

Fachwissenschaftliche Projektarbeit

Prognosemethoden und ihre Anwendbarkeit auf das Transportwesen

Technische Universität Dortmund

Fakultät Maschinenbau

Fachgebiet IT in Produktion und Logistik

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

Name: Kim Bianca Tscherner
Studiengang: Logistik, Master
Matrikelnummer: 158090

Name: Maurice Rudolph
Studiengang: Logistik, Master
Matrikelnummer: 134009

Ausgegeben am: 22.06.2017
Eingereicht am: 22.12.2017

Dortmund, im Dezember 2017

Inhaltsverzeichnis

| | |
|---|----|
| Abbildungsverzeichnis | 3 |
| Tabellenverzeichnis | 4 |
| 1 Einleitung | 5 |
| 2 Grundlage der Datenanalyse im Verkehrswesen | 7 |
| 2.1 Einführung in die Datenanalyse [MR] | 7 |
| 2.1.1 Datensammlung [MR] | 7 |
| 2.1.2 Datenaufbereitung [MR] | 15 |
| 2.1.3 Datenauswertung [MR] | 20 |
| 2.2 Methodische Grundlagen in Excel [MR] | 22 |
| 3 Prognoseverfahren im Verkehrswesen | 25 |
| 3.1 Überblick Verkehrsplanung [KT] | 25 |
| 3.2 Trendprognose [KT] | 26 |
| 3.2.1 Einfache Prognoseverfahren zur Zeitreihenanalyse [KT] | 26 |
| 3.2.2 Zeitreihendekomposition [KT] | 27 |
| 3.2.3 Eignung der Zeitreihenanalyse [KT] | 28 |
| 3.3 Verkehrsnachfragemodelle [KT] | 30 |
| 3.4 Angewendete Modelle [KT] | 32 |
| 4 Anwendung der Prognosemethoden auf ein Anwendungsbeispiel | 34 |
| 4.1 Beschreibung der Datengrundlage [MR] | 34 |
| 4.2 Auswahl der Prognosemethode [KT] | 35 |
| 4.3 Implementierung des Verfahrens in MS Excel [KT] | 36 |
| 4.4 Auswertung der Daten [MR] | 39 |
| 4.5 Fazit [KT] | 42 |
| 5 Zusammenfassung und Ausblick [KT] | 45 |
| Literaturverzeichnis | 47 |

Abbildungsverzeichnis

| | | |
|----------------|--|----|
| Abb. 1: | Übersicht der Befragungsverhalten | 14 |
| Abb. 2: | Übersicht Merkmale und Skalen | 16 |
| Abb. 3: | Entscheidungshilfe zur Auswahl eines Hochrechnungsverfahrens für Fahrzeugzählungen | 19 |
| Abb. 4: | Grundlegender Formelaufbau in MS Excel | 22 |
| Abb. 5: | Aufbau der Wenn-Formel in MS Excel | 23 |
| Abb. 6: | Beispielhafte Darstellung mehrerer Trendlinien | 38 |

Tabellenverzeichnis

| | | |
|-------------------|--|----|
| Tabelle 1: | Zuordnung der Kenngrößen zu den Erhebungsverfahren | 10 |
| Tabelle 2: | Merkmale der verschiedenen Hochrechnungsverfahren für Fahrzeugzählungen | 18 |
| Tabelle 3: | Zeitreihendekomposition Multiplikativ | 40 |
| Tabelle 4: | Zeitreihendekomposition Additiv | 41 |

1 Einleitung

Die Bedeutung des Verkehrswesens nimmt aktuell sowie in Zukunft stets zu. Aufgrund der Globalisierung und der wachsenden Vernetzung aller Akteure weltweit gilt dies auch für die Ansprüche und die damit verbundenen Anforderungen an das Transportwesen. Hinzu kommen Effekte wie der Güterstruktur- oder der Logistikeffekt, welche das Aufkommen auf allen Infrastrukturen erhöhen. Für die Planung und Organisation, sowohl von Infrastruktur als auch vom Verkehr allgemein, ist es wichtig zukünftige Entwicklungen abschätzen zu können. Dabei spielen sowohl langfristige als auch kurzfristige Vorhersagen eine entscheidende Rolle. Wird z.B. das Transportwesen betrachtet, so ist es für die wirtschaftlich beteiligten Akteure von Interesse die Entwicklungen auf vergleichsweise kürzere Zeit zu untersuchen, um so Unternehmenskennzahlen wie Größe des Fuhrparks oder auch Anzahl der Mitarbeiter an die Bedarfe anzupassen. Die Betreiber von Infrastrukturen hingegen, wie beispielsweise staatliche Institutionen, legen den Fokus bei Aspekten, wie der Erweiterung bestehender Infrastruktur oder dem Bau von neuen Netzen, auf die Untersuchung von Prognosen über mehrere Jahre oder Jahrzehnte hinaus. Alle Vorhersagen dienen dabei stets dem Zweck die Kosten für Aufwendungen zu minimieren und dabei gleichzeitig den Nutzen für die eigene Unternehmung zu maximieren.

Ziel dieser Arbeit ist die Vorstellung und Ausarbeitung einer ausgewählten Prognosemethode. Nach grundlegenden Erläuterungen zu Datengrundlagen sowie Prognoseverfahren wird die präferierte Methode in das Tabellenkalkulationsprogramm MS Excel implementiert. Mit dem so erstellten Tool werden real existierende Datensätze der Straßen-güterverkehrsstatistik untersucht. Es soll dabei ermittelt werden, wie genau die mit dem Programm prognostizierten Werte im Vergleich zu den realen Werten sind. Dadurch soll letztlich eine Aussage über die Eignung des gewählten Verfahrens in Bezug auf das Transportwesen getroffen werden.

Die Vorgehensweise der Arbeit baut sich zunächst auf den Grundlagen zu der Beschaffung von Daten auf, nimmt dann Bezug zu Prognoseverfahren im Allgemeinen und bringt im Anschluss daran beide Aspekte mit dem MS Excel Tool und den verwendeten Daten zusammen. Im Kapitel zu den Grundlagen der Datenanalyse im Verkehrswesen werden die Aspekte der Datenbeschaffung, der Datenaufbereitung sowie der Datenauswertung näher dargestellt. Bei allen drei Kategorien werden zunächst Grundlagen aus der Statistik und Marktforschung beschrieben, um im Anschluss die Gegebenheiten des Verkehrswesens näher vorzustellen.

Um einen generellen Einblick in die Funktion des für das Prognosetool verwendeten Tabellenkalkulationsprogrammes MS Excel zu bekommen, schließt sich nach den Daten Grundlagen eine Vorstellung des Programms sowie der grundlegenden und im Tool verwendeten Funktionen an. Das darauf folgende Kapitel gibt zunächst einen Überblick über die allgemeine Verkehrsplanung und die dort übliche Vorgehensweise in Bezug auf Prognoseverfahren. Zudem werden in diesem Kapitel die Trendprognose (wie die Zeitreihenanalyse und die Zeitreihendekomposition) und vorhandene Verkehrsmodelle (wie Verkehrsflussmodelle oder Verkehrsnachfragemodelle) vorgestellt. Im letzten Kapitel erfolgt dann die Anwendung des ausgewählten Prognoseverfahrens mithilfe von MS Excel auf die ausgewählten Datensätze. Hierbei wird neben der Erläuterung der genutzten Daten Grundlage auch das ausgewählte Prognoseverfahren vorgestellt und die anschließend erfolgte Implementierung in MS Excel. Danach erfolgt die Auswertung der Daten durch das Tool, um am Ende ein Fazit treffen zu können hinsichtlich der Genauigkeit und die damit verbundene Nutzbarkeit bezogen auf den vorliegenden Sachverhalt des Transportwesens. Das Ende der Arbeit bildet eine Zusammenfassung mit einem anschließenden Ausblick.

2 Grundlage der Datenanalyse im Verkehrswesen

2.1 Einführung in die Datenanalyse [MR]

Um Prognoseverfahren jeglicher Art nutzen und umsetzen zu können, bedarf es einer Datengrundlage. Der folgende Teil dieser Arbeit widmet sich zum einen dem Erzeugen einer solchen Datengrundlage und zum anderen der Auswertung dieser. Hierbei werden zunächst grundsätzliche Aspekte betrachtet, dann jedoch gezielt der Bereich des Verkehrswesens.

2.1.1 Datensammlung [MR]

Die aktuell stattfindende vierte industrielle Revolution bringt viele neue Aspekte mit sich. Einer der drei großen Hauptaspekte, neben dem cyber-physischen System und dem Internet der Dinge, lautet Big Data. Unter diesem Begriff versteht man „große Mengen an Daten (...) die mit speziellen Lösungen gespeichert, verarbeitet und ausgewertet werden“ (Gabler Wirtschaftslexikon 2017a). Die aus diesen Daten gewonnenen Informationen dienen anschließend unter anderem der Optimierung von Produktionen oder der Prognose in verschiedensten Bereichen.

Somit stellen sich anfangs zwei grundlegende Fragen. Die erste richtet sich nach dem Grund, also warum generell Daten erhoben werden müssen. Daten werden aus zwei Gründen erhoben. Entweder liegen zu einem Sachverhalt nicht genügend Informationen vor, um ihn ausreichend beurteilen zu können, oder aber es liegen gar keine Informationen vor. In beiden Fällen ist man somit gezwungen neue Informationen zu gewinnen durch Prüfung vorhandener Daten oder Neugewinnung eben jener (Schwarzbauer 2017, S. 1f.). Die zweite Frage bezieht sich auf die Methodik, also wie genau Daten erhoben werden. Der Bereich des Data Mining befasst sich beispielsweise mit dieser Frage, und darauf aufbauend wie die Datenmengen schnell und effizient ausgewertet werden können. In diesem Abschnitt geht es jedoch um die grundlegenden Möglichkeiten Daten zu erheben, welche Formen von Daten es gibt und welche Daten für das Verkehrswesen relevant sind.

Angefangen werden soll im Folgenden mit einer Beschreibung und Definition des Begriffes „Daten“. Laut deutscher Industrienorm 44300 sind Daten „Gebilde aus Zeichen oder kontinuierliche Funktionen, die aufgrund bekannter oder unterstellter Abmachungen Informationen darstellen, vorrangig zum Zwecke der Verarbeitung und als deren Ergebnis“ (Dierstein 2003).

Diese Definition zeigt bereits, dass Daten zu Informationen führen, und umgekehrt. Die folgende Definition des Begriffes „Information“ soll den Unterschied zwischen den beiden Begriffen Information und Daten verdeutlichen. Diese „sind zweckbezogenes Wissen und kennzeichnen die Datenmenge, die zur Lösung der definierten Probleme notwendig ist“ (Pepels und Bruns 2008, S.53). Grundsätzlich lässt sich jedoch keine einheitliche Definition für den Begriff „Daten“ festlegen. Je nach Untersuchungs- und Fachgebiet weichen diese voneinander ab oder unterscheiden sich gänzlich.

Um Daten generell zu erheben bieten sich viele Möglichkeiten. Generell gibt es zwei Varianten, die es zu unterscheiden gilt. Zum einen die Primärerhebung. Durch eine Primärerhebung können neue Daten gewonnen werden, welche aktuell in der entsprechenden Form oder Aktualität nicht vorliegen (vgl. Jahn et al. 2004, S.24). Die zweite Methode ist die der Sekundärerhebung. Diese „hat die Gewinnung von Informationen aus bereits vorliegenden Daten (...) zum Ziel.“ (Jahn et al. 2004, ebenda) Der Bezug zum Verkehrswesen lässt sich hier bereits gut darstellen. Für Verkehrserhebungen werden beispielsweise beide Verfahren benutzt. Primärstatistische Ansätze sind hier die Befragung von KFZ-Haltern, oder schlichtweg die Beobachtung des Verkehrsgeschehens und damit einhergehend die Zählung der Fahrzeuge. Im Bereich der Sekundärstatistik werden die Daten des Mautsystems, sowie die der KFZ-Hauptuntersuchung genutzt (Hautzinger et al., S. 12 f.).

Da sich die Sekundärerhebung mit dem Auswerten bereits vorhandener Daten befasst, wird im Folgenden der Fokus auf die Primärerhebung gelegt. Anfangs muss zwischen Deskriptiver Statistik und Inferenzstatistik unterschieden werden. Bei der Deskriptiven Statistik wird die Stichprobe auf ihre Eigenschaften hin untersucht um einen Überblick über die „Merkmalsausprägungen einzelner Variablen“ (Raab-Steiner und Benesch 2012, S.15) zu bekommen. Die Inferenzstatistik hingegen befasst sich mit dem Schluss von der Stichprobe auf die Grundgesamtheit. Die aus der Stichprobe gewonnen Erkenntnisse sollen somit repräsentativ die Allgemeinheit abbilden (vgl. Raab-Steiner und Benesch 2012, S.17 f.). Um Verkehrsdaten zu erheben, kann man entweder die Gesamtheit aller dafür infrage kommenden Objekte betrachten, oder man betrachtet lediglich eine Stichprobe. Untersucht man die Fahrleistung eines Unternehmens mit 50 Fahrzeugen ist eine Vollerhebung sinnvoll. Untersucht man jedoch die Fahrleistung innerhalb Deutschlands über einen Zeitraum von mehreren Monaten, ist es unmöglich alle Daten genau zu erfassen. Somit werden hier Stichproben betrachtet um Aussagen über die Gesamtheit treffen zu können. Die Güte der Stichprobe, also inwieweit sie Aussagekraft über die Gesamtheit hat, hängt dabei von ihrer Auswahl und Größe ab.

Beide sind Faktoren für den Sicherheitsgrad, mit dem die Genauigkeit in Bezug auf die Grundgesamtheit ausgedrückt wird (vgl. Pepels und Bruns 2008, S.97 f.). Der Umfang der Betrachtung bezogen auf eine Verkehrszählung wird im weiteren Verlauf noch erläutert.

Neben der Sekundärerhebung von Daten gibt es die bereits erwähnte primäre Erhebung. In der Marktforschung liegt der Fokus hierbei beispielsweise auf zwei Varianten. Es handelt sich dabei um die Befragung und um die Beobachtung. In der Literatur gibt es an manchen Stellen noch einen dritten Aspekt, das Experiment (vgl. Pepels und Bruns 2008, S.122). Da hier jedoch am Ende ebenfalls lediglich beobachtet und gemessen wird, beschränkt sich die Erläuterung im Folgenden auf die zwei anfangs erwähnten Möglichkeiten.

Die Befragung kann über verschiedene Wege erfolgen, beispielsweise telefonisch, schriftlich oder auch im Interview (vgl. Jahn et al. 2004, S.47). In allen Varianten wird ein Fragebogen ausgefüllt. Der Fragebogen ist eine quantitative Methode der Datenerhebung. „Als quantitative Methoden werden alle Vorgehensweisen, die zur numerischen Darstellung empirischer Sachverhalte dienen, verstanden, somit auch Erhebungen mittels standardisierter Fragebogen“ (Raab-Steiner und Benesch 2012, S.45). Die Gestaltung von Fragebogen ist komplex und hat keine standardisierte Anleitung. Grundlegend ist zu erwähnen, dass Aufbau und Auswahl der Fragen ausschlaggebend für die letztendliche Aussagekraft sind. Ebenso gibt es unterschiedliche Fragekategorien, wie beispielsweise offene und geschlossene Fragen, Alternativfragen, Dialogfragen, Skalenfragen usw. (vgl. Pepels und Bruns 2008, S.76). Die Komplexität jener Gestaltung und der Fragebogen an sich soll hier jedoch nicht Untersuchungsgegenstand sein. Im Verkehrswesen werden ebenfalls Befragungen von KFZ-Haltern durchgeführt um daraus Informationen über deren Verhalten zu gewinnen. Hierbei kann man vergleichsweise zwischen einer Erhebung tatsächlicher Ereignisse (Revealed Choice) und einem experimentellem Design (Stated Choice) unterscheiden. Bei der Revealed Choice „werden tatsächlich realisierte Wahlentscheidungen wie Wegekettwahl, Verkehrsmittelwahl und Routenwahl und die relevanten erklärenden Variablen wie Alter, Geschlecht, Zeit- und Geldaufwände für die verschiedenen Verkehrsmittel (...) erhoben“ (Martin Treiber). Dadurch erhält man Informationen die durch eine bloße Zählung nicht erreichbar wären. Bei der Stated Choice werden hingegen eher “Was wäre wenn?“ Fragen gestellt.

Die zweite Variante der Primärerhebung ist die Beobachtung/Messung. Generell gibt es bei jeder Messung zwei Hauptgütekriterien. Diese sind Reliabilität und Validität. Reliabilität bedeutet Verlässlichkeit, Validität Gültigkeit.

