

## **KIVAS**

**KI-gestützte Kurzzeitprognosen für die Verbesserung von Fahrzeugeinsatz-  
und Auslastungsplanungen  
im deutschen Straßengüterverkehr**

### **Schlussbericht**

# Zusammenfassung

## **Problemstellung**

LKWs stellen eine große Belastung für Umwelt, Infrastruktur und Verkehrssicherheit dar. Aus diesem Grund ist es wichtig, den Straßengüterverkehr langfristig zu verringern oder zumindest dessen Wachstum zu bremsen. Eine effiziente Möglichkeit Verkehre zu reduzieren ist, Leerfahrten und Fahrten mit geringer Auslastung zu vermeiden. Damit Straßengüterverkehrsunternehmen (StGV-Unternehmen) ihre Touren so planen können, dass ihre Flotte möglichst gut ausgelastet ist – also Leerfahrten vermieden werden – benötigen sie möglichst genaue Bedarfsprognosen. Die meisten StGV-Unternehmen schätzen diese Bedarfe heutzutage zumeist noch manuell ab. Die Unternehmen, die bereits datengetriebene Prognosen nutzen, arbeiten in der Regel aber lediglich auf Basis interner historischer Daten. Jedoch erhalten sie dadurch keine präzisen Kurzzeitprognosen, die mehrere Einflussgrößen berücksichtigen.

## **Vorgehen**

Im Forschungsprojekt »KIVAS« untersuchten wir daher eine Vielzahl an Einflussgrößen, wie Wettervorhersagen, kalendarische Effekte (z. B. Feiertage, Wochenenden, etc.), Konjunkturdaten sowie sozioökonomische Daten (z. B. Arbeitslosenstatistiken) auf ihren prädiktiven Mehrwert.

Um die Relevanz von Variablen einschätzen zu können, werden in der KI-Forschung sogenannte Feature-Selection Verfahren verwendet. Diese lassen sich wiederum in drei Klassen einteilen: Zum einen gibt es sogenannte Filter-Verfahren, die beispielsweise ein Korrelationsmaß nutzen. Variablen die im Untersuchungsprozess hoch korrelieren, werden als relevant eingestuft. Zum anderen werden Wrapper-Verfahren verwendet, bei denen das Prognosemodell als Blackbox betrachtet und anhand der Veränderung des Prognosefehlers die Relevanz geschätzt wird. Darüber hinaus gibt es noch Embedded-Verfahren, die bereits beim Berechnen der Prognose die Wichtigkeit der jeweiligen Variablen prüfen.

Im Projekt »KIVAS« verwendeten wir alle drei KI-Verfahren und verglichen deren Ergebnisse miteinander. Konkret nutzten wir Fast Correlation Based Filter (Filter-Verfahren), Sequential-Floating-Forward-Selection (Wrapper-Verfahren) und LASSO (Embedded-Verfahren).

## **Ergebnis**

Die Prognose verbessern können vor allem Features zur Abbildung komplexer Kalendereffekte, wie wöchentliche und jährliche Saisonalitäten sowie Feiertage. Darüber hinaus scheint die Temperatur und Regen das Frachtvolumen zu beeinflussen. Nur vereinzelt werden ökonomische Daten als relevant betrachtet. In der Prognosestudie weisen alle multivariaten Verfahren ähnliche Prognosegenauigkeiten auf. Es treten einzelnen Ausreißern nach unten bei der Kombination bestimmter Feature-Selection Verfahren und Prognoseverfahren auf. Alle multivariaten Verfahren erzielen eine höhere Prognosegenauigkeit als der univariate Benchmark. Im Rahmen der Vorstudie können wir deshalb konkludieren, dass externe Daten die Prognosegenauigkeit für Frachtvolumen erhöhen.

## Inhaltsverzeichnis

1 Kurze Darstellung.....	4
1.1 Aufgabenstellung.....	4
1.2 Voraussetzungen unter denen das Vorhaben durchgeführt wurde.....	5
1.3 Planung und Ablauf des Vorhabens .....	6
1.4 Wissenschaftlicher und technischer Stand an den angeknüpft wurde.....	6
1.5 Zusammenarbeit mit anderen Stellen .....	7
2 Eingehende Darstellung des Projektablaufs und der Projektergebnisse .....	8
2.1 Identifikation relevanter Einflussgrößen und Aufbau einer Datenbank.....	8
2.1.1 Literaturrecherche zu Prognosedeterminanten im Güterverkehr.....	9
2.1.2 Übersicht über potenzielle Einflussfaktoren .....	11
2.2 Aufstellen von Hypothesen und spezifisches Aufbereiten der Daten .....	12
2.2.1 Aufgestellte Hypothesen .....	12
2.2.2 Datenaufbereitung.....	13
2.2 Feature-Selection: Auswahl geeigneter Variablen.....	14
2.3 Evaluierung verschiedener Prognoseverfahrens für die betrachteten Zeitreihen.	20
2.4 Verwendete Softwarepakete .....	22
2.5 Interpretation der Ergebnisse .....	22
3 Die wichtigsten Positionen des zahlenmäßigen Nachweises .....	23
4 Notwendigkeit und Angemessenheit der geleisteten Arbeit .....	23
5 Voraussichtlicher Nutzen und Verwertbarkeit.....	23
6 Innovationen bei anderen Stellen während der Durchführung des Vorhabens .....	23
7 Öffentlichkeitswirksame Maßnahmen .....	24
Literaturverzeichnis .....	25

# 1 Kurze Darstellung

## 1.1 Aufgabenstellung

Das anhaltend starke Wachstum des Straßengüterverkehrs (SGV) ist ein wesentlicher Mitverursacher von Infrastruktur- und Umweltbelastungen sowie von Verkehrssicherheitsproblemen. Allerdings sind die Leistungen des SGV ein wichtiger Bestandteil der Wettbewerbsfähigkeit der deutschen Wirtschaft. Jedoch fallen jedes Jahr rund 151 Mio. Leerfahrten an, sodass nicht alle LKW-Fahrten optimal ausgelastet sind. Eine präzise und dynamische Prognose des zu erwartenden Sendungsaufkommens ist somit ein wichtiges Instrument für die SGV-Unternehmen, um die Auslastungsquote zu erhöhen, Wachstum über vorhandene Kapazitäten zu stemmen und dadurch auch einen Beitrag für die Umwelt zu leisten. Typischerweise erfolgt diese Prognose nur mit internen historischen Daten der SGV-Unternehmen und bleibt somit rückwärtsgerichtet. Die Hinzunahme von externen, führenden Einflussgrößen (z.B. Wettervorhersagen, Feiertage o.ä.) kann eine realistischere und präzisere Prognose ermöglichen.

Im Projekt KIVAS wurden dazu Sendungsmengen zweier Straßengüterverkehrsunternehmen betrachtet. Das eine Unternehmen (*im Folgenden Unternehmen A*) bedient den Stückgutmarkt, das andere Unternehmen (*im Folgenden Unternehmen B*) ist auf die Lebensmittelbranche, insbesondere gekühlte Lebensmittel spezialisiert.

Unternehmen A unterhält ein Liefernetzwerk von über ganz Deutschland verteilten Niederlassungen. Die zugestellten und abgeholten Lieferungen in einer spezifischen Niederlassung werden in verschiedenen Kategorien unterteilt:

- *Direct Load*: Die Ware wird vom Verloader abgeholt und direkt an den Empfänger zugestellt. Sendungen werden nicht im Hub gebündelt.
- *Ausgang Fernverkehr*: Die Sendung wird vom Verloader abgeholt und zu einer anderen Niederlassung transportiert. Von dort wird sie schließlich zugestellt.
- *Nahverkehr*: Die Sendung wird von einem Verloader abgeholt, in die lokale Niederlassung transportiert und dort mit anderen Sendungen gebündelt und anschließend zugestellt.

Die obigen Transportmodi sind in Abbildung 1 schematisch dargestellt.

Unternehmen B betreibt ebenfalls über ganz Deutschland verteilte Standorte. Auch dort werden die zugestellten und abgeholten Lieferungen in verschiedene Kategorien eingeteilt.

- *Sammelgut*: Die Sendungen verschiedener Versender bzw. Empfänger werden gebündelt.
- *Direct Load*: Die Ware wird vom Verloader abgeholt und direkt an den Empfänger zugestellt. Sendungen werden nicht im Hub gebündelt.
- *Full Truck Load*: Die Ware wird vom Verloader abgeholt und direkt an den Empfänger zugestellt. Dabei ist der LKW durch die versendete Ware voll ausgelastet.

