



# KI-basierte Analyse des Einflusses von Wetter auf die Fahrgeschwindigkeiten von Fernbussen und Lastkraftwagen

Anwendung von Methoden des maschinellen Lernens auf Umweltinformations- und Warnsysteme

David Plavcan, Eridy Lukau, Michael Klafft und Moritz Piening

## Zusammenfassung

Kenntnisse der Zusammenhänge zwischen Wetterbedingungen und Fahrgeschwindigkeiten verbessern die Planbarkeit und Reaktionsfähigkeit im Bereich der Güterlogistik und des öffentlichen Personenverkehrs und können so genutzt werden, um die Kundenzufriedenheit zu verbessern. Ziel ist es dabei, mit Hilfe möglichst genauer Wetterprognosen zu erwartende Beeinträchtigungen frühzeitig zu identifizieren und so die Auswirkungen derselben zu reduzieren. Dieser Beitrag untersucht den Einfluss des Wetters auf die Fahrgeschwindigkeiten von Lastkraftwagen und Fernbussen. Genauer werden zwei Ansätze aus dem Bereich der künstlichen Intelligenz angewandt, um zum einen wetterbedingte Ankunftszeitprognosen zu erstellen und zum anderen Frühwarnungen vor wetterbedingten Verzögerungen zu generieren. Die Ergebnisse der Modelle bestätigen einen negativen Einfluss von Neuschnee und

---

D. Plavcan  
UBIMET GmbH, Wien, Österreich  
E-Mail: [dplavcan@ubimet.com](mailto:dplavcan@ubimet.com)

E. Lukau (✉)  
Fraunhofer FOKUS, Berlin, Deutschland  
E-Mail: [eridy.lukau@fokus.fraunhofer.de](mailto:eridy.lukau@fokus.fraunhofer.de)

M. Klafft  
Jade Hochschule, Wilhelmshaven, Deutschland  
E-Mail: [michael.klafft@jade-hs.de](mailto:michael.klafft@jade-hs.de)

M. Piening  
Technische Universität Berlin, Berlin, Deutschland  
E-Mail: [moritz.piening@campus.tu-berlin.de](mailto:moritz.piening@campus.tu-berlin.de)

Starkregen auf die Fahrzeuggeschwindigkeiten, wohingegen andere Wetterparameter wie z. B. Wind nur einen marginalen Einfluss auf die Fahrgeschwindigkeiten im Untersuchungsgebiet hatten. Es konnte gezeigt werden, dass die genutzten Modelle geeignet sind, sowohl Navigationssysteme als auch intelligente Frühwarnsysteme zu verbessern.

---

### Schlüsselwörter

Maschinelles Lernen · Regression · Klassifikation · Logistik · Wetter · Expected Time of Arrival · Frühwarnsystem · Fahrgeschwindigkeit

---

## 1 Einleitung

Einer möglichst präzisen Vorhersage der Ankunftszeiten von Fahrzeugen kommt sowohl in der Straßenverkehrslogistik [1] als auch im öffentlichen Personenverkehr [2] eine große Bedeutung zu. Präzisere Prognosen erhöhen die Kundenzufriedenheit durch eine bessere Planbarkeit von Anschlussverbindungen, Reduzierung unnötiger Wartezeiten sowie im Logistikbereich durch die Möglichkeit einer effizienteren Bereitstellung von Ressourcen für Be- und Entladevorgänge. Zusätzlich hat in Schaltzentralen und im Notfallmanagement der betroffenen Unternehmen die Resilienz gegenüber Verkehrsstörungen eine hohe Priorität. Mitarbeiter und Mitarbeiterinnen profitieren daher von Entscheidungsunterstützungssystemen, welche frühzeitig auf erwartete Verzögerungen im Straßenverkehr hinweisen und es ermöglichen, rechtzeitig und vorbereitend darauf zu reagieren.

In der vorliegenden Studie werden wetterbedingte Verzögerungen bei Fahrten von LKW und Bussen mit zwei verschiedenen Ansätzen quantitativ untersucht.

Abschn. 2 verschafft einen Überblick über den aktuellen Stand des Wissens in Bezug auf Wetter und Verkehr bzw. den Einfluss des Wetters auf Geschwindigkeiten sowie die in der Forschungslandschaft für die Analyse dieser Zusammenhänge genutzten Ansätze.

Abschn. 3 beschreibt die grundlegende Datenauswahl sowie die benötigten Vorbereitungen, welche getroffen werden mussten, um beide Ansätze des maschinellen Lernens verfolgen zu können.

In Abschn. 4 wird zunächst analysiert wie verschiedene Wetterbedingungen die Fahrgeschwindigkeit beeinflussen, mit dem Ziel, diese Zusammenhänge für genauere Fahrgeschwindigkeitsprognosen zu nutzen.

In Abschn. 5 wird untersucht, ob wetterbedingte Verzögerungen auf Ebene einzelner Streckenabschnitte gelernt werden können, sodass auf Basis kurz- bis mittelfristiger Wettervorhersagen erwartete Verzögerungen ermittelt und darauf aufbauend Frühwarnungen für bestimmte Streckenabschnitte ausgespielt werden können.

Abschn. 6 fasst die Ergebnisse zusammen, diskutiert die dem Paper zugrunde liegende Hypothese und verschafft einen Ausblick auf weitere Arbeiten.

## 2 Stand des Wissens

Internationale Studien zeigen, dass starker Wind [3], schlechte Sicht [3] und starker Niederschlag [3, 4], insbesondere in Form von Schnee [4], zu einer Reduzierung der generellen Fahrgeschwindigkeit im Straßenverkehr führen. Dies legt den Schluss nahe, dass Wetter einen signifikanten Einfluss auf die Fahrgeschwindigkeiten – auch von Fernbussen und LKW – hat und die Einbeziehung von Wetter daher eine genauere Prognose der Ankunftszeiten ermöglicht.

Die Studienlage ist jedoch nicht einheitlich: so konnte [1] keine messbare Verbesserung der ETA (*Expected-Time-of-Arrival*)-Prognose von LKW durch Hinzunahme von Wetterdaten nachweisen, wobei die Parameter Schnee und Eis jedoch nicht berücksichtigt wurden und die Analyse auch nur auf einer einzigen Strecke in Schweden mit Daten von genau 5 Wetterstationen basierte. [5] konnte nur eine geringere Verbesserung der Genauigkeit von Stauprognosen durch Hinzunahme von Wetterfaktoren nachweisen, wobei sich die Analyse auf den Großraum Islamabad bezog, wo keine Schneefälle vorkommen. Allen Studien ist gemein, dass sie im Ausland durchgeführt wurden und daher nicht ohne weiteres auf den deutschen bzw. mitteleuropäischen Kontext übertragbar sind. Zudem bezieht sich keine der vorhandenen Studien explizit auf die Prognose der Fahrgeschwindigkeiten von Fernbussen. Die vorliegende Publikation wird diese beiden Forschungslücken adressieren.

Auch weitere Studien haben Zusammenhänge zwischen nachteiligem Wetter und verringerten Fahrgeschwindigkeiten aufzeigen können [6]. So wird in [7] ein verringerter Verkehrsfluss durch Regen, Schnee und Wind aufgrund von schlechten Sichtverhältnissen aufgezeigt. In [8] zeigen die Autoren signifikanten Einfluss von Regen auf Fahrgeschwindigkeiten bei geringem Verkehrsaufkommen und freier Fahrbahn (Autobahn) und gleichzeitig wenig Einfluss von Regen auf Fahrgeschwindigkeiten bei hohem Verkehrsaufkommen bzw. voller Straße.

