

Masterthesis

Konzept zur Integration von Verkehrsinformationen für ein Supply-Chain-Simulationsmodell

Simon Tietze
Matrikelnummer: 142825
Studiengang: Logistik

ausgegeben am: 20.05.2016

eingereicht am: 07.11.2016

Betreuer:
Univ.-Prof. Dr.-Ing. Markus Rabe
Astrid Klüter, M. Sc. Logistik

Inhaltsverzeichnis

Inhaltsverzeichnis	2
1 Einleitung	3
2 Einordnung der City-Logistik in den logistischen Kontext	6
2.1 Aktuelle Herausforderungen der City-Logistik	8
2.2 Forschungsansätze für die urbane Versorgung	10
2.3 Simulation als Analyseansatz für Probleme der City-Logistik	12
3 Grundlagen der Simulation	14
3.1 Vorgehensmodelle für die ereignisdiskrete Simulation	14
3.2 Datenbehandlung in der Simulation	19
3.3 Simulation im Rahmen logistischer Fragestellungen	24
4 Konzepte zur Integration von Verkehrsinformationen in die Simulation	26
4.1 Vorgehensweisen bei der Wegstreckeberechnung	27
4.2 Konzepte zur Integration weiterer Verkehrsinformationen im Rahmen der Fahrzeitberechnung	30
4.3 Vorgehen im Simulationswerkzeug <i>SimChain</i>	33
4.4 Identifikation der Forschungsbedarfe	34
5 Analyse der verschiedenen Vorgehensweisen zur Integration von Verkehrs- informationen	36
5.1 Das Attika als Untersuchungsgebiet für den Anwendungsfall	36
5.2 Untersuchung der Distanz-Metriken zur Approximation der Wegstrecke . . .	38
5.2.1 Berechnungen und Analyse für die Minkowski-Metrik	40
5.2.2 Berechnungen und Analyse für die gewichtete euklidische Distanz . .	46
5.2.3 Berechnungen und Analyse für die potenzierte euklidische Distanz .	50
5.2.4 Gegenseitiger Vergleich der verschiedenen Distanz-Metriken	54
5.3 Analyse und Weiterentwicklung der Vorgehensweisen zur zeitabhängigen Fahrzeitberechnung	58
6 Konzept zur Integration von Verkehrsinformationen in die Simulation	62
6.1 Einbeziehung der Wegstreckenapproximation in die Simulation	62
6.2 Einbeziehung der zeitabhängigen Fahrzeitberechnung in die Simulation . . .	66
7 Zusammenfassung und Ausblick	69

Literaturverzeichnis	76
Abbildungsverzeichnis	77
Tabellenverzeichnis	78
Abkürzungsverzeichnis	79
A Anhang: Programm zur Ermittlung der tatsächlichen Wegstrecken	80
B Anhang: Taxonomie zur Einordnung von Projekten der City-Logistik	84

1 Einleitung

Während Mitte des 20. Jahrhunderts noch über 70 % der Weltbevölkerung auf dem Land wohnten, stellt das Jahr 2007 einen Wendepunkt dar, da nun erstmals mehr Menschen in urbanisierten Räumen als in ländlichen Siedlungen lebten. Nach Einschätzung der UN wird sich diese Entwicklung fortsetzen, sodass im Jahr 2050 circa 75 % der Weltbevölkerung in Städten angesiedelt sein werden. (United Nations 2015, S. 7)

Obwohl bereits ein Großteil der Weltbevölkerung in Städten lebt, beanspruchen sie nur 2-3 % der Landfläche, wohingegen 80 % der Treibhausgase und 80 % des Ressourcenverbrauchs auf sie entfallen (Siedentop 2015, S. 17). Des Weiteren werden 40 % der CO₂-Emissionen, die beim Landtransport entstehen, in Städten verursacht (Lehmacher 2015, S. 10). Jedoch erwirtschaften Städte auch 80 % der globalen Wirtschaftsleistung und verzeichnen im Vergleich zu ländlichen Regionen einen höheren Anteil am globalen Wirtschaftswachstum (Graham 2014, S. 5). Darüber hinaus steigt in urbanisierten Räumen aufgrund von steigender Kaufkraft und verändertem Konsumverhalten der Bedarf an Fast Moving Consumer Goods (FMCG), den Gütern des täglichen Bedarfs. Somit stellen Städte die Netzknoten der Produktion, Distribution und Konsumption materieller Güter im globalen Stoffstromnetz dar (Siedentop 2015, S. 17).

Dies hat zur Folge, dass mit zunehmender Urbanisierung und unter Nutzung von heutigen Logistikkonzepten zur Ver- und Entsorgung des städtischen Raums mit einer Zunahme des Güterverkehrsaufkommens zu rechnen ist. Eine Herausforderung dabei ist, dass der Transport zur Befriedigung der gesteigerten Bedarfe auf begrenzter Infrastruktur erfolgen muss und dabei oftmals im Kontrast zu den Anforderungen an Versorgungskonzepten hinsichtlich ökonomischer, ökologischer und sozialer Nachhaltigkeit steht (Klumpp et al. 2013, S. 169). Dabei bieten Städte grundsätzlich auch wirtschaftliche Vorteile gegenüber ländlichen Bereichen, da hier in unterschiedlichem Maße Skaleneffekte aufgrund der Größe und Komprimiertheit erzielt werden können (Taubenböck und Wurm 2015, S. 7). Dies kann durch den pro-Kopf-Vergleich der Treibhausgas-Emissionen in amerikanischen Großstädten verdeutlicht werden. Diese sind in Städten mit hoher demographischer und baulicher Dichte wesentlich niedriger als in zersiedelteren und weniger stark konzentrierten Räumen (Taubenböck und Wurm 2015, S. 9). Um den Anforderungen an die Versorgungskonzepte urbanisierter Räume gerecht zu werden, müssen diese Effizienzvorteile konsequenter genutzt und in neue Logistikkonzepte integriert werden (Siedentop 2015, S. 18). Neue Konzepte im logistischen Bereich und besonders im Güterverkehr werden oft mittels Simulation analysiert, da sie ein probates Mittel zur Bewertung dynamischer Zusammenhänge im logistischen Kontext darstellt (Eley 2012; Klug 2010).

Das Simulationsvorgehen wird in den Prozess der Informationsgewinnung und der Modellbildung segmentiert (Rabe et al. 2007, S. 6). Dabei verteilt sich der Gesamtaufwand eines Projektes in gleichen Teilen auf beide Bereiche. Die Informationsgewinnung hat hierbei einen besonderen Stellenwert, da die Qualität der Eingangsdaten eines Simulationsmodells großen Einfluss auf die Aussagekraft der Simulationsergebnisse hat.

In verschiedenen Projekten der City-Logistik werden Konzepte für die Verteilung von Gütern z. B. des Bereiches der FMCG oder des E-Commerce in städtischen Gebieten simuliert und anschließend analysiert. Hierfür werden häufig datengetriebene Simulationsmodelle genutzt, wie bspw. im EU-Projekt U-TURN oder bei der Untersuchung der Profitabilität von Lieferaufträgen bei Cleophas und Ehmke (2014). Besondere Aufmerksamkeit wird oftmals der *letzten Meile* zuteil. Die *letzte Meile* wird hierbei als eines der wichtigsten Bestandteile einer Lieferkette wahrgenommen, ist jedoch häufig durch hohe Kosten und Ineffizienz geprägt (Gevaers et al. 2011). Somit ergibt sich gerade für die City-Logistik die Anforderung, Belieferungen in innerstädtischen Gebieten sorgfältig zu planen. Daraus resultiert für die Durchführung von Simulationsstudien innerhalb der City-Logistik in besonderem Maße, dass die einzelnen Transportrelationen und die benötigte Fahrzeit innerhalb eines Distributionsnetzwerkes detailliert abgebildet werden müssen. Nur so können Simulationsergebnisse produziert werden, die präzise genug sind, um ihre Implikationen auf die Realität zu übertragen. Vor dem Hintergrund eines datengetriebenen Simulationsmodells wird somit die Relevanz der Modelleingangsdaten hinsichtlich ihrer Qualität deutlich.

Das Ziel dieser Arbeit ist es, für die Durchführung von Simulationsstudien im verkehrslogistischen Bereich ein Konzept zur Integration von Verkehrsinformationen zu entwickeln. Im Rahmen des Informationsgewinnungsprozesses ist zu untersuchen, welche Informationen für die Simulation von Distributionsnetzwerken relevant sind und wie die entsprechenden Daten beschafft und aufbereitet werden können. Dabei wird konkret auf die Modellierung von Transportrelationen innerhalb von Distributionsnetzwerken eingegangen. Außerdem wird die Einbeziehung weiterer verkehrsrelevanter Informationen für die Fahrzeitberechnung thematisiert.

Zur Erreichung der beschriebenen Ziele soll das Vorgehen nach folgender Systematik erfolgen: Zu Beginn wird eine Einordnung der Problemstellung innerhalb des Themenbereichs *City-Logistik* vorgenommen und somit der allgemeine Rahmen für diese Ausarbeitung gelegt. Mit Bezug auf die City-Logistik werden anschließend derzeitige Herausforderungen und aktuelle Forschungsansätze herausgearbeitet. Auf diese Weise kann zunächst aufgezeigt werden, welche theoretischen Grundgedanken als Grundlage für diese Ausarbeitung dienen können.

Aktuelle Forschungsansätze im Bereich der City-Logistik wenden häufig die Simulation an, um verschiedene Umsetzungsszenarien zu vergleichen. Deshalb wird nachfolgend auf die Simulation im Allgemeinen und schließlich im Speziellen auf den Informationsgewinnungsprozess innerhalb der Simulation sowie der Einordnung in das Simulationsvorgehensmodell eingegangen. Es wird aufgezeigt, welche Rolle der Informationsgewinnungsprozess im Si-

mulationsvorgehensmodell einnimmt und welche Relevanz die Modelleingangsdaten für die Durchführung einer Simulationsstudie besitzt.

Da die Simulation von Wegstrecken und Fahrzeiten insbesondere für die City-Logistik eine wichtige Rolle einnimmt, wird eine Literaturrecherche durchgeführt. Dabei werden zunächst verschiedene Vorgehensweisen zur Bestimmung der Fahrdistanz sowie der anschließenden Berechnung der Fahrzeit analysiert. Außerdem wird auf das konkrete Vorgehen zur Wegstreckenapproximation und Fahrzeitberechnung im Simulationswerkzeug *SimChain* eingegangen.

Auf Basis der vorhergehenden Literaturrecherche werden anschließend die verschiedenen Vorgehensweisen zur Approximation der Wegstrecke auf den Anwendungsfall der Metropolregion Athen übertragen und analysiert. Dazu wird zunächst das Untersuchungsgebiet dargestellt und daraufhin das Vorgehen für die Analyse der Distanz-Metriken vorgestellt. Für die Durchführung der Simulationsstudie wurden im Vorfeld von Logistikdienstleistern Liefernetzwerk-Daten aus Athen zur Verfügung gestellt. Für die Darstellung des Distributionsnetzwerkes werden die Daten über das Liefernetzwerk zunächst aufbereitet, um sie anschließend in der Analyse verwenden zu können. Nach der Analyse der Distanz-Metriken wird außerdem eine Gegenüberstellung zum bisherigen Vorgehen im Simulationswerkzeug *SimChain* vorgenommen, um zu untersuchen, welche Vorgehensweise zur Wegstreckenapproximation am besten für das Distributionsnetzwerk in der Metropolregion Athen geeignet ist. Außerdem wird auf die zeitabhängige Fahrzeitberechnung in Distributionsnetzwerken eingegangen.

Um die gewonnenen Erkenntnisse nicht nur für den Anwendungsfall der Metropolregion Athen zu verwenden, sollen sie im Allgemeinen für datengetriebene Simulationsmodelle nutzbar gemacht werden. Dazu wird eine Vorgehensweise zur Integration der Verkehrsinformationen innerhalb der Simulation entworfen. Die Arbeit schließt mit einer Zusammenfassung der gewonnenen Erkenntnisse und einem Ausblick auf weitere Forschungsbedarfe in diesem Untersuchungsgebiet.

2 Einordnung der City-Logistik in den logistischen Kontext

Eine Beschreibung des Begriffes *City-Logistik* in der wissenschaftlichen Literatur geht mit unterschiedlichen Auffassungen bezüglich des Begriffsinhaltes einher. Erschwerend kommt hinzu, dass es nur wenig konkrete Definitionen dieses Begriffes gibt.

Taniguchi et al. (2001) definieren die City-Logistik als „...the process for totally optimizing the logistics and transport activities by private companies in urban areas while considering the traffic environment, traffic congestion and energy consumption within the framework of a market economy“ (Taniguchi et al. 2001, S. 2). Diese Definition ist im Vergleich zu anderen sehr allgemein formuliert. Als ein Grundgedanke wird von Taniguchi et al. (2001) ein umweltverträglicher und nachhaltiger Güterverkehr hervorgehoben. Crainic et al. (2009, S. 433) beschreiben das Hauptziel der City-Logistik wie folgt: „The general goal is to reduce the impact of freight-vehicle movements on city living conditions in terms of congestion, emissions, and pollution, while not penalizing the city social and economic activities.“ Auch hier wird der ökonomische und ökologische Gedanke der City-Logistik deutlich; darüber hinaus wird auch eine soziale Komponente in die Betrachtung einbezogen.

Wie die bei Crainic et al. (2009) adressierten Ziele erreicht werden können, wird bei Kaupp (1998) ersichtlich. Kaupp versteht unter dem Begriff City-Logistik „...die an ökonomischen und ökologischen Zielen ausgerichtete Planung, Steuerung und Kontrolle logistischer Leistungsprozesse in einem unternehmensübergreifenden Logistiksystem. Aufgabe der City-Logistik ist die kooperative Produktion von Logistikdienstleistungen, die eine Ver- und Entsorgung einer Stadt oder eines Ballungsraums sicherstellen“ (Kaupp 1998, S. 24). Die Begriffserklärung baut inhaltlich auf den Definitionen von Taniguchi et al. (2001) und Crainic et al. (2009) auf. Die ökonomische und ökologische Zielsetzung fungiert aber nur noch als Basis für die Aktivitäten innerhalb eines unternehmensübergreifenden Logistiksystems. Diese Ziele sollen mittels der Kooperation von Logistikdienstleistern erreicht werden. Durch die Integration des Kooperationsgedankens wird der Begriff City-Logistik wesentlich stärker eingegrenzt.

Schulte erläutert den Begriff City-Logistik zunächst als die Menge aller Tätigkeiten, „...die sich auf die bedarfsgerechte, nach Art, Menge, Zeit, Raum und Umweltfaktoren abgestimmte, effiziente Bereitstellung (bzw. Entsorgung) von Gütern in einer Stadt beziehen“ (Schulte 2013, S. 220). Auch hier wird wieder eine ökonomische und ökologische Zielsetzung angenommen. Zur Erreichung dieser Ziele wird von Schulte die Bündelung als wesentli-

ches Merkmal mit der City-Logistik zusammengeführt; demnach ist das Hauptmerkmal der City-Logistik die „...Bündelung der Auslieferungstouren und zwar sowohl die Bündelung der Lieferungen von verschiedenen Lieferanten für einen Empfänger als auch die Bündelung der Lieferverkehre für benachbarte Empfänger“ (Schulte 2013, S. 220). Bündelung bedeutet also sowohl eine Konsolidierung der Lieferungen von unterschiedlichen Logistikdienstleistern als auch die Koordinierung der Transportaktivitäten innerhalb der Stadt.

Von Strauß wird eine weitere Definition angeboten, in der zunächst auch auf das Konsolidierungs- und Koordinierungskonzept eingegangen wird. Demnach ist die City-Logistik als „... eine Bündelung gleichgerichteter Güterströme innerhalb eines Stadtgebietes definiert. Sie hat das Ziel eines effizienten und stadtverträglicheren Güterverkehrs und umfasst die Ausgestaltung aller dazu notwendigen infrastrukturellen, organisatorischen, informationstechnischen und personellen Komponenten“ (Strauß 1997, S. 28). Als Erweiterungen zur Begriffsbestimmung von Schulte (2013) werden die angedeuteten Aktivitätsbereiche konkretisiert. Dadurch wird der integrative Charakter der City-Logistik deutlich; ins Zentrum der Betrachtung werden neben der ökonomischen und ökologischen Zielsetzung die Konkretisierung aller infrastrukturellen, organisatorischen, informationstechnischen und personellen Komponenten gestellt. Dieser integrative Ansatz wurde bereits von Taniguchi et al. (2001) andeutungsweise aufgegriffen, der innerhalb der City-Logistik von vollständigen Optimierungen der logistischen Aktivitäten ausging.

Die Konzepte der Kooperation zwischen Unternehmen und Konsolidierung sowie Koordinierung von Warenströmen in sogenannten Güterverkehrszentren (GVZs) ist Bestandteil vieler weiterer Begriffsbeschreibungen der City-Logistik (Allen et al. 2012; Pamučar et al. 2016; Piontek 2013). Der Umsetzung dieser Konzepte wird eine hohe Relevanz attestiert, um die Ziele der City-Logistik zu erreichen. Crainic et al. (2009) setzen hierbei den Schwerpunkt auf den Ressourceneinsatz und erweitern die Leitidee der City-Logistik um den Grundgedanken, dass der Transport innerhalb der Städte zusätzlich in umweltfreundlichen Fahrzeugen erfolgen soll.

Durch die Darstellung der Begriffsbeschreibungen für die City-Logistik konnte aufgezeigt werden, dass es zwischen den einzelnen Begriffsbestimmungen viele Überschneidungen gibt. Trotzdem ist der Begriff selbst nicht eindeutig abgegrenzt, und es existieren viele verschiedene inhaltliche Schwerpunkte. Dies bezieht sich bspw. auch auf die Einbindung von verschiedenen städtischen Akteuren in die Problemlösung der City-Logistik. Anfangs soll deshalb zunächst noch einmal auf die konkreten Herausforderungen eingegangen werden, mit denen die City-Logistik zukünftig aufgrund demographischer und ökonomischer Entwicklungen konfrontiert sein wird. Im Anschluss daran wird die Kategorisierung der verschiedenen Forschungsansätze zur Begegnung der dargestellten Probleme der City-Logistik vorgestellt sowie exemplarisch Studien und Projekte aus den unterschiedlichen Bereichen angeführt. Abschließend wird auf Basis der vorherigen Kapitel und als Hinleitung auf das Thema dieser Arbeit konkret auf die Simulation als Werkzeug zur Problemlösung in der City-Logistik eingegangen.

2.1 Aktuelle Herausforderungen der City-Logistik

Städte werden als zentrale Knoten sowohl für die Produktion als auch für die Distribution und Konsumption materieller Güter betrachtet (Siedentop 2015, S. 17). Mit steigender Einwohnerzahl in Städten aufgrund der zunehmenden Urbanisierung ist zunächst ein starker Anstieg der Individualverkehre zu verzeichnen. Zusätzlich nimmt auch der Wirtschaftsverkehr innerhalb von Städten zu. Somit wird im Allgemeinen eine starke Zunahme von städtischen Verkehren erwartet. Dabei agieren die verschiedenen Akteure auf denselben infrastrukturellen Einrichtungen, deren Entwicklung der Urbanisierung nicht in angemessenem Maße folgen kann. Daraus ergibt sich für alle Akteure eine angespannte Verkehrssituation, die somit eine besondere Herausforderung für die Versorgung der städtischen Gesellschaft und Wirtschaft darstellt. (Lehmacher 2015, S. 7)

Im Bezug auf das eben Erwähnte ist die Urbanisierung nicht allein für die zunehmende Anspannung der Verkehrssituation verantwortlich. Stattdessen müssen weitere aktuelle Entwicklungen in die Betrachtung der sich verschärfenden Verkehrssituation einbezogen werden. In diesem Zusammenhang sei darauf hingewiesen, dass die Faktoren keine Allgemeingültigkeit für alle Regionen besitzen, sondern von der spezifischen regionalen Situation abhängig sind.

Generell kann der städtische Verkehr wie angedeutet in Privatverkehre und Wirtschaftsverkehre differenziert werden. Da insbesondere städtischen Güterverkehren im Rahmen der Privat- und Wirtschaftsverkehre eine besondere Affinität zu Konzepten der City-Logistik beigemessen wird, können sie als zentraler Gegenstand der City-Logistik betrachtet werden (Oexler 2002, S. 18-23). Dies wird damit begründet, dass die Ver- und Entsorgung der Stadt im Rahmen des Güterverkehrs als substanzielle Komponente der Wertschöpfungskette einen starken Einfluss auf die Produktivität und Effizienz der logistischen Aktivitäten hat (Lehmacher 2015, S. 7-8).

In Entwicklungsländern ist zu beobachten, dass mit zunehmender Urbanisierung und gesteigertem Einkommen insbesondere der Privatverkehr in städtischen Gebieten anwächst. Der Ursprung des Privatverkehrs liegt dabei in der Befriedigung privater Bedürfnisse (Oexler 2002). Da gleichzeitig die infrastrukturelle Gesamtsituation oftmals nur auf dem gegenwärtigen Zustand beibehalten werden kann, kommt es aufgrund der zunehmenden Pkw-Anzahl im Zusammenspiel mit weiteren Faktoren zu einem signifikanten Anstieg des privaten Verkehrs auf dem urbanen Straßennetz (Lehmacher 2015, S. 7). Ein weiterer Aspekt dieser Entwicklung ist, dass die zunehmende Kaufkraft in Entwicklungsländern in Kombination mit einer Bereitschaft zum Konsumieren zu einem veränderten Konsumverhalten führt. Es steigt die Nachfrage nach Gütern des täglichen Bedarfs, wobei in diesem Rahmen allgemeiner von FMCG gesprochen werden kann, wozu bspw. sowohl die Güter des täglichen Bedarfs als auch pharmazeutische und medizinische Produkte gehören. Da innerstädtischer Raum rar und teuer ist, ist für viele Geschäfte eine tägliche Belieferung zur Minimierung der Lagerhaltungsflächen notwendig. Daraus ergibt sich ein Anstieg der urbanen Wirtschaftsverkehre (Lehmacher 2015, S. 9). Wirtschaftsverkehre werden also durch die Konzepte der täglichen Belieferung aufgrund einer zunehmenden Urbanisierung

und durch gesteigerte Bedürfnisse der Konsumenten aufgrund eines veränderten Konsumverhaltens beeinflusst.

Das Thema City-Logistik hinsichtlich der urbanen Verkehrssituation wird zusätzlich durch den E-Commerce verschärft (Kiwitt 2010, S. 98). Aufgrund des expandierenden E-Commerce-Geschäfts kommt es zu einer Zunahme des Zustellvolumens in der Geschäftsbeziehung Business-to-Customer (B2C) und somit zu einem Anstieg des Wirtschaftsverkehrs. Diese Entwicklung wird durch den Trend zu kleineren Sendungen intensiviert. Der Zunahme der Wirtschaftsverkehre im Rahmen des E-Commerce wird entgegengehalten, dass sie Privatverkehre substituieren (Lehmacher 2015, S. 25). Trotzdem wird davon ausgegangen, dass die Entwicklung des E-Commerce in einer Gesamtbetrachtung zu einer Steigerung des städtischen Verkehrs führt (Kiwitt 2010, S. 98). Denn während es zu einer Verringerung des Anlieferungsvolumens führt, kommt es gleichzeitig zu einer Erhöhung der Anlieferhäufigkeit. Als Konsequenz führt dies im Bereich der City-Logistik unter Nutzung heutiger Logistikkonzepte zu einem Anstieg des innerstädtischen Verkehrs.

Die Steigerung der unterschiedlichen Verkehre führt zu einer erhöhten Beanspruchung der vorhandenen Infrastruktur. Diese kann aufgrund der Gegebenheiten in urbanisierten Räumen oftmals nicht kongruent zu den Verkehren weiterentwickelt werden (Lehmacher 2015, S. 7). Somit kommt es zwischen den unterschiedlichen Verkehren zu einer Konkurrenzsituation auf dem vorhandenen Straßennetz. In die Betrachtung der Verkehrssituation fließen neben den ökonomischen auch die ökologischen Ziele der City-Logistik mit ein (Lehmacher 2015, S. 7, 10). Zur Entschärfung der Verkehrssituation hinsichtlich der ökonomischen und ökologischen Nachhaltigkeit wurden in verschiedenen europäischen Ländern auf administrativer Ebene unterschiedliche Maßnahmen entwickelt.

In London wurde zur Vermeidung von Staus eine Staugebühr (Congestion Charge) eingeführt. Vor Einführung dieser Gebühr konnten die Auswirkungen der zunehmenden Verkehre in Londons Stadtzentrum gut beobachtet werden. So lag die durchschnittliche Geschwindigkeit bei 15 km/h, die man bereits vor über 150 Jahren erreicht hatte. Durch Einführung der Staugebühr im Stadtzentrum von London konnte die Durchschnittsgeschwindigkeit wieder signifikant gesteigert werden. (Lehmacher 2015, S. 8)

In Deutschland wurden zur Entlastung der Innenstädte sogenannte Umweltzonen eingeführt. Zugangsberechtigt sind nur noch Fahrzeuge mit einer entsprechenden Umweltplakette. Dieses Konzept bewirkt sowohl eine ökologische Entlastung als auch eine allgemeine Entspannung der Verkehrssituation in den Innenstädten. Dadurch wird sichergestellt, dass die Ver- und Entsorgung der Stadt aufrechterhalten werden kann.

Beide Konzepte versuchen, die ökonomischen und ökologischen Zielen der City-Logistik durch eine Restriktion der Verkehre zu erreichen. Klumpp et al. (2013, S. 169) verweisen darauf, dass jedoch umfassende Konzepte zur optimalen Versorgung fehlen. Des Weiteren wird von Siedentop (2015, S. 18) angeführt, dass Effizienzvorteile, die sich aufgrund der Kompaktheit urbaner Räume ergeben, konsequenter genutzt werden müssen, um städtische Versorgungskonzepte weiterzuentwickeln.

2.2 Forschungsansätze für die urbane Versorgung

Um den Herausforderungen der City-Logistik zu begegnen, werden in der Wissenschaft verschiedene Ansätze verfolgt. Zur Untersuchung dieser Ansätze gibt es eine Vielzahl von Forschungsprojekten, die sich hinsichtlich ihrer Schwerpunkte unterscheiden. Allein im aufgesetzten Forschungsprogramm *EffizienzCluster LogistikRuhr* befassen sich sechs Projekte mit den Problemen und Aufgaben der *urbanen Versorgung*; dabei besitzen sie alle eine unterschiedliche Ausrichtung. Die Bedeutung der City-Logistik und die Diversität der Forschungsansätze lässt sich auch durch die Arbeit von Benjelloun et al. (2010) verdeutlichen. Basierend auf bisherigen Einordnungen von City-Logistik-Konzepten entwickeln Benjelloun et al. (2010, S. 6218) eine dreischichtige Taxonomie zur Einordnung von City-Logistik-Projekten. Die erste Ebene beinhaltet zunächst fünf übergeordnete Kategorien zur Einordnung der Projekte. Auf der zweiten Ebene werden die übergeordneten Kategorien in verschiedene Kriterien unterteilt. Die unterste Ebene spezifiziert letztlich die Ausprägungen der Kriterien der zweiten Ebene. (Anhang B)

Im Rahmen dieser Taxonomie werden verschiedene Kriterien betrachtet, unter anderem der Status oder die Finanzierung der einzelnen City-Logistik-Konzepte. Da die Taxonomie fünf Kategorien, 22 Kriterien und 72 Kriterienausprägungen umfasst, wäre eine vollumfängliche Betrachtung für diese Arbeit nicht sachdienlich. Zur Darstellung der aktuellen Forschungsansätze ist die Einteilung der City-Logistik-Konzepte hinsichtlich ihrer Funktionalität zielführend. Die Kategorie Funktionalität wird auf der zweiten Ebene in die folgenden fünf Kriterien unterteilt (Benjelloun et al. 2010, S. 6221):

- Regulierung
- Verkehrsverlagerung
- Verkehrstelematik (Intelligent Transportation System - ITS)
- Konsolidierung
- Kooperation

Bei der Betrachtung des Kriteriums *Funktionalität* wird ersichtlich, dass Überschneidungen zu den aufgeführten Definitionen der City-Logistik und ihrem logistischen Grundverständnis vorliegen. Es wurde bereits darauf hingewiesen, dass den Konzepten der Konsolidierung und Kooperation eine besondere Relevanz im Rahmen der City-Logistik zugesprochen wird. Durch die Funktionalitätenbeschreibung innerhalb der Taxonomie von Benjelloun et al. (2010) wird jedoch ersichtlich, dass die beiden erwähnten Konzepte nicht das gesamte Spektrum der City-Logistik abdecken, sondern darüber hinaus weitere Forschungsansätze und City-Logistik-Konzepte etabliert wurden.

Im Rahmen der Regulierung werden auf kommunaler Ebene zwei Ansätze verfolgt. Durch die Einführung einer City-Maut soll in verkehrspolitischer Hinsicht die Straßennutzung in innerstädtischen Bereichen reduziert werden. Somit können Staus vermieden und die Reisegeschwindigkeit angehoben werden. Für die Umsetzung einer City-Maut existieren verschiedene Konzepte (vgl. Sammer 2012). Sammer (2012, S. 489-490) weist darauf hin,

dass die Umsetzung einer City-Maut aber nicht an den Problemursachen ansetzt und somit nur als ein Bestandteil der Problemlösung angesehen werden kann.

Ein weiterer Ansatz ist die Zugangsregulierung und -beschränkung zu innerstädtischen Bereichen. Umgesetzt werden diese Konzepte oftmals in Form von Umweltzonen für Innenstädte. Die Einrichtung dieser Umweltzonen resultierte aus umweltpolitischen Aspekten; inzwischen werden sie aber auch als Instrument für die Verkehrspolitik in Innenstädten verwendet (Cruz und Montenon 2016, S. 555). Die Zugangsregulierung und -beschränkung setzen ebenfalls nicht an den Problemursachen an und können daher ebenso nur als ein Bestandteil innerhalb eines Lösungskonzeptes betrachtet werden.

Eine Verkehrsverlagerung kann durch unterschiedliche Konzepte erreicht werden. Ein möglicher Ansatz ist der Einsatz von Güterstraßenbahnen auf dem vorhandenen Straßennetz. Zusätzlich können auch Untergrundbahnen und das zugehörige Schienennetz für den Gütertransport genutzt werden. Verschiedene Projekte haben in diesem Zusammenhang Konzepte mit speziell auf den Gütertransport ausgerichteten Untergrundbahnen entwickelt. Arvidsson und Browne (2013) stellen verschiedene Projekte vor, die sich mit dem Thema Güterstraßenbahnen befasst haben. Oftmals kam es aufgrund von unterschiedlichen Hindernissen jedoch nicht zu einer Realisierung der entwickelten Konzepte (Arvidsson und Browne 2013).

Im *EffizienzCluster LogistikRuhr* beschäftigen sich zwei Projekte mit dem Thema Verkehrsstelematik. Bisherige Navigationssysteme sind oftmals nicht den aktuellen Anforderungen im Straßengüterverkehr gewachsen. Die Probleme sind hierbei folgende: Oftmals verursacht das Umfahren eines Staus neue Störungen auf den angebotenen Ausweichrouten, oder besondere Restriktionen für den Straßengüterverkehr sind nicht in die Navigation integriert. Im Projekt *Dynamics in Navigation* soll die Routenempfehlung optimiert werden; unter Einbeziehung der globalen Konsequenzen einer Routenempfehlung für den Verkehr soll eine Nutzung der Restleistungsfähigkeit des vorhandenen Straßennetzes erreicht werden (EffizienzCluster LogistikRuhr 2016a). Im zweiten Projekt, der *Urban Business Navigation*, ist die Integration von Restriktionen und tagesaktuellen Routeninformationen bei der Routennavigation untersucht worden (EffizienzCluster LogistikRuhr 2016b).

Projekte aus dem Bereich der urbanen Versorgung befassen sich oftmals aufgrund ihrer engen Verflechtung sowohl mit dem Thema Konsolidierung als auch mit dem Thema Kooperation im städtischen Güterverkehr. Das Projekt *Urban Retail Logistics* aus dem EffizienzCluster LogistikRuhr beschäftigt sich mit Konzepten für einen Urban Hub (GVZ), der intelligenten Filialbelieferung und geräuscharmen Nachtbelieferung in städtischen Bereichen (EffizienzCluster LogistikRuhr 2016c). Bei der Untersuchung von GVZs lag das Interesse besonders auf dem umweltschonenden Umgang mit Ressourcen innerhalb der GVZs selbst.

