

Oliver Kunze

Tourenplanung für den eCommerce-Lebensmittel- Heimlieferservice



Oliver Kunze

**Tourenplanung für den
eCommerce-Lebensmittel-Heimlieferservice**

Wissenschaftliche Berichte des
Institutes für Fördertechnik und Logistiksysteme
der Universität Karlsruhe (TH)
Band 67

Tourenplanung für den eCommerce- Lebensmittel-Heimlieferservice

von
Oliver Kunze



universitätsverlag karlsruhe

Dissertation, Universität Karlsruhe (TH)

Fakultät für Maschinenbau, 2006

Referenten: Prof. Dr.-Ing. Kai Furmans, Prof. Dr. Werner Rothengatter

Impressum

Universitätsverlag Karlsruhe
c/o Universitätsbibliothek
Straße am Forum 2
D-76131 Karlsruhe
www.uvka.de



Dieses Werk ist unter folgender Creative Commons-Lizenz
lizenziert: <http://creativecommons.org/licenses/by-nc-nd/2.0/de/>

Universitätsverlag Karlsruhe 2006
Print on Demand

ISSN: 0171-2772
ISBN-13: 978-3-86644-082-1
ISBN-10: 3-86644-082-0

**Tourenplanung für den
eCommerce-Lebensmittel-Heimlieferservice**

Zur Erlangung des akademischen Grades eines

Doktors der Ingenieurwissenschaften

von der Fakultät für Maschinenbau
der Universität Karlsruhe (TH)

genehmigte

Dissertation

von

Dipl. Wi. Ing. Oliver Kunze

Tag der mündlichen Prüfung:

30.03.2006

Erster Gutachter:

Prof. Dr.-Ing. Kai Furmans

Zweiter Gutachter:

Prof. Dr. Werner Rothengatter

Dank & Widmung

Die vorliegende Arbeit entstand parallel zu meiner Tätigkeit bei der PTV AG in Karlsruhe als externe Dissertation am Institut für Fördertechnik und Logistik (IFL) der Universität Karlsruhe.

Herrn Prof. Dr.-Ing. Dieter Arnold danke ich für die wohlwollende Unterstützung in der Frühphase dieser Arbeit. Herrn Prof. Dr.-Ing. Kai Furmans danke ich besonders für die kritischen und konstruktiven Impulse. Herrn Prof. Dr. Werner Rothengatter danke ich für seine Bereitschaft, das Zweitgutachten für diese Arbeit zu erstellen.

Bedanken möchte ich mich außerdem bei meinem Kollegen, Herrn Werner Heid, der federführend bei der Konzeption und Implementierung der Verfahren, die in dieser Arbeit genutzt werden, mitgewirkt hat. Bei der Parametrisierung der verschiedenen Verfahren erhielt ich wertvolle Unterstützung von meinen Kollegen Simone Martin, Frank Radaschewski, Hansjörg Back und Marcus Eiser. Für die Freiräume zur Erstellung dieser Arbeit möchte ich mich bei der PTV AG und insbesondere bei Herrn Dr. Joachim Schmidt bedanken.

Beim Korrekturlesen haben mir Herr Dr. Dieter Wild und Herr Bertram Meyer wertvolle Hinweise gegeben – Euch dafür ein großes Danke.

Widmen möchte ich diese Arbeit meinen Eltern Jürgen und Marianne Kunze.

Karlsruhe im Dezember 2005

Oliver Kunze

Abkürzungen

#:	(alleinstehend) Anzahl; z.B. „# Fahrzeuge“ := „Anzahl Fahrzeuge“
#:	(als Suffix) Kennzeichen für einen Satz von Variationen einer Instanz (Instanzenklasse) z.B. R205# := Satz von Variation der Instanz R205
:	(alleinstehend) allgemeiner Platzhalter z.B. FPA/ := FPA/PPA oder FPA/REA
:	(als Suffix) Kennzeichen für einen Satz von ähnlichen Instanzen (Instanzengruppe) z.B. ZUE500 := Satz von ähnlichen Instanzen
*-DO:	<u>d</u> elivery notice at <u>o</u> rders placement
*-ST:	<u>s</u> tochastic travel <u>t</u> imes
*-STW:	<u>s</u> oft <u>t</u> ime <u>w</u> indows
*-UC:	<u>u</u> ncertain <u>c</u> ustomers
*-UCD:	<u>u</u> ncertain <u>c</u> ustomers
*-UD:	uncertain demand
AVIV:	<u>A</u> vis <u>v</u> orlauf
B2B:	<u>B</u> usiness <u>t</u> o <u>B</u> usiness (Firma zu Firma)
B2C:	<u>B</u> usiness <u>t</u> o <u>C</u> onsumer (Firma zu Konsument)
c.p.:	<u>c</u> eteris <u>p</u> aribus
CRM	<u>C</u> ustomer <u>R</u> elationship <u>M</u> anagement (Kundenbeziehungsmanagement)
CVRP:	<u>C</u> apacitated <u>V</u> ehicle <u>R</u> outing <u>P</u> roblem (Tourenplanungsproblem mit Kapazitätsbeschränkungen)
DO:	<u>D</u> elivery Notification at <u>O</u> rders Entry (Lieferzeitusage zum Bestellzeitpunkt)
DOE:	<u>D</u> esign of <u>e</u> xperiments
EFG1:	<u>E</u> rfüllungsgrad <u>1</u> (Erfüllung ja, nein)
EFG2:	<u>E</u> rfüllungsgrad <u>2</u> (Abweichung vom zugesagten Lieferzeitintervall)
FGV:	<u>F</u> ilialgeschäft <u>V</u> ariante
Filter:	Prozess zur Auftragsannahme bzw. -ablehnung
FPA:	<u>F</u> ilter vor <u>P</u> lanung <u>A</u> nsatz
KEP:	<u>K</u> urier <u>E</u> xpress <u>P</u> aket
KPI:	<u>k</u> ey performance <u>i</u> ndicator
LDL:	<u>L</u> ogistikdienst <u>l</u> eister
LTBG:	<u>L</u> iefer <u>t</u> ermin <u>b</u> eeinflussungsgrad
m-VRSP-*:	VRSP mit Fuhrparkbegrenzung auf <u>m</u> Fahrzeuge
MSA:	<u>M</u> ultiple <u>S</u> cenario <u>A</u> nsatz
TOP:	<u>T</u> eam <u>O</u> rienteering <u>P</u> roblem

OR:	<u>O</u> perations <u>R</u> esearch
OP:	<u>O</u> rienteering <u>P</u> roblem
ORA:	<u>O</u> nline <u>R</u> oute <u>P</u> lanning <u>A</u> pproach
PAA:	<u>P</u> lanung <u>a</u> nte <u>A</u> vis Ansatz
PERI:	<u>P</u> eripherie <u>e</u> ignung
PPA:	<u>P</u> lanung <u>p</u> ost <u>A</u> vis Ansatz
PRÄZ:	<u>P</u> rä <u>z</u> ision (Breite des zugesagten Zeitfensters)
r-opt:	eine Reihenfolgeoptimierungsheuristik
REA:	<u>R</u> eduzierter <u>E</u> rfüllungsgrad <u>A</u> nsatz
RLV:	<u>R</u> egional <u>l</u> ager <u>V</u> ariante
RTA:	<u>R</u> ahmen <u>t</u> our <u>A</u> nsatz
SAV:	<u>S</u> atelliten <u>V</u> ariante
SC:	<u>S</u> tochastic <u>C</u> ustomers (vgl. UC)
SCM:	supply chain management
TGTR	<u>T</u> ourgebiets <u>t</u> reue
TKEF:	<u>T</u> ransport <u>k</u> osteneffizienz
TLRZ:	<u>T</u> oler <u>z</u> anz
TSP:	<u>T</u> ravelling <u>S</u> alesman <u>P</u> roblem
TW:	Zeitfenster von engl. „ <u>t</u> ime <u>w</u> indow“
UC:	<u>U</u> ncertain <u>C</u> ustomers (vgl. SC)
VRSP:	<u>V</u> ehicle <u>R</u> outing and <u>S</u> cheduling <u>P</u> roblem
VRP:	<u>V</u> ehicle <u>R</u> outing <u>P</u> roblem

Variablen

#:	Anzahl
A:	Flächeninhalt
$A_{i,j}$:	Aggregat des Tourtemplates i an Stelle j der Aggregatfolge des Templates
$D(A_{i,j})$:	Bedienzeitfenster des Aggregats $A_{i,j}$
D :	Toleranz, d.h. der Kunde akzeptiert eine Lieferung gdw. $ta_j \in [tp_j - D, tp_j + D]$
d :	Präzision, d.h. eine Lieferung erfolgt rechtzeitig gdw. $ta_j \in [tp_j - d, tp_j + d]$
dur :	maximale Einsatzzeit eines Fahrzeugs
$E(\dots)$:	Erwartungswert von ...
$f(\dots)$:	Funktion von ...
fz_{ij} :	Fahrzeit von Kunde i zu Kunde j
G :	Gleichzeitigkeitsmaß

- i : Zählparameter
 j : Zählparameter
 K_{ges} : Gesamtkosten
 K_{transp} : Transportkosten
 $K_{opp(ERG2)}$: Opportunitätskosten für Verletzung des ERG2
 l_{ges} : Gesamttourlänge / Gesamttourkilometer
 M : Menge der verfügbaren Fahrzeuge
 M^∞ : sehr große Zahl
 m : Anzahl der eingesetzten Fahrzeuge
 $m_{max} = \lfloor M \rfloor$: Anzahl maximal verfügbarer Fahrzeuge
 $max(\dots)$: Maximum von ...
 n : (allgemein) eine ganze Zahl
 n : Anzahl der besuchten Kunden bzw. der besuchbaren Kunden ($n \leq N$)
 n° : Anzahl der mit einem definierten Fuhrpark maximal besuchbaren Kunden (Annahme: Alle Kunden liegen auf einem Punkt => keine Fahrzeiten nötig)
 N : Anzahl der eine Lieferung wünschenden Kunden
 O_r : Auftrag (Index denotiert lexikografisch die Auftragseingangsreihenfolge)
 Q : beliebige Teilmenge von Kunden; $Q \subseteq V - \{0\}$
 $p(i)$: Vorangänger („predecessor“) von i
 r : Radius
 R^2 : Korrelationskoeffizient
 S : Sättigung
 sz_j : Bediendauer des Kunden j
 $T(t)$: Anzahl der Touren, die zum Zeitpunkt t bereits begonnen haben, aber noch nicht beendet sind
 t : Zeit bzw. Zeitpunkt
 t_{ges} : Gesamttourfahrzeit
 t_j : Anfangszeitpunkt der Kundenbelieferung j in Bezug auf Tourstartzeit $t_0 = 0$
 Δt_j : Abweichung des Belieferungszeitpunkts vom zugesagten Zeitfenster
 ta_j : Avisierter Lieferzeitpunkt beim Kunden j
 $tf_j = ta_j - d$: Früheste Ankunftszeit beim Kunden j
 $to(j)$: Zeitpunkt der Bestellung (order placement)
 tp_j : bevorzugter Lieferzeitpunkt von Kunde j
 $ts_j = ta_j + d$: Späteste Ankunftszeit beim Kunden j
 TW : Zeitfenster von engl. „time window“; i.d.R. gilt: $TW = 2d$
 V : Menge aller zu besuchenden Kunden inkl. Depot (Index 0); $\lfloor V \rfloor = n + 1$

-
- $V-X^t$: Menge der bekannten Aufträge zum Zeitpunkt t
- wz_i : Wartezeit beim Kunden j
- X^t : Menge der unbekannteren Aufträge zum Zeitpunkt t
- x_{ij} : Binärvariable zur Nachfolgerbeziehung mit
 $x_{ij}=1$, falls Kunde j unmittelbar nach Kunde i besucht wird
 $x_{ij}=0$, sonst
- $x_{ij}(t)$: Binärvariable, die bestimmt, ob zur Zeit t ein Fahrzeug von i nach j unterwegs ist
- x : x -Koordinate im R^2
- y : y -Koordinate im R^2
- Δ : ein beliebiges Delta (ein beliebiger Unterschied zwischen zwei Größen)
- ε_j : Abweichung des Belieferungszeitpunkts für Kunden j vom zugesagten Zeitfenster
- λ : Wichtungparameter
- μ : Wichtungparameter
- ν : Wichtungparameter
- ϖ : Wichtungparameter

Inhalt

1	Einleitung	11
1.1	Motivation der Arbeit	12
1.2	Untersuchungskontext	13
1.3	Untersuchungsgegenstand	14
2	Zielsetzung, Vorgehen & Methodik	17
2.1	Zielsetzung.....	17
2.2	Vorgehen	18
2.3	Methoden	20
2.3.1	„ceteris paribus“-Vergleiche.....	20
2.3.2	„ceteris paribus“-Teilaspekt-Vergleiche.....	21
2.3.3	Nichtlineare Regression	22
2.3.4	Ordinale Wertprofile.....	22
2.3.5	Eingesetzte Software.....	23
3	Logistischer Kontext	25
3.1	Supply Chain Varianten	25
3.1.1	Die Regionallager-Variante RLV.....	26
3.1.2	Die Satelliten-Variante SAV	27
3.1.3	Die Filialgeschäft-Variante FGV.....	28
3.1.4	Vor- und Nachteilsbetrachtung	28
3.2	Kriterien zur Bewertung transportlogistischer Ansätze	29
3.2.1	Toleranz & Lieferterminbeeinflussungsgrad.....	30
3.2.2	Präzision	31
3.2.3	Erfüllungsgrad	32
3.2.4	Avisvorlauf.....	33
3.2.5	Tourgebietstreue	34
3.2.6	Peripherieeignung	34
3.2.7	Transportkosteneffizienz	35
3.3	Übersicht über bestehende transportlogistische Ansätze ...	35
3.3.1	Der Rahmentour-Ansatz RTA.....	36
3.3.2	Der „Planung ante Avis“ Ansatz PAA.....	38
3.3.3	Der „Planung post Avis“ Ansatz PPA.....	39

3.3.4	Der „Reduzierter Erfüllungsgrad“ Ansatz REA.....	40
3.3.5	Der „Filter vor Planung“ Ansatz FPA	41
3.3.6	Der „Online Tourenplanung“ Ansatz ORA	42
3.3.7	Weitere denkbare Ansätze	44
3.3.8	Gegenüberstellung der logistischen Ansätze	44
3.4	Fuhrparkausstattung	46
3.5	Bereitstellungsstrategien.....	46
3.5.1	Warenbereitstellung auf Rahmentourbasis.....	47
3.5.2	Warenbereitstellung auf Tourbasis.....	48
3.5.3	Warenbereitstellung auf Gebietsstellflächen	49
3.6	Zusammenfassung.....	51
4	Stand der Forschung	53
4.1	Einleitung	53
4.1.1	Klassifikation der VRSPs	54
4.1.2	Merkmalsklassen	54
4.1.3	Problemklassen der Transportplanung.....	55
4.1.4	Notation	57
4.1.5	Komplexität	59
4.2	Das VRSP	60
4.2.1	Verfahren für das VRSP.....	60
4.2.2	Verfahren für das VRSP-TW	60
4.2.2.1	Das Verfahren von Rochat & Taillard 1995	61
4.2.2.2	Das Verfahren von Taillard et al 1997	63
4.2.2.3	Das Verfahren von Shaw 1998	65
4.2.2.4	Die Verfahren von Homberger & Gehring 1999	66
4.2.2.5	Die Verfahren von Ibaraki et al 2001	69
4.2.2.6	Das Verfahren von Lau et al 2003	70
4.2.2.7	Das Verfahren von Tavares et al 2003	72
4.2.2.8	Das Verfahren von Bent & van Hentenryck 2004	74
4.2.2.9	Das Verfahren von Lim & Zhang 2005	75
4.2.2.10	Zusammenfassung	77
4.2.3	Verfahren für das VRSP-UD	77
4.2.3.1	Das exakte Verfahren von Gendreau et al	78
4.2.3.2	Die Tabu-Search Heuristik von Gendreau et al	78
4.2.3.3	Das Verfahren von Erera	79
4.3	Das D-VRSP	81
4.3.1	Abgrenzung nicht betrachteter D-VRSPs	82
4.3.2	Verfahren für das D-VRSP-UC.....	82
4.4	Das OP	84

4.4.1 Verfahren für das OP	84
4.4.1.1 Stochastischer Ansatz von Tsiligirides	84
4.4.1.2 Deterministischer Ansatz von Tsiligirides	85
4.4.1.3 Center of gravity heuristic von Golden	85
4.4.1.4 Weitere Verfahren für das Orienteering Problem	86
4.4.2 Verfahren für das TOP	86
4.5 Wertung der Literatur in Bezug auf Anwendungsfall	89
4.6 Zusammenfassung.....	91
5 Modellierung & Verfahren.....	93
5.1 Klassifikation als D-m-VRSP-TW-UC-DO	93
5.2 Formale Beschreibung des Optimierungsproblems	95
5.2.1 Logistische und geschäftsbedingte Nebenbedingungen	95
5.2.2 Charakter der Zielfunktion	96
5.2.3 Modellierung des Problems	97
5.2.4 Alternative Problemmodellierung	100
5.2.5 Zusammenfassung	102
5.3 Überlegungen zur Problemlösung.....	102
5.3.1 Überlegungen zur Ort-Zeit-Struktur	102
5.3.1.1 Gleichzeitigkeit	103
5.3.1.2 Implizite Sequenzialität	105
5.3.2 Überlegungen zur Sättigung	107
5.3.3 Überlegungen zu Verfahren & Parametern.....	107
5.3.4 Überlegungen zur Präzision	108
5.3.5 Überlegungen zur Erreichung des Optimierungsziels	108
5.3.6 Überlegungen zu kompakten Tourgebieten	109
5.3.6.1 Minimale & Maximale Entfernungen	109
5.3.6.2 Durchschnittsentfernungen	112
5.3.7 Überlegungen zu nichtkompakten Tourgebieten.....	116
5.3.8 Überlegungen zum Mehrfahrzeugproblem.....	119
5.3.8.1 Ein Labor Modell	119
5.3.8.2 Auswirkung Raum-Zeitstruktur auf Fahrzeuganzahl	119
5.3.8.3 Schlußfolgerung	124
5.3.9 Überlegungen zur Tourgebietstreue	125
5.4 Das ORA - Verfahren	125
5.4.1 Vorplanung.....	126
5.4.2 Vorschlags- & Buchungsfunktion	129
5.4.2.1 Initiale Buchung	129
5.4.2.2 Weitere Buchungen	130
5.4.2.3 Einfügen „auf Budget“ von Nachbaraggregat	131
5.4.3 Ex ante Kritik des Verfahrens	133

5.4.3.1	Nachteile	133
5.4.3.2	Vorteile	133
5.5	Das FPA-Verfahren.....	134
5.5.1	Nachteile.....	134
5.5.2	Vorteile.....	135
5.6	Zusammenfassung.....	135
6	Messungen und Ergebnisse.....	137
6.1	Exkurs 1 – Instanzen.....	139
6.1.1	Die Solomon Instanzen.....	139
6.1.2	Erweiterung der Solomon Instanzen	140
6.1.3	Die Instanzen ZUE500*.....	140
6.2	Exkurs 2 – Verfahren	142
6.2.1	HSav0.....	142
6.2.2	HSav1.....	142
6.2.3	HIns1	144
6.2.4	HSav1Ins2.....	145
6.3	Exkurs 3 – Wichtungsproblem	146
6.4	Der FPA/PPA-Ansatz (erweiterbare Flotte).....	149
6.4.1	Erste Ergebnisse.....	149
6.4.1.1	PRÄZ und Gesamttourkilometer	150
6.4.1.2	PRÄZ und Fahrzeuganzahl	153
6.4.2	Instanz- & verfahrensübergreifende Generalisierungsansätze.....	155
6.4.2.1	Instanzen Solomon R205# (Eigene Messungen)	156
6.4.2.2	Instanzen Solomon RC201# (Eigene Messungen)	159
6.4.3	Hypothesen zum FPA/PPA	161
6.4.4	Überprüfung der Hypothesen zum FPA/PPA	163
6.4.4.1	R1*	164
6.4.4.2	RC1*	166
6.4.4.3	R2*	168
6.4.4.4	RC2*	170
6.4.4.5	C1* & C2*	171
6.4.5	Zusammenfassung	171
6.4.6	Methodenkritik.....	172
6.5	Der FPA/REA-Ansatz (reduzierter Service)	174
6.5.1	FPA/REA als Relaxation ex ante.....	174
6.5.2	FPA/REA als Relaxation ex inter.....	174
6.5.3	FPA/REA mit „soft time windows“.....	175
6.6	Der ORA Ansatz (gesteuerte Nachfrage).....	177

6.6.1	Basisszenario SZ0.....	177
6.6.2	Auswirkungen der Kundentoleranz	179
6.6.2.1	Toleranz & Sättigung	179
6.6.2.2	Toleranz & Aggregatsgröße	182
6.6.2.3	Toleranz & raum-zeitliche Lage	183
6.6.2.4	Zusammenfassung zur „Toleranz“	183
6.6.3	Kosteneffekte der zugesagten Zeitfenster	183
6.6.3.1	Basisszenarioanalyse	185
6.6.3.2	Geschwindigkeit	189
6.6.3.3	Standzeit	191
6.6.3.4	Angebotsstruktur	193
6.7	Zusammenfassung und Kritik.....	195
6.7.1	Methodenkritik zur Analyse von FPA/PPA.....	196
6.7.2	Methodenkritik zur Analyse von ORA	197
7	Zusammenfassung, Kritik & Forschungsbedarf.....	199
7.1	Zusammenfassung der Arbeit	199
7.2	Kritik der Arbeit.....	202
7.3	Weiterer Forschungsbedarf.....	203
7.3.1	Logistik.....	203
7.3.2	BWL	203
7.3.3	OR.....	204
7.3.4	Angewandte Statistik	205
7.3.5	Interdisziplinär	205
7.4	Ausblick.....	205
8	Quellen	207
Anhang		

1 Einleitung

Durch kleine Sendungsgrößen und eine Vielzahl von Lieferorten ist die „last mile“ - Zustellung insbesondere im B2C Geschäft ein wesentlicher Erfolgs- bzw. Misserfolgsmoment. Dies gilt sowohl für die Anbieter von Heimlieferservices als auch für deren Transportlogistikpartner¹. Die Ansätze zur Umsetzung der Heimbelieferung und die entsprechenden logistischen Prozesse sind dabei sehr unterschiedlich².

Der e-Commerce Kunde sieht von diesen Prozessen nur wenig. Witten [Witten, 2001] vergleicht das Verhältnis von sichtbaren und unsichtbaren Prozessanteilen im e-Commerce mit einem Eisberg.

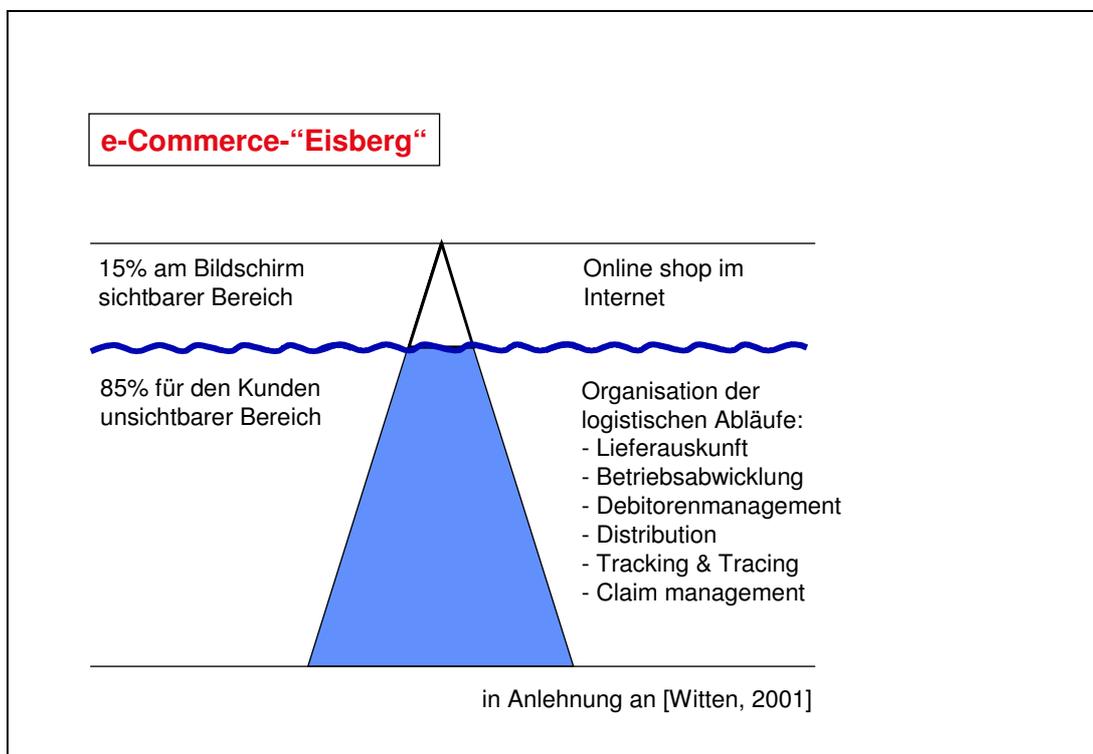


Abbildung 1: eCommerce Eisberg

Es ist zu vermuten, dass das „Platzen der e-Commerce Seifenblase“ im Herbst 2000³ zum Teil auch darauf zurückzuführen ist, dass der Unterwasserteil dieses Eisbergs (und hier insbesondere die „last mile“ - Transportlogistik für die Heimzustellung) im Vergleich zum Überwasserteil (Marketing, Produktspektrum, „web-shop-front-ends“, etc.) offenbar von einigen e-Commerce-Anbietern stark unterschätzt wurde. Am Beispiel der Firma Webvan (USA), die bei einem Jahresumsatz von 16,3 Mio. US\$ einen Verlust von 58 Mio. US\$ einfuhr [Slotta,

¹ Falls die last mile Zustellung durch einen Transportdienstleister erfolgt.

² Auf diese unterschiedlichen Prozesse und Prozessstrukturen wird im Kapitel 3 noch ausführlich eingegangen

³ Indikator für diesen Termin: siehe Kursverläufe von amazon, yahoo oder des Nasdaq im Jahr 2000

2001] wird plastisch, wie sich der Unterwasserteil dieses Eisbergs auf das kommerzielle Firmenergebnis auswirken kann.

Interessanterweise betrieben gleichzeitig andere Akteure im Bereich Heimzustellung diese Geschäft bereits seit Jahren erfolgreich und profitabel (Getränkeheimlieferservices, Möbelheimlieferservices, Versandhandelsheimlieferservices, ...), und wieder andere Akteure haben sich darauf konzentriert, die notwendigen logistischen Voraussetzungen zu schaffen, bevor sie öffentlichkeitswirksam tätig geworden sind.

Ein besonderer Faktor, der sowohl über Markterfolg und -misserfolg eines eCommerce-Lebensmittel-Heimlieferservice entscheidet, als auch die Kosten wesentlich beeinflusst, ist der **Lieferservicegrad**.

Für diese Arbeit sei der Lieferservicegrad folgendermaßen definiert:

*Definition „**Lieferservicegrad im weiteren Sinn**“:*

Der Lieferservicegrad im weiteren Sinn beschreibt die zeitlichen Dimensionen einer Lieferung in Hinblick auf

- ▶ *die Zeitspanne zwischen Bestellung und Lieferung*
- ▶ *die Avisierungsgenauigkeit der Lieferung (Breite des zugesagten Lieferzeitfensters)*
- ▶ *die Einhaltung des avisierten Lieferzeitfensters*
- ▶ *die Art und Weise, wie der Kunde das Lieferzeitfenster beeinflussen kann (Wunschlieferzeitpunkt, Auswahl aus verschiedenen Lieferzeitsvorschlägen, keine Beeinflussung des Lieferzeitfensters = „Zwangsbeglückung“ des Kunden)*

*Definition „**Lieferservicegrad im engeren Sinn**“ (Präzision):*

Der Lieferservicegrad im engeren Sinn sei die Avisierungsgenauigkeit der Lieferung (Breite des zugesagten Lieferzeitfensters). Er wird in dieser Arbeit auch „Präzision“ genannt.

1.1 Motivation der Arbeit

Die transportlogistischen Planungs- und Ausführungsprozesse bilden besonders im Heimlieferservice das wesentliche Bindeglied zwischen **Lieferservicegrad** und **Kosten**.

Slotta [Slotta, 2001] konstatiert, dass sowohl „... die Logistik über den Erfolg oder Misserfolg des elektronischen Handelns entscheidet ...“ als auch „... e-Commerce seinerseits [...] für die Logistikunternehmen zu einem entscheidenden Erfolgsfaktor avanciert“ ist.

Auf Grund dieser Interdependenz haben sowohl Heimlieferserviceanbieter (und insbesondere B2C-e-Commerce-Betreiber) als auch Logistik Service Betreiber (und hier insbesondere die KEP-Dienstleister) ein originäres Interesse an den für das jeweilige Geschäftsmodell optimierten Prozessen & Planungsverfahren.

Viele Veröffentlichungen beschreiben die Abhängigkeiten von Servicegrad (im oben definierten Sinn) und Transportkosten entweder gar nicht, nur implizit oder nur qualitativ. Oft ist der oben definierte Servicegrad nicht explizit beschrieben, sondern wird unter den Überbegriffen „Mehrwertleistungen“, „added value“ oder allgemein „Servicequalität“ subsummiert (vgl. [Gatzke, 2001]).

Diese Arbeit soll daher einen Beitrag zum Verständnis der Zusammenhänge zwischen Lieferservice und Transportkosten und zur **Quantifizierung** der Abhängigkeiten von Lieferservicegrad und Transportkosten für den Lebensmittelheimlieferservice leisten.

1.2 Untersuchungskontext

Der engere Kontext dieser Arbeit, die „last mile“ Transportlogistik, lässt sich in den weiteren Kontext des „supply chain managements“ (SCM) und des „customer relation managements“ (CRM) einordnen. In Anlehnung an die Übersicht von Witten [Witten, 2001] zum Zusammenspiel von SCM und CRM, lässt sich die „last mile“ Transportlogistik folgendermaßen positionieren:

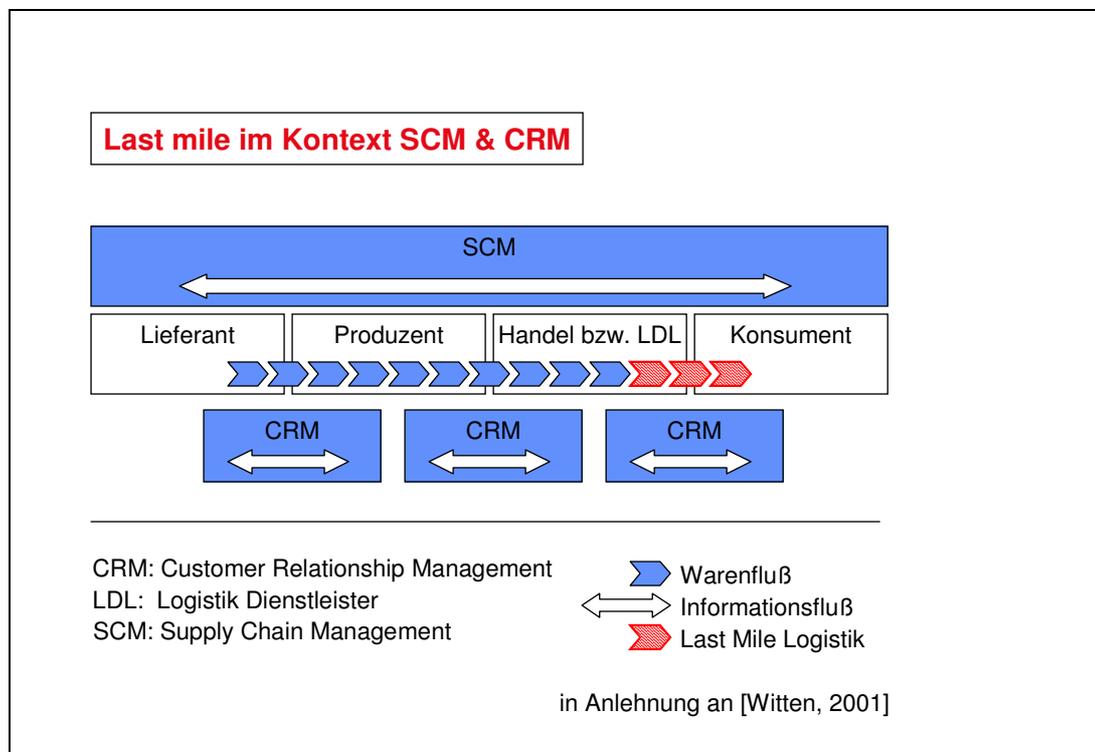


Abbildung 2: last mile Logistik im SCM-CRM-Kontext

Aus dieser vereinfachten Abbildung nach Witten [Witten, 2001] wird jedoch noch nicht intuitiv klar, dass drei Akteure, nämlich der Kunde, der Serviceanbieter und der Logistikdienstleister bei der Lieferzeitpunktzusage eng interagieren sollten, um Kosten-Service-optimierte „last mile“ Transporte zu ermöglichen⁴.

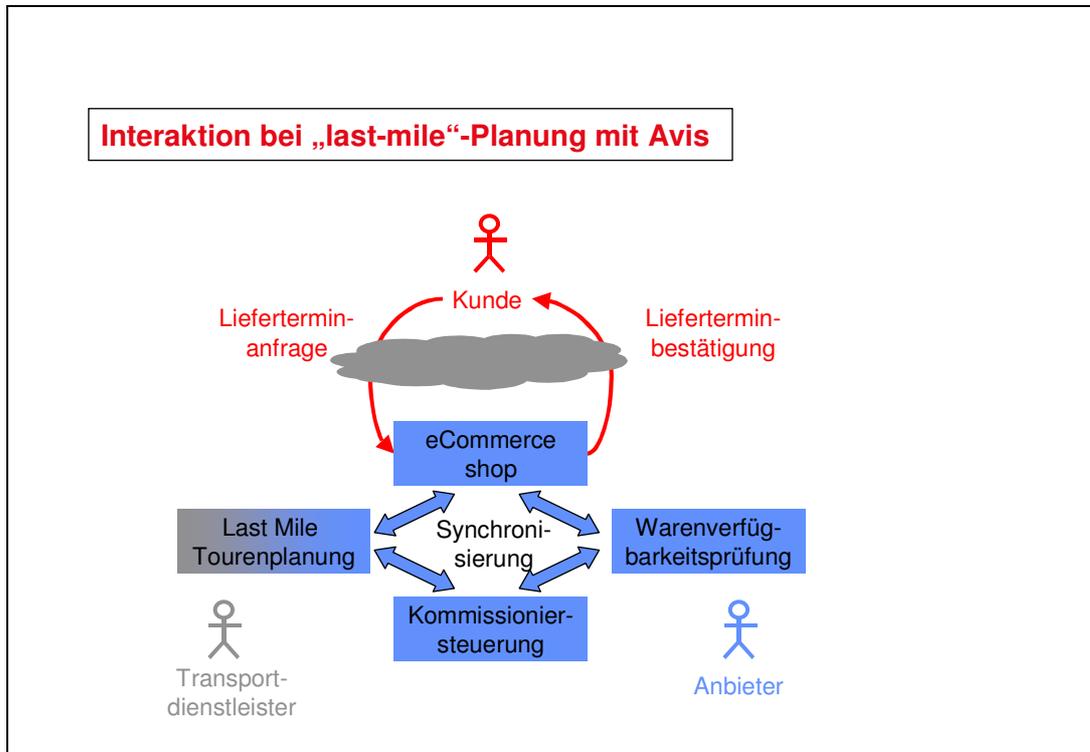


Abbildung 3: Interaktion Kunde, Transportdienstleister, Anbieter

In dieser Arbeit wird der Interaktion mit dem Kunden zum Bestellzeitpunkt ein besonderes Augenmerk geschenkt, da diesem Aspekt insbesondere in der OR-Literatur als Einflussfaktor und Optimierungsfreiheitsgrad bisher wenig Beachtung geschenkt wird.

Die Planung der „last mile“ Transporte ist ein notwendiger Schritt, der unabhängig von der jeweiligen „supply chain“ Ausprägung erfolgen muss, wenn die Zustellung nicht komplett an einen KEP-Dienstleister übergeben wird, der sie über sein Regelverkehrsnetz (d.h. ohne dedizierte „last mile“ Transporte) ausführt, wie im Beispiel der Bücher- oder CD-Zustellung von amazon.de.

1.3 Untersuchungsgegenstand

Der Untersuchungsgegenstand dieser Arbeit ist die Heimzustellung von Lebensmitteln. Der Untersuchungsschwerpunkt dieser Arbeit liegt dabei auf der Betrachtung der Interdependenz zwischen den „last mile“ Transportkosten und dem

⁴ Beispiele für eine diesbezüglich unzureichende Interaktion sind heute noch oft bei der Heimzustellung von Weiß- oder Braunware in praxi zu erleben.

Serviceangebot in Hinblick auf ein präzises Avis der Lieferung bereits zum Bestellzeitpunkt (Präzision⁵).

Der Untersuchungsschwerpunkt kann jedoch nicht losgelöst vom Untersuchungskontext (= eCommerce-Lebensmittel-Heimlieferservice) betrachtet werden. Daher werden im Folgenden immer wieder Querbeziehungen zum Untersuchungskontext hergestellt werden, insbesondere bei der Formulierung von Zielfunktionen, der Definition von Nebenbedingungen, der Bewertung von Verfahrensprinzipien, und der Bewertung von Ergebnissen.

⁵ Der Begriff Präzision ist ein zentraler Begriff, und wird unten noch präziser definiert.

2 Zielsetzung, Vorgehen & Methodik

Motiviert wurde diese Arbeit durch die Frage „Was kosten präzise vorausgesagte Zeitfenster?“ oder besser „Wie wirkt sich ein im Moment des Auftragseingangs präzise zugesagter Liefertermin auf die Transportkosten aus, und wie kann man einen solchen Service mit überschaubaren Kosten anbieten?“.

Intention dieser Arbeit ist es, Antworten auf diesen praxisrelevanten Fragenkomplex in der Logistik zu finden. Um einleitend einen Überblick zu schaffen, was warum wie untersucht wurde, werden in diesem Kapitel die Ziele dieser Arbeit, das Vorgehen und die angewandten Methoden beschrieben.

2.1 Zielsetzung

Die Ziele dieser Arbeit sind:

- ▶ **Verständnis** bezüglich der aus den vertrieblichen und transportlogistischen Rahmenbedingungen resultierenden Anforderungen an einen Lösungsansatz und dessen Freiheitsgrade
Dazu wird der Kontext des Problems ausgelotet, und verschiedene logistische Aspekte untersucht.
- ▶ **Analyse des Verhaltens von klassischen Verfahren** zum VRSP-TW⁶, unter engen Zeitfenstern
Um problemspezifische von verfahrensspezifischen Einflussfaktoren unterscheiden zu können ist ein gutes Verständnis für die Tourenplanung unter engen Zeitfenstern Voraussetzung.
- ▶ **Formale Modellierung** des speziellen transportlogistischen Problems der Lebensmittelheimzustellung mit präziser Lieferzusage zum Bestellzeitpunkt als OR⁷-Problem
Durch diese formale Formulierung wird die Basis für die Verfahrensauswahl geschaffen.
- ▶ **Herausarbeiten von Kriterien**, die bei der Problemlösung relevant sind
Dazu werden verschiedene ordinale Kriterien aus den logistischen und wirtschaftlichen Anforderungen und Rahmenbedingungen definiert, die einen Kriterienkatalog für die Verfahrensauswahl bilden.
- ▶ **Bewertung von Heuristiken** zur Lösung des Planungsproblems anhand dieses Kriterienkatalogs
Auf Grund der Bewertung der denkbaren Verfahren anhand der definierten

⁶ VRSP-TW := vehicle routing and scheduling problem with time windows / Tourenplanungsproblem mit Zeitfenstern

⁷ OR := operations research

Kriterien wird die Verfahrensauswahl für die tiefere Untersuchung begründet

- ▶ **Empirische Untersuchung** der Auswirkungen wesentlicher Einflussgrößen auf die Planungsergebnisse der ausgewählten Verfahren

So werden schließlich quantitative Aussagen zu den Kosten-Nutzen-Abhängigkeiten im Spannungsfeld von Servicegradzusagen (Lieferpräzision) und Transportkosten getroffen.

2.2 Vorgehen

Zur Erreichung dieser Ziele wurde diese Arbeit folgendermaßen strukturiert.

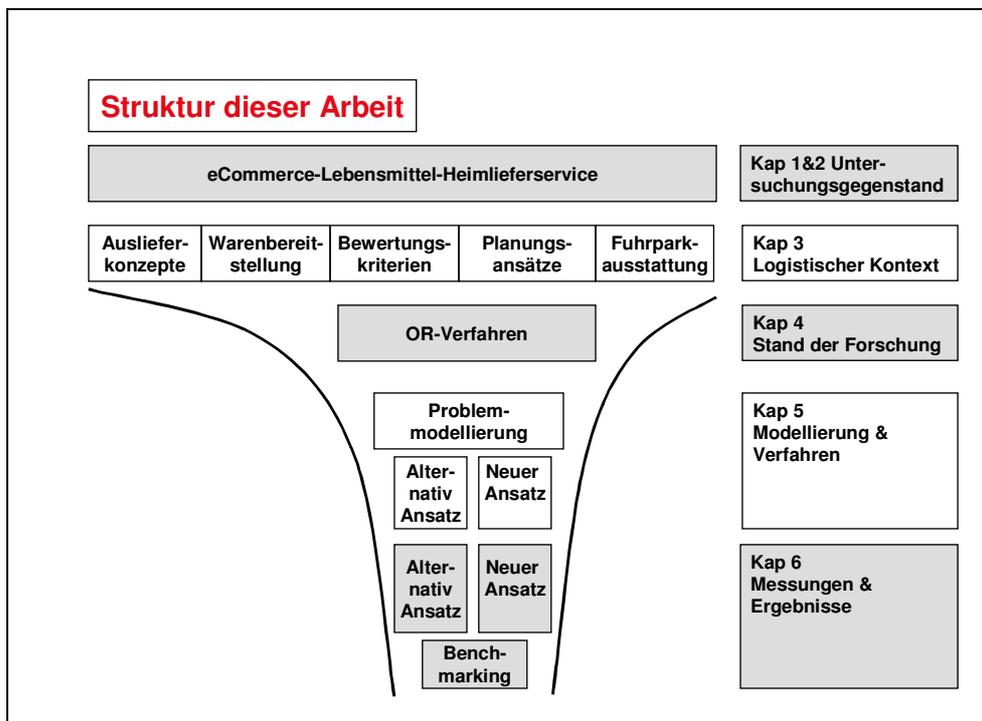


Abbildung 4: Struktur dieser Arbeit

Im **Kapitel 3 „Logistischer Kontext“** wird zunächst die transportlogistische Aufgabenstellung für den eCommerce-Lebensmittel-Heimlieferservice aus verschiedenen Sichten analysiert, und die Auswirkungen dieses Kontextes auf die Planung herausgearbeitet. Untersucht werden im einzelnen:

- ▶ **Auslieferungskonzepte** – hier werden verschiedene in der Praxis vorherrschende Varianten der „last mile supply chain“ betrachtet.
- ▶ **Warenbereitstellung** – hier wird die Interdependenz zwischen Tourenplanung und Warenbereitstellung untersucht.
- ▶ **Fuhrpark** – hier wird kurz auf typische Fuhrparkeigenschaften eingegangen, die wichtig für die Planungsfreiheitsgrade sind.

-
- ▶ **Bewertungskriterien** – hier werden sowohl Anforderungen des Marktes als auch logistische Anforderungen als Bewertungskriterien aufgestellt.
 - ▶ **Planungsansätze** – hier werden mögliche Planungsansätze kategorisiert, die später bei der Verfahrensauswahl wesentlich sind.

Kapitel 4 „Stand der Forschung“ zeigt eine Übersicht über mögliche Planungsverfahren und deren Prinzipien. Dabei werden zum einen die Planungsverfahren für den klassischen Planungsansatz erklärt und bewertet, und zum anderen wird aufgezeigt, dass es für den neuen Planungsansatz bisher noch keine maßgeschneiderten Verfahren gibt.

Im **Kapitel 5 „Modellierung & Verfahren“** wird das Problem unter Verwendung der Ergebnisse aus Kapitel 3 zunächst als neues formales OR-Problem D-m-VRSP-TW-UC-DO⁸ modelliert. Ausgehend von der Modellierung wird der neue Planungsansatz schrittweise abgeleitet und analysiert. Die Schritte dazu sind:

- ▶ Überlegungen an Laborbeispielen zum Verständnis des Problemtyps für das Ein-Fahrzeug-Problem und das Mehr-Fahrzeug-Problem
- ▶ Deduktion eines neuen Verfahrens (ORA⁹) aus den Überlegungen
- ▶ Theoretischer Nachweis der Eignung des neuen Verfahrens für ausgewählte Laborbeispiele
- ▶ Herausarbeitung der Einflussfaktoren, die die Ergebnisqualität des neuen Verfahrens determinieren
- ▶ Empirische quantitative Untersuchung dieser Einflussfaktoren (ceteris paribus)

In **Kapitel 6 „Messungen & Ergebnisse“** werden schließlich quantitative Ergebnisse für die zwei ausgewählten Planungsansätze generiert.

Für den klassischen Ansatz wird gezeigt, dass der typische Zusammenhang zwischen der Servicekenngröße „Avispräzision“ und den Transportkosten im Planungsansatz, und nicht in den Planungsverfahren begründet liegt.

Für den neuen Ansatz wird gezeigt, dass der für den klassischen Ansatz typische Zusammenhang zwischen der Servicekenngröße „Avispräzision“ und den Transportkosten durchbrochen werden kann.

Schließlich werden beide Ansätze (soweit möglich) miteinander verglichen.

⁸ D-m-VRSP-TW-UC-DO := dynamic m vehicle routing and scheduling problem with time windows, uncertain customers and delivery note at order placement.

⁹ ORA := Online Routing Approach

2.3 Methoden

Die folgenden Methoden wurden zur Erreichung der oben formulierten Ziele eingesetzt.

2.3.1 „ceteris paribus“-Vergleiche

Die wohl wichtigste Bewertungsmethode in dieser Arbeit ist die des „ceteris paribus¹⁰“-Vergleichs. Warum? In dieser Arbeit wird ein neues OR-Problem für einen neuen komplexen transportlogistischen Anwendungsfall definiert. Während sich für kleine, konstruierte Laborbeispiele Optimalität noch relativ einfach zeigen lässt, ist bereits für einfachste „real-life“ Szenarien der Nachweis von Optimalität aus drei Gründen schon nicht mehr ohne weiteres möglich.

Zum ersten wegen der *NP-schwierigen Komplexität* des Problems, die, wenn überhaupt, einen Optimalitätsnachweis nur nach extrem langen Rechenzeiten ermöglicht.

Zum zweiten wegen der *Vielfalt der Einflussfaktoren*, denn bereits die Änderung eines einzigen Einflussfaktors kann ein anderes Optimum bedeuten.

Zum dritten wegen des *Zielkonflikts verschiedener Optimierungsziele*, die nicht oder nur sehr schwer gegeneinander gewichtet werden können, beispielsweise wenn die einen Ziele kurzfristig sind (Maximiere die Anzahl heute zu bedienenden Kunden) und andere Ziele langfristig sind (Minimiere die Abweichung von zugesagten Lieferintervallen). Allein schon die quantitative Gewichtung dieser Ziele gegeneinander ist beliebig schwierig. Gleichzeitig sind die „Optimierungsobjekte“ inkompatibel (zum einen die kurzfristig stochastisch eingehenden, aber konkreten Bestellungen, die angenommen oder abgelehnt werden müssen, zum anderen der langfristige marketingstrategische Einflussfaktor „Glaubwürdigkeit bei Servicezeitzusagen zur langfristigen Kundenbindung“).

Der „ceteris paribus“-Vergleich ist eine Methode, die zwar keine allgemeingültigen Schlüsse über Optimalität zulässt, aber dafür quantitative Aussagen zur Ergebnisgüte für eine konkrete Problemausprägung in Relation zu anderen Ansätzen erlaubt. Dieses Prinzip wurde für den Themenkomplex der Tourenplanung von u.a. von [Solomon, 1984] durch die von ihm eingeführte Instanzenbildung¹¹ begründet, und wird in der einschlägigen Literatur extensiv genutzt.

¹⁰ ceteris paribus [lat.] – „unter sonst gleichen Bedingungen“, d.h. alle nicht explizit veränderten Einflussgrößen bleiben unverändert.

¹¹ Eine Solomon-Instanz ist ein Problem-Szenario mit definierten Eigenschaften, die ceteris paribus unverändert bleiben.

2.3.2 „ceteris paribus“-Teilaspekt-Vergleiche

Für direkte „ceteris paribus“-Vergleiche zur Bewertung von Tourenplanungsergebnissen müssen Referenzlösungen für ein definiertes Problem vorliegen. Bei neuen Problemen gibt es definitionsgemäß noch keine bestehenden Referenzlösungen, gegen die getestet werden kann. Möglicherweise liegen aber Ergebnisse vor (oder lassen sich generieren), die zumindest für einen Teilaspekt der Lösung als Referenz verwendet werden können. Am besten wird das durch folgendes Beispiel klar.

Beispiel:

Zu lösen sei ein Einfahrzeug-Orienteeing Problem OP^{12} . Eine Lösung L für das OP wurden gefunden, mit der ein „score“ von x bei einer Fahrzeit k erzielt werden kann. Diese Lösung L des OP kann ihrerseits als ein „Travelling Salesman Problem“ (TSP) mit x Stops betrachtet werden. Angenommen, die exakte Lösung L' des TSP verursacht eine Fahrzeit in der Höhe von k' .

Ist $k=k'$, so kann zwar nicht davon ausgegangen werden, dass L eine optimale Lösung des OP ist, aber es kann davon ausgegangen werden, dass L zumindest transportkosteneffizient ist, d.h. es gibt keine günstigere Lösung, die x Stops zu bedienen.

Ist $k>k'$, so kann davon ausgegangen werden, dass L nicht transportkosteneffizient ist (d.h. es gibt eine günstigere Lösung, die x Stops zu bedienen).

Das Verhältnis von k/k' ist somit ein Indikator dafür, ob die gefundene Lösung L noch verbessert werden könnte oder nicht. So lässt sich zwar nicht sagen, ob eine Lösung L optimal ist, aber $k/k' \gg 1$ ist ein guter Indikator dafür, dass L suboptimal ist, da L „ungenutztes“ Fahrzeitpotenzial aufweist.

Streng wissenschaftlich betrachtet ist der „ceteris paribus“-Teilaspekt-Vergleich somit zwar kein verlässliches Instrument zur Bestimmung der Lösungsgüte, wohl aber ein guter Indikator für eine schlechte Lösungsqualität. Erkenntnisse aus „ceteris paribus“-Teilaspekt-Vergleichen haben (wie beim Testen von Hypothesen, wo Hypothesen auch nicht empirisch belegt, wohl aber abgelehnt werden können) jedoch bereits einen wissenschaftlichen Aussagewert an sich.

¹² ZF: max. Anzahl Stops unter der Nebenbedingung: Tourdauer begrenzt.

2.3.3 Nichtlineare Regression

Im Kapitel 6 dieser Arbeit werden die in verschiedenen Quellen veröffentlichten Ergebnisse anderer Autoren aus einem neuen Blickwinkel mit Hilfe von nichtlinearer Regression ausgewertet, um von verfahrensspezifischen Eigenschaften abstrahieren und so problemspezifische Eigenschaften herausarbeiten zu können.

Dazu wird der Ansatz der „Nonlinear Least Squares Minimization“ mit Hilfe des Levenberg-Marquardt-Algorithmus verwendet, der auch bei nichtlinearem Kurvenverlauf geeignet ist, die Abweichungsquadrate zwischen Messwerten und Regressionsfunktion heuristisch zu minimieren. Der Levenberg-Marquardt-Algorithmus ist nach [Lourakis 2005] eine iterative Technik zur Bestimmung des Minimums einer multivariaten Funktion. Sie kann nach [Ranganathan 2004] als Kombination der Strategie „Folge auf der Minimumsuche dem steilsten fallenden Gradienten“¹³ und der Gauss-Newton-Methode gesehen werden.

Das zu lösende Problem ist dabei folgendes:

$$\min f(x) = \frac{1}{2} \sum_{j=1}^m r_j^2(x) = \frac{1}{2} \|r(x)\|^2 \text{ mit}$$

r_j : Funktion $\mathbb{R}^n \rightarrow \mathbb{R}$ (Residuen)

$r(x) = (r_1(x), r_2(x), \dots, r_m(x))$ (Residuenvektor)

$x = (x_1, \dots, x_n)$: Parameter-Vektor, der die (aktuelle) Belegung der Parameter darstellt

x_{\min} : gesuchtes Parametertupel mit $f(x_{\min}) = \min_x f(x)$

Details zur Lösung dieses Problems mit Hilfe des Levenberg-Marquardt-Algorithmus finden sich in [Levenberg 1944], [Marquardt 1963], [Ranganathan 2004], [Lourakis 2005] und [Roweis 2005].

2.3.4 Ordinale Wertprofile

Um den Transfer der theoretischen Überlegungen auf die logistische Praxis zu erreichen, werden ordinale Wertprofile nach [Arnold & Furmans, 2005, S267 ff] verwendet.

Diese Wertprofile bringen sowohl wissenschaftlich als auch pragmatisch gewonnene Erkenntnisse in eine Form, die den Entscheidern über logistische Prozesse helfen, die wesentlichen Weichenstellungen zu treffen.

¹³ „steepest descent“ [Lourakis 2005] bzw.

„vanilla gradient descent“ (vanilla [engl.] = einfach; ohne Schnickschnack) [Ranganathan 2004]

2.3.5 Eingesetzte Software

Zur Generierung quantitativer Planungsergebnisse wurden die folgenden Produkte der Firma PTV AG eingesetzt:

- ▶ für FPA (Details s.u.):
 - ▶ das Tourenplanungssystem „ptv INTERTOUR“
 - ▶ in Kombination mit einer Filterfunktion
- ▶ für ORA (Details s. unten):
 - ▶ eine Implementierung des ORA-Planungsverfahrens („ptv ZSR-Server“)
 - ▶ ein Tool zur Erzeugung der Vorplanung („ptv Map&Market“ zusammen mit „ptv INTERTOUR“)

Zur Berechnung der nichtlinearen Regressionen wurde die Software „OriginPro v7.5G“ der Firma OriginLab Corporation, eingesetzt. Die Software nutzt für die nichtlineare Regression den Levenberg-Marquardt-Algorithmus.

3 Logistischer Kontext

In diesem Kapitel sollen der logistische Kontext und die wesentlichen logistischen Aspekte aus der Praxis beschrieben werden, die bei der „last mile“-Abwicklung des Lebensmittelheimzustellservice berücksichtigt werden müssen.

3.1 Supply Chain Varianten

In der Praxis des Lebensmittelheimzustellservice gibt es in Hinblick auf die „last mile“ der „supply chain“ im wesentlichen drei Varianten. Diese sollen hier kurz vorgestellt werden.

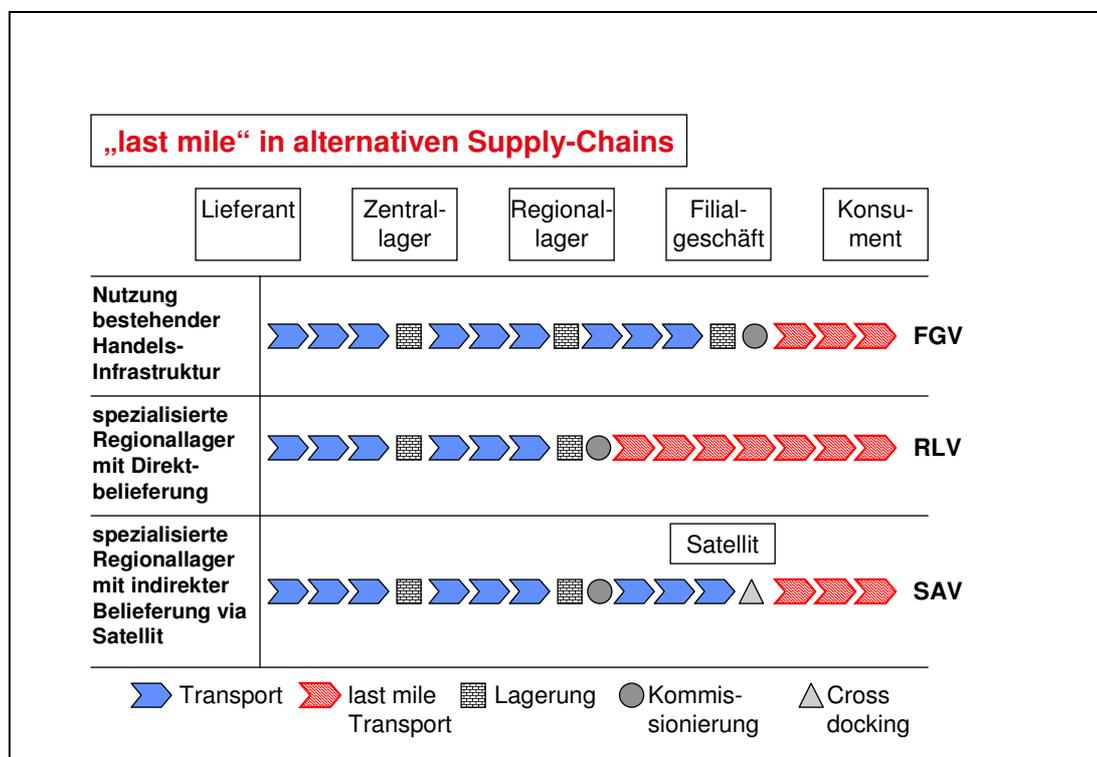


Abbildung 5: „last mile“ Logistik im Kontext verschiedener Supply-Chain Varianten der dedizierten Heimgestellung.

Die obenstehende Grafik skizziert drei wesentliche praxisrelevante Varianten der last mile Logistik, bei denen entweder der Anbieter oder der von ihm beauftragte Transportdienstleister den „last mile“-Transport organisiert. Diese sind die Filialgeschäftvariante **FGV**, die Regionallagervariante **RLV** und die Satellitenvariante **SAV**. Die Unterschiede zwischen den Varianten werden im Folgenden noch genauer herausgearbeitet.

Die Planungsaufgabenstellungen für die drei Varianten unterscheiden sich vor allem dadurch, zu welchem „Zeitpunkt“ in Bezug auf die Prozessstufen in der

„supply chain“ die „last mile“-Zustellung geplant wird.

Dieser Zeitpunkt ist entscheidend, da zwischen Bestellung und Lieferung die Durchführung der Planung, die Kommissionierung, das gegebenenfalls notwendige „cross docking“ und die „last mile“ Zustellung selbst erfolgen muss, und vorgelagerte Prozessschritte irreversible Rahmenbedingungen für die nachfolgenden Prozessschritte schaffen. Solche irreversiblen Rahmenbedingungen sind z.B. die **Lieferzeitusage** als Input für die Tourenplanung, die **Auftrags-Tourzuordnung** als Input für die tourbezogene Warenbereitstellung und das **Kommissionierende** als notwendige Bedingung für Tourstart.

3.1.1 Die Regionallager-Variante RLV

Bei der Regionallager Variante (RLV) erfolgt die Kommissionierung der einzelnen Kundenbestellungen in dedizierten Regionallagern (oder in dedizierten Bereichen bestehender Regionallager). Dort werden die Waren i.d.R. aus Durchlaufregallagern oder Bodenzeilenlagern¹⁴ in spezielle Ladungsträger¹⁵ kommissioniert.

Zur Gewährleistung der Kühlkette für frische und Tiefkühl-Produkte werden die drei Güterklassen (Trocken, Frisch & Tiefkühl) in getrennten Bereichen kommissioniert und zwischengelagert.

Vom Regionallager aus erfolgt die Zustellung direkt zum Endkunden entweder mit kleinen 3-Kammer-Spezialfahrzeugen mit fester oder flexibler Kühlzonentrennung im Laderaum, oder in ungekühlten Einkammerfahrzeugen mit entsprechender Isolierverpackung der temperatursensitiven Einzelsendungen (z.B. Trockeneis und Styropor).

¹⁴ Vgl. [Arnold, 1995], S. 154

¹⁵ Ladungsträger im Sinne von [Arnold, 1995], S. 62, hier meist Kunststoff- oder Blechkästen mit ebenem Boden.

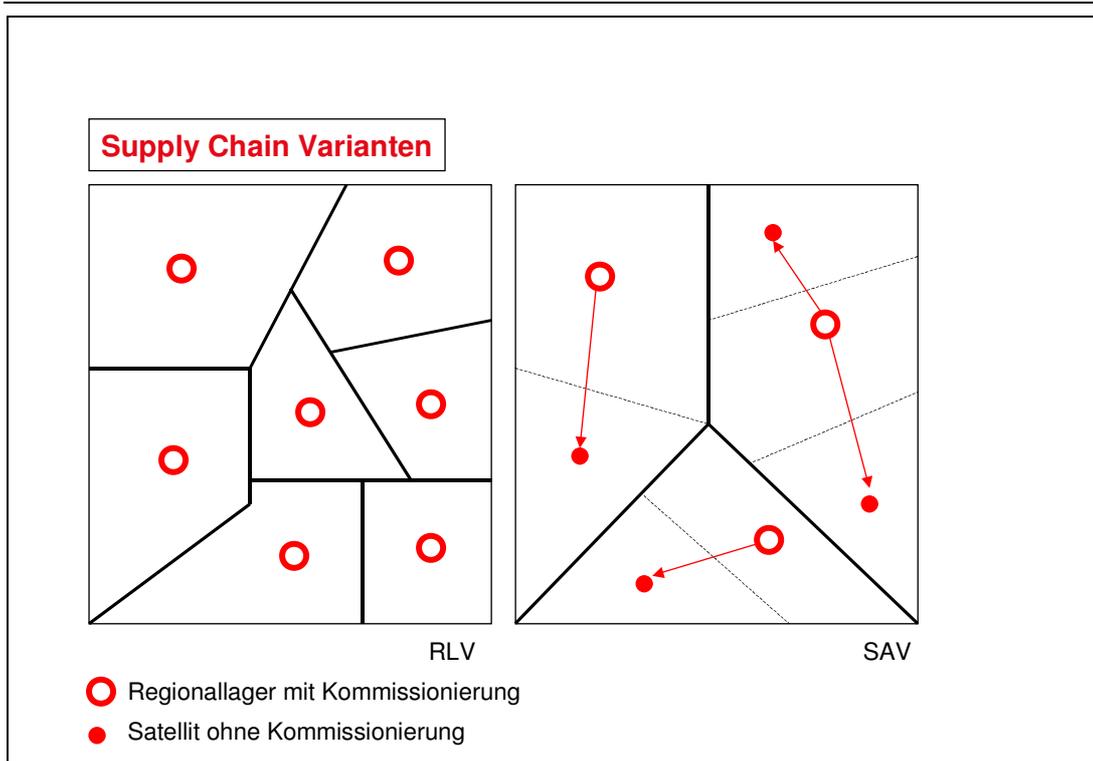


Abbildung 6: Prinzipskizze zum Unterschied zwischen Regionallager Variante (RLV) und Satelliten Variante (SAV).

Der Unterschied zwischen der RLV und der SAV liegt darin, dass bei der RLV die Sendungen direkt vom Regionallager aus an die Kunden gehen, während sie in der SAV gegebenenfalls noch über ein „cross docking“ in einem Satelliten umgeschlagen werden.

3.1.2 Die Satelliten-Variante SAV

Die Kommissionierung bei der Satelliten Variante (SAV) erfolgt i.d.R. auch in Regionallagern (s.o.). Allerdings ist die Anzahl solcher Regionallager in der SAV eher geringer als in der RLV, um die Anzahl der fixkostenintensiven Kommissionierinfrastruktureinrichtungen möglichst gering zu halten. Die stärkere Bündelung der Kommissionieraktivitäten rechtfertigt im Vergleich zur RLV i.d.R. höhere Investitionen in die Kommissionier- und Materialflusstechnik. Gleichzeitig wird ein größeres Zustellgebiet von einem Regionallager (= Kommissionierzentrum) aus versorgt, als in der RLV.

Je größer das Zustellgebiet, desto ineffizienter wird eine Direktzustellung der Endkunden vom Regionallager aus. Ab einer gewissen Zustellgebietsausdehnung ist es daher sinnvoll, Satelliten als „cross docking“ Punkte einzurichten, um Bündelungseffekte in der Transportlogistik zu erzielen. Die Endkundenbelieferung erfolgt in diesem Fall also zweistufig über die Satelliten, und somit umfasst die „last mile“ Logistik in der SAV nur die letzte dieser beiden Stufen.

3.1.3 Die Filialgeschäft-Variante FGV

Eine interessante Variante ist die Filialgeschäft Variante (FGV). Hier wird die bestehende Infrastruktur eines Filialgeschäftsnetzes für den Kommissionierprozess genutzt. Die Kommissionierung erfolgt wenig optimiert so, wie eine Hausfrau beim Einkaufen ihren Einkaufswagen füllt.

Durch die Nähe zum Endkunden, kann die last mile Zustellung gegebenenfalls sogar recht einfach durch PKW erfolgen, vorausgesetzt dass die gesetzlichen Vorschriften zur Aufrechterhaltung der Kühlkette erfüllt werden (z.B. durch Verwendung von gut isolierten Ladungsträgern während der PKW Zustellung).

3.1.4 Vor- und Nachteilsbetrachtung

Die Wahl einer angemessenen Variante hängt im Wesentlichen vom Umsatz bzw. den Umsatzprognosen ab.

In einer Markterschließungssituation mit vorerst geringen Umsätzen ist die **FGV** eine hochinteressante Möglichkeit, ohne hohe Investitionen und ohne lange Vorlaufzeiten einen Heimlieferservice anbieten zu können. Durch die kurze „time to market“ dieses Ansatzes konnte in Einzelfällen ein signifikanter strategischer Wettbewerbsvorteil (schnelle Marktdurchdringung) erzielt werden.

In Ballungszentren mit genügend Umsatzaufkommen ist sicher die **RLV** eine sinnvolle Alternative.

Bei sehr starken Umsätzen kann die **SAV** eine adäquate Lösung sein, die auch die Versorgung von dünn besiedelten Gebieten ermöglicht.

Diese Bewertung lässt sich unter folgenden Annahmen qualitativ illustrieren:

- ▶ Investitionskosten für Regionallager und Satelliten (KI fix):
 $KI\ fix\ (SAV) >^{16} KI\ fix\ (RLV) > KI\ fix\ (FGV)$
- ▶ Kommissionierkosten (KK var):
 $KK\ var\ (SAV) < KK\ var\ (RLV) < KK\ var\ (FGV)$
- ▶ Transportkosten (incl. „cross docking“ Kosten) (KT var)
 $KT\ var\ (SAV) < KT\ var\ (RLV) < KT\ var\ (FGV)$
- ▶ Die Kommissionier- und Transportkosten steigen linear¹⁷ mit dem Umsatz

Qualitativ, und nicht auf Messwerten beruhend, lässt sich dieser Sachverhalt somit folgendermaßen darstellen.

¹⁶ Diese Relation beruht auf der Annahme, dass höhere Investitionen bei der SAV den Automatisierungsgrad der Kommissionierung im Vergleich zur RLV steigern, und daher die Kommissionierstückkosten senken.

¹⁷ Diese Annahme wird aus Vereinfachungsgründen für die qualitative Betrachtung getroffen. In praxi ist eher mit einer degressiven Entwicklung der Kommissionier- und Transportstückkosten zu rechnen.

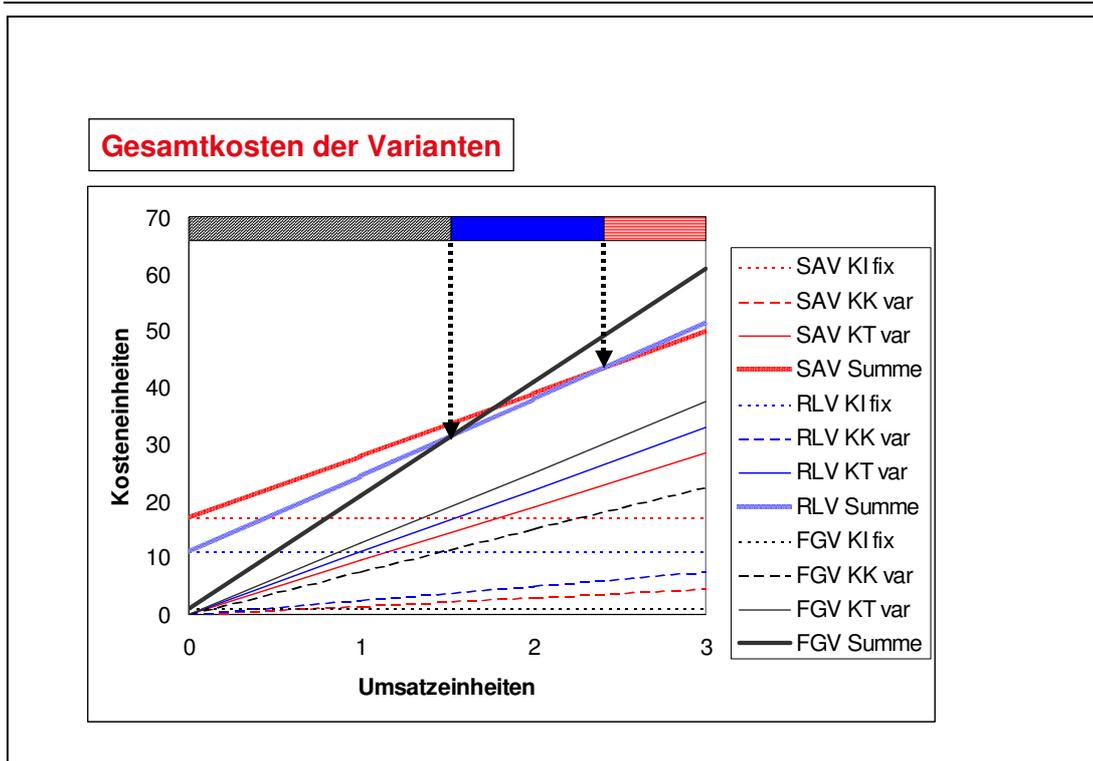


Abbildung 7: Qualitativer Variantenvergleich – Bereiche, in denen die einzelnen Varianten gesamtkostenoptimal sind - schwarz FGV, blau RLV, rot SAV.

Dieser Variantenvergleich ist bisher nicht mit Messwerten untermauert. Er kann daher nur als unbewiesene, wenngleich plausible Hypothese gewertet werden. Eine nähere Untersuchung dieses Themas würde den Rahmen dieser Arbeit sprengen, und ist daher an anderer Stelle zu führen.

3.2 Kriterien zur Bewertung transportlogistischer Ansätze

Neben den verschiedenen „supply chain“ Varianten, gibt es auch in der betrieblichen Praxis unterschiedliche Ansätze zur transportlogistischen Planung und Abwicklung der „last mile“ Transporte. Diese **Ansätze** werden später noch beschrieben und formal benannt.

Um diese verschiedenen Ansätze auf ihre Anwendbarkeit für spezielle Aufgabenstellungen hin zu untersuchen, und um die jeweiligen Vor- und Nachteile gegeneinander abzuwägen, werden hier zunächst für den Lebensmittelheimzustell-service eine Reihe von praxisrelevanten **Kriterien** formal definiert.

3.2.1 Toleranz & Lieferterminbeeinflussungsgrad

Ein potenzieller Kunde eines Lebensmittelheimzustelldienstes hat aus seiner privaten Tagesplanung heraus Restriktionen, die ihn flexibel oder unflexibel in Hinblick auf den Liefertermin machen. Beispielsweise wird er einen Lieferterminvorschlag um 09:30 ablehnen, wenn er von 09:00-10:00 bereits einen Zahnarzttermin vereinbart hat. Ein Kunde hat daher i.d.R. das Bedürfnis, den Liefertermin nach seinen Wünschen beeinflussen zu können. Daher werden die folgenden beiden Kriterien definiert.

Definition „Toleranz“

Toleranz (TLRZ) sei ein Kennwert, der die Toleranz des Kunden in Hinblick auf die Abweichung des zugesagten Lieferzeitpunkts von einem Wunschliefertermin des Kunden beschreibt.

Definition „Lieferterminbeeinflussungsgrad“

Lieferterminbeeinflussungsgrad (LTBG) sei ein Kennwert, der die Möglichkeit des Kunden zur aktiven Steuerung des Zustelltermins beschreibt.

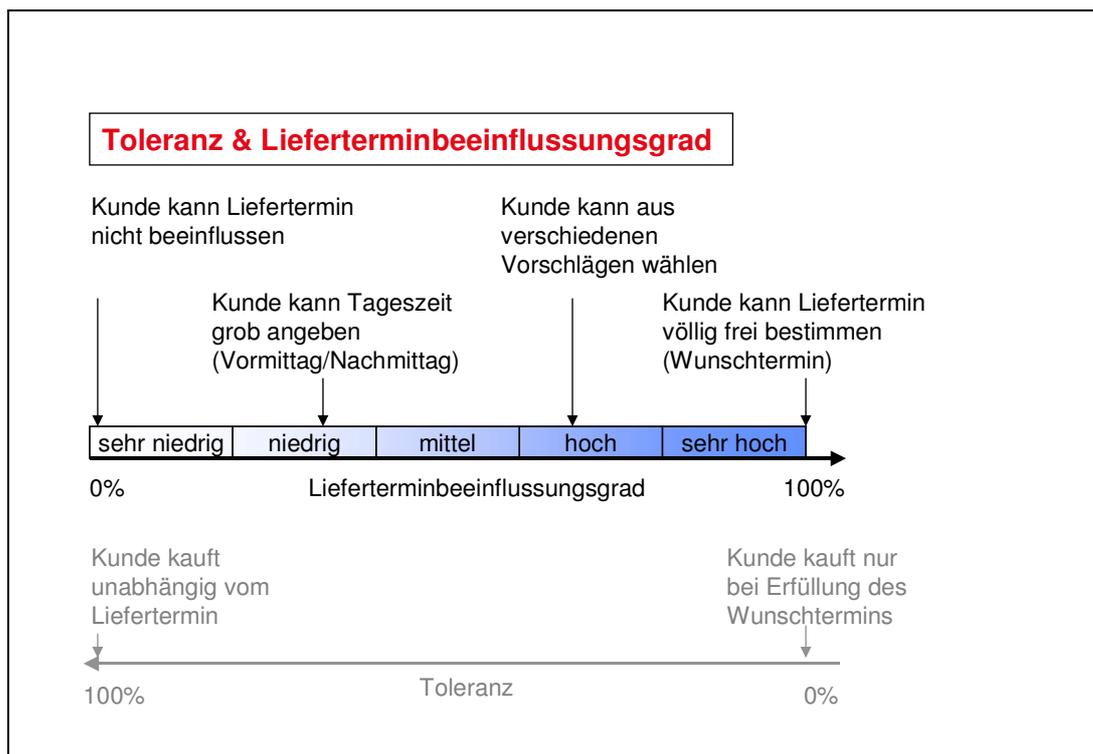


Abbildung 8: Qualitative Skala für Toleranz (TLRZ) und Lieferterminbeeinflussungsgrad (LTBG)

Die Kennwerte Toleranz & Lieferterminbeeinflussungsgrad sind insoweit reziprok zueinander, als für Kunden mit einer geringen Toleranz einen hohen Lieferterminbeeinflussungsgrad erforderlich ist, und Kunden mit einer hohen Toleranz trotz niedrigem Lieferterminbeeinflussungsgrad befriedigt werden können.

3.2.2 Präzision

Im Gegensatz zur Toleranz, die durch Ereignisse vor der Lieferterminvereinbarung determiniert wird, ist die Präzision der Lieferterminzusage ausschlaggebend dafür, wie verlässlich der Kunde seinen weiteren Tagesverlauf planen kann, nachdem die Lieferterminvereinbarung erfolgt ist.

Bei einer unpräzisen Lieferterminvereinbarung „zwischen 8:00 und 13:00 Uhr“ etwa, muss der Kunde sich den gesamten Vormittag freihalten.

Definition „Präzision“:

Präzision (PRÄZ) sei ein Kennwert zur Beschreibung der Genauigkeit der Lieferzeitpunktvorhersage durch den Anbieter.

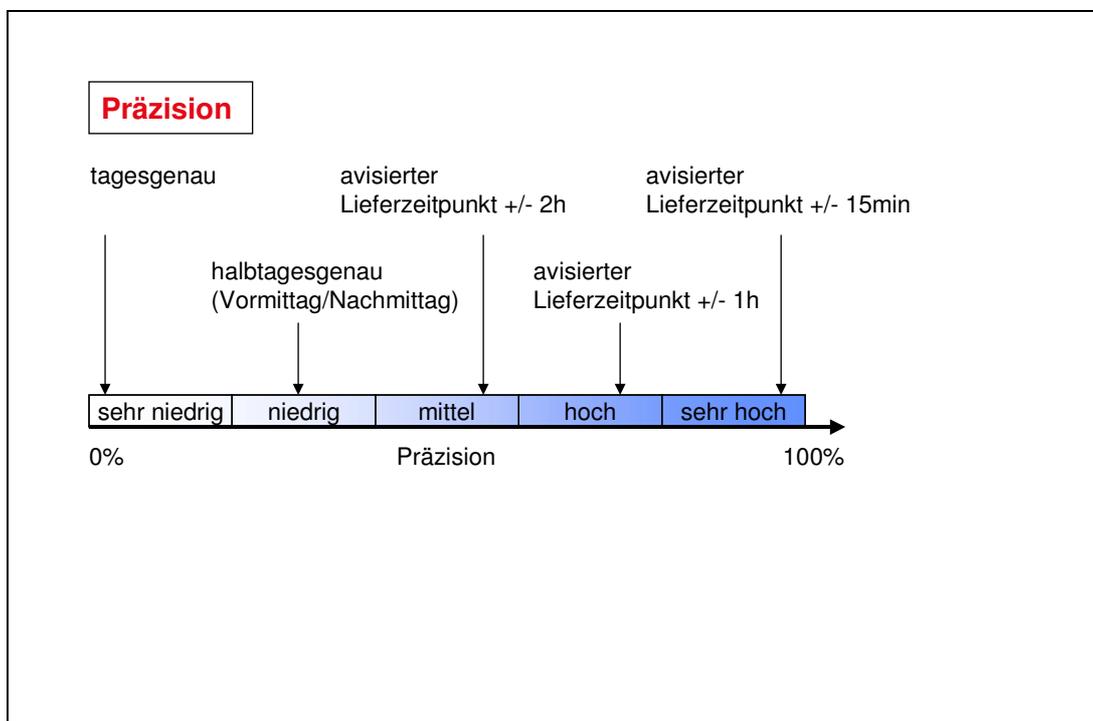


Abbildung 9: Qualitative Skala für Präzision (PRÄZ)

3.2.3 Erfüllungsgrad

Die „Verlässlichkeit“ des Lebensmittelheimzustellservice-Anbieters bei Lieferterminzusagen ist ein wesentlicher Faktor zur langfristigen Kundenbindung. Aus dieser Überlegung resultieren die folgenden zwei Kennwerte.

Definition „Erfüllungsgrad 1“

Der Erfüllungsgrad 1 (EFG1) sei ein Kennwert der beschreibt, ob ein angenommener Auftrag erfüllt wird oder nicht¹⁸.

Definition „Erfüllungsgrad 2“

Der Erfüllungsgrad 2 (EFG2) sei ein zur „Abweichung vom vorhergesagten Lieferzeitpunkt“ reziproker Kennwert.

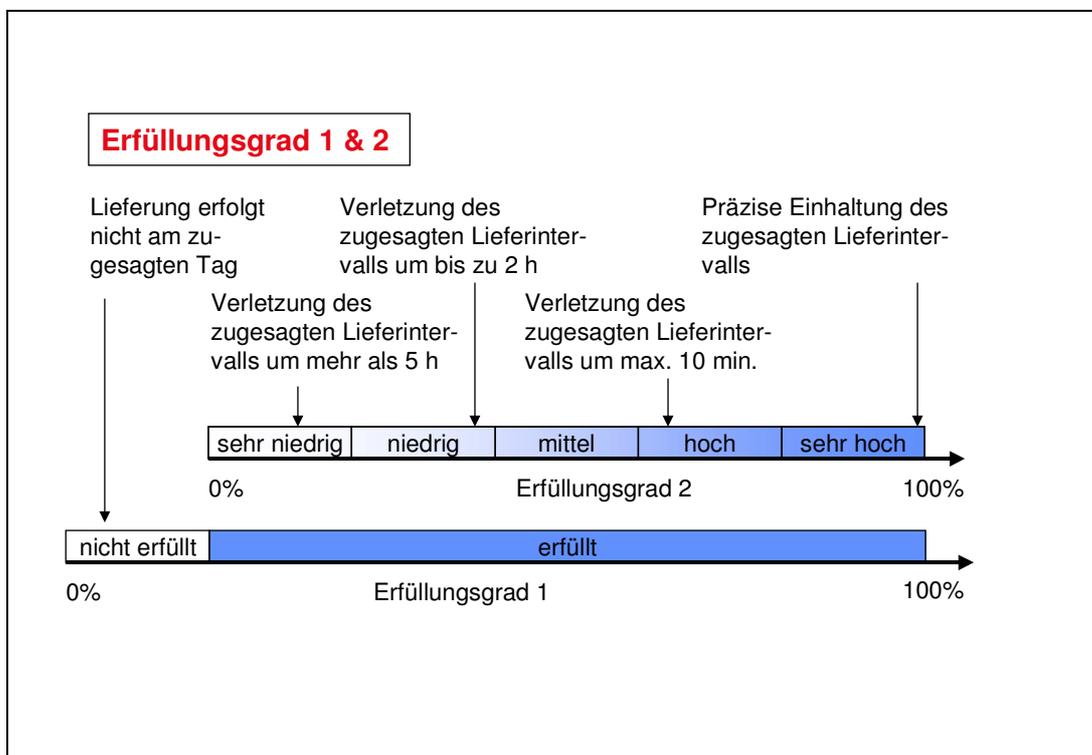


Abbildung 10: Qualitative Skala für Erfüllungsgrad (EFG1 & EFG2)

¹⁸ D.h. der EFG1 ist ein binärer Kennwert. Dabei ist die Nichterfüllung eines Auftrags kaum eine echte Handlungsoption, da der Imageschaden für den Anbieter beträchtlich wäre.

3.2.4 Avisvorlauf

Definition „Avisvorlauf“

Der Avisvorlauf (AVIV) sei ein Kennwert zur Bewertung des Zeitpunkt des Avises in Relation zum Bestell- und Lieferzeitpunkt.

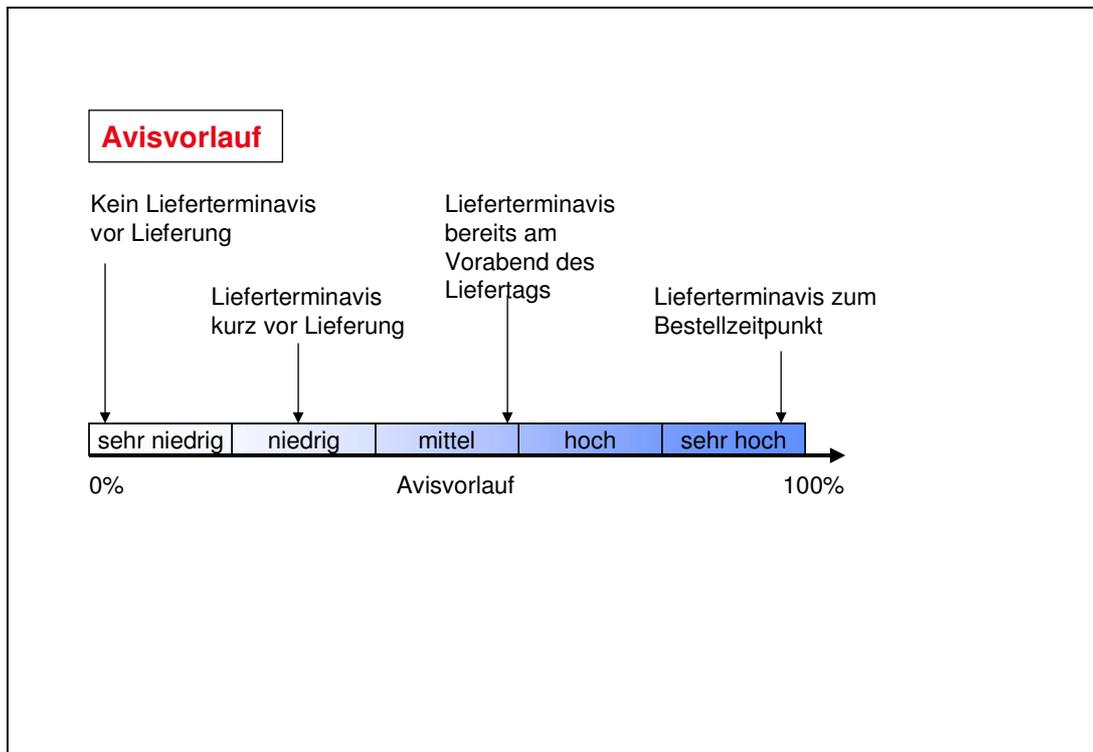


Abbildung 11: Qualitative Skala Avisvorlauf (AVIV)

3.2.5 Tourgebietstreue

*Definition „**Tourgebietstreue**“*

Tourgebietstreue (TGTR) sei ein binärer Kennwert, der aussagt, ob vorgeplante Tourgebiete bei der Tourenplanung eingehalten werden oder nicht.

TGTR = ja := vordefinierte Tourgebiete werden eingehalten

TGTR = nein := vordefinierte Tourgebiete werden nicht eingehalten.

Die TGTR ist in dreierlei Hinsicht wichtig:

Zum ersten ist sie aus **kommissioniertechnischer Sicht** ein wesentliches Kriterium, da bei einer TGTR=ja die Warenbereitstellung tourbezogen erfolgen kann, auch wenn die Tourenplanung noch nicht abgeschlossen ist. Bei einer TGTR=nein ist dies daher nicht möglich, da erst nach Abschluss der Tourenplanung feststeht, welcher Auftrag durch welche Tour bedient wird.

Zum zweiten aus Gründen der **Effizienz** bei der Tourdurchführung. Bei einer TGTR=ja können Stammfahrer auf Tourgebieten eingesetzt werden, die durch präzise Ortskenntnis die Touren in der Regel effektiver durchführen können (durch Ortskenntnis bedingtes Suchen entfällt), bei einer TGTR=nein dagegen nicht.

Zum dritten aus der **qualitativen Service Sicht**, denn bei einer TGTR=ja kann sich über die Zeit eine Fahrer-Kunden-Bindung entwickeln, da in diesem Fall Stammfahrer definierte Gebiete, und somit einen dezidierten Kundenkreis bedienen können.

3.2.6 Peripherieeignung

*Definition „**Peripherieeignung**“*

Peripherieeignung (PERI) sei ein Kennwert, der aussagt, ob ein Planungsverfahren nur für Gebiete mit hoher Kundendichte (Innenstadtgebiete) geeignet ist, oder ob es auch in Peripherie-Gebieten (ländliche Gegenden) mit geringerer Kundendichte brauchbare Ergebnisse erwarten lässt.

PERI = ja := auch für Peripheriegebiete geeignet

PERI = bedingt := nur bedingt für Peripheriegebiete geeignet

PERI = nein := nicht für Peripheriegebiete geeignet.

Bei der Festlegung des Zustellgebiets, die i.d.R. vornehmlich auf Basis des erwarteten Ertragspotenzials für die betrachteten Städte und Gemeinden erfolgt, sollte die Peripherieeignung des gewählten Planungsansatzes unbedingt mit berücksichtigt werden, um die Potenziale dann auch kostengünstig erschließen zu können.

3.2.7 Transportkosteneffizienz

Definition „Transportkosteneffizienz“

Transportkosteneffizienz (TKEF) sei ein Kennwert, der aussagt, wie Kosteneffizient die Last mile Transporte abgewickelt werden. Mögliche Messgrößen für die TKEF sind z.B.

- ▶ Anzahl mit definiertem Fuhrpark abgewickelter Aufträge
- ▶ Durchschnittliche Anzahl Aufträge pro Fahrzeug und Tag
- ▶ Durchschnittliche Anzahl km/ Auftrag
- ▶ Durchschnittliche Bedien-&Transportzeit / Auftrag
- ▶ Durchschnittliche Zustellkosten pro Auftrag

Die (Transportkosteneffizienz) TKEF ist stark abhängig von den oben genannten anderen Kriterien, insbesondere aber von der Präzision (PRÄZ) und dem Lieferterminbeeinflussungsgrad (LTBG).

Der quantitative Fokus dieser Arbeit liegt auf der Untersuchung der TKEF von der PRÄZ bei hohem LTBG (s. Kap. 6).

3.3 Übersicht über bestehende transportlogistische Ansätze

In der transportlogistischen Praxis gibt es verschiedene Ansätze, wie und wann die Planung im Prozesskontext bei Lieferterminavisierung erfolgt. Diese Ansätze werden im Folgenden idealisiert¹⁹ kurz vorgestellt.

Um diese Ansätze qualitativ miteinander zu vergleichen, werden die jeweiligen Eigenschaften der Ansätze mit Hilfe von „ordinalen Wertprofilen“ (s. [Arnold & Furmans, 2005]; S. 267 ff) dargestellt. Ordinale Wertprofile weisen je Kriterium eine Achse auf. Die Ausprägung des Kriteriums ist dann jeweils entlang der Achse angetragen.

Eine Gewichtung der Kriterien untereinander wird dabei nicht getroffen, daher sind die Winkel zwischen den Achsen identisch. Die Skalierung der Erfüllungswertachsen orientiert sich an den oben eingeführten ordinalen Merkmalsausprägungen [1:= nicht erfüllt, ..., 5:= voll erfüllt] bzw. [1:= nein; 3:= bedingt; 5:= ja].

Da die Transportkosteneffizienz (TKEF) in starkem Maße von der genauen Ausprägung des Anwendungsfalls abhängig ist (Geografie, Kundendichte, ...) und

¹⁹ Idealisiert daher, da es durchaus auch Mischformen der Ansätze geben kann.

sehr sensitiv auf kleinere Ausprägungsunterschiede²⁰ der Kriterien PRÄZ und ERG2 reagiert, wird dieser Kennwert an dieser Stelle bewusst noch nicht berücksichtigt.

Der Zusammenhang zwischen Transportkosten (TKEF), Lieferzeitfensterbreite und Planungsverfahren wird jedoch später noch eingehend und quantitativ untersucht.

3.3.1 Der Rahmentour-Ansatz RTA

Der Rahmentour-Ansatz erfordert keine tägliche Tourenplanung. Auf der Basis von Erfahrungswerten werden Gebiete definiert, die von einzelnen Fahrzeugen in Tagestouren bedient werden. Qua Gebietszugehörigkeit ist ein Auftrag der das Gebiet bedienenden Rahmentour zugeordnet.

Beispiel:

Ein Rahmentourplan wird auf den ersten 3 Ziffern der Postleitzahlen definiert:

- 80100-80199: Rahmentour 1

- 80200-80299: Rahmentour 2

- usw.

Soll eine Lieferung zu einer Adresse mit PLZ 80234 erfolgen, so wird der Kunde der Rahmentour 2 zugeordnet. Die Reihenfolge, in der die Sendungen auf einer Tour ausgeliefert werden bestimmt der Fahrer nach Gutdünken.

Dieses Verfahren eignet sich dann, wenn das Lieferaufkommen nur geringen Schwankungen unterworfen ist. Falls das Transportaufkommen einer Rahmentour an einem Tag die Fahrzeugkapazität oder den Arbeitszeitrahmens des Fahrers sprengt, so wird dieses Problem durch operatives ad hoc Umdisponieren behoben („exception management“). Beispielsweise kann der Disponent einen Auftrag aus der „überlaufenden“ Tour einer Nachbartour zuordnen oder diesen Auftrag bis zum Folgetag „liegen lassen“.

Der RTA findet in der Praxis beispielsweise Anwendung in der Zustellung von Weiß- und Braunware und Versandhandelsartikeln. Ein Vorteil dieses Ansatzes ist, dass auf ein Tourenplanungssystem verzichtet werden kann. Dieser Vorteil ist nicht zu unterschätzen, da Sendungen oft ohne vorausseilende²¹ und ohne vollständige²² Sendungsdaten abzuwickeln sind. Ein weiterer Vorteil ist die einfache Warenbereitstellung auf Basis der fixen Rahmentouren (hohe TGTR), die somit bereits vor Vorliegen aller Transportaufträge begonnen werden kann. Der RTA ist aus Sicht der effizienten Tourabwicklung auch für die Bedienung peripherer ländlicher Gebiete geeignet, solange keine exakten Wunschliefetermine

²⁰ kleinere Ausprägungsunterschiede = kaum erkennbar im ordinalen Wertprofil

²¹ d.h. die Daten sind für Tourenplanungszwecke verfügbar, bevor die physische Sendung eintrifft.

²² d.h. alle Sendungen sind durch vollständige Sendungsdatensätze so beschrieben, dass eine Tourenplanung möglich wird. Die wesentlichen Sendungsdatensätze sind: präzise Lieferanschrift, Auftragsgröße (Gewicht, Volumen, Anzahl Packstücke etc.) und Zustellzeitfenster

(Kombination von hohem LTBG und hoher PRÄZ) zugesagt werden müssen.

Die wesentlichen Nachteile dieses Ansatzes liegen zum einen in der mangelnden Flexibilität bei schwankendem Sendungsaufkommen. Ist das Sendungsaufkommen geringer als gewöhnlich, so sind die Rahmentouren nicht ausgelastet. Ist das Sendungsaufkommen höher als gewöhnlich, so laufen viele Rahmentouren über. Im günstigsten Fall kann dies durch geschickte Umdisposition bei deutlich steigendem Dispositionsaufwand und sinkender Transportkosteneffizienz (TKEF) abgefangen werden, im ungünstigsten Fall sinkt außerdem die Liefertreue (ERG1).

Zum andern ist es kaum möglich, den Kunden zum Bestellzeitpunkt genau über den zu erwartenden Liefertermin zu informieren (geringe PRÄZ), da die Auslieferreihenfolge erst nach Vorliegen aller Aufträge vom Disponenten oder vom Fahrer auf Basis seiner Gebietskenntnis festgelegt wird. Die bestmögliche Genauigkeit einer Lieferzeitpunktsvorhersage ist daher in der Regel auf grobe Aussagen wie „vormittags“ oder „nachmittags“ begrenzt.

Aus Kundensicht ist die systemimmanent fehlende Möglichkeit zur Lieferterminbeeinflussung durch den Kunden (sehr geringe LTBG) der wesentliche Nachteil dieses Ansatzes.

In praxi findet sich dieser Ansatz mit abnehmender Tendenz oft noch in der Möbelheimzustellung.

In einem ordinalen Wertprofil lässt sich der RTA folgendermaßen darstellen.

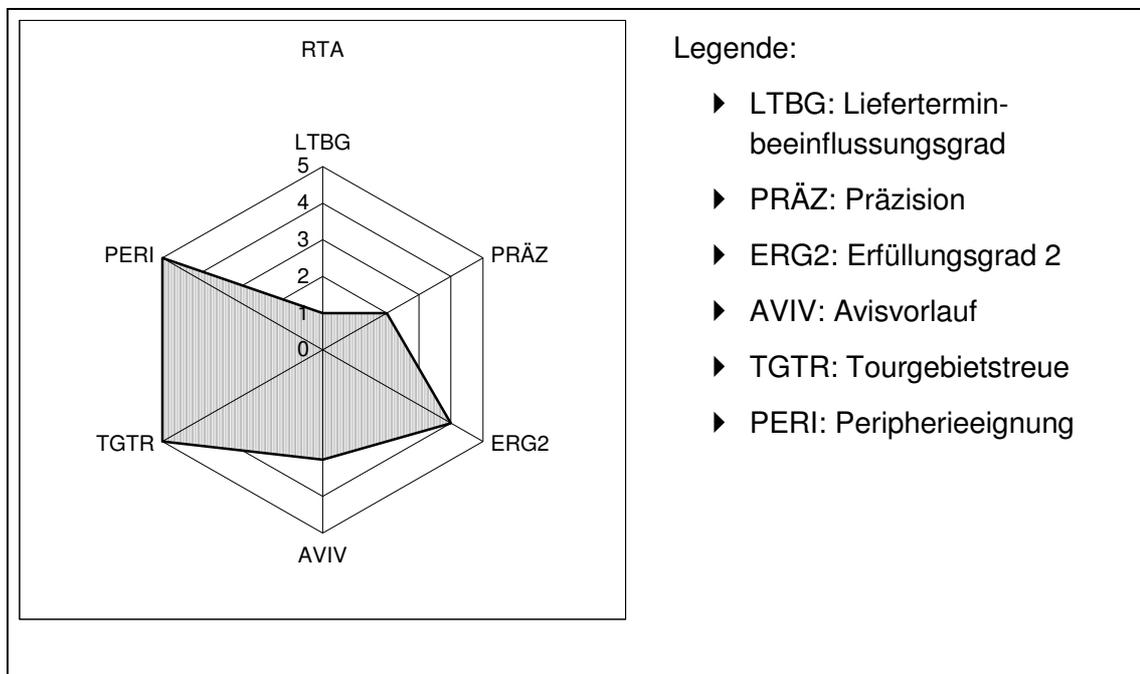


Abbildung 12: Ordinales Wertprofil RTA

3.3.2 Der „Planung ante Avis“ Ansatz PAA

Eine notwendige Bedingung für den PAA ist das Vorliegen vorausseilender und vollständiger Sendungsdaten.

Für einen definierten Liefertag werden im Vorfeld die Sendungsdaten gesammelt, und i.d.R. mit Hilfe eines Tourenplanungssystems optimiert zu Touren verplant. Vor der Tourenplanung sind die Sendungsaufträge i.d.R. nur auf einen Liefertag terminiert, jedoch nicht auf ein Lieferzeitfenster. Nachdem der Tourenplan vorliegt, lässt sich aus diesem Tourenplan für jede Sendung ein relativ präziser Lieferzeitpunkt ablesen²³.

Sobald die Tourenplanung abgeschlossen ist (geringer LTBG), können die Kunden jedoch präzise über den jeweils geplanten Liefertermin informiert werden (hohe PRÄZ). Aus Sicht des Kunden ist dieses Avis deshalb vorteilhaft, weil er sich so auf den Zeitpunkt der Lieferung einstellen kann.

Wegen der geringen Tourgebietstreue dieses Ansatzes (TGTR) muss es aus der Lager-Prozess-Sicht zeitlich²⁴ und räumlich²⁵ möglich sein, auf Basis von täglich anders laufenden Touren die Warenbereitstellung zu organisieren. Falls entsprechende Lagerprozesse den PAA ermöglichen, ist der PAA auch für ländliche Gebiete geeignet (hohe PERI).

Da die ex post avisierten Lieferzeitpunkte keine Planungsbeschränkung darstellen, sondern das Ergebnis der Planung sind, können die Touren in Hinblick auf die Ressourcennutzung optimiert und somit relativ kostengünstig abgewickelt werden. Der wesentliche Vorteil dieses Verfahrens ist daher die hohe Transportkosteneffizienz (TKEF) bei guter Präzision (PRÄZ) des Lieferavises.