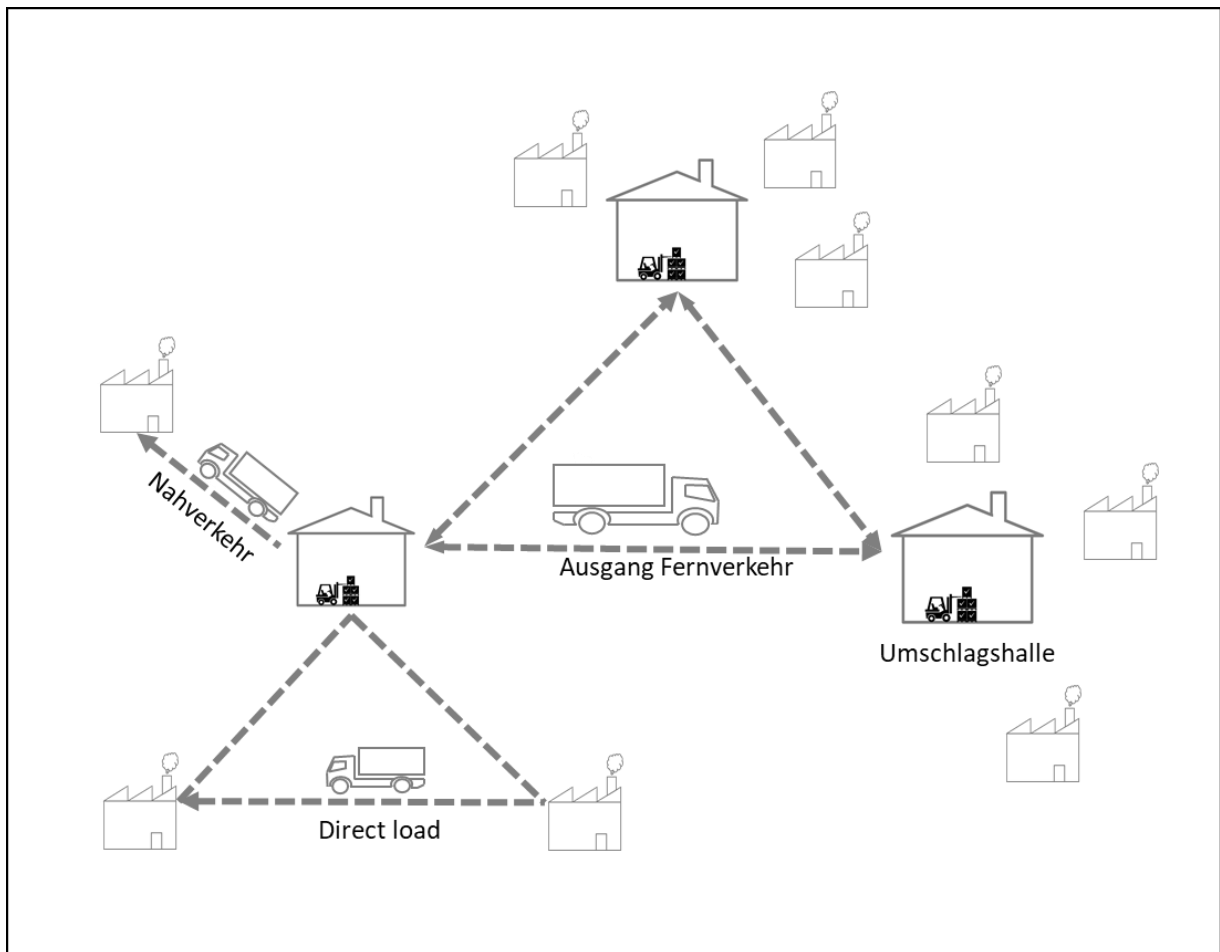


Abbildung 1: Schematische Darstellung des Liefernetzwerks von Unternehmen A

Um eine bessere Planung der zu allozierenden Ressourcen zu ermöglichen ist eine Prognose der Nachfrage nach Transportvolumen auf Basis der oben beschriebenen Transportmodi sowie auf Basis jeder einzelnen Niederlassung notwendig. Im Rahmen des Forschungsprojekts wählten wir einen Prognosehorizont von einem Tag auf Tagesebene. Es ist davon auszugehen, dass sich die wesentlichen Einflussfaktoren auf das Transportvolumen von Zeitreihe zu Zeitreihe unterscheiden.

## 1.2 Voraussetzungen unter denen das Vorhaben durchgeführt wurde

Das Projekt KIVAS wurde vom Bundesministerium für Verkehr und Infrastruktur im Rahmen des mFunds gefördert. Das Projekt war Teil der Förderlinie I, die Fördersumme betrug 99.475 €. Als Projektträger trat das Ministerium selbst auf.

Das Projektteam bestand aus der Fraunhofer Arbeitsgruppe für Supply Chain Services und dem Lehrstuhl für Logistik und Controlling an der Universität Regensburg. Darüber hinaus waren die beiden im obigen Abschnitt beschriebenen Straßenverkehrsgüterunternehmen als assoziierte Partner an dem Projekt beteiligt.

Das wissenschaftliche Institut für Infrastruktur und Kommunikationsdienste trat als Begleitforschung auf und organisierte insbesondere Veranstaltungen um sich mit anderen geförderten Projekten auszutauschen.

Der Zeitplan entsprechend des Antrags ging von einem Projektstart im Januar 2019 und vom Projektabschluss im September desselben Jahres aus. Dieser wurde eingehalten.

### 1.3 Planung und Ablauf des Vorhabens

Das Vorhaben gliederte sich in vier Arbeitspakete und drei Meilensteine. Untenstehende Abbildung 2 stellt den Projektablauf schematisch dar:

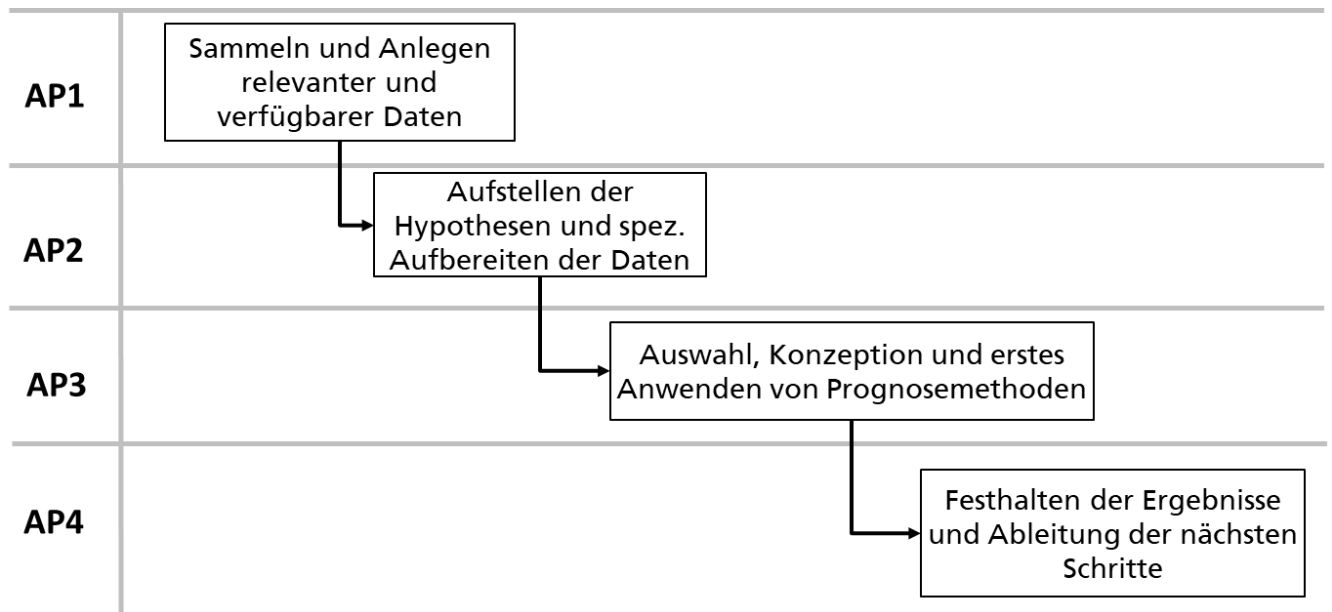


Abbildung 2: Projektablauf nach Arbeitspaketen

Es wurden drei Meilensteine festgelegt und erreicht:

- **Meilenstein 1 (31.02.2019):** Daten gesammelt und semantisch abgelegt
- **Meilenstein 2 (31.04.2019):** Hypothesenformulierung und Datenbereinigung abgeschlossen
- **Meilenstein 3 (31.07.2019):** Evaluierung der Prognosemethode abgeschlossen

Der Projektabschluss zum September 2019 konnte ebenfalls eingehalten werden.

### 1.4 Wissenschaftlicher und technischer Stand an den angeknüpft wurde

Die Prognose von Frachtvolumen ist dem Forschungsgebiet „Demand Forecasting“ zuzuordnen. Eine umfassende Übersicht über dieses Forschungsfeld würde über den Rahmen des vorliegenden Berichts hinausgehen, deshalb sei stattdessen auf die Übersichtsstudie von (Syntetos et al. 2016) verwiesen.

Spezifisch im Bereich Frachtvolumenprognose erforschten (Fite et al. 2002) den Nutzen ökonomischer Indizes auf das Frachtaufkommen eines großen amerikanischen StGV-Unternehmens. Sie werteten dabei eine Vielzahl an ökonomischen Indikatoren aus und überprüften den Zusammenhang mit dem Frachtvolumen anhand eines linearen Regressionsmodells. (Garrido und Mahmassani 2000) prognostizierten das Frachtvolumen eines ebenfalls amerikanischen StGV-Unternehmens auf Halbtages- und Niederlassungsebene. Sie verwendeten dafür ein multinomiales Probit Modell.

(Godfrey und Powell 2000) prognostizierten Frachtvolumen auf Tages- und Sendungsrelationsebene und entwickelten eine Methode exponentieller Glättung, welches ähnliche Performanz zeigte wie ein ARIMA Basismodell. (Zhou 2017) nutzte ein Ensemble aus einem ARIMA Modell und einem Multilayer Feed Forward Neural Network um das Transportvolumen eines LTL Transporteurs (= *Less than Truckload*) vorherzusagen. (Wang und Shi 2012) entwickelten ein hierarchisches Regressionsmodell um Frachtvolumen auf

Ebene von Straßen vorherzusagen. Sie nutzten dabei Kalendereffekte und sozioökonomische Daten als Prädiktoren. Keine der bisherigen Arbeiten hat bisher systematisch untersucht, welche Einflussgrößen auf das Frachtaufkommen spezifischer Unternehmen tatsächlich relevant sind.

## 1.5 Zusammenarbeit mit anderen Stellen

Im Rahmen verschiedener Veranstaltungen tauschten wir uns mit einigen anderen Forschungsprojekten des mFunds aus. Mit *Carrypicker* fand ein methodischer Austausch im Rahmen der mFund Konferenz in Berlin statt, mit *Cargonexx* im Rahmen der Transportlogistikmesse in München. Das Fraunhofer SCS organisierte gemeinsam mit der WIK-Begleitforschung eine KI-Tagung in Nürnberg.

## 2 Eingehende Darstellung des Projektablaufs und der Projektergebnisse

Für die Projektbearbeitung orientierten wir uns am CRISP-DM Referenzprozess. Diesen adaptierten wir, sodass sich das Projekt konkret in die folgenden Schritte gliederte:

- 1) *Identifikation relevanter Einflussgrößen und Aufbau einer Datenbank:* Mithilfe von Domänenwissen des Lehrstuhls der Uni Regensburg und des Fraunhofer SCS erstellen wir eine Sammlung geeigneter Prognosedeterminanten. Diese validieren wir im Gespräch mit Praktikern und mithilfe einer Literaturrecherche.
- 2) *Aufstellen von Hypothesen und spezifisches Aufbereiten der Daten:* Auf Basis der potenziell relevanten Einflussgrößen konstruieren die Forscher Hypothesen darüber welche potenziell wichtigen Einflussgrößen tatsächlich einen prädiktiven Mehrwert für die Prognose liefern. Anschließend bereiten sie die Daten so auf, dass die Hypothesen getestet werden können.
- 3) *Prüfung der Hypothesen:* Mittels verschiedener Feature-Selection Verfahren prüfen die Forscher die aufgestellten Hypothesen.
- 4) *Erste Prognosebenchmarks:* Die Forscher vergleichen verschiedene Machine-Learning-Algorithmen auf Ihre Performanz für die gegebene Problemstellung und verschiedene in Schritt (3) konstruierte Featuresets<sup>1</sup>.