In [9] wurden ähnliche Lösungsansätze verfolgt, wie es in dieser Studie der Fall ist. So wurden in den vorher genannten Studien die Wetterparameter: Nebel, Regen, Schnee, Überschwemmung und Eis, als für Fahrzeuge nachteilige Wetterbedingungen identifiziert. Auf Basis dieser Parameter wurde dann ein neuronales Netz trainiert, welches die Verkehrsgeschwindigkeit auf einem Streckenabschnitt prognostiziert. Auch wenn die Ergebnisse in [9] vielversprechend sind, wurden in der Studie lediglich 6 Monate historische Daten und auch nur ein einziger Streckenabschnitt untersucht. In der vorliegenden Studie werden Daten aus einem Zeitraum von knapp 3 Jahren für mehrere Streckenabschnitte auf Autobahnen durch Deutschland und Europa betrachtet, welche in der Straßengüterverkehrslogistik und von Fernbusunternehmen genutzt werden. Ebenso werden in dieser Studie im Vergleich zu [9] keine Deep-Learning-Methoden, sondern Methoden des klassischen maschinellen Lernens sowohl für die Regression als auch für die Klassifikation verwendet.

Als weiteren Beitrag zu Studien in diesem Feld präsentiert die vorliegende Studie die Betrachtung des Einflusses von Wetter auf Fahrgeschwindigkeiten – speziell von

Bussen und LKW – und untersucht die Integrierbarkeit von künstlicher Intelligenz in Navigationssysteme und intelligente Frühwarnsysteme.

---

### 3 Datengrundlage und Verarbeitung

Untersucht wurden Fahrten auf 12 ausgewählten Teststrecken mit einer Gesamtlänge von 5010 km pro Fahrtrichtung auf Autobahnen in Deutschland und Europa für den Zeitraum vom 01.01.2018 bis 16.07.2020. Insgesamt liegen Fahrspurdaten von ca. 1,1 Mio. Fahrten mit Wechselbrücken im Straßengüterverkehr sowie 39.190 Fernbusfahrten auf den o. g. Teststrecken vor. Insgesamt wurden für die Analyse 153,4 Mio. Fahrkilometer berücksichtigt (davon 144,8 Mio. km durch LKW und 8,6 Mio. km durch Fernbusse zurückgelegt). Für all diese Strecken liegen ebenfalls historische Wetteranalysen aus den Jahren 2018, 2019 und 2020 für die Parameter Temperatur, Luftfeuchtigkeit, Wind, Windböen, Niederschlag, Neuschnee, Schneehöhe und ein Glätteparameter in einer stündlichen Auflösung auf einen 1 km-Gitter vor.

Die Fahrten liegen als GPS-Punkte in einer zeitlichen Auflösung von 1 bis 5 min vor und müssen für die geplanten Untersuchungen zunächst auf Straßenabschnitte projiziert werden. Dafür werden Routendaten in sehr detaillierter Auflösung von teilweise wenigen Dekametern verwendet, entlang welcher die Fahrten aus den GPS-Punkten (re-) konstruiert werden. Dabei werden die Zeiten zwischen den GPS-Punkten entlang des Straßenverlaufs interpoliert. Im Anschluss werden für alle Straßensegmente aus deren Längen und den Durchfahrzeiten die Fahrgeschwindigkeiten berechnet. Diese tatsächlich gefahrenen Geschwindigkeiten werden in weiterer Folge *Track-Geschwindigkeiten* genannt.

Diese Track-Geschwindigkeiten können beispielsweise aufgrund von schwachem GPS-Signal fehlerhaft sein und müssen daher qualitätskontrolliert werden. Dafür werden alle unplausibel hohen Geschwindigkeiten ( $>105$  km/h bei Bussen und  $>95$  km/h bei LKW), aber auch sehr langsame ( $<3$  km/h) aus dem Datensatz komplett entfernt. Zusätzlich werden ungewöhnlich langsame Fahrten ( $<25$  km/h auf Autobahnen) bei denen anzunehmen ist, dass Wetter keinen Einfluss haben kann (kein Niederschlag, keine Glätte, nur schwacher Wind und relative Feuchte deutlich unter der Sättigung) ausgeschlossen, um die Daten um Staus aufgrund von anderen Ursachen wie z. B. Unfällen oder Baustellen zu bereinigen. Ebenfalls werden alle Segmente, welche durch Tunnels führen, entfernt, da hier die Fahrtstrecke witterungsgeschützt ist.

Neben diesen tatsächlich gefahrenen Track-Geschwindigkeiten werden für die Analyse zusätzlich *Routengeschwindigkeiten* verwendet. Die Routengeschwindigkeiten werden von einer Navigationssoftware verwendet, welche sich für jedes Segment aus gemittelten Geschwindigkeiten je Wochentag und Zeitpunkt (in 5 min Auflösung) ableiten. Somit berücksichtigen diese Routengeschwindigkeiten die üblichen Schwankungen der Verkehrsdichte je nach Tageszeit. Für LKW und Busse müssen diese Routengeschwindigkeiten mit der jeweiligen Maximalgeschwindigkeit (für Busse

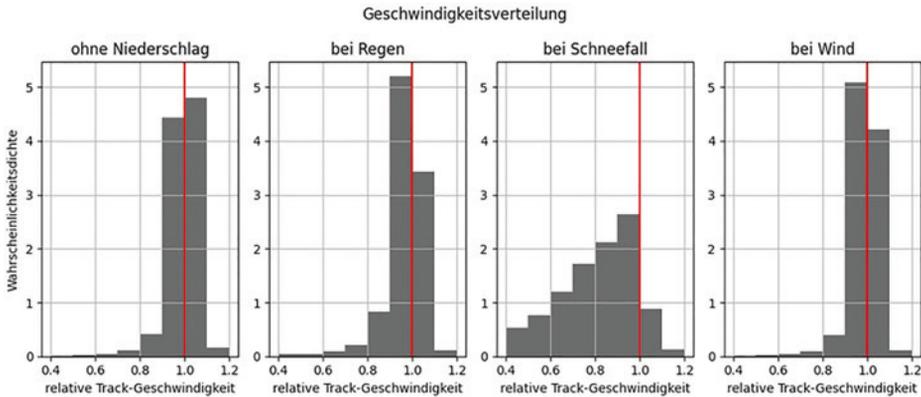
100 km/h und für LKW 90 km/h) nach oben begrenzt werden. Dabei ist darauf hinzuweisen, dass für LKW eigentlich eine Geschwindigkeitsbegrenzung von 80 km/h gilt, die genutzten Geschwindigkeitsdaten der LKW aber zeigen, dass in der Praxis meist mit knapp 90 km/h gefahren wird.

Ein Vergleich der Routengeschwindigkeiten mit den Track-Geschwindigkeiten zeigt im Median für die Mehrheit der Segmente eine gute Übereinstimmung, teilweise aber auch systematische Abweichungen. Zu diesen Abweichungen kommt es besonders an Steigungen, wobei die Routengeschwindigkeit die Track-Geschwindigkeit im Median häufig (je nach Segment) um 10 bis 30 km/h überschätzt. Da in dieser Studie die Routengeschwindigkeit als Referenz dienen soll, würden solche Segment-abhängigen Fehler die Ergebnisse und deren Interpretation verfälschen. Um diese Abweichungen zu korrigieren, wird je Segment der Bias der Routengeschwindigkeiten separat für LKW und Busse abgezogen. Dadurch wird je Segment eine *korrigierte Routengeschwindigkeit* eingeführt, deren Median dem Median der Track-Geschwindigkeiten entspricht. Folglich ist diese korrigierte Routengeschwindigkeit je Segment biasfrei, wobei weiterhin die tageszeitlichen und wochentäglichen Schwankungen der Routengeschwindigkeit berücksichtigt werden.