Aus Kundensicht stellen jedoch sowohl der relativ späte Zeitpunkt, zu dem dieses Avis gemacht werden kann (geringer AVIV), als auch der geringe LTBG Nachteile dar, die insbesondere in der Lebensmittelheimlieferung die Anwendung dieses Ansatzes verhindern.

Der PAA findet in praxi u.a. in der Küchenmöbel-Heimzustellung Verwendung.

In einem ordinalen Wertprofil lässt sich der PAA folgendermaßen darstellen.

²³ der Tourenplan weist für jede Sendung eine vorausberechnete Ankunftszeit aus

²⁴ physische Dauer der Warenbereitstellung

²⁵ z.B. Platz für Umlagern zwischen Kommissionierbereich und „Warenbereitstellungszone für den Tourabgang“, wenn Kommissionierung vor Abschluss der Tourenplanung erfolgen muss, und daher nicht tourbezogen kommissioniert werden kann.

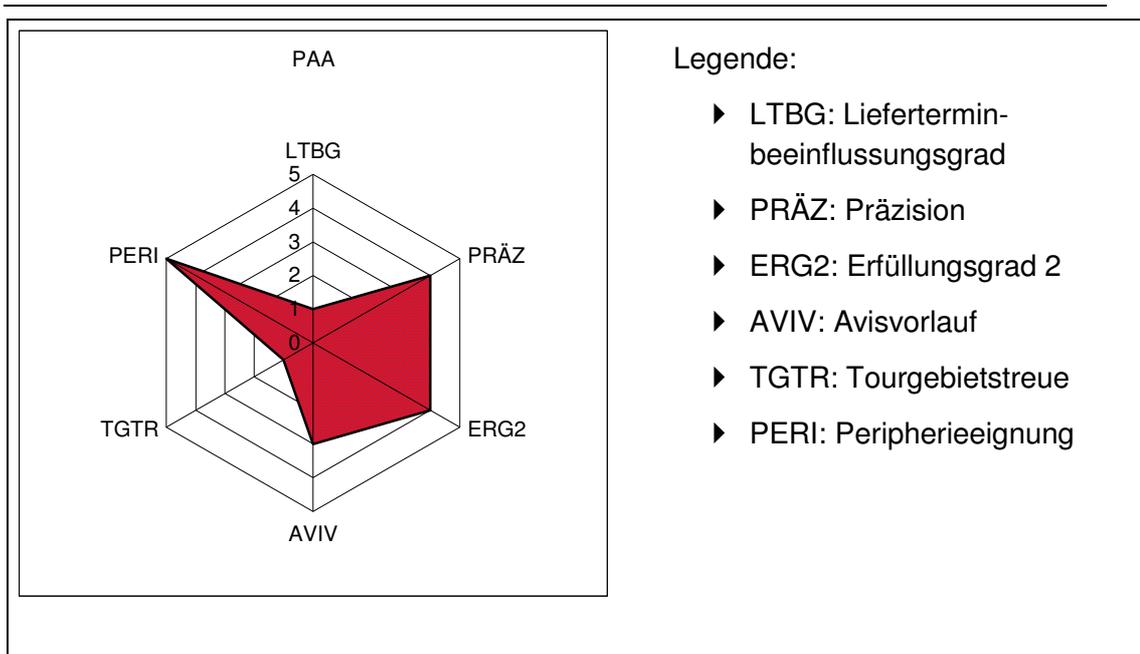


Abbildung 13: Ordinales Wertprofil PAA

3.3.3 Der „Planung post Avis“ Ansatz PPA

Im „Planung post Avis“ Ansatz werden zeitliche Avise extern generiert, d.h. sie sind für die transportlogistische Abwicklung unbeeinflussbare²⁶ Restriktionen. Dieser Ansatz entspricht aus OR-Sicht dem klassischen deterministischen VRSP-TW²⁷, da eine definierte Anzahl von Aufträgen mit definierten Zeitfenstern zu verplanen ist.

Der PPA findet sich in praxi z.B. dann, wenn der Versender die Lieferzeitfenster mit dem Kunden bilateral festgelegt hat, den Transport an einen Transportdienstleister ausgelagert hat („outsourcing“), und für den Transportdienstleister die Lieferzeitfenster somit extern vorgegeben sind.

Bei Konflikten mit zugesagten Zeitfenstern müssen ggf. zusätzliche Fahrzeugressourcen bereit gestellt werden.

In einem ordinalen Wertprofil lässt sich der PPA folgendermaßen darstellen.

²⁶ „unbeeinflussbar“ im Sinne von nicht operativ beeinflussbar.

²⁷ VRSP-TW: vehicle routing and scheduling problem with time windows

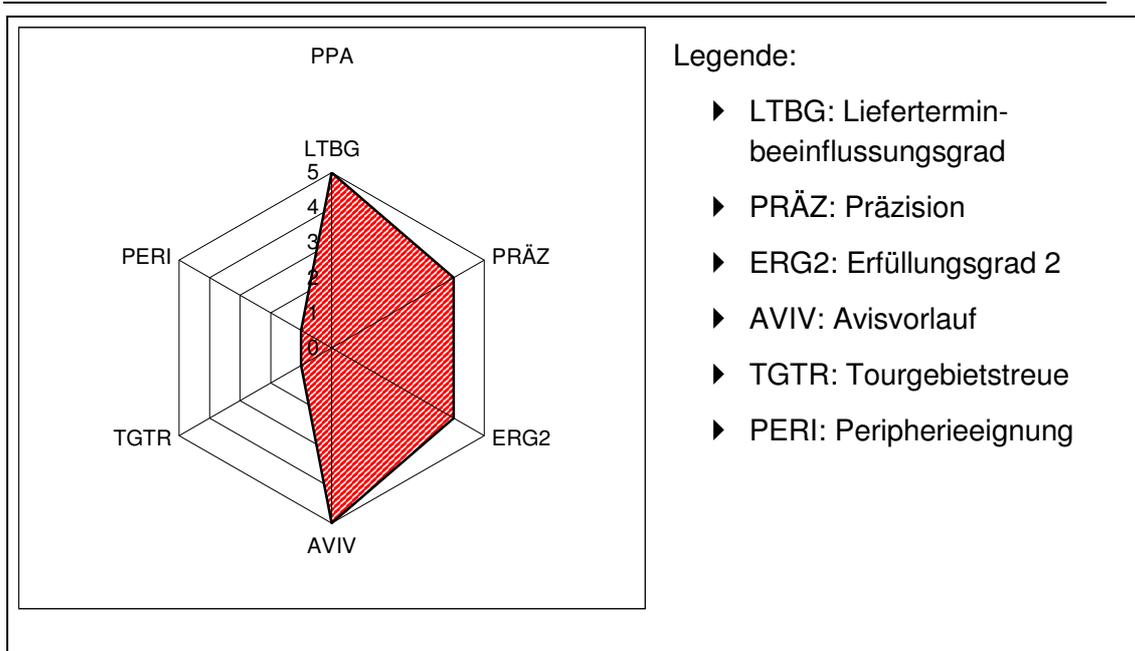


Abbildung 14: Ordinales Wertprofil PPA

3.3.4 Der „Reduzierter Erfüllungsgrad“ Ansatz REA

Der „Reduzierter Erfüllungsgrad“ Ansatz REA ist ein Ansatz, der definitionsgemäß akzeptiert, dass für den Fall dass mehr Bestellungen einlaufen als unter Einhaltung des ERG2 abgewickelt werden, der zugesagte ERG2 einfach verletzt werden kann.

Dieser Ansatz lässt sich marketingtechnisch schwer verkaufen, da der angebotene Service pointiert formuliert etwa folgender wäre: „Wir garantieren Ihnen ein definiertes Lieferzeitfenster, aber wenn's eng wird, kümmert uns die Garantie wenig“. Trotzdem ist dieser Ansatz in der Praxis anzutreffen – evtl. nicht als vom Marketing propagierte Methode, dafür aber als pragmatische Problemlösung in der logistischen Umsetzung.

Der Kunde wird diesen Ansatz akzeptieren, wenn er nur in Ausnahmefällen zur Anwendung kommt (abgedeckt z.B. durch Ausgabe von Verspätungsgutscheinen), oder wenn die Abweichungen „unerheblich“ sind (z.B. 5 min. zu spät bei 4h Lieferzeitfenster).

In praxi findet sich dieser Ansatz als „Notfallkonzept“. Aus Marketinggründen wird dieser aber sicher nicht propagiert.

In einem ordinalen Wertprofil lässt sich der REA deshalb nur schwer darstellen, da er irgendeinen anderen Ansatz als Basisansatz benötigt, von dem er sich dann nur im ERG2 signifikant unterscheidet. Daher wird hier der REA als Variante des PPA-Ansatzes beispielhaft dargestellt.

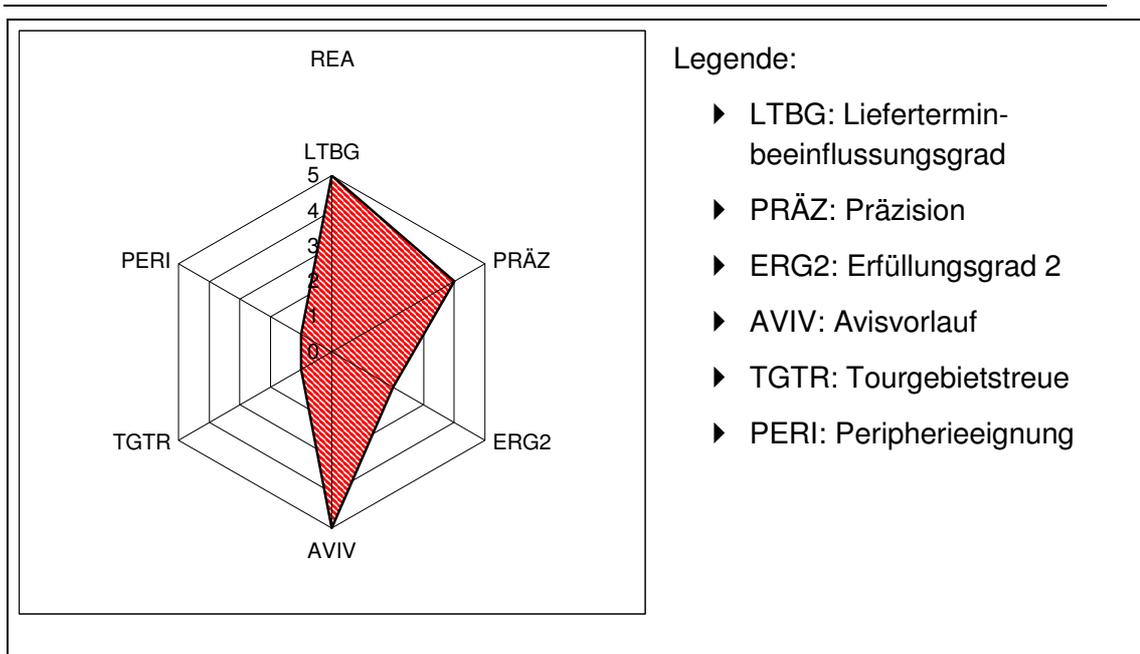


Abbildung 15: Ordinales Wertprofil REA (Beispiel)

3.3.5 Der „Filter vor Planung“ Ansatz FPA

Soll der Lieferzeitpunkt dem Kunden bereits bei der Bestellung möglichst präzise bekannt gemacht werden, so muss definiert werden, nach welcher Methode dieser Lieferzeitpunkt bestimmt werden soll. Dabei ist es wichtig, nicht zu vielen Kunden zur „gleichen Zeit“ eine Lieferung zuzusagen²⁸.

Ein möglicher Lösungsansatz für dieses Problem ist die einfache Filter-Regel: „Sage nicht mehr als k Kunden die selbe Tageszeit-Stunde²⁹ als Liefertermin zu“.

D.h. dieser Ansatz läuft zweistufig ab:

1. Zusage der Lieferzeiten in Übereinstimmung mit dem Filter:
Anzahl Lieferungen < k je Tageszeit-Stunde
2. Tourenplanung unter Berücksichtigung der zugesagten Lieferzeitfenster.

Dieser sehr einfache Ansatz kann dann funktionieren, wenn das Liefergebiet sehr kompakt ist, und letztendlich jedes Fahrzeug jeden Kunden bedienen kann. In diesem Fall können die Lieferzeiten „blind“ zugesagt werden, und die Wahrscheinlichkeit ist hoch, dass die zugesagten Liefertermine später auch eingehalten werden können, d.h. dass ein zulässiger Tourenplan gefunden wird, der diese engen zugesagten Lieferzeitfenster auch alle berücksichtigt.

Die wesentlichen Vorteile dieses Ansatzes sind der hohe Liefertermin-

²⁸ Würde 100 Kunden jeweils präzise 18:00 Uhr als Lieferzeitpunkt zugesagt, so wären 100 Fahrzeuge nötig, um diese Zusage genau einzuhalten.

²⁹ z.B. nicht mehr als 10 Kunden gleichzeitig die Lieferung zwischen 14:00 und 15:00 zusagen.

beeinflussungsgrad (LTBG) und die hohe Präzision (PRÄZ) gepaart mit kundenfreundlichem Avis bei Bestellung (hoher AVIV).

Die Nachteile dieses Ansatzes sind die geringe Eignung für nichturbane Gebiete (geringe PERI) und die geringe Tourgebietstreue (TGTR) mit den oben beschriebenen Implikationen für den Lagerbetrieb.

Der Erfüllungsgrad (ERG2) hängt dabei stark davon ab, wie „eng“ die zugesagten Zeitfenster sind, und wie viele Termine pro Stunde im Verhältnis zur Anzahl der verfügbaren Fahrzeuge zugesagt wurden³⁰.

In praxi findet sich dieser Ansatz auch im Lebensmittelheimlieferservice. In einem ordinalen Wertprofil lässt sich der FPA folgendermaßen darstellen.

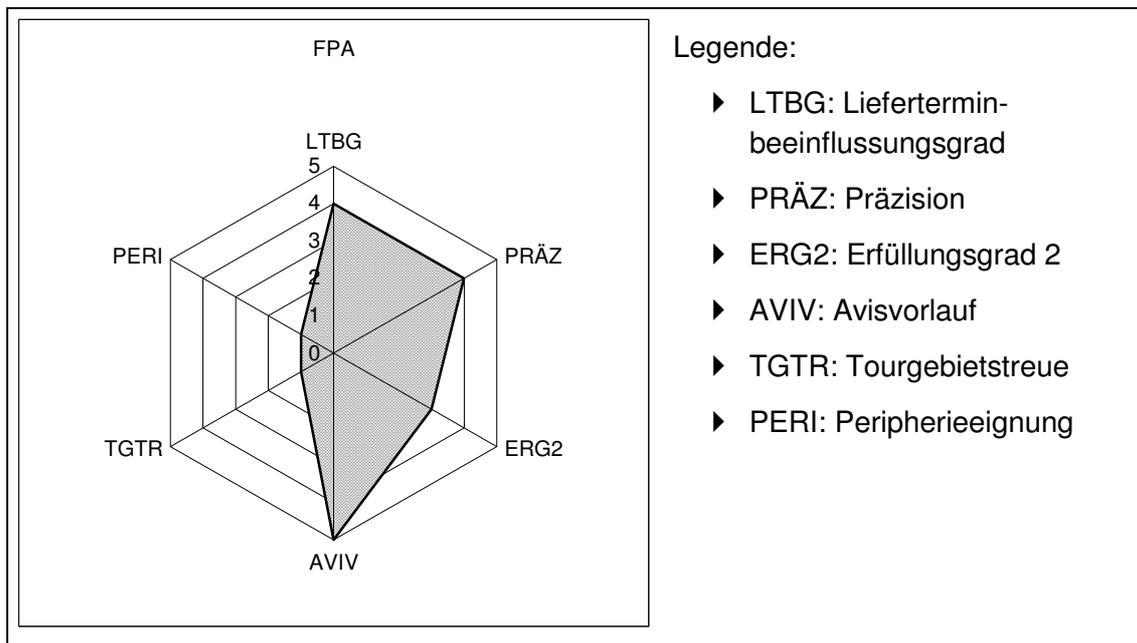


Abbildung 16: Ordinales Wertprofil FPA (Beispiel)

Das Kernproblem des FPA ist, den Filter gut zu kalibrieren. Wird der Filter zu restriktiv eingestellt (kleines k), so bleiben unnötig Lieferwünsche der Kunden unberücksichtigt, und Transportkapazitäten liegen brach. Wird der Filter zu großzügig eingestellt (großes k), so lassen sich ex post die zugesagten Zeitfenster nicht mehr oder nur durch Einsatz zusätzlicher Fahrzeuge einhalten. D.h. der FPA-Ansatz lässt sich „nach der Filterfunktion“ auf den PPA oder den REA zurückführen. Im Folgenden wird daher bei Bedarf nach FPA/PPA und FPA/REA unterschieden.

3.3.6 Der „Online Tourenplanung“ Ansatz ORA

Die wesentliche Idee dieses Ansatzes (Details s. Kap. 5.4 dieser Arbeit und [Kunze,

³⁰ z.B. um regelmäßig eine hohe Transportkosteneffizienz TKEF über einen hohen Fahrzeugauslastungsgrad zu erzielen

2003] sowie [Kunze, 2005]) ist ein vorab definiertes Touren-Template³¹. Auf Basis dieses Templates ist es möglich, den Kunden jeweils bei Bestellung eine Lieferzeitusage zu machen, auch wenn die eigentliche Tour noch gar nicht vollständig gebildet ist. Die Touren „wachsen“ sozusagen „entlang des Templates durch Aufnehmen von Aufträgen, die in der Nähe liegen“.

Damit der Kunde den Liefertermin nicht nur präzise zugesagt bekommt, sondern diesen auch selbst beeinflussen kann, sind die Templates so angelegt, dass Lieferregionen³² (z.B. Stadtteile) mehrfach bedient werden, und der Kunde daher mehrere Lieferzeitpunkte vorgeschlagen bekommt.

Dieser Prozess lässt sich automatisieren, und ist daher besonders für den Einsatz im Lebensmittelheimlieferservice (mit Bestellung im Internet oder über ein „call-center“) geeignet.

Der wesentliche Vorteil dieses Verfahrens ist, dass wie beim FPA bereits zum Bestellzeitpunkt (hohe AVIV) eine präzise Lieferzeitusage (hohe PRÄZ) gemacht werden kann, im Gegensatz zum FPA aber außerdem garantiert ist, dass diese Zusagen später auch eingehalten werden können (hoher ERG2), wenn nicht unvorhergesehene Verkehrsumstände dies verhindern.

Der Nachteil dieses Verfahrens ist, dass es nur bedingt für den Einsatz in nichturbanen Gebieten geeignet ist (mittlere PERI). Der Grund hierfür liegt darin, dass die Transportkosteneffizienz (TKEF) drastisch sinkt, wenn mit diesem Ansatz auch ländliche Gebiete bedient werden sollen³³.

Dieser Ansatz findet sich bereits in praxi im Lebensmittelheimlieferservice. In einem ordinalen Wertprofil lässt sich der ORA folgendermaßen darstellen.

³¹ template [englisch] = Muster, Vorlage; ein Touren-Template ist ein vom Planer auf Basis der erwarteten Auftragseingänge vorab definierter grober raum-zeitlicher Tourverlauf aller Touren im Planungsgebiet.

³² im folgenden auch Aggregate genannt (s.u.)

³³ Diese Aussage wird später noch transparenter, wenn das Verfahren detaillierter beschrieben wird. Im Großen und Ganzen kann man aber sagen, dass zeitlich eng begrenzte Lieferzusagen in ländlichen Gebieten wegen der fehlenden geografischen Synergien sehr teuer sind.

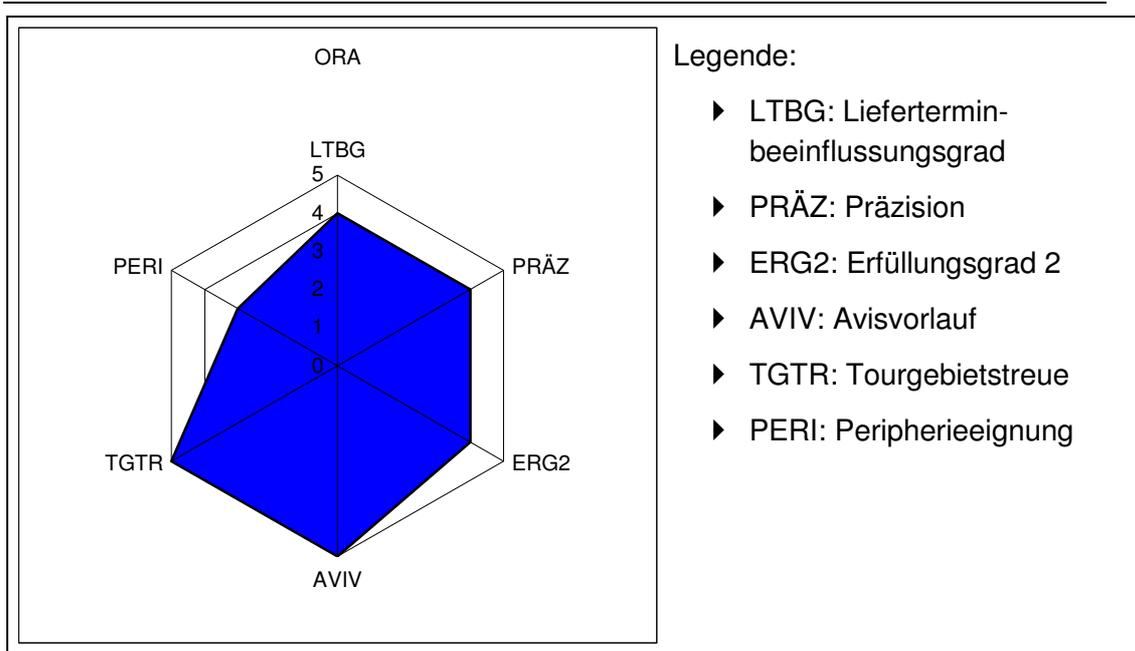


Abbildung 17: Ordinales Wertprofil ORA

3.3.7 Weitere denkbare Ansätze

Alle bisher hier vorgestellten Ansätze sind in der logistischen **Praxis** zu finden. Ein sehr interessanter theoretischer Ansatz, der prozesstechnisch als alternativer Ansatz zu ORA in Betracht kommen könnte, sich aber wesentlich von ORA unterscheidet ist der „multi scenario approach“ (MSA) von [Bent & van Hentenryck 2003], auf den später noch näher eingegangen wird. Aus der Literatur ist weder ersichtlich, ob MSA bereits in praxi eingesetzt wird oder nicht, noch ob MSA sich hinsichtlich der Antwortzeiten überhaupt für den Einsatz im Lebensmittelheimlieferservice mit online-Lieferzeit-Zusage zum Bestellzeitpunkt eignet.

3.3.8 Gegenüberstellung der logistischen Ansätze

Wie bereits geschildert, gibt es verschiedene Möglichkeiten auf den Zielkonflikt zwischen:

- ▶ Präzision der Vorhersage (PRÄZ),
- ▶ Ungewissheit der Nachfrage³⁴,
- ▶ Irreversibilität von einmal zugesagten Zeitfenstern,
- ▶ Einhalten der zugesagten Zeitfenster (ERG2),
- ▶ Ressourcenbegrenzung (Anzahl verfügbarer Fahrzeuge) und

³⁴ Ungewissheit in bezug auf Anzahl, geografische und zeitlicher Lage der Lieferwünsche

- ▶ angestrebter Transportkosteneffizienz (TKEF)

zu reagieren. Im Wesentlichen liegen die Reaktionsmöglichkeiten in der Steuerung des Nachfragevolumens³⁵ (ORA) und in der Anpassung des Transportleistungs-Angebots entweder durch ad hoc Erweiterung der Flotte (FPA-PPA) oder durch Reduzierung des Service³⁶ (FPA-REA - Verletzung von zugesagten Lieferzeitfenstern = Relaxation der zeitlichen Restriktionen der Nachfrage).

Die folgende Grafik illustriert diese drei Steuerungsprinzipien:

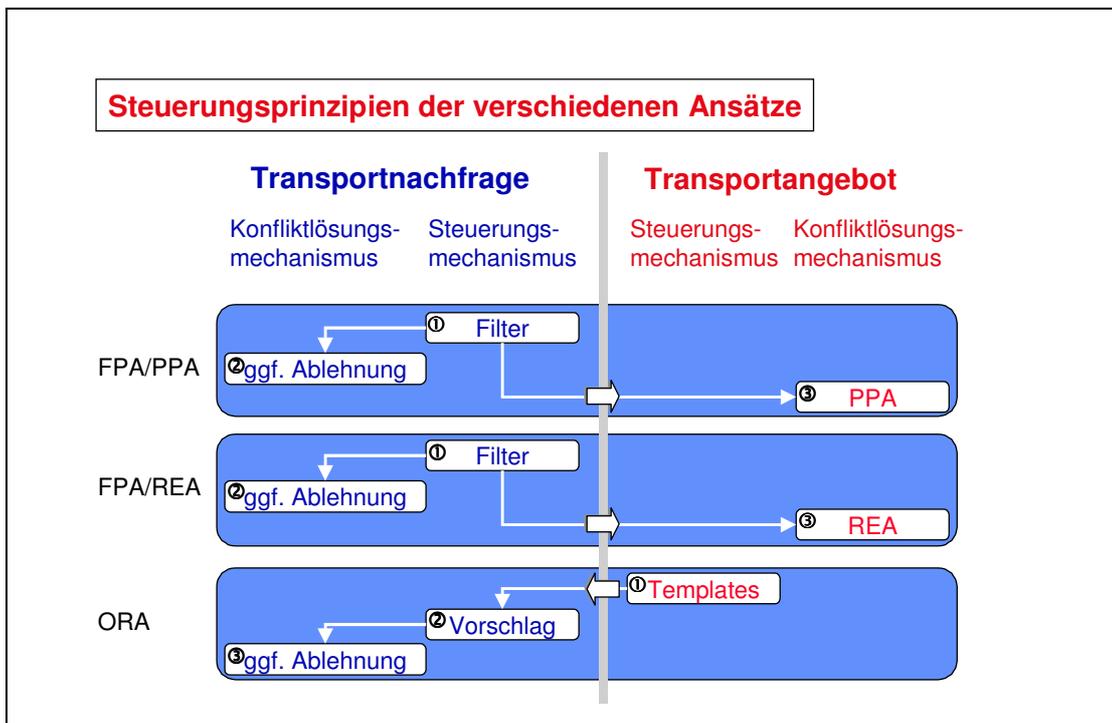


Abbildung 18: Steuerungsprinzipien

Für den FPA steuert der „Filter“ wie viele und welche Aufträge angenommen werden. Übersteigt die Transportnachfrage das Angebot trotz Filter, so kann entweder durch erhöhten Ressourceneinsatz (PPA) oder durch Zeitfensterverletzung (REA) reagiert werden.

Für den ORA steuert die Vorschlagsfunktion das Bestellverhalten der Kunden. Nur wenn kein für den Kunden akzeptabler Lieferterminvorschlag generiert wird, wird der Kunden eine Lieferung ablehnen (d.h. er wird nicht bestellen).

³⁵ Vgl. [Grupp 87; S30ff]: „A-Szenario“ = Ablehnung von Aufträgen

³⁶ Vgl. [Grupp 87; S33ff]: „V-Szenario“ = Verletzung von Zeitrestriktionen

3.4 Fuhrparkausstattung

Im Lebensmittel-Heimlieferservice werden, sofern die Sendungen nicht temperaturisoliert verpackt sind i.d.R. spezielle kleine Sprinter-Fahrzeuge mit drei Temperaturzonen eingesetzt.

Die drei Temperaturzonen sind notwendig, damit sowohl Trockensortimentsartikel, als auch frische Ware (Obst, Gemüse, Fleisch, Molkereiprodukte etc.) und Tiefkühlware ohne Unterbrechung der gesetzlich vorgeschriebenen Kühlkette transportiert werden können. Diese „Beschränkung des verfügbaren Fuhrparks“ ist für die Formulierung des Optimierungsproblems eine wichtige Nebenbedingung, auf die später noch näher eingegangen wird.

Sind die Sendungen hingegen temperaturisoliert verpackt (z.B. in Styropor und trockeneisgekühlt), so können auch Kleintransporter ohne Kühlung eingesetzt werden.

Kleine Fahrzeuge sind deshalb sinnvoll, da so zum einen die Fahrer bei der Zustellung einfacher parken können, und da zum anderen meist nicht das Fahrzeugvolumen, sondern der Zeitbedarf zur Abwicklung der Zustellungen der wesentliche tourbegrenzende Faktor ist. Die einzige Ausnahme bilden Getränkekastenlieferungen – daher ist bei vielen Heimlieferdiensten das Bestellvolumen an Getränkekästen beschränkt.

Da die Lieferadressen täglich wechseln, ist es außerdem sinnvoll, die Fahrzeuge mit Navigationssystemen auszustatten, damit zum einen neue Fahrer schnell produktiv werden können, es zu keinen signifikanten Verlusten bei Urlaubsvertretungen kommt und Staus gegebenenfalls umfahren werden können.

3.5 Bereitstellungsstrategien

Im Zusammenhang mit der TGTR wurde bereits kurz auf die Abhängigkeit zwischen Planungsverfahren und den lagerprozessbedingten Warenbereitstellungsmöglichkeiten eingegangen. Das Kernproblem hierbei ist die zeitliche Abhängigkeit der Prozesse Auftragseingang, Tourenplanung, Kommissionierung, Warenbereitstellung und Tourdurchführung.

Da dieser praxisrelevante Aspekt in der OR-Literatur oft unberücksichtigt bleibt, wird er hier explizit beschrieben.

Wesentlich für die Warenbereitstellung sind die zeitliche Dauer der Prozesse Auftragseingang, Kommissionierung und Tourenplanung, die Abhängigkeit der Tourenplanung vom Auftragseingang, und die Abhängigkeit der Kommissionierung

und Warenbereitstellung vom Ergebnis der Tourenplanung.

Es seien im Folgenden drei prinzipielle Fälle der Warenbereitstellung unterschieden:

- ▶ Fall 1: Warenbereitstellung auf Rahmentourbasis (hohe TGTR)
- ▶ Fall 2: Warenbereitstellung auf Tourbasis (geringe TGTR)
 - Fall 2a: Kommissionierung nach Tourenplanung und
 - Fall 2b: Kommissionierung parallel zur Tourenplanung
- ▶ Fall 3: Warenbereitstellung auf Gebietsstellflächen (unerhebliche TGTR)

3.5.1 Warenbereitstellung auf Rahmentourbasis

Die Abhängigkeiten zwischen den Prozessen Auftragseingang, Kommissionierung und Warenbereitstellung sind bei der Warenbereitstellung auf Rahmentourbasis sehr einfach.

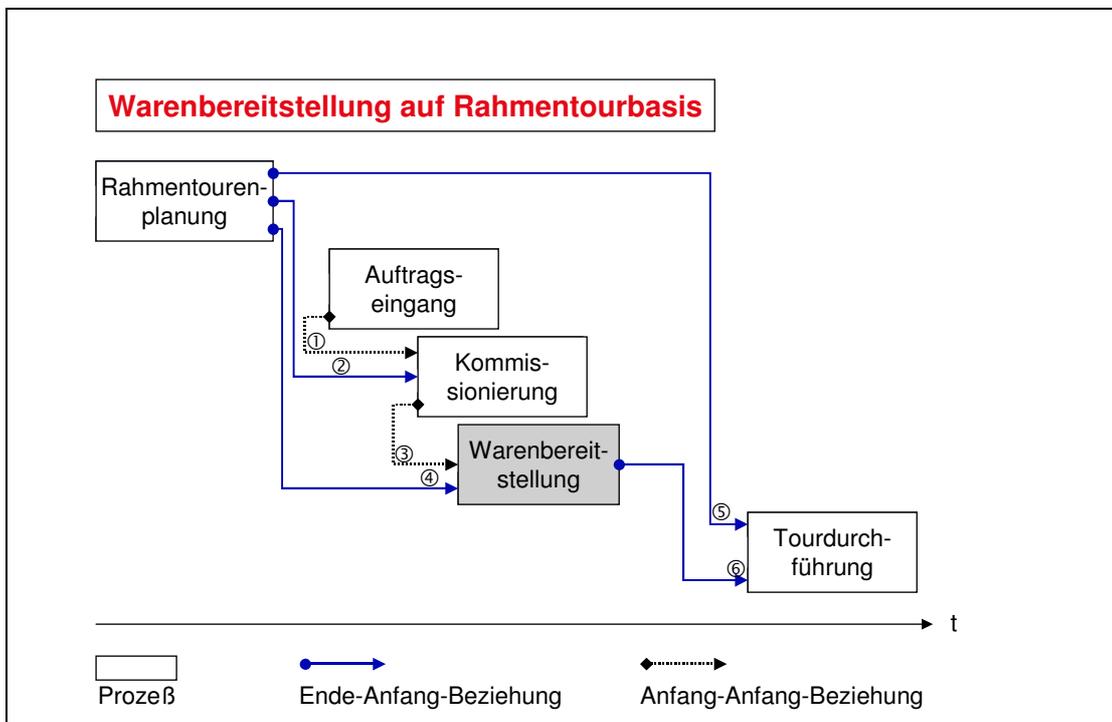


Abbildung 19: Prozessabhängigkeiten bei Rahmentourkonzept (z.B. RTA)

Da bei der Kommissionierung eines Auftrags die Rahmentourzuordnung bereits ex ante bekannt und unveränderlich ist ②, können beim Kommissionieren auf den Sendungen Labels aufgebracht werden, die die Rahmentournummer als Bereitstellungsflächenkennung ausweisen ③.

3.5.2 Warenbereitstellung auf Tourbasis

Soll die Warenbereitstellung tourbezogen erfolgen, so muss vor der Warenbereitstellung die Auftrags-Tour-Zuordnung bekannt sein.

Dies führt bei einem kontinuierlichen Auftragseingang dann zu einem Problem, wenn für eine gute Planung alle Aufträge bekannt sein müssten. Ein beispielsweise auf Basis von nur 60% der auszuführenden Aufträge optimierter Tourenplan wird durch die notwendige ex post Einbeziehung der restlichen Aufträge beliebig schlecht. Wird hingegen die Planung bei Vorliegen von 100% der auszuführenden Aufträge komplett neu erstellt, so wird sich die ursprüngliche **Auftrags-Tour-Zuordnung** der „60%-Planung“ i.d.R. signifikant **verändern**.

Aus der betrieblichen Ablaufsicht folgt aus einer veränderten Auftrags-Tour-Zuordnung ein Warenbereitstellungs-Problem, da Änderungen von Auftrags-Tour-Zuordnungen nach der Kommissionierung & Warenbereitstellung ein „Suchen und Umräumen“ im Warenausgang nach sich ziehen würde.

Soll für die Last-mile-Tourenplanung die „**Kommissionierung nach Tourenplanung**“ (z.B PAA oder FPA) zum Einsatz kommen, so ist wegen dem Bereitstellungsproblem sicher zu stellen, dass die Auftragsbasis für die Tourenoptimierung bereits möglichst alle auszuliefernden Aufträge umfasst.

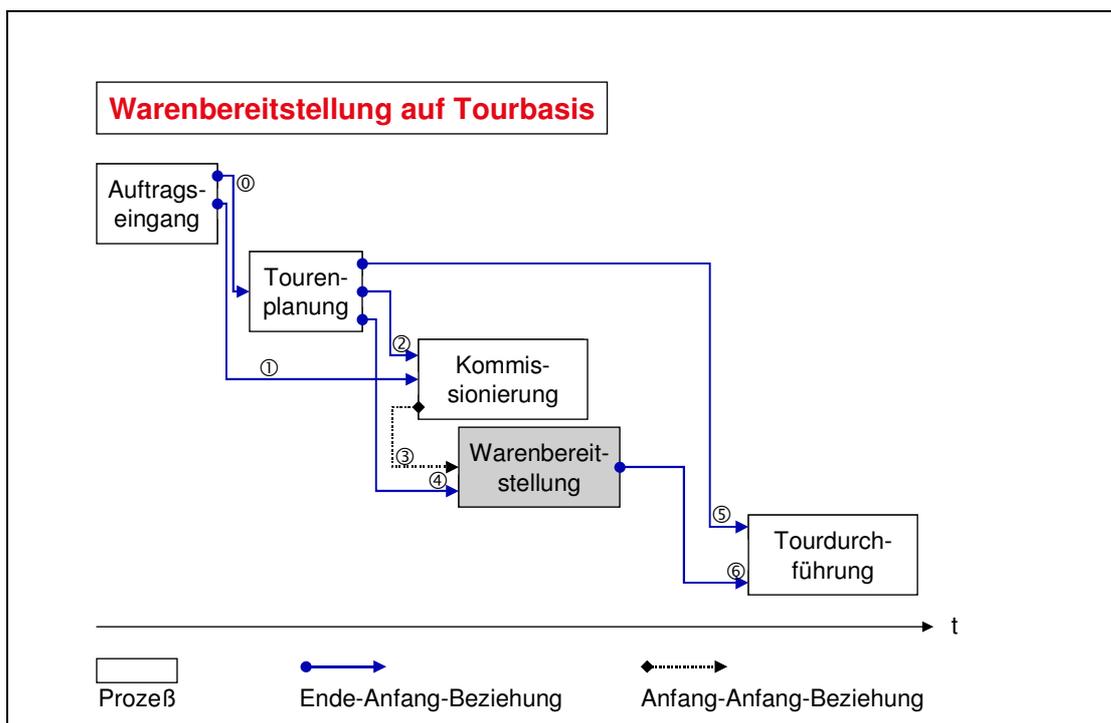


Abbildung 20: Prozessabhängigkeiten bei Kommissionierung nach Tourenplanung

Im Prozessbild wird deutlich, dass die Tourenplanung vor dem Kommissionierbeginn abgeschlossen sein muss ②, damit auf den Sendungslabells die richtige tourbezogene Warenbereitstellungszone als Eingangsgröße für den

Prozess der Warenbereitstellung aufgedruckt werden kann ③.

Wenn die Tourenplanung sukzessive (d.h. parallel zum Auftragseingang) erfolgt, und die „**Kommissionierung parallel zur Tourenplanung**“ erfolgen soll, so muss die Auftrags-Tour-Zuordnung **irreversibel** sein. Dieses Prinzip erfüllt z.B. der ORA-Ansatz.

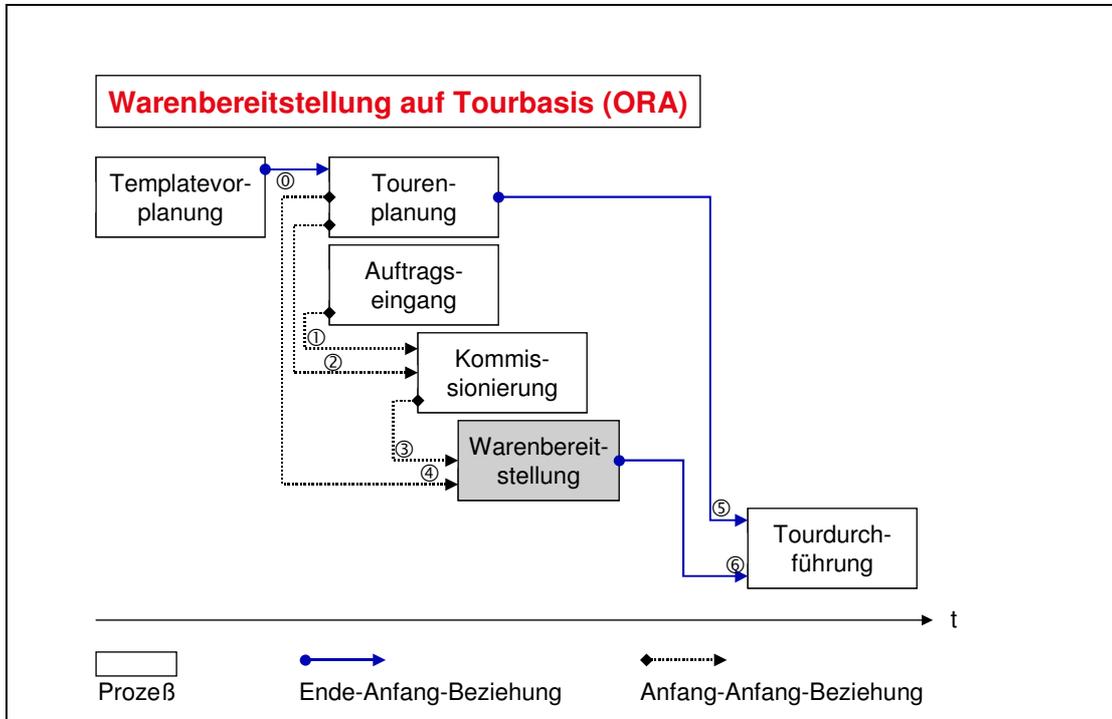


Abbildung 21: Prozessabhängigkeiten bei Kommissionierung parallel zur Tourenplanung

Wegen der irreversiblen Auftrags-Tour-Zuordnung müssen die Kommissionier- und Bereitstellungsprozesse nicht auf den Abschluss der vorgelagerten Prozesse (Auftragseingang und Tourenplanung) warten.

3.5.3 Warenbereitstellung auf Gebietsstellflächen

In manchen Anwendungsfällen (z.B. im Getränkeheimlieferservice) ist die Dauer des Kommissioniervorgangs so lang (z.B. wegen der großen physisch zu bewegendem Gewichte und Volumina), dass der Kommissionierbeginn i.d.R. nicht erst nach Auftragsannahmeschluss³⁷ und anschließender Tourenplanung liegen kann, da die verbleibende Zeitspanne bis zum Tourabgang zu kurz wäre.

Damit sowohl die Touren optimiert, als auch die nötigen Kommissionierzeiten berücksichtigt werden können, ist eine zeitliche Entzerrung der oben skizzierten Prozessabhängigkeiten notwendig.

³⁷ Aus Kundensicht soll der Auftragsannahmeschluss möglichst spät liegen.

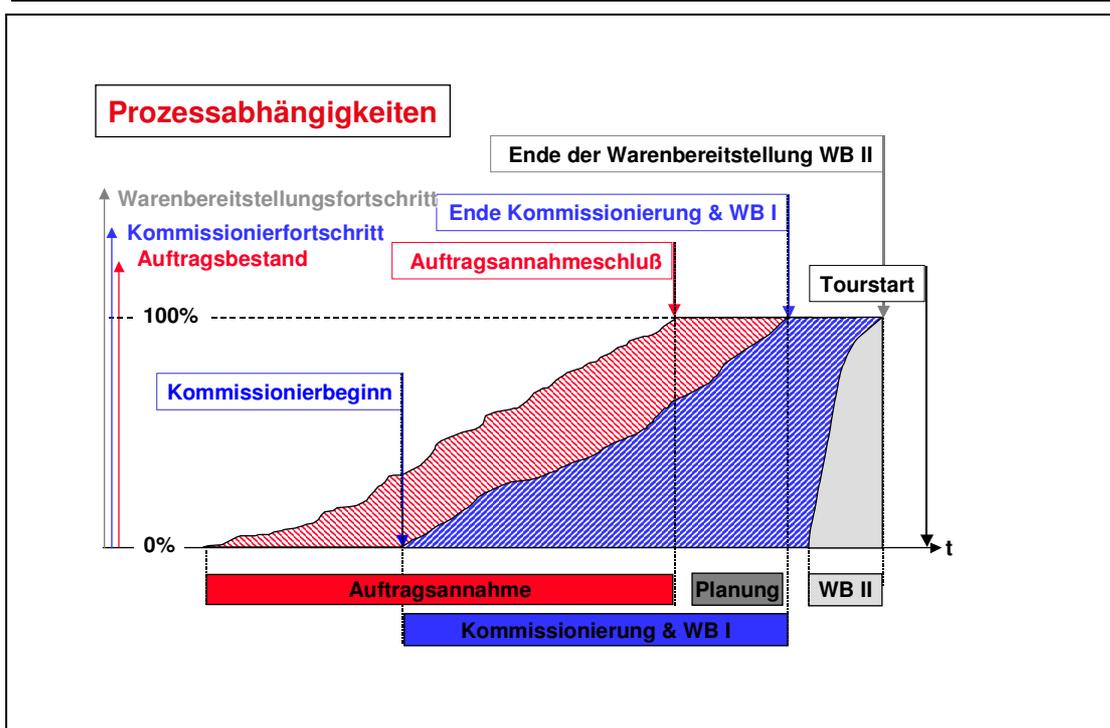


Abbildung 22: Prozessabhängigkeiten bei Nutzung von Gebietsstellflächen

Um diese Entzerrung zu ermöglichen, müssen die Waren bereits parallel zum Auftragseingang ohne exakte Kenntnis der Auftrags-Tour-Zuordnung kommissioniert werden.

Die vorkommissionierten Waren werden dann in sogenannten Gebietsstellflächen zwischengelagert (Warenbereitstellung WB I), da die tourspezifischen Bereitstellungsflächen noch nicht bekannt sind. Diese Zwischenlagerung auf Gebietsstellflächen erfolgt z.B. geordnet nach Postleitzahlen.

Nach Beendigung der Tourenplanung (wenn die Auftrags-Tour-Zuordnung bekannt ist) können die Waren aus den Gebietsstellflächen hin zu den tourspezifischen Bereitstellungsflächen umgelagert werden (Warenbereitstellung WB II).

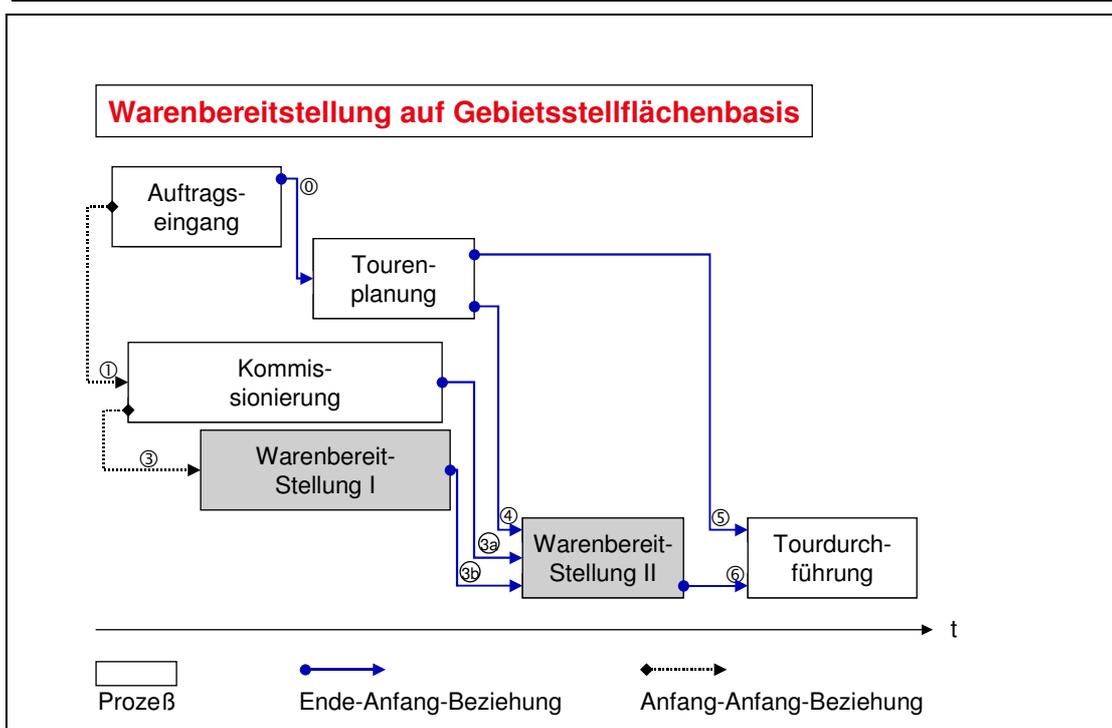


Abbildung 23: Prozessabhängigkeiten bei Nutzung von Gebietsstellflächen

3.6 Zusammenfassung

Die Marktbedürfnisse in Hinblick auf den Lieferservice sind je nach Anwendungsfall³⁸ unterschiedlich. Je nach dem, welche Lieferservice-Anforderungen aus den Marktbedürfnissen resultieren, müssen die entsprechenden logistischen Prozesse geschaffen werden.

Diese logistischen Prozesse (insbesondere die Teilprozesse der jeweils zugehörigen Warenbereitstellungsstrategie und ihre Abhängigkeit untereinander) sind ihrerseits entscheidend für die Einsetzbarkeit und den Erfolg von Planungsverfahren.

Beispiel: Aus OR-Sicht lässt sich ein Tourenplanungsproblem mit kontinuierlichem Auftragseingang als D-VRSP³⁹ modellieren. Wenn ein vorgeschlagenes OR-Verfahren zur Lösung dieses D-VRSP die Auftrags-Tour-Zuordnung immer wieder revidiert, um so bessere Ergebnisse zu erzielen, so ist die Anwendbarkeit des Verfahrens auf solche Bereitstellungsprozesse beschränkt, die mit sich verändernden Auftrag-Tour-Zuordnungen auch physisch umgehen können.

Um in der betrieblichen Praxis gute Ergebnisse im Spannungsfeld von Servicegrad und Kosten erzielen zu können müssen daher alle drei Elemente – die spezifischen

³⁸ Lebensmittel-Heimlieferservice, Getränke-Heimlieferservice, Lieferservice für Weiß- und Braunware, u.v.a.

³⁹ D-VRSP:= dynamic vehicle routing & scheduling problem

Marktanforderungen (zu definieren durch die Vertriebs- und Marketing Verantwortlichen), die logistischen Prozesse (zu definieren durch die entsprechenden Logistikverantwortlichen) und die Planungsverfahren (kontextspezifisch zu definieren durch die OR-Experten) aufeinander abgestimmt sein.

Dieser Aspekt ist wichtig, da insbesondere in der einschlägigen OR-Literatur der Schwerpunkt sehr oft auf den Verfahrensdetails, und weniger auf dem detaillierten logistischen Kontext liegt. Das ist zwar zulässig, aber die „logistische Forschung“ muss hier als Brückendisziplin die Vermittlung zwischen vertrieblichen und logistischen Anforderungen einerseits und den OR-Verfahren andererseits betreiben.

4 Stand der Forschung

Nachdem im vorangegangenen Kapitel vor allem die Aspekte der betrieblichen Praxis untersucht wurden, wird in diesem Kapitel der Stand der Technik in Hinsicht auf OR-Verfahren zusammengestellt, die zur Planung des Lebensmittelheim-lieferservice in Frage kommen können.

4.1 Einleitung

Das Problem des Heimlieferservicedienste mit Zeitzusagen zum Bestellzeitpunkt lässt sich als D-m-VRSP-TW-UC-DO (dynamic vehicle routing and scheduling problem with time windows, uncertain customers and delivery notification at order placement – fleet restricted to m vehicles) darstellen, wie später in dieser Arbeit gezeigt werden wird.

Um den Stand der Forschung zu diesem Problemtyp darzustellen, ist es sinnvoll, einleitend eine Übersicht über die benachbarten Problemklassen zu gewinnen.

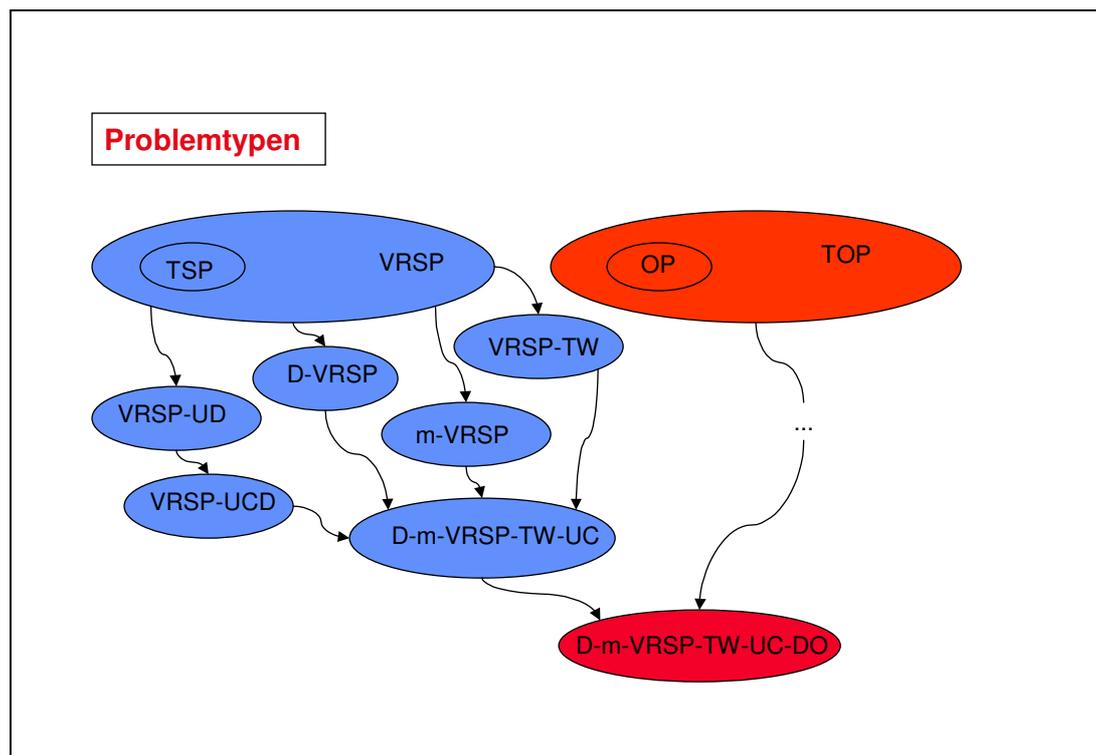


Abbildung 24: Problemtypen und Problemverwandschaften (Abkürzungen s.u.)

Die wesentlichen Verfahren zur „Tourenplanung im weiteren Sinne“ lassen sich einteilen in Verfahren zur Lösung von „vehicle routing and scheduling“-Problemen (VRSP) und „team orienteering“-Problemen (TOP).

Bei VRSP handelt es sich um Probleme, bei denen für eine Menge von Transportaufträgen ein Tourenplan erzeugt werden soll.

Bei TOP handelt es sich um Probleme, bei denen aus einer Anzahl von möglichen Lieferaufträgen diejenigen ausgewählt und in Touren verplant werden sollen, deren Abwicklung den höchsten Ertrag bringen. Es können daher durchaus auch Aufträge unerledigt bleiben. Einige Autoren sehen das TOP daher auch als eine Generalisierung des VRSPs.

4.1.1 Klassifikation der VRSPs

Tourenplanungsprobleme können ganz unterschiedliche Ausprägungen haben. In der Literatur werden die jeweiligen Tourenplanungsprobleme daher meist klassifiziert. Leider gibt es aber keine einheitliche Notation für diese Klassifikation.

Im Folgenden werden daher kurz zunächst die **Merkmalsklassen** für die Abgrenzung der unterschiedlichen Tourenplanungsprobleme aufgestellt, dann die **Problemklassen** in Hinblick auf die transportlogistischen Anwendungsfälle und auf formale OR-Probleme dargestellt, und schließlich eine einheitliche **Notation** für diese Arbeit eingeführt.

4.1.2 Merkmalsklassen

Schmidt ([Schmidt, 1994], S. 32f) führt in Anlehnung an [Bodin et al., 1983] die folgenden Merkmale zur Beschreibung eines Tourenplanungsproblems auf⁴⁰:

- ▶ **Anzahl der verfügbaren Fahrzeuge**
- ▶ Art der verfügbaren Fahrzeugflotte
- ▶ Anzahl der Depots
- ▶ Art der Nachfrage (deterministisch, stochastisch)
- ▶ Ort der Nachfrage
- ▶ Zugrundeliegender Graph⁴¹
- ▶ Kapazitätsbedingungen der verfügbaren Fahrzeuge
- ▶ Zeitbedingungen (**Kundenzeitschranken**, maximale Einsatzzeit)
- ▶ Art der Auftragsdurchführung
- ▶ Kostenstruktur
- ▶ Zielkriterien

⁴⁰ Die wichtigsten und in der OR-Literatur meist explizit ausgewiesenen Merkmale sind fettgedruckt.

⁴¹ symmetrisch vs. asymmetrisch; vgl. [Neumann, Morlock 1993, S.439 f]

Ein für diese Arbeit wesentliches zusätzliches Merkmal hinsichtlich der Bedienung von Aufträgen in zeitkritischen Konstellationen findet sich in [Grupp, 1983; S23 f] unter der abstrakten Bezeichnung:

- ▶ Szenario (**Ablehnung / Verspätung**)

Dieses Merkmal soll unterscheiden, ob entweder die Ablehnung von Aufträgen oder die Verletzung von vorgegebenen Zeitintervallen möglich sein soll, falls mit den verfügbaren Transportressourcen nicht alle Kunden innerhalb ihrer Kundenzeitschranken bedient werden können.

Mit Hilfe dieser und weiterer Merkmale können Tourenplanungsprobleme jeweils Anhand eines n-dimensionalen Merkmalsausprägungsvektors dargestellt werden.

4.1.3 Problemklassen der Transportplanung

Das Verständnis der jeweiligen Problemklassen erschließt sich aus **Sicht der Transportlogistik** am besten über die logistischen **Anwendungsfälle**⁴², welche durch diese Problemklassen jeweils beschrieben werden.

Einige beispielhafte Anwendungsfälle sind die Planung von:

- ▶ Einzeltouren mit mehreren Stops (Reihenfolgeoptimierung)
- ▶ umladungsfreien Komplettladungsverkehren
- ▶ umladungsfreien Teilladungsverkehren
- ▶ Vor- und Nachläufen beim gebrochenen expeditionellen Verkehr
- ▶ Ausliefertouren im Heimservice
- ▶ ...

Parallel dazu gibt es aus **Sicht des Operations Research (OR) abstrakte Problembeschreibungen**, die sich in der OR-Literatur etabliert haben. Diese sind

- ▶ Das „traveling salesman problem“ (TSP)
Das TSP lässt sich kurz beschreiben als die „Suche nach der distanzminimalen Rundreise durch alle gegebenen Auftragsorte“⁴³. Es entspricht dem Problem, das ein Handlungsreisender hat, der alle seine Kunden so besuchen möchte, dass der zurückgelegte Weg minimal ist. Das klassische TSP bezieht sich nur auf ein Fahrzeug, und berücksichtigt keine Zeitfenster an den Aufträgen.
- ▶ Das „vehicle routing problem“ (VRP)
Soll „eine gegebene Menge von Aufträgen mit einem oder mehreren Fahrzeugen transportiert werden, wobei die Gesamtfahrstrecke minimiert

⁴² [Schmidt, 1994, S.32f] unterscheidet in diesem Zusammenhang nach „logistischen Betriebsformen“

⁴³ vgl. [Grupp, 1983, S.1]

werden soll“ [Schmidt, 1994, S.32f], so handelt es sich um ein VRP. Der wesentliche Unterschied zum TSP ist, dass hier mehrere Fahrzeuge zur Verfügung stehen, die die Aufträge erledigen können. Das VRP berücksichtigt keine Zeitfenster an den Aufträgen.

- ▶ Das „vehicle scheduling problem“ (VSP)
Soll eine Mengen von Touren von einer Menge von Fahrzeugen abgewickelt werden, so muß eine Fahrzeugeinsatzplanung⁴⁴ (englisch: vehicle scheduling) erfolgen. Hierbei handelt es sich um ein VSP.
- ▶ Das „vehicle routing & scheduling problem“ (VRSP)
In praxi wird sowohl die Lösung des VRP als auch des VSP benötigt. Daher wird die Kombination beider Probleme meist als VSRP bezeichnet.
- ▶ Das „vehicle routing & scheduling problem with time windows“ (VRSP-TW)
Sollen beim VRSP Zeitfenster an den Aufträgen berücksichtigt werden, so spricht man vom VRSP-TW. Je nachdem wie im Konfliktfall mit diesen Zeitfenstern umgegangen wird, kann man weiter unterscheiden nach
 - ▶ order selection (m-VRSP-TW⁴⁵ bzw. VRSP-TW_{sel}⁴⁶)
einzelne Aufträge dürfen abgelehnt werden
 - ▶ soft time windows (VRSP-STW⁴⁷)
einzelne Aufträge dürfen zwar nicht abgelehnt werden, jedoch darf im Bedarfsfall das Zeitfenster verletzt werden (weiche Zeitfenster)

Im Folgenden wird unterschieden zwischen statischen und dynamischen Tourenplanungsproblemen. **Statische** Tourenplanungsprobleme sind Probleme, bei denen zum Planungszeitpunkt alle planungsrelevanten Größen (und insbesondere die zu planenden Transportaufträge) bekannt sind. **Dynamische** Tourenplanungsprobleme sind Probleme, bei denen zum jeweiligen Planungszeitpunkt die insgesamt zu planenden Transportaufträge noch nicht bekannt sind (rollierende Planung).

Außerdem wird unterschieden zwischen deterministischer und stochastischer Tourenplanung. **Deterministische** - oder klassische - Tourenplanung setzt voraus, dass die Eingangsgrößen des entsprechenden Optimierungsproblems vollständig bekannt sind. **Stochastische** Tourenplanung berücksichtigt auch Größen, die zum Planungszeitpunkt nicht bekannt sind, über die jedoch Annahmen über ihre Verteilung gemacht werden (z.B. geografisch gleichverteilt, zeitlich poissonverteilt, „im Gebiet x werden noch y Aufträge erwartet“, ...).

An dieser Stelle soll kurz eine Begriffsabgrenzung zwischen dynamischer und stochastischer Optimierung in Anlehnung an [Bent & van Hentenryck 2003]

⁴⁴ Ziel der Fahrzeugeinsatzplanung ist es zu bestimmen, welches Fahrzeug in welcher Reihenfolge welche Touren (job) abwickelt.

⁴⁵ m beschreibt die Anzahl der verfügbaren Fahrzeuge

⁴⁶ Der Zusatz „sel“ referenziert auf die $\alpha 7$ -Notation von [Domschke, 1997, S208]

⁴⁷ Der Zusatz „stw“ referenziert auf die $\alpha 5$ -Notation von [Domschke, 1997, S208]

erfolgen. Bei der *dynamischen Optimierung* sind die Kundenaufträge zum Planungszeitpunkt nicht alle bekannt, sondern laufen **inkrementell** über die Zeit ein. I.d.R. werden daher Tourenpläne für die bereits bekannten Aufträge generiert, und jeweils erweitert, wenn neue Aufträge eintreffen (Sukzessive Planung). Bei der *stochastischen Optimierung* wird hingegen ein Tourenplan **a priori erstellt**, der z.B. den Erwartungswert der Zielfunktion maximiert, d.h. dieser Tourenplan „antizipiert“ bereits die noch nicht bekannten Aufträge (ex ante Planung unter Ungewissheit).

4.1.4 Notation

Die Notation bezüglich der Planungsproblemklassen ist in der Literatur **nicht einheitlich**⁴⁸.

Eine **umfängliche Notation** der Problemklassen ist in [Domschke, 1997, S207f] zu finden. Ein klassisches Tourenplanungsproblem kann mit Hilfe dieser Notation zwar sehr detailliert als 4-Tupel beschrieben werden, aber diese Notation ist nicht „intuitiv verständlich“⁴⁹.

In der neueren Literatur hat sich daher eine **vereinfachte Notation** durchgesetzt die auch in dieser Arbeit verwendet wird. Sie ist in den folgenden Tabellen dargestellt. Außerdem wird in diesen Tabellen die OR-Sicht der logistischen Anwendungsfall-Sicht gegenübergestellt.

Notation	alternative Notationen	Beispiel für die Anwendung
TSP	VRP(1) ⁵⁰ , HRP ⁵¹	Handlungsreisendenplanung: Reihenfolgeoptimierung von Einzeltouren mit mehreren Stops
VRP	(M)TSP ⁵² , TPP ⁵³	Tourenplanung (ohne Fahrzeugeinsatzplanung) für Vorlauf im gebrochenen speditionellen Verkehr
VSP	-	Fahrzeugeinsatzplanung für Hauptlauf im gebrochenen Verkehr (Linienverkehr)
VRSP	-	Kombinierte Touren- und Fahrzeugeinsatzplanung

⁴⁸ Details finden sich u.a. in [Desrocher et al. 1990, S. 322f], [Schmidt, 1994, S. 32f], [Neumann, Morlock 1993, S. 439 f.], [Bodin et al., 1983, S. 63f], [Domschke, 1995, S. 207f] und [Grupp, 1983, S. 23 f.].

⁴⁹ Beispiele:

[1 | 1 | | |L] entspricht dem TSP,
[1,tw |M, cap,dur | |L] entspricht dem VRP-TW

⁵⁰ [Schmidt, 1994, S. 34] VRP für 1 Fahrzeug

⁵¹ [Neumann, Morlock 1993, S438 f.] HRP = Handlungsreisendenproblem

⁵² [Grupp, 1987, S10 f.] (M)TSP = multiple vehicle travelling salesman problem

⁵³ [Neumann, Morlock 1993, S438 f.] TPP = Tourenplanungsproblem

Notation	alternative Notationen	Beispiel für die Anwendung
VRSP-TW	TC-(M)TSP ⁵⁴	Nachlaufplanung beim gebrochenen speditionellen Verkehr mit Zeitfenstern
m-VRSP-TW	-	Auf m Fahrzeuge beschränkter Fuhrpark, (z.B. bei Spezialfahrzeugen) ohne die Möglichkeit der Ersatzfahrzeugbeschaffung

Tabelle 1: Notation klassischer Tourenplanungsprobleme

Damit dynamische Tourenplanungsprobleme schnell als solche erkenntlich werden, wird i.d.R. das Präfix „D-“ verwendet.

Notation	alternative Notationen	Beispiel für die Anwendung
D-*	-	Dynamische Planung
D-VRSP	-	Planung speditioneller Ladungsverkehre (FTL und LTL)
D-VRSP-STW ⁵⁵	TC-MTSP ^{v56}	Außendiensttechniker-Tourenplanung
D-m-VRSP-TW ⁵⁷	TC-MTSPA ⁵⁸	Home Delivery Tourenplanung mit Lieferavis zum Buchungszeitpunkt.

Tabelle 2: Notation dynamischer Tourenplanungsprobleme

Die Klassifikation der stochastischen VRSPs erfolgt i.d.R. nach der maßgeblichen stochastischen Einflußgröße. Diese Stochastizität kann sich beziehen auf

- ▶ Die **Liefermenge** (z.B. beim „Ab-Wagen-Verkauf“)

Dieser Problemtyp wird in der Literatur referenziert als **VRSP-UD** (Vehicle routing and scheduling with uncertain⁵⁹ demand)
- ▶ Die **Anzahl und Lage der Kunden** (z.B. wenn zum Tourenplanungszeitpunkt die genaue Lager aller Kunden deshalb noch nicht bekannt ist, weil noch nicht alle Bestellungen vorliegen)

Dieser Problemtyp wird in der Literatur referenziert als **VRSP-UC** (Vehicle routing and scheduling with uncertain customers)
- ▶ Die **Fahrzeit** (z.B. durch unvorhersehbare Staus)

VRSP-ST (Vehicle routing and scheduling with stochastic travel time)

⁵⁴ [Grupp, 1987, S11] TC-(M)TSP = Time-Constrained-(M)TSP

⁵⁵ STW:= soft time windows

⁵⁶ [Grupp, 1987, S1] TC-(M)TSPV = Verspätungsszenario, d.h. Verletzung der Zeitfenster möglich

⁵⁷ Dieses D-VRSP-Problem kann auch als D-TOP Problem gesehen werden (s.u.)

⁵⁸ [Grupp, 1987, S17] TC-(M)TSPA = Ablehnungsszenario, d.h. Ablehnung von Aufträgen möglich

⁵⁹ In der Literatur werden die Begriffe „uncertain“ und „stochastic“ oft synonym gebraucht.

Die Problemfälle der stochastischen Fahrzeit werden in dieser Arbeit nicht weiter untersucht

Notation	alternative Notationen	Beispiel für die Anwendung
VRSP-UD	VRPSD	Ab-Wagen Verkauf von Fahrzeugen mit Vollsortiment Stückgutsammeltouren bei Stammkunden
VRSP-UC	VRPSC	Home Delivery Tourenplanung mit Buchung über Internet oder Call-Center ohne zu erwartende Fahrzeugkapazitätsengpässe
VRSP-UCD	VRPSCD	Home Delivery Tourenplanung mit Buchung über Internet oder Call-Center mit zu erwartenden Fahrzeugkapazitätsengpässen
VRSP-ST	TD-VRSP ⁶⁰	Tourenplanung mit Berücksichtigung der zu erwartenden Staus

Tabelle 3: Notation stochastischer Tourenplanungsprobleme

4.1.5 Komplexität

VRP-Probleme gehören zur Klasse der NP-schwierigen Probleme (vgl. [Domschke, 1997, S.43]). Für die Klasse der NP-schwierige Probleme gilt, daß „bisher kein Algorithmus bekannt [ist], der eines der Probleme garantiert mit polynomialem Rechenaufwand optimal löst“ ([Domschke, 1995, S.37-38]).

Die Auswirkung dieser Erkenntnis ist, dass selbst für relativ kleine Probleme keine optimale Lösung in akzeptabler Zeit gefunden werden kann. Eine kleine Übersicht zeigt dies deutlich:

⁶⁰ Time dependent VRSP

Rechen- aufwand	Größe des Problems		
	n=10	n=30	n=50
n	10^{-5} sec	$3 \cdot 10^{-5}$ sec	$5 \cdot 10^{-5}$ sec
n^3	10^{-3} sec	$27 \cdot 10^{-3}$ sec	0,125 sec
n^5	0,1 sec	24,3 sec	5,2 min
2^n	10^{-3} sec	17,9 min	35,7 Jahre
3^n	0,059 sec	6,5 Jahre	$2 \cdot 10^{10}$ Jahre

Tabelle 4: Rechenaufwandsübersicht aus [Domschke, 1995, S.37-38]
Annahmen: n sei Indikator für Problemgröße; ein Rechenschritt dauert 10^{-6} sec.

Aus Sicht der logistischen Praxis besteht der Bedarf an schnellen Verfahren, die auch für größere Problemgrößen (mehrere hundert Lieferpunkte) gute Ergebnisse liefern. In der Literatur sind einige Näherungsverfahren (Heuristiken) zur Lösung des VRP beschrieben, auf die im Folgenden noch näher eingegangen wird.

[Solomon, 1986] hat für einige Heuristiken für das VRSP-TW nachgewiesen, dass diese mit einem „worst case“ Rechenaufwand von $O(n)$ eine Lösung finden. Im Vergleich zu exakten Verfahren, die einen „worst case“ Rechenaufwand von $O(x^n)$ erfordern ist dieser signifikante Performance-Vorteil dieser Heuristiken gegenüber exakten Verfahren ausschlaggebend für die Praxistauglichkeit.

4.2 Das VRSP

Das „vehicle routing and scheduling problem“ (VRSP) hat eine Fülle von Ausprägungen. In diesem Abschnitt wird zum einen ein kurzer Überblick über die gängigen Lösungsansätze gegeben, und zum anderen auf die Klasse der VRSP-UCD fokussiert, um diese vom Untersuchungsgegenstand dieser Arbeit, dem D-m-VRSP-TW-UC-DO abzugrenzen, welches später noch genau definiert wird.

4.2.1 Verfahren für das VRSP

Es gibt eine Fülle von Verfahren zur Lösung des klassischen (nicht stochastisch, nicht dynamisch) VRSP. Eine Auswahl der wichtigsten Verfahren (Basisroutinen und Meta-Steuerungsverfahren) sind im Anhang dieser Arbeit kurz beschrieben.

4.2.2 Verfahren für das VRSP-TW

Ein wesentlicher Aspekt dieser Arbeit widmet sich der Untersuchung von bekannten

Lösungen zum VRSP-TW anhand der Solomoninstanzen (s. 6.1.1).

[Solomon, 1984] setzte für die Lösung des VRSP-TW sowohl ein Savings-basiertes Verfahren als auch ein Insertion Verfahren ein. Die neueren Verfahren, die die verschiedenen Autoren dazu eingesetzt haben, werden hier sortiert nach Jahr der Veröffentlichung grob dargestellt.

4.2.2.1 Das Verfahren von Rochat & Taillard 1995

[Rochat & Taillard, 1995] versuchen das intrinsische Problem von „local search“ Heuristiken, nämlich die Sequenzialität der Verfahren und der dadurch verursachten langen Rechenzeiten durch Parallelisierung zu lösen.

Eine wesentliche Idee von [Rochat & Taillard, 1995] ist es, das Problem in Teil-Probleme⁶¹ zu zerlegen (**partitioning**), für jedes Teil-Problem eine Lösung zu suchen, die Teil-Lösungen zu einer Gesamtlösung zusammenzufügen, und diese Gesamtlösung wieder aufzuteilen usw.⁶². So lassen sich Rechenzeiten signifikant reduzieren.

Für ein so generiertes Teilproblem schlagen [Rochat & Taillard, 1995] folgendes Lösungsvorgehen vor (**diversification and intensification**):

Zuerst wird eine Menge I an Tourenplänen generiert. Anschließend werden alle Touren mit dem Zielfunktionswert der jeweiligen Lösung markiert, und nach ZF-Wert in einer Liste T sortiert. Anschließend werden Touren aus T stochastisch selektiert und in eine Menge S eingefügt - per Definition haben Touren mit einem niedrigeren ZF-Wert dabei eine höhere Auswahlwahrscheinlichkeit. Touren, die Stops enthalten, die bereits in S vorkommen, werden nicht in S aufgenommen. Können aus T keine Touren mehr in S übernommen werden und sind gleichzeitig noch nicht alle Stops in S enthalten, so wird S mit Stichtouren aufgefüllt (S'). Schließlich wird durch „local search“ die Lösung S' noch zu S'' verbessert. S'' fließt dann wieder in T ein usw. bis ein Abbruchkriterium erreicht ist.

⁶¹ mit jeweils 4-8 Touren bzw. 30-60 Stops pro Teil-Problem

⁶² Wie genau die Aufteilung jeweils erfolgt ist in [Rochat & Taillard, 1995] nicht genau beschrieben.

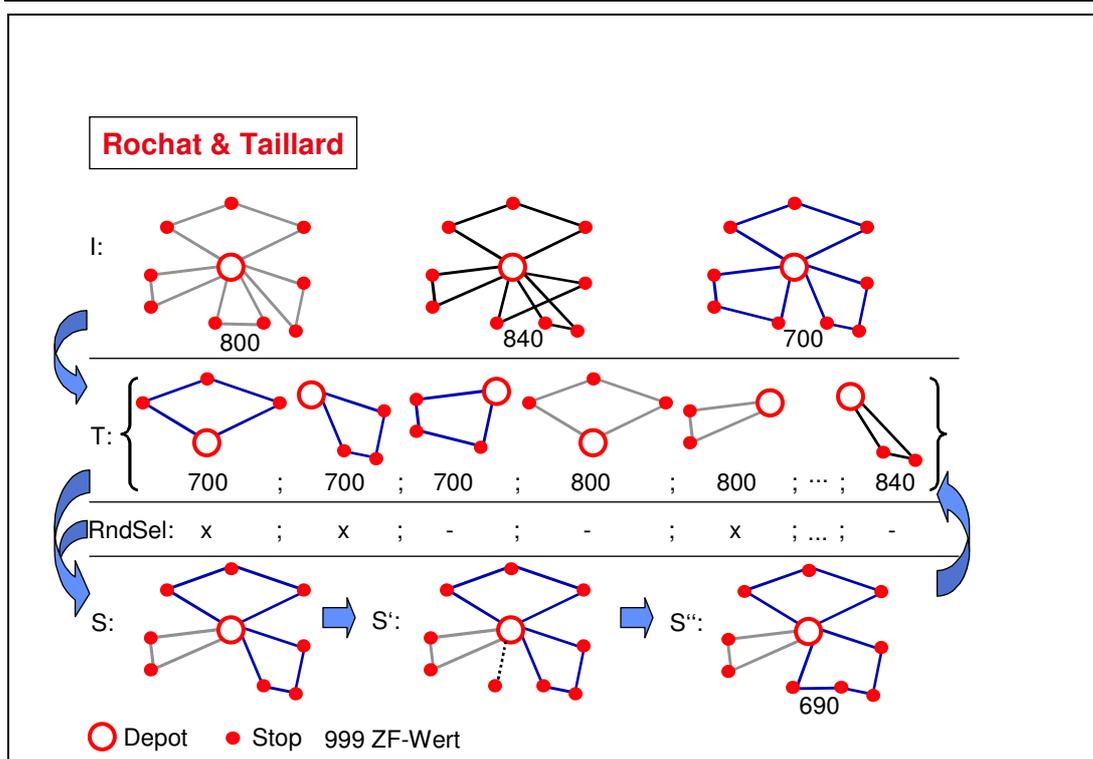


Abbildung 25: Verfahren von [Rochat & Taillard, 1995]

Im Pseudocode lässt sich das Vorgehen folgendermaßen beschreiben:

```

proc diversification and intensification algorithm
begin
    generate I different solutions with local search
    label each tour with the value63 of the solution it belongs to
    remove the tours having only one customer
    insert the remaining tours in a set of T tours
    repeat till stopping criterion is fulfilled
        sort T by increasing values of labels
        set T'=T; S:= ∅
        while T'≠ ∅ repeat
            chose t ∈ T' probabilistically based on its current relative
                evaluation
            set S:=S∪{t}
            remove from T' all tours including any customers belonging to t
        end while
        if customers are not covered by the tours of S, construct a feasible
            solution S' by including them using the partial solution S
        improve S' -> S'' by means of local search
        label tours of S''
        remove tours with only 1 customer
        insert the remaining tours in T
    end repeat
end
    
```

Abbildung 26: Prinzip des Verfahrens nach [Rochat & Taillard, 1995]

⁶³ „value“ ist hier der Zielfunktions-Wert der jeweiligen Lösung

Variablen & Funktionen:

I: Initialmenge an Tourenplänen

T: Liste von Touren, die in der aktuellen Lösungsmenge vorkommen;

T ist sortiert nach den ZF-Werten der zugehörigen Lösungen

T': Liste von Touren, die sukzessive geleert wird, bis keine Touren mehr in *T'* sind (Initiierung: $T' := T$)

S: sukzessive wachsende Teillösung des Problems

(Menge von Touren)

S': Vervollständigung von *S* und zulässige Lösung des Problems

(Menge von Touren)

S'': Verbesserung von *S'*

Der Initialisierungsschritt (*I*) und die Verbesserungsschritte (*S'*->*S''*) sind dabei am zeitaufwändigsten. Genau diese Schritte lassen sich jedoch parallelisieren. Die Kombination von Merkmalen verschiedener Lösungen in diesem Ansatz hat im Übrigen starke Ähnlichkeit mit genetischen Verfahren.

4.2.2.2 Das Verfahren von Taillard et al 1997

[Taillard et al, 1997] schlagen ein „tabu search“ Verfahren für das VRP-STW vor, und wenden dieses auch auf die klassischen Solomoninstanzen⁶⁴ für das VRP-TW an. Sie schlagen ein neues Austausch-Verfahren („cross exchange“ s. Anhang) vor, das auch von anderen Autoren aufgegriffen wird (s.u.).

```
proc Taillard et al Algorithm
begin
  generate set of s seed solutions by means of stochastic insertion heuristics
  apply tabu search to the s seed solutions and store resulting routes in
  adaptive memory AM
  repeat until stop criterion is met
    construct initial solution from routes in AM, and define this solution as
    current solution S
    for i:= 1 to W do
      decompose solution S into C disjoint subsets of routes
      apply a tabu search on each subset of routes
      reconstruct a complete solution by merging the new solutions for
      the C subsets, and define result as new current solution S
    end for
    store routes of S in AM
  end repeat
  apply post-optimization method to each route of S
end
```

Abbildung 27: Prinzip des Verfahrens nach [Taillard et al, 1997]

⁶⁴ Die Solomoninstanzen sind im Anhang dieser Arbeit beschrieben.

Variablen & Funktionen:

s: Anzahl der Ausgangslösungen

S: aktuelle Lösung

C: Teilmenge von Routen von S

AM: das adaptive memory (AM) ist eine sortierte Liste von Touren, die in den besten gefundenen Lösungen enthalten und nach der ZF-Güte ihrer zugehörigen Lösung sortiert sind

W: Obergrenze für Anzahl Iterationen

Bei der Initialisierung (initialization) des Verfahrens werden viele verschiedene Ausgangslösungen geringer Qualität generiert (reines Random Insertions-Verfahren), die dann jeweils durch tabu search verbessert werden.

Die Dekomposition (decomposition) in der for-Schleife erfolgt nach dem „Sweep“-Prinzip so, dass jedes „subset“ etwa gleich viele Touren enthält. Nach Lösung der einzelnen subset-Probleme werden die subset-Lösungen wieder zu einer Gesamtlösung zusammengesetzt (reconstruction)

Aus den im adaptive memory (AM) enthaltenen Touren werden nach dem Verfahren von [Rochat & Taillard, 1995] neue Lösungen kreiert.

Das eigentliche tabu search Verfahren arbeitet nach folgendem Prinzip:

```
proc tabu search (subset C)
begin
  set C*=C':= initial set of routes solving the subset-problem
  while stopping criterion is not met do
    generate neighbourhood of C' by applying cross-exchanges
    C':= best non-tabu solution in this neighbourhood
    if C' better than C* then
      improve C' by Solomons I1-insertion-procedure, if possible
      C*:=C':= improved C'
    end if
    update tabu list
  end while
  return C*
end
```

Abbildung 28: Prinzip des tabu search Verfahrens nach [Taillard et al, 1997]

Variablen & Funktionen:

C': aktuelle Lösung für das subset-Problem (current subset-solution)

C: beste bisher gefundenen Lösung für das subset-Problem (best overall subset-solution)*

4.2.2.3 Das Verfahren von Shaw 1998

[Shaw, 1998] schlägt ein Verfahren vor, bei dem das VRP-TW durch gezieltes Entfernen (Problemrelaxation) und Wiedereinfügen der Stops gelöst wird. Er nennt das Verfahren „large neighbourhood search“ (LNS). Motiviert wird das Verfahren dadurch, dass für das Wiedereinfügen von Stops in eine bestehende Lösung exakte „branch & bound“ Verfahren zur Verfügung stehen. Das Kernproblem in dem Verfahren von [Shaw, 1998] ist zu entscheiden, welche Stops jeweils entfernt werden sollen.

Das Entfernen von Kunden kann prinzipiell auf vielerlei Weise erfolgen („myriad such methods“). Shaws Idee ist es, falls irgendein Stop entfernt wird, auch die Stops zu entfernen, die in hohem Maße mit dem Stop „zusammenhängen“. [Shaw, 1998] schlägt daher eine „relatedness function“ $R(i,j)$ vor, die ein Maß dafür angibt, wie stark die Stops i und j zusammenhängen. Er benutzt die Funktion:

$$R(i, j) = \frac{1}{c_{i,j} + V_{i,j}} \text{ mit}$$

c_{ij} : = Maß für Entfernung zwischen i und j normalisiert auf $[0;1)$
 V_{ij} : = 1 falls i und j vorher in selber Tour waren; 0 sonst.

```
proc Remove Visits
begin
  remove a stop from I to R by random
  while |R| < r do
    select a stop v from R by random
    sort stops in I by relatedness  $R(i,v)$ 
    remove stop from sorted I at position  $\lceil |I| * \text{rnd}^D \rceil$  to R
  end while
end
```

Abbildung 29: Prinzip des Verfahrens nach [Shaw, 1998] – Entfernen von Stops (eigene Darstellung)

Variablen & Funktionen:

r: Anzahl Kunden, die entfernt werden sollen

I: Menge der Kunden im Plan

R: Menge der entfernten Kunden

v: selektierter Stop

rnd: Zufallszahl aus $[0;1)$

D: Exponent, der Zufälligkeit beeinflusst aus $[1;\infty)$:

1 => Zufallsauswahl,

∞ => erster Kandidat in sortierter Liste (mit höchster Relatedness zu v)

Das Entferungsverfahren entfernt somit im Prinzip insgesamt r Stops mit hohem Zusammenhang (relatedness) zu bereits entfernten Stops – dabei wird die Auswahl

jedoch durch eine Zufallsgröße gestört, um eine nichtdeterministische Auswahl zu ermöglichen (=>Vermeidung des Festfahrens des Verfahrens in Auswahl).

Die neighbourhood size des Verfahrens von Shaw wird durch r gesteuert. Nach [Shaw, 1998] sollte $r < 30$ gewählt werden. Er schlägt vor, r bei 1 starten zu lassen, und nur dann um 1 zu erhöhen, falls in einem Durchlauf keine Verbesserung der ZF gefunden werden konnte.

Das Einfügeverfahren ist ein branch-and-bound-Verfahren. Zur Performancesteigerung nutzt Shaw zum einen propagation rules (Prüfkriterien zum Ausschluß unzulässiger Lösungen), die den Suchbaum signifikant verkleinern. Zum anderen nutzt er für die Suche im Suchbaum Such-Heuristiken, die er im Domänenwissen über das Problem begründet (weitere Details s. [Shaw, 1998]).

4.2.2.4 Die Verfahren von Homberger & Gehring 1999

[Homberger & Gehring, 1999] schlagen zwei zweistufige Evolutions-Verfahren⁶⁵ ES1 und ES2 vor, die als vorrangiges Optimierungsziel die Minimierung der Fahrzeuganzahl und als nachrangiges die Minimierung der Tourkilometer haben.

Jedes Individuum einer Population ist dabei ein 3-Tupel aus einer gültigen Lösung und zwei Parametern:

<SV *SolutionVector* (gültige Lösung des VRSP-TW);
MC *MutationCode* (Mutationssvorschrift);
RE *RElimination* (Mutationsrichtung $\in [0;1]$)>.

Die Mutationsrichtung „0“ gibt an, dass bei fester Fahrzeuganzahl nur die km reduziert werden sollen, „1“ gibt an, dass außerdem versucht werden soll, die Anzahl der Fahrzeuge zu verringern. Die Mutationsrichtung wird zufällig bestimmt.

Im Verfahren ES1 (evolution strategy 1) ist der Mutationscode definiert als Folge von n zufälligen Mutationsschritten aus:

Or-opt-move (entspricht „string relocation“ - s. [Homberger & Gehring, 1999]),
2-opt*-move (entspricht „string cross“⁶⁶ - s. [Potvin & Rousseau 1995]) oder
1-interchange-move (s. [Osman 1993]).

Falls die „Mutationsrichtung = 1“ erfolgt zusätzlich der Schritt:

modified-Or-opt-move, dessen Ziel es ist, die „kürzeste“ Tour zu eliminieren.
D.h. ein Elternteil erzeugt durch Mutation ohne Rekombination Nachkommen.
(=RElimination)

⁶⁵ Genetische Verfahren mit problemorientierter Rekombination (d.h. nicht mit Chromosomen-Rekombination)

⁶⁶ s. Anhang

Im Verfahren ES2 (evolution strategy 2) wird im Gegensatz zu ES1 die Mutationsvorschrift (MutationCode MC) nicht als Folge von n zufälligen Mutationsschritten extern vorgegeben, sondern ist das Ergebnis einer Rekombination der MutationCodes zweier Eltern. Der so erzeugte MutationCode wird dann auf den SolutionVector (SV) eines dritten Elternteils angewendet. D.h. drei Elternteile erzeugen einen Nachkommen nach dem folgenden Schema:

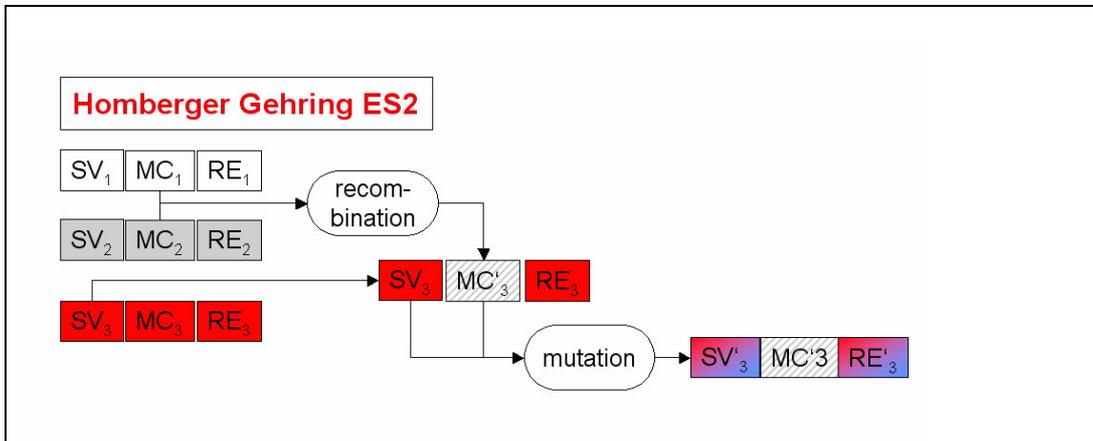


Abbildung 30: Homberger Gehring ES2

Der MutationCode ist dabei eine Folge, in der jeder der Stopindex genau zweimal vorkommt. Beim Abarbeiten des MutationCodes werden Stops erst entfernt, und später nach einem speziellen Insertionsverfahren wieder eingefügt.

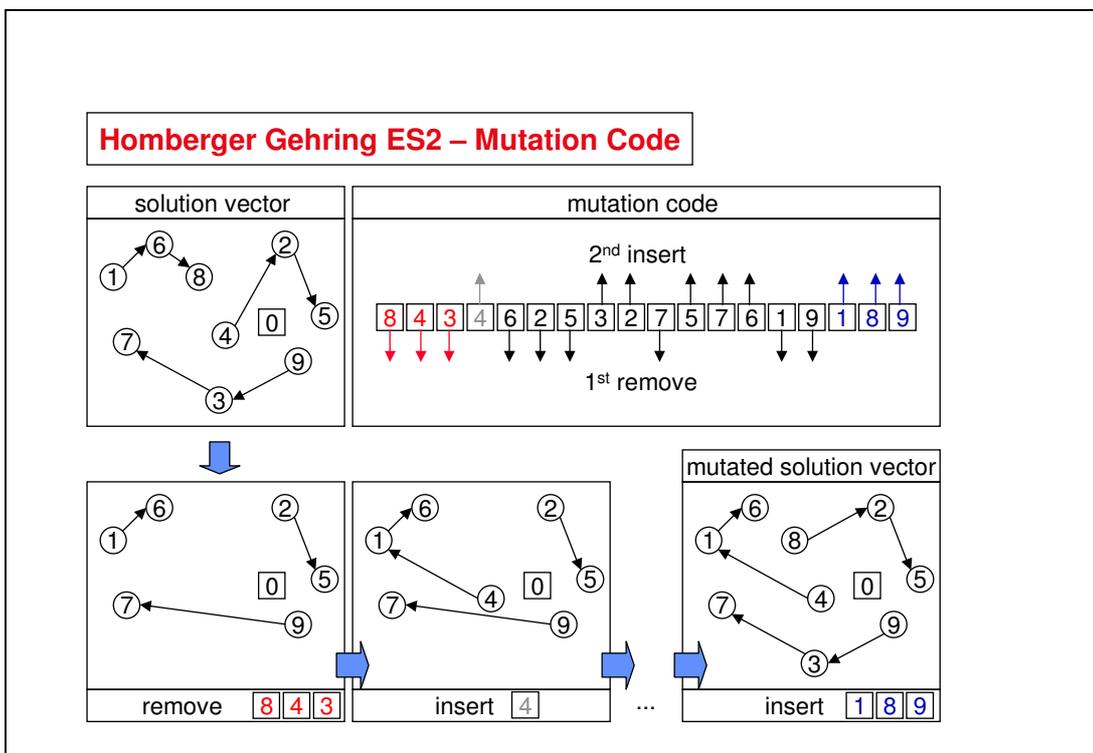


Abbildung 31: Homberger Gehring ES2 – Prinzip des Mutation Codes und Ablauf der Mutation

Da das Ziel der Optimierung sowohl die Minimierung der Fahrzeuganzahl als auch

die Minimierung der Gesamttourkilometer ist, sind für die Elterngeneration zwei Merkmale wünschenswert, nämlich zum einen **geringe Gesamttourkilometer** und zum anderen **kurze Resttouren**, die ggf. in späteren Schritten noch aufgelöst werden können.

Daher wird die Fitness eines Individuums entweder definiert als lexikografische Ordnung der Kriterien $\langle \text{vehicles}; \text{length} \rangle$ oder als Ordnung $\langle \text{vehicles}; \text{min. late} \rangle$ mit:

- ▶ $\langle \text{vehicles} \rangle$ - Anzahl Fahrzeuge
- ▶ $\langle \text{length} \rangle$ - gesamter Tourweg
- ▶ $\langle \text{min. late} \rangle$ - „minimale Verspätung“, mit der sämtliche Stops der Resttour in andere Touren eingefügt werden können (als Schätzer für „Auflösbarkeit“ kleiner Resttouren)

```

proc ES67
begin
  generate initial child population C(0) of  $\lambda$  3-tuples  $\langle \text{SV}, \text{MC}, \text{RE} \rangle$ :
    SV: by using random selection from a savings algorithm's savings list
    MC: by random generation
    RE: by random generation
  repeat
    sort C(t) by lexicographic fitness  $\langle \text{vehicles}; \text{min. late} \rangle$ 
    select first  $\kappa$  individuals from C(t) and include them in P(t+1)
    sort C(t) by lexicographic fitness  $\langle \text{vehicles}; \text{length} \rangle$ 
    select first  $\mu - \kappa$  individuals from C(t) and include them in P(t+1)
    if timeout goto (*)
    set t:=t+1
    generate C(t) = { $\lambda$  mutations out of P(t)} according to ES* mutation rule
  end repeat
  (*) return best SV from P(t)
end
    
```

Abbildung 32: Prinzip des Verfahrens von [Homberger & Gehring, 1999] – eigene Darstellung

Variablen & Funktionen:

t:= Generationszähler

P(t):= Eltern Population von 3-Tupeln $\langle \text{SV}, \text{MC}, \text{RE} \rangle$

C(t):= Kinder Population von 3-Tupeln $\langle \text{SV}, \text{MC}, \text{RE} \rangle$

μ := |P| Größe der Elternpopulation

λ := |C| Größe der Kindergeneration (vor Selektion)

$\kappa < \mu$:= Größe der ersten Teilselektion für neue Elterngeneration

$\mu - \kappa$:= Größe der zweiten Teilselektion für neue Elterngeneration

⁶⁷ ES= ES1 oder ES2 je nach verwendeter Mutationsvorschrift.

4.2.2.5 Die Verfahren von Ibaraki et al 2001

[Ibaraki et al 2001] schlagen drei verschiedene „dynamic programming“ - Heuristiken vor (MLS⁶⁸, ILS⁶⁹, AMLS⁷⁰), die im Kern jeweils auf ein spezielles „local search“ - Verfahren (LS) zugreifen. Sie formulieren das VRSP-TW als VRSP-STW, und bewerten Zeitrestriktionsverletzungen mit Strafkosten

Das LS sucht dabei eine „very large scale neighbourhood“ ab, und arbeitet folgendermaßen:

```
proc LS()
begin
  import solution
  (*)
  repeat
    improve solution by „intra route exchange“ (s. [Ibaraki et al 2001])
    improve solution by „2-opt**“ (s. [Potvin & Rousseau 1996])
    improve solution by „cross exchange“ (s. [Taillard et al 1997])
  until no improvement in 3 consecutive loops
  improve solution by „cyclic exchange“ (s. [Ibaraki et al 2001])
  if improvement achieved by cyclic exchange then goto (*)
  else
    export (improved) solution
end
```

Abbildung 33: Prinzip des LS von [Ibaraki et al 2001] (eigene Darstellung)

Im Folgenden ist die übergeordnete Heuristik ILS kurz skizziert⁷¹.

```
proc ILS
begin
  generate initial solution  $\sigma^o$  (initiale Ausgangslösung)
   $\sigma^{seed} := \sigma^o$  (aktueller local search Fokus)
   $\sigma^{best} := \sigma^o$  (beste bisher gefundenen Lösung)
  (*)
   $\sigma := LS(\sigma^{seed})$ 
  if best-eval( $\sigma$ ) < best-eval( $\sigma^{best}$ ) then  $\sigma^{best} := \sigma$ 
  if cost( $\sigma$ ) < cost( $\sigma^{seed}$ ) then  $\sigma^{seed} := \sigma$ 
  if termination-criterion=true then export  $\sigma^{best}$  & terminate
  else  $\sigma^{seed} := perturb(\sigma^{seed})$ 
  goto (*)
end
```

Abbildung 34: Prinzip des ILS von [Ibaraki et al 2001] (eigene Darstellung)

⁶⁸ MLS: multi start local search

⁶⁹ ILS: iterative local search

⁷⁰ AMLS: adaptive start local search

⁷¹ MLS und AMLS s. [Ibaraki et al 2001]

Variablen & Funktionen:

$cost(\sigma) = d(\sigma) + p(\sigma)$ *Kosten der Lösung inkl. Strafkosten mit*

$d(\sigma)$: *Tourlängenkosten*

$p(\sigma)$: *Strafkosten für Zeitfensterverletzungen*

$best\text{-}eval(\sigma)$: *Evaluationsfunktion für Lösung mit*

$$best\text{-}eval(\sigma) = \begin{cases} cost(\sigma) & \text{falls } p(\sigma) = 0 \\ cost(\sigma) + \beta & \text{sonst} \end{cases}$$

$perturb(\sigma)$: *Verändert die Lösung durch mehrfachen zufälligen „cross exchange“
Sie soll das Steckenbleiben in einem lokalen Optimum verhindern.*

Es ist zu beachten, dass $cost(\sigma)$ auch für solche Lösungen berechnet wird, die die Zeitfenster verletzen. Wird jedoch β groß genug gewählt (d.h. $\forall \sigma: \beta > cost(\sigma)$), so bleibt die Lösung mit kleinstem $best\text{-}eval(\sigma)$ stets gültig, wenn die Ausgangslösung σ° gültig war.

Anzumerken ist, dass das Verfahren relativ langsam ist - [Ibaraki et al 2001] geben für ihre Ergebnisse Rechenzeiten von 2.000 sec. (33min.) bis 15.000 sec. (4h 10 min.) für die Solomoninstanzen mit 100 Kunden an.

4.2.2.6 Das Verfahren von Lau et al 2003

Auch wenn [Lau et al 2003] noch keine Dynamik im VRSP-TW zulassen, so gehen sie bereits den Ideen „Maximiere Anzahl Kunden“, „Lasse Nichtbelieferungen zu“ und „Lasse Zeitrestriktionsverletzungen zu“ nach. Sie untersuchen das m-VRSP-TW.

Ein wesentlicher Aspekt ihrer Arbeit ist die Definition einer hierarchischen Gütefunktion um das Wichtungsproblem konkurrierender Ziele zu umgehen (s. Ausführungen zum Wichtungsproblem im Kapitel 6 dieser Arbeit). Sie definieren daher eine Gütefunktion mit der folgenden Hierarchie:

1. *Maximiere Anzahl bedienter Kunden*
2. *Minimiere Anzahl der Kunden, die zu spät beliefert werden*
3. *Minimiere Verspätungssumme*
4. *Minimiere Anzahl Fahrzeuge*
5. *Minimiere Gesamttourkilometer*

Im Rahmen eines iterativen Prozesses versuchen sie, je Schritt die gemäß der Gütefunktion beste Lösung zu finden. Sollte diese Lösung nicht zulässig sein, so wird die Anzahl an Fahrzeugen erhöht.

