falle einer Homogenisierung kann es notwendig werden, die Datenmodelle der zu verschmelzenden Bestände anzupassen, zu verändern oder vollkommen neu zu modellieren. Darüber hinaus müssen Zuweisungsvorschriften für Objektklassen bzw. deren Instanzen unter Einbeziehung aller Attribute erarbeitet und Alternativen entwickelt werden, wie mit nicht-zuweisbaren Objekten weiter zu verfahren ist.

4.3 Der Kostenrahmen zur Bereitstellung von Geodaten

Der erste Schritt in der Aufstellung eines Konzeptes zur Berechnung der Bereitstellungskosten von Geodaten für verkehrliche Anwendungen liegt in der Festlegung eines dafür geeigneten Kostenrahmens. Ausgangspunkt hierfür bilden die grundlegenden Normen zu Kosten und Kostenrechnung (*siehe Kapitel 3*). Die Grundlagen und zeitlichen Veränderungen an den in der industriellen Fertigung eingesetzten Kostenmodellen werden anhand der Arbeiten von [PICKEL, 1989, HILLEBRAND, 1991] und [EITRICH, 1996] motiviert. Da allerdings die Bereitstellung von Geodaten mit keinem Produktionsvorgang von Industriegütern verglichen werden kann, können auch keine produktionsgüterspezifischen Kostenmodelle wie von [PICKEL, 1989, KOELLE, 1971] dargestellt, verwendet werden. Vielmehr müssen eigene, speziell auf den Herstellungsprozess anwendungsspezifischer Geodaten zugeschnittene Kostenmodelle erarbeitet werden.

4.3.1 Grundlagen von Kostenmodellen

Ganz allgemein ist die Kostenermittlung durch Vorkalkulation ein Prognoseproblem. Zur Lösung dieses Problems verwendet die Systemforschung [KRALLMANN ET AL., 1999] ein Prognosemodell, hier entsprechend ein Kostenmodell. Es stellt die empirische Kostenbasis, also die Kostenwirklichkeit in einem Rechnungsmodell dar:

Durch die Modellbildung ist es möglich, die charakteristischen Tatbestände aus dem Unternehmensprozess herauszuheben, um so den komplexen Kausalzusammenhang auf ein vereinfachtes gedankliches Gebilde zu reduzieren. Charakteristisch für die Modellbildung ist, dass die Realität nur symbolisch und unter Beschränkung auf das Wesentliche abgebildet wird. Was bei dieser Abbildung als wesentlich bzw. unwesentlich angesehen wird, hängt jedoch grundsätzlich von der Zwecksetzung des Modells ab.

Das **Kostenprognosemodell** hat nach Modelldefinition (vgl. Kap2, [PICKEL, 1989]) die Aufgabe, Informationen über das zu fertigende Produkt in **Kosteninformationen** umzuwandeln. Es stellt den logischen und mathematischen Zusammenhang zwischen den Eingangsgrößen (variable Merkmale des Kalkulationsobjekts) und den Ausgangsgrößen (Kosten und Kostenanteile) her. Der mathematische Zusammenhang wird durch **Kostenfunktionen** beschrieben. Alle zu fertigenden Produkte (Kalkulationsobjekte) können durch charakteristische Merkmale beschrieben sein. Jedes einzelne Objekt wird folglich durch einen Satz von Ausprägungen dieser Merkmale gekennzeichnet.

4.3.2 Einflussgrößen und deren Zusammenhang in Kostenmodellen

Die Situation im Bereich der Geodatenbereitstellung (vgl. Kapitel 3) zeigt, dass derzeit noch ein hohes Maß an Expertenwissen für die anwendungsgerechte Herstellung und Nutzung von Geodaten notwendig ist. Dies beinhaltet auch die Problematik, dass die Konfektionierung von Geodatenbeständen für weitere Marktsektoren nur unzureichend erfolgen kann. Der Umfang und das Ausmaß an Investitionen zur Aufbereitung, Beschaffung oder Bereitstellung von anwendungstauglichen Geodaten sind in der Definitionsphase neuer, geodatenbasierter Anwendungen nicht, oder nur unzureichend abschätzbar und dann nur durch Experten möglich. Dieses hier entwickelte Kostenmodell soll helfen, die anfallenden Bereitstellungskosten der Geodaten für verkehrliche Anwendungen, wie am Beispiel des Geodatenaufbaus für die intermodale Routenberechnung in Kapitel 3 beschrieben, leichter abschätzen zu können.

Die Grundlage zur Herleitung eines solchen Kostenmodells ist der in Abbildung 4.3 dargestellte prinzipielle Zusammenhang zwischen Kosten, Kostenarten, Kostenparametern und Kostenfunktionen. Die Gesamtkosten, die bei einer Produktion oder Konstruktion anfallen, können einerseits gemäß Kostenarten klassifiziert und andererseits über Kostenfunktionen berechnet werden.

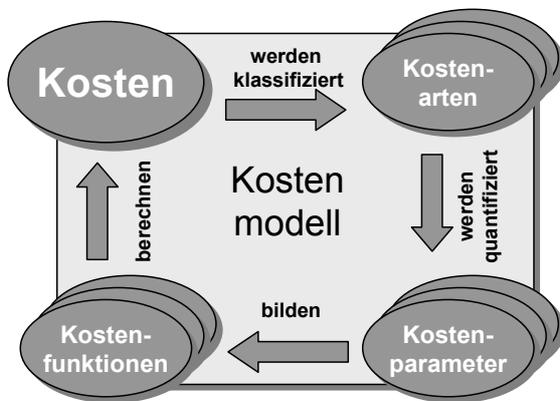


Abbildung 4.3: Grundprinzip Kostenmodell

Da sowohl die Klassifizierung der Kosten über Kostenarten, als auch deren Berechnung über Kostenfunktionen von den Kostenparametern abhängt, stellt die **Ermittlung dieser Parameter** die erste zentrale Herausforderung jeder modellgestützten Kostenprognose dar.

4.3.3 Kategorisierung kostenrelevanter Merkmale

Zur Beschreibung beliebiger Kalkulationsobjekte wäre eine unendliche Anzahl von Merkmalen nötig. Da es aber unmöglich ist, für eine unendliche Zahl von Merkmalen den funktionalen Zusammenhang zur Zielgröße (Kosten) herzuleiten, ist die Objektvielfalt, für die das Kostenmodell gelten soll, durch Clusterung so einzuschränken, dass nur eine begrenzte Zahl von Merkmalen für eine Objektgruppe variabel, ein Großteil der beschreibenden Merkmale aber aufgrund von Ähnlichkeiten für die Objektgruppe konstant ist.

Die variablen Merkmale einer Objektgruppe treten im Kostenmodell explizit auf (**explizite Merkmale**), die konstanten Merkmale sind im formulierten Zusammenhang (Algorithmus) zwischen Eingangsgröße und Zielgröße implizit verarbeitet (**implizite Merkmale**). Das Kostenmodell ist somit für Objekte und deren Varianten gültig, die mit den impliziten Merkmalen des Kostenmodells übereinstimmen und die durch Sätze von Ausprägungen der expliziten Merkmale beschrieben werden können. Dies gilt sowohl für analytisch ableitbare als auch für statistisch ermittelte Zusammenhänge. Eine so exakte Trennung der Merkmale ist in der Praxis jedoch nicht möglich, da man sich nur der variablen Merkmale der Objektgruppe bewusst ist und die Merkmale untereinander korrelieren. Die Übereinstimmung der impliziten Merkmale innerhalb der Gruppe ist deshalb schwer zu überprüfen. Je größer die Abweichung, desto größer wird der Kalkulationsfehler.

Eine sehr wichtige Klassifizierung von Merkmalen ist die **Unterteilung nach dem Skalentyp**. Dieser macht eine Aussage darüber, mit welchem Diskretisierungsgrad die Merkmalsausprägungen messbar sind. Die *Skalentypen* werden in

nicht-metrische Skalen (Nominalskala, Ordinalskala) und *metrische* Skalen (Intervallskala, Proportionalskala) unterteilt. Sie hängen vom unterschiedlichen Informationsgehalt der Daten ab und bieten eine unterschiedliche Anwendbarkeit von Rechenoperationen.

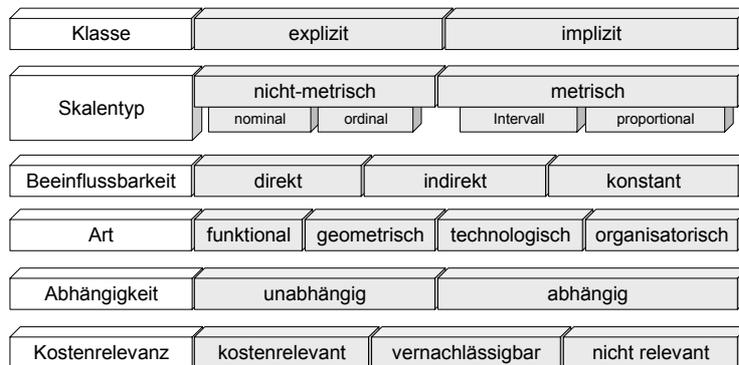


Abbildung 4.4: Kategorisierung kostenrelevanter Merkmale [PICHEL, 1989]

Kennzeichen von nominal skalierten Merkmalen ist die Benennung von qualitativen Eigenschaften zur Unterscheidung von Objekten. Beispiele aus der industriellen Produktion ist z.B. der Werkstoff eines Bauteils (z.B. Stahl Aluminium, Kunststoff) oder seine Hauptgeometrie (rotationssymmetrisch, kubisch). Qualitative Eigenschaften von Geodaten können mit Hilfe von Qualitätsmerkmalen wie *Verfügbarkeit*, *Vollständigkeit*, *Konsistenz*, *Korrektheit* und *Genauigkeit* (metrische und semantische) [WILTSCHKO, 2004] beschrieben werden.

Eine **nominale Skalierung** liegt z.B. im Falle der Zuweisung von einem der Attributwerte *Fußgängerzone* oder *Straßenverkehr* zum Attributtyp *Funktion* des Geometrieobjekts *Straße* vor. Es lassen sich keine arithmetischen Operationen ausführen, sondern lediglich die Gleichheit oder Ungleichheit von Merkmalsausprägungen der Objekte feststellen und durch Zählen ihre Häufigkeit ermitteln. **Ordinal skalierte** Merkmale erlauben zusätzlich die Festlegung einer Rangordnung, beispielsweise die Komplexität eines Bauteils (gering, mittel, hoch) oder die Qualitätseigenschaften eines Werkstoffes (sehr gut, gut, befriedigend, ..). Im Falle des Geometrieobjekts *Straße* [ADV, 1995] entspricht dies der Einteilung in verschiedene Straßenklassen. Ordinal skaliertes Merkmal ist hier der Attributtyp *Widmung* mit den entsprechend der administrativen Zuständigkeit und Ausbaustufe geordnet, unterschiedenen Attributwerten *Bundesautobahn*, *Bundesstraße*, *Landesstraße*, *Kreisstraße*, *Gemeindestraße*, *sonstige Straße*. Es sind ebenso wie bei der nominalen Skala keine arithmetischen Operationen möglich. Die **Intervallskala** stellt das nächst höhere Messniveau dar. Hierbei sind Aussagen über Differenzen sinnvoll, da diese Skala gleich große Abschnitte aufweist. Ein Beispiel aus der industriellen Fertigung ist die Schmelztemperaturskala unterschiedlicher Werkstoffe. Zur metrischen Beschreibung von Geodaten existiert die entsprechende Intervallskala in Form von Koordinatenachsen des räumlichen Bezugssystems. So werden z.B. beim Arbeitsvorgang *flächengleiche Flurstücksteilung* der Katastervermessung die Flurstücksflächen durch eine Intervallskala miteinander verglichen. Addition und Subtraktion von Merkmalsausprägungen sind erlaubt.

Proportional skalierbare Merkmale unterscheiden sich von der Intervallskala dadurch, dass ein natürlicher Nullpunkt existiert, der sich für das Merkmal als „nicht vorhanden“ deuten lässt. Dadurch werden auch Division und Multiplikation von Merkmalsausprägungen, also Aussagen über die zwei miteinander verglichenen Zahlen sinnvoll. Das Verhältnis der Eigenschaften zweier Produkte (z.B. Durchflussmengen oder Einwohnerdichte pro qkm) kann mit der Proportionalskala durch Division hergestellt werden.

Die wichtigste Gliederungsart im Hinblick auf die spätere Anwendbarkeit und Akzeptanz des Kostenmodells ist die Einteilung von Merkmalen bezüglich der **Beeinflussbarkeit der Ausprägungen** durch den Produktentwickler. Die **direkte Beeinflussbarkeit** setzt voraus, dass dem Entwickler die Merkmalsausprägungen bekannt sind. Dies kann sich mit dem Fortschritt des Veredelungsprozesses ändern, d.h. dass beispielsweise die Anzahl der Instanzen einer Objektklasse in der frühen Konzeptphase noch unbekannt ist und diese sich erst im Laufe der Ausarbeitungsphase ermitteln lässt. Vom Strategieentwickler **nicht beeinflussbare Merkmale** sind z.B. die Losgröße (Menge einer Produktart, die zu fertigen bzw. zu beschaffen ist) sowie konstante Merkmalsausprägungen, die durch das Pflichtenheft der Kundenanforderung festgeschrieben sind. Die Losgröße eines Geodatensatzes entspricht dem Volumen des zu fertigenden bzw. zu beschaffenden Datenbestands bzgl. eines bestimmten Objektthemenbereichs. Das Volumen kann durch Parameter wie

z.B. durch die Gesamtsumme der jeweiligen Anzahl von Instanzen einer Objektklasse, durch den Attributierungsgrad eines Attributtyps sowie durch die Anzahl von darin enthaltenen Objektrelationen angegeben werden. Mit der Merkmalbeeinflussbarkeit hängt die Unterscheidung in die **Merkmalearten** funktional (Erfassungsmethoden), geometrisch (räumliche Ausdehnung), technologisch (Genauigkeitsanforderungen) und organisatorisch (Auslastung Betriebsmittel) eng zusammen. Die Objektmerkmale können weiterhin nach ihrer **gegenseitigen Abhängigkeit** unterteilt werden. Die Genauigkeitsanforderungen an Geodaten sind, abgesehen von der Spezifikation im Pflichtenheft des Kunden, **unabhängige** Merkmale, während die zu wählenden Erfassungsmethoden von den Genauigkeitsanforderungen einerseits und dem Automatisierungsgrad des Erfassungsprozesses andererseits **abhängig** sind. Die letzte und zugleich wichtigste Merkmalgliederung für ein Kostenmodell ist die Einteilung nach der **Kostenrelevanz**, also die Beantwortung der Frage, ob die Parameter den Ressourcenverbrauch beeinflussen oder nicht. Dazwischen befinden sich die Merkmale, die die Kosten nur marginal beeinflussen, deren Erfassung aber im Verhältnis dazu unwirtschaftlich erscheint.

4.3.4 Arten von Kostenmodellen

Bei den meisten existierenden Kostenmodellen (Suchkalkulation, Analogieverfahren, etc.) der industriellen Fertigung und der Konstruktion standen der Fertigungsablauf und das betriebliche Kostenrechnungsschema lange Zeit im Hintergrund [PICKEL, 1989; EITRICH, 1996]. Sie verwendeten hauptsächlich geometrische, objektspezifische Merkmale, während die fertigungstechnischen und betriebspezifischen Einflüsse und Randbedingungen als Konstanten in das Kostenmodell integriert wurden. Durch diesen Aufbau waren die Kostenmodelle in ihrer Gültigkeit stark eingeschränkt und nicht in der Lage, für verschiedene andere Entscheidungssituationen Kosteninformationen zu liefern.

Zur Ermittlung entscheidungsrelevanter Kosteninformationen wurde z.B. von [PICKEL, 1989] neben dem mathematisch-objektspezifischen Teil der betriebliche (prozessspezifische) Kostenrechnungsteil sowie ein logischer Teil zur Steuerung mit in den formalen Aufbau des Kostenmodells integriert. Folgende Abbildung zeigt die Methoden zur Ermittlung der Kosteninformationen hinsichtlich objektspezifischer und prozessspezifischer Modellbildung.

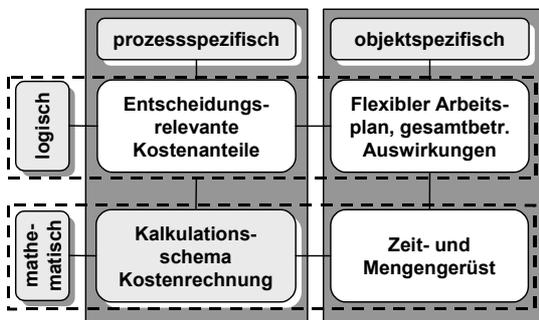


Abbildung 4.5: Einteilung der Kostenmodelle [PICKEL, 1989]

Logische Modelle befassen sich mit logischen Abläufen, z.B. Wenn-Dann-Beziehungen. Mathematische Modelle sind durch die Verknüpfung von Variablen und Konstanten gekennzeichnet (Kostenfunktionen). Mathematisch-logische Modelle sind eine Mischform dieser beiden Arten. Beim Kostenmodell, das für verschiedene Entscheidungssituationen relevante Kostenaussagen liefern soll, liegt eine Mischform zwischen einem mathematischen und einem logischen Modell vor. Das Kostenmodell bildet einen objektspezifischen und einen prozessspezifischen Teil der realen Kostenwirklichkeit ab. Diese Trennung ist sinnvoll, da so die Anforderungen und Aufgaben des Kostenmodells präziser abgegrenzt werden können. Beide Teile können weiter in einen logischen (organisatorischen) und einen mathematischen (funktionalen) Bereich untergliedert werden.

Der objektspezifische Teil des Kostenmodells beinhaltet die festen Parameter des Kalkulationsobjekts sowie die objektspezifischen Randbedingungen und verarbeitet die variablen, objektbeschreibenden Merkmale. Der prozessspezifische Teil des Kostenmodells berücksichtigt die betrieblichen Randbedingungen, die Betriebssituation, die Kosten- und Leistungsrechnung und ist für die Bereitstellung der entscheidungsrelevanten Kosteninformationen verantwortlich. Dementsprechend können die Eingangsgrößen in Merkmale zur Beschreibung von Kalkulationsobjekte (Objektmerkmale), Betriebssituation (Betriebsmerkmale), Erfassungssituation (Erfassungsmerkmale) und Entscheidungssituation (Steuerparameter) gegliedert werden. Als zentrale Herausforderung bei der Erstellung eines Kostenmodells ist die Bildungs-

vorschrift der Kostenfunktionen aus den Kostenparametern zu nennen. Während bei den **objektspezifischen Kostenmodellen** Objekteigenschaften zur Bildung funktionaler Zusammenhänge herangezogen werden, sind bei **prozessspezifischen Kostenmodellen** die funktionalen Zusammenhänge aus den Fertigungs- bzw. Konstruktionsprozessen abzuleiten. Überträgt man dieses Grundprinzip eines Kostenmodells auf die Aufgaben der Geodatenbereitstellung, wie sie in Abschnitt 4.2 allgemein erläutert sind, und berücksichtigt man die Modelleinteilung nach [EITRICH, 1996], so kann ein Kostenmodell, welches das Ziel verfolgt, die Bereitstellungskosten von Geodaten für verkehrliche Anwendungen prognostizieren zu wollen, nur als **mathematisch, prozessspezifisches Kostenmodell** aufgestellt werden.

4.3.5 Prozessspezifisches Kostenmodell der Geodatenbereitstellung

Aufbauend auf den vorangegangenen Abschnitt, der das Grundprinzip von Kostenmodellen erläutert hat, folgt hier dessen Adaption zu einem **Kostenmodell für die Geodatenbereitstellung**. Als Grundlage dient der in Abschnitt 4.2 dargestellte prinzipielle Ablauf der Bereitstellung von Geodaten. Dieser ist in die Hauptprozesse Analyse, Datenerhebung, Integration und Homogenisierung unterteilt.

Da der Zweck und die Notwendigkeit der Geodatenbereitstellung unterschiedlich sein kann, spielt die Vorgehensweise (Veredelung, Aktualisierung, Neuerfassung), wie die Geodatenbereitstellung durchgeführt wird, eine entscheidende Rolle. Die Vorgehensweise, nachfolgend *Aufbaumöglichkeit des Geodatenbestands* genannt, bildet sozusagen den funktionalen Rahmen, welcher die Diskretisierung der Hauptprozesse in Teilprozesse und deren Zusammenstellung zu Kostenfunktionen im Rahmen des Kostenmodells bestimmt. Maßgeblichen Einfluss auf das hier definierte prozessspezifische Kostenmodell haben somit die *gewählte Aufbaumöglichkeit des Geodatenbestands* (siehe Abschnitt 4.4) und die *allgemeinen, natürlichen Kostenarten* (Personal, Material, Kapital und Fremdleistungen). Dieser Zusammenhang ist in nachfolgender Abbildung 4.6 verdeutlicht.

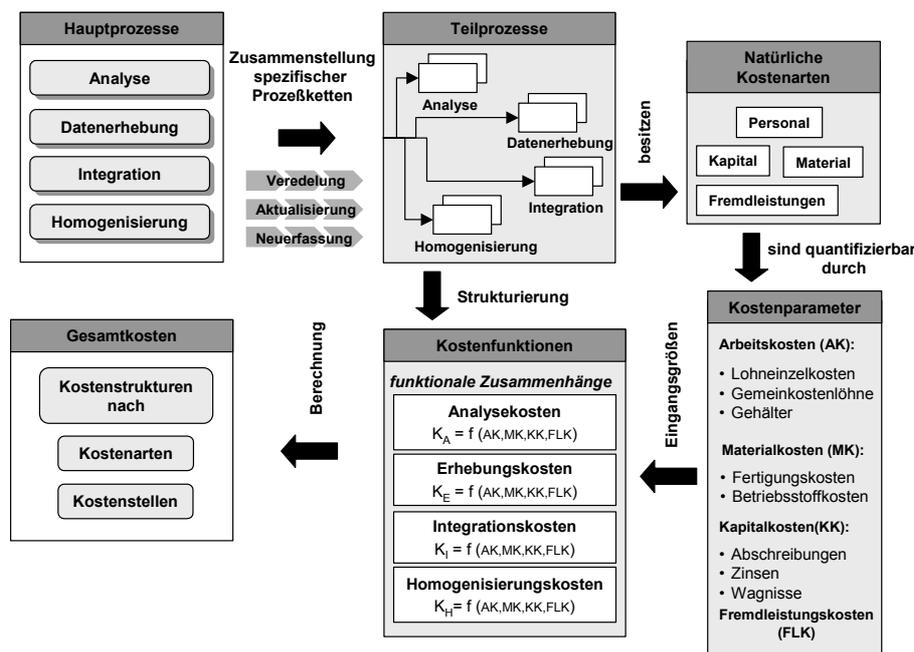


Abbildung 4.6: Kostenmodell für die Bereitstellung von Geodaten

Mit dieser Strukturierung der Teilprozesse und den natürlichen Kostenarten (vgl. [WARNECKE, 1996], [OLFERT, 2001]), die durch die Kostenparameter Personal, Material, Kapital und Fremdleistungen quantifiziert werden, können die allgemeinen vier Kostenfunktionen (K_A, K_E, K_I, K_H), die in der Summe das prozessspezifische Kostenmodell ergeben, zusammengestellt werden. Die **Ergebnisse der Kostenfunktionen** dienen daraufhin zur Aufstellung der **Kostenstrukturen**, welche ihrerseits bezüglich **Kostenarten und Kostenstellen** unterschieden werden können. Mit Hilfe dieser Kostenstrukturen sind anschließend Aussagen über die Wirtschaftlichkeit der betrachteten Aufbaumöglichkeiten zu treffen. Die Strukturierung der Prozesszusammenhänge wird darin als Bildungsvorschrift der funktionalen Zusammenhänge zwischen den natürlichen Kostenarten verwendet.

4.4 Aufbaumöglichkeiten von Geodatenbeständen

Die jeweilige *Aufbaumöglichkeit eines Geodatenbestandes* ist die strukturbestimmende Eingangsgröße des Kostenmodells. Aus diesem Grund ist die Erarbeitung einer Methodik, die zur Herleitung dieser Aufbaumöglichkeit dient, notwendig, um die spezifischen Zusammenhänge zwischen Kostenparametern und Kostenfunktionen herstellen zu können.

Ein Hilfsmittel dieser Methodik ist ein in diesem Abschnitt dargestelltes Entscheidungsdiagramm, mit dessen Hilfe die Aufbaumöglichkeiten von Geodatenbeständen festgelegt werden können. Dieses Entscheidungsdiagramm ist für sämtliche Aufgaben, die unter dem Begriff der Geodatenbereitstellung zusammengefasst sind, anwendbar.

4.4.1 Prinzipielle Aufbaumöglichkeiten

Die Aufbaumöglichkeiten orientiert sich an den beiden prinzipiellen Alternativen des *Zukaufs von Daten* aus Fremdbeständen und der *eigenen Erfassung*. Sie wird darüber hinaus maßgeblich von der Existenz *eigener Datenbestände* beeinflusst.

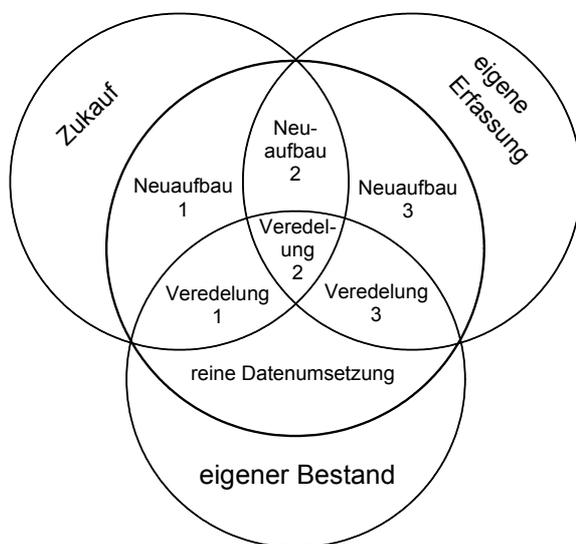


Abbildung 4.7:
Aufbaumöglichkeiten von Geodatenbeständen

Abbildung 4.7 stellt den Zusammenhang zwischen dem aufzubauenden Zieldatenbestand, der dafür notwendigen Arbeiten und der Ausgangssituation graphisch dar. Der Zieldatenbestand kann demzufolge durch 7 unterschiedliche Möglichkeiten realisiert werden. Diese lassen sich in die beiden Gruppen:

- des **Neuaufbaus** und
- der **Veredelung**

von Daten unterscheiden. Die Einteilung in diese Gruppen ist von der Ausgangssituation, also von der Existenz eigener Datenbestände abhängig. Die Aufbaumöglichkeit der *reinen Datenumsetzung* ist ein Sonderfall und wird in nachfolgendem Abschnitt näher beschrieben.

Die **Möglichkeiten des Neuaufbaus** können zur Bereitstellung von Geodaten Verwendung finden, wenn keine eigenen Datenbestände vorhanden sind. Sie sind in drei unterschiedliche Varianten (1,2,3) unterteilt. Die Variante 1 (N1) beschreibt den Neuaufbau unter alleiniger Hilfe von Fremddatenbeständen mittels Zukauf. Variante 2 (N2) geht von einer Kombination des Zukaufs und eigener Erfassungsleistung aus, während die Variante 3 (N3) den Neuaufbau des Datenbestands vollständig durch eigene Erfassungsleistungen realisiert.

Die **Möglichkeiten der Veredelung** greifen auf prinzipiell nutzbare eigene Geodatenbestände zurück und sind ebenso wie die Neuaufbaumöglichkeiten in drei unterschiedliche Varianten zu unterscheiden. Die erste Variante (V1) geht von der Bereitstellung des Zieldatenbestandes von einer Zusammenführung (Integration /Homogenisierung) des eigenen Bestandes mit Fremdbeständen mittels Zukauf aus. Variante 2 (V2) stellt den Zieldatenbestand durch eigene Daten, den Zukauf von Fremddaten und eigener Erfassungsleistung sicher. Die Veredelungs-Variante 3 (V3) realisiert den Zieldatenbestand über eine Kombination des eigenen Bestands mit eigener Erfassungsleistung.

4.4.2 Entscheidungsdiagramm zur Herleitung der Aufbaumöglichkeiten

Aufbauend auf der Beschreibung der verschiedenen Möglichkeiten der Bereitstellung von Geodaten wird in diesem Abschnitt ein Entscheidungsdiagramm (Abbildung 4.8) vorgestellt, welches zur Herleitung der problemspezifischen Aufbaumöglichkeiten als Hilfe herangezogen werden kann. Die graphische Darstellung ist an die UML - Modellierungsvorschrift [RUMBOUGH, 1993], [OESTERREICH, 1997] des Zustandsdiagramms angelehnt.

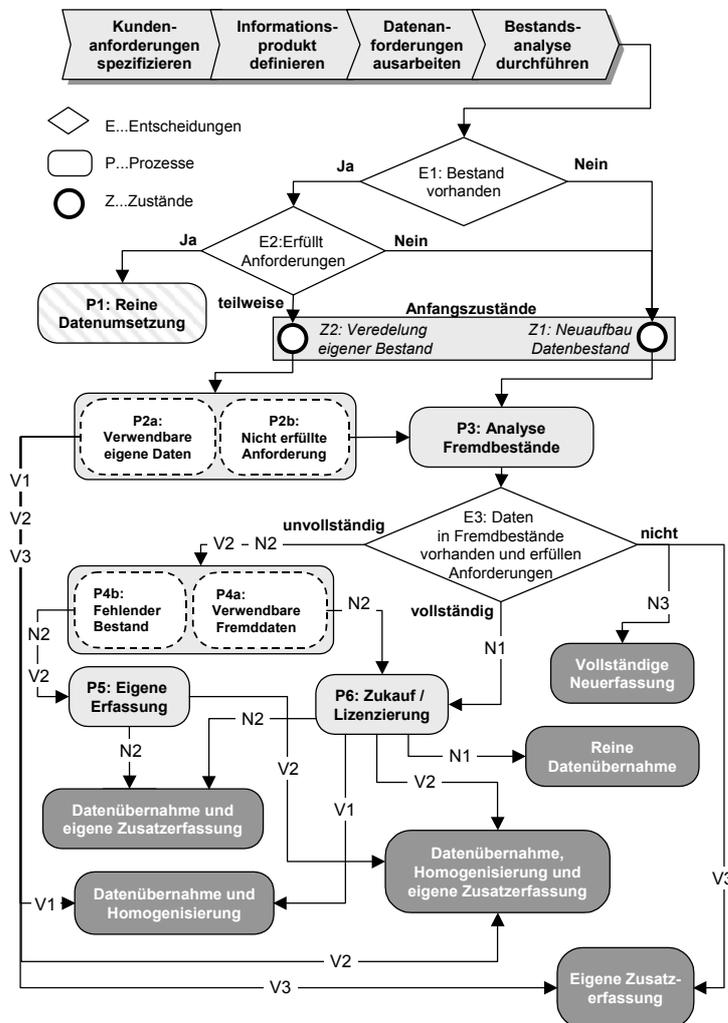


Abbildung 4.8: Entscheidungsdiagramm für Aufbaumöglichkeiten von Geodatenbeständen

Dieses Entscheidungsdiagramm ist aus Sicht eines Unternehmens zu sehen, welches für ein geplantes neues Informationsprodukt von Ausgangssituationen (hier: *Anfangszuständen*) ausgehen und *Entscheidungen* treffen muss, sowie Arbeiten in Form von *Prozessen* durchzuführen hat. Die Aufbaumöglichkeiten werden hier unabhängig von der Arbeitszuweisung, etwa auf unternehmensinterne Abteilungen oder Fremdfirmen in Form von Subunternehmern allgemeingültig aufgestellt.

Am Anfang steht der in 4 Teilprozesse differenzierte *Analyseprozess*. Dieser ist in die Spezifikation der Kundenanforderungen, die Definition des Informationsproduktes, die Ausarbeitung der Datenanforderungen und in die Durchführung der Bestandsanalyse unterteilt. Das Ergebnis des Analyseprozesses lässt eine erste Unterscheidung gemäß der in Abschnitt 4.4.1 definierten Gruppeneinteilung zu. Sie werden als die Anfangszustände:

- Z1: Neuaufbau des Datenbestands und
- Z2: Veredelung des eigenen Bestandes

der Aufbaumöglichkeiten bezeichnet. Der erste Ausgangszustand (Z1), *Neuaufbau des Datenbestands*, resultiert aus den Entscheidungen E1 und E2 für die Fälle, wenn entweder kein eigener Bestand vorhanden ist oder der vorhandene die

spezifizierten Anforderungen nicht erfüllt. Der zweite Ausgangszustand (Z2), *Veredelung des eigenen Bestandes*, ergibt sich, wenn der eigene Datenbestand die neuen Anforderungen nur zum Teil erfüllt.

Im Sinne der Vollständigkeit des Entscheidungsbaums ist der Prozess der **reinen Datenumsetzung**, (P1) (vgl. auch Abbildung 4.7), als Folge von erfüllten Anforderungen durch den eigenen Datenbestand (E2) mit aufgeführt. Als reine Datenumsetzung wird dabei nur die Übertragung von Geodaten aus einer bestehenden Datenbankstruktur in eine andere verstanden, wobei von eindeutigen Abbildungsvorschriften der Dateninhalte ausgegangen wird. Von einer reinen Datenumsetzung kann nur dann gesprochen werden, wenn keinerlei Veränderungen am Datenmodell, der Datenstruktur und dem Inhalt (Objekte, Attribute und Relationen) vorgenommen werden. Das Fehlen wichtiger Attribute oder Objektrelationen wird bei der Entscheidung E2 beurteilt und mündet, je nach Ergebnis, in einem der beiden Ausgangszustände Z1 oder Z2. Die Herleitung der verbleibenden 6 Aufbaumöglichkeiten mit Hilfe des Entscheidungsdiagramms ist in folgenden Abschnitten, entsprechend der Anfangszustände, dargestellt.

4.4.3 Varianten des Neuaufbaus von Geodatenbeständen

Wurde der Anfangszustand Z1 festgestellt, so erfolgt daraufhin die Suche nach Fremddatenbeständen (P3) sowie die damit verbundene inhaltliche Analyse und Aufwandsbewertung mit dem Ziel, sich diese Bestände nutzbar zu machen. Die Entscheidung (E3) *-Daten in Fremdbeständen vorhanden-* differenziert zwischen *vorhandenen (Ja)*, *teilweise oder nicht vorhanden (Nein)*. Diesbezüglich werden die nachfolgenden Varianten unterschieden. Die Problematik, dass benötigte Daten in verfügbaren Fremdbeständen zwar prinzipiell vollständig vorhanden sind, diese aber nicht ohne weitere Konvertierungsmaßnahmen sofort und vollständig nutzbar gemacht werden können, ist eine Aufgabe der Integration und Homogenisierung. Sie hat keinen Einfluss auf die grundlegende Existenz der 6 Aufbaumöglichkeiten, kann aber unter Umständen zu einem Wechsel der Vorgehensweise führen, wenn im Zuge der Integration und Homogenisierung erkannt wird, dass eine Konvertierung der prinzipiell vorhandenen Daten in prinzipiell nutzbare nicht mit demselben Aufwand (Kosten und Arbeit), wie durch eine Neuerfassung realisiert werden kann.

Neuaufbau 1: „Vorhandene Daten in Fremdbeständen“

Bei der Aufbaumöglichkeit des *Neuaufbaus 1* geht man von vollständig nutzbaren Daten in vorhandenen Fremdbeständen aus, die mittels Zukauf und ggf. Lizenzierung bereitgestellt werden können.



Abbildung 4.9: Neuaufbau 1

Es findet eine **reine Datenübernahme** statt. Eine eigenständige Erfassung zusätzlicher Daten und Informationen findet nicht statt. Der Aufbau des applikationsspezifischen Datenbestands geschieht durch reine Integration und Homogenisierung bereits vorhandener Daten. Das Datenmodell steht in diesem Fall fest. Mögliche Änderungen oder Erweiterungen sind mit einem hohen Aufwand verbunden und werden in der Regel nicht durchgeführt.

Neuaufbau 2: „Unvollständig vorhandene Daten in Fremdbeständen“

Bei der Aufbaumöglichkeit des *Neuaufbaus 2* findet sowohl der Zukauf und Lizenzierung von Fremdbeständen als auch eine eigenständige Erfassung fehlender Daten statt. Sie führt deshalb zur Aufgabe der **Datenübernahme und eigener Zusatzerfassung**.

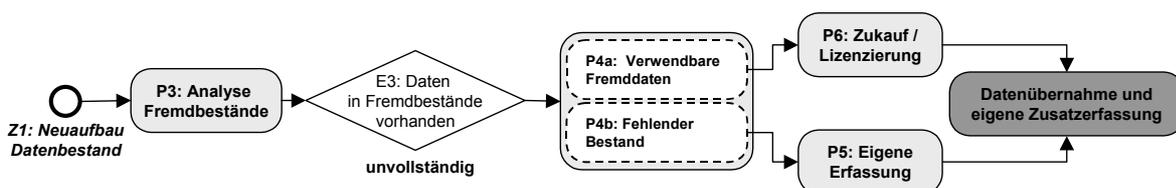


Abbildung 4.10: Neuaufbau 2

Bei dieser Aufbaumöglichkeit müssen geeignete, speziell auf die fehlenden Daten ausgerichtete, Erfassungsverfahren eingesetzt werden. Das Datenmodell kann in der Regel vom Fremddatenbestand übernommen werden. Eine Änderung bzw. Anpassung des Fremddatenmodells kann hierbei nicht vollständig ausgeschlossen werden. Die notwendige Modelländerung ergibt sich aus den Anforderungen der Anwendung. Das eingesetzte Erfassungsverfahren für die eigene Zusatzerfassung ist dann entsprechend auf das Modell anzupassen.

Neuaufbau 3: „Keine nutzbaren Fremdbestände“

Die Aufbaumöglichkeit des *Neuaufbaus 3* geht davon aus, dass keinerlei Daten in Fremdbeständen vorhanden sind, die für die Anwendung genutzt werden können. Eine vollständige Neuerfassung der Daten wird notwendig.

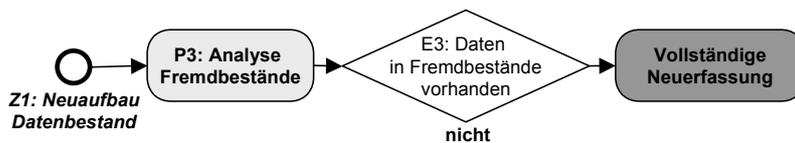


Abbildung 4.11: Neuaufbau 3

Das Datenmodell muss entsprechend den Anforderungen der Anwendung neu aufgestellt werden. Der diesbezügliche Aufwand ist als sehr hoch zu bezeichnen. Diese Aufbaumöglichkeit entspricht dem Vorgang der Datenersterfassung. Hierbei sind alle Schritte von der Anforderungsanalyse, der Spezifikation des Informations- und Datenmodell, der prototypischen Bestandserfassung, der Überprüfung und Bewertung der gewählten Erfassungsmethode bis hin zur vollständigen Datenerfassung durchzuführen. Eine Integration und Homogenisierung im Sinne der Zusammenführung zweier auf unterschiedlichen Datenmodellen basierenden Fremdbeständen findet hier nicht statt.

4.4.4 Varianten der Veredelung eigener Bestände

Ausgehend vom Anfangszustand *Z2 – Veredelung eigener Bestand* wird zuerst der eigene Bestand bezüglich *verwendbare eigenen Daten* (P2a) untersucht und daraus die *nicht erfüllten Anforderungen* (P2b) abgeleitet. Die Analyse der Fremddatenbestände (P3) findet daraufhin speziell vor dem Hintergrund der nicht erfüllten Anforderungen statt. Es wird konkret untersucht, ob fehlende Daten prinzipiell aus Fremdbeständen entnommen werden können. Die Entscheidung (E3) differenziert hier zwischen *verwendbaren*, *unvollständig verwendbaren* und *nicht verwendbaren* Daten aus Fremdbeständen.

Veredelung 1: „Verwendbare Daten in Fremdbeständen“

Bei der Aufbaumöglichkeit der *Veredelung 1* wird von vollständig verwendbaren Daten in Fremdbeständen ausgegangen. Hier kommt es zum Vorgang der Datenübernahme mit anschließender Integration oder Homogenisierung mit den vorhandenen eigenen Beständen.

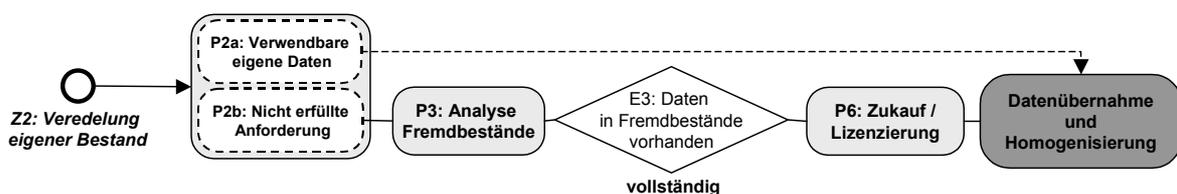


Abbildung 4.12: Veredelung 1

Die Beibehaltung des eigenen Datenmodells ist wahrscheinlich, kann aber nicht grundsätzlich vorausgesetzt werden. Zentrale Aufgabe dieser Aufbaumöglichkeit ist die Integration und Homogenisierung des eigenen Bestands mit Daten aus Fremdbeständen unter vorhergehender Eignungsprüfung des verwendeten Datenmodells.