Beide „sind Basis jedes Untersuchungsansatzes“ (Jahn et al. 2004, S.65). Ein Wert gilt dann als verlässlich, wenn er unter gegebenen Bedingungen reproduzierbar ist. Gültigkeit liegt dann vor, wenn das Ergebnis der Messung bzw. Betrachtung „den Sachverhalt, der ermittelt und interpretiert werden soll, tatsächlich wieder gibt“ (Jahn et al. 2004, ebenda). Die Messung als solches kann in vielen Arten erfolgen. Die gewählte Methode hängt dabei stets vom Sachverhalt, den Umständen und den zur Verfügung stehenden Mitteln ab. Bezieht man sich auf den Verkehr und sollen Daten zum Verkehrsaufkommen oder Verkehrsflusszahlen ermittelt werden, gibt es auch hier verschiedene Möglichkeiten, die man nutzen kann. Die am meisten verwendeten sind hierbei Dauerzählstellen, manuelle Zählungen oder die Auswertung von Videoaufzeichnungen.

Im Verkehrswesen gibt es nicht zwei, sondern vier Möglichkeiten, die im Fokus der Erhebung primärer Daten stehen. Diese sind die Beobachtung, die Messung, die Zählung und die Befragung. Allen Erhebungen liegt die Tatsache zugrunde, dass Ortsveränderungen von Personen oder Fahrzeugen erfasst werden sollen. Bei einer Verkehrsuntersuchung handelt es sich „also um „Inzidenzstudie(n)“, d.h. um eine zeitraumbezogene Erfassung von Ereignissen (Ortsveränderungen) und deren Eigenschaften“ (Bäumer und Hautzinger 2010, S. 18). Der Begriff der Inzidenzstudien stammt aus der Epidemiologie und beschreibt die Zahl von Neuerkrankungen innerhalb eines bestimmten Zeitraumes (Bäumer und Hautzinger 2010, S. 18). Schließlich gibt es noch qualitative Erhebungsverfahren. Die daraus gewonnenen Daten sind jedoch für diese Arbeit nicht von Relevanz, weshalb sie nicht weiter thematisiert werden.

| Kenngrößen | | Erhebungsmethode | | | |
|---|---|------------------|---------|-------------|-----------|
| | | Zählung | Messung | Beobachtung | Befragung |
| Verhaltensmerkmale | z.B. Verkehrsmittelwahl | | | | x |
| Aggregierte Verkehrskenngrößen | Verkehrsaufkommen | x | | | x |
| | Verkehrsleistung | | | | x |
| Mikroskopische Kenngrößen des Verkehrsablaufs | Zeitlücke Abstand Geschwindigkeit | | x | | |
| Makroskopische Kenngrößen des Verkehrsablaufs | Verkehrsstärke | x | | | |
| | Verkehrsdichte | o | o | x | |
| | Geschwindigkeit | o | x | | |
| Einstellungen/ Präferenzen | z.B. bevorzugte Verkehrsmittel | | | | x |
| sichtbare Verhaltensweisen | z.B. Rotläufer | | | x | |

Tabelle 1: Zuordnung der Kenngrößen zu den Erhebungsverfahren (Forschungsgesellschaft für Straßen- und Verkehrswesen 2012, S.11)

Die Tabelle zeigt den Zusammenhang der in der Regel verwendeten Kenngrößen und den vier vorgestellten Erhebungsmethoden. Eine Kennzeichnung mit einem X gibt an, dass die Kenngröße mit der ausgewählten Methode komplett erfasst werden kann. Ein O stellt eine eingeschränkte Erfassung dar. Ist eine Zelle gar nicht gefüllt, so kann der gesuchte Wert mit jener Methode nicht ermittelt werden (vgl. Forschungsgesellschaft für Straßen- und Verkehrswesen 2012, S.11).

Die erste der vier genannten Varianten ist die Beobachtung. „Beobachtungen dienen der planmäßigen Erfassung äußerer Merkmale und sichtbarer Verhaltensweisen im Straßenraum“ (Forschungsgesellschaft für Straßen- und Verkehrswesen 2012, S.56). Bei Verkehrsbeobachtungen steht weniger die Protokollierung der Anzahl von Merkmalsträgern im Fokus, sondern die Art und Weise wie jene agieren, also wann und wie sie sich fortbewegen. Anwendungsgebiete von solchen Untersuchungen liegen z.B. in der Verkehrssicherheit, wenn beispielsweise das Verhalten von Verkehrsteilnehmern an kritischen Punkten wie Kreuzungen analysiert werden soll, oder auch in der Analyse von komplexen Abläufen und Wechselwirkungen zwischen den Verkehrsteilnehmern (vgl. Forschungsgesellschaft für Straßen- und Verkehrswesen 2012, ebenda). Geläufige Formen der Verkehrsbeobachtung sind beispielsweise Verkehrskameras in Städten, die Verkehrsfluss überwachen und protokollieren, oder auch Luftaufnahmen via Helikopter um Staus auf Bundesautobahnen zu lokalisieren.

Bei der zweiten der vier genannten Erhebungsformen handelt es sich um die Messung. „Mit Messungen werden Merkmale erfasst, die stetige Werte (z.B. Geschwindigkeiten) annehmen können“ (Forschungsgesellschaft für Straßen- und Verkehrswesen 2012, S.49).

Das bedeutet, dass im Gegensatz zur Beobachtung und Zählung nicht nur ganze Werte erfasst werden können. Messgeräte sind von zentraler Bedeutung bei Messungen. Wichtig hierbei ist, dass jene von den Verkehrsteilnehmern nicht als solche wahrgenommen werden, da die sonst daraus resultierenden Reaktionen die Messung verfälschen würden (vgl. Forschungsgesellschaft für Straßen- und Verkehrswesen 2012, ebenda). Einsatzmöglichkeiten liegen erneut in der Verkehrssicherheit, wie Geschwindigkeitskontrollen, Messungen von Wartezeiten an Knotenpunkten, oder auch die Ermittlung von „Reisezeiten und -geschwindigkeiten für Streckenabschnitte bzw. Routen“ (Forschungsgesellschaft für Straßen- und Verkehrswesen 2012, ebenda). Innerhalb der Messungen kann zwischen punktuellen Messungen und linienhaften Messungen unterschieden werden.

Während bei punktuellen Messungen lediglich an einem Ort beispielsweise eine Geschwindigkeit gemessen wird, kann bei linienhaften Messungen die Geschwindigkeit über einen Streckenabschnitt ermittelt werden.

Die Zählung ist neben der Befragung eines der wichtigsten Erhebungsverfahren im Verkehrswesen. Mithilfe der Zählung kann, wie in Tabelle 1 zu sehen, das Verkehrsaufkommen und die Verkehrsstärke erfasst werden, letztere auch einzig durch Zählung. „Mit Zählungen werden quantitativ erfassbare Ortsveränderungen von Personen und/oder Fahrzeugen auf den Verkehrswegen eines Planungsraumes erfasst“ (Forschungsgesellschaft für Straßen- und Verkehrswesen 2012, S.25). Es gibt vier Hauptarten von Zählungen, die sich durch die zu erfassenden Merkmalsträger sowie den Zustand der Fortbewegung unterscheiden. Bei den Varianten handelt es sich um Zählungen von: Fußgängern und Radfahrern, Personen im ÖPNV, Fahrzeuge im ruhenden Verkehr sowie dem generellen Kraftfahrzeugverkehr (Forschungsgesellschaft für Straßen- und Verkehrswesen 2012, ebenda).

Bei Zählungen von Fußgängern und Radfahrern müssen alle Verkehrsteilnehmer innerhalb der Betrachtung erfasst werden. Es kann hierbei nicht, wie bei Fahrzeugen, nach Gewicht oder Typ unterschieden werden. Das bei dieser Zählung untersuchte Aufkommen ist sehr stark von äußeren Umständen, wie der Witterung abhängig. Dadurch kann es zu saisonalen, aber auch zu abrupten Schwankungen kommen (vgl. Forschungsgesellschaft für Straßen- und Verkehrswesen 2012, S.29). Die ermittelten Daten dienen der Dimensionierung von Rad- sowie Fußwegen, der Beurteilung der aktuellen Gegebenheiten oder auch einer Trendanalyse für Prognosen.

Die Zählung von Personen im öffentlichen Personennahverkehr kann grundsätzlich an zwei verschiedenen Orten, den Haltestellen oder innerhalb der Fahrzeuge, erfolgen. Oftmals erfolgt neben der Zählung simultan eine Befragung zum Nutzungsverhalten der Fahrgäste. Die Zählungen können automatisch oder manuell erfolgen. Ermittelt werden sollen Auslastungsgrade der Haltestellen sowie der Fahrzeuge. Dadurch lassen sich Informationen über das Nutzungsverhalten der Kunden gewinnen, sowie über die Rentabilität von Linien oder auch Nutzungszeiten (vgl. Forschungsgesellschaft für Straßen- und Verkehrswesen 2012, S.30 f.).

Fahrzeuge im ruhenden, oder auch parkenden, Verkehr zu zählen dient den verantwortlichen Betreibern vor allem zur Ermittlung von Kennwerten wie Parkdauer, Umschlaggrad sowie der Auslastung der Parkanlage.

Zusätzlich können noch Unterscheidungen hinsichtlich der Art des Parktickets (Dauerparker, Kurzzeitparker, Anwohner etc.) oder auch des genutzten Parkplatzes (Frauenparkplatz, Behindertenparkplatz etc.) getroffen werden. Das Resultat der Zählung hängt dabei stark von der Zeit ab. Tag, Uhrzeit und äußere Umstände (z.B. umliegende Veranstaltungen) haben großen Einfluss auf das Ergebnis. Daher muss die Aufgabenstellung vorab klar definiert sein um Verfälschungen zu vermeiden (vgl. Forschungsgesellschaft für Straßen- und Verkehrswesen 2012, S.40 f.).

Zählungen des Kraftfahrzeugverkehrs liefern als einziges Verfahren Daten zur Verkehrsstärke, wie man Tabelle 1 entnehmen kann. Diese Daten sind Grundlage für die Planung von Straßenverkehrsanlagen, für betriebs- und volkswirtschaftliche Verkehrsuntersuchungen, von Untersuchungen von Entwicklungen innerhalb des Verkehrs und für viele weitere Analysen (vgl. Forschungsgesellschaft für Straßen- und Verkehrswesen 2012, S.25). Innerhalb der Zählung des Kraftfahrzeugverkehrs wird zwischen drei Methoden der Zählung unterschieden. Dabei handelt es sich um die Zählung an einem Querschnitt, die Zählung an einem Knotenpunkt sowie die Zählung innerhalb des Netzes (vgl. Forschungsgesellschaft für Straßen- und Verkehrswesen 2012, S.33). Die Zählung an einem Querschnitt ist dabei von größerer Relevanz, da sich mit ihr der durchschnittliche tägliche Verkehr (DTV) erfassen lässt. Diese Kennzahl bildet die Grundlage für viele Untersuchungen und Berechnungen innerhalb des Verkehrswesens und ist vor allem für die Planung und Dimensionierung von Verkehrsanlagen von Bedeutung (vgl. Forschungsgesellschaft für Straßen- und Verkehrswesen 2012, S.34). Die erforderlichen Daten zur Anzahl der Fahrzeuge, Fahrzeugtyp, Fahrtrichtung und Fahrstreifenwahl werden dabei sowohl automatisch als auch manuell erhoben.

Beide Varianten sind notwendig, um im Zuge der Datenaufbereitung Hochrechnungsverfahren durchführen zu können, aus denen sich später belastbare Zahlen zum jährlichen Verkehr ableiten lassen. Neben der Unterscheidung nach dem Ort der Erfassung gibt es die bereits erwähnte Differenzierung nach Fahrzeugklassen.

Laut der Forschungsgesellschaft für Straßen- und Verkehrswesen wird in der Empfehlung für Verkehrserhebungen eine Einteilung in sieben Fahrzeugklassen für Straßenverkehrszählungen vorgenommen (vgl. Forschungsgesellschaft für Straßen- und Verkehrswesen 2012, S.33). Das Bundesministerium für Verkehr und digitale Infrastruktur legt hingegen eine Unterteilung in sechs Unterkategorien fest, die sich anschließend zu Leicht- und Schwerverkehr zusammenfassen lassen (Bundesministerium für Verkehr und digitale Infrastruktur 2015, S. 8 f.).

In der Straßengüterverkehrsstatistik des statistischen Bundesamtes hingegen werden nur zwei Fahrzeugklassen nach dem Kriterium „zulässiges Gesamtgewicht“ unterschieden (Destatis 2017a). Die Einteilungen können sich somit leicht bis sehr stark voneinander unterscheiden und müssen im Vorfeld klar definiert werden. Neben jener Bestimmung der Fahrzeugklassen werden vom Bundesamt für Verkehr und digitale Infrastruktur beispielsweise ebenso klare Vorgaben erlassen, wann, wie und wo Daten erhoben werden. Die Zeit in der eine Zählung erfolgt, unterscheidet sich dabei nicht nur von der Stärke des durchschnittlichen täglichen Verkehrs, welcher eine Unterscheidung von Straßen über und unter einem Wert von 7.000 Kfz/24h vorsieht, sondern auch in welchem Bundesland die Zählung erfolgt (Bundesministerium für Verkehr und digitale Infrastruktur 2015).

Die vierte und letzte Möglichkeit der primären Datenerhebung im Straßenverkehr ist die Befragung. Analog zu den bereits vorgestellten Verfahren lässt sich dieses ebenfalls in verschiedene Aspekte unterteilen. Die Unterscheidung der Befragung erfolgt nach dem Aspekt der Lokalität, welcher in der folgenden Abbildung dargestellt wird.

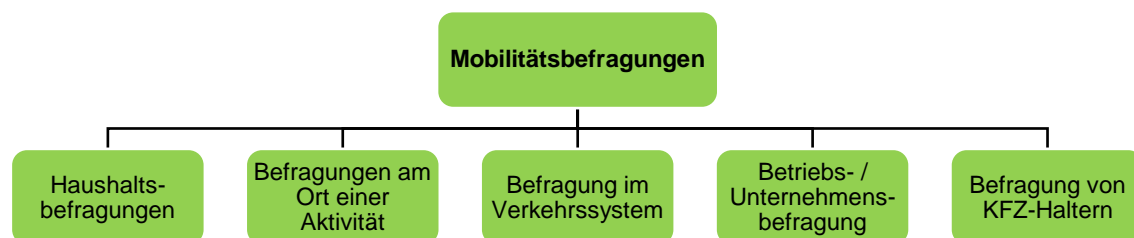


Abb. 1: Übersicht der Befragungsverhalten (Forschungsgesellschaft für Straßen- und Verkehrswesen 2012, S.63)

Für die Befragung an sich gilt es im Vorfeld die Parameter der Art und Weise sowie des Inhaltes zu definieren. Erhebungen können dabei schriftlich, telefonisch, persönlich oder auch via Internet durchgeführt werden (vgl. Forschungsgesellschaft für Straßen- und Verkehrswesen 2012, S.65). Die generelle Komplexität von Fragebögen wurde anfangs dieses Kapitels bereits erläutert und lässt sich ebenso auf Verkehrsbefragungen übertragen. Von den fünf in Abb.1 aufgezeigten Varianten der Befragung ist die Betriebs- bzw. Unternehmensbefragung von besonderem Stellenwert, da diese Aufschluss über den Wirtschaftsverkehr liefert. Die daraus gewonnenen Daten geben Aufschluss über die zurückgelegten Strecken und die dabei transportierten Waren (vgl. Forschungsgesellschaft für Straßen- und Verkehrswesen 2012, S.83).

Das Produkt dieser Faktoren bildet die Beförderungsleistung in Tonnenkilometern. Diese Werte sind Hauptinhalt der Straßengüterverkehrsstatistik (Destatis 2017b).

2.1.2 Datenaufbereitung [MR]

Im Anschluss an jede Datenerhebung folgt die Datenaufbereitung. Je nach Umfang der gewonnenen Daten kann diese mehr oder weniger umfangreich ausfallen. Die Definition unterscheidet sich dabei teils ebenso stark. „Datenaufbereitung ist die Säuberung der Daten von Tipp- oder Formatierungsfehlern“ (Onpulsion Wirtschaftslexikon 2017). Dieser Aspekt gibt jedoch nur einen Teil des Gesamten wieder. Neben der Feststellung von fehlerhaften Datensätzen, gibt es weitere Faktoren wie die Kontrolle und Ordnung der Daten, die Kodierung von Merkmalsausprägungen oder die generelle Datenanalyse (Gabler Wirtschaftslexikon 2017b).