Im Pseudocode lässt sich dieses Verfahren (OV) folgendermaßen beschreiben:

```

proc OV – Algorithm A
begin
  initialize numVeh
  until holding list is empty or numVeh=m
    set Count=0
    while Count≤CountLimit
      call TS based on numVeh vehicles
      if better solution found then set Count=0
      else set Count=Count+1
    end while
    set numVeh = min(numVeh+Stepsize, m)
  end until
end

```

Abbildung 35: Prinzip des OV nach [Lau et al 2003]

Variablen & Funktionen:

holding list: Liste von Stops, die keiner Tour zugeordnet sind

TS: ein Schritt in einer TabuSearch Heuristik

numVeh: Fahrzeuganzahl

Count: Zählvariable für die TS-Aufrufe mit Obergrenze CountLimit

Stepsize: Schrittweite für Erhöhung der Fahrzeuganzahl

D.h. der Algorithmus A erhöht in der äußeren Schleife sukzessive die Anzahl der Fahrzeuge, und in der inneren Schleife versucht er, die ZF zu verbessern. Das Kernverfahren ist dabei eine Standard-Tabu-Search-Heuristik TS, die zwar nicht in detail definiert wird, aber die oben beschriebene Gütefunktion verwendet, und mit Hilfe der folgenden 3 Schritte operiert:

„relocate from holding list“ (Einfügen eines unverplanten Stops in eine Tour)

„relocate to holding list“ (Lösen eines Stops aus seiner Tour), und

„exchange with holding list“ (Vertauschen eines unverplanten Stops mit einem Stop, der einer Tour zugeordnet ist).

Neben dem Algorithmus A schlagen [Lau et al 2003] noch einen Algorithmus B vor, der als Schale um Algorithmus A sukzessive die Zeitfenster aufweitet, bis alle Stops verplant sind⁷².

Das Verfahren von Lau et al ist deshalb interessant, da es eigentlich ein m-VRSP-TW zu lösen versucht. Ist die Anzahl der verfügbaren Fahrzeuge m groß genug, so liefert das Verfahren eine Lösung für das entsprechende VRSP-TW bei dem möglichst wenige Fahrzeuge eingesetzt werden. Ist die Anzahl der verfügbaren Fahrzeuge m jedoch zu klein, so liefert das Verfahren entweder eine Lösung für das m-VRSP-STW (falls die Zeitfenster verletzt werden, aber alle Kunden bedient werden) oder für das TOP (falls nicht alle Kunden bedient werden und somit übrig bleiben). Die Dynamik im Planungsproblem des Lebensmittelheimlieferservice berücksichtigen Lau et al. allerdings nicht.

⁷² Diese Idee wird in dieser Arbeit bei der Analyse des VRSP-TW aufgegriffen.

4.2.2.7 Das Verfahren von Tavares et al 2003

Tavares veröffentlicht in [Tavares 2004] Lösungen zu den Solomoninstanzen. Dort nennt er das „genetic vehicle representation“-Verfahren GVR als das Verfahrensprinzip, mit dem die Ergebnisse generiert wurden. Das Verfahren selbst nennt er evolutionary computation EC und es ist beschrieben in [Tavares, Pereira et al 2003] und in [Pereira, Tavares et al 2002].

Tavares et al. [Tavares, Pereira et al 2003] schlagen einen genetischen Ansatz vor, bei dem das Chromosom einer Lösung unmittelbar eine Codierung der Touren darstellt:

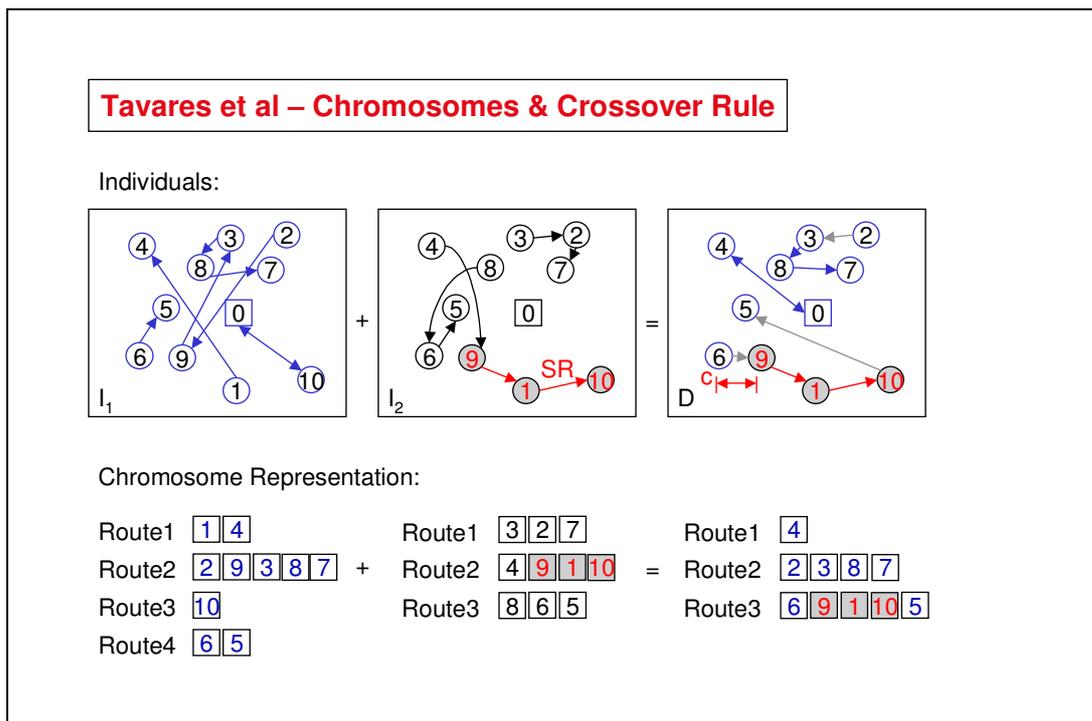


Abbildung 36: Prinzip der Codierung und Rekombination EC nach [Tavares 2004], [Tavares, Pereira et al 2003] und [Pereira, Tavares et al 2002]

Als Rekombinationsregel schlagen [Tavares, Pereira et al 2003] folgendes Teilverfahren vor:

```

proc crossover
begin
  for each individual  $I_1$  from the selected set S repeat
    randomly select another Individual  $I_2$  from S
    from the genetic material of  $I_2$  randomly select
      a sub route  $SR = \{a_1, a_2, \dots, a_n\}$ 
    find customer c, not belonging to SR, that is geographically close to  $a_1$ 
    Insert SR into the genetic material of  $I_1$  in such a way, that  $a_1$  is placed
      immediately after c
    From the original genetic material of  $I_1$  remove all duplicated customers
      that also appear on SR, obtaining a descendant D
  end for
end

```

Abbildung 37: Prinzip des Rekombinations-Teilverfahrens nach [Tavares 2004], [Tavares, Pereira et al 2003] und [Pereira, Tavares et al 2002]

Variablen & Funktionen:

I : Individuum der Elterngeneration = Lösung des VRP-TW

S : Menge von Lösungen; $S = \{I_1, I_2, \dots, I_m\}$

SR: „sub route“ Teilstück einer Tour

a_i : Stop in einem Teilstück einer Tour

c: Kunde, der dem aktuellen a_1 am nächsten liegt

D: „Nachkomme“ = Individuum der Kindergeneration

So erzeugte Nachkommen D können außerdem mutiert werden. Dazu kommen die Mutationsverfahren:

- ▶ Swap (Austausch zweier Stops)
- ▶ Inversion (Umdrehen der Tourreihenfolge)
- ▶ Insertion (Entfernen eines Stops und einfügen an anderer Stelle)
- ▶ Displacement (Entfernen einer Teilsequenz einer Tour und Einfügen an anderer Stelle in selber oder anderer Tour)

Die beiden letzten Mutationsverfahren können die Anzahl Touren jeweils um 1 erhöhen oder verringern. Das übergeordnete Planungsverfahren EC lässt sich im Pseudocode folgendermaßen darstellen:

```

proc EC
begin
  create initial set of solutions S;  $|S|=N$ 
  for  $i := 1$  to max. iterations do
    create N descendants by random application of a genetic
      operator  $\in \{\text{crossover, swap, inversion, insertion, displacement}\}$ 
      with predefined selection probability per genetic operator
  end for
  return best found solution
end

```

Abbildung 38: Grobes Prinzip des EC-Verfahrens nach [Tavares, Pereira et al 2003] Kapitel 3.2. – eigene Darstellung

[Tavares, Pereira et al 2003] geben keinen Hinweis darauf, wie sie die Lösungsgüte beurteilen (Anzahl Fahrzeuge vs. Gesamttourkilometer). Aus den von ihnen veröffentlichten Ergebnissen lässt sich jedoch ablesen, dass ihr Verfahren der Gesamttourkilometerreduktion mehr Gewicht beimisst als der Reduktion der benötigten Anzahl an Fahrzeugen.

Erwähnenswert ist außerdem, dass [Tavares, Pereira et al 2003] darauf hinweisen, dass dieser „straightforward approach“ des EC-GVR i.d.R. zu früh konvergiert, und die Lösungsgüte für das VRSP-TW aus ihrer Sicht nicht befriedigend ist.

4.2.2.8 Das Verfahren von Bent & van Hentenryck 2004

[Bent & van Hentenryck 2004] setzen eine Kombination⁷³ eines „simulated annealing“ Verfahren (SA) zur Minimierung der Fahrzeuganzahl (vorrangiges Optimierungsziel) und eines „large neighbourhood search“ Verfahrens (LNS) zur Tourwegminimierung (nachrangiges Optimierungsziel) ein.

Die durch das SA sukzessive eingeschränkte Zufalls-Suche wird durch eine Zielfunktion $e(\sigma)$ ⁷⁴ bewertet, die

- ▶ die Anzahl benötigter Fahrzeuge minimiert,
- ▶ statt gleichausgelasteter Touren tendenziell eher maximal volle Touren anstrebt, auch wenn einzelne kleine „Resttouren“ entstehen, die dann ggf. später völlig aufgelöst werden und
- ▶ solche „Resttouren“ bevorzugt, deren Stops mit minimalen Zeitfensterverletzungen⁷⁵ in andere Touren verschoben werden könnten.

Die möglichen zufälligen Schritte der Zufalls-Suche werden dabei durch folgende Operatoren \circ generiert:

- ▶ **„two exchange“** (Austausch zweier Stops in einer Tour und Invertierung der Zwischensequenz⁷⁶)
- ▶ **„or exchange“** (Entfernen einer Stopsequenz und Einfügen in eine beliebige Tour),
- ▶ **„relocation“** (Verschieben eines Stops innerhalb seiner Tour oder in eine andere Tour),
- ▶ **„exchange“** (Vertauschen zweier Stops) und
- ▶ **„crossover“** (Vertauschen der Nachfolger zweier Stops).

⁷³ Two-Stage Hybrid Local Search

⁷⁴ σ := gültige Lösung; $E(\sigma)$:= Zielfunktionswert der Lösung

⁷⁵ Diese minimalen Zeitfensterverletzungen sind ein grober Schätzer dafür, ob die Resttour aufgelöst werden kann, oder nicht.

⁷⁶ dies ist der selbe Schritt der auch beim 2-opt Verfahren durchgeführt wird.

Das SA-Verfahren selbst lässt sich im Pseudocode folgendermaßen darstellen.

```

proc route minimize
begin
  generate initial solution  $\sigma^\circ$  (initiale Ausgangslösung)
   $\sigma^b := \sigma^\circ$  (derzeit beste Lösung)
  while (time < time limit) do
     $\sigma := \sigma^b$  (aktuelle Lösung)
    t := starting temperature
    while (time < time limit) and (temperature > temperature limit) do
      for (i:=1; i ≤ maxIterations; i++) do
        o := random (operators)
        c := random (customers)
        generate  $\langle \sigma_1, \dots, \sigma_s \rangle = N(o, c, \sigma)$ 
        if  $e(\sigma_1) < e(\sigma^b)$  then  $\sigma^b := \sigma_1; \sigma := \sigma_1$ 
        else
           $\sigma := \text{randomselect}(\langle \sigma_1, \dots, \sigma_s \rangle, \sigma, t)$ 
        end for
        t :=  $\alpha \times t$ 
      end while
    end while
end

```

Abbildung 39: Prinzip des SA „route minimize“ von [Bent & van Hentenryck 2004] (eigene Darstellung)

Variablen & Funktionen:

$e(\sigma)$ = Zielfunktionswert der Lösung σ

o : ausgewählter Operator

c : ausgewählter Stop

$\langle \sigma_1, \dots, \sigma_s \rangle = N(o, c, \sigma)$: nach $e(\sigma)$ sortierte Nachbarschaft von σ , die durch Anwendung des Operators o auf den Stop c erreicht werden kann; d.h. $e(\sigma_i) < e(\sigma_j)$ gdw. $i < j$

$\text{randomselect}(\langle \sigma_1, \dots, \sigma_s \rangle, \sigma, t)$: selektiert in Abhängigkeit von t und der Güte von $e(\sigma_i)$ einen neuen Vektor σ nach der Vorschrift:

$r := \lfloor \text{random}([0;1])^\beta \times s \rfloor$

$\Delta := e(\sigma) - e(\sigma_r)$

if $\Delta \geq 0$ **then** $\sigma := \sigma_r$

else { **if** $\text{random}([0;1]) \leq e^{-\Delta/t}$ **then** $\sigma := \sigma_r$ }

d.h. falls das aktuelle σ durch das zufällig ausgewählte σ_r verbessert werden kann, wird σ_r zur neuen aktuellen Lösung σ . Falls nicht, hängt es von der aktuellen Temperatur ab, ob σ aktualisiert wird.

Im Anschluss an diese Phase, die die Anzahl Fahrzeuge minimiert, wird dann ein LNS-Verfahren zur Tourkilometerminimierung eingesetzt.

4.2.2.9 Das Verfahren von Lim & Zhang 2005

[Lim & Zhang 2005] schlagen ein zweistufiges Verfahren vor, das sich sowohl für klassisch lösbare VRSP-TW eignet (alle Kunden werden bedient), als auch für m-VRSP-TWs, bei denen nicht alle Kunden bedient werden. Sie nutzen das Konzept

eines „ejection pools“ (EP), der temporär nicht bediente Stops beinhaltet (vgl. „holding list“-Konzept von [Lau et al 2003]).

Als Eröffnungsverfahren nutzen sie eine Construct/Analyse/Prioritize-Heuristik, die pro Iteration ein Insertions-Verfahren aufruft (construct), die jeweilige Lösung untersucht (analyse) und dann neue Prioritäten für die Insertions-Reihenfolge des nächsten Iterationsschritts definiert. Die neue Einfüge-Priorität eines Stops ist umgekehrt proportional zur Auslastung der Tour, in der der Stop in der Vorgängeriteration verplant wurde.

Auf der so gefundenen Eröffnungslösung setzt dann das eigentliche Verfahren („procedure maximize customers“) auf, dass entscheidet, welche Kunden bedient werden.

```
proc maximize customers
begin
  delete extra routes with fewest customers from routes
  add the customers from deleted routes into the EP
  while (EP is not empty) and (not to terminate) do
    select a customer u from the EP
    determine the insertion position of u in routes
    remove u from EP and insert it into the target route
    while (target route is not feasible) do
      remove a customer from the route and add it to the EP
    end while
    apply exchange, relocate, 2-opt, or-opt (to single routes) and
      2-opt*, cross to routes (to pairs of routes)
  end while
end
```

Abbildung 40: Prinzip der Prozedur „maximize customers“ nach [Lim & Zhang 2005]

Variablen & Funktionen:

EP: Liste von Stops, die keiner Tour zugeordnet sind

routes: Menge aller Touren

Die „local search“ Verbesserungsschritte am Ende der Prozedur haben zum Ziel die gewichtete Summe der „re-insertion costs“ für die Stops in EP zu minimieren und die „maximum free time“ einer Tour (bei „single route moves“) bzw. eines Tourenpaars (bei „pair of route moves“) zu maximieren.

In einem nachgeschalteten Schritt werden dann die Gesamttourkilometer minimiert („procedure reduce distances“).

4.2.2.10 Zusammenfassung

Ein Grossteil der oben vorgestellten Verfahren⁷⁷ hat mindestens eine „beste bekannte heuristische Lösung“ für eine der Solomoninstanzen beigesteuert⁷⁸. Es kann daher davon ausgegangen werden, dass sie (richtig parametrisiert) recht gute Lösungen für das VRSP-TW erzeugen können.

Warum wurde den Verfahren zum VRSP-TW jedoch so viel Raum an dieser Stelle eingeräumt?

Die Ergebnisse dieser Verfahren werden für ausgewählte Standard-Probleme in Kapitel 6 miteinander verglichen. Um diese Vergleiche richtig interpretieren zu können, ist ein Grundverständnis über die zugrundeliegenden Verfahren notwendig. Daher wurden zumindest die Grundzüge der Verfahren an dieser Stelle aufgezeigt. Für darüber hinausgehende Details sei auf die jeweils genannten Primärquellen verwiesen.

4.2.3 Verfahren für das VRSP-UD

Im Gegensatz zum deterministischen VRSP sind beim stochastischen VRSP eine oder mehrere Größen nichtdeterministisch.

Im Folgenden werden einige Verfahren für das VRSP-UD (uncertain⁷⁹ demand) aufgezeigt. Dabei sind die wesentlichen Ziele der jeweiligen Verfahren je nach Problemklasse sehr unterschiedlich:

Bei den VRSP-UD stehen meist wegen der begrenzten Fahrzeugkapazität (z.B. Fahrzeug ist „voll“ bei Warensammlungen) die **Minimierung der zusätzlichen Depotfahrten** im Fokus der Optimierung, da jeweils dann eine zusätzliche Depotfahrt nötig ist, wenn die Fahrzeugkapazitätsgrenze ungeplant überschritten wird.

Dieses Problem tritt typischerweise bei Abholverkehren im Teilladungstransport auf. Hierbei werden i.d.R. Stammkunden bedient, von denen aber ex ante nicht klar ist, wie groß die Mengen sind, die sie versenden möchten.

Das Kernproblem dieses stochastischen VRSP-Typs ist, dass wegen der unbekanntem Sammelmengen ex ante nicht klar ist, wann ein Fahrzeug voll wird, und daher zur Entladung zum Depot zurückkehren muss. Es ist daher anzustreben, dass das Fahrzeug um so dichter am Depot ist, je wahrscheinlicher es ist, dass es voll wird.

⁷⁷ [Rochat & Taillard 1995], [Taillard et al 1997], [Shaw 1998], [Homberger & Gehring 1999], [Ibaraki et al 2001], [Bent & van Hentenryck 2004]

⁷⁸ s. [Solomon 2003]

⁷⁹ einige Autoren verwenden den Begriff „stochastic“ statt „uncertain“

4.2.3.1 Das exakte Verfahren von Gendreau et al

Im Gegensatz zum reinen VRSP-UD kann es beim VRSP-UDC auch vorkommen, dass ein Kunde an einem Tag nichts zu versenden hat, und er dies erst sehr kurzfristig mitteilt. [Gendreau et al; 1995] und [Gendreau et al; 1996] gehen daher davon aus, dass zumindest bei Erreichen des Vorgängerkunden bekannt ist, dass ein Kunde nicht bedient werden muss, und dieser Kunde aus der Tour entfallen kann.

[Gendreau et al; 1995] beschreiben ein exaktes Verfahren für dieses Problem, welches sie als modifizierten „branch and bound solver for mixed integer programs“ beschreiben (Land-Powell BB program).

Sie zeigen, dass die Rechenzeit dieses Verfahrens insbesondere bei einer Vielzahl von stochastische Kunden (d.h. Kunden, die ggf. gar nicht bedient werden) sehr schnell wächst:

n	Rechenzeit [sec.]	Rechenzeit [sec.]
	Exaktes Verfahren	Heuristik
6	0.28	2.26
7	2.64	2.75
8	28.22	4.72
9	308.76	6.52
10	4305.75	8.27
11	53903.64	15.46

Tabelle 5: Rechenaufwandsübersicht aus [Gendreau et al; 1996; Table 1]

n = Anzahl Kunden (incl. Depot); Angestrebter Fahrzeugbelegungskoeffizient: $f=0.75$

Es eignet sich daher für kleine Problemgrößen und zur Kalibrierung von Heuristiken. Bei größeren Problemgrößen schlagen [Gendreau et al; 1996] daher eine Heuristik nach dem Tabu-Search-Verfahren vor (s.u.).

4.2.3.2 Die Tabu-Search Heuristik von Gendreau et al

Für eine spezifische Ausprägung des VRSP-UDC bzw. VRP-SDC⁸⁰, nämlich die Abholung ungewisser Mengen bei einer **bekannt**en Anzahl Abholpunkte entwickelten [Gendreau et al; 1996] das auf diesen Fall zugeschnittene Verfahren TABUSTOCH.

⁸⁰ Die von Gendreau verwendete Notation VRP-SDC (stochastic customers) läßt vermuten, dass es sich um zufällige Kundenlokationen handeln könnte. In Wirklichkeit ergibt sich die Stochastizität der Kunden in diesem Fall jedoch lediglich dadurch, dass die Abholmenge auch Null sein kann, und in diesem Fall eine Anfahrt entfallen kann. Somit ist hier nicht der allgemeine Fall von stochastischen Anfahrtstellen, sondern nur der einer stochastischen Auswahl aus einer vordefinierten Menge von Anfahrtstellen abgedeckt.

Gendreau et al haben das **Domänenwissen für diesen Spezialfall** exzessiv genutzt, und in die Zielfunktion, die strafkostenbehafteten erweiterten Zielfunktion, die Nachbarschaftsdefinition, die „tabu moves“ und die „tabu move“-Bewertung eingebaut (die Zielfunktion berücksichtigt z.B. die zu erwartende Depotrückfahrt bei Erreichung der Fahrzeugkapazitätsgrenze). Es ist daher davon auszugehen, dass das Verfahren auf strukturell andere Aufgabenstellungen (wie das in dieser Arbeit untersuchte) nicht ohne weiteres übertragbar ist.

4.2.3.3 Das Verfahren von Erera

Zur selben Aufgabenstellung VRSP-UD hat [Erera 2000] ein anderes heuristisches Verfahren entwickelt.⁸¹:

- ▶ Gegeben: Depot, Fuhrpark mit räumlicher Kapazität, Kundennachfrage als Zufallsgröße mit bekannter Verteilung, Punkt zu Punkt Transportkosten
- ▶ Gesucht: Operative Strategie zur Minimierung der zu erwartenden Kosten
- ▶ u.d.N.:
 - ▶ alle Kunden werden bedient
 - ▶ keine Fahrzeugkapazitätsgrenzen werden verletzt

Außerdem ist die definitive Nachfrage einzelner Kunden zum Planungszeitpunkt unbekannt (uncertain).

Sie schlagen für dieses - vor allem von Kapazitätsbeschränkungen abhängige - Problem ein Regionen-Verfahren vor, das zwischen einer Nahregion und einer Fernregion unterscheidet. In der Fernregion werden die Touren vorgeplant, in der Nahregion werden die Fahrzeuge dynamisch in Abhängigkeit von der aktuellen Fahrzeugbeladung den jeweiligen Tortenstück-Sektoren zugewiesen.

⁸¹ [Erera & Daganzo, 2000] klassifizieren den von ihnen untersuchten Anwendungsfall als Problemtyp $VRP_{sc}(UD)$
 $VRP_{sc}(UD)$:= Vehicle Routing Problem – space capacity constraint – with uncertain demand

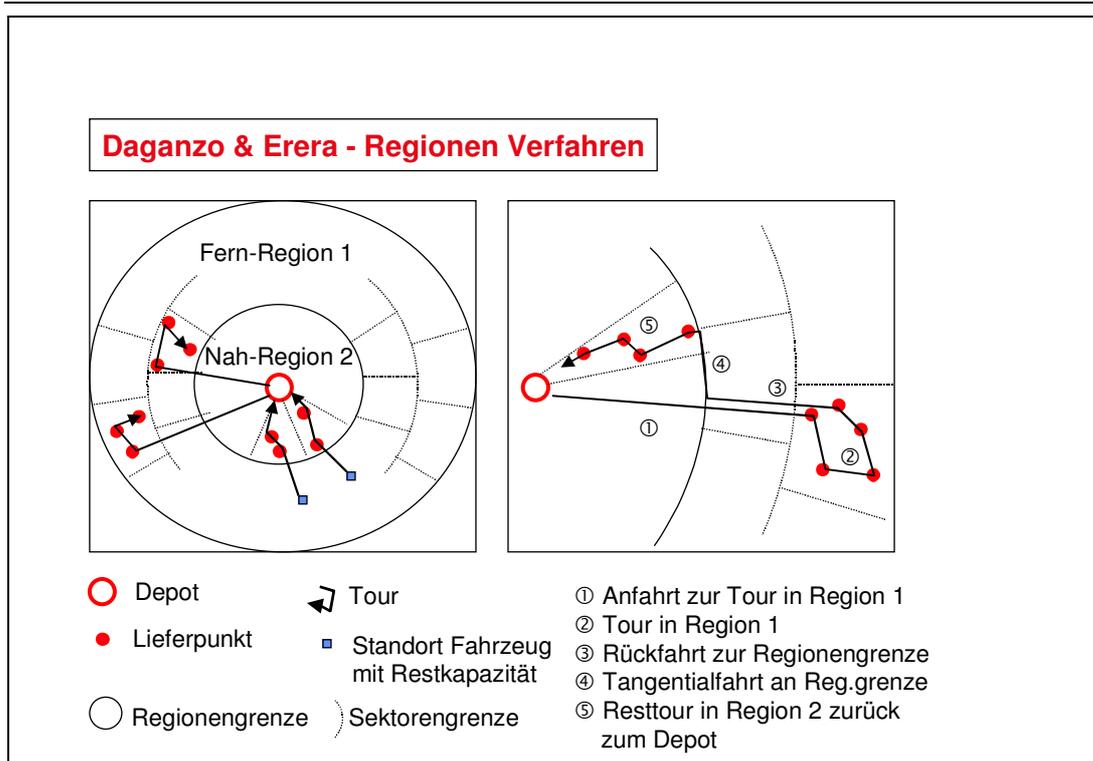


Abbildung 41: Regionen-Verfahren nach Daganzo & Erera

Ziel dieses Verfahrens ist es, die zu erwartenden vorzeitigen Rückfahrten zum Depot, die sich ergeben, wenn die auf dem Fahrzeug geladene Ware „abverkauft“ wurde zu minimieren, und die Fahrzeuge mit entsprechenden Restkapazitäten (= noch geladene Ware) nach Bedienung der Außenregion 1 möglichst geschickt den Sektoren der Innenregion 2 zuzuweisen.

Durch Variation des Abstands r vom Depot zur Regionengrenze ergeben sich unterschiedliche Ergebnisse. Durch Simulation wird der Abstand r ermittelt, der die besten Ergebnisse liefert.

Das Verfahren umfasst neben dem beschriebenen Verfahrensschritten auch Verfahrensschritte für die Fälle, dass in den Regionen 1 oder 2 Kunden übrig bleiben („overflows“):

- ▶ Restkunden in der Fern-Region 1 werden durch Fahrzeuge an der Regionengrenze bedient
- ▶ Restkunden in der Nah-Region 2 werden durch zusätzliche „Besentouren“ bzw. „sweep tours“ bedient.

Um zu überprüfen, ob dieses Verfahren auch für den Einsatz im Kontext des D-m-VRSP-TW-UC-DO einsetzbar ist, wird die Ausprägung des von [Erera 2000] gelösten Problemtyps VRSP-UD und die Ausprägung des in dieser Arbeit untersuchten Problemtyps D-VRSP-TW-UC gegenübergestellt.

Aspekt	Ausprägung des VRSP-UD [Erera 2000]	Ausprägung des D-VRSP-TW-UC [diese Arbeit]
Erfüllungsorte	Abholorte	Lieferorte
Stochastizität der Kunden (SC)	Aus n bekannten Erfüllungsorten müssen $m < n$ Erfüllungsorte angefahren werden	Weder die Anzahl noch die Lage der Erfüllungsorte ist a priori bekannt ⁸²
Relevanz der Kapazitätsbeschränkung (C) der Fahrzeuge und der Nachfrage (D)	C ist der wesentliche Einflussfaktor des Problems, da bei Überschreitung von C eine zusätzliche Depotanfahrt anfällt	C kann als Einflussfaktor des Problems vernachlässigt werden ⁸³ ($D \ll C$)
Relevanz der Zeitfensterbreite (TW)	TW muss i.d.R. so groß sein, dass eine zwischen-geschaltete Depotanfahrt im Prinzip möglich ist (sonst würde das Kernverfahren von Erera nicht greifen, und es müssten ausschließlich „sweep tours“ gefahren werden).	TW kann beliebig klein werden.

Tabelle 6: Vergleich der untersuchten Ausprägungen der Problemtypen

Wegen dieser grundsätzlichen Unterschiede wird das Verfahren nicht weiter betrachtet.

4.3 Das D-VRSP

Auch das D-VRSP umfasst eine Fülle von Ausprägungen (Unterklassen). In diesem Abschnitt werden daher die nicht betrachteten Ausprägungen des D-VRSP kurz abgegrenzt, und dann auf ein neueres Verfahren zum D-VRSP-UC näher eingegangen.

⁸² Formell ließe sich dieser Unterschied auflösen, wenn man beispielsweise alle postalisch bekannten Adressen im Liefergebiet als mögliche Erfüllungsorte n^* interpretiert, und die zu beliefernden Adressen als m^* . Für diese Interpretation würde dann jedoch gelten:
 $m/n \gg m^*/n^*$.

⁸³ Bei engen Zeitfenstern dominieren diese die Fahrzeugkapazität als tourbegrenzenden Faktor (Touren sind kapazitiv nicht voll ausgelastet).

4.3.1 Abgrenzung nicht betrachteter D-VRSPs

Die Dynamik von D-VRSPs kann sich nach [Larsen 2001, S.15ff] auf verschiedene Anwendungsfälle beziehen, die jeweils unterschiedliche Grade an Dynamik (Anteil der Aufträge, die ex ante nicht bekannt sind, sondern erst sukzessive bekannt werden) aufweisen:

- ▶ **Travelling Repairman** – Zuordnung von Serviceaufträgen zu Servicetechnikern
- ▶ **Courier Mail Services** – Abholung von Eilsendungen bei Bedarf
- ▶ **Distribution of Heating Oil** – Heizöldistribution an private Haushalte
- ▶ **Dynamic Dial a Ride Systems** – Personentransporte nach Bedarf (i.d.R. mit signifikanter Vorankündigungsfrist)
- ▶ **Taxi Cab Services** - Personentransporte nach Bedarf (i.d.R. ohne signifikante Vorankündigungsfrist)
- ▶ **Emergency Services** – Einsatzplanung von Polizei, Feuerwehr, Ambulanz (meist ein reines „scheduling problem“, da i.d.R. nur ein Auftrag pro Einsatz auszuführen ist)

Diese Anwendungsfälle werden in dieser Arbeit nicht betrachtet, da ihr Problemcharakter nicht mit dem des Lebensmittelheimlieferservice übereinstimmt.

4.3.2 Verfahren für das D-VRSP-UC

Bei den D-VRSP-UC geht es nach [Bent & van Hentenryck 2003, Kapitel 2] vor allem darum, durch geschicktes Auswählen bzw. Ablehnen von Aufträgen den Erwartungswert der **Anzahl der Lieferungen** / Abholungen zu **maximieren**⁸⁴.

Erst in jüngster Zeit haben [Bent & van Hentenryck 2003] ein Verfahren zur Lösung des D-VRSP-UC veröffentlicht. Sie nennen es MSA (Multi Scenario Approach).

Im Gegensatz zu klassischen Greedy-Heuristiken, die zur Lösung von nicht-stochastischen dynamischen VRSP eingesetzt werden und nur auf Basis der aktuell bekannten Kundenaufträgen Ergebnisse generieren, versucht der MSA Informationen über die noch zu erwartenden Kundenaufträge in der Lösungsfindung zu berücksichtigen.

Die Kernidee des MSA ist es, neue Tourenpläne für Szenarien zu generieren, die jeweils sowohl die bereits bekannten, als auch „künftige“ Aufträge enthalten. „Künftige“ Aufträge werden einfach zufällig entsprechend der jeweiligen

⁸⁴ Es wäre auch denkbar, statt der Anzahl der Kunden den Erwartungswert der jeweiligen Erträge der Lieferungen zu maximieren. Spätestens in diesem Fall liegt das Problem in der Klasse der Multiple-Vehicle-Orienteeing bzw. Team Orienteering Probleme (TOP).

Verteilungsfunktion generiert. So werden mehrere mögliche Auftrags-Szenarien erzeugt, für die jeweils ein entsprechender Tourenplan σ generiert wird. Aus diesem Tourenplan σ werden nun alle „künftigen“ Aufträge wieder entfernt, und nur die bekannten Aufträge verbleiben in den Touren. Dieser reduzierte Tourenplan sei σ^- , und hat somit definitionsgemäß noch Raum für die de facto unbekannt künftigen Aufträge. Für alle Auftrags-szenarien werden die jeweiligen reduzierten Tourenpläne σ^- in der Menge S_t zusammengefasst.

Anschließend wird durch eine Konsensfunktion („consensus function“) derjenige Plan σ^* aus der Menge S_t ausgewählt, der allen anderen Plänen am ähnlichsten ist. Dieses Vorgehen entspricht einer „least commitment“ Strategie (d.h. wähle den Plan aus, der sich am wenigsten festlegt).

So gibt es zu jedem Zeitpunkt t jeweils eine Menge mit möglichen Plänen S_t und einen ausgewählten Plan σ^* . Aufbauend auf diesem Grundprinzip steuert eine Metasteuerung 4 event-gesteuerte Prozesse:

- ▶ **Timeout:** Ein Plan σ^- wird zum Zeitpunkt t ungültig, wenn t die letztmögliche Abfahrt eines Fahrzeugs zur Erreichung eines Kunden r ist. Ein ungültiger Plan σ^- wird aus S_t entfernt.
- ▶ **Plan-Erzeugung:** Wird zu einem Zeitpunkt t ein neuer Plan σ^- generiert, so wird er zu S_t hinzugefügt, und σ^* wird neu berechnet.
- ▶ **Kundenanfrage:** Gelangt ein neuer Auftrag r ins System, so ermittelt MSA, in welchen der Pläne σ^- dieser Auftrag r eingefügt werden kann. Kann dieser Auftrag in keinen der Pläne eingefügt werden, so wird dieser Auftrag abgelehnt. Ansonsten ergibt sich S_t neu als die Menge der σ^- , in die r eingefügt werden kann.
- ▶ **Fahrzeugabfahrt:** Wenn σ^* festlegt, dass ein Fahrzeug v zur Zeit t vom Kunden r abfahren muss, so werden alle σ^- aus S_t entfernt, die mit dieser Abfahrt inkompatibel sind.⁸⁵

Mit fortschreitendem t werden die beschriebene Prozesse sukzessiv angestoßen. Die Touren werden nach dem jeweils zur Zeit t gültigen Plan σ^* durchgeführt.

Dieses Verfahren scheint prinzipiell auch für den in dieser Arbeit untersuchten Fall des D-m-VRSP-TW-UC-DO geeignet zu sein. Es setzt aber zum einen voraus, dass verlässliche stochastische Informationen über die zu erwartenden Kundenanfragen verfügbar sind. Zum anderen nutzt das Verfahren nicht den Freiheitsgrad, die Lieferzeitzusagen für die Kunden zeitlich so zu steuern, dass eine gute Fahrzeugauslastung erzielt werden kann, es entscheidet binär über Annahme oder Ablehnung eines Kunden.

⁸⁵ Muß das Fahrzeug spezifische Waren für die jeweiligen Kunden mitführen, so ist die Fahrzeugabfahrt am Depot (ggf. abzüglich einer definierten Rüstzeit) ausschlaggebend für diesen Prozessschritt.

4.4 Das OP

Ursprünglich resultiert das Orienteering-Problem (OP) aus der Aufgabenstellung des Orientierungslaufs (daher auch die Namensgebung). Dort ist es Aufgabe der Teilnehmer (meist in einem Waldgebiet ausgerüstet mit Karte und Kompass) ausgehend von einem Ausgangspunkt eine Teilmenge von Kontrollpunkten zu besuchen, deren jeweiliger Besuch dem Team Bonuspunkte bringt, und anschließend den Zielort aufzusuchen. Dabei gibt es ein Zeitlimit, in der das Ziel erreicht werden muss.

Das Team, das am meisten Bonuspunkte sammelt, gewinnt. Teams, die das Zeitlimit überschreiten, werden disqualifiziert.

Dieses Problem lässt sich als Generalisierung des TSP (GTSP) formulieren und ist somit NP-schwierig (s. [Golden et al., 1987]):

$$ZF: \max \sum_{i \in B(j)} s(i) \quad u.d.N. \quad t(j) \leq t_{max}$$

Variablen & Funktionen:

$i = 1..n$ Knoten in der euklidischen Ebene ($i=1$: Startpunkt; $i=n$: Zielpunkt)

$s(i) \geq 0$ ($\forall i = 2, \dots, n-1$) & $s(1) = s(n) = 0$: der „Score“ bzw. Ertrag des Knotens i
(im obigen Fall z.B. die Anzahl der jeweiligen Bonuspunkte)

$B(j)$:= Menge der in einer Tour j besuchten Knoten

$t(j)$:= minimale Dauer der Tour j durch die jeweiligen Punkte $B(j)$
mit Start $i=1$ und Ende $i=n$

t_{max} := maximal zulässige Tourdauer

In dieser klassischen Form entspricht das Orienteering-Problem einer Aufgabenstellung, in der 1 Fahrzeug aus $n-2$ Aufträgen diejenigen innerhalb einer maximalen Tourdauer t_{max} bedienen soll, deren gesamter Ertrag maximal ist.

4.4.1 Verfahren für das OP

4.4.1.1 Stochastischer Ansatz von Tsiligirides

[Tsiligirides, 1984] schlägt zum einen einen stochastischen Ansatz vor, bei dem basierend auf einer Monte Carlo Methode erst eine große Anzahl von möglichen Touren generiert wird. Die Tourbildung erfolgt dabei nach folgendem Muster:

-
- ▶ Ein Eröffnungsverfahren bildet eine Initialtour (z.B. die Trivial-Tour [1,n])
 - ▶ Für jede Station der bestehenden Tour wird ein Begehrlichkeits-Maß $A(j)$ (desirability) für all diejenigen Stationen j gebildet, die der Tour noch nicht angehören:
 $A(j) = (s(j) / t(\text{last},j))^{4.0}$
mit $t(\text{last},j) :=$ Fahrzeit vom gewählten Vorgänger-Knoten „last“ zum Knoten j
 - ▶ Aus den 4 größten Werten von $A(j)$ wird *zufällig* einer ausgewählt, und in die Tour eingefügt (wenn die maximale Fahrzeit nicht überschritten wird).
 - ▶ So wird sukzessive weiter verfahren, bis keine weiteren Knoten mehr in eine Tour eingefügt werden können.

So wird eine Vielzahl von Touren basierend auf einem Zufallsverfahren generiert aus denen schließlich die jeweils beste Tour ausgewählt wird.

4.4.1.2 Deterministischer Ansatz von Tsiligirides

[Tsiligirides, 1984] schlägt zum anderen einen deterministischen Ansatz vor, bei dem das Tourgebiet in Sektoren geteilt wird, die auf konzentrischen Kreisen basieren. Touren werden innerhalb dieser Sektoren gebildet um die Fahrzeit zu minimieren. Durch Variation der Radien und durch Rotation der Achsen werden verschiedene Ergebnisse für jedes t_{\max} generiert.

4.4.1.3 Center of gravity heuristic von Golden

[Golden et al., 1987] schlägt einen 3 stufigen Tourschwerpunkts-Ansatz („Center of gravity“) vor.

- ▶ Im ersten Schritt wird durch eine Einfüge-Heuristik eine Tour erzeugt, die
 - ▶ in $i=1$ beginnt und in $i=n$ endet,
 - ▶ einen relativ hohen Ertrag hat, und
 - ▶ kürzer als t_{\max} ist.
- ▶ Im zweiten (Verbesserungs-) Schritt wird ein Reihenfolgeoptimierungsverfahren (z.B. 2-opt) eingesetzt, um die Tourdauer zu verkürzen. Die entstehende Tour sei L .
- ▶ Im dritten („center of gravity“-) Schritt wird
 - ▶ (1) zuerst der nach scores s gewichtete geografische Schwerpunkt g der Tour L bestimmt:

$$g = (\bar{x}, \bar{y}) \quad \text{mit} \quad \bar{x} = \frac{\sum_{i \in L} s(i)x(i)}{\sum_{i \in L} s(i)}; \quad \bar{y} = \frac{\sum_{i \in L} s(i)y(i)}{\sum_{i \in L} s(i)}$$

- ▶ (2) Anschließend werden für alle $i \notin L$ die Abstände $a(i)$ zu diesem geografischen Tourschwerpunkt g berechnet
- ▶ (3) Dann wird eine absteigend nach den Verhältnissen $s(i)/a(i)$ sortierte Liste der Knoten $i \notin L$ generiert, aus der sukzessive die Knoten i in die Tour L eingefügt werden (u.d.N., dass jeweils t_{\max} nicht überschritten wird)
- ▶ (4) Anschließend wird die entstehende Tour reihenfolgeoptimiert. Die entstehende Tour sei die Tour L_1' .
- ▶ (5) Für L_1' wird nun der neue Tourschwerpunkt $g(L_1')$ berechnet.

Der „center of gravity“-Schritt wird nun jeweils neu von (1) bis (5) mit Tour L aber mit $g := g(L_{i+1}')$ durchlaufen. Die resultierenden Tour seien L_2', L_3' u.s.w.

Sobald für einige $q > p$: $L_p' = L_q'$ (d.h. keine Veränderungen mehr), terminiert das Verfahren.

Variablen & Funktionen:

- L : Tour
- s : score/Ertrag;
- $s(i)$: score des Stops i
- g : nach scores gewichteter geografischer Schwerpunkt einer Tour
- $a(i)$: Abstand des Stops i von g
- t_{\max} : Maximale Tourdauer von L

[Golden et al., 1987] konnte für einige Beispiele zeigen, dass die „center of gravity“ Heuristik sowohl dem stochastischen als auch dem deterministischen Ansatz von Tsiligirides meist überlegen ist.

4.4.1.4 Weitere Verfahren für das Orienteering Problem

Weitere Verfahren für Varianten des Orienteering Problems werden u.a. genannt in [Chao, Golden et al, 1996a] - ein bzgl. Rechenzeitaufwand relativ schnelles heuristisches Verfahren zur Lösung des OP) und [Fomin et al., 2002] - Näherungsansätze für das dynamische Orienteering Problem (time depending Orienteering TDO) mit einem Fahrzeug, unter Berücksichtigung von tageszeitabhängigen Entfernungen zwischen den einzelnen Anfahrtstellen.

4.4.2 Verfahren für das TOP

Das klassische Orienteering Problem berücksichtigt nur die Auswahl von m aus n ($m \leq n$) Besuchspunkten (Knoten) für max. *eine* Tour. [Chao, Golden et al, 1996b], definieren eine Verallgemeinerung des OP zum team orienteering problem (TOP)

bei dem mehr als eine Tour die möglichen Besuchspunkte bedienen kann. Das TOP verhält sich somit zum OP wie das VRSP zum TSP.

[Chao, Golden et al, 1996b] schlagen eine Heuristik zur Lösung dieses Problems vor die im Wesentlichen aus folgenden Schritten besteht (vereinfachte Darstellung):

- ▶ Der geografische Suchraum der Lösung wird beschränkt auf die N Punkte innerhalb einer Erreichbarkeitsellipse, deren Foci jeweils durch den Start- und Endpunkt der Touren, und deren Großradius durch die maximale Tourdauer T_{max} determiniert wird.
- ▶ In einem Initialisierungsschritt werden M initiale Touren folgendermaßen gebildet:
Aus der Menge von L Punkten ($M < L \leq N$) innerhalb der Erreichbarkeitsellipse mit größtem Abstand zu den Foci wird jeweils ein Punkt ausgewählt, mit dem Start- und Endpunkt verbunden, und mittels einer greedy-Heuristik solange mit weiteren Tourpunkten aufgefüllt, bis die maximale Tourdauer T_{max} erreicht ist u.s.w.
Bei definierten L und M gibt es also „ L über M “ *mögliche Ausgangslösungen*. Aus diesen wird diejenige Lösung (= Tourenplan mit M Touren =: $paths_{top}$) als *Initiallösung* ausgewählt, die den höchsten Score erzielt.
Alle weiteren Punkte werden anschließend auch zu Touren verplant – die Menge dieser Touren sei $paths_{ntop}$
- ▶ Die Ausgangslösung wird anschließend sukzessive verbessert durch folgende drei Teilverfahren, die über eine zyklische Metasteuerung (2 geschachtelte FOR-Schleifen) nacheinander angestoßen werden:
 - ▶ Tauschverfahren (two-point-exchange): jeweils ein Punkt wird aus dem Tourenplan $paths_{top}$ entfernt und „gültig“ in den Tourenplan $paths_{ntop}$ eingefügt. Gleichzeitig wird ein Punkt aus dem Tourenplan $paths_{ntop}$ entfernt und „gültig“ in den Tourenplan $paths_{top}$ eingefügt. Steigt so der Score von $paths_{top}$, so wird die aktuelle Lösung aktualisiert.
 - ▶ Einfügeverfahren (one point movement): ein Punkt wird aus einer Tour entfernt und in eine andere Tour eingefügt. Steigt so der Score von $paths_{top}$, so wird die aktuelle Lösung aktualisiert
 - ▶ Reihenfolgeoptimierung über die bestehenden Touren: hier: 2-opt

Variablen & Funktionen:

N : Anzahl der Punkte innerhalb der Erreichbarkeitsellipse

L : Verfahrensparameter – eine definierte Anzahl von Punkten, die von den Foci „weitest entfernt“ sind ($L \leq N$)

M : Verfahrensparameter - Anzahl von betrachteten Touren ($M < L$)

T_{max} : maximale Tourdauer

$paths_{top}$: M Touren mit jeweils höchsten (top) scores

$paths_{ntop}$ restliche Touren mit scores kleiner als der der top M Touren

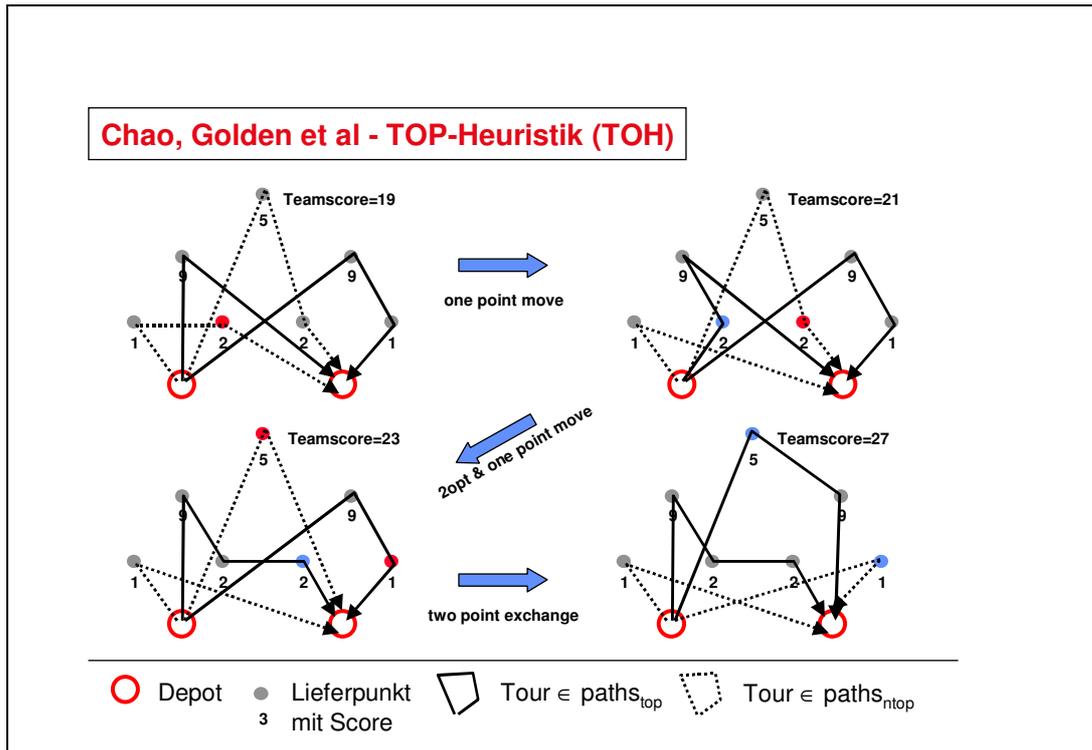


Abbildung 42: TOP-Heuristik (TOH) aus [Chao, Golden et al, 1996b]

Dieses Verfahren ist eine einfache Heuristik zur Lösung eines TOP. Um auch mit engen Zeitfenstern umzugehen, müsste das Verfahren noch signifikant weiterentwickelt werden. Um es für den Anwendungsfall „eCommerce-Lebensmittelheimzustellung“ anzupassen, müsste außerdem die Dynamik der Aufgabenstellung berücksichtigt werden. Es wird daher in dieser Arbeit nicht weiter untersucht.

Das Verfahren von [Chao, Golden et al, 1996b] ist sicher ein wichtiger erster Schritt zur Lösung des TOP. Die Forschung zur Lösung der TOP ist aber leider noch nicht so weit fortgeschritten, wie die exzessive Forschung auf dem Gebiet der VRSP. Es wäre begrüßenswert, wenn die Forschung zum TOP ähnlich intensiv betrieben würde, wie die zum VRSP. Gerade für den Anwendungsfall „eCommerce-Lebensmittelheimzustellung“ wären interessante Impulse aus dieser Forschungsrichtung zu erwarten.

4.5 Wertung der Literatur in Bezug auf Anwendungsfall

Gibt es bereits beschriebene Ansätze für den Anwendungsfall Lebensmittelheimlieferservices **mit Lieferzeitzusage zum Bestellzeitpunkt**?

Meist orientiert sich die OR-Literatur nicht an Anwendungsfällen, sondern an OR-Problemklassen. Greifen wir daher nochmals auf die im vorangegangenen Kapitel definierten Planungsansätze zurück, stellen diese den zugehörigen OR-Planungsproblemtypen gegenüber, und leiten daraus ab, ob der Ansatz für den eCommerce-Lebensmittelheimlieferservice geeignet ist.

Planungsproblem	Planungsansatz	Bemerkung	Geeignet?
VRSP	PAA	Hier werden alle Aufträge gesammelt, Touren ohne Zeitfenster geplant, und dann erst Lieferzeitfenster kommuniziert	nein
VRSP-TW	PPA	Hier ist der Fuhrpark nicht begrenzt (unrealistisch), und es gibt keine Gewähr, dass nicht alle Kunden zum selben Zeitpunkt beliefert werden wollen	nein
m-VRSP-TW	-	Ohne eine Begrenzung der Anzahl Kunden ist die Lösbarkeit des m-VRSP-TW nicht ex ante garantierbar, d.h. der ERG1 ist gefährdet	nein
m-VRSP-STW	REA	Ohne eine Begrenzung der Anzahl Kunden sind Zeitfensterverletzungen nicht eingrenzbar, d.h. der ERG2 ist gefährdet	nein
Filter + VRSP-TW	FPA/PPA	Die Fahrzeuganzahl ist der Freiheitsgrad zur Konfliktlösung. Die richtige Filterkalibrierung ist daher entscheidend	ja
Filter + m-VRSP-STW	FPA/REA	Die Zeitfensterverletzungen sind der Freiheitsgrad zur Konfliktlösung. Die richtige Filterkalibrierung ist daher entscheidend	bedingt
D-m-VRSP-TW-UC-DO	ORA	Durch die Vorschlagsfunktion wird die Kundenflexibilität als zusätzlicher Freiheitsgrad genutzt.	ja

Tabelle 7: Gegenüberstellung der transportlogistischen Ansätze und der zugehörigen OR-Probleme

Die klassischen nichtdynamischen Ansätze zur Lösung des **VRSP-TW** sind in der OR Literatur hinlänglich beschrieben. Wesentliche Erkenntnisse über den FPA/PPA

werden daher im Folgenden auch aus der Literatur entnommen und neu interpretiert.

Zeitfensterverletzungen sind in der Praxis des eCommerce-Lebensmittelheimlieferservice nicht erwünscht (auf wenn sie in Ausnahmefällen in Kauf genommen werden) – die Verfahren zu ***-STW** werden in dieser Arbeit daher nicht näher betrachtet.

Die Literatur zu den **D-VRSP** ist bei weitem nicht so umfangreich wie die zum nichtdynamischen VRSP. Sie bezieht sich i.d.R. meist auf speditionelle „pickup & delivery“-Verkehre (LTL⁸⁶-Verkehre) oder auf die ad-hoc-Einsatzplanung (d.h. das Fahrzeug ist bereits unterwegs). In der neueren Literatur finden sich außerdem einige Hinweise darauf, dass es noch signifikanten Forschungsbedarf für Aufgabenstellungen mit dynamischem Auftragseingang und stochastischen Kunden gibt:

- ▶ [Gendreau et al., 2001] konstatieren, dass im Gegensatz zu den deterministischen CVRP⁸⁷ relativ wenig Aufwand für die Untersuchung von stochastischen CVRP getrieben wird.
- ▶ [Bent & van Hentenryck 2003], weisen darauf hin, dass die wenigen existenten Veröffentlichungen entweder auf statische stochastische oder auf dynamische Probleme (ohne stochastische Informationsauswertung) fokussieren („... very few papers where stochastic information is used to solve dynamic vehicle routing problems ...“).
- ▶ [Ghiani et al. 2004] sehen den Forschungsbereich der real-time VRPs als derzeit unorganisiert („disorganized“) und erkennen signifikante Forschungs-Defizite („major issues have not been addressed, yet“). Sie fordern daher dazu auf, dass Heuristiken zur Tourenbildung mit einer Vorausschau-Möglichkeit („look ahead capability“) entwickelt werden sollten.

Auch die folgenden Übersichten im www über die verschiedenen VRSPs weisen derzeit⁸⁸ noch keine entsprechenden Links auf das D-VRSP-TW für den Anwendungsfall des eCommerce-Lebensmittelheimlieferservice aus.

- ▶ The VRP-Web: www.neo.lcc.uma.es/radi-aeb/WebVRP
- ▶ VRPTW benchmark problems:
<http://w.cba.neu.edu/~msolomon/problems.htm>

Nach Kenntnis des Autors gibt es bisher bis auf [Bent & van Hentenryck 2003] keine Veröffentlichungen bezüglich dynamischer Probleme mit begrenztem Fuhrpark, und engen Zeitfenstern, die für den Anwendungsfall des eCommerce-Lebensmittelheimlieferservice geeignet erscheinen.

⁸⁶ LTL:= less than full truck load

⁸⁷ CVRP = Capacitated VRP

⁸⁸ Stand August 2005

Ein dynamisches Verfahren, das darüber hinaus die Zeitfenster (TW) der Kundenanfragen statt als gegeben zu verwenden durch Vorschlagsfunktionen steuert, und so zusätzlichen Optimierungsspielraum schafft, ist nach Kenntnis des Autors bisher nicht anderswo veröffentlicht. Solch ein neues Verfahren wird in dieser Arbeit vorgestellt, und eröffnet die neue Problemklasse **D-m-VRSP-TW-UC-DO**.

Schließlich ist auch die Literatur zum **TOP** zu berücksichtigen. Eine wesentliche Eigenschaft des klassischen (also nicht dynamischen) TOP ist jedoch, dass zum Zeitpunkt der Entscheidung wer bedient wird und wer nicht, alle möglichen Stops bekannt sein müssen. Verfahren zur Lösung des klassischen TOP sind daher für den Anwendungsfall des eCommerce-Lebensmittelheimlieferservice ungeeignet, denn ein Kunde will bei Bestellung eine verlässliche Bestellzusage haben. Würde das TOP zum D-TOP erweitert, wären die Ansätze für den hier untersuchten Anwendungsfall interessant. Veröffentlichungen zum **D-TOP** gibt es nach Kenntnis des Autors jedoch derzeit nicht⁸⁹.

4.6 Zusammenfassung

In den letzten 20 Jahren seit der Arbeit von [Solomon 1984] sind signifikante Fortschritte im Bereich der Lösung von VRSP-TW gelungen. Aber trotz über 20 Jahren Forschung sind selbst für die im Vergleich zu praxisrelevanten Problemen eher überschaubaren Solomon-Referenzprobleme mit je nur 100 Kunden bis heute nicht alle optimal gelöst. Daher werden verstärkt **Heuristiken** untersucht. Verschiedenste Meta-Heuristiken, die **Domain-Knowhow** inkorporieren sind dabei „allgemeinen“ Lösungsstrategien meist überlegen. Gute Metaheuristiken kommen dabei (zumindest für die bekannten Solomoninstanzen) dicht an die Optimalität heran (s. Kapitel 6 dieser Arbeit).

Aber selbst die aus OR-Sicht als „optimal bzgl. der Tourlänge“ ausgewiesenen Ergebnisse sind aus logistischer Sicht oft unbefriedigend⁹⁰, da diese Optimalität durch den Bedarf an zusätzlichen Fahrzeugen erkauft wird. Das Kerndilemma der Forschung liegt darin, dass „real life“ Probleme einen Kontext haben, aus dem heraus erst die Güte einer Lösung beurteilt werden kann, OR Probleme hingegen **kontextfrei Optimalität** anstreben & ausweisen. Dieses Problem wurde erkannt, und führte in den letzten Jahren zur „Variantenbildung“ der VRSP-Problemklassen.

Diese Varianten standen in den letzten Jahren im Fokus der Forschung. Eine

⁸⁹ Auch in neuesten Veröffentlichungen zum TOP sind keine Literaturverweise auf ein D-TOP zu finden. S. z.B. [Archetti et al 2005] oder [Geem et al 2005]

⁹⁰ Beispiele lt. [Solomon, 2003]:

Instanz R109: km-optimale Lösung:	1146,90 km; benötigt Fahrzeuge: 13
bekannte andere Lösung:	1194,73 km; benötigt Fahrzeuge: 11
Instanz R201: km-optimale Lösung:	1143,20 km; benötigt Fahrzeuge: 8
bekannte andere Lösung:	1252,37 km; benötigt Fahrzeuge: 4

Vielzahl von spezifischen kontextsensitiven Problemen wurde definiert und Lösungen vorgeschlagen. Dabei ist die Kontextsensitivität der Lösungsansätze gut für die Problemlösung des spezifischen Problems, aber schlecht bezüglich der Übertragbarkeit auf andere Probleme.

Der Bereich der D-m-VRSP-TW-UC-DO ist ein neuer Forschungsbereich, der bei Vorliegen der Möglichkeit zur „Nichtbedienung“ ausgewählter Kunden sowohl den D-m-VRSP Problemen als den D-TOP-Problemen zugeordnet werden kann. Es ist zu erwarten, dass es wichtige neue Anstöße geben könnte, falls die Forschung zum TOP sich auch zu dynamisch-stochastischen Aspekten hinwendet.

5 Modellierung & Verfahren

In diesem Kapitel wird ausgehend von den Anforderungen aus der betrieblichen Praxis das Optimierungsproblem des „eCommerce-Lebensmittelheimzustellservice mit Lieferzeitzusage zum Bestellzeitpunkt“ zunächst klassifiziert und dann formal beschrieben.

Anschließend wird ein Verfahren zur Lösung dieses Problems hergeleitet.

5.1 Klassifikation als D-m-VRSP-TW-UC-DO

Aus dem Kontext der Planungsaufgabe wird schnell klar, dass die Lieferanfragen nicht auf einen Schlag, sondern sukzessive eintreffen. Damit ist das Problem ein dynamisches Tourenplanungsproblem (**D-VRSP**).

Der Fuhrpark ist begrenzt, daher handelt es sich um ein D-**m**-VRSP.

Außerdem sind Zeitfenster zu berücksichtigen: D-m-VRSP-**TW**.

Weiterhin sind ex ante die Kunden unbekannt, daher handelt es sich um ein D-m-VRSP-TW-**UC**.

Der wesentliche neue Aspekt ist hier die Bestellzeitzusage zum Bestellzeitpunkt. Wichtig ist dabei zum einen, dass einmal getroffene Lieferzeitzusagen auch eingehalten werden müssen (Irreversibilität). Zum anderen wird das Lieferzeitfenster erst bei Bestellung fixiert. Zum Dritten kann ein Lieferwunsch bei Bestellung auch abgelehnt werden. Diese drei Aspekte werden durch die neue Verfeinerung der Problemklasse um die Endung **-DO** beschrieben.

D-m-VRSP-TW-UC-DO:

dynamic m vehicle routing and scheduling problem with time constraints and uncertain customers with delivery notification at orders entry.

Nach [Domschke 1997, S. 207f]⁹¹ lässt sich ein Tourenplanungsproblem als 4-Tupel $[\alpha | \beta | \gamma | ZF]$ klassifizieren mit

- ▶ α =Depot- und Kundencharakteristik ($\alpha_1, \dots, \alpha_7$)
- ▶ β = Fahrzeugcharakteristik (β_1, \dots, β_5)
- ▶ γ = Problem- oder Zusatzcharakteristik ($\gamma_1, \dots, \gamma_3$)
- ▶ ZF= Zielfunktion $\in \{L, FZ, M, FK\}$ ⁹²

⁹¹ Angelehnt an [Desrochers et al. 1990]

⁹² L:= min. Entfernungen; FZ:= min. Fahrzeiten; M: min. Anzahl Fahrzeuge; FK: min. Fahrtkosten

Durch eine Erweiterung dieser Klassifikation in Hinblick auf die Zusatzcharakteristik und die Zielfunktionsklasse ist die formale Klassifikation des *Problems mit Lieferzeitzusagen zum Bestellzeitpunkt* möglich.

Sei $\gamma_4 \in \{ \circ, d \}$ ein neues Attribut, das die Avispräzision beschreibt

\circ *kein Avis zum Bestellzeitpunkt*

d *ein Lieferzeitintervall definierter Breite d wird zum Bestellzeitpunkt zugesagt.*

Sei X eine neue Zielfunktionsklasse

X *Maximiere dynamisch zu jedem Zeitpunkt t den Erwartungswert der Anzahl der möglichen Auslieferungen $E_t(n)$.*

Dann lässt sich das Problem auch klassifizieren als

[1, *tw*, *sel* / *M*, *tw*, *dur*, *R/F* / *d* / *X*]

1: *1-Depot; tw: Kunden mit Zeitfenstern; sel: es sind nicht alle Kunden zu bedienen*

M: *M Fahrzeuge; tw: Fahrzeuge mit Einsatzzeitbegrenzung*

dur: *identische Zeitbeschränkung für die Fahrzeuge; R/F: mehr als 1 Route pro Fahrzeug erlaubt*

d: *Lieferavis zum Bestellzeitpunkt*

X: *maximiere $E_t(n)$*

Hinweis: Im Vorlaufenden wurde stillschweigend (und im Einklang mit der aktuellen Literatur) davon ausgegangen, dass es sich bei dem Problem um ein VRSP handelt. Es sei an dieser Stelle jedoch angemerkt, dass sich das oben formulierte Problem auch als spezielles „team orienteering problem“ TOP (s. [Chao, Golden et al, 1996b]) beschreiben lässt, da nicht alle Anfragen bedient werden müssen, und für jeden besuchten Kunden ein „score“ erzielt wird⁹³. Alle weiteren oben geschilderten Eigenschaften des Problems blieben von dieser Überlegung unberührt, so dass das Problem auch als

D-m-TOP-TW-UC-DO

dynamic m vehicle team orienteering problem with time windows, uncertain customers with delivery notification at order entry

beschrieben werden kann.

⁹³ Wenn alle Kunden gleich behandelt werden, ist der „score“ pro Kunde einheitlich zu wählen (z.B. $\text{score}(i)=1 \forall i$).

5.2 Formale Beschreibung des Optimierungsproblems

Die Tourenplanung für Heimservicedienste mit Zeitzusagen zum Bestellzeitpunkt hat eine Reihe logistischer und ökonomischer Anforderungen zu erfüllen. Um diese Anforderungen klar herauszuarbeiten, sollen diese zunächst als Optimierungsproblem mit Zielfunktion und Nebenbedingungen formalisiert werden.

5.2.1 Logistische und geschäftsbedingte Nebenbedingungen

Die wesentliche Nebenbedingung im Heimlieferservice ist die Einhaltung des zugesagten Lieferzeitpunkts. Da ein Lieferzeitpunkt nur schwerlich auf die Minute genau eingehalten werden kann, ist es sinnvoll, von einem **Lieferzeitintervall** zu sprechen (gegebenenfalls definiert durch den Lieferzeitpunkt und einen zeitlichen Radius um diesen Zeitpunkt, z.B. 12:30 +/- 15 Minuten = [12:15;12:45]). Der Lieferzeitintervallradius entspricht der oben definierten Kenngröße PRÄZ.

Eine weitere wichtige Nebenbedingung ist außerdem der **begrenzte Fuhrpark** (i.d.R. Spezialfahrzeuge, die bei kurzfristigem Bedarf nicht ad hoc besorgt werden können).

Ein wichtiges Marketingwerkzeug kann die langfristige **Fahrer-Kunden-Bindung** sein. Ohne entsprechende empirische Untersuchungen ist es schwierig zu beurteilen, wie sich die Fahrer-Kunden-Bindung langfristig auf das Bestellverhalten der Kunden auswirkt. Trotzdem wird dieser Aspekt von Betreiber von Heimlieferservices gerne gesehen, auch wenn er sich nur schwer quantifizieren lässt. Aus logistischer Sicht ist es aus zwei Gründen sinnvoll, wenn ein **Fahrer Gebietskenntnisse** aufbaut. Erstens kann er nichtmodellierte Verkehrseinflüsse (Staus, Bauarbeiten) berücksichtigen, und seine Touren so effizienter abwickeln. Zweitens kann er den Orientierungsaufwand (Adresse suchen, geeignete Parkmöglichkeit suchen, geschicktes Aufteilen der Liefermenge in handliche Einheiten bei weiteren Fuß- und Treppenwegen, etc.), der beim Neukundenbesuch anfällt, minimieren. Sowohl die Fahrer-Kunden-Bindung, als auch die Fahrergebietskenntnis, als auch die oben bereits diskutierte **Warenbereitstellung** sind sehr eng mit der TGTR korreliert. Ohne eine Quantifizierung dieser drei Aspekte (Fahrer-Kunden-Bindung und Fahrer-Gebietskenntnisse und Warenbereitstellung), kann die TGTR nur schwer als Nebenbedingung formuliert werden. Auch eine Formulierung innerhalb der Zielfunktion ist nicht ohne weiteres möglich. Das Ziel der TGTR lässt sich allerdings trotzdem verwirklichen, wenn diesem Aspekt im Verfahrensdesign Rechnung getragen wird.

5.2.2 Charakter der Zielfunktion

Die wichtigste Frage ist: „Was genau ist zu optimieren?“. Man kann davon ausgehen, dass das primäre Ziel eines Anbieters von „Heimservicediensten mit Zeitzusagen zum Bestellzeitpunkt“ ist, einen **möglichst hohen Gewinn** zu erzielen.

Angenommen, dem Anbieter steht ein definiertes Investitionsvolumen für die Einrichtung eines derartigen Lieferservices zur Verfügung um z.B. den Kommissionierbetrieb zu gewährleisten, eine dedizierte Fahrzeugflotte anzuschaffen, und die entsprechenden Fahrer zu beschäftigen, (d.h. der Kostenanteil ist in erster Näherung konstant⁹⁴), so ist die **Umsatzmaximierung** ein sinnvoller Ansatz, um das eigentliche Ziel der Gewinnmaximierung zu erreichen.

Angenommen, der Umsatz ist langfristig proportional zur Anzahl der bedienten Kunden, und aus Gründen der Kundenbindung soll der angebotene Service nicht vom einmaligen Bestellvolumen eines Kunden abhängig gemacht werden (d.h. ein Kunde mit aktuell großem Warenbestellwert kann den Liefertermin nicht anders beeinflussen als ein Kunde mit aktuell kleinem Warenbestellwert), so ist die **Maximierung der Anzahl der Kunden**, die mit den verfügbaren Betriebsmitteln bedient werden, eine sinnvolle Zielfunktion.

Da weder die Anzahl, noch die Lage der Kunden ex ante bekannt ist, sollte das gewählte Planungsverfahren den **Erwartungswert der Anzahl der bedienbaren Kunden** maximieren.

Hinweis: Die obigen Überlegungen sind aus der betrieblichen Praxis motivierte Annahmen, die für diese Arbeit getroffen werden. Natürlich sind auch andere Zielfunktionen denkbar.

⁹⁴ Kosten sind aus logistischer Sicht komplexer als aus OR-Sicht. Allein die Frage ob eine Gesamtkostenbetrachtung (inkl. Fixkosten) oder eine Grenzkostenbetrachtung (nur variable Kosten exkl. Fixkosten) erfolgen soll, ist hochgradig relevant für die Optimierungsergebnisse.

Ist die Anzahl der verfügbaren Fahrzeuge kurzfristig fix, so kann die Fahrzeuganzahl als Restriktion behandelt werden. Außerdem fallen in diesem Fall die variablen km-Kosten im Verhältnis zu den Fixkosten (Flottenbereitstellung und Fahrerlöhne) kaum ins Gewicht.

Wäre die Anzahl der verfügbaren Fahrzeuge kurzfristig variabel, so müsste die Zielfunktion statt der Maximierung der Aufträge die Maximierung eines Ertrag/Kosten-Verhältnisses enthalten.

Wäre die Anzahl der verfügbaren Fahrzeuge mittelfristig variabel, so käme die Angebots-Nachfrage als zusätzliche Variable mit ins Spiel, die die „real world“-Optimierungsaufgabe noch komplexer werden lässt.

5.2.3 Modellierung des Problems

Nach der Klassifikation des Problems soll es nunmehr formal als Optimierungsproblem beschrieben werden. Die Variablenbezeichnungen sind in Anlehnung an [Domschke 1997, S. 213f] gewählt und um jeweils neue Variablen erweitert. Es sei:

n : Anzahl der besuchten Kunden

x_{ij} : Binärvariable zur Nachfolgerbeziehung mit
 $x_{ij}=1$, falls Kunde j unmittelbar nach Kunde i besucht wird
 $x_{ij}=0$, sonst

$x_{ij}(t)$: Binärvariable, die bestimmt, ob zur Zeit t ein Fahrzeug von i nach j unterwegs ist

fz_{ij} : Fahrzeit von Kunde i zu Kunde j

sz_j : Bediendauer des Kunden j

ta_j : Avisierter Lieferzeitpunkt beim Kunden j

d : Präzision

$tf_j = ta_j - d$: Früheste Ankunftszeit beim Kunden j

$ts_j = ta_j + d$: Späteste Ankunftszeit beim Kunden j

t_j : Anfangszeitpunkt der Kundenbelieferung j in Bezug auf Tourstartzeit $t_0 = 0$

M^∞ : sehr große Zahl

V : Menge aller zu besuchenden Kunden inkl. Depot (Index 0); $|V| = n+1$

Q : beliebige Teilmenge von Kunden; $Q \subseteq V - \{0\}$

M : Menge der verfügbaren Fahrzeuge

$m_{max} = \lfloor M \rfloor$: Anzahl maximal verfügbarer Fahrzeuge

m : Anzahl der eingesetzten Fahrzeuge

$T(t)$: Anzahl der Touren, die zum Zeitpunkt t bereits begonnen haben, aber noch nicht beendet sind

dur : maximale Einsatzzeit eines Fahrzeugs

X^t : Menge der unbekanntem Aufträge zum Zeitpunkt t

$V - X^t$: Menge der bekannten Aufträge zum Zeitpunkt t

tp_j : bevorzugter Lieferzeitpunkt von Kunde j

D : Toleranz, d.h. der Kunde akzeptiert eine Lieferung gdw. $ta_j \in [tp_j - D; tp_j + D]$

$to(j)$: Zeitpunkt der Bestellung („order placement“)

wz_j : Wartezeit beim Kunden j

Das Optimierungsproblem lässt sich somit formulieren als:

$$(5.1.) \max n$$

u.d.N.

$$(5.2.) \sum_{j=0}^n x_{ij} = 1 \quad \text{für } i = 1, \dots, n$$

$$(5.3.) \sum_{i=0}^n x_{ij} = 1 \quad \text{für } j = 1, \dots, n$$

$$(5.4.) \sum_{i \in Q} \sum_{j \in Q} x_{ij} \leq |Q| - 1 \quad \text{für alle Teilmengen } Q \subseteq V - \{0\} \text{ mit } |Q| \geq 1$$

$$(5.5.) tf_j \leq t_j \leq ts_j \quad \text{für } j = 1, \dots, n$$

$$(5.6.) t_j \geq t_i + sz_i + fz_{ij} - (1 - x_{ij})M^\infty \quad \text{für } i = 0, \dots, n \text{ und } j = 1, \dots, n$$

$$(5.7.) t_j + sz_j + fz_{j0} - (1 - x_{j0})M^\infty \leq dur \quad \text{für } j = 1, \dots, n$$

$$(5.8.) T(t) \leq m_{\max} \quad \text{für } t \geq t_0$$

$$(5.9.) \left\{ \begin{array}{l} to(j) > to(i) \Leftrightarrow j \in X^{to(i)} \wedge \\ to(j) \leq to(i) \Rightarrow j \in V - X^{to(i)} \end{array} \right\} \forall i, j = 1 \dots n$$

$$(5.10.) \left\{ \begin{array}{l} t = to(j) \Rightarrow ta_j \in [tp_j - D; tp_j + D] \wedge \\ t > to(j) \Rightarrow t_j \in [ta_j - d; ta_j + d] \end{array} \right\} \forall j = 1 \dots n$$

Erläuterung der Nebenbedingungen:

(5.2.) & (5.3) stellen sicher, dass jeder Kunde genau einmal verlassen bzw. erreicht wird.

(5.4.) gewährleistet, dass keine Kurzzyklen ohne Depot entstehen

(5.5.) gewährleistet, dass die avisierten Zeitfenster eingehalten werden

(5.6.) gewährleistet, dass „zeitlich konsistente Routen“ entstehen, die sowohl Fahr- (fz), und Stand- (sz) als auch implizit Wartezeiten (\geq) beinhalten können. Diese Bedingung greift nur für die Kundenpaare i, j die direkt hintereinander angefahren werden (d.h. $x_{ij}=1$). Für diese Kundenpaare wird der letzte Subtrahend $(1-x_{ij})M^\infty$ Null und fällt somit weg. Für alle anderen Kundenpaare wird dieser letzte Subtrahend mit hinreichend großem M^∞ so groß, dass die Größer-Bedingung immer gilt und somit nicht greift.

(5.7.) gewährleistet, dass die maximale Tourdauer eingehalten wird.

(5.8.) gewährleistet, dass die Anzahl verfügbarer Fahrzeuge nicht überschritten wird. Dabei sei

$$T(t) := \sum_{j=0}^n \sum_{i=0}^n x_{ij}(t)$$

mit $x_{ij}(t) := \begin{cases} 1: \text{ein Fahrzeug ist zur Zeit } t \text{ unterwegs von } i \text{ nach } j \\ 0: \text{sonst} \end{cases} \quad \forall i = 0..n \quad \forall j = 0..n \mid (i, j) \neq (0,0)$

und $x_{00}(t) := 0$

und $x_{ij} := \max(x_{ij}(t)) \quad \forall i \neq j$

D.h. falls ein Fahrzeug zur Zeit t beim Kunden i steht, ist $i=j$. Außerdem werden Be- und Entladezeiten am Depot vernachlässigt ($i=j=0$).

Die in der Literatur an dieser Stelle oft verwendete Formulierung $\sum_j x_{0j} = m$

ist zu scharf, falls Mehrfacheinsätze der Fahrzeuge zulässig sind, da diese Formulierung die Anzahl der Touren auf die Anzahl der verfügbaren Fahrzeuge beschränkt.

(5.9.) ist die Unbestimmtheitsaussage, die besagt, dass beim Eingehen des Auftrags i alle Aufträge $j \leq i$ bekannt, aber die Aufträge $j > i$ noch unbekannt sind. Sie repräsentiert den dynamischen Aspekt des Optimierungsproblems

(5.10.) zeigt den Freiheitsgradverlust bei der Zustellzeitzusage:

- ▶ vor der Bestellung liegt die mögliche Lieferzeit in einem Intervall um den bevorzugten Lieferzeitpunkt tp_j mit Radius = Toleranz D
- ▶ nach der Bestellung liegt die mögliche Lieferzeit in einem Intervall um den zugesagten Lieferzeitpunkt ta_j mit Radius = Präzision d

Diese Modellierung ist die Basis für den in Kap. 5.4. dargestellten Lösungsansatz ORA.

Hinweis: Die in (5.1.) formulierte Zielfunktion ist Ausdruck einer Politik, die davon ausgeht, dass die variablen Transportkosten im Verhältnis zum Umsatz oder zu den Fixkosten nur eine untergeordnete Rolle spielen. Sollten die variablen Transportkosten oder die Fahrzeuganzahl in die Zielfunktion eingehen, so wären die verschiedenen Ziele entweder gegeneinander zu wichten oder in eine Hierarchie zu bringen. Denkbare Alternativen zu (5.1.) wären daher beispielsweise:

(5.1.a) $\max (\lambda n - \mu m - v t_{\text{ges}})$
 λ, μ, v : Wichtungparameter
 m : Anzahl eingesetzter Fahrzeuge
 t_{ges} : Gesamttourdauer

falls die Wichtungparameter irgendwie sinnvoll bestimmt werden könnten (dies

stellt in der Praxis ein größeres Problem dar⁹⁵), oder

$$(5.1.b) \quad \max \langle n; -m; -t_{\text{ges}} \rangle \\ \langle \dots; \dots; \dots \rangle: \text{Lexikografische Ordnung}$$

falls eine derartige Zielhierarchie aufgestellt werden kann (auch dies stellt in der Praxis ein größeres Problem dar⁹⁶).

5.2.4 Alternative Problemmodellierung

Das Planungsproblem im eCommerce-Lebensmittel-Heimlieferservice lässt sich aber auch anders modellieren. Geht man davon aus, dass die Menge der zu bedienenden Kunden ex ante vorgegeben ist, so liegt ein klassisches VRSP-TW vor, das beispielweise die Minimierung der Anzahl der einzusetzenden Fahrzeuge m zum Ziel hat:

$$(5.1'.) \quad \min m$$

u.d.N.

$$(5.2'.) \quad \sum_{j=0}^n x_{ij} = 1 \quad \text{für } i = 1, \dots, n$$

$$(5.3'.) \quad \sum_{i=0}^n x_{ij} = 1 \quad \text{für } j = 1, \dots, n$$

$$(5.4'.) \quad \sum_{i \in Q} \sum_{j \in Q} x_{ij} \leq |Q| - 1 \quad \text{für alle Teilmengen } Q \subseteq V - \{0\} \text{ mit } |Q| \geq 1$$

$$(5.5'.) \quad tf_j \leq t_j \leq ts_j \quad \text{für } j = 1, \dots, n$$

$$(5.6'.) \quad t_j \geq t_i + sz_i + fz_{ij} - (1 - x_{ij})M^\infty \quad \text{für } i = 0, \dots, n \text{ und } j = 1, \dots, n$$

$$(5.7'.) \quad t_j + sz_j + fz_{j0} - (1 - x_{j0})M^\infty \leq dur \quad \text{für } j = 1, \dots, n$$

$$(5.8'.) \quad T(t) \leq m \quad \text{für } t \geq t_0$$

$$(5.9'.) \quad m \leq m_{\max}$$

Dabei ist insbesondere (5.9.') keineswegs eine triviale Nebenbedingung, da für diese Modellierung gewährleistet sein muss, dass die ex ante determinierte Anzahl n der zu bedienenden Aufträge sich überhaupt mit m_{\max} Fahrzeugen bedienen lässt.

Diese Modellierung ist die Basis für den Einsatz des FPA/PPA (s. 3.3.3. & 3.3.5.), für den auch alternative Zielfunktionen denkbar sind, wie z.B.

⁹⁵ Was wäre besser, die Bedienung von 100 Kunden mit 10 Fahrzeugen und 100 h oder die Bedienung von 99 Kunden mit 9 Fahrzeugen und 105 h Gesamttourdauer?

⁹⁶ Beispielsweise wäre mit (5.1.b) die Bedienung von 100 Kunden mit 10 Fahrzeugen „optimaler“ als die Bedienung von 99 Kunden mit 5 Fahrzeugen.

(5.1'.a) $\min (\mu m + v t_{\text{ges}})$
 μ, v : Wichtungparameter
 m : Anzahl eingesetzter Fahrzeuge
 t_{ges} : Gesamttourdauer

bei der das Optimierungsziel die Minimierung sowohl der Fahrzeuge als auch der Gesamttourdauer ist, oder

(5.1'.b) $\min \langle m; t_{\text{ges}} \rangle$
 $\langle \dots; \dots \rangle$: Lexikografische Ordnung.

bei der es zuerst darum geht, die Anzahl der eingesetzten Fahrzeuge zu minimieren, und dann die Gesamttourdauer zu minimieren. Statt t_{ges} (Gesamttourdauer) könnte auch die Minimierung von l_{ges} (Gesamttourkilometer) in die Zielfunktion eingehen.

Ist die Anzahl der verfügbaren Fahrzeuge begrenzt, und dürfen die zugesagten Zeitfenster verletzt werden, um bei knappen Ressourcen einen zulässigen Tourenplan zu finden, so lässt sich das Problem auch als m-VRSP-STW (VRSP mit „soft time windows“) modellieren:

$$(5.1'') \min \sum_j \varepsilon_j$$

u.d.N.

$$(5.2'') \sum_{j=0}^n x_{ij} = 1 \quad \text{für } i = 1, \dots, n$$

$$(5.3'') \sum_{i=0}^n x_{ij} = 1 \quad \text{für } j = 1, \dots, n$$

$$(5.4'') \sum_{i \in Q} \sum_{j \in Q} x_{ij} \leq |Q| - 1 \quad \text{für alle Teilmengen } Q \subseteq V - \{0\} \text{ mit } |Q| \geq 1$$

$$(5.5'') t f_j - \varepsilon_j \leq t_j \leq t s_j + \varepsilon_j \quad \text{für } j = 1, \dots, n$$

$$(5.6'') t_j \geq t_i + s z_i + f z_{ij} - (1 - x_{ij}) M^\infty \quad \text{für } i = 0, \dots, n \text{ und } j = 1, \dots, n$$

$$(5.7'') t_j + s z_j + f z_{j0} - (1 - x_{j0}) M^\infty \leq dur \quad \text{für } j = 1, \dots, n$$

$$(5.8'') T(t) \leq m \quad \text{für } t \geq t_0$$

$$(5.9'') m \leq m_{\max}$$

$$(5.10'') \varepsilon_i \leq \varepsilon_{\max}$$

Dabei ist ε_i die Verletzung des zugesagten Lieferzeitfensters des Kunden i , und ε_{\max} die maximal zulässige Verletzung eines zugesagten Lieferzeitfensters. Auch hier sind Zielfunktionen mit Wichtung mehrerer Optimierungsziele oder mit lexikografischer Ordnung der verschiedenen Optimierungsziele denkbar.

Diese Modellierung ist die Basis für den Ansatz FPA/REA (s.3.3.4. & 3.3.5.)

5.2.5 Zusammenfassung

Das Optimierungsproblem unterscheidet sich je nach gewähltem logistischen (bzw. marketingtechnischen) Lösungsansatz.

Die obige Modellierung des D-m-VRSP-TW-UC-DO ist ein neuer Ansatz. Aus OR-Sicht ist die Maximierung der Zielfunktion $\max n$ nicht trivial, da

- ▶ die Anzahl aller Bestellungen n zu maximieren ist, und gleichzeitig
- ▶ die Bestellungen sequentiell eingehen, und für die Bestellung j nur die bisherigen Bestellungen $1, \dots, j-1$ bekannt sind, nicht jedoch die noch ausstehenden Bestellungen.

Das eigentliche Problem ist damit noch nicht ohne Weiteres lösbar. Bevor ein Lösungsansatz vorgeschlagen wird, werden im Folgenden daher einige prinzipielle Überlegungen angestellt. Aus diesen Überlegungen wird anschließend der Lösungsansatz von ORA abgeleitet.

5.3 Überlegungen zur Problemlösung

Damit ein Verfahren zur Optimierung der Zielfunktion $\max n$ gefunden werden kann, sollte klar sein, wovon die Erreichung dieses Optimierungsziels abhängt. Als Arbeitshypothese soll b.a.w. gelten:

$$H_A: n = f(\text{Ort-Zeit-Struktur}; \text{Sättigung}, \text{Verfahren\&Parameter}; \text{Präzision})^{97}$$

u.d.N.: Anzahl verfügbare Fahrzeuge = konstant

Diese Arbeitshypothese gilt, wenn folgende Teilhypothesen gelten:

$$H_{A1}: n = f(\text{Ort-Zeit-Struktur})$$

$$H_{A2}: n = f(\text{Sättigung})$$

$$H_{A3}: n = f(\text{Verfahren\&Parameter})$$

$$H_{A4}: n = f(\text{Präzision})$$

5.3.1 Überlegungen zur Ort-Zeit-Struktur

Die Nachfrage in Hinblick auf

- ▶ die Anzahl der Nachfragepunkte,
- ▶ die geografische Lage der Nachfragepunkte und

⁹⁷ Diese Begriffe werden unmittelbar im Anschluss jeweils näher definiert und beleuchtet.

- ▶ die beschränkter zeitlicher Toleranz der Kunden,

lässt sich als dreidimensionale (x,y,t) Ort-Zeit-Struktur modellieren. In der folgenden Grafik ist der „Wunschlieferzeitpunkt“ als Punkt, und die Toleranz D um den Lieferzeitpunkt als gestrichelter Pfeil dargestellt.

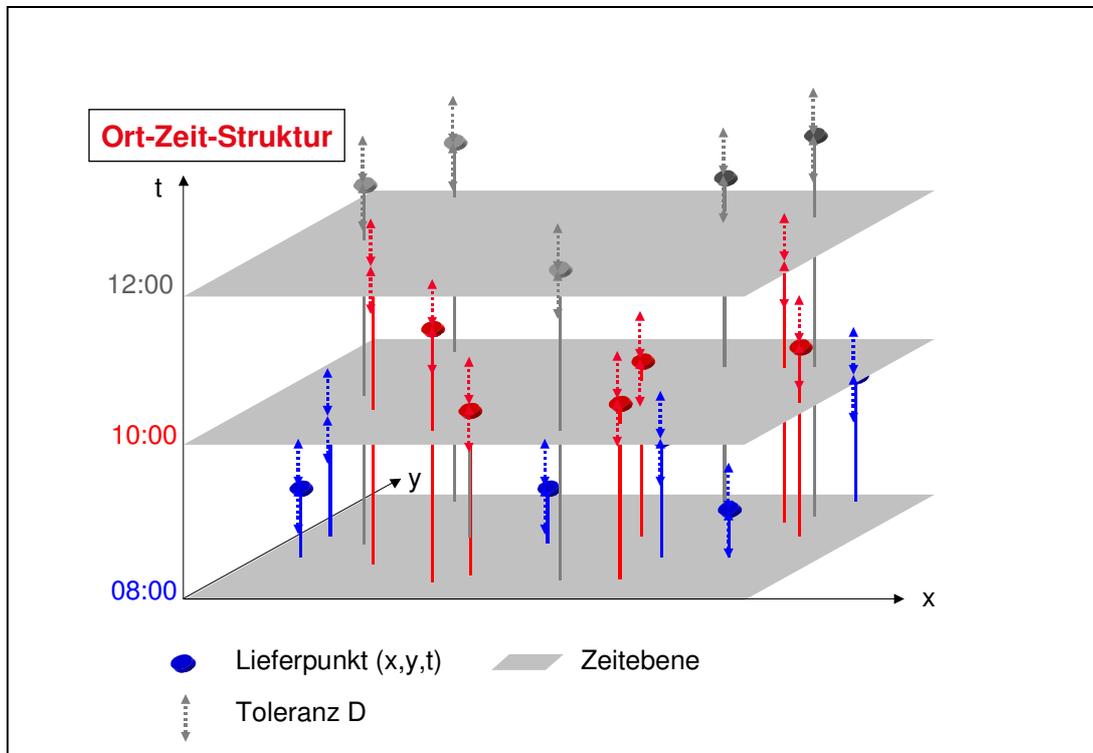


Abbildung 43: Ort-Zeit-Struktur der Nachfrage

Im folgenden wird die Hypothese

$$H_{A1}: n = f(\text{Ort-Zeit-Struktur})$$

unter den Aspekten der Gleichzeitigkeit und der „impliziten Tourstruktur“ untersucht.

5.3.1.1 Gleichzeitigkeit

Die Anzahl der verfügbaren Fahrzeuge beschränkt die Anzahl der Lieferungen, die an verschiedenen geografischen Orten „zum selben Zeitpunkt“ zugesagt werden können. Dabei ist der Begriff „zum selben Zeitpunkt“ etwas ungenau und soll daher durch ein formales Maß zur „Gleichzeitigkeit“ ersetzt werden.

Seien die in einem beliebigen Lieferszenario zu bedienenden n Stops sortiert nach aufsteigender frühester Lieferzeit t_f :

$$t_{f_i} < t_{f_j} \Leftrightarrow i < j \quad (5.11)$$

Sei dann das Gleichzeitigkeitsmaß G die maximale Anzahl der Belieferungen im Lieferszenario, die gleichzeitig stattfinden:

$$G = 1 + \max(k) \mid tf_{j-k} + sz_{j-k} > ts_j \quad \forall k=1..j-1, \forall j=1..n \quad (5.12)$$

D.h. von keinem der k Vorgänger i von j in der nach tf sortierten Liste kann j ohne Lieferzeitintervallverletzung angefahren werden, selbst wenn die Fahrzeit von i nach j vernachlässigt wird.

Da G definitionsgemäß eine Untergrenze der benötigten Anzahl von Fahrzeugen m für ein definiertes Lieferszenario ist, dürfen keinesfalls Lieferzeitintervalle so zugesagt werden, so dass $G > m$ ist (Vermeidung zu vieler gleichzeitiger Lieferungen⁹⁸).

In der Untersuchung von Variationen der Solomoninstanz C107 wird der Effekt der Gleichzeitigkeit klar deutlich. Schrumpft die Toleranz auf wenige Minuten (d.h. $tf_{j-k} \rightarrow ts_j$), so werden die rot gekennzeichneten Lieferpunkte im R^3 „gleichzeitig“.

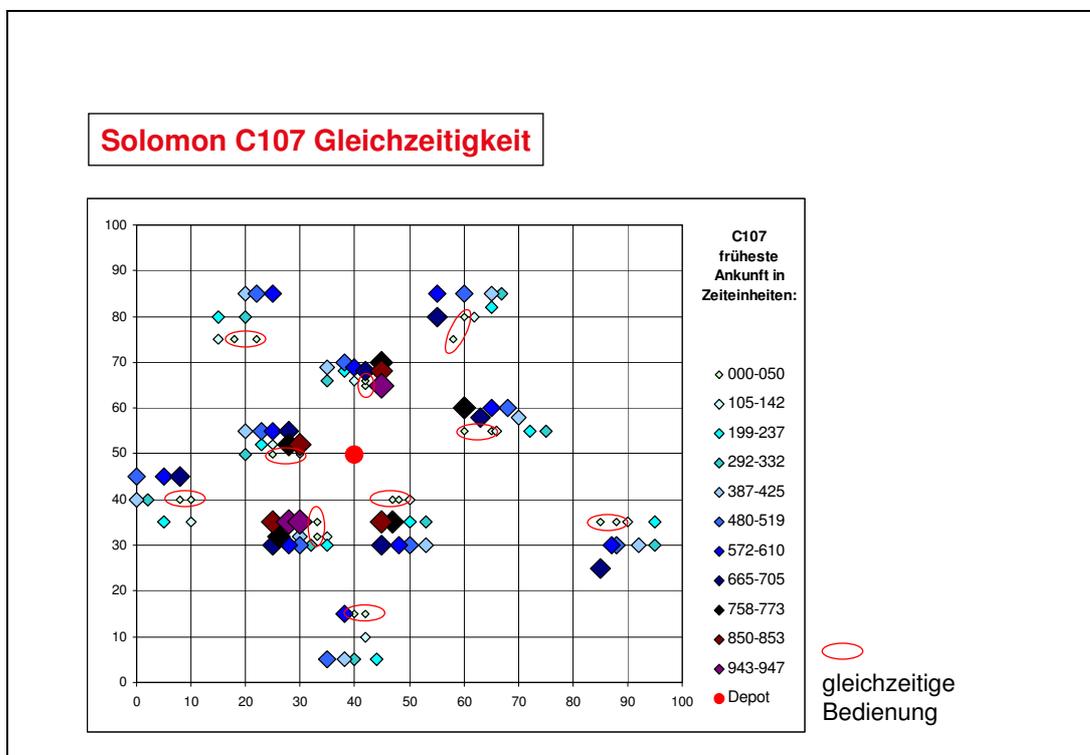


Abbildung 44: Ort-Zeit-Struktur der Nachfrage – Gleichzeitigkeit

Dies schlägt sich erwartungsgemäß unabhängig vom Planungsverfahren in der Anzahl der benötigten Fahrzeuge nieder.

⁹⁸ Ein pragmatischer Lösungsansatz aus der Praxis für das D-m-VRSP-TW-UC-DO ist geprägt durch diese Idee der Vermeidung von Gleichzeitigkeit. Dabei werden unabhängig von der geografischen Lage der Kunden für eine definierte Zeitspanne maximal p Lieferzusagen getroffen (z.B. max. 10 Lieferungen pro Tageszeitstundenintervall). Bei Einführung des Lieferservice wird p relativ klein gewählt. Sobald sich das Zustellsystem eingespielt hat, wird p solange langsam erhöht, wie die eingehenden Aufträge noch mit den verfügbaren m Fahrzeugen bedient werden können, ohne die Lieferzeitzusagen signifikant zu verletzen.

Dieser pragmatische Ansatz vertraut darauf, dass für das so erzeugte VRSP-TW irgendeine Lösung mit max. m Fahrzeugen gefunden werden kann. Gelingt das immer leicht, kann p erhöht werden, gelingt das regelmäßig nicht mehr, wird entweder p gesenkt oder m erhöht (weitere Fahrzeuge werden angeschafft). Obwohl dieses Vorgehen aus OR-Sicht ein reiner try&error-Ansatz ist, findet er doch Anwendung in der Praxis.

Analyse C107# - Anzahl Fahrzeuge

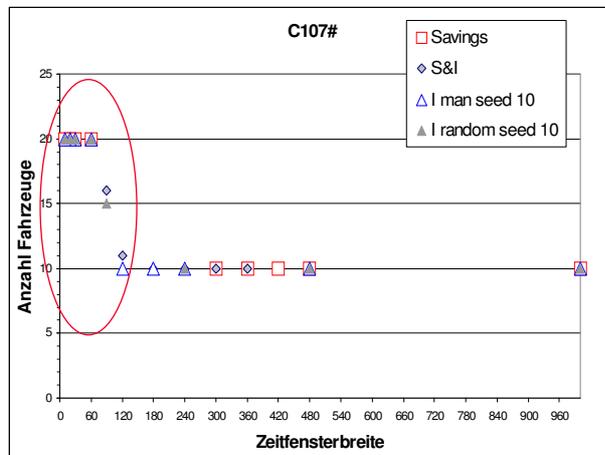


Abbildung 45: Zusammenhang Anzahl Fahrzeuge – Gleichzeitigkeit

Untersuchung mit 4 verschiedenen Verfahren:

Savings: Savingsverfahren

S&I: Kombination aus Savings & Insertionverfahren

I man seed 10: Insertion mit manueller Festlegung der ersten 10 seed-Kunden

I random seed 10: Insertion mit zufälliger Festlegung der ersten 10 seed-Kunden

Das Beispiel C107# zeigt, daß bei enger werdenden Zeitfenstern die „Gleichzeitigkeit“ die Anzahl der benötigten Fahrzeuge steigen läßt, da die Fahrzeuge nicht gleichzeitig an hier 20 verschiedenen Orten (♦ in Grafik: „Ort-Zeit-Struktur der Nachfrage – Gleichzeitigkeit“) sein können. Bei ex ante fixer Fahrzeuganzahl hat die Gleichzeitigkeit im Umkehrschluss Auswirkungen auf die Anzahl der bedienbaren Kunden. Die Hypothese $H_{A1}: n = f(\text{Ort-Zeit-Struktur})$ ist somit durch das obige Beispiel untermauert.

5.3.1.2 Implizite Sequenzialität

Def. „Implizite Sequenzialität“

„Implizite Sequenzialität“ sei eine Eigenschaft eines Transportauftragsbestands, die einem raum-zeitlichen Tourverlauf durch den Transportauftragsbestand entspricht. Diese implizite Sequenzialität muss nicht bekannt sein.

Die Betrachtung von Variationen der Solomon-Instanz C107 zeigt erwartungsgemäß den Effekt der „Implizite Sequenzialität“ auf die Anzahl benötigter Fahrzeuge (und der Transportkosten von der Lieferzeitintervallbreite). Bis zur Erreichung der „Gleichzeitigkeitsrestriktion“ (s.o.) verlaufen alle 3 Kenngrößen m ,

l_{ges} und t_{ges} konstant (d.h. unabhängig von der Lieferzeitintervallbreite). Dieser Effekt lässt sich leicht erklären, da Solomon die Instanz entlang einer vordefinierten Tourstruktur (= implizite Sequenzialität) definiert hat, von der er annimmt, sie sei optimal.

