

# Quantitativer Einfluss gesellschaftlicher Trends auf den Fahrzeugabsatz in Deutschland

Christian Rühl

Institut für Automobil Forschung (im RIF e.V.), Joseph-von-Fraunhofer-Straße 20,  
44227 Dortmund, C.Ruehl@automobil-forschung.org

Prof. Dr. Niels Biethahn

Institut für Automobil Forschung (im RIF e.V.) und Professor für Unternehmenssteuerung und Projekt Manager für das Thema Automotive Management, BITS Business and Information Technology School GmbH, Reiterweg 26b, 58535 Iserlohn,  
Niels.Biethahn@bits-iserlohn.de

1	Einleitung .....	193
2	Künstliche Neuronale Netze als Instrument für die Prognose des Fahrzeugabsatzes.....	195
3	Zusammenfassung der Ergebnisse und Ausblick.....	208
4	Literatur .....	210

## Abstract

*In der vorliegenden Ausarbeitung wird untersucht, welchen quantitativen Einfluss eine umfangreiche Anzahl qualitativer Einflussfaktoren auf den Fahrzeugabsatz pro Fahrzeugklasse in Deutschland hat. Aufbauend dafür werden die Prognosefähigkeiten von ausgewählten Verfahren der Künstlichen Neuronalen Netze beschrieben und mit herkömmlichen statistischen Methoden aus Vergleichsuntersuchungen dargestellt. Ausgehend von einer vorausgegangen Arbeit der Autoren wird mit Hilfe von einem Künstlichen Neuronalen Netz der spezifische Fahrzeugabsatz auf Grundlage von umfangreichen Datensätzen prognostiziert<sup>1</sup>. Die Untersuchung bestätigt, dass Künstliche Neuronale Netze für Prognosen verwendet werden können, sofern der Datensatz für einen unterunterbrochenen und langfristigen Zeitraum zur Verfügung steht. Darüber hinaus zeigt die Untersuchung, dass bei komplexen Problemen mit vielen Einflussfaktoren eine große Anzahl von Eingangsvariablen genutzt und mit Hilfe von weiteren Verfahren fallweise auf latente Variable übergeleitet werden können.*

*JEL Classification: C53 ( Forecasting and Prediction Methods; Simulation); L62 (Automobiles; Other Transportation Equipment)*

*Keywords: Künstliche Neuronale Netze (KNN), Automotive, Fahrzeugabsatz, Prognosetool*

---

<sup>1</sup> Ein besonderer Dank gilt Herrn Dr. Frank Buckler von NEUSREL Causal Analytics für die fachliche und simulationsgesteuerte Unterstützung bei der Analyse der Datensätze im Künstlichen Neuronalen Netz.

## 1 Einleitung

### 1.1 Ausgangspunkt der Untersuchung

Ausgangspunkt dieser Untersuchung ist das Paper von Rühl und Biethahn (2013), in dem gesellschaftlichen Trends untersucht wurden, die einen generellen Einfluss auf die weltweite Automobilbranche haben. Aufbauend auf einer Vielzahl qualitativer und quantitativer Analysen wurden während der Untersuchung Einflussfaktoren auf den weltweiten Besitz privater PKW ermittelt. Diesen Einflussfaktoren wurden schlussendlich die Eigenschaft zugeschrieben, entweder einen negativen oder positiven Einfluss Besitz privater PKW zu haben (siehe Abbildung 20).

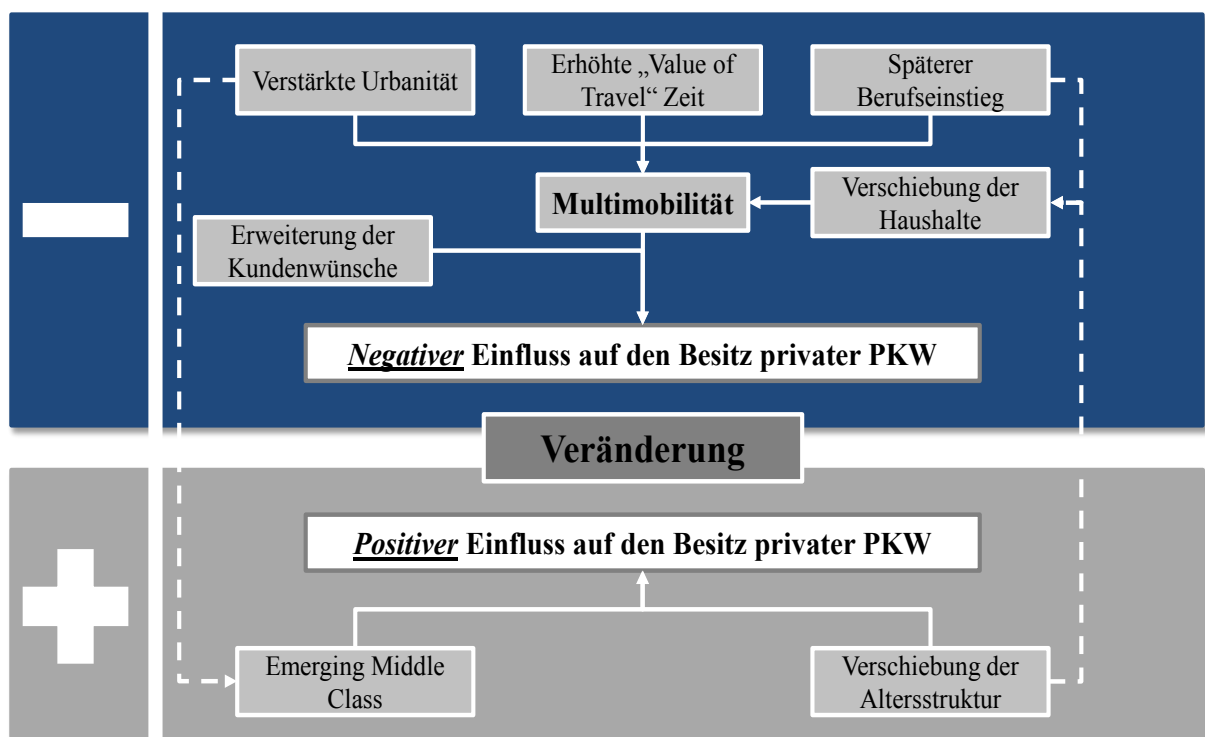


Abbildung 20: Grafische Zusammenfassung der Studie "Einfluss gesellschaftlicher Trends auf die Automobilbranche"<sup>2</sup>

Die vorliegende Untersuchung nutzt diese Erkenntnisse und setzt auf den Ergebnissen auf. Ziel der vorliegenden Untersuchung ist es, Rückschlüsse über die reale Höhe der spezifischen Einflussfaktoren auf den Fahrzeugabsatz zu ziehen und eine rückwirkende als auch eine zukünftige Betrachtung zu ermöglichen.

In jeder Hinsicht sind Prognosen als Planungsgrundlage für strategische Entscheidungen in marktorientierten Unternehmen zu verstehen. Deshalb sollten Prognosen nicht

<sup>2</sup> Rühl und Biethahn (2013).

auf intuitiven ökonomischen Schätzungen beruhen, sondern auf Modellen mit möglichst hoher Prognosegenauigkeit.<sup>3</sup>

### 1.2 Herangehensweise

Aufgrund der Vielzahl von gesellschaftlichen Trends und deren Auswirkungen auf die Einflussfaktoren sowie einem bisher nicht wissenschaftlich definierten quantitativen Kausalzusammenhang der Trends auf den Fahrzeugabsatz, konnte eine schlüssige Festlegung von Zusammenhängen und deren Ableitung der Zukunftsprognosen nicht systematisch festgelegt werden (siehe Abbildung 21). Die Vielzahl der gesellschaftlichen Trends und deren Auswirkung auf einen oder mehrere Einflussfaktoren stellt dabei eine wesentliche Herausforderung dar.