All diese Geschwindigkeitsdaten liegen bisher auf sehr feingliedrigen Straßensegmenten vor mit einer mittleren Segmentlänge von 0,24 km. Um die Datenmenge zu reduzieren und zufällige, kurzfristige Geschwindigkeitsfluktuationen auszugleichen, werden die Segmente auf eine Länge in der Größenordnung von Kilometern zusammengefügt, wobei berücksichtigt wird, dass nur zusammenhängende Abschnitte mit ähnlichen Steigungen und Höhenlagen als ein ganzes Segment zusammengefasst werden. Dadurch ergibt sich für die Untersuchung eine mittlere Segmentlänge von 1,4 km, was auch in etwa einer minütlichen Auflösung der GPS-Daten entspricht (bei 90 km/h legt ein Fahrzeug in einer Minute 1,5 km zurück).

Neben dieser räumlichen Aggregation werden die Fahrten auch zeitlich zusammengefasst. Es kommt häufig vor, dass LKW nur kurz hintereinanderfahren und dadurch unnötig redundante Daten erzeugen. Bei der statistischen Analyse können solche stark miteinander korrelierte Datenpunkte zu einer überproportional hohen Gewichtung einer einzelnen Situation führen was die Ergebnisse verfälschen würde. Deshalb werden alle Fahrten je Segment innerhalb von 20-minütigen Zeitfenstern zusammengefasst.

Da Geschwindigkeiten je nach Streckenabschnitt und Fahrzeugtyp variieren, sind die absoluten Geschwindigkeiten der unterschiedlichen Segmente nicht vergleichbar und müssen für die (gemeinsame) Analyse normiert werden. Dafür werden *relative Track-Geschwindigkeiten* verwendet, welche durch das Verhältnis der Track-Geschwindigkeiten zu korrigierten Routengeschwindigkeiten definiert werden. Durch dieses Vorgehen werden sowohl Segmente mit unterschiedlichen Routengeschwindigkeiten als auch beider Fahrzeugtypen (Busse und LKW) mit unterschiedlichen Geschwindigkeiten miteinander vergleichbar gemacht und können gemeinsam untersucht und modelliert werden. Die Daten sind jedoch so aufbereitet, dass für das Training der Modelle Busse und LKW trotz Vergleichbarkeit getrennt betrachtet werden können.



**Abb. 1** Normierte Häufigkeitsverteilungen der relativen Track-Geschwindigkeit (einheitenlos) je Wetter: ohne Niederschlag, bei Regen (kein Schneefall und Niederschlag  $>2$  mm/h), bei Schneefall (Neuschnee  $>1$  cm/h) und bei Wind (kein Niederschlag und Windgeschwindigkeit  $>10$  m/s)

Die Häufigkeitsverteilungen der relativen Track-Geschwindigkeiten je Wetterklasse sind in Abb. 1 für alle verwendeten Daten dargestellt.

## 4 Wetterabhängige Routengeschwindigkeit

Nach abgeschlossener Datenvorbereitung wird ein Regressionsmodell trainiert, welches die Fahrgeschwindigkeit in Abhängigkeit von Wetter- und Routenvariablen beschreibt. Im Anschluss werden insbesondere die Einflüsse der Wettervariablen auf die Fahrgeschwindigkeit quantitativ untersucht. Zusätzlich wird ausgewertet ob die modellierten Routengeschwindigkeiten die tatsächlichen Fahrgeschwindigkeiten genauer prognostizieren können als die wetterunabhängigen und Bias-korrigierten Routengeschwindigkeiten (siehe Abschn. 3).

### 4.1 Modellierung eines wetterabhängigen Verzögerungsfaktors für die Routengeschwindigkeit

Um die Routengeschwindigkeit wetterabhängig zu modellieren, wird ein Regressionsmodell entwickelt, welches den allgemeinen Zusammenhang zwischen Wetter und Fahrgeschwindigkeit erfassen soll. Dieses Regressionsmodell beschreibt die relative Track-Geschwindigkeit, welche dem Verhältnis der Track-Geschwindigkeit zur korrigierten Routengeschwindigkeit entspricht. Diese relative Track-Geschwindigkeit kann auch als Verzögerungsfaktor interpretiert werden, wobei der Wert 1 der vollen

Routengeschwindigkeit entspricht und Werte unterhalb von 1 die entsprechende Verlangsamung angeben. In der Anwendung prognostiziert das Modell diesen Faktor, welcher in weiterer Folge in die Routenberechnung mit einfließt und entsprechend die Routengeschwindigkeit um den Wettereinfluss verringert.

Als Prädiktoren werden neben den oben beschriebenen Wettervariablen zusätzliche Attribute der Straßensegmente verwendet, wie Steigung, Kurvenradien, Brückenindikator und Routengeschwindigkeiten. Zusätzlich wird eine binäre Variable für die Unterscheidung Bus/LKW und die Uhrzeit eingeführt.

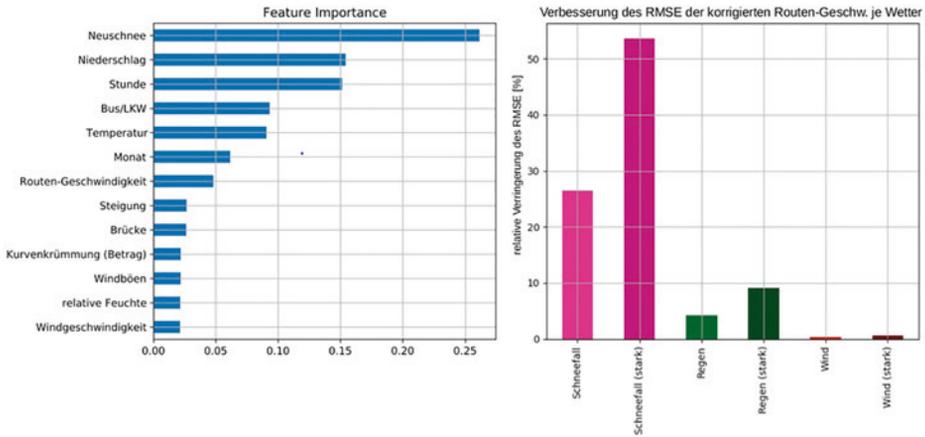
Damit die Prognosen unabhängig verifiziert werden können, werden die Daten in Trainings- und Testdaten aufgeteilt. Dafür werden aus dem Datensatz sowohl zufällige, ca. 10 km lange und zusammenhängende Streckenabschnitte für alle Fahrten (und somit für den gesamten Zeitraum), als auch 5-Tagesblöcke von den Trainingsdaten separiert und bilden einen unabhängigen Testdatensatz. Alle präsentierten Ergebnisse werden ausschließlich auf diesem Testdatensatz ausgewertet, welcher nicht für das Training des Modells verwendet wird.

Als statistisches Modell wurden *Gradient Boosted Trees* gewählt, welche besonders gut geeignet sind, da diese sowohl nichtlineare Zusammenhänge als auch mehrfache Abhängigkeiten gut erfassen können. Trainiert wird ein universelles Modell mit allen Segmenten aus dem Trainings-Datensatz, welche Streckenabschnitte verschiedener Autobahnen abdecken. Anschließend werden mit diesem Modell „Prognosen“ für Routengeschwindigkeiten der Testdaten berechnet und analysiert.