In Kapitel 2.1.1 wurde bereits auf den Unterschied zwischen deskriptiver und Inferenzstatistik hingewiesen, welcher hier erneut getroffen wird. Mithilfe von deskriptivstatistischen Methoden sollen die vorhandenen Informationen auf das Wesentliche reduziert werden. Neben der grafischen Auswertung und Darstellung der Daten wird hier der Bezug zu statistischen Kenngrößen hergestellt (vgl. Raab-Steiner und Benesch 2012, S.109). Beispiele sind der Mittelwert, der Median, die Varianz, die Standardabweichung oder der Modus (auch Modalwert), welcher „der am häufigsten auftretende Wert in einer Stichprobe“ (Raab-Steiner und Benesch 2012, S.102) ist. In der Inferenzstatistik wird hingegen von den vorliegenden Daten der Stichprobe versucht auf die Gesamtheit Rückschlüsse treffen zu können. Dies wird mithilfe von Hypothesen bewerkstelligt. Unterschieden wird dabei zwischen der Nullhypothese und der Forschungshypothese. Erstere postuliert, dass die erhobenen Daten und Variablen in keinem Zusammenhang stehen. Letztere hingegen sagt das Gegenteil aus (vgl. Raab-Steiner und Benesch 2012, S.111).

Um feststellen zu können, welche der Hypothesen zutrifft, werden statistische Tests vorgenommen, um Rückschlüsse auf die Allgemeinheit zulassen zu können. Beispiele sind die Erhebung neuer Daten mittels des mehrfach erwähnten Fragebogen, oder der Datenaufbereitung durch deskriptive Methoden (vgl. Raab-Steiner und Benesch 2012, S. 113).

Spätestens im Zuge der Datenaufbereitung sollte eine Unterscheidung der erhobenen Merkmale innerhalb des Datensatzes erfolgen. Die folgende Abbildung zeigt eine Kategorisierung jener.

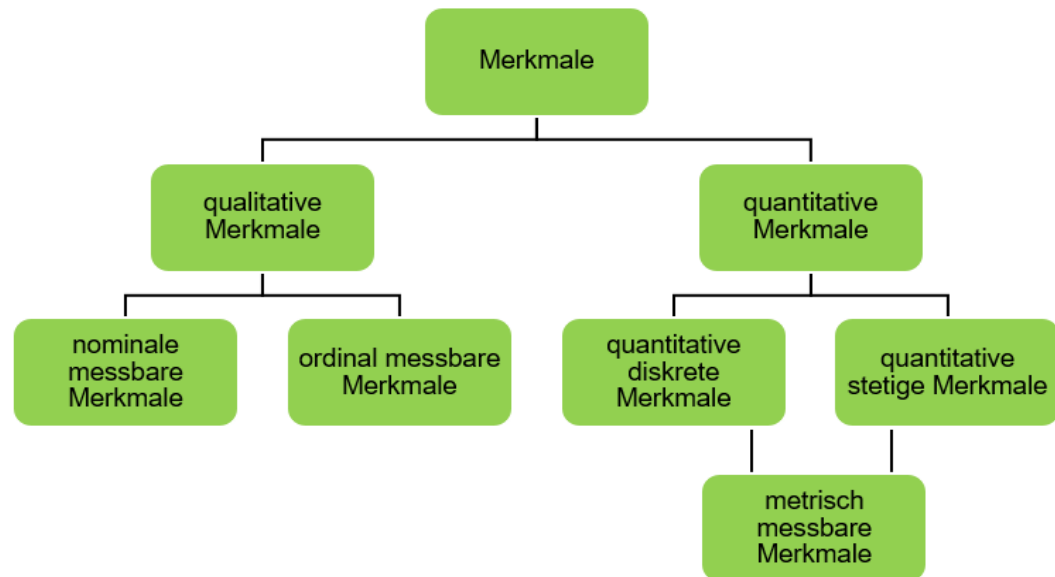


Abb. 2: Übersicht Merkmale und Skalen (Pepels und Bruns 2008, S.416)

Grundsätzlich lassen sich qualitative und quantitative Merkmale voneinander unterscheiden. Qualitative Merkmale werden nicht durch Zahlen dargestellt. Sie haben somit einen absoluten Zustand und können nicht in Rangfolgen gebracht werden (vgl. Pepels und Bruns 2008, S.414). Eine Skala, welche lediglich die Ausprägungen sowie deren Unterschiede aufzeigt wird als Nominalskala bezeichnet (Gabler Wirtschaftslexikon 2017d). Die qualitativen Merkmale werden letztlich noch in nominal und ordinal messbar unterschieden. Während bei nominal messbaren Merkmalen die Ausprägungen gleichberechtigt nebeneinander stehen (z.B. Farben oder Geschlecht) können bei ordinal messbaren Merkmalen durchaus Rangfolgen gebildet werden (z.B. Schulnoten oder Dienstgrade). Quantitative Merkmale nehmen reelle Zahlenwerte an und können somit stets in Beziehung gesetzt, verglichen oder geordnet werden.

Diese Skalen, welche jene Werte wiedergeben, werden als Kardinalskalen bezeichnet. (vgl. Pepels und Bruns 2008, S.415). Sie sind generell metrisch messbar, jedoch können sie diskret oder stetig sein. Diskret sind sie dann, wenn nur bestimmte Zahlenwerte möglich sind, etwa die Anzahl von Einwohnern in einer Stadt. Stetige Werte hingegen können jeden Wert, der theoretisch möglich ist, annehmen (beispielsweise das Körpergewicht).

Ein in der Statistik weit verbreitetes System zur Datenaufbereitung und auch Auswertung, insbesondere für Fragebögen, ist das sogenannte SPSS. Die Abkürzung stand ursprünglich für "Statistical Package for the Social Sciences".

Durch die Firmenübernahme von IBM blieb zwar die Abkürzung bestehen, der Name jedoch wurde in PASW Statistics geändert, wobei PASW für "Predictive Analytics Software" steht (vgl. Raab-Steiner und Benesch 2012, S.67). Folgende Auflistung zeigt einen Teil der mit SPSS möglichen Anwendungen:

- Kodieren der Datensätze
- Sortieren und Verbinden der Daten
- Berechnung von statistischen sowie anderen Parametern
- Visualisierung durch Tabellen und Diagramme
- Kontingenz- sowie lineare Regressionsanalyse
- Faktor und Clusteranalyse
- Forecasting Modelle wie exponentielle Glättung oder ARIMA
- Conjoint Analysen (vgl. Pepels und Bruns 2008, S.540 f.)
- Fehler- und Plausibilitätsprüfung des Datensatzes (vgl. Raab-Steiner und Benesch 2012, S.79 f.)

Das Layout des Programms ist vergleichbar mit dem Layout von MS Excel. Der Ablauf der Implementierung lässt sich grob in drei Schritte gliedern. Anfangs müssen die Fragebögen vorbereitet werden. Dazu gehört unter anderem eine fortlaufende Nummerierung der Bögen. Des Weiteren werden im Programm selbst die Fragen in einem Codeplan strukturiert. Dabei werden die Fragen mit Variablennamen versehen und die später eingetragenen Merkmalsausprägungen mit Kodenummern (vgl. Raab-Steiner und Benesch 2012, S.72). Im zweiten Schritt werden die Informationen der Erhebung in das Programm übernommen und auf Fehler sowie ihre Plausibilität hin überprüft. Im letzten Schritt erfolgt dann die Auswertung. Der Vorteil von SPSS gegenüber Excel ist der, dass es ein rein statistisches Programm ist und sich zudem mit anderen Statistikprogrammen wie R (statistisches Datenanalyseprogramm) verknüpfen lässt. Dadurch bietet es dem Anwender mehr Möglichkeiten der Ausarbeitung.

Unabhängig von Merkmalsvarianten oder dem zur Aufarbeitung genutzten System ist die Fehler- und Plausibilitätsprüfung ein wichtiger Schritt vor der Auswertung. Das Spektrum der Fehler reicht dabei von einfachen, wie fehlenden Werten oder der Verwendung von Buchstaben statt Zahlen, bis hin zu komplexen Fehlern welche sich erst in der Analyse von Ganglinien oder der Gesamtheit der Werte feststellen lassen. Im Verkehrswesen können von der Erfassung der Daten bis hin zur Auswertung stets Fehler auftreten. Ursachen sind beispielsweise schlecht geschultes Personal bei Zählungen, schlecht oder gar nicht funktionierende Technik, falsche Zählzeiten oder Orte, usw..

Im Zuge der Datenaufbereitung werden die Datensätze auf diese und weitere Fehler hin untersucht um letztlich möglichst genaue Aussagen treffen zu können.

Die Form der Datenaufbereitung hängt oftmals von der untersuchten Fragestellung ab, sowie von der weiteren Verwendung. Bezogen auf das Verkehrswesen spielt hier die Hochrechnung von Verkehrsdaten eine wichtige Rolle. Dabei werden Werte, die bei Zählungen oder Befragungen aus einer Stichprobe erhoben wurden, hochgerechnet, um dadurch Rückschlüsse zu Werten aller entsprechenden Verkehrsteilnehmer treffen zu können.

Die für eine Hochrechnung erforderlichen Daten werden mit Hilfe von Zählungen gewonnen. Unterschieden wird dabei zwischen den Daten von manuellen Zählungen, und jenen von den aktuell 1.736 Dauerzählstellen auf Bundesautobahnen und Bundesstraßen (Geo-Basis-DE / BKG 2017). Einzelheiten zur Zählung wurden im vorangegangenen Kapitel bereits behandelt. Die folgende Tabelle zeigt nun die verschiedenen Hochrechnungsverfahren hinsichtlich ihrer Verwendung in Bezug auf die Lage der Zählstelle, die Zähldauer und den Zeitpunkt einer möglichen Hochrechnung.

| Verfahren | Lage der Zählstelle | | Zählzeit, Zähldauer | | Zeitpunkt der HR | |
|-----------|---------------------|-----------|---------------------|----------|------------------|------------|
| | innerorts | außerorts | genau definiert | flexibel | sofort | Jahresende |
| Arnold | X | | X | | X | |
| HBS 2001 | X | | X | | X | |
| HRDTV-Win | X | X | | X | X | |
| Lensing | | X | X | | X | |
| SVZ | | X | | X | | X |

Tabelle 2: Merkmale der verschiedenen Hochrechnungsverfahren für Fahrzeugzählungen (Forschungsgesellschaft für Straßen- und Verkehrswesen 2012, S.113)

Das erste der fünf aufgelisteten Verfahren stammt aus dem Jahr 2005 und wurde für die damalige Straßenverkehrszählung (SVZ) entwickelt. Grundgedanke dabei ist die Zuordnung von manuellen Zählstellen zu Dauerzählstellen. Ist das Verkehrsaufkommen einer manuellen Stelle ähnlich einer oder mehreren automatischen, erfolgt eine entsprechende Verknüpfung. Im Anschluss wird in zwei Schritten das Gesamtaufkommen hochgerechnet. Zuerst wird das Ergebnis der manuellen Erfassung einiger Stunden mit einem Stunden/Tag Faktor auf den Tag hochgerechnet.

Daraus resultieren mehrere Tageswerte für jede der sieben Fahrzeugklassen. Anschließend werden im zweiten Schritt die Tageswerte durch Tag/Jahr Faktoren auf das gesamte Jahr übertragen.

Resultat sind die durchschnittliche tägliche Verkehrsstärke aller Tage des Jahres (DTV) und die Bemessungsverkehrsstärke (MSV). Da dieses Verfahren die Faktoren der Dauerzählstellen benötigt, welche jedoch erst am Ende eines Jahres vorliegen, ist es in der Anwendung limitiert und eine Hochrechnung kann, im Gegensatz zu allen anderen Methoden, erst am Jahresende erfolgen (Bäumer und Hautzinger 2010, S. 41 f.).

Das Computerprogramm HRDTV-Win stellt das zweite Verfahren dar. Hierbei werden ebenfalls erst die gezählten Werte auf den Tag, und anschließend auf das Jahr hochgerechnet. Unterschied ist jedoch, dass die Faktoren nicht erst am Ende des jeweiligen Jahres vorliegen, sondern dauerhaft. Grund dafür ist die Verwendung von Werten aus vergangenen Jahren, welche jedoch ständig angepasst und aktualisiert werden. Dadurch kann die Hochrechnung jederzeit erfolgen und sich durch entsprechende Korrekturen auch für Ferien, Feiertage usw. anpassen. Das mit HBS 2001 dargestellte Verfahren ist eine minimal veränderte Variante des vorangegangenen. Im Handbuch zur Bemessung von Straßenverkehrsanlagen (HBS) wurde dieses Verfahren mit aufgenommen. Es ist für den Gebrauch innerorts vorgesehen und benötigt statt der bisherigen zwei Schritte insgesamt sechs um die DTV sowie die MSV zu berechnen. Das Vorgehen hierbei ist lediglich kleinschrittiger, da vom Tag erst auf die Woche hochgerechnet wird, anschließend noch Monat oder Halbmonat und letztlich erst auf das Jahr (Bäumer und Hautzinger 2010, S. 42 f.).

Verfahren Nummer vier ist ein Hochrechnungsverfahren nach Arnold, Hedeler, Wöppel und Dahme aus dem Jahre 2008 und dient der Hochrechnung für Kurzzeitzählungen welche innerorts stattfinden. Auch hier wird ein zweistufiger Ansatz verfolgt. Anders als bei den bereits erwähnten Varianten findet hier eine Unterscheidung der Verkehrsteilnehmer in KFZ und Schwerverkehr statt. Berücksichtigt werden zudem Wochen- und Saisonfaktoren. Zusätzlich kann die Genauigkeit der Berechnung über gegebene Koeffizienten angepasst und verbessert werden.

Trotz der Unterteilung wird das Verfahren in der Praxis als einfach und praktisch klassifiziert (Bäumer und Hautzinger 2010, S. 42 f.).

Das fünfte und letzte vorgestellte Verfahren ist das faktorbasierte Hochrechnungsverfahren nach Lensing, Mavridis und Täubner aus dem Jahre 2001. Nachdem die Zählung zu empfohlenen Zeiten durchgeführt wurde, werden die Werte erst auf den Tag, anschließend auf die Woche oder den Monat und schließlich auf das Jahr hochgerechnet.

Es kann zudem eine Unterteilung der Fahrzeugklassen erfolgen, ist jedoch nicht zwingend. Aufgrund der drei Stufen dieses Modells, sowie weiterer Faktoren die vorausgesetzt werden, wie den Sonntagsfaktor, werden viele Informationen benötigt um eine genaue Berechnung durchführen zu können ohne dabei die entsprechenden Faktoren zu vereinfachen.

Die Verfahren unterscheiden sich somit stark in ihrer Komplexität sowie in ihrer Anwendbarkeit im Straßenverkehr. Auf der folgenden Abbildung lässt sich der Weg der Entscheidungsfindung hinsichtlich des Verfahrens in Abhängigkeit der gegebenen Umstände nochmals nachvollziehen.

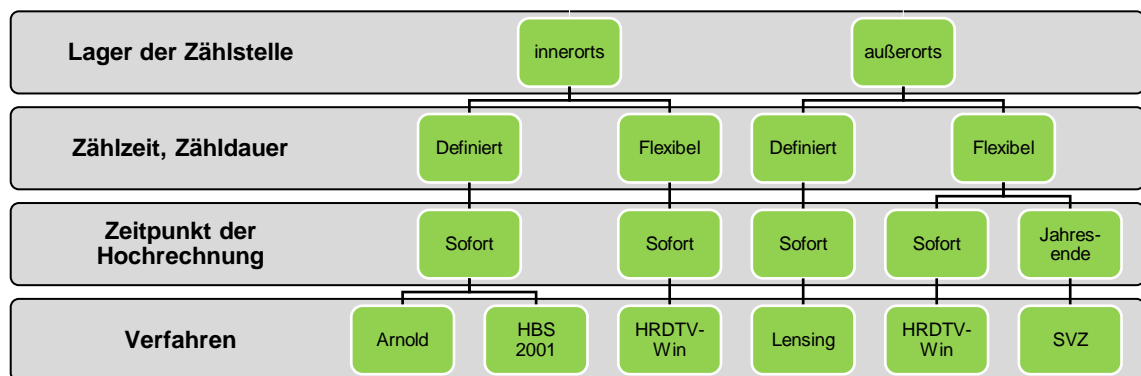


Abb. 3: Entscheidungshilfe zur Auswahl eines Hochrechnungsverfahrens für Fahrzeugzählungen (Bäumer und Hautzinger 2010, S.46)

2.1.3 Datenauswertung [MR]

Eine empirische Arbeit gliedert sich in mehrere Abschnitte. Den Anfang bildet das Problem, welches es zu lösen oder zu untersuchen gilt. Im Anschluss daran erfolgt die Methodik der Problemlösung, samt Verfahren und Durchführung. Ist dies geschehen, werden die daraus gewonnenen Ergebnisse den Rezipienten vorgestellt.

Die Datenauswertung ist nichts anderes als jene Ergebnisdarstellung (vgl. Raab-Steiner und Benesch 2012, S.166 f.).

Die geläufigsten Formen der Ergebnisdarstellung von statistischen Untersuchungen sind Tabellen und Diagramme. Die Tabellen unterscheiden sich dabei von jenen, welche in der Phase der Datenaufbereitung verwendet wurden. Form und Layout sind hierbei ebenso entscheidend wie der Inhalt.