### 2.1 Identifikation relevanter Einflussgrößen und Aufbau einer Datenbank

Um eine sinnvolle Prognose für Sendungsvolumina im Straßengüterverkehr konstruieren zu können, sollten unternehmensexterne Daten gesammelt und ausgewertet werden. Für diese Daten wurden zunächst grundsätzliche Anforderungen definiert. Alle zu verwendenden Daten müssen flächendeckend für ganz Deutschland erhoben werden. Zudem muss die Erhebung der Daten in einer konsistenten Taktung erfolgen. Außerdem müssen die Daten, um sich für die Verwendung zu qualifizieren, frei zugänglich sein. Zuletzt muss die Nachvollziehbarkeit der Erhebung der Daten gewährleistet sein.

Zu Beginn des Forschungsvorhabens stellten wir eine erste Auswahl von möglichen Prognosedeterminanten auf, die wir nachfolgend durch eine Literaturrecherche verifizierten. Als mögliche Einflussgrößen identifizierten wir im ersten Schritt:

- Wettervorhersagen
- Baustellen auf Autobahnen
- Arbeitsmarktdaten
- Verbraucherpreisindizes
- Bevölkerungsdaten
- Aggregate zu privaten oder staatlichen Konsumausgaben
- Bruttoinvestitionen
- Bauinvestitionen
- Import- und Exportkenngrößen
- Branchenentwicklungen
- Bruttowertschöpfungsentwicklungen

Um die Validität dieser Determinatenvorschläge zu überprüfen, führten wir eine Literaturrecherche durch.

---

<sup>1</sup> Ein Featureset bezeichnet eine Auswahl an Features. Ein Feature bezeichnet eine unabhängige Variable bzw. eine Einflussgröße.



## 2.1.1 Literaturrecherche zu Prognosedeterminanten im Güterverkehr

Die Literaturrecherche erfolgte unter Annahme des in Abbildung 1 dargestellten Modells. Wir gehen also davon aus, dass der Güterfluss bei einem Versender beginnt, der grundsätzlich zwischen einer Vielzahl von Beförderungsmöglichkeiten wählen kann. Dabei können die Möglichkeiten sowohl durch ihren spezifischen Beförderungsmodus (Straßenverkehr, Schienenverkehr, Binnenschifffahrt...) als auch durch die innerhalb des Beförderungsmodus konkurrierenden Anbieter unterschieden werden.

Anhand des Modells können die Veränderungen der Sendungsvolumina immer einer bestimmten Menge an Akteuren zugeschrieben werden. Volumenschwankungen können damit vom Versender, der Konkurrenz (sowohl innerhalb als auch außerhalb des Straßenverkehrsmodus), dem Handel oder den Konsumenten induziert werden. Aus diesem Grund werden Determinanten als „versenderbezogen“, „konkurrenzbezogen“, „handelbezogen“ oder „konsumentenbezogen“ klassifiziert.

Im Folgenden liegt der Fokus auf der Kette „Versender – Betrachteter Logistic Service Provider (LSP) – Handel – Konsumenten“.

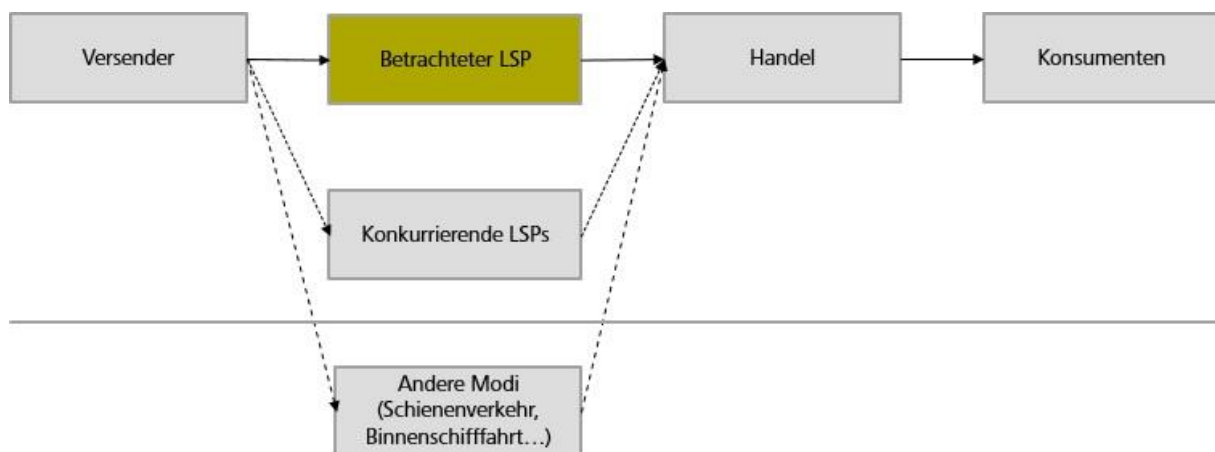


Abbildung 3: Güterströme und beteiligte Akteure

Im Rahmen der Literaturrecherche suchten wir vor allem Beiträge, die sich mit der Prognose von Frachtvolumina im LKW-Verkehr beschäftigen. Zusätzlich wurde die in der Literatur gestützte Annahme getroffen, dass sich der LKW-Ladungsverkehr vor allem durch das Auffüllen von reduziertem Lagerbestand ergibt (Lahiri et al. 2004, S. 103), was wiederum in Folge von Produktverkäufen nötig ist. Somit hängt das Frachtvolumen unmittelbar mit den Verkaufszahlen transportierter Produkte zusammen. Aus diesem Grund wurden in der Literaturrecherche auch Beiträge zu Prognosedeterminanten von Produktverkaufszahlen berücksichtigt.

Die in Kapitel 2 genannten Determinanten lassen sich inhaltlich in vier verschiedene Kategorien einordnen. Wettervorhersagen sind der Klasse der meteorologischen Einflüsse zuzuordnen. Baustellen auf Autobahnen spiegeln einen direkten Einfluss auf die Abwicklung von Transportdienstleistungen wider, es handelt sich also um spezifische Einflüsse. Als soziodemografische Einflüsse lassen sich alle Daten rund um die Bevölkerung eines Landes einordnen. Die restlichen Determinanten können dem Feld der makroökonomischen Einflüsse zugewiesen werden.

### *Meteorologische Einflüsse*

In der Literatur finden sich Ansätze für Abhängigkeiten von Wetterdaten und Sendungsvolumina. Dabei wird herausgestellt, dass in den Sommermonaten die durchschnittlich transportierten Mengen, gemessen am Sendungsgewicht, höher sind, als in den kalten Monaten (Wang und Shi 2012, S. 2756). Außerdem findet man bei der Recherche

Evidenz für den Zusammenhang von Wetterdaten und Absatzzahlen bestimmter Produkte (Bertrand et al. 2015, S. 261). Dabei beschränkt sich die Abhängigkeit nicht nur auf die Ausprägung der Temperatur (Ramanathan und Muyltermans 2010, S. 538; Arunraj und Ahrens 2016, S. 750; Ramanathan und Muyltermans 2010, S. 540; Bertrand et al. 2015, S. 269; Bahng und Kincade 2012, S. 424), sondern besteht auch zwischen den Produktabsatzzahlen und den Faktoren Sonnenscheinstunden (Murray et al. 2010, S. 517; van Donselaar et al. 2016, S. 70; Dekker et al. 2004, S. 161), Luftfeuchtigkeit und Schneefall (Murray et al. 2010, S. 517; Arunraj und Ahrens 2016, S. 748). Meteorologische Einflüsse können sowohl als handelbezogen als auch als konsumentenbezogen klassifiziert werden.

### *Spezifische Einflüsse*

Es konnte ein Einfluss des Schienengüterverkehrs auf die LKW-Sendungsvolumina gefunden werden. Nach Forkenbrock (2001, S. 323) können 36,4% des gesamten LKW- und Schienengüterverkehr durch die jeweils andere Transportart substituiert werden, was zu einer Wechselmöglichkeit für Versender führt, um auf Probleme im Schienen- oder Straßenverkehr zu reagieren. Außerdem könnten Benzinpreise einen Einfluss auf die Entwicklung der Sendungsvolumina haben (Lahiri und Yao 2006, S. 886). Bei den spezifischen Einflüssen handelt es sich um konkurrenzbezogene Determinanten.

### *Makroökonomische Einflüsse*

Literaturbeiträge erwähnen den Zusammenhang von Transportvolumina und der wirtschaftlichen Lage eines Landes (Lahiri et al. 2004, S. 110; Lahiri und Yao 2006, S. 873). Dabei können zur Messung der Abhängigkeiten viele verschiedene Faktoren herangezogen werden, da eine höhere wirtschaftliche Aktivität oft in einem höheren Güterfluss resultiert (Fite et al. 2002, S. 301; Zhou 2017, S. 2). Verwendete Größen, um Transportvolumina zu prognostizieren, sind das Bruttonettoprodukt (Saidi und Hammami 2017, S. 422; Kim et al. 2011, S. 349; Guo et al. 2013) und der Verbraucherpreisindex (Guo et al. 2013). Außerdem werden das Wirtschaftswachstum (Kim et al. 2011, S. 349; Lahiri und Yao 2006, S. 886; Fite et al. 2002, S. 301; Saidi und Hammami 2017, S. 425; Kim et al. 2011, S. 343), der Leitzins (Lahiri und Yao 2006, S. 886) und der Wechselkurs einzelner Währungen als Einflussfaktoren auf die Transportbranche genannt (Kim et al. 2011, S. 343). Wegweisend für die Entwicklung von Produktabsatzzahlen kann auch die Wertpapierrendite ausgewählter Märkte sein (Osadchiy et al. 2013, S. 1065–1066). Makroökonomische Einflüssen werden als handel- und konsumentenbezogen eingestuft.