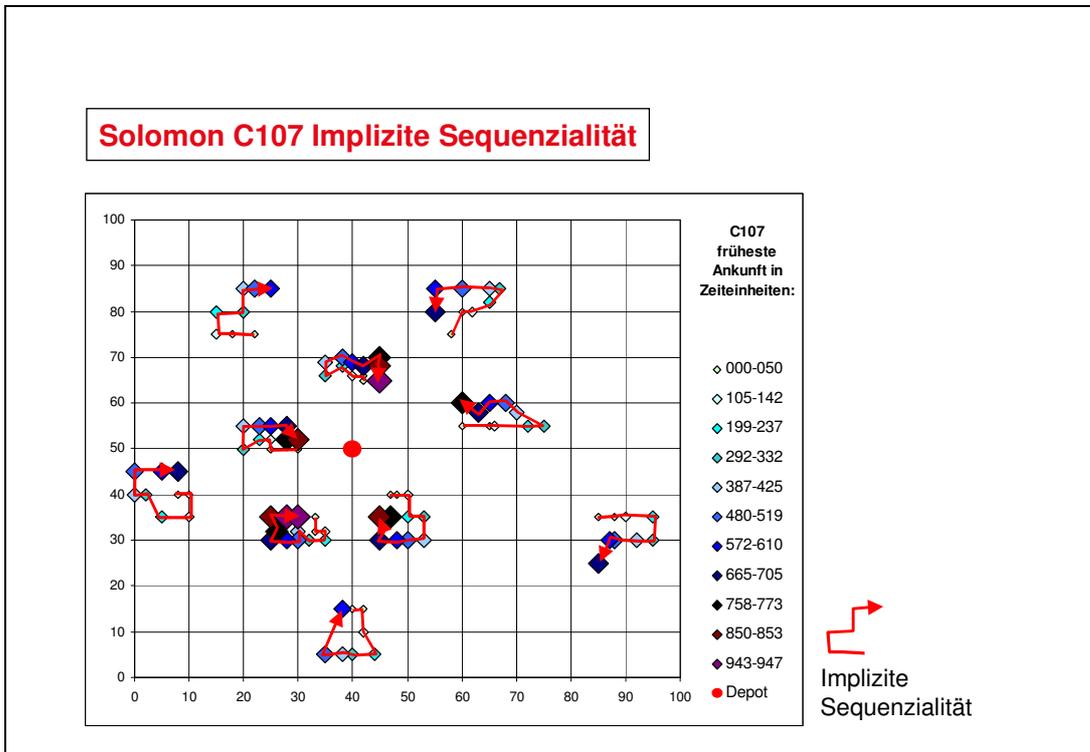


Abbildung 46: Ort-Zeit-Struktur der Nachfrage – Implizite Sequenzialität

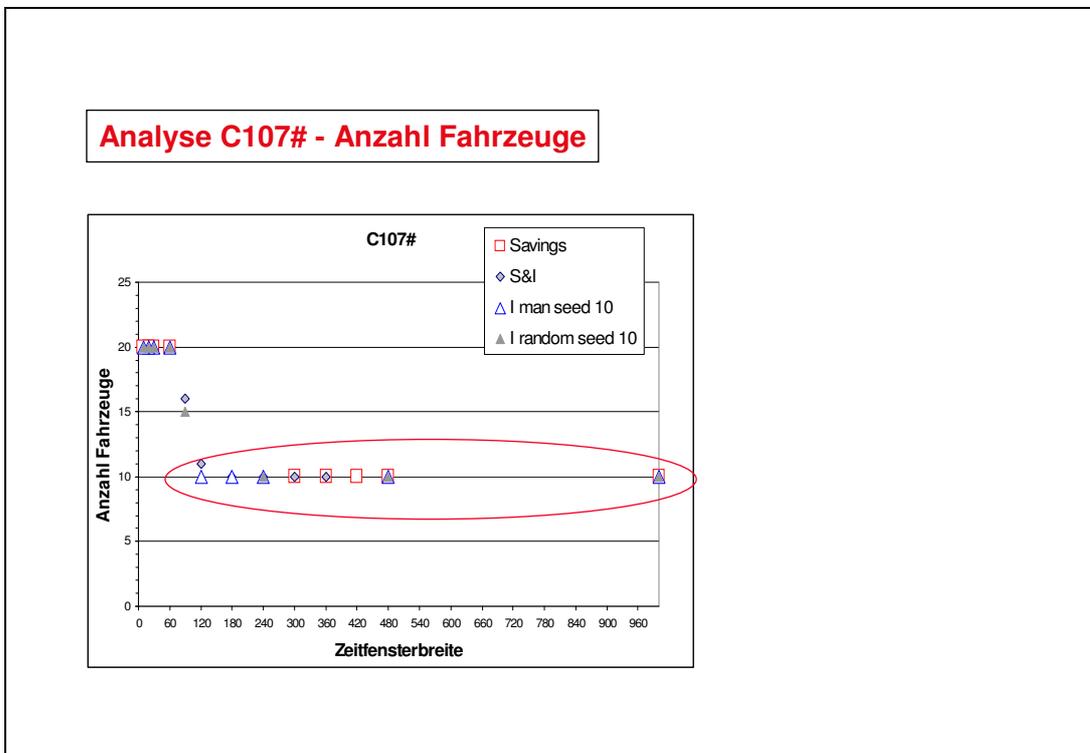


Abbildung 47: Zusammenhang „Anzahl Fahrzeuge – Zeitfensterbreite“ bei Impliziter Sequenzialität

5.3.2 Überlegungen zur Sättigung

Def. „Sättigung“

Sättigung S sei das Verhältnis zwischen der nachfragebedingt erwarteten Obergrenze von Lieferanfragen $E(\max(N))$ und der unter besten Bedingungen von den zur Verfügung stehenden Ressourcen ($m :=$ Anzahl Fahrzeuge) umsetzbaren Aufträgen n .

$$S := \frac{E(\max(N))}{\max n(m)} \quad (5.13)$$

Außerdem sei:

$$\begin{aligned} \text{„Untersättigung“} &\Leftrightarrow S \ll 1 \Leftrightarrow \\ &\Leftrightarrow E(\max(N)) \ll \max n(m) \end{aligned} \quad (5.14)$$

$$\begin{aligned} \text{„Sättigung“} &\Leftrightarrow S \approx 1 \Leftrightarrow \\ &\Leftrightarrow E(\max(N)) \approx \max n(m) \end{aligned} \quad (5.15)$$

$$\begin{aligned} \text{„Übersättigung“} &\Leftrightarrow S \gg 1 \Leftrightarrow \\ &\Leftrightarrow E(\max(N)) \gg \max n(m) \end{aligned} \quad (5.16)$$

Es ist leicht zu zeigen, dass die Sättigung wesentlichen Einfluss auf den Zielwert der Zielfunktion $\max n$ hat. Es gilt:

$$n \leq E(\max(N)) = S \cdot \max n(m) \quad (5.17)$$

D.h. der maximale Zielfunktionswert muss kleiner sein als die erwartete Anzahl von Aufträgen. Außerdem gilt:

$$n \leq \max n(m) = \frac{E(\max(N))}{S} \quad (5.18)$$

D.h. der maximale Zielfunktionswert muss auch kleiner sein als die maximale Anzahl von Aufträgen, die mit m Fahrzeugen bewältigt werden kann. Somit ist die Hypothese H_{A2} : $n = f(\text{Sättigung})$ bewiesen.

5.3.3 Überlegungen zu Verfahren & Parametern

Die Hypothese H_{A3} : $n = f(\text{Verfahren \& Parameter})$ ist trivialerweise richtig, solange es sich nicht um exakte Verfahren handelt. Dies gilt sowohl für die unterschiedlichen denkbaren Verfahren an sich, als auch für unterschiedliche Parametrierungen eines Verfahrens, wenn diese unterschiedliche Lösungen generieren. An dieser Stelle soll diese knappe Aussage genügen, da auf diesen Aspekt später noch eingehend eingegangen wird (s. Kap 6)

5.3.4 Überlegungen zur Präzision

Die Hypothese H_{A4} : $n = f(\text{Präzision})$ ist leicht einsichtig, da die Präzision die wesentliche Restriktion im Planungsproblem darstellt. Eine kurze Extremwertbetrachtung zeigt dies:

- ▶ Sei die Gesamtheit aller potenzieller Kunden die Menge $\{A,B\}$ mit den beiden Kunden A und B
- ▶ Sei die raum-zeitliche Lage der Kunden A und B so, dass sie beide mit einem Fahrzeug in der Reihenfolge $[A,B]$ bedient werden können.
Die minimale Fahrzeit zwischen ihnen betrage fz_{AB}
Die avisierten Lieferzeitpunkte seien ta_A und ta_B mit
 $ta_B = ta_A + sz_A + fz_{AB} - \Delta$
Das Lieferzeitfenster sei definiert über die Lieferzeitpräzision $[ta_i - d; ta_i + d]$
 Δ sei eine beliebige positive Zeitdauer.
- ▶ Eine Bedienung in der Reihenfolge $[A,B]$ ist möglich, gdw:
 $ta_B + d \geq ta_A - d + sz_A + fz_{AB} \Leftrightarrow$
 $ta_A + sz_A + fz_{AB} - x + d \geq ta_A - d + sz_A + fz_{AB} \Leftrightarrow$
 $- \Delta + d \geq -d \Leftrightarrow$
 $2d \geq \Delta$
- ▶ D.h. wenn $2d \geq \Delta$ dann können beide Kunden mit einem Fahrzeug in der Reihenfolge $[A,B]$ bedient werden, sonst nicht. Gleiches gilt analog für die Reihenfolge $[B,A]$
- ▶ Somit ist bei der Verfügbarkeit von nur einem Fahrzeug
 $n = 2$ für $d \geq \Delta/2$ und
 $n = 1$ für $d < \Delta/2$

Die Präzision der Lieferzeitzusage ist somit ein wesentlicher Einflussfaktor für den Zielfunktionswert n .

5.3.5 Überlegungen zur Erreichung des Optimierungsziels

Da die Ressourcen zur Transportabwicklung (gesamtes Einsatzzeitpotenzial für eine gegebene Anzahl von Fahrzeugen) beschränkt sind, teilt sich dieses Einsatzzeitpotenzial auf folgende Einsatzzeitklassen auf:

- ▶ Servicezeiten (sz),
- ▶ Fahrzeiten (fz), und
- ▶ Wartezeiten (wz).

Formal bedeutet dies:

$$\sum_{i=1}^n s z_i + \sum_{i=1}^n f z_{p(i),i} + \sum_{\forall i=p(0)} f z_{i,0} + \sum_{i=1}^n w z_i \leq m \cdot dur \quad (5.19)$$

$\forall i=p(0) :=$ alle Vorangänger von „0“⁹⁹ im Tourenplan
d.h. alle jeweils letzten Stops der Touren

Da die Summe der Servicezeiten $s z_i$ linear mit dem Zielfunktionswert n wächst, muss zur Maximierung von n die Summe der Fahrzeiten $f z$ minimiert werden, um Zeitpotenzial für die Bedienung möglichst vieler Kunden zu schaffen. Eine notwendige Voraussetzung zur Erreichung von (5.1.) im Übersättigungsfall ist daher die Zielfunktion ZF'

$$ZF' = \min \sum_{i=1}^n f z_{p(i),i} + \sum_{\forall i=p(0)} f z_{i,0} \quad (5.20)$$

bzw. die Minimierung der durchschnittlichen Fahrzeit zwischen zwei Kunden ZF'' .

$$ZF'' = \min \overline{f z} \quad (5.21)$$

Im folgenden werden daher Überlegungen zur Abschätzung der minimalen, maximalen und durchschnittlichen Fahrzeit zwischen zwei Kunden und der daraus resultierenden Anzahl der bedienbaren Punkte angestellt.

5.3.6 Überlegungen zu kompakten Tourgebieten

Die Erwartungswerte für die Entfernung zwischen zwei Tourpunkten sind eng korreliert mit der erwarteten Anzahl der bedienbaren Kunden. Um die Signifikanz dieser Überlegung auszuloten, werden im Folgenden die Fragen untersucht „Wie weit liegen „best case“ und „worst case“ in Bezug auf mittlere Entfernungen zwischen zwei Kunden auseinander?“ und „Welcher Mittelwert ist für die mittlere Entfernungen zwischen zwei Kunden zu erwarten?“.

Natürlich hängt die Antwort auf diese Fragen von vielen Faktoren ab. Aber in einigen idealisierten Fällen lassen sich diese Fragen gut beantworten. Daher werden im Folgenden ein paar konstruierte Laborbeispiele untersucht.

5.3.6.1 Minimale & Maximale Entfernungen

Zuerst soll der Unterschied zwischen „best cases“ und „worst cases“ von Durchschnittsentfernungen innerhalb einer Tour betrachtet werden. Hierzu werden folgende vereinfachende Annahmen getroffen:

⁹⁹ 0 steht für das Depot

- ▶ Das Depot und die Kunden liegen jeweils auf den Ecken eines n-Ecks.
- ▶ Die Anzahl der Tourpunkte (Depot & Summe aller Stops) sei ungerade.

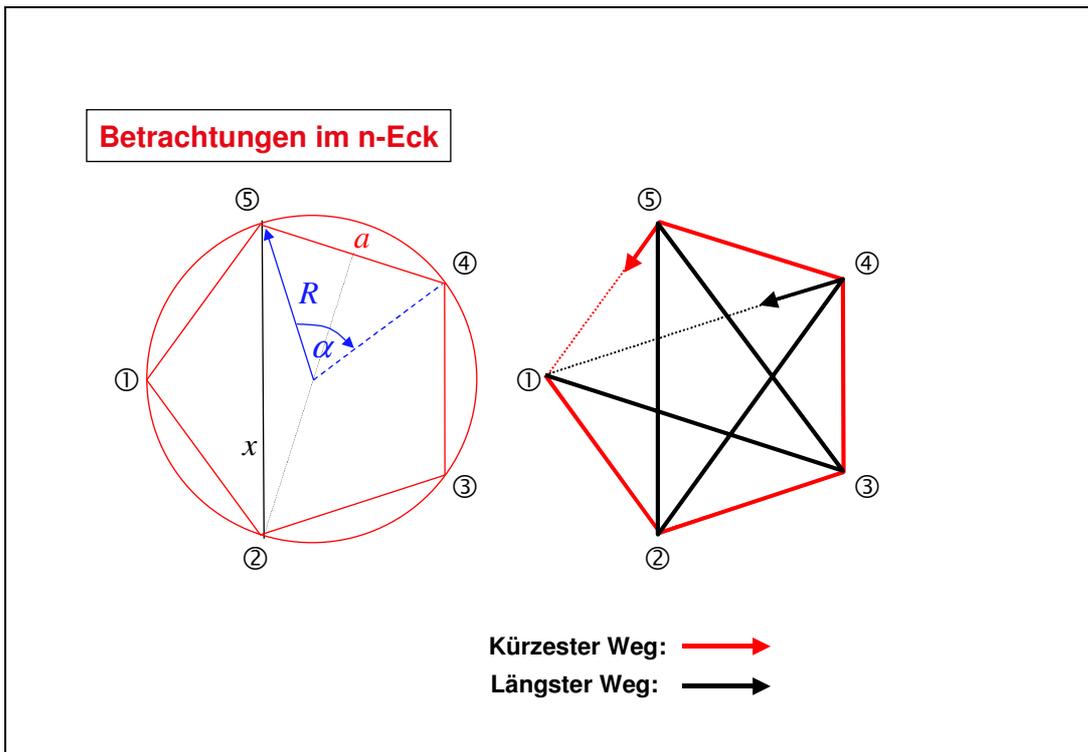


Abbildung 48: Kürzeste und längste Wege im n-Eck

Der kürzeste Weg läuft dabei entlang der Kanten a , der längste Weg jeweils über die Diagonalen x .

Folgende Größen seien definiert:

n : Anzahl Ecken (=Summe der Tourpunkte)

a : Kantenlänge des n-Ecks

x : Diagonalenlänge des n-Ecks (= längste Verbindung zwischen zwei Ecken)

L_n^a : Weglänge best case: $L_n^a = a \cdot n$
(Verbindung aller Ecken entlang der Kanten zu einer Tour)

L_n^x : Weglänge worst case: $L_n^x = x \cdot n$
(Verbindung aller Ecken entlang der Diagonalen zu einer Tour)

α : Innenwinkel

R : Umkreisradius

Nach [Bronstein & Semendjajew 1991, S.193] gilt:

$$\alpha_n = 360^\circ / n = 2\pi / n \quad (5.22)$$

$$R_n = \frac{a_n}{2\sin(\alpha_n/2)} \quad (5.23)$$

$$\begin{aligned} x_n^2 &= \left(\frac{a_n}{2}\right)^2 + \left(R_n + R_n \cos\left(\frac{\alpha_n}{2}\right)\right)^2 \Leftrightarrow \\ x_n^2 &= \left(\frac{a_n}{2}\right)^2 + \left(R_n \left(1 + \cos\left(\frac{\alpha_n}{2}\right)\right)\right)^2 \Leftrightarrow \\ x_n^2 &= \left(\frac{a_n}{2}\right)^2 + R_n^2 \left(1 + \cos\left(\frac{\alpha_n}{2}\right)\right)^2 \Rightarrow \\ x_n &= \sqrt{\left(\frac{a_n}{2}\right)^2 + R_n^2 \left(1 + \cos\left(\frac{\alpha_n}{2}\right)\right)^2} \end{aligned} \quad (5.24)$$

$$\lim_{n \rightarrow \infty} x_n = \lim_{a_n \rightarrow 0 \wedge \alpha_n \rightarrow 0} x_n = \sqrt{0 + R_n^2 (1+1)^2} = 2R_n \quad (5.25)$$

Für das Verhältnis der Längen des kürzesten zum längsten möglichen Weg gilt für **große n**:

$$\lim_{n \rightarrow \infty} \frac{L_n^a}{L_n^x} = \lim_{n \rightarrow \infty} \frac{a_n \cdot n}{x_n \cdot n} \quad (\text{mit 5.25})$$

$$= \lim_{n \rightarrow \infty} \frac{a_n}{2R_n} = \quad (\text{mit 5.23})$$

$$= \lim_{n \rightarrow \infty} \frac{a_n}{\left(\frac{2a_n}{2\sin(\alpha_n/2)}\right)} = \lim_{n \rightarrow \infty} \sin(\alpha_n/2) = \quad (\text{mit 5.22})$$

$$= \lim_{n \rightarrow \infty} \sin\left(\frac{2\pi}{n}\right) = \lim_{n \rightarrow \infty} \sin\left(\frac{\pi}{n}\right) = \frac{\pi}{n} \quad (5.26)$$

D.h. je größer die Anzahl Stops n desto größer wird die Differenz zwischen dem theoretisch besten und schlechtesten Weg.

Wie sieht das Verhältnis für **kleine n** aus? Eine kleine Messreihe gibt dazu Aufschluss:

n	π/n	$\frac{L_n^a}{L_n^x}$	3/n	(5.27)
n=3	1,047	1	1	
n=5	0,628	0,615	0,600	
n=7	0,449	0,444	0,429	
n=9	0,349	0,347	0,333	
n=21	0,150	0,149	0,143	
n=99	0,032	0,032	0,030	

D.h. für das Verhältnis der Längen des kürzesten zum längsten möglichen Weg gelten die Ober- und Untergrenzen:

$$\frac{\pi}{n} \geq \frac{L_n^a}{L_n^x} \geq \frac{3}{n} \quad (5.28)$$

Damit ist gezeigt, dass die Spanne der durchschnittlichen Entfernungen zwischen zwei Stops zwischen dem „best case“ und dem „worst case“ extrem viel Optimierungsspielraum lässt.

Obwohl diese Erkenntnis wichtig ist, hilft sie in Hinblick auf die Verfahrenssuche bisher noch nicht viel weiter. Wenden wir uns daher jetzt den Durchschnittsentfernungen in kompakten Tourgebieten zu.

5.3.6.2 Durchschnittsentfernungen

Für die folgenden Überlegungen seien folgende Größen definiert:

d_q° : durchschnittliche Entfernung zwischen zwei Punkten im Einheitsquadrat (Fläche $A=1$; Kantenlänge $d=1$).

$d_q^{\#(n,k)}$: durchschnittliche Entfernung zwischen zwei Punkten auf einem speziellen „random walk“ im Einheitsquadrat (Fläche $A=1$; Kantenlänge $d=1$), das in k^2 Teilquadrate unterteilt ist und n Punkte enthält, wobei maximal 2 der Punkte eines Teilquadrats mit Punkten anderer Teilquadrate verbunden sein dürfen (s. Grafik unten)

$d_{iq}^\circ(k)$: durchschnittliche Entfernung zwischen zwei Punkten im Teilquadrat (Fläche $A=1/k^2$; Kantenlänge $d=1/k$).

D_k° : durchschnittliche Entfernung zwischen zwei Punkten im Einheitskreis mit $R=1$ (Fläche $A=R^2\pi=\pi$; Radius $R=1$).

d_k° : durchschnittliche Entfernung zwischen zwei Punkten im Kreis mit $A=1$ (Fläche $A=1$; Radius $R=1/\sqrt{\pi}$).

L Länge eines geschlossenen Random Walks

In einem Einheitsquadrat mit Kantenlänge 1 ist die durchschnittliche Entfernung zwischen zwei Punkten¹⁰⁰ (x,y) und (i,j)

$$d^{\varnothing}_q = \int_{x=0}^1 \int_{y=0}^1 \int_{i=0}^1 \int_{j=0}^1 \sqrt{(i-x)^2 + (j-y)^2} dx dy di dj \quad (5.29)$$

Für das Einheitsquadrat ist nach einer Berechnung nach der Monte-Carlo-Methode mit 4000 Punktepaaren $d^{\varnothing}_q = 0,52$. Der durchschnittliche Abstand von zwei Punkten in einem geschlossenen „random walk“ im Quadrat ist somit auch d^{\varnothing}_q .

Wird das Quadrat jedoch in k^2 Teilquadrate unterteilt, so dass sich im Teilquadrat im Schnitt n/k^2 Punkte befinden, so ist der durchschnittliche Abstand von zwei Punkten in einem geschlossenen „random walk“, der jedes Teilquadrat nur einmal durchläuft“ $d^{\varnothing\#}_q(n,k)$ signifikant kleiner als d^{\varnothing}_q .

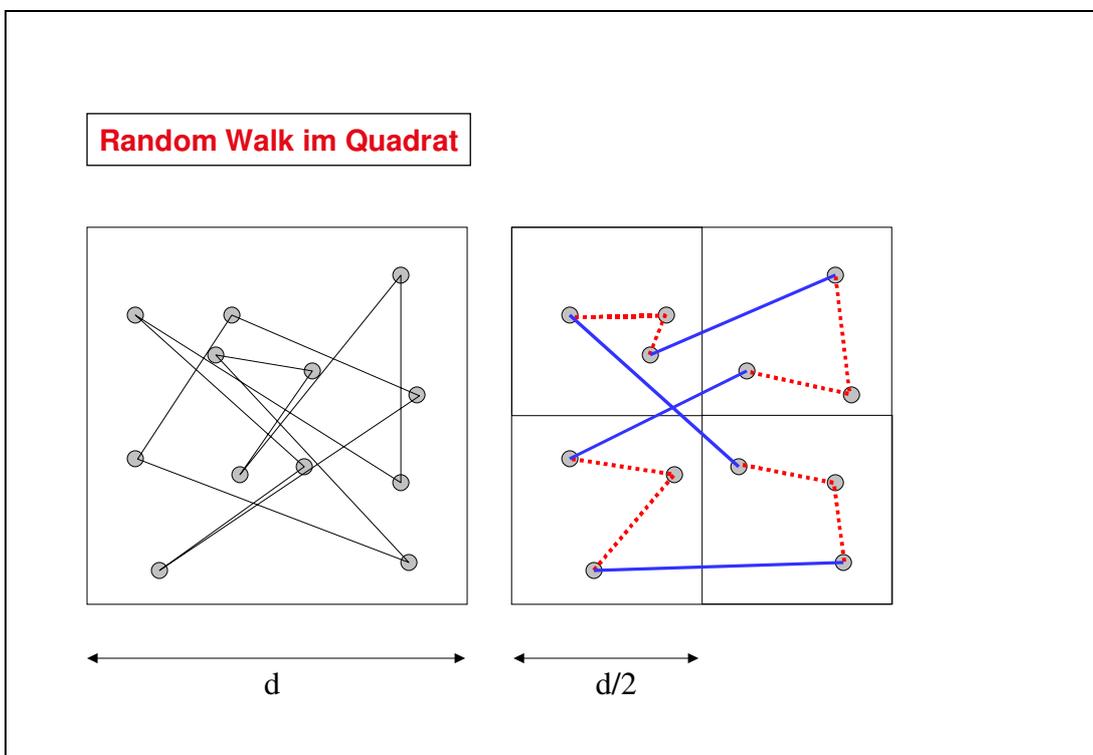


Abbildung 49: Random Walks
 $k=1; k^2=1; n=12$

$k=2; k^2=4; n=12$

Der Unterschied zwischen $d^{\varnothing\#}_q(n,k)$ und d^{\varnothing}_q lässt sich folgendermaßen konservativ abschätzen. Der „random walk“ L setzt sich zusammen aus:

- ▶ k^2 mal der durchschnittliche Entfernung zwischen zwei Punkten im Gesamtgebiet (d^{\varnothing}_q), und
- ▶ $k^2(n/k^2-1)$ mal der durchschnittlichen Entfernung zwischen zwei Punkten in einem Teilquadrat ($d^{\varnothing}_{iq}(k) = 1/k d^{\varnothing}_q$)

¹⁰⁰ im kartesischen Koordinatensystem

$$L = k^2 d_q^\circ + k^2 (n/k^2 - 1) d_{iq}^\circ(k) \quad (5.30)$$

$$d_q^{\circ\#}(n,k) = L/n \quad (5.31)$$

$$d_q^{\circ\#}(n,k) = 1/n (k^2 d_q^\circ + (n-k^2) 1/k d_q^\circ) \quad (5.32)$$

$$d_q^{\circ\#}(n,k) = 1/n k^2 (d_q^\circ + (n-1)/k d_q^\circ) \quad (5.33)$$

$$d_q^{\circ\#}(n,k) = \frac{k^2 - k + \frac{n}{k}}{n} d_q^\circ \quad (5.34)$$

Für das obige Beispiel gilt somit:

$$d_q^{\circ\#}(n=12, k=2) \leq 0,667 d_q^\circ$$

$$d_q^{\circ\#}(n=100, k=2) \leq 0,520 d_q^\circ$$

$$d_q^{\circ\#}(n=1000, k=3) \leq 0,339 d_q^\circ$$

Allgemein gilt für „viele Stops“:

$$\lim_{n \rightarrow \infty} d_q^{\circ\#} \leq \lim_{n \rightarrow \infty} \frac{k^2 - k + \frac{n}{k}}{n} d_{\text{quadrat}(A=1)}^\circ = \frac{1}{k} d_q^\circ \quad (5.35)$$

Übertragen auf die Tourenplanung für ein quadratisches Gebiet mit einem Fahrzeug heißt das: durch die Zerlegung des Quadrats in k^2 Teilquadrate kann die durchschnittliche Entfernung zwischen zwei Punkten um den Faktor $1/k$ gesenkt werden, falls $n \rightarrow \infty$.

In der Praxis wird n nicht gegen unendlich gehen, da die Anzahl von Lieferwünschen pro Teilquadrat (Region) beschränkt ist. Trotzdem ist dieses Prinzip der Senkung der Durchschnittsentfernungen durch „Clustering“ interessant, und wird im später vorgestellten Verfahren genutzt werden.

Betrachten wir nun noch kurz, ob die geografische Form eines kompakten Gebiets einen entscheidenden Einfluss auf die zu erwartenden Durchschnittsentfernungen hat. Im Einheitskreis mit Radius $R=1$ beträgt die durchschnittliche Entfernung D_k° zwischen zwei Punkten¹⁰¹ (a,α) und (b,β)

$$D_k^\circ = \int_{a=0}^1 \int_{b=0}^1 \int_{\alpha=0}^{2\pi} \int_{\beta=0}^{2\pi} \frac{1}{4\pi^2} \sqrt{b^2 + a^2 - 2ab \cos(\alpha - \beta)} \quad (5.36)$$

Für den Einheitskreis ist nach einer eigenen Berechnung nach der Monte-Carlo-Methode mit 4000 Punktpaaren $D_k^\circ = 0,90$. Um diesen Wert sinnvoll mit dem des

¹⁰¹ im Polarkoordinatensystem

Einheitsquadrats vergleichen zu können, muss noch d_k^{\varnothing} für den Kreis mit Flächeninhalt 1 berechnet werden.

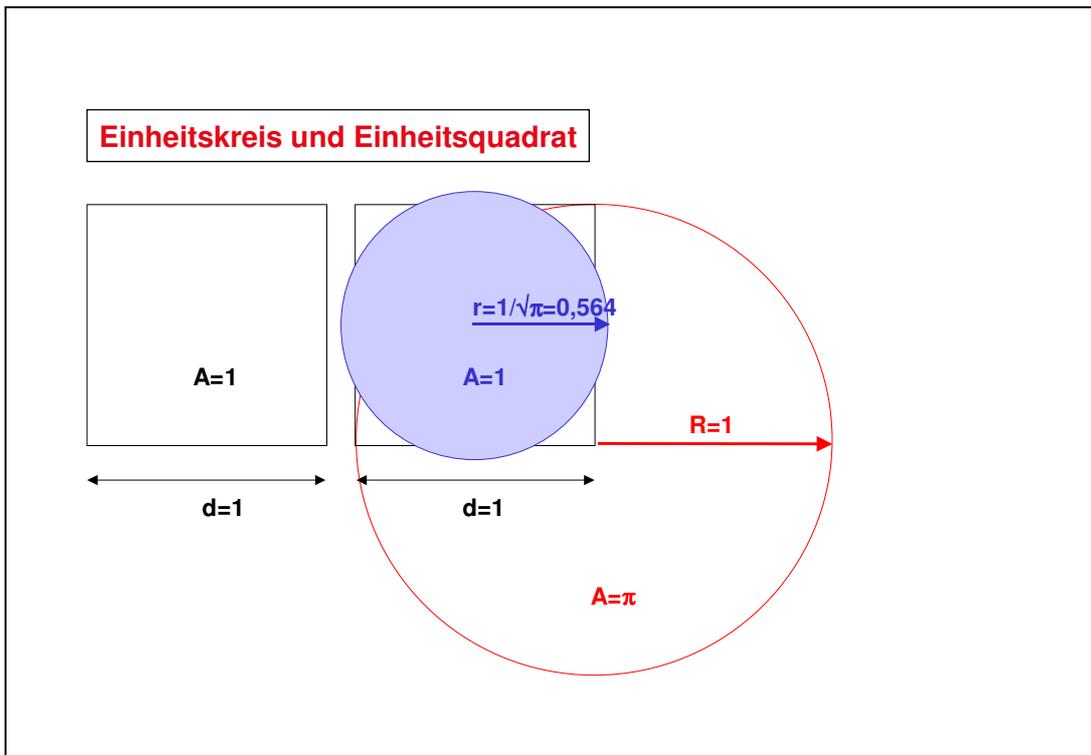


Abbildung 50: Vergleich Einheitsquadrat, Einheitskreis (R=1) und Kreis (A=1)

Es gilt:

$$\text{Einheitskreis (mit Radius } R=1\text{): } A' = R^2 \pi = \pi \quad (5.37)$$

$$\text{Kreis (mit Flächeninhalt } A=1\text{): } A = r^2 \pi = 1 \quad (5.38)$$

$$1/\pi = r^2 \quad (5.39)$$

$$r = 1/\sqrt{\pi} \quad (5.40)$$

$$d_k^{\varnothing} = D_k^{\varnothing} \cdot 1/\sqrt{\pi} = 0,90 \cdot 1/\sqrt{\pi} = 0,51 \quad (5.41)$$

D.h. der Erwartungswert der Entfernung zweier Punkte im Quadrat unterscheidet sich nicht wesentlich von dem zweier Punkte im Kreis gleichen Flächeninhalts.

Die vorangegangenen Überlegungen zeigen zwei wesentliche Ergebnisse für das 1-Fahrzeug-Problem. Zum einen konnte gezeigt werden, dass in der zu erwartenden durchschnittlichen Entfernung zwischen zwei Stops ein signifikantes Optimierungspotenzial liegt. Zum anderen konnte gezeigt werden, dass durch Unterteilung kompakter Tourgebiete in kleinere Untereinheiten die Durchschnittsentfernungen zwischen zwei Stops gesenkt werden können.

5.3.7 Überlegungen zu nichtkompakten Tourgebieten

Während Innenstadtgebiete grob durch kompakte Tourgebiete approximiert werden können, sind Tourgebiete, die auch die Randgemeinden abdecken i.d.R. nicht kompakt, sondern weisen bereits eine räumliche Clusterung in sich auf.

Sei vereinfachend folgende Tourgebietsstruktur angenommen:

- ▶ alle potenziellen Kunden und das Depot liegen im \mathbb{R}^2 „äquidistant“ auf 5 lokalen Kreisumfängen, und
- ▶ die lokalen Kreise liegen auf einem übergeordneten Kreisbogensegment mit 5 lokalen Kreisen.

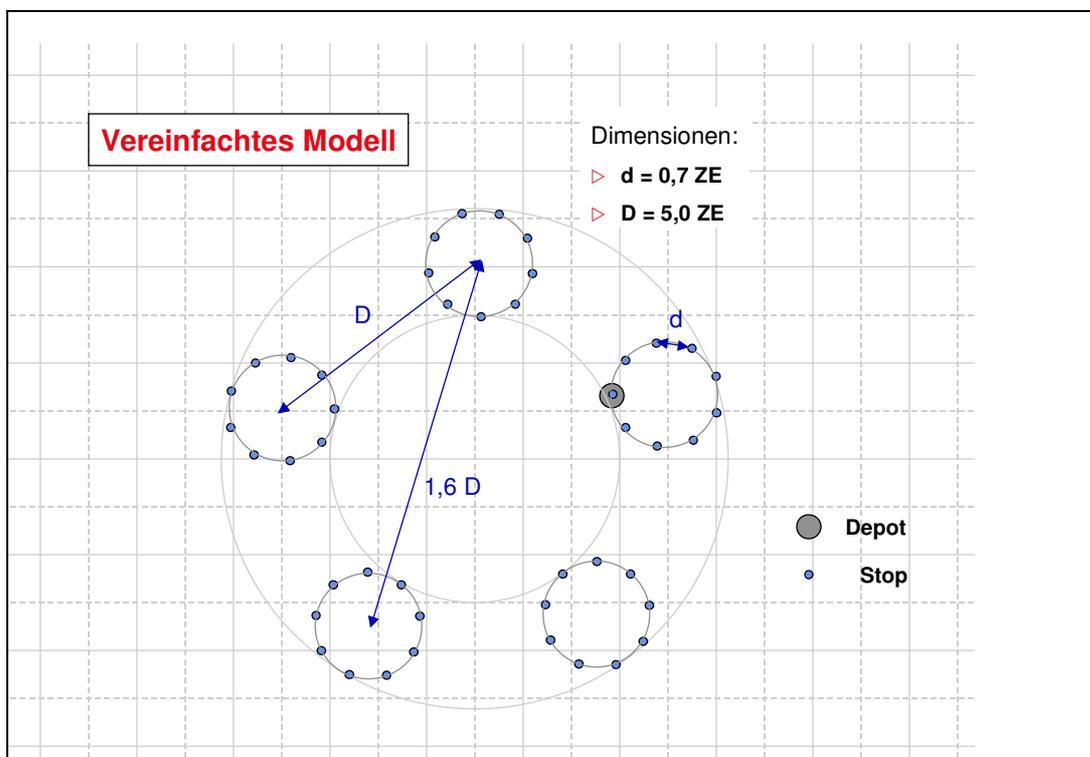


Abbildung 51: Vereinfachte Tourgebietsstruktur mit 5 Clustern

Der „best case“ Tourverlauf ist intuitiv einsichtig von folgender Struktur¹⁰²:

¹⁰² Möglicherweise lassen sich noch kleinste Reihenfolgeverbesserungen der Art [...8,17,9,10,...,16,47,18,19, ...] erzielen, aber die Tourstruktur ist sicher ein „best case“.

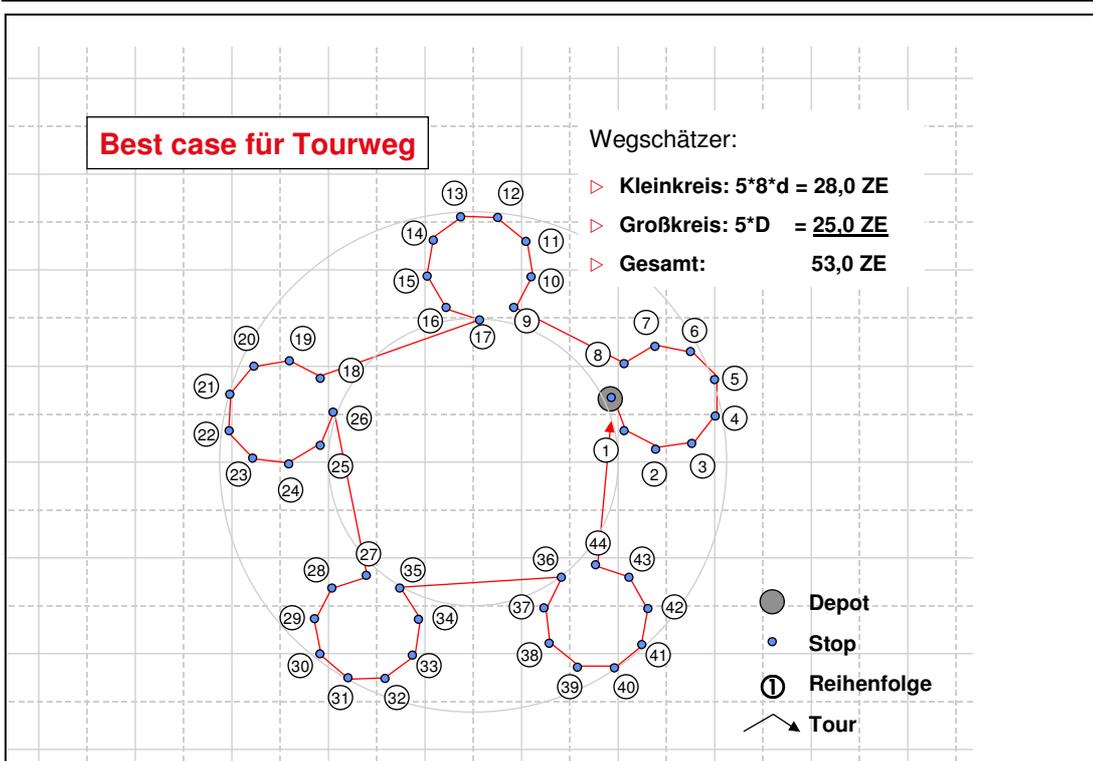


Abbildung 52: „best case“ Tourverlauf bei Tourgebietsstruktur mit 5 Clustern

Der „worst case“ Tourverlauf ist intuitiv einsichtig von folgender Struktur:

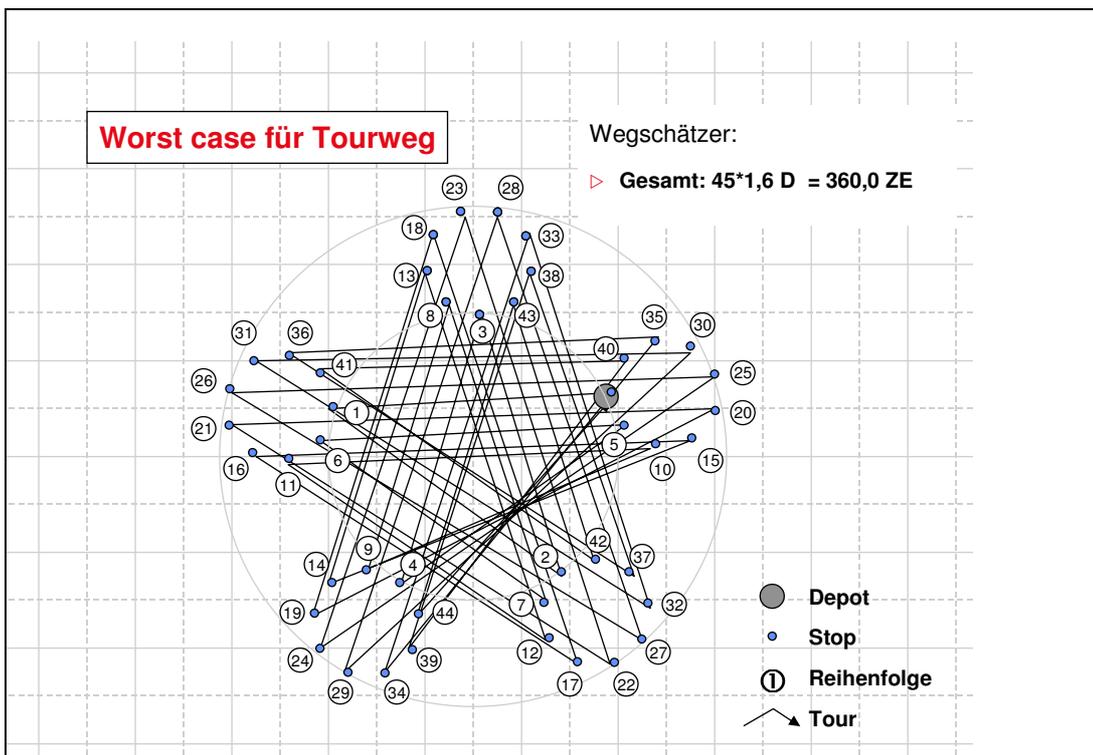


Abbildung 53: „worst case“ Tourverlauf bei Tourgebietsstruktur mit 5 Clustern

Für das oben gezeigte Beispiel gilt, dass die Durchschnittsentfernung für 2 Stops im „worst case“ ca. 6,8 (=360/53) mal so lang ist, wie im „best case“.

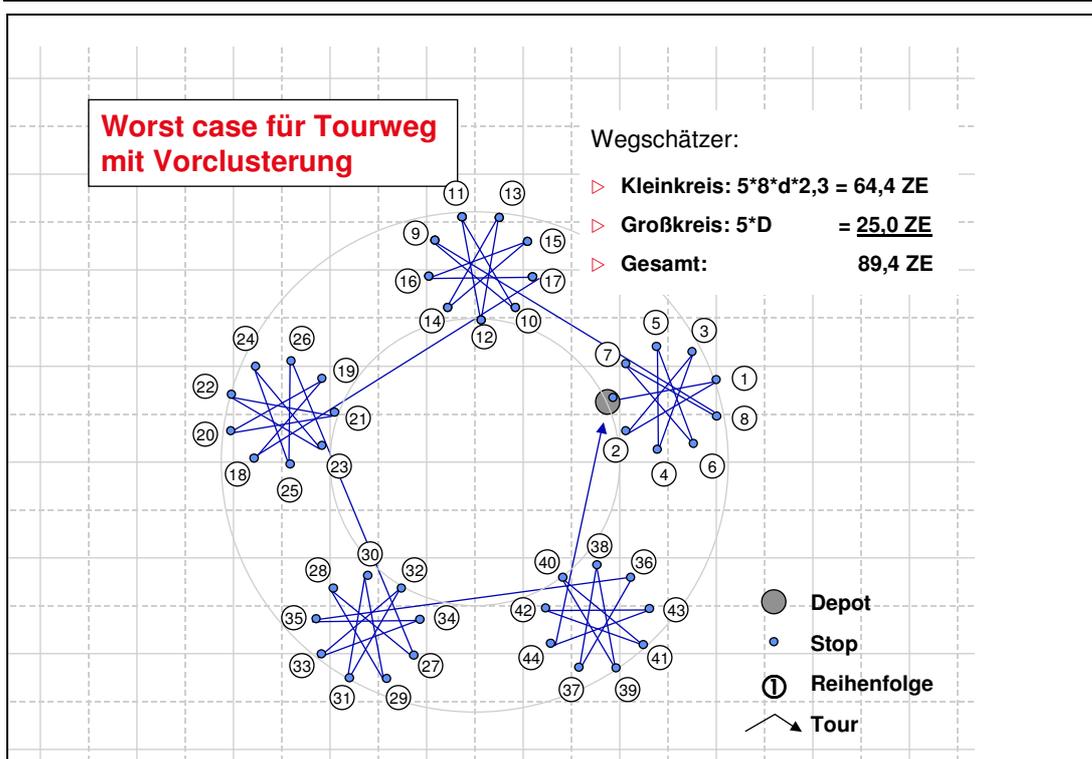


Abbildung 54: worst case Tourverlauf bei „Vorclustering“

Durch eine Vor-Clustering kann das „worst case“/„best case“-Verhältnis auf 1,7 (=89,4/53) signifikant verbessert werden.

Aus den vorangegangenen Überlegungen können folgende aussagenlogische Schlüsse für die Tourenplanung **unter engen Zeitfensterrestriktionen** gezogen werden:

Wenn ein „großes Tourgebiet“ (A) zu bedienen ist, und keine „implizite Sequenzialität“ in der Kundenstruktur (B) vorliegt, ist ein „großer Servicezeitverlust durch die entstehenden Fahrzeiten“ (C) zu erwarten.
 $[A \wedge \neg B \Rightarrow C]$

Durch Negation dieses Satzes folgt:

Soll ein großer Servicezeitverlust durch die entstehenden Fahrzeiten vermieden werden ($\neg C$), so ist entweder das Tourgebiet zu verkleinern ($\neg A$), oder eine implizite Sequenzialität zu gewährleisten (B). $[\neg C \Rightarrow \neg A \vee B]$

Da ein Fahrer im Rahmen seiner Arbeitszeit entweder fahren, oder Kunden beliefern kann, (und daher die Summe aus Servicezeiten und Fahrzeiten konstant ist¹⁰³) folgen die Korollare:

¹⁰³ Wartezeiten können vereinfachend als besondere unproduktive Fahrzeiten gewertet werden.

Wenn ein „großes Tourgebiet“ (A) zu bedienen ist und keine „implizite Sequenzialität“ in der Kundenstruktur (B) vorliegt, ist eine „geringe Anzahl von bedienbaren Kunden“ (D) zu erwarten. $[A \wedge \neg B \Rightarrow D]$

und

Soll eine große Anzahl von bedienbaren Kunden erzielt werden ($\neg D$), so ist entweder das Tourgebiet zu verkleinern ($\neg A$), oder eine implizite Sequenzialität zu gewährleisten (B). $[\neg D \Rightarrow \neg A \vee B]$

5.3.8 Überlegungen zum Mehrfahrzeugproblem

Die bisherigen Überlegungen gingen von einem Ein-Fahrzeug-Problem aus. Anhand des folgenden Modells werden Überlegungen zum Mehrfahrzeugproblem angestellt.

5.3.8.1 Ein Labor Modell

Für die weiteren Betrachtungen sei ein einfaches Labor-Modell A durch folgende Annahmen definiert:

- ▶ Sei das geografische Liefergebiet eine quadratische Fläche mit Kantenlänge 6 Längeneinheiten (LE)
- ▶ Sei der Lieferzeithorizont 4 Zeiteinheiten (ZE)
- ▶ Sei die maximale Geschwindigkeit der Fahrzeuge 1 LE / ZE
- ▶ Seien Wartezeiten in den Touren zulässig
- ▶ Seien die Lieferungen punktgenau zu einem definierten Lieferzeitpunkt durchzuführen.
- ▶ Seien insgesamt 12 Kunden zu bedienen, jeweils 3 zu den Zeiten 1,2,3,4
- ▶ Seien die Bedienzeiten = 0
- ▶ Seien die jeweiligen Depotan- und -abfahrten nicht zu betrachten.

5.3.8.2 Auswirkung Raum-Zeitstruktur auf Fahrzeuganzahl

Die folgenden 3 Beispiele zeigen, dass die zeitlich-geografische Lage der Lieferpunkte die fahrzeugminimale Bedienung der Lieferpunkte entscheidend beeinflusst.

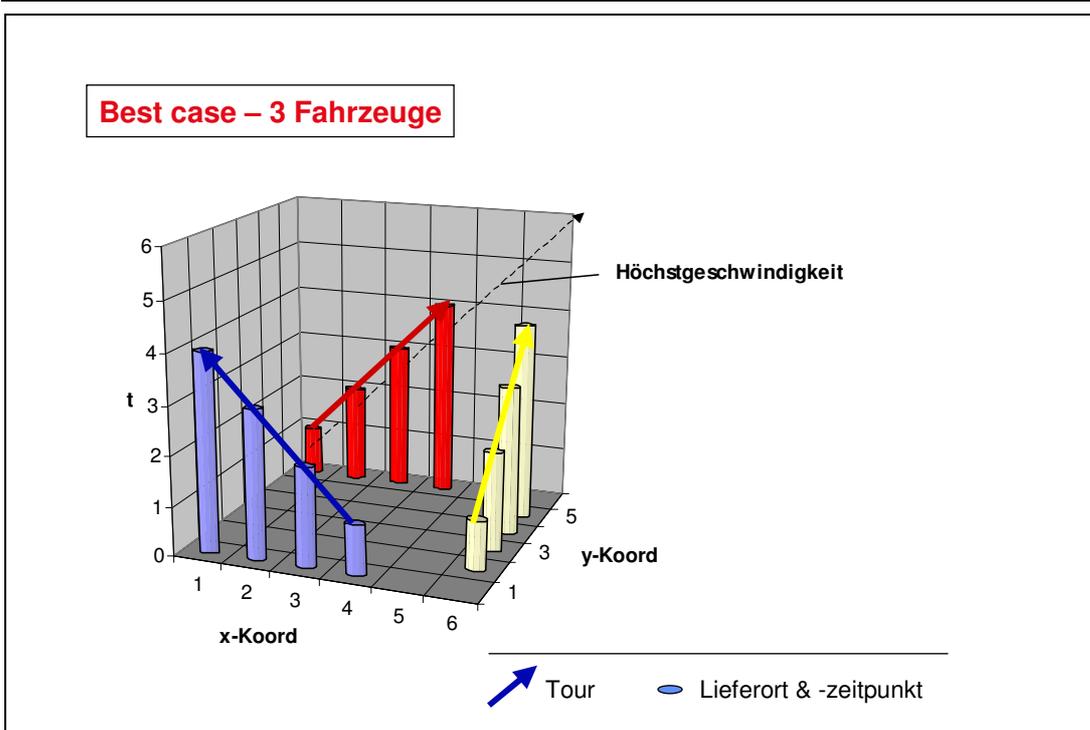


Abbildung 55: Ein „best case“ Szenario für das Modell A

In diesem beispielhaften „best case“ Szenario liegen die Lieferorte und –zeitpunkte so, dass eine zulässige Lösung mit nur 3 Fahrzeugen erzielt werden kann.

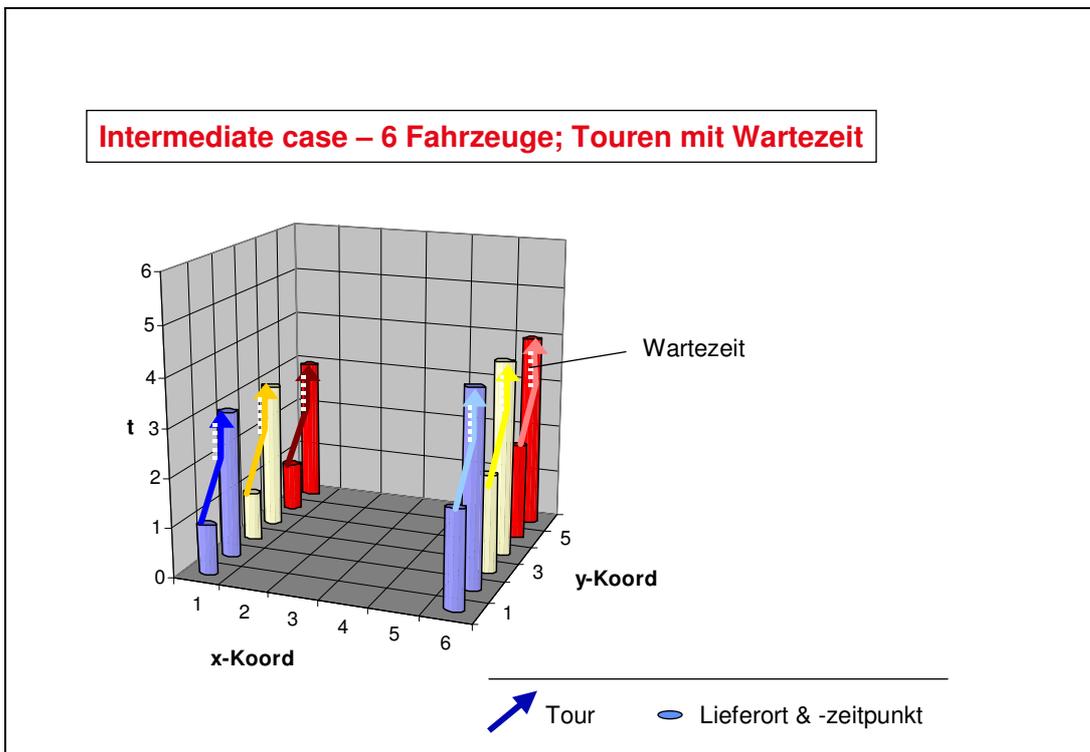


Abbildung 56: Ein „intermediate case“ Szenario für das Modell A

In diesem beispielhaften „intermediate case“ Szenario liegen die Lieferorte und –zeitpunkte so, dass eine zulässige Lösung mit nur 6 Fahrzeugen erzielt werden kann, wenn die Touren Wartezeiten beinhalten.

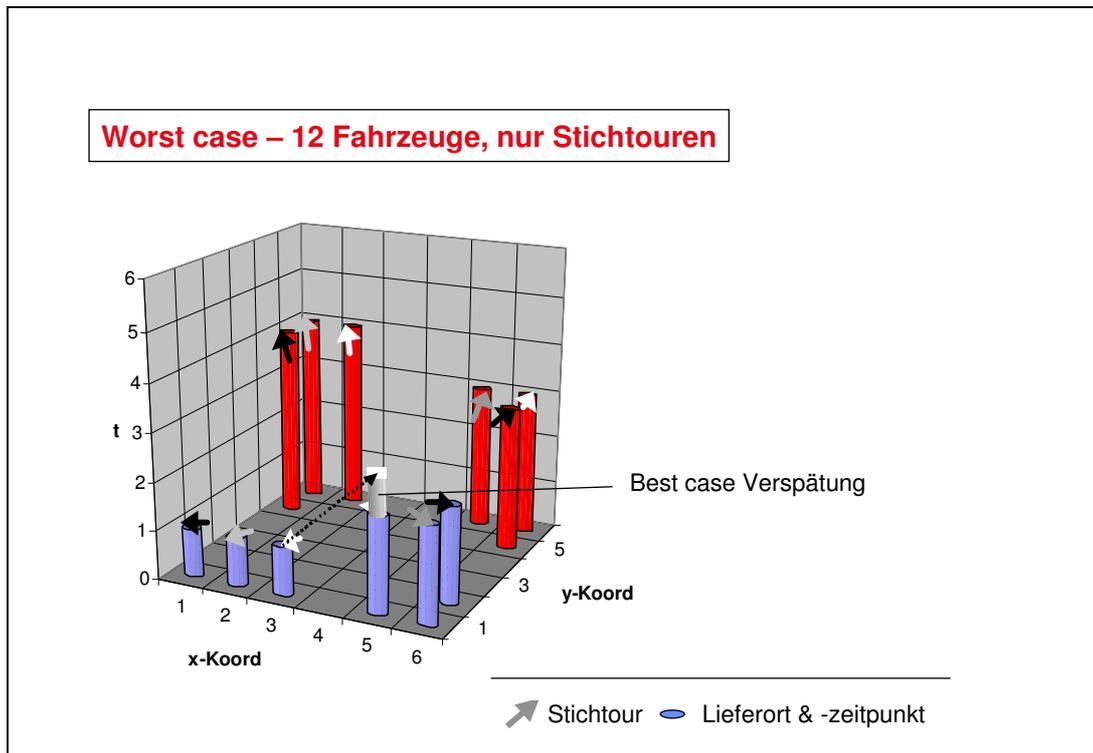


Abbildung 57: Ein „worst case“ Szenario für das Modell A

In diesem beispielhaften „worst case“ Szenario liegen die Lieferorte und Lieferzeitpunkte so, dass eine zulässige Lösung nur mit 12 Fahrzeugen erzielt werden kann, von denen jedes einen Kunden in einer Stichtour bedient. Kombinationsmöglichkeiten von mindestens 2 Kunden in einer Tour gibt es nicht, da die Entfernungen zwischen allen Kunden jeweils so groß sind, dass sie ohne Überschreitung der Fahrzeughöchstgeschwindigkeit nicht so überbrückt werden können, dass die Lieferzeitvorgaben irgendeines anderen Lieferorts eingehalten werden könnten.

Vermutung:

Für das konstruierte Beispiel gibt es keine zulässigen Touren, die mehr als einen Kunden bedienen.

Anschauliches Beispiel:

Die Entfernung zwischen den Punkten (3,1) und (5,1) beträgt 2 LE, die Fahrzeit somit auch 2 ZE. Bei einer Abfahrt an (3,1) zur Zeit 1 kann (5,1) somit frühestens zur Zeit 3 erreicht werden. Die Lieferung muss aber bereits zur Zeit 2 erfolgen – d.h. im best case würde das Fahrzeug mit einer Verspätung von 1 ZE an (5,1) eintreffen.

Beweis:

Def. Seien für diesen Beweis folgende Größen definiert:

(i,j) eine Tour mit den beiden Lieferpunkten i und j

$FAb(i)$ die früheste Abfahrtszeit im Punkt i

$FAn(j)^i$ die früheste Ankunftszeit in j u.d.N. dass i der Vorgänger von j ist

$SAn(j)$ die späteste Ankunftszeit im Punkt j

$E(i,j)$ die Fahrzeit zwischen i und j

Zu zeigen:

$$\neg \exists (i,j) \text{ mit } FAn(j)^i < SAn(j) \quad \forall i,j \in \{a,b,c,d,e,f,g,h,i,j,k,l\}, i \neq j \Leftrightarrow$$

$$\neg \exists (i,j) \text{ mit } (FAb(i) + E(i,j)) < SAn(j) \quad \forall i,j \in \{a,b,c,d,e,f,g,h,i,j,k,l\}, i \neq j \Leftrightarrow$$

$$\min (FAb(i) + E(i,j)) > SAn(j) \quad \forall i,j \in \{a,b,c,d,e,f,g,h,i,j,k,l\}, i \neq j$$

Für dieses kleine Labor-Beispiel lässt sich dieser Beweis leicht durch vollständige Enumeration führen:

Entfernungsmatrix E (von?nach?)													geografische Lage						Koordinaten			
	a	b	c	d	e	f	g	h	i	j	k	l	1	2	3	4	5	6	x	y		
a	0,00	1,00	2,00	4,00	5,00	5,10	6,40	7,07	6,40	5,10	5,00	4,00	6	k	j			i	h	a	1	1
b	1,00	0,00	1,00	3,00	4,00	4,12	5,66	6,40	5,83	5,00	5,10	4,12	5	l					g	b	2	1
c	2,00	1,00	0,00	2,00	3,00	3,16	5,00	5,83	5,39	5,10	5,39	4,47	4							c	3	1
d	4,00	3,00	2,00	0,00	1,00	1,41	4,12	5,10	5,00	5,83	6,40	5,66	3							d	5	1
e	5,00	4,00	3,00	1,00	0,00	1,00	4,00	5,00	5,10	6,40	7,07	6,40	2						f	e	6	1
f	5,10	4,12	3,16	1,41	1,00	0,00	3,00	4,00	4,12	5,66	6,40	5,83	1	a	b	c		d	e	f	6	2
g	6,40	5,66	5,00	4,12	4,00	3,00	0,00	1,00	1,41	4,12	5,10	5,00								g	6	5
h	7,07	6,40	5,83	5,10	5,00	4,00	1,00	0,00	1,00	4,00	5,00	5,10								h	6	6
i	6,40	5,83	5,39	5,00	5,10	4,12	1,41	1,00	0,00	3,00	4,00	4,12								i	5	6
j	5,10	5,00	5,10	5,83	6,40	5,66	4,12	4,00	3,00	0,00	1,00	1,41								j	2	6
k	5,00	5,10	5,39	6,40	7,07	6,40	5,10	5,00	4,00	1,00	0,00	1,00								k	1	6
l	4,00	4,12	4,47	5,66	6,40	5,83	5,00	5,10	4,12	1,41	1,00	0,00								l	1	5

Früheste-Abfahrtszeiten-Matrix FAb (von?nach?)													zeitliche Lage						
	a	b	c	d	e	f	g	h	i	j	k	l	1	2	3	4	5	6	
a	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	6	4	4			3	3
b	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	5	4					3
c	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	4						
d	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	3						
e	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2						2
f	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	1	1	1	1		2	2
g	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00							
h	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00							
i	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00	3,00							
j	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00							
k	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00							
l	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00	4,00							

Früheste-Ankunftszeiten-Matrix FAn (von?nach?)													min(frühesteAnkunft)		SpätesteAnkunft:		
	a	b	c	d	e	f	g	h	i	j	k	l	a	b	a	b	
a		2,00	3,00	5,00	6,00	6,10	7,40	8,07	7,40	6,10	6,00	5,00	a	2,00	>	a	1,00
b	2,00		2,00	4,00	5,00	5,12	6,66	7,40	6,83	6,00	6,10	5,12	b	2,00	>	b	1,00
c	3,00	2,00		3,00	4,00	4,16	6,00	6,83	6,39	6,10	6,39	5,47	c	2,00	>	c	1,00
d	6,00	5,00	4,00		3,00	3,41	6,12	7,10	7,00	7,83	8,40	7,66	d	3,00	>	d	2,00
e	7,00	6,00	5,00	3,00		3,00	6,00	7,00	7,10	8,40	9,07	8,40	e	3,00	>	e	2,00
f	7,10	6,12	5,16	3,41	3,00		5,00	6,00	6,12	7,66	8,40	7,83	f	3,00	>	f	2,00
g	9,40	8,66	8,00	7,12	7,00	6,00		4,00	4,41	7,12	8,10	8,00	g	4,00	>	g	3,00
h	###	9,40	8,83	8,10	8,00	7,00	4,00		4,00	7,00	8,00	8,10	h	4,00	>	h	3,00
i	9,40	8,83	8,39	8,00	8,10	7,12	4,41	4,00		6,00	7,00	7,12	i	4,00	>	i	3,00
j	9,10	9,00	9,10	9,83	###	9,66	8,12	8,00	7,00		5,00	5,41	j	5,00	>	j	4,00
k	9,00	9,10	9,39	###	###	###	9,10	9,00	8,00	5,00		5,00	k	5,00	>	k	4,00
l	8,00	8,12	8,47	9,66	###	9,83	9,00	9,10	8,12	5,41	5,00		l	5,00	>	l	4,00

$FAn = E + FAb$													qed			
-----------------	--	--	--	--	--	--	--	--	--	--	--	--	-----	--	--	--

Abbildung 58: Beweis der Unkombinierbarkeit im „worst case“ Szenario für das Model A

D.h. das „worst case“ Szenario des Model A benötigt 12 Fahrzeuge.

Durch eine veränderte zeitliche Lage der Lieferpunkte bei gleicher geografischer Lage sinkt die Anzahl der benötigten Fahrzeuge signifikant auf 4 (s. folgende Grafik).

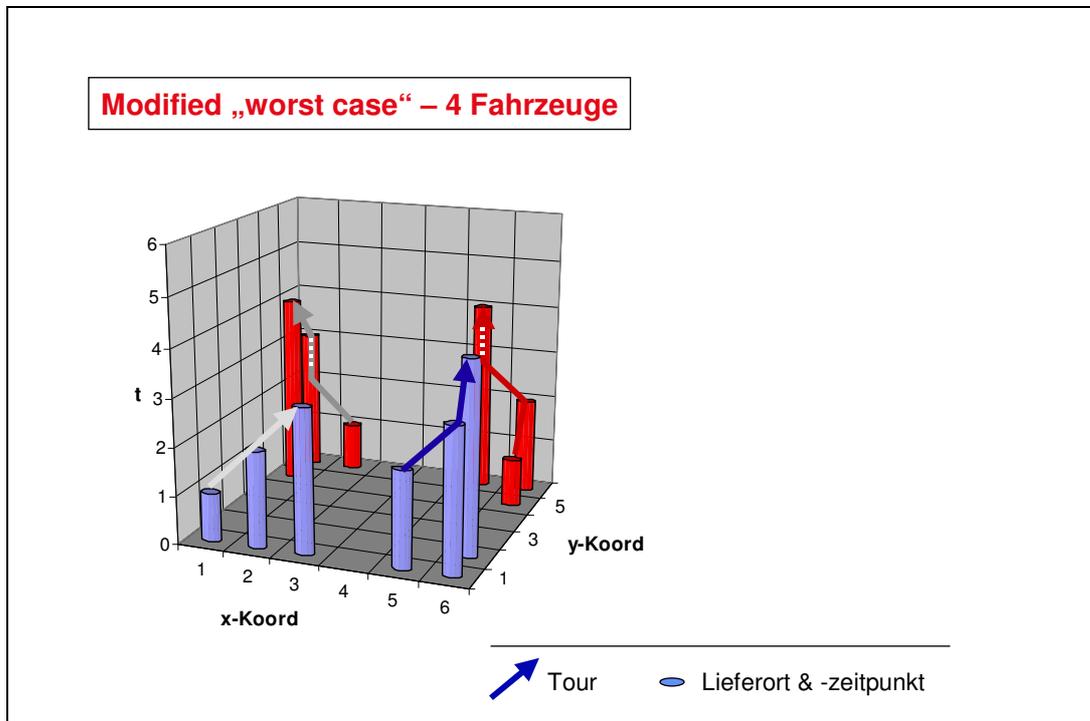


Abbildung 59: Ein zeitlich modifiziertes „worst case“ Szenario für das Modell A

Das „worst case“ Szenario und das modifizierte „worst case“ Szenario haben also sowohl eine identische geografische Verteilung der Lieferpunkte als auch eine identische zeitliche Verteilung der Lieferpunkte¹⁰⁴. Trotzdem unterscheidet sich deren jeweils in Hinsicht auf die benötigte Anzahl Fahrzeuge bestmögliche Lösung signifikant voneinander.

5.3.8.3 Schlußfolgerung

Was bedeutet dies für die in dieser Arbeit untersuchte Aufgabenstellung?

Eine rein zeitliche oder rein geografische Betrachtung kann keinen Aufschluss darüber geben, ob ein „günstiger“ Tourenplan möglich ist, oder nicht. Vielmehr determiniert die *spezielle Kombination der zeitlich-geografischen Lage* der Besuchspunkte die bestmögliche Belieferungsalternative bei sehr engen Lieferzeitfenstern.

Im Modell wurde gezeigt, wie in Anhängigkeit von dieser Raum-zeitlichen Lage die

¹⁰⁴ Die Kombinationen vor Ort und Zeit sind in beiden Szenarien jedoch verschieden

Anzahl der benötigten Fahrzeuge variiert. Ist jedoch die Anzahl der Fahrzeuge ex ante beschränkt, und sind Verletzungen von einmal zugesagten Lieferzeitfenstern auszuschließen, so folgt unmittelbar, dass die Anzahl der bedienbaren Kunden von ihrer *speziellen zeitlich-geografischen Lage* abhängt.

Für die Lösung eines derartigen Problems müssen daher drei Kernpunkte erfüllt sein:

- ▶ Existenz einer zulässigen Lösung:
Die *zeitlich-geografische Lage* der Kunden muss ex ante eine servicevereinbarungsgemäße Bedienung erlauben (d.h. eine Lösung des Tourenplanungsproblems mit einer ex ante beschränkten Anzahl von Fahrzeugen muss überhaupt möglich sein).
- ▶ Transportkosteneffizienz der Lösung
Die *zeitlich-geografische Lage* der Kunden muss es ex ante erlauben, die Belieferung der Kunden mit betriebswirtschaftlich vertretbaren Mitteln durchzuführen.
- ▶ Algorithmisches Vermögen:
Die eingesetzten Verfahren müssen diese zulässige und effiziente Lösung auch *finden* können.

5.3.9 Überlegungen zur Tourgebietstreue

Wie im Kapitel 3 bereits diskutiert ist in der Praxis die Tourgebietstreue (TGTR) ein in mehrerlei Hinsicht gewünschter Aspekt (Fahrer-Kunden-Bindung, Ortskenntnis der Fahrer, Warenbereitstellung).

Der Aspekt der TGTR ist jedoch nur schwer als Nebenbedingung oder Zielfunktionsterm zu modellieren, und findet daher in der Literatur zu Optimierungsproblemen wenig Beachtung. Er ist für den logistischen Anwendungsfall der Lebensmittelheimbelieferung jedoch entscheidend, und wurde daher bei der Verfahrenskonstruktion für den ORA-Ansatz berücksichtigt.

5.4 Das ORA - Verfahren

Aufbauend auf den oben geschilderten Überlegungen wurde das ORA-Verfahren entworfen. Das Verfahren besteht aus zwei wesentlichen Komponenten. Zum einen einer Vorplanungsmethodik zur Erzeugung von Tour-Templates und zum andern aus einem sukzessiven Verfahren zur Entscheidung über Auftragsannahme bzw. -ablehnung und zur Tourenplanung.

Im Vorplanungsschritt werden geografische Tourgebiete für definierte Zeiträume

geplant. Die Sequenz dieser Tourgebiete sei im Folgenden „Tourtemplate“¹⁰⁵ genannt. Kommen im Lauf des Verfahrens Anfragen zur Bedienung eines Lieferpunktes (x,y) zur Zeit t an das System, so entscheidet das Verfahren ob, wo und in welches der durch die jeweiligen Tourgebiete aufgespannten Tour-Templates der Lieferpunkt (x,y,t) eingefügt wird. Das Ergebnis dieses Einfügens ist dann eine Tour (s. Grafik).

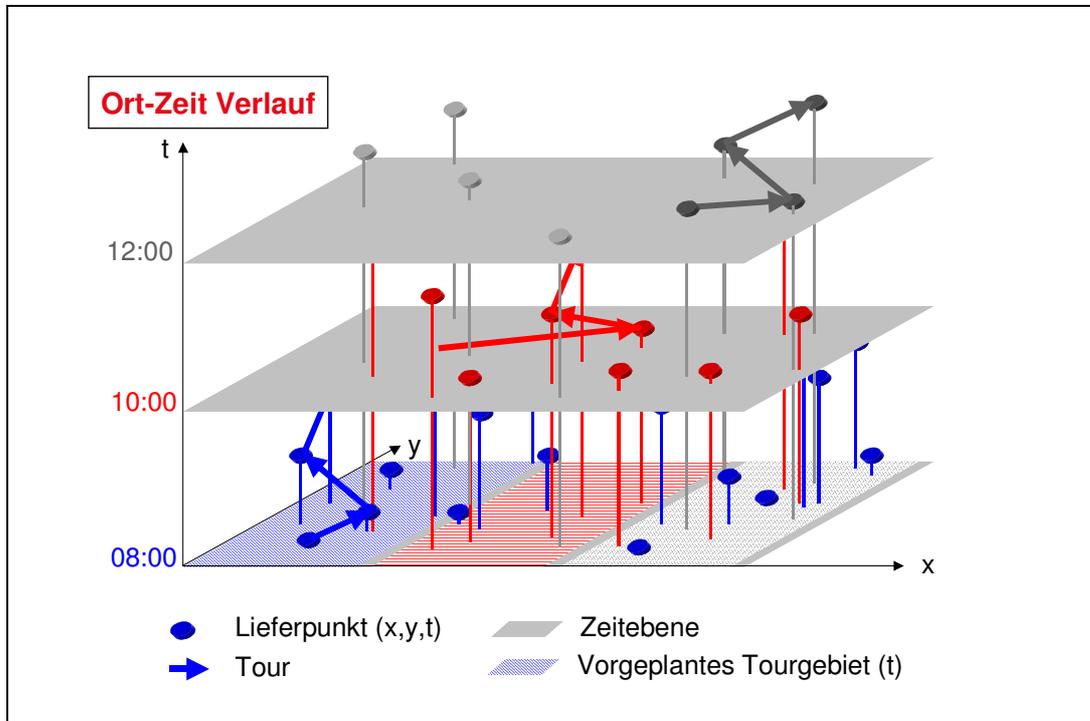


Abbildung 60: Prinzipskizze von ORA –in der Grafik sind die Lieferzeitwünsche t , und nicht die Bestellzeitpunkte dargestellt.

Das gesamte Liefergebiet wird durch mehrere Tour-Templates abgedeckt. Die Anzahl der Tour-Templates entspricht der Anzahl der ex ante geplanten Touren. Ein Fahrzeug kann ggf. auch mehrere Touren hintereinander bedienen. Ein geografisches Gebiet kann in mehreren Templates vorkommen, so dass ein Gebiet zu verschiedenen Tageszeiten mehrfach bedient werden kann.

5.4.1 Vorplanung

Die Aufgabe der Vorplanung ist die Festlegung von Tourgebieten und die Generierung der jeweiligen Tourtemplates auf Basis der Tourgebiete.

Def.: Tourgebiet (Aggregat)

Ein Tourgebiet (Aggregat) sei ein geografisches Gebiet, in dem sich ein Fahrzeug für eine definierte Zeitdauer aufhalten soll. Ex ante ist die Dauer

¹⁰⁵ template (engl.) = Muster, Vorlage

dieses Aufenthalts definiert, nicht jedoch das Zeitfenster dieses Aufenthalts¹⁰⁶. Die Aufenthaltsdauer stellt ein Zeitpotenzial dar, das sowohl die späteren Bedienzeiten (Standzeiten beim Kunden) als auch die Fahrzeiten zwischen den Kundenbesuchen benötigt wird. Wird dieses Potenzial nicht ausgenutzt, entstehen unproduktive Wartezeiten.

Def.: Tourtemplate (Rahmentour)

Ein Tourtemplate sei die zeitliche Aneinanderreihung mehrerer Tourgebiete, die von einem Fahrzeug bedient werden sollen. Durch die Reihung werden die Aufenthalts-Zeiträume relativ zum Beginn des Tourbeginns determiniert. Für die Fahrzeiten zwischen zwei Tourgebieten werden die jeweiligen Fahrzeiten zwischen den Tourgebietsreferenzpunkten (z.B. Flächenschwerpunkt des Tourgebiets) zueinander berechnet, und bei der zeitlichen Reihung der Tourgebiete berücksichtigt.

Für das hier untersuchte ORA Verfahren wurde ein pragmatisches Verfahren zur Vorplanung (Tourtemplatedefinition) gewählt.

Als kleinste Tourgebiete wurden einzelne Postleitzahlengebiete verwendet. Als größte Tourgebiete wurden Cluster von bis zu 4 benachbarten Postleitzahlgebieten verwendet. Die Verwendung von Postleitzahlgebieten hat den Vorteil der einfachen Implementierung im logistischen Prozess, da jede Adresse über die Postleitzahl bereits eindeutig einem Tourgebiet zugeordnet werden kann¹⁰⁷. Außerdem sind Postleitzahlgebiete in der Regel auch hinreichend kompakt.

Das Prinzip eines Tourtemplates ist in der folgenden Grafik dargestellt.

¹⁰⁶ Das Zeitfenster wird erst dadurch festgelegt, dass das Aggregat einem Tourtemplate zugeordnet wird (Ergebnis der Vorplanung).

¹⁰⁷ Im Prinzip könnte ein Tourgebiet auch beliebig definiert werden, z.B. als Rechteck oder Polygon, und eine entsprechende Zuordnung der Kunden zum Tourgebiet könnte über die entsprechenden GPS-Koordinaten erfolgen.

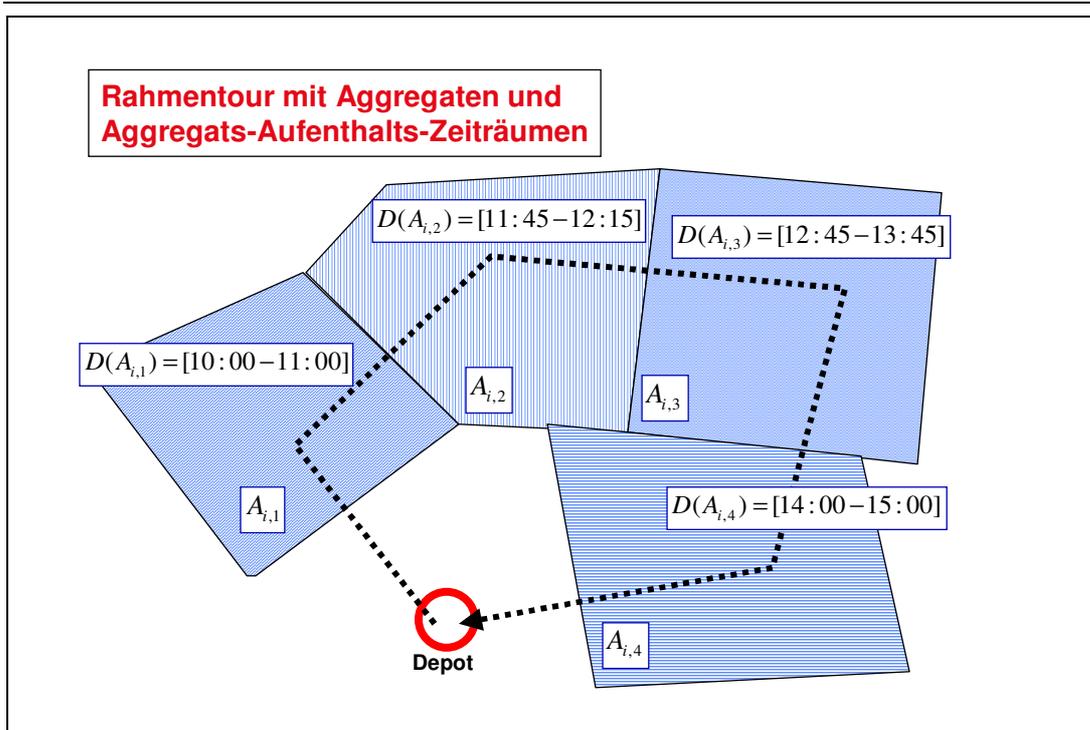


Abbildung 61: Tourtemplate (Rahmentour) und Tourgebiete (Aggregate).

Variablendefinition:

$A_{i,j}$: *Aggregat des Templates i an Stelle j der Aggregatfolge des Templates*

$D(A_{i,j})$: *Bedienzeitenfenster des Aggregats $A_{i,j}$*

O_r : *Auftrag (order) – der Index gibt die Auftragseingangsreihenfolge lexikografisch an (z.B. O_r wurde vor O_s eingebucht)*

Das Template der Tour i wird durch vier Aggregate $A_{i,1}$ bis $A_{i,4}$ aufgespannt. Die Bedienzeitenfenster der Aggregate $D(A_{i,j})$ ergeben sich aus den Fahrzeiten zwischen den Aggregaten und den Bedienzeitdauern der Vorgängeraggregate.

Die Tourtemplatedefinition kann im Prinzip beliebig erfolgen. In dieser Arbeit wurden die Aggregatsreferenzpunkte als Stops für ein klassisches VRSP betrachtet, und mit klassischen VRSP-Verfahren zu Tourtemplates kombiniert. Wenn aus Nachfragesicht „Präferenzzeiten“ für gewisse Tourgebiete gegeben wären (z.B. die Präferenz zur morgentlichen Belieferung eines bestimmten Wohngebiets), so könnte die Vorplanung als VRSP-TW modelliert und gelöst werden. In praxi hat sich jedoch gezeigt, dass die „Optimalität“ der Vorplanung gegenüber der „Nachfragegerechtigkeit“ in den Hintergrund tritt. Der Planer definiert daher i.d.R. zusammen mit dem Marketing die Tourtemplates auf Basis von Erfahrungs- bzw. Erwartungswerten in Hinblick auf die raum-zeitliche Nachfragestruktur¹⁰⁸.

Die folgende Grafik zeigt beispielhaft die Tourtemplatestruktur für ein Liefergebiet.

¹⁰⁸ Die Auswirkungen der Vorplanung auf die Ergebnisse werden in dieser Arbeit später noch quantitativ untersucht.

Beispiel Tourtemplates

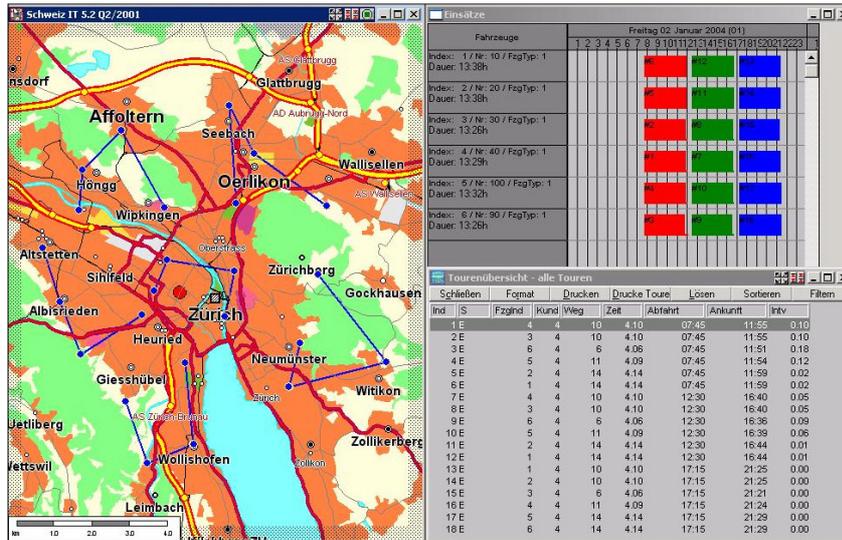


Abbildung 62: Tourtemplates – dargestellt als Folge von Aggregatsreferenzpunkten. Jede geografische Tour wird 3x täglich gefahren => 18 Tourtemplates.

5.4.2 Vorschlags- & Buchungsfunktion

Die Vorschlagsfunktion schlägt auf Basis der Tourtemplates und der bisher bestätigten Buchungen jeweils weitere mögliche Lieferzeitpunkte vor, aus denen der Kunde sich den ihm am besten gelegenen aussuchen kann. Der vom Kunden ausgewählte Liefertermin wird dem Kunden bei der Buchung bestätigt, und die entsprechende Lieferung wird dann fest in die Tour eingebucht.

5.4.2.1 Initiale Buchung

Auf Basis des Tourtemplates wird der erste Liefervorschlag für eine Tour generiert. Dazu wird das Aggregat gesucht, in das die Lieferadresse fällt. Der zeitliche Mittelpunkt des Aggregatsbedienzeitfensters wird dem Kunden als Lieferzeitvorschlag bekannt gegeben¹⁰⁹.

Wenn sich der Kunde für einen Lieferzeitpunkt entschieden hat, wird die Lieferung entsprechend „eingebucht“.

¹⁰⁹ Da eine Lieferadresse durch mehreren Tourtemplates bedient werden kann (z.B. täglich einmal morgens, einmal mittags und einmal abends), wird der Kunde in der Regel mehrere Lieferzeitvorschläge erhalten.

5.4.2.2 Weitere Buchungen

Zur einfacheren Verständlichkeit des Verfahrens soll an dieser Stelle der Begriff der „aktuellen Tour“ eingeführt werden.

Def.: aktuelle Tour

Unter der „aktuellen Tour“ ist die Tour zu verstehen, die sukzessive bis zum Zeitpunkt t auf Basis des entsprechenden Tourtemplates und der bestätigten Lieferungen generiert wurde.

Ist in einer „aktuellen Tour“ das Aggregat, in das eine neue Kundenanfrage fällt bisher nicht bebucht worden, so wird für den ersten Kunden im Aggregat i.d.R. (Ausnahmen s.u.) der zeitliche Aggregatsmittelpunkt als Liefertermin vorgeschlagen.

Ist in einer „aktuellen Tour“ das Aggregat, in das eine neue Kundenanfrage fällt bereits bebucht worden, so wird mit einem Insertion-Verfahren versucht, den Kunden bestmöglich (d.h. umwegminimal u.d.N. der Einhaltung aller bereits für die aktuelle Tour bestätigten Zeitfenster) eingefügt. Der resultierende Liefertermin wird dem Kunden vorgeschlagen. Ist es nicht möglich, die Kundenanfrage ohne Verletzung der zugesagten Zeitfenster in die aktuelle Tour einzufügen, so wird für das entsprechende Tourtemplate kein Liefervorschlag generiert.

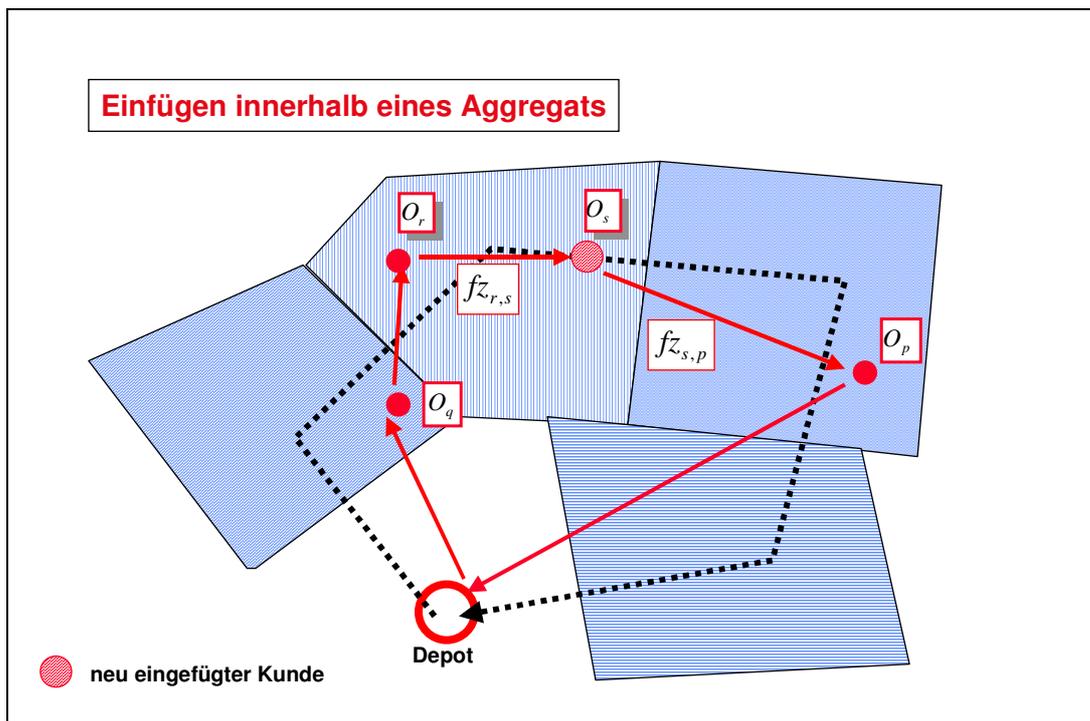


Abbildung 63: Einfügen des Kunden O_s in eine aktuelle Tour – geografische Darstellung.

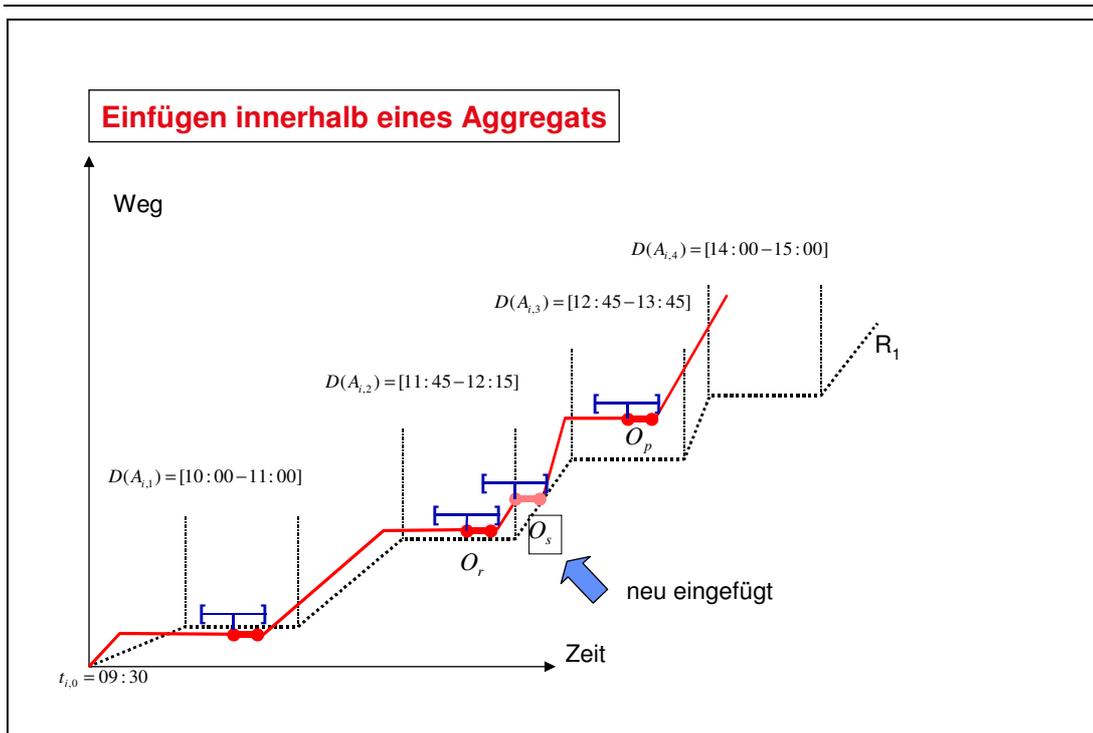


Abbildung 64: Einfügen des Kunden O_s in eine aktuelle Tour –Weg-Zeit-Diagramm.

Auch hier wird die Lieferung in die aktuelle Tour eingebucht, wenn der Kunde sich für den entsprechenden Liefertermin entschieden hat.

5.4.2.3 Einfügen „auf Budget“ von Nachbaraggregat

Da die Nachfrage für einige Aggregate stärker oder schwächer ausfallen kann als erwartet, erlaubt das Verfahren Buchungen „auf Budget des Nachbaraggregats“.

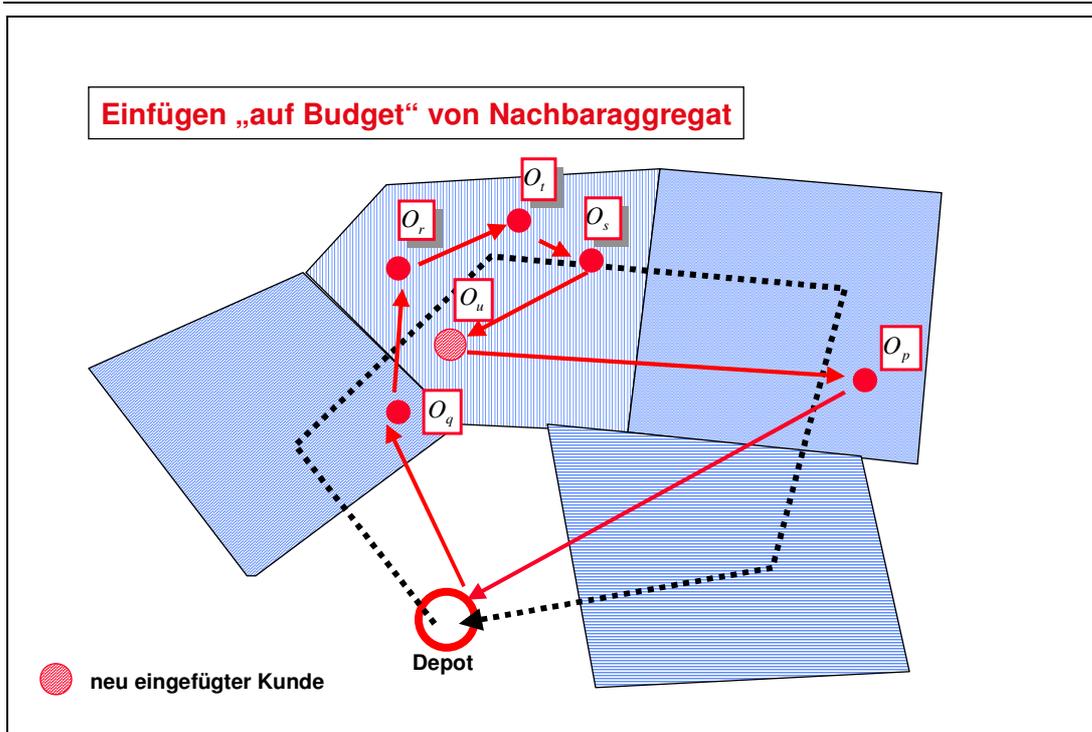


Abbildung 65: Einfügen eines Kunden in eine aktuelle Tour – geografische Darstellung

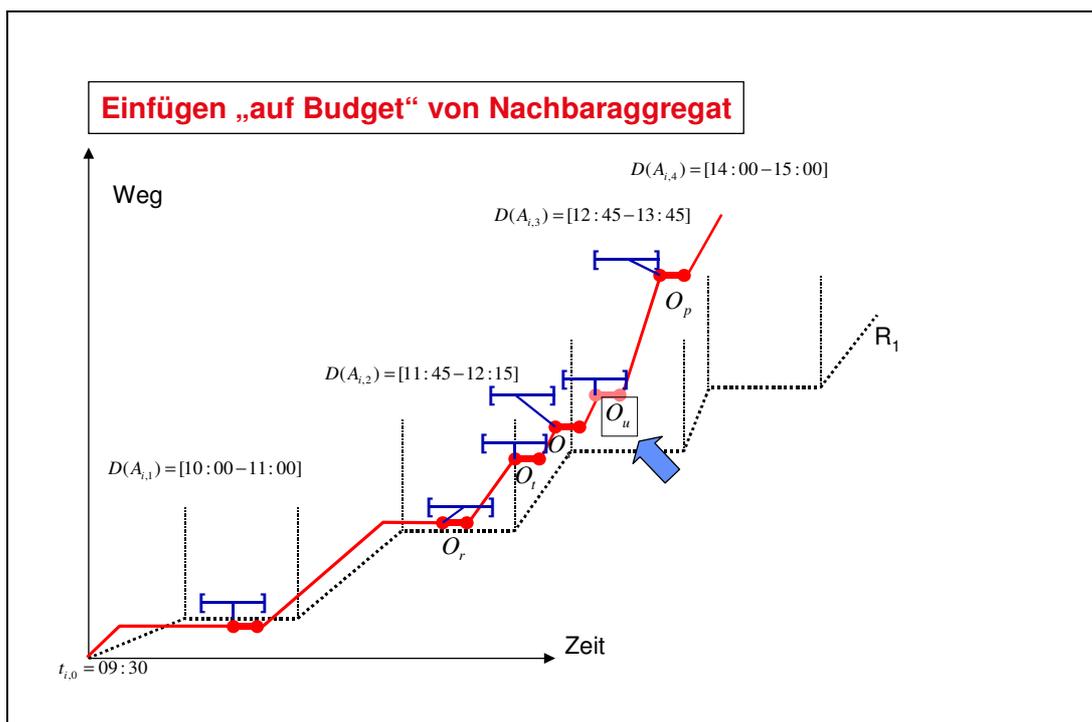


Abbildung 66: Einfügen „auf Budget des Nachbaraggreats“ – Weg-Zeit-Diagramm.

In der Weg-Zeit-Darstellung wird klar, dass das ursprüngliche Zeitbudget für die Buchung im zweiten Aggregat nicht ausreicht. Da zum Zeitpunkt der Buchungsanfrage im Nachbaraggreat noch kaum Buchungen vorliegen, und alle zugesagten Zeitfenster der aktuellen Tour eingehalten werden können, schöpft der Liefervorschlag für O_u aus dem Zeitbudget des Nachbaraggreats.

An dieser Stelle greift jetzt auch die oben angesprochene Ausnahme bei der ersten Buchung in einem Aggregat – wurde beispielsweise eine Buchung im Vorgängeraggregat so eingebucht, dass für die erste Buchung im Nachfolgeraggregat nicht mehr der zeitliche Aggregatsmittelpunkt als Liefertermin vorgeschlagen werden kann, so wird für die erste Buchung in einem Aggregat natürlich ein späterer Termin vorgeschlagen.

5.4.3 Ex ante Kritik des Verfahrens

Wie jede Heuristik hat das ORA-Verfahren verschiedene intrinsische Eigenschaften, die an dieser Stelle kurz kritisch beleuchtet werden sollen.

5.4.3.1 Nachteile

Die Kerneigenschaft des Verfahrens ist zugleich dessen größter intrinsischer Nachteil, nämlich die Abhängigkeit von der Vorplanung, insbesondere

- ▶ von der Gesamtheit des erwarteten Auftragspotenzials
- ▶ von den Annahmen über die raum-zeitliche Nachfragestruktur, und
- ▶ von der Granularität der Vorplanung (wenige große Aggregate vs. viele kleine Aggregate)

Die Ergebnisqualität des Verfahrens ist somit stark determiniert durch die „Güte“ der Vorplanung.

5.4.3.2 Vorteile

Demgegenüber stehen die folgenden augenscheinlichen Vorteile des Verfahrens.

Das Verfahren ist sehr einfach und damit transparent. Es ist daher leicht nachzuvollziehen warum das Verfahren welche Ergebnisse generiert, und leicht vorherzusagen, wie sich das Verfahren unter gewissen Rahmenbedingungen verhalten wird.

Das Verfahren ist außerdem in gewissem Maße robust – wenn die Tourtemplates nicht der aktuellen Nachfragestruktur entsprechen, so kann dieser „Prognosefehler“ durch die Nutzung von „Zeitbudgets von Nachbaraggregaten“ zumindest teilweise „geheilt“ werden.

Das Verfahren ist schließlich leicht adaptierbar – wenn die Tourtemplates nicht der aktuellen Nachfragestruktur entsprechen, so können entsprechende Erkenntnisse über die raum-zeitliche Nachfragestruktur leicht genutzt werden, indem die

Tourtemplates der Vorplanung entsprechend angepasst werden.

5.5 Das FPA-Verfahren

Das FPA - Verfahren soll an dieser Stelle kurz beschrieben werden, da es zum einen in der Praxis Einsatz findet, zum anderen in dieser Arbeit die Basis für den „ceteris paribus“-Teilaspekt-Vergleich in Bezug auf die ORA-Ergebnisse liefert.

Die Kernidee des FPA-Verfahrens ist die Rückführung des D-m-VRSP-TW-UC-DO auf ein klassisches VRSP-TW mit Hilfe einer vorgeschalteten Filterfunktion.

Def.: Filterfunktion

Eine Filterfunktion ist eine Funktion, die entscheidet, ob eine Lieferanfrage angenommen oder abgelehnt wird. Das Ziel der Filterfunktion ist zum einen die Begrenzung der „Gleichzeitigkeit“ der Bestellungen, zum anderen die Begrenzung der zugesagten Lieferungen auf eine Anzahl, die mit den verfügbaren Fahrzeugen ohne große Verletzung der zugesagten Lieferzeitrestriktionen transportlogistisch erfüllt werden kann.

Eine solche Filterfunktion kann beispielsweise die „maximale Anzahl zugesagter Lieferungen (=Kontingent) für eine Tageszeitstunde¹¹⁰“ sein. Wie groß diese maximale Anzahl sein kann, ist Abhängig von der Anzahl verfügbarer Fahrzeuge. Die maximale Anzahl akzeptierter Aufträge ist i.d.R. ein „Erfahrungswert“, der sich über die Zeit einpendelt.

Ist eine derartige Filterfunktion im Bestellportal eines online-shops integriert, so kann die Bestellung losgelöst von der Tourenplanung erfolgen. Die Lieferzeitbestätigung erfolgt „blind“ (z.B. auf Basis der Lieferkontingente pro Tageszeitstunde). Nach Auftragsannahmeschluss werden alle Lieferungen als Input in ein klassisches Tourenplanungssystem übergeben, das dann da entsprechende VRSP-TW zu lösen versucht.

5.5.1 Nachteile

Dieses Vorgehen kann nicht garantieren, dass es für das VRSP-TW bei beschränkter Fahrzeuganzahl eine gültige Lösung gibt. Dies gilt umso mehr, je präziser die Lieferterminavisierung erfolgt. Daher wird das VRSP-TW Problem i.d.R. bei Bedarf durch Aufweitung der bereits zugesagten (und daher eigentlich nicht mehr revidierbaren) Lieferzeitfenster relaxiert, bis eine Lösung des relaxierten VRSP-TW gefunden ist. Die Überschreitung der Lieferzeitfenster (vgl. die im Vorlaufenden definierte „Präzision“) ist der Preis in Hinblick auf Lieferserviceglaubwürdigkeit, der bei Einsatz dieses einfachen Verfahrens zu

¹¹⁰ Diese einfachste Filterfunktion findet in der Praxis Anwendung im Lebensmittelheimlieferservice

zahlen ist.