## 4.2 Wettereinflüsse und die Prognose der Fahrgeschwindigkeit

Für die Untersuchung der Wettereinflüsse auf die Fahrgeschwindigkeit werden mittlerer quadratischer Fehler (RMSE [10]) je Wetterklasse sowohl für die wetterunabhängigen und korrigierten, als auch für die wetterabhängigen Routengeschwindigkeiten berechnet und miteinander verglichen. Die aus diesem Vergleich resultierenden relativen Verbesserungen der RMSE durch die wetterabhängige Routengeschwindigkeitsmodellierung sind in Abb. 2 rechts dargestellt. Dieser Vergleich zeigt, dass besonders bei Schneefall (je nach Schneefallintensität) die wetterabhängige Modellierung zu deutlichen Verbesserungen von ca. 25 bis über 50 % führt. Bei Regen fallen die Verbesserungen mit ca. 4 bis 8 % deutlich geringer aus. Hingegen sind bei (starkem) Wind nur sehr marginale Verbesserungen zu erkennen.

Die Analyse der Prädiktoren (siehe Abb. 2 links) bestätigt, dass die Parameter Neuschnee und Niederschlag deutlich stärker die Fahrgeschwindigkeit beeinflussen als Wind und Windböen. Auffällig ist, dass Straßenparameter wie Steigung und Kurven eine geringe Bedeutung haben.



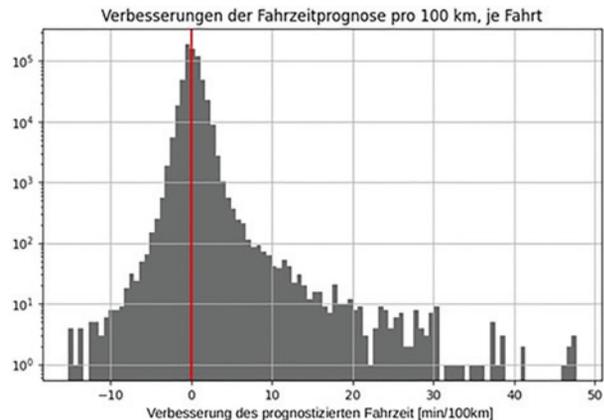
**Abb. 2** Gradient-Boosted-Tree-Modell: Relative Bedeutung der verwendeten Prädiktoren (links) und die erzielte, relative Reduktion der Wurzel des mittleren quadratischen Fehlers (RMSE) der korrigierten Routengeschwindigkeit, je Wetterklasse (rechts)

### 4.3 Verbesserung der Fahrzeitprognosen

Genauere Geschwindigkeitsprognosen führen zu genaueren Berechnungen der Fahrzeiten und der Ankunftszeiten (Estimated time of arrival, ETA). Für die Analyse der Wettereinflüsse auf Fahrzeiten werden für jede Fahrt die Fahrzeiten jeweils mit wetterunabhängigen und wetterabhängigen Routengeschwindigkeiten berechnet und mit den tatsächlichen Fahrzeiten verglichen. Anschließend wird betrachtet, wie sich die Abweichungen der beiden Routengeschwindigkeiten von den tatsächlichen Fahrzeiten unterscheiden.

Die Häufigkeitsverteilung dieser Unterschiede ist in Abb. 3 dargestellt, wobei die Unterschiede auf 100 km normiert sind, um Fahrten mit unterschiedlichen Distanzen

**Abb. 3** Häufigkeitsverteilung der Verbesserungen der Fahrzeitprognosen durch wetterabhängige Modellierung der Routengeschwindigkeit. Dargestellt sind die Verbesserungen in Minuten je Fahrt, normiert auf 100 km Strecke und auf der X-Achse die Anzahl der Fahrten, für die diese Verbesserung/ Verschlechterung zutrif



vergleichbar zu machen. Es fällt auf, dass die Anzahl der Fahrten mit verbesserten Fahrzeitprognosen deutlich diejenigen Fahrten überwiegt, bei denen es zu geringen Verschlechterungen der Prognosen durch die Hinzunahme von Wetter kommt. Bei vielen Fahrten lässt sich die ETA um 15 bis 30 min je 100 km Strecke genauer prognostizieren. Bei einzelnen Fahrten erreicht die Verbesserung sogar Werte um 45 min je 100 km. Wie zu erwarten ist, wird bei der Mehrheit der Fahrten die Fahrzeit kaum verändert da insgesamt Niederschlag und insbesondere Schneefall seltene Ereignisse darstellen.

---

## 5 Wetterabhängige Verzögerungswarnungen

Ziel der hier diskutierten Klassifikationsmethodik ist es, Warnungen auszulösen, wenn anhand einer Wettervorhersage erkannt wird, dass ein Streckenabschnitt aufgrund des Wetters zu einer langsameren Geschwindigkeit führen wird. Der zu trainierende Klassifikator soll lernen die wetterbedingten Eigenschaften so einzuschätzen, dass auf Basis einer Wettervorhersage vorausgesagt werden kann, ob eine Fahrt auf besagtem Abschnitt verlangsamt wird oder nicht.

Als Messgröße für die Verlangsamung betrachten die Autoren das Verhältnis zwischen Track-Geschwindigkeit und korrigierter Routen-Geschwindigkeit. Der Schwellwert für die Verlangsamung liegt hier bei 0,90, was einer Verlangsamung von 10 % entspricht. Erkennt man also auf einem Segment eine Geschwindigkeit, welche mindestens 10 % langsamer ist als die erwartete Geschwindigkeit auf diesem Segment, so wird dies als „Slow-Trip“ bezeichnet. Wird dieser Slow-Trip bei nachteiligem Wetter erkannt, so gilt dieser als wetterbedingt verlangsamt. Die Geschwindigkeiten auf Fahrten von LKW und Bussen werden hierbei getrennt betrachtet.

Die Autoren verstehen eine korrekte streckenabschnittsbezogene Prognose des Klassifikator-Modells entweder als korrekte Warnung oder korrekte Nicht-Warnung.

**Eine korrekte Warnung gilt dann**, wenn erkannt wird, dass nachteiliges Wetter in der aktuellen Wettervorhersage auf dem Streckenabschnitt zu Verlangsamung führen wird, wenn sehr häufig in der Vergangenheit wiederholt bei vergleichbaren Wetterverhältnissen langsamer gefahren wurde.

**Eine korrekte Nicht-Warnung gilt dann**, wenn erkannt wird, dass nachteiliges Wetter in der aktuellen Wettervorhersage auf dem Streckenabschnitt nicht zu Verzögerungen führen wird, wenn sehr selten in der Vergangenheit bei vergleichbaren Wetterverhältnissen langsamer gefahren wurde.

Das Klassifikator-Modell soll das Verhalten der Abschnitte auf Basis der Vergangenheit erlernen. Die Kombination aus verlangsamerter Fahrt bei nachteiligem Wetter bezeichnen wir als *Slow-Trip*. Die Kombination aus nichtverlangsamerter Fahrt bei nachteiligem Wetter bezeichnen wir hingegen als *Normal-Trip*.