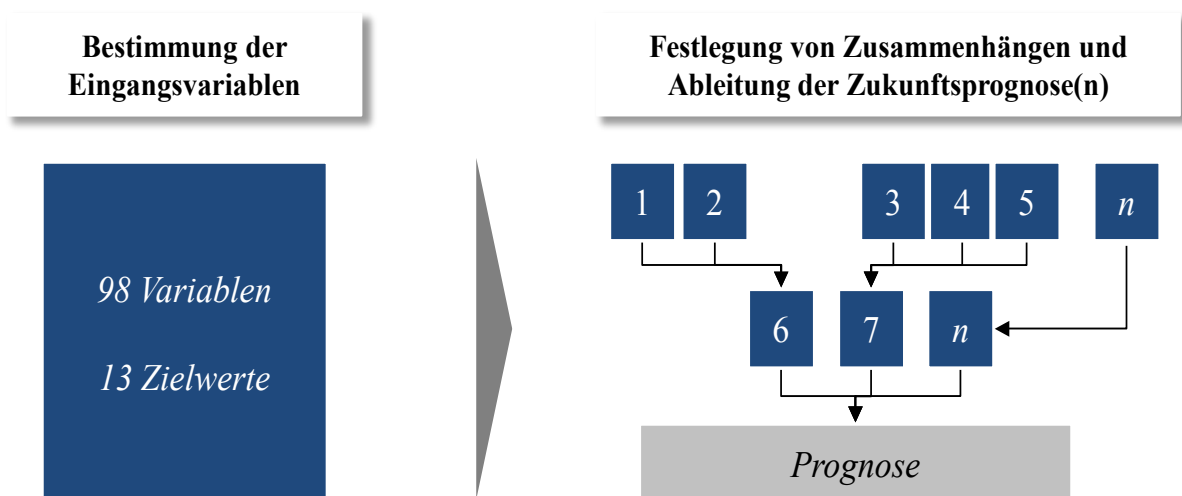


Abbildung 21: Prinzipielle Herangehensweise der Untersuchung

Als Folge einer bisher fehlenden Systematik wurde die Analyse des quantitativen Einflusses gesellschaftlicher Trends auf den Fahrzeugabsatz mit Hilfe eines Künstlichen Neuronalen Netzes (*KNN*) untersucht. Auch wenn die Studie von Rühl und Biethahn (2013) einen globalen Betrachtungsrahmen anlegt, wurde für die vorliegende Ausarbeitung der deutsche Fahrzeugmarkt ausgewählt. Dies diente der Komplexitätsreduktion der Prognosebildung. Gleichzeitig soll damit das Gesamtverfahren überprüft werden.

<sup>3</sup> Vgl. Hülsmann et al. (2012).

## 2 Künstliche Neuronale Netze als Instrument für die Prognose des Fahrzeugabsatzes

### 2.1 Generelle Nutzung von Künstliche Neuronale Netzen

KNN bestehen aus vielen kleinen Verarbeitungseinheiten, die die biologischen Fähigkeiten des menschlichen Gehirns abbilden. Sie vereinen Eigenschaften wie Lern- und Verallgemeinerungsfähigkeit und assoziative Speicherung von Informationen.<sup>4</sup> Das erste mathematische Modell von einem Neuron wurde von McCulloch und Pitts 1943 entworfen.<sup>5</sup> Neuronen sind dabei die grundsätzliche Verarbeitungseinheit in einem KNN. Es besteht aus einer Reihe von Eingabekanälen, die über Input-, Aktivierungs- und Outputfunktion in einem Ausgabekanal münden.<sup>6</sup> Die Funktion des Neurons übernimmt damit die „*nichtlineare Transformation des Gesamtinputs*“<sup>7</sup>.

Durch die Fähigkeit des Lernens von Beispielen können mit KNN auch nichtlineare Probleme lösen. Besonders geeignet sind KNN bei Aufgaben, für die eine hohe Anzahl an Daten vorliegen und bei denen die Beziehung zwischen den Daten nur schwer zu beschreiben oder unbekannt ist sowie für die Lösung Wissen benötigt, welches nicht grundlegend spezifizierbar ist.<sup>8</sup> Durch die Abstraktion der Gehirnfunktionen in KNN benötigen sie auch einen spezifischen Lernprozess, um Probleme zu verstehen und in eine Ausgabefunktion zu überführen. Dadurch ist es essentiell, dass der Datensatz das konkrete Problem bestmöglich abbildet.<sup>9</sup> *Eine Konkretisierung der Funktion von KNN wird an dieser Stelle nicht vorgenommen, kann aber in einschlägige Literatur zum Thema nachgelesen werden.*

Lackes und Mack gehen davon aus, dass KNN auch in der Unternehmensplanung generell gewinnbringend eingesetzt werden können. Sie verweisen aber darauf, dass KNN mit Akzeptanzproblemen in betriebswirtschaftlichen Problemstellungen zu kämpfen haben. Hintergrund dazu sei, dass KNN einem Black-Box-Modell ähneln und die Ergebnisse von KNN logisch häufig nur für Experten nachvollziehbar seien.<sup>10</sup>

Die nachfolgende Tabelle gibt einen beispielhaften Einblick über die generellen Einsatzbereiche von KNN bei betriebswirtschaftlichen Fragestellungen.

Funktionsbereiche	Anwendung Neuronaler Netze
<b>Absatz-, Marktfor- schung</b>	<ul style="list-style-type: none"> <li>▪ Analyse des Konsumverhaltens</li> <li>▪ Identifikation und Analysen von Marktsegmenten</li> <li>▪ <i>Umsatzprognosen für Produkt- und Preisplanung</i></li> <li>▪ Unterstützung bei der Entwicklung von Preisstrategien</li> </ul>

<sup>4</sup> Vgl. Chauhan Manish Kumar und Mittal M. L. (2012); Zhang, Patuwo und Hu (1998).

<sup>5</sup> In Chauhan Manish Kumar und Mittal M. L. (2012): McCulloch W. S. and Pitts W., A Logical Calculus of Ideas Immanent in Nervous Activity, Bull. Mathematical Bio-physics, 5, 115-133 (1943).

<sup>6</sup> Vgl. Lackes und Mack (2000), S. 3-4.

<sup>7</sup> Lackes und Mack (2000), S. 4.

<sup>8</sup> Vgl. Zhang, Patuwo und Hu (1998).

<sup>9</sup> Vgl. Pradhan und Kumar (2008).

<sup>10</sup> Lackes und Mack (2000), S. 4.

<b>Materialwirtschaft, Produktion und Logistik</b>	<ul style="list-style-type: none"> <li>▪ <i>Materialbedarfsplanung</i></li> <li>▪ Identifikation von Störungsfällen in der Fertigung</li> <li>▪ Unterstützung einer kostenorientierten Qualitätssicherung</li> <li>▪ Auftragsreihenfolgeplanung</li> </ul>
<b>Rechnungswesen</b>	<ul style="list-style-type: none"> <li>▪ Kostenplanung im innerbetrieblichen Rechnungswesen</li> <li>▪ Ermittlung konstruktionsbegleitender Kalkulationsdaten</li> <li>▪ Bilanzanalyse und –auswertung</li> </ul>
<b>Personalwesen</b>	<ul style="list-style-type: none"> <li>▪ Personalbeurteilung für die Einstellung, Versetzung, Entlassung</li> <li>▪ <i>Personaleinsatz- und Schichtplanung</i></li> </ul>
<b>Finanzwesen</b>	<ul style="list-style-type: none"> <li>▪ <i>Aktienprognosen für Investitions- und Finanzierungentscheidungen</i></li> <li>▪ <i>Devisenkursprognosen für Investitions- und Finanzierungentscheidungen</i></li> <li>▪ Kreditwürdigkeitsprüfungen von Privat- und Firmenkunden</li> </ul>

Tabelle 10: Anwendungen Neuronaler Netze in der Unternehmensplanung<sup>11</sup>

Bereits die Anwendungsbereiche in Tabelle 10 verdeutlichen, dass KNN für Prognoseaufgaben eingesetzt werden können (*kursiv* markiert). Im kommenden Kapitel wird deshalb vertieft auf die Möglichkeiten von KNN als Prognoseinstrument eingegangen.

## 2.2 Künstliche Neuronale Netze als Prognoseinstrument

KNN als Prognoseinstrument wurden 1987 und 1988 von Lapedes und Farber das erste Mal erfolgreich angewendet.<sup>12</sup> In den folgenden Jahren wurden KNN als Prognoseinstrument in vielen Bereichen der Ökonomie angewendet, unter anderem für die Prognosen von Wechselkursen, Zinssätzen, Industrieproduktion, Bonds, Derivaten und Inflation.<sup>13</sup> Brühl et al. gehen davon aus, dass besonders durch die Weiterentwicklung von mathematischen Algorithmen und die Auswertungsmöglichkeiten durch Computer die Einsatzgebiete von verbesserten Methoden – wie KNN sie darstellen – unterstützt wurden und damit verfügbare Datensätze in höhere Qualität als bisher mit herkömmlichen statistischen Verfahren untersucht werden konnten.<sup>14</sup>

Nach der stärkeren Verbreitung von KNN als Prognoseinstrument stellten einige Autoren (Aday und Collopy (1998); Zhang, Patuwo und Hu (1998)) in Frage, ob KNN in der Prognosequalität gegenüber alternativen Methoden überhaupt überlegen sind. In

<sup>11</sup> Lackes und Mack (2000), S. 5.

<sup>12</sup> Vgl. Zhang, Patuwo und Hu (1998).