Veredelung 2: „Unvollständig verwendbare Daten in Fremdbeständen“

Die Aufbaumöglichkeit der *Veredelung 2* ist bezüglich Arbeitsumfang eine der aufwendigsten Vorgehensweisen. Neben den eigenen Beständen werden Fremdbestände zugekauft und darüber hinaus die dann noch fehlenden Daten durch eigene Erfassung zu einem Zieldatenbestand zusammengeführt. Es handelt sich um die **Datenübernahme mit Integration / Homogenisierung und eigener Zusatzerfassung**.

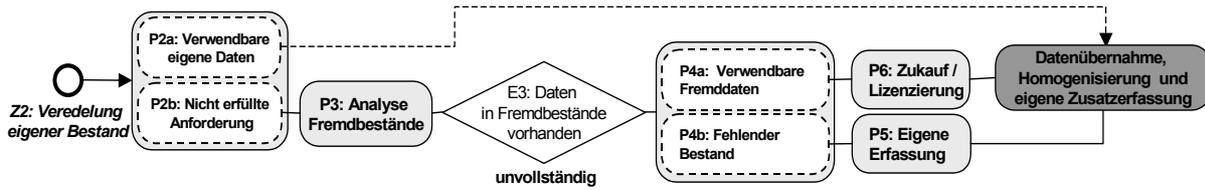


Abbildung 4.13: Veredelung 2

Das Datenmodell des Zieldatenbestandes kann, wie bei der *Veredelung 1*, in der Regel aus dem eigenen Modell übernommen oder abgeleitet werden. Dies ist aber auch nicht als allgemeingültig anzunehmen. Die Aufgaben der Integration und Homogenisierung sind ebenfalls mit denen der Veredelung 1 vergleichbar. Darüber hinaus müssen geeignete Erfassungsverfahren eingesetzt werden, damit deren Ergebnisse unmittelbar in den Integrations- und Homogenisierungsvorgang einfließen können.

Veredelung 3: „Nicht verwendbare Daten in Fremdbeständen“

Sind in Fremdbeständen keine verwendbaren Daten enthalten, so sind die fehlenden Daten hier durch geeignete Erfassungsverfahren bereitzustellen.

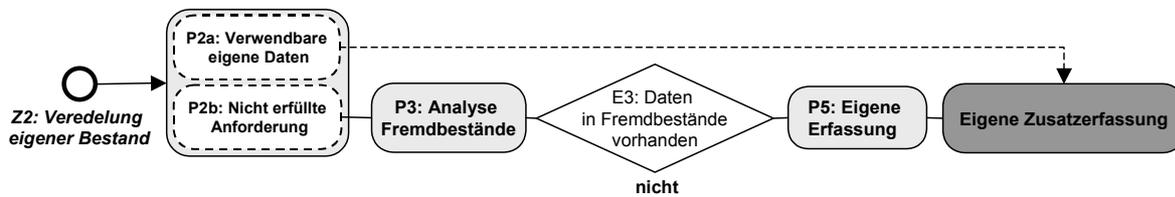


Abbildung 4.14: Veredelung 3

Eine Integration und Homogenisierung findet nicht statt, unter Voraussetzung, dass die Erfassungsverfahren auf das existierende Datenmodell angepasst wurden. Es handelt sich hier um die reine **Eigene Zusatzerfassung**.

4.5 Die Kostenstufen der Aufbaumöglichkeiten

Da die Aufbaumöglichkeiten von unterschiedlichen, nacheinander folgenden Teilprozessen geprägt sind, die in ihrer Anzahl und Kostenauswirkung stark unterschiedlich sein können, werden die Aufbaumöglichkeiten in diesem Abschnitt als Entwicklungsstufen mit den zugehörigen Kosten dargestellt. Durch diese Darstellungsform wird gleichzeitig die damit verbundene Qualitätsverbesserung deutlich.

Neuaufbau 1:

Reiner Datenzukauf

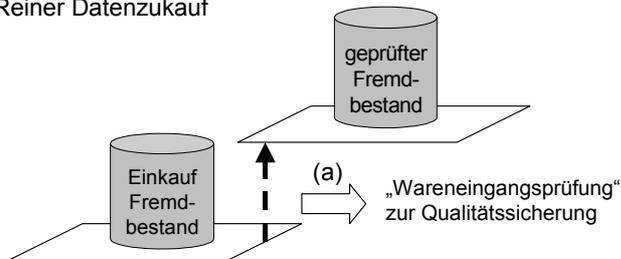


Abbildung 4.15: Kostenstufe Neuaufbau 1

Im Fall des **Neuaufbaus 1** (Abbildung 4.15), dem reinen Datenzukauf, erfolgt nach dem Dateneinkauf des Fremdbestandes die Qualitätssicherungsmaßnahme in Form einer „Wareneingangsprüfung“ (a). Damit existieren für diese Möglichkeit die Teilprozesskosten des Einkaufs K_{Ein} , die Kosten des Prüfvorgangs $K_{Prüf}^{Ein}$ sowie die damit verbundenen Maßnahmen zur Fehlerverbesserung $K_{Verbess}^{Ein}$. Es ergibt sich folgende allgemeine Kostenformel des Neuaufbaus 1:

$$K_{N1} = K_{Ein} + K_{Prüf}^{Ein} + K_{Verbess}^{Ein} \quad (4.1)$$

Geht man hingegen vom **Neuaufbau 2**, dem Zukauf und der eigenen Erfassungsleistung aus (Abbildung 4.16), dann wird der geprüfte Fremdbestand auf Grund der Forderung nach räumlicher oder inhaltlicher Erweiterung des Datenbestandes (b) in einen vorerst ungeprüften, veredelten Bestand überführt.

Neuaufbau 2:

Zukauf und eigene Zusatzerfassung

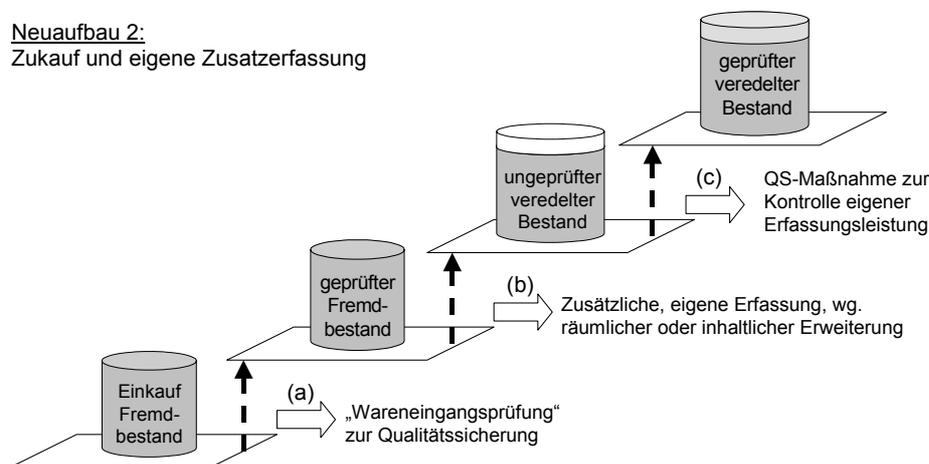


Abbildung 4.16:
Kostenstufen Neuaufbau 2

Nach Abschluss der eigenen Zusatzerfassung erfolgt die Qualitätssicherungsmaßnahme zur Kontrolle der eigenen Erfassungsleistung (c). Diese beschränkt sich auf den eigenen, nacherfassten Datenumfang. Die allgemeinen Gesamtkosten ergeben sich aus den Teilprozesskosten des Neuaufbaus 1 sowie den Kosten der eigenen Erfassungsleistung K_{EigErf} , den Kosten der Prüfung $K_{Prüf}^{EigErf}$ sowie den Kosten der Verbesserung $K_{Verbess}^{EigErf}$ zu:

$$K_{N2} = K_{Ein} + K_{Prüf}^{Ein} + K_{Verbess}^{Ein} + K_{EigErf} + K_{Prüf}^{EigErf} + K_{Verbess}^{EigErf} \quad (4.2)$$

Für den **Neuaufbau 3**, also der vollständigen Neuerfassung (Abbildung 4.17), sind die Teilprozesse der Anforderungsanalyse, dem Systemkonzept bis hin zur Modellierung inklusive der Erarbeitung von Erfassungsvorschriften vollständig anzusetzen (d).

Neuaufbau 3:
Vollständige Neuerfassung

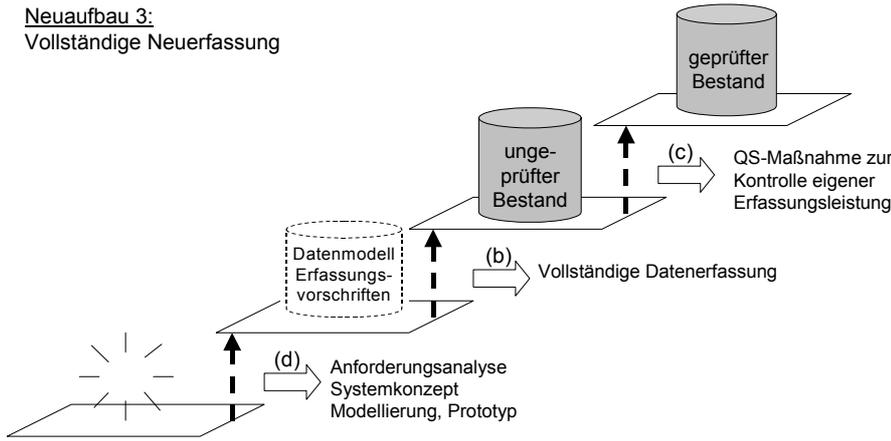


Abbildung 4.17:
Kostenstufen Neuaufbau 3

Diese Einzelvorgänge sollen hier durch den Kostenfaktor der Vorbereitung der vollständigen Neuerfassung K_{Vorb} repräsentiert werden. Auf die Kostenermittlung wird im folgenden Abschnitt näher eingegangen. Unter Berücksichtigung der Kosten der vollständigen eigenen Erfassungsleistung mit dem Kostenfaktor K_{EigErf} und den QS-Maßnahmen zu deren Kontrolle ergibt sich die folgende Kostenfunktion des Neuaufbaus 3:

$$K_{N3} = K_{Vorb} + K_{EigErf} + K_{Prüf}^{EigErf} + K_{Verbess}^{EigErf} \tag{4.3}$$

Im Falle einer **Veredelung 1** (Abbildung 4.18), also dem Zukauf von Fremddaten und der notwendigen Integration und Homogenisierung mit dem eigenen Bestand, erfolgt für beide Datenbestände unabhängig die Qualitätssicherungsmaßnahmen (a), bevor mit der Integration bzw. Homogenisierung der beiden geprüften Bestände ($K_{Int\&Hom}$) begonnen wird.

Veredelung 1:
Zukauf + eigener Bestand

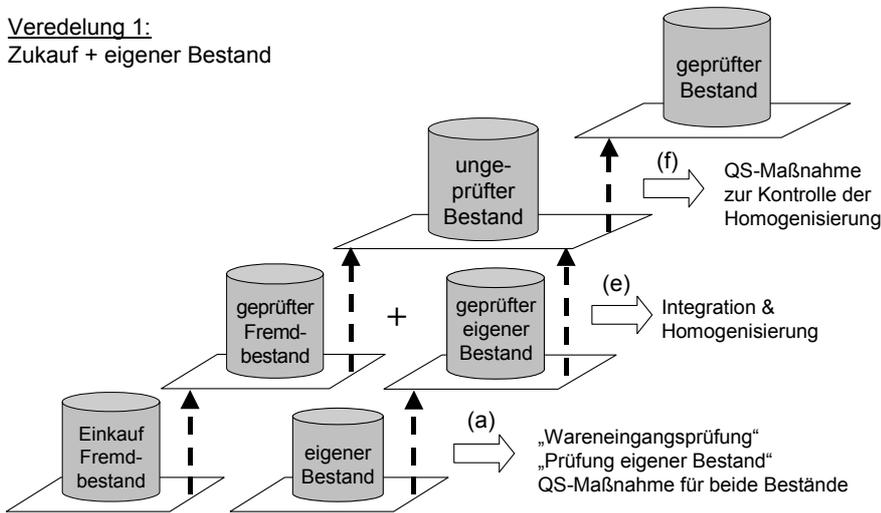


Abbildung 4.18:
Kostenstufen Veredelung 1

Eine abschließende Qualitätssicherungsmaßnahme zur ausschließlichen Kontrolle der Homogenisierung garantiert einen geprüften und homogenisierten Datenbestand. Es ergeben sich die Kosten zu:

$$K_{V1} = \{ K_{Ein} + K_{Prüf}^{Ein} + K_{Verb}^{Ein} \} + \{ K_{Prüf}^{EigBest} + K_{Verbess}^{EigBest} \} + K_{Int\&Hom} + K_{Prüf}^{Int\&Hom} + K_{Verbess}^{Int\&Hom} \tag{4.4}$$

Veredelung 2:
Zukauf + eigener Bestand +
eigene Zusatzerfassung

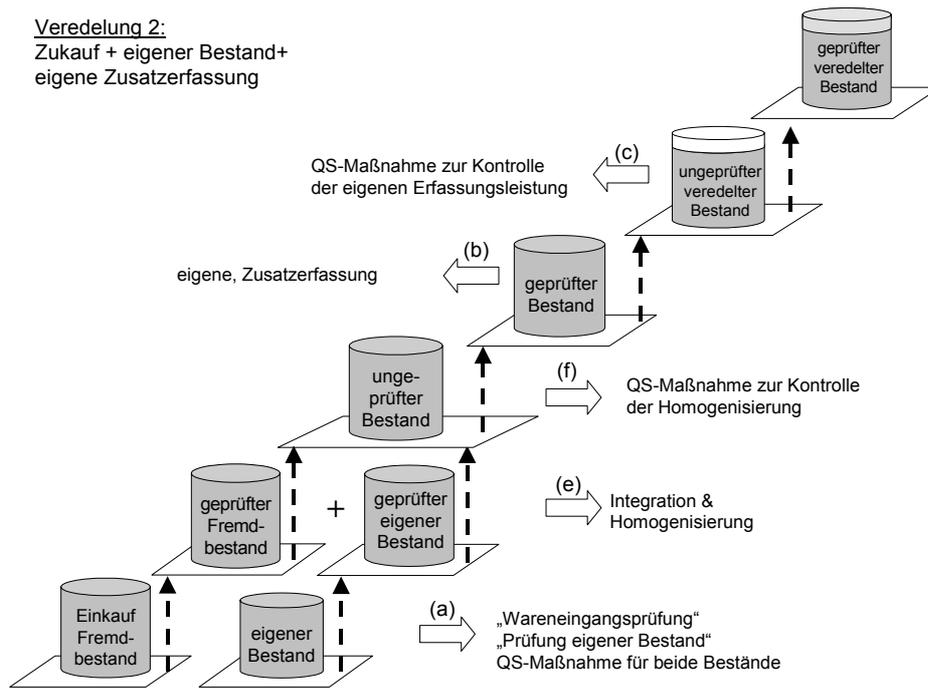


Abbildung 4.19:
Kostenstufen Veredelung 2

Erweitert man die Veredelung 1 um die zusätzliche eigene Erfassungsleistung (b), dann ergibt sich daraus die Möglichkeit der **Veredelung 2** (Abbildung 4.19). Konkret bedeutet dies eine Erweiterung der Kostenfunktion (4.4) um die Kostenfaktoren der eigenen zusätzlichen Erfassung (b) und der Qualitätssicherungsmaßnahme zur Kontrolle der eigenen Erfassungsleistung (c). Die Kostenfunktion der Veredelung 2 lautet daher:

$$K_{V2} = K_{V1} + K_{EigErf} + K_{Prüf}^{EigErf} + K_{Verbess}^{EigErf} \tag{4.5}$$

Aus der Anzahl der Kostenfaktoren aus (4.5) kann hier allerdings nicht automatisch auf die betragsmäßig größten Kosten geschlossen werden.

Veredelung 3:
eigener Bestand +
eigene Zusatzerfassung

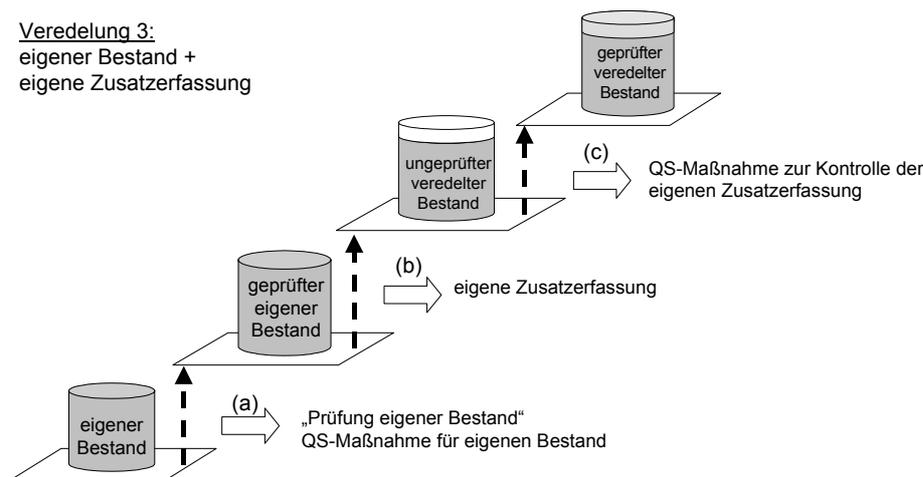


Abbildung 4.20: Kostenstufen Veredelung 3

Verwendet man den existierenden eigenen Bestand weiter und veredelt diesen durch eigenen zusätzliche Erfassungsleistung (b) (**Veredelung 3**) so erhält man die Kostenfunktion über die Summe der Qualitätssicherungskosten des eigenen Anfangsbestandes $K_{Prüf}^{EigBest}$, den Kosten zur Fehlerverbesserung $K_{Verbess}^{EigBest}$, den Kosten der eigenen Erfassungsleistung K_{EigErf} , sowie den Kosten der Qualitätssicherung der eigenen Erfassungsleistung $K_{Prüf}^{EigErf}$ und $K_{Verbess}^{EigErf}$:

$$K_{V3} = K_{Prüf}^{EigBest} + K_{Verbess}^{EigBest} + K_{EigErf} + K_{Prüf}^{EigErf} + K_{Verbess}^{EigErf} \quad (4.6)$$

Die in diesen Formeln dargestellten Kostenfaktoren, nämlich:

1. der Zukauf von Fremddaten,
2. die eigene Zusatzerfassung,
3. die Vollständige Erfassung,
4. die Integration und Homogenisierung, sowie
5. die Qualitätssicherungsmaßnahmen

und deren Ermittlungsmethodik sind zentraler Gegenstand des in Abbildung 4.6 dargestellten Kostenmodells. Die Bewertung des Kostenfaktors von Qualitätssicherungsmaßnahmen bei Geodaten wird auf Grund ihrer Komplexität ausführlich in Kapitel 5 betrachtet.

Bei den hier betrachteten Kostenstufen werden die Maßnahmen zur Qualitätssicherung explizit pro Einzelstufe angesetzt und nicht als eine Art ganzheitlichen Qualitätssicherungsmaßnahme am Ende der jeweiligen Aufbaumöglichkeit betrachtet. Mehrere stufenspezifische QS-Maßnahmen besitzen gegenüber einer ganzheitlichen QS-Maßnahme die Vorteile, dass Qualitätsverletzungen frühzeitig erkannt und korrigiert werden können und es damit nicht zu einer Verschleppung oder gar zu einer Fortpflanzung möglicher fehlerhafter Elemente kommt. Bei einer ganzheitlichen QS-Maßnahme ist es ungleich schwieriger die Fehlerursache zurückzuverfolgen, da man in einem solchen Fall sämtliche Stufen rückwärts durchlaufen müsste, um die Fehlerquelle identifizieren zu können. So hat es sich z.B. durch den im Rahmen des Aufbaus der digitalen Straßenkarte für Europa (EDRM) durchgeführten Benchmark-Test herausgestellt, dass stufenweise Qualitätssicherungsmaßnahmen wesentliche Kosten- und Qualitätsvorteile gegenüber ganzheitlichen, am Ende durchgeführten QS-Maßnahmen besitzen [KRONJÄGER, 1992].

Eine ganzheitliche QS-Maßnahme würde nur im direkten Kostenvergleich mit den einzelnen stufenspezifischen QS-Maßnahmen einen Vorteil bringen. Jedoch die Folgekosten, die ein unentdecktes fehlerhaftes Element bei einer ganzheitlichen QS-Maßnahme nach sich zieht, übersteigen die Kosten bei den einzelnen stufenspezifischen Maßnahmen im ungünstigsten Fall um das Doppelte [MÖHLENBRINK, 1992].

4.6 Das Berechnungsverfahren zur Kostenprognose

Als Grundlage zur Ermittlung der Kosten der Erfassungsleistungen dient in diesem Abschnitt einerseits die Honorarordnung für Architekten und Ingenieure (HOAI) sowie die speziell für die Erhebung von Geodaten formulierte Arbeitshilfe für die vermessungstechnischen Leistungen (AHO) der HOAI. Diese werden in den nachfolgenden Abschnitten als Grundlage des Kostenmodells und zur Berechnung der Erfassungsleistungen erläutert.

4.6.1 Grundlagen der HOAI und AHO

Als Grundlage dient in dieser Arbeit der, in der HOAI festgelegte Rahmen zur Leistungsverrechnung in der aktuellen Version vom 01. Januar 1996 mit den Änderungen zur Umstellung der HOAI auf Euro, veröffentlicht im Bundesgesetzblatt 2001 [BGBl, 2001] vom 14. November 2001 (Teil 1, Nr.58, Art. 5). Die HOAI ist eine Rechtsverordnung mit Gesetzescharakter.

Nach § 2 Leistungen der HOAI gliedern diese sich in Grundleistungen und Besondere Leistungen:

§ 2, Abs. 2: Grundleistungen umfassen die Leistungen, die zur ordnungsgemäßen Erfüllung eines Auftrages im allgemeinen erforderlich sind. Sachlich zusammengehörige Grundleistungen sind zu jeweils in sich abgeschlossenen Leistungsphasen zusammengefasst.

§ 2, Abs. 3: Besondere Leistungen können zu den Grundleistungen hinzu oder an deren Stelle treten, wenn besondere Anforderungen an die Ausführung des Auftrags gestellt werden, die über die allgemeinen Leistungen hinausgehen oder diese ändern. Sie sind in den Leistungsbildern nicht abschließend aufgeführt. Die Besonderen Leistungen eines Leistungsbildes können auch in anderen Leistungsbildern oder Leistungsphasen vereinbart werden, in denen sie nicht aufgeführt sind, soweit sie dort nicht Grundleistungen darstellen.

Die in der HOAI aufgeführten jeweiligen Leistungen gliedern sich in verschiedene Themenbereiche. Der für diese Arbeit relevante Themenbereich *Vermessungstechnische Leistungen* findet sich in Teil XIII (§96-§100). Die darin beschriebenen Grundleistungen sind in insgesamt 7 Leistungsphasen unterteilt.

Diese Aufgaben repräsentieren die **Grundleistungen** zur Einführung eines Geoinformationssystems innerhalb eines Unternehmens sowie zum Aufbau einer Datengrundlage. Sie sind in ihren wesentlichen Arbeitsschritten mit denen in [BEHR, 1998] dargestellten Komponenten des GIS-Managements identisch. Über diese Grundleistungen hinaus werden so genannte „Besondere Leistungen“ definiert, welche je nach Aufgabenstellung, Anforderungskatalog und Schwierigkeitsgrad variieren können.

Table 4-1: Die sieben Leistungsphasen der HOAI-AHO

1. Grundlagen-ermittlung	Einholen von Informationen zur Klärung der Aufgabestellung. Ermitteln des Leistungsumfangs in Abhängigkeit von den Rahmenbedingungen
2. Ist-Analyse	Erheben der vorhandenen Karten- und Planwerke. Begutachten der vorhandenen Bestandsdokumentation. Ermittlung der Organisationsstruktur, der Informationserzeugung und -wege. Ermitteln der Nutzer, des Umfangs, der Häufigkeit der bisherigen Nutzung. Aufstellen Schwachstellenkatalogs, Ermitteln der Mengengerüste.
3. Soll-Konzept	Erarbeitung eines Soll-Konzeptes. Formulieren der Anforderungen. Entwicklung von Zielvorstellungen. Formulieren der Ansprüche an die Geometrie und Sachdaten und schließlich die Abstimmung des Sollkonzepts mit dem Auftraggeber.
4. Datenmodell	Definieren eines Datenmodells durch Entwerfen einer logischen Objektstrukturierung, Aufstellen der Schlüsselarchitektur, Darstellen der Datenrelationen, Entwerfen von Referenztabellen, Festlegung von Plausibilitätskontrollen und das Erstellen eines Konzepts zur integrierten Verarbeitung von Geometrie und Sachdaten.

5. Systemkonzept	Konzipieren und Planen des Informationssystems durch Beschreiben der Umgebung, der Datenbankkonzeption, der Benutzerorganisation, der Rechner- und Betriebssystemarchitektur, der Datensicherung und des Datenschutzes, sowie der Datenkommunikation.
6. Kostenanalyse / Terminplan	Durchführen einer Kostenanalyse durch Bewerten der konventionellen Planerstellung und Karteiverwaltung. Bewerten der digitalen Informationsverarbeitung. Aufstellen eines Terminplans
7. Datenerfassung und Bestandsmodell	Erstellen des geometrischen Bestandsmodells durch graphische Digitalisierung und/oder Konstruktion. Übergabe des Bestandsmodells in einfacher analoger Form und auf Datenträger.

Bei diesen Besonderen Leistungen handelt es sich im wesentlichen um:

- *Die Erarbeitung von einer oder mehreren Varianten / Alternativen zu dem Sollkonzept. Der Festlegung von Auswirkungen auf die Unternehmensorganisation und die Arbeitsabläufe. Die Erarbeitung eines Konzepts für ein Testgebiet sowie die Durchführung des Feldvergleichs und der Einmessarbeiten und die Überarbeitung der fachbezogenen Einträge und das Erstellen der Digitalisierungsvorlage für das Testgebiet.*
- *Das Vorbereiten und Begleiten der Systemausschreibung für Hard- und Software, die Mitwirkung bei der Vergabe und bei der Systemeinführung sowie das Erstellen eines Schulungskonzepts.*
- *Die Aufbereitung der Vorlagen für die Digitalisierung, das Berichtigen der zur Verfügung gestellten Ausgangsdaten, die graphische Aufbereitung und Ausgabe analoger Karten- und Planwerke, Karteien, Listen und Verzeichnisse aus dem Bestandsmodell.*
- *Die Ableitung von thematischen Karten aus den Informationen des Bestandsmodells, die Schaffung der Kompatibilität zu anderen Informationssystemen oder Datenbeständen und Datenübernahme.*
- *Das Erfassen weiterer fachspezifischer alphanumerischer Daten sowie die Übernahme eines auftraggeberspezifischen Datenmodells mit Symbolbibliothek und Prozeduren.*

Die Kosten bzw. Honorarermittlung der **besonderen Leistungen** sind nicht über die Ermittlungsverfahren der Honorare für die Grundleistungen abgedeckt, sondern müssen gesondert vertraglich vereinbart werden. Hierfür eignet sich die Abrechnung nach dem Zeithonorar, worin die minimal und maximal zu verrechnenden Stundensätze des Auftragnehmers (€ 38,00- €82,00), seiner technischen Mitarbeiter (€ 36,00- €59,00) und seiner technischen Zeichner, bzw. Mitarbeiter mit vergleichbarer Qualifikation (€ 31,00- €43,00) festgelegt sind.

Das Honorar für die 7 beschriebenen Phasen der Grundleistungen richtet sich nach den anrechenbaren Verrechnungseinheiten und teilt sich entsprechend den folgenden Prozentsätzen auf die ersten 6 Phasen auf. Die *Phase 7: Datenerfassung und Bestandsmodell* erfolgt in einer getrennten Abrechnung.

- Grundlagenermittlung 3%
- Ist-Analyse 23%
- Soll-Konzept 27%
- Datenmodell 20%
- Systemkonzept 18%
- Kostenanalyse / Terminplan 9%

In Abschnitt 4 des „*Leistungsbilds und Honorierung von geographisch-geometrischen Datenbasen für raumbezogene Informationssysteme*“ der Arbeitshilfe [AHO, 2001] werden die anrechenbaren Verrechnungseinheiten für die Grundleistungen in Abhängigkeit von der Gesamtfläche des Bearbeitungsgebiets und den Themenbereichen ermittelt. Sie dient im Folgenden als Grundlage zur Erstellung der HOAI basierten erweiterten Kostenprognose.

4.6.2 Die Kostenfunktionen und Ermittlungsmethodiken

Die Kostenfunktionen des Prognosemodells ergeben sich unter Einfluss der Aufbaumöglichkeiten (Neuaufbau, Veredelung) sowie den darin enthaltenen Arbeitstätigkeiten gemäß Abbildung 4.21.

	Vorbereitung	Zukauf von Daten			Eigene Erfassung			Eigener Bestand		Integration und Homogenisierung		
		Preis	Prüfung	Korrektur	Kosten	Prüfung	Korrektur	Prüfung	Korrektur	Kosten	Prüfung	Korrektur
N1	K_{Vorb}	K_{Kauf}	$K_{Kauf}^{Prüf}$	K_{Kauf}^{Verb}	-	-	-	-	-	-	-	-
N2	K_{Vorb}	K_{Kauf}	$K_{Kauf}^{Prüf}$	K_{Kauf}^{Verb}	K_{EigErf}	$K_{EigErf}^{Prüf}$	K_{EigErf}^{Verb}	-	-	-	-	-
N3	K_{Vorb}	-	-	-	K_{EigErf}	$K_{EigErf}^{Prüf}$	K_{EigErf}^{Verb}	-	-	-	-	-
V1	K_{Vorb}	K_{Kauf}	$K_{Kauf}^{Prüf}$	K_{Kauf}^{Verb}	-	-	-	$K_{EigBest}^{Prüf}$	$K_{EigBest}^{Verb}$	K_{I+H}	$K_{I+H}^{Prüf}$	K_{I+H}^{Verb}
V2	K_{Vorb}	K_{Kauf}	$K_{Kauf}^{Prüf}$	K_{Kauf}^{Verb}	K_{EigErf}	$K_{EigErf}^{Prüf}$	K_{EigErf}^{Verb}	$K_{EigBest}^{Prüf}$	$K_{EigBest}^{Verb}$	K_{I+H}	$K_{I+H}^{Prüf}$	K_{I+H}^{Verb}
V3	K_{Vorb}	-	-	-	K_{EigErf}	$K_{EigErf}^{Prüf}$	K_{EigErf}^{Verb}	$K_{EigBest}^{Prüf}$	$K_{EigBest}^{Verb}$	-	-	-
Methodik	Arbeitszeitmodell	Fixkosten	Kosten der Qualitätssicherung		HOAI basierte erweiterte Prognose	Kosten der Qualitätssicherung		Kosten der Qualitätssicherung		Arbeitszeitmodell (komplex)	Kosten der Qualitätssicherung	

Abbildung 4.21: Kostenfunktionen und korrespondierende Ermittlungsmethodik

Die Kostenfunktionen der jeweiligen Aufbaumöglichkeiten ergeben sich aus zeilenweiser Addition der Teilkostenparameter und kombinieren diese entsprechend ihren Arbeitstätigkeiten unterschiedlich. Deutlich hervor tritt in dieser Darstellung der große Kostenblock der Qualitätssicherung auf den wegen seiner Komplexität erst im folgenden Kapitel 5 eingegangen werden soll.

Darüber hinaus zeigt Abbildung 4.21 die zu den Kostenfunktionen korrespondierende Ermittlungsmethodik. Während die Kosten der Arbeitsvorbereitung sowie der Integration und Homogenisierung weitestgehend über ein Arbeitszeitmodell prognostiziert werden können, kommt zur Ermittlung der Erfassungskosten die Methodik der HOAI basierten erweiterten Prognose zum Einsatz.

Die nachfolgenden beiden Abschnitte erläutern deshalb die Methodik der HOAI basierten erweiterten Prognose sowie das Arbeitszeitmodell und geben die damit verbundenen Wertermittlung an. Sie ist wichtiger Bestandteil des Kostenprognosemodells.

4.6.2.1 Die HOAI basierte erweiterte Prognose

Das Rechenverfahren zur Ermittlung der Kosten ist abhängig von den Aufbaumöglichkeiten (Vollständige Neuerfassung, Reine Datenübernahme, Datenübernahme und eigene Zusatzerfassung), den Qualitätsforderungen (ausgedrückt durch komplementäre Qualitätsparameter), den betrachteten Themenbereichen (z.B. Straßen, Schienen, Gebiete, usw.) sowie der Datenstruktur (Objekte, Attribute und Relationen). Das hier vorgestellte Rechenverfahren differenziert in erster Linie nach den Aufbaumöglichkeiten und berücksichtigt die weiteren Einflussfaktoren durch die Variation der Funktionen.

Die Berechnung der Grundleistungen der Phasen 1-6 und 7 dient zur Ermittlung der Verrechnungseinheiten für die Abschätzung der Kosten der **Vollständigen Neuerfassung** und trägt im Rahmen der Kostenermittlung bei der *Daten-*

übernahme und eigene Zusatzerfassung dazu bei, die Kosten der eigenen Erfassungsleitung (Verrechnungseinheiten der Phase 7) prognostizieren zu können.

Verrechnungseinheiten der Leistungsphasen 1-6 ($V_E^{LPh 1-6}$)

Die Berechnung der Verrechnungseinheiten für die Leistungsphasen 1-6 ist mit Hilfe der AHO durchzuführen wird in nachstehender Formel ausgedrückt.

$$\begin{aligned}
 V_E^{LPh 1-6} &= ha \cdot 1 \text{ VE} \\
 &+ n^{BK} \cdot 9.000 \text{ VE} \\
 &+ x^{LK} \cdot 12.000 \text{ VE} \\
 &+ x^{FL} \cdot 6.000 \text{ VE} \\
 &+ n^{PO} \cdot 4.000 \text{ VE}
 \end{aligned}
 \left| \begin{array}{l}
 ha = \text{Fläche des Gebiets in Hektar} \\
 n^{BK} = \text{Anzahl notwendiger Basiskarten,} \\
 x^{LK} = \text{Wahrheitswert Leitungskataster, Typ: BOOLEAN} \\
 x^{FL} = \text{Wahrheitswert Flächenelement, Typ: BOOLEAN} \\
 n^{PO} = \text{Anzahl Themengebieten mit unterschiedlichen Punktobjekten.}
 \end{array} \right.$$

Da Straßen bzw. Schienenelemente in der Regel nicht mit derselben Häufigkeit innerhalb eines räumlichen Gebiets, wie die Elemente des Leitungskatasters vorliegen, aber im Vergleich zu den Verrechnungseinheiten der Emissionen / Immissionen (hier als Flächenelement eingeführt) deutlich häufiger vorkommen, werden für geometrische Linienelemente (Straßen- und Schienenelemente) als Verrechnungseinheiten 10.000 VE angesetzt. Damit erweitert sich die Formel um die Komponente:

$$+ x^{LE} \cdot 10.000 \text{ VE}
 \left| \begin{array}{l}
 x^{LE} = \text{Wahrheitswert Themengebieten mit unterschiedlichen Linienelementen; Typ BOOLEAN}
 \end{array} \right.$$

Damit ergibt sich die Formel zur Berechnung der **Verrechnungseinheiten der Leistungsphasen 1-6** für die *Vollständige Neuerfassung* zu:

$$V_E^{LPh 1-6} = ha \cdot 1 \text{ VE} + n^{BK} \cdot 9.000 \text{ VE} + x^{LK} \cdot 12.000 \text{ VE} + x^{LE} \cdot 10.000 \text{ VE} + x^{FL} \cdot 6.000 \text{ VE} + n^{PO} \cdot 4.000 \text{ VE}$$

Mit Hilfe dieser Formel wird gewissermaßen der Sockelbetrag ermittelt, welcher sich durch die „äußeren“ Rahmenbedingungen der Datenerfassung ergibt. Unter den äußeren Rahmenbedingungen sind die Einflussfaktoren der Gebietsgröße, die zu erfassenden Objektthemenbereiche sowie die Eigenschaften (Punkt, Linie, Fläche) der zu erfassenden Objekte zu verstehen. Die Gebietsgröße wird als Einflussfaktor dieser Leistungsermittlung explizit zweimal berücksichtigt. Einerseits dient sie als Faktor zur Ermittlung der Grundleistungen und wird, wie unter Punkt b) beschrieben auch zur Ermittlung des Aufwands gemäß Leistungsphase 7 herangezogen.

Verrechnungseinheiten der Leistungsphase 7 – Erfassungsvorlagen ($V_{E,ErFVlg}^{LPh 7}$)

Neben den äußeren Rahmenbedingungen existieren konkrete Kostenfaktoren wie z.B. die Größe des zu erfassenden Gebiets, die Erfassungsmethode der Digitalisierung oder die Anzahl benötigter Vorlagen zu Digitalisierung sowie aller gefertigten Handskizzen. Entsprechend der AHO ergibt sich folgende Formel zur Berechnung der Komponente der Verrechnungseinheiten der Erfassungsvorlagen für die Leistungsphase 7:

$$\begin{aligned}
 V_{E,ErFVlg}^{LPh 7} &= ha \cdot 20 \text{ VE} \\
 &+ n^{DV} \cdot 70 \text{ VE} \\
 &+ n^{SZ} \cdot 20 \text{ VE}
 \end{aligned}
 \left| \begin{array}{l}
 ha = \text{Fläche des Gebiets in Hektar} \\
 n^{DV} = \text{Anzahl benötigter Digitalisierungs-Vorlagen (z.B. Orthobilder),} \\
 n^{SZ} = \text{Anzahl benötigter Skizzen (z.B. Feldbücher, Handskizzen)}
 \end{array} \right.$$

Verrechnungseinheiten der Leistungsphase 7 – Objekterfassung ($V_{E,ObjErf}^{LPh 7}$)

Um die Berechnung der Verrechnungseinheiten für die Objekterfassung möglichst auf alle Themenbereiche anwenden zu können, muss die Objektstrukturierung (Art und Anzahl geometrischer Primitive, Attributtypen und -anzahl, Relationen) berücksichtigt und in das Berechnungsverfahren integriert werden.

$$\begin{array}{l}
 V_{E,ObjErf}^{LPh 7} = n_{pkt}^{obj} \cdot 1 \text{ VE} \\
 + n_{codAtt}^{obj} \cdot 1 \text{ VE} \\
 + n_{ZeiAtt}^{obj} \cdot 2,5 \text{ VE} \\
 + 1,5 \text{ VE} \\
 + n_{rel}^{obj} \cdot 3 \text{ VE}
 \end{array}
 \left|
 \begin{array}{l}
 n_{pkt}^{obj} = \text{Anzahl geometrischer Punkte zur Objektbildung} \\
 n_{codAtt}^{obj} = \text{Anzahl Attribute (Typ: INTEGER) pro Objekt.} \\
 n_{ZeiAtt}^{obj} = \text{Anzahl Attribute (Typ:STRING) pro Objekt} \\
 \text{Pauschale für Objektbildung} \\
 n_{rel}^{obj} = \text{Anzahl der Objektrelationen}
 \end{array}
 \right.$$

Gemäß der AHO besteht die Möglichkeit die Digitalisierung von Objektpunkten im Rahmen eines Spektrums von 1,0 VE bis 3,0 VE unterschiedlich und je nach Schwierigkeitsgrad der Punktdigitalisierung zu bewerten. In dieser Formel wird deshalb der niedrigste Schwierigkeitsgrad (1VE) angesetzt, da z.B. für den Aufbau von digitalen Geodatenbeständen für verkehrliche Anwendungen (z.B. GDF-Straßendaten) oder zur Beschreibung der topographischen Gebietseigenschaften (ATKIS-Daten) in der Regel von der Digitalisierung der Planvorlage ohne geometrische Konstruktion ausgegangen wird. Betrachtet man hingegen die Digitalisierung mit konstruktiven Aufgaben, z.B. Erfassung der ALK, so ist die Formel entsprechend den genannten Verrechnungseinheiten anzupassen.

Unter der Anzahl n_{codAtt}^{obj} (Attribute vom Typ INTEGER) werden alle Attribute zusammengefasst, die numerische oder über einen Objektschlüssel codierte Attributwerte besitzen. Als Attribute vom Typ STRING werden z.B. Zeichenketten (Straßenname, Eigennamen von Gebieten etc.) verstanden, deren Erfassungsaufwand im Verhältnis zu codierten Attributen als größer zu bewerten ist. Der Aufwand der Objektbildung zur Realisierung der Topologie wird mit einer Pauschale von 1,5 VE und die notwendigen Relationen n_{rel}^{obj} mit 3,0VE berücksichtigt. Diese 3 VE entstammen der Differenz zwischen 4,5VE für eine logische Objektverknüpfung und den 1,5 VE für die allgemeine Objektbildung ohne Verknüpfung.