Die Tabelle sollte in sich verständlich und lesbar sein, sodass es theoretisch keiner weiteren Erläuterung bedarf. Erklärungen zu den Werten erfolgen durch die Überschrift sowie den Zeilen- und Spaltentexten. Abgesehen von gängigen Abkürzungen sollten alle weiteren Abkürzungen vermieden werden. Weitere ausführende Inhalte sollten in einem Begleittext oder entsprechenden Fußnoten beigefügt werden. Weitere Aspekte wie ein einheitliches Farbschema oder gleiche Schriftgrößen sind dabei selbstverständlich (vgl. Pepels und Bruns 2008, S.579 f.).

Neben Tabellen liefern Diagramme schnelle und verständliche Informationen über Sachverhalte. Unterschieden wird dabei zwischen drei Haupttypen. Es handelt sich dabei um Punktdiagramme, Liniendiagramme und Flächendiagramme.

Punktdiagramme dienen der Veranschaulichung von „Häufigkeit und Streuung“ (Pepels und Bruns 2008, S.582). Jeder Punkt im Diagramm entspricht dabei einem statistischen Wert. Die Punkte können dabei noch durch Linien verbunden werden um Zusammenhänge besser darzustellen. Auch hier gibt es verschiedene Variationen der Darstellung, welche von der bloßen Darstellung von Zeitreihen, bis hin zu Spider Charts reichen, mit denen multivariate Zusammenhänge dargestellt werden können.

Bei Liniendiagrammen werden Linien parallel zu einer der Achsen ins Koordinatensystem übernommen. Die Länge einer solchen Linie gibt Aufschluss über die statistische Größe welche mit ihr dargestellt werden soll. Die dritte und letzte Variante sind Flächendiagramme. Hierbei stellt die Größe der Fläche den statistischen Wert oder dessen Verhältnismäßigkeit dar. Meist werden dabei Rechtecke, Kreise sowie deren Variationen und Teilflächen verwendet.

Mit Flächendiagrammen können sowohl nominale als auch ordinale Merkmalsausprägungen dargestellt werden. Zudem lassen beispielsweise Kreisdiagramme eine wertungs- und rangfolgenfreie Darstellung der Merkmalsausprägungen zu (vgl. Pepels und Bruns 2008, ebenda f.).

Ergebnisdarstellungen von Verkehrsuntersuchungen nutzen ebenfalls die vorgestellten Varianten. Beispiele sind der Straßengüterverkehrsbericht, in dem die Masse der beförderten Güter sowie die Beförderungsleistung tabellarisch dargestellt wird, oder auch der Modal-Split, welcher häufig in Form von Balken-, also Flächendiagrammen dargestellt wird.

2.2 Methodische Grundlagen in Excel [MR]

MS Excel ist eines von vielen Tabellenkalkulationsprogrammen. Solche Programme dienen der Bearbeitung von Problemen und Daten, welche in Tabellenform vorliegen. Der Aufbau ist generell immer gleich. Auf einem Arbeitsblatt werden Zeilen und Spalten dargestellt in welchen sich die Zellen befinden, in denen die Werte und Rechenoperationen eingetragen und ausgeführt werden können (Gabler Wirtschaftslexikon 2017e). Beispiele solcher Programme sind Calc als Teil von OpenOffice, Numbers von Apple oder eben Microsoft Excel als Teil von Microsoft Office. MS Excel ist eines der am häufigsten verwendeten Tabellenkalkulationsprogramme. Seine Anwendung „reicht dabei vom privaten Haushaltsbuch bis hin zu professionellen Nutzung im Rahmen von Forschung und Ökonomie. Mit MS Excel lassen sich einfach, jedoch auch komplexe Analysen anstellen, die dank zahlreicher Darstellungsmöglichkeiten schnell und einfach grafisch aufbereitet werden können. Hinzu kommt, dass MS Excel Formeln und Funktionen anbietet, die für eine gezielt statistische Anwendung nützlich und hilfreich sind“ (Feiks 2016, S.25).

Eine MS Excel Datei wird auch als Arbeitsmappe bezeichnet. Diese bestehen aus einem oder mehreren Tabellenblättern. Die Tabellenblätter enthalten Zeilen und Spalten. Zeilen werden mit aufsteigenden Nummern und Spalten mit Buchstaben betitelt. Dadurch kann jede Zelle in einem Tabellenblatt eindeutig zugeordnet werden (beispielsweise Inhalt aus Spalte B in Zeile 4 befindet sich also in Zelle B4). Inhalt der Zellen können Zahlen, Texte oder Funktionen und Formeln sein. Die Formatierung des Layouts einzelner Tabellenteile oder der gesamten Datei ist dabei variabel anpassbar.

Kern der Arbeit mit MS Excel sind die Funktionen und Formeln. Eine Formel innerhalb einer Zelle wird stets mit dem Gleichheitszeichen eröffnet. Dadurch wird dem Programm signalisiert dass im Anschluss eine Formel oder Rechenoperation folgt. Die Grundrechenarten bilden die einfachsten Operatoren und Formeln.

Innerhalb einer Rechnung können sowohl Konstanten verwendet werden, als auch Inhalte aus anderen Zellen. Es handelt sich dann um sogenannte Zellbezüge (vgl. Wies 2013, S.32 f.). Eine einfache Formel könnte demnach wie folgt aussehen:

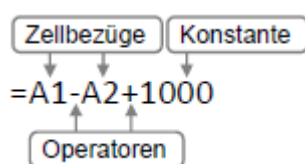


Abb. 4: Grundlegender Formelaufbau in MS Excel (Wies 2013, S.32)

Neben den Grundrechenarten verfügt MS Excel über ein Portfolio von rund 460 Formeln und Funktionen die bei einer Bearbeitung Anwendung finden können. Eine der einfachsten ist die Funktion "Summe". Nach der Einleitung mit dem Gleichheitszeichen folgt wie bei vielen Funktionen der Name, also hier „=summe()“. Im Anschluss öffnet sich eine, manchmal auch mehrere Klammern, in welche die für die Funktion relevanten Daten eingetragen werden können. Mit der Summenfunktion lassen sich sowohl einzelne Werte aufsummieren, als auch ganze Zeilen, Spalten oder Flächen. Funktionen können innerhalb MS Excels auch verschachtelt werden, wodurch sie in- oder nacheinander ausgeführt werden. Ein oft verwendetes Beispiel für eine solche verschachtelte Funktion ist die Wenn-Funktion. Diese logische Funktion dient der Überprüfung von Zelleninhalten auf bestimmte definierte Bedingungen und führt anschließend eine von zwei Aktionen aus, welche vom überprüften Inhalt abhängig sind (vgl. Feiks 2016, S.32 f.). Der Aufbau gestaltet sich dabei folgendermaßen:

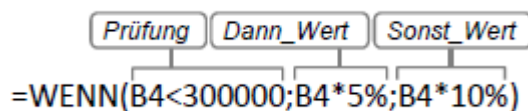


Abb. 5: Aufbau der Wenn-Formel in MS Excel (Wies 2013, S.88)

Im ersten Argument, der Prüfung, wird eine Konstante oder ein Zellenbezug auf eine Bedingung hin untersucht. Dabei kann es sich um simple Umstände wie kleiner oder größer als handeln, jedoch genauso kann die Bedingung ebenfalls eine Funktion sein. Ist die Bedingung erfüllt, wird der Inhalt von Dann_Wert, andernfalls der Sonst_Wert wiedergegeben. Beide Werte können, wie auch die Bedingung, erneut aus Formeln oder Rechenoperationen bestehen (vgl. Wies 2013, S.88) .

Durch die Verkettung von Wenn-Funktionen lässt sich beispielsweise ein Zelleninhalt auf beliebig viele Bedingungen hin untersuchen.

Neben den vorhandenen Formeln und Funktionen bietet MS Excel dem Anwender die Möglichkeit selbst Abläufe und Routinen zu erstellen, mit Hilfe von VBA und Makros. VBA ist eine Abkürzung und steht für „Visual Basic for Applications“. Es handelt sich dabei um eine objektorientierte Programmiersprache. Durch diese Programmierung innerhalb MS Excels gibt es die Möglichkeit andere Office Anwendungen zu integrieren, eigene Funktionen zu erstellen oder Visualisierung und Abläufe zu automatisieren.

Makros hingegen sind Routinen, die sich zwar ebenfalls programmieren, aber auch aufzeichnen lassen. Die Ausführung eines Vorgangs oder auch einer Abfolge von Vorgängen wird dabei durch das Programm dokumentiert, gespeichert und kann schließlich für zukünftige Arbeiten automatisch wiederholt werden. Weitere Ausführungen sind wenig sinnvoll, da die Systemkomponenten allein sehr umfangreich sind und in dieser Abhandlung keinerlei Verwendung finden.

Zusätzlich bietet MS Excel dem Anwender eine Auswahl von Auswertungs- und Prognoseverfahren, welche sich ohne weitere Berechnungen der Daten umsetzen lassen. Mithilfe von Trendlinien können in Datensätzen Zusammenhänge in Form von Verläufen aufgezeigt werden, welche sich wiederum fortführen lassen und so zur Prognose genutzt werden können. Nach der Erstellung eines Diagramms beispielsweise zu einer Zeitreihe mit den jeweiligen Werten, besteht die Möglichkeit, im Diagrammmenu eine Trendlinie einzufügen. Hierbei stehen dem Anwender unterschiedliche Varianten zur Verfügung, wie zum Beispiel eine exponentielle, lineare, polynomische oder logarithmische Trendlinie. Diese unterscheiden sich im Anwendungsbereich und im Layout der Formeln. So eignen sich logarithmische Trendlinien für Werte, die schnell ansteigen oder abnehmen, während polynomische für Datensätze sinnvoll sind in denen die Daten schwanken (Microsoft Office 2017b). Die Berechnung der Funktionsformeln erfolgt stets nach dem gleichen Muster und ist lediglich auf den jeweiligen Formeltypus angepasst. Die Methode der kleinsten Quadrate berechnet hierbei die benötigten Variablen (Microsoft Office 2017a). Durch dieses Verfahren werden die Parameter einer gesuchten Funktion in Abhängigkeit der Minimierung der Summe der quadrierten Störfaktoren geschätzt. Sie ist eine der oft gebrauchten Methoden um Regressionsmodelle zu bestimmen (Gabler Wirtschaftslexikon 2017c).

Letztlich beinhaltet MS Excel eine Vielzahl von grafischen Variationen zur Veranschaulichung von Daten. Die in Kapitel 2.1.3 vorgestellten Typen finden sich dabei in mannigfaltigen Differenzierungen hier wieder und werden somit nicht weiter ausgeführt.

3 Prognoseverfahren im Verkehrswesen

3.1 Überblick Verkehrsplanung [K7]

Im folgenden Abschnitt wird die Verkehrsplanung vorgestellt. Aufgabe der Verkehrsplanung ist es, Aussagen über mögliche zukünftige Verkehrsentwicklungen zu treffen. Da die genutzten Anlagen langlebig sind, aber auch einen hohen Wert besitzen, wird ein langfristiger Planungshorizont unterstellt (Schnabel und Lohse 2011, S. 1). Dieser wird in der Verkehrsplanung unterschiedlich definiert. Während Schnabel und Lohse (2011, S. 5) 15 bis 25 Jahre als langfristigen Planungshorizont definieren, richtet sich Köhler (2014, S. 57) nach fünf bis 20 Jahren. Bei Zeiträumen bis zu fünf Jahren fallen Verkehrsentwicklungsprognosen zwar vergleichsweise genau aus, jedoch sind die Möglichkeiten der Einflussnahme auf den Verkehrsanlagenbestand begrenzt. Hier wird mit bekannten Anlagen der Verkehr prognostiziert. Bei der Mittelfristprognose, die einen Planungshorizont von bis zu zehn Jahren umfasst, nehmen die Einflussmöglichkeiten auf den Anlagenbestand zu, während die Prognosegenauigkeit abnimmt. Die Langfristprognose, ein Zeitraum bis zu 20 Jahren, ist mit hoher Prognoseunsicherheit verbunden. Die Prognose ist hierbei am ungenauesten, jedoch hat der Planer hier die meisten Einflussmöglichkeiten (Köhler 2014, S. 57).

Die Verkehrsplanung gliedert sich nach Lohse und Schnabel (2011, S. 5) in drei Teilschritte: Zunächst erfolgt eine Analyse, die Informationen über das betrachtete System liefert. Die Erkenntnisse aus diesem Schritt dienen als Basis für die spätere Prognose. Das System wird bewertet und wissenschaftlich begründete Verfahren oder Modelle werden erprobt. Der zweite Schritt umfasst die Prognose im engeren Sinne. Hier wird das zu erwartende Verkehrsgeschehen vorausgesagt. Die in der Analyse verwendeten Verfahren werden hierfür herangezogen. Der letzte Schritt ist die Therapie oder Maßnahmenuntersuchung. Dabei werden konkrete planerische Gestaltungsvorschläge erarbeitet.

Der Untersuchungsbereich gliedert sich nach Köhler (2014, S. 51) in zwei Arten. Die zu untersuchende Raumeinheit wird als Untersuchungsgebiet bezeichnet. Dabei ist eine Fläche gemeint, die abtrennbar ist, z.B. das Gebiet einer Stadt. Werden diese Gebietsgrenzen überschritten, wird der sogenannte Untersuchungsraum oder auch erweiterbares Planungsgebiet betrachtet. Da Verkehrsströme nicht nur im Inneren des Untersuchungsgebietes entstehen, sondern auch der Durchgangsverkehr nicht außer Acht gelassen werden darf, muss Letzteres ebenfalls in die Planung mit einfließen.

3.2 Trendprognose [KT]

Die sogenannte Trendprognose ist die einfachste Methode der Verkehrsplanung. Die Planungsgrundlage stellt dabei die Entwicklung einer interessierenden Größe im Zeitverlauf dar. Aus dieser Datengrundlage kann der Planer die zukünftige Entwicklung mithilfe linearer, exponentieller oder logistischer Funktionen ableiten (Köhler 2014, 57).

Eine Art der Trendprognose ist die Zeitreihenanalyse. Es wird zwischen unmittelbarer und mittelbarer Zeitreihenprognose unterschieden. Die unmittelbare Zeitreihenprognose ist ein Instrument, welches in Ausnahmefällen angewendet werden kann, jedoch ist sie vornehmlich für Kurzzeitprognosen geeignet. Vorausgesetzt wird bei dieser Art der Prognose, dass eine ausreichende Menge an statistischen Daten vorliegt und der Funktionstyp erkennbar ist (Köhler 2014, S. 57).

3.2.1 Einfache Prognoseverfahren zur Zeitreihenanalyse [KT]

Ist die bisherige Entwicklung nicht durch eine analytische Funktion erkennbar, wird häufig der gleitende Durchschnitt verwendet. Dabei wird eine unregelmäßige, aber dennoch einem Trend folgende Entwicklung unterstellt (Köhler 2014, S. 57).

Bei dieser Methode werden, anders als bei der Bildung eines arithmetischen Mittels der Zeitreihe, nur die letzten n Perioden in die Berechnung mit einbezogen. Ziel ist es, zufällige Schwankungen in der Zeitreihe zu eliminieren (Buzacott et al. 2010, S. 34). Die Formel lautet:

$$\hat{x}_t = \sum_{t'=t-n+1}^T \frac{x_{t'}}{n}$$

Wird für n ein kleiner Wert festgelegt, so wird die Prognose naiver. Eine naive Prognose berücksichtigt nur den Wert aus der letzten Periode und ist im Fall des gleitenden Durchschnitts $n = 1$. Ist n ein hoher Wert, so werden bei der Prognose zusätzlich viele Vergangenheitswerte berücksichtigt. Bei $n = T$ wird das arithmetische Mittel der gesamten Datenreihe gebildet, was bei hohen Schwankungen in der Zeitreihe zu einer ungenauen Prognose führt (Hillier et al. 2002, S. 661).

Je nachdem, wie groß n gewählt wird, kann die Prognose nicht in $t = 1$ beginnen, sondern in $n + 1$.

Die berücksichtigten Werte für die Berechnung werden alle gleich stark gewichtet, d.h. die aus der letzten Periode genauso stark, wie die aus einer der älteren Perioden. Das stellt einen Nachteil der Prognose basierend ausschließlich auf dem gleitenden Durchschnitt dar (Hillier et al. 2002, S. 661f.).