## 5.1 Modellierung der Klassen und Aufteilung von Wetterdaten

Für die Klassifikation werden mehrere Segmente auf einer Route als Streckenabschnitte zusammengefasst. Für jeden Streckenabschnitt wird ein binärer Klassifikator als Modell für zwei Klassen trainiert: Die eine Klasse beschreibt wetterbedingt verlangsamten Fahrten auf diesem Abschnitt (Slow-Trip). Die andere Klasse beschreibt alle nicht verlangsamten Fahrten bei normalem Wetter, nicht verlangsamte Fahrten bei nachteiligem Wetter und verlangsamte Fahrten bei normalem Wetter (Normal-Trip).

Damit die Fahrten im Datensatz gutem bzw. schlechtem Wetter zugeordnet werden können, müssen ebenso die Wetterdaten in gut oder schlecht unterteilt werden. Es wäre jedoch hier keine gute Praxis eigene Aufteilungsregeln „per Hand“ zu definieren um die Wetterdaten aufzuteilen. Daher erfolgt das Aufteilen der Wetterdaten in zwei Gruppen (nachteiliges Wetter und normales Wetter) durch Anwendung einer als „*Unsupervised Clustering*“ bekannten Methode mithilfe des „Isolation-Forest“-Algorithmus. Hierbei soll der Algorithmus es eigenständig (daher unsupervised) schaffen, die Daten in die zwei Gruppen aufzuteilen. Der Algorithmus bildet aufgrund von in den Daten erkannten Ausreißern zwei Wetter-Gruppen. Auf Basis dieser beiden Wetter-Gruppen wurden nun alle Fahrten untersucht, die innerhalb dieser Gruppen stattgefunden haben.

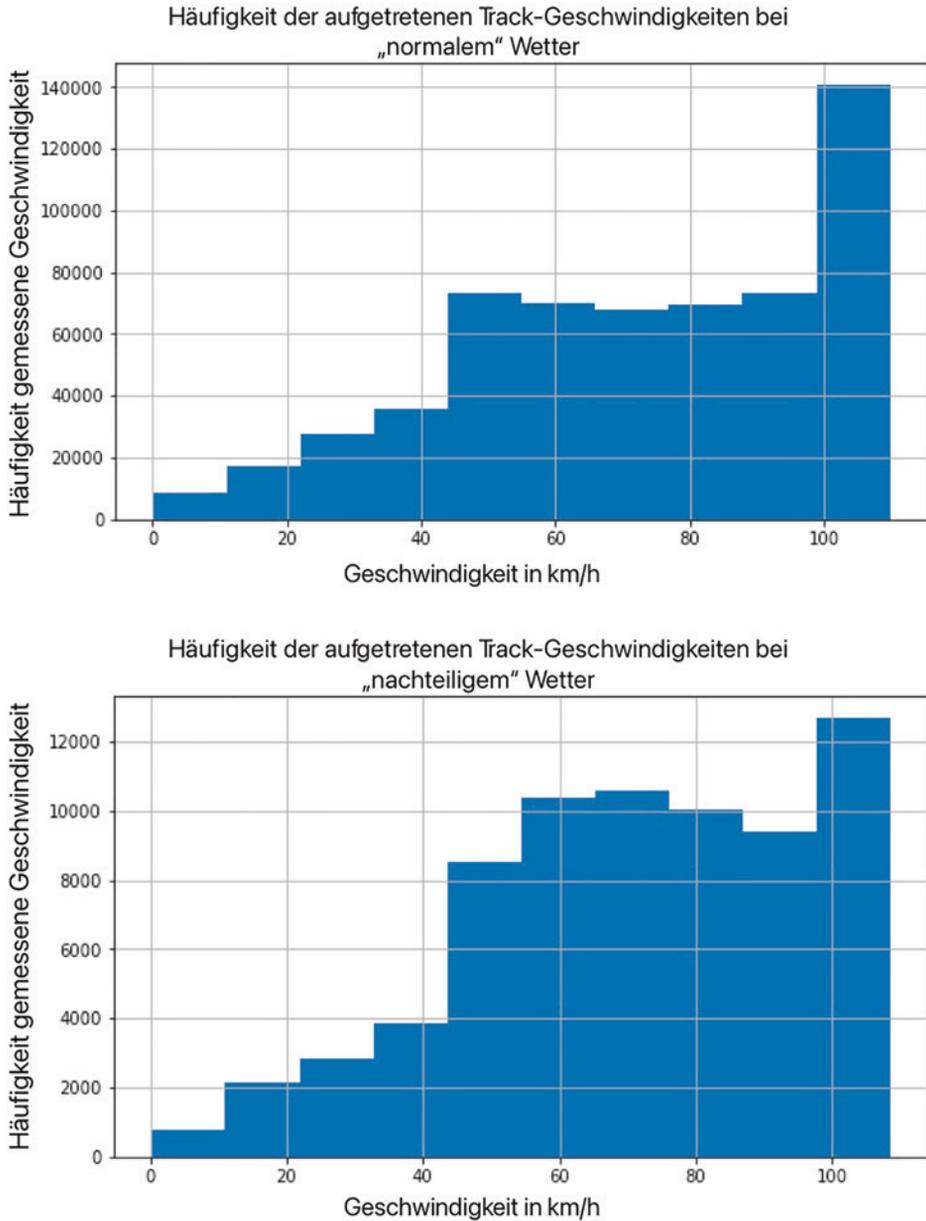
Der Erfolg dieser Methode wird in Abb. 4 ersichtlich. Hier zeigt sich eindeutig eine größere Anzahl an Fahrten je Segment mit niedrigeren Geschwindigkeiten (zwischen 40 und 90 km/h) in der Gruppe des nachteiligen Wetters im Vergleich zu den Fahrten in der Gruppe des normalen Wetters. Es lässt sich demnach durch die Gruppierung des Wetters die Aussage bestätigen, dass bei nachteiligem Wetter häufiger deutlich langsamer gefahren wurde als bei normalem Wetter. Weiterhin kann nun ein Klassifikator lernen ob ein neu gesehenes Wetter nachteilig für das Fahren ist oder nicht.

## 5.2 Auswahl der algorithmischen Klassifikatoren

Als mögliche Klassifikatoren aus dem klassischen maschinellen Lernen wurden unterschiedliche Algorithmen in Erwägung gezogen, trainiert und getestet. Die Evaluation erfolgte anhand der Metriken zur Trefferquote (Recall) und Präzision (Precision).

Die Trefferquote des Modells gibt eine Aussage darüber, wie häufig das Modell aus einem gesehenen Datensatz tatsächlich einen wetterbedingten Slow-Trip richtig erkennt und einordnet. Diese Metrik ist im Bereich der Warnsysteme dahingehend von großer Bedeutung, dass sie für die Häufigkeit steht, in der eine Warnung korrekterweise ausgelöst wird. Je höher die Trefferquote, umso häufiger werden Slow-Trips richtig vorhergesagt und somit wird korrekt gewarnt.

Die Präzision des Modells gibt eine Aussage darüber, wie präzise die Aussage des Modells ist, also wie häufig das Modell richtig liegt, wenn es einen Slow-Trip erkennt. Je höher die Präzision, umso genauer ist das Warnsystem bzw. umso seltener tendiert das Warnsystem zur Überwarnung.



**Abb. 4** Häufigkeitsverteilung der gemessenen Geschwindigkeiten bei normalem Wetter und nachteiligem Wetter. Es wird deutlich, dass bei nachteiligem Wetter mehr Geschwindigkeiten zwischen 40 und 90 km/h gefahren wurden als bei normalem Wetter

Untersucht wurden für diese Studien die Algorithmen: Decision Tree, Random Forest, K-Nearest Neighbours, Ada Boost, QDA (Quadratic Discriminant Analysis) und Naive Bayes. Letzterer scheint die beste Performance zu bieten und lieferte in allen Tests im Vergleich zu den anderen Algorithmen sowohl die beste Trefferquote als auch die besten Präzisionswerte auf Basis des verwendeten Datensatzes.