<sup>13</sup> Vgl. Carvalhal et al. (2007): „[...] *exchange rates* (Verkooijen (1996) and Hu et al. (1999)), *economic growth* (Stock and Watson (1998) and Tkacz (1999)), *industrial production* (Moody, Levin and Rehfuss (1993), *stock performance* (Refenes, Zapranis and Francis (1994) and Wong and Long (1995)), *stock volatility* (Donaldson and Kamstra (1997)), *derivatives* (Hutchinson et al. (1994) and Carverhill and Cheuk (2003)), *futures trading volume* (Kaastra and Boyd (1995)), *international equity prices* (Cogger et al. (1997)), *bonds* (Desai and Bharti (1998)), *dividends* (Donaldson and Kamstra (1996)), *initial public offerings* (Jain and Nag (1993)), *credit scoring of loan applications and bond ratings* (Swales and Yoon (1992)), *oil price futures* (Moshiri (2004)), *interest rates* (Wood and Dasgupta (1995)) and *inflation* (Chen et al. (2001), Marcellino (2002), Franck and Krausz (2004) and McNelis and McAdam (2004))”

<sup>14</sup> Vgl. Brühl et al. (2009); Pradhan und Kumar (2008).

der Untersuchung von Zhang, Patuwo und Hu (1998)<sup>15</sup> wurde die generelle Nutzung von KNN als Prognoseinstrument mit statistischen Methoden verglichen und auf Grundlage von bis diesem Zeitpunkt verfügbaren Studien evaluiert. Demnach waren nicht alle KNN-Prognosen durchgehend besser als vergleichbare statistische Methoden. Nichtsdestotrotz gehen die Autoren davon aus, dass in den Fällen von keiner Überlegenheit der KNN-Prognosequalität über lineare Methoden bereits der untersuchte Datensatz linear war und daraus folgend keine besseren Ergebnisse erwartet werden können. Schlussendlich würden KNN einzigartige Vorteile bieten (u.a. Adaptierbarkeit, Untersuchung von nicht-linearen Problemen), die sie zu einem hilfreichen Instrument für Prognose machen – auch wenn ein hoher Grad von Unsicherheit einhergeht.

Aday und Collopy (1998) greifen in ihrer Untersuchung die Problematik auf, dass in einigen Studien die KNN-Prognose einerseits nicht ausreichend mit Alternativmethoden verglichen wurde oder andererseits die Implementation des Vorgehens ungenügend war. Dazu wurden 48 Studien aus dem Zeitraum von 1988 bis 1994 miteinander verglichen. In elf Studien war sowohl der Vergleich mit den Alternativmethoden als auch die Implementation validiert. In weiteren elf Studien konnte keine ausreichende Implementationsqualität festgestellt werden, doch fielen die Ergebnisse positiv aus. Somit konnten nach der ersten Analyse insgesamt 22 valide Studien festgestellt werden. Aus diesen 22 Studien unterstützten 18 die Aussage, dass KNN das Potential für Prognosen aufweisen und damit dafür eingesetzt werden können. Unter diesen Umständen kann davon ausgegangen werden, dass KNN als Prognoseinstrument effiziente Ergebnisse liefern, sofern die Modelle fehlerfrei implementiert und validiert sind.<sup>16</sup>

Die Überlegenheit von KNN für Prognose liegt laut einer Untersuchung von Pradhan und Kumar, in der sie Prognosen über die wirtschaftliche Entwicklung mit Hilfe von KNN untersuchten, in der Fähigkeit, dass KNN auch nicht offensichtliche Informationen hinter Einflussfaktoren interpretieren können.<sup>17</sup> Dies unterstützt die Aussage von Donaldson und Kamstra, die KNN die nötige Flexibilität zuschreiben, um versteckte nicht-lineare Beziehungen zwischen Kennwerten herzustellen.<sup>18</sup>

Sollte eine signifikante nicht-lineare Beziehung vorliegen, so sind KNN-Prognosen anderer Methoden überlegen. Die zeigt die Untersuchung von Carvalhal und Ribeiro, in der Prognosen über latein-amerikanische Börsenindizes mit Hilfe von KNN zu den Methoden *Random Walk*, *ARMA*<sup>19</sup> und *GARCH*<sup>20</sup> verglichen wurden.<sup>21</sup> Unterstützt

<sup>15</sup> Vgl. Zhang, Patuwo und Hu (1998).

<sup>16</sup> Vgl. Aday und Collopy (1998).

<sup>17</sup> Vgl. Pradhan und Kumar (2008).

<sup>18</sup> Vgl. Donaldson und Kamstra (1996). *Anmerkung*: Die Autoren untersuchten die Nutzung von KNN in Kombination mit Prognose für Zeitreihen im Bereich der Börsenprognose in den USA, Kanada, Japan und Großbritannien.

<sup>19</sup> Auto-Regressive Moving Average.

<sup>20</sup> Generalized Auto-Regressive Conditional Heteroskedasticity.

<sup>21</sup> Vgl. Carvalhal und Ribeiro (2007).

wird dies weiterhin durch die Ergebnisse von Kumar<sup>22</sup>, auf die im nächsten Kapitel weiter eingegangen wird.

### 2.3 Spezifische Nutzung von Künstlichen Neuronalen Netzen als Prognosetool für den Fahrzeugabsatz in Deutschland

#### 2.3.1 Prognose des Fahrzeugabsatzes in der Automobilindustrie

Methoden über Prognosen für den Fahrzeugabsatz werden ab 1957 in der Literatur erwähnt. In zwei Untersuchungen geht hervor, dass Nerlove (1957) einfache Extrapolationen von Zeitreihen verwendet hat.<sup>23</sup> Lewandowski hat Absatzprognosen im Allgemeinen und fahrzeugabsatz-spezifische Prognosen zwischen 1970 und 1980 untersucht.<sup>24</sup> Im nächsten Schritt haben sich die Autoren Prevedouros und Ann (1998) sowie Dargay und Gately (1999) mit kausalen Beziehungsmodellen zwischen Fahrzeugabsatz und sozioökonomischen Faktoren beschäftigt.<sup>25</sup>

Das Fraunhofer Institut setzt für die Bildung eines angemessenen Modells – unabhängig der konkreten Methodik – voraus, dass ausschließlich nützliche Parameter gewählt werden. Dazu würde unter anderem der Fahrzeugabsatz, das Bruttoinlandsprodukt, der Verbraucherpreisindex, die Arbeitslosenrate, die Investitionsmenge der Industrie und einige weitere zählen. Diese müssen durch Experten in Wissenschaft und Wirtschaft bestätigt werden sowie durch mathematische Modelle ausgearbeitet werden (siehe Abbildung 22).<sup>26</sup>

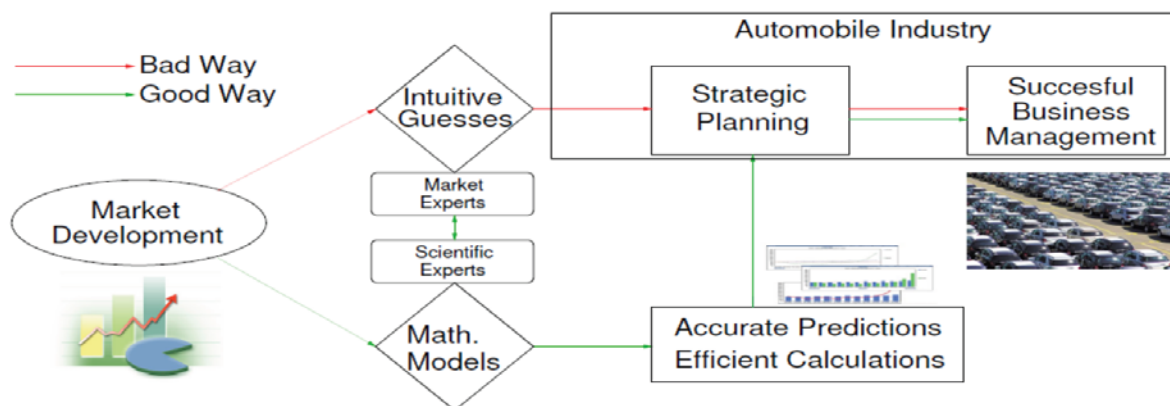


Abbildung 22: Prozess zur Prognosebildung nach dem Fraunhofer Institut<sup>27</sup>

<sup>22</sup> Vgl. Kumar et al. (2012).

<sup>23</sup> Vgl. Shababuddin (2009).

<sup>24</sup> Vgl. in Brühl et al. (2009): Lewandowski, R. (1974): Prognose- und Informationssysteme und ihre Anwendungen. de Gruyter, Berlin **und** Lewandowski, R. (1980): Prognose- und Informationssysteme und ihre Anwendungen Band II. de Gruyter, Berlin.

<sup>25</sup> Vgl. Shababuddin (2009).

<sup>26</sup> Vgl. Fraunhofer SCAI (2013).

<sup>27</sup> Fraunhofer SCAI (2013).

Aktuelle Untersuchungen knüpfen daran an und versuchen mit unterschiedlichen Verfahren für separate Märkte Prognosemethoden zu evaluieren. Auffällig ist dabei, dass es keine einheitliche Basis für die Einflussfaktoren gibt.