Die exponentielle Glättung beugt diesem Nachteil vor, indem ein Glättungsfaktor eingeführt wird. Dieser liegt zwischen 0 und 1 und wird als α bezeichnet. Die Formel für die exponentielle Glättung 1. Ordnung lautet:

$$\hat{x}_t = \alpha \cdot x_{t-1} + (1 - \alpha) \cdot \hat{x}_{t-1}$$

Dabei ist \hat{x}_{t-1} der prognostizierte Wert der Periode t und beinhaltet somit auch alle Werte aus den vorhergegangenen Perioden, während x_{t-1} der Realwert der vorhergegangenen Periode beschreibt. Wird α klein gewählt, so werden aktuelle Werte nur schwach und alte Werte stärker berücksichtigt. Die Prognose reagiert nicht auf Schwankungen. Bei der Wahl eines großen α werden die aktuellen Werte stärker berücksichtigt und die alten Werte schwächer. Hier sind sehr starke Schwankungen zu beobachten (Bamberg et al. 2012, S. 201f.). In der Praxis erwies sich ein α zwischen 0,1 und 0,3 als sinnvoll (Hillier et al. 2002, S. 663).

3.2.2 Zeitreihendekomposition [KT]

Beide oben genannten Verfahren sind geeignet, wenn eine konstante Zeitreihe vorliegt. Sind saisonale Faktoren zu berücksichtigen, kann die Zeitreihendekomposition eine gute Prognose liefern. Dabei wird die Zeitreihe in Komponenten mit einer regelmäßigen Charakteristik zerlegt. Zufällige Schwankungen können so eliminiert werden und die regelmäßigen Komponenten können später zu einer Vorhersage kombiniert werden. Es wird unterschieden zwischen der multiplikativen Kombination (wenn die Annahme besteht, dass die saisonalen Effekte im Laufe der Zeit zunehmen) und der additiven Kombination, bei der die Annahme besteht, dass die saisonalen Effekte unabhängig vom Zeitverlauf sind (Buzacott et al. 2010, S. 40; Lippe 1993, S. 403).

Die Schätzgleichungen für das multiplikative und das additive Modell lauten:

Multiplikativ: $\hat{x}_t = TR_t \cdot s_t$

Additiv: $\hat{x}_t = TR_t + s_t$

Dabei ist TR_t die Trendfunktion der Prognose in Zeitpunkt t , während s_t mittelfristige zyklische Schwankungen, eine Saisonkomponente und eine irreguläre Komponente beinhaltet. Im Folgenden wird s_t saisonale Komponente genannt (Buzacott et al. 2010, S. 40; Bamberg et al. 2012, S. 59).

Für die Zeitreihendekomposition existiert eine systematische Vorgehensweise. Im ersten Schritt wird die Glättungskomponente \check{x}_t bestimmt, um saisonale Schwankungen zu glätten. Diese ist ein zentrierter gleitender Durchschnitt über die Anzahl der Saisonperioden.

Es wird unterschieden zwischen gerader und ungerader Saisonperiodenzahl. Die Anzahl der Saisonperioden wird durch Q bezeichnet. Für eine ungerade Saisonperiodenzahl gilt die Formel $\check{x}_t = \frac{1}{Q} \cdot \sum_{t'=t-Q'}^{t+Q'} x_{t'}$, mit $Q' = \frac{Q-1}{2}$. Die Formel für die gerade Saisonperiodenzahl lautet $\check{x}_t = \frac{1}{Q} \cdot \left(\frac{x_{t-Q \cdot 0,5} + x_{t+Q \cdot 0,5}}{2} \sum_{t'=t-Q'}^{t+Q'} x_{t'} \right)$, mit $Q' = \frac{Q}{2} - 1$ (Buzacott et al. 2010, S. 41).

Im nächsten Schritt werden die saisonale und die irreguläre Komponente berechnet. Bei der multiplikativen Zeitreihendekomposition ist dies die Differenz aus dem Realwert und der Glättungskomponente $sy_t = \frac{x_t}{\check{x}_t}$, bei der additiven Variante ist es die Differenz aus diesen beiden Werten $sy_t = x_t - \check{x}_t$ (Buzacott et al. 2010, S. 41; Bamberg et al. 2012, S. 63).

Des Weiteren werden die Saisonfaktoren berechnet, um irreguläre Schwankungen zu eliminieren. Dazu werden saisonperiodenbezogene Mittelwerte $sy_q = \frac{1}{Q} \cdot \sum_{a=1}^A sy_{aq}$ ermittelt. Diese Werte sind jedoch noch nicht standardisiert, sie ergeben in Summe noch nicht Q (Buzacott et al. 2010, S. 42).

Die Standardisierung unterscheidet sich erneut beim multiplikativen und additiven Verfahren. Für die multiplikative Prognose lautet die Formel $s_q = sy_q \cdot \frac{Q}{\sum_{q=1}^Q sy_q}$, für die additive Zeitreihendekomposition $s_q = sy_q - \frac{\sum_{q=1}^Q sy_q}{Q}$ (Buzacott et al. 2010, S. 41; Bamberg et al. 2012, S. 63).

Für die Bestimmung der Trendfunktion kann ein Regressionsverfahren verwendet werden. Wird ein linearer Trend angenommen, wird die lineare Regression mit der Funktion $TR_t = a + b \cdot t$ zur Bestimmung der Parameter a und b verwendet (Buzacott et al. 2010, S. 42).

Die ermittelten Komponenten können nun in die Formeln $\hat{x}_t = TR_t \cdot s_t$ (multiplikativ) und $\hat{x}_t = TR_t + s_t$ (additiv) eingesetzt werden, damit die Prognosewerte berechnet werden können (Buzacott et al. 2010, S. 42).

3.2.3 Eignung der Zeitreihenanalyse [KT]

Ob sich ein Verfahren für eine Prognose eignet, kann mit verschiedenen Kennzahlen ermittelt werden. Im Folgenden werden die Kennzahlen mittlere absolute Abweichung, mittlere quadratische Abweichung und Theil's Ungleichheitskoeffizient erläutert.

In der Berechnung dieser Kennzahlen findet der Prognosefehler e_t Einsatz. Dieser kann als Differenz, als absolute Abweichung oder als quadratischer Wert angegeben werden (Buzacott et al. 2010, S. 45). Da sich bei einer Betrachtung des arithmetischen Mittels der Differenz das Ergebnis zu Null ausgleicht, ist die Verwendung der absoluten Abweichung sinnvoller. Die Formel für die mittlere absolute Abweichung (MAA) lautet somit $MAA = \frac{1}{T} \cdot \sum_t |x_t - \hat{x}_t|$ (Bleymüller et al. 2015, S. 27; Puhani 2012, S. 41).

Eine weitere Möglichkeit des Ausgleichs auf Null vorzubeugen, ist die Aufhebung des Vorzeichens durch Quadrierung und somit die Bestimmung der mittleren quadratischen Abweichung (MQA) durch $MQA = \frac{1}{T} \cdot \sum_t (x_t - \hat{x}_t)^2$ (Müller und Poguntke 2015, S. 18). Oft wird auch die Quadratwurzel der MQA berechnet, um die Abweichung in dieselbe Dimension wie die prognostizierten Werte zurückzuführen (Theil 1966, S. 27).

Ferner eignet sich der Ungleichheitskoeffizient von Theil (1966, S. 28) mit der Formel $TU = \sqrt{\frac{\sum_t (x_t - \hat{x}_t)^2}{\sum_t (x_t - x_{t-1})^2}}$ zur Untersuchung der Prognosegüte. Dabei wird das Verhältnis der Prognose zur naiven Prognose betrachtet. Ist $TU = 1$, so ist die Prognose naiv; bei einem Wert von $TU = 0$ liegt eine „perfekte“ Prognose vor. Ist $TU > 1$, dann ist die Abweichung des Prognosewert zum Realwert so groß, dass in diesem Fall die naive Prognose besser wäre (Theil 1966, S. 28)

Eine unterstellte geringe Eignung der unmittelbaren Zeitreihenprognose ist darauf zurückzuführen, dass es sich um eine stark vereinfachte Methode handelt. Der einzige betrachtete Parameter ist die Zeit, von der die Entwicklung direkt abhängig ist. Bei der mittelbaren Zeitreihenprognose werden zeitvariable Einflussgrößen mit einbezogen. Auf Basis der Kurzzeitprognose kann dann eine Langzeitprognose aus einer rekursiven Anwendung ermittelt werden (Schnabel und Lohse 2011, S. 84).

Zeitreihenanalysen sind univariat, d.h. sie betrachten nur die Zeitreihe und die dazugehörigen Datenreihen, nicht aber andere Einflussgrößen. Dafür sind kausale Verfahren geeignet. Ein Beispiel für ein solches Verfahren ist die Regressionsanalyse (Bamberg et al. 2012, S. 201).

3.3 Verkehrsnachfragemodelle [KT]

Verkehrsnachfragemodelle oder kurz Verkehrsmodelle bilden das Verhalten der Verkehrsnachfrager ab und werden anhand mehrerer Kriterien unterschieden. Dabei findet eine mathematische Verknüpfung von Wirtschafts-, Siedlungs-, Haushalts- und Verkehrsstrukturdaten und Informationen zum Verkehrsgeschehen mithilfe formalisierter Verhaltenshypothesen statt. Die Verkehrsmodellierung ist der Kern der Verkehrsplanung und ist Bestandteil der Analyse und Prognose.

Eine mögliche Einteilung erfolgt nach dem Entscheidungsablauf. Bei der sequenziellen Modellierung wird eine Kette von Submodellen erzeugt. Die Analyse und Prognose der Verkehrserzeugung, Verkehrsverteilung, Verkehrsmittelwahl und Verkehrswegwahl (siehe Abschnitt) erfolgt schrittweise. Simultane Modelle fassen mehrere dieser Submodelle zusammen.

Des Weiteren können Verkehrsmodelle nach der Aggregationsstufe der Datenbasis eingeteilt werden. In aggregierten Modellen wird nach Fahrzwecken der Verkehrsteilnehmer (z.B. Freizeit- oder Berufsverkehr) oder nach Wirtschaftsbereichen sowie Gütergruppen differenziert. Ausgegangen wird hierbei von verkehrsbestimmten Strukturdaten der Verkehrszellen als verkehrsverhaltensbestimmte Einflussgrößen. Disaggregierte Verkehrsmodelle betrachten relativ homogene Verkehrsteilnehmergruppen, wie z.B. sozioökonomische Gruppen (Gabler Wirtschaftslexikon).

Ferner kann eine Unterscheidung zwischen mikroskopischen und makroskopischen Modellen stattfinden. Im ersten Fall wird jeder Weg des einzelnen Verkehrsteilnehmers oder der Bewegungsablauf der einzelnen Fahrzeuge vereinfacht dargestellt. Verwendet werden mikroskopische Modelle für die Simulation des Verkehrsablaufs auf Straßen oder Schienenwegen (z.B. für die Untersuchung von Lichtsignalanlagen) (Köhler 2014, S. 56).

Makroskopische Modelle werden zu den aggregierten Modellen zugeordnet und als „klassische“ Modelle bezeichnet (Aberle 2000, S. 422). Die Vorgehensweise ist auf zwei Arten möglich, zum einen mit dem top-down Ansatz, mit welchem ein gesamträumlicher Prognosewert auf aggregiertem Niveau ermittelt wird und zum anderen mit einem bottom-up Ansatz, mit welchem zunächst kleinere Einheiten prognostiziert werden, die später zu einem makroskopischen Ergebnis zusammengefasst werden (Stock und Bernecker 2014, S. 200).

In der Verkehrsnachfragemodellierung wird seit den 1960er Jahren der sogenannte Vier-Stufen-Ansatz verwendet (Ortúzar und Willumsen 2011, S. 21). Dazu gehören die Verkehrserzeugung, die Verkehrsmittelwahl, die Verkehrsverteilung und die Verkehrsumlegung. Im Folgenden werden diese Aspekte anhand eines simulativen Verkehrsmodells erläutert.

Bei der Verkehrserzeugung wird das Verkehrsaufkommen eines Gebietes ermittelt. Dieses hängt von der Art und dem Maß der baulichen Nutzung auf dem Gebiet ab (Köhler 2014, S. 34).

Die Verkehrsmittelwahl wird durch ein Modal Split dargestellt. Die bekannten Wege werden unterschiedlichen Verkehrsmitteln zugeordnet. Dieser Schritt wird jedoch nur bei Verkehrsteilnehmern angewendet, die zwischen unterschiedlichen Verkehrsmitteln wählen können. Im Modell wird angenommen, dass ein Verkehrsteilnehmer jeder Alternative einen Nutzen zuordnet. Die Wahl fällt dann auf die Alternative mit dem größten Nutzen, jedoch wird vorausgesetzt, dass der Verkehrsteilnehmer vollständige Informationen über das Verkehrssystem besitzt. Um das Modell implementieren zu können, muss der Nutzen der jeweiligen Alternativen quantifizierbar sein (Köhler 2014, S. 40f).

Der Aspekt der Verkehrsverteilung ist die Zuordnung von Zielen zu den gebildeten Wegen aus dem Schritt der Verkehrserzeugung. Dies geschieht in Abhängigkeit von der beabsichtigten Aktivität. Diese Modellierung kann vor oder nach der Verkehrsmittelwahl angewendet werden. Die Wahl des Zielorts ist abhängig von der Entfernung zwischen Quelle und Ziel, der Attraktivität des Ziels, dem Zielpotential konkurrierender Zielorte, sowie der Widerstandsempfindlichkeit¹ (Köhler 2014, S. 43).

Die Schritte Verkehrserzeugung, Verkehrsmittelwahl und Verkehrsverteilung laufen nicht unabhängig voneinander ab. Für den Verkehrsteilnehmer besteht eine enge Verknüpfung der Prozesse. Die Wahl des Verkehrsmittels ist z.B. vom gewähltem Ziel und dessen Erreichbarkeit mit den zur Verfügung stehenden Verkehrsmitteln abhängig. Außerdem kann die Wahl des ersten Verkehrsmittels die Wahl des zweiten bestimmen. Es gilt, dass das Ziel des n -ten Weges die Quelle des Weges $n + 1$ ist (Köhler 2014, S. 44f).

¹ Als Widerstand wird die Verbindungsqualität eines Weges zwischen einer Quelle i und einem Ziel j definiert. Als Einflussgrößen fließen die Reisezeit oder -kosten in die Widerstandsfunktion ein. (Köhler 2014, S. 43)

Der letzte Schritt ist die Verkehrsumlegung oder Routenwahl. Für die ermittelten Wege werden Routen im Verkehrsnetz mit den zur Verfügung stehenden Verkehrsmitteln bestimmt. Aufgabe der Routenwahl ist es, die zeitkürzeste Route zwischen einer Quelle i und dem Ziel j zu finden. Da die Routenwahl ein eigenständiges Teilmodell ist, wird diese in dieser Arbeit nicht weiter betrachtet (Köhler 2014, S. 45).

Ein Verkehrsflussmodell kann als Teil der Verkehrsnachfragemodellierung verstanden werden. Die Modellierung kann z.B. für die Simulation von Verkehrsflüssen bei der Routenwahl verwendet werden. Die Modelle bilden Verkehrsbewegungen ab: Die Verkehrsteilnehmer bewegen sich vom Start zum Ziel auf einer gegebenen Route durch das Verkehrsnetz. Dabei beschreiben mikroskopische Modelle die Interaktion zwischen einzelnen Fahrzeugen, während bei der makroskopischen Betrachtung Kenngrößen wie Verkehrsstärke und Verkehrsdichte ermittelt werden (Aleksić 2012, S. 16).

3.4 Angewendete Modelle [KT]

Viele Modelle, die in der Literatur zu finden sind, werden für eine kurzfristige Prognose des Verkehrsstatus für intelligente Transportsysteme (ITS) verwendet. Han und Song (2003) geben eine Übersicht über die gängigen grundlegenden Verfahren an.

Ein Verfahren, welches als optimale Prognose gilt, ist das ARIMA Modell. Dieses basiert auf der Box-Jenkins Methode (Bamberg et al. 2012, S. 206; Han und Song 2003, S. 216). Sie gilt als optimal, da diese Methode als einzige Methode eine sehr geringe mittlere quadratische Abweichung besitzt, allerdings gilt dies nur, wenn die historischen Daten eine lineare Form aufweisen (Han und Song 2003, S. 217).

Die Box-Jenkins Methode ist außerdem sehr komplex, lässt sich aber mit Computer Software lösen (Hillier et al. 2002, S. 666).

Ein Modell, welches auf dem Verfahren des ARIMA Modells basiert, ist das Verfahren von Wagner-Muns et al. (2017). Es verwendet ein SARIMA Modell, welches eine weiterentwickelte Form von ARIMA darstellt.

Zusätzlich wird die Prognose durch die sogenannte Functional Principal Component Analysis, ein Teilgebiet der Functional Data Analysis, ergänzt und miteinander verglichen. Ziel ist es bei der Arbeit von Wagner-Muns et al. (2017), das Verkehrsaufkommen eines bestimmten Wochentages zu prognostizieren. Der Datensatz besteht aus einer Zeitreihe in 15 Minuten Intervallen, dennoch wird mit dem verwendeten Modell gezeigt, dass eine langfristige Prognose für einen Tag möglich ist.