### *Soziodemografische Einflüsse*

Ebenfalls zur Prognose verwendete Daten sind Einwohnerzahlen eines bestimmten Gebiets (Garrido und Mahmassani 2000, S. 408). Dabei handelt es sich um eine handelsbezogene und konsumentenbezogene Determinante.

### *Kalendarische Einflüsse*

Zusätzlich zu den anfänglich festgelegten Determinanten und ihren Klassifikationen können kalendarische Einflüsse als Verbesserungspotenziale von Prognosen in der Literatur gefunden werden. Dabei wird zwischen periodischen und nicht-periodischen Effekten unterschieden (Godfrey und Powell 2000, S. 465). Periodische Effekte beziehen sich dabei auf die Anzahl der Tage eines Monats, die Struktur der Wochentage (Alp Ertem 2011, S. 3) oder das generelle Auftreten von Saisonalitäten in Bezug auf einzelne Produktgruppen (Ramanathan und Muyltermans 2010, S. 543). Feiertage und Schulferien kommen sowohl nicht-periodisch als auch periodisch vor, da z.B. Weihnachten und Neujahr immer an einem fixen Kalenderdatum stattfinden, wohingegen Ostern und die Schulferien des jeweiligen Bundeslandes zu unterschiedlichen Zeitintervallen auftreten können. Feiertage und Schulferien können Produktverkäufe und folglich Sendungsvolumina erheblich beeinflussen (Aburto und Weber

2007, S. 143; Ramanathan und Muyldermans 2010, S. 540; Taylor 2007, S. 159). Kalendarische Einflüsse können sowohl versenderbezogen, handelsbezogen als auch konsumentenbezogen auftreten.

Es gilt nun die erarbeiteten Ergebnisse auf eine Sammlung von Determinanten zu reduzieren, die im vorliegenden Fall möglicherweise einen signifikanten Einfluss auf die gemessenen Sendungsvolumina haben.

### 2.1.2 Übersicht über potenzielle Einflussfaktoren

Wir identifizierten mögliche Features auf zwei Wegen: Zum einen könnten Features, welche andere Forscher genutzt haben, relevant sein. Diese können wir aus der Literaturanalyse ableiten. Zum anderen prüften wir, ob diese Features aus der Literatur auch für den vorliegenden Anwendungsfall relevant sein können. Zusätzlich ergänzten wir weitere Features auf Basis unseres eigenen Verständnisses der Domäne. Anschließend prüften wir visuell ob die so identifizierten Features Einfluss auf das Frachtvolumen haben können. Beispielsweise Einflussgrößen, die sich im Betrachtungszeitraum kaum ändern – wie der Leitzins –, schlossen wir in diesem dritten Schritt aus. In den untenstehenden Tabellen geben wir eine Übersicht, welche Features wir in welchem Schritt für relevant halten. Aus den relevanten Features leiteten wir dann die im nächsten Kapitel beschriebenen Hypothesen ab.

#### *Meteorologische Einflüsse*

<b>Bezeichnung</b>	<b>Literatur</b>	<b>Domänen- verständnis</b>	<b>Visuelle Analyse</b>
<b>Temperatur</b>	X	X	X
<b>Sonnenscheindauer</b>	X	X	
<b>Luftfeuchtigkeit</b>	X	X	
<b>Niederschlag</b>	X	X	X

#### *Spezifische Einflüsse*

<b>Bezeichnung</b>	<b>Literatur</b>	<b>Domänen- verständnis</b>	<b>Visuelle Analyse</b>
<b>Ausfall Schienenverkehr</b>	X		
<b>Dieselpreise</b>	X	X	

### *Makroökonomische Einflüsse*

<b>Bezeichnung</b>	<b>Literatur</b>	<b>Domänen- verständnis</b>	<b>Visuelle Analyse</b>
<b>Arbeitsmarktdaten</b>		X	X
<b>Verbraucherpreisindizes</b>	X	X	
<b>Import- / Exportkenngrößen</b>		X	X
<b>Branchenentwicklung</b>		X	
<b>Bruttowertschöpfungsentwicklungen</b>		X	X
<b>Wirtschaftswachstum</b>	X	X	
<b>Leitzins</b>	X		
<b>Tauschrate der Top-10 Import- /Exportpartner außerhalb Eurozone</b>	X	X	
<b>Wertpapierrendite</b>	X	X	

### *Soziodemografische Einflüsse*

<b>Bezeichnung</b>	<b>Literatur</b>	<b>Domänen- verständnis</b>	<b>Visuelle Analyse</b>
<b>Bevölkerungszahl</b>	X	X	

### *Kalendarische Einflüsse*

<b>Bezeichnung</b>	<b>Literatur</b>	<b>Domänen- verständnis</b>	<b>Visuelle Analyse</b>
<b>Anzahl Tage Monat</b>	X	X	
<b>Wochentage</b>	X	X	X
<b>Feiertage</b>	X	X	X
<b>Schulferien</b>	X	X	X
<b>Jährliche Saisonalität</b>		X	X
<b>Zeit vor Weihnachten</b>		X	X

## 2.2 Aufstellen von Hypothesen und spezifisches Aufbereiten der Daten

Die Einflussgrößen, die wir im Rahmen der visuellen Analyse für potenziell relevant hielten, untersuchten wir auf ihren Mehrwert für die Prognose. Im Folgenden sind unsere aufgestellten Hypothesen und die darauffolgende Datenaufbereitung beschrieben.

### 2.2.1 Aufgestellte Hypothesen

Unsere Hypothese ist, dass folgende unternehmensexterne Daten einen prädiktiven Mehrwert für die kurzfristige Prognose von Frachtvolumen liefern:

#### *Kalendarische Effekte:*

- Feiertage
- Schulferien
- Wochentage
- Zeit vor Weihnachten

#### *Wettereffekte:*

- Temperatur
- Regen

### *Volkswirtschaftliche Effekte:*

- IFO – Geschäftsklimaindex
- Handelsstatistik Import
- Handelsstatistik Export
- Bruttoinlandsprodukt
- Arbeitslosenstatistik

Darüber hinaus konstruieren wir aus den Zeitreihen des Transportvolumens noch abgeleitete Variablen, deren Mehrwert für die Prognose wir ebenfalls testen:

- Zeitversetzte Beobachtungen des Transportvolumens
- Gleitender Mittelwert des Transportvolumens der letzten von Tage
- Fourierterme zur Abbildung der wöchentlichen Saisonalität
- Fourierterme zur Abbildung der jährlichen Saisonalität

## 2.2.2 Datenaufbereitung

Da die Prognose letztlich auf Ebene der Niederlassung bzw. des entsprechenden Transportmodus sowie auf Tagesebene erfolgen soll müssen die relevanten externen Daten auf die Schlüsselspalten „Datum“ und „Niederlassung“ aggregiert werden.

Bei den Variablen, welche das Wetter beschreiben beziehen sich die Daten dann jeweils auf die der Niederlassung nächstgelegene Wetterstation<sup>2</sup>. Feiertage und Ferientage werden entsprechend des Bundeslands der Niederlassung zugeordnet und mit „dummy Variablen“<sup>3</sup> codiert.

Volkswirtschaftliche Indikatoren werden teilweise nur monatlich bzw. je Quartal erhoben und werden deshalb auf tagesebene disaggregiert.

### *Zeitlicher Versatz*

Um verschiedene Prognosehorizonte abzubilden werden sogenannte „lag“-Variablen eingeführt, also der zeitliche Versatz von unabhängigen Variablen<sup>4</sup> um den Prognosehorizont. Bei Variablen, welche für die Zukunft bekannt sind, wie beispielsweise bei Feiertagen, oder in begrenztem Umfang dem Wetter, werden darüber hinaus „lead“-Variablen eingeführt, welche beschreiben, dass ein bekanntes Ereignis in der Zukunft einen Einfluss auf die aktuelle Periode hat. Auch die Zielvariablen an sich werden so versetzt, um die Autokorrelation der Variablen abzubilden.

### *Umgang mit fehlenden Werten*

Die im Projekt zur Prognose verwendeten Verfahren können fehlende Werte nicht berücksichtigen. Um diese Verfahren trotzdem anwenden zu können, müssen fehlende Werte behandelt werden. Denkbar sind dabei drei Strategien: Das Löschen von ganzen Zeilen mit fehlenden Werten, das Löschen von ganzen Spalten mit fehlenden Werten oder die Ersetzung fehlender Werte.

Entsprechend unseres Domänenwissens behandelten wir fehlende Werte je nach Variable unterschiedlich. Fehlende Werte beim Transportaufkommen je Tag wurden durch 0 ersetzt – an dem spezifischen Tag wird bei fehlenden Aufträgen, davon ausgegangen, dass tatsächlich

---

<sup>2</sup>Eine Liste der Wetterstationen des Deutschen Wetterdiensts ist hier zu finden: <https://www.dwd.de/DE/leistungen/klimadatendeutschland/stationsuebersicht.html>

<sup>3</sup> Dies beschreibt die Darstellung einer ursprünglich kategorialen Variable mittels 1 (tritt ein) oder 0 (tritt nicht ein)

<sup>4</sup> Unabhängige Variablen beschreiben in diesem Kontext vermutete Einflussgrößen auf das Transportvolumen, wie beispielsweise die Temperatur

kein Transportvolumen transportiert wurde. Diese Annahme hielt einer visuellen Plausibilitätsprüfung stand.

Außerdem treten bei den Wetterdaten vereinzelt fehlende Werte auf. Bei diesen gehen die Forscher davon aus, dass sie sich durch andere Wettervariablen näherungsweise modellieren lassen. Beispielsweise bedeutet das, dass die Luftfeuchtigkeit anhand der Temperatur und des Regenfalls geschätzt wird. Konkret wurde multiple Imputation (predictive mean matching) auf den das Wetter beschreibenden Variablen durchgeführt. Dafür verwendeten die Forscher das MICE Paket (van Buuren und Groothuis-Oudshoorn 2011), welches den Algorithmus von (Little und Rubin 1987) implementiert.

Durch die Einführung der „lag“- und „lead“-Variablen wurden ebenfalls fehlende Werte am Anfang bzw. Ende der Zeitreihe erzeugt. Die entsprechenden Zeilen wurden gelöscht.

### *Behandlung von Wochenenden*

Bedarf an Transportvolumen fällt regelmäßig nur an Werktagen an – entsprechend wurden Wochenenden aus der Zeitreihe entfernt.