Shababuddin verwendet in seiner Untersuchung über den Fahrzeugabsatz ausländischer und nationaler Fahrzeuge in den USA ökonomische und demografische Daten. In seiner Analyse mit Hilfe schrittweiser Regressionen konnte festgestellt werden, dass ökonomische Faktoren mit dem Fahrzeugabsatz ausländischer Marken korrelieren, für nationale Marken konnte keine Korrelation festgestellt werden.<sup>28</sup> In der Untersuchung von Muhammad et al. über den Zusammenhang von makroökonomischen Variablen und dem Fahrzeugabsatz in Malaysia konnte keine langfristige Wechselwirkung festgestellt werden.<sup>29</sup> Brühl und seine Kollegen konnten allerdings bei einem Vergleich verschiedener Prognosemodelle für den deutschen Fahrzeugmarkt feststellen, dass nicht-lineare Modelle den linearen Modellen überlegen sind.<sup>30</sup> In einer neueren Untersuchung wurde – nach bestem Wissen der Autoren – bisher das erste Mal ein Künstliches Neuronales Netz zur Prognose von Absatzzahlen in der Automobilindustrie (hier in Indien) genutzt. Die Ergebnisse zeigen, dass auch hier KNN dem Modell der klassischen Zeitreihenanalysen in ihrer Prognosequalität und Genauigkeit überlegen sind.<sup>31</sup>

Rückblickend zeigen alle Untersuchungen, dass bisher keine allgemeingültigen Parameter vorliegen, auf deren Grundlage der Fahrzeugabsatz prognostiziert werden kann (siehe Tabelle 11). Darüber hinaus geht das Fraunhofer Institut davon aus, dass alle externen Faktoren die den Fahrzeugabsatz beeinflussen, für die Prognose mit validen zukünftigen Werten hinterlegt werden müssen. Darum ist eine Einbeziehung von Experten als auch die Integration von weiteren Prognosemodellen für die Einflussfaktoren essentiell.<sup>32</sup> Eine Vielzahl von Autoren ist sich jedoch einig, dass unerwartete Entwicklungen (z.B. die Umweltprämie in Deutschland<sup>33</sup>) nicht in Modellen hinterlegt werden können.<sup>34</sup>

Shababuddin (2009)	Brühl et al. (2009)	Muhammad et al. (2013)
<u>Fahrzeugabsatz</u>	<u>Fahrzeugabsatz</u>	<u>Fahrzeugabsatz</u>
BIP (Bruttoinlandsprodukt)	BIP	industrieller Produktionsindex

<sup>28</sup> Vgl. Shababuddin (2009).

<sup>29</sup> Vgl. Muhammad et al. (2013). *Anmerkung:* Zur Analyse wurde die VAR (Vector Autoregressive) Methode angewendet.

<sup>30</sup> Vgl. Brühl et al. (2009).

<sup>31</sup> Vgl. Chauhan Manish Kumar und Mittal M. L. (2012).

<sup>32</sup> Vgl. Fraunhofer SCAI (2013).

<sup>33</sup> Vgl. Bundesamt für Wirtschaft und Ausfuhrkontrolle (2010).

<sup>34</sup> Vgl. Fraunhofer (2013); Hülsmann et al. (2012); Brühl (2009).



<b>BNP (Bruttonationalprodukt)</b>	Verbraucherpreisindex	Verbraucherpreisindex
<b>Geldmenge M1</b>	Zinssatz	Zinssatz
<b>Geldmenge M2</b>	Benzinpreis	Ölpreis
<b>Geldmenge M3</b>	Investitionsnachfrage der Industrie	
<b>Konjunkturellen Frühindikatoren</b>	Verfügbares Einkommen	
<b>Diskontierungssatz</b>	Arbeitslosigkeitsrate	
<b>Bevölkerungsmenge</b>	Privater Konsum	
<b>Dauerhaft industrielle Nachfrage</b>	Latenter Ersatzbedarf	
<b>Dauerhaft privater Konsum</b>	Modellpolitik	
<b>Wechselnd industrielle Nachfrage</b>		
<b>Wechselnder privater Konsum</b>		

Tabelle 11: Gegenüberstellung der untersuchten Einflussfaktoren<sup>35</sup>

### 2.3.2 Generierung der Datenbasis

Die Studienergebnisse von Rühl und Biethahn (2013) zeigen, dass die Entwicklung des Fahrzeugabsatzes nicht ausschließlich von ökonomischen Kenngrößen abhängt. Aus diesem Grund wurden im Gegensatz zur Vorgehensweise von verschiedenen, bereits diskutierten Untersuchungen (Fraunhofer SCAI (2013); Shababuddin (2009); Hülsmann et al. (2012); Muhammad et al. (2013); Brühl et al. (2009)) – die ausschließlich externe wirtschaftliche Parameter als Grundlage nutzen – in dieser Untersuchung auch demografische Faktoren integriert. Wurden in anderen Studien maximal zwölf Eingangsvariablen benutzt, sind für diese Simulation im KNN 98 Eingangsvariablen genutzt worden. Die hohe Anzahl der Eingangsvariablen und verfügbaren Datensätze ist relevant, da KNN erst mit einem Datensatz valide Ergebnisse liefern, der das konkrete Problem möglichst gut beschreibt.<sup>36</sup> Weiterhin wurden 13 Zielvariablen (Fahrzeugklassen) untersucht, während in anderen Untersuchungen ausschließlich der Gesamtfahrzeugabsatz als Zielvariable festgelegt wurde.

Folgende Fahrzeugklassen wurden betrachtet:

- [1] Gesamtanzahl Neuzulassungen<sup>37</sup>
- [2] Minis
- [3] Kleinwagen
- [4] Kompaktklasse
- [5] Mittelklasse
- [6] Obere Mittelklasse
- [7] Oberklasse
- [8] Geländewagen
- [9] Sportwagen
- [10] Cabriolets

<sup>35</sup> Vgl. Shababuddin (2009); Brühl et al. (2009); Muhammad et al. (2013).

<sup>36</sup> Vgl. Pradhan und Kumar (2008).

<sup>37</sup> Ohne Wohnmobile.

- [11] Vans Gesamt
- [12] Utilities
- [13] Sonstige

### 2.3.3 Vorgehen zur Datenaufbereitung

Aus der Vielzahl von Einflussfaktoren, deren konkreten Zusammenhänge und Einflusswichtigkeiten auf der Fahrzeugabsatz sowie einer resultierenden komplexen Modellbildung wurde die Prognosebildung mit der Hilfe eines KNN durchgeführt. Nichtsdestotrotz war auch für die Simulation in einem KNN die historische Datenverfügbarkeit wichtig, als auch eine geringe Anzahl von Eingangsvariablen.

*Folgende Schritte zur Datenaufbereitung wurden vorgenommen:* Aus allen verfügbaren und sinnvollen Datensätzen, wurde die Schnittmenge der verfügbaren Jahre ermittelt, da teilweise Datensätze über weitere Jahre verfügbar waren, andere allerdings nur für einzelne Jahre. Die Schnittmenge für die Eingangsvariablen waren die Jahre 2005 bis 2010, für die Zielvariablen die Jahre 2004 bis 2012.

Um die sinnvolle Simulation im KNN zu gewährleisten, musste die Anzahl der Eingangsvariablen von 98 Stück<sup>38</sup> auf unter 15 gesenkt werden. Dazu wurden latente Variablen eingeführt, die eine Aggregation von verschiedenen Variablen darstellen. Latente Variablen haben damit keine direkte Aussage, sondern sind die Summe aus den Produkten der Variablen und deren jeweiligen Gewichtung pro Jahr.

Folgende elf (latente) Variablen wurden aufgrund inhaltlicher Gemeinsamkeiten gebildet:

- [1] Haushaltsgröße
- [2] Ausstattungsgrad
- [3] Ausstattungsbestand (*keine latente Variable, besteht nur aus einer Eingangsvariable*)
- [4] Haushaltsnettoeinkommen
- [5] Verkehrsausgaben
- [6] Demografie
- [7] BIP (*keine latente Variable, besteht nur aus einer Eingangsvariable*)
- [8] Studentische Entwicklung
- [9] Ausbildungsentwicklung
- [10] Pkw-Verfügbarkeit
- [11] Führerscheinbesitz

Jede latente Variable besteht aus zwei bis 27 realen Variablen. Um diese realen in latente Variablen zu überführen, wurden zwei Prozesse angewendet. Im ersten Schritt wurde überprüft, welche Variablen in einer Abhängigkeit zueinander stehen. Dazu

---

<sup>38</sup> *Anmerkung.* Anzahl der Variablen, die nach der Festlegung der Datenverfügbarkeit zwischen 2005 und 2010 zur Simulation genutzt werden konnten.

wurde pro Jahr das Verhältnis aller Variablen zueinander gebildet und der Maximal- und Minimalverhältniswert für jedes Jahr gebildet. Lag das Verhältnis innerhalb von sechs Prozent, wurde die jeweilige Variable ausgeschlossen.

Im zweiten Schritt musste nun eine Gewichtung für jede Variable berechnet werden, um diese in einer latenten Variablen zusammen zu fassen. Dazu wurde das Modell des *Analytic Hierarchy Processes* (AHP – Analytischer Hierarchieprozess) genutzt.<sup>39</sup> Es wurde exemplarisch untersucht, ob Variable x im Vergleich zu Variable y einen größeren Einfluss auf den Autoabsatz (das Ziel der Untersuchung) hat. Hierbei wurden subjektive Tendenzen integriert, da eine Skala – bei gleichem oder größeren Einfluss zwischen eins (=gleichen Einfluss) und zehn (=stärkeren Einfluss) sowie bei kleiner Einfluss die Gegenoperation (ein dividiert durch x) – als Grundlage gewählt wurde. Allerdings wurde dieser Prozess durch beide Autoren sowie durch Prof. Dr.-Ing. Reinhard Kolke<sup>40</sup> durchgeführt, um Fehlinterpretationen zu vermeiden. Die Matrizen wurden multipliziert, bis weitere Iterationen keine Veränderungen bewirkten. Damit konnten die Wichtigkeiten der Einflussfaktoren berechnet werden und mit den spezifischen Jahreswerten zur jeweiligen latenten Variable zusammengefasst werden.