Ein alternativer Ansatz liegt dem EU-Projekt U-TURN zugrunde. Hierbei sollen neue kollaborative Wege für die Güterverteilung in europäischen Städten erforscht werden. Dabei werden die Anforderungen und Bedürfnisse für den Lebensmitteltransport besonders in die Betrachtung einbezogen. Neben dem Aufbau einer Kollaborationsplattform zur Unter-

stützung der koordinierenden und konsolidierenden Aktivitäten von Unternehmen wird ein Simulationsmodell entwickelt. Die Simulation soll genutzt werden, um Effizienz, Effektivität und umweltbezogene Wirkung eines logistischen Kollaborationskonzeptes abschätzen zu können. Die Ergebnisse der Simulation können genutzt werden, um verschiedene Alternativen untereinander zu vergleichen und eine Grundlage für eine Entscheidungsfindung zu schaffen.

2.3 Simulation als Analyseansatz für Probleme der City-Logistik

Logistischen Problemstellungen liegen oftmals komplexe Zusammenhänge zugrunde. Somit eignet sich das Werkzeug der Simulation besonders für die Analyse von logistischen Systemen. Dies ist auch darauf zurückzuführen, dass die Simulation eine besondere Vielfältigkeit hinsichtlich ihres Einsatzes in der Logistik bietet; so können einzelne Standorte und Relationen bis hin zu ganzen Supply-Chains und Transportnetzwerken abgebildet werden. Durch den Simulationseinsatz für logistische Fragestellungen kann eine Grundlage mit verschiedenen Kenngrößen geschaffen werden, um die Auswirkungen von Handlungsalternativen aufzuzeigen und den Prozess der Entscheidungsfindung positiv zu unterstützen. (Eley 2012, S. 5-8).

Eley (2012, S. 6) verweist darauf, dass der Einsatz von Simulation für verschiedenste Bereiche logistischer Fragestellungen geeignet ist. Auch für Problemstellungen der City-Logistik wird die Simulation angewendet (Chow et al. 2010). Dabei können unterschiedliche innerstädtische Logistikkonzepte bewertet und verglichen sowie Frachtaufkommen im urbanen Raum prognostiziert werden. Für die Probleme diverser Themenbereiche der City-Logistik wird die Simulation als Analysewerkzeug im Rahmen der Entscheidungsfindung verwendet.

Für die Darstellung logistischer Zusammenhänge gibt es grundsätzlich verschiedene Modellierungsmethoden. Unterschieden werden die Vorgehensweisen dabei in verhaltensgestützte und optimierende Methoden (Chow et al. 2010). Für die Problemstellung der City-Logistik existieren wiederum spezifische Modellierungsmethoden, die entweder einen verhaltensgestützten oder optimierenden Schwerpunkt besitzen. Chow et al. (2010) stellen dar, dass integrative Ansätze für die Modellierung der City-Logistik neue Herausforderungen mit sich bringen. Diese neue Problemstellung wird von Cleophas und Ehmke (2014) am Beispiel des E-Commerce thematisiert.

Die Branchen des E-Commerce und der *Kurier-Express-Paket-Dienste* als essenzielle Elemente der City-Logistik sind eng miteinander verzweigt. Der E-Commerce stellt dabei einen stark wachsenden Wirtschaftsbereich dar (Müller-Steinfahrt und Klaus 2014), der jedoch bei der Frei-Haus-Belieferung auf der *letzten Meile* unter hohen Effizienzverlusten leidet (Cleophas und Ehmke 2014, S. 173). Cleophas und Ehmke (2014) widmen sich dieser Problemstellung der City-Logistik. Dabei wird versucht, die bisher in der Logistik separat untersuchten Konzepte der Tourenplanung und Auftragsannahme innerhalb der City-Logistik zusammenzuführen. Hierzu wird ein iterativer Lösungsansatz erarbeitet, der

die Vorhersage und Annahme des Auftragsaufkommens sowie die Tourenplanung in einen gemeinsamen Prozess integriert.

Zur Untersuchung und anschließenden Bewertung des entwickelten Konzeptes wird eine Simulationsstudie durchgeführt. Für die Untersuchung verschiedener Vorgehensweisen werden vier Szenarien entwickelt. Zum Vergleich untereinander und zur Beurteilung der unterschiedlichen Vorgehensweisen werden die Szenarien hinsichtlich ihres wertbasierten Potenzials für die Gewinnmaximierung, des Einflusses der Prognosegüte und der Wirkung der Nachfragestruktur analysiert. Durch die Arbeit von Cleophas und Ehmke (2014) wird die Problematik der wertbasierten Auftragserfüllung adressiert und eine mögliche Lösungsstrategie dargestellt. Der iterative Lösungsansatz stellt dabei nur eine mögliche Strategie dar; auch eine parallele Integration wäre als Forschungsansatz denkbar.

Chow et al. (2010) weisen darauf hin, dass besonders Modellen für die City-Logistik eher optimierende als verhaltensgestützte Methoden zugrunde liegen. Dies wurde bereits in der Definition von Taniguchi et al. erwähnt, die bei der Beschreibung der City-Logistik vom „...process for totally optimizing the logistics and transport activities [...] in urban areas...“ (Taniguchi et al. 2001, S. 2) ausgehen. Dabei ist der Blick vor allem auf ökonomische Zielsetzungen ausgerichtet. Bei der Optimierung der City-Logistik werden sowohl der Themenbereich der Tourenplanung als auch der Standortprobleme im Rahmen der Netzwerkplanung betrachtet. Letzterem widmet sich unter anderem auch das EU-Projekt U-TURN, innerhalb dessen durch die Simulation verschiedene Logistikkonzepte für die urbane Versorgung verglichen werden sollen. Ein Schwerpunkt liegt dabei auf der Kollaboration zur Versorgung innerstädtischer Bereiche bei vorhergehender Konsolidierung von Frachtaufkommen in GVZs.

3 Grundlagen der Simulation

Im vorhergehenden Kapitel wurde eine Einordnung des Untersuchungsgegenstandes in die Thematik der City-Logistik vorgenommen. Zur Problemanalyse wird dabei oft das Werkzeug der Simulation angewandt. Im Allgemeinen kann zunächst zwischen diskreter und kontinuierlicher Simulation differenziert werden. Der diskreten Simulation liegt zugrunde, dass sich der Zustand des Simulationsmodells nur sprunghaft durch eintretende Ereignisse zu diskreten Zeitpunkten während der Simulation transformieren kann. Demgegenüber verändert sich bei der kontinuierlichen Simulation der Modellzustand stetig über die Simulationszeit. (VDI 3633: 2013)

Eine besondere Form der diskreten Simulation ist die ereignisorientierte Simulation; dabei terminieren die Ereignisse ihren Eintrittszeitpunkt selbst. Eine Änderung des Modellzustands ist somit durch das Eintreten eines Ereignisses bestimmt. Gerade im logistischen Bereich kommen häufig ereignisorientierte Simulationsmodelle zum Einsatz, so auch bspw. im EU-Projekt U-TURN oder bei der Untersuchung von Lieferaufträgen bei Cleophas und Ehmke (2014). Nachfolgend soll deshalb zunächst auf die Simulation eingegangen und zusätzlich dargestellt werden, wie das Vorgehen für Simulationsstudien der ereignisorientierten Simulation erfolgen kann. Hinsichtlich der Zielsetzung dieser Arbeit soll anschließend die Informations- und Datengewinnung im Rahmen von Simulationsvorgehensmodellen für die ereignisorientierte Simulation thematisiert werden.

3.1 Vorgehensmodelle für die ereignisdiskrete Simulation

Der Verein Deutscher Ingenieure (VDI) hat sich in seiner Richtlinie 3633 konkret mit der Durchführung von Simulationsstudien für Logistik-, Materialfluss- und Produktionssystemen beschäftigt. Dabei werden die Rahmenbedingungen für die Simulation im Allgemeinen festgelegt; dies beginnt mit der Definition sämtlicher Begriffe im Kontext der Simulation. Die Simulation selbst wird dabei wie folgt definiert (VDI 3633: 2014, S. 3):

Nachbilden eines Systems mit seinen dynamischen Prozessen in einem experimentierbaren Modell, um zu Erkenntnissen zu gelangen, die auf die Wirklichkeit übertragbar sind; insbesondere werden die Prozesse über die Zeit entwickelt.

In dieser Definition sind die drei wesentlichen Komponenten der Simulation inkludiert, nämlich die Modellbildung, das Experimentieren und die Adaption der gewonnenen Erkenntnisse auf das reale System. Innerhalb der Modellbildung erfolgt zunächst eine Abstraktion des realen Systems und anschließend eine Abbildung ebendieses in einem Mo-

dell. Dabei sollten stets die sechs Grundsätze der Modellierung berücksichtigt werden (vgl. Becker et al. 2012, S. 31-32). Das gesamte Vorgehen zur Modellbildung wurde in der wissenschaftlichen Literatur an unterschiedlichen Stellen beschrieben.

Banks et al. (1988) legen zunächst dar, dass Simulationsvorgehensmodelle sich grundsätzlich aus den folgenden fünf Phasen zusammensetzen:

- Aufgabenanalyse
- Modellformulierung
- Modellimplementierung
- Modellüberprüfung
- Modellanwendung

Die verschiedenen Vorgehensmodelle sind oftmals wesentlich detaillierter beschrieben und verwenden andere Begrifflichkeiten für die Beschreibung der einzelnen Phasen. Sie unterscheiden sich untereinander stark hinsichtlich ihres Umfangs und der ausgestalteten Komplexität. Gemeinsamkeiten und Unterschiede von Simulationsvorgehensmodellen sollen nachfolgend an drei Beispielen dargestellt werden.

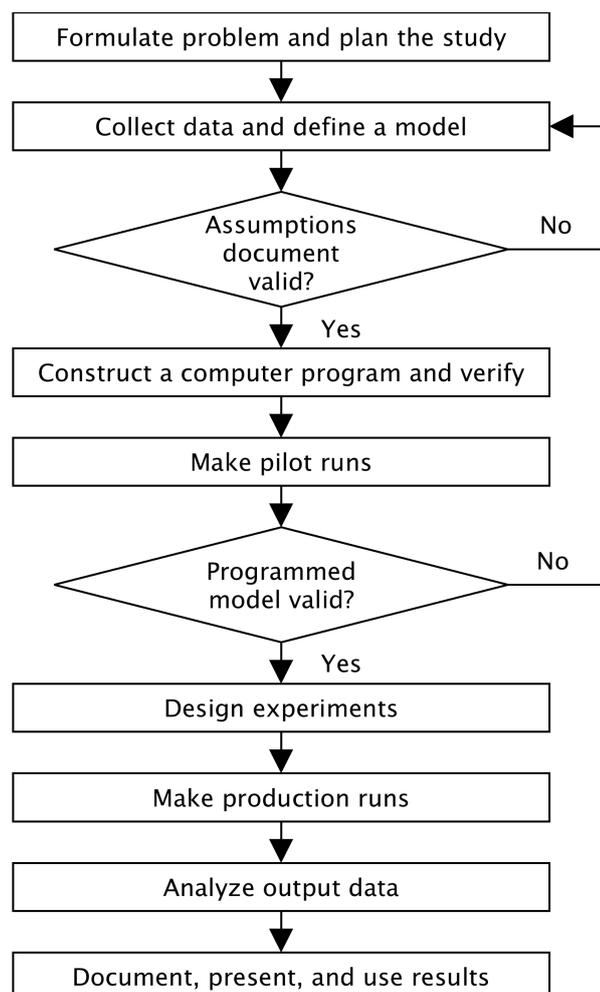


Abbildung 3.1: Simulationsvorgehensmodell (nach Law und Kelton 1991, S. 67)

Das Simulationsvorgehensmodell von Law und Kelton (1991, S. 67) in Abbildung 3.1 beschreibt ein sehr kompaktes und geradliniges Vorgehen für die Modellerstellung und Durchführung einer Simulationsstudie; die verschiedenen Elemente bauen strikt aufeinander auf. Wenn auch nicht mit denselben Begriffen lassen sich die fünf Phasen nach Banks et al. (1988) auch im Vorgehensmodell von Law wiederfinden. Die Modellüberprüfung wurde dabei durch eine Schleife innerhalb des Vorgehensmodells realisiert. Die Modellüberprüfung ist deshalb so wichtig, weil ein Modell, welches die Validierung nicht erfolgreich besteht, bei seiner Implementierung zu falschen Ergebnissen führen kann. Eine ausgegebene Handlungsempfehlung auf Basis falscher Ergebnisse birgt hohe Gefahren und kann letztendlich sehr kostspielig für ein Unternehmen sein.

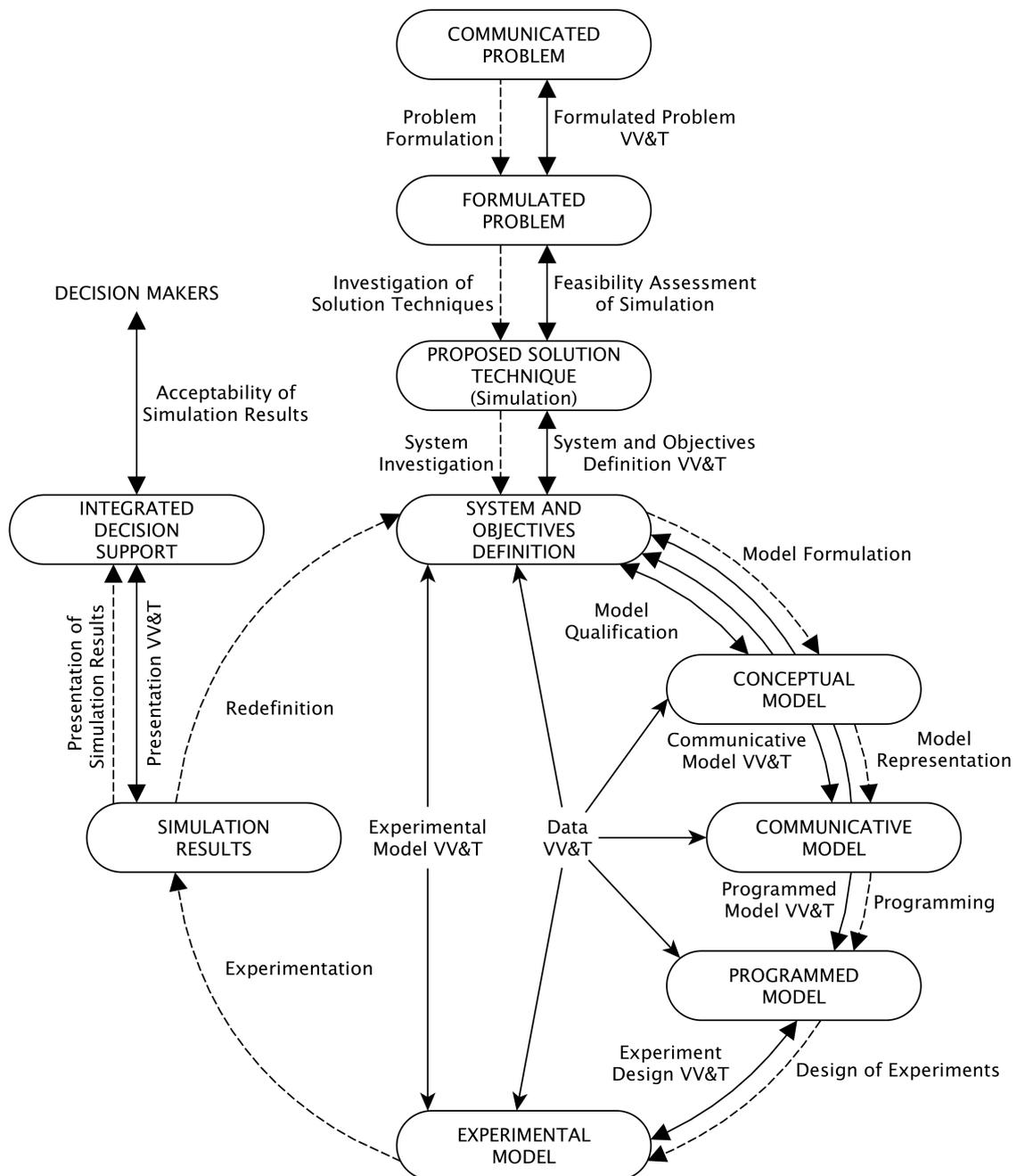


Abbildung 3.2: Simulationsvorgehensmodell (nach Balci 1998, S. 337)

Ein wesentlich umfassenderes Vorgehen wird von Balci (1998) präsentiert. Mit seinen zehn Prozessschritten weist es wesentlich mehr Bestandteile als die fünf Grundelemente nach Banks et al. (1988) auf. Es wurden verschiedene Prozessschritte in das Vorgehensmodell aufgenommen, die den Bezug zur Vor- und Nachbereitung der Simulationsstudie herstellen sowie die Nutzung der Simulationsergebnisse einbeziehen. Des Weiteren wurden verschiedene Aspekte der Verifikation & Validierung (V&V) konkretisiert und in das Simulationsvorgehensmodell integriert.

Ein grundlegender Unterschied zum Modell von Law und Kelton ist bei Balci die konkrete Einbeziehung der Daten in das Vorgehensmodell. In Laws Vorgehensmodell besteht die Einbeziehung der Daten im Wesentlichen nur aus der Datensammlung (Collect data, siehe Abbildung 3.1). Als Andeutung auf die Relevanz der Daten für das Simulationsmodell kann bei Law und Kelton (1991, S. 68) der Hinweis verstanden werden, die Daten mit dem entwickelten Simulationsmodell abzustimmen. Aufgrund der Bedeutung der Daten für die Simulationsstudie ist dieses Vorgehen nicht ausreichend. Zwar wird bei Balci die Datenerfassung nicht als konkreter Schritt in das Vorgehensmodell integriert; trotzdem werden bei Balci gegenüber Law und Kelton die Behandlung der Daten als Daten-Verifikation, Validierung und Test (VV&T) konkreter in das Simulationsvorgehensmodell eingebunden. Dazu wird die Überprüfung der für das Modell zu benutzenden Daten in den gesamten Prozess der Modellbildung eingebunden (Data VV&T). Es findet also eine durchgängige Überprüfung der Daten in Abstimmung mit dem entwickelten Modell statt. Somit soll erreicht werden, dass bei der Durchführung von Simulationsexperimenten die Ergebnisse nicht durch falsche Modelleingangsdaten verfälscht werden.

Von Rabe et al. (2007) wird ein weiteres Vorgehensmodell angeboten, welches sich auf die ereignisorientierte Simulation bezieht (Abbildung 3.3). Als Unterschied zu den Vorgehensmodellen von Law und Kelton (1991) und Balci (1998) zeichnet sich das Simulationsvorgehensmodell von Rabe et al. (2007, S. 5) besonders durch die parallele Handhabung der Modellbildung und der Datengewinnung aus. Anlass für die Trennung war, dass einerseits die Simulation signifikant abhängig ist von den verwendeten Modelleingangsdaten und andererseits der Aufwand für die Datengewinnung oftmals unterschätzt wird. Dabei kann die Bereitstellung der richtigen Daten bis zu 50 % des Projektaufwandes betragen. Darüber hinaus ist für die Autoren neben der gesonderten Handhabung von Datengewinnung und Modellbildung auch eine Überprüfung der Daten aufgrund ihrer Relevanz innerhalb des gesamten Prozesses unerlässlich. Dies bestätigt die Vorgehensweise von Balci die Daten in den gesamten Prozess des Simulationsvorgehensmodells einzubeziehen. Zusätzlich wurden bei Rabe et al. für die einzelnen Phasen im Rahmen der Modellbildung und der Datengewinnung Phasenergebnisse eingeführt. Dabei orientieren sich Rabe et al. am V&V-Vorgehensmodell von Brade (vgl. 2003, S. 62). Die Phasenergebnisse sind die Basis für eine durchgängige Überprüfung innerhalb des Simulationsvorgehensmodells. Die Datengewinnung wird dabei in zwei Phasen gegliedert (Abbildung 3.3):

- Datenbeschaffung
- Datenaufbereitung

Im Rahmen der *Datenbeschaffung* geht es zunächst einmal um die Bereitstellung der Daten für die Simulation. Art und Umfang der zu beschaffenden Daten sind bereits in der Aufgabenspezifikation und dem Konzeptmodell konkretisiert und dienen somit als Grundlage für die Datenbeschaffung. Diese erforderlichen Daten können bspw. aus Systemen der Betriebsdatenerfassung (BDE) oder Maschinendatenerfassung (MDE) extrahiert werden. Das Phaseergebnis dieses Prozessschrittes sind die *Rohdaten*, die nach Möglichkeit in unaggregierter Form für den Prozessschritt der *Datenaufbereitung* zur Verfügung gestellt werden sollen.

Im Prozess der *Datenaufbereitung* werden von Simulationsfachleuten daraufhin die zur Verfügung gestellten Rohdaten in eine Form überführt, die die Nutzung der Daten innerhalb der Simulation ermöglicht. Hierbei können die relevanten Daten für die Simulationsstudie gefiltert werden oder eine Extrahierung von Informationen aus den zur Verfügung gestellten Daten erfolgen. Das Ergebnis dieser Phase sind die *aufbereiteten Daten*.

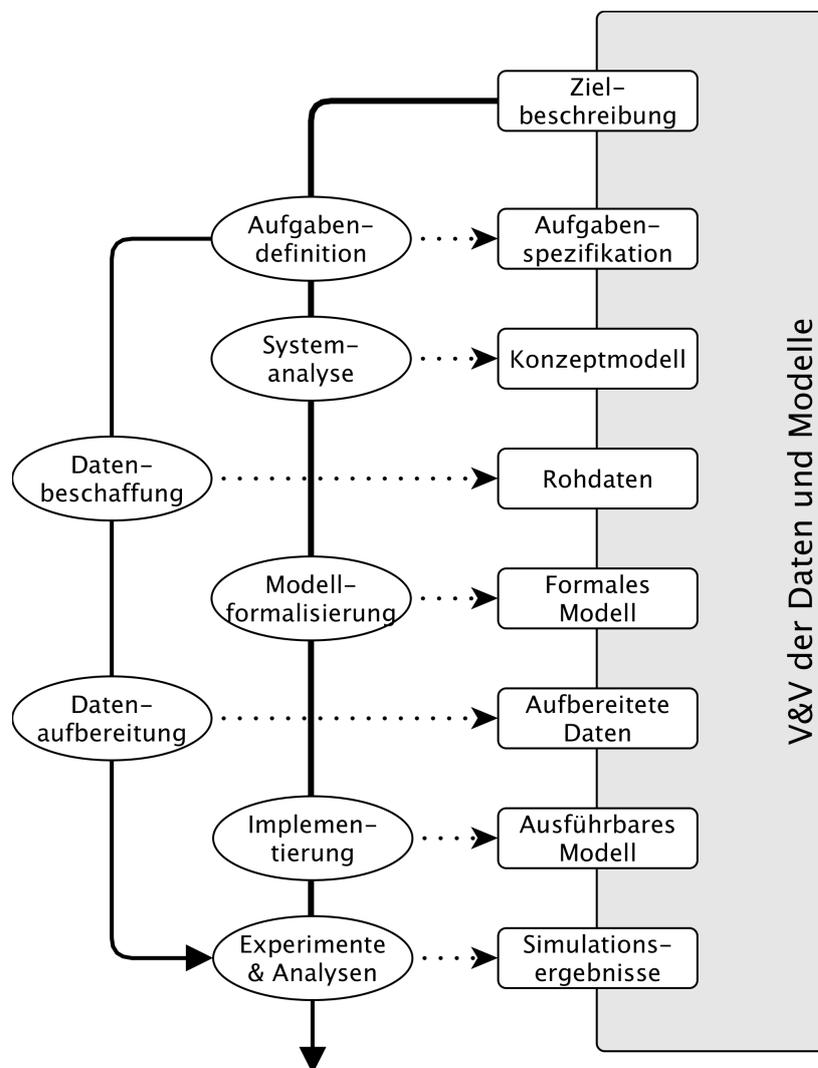


Abbildung 3.3: Simulationsvorgehensmodell (nach Rabe et al. 2007, S. 5)

Eine zeitliche Einordnung der Phasen innerhalb des Simulationsvorgehensmodells ist bei Rabe et al. (2007) nicht stringent vorgegeben, wie die Darstellung in Abbildung 3.3 vermuten lässt. Für das ausführbare Modell müssen aber zwingenderweise die *aufbereiteten Daten* als Phasenergebnis der Datenaufbereitung zur Verfügung stehen (Bernhard et al. 2007, S. 5). Aufbauend auf dem Vorgehen von Brade (2003, S. 62) zur V&V des Prozesses der Modellerstellung wird dieses Vorgehen von Rabe et al. (2007, S. 118-120) auf die Schritte der Datengewinnung in Abstimmung mit den Phasen und den Phasenergebnissen der Modellerstellung erweitert.

Durch den Vergleich der verschiedenen Simulationsvorgehensmodelle wurden die Unterschiede hinsichtlich ihrer individuellen Ausgestaltung sichtbar. Sie verfügen über eine unterschiedliche Anzahl an Prozessschritten und divergieren auch hinsichtlich des Detaillierungsgrades der einzelnen Phasen und Prozessschritte. Außerdem ist ersichtlich, dass die Datengewinnung und die Abstimmung mit dem entwickelten Simulationsmodell immer mehr an Bedeutung innerhalb des Simulationsvorgehensmodells gewinnt. Während bei Law und Kelton (1991) die Daten eher eine nebensächliche Bedeutung hatten und im Allgemeinen in den Modellbildungsprozess integriert waren, wurde bei Balci (1998) bereits stärker auf die Daten innerhalb der Simulation eingegangen. Schließlich werden sie bei Rabe et al. (2007) aufgrund ihrer Relevanz für das Simulationsmodell in einem separaten Prozess innerhalb des Simulationsvorgehensmodells behandelt.

3.2 Datenbehandlung in der Simulation

Wie bereits im vorherigen Kapitel dargestellt, gibt es verschiedene Vorgehensmodelle zur Durchführung von Simulationsstudien. Diese unterscheiden sich in ihrem Umfang und Detaillierungsgrad und ganz speziell auch in der Integration der Daten- bzw. Informationsgewinnung in das Vorgehensmodell selbst. Welche Relevanz die Daten innerhalb des Simulationsvorgehens besitzen, wurde von Rabe et al. (2007) für die ereignisdiskrete Simulation dargestellt. Darauf basierend wurde ein Simulationsvorgehensmodell entwickelt, welches die Modellbildung und Datengewinnung voneinander trennt, um die Bedeutung der Daten für die ereignisdiskrete Simulation herauszustellen. Des Weiteren fand eine Integration von Phasenergebnissen für eine durchgehende V&V in das Vorgehensmodell statt (Rabe et al. 2007, S. 119).

Die Behandlung der Daten innerhalb des Simulationsvorgehensmodells von Rabe et al. (2007, S. 5) wird in die Schritte *Datenbeschaffung* und *Datenaufbereitung* unterteilt. Dieser Ansatz wird von Jodin et al. (2009, S. 5) aufgegriffen und erweitert. Dabei werden die Schritte der Datengewinnung in den ausgedehnten Prozess der Informationsgewinnung integriert.

Es erfolgt also eine explizite Differenzierung zwischen Daten und Informationen. Das Verständnis für Daten und Informationen in dieser Arbeit orientiert sich hinsichtlich des Zusammenhangs zwischen Zeichen, Daten und Informationen an Leimeister (2015, S. 24-25). Die Basis für die Darstellung von Daten sind zunächst einmal Zeichen, die erst unter

Berücksichtigung einer definierten Syntax zu Daten werden. Daten dienen dann in Abhängigkeit des zu beschreibenden Zusammenhangs der Repräsentation und Darstellung von Informationen. Diese Informationen lassen sich aus den Daten selbst zunächst nicht schlussfolgern.

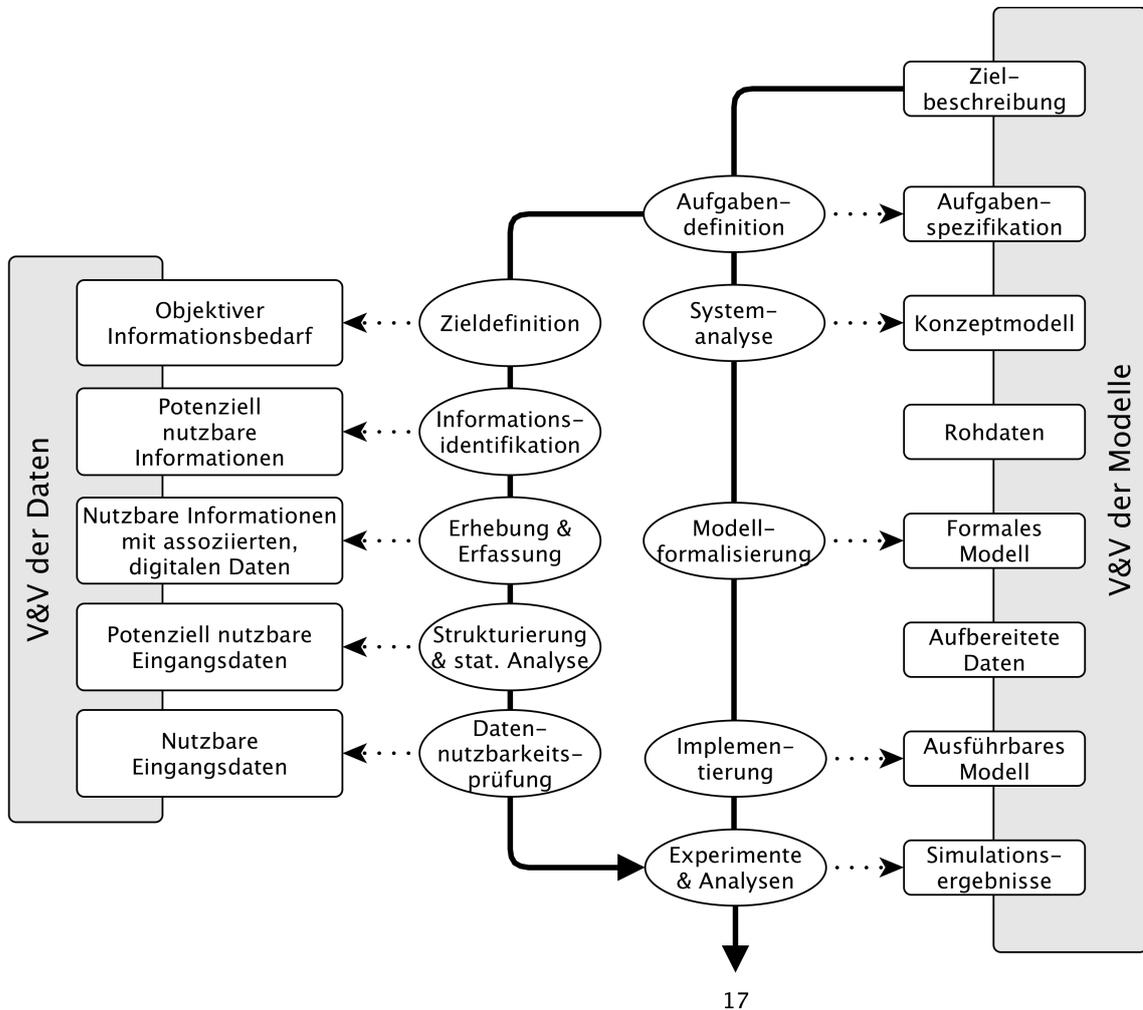


Abbildung 3.4: Simulationsvorgehensmodell (nach Jodin et al. 2009, S. 5)

Durch die Unterscheidung zwischen Daten und Informationen ist auch der Informationsgewinnungsprozess nicht nur eine einfache Ausgestaltung des Datengewinnungsprozesses. So wird durch die Informationsgewinnung das Spektrum der Aufgaben im Rahmen der Datengewinnung erweitert und weiterentwickelt. Es ist in erster Linie nicht mehr relevant, welche Daten benötigt werden, sondern wie sich der Informationsbedarf für den Gegenstand der Simulation darstellt. Aufbauend auf dieser Sichtweise wird die Datengewinnung durchgeführt.

Diese Sichtweise auf den Informationsbedarf für die Durchführung einer Simulationsstudie unterscheidet sich grundsätzlich von der integrierten Datengewinnung im Simulationsvorgehensmodell von Rabe et al. (2007). Der Prozess der Datengewinnung wird durch den neuen Prozess der Informationsgewinnung detaillierter dargestellt und erweitert. Der eigentliche Prozess der Datenbeschaffung wird ausgestaltet und gliedert sich anschließend

in die Phasen der Informations- und Datenerhebung sowie die Datenerfassung. Die Datenaufbereitung schließt eine Datenstrukturierung und eine statistische Analyse ein.