### *Stationarität*

Die verwendeten Algorithmen zur Zeitreihenprognose (Random Forest, Support Vector Regression, ARIMAX) erfordern üblicherweise stationäre Zeitreihen. Falls die Stationarität nicht gegeben ist, kann gegebenenfalls differenziert werden, um Stationarität herzustellen. Zur Prüfung, ob die gegebenen Zeitreihen stationär sind, nutzten wir den Augmented Dickey-Fuller (ADF) Test und den Kwiatowski-Phillips-Schmidt-Shin (KPSS) Test. Lediglich bei den ökonomischen Zeitreihen stellten beide Tests Instationarität fest. Allerdings ist es aus dem Domänenkontext heraus trotzdem sinnvoller, dass das Niveau der ökonomischen Zeitreihen Einfluss auf das Frachtvolumen hat und weniger die Änderung desselben. Deshalb verzichteten wir auch bei diesen auf Differenzierung.

## 2.3 Feature-Selection: Auswahl geeigneter Variablen

Um zu prüfen, welche Variablen für eine Prognosefragestellung tatsächlich relevant sind, werden Feature-Selection Verfahren genutzt. Dabei unterscheidet man drei verschiedene Ansätze: Filter-, Wrapper- und Embedded-Verfahren. Filter-Verfahren nutzen Metriken, wie beispielsweise den Pearson-Korrelations-Koeffizienten und Grenzwerte. Damit lassen sich allerdings nur Zusammenhänge erfassen, welche durch die entsprechende Metrik abgebildet werden. Einige Metriken erfassen beispielsweise lediglich lineare Zusammenhänge. Zudem ist die Festlegung des Grenzwerts häufig willkürlich. Eine Lösung für dieses Problem bieten Wrapper-Verfahren. Dabei wird ein bestimmter Prognosealgorithmus als Blackbox gesehen und es werden iterativ Variablen zum ursprünglichen Variablenset hinzu- oder weggenommen. Die Änderung des Prognosefehlers auf einem Validierungsset gibt dann Aufschluss über den Nutzen dieser spezifischen Variable. Als Blackbox kann dabei ein Algorithmus genommen werden, welcher alle möglichen Funktionen approximieren kann, wie beispielsweise baumbasierte Verfahren, Neuronale Netze oder Support Vector Regression (mit nichtlinearem Kernel). So ist es möglich auch relevante Features zu finden die in nichtlinearem Verhältnis zur Zielgröße stehen. Dies vergrößert allerdings den Suchraum stark, sodass Wrapper-Verfahren zu Overfitting neigen. Overfitting beschreibt hier das Selektieren von Features welche zwar zufällig auf dem Validierungsset den Prognosefehler verkleinern, allerdings keinen echten Mehrwert für die Prognose bieten. Embedded-Verfahren führen inhärent eine Variablenselektion durch. In der Regel werden dafür Strafterme genutzt, die beim Trainieren des Modells bereits die Summe der gewählten Parameter bestrafen.

Jede der beschriebenen Verfahrensklassen hat spezifische Vor- und Nachteile und es ist nicht möglich mithilfe der Anwendung einer dieser Methoden final den Nutzen einzelner Features zu beurteilen. Zunehmend werden deshalb auch Feature-Selection Ensembles genutzt, welche

verschiedene Feature-Selection Verfahren miteinander kombinieren und daraus ein Ergebnis konstruieren.

In unserem Forschungsdesign nutzten wir folgende drei Feature-Selection Verfahren: Fast Correlation Based Filter (FCBF, ein Filter-Verfahren), Sequential-Floating-Forward-Selection (SFFS, ein Wrapper-Verfahren) und Least Absolute Shrinkage and Selection Operator (LASSO, ein Embedded-Verfahren). Das SFFS-Verfahren kombinierten wir darüber hinaus mit drei verschiedenen Machine-Learning-Algorithmen: ARIMAX, Random Forest (RF) und Support Vector Regression (SVR). Die drei Algorithmen nutzten wir anschließend auch zur Prognosestudie. Neuronale Netze innerhalb des SFFS-Verfahrens zu nutzen ist nicht möglich, da eine Vielzahl von Modellen je Iteration trainiert werden muss und deshalb die Laufzeit impraktikabel ist.

Zur Einführung in Verfahren zur Feature-Selection verweisen wir auf (Chandrashekar und Sahin 2014). Das FCBF Verfahren wurde von (Yu und Liu 2003) entwickelt. (Pudil et al. 1994) publizierten das SFFS-Verfahren. Zum Problem des Overfittings bei Wrapper-Verfahren sei auf (Loughrey und Cunningham 2005) verwiesen. Für das LASSO-Verfahren empfehlen wir (James et al. 2013), bzw. (Hastie et al. 2017) für eine tiefergehende Einführung. Diese Autoren beschreiben auch die genannten Prognosealgorithmen und speziell (James et al. 2013) gibt auch eine gute Einführung zu Machine Learning im Allgemeinen.

### *Durchführung der Feature-Selection*

Im Rahmen der Studie beschränkten wir uns auf sechs Niederlassungen (3 je untersuchtem Unternehmen) und spezifisch auf die Abholung von den Endkunden (Ausgang Fernverkehr, bzw. Sammelgut). Für diese sechs Zeitreihen untersuchten wir die unter 2.2.1 aufgestellten Hypothesen. Konkret untersuchten wir die in Tabelle 1 aufgelisteten Variablen.

<b>Variable</b>	<b>Lags/Leads</b>	<b>Erläuterung</b>
<b>Transportvolumen</b>	1:5	Zeitversetzte Zielvariable (Nicht genutzt für die ARIMA(X) Modelle)
<b>Rollierender Mittelwert Transportvolumen</b>	1:5	Rollierender Mittelwert der letzten fünf Werte
<b>Feiertage</b>	-5:5	Binäre Variable die 1 ist, wenn in der Region der spezifischen Niederlassung ein Feiertag ist.
<b>Ferientage</b>	0	Binäre Variable die 1 ist, wenn in der Region der spezifischen Niederlassung Schulferien sind.
<b>Wochentage</b>	0	Binäre Variable die Montag bis Freitag beschreibt
<b>Zeit vor Weihnachten</b>	0	Binäre Variable die die Tage von 01/12 bis 23/12 beschreibt.
<b>Fourier Terme für die wöchentliche Saisonalität</b>	0	Fourierterme zur Beschreibung der wöchentlichen Saisonalität (Ordnung 1 und 2)
<b>Fourier Terme für die jährliche Saisonalität</b>	0	Fourierterme zur Beschreibung der jährlichen Saisonalität (Ordnung 1 bis 10)
<b>Temperatur</b>	-5:5	Mittlere Tagestemperatur, gegeben durch die nächste Wetterstation zur untersuchten Niederlassung
<b>Tagessumme Regen</b>	-5:5	Regensumme an diesem Tag, gegeben durch die nächste Wetterstation zur untersuchten Niederlassung
<b>ifo Geschäftsklima-index</b>	1 Monat	IFO Geschäftsklimaindex, herausgegeben vom IFO Institut, basierend auf der Befragung eines Panels.
<b>Handelsstatistik import</b>	2 Monate	Ökonomischer Wert aller nach Deutschland importierten Produkte
<b>Handelsstatistik</b>	2 Monate	Ökonomischer Wert aller aus Deutschland exportierten Produkte

<b>export</b>		
<b>Bruttoinlandsprodukt</b>	135 Tage	Ökonomischer Wert aller Endprodukte und Dienstleistungen in Deutschland.
<b>Arbeitslosenrate</b>	1 Monat, 5 Tage	Anzahl der Arbeitslosen in Deutschland

*Tabelle 1: Auf den prädiktiven Mehrwert untersuchte Variablen*

### *Ergebnisse der Feature-Selection*

Im Folgenden stellen wir die aggregierten Ergebnisse der Variablenselektion und die einzelnen Ergebnisse je untersuchter Niederlassung bzw. Zeitreihe dar. Links sind dabei jeweils die untersuchten Variablen aufgeführt. Die Spaltenüberschriften bezeichnen die verschiedenen Verfahren zur Feature-Selection. Ein „X“ kennzeichnet, wenn ein bestimmtes Feature gewählt wurde. Bei der über alle Niederlassungen aggregierten Tabelle bezeichnen die Zahlen wie häufig ein Feature ausgewählt wurde (maximal 6-, minimal 0-mal).

### **Übersicht über die gewählten Features, aggregiert über alle Niederlassungen:**

Feature	FCBF	SFFS - ARIMAX	SFFS - RF	SFFS - SVR	LASSO	Summe
<b>Transportvolumen</b>	3	0	4	5	6	<b>18</b>
<b>Rollierender Mittelwert Transportvolumen</b>	4	0	2	4	6	<b>16</b>
<b>Feiertage</b>	6	6	6	6	6	<b>30</b>
<b>Ferientage</b>	1	2	0	0	1	<b>4</b>
<b>Wochentage</b>	0	4	2	2	6	<b>14</b>
<b>Zeit vor Weihnachten</b>	0	2	0	0	5	<b>7</b>
<b>Fourier Terme wöchentliche Saisonalität</b>	1	6	3	4	6	<b>20</b>
<b>Fourier Terme jährliche Saisonalität</b>	4	6	5	2	6	<b>23</b>
<b>Temperatur</b>	0	5	3	3	4	<b>15</b>
<b>Tagessumme Regen</b>	0	6	2	2	6	<b>16</b>
<b>ifo Geschäftsklimaindex</b>	1	3	0	0	4	<b>8</b>
<b>Handelsstatistik import</b>	0	2	0	0	5	<b>7</b>
<b>Handelsstatistik export</b>	0	4	1	0	1	<b>6</b>
<b>Bruttoinlandsprodukt</b>	0	1	1	0	3	<b>5</b>
<b>Arbeitslosenrate</b>	0	0	1	0	3	<b>4</b>
<b>Summe</b>	<b>20</b>	<b>47</b>	<b>30</b>	<b>28</b>	<b>68</b>	

*Tabelle 2: Übersicht über die gewählten Features je Methode über alle Niederlassungen aggregiert*