## 2.4 Ergebnisse

### 2.4.1 Erklärungsgüte der Ergebnisse

Nach der Simulation der (latenten) Variablen im KNN und die Überprüfung durch rückwirkende Prognosen auf Grundlage der KNN-Ergebnisse konnte die Erklärungsgüte für die einzelnen Fahrzeugklassen festgestellt werden. Die Ergebnisse deuten drauf hin, dass die Resultate zu einem hohen Maße die Wirklichkeit widerspiegeln. Sechs von 13 Fahrzeugklassen weisen eine Erklärungsgüte größer 80 Prozent auf, vier weitere Fahrzeugklassen können mit einer Erklärungsgüte größer 70 Prozent noch erklärt werden. Am besten wird die Fahrzeugklasse der *Minis* beschrieben. Die Fahrzeugkategorie der *Cabriolets* wurde ab 2007 durch das Kraftfahrtbundesamt in Deutschland nicht mehr geführt und konnte deshalb nicht simuliert werden. Weiterhin wurde die Fahrzeugkategorie der *Sportwagen* erst ab 2007 geführt, weshalb hier eine geringe Erklärungsgüte in Höhe von 36 Prozent vorzufinden ist (siehe Tabelle 12).

Zielwert	Erklärungsgüte	Weiterführende Auswertung
Gesamtanzahl Neuzulassungen <sup>41</sup>	74 %	Ja
Minis	96 %	Ja

<sup>39</sup> Vgl. Saaty (2008).

<sup>40</sup> Prof. Dr.-Ing. Kolke ist Leiter für Test und Technik im ADAC Technik Zentrum in Landsberg am Lech. Darüber hinaus ist er Hochschullehrer für Automotive Technology Management – Zukunftsmobilität an der *Business and Information Technology School (BiTS)* in Iserlohn.

<sup>41</sup> Anmerkung: Ohne Wohnmobile.

<b>Kleinwagen</b>	89 %	Ja
<b>Kompaktklasse</b>	90 %	Ja
<b>Mittelklasse</b>	93 %	Ja
<b>Obere Mittelklasse</b>	93 %	Ja
<b>Oberklasse</b>	88 %	Ja
<b>Geländewagen</b>	73 %	Ja
<b>Sportwagen</b>	36 %	Nein
<b>Cabriolets</b>	<i>nicht verfügbar</i>	Nein
<b>Vans Gesamt</b>	76 %	Ja
<b>Utilities</b>	71 %	Ja
<b>Sonstige</b>	3 %	Nein

Tabelle 12: Erklärungsgüte je Zielvariable nach der Simulation

In der gesamten Studie wurden alle Fahrzeugklassen mit einer Erklärungsgüte größer 70 Prozent weiter untersucht. Im folgenden Kapitel wird an Hand der Fahrzeugklasse *Obere Mittelklasse* das weitere Vorgehen und die Ergebnisse vorgestellt.

#### 2.4.2 Detaillierergebnisse der Oberen Mittelklasse

Die Ergebnisse für die Fahrzeugklasse der *Oberen Mittelklasse* deuten mit einer Erklärungsgüte in Höhe von 93 Prozent darauf hin, dass eine Prognose relativ genau sein sollte.<sup>42</sup> Gleichzeitig stellt die *Obere Mittelklasse* mit den Fahrzeugen wie Audi A6, BMW 5er und Mercedes E-Klasse eine relative Wichtigkeit für den deutschen Automobilbau dar.

Mit den Ergebnissen aus dem KNN war es möglich, für jede (latente) Variable eine Funktion zweiten Grades zu bilden (siehe Formel 1). Diese bildet den Einfluss jeder latenten Variablen pro Betrachtungsjahr ab.

*Formel 1:*

$$F(x)_n = ax^2 + bx + c$$

*Mit:*

$n$  = Betrachtungsjahr

$F(x)_n$  = Einfluss der latenten Variable

$x$  = latente Variable im Betrachtungsjahr

$a$ ,  $b$  und  $c$  = Konstanten (unterschiedlich je latenter Variable)

<sup>42</sup> Unter der Annahme keiner unerwarteten Entwicklungen.

In Abbildung 23 wird deutlich, dass sich der reale Gesamteinfluss pro Betrachtungsjahr aus der Summe der Einzeleinflüsse ausgehend von den latenten Variablen zusammensetzt. Es ist erkennbar, dass in den Jahren 2005 alle Einflussfaktoren eine positive Entwicklung begünstigten. In 2008 und 2009 werden auch mit Hilfe der rückwirkend betrachteten Einflussfaktoren die Auswirkungen der Finanzkrise deutlich.

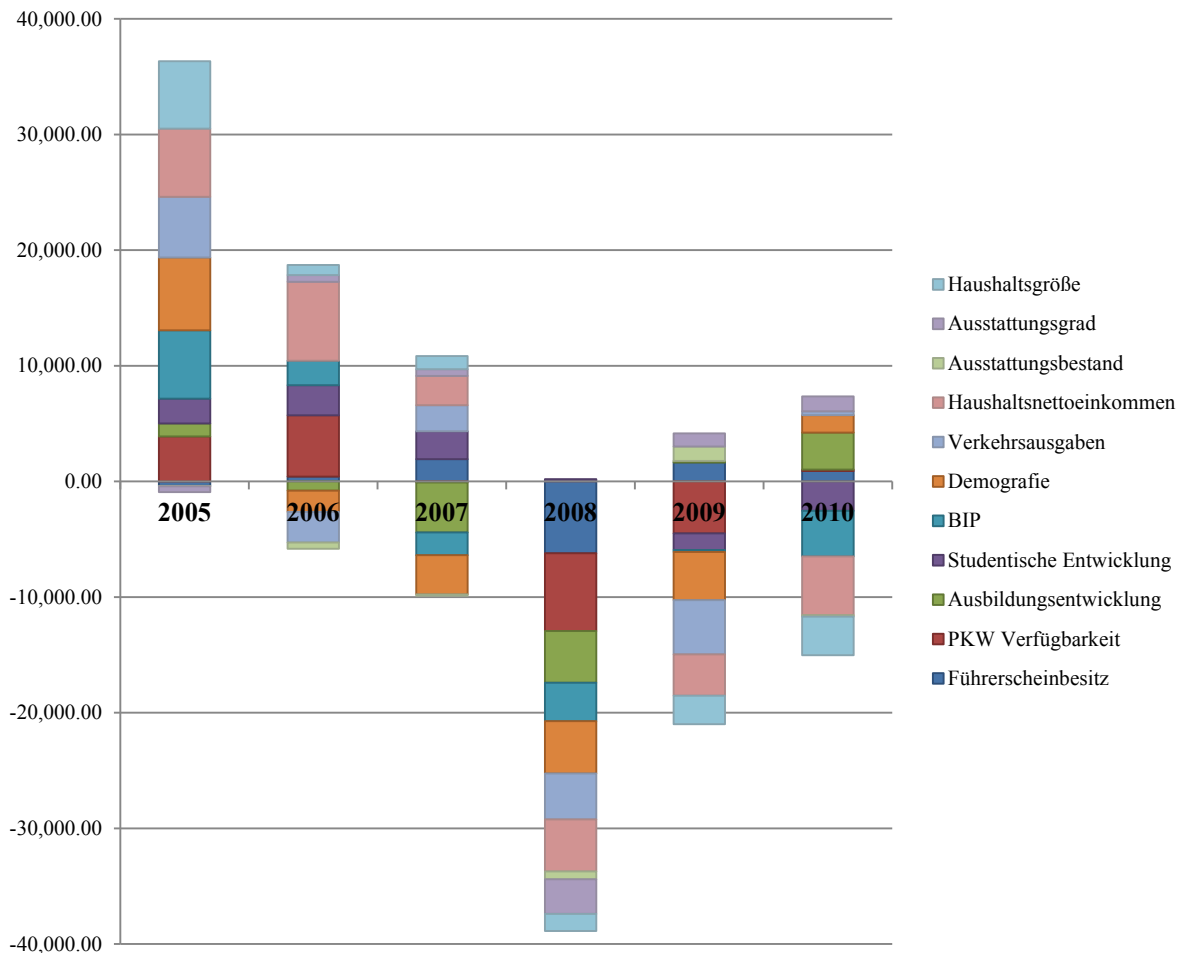


Abbildung 23: Einfluss der latenten Variablen (in Stück) pro Jahr

Abbildung 24 bringt die realen Auswirkungen (in Fahrzeugabsatz) aus Abbildung 23 in ein relatives Verhältnis zum Gesamteinfluss. Die Erkenntnisse sind ähnlich der aus Abbildung 23 zu deuten.