Durch die Integration der Informationsgewinnung als eigenständigen Prozess innerhalb des Simulationsvorgehensmodells soll erreicht werden, den Anforderungen an die Qualität der Modelleingangsdaten für die Simulation gerecht zu werden. Grundsätzlich ist festzuhalten, dass einerseits der Aufwand für die Informationsgewinnung oftmals unterschätzt wird und andererseits die Güte der Simulation und der Simulationsergebnisse in direkter Abhängigkeit zur Qualität der verwendeten Informationen und Eingangsdaten steht. Durch den separierten Informationsgewinnungsprozess soll diesen beiden Herausforderungen begegnet werden.

In Analogie zum Prozess der Modellbildung sind den einzelnen Prozessschritten des Informationsgewinnungsprozesses jeweils Phaseergebnisse zugeordnet (Abbildung 3.4). Entsprechend der durchgehenden V&V in der Modellbildung auf Basis der Phaseergebnisse erfolgt auch in der Informationsgewinnung anhand der Phaseergebnisse entsprechend eine kontinuierliche V&V. Somit soll die Qualität der verwendeten Informationen und Eingangsdaten für die Simulationsstudie gewährleistet werden.

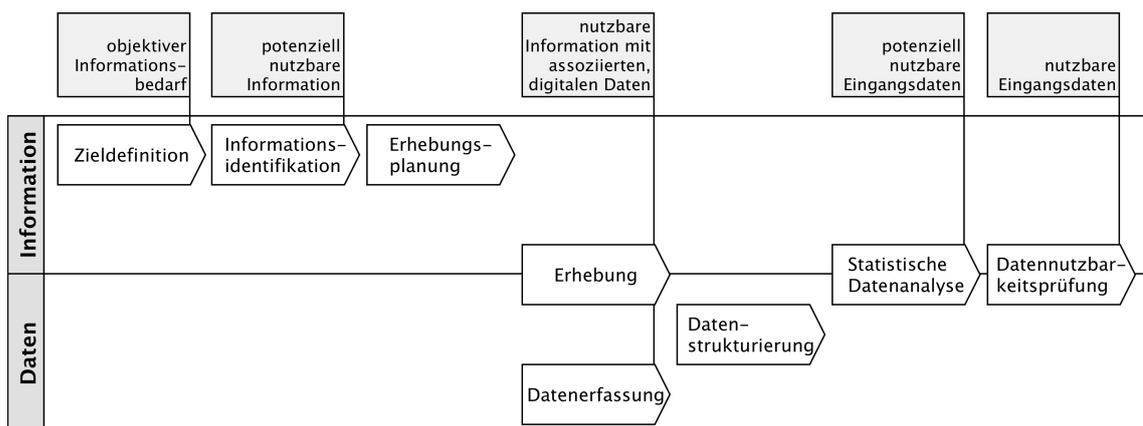


Abbildung 3.5: Ausführliche Darstellung des Informationsgewinnungsprozesses (nach Bernhard et al. 2007, S. 7)

Eine ausführlichere Darstellung des Informationsgewinnungsprozesses ist durch das prozessorientierte Vorgehensmodell von Bernhard et al. (2007, S. 7) gegeben (Abbildung 3.5). Die nachfolgende Beschreibung des Informationsgewinnungsprozesses am prozessorientierten Vorgehensmodell zeigt auf, dass im Prozess verschiedene Methoden unterschiedlicher Disziplinen angewandt werden. Die Differenzierung zwischen Informations- und Datensicht ermöglicht dabei die Darstellung des gesamten Gestaltungsrahmens beginnend mit der Zieldefinition und Informationsidentifikation über die Datenerhebung und -strukturierung bis hin zu den real nutzbaren Modelleingangsdaten.

Der Informationsgewinnungsprozess wird mit der Zieldefinition eröffnet. In Hinblick auf das Untersuchungsobjekt soll dabei zuerst der notwendige, objektive Informationsbedarf abgeleitet werden. Aufbauend auf einer Analyse der Aufgabenstellung findet eingangs eine Detaillierung ebendieser statt, sodass eine einheitliche Basis für die weiteren Aktivitäten im Informationsgewinnungsprozess geschaffen wird. Hierbei werden zunächst die zu

modellierenden Prozesse identifiziert und anschließend die Relevanz und Granularität hinsichtlich des Untersuchungsgegenstandes spezifiziert. Darauf aufbauend müssen die Ziele für die Informationsgewinnung, der objektive Informationsbedarf, definiert werden. Dies ist ein theoretischer Bedarf, der sich aus dem Prozess und dem damit zusammenhängenden Logistikdatensatz ergibt. Der Logistikdatensatz beinhaltet in standardisierter Form alle notwendigen Daten, die für die Darstellung von Logistikprozessen notwendig sind (Hömberg et al. 2007, S. 13-14).

Der Zieldefinition schließt sich die Informationsidentifikation an. Vorbereitend soll der objektive Informationsbedarf der sich aus der Zieldefinition ergab in Abstimmung mit dem subjektiven Informationsbedarf und dem Informationsangebot in die potenziell nutzbaren Informationen überführt werden. Dazu wird zunächst durch eine Informationsbedarfsanalyse basierend auf dem objektiven Informationsbedarf und unter Berücksichtigung des Logistikdatensatzes der subjektive Informationsbedarf abgeleitet. Es findet somit eine Übertragung des Logistikdatensatzes auf den abzubildenden Prozess statt. Nachfolgend wird eine Informationsangebotsanalyse durchgeführt, in der zunächst alle potenziell möglichen Informationsquellen, die die identifizierten Informationen aus der Bedarfsanalyse enthalten können, erfasst und anschließend bewertet werden. In einer Nutzbarkeitsprüfung wird dann ein Abgleich des objektiven und subjektiven Informationsbedarfs mit dem Informationsangebot vorgenommen, woraus demzufolge als Phasenergebnis der Informationsidentifikation die potenziell nutzbaren Informationen resultieren.

Der nächste Schritt im Informationsgewinnungsprozess ist die Erhebungsplanung. Sie zielt auf eine vollumfängliche Vorbereitung der Datenerhebung ab. Dazu gehört zunächst die Auswahl angemessener Informationsquellen hinsichtlich des subjektiven Informationsbedarfs und darauf aufbauend die Auswahl von Erhebungsmethoden unter Berücksichtigung eines beschränkten Mitteleinsatzes. Der Fokus der Erhebungsplanung liegt dabei auf der Sicherstellung durch Prüfung, dass die eingesetzten Methoden zu validen und verlässlichen Ergebnissen führen.

Der Erhebungsplanung schließen sich folgerichtig die Erhebung und Datenerfassung an. Damit sollen die in der Informationsidentifikation ermittelten potenziell nutzbaren Informationen mit den in der Erhebungsplanung festgelegten Methoden aus den ausgewählten Quellen extrahiert werden. Im Rahmen des Informationsgewinnungsprozesses müssen die Daten in digitaler Form erfasst werden. Falls dabei Daten in offener Form über nicht-standardisierte Fragebögen oder Expertenbefragungen erfasst wurden, müssen diese in digitale Daten überführt werden. Des Weiteren kann es notwendig sein, weitere Prozessdaten für den Untersuchungsgegenstand der Simulation zu erheben. In einer anschließenden Validierung werden sowohl die angewandten Methoden als auch die erhobenen Daten ein weiteres Mal einer Prüfung unterzogen.

Als Ergebnis der Erhebung und Datenerfassung liegen die erfassten Daten in digitaler Form vor. Diese müssen als Nächstes im Rahmen der Datenstrukturierung in eine für die Analyse verwendbare Form überführt werden. Innerhalb der Datenstrukturierung finden eine syntaktische Formatierung, eine Definition von Relationen, eine Fehlerbereinigung

und eine Anpassung der Granularität statt. Die Reihenfolge dieser Prozessschritte ist nicht obligatorisch und kann situationsabhängig angepasst werden.

Die syntaktische Formatierung harmonisiert zuerst die vorliegenden Formate und passt sie den erfassten Daten characterspezifisch an (z. B. Boolesche Variable, Integer, Datum). Nach erfolgreicher syntaktischer Formatierung werden die Datensätze zentral in einer Datenbank gespeichert. Die anschließende Definition von Relationen widmet sich der Darstellung von Zusammenhängen zwischen verschiedenen Variablen oder Tabellen. Ggf. können dabei auch Daten, die in mehreren Tabellen erfasst wurden, in eine Tabelle zusammengeführt werden. Bei der Darstellung von Zusammenhängen zwischen Daten und Tabellen muss kontextabhängiges Hintergrundwissen mitberücksichtigt werden. Bei der Fehlerbereinigung werden die Einträge hinsichtlich ihrer logischen Übereinstimmung analysiert. Mögliche als *einfach* angesehene Fehler, wie bspw. nicht-ganzzahlige Stückzahlen oder fehlende Werte, können ohne größeren Aufwand korrigiert werden. Demgegenüber stellt sich die Identifikation von Ausreißern komplizierter dar. Hier kommen Verfahren der deskriptiven Statistik zum Einsatz. Als Beispiel für die Anpassung der Granularität sei die Festlegung von Nachkommastellen erwähnt. Dabei kann der Informationsgehalt niemals erhöht werden. Den Abschluss der syntaktischen Formatierung bildet eine Plausibilitätsprüfung. Dabei wird sowohl die Vollständigkeit als auch die Konsistenz der Daten überprüft. Des Weiteren muss kontrolliert werden, ob die ermittelte Datenlage die Anforderungen an den subjektiven Informationsbedarf, der sich aus der Informationsidentifikation ergab, erfüllt.

Der nächste Schritt der Informationsgewinnung ist die statistische Analyse, um aus den strukturierten Daten Modelleingangsdaten zu entwickeln. Dazu wird zunächst eine deskriptive Analyse durchgeführt. Durch Lage- und Streuungsparameter, deren Auswahl sich am subjektiven Informationsbedarf orientiert, können verschiedene Zusammenhänge zwischen den Daten aufgezeigt werden. Somit kann ein Verständnis über wesentliche Charakteristika der erhobenen Daten geschaffen werden. Darauf aufbauend findet eine Auswahl und Anwendung von statistischen Verfahren zur Beschreibung dieser Zusammenhänge statt. Abschließend muss auch in diesem Prozessschritt eine Verifikation des Phasenergebnisses stattfinden.

Im letzten Prozessschritt der Informationsgewinnung findet eine Datennutzbarkeitsprüfung statt. Diese ist stark mit dem vorherigen Schritt der statistischen Analyse verbunden. Bei einer anstehenden Validierung der Daten wird über die allgemeine Richtigkeit der Daten hinaus überprüft, ob die aufbereiteten Daten auch den Anforderungen an den subjektiven Informationsbedarf genügen. Letztendlich muss untersucht werden, wie die Daten in das Simulationsmodell eingebunden werden können. Als Ergebnis des Informationsgewinnungsprozesses ergeben sich die nutzbaren Eingangsdaten, die für die Durchführung der Simulationsstudie verwendet werden können.

3.3 Simulation im Rahmen logistischer Fragestellungen

In der Verkehrssimulation werden zunächst einmal nicht logistische Fragestellungen behandelt, sondern die Untersuchung der Verkehrsdynamik selbst angestrebt. Der Verkehr kann dabei auf unterschiedlichen Ebenen, die sich hinsichtlich ihres Detaillierungsgrades unterscheiden, simuliert und analysiert werden. Während auf Mikroebene der Verkehr auf dem Level der einzelnen Verkehrsteilnehmer (Fahrzeuge, Fußgänger, Fahrradfahrer etc.) betrachtet wird, findet auf Makroebene eine Beschreibung des Verkehrssystemverhaltens durch wenige Variablen auf übergeordneter Ebene statt. In Makromodellen werden dann bspw. Berechnungen auf Basis von Durchschnittsgeschwindigkeiten und Verkehrsdichte durchgeführt. (Dallmeyer 2014, S. 15)

Während die Verkehrssimulation oftmals verschiedene Arten von Verkehrsteilnehmern betrachtet, konzentriert sich die Simulation im logistischen Kontext grundsätzlich auf wirtschaftliche Implikationen (Klug 2010, S. 28). Auch im Umfeld logistischer Problemstellungen wird oft das Werkzeug der Simulation angewandt. Dabei eignen sich gerade Planungsaufgaben aus dem logistischen Bereich aufgrund ihrer Komplexität für die Lösung durch Simulation. Wie bereits kurz dargestellt, verweist Eley (2012) darauf, dass sich die Simulation für verschiedenste Bereiche im logistischen Kontext eignet.

Die verschiedenen Logistikmodelle lassen sich nach Klug (2010, S. 27) entsprechend folgender Merkmale differenzieren:

- Detaillierungs- und Abstraktionsgrad (Mikro-/Meso-/Makromodelle)
- Berücksichtigung des Zeitverhaltens (statische/dynamische Modelle)
- Optimalitätsanspruch (heuristische/optimierende Verfahren)

Die Gestaltung eines Planungssystems entsprechend dieser Merkmale hängt maßgeblich von der jeweiligen Aufgabenstellung und der jeweiligen Phase im Projekt ab. Ähnlich wie für Verkehrsmodelle wird ebenso für Logistikmodelle eine Differenzierung hinsichtlich der Modellgranularität vorgenommen. Somit können entweder einzelne Bereiche innerhalb eines Standorts simuliert werden oder das Distributionsnetzwerk aus dem Zusammenspiel vieler verschiedener Standorte Gegenstand der Simulationsstudie sein.

Ein Hauptvorteil hierarchisch aufgebauter Logistikmodelle mit Makro-, Meso- und Mikroebene besteht darin, dass Modelle zunächst auf übergeordneter Ebene entwickelt und anschließend für untergeordnete Ebenen verfeinert werden. Außerdem kann über hierarchische Logistikmodelle gesteuert werden, dass Bereiche mit starkem Logistikbezug detaillierter dargestellt werden und Komplexe mit geringem oder keinem Logistikbezug ausgespart werden können. (Klug 2010, S. 28)

Nach Klug (2010, S. 28) werden auf Makroebene zunächst die Zusammenhänge bzgl. des Material- und Informationsflusses innerhalb eines Distributionsnetzwerkes abgebildet. Dabei zielt die Simulation nicht auf die Optimierung einzelner Standorte ab, sondern beschäftigt sich mit strategischen Aufgaben wie Standortplanung oder Kapazitätsplanung für einzelne Standorte.

Demgegenüber werden in Mikromodellen einzelne Komponenten des Makromodells detailliert nachgestaltet. Da dafür bereits hohe Detailkenntnis erforderlich ist, eignen sich Mikromodelle für eine spätere Verwendung in Planungsprozessen. Auf Mikroebene können Liefer- und Bestellzyklen optimal geplant oder wirtschaftliche Losgrößen und Bestandsparameter ermittelt werden. (Klug 2010, S. 28)

4 Konzepte zur Integration von Verkehrsinformationen in die Simulation

In den vorherigen Kapiteln wurde zunächst der Begriff der City-Logistik konkretisiert und aktuelle Herausforderungen und Forschungsansätze innerhalb der City-Logistik dargestellt. Da für die Lösung von Problemen der City-Logistik die Simulation häufig zum Einsatz kommt, wurde anschließend ein kurzer Überblick über die Simulation gegeben. Dabei wurde konkret auf den Informationsgewinnungsprozess innerhalb der Simulation und auf seine separate Behandlung im Simulationsvorgehensmodell eingegangen. Die separate Behandlung der Informationsgewinnung wird damit begründet, dass der Gewinnung von Eingangsdaten für die Durchführung von Simulationsstudien bisher nicht die Aufmerksamkeit zuteil kam, die aufgrund ihres Stellenwertes angemessen wäre.

In der wissenschaftlichen Literatur wird darauf verwiesen, dass zuverlässige Ergebnisse und gesicherte Übertragungen auf die Wirklichkeit aus der Simulation nur dann abgeleitet werden können, wenn eine Informations- und Datenlage für die Durchführung der Simulation in hinreichender Güte vorliegt. Jodin et al. (2009, S. 2) sprechen in diesem Kontext von der richtigen Qualität, Quantität und Granularität der Modelleingangsdaten. Diesem Anspruch gerecht zu werden, erfordert die separate Informationsgewinnung im Rahmen des Simulationsvorgehensmodells. Die Forderung nach einer zielorientierten Behandlung von Informationen und Daten im Rahmen der Simulation wird von Gabler et al. (2013, S. 34) im Besonderen für die City-Logistik adressiert. Die Gewinnung der Eingangsdaten für das Simulationsmodell zur Lösung von Problemen innerhalb der Distributionslogistik stellt dabei eine besondere Herausforderung dar. Die vorliegenden Daten sind zunächst einmal nicht geeignet, um sie in unmittelbarer Form innerhalb der Simulationsstudie zu verwenden. Dies ist darauf zurückzuführen, dass die bereits vorliegenden Daten nicht für die definierte Problemstellung erhoben wurden. Deshalb ist es notwendig, die Daten auf ihre Verwendbarkeit innerhalb der Simulationsstudie und dem entwickelten Simulationsmodell zu überprüfen und im Allgemeinen durch Erhebung weitere Daten in den vorhandenen Datenbestand zu integrieren.

Diese Arbeit beschäftigt sich im Rahmen der verkehrslogistischen Simulation mit Modelleingangsdaten für die Modellierung von Entscheidungsproblemen der Distributionslogistik. Neben der Verhaltensmodellierung und der Ableitung einer entsprechenden Nachfragestruktur liegt bei Problemen der City-Logistik die besondere Herausforderung in der Gewinnung der Eingangsdaten für die Modellierung von Distributionsnetzwerken. Dabei sollen neben der Simulation der Distanzen innerhalb des Netzwerkes selbst auch weitere Verkehrsinformationen in die Simulation einbezogen werden. Hier kann der Logistikda-

tensatz zunächst eine grobe Richtung vorgeben, welche Informationen für die Darstellung eines Prozesses relevant sind.

Nachfolgend soll somit zunächst untersucht werden, wie zum Zeitpunkt der Betrachtung die Simulation von Relationen vorgenommen wird. Die unterschiedlichen Ansätze werden dabei untereinander verglichen. Des Weiteren wird eine Literaturrecherche durchgeführt, um herauszufinden, welche Verkehrsinformationen gegenwärtig bereits in die Simulation von Distributionsnetzwerken einbezogen werden und wie dies konkret umgesetzt wird. Darauf aufbauend wird analysiert, wie sich die Erkenntnisse aus der Literaturrecherche auf die Simulation innerhalb der City-Logistik übertragen lassen. Außerdem wird konkret auf das entsprechende Vorgehen im Simulationswerkzeug *SimChain* eingegangen.

4.1 Vorgehensweisen bei der Wegstreckenberechnung

Für die Berechnung von Wegstrecken zwischen Netzwerkknoten existieren verschiedene Verfahren, die in unterschiedlichen Forschungsbereichen angewandt werden. Die Versorgungsforschung als Fachrichtung der Gesundheitssystemforschung beschäftigt sich unter anderem mit der Erreichbarkeit von Einrichtungen des Gesundheitswesens, wie bspw. Krankenhäuser oder Arztpraxen. Die Netzwerkknoten stellen dabei die einzelnen Wohnsitze der eventuellen Patienten innerhalb eines bestimmten Gebiets dar. Dabei werden die Netzwerkknoten nicht untereinander verbunden, sondern jeweils einer speziellen Einrichtung des Gesundheitswesens zugeordnet. Somit wird die Erreichbarkeit der Einrichtungen im vorher abgesteckten Gebiet analysiert.

Für die Untersuchung der Erreichbarkeit ist die Entfernung zwischen den Wohnsitzen eventueller Patienten und der jeweiligen Einrichtung des Gesundheitswesens relevant. Die Entfernung kann durch die euklidische Distanz (Luftlinie) oder die Manhattan-Distanz (absolute Differenz der Einzelkoordinaten zweier Punkte) abgeschätzt oder auf Basis genauerer Berechnung entweder über den *kürzesten Pfad* oder den *zeitabhängigen kürzesten Pfad* berechnet werden (Apparicio et al. 2008). Die euklidische Distanz (4.1) und die Manhattan-Distanz (4.2) zwischen zwei Netzwerkknoten lassen sich nach folgenden Vorschriften auf Basis ihrer kartesischen Koordinaten (x_i, y_i) und (x_j, y_j) berechnen:

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (4.1)$$

$$d_{ij} = |x_i - x_j| + |y_i - y_j| \quad (4.2)$$

Die Ermittlung des *kürzesten Pfades* und des *zeitabhängigen kürzesten Pfades* über Graphsuchalgorithmen ist demgegenüber wesentlich komplexer, da hierzu umfangreiches Kartenmaterial mit zusätzlichen Informationen über Fahrtrichtung, Tempolimits etc. für jeden Streckenabschnitt benötigt wird (Apparicio et al. 2008).

Die verschiedenen Vorgehensweisen haben je nach Anwendungsgebiet ihren jeweiligen Nutzen; jedoch ist festzuhalten, dass die euklidische Distanz zur Abschätzung der zurückzule-

genden Wegstrecke auf verkehrslogistischer Infrastruktur durch Autos, Lkws oder öffentliche Transportmittel nicht zielführend ist. Aparicio et al. (2008) untersuchen in ihrem Anwendungsfall die Erreichbarkeit von mehr als 600 Einrichtungen des Gesundheitswesens durch die Bewohner einer Stadt. Dabei wird festgestellt, dass zwischen den verschiedenen Distanzmaßen eine gewisse Korrelation besteht und bspw. die euklidische Distanz bei der Untersuchung der bloßen Erreichbarkeit von Einrichtungen des Gesundheitswesens ähnlich gute Ergebnisse liefert wie die tatsächliche Wegstrecke oder die Manhattan-Distanz. Trotzdem ergeben sich im urbanen Raum bei der Wegstreckenberechnung Differenzen, die in Abhängigkeit des gewählten Distanzmaßes lokal sehr stark variieren können.

Häufig wird im Rahmen der Netzwerkforschung die Verwendung von Methoden der Graphentheorie für den Einsatz innerhalb der logistischen Simulation diskutiert. Dabei kommen Algorithmen für die Berechnung des *kürzesten Pfades* oder des *zeitabhängigen kürzesten Pfades*, wie bspw. der Dijkstra-Algorithmus, zum Einsatz. Mit den Algorithmen der Graphentheorie können Wegstrecken entsprechend der verkehrslogistischen Infrastruktur relativ genau bestimmt werden. Der kritische Punkt bei diesen Algorithmen ist, dass sie entweder schnell sind, dann aber hohe Speicherkapazitäten benötigen, oder umgekehrt (Alanis 2014, S. 1931). Eine Approximation der Wegstrecke anstelle einer genauen Bestimmung des kürzesten Pfades innerhalb eines Distributionsnetzwerkes würde hierbei zu einer signifikanten Verkürzung der Laufzeit während der Simulation führen.

Um die aufwendige Streckenberechnung über Graphsuchalgorithmen zu umgehen und trotzdem eine ungefähre Abschätzung der zurückzulegenden Distanz zu bekommen, wird in der wissenschaftlichen Literatur immer wieder die gewichtete euklidische Distanz zwischen zwei Netzwerkknoten herangezogen. Hoerstebroek (2014, S. 74) nimmt in seiner Analyse der Elektromobilität bspw. eine Skalierung der Luftlinie um einen Faktor $g = 0,775$ vor, was ungefähr einer Gewichtung mit dem Faktor $h \approx 1,3$ entspricht. Dieses Vorgehen wird durch den Verweis darauf begründet, dass mehrere Experimente durchgeführt wurden und dieser Faktor als geeignet angesehen wurde. Die Entscheidung für eine Approximation der Wegstrecke anstelle der Verwendungen von exakteren Graphsuchalgorithmen wurde damit gerechtfertigt, dass dadurch rechenintensive Routingabfragen innerhalb der Simulation vermieden werden können.

Auch Cleophas und Ehmke (2014) verwenden bei der Analyse der wertbasierten Auftragserfüllung in ihrem Simulationsmodell dieses Vorgehen. Die euklidischen Distanzen zwischen zwei Netzknoten werden dazu mit dem Faktor $h = 1,5$ multipliziert (Cleophas und Ehmke 2014, S. 180). Damit kann nach Auskunft der Autoren eine grobe Approximation der kürzesten Pfade innerhalb der zur Verfügung stehenden Verkehrsinfrastruktur im Anwendungsfall der Stuttgarter Innenstadt erfolgen. Eine weitere Begründung für die Approximation der Wegstrecke anstelle der Anwendung genauerer Verfahren wie Graphsuchalgorithmen fehlt bei Cleophas und Ehmke (2014).

Einerseits ist die Berechnung der Strecken mit Graphsuchalgorithmen innerhalb der Simulation sehr zeitaufwendig, andererseits resultieren aus einer zu starken Vereinfachung der Abschätzung in der Wegstreckenberechnung mit der euklidischen Distanz oder der

Manhattan-Distanz zu große Abweichungen von den realen Wegstrecken. Somit stellt die gewichtete euklidische Distanz einen Mittelweg zwischen diesen beiden Ansätzen dar:

$$d_{ij}(h) = h \cdot \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (4.3)$$

Eine weitere Alternative zur simplen Approximation von Wegstrecken wird von Morris und Love (Morris und Love 1972, S. 62) angeführt. Neben der gewichteten euklidischen Distanz wird von ihnen auch die potenzierte euklidische Distanz vorgestellt, die sich über folgende Berechnungsvorschrift bestimmen lässt:

$$d_{ij}(r) = \left(\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \right)^z \quad (4.4)$$

Demgegenüber versuchen genauere Approximationsverfahren die Wegstrecken exakter abzuschätzen. Eine mögliche Form zur genaueren Abschätzung der Wegstrecke wird mit folgender Formel beschrieben (Berens und Körling 1985, S. 54):

$$d_{ij}(\alpha, p, s) = \alpha \cdot \left((x_i - x_j)^p + (y_i - y_j)^p \right)^{\frac{1}{s}} \quad (4.5)$$

Dabei stellen (x_i, y_i) und (x_j, y_j) die kartesischen Koordinaten der Netzwerkknoten und α , p und s spezifische Parameter dar. Sowohl die euklidische Distanz ($\alpha = 1$, $p = s = 2$) als auch die Manhattan-Distanz ($\alpha = 1$, $p = s = 1$) lassen sich mit dieser Formel darstellen. Darüber hinaus können über die Formel (4.5) aber auch besondere geographische und verkehrsinfrastrukturelle Charakteristika einer Region durch die individuelle Anpassung der Parameter α , p und s in die Berechnung der Wegstrecke integriert werden. Eine Optimierung der Parameter α , p und s kann dann bspw. über eine Regressionsanalyse erfolgen.

Dieser Ansatz wurde von Bertazzon et al. (2009) aufgegriffen. Zur Kalkulation von Wegstrecken im Rahmen der Versorgungsforschung wurde die Minkowski-Distanz verwendet. Die Entwicklung dieses Ansatzes basiert dabei auf der Annahme, dass die reale Wegstrecke innerhalb einer komplexen urbanen Verkehrsinfrastruktur oftmals eine Kombination aus Straßenverläufen entsprechend der euklidischen Distanz und der Manhattan-Distanz sowie kurvenförmigen Streckenabschnitten darstellt. Die Minkowski-Distanz selbst repräsentiert als Verallgemeinerung der euklidischen Distanz und der Manhattan-Distanz eine Metrik im normierten Vektorraum, die als allgemeine Distanz-Metrik verstanden werden kann. Darin sind die euklidische Distanz und die Manhattan-Distanz nur spezielle Fälle.

Die Minkowski-Distanz lässt sich zunächst über folgende Berechnungsvorschrift bestimmen (siehe Bertazzon et al. 2009):

$$d_{ab} = \left(\sum_{i=1}^n |a_i - b_i|^p \right)^{\frac{1}{p}} \quad (4.6)$$

für $A = (a_1, a_2, \dots, a_n)$ und $B = (b_1, b_2, \dots, b_n) \in \mathbb{R}^n$

Für den zweidimensionalen euklidischen Raum kann die euklidische Distanz, die Manhattan-Distanz und die Minkowski-Distanz aus (4.6) über folgende Vereinfachung berechnet werden:

$$d_{ab} = \left[(a_i - b_i)^p + (a_j - b_j)^p \right]^{\frac{1}{p}} \quad (4.7)$$

Für $[p = 2]$ in (4.7) ergibt sich die euklidische Distanz, für $[p = 1]$ die Manhattan-Distanz und für alle anderen Werte $[1 < p < 2]$ resultiert aus (4.7) ein Spektrum an Minkowski-Distanzen.

Auf Basis dieser vereinfachten Formel der Minkowski-Distanz wird von Bertazzon et al. (2009) der Parameter p optimiert, um eine bestmögliche Approximation für die Wegstreckenberechnung innerhalb einer vorhandenen Verkehrsinfrastruktur zu erreichen. Dabei wird sowohl die kürzeste Wegstrecke als auch die kürzeste Fahrtzeit betrachtet. Die abgeleiteten Werte für p sind dabei sehr spezifisch für eine Stadt und deren charakteristische Verkehrsinfrastruktur. Dies wird bspw. am Anwendungsfall der Stadt Calgary, Kanada, bei Apparicio et al. (2008) deutlich, bei dem die Stadt durch einen Fluss geteilt ist.

Diese Idee wird auch von Brimberg et al. (2007) aufgenommen und erweitert. Sie weisen darauf hin, dass die Parameter für die Approximation von Wegstrecken basierend auf der euklidischen Distanz oder der Minkowski-Distanz im wesentlichen Maße von den infrastrukturellen Besonderheiten eines Distributionsnetzwerkes abhängig sind. Dabei wird die gewichtete Minkowski-Distanz zugrunde gelegt:

$$d_{ab} = k \cdot \left[(a_i - b_i)^p + (a_j - b_j)^p \right]^{\frac{1}{p}} \quad (4.8)$$

Die gewichtete Minkowski-Distanz nach Brimberg et al. (2007) integriert einen gewichtenden Faktor in die Minkowski-Distanz. Somit soll erreicht werden, geographische und infrastrukturelle Charakteristika einer Region besser abbilden zu können.

4.2 Konzepte zur Integration weiterer Verkehrsinformationen im Rahmen der Fahrtzeitberechnung

Neben der Approximation der Wegstrecke zwischen zwei Netzwerkknoten ist bei der Analyse von Distributionsnetzwerken die Simulation der Fahrtzeit von zentraler Bedeutung. In den vorgestellten Ansätzen zur Abschätzung der Wegstrecken zwischen zwei Netzwerkknoten werden bisher keine weiteren Verkehrsinformationen in die Berechnung einbezogen. Einzig Bertazzon et al. (2009, S. 5) gehen bei der Optimierung der Minkowski-Distanz auch auf die Fahrtzeit ein. Dafür entwickeln sie einen spezifischen Wert für den Parameter p , um im Anwendungsfall bestmöglich die Fahrtzeit abzuschätzen. Dabei wird von einer Durchschnittsgeschwindigkeit von 60 km/h ausgegangen. Darüber hinaus werden keine weiteren Verkehrsinformationen, wie bspw. Staus zu stark frequentierten Zeiten, wetterbedingte oder klimatische Einflüsse, Lichtsignalanlagen etc., in die Berechnung integriert. Bei Ber-

tazon et al. (2009) ergibt sich aus dem Untersuchungsgegenstand diese Anforderung nun einmal nicht, da die zeitliche Verteilung der Fahrten nicht determiniert werden kann. Dies ist auch damit begründet, dass es sich beim Untersuchungsgegenstand nicht um ein Distributionsnetzwerk im Rahmen logistischer Fragestellung handelt.

Im logistischen Kontext scheint die Simulation der Fahrzeit auf Basis einer zugrunde gelegten Durchschnittsgeschwindigkeit durchaus üblich zu sein (Vidalakis et al. 2011). Dabei wird für die Simulation der Fahrzeiten die Notwendigkeit einer Berechnung der Fahrzeit, abhängig von der zeitlichen Verteilung über den Tag, von vielen Autoren hervorgehoben (Ehmke et al. 2012, S. 195). Eine Fahrzeitberechnung, die auf einer vorgegebenen Durchschnittsgeschwindigkeit basiert, scheint demnach nicht mehr hinreichend genau für die Simulation logistischer Fragestellungen zu sein. Dies ist darauf zurückzuführen, dass der Verkehr von vielen Faktoren beeinflusst wird; dazu zählen unter anderem Stau oder variierende Geschwindigkeiten zwischen Tages- und Nachtzeit. Außerdem sind Abweichungen von der Durchschnittsgeschwindigkeit in Abhängigkeit von Wochentagen, Jahreszeiten oder Wetterbedingungen zu erwarten (Eglese et al. 2010). Eine Integration von Verkehrsinformationen in die Simulation von Fahrzeiten innerhalb eines Distributionsnetzwerkes ist demnach unbedingt erforderlich.