In Praxi ist der Balanceakt zwischen „zu vielen zugesagten Lieferungen“ (kurzfristige Deckungsbeitragssteigerung einhergehend mit mittelfristigem Verlust des Kundenvertrauens in den zugesagten Lieferservice¹¹¹) und „zu wenig zugesagten Lieferungen“ (unnötige Beschneidung des möglichen Ertrages für Garantie des zugesagten Liefertermine) erfolgsentscheidend. Im Prinzip gilt: je enger die zugesagten Lieferzeitfenster sind (hohe Präzision), desto schwieriger ist diese Balance zu halten.

5.5.2 Vorteile

Die Vorteile dieses Verfahrens sind zum einen die einfache Integrierbarkeit in ein Bestellsystem (es sind keine online-Schnittstellen zwischen Shop und Planungssystem zu bedienen).

Zum anderen kann dieser Ansatz bei einer geringen „Präzision“ (= große Lieferzeitfenster) recht gute Ergebnisse liefern, da gängige Systeme zur Lösung von VRSP-TW-Problemen bei weiten Zeitfenstern die Transportkosten gut „minimieren“ können.

Schließlich ist das Verfahren flexibel in Hinblick auf kurzfristigen Nachfragerückgang. Bei kurzfristig unerwartet niedriger Nachfrage (z.B. Bestellrückgang bei Ferienbeginn) kann angebotsseitig einfach durch Nichtnutzung von Fahrzeugkapazitäten reagiert werden (kurzfristige Stilllegung von Fahrzeugen bei Beibehaltung einer guten Tourauslastung)

5.6 Zusammenfassung

In diesem Kapitel wurde das Problem der Tourenplanung für den eCommerce-Lebensmittelheimzustellservice zunächst klassifiziert (neue Problemklasse: D-m-VRSP-TW-UC-DO) und dann formal für drei in der Praxis eingesetzte Ansätze (ORA, FPA/PPA und FPA/REA) modelliert. Die Modellierung für ORA stellt dabei ein neues Optimierungsproblem auf, die Modellierung für FPA/PPA und FPA/REA sind klassische Probleme. Jede Modellierung beruht dabei auf Anforderungen aus der Praxis. Das Kernproblem des neuen D-m-VRSP-TW-UC-DO ist die Unkenntnis über die jeweils noch zu erwartenden Aufträge, wenn bei der Bestellung enge Lieferzeitpunkte zugesagt werden müssen.

Anschließend wurden im wesentlichen zwei grundverschiedene Möglichkeiten untersucht, mit diesem Problem umzugehen, nämlich zu warten, bis eine Obergrenze von n Aufträge eingegangen ist, und dann optimieren, (Filter-Ansatz

¹¹¹ Einhaltung des bestätigten der Liefertermins

FPA/PPA und FPA/REA) bzw. den Tourenplan sukzessive aufbauen (ORA). Beide Lösungsansätze finden sich in der betrieblichen Praxis im Einsatz.

Dann wurden Überlegungen zu verschiedenen Aspekten des Problems angestellt, und mit kleinen Laborbeispielen veranschaulicht. Die wesentlichen neuen Aspekte sind der Nutzen der Vermeidung von Gleichzeitigkeit und den Nutzen von impliziten Tourstrukturen.

Aus diesen Überlegungen heraus wurde dann der Lösungsansatz ORA abgeleitet und erklärt. Abschließend wurde der alternative Planungsansatz FPA erklärt.

6 Messungen und Ergebnisse

Im Fokus dieses Kapitels steht der **Einfluss des zugesagten Lieferzeitfensters (PRÄZ) auf die Transportkosteneffizienz (TKEF)**, die sich je nach Untersuchungskontext in den Kenngrößen Gesamttourkilometer, Anzahl benötigter Fahrzeuge, Durchschnittliche Anzahl Kunden pro Fahrzeug, etc. widerspiegelt.

In diesem Kapitel werden **zwei Ansätze** näher untersucht, die sowohl die Nachfrage steuern, als auch die zugesagten Lieferzeitfenster als Restriktionen einhalten:

- ▶ **ORA** „Online Tourenplanung“ Ansatz (Steuerung des Nachfragevolumens)
- ▶ **FPA/PPA** „Filter vor Planung mit Planung post Avis“ Ansatz (Anpassung des Angebots durch erweiterte Fahrzeugflotte)

Beide Ansätze sind in der betrieblichen Praxis zu finden. Der dritte Ansatz

- ▶ **FPA/REA** „Filter vor Planung mit reduziertem Erfüllungsgrad“ Ansatz (Anpassung des Angebots durch Verringerung des Service)

kommt in der betrieblichen Praxis höchstens in Notfällen zum Einsatz, wenn mit FPA/PPA keine zulässige Lösung gefunden werden kann. Er wird daher auch nur kurz angerissen.

Nicht näher untersucht werden der

- ▶ **PAA** „Planung ante Avis“
Dieser Ansatz erlaubt definitionsgemäß keine präzise Lieferzeitzusage zum Bestellzeitpunkt, da die Lieferzeitzusage erst nach der Planung, und diese erst nach vorliegen aller Lieferaufträge erfolgt)
- ▶ **RTA** „Rahmentour Ansatz“,
Der RTA wird nicht näher untersucht, da in praxi ein rein geografisch orientierter Ansatz ohne Berücksichtigung der zeitlichen Dimension bei engen und präzise einzuhaltenden Zeitfenstern bei stochastischer Nachfrage wenig erfolgversprechend ist.

Die quantitative Analyse der beiden untersuchten Ansätze erfolgt **zunächst disjunkt** voneinander. Zur Analyse des PPA Ansatzes wurden eigene Messungen durchgeführt, und dann anhand von veröffentlichten Ergebnissen aus der Literatur validiert. Die Ergebnisse zum ORA-Ansatz wurden vollständig selbst generiert.

Abschließend werden Teilaspekte der Ergebnisse der Ansätze FPA/PAA und ORA soweit möglich miteinander verglichen. Die **Vergleichbarkeit der Ergebnisse** ist ein nichttriviales Problem, denn die verschiedenen Ansätze verfolgen die Lösung von unterschiedlichen OR-Planungsproblemen, wie die folgende Grafik zeigt.

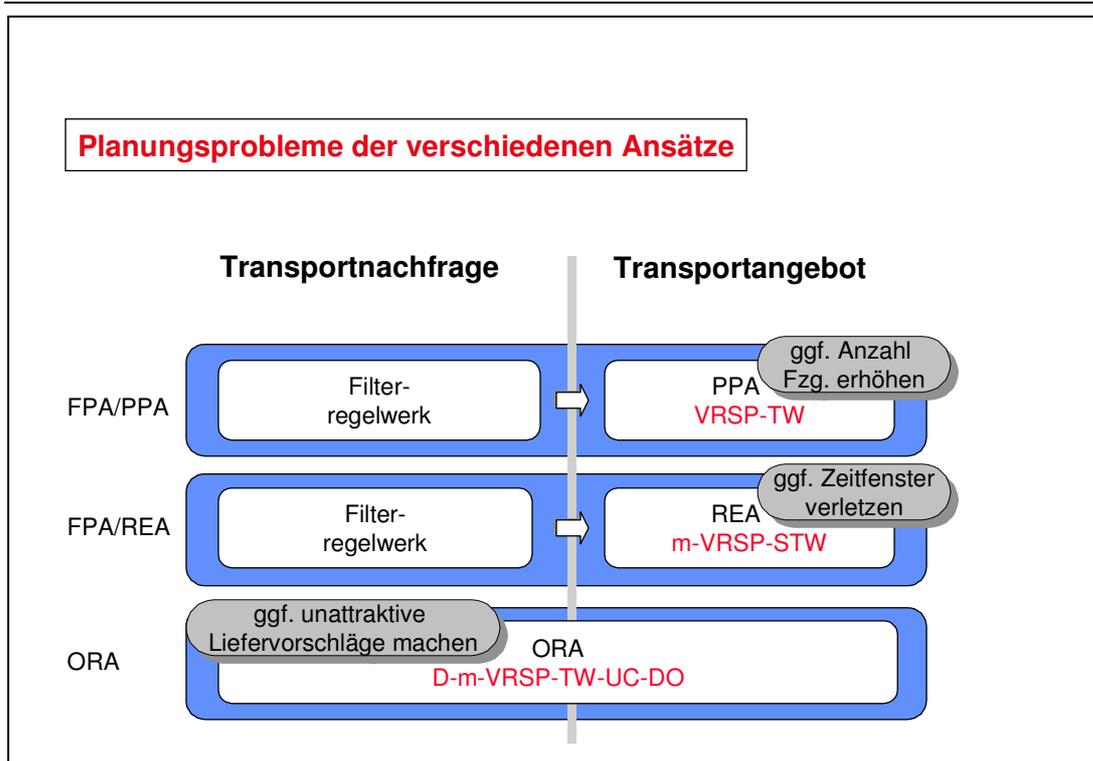


Abbildung 67: OR-Planungsprobleme (rot) der verschiedenen Planungsansätze mit jeweiligem Konfliktlösungsmechanismus bei engen Ressourcen (grau)

Während der ORA-Ansatz in einem einzigen Verfahren Kunden annimmt bzw. ablehnt und Touren plant, erfolgt die Annahme bzw. Ablehnung von Kundenanfragen im FPA-Ansatz über den vorgelagerten Filter anhand eines beliebigen Regelwerks, und die Tourenplanung als getrennter nachgelagerter Schritt.

Die Filterfunktion des FPA wird in dieser Arbeit nicht näher betrachtet, denn in der betrieblichen Praxis ist das Einstellen des Filters „Erfahrungssache“, d.h. es ein iterativer Lernprozess¹¹², der die Einflussgrößen des raum-zeitlichen Nachfrageverhaltens und das verfügbare Transportleistungsangebot berücksichtigen muss. Die Ergebnisqualität des FPA/PPA hängt jedoch entscheidend von einer guten Einstellung des Filters ab.

Wie können dann aber ORA & FPA sinnvoll quantitativ miteinander verglichen werden?

Ein direkter und vollständiger Vergleich zwischen ORA und FPA wird in dieser Arbeit aus den oben genannten Gründen nicht angestellt. Stattdessen wurde der methodische Ansatz des „ceteris paribus“-Teilaspekt-Vergleichs gewählt, der zumindest Aussagen über die Transportkosteneffizienz beider Ansätze ermöglicht.

¹¹² Die Anzahl der Aufträge, die der Filter akzeptieren soll wird i.d.R. zuerst mit Blick auf die verfügbaren Fahrzeuge und die im Serviceangebot zugesagte PRÄZ geschätzt, dann in der Anlaufphase etwas vorsichtiger eingestellt, und dann im laufenden Betrieb langsam sukzessive erhöht, bis entweder eine befriedigende Fahrzeugauslastung erreicht wird (PPA), oder die Anzahl und Summe der Zeitfensterverletzungen unakzeptabel wird (REA).

Um vergleichbare Ergebnisse zu erzielen, wurde daher die systemimmanente Filterfunktion von ORA als Filter für FPA genutzt, um den Input für die zweite Stufe des FPA zu generieren.

Bevor auf die einzelnen Messungen eingegangen wird, werden in drei Exkursen die wesentlichen Grundlagen für die quantitativen Untersuchungen eingeführt.

6.1 Exkurs 1 – Instanzen

Die folgenden Ergebnisse basieren auf sogenannten „Instanzen“. Daher vorab eine kurze Definition des Instanz-Begriffs für diese Arbeit:

Definition

Eine „**Instanz**“ sei ein definierter Satz von **Lieferaufträgen** (die sich durch ihre räumliche und zeitliche Lage und die jeweilige Liefermenge¹¹³ auszeichnen) und von **Restriktionen** (insbesondere einer Vorschrift zur Entfernungsberechnung).

Eine Instanz beschreibt damit ein **Tourenplanungsproblem**, das von verschiedenen **Verfahren** (s.u.) gelöst werden kann.

Definition

Eine „**Instanzenklasse**“ sei ein definierter Satz von **Variationen einer Instanz**, die sich weder in ihrer räumlichen, noch in ihrer absoluten zeitlichen Lage unterscheiden, wohl aber in ihrer Zeitfensterbreite. Eine Instanzenklasse wird durch das Suffix „#“ gekennzeichnet

Definition

Eine „**Instanzengruppe**“ sei ein definierter Satz von **ähnlichen Instanzen** die sich nicht strukturell¹¹⁴ in ihrer räumlichen und in ihrer zeitliche Lage unterscheiden. Eine Instanzengruppe wird durch das Suffix „*“ gekennzeichnet

6.1.1 Die Solomon Instanzen

[Solomon, 1984] hat eine Reihe von Referenzdatensätzen (Instanzen) eingeführt, an der sich bis heute die wesentlichen Verfahren zum VRSP-TW messen. Eine detaillierte Übersicht über die Solomoninstanzen ist im Anhang zusammengestellt. Die folgende Grafik zeigt die raum-zeitliche Verteilung der Lieferpunkte für einige ausgewählte Solomoninstanzen.

¹¹³ Die Liefermenge wird in dieser Arbeit meist vernachlässigt, da in der betrieblichen Praxis i.d.R. die Liefermenge (im Gegensatz zu den Lieferzeitfenstern) keine relevante Restriktion darstellt.

¹¹⁴ Strukturelle Unterschiede wären z.B. signifikant unterschiedliche Cluster-Strukturen in zwei Instanzen

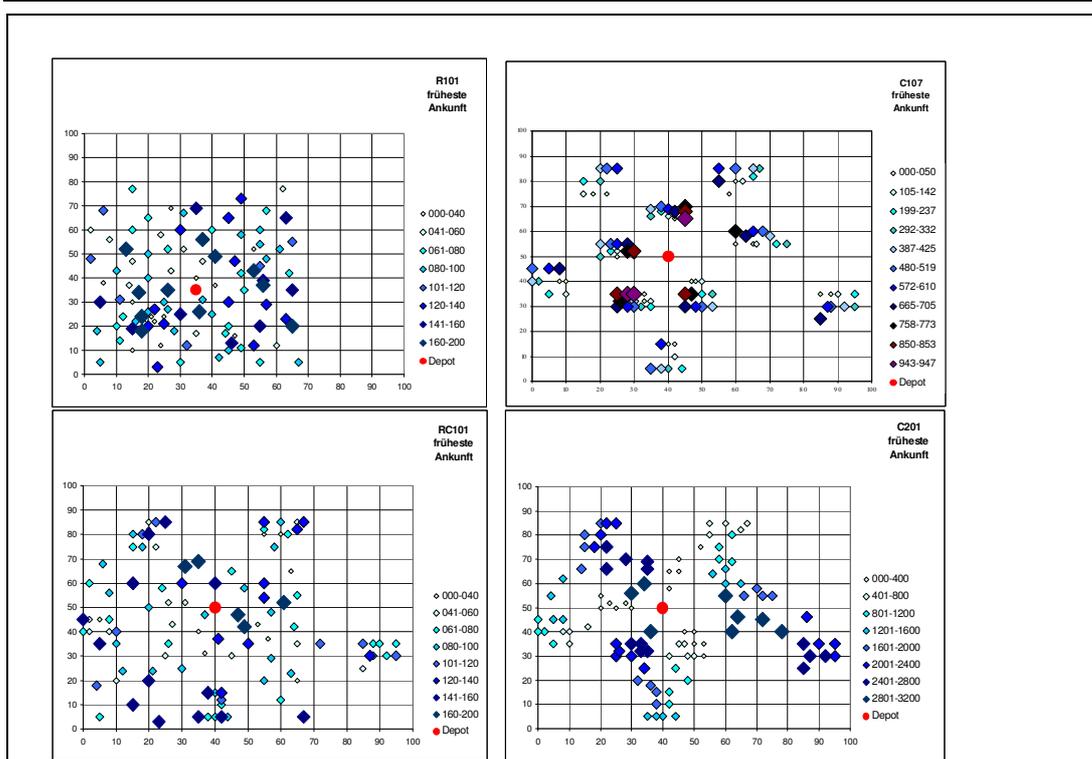


Abbildung 68: Raum-zeitliche Lage der Lieferpunkte in ausgewählten Solomoninstanzen

Die geografische Lage wird durch die Lage im (x,y)-Koordinatensystem dargestellt. Die zeitliche Lage wird durch Farbe und Form der Lieferpunkte dargestellt (Referenzwert ist die frühest mögliche Ankunft beim Kunden gemessen in Zeiteinheiten).

6.1.2 Erweiterung der Solomon Instanzen

In dieser Arbeit wurden ausgewählte Solomon-Instanzen hinsichtlich ihrer Zeitfensterbreite modifiziert. Diese Erweiterungen basieren jeweils auf einer Solomoninstanz, von der ausgehend die Zeitfenster sukzessive verringert bzw. verbreitert wurden. Diese erweiterten Instanzen wurden durch das Zeichen „#“ gekennzeichnet (z.B. „R205#“).

6.1.3 Die Instanzen ZUE500*

Die Instanzen ZUE500e, ZUE500f und ZUE500g, wurde für diese Arbeit generiert. Sie umfassen jeweils 500 Stops im Großraum Zürich. Details zu diesen Instanzen sind im Anhang aufgeführt. Die folgende Grafik zeigt die raum-zeitliche Verteilung der Lieferpunkte für die Instanz ZUE500e.

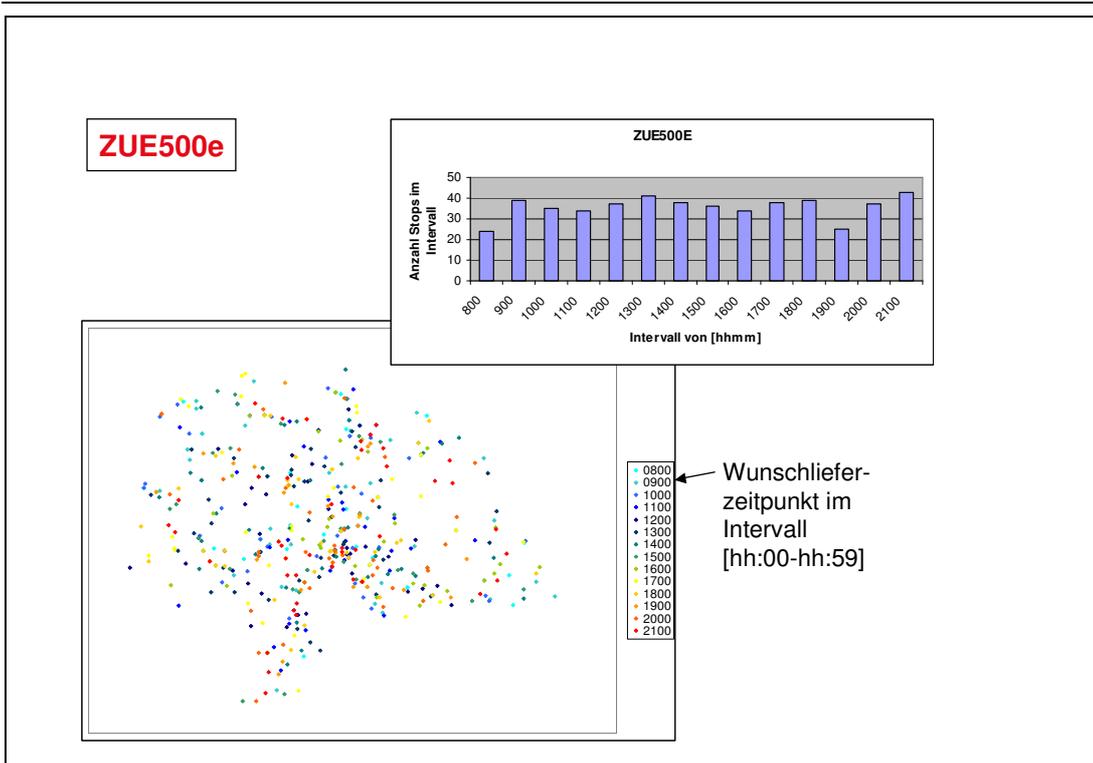


Abbildung 69: Raum-zeitliche Lage der Lieferpunkte in der Instanz ZUE500e
 Farbe zeigt zeitliche Lage im Lieferzeitpunktintervall [hh:00-hh:59]

6.2 Exkurs 2 – Verfahren

In den folgenden Untersuchungen zum PAA kommen verschiedene heuristische Verfahren zum Einsatz, die hier kurz eingeführt werden.

Definition

Ein „Verfahren“ sei ein Algorithmus zur Erzeugung einer gültigen Lösung für ein Planungsproblem. Der Input für ein Verfahren ist eine Instanz.

6.2.1 HSav0

```
proc HSav0
begin
  Plane mit klassischem Savings und Parametertupel  $P^i=(\alpha^i, \beta^i, \max$ 
  Wartezeit/Stop) ohne Einschränkung der Anzahl der verfügbaren
  Fahrzeuge
  Plane Fahrzeugeinsätze (Nachoptimierung)
  Gebe das so erzielte Ergebnis aus
end
```

Abbildung 70: Algorithmus „HSav0“

Dieses Verfahren beinhaltet als Savings-basierte Greedy-Heuristik keinen Mechanismus, der die Anzahl Fahrzeuge minimiert.

6.2.2 HSav1

Wie im Anhang dargestellt¹¹⁵, hat der klassische Savings (ohne Begrenzung der Fahrzeuganzahl) den Nachteil, dass durch die Tourbildung ggf. die Anzahl benötigter Fahrzeuge durch eine nachgeschaltete Fahrzeugeinsatzplanung nicht immer sinnvoll minimiert werden kann.

Um diesem Phänomen zu begegnen wurde eine pragmatische Heuristik „HSav1“ entwickelt, deren Kern das Savingsverfahren ist:

¹¹⁵ Ergebnisdetails der Untersuchung zur Instanz ZUE500E / Entwicklung der Fahrzeuganzahl

Algorithmus „HSav1“

```
proc HSav1
begin
  Setze Anzahl Fahrzeuge m:=1
  repeat
    for i=1..k
      Plane mit klassischem Savings und Parametertupel
       $P^i=(\alpha^i, \beta^i, \text{max Wartezeit } tw^i/\text{Stop})$  unter Einschränkung
      der Anzahl der verfügbaren Fahrzeuge auf m
      Speichere Ergebnis in Liste
    next i
  if
  (Liste enthält nur Ergebnisse, die unverplante Stops aufweisen)
  then
    Lösche Liste
    m:=m+1
  else
    Selektiere das Ergebnis aus Liste, bei dem am wenigsten
    km generiert wurden
    Plane Fahrzeugeinsätze (Nachoptimierung)
    Gebe das so erzielte Ergebnis aus
  ABBRUCH
until ABBRUCH
end
```

Abbildung 71: Algorithmus „HSav1“

Die Heuristik versucht also bei begrenzter Anzahl verfügbarer Fahrzeuge m mit Hilfe des Savings und verschiedenen Parametereinstellungen P^i alle Stops zu verplanen. Ist dies nach Durchprobieren von k Parametertupeln nicht gelungen, so wird die Fahrzeuganzahl um 1 erhöht.

6.2.3 Hlns1

Hlns1 baut im Kern auf einem Insertionsverfahren auf.

```
proc Hlns1
begin
  Setze Anzahl Fahrzeuge m:=1
  repeat
    for i=1..k
      Definiere Seed-Stop-Tupel  $(x^i_1, \dots, x^i_m)$ 
      Plane Stichtouren zu den m Seed Stops
      Füge restliche Stops mit klassischem Insertions-
      Verfahren in Liste ein
      Speichere Ergebnis in Liste
    next i
  if
    (Liste enthält nur Ergebnisse, die unverplante Stops aufweisen)
  then
    Lösche Liste
    m:=m+1
  else
    Selektiere das Ergebnis aus Liste, bei dem am wenigsten
    km generiert wurden, und gebe es als heuristisches
    Ergebnis aus
  ABBRUCH
until ABBRUCH
end
```

Abbildung 72: Algorithmus „Hlns1“

Ähnlich zu HSav1 wird hier die Fahrzeuganzahl m sukzessive erhöht, bis alle Kunden verplant sind.

6.2.4 HSav1Ins2

```
proc HSav1Ins2
begin
  Setze Anzahl Fahrzeuge m:=1
  repeat
    for i=1..k
      Plane mit klassischem Savings und Parametertupel
       $P^i=(\alpha^i, \beta^i, \max \text{Wartezeit}^i/\text{Stop})$  unter Einschränkung der
      Anzahl der verfügbaren Fahrzeuge auf n
      for j=1..t
        Optimiere Reihenfolge mit 3-opt-Verfahren
        Füge möglicherweise übriggebliebene Stops mit
        Insertionverfahren ein
      next j
      Speichere Ergebnis in Liste
    next i
    if
      (Liste enthält nur Ergebnisse, die unverplante Stops aufweisen)
    then
      Lösche Liste
      m:=m+1
    else
      Selektiere das Ergebnis aus Liste, bei dem am wenigsten
      km generiert wurden, und gebe es als heuristisches
      Ergebnis aus
      ABBRUCH
    until ABBRUCH
  end
```

Abbildung 73: Algorithmus „HSav1Ins2“

Im Vergleich zu „HSav1“ nutzt „HSav1Ins2“ zusätzlich das Insertion-Verfahren in t Nachoptimierungsschritten, falls nur „wenige“ Stops beim Savings übrigbleiben, bevor es die Fahrzeuganzahl um eins erhöht.

6.3 Exkurs 3 – Wichtungsproblem

An dieser Stelle soll vorab noch eine wichtige Frage betrachtet werden: Sind die Anzahl benötigter Fahrzeuge und die Gesamttourkilometer irgendwie voneinander abhängig?

Diese Frage ist deshalb wichtig, weil aus einer **logistisch-ökonomischen Sicht** ein Gesamtkostenoptimum (also ein Kostenminimum aus fixen und variablen Kosten) anzustreben ist. Die Anzahl Fahrzeuge ist bei eigenem Fuhrpark ein **fixer** Kostenfaktor (Fzg-Afa), während die Tourkilometer (und damit die Benzinkosten und die Fahrerstundenkosten) ein **variabler** Kostenfaktor sind.

Wenn die beiden Kostenfaktoren nicht unabhängig voneinander sind (was wahrscheinlich ist) – wie kann eine Bewertung der beiden Kostenfaktoren „Anzahl Fahrzeuge (=>Fixkosten) und „Anzahl Tourkilometer“ (=>variable Kosten) dann sinnvoll erfolgen?

Eine gemeinsame Bewertungsfunktion muss beide Kostengrößen gegeneinander sinnvoll gewichten, obwohl sie jeweils andere zeitliche Dimensionen haben:

Beispiel:

$$\min \text{Gesamtkosten/Tag} = \min (\text{VarKosten}(l_{\text{ges}})/\text{Tag} + \lambda \text{FixKosten}(m)/\text{Jahr}),$$

mit

λ [Jahr/Tag]: Gewichtungsfaktor

m : Anzahl Fahrzeuge

l_{ges} : Anzahl Gesamttourkilometer

Eine generalisierte kontextunabhängige Gewichtung (d.h. unabhängig vom logistischen Anwendungsfall und den spezifischen Kostenfunktionen) ist nur schwer zu treffen, wie die folgenden Beispiele zeigen.

Beispiel1:

Fuhrpark (fix)	Personal (fix)	Km (var)
Eigener Fuhrpark mit 3-Temperaturzonen-Kühl-LKW	Festangestellter Fahrerstamm; Einschichtbetrieb	Spritverbrauch: 14l / 100 km
Anschaffungswert 50.000 € / Stück Afa über 3 Jahre	Lohnkosten 50.000 € p.a. / Fahrer	Benzinkosten: 1,25 € / l
250 Betriebstage p.a	250 Betriebstage p.a	Wartungsumlage: 0,10€ / km
Fixkostenanteil 66,7€/Tag	Fixkostenanteil 200€/Tag	Variable Kosten: 0,275 € / km
Fixkosten: 266,7 € / Fzg		
=> 1 eingespartes Fzg ist äquivalent zu eingesparten		969,8 km / Tag

Tabelle 8: Wichtungsproblem – Beispiel 1

Beispiel2:

Fuhrpark (fix)	Personal (fix)	Km (var)
Eigener PKW Fuhrpark	Festangestellter Fahrerstamm; Einschichtbetrieb	Spritverbrauch: 8 l / 100 km
Anschaffungswert 21.000 € / Stück Afa über 3 Jahre	Lohnkosten 40.000 € p.a. / Fahrer	Benzinkosten: 1,25 € / l
250 Betriebstage p.a	250 Betriebstage p.a	Wartungsumlage: 0,05 € / km
Fixkostenanteil 28 €/Tag	Fixkostenanteil 160 € /Tag	Variable Kosten: 0,15 € / km
Fixkosten: 188,00 € / Fzg / Tag		
=> 1 eingespartes Fzg ist äquivalent zu eingesparten		1253,3 km / Tag

Tabelle 9: Wichtungsproblem – Beispiel 2

Für die Interpretations-Sicht der OR-Probleme heißt das, die Parametrisierung der Gewichtung von fixen und variablen Kosten zueinander in der Zielfunktion ist entscheidend.

Noch deutlicher wird dies im nächsten Beispiel, bei dem lediglich kostenrechnerisch

die Jahreslohnkosten auf Stundenlöhne umgelegt wurden und die km vereinfachend als Vielfache der geleisteten Fahrerstunden approximiert wurden.

Beispiel2a:

Fuhrpark (fix)	Personal (var)	Km (var)
Eigener PKW Fuhrpark	Fahrerentlohnung auf Stundenbasis	Spritverbrauch: 8 l / 100 km
Anschaffungswert 21.000 € / Stück Afa über 3 Jahre	Lohnkosten: 20 € / h (entspricht 40.000 € p.a. ¹¹⁶)	Benzinkosten: 1,25 € / l
250 Betriebstage p.a	Durchschnittsgeschwindigkeit: 30km/h ¹¹⁷	Wartungsumlage: 0,05€ /km
	Var Kostenanteil: 0,67 € / km	Var Kostenanteil: 0,15 € / km
Fixkostenanteil: 28 €/Fzg	Variable Kosten: 0,82 € / km	
=> 1 eingespartes Fzg ist äquivalent zu eingesparten		34,1km / Tag

Tabelle 10: Wichtungsproblem – Beispiel 2a

D.h. nur weil per Umlageverfahren die Jahreslohnkosten auf Stundenlöhne umgelegt wurden, verschiebt sich die Gewichtung drastisch. Außerdem ist in der physischen Welt der verfügbare Fuhrpark oft nicht nur ein Kostenfaktor in der Zielfunktion, sondern auch eine wesentliche Nebenbedingung. (im Beispiel 1 ist eine kurzfristige Beschaffung der Spezialfahrzeuge nicht ohne weiteres möglich; während dies im Beispiel 2 einfacher sein dürfte). Eine generelle Wichtung der km gegenüber der Anzahl benötigter Fahrzeuge bleibt somit schwierig (Wichtungsproblem). Aus diesem Grund werden im Folgenden die Kostenfaktoren „km“ und „Anzahl Fahrzeuge“ **prinzipiell getrennt voneinander** betrachtet. Auf relevante Interdependenzen zwischen beiden Größen wird jeweils hingewiesen.

Anmerkung: Einige Autoren, wie z.B. [Bent & van Hentenryck 2004] oder [Hombberger & Gehring, 1999] haben das Wichtungs-Problem folgendermaßen umgangen: Für die „Gütebewertung“ einer Lösung haben sie eine lexikografische Gütefunktion definiert, die vorrangig die Anzahl benötigter Fahrzeuge, und erst nachrangig die Gesamttourkilometer berücksichtigt. Andere Autoren hingegen, z.B. [Solomon, 2003,] fokussieren nur auf die Gesamttourkilometer.

¹¹⁶ bei 250 Arbeitstagen und 8h/d

¹¹⁷ Annahmen:

- 1) langfristige Durchschnittsgeschwindigkeit für Innenstadtauslieferung beträgt 30 km/h
- 2) Wartezeiten werden nicht entlohnt
- 3) Servicezeiten werden nicht berücksichtigt, da sie in beiden Fällen anfallen

6.4 Der FPA/PPA-Ansatz (erweiterbare Flotte)

Im folgenden soll der FPA/PPA-Ansatz in Hinblick auf den Zusammenhang von zugesagten Lieferzeitfenstern und induzierten Transportkosten quantitativ untersucht werden.

Methodisch werden dazu in einem ersten Schritt ausgehend von eigenen Stichprobenuntersuchungen an drei¹¹⁸ Instanzengruppen Hypothesen aufgestellt, die dann in einem zweiten Schritt unter zu Hilfenahme von in der Literatur veröffentlichten Ergebnissen verifiziert werden.

6.4.1 Erste Ergebnisse

In einem ersten Schritt werden hier für *ein* Verfahren (HSav1) und *eine* Instanzengruppe (ZUE500*) erste Ergebnisse zum Zusammenhang zwischen PRÄZ und Gesamttourkilometern und der Anzahl benötigter Fahrzeuge beschrieben.

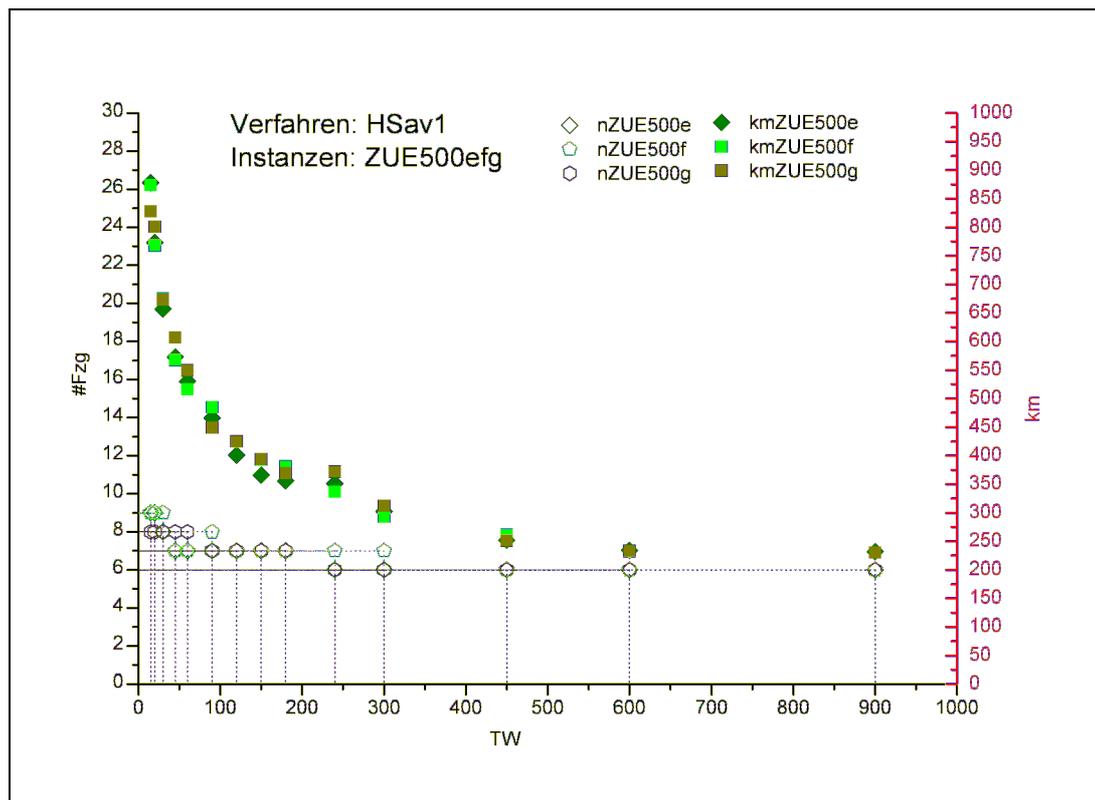


Abbildung 74: Entwicklung der km \blacklozenge und Fahrzeuganzahl \diamond in den Instanzen ZUE500efg mit Verfahren HSav1 ($\alpha=1,1$; $\beta=-2,0$; max. Wartezeit=5min)
TW:= Zeitfensterradius in Minuten (time window)

¹¹⁸ ZUE500*, R205#, und RC201#

In der obigen Grafik kann man zunächst erkennen, dass sich die Ergebnisse der drei verschiedenen Instanzen kaum unterscheiden. Da diese Beobachtung auch in allen anderen untersuchten Fällen gemacht werden konnte, soll sie hier festgehalten werden:

Beobachtung:

Die Art der Abhängigkeit zwischen km und Zeitfensterradius unterscheidet sich zwischen Instanzen dann nicht, wenn keine strukturellen Instanz-Differenzen¹¹⁹ vorliegen, d.h. wenn insbesondere keine unterschiedliche geografische Verteilung und keine signifikant unterschiedliche Clusterung der Zeitfenster (Gleichzeitigkeit) vorliegt.

Aus diesen ersten Messreihen zur Instanzengruppe „ZUE500*“ und dem Verfahren „HSav1“ ergeben sich folgende erste Ergebnisse zur Abhängigkeit der Tourkilometer von der Zeitfensterbreite:

- ▶ Sowohl die Fahrzeuganzahl, als auch die Gesamttourkilometeranzahl steigt mit enger werdenden Zeitfenstern.
- ▶ Der Steigungsverlauf ist bei den Fahrzeugen und Gesamttourkilometern unterschiedlich.

6.4.1.1 PRÄZ und Gesamttourkilometer

Eine augenscheinliche Frage, die sich aus dem visuellen Eindruck der obigen Grafik ergibt ist:

Frage: Mit welcher Funktion lässt sich die Abhängigkeit der Gesamttourkilometer von der Zeitfensterbreite (PRÄZ) funktional beschreiben?

Es gibt für eine Schar von Messpunkten meist mehrere Möglichkeiten, Funktionen (mit den zugehörigen Parametern) zu finden, die die Abhängigkeit der y von den x-Werten für einen eingeschränkten Wertebereich funktional beschreibt. Eine etablierte Methode dazu ist die nichtlineare Regression mit Hilfe des Levenberg-Marquardt-Algorithmus.

Die Entscheidungen, die bei dieser Regressionsfunktionssuche zu treffen sind, sind im folgenden dargestellt.

¹¹⁹ d.h. für ZUE500: Die Instanzen E,F und G unterscheiden sich zwar in den jeweiligen Paarungen (geografische Lage, Lieferzeitpunkt) einzelner Lieferpunkte, jedoch nicht in der geografischen Streuung und nicht strukturell in der zeitlichen Streuung.

Für welchen Wertebereich soll die Funktion gelten?

Annahmen:

- ▶ Der Wertebereich soll bei einem Zeitfensterradius von $[\pm 15 \text{ min}]$ beginnen¹²⁰
- ▶ Der Wertebereich soll mindestens bis zu einem Zeitfensterradius von $[\pm 300 \text{ min}] = [\pm 5 \text{ h}]$ gehen
- ▶ Außerhalb des so definierten Wertebereichs soll der Funktionsverlauf nicht völlig unplausibel sein

Welcher Funktionstyp kommt in Frage?

Annahmen:

- ▶ Es soll eine „einfache“ Funktion sein, die dem beobachteten Verlaufsprinzip nicht prinzipiell widerspricht
- ▶ Es soll eine „durchgängige“ Funktion sein, d.h. sie soll nicht abschnittsweise definiert sein.

Wie viele Freiheitsgrade (Parameter) soll die Funktion haben?

Annahmen:

- ▶ Möglichst wenige, da ein Funktionstyp mit vielen Freiheitsgraden schwer zu parametrisieren ist, wenn Vorhersagen getroffen werden sollen,
- ▶ aber genug, um eine hinreichende Genauigkeit für die nichtlineare Regression zu erreichen.

Auf Basis dieser Annahmen wurden Regressionsversuche mit folgenden (im Wertebereich monoton fallenden) Funktionstypen unternommen:

▶ Funktionen mit 2 Parametern

▶ Exponentiell (Exp2Pmod1) $y(x) = ae^{bx}$ (6.1.)

▶ Potenziell (Allometric1) $y(x) = ax^b$ (6.2.)

▶ Rational (Reciprocal) $y(x) = \frac{1}{a+bx}$ (6.3.)

▶ Rational (Reciprocal_ok1) $y(x) = a + \frac{1}{bx}$ (6.4.)

▶ Rational (Rational4_ok3) $y(x) = a + \frac{b}{x+0.1407a}$ (6.5.)

▶ Funktionen mit 3 Parametern

▶ Exponentiell (YldFert1) $y(x) = a + be^{-kx}$ (6.6.)

¹²⁰ höhere Präzisionen scheinen für die Praxis nicht sinnvoll, da Verkehrsstörungen und unplanbare Ereignisse wie unerwartet hohe Lieferzeiten bei Vorgängerstops eine präzisere Vorausschau kaum sinnvoll erscheinen lassen

▶ Exponentiell (Exp3P1) $y(x) = ae^{\frac{b}{x+c}}$ (6.7.)

▶ Potenziell (Allometric2) $y(x) = a + bx^c$ (6.8.)

▶ Potenziell (Harris) $y(x) = \frac{1}{a + bx^c}$ (6.9.)

▶ Rational (Rational0) $y(x) = \frac{b + cx}{1 + ax}$ (6.10.)

▶ Rational (Rational4) $y(x) = c + \frac{b}{x + a}$ (6.11.)

Die Ergebnisse der Fittingversuche sind im Anhang zusammengefasst.

Von den untersuchten 3-parametrischen Funktionen liefern mehrere ähnlich gute Ergebnisse (s. Anhang).

Von den untersuchten 2-parametrischen Funktionen scheint „**Allometric1**“ ein geeigneter Kandidat zu sein, der die Abhängigkeit der km von der Zeitfensterbreite (zumindest für die Instanzen ZUE500efg und das „HSav1“-Verfahren) gut beschreibt.

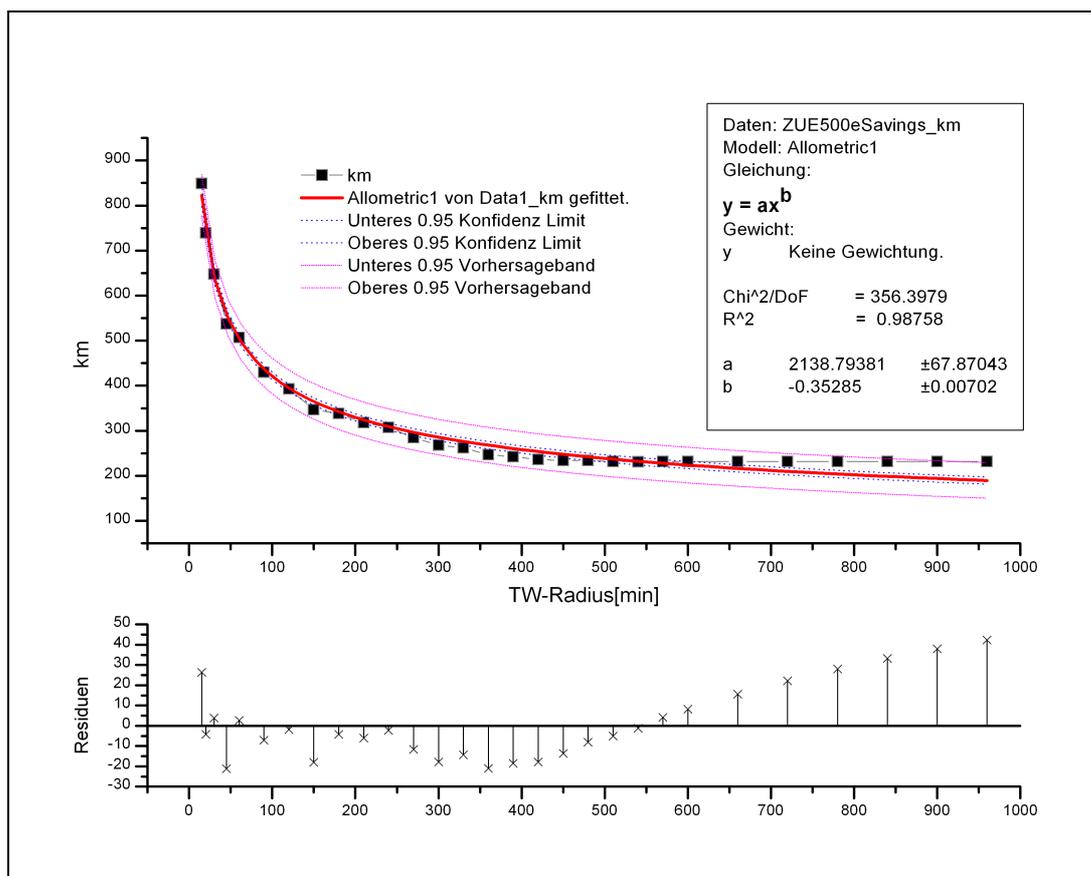


Abbildung 75: Abhängigkeit der Tourkilometer vom Zeitfensterradius

Die Abbildung zeigt, dass der Zusammenhang zwischen km und Zeitfensterbreite mit einem R^2 von 0,988 für den untersuchten Fall recht gut beschrieben werden kann¹²¹. Die Frage, die aus diesem Stichprobenergebnis erwächst ist:

Frage:

*Gilt dieser Funktionstyp möglicherweise universell auch für andere **Verfahren** und **Instanzen**?*

Um diese Frage zu beantworten, wurden unten weitere Verfahren anhand weiterer Instanzen untersucht (s. 6.4.2).

6.4.1.2 PRÄZ und Fahrzeuganzahl

Bei der Untersuchung verschiedener Verfahren anhand der Instanzengruppe ZUE500* fiel auf, dass die Fahrzeuganzahl nicht immer monoton mit abnehmender PRÄZ steigt.

Beim Einsatz des auf dem einfachen Savings basierten Verfahrens „HSav0“ (ohne Begrenzung der Fahrzeuganzahl) ist die Fahrzeuganzahl im Gegensatz zu den Gesamttourkilometern beispielsweise nicht monoton mit dem Zeitfensterradius korreliert (s. Grafik: □).

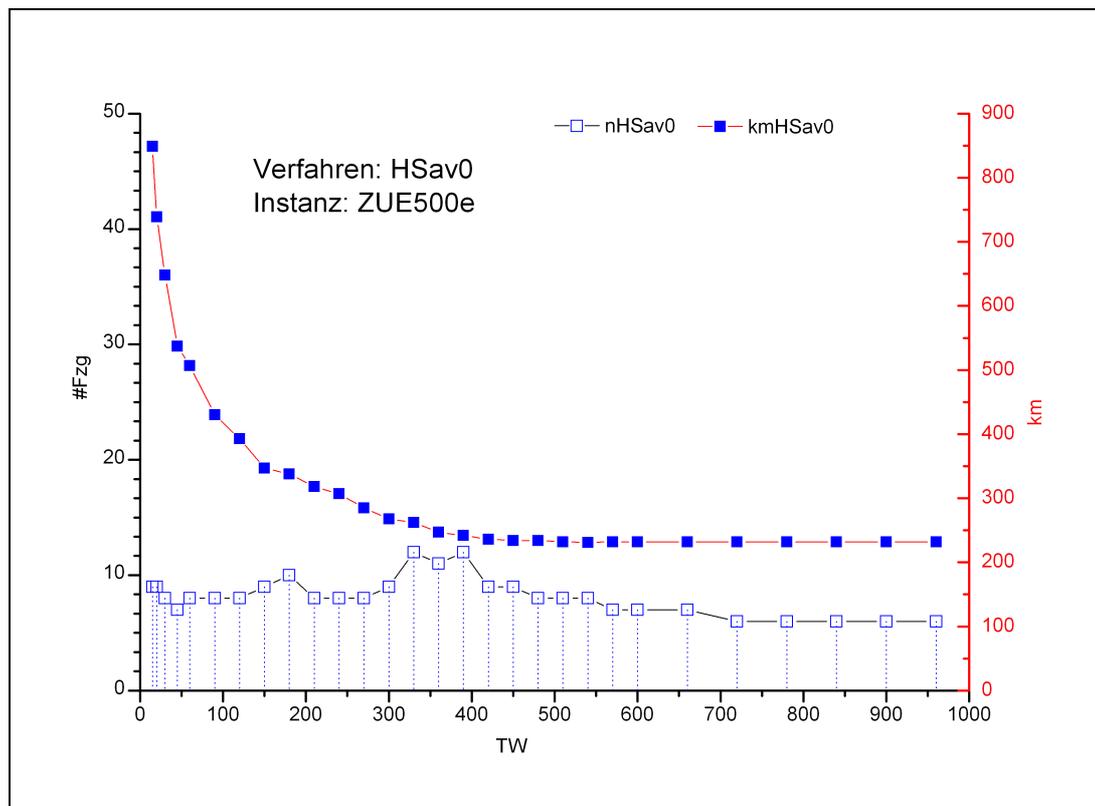


Abbildung 76: Nichtmonotone Abhängigkeit der Fahrzeuganzahl vom Zeitfensterradius

¹²¹ Hinweis: Dies ist ein Einzelfall-Zwischenergebnis, dass sich nicht gut verallgemeinern lässt (Kap. 6.4.2.-6.4.4.)

Erkenntnis: Es gibt Verfahren (z.B. „HSav0“), die keine „quasimonotone“¹²² Korrelation“ für die Abhängigkeit zwischen der Anzahl benötigter Fahrzeuge von der Lieferzeitfensterbreite TW aufweisen.

Für das Verfahren „HSav0“ ist das leicht zu erklären, denn das klassische Savingsverfahren hat in der Zielfunktion keinen Anreiz zur Fahrzeugreduzierung. (Details zu diesem Verhalten sind im Anhang am Beispiel der Instanz ZUE500e näher dargestellt). Selbst eine nachgeschaltete Fahrzeugeinsatzoptimierung kann die Fahrzeuganzahl bei mittelgroßen Zeitfenstern nicht mehr unter die Anzahl der von Verfahren simultan aufgebauten relativ langen Touren senken.

Durch eine heuristische „Hülle“ um das klassischen Savings Verfahren, wie sie im Verfahren „HSav1“ realisiert wurde, kann dieses Phänomen eliminiert werden, indem sukzessive die Anzahl der verfügbaren Fahrzeuge erhöht wird, bis alle Stops verplant werden können (= sukzessive Relaxation der Nebenbedingungen in Bezug auf die verfügbaren Fahrzeuge).

Mit Hilfe der Hülle des Verfahrens HSav1 um den Savings-Kern kann eine quasimonotone Abhängigkeit der Anzahl benötigter Fahrzeuge von der PRÄZ erreicht werden. Der typischer Verlauf ist in der folgenden Grafik dargestellt.

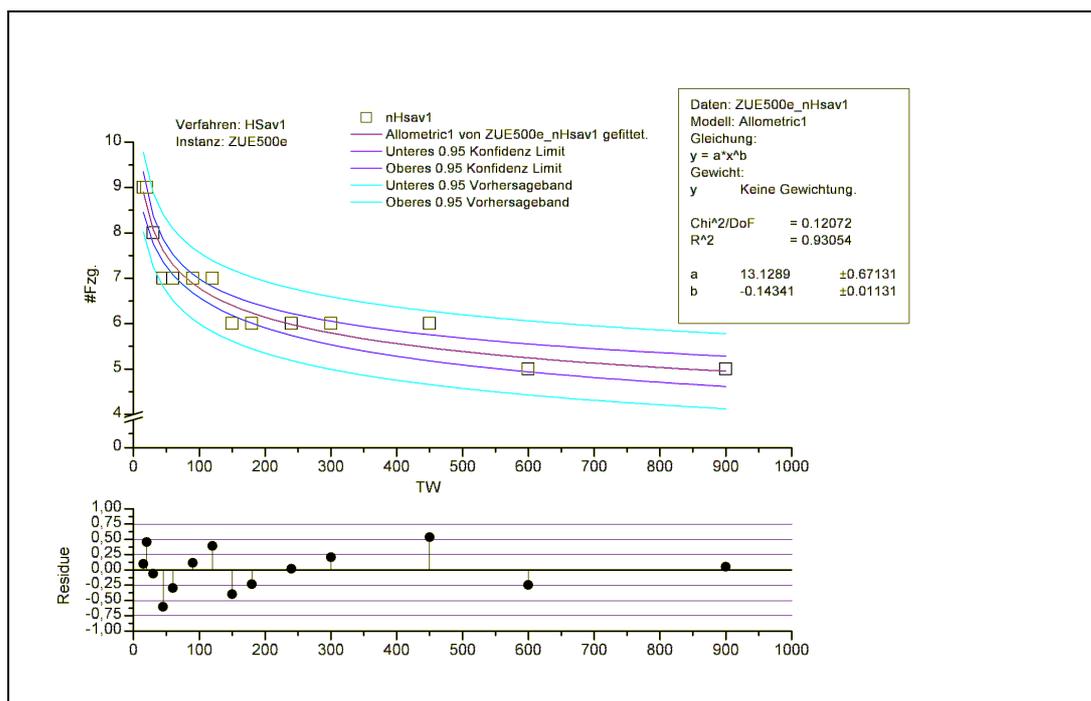


Abbildung 77: Abhängigkeit der Anzahl benötigter Fahrzeuge vom Zeitfensterradius am Beispiel der Instanz ZUE500e mit dem Verfahren HSav1 (alpha=1,3; beta=2,0; max. Wartezeit = 5min)

Obwohl die Fahrzeuganzahl ganzzahlig ist, scheint sich auch ihre Abhängigkeit von der Zeitfensterbreite durch eine allometrische Funktion recht gut beschreiben zu

¹²² „quasimonoton“ := im Prinzip monoton mit eventuellen stochastischen Ausreißern; „monoton“ gemäß [Bronstein & Semendjajew 1991] S. 245.

lassen¹²³. Ein etwas erweiterter Test mit anderen Parametern und über alle 3 Instanzen ZUE500efg hinweg bestätigt dies:

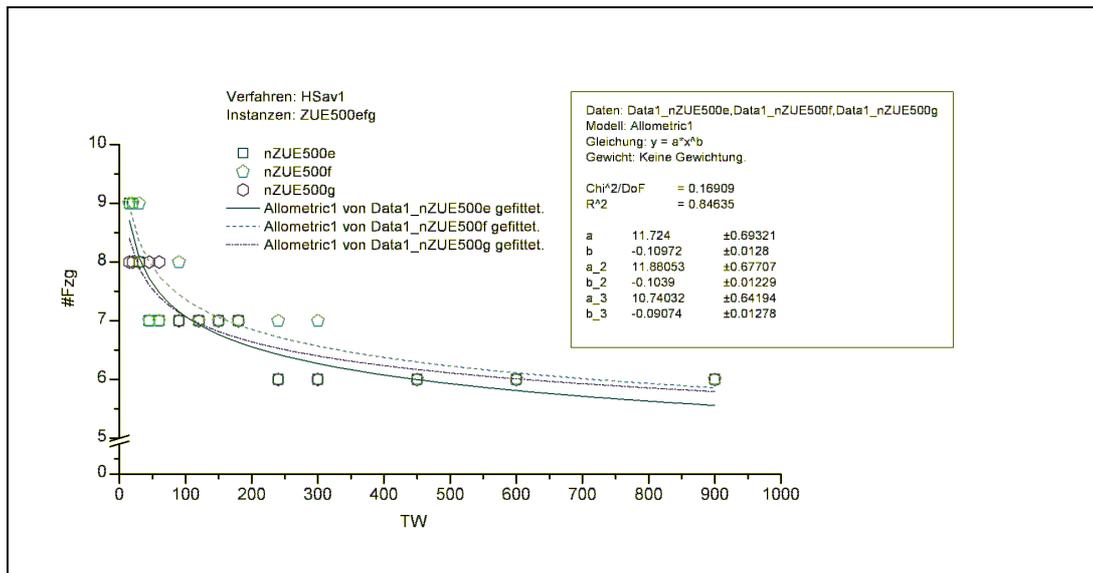


Abbildung 78: Abhängigkeit der Anzahl benötigter Fahrzeuge vom Zeitfensterradius am Beispiel der Instanzen ZUE500efg mit dem Verfahren HSav1 (alpha=1,1; beta=-2,0; max. Wartezeit =5min)

6.4.2 Instanz- & verfahrensübergreifende Generalisierungsansätze

Die obigen Erkenntnisse bezüglich der allometrischen Abhängigkeit der Gesamttour-km und der Fahrzeuge von der Zeitfensterbreite sind bisher nur stichprobenhaft für ausgewählte Verfahren und ausgewählte Instanzen gezeigt worden.

Um dieses Phänomen Instanz- und Verfahrens-übergreifend generalisieren zu können, müsste eine sehr große Anzahl von Messungen durchgeführt werden. Außerdem hängen die Ergebnisse nicht nur von den Verfahren selbst, sondern auch von den jeweiligen Parametrisierungen der Verfahren ab. Daher müssten für jedes Verfahren und jede Instanz die richtigen Parametrisierungen jeweils erst gefunden werden – besonders bei neuen Verfahren ist dies oft nicht trivial. Eine solche nachhaltig fundierte Generalisierung kann im Rahmen dieser Arbeit nicht erbracht werden.

Dennoch soll hier der Versuch einer Generalisierung gemacht werden. Dazu wurden zunächst weitere Messungen mit verschiedenen Verfahren an den Solomoninstanzklassen R205# und RC201# vorgenommen. Im Anschluss wird diese Generalisierung durch Auswertung von Literaturergebnissen noch erweitert.

¹²³ Durch die Ganzzahligkeit der Fahrzeuganzahl ist das R² erwartungsgemäß nicht sehr hoch, aber immer noch deutlich über 0,8. Außerdem ist die Abweichung vom Regressionsfunktionsverlauf stets kleiner 1 Fzg.

6.4.2.1 Instanzen Solomon R205# (Eigene Messungen)

Noch könnte es sein, dass die Instanzgruppe ZUE500* oder das Savings-Verfahren allein diese Funktionsverläufe bedingen. Daher wird das Phänomen nun an zwei weiteren Instanzklassen untersucht, die auf den Solomoninstanzen basieren.

Sei R205# ein Satz von Variationen der Solomoninstanz R205 bei der die ursprünglichen Zeitfenster von 240 min. um ihren Mittelpunkt herum verengt und erweitert werden, so dass die Zeitfenster die Breiten „#“ \in (10, 20, 30, 60, 90, 120, 180, 240, 300, 360, 420, 480, 1000) ZE aufweisen.

Für diese Variationen der Solomoninstanz R205 wurden mit 3 Verfahren Tourenpläne erzeugt, und deren Gesamtkilometer verglichen. Die eingesetzten Verfahren waren:

- ▶ HSav1
"best of 2" - Ergebnisse eines klassischen Savings-Verfahrens mit den Parametersätzen (α, β, tw) mit $\alpha \in \{1,3; 1,2\}$; $\beta \in \{-0,2; 0,0\}$ und $tw [\text{min}] \in \{7; 20; 60\}$
- ▶ HSav1Ins2:
wie HSav1, aber mit nachgeschaltetem "Insertion"-Verfahren (Einfügekriterium: minimaler Umweg), falls durch das reine „Savings“-Verfahren nicht alle¹²⁴ Kunden verplant werden konnten.
- ▶ HIns1:
"best of 10" – Ergebnisse eines klassischen Insertion Verfahrens mit jeweils 10 n-Tupeln von n „seed-trips“ zu n zufällig ausgewählten Kunden, wobei n der Anzahl der verfügbaren Fahrzeuge entspricht.

Für alle Verfahren wurde die Anzahl der verfügbaren Fahrzeuge n so lange erhöht, bis erstmals alle Kunden verplant werden konnten.

¹²⁴ Die Mindestanzahl der bereits durch den Savings verplanten Kunden lag hierbei innerhalb [95;100] %

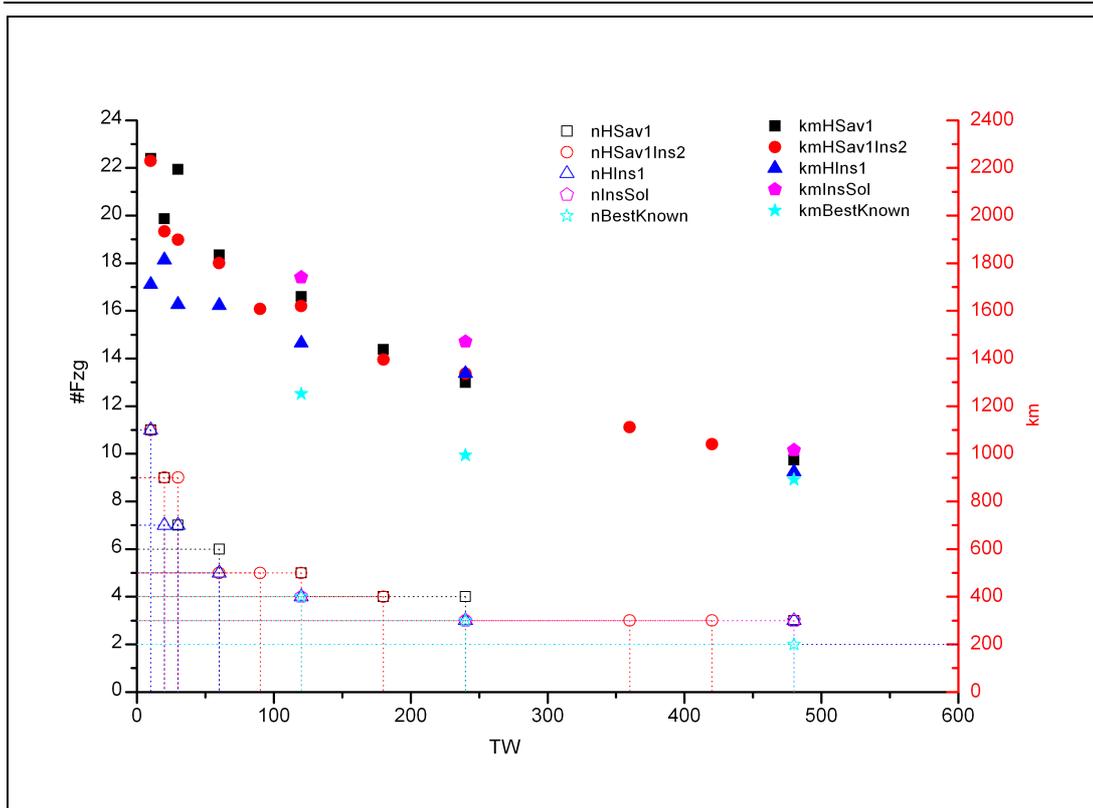


Abbildung 79: Tourlänge & Fahrzeuganzahl in Abhängigkeit von der Zeitfensterbreite TW für R205#

Im unteren Bereich der Grafik ist die Anzahl der benötigten Fahrzeuge dargestellt, im oberen Bereich der Grafik die Anzahl der benötigten Tourkilometer. Damit diese Messungen nicht im „luftleeren Raum“ hängen, wurden in der Darstellung zusätzlich 6 Referenzpunkte eingezeichnet, die die bekannten „**best known**“ Ergebnisse (★), und die Ergebnisse des Solomon Insertion Verfahrens **InsSol** (◆) zu den Instanzen R201, R205 und R211 angeben. Auch wenn R201 und R211 nicht 1:1 zu R205# passen, so sind diese „Referenzinstanzen“ zu der Instanzklasse R205# zumindest recht gut vergleichbar, wie die folgende Tabelle zeigt:

Instanz	Zeitfensterbreite Vorgabe [ZE]	Zeitfensterbreite MW [ZE]	Zeitfensterbreite SD	Symbole für "best Known" & "InsSol"	Symbole für Messwerte R205#
R201 ¹²⁵	120	115,96	35,78	★ ◆	
R205# (120)	120	120,00	0,00		■ ● ▲
R205 ¹²⁶	240	240,00	0,00	★ ◆	
R205# (240)	240	240,00	0,00		■ ● ▲
R211 ¹²⁷	480	471,94	71,67	★ ◆	
R205# (480)	480	480,00	0,00		■ ● ▲

Tabelle 11: Solomon-Referenzinstanzen

Auch für die Variation der Solomon Instanz R205# lässt sich zeigen, dass eine allometrische Abhängigkeit der Fahrzeuganzahl von der Zeitfensterbreite vorliegt.

¹²⁵ Die Instanz R205 (120) ist in etwa vergleichbar mit R201, da sie eine ähnliche mittlere Zeitfensterbreite aufweist, und die Streuung der Zeitfensterbreite minimal ist.

¹²⁶ Die Instanz R205 ist selbst Element des Satzes R205#.

¹²⁷ Die Instanz R205 (480) ist in etwa vergleichbar mit R211, da sie eine ähnliche mittlere Zeitfensterbreite aufweist, und die Streuung der Zeitfensterbreite minimal ist.

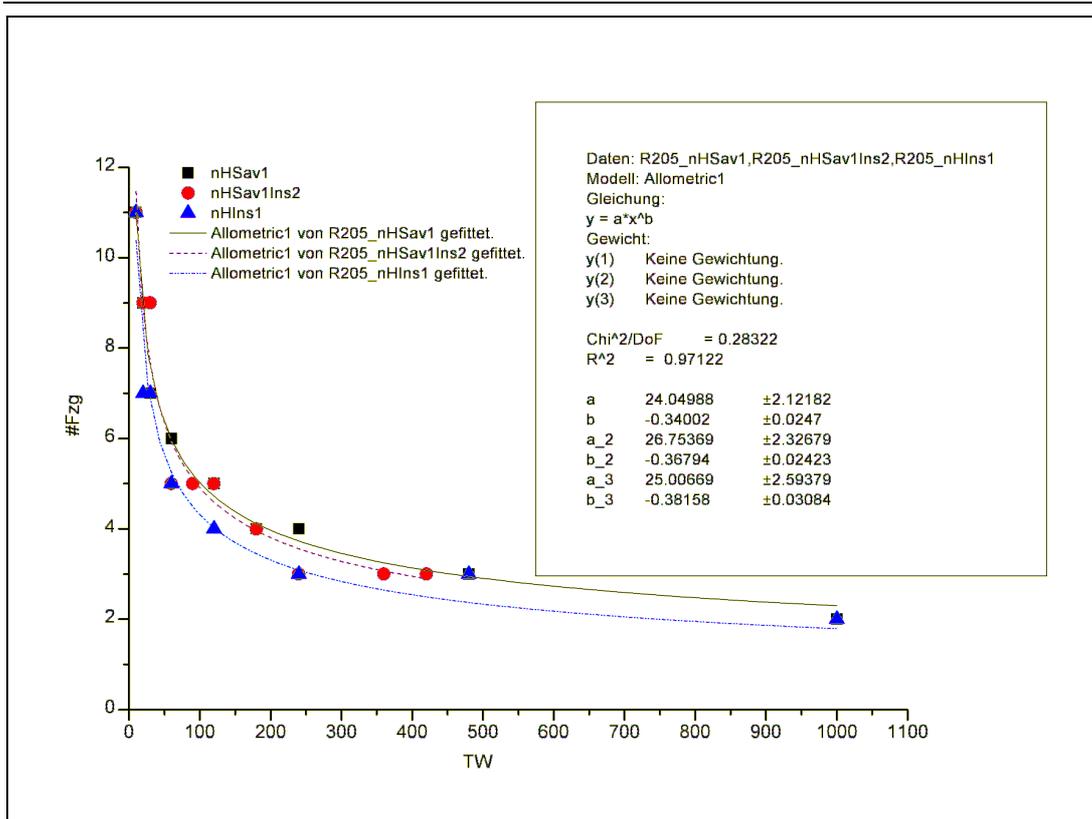


Abbildung 80: Allometrische Regression zur Fahrzeuganzahl für R205#

Allerdings lässt sich die Annahme einer allometrischen Abhängigkeit der Tourkilometer von der Zeitfensterbreite TW für diese Instanzklasse nicht gut bestätigen, denn ein Funktionsverlauf exponentieller Art (YldFert1) liefert deutlich bessere Ergebnisse im nichtlinearen Fitting sowohl für den HSav1 als auch für HIns1.

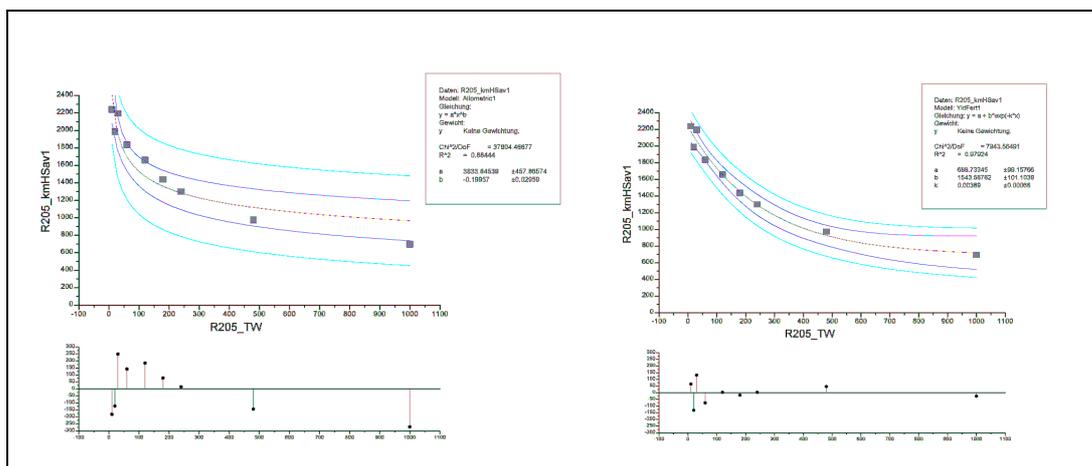


Abbildung 81: Vergleich der Fittingversuche „Allomtric1“ ($R^2=0,884$) und „YldFert1“ ($R^2=0,979$) für das Verfahren HSav1

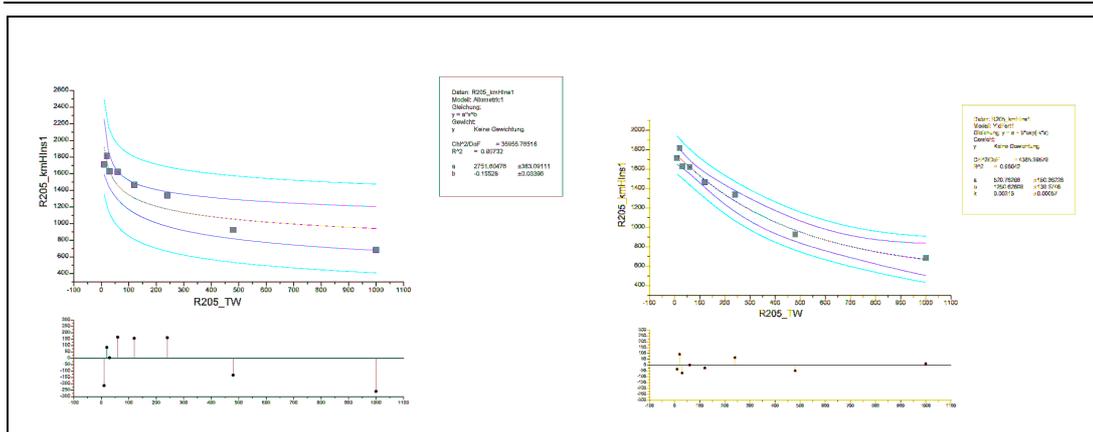


Abbildung 82: Vergleich der Fittingversuche „Allometric1“ ($R^2=0,807$) und „YldFert1“ ($R^2=0,980$) für das Verfahren HIns1

Die Vermutung aus Kap. 6.4.1., dass die Gesamttourkilometer durch eine 2-parametrische allometrische Funktion beschreiben werden können, lässt sich also nicht generalisieren. Eine Beschreibung durch die 3-parametrische Funktion YldFert1 ist hingegen auch mit den Ergebnissen aus 6.4.1. kompatibel.

6.4.2.2 Instanzen Solomon RC201# (Eigene Messungen)

Für die Instanz RC201# wurden Messungen für die drei Verfahren „HSav1“, „HSav1Ins2“ und „HIns1“ durchgeführt.

Anschließend wurden allometrische Regressionsversuche bzgl. der Anzahl Fahrzeuge und „YldFert1“-Regressionsversuche bzgl. der Tourkilometer unternommen. Die Ergebnisse sind in den folgenden Charts dargestellt.

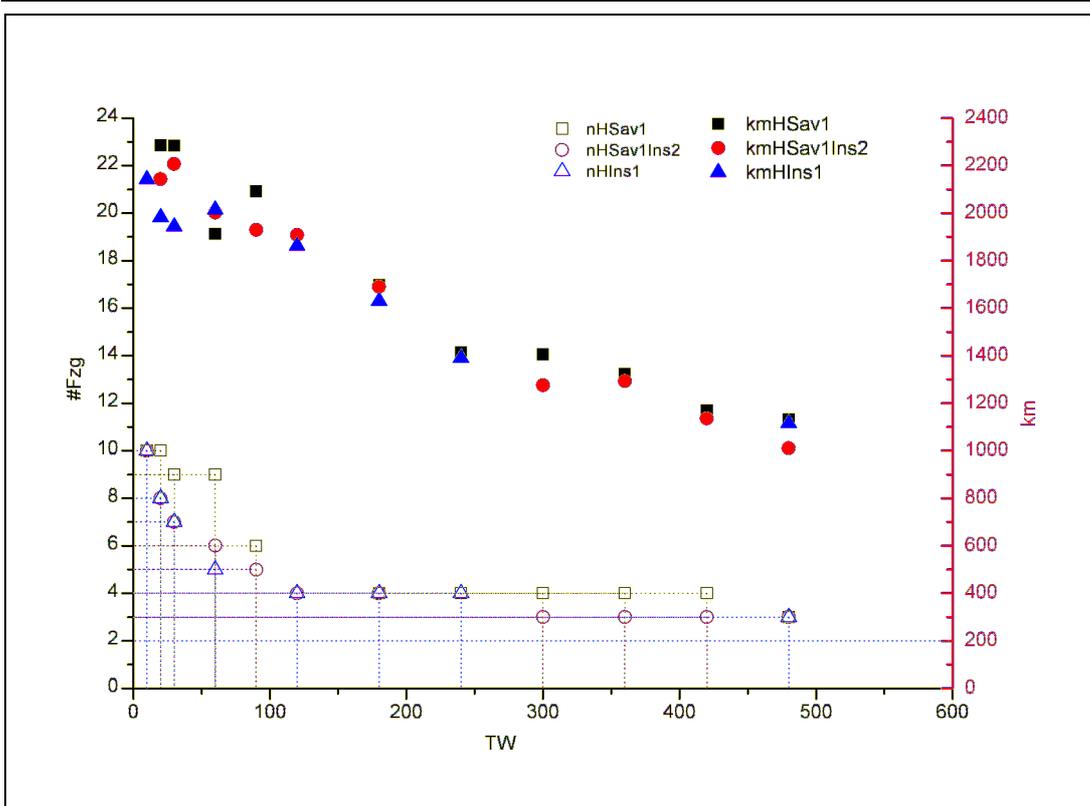


Abbildung 83: Instanz RC201#

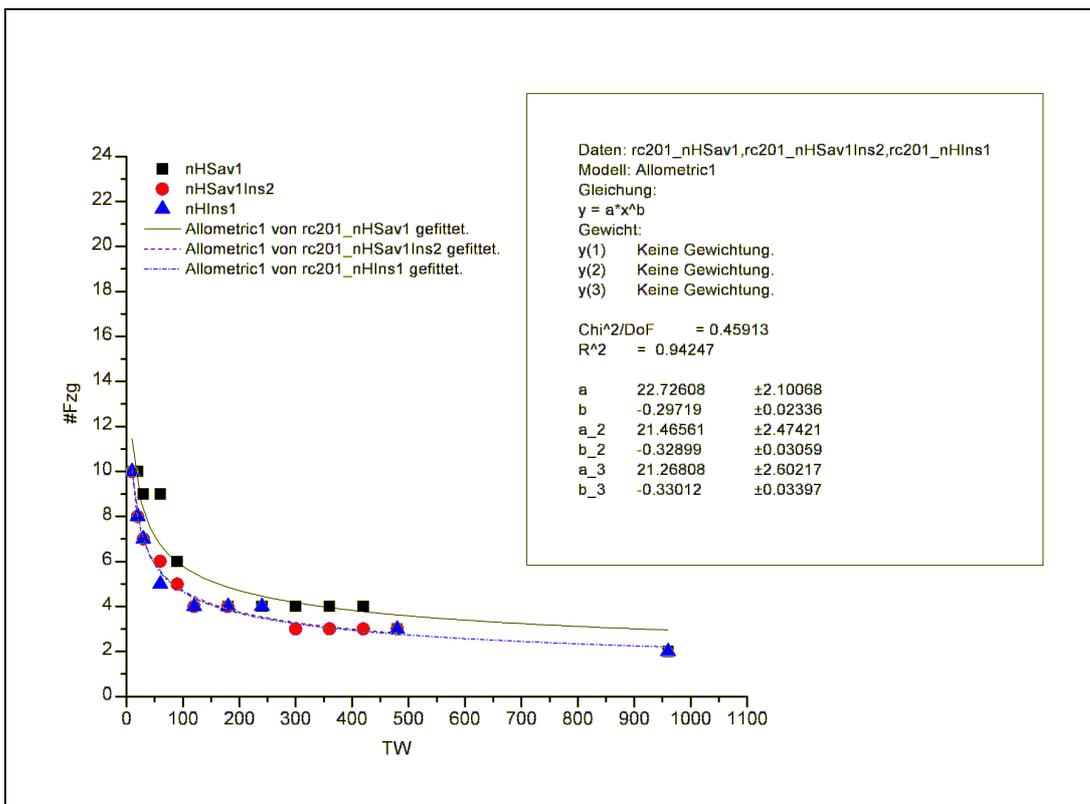


Abbildung 84: Instanz RC201# - Anzahl Fahrzeuge

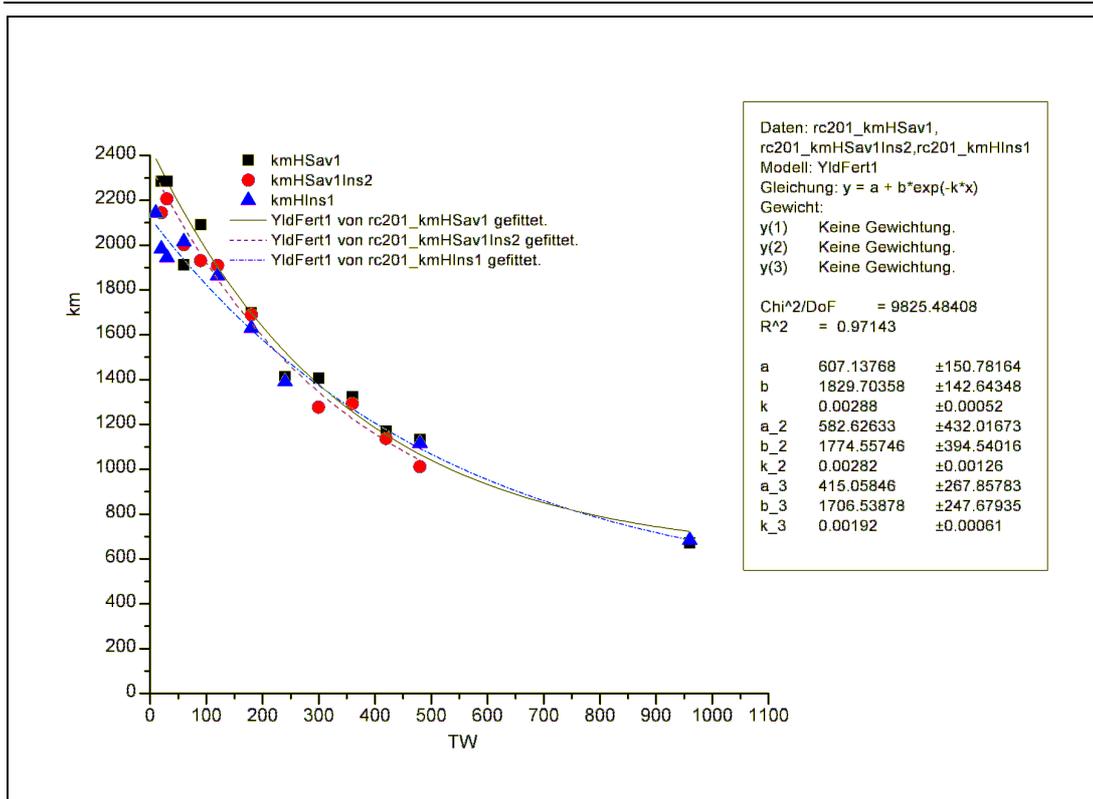


Abbildung 85: Instanz RC201# - Tourkilometer

6.4.3 Hypothesen zum FPA/PPA

Aus den vorgehenden Überlegungen lassen sich folgende Hypothesen ableiten:

*H1a: Für einen definierten Satz von bereits bestätigten Lieferungen, der keine ex ante erkennbare implizite Tourstruktur aufweist, wächst die Anzahl der benötigten **Fahrtkilometer** im Mittel quasimonoton mit zunehmend enger werdenden Zeitfenstern unabhängig vom eingesetzten Verfahren zur Tourenbildung.*

*H1b: Für einen definierten Satz von bereits bestätigten Lieferungen, der keine ex ante erkennbare implizite Tourstruktur aufweist, wächst die Anzahl der benötigten **Fahrzeuge** im Mittel quasimonoton mit zunehmend enger werdenden Zeitfenstern unabhängig vom eingesetzten Verfahren zur Tourenbildung, wenn im Verfahren einen Anreiz zur Fahrzeugminimierung implementiert ist.*

Warum ist die Formulierung dieser Hypothesen so gewählt? Die wesentlichen Kostentreiber in der Auslieferung sind die Fixkosten (der wesentliche Transportfixkostenindikator ist die Anzahl der Fahrzeuge) und die variablen Kosten (ein wesentlicher Indikator für variable Transportkosten ist die Anzahl von den gefahrenen Kilometern). Nun ließe sich die Abhängigkeit der Kosten von den

Zeitfensterenge über ein gewichtetes Mittel von Fixkosten und variablen Kosten untersuchen. Allerdings wäre dann der gewählte Gewichtungsfaktor entscheidend für die Aussagekraft der Hypothese. Um das Wichtungproblem zwischen variablen und fixen Kosten auszuklammern, werden in den Hypothesen beide Aspekte einzeln betrachtet.

Die „wenn“-Bedingung in der Hypothese zur Fahrzeuganzahl ist motiviert durch die Messergebnisse zum HSav0.

Schärfer lassen sich die beiden Hypothesen formulieren als:

*H2a: Für einen definierten Satz von bereits bestätigten Lieferungen, der keine ex ante erkennbare implizite Tourstruktur aufweist, wächst die Anzahl der benötigten **Fahrkilometer** im Mittel quasimonoton mit zunehmend enger werdenden Zeitfenstern unabhängig vom eingesetzten Verfahren zur Tourenbildung, nach dem exponentiellen YldFert¹²⁸-Funktionstyp*

$$l_{ges}(TW) = a + be^{-kTW}$$

mit

l_{ges} = Gesamttourkilometer

TW = Zeitfensterbreite

a, b, k : Parameter des Funktionsverlaufs

*H2b: Für einen definierten Satz von bereits bestätigten Lieferungen, der keine ex ante erkennbare implizite Tourstruktur aufweist, wächst die Anzahl der benötigten **Fahrzeuge** im Mittel quasimonoton mit zunehmend enger werdenden Zeitfenstern unabhängig vom eingesetzten Verfahren zur Tourenbildung, wenn im Verfahren einen Anreiz zur Fahrzeugminimierung implementiert ist, nach dem allometrischen Funktionstyp*

$$m(TW) = A TW^B$$

mit

m = Anzahl Fahrzeuge

TW = Zeitfensterbreite

A, B : Parameter des Funktionsverlaufs

Über konkrete Werte von a , b , k , A und B wird hier noch nichts ausgesagt. Die Hypothesen beziehen sich lediglich auf den Funktionstyp.

H3a/b: Die Parameter ergeben sich jeweils aus Überlagerung von Problemspezifika und Verfahrensspezifika

Diese Hypothese ist eine reine Existenzhypothese, d.h. sie postuliert die Abhängigkeit der Parameter vom Problem an sich und dem Verfahren, sagt aber noch nichts über die Art der Abhängigkeit.

¹²⁸ „YldFert“ – ein Funktionstyp, der oft zur Darstellung der Korrelation von Düngereinsatz und Ertrag (Yield Fertilizer) sowie zur Beschreibung von Lernkurven in der Psychologie eingesetzt wird. Terminologie „YldFert1“ nach der Benennung in Origin Pro.

H4: Die beiden Größen $I_{ges}(TW)$ und $m(TW)$ dürfen nicht unabhängig voneinander betrachtet werden.

6.4.4 Überprüfung der Hypothesen zum FPA/PPA

Die dem PPA-Ansatz zugrundeliegenden Verfahren zur Lösung des VRSP-TW sind in der OR-Literatur intensiv untersucht worden. Da dieser Ansatz bedeutet, einen Satz von zeitlich terminierten Lieferaufträgen mit einer nach oben unbeschränkten Flotte zu befriedigen, kann die gesamte Literatur zum VRSP-TW (bei der die Anzahl der verfügbaren Fahrzeuge keine maßgeblich zum tragen kommende Nebenbedingung darstellt) herangezogen werden, um die Auswirkungen auf die Transportkosten zu untersuchen. Aus diesem Grund werden in dieser Arbeit diese veröffentlichten Ergebnisse unter dem Blickwinkel der Service-Kosten-Interdependenz neu ausgewertet.

Um die veröffentlichten Ergebnisse zu den Solomoninstanzen im Sinne dieser Arbeit interpretieren zu können, muss eine wesentliche Annahme getroffen werden:

Annahme:

Die Art der Abhängigkeit zwischen km und Zeitfensterradius bzw. Anzahl Fahrzeuge und Zeitfensterradius in Hinblick auf ein Verfahren unterscheidet sich zwischen Instanzen dann nicht, wenn keine strukturellen Differenzen zwischen den Instanzen vorliegen (d.h. insbesondere keine unterschiedliche geografische Verteilung, keine signifikant unterschiedliche Clusterung der Zeitfenster und keine signifikant großen Unterschiede in den Zeitfenstern)

Diese Annahme wird gestützt durch die Ergebnisse der Untersuchungen zu ZUE500*, R205# und RC201#.

Die folgenden Solomoninstanzen erfüllen diese Annahme jeweils für ihre Instanzgruppe:

Instanz- gruppe	Instanz	Zeitfenster-breite Vorgabe [ZE]	Zeitfenster- breite MW [ZE]	Zeitfenster-Stand- zeit [ZE]	Implizite Tourstruktur
R1*	R101	10	10,00	0,00	10nicht sortiert
	R105	30	30,00	0,00	10nicht sortiert
	R109	60	58,89	8,93	10nicht sortiert
	R112	120	117,64	17,45	10nicht sortiert
RC1*	RC101	30	30,00	0,00	10nicht sortiert
	RC106	60	60,00	0,00	10nicht sortiert
	RC108	120	112,33	30,80	10nicht sortiert

Instanz- gruppe	Instanz	Zeitfenster-breite Vorgabe [ZE]	Zeitfenster- breite MW [ZE]	Zeitfenster-Stand- breite SD zeit [ZE]	Implizite Tourstruktur
R2*	R201	120	115,96	35,78	10nicht sortiert
	R205	240	240,00	0,00	10nicht sortiert
	R211	480	471,94	71,67	10nicht sortiert
RC2*	RC201	120	120,00	0,00	10nicht sortiert
	RC206	240	240,00	0,00	10nicht sortiert
	RC208	480	471,93	71,67	10nicht sortiert
C1*	C101	60	60,76	10,53	90nach Clustern sortiert
	C105	120	121,61	20,98	90nach Clustern sortiert
	C107	180	180,00	0,00	90nach Clustern sortiert
	C108	240	234,28	41,96	90nach Clustern sortiert
	C109	360	360,00	0,00	90nach Clustern sortiert
C2*	C201	160	160,00	0,00	90nach Clustern sortiert
	C205	320	320,00	0,00	90nach Clustern sortiert
	C206	480	486,64	83,99	90nach Clustern sortiert
	C208	640	640,00	0,00	90nach Clustern sortiert

Tabelle 12: Übersicht Solomon Instanzen

Methodischer Hinweis: Für die Hypothesenvalidierung an den Instanzgruppen RC1*, R2* und RC2* gibt es ein Bestimmtheitsproblem. Da die Instanzgruppen jeweils nur 3 Stützpunkte beinhalten, die „YldFert1“-km-Regression aber 3 Parameter benötigt, gilt, solange die Messpunkte einen von unten konvexen¹²⁹ Funktionsverlauf aufweisen: die 3 Parameter lassen sich so bestimmen, dass ein $r^2=1.0$ erreicht werden kann. Auch wenn der Aussagegehalt der Hypothesenüberprüfung an diesen 3 Instanzgruppen für YldFert1 deshalb geringer als wünschenswert ist, werden diese der Vollständigkeit halber betrachtet.

6.4.4.1 R1*

Die in der Literatur veröffentlichten Ergebnisse zur Instanzgruppe R1* stützen die obigen Hypothesen:

¹²⁹ „konvex von unten“ gemäß [Bronstein & Semendjajew 1991] S.270

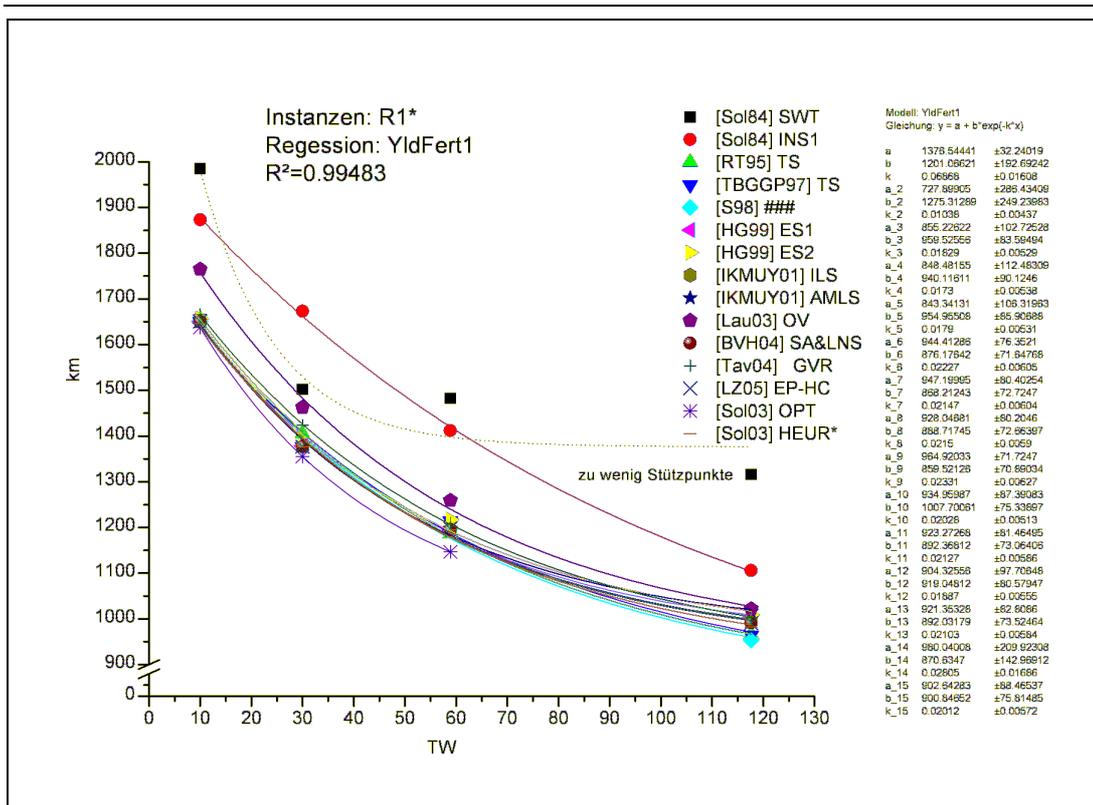


Abbildung 86: Instanz R1* km

- ▶ H1a/b: Sowohl der Tourkilometerverlauf $I_{ges}(TW)$ als auch der Fahrzeugverlauf $m(TW)$ steigen mit abnehmender Zeitfensterbreite
- ▶ H2a/b: Der Tourkilometerverlauf $I_{ges}(TW)$ lässt sich mit dem Funktionstyp *YldFert1* für jeden untersuchten Algorithmus gut beschreiben. Der Fahrzeugverlauf $m(TW)$ lässt sich mit dem Funktionstyp *Allometric1* für jeden untersuchten Algorithmus gut beschreiben
- ▶ H3a: Die problemspezifische km-Charakteristik wird sichtbar anhand des Funktionsverlaufs von [Sol03] OPT (= nachgewiesene optimale Lösungen für die Instanzen bzgl. minimaler Tourkilometer) – „blaue Linie als Untergrenze“
- ▶ H3b: Die problemspezifische Fahrzeug-Charakteristik lässt sich annähern anhand des Funktionsverlaufs von [Sol03]Heur* (= beste heuristische Lösungen für die Instanzen bzgl. minimaler Tourkilometer) – „orange Linie als approximative Untergrenze“
- ▶ H3a/b: Die verfahrensspezifische Charakteristik, die die problemspezifische Charakteristik überlagert, wird klar erkenntlich.¹³⁰

¹³⁰ Die Messwerte bezüglich des ursprünglich von Solomon eingesetzten Verfahrens „[Sol84] SWT“ fallen als einzige signifikant aus der Reihe. Es ist stark zu vermuten, dass der „Positiv-Ausreißer-Messwert“ (30;1500) daher rührt, dass gleichzeitig die Fahrzeuganzahl einen „Negativ-Ausreißer“ aufweist (s. folgende Grafik). Trotzdem wäre zu prüfen,

- ▶ ob die verfahrensspezifische Charakteristik, d.h. der Regressionskurvenverlauf für „[Sol84] SWT“ tatsächlich prinzipiell anders aussieht wie es in der Grafik erscheint, oder
- ▶ ob die verfahrensspezifische Charakteristik für „[Sol84] SWT“ eine große Streuung aufweist, und daher die geringe Anzahl von Stützpunkten zusammen mit dieser großen Streuung um die Verfahrenscharakteristik das Regressionsergebnis bei nur 4 Stützpunkten stark verfälscht

- ▶ H4: Eine Tourkilometerminimierung geht nicht notwendigerweise einher mit einer Minimierung der Fahrzeuganzahl (s. Verlauf der blauen Linen zu [Sol03]OPT)

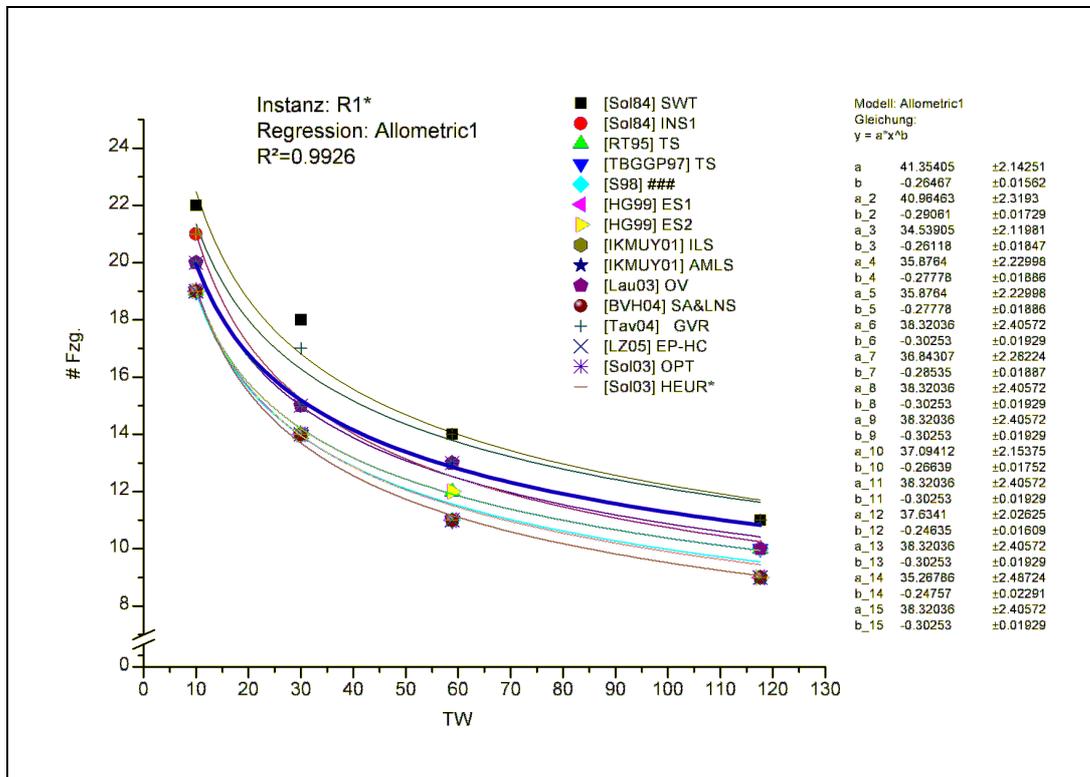


Abbildung 87: Instanz R1* Fzg

Der „atypische“ Verlauf der beiden Kurven von [Sol84]SWT ist zum einen dadurch gekennzeichnet, dass es zu wenige Stützpunkte gibt, um Ausreißer klar zu ermitteln. Zum anderen ist [Sol84]SWT ähnlich wie HSav0 ein Savings-basierten Verfahren ohne Anreizfunktion zur Fahrzeugminimierung. Der „außergewöhnlich gute“ Wert von $l_{ges}(TW=30)=1500 \text{ km}$ wird durch den „ziemlich schlechten Wert“ $m(TW=30)=18$ erreicht.

6.4.4.2 RC1*

Auch die in der Literatur veröffentlichten Ergebnisse zur Instanzgruppe RC1* stützen die obigen Hypothesen, wenn auch der Aussagewert bzgl. der $l_{ges}(TW)$ und $m(TW)$, wie im Hinweis oben wegen der Überparametrisierung in Bezug auf die wenigen Stützpunkte erwähnt, relativ gering ist.

Eine solche Prüfung müsste mit dem Originalverfahren SWT von Solomon erfolgen, und ähnlich aufgebaut sein wie z.B. die eigenen Messungen zu R201# in dieser Arbeit. Ohne eine solche Prüfung sind die Ergebnisse zu „[Sol84] SWT“ mit Vorsicht zu interpretieren.

Trotz der wenigen Stützpunkte lässt sich die km-Problemcharakteristik (Untergrenze der Funktionsverläufe) gut erkennen, und auch die Unterschiede zwischen den jeweiligen Verfahrenscharakteristiken werden deutlich.