Ebenfalls für die sehr kurzfristige Prognose werden Modelle mit neuronalen Netzen verwendet (Han und Song 2003, S. 217). Neuronale Netze bestehen aus künstlichen Neuronen, welche einfache Elemente darstellen. Diese Elemente nehmen Daten aus verschiedenen Quellen auf und berechnen mithilfe einer Transferfunktion von den Inputs abhängige Outputs. Durch gewichtete Verbindungen werden diese Neuronen wieder zusammengesetzt. (Chrobok 2005, S. 32)

Die neuronalen Netzwerke werden aus dem Grund verwendet, dass Verkehrsdaten von stochastischer Natur sind und besonders Kurzzeitprognosen keine linearen Charakteristika aufweisen (Han und Song 2003, S. 217). Unter Wissenschaftlern besteht die Diskussion, dass neuronale Netzwerke Vorteile gegenüber dem ARIMA Modell haben, was jedoch nicht immer der Fall ist (Han und Song 2003, S. 217; Chang et al. 2000, S.). Lv et al. (2014) verwenden in ihrem Modell neuronale Netzwerke, welche aus übereinander angeordneten Autoencodern, welche ihren Input reproduzieren, bestehen. Dadurch entsteht ein sogenanntes „deep network“ welches durch eine Greedy Heuristik Eigenschaften von Verkehrsflüssen erlernt, welche später für die Prognose verwendet werden können.

Chrobok (2005, S. 48) beschreibt in seiner Arbeit außerdem die Möglichkeit Kurz- und Langzeitprognosen zu kombinieren, indem z.B. ein ARIMA Modell in Kombination mit einer Heuristik ausgewertet wird.

Die Modelle und Verfahren sind nicht unmittelbar miteinander vergleichbar, da unterschiedliche Datensätze verwendet werden (Chrobok 2005, S. 181).

4 Anwendung der Prognosemethoden auf ein Anwendungsbeispiel

4.1 Beschreibung der Datengrundlage [MR]

Als Datengrundlage dienen zwei Datensätze aus der Straßengüterverkehrsstatistik des statistischen Bundesamtes. Inhaltlich listen sie die Menge der beförderten Güter sowie die Beförderungsleistung auf. Der Zeitraum umfasst dabei die Jahre 1998 bis 2014. Jedes Jahr ist dabei in zwölf Monate unterteilt.

Neben dem Zeitraum, aus welchem die entsprechenden Werte stammen, sind die Daten in sechs Kategorien mit jeweils zwei Unterkategorien unterteilt. Die betrachteten Verkehre lauten dabei: Gewerblicher und Werkverkehr insgesamt, Binnenverkehr, Grenzüberschreitender Verkehr, Grenzüberschreitender Versand, Grenzüberschreitender Empfang und Kobotage. Die zwei Unterteilungen sind zum einen die beförderte Gütermenge in je 1.000 Tonnen, sowie die Beförderungsleistung in Millionen Tonnenkilometern. Der zweite Datensatz beinhaltet die gleichen Kategorien, jedoch aufgeteilt nach gewerblichen Verkehr und Werkverkehr. Laut Bundesamt für Güterverkehr handelt es sich bei einem Transport um Werkverkehr, „wenn die Beförderung für eigene Zwecke mit Kraftfahrzeugen, die einschließlich Anhänger ein höheres zulässiges Gesamtgewicht als 3,5 Tonnen haben, durchgeführt werden“ (Bundesamt für Güterverkehr 2017). Zudem müssen weitere Kriterien erfüllt sein. Die transportierten Güter müssen in einem direkten Zusammenhang zum Unternehmen stehen (Güter werden vom Unternehmen gekauft, verkauft, gemietet usw.). Der Transport der Güter hängt dabei ebenso vom Unternehmen ab (Versand, Anlieferung usw.). Schließlich muss der Transport mit eigenen Fahrzeugen und eigenem Personal erfolgen und darf maximal als Hilfstätigkeit im Unternehmen angesehen werden (Bundesamt für Güterverkehr 2017). Wird eine oder mehrere der genannten Bedingungen nicht erfüllt, so handelt es sich auch nicht um Werkverkehr, sondern um gewerblichen Verkehr.

Die Erhebung der vorgestellten Daten erfolgt durch Unternehmensbefragungen. Dabei handelt es sich um eine repräsentative Stichprobe die auf maximal fünf Promille der Gundgesamtheit zurückgreift. Befragt werden Unternehmen, die „gewerblichen Straßengüterverkehr oder Werkverkehr im Sinne des Güterkraftverkehrsgesetzes betreiben“ (Destatis 2017a)

Die Auskunft muss dabei vom Fahrzeughalter oder -besitzer gemäß § 868 BGB erteilt werden. Der Zeitraum welcher in der Befragung betrachtet wird ist die sogenannte „Halbwoche“. Diese bietet zwei Betrachtungszeiträume.

Der erste Zeitraum umfasst alle Transporte von Sonntagabend 22.00 Uhr bis Mittwochabend 24.00 Uhr, während der zweite Zeitraum die Zeitspanne von Donnerstagmorgen 0.00 Uhr bis Sonntag 22.00 beinhaltet. Wird ein Fahrzeug für die Befragung ausgesucht, so ist es für ein Jahr gesperrt in Bezug auf weitere Erhebungen. Ausnahme bilden hierbei Fahrzeuggruppen die im Vergleich unterbesetzt sind. Die Untersuchung bezieht sich generell nur auf inländische Kraftfahrzeuge (Destatis 2017a). Dies zeigt zugleich einen Nachteil der Statistik auf, da ausländische Fahrzeuge nicht erfasst werden.

4.2 Auswahl der Prognosemethode [KT]

Da das Ziel der Arbeit darin besteht, anstelle von Verkehrsmodellen klassische Prognosemethoden auf ihre Eignung im Verkehrswesen zu überprüfen, wird auch ein solches Verfahren hier verwendet. Dabei ist es zunächst relevant zu wissen, wie sich der Verlauf der Daten in der Vergangenheit entwickelt hat, also ob diese z.B. einem Trend oder einer Saisonalität folgen. Hierfür bietet sich die Ermittlung des Störpegels an (Buzacott et al. 2010, S. 65).

Der Störpegel wird durch die Formel $SP = \frac{MAA}{\hat{x}}$ beschrieben, wobei für die mittlere absolute Abweichung MAA gilt: $MAA = \frac{1}{T} \cdot \sum_{t=1}^T |x_t - \hat{x}|$, während $\hat{x} = \frac{1}{T} \cdot \sum_{t=1}^T x_t$ das arithmetische Mittel beschreibt. Ist $SP > 0,5$, dann gilt die Zeitreihe als unregelmäßig (Buzacott et al. 2010, S. 65).

Zunächst wird der Datensatz, in dem keine Unterscheidung zwischen gewerblichen Verkehr und Werkverkehr stattfindet, angepasst. Für jeden Verkehrsweg wird erst das arithmetische Mittel mit der Excel-Formel „MITTELWERT“ und die mittlere absolute Abweichung mit der Formel „MITTELABW“ gebildet. Anschließend wird der Störpegel ermittelt, indem der Quotient der beiden ersten Werte gebildet wird. Analog dazu wird auch der Datensatz, der den Unterschied zwischen gewerblichem Güterverkehr und Werkverkehr beschreibt, bearbeitet.

Wird der Störpegel betrachtet, ist erkennbar, dass dieser zwischen 0,1481 und 0,1705 bei dem Datensatz, der auf die Unterscheidung zwischen dem gewerblichen Güterverkehr und dem Werkverkehr verzichtet, schwankt. Im anderen Datensatz schwankt dieser zwischen 0,1698 und 0,507. Die obere Schranke gehört zu den Werten der Kabotage im Werkverkehr. Da hier sehr wenig Daten vorliegen und somit nicht aussagekräftig sind, wird dieser Datensatz vernachlässigt. Somit bewegen sich alle Störpegel unterhalb der Grenze von 0,5, weshalb eine gewisse Saisonalität in bestimmten Monaten angenommen werden kann.

Damit eignet sich die Zeitreihendekomposition zur Prognose des Güterverkehrs nach 2013.

4.3 Implementierung des Verfahrens in MS Excel [KT]

Excel bietet viele Funktionen, die als Hilfsmittel für Fragestellungen, wie die hier betrachtete, verwendet werden können. Die Verwendung von MS Excel erfordert keine tiefgehenden Informatik- oder Mathematikkenntnisse und das Programm hat, verglichen mit Instrumenten zum Programmieren, eine breite Nutzerbasis. Demnach sollte ein Tool, welches mit Excel entwickelt wird, auch einfach zu handhaben sein. Ziel soll also sein, das Tool übersichtlich und leicht verständlich zu gestalten. Um das zu erreichen, sollte die Anzahl der Arbeitsblätter und Hilfsmittel gering gehalten werden. Die Rechnungen sollten nachvollziehbar sein, damit geringfügige Anpassungen auch nachträglich vom Nutzer vorgenommen werden können.

Da es einen Unterschied zwischen der multiplikativen und der additiven Zeitreihendekomposition gibt, werden hier beide Verfahren implementiert, um einen Vergleich ziehen zu können. Es werden ähnlich genaue Prognosewerte erwartet.

Zunächst erfolgt im ersten Arbeitsblatt die Dateneingabe. Diese gilt für das gesamte Tool. Damit das Tool für beliebige Daten verwendet werden kann, werden den verwendeten Datensätzen Werte für die Periode t zugeteilt. So ist bei den verwendeten Daten der Januar 1998 $t = 1$, und so entspricht z.B. $t = 165$ dem Monat September im Jahr 2011. Zusätzlich ist eine nochmalige Berechnung des Störpegels (s. Abschnitt 4.2) zu finden, um zu beurteilen, ob die Zeitreihendekomposition für jeden beliebigen Datensatz, der eingepflegt wird, geeignet ist.

Im nächsten Arbeitsblatt wird der Datensatz, der eine Zeitreihe darstellt, in einem Diagramm abgebildet. Es werden nur die Realwerte wiedergegeben. Hierbei kann auch noch mal die Saisonalität sowie ein möglicher Trend grafisch überprüft werden.

Im Arbeitsblatt „Zeitreihendek. (mult.)“ befindet sich der Kernteil des Tools. Die Daten werden aus dem Arbeitsblatt „Dateneingabe“ übernommen. In einer Zelle wird der Saisonfaktor $Q = 12$ eingetragen. Da das Jahr 2014 nur als Vergleichswert gilt, werden für das Tool 192 Daten berücksichtigt. Die dritte Spalte gibt den berechneten Prognosewert an. Dieser ist unmittelbar neben der Spalte der Realwerte platziert, damit diese direkt verglichen werden können. Die folgenden Spalten beschreiben die Glättungskomponente, die irreguläre Komponente, die Trendfunktion und den saisonalen Faktor.

Da eine gerade Saisonperiodenanzahl vorliegt, wird die folgende Formel für den Glättungsfaktor verwendet: $\tilde{x}_t = \frac{1}{Q} \cdot \left(\frac{x_{t-Q \cdot 0,5} + x_{t+Q \cdot 0,5}}{2} + \sum_{t'=t-Q}^{t+Q'} x_{t'} \right)$, mit $Q' = \frac{Q}{2} - 1$. Somit ist die erste Glättungskomponente die, die für den Juli des Jahres 1998 ($t = 7$) gilt. Außerdem endet die Glättungskomponente bei $t = 186$, da die Realwerte für das Jahr 2014 nicht mehr berücksichtigt werden.

Die irreguläre Komponente bei der multiplikativen Zeitreihendekomposition wird berechnet durch $s_{yt} = \frac{x_t}{\tilde{x}_t}$, ebenfalls nur für $t = 7, \dots, 186$. Bevor die Trendfunktion und die saisonale Komponente bestimmt werden, werden die irregulären Komponenten in eine Matrix übertragen. In dieser Matrix stehen die Jahre in den Spalten und die einzelnen Monate in den Zeilen. Somit werden die einzelnen irregulären Komponenten eines Monats für alle Jahre in einer Zeile abgebildet. In zwei neuen Spalten wird aus diesen Komponenten die Summe pro Zeile und der Mittelwert über die verfügbaren Daten gebildet. Obwohl im $a = 16$ gilt, wird letzterer nur über 15 Werte gebildet, da in der ersten Spalte sechs Werte des ersten Halbjahres und in der letzten Spalte die letzten sechs Werte nicht berechnet werden konnten. Wenn die Summe nicht $Q=12$ ergibt, müssen die einzelnen Werte der nicht-standardisierten saisonalen Komponente durch die Formel $s_q = sy_q \cdot \frac{Q}{\sum_{q=1}^Q sy_q}$ bereinigt werden. Die Summe aus diesen Werten ergibt Q . Damit werden im nächsten Schritt die Saisonfaktoren übertragen. Der Wert aus der ersten Zeile der standardisierten saisonalen Komponente gilt somit immer für den Januar eines jeweiligen Jahres, der zweite Wert für den Februar usw.

Die Berechnung der Trendfunktion kann an beliebiger Stelle nach der Bestimmung der Glättungskomponente erfolgen. Es wird ein Diagramm aus den Daten der Glättungskomponente gebildet, welches im Arbeitsblatt „D Trendlinie“ zu finden ist. Dieses wird durch eine Trendlinie ergänzt. Bei der Implementierung des Tools standen mehrere Trendlinienfunktionen zur Auswahl, von denen die Beste für das Tool gewählt worden ist. Durch Vergleich der einzelnen Prognosefehler lässt sich eine Auswahl bezüglich der verwendeten Trendfunktion treffen. Die für den weiteren Verlauf zu verwendete Trendlinienfunktion wird ausgewählt aus einer linearen Trendlinie, Polynomfunktionen zweiten, dritten und vierten Grades und einer exponentiellen Trendline. Die Funktionen, sowie das Diagramm der Glättungskomponenten sind in Abb. 6 abgebildet.

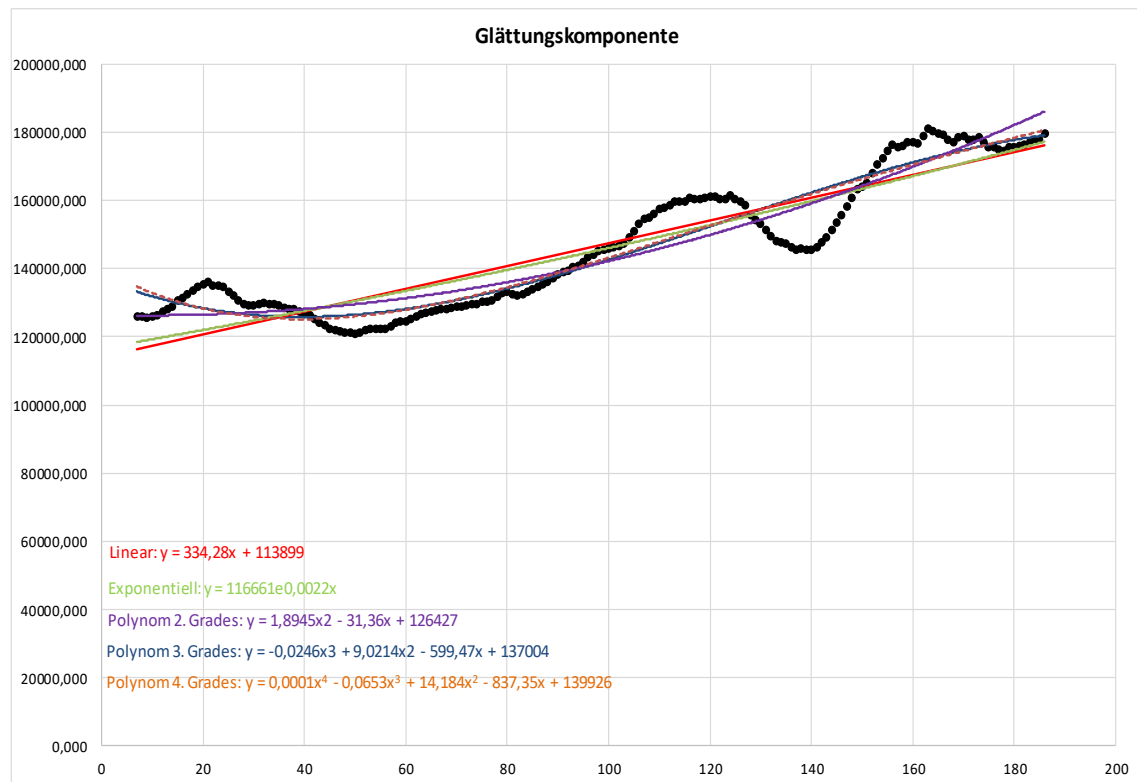


Abb. 6: Beispielhafte Darstellung mehrerer Trendlinien

Die mittlere absolute Abweichung wird analog zur Bestimmung des Störpegels bestimmt. Allerdings wurde zunächst der Betrag zwischen dem Realwert und dem prognostizierten Wert und daraus der Mittelwert bestimmt. Da die Ermittlung des Störpegels nur eine schnelle Überprüfung ist, ob die Zeitreihendekomposition für die benötigten Zwecke geeignet ist, wurde dieser in Abschnitt 4.2 zusammengefasst ermittelt. Bei der tatsächlichen Prognoserechnung ist es besser, diese Berechnung aufzuteilen, damit mögliche Fehler sowie Unstimmigkeiten in den Datenreihen schnell ermittelt werden können. Zusätzlich werden hier auch die mittlere quadratische Abweichung und der Ungleichheitskoeffizient nach Theil ermittelt, die eine Aussage über die Güte der Prognose liefern können. Für die MAA wird die Differenz aus Realwert und prognostiziertem Wert quadriert. Der TU wird berechnet, indem noch eine Spalte für die naive Prognose aus Vorgänger und Realwert gebildet und dieser Wert quadriert wird. Diese Datenreihen können nun zu den drei Werten für den Prognosefehler zusammengefasst werden. Die mittlere quadratische Abweichung wird durch den Mittelwert aller quadrierten Differenzen bestimmt, während beim TU die Wurzel aus dem Quotienten der Summe der quadrierten naiven Prognose und der Summe der quadrierten Differenz das benötigte Ergebnis liefert.