### Gewählte Features für das Unternehmen A – Niederlassung 1:

Feature	FCBF	SFFS - ARIMAX	SFFS - RF	SFFS - SVR	LASSO
Transportvolumen		-	X	X	X
Rollierender Mittelwert Transportvolumen	X	-	X	X	X
Feiertage	X	X	X	X	X
Ferientage	X	X			
Wochentage		X			X
Zeit vor Weihnachten					X
Fourier Terme wöchentliche Saisonalität		X			X
Fourier Terme jährliche Saisonalität	X	X	X		X
Temperatur		X		X	
Tagessumme Regen		X			X
ifo Geschäftsklimaindex	X				X
Handelsstatistik import					
Handelsstatistik export		X			
Bruttoinlandsprodukt		X			
Arbeitslosenrate					

Tabella 3: Gewählte Features je Verfahren für Unternehmen A - Niederlassung 1

### Gewählte Features für das Unternehmen A – Niederlassung 2:

Feature	FCBF	SFFS - ARIMAX	SFFS - RF	SFFS - SVR	LASSO
Transportvolumen	X	-		X	X
Rollierender Mittelwert Transportvolumen	X	-	X	X	X
Feiertage	X	X	X	X	X
Ferientage					
Wochentage			X	X	X
Zeit vor Weihnachten		X			X
Fourier Terme wöchentliche Saisonalität		X			X
Fourier Terme jährliche Saisonalität		X	X		X
Temperatur		X	X		X
Tagessumme Regen		X	X		X
ifo Geschäftsklimaindex		X			
Handelsstatistik import		X			X
Handelsstatistik export		X	X		
Bruttoinlandsprodukt					
Arbeitslosenrate					

Tabella 4: Gewählte Features je Verfahren für Unternehmen A - Niederlassung 2

### Gewählte Features für das Unternehmen A – Niederlassung 3:

Feature	FCBF	SFFS - ARIMAX	SFFS - RF	SFFS - SVR	LASSO
Transportvolumen	X	-	X	X	X
Rollierender Mittelwert Transportvolumen	X	-			X
Feiertage	X	X	X	X	X
Ferientage					
Wochentage					X
Zeit vor Weihnachten					
Fourier Terme wöchentliche Saisonalität		X	X	X	X
Fourier Terme jährliche Saisonalität	X	X	X	X	X
Temperatur		X	X		
Tagessumme Regen		X		X	X
ifo Geschäftsklimaindex		X			X
Handelsstatistik import					X
Handelsstatistik export					
Bruttoinlandsprodukt					
Arbeitslosenrate			X		

Tabelle 5: Gewählte Features je Verfahren für Unternehmen A - Niederlassung 3

### Gewählte Features für das Unternehmen B – Niederlassung 1:

Feature	FCBF	SFFS - ARIMAX	SFFS - RF	SFFS - SVR	LASSO
Transportvolumen		-	X		X
Rollierender Mittelwert Transportvolumen		-			X
Feiertage	X	X	X	X	X
Ferientage		X			X
Wochentage		X			X
Zeit vor Weihnachten					X
Fourier Terme wöchentliche Saisonalität		X	X	X	X
Fourier Terme jährliche Saisonalität		X	X	X	X
Temperatur		X			X
Tagessumme Regen		X		X	X
ifo Geschäftsklimaindex					
Handelsstatistik import		X			X
Handelsstatistik export		X			X
Bruttoinlandsprodukt					X
Arbeitslosenrate					X

Tabelle 6: Gewählte Features je Verfahren für Unternehmen B - Niederlassung 1

### Gewählte Features für das Unternehmen B – Niederlassung 2:

Feature	FCBF	SFFS - ARIMAX	SFFS - RF	SFFS - SVR	LASSO
Transportvolumen	X	-	X	X	X
Rollierender Mittelwert Transportvolumen	X	-		X	X
Feiertage	X	X	X	X	X
Ferientage					
Wochentage		X		X	X
Zeit vor Weihnachten		X			X
Fourier Terme wöchentliche Saisonalität		X	X	X	X
Fourier Terme jährliche Saisonalität	X	X			X
Temperatur			X	X	X
Tagessumme Regen		X	X		X
ifo Geschäftsklimaindex		X			X
Handelsstatistik import					X
Handelsstatistik export					
Bruttoinlandsprodukt			X		X
Arbeitslosenrate					X

Tabella 7: Gewählte Features je Verfahren für Unternehmen B - Niederlassung 2

### Gewählte Features für das Unternehmen B – Niederlassung 3:

Feature	FCBF	SFFS - ARIMAX	SFFS - RF	SFFS - SVR	LASSO
Transportvolumen		-		X	X
Rollierender Mittelwert Transportvolumen		-		X	X
Feiertage	X	X	X	X	X
Ferientage					
Wochentage		X	X		X
Zeit vor Weihnachten					X
Fourier Terme wöchentliche Saisonalität	X	X		X	X
Fourier Terme jährliche Saisonalität	X	X	X		X
Temperatur		X		X	X
Tagessumme Regen		X			X
ifo Geschäftsklimaindex					X
Handelsstatistik import					X
Handelsstatistik export		X			
Bruttoinlandsprodukt					X
Arbeitslosenrate					X

Tabella 8: Gewählte Features je Verfahren für Unternehmen B - Niederlassung 3

## 2.4 Evaluierung verschiedener Prognoseverfahrens für die betrachteten Zeitreihen

Für jedes Featureset und jede untersuchte Zeitreihe untersuchen wir nun die Prognosegenauigkeit der drei erwähnten Prognosealgorithmen.

### *Hyperparametertuning der Verfahren*

ARIMAX Modelle, Random Forests und Support Vector Regression erfordern die Festlegung sogenannter Hyperparameter. Die richtige Festlegung dieser Parameter ist entscheidend für die Prognosegenauigkeit der Algorithmen. Um die Parameter für RF und SVR herauszufinden führen wir zunächst eine grobe „grid search“ über die verschiedenen Parameter der Verfahren durch. In dem vielversprechendsten Gebiet der groben Suche verfeinern wir das Gitter und suchen erneut die besten Parameter. Dieses Parameterset nutzen wir dann zur Prognose. Die Parameter für die ARIMAX Modelle schätzen wir mithilfe der „auto.arima“-Funktion des „forecast“-Pakets in R. Dieses minimiert das „Akaike information criterion“ (AIC).

### *Berechnung der Prognosefehler*

Die Beurteilung erfolgt auf einem Testset, also einem Datensatz welchen wir nicht zur Feature-Selection genutzt haben. Darüber hinaus testen wir die Ergebnisse gegen ein univariates ARIMA Modell, welches mit der „auto.arima“ Funktion des „forecast“-Pakets in R geschätzt wurde. Zur Beurteilung von Prognosefehlern bei der Zeitreihenprognose sind verschiedene Maße gängig, jedes einzelne hat dabei Vor- und Nachteile. Wir beurteilen die Prognosegenauigkeit unserer Verfahren deshalb nach drei gängigen, sich ergänzenden Prognosefehlern: Dem Root Mean Squared Error (RMSE), um einen absoluten Prognosefehler darzustellen. Dem sMAPE (scaled Mean Absolute Percentage Error), um einen relativen und dennoch unverzerrten Prognosefehler darzustellen und dem MASE (Mean Absolute Scaled Error), um die Prognosegenauigkeit im Verhältnis zur naiven Prognose zu messen. Die naive Prognose nimmt den Wert der Vorperiode als Prognose für die nächste Periode. Für genauere Ausführungen zu den Maßzahlen zu Prognosefehlern empfehlen wir (Makridakis 1993) und (Hyndman und Koehler 2006). Darüber hinaus nutzten wir Time-Series-Cross-Validation zur Beurteilung des Prognosefehlers, wie von (Bergmeir und Benítez 2012) empfohlen.

### Ergebnisdarstellung

Tabelle 9 gibt eine Übersicht über die erzielten Prognosefehler je Feature-Selection und Prognose Verfahren. Die Werte sind jeweils gemittelt über alle Niederlassungen.

		sMAPE	MASE	RMSE
<b>FCBF</b>	<b>ARIMAX</b>	12.80%	56.21%	60,689
	<b>RF</b>	13.99%	63.74%	64,464
	<b>SVR</b>	13.60%	63.05%	67,182
	<b>Overall</b>	<b>13.46%</b>	<b>61.00%</b>	<b>64,112</b>
<b>LASSO</b>	<b>ARIMAX</b>	14.15%	56.67%	45,971
	<b>RF</b>	12.13%	52.09%	51,571
	<b>SVR</b>	13.16%	56.17%	57,130
	<b>Overall</b>	<b>13.02%</b>	<b>54.76%</b>	<b>52,256</b>
<b>SFFS mit RF</b>	<b>ARIMAX</b>	12.96%	56.71%	61,009
	<b>RF</b>	12.85%	58.16%	59,464
	<b>SVR</b>	15.77%	75.53%	87,528
	<b>Overall</b>	<b>13.86%</b>	<b>63.47%</b>	<b>69,334</b>
<b>SFFS mit SVR</b>	<b>ARIMAX</b>	12.79%	56.18%	60,157
	<b>RF</b>	12.56%	54.21%	53,028
	<b>SVR</b>	12.95%	54.92%	58,292
	<b>Overall</b>	<b>12.77%</b>	<b>55.11%</b>	<b>57,159</b>
<b>SFFS mit ARIMAX</b>	<b>ARIMAX</b>	<b>12.85%</b>	<b>56.15%</b>	<b>58,610</b>
<b>Univariat</b>	<b>ARIMA</b>	<b>14.59%</b>	<b>80.36%</b>	<b>106,251</b>

Tabelle 9: Übersicht über die Prognosefehler je Feature-Selection und Prognoseverfahren

## 2.5 Verwendete Softwarepakete

Die Datensammlung des Projekts programmierten wir in Python, die statistischen Analysen in R. Die Daten hielten wir in einer Postgres Datenbank. Wir nutzten eine Vielzahl an frei verfügbaren Softwarepaketen und möchten uns bei den Entwicklern dieser Pakete herzlich bedanken. Zur allgemeinen Datenmanipulation nutzten wir „data.table“ (Dowle und Srinivasan 2019), „dplyr“ (Wickham et al. 2019b) und „zoo“ (Zeileis et al. 2019), um speziell mit Zeitreihen umzugehen. Zum Umgang mit Datumswerten nutzten wir „lubridate“ (Spinu et al. 2018). Für die Anbindung der Datenbank an R nutzten wir „RPostgreSQL“ (Conway et al. 2017). Zur Parallelisierung extrem rechenintensiver Algorithmen verwendeten wir „doParallel“ (Corporation und Weston 2019) und „foreach“ (Microsoft und Weston 2019). Für die Breimann-Implementierung des Random Forests nutzten wir das Paket „randomForest“ (Fortran original by Leo Breiman et al. 2018), für die Support Vector Regression das Paket „e1071“ (Meyer et al. 2019) und für das ARIMA(X) Modell das „forecast“ Paket (Hyndman und Khandakar 2008). Die LASSO Feature-Selection implementierten wir mithilfe des Pakets „glmnet“ (Friedman et al. 2019). Fehlende Werte ersetzten wir mithilfe des Pakets „mice“ (van Buuren und Groothuis-Oudshoorn 2019). Das FCBF-Verfahren (Yu und Liu 2003) und das SFFS-Verfahren (Pudil et al. 1994) implementierten wir selbst auf Basis der jeweiligen Publikationen. Die Daten visualisierten wir mithilfe von „ggplot2“ (Wickham et al. 2019a). Um die Daten deskriptiv zu analysieren nutzten wir darüber hinaus „Summarytools“ (Comtois 2019). Prognosefehler berechneten wir mit dem „Metrics“-Paket (Hamner und Frasco 2018).