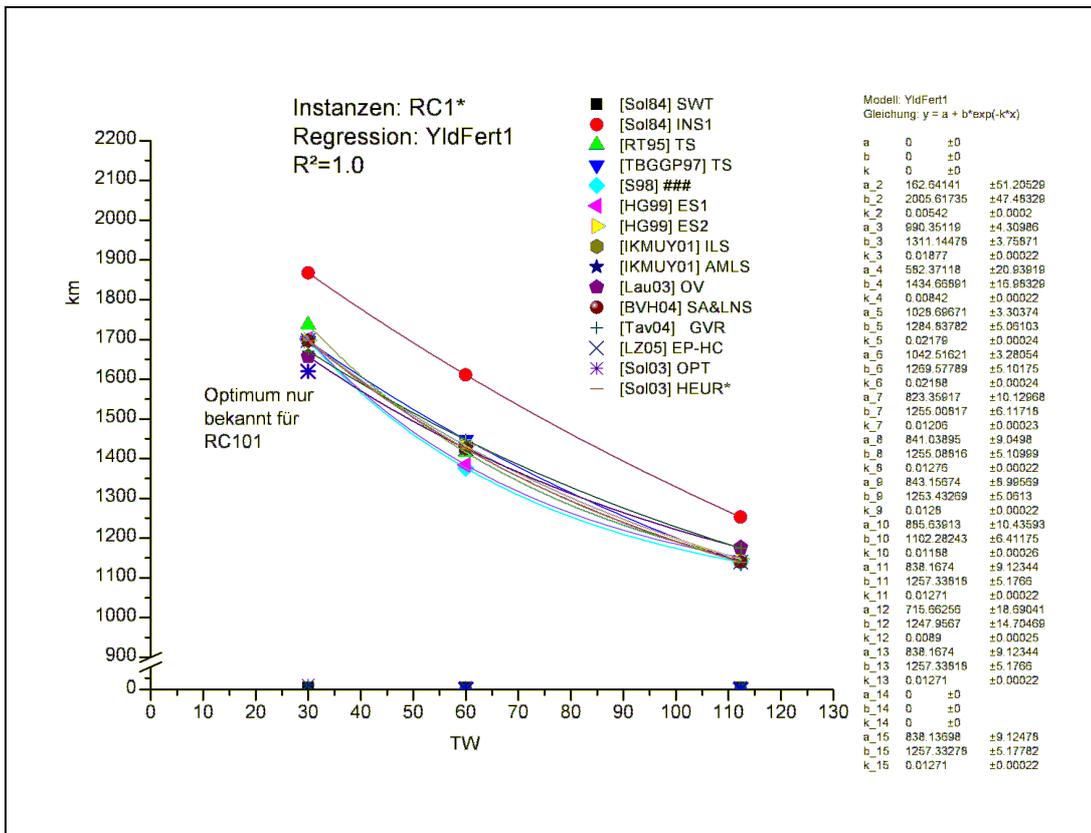


Abbildung 88: Instanz RC1* km

Das km-Optimum ist derzeit nur für RC101, nicht aber für RC106 oder RC108 bekannt (s. [Solomon 2003]), daher ist nur ein Wert * für das Optimum in beiden Grafiken dargestellt.

Bezüglich des Zusammenhangs „Fahrzeuganzahl ⇔ km“ wird auch hier wieder deutlich, dass „optimale“ Tourkilometer durch ein „Mehr“ an eingesetzten Fahrzeugen erkauf werden können.

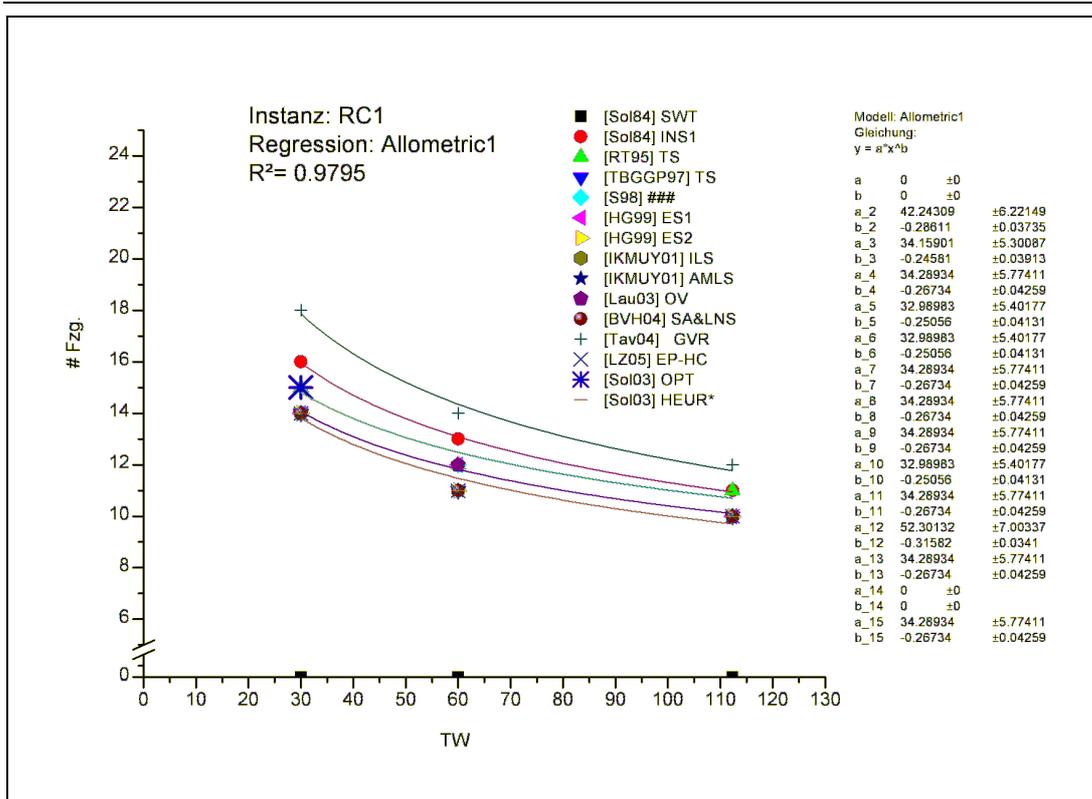


Abbildung 89: Instanz RC1* Fzg

6.4.4.3 R2*

Auch die Instanzgruppe R2* stützt die Hypothesen H1-H4 im wesentlichen. Die meisten Verfahren zeigen bzgl. der Tourkilometer einen sehr ähnlichen YldFert1-Verlauf knapp oberhalb des bekannten Optimums für R201.

Erwähnenswert sind die abweichenden fast linearen Verläufe der Verfahren [Sol 84], [RT95], [HG99] und [Lau 03]. Es ist stark zu vermuten, dass auch hier die Korrelation mit der Fahrzeuganzahl zum tragen kommt (Hypothese 4), denn ob bei $TW=480$ zwei oder drei Fahrzeuge zu Einsatz kommen, macht einen großen Unterschied.

Auch hier werden die „optimalen Tourkilometer“ * für die Instanz R201 durch den Einsatz zusätzlicher Fahrzeuge erkauft (8 Fahrzeuge statt 4 !).

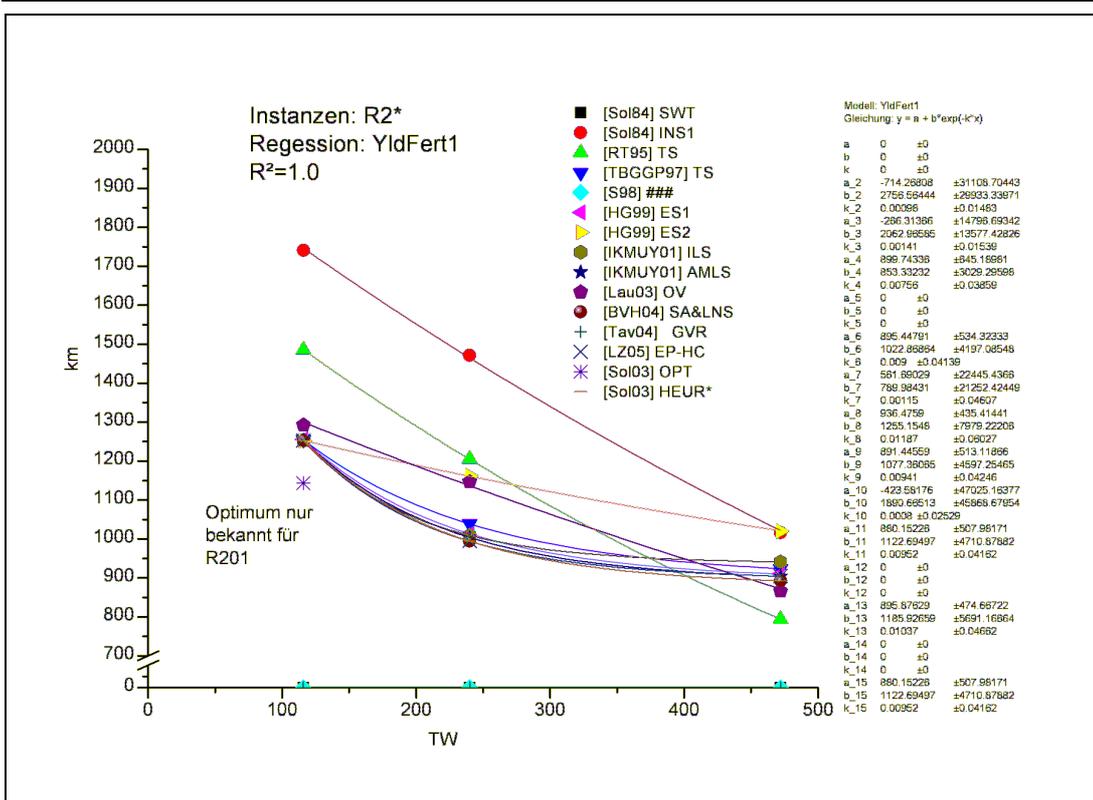


Abbildung 90: Instanz R2* km

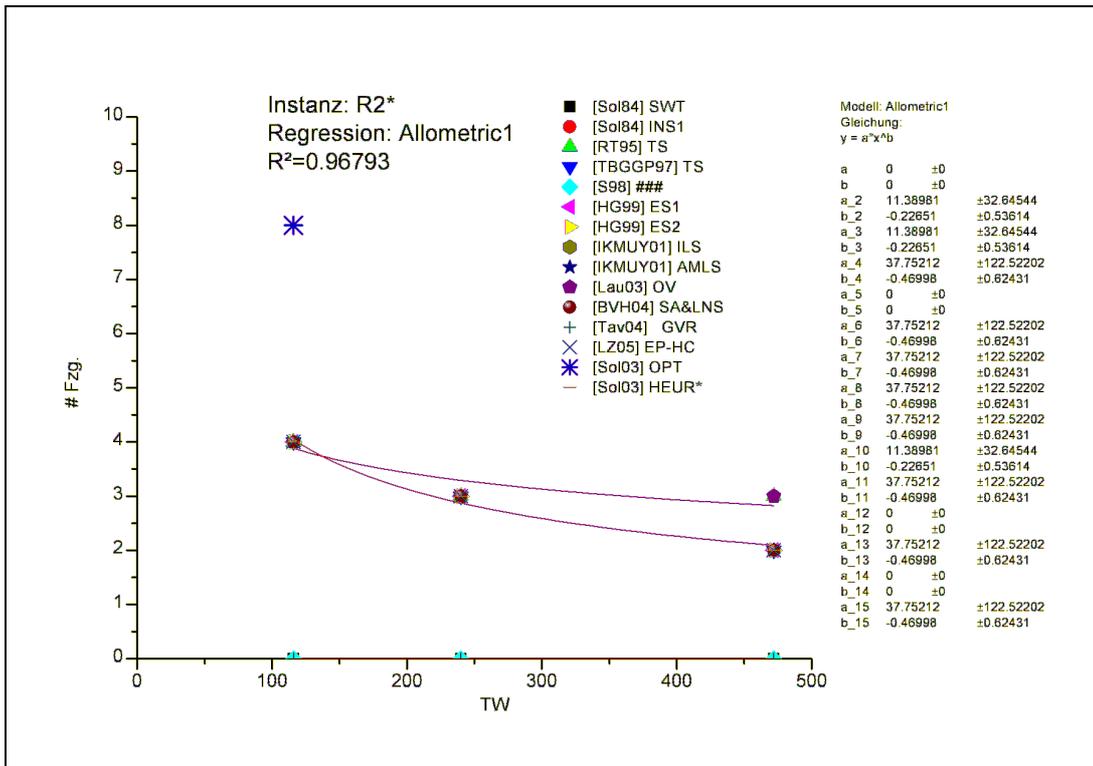


Abbildung 91: Instanz R2* Fzg.

6.4.4.4 RC2*

Auch die Instanzgruppe RC2* stützt die Hypothesen H1-H4 im wesentlichen. Einzig das Verfahren von [BVH04] zeigt nicht den erwarteten von unten konvexen Verlauf, und widerspricht daher der Hypothese H2a.

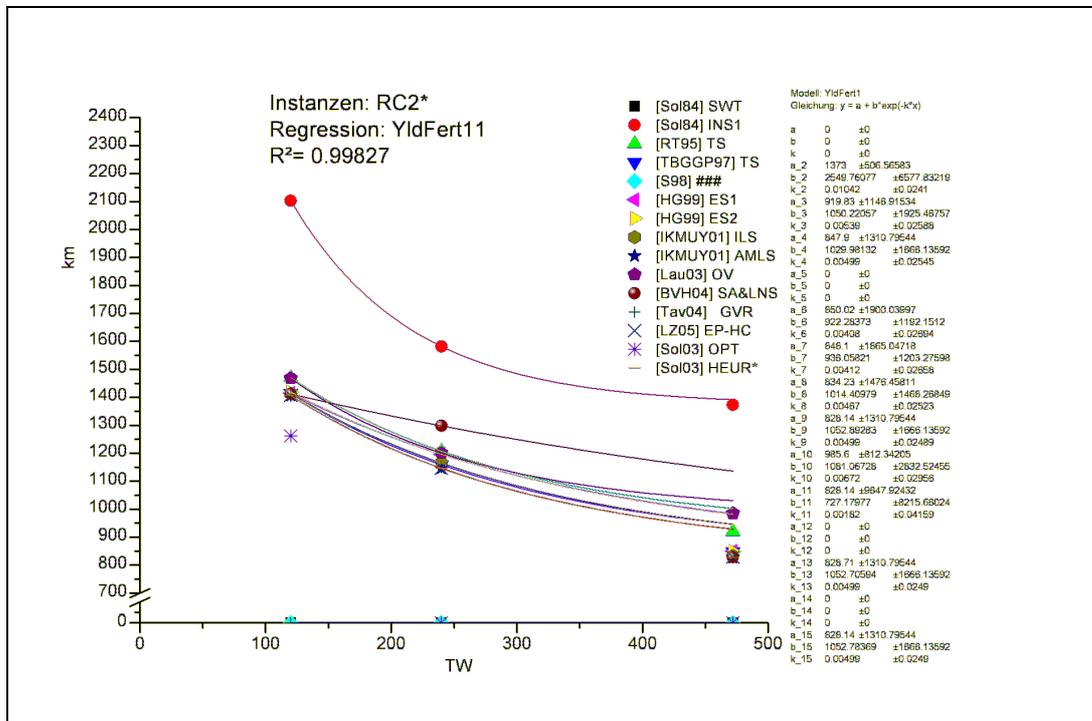


Abbildung 92: Instanz RC2* km

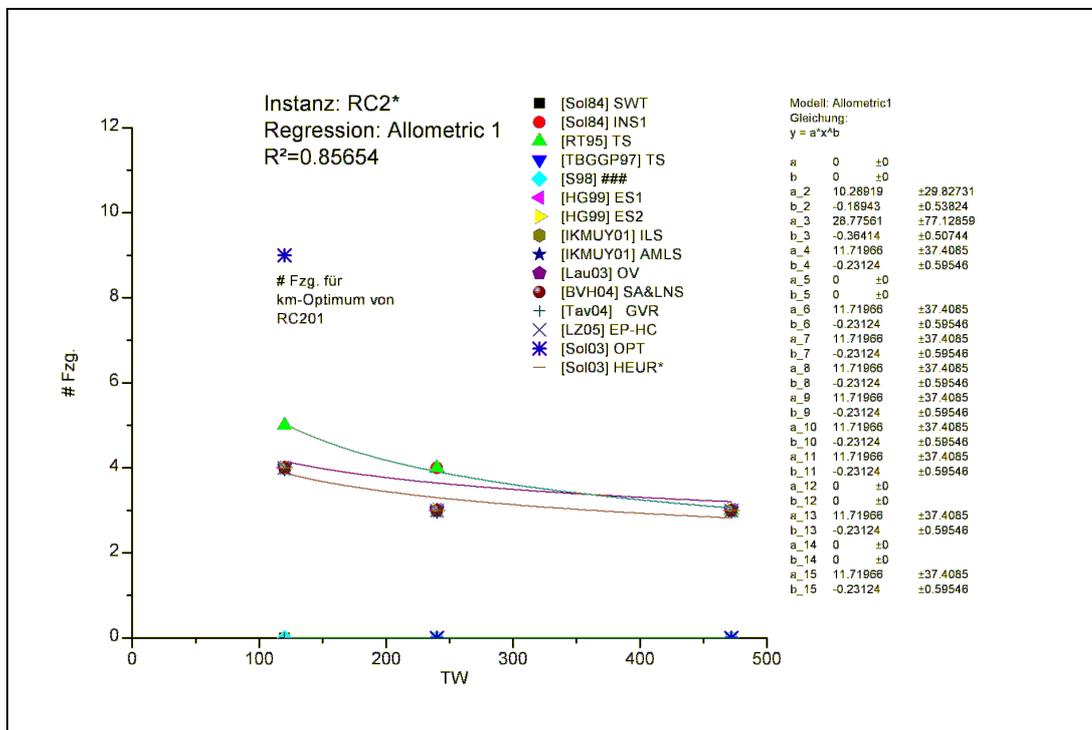


Abbildung 93: Instanz RC2* Fzg

6.4.4.5 C1* & C2*

Die Instanzen C1* und C2* bestätigen die einschränkende Bedingung der Hypothesen H1a und H1b „..., der keine ex ante erkennbare implizite Tourstruktur aufweist ...“.

Sowohl C1* als auch C2* weisen eine sehr deutliche implizite Tourstruktur auf. Daher ist der Verlauf der Tourkilometer und der der benötigten Fahrzeuge unabhängig vom eingesetzten Verfahren fast konstant.

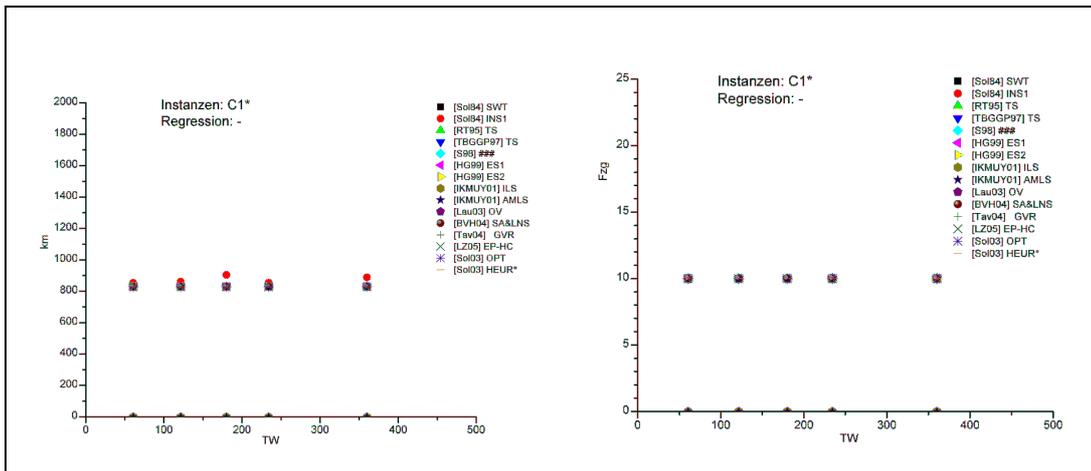


Abbildung 94: Instanz C1* km & Fzg

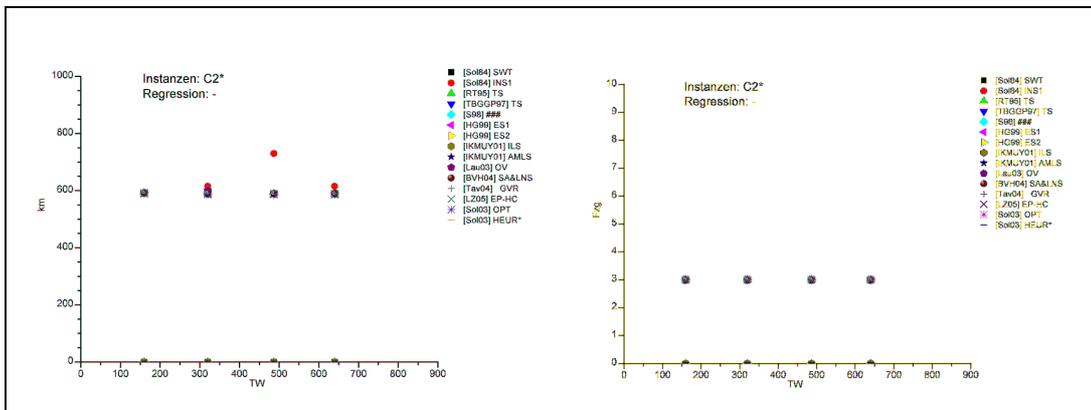


Abbildung 95: Instanz C2* km & Fzg

6.4.5 Zusammenfassung

In einfachen Worten ausgedrückt ist das wichtigste Ergebnis der Untersuchung zum FPA/PPA, dass, egal, wie gut die Verfahren zur Lösung des aufgestellten VRSP-TW sind, das dahinterliegende **Problemparadigma** (nämlich das **Ansteigen der Transportkosten bei enger werdenden Zeitfenstern**) bestehen bleibt. Dieses Paradigma ist unabhängig vom eingesetzten Verfahren. Es ist auch unabhängig von der untersuchten Instanz, solange es **keine implizite Tourstruktur** in der

durch den Filter angenommenen Auftragsmenge gibt. Soll dieses Paradigma durchbrochen werden, so muss ein völlig anderer Problemlösungsansatz verfolgt werden¹³¹.

Methodisch wurde im vorlaufenden Kapitel ausgehend von eigenen Messungen 4 Hypothesen aufgestellt. Die Hypothesen wurden an eigenen Datensätzen und an aus der Literatur verfügbaren Datensätzen validiert. Es gab bei der Validierung keine Hinweise darauf, dass die Hypothesen nicht halten. Ein Beweis der Gültigkeit der Hypothesen konnte allerdings nicht erbracht werden.

Eine weitergehende Erhärtung oder Widerlegung der aufgestellten Hypothesen ist die Aufgabe weiterführender Forschung.

6.4.6 Methodenkritik

Diese Hypothesen wurden lediglich an einzelnen Beispielen und anhand von in der Literatur veröffentlichter Ergebnisse validiert. Eine darüber hinausgehende Validierung der Hypothesen mit statistischen Methoden wurde aus folgenden Gründen nicht betrieben:

Überlagerung von Problem- und Verfahrensspezifika

(=> Hypothese 3)

Aus den veröffentlichten Ergebnissen zu den verschiedenen Verfahren wird klar, dass sich Problem- und Verfahrensspezifika überlagern. Eine nachhaltige Untersuchung der Problemspezifika mit heuristischen Methoden¹³² muss daher mehrere Verfahren umfassen, damit die Problemspezifika hervortreten¹³³. Damit müssten weitergehende Messungen sehr umfangreich angelegt werden (mehrere Probleminstanzen × mehrere Verfahren).

Interdependenz zwischen Tourkilometern und Fahrzeuganzahl

(=> Hypothese 4)

Ohne das zugrundeliegende Wichtungsproblem zu lösen, bleibt es schwierig ein „Gesamtoptimum“ zu finden. Bei Betrachtung der beiden Größen „nebeneinander“ (wie hier geschehen) müssen einzelne Ergebnisse bezüglich der einen Größe oft mit den Ergebnissen bezüglich der anderen Größe abgeglichen werden¹³⁴. Auf dem Validierungsniveau dieser Arbeit ist dies noch durch Einzelfallbetrachtungen möglich. Bei weitergehender statistischer Analyse wird dieser Abgleich zunehmend schwerer.

Validierung des Kurventyps:

¹³¹ z.B. ORA s.u.

¹³² Heuristiken deshalb, da exakte Lösungen der Probleme sehr schwierig sind (s. noch nicht exakt gelöste Solomoninstanzen)

¹³³ s. Konvergenz der guten Verfahren nahe den exakten Lösungen bei den Solomon Instanzen

¹³⁴ s. beste bekannte heuristische Lösung zu RC2* in Bezug auf km vs. dazu notwendiger vergleichsweise hoher Fahrzeugbedarf.

(=> Hypothese 2)

Die Anzahl der möglichen Regessionsfunktionen, die mit wachsendem TW asymptotisch verflachen, ist fast beliebig groß.

Die Anzahl der Regressionsfunktionsparameter von 2 bzw. 3 scheint auf Grund der eigenen Messwerte ausreichend, um ein gutes Fitting zu erreichen.

Die zur Validierung verfügbaren Literaturergebnisse weisen hingegen nur 3 bzw. 4 Stützpunkte auf. Bei 2 bzw. 3 Parametern sind dies für eine harte Validierung zu wenige Stützpunkte. Um hier härtere Ergebnisse zu erzielen, müssten auch hier weitergehende Messungen sehr umfangreich angelegt werden (insbesondere müssten die relevanten Verfahren implementiert und richtig parametrisiert werden).

Nach Ansicht des Autors ist die Aussageschärfe, die sich mit der gewählten Methodik erzielen lässt, ausgeschöpft. Verfeinerte statistische Methoden ließen sich zwar anwenden - damit diese aber auch einen zusätzlichen wissenschaftlichen Erkenntnisgewinn bringen, müsste vorher eine deutlich breitere Datenbasis geschaffen werden. Außerdem ist die Formulierung der Hypothesen eine Gratwanderung zwischen Trivialität und Falschaussage. Wenn in den Hypothesen beispielsweise von „verfahrensunabhängig“ gesprochen wird, so gilt dies nur für dem Problem angemessene Verfahren. Es lassen sich sicher leicht irgendwelche Tourenbildungsverfahren entwerfen, die beliebig schlechte Ergebnisse erzielen, und den Hypothesen widersprechen. In diesem Sinne ist die Formulierung der Hypothesen angreifbar. Neben dieser prinzipiellen Überlegungen gibt es auch noch kritikwürdige Details am Vorgehen:

- ▶ Die Wahl der selbstdefinierten Messpunkte erfolgte „intuitiv“, und nicht nach formalen DOE¹³⁵-Gesichtspunkten.
- ▶ Die Wahl der Messpunkte bei Auswertung der Literaturergebnisse richtete sich nach dem, was verfügbar war.
- ▶ Beim einfachen Fitting wurden alle Messpunkte gleichgewichtet (=> Messpunktauswahl beeinflusst Ergebnisse)
- ▶ Beim multiplen Fitting wurden alle OR-Verfahren bei der Berechnung des R^2 gleichgewichtet (=> OR-Verfahrensauswahl beeinflusst Ergebnisse)
- ▶ Beim 2 parametrischen Fitting ohne Konstante (z.B. Allometric1) müssen die beiden Fittingparameter nicht nur die Krümmung, sondern auch das asymptotische Verhalten¹³⁶ im Wertebereich modellieren. Daher ist der Wertebereich, für den das Fitting verlässliche Werte liefert nach rechts begrenzt.
- ▶ Ob die vorgeschlagenen Regressionen die bestmöglichen sind, wurde nicht formal untersucht.

¹³⁵ DOE:= design of experiments

¹³⁶ horizontale Asymptote für $TW \rightarrow \infty$

6.5 Der FPA/REA-Ansatz (reduzierter Service)

Der „Reduzierter Erfüllungsgrad Ansatz“ (REA) als Prozessschritt nach dem Filter lässt bei Ressourcenknappheit Verletzungen der zugesagten Zeitfenster zu (ERG2-Reduktion). Es gibt mehrere Möglichkeiten, wie diese Restriktionsverletzungen „genutzt“ werden können.

Zum einen können alle Zeitfenster vor der Planung um x Minuten aufgeweitet werden (Relaxation ex ante => VRSP-TW).

Zum anderen können „einiger wenige“ Zeitfenster innerhalb des Verfahrens bei Bedarf aufgeweitet werden (Relaxation ex inter => VRSP-TW mit kontextsensitiver TW-Nebenbedingungsaufweichung).

Zum dritten kann ein Tourenplanungsverfahren eingesetzt werden, dass die Zeitfensterverletzung in der Zielfunktion mit Strafkosten versieht. (Soft time windows => VRSP-STW)

6.5.1 FPA/REA als Relaxation ex ante

Setzt man den FPA/REA planungstechnisch so um, dass die Zeitfenster nach dem Filter und vor der Planung entsprechend relaxiert werden, so lässt sich der FPA/REA vollständig auf den FPA/PPA zurückführen. Die Ergebnisse können je nach „erlaubter“ ZF-Verletzung direkt aus den obenstehenden Diagrammen abgelesen werden (Linksverschiebung der Kurven im Diagramm).

Somit gelten die obenstehenden Untersuchungen rein kostenseitig auch für den FPA/REA (mit entsprechendem x-Achsen-Offset). Nicht enthalten sind bei diesem Ansatz jedoch die Opportunitätskosten für die Nichteinhaltung des ERG2.

6.5.2 FPA/REA als Relaxation ex inter

Ein Ansatz, der die Zeitfensterrestriktionsrelaxation nicht ex ante vor der Planung, sondern während der Planung im Algorithmus durchführt, ist der von Lau et al [Lau, 2003] vorgeschlagene.

Lau et al haben die Idee der Zeitfensterrestriktionsaufweichung an die Optimierung der „customer packing density“ (Maximierung der Packungsdichte der Kunden auf einer Tour = Maximierung der Kunden pro Fahrzeug) gekoppelt.

Ausgehend von einer Lösung ohne Zeitfensterrestriktionsverletzung (erzeugt durch

„algorithm A“¹³⁷) reduzieren sie die Anzahl der Fahrzeuge unter Inkaufnahme von sowohl Zeitfensterverletzungen (Relaxation des EFG2), als auch von „Nichtbelieferungen“ (Relaxation des EFG1) mit dem Ziel der Fahrzeugauslastungsoptimierung („algorithm B“).

Auch wenn Lau et al verfahrensseitig keine explizite Wichtung zwischen „customer packing density“ und EFG1 & EFG2 herstellen, muss diese Wichtung jedoch genau dann erfolgen, wenn die Lösungsgüte ihrer „algorithm B“-Ergebnisse mit anderen Lösungen verglichen werden soll.

Beispiel: [Lau et al 2003] erreichen für die Instanzengruppe C1* im Mittel eine Reduktion der benötigten Fahrzeuge um 1 (= -10%) durch:

- ▶ (EFG1↓) Nichtbelieferung von durchschnittlich 2 Kunden & (EFG2↓) Öffnung aller Zeitfenster nach hinten $[t_{f_i}; t_{s_i}] \Rightarrow [t_{f_i}; \infty]$ oder durch
- ▶ (EFG1↓↓) Nichtbelieferung von durchschnittlich 5 Kunden

Die ungelöste Wichtungsproblematik macht einen quantitativen Ergebnisvergleich unmöglich. Daher wird dieser Ansatz hier nicht weiter untersucht.

6.5.3 FPA/REA mit „soft time windows“

Beim diesem Ansatz sollen „soft time windows“ den Freiheitsgrad für die Planung nach dem Filterprozess schaffen, um eine gültige Lösung zu erzielen, falls der Filter „zu viele“ Aufträge durchgelassen hat.

In diesem Fall müssen Opportunitätskosten für die Zeitfensterverletzungen definiert werden. Dazu muss sowohl ein **Opportunitätskostenmaß** $K_{opp(ERG2)}$ als auch ein **Wichtungsfaktor** ω für die zu bestimmenden Opportunitätskosten gegenüber den Transportabwicklungskosten definiert werden.

Die mathematische Definition eines Wichtungsfaktors ist einfach:

$$K_{ges} = K_{transp} + \omega K_{opp(ERG2)}$$

d.h. die Gesamtkosten (K_{ges}) setzen sich zusammen aus den Transportkosten (K_{transp}) und den gewichteten Opportunitätskosten aus der Verletzung des ERG2 ($K_{opp(ERG2)}$). ω ist dabei der entsprechende Wichtungsfaktor. So leicht es ist, diesen Faktor mathematisch zu definieren, so schwer ist es, diesen Faktor in der Praxis richtig festzulegen.

Das Opportunitätskostenmaß $K_{opp(ERG2)}$ wird aus Marketing-Sicht sowohl das Ausmaß der individuellen Zeitfensterverletzungen berücksichtigen müssen, als auch die Anzahl der Zeitfensterverletzungen insgesamt. Eine Vorschlag für das

¹³⁷ Die Ergebnisse aus Ihrem Verfahren „algorithm A“ (ohne Zeitfensterrestriktionsrelaxation) bzgl. der Solomoninstanzen sind in dieser Arbeit bereits dargestellt worden (s.o. PPA-Ansatz).

Opportunitätskostenmaß $K_{opp(ERG2)}$ könnte daher sein:

$$K_{opp(ERG2)} = \sum_i \omega_n B_i + \omega_t (\varepsilon_i)^{\omega_{t,exp}}$$

$$B_i = \begin{cases} 1 & \forall \varepsilon_i \neq 0 \\ 0 & \forall \varepsilon_i = 0 \end{cases}$$

$$\varepsilon_i = \max(0; t_i - ts_i; tf_i - t_i)$$

ω_n = Gewicht für Zeitfensterverletzung (j/n)

ω_t = Gewicht für Betrag Zeitfensterverletzung

$\omega_{t,exp}$ = Progressionsexponent

Dieses Maß berücksichtigt sowohl die Anzahl der Zeitfensterverletzungen, als auch deren jeweilige individuelle Dauer (ggf. mit Progression).

Rein mathematisch lässt sich eine solche Formel leicht definieren und in eine lexikografische oder gewichtete Zielfunktion einbauen, aber auch hier stellt sich in der betrieblichen Praxis das Wichtungsproblem („Sollen lieber 2 Kunden um je 10 Min. verspätet beliefert werden, oder lieber nur ein Kunde mit 20 Min. Verspätung?“).

Da es in praxi kaum möglich ist, das Wichtungsproblem auf Grund von Problemdomänenkriterien eindeutig zu lösen, sind alle diesbezüglichen Optimierungsansätze einer gewissen „Wichtungswillkür“ unterworfen.

Eine nähere Untersuchung des FPA/REA mit „soft time windows“ erfolgt an dieser Stelle daher nicht.

6.6 Der ORA Ansatz (gesteuerte Nachfrage)

Bei allen bisherigen Ansätzen wurde ein wichtiges Optimierungspotenzial nicht ausgeschöpft, nämlich die „Flexibilität der Kunden bezüglich der tageszeitlichen Lage der Zustellungszeit bei unveränderten Anforderungen an den ERG2“, oder kurz die Toleranz (TLRZ). Umgangssprachlich heißt Toleranz aus Kundensicht formuliert: „Egal ob die Lieferung um 11:00 Uhr oder um 14:30 Uhr kommt – Hauptsache, die zugesagte Zustellzeit wird pünktlich erfüllt“.

Die Nutzung dieses Potenzials durch Steuerung der zugesagten Zeitfenster in einem Dialog mit dem Kunden ist die Kernidee des ORA-Ansatzes.

Im Folgenden werden die Eigenschaften des ORA in Bezug auf verschiedene Parametrisierungen und Nachfragesituationen näher untersucht.

6.6.1 Basisszenario SZ0

Die Vielzahl der zu untersuchenden interagierenden Größen lässt es sinnvoll erscheinen, ein Basisszenario zu definieren, um die ceteris paribus Einflüsse der einzelnen Einflussfaktoren herauszuarbeiten, bevor Untersuchungen zu den Transportkosten vorgenommen werden.

Def. Nachfrage:

Unter Nachfrage sei im Folgenden die Gesamtheit aller Lieferwünsche definiert. Die Nachfrage wird dabei insbesondere beschrieben durch die Größen: geografisch-zeitliche¹³⁸-Lage der Nachfrage und die erforderliche Standzeit (Servicezeit), die zur Abwicklung der Lieferung am Zustellort erforderlich ist.

Def. Angebot:

Das Angebot beim ORA-Ansatz sei im Folgenden die Gesamtheit aller zur Verfügung stehenden Fahrzeugressourcen, und deren (bei ORA ex ante definierten) Einsatzstruktur. Das Angebot wird dabei insbesondere beschrieben durch die Größen: Anzahl Fahrzeuge, Geschwindigkeit der Fahrzeuge, Tourstruktur

Def. Basisszenario SZ0:

Nachfrage:

ZUE500G mit 35-36Stops/h;

Standzeit/Stop: 10 min.

Angebot:

PRÄZ=30min. (Präzision der Terminusage)

Transportressourcen: 6 Fahrzeuge; mit je 3 Einsätzen à max. 4,5 h Tourdauer;

¹³⁸ Die zeitliche Lage der Nachfrage sei beschrieben als ein „Wunschlieferzeitpunkt“.

Planungsbasis: 24 geografische Aggregate
 (pro Tour werden 4 Aggregate bedient)
 Geschwindigkeit: 20km/h (Die Entfernungsrechnung basiert dabei auf
 Luftlinien. Die Fahrzeiten werden minutengenau gerundet.)

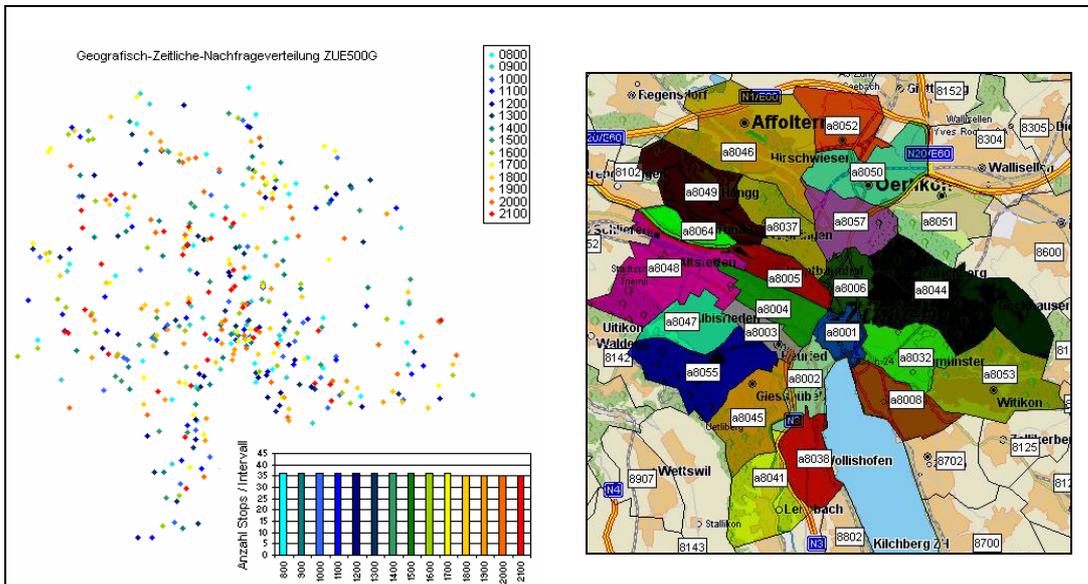


Abbildung 96: Nachfrage: Raum-zeitliche Verteilung

Angebot: Postleitzahl-basierte Aggregatstruktur

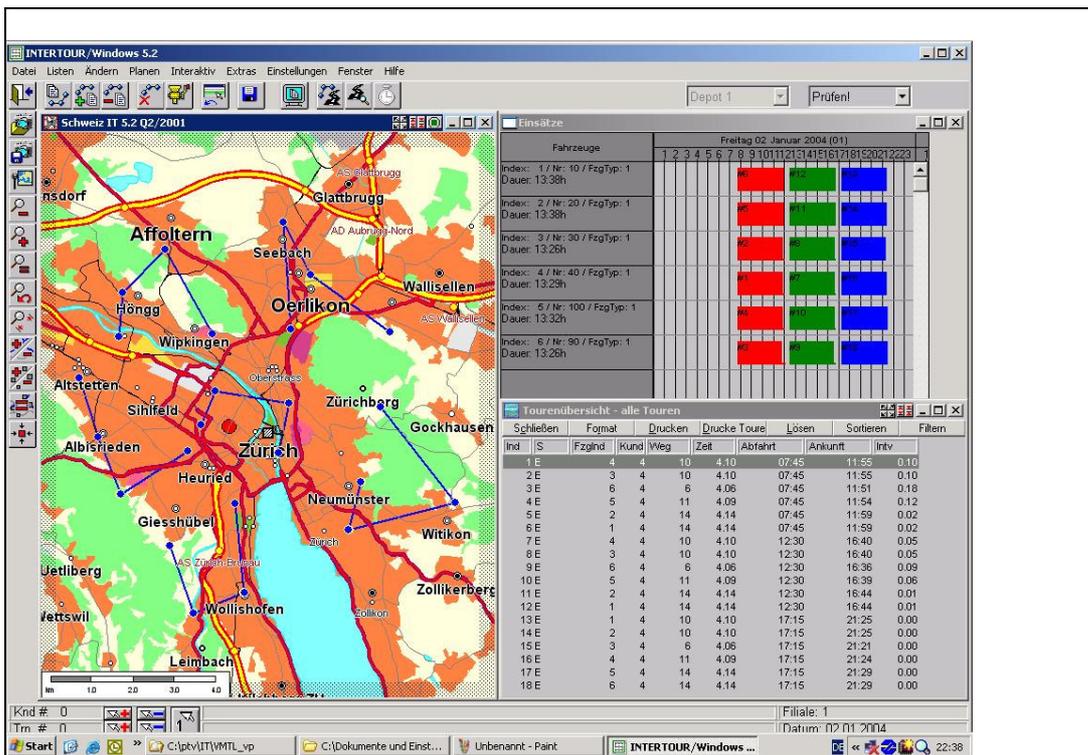


Abbildung 97: Angebot: Tourstruktur des Basisszenarios (Aggregate durch Punkt repräsentiert) – jede Tour wird 3x täglich durchgeführt

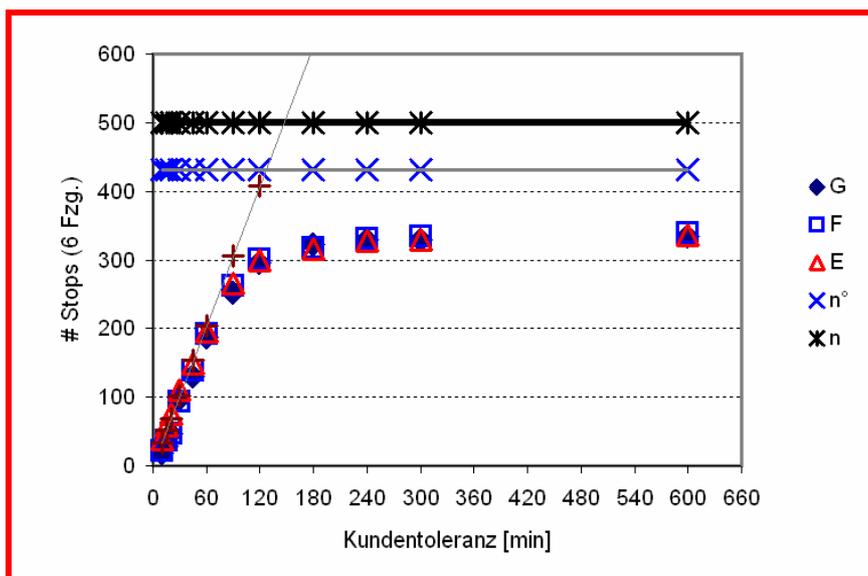
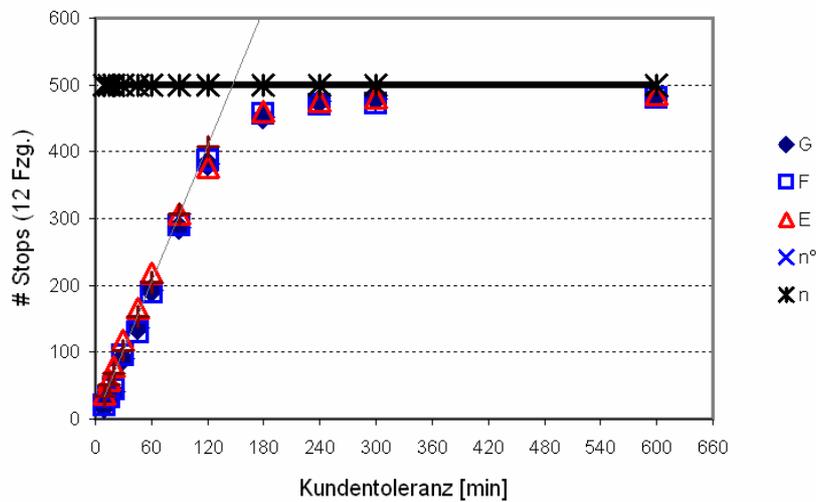
Ausgehend vom Basisszenario SZ0 lassen sich nun ceteris paribus Untersuchungen durchführen.

6.6.2 Auswirkungen der Kundentoleranz

Die Kundentoleranz als Nachfragegröße ist ein Maß für die Flexibilität der Kunden hinsichtlich ihres bevorzugten Lieferzeitwunsches. Da ORA auf der Idee basiert, diese Toleranz als Optimierungspotenzial zu nutzen, sollen hier die Auswirkungen der Toleranz im Kontext mit verschiedenen anderen Einflussgrößen untersucht werden.

6.6.2.1 Toleranz & Sättigung

In den folgenden 4 Grafiken wird die Abhängigkeit der bedienten (bzw. abgelehnten) Kunden von der Nachfrage-Einflussgröße Toleranz für unterschiedliche Angebote dargestellt. Das Basisszenario SZ0 ist rot eingrahmt.



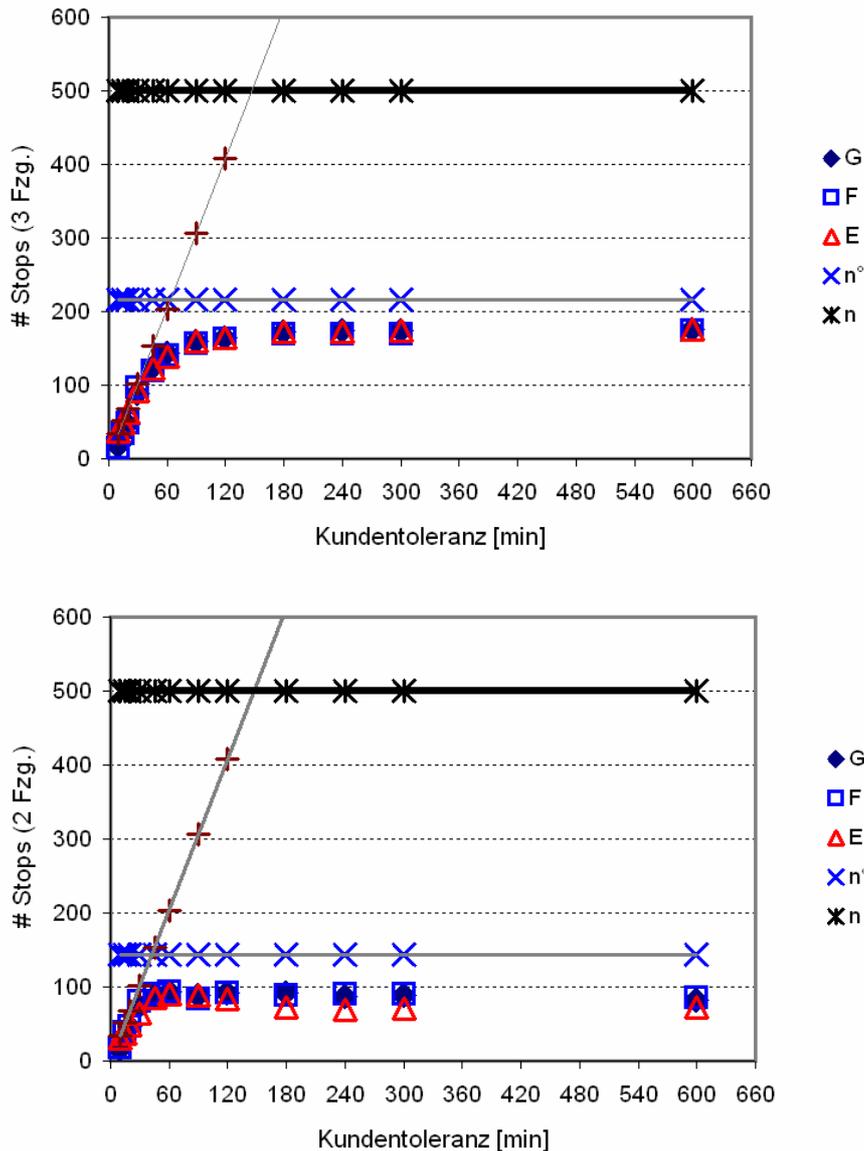


Abbildung 98: Sättigungsverhalten

Die Grafiken zeigen jeweils die ORA-Ergebnisse bzgl. der 3 Instanzen ZUE500E, F und G. Die Anzahl verplanter Stops ist naturgemäß nach oben bei 500 begrenzt, da die Nachfrage-Instanzen $n=500$ Lieferanfragen enthalten. Eine weitere obere Schranke ist n° . n° zeigt die maximale Anzahl der durch die jeweils verfügbaren Fahrzeuge bedienbaren Stops für den Fall, dass nur Standzeiten, aber keine Fahrzeiten anfallen (d.h. pro Fahrzeug max. 6 Stops / h, falls die Standzeit je 10 min. beträgt.)

Aus Sicht der verfügbaren Fahrzeuge kann man die Kurvenverläufe als „Sättigungskurven“¹³⁹ (bzgl. der Auslastung der verfügbaren Fahrzeuge) interpretieren.

¹³⁹ Definition von „Sättigung“, „Übersättigung“ und „Untersättigung“ s. 5.3.2.

Die jeweils waagrecht verlaufende rechte Seite der Kurven zeigt die ORA-verfahrensbedingte Obergrenze der Fahrzeugauslastung auf, d.h. selbst bei steigender Toleranz der 500 Kunden bzgl. Abweichungen vom Wunschlieferzeitpunkt (dies entspricht einer steigenden raum-zeitlichen Nachfrage) können nicht mehr Stops verplant werden – der Fuhrpark ist „voll ausgelastet“. („Übersättigungsbereich“)

Im linken Bereich, wo die Kurven linear mit der Steigung von ca. 3 ansteigen ist augenscheinlich die raum-zeitliche Nachfrage (definiert als Tupel von Ort, Wunschzeitpunkt & Toleranz) kleiner als das Angebot. („Untersättigungsbereich“)

Als Indiz für die Güte des Verfahrens kann der Abstand der waagrecht Sättigungsasymptote zu n° gewertet werden. Dieser Abstand beinhaltet das „Optimierungspotenzial“ an Fahrzeit, das alternativ bei besserer Reihenfolgeplanung als Servicezeit genutzt werden könnte. Dieser Abstand ist sowohl prozentual als auch absolut bei 3 Fahrzeugen am geringsten.

Bevor weitere Einflussfaktoren analysiert werden, soll an dieser Stelle noch die „Sättigung“ als wichtiger Zusammenhang zwischen Angebot und Nachfrage beschrieben werden.

Die Sättigung ist zum einen maßgeblich für die Vergleichbarkeit der Ergebnisse zwischen ORA und anderen Verfahren. Zum anderen ist eine starke „Übersättigung“ ein Indiz für einen geringen ERG1 (d.h. viele Anfragen werden abgelehnt). Es sei davor gewarnt, den ORA-Ansatz mit anderen Ansätzen zu vergleichen, ohne den ERG1 zu berücksichtigen.

Aus kommerzieller Sicht ist darauf zu achten, dass ORA in einer „angemessenen“ Sättigung operiert. „Angemessen“ ist dabei von Vertriebsseite zu definieren. Bei Aufsetzen eines neuen Lieferservice kann z.B. eine Untersättigung erwünscht sein (damit Erstkunden einen hohen ERG1 und ERG2 wahrnehmen, und dies schätzen lernen), während bei einem etablierten Lieferservice aus Kostenoptimierungssicht ggf. eine leichte Übersättigung angestrebt werden könnte.

6.6.2.2 Toleranz & Aggregatsgröße

Eine wichtige Frage für die ORA –Vorplanung ist die nach einer guten Wahl der Aggregatsgröße. Es wurden verschiedene Untersuchungen vorgenommen, bei denen ein Aggregat jeweils einen, zwei oder vier Postleitzahlbereiche umfasste.

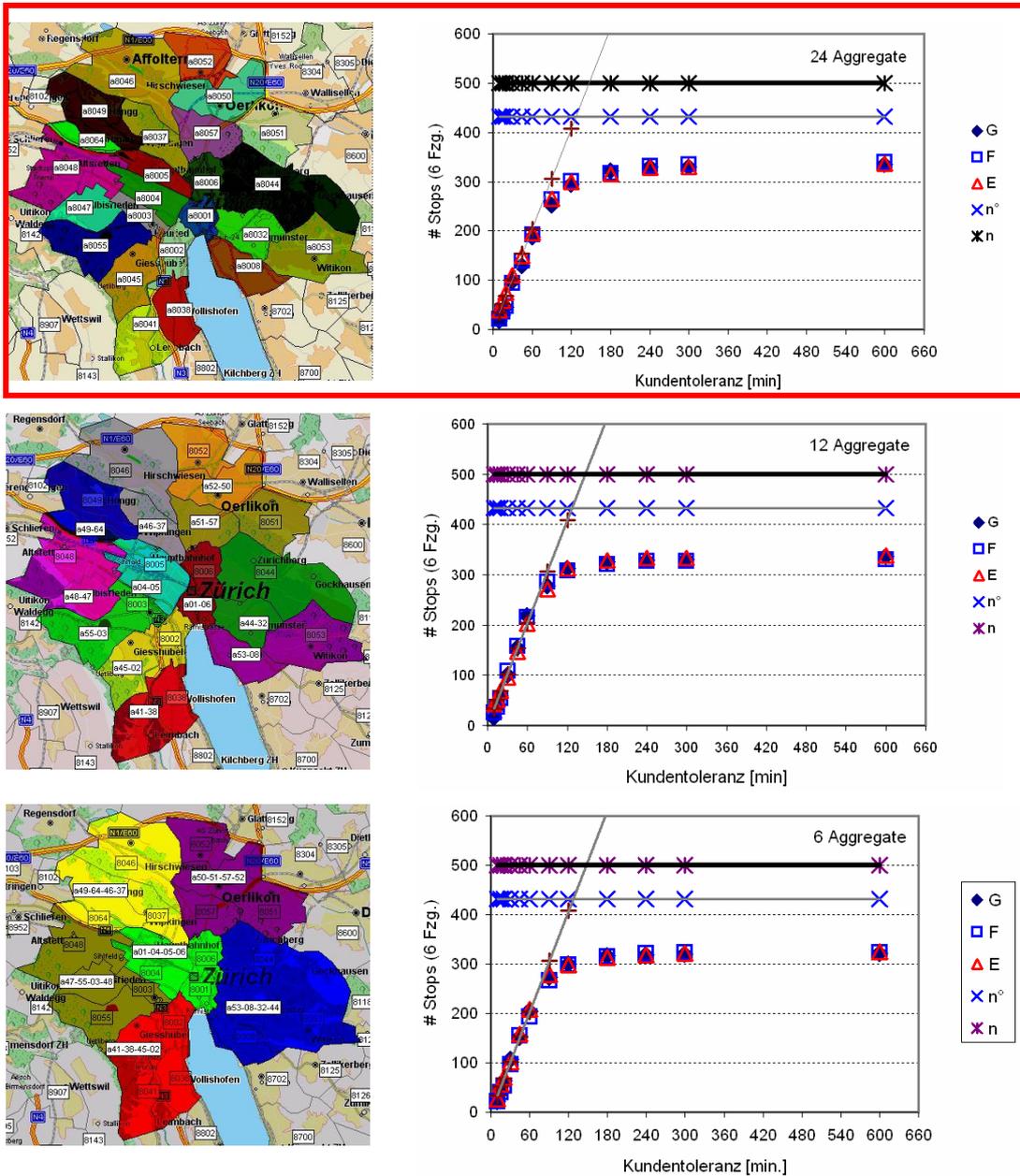


Abbildung 99: Auswirkungen der Aggregatsgröße

Der Übergang von der linearen Steigung zur waagrechten Sättigungsasymptote scheint bei wachsender Aggregatsgröße zwar einen ausgeprägteren Knick aufzuweisen¹⁴⁰, aber ansonsten scheint die Aggregatsgröße zumindest für die

¹⁴⁰ was sich durch das größere Potenzial an möglichen Einfügekandidaten pro Aggregat erklären ließe.

untersuchten Beispiele keinen signifikanten Einfluss auf den Zusammenhang von Toleranz und Anzahl bedienter Kunden zu haben.

6.6.2.3 Toleranz & raum-zeitliche Lage

Sehr interessant ist, dass die Art der raum-zeitlichen Verteilung (Instanzunterschiede von *E,*F oder *G) kaum eine Rolle spielt. Dies lässt auf eine große **Verfahrensrobustheit** bzgl. Toleranz bei raum-zeitlich gleichverteilter Nachfrage schließen.

6.6.2.4 Zusammenfassung zur „Toleranz“

Im Gegensatz zu anderen Tourenplanungsaufgaben ist die Nachfrage nicht mehr allein durch die Anzahl der Lieferwünsche, sondern auch durch die Toleranz der Kunden zu kennzeichnen¹⁴¹. Die Toleranz ist daher ein wichtiger Parameter der Nachfrage.

Es konnte gezeigt werden, dass ORA c.p. mit wachsender Toleranz eine wachsende Anzahl von Kunden bedienen kann. Dieses Wachstum ist nach oben beschränkt entweder durch die „Sättigung“ (Grenze = Fahrzeugauslastung) oder die Anzahl der zur Verfügung stehenden Lieferwünsche (Grenze = Nachfragepotenzial).

Performance-Aussagen für ORA sind daher immer auf ein Toleranzniveau zu beziehen. Um ungewollte Überlagerungen des Einflusses der Toleranz auf andere Einflussfaktoren bei c.p. Analysen auszuschließen, wird im Folgenden darauf geachtet, dass die **Toleranz im Sättigungsbereich** liegt.

Diese Einschränkung (nämlich dass die folgenden c.p. Betrachtungen nur im Sättigungsbereich angestellt werden) ist zunächst eine methodische Entscheidung. Für die Interpretation der folgenden Ergebnisse heißt das, dass diese Ergebnisse nur dann gelten, wenn Kunden eine große Toleranz aufweisen.

6.6.3 Kosteneffekte der zugesagten Zeitfenster

Nach den Untersuchungen zur Toleranz wenden wir uns nun den Kosteneffekten in Abhängigkeit von der Präzision (PRÄZ) der zugesagten Zeitfenster zu. Um nicht im „luftleeren Raum“ zu messen, wird dazu die Methode des „ceteris paribus Teilaspektvergleichs“ eingesetzt.

¹⁴¹ Aus Sicht des Serviceanbieters können „wenige Kunden mit hoher Toleranz“ und „viele Kunden mit geringer Toleranz“ ähnliche Fahrzeugauslastungsgrade erzielen. Der ERG1 unterscheidet sich in beiden Fällen jedoch signifikant.

Durch Lösung des D-m-VRSP-TW-UC-DO definiert ORA als „Teilergebnis“ auch eine Menge von zu bedienenden Kunden mit den jeweils zugesagten Zeitfenstern. Dieses „Teilergebnis“ ist vergleichbar mit dem Output des Filters im FPA/PPA-Ansatz, und es spannt ein klassisches VRSP-TW auf. In Bezug auf dieses VRSP-TW können die Ergebnisse von ORA mit klassischen Planungsansätzen verglichen werden. Der „ceteris paribus Teilaspektvergleich“ bezieht sich also nicht auf das eigentliche D-m-VRSP-TW-UC-DO, sondern nur auf das abgeleitete VRSP-TW, das von ORA selbst aufgespannt wird.

Ließe sich eine exakte Lösung des jeweils abgeleiteten VRSP-TW leicht generieren, so wäre eine scharfe Güteaussage bzgl. des VRSP-TW möglich. Da das VRSP-TW jedoch NP-schwierig ist, und selbst für die Solomoninstanzen mit jeweils nur 100 Stops bis heute¹⁴² nicht alle exakten Lösungen bekannt sind, sind exakte Vergleichsgrößen nur äußerst schwer zu erzeugen. Für diese Arbeit wurde methodisch daher eine „Obergrenze“ und eine „grobe Untergrenze“ für die Lösungsqualität des VRSP-TW generiert, die dann für den „ceteris paribus Teilaspektvergleich“ herangezogen wurde.

Als „**Obergrenze**“ für die Lösungsgüte kann bereits eine beliebige **zulässige Lösung des VRSP-TW** dienen (eine gültige Lösung ist immer eine Obergrenze für die Lösungsgüte, die optimale Lösung wird auf jeden Fall besser sein).

Eine entsprechende „Untergrenze“ für die Lösung neuer „real-life“ Probleme ist hingegen wiederum nicht einfach zu erhalten. Eine optimale Lösung des um die Zeitfensterrestriktion TW relaxierte VRSP wäre eine echte Untergrenze. Aber auch das VRSP ist noch NP-schwierig. Eine „gute“ Lösung des um die Zeitfensterrestriktion relaxierte VRSP kann zwar als „approximative Untergrenze“ herangezogen werden. Sobald die Lösung des VRSP jedoch nicht optimal ist, ist nicht mehr gewährleistet, dass sie eine wirkliche Untergrenze für das zugehörige VRSP-TW darstellt. Daher kann zwar mit hoher Wahrscheinlichkeit davon ausgegangen werden, dass eine heuristisch gefundene „gute“ **Lösung des VRSP** auch eine „**approximative Untergrenze**“ für die Lösung des zugehörigen VRSP-TW mit engen Zeitfenstern darstellt - eine Gewähr dafür gibt es jedoch nicht.

Für diese Arbeit wurde trotz dieser Bedenken die approximative Untergrenze als Vergleichsgröße verwendet, denn je enger die Zeitfenster sind, desto stärker ist die Problemrelaxation durch Weglassen der Zeitfenster.

Als Verfahren für die Generierung der Obergrenze und der approximativen Untergrenze wurde hier das **Savings**-Verfahren gewählt, da es für ein VRSP ohne TW als Standardverfahren anerkannt ist, und i.d.R. eine gute heuristische Lösung erzeugt. Außerdem kann es schnell eine gültige Lösung für ein VRSP-TW erzeugen, auch wenn es i.d.R. keine allzu guten Ergebnisse generiert.

¹⁴² Stand 10. Juni 2003 s. [Solomon, 2003]

Die Obergrenze wird im Folgenden durch **SmZF** (Savings mit Zeitfenstern), die approximative Untergrenze durch **SoZF** (Savings ohne Zeitfenster) denotiert.

Die Einheit der PRÄZ ist im Folgenden [min.].

6.6.3.1 Basisszenarioanalyse

Eine der wichtigsten Erkenntnisse wird bereits aus der folgenden Grafik deutlich: Für ORA ist sowohl die Anzahl der Fahrzeuge als auch die Anzahl der Touren konstant, da ex ante durch die Vorplanung vorgegeben.

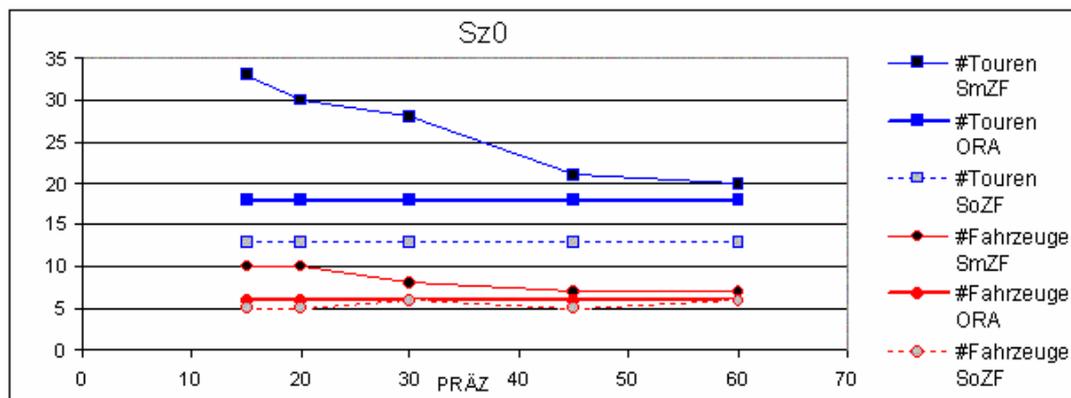


Abbildung 100: „ceteris paribus Teilaspektvergleich“ bzgl. Anzahl Touren und Anzahl Fahrzeuge für Szenario Sz0
x-Wert: PRÄZ; y-Wert: Anzahl Fahrzeuge bzw. Anzahl Touren

Außerdem wird ersichtlich, dass das durch ORA aus dem D-m-VSPR-TW-UC-DO generierte VRSP-TW in Bezug auf die Anzahl benötigter Fahrzeuge recht gut gelöst wurde (in zwei Fällen benötigt der Savings ohne Zeitfenster die selbe Anzahl von Fahrzeugen, in 3 Fällen nur eins weniger).

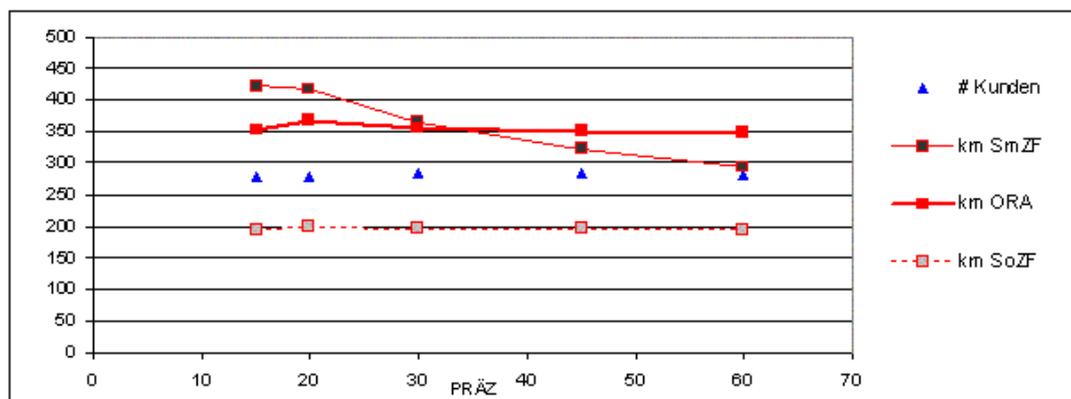


Abbildung 101: „ceteris paribus Teilaspektvergleich“ bzgl. Tourkilometern für SZ0
x-Wert: PRÄZ; y-Wert: Anzahl bedienter Kunden bzw. Gesamttourkilometer

In dieser Abbildung wird zunächst deutlich, dass die Anzahl der bedienten Kunden unabhängig von der zugesagten Präzision ist (ein Indiz dafür, dass die Toleranz sehr hoch ist – in diesem und in den folgenden Beispielen wurde die Toleranz auf +/- 900 min. festgesetzt).

Die Gesamttourkilometer sind erwartungsgemäß sowohl für die Ergebnisse von ORA, als auch für die SoZF-Vergleichsgröße nahezu konstant, und sinken erwartungsgemäß für die SmZF- Vergleichsgröße mit größer (größer) werdenden Werten für PRÄZ. Interessant ist, dass ORA für sehr genaue (kleine) PRÄZ-Werte immer noch besser ist als SmZF. Auch wenn für gröbere PRÄZ-Werte SmZF in Hinblick auf die Tourkilometer leicht bessere Ergebnisse liefert als ORA, so ist zu Berücksichtigen, dass SmZF dafür deutlich mehr Fahrzeuge benötigt¹⁴³.

Übertragen in die physische Welt bedeuten die Ergebnisse zu SZ0: Würde in der betrieblichen Praxis ORA lediglich als Filter des FPA/PPA eingesetzt, der Aufträge für maximal 6 Fahrzeuge akzeptieren sollte, und würde dann versucht mit einem Savings-Verfahren die Touren zu planen, so würde der Savings keine Lösung mit 6 Fahrzeugen finden, da das Savings-Verfahren die von ORA generierte implizite Tourstruktur in den Aufträgen nicht erkennen würde. Bei PRÄZ=60 min. würde der Savings jedoch unter Einsatz von 7 (statt 6) Fahrzeugen 60 km weniger Gesamttourkilometer generieren.

Betrachten wir nun die Verteilung der Zeitanteile (Fahrzeit, Servicezeit und Wartezeit).

¹⁴³ Hier kommt das oben diskutierte Wichtungsproblem zwischen Fix- und variablen Kosten zum Tragen.

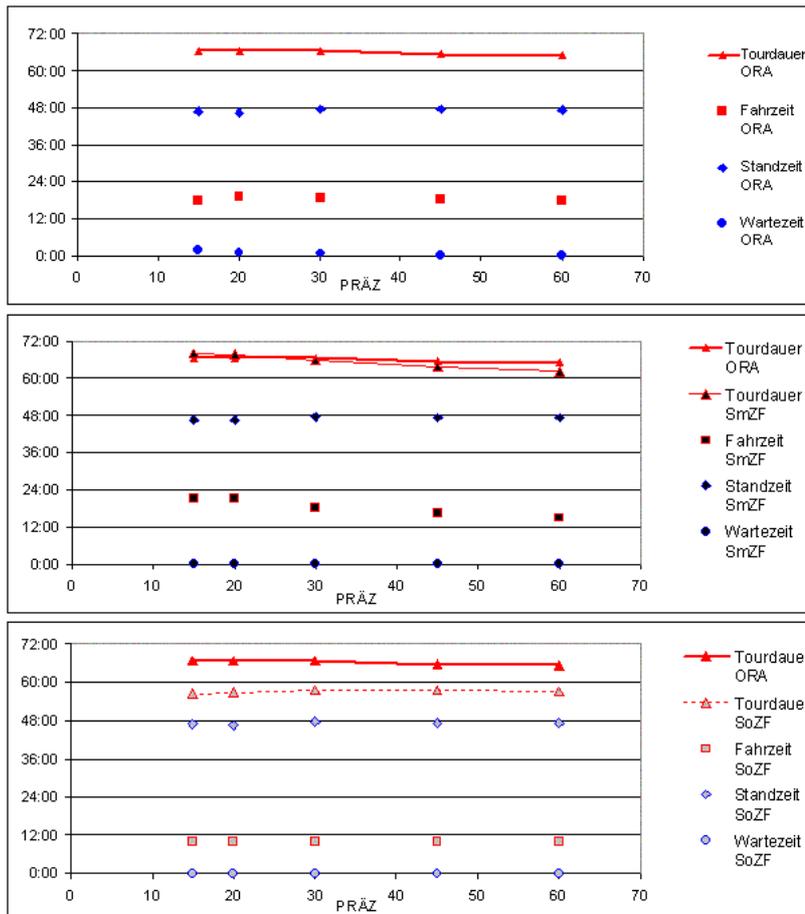


Abbildung 102: „ceteris paribus Teilaspektvergleich“ bzgl. der Zeitanteile
x-Wert: PRÄZ; y-Wert: Zeit

Interessant ist, dass bei ORA noch ein kleiner messbarer „Rest“ an Wartezeiten zu verzeichnen ist – dies ist sozusagen „ungenutztes Servicepotenzial“.

Entscheidend ist in dieser Darstellung, dass die Unterschiede in der Gesamttourdauer zwischen ORA und SoZF relativ gering ausfallen. Dies liegt u.a. daran, dass den geringen Fahrzeitanteilen relativ große Servicezeitanteile gegenüberstehen. Dieser Effekt wird weiter unten durch Variation der Servicezeitanteile und Geschwindigkeiten noch näher untersucht.

Aus den obigen Messwerten lassen sich auch die beiden KPI („key performance indicators) „durchschnittliche Anzahl Kunden pro eingesetztem Fahrzeug“ und „durchschnittlicher Gesamtzeitaufwand pro Zustellung“ ableiten.

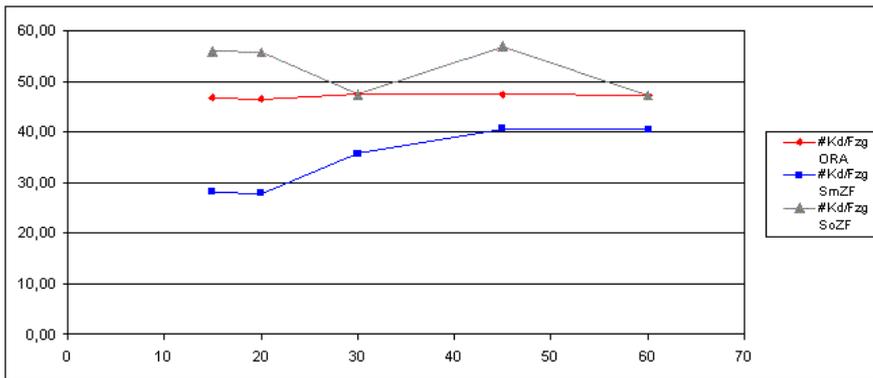


Abbildung 103: KPI durchschnittliche Anzahl Kunden / Fahrzeug
x-Wert: PRÄZ; y-Wert: Anzahl Fahrzeuge

Im Hinblick auf den KPI „Anzahl Kunden/Fahrzeug“ kommt ORA hier bereits relativ dicht an eine Planung ohne Zeitfenster heran.

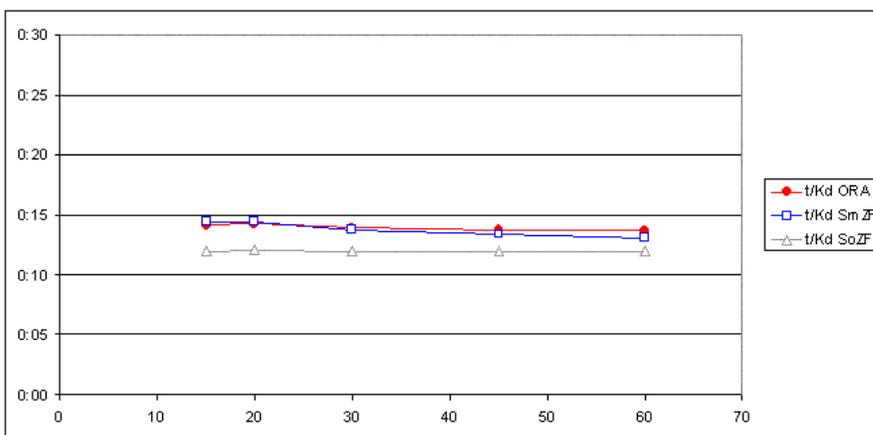


Abbildung 104: KPI durchschnittlicher Gesamtzeitaufwand pro Zustellung

Im Hinblick auf den KPI „Zeitaufwand/Kunde“ ist ORA nur bei sehr engen Zeitfenstern geringfügig besser als der Savings mit Berücksichtigung der Zeitfenster.

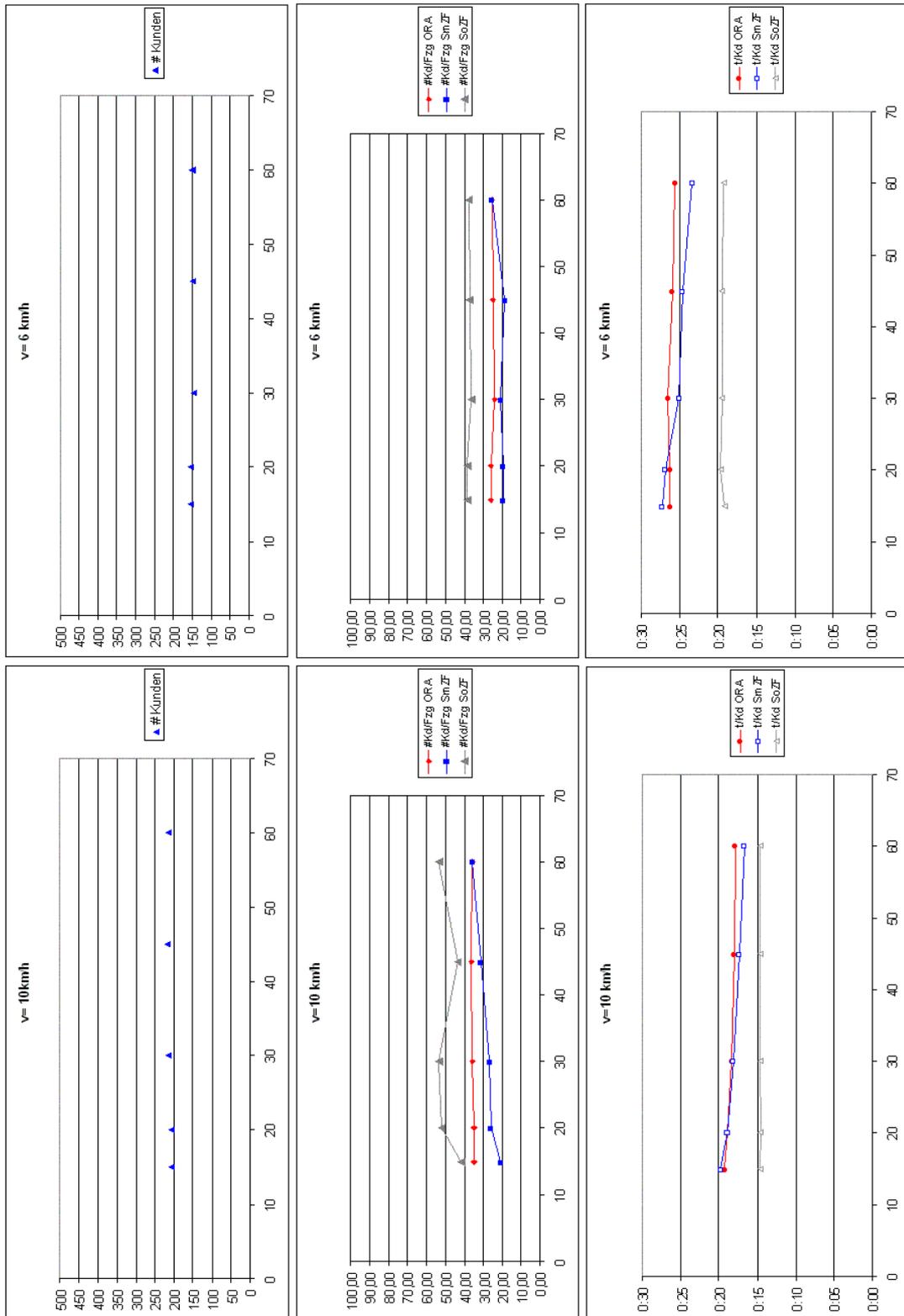
Zusammenfassend kann für das Szenario Sz0 gesagt werden:

- ▶ Bezüglich der Anzahl benötigter Fahrzeuge liefert ORA im betrachteten Wertebereich recht gute Ergebnisse.
- ▶ Bezüglich der Gesamttourdauer und der Gesamttourkilometer ist ORA hinsichtlich der PRÄZ relativ unsensibel. Je unpräziser Termine zugesagt werden, desto schlechter wird ORA daher voraussichtlich abschneiden.

Erinnerung: Die Grafiken stellen keinen Vergleich zwischen verschiedenen Ansätzen dar – sie zeigen nur „ceteris paribus Teilaspektvergleiche“ für das durch ORA generierte VRSP-TW.

6.6.3.2 Geschwindigkeit

Wie in den vorangegangenen Messungen erwähnt ist das Verhältnis von Servicezeit zu Fahrzeit eine wichtige Größe, die hier näher untersucht wird.



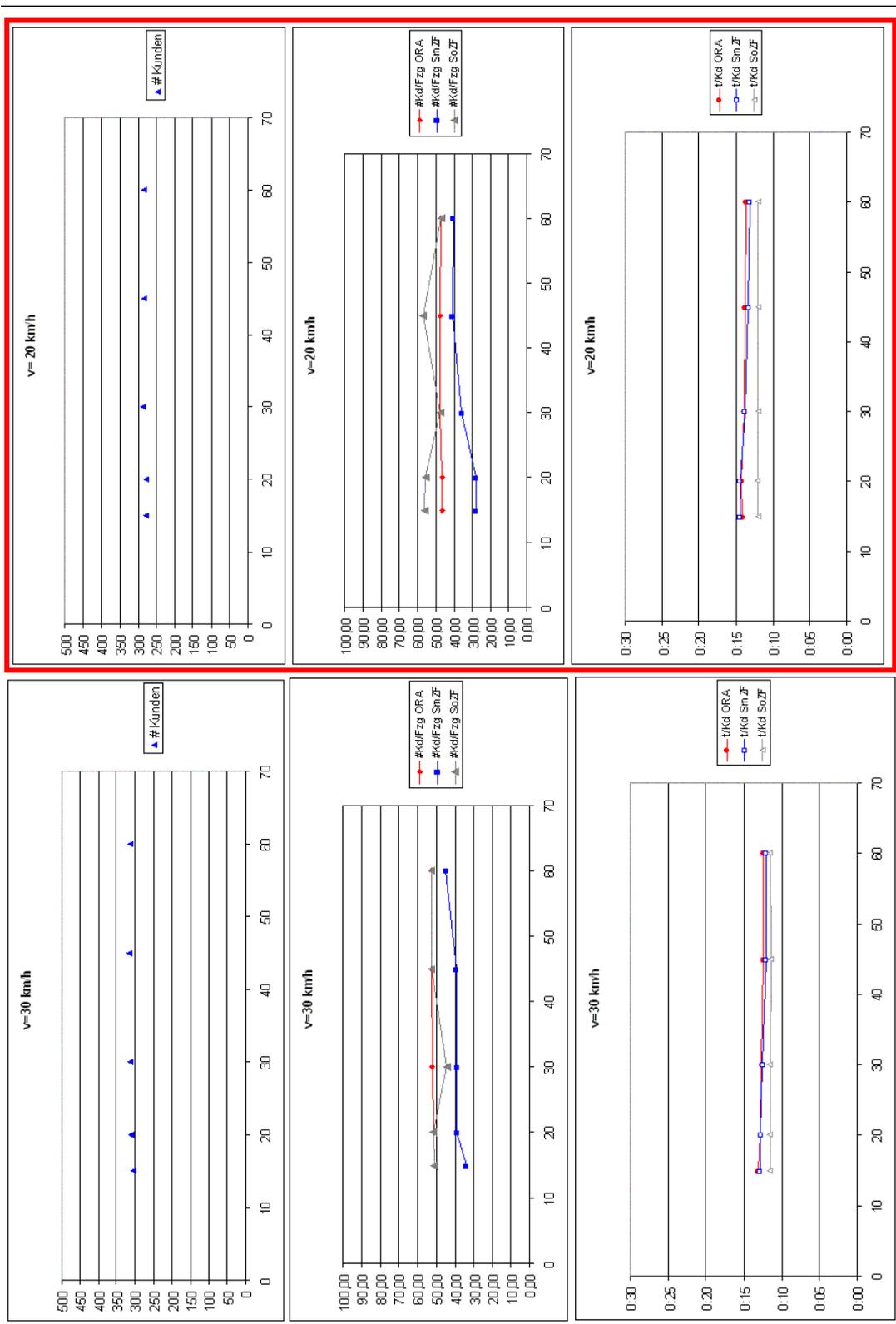


Abbildung 105: Einfluss der Durchschnittsgeschwindigkeit
 x-Achse: PRÄZ; y-Achse s. Legende

Ausgehend vom Basisszenario SZ0 (rot markiert), wurde die Durchschnittsgeschwindigkeit variiert. Erwartungsgemäß fällt mit sinkender

Geschwindigkeit die Anzahl der Lieferungen.

Mit steigender Geschwindigkeit schrumpft der Abstand zwischen den Vergleichsgrößen SmZF und SoZF.

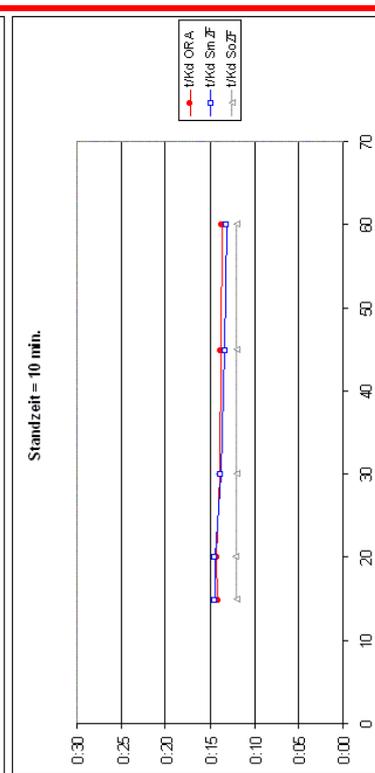
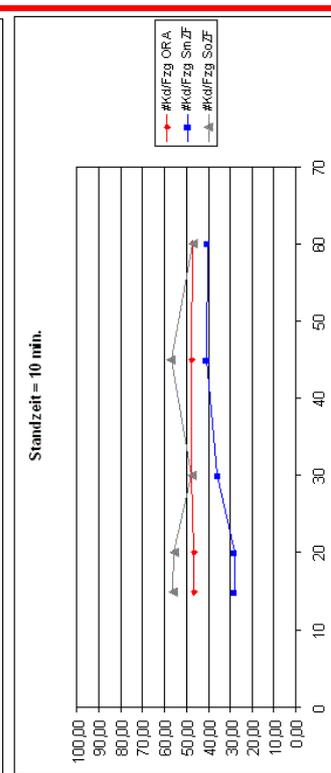
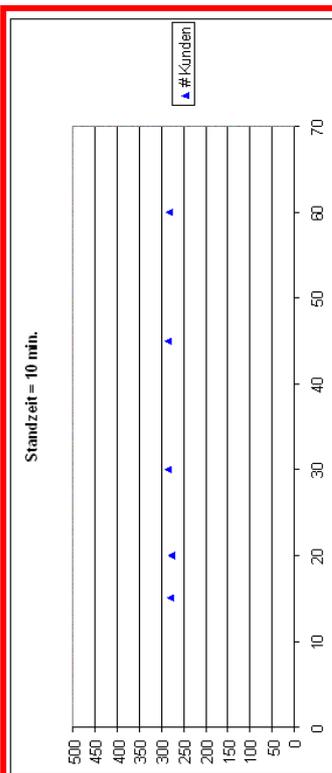
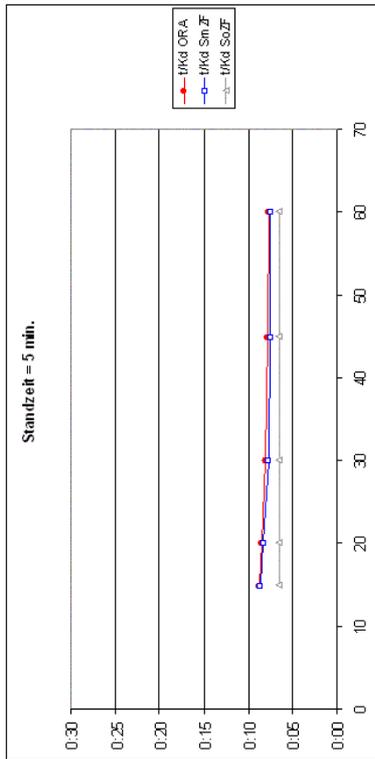
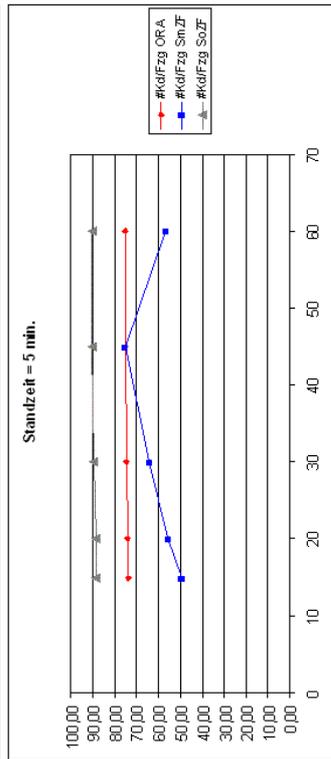
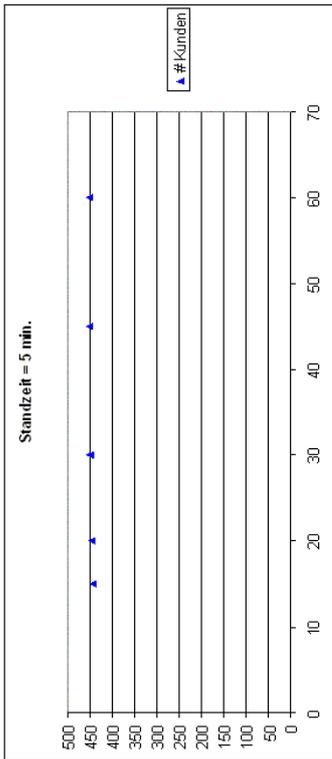
Interessant ist, dass für ORA die Kenngröße „Anzahl Kunden/Fahrzeug“ mit steigender Geschwindigkeit kontinuierlich steigt, während sie für SoZF beim Übergang von 20km/h auf 30 km/h zurückgeht. Der Grund hierfür liegt in der ganzzahligen Fahrzeuganzahl auf kleinem Niveau und der Eigenschaft des Savings die km und nicht unbedingt auch die Anzahl Fahrzeuge zu optimieren¹⁴⁴. In dem Beispiel (v=30km/h; PRÄZ=30 min.) braucht der Savings trotz Zeitfensterrelaxation ein Fahrzeug mehr als ORA.

Anmerkung: Vorsicht bei der Ergebnisinterpretation – auch wenn die Anzahl (= Mächtigkeit der Auftragsmenge) der von ORA bedienten Aufträge kaum schwankt, kann die Menge der Aufträge für jede Ausprägung von PRÄZ unterschiedlich sein. D.h. es ist nicht auszuschließen, dass das von ORA bei gröberer (=größerer) PRÄZ generierte relaxierte VRSP mehr Fahrzeuge benötigt als eins mit stärkerer (=kleinerer) PRÄZ. Insofern sind die „Schwankungen“ der Performanceindikatoren eher natürlich, und die „Konstantheit“ der ORA Performanceindikatoren das besondere.

6.6.3.3 Standzeit

Die Ergebnisse der Standzeitvariation sind in der folgenden Grafik dargestellt. Auch hier wurde ausgehend vom Basisszenario die Standzeit variiert.

¹⁴⁴ Auch hier treffen wir wieder auf das Wichtungsproblem.



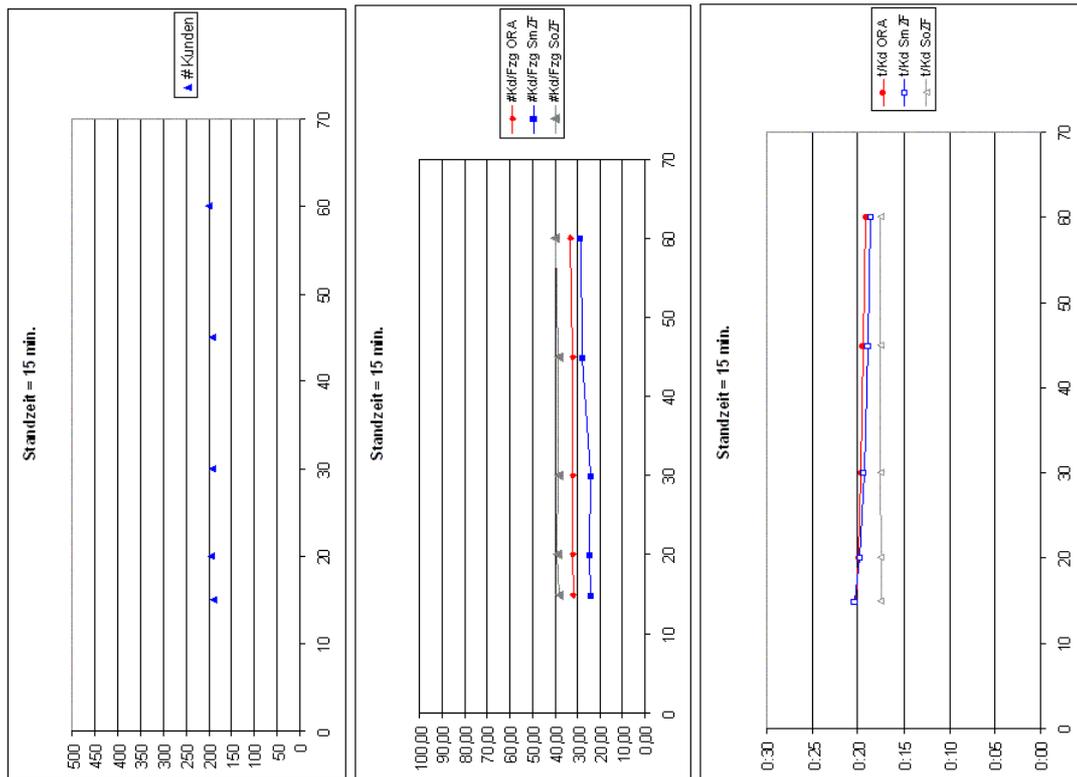


Abbildung 106: Einfluss der Standzeit (x-Achse: PRÄZ; y-Achse s. Legende)

Veränderte Standzeiten wirken sich erwartungsgemäß sehr ähnlich aus wie Änderungen der Geschwindigkeit.

Anmerkung: bei einer Standzeit von 5 min. operiert ORA nicht im Sättigungsbereich¹⁴⁵. Daher zeigt auch die SoZF-Vergleichsgröße bei einer Standzeit von nur 5 min. eine deutlich bessere Performance.

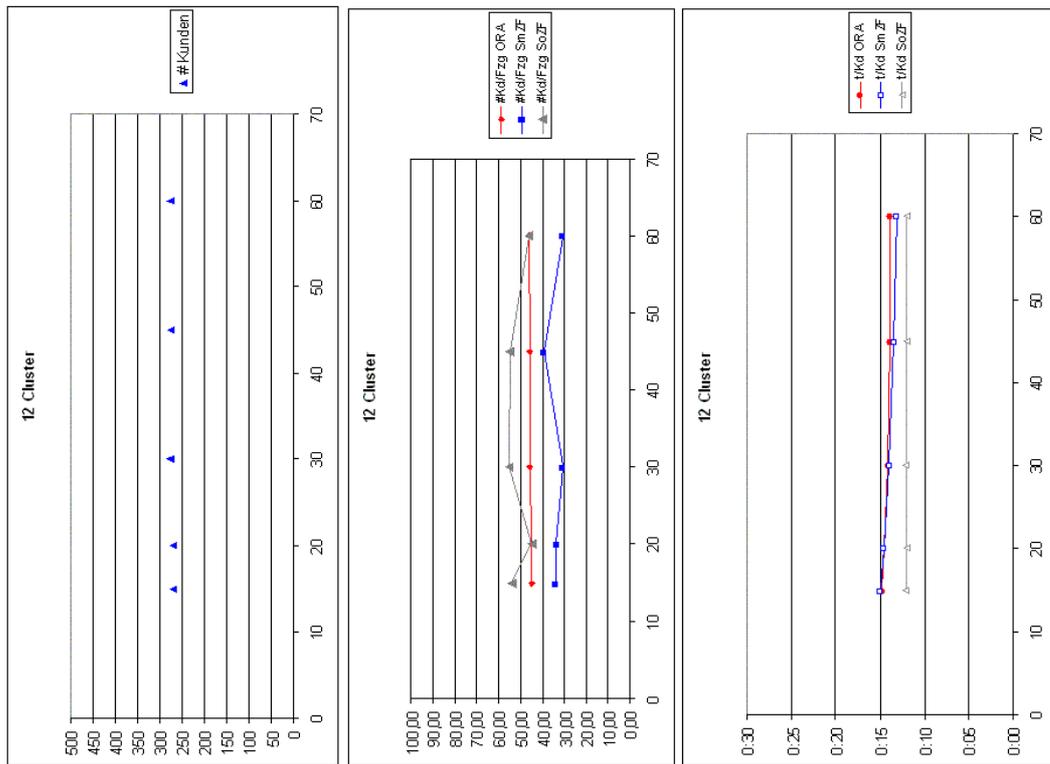
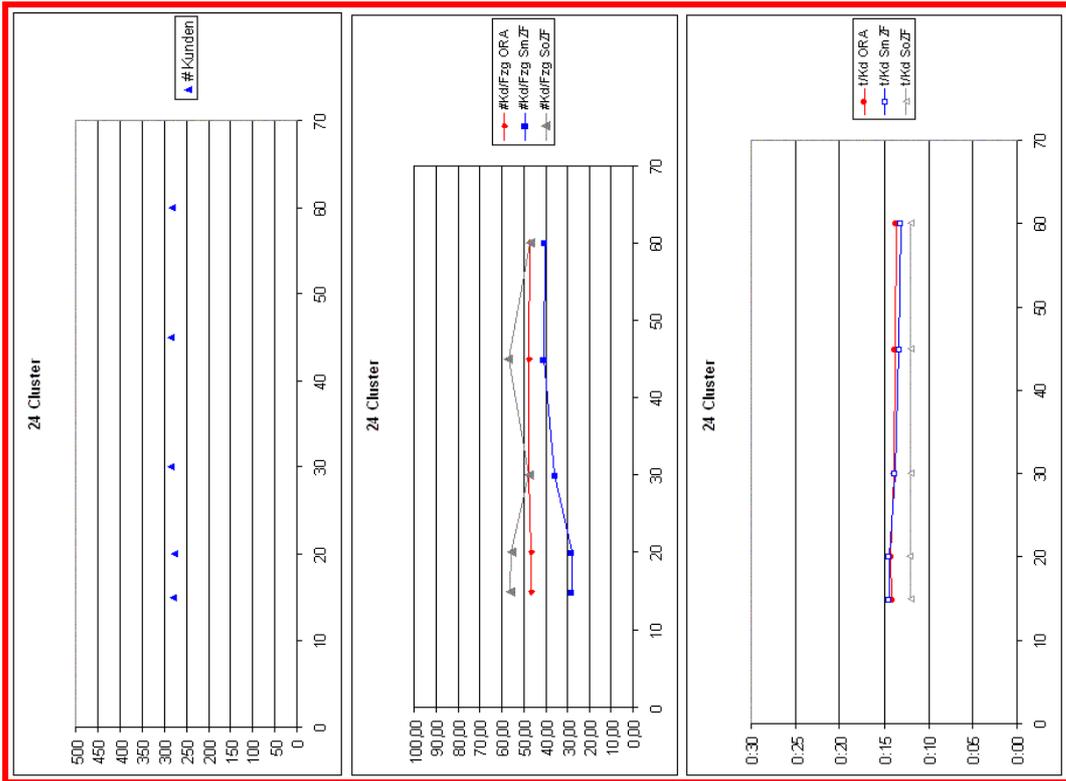
6.6.3.4 Angebotsstruktur

Im Vorlauf wurde bereits die Auswirkung der geografischen Angebotsstruktur¹⁴⁶ auf den Sättigungsverlauf (als Funktion der Toleranz) untersucht.

Hier wird nun die Auswirkung der geografischen Angebotsstruktur auf die Performance von ORA für eine definierte Toleranz (hier +/- 900 min.) untersucht.

¹⁴⁵ Weitergehende Detailuntersuchungen zeigen u.a., dass bei feiner (=kleiner) PRÄZ signifikante Wartezeitanteile entstehen.