Der E-Commerce als zentraler Bestandteil der City-Logistik ist mit der Problematik konfrontiert, dass Lieferaufträge besonders auf der Belieferung der *letzten Meile* großem Kostendruck ausgesetzt sind. Die Belieferung der *letzten Meile* ist dabei ein wichtiger Bestandteil der Lieferkette; jedoch kommt es gerade in diesem Bereich oftmals zu einer unwirtschaftlichen Realisierung des Lieferverkehrs (Gevaers et al. 2011). Eine optimierte Planung für die Lieferverkehre innerhalb der City-Logistik unter Einbeziehung von Verkehrsinformationen könnte die Realisierung der Transporte auf der *letzten Meile* verbessern und somit wirtschaftlicher gestalten.

In der wissenschaftlichen Literatur wurde Konzepten zur Integration von Verkehrsinformationen in die Simulation von Distributionsnetzwerken kaum Aufmerksamkeit geschenkt. Einige Autoren haben sich aber mit dieser Problemstellung im Rahmen der Routenplanung in Gebieten unterschiedlicher geographischer Ausbreitung beschäftigt:

- van Woensel et al. (2008) setzen sich mit dem Transport als wichtige Komponente bezüglich der Wettbewerbsfähigkeit von Supply-Chains auseinander. Zur Berechnung zeitabhängiger Fahrzeiten entwickeln sie einen Ansatz auf Basis der Warteschlangentheorie. Dabei konzentrieren sie sich auf interstädtische Verkehre in Distributionsnetzwerken mit großer Ausdehnung.
- Eglese et al. (2010) haben eine Heuristik auf Basis einer Tabu-Suche zur Minimierung der zeitabhängigen Fahrzeit entwickelt. Die Planung wird hierzu auf eine Lkw-Flotte eines Elektro-Großhändlers übertragen. Im Vergleich zu konventionellen Methoden ohne Einbeziehung zeitabhängiger Fahrzeiten kann mit diesem Vorgehen am Anwendungsbeispiel dargestellt werden, dass die Integration zeitabhängiger Fahrzeiten in die Routenplanung zu einer CO₂-Reduktion von bis zu 7% führen kann.

- Kritzinger et al. (2012) adaptieren den Dijkstra-Algorithmus zur Kalkulation zeitabhängiger Fahrzeiten. Zusätzlich konzentrieren sich Kritzinger et al. auf die Umsetzung einer effizienten Datenstruktur, um die Laufzeit bei der Routenplanung gering zu halten. Die Ergebnisse werden auf den Anwendungsfall der Stadt Wien übertragen und beziehen sich konkret auf die City-Logistik.
- Ehmke et al. (2012) haben einen Data-Mining-Ansatz in die Berechnung der zeitabhängigen Fahrzeit integriert. Auf Basis von GPS-Daten einer Taxi-Flotte werden auf der ersten Aggregationsebene zunächst für die einzelnen Streckenabschnitte eines Straßennetzes die durchschnittlich erreichten Geschwindigkeiten ermittelt. Darauf aufbauend werden auf der zweiten Aggregationsebene Cluster über die Streckenabschnitte gebildet. Diese Cluster orientieren sich an der Variation der tagesbezogenen Durchschnittsgeschwindigkeiten auf den einzelnen Streckenabschnitten. Hierbei kommt ein Data-Mining-Ansatz zur Anwendung, um die Menge der Daten handhaben zu können. Als Ergebnis werden Korrekturfaktoren auf Tages- und Stundenbasis ermittelt, die anschließend in die Berechnung der Fahrzeit einbezogen werden. Das entwickelte Konzept bezieht sich konkret auf die Berechnung zeitabhängiger Fahrzeiten für innerstädtische Verkehre und wird an drei Szenarien in der Stadt Stuttgart analysiert. Cleophas und Ehmke (2014) integrieren diesen Ansatz daraufhin im Rahmen einer Fragestellung des E-Commerce in der City-Logistik in die Simulation.

Die vier beschriebenen Ansätze analysieren die zeitabhängige Fahrzeit in unterschiedlichen geographischen Gebieten und mit abweichenden Streckenlängen. Bei van Woensel et al. (2008) und Eglese et al. (2010) beziehen sich die Fallstudien bspw. auf interstädtische Verbindungen eines Distributionsnetzwerkes. Demgegenüber wird von Kritzinger et al. (2012) und Ehmke et al. (2012) die Berechnung der zeitabhängigen Fahrzeit im Kontext der City-Logistik analysiert. Dabei wird von Ehmke et al. bei der Berechnung der zeitabhängigen Fahrzeit sowohl die Abhängigkeit vom Wochentag als auch von der Tageszeit adressiert.

In die Berechnung der zeitabhängigen Fahrzeit fließen bei Ehmke et al. auf Basis der GPS-Daten von Taxi-Flotten zunächst nur tages- und nachtzeitabhängige sowie stauabhängige Beeinflussungen der Fahrzeit ein. Auch externe Einflüsse wie Tempolimits werden hierbei indirekt erfasst und berücksichtigt. Eine konkrete Betrachtung von Witterungsverhältnissen oder klimatischen Rahmenbedingungen im Straßenverkehr als beeinflussende Faktoren der zeitabhängigen Fahrzeit findet hier nicht separat statt. Der Einfluss von Wetterlagen und klimatischen Bedingungen auf die zeitabhängige Fahrzeit ist aber unbestritten (Lam et al. 2008).

Lam et al. (2008) heben hervor, dass die zeitabhängige Fahrzeit auch durch Faktoren wie das Wetter beeinflusst wird. In ihrer Untersuchung konzentrieren sich Lam et al. auf die Beeinflussung durch Niederschläge. Sie stellen dar, dass die Routenplanung bisher oftmals ausschließlich auf Basis von Durchschnittsgeschwindigkeiten erfolgt. Um die Beeinflussung durch Niederschläge bei der Routenplanung zu berücksichtigen, wurde ein Modell entwickelt, das abhängig von den erwarteten Witterungsverhältnissen Pufferzeiten für die zeit-

abhängigen Fahrzeiten zur rechtzeitigen Belieferung in die Routenplanung integriert. Zur Ermittlung dieser Pufferzeiten dient die Wettervorhersage für den darauffolgenden Tag.

Neben der dargestellten Integration von Witterungsverhältnissen bei der Routenplanung sind weitere Belege für die Berücksichtigung des Wetters in der Verkehrssimulation noch verstärkt im Bereich der Mikrosimulation zu finden (Dallmeyer 2014, S. 205). Da sich die Mikrosimulation auf die einzelnen Verkehrsteilnehmer selbst konzentriert, ist dieser Ansatz innerhalb der makroskopischen Verkehrssimulation im Rahmen der City-Logistik für die Aufgabenstellung dieser Arbeit nicht zielführend.

4.3 Vorgehen im Simulationswerkzeug *SimChain*

Das Simulationswerkzeug *SimChain* wurde im Projekt e-SAVE verwendet. Hierbei wurden verschiedene Gestaltungsszenarien einer Supply-Chain durch die Simulation hinsichtlich ihrer Energieeffizienz analysiert. Die Wegstrecken zwischen den einzelnen Netzwerkknoten der Supply-Chain wurden dabei wie folgt abgeschätzt:

1. Schritt: Berechnung der Luftlinie zwischen zwei Netzwerkknoten auf Basis ihrer Geokoordinaten
2. Schritt: Gewichtung der Luftlinie mit dem Faktor $g = 1, 3$

Die Berechnung der Wegstrecke erfolgt also auf Basis der gewichteten euklidischen Distanz. Das im Anschluss an e-SAVE realisierte EU-Projekt U-TURN greift ebenfalls auf das Simulationswerkzeug *SimChain* zurück. Hierbei werden aber nicht mehr Supply-Chains untersucht, sondern innerstädtische Problemstellungen analysiert. Innerhalb der City-Logistik sollen mit *SimChain* verschiedene Szenarien für unterschiedliche Belieferungskonzepte analysiert werden. Auch hier erfolgt die Berechnung der Wegstrecken im Distributionsnetzwerk auf Basis der Luftlinie zwischen zwei Netzwerkknoten und einer anschließenden Gewichtung ebendieser mit dem Faktor $g = 1, 3$.

Für die Kalkulation der Fahrzeit wird innerhalb des Simulationswerkzeugs *SimChain* zunächst die ermittelte Wegstrecke zugrunde gelegt. Die Fahrzeitberechnung erfolgt dann über folgende Formel auf Basis einer Durchschnittsgeschwindigkeit:

$$tt_{ab} = \frac{d_{ab}}{v} \quad (4.9)$$

Für die Berechnung der Fahrzeit wurde im Projekt e-SAVE innerhalb des Simulationswerkzeugs *SimChain* für den Anwendungsfall eine konstante Durchschnittsgeschwindigkeit von $v = 60 \text{ km/h}$ angenommen. Bisher wurde das Simulationswerkzeug *SimChain* mit gleichen Parametern im EU-Projekt U-TURN eingesetzt. Besonders vor dem Hintergrund der City-Logistik muss dieses Vorgehen kritisch hinterfragt werden. Gerade in Innenstädten müssen deutlich geringere Durchschnittsgeschwindigkeiten angenommen werden. So konnten bspw. für einige europäische Metropolen folgende Durchschnittsgeschwindigkeiten ermittelt werden (Statista 2008):

- London: 19 km/h
- Berlin: 24 km/h
- Paris: 31 km/h
- München: 32 km/h

Der Vergleich zeigt auf, dass die Durchschnittsgeschwindigkeiten einerseits wesentlich geringer sind als die Durchschnittsgeschwindigkeit, die in *SimChain* bisher verwendet wurde. Andererseits wird auch deutlich, dass die Durchschnittsgeschwindigkeiten zwischen den europäischen Metropolen stark variieren.

4.4 Identifikation der Forschungsbedarfe

Durch die Literaturrecherche konnte gezeigt werden, wie die Wegstreckenberechnung bisher erfolgt und welche Schwachstellen existieren. Es ist festzuhalten, dass bei der Approximation der Wegstrecke derzeit häufig nicht die speziellen infrastrukturellen Charakteristika des Distributionsnetzwerkes in die Betrachtung einbezogen werden. Dieses Vorgehen wird in der wissenschaftlichen Literatur kritisch betrachtet. So weisen Brimberg et al. (2007) darauf hin, dass die Parameter für die Approximation von Wegstrecken basierend auf der euklidischen Distanz oder der Minkowski-Distanz im wesentlichen Maße von den infrastrukturellen Besonderheiten eines Distributionsnetzwerkes abhängig sind. Eine Verwendung gleicher Werte für unterschiedliche Systeme kann somit zu falschen Ergebnissen in der Simulation führen. Dabei wird von Cleophas und Ehmke (2014) aufgezeigt, dass gerade im E-Commerce die innerstädtischen Transporte unter hohen Effizienzverlusten leiden; somit ergibt sich die Anforderung, gerade diese innerhalb der Simulation möglichst gut abzubilden, um valide Simulationsergebnisse zu erzeugen.

In der Literatur existieren folgende Ansätze:

- Minkowski-Distanz
- gewichtete euklidische Distanz
- potenzierte euklidische Distanz

Darüber hinaus existieren weitere Approximationsmethoden unter Einbeziehung gewichteter Faktoren oder der unterschiedlicher Variation von Parametern in den aufgeführten Metriken (S. 29).

Neben der Abschätzung der Wegstrecke wurde die Relevanz der Fahrzeitberechnung für die Simulation herausgestellt. Die Literaturrecherche hat aufgezeigt, dass die Fahrzeitberechnung momentan häufig nur auf Basis einer zugrunde gelegten Durchschnittsgeschwindigkeit erfolgt. Zusätzlich wird diese zugrunde gelegte Durchschnittsgeschwindigkeit auch für verschiedene Distributionsnetzwerke verwendet. Diverse Autoren verweisen darauf, dass die Fahrzeit aber über den Tag und in Abhängigkeit des Wochentages stark variieren kann. Alternative Ansätze versuchen diese Schwankungen in die Fahrzeitberechnung einzubeziehen.

Außerdem variiert die Durchschnittsgeschwindigkeit in Abhängigkeit des Distributionsnetzwerkes. Für weite Strecken in großen Netzwerken, deren Relationen über Schnellstraßen oder Autobahnen abgebildet werden können, kann möglicherweise eine höhere Durchschnittsgeschwindigkeit angenommen werden als für ein Distributionsnetzwerk im Bereich des E-Commerce oder im Allgemeinen der City-Logistik, wo oftmals kurze Relationen bedient werden und restriktive Geschwindigkeitsbegrenzungen existieren. Gerade vor dem Hintergrund unwirtschaftlicher Transporte scheint eine exaktere Fahrzeitberechnung im Rahmen der City-Logistik von zentraler Bedeutung zu sein.

5 Analyse der verschiedenen Vorgehensweisen zur Integration von Verkehrsinformationen

In Kapitel 4 konnte zunächst dargestellt werden, wie Verkehrsinformationen derzeit in die Simulation einbezogen werden. Darüber hinaus konnte herausgearbeitet werden, welche weiteren Einflüsse sich auf Wegstrecke und Fahrzeit auswirken. Nachfolgend soll deshalb sowohl auf die Approximation von Wegstrecken als auch auf die Berechnung einer zeitabhängigen Fahrzeit eingegangen werden.

Als Basis für die Berechnung der Fahrzeit wird die zurückzulegende Wegstrecke benötigt. Aus diesem Grund werden zunächst die unterschiedlichen Distanz-Metriken untersucht. Das Ziel dieser Untersuchung ist herauszufinden, welche Distanz-Metrik am geeignetsten ist, die Wegstrecke bestmöglich abzuschätzen. Diese Vorgehensweise wird auch dem bisherigen Vorgehen im Simulationswerkzeug *SimChain* gegenübergestellt.

Als zweiter Schritt wird sich der zeitabhängigen Fahrzeit gewidmet. Dabei wird aufbauend auf den gewonnenen Erkenntnissen aus Abschnitt 4.4 ein neues Konzept zur Berechnung der zeitabhängigen Fahrzeit auf Basis approximierter Strecken entworfen. Es wird versucht, neben verkehrstechnischen Einflüssen auch klimaabhängige Faktoren in die Berechnung einzubeziehen.

Mit beiden Schritten kann eine bessere Approximation der Wegstrecke und eine genauere Berechnung der Fahrzeit in Distributionsnetzwerken stattfinden. Dies trägt dazu bei, die Ergebnisse der Simulation auf Basis geeigneterer Modelleingangsdaten zu verbessern.

5.1 Das Attika als Untersuchungsgebiet für den Anwendungsfall

Das Untersuchungsgebiet dieser Arbeit erstreckt sich über die Metropolregion Athen, das Attika, welches Athen als bevölkerungsreichste Stadt Griechenlands umfasst und das geographisch größte und bevölkerungsreichste Ballungsgebiet in Griechenland ist. Athen erstreckt sich über ein Gebiet von 39 km^2 und umfasst eine Bevölkerung von 665 000 Menschen. Demgegenüber umfasst das Attika eine Fläche, die mit 3000 km^2 ca. 75-mal so groß wie Athen ist und von rund 5,5-mal so vielen Menschen wie Athen bevölkert wird (3 750 000).

Für die nachfolgende Untersuchung werden unterschiedliche Informationen und Daten verwendet, wofür von verschiedenen Logistikdienstleistern aus Griechenland umfangreiche

Datensätze über ihre jeweiligen Distributionsnetzwerke zur Verfügung gestellt wurden. Die Logistikdienstleister sind dabei in diversen Bereichen tätig, wie bspw. der Warenauslieferung im Rahmen des E-Commerce oder der Belieferung von Einzelhandelsgeschäften. Neben den Transportaufträgen und den einzelnen Auslieferungstouren wurden von Seiten der Logistikdienstleister auch spezifische Daten über die eingesetzten Fahrzeuge erfasst.

Des Weiteren wurde vom Lehrstuhl IT in Produktion und Logistik (ITPL) eine Konvertierungsdatei bereitgestellt, um PLZs in Geokoordinaten zu überführen. Dies ist notwendig, da in den Daten der Auslieferungstouren der griechischen Logistikdienstleister keine genauen Adressen der jeweiligen Belieferungsorte angegeben sind, sondern lediglich die PLZs für die entsprechenden Gebiete, in denen sich die Belieferungsorte befinden.

Außerdem wurde auf das Straßenverkehrsnetz der Metropolregion Athen über den freizugänglichen Kartendienst *OpenStreetMap (OSM)* zurückgegriffen. Somit konnten die echt gefahrenen Distanzen zwischen den jeweiligen Geokoordinaten als Vergleichswerte für die nachfolgende Untersuchung näherungsweise ermittelt werden.



Abbildung 5.1: Darstellung des Untersuchungsgebietes

Eine Extrahierung der relevanten Auslieferungstouren für die Metropolregion Athen aus den zur Verfügung gestellten Liefernetzwerkdaten kann auf Basis der griechischen PLZs geschehen. Das Attika umfasst zunächst die PLZs von 10XXX bis 19XXX. Da das zum Attika gehörende Gebiet mit den PLZs 18XXX mit dem restlichen Teil des Attikas geographisch unverbunden ist, wird dieses von der weiteren Betrachtung innerhalb der Abschätzung von Wegstrecken ausgeschlossen. Des Weiteren werden keine Inseln der Metropolregion Athen in die Analyse mit aufgenommen, da der Fokus der Untersuchung

auf der Approximation der Wegstrecken im Straßenverkehr liegt. Das Untersuchungsgebiet und die PLZ-Gebiete sind zur Veranschaulichung nochmals in Abbildung 5.1 dargestellt.

5.2 Untersuchung der Distanz-Metriken zur Approximation der Wegstrecke

Wie in Abschnitt 4.1 dargestellt, existieren verschiedene Metriken zur Abschätzung von Wegstrecken. Für einige Bereiche eines Gebietes scheinen die euklidische oder die Manhattan-Distanz eine gute Approximation der Wegstrecke zu liefern. Dies ist darauf zurückzuführen, dass dem Untersuchungsgebiet ein bestimmtes infrastrukturelles Muster zugrunde liegt. Für Bereiche, die allerdings von diesem Muster abweichen, gerade bei historisch gewachsenen Städten, führt diese Vorgehensweise möglicherweise zu starken Fehlern in der Approximation.

Ein Vorteil der euklidischen oder der Manhattan-Distanz liegt in ihren einfachen Anwendungen; demgegenüber sind genauere Verfahren, bspw. auf Basis von Suchalgorithmen, wesentlich aufwendiger in der Umsetzung. Durch die Literaturrecherche konnte herausgestellt werden, dass weitere einfach zu berechnende Verfahren existieren, um Wegstrecken abzuschätzen (Abschnitt 4.1). Dazu zählen bspw. die gewichtete euklidische Distanz oder die Minkowski-Distanz. Da euklidische oder Manhattan-Distanz bei der Approximation von Wegstrecken ungenaue Ergebnisse liefern, soll untersucht werden, ob sich andere Methoden besser dazu eignen, die Wegstrecke zu approximieren.

Dazu wurden zunächst 6 779 unterschiedliche Verbindungen innerhalb der Liefernetzwerke der griechischen Logistikdienstleister untersucht. Jeder ausgewählte Netzwerkknoten ist dabei über den Längen- und Breitengrad determiniert. Da die Netzwerkknoten nur über ihre PLZ definiert sind, wurde die Angabe des Längen- und Breitengrades für ein PLZ-Gebiet über den geographischen Mittelpunkt als geometrischer Flächenschwerpunkt umgesetzt.

Die weiteren Berechnungen erfolgten auf Basis des geodätischen Referenzsystems *World Geodetic System 1984 (WGS 84)*, welches als Ausgangsbasis für die Berechnung von Positionierungen auf der Erde verwendet wird und dessen Referenzellipsoid die Erde nicht als Kugel zugrunde legt, sondern eine Abplattung von den Polen her einbezieht. Außerdem wurden Angaben zum Längen- und Breitengrad, die bisher in Graden angegeben waren, in Bogenmaße überführt.

Bei der Bestimmung der zurückzulegenden Strecken im Verkehrsnetz der Metropolregion Athen wurde auf den frei nutzbaren Routing-Dienst des Geographischen Instituts der Universität Heidelberg zurückgegriffen. Dieser Routing-Dienst benutzt das Application Programming Interface (API) des frei verwendbaren Online-Routingplaners OpenRoute-Service (ORS), der wiederum das Kartenmaterial des frei nutzbaren Online-Kartendienstes OSM verwendet. Die Abfrage erfolgt dabei über eine URL, in der Angaben für die Routenberechnung durch verschiedene Parameter spezifiziert werden können.

Mit den Parametern *start* und *end* werden die Start- und Endpunkte einer Verbindung mit Längen- und Breitengraden angegeben. Da keine speziellen Informationen über den verwendeten Fuhrpark vorliegen, wurde außerdem für die Routenfindung der Parameter *routepref* mit *Car* konkretisiert. Somit wird eine Routenplanung für Pkws durchgeführt. Da darüber hinaus bisher keine fahrzeitbeeinflussenden Faktoren in die Betrachtung mit einbezogen wurden, wird zusätzlich der Parameter *weighting* mit *shortest* angegeben, sodass für die Verbindung zwischen den beiden Netzwerkknoten die kürzeste Strecke auf Basis der Fahrdistanz berechnet wird.

Zur automatisierten Abfrage der Routinginformationen wurde ein Shell-Skript entworfen. Somit können Start- und Endpunkte einer Verbindung automatisch aus einer Datei mit dem Format *Comma-separated values (CSV)* ausgelesen werden. Zusätzlich wird eine Datei mit dem Format *Extensible Markup Language (XML)* für die einzelnen Routen angelegt. In dieser Datei sind einerseits die einzelnen Wegpunkte der Route gespeichert, sodass diese über ein Geoinformationssystem (GIS) visualisiert werden können; andererseits ist auch die Gesamt-Fahrdistanz der Route erfasst. Über das Skript wird die Gesamt-Fahrdistanz der Route aus der XML-Datei ausgelesen und anschließend an die CSV-Datei zurückgegeben und gespeichert. Außerdem wird in der CSV-Datei über den Index erfasst, welche XML-Datei zu welchem Tabelleneintrag gehört. Das Shell-Skript, die CSV-Datei, die verwendete URL und die zu spezifizierenden Parameter sowie die ausgegebene XML-Datei sind exemplarisch in Anhang A dargestellt.

Wie in Kapitel 4.1 beschrieben, kann zur Berechnung der euklidischen, der Manhattan- und der Minkowski-Distanz Gleichung 4.7 verwendet werden:

$$d_{ab} = [(a_i - b_i)^p + (a_j - b_j)^p]^{\frac{1}{p}} \quad (4.7)$$

Dabei gibt d die Distanz zwischen den beiden Netzwerkknoten A und B an. Die Netzwerkknoten werden dabei durch die geographischen Koordinaten a_i und a_j bzw. b_i und b_j bestimmt. Der zu spezifizierende Parameter p wird für die euklidische Distanz mit dem Wert 2 und für die Manhattan-Distanz mit dem Wert 1 angegeben. Für $[1 < p < 2]$ ergibt sich ein Spektrum an Minkowski-Distanzen. Für eine bestmögliche Approximation der Wegstrecke muss folglich p im Intervall $(1, 2)$ entsprechend angepasst werden.

Für die Berechnung der gewichteten und potenzierten euklidischen Distanz werden folgende Formeln angewandt:

$$d_{ab} = g \cdot [(a_i - b_i)^2 + (a_j - b_j)^2]^{\frac{1}{2}} \quad (5.1)$$

$$d_{ab} = ([(a_i - b_i)^2 + (a_j - b_j)^2]^{\frac{1}{2}})^z \quad (5.2)$$

Auch die Parameter g und z müssen nachfolgend für eine Approximation der Wegstrecke angepasst werden. Zur Analyse der Distanz-Metriken werden in der empirischen und theoretischen Forschung zwei Kriterien angewandt (Morris und Love 1972). Mit dem Ver-

gleich der Kriterien für die Anpassung der Parameter kann der beste Wert für die Metrik bestimmt werden. Als Kriterien werden folgende zwei Formeln verwendet, die nachfolgend in Abhängigkeit des Metrik-Parameters minimiert werden sollen:

$$AD = \sum_{l=m=1}^n |d_{lm}(A, B) - A_{lm}(A, B)| \quad (5.3)$$

Das erste Kriterium (Gleichung 5.3) summiert die absoluten Abweichungen. Die absolute Abweichung berechnet sich aus dem Betrag der Differenz zwischen ermittelter und approximierter Wegstrecke. Dieses Kriterium ist deswegen geeignet, da lange Wegstrecken mit höherer Genauigkeit als kurze Wegstrecken abgeschätzt werden müssen. Dies lässt sich dadurch exemplarisch darstellen, dass eine Abweichung der approximierten Wegstrecke von 50 % zu einer tatsächlichen Wegstrecke von 1 km vergleichsweise wenig ins Gewicht fällt als zu einer tatsächlichen Wegstrecke von 100 km. Ein Vorteil dieses Kriteriums besteht darin, dass reale Abweichungen im Gegensatz zu proportionalen Abweichungen betrachtet werden.

Das zweite Kriterium (Formel Gleichung 5.4) repräsentiert die gewichtete Summe der quadrierten absoluten Abweichungen:

$$SD = \sum_{l=m=1}^n (d_{lm}(A, B) - R_{lm}(A, B))^2 / R_{lm}(A, B) \quad (5.4)$$

Im Zähler steht dabei die quadrierte absolute Abweichung zwischen tatsächlicher und approximierter Wegstrecke, im Nenner die tatsächliche Wegstrecke. Somit können Verzerrungen für lange Distanzen ausgeglichen werden; kurze Distanzen müssen mit einer ähnlichen Genauigkeit wie lange Distanzen approximiert werden. Das zweite Kriterium misst folglich die Anpassungsgüte der Distanz-Metrik auf eine grundlegend andere Art und Weise als das erste Kriterium und ist als dimensionslose Größe zu verstehen.

Mit diesen beiden Kriterien können die Parameter, mit denen die Distanz-Metriken die beste Wegstrecken-Approximation ermöglichen, bestimmt werden. Da sich die aufgestellten Kriterien in ihrer Berechnungsweise fundamental voneinander unterscheiden, ist nicht vorausgesetzt, dass beide Kriterien den gleichen Wert für den Parameter der Distanz-Metrik ermitteln.

Zusätzlich zu den beiden aufgeführten Kriterien werden anhand verschiedener Methoden der deskriptiven Statistik die Berechnungen der unterschiedlichen Distanz-Metriken angewandt. Somit können die Ergebnisse der alternativen Vorgehensweisen zur Approximation der Wegstrecke verglichen werden. (McGrew et al. 2014)

5.2.1 Berechnungen und Analyse für die Minkowski-Metrik

Für die Minkowski-Distanz ist die Auswertung des ersten Kriteriums in Abbildung 5.2 und des zweiten Kriteriums in Abbildung 5.3 dargestellt. In diesen Abbildungen ist zu erkennen, dass die Werteverläufe beider Kriterien ein lokales Minimum besitzen. Diese

sind für beide Kriterien dahingehend identisch, dass sie sich im Anwendungsfall für den gleichen p -Wert einstellen. Für p ergibt sich anhand beider Kriterien der Wert 1,15. Dieser Wert liegt nah bei 1, wobei darauf hingewiesen sei, dass mit dem Wert $[p = 1]$ über die Minkowski-Metrik die Manhattan-Distanz abgebildet wird.

Das lokale Minimum des ersten Kriteriums liegt bei $AD(p = 1,15) = 6017,5 \text{ km}$ und für das zweite Kriterium bei $SD(p = 1,15) = 995$. Im Vergleich zur Approximation der Wegstrecke über die euklidische Distanz ($p = 2$) werden große Differenzen sichtbar. Die Summe der absoluten Abweichung gegenüber der ermittelten Wegstrecke liegt für die euklidische Distanz mit einem Wert von $11980,9 \text{ km}$ fast doppelt so hoch wie für die optimierte Minkowski-Distanz. Auch für das zweite Kriterium ergibt sich ein um ein Vielfaches erhöhter Wert mit $SD(p = 2) = 2735,6$. Anhand dieser Kriterien scheint eine Abschätzung der Wegstrecke durch die Minkowski-Metrik besser geeignet zu sein als durch die euklidische Distanz.

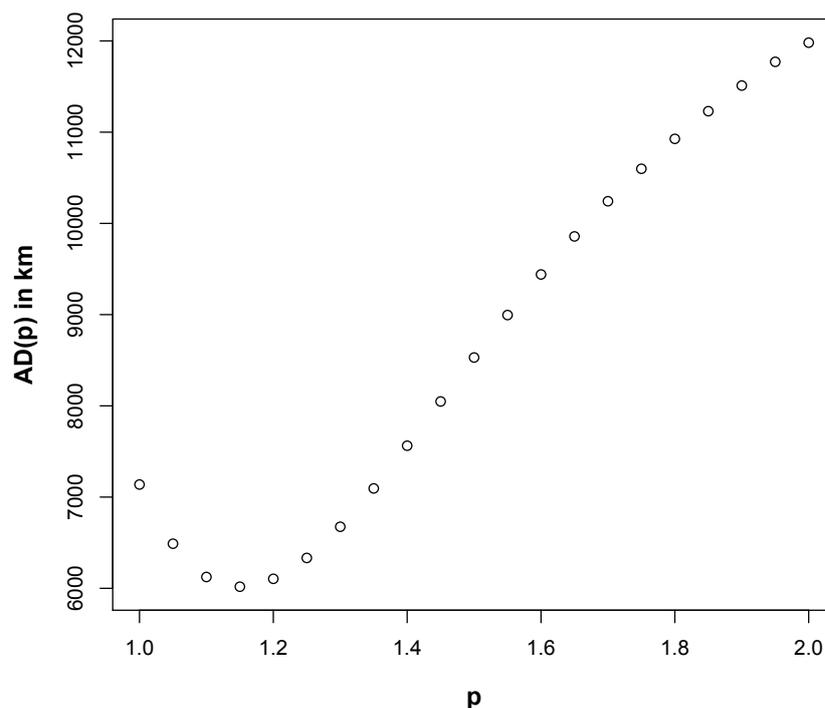


Abbildung 5.2: Auswertung des Kriteriums AD für die Minkowski-Metrik

Für die Manhattan-Distanz können das erste und zweite Kriterium mit $AD(p = 1) = 7137,8 \text{ km}$ und $SD(p = 1) = 1241,6$ angegeben werden. Im Vergleich zur Kriterienauswertung für die euklidische Distanz liegen die Werte für die Manhattan-Distanz relativ nah an den Ergebnissen der optimierten Minkowski-Distanz. Obwohl die Parameterwerte für die optimierte Minkowski-Distanz und die Manhattan-Distanz mit $[p = 1,15]$ und $[p = 1]$ nah beieinander liegen, werden beim Vergleich der Kriterien-Auswertungen trotzdem deutliche Unterschiede sichtbar. Durch die Minkowski-Distanz können die absoluten Abweichungen (erstes Kriterium) gegenüber der Manhattan-Distanz bereits um über

1 000 *km* reduziert werden. Durch die Betrachtung des zweiten Kriteriums kann außerdem festgestellt werden, dass sich über die optimierte Minkowski-Distanz sowohl kurze als auch lange Strecken ausgewogener approximieren lassen als mit der Manhattan-Distanz.

Die relative Nähe zwischen den Parameterwerten der Manhattan-Distanz und der optimierten Minkowski-Distanz lässt zunächst vermuten, dass dem Straßennetz in Teilen eher ein quadratisches Muster zugrunde liegt. Die beiden aufgestellten Kriterien versuchen, wie bereits auf S. 40 dargestellt, verschiedene Sachverhalte zu bewerten. Da sich sowohl für das erste als auch für das zweite Kriterium das lokale Minimum für den gleichen Parameterwert $[p = 1, 15]$ einstellt, deutet dies auf eine eher dispersive Struktur des Straßennetzes hin. Dadurch, dass das erste Kriterium und zweite Kriterium ihr Minimum im gleichen Parameterwert besitzen, lässt sich schließen, dass keine ausgewogenere Approximation von kurzen und langen Distanzen im Anwendungsfall realisiert werden kann.