4.6.2.2 Prozessspezifisches Arbeitszeitmodell zur Leistungsermittlung

Die HOAI und die zugehörige Erläuterung AHO zählt den Fall der Datenhomogenisierung (Kompatibilität Datenbestände und Datenübernahme) zu den *Besonderen Leistungen* und nicht zu den *Grundleistungen*. Deswegen ist die Aufgabe der Datenhomogenisierung nicht durch Verrechnungseinheiten beschreibbar.

Da die *Besonderen Leistungen* in der HOAI über vertragliche Vereinbarungen geregelt und über die geleistete Arbeitszeit abgerechnet werden, ist hier das Aufstellen eines **Arbeitszeitmodells** notwendig, um die Leistungsermittlung der Datenhomogenisierung durchführen zu können. Grundlage eines solchen Arbeitszeitmodells ist der in einzelne Phasen zu unterteilende **Prozessablauf der Datenhomogenisierung**, um so jede einzelne Phase als Kostenstelle betrachten zu können.

Der Prozessablauf der Datenhomogenisierung / -konvertierung kann laut [BEHR, 1998] (Verfahrensablauf der Datenkonvertierung) in die Phasen: *Analyse der Eingangsdaten* und der *Objektstruktur des Zielsystems*, *Aufstellen von Abbildungsregeln*, *Konvertierung eines Testbestandes*, *Prüfung und Modifikation* und in die *Konvertierung des Gesamtbestandes* unterschieden werden. Unter Berücksichtigung, dass mehrere Datenbestände zusammen in ein Zielsystem konvertiert werden können und die vollständige Homogenisierung des Gesamtbestandes in der Regel nicht ohne manuelle Nachbearbeitung durchführbar ist, wurde der von BEHR definierte Verfahrensablauf erweitert und dient hier fortan als Grundlage für die Quantifizierung des Zeitbedarfs der einzelnen Prozessphasen.

Der Verfahrensablauf der Konvertierung von Daten in das Zielformat ist genau dann als einfach zu bezeichnen, wenn bei der Parametrisierung direkte Beziehungen zwischen den Datenelementen über einfache Umsetzungsregeln hergestellt werden können. Er gestaltet sich dagegen schwierig, wenn dafür komplexe Algorithmen verwendet werden müssen, welche die:

- Datenelemente zusammenfassen (z.B. Objektbildung),
- Objekte klassifizieren (Gebiets- oder Themenbereichszugehörigkeit),
- Positionsangaben von Objekten transformieren, sowie
- Informationen ergänzen.

Folgende Abbildung verdeutlicht den Ablauf der Datenkonvertierung in ein Zielsystem und berücksichtigt in Form eines Arbeitszeitmodells den geschätzten Arbeitszeitverbrauch innerhalb der jeweiligen Prozessphase. Die nachfolgend angesetzten Arbeitszeitwerte erheben nicht den Anspruch auf Allgemeingültigkeit. Sie wurden lediglich zur Veranschaulichung des Kostenrahmens festgelegt. Ihre Dimension orientiert an Erfahrungswerten aus bekannten Literaturstellen [BEHR, 1998; HAKE, 1994 ; BILL, 1999]. Es verbleibt die Aufgabe der praktischen Kostenprognose diese Werte á priori unternehmensspezifisch zu ermitteln.

Prozesse	Arbeitszeit	Formel
<div style="border: 1px solid black; padding: 2px; margin-bottom: 2px;">Analyse Eingangsdaten</div> <div style="border: 1px solid black; padding: 2px;">Analyse Struktur Zielsystem</div>	n^{Fusion} Zeitfaktor Analyse pro Datensatz:	$0,5PM + 0,1h n^{Fusion}$
<div style="border: 1px solid black; padding: 2px;">Aufstellen der Abbildungsregeln</div>	Zeitfaktor pro Datensatz: Arbeitsgruppe mit 3P	$1,5PM + 0,3h n^{Fusion}$
<div style="border: 1px solid black; padding: 2px;">Entwicklung Konverter</div>	Konst. Zeitfaktor Entwicklung:	$6PM$
<div style="border: 1px solid black; padding: 2px;">Konvertierung Testbestand</div>	kein Kostenfaktor angesetzt	
<div style="border: 1px solid black; padding: 2px;">Prüfung und Modifikation</div>	Konst. Zeitfaktor Prüfung: Variabler n^{Fusion} -Anteil:	$2PM + 0,1h n^{Fusion}$
<div style="border: 1px solid black; padding: 2px;">Konvertierung Gesamtbestand</div>	kein Kostenfaktor angesetzt	
<div style="border: 1px solid black; padding: 2px;">Manuelle Nachbearbeitung</div>	Autom. Konvertierungsquote Anzahl Datenrecords Arbeitsleistung	$\frac{(1-ag) \times n^{Fusion}}{25 \text{ DatRec/h}}$ $ag \{ 75\%-100\% \}$ n^{Fusion}

Abbildung 4.22: Arbeitszeitmodell zur Leistungsermittlung der Datenmigration

Erläuterungen der Einzelschritte der Datenkonvertierung:

- Im ersten Schritt des Ablaufs der Datenkonvertierung muss eine detaillierte **Analyse** der umzusetzenden Datenbestände und des Datenmodells im Zielsystem (Zielbestand) vorgenommen werden. Diese Analyse umfasst alle darin enthaltenen Objektklassen, den Attributumfang, die Art und der Umfang der Graphikinformationen sowie die Topologie. Die Analyse des Eingangsbestandes kann auf der Grundlage vorliegender Strukturdefinitionen des Ausgangssystems (Quellbestand), der Formatbeschreibung oder durch DV-gestützte Analyse des Datenbestandes erfolgen. Die Objektstruktur des Zielsystems kann der Schnittstellendefinition und dem Datenmodell der Zielanwendung entnommen werden.
- Für die Datenumsetzung muss im nächsten Schritt für jedes Datenelement im Quellbestand eine Entsprechung im Zielsystem gefunden werden. Diese Entsprechung wird in Form von **Abbildungsvorschriften** festgelegt. Diese umfasst die Vorgaben für die Überführung der Datenelemente bzw. Objekte, der Attribute, der Geometrie und der durch den Konverter zu generierenden Zusatzinformationen. Der Aufwand für diese Festlegungen und die notwendigen, iterativ durchzuführenden Testläufe ist hoch.
- Die Abbildungsvorschriften werden durch den **Konverter** ausgewertet und sind zweckmäßigerweise zunächst auf einen Testbestand anzuwenden. Musterdatenbestände decken oft nicht alle Fälle ab, so dass sich der Aufwand der systematischen Erstellung eigener Testdatensätze lohnt.

- Der Erfolg bzw. Misserfolg der Datenkonvertierung hängt vom Automatisierungsgrad des Konverters ab. Je höher dieser Grad ausfällt, desto geringer ist der Nachbearbeitungsaufwand, bzw. desto weniger Iterationsschritte, sind bei der Erstellung des Konverterprogramms notwendig.
- Die Anzahl nachzubearbeitender Datensätze (Datenrekords) ergibt sich aus der komplementären Umsetzungsquote und deren Gesamtanzahl. Um den Nachbearbeitungsaufwand in Stunden ausdrücken zu können, wird eine stündliche Arbeitsleistung von 25 Datenrekords bei manueller Nachbearbeitung angesetzt.

Erläuterung des Arbeitszeitmodells:

Die in Abbildung 4.22 geschätzten Zeitfaktoren sind in Personenmonate (PM) aufgeführt. Legt man zur Umrechnung der Personenmonate in Arbeitsstunden eine Wochenarbeitszeit von 37,5 Stunden und eine durchschnittliche Anzahl von 21 Arbeitstagen eines Monats zugrunde, dann ergibt sich eine monatliche Stundenleistung von rund 158 Stunden.

Die in Abbildung 4.22 festgelegten Arbeitszeitfaktoren werden ohne den direkten Bezug zu einer praktisch durchgeführten Datenkonvertierung festgelegt und stellen damit keine Erfahrungswerte unternehmensinterner Kostenrechnung dar. Diese Festlegung geschieht deshalb, da der Prozess der Datenkonvertierung, und die darin zu definierenden Abbildungsvorschriften sowie die Umsetzungsquote eines Konverterprogramms von zu vielen externen, global nicht abschätzbaren Unbekannten abhängt.

Damit können aus der im Rahmen dieser Arbeit ermittelten Gesamtkostenstruktur des Aufbaus von Datenbeständen nur relative Kostenaussagen getroffen werden. Eine absolute Kostenaussage auf der Basis dieser festgelegten Arbeitszeiten der Datenkonvertierung kann nicht gegeben werden. Dies ist auch nicht zwingend erforderlich, da für die Aussagen über mögliche Konkurrenzsituationen der Aufbaumöglichkeiten eine relative Kostenaussage genauso aussagekräftig ist wie ein absolute. Wichtig dabei ist nur, dieses Arbeitszeitmodell über die Kostenberechnungen (Kap. 5) hinweg konstant zu halten.

Die in diesem Arbeitszeitmodell enthaltenen Variablen sind die Anzahl der Datensätze (n^{Fusion}), die mit Hilfe eines Konverters zu einem gemeinsamen Datenbestand zusammenzuführen sind, und der Automatisierungsgrad (ag) des Konverterprogramms. Dabei wirkt die Anzahl n^{Fusion} auf den Zeitfaktor der Analyse der Eingangsdaten, da diese entsprechend n^{Fusion} -fach durchgeführt werden muss. Es resultiert die Arbeitszeit der Analyse unter Berücksichtigung eines variablen Anteils von 0,1h zu:

$$0,5\text{PM} + 0,1\text{h } n^{\text{Fusion}}$$

Das Aufstellen der Abbildungsregeln ist ebenfalls von der Anzahl n^{Fusion} abhängig, da die Datenbestände gleichen Typs miteinander verglichen werden müssen, bevor eine Abbildungsvorschrift in das Zielsystem erfolgen kann. Die Arbeitszeit ergibt sich mit demselben Zeitfaktor der Analyse (0,5PM) sowie einer Arbeitsgruppe von z.B. 3 Personen zu:

$$1,5\text{PM} + 0,3\text{h } n^{\text{Fusion}}$$

Die Umsetzung des Konverters (Programmiertätigkeit) wird, wenn die Abbildungsregeln feststehen, mit einem konstanten Zeitwert von 6 PM angesetzt. Ein variabler Zeitanteil ist für diese Tätigkeit nicht anzusetzen, da der Arbeitsaufwand der Programmierung für den vorliegenden Fall vom Datenvolumen weitestgehend unabhängig ist.

Der Zeitaufwand zur Prüfung der konvertierten Testdatenbestände und die Verbesserung des Konverterprogramms lässt sich durch dieselbe Vorgehensweise abschätzen, wie der erstmalige Umsetzungsaufwand des Konverters. Der Modifikationsaufwand des Konverterprogramms wird über einen konstanten Anteil von 2PM und einen $0,5 n^{\text{Fusion}}$ -variablen Anteil angesetzt:

$$2\text{PM} + 0,1\text{h } n^{\text{Fusion}}$$

Der Aufwand der manuellen Nachbearbeitung ist maßgeblich vom Automatisierungsgrad ag abhängig. Mit einer Produktivität von 25 zu homogenisierenden Datensätzen pro Arbeitsstunde ergibt sich der Aufwand zur manuellen Nachbearbeitung der Homogenisierung zu:

$$(1-ag) n^{\text{Fusion}} 1/25$$

Der Wertebereich der Konvertierungsquote hängt stark vom Datenmodell und dessen –struktur ab. Die Auswirkungen verschiedener Konvertierungsquoten auf die Kosten der Aufbaumöglichkeiten ist deshalb zentraler Gegenstand der Betrachtungen in Kapitel 5. Der Arbeitszeitfaktor der Nachbearbeitung ergibt sich damit unter der weiteren Annahme einer manuellen Konvertierung von 25 Datenrekords / Stunde. Aus diesen Einzelzeitkomponenten zur Leistungsberechnung über das Arbeitszeithonorar der HOAI ergibt sich, unterteilt in konstante und variable Zeitanteile, die folgende Formel zur Berechnung des Arbeitszeitaufwandes der Homogenisierung von Datenbeständen:

$$10PM + \frac{0,5h}{25} \cdot (1 - ag) \cdot n^{Fusion}$$

Zur Berechnung des Honorars gibt die HOAI für unterschiedlich qualifizierte Bearbeiter minimal und maximal zu verrechnende Stundensätze an. Für die weitere Berechnung wird jeweils der mittlere Stundensatz für den Auftragnehmer von €60,00, den Mitarbeiter der Qualifikationsstufe 1 von €48,00 und für einen Mitarbeiter der Qualifikationsstufe 2 von €37,00 angesetzt. Ferner wird davon ausgegangen, dass sich die Summe der monatlichen Stundenleistung wie folgt aufteilt.

Tabelle 4-2: Bearbeiterprofil mit Stundensatz und Arbeitsanteil

Person	Profil	Stundensatz	Arbeitsanteil
Auftragnehmer	Leiter der Arbeitsgruppe	€ 60,00	10%
Mitarbeiter, Qualifikationsstufe 1	Technischer Mitarbeiter, Vertreter des Leiters	€ 48,00	60%
Mitarbeiter, Qualifikationsstufe 2	Technischer Zeichner, gelernte Fachkraft	€ 37,00	30%

Anhand dieses Profils der am Prozess beteiligten Personen und einer monatlichen Stundenleistung von 158h folgt die Kostenfunktion für die Datenhomogenisierung nach diesem Arbeitszeitmodell:

$$46 \frac{\text{€}}{h} \left(1580h + \frac{0,5h}{25} \cdot (1 - ag) \cdot n^{Fusion} \right)$$

Es sei hier ausdrücklich noch einmal betont, dass das hier definierte Arbeitszeitmodell, nicht den Anspruch auf Allgemeingültigkeit und universelle Anwendbarkeit zur Kostenabschätzung erhebt. Vielmehr sind die Arbeitszeitfaktoren durch unternehmensinterne Kostenstellenverantwortliche entsprechend anzupassen. Zum Zwecke einer relativen und vergleichenden Aussage bzgl. verschiedener Aufbaumöglichkeiten sind die absoluten Zahlen allerdings weniger entscheidend, als die Konstanz, mit der die Arbeitszeitwerte innerhalb der Gesamtkostenschätzung gehalten werden. Das daraus resultierende Ergebnis liefert dann zwar keine absolute sondern eine relative Kostenaussage, welche aber für einen Vergleich der Aufbaumöglichkeiten ausreichend ist. Damit steht mit dieser Formel ein Instrument zur Verfügung, mit dem die Besonderen Leistungen (z.B. nach HOAI) über die Arbeitszeit abgerechnet werden können und das die Anzahl der Datenbestände gleichermaßen berücksichtigt wie die Automatisierbarkeit der Datenumsetzung in Form des Automatisierungsgrades.

4.7 Validierung des Prognosemodells

Unter Zuhilfenahme der realen Kostenanalyse des intermodalen Routenplaners und den damit verfügbaren Kostenparametern sowie deren realistischen Werten erfolgt die Kalibration und Validierung der synthetischen Modellbildung. Die Vorgehensweise zur Kalibration und Validierung kann anhand Abbildung 4.23 verdeutlicht werden:

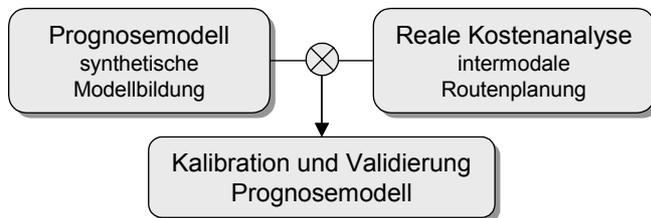


Abbildung 4.23: Vorgehensweise Modellkalibration

4.7.1 Datenbankbasierte Modellumsetzung

Diese Kalibration und Validierung wurde durch die Modellumsetzung in Form einer Datenbank ermöglicht. Als Umsetzungshilfsmittel fiel die Wahl auf die Standarddatenbank MSAccess™ von Microsoft. Ihr Funktionsprofil erlaubte es in kürzester Zeit ein funktionierendes, erweiterbares sowie leicht pflegbares Prognosewerkzeug zu erstellen. Mit Hilfe mehrerer formularbasierter Abfragen und relational verknüpften Datenbanktabellen entstand eine, in Arbeitsphasen unterteilte, Kostenprognosedatenbank.

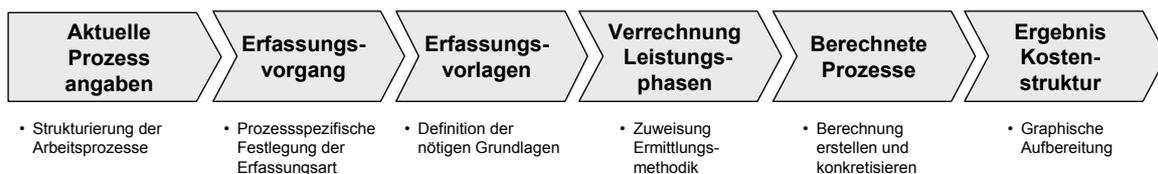


Abbildung 4.24: Arbeitsphasen der Kostenprognosedatenbank

Stellvertretend für alle Arbeitsphasen der Prognosedatenbank sollen zumindest die Datenbankformulare der Phasen *Erfassungsvorgang* und *Verrechnung Leistungsphasen* dargestellt werden.

modellbasierte Kostenberechnung für den Aufbau von Geodatenbeständen

Aktuelle Prozessangaben | Erfassungsvorgang | Erfassungsvorgang | Verrechnung der Prozesse nach Leistungsphasen | Berechnete Prozesse | Kostenstruktur

Definition Erfassungsvorgang, Erfassungsgegenstand, Attribute und Relationen

Aktiver Hauptprozess: Nachbearbeitung der digitalen Straßenkarte
 Aktiver Teilprozess: Erfassen von Fußwegen

Erfassungsvorgang

Setze den Erfassungsvorgang: Bildschirmdigitalisierung

definierte Erfassungsvorgänge:

- Digitalisierung
- Bildschirmdigitalisierung
- Sachdatenerfassung
- Vor-Ort Erfassung
- interne Erfassung (unternehmer existierende Daten bearbeiten)

Datensatz: 1 | 1 | von 13

Gegenstand der Erfassung

Definiere Elementtyp: Kante

definierte Elementtypen:

- Punkt
- Kante
- Attribut
- Handskizze

Elementname (optional): Fußweg

Setze Anzahl zu erfassende Elemente: 2751

Angaben zu Attributen:

Element besitzt Attribute: mit Attribute ohne Attribute

Wettermittlung: manuell automatisch

Anzahl Attribute pro Element: 4

Attributnamen: Klasse, Niveau, Hin- und Rückweg

Angaben zu Relationen:

Objekte verknüpft: mit Relationen ohne Relationen

Abbildung 4.25: Formular Erfassungsvorgang

Im Rahmen dieses Formulars ist es möglich, entsprechend den zuvor definierten Teilprozessen, die Art und den Gegenstand der Datenerfassung zuzuweisen.

So wird z.B. der Elementtyp (Punkt, Kante, Attribut, etc), sowie weitere Informationen wie Anzahl, Attribute, deren Wertermittlung und auch mögliche Objektkrelationen angegeben.

Diese Festlegung des Erfassungsvorgangs geschieht für alle definierten Teilprozesse. Sie dient zur Quantifizierbarkeit der Erfassungstätigkeiten und ist Grundlage für die Zuweisung zu Kostenfunktionen.

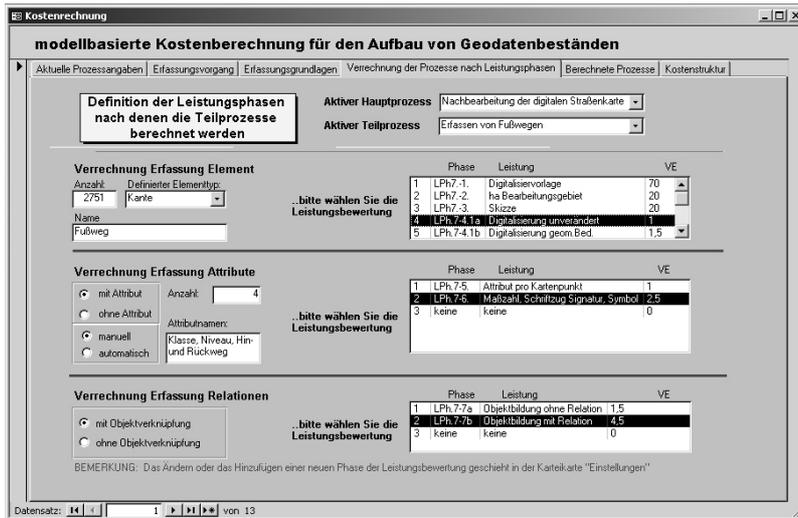


Abbildung 4.26: Zuweisung der Verrechnungsmethodik

Dieses Formular ermöglicht die Zuweisung der zu erfassenden Elemente mit der entsprechend kostenmäßigen Leistungsbewertung, also entweder nach HOAI basierter erweiterter Prognose oder nach dem Arbeitszeitmodell.

Grundsätzlich wird hier in die Erfassung von Elementen (Objekten) deren Attribute sowie in die Objektrelationen unterschieden.

Diese Zuweisung ist die fehlende Verbindung zwischen den klar zu beschreibenden Arbeitstätigkeiten und der entsprechenden Leistungsbewertung.

Nach Bearbeitung der formularbasierten Arbeitsphasen erfolgt die automatische Prognose der Kosten sowie deren graphische Aufbereitung in Form einer Kostenstruktur.

4.7.2 Kalibration und Validierung durch reale Kostenanalyse

Durch die im Rahmen der Mobilist Begleitforschung durchgeführte Kostenanalyse der intermodalen Routensuche, dargestellt in Kapitel 3 dieser Arbeit, stehen zur Modellkalibration und -validierung die notwendigen Eingangsparameter zur Verfügung. Diese wurden in die formularbasierte Kostenprognosedatenbank eingepflegt und die Kostenschätzung durchgeführt. Das Ergebnis dieses Vergleichs ist in Abbildung 4.27 dargestellt.

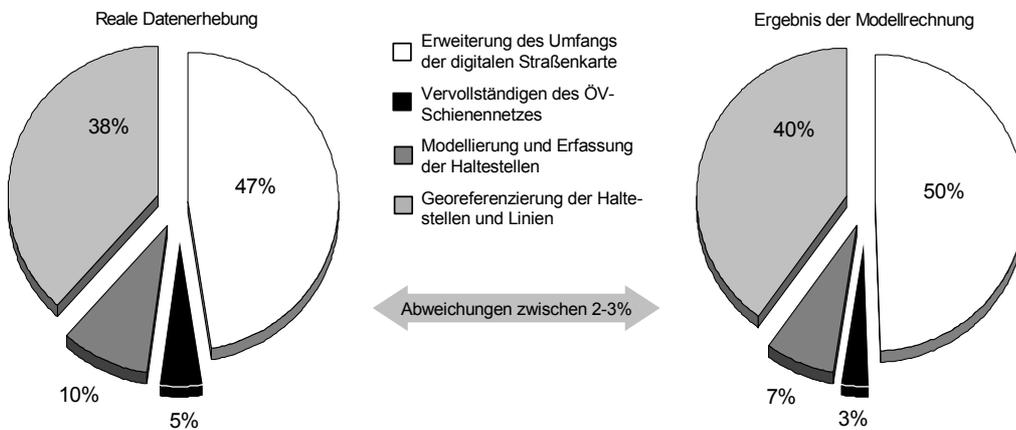


Abbildung 4.27: Vergleich der realen mit der prognostizierten Kostenstruktur

Gegenstand dieses Vergleichs sind die reinen Nachbearbeitungstätigkeiten zum Aufbau der Geodatengrundlage ohne die Fixkosten der digitalen Straßenkarte oder den Kostenanteilen der 30%igen, bzw. 100%igen Überprüfung.

Klar zu erkennen ist die deutliche Übereinstimmung zwischen den Kostenanteilen der realen Datenerhebung und dem Ergebnis der Modellrechnung. Es ergeben sich geringfügige Abweichungen zwischen 2-3%. Damit kann dem Modell zumindest für diese typische Aufgabe des Geodatenaufbaus eine gute Funktionsweise attestiert werden. Es gilt hier zu betonen, dass vor einer weiteren Anwendung dieser Kostenprognosedatenbank entsprechend fallspezifische Einstellungen und Vorarbeiten zu leisten sind, um auf dieselbe Prognosegenauigkeit zu gelangen. Eine direkte Übertragbarkeit der Kostenprognose würde der Einzigartigkeit jedes Geodatenaufbaus nicht gerecht werden und zu verfälschten Ergebnissen führen.

4.8 Bemerkungen zum Ermittlungskonzept der Bereitstellungskosten von Geodaten

Das Ziel dieses Kapitels war es, ein Konzept zu entwickeln, mit dessen Hilfe die Bereitstellungskosten von Geodaten vor dem Hintergrund einer definierten Anwendung bereits im Vorfeld prognostiziert werden können. Es wurde bei diesem Kapitel von einer Situation ausgegangen, die sich einem Unternehmen einstellt, wenn dieses beabsichtigt, sich mit Unterstützung von Geodaten ein neues Geschäftsfeld in Form einer Informationsdienstleistung zu erschließen.

Das in diesem Kapitel entwickelte Konzept greift bereits zu Beginn, also zum Zeitpunkt der Zielsetzungsspezifikation der Informationsdienstleistung ein und ermöglicht es so, mit Hilfe eines aufbauprozessbegleitenden Kostenmodells die Teilprozesskosten explizit ermitteln und mögliche Kostenrisiken vor Durchführung erkennen zu können. Mit Hilfe der Beschreibung der prinzipiell durchführbaren Möglichkeiten des Geodatenaufbaus (Neuaufbau bzw. Veredelung), die für sich genommen durch jeweils eine Prozesskette darstellt werden können und das dazugehörige Entscheidungsdiagramm wurden die strukturellen und inhaltlichen Voraussetzungen zur Ermittlung der Aufbauprozesskosten gelegt.

Dieses Konzept basiert im Wesentlichen auf einem prozessspezifischen mathematischen Kostenmodell, welches speziell auf den Vorgang der Geodatenbereitstellung angepasst worden ist. Durch eine Unterteilung der insgesamt sechs verschiedenen Aufbaumöglichkeiten in ihre einzelnen Kostenstufen ist das hier entwickelte Konzept darauf vorbereitet, die Kostenarten entsprechend ihrer Entstehung auf die Kostenstellen (Teilprozesse) bzw. auf ihre Kostenauswirkung (Kostenträger) angeben zu können. Diese Untergliederung ist zur betriebswirtschaftlichen Beurteilung der Vorgehensweise von entscheidender Bedeutung. Aus diesem Grund ist das Konzept an der Schnittstelle zwischen technischer Realisierung und betriebswirtschaftlicher Kostenrechnung angesiedelt. Es liefert die zur betriebswirtschaftlichen Kostenberechnung benötigten Eingangsdaten.

Mit Hilfe der realen Kostenanalyse des intermodalen Routenplaners gelang die Validierung des Kostenprognosemodells, also der synthetischen Modellbildung mit einer Abweichung von nur 2-3% der prognostizierten zu den realen Kostenanteilen. Dies wurde erreicht durch die Umsetzung des Prognosemodells in Form einer MSAccess™ Datenbank. Mit deren Hilfe man, entsprechend Arbeitsphasen folgend, den Gesamtvorgang des Geodatenaufbaus hinsichtlich Art, Menge und Elementtyp formularbasiert in die Datenbank einpflegen kann, um so zu einer Kostenprognose zu gelangen.

Wie bereits in Abbildung 4.21 deutlich wurde, sind die Qualitätssicherungsmaßnahmen einer der Hauptkostenfaktoren. Es ist deshalb im folgenden Kapitel 5 das Ziel, die Kostenauswirkungen von Qualitätssicherungsmaßnahmen zu ermitteln und diejenigen Einflussfaktoren herauszufinden, die für eine mögliche Konkurrenz der Aufbaumöglichkeiten sorgen. Das Ziel hierbei ist es, angegeben zu können, ab welchen Parameterwerten eine Aufbaumöglichkeit einer anderen vorzuziehen ist.

5 Qualitätssicherung von Geodaten

Das Ziel dieses Kapitels ist die Integration von Qualitätsforderungen in das, in Kapitel 4 beschriebene, Konzept zur Ermittlung der Bereitstellungskosten von Geodaten, die Ableitung von Kostenfunktionen zur Beschreibung des fehlerabhängigen Kostenverlaufs sowie generell die Abhängigkeiten zwischen Qualitätsforderungen und den damit unwillkürlich verbundenen Kosten aufzuzeigen. Darüber hinaus sollen Aussagen getroffen werden, um sagen zu können, von welchen Einflussfaktoren eine Konkurrenz zwischen den Kostenverläufen unterschiedlicher Aufbaumöglichkeiten abhängt. Es wird beabsichtigt, anhand dieser Konkurrenz diejenige Aufbaumöglichkeit ermitteln zu können, welche sich im Vergleich zu den anderen am kostengünstigsten verhält.

Da die Qualitätsforderungen, die sich an einen aufzubauenden Datenbestand richten, nur über Qualitätssicherungsmaßnahmen wie z.B. Vollprüfungen des Bestandes oder Stichprobenprüfungen durchgeführt werden können und diese nur im umfassenden Rahmen des Qualitätsmanagements von Geodaten angegeben werden können, soll im ersten Teil dieses Kapitels eine Abgrenzung zwischen den Begriffen des Qualitätsmanagements und der Qualitätssicherung dargestellt werden.

Im Anschluss daran folgt die Ableitung geeigneter, kostenoptimaler Stichprobenprüfpläne, welche die Qualitätsforderungen seitens des Produzenten von Geodaten ebenso berücksichtigen wie die der Abnehmer (Nutzer) dieser Daten. Daraufhin werden die Kostenfunktionen, die sich durch die Qualitätssicherungsmaßnahmen ergeben, aufgestellt und eine strukturierte Vorgehensweise zur Anwendung der Berechnungsmethodik angegeben. Anschließend werden mit Hilfe eines Modelldatenbestandes die Kostenfunktionen der Aufbaumöglichkeiten in Abhängigkeit der Kosteneinflussfaktoren berechnet. Es sind diejenigen Einflussfaktoren dargestellt, unter deren Variation sich unterschiedlichen Konkurrenzsituationen bei Aufbaumöglichkeiten ergeben.

Insgesamt dient dieses Kapitel zur Darlegung und Erläuterung der Analyse- und Berechnungsmethodik zur individuellen, problemspezifischen Kostenprognose, um damit ein Dienstleistungsunternehmen in die Lage zu versetzen, die günstigste Aufbaumöglichkeit von Geodatenbeständen identifizieren zu können. Anhand mehrerer Beispiele werden die Konkurrenzsituationen abschließend berechnet und detailliert erläutert.

5.1 Qualitätsmerkmale als Beurteilungskriterien des Geodatenaufbaus

Um Entscheidungen über die Verwendbarkeit von Datenbeständen treffen zu können, müssen geeignete Kriterien fest stehen, deren Wertebereiche bekannt sind und für die Grenz- bzw. Schwellwerte existieren. Nur so kann die Entscheidung, ob eine *Reine Datenübernahme*, eine *Datenübernahme und Zusatzerfassung (mit oder ohne Homogenisierung)* oder eine *Vollständige Neuerfassung* durchzuführen ist, getroffen werden. Neben den Kosten der Aufbaumöglichkeiten kann grundsätzlich die Qualität von Geodaten als Entscheidungskriterium herangezogen werden. Sie lässt sich nach DIN EN ISO 9000 allgemeingültig über verschiedene Merkmale und deren Ausprägung beschreiben.

Die in diesem Abschnitt verwendeten **Qualitätsmerkmale für die Bereitstellung von Geodaten** gehen dabei im wesentlichen auf die Arbeit von [WILTSCHKO, 2004] zurück. Da WILTSCHKO diese Qualitätsmerkmale zur Beschreibung der Informationsqualität informationsverarbeitender Systeme aufstellt, hier aber das Problem der Geodatenbereitstellung im Mittelpunkt der Betrachtung steht, müssen die Qualitätsmerkmale bezüglich ihrer Auswirkung auf die Aufbaumöglichkeiten neu betrachtet werden. Einen wesentlichen Einfluss auf die nachfolgende Zusammenstellung der Qualitätsmerkmale hat insbesondere die Arbeit von [JOOS, 2000], welche einen Beitrag zur Qualität objektstrukturierter Geodaten liefert. Unter Berücksichtigung beider Arbeiten entsteht die Definition von **Merkmalsätzen als Entscheidungskriterien** für die Bewertung der Aufbaumöglichkeiten

Neben der Definition dieser Merkmalsätze und der Festlegung ihrer Wertebereiche steht darüber hinaus die Festlegung von **Grenzwerten bzw. Schwellwerten** für ihre Parameterwerte im Vordergrund. Die Gesamtheit dieser Schwellwerte kann sozusagen als wirtschaftliche Grenze interpretiert werden, bei der eine gewählte Aufbaumöglichkeit gegenüber einer Alternative teurer wird. Die Herleitung einer solchen Grenze ist eine der zentralen Aufgaben des **prozessorientierten Kostenmodells**. Diese wirtschaftliche Grenze, die zum Wechsel der Aufbaumöglichkeit oder zur Änderung der Zieldefinition führen kann führt, ist von der geforderten Qualität des Zieldatenbestands (Qualitätsforderung) ebenso

abhängig, wie von dem Qualitätsniveau, welches sich durch die Überführung von vorhandenen Datenbeständen in das Zieldatenmodell ergibt.

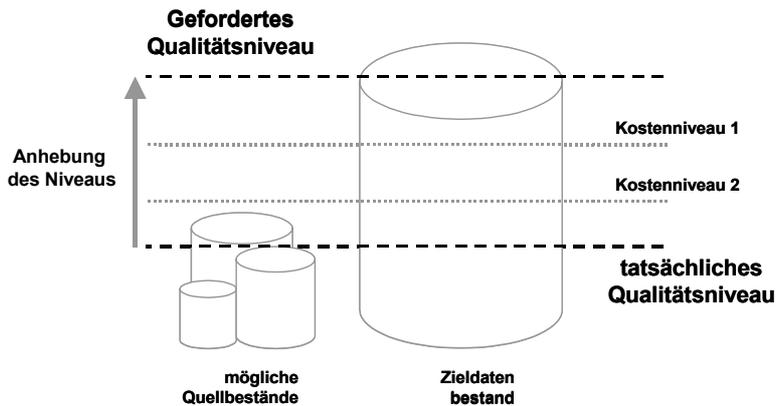


Abbildung 5.1: Qualitätsforderung und Qualitätsniveau

Es ist deshalb das zentrale Ziel dieser Arbeit eine Methodik zu entwickeln, die es gestattet festzustellen, welche Qualitätsmerkmale und Kostenparameter den größten Einfluss auf das Kostenverhalten der Aufbaumöglichkeiten besitzen und ab welchen Parameterwerten eine Aufbaumöglichkeit einer anderen hinsichtlich ihres Kostenverlaufs vorzuziehen ist.

5.1.1 Allgemeine Betrachtung verwendeter Qualitätsmerkmale

Die Beschreibung der Qualität von Daten und Informationen kann allgemein über einen Satz von Qualitätsmerkmalen durchgeführt werden. Diese Qualitätsmerkmale variieren entsprechend ihrer Verwendung und können sich auch innerhalb verschiedener Fachdisziplinen stark unterscheiden. Bei diesen Qualitätsmerkmalen handelt es sich um

- die **Zuverlässigkeitsmerkmale**: *Verfügbarkeit* und *Aktualität*,
- die **Integritätsmerkmale**: *Vollständigkeit*, *Konsistenz* und *Korrektheit* und um
- die **Genauigkeitsmerkmale**: *metrische* und *semantische Genauigkeit*.

Die folgende Definition der Qualitätsmerkmale gibt einen kurzen Einstieg in die Beschreibung der Qualität mittels Merkmale. Für eine detailliertere Betrachtung sei hier ausdrücklich auf die Arbeiten von WILTSCHKO und JOOS verwiesen.

Verfügbarkeit: Die Verfügbarkeit gibt das Ausmaß des Vorhandenseins der Information zu einem definierten Zeitpunkt an einem bestimmten Ort an. Die Verfügbarkeit und die Vollständigkeit beschreiben beide die Existenz von Daten und Informationen. Die Verfügbarkeit einer einzelnen Information oder eines Datensatzes lässt sich durch einen Wahrheitswert (Booleschen Wert) ausdrücken.

Aktualität: Die Aktualität gibt das Ausmaß der Übereinstimmung der Information mit der sich zeitlich ändernden konzeptionellen Realität¹ an. Eine Aktualisierung wird notwendig, wenn sich die konzeptionelle Realität ändert. Dieser Vorgang wird umgangssprachlich auch als „Datenalterung“ bezeichnet.

Vollständigkeit: Die Vollständigkeit gibt das Ausmaß des Vorhandenseins sämtlicher zur Beschreibung der konzeptionellen Realität erforderlichen Daten und Informationen an. Genauso wie die Konsistenz und die Korrektheit beschreibt die Vollständigkeit die Verwendbarkeit von Daten bzw. Informationen. Für Geodatenbestände bedeutet dies, dass die Vollständigkeit gewahrt ist, wenn die reale Welt durch Objektinstanzen von Objektklasse vollständig beschrieben ist und die darin enthaltenen Attribute mit Werten besetzt sowie die im Modell vorgesehenen Objektrelationen vollständig angelegt sind.

Konsistenz: Die Konsistenz gibt das Ausmaß der Übereinstimmung der Information mit dem Informationsmodell an. Um die Konsistenz von Daten und Informationen zu wahren ist lediglich deren Speicherung gemäß

¹ Konzeptionelle Realität ist die Abbildung von Phänomene der realen Welt in Form von Datenmodellen und –strukturen

dem zugrunde liegenden Informationsmodell zu gewährleisten. Die Konsistenz ist Grundvoraussetzung für einen widerspruchsfreien Datenbestand und deshalb speziell für den Prozess der Geodatenbereitstellung von besonderer Bedeutung.

Korrektheit: Die Korrektheit gibt das Ausmaß der Übereinstimmung der Daten bzw. Informationen mit der konzeptionellen Realität bei vorausgesetzter Aktualität an. Mit der Korrektheit wird, wie mit der Aktualität, das Freisein von Fehlern eines Datenbestands beschrieben. Bei der Korrektheit liegt dabei das Augenmerk auf die korrekte Erfassung der Daten bzw. Informationen bzgl. der konzeptionellen Realität. Eine Verletzung der Korrektheit liegt z.B. bei groben oder systematischen Fehlern vor.

Metrische Genauigkeit: Die metrische Genauigkeit gibt den Zusammenhang zwischen dem gemessenen und den wahren, bzw. dem plausibelsten Wert an. Die metrische Genauigkeit kann auf kontinuierlich darstellbare Werte angewendet werden. Beispiele hierfür sind die Lagegenauigkeit von Geometrieobjekten oder auch die Genauigkeit von metrisch darstellbaren Attributen der Objekte. Als Maß für die metrische Genauigkeit wird die empirische Standardabweichung definiert.

Semantische Genauigkeit: Die Semantische Genauigkeit gibt den Diskretisierungsgrad der Objekt- und Sachdaten an. Sie wird auf diskret darstellbare Werte bezogen. Im Gegensatz zur metrischen Genauigkeit ist weniger die Messgenauigkeit als vielmehr der Abstraktionsgrad bei der Modellierung der realen Welt ausschlaggebend. Dieser Abstraktionsgrad wird durch den vorhandenen Skalentyp der Attributwerte und der Vielfältigkeit des Attributkatalogs beschrieben.

5.1.2 Qualitätsmerkmale als Entscheidungskriterien

Die nachfolgende Einteilung der Qualitätsmerkmale muss vor dem Hintergrund beachtet werden, dass ein Datenbestand nicht nach seinen eigenen, impliziten Qualitätseigenschaften (Parameterwerte) beurteilt wird, sondern danach, welche Qualität sich bei seiner Transformation in den Zielbestand ergibt.

5.1.2.1 Einteilung der Qualitätsmerkmale

Die Einteilung der Qualitätsmerkmale geschieht vor dem Hintergrund der Reihenfolge ihrer Bewertung. Es ergeben sich die beiden Gruppen der *grundlegenden und inhaltlichen Merkmale*. Zu den grundlegenden Merkmalen zählen die Verfügbarkeit, Konsistenz und Genauigkeit. Sie sind als notwendige Voraussetzungen zu betrachten. Die inhaltlichen Merkmale umfassen die Vollständigkeit, Korrektheit und die Aktualität. Durch sie werden die inhaltlichen Voraussetzungen der Verwendbarkeit festgelegt. So wird z.B. eine Überprüfung der Vollständigkeit, der Korrektheit oder der Aktualität eines Datenbestands hinfällig, wenn dieser nicht verfügbar ist, inakzeptable Konsistenzverletzungen aufweist oder gegen Genauigkeitsforderungen verstößt, die in keinem wirtschaftlichen Rahmen nachbearbeitet werden können.