Da bei der Polynomfunktion 3. Grades mit $y = -0,0246 \cdot x^3 + 9,0214 \cdot x^2 - 599,47 \cdot x + 137004$ die Prognosefehler für den für die Implementierung beispielhaften verwendeten Datensatz am geringsten sind und der TU Wert nur 0,57702 beträgt, wird zunächst diese Trendlinie für das Tool gewählt. Die zweitbeste Prognose liefert die Polynomfunktion 4. Grades, die lineare und exponentielle Trendlinie werden nicht mehr weiter betrachtet. Dabei ist die Trendfunktion nur von Periode t abhängig. Der prognostizierte Wert wird nun aus dem Produkt aus der Trendfunktion und dem Saisonfaktor ermittelt (s. Abschnitt 3.2.2) Die Vorhersagequalität des Tools für die jeweils verwendeten Werte kann nun über die Prognosefehler (s.o.) abgelesen werden.

Für das additive Modell verläuft die Implementierung analog. Die Formeln für die Berechnung der irregulären Einflüsse, zur Berechnung der standardisierten saisonalen Komponente und des Prognosewertes wurden jeweils durch die Formel der additiven Berechnung (siehe Abschnitt 3.2.2) ersetzt.

In den Arbeitsblättern „D Zeitreihendek. (mult.)“ und „D Zeitreihendek. (add.)“ sind jeweils die beiden Zeitreihen mit ihren Realwerten und den prognostizierten Werten abgebildet. Ist eine Überlappung erkennbar, so ist eine hohe Güte der Prognose ebenfalls erkennbar.

4.4 Auswertung der Daten [MR]

Um die in Abschnitt 4.1 beschriebenen Datensätze mit dem in Abschnitt 4.3 vorgestelltem Tool auswerten zu können, werden lediglich zwei Arbeitsschritte benötigt. Nachdem die Datenreihe eingefügt wurde, wird diese automatisch grafisch ausgewertet. Dieser Graph liefert im Anschluss die Trendlinie sowie deren Funktionsgleichung, welche in der multiplikativen und der additiven Zeitreihendekomposition übernommen werden muss. Der Grad der polynomischen Gleichung kann dabei zwischen dem zweiten Grad und dem vierten Grad verändert werden. Höhere Grade sind unbrauchbar, da sie keine verwertbaren Ergebnisse liefern. Danach muss die gewählte Gleichung in der multiplikativen und der additiven Zeitreihendekomposition eingefügt werden. In der Spalte „Trend“ wird die Gleichung eingefügt, mit einem Zellbezug zur Zeitreihe, also zur Spalte „t“, versehen und auf die danach folgenden Zellen innerhalb der Spalte „Trend“ kopiert. Das Tool passt schließlich alle anderen Gleichungen entsprechend an und gibt sowohl die Prognosewerte, als auch die bereits erläuterten Werte der mittleren absoluten Abweichung, der mittleren quadratischen Abweichung sowie Theil's Ungleichheitskoeffizient wieder.

Um einen Datensatz verwenden zu können, muss dieser lückenlos und vollständig sein, da sonst keine genauen Werte ermittelt werden können und das Tool verfälschte Ergebnisse liefert. Aufgrund dieses Kriteriums konnten einige Datenreihen nicht verwendet werden. Bei insgesamt drei Datenreihen fehlen zwischen einem und vier der insgesamt 204 Werte. Um eine Auswertung dennoch zu gewährleisten, werden diese durch die Ermittlung der prozentualen Veränderung der Vorperioden und deren anschließende Verrechnung mit dem Wert des Vorjahres ergänzt.

Die Werte können dabei nicht einfach ausgelassen werden, da selbst das Fehlen eines Wertes zu starken Schwankungen aller Variablen führt. Somit können von den am Anfang 14 zur Verfügung stehenden Datenreihen zehn ausgewertet werden. Es werden aus den in Abschnitt 4.1 vorgestellten Datensätzen nur jene Datenreihen verwendet, die sich inhaltlich mit der beförderten Gütermenge befassen. Die Menge der beförderten Güter lässt eher Rückschlüsse auf die für den Transport notwendigen Fahrzeuge zu, als die Beförderungsleistung. Die Ergebnisse der Auswertung sind wie folgt dargestellt:

| | | Zeitreihendekomposition - Multiplikativ | | | | | | | | |
|-----------------------------|---------------------------------------|---|---------|---------|----------------------------------|--------|---------|---------------------------------|---------|---------|
| | | TU-Wert | | | Prozentuale Abweichung Insgesamt | | | Prozentuale Abweichung Prognose | | |
| | | Grad 2 | Grad 3 | Grad 4 | Grad 2 | Grad 3 | Grad 4 | Grad 2 | Grad 3 | Grad 4 |
| Verkehre Gesamt | gew.Verkehr und Werkverkehr insgesamt | 0,45040 | 0,44857 | 0,44630 | 5,955% | 5,943% | 5,941% | 5,755% | 5,802% | 6,312% |
| | Binnenverkehr | 0,44206 | 0,44124 | 0,44052 | 5,953% | 5,945% | 5,993% | 5,882% | 5,881% | 6,629% |
| | Grenzüberschreitender Verkehr | 0,68130 | 0,67360 | 1,45526 | 8,017% | 7,849% | 13,400% | 5,429% | 9,750% | 36,825% |
| | Grenzüberschreitender Versand | 0,66298 | 0,65751 | 1,90869 | 7,801% | 7,604% | 14,732% | 6,272% | 7,110% | 41,549% |
| | Grenzüberschreitender Empfang | 0,77464 | 0,78571 | 1,45212 | 8,414% | 8,816% | 13,107% | 4,072% | 17,120% | 32,920% |
| | Kabotage | - | - | - | - | - | - | - | - | - |
| Gewerblicher Verkehr | Verkehr Insgesamt | 0,48490 | 0,47157 | 0,46938 | 6,070% | 5,871% | 5,887% | 5,399% | 5,121% | 5,173% |
| | Binnenverkehr | 0,47863 | 0,46557 | 0,46424 | 6,136% | 5,951% | 5,974% | 5,770% | 5,345% | 5,288% |
| | Grenzüberschreitender Verkehr | 0,71847 | 0,72084 | 0,98431 | 7,728% | 7,459% | 9,333% | 5,593% | 4,948% | 20,979% |
| | Kabotage | - | - | - | - | - | - | - | - | - |
| Werkverkehr | Verkehr Insgesamt | 0,51649 | 0,52015 | 0,51714 | 8,273% | 8,159% | 8,153% | 9,830% | 7,187% | 8,950% |
| | Binnenverkehr | 0,51269 | 0,51661 | 0,52043 | 8,284% | 8,157% | 8,488% | 9,886% | 6,896% | 12,833% |
| | Grenzüberschreitender Verkehr | - | - | - | - | - | - | - | - | - |
| | Kabotage | - | - | - | - | - | - | - | - | - |

Tabelle 3: Zeitreihendekomposition Multiplikativ

| | | Zeitreihendekomposition - Additiv | | | | | | | | |
|------------------------------|---|-----------------------------------|---------|---------|-------------------------------|--------|--------|---------------------------------|--------|--------|
| | | TU-Wert | | | Prozentuale Abweichung Gesamt | | | Prozentuale Abweichung Prognose | | |
| | | Grad 2 | Grad 3 | Grad 4 | Grad 2 | Grad 3 | Grad 4 | Grad 2 | Grad 3 | Grad 4 |
| Verkehre Gesamt | gew.Verkehr und Werkverkehr insge- samt | 0,45022 | 0,44868 | 0,44744 | 5,93% | 5,93% | 5,97% | 5,12% | 5,23% | 6,00% |
| | Binnenverkehr | 0,44286 | 0,44219 | 0,44297 | 5,95% | 5,95% | 6,05% | 5,10% | 5,14% | 6,30% |
| | Grenzüberschreitender Verkehr | 0,69192 | 0,68533 | 1,45549 | 8,41% | 8,27% | 13,77% | 5,71% | 9,98% | 36,96% |
| | Grenzüberschreitender Versand | 0,67017 | 0,66586 | 1,90677 | 8,14% | 7,94% | 15,06% | 6,37% | 7,23% | 41,67% |
| | Grenzüberschreitender Empfang | 0,77905 | 0,79082 | 1,45160 | 8,59% | 8,99% | 13,24% | 4,28% | 17,46% | 32,99% |
| | Kabotage | - | - | - | - | - | - | - | - | - |
| Gewerbli- cher Verkehr | Verkehr Insgesamt | 0,49589 | 0,48498 | 0,48385 | 6,18% | 6,03% | 6,03% | 4,28% | 4,24% | 4,47% |
| | Binnenverkehr | 0,49041 | 0,48016 | 0,48042 | 6,25% | 6,16% | 6,18% | 4,38% | 4,35% | 4,56% |
| | Grenzüberschreitender Verkehr | 0,72568 | 0,72896 | 0,98951 | 7,90% | 7,75% | 9,63% | 5,87% | 5,20% | 21,06% |
| | Kabotage | - | - | - | - | - | - | - | - | - |
| Werkverkehr | Verkehr Insgesamt | 0,57712 | 0,58019 | 0,57767 | 10,08% | 10,13% | 10,31% | 19,80% | 18,07% | 19,76% |
| | Binnenverkehr | 0,57553 | 0,57882 | 0,58176 | 10,13% | 10,18% | 10,73% | 19,97% | 18,11% | 23,33% |
| | Grenzüberschreitender Verkehr | - | - | - | - | - | - | - | - | - |
| | Kabotage | - | - | - | - | - | - | - | - | - |

Tabelle 4: Zeitreihendekomposition Additiv

Der Aufbau beider Tabellen ist identisch. Die Zeilen sind aufgeteilt auf die Unterkategorien der zwei verwendeten Datensätze. Protokolliert werden dabei der TU, die prozentuale Abweichung der gesamten Datenreihe im Vergleich zu den prognostizierten Werten, sowie die prozentuale Abweichung bezogen auf eine beispielhafte Prognose.

Der TU-Wert gibt wie bereits im Vorfeld erläutert das Verhältnis zwischen der vorliegenden Prognose und der „naiven“ Prognose wieder. Je mehr dieser Wert gegen null geht, desto besser ist die vorliegende Prognose. Die „Prozentuale Abweichung Gesamt“ gibt den Mittelwert der prozentualen Abweichung der verwendeten Datenreihe im Vergleich zu den durch das Tool berechneten Werten wieder. Um eine Prognose für einen Zeitraum zu simulieren dienen die Werte des Jahres 2014 bei allen Datenreihen als Beispiel. Die „Prozentuale Abweichung Prognose“ gibt die Unterschiede der zwischen Prognose und jenen Werten wieder.

Der Vergleich der prozentualen Abweichungen wird dem der absoluten sowie quadratischen Abweichung vorgezogen. Da die Datensätze bezogen auf die Größenordnung der Werte sich stark unterscheiden, ist der Vergleich der absoluten sowie quadratischen Abweichung zwar möglich, jedoch veranschaulicht nicht direkt die Aussage über ihre Güte.

Durch die Variation des Polynomgrades der Trendlinienfunktion unterscheiden sich die berechneten Werte voneinander. Die jeweils grün unterlegten Zellen geben den geringsten und damit besten Wert der drei Formeln wieder. Der TU unterscheidet sich größtenteils kaum. Von 60 Werten weichen acht mit mehr als 15 Prozent von ihren Vergleichswerten in der gleichen Reihe ab. Bei allen anderen Reihenwerten liegt der Unterschied im Durchschnitt bei unter einem Prozent.

Die prozentuale Abweichung der gesamten Datenreihen in Bezug auf die Prognosewerte unterscheidet sich ebenfalls nur marginal. Die Differenz beträgt abgesehen von drei Werten stets unter zwei Prozent. Insgesamt beträgt die jeweils minimalste Abweichung nie mehr als zehn Prozent.

Die Werte der beispielhaften Prognose unterscheiden sich stärker als die der zwei vorangegangenen Datentypen. Dies ist auf den Datenumfang der dabei betrachtet wird zurückzuführen, welcher hier geringer ist und dadurch die Polynomgrade sich in Bezug auf die Genauigkeit stärker voneinander abgrenzen. Abgesehen vom Werkverkehr sind die prognostizierten Werte auch hier sehr nah an den realen Daten.

4.5 Fazit [KT]

Das in Abschnitt 4.3 implementierte Modell zeigt, dass die Zeitreihendekomposition sehr gute Prognosen liefern kann. Es stellt sich jedoch die Frage, wie das Modell in Hinblick auf die genutzten Daten geeignet ist. Dazu sollte auch der Zeitaufwand der Datenauswertung berücksichtigt werden.

Das vorgestellte Tool lässt sich anwenden, wenn Anpassungen vorgenommen werden können, d.h. dann, wenn das Wissen über zu tätige Anpassungen vorhanden ist und der Zeitaufwand für eine Anpassung in einem angemessenen Rahmen liegen.

Das Tool ist für Datenreihen vorgesehen, die eine Saison von 12 Monaten darstellen. So kann es nicht für saisonale Schwankungen alle vier Perioden (z.B. Quartal) oder alle sieben Perioden (z.B. Woche) verwendet werden. Jedoch ist dies kein Nachteil, da die Verwendung des Tools meist den Zweck hat, eine bestimmte Gruppe von Datenreihen auszuwerten.

Was jedoch definitiv bei jeder neuen Datenreihe angepasst werden muss, ist die Trendfunktion. Das wäre nicht nötig, wenn auch eine Regressionsanalyse, welche die Parameter bestimmt, implementiert wäre. Das gestaltet sich bei einer linearen Regression einfacher als bei polynomischen Trendfunktionen, sodass für dieses Tool die Hilfsmittel von Excel verwendet werden, um die jeweiligen Trendfunktionen aufgrund den sich jedes Mal veränderten Glättungskomponenten zu ermitteln. Die Funktion muss dann manuell neu eingetragen werden.

Des Weiteren müssen leere Felder eliminiert werden bzw. durch geschätzte Werte ersetzt werden. Viele Datensätze werden dadurch unbrauchbar, da sich unter den Vorgängerdaten z.T. nicht genügend Informationen für die Schätzung der Werte vorfinden lassen. Außerdem sind die berechneten Werte bereits geschätzte Werte und keine tatsächlichen Werte mehr. Die Prognose beruht im schlimmsten Fall auf Werten, die keine Realität abbildet. Letztendlich ist es zeitaufwändig, die Werte durchzugehen und jeden einzelnen fehlenden Wert „per Hand“ zu ermitteln.

In der Auswertung wurde jeweils nur das Jahr 2014 zum Vergleich der Prognosewerte hinzugezogen. Grundsätzlich ist das Tool aber auch für längere Prognosen geeignet, jedoch ist zu bedenken, dass die Prognose dann immer auf denselben Daten aufbaut. Je nachdem wie die Wirtschaftssituation in der zu untersuchenden Region aussah, kann das zu starken Beeinflussungen der Trendfunktion führen. Ein Vorschlag dem vorzubeugen, ist es, die Prognose rollierend zu gestalten. Als Richtwert wird die Zeitreihe für die nächsten 10-15 Jahre (120-180 Werte) prognostiziert und jedes Jahr, wenn wieder ein neuer Datensatz vorhanden ist, wird dieser in die Prognose eingebaut. Dafür wird der älteste Datensatz entfernt.