### 5.3 Auswertung des Wahrheitsgehalts des Klassifikators

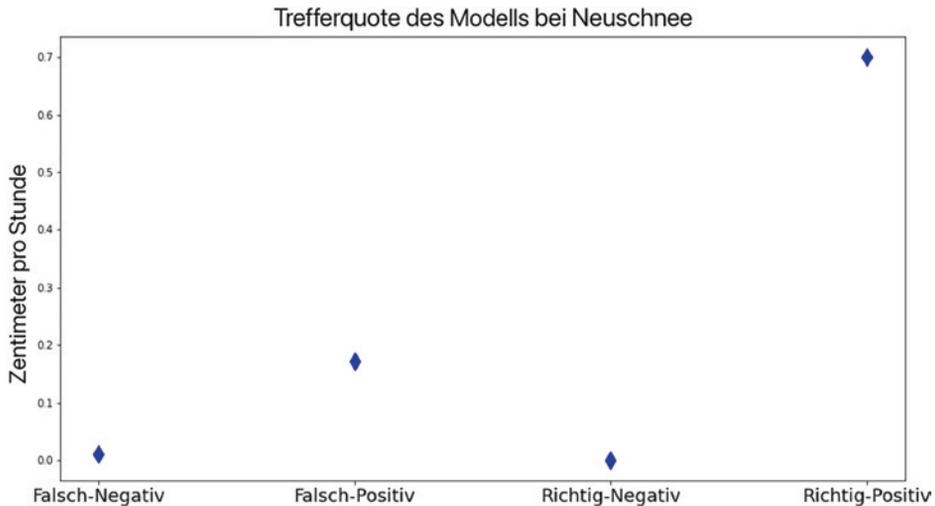
Des Weiteren betrachten die Autoren, wie präzise der Naive-Bayes-Algorithmus für jedes einzelne Streckensegment Prognosen generiert. Die Prognosen betreffen eine der beiden Klassen Slow-Trip (Verlangsamung auf dem Segment durch Wetter) oder Normal-Trip (keine Verlangsamung auf dem Segment durch Wetter).

Um eine Aussage darüber treffen zu können, wie präzise die Modelle pro Segment agieren, werden die Modelle nach dem Training auf einen Test-Datensatz angewandt. Hierbei bekommt ein Modell ein vorher unbekanntes Wetterphänomen zu sehen, woraufhin der Wettereinfluss auf das örtlich betroffene Segment bewertet wird. Die prognostizierte Klasse wird dann mit der tatsächlichen Klasse verglichen, um so zu erkennen wie viele richtige oder falsche Annahmen getroffen wurden. Die Annahmen oder auch Bewertungsfälle, welche das Modell trifft, sind die folgenden:

- **Richtig Positiv:** Das Modell prognostiziert aufgrund des gesehenen Wetters für ein Segment richtigerweise einen Slow-Trip (mit Warnung als Folge).
- **Richtig Negativ:** Das Modell prognostiziert aufgrund des gesehenen Wetters für ein Segment richtigerweise einen Normal-Trip, erwartet also keine Verlangsamung. Eine Warnung bleibt aus.
- **Falsch Positiv:** Das Modell prognostiziert fälschlicherweise einen durch Wetter verursachten Slow-Trip. Es wird eine Fehlwarnung generiert.
- **Falsch Negativ:** Das Modell erkennt einen durch Wetterverursachten Slow-Trip nicht und prognostiziert fälschlicherweise einen Normal-Trip. Die notwendige Warnung fällt demnach aus.

Die Autoren betrachten zur Auswertung der Aussagekraft die Trefferquote des Modells in Bezug auf ein ausgewähltes Wetterphänomen. Dies ist für das Phänomen Neuschnee zu sehen in Abb. 5.

Es ist zu erkennen, dass eine Richtig-Positiv-Aussage des Modells bei durchschnittlich 0,7 cm Neuschnee pro Stunde erfolgt. Es ist ebenso zu erkennen, dass eine Richtig-Negativ-Aussage des Modells durchschnittlich bei einem Neuschnee Wert von ca. 0,1 cm pro Stunde erfolgt. Man kann demnach sagen, dass das Modell bei einer hohen Menge an Neuschnee richtig warnt, bzw. bei geringen Mengen an Neuschnee richtig *nicht* warnt. Es ist ebenso zu erkennen, dass Falsch-Positiv Aussagen des Modells (Fehlwarnung) bei durchschnittlich bis zu 0,2 cm pro Stunde erfolgen. Bei einer Neuschnee-



**Abb. 5** Nach Anwendung der Testdaten wurden die Aussagen des Modells nach Bewertungsfall sortiert. Innerhalb dieser Fälle wurde dann gemessen bei welchem Wetterwert im Durchschnitt die Aussage getroffen wurde. Die beste Trefferquote lässt sich bei durchschnittlich 0,7 cm/Std. Neuschnee feststellen

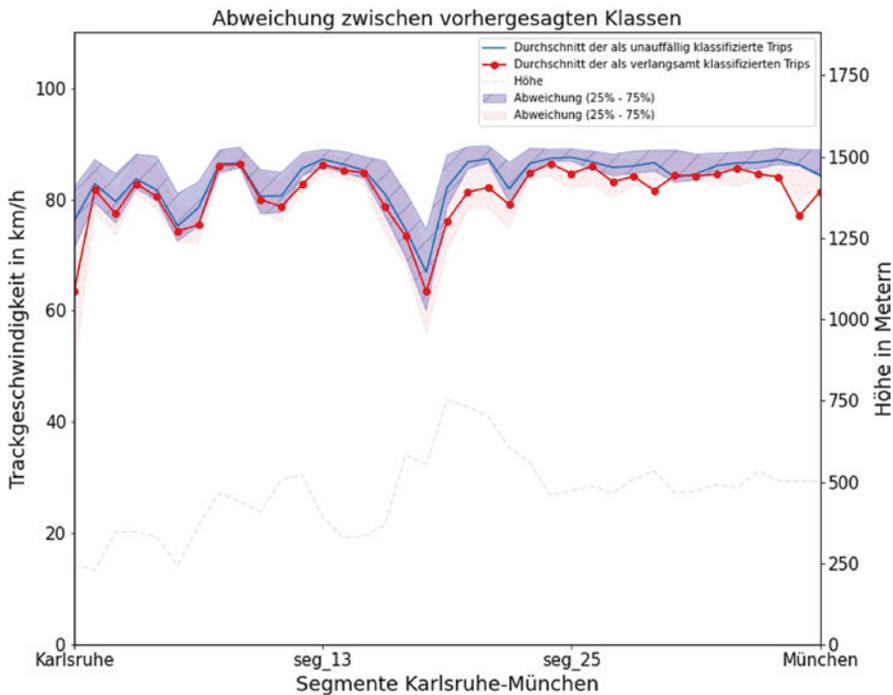
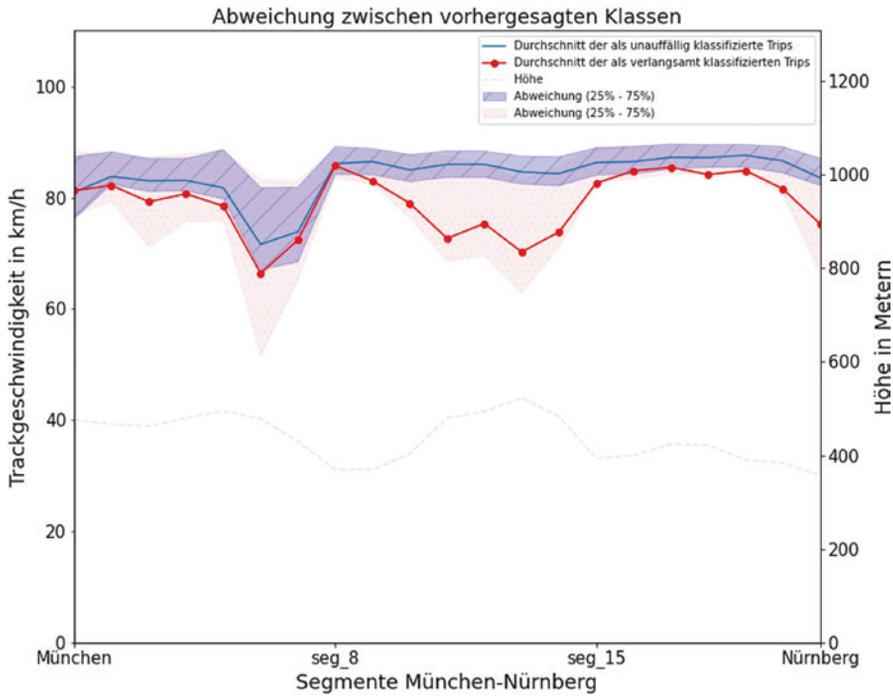
menge bis zu diesem Wert sind demnach Unsicherheiten des Modells zu erwarten, die Situation richtig einzuschätzen. Je mehr Schnee fällt umso eher warnt das Modell. Dies hat ebenso zur Folge, dass Situationen, in denen schon sehr wenig Schneefall zu Verzögerungen führen kann, nicht immer erkannt werden. Dies ist erkennbar beim durchschnittlichen Schneefall von 0,1 cm pro Stunde. Verlangsamungen, die bereits in diesem Bereich stattfinden, können durch das Modell nur schwer erkannt werden.