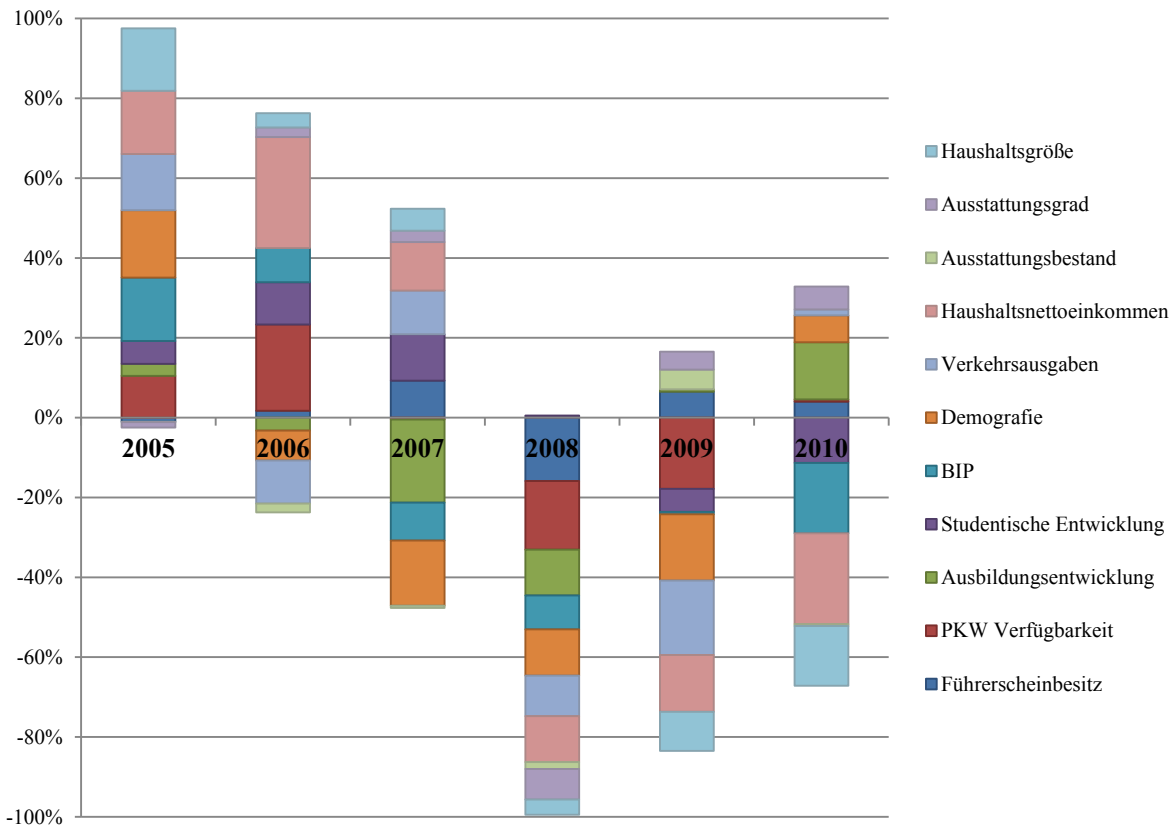


Abbildung 24: Einfluss der latenten Variablen (in Prozent) pro Jahr

Die Darstellung in Abbildung 25 wiederum zeigt den Einfluss der jeweiligen Variablen über den kompletten Betrachtungszeitraum. Auffällig dabei ist, dass nicht – wie in vielen Studien üblich – die Entwicklung des BIP den größten Einfluss auf den Fahrzeugabsatz hat, sondern das verfügbare Nettoeinkommen. Ausstattungsbestand und Ausstattungsgrad wiederum zeigen nur einen geringen Einfluss auf den Fahrzeugabsatz. Der Grund hierfür könnte sein, dass beide (latenten) Variablen schon durch das Haushaltsnettoeinkommen und weitere Faktoren beeinflusst sind und damit in gewissem Umfang korrelieren.

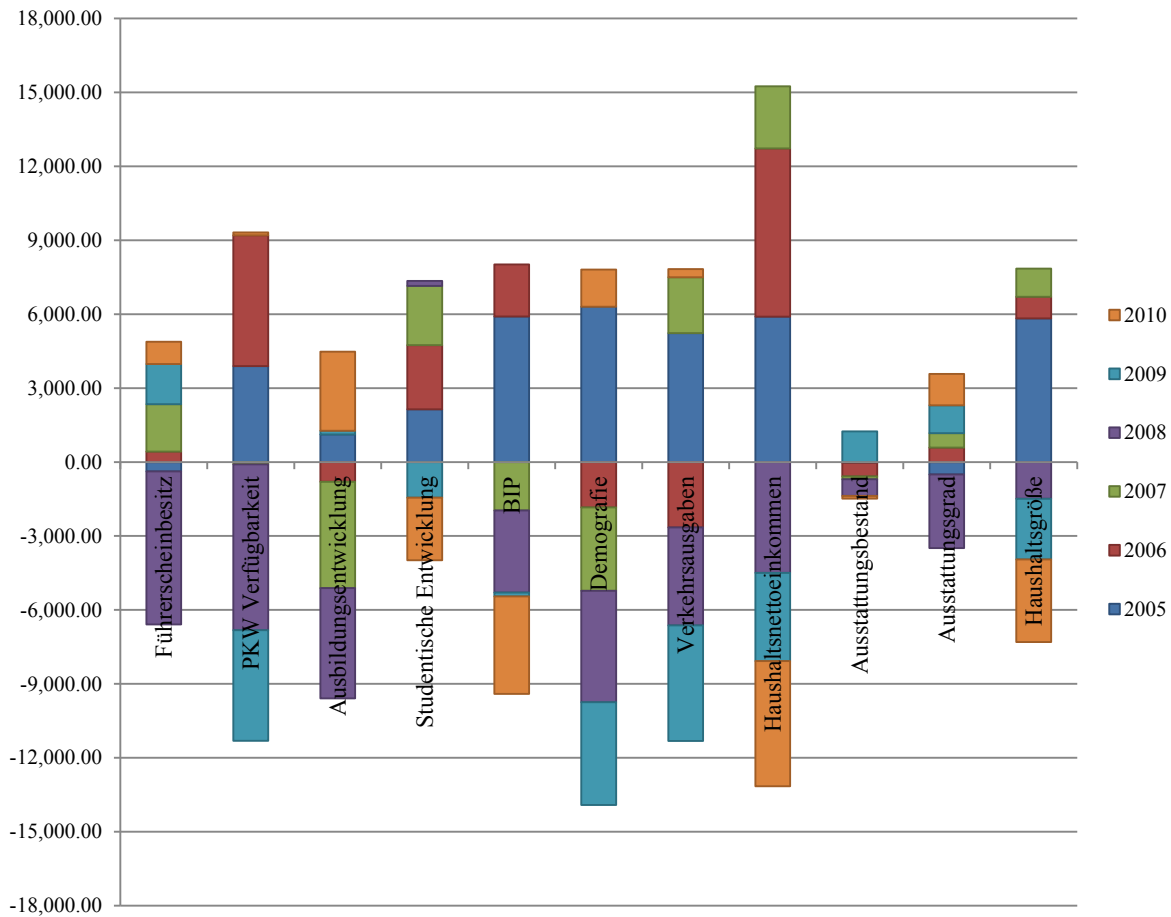


Abbildung 25: Einfluss der latenten Variablen (in Stück) zwischen 2005 und 2010

Um zu untersuchen, wie realitätsnah die Ergebnisse sind, wurde im nächsten Schritt eine rückwirkende Prognose aus der Summe der Einflüsse der (latenten) Variablen erstellt. Dazu wurde die durchschnittliche reale Zulassung der Jahre 2005 bis 2010 als Grundlage genommen und um die Summe der Einflüsse der (latenten) Variablen abgewertet durch ein Bestimmtheitsmaß  $R^2$  addiert (siehe Formel 2 und Formel 3).

Formel 2:

$$Prognose_n = \emptyset Zulassungen_{2005-2010} + \sum_1^i Einfluss\ der\ latenten\ Variable_{n-1} * R^2$$

Formel 3:

$$R^2 = \frac{Var (Neuzulassungen)_{2006-2011} - Var (Prognosefehler)_{2006-2011}}{Var (Neuzulassungen)_{2006-2011}}$$

Mit:

$n = Jahr$

$i = (Latente) Variablen\ 1\ bis\ 11$

Prognosefehler = Differenz zwischen realer Entwicklung und Prognose ohne  $R^2$

Für die Jahre 2006 bis 2011 wird in Abbildung 26 deutlich, dass die generierten Werte durch das KNN sehr stark der Realität entsprechen. Während in Jahr 2008 die höchste Abweichung von -12,1 Prozent zu verzeichnen ist, zeigen die anderen Jahren – besonders 2006 – wie genau die Funktion die Entwicklung abbilden kann.