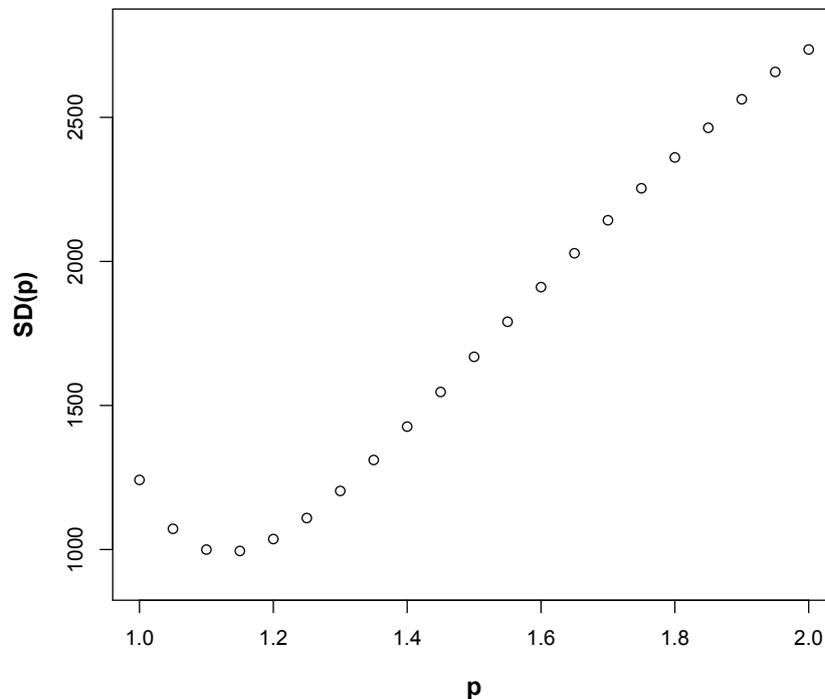


Abbildung 5.3: Auswertung des Kriteriums SD für die Minkowski-Metrik

In Tabelle 5.1 sind für die Minkowski-Distanz mit dem Parameterwert $[p = 1, 15]$ sowie für die Vergleichs-Distanzen (tatsächliche Wegstrecke, euklidische Distanz, Manhattan-Distanz) verschiedene Kennzahlen der deskriptiven Statistik aufgelistet. Die tatsächliche Wegstrecke weist zunächst einmal eine Spannweite von 0,4 *km* bis 78,3 *km* und ein arithmetisches Mittel von 8,9 *km* auf. Verglichen mit den anderen Distanzmaßen verfügt die tatsächliche Wegstrecke über das zweithöchste arithmetische Mittel und die zweithöchste Standardabweichung.

Die Standardabweichung ist im Allgemeinen sehr hoch. Dies bedeutet, dass das arithmetische Mittel selbst keine besonders hohe Repräsentativität für die Verteilung besitzt. Da

aber sowohl kurze als auch lange Relationen abgebildet werden, wird nicht erwartet, dass das arithmetische Mittel repräsentativ für die Verteilung ist. Wichtiger ist, dass die verschiedenen Distanz-Metriken ähnliche Mittelwerte und Standardabweichungen generieren, da dies einen Schluss über die grundsätzliche Verteilung zulassen würde.

Tabelle 5.1: Übersicht statistischer Kennzahlen für die Analyse der optimierten Minkowski-Metrik (in *km*; R=Spannweite, \bar{x} =arithmetisches Mittel, s=Standardabweichung)

Distanz-Metrik	Min	Max	R	\bar{x}	s
tatsächliche Wegstrecke	0,4	78,3	77,9	8,9	9,1
euklidische Distanz	0,2	59,4	59,2	7,1	7,3
Manhattan-Distanz	0,2	83,9	83,7	9,0	9,3
Minkowski-Distanz ($p=1,15$)	0,2	76,7	76,5	8,4	8,6

Mit der euklidischen Distanz findet eine durchgängig unterbewertete Approximation der Wegstrecke statt. Dies lässt sich gut an den in Tabelle 5.1 dargestellten Kennzahlen verdeutlichen. Mit einem Minimum von 0,2 *km* und einem Maximum von 59,4 *km* wird mit der euklidischen Distanz eine wesentlich kleinere Spannweite von 59,2 *km* erzeugt als bei der tatsächlichen Wegstrecke mit einer Spannweite von 77,9 *km*. Außerdem ist die euklidische Distanz durch ein wesentlich kleineres arithmetisches Mittel von 7,1 *km* und einer deutlich kleineren Standardabweichung von 7,3 *km* gegenüber der tatsächlichen Wegstrecke mit einem arithmetischem Mittel von 8,9 *km* und einer Standardabweichung von 9,1 *km* gekennzeichnet.

Demgegenüber führt die Anwendung der Manhattan-Metrik im Vergleich zur tatsächlichen Wegstrecke zu einer überbewerteten Approximation der Wegstrecke. Verglichen mit der tatsächlichen Wegstrecke wird mit der Manhattan-Distanz zwar ein kleineres Minimum von 0,2 *km* (tatsächliche Wegstrecke: 0,4 *km*), dafür aber ein erhöhtes Maximum von 83,9 *km* (tatsächliche Wegstrecke: 78,3 *km*) generiert. Dies führt wiederum zu einer erhöhten Spannweite von 83,7 *km* relativ zur Spannweite der tatsächlichen Wegstrecke von 77,9 *km*. Die Standardabweichung von 9,3 *km* ist im Vergleich zu allen anderen Distanzen die größte Ausprägung. Zusätzlich wird durch die Approximation der Wegstrecken auf Basis der Manhattan-Metrik das höchste arithmetische Mittel von 9,0 *km* erzeugt, welches relativ nah am arithmetischem Mittel der tatsächlichen Wegstrecke von 8,9 *km* liegt.

Wie bereits erwähnt, liegt die optimierte Minkowski-Distanz mit dem Parameterwert [$p = 1, 15$] in relativer Nähe zur Manhattan-Distanz, die sich über die Minkowski-Distanz mit dem Parameterwert [$p = 1$] darstellen lässt. Trotzdem können abweichend von der Manhattan-Distanz kleine Vorteile bei der Approximation der Wegstrecken auf Basis der Minkowski-Distanz mit optimierten Parameterwerten erzielt werden, da die Manhattan-Distanz die Wegstrecke überbewertete. Zunächst einmal wird durch die Minkowski-Distanz ($p = 1, 15$) eine kleinere Spannweite von 76,5 *km* erzeugt. Das arithmetische Mittel weicht mit 8,4 *km* etwas vom arithmetischem Mittel der tatsächlichen Wegstrecke ab. Bei der Betrachtung der Standardabweichung ist die Minkowski-Distanz mit einer Standardabwei-

chung von 8,6 km leicht reduziert gegenüber der Standardabweichung für die tatsächliche Wegstrecke mit 9,1 km. Bei Betrachtung von arithmetischem Mittel und Standardabweichung scheint die optimierte Minkowski-Distanz zunächst keine eindeutigen Vorteile gegenüber der Manhattan-Distanz zu generieren. Dies ist auch damit zu begründen, dass sowohl Manhattan-Distanz als auch Minkowski-Distanz bewertet am arithmetischen Mittel und an der Standardabweichung ähnliche Verteilungen wie die tatsächliche Wegstrecke erzeugt.

Da im Datensatz zur Berechnung der aufgestellten Kriterien AD und SD die Häufigkeitsverteilung über die Anzahl an Fahrten auf den einzelnen Distanzen nicht beachtet wurde, ist eine Ermittlung der Parameterwerte auf Basis der realen Abweichungen mit demselben Datensatz als zusätzliches Kriterium irreführend. Nachdem über das erste und zweite Kriterium die absoluten Abweichungen bereits in die Analyse eingeschlossen wurden, sollen nachfolgend die realen Abweichungen zwischen der ermittelten Wegstrecke und den Vergleichsdistanzen sowie der optimierten Minkowski-Distanz untersucht werden.

Tabelle 5.2: Statistische Kennzahlen für die Untersuchung der realen Abweichungen der Minkowski-Metrik (in km)

Distanz-Metrik	\bar{x}	$\tilde{x}_{0,5}$	$\tilde{x}_{0,25}$	$\tilde{x}_{0,75}$
euklidische Distanz	1,8	1,1	0,7	1,8
Manhattan-Distanz	-0,2	-0,1	-0,8	0,4
Minkowski-Distanz (p=1,15)	0,4	0,4	-0,2	0,7

In Tabelle 5.2 sind das arithmetische Mittel, der Median der realen Abweichungen sowie unteres und oberes Quantil für die verschiedenen Distanz-Metriken aufgeführt. Bei Betrachtung der Kennzahlen für die einzelnen Distanz-Metriken fällt zunächst auf, dass mit der euklidischen Distanz eine grundlegend andere Verteilung erzeugt wird als für die Manhattan- oder die optimierte Minkowski-Distanz. Offensichtlich ist, dass die Anwendung der euklidischen Distanz stets eine Unterbewertung der Wegstrecke zur Folge hat. Durch die statistischen Kennzahlen wird zusätzlich deutlich, dass durch die euklidische Distanz als Approximationsverfahren eine relativ starke Unterbewertung von durchschnittlich 1,8 km der Wegstrecke je Relation erfolgt.

Beim Vergleich der arithmetischen Mittel der verschiedenen Metriken wird deutlich, dass die Manhattan-Distanz mit einem Wert von $\bar{x} = -0,2$ dem Nullpunkt am nächsten liegt. Dagegen liegt das arithmetische Mittel für die optimierte Minkowski-Distanz mit 0,4 km doppelt so weit vom Nullpunkt entfernt. Mit der Manhattan-Distanz werden Über- und Unterbewertung während der Approximation der Wegstrecke im Vergleich zu den anderen Distanz-Metriken somit am besten ausgeglichen.

Beim Vergleich von arithmetischem Mittel und Median der Manhattan-Distanz wird jedoch ersichtlich, dass das arithmetische Mittel mit einem Wert von $-0,2$ km gegenüber dem Median von $-0,1$ km leicht verzerrt wird. Demgegenüber stimmt für die Minkowski-Distanz das arithmetische Mittel mit dem Median im Wert von 0,4 km überein und wird weniger durch starke Abweichungen beeinflusst.

Im Boxplot-Diagramm (Abbildung 5.4) werden Lage und Streuung für die realen Abweichung veranschaulicht. Zunächst ist die grundlegend unterschiedlich erzeugte Verteilung zwischen der euklidischen Distanz und der Manhattan- sowie der optimierten Minkowski-Distanz ersichtlich. Aber auch Manhattan- und Minkowski-Distanz weisen Unterschiede auf, wenn auch weniger deutlich ausgeprägt.

Bei Betrachtung des Boxplot-Diagramms wird ersichtlich, dass relativ viele Ausreißer existieren. Für die Manhattan- und die Minkowski-Distanz werden circa 10 % der Werte als Ausreißer markiert; 90 % der Abweichungen liegen also zwischen oberem und unterem Whisker. Während sich die Ausreißer nach oben und unten für die optimierte Minkowski-Distanz auszugleichen scheinen, da arithmetisches Mittel und Median mit $0,4 \text{ km}$ den gleichen Wert annehmen, werden durch die Manhattan-Distanz wesentlich mehr Ausreißer nach unten generiert. Dies bestätigt die Vermutung, dass das arithmetische Mittel für die Manhattan-Distanz unter dem Einfluss von Ausreißern vom Median abweicht.

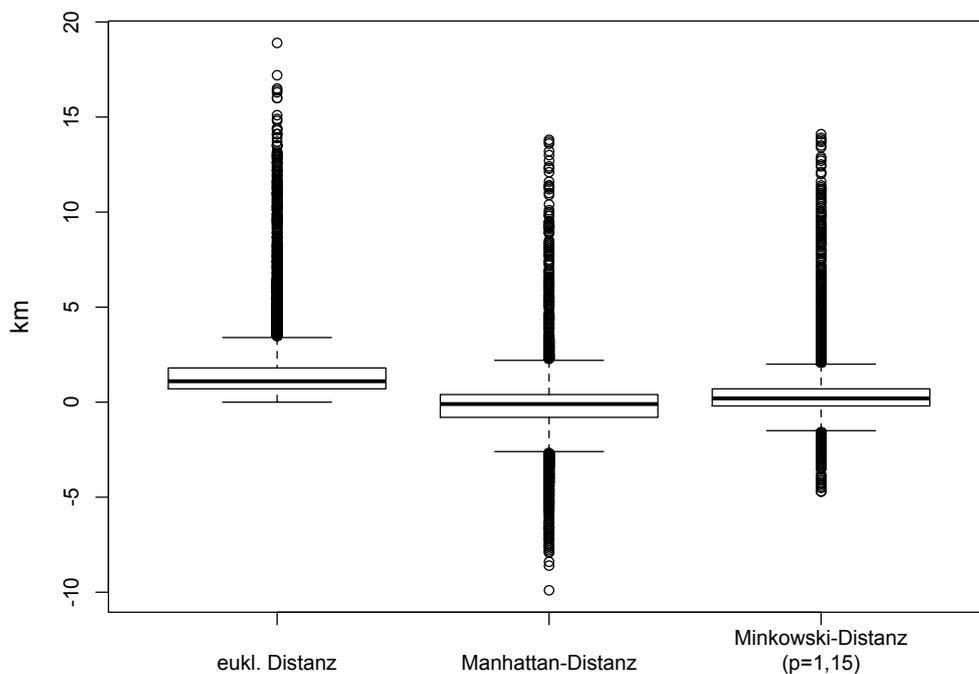


Abbildung 5.4: Boxplot der realen Abweichungen für die Analyse der Minkowski-Metrik

Für die Minkowski-Distanz liegen die mittleren 50 % der Abweichungen im Bereich von $\tilde{x}_{0,25} = -0,2 \text{ km}$ und $\tilde{x}_{0,75} = 0,7 \text{ km}$ mit einer Spannweite von $0,9 \text{ km}$. Demgegenüber generiert die Manhattan-Distanz für die mittleren 50 % der Abweichungen mit einer Spannweite von $1,2 \text{ km}$ zwischen unterem und oberem Quantil ein schlechteres Resultat. Auch die Spannweite zwischen oberem und unterem Whisker liegen für die Minkowski-Distanz mit $3,5 \text{ km}$ deutlich unter der Spannweite für die Manhattan-Distanz von $4,8 \text{ km}$.

Durch die beiden Kriterien $AD(p)$ und $SD(p)$ konnte zunächst dargestellt werden, dass herkömmliche Approximationsverfahren wie die euklidische Distanz oder die Manhattan-

Distanz ungenaue Resultate liefern. Demgegenüber kann die alternative Methode auf Basis der Minkowski-Distanz bessere Ergebnisse bei der Approximation der Wegstrecke erzielen. Sowohl das erste als auch das zweite Kriterium liefern für die optimierte Minkowski-Distanz wesentlich bessere Ergebnisse als die Vergleichsdistanzen. Ein anschließender Vergleich verschiedener Kennzahlen der deskriptiven Statistik konnte diesen Eindruck nicht bestätigen. Offensichtlich ist, dass die euklidische Distanz immer zu einer Unterbewertung der Wegstrecke führt, die Manhattan-Distanz aber ähnliche Werte bei arithmetischem Mittel und Standardabweichung im Vergleich zur tatsächlichen Wegstrecke erzeugt. Des Weiteren konnte kein Vorteil der optimierten Minkowski-Distanz gegenüber der Manhattan-Distanz herausgestellt werden. Die anschließende Untersuchung der realen Abweichungen zeigte, dass die Manhattan-Distanz mit einem Wert von $-0,2 \text{ km}$ dasjenige arithmetische Mittel generiert, welches dem Nullpunkt am nächsten liegt. Jedoch liegt der Vorteil der optimierten Minkowski-Distanz darin, dass arithmetisches Mittel und Median übereinstimmen und somit Ausreißer weniger Einfluss auf die Berechnung haben.

5.2.2 Berechnungen und Analyse für die gewichtete euklidische Distanz

Als Nächstes soll die gewichtete euklidische Distanz untersucht werden. Dabei wird die euklidische Distanz ausschließlich mit dem Faktor g gewichtet, also eine prozentuale Anpassung der euklidischen Distanz vorgenommen. Die Auswertungsergebnisse der beiden oben angeführten Kriterien AD und SD für die gewichtete euklidische Distanz sind zunächst in den Abbildungen 5.5 und 5.6 dargestellt.

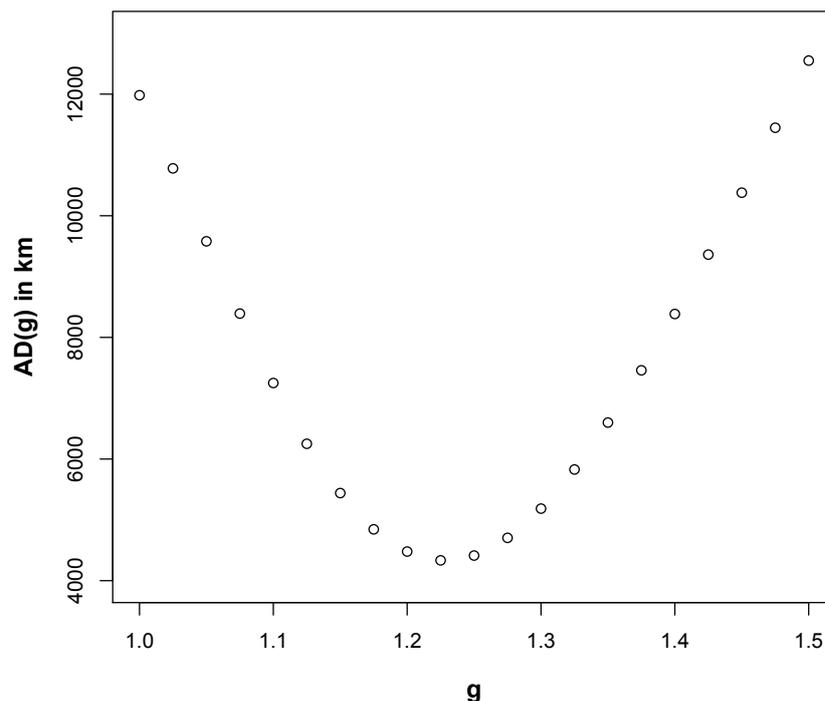


Abbildung 5.5: Auswertung des Kriteriums AD für die gewichtete euklidische Distanz

Wie bereits für die optimierte Minkowski-Distanz existieren auch für die gewichtete euklidische Distanz je ein Minimum für die Kriterien AD und SD . Für g ergibt sich anhand des ersten Kriteriums der Wert 1,225 und mittels des zweiten Kriteriums der Wert 1,25. Mit dem Wert $[g = 1]$ würde die euklidische Distanz ungewichtet bleiben. Durch die Darstellung in den Abbildungen 5.5 und 5.6 ist jedoch gut zu erkennen, dass die einfache euklidische Distanz gegenüber der optimierten gewichteten euklidischen Distanz zu wesentlich stärkeren Abweichungen führt.

Das lokale Minimum des ersten Kriteriums liegt bei $AD(g = 1,225) = 4\,334,0\text{ km}$, wohingegen sich für das zweite Kriterium das lokale Minimum bei $SD(g = 1,25) = 539,6$ einstellt. Für die euklidische Distanz können das erste und zweite Kriterium mit $AD(g = 1) = 11\,980,9\text{ km}$ und $SD(g = 1) = 2\,735,6$ angegeben werden und liegen verglichen mit der optimierten gewichteten euklidischen Distanz um ein Vielfaches erhöht. Auch im Vergleich zur Manhattan-Distanz werden mit $AD(p = 1) = 7\,137,8\text{ km}$ bzw. $SD(p = 1) = 1\,241,6$ (S. 41; S. 42) große Abweichungen sichtbar. Die gewichtete euklidische Distanz generiert gegenüber der Manhattan- und der euklidischen Distanz als Vergleichsdistanzen wesentlich niedrigere Werte.

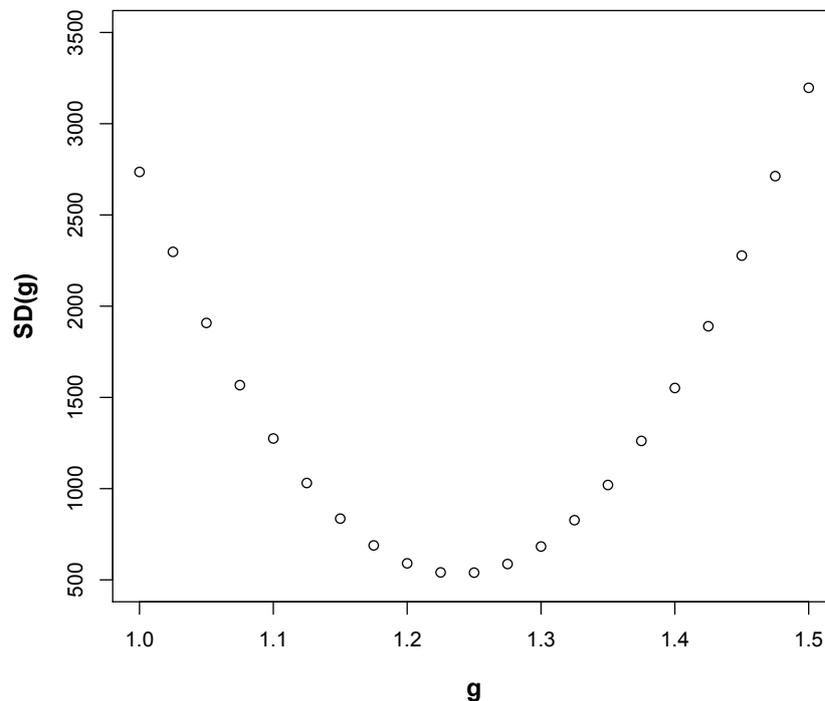


Abbildung 5.6: Auswertung des Kriteriums SD für die gewichtete euklidische Distanz

Für einen näheren Vergleich der Distanz-Metriken sind in Tabelle 5.3 verschiedene Kennzahlen der deskriptiven Statistik für die gewichteten euklidischen Distanzen ($g = 1,225$; $g = 1,25$) sowie für die zu vergleichenden Distanzen (tatsächliche Wegstrecke, euklidische Distanz, Manhattan-Distanz) aufgeführt.

Wie bereits im vorherigen Kapitel dargestellt, führt die Manhattan-Distanz zu einer Überbewertung und die euklidische Distanz zu einer Unterbewertung der Wegstrecke. Bei Betrachtung der Spannweiten R wird deutlich, dass die mit $[g = 1, 25]$ gewichtete euklidische Distanz eine Spannweite produziert, die mit einem Wert von $74,0 \text{ km}$ der Spannweite der tatsächlichen Wegstrecke von $77,9 \text{ km}$ am nächsten liegt. Trotzdem ist festzuhalten, dass sowohl für $[g = 1, 225]$ als auch für $[g = 1, 25]$ ein kleineres Minimum, ein kleineres Maximum und eine kleinere Spannweite gegenüber der tatsächlichen Wegstrecke generiert werden. Jedoch fällt beim Vergleich des arithmetischen Mittels und der Standardabweichung auf, dass die Ergebnisse der mit $[g = 1, 25]$ gewichteten euklidischen Distanz identisch mit den Ergebnissen der tatsächlichen Wegstrecke sind. Somit wird die Verteilung der tatsächlichen Wegstrecke sehr gut durch die mit $[g = 1, 25]$ gewichtete euklidische Distanz adaptiert.

Tabelle 5.3: Übersicht statistischer Kennzahlen für die Analyse der gewichteten euklidischen Distanz (in km ; R =Spannweite, \bar{x} =arithmetisches Mittel, s =Standardabweichung)

Distanz-Metrik	Min	Max	R	\bar{x}	s
tatsächliche Wegstrecke	0,4	78,3	77,9	8,9	9,1
euklidische Distanz	0,2	59,4	59,2	7,1	7,3
Manhattan-Distanz	0,2	83,9	83,7	9,0	9,3
gew. eukl. Distanz ($g=1,225$)	0,2	72,4	72,2	8,7	8,9
gew. eukl. Distanz ($g=1,25$)	0,2	74,2	74,0	8,9	9,1

Bei der Untersuchung der realen Abweichungen der verschiedenen Distanz-Metriken von der tatsächlichen Wegstrecke wird wieder auf das arithmetische Mittel, den Median als 0, 5-Quantil und das untere ($\tilde{x}_{0,25}$) sowie das obere Quantil ($\tilde{x}_{0,75}$) zurückgegriffen (Tabelle 5.4). Beim Vergleich der arithmetischen Mittel wird zunächst deutlich, dass die Werte für die Manhattan-Distanz und die mit $[g = 1, 225]$ gewichtete euklidische Distanz ähnlich weit von Null entfernt sind. Während das arithmetische Mittel $\bar{x} = -0,2 \text{ km}$ zum Median $\tilde{x}_{0,5} = -0,1 \text{ km}$ für die Manhattan-Distanz nach unten verzerrt wird, wird es für die gewichtete euklidische Distanz ($g = 1, 225$) nach oben verzerrt. Die Manhattan-Distanz generiert für die mittleren 50 % der Abweichungen eine Spannweite von $1,2 \text{ km}$, während die mit $[g = 1, 225]$ gewichtete euklidische Distanz eine halb so große Spannweite von $0,6 \text{ km}$ erzeugt. Während die mit $[g = 1, 25]$ gewichtete euklidische Distanz gegenüber der mit $[g = 1, 225]$ gewichteten euklidischen Distanz eine leicht erhöhte Spannweite von $0,7 \text{ km}$ hervorbringt, liegt das arithmetische Mittel bei $0,0 \text{ km}$ und stimmt zusätzlich mit dem Median überein. Die mit $[g = 1, 25]$ gewichtete euklidische Distanz erzeugt hier bessere Ergebnisse als die zu vergleichenden Distanzen.

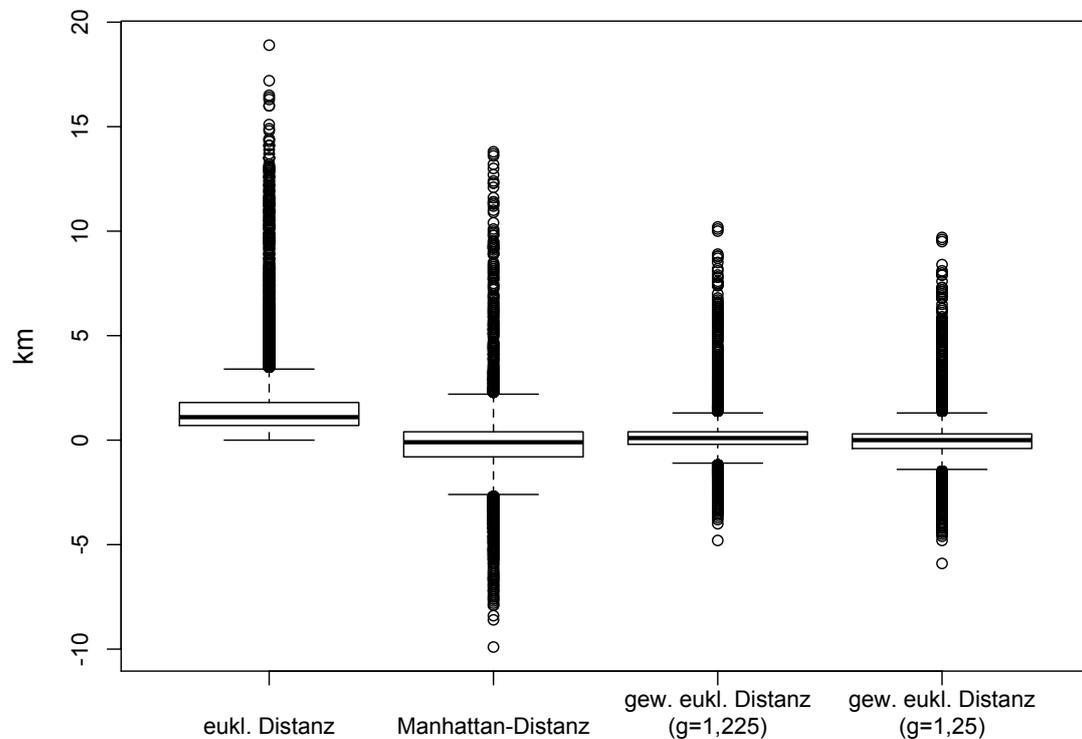
Für die Manhattan-Distanz wird zwischen oberem und unterem Whisker eine Spannweite von $4,8 \text{ km}$ aufgespannt und circa 10 % der realen Abweichungen als Ausreißer deklariert. Demgegenüber werden durch die gewichteten euklidischen Distanzen zwar wesentlich kleinere Spannweiten zwischen oberem und unterem Whisker von $2,4 \text{ km}$ ($g = 1, 225$) und

Tabelle 5.4: Statistische Kennzahlen für die Untersuchung der realen Abweichungen der gewichteten euklidischen Distanz (in *km*)

Distanz-Metrik	\bar{x}	$\tilde{x}_{0,5}$	$\tilde{x}_{0,25}$	$\tilde{x}_{0,75}$
euklidische Distanz	1,8	1,1	0,7	1,8
Manhattan-Distanz	-0,2	-0,1	-0,8	0,4
gew. eukl. Distanz ($g=1,225$)	0,2	0,1	-0,2	0,4
gew. eukl. Distanz ($g=1,25$)	0,0	0,0	-0,4	0,3

2,7 *km* ($g = 1,25$) generiert, jedoch werden für $[g = 1,25]$ mit circa 11,4 % und für $[g = 1,225]$ mit circa 12,6 % gegenüber circa 10 % für die Manhattan-Distanz in geringfügigem Maße etwas mehr der realen Abweichungen als Ausreißer gekennzeichnet.

Im Boxplot wird noch einmal die Lage und Streuung der realen Abweichungen graphisch dargestellt. Hier kann einmal mehr verdeutlicht werden, dass die gewichteten euklidischen Distanzen im Vergleich zur euklidischen und Manhattan-Distanz wesentlich kleinere Streuungen erzeugen. Die Betrachtung der Ausreißer zeigt deutlich, dass durch die Manhattan-Distanz im Vergleich zu den gewichteten euklidischen Distanzen sowohl nach oben als auch nach unten wesentlich mehr Ausreißer erzeugt werden, die besonders große Abweichungen zur tatsächlichen Wegstrecke aufweisen.

**Abbildung 5.7:** Boxplot der realen Abweichungen für die Analyse der gewichteten euklidischen Distanz

Anhand der aufgestellten Kriterien AD und SD konnte zunächst dargestellt werden, dass die gewichtete euklidische Distanz wesentlich bessere Resultate bei der Wegstreckenabschätzung erreicht als die Vergleichsdistanzen (euklidische Distanz, Manhattan-Distanz). Bei der Betrachtung des arithmetischen Mittels und der Standardabweichung konnte dieser Eindruck zunächst zumindest gegenüber der Manhattan-Distanz nicht besonders deutlich wahrgenommen werden. Hier wurden ähnliche Werte sowohl gegenüber der tatsächlichen Wegstrecke als auch gegenüber den gewichteten euklidischen Distanzen erzeugt. Der Vergleich der realen Abweichungen zeigte dann aber, dass die gewichteten euklidischen Distanzen sowohl zwischen den Quantilen als auch zwischen den Whiskern bei leicht erhöhter Ausreißeranzahl kleinere Spannweiten generieren. Zusätzlich hat die mit $[g = 1, 25]$ gewichtete euklidische Distanz ein arithmetisches Mittel von $0,0 \text{ km}$ erzeugt, das zusätzlich mit dem Median übereinstimmt. Somit scheinen sich Über- und Unterbewertung der Wegstrecke bei der mit $[g = 1, 25]$ gewichteten euklidischen Distanz auszugleichen.

5.2.3 Berechnungen und Analyse für die potenzierte euklidische Distanz

Als weitere Approximationsmethode gilt die potenzierte euklidische Distanz. Dabei wird die euklidische Distanz mit einem reellen Exponenten potenziert. Für die aufgestellten Kriterien AD und SD sind die Ergebnisse zunächst in den Abbildungen 5.8 und 5.9 dargestellt.

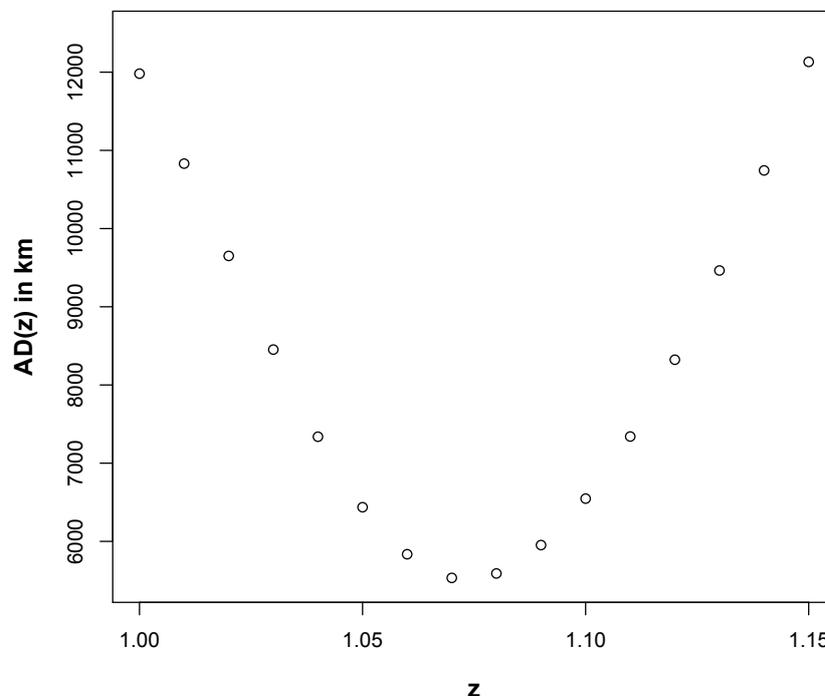


Abbildung 5.8: Auswertung des Kriteriums AD für die potenzierte euklidische Distanz

Ebenso wie für die gewichtete euklidische und die Minkowski-Distanz existiert für die potenzierte euklidische Distanz für beide Kriterien jeweils ein lokales Minimum. Das

Minimum beider Kriterien stellt sich für nah beieinander liegende Parameterwerte von $[z = 1,07]$ und $[z = 1,08]$ ein.