Abschließend lässt sich daher sagen, dass der Klassifikator durchaus korrekte Frühwarnungen aussprechen kann bzw. korrekte *nicht*-Warnungen, wenn eine gewisse Menge an Schnee fällt.

## 5.4 Anwendung des Klassifikators auf Beispielstrecken

Im vorherigen Abschnitt konnte gezeigt werden, dass ein Modell Verzögerungen aufgrund von Neuschnee gut erkennen und vorhersagen kann. Folglich wird nun die Warnfähigkeit des Modells pro Segment im Verlauf einer kompletten Route betrachtet.

Abb. 6 präsentiert die Abweichung zwischen den vorhergesagten Klassen. Zu sehen ist hier der Durchschnitt der als unauffällig klassifizierten Trips (Normal-Trips) und der Durchschnitt der verlangsamten Trips (Slow-Trips). Ersteres wird mit durchgezogener, letzteres wird mit durchgezogen-gepunkteter Linie dargestellt. Die Höhe der Linien ist immer bezogen auf die Geschwindigkeit für jedes Segment auf der Route zwischen Start und Ziel. Auf jedem Segment auf dem sich die Slow-Trip Linie unterhalb



**Abb. 6** Standardabweichung zwischen den Klassen Normal-Trip und Slow-Trip (durchgezogene Linie: Fahrten ohne Verzögerung bzw. Normal Trip, durchgezogen-gepunktete Linie: durch das Modell erkannte Fahrten mit wetterbedingter Verzögerung nzw. Slow Trip; Je weiter beide Linien auf einer Strecke auseinanderliegen, umso besser sind die Frühwarnungen der Modelle)

der Normal-Trip Linie befindet, hat das für dieses Segment trainierte Modell die Verlangsamung korrekt erkannt und gewarnt. Liegen beide Linien aufeinander, kann für dieses Segment kein eindeutiger Unterschied bei winterlichem und nicht-winterlichem Wetter gelernt bzw. erkannt werden.

Je weiter die Linien auseinanderliegen (mit der Slow-Trip-Linie unterhalb der Normal-Trip-Linie), umso besser ist das Modell in der Lage, wetterbedingte Verlangsamungen eindeutig zu erkennen.

Betrachtet man nun jeden Punkt auf der Slow-Trip-Linie als Warnung eines der Modelle pro Segment, lässt sich feststellen, dass auf der Strecke zwischen München und Nürnberg (größtenteils Autobahn A6) ein Frühwarnsystem den Einfluss des Wetters sehr gut einschätzen kann. Jedes Mal, wenn eines der segmentbezogenen Modelle einen Slow-Trip auf einem Segment bei winterlichen Verhältnissen prognostiziert, liegt die tatsächlich gemessene Geschwindigkeit auf diesem Segment bei winterlichen Verhältnissen tatsächlich unterhalb der als normal gewerteten Geschwindigkeit. Am besten funktionieren die Modelle für alle Segmente zwischen Segment 8 und Segment 15. Hier wird bei winterlichen Verhältnissen sehr viel langsamer gefahren, was durch die Modelle klar wiedergegeben wird. Die Autobahn A6 scheint demnach für Bus und LKW bei winterlichen Verhältnissen anfälliger zu sein. Diese Anfälligkeit lässt sich trainieren und anwenden.

Im Gegensatz zur Strecke zwischen München und Nürnberg, zeigt das Beispiel der Strecke zwischen Karlsruhe und München (größtenteils Autobahn A8) eine andere Aussagekraft des Modells. Hier ist zu erkennen, dass die Linien oftmals sehr nahe beieinanderliegen, sich teilweise sogar überlappen. Die Modelle schlagen demnach je Segment bei ähnlichen Geschwindigkeiten aus. Hier ist demnach auf den meisten Segmenten kein großer Geschwindigkeitsunterschied bei winterlichen Verhältnissen zu erwarten. Die Aussagen der Modelle bezüglich Slow-Trips, treffen daher auch nicht immer eindeutig den Bereich der Verlangsamung.

Da die Autobahn A8 in den Daten kaum eindeutige Geschwindigkeitsunterschiede bei winterlichen Verhältnissen aufweist, können diese auch nicht von den Modellen gelernt und angewendet werden. Daher ist basierend auf dieser Datenlage mit wesentlich mehr Fehlmeldungen zu rechnen.

---

## 6 Diskussion und Schlussfolgerung

Die Autoren haben gezeigt, dass auf mitteleuropäischen Autobahnen das Wetter einen Einfluss auf die Fahrgeschwindigkeit von LKW und Bussen hat. Mit statistischer Modellierung und maschinellem Lernen lassen sich diese Wettereinflüsse bestimmen und die Zusammenhänge mit Routen- und Fahrtgeschwindigkeiten quantifizieren. In weiterer Folge können diese gelernten Zusammenhänge mit Wetterprognosedaten auf zukünftige Routenberechnungen und Warnsysteme übertragen werden. Die Berücksichtigung all dieser Wettereinflüsse erhöht auch die Genauigkeit der Fahrzeitprognose und somit auch der ETA.