Die verwendeten Daten stellen nicht unmittelbar die Verkehrsnachfrage dar, sondern nur den Straßengüterverkehr. Aus der Gütermenge ist jedoch eine Schlussfolgerung auf das LKW-Aufkommen in den nächsten Jahren möglich. Für die Planung der Städte- und Autobahninfrastruktur oder Autobahnbrücken ist die beförderte Gütermenge eine bedeutende Kennzahl. Die Fragestellungen die sich daraus ableiten und die mithilfe eines solchen Verfahrens beantwortet werden können, sind ob Städte und Autobahnen mit einer zusätzlichen Belastung durch eine höhere Gütermenge und somit eine höhere Anzahl von LKW umgehen können und ob Autobahnbrücken eine solche Mehrbelastung tragen können.

Aufgrund dieser Datengrundlage eignen sich die Modelle, die in den Abschnitten 3.3 und 3.4 erläutert werden, nicht. Diese sind für tatsächliche Verkehrsdaten geeignet. Damit kann auf klassische Verfahren wie die Zeitreihendekomposition zurückgegriffen werden.

Die Untersuchung der klassischen Verfahren war das ursprüngliche Ziel der Arbeit, welches hiermit erfüllt wurde. Bei der Anwendung des Tools zu Prognosezwecken sind jedoch die bereits erläuterten Einschränkungen zu beachten.

5 Zusammenfassung und Ausblick [KT]

Ziel dieser Projektarbeit war es, ein Prognoseverfahren auf die Eignung im Verkehrswesen zu untersuchen.

Zunächst wurden Grundlagen zur Datenanalyse in Form einer geeigneten Datenbasis geschaffen. Nach einer Einführung und der Erläuterung der Punkte Datensammlung, Datenaufbereitung und Datenauswertung, besonders bei der Handhabung im Verkehrswesen, wurden auch methodische Grundlagen in MS Excel vorgestellt, da das Tabellenkalkulationsprogramm eine Basis für die Untersuchung des Prognoseverfahrens darstellt.

Weiter wurden Prognoseverfahren im Transportwesen betrachtet. Zunächst wurde ein Überblick über die Verkehrsplanung, bei welcher Verkehrsprognosen notwendig sind, gegeben. Es folgte eine Beschreibung der Trendprognose, wobei drei mathematische Verfahren detailliert dargestellt wurden. Der gleitende Durchschnitt als sehr einfaches Verfahren findet in ähnlicher Form Anwendung in der Zeitreihendekomposition. Ebenfalls wurde die exponentielle Glättung betrachtet, da dieses Verfahren einen Vorteil gegenüber dem gleitenden Durchschnitt hat. Die Zeitreihendekomposition als letztes Beispiel wurde in Kapitel 4 angewendet. Weiter wurde ein kurzer Einblick in die Verkehrsmodellierung und in Modelle gegeben, die aktuell für die kurzfristige Prognose von Verkehrsdaten angewendet wurde. Da diese Modelle nicht auf den vorhandenen Datensatz anwendbar sind, wurde auf eine detaillierte Betrachtung verzichtet.

Kapitel 4 gibt das Anwendungsbeispiel wieder. Es wurden Daten zur Güterverkehrsmenge im Straßenverkehr aus den Jahren 1998 bis 2013 verwendet, um eine Prognose für das Jahr 2014 zu ermitteln und diese mit den tatsächlichen Daten aus 2014 zu vergleichen. Dazu wurde der Verkehr in Binnenverkehr, grenzüberschreitenden Verkehr und Kabotage aufgeteilt. In einem zweiten Datensatz fand noch eine Unterscheidung in gewerblichen Verkehr und Werkverkehr statt. Anhand des Störpegels aller Zeitreihen wurde angenommen, dass die Zeitreihendekomposition ein geeignetes Verfahren ist, diese Prognose durchzuführen. Das Verfahren wurde in der multiplikativen und der additiven Form in MS Excel implementiert, danach wurden die Daten ausgewertet. Dabei wurde ausschließlich die Prognosegüte betrachtet und anhand von Theil's Ungleichheitskoeffizienten und der prozentualen Abweichung von den Prognosewerten im Vergleich zum ursprünglichen Datensatz, sowie zu dem Datensatz aus 2014, aufgezeichnet.

Dabei wurde festgestellt, dass die Prognosegüte sehr hoch ist, dennoch eignet sich das Verfahren mit dem Hilfsmittel MS Excel nur bedingt für die Prognose der Gütermenge, da bei jedem neu eingefügten Datensatz Anpassungen vorgenommen werden müssen.

Um dies zu verbessern, könnte eine Regressionsanalyse, die Parameter für die Polynomfunktionen zweiten, dritten und vierten Grades ermittelt, implementiert werden. Weiter können selbstlernende Algorithmen, wie sie auch von Lv et al. (2014) vorgestellt werden, verwendet werden, um Zeitreihen in der Zukunft zu prognostizieren. Die Modelle selbst basieren auf mathematischen Regeln und Verfahren, die bereits in der Vergangenheit getestet worden sind, sodass die Forschung sich darauf konzentriert diese Verfahren neu zu verknüpfen und dadurch die Prognosegüte zu erhöhen.

Da die Weltbevölkerung wächst (Bundeszentrale für politische Bildung 2017a) und Schwellenländer in Zukunft wohlhabender werden (Bundeszentrale für politische Bildung 2017b), wird auch der Bedarf an moderner Infrastruktur steigen. Daher wird die Prognose des Verkehrs auch in Zukunft eine große Rolle in Gesellschaft und Forschung spielen.

Literaturverzeichnis

Aberle, Gerd (2000): Transportwirtschaft. Einzelwirtschaftliche und gesamtwirtschaftliche Grundlagen. 3., überarb. Aufl. München: Oldenbourg.

Aleksić, Mario (2012): Dynamische Umlegung von Quelle-Ziel-Matrizen mit einem makroskopischen Verkehrsflussmodell. Wiesbaden: Vieweg+Teubner Verlag.

Bamberg, Günter; Baur, Franz; Krapp, Michael (2012): Statistik. 17., überarb. Aufl. München: Oldenbourg.

Bäumer, Marcus; Hautzinger, Heinz (Hg.) (2010): Ermittlung von Standards für anforderungsgerechte Datenqualität bei Verkehrserhebungen. [Bericht zum Forschungsprojekt FE 01.169/2007/CGB]. Bundesanstalt für Straßenwesen. Bremerhaven: Wirtschaftsverl. NW Verl. für neue Wiss (Berichte der Bundesanstalt für Straßenwesen Verkehrstechnik, 200).

Bleymüller, Josef; Weißbach, Rafael; Gehlert, Günther; Gülicher, Herbert (2015): Statistik für Wirtschaftswissenschaftler. 17. Aufl. München: Franz Vahlen. Online verfügbar unter <http://dx.doi.org/10.15358/9783800649617>.

Bundesamt für Güterverkehr (2017): Definition Werkverkehr. Online verfügbar unter <https://www.bag.bund.de/SharedDocs/FAQ/DE/Werkverkehr/Werkverkehr.html;jsessionid=9E5F94B014EA7DC72044F969D1E4D8B8.live21304?nn=13120>, zuletzt geprüft am 21.12.17.

Bundesministerium für Verkehr und digitale Infrastruktur (2015): Richtlinien für die Straßenverkehrszählung im Jahre 2015 auf den Bundesfernstraßen. Online verfügbar unter http://www.vm.nrw.de/verkehr/strasse/Strassenverkehr/Verkehrszaehlungen/Manuelle_Zaehlung/SVZ_2015.pdf, zuletzt geprüft am 21.12.17.

Bundeszentrale für politische Bildung (2017a): Bevölkerungsentwicklung. Online verfügbar unter <http://www.bpb.de/nachschlagen/zahlen-und-fakten/globalisierung/52699/bevoelkerungsentwicklung>, zuletzt geprüft am 21.12.17.

Bundeszentrale für politische Bildung (2017b): Schwellenländer. Online verfügbar unter <http://www.bpb.de/nachschlagen/lexika/lexikon-der-wirtschaft/20580/schwellenlaender>, zuletzt geprüft am 21.12.17.

Buzacott, John A.; Corsten, Hans; Gössinger, Ralf; Schneider, Herfried (2010): Produktionsplanung und -steuerung. Grundlagen, Konzepte und integrative Entwicklungen. München: Oldenbourg Verlag.

Chang, S. C.; Kim, R. S.; Kim, S. J.; Ahn, M. H. (2000): Traffic-flow forecasting using a 3-stage model. In: Proceedings of the IEEE Intelligent Vehicles Symposium 2000 (S. 451–456)

Chrobok, Roland (2005): Theory and Application of Advanced Traffic Forecast Methods. unv. Diss. Universität Duisburg-Essen.

Destatis (2017a): Berichte. Online verfügbar unter https://www-genesis.destatis.de/genesis/online/data;jsessionid=950086308054ACE5D54828598C8964AF.tomcatGO_1_1?operation=statistikLangtext&levelindex=0&levelid=1513808410115&index=2, zuletzt geprüft am 21.12.17.

Destatis (2017b): Tabellen. Online verfügbar unter https://www-genesis.destatis.de/genesis/online/data;jsessionid=950086308054ACE5D54828598C8964AF.tomcat_GO_1_1?operation=statistikAbruftabellen&levelindex=0&levelid=1513808528302&index=2, zuletzt geprüft am 21.12.17.

Dierstein, Rüdiger (2003): DIN 44 300. Begriffe - Definitionen und Erläuterungen. Ulmer Akademie für Datenschutz und IT-Sicherheit e.V. Oberpfaffenhofen. Online verfügbar unter <http://wwwbayer.in.tum.de/lehre/WS2003/ITS-dierstein/DefDV03.pdf>, zuletzt geprüft am 21.12.17.

Feiks, Markus (2016): Datenerhebung mit Excel. Eine Anleitung zur Umsetzung von Inhaltsanalysen und Befragungen. Wiesbaden: Springer VS.

Forschungsgesellschaft für Straßen- und Verkehrswesen (2012): Empfehlungen für Verkehrserhebungen. EVE. Ausg. 2012. Köln: FGSV-Verl. (FGSV R2 - Regelwerke, 125).

Gabler Wirtschaftslexikon (2017a): Big Data. Definition. Online verfügbar unter <http://wirtschaftslexikon.gabler.de/Archiv/-2046774198/big-data-v5.html>, zuletzt geprüft am 21.12.17.

Gabler Wirtschaftslexikon (2017b): Datenaufbereitung. Online verfügbar unter <http://wirtschaftslexikon.gabler.de/Archiv/57520/aufbereitung-v11.html>, zuletzt geprüft am 21.12.17.

Gabler Wirtschaftslexikon (2017c): Kleinste Quadrate. Online verfügbar unter <http://wirtschaftslexikon.gabler.de/Archiv/57785/kleinstquadratmethode-gewoehnliche-v14.html>, zuletzt geprüft am 21.12.17.

Gabler Wirtschaftslexikon (2017d): Nominalskala. Online verfügbar unter <http://wirtschaftslexikon.gabler.de/Archiv/11049/nominalskala-v9.html>, zuletzt geprüft am 21.12.17.

Gabler Wirtschaftslexikon (2017e): Tabellenkalkulationssystem. Online verfügbar unter <http://wirtschaftslexikon.gabler.de/Archiv/56394/tabellenkalkulationssystem-v12.html>, zuletzt geprüft am 21.12.17.

Geo-Basis-DE / BKG (2017): Anzahl Dauerzählstellen. Online verfügbar unter http://www.bast.de/DE/Verkehrstechnik/Fachthemen/v2-verkehrszaehlung/Aktuell/zaehl_aktuell_node.html, zuletzt geprüft am 21.12.17.

Han, Chao; Song, Su (2003): A review of some main models for traffic flow forecasting. In: *IEEE Trans. Intell. Transport. Syst.*, S. 216-219.

Hautzinger, Heinz; Fuchs, Mario; Schmidt, Jochen; Stock, Wilfried: Laufende Hochrechnung der Kfz- Fahrleistungen auf der Basis von Hauptuntersuchungsdaten. Institut für angewandte Verkehrs- und Tourismusforschung. Online verfügbar unter https://www.kba.de/DE/Statistik/Kraftverkehr/hochrechnung_hu_daten_pdf.pdf?__blob=publicationFile&v=1.

Hillier, Frederick S.; Lieberman, Gerald J.; Bauer, Gabriele (2002): Operations Research. Einführung. 5. Aufl., München: Oldenbourg.

Jahn, Detlev; Sarcander, Steve; Wagner, Ralf (2004): Datenerhebung. Frankfurt a.M.: Dt. Fachverl. (p&a Lehrbuch, Bd. 1).

Köhler, Uwe (2014): Einführung in die Verkehrsplanung. Grundlagen, Modellbildung, Verkehrsprognose, Verkehrsnetze. Stuttgart: Fraunhofer IRB Verlag.

Lippe, Peter von der (1993): Deskriptive Statistik. Stuttgart: Fischer.

Ly, Yisheng; Duan, Yanjie; Kang, Wenwen; Li, Zhengxi; Wang, Fei-Yue (2014): Traffic Flow Prediction With Big Data. A Deep Learning Approach. In: *IEEE Trans. Intell. Transport. Syst.*, Vol. 16 (2), S. 1-9.

Martin Treiber: Datengewinnung: Verkehrserhebungen. Online verfügbar unter http://www.mtreiber.de/Vkoek_Ma_Skript/vkoek_Teil3.pdf, zuletzt geprüft am 22.10.2017.

Microsoft Office (2017a): Gleichungen zum Berechnen von Trendlinien. Online verfügbar unter <https://support.office.com/de-de/article/Gleichungen-zum-Berechnen-von-Trendlinien-12cfdaa5-0652-436f-839c-0561e8620ba5>, zuletzt geprüft am 21.12.17.

Microsoft Office (2017b): Hinzufügen einer Trend- oder Durchschnittslinie zu einem Diagramm. Online verfügbar unter <https://support.office.com/de-de/article/Hinzufügen-einer-Trend-oder-Durchschnittslinie-zu-einem-Diagramm-fa59f86c-5852-4b68-a6d4-901a745842ad>, zuletzt geprüft am 21.12.17.

Müller, Michael; Poguntke, Werner (2015): Basiswissen Statistik. 2. Auflage. Dortmund: W3L-Verlag.

Onpulson Wirtschaftslexikon (2017): Datenaufbereitung. Online verfügbar unter <http://www.onpulson.de/lexikon/datenaufbereitung/>, zuletzt geprüft am 21.12.17.

Ortúzar, Juan de Dios; Willumsen, Luis G. (2011): Modelling Transport. Chichester, UK: John Wiley & Sons, Ltd.

Pepels, Werner; Bruns, Jürgen (Hg.) (2008): Verfahren, Datenauswertung, Ergebnisdarstellung. 2. Aufl. Düsseldorf: Symposion Publ.

Puhani, Josef (2012): Statistik. Einführung mit praktischen Beispielen, 12. Aufl. Eibelsstadt: Lexika-Verlag.

Raab-Steiner, Elisabeth; Benesch, Michael (2012): Der Fragebogen. Von der Forschungsidee zur SPSS-Auswertung. 3., aktualisierte und überarb. Aufl. Wien: Facultas-Verl. (UTB Schlüsselkompetenzen, 8406). Online verfügbar unter <http://www.utb-studie-book.de/9783838584966>.

Schnabel, Werner; Lohse, Dieter (2011): Grundlagen der Straßenverkehrstechnik und der Verkehrsplanung. Band 2 ; Verkehrsplanung. 3. Aufl. Berlin, Wien, Zürich: Beuth Verlag GmbH.

Schwarzbauer, Peter (2017): Methoden der Datenerhebung. Universität für Bodenkultur Wien. Wien. Online verfügbar unter http://www.wiso.boku.ac.at/fileadmin/data/H03000/H73000/H73500/_TEMP_/Lehrmaterial/Datenerhebung/datenerhebung_kap5.pdf, zuletzt geprüft am 21.12.17.

Gabler Wirtschaftslexikon: Definition Verkehrsmodelle. Gabler Wirtschaftslexikon. Online verfügbar unter <http://wirtschaftslexikon.gabler.de/Archiv/78687/verkehrsmodelle-v6.html>, zuletzt geprüft am 14.11.2017.

Stock, Wilfried; Bernecker, Tobias (2014): Verkehrsökonomie. Eine volkswirtschaftlich-empirische Einführung in die Verkehrswissenschaft. 2. Aufl. 2014. Wiesbaden: Springer Gabler.

Theil, Henri (1966): Applied economic forecasting., 2. Aufl., Amsterdam: North-Holland Publ. Co.

Wagner-Muns, Isaac Michael; Guardiola, Ivan G.; Samaranyake, V. A.; Kayani, Wasim Irshad (2017): A Functional Data Analysis Approach to Traffic Volume Forecasting. In: *IEEE Trans. Intell. Transport. Syst.* (99), S. 1–11.

Wies, Peter (2013): Excel 2013 Grundlagen. Neu: HERDT BuchPlus. Bodenheim: Herdt.
Online verfügbar unter https://herdt-campus.de/product/EX2013_NEU.