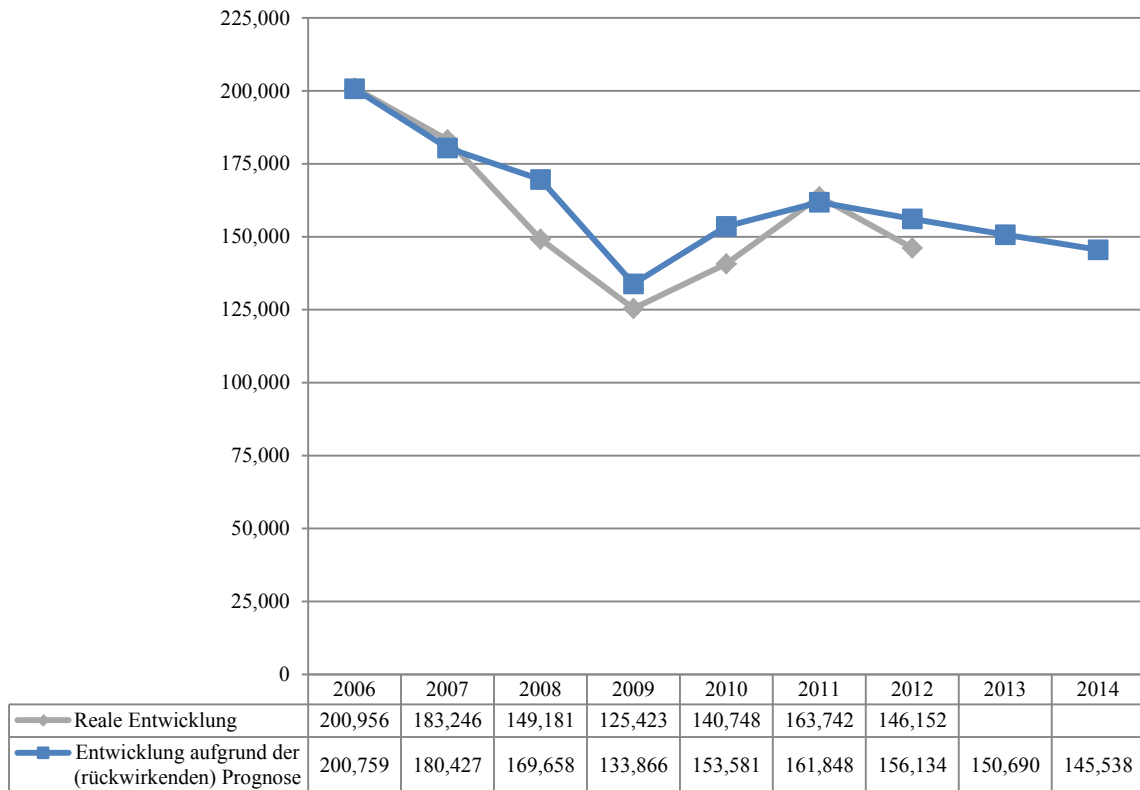


Abbildung 26: Vergleich der realen Entwicklung und der rückwirkende Prognose für die Jahre 2006 bis 2011 der *Oberen Mittelklasse* sowie (rückwirkende) Prognosen für 2012 bis 2014

Um eine Prognose bilden zu können, benötigt es die Fortschreibung der Eingangswerte in das Modell und damit die Prognose über die zukünftige Entwicklung der (latenten) Variablen. Bereits weitere Autoren (u.a. Fraunhofer (2013)) machten deutlich, dass dazu weitere Prognosemodelle angewendet werden müssen, um schlussendlich die nicht-lineare Zukunftsprognose erstellen zu können. Da die Eingangsvariablen für den Zeitraum von 2005 bis 2010 verfügbar waren, wurden die Werte mit der Compounded Annual Growth Rate (CAGR, 2005 bis 2010) fortgeschrieben. Mit dieser Methode konnten die Werte für 2012 bis 2014 prognostiziert werden (siehe Abbildung 26), wobei 2012 bereits einer rückwirkenden Prognose mit fortgeschriebenen Variablen entspricht. Es wird für 2014 auf dieser Grundlage mit einem Fahrzeugabsatz der *Oberen Mittelklasse* in Höhe von 145.538 Fahrzeugen gerechnet.

*Für alle weiteren Fahrzeugklassen wurde das identische Vorgehensmodell gewählt. Die zusammengefassten Ergebnisse werden in Kapitel 3 näher beschrieben.*



### 3 Zusammenfassung der Ergebnisse und Ausblick

Abbildung 27 zeigt die rückwirkenden und zukünftigen Prognosen auf Grundlage der KNN-Ergebnisse für alle betrachteten Fahrzeugklassen. Es ist deutlich zu sehen, dass die Absatzzahlen der Fahrzeugklassen *Kompaktklasse*, *Kleinwagen* und *Minis* auch bei der rückwirkenden Prognose einen positiven Ausschlag für 2009 und für 2010 einen Rückgang beschreiben. Es zeigt, dass zumindest rückwirkend auch unterwarteten Einflüsse wie hier die Abwrackprämie in Deutschland 2009 dargestellt sind. Dies deutet auch darauf hin, dass die Untersuchung relevante Ergebnisse geliefert hat.

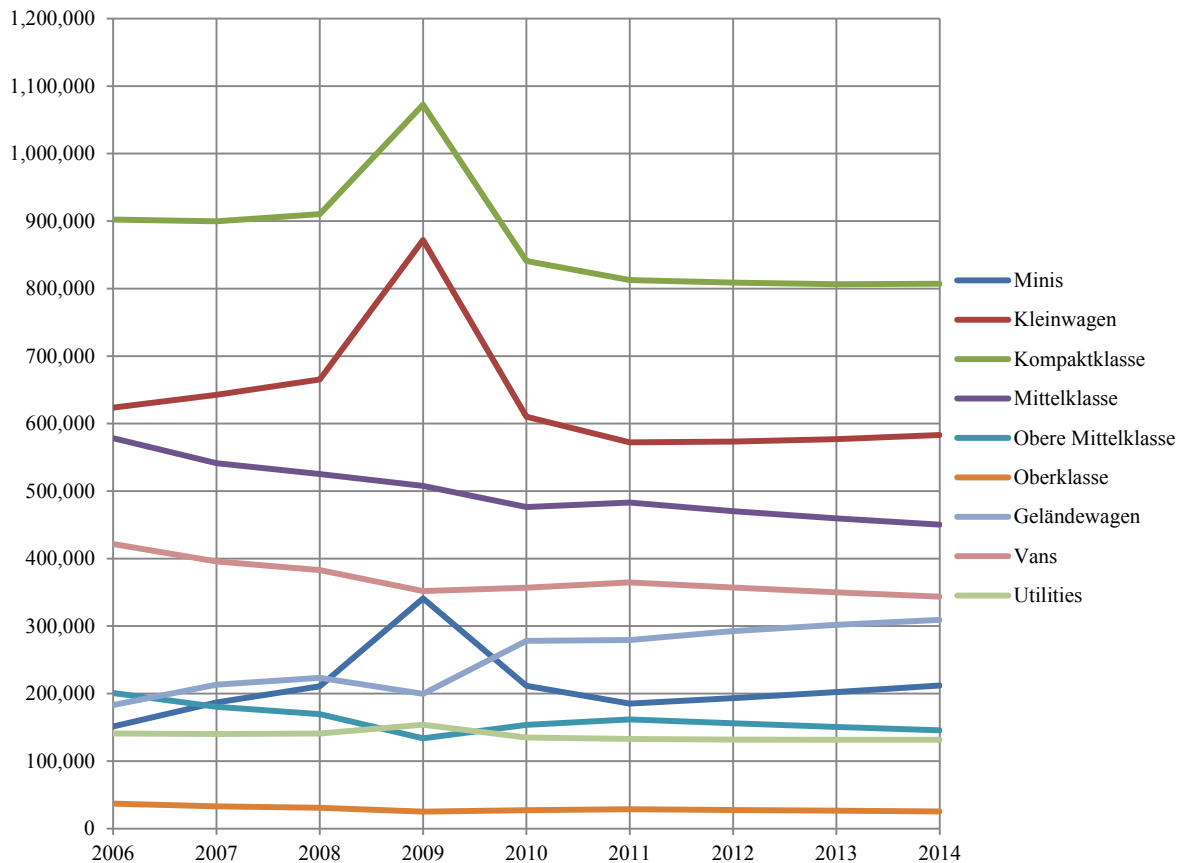


Abbildung 27: Rückwirkende und zukünftige Prognose pro Fahrzeugklasse für 2006 bis 2014

Bei der Prognose über den Gesamtfahrzeugmarkt ergeben sich zwei differenzierte Ergebnisse. Einerseits können alle Einzelprognose summiert und aufaddiert werden. Dadurch werden einzelne Fahrzeugklassen, die in der Simulation auf Grund ungenügender Ergebnisqualität nicht weiter untersucht wurden (siehe auch Tabelle 12), vernachlässigt. Nichtsdestotrotz ergibt sich für die berücksichtigten Fahrzeugklassen eine Verringerung der Fahrzeugverkäufe von 2012 bis 2014 um 0,1 Prozent (siehe Abbildung 28). Weiterhin wurde auch der Gesamtfahrzeugmarkt unabhängig der einzelnen Fahrzeugklassen untersucht und mit identischem Vorgehensmodell beschrieben. Hier sind alle Fahrzeugklassen enthalten und es ist einem Rückgang zwischen 2012 und 2014 in Höhe von 0,47 Prozent erwartet.