Für das Kriterium AD liegt das Minimum für den Parameterwert $[z = 1,08]$ bei $5\,532,9\text{ km}$ und für das Kriterium SD für den Parameterwert $[z = 1,08]$ bei $896,3$. Verglichen mit der euklidischen Distanz ($z = 1$) werden für beide Kriterien große Differenzen sichtbar. Die Werte für beide Kriterien sind für die euklidische Distanz gegenüber der potenzierten euklidischen Distanz um ein Vielfaches erhöht. Auch der Vergleich mit der Manhattan-Distanz ist eindeutig. Für beide Kriterien können mit der potenzierten euklidischen Distanz wesentlich geringere Abweichungen generiert werden als mit der Manhattan-Distanz ($AD(p = 1) = 7\,137,8\text{ km}$ (S. 41); $SD(p = 1) = 1\,241,6$ (S. 42)).

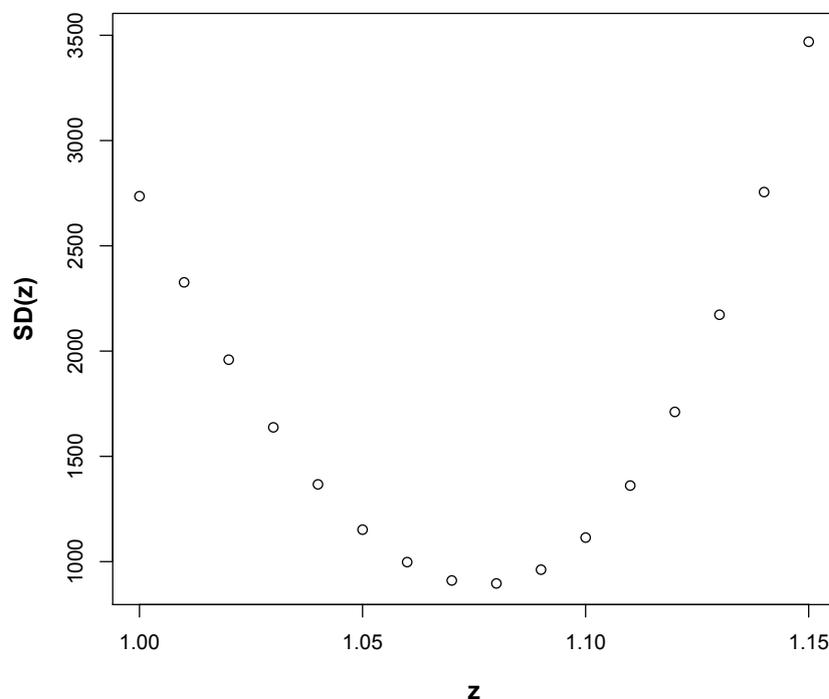


Abbildung 5.9: Auswertung des Kriteriums SD für die potenzierte euklidische Distanz

Genauso wie für die gewichtete euklidische und die Minkowski-Distanz werden nachfolgend verschiedene statistische Kennzahlen für die weitere Untersuchung der potenzierten euklidischen Distanz ($z = 1,07$, $z = 1,08$) herangezogen. Diese sind in Tabelle 5.5 dargestellt. Vergleichend dazu werden wieder die Kennzahlen für die tatsächliche Wegstrecke und die Vergleichsdistanzen (euklidische und Manhattan-Distanz) aufgeführt. Dabei produziert die potenzierte euklidische Distanz mit dem Parameterwert $[z = 1,08]$ zunächst diejenige Spannweite, welche mit $78,8\text{ km}$ der Spannweite der tatsächlichen Wegstrecke am nächsten liegt. Auch der Wert des arithmetischen Mittels liegt für die potenzierte euklidische Distanz ($z = 1,08$) relativ nah am Wert der tatsächlichen Wegstrecke. Jedoch ist die Standardabweichung gegenüber der tatsächlichen Wegstrecke mit $9,6\text{ km}$ zu $9,1\text{ km}$ leicht erhöht. Es entsteht der Eindruck, dass auch die Manhattan-Distanz zunächst ähnli-

che Werte wie die potenzierte euklidische Distanz generiert. Sowohl die Manhattan-Distanz als auch die potenzierte euklidische Distanz erzeugen eine der tatsächlichen Wegstrecke sehr ähnliche Verteilung.

Tabelle 5.5: Übersicht statistischer Kennzahlen für die Analyse der potenzierten euklidischen Distanz (in *km*; R=Spannweite, \bar{x} =arithmetisches Mittel, s=Standardabweichung)

Distanz-Metrik	Min	Max	R	\bar{x}	s
tatsächliche Wegstrecke	0,4	78,3	77,9	8,9	9,1
euklidische Distanz	0,2	59,4	59,2	7,1	7,3
Manhattan-Distanz	0,2	83,9	83,7	9,0	9,3
pot. eukl. Distanz ($z=1,07$)	0,2	79,0	78,8	8,4	9,3
pot. eukl. Distanz ($z=1,08$)	0,2	82,3	82,1	8,6	9,6

In Tabelle 5.6 sind zunächst wieder die statistischen Kennzahlen für die Untersuchung der realen Abweichungen für die jeweiligen Distanz-Metriken erfasst. Die arithmetischen Mittel für die potenzierten euklidischen Distanzen liegen mit $0,5 \text{ km}$ ($z = 1,07$) und $0,3 \text{ km}$ ($z = 1,08$) gegenüber der Manhattan-Distanz mit $-0,2 \text{ km}$ etwas weiter von Null entfernt. Während für eine Potenzierung der euklidischen Distanz mit [$z = 1,07$] das arithmetische Mittel mit dem Median im Wert von $0,5 \text{ km}$ übereinstimmt, wird sowohl für die Manhattan-Distanz als auch für die mit [$z = 1,08$] potenzierte euklidische Distanz eine leichte Verzerrung festgestellt.

Während aber für die Manhattan-Distanz zwischen oberem und unterem Quantil ein Intervall von $1,2 \text{ km}$ aufgespannt wird, erzeugen die potenzierten euklidischen Distanzen kleinere Intervalle von $0,6 \text{ km}$ ($z = 1,07$) und $0,7 \text{ km}$ ($z = 1,08$). Auch zwischen den Whiskern werden wesentlich kleinere Intervalle ausgespannt. Die Manhattan-Distanz verfügt zwischen den Whiskern über eine Spannweite von $4,8 \text{ km}$ gegenüber einer halb so großen Spannweite von $2,4 \text{ km}$ für beide potenzierten euklidischen Distanzen. Dabei werden für die Manhattan-Distanz aber mit circa 10 % etwas weniger Abweichungen als Ausreißer deklariert als für die potenzierten euklidischen Distanzen mit jeweils 12,4 %.

Tabelle 5.6: Statistische Kennzahlen für die Untersuchung der realen Abweichungen der potenzierten euklidischen Distanz (in *km*)

Distanz-Metrik	\bar{x}	$\tilde{x}_{0,5}$	$\tilde{x}_{0,25}$	$\tilde{x}_{0,75}$
euklidische Distanz	1,8	1,1	0,7	1,8
Manhattan-Distanz	-0,2	-0,1	-0,8	0,4
pot. eukl. Distanz ($z=1,07$)	0,5	0,5	0,2	0,8
pot. eukl. Distanz ($z=1,08$)	0,3	0,4	0,0	0,7

Abschließend sind Lage und Streuung der realen Abweichung in einem Boxplot dargestellt (Abbildung 5.10). Bei Betrachtung der Ausreißer ist zu erkennen, dass für beide potenzierten euklidischen Distanzen die Ausreißer gegenüber den Vergleichsdistanzen nach oben

reduziert werden. Diese Beobachtung ist in der Deutlichkeit für die Ausreißer nach unten nicht zu erkennen.

Anhand des ersten Kriteriums konnte zunächst nachgewiesen werden, dass die Abweichungen durch die potenzierte euklidische Distanz wesentlich niedriger sind als für die euklidische Distanz und in gesteigerter Form für die Manhattan-Distanz. Auch das zweite Kriterium lieferte diesbezüglich ein eindeutiges Ergebnis. Der Vergleich der verschiedenen statistischen Kennzahlen konnte zunächst keine weiteren Vorteile der potenzierten euklidischen Distanz gegenüber der Manhattan-Distanz herausstellen. Sowohl für das arithmetische Mittel als auch für die Standardabweichungen wurden ähnliche Werte wie für die tatsächliche Wegstrecke und auch für die Manhattan-Distanz erzeugt. Bei der Untersuchung der realen Abweichungen zwischen den Distanz-Metriken und der tatsächlichen Wegstrecke ergab sich, dass die potenzierte euklidische Distanz gegenüber der Manhattan-Distanz leichte Vorteile besitzt, da ein Großteil der Abweichungen in wesentlich kleineren Intervallen liegt, auch wenn dadurch geringfügig mehr Ausreißer deklariert werden als für die Manhattan-Distanz. Somit wird deutlich, dass die potenzierte euklidische Distanz sowohl für den Parameterwert $[z = 1,07]$ als auch für $[z = 1,08]$ bessere Ergebnisse generiert als die Manhattan-Distanz.

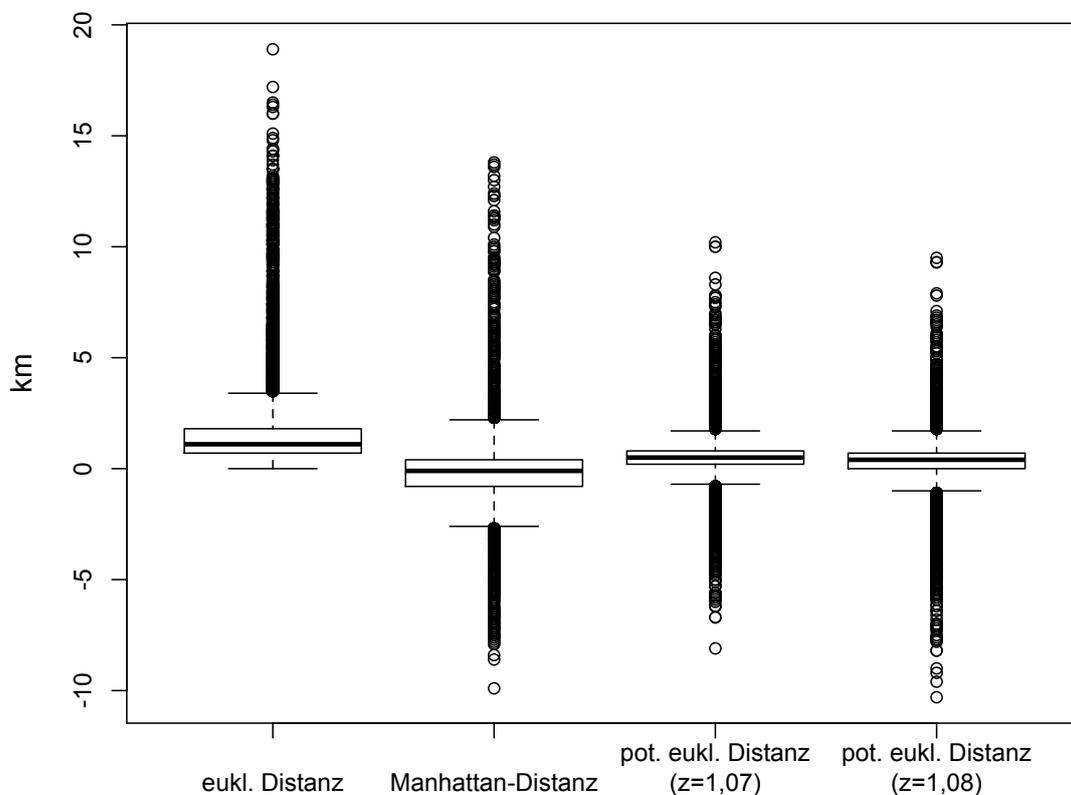


Abbildung 5.10: Boxplot der realen Abweichungen für die Analyse der potenzierten euklidischen Distanz

5.2.4 Gegenseitiger Vergleich der verschiedenen Distanz-Metriken

Die durchgeführte Analyse am Anwendungsfall der Metropolregion Athen hat zunächst gezeigt, dass Approximationsmethoden existieren, die geeignet sind, die tatsächliche Wegstrecke abzuschätzen. Vergleichsmetriken wie die euklidische Distanz und die Manhattan-Distanz wurden dabei von allen verglichenen Metriken übertroffen. Dies zeigte sich zunächst am Vergleich der beiden aufgestellten Kriterien AD und SD . Für beide Kriterien konnten die optimierten Minkowski-Distanz, die optimierte gewichtete euklidische Distanz und die optimierte potenzierte euklidische Distanz gegenüber den Vergleichsdistanzen wesentlich bessere Ergebnisse erzielen.

Die anschließende Analyse der approximierten Wegstrecken hat zunächst in weiten Teilen keine weiteren Indizien geliefert, die auf besondere Vorteile der alternativen Distanz-Metriken hindeuten. Zumindest im Vergleich zur Manhattan-Distanz wurden ähnliche Spannweiten, arithmetische Mittel und Standardabweichungen generiert, ebenso wie für die Minkowski-Distanz und die gewichtete und potenzierte euklidische Distanz. Bei der Betrachtung der realen Abweichungen zwischen den approximierten und den tatsächlichen Wegstrecken wurde dann aber wieder der Eindruck aus der Analyse der Kriterien AD und SD bestätigt, dass die alternativen Distanz-Metriken in einigen Bereichen bessere Ergebnisse erzielen.

Nachfolgend sollen deshalb die alternativen Distanz-Metriken sowohl untereinander als auch mit dem bisherigen Vorgehen im Simulationswerkzeug *SimChain* auf Basis einer Gewichtung der euklidischen Distanz mit dem Faktor $[g = 1, 3]$ gegenübergestellt und analysiert werden. Dazu werden abermals die Kennzahlen der deskriptiven Statistik verglichen. Außerdem werden für die Distanz-Metriken in Tabelle 5.7 die beiden definierten Kriterien AD und SD aufgeführt.

Tabelle 5.7: Übersicht statistischer Kennzahlen zum Vergleich der Distanz-Metriken (in km , außer SD)

Distanz-Metrik	Min	Max	R	\bar{x}	s	AD	SD
tatsächliche Wegstrecke	0,4	78,3	77,9	8,9	9,1	-	-
Minkowski-Distanz ($p=1,15$)	0,2	76,7	76,5	8,4	8,6	6 017,5	995,0
gew. eukl. Distanz ($g=1,225$)	0,2	72,4	72,2	8,7	8,9	4 334,0	540,8
gew. eukl. Distanz ($g=1,25$)	0,2	74,2	74,0	8,9	9,1	4 413,1	539,6
pot. eukl. Distanz ($z=1,07$)	0,2	79,0	78,8	8,4	9,3	5 532,9	910,5
pot. eukl. Distanz ($z=1,08$)	0,2	82,3	82,1	8,6	9,6	5 589,5	896,3
gew. eukl. Distanz ($g=1,3$)	0,3	77,2	76,9	9,2	9,5	5 184,8	682,7

In Tabelle 5.7 ist zu sehen, dass die verschiedenen Distanz-Metriken inklusive der mit $[g = 1, 3]$ gewichteten euklidischen Distanz zunächst einmal ähnliche Lageparameter erzielen. Die Spannweite der erzeugten Distanzen liegt zwischen $72,2 km$ und $82,1 km$ und damit in relativer Nähe zur Spannweite der tatsächlichen Wegstrecke von $77,9 km$ bei ähnlichen Minima und Maxima. Beim Vergleich von arithmetischem Mittel und Stan-

dardabweichung wird ersichtlich, dass die mit $[g = 1, 25]$ gewichtete euklidische Distanz Werte generiert, die mit den Werten der tatsächlichen Wegstrecke übereinstimmen. Für alle anderen Distanz-Metriken weichen diese leicht von den Werten für die tatsächliche Wegstrecke ab. Die mit $[g = 1, 3]$ gewichtete euklidische Distanz erzeugt dabei das größte arithmetische Mittel und die zweitgrößte Standardabweichung.

Der Vergleich der beiden Kriterien AD und SD zeigt, dass beide gewichteten euklidischen Distanzen zunächst die niedrigsten Werte erzielen. Sowohl für die Minkowski-Distanz als auch für die potenzierten euklidischen Distanzen werden für beide Kriterien wesentlich höhere Werte erzeugt. Für die mit $[g = 1, 225]$ und $[g = 1, 25]$ gewichteten euklidischen Distanzen liegen die Werte für die Auswertung des zweiten Kriteriums mit 540,8 und 539,6 sehr nah beieinander. Für das erste Kriterium gibt es aber geringfügig größere Differenzen von 4334,0 km ($g = 1, 225$) zu 4413,1 km ($g = 1, 25$).

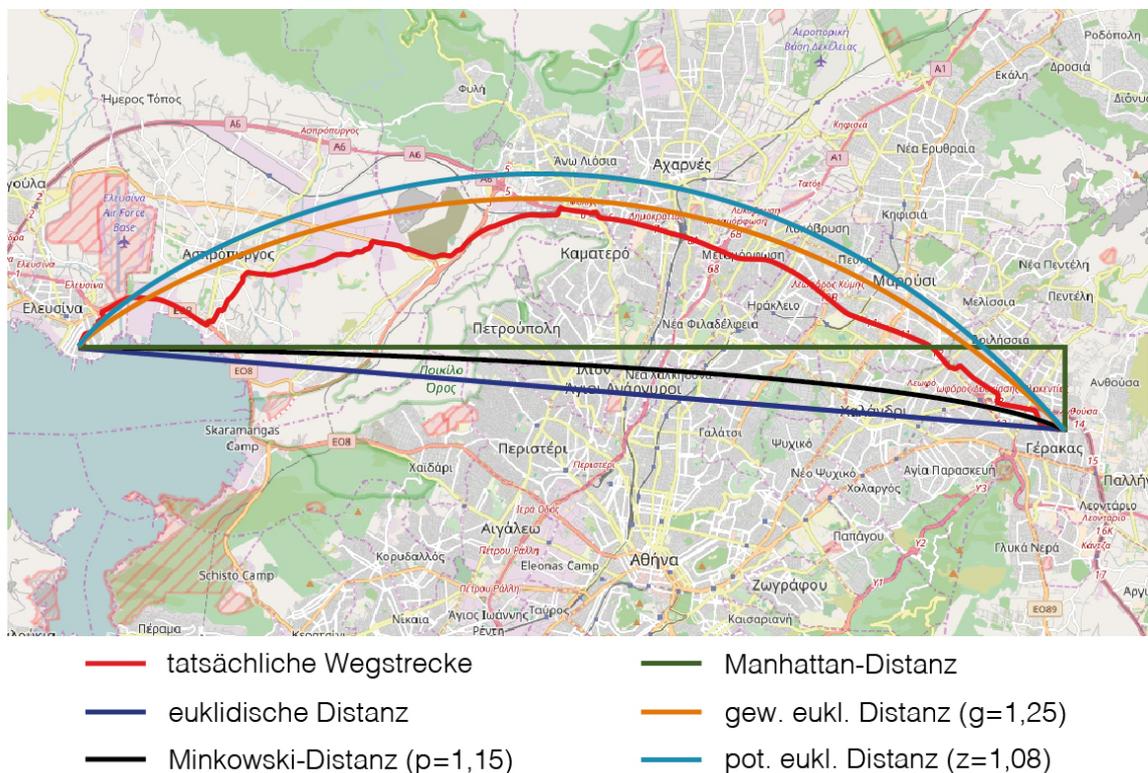


Abbildung 5.11: Beispielroute in der Metropolregion Athen

Die mit $[g = 1, 225]$ und $[g = 1, 25]$ gewichteten euklidischen Distanzen generieren bei Betrachtung der definierten Kriterien zunächst einmal die beste Approximation der Wegstrecke. Dies ist auch exemplarisch in Abbildung 5.11 dargestellt. Hier ist gut ersichtlich, dass die mit $[g = 1, 25]$ gewichtete euklidische Distanz die tatsächliche Wegstrecke sehr gut nachzeichnet. Demgegenüber führen euklidische Distanz und Manhattan-Distanz zu einer starken Unterbewertung der Wegstrecke.

Die Werte in Tabelle 5.7 zeigen, dass die Optimierung der gewichteten euklidischen Distanz darüber hinaus bessere Ergebnisse liefert als das bisherige Vorgehen im Simulationswerkzeug *SimChain* mit dem Gewichtungsfaktor $[g = 1, 3]$. Bereits der Vergleich von arithmetischem Mittel und Standardabweichung zeigte auf, dass die Werte für den Faktor $[g = 1, 3]$

aufgrund der höheren Ausprägung gegenüber den optimierten gewichteten euklidischen Distanzen leicht erhöht sind. Außerdem sind beide Kriterien für die mit $[g = 1, 3]$ gewichtete euklidische Distanz im Verhältnis zur optimierten gewichteten euklidischen Distanz wesentlich höher.

Im nächsten Schritt soll untersucht werden, ob der bisherige Eindruck durch die Analyse der realen Abweichungen zwischen den Distanz-Metriken und der tatsächlichen Wegstrecke bestätigt werden kann. Wie bereits in Unterabschnitt 5.2.2 dargestellt und auch hier noch einmal in Tabelle 5.8 aufgeführt, generiert die mit $[g = 1, 25]$ gewichtete euklidische Distanz zunächst ein arithmetisches Mittel von $0,0 \text{ km}$, welches mit dem Median übereinstimmt. Das arithmetische Mittel wird somit nicht von den Ausreißern beeinflusst, sondern die Ausreißer gleichen sich gegenseitig aus. Dies ist bspw. bei der Minkowski-Distanz oder der mit $[g = 1, 3]$ gewichteten euklidischen Distanz nicht der Fall. Ausreißer verzerren das arithmetische Mittel bei der optimierten Minkowski-Distanz nach oben und bei der gewichteten euklidischen Distanz ($g = 1, 3$) nach unten, sodass das arithmetische Mittel nicht mit dem Median übereinstimmt.

Ein weiterer Nachteil des bisherigen Vorgehens wird bei der Untersuchung der Quantile sichtbar. So liegen für die mit $[g = 1, 25]$ gewichtete euklidische Distanz die mittleren 50 % in einem Intervall von $0,7 \text{ km}$, während für den Gewichtungsfaktor $[g = 1, 3]$ ein Intervall von $1,0 \text{ km}$ aufgespannt wird und somit die höchste Spannweite zwischen unterem und oberem Quantil für alle Distanz-Metriken erzeugt.

Tabelle 5.8: Statistische Kennzahlen für die Untersuchung der realen Abweichungen der Distanz-Metriken (in km)

Distanz-Metrik	\bar{x}	$\tilde{x}_{0,5}$	$\tilde{x}_{0,25}$	$\tilde{x}_{0,75}$
Minkowski-Distanz ($p=1,15$)	0,4	0,2	-0,2	0,7
gew. eukl. Distanz ($g=1,225$)	0,2	0,1	-0,2	0,4
gew. eukl. Distanz ($g=1,25$)	0,0	0,0	-0,4	0,3
pot. eukl. Distanz ($z=1,07$)	0,5	0,5	0,2	0,8
pot. eukl. Distanz ($z=1,08$)	0,3	0,4	0,0	0,7
gew. eukl. Distanz ($g=1,3$)	-0,4	-0,2	-0,8	0,2

Im Boxplot-Diagramm (Abbildung 5.12) werden wiederum Lage und Streuung der Abweichungen ausgewählter Distanz-Metriken abgebildet. Neben der Minkowski-Distanz wird sich aufgrund der Ähnlichkeit der Ergebnisse bei der anfänglichen Kriterienauswertung für die gewichteten euklidischen Distanzen auf den Parameterwert $[g = 1, 25]$ konzentriert. Aus gleichem Grund wird nachfolgend für die potenzierte euklidische Distanz der Parameterwert $[z = 1, 08]$ betrachtet. Außerdem wird die mit $[g = 1, 3]$ gewichtete euklidische Distanz in die Betrachtung mit aufgenommen, da so im Simulationswerkzeug *SimChain* verfahren wird.

Auch die Whisker spannen für die mit $[g = 1, 3]$ gewichtete euklidische Distanz mit $4,0 \text{ km}$ das größte Intervall im Vergleich zu allen anderen Distanz-Metriken aus. Bis auf die Minkowski-Distanz liegt die Spannweite zwischen den Whiskern für die anderen Distanz-

Metriken unter $3,0 \text{ km}$. Dies wirkt sich schließlich auch auf die Ausreißer aus. Während für die anderen Distanz-Metriken immer mehr als 10 % der Daten als Ausreißer deklariert sind, werden für die mit $g = 1,3$ gewichtete euklidische Distanz lediglich 7,7 % als Ausreißer deklariert. Dies ist aber bereits auf die erhöhte Spannweite zwischen unterem und oberem Quantil zurückzuführen.

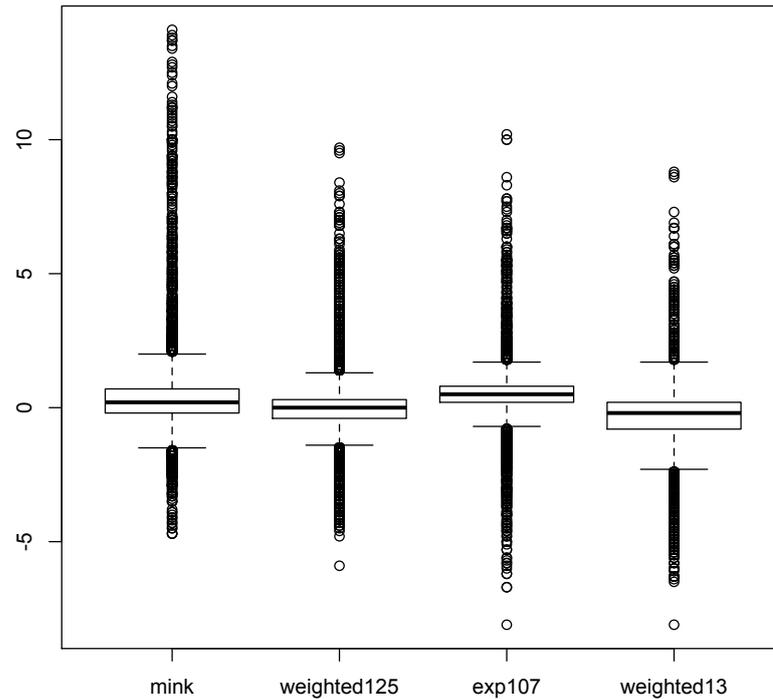


Abbildung 5.12: Boxplot der realen Abweichungen für den Vergleich der Distanz-Metriken

Die Gegenüberstellung der optimierten Distanz-Metriken hat gezeigt, dass die gewichtete euklidische Distanz ($g = 1,25$) zunächst die besten Ergebnisse generiert. Dies wurde besonders durch die Auswertung der Kriterien AD und SD deutlich. Die statistischen Kennzahlen zeigten daraufhin, dass durch die gewichtete euklidische Distanz eine ähnliche Verteilung erzeugt wird, wie die tatsächliche Wegstrecke vorgibt.

Darüber hinaus konnte dargestellt werden, dass die bisherige Vorgehensweise auf Basis einer Gewichtung mit dem Faktor $[g = 1,3]$ gewisse Schwächen gegenüber der optimierten gewichteten euklidischen Distanz mit einem Parameterwert von $[g = 1,25]$ aufweist. Die optimierte euklidische Distanz erzeugt sowohl für die Kriterien AD und SD als auch bei der Untersuchung der statistischen Kennzahlen bessere Ergebnisse als die mit $[g = 1,3]$ gewichtete euklidische Distanz. Es sei darauf verwiesen, dass auch andere Gewichtungsfaktoren, die in der Literatur für den Einsatz in der Simulation erwähnt wurden, hier als pauschale Faktoren nicht die optimale Lösungen darstellen.

Die Wahl der entsprechenden Parameter ist also grundsätzlich abhängig vom jeweiligen Distributionsnetzwerk. Wie bereits in Kapitel 4 dargestellt, ist für die City-Logistik die Belieferung auf der *letzten Meile* eine besondere Herausforderung. Dies wurde damit begründet, dass sich hier oft hohe Kosten und Ineffizienzen ergeben. Besonders vor dem Hintergrund unwirtschaftlicher Transporte scheint eine Anpassung der Parameterwerte für den Anwendungsfall notwendig. Nur so kann garantiert werden, dass sich die Ergebnisse aus der Simulation auf die Realität übertragen lassen.

5.3 Analyse und Weiterentwicklung der Vorgehensweisen zur zeitabhängigen Fahrzeitberechnung

Die Literaturrecherche in den Abschnitten 4.2 und 4.3 hat gezeigt, dass in urbanisierten Räumen relativ niedrige Durchschnittsgeschwindigkeiten für den motorisierten Verkehr angenommen werden müssen. Außerdem wurde dargestellt, dass die Durchschnittsgeschwindigkeiten nicht für alle Städte gleich sind, sondern bereits hier stark schwanken. Des Weiteren wurde klar, dass die Geschwindigkeiten zusätzlich in Abhängigkeit von zeitlichen Faktoren gewissen Schwankungen unterworfen sind. So kann nachts ggf. eine höhere Durchschnittsgeschwindigkeit erreicht werden als tagsüber; die Durchschnittsgeschwindigkeit wird dabei zusätzlich durch Berufsverkehre in den Morgen- und Abendstunden beeinflusst. Auch zwischen unterschiedlichen Wochentagen zeichnen sich Schwankungen der Durchschnittsgeschwindigkeit ab. Daneben werden besonders in urbanisierten Räumen häufig Tempolimits zur Verkehrsberuhigung eingesetzt, die die Durchschnittsgeschwindigkeit weiter verringern können. Insbesondere die Annahme einer Durchschnittsgeschwindigkeit für alle Fahrten wird vor dem Hintergrund, dass die Belieferung auf der *letzten Meile* gerade im Bereich der City-Logistik oftmals sehr unwirtschaftlich ist, kritisch betrachtet. Damit die Ergebnisse der Simulation auf die Realität übertragbar sind, muss gerade im Bereich der City-Logistik das Vorgehen zur Berechnung der Fahrzeit genauer abgebildet werden.

In Abschnitt 4.2 wurde zur Ermittlung der zeitabhängigen Differenzen der Durchschnittsgeschwindigkeit ein Data-Mining-Ansatz angeführt. Dabei wird eine große Menge an Verkehrsdaten aus dem entsprechenden Gebiet des Distributionsnetzwerks benötigt. Wie in Abschnitt 4.2 erwähnt, verwenden Ehmke et al. (2012) dazu die GPS-Daten einer Taxi-Flotte. Die Ergebnisgüte könnte weiter verbessert werden, wenn eine auf den Anwendungsfall bezogene Datengrundlage für den Data-Mining-Ansatz verwendet wird. Hierfür würden sich bspw. die GPS-Daten über den Fuhrpark der Zustellungsfahrzeuge von entsprechenden Logistikdienstleistern anbieten. Ein exemplarischer Unterschied zu den Daten der Taxi-Flotte können z. B. abweichende Beschleunigungs- und Bremsrampen von Zustellfahrzeugen im Vergleich zu Pkws sein. Gerade für den Bereich des E-Commerce kann nicht immer davon ausgegangen werden, dass Zustellfahrzeuge bis zur zulässigen Höchstgeschwindigkeit beschleunigen können. Dies ist darauf zurückzuführen, dass häufig nur kurze Strecken zwischen den einzelnen Stopps einer Auslieferungstour liegen können. Dies führt wiederum dazu, dass sich die Geschwindigkeit eines Zustellfahrzeuges stark von der

Durchschnittsgeschwindigkeit eines Taxis unterscheiden kann. Somit könnte mit spezifischeren Ausgangsdaten eine genauere Qualität der Modelleingangsdaten für die Simulation der Fahrzeit erzeugt werden.