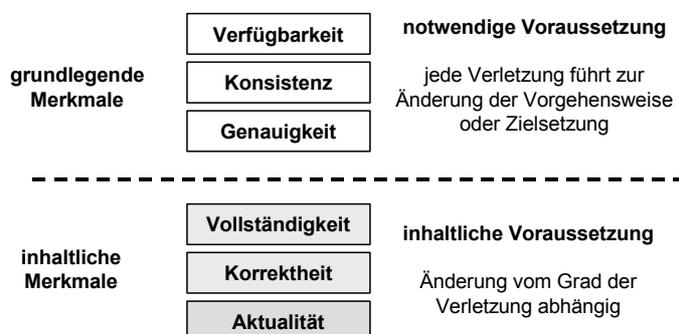


Abbildung 5.2: grundlegende und inhaltliche Qualitätsmerkmale

Sind die grundlegenden Qualitätsmerkmale durch den betrachteten Datenbestand jedoch erfüllt, bzw. nur geringfügig verletzt, so können in einem weiteren Schritt die inhaltlichen Qualitätsmerkmale gegenüber der Qualitätsforderung beurteilt werden.

5.1.2.2 Grundlegende Qualitätsmerkmale eines Datenbestands

Das Qualitätsmerkmal **Verfügbarkeit** (QM^{Vb}) beschreibt die prinzipielle Existenz eigener oder fremder Datenbestände, die zum Zwecke des Aufbaus applikationsspezifischer Datenbestände herangezogen werden können. Bei dieser Beurteilung stehen ausschließlich rechtliche (Nutzungsrechte, Lizenzvereinbarungen) und administrative Regeln im Vordergrund. Im Rahmen einer ersten Entscheidung genügt es folglich, die Verfügbarkeit als BOOLESCHE WERT auszudrücken:

$$QM^{Vb} = \begin{cases} 1 & \text{verfügbar} \\ 0 & \text{nicht verfügbar} \end{cases}$$

Grundsätzlich muss bei der hier zu Grunde liegenden Betrachtungsweise zwischen der Verfügbarkeit des Zieldatenbestandes für Informationsdienstleistungen am Ende der Bereitstellungskette und der Verfügbarkeit von Daten und Informationen aus eigenen und fremden Beständen getrennt werden. Da die Verfügbarkeit im ersten Fall hauptsächlich von Kommunikationswegen und –protokollen (zwischen Anbieter und Kunden) abhängig ist und damit ein rein technisches Realisierungsproblem darstellt, wird im weiteren auf eine nähere Betrachtung dieses Begriffsverständnisses der Verfügbarkeit verzichtet.

Die **Konsistenz** (QM^{Kz}) indiziert die Widerspruchsfreiheit eines Datenbestands. Verletzungen der Konsistenz treten auf, wenn z.B. Daten nicht im für sie vorgesehenem Format gespeichert sind. Eine Prüfung der Konsistenz von Datenbeständen wird gegenüber dem zugrunde liegenden Informationsmodell durchgeführt und kommt deshalb ohne Referenzdaten aus. Werden Datenbestände miteinander vereinigt, verschmolzen oder innerhalb einer Plattform integriert, so müssen notwendigerweise auch die Informationsmodelle beurteilt und ggf. angepasst werden. Da eine Konsistenzbewertung immer nur vor dem Hintergrund eines definierten Informationsmodells stattfinden kann, ist eine Weitergabe konsistenter Eigenschaften von Objekten oder Objektklassen im Sinne einer Vererbung zwischen Quell- und Zieldatenbeständen nicht möglich.

Unter der Voraussetzung, dass diese Programme fehlerfrei und zuverlässig arbeiten, ergibt sich ein Zieldatenbestand, dessen Inhalt (Menge) dem des Quellbestandes entspricht. Ein möglicherweise auftretender Mangel innerhalb der Konvertierungsprogramme (Datentransformationen) wirkt sich dann negativ auf die Vollständigkeit, Korrektheit und die semantische Genauigkeit, aber nicht auf die Konsistenz des Zielbestandes aus. Dieses scheinbare Paradoxon löst sich schnell auf, wenn man bedenkt, dass bei einer Konvertierung eines Attributwertes nur die beiden Möglichkeiten bestehen, dass dieser Wert gemäß dem Informationsmodell des Zieldatenbestandes abbildbar ist oder nicht. Stehen einer Abbildung andere Wertebereiche oder Typdefinitionen entgegen, als sie durch eine Formatkonvertierung lieferbar sind, so fehlen diese Attributwerte im Zielbestand schlichtweg. Damit ist dann die Vollständigkeit und nicht die Konsistenz des Zielbestandes verletzt.

Zur **Kontrolle der Konsistenz** kann auf, ebenfalls in [JOOS, 2000] dargestellte und zitierte Verfahren von z.B. [PLÜMER UND GRÖGER 1996] oder [KAINZ, 1995] gleichermaßen verwiesen werden, wie auf deren Darstellung zur logischen und konzeptionellen Konsistenzprüfung in Geoinformationssystemen. Wichtig an dieser Stelle ist die **dokumentierte Existenz von Prüfverfahren**, mit denen Inkonsistenzen innerhalb Geodatenbeständen erkannt und zur Nachbearbeitung durch den Bearbeiter markiert werden können. Diese Prüfverfahren können auf einzelne Objekte, Objektklassen und Themenbereiche sowie auf einen kompletten Datenbestand angewendet werden.

Das Vorhandensein von Verfahren zur Konsistenzprüfung von Datenbeständen einerseits und die Auswirkung von Konsistenzverletzungen auf die Vollständigkeit und die Korrektheit der transformierten Quellbestände in einen Zieldatenbestand andererseits, sind die Gründe, weshalb hier die Konsistenz als *grundlegendes Merkmal* und nicht als *inhaltliches Merkmal* festgelegt wird. Durch diese Existenz von Konsistenzprüfroutinen kann eigentlich nur von den beiden Zuständen des ungeprüften (Konsistenz ist zu einem bestimmten Prozentsatz verletzt) und geprüften (keine Konsistenzverletzung) Datenbestand ausgegangen werden. Eine Konsistenzprüfung ist somit quasi Voraussetzung zur Nutzung des Bestandes. Dabei bezieht sich die Konsistenzangabe nicht auf eine einzelne Objektinstanz, sondern auf Objektklassen bzw. zu Themenbereiche zusammengefasste Objektklassen. Eine Konsistenzverletzung eines Quellbestandes ist nur dann ein kritischer Kostenfaktor, wenn auch gleichzeitig Datentransformationsverfahren versagen und zu Vollständig-

keits- und Korrektheitsverletzungen des Zieldatenbestandes führen. Dies wiederum führt aber dazu, den betrachteten Datenbestand als untauglich zu bewerten und ihn nicht zum Aufbau eines Zieldatenbestands heranzuziehen.

Unter Zuhilfenahme binärer Zustandsvariablen, die die Quelldatenkonsistenz (QM^{KZ}_{quell}) und das Ergebnis einer Datentransformation (DT) angeben, lässt sich deren Zusammenwirken durch die Wahrheitsfunktion der **Konjunktion** [BRONSTEIN, 1991] beschreiben. Es ergeben die beiden binären Zustandsvariablen wie folgt:

$$QM^{KZ}_{\text{quell}} = \begin{cases} 1 & \text{konsistent} \\ 0 & \text{nicht konsistent} \end{cases} \quad DT = \begin{cases} 1 & \text{fehlerfrei} \\ 0 & \text{nicht fehlerfrei} \end{cases}$$

Das Ergebnis wird als *Wahrheitswert der Quelldatenkonsistenz und Datentransformation* (W^{QKz}_{DT}) bezeichnet.

$$W^{QKz}_{DT} = QM^{KZ}_{\text{quell}} \wedge DT$$

Tabelle 5-1 stellt die Wahrheitstafel für die Operationen Konjunktion und Disjunktion mit zwei binären Zustandsvariablen x_1 und x_2 dar. Das Ergebnis dieser Konjunktion kann nicht als endgültige Entscheidungsgrundlage über die Verwendung des betrachteten Datenbestands herangezogen werden. Das Konjunktionsergebnis „1“ ist lediglich notwendige Voraussetzung für die reine Datenumsetzung.

Tabelle 5-1: Wahrheitstafel

Variable		Konjunktion	Disjunktion
x_1	x_2	$x_1 = x_1 \wedge x_2$	$x_1 = x_1 \vee x_2$
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	1

$$\text{mit der Positiv-Logik } X_i = \begin{cases} 1 & \text{erfüllt} \\ 0 & \text{nicht erfüllt} \end{cases}$$

Ferner kann für den Fall der Quelldateninkonsistenz ($QM^{KZ}_{\text{quell}} = 0$) und dem Versagen von Datentransformationen ($DT=0$) von einem Nicht-Verwenden des Datenbestands ausgegangen werden.

Der Fall einer vorliegenden Quelldateninkonsistenz und funktionierenden Datentransformationsverfahren führt dazu, dass die inkonsistenten Daten aus dem Quellbestand nicht in den Zielbestand überführt werden können, sofern die Transformationsverfahren nicht über Konsistenzprüfroutinen verfügen. Es kommt zu einer Verletzung der Vollständigkeit und Korrektheit der Daten im Zielbestand. Dieselbe Verletzung tritt für den Fall auf, wenn konsistente Quelldaten vorliegen aber die Datentransformation nicht fehlerfrei arbeitet. Durch die Festlegung einer binären Zustandsgröße für die Konsistenz, die auch über einen Relativwert angegeben werden könnte, wird impliziert, die Konsistenzprüfung der Daten und deren Herstellung in der Verantwortung des Datenproduzenten zu belassen und nicht dem Datennutzer zu überlassen.

Das Qualitätsmerkmal **Genauigkeit** (QM^{Gk}) muss in die metrische (QM^{Gk}_{metr}) und semantische Genauigkeit (QM^{Gk}_{sem}) unterschieden werden. Während bei der metrischen Genauigkeit die empirische Standardabweichung als Maß herangezogen werden kann, verbleibt für die semantische Genauigkeit nur der Vergleich zwischen der Anzahl möglicher Werte eines Attributs im Quellbestand gegenüber der Forderung im Zieldatenbestand.

Die **metrische Genauigkeit** wirkt damit als **selektives Merkmal** („Kill-Kriterium“). Eine Verletzung der Genauigkeitsforderung führt unmittelbar zur Korrektheitsverletzung. Liegen z.B. die Koordinatenwerte eines Gebäudepunktes nur mit einer Standardabweichung von $s_{x,y} = \pm 8\text{cm}$ (i.S.d. Helmert'schen Punktfehlers) vor und existiert die Genauigkeitsforderung von $s_{x,y} = \pm 5\text{cm}$, so sind diese Koordinatenwerte als inkorrekt zu werten. Geht man ferner davon aus, dass für alle metrisch zu beschreibenden Objekte eines Geodatenatzes eine bestimmte Genauigkeitsgrenze (Schwellwert) existiert, dann ist dieser Datensatz genau dann als nicht brauchbar zu definieren, wenn eine höhere Genauigkeitsforderung an denselben gestellt wird und damit der Schwellwert überschritten ist. Eine **Verbesserung der metrischen Genauigkeit** ist dann nur durch eine **vollständige Neuerfassung** der Geometrie zu erreichen.

Eine Verletzung der **semantischen Genauigkeit** liegt vor, wenn der Diskretisierungsgrad eines Attributwertes des Quellbestands geringer ist als der geforderte Diskretisierungsgrad desselben Attributwertes im Zielbestand. Eine diesbezügliche Verletzung liegt ebenfalls vor, wenn sich die Typen der Attributwerte ändern.

Liegen aber z.B. für Waldflächen eines GIS Altersindices in 10-jährigen Einteilungsstufen bis zu einem Maximalwert von 170 Jahre vor, so können dem Attribut *<Alter des Bestands>* für ein Umringpolygon einer Waldbestandsfläche die Attributwerte von 1 bis 17 zugewiesen werden. Wird nun eine semantische Genauigkeit in 5-jährigen Einteilungsstufen gefordert, so muss zwangsläufig eine Neuerfassung durchgeführt werden. Diese Neuerfassung beinhaltet in diesem Fall nicht nur die Attributwerte, also die Neuklassifizierung des Waldbestands in 5-Jahres-Schritten, sondern zieht auch eine vollständige Neuerfassung der Umringpolygone der Bestandsflächen nach sich. Das bedeutet, dass nicht nur eine Steigerung der metrischen, sondern auch die der semantischen Genauigkeitsforderung unter Umständen als selektives Merkmal betrachtet werden muss.

Mit Hilfe des in Abbildung 5.3 dargestellten Entscheidungsbaums können die Ergebnisse der Genauigkeitsprüfung nachvollzogen werden. Darin wird deutlich, dass eine Verletzung der metrischen Genauigkeitsforderung unmittelbar zur *Vollständigen Neuerfassung* oder zumindest zur *eigenen Zusatzerfassung* führt. Zu einer Prüfung weiterer Qualitätsmerkmale kommt es nur dann, wenn die metrische und semantische Genauigkeitsforderung erfüllt ist oder die Verletzung der semantischen Forderung durch entsprechende Formatkonvertierungsprogramme korrigiert werden kann.

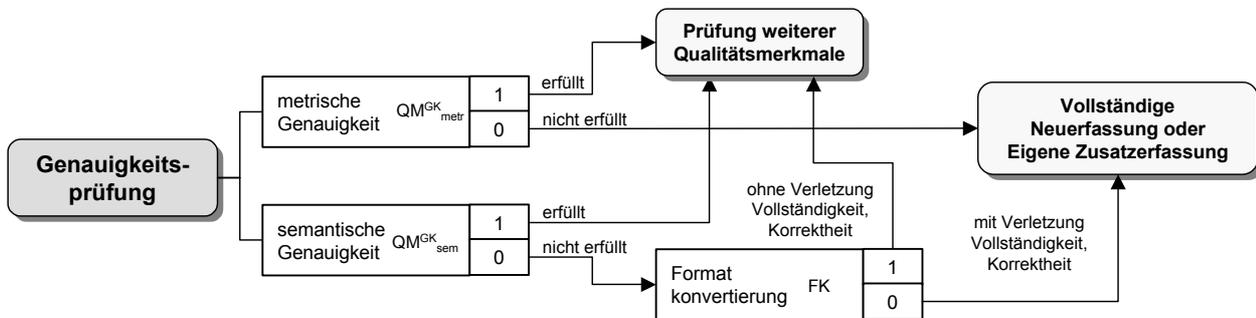


Abbildung 5.3: Ergebnisse der Genauigkeitsprüfung

Wird trotz Formatkonvertierung die semantische Genauigkeitsforderung nicht erfüllt, so wirkt dies unmittelbar auf die Vollständigkeit und Korrektheit. Damit ist das Ergebnis der Genauigkeitsprüfung gleichermaßen von der Beurteilung der metrischen als auch der semantischen Genauigkeit abhängig.

Mit Hilfe der **Disjunktion** kann das Zusammenwirken der Zustandsvariablen QM_{sem}^{GK} und FK mathematisch beschrieben werden. Eine semantische Genauigkeitsverletzung $QM_{sem}^{GK} = 0$ kann durch eine fehlerfreie und vollständige Formatkonvertierung $FK=1$ kompensiert werden. Da eine Verletzung der metrischen Genauigkeit mit demselben Gewicht das Resultat der Genauigkeitsprüfung beeinflusst, ergibt sich der **logische Zusammenhang der Genauigkeitsprüfung** als Kombination von Konjunktion und Disjunktion der Zustandsvariablen zu:

$$QM^{GK} = QM_{metr}^{GK} \wedge (QM_{sem}^{GK} \vee FK)$$

Damit können die möglichen Varianten der binären Zustandsvariablen und die daraus folgende Auswirkung folgendermaßen dargestellt werden (Tabelle 5-2):

Tabelle 5-2: Auswirkungen

QM ^{GK} _{metr}	QM ^{GK} _{sem}	FK	QM ^{GK}	Bemerkungen und Auswirkungen
1	0	1	1	Erfüllte metrische Genauigkeit. Nicht erfüllte semantische Genauigkeit, aber durch erfolgreiche Formatkonvertierung kompensierbar. Prüfung weiterer Merkmale
1	1	1	1	Erfüllte metrische Genauigkeit. Erfüllte semantische Genauigkeit und erfolgreiche Formatkonvertierung. (rein theoretischer Fall) Prüfung weiterer Merkmale
1	1	0	1	Erfüllte metrische Genauigkeit. Erfüllte semantische Genauigkeit keine Formatkonvertierung notwendig. Prüfung weiterer Merkmale
1	0	0	0	Erfüllte metrische Genauigkeit. Die verletzte semantische Genauigkeit kann durch keine Formatkonvertierung kompensiert werden. Vollständige Neuerfassung oder Eigene Zusatzerfassung
0	0	1	0	Die verletzte semantische Genauigkeit kann durch Formatkonvertierung kompensiert werden. Wegen nicht erfüllter metrischer Genauigkeit kommt es zur: Vollständigen Neuerfassung oder Eigenen Zusatzerfassung
0	1	1	0	Erfüllte semantische Genauigkeit macht Formatkonvertierung nicht zwingend notwendig. Wegen nicht erfüllter metrischer Genauigkeit kommt es zur: Vollständige Neuerfassung oder Eigene Zusatzerfassung
0	1	0	0	Erfüllte semantische Genauigkeit macht Formatkonvertierung nicht zwingend notwendig. Wegen nicht erfüllter metrischer Genauigkeit kommt es aber zur: Vollständigen Neuerfassung oder Eigenen Zusatzerfassung
0	0	0	0	Die verletzte semantische Genauigkeit kann nicht durch Formatkonvertierung kompensiert werden. Auch wegen nicht erfüllter metrischer Genauigkeit kommt es zur: Vollständigen Neuerfassung oder Eigenen Zusatzerfassung

Ist die metrische Genauigkeit verletzt, so kann dies nur durch eine *Vollständige Neuerfassung* kompensiert werden. Eine semantische Genauigkeitsverletzung hat im schlimmsten Fall auch eine *Vollständige Neuerfassung* der Daten, in jedem Fall aber eine *eigene Zusatzerfassung* zur Folge.

5.1.2.3 Inhaltliche Qualitätsmerkmale eines Datenbestands

Die *Vollständigkeit*, *Korrektheit* und *Aktualität* eigener oder fremder Datenbestände sind die *inhaltlichen Qualitätsmerkmale*, die als nachgeordnete Beurteilungskriterien zur Entscheidung über die Aufbaumöglichkeiten herangezogen werden. Da sich alle Qualitätsmerkmale immer nur vor dem Hintergrund des zum Datenbestand gehörigen Informationsmodells bestimmen lassen und zu einem neuen Datenbestand zwangsläufig auch ein anderes Informationsmodell gehört, beziehen sich die nachfolgenden Qualitätsforderungen grundsätzlich auf das Informationsmodell des Zieldatenbestands. Diese Forderungen beschreiben gewissermaßen den **Sollzustand** und das Ziel, welches durch den Aufbau eines Geodatenbestandes erreicht werden soll. Über einen Vergleich der Qualitätsforderungen mit denen aus eigenen und fremden Beständen bekannten Qualitätsmerkmalwerten ist ein **Zielerreichungsgrad** formulierbar. Dieser gibt an, zu wie viel Prozent die Qualitätsforderungen durch verfügbare Geodatenbestände erfüllt werden können. In folgenden Abschnitten werden deshalb diese inhaltlichen Merkmale eingehend definiert und Möglichkeiten aufgezeigt, wie diese zu quantifizieren sind.

Das Qualitätsmerkmal **Vollständigkeit** (QM^{Vk}) kann als Relativwert ausgedrückt werden, der angibt, zu wie viel Prozent die theoretische Vollständigkeitsforderung des Zieldatenbestands erfüllt ist. Maßstab zur Beurteilung der Vollständigkeit ist nicht die reale Welt oder die konzeptionelle Realität des untersuchten Bestandes, sondern die dem Zieldatenbestand zugrundeliegende konzeptionelle Realität sowie die räumlichen Gebietsgrenzen. Sind alle geforderten Objektelemente gemäß Menge (Umfang) und Struktur (Datenmodell) im Zielbestand enthalten, so ist die Vollständigkeitsforderung erfüllt.

Eine Verletzung der Vollständigkeit liegt vor, wenn Objektinstanzen, die gemäß Umfang vorhanden sein müssten, fehlen, wenn Attribute der Objektklassen nicht mit Attributwerten besetzt sind oder wenn definierte Objektrelationen nicht gebildet wurden. Fehlt eine gesamte Objektinstanz, dann fehlen auch die entsprechende Anzahl an Attributwerten oder Objektrelationen. Die Vollständigkeitsforderung kann entsprechend Geometrie, Attributen und Relationen getrennt werden:

<p>Geometrie:</p> $QM^{VK}_{geo} = \frac{1}{N_G} \sum n_G$	<p>Attribute:</p> $QM^{VK}_{att} = \frac{1}{N_A} \sum n_A$	<p>Relationen:</p> $QM^{VK}_{rel} = \frac{1}{N_R} \sum n_R$
Mit:		
N_G ...Sollanzahl Objektgeometrien	$\sum n_G$...Summe vorhandener Objektgeometrien	
N_A ...Sollanzahl Attribute	$\sum n_A$...Summe vorhandener Attribute	
N_R ...Sollanzahl Relationen	$\sum n_R$...Summe vorhandener Relationen	

Die hier formulierte Vollständigkeitsforderung bezieht sich auf die Betrachtung der Objektgeometrie als Ganzes. Eine Unterscheidung der Vollständigkeitsforderung der Objektgeometrie in geometrische Primitive (Punkt, Linie, Fläche) findet nicht statt.

Fehlt demzufolge bei einer Objektinstanz eines der notwendigen geometrischen Primitive, so wird dieses Fehlen als eine Vollständigkeitsverletzung der gesamten Objektgeometrie gewertet. Fehlt die Objektgeometrie, so muss auf ein Fehlen des gesamten Objekts geschlossen werden, auch wenn die entsprechenden Objektattribute und Objektrelationen in der Datenbank vorhanden sind. Die dem fehlenden bzw. unvollständigen Objekt theoretisch verbleibenden Attribute und Relationen können erst nach Wiederherstellung des Objekts genutzt werden.

Die **Korrektheit** des Zieldatenbestands ($QM^{Kh, Ziel}$), die durch Homogenisierung von Quelldaten entsteht, hängt bei gleicher konzeptioneller Realität direkt von der Korrektheit der Quelldaten ($QM^{Kh, Quell}$) ab. Ein Fehler im Quellbestand wird ohne Maßnahmen zur Korrektheitsprüfung (z.B. Stichprobenprüfung, 100%-Prüfung) zu einem Fehler im Zielbestand. Selbst bei unterschiedlichen konzeptionellen Realitäten, welche den beiden Datenbeständen zu Grunde liegen, muss von einer Übertragung des Fehlers ausgegangen werden. Es kommt zu einer **Vererbung der Korrektheit**.

Eine trotz Formatkonvertierung nicht erkannte semantische Genauigkeitsverletzung wirkt prinzipiell auf die Korrektheit (unter Umständen auch auf die Vollständigkeit), da durch fehlerhafte Zuordnungsvorschriften des Quelldatenmodells bei Überführung in das Zieldatenmodell Interpretationsfehler entstehen können. Dies betrifft z.B. Objektattribute, die im Quellbestand mit niedrigem Diskretisierungsgrad vorhanden sind und in einen Zieldatenbestand überführt werden müssen, der einen höheren Diskretisierungsgrad für dasselbe Objektattribut vorschreibt.

Beispiel: Wird eine Straße durch das semantische Attribut *Nutzung* mit den möglichen Attributwerten: *Autobahn, Landstraße, Feldweg* in einem Quelldatenbestand beschrieben, so führt eine Klassifizierung der *Nutzung* im Zielbestand durch die Attributwerte: *Autobahn, Bundesstraße, Kreisstraße, Landstraße, Gemeindestraße, Feldweg, Wirtschaftsweg* automatisch zur Verletzung der Korrektheit bzw. bei nicht eindeutiger Entscheidungsmöglichkeit zu einem nicht zugewiesenen Attributwert, also zur Verletzung der Vollständigkeit.

Die Vollständigkeits- und die Korrektheitsforderung kann in diesem Beispiel nur durch Nachbearbeitung sämtlicher Straßen erfüllt werden. Eine Teilautomatisierung der Attributwertzuweisung kann aber geschehen, wenn beide Erfassungsvorschriften des Quell- und des Zieldatenbestandes eine Straße gemäß ihren Eigenschaften als *Autobahn* klassifizieren würden. Für Straßen die im Quellbestand als *Landstraßen* bezeichnet sind können im Zielbestand die Attributwerte *Kreisstraße, Landstraße* oder *Gemeindestraße* zugewiesen bekommen. Gelingt es jedoch Vorschriften zu erstellen, die mit Hilfe weiterer Objekteigenschaften die Attributzuweisung vornehmen, so ist von einer Steigerung der Korrektheit bzw. Vollständigkeit auszugehen.

Dieses Beispiel und die Erläuterung zeigt, dass der Einfluss einer Konsistenzverletzung auf die Korrektheit von vielen Rahmenbedingungen abhängen, die erst bei der tatsächlichen Durchführung der Datenübernahme beschrieben werden können. Der Einfluss der Konsistenz auf die Korrektheit ist damit zwar prinzipiell existent, findet aber hier wegen der

Wahrung der Übersichtlichkeit keine explizite Berücksichtigung im Korrektheitsmaß. Der Einfluss ist darin implizit vorhanden.

Damit ist das Korrektheitsmaß des Zieldatenbestands von der Korrektheit und von der Vollständigkeit des Quellbestandes abhängig, da die vollständigen Daten um inkorrekte Anteile noch vermindert werden. Die inhaltlichen Qualitätsmerkmale werden im Gegensatz zu den grundlegenden Merkmalen mit Hilfe von Relativwerten angegeben. Damit ergeben sich die Korrektheitsmaße des Zieldatenbestands aus dem Produkt der Korrektheit und der Vollständigkeit des Quelldatenbestandes der entsprechenden Objektelemente:

$$\text{Geometrie: } \quad \text{QM}^{\text{Kh,Ziel}}_{\text{geo}} = \text{QM}^{\text{VK}}_{\text{geo}} \cdot \text{QM}^{\text{Kh,Quell}}_{\text{geo}}$$

$$\text{Attribute: } \quad \text{QM}^{\text{Kh,Ziel}}_{\text{att}} = \text{QM}^{\text{VK}}_{\text{att}} \cdot \text{QM}^{\text{Kh,Quell}}_{\text{att}}$$

$$\text{Relationen: } \quad \text{QM}^{\text{Kh,Ziel}}_{\text{rel}} = \text{QM}^{\text{VK}}_{\text{rel}} \cdot \text{QM}^{\text{Kh,Quell}}_{\text{rel}}$$

Die **Aktualität** (QM^{Ak}) gibt an, wie viel von den verfügbaren Objekten gegenüber der konzeptionellen Realität aktuell bzw. veraltet sind. Hier muss zur Bewertung in statischen Daten mit kleiner bis geringer Veränderungsrate pro Zeiteinheit und in dynamische Daten mit mittlerer bis hoher Veränderungsrate unterschieden werden. Während statische Daten in der Regel über Fortführungsprozesse aktuell gehalten werden können, müssen dynamische Daten durch geeignete Informations- und Kommunikations-Technologien ständig aktualisiert werden. Dabei ist in diskrete und ereignisbezogene Aktualisierung zu unterscheiden. Die Fahrplanwechsel von ÖV-Unternehmen, zweimal im Jahr, zählen z.B. zu den diskreten Aktualisierungen. Das Intervall zwischen zwei diskreten Zeitpunkten bestimmt die Gültigkeitsperiode. Stauinformationen auf Autobahnen oder Betriebsstörungen besitzen keinen vorab definierbaren Zeitpunkt des Ereigniseintritts und können nur durch „Online“-Detektoren erfasst werden. Eine Aktualitätsforderung an einen Datenbestand kann nur durch ein, entsprechend der Art und der Änderungszyklen, angepasstes Datenfortführungsverfahren erreicht werden. Im Falle amtlicher Katasterdaten liegen z.B. über Veränderungsnachweise und die vermessungstechnischen Unterlagen sämtliche zur Dokumentation der Fortführung des Liegenschaftskatasters notwendigen Unterlagen vor. Das Aktualitätsmaß des Zieldatenbestandes lässt sich wie das Vollständigkeitsmaß über das Produkt der Aktualität und der Vollständigkeit der Quelldaten beschreiben:

$$\begin{array}{lll} \text{Geometrie:} & \text{Attribute:} & \text{Relationen:} \\ \text{QM}^{\text{Ak}}_{\text{geo}} = \frac{1}{N_G} \sum n_G \cdot \text{QM}^{\text{VK}}_{\text{geo}} & \text{QM}^{\text{Ak}}_{\text{att}} = \frac{1}{N_A} \sum n_A \cdot \text{QM}^{\text{VK}}_{\text{att}} & \text{QM}^{\text{Ak}}_{\text{rel}} = \frac{1}{N_R} \sum n_R \cdot \text{QM}^{\text{VK}}_{\text{rel}} \end{array}$$

mit:

N_G ...Sollanzahl Geometrie	$\sum n_G$...Summe aktueller Objektgeometrien
N_A ...Sollanzahl Attribute	$\sum n_A$...Summe aktueller Attribute
N_R ...Sollanzahl Relationen	$\sum n_R$...Summe aktueller Relationen

Die Aktualitätsmaße der Geometrie, der Attribute und der Relationen besitzen einen Wertebereich zwischen $[0, \dots, 1]$ und geben an, mit welcher Aktualität die vorhandenen Objektelemente (Vollständigkeit) vorliegen. Eine Verletzung der Aktualitätsforderung führt zur Nicht-Verfügbarkeit der benötigten Information. Das Aktualitätsmaß des Zieldatenbestands ist bei der Übernahme von Daten aus Quellbeständen von der Aktualität der Quelldaten und dem Vollständigkeitsmaß abhängig. Um zur Aktualität des Zieldatenbestands zu kommen, muss die Aktualität des Quellbestandes mit dessen Vollständigkeitsmaß multipliziert werden, da nur diejenigen Elemente aktuell bzw. veraltet sein können, die auch vollständig vorliegen. Damit stehen die inhaltlichen Qualitätsmerkmale (QM), welche zur Prüfung innerhalb des Entscheidungsdiagramms genutzt werden sowie deren Wertebereiche fest. Es sind die:

- Vollständigkeit $[0,0 \dots 0,999]$
- Korrektheit $[0,0 \dots 0,999]$
- Aktualität $[0,0 \dots 0,999]$

Diese Merkmale sind, wie dargestellt, bezüglich Objektgeometrie, Objektattribute und Relationen gesondert zu bewerten und anzugeben. Da eine 100 %ige Vollständigkeit, Korrektheit bzw. Aktualität theoretisches Ziel ist, aber praktisch nie erreicht werden kann, wird hier eine Intervallgrenze des Wertebereichs von 99,9% festgelegt. Nach diesen Ausführungen sind die drei Merkmale Vollständigkeit, Korrektheit und Aktualität keine voneinander unabhängigen Merkmale. Vielmehr ist bei der Analyse des Zieldatenbestands stufenweise vorzugehen, um eine Aussage darüber zu treffen, wie viele vollständig vorhandene Datenelemente auch korrekt bzw. aktuell sind.

5.1.3 Qualitätsforderung an den Zieldatenbestand

Mit Hilfe dieser inhaltlichen Qualitätsmerkmale (QM) kann die Qualitätsforderung (QF) an den Zieldatenbestand formuliert werden. Die Qualitätsforderung an den Zieldatenbestand besteht aus jeweils einem Merkmalsatz von Qualitätseinzelforderungen (QeF) pro Objektklasse des Zieldatenmodells. Die Forderung richtet sich speziell an die Objektklassen, da einzelne Objektinstanzen oder übergeordnete Objektbereiche entweder eine zu spezielle oder zu unscharfe Festlegung der Qualitätsforderung zulassen würden. Die Qualitätsforderung an den Zieldatenbestand ist eine abstrakte Größe ohne eigenständige Aussagekraft. Diese ergibt sich erst durch eine Diskretisierung der Qualitätsforderungen auf die *Geometrien, Attribute und Relationen* der betrachteten Objektklasse.

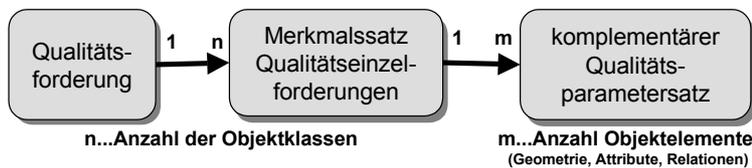


Abbildung 5.4: Diskretisierung der Qualitätsforderung

Das Ziel einer solchen Diskretisierung der *Merkmalsätze der Qualitätseinzelforderungen* (QeF) ist es, bezüglich den Elementen einer Objektklasse konkrete Aussagen über den Homogenisierungs- und Nacherfassungsaufwand geben zu können. Dafür werden hier **komplementäre Qualitätsparameter (kQP)** eingeführt, die als zweiter Diskretisierungsschritt (vgl. Abbildung 5.4) der abstrakten Qualitätsforderung zu verstehen sind. Sie ergeben sich grundsätzlich aus dem Zusammenhang:

$$kQP = 1 - QM$$

Diese komplementären Qualitätsparameter beschränken sich dabei auf die inhaltlichen Qualitätsmerkmale. Mit deren Hilfe wird die Differenzierung der Qualitätsforderungen entsprechend der Anzahl der Objektelemente einer Objektklasse ermöglicht. Ist eine Objektklasse beispielsweise mit einem Geometrieelement modelliert und besitzt 4 metrische und 6 semantische Attribute sowie 3 Objektrelationen, dann setzt sich der Merkmalsatz der Qualitätsforderung dieser Objektklasse aus $m = 14$ komplementären Qualitätsparametersätzen zusammen.

Diese Qualitätsparameterwerte sind relative Maßzahlen, die die Qualität (Vollständigkeit, Korrektheit, Aktualität) im Verhältnis zur zu Grunde gelegten inhaltlichen Forderung des Zieldatenbestandes beschreiben. Was mit den Qualitätsmerkmalen beschrieben werden soll, legt letztendlich die Datenspezifikation der Anwendung fest. Aus den geplanten Anwendungen und deren Rahmenbedingungen werden Sollgrößen abgeleitet, die als Maß für eine vollständige Einhaltung der Qualitätsforderungen herangezogen werden. So ist z.B. eine Sollgröße für die verkehrliche Anwendung des Pkw-Routings die Anzahl aller öffentlicher und befahrbarer Straßen, bezogen auf eine räumliche Einheit. Eine diesbezüglich weitere Sollgröße, z.B. für die metrische Lagegenauigkeit derselben Straße, wird durch die Erfassungsgenauigkeit primärer oder sekundärer Quellen beeinflusst und liegt in Navigationsdatensätzen (GDF) derzeit mit einer Genauigkeit von +/- 3 bis 5m vor.

An diesen Sollgrößen orientieren sich die Qualitätsforderungen. Werden die Sollgrößen erreicht, so sind die Qualitätsforderungen erfüllt. Um aber zu den komplementären Qualitätsparametern zu kommen, müssen die vorhandenen Quellbestände den Sollgrößen gegenübergestellt werden. Aus dieser Gegenüberstellung lässt sich dann das Qualitätsniveau ermitteln, welches beim Aufbau des Zieldatenbestands aus Quellbeständen entstehen würde. Das erreichbare Qualitäts-

niveau einer Objektklasse lässt sich mit den Merkmalsätzen und deren Parameterwerten der Qualitätsmerkmale zahlenmäßig ausdrücken.

Abbildung 5.5 verdeutlicht exemplarisch die Qualitätseinzelforderungen an Elemente der Objektklassen. In dieser Abbildung ist auch das Qualitätsniveau für die Geometrie der Objektklasse 1 (OK1), wie es sich möglicherweise durch eine Datenübernahme aus Quellbeständen ergibt, dargestellt. Darin sind die Qualitätseigenschaften der Quellbestände von den Qualitätsforderungen des Zieldatenbestands grundsätzlich zu unterscheiden, da die konzeptionelle Realität zwischen Ziel- und Quellbeständen unterschiedlich sein kann. So kann z.B. bei einer Aktualität von 99,9%, bezogen auf den Quellbestand, nicht automatisch auf dieselbe Aktualität im Zieldatenbestand geschlossen werden. Bewertet man die Aktualität des Quellbestandes vor dem Hintergrund der konzeptionellen Realität des Zieldatenbestandes, so können sich wesentlich geringere Werte für die einzelnen Qualitätsmerkmale ergeben.

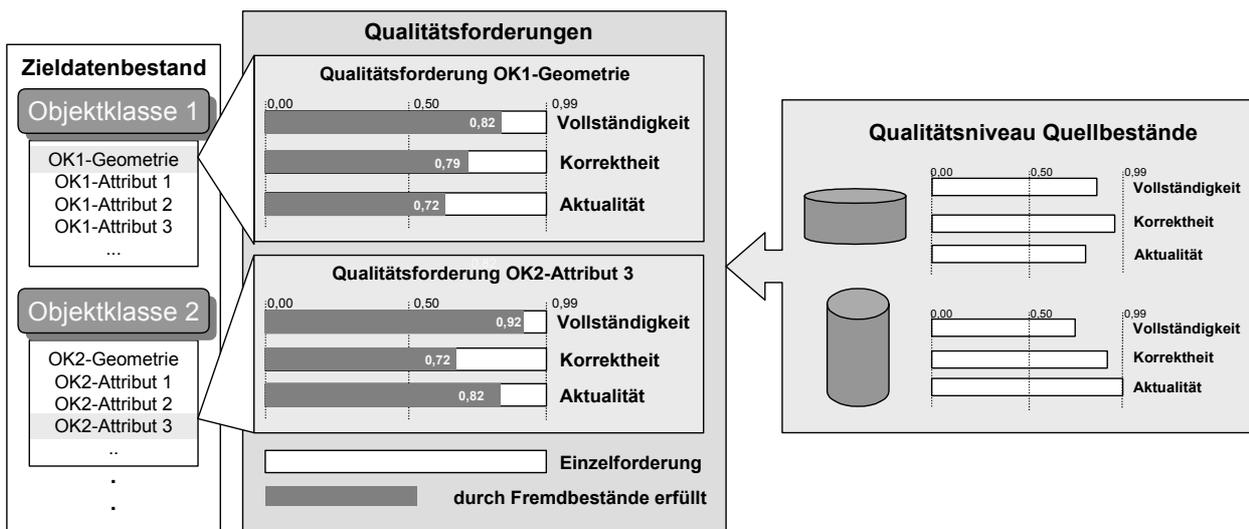


Abbildung 5.5: Qualitätsforderungen und erreichbares Qualitätsniveau aus Quellbeständen

Der in Abbildung 5.5 dargestellte Übergang vom Qualitätsniveau der Quellbestände zur Qualitätsforderung des Zieldatenbestands drückt beispielhaft für die Geometrie der OK1 aus, dass diese zu 82% vollständig, zu 79% korrekt und zu 72% aktuell aus Fremdbeständen entnommen werden kann. Das bedeutet umgekehrt, dass für die Objektklasse 1 des Zieldatenbestandes ein Nachbearbeitungsaufwand bezüglich der Vollständigkeit von 18%, für die Korrektheit von 21% und für die Aktualität von 28% der Objektgeometrie notwendig ist, um die Qualitätseinzelforderung zu erfüllen. Damit steht konkret eine Qualitätseinzelforderung in Form komplementärer Qualitätsparameterwerte für die Objektgeometrie der Objektklasse 1 zur Verfügung. Durch weitere Gegenüberstellungen für die einzelnen Elemente der Objektklasse ergibt sich ein Gesamtbild in Form nachfolgender Tabelle des Merkmalsatzes der Qualitätseinzelforderung an die Objektklasse 1 des Zieldatenbestandes.

Tabelle 5-3: Beispiel für einen Merkmalsatz der komplementären Qualitätsparameter

Objektklasse 1	Unvollständigkeit	Inkorrektheit	Nicht aktuell
	kQP ^V	kQP ^{KR}	kQP ^A
Geometrie	18%	21%	28%
Attribut 1	23%	16%	8%
Attribut 2	80%	33%	14%

Diese Tabelle stellt die Basis zur weiteren Ermittlung des Aufwandes der Geodatenbereitstellung dar. Mit ihrer Hilfe lassen sich die Entscheidungen innerhalb der Aufbaumöglichkeiten begründen.