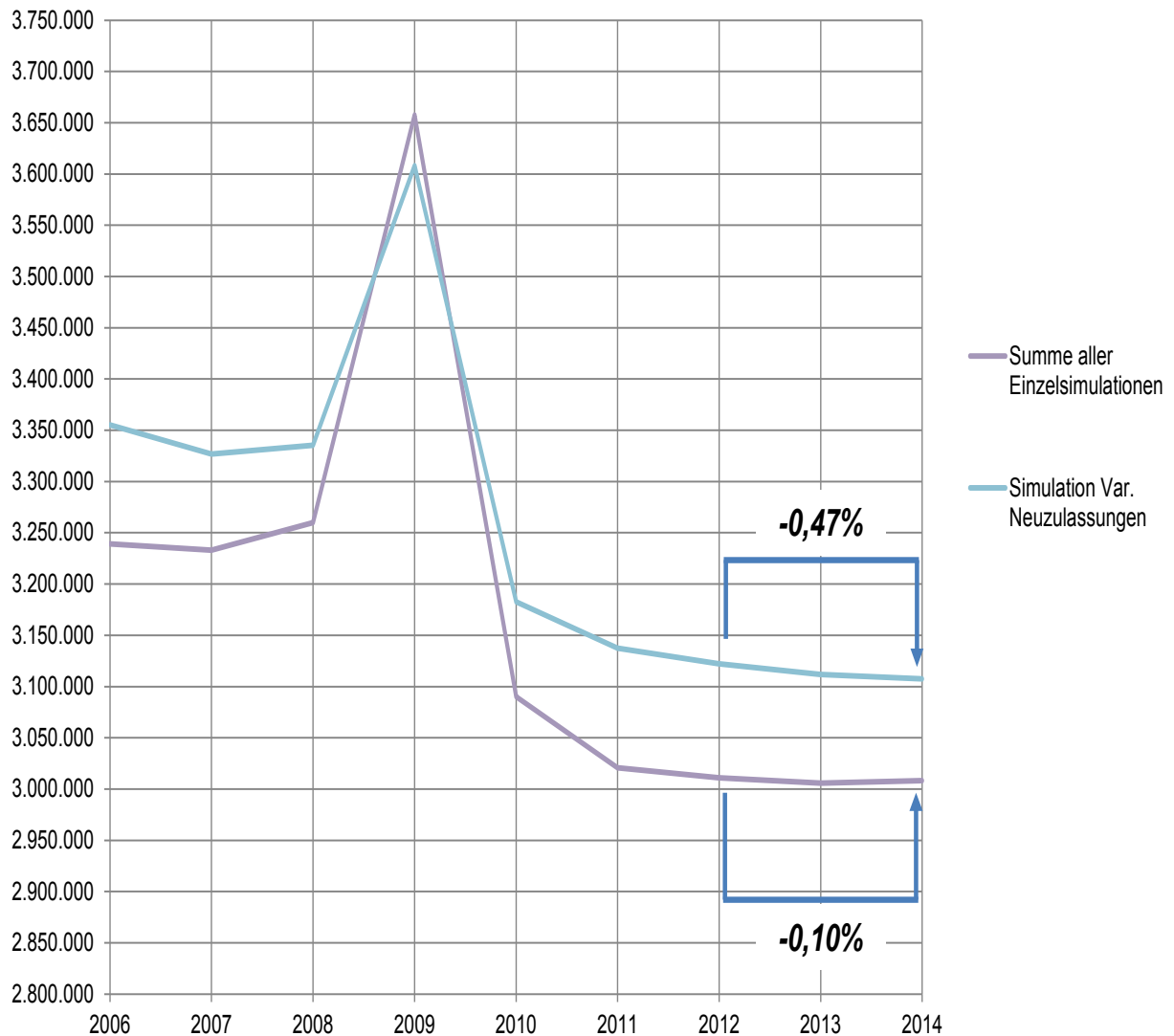


Abbildung 28: Rückwirkende und zukünftige Prognose der Gesamtzulassungen auf Grundlage der Summe aller Einzelsimulationen und der Gesamtmarktsimulation für 2006 bis 2014

Insgesamt bestätigt die Untersuchung, dass KNN für Prognosen verwendet werden können, sofern der Datensatz für einen unterunterbrochenen und langfristigen Zeitraum zur Verfügung steht. Darüber hinaus zeigt die Untersuchung, dass bei komplexen Problemen mit vielen Einflussfaktoren eine große Anzahl von Eingangsvariablen genutzt und mit Hilfe von weiteren Verfahren fallweise auf latente Variable übergeleitet werden können.

Vor dem Hintergrund, dass langfristige Prognosen auch mit diesem Modell umso ungenauer werden, desto weiter sie in der Zukunft liegen, benötigt der Datensatz weitere Aktualisierungen. Es besteht darüber hinaus die Möglichkeit, die Prognose der Eingangsvariablen mit weiteren Verfahren durchzuführen (hier wurde die CAGR genutzt) und zusätzlich weitere Eingangsvariablen sinnvoll zu ergänzen.

#### 4 Literatur

- Aday, M.; Collopy, F. (1998) *How Effective are Neural Networks at Forecasting and Prediction? A Review and Evaluation*. Journal of Forecasting. Vol. 17, S. 481-495. 1998.
- Brühl, B.; Hülsmann, M.; Borscheid, D.; Friedrich, C. M.; Reith, D. (2009) *A Sales Forecast Model for the German Automobile Market Based on Time Series Analysis and Data Mining Methods*. In P. Perner (Ed.): *Advances in Data Mining*. Applications and Theoretical Aspects: 9th Industrial Conference, ICDM 2009, Leipzig, Germany, July 20 - 22, 2009. Proceedings, S. 146-160. 2009.
- Bundesamt für Wirtschaft und Ausfuhrkontrolle (2010) *Abschlussbericht – Umweltprämie*. URL: [http://www.bafa.de/bafa/de/wirtschaftsfoerderung/umweltpraemie/publikationen/ump\\_abschlussbericht.pdf](http://www.bafa.de/bafa/de/wirtschaftsfoerderung/umweltpraemie/publikationen/ump_abschlussbericht.pdf) (zuletzt aufgerufen am 05.11.2013). Eschborn. November, 2010.
- Carvalho, A.; Ribeiro, T. (2007) *Do Artificial Neural Networks Provide Better Forecasts? Evidence from Latin American Stock Indexes*. Latin American Business Review. Vol. 8(3). 2007.
- Chauhan Manish Kumar; Mittal M. L. (2012) *Application of Artificial Neural Networks for Sales Forecasting in an Indian Automobile Manufacturing Company*. Advances In Management. Vol. 5 (9). September 2012.
- Donaldson, R. G.; Kamstra, M. (1996) *Forecast Combining with Neural Networks*. Journal of Forecasting, Vol. 15, S. 49-61 (1996).
- Fraunhofer SCAI (2013) *Automotive Sales Forecast Models*. URL: <http://www.scai.fraunhofer.de/en/about-us/staff/reith/automotive-sales-forecast-models.html> (zuletzt aufgerufen am 22.10.2013).
- Hülsmann, M.; Borscheid, D.; Friedrich, C.M.; Reith, D. (2012) *General Sales Forecast Models for Automobile Markets and their Analysis*. Transactions on Machine Learning and Data Mining. Vol. 5, No. 2, S. 65-86. 2012.
- Lackes, R.; Mack, D. (2000) *Neuronale Netze in der Unternehmensplanung*. München: Vahlen, 2000.
- Muhammad, F.; Hussin, M. Y. M.; Razak, A. A.; Rambeli, N.; Tha, G. P. (2013) *The Relationship between Macroeconomic Variables and Passenger Vehicle Sales in Malaysia*. Business and Economic Research. Vol. 3, No. 2. ISSN 2162-4860. 2013.
- Pradhan, R. P.; Kumar, A. (2008) *Forecasting economic growth using an Artificial Neural Network Model*. Journal of Financial Management and Analysis. Vol. 21(1), S. 24-31. 2008.
- Rühl, C.; Biethahn, N. (2013): *Metastudie zum Einfluss gesellschaftlicher Trends auf die Automobilindustrie*. In: Biethahn, N.; Sucky, E.; Werner, J. (2013): *Tagungsband zur Konferenz Mobility in a Globalised World*: Vol. II. Voraussichtlich Bamberg, Dezember 2013.
- Saaty, T.L. (2008) *Decision making with the analytic hierarchy process*. International Journal Services Sciences. Vol. 1, Nr. 1, S. 83-98. 2008.
- Shababuddin, S. (2009) *Forecasting automobile sales*. Management Review News. Vol. 32 (7): 13 Emerald Publishing. June 19, 2009.
- Zhang, G.; Patuwo, B. E.; Hu, M. Y. (1998) *Forecasting with artificial neural networks: The state of the art*. International Journal of Forecasting. Vol. 14, S. 35-62. 1998.