¹⁴⁶ Die geografische Angebotsstruktur wird definiert durch die Aggregate (bzw. Cluster)



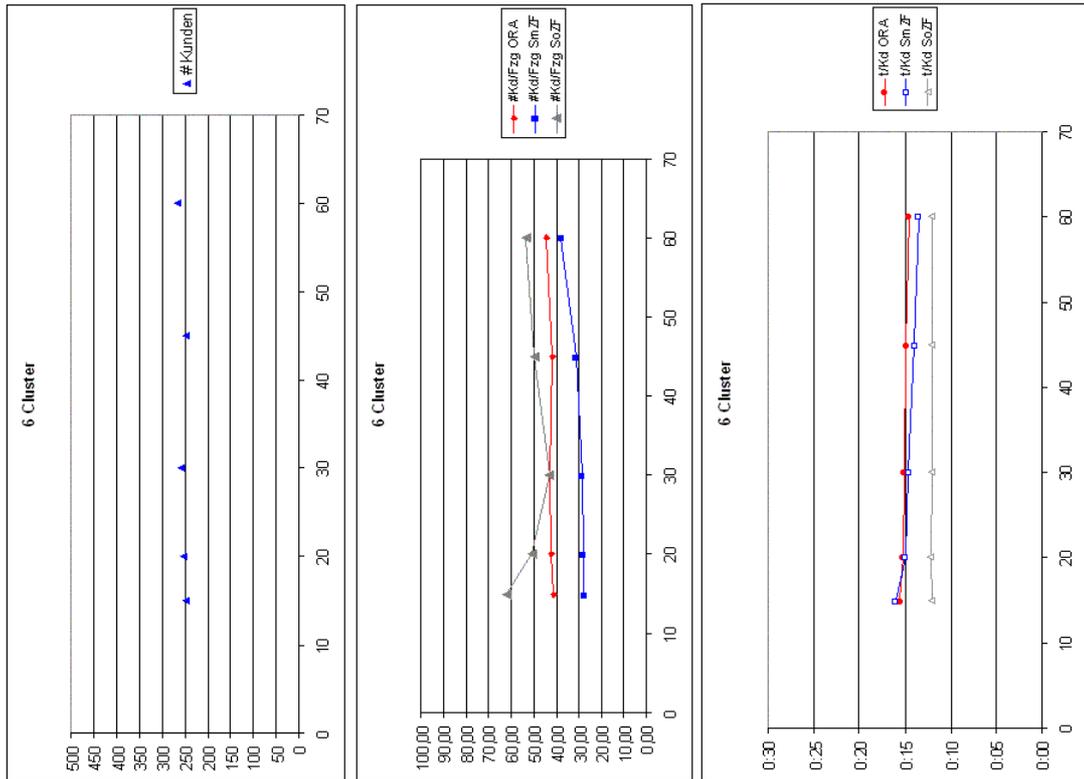


Abbildung 107: Einfluss der Angebotsstruktur (Anzahl Aggregate bzw. Cluster¹⁴⁷)
(x-Achse: PRÄZ; y-Achse s. Legende)

Aus den Messwerten der drei betrachteten Szenarios lässt sich folgendes schließen:

Je kleiner die Clusteranzahl¹⁴⁸, desto weniger Kunden werden von ORA akzeptiert, und desto schlechter ist daher auch die Performance bzgl. der KPI #Kunden/Fahrzeug.

Je kleiner die Clusteranzahl, desto mehr km werden generiert – dies ist einleuchtend, da mit steigender Clustergröße die möglichen Fahrwege innerhalb des Clusters tendenziell länger werden.

6.7 Zusammenfassung und Kritik

In diesem Kapitel wurden im wesentlichen die zwei Ansätze FPA/PPA („Filter mit Planung post Avis“ Ansatz) und ORA („Online Tourenplanung“ Ansatz) zur Lösung des D-m-VRSP-TW-UC-DO quantitativ näher untersucht. Nur kurz angerissen wurde der FPA/REA („Filter mit reduziertem Erfüllungsgrad“ Ansatz)¹⁴⁹.

¹⁴⁷ Aggregate und Cluster wird hier synonym verwendet.

¹⁴⁸ d.h. je größer die Clusterausdehnung.

¹⁴⁹ Der FPA/REA wurde in der Variante „Relaxation ex ante“ auf den relaxierten FPA/PPA zurückgeführt.

Für den FPA/PPA wurde gezeigt, dass die Anzahl benötigter Fahrzeuge allometrisch und die Tourkilometer exponentiell mit der Verengung des zugesagten Lieferzeitfensters (PRÄZ) zunehmen.

Für das neue Verfahren ORA wurde gezeigt, dass die Ergebnisse des Verfahrens weitgehend unabhängig von der zugesagten Zeitfensterbreite (PRÄZ) sind.

In Folgenden soll das gewählte methodische Vorgehen nochmals zusammengefasst und kritisch betrachtet werden.

6.7.1 Methodenkritik zur Analyse von FPA/PPA

Ein Problem der Aufgabenstellung „Untersuchung von Kosteneffekten“ ist die Vielzahl der möglichen Lösungsansätze. Daher wurde versucht, für ausgesuchte¹⁵⁰ bekannte Verfahren die **problemspezifischen** und die **verfahrensspezifischen Effekte voneinander zu trennen**, um so die problemspezifischen Charakteristika herauszuarbeiten.

Auf der Suche nach den Problemcharakteristika wurden zunächst **eigene Messungen** (Stichproben) angestellt, auf Grund dieser Stichproben **Hypothesen** formuliert, und diese Hypothesen dann anhand von veröffentlichten Ergebnissen **grob validiert**. Nachhaltig statistisch signifikant abgesicherte Aussagen können so zwar nicht getroffen werden¹⁵¹, aber die gezeigte Widerspruchsfreiheit zu veröffentlichten Ergebnissen ist doch ein wichtiges Indiz für die Validität der aufgestellten Hypothesen.

Für die Regressionsanalyse sind zwei Fragen entscheidend: mit **wie vielen Parametern** erfolgt die nichtlineare Regression, und an **wie vielen Stützpunkten** wird sie festgemacht. Das Dilemma dabei ist:

- ▶ 3-parametrische Funktionen liefern für die untersuchten Kurventypen bessere Fittings als 2-parametrische, da sie durch einen Parameter (konstantes Glied) die wagrechte Asymptote und durch 2 weitere Parameter die Krümmung modellieren können.
- ▶ Die Anzahl der Messpunkte aus den verwerteten Literaturergebnisse gibt mit je nur 3 bzw. 4 Messpunkten nicht genug Validierungsmasse her.

Da die Messpunkte aber nicht ohne weiteres ergänzt werden können, musste in dieser Arbeit ein Kompromiss gefunden werden zwischen „einigermaßen aussagekräftigem R^2 “ und „möglichst wenigen Parametern“.

¹⁵⁰ Natürlich ist schon allein die Verfahrensauswahl wichtig – hier wurde pragmatisch vorgegangen, und all die Verfahren miteinbezogen, die von Solomon selbst verwendet wurden oder die derzeit eine „beste Lösung“ für die Solomoninstanzen beigesteuert haben.

¹⁵¹ Die Vielzahl der dazu nötigen Untersuchungen an verschiedenen Instanzen und mit verschiedenen Verfahren hätte den Rahmen dieser Arbeit bei weitem gesprengt.

6.7.2 Methodenkritik zur Analyse von ORA

ORA ist ein neues Verfahren zur Lösung des D-m-VRSP-TW-UC-DO, das in dieser Arbeit erstmals im Detail beschrieben und in Hinblick auf sein Verhalten unter verschiedenen Nebenbedingungen (Standzeit, Geschwindigkeit, Anzahl Cluster) ausgewertet wurde¹⁵².

Ein direkter Vergleich von ORA mit anderen Verfahren zur Lösung des D-m-VRSP-TW-UC-DO ist aus zwei Gründen nicht möglich:

- ▶ Es gibt bisher **keine Referenzinstanzen** für das D-m-VRSP-TW-UC-DO
- ▶ Es gibt bisher **keine beschriebenen Verfahren**¹⁵³ zur Lösung des D-m-VRSP-TW-UC-DO

Daher wurde die Methode des „**ceteris paribus Teilaspektvergleichs**“ angewendet, mit der die Lösungsgüte von ORA in Bezug auf das klassische **VRSP-TW**¹⁵⁴ anhand einer **Obergrenze** (Heuristische Lösung des VRSP *ohne* Zeitfenster) und einer „**approximativen Untergrenze**“ (Heuristische Lösung des VRSP-TW) abgeschätzt wurde.

Damit in Zukunft eine **Referenzinstanz** für die Lösung des D-m-VRSP-TW-UC-DO bereit steht, um die hier vorgestellten Ergebnisse in Zukunft mit anderen Verfahren vergleichen zu können, wird die Testinstanzgruppe ZUE500* **veröffentlicht** (s. www.ifl.uni-karlsruhe.de).

¹⁵² Erste Ergebnisse zu ORA wurden vorab veröffentlicht in [Kunze 2003] und [Kunze 2005]

¹⁵³ In der neueren Literatur finden sich zunehmend Verweise auf Verfahren zur Lösung des D-VRSP, deren prinzipielle Einsetzbarkeit zur Lösung des D-m-VRSP-TW-UC-DO nicht ausgeschlossen werden kann. [Dorer & Calisti, 2005] stellen z.B. ein agentenbasiertes Verfahren zur Lösung eines D-m-PDPSTW Problems (dynamic multi vehicle pickup and delivery problem with soft time windows) vor. Es besteht Forschungsbedarf, ob dieses Verfahren auch zur Lösung des D-m-VRSP-TW-UC-DO eingesetzt werden kann.

¹⁵⁴ Das klassische VRSP-TW wird dadurch erzeugt, dass ORA aus den dynamisch eingehenden Lieferanfragen eine Teilmenge auswählt und dabei auch die Zeitfenster bestätigt. Die Gesamtheit der von ORA ausgewählten Lieferaufträge spannen das klassische VRSP-TW auf, welches als Basis für den Vergleich mit anderen Verfahren dient.

7 Zusammenfassung, Kritik & Forschungsbedarf

Diese Arbeit ist vom Kern her eine transportlogistische Forschungsarbeit, d.h. sie ist weder eine OR-Arbeit, noch eine Statistik-Arbeit, auch wenn sie sich beider Disziplinen bedient.

Die wesentlichen Untersuchungsaspekte sind in der folgenden Grafik nochmals zusammengefasst.

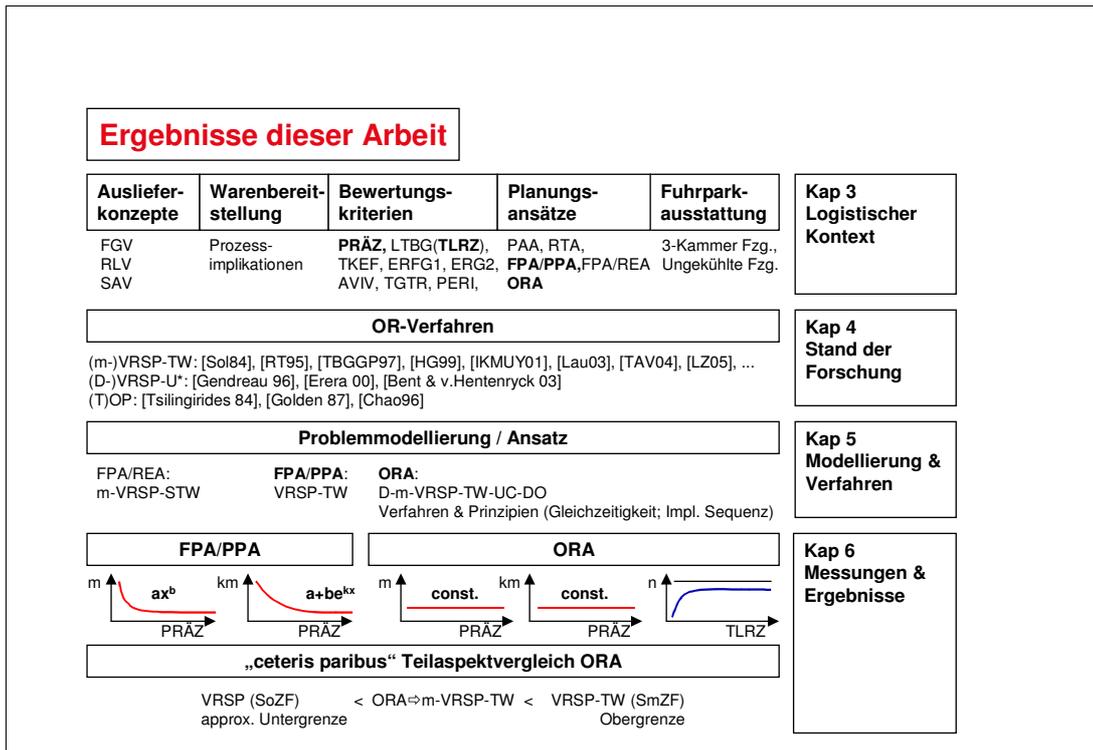


Abbildung 108: Ergebnisübersicht

7.1 Zusammenfassung der Arbeit

Gegenstand dieser Arbeit ist das praxisrelevante Problem der Tourenplanung für den eCommerce-Lebensmittelheimzustellservice.

Eingangs (**Kapitel 3**) wurden daher die logistischen und servicerelevanten Aspekte aus der Praxis beleuchtet. Aus diesen wurden zum einen Bewertungskriterien für die Planungsergebnisse abgeleitet (Lieferterminavispräzision, Lieferterminbeeinflussungsgrad (⇔Toleranz), Transportkosteneffizienz, Serviceerfüllungsgrad, Avisvorlauf, Tourgebietstreue und Peripherieeignung).

Daneben wurden die Prozesszusammenhänge zwischen Auftragseingang, Kommissionierung, Warenbereitstellung, und der Tourenplanung dargestellt. Diese

Zusammenhänge sind wichtig, da die Dynamik des Auftragseingangs zusammen mit dem Zeitbedarf für die Kommissionierung besondere Anforderungen an die Planungsergebnisse stellt (TGTR). Außerdem ist die Dynamik des Auftragseingangs entscheidend für die Wahl des Planungsansatzes und des Planungsverfahrens (Werden beispielsweise nichtdynamische Optimierungsverfahren zur Tourenplanung eingesetzt, so muss das dynamische Planungsproblem durch einen Auftragsannahme/-ablehnungs-Prozess (Filter) in ein nichtdynamisches überführt werden.)

Durch eine kurze Betrachtung der Ausstattungsanforderungen an die Auslieferungsfahrzeuge wurde gezeigt, dass der Auslieferungspark i.d.R. begrenzt ist - für die spätere Formulierung des OR-Problems ist das entscheidend (s. D-m-VRSP-TW-UC-DO).

Schließlich wurden die in der betrieblichen Praxis gängigen Planungsansätze für diese Arbeit klassifiziert und benannt (PPA, PAA, RTA, FPA, ORA, REA), und auf ihre Eignung für das zu lösende Problem hin untersucht. Aus Basis dieser Untersuchung und auf Basis der in der betrieblichen Praxis zu findenden Planungsansätze wurden die weiteren Analysen auf die beiden Ansätze FPA/PPA und ORA fokussiert (FPA/REA wurde in der Relaxation-ex-ante-Ausprägung auf den FPA/PPA zurückgeführt und ansonsten nur am Rand betrachtet, da sich dieser Ansatz nur als „Notfallansatz“ eignet, und nicht nach außen „verkaufen“ lässt).

Nach diesen vornehmlich praxisorientierten Betrachtungen wurden verschiedene Planungsverfahren aus der Literatur untersucht (**Kapitel 4**), die sich für die beiden Planungsansätze eignen. Während es für den FPA/PPA-Ansatz eine Fülle von Verfahren gibt (diese Verfahren wurde jeweils kurz zusammengefasst, um die Verfahrensprinzipien herauszuarbeiten), gibt es bisher keine beschriebenen Verfahren für den ORA-Ansatz der die Nachfrage aktiv steuert (Verfahren, die verwandte dynamische, aber andersartige Problemstellungen lösen wurden kurz skizziert.)

Im Anschluss (**Kapitel 5**) wurden die Planungsprobleme für den ORA und den FPA/PPA als OR-Probleme formalisiert. Dabei wurde eine neue OR-Problemklasse, nämlich die Klasse der D-m-VRSP-TW-UC-DO definiert. Diese Klasse wurde darüber hinaus in die Reihe der bekannten OR-Probleme eingereiht, und die Überschneidung mit der bisher wenig untersuchten Klasse der D-TOP-Probleme wurde herausgearbeitet. Aufbauend auf

- ▶ der neuen OR-Problem-Formalisierung,
- ▶ den neu eingeführten Prinzipien „Vermeidung von Gleichzeitigkeit“ und „Erzielung einer impliziten Tourstruktur“ und
- ▶ einigen theoretischen Überlegungen an akademischen Beispielen

wurde das ORA-Verfahren hergeleitet, erläutert und einer kritischen ex ante

Wertung unterworfen. Daneben wird das Prinzip des FPA/PPA mit dem notwendigen zugehörigen Filter-Prozess erläutert.

Im **Kapitel 6** werden schließlich quantitative Untersuchungen für beide Ansätze angestellt. Die wesentliche Frage dabei ist, wie die Abhängigkeit der Transportkosten (gemessen über die Indikatoren Anzahl Fahrzeuge und Gesamttourkilometer) von der im Lieferservice zugesagten Präzision (PRÄZ) aussieht.

Zuerst wird das Verhalten von FPA/PPA auf diese Frage hin untersucht. Es wird gezeigt, dass unabhängig vom eingesetzten Verfahren die Transportkostenindikatoren „Anzahl Fahrzeuge“ und „Gesamttourkilometer“ mit enger werdender Präzision steigen, solange keine erkennbare implizite Sequenzialität in den Aufträgen vorliegt. Um dies zu zeigen wurden zwei wesentliche Methoden eingesetzt:

- ▶ Um den ansatzimmanenten Effekt von den überlagernden verfahrensspezifischen Effekten zu trennen, wurde die Methode der linearen Regression eingesetzt.
- ▶ Um die aus eigenen Messungen resultierenden Ergebnisse zu validieren, wurden in der Literatur veröffentlichte Ergebnisse in Bezug auf Referenzinstanzen des VRSP-TW herangezogen (Solomoninstanzen).

Anschließend wird für ORA gezeigt, dass die Transportkostenindikatoren „Anzahl Fahrzeuge“ und „Gesamttourkilometer“ mit enger werdender Präzision konstant bleiben, da ORA die implizite Tourstruktur herstellt (und Gleichzeitigkeit vermeidet).

Zuvor wurde noch eine wichtige Voraussetzung von ORA, die Toleranz der Kunden (TLRZ) bzgl. eines möglichen Liefertermins untersucht. Dieser Aspekt ist deshalb wichtig, da ORA sich gerade dadurch auszeichnet, dass es die Toleranz der Kunden als zusätzlichen Freiheitsgrad der Optimierung nutzt.

Abschließend wird die Transportkosteneffizienz von ORA der von FPA/PPA mit Hilfe von „ceteris paribus Teilaspektvergleichen“ gegenübergestellt. Dazu wird postuliert, dass die von ORA akzeptierten Aufträge dem Filterergebnis von FPA/PPA entsprechen. Zum anderen wird eine Ober- und eine approximative Untergrenze für die Bewertung von ORA gegenüber dem FPA/PPA definiert. Es wird gezeigt, dass ORA bei engen Zeitfenstern insbesondere im Hinblick auf die Anzahl der benötigten Fahrzeuge gute Ergebnisse erzielt.

Da es bisher für das OR-Problem des D-m-VRSP-TW-UC-DO keine Referenzinstanzen gibt, werden dazu drei neue Test-Instanzen vorgeschlagen (ZUE500e, ZUE500f und ZUE500g) anhand derer in Zukunft direkte Vergleiche mit ORA möglich werden.

7.2 Kritik der Arbeit

Nach [Popper 1994, S. 21] gibt es ein dreistufiges wissenschaftliches Schema:

1. Problem
2. Lösungsversuch
 - 2.1. Beobachtung / Sinneswahrnehmung
 - 2.2. Lösungsversuch
 - 2.3. Theorie / Hypothese
3. Elimination

Für die Betrachtung des FPA/PPA wurde ein Problem eingeführt (= Stufe 1), und verschiedene Messungen (= Stufe 2.1 Beobachtungen / Datenauswertung aus Literatur) dazu vorgenommen. Bei der Suche nach problem- und verfahrensspezifischen Eigenschaften des FPA/PPA wurde in dieser Arbeit versucht aus Messreihen mit statistischen Methoden (nichtlineare Regressionsversuche) Hypothesen abzuleiten (= Stufe 2.2. / 2.3.). Diese Hypothesen wurden teilweise wieder eliminiert (= Stufe 3 verworfene Regressionsversuche¹⁵⁵).

Da die Ergebnisse mit statistischen Methoden aus den verfügbaren Daten mittels nichtlinearer Regression gewonnen wurden, würde das Testen der so gewonnenen Hypothesen an ebendiesen Daten keinen Mehrwert bringen. Das Testen von Hypothesen zum FPA/PPA macht dann Sinn, wenn sich die in dieser Arbeit aufgestellten Hypothesen an anderen, neuen Daten¹⁵⁶ testen lassen.

Für die Betrachtung des ORA ist in dieser Arbeit der Weg bis zur Stufe 2.3 gegangen worden. Als Basis für die Stufe 3 werden die Ergebnisse zu ORA hier zusammen mit den Test-Instanzen ZUE500* veröffentlicht, damit die weitere Forschung ggf. Schwachpunkte aufdecken und eliminieren kann. Die Stufe 3 wurde in dieser Arbeit für ORA somit noch nicht erklommen.

Voraussetzung für das künftige Testen von Hypothesen bzgl. des ORA ist das Vorliegen von nachhaltigen Untersuchungen für das D-m-VRSP-TW-UC-DO in Bezug auf standardisierte Testinstanzen und in Bezug auf verschiedene andere Verfahren. Erst dann besteht die Chance, Verfahrenseffekte klar von Problemeffekten zu trennen.

¹⁵⁵ z.B. die Annahme, die Abhängigkeit der Gesamtkilometer von der Präzision ließe sich auch als allometrische Funktion darstellen.

¹⁵⁶ Um Hypothesen statistisch sauber testen zu können, braucht es eine entsprechend große Anzahl von Testinstanzen. Schon der Versuch der Extraktion der Problemspezifika beim PPA über verschiedene Verfahren hinweg hat durch die Literaturlauswertung auf einige Mannjahre Forschung zum VRSP-TW zurückgegriffen, und ist in seiner Aussagegültigkeit absolut gesehen immer noch verbesserungswürdig (Basis sind allein die wenigen Solomoninstanzen, mit einer einzigen Durchschnittsgeschwindigkeit und mit einer uneinheitlichen Standzeit je nach Szenario)

7.3 Weiterer Forschungsbedarf

Zur Beantwortung der eingangs gestellten einfachen Frage „Was kosten präzise vorausgesagte Zeitfenster?“ bzw. „Wie hängen die Transportkosten von enger werdenden Zeitfenstern ab?“ hat diese Arbeit einen Beitrag geleistet. Trotzdem bleiben viele Fragen offen. Es besteht daher noch folgender signifikanter Forschungsbedarf über verschiedene Disziplinen hinweg.

7.3.1 Logistik

Die Aufgabe der „Logistischen Forschung“ ist es, die Optimierungsprobleme und insbesondere deren Freiheitsgrade¹⁵⁷ aus logistischer Sicht klarzu formulieren und zu modellieren¹⁵⁸, und schließlich OR-Verfahrensergebnisse in die betriebliche Praxis umzusetzen¹⁵⁹. Die präzise Beschreibung der möglichen **Planungsansätze** spielt dabei eine entscheidende Rolle. Es sollte daher angestrebt werden, die Planungsansätze ebenso gut zu **klassifizieren**, wie dies im Bereich der OR-Problem-Klassifikation bereits erfolgt ist. Ein aller erster (sicher verbesserungswürdiger) Ansatz wurde hier mit der RTA, ORA, FPA, PPA, ... – Notation gemacht.

Die in dieser Arbeit aufgestellten Hypothesen zum **Zusammenhang** der **Lieferavispräzision** auf der einen Seite, und den **Gesamttourkilometern** und der benötigten **Fahrzeuganzahl** auf der anderen Seite, sind durch weitergehende Untersuchungen zu festigen oder zu widerlegen. Bei diesen Untersuchungen muss jeweils zwischen Problem- und Verfahrensspezifika unterschieden werden. Außerdem ist das Wichtungsproblem (Fahrzeuge vs. km) dabei jeweils zu betrachten. Diese Arbeit ist hierfür erst ein Ausgangspunkt.

Eine weitere offene Frage aus dieser Arbeit ist, wie die gefundenen **Regressionsparameter**, die den Zusammenhang zwischen Lieferzeitfensterbreite und Gesamttourkilometern bzw. Fahrzeuganzahl beschreiben, aus logistischer Sicht **interpretiert** werden können, falls sie verfahrensübergreifend, also problemspezifisch ermittelt werden können¹⁶⁰.

7.3.2 BWL

Die eigentliche Optimierungsfrage aus dieser Arbeit ist nicht nur eine

¹⁵⁷ wie zum Beispiel die „Toleranz“ der Kunden als nutzbaren Freiheitsgrad erkennen und formal beschreiben.

¹⁵⁸ Definition des Anwendungsfalls und des zugehörigen OR-Problems (ZF und NB)

¹⁵⁹ z.B. „Wie lässt sich ein optimiertes Tourenplanungsergebnis umsetzen, das erst zu einem Zeitpunkt erzielbar ist, an dem die Kommissionierung bereits abgeschlossen sein müsste?“

¹⁶⁰ Die Abweichung der verfahrensspezifischen Regressionsparameter von den problemspezifischen Regressionsparametern könnten dann als scharfes quantitatives Maß für die Verfahrensgüte in Bezug auf das Planungsproblem genutzt werden.

transportlogistische, sondern auch eine betriebswirtschaftliche. Wesentlicher betriebswirtschaftlicher Forschungsbedarf besteht daher zum einen bei der Konkretisierung und **Quantifizierung** der **Opportunitätskosten für die Verletzung zugesagter Zeitfenster** (ERG2-Verletzungen). Ob ein quantitatives Maß für diese Opportunitätskosten so definiert werden kann, dass es in die Zielfunktion der täglichen Tourenoptimierung eingehen kann, und wie dieses Maß sinnvoll bestimmt werden kann ist zu untersuchen.

Zum anderen sollten statistisch fundierte **Marktanalysen zur Toleranz** der Kunden in Bezug auf Lieferterminvorschläge durchgeführt werden, damit die in dieser Arbeit theoretisch postulierte und im Praxiseinsatz genutzte Größe Toleranz (TLRZ) als wichtiger Freiheitsgrad der Optimierung „berechenbarer“ wird.

Sowohl Marktanalysen zum **Nutzen** einer hohen **Lieferterminpräzision** (PRÄZ) für den Kunden (verlässlicher Liefertermin) als auch Untersuchungen zum Nutzen einer hohen PRÄZ für den Lieferanten (Minimierung der Kosten¹⁶¹ für „Nichtantreffen“ des Kunden) können helfen, das betriebswirtschaftliche Optimum für die Präzision der Lieferterminzusage zu bestimmen.

Schließlich kann die betriebswirtschaftliche Forschung helfen, die Parametrisierung für das oben diskutierte **Wichtungsproblem** (variable km-Kosten vs. Fahrzeug-Fixkosten) zu liefern, die in der betrieblichen Praxis benötigt wird, um die Optimierungsverfahren richtig zu steuern.

7.3.3 OR

Unabhängig davon, welcher Tourenplanungsansatz für den eCommerce-Lebensmittelheimzustellservice gewählt wird, besteht in der betrieblichen Praxis ein Bedarf an **hochperformanten Verfahren**, die auch bei „**real life**“-**Problemgrößen** (z.B. 1000 Aufträge, 100 Fahrzeuge) in Sekunden möglichst gute Ergebnisse liefern.

Da Optimalität für „real life“-Probleme nur schwer nachweisbar ist, werden außerdem verlässliche „**Brauchbarkeitsindikatoren**“ gebraucht, die für konkrete Anwendungsfälle die **Verfahrenseignung** unter Berücksichtigung der jeweiligen logistischen Nebenbedingungen **quantifizieren**. Referenzprobleme (**Instanzen**) für konkrete Anwendungsfallklassen können helfen, diese „**Brauchbarkeitsindikatoren**“ zu generieren.

Die **Freiheitsgrade** bei der Planung mit engen Zeitfenstern sind je nach Anwendungsfall unterschiedlich. [Grupp, 1987] beschreibt schon 1987 den Freiheitsgrad der „**Ablehnung von Aufträgen**“ (Ablehnungsszenario) als Alternative zur Serviceaufweichung (Verspätungsszenario). Die neuere OR-Forschung hat das

¹⁶¹ Kosten für die physische Zweitzustellung, Administrative Kosten, ggf. Umsatzverlust, etc.

Thema „Verspätung“ in Form der VRSP-**STW** aufgenommen. Der Freiheitsgrad „Ablehnung“ ist ein wesentlicher Bestandteil des **TOP**. Die dynamischen Aspekte von Planungsproblemen werden in der Forschung zu den **D-m-VRSP** untersucht. Daher sind für die Forschung in Bezug auf die Tourenplanung für den eCommerce-Lebensmittelheimzustellservice künftige Ergebnisse zu allen drei Problemklassen relevant.

Konkreter Forschungsbedarf besteht in Hinblick auf **neue Verfahren** zur Lösung des in dieser Arbeit neu formulierten D-m-VRSP-TW-UC-DO. Forschungsbedarf, der sich aus dieser Arbeit darüber hinaus ableitet, ist die **Validierung** (oder Elimination) der 4 **Hypothesen** an weiteren Instanzen und mit weiteren Verfahren.

7.3.4 Angewandte Statistik

Die angewandte Statistik kann wichtige methodische Beiträge zur Erkenntnisfindung im Kontext dieser Arbeit leisten. Forschungsbedarf besteht zum einen hinsichtlich einer verbesserten Methodik zur **multivariaten Interdependenzanalyse** der verschiedenen Einflussfaktoren – der in dieser Arbeit eingesetzte „ceteris paribus“-Teilaspektvergleich ist erst ein aller erster Schritt.

Wenn die Interdependenzen der Einflussgrößen über die bisherigen „ceteris paribus“-Untersuchungen hinaus klarer sind, kann die angewandte Statistik außerdem einen wichtigen Beitrag bei der Erzeugung **aussagekräftiger Referenzinstanzen** leisten (Design of Experiments (DOE) unter Berücksichtigung der Einflussfaktorinterdependenzen).

Schließlich besteht Forschungsbedarf in Bezug auf die schärfere **Trennung von Verfahrens- und Problemcharakteristika** immer dann, wenn nur heuristische Lösungen (und keine optimalen Lösungen) für Probleminstanzen bekannt sind.

7.3.5 Interdisziplinär

Der im Vorangegangenen für die einzelnen Disziplinen aufgezeigte Forschungsbedarf (in Bezug auf die Tourenplanung im eCommerce-Lebensmittelheimlieferservice) entfaltet seinen Nutzen für die betriebliche Praxis dann, wenn er interdisziplinär synchronisiert wird.

7.4 Ausblick

eCommerce-Heimlieferservices sind hochwertige Dienstleistungen, die dem Kunden das physische „Einkaufen gehen“ abnehmen. Die i.d.R. vom Kunden selbst erbrachten Kommissionierleistungen (sich das aus den Regalen nehmen, was man

kaufen möchte) und „last mile“-Transportleistungen (die gekaufte Ware nach Hause transportieren) werden hier vom Anbieter und den ggf. beauftragten Transportdienstleistern geleistet.

Je nach Produkt ergeben sich dabei unterschiedliche Anforderungen an die physische Zustellung. Während beispielsweise Bücher und CDs über die Netze der KEP-Dienstleister kostengünstig ausgeliefert werden können, und keine physische Begegnung mit dem Warenempfänger notwendig ist (Briefkastenzustellung), und daher keine präzise Lieferterminvereinbarung notwendig ist, müssen Lebensmittel, Getränke, Möbel und Weißware bei der Heimzustellung übergeben werden, was eine Lieferterminvereinbarung erfordert. (Eine Alternative zur dieser persönlichen Warenempfangnahme ist die Zustellung an Paket- oder Packstationen.)

Ob und wie ein solcher Service erfolgreich angeboten, vom Kunden akzeptiert und mikroökonomisch gewinnbringend umgesetzt wird, regeln marktwirtschaftliche Prozesse. (Einige Anbieter sind am eCommerce-Eisberg zerschellt, andere haben es verstanden, den eCommerce-Eisberg rechtzeitig zu orten und logistisch und betriebswirtschaftlich entsprechend richtig zu navigieren.)

Aus makroökonomischer und verkehrswissenschaftlicher Sicht verlagert der eCommerce-Heimlieferservice Teile des Individualverkehrs (nämlich die Einkaufsfahrten) in den gewerblichen Güterverkehr. Ob das Gesamtverkehrsaufkommen dadurch steigt, sinkt oder unverändert bleibt hängt stark von den jeweiligen Konsolidierungspotenzialen ab (aus Kundensicht die Konsolidierung beim Einkauf mehrerer Güter, aus Anbietersicht die Konsolidierung mehrerer Kunden auf einer Liefertour).

Gerade weil die makroökonomischen und verkehrswissenschaftlichen Aspekte des eCommerce-Heimlieferservice nicht Gegenstand dieser Arbeit sind, die Durchführung des eCommerce-Heimlieferservice jedoch einen Beitrag zum Güterverkehrsaufkommen leistet, soll diese Arbeit zumindest mit diesem Ausblick geschlossen werden.

8 Quellen

[Arnold & Furmans, 2005] Arnold, Dieter; Furmans, Kai (2005). *Materialfluss in Logistiksystemen*, Springer Verlag, Berlin Heidelberg

[Archetti et al 2005] Archetti, C., Hertz, A. Speranza, M.G. (2005) *Metaheuristics for the team orienteering problem*, paper posted for Optimization Days 2005, Montreal, May 9-11, Groupe d'Etudes et de Recherche en Analyse des Decisions (GERAD)

[Baumung, 2003] Baumung, Martin (2003) *Tourenplanung mit alternativen Lieferorten - Entwicklung und Implementierung verschiedener Lösungsverfahren* Diplomarbeit, Institut für Wirtschaftstheorie und Operations Research, Universität Karlsruhe (TH)

[Bent & van Hentenryck 2003] Bent, Russel W., Van Hentenryck, Pascal (2003) *Scenario Based Planning for Partially Dynamic Vehicle routing with Stochastic Customers*, www publication <http://www.cs.brown.edu/people/rbent/ORsr.pdf> of Department of Computer Science, Brown University; (Publication in Operations Research to appear)

[Bent & van Hentenryck 2004] Bent, Russel W., Van Hentenryck, Pascal (2004) *A Two Stage Hybrid Local Search for the Vehicle Routing Problem with Time Windows* Transportation Science, Vol. 38, No. 4, November 2004, S. 515-530

[Blum 2004] Blum, Christian (2004) *Theoretical and practical aspects of ant colony optimization*, Dissertation, IRIDIA Institut de Recherche Interdisciplinaires et de Developpements en Intelligence Artificielle, Faculté de Sciences Appliquées; Université Libre de Bruxelles

[Bodin et al., 1983] Bodin, L.D., Golden, B.L., Assad, A.A.; Ball, M.O. (1983) *Routing and scheduling of vehicles and crews – The state of the art*, Computers and Operations Research 10

[Bronstein & Semendjajew 1991] Bronstein I.N. & Semendjajew K.A. (1991) *Taschenbuch der Mathematik*, B.G. Teubner Verlagsgesellschaft, Stuttgart, Leipzig

[Bräysy & Gendreau 2001] Bräysy, O. und Gendreau, M. (2001). *Metaheuristics for the vehicle routing problem with time windows*. Technical Report STF42 A01025, SINTEF Applied Mathematics, Department of Optimisation, Oslo, Norway

[BVH04] s. [Bent & van Hentenryck 2004]

[Chao, Golden et al, 1996a] Chao, I-Ming, Golden, Bruce L., Wasil, Edward A. (1996) *A fast and effective heuristic for the orienteering problem*, European Journal of Operations Research 88, S. 475-489

[Chao, Golden et al, 1996b] Chao, I-Ming, Golden, Bruce L., Wasil, Edward A. (1996) *The team orienteering problem*, European Journal of Operations Research 88, S. 464-474

[Daduna 2002] Daduna, Joachim R. (2002), *Distribution planning with specific delivery time restrictions for the handling of electronic customer orders in food/non-food retail trade*, Quantitative approaches to distribution logistics and supply chain management (Editors: Klose,A., Speranza, M.G., van Wassenhove L.N.), Springer-Verlag Berlin, Heidelberg, New York, London, Paris, Tokyo, Hong Kong

[Daganzo & Erera, 1999] Daganzo, C.F., Erera, A. (1999) *On planning and design of logistic systems for uncertain environments*, New Trends in Distribution Logistics (Editors: M.G. Speranza, P.Stahly) – Lecture Notes in Economics and Mathematical Systems, V.480, Springer-Verlag Berlin, Heidelberg, New York, London, Paris, Tokyo, Hong Kong

[Desrochers et al., 1990] Desrochers, M., Lenstra, J.K, Savelsbergh M.W.P. (1990) *A classification scheme for vehicle routing and scheduling problems* , European Journal of OR 46

[Domschke, 1995] Domschke, Wolfgang, *Logistik-Band 1 Transport*, R.Oldenbourg Verlag, München

[Domschke, 1997] Domschke, Wolfgang, *Logistik-Band 2 Rundreisen und Touren*, R.Oldenbourg Verlag, München

[Dorer & Calisti, 2005] Dorer, Klaus und Calisti, Monique (2005) *An adaptive Approach to Dynamic Transport Optimization*, Applications of Agent Technology in Traffic and Transportation (Editors: F. Klügl, A. Bazzan, S. Ossowski), Birkhäuser Verlag, Basel, Boston, Berlin

[Erera, 2000] Erera A.L. (2000) *Design of vehicle routing systems for uncertain environments (title approximate)*, PhD thesis, Department of Industrial Engineering and Operations Research, University of California, Berkeley, C.A.

[Erera & Daganzo, 2000] Erera Alan L.; Daganzo Carlos F. (2000) *Coordinated vehicle routing with uncertain demand*; presentation at ROUTE 2000 – International workshop on vehicle routing, Skodsborg (denmark), August 16-19; 2000

[Fomin et al., 2002] Fomin, Fedor V., Lingas, Andrzej (2002). *Approximation algorithms for time dependent orienteering* Information Processing letters 83 (2002), S. 57-62

[Gatzke, 2001].Gatzke, Eckard (2001). Der KEP-Markt wandelt sich zu KEAV: Kurier, Express und Added Value, in *Logistik im E-Zeitalter* (Herausgeber: H.Baumgarten), Frankfurter Allgemeine Zeitung Verlagsbereich Buch, Frankfurt (Main)

[Geem et al 2005] Geem, Z.W., Tseng, C.L., Park, Y (2005). *Harmony search for Generalized Orienteering Problem: Best Touring in China*, Springer Lecture Notes in Computer Science, 3412 August 2005, S741-750.

[Gendreau et al; 1992] Gendreau, M., Hertz, A., und Laporte, G. (1992). *New insertion and postoptimization procedures for the traveling salesman problem*. Operations Research, 40:1086–1094.

[Gendreau et al; 1995] Gendreau, Michel; Laporte, Gilbert; Séguin, René (1995), *An exact algorithm for the vehicle routing problem with stochastic demands and customers*, Transportation science 29-2, S.143-155

[Gendreau et al; 1996] Gendreau, Michel; Laporte, Gilbert; Séguin, René (1996), *A Tabu search heuristic for the vehicle routing problem with stochastic demands and customers*; Operations Research 44-3, S.469-477

[Gendreau et al., 2001] Gendreau, Michel; Pelleu, Josephine; Séguin, René; Soriano, Patrick; (2001) *Improving Robustness of a Tabu Search Algorithm for Stochastic Vehicle Routing Problems via Diversification*, 4th Metaheuristics International Conference MIC'2001, Porto Portugal, July 16-20, 2001

[Ghiani et al. 2004] Ghiani, Gianpaolo; Guerriero, Francesca; Laporte, Gilbert; Musmanno, Roberto (2004) *Real-Time Vehicle Routing: Solution concepts, Algorithms and Parallel Computing Strategies*, European Journal of Operations Research 151 (1), S.1-11

[Golden et al., 1987] Golden, Bruce L.; Levy, Larry; Vohra, Rakesh (1987) *The orienteering problem*, Naval Research Logistics, Vol. 34, S. 307-318

[Grupp, 1987] Grupp, Bernhard (1987) *Logistische Tourenplanung unter Zeitrestriktionen*, Dissertation - Wirtschaftswissenschaftliche Fakultät der Universität Regensburg

[HG99] s. [Homberger & Gehring, 1999]

[Homberger & Gehring, 1999] Homberger J., Gehring H. (1999), *Two Evolutionary Meta Heuristics for the VRP with Time Windows*, INFOR, Vol 37, 297-318

[Ibaraki et al 2001] Ibaraki T., Kubo M., Masuda T., Uno T., Yagiura M. (2001) *Effective Local Search Algorithms for the VRP with General Time Windows*, Working Paper, Department of Applied Mathematics and Physics, Kyoto University, Japan

[IKMUY01] s. [Ibaraki et al 2001]

[Kunze, 2003] Kunze, Oliver (2003), *A new interactive approach on route planning with tight delivery time windows*, Logistics Systems for Sustainable Cities;

Proceedings of the 3rd International Conference on City Logistics, Madeira, Portugal, 25-27 June, 2003, Eiichi Taniguchi & Russel G. Thompson (editors); Elsevier Ltd. Oxford

[Kunze, 2005] Kunze, Oliver (2005), *Ein praxistauglicher Ansatz zur Lösung eines spezifischen D-VRSP-TW-UC*, Operations Research Proceedings 2004, H. Fleuren, D. den Hertog, P. Kort (editors), Springer Verlag, Berlin Heidelberg New York

[Lau et al 2003] Lau H.C., Si M., Kwong M.T. (2003), *Vehicle Routing problem with Time Windows and a Limited Number of Vehicles*, European Journal of Operational Research 148 (2003) 559-569

[Lau03] s. [Lau et al 2003]

[Levenberg 1944] Levenberg, K. (1944), *A Method for the Solution of Certain Non-linear Problems in Least Squares*, Quarterly of Applied Mathematics, 2(2) 1944, 164-168

[Lim & Zhang 2005] Lim A. & Zhang X. (2005), *A Two Stage Hybrid Local Search for the Vehicle Routing Problem with Time Windows*, Proceedings of the 38th Hawaii International Conference on System Sciences

[Lourakis 2005] Lourakis, Manolis I. A. (2005) *A Brief Description of the Levenberg-Marquardt Algorithm Implemented by `levmar`*, Institute of Computer Science Foundation of Research and Technology – Hellas (FORTH), www publication www.ics.forth.gr/~lourakis/levmar

[LZ05] s. [Lim & Zhang 2005]

[Marquardt 1963] Marquardt, D.W. (1963) *An Algorithm for Least-Squares Estimation of Nonlinear Parameters*, SIAM Journal of Applied Mathematics, 11(2) 1963, 431-441

[Maziejewski, 1992] Maziejewski, Stefan (1992), *The Vehicle Routing and Scheduling Problem with Time Window Constraints using genetic Algorithms*, Diploma Thesis, Institut für Logik, Komplexität und Deduktionssysteme, Universität Karlsruhe (TH)

[Mole, Jameson 1976] Mole, R. und Jameson, S. (1976). *A sequential route-building algorithm employing a generalized savings criterion*. Operational Research Quarterly, 27

[Neumann, Morlock 1993] Neumann, Klaus und Morlock, Martin (1993), *Operations Research*, Carl Hanser Verlag, München, Wien

[Osman, 1993] Osman, I. (1993). *Metastrategy simulated annealing and tabu search algorithms for the vehicle routing problem*. Annals of Operations Research,

41:421–451.

[Paessens, 1987], Paessens, H. (1987) *Tourenplanung mit TourMaster*, Oldenbourg Verlag, München, Wien

[Pereira, Tavares et al 2002] Pereira, F.B., Tavares, J., Machado P., Costa, E. *GVR: a New Genetic Representation for the Vehicle Routing Problem*, in Proceedings of the 13th Irish Conference on Artificial Intelligence and Cognitive Science (ACIS 2002), pp 95-102, Limerick, Ireland, 12-13 September, 2002

[Popper 1994] Popper, Karl R. (1994) *Alles im Leben ist Problemlösen*, Piper Verlag GmbH, München

[Potvin & Rousseau 1995] Potvin, J.Y. & Rousseau, J.M. (1995) *An exchange heuristic for routing problems with time windows* Journal of the Operations Research Society, 46, 1433-1446

[Potvin & Bengio 1996] Potvin, J. und Bengio, S. (1996). *The vehicle routing problem with time windows, part II: Genetic search*. INFORMS Journal on Computing, 8(2):165–172.

[Ranganathan 2004] Ranganathan, Ananth (2004) „The Levenberg-Marquardt Algorithm“, *A tutorial by Ananth Ranganathan*, www publication <http://www.cc.gatech.edu/~ananth/lmtut.pdf>

[Rasch et al 1999] Rasch, D., Verdooren, L.R., Gowers J.I. (1999) *Grundlagen der Planung und Auswertung von Versuchen und Erhebungen*, R.Oldenbourg Verlag München, Wien

[Rochat & Taillard, 1995] Rochat, Yves & Taillard, Eric D. (1995) *Probabilistic Diversification and Intensification in Local Search for Vehicle Routing*, Journal of Heuristics 1, 1995, S. 147-167

[Roweis 2005] Roweis, Sam, *Levenberg Marquardt Optimization* (wahrscheinlich 2005) www publication www.neuron.eng.wayne.edu/ECE512/lm.ps

[RT95] s. [Rochat & Taillard, 1995]

[S98] s. [Shaw, 1998]

[Schmidt,1994] Schmidt, Joachim (1994) *Die Fahrzeugeinsatzplanung im gewerblichen Güterfernverkehr – Ein graphentheoretischer Ansatz zur Planung von Fahrzeugumläufen*, Peter Lang Europäischer Verlag der Wissenschaften, Frankfurt (Main), Berlin, Bern, New York, Paris, Wien

[Shaw, 1998], Shaw, Paul (1998) *Using Constraint Programming and Local Search Methods to Solve VRPs*, Principles and Practice of Constraint Programming –

CP98, Lecture Notes in Computer Science, 417-431, M.Mahler & J.F. Puget (eds), Springer-Verlag, New York

[Slotta, 2001] Slotta, Gerd (2001) *Vom Speditionsunternehmen zum E-Logistiker*, in Logistik im E-Zeitalter (Herausgeber: H. Baumgarten), Frankfurter Allgemeine Zeitung Verlagsbereich Buch, Frankfurt (Main)

[Sol84] s. [Solomon, 1984]

[Sol03] s. [Solomon, 2003]

[Solomon, 1984] Solomon, Marius M. (1984) *Vehicle Routing and Scheduling with Time Constraints: Models and Algorithms*, Dissertation, Department of Decision Science; The Wharton School, University of Pennsylvania, Philadelphia

[Solomon, 1986] Solomon, Marius M. (1986) *On the worst case performance of some heuristics for the vehicle routing and scheduling problem with time window constraints*, Networks, Vol. 16, S.161-174

[Solomon, 2003] Solomon, Marius M. (2003) *VRPTW Benchmark Problems (updated June 10, 2003)* <http://w.cba.neu.edu/~msolomon> Northeastern University, Boston, Massachusetts

[Sol03] = [Solomon, 2003]

[Taillard et al, 1997], Taillard E., Badeau P., Gendreau M., Geurtin F., Potvin J.Y. (1997) *A Tabu Search Heuristic for the VRP with Time Windows*, Transportation Science, 31, 170-186

[TAV04] s. [Tavares 2004]

[Tavares 2004] Tavares, Jorge (2004) *Vehicle Routing Problem – The Evolutionary Way*, <http://eden.dei.uc.pt/~jast/vrp/solomon100.html> Centre for Informatics and Systems, University of Coimbra

[Tavares, Pereira et al 2002] Tavares, J., Pereira, F.B., Machado P., Costa, E. *Crossover and Diversity: A Study about GVR*, in Proceedings of the Analysis and Design of Representations and Operations (ADoRo 2003) a bird-of-a-feather workshop at the 2003 Genetic and Evolutionary Computation Conference (GECCO 2003), Chicago, Illinois, USA 12-16 July 2003

[TBGGP97] s. [Taillard et al, 1997]

[Tsiligirides, 1984] Tsiligirides, T. (1984). *Heuristic Methods Applied to Orienteering*, Journal of the Operations Research Society 35(9), S. 797-809

[Wendt, 1995] Wendt, Oliver (1995). *Tourenplanung durch Einsatz naturanaloger*

Verfahren, Deutscher Universitäts Verlag / Gabler Verlag, Wiesbaden,

[Witten, 2001] Witten, Peer (2001). Internationale Logistik als Tool für strategischen Markterfolg, in *Logistik im E-Zeitalter (Herausgeber: H. Baumgarten)*, Frankfurter Allgemeine Zeitung Verlagsbereich Buch, Frankfurt (Main)

[Wildemann, 1997] Wildemann, Horst (1997). *Logistik Prozessmanagement*, TCW Transfer-Centrum-Verlag GmbH, München

Anhang

A	Anhang A – Verfahren für das VRSP.....	3
A.1	Klassische Verfahrensansätze für VRSPs.....	3
A.1.1	Konstruktive Verfahren.....	3
A.1.2	Verbesserungsverfahren.....	7
A.2	Metaverfahren für VRSPs	10
A.2.1	Evolution – Genetische Algorithmen	11
A.2.2	Zoologie – Ameisenalgorithmen.....	17
A.2.3	Werkstoffkunde – Simulated Annealing.....	20
A.2.4	Erderkundung – Tabu Search	21
A.2.5	Bewertung der Metaheuristiken für diese Arbeit	22
B	Anhang B – Instanzen	23
B.1	Die Solomon Instanzen.....	23
B.2	Die Instanzengruppe ZUE500*.....	25
B.2.1	Rahmenbedingungen.....	25
B.2.2	Ort-Zeitliche Stopplage (Übersicht)	26
C	Anhang C - Untersuchungen zur Instanz ZUE500E.....	29
C.1	Entwicklung der Fahrzeuganzahl	29
C.1.1	HSav0 - Ohne Fuhrparkbegrenzung	29
C.1.2	HSav1 - Mit Fuhrparkbegrenzungsaufweitung	33
C.2	Entwicklung der Tourstruktur.....	35
D	Anhang D – Beispiel km-Regression	39
D.1	Km-Regressionsversuche für ZUE500e	39
D.1.1	2-Parametrische Funktionen.....	39
D.1.2	3-Parametrische Funktionen.....	42
D.1.3	Zwischenergebnis	45
D.2	Km-Regressionsversuche für ZUE500efg	46
D.3	Km-Regressionsversuche für R205#.....	47

A Anhang A – Verfahren für das VRSP

A.1 Klassische Verfahrensansätze für VRSPs

Im folgenden werden die für diese Arbeit wesentlichen konstruktiven Verfahren und Verbesserungsverfahren kurz erläutert.

A.1.1 Konstruktive Verfahren

Konstruktive Verfahren erzeugen zu einem definierten Problem eine gültige Ausgangslösung, die ggf. im Anschluss noch durch ein Verbesserungsverfahren modifiziert wird. In dieser Arbeit werden insbesondere das Savings- und das Insertions-Verfahren genutzt.

A.1.1.1 Das Savings-Verfahren

Das Savings-Verfahren zählt nach [Domschke, 1997, S.243] zu den „am meisten diskutierten“ und dem „am häufigsten in der Praxis eingesetzte[n] Verfahren zur Lösung knotenorientierter VRP ...“.

Die Grundidee dieses Verfahren ist intuitiv leicht verständlich.

Grundidee:

Es sind 2 Kunden i und j vom Depot 0 aus zu bedienen. Werden diese Kunden jeweils über Stichtouren bedient, so werden in der ersten Tour $(0,i,0)$ die Strecken $[0,i]$ und $[i,0]$, und in der zweiten Tour $(0,j,0)$ die Strecken $[0,j]$ und $[j,0]$ zurückgelegt. Werden beide Kunden hingegen in einer kombinierten Tour $(0,i,j,0)$ bedient, so sind insgesamt nur die drei Strecken $[0,i],[i,j]$ und $[j,0]$ zurückzulegen. Die km-Ersparnis (engl. saving) ist somit:

$$[0,i]+[i,0]+[0,j]+[j,0] - ([0,i]+[i,j]+[j,0]) = [i,0]+[0,j] - [i,j].$$

Das selbe Ergebnis ergibt sich beim Verschmelzen zweier beliebiger Touren an den Randpunkten i und j .

Formal lässt sich dieses Verfahren in Anlehnung an [Neumann, Morlock 1993, S472] folgendermaßen beschreiben.

Es seien:

- ▶ $i, j :=$ Knoten (Lieferpunkte, Kunden)
- ▶ $0 :=$ Knoten (Depot)
- ▶ $(0, i, 0) :=$ Stichtour (Pendeltour) vom Depot 0 zum Lieferpunkt j und zurück

- ▶ $(0, \dots, i, \dots, j, \dots, 0) :=$ Rundtour um das Depot 0 mit den Lieferpunkten i und j
- ▶ $L :=$ eine nach nichtwachsenden Werten geordnete Liste
- ▶ $t_{ij} :=$ Maß für die Entfernung von Lieferpunkt i zu Lieferpunkt j (z.B. Fahrzeit von i zu j , oder Entfernung zwischen i und j , oder gewichtetes gemischtes Maß für Fahrzeit und Entfernung zwischen i und j)
- ▶ $n :=$ die Anzahl der zu bedienenden Lieferpunkte
- ▶ $s_{ij} := t_{i0} + t_{0j} - t_{ji}$ (saving)
Dieser Wert beschreibt die Einsparung (z.B. km-Ersparnis) bei Verschmelzung der beiden Touren $(0, \dots, i, 0)$ und $(0, j, \dots, 0)$ zur Tour $(0, \dots, i, j, \dots, 0)$.
- ▶ Zulässiges Saving: Das Saving s_{ij} heiÙe „zulässig“ gdw. die Verbindung der zugehörigen Touren $(0, \dots, i, \dots, 0)$ und $(0, \dots, j, \dots, 0)$ möglich ist, d.h.
 - ▶ i und j sind Tourrandpunkte unterschiedlicher Touren
 $[(0, \dots, i, \dots, 0) = (0, \dots, i, 0) \& (0, \dots, j, \dots, 0) = (0, j, \dots, 0)$ oder
 $(0, \dots, i, \dots, 0) = (0, i, \dots, 0) \& (0, \dots, j, \dots, 0) = (0, \dots, j, 0)]$
 und $(0, \dots, i, 0) \neq (0, j, \dots, 0)$
 - ▶ die Kapazitäts- und Zeitrestriktionen der durch Verschmelzung an den Tourrandpunkten i und j entstehenden Tour werden nicht verletzt.

Algorithmus

Schritt 1

Bilde die Pendeltouren $(0, i, 0)$ für $i=1, \dots, n$.

Berechne für alle $i, j=1, \dots, n$ mit $i \neq j$ die savings s_{ij} , und speichere die positiven savings s_{ij} in einer nach nichtwachsenden Werten geordnete Liste L .

Schritt 2

Falls $L = \emptyset$ terminiere

Andernfalls entferne das erste (d.h. das größte) Element aus L , etwa s_{ij} .

Ist s_{ij} ein zulässiges Saving, so verschmelze die entsprechenden beiden Touren mit der Verbindung (i, j)

Gehe zu Schritt 2

Beim Ablauf des Savings-Algorithmus wachsen die Touren i.d.R. von außen nach innen, da die peripheren Lieferpunkte meist das höchste Saving zueinander aufweisen. Durch Einführung sogenannter „Tourformparameter“ (z.B. α) in die Savingsfunktion (z.B. $s_{ij} := t_{i0} + t_{0j} - \alpha t_{ji}$) lässt sich die Tourstruktur steuern (Zwiebelschalentourstruktur vs. Strahlentourstruktur).

A.1.1.2 Insertion Verfahren

Ein weiteres elementares konstruktives Tourenplanungsverfahren ist das Einfügeverfahren (Insertion Verfahren). Beispielhaft sei hier das Verfahren von Mole und Jameson beschrieben [Mole, Jameson 1976, S 503–511].

Es seien:

- ▶ c_{ij} := Kostenmaß für die Strecke zwischen den Lieferpunkten i und j (z.B. Strecke in km, Fahrzeit in Minuten, gewichtetes Mittel aus Fahrzeit und Strecke, ...)
- ▶ λ, μ Gewichtungsfaktoren
- ▶ $\alpha(i, g, j) = c_{ig} + c_{gj} - \lambda c_{ij}$: Kostenmaß für das Einfügen des Lieferpunktes g in eine bestehende Tour zwischen die Lieferpunkte i und j
- ▶ $\beta(i, g, j) = \mu c_{0g} - \alpha(i, g, j)$: Opportunitätsmaß für die Direktbelieferung bei Nichteinfügen in eine bestehende Tour¹
- ▶ N := Menge aller Lieferpunkte mit $N = T \cup A \wedge T \cap A = \emptyset$
- ▶ T := Menge aller bereits in einer Tour verplanten Lieferpunkte
- ▶ A := Menge aller noch nicht verplanten Lieferpunkte
- ▶ g := der Index des aktuell betrachteten Lieferpunktes

Algorithmus

Wähle ein $g \in A$

Schritt 1 (*)

Bilde Pendeltour $(0, g, 0)$
Lösche g aus A , und füge g in T ein.
Terminiere, falls $A = \emptyset$

Schritt 2 (**)

Berechne für jedes $g \in A$ die minimalen Einfügekosten

$$\alpha'(i_g, g, j_g) = \min_{i, j \in T} \alpha(i, g, j)$$

und den zugehörigen Wert

$$\beta'(i_g, g, j_g) = \mu c_{0g} - \alpha'(i_g, g, j_g).$$

Falls möglich: Füge den Ort g^* , für den gilt

$$\beta'(i_{g^*}, g^*, j_{g^*}) = \max_{i, j \in T, g \in A} \beta'(i, g, j)$$

zwischen die Lieferpunkte i_{g^*} und j_{g^*} ein - dies ist der Lieferpunkt mit den höchsten Direktbelieferungs-Opportunitätskosten. Lösche g aus A , und füge g in T ein.

Falls Einfügen nicht möglich:
gehe zu (*)

Schritt 3

Führe Reihenfolgeoptimierung für die Tour durch, in die g eingefügt wurde.
Ist $A \neq \emptyset$, gehe zu (**), sonst terminiere.

Durch die entsprechende Wahl der Parameter λ und μ können verschiedenartige

¹ Anmerkung: Damit die beiden Fahrtanteile vom und zum Depot berücksichtigt werden, ist μ so zu wählen, dass es den Faktor 2 bereits beinhaltet.

Einfügeregeln festgelegt werden:

- ▶ soll jeweils der Lieferpunkt eingefügt werden, welcher die minimalen zusätzlichen Kosten durch Einfügen zwischen ein aktuelles Nachbarpaar i und j verursacht, so ist $\lambda = 1$ und $\mu = 0$ zu wählen,
- ▶ soll jeweils der Lieferpunkt eingefügt werden, der die Kosten zu seinen neuen zwei benachbarten Lieferorten minimiert, so ist $\lambda = \mu = 0$ zu wählen,
- ▶ soll jeweils der Lieferpunkt eingefügt werden, der am weitesten vom Depot entfernt ist, so ist $\lambda > 0$ und $\mu = \infty$ zu wählen.

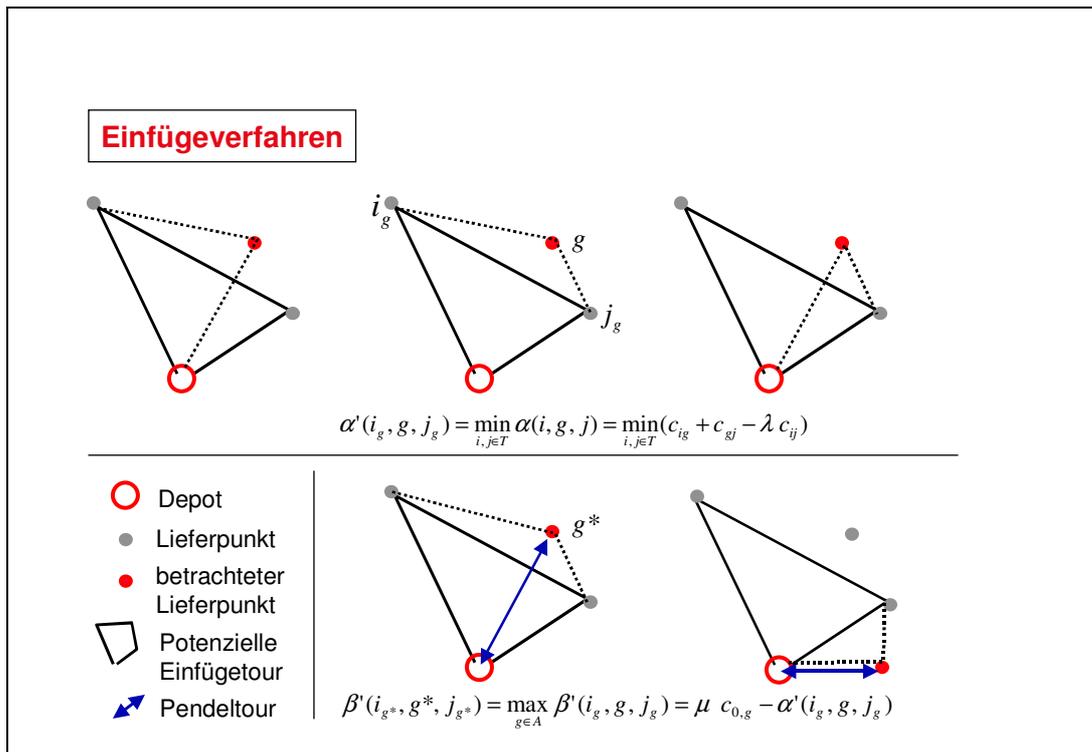


Abbildung 1: Beispiel zum Insertion Verfahren (Prinzipiskizze)

A.1.1.3 Verfahren für enge Zeitfenster nach Daduna

Der Erfolg von Tourenplanungsverfahren bei Vorliegen von engen Zeitfenstern ist aus praktischer Sicht oft unbefriedigend. Das Kernproblem liegt darin, dass formal unzulässige (aber eventuell akzeptierbare) Lösungen, abgelehnt werden. Eine Lösung ist ggf. akzeptabel, falls „einzelne“ Zeitrestriktionen „minimal“ verletzt werden.

[Daduna 2002] schlägt daher ein Verfahren vor, dass die dezidierte Verletzung einzelner Zeitfenster erlaubt, wenn dadurch die Anzahl der benötigten Fahrzeuge signifikant gesenkt werden kann (d.h. es wird ein VRSP-STW gelöst).

In einem ersten Schritt wird eine Anfangslösung generiert, die die Fahr- und

Wartezeiten minimiert², jedoch kein Augenmerk auf die benötigte Anzahl Fahrzeuge legt.

In einem zweiten Schritt wird diese gültige Lösung durch sukzessives Aufweichen der Zeitfenster (Verletzung der ursprünglichen Restriktionen) in Hinsicht auf die Anzahl der benötigten Fahrzeuge verbessert. Die Zeitfensterverletzungen werden dabei mit Strafkosten belegt.

Wie bei allen „Strafkostenansätzen“ für Zeitfensterverletzungen ist die Parametrisierung der Strafkosten ein zu lösendes Problem. Aus dem von [Daduna 2002] genannten Beispiel geht implizit³ hervor, dass die Strafkosten stark progressiv mit zunehmender Verletzung der ursprünglichen Zeitfenster steigen sollten.

A.1.2 Verbesserungsverfahren

Verbesserungsverfahren sind Verfahren, die ausgehend von einer bereits bestehenden Ausgangslösung versuchen ein bestimmtes Zielkriterium zu verbessern.

Aus der Fülle von denkbaren Verbesserungsverfahren soll im Folgenden insbesondere ein Verbesserungsverfahren zur Reihenfolgeoptimierung dargestellt werden.

A.1.2.1 Prinzip von Verbesserungsverfahren

Verbesserungsverfahren arbeiten häufig nach dem Prinzip der **lokalen Suche**:

„Hierbei versucht man eine bekannte zulässige Lösung x zu verbessern, indem man x mit allen Elementen einer „Umgebung“ $U(x)$ von x vergleicht und hieraus eine zulässige Lösung mit (bei einem Minimumproblem) kleinstem Zielfunktionswert auswählt. $U(x)$ ist dabei eine Teilmenge des zulässigen Bereichs mit der Eigenschaft, dass $x' \in U(x)$ genau dann gilt, wenn $x \in U(x')$ ist. Bei der lokalen Suche geht man also wie folgt vor: Sei F die zu minimierende Zielfunktion. Dann generiert man so lange zulässige Lösungen von $x' \in U(x)$ und setzt für den Fall $F(x') < F(x)$ $x := x'$, bis $F(x') \geq F(x)$ für alle $x' \in U(x)$ ist.“ [Neumann, Morlock 93; S. 403f]

Im unten genannten Fall der Reihenfolgeoptimierung wäre:

- ▶ x : die Ausgangslösung der Tourreihenfolge aus dem vorgeschalteten Tourbildungsverfahren.

² durch Lösung des klassischen linearen Zuordnungsproblems nach der „Ungarischen Methode“ gemäß [Domschke, 1995, S208 ff] mit Modifikationen

³ 1 min. Verletzung: 1 Strafkosteneinheit;
2 min. Verletzung: 10 Strafkosteneinheiten

- ▶ $U(x)$: die Menge aller gültigen Reihenfolgen über die Stops der Tour x , die nach den Regeln des r -opt erzeugt werden können.
- ▶ $x' \in U(x)$: eine gültige Tourreihenfolge, die sich durch r -opt erzeugen lässt.

A.1.2.2 Das r -opt-Verfahren

r -opt-Verfahren sind klassische Verbesserungsverfahren zur Lösung des TSP. Sie setzen voraus, dass bereits eine gültige Lösung für das TSP vorliegt. Liegt diese Lösung nicht vor, so kann sie beispielsweise mit dem Algorithmus „der sukzessiven Einbeziehung von Knoten“ [Neumann, Morlock 1993, S.445 f.] oder dem Verfahren von Christofides [Neumann, Morlock 1993, S.448 f.] erzeugt werden.

In der Tourenplanung wird das r -opt-Verfahren oft im Rahmen der Dekompositionsstrategie als Verbesserungsverfahren nach einem vorhergehenden Zuordnungsverfahren (z.B. Sweep oder Savings s.u.) eingesetzt. Eine verständliche Beschreibung von r -opt-Verfahren findet sich in [Domschke, 1997, S.117]:

*„Ein **r -optimales Verfahren** ist ein Vertauschungsverfahren bei dem in jeder Iteration versucht wird, die aktuelle Rundreise ρ durch Austausch von r Kanten gegen r „andere“ Kanten (r -Tausch) zu verbessern. Dabei müssen die Menge der aus ρ zu entfernenden und die Menge der dafür aufzunehmenden Kanten nicht disjunkt sein; es können auch weniger als r Kanten vertauscht werden. Eine Rundreise ρ heißt r -optimal, wenn es nicht möglich ist, eine kürzere Rundreise als ρ durch einen r -Tausch zu finden.“*

Formal lässt sich dieses Verfahren⁴ für den Fall $r=2$ in Anlehnung an [Neumann, Morlock 1993, S452] folgendermaßen beschreiben:

Es seien:

- ▶ $v, \mu :=$ Zählindizes
- ▶ $i_v :=$ Knoten
- ▶ $[i_v, i_\mu] :=$ Kante zwischen den Knoten i_v und i_μ
- ▶ $c_{i_v, i_\mu} :=$ Länge der Kante zwischen den Knoten i_v und i_μ
- ▶ $K = [i_1, \dots, i_n, i_1] :=$ Hamiltonscher Kreis⁵

⁴ Die Verfahrensvarianten und –details sind ausführlich beschrieben in [Domschke, 1997, S.117 f.] und [Neumann, Morlock 1993, S.451 f.].

⁵ [Neumann, Morlock 1993, S.439 f.]

Algorithmus

(*) Für $v=1, \dots, n-2$

Für $\mu=v+2, \dots, n-1$

Falls $c_{i_v i_{v+1}} + c_{i_\mu i_{\mu+1}} > c_{i_v i_\mu} + c_{i_{v+1} i_{\mu+1}}$,

ersetze in $K=[i_1, \dots, i_n, i_1]$ die Kanten $[i_v, i_{v+1}], [i_\mu, i_{\mu+1}]$ durch $[i_v, i_\mu], [i_{v+1}, i_{\mu+1}]$,

nummeriere um, so dass wieder $K=[i_1, \dots, i_n, i_1]$ gilt,

gehe zu (*)

Falls $c_{i_v i_{v+1}} + c_{i_n i_1} > c_{i_v i_n} + c_{i_{v+1} i_1}$,

ersetze in $K=[i_1, \dots, i_n, i_1]$ die Kanten $[i_v, i_{v+1}], [i_n, i_1]$ durch $[i_v, i_n], [i_{v+1}, i_1]$,

nummeriere um, so dass wieder $K=[i_1, \dots, i_n, i_1]$ gilt,

gehe zu (*)

Terminiere, wenn in einer gesamten v -Schleife keine Verbesserung mehr erzielt wurde.

Um ein intuitives Verständnis für die Arbeitsweise des Verfahrens zu erzielen, soll die folgende Grafik dienen.

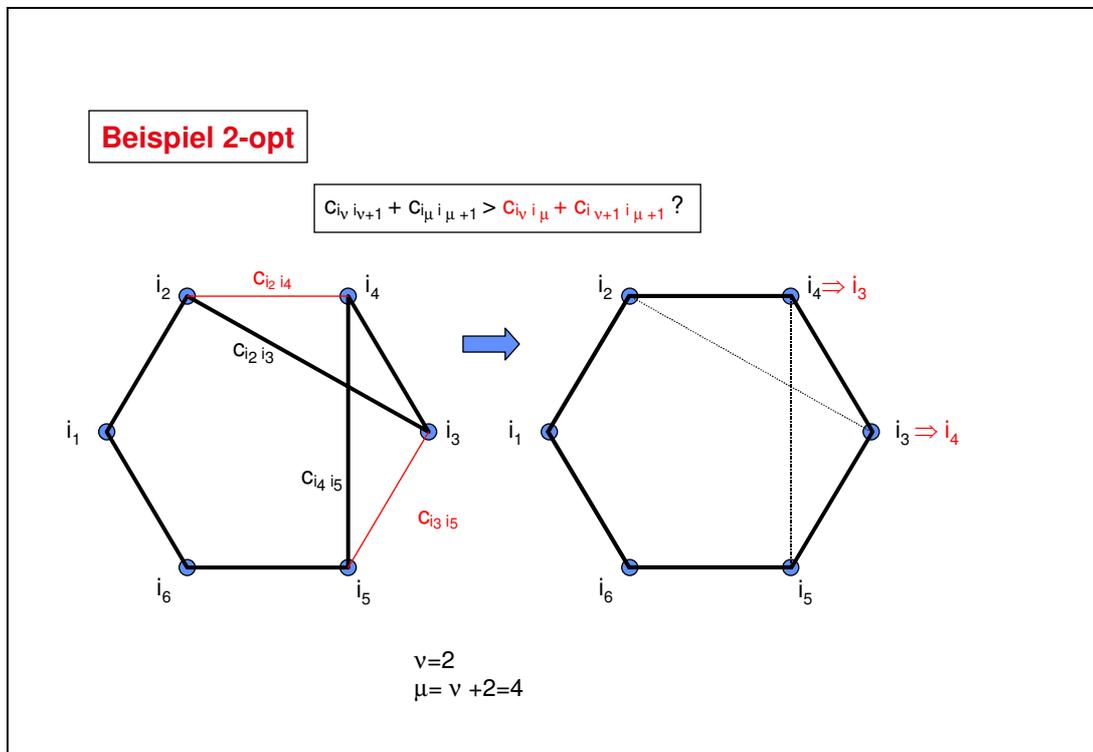


Abbildung 2: Beispiel zum r -opt Verfahren ($r=2$)

A.1.2.3 Tourübergreifende Austauschverfahren

Im Gegensatz zum vorgehend beschriebenen Reihenfolgeoptimierungsverfahren ist es das Ziel von tourübergreifenden Austauschverfahren, nicht nur Stopps innerhalb einer Tour, sondern Tourteilstücke (Strings) über Touren hinweg zu tauschen.

Die wesentlichen Verfahren hierzu sind:

- ▶ **String Relocation**
Hierbei wird aus einer Tour ein Teilstück herausgeschnitten und in eine andere Tour eingefügt. (Die Länge des Teilstücks kann dabei variieren)
- ▶ **String Cross**
Hierbei werden die Touren an jeweils einer Stelle durchtrennt, und die jeweils zweiten Tourhälften überkreuz an die ersten Tourhälften angeschlossen.
- ▶ **String Exchange**
Hierbei werden zwei Teilstücke zwischen zwei Touren ausgetauscht. (Die Länge der Teilstücke kann dabei variieren. Sind die Teilstücke nur 1 Element groß, so werden einzelne Stops zwischen Touren getauscht.)

Durch Anwendung dieser Operationen auf die „Leere Tour“ (Tour ohne Stopps) können neue Touren entstehen. Umfasst ein Tourteilstück alle Stops einer Tour können bei der String Relocation Touren aufgelöst werden.

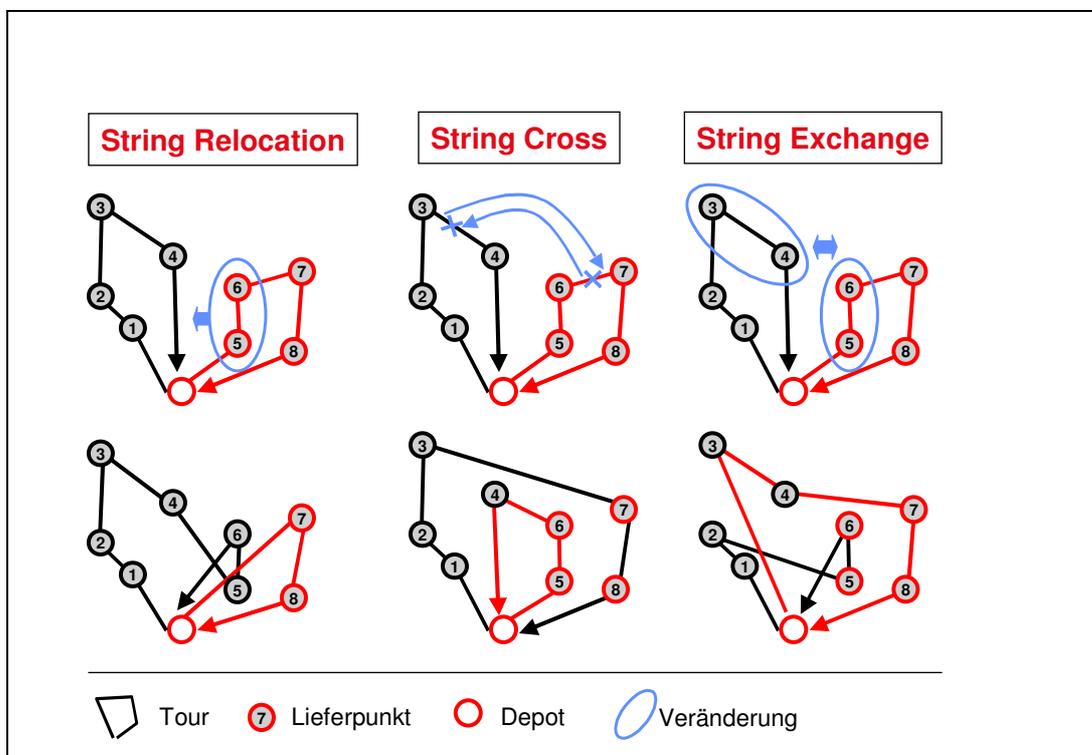


Abbildung 3: Tourübergreifende Austauschverfahren

Weitere tourübergreifende Austauschverfahren sind u.a. beschrieben in [Baumung, 2003] und [Gendreau et al; 1992].

A.2 Metaverfahren für VRSPs

Nach [Blum 2004, S.4] ist bis heute der Begriff „Metaheuristik⁶“ nicht allgemein

⁶ Aus dem Griechischen: μετά (dahinter, darüber, auf einem übergeordneten Niveau) und εвриσκειν (finden)

anerkannt definiert. Bevor daher in diesem Kapitel auf die Funktionsweise von Metaverfahren eingegangen wird, soll einleitend kurz definiert werden, was in dieser Arbeit unter Metaverfahren (bzw. Metaheuristiken) zu verstehen ist.

Definition „Metaverfahren“:

*Metaverfahren (oder Metaheuristiken) sind iterative übergeordnete Verfahren, die untergeordnete Heuristiken (auf einer „Meta-Ebene“) steuern. Die Steuerungsprinzipien der Metaebene sind dabei ex ante bekannt und haben i.d.R. eine Entsprechung in der realen „Welt 1“⁷. Die Steuerungsprinzipien entstammen dabei jedoch nicht der Problem-Domäne, d.h. sie sind nicht problemspezifisch. Sind die Steuerungsprinzipien aus der Natur „abgeschaut“, so spricht man auch von „**naturalogen Verfahren**“. Die Hauptaufgabe des Metaverfahrens ist dabei, das Festfahren der untergeordneten Heuristik in einem lokalen Optimum zu verhindern.*

Zur besseren Verständlichkeit werden für die im Folgenden dargestellten Verfahren die Domänen des Lösungsprinzips in der Überschrift kurz genannt.

A.2.1 Evolution – Genetische Algorithmen

Kernidee

Genetische Algorithmen (GAs) ahmen die genetische Evolution nach. Die genetische Evolution basiert auf einfachen Prinzipien, nämlich:

- ▶ **Codierung** von Körpereigenschaften (Erbanlagen) in Genen
- ▶ **Vererbung** von Erbanlagen von den Eltern auf die Kinder durch die **Kombination** der Erbanlagen der Eltern zu den Erbanlagen der Kinder
- ▶ **Mutation**, d.h. zufällige Veränderung einzelner Erbanlagen
- ▶ **Selektionsprozesse**, die die Fortpflanzungswahrscheinlichkeit von Eltern mit hoher „**Fitness**“ steigern und die von „geschwächten“ Eltern senken („survival of the fittest“)

Umsetzung im Algorithmus

Diese Prinzipien lassen sich als Algorithmus implementieren. Eine Formulierung als Pseudocode nach [Maziejewski 1992, S.19 ff] stellt dieses algorithmische Prinzip kurz und prägnant dar.

⁷„Welt 1“ nach [Popper 1994]

```

proc Genetic Algorithm
begin
  initialize (population)
  improve (population)
  repeat
    select (population, mates)
    if (random flip) then
      crossover (mates, offspring)
    else
      mutate (mate, offspring)
    improve (offspring)
    include (population, offspring)
  until termination
end

```

Abbildung 4: Prinzip des Genetischen Algorithmus nach [Maziejewski 1992]

Wichtig bei diesem Verfahren ist die sogenannte **Fitness** – ein Wert, der ein sinnvolles Maß für die Güte der jeweiligen Lösung (**Individuum**) darstellt. Außerdem wird i.d.R. das Ergebnis eines Iterationsschritts als **Generation** bezeichnet. Die Ausprägung der Lösung (z.B. der Tourenplan) wird i.d.R. als **Chromosom** bezeichnet.

initialize (population): Zuerst ist eine Menge möglichst vielfältiger (großer **Gen-Pool**) zulässiger Anfangslösungen (**Population**) zu generieren. Im Anwendungsfall für VRPs kann dies beispielsweise durch Einsatz einer Insertions-Heuristik mit verschiedenen Initialisierungen erfolgen. Auch die Savings-Heuristik wäre im Prinzip geeignet – da der Savings aber i.d.R. strukturähnliche Lösungen liefert, sollten nicht ausschließlich durch den Savings generierte Individuen die Ausgangspopulation ausmachen, da sonst der Gen-Pool zu homogen wäre.

improve (population): Die Lösungen der Ausgangspopulation können anschließend durch Verbesserungsverfahren (z.B. r-opt) noch verbessert werden, bevor das eigentliche genetische Verfahren beginnt.

select (population, mates): Aus der Population werden 2 **Eltern** (mates) ausgewählt, die einen **Nachkommen** (offspring) erzeugen sollen. Dabei ist das Selektionskriterium entscheidend.

- ▶ Eine zufällige Selektion (random selection) erhält zwar lange einen heterogenen Gen-Pool, lässt aber keine schnelle Erzeugung von Nachkommen mit höherer Fitness erwarten.
- ▶ Eine Selektion der Eltern nach der höchsten Fitness (best selection) lässt umgekehrt schnell eine Erzeugung von Nachkommen mit höherer Fitness erwarten, dafür verliert der Gen-Pool auch schnell an Vielfalt, und somit könnte das Verfahren leicht in einem lokalen Optimum enden.
- ▶ Daher wird an dieser Stelle oft eine gewichtete Mischung selektiert (harmonic selection), d.h. Individuen mit höherem Fitness-Grad werden mit einer

höheren Wahrscheinlichkeit ausgewählt als Individuen mit niedrigem Fitness-Grad

crossover (mates, offspring): Aus den beiden Eltern-Lösungen wird nun nach einer definierten Regel ein Nachkomme generiert (**Rekombination**). Dieser zentrale Schritt im Verfahren muss aus den zulässigen Eltern-Ausgangslösungen eine zulässige Nachkommens-Lösung generieren und ist daher nichttrivial. Im Prinzip gibt es zwei Ideen zu dieser Rekombination:

- ▶ Die Chromosomen-Rekombination (string oriented crossover)
In der Biologie sind die physischen Eigenschaften eines Individuums durch dessen Chromosomen (Informationssequenzen) codiert. Durch Kombination von Informationsteilsequenzen der Eltern entstehen die Chromosome der Nachkommen. (Details s. [Wendt, 1995], Kap. 4.1.2.)
In Analogie dazu wird daher die Ausprägung der Lösung als Vektor (string) dargestellt (**Codierung**). Da dieser die Lösung beschreibende Vektor jedoch keine Information über die Zulässigkeit der Lösung beinhaltet, stellt i.d.R. eine reine Rekombination der Elternchromosomen keine gültige Lösung dar. Durch **Reparaturverfahren** lässt sich dieses Problem zwar lösen [Maziejewski 1992, S.23f]. Eine intuitiv einleuchtende Vererbung von Lösungseigenschaften von den Eltern auf die Nachkommen ist aber nur schwer zu erzielen.
- ▶ Die problemorientierte Rekombination (problem oriented crossover)
Bei diesem Ansatz wird von der reinen Rekombination der Elternchromosomen abgewichen, und das Domänenwissen über das zugrundeliegende Problem (hier ein VRP) genutzt, um eine sinnvolle Rekombination aus den Lösungsmerkmalen der Eltern zu erzielen

Beispiele:

Für die reine **Chromosomen-Rekombination** werden zunächst die Lösungseigenschaften in Chromosomen codiert. In diesem Fall beschreibt das Chromosom für jeden Lieferpunkt den jeweiligen Nachfolger (Element i des Vektors beschreibt den Nachfolger des Lieferpunkts i). Durch den Rekombinationsvektor werden aus den Chromosomen der Elternteile diejenigen Teilsequenzen bestimmt, die rekombiniert werden sollen.

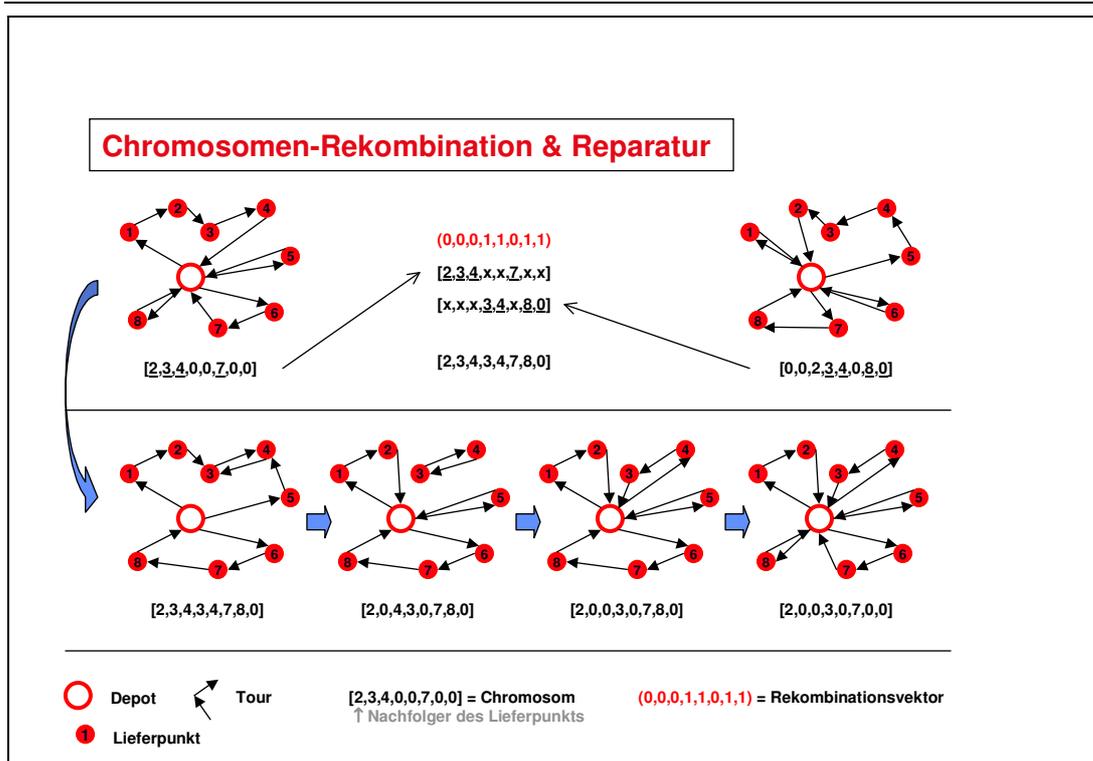


Abbildung 5: Chromosomen-Rekombination

Die so entstehende Lösung ist noch keine zulässige Lösung, und muss daher noch nach folgendem Muster repariert werden:

1. Entferne mehrfach auftretende Lieferpunkte aus dem Chromosom, und ersetze sie durch 0 (=Depot)
2. Eliminiere Zyklen ohne Depot
3. Splitte Touren auf, die den Restriktionen (z.B. max. Tourdauer) nicht genügen

Für die **problemorientierte Rekombination** werden die Lösungseigenschaften nicht in Chromosomen codiert. In diesem Beispiel werden die Lösungseigenschaften als vollständige Liste der realisierten Verbindungen gespeichert, die nicht das Depot enthalten. Die Rekombination der Eltern-Lösungen erfolgt problemorientiert in Anlehnung an den Savings-Algorithmus:

- ▶ Vereinige die Liste der realisierten Verbindungen beider Elternteile
- ▶ Sortiere die Liste nach aufsteigender Kantenlänge
- ▶ Initiere die Nachkommenslösung durch einen Tourenplan mit ausschließlich Stichtouren
- ▶ Modifiziere die Initiallösung durch sukzessives Einfügen der Kanten aus der sortierten Liste
- ▶ Eine Kante kann nicht in den aktuellen Tourenplan eingefügt werden, wenn der Ausgangspunkt der Kante bereits einen Nachfolger ungleich dem Depot hat – in diesem Fall wird die betreffende Kante in der Liste übergangen.

Weitere Rekombinationsmethoden finden sich u.a. in [Potvin & Bengio 1996] als *sequence-based crossover (SBX)* und *route-based crossover (RBX)*.

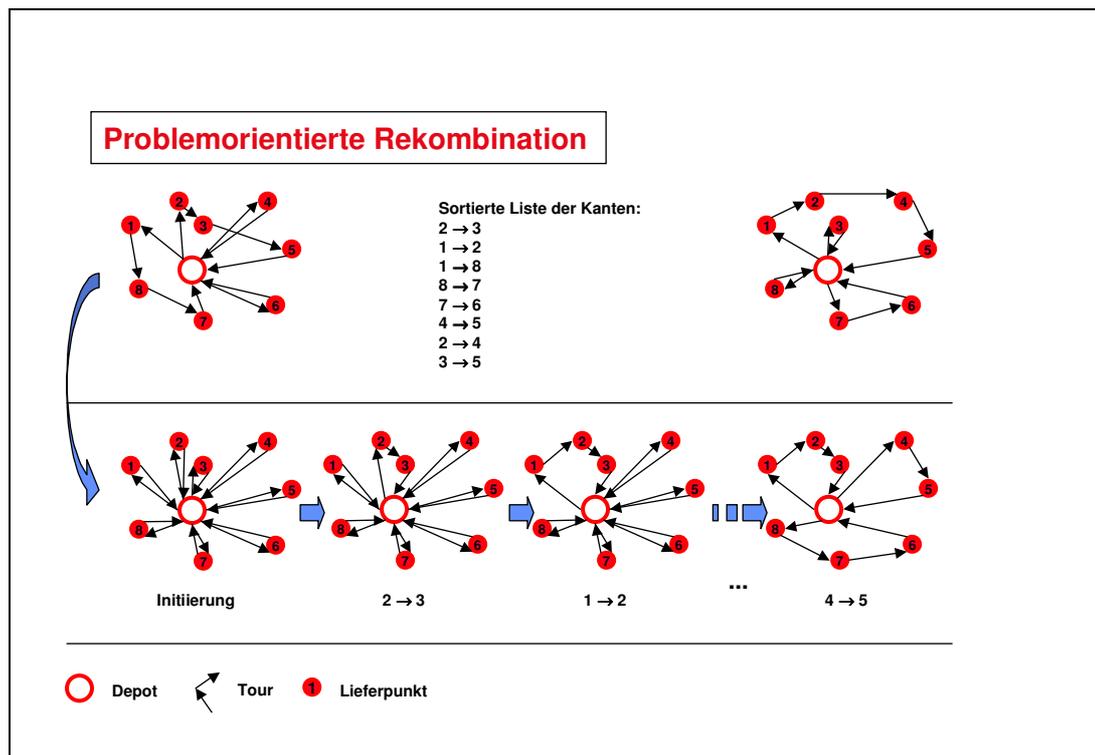


Abbildung 6: Problemorientierte Rekombination

mutate (mate, offspring): Alternativ zur Rekombination kann eine zufällige Mutation eines Elements der Elterngeneration erfolgen. Die Mutation trägt dazu bei, den Gen-Pool zu erweitern. Eine einfache Mutationsregel ist z.B. das zufällige Ersetzen von 20% der Chromosomeneinträge (auch hier ist das oben beschriebene Reparaturverfahren analog anzuwenden, um eine zulässige Lösung zu erhalten). Weitere Rekombinationsmethoden finden sich u.a. in [Potvin & Bengio 1996] als *1M-* und *2M-Mutationsoperatoren*.

Ob die Nachkommens-Lösung durch eine Rekombination zweier Eltern-Lösungen oder eine Mutation entsteht wird durch eine Mutationsrate festgelegt. Gemäß dieser Mutationsrate wird zufällig entschieden, ob eine Mutation oder eine Rekombination erfolgt.

improve (offspring): Die Qualität der Nachkommens-Lösung kann anschließend durch Verbesserungsverfahren (z.B. r-opt) noch verbessert werden.

include (population, offspring): Damit eine Degeneration des Genpools vermieden wird, ist jeweils zu entscheiden, ob der Nachkomme in die Population aufgenommen wird. [Maziejewski 1992, S.27] schlägt ein Ähnlichkeitsmaß zwischen den Individuen vor, und erlaubt eine **Aufnahme** in die Population nur dann, wenn der Nachkomme eine minimale Verschiedenheit zu den bereits in der Population vertretenen Lösungen aufweist, um eine Degeneration des Genpool zu

vermeiden.

termination: Das Verfahren terminiert i.d.R. mit Erreichen einer definierten Anzahl von Generationen. Es ist auch denkbar, dass das Ausbleiben einer Verbesserung des besten Zielfunktionswerts einer Population über eine definierte Anzahl von Generationen das Verfahren terminiert.

Weitere Überlegungen

Eine etwas andere Variante des Genetischen Algorithmus in Hinblick auf die Selektion wird von [Wendt, 1995] beschrieben. Im Gegensatz zu [Maziejewski 1992], der kein Ausschlusskriterium für die Population beschreibt, und bei dem daher zu erwarten ist, dass die Population zwar gebremst durch das Aufnahmekriterium der *include (population, offspring)*-Vorschrift, aber kontinuierlich wächst, schlägt [Wendt, 1995] einen den natürlichen Auslese in der Natur entsprechenden **Ausschluss** von Individuen aus der Population vor, so dass die Populationsgröße konstant bleibt. Die Entscheidung über den Ausschluss erfolgt direkt oder indirekt über die Fitness der jeweiligen Individuen.

[Wendt, 1995, Kap. 4.3.2.] weist in diesem Zusammenhang auf folgende Analogie zur Biologischen Evolution hin:

„Werden Individuen mit hoher Fitness zu häufig zur Produktion von Nachkommen herangezogen, so können kleine, aber wichtige Schemata (Anm.: Teile der Ausprägung der Lösung) in weniger guten Individuen verloren gehen, da diese nicht genug Möglichkeiten erhalten, die Information weiterzugeben. Man spricht in diesem Fall von verfrühter Konvergenz (premature convergence). Die Population degeneriert auf Grund mangelnder Diversifikation des Gen-Pools. Wird dagegen ein zu geringer Selektionsdruck gewählt, so ähnelt die Suche einer Monte-Carlo-Simulation, und ist somit bei großen Suchräumen völlig ineffizient.“

Schließlich soll noch auf einen wesentlichen Aspekt hingewiesen werden.

Die Effizienz von genetischen Algorithmen hängt wesentlich von der jeweils gewählten Modellierung ab. Die rekombinationsfähige Codierung von Tourenplänen durch Chromosomen bzw. die gewählten Rekombinations- und Reparaturverfahren sowie die Mutationsvorschriften müssen das Domänenwissen über die Aufgabenstellung gut abbilden („Without problem specific knowledge, the GA was really useless as one could see in the test runs“ [Maziejewski 1992, S. 55]).

D.h. um diesen Ansatz näher für den Anwendungsfall des eCommerce homedelivery im Lebensmittelhandel untersuchen zu können, müsste erst ein Weg gefunden werden, das Domänenwissen nachhaltig in einem spezifischen GA zu manifestieren.

A.2.2 Zoologie – Ameisenalgorithmen

Kernidee

Ameisen suchen durch „random walks“ immer wieder die Nestumgebung nach neuer Nahrung ab. Dabei können sie entlang ihrer eigenen Laufwege sogenannte „Pheromone“ absondern, und so „Pheromonspuren“ legen.

Haben sie irgendwo Nahrung gefunden, so bewerten sie Qualität und Quantität der Nahrungsquelle, und tragen ein Stück der Nahrung ins Nest zurück. Die Stärke der Pheromonabgabe auf dem Rückweg ist abhängig von der Qualität und Quantität des Nahrungsvorkommens.

Stoßen andere Ameisen auf eine derartige Duftspur, wächst die Wahrscheinlichkeit, dass sie dieser Spur folgen, mit der Pheromonintensität. Wenn mehrere Ameisen auf dem Rückweg von der selben vielversprechende Nahrungsquelle sind, so verstärkt sich die Pheromonkonzentration auf diesem Weg. Wird ein Weg länger nicht benutzt, so verflüchtigt sich die Pheromonkonzentration. Nach diesem Prinzip bilden sich die in der Natur beobachtbaren Ameisenstraßen.

Die Idee von Ameisenalgorithmen (ACOs⁸) ist es, dieses Prinzip nachzuahmen.

Umsetzung im Algorithmus

Der Algorithmus kann frei nach [Blum 2004] folgendermaßen beschrieben werden.

```
proc Ant Algorithm  
begin  
  repeat while termination conditions not met  
    Schedule activities  
      AntBasedSolutionConstruction()  
      PheromoneUpdate()  
      DaemonActions() {optional}  
    end Schedule activities  
  end repeat  
end
```

Abbildung 7: Prinzip des Ameisenalgorithmus

Die einzelnen Prozessschritte sehen dabei folgendermaßen aus.

AntBasedSolutionConstruction(): Ein konstruktives Verfahren generiert eine Menge von Lösungen

- ▶ durch Rekombination von Komponenten von bereits bekannten Lösungen (dies entspricht dem probabilistischen Folgen von Pheromonspuren) und

⁸ Ant colony algorithms

- ▶ durch Generierung neuer Lösungen (dies entspricht der weiteren Erkundung des Lösungsraums durch random walks).

Die Pheromonstärke, die den Lösungskomponenten anhaftet, entspricht dabei der „Güte“ der zugehörigen bereits bekannten Lösung. Die Wahrscheinlichkeit, mit der eine Lösungskomponente bei der Rekombination berücksichtigt wird, steigt mit wachsender Pheromonstärke.

PheromoneUpdate(): Zuerst werden alle bisherigen Pheromonlevel gesenkt (dies entspricht der natürlichen Pheromonreduktion entlang der Wege durch Verflüchtigung über die Zeit). Anschließend werden einige (oder alle) der aktuellen (und ggf. auch älteren) Lösungen ausgewählt, und das Pheromonlevel ihrer jeweiligen Lösungskomponenten erhöht. Je besser eine Lösung ist, desto stärker steigt der Pheromonlevel der jeweiligen Lösungskomponenten.

DaemonActions(): DaemonActions sind Aktionen die zentralseitig (d.h. nicht durch einzelne Ameisen) ausgeführt werden. Diese DaemonActions können beispielsweise

- ▶ lokale Suchverfahren sein, die die gefundenen Lösungen verbessern oder
- ▶ das Sammeln von globalen Informationen (z.B. der Lösungsvergleich), und
 - ▶ die gezielte zusätzliche Verstärkung der Pheromonlevels guter Lösungen (Verstärkung der Konvergenz)
 - ▶ die Dämpfung von Pheromonlevelunterschieden (Aufrechterhaltung der Diversität des Lösungsraums)

Um das Prinzip von Ameisenalgorithmen zu verdeutlichen, soll das folgende Beispiel dienen, das eine Anwendung im Bereich der Tourenplanung skizziert.

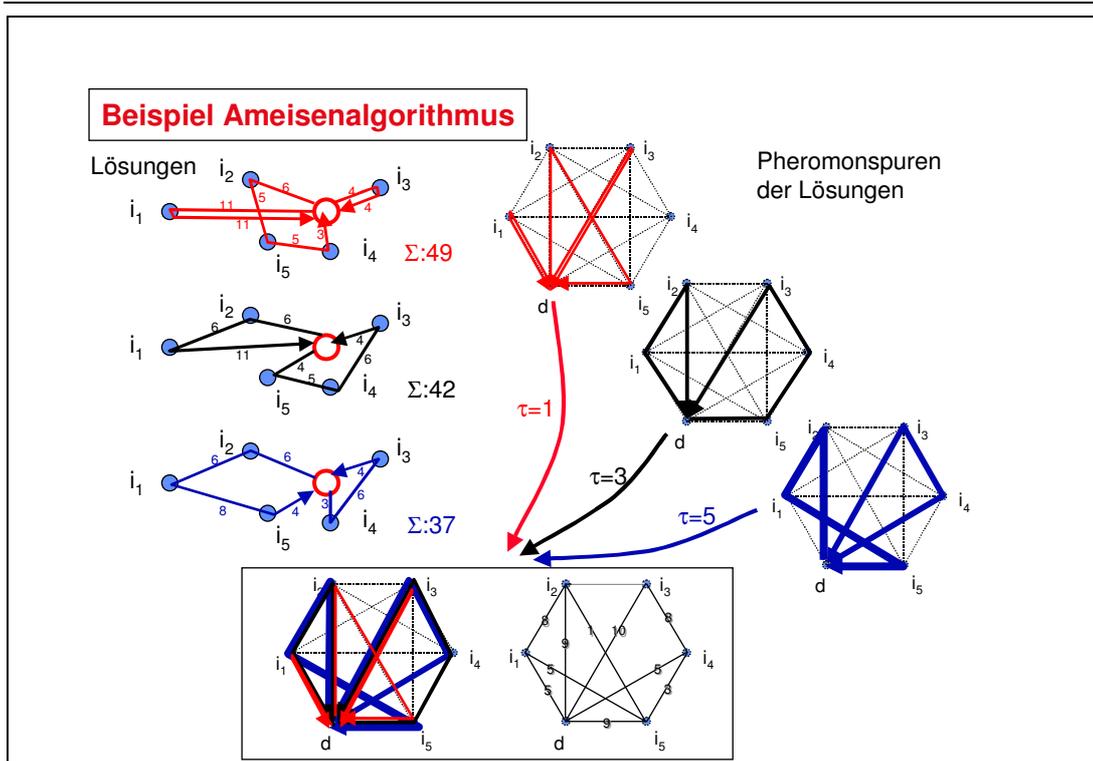


Abbildung 8: Ameisenalgorithmus

Durch ein beliebiges Verfahren wurden die drei Lösungen für ein Tourenplanungsproblem mit 5 Stops und einem Depot generiert (Abbildung oben links).