5.2 Fehlerrate in Abhängigkeit inhaltlicher Qualitätsmerkmale

In diesem Abschnitt wird der Zusammenhang zwischen den inhaltlichen Qualitätsmerkmalen Vollständigkeit, Korrektheit und Aktualität und der Fehlerrate der Grundgesamtheit der Elemente des Geodatenbestandes betrachtet. Die Fehlerrate der Grundgesamtheit soll fortan als das Gütekriterium zur Qualitätsbeurteilung betrachtet werden. Eine einzelne, nacheinanderfolgende Beurteilung der inhaltlichen Qualitätsmerkmale ist zwar prinzipiell durchführbar und auch mit der nachfolgenden Methodik zur Berechnung des Kosteneinflusses vereinbar, jedoch die Zusammenfassung zu einem Qualitätsmerkmal vereinfacht die Qualitätsbetrachtung des Geodatenbestandes sowie die darauf aufbauende Modellrechnung erheblich. Für die Ableitung der Fehlerrate soll hier von den Qualitätsmerkmalen Vollständigkeit, Korrektheit und Aktualität ausgegangen werden.

Die Zusammenfassung der Qualitätsmerkmale geschieht auch vor dem Hintergrund, dass die Qualität eines Datenbestandes vor dessen Prüfung als unbekannt angenommen werden muss. So kann nicht davon ausgegangen werden, für einen vorliegenden Datenbestand explizit die Fehlerrate bzw. die Merkmalswerte der Vollständigkeit, Korrektheit und Aktualität angeben zu können. Sie ist Gegenstand der Prüfung. Der Zweck einer solchen Prüfung ist jedoch nicht nur das alleinige Ermitteln der Fehlerrate, sondern die damit eng verbundene Aussage über die Erfüllung einer bestehenden Qualitätsforderung. Unter Verwendung der inhaltlichen Qualitätsmerkmale kann die Zustandsgröße der Datenqualität eines Elements als Vektor wie folgt angegeben werden:

$$Q_E = \begin{pmatrix} Q_E^V \\ Q_E^K \\ Q_E^A \end{pmatrix} \quad \begin{array}{l} Q_E^V \dots \text{Qualitätsmaß Vollständigkeit} \\ Q_E^K \dots \text{Qualitätsmaß Korrektheit} \\ Q_E^A \dots \text{Qualitätsmaß Aktualität} \end{array} \quad (5.1)$$

Ferner seien:

n	sei die betrachtete Anzahl Datenelemente	
n_v	seien die vollständig vorhandenen Daten mit	$n_v = Q_E^V n$
n_{nv}	seien die fehlenden Elemente	$n_{nv} = n - n_v$ bzw. $n_{nv} = (1 - Q_E^V) n$
n_{nk}	seien die nicht korrekten Elemente	$n_{nk} = (1 - Q_E^K) n_v$
n_{na}	seien die nicht aktuellen Elemente	$n_{na} = (1 - Q_E^A) n_v$

Damit kann die Fehlerquote aus den Qualitätsmerkmalen abgeleitet werden. Der Zusammenhang ergibt sich zu:

$$\begin{aligned} F_q &= (1 - Q_E^V) + Q_E^V (1 - Q_E^K) + Q_E^A (1 - Q_E^A) \\ &= 1 + Q_E^V - Q_E^V (Q_E^K + Q_E^A) \end{aligned} \quad (5.2)$$

Dieser Zusammenhang gilt für den Fall, dass die Korrektheit und Aktualität unabhängig voneinander betrachtet werden, obwohl eine Verletzung der Aktualität eine Verletzung der Korrektheit mit sich bringt bzw. dessen Ursache ist.

Eine Rückführung der gesuchten Qualitätsmerkmalswerte aus der Fehlerrate ist ungleich schwerer, bzw. nicht eindeutig anzugeben. So kann selbst bei Kenntnis der Vollständigkeit nur ein gemeinsamer prozentualer Anteil für die Aktualität und Korrektheit aus dieser Formel abgeleitet werden. Die endgültige Angabe der Merkmalswerte ist nur nach erfolgter Prüfung möglich.

5.3 Der Begriff der Qualitätssicherung

Bis zum Jahr 1993 wurde der Begriff der Qualitätssicherung (QS) als Oberbegriff für sämtliche qualitätsbezogenen Maßnahmen und Zielsetzungen verwendet. Aufgrund einer Empfehlung der Deutschen Gesellschaft für Qualität (DGQ), wegen eines international bedingten Bedeutungswandels [DGQ11-04, 1994], welcher beinhaltet, den Begriff der Qualitätssicherung zur Vermeidung von Verwechslungen nicht mehr allgemeingültig zu verwenden, wird seit dieser Zeit die „Gesamtheit der qualitätsbezogenen Tätigkeiten und Zielsetzungen“ in der Literatur unter dem Oberbegriff des **Qualitätsmanagements** zusammengefasst.

Dieses Qualitätsmanagement lässt sich in die Funktionen Qualitätsplanung, Qualitätslenkung, Qualitätsprüfung, Qualitätsverbesserung, Qualitätsmanagementdarlegung und Qualitätsaudit unterscheiden.

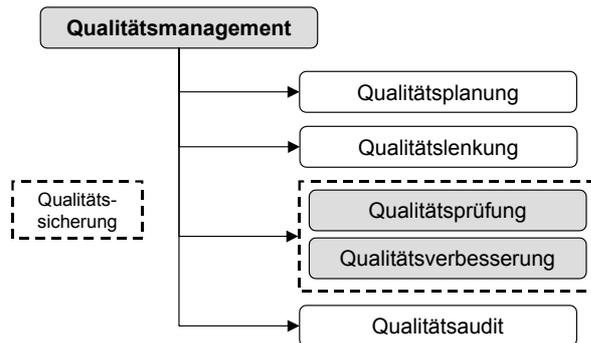


Abbildung 5.6: Abgrenzung Qualitätssicherung [DGQ11-04,1994]

Der Begriff der Qualitätssicherung umfasst darin die Funktionen der Qualitätsprüfung und der Qualitätsverbesserung. Der Begriff der Qualitätssicherung wird nach DIN 8402 in den Fällen weiterverwendet, in denen konkrete Tätigkeiten angesprochen werden, die dazu dienen, bei der Unternehmensleitung oder bei Kunden und anderen angemessenes Vertrauen zu schaffen, damit eine Einheit die Qualitätsforderung erfüllt.

Die Aufgabe der **Qualitätsplanung** ist das Auswählen, Klassifizieren und Gewichten der Qualitätsmerkmale sowie das Konkretisieren der Qualitätsforderungen unter Berücksichtigung von Anspruchsniveau und Realisierungsmöglichkeiten [GEIGER, 1998]. Die **Qualitätslenkung** überwacht und korrigiert die Realisierung einer Einheit mit dem Ziel, die Qualitätsforderung zu erfüllen. Dabei werden die Ergebnisse von **Qualitätsprüfungen** mit den Vorgaben aus der Qualitätsplanung verglichen und bei Abweichungen Korrekturmaßnahmen durchgeführt. Durch die Qualitätsprüfung wird festgestellt, inwieweit eine Einheit die Qualitätsforderung erfüllt. Unter der **Qualitätsverbesserung** werden alle Maßnahmen zur Steigerung von Effektivität und Effizienz in Tätigkeiten und Prozessen verstanden. Die Qualitätsverbesserung nutzt sowohl dem Kunden als auch dem Unternehmen.

Die Qualitätsplanung, -prüfung, -lenkung und -verbesserung bilden zusammen einen operativen und evolutionären Regelkreis [DGQ11-04,1994]. Beim operativen Regelkreis wird bei festgestellten Abweichungen lenkend in Fertigungsprozesse eingegriffen, um die Entstehung von Fehlern zu verhindern oder um das Auftreten weiterer Fehler zu unterbinden.

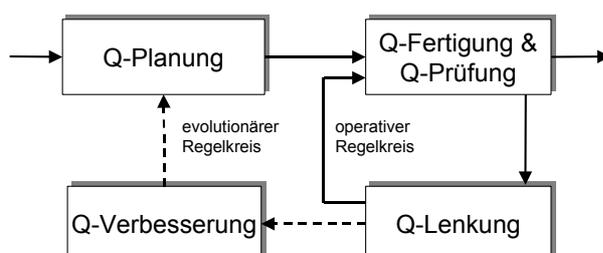


Abbildung 5.7: Operativer und evolutionärer Qualitätsregelkreis [DGQ11-04, 1994]

Der evolutionäre Regelkreis führt auf Basis von Erkenntnissen und Resultaten aus der Fertigung zu Verbesserungen, die wieder in die Qualitätsplanung einfließen. Ziel der evolutionären Regelung ist das Verhindern von Fehlern in der Zukunft.

Im folgenden werden die **Kosten der Qualitätsprüfung und Qualitätsverbesserung** (Fehlerkorrektur) von Geodatenbeständen ohne den umfassenden Rahmen des Qualitätsmanagements betrachtet, da die Kosten dafür in allen Bereichen einer Organisation und bei allen Tätigkeiten anfallen. Diese Kostenanteile können prinzipiell nicht alle erfasst werden und sie sind nicht für die betriebswirtschaftliche Aufgabe des Vergleichs von Kosten und Aufwand mit dem Ertrag geeignet. [GEIGER, 1998] stellt fest, dass es z.B. unmöglich ist, den qualitätsbezogenen Kostenanteil bei Tätigkeiten zu isolieren und zu erfassen. Dabei greift er unter anderem in seiner Begründung auf die in der DIN 55350-11 enthaltene Formulierung zurück:

„Die Beurteilung der Qualitätskosten dient...nicht der Erfassung der Kosten der Qualitätssicherung oder der Ermittlung der Kosten für die Qualität einer Einheit.“

Aus diesem Grund ist eine nähere, begriffliche Definition des Zusammenhangs zwischen Qualität und Kosten notwendig. Es wird deshalb fortan von **qualitätsbezogenen Kosten** gesprochen. Aus [DGQ11-04] stammt folgende Definition:

Qualitätsbezogene Kosten sind Kosten, die vorwiegend durch Qualitätsforderungen verursacht sind. Das heißt es sind Kosten die:

- durch Tätigkeiten der Fehlerverhütung,
- durch planmäßige Qualitätsprüfungen,
- durch intern oder extern festgestellte Fehler sowie
- durch die externe QM-Darlegung verursacht sind.

Diese vier Punkte sind in DIN 55350-11 (1995-08) als Elemente der qualitätsbezogenen Kosten definiert (kurz: QK-Elemente). Diese vier QK-Gruppen sind in Abbildung 5.8 dargestellt.

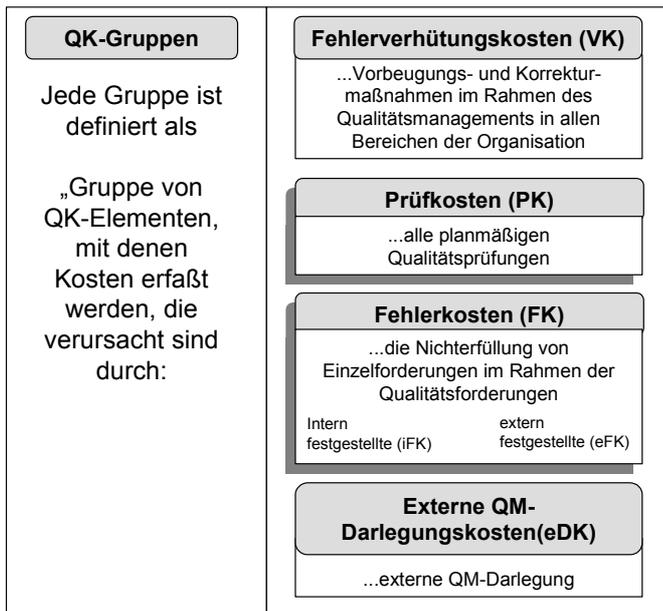


Abbildung 5.8: Die vier QK-Gruppen (Gruppen qualitätsbezogener Kosten)

Die weiter folgende Betrachtung der Qualitätssicherung beschränkt sich auf die beiden QK-Elemente: Prüfkosten (PK) und Fehlerkosten (FK)

5.4 Möglichkeiten zur Qualitätssicherung von Geodatenbeständen

Grundsätzlich kann, wie in Abschnitt 0 erläutert, die Qualität von Geodatenbeständen durch einen Satz von Merkmalen und deren Ausprägung beschrieben werden. Diese Merkmale lassen sich entsprechend ihres Einflusses auf die Wahl der Aufbaustrategie in grundlegende und inhaltliche Merkmale unterscheiden. Da die grundlegenden Merkmale (Verfügbarkeit, Konsistenz, Genauigkeit) die notwendige Voraussetzung der Nutzung des Datenbestandes sind und eine Verletzung automatisch zu einer Änderung der Aufbaustrategie (Nicht-Nutzung des Bestandes) führt, stellen die inhaltlichen Qualitätsmerkmale (Vollständigkeit, Korrektheit und Aktualität) diejenigen Merkmale dar, anhand derer die Möglichkeiten zur Qualitätssicherung hier untersucht werden sollen. Prinzipiell existieren für Geodatenbestände die Möglichkeiten zur Qualitätssicherung in Form von:

- Vollprüfungen und
- Stichprobenprüfungen.

Beide Qualitätssicherungsmöglichkeiten können nur mit Hilfe von Referenzdaten, die ein höheres Qualitätsniveau besitzen, durchgeführt werden. Inwieweit die zu prüfenden Daten mit den Referenzdaten im Rahmen der geforderten Qualitätsziele übereinstimmen, kann nur dann automatisch durchgeführt werden, wenn erstens die Referenzdaten digital vorliegen und zweitens diese auch nach dem gleichen Datenmodell erfasst worden sind. Bei einer solchen digitalen Prüfung sind Algorithmen notwendig, die identische Objekte in beiden Datenbeständen einander zuordnen und entsprechend der Qualitätskriterien miteinander vergleichen können.

Durch die unterschiedliche räumliche und inhaltliche Dichte der Geodatenobjekte innerhalb des abdeckenden Gebietes der Grundgesamtheit, stellt sich die Frage nach der richtigen Auswahl der Stichproben. So kann z.B. aus einer Stichprobenprüfung, welche Geodatenobjekte überwiegend aus ländlichen Gebieten enthält, nicht auf die Qualität der Geodatenobjekte in dicht besiedelten Gebieten geschlossen werden. Es ist demzufolge zweckmäßig, aus einer entsprechend großen Grundgesamtheit, die bei Geodatenbeständen vorhanden ist, entsprechende **Los**e auszuwählen, die hinsichtlich ihrer Metadaten (Objektdichte, Objektthemenbereiche, Größe, etc.) eine miteinander vergleichbare Homogenität aufweisen.

JOOS gibt in seiner Arbeit die Möglichkeiten zur Auswahl dieser Lose über die Verfahren des *Ziehens von Objektidentifikatoren*, des *Landkartenverfahrens mit Koordinaten oder Linienauswahl* und die *Flächenstichprobe* als Möglichkeiten an. Auf die Auswahlverfahren geeigneter Lose zur Berücksichtigung der unterschiedlichen räumlichen und inhaltlichen Dichte einer Grundgesamtheit von Geodatenobjekten soll hier nicht weiter eingegangen werden. Sie werden als bekannt vorausgesetzt und auf die Literatur verwiesen. (z.B. [MENGENS, 1973], [CASPARY, 1992], [LAWFORD, 1995]).

Diese Arbeit geht fortan zur Beurteilung der Geodatenqualität nicht von der Grundgesamtheit des gesamten Geodatenbestand aus, sondern betrachtet über die Auswahl von homogenen Losen deren Umfang als Grundgesamtheit aus der zur Prüfung Stichproben entnommen werden (Abbildung 5.9).

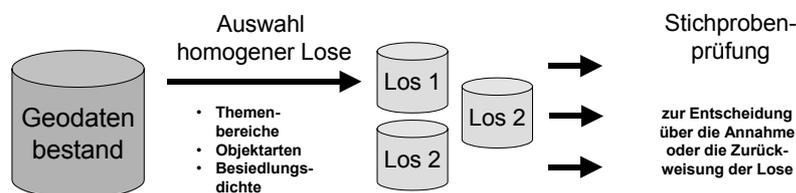


Abbildung 5.9:
Auswahl von Losen zur Stichprobenprüfung

Die Elemente der Gesamtheit können hinsichtlich verschiedener Merkmale X , die ihrerseits durch einer reelle Zahl x charakterisiert sind, untersucht werden. Die Verteilung dieses Merkmals in der Grundgesamtheit wird durch die Verteilungsfunktion $F(x)$ beschrieben.

Wird jedes Geodatenelement bezüglich der inhaltlichen Qualitätsmerkmale Vollständigkeit, Korrektheit und Aktualität bewertet, dann lassen sich die Objekte der Grundgesamtheit in fehlerhafte und fehlerfreie Objekte einteilen und über die Merkmalausprägung x beschreiben ($x=1$ fehlerbehaftet; $x=0$ fehlerfrei).

5.4.1 Geeignete Verteilfunktion bei der Stichprobenprüfung von Geodaten

Es stellt sich hier die Frage nach der Gesetzmäßigkeit der Verteilung fehlerhafter Objekte in einer Grundgesamtheit. Entsprechend dieser Gesetzmäßigkeit ist die Hypergeometrische Verteilung heranzuziehen (vgl. [FISZ, 1970], [SACHS, 1968], [RASCH, 1968]).

Hypergeometrische Verteilung: Werden Stichproben durch „Ziehen ohne Zurücklegen“ entnommen, dann ist die hypergeometrische Verteilung zu verwenden. Die Wahrscheinlichkeit innerhalb einer Stichprobe ein weiteres fehlerhaftes Objekt zu finden, ist deshalb davon abhängig, wie viele Objekte unter den zuvor untersuchten Objekten schon fehlerhaft waren. Die Wahrscheinlichkeitsfunktion der Hypergeometrischen Verteilung ergibt sich aus dem Quotienten der für das Ereignis $X_n=x$ günstigsten Fälle mit der Gesamtzahl der möglichen Fälle zu:

$$P(X_n = x) = \frac{\binom{X}{x} \binom{N-X}{n-x}}{\binom{N}{n}}, \text{ mit } 0 \leq x \leq n; x \leq X; n-x \leq N-X \tag{5.3}$$

Intensiv angewendet wird diese Verteilung bei Problemen, die mit der Qualitätsüberwachung zusammenhängen [Sachs,1968]. Sie ist deshalb diejenige Wahrscheinlichkeitsfunktion, welche fortan den Berechnungen zu Grunde gelegt wird. Da bei einer großen Grundgesamtheit des Loses N die Wahrscheinlichkeitsfunktion wegen der Fakultätsberechnung numerische Schwierigkeiten mit sich bringt, wird auf die in [FISZ,1970] hergeleitete Produktformel (5.7) zurückgegriffen.

$$P(X_n = x) = \prod_{i=0}^{x-1} \frac{n-i}{x-i} \cdot \prod_{i=0}^{x-1} \frac{x-i}{N-n+x-i} \cdot \prod_{i=0}^{n-x-1} \frac{N-X-i}{N-i}, \text{ mit } i, n, x \ll N, X \tag{5.4}$$

In folgender Abbildung 5.10 sind die Werte der Wahrscheinlichkeitsfunktion der hypergeometrischen Verteilung für das Antreffen von x fehlerhaften Elementen in Abhängigkeit von der Grundgesamtheit und des Stichprobenumfangs dargestellt. Für dieses Beispiel wurde der Stichprobenumfang mit n = 2.000 konstant gehalten. Der Umfang der Grundgesamtheit variiert von N = 3.000 bis 100.000. Gleichzeitig wurde von einem konstanten Fehleranzahl M = 16 der Grundgesamtheit ausgegangen. Anhand folgender Abbildung kann gezeigt werden, dass je kleiner das Verhältnis zwischen Stichprobe und Grundgesamtheit ist, um so unwahrscheinlicher ist es, einen Fehler durch die Stichprobe aufzudecken.

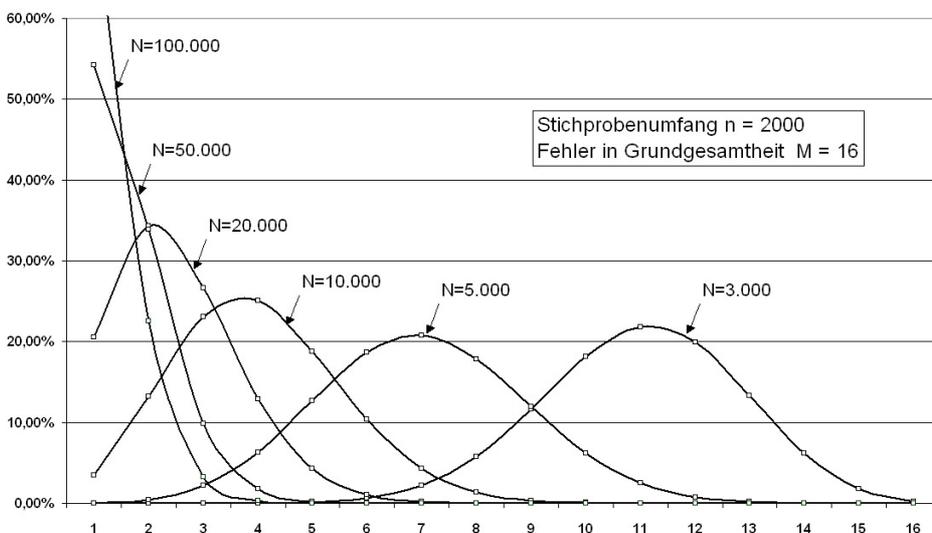


Abbildung 5.10: Die Werte der Wahrscheinlichkeitsfunktion der hypergeometrischen

Nähert sich hingegen der Stichprobenumfang der Grundgesamtheit an, dann ist die theoretischen Wahrscheinlichkeit die 16 fehlerhaften Objekte zu identifizieren bei 100 %. Anders ausgedrückt bedeutet dies für die Qualitätsforderung

von 100 % einen Stichprobenumfang entsprechend der Grundgesamtheit. In diesem Fall spricht man von einer Vollprüfung des gesamten Geodatenbestandes. Restfehler können allerdings trotz Vollprüfungen nicht vollständig ausgeschlossen werden.

5.4.2 Ergebnisbewertung der Stichprobenuntersuchung

Das statistische Instrument von Abnahmeprüfungen die eingesetzt werden, um über die Annahme oder Rückweisung eines Loses zu entscheiden, sind Stichprobenprüfpläne, die entweder einfache, doppelte oder mehrfache Stichproben in die Entscheidung mit einbeziehen. Diese Stichprobenprüfpläne entsprechen einem Parametertest mit den Hypothesen:

Nullhypothese	H_0 : das Prüflös genügt den Anforderungen
Alternativhypothese	H_A : das Prüflös genügt nicht den Anforderungen

Die Anwendung eines solchen Prüfplans entspricht der Anwendung eines Parametertests mit der Nullhypothese H_0 : $p \leq p_0$ gegen die Alternativhypothese H_A : $p > p_0$ wobei p_0 die so genannte kritische Qualitätslage (maximal zulässige Ausschussquote) definiert. Man entscheidet sich für die Nullhypothese, wenn die Testfunktion $X_n^T < C$ ausfällt und gegen H_0 wenn die Testfunktion $X_n^T > C$ ist. Der Ablehnbereich des Tests ist also $CR = \{c+1, c+2, \dots\}$, da eine große Zahl fehlerhafter Einheiten in der Stichprobe auch für eine hohe Ausschussquote im Prüflös entspricht. Von einem statistischen Test muss verlangt werden, dass er angibt, für welche Werte p die Nullhypothese H_0 angenommen und für welche Werte sie abgelehnt werden soll. Der Test muss für die Realisationen von p einen Annahmehereich und einen Ablehnungsbereich bezüglich H_0 angeben. Diesen Bereich nennt man auch den kritischen Bereich eines Hypothesentests.

Durch diesen Bereich entsteht eine Unsicherheit, ob eine für richtig oder falsch angenommene Nullhypothese auch tatsächlich richtig oder falsch ist. Man spricht von einem Fehler 1. Art, wenn man H_0 verwirft, obwohl sie richtig ist und von einem Fehler 2. Art wenn man sie annimmt, obwohl sie falsch ist. Die Wahrscheinlichkeit, mit der die Ablehnung der Nullhypothese in einem Test fälschlicherweise erfolgt, wird auch als Irrtumswahrscheinlichkeit oder als Signifikanzniveau α bezeichnet. Die Wahrscheinlichkeit, mit der die Nullhypothese fälschlicherweise als richtig angenommen wird, wird mit β (Operationscharakteristik) bezeichnet. Folgende Tabelle fasst die Entscheidungen bei einem statistischen Test zusammen.

Tabelle 5-4: Entscheidungen bei einem statistischen Test [RASCH, 1968]

Wahrer Sachverhalt	Entscheidung	Ergebnis der Entscheidung	Wahrscheinlichkeit des Ergebnisses
H0 richtig HA falsch	H0 angenommen (HA abgelehnt)	Richtige Entscheidung	Sicherheitswahrscheinlichkeit $1 - \alpha$
	H0 abgelehnt (HA angenommen)	Fehler 1. Art	Irrtumswahrscheinlichkeit, Risiko erster Art, α
H0 falsch HA richtig	H0 angenommen (HA abgelehnt)	Fehler 2. Art	Risiko zweiter Art, Signifikanzniveau, β
	H0 abgelehnt (HA angenommen)	Richtige Entscheidung	Gütefunktion $1 - \beta$

5.5 Die Operationscharakteristik

Ein solcher statistischer Test wird durch seine Operationscharakteristik (OC) vollständig beschrieben. Die Funktion der Operationscharakteristik $L(p)$ gibt in Abhängigkeit von der wahren Qualitätslage im Los die Wahrscheinlichkeit der Annahme des Loses an. Die OC-Funktion definiert sich nach [FISZ,1970] in (5.5) und ergibt sich in Abbildung 5.11 für zwei unterschiedliche Prüfpläne.

$$L_{N,n,C}(p) = 1 - p(x) \tag{5.5}$$

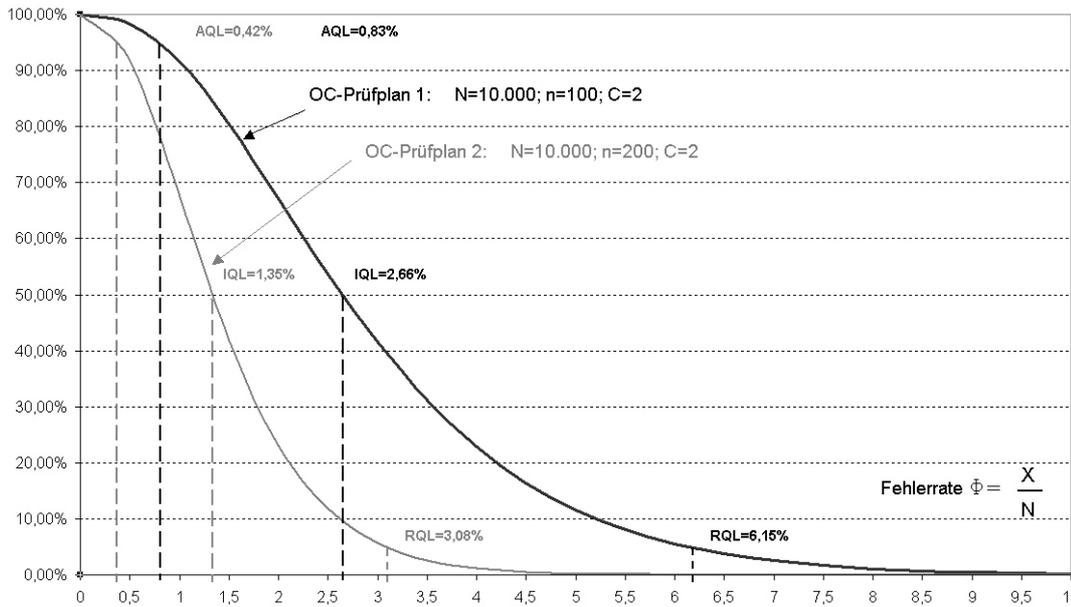


Abbildung 5.11. Operationscharakteristik

In der Abbildung 5.11 sind die drei Kennwerte der Operationscharakteristik nach DIN 55350 angegeben. Diese sind:

AQL: acceptable quality level, annehmbare Qualitätsgrenzlage: Damit wird diejenige (gute) Qualitätslage eines Loses bezeichnet, bei der die Annahmewahrscheinlichkeit für das Los einen zwischen Konsumenten und Produzenten zum Schutz des Produzenten vereinbarten hohen Wert hat. Das Produzentenrisiko ist das Zurückweisen eines Loses mit tolerierbarer Ausschussquote (Fehler 1.Art).

RQL: rejectable quality level, rückzuweisende Qualitätsgrenzlage: Dies ist die (schlechte) Qualitätslage eines Loses, bei der die Annahmewahrscheinlichkeit für das Los einen zwischen Konsumenten und Produzenten zum Schutz des Konsumenten vereinbarten niedrigen Wert hat (Fehler 2.Art).

IQL: indifferent quality level, indifferente Qualitätslage: Die indifferente Qualitätslage $p_{0,5}$ ist die Stelle, an der die Annahme- und Ablehnwahrscheinlichkeit eines Loses gerade 0,5 beträgt. Der 50%-Punkt $p_{0,5}$ als derjenige Punkt, für den die Wahrscheinlichkeit für die Annahme gleich der Wahrscheinlichkeit für die Ablehnung des Loses ist, zeigt an, welche Ausschussanteile p (nämlich $p < p_{50\%}$) mit größter Wahrscheinlichkeit eine Annahme und welcher Ausschussanteil p (nämlich $p > p_{50\%}$) mit größter Wahrscheinlichkeit eine Ablehnung des Loses bewirkt.

Zur Berechnung der Operationscharakteristik $L_{N,n,C}(\Phi_\alpha)$ einer Stichprobe für die Fehlerquote Φ_α % eignet sich zweckmäßigerweise die Produktformel:

$$L_{N,n,C}(\Phi_\alpha) = \sum_{i=0}^C \left(\prod_{k=0}^{i-1} \frac{n-k}{i-k} \cdot \prod_{k=0}^{i-1} \frac{(N \cdot \Phi_\alpha) - 1}{N - n + i - k} \cdot \prod_{k=0}^{n-i-1} \frac{N - (N \cdot \Phi_\alpha) - k}{N - k} \right) \tag{5.6}$$

Der hauptsächliche Vorteil dieser Schreibweise ist das Umgehen eines numerischen Problems der Berechnung der Fakultät für große Grundgesamtheiten. Damit lässt sich diese Formel leicht in ein Programmsystem zur automatischen Berechnung einführen.

Abbildung 5.11 zeigt die Operationscharakteristiken für eine einfache Stichprobenprüfung (2 Prüfpläne) mit einer Grundgesamtheit von $N = 10.000$, einer Annahmezahl $C = 2$ und den unterschiedlichen Stichprobenumfängen von $n = 100$ und $n = 200$. Der Wert der Annahmezahl C ergibt sich über die hypergeometrische Verteilung mit der Fehlerquote Φ_α .

Die im ersten Prüfplan berechnete rückzuweisende Qualitätsgrenzlage hat mit $RQL = 6,15\%$ einen relativ hohen Wert gegenüber dem des zweiten Prüfplans mit $RQL = 3,08\%$. Damit ist der erste Prüfplan für einen Abnehmer von Geodaten schlechter gegenüber dem zweiten Prüfplan mit einem doppelt so großen Stichprobenumfang. Je steiler die OC-Kurve eines Prüfplans ausfällt, desto günstiger ist dies für einen Fehler 2. Art. Es kommt dadurch zur Verringerung des so genannten Konsumentenrisikos.

Anhand dieses Vergleichs der beiden OC-Kurven kann gezeigt werden, welche Konsequenz aus einer festgelegten Fehlerrate resultiert. Die Senkung der rückzuweisenden Qualitätsgrenzlage kann über die Erhöhung des Stichprobenumfangs des Prüfplans erreicht werden. Ob nun dieser funktionale Zusammenhang, so wie hier zu vermuten wäre, linear verläuft, ist Gegenstand des folgenden Abschnitts.

5.5.1 Einfluss der Fehlerrate auf den Stichprobenumfang

Es stellt sich demzufolge die Frage nach dem notwendigen Umfang der Stichprobe, der sich bei einem gegebenen Signifikanzniveau aus einer maximal tolerierbaren Fehlerrate ergibt. Anders ausgedrückt bedeutet dies: Wie groß ist der Stichprobenumfang zu wählen, um bei der Qualitätssicherung von Geodatenbeständen eine vorgegebene Fehlerrate garantieren zu können. Das hier vorliegende Ziel ist es, den Stichprobenumfang n und die Annahmezahl C , also den Prüfplan (n, C) so zu bestimmen, dass die Operationscharakteristik (ungefähr) durch die zwei vorgegebenen Punkte $\Phi_{1-\alpha}, \Phi_\beta$ geht. Dazu gehört, dass man diese Annahmeregrenzen vorschreibt. Man sucht also einen möglichst kleinen Stichprobenumfang n und eine Annahmezahl C so, dass Lose mit z. B. 1 % Fehleranteil mit nicht weniger als 90 % Wahrscheinlichkeit angenommen werden und dass Lose mit 3 % Fehleranteil mit nicht mehr als 10 % Wahrscheinlichkeit angenommen werden. Mit der Losgröße N , sind damit die vier reellen Zahlen $\alpha, \beta, \Phi_{1-\alpha}, \Phi_\beta$ gegeben, die den folgenden Bedingungen genügen sollen:

$$1 > 1 - \alpha > 0,5 > \beta > 0 \quad 0 < \Phi_{1-\alpha} < \Phi_\beta < 1 \quad \Phi_\beta - \Phi_{1-\alpha} > \frac{1}{N}$$

Die Berechnung der Parameter des Stichprobenplans C und n kann über die inverse Funktion der χ^2 -Verteilung erfolgen [UHLMANN, 1982]. Für die χ^2 -Verteilung gilt für $x > 0$ die Dichtefunktion mit dem Freiheitsgrad f :

$$g_f(x) = \frac{e^{-\frac{x}{2}} x^{\frac{f}{2}-1}}{2^{\frac{f}{2}} \cdot \Gamma\left(\frac{f}{2}\right)} \quad (5.7)$$

Mit Substitution von $t = \frac{x}{2}$ ergibt sich die von t abhängige Formel der Dichteverteilung:

$$g_f(t) = \frac{t^{\frac{f}{2}-1} e^{-t}}{2 \cdot \Gamma\left(\frac{f}{2}\right)} \quad (5.8)$$

Unter Verwendung der Reihenentwicklung aus [FICHTENHOLZ, 1973] für den Zähler von (5.8):

$$t^{\frac{f}{2}-1} e^{-t} = \left(1 - \frac{t}{1!} + \frac{t^2}{2!} - \frac{t^3}{3!} + \dots\right) \cdot t^{\frac{f}{2}-1} = \sum_{i=0}^{\infty} \frac{(-1)^i \cdot t^{\left(\frac{f}{2}-1+i\right)}}{i!} \quad (5.9)$$

und der Reihenentwicklung von $n!$ mit:

$$n! \sim \sqrt{2\pi \cdot n} \left(\frac{n}{e}\right)^n \left[1 + \frac{1}{12n} + \frac{1}{288n^2} - \frac{139}{51840n^3} - \frac{571}{2488320n^4} + \dots\right] \quad (5.10)$$

sowie der sogenannten *Stirlingschen Reihe* der Gammafunktion $\Gamma(a)$ mit:

$$\begin{aligned} \ln \Gamma(a) \sim \ln \sqrt{2\pi} + \left(a - \frac{1}{2}\right) \ln a - a + \frac{B_1}{1 \cdot 2} \frac{1}{a} - \frac{B_2}{3 \cdot 4} \frac{1}{a^3} + \dots \\ + (-1)^{m-1} \frac{B_m}{(2m-1)2m} \frac{1}{a^{2m-1}} + \dots \end{aligned} \quad (5.11)$$

ergibt sich mit den Bernoullischen Zahlen $B_1 = \frac{1}{6}$, $B_2 = \frac{1}{30}$, $B_3 = \frac{1}{42}$, $B_4 = \frac{1}{30}$, $B_5 = \frac{5}{66}$, ... für die allgemein gilt:

$$B_n = \frac{2(2n)!}{(2\pi)^{2n}} \cdot \sum_{m=1}^{\infty} \frac{1}{m^{2n}} \quad (5.12)$$

eine programmtechnisch handhabbare Funktion der χ^2 -Verteilung für numerisch große Werte zu:

$$g_f(t) = \left(\sum_{i=0}^{\infty} \frac{(-1)^i t^{\left(\frac{f-1-i}{2}\right)}}{\sqrt{2\pi} \left(\frac{i}{e}\right)^i} \right) \cdot \frac{1}{e^{h\left(\frac{f}{2}\right)}} \quad (5.13)$$

und der darin enthaltenen substituierten *Stirlingschen Reihe* der Gammafunktion $h\left(\frac{f}{2}\right) = \ln \Gamma\left(\frac{f}{2}\right)$.

Der in [UHLMANN, 1982] dargestellte Beweis, den Freiheitsgrad f durch $2(C+1)$ ersetzen zu können sowie der Zusammenhang zwischen Fehlerrate Φ und der inversen χ^2 -Verteilung mit:

$$2n\Phi = \frac{I}{g_{2(C+1)}(t)} \quad (5.14)$$

führt auf die zu lösende Gleichung für den Stichprobenplan, in dem die beiden Bedingungen für einen Fehler 1. und 2. Art eingehalten werden müssen zu:

$$L_{N,n,C}(\Phi_{1-\alpha}) \geq 1 - \alpha \quad (5.15) \qquad L_{N,n,C}(\Phi_{\beta}) \leq \beta \quad (5.16)$$

Unter Berücksichtigung des Zusammenhangs zwischen der Operationscharakteristik und der inversen χ^2 -Verteilung (Beweis siehe [UHLMANN, 1982]) ergibt sich der Zusammenhang zwischen den Irrtumswahrscheinlichkeiten α, β , der Annahmezahl C und dem Stichprobenumfang n zu:

$$\frac{I}{g_{2(C+1)}(2\beta) \cdot 2\Phi_{\beta}} \leq n \leq \frac{I}{g_{2(C+1)}(2(1-\alpha)) \cdot 2\Phi_{1-\alpha}} \quad (5.17)$$

An dieser Formel ist zu erkennen, dass die maßgebliche Abhängigkeit durch das Verhältnis der Fehlerrate der Nullhypothese H_0 : $\Phi_{1-\alpha}$ und der Alternativhypothese H_A : Φ_{β} existiert.

Aus Abbildung 5.12 kann erstens abgelesen werden, dass eine Senkung der Fehlerrate der Grundgesamtheit einen exponentiell wachsenden Stichprobenumfang zur Folge hat und zweitens dieser Stichprobenumfang ansteigt, je größer das Verhältnis zwischen der Fehleranzahl M_A der Hypothese H_A und der Fehleranzahl M_0 der Hypothese H_0 wird.

Diese Kurven sind bei der Festlegung der einzuhaltenen Fehlergrenze und der maximal zu tolerierenden Fehleranzahl in einer Grundgesamtheit vor dem Hintergrund des Prüfaufwandes zu berücksichtigen. Aus diesen Kurven lässt sich ebenfalls ohne viel Mühe ablesen, dass Qualitätssicherungsmaßnahmen bis zu einer Fehlerrate von 0,5% in diesem Beispiel mit einem vertretbaren Prüfaufwand zu garantieren sind. Ab dieser Grenze sieht man allerdings das exponentielle Wachstum des Stichprobenumfangs.

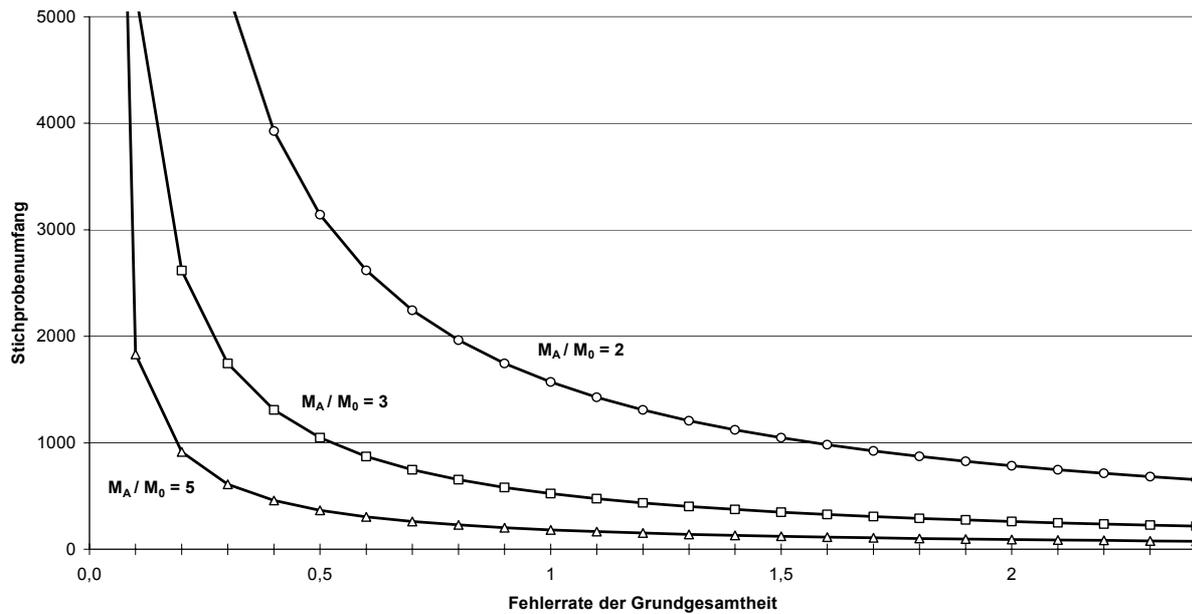


Abbildung 5.12: Stichprobenumfang in Abhängigkeit der Fehlerrate der Grundgesamtheit

Anders ausgedrückt bedeutet dies, dass zur Gewährleistung der fehlenden 0,5% auf die 100%ige Qualitätssicherungsforderung mit einem exponentiell steigenden Überprüfungsaufwand zu rechnen ist. Dieser bereits von [JOOS, 2000] aufgezeigte Zusammenhang wurde hier als Basis zur weiteren Begründung der Abhängigkeit der Kosten von den Qualitätsforderungen an den Zieldatenbestand noch einmal dargestellt. Der Unterschied zu JOOS ist die Adaption der hergeleiteten Berechnungsformeln zur Umsetzung in ein Programmsystem zur Aufwandsschätzung der Qualitätssicherung mittels Stichproben.