## 2.6 Interpretation der Ergebnisse

Über die verschiedenen Feature-Selection Verfahren und Zeitreihen hinweg betrachtet lassen sich Hinweise darauf ableiten, mittels welcher externer Variablen sich kurzfristige Frachtvolumenprognosen verbessern lassen. Wichtig sind dabei vor allem komplexe Kalendereffekte, wie wöchentliche und jährliche Saisonalitäten sowie Feiertage. Darüber hinaus scheint die Temperatur und Regen das Frachtvolumen zu beeinflussen. Nur vereinzelt werden ökonomische Daten als relevant betrachtet. Da Feature-Selection Verfahren zu Overfitting neigen kann es sich dabei auch um Artefakte handeln.

In der Prognosestudie weisen alle multivariaten Verfahren ähnliche Prognosegenauigkeiten auf. Es treten einzelne Ausreißer nach unten bei der Kombination bestimmter Feature-Selection und Prognoseverfahren. Alle multivariaten Verfahren erzielen eine höhere Prognosegenauigkeit als der univariate Benchmark mit „auto.arima“. Im Rahmen der Vorstudie können wir deshalb konkludieren, dass externe Daten die Prognosegenauigkeit für Frachtvolumen erhöhen.

Um die Ergebnisse weiter zu validieren und neue Erkenntnisse zur Frachtvolumenprognose zu generieren, ist weitere Forschung notwendig. Insbesondere ist die Skalierung der Verfahren auf viele Zeitreihen notwendig, um die Prognoseverfahren praktisch einzusetzen. Da die Zeitreihen allerdings sehr heterogen sind, müssen dazu Auto-ML Verfahren eingesetzt werden. Zudem sollte ein längerer Prognosehorizont betrachtet werden, nämlich mindestens über eine Woche, um praktischen Nutzen für die Allokation von Ressourcen zu gewährleisten. Methodisch können wir bestätigen, dass keine der untersuchten Feature-Selection Verfahren eindeutige Schlüsse über die Relevanz von Einflussgrößen zulässt – insbesondere, wenn sehr viele Einflussgrößen untersucht werden. Stabilere Ergebnisse lassen sich durch die Kombination von mehreren Feature-Selection Verfahren erzielen. Allerdings sollte auch in diesem Fall die Relevanz von Einflussgrößen im Wechselspiel mit Domänenexperten beurteilt werden.

### 3 Die wichtigsten Positionen des zahlenmäßigen Nachweises

Im Projekt fielen lediglich Personal- und Reisekosten an. Wesentliche Aufwände entstanden dabei durch die Literaturrecherche, Datensammlung und Aufbereitung. Des Weiteren entstanden wesentliche Aufwände durch die Implementierung und Auswertung der Feature-Selection und Prognosestudie.

### 4 Notwendigkeit und Angemessenheit der geleisteten Arbeit

Wie schon beim Aufzeigen des Wissensstandes angedeutet, wurde bisher nicht systematisch untersucht welche Unternehmensexternen Einflussgrößen das Transportvolumen beeinflussen. Die Untersuchung dieses Sachverhalts ist zeitaufwändig und erfordert hochqualifiziertes, kostenintensives Personal. Demgegenüber stand ein unsicheres Ergebnis des Projekts. Andererseits können die Ergebnisse des Forschungsprojekts unternehmensübergreifend Mehrwerte bringen. Entsprechend war öffentliche Forschungsförderung zur bearbeiteten Forschungsfrage notwendig.

Die entstandenen Arbeitsaufwände entsprechen im Wesentlichen denen in der Vorhabenbeschreibung. Insbesondere die Integration der externen Daten miteinander, sowie die Bestimmung relevanter Einflussfaktoren war wie erwartet sehr zeitaufwändig.

### 5 Voraussichtlicher Nutzen und Verwertbarkeit

Bessere Prognosen des Frachtvolumens sind erforderlich, um Leerfahrten zu reduzieren und den Unternehmen eine akkuratere Ressourcenallokation zu ermöglichen. Auf Basis der Ergebnisse gehen die Forscher davon aus, dass dies grundsätzlich möglich ist – eine entsprechende Integration von Prognose, Tourenplanung und Ressourcenallokation im nächsten Schritt vorausgesetzt.

Unmittelbare Verwendung finden die Ergebnisse im Rahmen des publizierten Whitepapers zum Thema Transportvolumenprognose und in folgenden Industrieprojekten. Insbesondere den beteiligten Partnern werden die Ergebnisse zur weiteren Verwendung zur Verfügung gestellt. Im Rahmen einer wissenschaftlichen Publikation werden die Ergebnisse anderen Forschern zur Verfügung gestellt.

Darüber hinaus ist die Anschlussverwertung im Rahmen eines folgenden Forschungsprojekts zu empfehlen. Bei diesem sollte der Fokus darauf liegen mithilfe der erarbeiteten Prognosen ein Verfahren zu entwickeln um Leerfahrten zu reduzieren. Dazu ist gegebenenfalls eine Erweiterung der Prognose um weitere Aggregationsebenen mit einhergehender Prüfung der bereits gesammelten externen Variablen sowie die entsprechende Anpassung der Verfahren zur Tourenplanung und Ressourcenallokation notwendig.

### 6 Innovationen bei anderen Stellen während der Durchführung des Vorhabens

Im wissenschaftlichen Diskurs wurden keine neuen Publikationen zu dem betreffenden Thema veröffentlicht. Bei den einzelnen Unternehmen gibt es vereinzelt Bestrebungen den Informationsfluss zwischen Verladern und StGV-Unternehmen zu verbessern. Dies verbessert voraussichtlich die Prognose des Transportvolumens und kann als Ergänzung zu dem datengetriebenen Ansatz des vorliegenden Forschungsprojekts betrachtet werden.

## 7 Öffentlichkeitswirksame Maßnahmen

Um die Forschungsergebnisse in die Unternehmen zu tragen wurden verschiedene öffentlichkeitswirksame Maßnahmen durchgeführt. Die Forscher präsentierten das Projekt auf der Transport Logistik Messe in München, sowie auf drei verschiedenen mFund Veranstaltungen. Darüber hinaus vermarkteten sie das Forschungsprojekt auf dem Deutschen Logistikkongress vom 23. bis 25. Oktober in Berlin und der Supply & Procure Conference am 19.02.2020. Überdies präsentierten wir das Projekt in verschiedenen Fachkreisen, wie beispielsweise der ERFA-Gruppe Sanitärgrößhandel.

Dem wissenschaftlichen Publikum wurde das Vorgehen und die Ergebnisse auf dem International Symposium of Forecasting in Thessaloniki und auf der Operations Research Konferenz in Dresden präsentiert. Eine wissenschaftliche Publikation ist außerdem in Arbeit.

Des Weiteren publizierten die Forscher zum Projektabschluss einen Artikel in der renommierten Fachzeitschrift „Lebensmittelzeitung“ und ein Whitepaper welches sich gezielt an Entscheider und Praktiker in den Straßengüterverkehrsunternehmen wendet. Dieses stellten wir auf dem Logistik Forum am 19.11.2019 in Nürnberg vor.

Bei verschiedenen persönlichen Terminen mit Verladern und Logistikunternehmen stellte das Fraunhofer SCS das Projekt vor und diskutierte mit den Ansprechpartnern im Unternehmen den Mehrwert.