Die Komponenten (Teiltourstücke) dieser Lösungen (Touren) sind als Kanten eines Grafen dargestellt (Abbildung oben rechts).

Für jede Lösung kann als Gütemaß beispielsweise die Gesamttourweglänge (Σ) berechnet werden. Je nach „Güte“ der Lösung, werden den Lösungskomponenten Pheromonlevel (τ) zugeordnet.

Beispielsweise durch Summierung der Pheromonlevel aus den einzelnen Lösungen werden die Pheromonlevel für die Lösungskomponenten bestimmt (Abbildung unten).

Für die nächste Iteration werden in der *AntBasedSolutionConstruction()* nun probabilistisch die Kanten öfter zur Konstruktion neuer Lösungen genutzt, deren Pheromonlevel am höchsten ist.

A.2.3 Werkstoffkunde – Simulated Annealing

Kernidee

Bei der Abkühlung von Schmelzen findet ein Erstarrungsprozess statt, der nach folgenden wesentlichen Prinzipien abläuft:

- ▶ Die Schmelze strebt einem möglichst niedrigen **Energieniveau** zu.
- ▶ Je höher die **Temperatur** der Schmelze ist, desto größer ist Ihre „**Beweglichkeit**“ (flüssig), je niedriger die Temperatur, desto starrer ist sie (fest). D.h. bei hohen Temperaturen lässt sich die Schmelze noch „umformen“.

Der Prozess des Abkühlens und Erstarrens wird durch das Verfahren Simulated Annealing (SA) nachempfunden.

Umsetzung im Algorithmus

Ein Steuerungsparameter T simuliert die Temperatur der Schmelze, und sinkt sukzessive mit Fortschritt des Verfahrens.

Ausgehend von einer Ausgangslösung L mit einem Energieniveau E wird zufällig eine Lösung L' generiert (Energieniveau E'), die sich nur wenig von der Ausgangslösung L unterscheidet. Das Energieniveau sei umgekehrt proportional zur Zielfunktionsgüte (hohes Energieniveau \Leftrightarrow niedrige Zielfunktionsgüte).

Bei einer Verringerung des Energieniveaus beim Übergang von L zu L' wird die Lösung L durch L' ersetzt.

Bei einer Erhöhung des Energieniveaus wird die Lösung L nur mit einer Wahrscheinlichkeit $p(T, \Delta E)$ durch L' ersetzt. Das von der aktuellen Temperatur T und dem Energieniveauezuwachs $\Delta E = E' - E$ abhängige $p(T, \Delta E)$ sinkt mit steigendem ΔE und fallendem T z.B. nach der Formel [Wendt, 1995; S 116]

$$p(T, \Delta E) = e^{-\frac{\Delta E}{T}}$$

D.h. ein Übergang zu einer schlechteren Lösung kann zwar erfolgen, die Wahrscheinlichkeit einer solchen Verschlechterung ist aber umso geringer je geringer die Temperatur ist, und je stärker das Energieniveau steigen würde.

Anschließend wird das Verfahren mit dem neuen L und abgesenktem T wiederholt, und terminiert nach n Iterationen oder wenn das Energieniveau über k Iterationen hinweg nicht mehr sinkt.

Die Extremwertbetrachtung dieses Vorgehens macht das Prinzip noch klarer.

Würde jede Lösung akzeptiert ($p(T, \Delta E) := 1$), so entspräche das Verfahren einer ungerichteten Zufallssuche. Würden nur bessere Lösungen akzeptiert, würde das Verfahren zielstrebig auf das nächstliegende lokale Optimum zusteuern, und dort terminieren.

Einsatz in der Tourenplanung

Neben der geschickten Parametrisierung des Verfahrens (Dimensionierung von T , Definition von ΔE , ...) ist es wesentlich, *eine Lösung L' zu generieren, die sich nur wenig von der Lösung L unterscheidet*. Ein möglicher Ansatzpunkt ist zum Beispiel die Anwendung der oben beschriebenen Verfahren zum Austausch von Stops zwischen Touren (string cross, string relocation, string exchange). Außerdem sollten an dieser Stelle Verfahren eingesetzt werden, die sowohl zusätzliche Touren entstehen lassen können als auch bestehende Touren auflösen können. So umfasst die Lösungssuche auch Lösungen, bei der die Anzahl der eingesetzten Fahrzeuge nicht der Ausgangslösung entspricht.

Ein Verfahren für den Einsatz eines Simulated-Annealing-Verfahrens für den Anwendungsfall der Tourenplanung findet sich u.a. in [Osman, 1993].

A.2.4 Erderkundung – Tabu Search

Dieses Verfahren lehnt sich an die Erkundung unbekannter geografischer Gegenden an.

Kernidee

„Tabu search“ (TS) heißt auf Deutsch Tabu-Suche oder besser „Suchen mit Tabus“. Ausgehend von einer Ausgangslösung werden in einer definierten „Nachbarschaft“ neue Lösungen gesucht. Wird eine bessere Lösung gefunden, so wird die bisherige Lösung aufgegeben, und von der neuen Lösung aus weitergesucht (tabu move). Auch Lösungen, die **außerhalb des** durch Nebenbedingungen definierten **Zulässigkeitsbereichs** liegen, können durchlaufen werden, werden jedoch mit **Strafkosten** S belegt.

Bei der Bewertung einer Lösung wird unterschieden zwischen der **Zielfunktion** T und der **strafkostenbehafteten erweiterten Zielfunktion** $F = T + S$.

Damit bei der Suche **keine Zyklen** entstehen, werden für eine gewisse Anzahl von Iterationen bereits durchsuchte Teile des Lösungsraums als „**tabu**“ erklärt, und dürfen vorerst nicht erneut durchsucht werden.

Umsetzung im Algorithmus

Ein wesentlicher Erfolgsfaktor für den erfolgreichen Einsatz von Tabu Search zur Lösung von Tourenplanungsproblemen ist auch hier die Einbettung problemspezifischen know-hows (Domänenwissen) in den Tabu-Search-Algorithmus. Der in verschiedenen Veröffentlichungen jeweils erzielte Erfolg schwankt bereits in Hinblick auf deterministische Tourenplanungsprobleme erheblich („Applying tabu search to a particular problem requires embedding in the algorithm a fair amount of problem specific knowledge. Thus, tabu search has been applied to the deterministic VRP by several authors, but in a varying degree of success“ [Gendreau et al; 1996, S. 471]).

Eine Umsetzung des Algorithmus sollte daher spezifisch auf den Anwendungsfall zugeschnitten sein. Auf eine allgemeingültige Beschreibung, die über die oben beschriebene Kernidee hinausgeht, wird daher an dieser Stelle verzichtet.

A.2.5 Bewertung der Metaheuristiken für diese Arbeit

Eine umfassende Übersicht von [Bräysy & Gendreau 2001] über den Einsatz von Metaheuristiken für VRSP-TWs beschränkt sich auf die Auswertung von Benchmarks zu den Solomoninstanzen ohne speziellen Fokus auf

- ▶ die Abhängigkeiten in Hinblick auf die Präzision
- ▶ die Möglichkeiten zur Nichtbedienung ausgewählter Kunden (rejection)

Sie kommt zu dem Ergebnis, dass Metaheuristiken zwar mittlerweile oft besser als traditionelle Verfahren sind, sie aber **mehr Rechenzeit benötigen**, komplexer zu implementieren und **schwerer zu kalibrieren** sind. „... The quality of the solutions obtained with different metaheuristic techniques is often much better compared to traditional construction heuristics and local search algorithms. At the same time, metaheuristics require more CPU-time and are more complex to implement and calibrate. ...“ [Bräysy & Gendreau 2001, S.31]. Insbesondere die Kalibrierungsnotwendigkeit stellt in der Praxis hohe Anforderungen

- ▶ sowohl an das Domänenwissen bzgl. des logistischen Problems,
- ▶ als auch an das Verfahrensverständnis bzgl. der verwendeten Metaheuristik.

In den einschlägigen Quellen findet sich nach Kenntnis des Autors jedoch **kein Hinweis** darauf, dass es bereits systematische Untersuchungen gibt, die die Einsatzfähigkeit von Metaheuristiken auf **dynamisch-stochastische** Anwendungsfälle untersuchen.

B Anhang B – Instanzen

B.1 Die Solomon Instanzen

[Solomon, 1984] hat eine Reihe von Referenzdatensätzen (Instanzen) eingeführt, an der sich bis heute die wesentlichen Verfahren zum VRSP-TW messen. Eine Übersicht über die Solomoninstanzen ist in der folgenden Tabelle zusammengestellt.

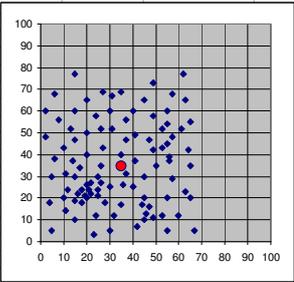
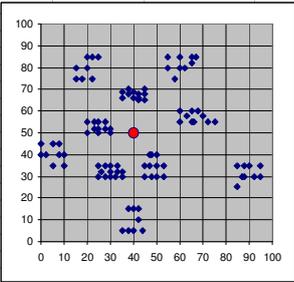
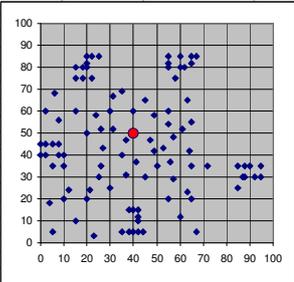
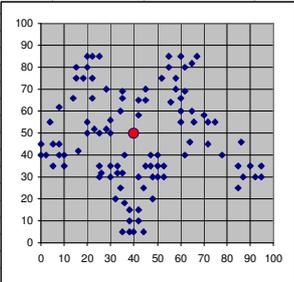
Beschreibung der Solomon Instanzen									
Beschreibung der Geografie									
R	Zufallskordinaten (Randomly Generated Coordinates)								
C1	Koordinatencluster (clustered coordinates) 10 Cluster; Intervalle entlang einer 3opt-Route innerhalb eines jeden Clusters generiert; Zeitfenstermitte = Ankunftszeit in Clusterroute								
C2	Koordinatencluster (clustered coordinates) 3 Cluster; Abgeleitet aus C1 durch Lagerverschiebung einiger Stops								
RC	Mischung zwischen R und C (semi clustered) 50% der Daten aus C (5 äußere Cluster + 7 Stops) 50% der Daten aus R								
		Zeitrestriktionen			Mengenrestriktionen			Fahrzeugrestriktionen	
Typ	#Stops	ZE/Stop	max ZE/Tour	Lieferhorizont [ZE]	Lieferhorizont [dd hh mm]	ME/Stop	max ME/Fahrzeug	max # Fahrzeuge	min # Fahrzeuge ohne Zeitfenster
R1	100	10	230	230	3h 50 min	[1:41] step 1	200	25	10
C1	100	90	1236	1236	20h 36 min	[10:50] step 10	200	25	10
RC1	100	10	240	240	4h 00 min	[3:40] step 1	200	25	10
R2	100	10	1000	1000	16h 40 min	[1:41] step 1	1000	25	2
C2	100	90	3390	3390	2d 8h 30 min	[10:50] step 10	700	25	3
RC2	100	10	960	960	16h 00 min	[3:40] step 1	1000	25	2
Einheiten & Geschwindigkeiten									
ME	Mengeneinheit								
ZE	Zeiteinheit; Annahme: 1 ZE = 1 min								
WE	Wegeinheit; Annahme: 1 WE = 1km								
v.	Definitionsgemäß sind in den Solomon Instanzen Zeiteinheiten = Wegeinheiten => Durchschnittsgeschwindigkeit = 60km/h								
<div style="display: flex; flex-wrap: wrap;"> <div style="width: 50%;"> <p>R1&R2</p>  </div> <div style="width: 50%;"> <p>C1</p>  </div> <div style="width: 50%;"> <p>RC1 & RC2</p>  </div> <div style="width: 50%;"> <p>C2</p>  </div> </div>									

Abbildung 9: Die Solomon Instanzen – geografische Lage

Das besondere an den Solomon-Instanzen aus Sicht dieser Arbeit ist, dass sie nicht nur extensiv in der Literatur als Benchmark-Instanzen für neue Verfahren verwendet werden, sondern sich einige der Instanzen durch ihre besondere zeitliche Struktur zur Untersuchung der Auswirkung der Lieferzeitfensterbreite auf die Transportkosten besonders gut eignen.

Es sind dies die folgenden Instanzen, die bei 100% der Lieferpunkte Zeitfenster⁹ aufweisen:

Instanz	Zeitfensterbreite Vorgabe [ZE]	Zeitfensterbreite MW [ZE]	Zeitfensterbreite SD	Stan dzeit	Implizite Tourstruktur	Kapazitätsbeschränkt
R101	10	10,00	0,00	10	nicht sortiert	ja
R105	30	30,00	0,00	10	nicht sortiert	ja
R109	60	58,89	8,93	10	nicht sortiert	ja
R112	120	117,64	17,45	10	nicht sortiert	ja
C101	60	60,76	10,53	90	nach Clustern sortiert	ja
C105	120	121,61	20,98	90	nach Clustern sortiert	ja
C107	180	180,00	0,00	90	nach Clustern sortiert	ja
C108	240	234,28	41,96	90	nach Clustern sortiert	ja
C109	360	360,00	0,00	90	nach Clustern sortiert	ja
RC101	30	30,00	0,00	10	nicht sortiert	ja
RC106	60	60,00	0,00	10	nicht sortiert	ja
RC108	120	112,33	30,80	10	nicht sortiert	ja
R201	120	115,96	35,78	10	nicht sortiert	nein
R205	240	240,00	0,00	10	nicht sortiert	nein
R211	480	471,94	71,67	10	nicht sortiert	nein
C201	160	160,00	0,00	90	nach Clustern sortiert	nein
C205	320	320,00	0,00	90	nach Clustern sortiert	nein
C206	480	486,64	83,99	90	nach Clustern sortiert	nein
C208	640	640,00	0,00	90	nach Clustern sortiert	nein
RC201	120	120,00	0,00	10	nicht sortiert	nein
RC206	240	240,00	0,00	10	nicht sortiert	nein
RC208	480	471,93	71,67	10	nicht sortiert	nein

Abbildung 10: Die Solomoninstanzen - Kennwerte

Da im Hauptteil dieser Arbeit mehrfach auf die Relevanz von „impliziten Tourstrukturen“ eingegangen wurde, werden auch die Solomon-Instanzen dahingehend kurz untersucht. Dazu wird die zeitliche Lage grafisch durch die zwei Attribute Farbintensität und Größe dargestellt (je früher, desto kleiner und heller, je später, desto größer und dunkler).

Durch diese einfache grafische Darstellung ist klar zu erkennen, dass in den Instanzen C1* (10 Touren) und C2* (3 Touren) eine implizite Tourstruktur vorliegt, in den Instanzen R1* und RC1* jedoch nicht.

⁹ Ein kleiner Schönheitsfehler dabei ist, dass nicht alle Zeitfenster einer Instanz identisch groß sein müssen, sondern nur einen definierten Mittelwert (MW) aufweisen müssen. Bei der Generierung der Instanzen ist Solomon jedoch von einem „Zielwert“ der Zeitfensterbreite (Vorgabe) ausgegangen. Für jede Instanz ist daher die entsprechende Standardabweichung (SD) der Streuung der Zeitfensterbreite angegeben.

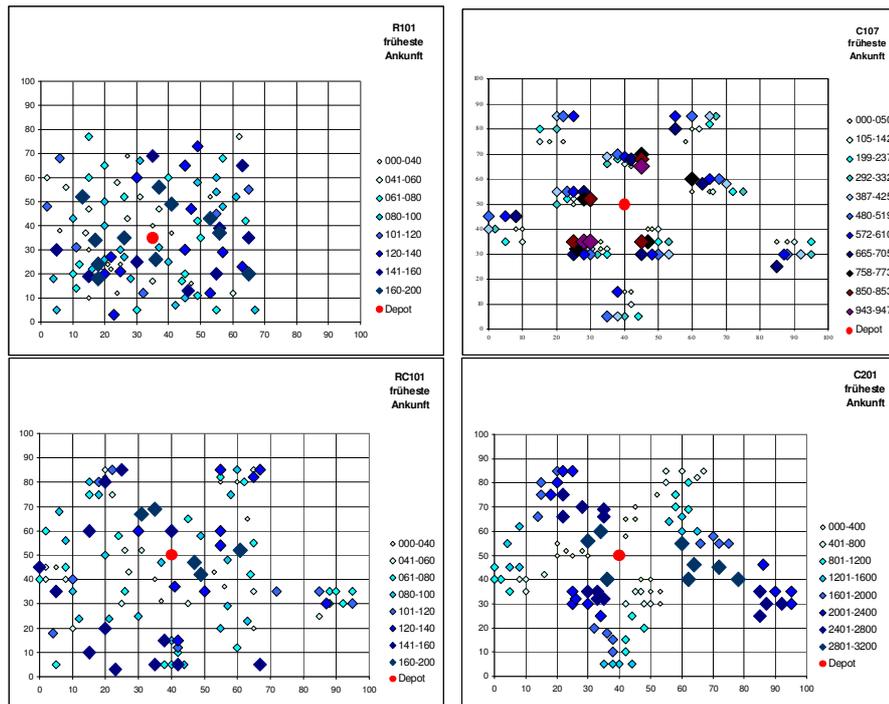


Abbildung 11: Die Solomoninstanzen - zeitlich geografische Lage; Farbcode beschreibt früheste Ankunftszeit in Zeiteinheiten

B.2 Die Instanzengruppe ZUE500*

Die Instanzengruppe ZUE500* wurde für diese Arbeit als neue Benchmarkinstanz generiert. Sie wird veröffentlicht unter www.ifl.uni-karlsruhe.de. Im Folgenden sind einige Details zu dieser Instanzengruppe kurz zusammengestellt.

B.2.1 Rahmenbedingungen

- ▶ Koordinatensystem: Geodetisch (Superconform)
- ▶ Depotlage:
XCoord= 8,52686° E (5089493)
YCoord= 47,37810° N (3167018)
- ▶ Depotöffnungszeiten: 08:00 - 22:00
- ▶ Max. Tourdauer: 08:00h
- ▶ Max. Einsatzdauer eines Fahrzeugs: 16:00h
- ▶ Max. Einsätze: 10
- ▶ Max. Stops/Tour: 100
- ▶ Kapazitätsbegrenzung: keine

- ▶ Entfernungsrechnung: Luftlinie¹⁰
- ▶ Fahrzeitberechnung: 20 km/h¹¹ (1km = 3 min.)
- ▶ Standzeit/Stop: 5 min.
- ▶ Min. Entfernung zwischen 2 Stops:
Min. Weg: 333 m
Min. Fahrzeit: 60 sec
- ▶ Keine Wendezeit zwischen 2 Touren
- ▶ Keine Pausen

B.2.2 Ort-Zeitliche Stopplage (Übersicht)

Die geografische Verteilung der Stops orientiert sich an der Siedlungsstruktur im Großraum Zürich. Die Stopstandorte wurden zufällig so gewählt, dass sie in den Siedlungsgebieten liegen (Zufallsauswahl aus den Straßenreferenzkoordinaten¹² im Großraum Zürich; Postleitzahlgebiet 80**).

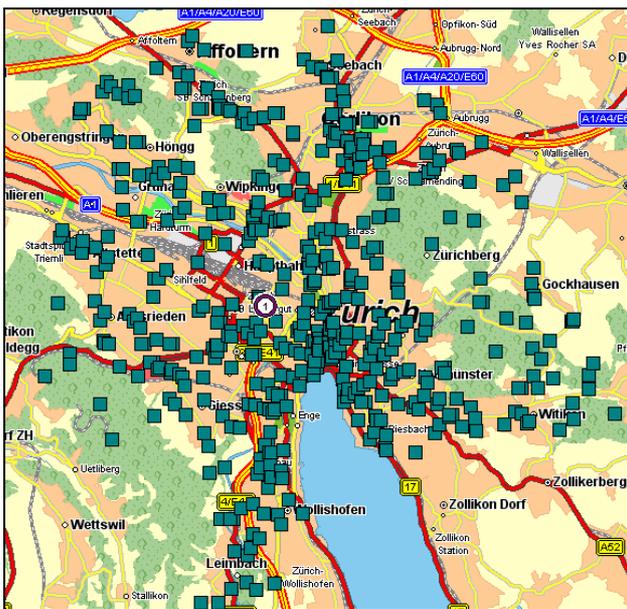


Abbildung 12: ZUE500* - geografische Verteilung

¹⁰ Annahme in Anlehnung an die Solomon-Instanzen. Diese Annahme ist zwar nicht sehr realitätsnah, da die km auf einem Straßennetz länger sind als Luftlinienentfernungen. Im Gegensatz zu netzbasierten Entfernungen sind Luftlinienentfernungen aber einfach zu berechnen. Außerdem können so unterschiedliche Netztopologien als ergebnisbeeinflussende Größen ausgeschlossen werden.

¹¹ Diese Annahme ist sehr grob, aber für Großstadteinzugsgebiete im Mittel noch einigermaßen realistisch, wenn man bedenkt, dass bereits die Entfernungen durch Luftlinien approximiert werden. Auch hier soll die Instanz möglichst einfach ausgewertet werden können, so dass diese problemvereinfachende Annahme sinnvoll ist.

¹² Ziel dieses Vorgehens ist es, eine Lieferpunktstruktur zu generieren, die dem einer europäischen Stadt nahe kommt. Die Zufallsauswahl der Straßenreferenzpunkte wurde nicht durch soziodemografische Einwohnerzahlen gewichtet, und hat dazu geführt, dass auch einige „Waldwege“ mit aufgenommen wurden. Da Entfernungen für die Instanz per Luftlinie berechnet werden, ist dieser Effekt jedoch unerheblich.

Die zeitliche Verteilung der Stopflieferzeitpunkte liegt zwischen 08:00 und 22:00 Uhr. Die Zeitpunktsgenerierung basiert auf einer Gleichverteilungsannahme.

Es wurden 3 Instanzen zur Klasse ZUE500*¹³ generiert, die sich nur in der zeitlichen Lage des Lieferzeitpunkts, nicht aber in der geografischen Verteilung unterscheiden. Das Ergebnis der Zeitpunktdistribution ist in der folgenden Übersicht dargestellt:

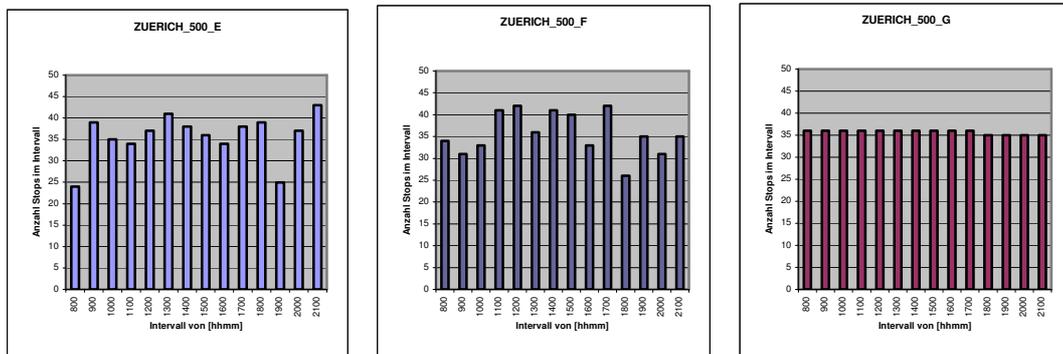
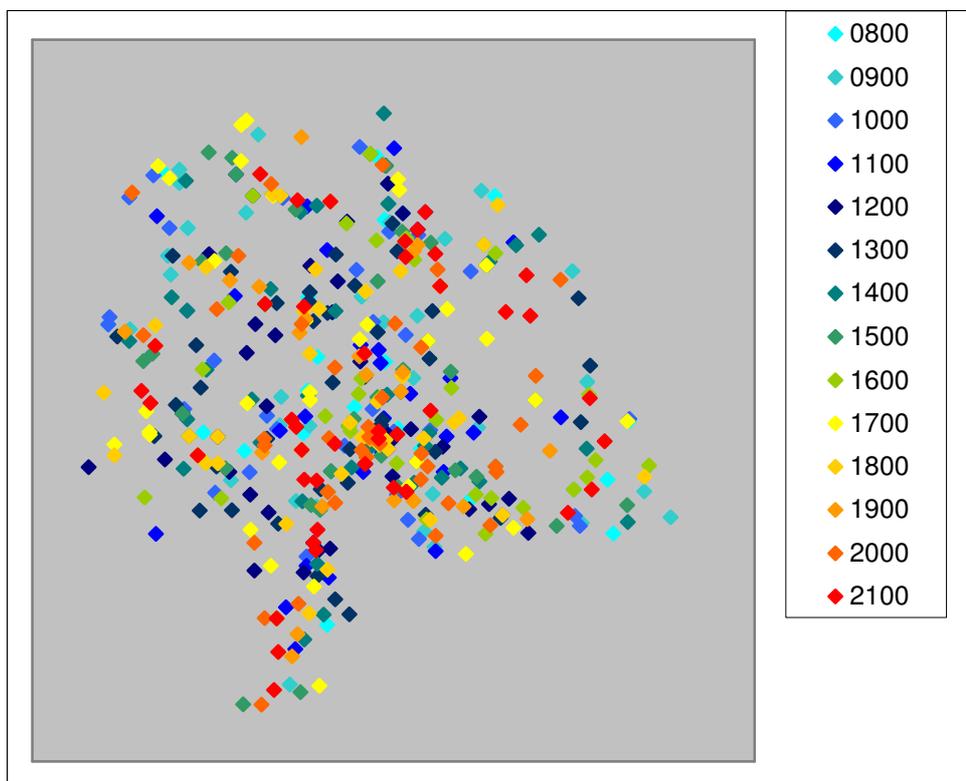


Abbildung 13: ZUE500 - zeitliche Verteilung

Das Ergebnis der Zeitpunktdistribution für die Instanzgruppe ZUE500* ist in den folgenden Übersichten dargestellt:



¹³ ZUE500_E, ZUE500_F und ZUE500_G

Abbildung 14: ZUE500E – zeitlich geografische Lage

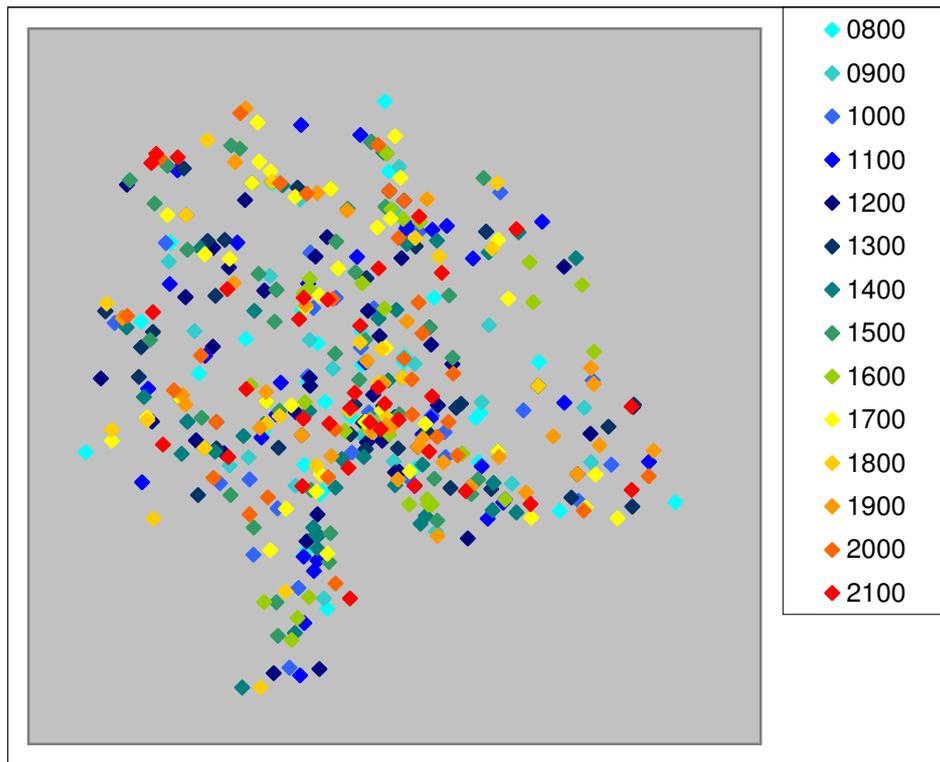


Abbildung 15: ZUE500F – zeitlich geografische Lage

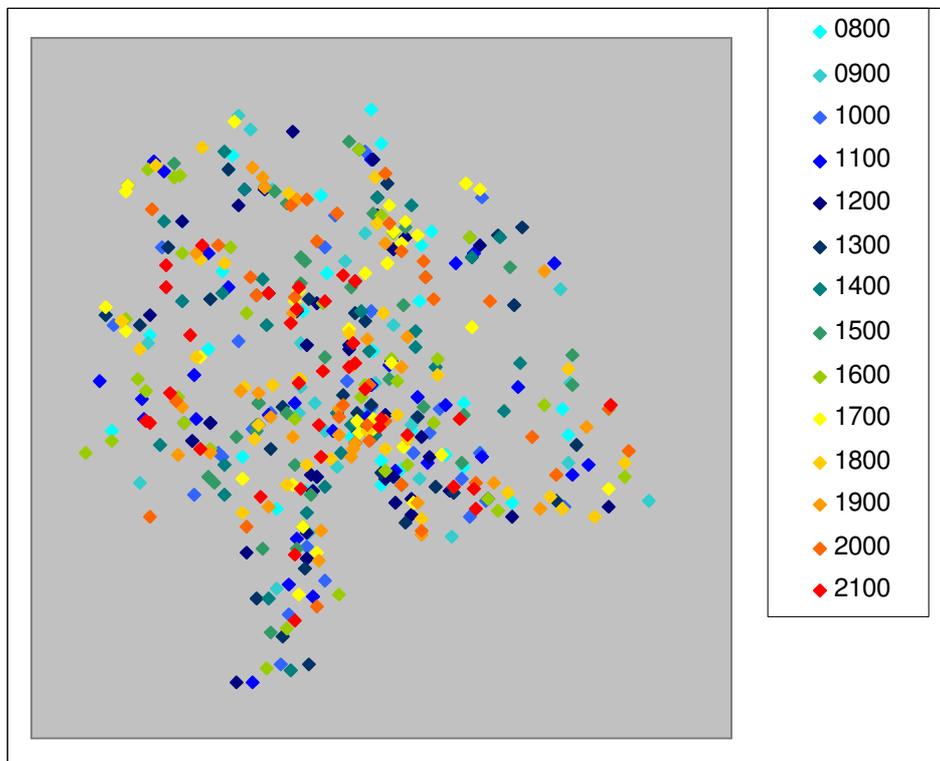


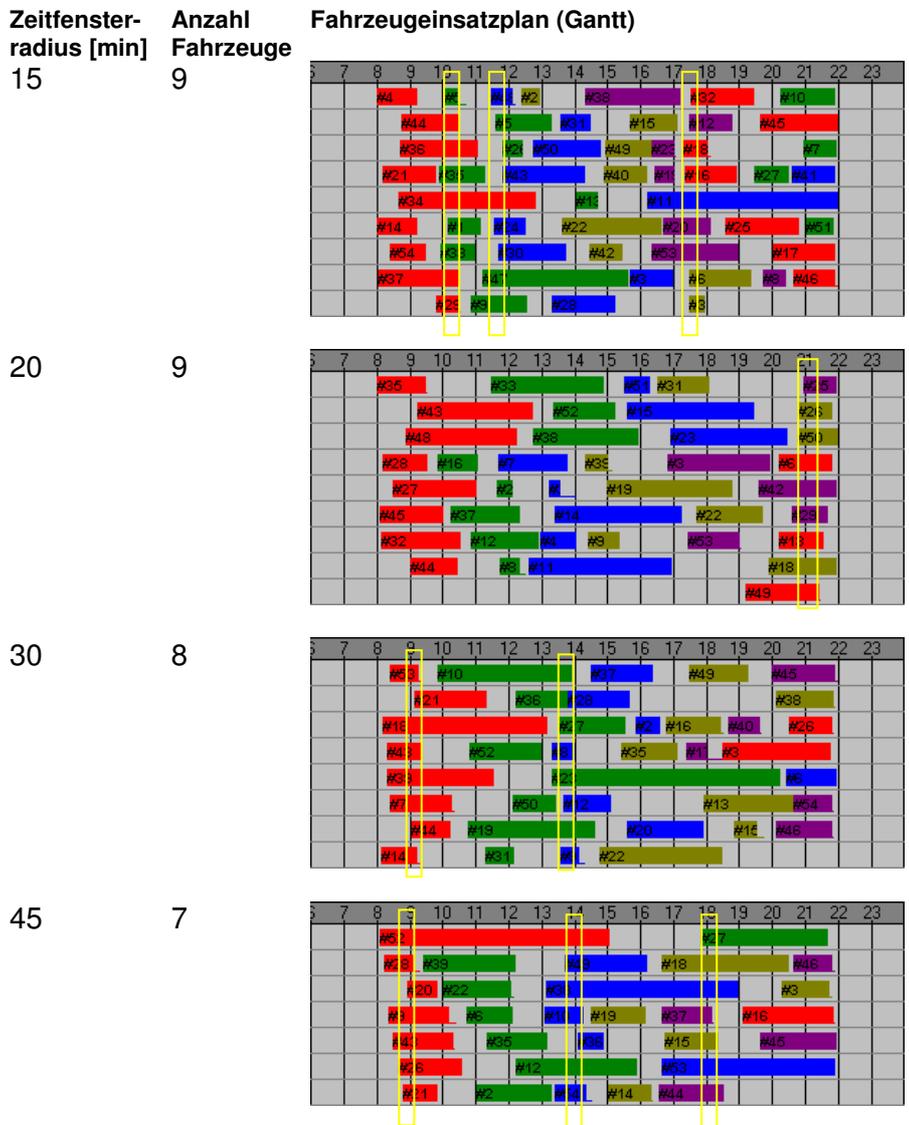
Abbildung 16: ZUE500G – zeitlich geografische Lage

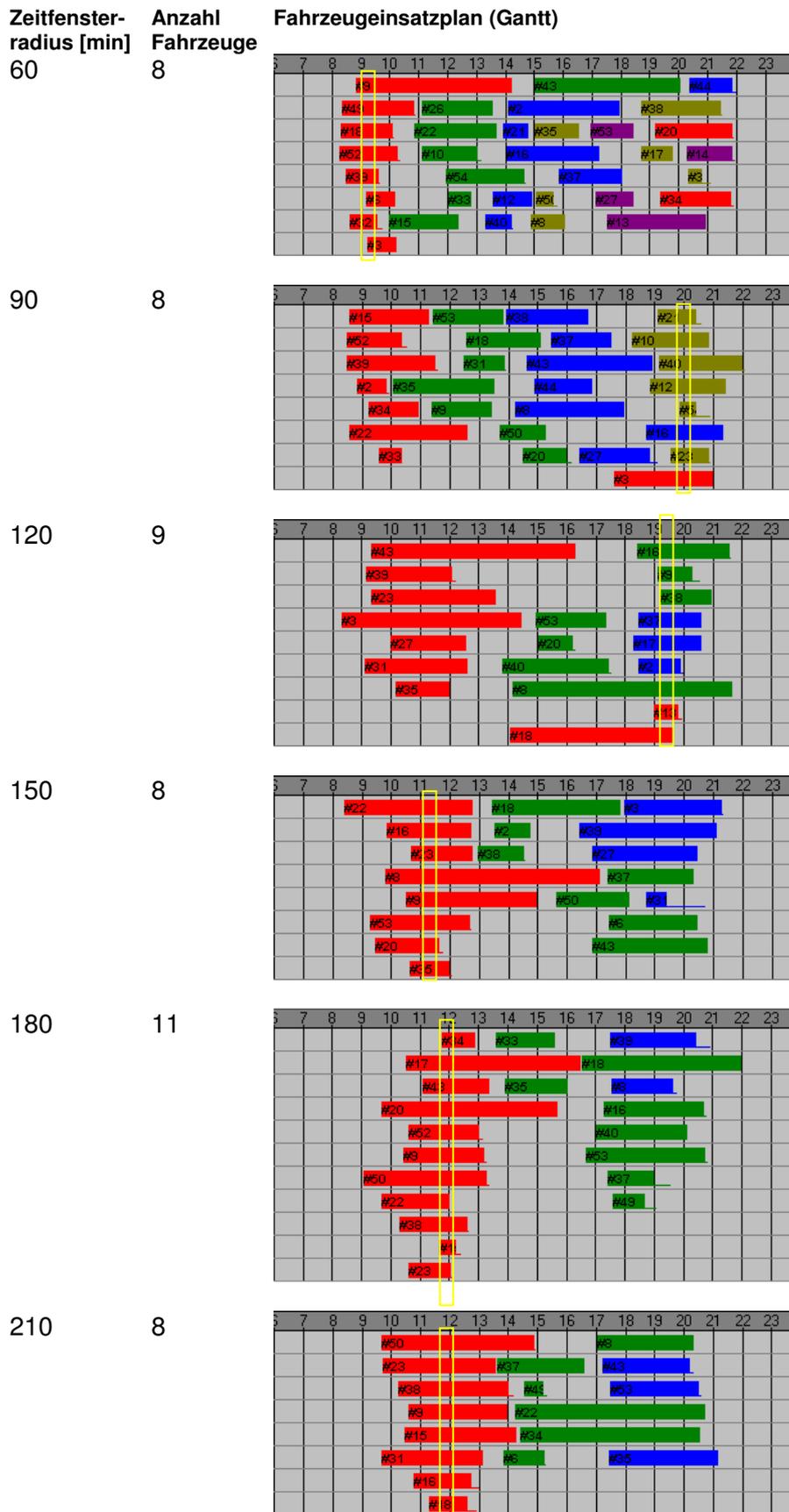
C Anhang C - Untersuchungen zur Instanz ZUE500E

C.1 Entwicklung der Fahrzeuganzahl

C.1.1 HSav0 - Ohne Fuhrparkbegrenzung

- ▶ Instanz: ZUE500E
- ▶ Verfahren: HSav0: Savings & nachgeschaltete Einsatzplanung
- ▶ Parameter: alpha: 1,3; beta=0,2; max Wartezeit=5min

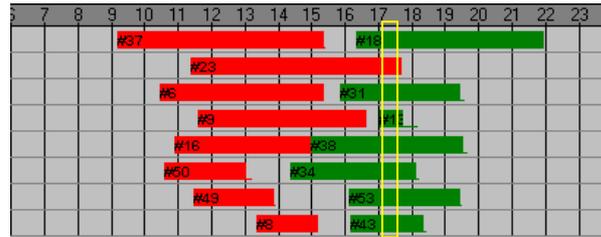




Zeitfenster-
radius [min] Anzahl
Fahrzeuge Fahrzeugeinsatzplan (Gantt)

240

8



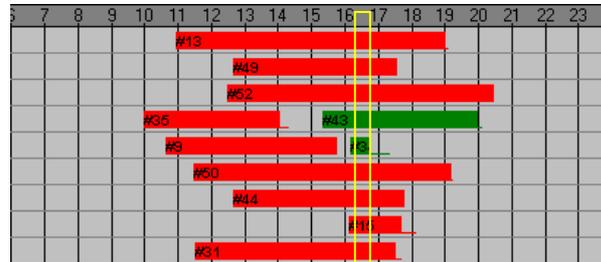
270

8



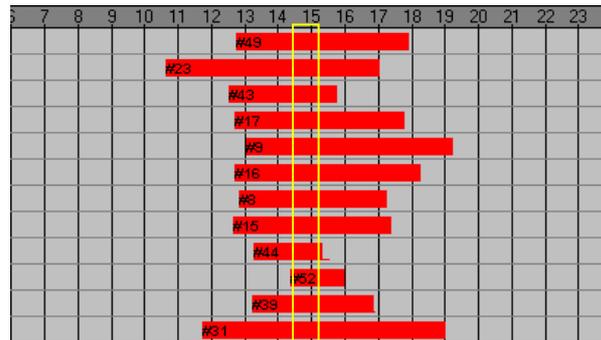
300

9



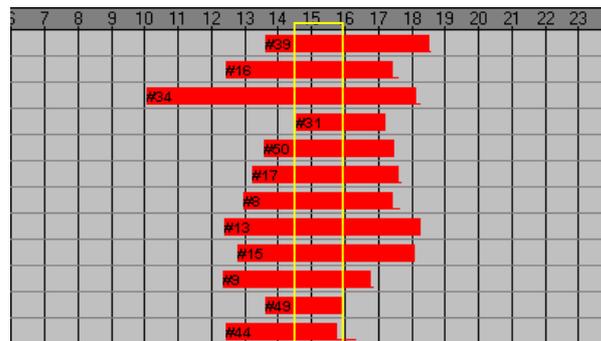
330

12

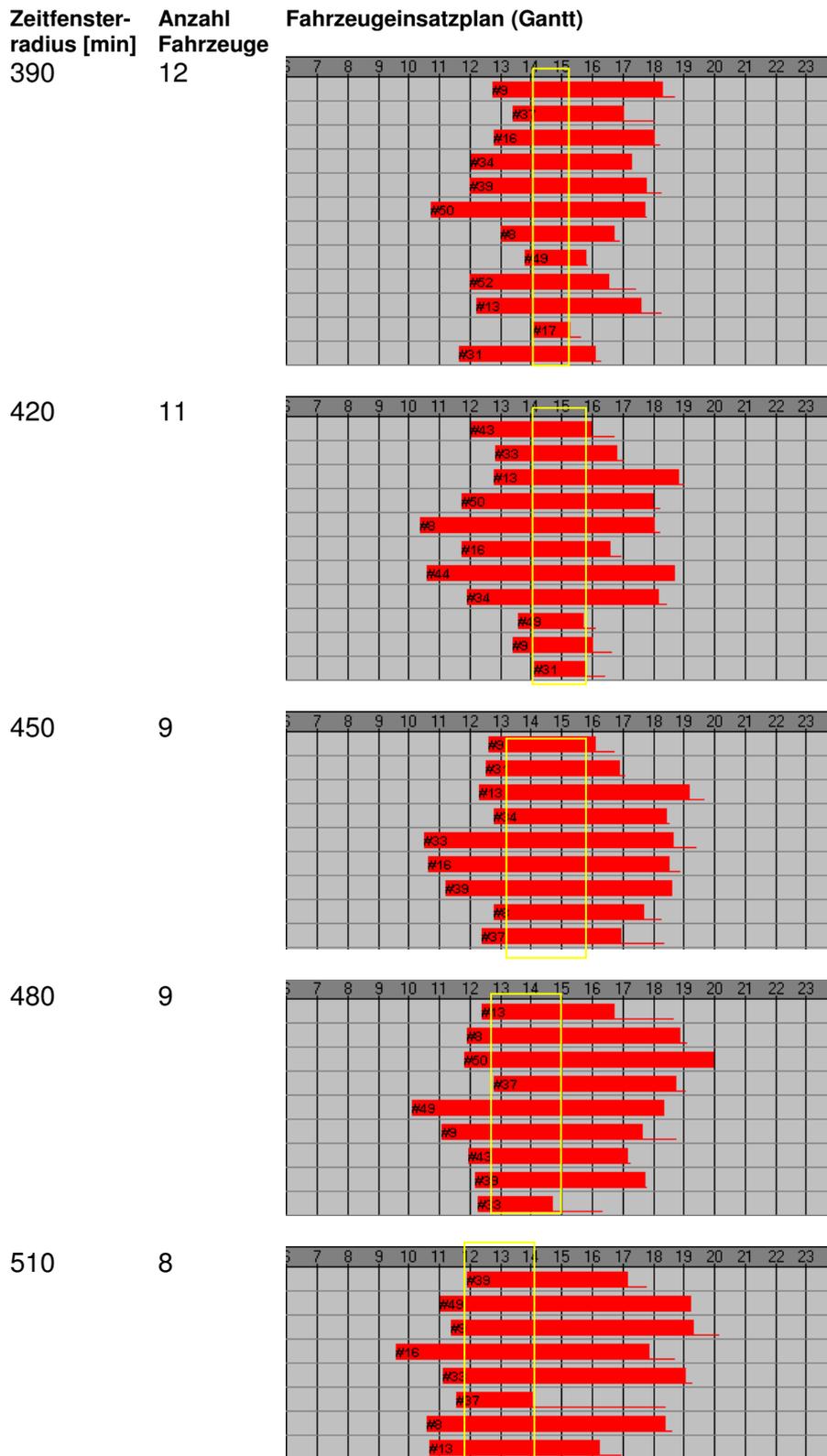


360

12



Anhang C - Untersuchungen zur Instanz ZUE500E



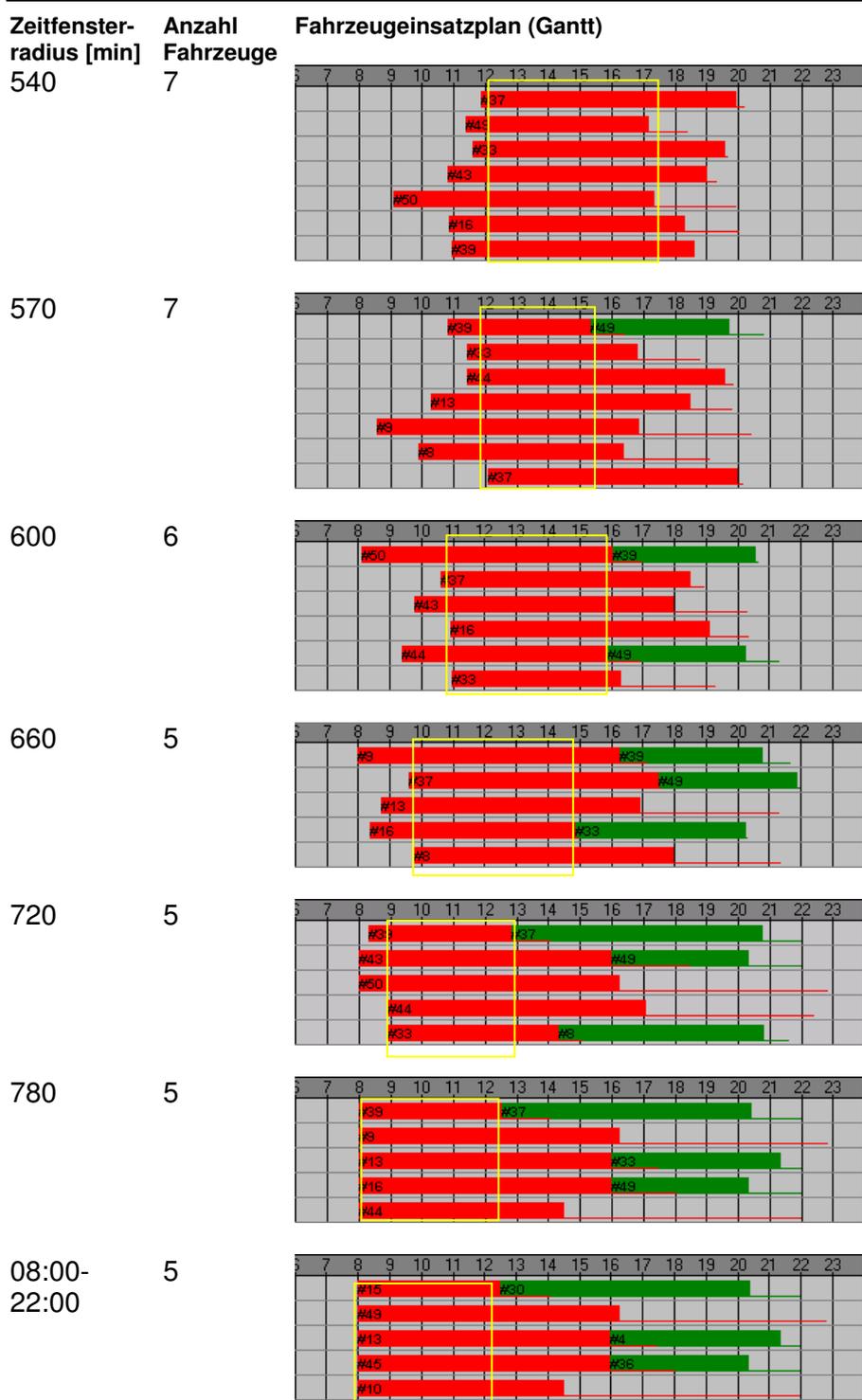
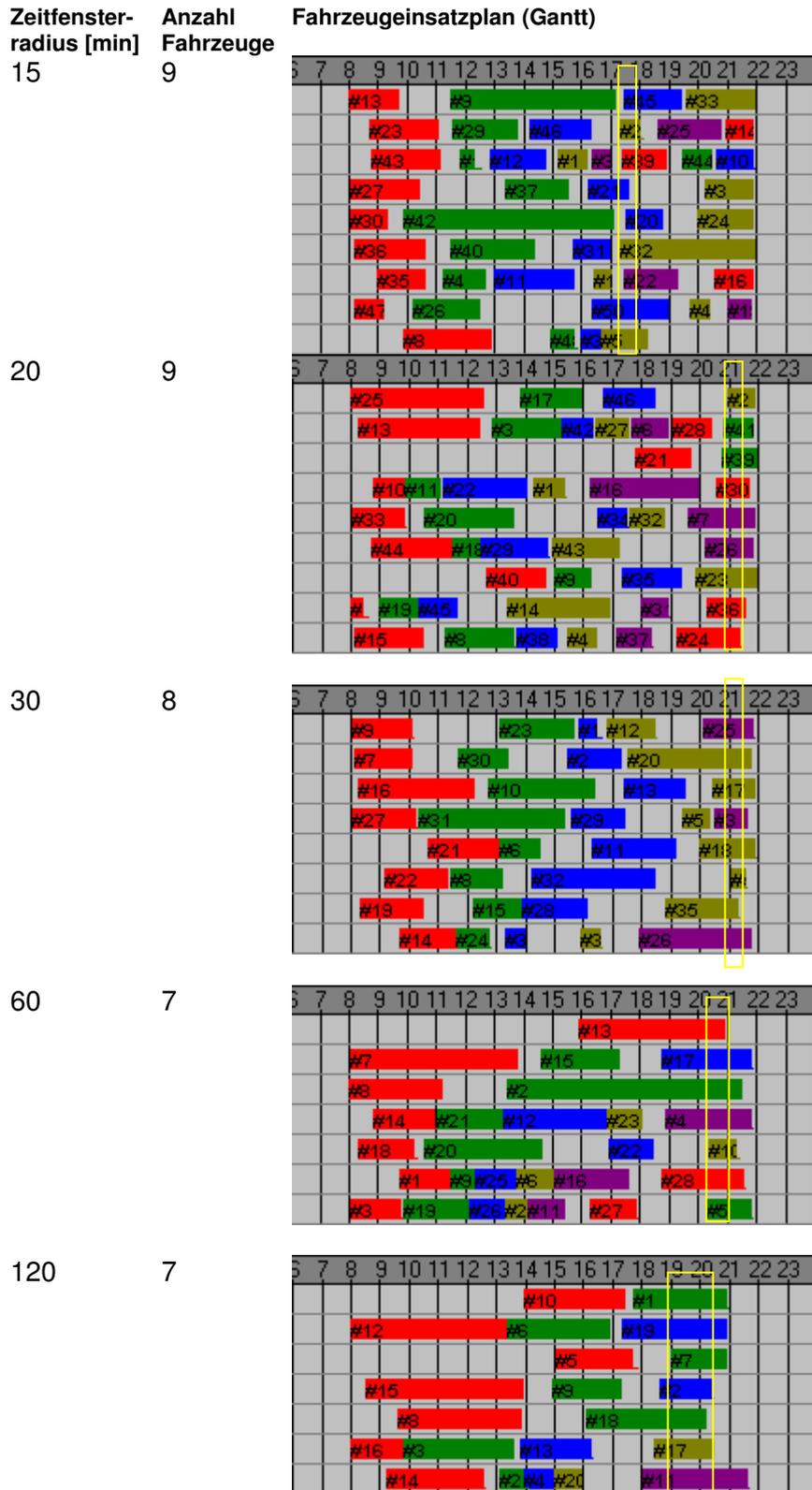


Abbildung 17: Entwicklung der Fahrzeuganzahl

C.1.2 HSav1 - Mit Fuhrparkbegrenzungsaufweitung

- ▶ Instanz: ZUE500E
- ▶ Verfahren: HSav1 mit sukzessive aufgeweiteter Fahrzeugverfügbarkeit

► Parameter: alpha: 1,3; beta=0,2; max Wartezeit=5min



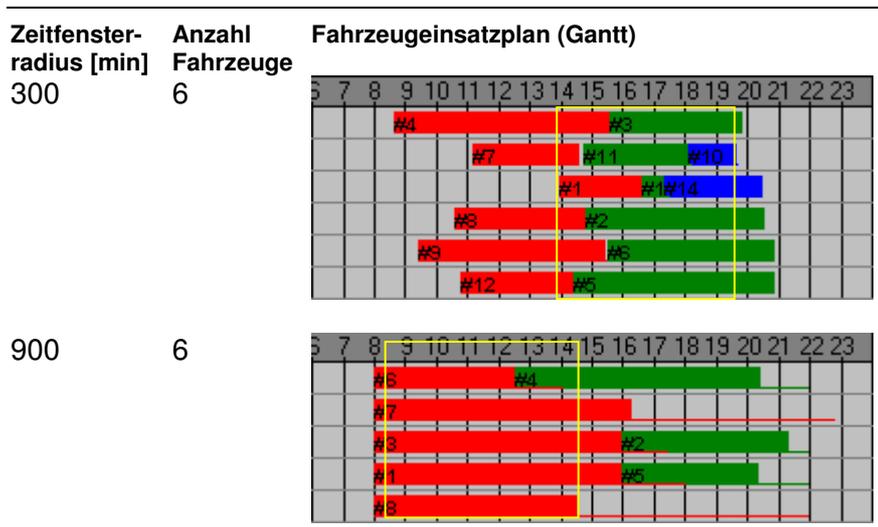
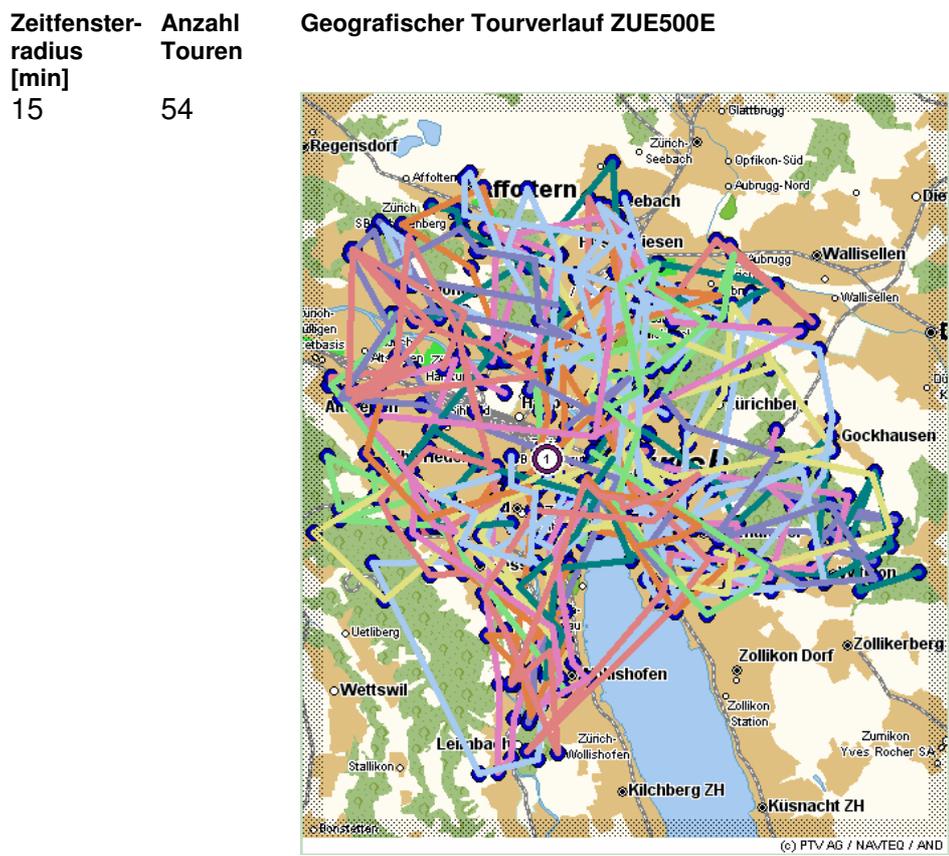


Abbildung 18: Entwicklung der Fahrzeuganzahl für HSav1

C.2 Entwicklung der Tourstruktur

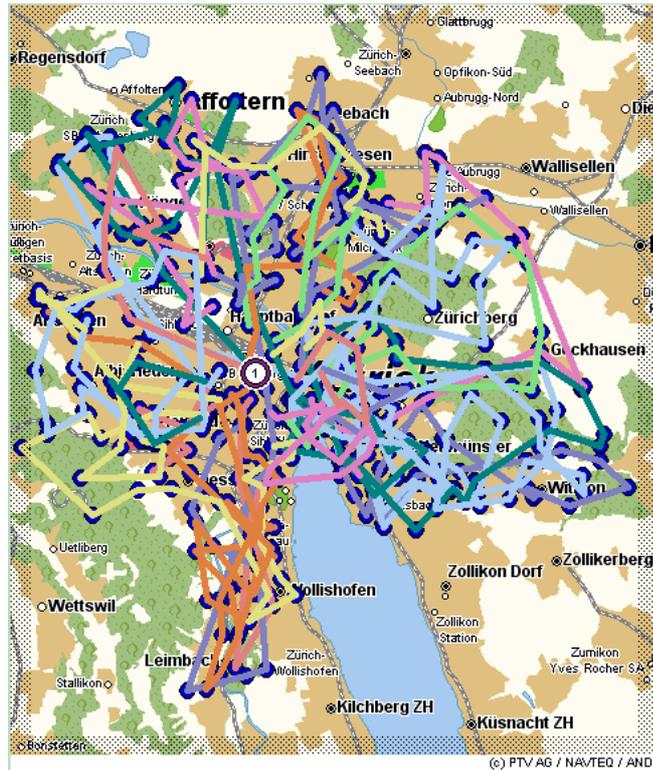
- ▶ Instanz: ZUE500E
- ▶ Verfahren: HSav0
- ▶ Parameter: alpha: 1,3; beta=0,2; max. Wartezeit=5min



Zeitfenster-
radius
[min]
60

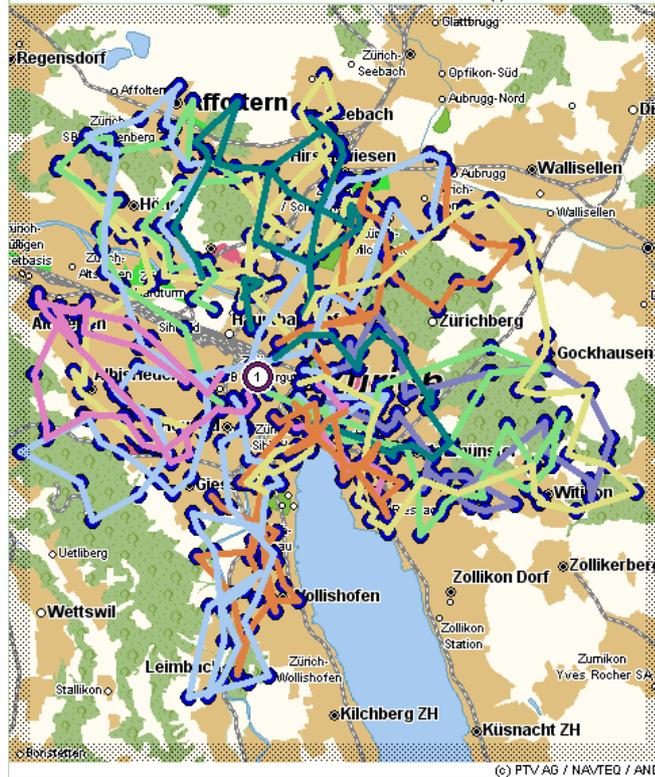
Anzahl
Touren
34

Geografischer Tourverlauf ZUE500E



150

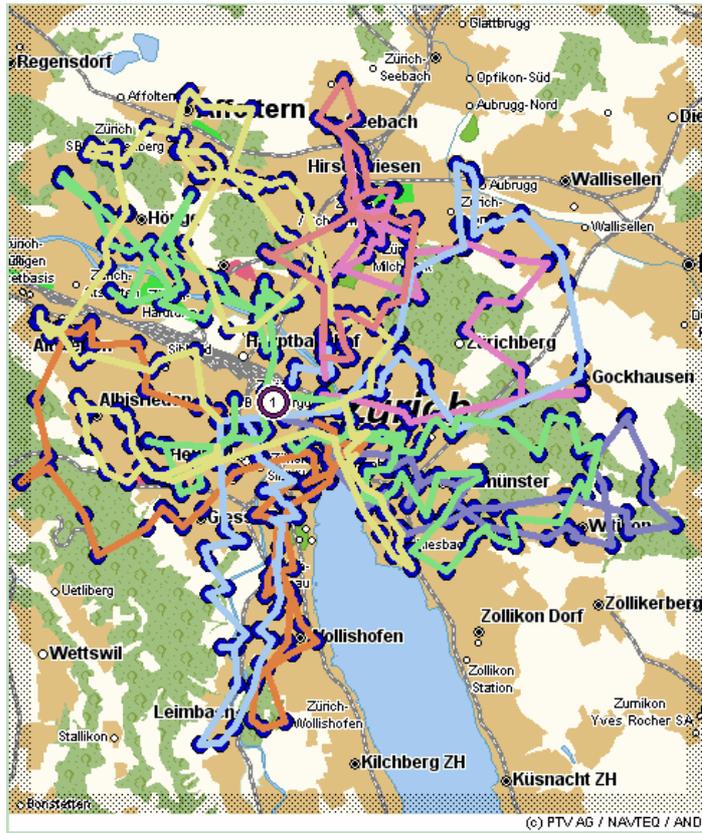
19



Zeitfenster-
radius
[min]
300

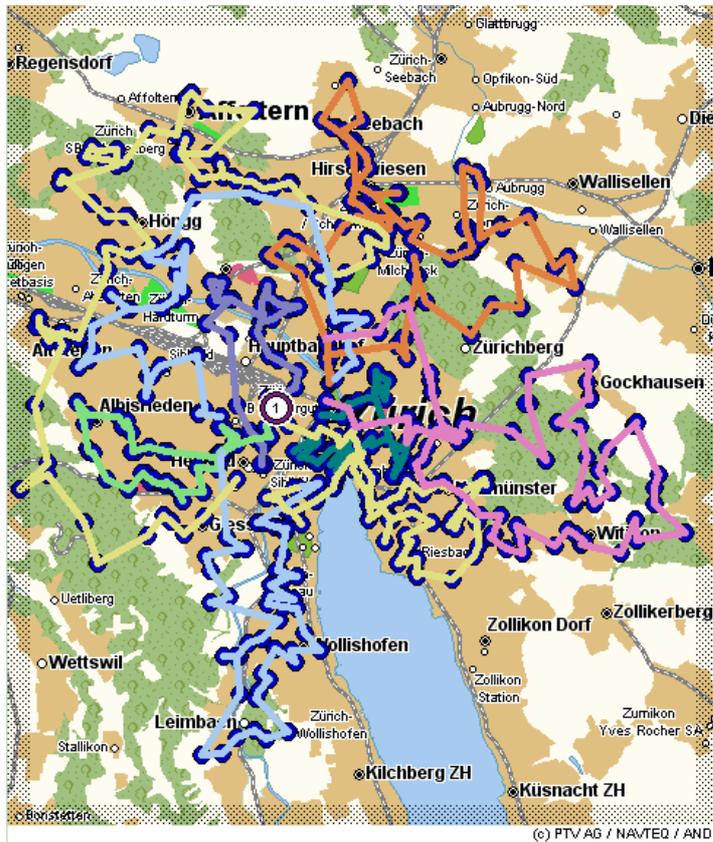
Anzahl
Touren
11

Geografischer Tourverlauf ZUE500E



450

9



Zeitfenster- radius [min]	Anzahl Touren	Geografischer Tourverlauf ZUE500E
---------------------------------	------------------	-----------------------------------

08:00- 22:00	8	
-----------------	---	--

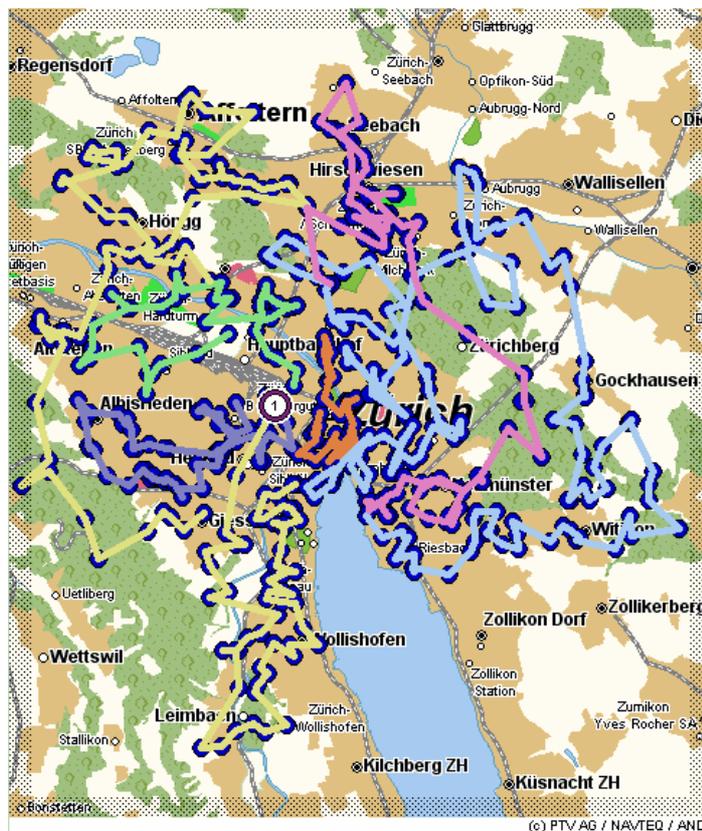


Abbildung 19: Entwicklung der Tourstruktur

D Anhang D – Beispiel km-Regression

D.1 Km-Regressionsversuche für ZUE500e

Die folgenden Funktionstypen wurden mit Hilfe nichtlinearer Regression darauf untersucht, wie gut sie die Abhängigkeit der km von der Zeitfensterbreite (Radius) am Beispiel der Instanz ZUE500E und dem Savings-Algorithmus approximieren.

Beim Fitting wurden keine Wichtungen der Messpunkte nach x- oder y-Wert gegeneinander vorgenommen, d.h. jeder Messpunkt geht mit dem gleichen Gewicht in die Fitting-Zielfunktion ein.

D.1.1 2-Parametrische Funktionen

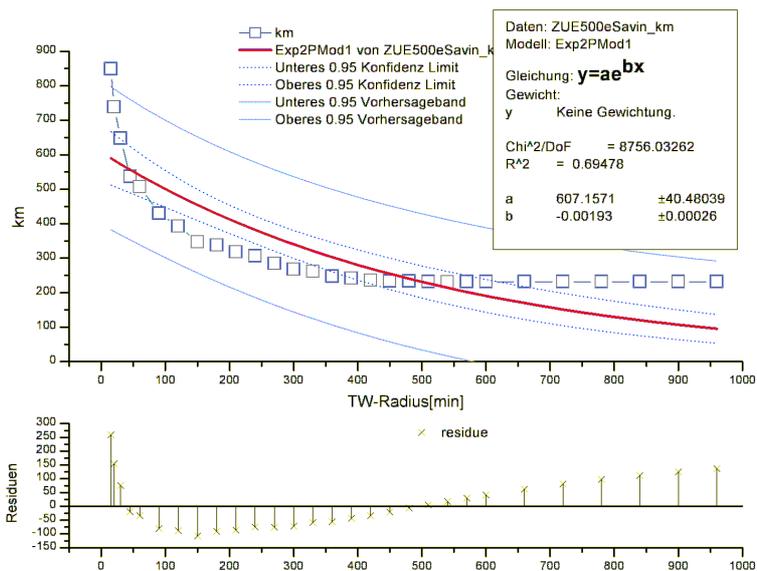


Abbildung 20: Exp2PMod1

Bei der 2-parametrischen Exponentialfunktion Exp2PMod1 fehlt das konstante Glied. Daher ist ein gutes Fitting über den gesamten Wertebereich [15;960] mit dieser Funktion nicht zu erreichen.

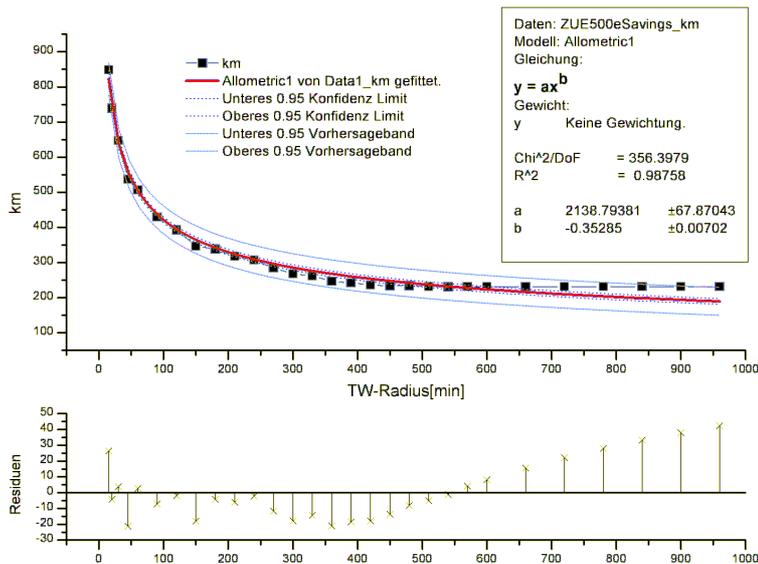


Abbildung 21: Allometric1

Mit einem $R^2=0,98758$ ist das Fitting für die Potenzfunktion „Allometric1“ bereits mit 2 Parametern recht gut gelungen.

Bei großen Zeitfensterbreiten scheint es aus den Meßwerten heraus eine wagrechte Asymptote zu geben. Dieses Phänomen wird durch die Funktion „Allometric1“ nicht exakt abgebildet.

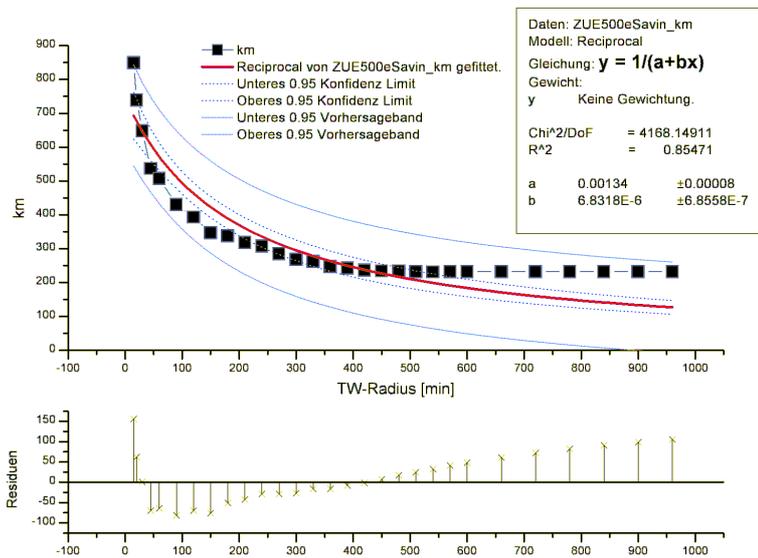


Abbildung 22: Reciprocal

Das Fitting-Ergebnis dieses Funktionstyps ist nicht befriedigend, da im linken und im rechten Wertebereich die Werte signifikant „unterschätzt“ werden, während im mittleren Wertebereich die Werte signifikant „überschätzt“ werden.

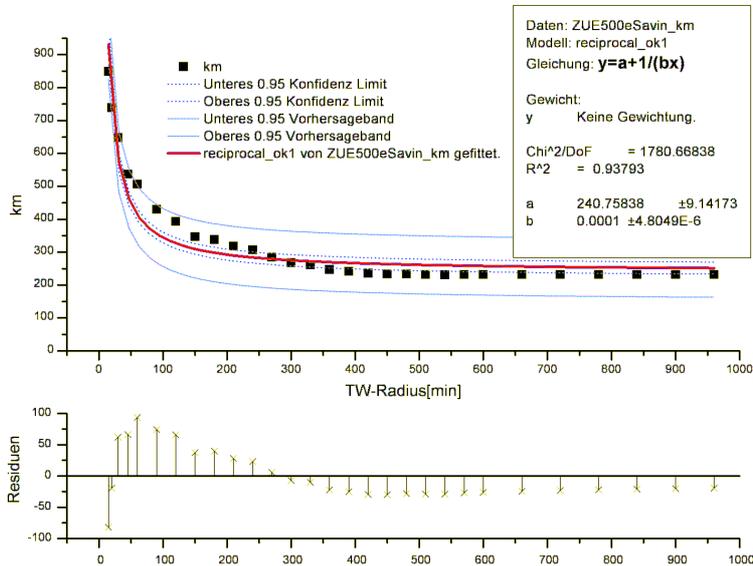


Abbildung 23: Reciprocal_ok1

Der Versuch, durch ein Konstantes Glied in der Funktion die wagrechte Asymptote abzubilden ist hier zwar gelungen. Nicht befriedigend ist dafür der „Knick“ der Funktion im Bereich $x=50\text{min}$, der im mittleren Wertebereich [200;400] eine zu schwache Steigung, dafür im linken Wertebereich [15;60] eine zu starke Steigung induziert.

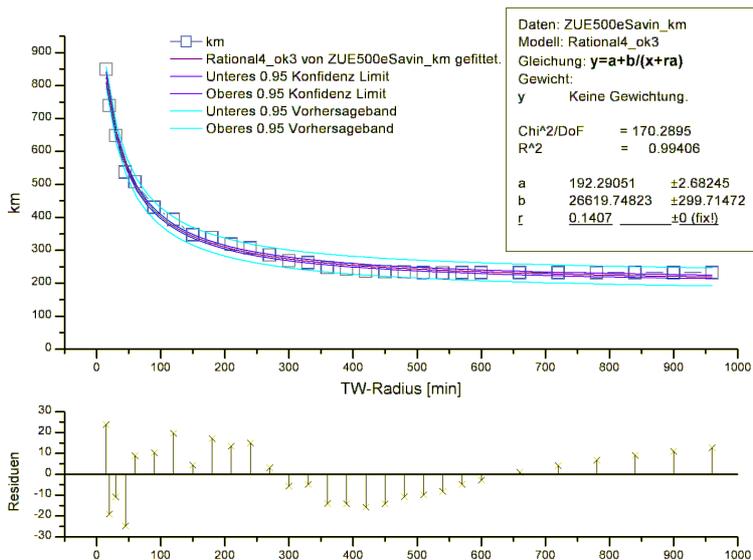


Abbildung 24: Rational4_ok3

Diese Funktion ist das Resultat einiger Try&Error-Versuche bzgl. der Konstante „r“. Die Konstante „r“ wurde letztlich so gewählt, dass R^2 für eine zweiparametrische Funktion sehr dicht an 1 herankam ($R^2=0,994$).

D.1.2 3-Parametrische Funktionen

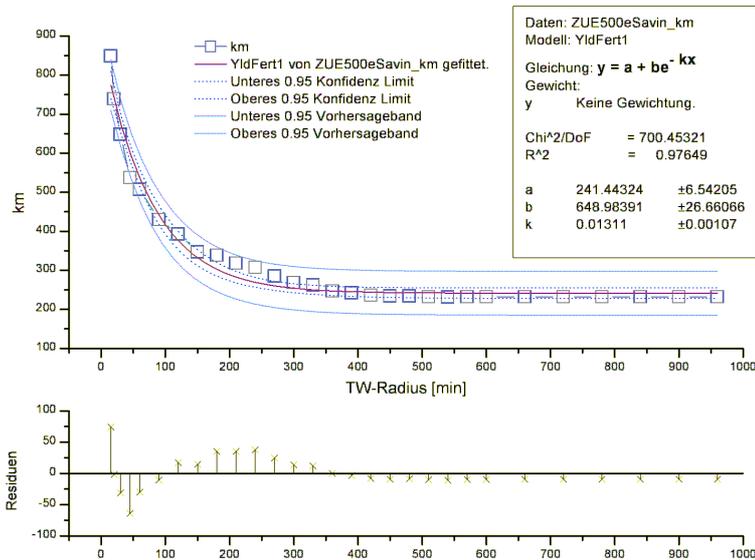


Abbildung 25: YldFert1

Diese Funktion wurde motiviert aus dem „Mißerfolg“ der Funktion Exp2PMod1. Daher wurde hier ein konstantes Glied hinzugefügt.

Diese Funktion ist aus zwei Komponenten zusammengesetzt - zum einen aus einem konstantes Glied, mit dem die wagrechte Asymptote abgebildet wird, zum anderen aus einer zweiparametrischen Exponentialfunktion, die die Steigung im linken Wertebereich gut annähert.

Das relativ großen R^2 von 0,976 wird jedoch zum Preis von 3 Parametern erkauft.

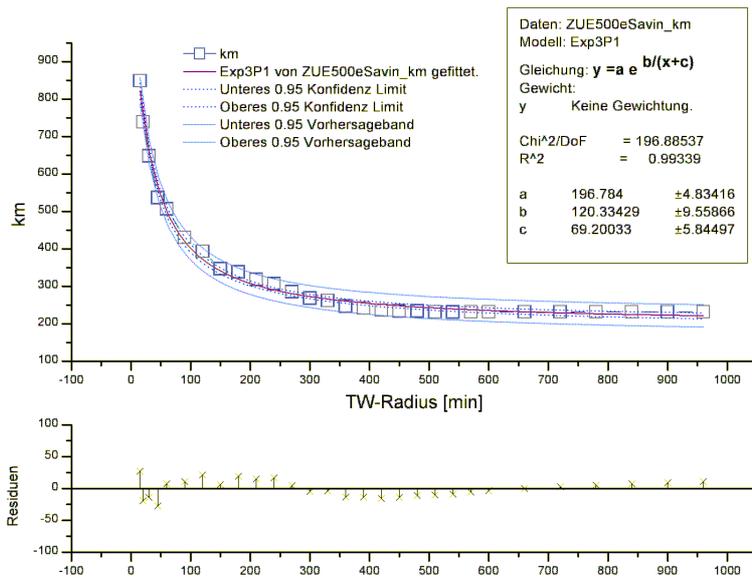


Abbildung 26: Exp3P1

Die Funktion „Exp3P1“ erzielt ein außerordentlich gutes R^2 von über 0,99.

Auch ohne konstantes Glied zeigt sie im rechten Wertebereich recht gutes asymptotisches Verhalten.

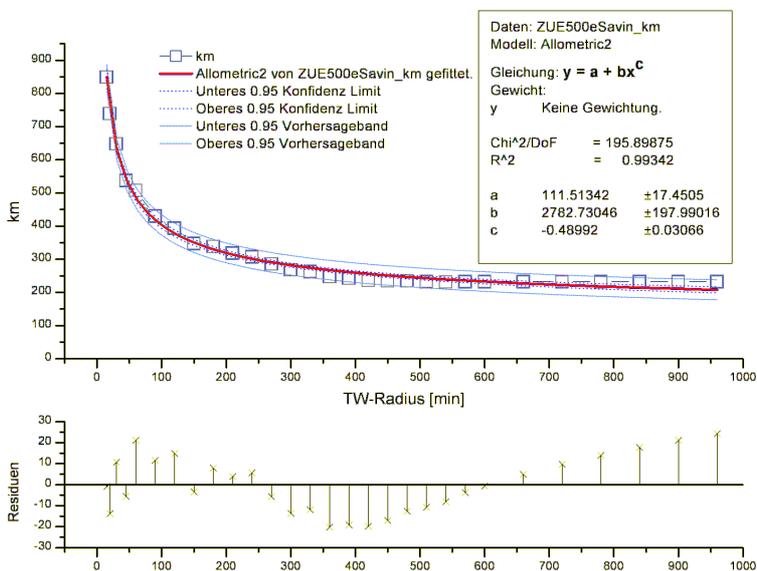


Abbildung 27: Allometric2

Die Funktion „Allometric2“ unterscheidet sich nur von der Funktion „Allometric1“ durch ein konstantes Glied, dessen Einführung R^2 auf über 0,99 hebt.

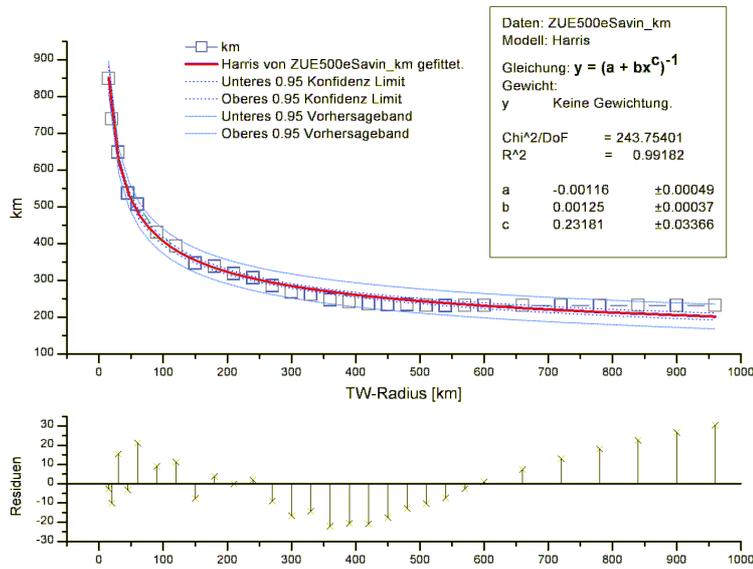


Abbildung 28: Harris

Die „Harris“-Funktion ist auch eine gute Näherung für den Verlauf der Messpunkte.

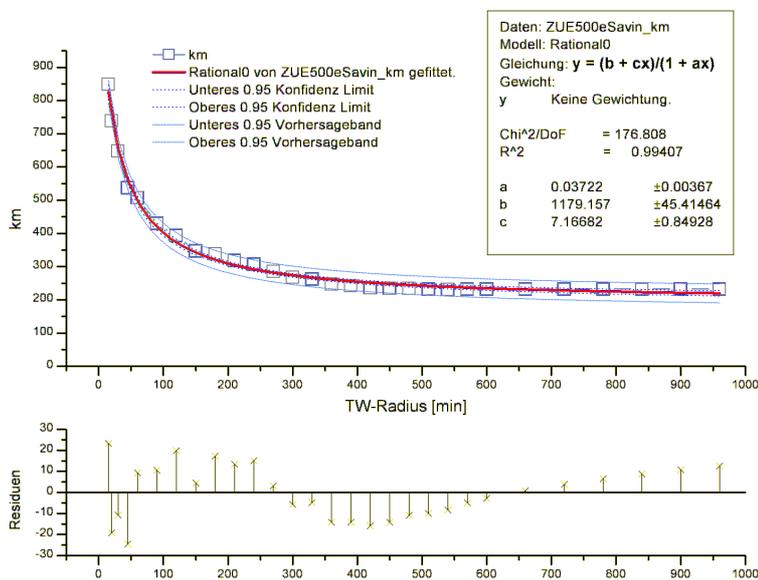


Abbildung 29: Rational0

Auch „Rational0“ weißt einen sehr hohen R²-Wert aus.

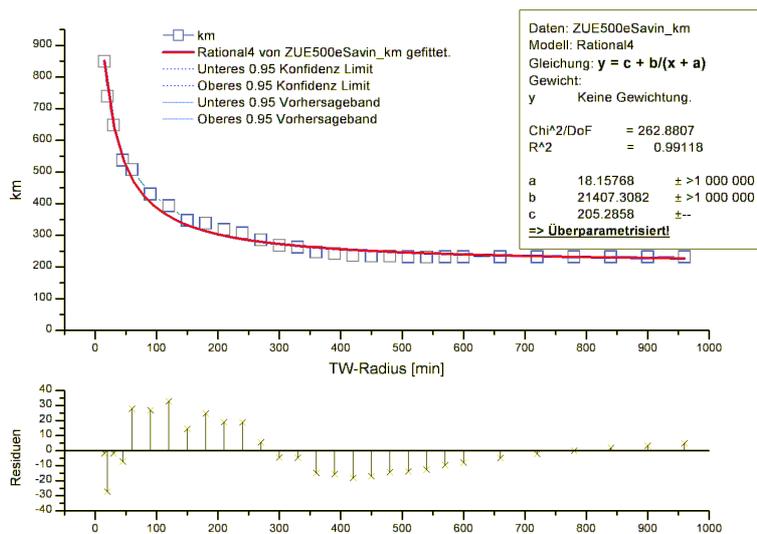


Abbildung 30: Rational4

Das NLSF (nonlinear least square fitting) aus Origin v7.5 weist hier eine starke Abhängigkeit der Parameter voneinander aus, so dass die Annahme begründet ist, die Funktion ist "überparametrisiert". Diese Überparametrisierung zeigt sich eindeutig beim Fittingversuch mit der Funktion „Rational4“. Der Fehlerbereich der einzelnen Parameter ist so groß, dass er nicht mehr ausgewiesen wird.

Aus diesem Ergebnis heraus wurde der Versuch unternommen, einen Parameter durch eine Konstante zu ersetzen (s. o.: „Rational4_ok3“):

$$y(x) = c + \frac{b}{x + a}$$

$$\Downarrow$$

$$y(x) = a + \frac{b}{x + 0.1407a}$$

Die Konstante 0.1407 wurde dabei empirisch durch „try&error“ gefunden.

D.1.3 Zwischenergebnis

In dieser Arbeit werden nach Möglichkeit die 2-parametrischen den 3-parametrischen Funktionen vorgezogen, um eine unnötige „Überparametrisierung“ zu vermeiden.

Von den 2-parametrischen Funktionen ist „Rational4_ok3“ mit dem „Makel“ des vom Himmel gefallenen Parameters behaftet. „Allometric1“, mit einem R²=0,98758 scheint aber zumindest ein nicht ungeeigneter Kandidat zu sein, um diesen Verlauf zu beschreiben.

Dass es noch „bessere“ 2-parametrische Funktionen gibt, die die Abhängigkeit der km von der Zeitfensterbreite für den Savings in der Instanz ZUE500E beschreiben, kann nicht ausgeschlossen werden.

D.2 Km-Regressionsversuche für ZUE500efg

Im folgenden wird kurz beleuchtet, ob und wie gut sich die Ergebnisse bzgl. des „HSav1“ in den Instanzen ZUE500E, ZUE500F, ZUE500G durch allometrische bzw. exponentielle Regression des Typs „YldFert1“ beschreiben lassen. Im Folgenden ist dargestellt, was die Ergebnisse der jeweiligen Regressionen sind.

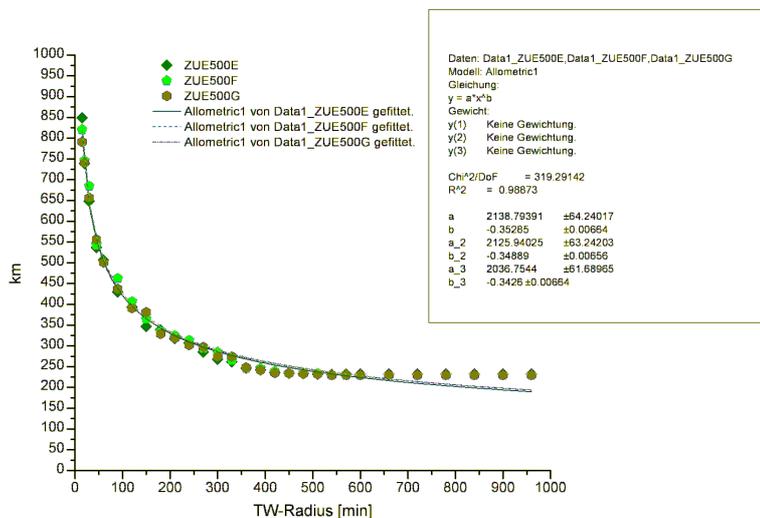


Abbildung 31: Allometrische Regression unabhängig je Instanz

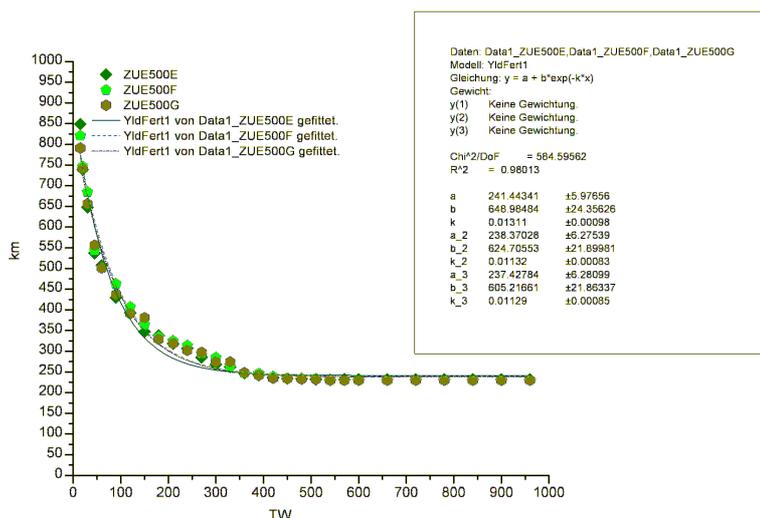
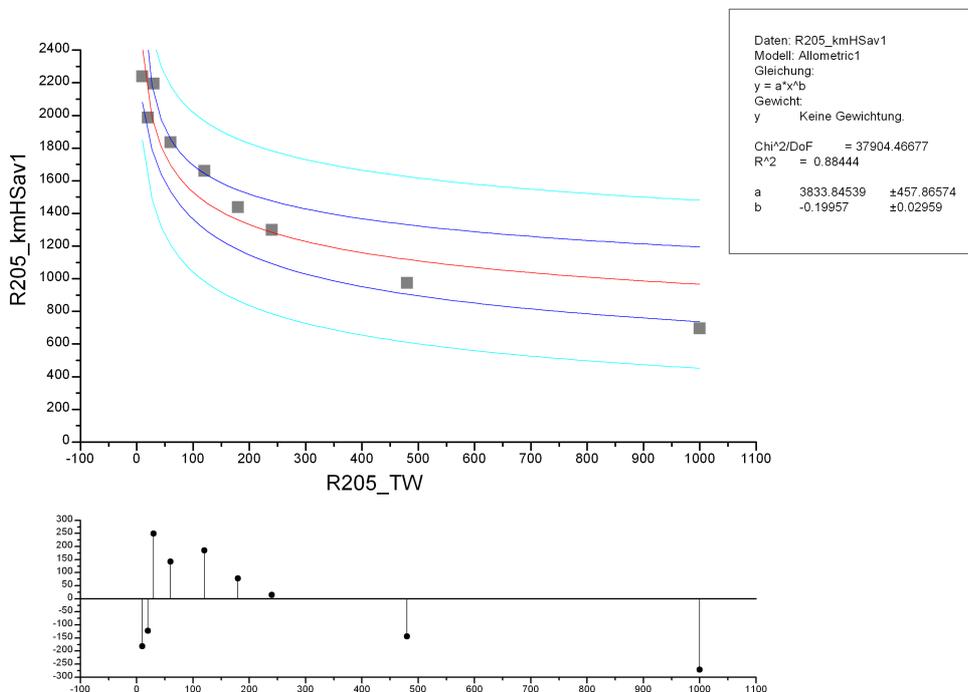
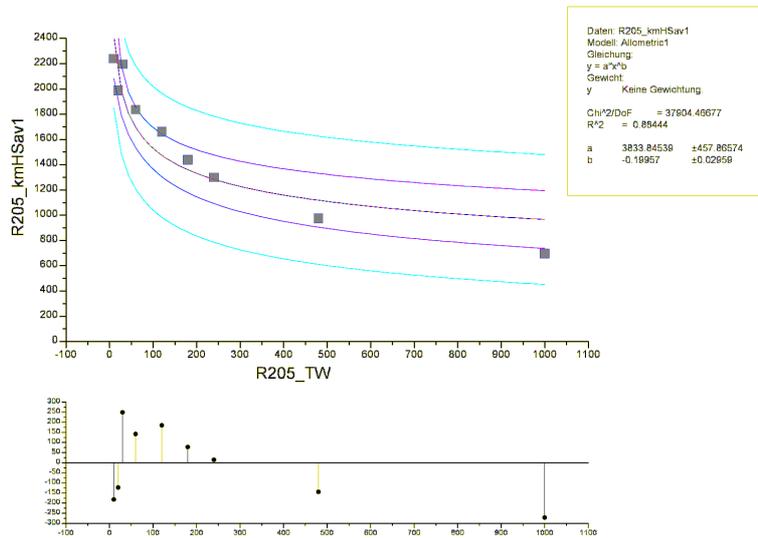


Abbildung 32: Exponentielle Regression unabhängig je Instanz

D.3 Km-Regressionsversuche für R205#

Die allometrischen Regressionsversuche bzgl. der Tourkilometer für die Instanz R205# waren nicht befriedigend, da kein R^2 größer 0,9 erreicht werden konnte. Die Ergebnisse der Regressionsversuche sind in den beiden folgenden Grafiken dargestellt.



Motiviert durch die schlechten Ergebnisse der allometrischen Regression für die Abhängigkeit der Tourkilometer von der Zeitfensterbreite wurden weitere Regressionsversuche mit der oben vorgestellten exponentiellen Funktion „YldFert1“ angestellt, die deutlich bessere Ergebnisse lieferten.

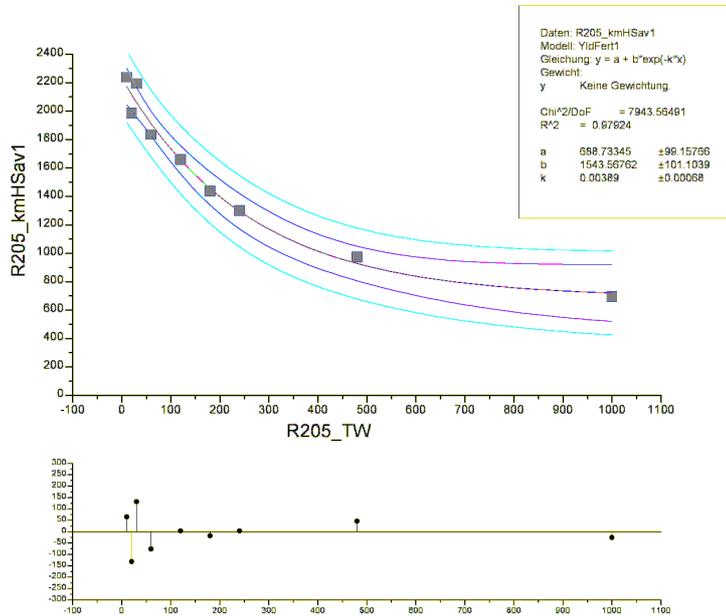


Abbildung 35: Exponentielle Regression Tourkilometer R205# (Savings)

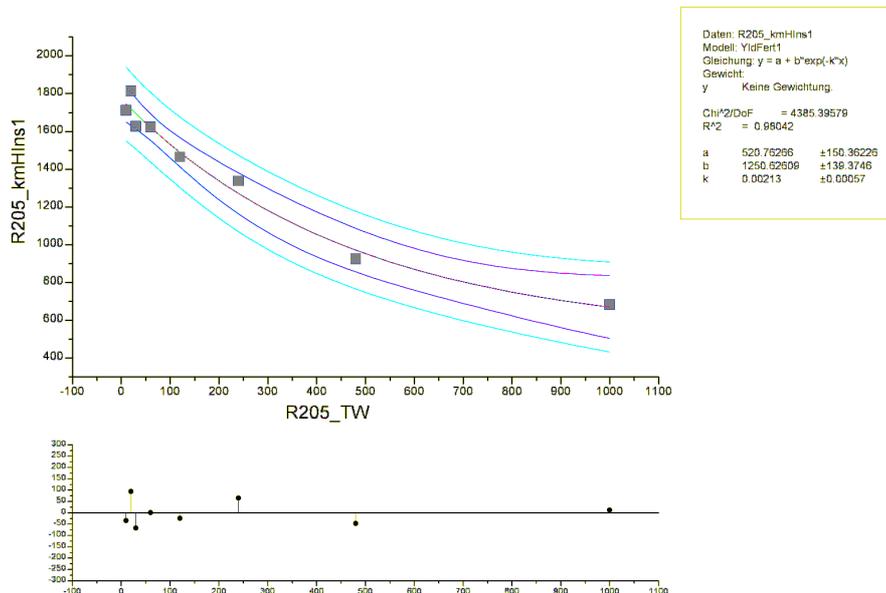


Abbildung 36: Exponentielle Regression Tourkilometer R205# (Insertion)

Universität Karlsruhe (TH)

Institut für Fördertechnik
und Logistiksysteme (IFL)



In dieser Arbeit wird das praxisrelevante Problem der Tourenplanung im eCommerceland-Lebensmittel-Heimlieferservice im Kontext transportlogistischer und betriebswirtschaftlicher Anforderungen untersucht. Ausgehend von der realen Aufgabenstellung wird eine neue Variante des Tourenplanungsproblems mathematisch modelliert - das D-m-VRSP-TW-UC-DO (dynamic m vehicle routing and scheduling problem with time constraints and uncertain customers with delivery notification at order entry).

Verschiedene in der Praxis gebräuchliche klassische Ansätze zur Lösung ähnlicher Probleme werden vorgestellt und einem neuen Planungsansatz gegenübergestellt. Für die Bewertung ausgewählter Verfahren wurden geeignete Solomon-Instanzen verwendet und wo nötig neue realitätsnahe Probleminstanzen generiert. Anhand dieser wird quantitativ gezeigt, wie die Transportkosten bei klassischen Ansätzen mit enger werdenden Lieferzeitfenstern steigen, während der neue Ansatz systembedingt auf enger werdende Lieferzeitfenster nur wenig sensitiv reagiert.

ISSN: 0171-2772

ISBN-13: 978-3-86644-082-1

ISBN-10: 3-86644-082-0

www.uvka.de