5.5.2 Kostenoptimale Stichprobenprüfpläne

Anders als bei der Produktion von Massengütern, bei denen ein fehlerhaftes Los zur Zurückweisung führt und es bei mehrfacher Loszurückweisung schlimmstenfalls zum Wechsel des Zulieferers kommt, müssen bei den Handelsbeziehungen mit Geodaten detaillierte vertragliche Vereinbarungen getroffen werden, die genau regeln, wie mit zurückgewiesenen Loses umzugehen ist. Oftmals existiert aufgrund der Monopolstellung von Geodatenanbieter und –abnehmer keine alternative Einkaufsmöglichkeit.

Übertragen auf die Situation, die bei einer Ablehnung eines Loses entsteht, bedeutet dies, dass die Zurückweisung eines Loses zwangsläufig eine Verbesserung durch den Geodatenanbieter zur Folge haben muss. Da bei Stichprobenprüfungen allgemein von einer Qualitätsaussage der Stichprobe auf die Grundgesamtheit geschlossen wird, bleiben existierende Restfehler in einem angenommenen Los unentdeckt. Man spricht vom sogenannten **Durchschlupf**. Er bezeichnet jenen Restfehleranteil, welcher im Mittel bei vielen Loses mit gleichem Fehleranteil Φ unentdeckt bleibt.

Neben der Operationscharakteristik wird auch der Durchschlupf Φ_u (u..unvermeidbar) zur Bewertung der Stichprobenanweisung herangezogen. Nachfolgendes Beispiel soll dies verständlicher machen:

Es werden 100 Lose mit einem jeweiligen Umfang von 1000 Geodatenelementen, welche alle einen Fehleranteil von $\Phi = 2\%$ besitzen, mit einem Stichprobenplan von $n=55$ und einer Annahmezahl $C=1$ untersucht. Mit Hilfe der inversen χ^2 -Verteilung ergibt sich die Anzahl der angenommenen Lose zu 70 und die der zurückgewiesenen Lose zu 30. Die 70 angenommenen Lose enthalten somit 70.000 Geodatenelemente, von denen im Rahmen der Prüfung $70 \cdot 55 = 3850$ Elemente geprüft und die fehlerhaften Elemente verbessert wurden. Es verbleiben $70.000 - 3.850 = 66.150$ ungeprüfte Elemente, darunter 2% fehlerhafte, d.h. 1323 Elemente. Die 30 zurückgewiesenen Lose werden 100% geprüft, so dass die fehlerhaften Geobjekte korrigiert wurden. Letztendlich sind von den $100 \cdot 1000 = 100.000$ Einheiten, die der Abnehmer erhält, im Mittel 1323 Einheiten fehlerhaft.

Der Durchschlupf ist demnach das Verhältnis der Anzahl der verbleibenden fehlerhaften Einheiten $\Phi_u = \frac{1.323}{100.000} \approx 1,3\%$ nach der Prüfung zur Anzahl der insgesamt gelieferten Einheiten. Interessant ist hierbei immer nur der Höchstwert des mittleren Durchschlupfs der sich mit:

$$\Phi_u = \underset{x=0,1,\dots,N}{Max} \frac{X_u}{N} L_{N,n,C} \left(\frac{X_u}{N} \right) \quad (5.18)$$

ergibt. Das Lösen dieser Gleichung führt auf eine Extremalaufgabe, welche die ausschließliche Abhängigkeit des Durchschlupfs Φ_u von der Annahmezahl C beweist.

Die seitherige Betrachtung der Prüfpläne ließ den Kostenaspekt weitestgehend außer Acht. Es wurden lediglich die Vorgabe zweier Punkte der OC bzw. der Steilheit und des Indifferenzpunktes berücksichtigt. Um Prüfpläne kostengünstig bzw. **kostenoptimal** zu gestalten, sind Annahmen über die entstehenden Kosten notwendig. Diese Annahmen betreffen diejenigen Kosten, die durch die Annahme eines Loses, dessen Ablehnung und bei der Kontrolle auftreten. Wie bereits in den vorangegangenen Abschnitten werden hier die folgende Abkürzungen verwendet:

- N...Umfang des Loses (Umfang der Grundgesamtheit)
- M...Anzahl der fehlerhaften Elemente des Loses (der Grundgesamtheit)
- $\Phi = \frac{M}{N}$...Ausschussanteil des Loses (Fehlerquote)
- n...Stichprobenumfang
- x...Anzahl der fehlerhaften Elemente des Loses
- C Annahmezahl
- (n,C)...Prüfplan

Kostenfaktor 1: Annahme des Loses ohne Kontrolle

Würde man das Los ohne Kontrolle annehmen, so seien die entstehenden Kosten $K^{AnoK}(\Phi)$ eine lineare Funktion, abhängig vom Ausschussanteil (Fehlerquote) Φ des Loses mit:

$$K^{AnoK}(\Phi) = a_1 + b_1 \Phi \quad (5.19)$$

Diese Funktion setzt sich, anders ausgedrückt, aus dem Gewinn, der angenommenen guten Elemente (N-M) und dem Verlust, welcher sich für M angenommene falsche Elemente ergibt zusammen. Führt man für den Gewinn bzw. für den Verlust einzelne Kostenfaktoren ein, die mit:

g_{korr}^{An} Gewinn durch ein angenommenes korrektes Datenelement und

v_{falsch}^{An} ...Verlust durch ein angenommenes falsches Datenelement

bezeichnet werden vgl. [UHLMANN, 1982], dann ergeben sich die entstehenden Kosten bei der Annahme des Loses ohne Kontrolle zu:

$$K^{AnoK}(\Phi) = (N - M)g_{korr}^{An} + Mv_{falsch}^{An} = \dots = Ng_{korr}^{An} + (v_{falsch}^{An} - g_{korr}^{An})N\Phi \quad (5.20)$$

Daraus zu sehen sind die Kostenfaktoren mit $a_1 = Ng_{korr}^{An}$ und $b_1 = (v_{falsch}^{An} - g_{korr}^{An})N$.

Kostenfaktor 2: Ablehnung des Loses ohne Kontrolle

Würde man demgegenüber ein Los ohne Kontrolle ablehnen, dann seien die entstehenden Kosten $K^{AboK}(\Phi)$ ebenfalls eine lineare Funktion, abhängig vom Ausschussanteil (Fehlerquote) Φ des Loses mit:

$$K^{AboK}(\Phi) = a_2 + b_2 \Phi \quad (5.21)$$

Damit entsteht der Gewinn aus den M abgelehnten falschen Elementen und der Verlust aus den $(N-M)$ abgelehnten richtigen Elementen. Nach Einführung der entsprechenden Kostenfaktoren:

$$\begin{aligned} g_{falsch}^{Ab} & \text{ Gewinn durch ein abgelehntes falsches Datenelement und} \\ v_{korr}^{Ab} & \dots \text{ Verlust durch ein abgelehntes richtiges Datenelement} \end{aligned}$$

ergibt sich daraus die Formel für die entstehenden Kosten bei der Ablehnung eines Loses ohne Kontrolle zu:

$$K^{Abok}(\Phi) = M g_{falsch}^{Ab} + (N - M) v_{korr}^{Ab} = \dots = N v_{korr}^{Ab} + (g_{falsch}^{Ab} - v_{korr}^{Ab}) N \Phi \quad (5.22)$$

Auch hieraus ergeben sich a_2 und b_2 zu: $a_2 = N v_{korr}^{Ab}$ und $b_2 = (g_{falsch}^{Ab} - v_{korr}^{Ab}) N$.

Die beiden Fragestellungen, nämlich ob der Verlust bei Annahme eines falschen Datenelements betragsmäßig (umgekehrtes Vorzeichen) gleich groß ist wie der Gewinn bei Ablehnung eines falschen Datenelements und ob der Gewinn bei Annahme eines richtigen Datenelements betragsmäßig gleich groß ist (umgekehrtes Vorzeichen) wie der Verlust bei Ablehnung eines richtigen Datenelements, können i.d.R. positiv beantwortet werden. Dies führt zu den beiden Thesen:

These 1: Der Verlust, ein falsches Datenelement zu akzeptieren, ist betragsmäßig gleich groß wie der Gewinn, dasselbe Datenelement zu verwerfen (als falsch zu identifizieren).

These 2: Der Verlust, ein richtiges Datenelement fälschlicherweise abzulehnen, ist betragsmäßig gleich dem Gewinn, dasselbe Datenelement richtigerweise anzunehmen.

Unter Annahme der Richtigkeit beider Thesen ergeben sich für die Formeln (5.20) und (5.22) folgende Vereinfachungen mit dem Zusammenhang zwischen den Kostenfaktoren:

$$v_{falsch}^{An} = -g_{falsch}^{Ab} \qquad v_{korr}^{Ab} = -g_{korr}^{An}$$

zu den Formel für die Kosten der Annahme bzw. Ablehnung des Loses ohne Kontrolle mit:

$$K^{Anok}(\Phi) = N g_{korr}^{An} + (-g_{falsch}^{Ab} - g_{korr}^{An}) N \Phi \quad (5.23)$$

$$K^{Abok}(\Phi) = -N g_{korr}^{An} + (g_{falsch}^{Ab} - g_{korr}^{An}) N \Phi \quad (5.24)$$

Betrachtet man für die Funktionen (5.23) und (5.24) die beiden extremalen Möglichkeiten, dass erstens das Los keine fehlerhaften Objekte enthält ($\Phi = 0$) und zweitens das Los ausschließlich aus fehlerhaften Objekten besteht ($\Phi = 1$), erhält man durch die Extremwertbetrachtung mit:

Kostenfunktion für $\Phi = 0$:

$$K^{Anok}(0) = N g_{korr}^{An} \qquad K^{Abok}(0) = -N g_{korr}^{An}$$

Kostenfunktion für $\Phi = 1$:

$$\begin{aligned} K^{Anok}(1) &= N g_{korr}^{An} + (-g_{falsch}^{Ab} - g_{korr}^{An}) N & K^{Abok}(1) &= -N g_{korr}^{An} + (g_{falsch}^{Ab} - g_{korr}^{An}) N \Phi \\ &= -N g_{korr}^{An} & &= N g_{falsch}^{Ab} \end{aligned}$$

. In allen praktisch relevanten Fällen wird der Gewinn bei einem Los, das nur aus fehlerfreien Datenelementen besteht, bei der Annahme größer sein als bei der Ablehnung:

$$K^{Anok}(0) > K^{Abok}(0) \text{ mit } N g_{korr}^{An} > -N g_{korr}^{An} \quad (5.25)$$

Andererseits wird der Gewinn bei einem Los, das nur aus fehlerbehafteten Datenelementen besteht, bei der Annahme des Loses kleiner sein als bei der Ablehnung.

$$K^{Anok}(1) < K^{Abok}(1) \text{ mit } -N g_{falsch}^{Ab} < N g_{falsch}^{Ab} \quad (5.26)$$

Daher existiert genau ein Ausschussanteil Φ_0 , für den der Gewinn bei Annahme des Loses genauso groß ist, wie bei der Ablehnung: $K^{Anok}(\Phi_0) = K^{Abok}(\Phi_0)$. Durch Einsetzen der entsprechenden Kostenfaktoren a_1, b_1, a_2, b_2 folgt der Zusammenhang:

$$\Phi_0 = \frac{a_2 - a_1}{b_1 - b_2} = \frac{(-Ng_{korr}^{An}) - Ng_{korr}^{An}}{(-g_{falsch}^{Ab} - g_{korr}^{An})N - (g_{falsch}^{Ab} + g_{korr}^{An})N} = \frac{-2Ng_{korr}^{An}}{-2Ng_{falsch}^{Ab} - 2Ng_{korr}^{An}} = \frac{g_{korr}^{An}}{g_{falsch}^{Ab} + g_{korr}^{An}} \quad (5.27)$$

Damit kann festgehalten werden, dass der Ausschussanteil Φ_0 nur vom Gewinn eines angenommenen, korrekten Datenelement und dem Gewinn durch ein abgelehntes falsches Datenelement (unter Annahme der These 1 und 2) abhängt.

Würde man den Ausschussanteil Φ des Loses kennen, so würde man – um einen möglichst kleinen Verlust zu erleiden, also einen möglichst großen Gewinn zu erzielen – das Los annehmen, wenn $\Phi < \Phi_0$ ist und das Los ablehnen, wenn $\Phi > \Phi_0$ ist. Selbst bei Kenntnis von Φ blieben also unvermeidbare Kosten für:

$$K_u(\Phi) = \begin{cases} K^{Anok}(\Phi) & , \text{ falls } \Phi \leq \Phi_0 \\ K^{Abok}(\Phi) & , \text{ falls } \Phi \geq \Phi_0 \end{cases} \quad (5.28)$$

Da aber der Ausschussanteil Φ in der Regel unbekannt ist, wird man eine Zufallsstichprobe vom Umfang n ziehen. Mit einer angenommenen, linearen Funktion der Prüfkosten $d_1n + d_2$ ergeben sich die zu erwartenden unvermeidbaren Gesamtkosten $K_g(\Phi)$ für den Fall, dass die Prüfung $x \leq C$ ausfällt zu:

$$K_g(\Phi) = \left\{ Ng_{korr}^{An} + (-g_{falsch}^{Ab} - g_{korr}^{An})N\Phi \right\} \cdot L_{n,C}(\Phi) + \left\{ Ng_{korr}^{An} + (g_{falsch}^{Ab} + g_{korr}^{An})N\Phi \right\} \cdot (1 - L_{n,C}(\Phi)) + d_1n + d_2 \quad (5.29)$$

Für den Fall, dass die Thesen 1 und 2 nicht zutreffen, also dass der Verlust, ein falsches Datenelement zu akzeptieren, betragsmäßig nicht dem Gewinn dasselbe Datenelement zu verwerfen, entspricht, und der Verlust ein richtiges Datenelement fälschlicherweise abzulehnen, betragsmäßig nicht dem Gewinn, dasselbe Datenelement richtigerweise anzunehmen entspricht, gilt die folgende allgemeine Kostenfunktion $K_g(\Phi)$ zur Berechnung **kostenoptimaler Stichprobenprüfpläne**:

$$K_g(\Phi) = \left\{ Ng_{korr}^{An} + (v_{falsch}^{An} - g_{korr}^{An})N\Phi \right\} \cdot L_{n,C}(\Phi) + \left\{ Nv_{korr}^{Ab} + (g_{falsch}^{Ab} - v_{korr}^{Ab})N\Phi \right\} \cdot (1 - L_{n,C}(\Phi)) + d_1n + d_2 \quad (5.30)$$

Folgende Abbildung 5.13 zeigt den Kostenverlauf der Stichprobenprüfung in Abhängigkeit der Fehlerrate Φ für drei unterschiedliche Varianten der Gewinne und Verluste bei angenommenen bzw. abgelehnten richtigen oder falschen Elementen der Stichprobe.

Variante 1:

$$g_{korr}^{An} = 100 \text{ WE}; \quad v_{falsch}^{An} = 140 \text{ WE} \\ g_{falsch}^{Ab} = 140 \text{ WE}; \quad v_{korr}^{Ab} = 100 \text{ WE}$$

$$q = \frac{g_{korr}^{An}}{v_{falsch}^{An}} = \frac{v_{korr}^{Ab}}{g_{falsch}^{Ab}} = 0,71$$

Bei dieser Variante wird davon ausgegangen, dass der Gewinn, ein korrektes Element anzunehmen, geringer ist als der entstehende Verlust bei der Annahme eines fehlerhaften Elements sowie der resultierende Gewinn bei Ablehnung eines falschen Elements größer ist als der Verlust, ein korrektes Element abzulehnen. Die Kostenfaktoren stehen zueinander im Verhältnis mit dem Quotienten $q=0,71$.

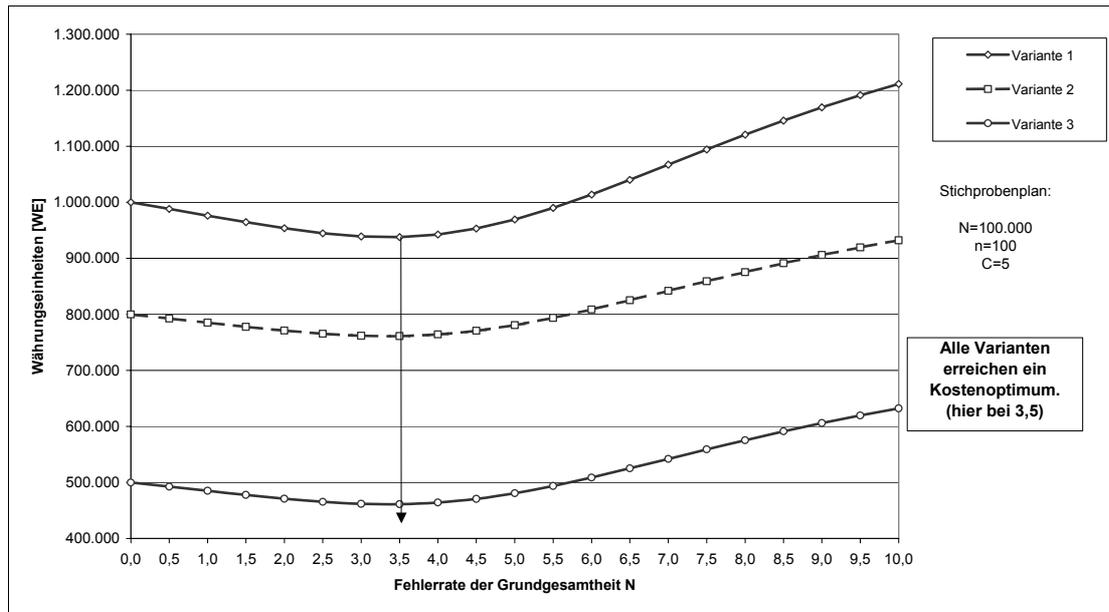


Abbildung 5.13: Verlauf der Kostenfunktion $K_g(\Phi)$ für kostenoptimale Stichprobenprüfpläne

Variante 2:

$$g_{korr}^{An} = 80 \text{ WE}; \quad v_{falsch}^{An} = 70 \text{ WE}$$

$$g_{falsch}^{Ab} = 70 \text{ WE}; \quad v_{korr}^{Ab} = 80 \text{ WE}$$

$$q = \frac{g_{korr}^{An}}{v_{falsch}^{An}} = \frac{v_{korr}^{Ab}}{g_{falsch}^{Ab}} = 1,14$$

Bei dieser Variante wird von einem Gewinn für ein angenommenes korrektes Datenelement ausgegangen der geringfügig größer ist, als der entstehende Verlust ein fehlerhaftes Element fälschlicherweise anzunehmen.

Variante 3:

$$g_{korr}^{An} = 50 \text{ WE}; \quad v_{falsch}^{An} = 100 \text{ WE}$$

$$g_{falsch}^{Ab} = 100 \text{ WE}; \quad v_{korr}^{Ab} = 50 \text{ WE}$$

$$q = \frac{g_{korr}^{An}}{v_{falsch}^{An}} = \frac{v_{korr}^{Ab}}{g_{falsch}^{Ab}} = 0,5$$

Diese Variante geht von einem Verhältnis zwischen den Kostenfaktoren von 0,5 aus. Damit wird der Verlust, ein fehlerhaftes Datenelement fälschlicherweise anzunehmen doppelt so hoch bewertet wie der Gewinn, ein richtiges Datenelement anzunehmen.

Zusammenfassend gilt für die Varianten: Alle Varianten sind bei einer Fehlerrate in diesem Beispiel von 3,5% kostenoptimal. Ihr Verlauf und ihre Steigungswerte sind dabei über die Verhältnisse q definiert. Die Kostensteigerung für zunehmende Fehlerraten ist maßgeblich von der Wahrscheinlichkeit abhängig, einen höheren Verlust durch fehlerhafte Elemente in Kauf zu nehmen.

Mit Hilfe dieser Herleitungen steht nun für die kostenmäßige Betrachtung eine Formel zu Verfügung, welche die Prüfkosten, den damit verbundenen Aufwand, die Fehlerquote sowie die einzelnen Kostenfaktoren berücksichtigt, die bei Annahme oder Ablehnung eines richtigen oder falschen Datenelements auftreten können. Damit kann entschieden werden, **welcher Prüfplan sich** unter gegebenen Kostenfaktoren **kostenoptimal verhält**. Dies ist eine Grundlage im Rahmen der Entscheidung, ob die Qualitätsforderungen an einen Datenbestand mittels Stichprobenprüfung oder Vollprüfung stattfinden kann.

5.5.3 Stichprobenprüfkosten unter Vollprüfungsrisiko

In diesem Abschnitt soll die Frage beantwortet werden, wie sich die resultierenden Kosten einer Stichprobenprüfung unter dem Einfluss des Risikos einer Vollprüfung verhalten. Durch die Operationscharakteristik ist festgelegt, wie sich die Wahrscheinlichkeiten der Losannahme in Bezug auf die Fehlerrate verhalten. Bei steigender Fehlerrate sinkt entsprechend die Annahmewahrscheinlichkeit. Im umgekehrten Verhältnis dazu steigt das Risiko einer Vollprüfung des Datenbestandes. Aus diesem Grund kann die folgende Annahme getroffen werden:

Die **Kosten des Risikos der Vollprüfung** K_{Rvpr} beschreiben diejenigen Kosten, die entstehen, wenn ein Prüflös zurückgewiesen wird und deshalb einer Vollprüfung (100%-Prüfung) zu unterziehen ist. Da die Wahrscheinlichkeit, das Prüflös zurückzuweisen in Abhängigkeit mit der ansteigenden Fehlerrate zunimmt, so kann die Komponente des Vollprüfungsrisikos mit einem Kostenparameter $k_{prüf}^{100\%}$ (Prüfungskosten - PK) folgendermaßen angegeben werden:

$$K_{Rvpr} = \frac{N^G}{N^L} \cdot k_{prüf}^{100\%} \cdot (N^L - n) \cdot (1 - L_{n,C}(\Phi)) \tag{5.31}$$

Der Kostenparameter $k_{prüf}^{100\%}$ beschreibt die Einzelkosten pro Element, die anfallen, um ein Prüflös vom Umfang N^L vollständig zu überprüfen. In folgender Abbildung 5.14 sind die Kurven ($K = K_{tPR} + K_{Rvpr}$) der entstehenden Kosten, trotz Datenprüfung, mit den Kosten des Risikos der Vollprüfung in Abhängigkeit des Kostenparameters $k_{prüf}^{100\%}$ und dem mittleren Verlust durch ein fehlerhaftes Element v_{FEI} (Fehlerkosten - FK) dargestellt.

Die Kostenverläufe lassen sich in zwei Scharen unterteilen. Während bei der ersten Schar (schwarze Kurven) der mittlere Verlust, der einer Anwendung durch ein fehlerhaftes Element entstünde, konstant 200 WE beträgt, variiert das Kostenrisiko der 100 %-igen Losprüfung zwischen 2 und 5 WE pro Element. Bei diesen Verläufen bestimmt das Kostenrisiko vollständiger Losprüfung maßgeblich den Gesamtkostenverlauf bei steigender Fehlerrate. Bei der zweiten Schar (graue Kurven) wurde das Kostenrisiko der Vollprüfung auf dem konstanten Wert von 5 WE pro Element belassen und der mittlere Verlust zwischen 100 und 250 WE variiert. Deutlich zeigt sich, dass der mittlere Verlust für den Anfangsverlauf der Kostenfunktion verantwortlich ist, und dass die Kurvenverläufe der zweiten Schar bei steigender Fehlerraten ineinander übergehen. Sie erreichen alle denselben Wert für das Risiko einer 100 %-igen Prüfung. Sämtliche Verläufe zeigen sich in Ihrem Trend trotzdem abnehmend, da mit steigender Fehlerrate zwar das Vollprüfungsrisiko ebenfalls steigt, aber gleichzeitig damit verbunden ist, dass keine fehlerhaften Objekte mehr im Datenbestand existieren, die bei der Anwendung Kosten verursachen können.

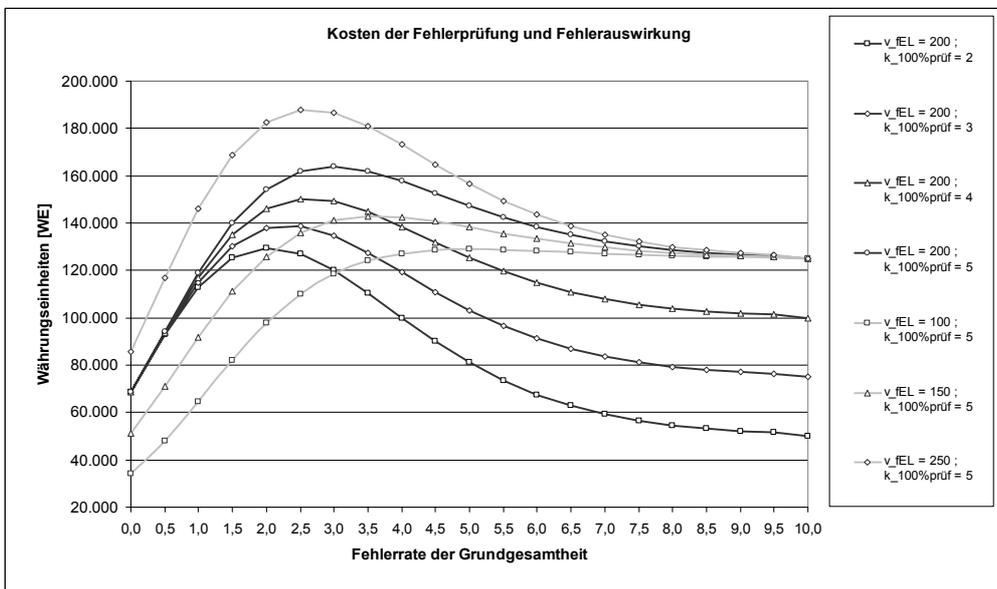


Abbildung 5.14: Scharen möglicher Kostenverläufe

Die gezeigten Kostenverläufe lassen den Schluss zu, dass die größten Kosten bei Fehlerraten zwischen 1,5 und 4,0% zu erwarten sind und man sich berechtigt fragen müsste weshalb denn eigentlich eine Stichprobenprüfung Sinn macht, wenn die Kurvenverläufe sowieso bei steigender Fehlerraten geringere Kosten aufweisen (trotz Risiko der 100%-igen Prüfung). Diese scheinbar naheliegende Schlussfolgerung wird allerdings widerlegt, wenn man bedenkt, dass der Kostenverlauf ohne Fehlerprüfung eine lineare Funktion in Abhängigkeit der Fehlerrate wäre und es somit durch Fehlerprüfung zur Amortisation der Kosten kommt (vgl. Abbildung 5.15).

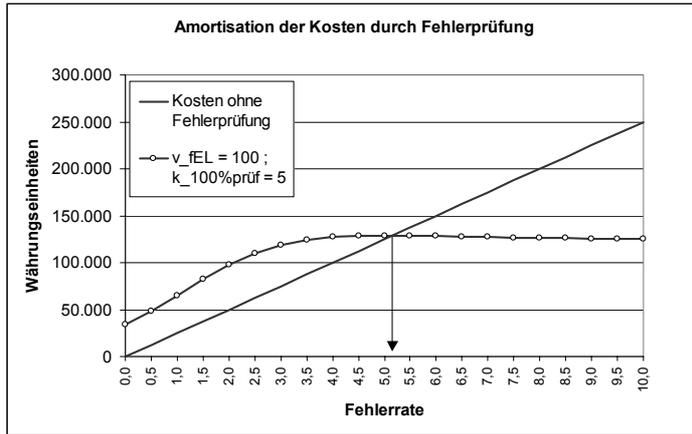


Abbildung 5.15: Amortisation der Kosten durch Fehlerprüfung

Aus den in Abbildung 5.14 dargestellten Kostenverläufe lassen sich aber andere, wesentlich wichtigere Schlussfolgerungen ziehen, wenn man das Verhältnis q zwischen dem mittleren Verlust und dem Kostenfaktor, der bei einer 100%-Prüfung entsteht, betrachtet. In folgender Abbildung 5.16 sind die Kostenverläufe für die angegebenen Verhältnisse q dargestellt.

$$q = \frac{k_{\text{prüf}}^{100\%}}{v_{\text{fEL}}} = \frac{\text{Prüfkosten}}{\text{Fehlerkosten}}$$

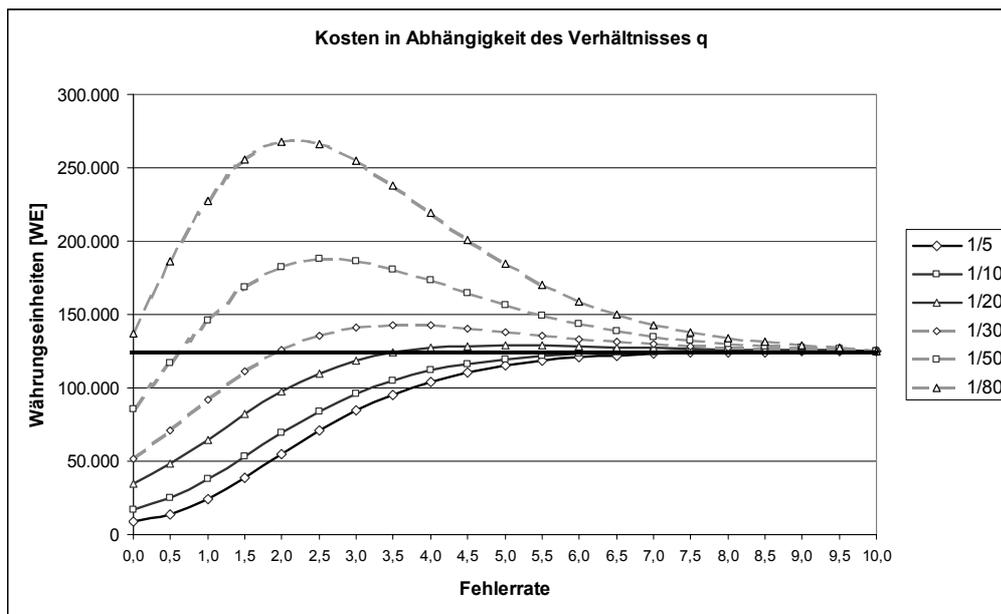


Abbildung 5.16: Kostenverlauf der Prüfkosten zu Fehlerkosten

Es zeigt sich, dass eine Stichprobenprüfung für die Verhältnisse $q=1/5$; $1/10$ und $1/20$ zum Erreichen einer möglichst geringen Fehlerrate günstiger ist als eine 100 %-Prüfung (schwarze Kurven). Werden die **Fehlerkosten**, die durch ein fehlerhaftes Datenelement ausgelöst werden, **größer als das 20ig-fache der Prüfungskosten** pro Element, so können die resultierenden **Kosten der Stichprobenprüfung die Kosten der Vollprüfung übersteigen**.

Eine **Vollprüfung** ist hingegen **einer Stichprobenprüfung** in diesem Beispiel **für die 80ig-fachen Kosten** in jedem Fall **vorzuziehen**, da hier selbst bei einer Fehlerrate von 0 % der Kosteneinfluss der unvermeidbar fehlerhaften Datenelemente trotz Stichprobenprüfung größer wird als die Kosten der Vollprüfung.

Anhang dieses Verhältnisses kann unter Kenntnis der Kostenparameter **angegeben werden**, ob zum Erreichen einer geforderten Fehlerrate **eine Stichprobenprüfung** genügt oder ob statt dessen eine **vollständige Datenprüfung** durchzuführen ist. Es zeigt sich, dass eine Stichprobenprüfung nicht immer einen Kostenvorteil gegenüber der Vollprüfung bringt.

5.6 Die Kostenfunktionen der Qualitätssicherungsmaßnahmen

In diesem Abschnitt sind die vollständigen Kostenfunktionen der Qualitätssicherungsmaßnahmen (Prüfung, Fehlerkorrektur) dargestellt. Ihre Zusammenstellung (Struktur) basiert dabei auf den in Kapitel 4 Abschnitt 4.5.4 allgemein formulierten Kostenstufen der Aufbaustrategien von Geodatenbeständen. Die Kostenfunktion der Qualitätssicherungsmaßnahmen besteht aus den beiden Hauptkomponenten der Fehlerprüfung K_{FP} und der Fehlerkorrektur K_{FK} :

$$K_{QS} = K_{FP} + K_{FK} \quad (5.32)$$

Diese nun folgenden Funktionsdefinitionen verwenden eine Vielzahl an Kostenfaktoren, welche deshalb in folgender Tabelle 5-5 übersichtlich dargestellt und die für die nachfolgende Modellrechnung, benötigten Werte festgelegt sind.

Tabelle 5-5: Definition und Festlegung der Kostenfaktoren

Kostenfaktoren	Beschreibung	Festgelegte Werte in Währungseinheiten (WE)
k_{fix}	Fixe Kosten des Dateneinkaufs.	abhängig von Objektanzahl
k_{var}	Variable Kosten des Dateneinkaufs (Lizenzgebühren)	1.000 WE
pr_{fix}	Fixe Prüfkosten (Rechner, Personal, etc.)	15.000 WE
pr_{var}	Variable Prüfkosten pro Element (Stichprobe)	2-5 WE
h_{fix}	Fixe Homogenisierungskosten (zusätzlich zu pr_{fix})	5.000 WE
h_{var}	Variable Homogenisierungskosten pro Element	2-5 WE
fk_{fix}	Fixe Kosten der Fehlerkorrektur (zusätzlich zu pr_{fix} und h_{fix})	5.000 WE
fk_{var}	Variable Kosten der Fehlerkorrektur pro Element	2-5 WE
v_{fel}	Verlust der durch ein fehlerhaftes Element entsteht.	25-400 WE
$k_{prüf}^{100\%}$	Kostenfaktor zur 100%igen Prüfung eines Elements	2-5 WE

Ebenfalls wegen des Aspektes der einheitlichen Betrachtung und Bewertung wird zunächst der Stichprobenprüfplan mit $N=10.000$; $n=100$ und $AC=2$ für folgende Modellrechnungen festgehalten. Die daraus resultierenden unvermeidbar fehlerhaften Elemente M_u ergeben sich zu $\Phi_u = 1,371\%$ ($M_u=N \Phi_u$).

5.6.1 Kostenfunktion der Fehlerprüfung K_{FP}

Die Kosten, die durch die Maßnahmen der Fehlerprüfung entstehen, bestehen ihrerseits aus den Einzelkomponenten:

- der Prüfkosten, d.h. der Aufwand zur Durchführung der Stichprobenprüfung $K_{prüf}$ und
- den Kosten, die trotz Datenprüfung entstehen (fehlerhafte Objekte bei Losannahme, unvermeidbare Fehler aufgrund des Durchschlupfs) K_{tPr} .

Die **Prüfkosten** lassen sich als lineare Funktion auffassen, die eine Kostensteigerung entsprechend des Stichprobenumfangs n erfahren. Sie ergeben sich zu:

$$K_{prüf} = pr_{fix} + n \cdot pr_{var} \quad (5.33)$$

worin die konstanten (fixen) Prüfkosten pr_{fix} und die variablen Prüfkosten pr_{var} pro Element enthalten sind.

Die **Kosten trotz Datenprüfung** K_{tPr} entstehen erstens durch die unvermeidbaren Fehler (Durchschlupf), die sich aus dem Design des Stichprobenprüfplans ergeben und zweitens aus dem Zusammenhang zwischen sinkender Fehlerrate und damit verbundener Wahrscheinlichkeitssteigerung, das Prüflös und damit entsprechend fehlerhafte Elemente zu akzeptieren. Die Kosten lassen sich durch einen Parameter v_{fel} angeben, der für den mittleren Verlust steht, der durch

ein fehlerhaftes Element im Datenbestand ausgelöst wird. Dieser mittlere Verlust ist von der Art der Anwendung, die mit diesen Geodaten arbeitet abhängig. Damit kann die Komponente der Kosten, die trotz Datenprüfung entstehen folgendermaßen angegeben werden:

$$K_{tPr} = \frac{N^G}{N^L} \cdot \left\{ v_{fEl} \cdot \left(N^L \Phi \cdot L_{n,C}(\Phi) + N^L \Phi_u - N^L \Phi_u \cdot L_{n,C}(\Phi) \right) \right\} \quad (5.34)$$

$\frac{N^G}{N^L}$ Da bei großen Grundgesamtheiten N^G , deren Qualität ermittelt werden soll, auf das Losprüfverfahren zurückgegriffen wird, dient dieser Quotient zur Übertragung des Resultats von der Losprüfung N^L auf die Grundgesamtheit.

$N^L \Phi \cdot L_{n,C}(\Phi)$ Die Wahrscheinlichkeit, das Prüflös zu akzeptieren steigt mit sinkender Fehlerrate entsprechend der OC-Kurve. Durch ein akzeptiertes Los, akzeptiert man allerdings auch entsprechende fehlerhafte Elemente $M(N^L \Phi)$.

$N^L \Phi_u$ Da das Prüflös über einen Stichprobenplan geprüft wird und dieser unweigerlich einen Durchschluß (unvermeidbarer Fehler) besitzt, wird dieser hier als konstanter Faktor eingeführt.

$N^L \Phi_u \cdot L_{n,C}(\Phi)$ Dieser konstante Faktor $N^L \Phi_u$ mindert sich entsprechend der Operationscharakteristik, da für steigende Fehlerraten die Wahrscheinlichkeit abnimmt, das Prüflös zu akzeptieren und so den durch den Stichprobenplan auftretenden unvermeidbaren Fehleranteil nicht mit einfließen lässt.

Damit setzen sich die Kosten der Fehlerprüfung aus den Prüfkosten der Stichprobenprüfung und den resultierenden Kosten der fehlerhaften Elemente trotz Stichprobenprüfung zusammen. Sie ergeben sich mit:

$$K_{FP} = K_{prüf} + K_{tPR} \quad (5.35)$$

5.6.2 Kostenfunktion der Fehlerkorrektur

Zusätzlich zu den Kosten der Fehlerprüfung sind auch die Kosten der Fehlerkorrektur im Rahmen der Qualitätssicherungskosten zu berücksichtigen. Diese Fehlerkorrekturkosten sind umso größer, je größer die Wahrscheinlichkeit ist, das Prüflös zurückzuweisen. Sie verhalten sich folglich entsprechend $(1 - L_{n,C}(\Phi))$. Unter Verwendung von Kostenparametern für die fixen Fehlerkorrekturkosten fk_{fix} und die variablen Fehlerkorrekturkosten fk_{var} ergibt sich der formelmäßige Zusammenhang der Fehlerkorrekturkosten zu:

$$K_{FK} = \frac{N^G}{N^L} \left\{ \left(fk_{fix} + fk_{var} \left(N^L \Phi + N^L \Phi_u \right) \right) \cdot \left(1 - L_{n,C}(\Phi) \right) \right\} \quad (5.36)$$

Die Tatsache, dass bei einer Fehlerkorrektur immer erst eine Fehlerprüfung vorangehen muss, ist kostenmäßig über das Risiko einer Vollprüfung berücksichtigt und wird deshalb in dieser Formel nicht mit eingeführt. Abbildung 5.17 zeigt den Kostenverlauf der Fehlerkorrektur in Abhängigkeit der Fehlerrate.

Der Kostenverlauf der Fehlerkorrektur verhält sich entsprechend der inversen OC-Kurve. Der Funktionsverlauf wird für große Fehlerraten (hier ab ca. 7 %) linear, da hier die Wahrscheinlichkeit das Prüflös zu akzeptieren gegen 0 strebt und sich die Korrekturkosten fortan proportional zur Fehlerrate verhalten. Hieraus kann abgelesen werden, dass die Fehlerkorrekturkosten bei sinkender Fehlerrate abnehmen und konsequenterweise bei einer Fehlerrate von 0 % zu Null werden. Bei einem angenommenen, vollständig korrekten Datenbestand existieren logischerweise auch keine Korrekturkosten.