Wie bereits durch die Problematik der Beschleunigungs- und Bremsrampen im vorherigen Absatz angedeutet, werden Unterschiede der Durchschnittsgeschwindigkeiten erwartet. Dazu werden nach dem in Abschnitt 4.2 dargestellten Ansatz zunächst auf der ersten Aggregationsebene für die einzelnen Streckenabschnitte eines Straßennetzes die Durchschnittsgeschwindigkeiten ermittelt. Auf der zweiten Aggregationsebene werden dann die Wegstrecken hinsichtlich der Variation der Durchschnittsgeschwindigkeit in Clustern gruppiert. Die Analyse erfolgt für 24x7 Zeitabschnitte anhand der stündlichen Unterteilung für jeden Wochentag. Die Unterteilung der Wochentage in entsprechende Zeitabschnitte muss nicht anhand der Stundenanzahl von 24 erfolgen, sondern kann auch in abweichender Granularität vorgenommen werden (vgl. Eglese et al. 2006).

Die Berechnung der Fahrzeit erfolgt dann auf Basis der Cluster-Unterteilung für die einzelnen Streckenabschnitte entsprechend der zeitlichen Ansetzung der einzelnen Tour. Dieses Vorgehen ist für die Fahrzeitberechnung auf Basis von approximierten Wegstrecken nicht praktikabel, da die einzelnen Straßenabschnitte und die Profile der aggregierten Durchschnittsgeschwindigkeiten nicht integriert werden können. Aus diesem Grund muss dieser Ansatz an die Berechnung mit approximierten Wegstrecken angepasst werden.

Da bei der Approximation der Wegstrecke zwischen zwei Punkten, die anhand ihrer Geokoordinaten spezifiziert sind, keine Informationen über die einzelnen zurückzulegenden Straßenabschnitte beschrieben werden, soll im nachfolgenden Ansatz die Analyse der GPS-Daten auf Basis von Quelle-Senke-Verbindungen einer Relation erfolgen. Da bei längeren Wegstrecken häufiger Schnellstraßen als bei kurzen Wegstrecken genutzt werden können, kann bei langen Wegstrecken ggf. eine höhere Durchschnittsgeschwindigkeit erzielt werden. Somit wird vorgeschlagen, zunächst eine Clusteranalyse der Wegstrecken vorzunehmen. Dabei werden die Wegstrecken entsprechend ihrer Länge in einheitliche Gruppen aufgeteilt. Hierbei wird angenommen, dass die Durchschnittsgeschwindigkeiten auf den Strecken innerhalb eines gebildeten Clusters korrelieren. Anschließend erfolgt innerhalb der Cluster die Ermittlung der Durchschnittsgeschwindigkeit zunächst für die einzelnen Wochentage und darauf aufbauend für die jeweiligen Zeitabschnitte eines Wochentages.

Die Ermittlung der zeitabhängigen Fahrzeit erfolgt danach entsprechend der Länge der approximierten Wegstrecke, der zeitlichen Ansetzung einer Tour und des entsprechenden Geschwindigkeitsprofils für das jeweilige Cluster der approximierten Wegstrecke. Somit können zunächst einmal Unterschiede hinsichtlich der Durchschnittsgeschwindigkeit über den Tag verteilt abgebildet werden. Die Erfassung von Geschwindigkeitsverringerungen während der Hauptverkehrszeiten trägt dazu bei, die Fahrzeitberechnung zu verbessern. Die Berechnung erfolgt anhand von Korrekturfaktoren gemäß des zugehörigen Geschwindigkeitsprofils. Dabei wird von einer allgemeinen Durchschnittsgeschwindigkeit v_a ausgegangen, die dann mit entsprechenden Korrekturfaktoren c_d zunächst für den Wochentag

und anschließend mit c_{dh} für den entsprechenden Zeitabschnitt des Wochentages korrigiert wird. Somit ergibt sich die zeitabhängige Fahrzeit entsprechend folgender Formel:

$$t_{ij}^{dh} = d_{ij} \cdot v_a(d_{ij}) \cdot c_d \cdot c_{dh} \tag{5.5}$$

Wie die schwankenden Verläufe der Durchschnittsgeschwindigkeit auf Stunden- und Tagesbasis aussehen können, ist exemplarisch in den Abbildungen 5.13 und 5.14 dargestellt.

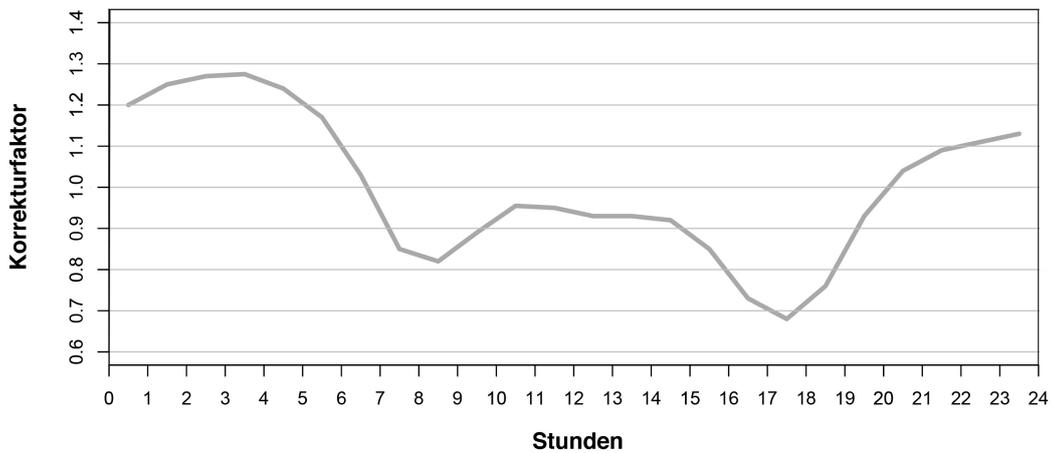


Abbildung 5.13: Exemplarische Darstellung der stündlichen Korrekturfaktoren für ein Cluster für einen Wochentag (in Anlehnung an Ehmke et al. 2012)



Abbildung 5.14: Exemplarische Darstellung der täglichen Korrekturfaktoren für ein Cluster (in Anlehnung an Ehmke et al. 2012)

Ein Vorteil der Gleichung 5.5 ist, dass der Aggregationsgrad durch Auslassen oder Hinzufügen von Korrekturfaktoren auf simple Art und Weise angepasst werden kann. Wie bereits in Abschnitt 4.2 dargestellt, wird die Durchschnittsgeschwindigkeit nicht nur durch tages- und nachtzeitabhängige Faktoren sowie Hauptverkehrszeiten beeinflusst, sondern ist darüber hinaus auch von Witterungsbedingungen abhängig. Diese klimabezogenen Abwei-

chungen von der Durchschnittsgeschwindigkeit werden bisher nicht durch die Gleichung 5.5 zur Berechnung der zeitabhängigen Fahrzeit erfasst.

Insbesondere zwischen den verschiedenen Jahreszeiten sind witterungsbedingte Abweichungen der Durchschnittsgeschwindigkeit zu erwarten. Da sich die Formel zur zeitabhängigen Fahrzeitberechnung durch Hinzufügen von Korrekturfaktoren sehr einfach anpassen lässt, wird vorgeschlagen, einen weiteren Faktor c_m zur Differenzierung der Monate in die Formel zu integrieren. Somit können nicht nur tages- und nachtzeitabhängige Einflüsse sowie Hauptverkehrszeiten erfasst, sondern darüber hinaus auch übergreifende Einflüsse klimatischer Bedingungen auf die Durchschnittsgeschwindigkeit abgebildet werden. Schwankungen der Durchschnittsgeschwindigkeit zwischen verschiedenen Monaten sind aber ggf. nicht ausschließlich auf klimatische Rahmenbedingungen zurückzuführen, sondern können darüber hinaus bspw. auch durch Reise- und Urlaubszeiten beeinflusst werden. Somit wird zwar mit dem Faktor c_m eine genauere Fahrzeitberechnung ermöglicht, diese ist aber nicht ausschließlich auf klimatische Bedingungen zurückzuführen. Durch die Integration des Faktors c_m ergibt sich folgende Formel:

$$t_{ij}^{mdh} = d_{ij} \cdot v_a(d_{ij}) \cdot c_m \cdot c_d \cdot c_{dh} \quad (5.6)$$

Abhängig vom Untersuchungsgegenstand einer Simulationsstudie kann diese Formel individuelle hinsichtlich ihres Detaillierungsgrades angepasst werden.

6 Konzept zur Integration von Verkehrsinformationen in die Simulation

In Abschnitt 5.2 wurde zunächst untersucht, wie eine Approximation von Wegstrecken erfolgen kann. Dabei hat sich herausgestellt, dass im Anwendungsfall die Approximation der Wegstrecke auf Basis der mit $[g = 1, 25]$ gewichteten euklidischen Distanz die besten Ergebnisse generiert. Darauf aufbauend wurde ein Konzept zur zeitabhängigen Fahrzeitberechnung auf Basis von approximierten Wegstrecken entwickelt. Hierzu wurde ein bestehender Data-Mining-Ansatz auf Basis von Streckenabschnitten eines Straßennetzes auf die Gegebenheiten der approximierten Wegstrecken angepasst.

Nachfolgend wird untersucht, wie die Informationen zur Approximation der Wegstrecke und zur Fahrzeitberechnung in die Simulation integriert werden können. Zunächst soll auf die Integration der Verkehrsinformationen innerhalb des Informationsgewinnungsprozesses eingegangen werden. Dabei soll auch betrachtet werden, inwieweit die Informationsgewinnung auf Basis gegebener Daten automatisiert innerhalb des Simulationswerkzeuges erfolgen kann. Darüber hinaus müssen weitere Möglichkeiten zur Einbeziehung von Verkehrsinformationen betrachtet werden, für den Fall, dass dies nicht innerhalb des Simulationswerkzeuges geleistet werden kann.

6.1 Einbeziehung der Wegstreckenapproximation in die Simulation

Prinzipiell kann das Simulationsvorgehen in die Prozesse der Modellbildung und der Informationsgewinnung unterteilt werden (siehe Abschnitt 3.1). Wie die Gewinnung der benötigten Informationen für die Approximation von Wegstrecken erfolgen kann, wurde grundsätzlich bereits in Kapitel 5 dargestellt. Bei der separaten Gewinnung von Modelleingangsdaten für die Approximation von Wegstrecken ist vorteilhaft, dass die Modellbildung für die Simulation des Distributionsnetzwerkes nicht strukturell darauf aufbaut. Wie bereits in Abschnitt 3.2 dargestellt, werden für das ausführbare Modell die *nutzbaren Eingangsdaten* als Phasenergebnisse der *Datennutzbarkeitsprüfung* benötigt. Die vorherigen Schritte der Informationsgewinnung können aber nach der Bestimmung des *objektiven Informationsbedarfs* in der *Zieldefinition* weitgehend unabhängig vom Prozess der Modellbildung erfolgen. Durch die Parallelisierung von Prozessschritten innerhalb des Simulationsvorgehens kann somit der zeitliche Umfang der Simulationsstudie verdichtet werden.

Neben der Approximation von Wegstrecken besteht grundsätzlich auch die Möglichkeit, über Methoden der Graphentheorie die Wegstrecke zwischen zwei Netzwerkknoten exakter zu ermitteln. Dabei ergibt sich für dafür eingesetzte Algorithmen, bspw. den Dijkstra-Algorithmus, wie in Abschnitt 4.1 angeführt ein Trade-off zwischen den Faktoren Speicherplatz und Laufzeit während der Simulation. Eine Entscheidung für die Approximation von Wegstrecken wäre aus folgendem Grund vorteilhaft: Die Auswahl ein entsprechendes Approximationsverfahren und damit einhergehend die Ermittlung der Parameterwerte für das ausgewählte Approximationsverfahren könnte innerhalb des Informationsgewinnungsprozesses der Simulation vorgelagert werden. Die Einbeziehung des ausgewählten Approximationsverfahrens in das Simulationsmodell und die darauf basierende Berechnung der Wegstrecke innerhalb der Simulation selbst wären anschließend wesentlich zeitökonomischer und bei gleichzeitig geringerer Speicherplatznutzung zu realisieren als die Verwendung von Graphsuchalgorithmen.

Auch wenn die Auswahl des Approximationsverfahrens und die damit einhergehende Ermittlung der notwendigen Parameterwerte grundsätzlich der Simulation vorgelagert sind, kann das Auswahlverfahren als Teil des Informationsgewinnungsprozesses partiell in das Simulationsprogramm integriert werden. Denkbar wäre zunächst, die Ermittlung der tatsächlichen Wegstrecken als Vergleichsdistanz durch das bereits in Abschnitt 5.2 verwendete Skript in das Simulationsprogramm einzubetten. Ob und in welchem Umfang zunächst die Vergleichsdistanzen durch Verwendung des Shell-Skriptes ermittelt werden müssen, ist grundlegend von der vorherigen Datenlage abhängig. Diese wird zunächst innerhalb der ersten Schritte des Informationsgewinnungsprozesses bestimmt. Des Weiteren kann die statistische Analyse innerhalb des Simulationsprogramms automatisiert werden.

Daher wird zunächst auf die datenvorverarbeitenden Schritte innerhalb des Informationsprozesses eingegangen. Darauf aufbauend findet dann eine nähere Darstellung des Konzeptes zur Integration der Methoden zur Wegstreckenapproximation in das Simulationstool statt.

Der Informationsgewinnungsprozess beginnt mit der Zieldefinition und der darauf aufbauenden Informationsidentifikation. Über den Logistikdatensatz können innerhalb der Zieldefinition zunächst alle notwendigen Daten zur Darstellung des entsprechenden Prozesses hergeleitet werden. Wie eingangs erwähnt, ist neben der Verwendung von Graphsuchalgorithmen für die Bestimmung der Wegstrecke auch die Anwendung von Approximationsmethoden in der Literatur beschrieben. Durch die Literaturrecherche in Abschnitt 4.1 konnte dargestellt werden, welche Relevanz eine sorgfältige Bestimmung von Wegstrecken zwischen zwei Netzwerkknoten auf Basis von Approximationsverfahren für die Simulation besitzt.

Für die Approximation der Wegstrecken zwischen zwei Knoten eines Distributionsnetzwerkes müssen zunächst die Netzwerkknoten selbst konkretisiert werden. Dies kann durch Adressen oder geographische Koordinaten erfolgen. In der sich anschließenden Erhebungsplanung innerhalb des Simulationsvorgehensmodells müssen deshalb zunächst mögliche Quellen identifiziert werden, aus denen diese Daten extrahiert werden können. Grundsätz-

lich muss davon ausgegangen werden, dass diese Daten bisher nicht in einheitlicher und zusammenhängender Form innerhalb eines Unternehmens erfasst wurden. Mit Blick auf den Untersuchungsgegenstand sind entsprechende Daten über ein Distributionsnetzwerk in eingesetzten Routing- oder Tourenplanungsprogrammen zu finden. Je nach Umfang und Funktionalität der Planungswerkzeuge sind dabei bereits auch zurückgelegte Strecken zwischen den Netzwerkknoten erfasst.

Für den Fall der Wegstreckenapproximation wird angenommen, dass die Daten zur Beschreibung eines Distributionsnetzwerkes grundsätzlich in geschlossener und digitaler Form vorliegen. Somit schließt sich folgerichtig die Datenstrukturierung an.

Im Rahmen der Datenstrukturierung müssen die aus verschiedenen Quellen extrahierten Datensätze in einer Datenbank konsolidiert werden. Wichtig ist hierbei, dass Zusammenhänge zwischen Daten und Tabellen besonders vor dem Hintergrund unterschiedlicher Quellen deutlich dargestellt werden. Anschließend muss im Rahmen der Datenstrukturierung eine Korrektur der Daten erfolgen, wozu unter anderem eine Fehlerbereinigung notwendig ist. Dazu gehört auch eine Anpassung der Granularität. Im Anwendungsfall der Metropolregion Athen stellt sich dies wie folgt dar: Bei km-Angaben würden zwei oder mehrere Dezimalstellen innerhalb der Wegstreckenapproximation fälschlicherweise eine Genauigkeit imitieren, die auf Basis der zur Verfügung gestellten Daten der Logistikdienstleister zur Beschreibung des Distributionsnetzwerkes (PLZ-Gebiete) nicht gegeben ist.

Dem Schritt der Datenstrukturierung schließt sich im Informationsgewinnungsprozess die statistische Analyse an. Diese soll wie anfänglich erwähnt innerhalb des Simulationswerkzeuges stattfinden. Innerhalb der Analyse der Distanz-Metriken zur Approximation von Wegstrecken in Abschnitt 5.2 wurden zunächst die tatsächlichen Wegstrecken über die Ausführung eines Shell-Skriptes ermittelt. Dies war notwendig, da die bereitgestellten Daten der Logistikdienstleister keine Informationen über die zurückgelegten Strecken zwischen den einzelnen Netzwerkknoten beinhalteten.

Zur Bestimmung der tatsächlichen Wegstrecke auf Basis des Shell-Skriptes muss eine Liste übergeben werden, die mindestens folgende Angaben enthält:

- Index für die Quelle-Senke-Verbindung
- geographische Breite der Quelle in Dezimalschreibweise
- geographische Länge der Quelle in Dezimalschreibweise
- geographische Breite der Senke in Dezimalschreibweise
- geographische Länge der Senke in Dezimalschreibweise

Das Shell-Skript liest zunächst die geographischen Koordinaten jeder Quelle-Senke-Verbindung aus (Zeile 25-28, Anhang A) und übergibt sie der URL zur Berechnung der Route (Zeile 30, Anhang A). Die zurückgegebene XML-Datei beinhaltet dann die Route, aus der schließlich die Länge der Wegstrecke ausgelesen und in der Tabelle gespeichert wird (Zeile 34-35, Anhang A).

Das beschriebene Vorgehen zur Ermittlung der tatsächlichen Wegstrecke zwischen zwei Netzwerkknoten ist nur notwendig, wenn diese Daten nicht aus den Routing- und Tourenplanungsprogrammen extrahiert werden können. Die verwendete URL-Vorlage des Geographischen Instituts der Universität Heidelberg bietet dabei die Möglichkeit, weitere Parameter zu spezifizieren (siehe Anhang A). Dabei können explizit Parameter für das Routing von Lkws einbezogen werden. Dazu können Angaben über maximal zulässige Höhe, Breite, Gewicht sowie Achslast bestimmt werden. Für die Ermittlung der entsprechenden Lkw-Route durch das Ausführen des Shell-Skriptes gäbe es die Möglichkeit, die übergebene Liste mit den geographischen Koordinaten der Quelle und Senke einer Verbindung um weitere Attribute zur Spezifizierung des Fuhrparks zu ergänzen. Das Shell-Skript müsste dann folglich erweitert werden, um die entsprechenden Attribute auszulesen und der URL in der richtigen Reihenfolge zu übergeben.

Auf Basis dieser Daten kann anschließend die statistische Analyse durchgeführt werden. Die in Kapitel 5 durchgeführte Analyse hat gezeigt, dass die beiden angeführten Kriterien AD und SD (Abschnitt 5.2) deutlich aufzeigen, welche Distanz-Metrik die tatsächliche Wegstrecke am besten approximiert. Deshalb wäre eine Möglichkeit, die statistische Analyse in das Simulationswerkzeug zu integrieren und sich bei der Bewertung der verschiedenen Distanz-Metriken auf die angeführten Kriterien AD und SD zu beschränken. Für eine ausführlichere Analyse der Approximationsverfahren könnten auch weitere Kriterien im Rahmen einer Nutzwertanalyse über gewichtete Faktoren einbezogen werden. Als Ergebnis der statistischen Analyse wird ein Approximationsverfahren ausgewählt und der zu verwendende Parameter für das Distributionsnetzwerk identifiziert. Für den Fall, dass eine Bewertung auf Basis der Kriterien AD und SD zu keinem eindeutigen Ergebnis führt, müssten weitere Entscheidungskriterien spezifiziert werden. Für das angeführte Beispiel wäre denkbar, das arithmetische Mittel der optimierten Parameterwerte der Kriterien AD und SD zu bilden.

Im untersuchten Anwendungsfall der Metropolregion Athen waren in den zur Verfügung gestellten Daten der Logistikdienstleister keine Informationen über die verwendeten Fahrzeugtypen oder sonstige Spezifikationen des Fuhrparks angegeben. Somit fußt die in Kapitel 5 getätigte Analyse zur Approximation von Wegstrecken auf den Standardwerten des Routings innerhalb des frei nutzbaren Online-Kartendienstes OSM. Wenn jedoch genauere Informationen über eingesetzte Fahrzeuge ermittelt werden können, sollten diese auch in das Routing zur Ermittlung der tatsächlichen Wegstrecke einbezogen werden. Dabei muss beachtet werden, dass sich abhängig von der Spezifikation des Fuhrparks unterschiedliche Distributionsnetze ergeben.

Das Ergebnis der statistischen Analyse muss abschließend an das Simulationsmodell übergeben werden. Dabei müssen sowohl die Berechnungsvorschrift für das ausgewählte Approximationsverfahren als auch der ermittelte Parameter für die Distanz-Metrik in das Simulationsmodell integriert werden. Hiermit zeigt sich deutlich die Schnittstelle zum Simulationsmodell. Die vorherigen Schritte waren eindeutig dem Informationsgewinnungsprozess zuzuordnen. Alle weiteren Schritte sind aber von der Informationsgewinnung losgelöst und

weisen einen deutlichen Bezug zum Simulationsmodell selbst auf. Die Approximation der Wegstrecken erfolgt schließlich während der Durchführung der Simulation.

6.2 Einbeziehung der zeitabhängigen Fahrzeitberechnung in die Simulation

In Abschnitt 5.3 wurde zunächst dargestellt, welche Daten für eine Analyse zur zeitabhängigen Fahrzeitberechnung benötigt werden. Des Weiteren wurde beschrieben, wie die Analyse zur Ermittlung der Korrekturfaktoren zur zeitabhängigen Fahrzeitberechnung erfolgen kann. Da die Vorverarbeitung der Daten auf Basis eines Clustering-Verfahrens wesentlich aufwendiger ist als bei der Wegstreckenapproximation, kann im Gegensatz zur Analyse der Approximationsverfahren für die Wegstreckenberechnung keine Automatisierung innerhalb des Simulationsprogramms vorgenommen werden. Trotzdem können die Prozesse der Modellbildung und der Informationsgewinnung parallel ausgeführt werden, was wiederum zu einer zeitlichen Verdichtung des Umfangs einer Simulationsstudie führt. Die ersten Schritte des Informationsgewinnungsprozesses sind die Zieldefinition und die Informationsidentifikation. Der Logistikdatensatz bildet dabei die Grundlage zur Abbildung aller logistischen Prozesse. Die Literaturrecherche in Abschnitt 4.2 hat einerseits aufgezeigt, dass die Berechnung der Fahrzeit innerhalb der Simulation basierend auf einer Durchschnittsgeschwindigkeit erfolgt. Andererseits konnte aufgezeigt werden, dass die Belieferung besonders auf der *letzten Meile* und insbesondere im Bereich der City-Logistik unter hohen Effizienzverlusten leidet. Daraus ergibt sich die Anforderung, die Berechnung der Fahrzeit in Abhängigkeit der zeitlichen Verteilung der Transporte genauer abzubilden und schließlich weitere Verkehrsinformationen in die Fahrzeitberechnung zu integrieren. Wie genau die Fahrzeitberechnung erfolgen muss, ist dabei immer abhängig vom Untersuchungsgegenstand der Simulationsstudie. Nachfolgend wird beschrieben, welche datenvorverarbeitenden Schritte neben der Zieldefinition und Informationsidentifikation innerhalb des Informationsgewinnungsprozesses notwendig sind. Darauf aufbauend wird aufgezeigt, wie die aufbereiteten Daten in das Simulationsmodell einbezogen werden können.

Im dargestellten Ansatz von Cleophas und Ehmke (2014) wurden die Korrekturfaktoren zur fahrzeitabhängigen Fahrzeitberechnung durch einen Data-Mining-Ansatz auf Basis der GPS-Daten einer Taxi-Flotte bestimmt. In Abschnitt 5.3 konnte aufgezeigt werden, dass die Verwendung anwendungsbezogener GPS-Daten aus dem Fuhrpark von Logistikdienstleistern ggf. zu einer geeigneteren Berechnung der Fahrzeit führen kann. Um den in Abschnitt 5.3 dargestellten Ansatz zur zeitabhängigen Fahrzeitberechnung zu realisieren, werden zunächst einmal Daten benötigt. Grundsätzlich ist auch hier davon auszugehen, dass die benötigten Daten bisher nicht in einheitlicher Form für den Untersuchungsgegenstand der Simulationsstudie innerhalb des Unternehmens erfasst wurden. Deshalb muss zunächst im Rahmen der Erhebungsplanung eine Identifikation möglicher Datenquellen im Unternehmen vorgenommen werden.

Für die Umsetzung des dargestellten Data-Mining-Ansatzes werden unterschiedliche Daten benötigt:

- Quelle-Senke-Verbindung
- Länge der zurückgelegten Strecke auf der entsprechenden Verbindung
- benötigte Fahrzeit für die entsprechende Verbindung

Für die Darstellung des Distributionsnetzwerkes über Quelle-Senke-Verbindungen kann auf die in Abschnitt 6.1 dargestellten Datenquellen wie Routing- oder Tourenplanungsprogramme zurückgegriffen werden. Im besten Fall sind dabei auch schon die Länge der zurückgelegten Strecken auf den Quelle-Senke-Verbindungen erfasst. Falls dies nicht der Fall ist, müssen weitere Daten wie bspw. GPS-Daten der Transportfahrzeuge in die Betrachtung einbezogen werden. Diese Daten lassen sich möglicherweise aus heutzutage vorgeschriebenen digitalen Tachographen extrahieren.

Digitale Tachographen erfassen unter anderem die Fahrzeiten, die gefahrenen Geschwindigkeiten und auch die Ortung des Fahrzeuges über GPS. Somit können digitale Tachographen eine mögliche Quelle innerhalb des Unternehmens für die Ermittlung der benötigten Daten zur Bestimmung der Wegstreckenlänge und der benötigten Fahrzeit einer Quelle-Senke-Verbindung sein. Dabei ist zu beachten, dass die Daten aus unterschiedlichen Quellsystemen kommen und somit in weiteren Schritten zusammengeführt werden müssen. Außerdem ist ein sensibler Umgang mit personenbezogenen Daten, die über einen digitalen Tachographen erfasst werden, notwendig.

Nachdem die Daten im Rahmen der Erhebung und Datenerfassung aus den unterschiedlichen Quellsystemen extrahiert wurden, müssen diese während der Datenstrukturierung weiter aufbereitet werden. Grundsätzlich ist über die Ortung per GPS ersichtlich, wo sich ein Transportfahrzeug bewegt hat. Diese müssen anschließend aber noch den einzelnen Quelle-Senke-Verbindungen zugeordnet werden, um schließlich die Wegstreckenlänge einer Quelle-Senke-Verbindung zu bestimmen. Wenn die GPS-Daten bspw. einem bestimmten Fahrzeug zugeordnet sind, muss vorher zunächst das Fahrzeug der entsprechenden Quelle-Senke-Verbindung zugeordnet werden. Noch komplizierter wird es, wenn bspw. die GPS-Daten personalisiert über den Tachographen für einen Fahrer erfasst werden und somit in einem noch weiter vorgelagerten Schritt zunächst der Fahrer dem entsprechenden Fahrzeug zugeordnet werden muss. Besonders vor dem Hintergrund unterschiedlicher Datenquellen ist die richtige Bestimmung der Wegstreckenlänge mit hohem Aufwand verbunden. Wichtig hierbei ist, dass bei Konsolidierung der unterschiedlichen Tabellen und Daten in einer Datenbank die Relationen richtig dargestellt werden. Bei der Fehlerbereinigung im Rahmen der Datenstrukturierung ist auch darauf zu achten, dass unvollständige oder fehlerhafte Datensätze sowie Ausreißer entfernt werden. Als Ergebnis der Datenstrukturierung liegen schließlich die *strukturierten Daten* in geschlossener und digitaler Form vor.

Innerhalb der statistischen Analyse soll aufbauend auf den *strukturierten Daten* als Phasenergebnis der Datenstrukturierung ein Data-Mining-Ansatz umgesetzt werden. Ziel des Data-Mining-Ansatzes ist, Korrekturfaktoren für die Berechnung einer zeitabhängigen Fahrzeitberechnung zu bestimmen. Das grundsätzliche Vorgehen ist dabei bereits in Ab-

schnitt 5.3 dargestellt. Das Ergebnis des Data-Mining-Ansatzes sind je nach Ausrichtung der Simulation und des zu untersuchenden Simulationsgegenstandes monatliche, tägliche oder stündliche Faktoren für die Korrektur der allgemeinen Durchschnittsgeschwindigkeit. Im Allgemeinen wird der Prozess der Informationsgewinnung unabhängig vom Prozess der Modellbildung durchgeführt. Trotzdem muss im Modell die Möglichkeit zur Integration der Korrekturfaktoren zur Berechnung der fahrzeitabhängigen Fahrzeit integriert werden. Dazu muss die in Abschnitt 5.3 entwickelte Berechnungsvorschrift (5.6) zur Ermittlung der zeitabhängigen Fahrzeit in das Simulationsmodell integriert werden. Die Berechnung der Fahrzeit schließt dann die Korrekturfaktoren mit ein. Somit werden während der Simulation der Fahrzeit in Abhängigkeit zur zeitlichen Verteilung des abzubildenden Transportes die entsprechenden Korrekturfaktoren aus den Modelleingangsdaten ausgelesen. So kann eine genauere Darstellung der Transporte innerhalb der Simulation erfolgen.

Die Integration von Verkehrsinformationen und darauf aufbauend die möglichst genaue Abbildung der Fahrzeit in der Simulation war im Rahmen dieser Arbeit durch die speziellen Gegebenheiten der City-Logistik und der Ineffizienz der Belieferung auf der *letzten Meile* motiviert. Dies ist aber ggf. nicht für jede Simulationsstudie relevant und muss mit dem entsprechenden Grad der Darstellung im Simulationsmodell abgestimmt werden. Ein Vorteil des in Abschnitt 5.3 entwickelten Vorgehens zur Berechnung einer zeitabhängigen Fahrzeit ist, dass die Berechnung der Fahrzeit durch Hinzufügen oder Auslassen von Korrekturfaktoren sehr einfach angepasst werden kann. Somit kann das entsprechende Vorgehen auf simple Art und Weise an den Gegenstand der Simulation angeglichen werden.

Die Berechnung der Fahrzeit erfolgt während der Simulation. Dazu wird dann auf die approximierte Wegstrecke, die allgemeine Durchschnittsgeschwindigkeit sowie die Geschwindigkeitsprofile entsprechend der zeitlichen Ansetzung einer Relation zurückgegriffen.

7 Zusammenfassung und Ausblick

Zunächst wurden in Kapitel 4 aufbauend auf den theoretischen Grundlagen die Forschungsbedarfe zur Integration von Verkehrsinformationen in die Simulation dargestellt. Die Notwendigkeit zur Integration weiterer Verkehrsinformationen im Rahmen der City-Logistik ergab sich aus der mangelhaften Effizienz der Belieferung auf der *letzten Meile* in urbanisierten Räumen. Die Integration von Verkehrsinformationen wurde dabei in die Gebiete der Wegstreckenapproximation und der Berechnung einer zeitabhängigen Fahrzeit unterteilt.

Die Analyse der Verfahren zur Wegstreckenapproximation wurde am Anwendungsfall der Metropolregion Athen vorgenommen. Die Untersuchung hat zunächst gezeigt, dass unterschiedliche alternative Verfahren grundsätzlich bessere Ergebnisse bei der Wegstreckenapproximation liefern als die Ermittlung der Wegstrecke auf Basis der euklidischen oder der Manhattan-Distanz. Im Anwendungsfall liefert bspw. bei Betrachtung der aufgestellten Kriterien und statistischer Lage- und Streuungsparameter die mit $[g = 1, 25]$ gewichtete euklidische Distanz das beste Ergebnis bei der Wegstreckenapproximation.

Das zweite Themengebiet dieser Arbeit widmete sich der Berechnung einer zeitabhängigen Fahrzeit. Auch dieses war wie die Analyse der Verfahren zur Wegstreckenapproximation durch die Ineffizienz der Belieferung auf der *letzten Meile* motiviert. Aufbauend auf einem Data-Mining-Ansatz aus der Literatur, der sich auf die ermittelten Durchschnittsgeschwindigkeiten eines Straßenabschnitts gründete, wurde ein Ansatz entwickelt, der auf approximierten Wegstrecken basiert. Die Berechnung einer zeitabhängigen Fahrzeit erfolgt dabei unter Einbeziehung von Faktoren zur Korrektur der Durchschnittsgeschwindigkeit auf monatlicher, täglicher und stündlicher Basis.