## Literaturverzeichnis

- Aburto, Luis; Weber, Richard (2007): Improved supply chain management based on hybrid demand forecasts. In: *Applied Soft Computing* 7 (1), S. 136–144. DOI: 10.1016/j.asoc.2005.06.001.
- Alp Ertem, Mustafa (2011): Forecasting Day of Week Volume Fluctuations In the Intermodal Freight Transportation. In: Proceedings of the 2011 Industrial Engineering Research Conference.
- Arunraj, Nari Sivanandam; Ahrens, Diane (2016): Estimation of non-catastrophic weather impacts for retail industry. In: *Intl J of Retail & Distrib Mgt* 44 (7), S. 731–753. DOI: 10.1108/IJRDM-07-2015-0101.
- Bahng, Youngjin; Kincade, Doris H. (2012): The relationship between temperature and sales. In: *Intl J of Retail & Distrib Mgt* 40 (6), S. 410–426. DOI: 10.1108/09590551211230232.
- Bergmeir, Christoph; Benítez, José M. (2012): On the use of cross-validation for time series predictor evaluation. In: *Information Sciences* 191, S. 192–213. DOI: 10.1016/j.ins.2011.12.028.
- Bertrand, Jean-Louis; Brusset, Xavier; Fortin, Maxime (2015): Assessing and hedging the cost of unseasonal weather: Case of the apparel sector. In: *European Journal of Operational Research* 244 (1), S. 261–276. DOI: 10.1016/j.ejor.2015.01.012.
- Chandrashekar, Girish; Sahin, Ferat (2014): A survey on feature selection methods. In: *Computers & Electrical Engineering* 40 (1), S. 16–28. DOI: 10.1016/j.compeleceng.2013.11.024.
- Comtois, Dominic (2019): summarytools: Tools to Quickly and Neatly Summarize Data. Online verfügbar unter <https://CRAN.R-project.org/package=summarytools>.
- Conway, Joe; Eddelbuettel, Dirk; Nishiyama, Tomoaki; Prayaga, Sameer Kumar; Tiffin, Neil (2017): RPostgreSQL: R Interface to the 'PostgreSQL' Database System. Online verfügbar unter <https://CRAN.R-project.org/package=RPostgreSQL>.
- Corporation, Microsoft; Weston, Steve (2019): doParallel: Foreach Parallel Adaptor for the 'parallel' Package. Online verfügbar unter <https://CRAN.R-project.org/package=doParallel>.
- Dekker, Mark; van Donselaar, Karel; Ouwehand, Pim (2004): How to use aggregation and combined forecasting to improve seasonal demand forecasts. In: *International Journal of Production Economics* 90 (2), S. 151–167. DOI: 10.1016/j.ijpe.2004.02.004.
- Dowle, Matt; Srinivasan, Arun (2019): data.table: Extension of 'data.frame'. Online verfügbar unter <https://CRAN.R-project.org/package=data.table>.
- Fite, Jonathon T.; Don Taylor, G.; Usher, John S.; English, John R.; Roberts, John N. (2002): Forecasting freight demand using economic indices. In: *Int Jnl Phys Dist & Log Manage* 32 (4), S. 299–308. DOI: 10.1108/09600030210430660.
- Forkenbrock, David J. (2001): Comparison of external costs of rail and truck freight transportation. In: *Transportation Research Part A: Policy and Practice* 35 (4), S. 321–337. DOI: 10.1016/S0965-8564(99)00061-0.
- Fortran original by Leo Breiman; Cutler, Adele; R port by Andy Liaw; Wiener, Matthew (2018): randomForest: Breiman and Cutler's Random Forests for Classification and Regression. Online verfügbar unter <https://CRAN.R-project.org/package=randomForest>.
- Friedman, Jerome; Hastie, Trevor; Tibshirani, Rob; Narasimhan, Balasubramanian; Simon, Noah (2019): glmnet: Lasso and Elastic-Net Regularized Generalized Linear Models. Online verfügbar unter <https://CRAN.R-project.org/package=glmnet>.
- Garrido, Rodrigo A.; Mahmassani, Hani S. (2000): Forecasting freight transportation demand with the space–time multinomial probit model. In: *Transportation Research Part B: Methodological* 34 (5), S. 403–418. DOI: 10.1016/S0191-2615(99)00032-6.
- Godfrey, Gregory A.; Powell, Warren B. (2000): Adaptive estimation of daily demands with complex calendar effects for freight transportation. In: *Transportation Research Part B:*

*Methodological* 34 (6), S. 451–469. DOI: 10.1016/S0965-8564(99)00032-4.

Guo, Z. X.; Wong, W. K.; Li, Min (2013): A multivariate intelligent decision-making model for retail sales forecasting. In: *Decision Support Systems* 55 (1), S. 247–255. DOI: 10.1016/j.dss.2013.01.026.

Hamner, Ben; Frasco, Michael (2018): Metrics: Evaluation Metrics for Machine Learning. Online verfügbar unter <https://CRAN.R-project.org/package=Metrics>.

Hastie, Trevor; Tibshirani, Robert; Friedman, Jerome H. (2017): The elements of statistical learning: Data mining, inference, and prediction. Second edition, corrected at 12th printing 2017. New York, NY: Springer (Springer series in statistics).

Hyndman, Rob J.; Khandakar, Yeasmin (2008): Automatic time series forecasting: the forecast package for R. In: *Journal of Statistical Software* 26 (3), S. 1–22. Online verfügbar unter <http://www.jstatsoft.org/article/view/v027i03>.

Hyndman, Rob J.; Koehler, Anne B. (2006): Another look at measures of forecast accuracy. In: *International Journal of Forecasting* 22 (4), S. 679–688. DOI: 10.1016/j.ijforecast.2006.03.001.

James, Gareth; Witten, Daniela; Hastie, Trevor; Tibshirani, Robert (2013): An Introduction to Statistical Learning: With Applications in R. New York, NY: Springer (Springer Texts in Statistics, 103). Online verfügbar unter <http://dx.doi.org/10.1007/978-1-4614-7138-7>.

Kim, Jong-Kil; Pak, Ji-Yeong; Wang, Ying; Park, Sung-II; Yeo, Gi-Tae (2011): A Study on forecasting container volume of port using SD and ARIMA. In: *Journal of Navigation and Port Research* 35 (4), S. 343–349. DOI: 10.5394/KINPR.2011.35.4.343.

Lahiri, Kajal; Yao, Vincent Wenxiong (2006): Economic indicators for the US transportation sector. In: *Transportation Research Part A: Policy and Practice* 40 (10), S. 872–887. DOI: 10.1016/j.tra.2006.02.006.

Lahiri, Kajal; Yao, Wenxiong; Young, Peg (2004): Transportation and the Economy: Linkages at Business-Cycle Frequencies. In: *Transportation Research Record* 1864 (1), S. 103–111. DOI: 10.3141/1864-14.

Little, Roderick J.; Rubin, Donald B. (1987): Statistical analysis with missing data. New York: Wiley (Wiley series in probability and mathematical statistics).

Loughrey, John; Cunningham, Pádraig (2005): Overfitting in Wrapper-Based Feature Subset Selection: The Harder You Try the Worse it Gets. In: Max Bramer, Frans Coenen und Tony Allen (Hg.): Research and Development in Intelligent Systems XXI, Bd. 97. London: Springer London, S. 33–43.

Makridakis, Spyros (1993): Accuracy measures: theoretical and practical concerns. In: *International Journal of Forecasting* 9 (4), S. 527–529. DOI: 10.1016/0169-2070(93)90079-3.

Meyer, David; Dimitriadou, Evgenia; Hornik, Kurt; Weingessel, Andreas; Leisch, Friedrich (2019): e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. Online verfügbar unter <https://CRAN.R-project.org/package=e1071>.

Microsoft; Weston, Steve (2019): foreach: Provides Foreach Looping Construct. Online verfügbar unter <https://CRAN.R-project.org/package=foreach>.

Murray, Kyle B.; Di Muro, Fabrizio; Finn, Adam; Popkowski Leszczyc, Peter (2010): The effect of weather on consumer spending. In: *Journal of Retailing and Consumer Services* 17 (6), S. 512–520. DOI: 10.1016/j.jretconser.2010.08.006.

Osadchiy, Nikolay; Gaur, Vishal; Seshadri, Sridhar (2013): Sales Forecasting with Financial Indicators and Experts' Input. In: *Prod Oper Manag* 53 (1), n/a-n/a. DOI: 10.1111/poms.12022.

Pudil, P.; Novovičová, J.; Kittler, J. (1994): Floating search methods in feature selection. In: *Pattern Recognition Letters* 15 (11), S. 1119–1125. DOI: 10.1016/0167-8655(94)90127-9.

Ramanathan, Usha; Muyltermans, Luc (2010): Identifying demand factors for promotional

planning and forecasting: A case of a soft drink company in the UK. In: *International Journal of Production Economics* 128 (2), S. 538–545. DOI: 10.1016/j.ijpe.2010.07.007.

Saidi, Samir; Hammami, Sami (2017): Modeling the causal linkages between transport, economic growth and environmental degradation for 75 countries. In: *Transportation Research Part D: Transport and Environment* 53, S. 415–427. DOI: 10.1016/j.trd.2017.04.031.

Spinu, Vitalie; Grolemond, Garrett; Wickham, Hadley (2018): lubridate: Make Dealing with Dates a Little Easier. Online verfügbar unter <https://CRAN.R-project.org/package=lubridate>.

Syntetos, Aris A.; Babai, Zied; Boylan, John E.; Kolassa, Stephan; Nikolopoulos, Konstantinos (2016): Supply chain forecasting: Theory, practice, their gap and the future. In: *European Journal of Operational Research* 252 (1), S. 1–26. DOI: 10.1016/j.ejor.2015.11.010.

Taylor, James W. (2007): Forecasting daily supermarket sales using exponentially weighted quantile regression. In: *European Journal of Operational Research* 178 (1), S. 154–167. DOI: 10.1016/j.ejor.2006.02.006.

van Buuren, Stef; Groothuis-Oudshoorn, Karin (2011): mice: Multivariate Imputation by Chained Equations in R. In: *Journal of Statistical Software* 45 (1), S. 1–67. DOI: 10.18637/jss.v045.i03.

van Buuren, Stef; Groothuis-Oudshoorn, Karin (2019): mice: Multivariate Imputation by Chained Equations. Online verfügbar unter <https://CRAN.R-project.org/package=mice>.

van Donselaar, K. H.; Peters, J.; Jong, A. de; Broekmeulen, R.A.C.M. (2016): Analysis and forecasting of demand during promotions for perishable items. In: *International Journal of Production Economics* 172, S. 65–75. DOI: 10.1016/j.ijpe.2015.10.022.

Wang, Hui Li; Shi, Zhong Ke (2012): Hierarchical Regression Model for Analyzing Truck Freight Demand. In: *AMM* 178-181, S. 2752–2756. DOI: 10.4028/www.scientific.net/AMM.178-181.2752.

Wickham, Hadley; Chang, Winston; Henry, Lionel; Pedersen, Thomas Lin; Takahashi, Kohske; Wilke, Claus et al. (2019a): ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics. Online verfügbar unter <https://CRAN.R-project.org/package=ggplot2>.

Wickham, Hadley; François, Romain; Henry, Lionel; Müller, Kirill (2019b): dplyr: A Grammar of Data Manipulation. Online verfügbar unter <https://CRAN.R-project.org/package=dplyr>.

Yu, Lei; Liu, Huan (2003): Feature selection for high-dimensional data: A fast correlation-based filter solution. In: *Proceedings of the 20th international conference on machine learning (ICML-03)*, S. 856–863.

Zeileis, Achim; Grothendieck, Gabor; Ryan, Jeffrey A. (2019): zoo: S3 Infrastructure for Regular and Irregular Time Series (Z's Ordered Observations). Online verfügbar unter <https://CRAN.R-project.org/package=zoo>.

Zhou, Ling (2017): Prediction of a service demand using combined forecasting approach. In: *J. Phys.: Conf. Ser.* 887, S. 12075. DOI: 10.1088/1742-6596/887/1/012075.