Aus dieser Abbildung kann darüber hinaus auch die Wirkung der Stichprobenprüfung auf die Kosten abgelesen werden. Da man berechtigt annehmen darf, dass sich die Fehlerkorrekturkosten ohne Stichprobenprüfung (Vollprüfung) aus einem konstanten Anteil und einem von der Fehlerrate abhängigen variablen Anteil zusammensetzen und sich linear verhalten würden, erkennt man den deutlichen Kostenvorteil durch Stichprobenprüfung im Bereich von 0 % - 7 % der Fehlerrate. Es zeigt sich der charakteristische Verlauf der inversen Operationscharakteristik (Gütefunktion).

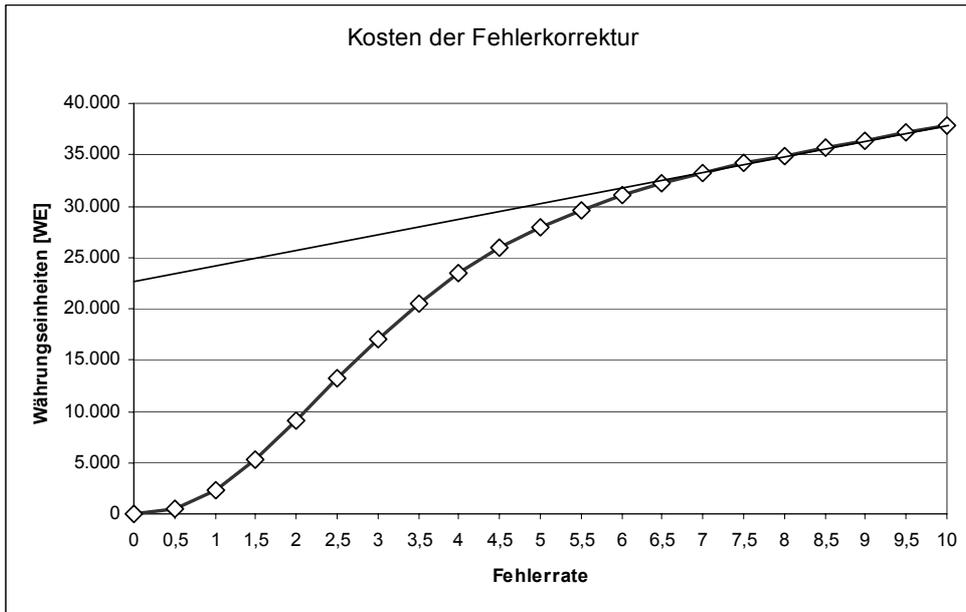


Abbildung 5.17: Kosten der Fehlerkorrektur in Abhängigkeit der Fehlerrate

5.6.3 Kostenfunktionen der Aufbaumöglichkeiten

Die Grundlage zur Bildung der Gesamtkostenfunktionen bezüglich der einzelnen Aufbaumöglichkeiten sind die bereits in Abschnitt 5.5.3 dargestellten Kostenfunktionen. Sie werden in diesem Abschnitt auf die Aufbaumöglichkeiten übertragen.

Tabelle 5-6: Parameter der Gesamtkostenfunktionen

	Kostenrelevante Parameter	Beschreibung
Kauf Fremd- daten	N^{KF}	Anzahl der Datenelemente durch den Kauf von Fremdbestände
	n^{KF}	Stichprobenumfang der Datenelemente aus Fremdbeständen
	Φ^{KF}	Fehlerquote des Fremddatenbestandes
	M^{KF}	Fehlerhafte Elemente des eingekauften Datenbestandes ($N^{KF} \Phi^{KF}$)
Eigene Be- stände	N^{eB}	Anzahl der Datenelemente durch eigene Datenbestände
	n^{eB}	Stichprobenumfang der Datenelemente aus eigenen Datenbeständen
	Φ^{eB}	Fehlerquote des eigenen Datenbestands
	M^{eB}	Fehlerhafte Elemente des eigenen Datenbestands ($N^{eB} \Phi^{eB}$)
Homo- genisierung	N^{ih}	Anzahl der Datenelemente die homogenisiert werden müssen
	n^{ih}	Stichprobenumfang der Datenelemente zur Prüfung der Homogenisierung
	Φ^{ih}	Fehlerquote der Homogenisierung
	M^{ih}	Fehlerhafte Elemente der Homogenisierung ($N^{ih} \Phi^{ih}$)
Eigene Erfas- sung	N^{eE}	Anzahl der Datenelemente die selbst zu erfassen sind
	n^{eE}	Stichprobenumfang der Datenelemente zur Prüfung eigener Erfassung
	Φ^{eE}	Fehlerquote der eigenen Erfassung
	M^{eE}	Fehlerhafte Elemente bei der eigenen Erfassung ($N^{eE} \Phi^{eE}$)

5.6.3.1 Gesamtkostenfunktion des Neuaufbaus 1

Die Gesamtkostenfunktion der Neuaufbastrategie 1, dem reinen Dateneinkauf setzt sich aus den Einkaufskosten K^{KF} , sowie den Kosten der Qualitätssicherungsmaßnahmen wie folgt zusammen:

$K_{N1}^g =$	Einkauf Fremdbestand	K^{KF}	$= k_{fix} + d \cdot k_{var}$, Parameterwerte lt. Modellrechnung
		$+ K_{QS}^{KF}$	$K_{prüf}^{KF} = pr_{fix} + n \cdot pr_{var}$
		K_{tPr}^{KF}	$= \frac{N^{KF}}{N^L} \left\{ v_{fEl} (N^L \Phi \cdot L_{n,C}(\Phi) + N^L \Phi_u - N^L \Phi_u \cdot L_{n,C}(\Phi)) \right\}$
		K_{FK}^{KF}	$= \frac{N^{KF}}{N^L} \left\{ (fk_{fix} + fk_{var} (N^L \Phi + N^L \Phi_u)) \cdot (1 - L_{n,C}(\Phi)) \right\}$

Die Einkaufskosten K^{KF} setzen sich aus dem fixen Anteil (Einmalzahlung) k_{fix} und dem variablen Anteil (Lizenzgebühren) k_{var} zusammen. Der Parameter d steht für die Zahlungshäufigkeit der Lizenzgebühren in einem Jahr. Derselbe lineare Zusammenhang kann für die Kosten der Datenprüfung pr_{fix} , pr_{var} angenommen werden, die im Kostenfaktor der Qualitätssicherung Berücksichtigung finden. Für sämtliche nachfolgenden Kostenfaktoren des Dateneinkaufs bzw. den der eigenen Erfassungsleistungen werden im Rahmen dieser Modellrechnung Stückkosten pro Element angesetzt. Die daraus resultierenden Kostenverläufe der Aufbaustrategien lassen sich dann, entsprechend den Verhältniss $\frac{K^{KF}}{K^{eE}}$ bewerten. Die drei Komponenten der Prüfung, der entstehenden Kosten durch unvermeidbare Fehler trotz kostenoptimaler Stichprobenpläne und den Kosten der Fehlerkorrektur sind entsprechend Formel (5.44) zum Faktor K_{QS}^{KF} zusammengefasst.

5.6.3.2 Gesamtkostenfunktion des Neuaufbaus 2

Wird ein Datenbestand durch den Zukauf von Fremdbeständen und eigener Erfassungsleistung aufgebaut, so folgt daraus die Kostenformel:

$K_{N2}^g =$	Einkauf Fremdbestände	K^{KF}	$= k_{fix} + d \cdot k_{var}$, Parameterwerte lt. Modellrechnung
		$+ K_{QS}^{KF}$	$K_{prüf}^{KF} = pr_{fix} + n \cdot pr_{var}$
		K_{tPr}^{KF}	$= \frac{N^{KF}}{N^L} \left\{ v_{fEl} (N^L \Phi \cdot L_{n,C}(\Phi) + N^L \Phi_u - N^L \Phi_u \cdot L_{n,C}(\Phi)) \right\}$
		K_{FK}^{KF}	$= \frac{N^{KF}}{N^L} \left\{ (fk_{fix} + fk_{var} (N^L \Phi + N^L \Phi_u)) \cdot (1 - L_{n,C}(\Phi)) \right\}$
	Eigene Erfassung	$+ K^{eE}$	Parameterwerte siehe Modellrechnung
		$+ K_{QS}^{eE}$	$K_{prüf}^{eE} = pr_{fix} + n \cdot pr_{var}$
		K_{tPr}^{eE}	$= \frac{N^{eE}}{N^L} \left\{ v_{fEl} (N^L \Phi \cdot L_{n,C}(\Phi) + N^L \Phi_u - N^L \Phi_u \cdot L_{n,C}(\Phi)) \right\}$
		K_{FK}^{eE}	$= \frac{N^{eE}}{N^L} \left\{ (fk_{fix} + fk_{var} (N^L \Phi + N^L \Phi_u)) \cdot (1 - L_{n,C}(\Phi)) \right\}$

5.6.3.3 Gesamtkostenfunktion des Neuaufbaus 3

Für den Fall des vollständigen Neuaufbaus durch eigene Erfassungsleistung kann von folgender Gesamtformel ausgegangen werden:

$K_{N3}^g =$	Vollständiger Neuaufbau	K^{eE}	Parameterwerte siehe Modellrechnung
		$+ K_{QS}^{eE}$	$K_{prüf}^{eE} = pr_{fix} + n \cdot pr_{var}$
			$K_{tPr}^{eE} = \frac{N^{eE}}{N^L} \left\{ v_{jEl} \left(N^L \Phi \cdot L_{n,C}(\Phi) + N^L \Phi_u - N^L \Phi_u \cdot L_{n,C}(\Phi) \right) \right\}$
			$K_{FK}^{eE} = \frac{N^{eE}}{N^L} \left\{ (fk_{fix} + fk_{var} (N^L \Phi + N^L \Phi_u)) \cdot (1 - L_{n,C}(\Phi)) \right\}$

Auch bei dieser Strategie N3 berechnen sich die Kosten zur Qualitätssicherung nach demselben Schema wie bei N1 und N2, da auch bei eigener Erfassung mit fehlerhaft erfassten Elementen gerechnet werden muss. Der Prüfaufwand ist unabhängig davon, ob es sich beim Prüfungsgegenstand um fremde oder um eigene Daten handelt. Es kann lediglich davon ausgegangen werden, dass der Aufwand der Prüfungsvorbereitung bei Fremdbeständen größer ist als bei der Prüfung eigener oder eigenständig erfasster Daten. Da dieser Aufwand nur individuell ermittelt werden kann, werden die Fixkosten der Datenprüfung pr_{fix} für die Modellrechnungen aller Aufbaumöglichkeiten als konstant angenommen. Diese sind bei sich ändernden Rahmenbedingungen ebenfalls individuell anzupassen.

5.6.3.4 Gesamtkostenfunktion der Veredelung 1

Wird ein Datenbestand durch die Verwendung von eigenen Daten und durch Datenzukauf realisiert, so sind zusätzlich Homogenisierungskosten zu berücksichtigen, die hier auch linear mit fixem und variablem Anteil (30%) angesetzt werden. Berücksichtigt man darüber hinaus noch einen Automatisierungsgrad (ag) der Homogenisierung, dann reduzieren sich die Datensätze um den Faktor $(0,3N^{eB})(1-ag)$. Dieser Automatisierungsgrad ist ebenfalls situationsabhängig und wird für diese Modellrechnung zunächst mit $ag=95\%$ festgehalten. Die Qualitätssicherungsmaßnahme umfasst allerdings alle $0,3N^{eB}$ zu homogenisierenden Datensätze.

$K_{V1}^g =$	Einkauf Fremdbestände	K^{KF}	$= k_{fix} + d \cdot k_{var}$, Parameterwerte siehe Modellrechnung
		$+ K_{QS}^{KF}$	$K_{prüf}^{KF} = pr_{fix} + n \cdot pr_{var}$
			$K_{tPr}^{KF} = \frac{N^{KF}}{N^L} \left\{ v_{jEl} \left(N^L \Phi \cdot L_{n,C}(\Phi) + N^L \Phi_u - N^L \Phi_u \cdot L_{n,C}(\Phi) \right) \right\}$
			$K_{FK}^{KF} = \frac{N^{KF}}{N^L} \left\{ (fk_{fix} + fk_{var} (N^L \Phi + N^L \Phi_u)) \cdot (1 - L_{n,C}(\Phi)) \right\}$
	Eigener Bestand	$+ K_{QS}^{eB}$	$K_{prüf}^{eB} = pr_{fix} + n \cdot pr_{var}$
			$K_{tPr}^{eB} = \frac{N^{eB}}{N^L} \left\{ v_{jEl} \left(N^L \Phi \cdot L_{n,C}(\Phi) + N^L \Phi_u - N^L \Phi_u \cdot L_{n,C}(\Phi) \right) \right\}$
			$K_{FK}^{eB} = \frac{N^{eB}}{N^L} \left\{ (fk_{fix} + fk_{var} (N^L \Phi + N^L \Phi_u)) \cdot (1 - L_{n,C}(\Phi)) \right\}$
	Homogenisierung	$+ K^{ih}$	$= h_{fix} + 0,3N^{eB}(1-ag) \cdot h_{var}$
		$+ K_{QS}^{ih}$	$K_{prüf}^{ih} = pr_{fix} + n \cdot pr_{var}$
			$K_{tPr}^{ih} = \frac{N^{ih}}{N^L} \left\{ v_{jEl} \left(N^L \Phi \cdot L_{n,C}(\Phi) + N^L \Phi_u - N^L \Phi_u \cdot L_{n,C}(\Phi) \right) \right\}$
			$K_{FK}^{ih} = \frac{N^{ih}}{N^L} \left\{ (fk_{fix} + fk_{var} (N^L \Phi + N^L \Phi_u)) \cdot (1 - L_{n,C}(\Phi)) \right\}$

Die in dieser Gesamtkostenfunktion enthaltene Teilkomponente der Homogenisierung K^{ih} wurde analog zu dem in Kap.4, Abs. 4.6.3.2 formulierten Arbeitszeitmodell zur Abschätzung der Homogenisierungskosten mit Hilfe der HOAI aufgestellt. Dadurch ist die vollständige Kompatibilität hergestellt und es können sämtliche Anwendungsfälle, die in der HOAI definiert sind, auch in diese Gesamtkostenbetrachtung überführt werden.

5.6.3.5 Gesamtkostenfunktion der Veredelung 2

Der Aufbau eines Datenbestandes durch Homogenisierung des eigenen Bestandes mit zugekauften Fremddaten und eigener Erfassungsleistung erweitert die Gesamtkostenfunktion K_{V1}^g um die Kostenfaktoren der eigenen Erfassungsleistung zu:

$K_{V2}^g =$	Einkauf Fremdbestände	K^{KF}	$= k_{fix} + d \cdot k_{var}$, Parameterwerte siehe Modellrechnung
		$+ K_{QS}^{KF}$	$K_{prüf}^{KF} = pr_{fix} + n \cdot pr_{var}$
			$K_{tPr}^{KF} = \frac{N^{KF}}{N^L} \{v_{jEl} (N^L \Phi \cdot L_{n,C}(\Phi) + N^L \Phi_u - N^L \Phi_u \cdot L_{n,C}(\Phi))\}$
			$K_{FK}^{KF} = \frac{N^{KF}}{N^L} \{fk_{fix} + fk_{var} (N^L \Phi + N^L \Phi_u)\} \cdot (1 - L_{n,C}(\Phi))\}$
	Eigener Bestand	$+ K_{QS}^{eB}$	$K_{prüf}^{eB} = pr_{fix} + n \cdot pr_{var}$
			$K_{tPr}^{eB} = \frac{N^{eB}}{N^L} \{v_{jEl} (N^L \Phi \cdot L_{n,C}(\Phi) + N^L \Phi_u - N^L \Phi_u \cdot L_{n,C}(\Phi))\}$
			$K_{FK}^{eB} = \frac{N^{eB}}{N^L} \{fk_{fix} + fk_{var} (N^L \Phi + N^L \Phi_u)\} \cdot (1 - L_{n,C}(\Phi))\}$
	Homogenisierung	$+ K^{ih}$	$= h_{fix} + 0,3 N^{eB} (1 - ag) \cdot h_{var}$
		$+ K_{QS}^{ih}$	$K_{prüf}^{ih} = pr_{fix} + n \cdot pr_{var}$
			$K_{tPr}^{ih} = \frac{N^{ih}}{N^L} \{v_{jEl} (N^L \Phi \cdot L_{n,C}(\Phi) + N^L \Phi_u - N^L \Phi_u \cdot L_{n,C}(\Phi))\}$
			$K_{FK}^{ih} = \frac{N^{ih}}{N^L} \{fk_{fix} + fk_{var} (N^L \Phi + N^L \Phi_u)\} \cdot (1 - L_{n,C}(\Phi))\}$
Eigene Erfassung	$+ K^{eE}$	Parameterwerte siehe Modellrechnung	
	$+ K_{QS}^{eE}$	$K_{prüf}^{eE} = pr_{fix} + n \cdot pr_{var}$	
		$K_{tPr}^{eE} = \frac{N^{eE}}{N^L} \{v_{jEl} (N^L \Phi \cdot L_{n,C}(\Phi) + N^L \Phi_u - N^L \Phi_u \cdot L_{n,C}(\Phi))\}$	
		$K_{FK}^{eE} = \frac{N^{eE}}{N^L} \{fk_{fix} + fk_{var} (N^L \Phi + N^L \Phi_u)\} \cdot (1 - L_{n,C}(\Phi))\}$	

Auch hier führt die notwendige Homogenisierung zu einer Vergrößerung der Gesamtkostenformel. Bei sämtlichen Fixkosten der Prüfung, Verbesserung und Homogenisierung ist zu beachten, dass gewisse Redundanzen auftreten. So dürfen z.B. Rechner (Hardware) nur insoweit bei den Einzelposten auftreten, wie sie speziell dafür benötigt werden um sie z.B. nicht bei Prüfung und Homogenisierung doppelt zu verrechnen. Diese innerbetriebliche Struktur kann durch dieses Schema zwar abgebildet, deren zahlenmäßigen Werte aus externer Sicht dieser Modellierung jedoch nicht berücksichtigt werden.

5.6.3.6 Gesamtkostenfunktion der Veredelung 3

Wird der Datenbestand durch eigene, vorhandene Daten und durch eigene Erfassungsleistung aufgebaut, so kann die Kostenfunktionen folgendermaßen dargestellt werden:

$K_{V3}^g =$	Eigener Bestand	$+ K_{QS}^{eB}$	$K_{\text{prüf}}^{eB} = pr_{\text{fix}} + n \cdot pr_{\text{var}}$
			$K_{\text{tPr}}^{eB} = \frac{N^{eB}}{N^L} \left\{ v_{\text{fEl}} \left(N^L \Phi \cdot L_{n,C}(\Phi) + N^L \Phi_u - N^L \Phi_u \cdot L_{n,C}(\Phi) \right) \right\}$
			$K_{\text{FK}}^{eB} = \frac{N^{eB}}{N^L} \left\{ (fk_{\text{fix}} + fk_{\text{var}} (N^L \Phi + N^L \Phi_u)) \cdot (1 - L_{n,C}(\Phi)) \right\}$
	Eigene Erfassung	$+ K_{QS}^{eE}$	Parameterwerte siehe Modellrechnung
		$+ K_{QS}^{eE}$	$K_{\text{prüf}}^{eE} = pr_{\text{fix}} + n \cdot pr_{\text{var}}$
			$K_{\text{tPr}}^{eE} = \frac{N^{eE}}{N^L} \left\{ v_{\text{fEl}} \left(N^L \Phi \cdot L_{n,C}(\Phi) + N^L \Phi_u - N^L \Phi_u \cdot L_{n,C}(\Phi) \right) \right\}$
		$K_{\text{FK}}^{eE} = \frac{N^{eE}}{N^L} \left\{ (fk_{\text{fix}} + fk_{\text{var}} (N^L \Phi + N^L \Phi_u)) \cdot (1 - L_{n,C}(\Phi)) \right\}$	

Diese Gesamtkostenfunktionen der sechs Aufbaustrategien sollen anhand des Modelldatenbestands (Abschnitt 5.7.1), den definierten Verfahrensanteilen (Tabelle 5-7) und den festgelegten Werten der Kostenfaktoren (Tabelle 5-6) berechnet werden.

5.6.4 Vorgehensweise zur Anwendung dieser Methodik

Die beschriebene Methodik zur Berechnung der Kosten der Qualitätssicherungsmaßnahmen besteht aus mehreren Einzelschritten, welche nacheinander abgearbeitet werden müssen, um die entsprechend richtigen Kostenparameter in die Kostenfunktionen der Aufbaumöglichkeiten einfließen lassen zu können. In folgender Abbildung 5.18 sind deshalb die fünf wesentlichen Schritte dargestellt.

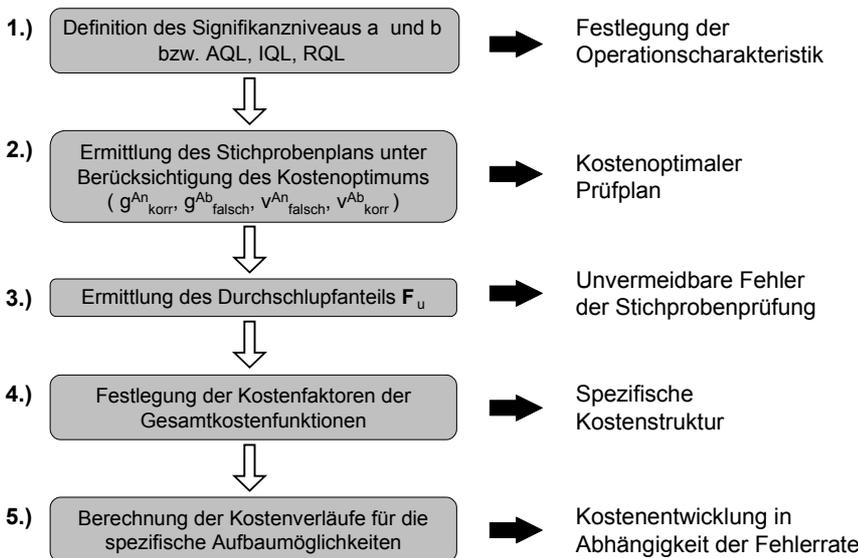


Abbildung 5.18: Entscheidungsfolge zur Berechnung der Kostenverläufe der spezifischen Aufbaumöglichkeiten

In dem **ersten Schritt** sind die Parameter des, mit dem Stichprobenprüfverfahren verbundenen, Hypothesentests (Signifikanzniveau, Fehler 1.Art und Fehler 2.Art) festzulegen. Darin gehen die als tolerierbaren Fehlergrenzen der Grundgesamtheit ein. Durch diese Festlegung kann nun derjenige **Stichprobenprüfplan** mit samt zugehöriger **Operationscharakteristik** festgelegt werden, um den Geodatenbestand nach dem Losprüfverfahren auf fehlerbehaftete Elemente hin zu überprüfen und zu korrigieren.

Die Ermittlung bzw. die Überprüfung des gefundenen Stichprobenplans hinsichtlich des zu erwartenden **Kostenoptimums** ist Gegenstand des **zweiten Schrittes** der Entscheidungsfolge. Stellt sich heraus, dass der in 1.) ermittelte Stichprobenplan dem Kostenoptimum nicht entspricht, so ist die Ermittlung unter Berücksichtigung von 2.) zu wiederholen.

Da sämtliche Stichprobenprüfverfahren einen bestimmten **Durchschlupfanteil** fehlerhafter Elemente besitzen, sozusagen die unvermeidbaren Fehler bei der Stichprobenprüfung, ist für den berechneten Stichprobenplan der entsprechende Anteil im **dritten Schritt** zu ermitteln.

Im **vierten Schritt** ist die **spezifische Kostenstruktur** in Form der Festlegung der Kostenfaktoren der Gesamtkostenfunktionen zu ermitteln. Dabei handelt es sich im wesentlichen um die in Tabelle 5-5 genannten Kostenparameter. Darüber hinaus sind die Verfahrensanteile innerhalb einer Aufbaustrategie festzulegen. Damit werden diejenigen Anteile festgelegt, die angeben, zu wie viel Prozent die Elemente zugekauft, bzw. selbst erfasst werden müssen. Für die in dieser Arbeit durchgeführte Modellrechnung finden sich die angesetzten Verfahrensanteile in Tabelle 5-7.

Im **fünften** und letzten **Schritt** können die Kostenparameterwerte in die Kostenfunktionen übertragen werden und die strategieabhängigen Kostenverläufe mit einander verglichen werden. Durch Variation lassen sich die Grenzwerte einzelner Kosteneinflussparameter bestimmen, ab denen ein Wechsel der Ausbaumöglichkeit kostenmäßig möglich ist.

Ohne eine solche Festlegung der Objektstruktur, dem aufzubauenden Datenvolumen und den Verfahrensanteilen der Aufbaumöglichkeiten als Gewichtungsfaktoren der Kostenrechnung könnte eine nachfolgende Modellrechnung gar nicht stattfinden. Gleichzeitig legt diese Festlegung die ersten maßgeblichen Faktoren offen, die Einfluss auf die Kostenrechnung haben. Überträgt man die Verfahrensanteile auf die über die verschiedenen Strategien aufzubauenden Datenbestände, so ergeben sich die einzelnen Anteile entsprechend folgender Tabelle 5-8.

Tabelle 5-8: Aufteilung der Elementanzahl entsprechend dem Verfahrensanteil

Strategien	Anteile			eigenständig zu erfassende Elemente				Elemente aus eigenem Bestand				Elemente per Zukauf Fremddaten			
	p^{eE}	p^{eB}	p^{KF}	n^{eE}_{geo}	n^{eE}_{att}	n^{eE}_{rel}	Summe	n^{eB}_{geo}	n^{eB}_{att}	n^{eB}_{rel}	Summe	n^{KF}_{geo}	n^{KF}_{att}	n^{KF}_{rel}	Summe
N1			1									100.000	100.000	50.000	250.000
N2	0,5		0,5	50.000	50.000	25.000	125.000					50.000	50.000	25.000	125.000
	0,1		0,9	10.000	10.000	5.000	25.000					90.000	90.000	45.000	225.000
	0,5		0,5	50.000	50.000	25.000	125.000					50.000	50.000	25.000	125.000
	0,9		0,1	90.000	90.000	45.000	225.000					10.000	10.000	5.000	25.000
N3	1			100.000	100.000	50.000	250.000								
V1		0,5	0,5					50.000	50.000	25.000	125.000	50.000	50.000	25.000	125.000
		0,1	0,9					10.000	10.000	5.000	25.000	90.000	90.000	45.000	225.000
		0,5	0,5					50.000	50.000	25.000	125.000	50.000	50.000	25.000	125.000
		0,9	0,1					90.000	90.000	45.000	225.000	10.000	10.000	5.000	25.000
V2	0,2	0,45	0,35	20.000	20.000	10.000	50.000	45.000	45.000	22.500	112.500	35.000	35.000	17.500	87.500
V3	0,5	0,5		50.000	50.000	25.000	125.000	50.000	50.000	25.000	125.000				
	0,1	0,9		10.000	10.000	5.000	25.000	90.000	90.000	45.000	225.000				
	0,5	0,5		50.000	50.000	25.000	125.000	50.000	50.000	25.000	125.000				
	0,9	0,1		90.000	90.000	45.000	225.000	10.000	10.000	5.000	25.000				

Durch diese Aufteilung der Elementanzahl verfügen die unterschiedlichen Kostenfunktionen innerhalb dieser Modellrechnung über die notwendigen Eingangsparameterwerte. Um daraus die Gesamtkosten pro Strategie berechnen zu können, sind Angaben über die Kostenstrukturen des Dateneinkaufs und der eigenen Erfassungsleistung notwendig. Diese Angaben gestalten sich für die hier angesetzte Modellrechnung als schwierig. Während zur Abschätzung der eigenen Erfassungsleistung das Berechnungsverfahren in Kap. 4 zur Hilfe herangezogen werden kann, existiert für die Ermittlung der Einkaufskosten keine vergleichbare Vorgehensweise. Sie können sich deshalb für diese Modellrechnung nur an plausiblen, aber geschätzten Parameterwerten orientieren. Von besonderem Interesse bei der Ermittlung der Einkaufspreisstruktur sind hierbei die Abstufungen der Kosten entsprechend der Objektanzahl und -thematik. Oftmals wird bei amtlichen Datenbeständen die Gebietsgröße, Thematik und die geplante Nutzung an verschiedenen Rechnern der Kostenberechnung über Gebührenverordnungen zu Grunde gelegt. In dieser Gebührenverordnung ist i.d.R. auch eine degressive Kostensteigerung für die Abnahme von großen Datenbeständen enthalten.

Da es im Rahmen dieser Modellrechnung nahezu unmöglich erscheint, sämtliche Abgabe- und Nutzungsbedingungen von digitalen Daten aus amtlichen oder privaten Beständen in Form einer Preisstruktur abzubilden, wird zur weiteren Modellrechnung von einem linearen Verhältnis zwischen Objektanzahl und Verkaufspreis ausgegangen. Es sei für Anwender dieser Methodik und des darauf aufbauenden Berechnungsverfahrens ausdrücklich darauf hingewiesen, dass es zur spezifischen Kostenberechnung unumgänglich ist, diese Kostenstruktur des Dateneinkaufs je nach betrachtetem Fall explizit zu berücksichtigen. Ferner werden für die Aufwendungen der eigenen Erfassungsleistung ebenfalls lineare Abhängigkeiten zwischen Objektanzahl und Kosten innerhalb dieser Modellrechnung vorausgesetzt. Dies dient zum einen für eine einheitliche Betrachtung, da das Berechnungsverfahren nach Kap.4 von Verrechnungseinheiten (VE) ausgeht, die über eine entsprechende Tabelle in konkrete Euro-Beträge umzurechnen sind und damit die Vergleichbarkeit mit dem hier aufgestellten Verfahren der Kostenberechnung der Qualitätssicherung über Währungseinheiten (WE) nicht gegeben ist. Zum zweiten kann so das Verhältnis zwischen dem Aufwand der eigenen Erfassung eines Elements und dem Preis des hypothetischen Einkaufs desselben Elements besser ausgedrückt werden.

Bei sämtlichen Vereinheitlichungen ist zu bedenken, dass diese nicht den Anspruch auf Allgemeingültigkeit erheben und dass mit diesen Vereinheitlichungen eine Vielzahl an situationsspezifischen und besonders kostenverursachenden Faktoren pauschalisiert werden. Auch hier sei dem Anwender dieser Methodik nahegelegt diese Kostenfaktoren für seinen spezifischen Fall explizit zu erheben und mit in die Berechnung einfließen zu lassen.

Aus den dargelegten Gründen werden die Verhältnisse zwischen den eigenen Erfassungskosten pro Element und die Einkaufskosten pro Element für die Modellrechnung variiert. Damit steht bereits zu Beginn ein wesentlicher Einfluss-

faktor auf die Wahl der Aufbaumöglichkeit fest. Die Auswirkungen dieses Einflussfaktors werden im Anschluss an die nachfolgenden Gesamtkostenfunktionen diskutiert.

5.7.2 Die Kostenverläufe der Aufbaumöglichkeiten

Wie bereits in Abschnitt 5.6.3.1 kurz angedeutet, werden hier die Kostenverläufe der Aufbaumöglichkeiten in Abhängigkeit des Verhältnisses zwischen den Kosten der eigenen Erfassung und den Kosten des Einkaufs (K^{EF}/K^{KF}) dargestellt. Dies dient dem Zweck, um daraus anhand des variierten Kostenverhältnisses Aussagen über die mögliche Konkurrenz der Aufbaumöglichkeiten machen zu können. Das Ziel dabei ist es, zu zeigen, dass sich bestimmte Aufbaumöglichkeiten im Vergleich zueinander kostengünstiger verhalten können.

Für den folgenden beispielhaft gezeigten Kostenverlauf (Abbildung 5.19) wird die in Abbildung 5.11 dargestellte OC-Kurve des Prüfplans 1 ($N = 10.000$; $n = 100$; $AC = 2$, d. h.: $AQL = 0,83\%$, $IQL = 1,35\%$, $RQL = 6,15\%$) sowie ein Verhältnis $K^{EF}/K^{KF} = 0,5 / 1$ zu Grunde gelegt. Das Festhalten der Operationscharakteristik macht auch vor dem Hintergrund Sinn, dass gleichgültig ob eigene Daten, Fremddaten oder zu homogenisierende Daten geprüft und verbessert werden sollen, von einer konstanten Testschärfe und einem konstanten Indifferenzpunkt IQL auszugehen ist. Anders ausgedrückt bedeutet dies, dass man bemüht sein wird, die „Qualität“ des Prüf- und Verbesserungsvorgangs konstant zu halten.

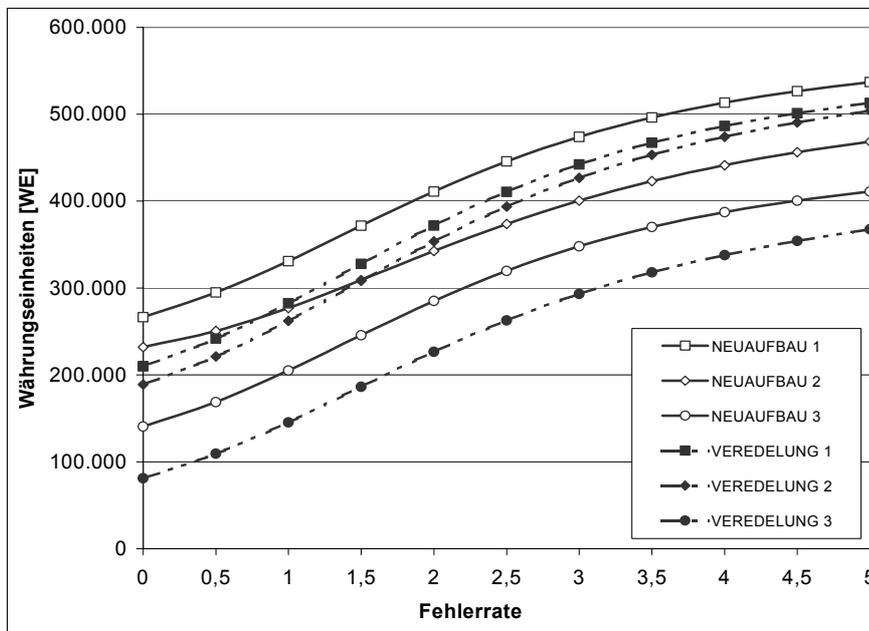


Abbildung 5.19:
Kostenverläufe
konkurrierender Aufbaumöglichkeiten

Anhand dieser Abbildung 5.19 lässt sich die Wirkungsweise der beschriebenen Methodik beweisen. Deutlich zu erkennen sind die existierenden Konkurrenzsituationen der Aufbaumöglichkeiten bei den Schnittpunkten ihrer Kostenverläufe. Diese ergeben sich hier bei einer Fehlerrate von 0,75% und 1% wobei dort mit zunehmender Fehlerrate der *Neuaufbau 2* kostengünstiger verläuft als die *Veredelung 1* und *Veredelung 2*. Anders ausgedrückt bedeutet dies, dass mit Hilfe dieser Methodik und den entsprechenden fallspezifischen Kostenparametern eine Kostenverlaufsprognose erstellt werden kann, die in der Lage ist vorauszusagen, welche der in Frage kommenden Aufbaumöglichkeiten die kostengünstigere ist.

Ferner kann hier die Aussage getroffen werden, dass es Situationen geben kann in denen der Zukauf von Daten und die Eigene Datenerfassung (Neuaufbau 2) günstiger werden kann, als an einem alten, eigenen Bestand festzuhalten und zu versuchen, diesen per Datenzukauf (Veredelung 1) zu veredeln. Diese Situation stellt sich in diesem Beispiel genau bei einer Fehlerrate von 0,75% ein. Das bedeutet, dass ein Unternehmen, welches mit einer Fehlerrate kleiner 0,75% auskommt mit dem Zukauf von Daten und Eigener Erfassung besser beraten ist.

Über dieses Beispiel hinaus existiert eine Vielzahl an Variationsmöglichkeiten der Einflussparameter auf diese Kostenverlaufsprognose, die im Rahmen dieser Arbeit nicht umfassend diskutiert werden können. Deshalb beschränkt sich die Betrachtung auf die folgenden Varianten:

a. Verhältnis eigene Erfassung zu Fremddateneinkauf 1:2 mit 50% automatischer Homogenisierung

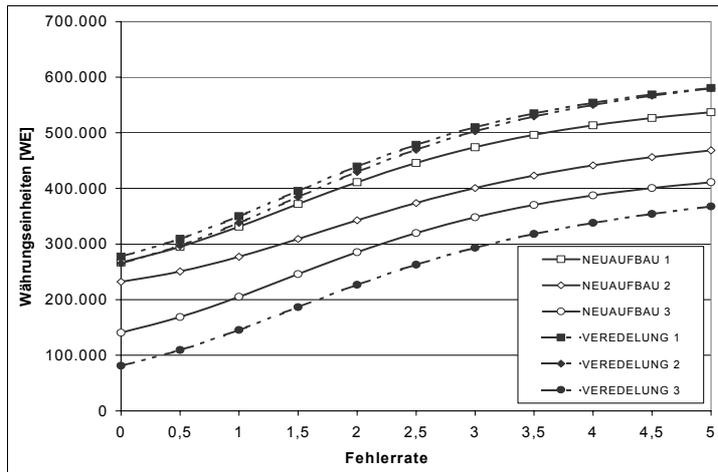


Abbildung 5.20:
Kostenverläufe bei automatischer Homogenisierung von 50%

Erwartungsgemäß tritt eine Kostensteigerung bei den Strategien der Veredelung 1 und Veredelung 2 ein. Interessant hierbei ist die Aussage, dass unabhängig von der Fehlerrate, die Kosten des Neuaufbaus, also des reinen Datenzukaufs trotz dem Verhältnis 0,5:1 günstiger ist als eine Veredelung 1 bzw. Veredelung 2. Anders ausgedrückt bedeutet dies, dass bei einem relativ schlechten Automatisierungsgrad der Homogenisierung von „nur“ 50 % selbst dann der Einkauf von Fremddaten günstiger ist, wenn dieser pro Element doppelt so teuer sind, wie die Kosten der eigenen Erfassungsleistung.

b. Kostenverläufe unter Variation der Operationscharakteristik

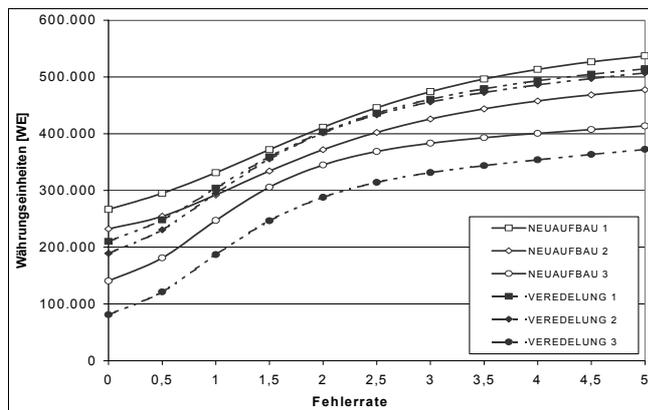


Abbildung 5.21: Kostenverläufe bei Prüfplan 1

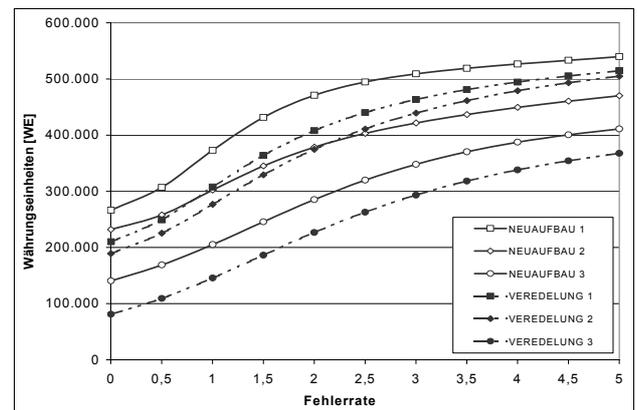


Abbildung 5.22: Kostenverläufe bei Prüfplan 2

Anhand dieser beiden Abbildungen kann verdeutlicht werden, dass sich der Einfluss der Stichprobenprüfung auf einen Bereich der Fehlerraten zwischen ca. 0,5 % und 5 % beschränkt. Man erkennt bei den Kurvenverläufen, bei denen die Prüfverfahren einen schärferen Prüfplan 2 zu Grund gelegt bekommen haben eine steilere Flanke des Kostenanstiegs zwischen ca. 0,5% und 3 % Fehlerrate. Vergleicht man die Kurvenverläufe beider Varianten miteinander, so erkennt man, dass alle Kurvenverläufe der jeweiligen Strategie vom selben Kostenwert ausgehen und im selben Kostenwert enden. Der Kostenanstieg bei schärferem Prüfplan begründet sich auf dem gestiegenen Stichprobenumfang und der damit auch gestiegenen Wahrscheinlichkeit, mehr fehlerhafte Elemente aufzudecken und korrigieren zu müssen.