Aufbauend auf der vorhergehenden Analyse wurde in Kapitel 6 untersucht, wie die Wegstreckenapproximation und die zeitabhängige Fahrzeitberechnung im Rahmen des Simulationsvorgehensmodells in die Simulation einbezogen werden können. Hierbei wurde konkret auf die datenvorverarbeitenden Schritte innerhalb des Informationsgewinnungsprozesses eingegangen. Für die Wegstreckenapproximation können Schritte teilweise innerhalb des Simulationswerkzeuges selbst automatisiert und parallel zur Modellbildung umgesetzt werden. Hierbei konnte deutlich die Grenze zum Simulationsmodell aufgezeigt werden. Demgegenüber ist die Umsetzung eines Data-Mining-Ansatzes für die Ermittlung der Korrekturfaktoren zur zeitabhängigen Fahrzeitberechnung im Rahmen des Informationsgewinnungsprozesses wesentlich aufwendiger als die Analyse von Verfahren zur Wegstreckenapproximation. Es konnte dargestellt werden, dass die Prozesse der Informationsgewinnung und der Modellbildung grundsätzlich parallel ablaufen können, aber dass hier keine Integration von Analyseschritten in das Simulationswerkzeug möglich ist, wie dies für die

Wegstreckenapproximation möglich war. Abschließend wurde gezeigt, auf welche Weise die Integration von Korrekturfaktoren in das Simulationsmodell erfolgen kann. Hierbei konnte die einfache Adaption der Korrekturfaktoren an den Untersuchungsgegenstand der Simulationsstudie dargestellt werden.

Hinsichtlich der Analyse der Minkowski-Metrik bleibt aber folgender Gesichtspunkt offen: Wie wirkt sich die Rotation der Referenzachsen auf die Ergebnisse der Minkowski-Metrik aus. Es besteht die Vermutung, dass das Rotieren des Koordinatensystems bspw. die Ergebnisauswertung der aufgestellten Untersuchungskriterien AD und SD (S. 40) signifikant beeinflusst. Diese Rotation ist für die gewichtete und potenzierte euklidische Distanz irrelevant, da die Länge der euklidischen Distanz unabhängig von der Ausrichtung des zugrunde liegenden Koordinatensystems ist. Da die Analyse der Approximationsverfahren aber durch eine möglichst genaue Abschätzung der Wegstrecke motiviert war, muss diese Problematik in die Betrachtung aufgenommen werden. Die Bewertung der Approximationsverfahren wäre in diesem Fall wesentlich aufwendiger als das in dieser Arbeit gewählte Vorgehen. Folglich müsste auch untersucht werden, inwieweit die Bewertung der Approximationsmethoden weiterhin automatisiert im Simulationswerkzeug vorgenommen werden könnte. Des Weiteren muss untersucht werden, wie genau die Ergebnisse des Informationsgewinnungsprozesses zur Analyse der Approximationsverfahren Eingang in das Simulationsmodell finden.

Für die Integration einer fahrzeitabhängigen Fahrzeitberechnung in die Simulation wurde in den Abschnitten 5.3 und 6.2 zunächst ein Vorgehen entworfen und anschließend dargestellt, wie die entsprechenden Korrekturfaktoren der Durchschnittsgeschwindigkeit im Rahmen des Informationsgewinnungsprozesses in die Simulation aufgenommen werden können. Auch wurde gezeigt, welche Daten für die Umsetzung des dargestellten Data-Mining-Ansatzes benötigt werden. Das entwickelte Vorgehen stellt sich bisher aber als überaus theoretisches Konzept dar. Deshalb könnte bezüglich dieses Themenbereiches in weiteren Betrachtungen die Umsetzung des in dieser Arbeit entwickelten Vorgehens untersucht werden. Von besonderem Interesse für die Integration der fahrzeitabhängigen Fahrzeitberechnung in die Simulation wäre in diesem Fall die Ermittlung der notwendigen Ausgangsdaten zur Umsetzung des Data-Mining-Ansatzes. Auch für diesen Themenbereich sollte untersucht werden, wie die Erkenntnisse aus dem Informationsgewinnungsprozess anschließend konkret in das Simulationsmodell integriert werden können.

Bei der Beschreibung des Distributionsnetzwerkes wurden zunächst nur die Verbindungen zwischen den Netzwerkknoten selbst betrachtet. Für eine nähere Beschreibung des Netzwerkes könnten aber auch die auftretenden Häufigkeiten der gewählten Verbindungen relevant sein. Auch dieser Aspekt könnte Bestandteil zukünftiger Untersuchungen sein. Zur Analyse der Distanz-Metriken wurden im Anwendungsfall alle verfügbaren Daten aus den zur Verfügung gestellten Datensätzen verwendet. Grundsätzlich können allgemein gültige Aussagen über ein Distributionsnetzwerk erst getroffen werden, wenn eine bestimmte Menge an Ausgangsdaten für die Analyse des Netzwerkes zur Verfügung stehen. Dies gilt sowohl für die Analyse der Distanz-Metriken als auch für die Untersuchung der zeitabhängigen Fahrzeit. Um valide Modelleingangsdaten im Informationsgewinnungsprozess

zu generieren, muss geprüft werden, wieviele Daten notwendig sind, um allgemeingültige Aussagen aus der Analyse treffen zu können. Dies ist wichtig, damit die Ergebnisse der Simulation auf die Realität übertragbar sind.

Literatur

- Alanis, Ramon (2014). „Efficient storage of transport network routes for simulation models“. In: *Proceedings of the 2014 Winter Simulation Conference*. Hrsg. von A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley und J. A. Miller, S. 1931–1942.
- Allen, Julian, Michael Browne, Allan Woodburn und Jacques Leonardi (2012). „The Role of Urban Consolidation Centres in Sustainable Freight Transport“. In: *Transport Reviews* 32.4, S. 473–490.
- Apparicio, Philippe, Mohamed Abdelmajid, Mylène Riva und Richard Shearmur (2008). „Comparing alternative approaches to measuring the geographical accessibility of urban health services: Distance types and aggregation-error issues“. In: *International Journal of Health Geographics* 7.1.
- Arvidsson, Niklas und Michael Browne (2013). „A review of the success and failure of tram systems to carry urban freight: the implications for a low emission intermodal solution using electric vehicles on trams“. In: *European Transport* 54.5.
- Balci, Osman (1998). „Verification, Validation, and Testing“. In: *Handbook of simulation: principles, methodology, advances, applications and practice*. Hrsg. von Jerry Banks. New York: Wiley. Kap. 10, S. 355–393.
- Banks, Jerry, Danie Gerstein und Sean J. Searles (1988). „Modeling processes, validation, and verification of complex simulation: A survey“. In: *Methodology and validation : proceedings of the Conference on Methodology and Validation, 1987, 6 - 9 April 1987, Orlando, Florida*. Hrsg. von Osman Balci. San Diego: Society for Computer Simulation, S. 13–18.
- Becker, Jörg, Wolfgang Probandt und Oliver Vering (2012). *Grundsätze ordnungsmäßiger Modellierung*. Berlin, Heidelberg: Springer.
- Benjelloun, Abderrahim, Teodor Gabriel Crainic und Yvon Bigras (2010). „Towards a taxonomy of City Logistics projects“. In: *Procedia Social and Behavioral Sciences* 2.3, S. 6217–6228.
- Berens, Wolfgang und Franz-Josef Körling (1985). „Estimating road distances by mathematical functions“. In: *European Journal of Operational Research* 21.1, S. 54–56.
- Bernhard, Jochen, Dirk Jodin, Kay Hömberg, Sonja Kuhnt, Christoph Schürmann und Sigrid Wenzel (2007). *Vorgehensmodell zur Informationsgewinnung – Prozessschritte und Methodennutzung*. Technical Report 06008. Dortmund: Universität Dortmund.
- Bertazzon, Stefania, Rizwan Shahid, Merrill L. Knudtson und William A. Ghali (2009). „Comparison of distance measures in spatial analytical modeling for health service planning“. In: *BMC health service research* 9.1.

- Brade, Dirk (2003). „A generalized process for the verification and validation of models and simulation results“. Dissertation. Neubiberg: Universität der Bundeswehr.
- Brimberg, Jack, Robert F. Love und John H. Walker (2007). „Estimation of travel distances with the weighted lp-norm: Some empirical results“. In: *Journal of Transport Geography* 15.1, S. 62–72.
- Chow, Joseph Y. J., Choon Heon Yang und Amelia C. Regan (2010). „State-of-the-art of freight forecast modeling: lessons learned and the road ahead“. In: *Transportation: Planning - Policy - Research - Practice* 37.6, S. 1011–1030.
- Cleophas, Catherine und Jan Fabian Ehmke (2014). „Wann sind Lieferaufträge profitabel?“ In: *Wirtschaftsinformatik* 56.3, S. 173–184.
- Crainic, Teodor Gabriel, Nicoletta Ricciardi und Giovanni Storchi (2009). „Models for Evaluating and Planning City Logistics Systems“. In: *Transportation Science* 43.4, S. 432–454.
- Cruz, Cecilia und Antoine Montonen (2016). „Implementation and impacts of low emission zones on freight activities in Europe: Local schemes versus national schemes“. In: *Transportation Research Procedia* 12, S. 544–556.
- Dallmeyer, Jörg (2014). *Simulation des Straßenverkehrs in der Großstadt*. Berlin, Heidelberg: Springer.
- EffizienzCluster LogistikRuhr (2016a). *Dynamics in Navigation*. URL: http://www.effizienzcluster.de/de/spitzencluster_wettbewerb/projekt.php?proPid=9.
- (2016b). *Urban Business Navigation – branchenindividuelle Navigation*. URL: http://www.effizienzcluster.de/de/spitzencluster_wettbewerb/projekt.php?proPid=32.
- (2016c). *Urban Retail Logistics*. URL: http://www.effizienzcluster.de/de/spitzencluster_wettbewerb/projekt.php?proPid=28.
- Eglese, R., W. Maden und D. Black (2010). „Vehicle Routing and Scheduling with Time-Varying Data: A Case Study“. In: *The Journal of the Operational Research Society* 61.3, S. 515–522.
- Eglese, Richard, Will Maden und Alan Slater (2006). „A Road Timetable to aid vehicle routing and scheduling“. In: *Computers & Operations Research* 33.12, S. 3508–3519.
- Ehmke, Jan Fabian, André Steinert und Dirk Christian Mattfeld (2012). „Advanced routing for city logistics service providers based on time-dependent travel times“. In: *Journal of Computational Science* 3.1, S. 193–205.
- Eley, Michael (2012). *Simulation in der Logistik*. Berlin, Heidelberg: Springer.
- Gabler, Manuel, Stefan Schröder, Hanno Friedrich und Gernot Liedtke (2013). „Generierung der Nachfragestrukturen für die mikroskopische Simulation des städtischen Distributionsverkehrs im Lebensmittelhandel“. In: *Wirtschaftsverkehr 2013*. Hrsg. von Uwe Clausen und Carina Thaller. Berlin, Heidelberg: Springer. Kap. 3, S. 32–48.
- Gevaers, Roel, Eddy van de Voorde und Thierry Vanelslander (2011). „Characteristics and Typology of Last-mile Logistics from an Innovation Perspective in an Urban Context“. In: *City distribution and urban freight transport: multiple perspectives*. Hrsg. von Cathy Macharis und Sandra Melo. Cheltenham: Elgar. Kap. 3.

- Graham, Phena (2014). *The Mobility Opportunity – Improving public transport to drive economic growth*. Report. Siemens AG.
- Hoerstebroek, Tim (2014). „Strategische Analyse der Elektromobilität in der Metropolregion Bremen/Oldenburg – Multi-Agenten basierte Simulation alternativer Antriebssysteme“. Diss. Oldenburg: Carl von Ossietzky Universität Oldenburg.
- Hömberg, Karl, Jan Hustadt, Dirk Jodin, Joachim Kochsiek, Lard Nagel und Iwo Riha (2007). *Basisprozesse für die Modellierung in großen Netzen der Logistik*. Forschungsbericht 07004. Universität Dortmund.
- Jodin, Dirk, Sonja Kuhnt und Sigrid Wenzel (2009). „Methodennutzung zur Informationsgewinnung in großen Netzen der Logistik“. In: *Große Netze der Logistik*. Hrsg. von Peter Buchholz und Uwe Clausen. Berlin, Heidelberg: Springer. Kap. 1, S. 1–18.
- Kaupp, Martin (1998). *City-Logistik als kooperatives Güterverkehrs-Management*. Wiesbaden: Deutscher Universitäts-Verlag.
- Kiwitt, Petra (2010). „City Logistics – Distribution in wachsenden Ballungsgebieten“. In: *Delivering Tomorrow – Zukunftstrend Nachhaltige Logistik*. Hrsg. von Christof E. Ehrhart. Bonn: Deutsche Post AG, S. 97–100.
- Klug, Florian (2010). *Logistikmanagement in der Automobilindustrie*. Berlin Heidelberg: Springer.
- Klumpp, Matthias, Sascha Bioly und Hella Abidi (2013). „Zur Interdependenz demografischer Entwicklung, Urbanisierung und Logistiksystemen“. In: *Wirtschaftliche Implikation des demografischen Wandels – Herausforderungen und Lösungsansätze*. Hrsg. von Michael Göke und Thomas Heupel. Berlin Heidelberg: Springer. Kap. 11, S. 161–178.
- Kritzinger, Stefanie, Karl F. Doerner, Richard F. Hartl, Günter Kiechle, Horst Stadler und Senthana Sirpi Manohar (2012). „Using traffic information for time-dependent vehicle routing“. In: *Procedia - Social and Behavioral Sciences* 39, S. 217–229.
- Lam, William H. K., Hu Shao und Agachai Sumalee (2008). „Modeling impacts of adverse weather conditions on a road network with uncertainties in demand and supply“. In: *Transportation Research Part B* 42.10, S. 890–910.
- Law, Averill M. und Walter David Kelton (1991). *Simulation modeling and analysis*. New York: McGraw Hill.
- Lehmacher, Wolfgang (2015). *Logistik im Zeichen der Urbanisierung – Versorgung von Stadt und Land im digitalen Zeitalter*. Berlin, Heidelberg: Springer.
- Leimeister, Jan Marco (2015). *Einführung in die Wirtschaftsinformatik*. Berlin, Heidelberg: Springer.
- McGrew, J. Chapman, Arthur J. Lembo und Charles B. Monroe (2014). *An Introduction to Statistical Problem Solving in Geography*. Illinois: Waveland Press.
- Morris, J. G. und R. F. Love (1972). „Modelling Inter-City Road Distance by Mathematical Functions“. In: *Journal of the Operational Research Society* 23.1, S. 61–71.
- Müller-Steinfahrt, Ulrich und Peter Klaus (2014). *Die Top 100 der Logistik: Marktgrößen, Marktsegmente und Marktführer in der Logistik-Dienstleistungswirtschaft 2014/2015*. Hamburg: DVV Media Group.

- Oexler, Petra (2002). *Citylogistik-Dienste – Präferenzanalysen bei Citylogistik-Akteuren und Bewertung eines Pilotbetriebs dargestellt am Beispiel der dienstleistungsorientierten Citylogistik Regensburg*. Wirtschaft & Raum. München: VVF.
- Pamučar, Dragan, Ljubormir Gigović, Goran Ćirović und Miodrag Regodić (2016). „Transport spatial model for the definition of green routes for city logistics centers“. In: *Environmental Impact Assessment Review* 56, S. 72–87.
- Piontek, Jochem (2013). *Bausteine des Logistikmanagements*. Herne: NWB-Verlag.
- Rabe, Markus, Sven Spieckermann und Sigrid Wenzel (2007). *Verifikation und Validierung für die Simulation in Produktion und Logistik – Vorgehensmodelle und Techniken*. Berlin, Heidelberg: Springer.
- Sammer, Gerd (2012). „Wirkung und Risiken einer City-Maut als zentrale Säule eines städtischen Mobilitätskonzepts“. In: *Zukünftige Entwicklung in der Mobilität - Betriebswirtschaftliche und technische Aspekte*. Hrsg. von Heike Proff, Jörg Schönharting, Dieter Schramm und Jürgen Ziegler. Berlin, Heidelberg: Springer, S. 479–492.
- Schulte, Christof (2013). *Logistik: Wege zur Optimierung der Supply Chain*. Vahdens Handbücher der Wirtschafts- und Sozialwissenschaften. München: Vahlen.
- Siedentop, Stefan (2015). „Ursachen, Ausprägungen und Wirkungen der globalen Urbanisierung – ein Überblick“. In: *Globale Urbanisierung – Perspektive aus dem All*. Hrsg. von Hannes Taubenböck, Michael Wurm, Thomas Esch und Stefan Dech. Berlin, Heidelberg: Springer. Kap. 3, S. 11–22.
- Statista (2008). *Average speed in Europe's 15 most congested cities in 2008*. URL: <https://www.statista.com/statistics/264703/average-speed-in-europes-15-most-congested-cities/>.
- Strauß, Susanne (1997). *City-Logistik – Ein Instrument zur Verringerung des städtischen Güterverkehrs*. Schriftenreihe Verkehr. Clausthal-Zellerfeld: Papierflieger.
- Taniguchi, Eiichi, Russell G. Thompson, Tadashi Yamada und Ron van Duin (2001). *City Logistics – Network modelling and intelligent transport systems*. Bingley: Emerald Group Publishing.
- Taubenböck, Hannes und Michael Wurm (2015). „Globale Urbanisierung – Markenzeichen des 21. Jahrhunderts“. In: *Globale Urbanisierung – Perspektive aus dem All*. Hrsg. von Hannes Taubenböck, Michael Wurm, Thomas Esch und Stefan Dech. Berlin, Heidelberg: Springer. Kap. 2, S. 5–10.
- Taubenböck, Hannes, Michael Wurm, Thomas Esch und Stefan Dech, Hrsg. (2015). *Globale Urbanisierung – Perspektive aus dem All*. Berlin, Heidelberg: Springer.
- United Nations (2015). *World Urbanization Prospects: The 2014 Revision*. Techn. Ber. New York: Department of Economics und Social Affairs.
- van Woensel, T., L. Kerbach, H. Peremans und N. Vandaele (2008). „Vehicle routing with dynamic travel times: A queueing approach“. In: *European Journal of Operational Research* 186.3, S. 990–1007.
- VDI 3633: 2013 (2013). *Simulation von Logistik-, Materialfluss- und Produktionssystemen – Begriffe*.

-
- VDI 3633: 2014 (2014). *Simulation von Logistik-, Materialfluss- und Produktionssystemen – Blatt 1: Grundlagen*.
- Vidalakis, Christos, John E. Tookey und James Sommerville (2011). „Logistics simulation modelling across construction supply chains“. In: *Construction Innovation* 11.2, S. 212–228.

Abbildungsverzeichnis

Abb. 3.1:	Simulationsvorgehensmodell (nach Law und Kelton 1991, S. 67)	15
Abb. 3.2:	Simulationsvorgehensmodell (nach Balci 1998, S. 337)	16
Abb. 3.3:	Simulationsvorgehensmodell (nach Rabe et al. 2007, S. 5)	18
Abb. 3.4:	Simulationsvorgehensmodell (nach Jodin et al. 2009, S. 5)	20
Abb. 3.5:	Ausführliche Darstellung des Informationsgewinnungsprozesses (nach Bernhard et al. 2007, S. 7)	21
Abb. 5.1:	Darstellung des Untersuchungsgebietes	37
Abb. 5.2:	Auswertung des Kriteriums AD für die Minkowski-Metrik	41
Abb. 5.3:	Auswertung des Kriteriums SD für die Minkowski-Metrik	42
Abb. 5.4:	Boxplot der realen Abweichungen für die Analyse der Minkowski-Metrik	45
Abb. 5.5:	Auswertung des Kriteriums AD für die gewichtete euklidische Distanz .	46
Abb. 5.6:	Auswertung des Kriteriums SD für die gewichtete euklidische Distanz . .	47
Abb. 5.7:	Boxplot der realen Abweichungen für die Analyse der gewichteten eukli- dischen Distanz	49
Abb. 5.8:	Auswertung des Kriteriums AD für die potenzierte euklidische Distanz .	50
Abb. 5.9:	Auswertung des Kriteriums SD für die potenzierte euklidische Distanz .	51
Abb. 5.10:	Boxplot der realen Abweichungen für die Analyse der potenzierten eu- klidischen Distanz	53
Abb. 5.11:	Beispielroute in der Metropolregion Athen	55
Abb. 5.12:	Boxplot der realen Abweichungen für den Vergleich der Distanz-Metriken	57
Abb. 5.13:	Exemplarische Darstellung der stündlichen Korrekturfaktoren für ein Cluster für einen Wochentag (in Anlehnung an Ehmke et al. 2012)	60
Abb. 5.14:	Exemplarische Darstellung der täglichen Korrekturfaktoren für ein Clus- ter (in Anlehnung an Ehmke et al. 2012)	60

Tabellenverzeichnis

Tab. 5.1:	Übersicht statistischer Kennzahlen für die Analyse der optimierten Minkowski-Metrik	43
Tab. 5.2:	Statistische Kennzahlen für die Untersuchung der realen Abweichungen der Minkowski-Metrik	44
Tab. 5.3:	Übersicht statistischer Kennzahlen für die Analyse der gewichteten euklidischen Distanz	48
Tab. 5.4:	Statistische Kennzahlen für die Untersuchung der realen Abweichungen der gewichteten euklidischen Distanz	49
Tab. 5.5:	Übersicht statistischer Kennzahlen für die Analyse der potenzierten euklidischen Distanz	52
Tab. 5.6:	Statistische Kennzahlen für die Untersuchung der realen Abweichungen der potenzierten euklidischen Distanz	52
Tab. 5.7:	Übersicht statistischer Kennzahlen zum Vergleich der Distanz-Metriken .	54
Tab. 5.8:	Statistische Kennzahlen für die Untersuchung der realen Abweichungen der Distanz-Metriken	56

Abkürzungsverzeichnis

API	Application Programming Interface
BDE	Betriebsdatenerfassung
CSV	Comma-separated values
FMCG	Fast Moving Consumer Goods
GIS	Geoinformationssystem
GVZ	Güterverkehrszentrum
ITPL	IT in Produktion und Logistik
MDE	Maschinendatenerfassung
ORS	OpenRouteService
OSM	OpenStreetMap
VDI	Verein Deutscher Ingenieure
V&V	Verifikation & Validierung
VV&T	Verifikation, Validierung und Test
WGS 84	World Geodetic System 1984
XML	Extensible Markup Language

A Anhang: Programm zur Ermittlung der tatsächlichen Wegstrecken

Shell-Programm zur Ermittlung der tatsächlichen Wegstrecken

```
1  #!/bin/bash
2
3  # OpenRouteService API
4  # Source: http://wiki.openstreetmap.org/wiki/OpenRouteService
   #Direct_Routing_Request_.28via_GET.29
5  urlTemplate="http://opensls.geog.uni-heidelberg.de/route?
   api_key=ee0b8233adff52ce9fd6afc2a2859a28&start=%s,%s&end=%
   s,%s&via=&lang=de&distunit=KM&routepref=Car&weighting=
   Shortest&avoidAreas=&useTMC=false&noMotorways=false&
   noTollways=false&noUnpavedroads=false&noSteps=false&
   noFerries=false&instructions=false "
6
7  csv=$1
8  out=$2
9  outtmp=$2.$$
10
11 # Temporary CSV file , to be merged w/ the original one.
12 echo "Distance;Route" > $outtmp
13
14 function getCoordinate {
15     echo $1 | csvcut -d ';' -c $2 | tail -n 1
16 }
17
18 header='head -1 $csv '
19 for line in `tail -n +2 $csv`;
20 do
21     current='echo -e $header "\n" $line '
22     ID='getCoordinate "$current" ID '
23     echo "Processing entry ID #$ID."
24
25     startLong='getCoordinate "$current" SourceCdtEast '
```

```

26     startLati='getCoordinate "$current" SourceCdtNorth '
27     endLong='getCoordinate "$current" SinkCdtEast '
28     endLati='getCoordinate "$current" SinkCdtNorth '
29
30     url='printf $urlTemplate $startLong $startLati
        $endLong $endLati '
31     outroute=$ID.route.xml
32     curl -s "$url" > $outroute
33
34     distance='xmlstarlet sel -N "xls:http://www.opengis.
        net/xls" -t -m "//xls:TotalDistance" -v "./@value"
        -n $outroute '
35     echo $distance ";" $outroute >> $outtmp
36 done
37
38 # Merge the original w/ the temp file containing the added
        columns 'distance' and 'route'.
39 csvjoin -d ";" $csv $outtmp > $out
40 rm -f $outtmp

```

Verwendete URL innerhalb des Shell-Skriptes und zu spezifizierende Parameter:

URL:

```

http://openls.geog.uni-heidelberg.de/route?api_key=
ee0b8233adff52ce9fd6afc2a2859a28&start=%s,%s&end=%s,%s&via=&lang=de&
distunit=KM&routepref=Car&weighting=Shortest&avoidAreas=&useTMC=false&
noMotorways=false&noTollways=false&noUnpavedroads=false&noSteps=false&
noFerries=false&instructions=false

```

Parameter:

- start: geographische Länge und Breite des Startpunktes
- via: geographische Länge und Breite eines Zwischenpunktes
- end: geographische Länge und Breite des Endpunktes
- routepref: Routenprofil (*Car* - Auto, *Bicycle* - Fahrrad, *HeavyVehicle* - Lkw etc.)
- weighting: bevorzugte Routingmethode (fastest, shortest, recommended)
- distunit: Maßeinheit für die Entfernung (standardmäßig in *km*, ansonsten *M* für Meter oder *MI* für Meilen)
- noMotorways: Autobahnen und Schnellstraßen umfahren? (false, true)
- noTollways: Mautstraßen umfahren? (false, true)
- noFerries: Fähren meiden? (false, true)

- noUnpavedroads (nur für Fahrrad): Unbefestigte Straßen umfahren? (false, true)
- noSteps (nur für Fahrrad): Treppen umfahren? (false, true)
- instructions: Ausgabe der Routenanweisung? (false, true)
- lang: Sprache der Routenanweisung
- maxspeed: Angabe der Maximalgeschwindigkeit

Weitere Parameter für das Routing schwerer Fahrzeuge:

- hazardous: Transport von Gefahrgütern? (false, true)
- value_weight: Maximales Gewicht in Tonnen
- value_height: Maximale Höhe in Metern
- value_length: Maximale Länge in Metern
- value_axleload: Maximale Achslast in Tonnen

Exemplarische Darstellung der CSV-Datei in Tabellenform:

ID	Sink	Sink	Source	Source	Distance	Route
	CdtNorth	CdtEast	CdtNorth	CdtEast		
15	38,0170492	23,8593451	38,0375283	23,5442699	33,1	15.route.xml
18	38,0724807	23,7470526	38,0078308	23,6823818	11,0	18.route.xml
21	37,9471012	23,7015776	38,0375283	23,5442699	23,8	21.route.xml
22	38,0282618	23,7876589	38,0375283	23,5442699	27,6	22.route.xml
23	38,0602000	23,7920834	38,0282618	23,7876589	5,7	23.route.xml
...

Exemplarische Ausgabe der XML-Datei:

```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <xls:XLS xmlns:xls="http://www.opengis.net/xls" xmlns:xsi="
   http://www.w3.org/2001/XMLSchema-instance" xmlns:gml="http
   ://www.opengis.net/gml" version="1.1" xsi:schemaLocation="
   http://www.opengis.net/xls http://schemas.opengis.net/ols
   /1.1.0/RouteService.xsd">
3 <xls:ResponseHeader xsi:type="xls:ResponseHeaderType"/>
4 <xls:Response xsi:type="xls:ResponseType" requestID
   ="123456789" version="1.1" numberOfResponses="1">
5 <xls:DetermineRouteResponse xsi:type="xls:
   DetermineRouteResponseType">
6 <xls:RouteSummary>
7 <xls:TotalTime>PT1M57S</xls:TotalTime>

```

```
8      <xls:TotalDistance uom="KM" value="1.4"/>
9      <xls:ActualDistance uom="KM" value="0.0"/>
10     <xls:BoundingBox srsName="EPSG:4326" >
11         <gml:pos >23.7171017 37.9783380 </gml:pos>
12         <gml:pos >23.7272497 37.9826174 </gml:pos>
13     </xls:BoundingBox>
14 </xls:RouteSummary>
15 <xls:RouteGeometry>
16     <gml:LineString srsName="EPSG:4326" >
17         <gml:pos >23.727078 37.9784336 </gml:pos>
18         <gml:pos >23.7267922 37.978338 </gml:pos>
19         <gml:pos >23.7266523 37.9783467 </gml:pos>
20         <gml:pos >23.7265717 37.9783583 </gml:pos>
21         <gml:pos >23.7271595 37.9807043 </gml:pos>
22         <gml:pos >23.7272497 37.9810776 </gml:pos>
23         <gml:pos >23.7245712 37.9815121 </gml:pos>
24         <gml:pos >23.7232066 37.9817081 </gml:pos>
25         <gml:pos >23.7231085 37.9817859 </gml:pos>
26         <gml:pos >23.722999 37.981828 </gml:pos>
27         <gml:pos >23.7208973 37.9808984 </gml:pos>
28         <gml:pos >23.7207304 37.9808276 </gml:pos>
29         <gml:pos >23.7206831 37.9808496 </gml:pos>
30         <gml:pos >23.7197473 37.9821119 </gml:pos>
31         <gml:pos >23.7193657 37.9826174 </gml:pos>
32         <gml:pos >23.7187592 37.982396 </gml:pos>
33         <gml:pos >23.7171456 37.981676 </gml:pos>
34         <gml:pos >23.7171017 37.9816567 </gml:pos>
35     </gml:LineString>
36 </xls:RouteGeometry>
37 </xls:DetermineRouteResponse>
38 </xls:Response>
39 </xls:XLS>
```

B Anhang: Taxonomie zur Einordnung von Projekten der City-Logistik

1. Description

- Objectives
 - Economics
 - Environmental
 - Congestion
 - Energy consumption
- Status
 - Study
 - Pilot/Experimental
 - Permanent
 - Temporary
- Evaluation tools for City Logistics initiatives
 - Financial analysis
 - Traffic analysis
 - Environmental assessment
- Stakeholders
 - City
 - Government
 - Shippers and consignees
 - Carriers
 - Others
- Project initiator
 - Public Sector
 - Private Sector

2. Business model

- Infrastructure financing
 - Public

- Private
- Operation financing
 - Autonomous
 - Initial
 - Unconditional subsidy
 - Conditional subsidy
- Management
 - Public
 - Private
 - Concession
- Competitive advantage
 - None
 - Partial
 - Total

3. Functionality

- Consolidation
 - None
 - Single-level
 - Two-level
- Modal shift
 - Underground system
 - Pipeline
 - Water
 - Rail
- Regulation
 - Time windows
 - Lane and space use
 - Environmental standards
 - Weight/volume/load factor
 - Access charges
- Intelligent Transportation Systems (ITS)
 - Advanced Traffic Management Systems (ATMS), Advanced Traveller Information Systems (ATIS)
 - E-payment and Automatic Vehicle Identification (AVI)

- Advanced Fleet Management Systems (AFMS)

- Cooperation
 - Carriers
 - Shippers/Consignees

4. Scope

- Geographic coverage
 - Corridor
 - Shared zone
 - Parcelled zone
 - Town
- Transport mode
 - Unimodal
 - Multimodal
- Product
 - Product specific
 - Non-product-specific
- Customers
 - Residents
 - Commercial
 - Others
- Services
 - Delivery
 - Pick up
 - Delivery and pick up
 - Additional logistics services

5. Technology

- Vehicles
 - Standard
 - Low-emission
 - Clean
- Information and communication
 - Internet and Electronic Data Interchange (EDI)
 - Wireless communications
 - Computational technologies

- Sensing technologies
- Decision technologies
 - Strategic/evaluation
 - Tactic/exploitation
 - Real-time

Eidesstattliche Versicherung

Name, Vorname

Matrikelnummer

Ich versichere hiermit an Eides statt, dass ich die vorliegende Masterarbeit mit dem Titel

Konzept zur Integration von Verkehrsinformationen für ein Supply-Chain-Simulationsmodell

selbstständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Belehrung:

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden. (§ 63 Abs. 5 Hochschulgesetz - HG -)

Die Abgabe einer falschen Versicherung an Eides statt wird mit Freiheitsstrafe bis zu 3 Jahren oder mit Geldstrafe bestraft.

Die Technische Universität Dortmund wird ggf. elektronische Vergleichswerkzeuge (wie z. B. die Software „turnitin“) zur Überprüfung von Ordnungswidrigkeiten in Prüfungsverfahren nutzen.

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

Ort, Datum

Unterschrift