Während bei der Variante 1 die beiden Konkurrenzsituationen zwischen dem Neuaufbau 2 und der Veredelung 1 bzw. Veredelung 2 bei ca. 0,7 % bzw. 0,9 % Fehlerraten auftreten und der Kostenverlauf von Veredelung 1 und 2 nahezu identisch ist, ergibt sich bei Variante 2 ein anderes Bild. Darin erkennt man eine deutliche Konkurrenz zwischen Neuaufbau 2 und Veredelung 1 bei ca. 1 % Fehlerrate bzw. zwischen Neuaufbau 2 und Veredelung 2 bei ca. 2 % Fehlerrate. Ab diesen 2 % ist der Datenzukauf mit eigener Zusatzerfassung günstiger als die beiden Veredelungsmöglichkeiten 1 und 2.

c. Kostenverläufe unter Variation der Homogenisierungsquote

Bisher wurden sämtliche Betrachtungen und Interpretationen der Kostenverläufe mit einem 30%-igen Anteil notwendigerweise zu homogenisierenden Datenelementen durchgeführt. Hier wird nun dieser Anteil variiert, um dessen Auswirkung interpretieren zu können. Dies ist deshalb auch von besonderem Interesse, da im Marktsektor der Geoinformationsverarbeitung die Homogenisierung von Datenbestände nicht mehr alleine als das Mittel zur Nutzung verschiedener Datenbestände angesehen wird, sondern gerade durch die Definition eines strukturierten Austauschs der Daten, wie er in den Normungsgremien (DIN, CEN, ISO, 1998) oder von privatwirtschaftlichen Organisationen [OGC, 2000] erarbeitet wird, die notwendigen technischen und strukturellen Voraussetzungen geschaffen wurden bzw. sich im Aufbau befinden, um große Datenbestände miteinander kombinieren zu können.

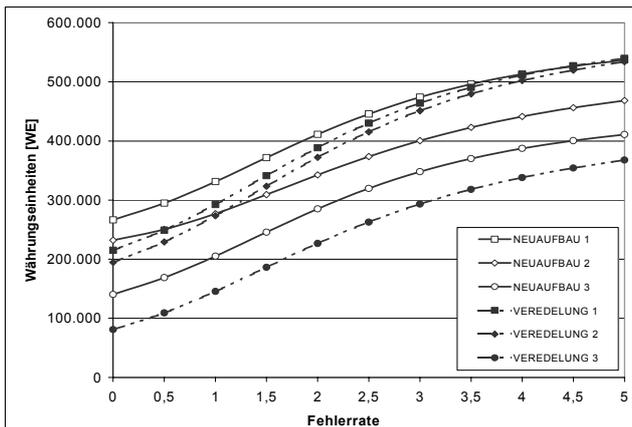


Abbildung 5.23:
Kostenverläufe bei 50%iger Homogenisierungsquote

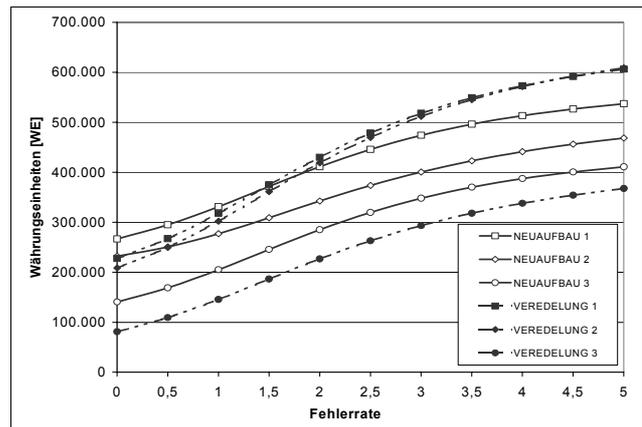


Abbildung 5.24:
Kostenverläufe bei 100%iger Homogenisierungsquote

Diese Kombination lässt sich als Integration unterschiedlicher Datenbestände in z.B. einem Geoinformationssystem interpretieren, welches über eine Metadatenebene und darin enthaltene und miteinander verknüpfte Objektidentifikatoren zugriff auf Daten in unterschiedlichen Beständen besitzen, ohne dass diese Bestände notwendigerweise zu homogenisieren sind. Dabei ist das Herstellen einer solchen Metadatenebene bzw. die Realisierung der Verknüpfung thematisch identischer Objektidentifikatoren aus verschiedenen Beständen nichts anderes als der Vorgang der Homogenisierung mit dem Unterschied, dass dabei nur ein Bruchteil der Datenmenge physisch bearbeitet werden muss und gerade deshalb mit einer Aufwandsreduktion zu rechnen ist. Das bedeutet, dass mit der hier angesetzten Variation der zu homogenisierenden Datenbestände und den aufgestellten Kostenfunktionen auch dieser Entwicklung Rechnung getragen werden kann. Im folgenden sollen deshalb die Kostenverläufe mit den Anteilen 10%, 50% und 100% berechnet werden.

Vergleicht man die Kostenverläufe der Veredelung 1 und 2, für die 50%ige und 100%ige Homogenisierungsquote (Abbildung 5.23 und Abbildung 5.24) so erkennt man die Kostenauswirkung, ausgelöst durch den höheren Homogenisierungsaufwand. Bei der notwendigen 100%igen Homogenisierung aller Datenelemente, werden die beiden Veredelungsvarianten 1 und 2 bei einer Fehlerquote von ca. 1,5-1,75% teurer als der Neuaufbau 1, dem reinen Datenzukauf.

Steht man also vor der Situation, dass 40% der Daten aus eigenen Beständen übernommen und 60% der Daten zugekauft werden können und die Forderung nach vollständiger Homogenisierung der Daten (100%) besteht, dann kann ab einer Fehlerrate von 2% aus Abbildung 5.24 abgelesen werden, dass eine Homogenisierung im Vergleich zum Neuaufbau 1, dem reinen Datenzukauf, teurer wird.

Geht man allerdings nur von Homogenisierungsquoten von 10% bis 50% aus, was dann auch als Integrationslösung im Sinne einer Metadatenverknüpfung unterschiedlicher Datenbestände interpretiert werden kann, so sieht man, dass beide Veredelungsvarianten mit zunehmender Fehlerrate günstiger bleiben als der reine Datenzukauf. In diesem Fall lohnt sich die Homogenisierung. Diese Kostenverlaufsabschätzung ist damit gleichzeitig der Nachweis für die unter Umständen wirtschaftlichere Integrationslösung zweier unterschiedlicher Datenbestände, wenn nicht die Forderung besteht, sämtliche Datenelemente miteinander vergleichen und zu einem Gesamtdatenbestand verschmelzen zu müssen. Diese Methodik bietet damit ein Werkzeug an, abschätzen zu können, ob eine Integrationslösung einer kostenintensiven Datenhomogenisierung vorzuziehen ist.

5.8 Ergebnisse der dargestellten Methodik des Kostenvergleichs der Aufbaustrategien

In den vorangegangenen Abschnitten wurden die Einflussfaktoren auf die unterschiedlichen Kostenverläufe der Aufbaumöglichkeiten beschrieben und die sich daraus ergebenden Konkurrenzsituationen dargestellt. In diesem Abschnitt sollen nun die Ergebnisse der dargestellten Analysemethodik zusammengefasst werden, um einen Überblick über die Einflussfaktoren der Kostenverläufe zu geben. Neben den festgehaltenen Kostenparameterwerten (siehe Tabelle 5-5) besitzen die Faktoren:

- das **Kostenverhältnis** zwischen eigener Erfassungsleistung und Kosten des Zukaufs von Daten bezogen auf das einzelne Datenelement,
- der **Automatisierungsgrad** der Homogenisierung bei der Datenveredelung zwischen vorhandenen eigenen Beständen und zugekauften Fremdbeständen,
- die **Stichprobenprüfpläne** (bzw. die damit definierten Operationscharakteristiken) innerhalb der unterschiedlich notwendigen Qualitätssicherungsmaßnahmen der Aufbaustrategien,
- die **Verfahrensanteile** zur wie viel Prozent Daten zugekauft, selbst erfasst oder dem eigenen, wenn vorhandenen Bestand entnommen werden können sowie
- die **Homogenisierungsquote**, der notwendigerweise zu verschmelzenden Datenelemente,

den größten Einfluss auf den Verlauf der Kostenfunktionen und auf die jeweiligen Konkurrenzsituationen untereinander.

Grundsätzlich wurde deutlich gemacht, dass die **Aufbaumöglichkeiten** maßgeblich vom **Kostenverhältnis** zwischen den eigenen Erfassungskosten pro Element und den Kosten für den Fremddateneinkauf **abhängig** sind. Eine solche Aussage hätte allerdings alleine schon aufgrund allgemeiner Plausibilität auch ohne die hier dargestellte Methodik getroffen werden können. Der wesentliche Vorteil hierbei zeigt sich jedoch in den **Konkurrenzsituationen** zwischen dem **Neuaufbau** und der **Veredelung**. So zeigt sich, dass sich unter Verwendung bekannter Kosteneinflussfaktoren durchaus unter bestimmten Annahmen wie z. B. der Fehlerrate die Veredelung als kostengünstiger als ein Neuaufbau herausstellen kann.

Es hat sich weiter herausgestellt, dass bei einem **schlechten Automatisierungsgrad** der Homogenisierung von 50% der **vollständige Dateneinkauf** inklusive Qualitätssicherungsmaßnahmen selbst dann **günstiger** als **alle Veredelungsstrategien** ist, wenn die Kosten des Dateneinkaufs pro Element doppelt so teuer sind als die eigenen, auf das einzelne Element umgelegte Kosten der eigenen Erfassungsleistung. Erst ab einem Automatisierungsgrad von ca. 75 % werden die Veredelungskosten für kleine Fehlerraten günstiger als eine umfassende Fremddatenbeschaffung.

Bei einem solchen Fall kann den Entscheidungsträgern eines Unternehmens durch diese Analysemethodik eine aussagekräftige Entscheidungshilfe an die Hand gegeben werden, um zwischen der teureren Veredelung eigener Beständen und der im Vergleich dazu günstigeren Variante des vollständigen Dateneinkaufs begründet entscheiden zu können. Da die einmalig getätigte Investition in den Aufbau eines eigenen Bestandes zum Grundkapital eines Unternehmen gezählt werden, tut man sich generell schwer, dieses Grundkapital sozusagen für ein neues Geschäftsfeld unberücksichtigt zu lassen oder gar vollständig zu ersetzen.

Aber auch die unterschiedlichen Qualitätsanforderungen an die Aufbaumöglichkeiten lassen sich mittels Stichprobenprüfpläne berücksichtigen, deren kostenmäßiger Einfluss über die Operationscharakteristik, also der inversen Gütefunktion, in die Kostenverläufe mit eingeht. Dabei zeigte sich eine deutliche Abhängigkeit der Verlaufskurven vom zugrunde gelegten Stichprobenplan. Allgemeingültig lässt sich die Vermutung bestätigen, dass eine höhere Qualitätsforderung auch automatisch mehr Kosten verursacht, da einerseits ein erhöhter Prüfaufwand, aber auch gleichzeitig ein höherer Aufwand der Fehlerkorrektur damit verbunden ist. Damit liegt der Vorteil dieser Methodik klar auf der Hand. Ein Unternehmen, welches sich ein neues Geschäftsfeld, das auf die Nutzung von Geodaten basiert, aufbauen möchte, ist notwendigerweise darauf angewiesen, sich einen Überblick über die qualitätsforderungsabhängigen Kosten zu verschaffen. Hierfür stellt die dargestellte Methodik durch die Ermittlung des kostenoptimalen Stichprobenplans, unter Berücksichti-

gung der Testschärfe sowie den Zusammenhang zwischen dem notwendigen Stichprobenumfang und der maximal tolerierbaren Fehlerrate, die benötigten Werkzeuge bereit.

Weiter wurde gezeigt, dass die Kostenverläufe durch die **Verfahrensanteile**, also dem prozentualen Anteil wie viel Datenelemente zugekauft, selbst erfasst oder aus eigenen Beständen übernommen werden können, abhängen. Darin wurde nachgewiesen, dass der maßgebliche Träger der Kostensteigerung die notwendige Homogenisierung der Daten ist. Dies wird durch die Variation der **Homogenisierungsquote** noch deutlicher herausgearbeitet. Dort ist klar zu erkennen, dass unter den angenommenen Rahmenbedingungen ab einer Homogenisierungsquote von 50% der Datenelemente die Datenhomogenisierung teurer wird als der reine Zukauf von Fremddatenbeständen. Umgekehrt lässt sich damit auch begründen, dass eine Integrationslösung, also die Verknüpfung unterschiedlicher Datenbestände auf Metadatenebene unterhalb der 50%igen Homogenisierungsquote günstiger ist. Können z.B. unterschiedliche Datenbestände auf Objektebene miteinander verknüpft werden ohne, dass dabei die Datenmodelle verändert oder angepasst werden müssen, so kann für den Fall, dass diese Objektverknüpfung nur 50% des Gesamtdatenvolumens ausmacht, eine Integrationslösung günstiger ausfallen als die Vollständige Datenhomogenisierung.

Es bleibt allerdings kritisch anzumerken, dass schon alleine die Objektverknüpfung der größte Kostenfaktor sein wird und dass eine daraufhin mögliche, softwareunterstützte Attributwerthomogenisierung weitestgehend automatisiert werden kann. Fehlerhafte oder nicht mögliche Wertezuordnungen sind dann entsprechend zu kennzeichnen.

5.9 Nutzen für künftige Informationsdienste

Der Nutzen dieser Methodik für künftige Informationsdienstleistungen lässt sich anhand folgender Punkte festhalten:

- Es wurde ein praxistaugliches, generisches Kostenmodell aufgestellt und an einem konkreten Beispiel verifiziert.
- Die Qualitätssicherung von Geodaten durch Stichprobenprüfung ist kostenspezifisch und systematisch planbar.
- Es wurde nachgewiesen, dass bei sicherheitsrelevanten Anwendungen eine Vollprüfung der Information schon allein aus wirtschaftlichen Gesichtspunkten unvermeidlich ist.
- Mit dieser Arbeit existiert eine Entscheidungsmethodik für den Aufbau von Geodatenbeständen zur Kombination des Datenzukaufs mit der eigenen Erfassungsleistung und eigenen vorhandenen Beständen.

Die in diesem Abschnitt 5.5 dargestellte Methodik der Analyse von Kostenverläufen der Aufbaumöglichkeiten lassen sich unter Kenntnis der Kosteneinflussfaktoren **auf sämtliche Fragestellungen des Bestandsaufbaus von Geodaten anwenden**. Es bleibt zu bemerken, dass aufgrund der hier durchgeführten Modellrechnung keine unmittelbaren Schlüsse bezüglich eines konkreten Falls gezogen werden können. Vielmehr verbleibt es Aufgabe bei Anwendung dieser Methodik, die Einflussfaktoren problemspezifisch zu erheben und entsprechend in den Kostenfunktionen zu berücksichtigen. An einer, dieser Methodik vorangestellten, situationsspezifischen Erhebung der Einflussfaktoren führt kein Weg vorbei.

Wesentlich bedeutsamer als die Aussage, **ab welcher Fehlerrate** bzw. welchen Kosteneinflussfaktoren oder der **Testschärfe der Stichprobenprüfung** sich die Kostenverläufe der sechs Aufbaumöglichkeiten schneiden, ineinander fallen oder weit auseinander liegen, ist **die hiermit dargelegte Methodik**, mit der gezeigt worden ist, wie eine Entscheidung über **den kostengünstigsten Aufbau von Geodatenbeständen** unter Einhaltung einer bestehenden **Qualitätsforderung**, getroffen werden kann.

Damit fügt sich diese Methodik in das in Kapitel 4 dargestellte Entscheidungsdiagramm ein. Darin mussten die Entscheidungen, ob ein Bestand die an ihn gestellten Anforderungen erfüllt (E2) bzw. ob die Daten in Fremdbeständen verfügbar (E3) sind nicht nur aus Sicht der reinen technischen Machbarkeit beurteilt werden, sondern auch unter dem vergleichende Aspekt der Wirtschaftlichkeit. Mittels diese Methodik kann somit umfassende Hilfestellung geleistet werden, um die Entscheidung bezüglich der zu wählenden Aufbaumöglichkeit treffen zu können.

6 Ergebnisbewertung und Ausblick

Die Möglichkeiten des Einsatzes raumbezogener Daten weiten sich von den bisher klassischen Einsatzfeldern der Bestandserfassung und Verwaltung sowie der raumbezogenen Planung in neu entstehende Marktbereiche von z.B. dem Informationsdienstleistungen aus. So unterstützen konfektionierte Geodaten Unternehmensentscheidungen in den Bereichen der Standortentscheidungen bei Unternehmensentscheidungen in Form räumlicher Kaufkraft- oder Absatzmarktanalysen. Ebenso dienen sie zum flexibleren, zeit- und kostensparenden Einsatz von Produktionsgütern in den Sektoren der Transport- und Auslieferungslogistik durch die Unterstützung mit Programmsystemen zur Disposition und Tourenplanung.

Damit kommt der anwendungsgerechten Konfektionierung und Bereitstellung von Geodaten eine besondere Bedeutung zu. Ihren Marktwert erhalten Geobasisdaten amtlicher Anbieter erst durch die anforderungsspezifische Veredelung durch weitere thematische Informationen. Diese Konfektionierung der Daten bildet damit den unverzichtbaren Arbeitsschritt zur Realisierung raumbezogener Informationsdienstleistungen oder von speziellen Geodatenprodukten. Dieser Arbeitsschritt repräsentiert damit einen Wertschöpfungsvorgang. Ohne eine solche Wertschöpfung haben die Geo(basis)daten einen geringen Nutzen für den Kunden. Hohe Preise und restriktive Lizenzvereinbarungen amtlicher Geodatenhersteller verhindern derzeit das stärkere Wachstum des Geodatenmarktes. Flexiblere Tarifierungsstrategien können helfen, Geoinformationen dem Kunden ihrem Nutzwerte entsprechend bereitzustellen. Die situationsbedingte Unterscheidung des Nachfrageverhaltens der Kunden in Kunden mit elastischer und unelastischer Nachfrage sowie eine Nutzwertanalyse ist als zentrale Forderung an staatliche Geobasisdatenproduzenten zu richten.

Zwar wird eine Gratisabgabe, bei der sämtliche Erfassungs- und Aktualisierungskosten durch den staatlichen Haushalt finanziert werden, die Verbreitung von Geodaten anfänglich fördern und damit den gesamtwirtschaftlich Nutzen kurzfristig erhöhen, aber gleichzeitig umgeht man damit den Marktwertindex „Preis“. Es kommt zu falschen oder fehlenden effizienten Preissignalen, was langfristig zu einer notwendigen Erhaltung der Geodaten durch allgemeine Steuermittel führt. Alternativen hierfür stehen über die Möglichkeiten der geringfügigen Kostenbeteiligung oder der Tarifierungsstrategie des anteiligen Investitionsrückflusses zur Verfügung. Erst in Verbindung mit der situationsbedingten Unterscheidung des Nachfrageverhaltens des Kunden und der Frage nach der Höhe des Betrages, die der Kunde bereit ist, für qualitative Geoinformationen zu zahlen, werden sich diese Tarifierungsstrategien weiter ausgestalten lassen.

Trotz der steigenden Bedeutung interoperabler Systeme, die das Ziel haben, Geoinformationen über den Zugriff auf unterschiedliche räumlich und physisch von einander getrennte Geobasisdaten generieren zu können, kann auf eine Integrations- und Homogenisierungslösung in vielen Anwendungsbereichen nicht verzichtet werden. So wäre z.B. die Informationsauskunft eines intermodalen Routenplaners ohne die geschaffenen semantischen Relationen zwischen den Haltestellenobjekten und den digitalen Straßenkanten der routingfähigen Straßenkarte gar nicht möglich. Nur durch diese Veredlung der Geobasisdaten (Straßenkarte) um eigene Datenbestände (Haltestellendaten) und die Arbeit der zusätzlichen Datenerfassung (Park&Ride-Plätze, Fußgängerzugänge) ist dieses internetbasierte Auskunftssystem in der Lage, kundenspezifische Informationen zu generieren. Die Betrachtung der Kostenstruktur zu Realisierung einer solchen Geodatengrundlage hat gezeigt, dass die Kosten der qualitätssichernden Maßnahmen das Niveau der Erfassungskosten erreichen oder dieses sogar noch übersteigen kann.

Gerade für die Abschätzung der anfallenden Kosten ist eine modellbasierte Kostenprognose wichtiges Instrument zur Entscheidungsfindung. Das in dieser Arbeit aufgestellten Kostenmodell und das darauf aufbauende Berechnungsverfahren lässt sich bei sämtlichen Aufbaumöglichkeiten von Geodatenbeständen anwenden. So können z.B. viele Datenbestände bereits von Drittfirmen eingekauft werden, was eine umfängliche eigene Erfassung unnötig macht. Insgesamt lassen sich sechs verschiedenen Möglichkeiten des Geodatenaufbaus identifizieren, die in ihrem Ablauf differenzierte Kostenstufen enthalten. Daten können eingekauft, mit eigenen Beständen homogenisiert oder durch zusätzlichen eigene Erfassungsleistung veredelt werden. Es stellt sich dabei grundsätzlich die Frage, welche der Aufbaumöglichkeiten unter Berücksichtigung der Qualitätsforderungen erfüllt werden können.

Die Qualitätsforderung, die sich mit Hilfe von Merkmalen an den Geodatenbestand richtet, kann über Maßnahmen der Qualitätssicherung gewährleistet werden. Die hier betrachtete Maßnahme ist die Überprüfung der Geodaten mittels Stichproben. So ist in dieser Arbeit ein Verfahren aufgezeigt, mit dem das Kostenverhalten der Stichprobenprüfung für

die existierenden sechs Möglichkeiten des Geodatenaufbaus berechnet werden kann, um daraus die jeweils günstigste Möglichkeit zu identifizieren. Durch die transparente Darstellung der maßgeblichen Kosteneinflussparameter und die sich ergebenden Kostenfunktionen kann diese Methodik bei allen Aufgaben der modellbasierten Kostenschätzung angewendet werden. Vielmehr werden Grenzwerte für die Kosteneinflussparameter angegeben, ab denen sich die Kostenverläufe der Stichprobenprüfungen schneiden, also sich die Aufbaumöglichkeiten gegenseitig konkurrieren.

Mit dieser Methodik steht damit nicht nur ein Werkzeug zur modellbasierten Kostenprognose zur Verfügung, sondern es bietet die Möglichkeit, anhand der Variation der Einflussparameter, wie z.B. der Automatisierungsgrad der Homogenisierung, die Schärfe der Stichprobenprüfung, die maximal tolerierbare Fehlerrate der betrachteten Grundgesamtheit sowie die Verfahrensanteile, zu wie viel Prozent Fremddaten eingekauft, selbst erfasst oder mit dem eigenen Bestand homogenisiert werden müssen, diejenige Aufbaumöglichkeit des Geodatenbestandes auszurechnen, bei der sich die Kosten am geringsten gegenüber allen anderen ergeben.

Gerade bei der Realisierung der Geodatenbasis für die intermodale Routenberechnung zeigt sich der entstehende Mehrwerteffekt für den Verkehrs- und Tarifverbund. Der Schwerpunkt des künftigen Einsatzes der modellierten Linien des ÖPNV-Netzwerks liegt im Ersatz unterschiedlicher Printmedien im Bereich von Verkehrslinienplänen, welche heute arbeitsintensiv von Hand hergestellt werden. Durch diesen Aufbau der Geodatengrundlage können nicht nur reine Reiseauskunftssysteme mit innovativen Informationen versorgt werden. Die exakten Streckenlängen der einzelnen Linien lassen sich optimal für die Zwecke der Betriebsleistungsstatistik und als Planungsgrundlage für zukünftige Linien verwenden. Eine aufwendige Erhebung vor Ort entfällt. Des Weiteren können mit Hilfe der neuen Technik Pläne und Karten mit variablen Inhalten und Maßstäben flexibel erzeugt werden. Darüber hinaus erfolgt die Umsetzung der GIS- und ÖV-Daten in die Pläne praktisch ohne Informationsverlust, da die bisherigen Abstimmungen zwischen den verschiedenen Zuständigkeitsbereichen des VVS entfallen. Und schließlich besteht hier ein großes Einsparpotential, indem durch den Wegfall von externen Kartographen sowie aufwendige Drucktechnik ebenfalls Kosten reduziert werden.

7 Literatur

- Adv, (1995): ATKIS-Objektartenkatalog, Druck und Vertrieb: Landesvermessungsamt Nordrhein-Westfalen. www.adv-online.de.
- AHO (2001): Ausschuss der Ingenieurverbände und Ingenieurkammern für die Honorarordnung e.V., Arbeitshilfe für Teil XIII der HOAI. www.aho.de
- Barthelme, N. (1989): GIS-Technologie. Geoinformationssysteme, Landinformationssysteme und ihre Grundlagen, Berlin, Springer Verlag.
- Barthelme, N. (2000): Geoinformatik. Modelle, Strukturen, Funktionen. 3., erweiterte und aktualisierte Auflage, Berlin, Springer Verlag.
- Behl, M. (1981): Die messende Prüfung bei Abweichung von der Normalverteilungsannahme, In Zeitschrift Metrika 28.
- Behr, F.-J (1998): Strategisches GIS-Management: Grundlagen und Schritte zur System Einführung, Wichmann Verlag, Heidelberg.
- Bellmann, R. (1958): On a routing problem, Quart. Appl. Math., 16(1958), S.87-90.
- Bernhard, U. (2002): GIS-Technologien in der New Economy: -Markttransparenz durch Geoinformationssysteme, Heidelberg, Wichmann.
- BGBL (2001): Bundesgesetzblatt. Änderung der Honorarordnung für Architekten und Ingenieure zur Umstellung auf Euro. Jahrgang 2001, Teil 1, Nr.58, Bonn am 14. November 2001
- Bill, R. (1996): Grundlagen der Geoinformationssysteme, Band 2, 2. Auflage, Herbert Wichmann Verlag, Karlsruhe.
- Bracht, A. (2002): Dienstportale in MOBILIST- Ein neuer Weg zur Mobilität. In: Tagungsdokumentation zur BMBF-Tagung: Leitprojekte Mobilität in Ballungsräumen, Berlin.
- Bronstein, I.N., (1991): Taschenbuch der Mathematik, 25. Auflage, Teubner Verlags-GmbH, Stuttgart
- Buhmann, E., Wiesel, J. (2003): GIS-Report 2003 – Software, Daten, Firmen, Karlsruhe, Bernhard Harzer Verlag GmbH.
- Burrough, P.A. (1979): Principles of Geographical Information Systems for land resources assessment, Clarendon Press.
- Caspary, W. (1992): Qualitätsmerkmale von Geodaten, Zeitschrift für Vermessungswesen, Heft 8/9, S.360-367.
- CEN (1995): Geographic Data Files – Version 3.0, Europäisches Komitee für Normung, CEN TC 278. (www.ertico.com/links/gdf/gdfdocs/gdfdoc.htm)
- Coarse, R. (1991): The Nature of the Firm, In: Origins, Evolution, and Development, New York-Oxford, S34-74.
- Cowen, D.J. (1988): GIS versus CAD versus DBMS: what are the differences?, In Photogrammetric Engineering an Remote Sensing 54.
- Dijkstra, E.W. (1959): A note on two problems in connection with graphs, Num. Math.,1(1959), S.269-271.
- Dueker, K.J. (1979): Land resource information systems: a review of fifteen years of experience. In Geo-Processing 1.
- Eitrich, O. (1996): Prozessorientiertes Kostenmodell für die entwicklungsbegleitende Vorkalkulation, Dissertation, Forschungsberichte aus dem Institut für Werkzeugmaschinen und Betriebstechnik der Universität Karlsruhe, Karlsruhe.
- Fiehn, D. (2002): Verteilte Berechnung einer Fahrtauskunft durch kooperierende Routingdienstleister, Dissertation, Verlag im Internet GmbH, www.dissertation.de.

- Fichtenholz, G.M. (1973): Differential- und Integralrechnung, 8.Auflagen, VEB Deutscher Verlag der Wissenschaften, Berlin.
- Fisz, M. (1970): Wahrscheinlichkeitsrechnung und mathematische Statistik, 5.Auflage VEB Deutscher Verlag der Wissenschaften, Berlin.
- Fornefeld, M., Oefinger, P. (2001): Aktivierung des Geodatenmarktes in Nordrhein-Westfalen. Marktstudie der Micus Management Consulting GmbH. Im Auftrag der Staatskanzlei des Landes Nordrhein-Westfalen. Düsseldorf, www.micus.de.
- Freidank, C.-C. (1997): Kostenrechnung: Einführung in die betrieblichen, theoretischen, verrechnungstechnischen Grundlagen des betrieblichen Rechnungswesen. 6., vollst. überarb.u.erg. Auflage, München, Wien, Oldenburg Verlag.
- Gabriel, P., Wagner, R. (2001): Der deutsche Markt für Geodaten und seine Unterstützung durch elektronische Marktplätze – Status Quo und Perspektiven. In: Strobl, Blaschke, Griesebner (Hrsg.): Angewandte Geographische Informationsverarbeitung XIII, Beiträge zum AGIT-Symposium 2001, Salzburg, Wichmann.
- Geiger, W. (1998): Qualitätslehre, 3.Auflage, Verlag Vieweg, Braunschweig.
- Göpfert, W. (1991): Raumbezogene Informationssysteme, Heidelberg, Wichmann.
- GZGI (2002): Gesellschaft zur Zertifizierung von Geo-Information GbR, Beschreibung des Prüfverfahrens für Geodatenprodukte. www.gzgi.de.
- Hake, G., Grünreich, D. (1994): Kartographie, 7. Auflage, Walter De Gruyter, Berlin, ISBN 3-11-013398-9.
- Hautzinger et al., (2001): 3. Evaluationsbericht der AG Wissenschaft für das BMBF Leitprojekt MOBILIST im Auftrag des Bundesministerium für Bildung und Forschung. Stuttgart, Heilbronn, Tübingen (unveröffentlicht).
- Herdeg, E. (1996): Perspektiven und Chancen bei der Vermarktung von amtlichen Geodaten. In: Zeitschrift für Vermessungswesen (ZfV), Heft 8/1996, S.378-387.
- Hillebrand, A. (1991): Ein Kosteninformationssystem für Neukonstruktion mit der Möglichkeit zum Anschluss an ein CAD-System, Dissertation, Hanser-Verlag, München, Wien, ISBN 3-446-16267-4.
- HOAI (1996): Honorarordnung für Architekten und Ingenieure.
- Jermann, J. (2002): GIS-gestützte Modellierung von Anmarschwegen auf Haltestellen des öffentlichen Verkehrs. In: Strobl, J. et al (Hrsg.): Angewandte Geographische Informationsverarbeitung. XIV Beiträge zum AGIT-Symposium 2002, Heidelberg, Wichmann Verlag.
- Joos, G. (2000): Zur Qualität objektstrukturierter Geodaten. Dissertation, Schriftenreihe der Universität der Bundeswehr München, Heft 66, Neubiberg.
- Jost, P.-J.(Hrsg.) (2001): Der Transaktionskostenansatz in der Betriebswirtschaftslehre. Verlag Schäffer-Poeschel, Stuttgart, ISBN 3-7910-1776-4.
- Jost, P.-J. (2001): Der Transaktionskostenansatz im Unternehmenskontext, In Jost, P.-J.(Hrsg.) (2001): Der Transaktionskostenansatz in der Betriebswirtschaftslehre, S. 9-35, Verlag Schäffer-Poeschel, Stuttgart.
- Kainz, W. (1995): Logical consistency, In: Guptil, S.C., Morrison, J.L. (Hrsg.): Elements of spatial Data Quality. Elsevier Science, Kidlington, Tarytown, Tokyo.
- Keller, M., et al. (2001): Tarifierungsstrategien für Geodaten. INFRAS, Arbeitsgruppe Geographische Informationssysteme der Schweizerischen Informatikkonferenz. Bern, Zürich.

- Koelle, Dietrich E. (1971): Statistisch-analytische Kostenmodelle für die Entwicklung und Fertigung von Raumfahrtgerät. Dissertation an der TU München und in Zeitschrift Raumfahrtforschung (Sonderdruck) Heft 5 (1971) und Heft1 (1972), München.
- Krallmann, H., et al. (1999): Systemanalyse im Unternehmen. Partizipative Vorgehensmodelle, objekt- und prozessorientierte Analysen flexible Organisationsarchitekturen. 3.Auflage, Oldenbourg Verlag, München Wien. ISBN 3-486-25124-4.
- Kronjäger, W. (1992): Results of Benchmark Test – Data Acquisition, Proceedings Internationaler Workshop: European Digital Road Map, 2.-3. März 1992, unveröffentlicht.
- Lawford, G.J. (1995): Geodata quality validation. Cartography, Vol. 24, No.2.
- Longley, P.A., et al. (1999): Geographical Information Systems: Principles, techniques, applications and management (vol.1+2). New York, John Wiley & Son.
- Maguire, D.J., Dangermond, J. The functionality of GIS, In : Maguire, D.J, Goodchild, M., Rhind, D. (ed.): Geographical Information Systems. Longman Scientific & Technical (1991):
- Menges, G., Skala, H.J.(1973): Grundriss der Statistik 2: Daten, Westdeutscher Verlag, Opladen.
- MentzDV : Internetseite der Firma MentzDV GmbH, www.mentzdv.de.
- Mobilist, (2002): Tagungsdokumentation zur BMBF-Tagung: Leitprojekte Mobilität in Ballungsräumen, Berlin.
- Möhlenbrink, W. (1992): EDRM-Further Strategy, Outlook, Proceedings Internationaler Workshop: European Digital Road Map, 2.-3. März 1992, unveröffentlicht.
- Möhlenbrink, W. (1998): Traffic-Guidance and Information Systems – New Technologies for the Information Market, In Kahmen, H., Brückl, E., Wunderlich, Th. (Eds.): Proceedings of the IAG SC4 Symposium in Eisenstadt/Austria.
- Moore, E.F. (1957): The shortest path through an maze, 1957 Harvard University Press, 1959.
- Oesterreich, B. (1997): Objektorientierte Softwareentwicklung mit der Unified Modelling Language. 3.Auflage, Oldenbourg Verlag, München.
- Olfert, K., Rahn, H.-J. (2001): Lexikon der Betriebswirtschaftslehre - Kompendium der praktischen Betriebswirtschaft. 4., überarbeitete und aktualisierte Auflage, Kiehl Verlag, Ludwigshafen.
- Peuquet, D.J. (1984): Data structures for a knowledged-based GIS. In: Proceedings of the International Symposium on spatial data handling, Zürich, Geographisches Institut der Universität Zürich-Irchel.
- Pickel, H. (1989): Kostenmodell als Hilfsmittel zum kostengünstigen Konstruieren. Hanser Verlag, München, ISBN 3-446-15770-0.
- Plümer, L., Gröger, G. (1996): Nested Maps - a Formal, Provably Correct Object Model for Spatial Aggregates. Proceedings of the 4th ACM Workshop on Advances in GIS, Rockwell, Maryland.
- Plümer, L. (1996): Zur Prüfung der Konsistenz von Geometrie und Topologie in Landkarten. Nachrichten aus dem Karten- und Vermessungswesen. Verlag des Instituts für Angewandte Geodäsie, Frankfurt a.M.
- Prinz, Th. (2002): GIS als Instrument zur Standortoptimierung im öffentlichen Personennahverkehr, In: Strobl, J. et al (Hrsg.): Angewandte Geographische Informationsverarbeitung. XIV Beiträge zum AGIT-Symposium 2002. Wichmann-Verlag. Heidelberg.
- Rasch, D. (1968): Elementare Einführung in die mathematische Statistik, VEB Deutscher Verlag der Wissenschaften, Berlin, Leipzig.
- Rumbough, J. (1993): Objektorientiertes Modellieren und Entwerfen, Hanser Verlag, München.

- RVP (2001): Regionalverkehrsplan der Region Stuttgart. Herausgegeben durch den Verband der Region Stuttgart, Stuttgart, www.region-stuttgart.org.
- Sachs, L (1968): Statistische Auswertemethoden, Springer Verlag, Berlin, Heidelberg, New York.
- Schmidt, A. P.(1999): Der Wissensnavigator – Das Lexikon der Zukunft, Deutsche Verlags-Anstalt DVA. Stuttgart, München.
- Schilcher, M., et al..(1996): Geoinformationssysteme – Zwischenbilanz einer stürmischen Entwicklung. In: Zeitschrift für Vermessungswesen (ZfV), Heft 8, S.363-377.
- Senkler, K., Remke, A. (2002): Trading Geoinformation within an Infrastructure for distributed GI Services, In: Zeitschrift Geoinformationssysteme 9/2002, S.12-17.
- Stanek, H., et al. (1995): Modellierung und Normierung von Datenqualität im GIS, In: Dollinger, F. Strobl, J.(Hrsg.): Angewandte Geographische Informationstechnologie VII, Beiträge zum GIS-Symposium 1995, Salzburg.
- Stark, M. (2001): Informationsanforderung an die intermodale Routensuche zur Verknüpfung der Verkehrsnetze. Beitrag zur DGON Jahreshauptversammlung 2001, Mobilität und Sicherheit, S.116-125, Wolfsburg.
- Stark, M. (2001a): Kostenprofile für die Beschaffung am Markt verfügbarer digitaler Datenbestände. Bericht zum Forschungsprojekt M21-Einführung neuer telematikbasierter Informationsdienstleistungen und Untersuchung ihrer Wirksamkeit im Auftrag des Umwelt- und Verkehrsministeriums Baden-Württemberg und der DaimlerChrysler AG, Stuttgart, 2001, unveröffentlicht.
- Stark, M., Torlach, V. (2003): Realisierungsaufwand zur Herstellung einer Geodatenbasis für die intermodale Routenberechnung zwischen Individual- und öffentlichem Personennahverkehr. In: Strobl, J. et al (Hrsg.): Angewandte Geographische Informationsverarbeitung XV. Beiträge zum AGIT-Symposium 2003, Heidelberg, Wichmann Verlag 2003.
- Stat.LaBw (2003): Statistisches Landesamt Baden-Württemberg, Internet <http://www.statistik.baden-wuerttemberg.de>.
- TeleAtlas (2000): MultiNet - Standard Data Model Description, Version 3.0, TeleAtlasBV, Gent, Belgien, interne Spezifikation, unveröffentlicht.
- Uhlmann, W. (1982): Statistische Qualitätskontrolle, Teubner Studienbücher, 2. Auflage, Teubner Verlag, Stuttgart.
- VDI (1987): Wirtschaftliche Entscheidungen beim Konstruieren - Methoden und Hilfen. VDI-Richtlinie 2235, VDI-Verlag, Düsseldorf.
- Warnecke, H.J., et al.. (1996): Kostenrechnung für Ingenieure, 5.überarb.u.erw. Auflage, München, Wien, Hanser Verlag.
- Wieser, E. (1989): Systemanalytische Aspekte kommunaler Landinformationssysteme, Dissertation, Deutsche Geodätische Kommission, DGK-Heft 350, Reihe C, Bayerische Akademie der Wissenschaften, München.
- Wiltschko, T. (2004): Sichere Information durch infrastrukturgestützte Fahrerassistenzsysteme zur Steigerung der Verkehrssicherheit an Straßenknotenpunkten. Dissertation. Fortschritt-Bericht VDI, Reihe 12, Nr. 570. Düsseldorf. VDI-Verlag.

Danksagung

Diese Arbeit entstand während meiner sechsjährigen Tätigkeit als wissenschaftlicher Angestellter am Institut für Anwendungen der Geodäsie im Bauwesen der Universität Stuttgart unter Leitung von Prof. Dr.-Ing. Wolfgang Möhlenbrink.

Mein besonders herzlicher Dank gilt Herrn Prof. Dr.-Ing. Wolfgang Möhlenbrink für die zahlreichen und gewinnbringenden Diskussionen, für sein Verständnis, sein Vertrauen und seine aktive Unterstützung, die er mir im Laufe dieser Arbeit zu Teil werden ließ. Ebenso möchte ich mich bei Herrn Prof. Dr.-Ing. Markus Friedrich für die bereitwillige Übernahme des Koreferates bedanken.

Allen Kolleginnen und Kollegen des Instituts für Anwendungen der Geodäsie im Bauwesen sei für viele hilfreiche Diskussionen sowie für die freundschaftliche Arbeitsatmosphäre herzlich gedankt. Abschließend danke ich für alle Unterstützungen seitens meiner Familie, die mir mit ihrem Verständnis und Rat sehr geholfen haben, diese Arbeit abzuschließen.

Lebenslauf

Name: Martin Alexander Stark
 Geburtsdatum: 28.06.1972
 Geburtsort: Stuttgart
 Familienstand: ledig

Ausbildung

1982 – 1991: Allgemeine Hochschulreife am Georgii-Gymnasium in Esslingen am Neckar
 1992 – 1997: Studium des Vermessungswesens an der Universität Stuttgart
 29.08.1997: Abschluss Diplomingenieur Vermessungswesen (Geodäsie und Geoinformatik)
 01.07.1998: Auszeichnung der Diplomarbeit durch den Förderverein des Studiengangs Vermessungswesen

Berufstätigkeit

09/97 – 08/03: Wissenschaftlicher Angestellter an der Universität Stuttgart, Institut für Anwendungen der Geodäsie im Bauwesen.
 Seit 15.09.2003 Mitarbeiter der Dornier Consulting Automotive Systems GmbH, Sindelfingen.