Für die Prognose der wetterabhängigen Routengeschwindigkeit wird ein universelles, segment-unabhängiges Modell entwickelt. Durch die Unterteilung aller Segmente in Trainings- und Testdaten konnte gezeigt werden, dass sich (zumindest innerhalb von Mitteleuropa) die Ergebnisse auch auf neue und unbekannte Autobahnabschnitte übertragen lassen. Des Weiteren kann der gesamte (Trainings-)Datensatz für die Modelloptimierung genutzt werden. Dadurch kann das Modell bereits mit dem nur knapp 3-jährigen Datensatz auch die Wettereinflüsse von seltenen Ereignissen wie z. B. starken Schneefällen lernen.

Für die Frühwarnung vor wetterbedingten Verzögerungen wurde ein multi-model Ansatz gewählt, der für jeden Streckenabschnitt ein spezifisches eigenes eigenständiges Modell trainiert, welches den in der Vergangenheit gemessenen Einfluss winterlichen Wetters auf diesen Abschnitt erlernt. Hier lernen die Modelle nicht nur regionale Unterschiede, sondern auch lokal-spezifische Effekte und individuelle nicht auf andere Regionen übertragbare Eigenschaften eines Streckenabschnittes kennen, welche den Einfluss des Wetters verstärken oder abschwächen. Hier konnten vor allem bei Neuschnee (also aktiv fallendem Schnee) für LKW und Busse sehr gute Warnfähigkeiten erzielt werden, wenn negative Einflüsse auf die Geschwindigkeit durch Wetter auch in der Vergangenheit aufgetreten sind.

## 6.1 Diskussion

Sowohl die explorativen Datenanalysen (vgl. Abb. 1) als auch die Analyse der Modellergebnisse (vgl. Abb. 2) zeigen, dass Wetterparameter unterschiedlich stark die Fahrgeschwindigkeit beeinflussen. Erwartungsgemäß ist bei Schneefall der Einfluss am größten, wobei auch die Varianz der relativen Track-Geschwindigkeiten (meist zw. 0,4 und 1,0) am größten ist. Diese Schwankungen lassen sich primär mit der Schneefallintensität erklären. Da zur Modellierung ein Baum-basiertes Modell des maschinellen Lernens verwendet wird, können sowohl nicht-lineare Abhängigkeiten der Schneefallintensität, als auch kombinierte Einflüsse von mehreren Variablen vom Modell erfasst werden. Es muss also auch berücksichtigt werden, dass kombinierte Effekte mit Wind (Schneeverwehungen) bzw. Temperatur (abgesenkter Gefrierpunkt von Wasser durch Salzstreuen) ebenfalls eine wichtige Rolle spielen.

Regen führt meist zu einer geringen, aber signifikanten Verlangsamung um ca. 5 %. Im Gegensatz dazu zeigt Wind – abgesehen in Kombination von Schneefall – einen vernachlässigbar geringen Einfluss auf die Fahrgeschwindigkeit.

Neben Schneefall und Regen ist auch anzunehmen, dass reduzierte Sichtweiten durch Nebel die Fahrgeschwindigkeit verringern. Allerdings konnte dieser Zusammenhänge in dieser Arbeit nicht untersucht werden, weil keine Daten vom Nebel oder der Sichtweite verfügbar sind. Da es sich bei Nebel um ein sehr lokales Wetterphänomen handelt wären für solch eine Untersuchung Mess- oder Beobachtungsdaten direkt von den Autobahnen notwendig.

Zusätzlich zum Wetter wurden auch Routenparameter und ein binärer Parameter zur Unterscheidung von Bus und LKW bei der Modellierung des wetterabhängigen Verzögerungsfaktors verwendet. Die Analyse des trainierten Modells (siehe Abb. 2 links) deutet einen relativ großen Einfluss des Prädiktors zur Unterscheidung LKW-Bus an, während die Routenparameter (Steigung, Brücke und Kurvenkrümmung) geringere Einflüsse zeigen. Zur weiteren Untersuchung wird auch ein Modell ohne Routenparameter und ohne den Parameter zur Unterscheidung von Bus-LKW trainiert. Durch das Entfernen dieser Prädiktoren verringern sich die in Abb. 2 rechts gezeigten Verbesserungen bei Schneefall und Regen geringfügig um jeweils ca. 2 Prozentpunkte. Diese nur geringe Verschlechterung deutet darauf hin, dass bei der Modellierung der Wettereinflüsse (unter Verwendung der relativen-Track-Geschwindigkeit, welche implizit die absoluten Geschwindigkeitsunterschiede berücksichtigt) eine Unterscheidung zwischen Bus und LKW nicht notwendig erscheint.

## 6.2 Ausblick

Für verbesserte Fahrzeitprognosen zukünftiger Fahrten lassen sich wetterabhängige Verzögerungsfaktoren mit Wetterprognosedaten berechnen und in ein Routingsystem integrieren. Für eine verbesserte Resilienz vor Wettereinflüssen im Straßenverkehr lassen sich Frühwarnsysteme in Entscheidungsunterstützungssysteme und Dashboards integrieren. Die in dieser Studie gezeigten Ansätze zeigen beide eine starke Relevanz für aktuelle Fragestellungen im Logistikbereich. Bis es jedoch zur unmittelbaren Anwendbarkeit der Lösungsansätze kommt, bedarf es noch weiterer Untersuchungen mit unterschiedlichen Datenzusammensetzungen und Einflussgrößen. Auch muss untersucht werden inwieweit regionale Unterschiede stärker berücksichtigt werden müssen.

Da die Studie aufzeigt, dass über beide Ansätze hinweg Schneefall mit dem größten Einfluss hervortritt, kann der größte Nutzen für Applikationen auf Basis dieser Technologie in schneereichen Regionen wie beispielsweise in höher gelegenen Mittelgebirgsregionen oder bei Fahrten im Bereich der Alpen erwartet werden.

**Danksagung** Die Autoren möchten sich an dieser Stelle recht herzlich bei DB Schenker und FlixBus für die Bereitstellung der Fahrzeugdaten von Reisebussen und Lastkraftwagen als auch bei der Firma INFOWARE Informationstechnik GmbH für die Aufbereitung der Fahrzeug- und Streckendaten bedanken. Weiterhin geht ein großer Dank an die UBIMET GmbH aus Wien für die Bereitstellung der historischen und aktuellen Wetterdaten sowie an das Fraunhofer FOKUS Data Analytics Center in Berlin für das zur Verfügung stellen von Rechenkapazitäten. Diese Publikation wurde im Rahmen des mFUND-Programms vom Bundesministerium für Digitales und Verkehr kofinanziert (FKZ 19F2085A, 19F2085C, 19F2085D, 19F2085E, 19F2085G). Weitere Informationen zum Projekt MeteoValue Live sind zu finden unter (<https://www.meteo-value-live.de/>) sowie in den Ausführungen des BMDV auf [11].

## Literatur

1. Konstantinou, K. (2019). *Calculation of estimated time of arrival using artificial intelligence*. Chalmers University of Technology.
2. Rashid, O., Coulton, P., Edwards, R., Fisher, A., & Thompson, R. (2005). Mobile information systems providing estimated time of arrival for public transport users. *61st Vehicular Technology Conference. IEEE* (S. 2765–2769).
3. Agarwal, M., Maze, T. H., & Souleyrette, R. (2005). Impacts of weather on urban freeway traffic flow characteristics and facility capacity. *Proceedings of the 2005 mid-continent transportation research symposium, Iowa State University, Ames*.
4. Kyte, M., Khatib, Z., Shannon, P., & Kitchener, F. (2000). Effect of environmental factors on free-flow speed. In W. Brilon (Hrsg.), *Proceedings 4th International symposium on highway capacity, Transportation Research Board* (S. 108–118).
5. Zafar, N., & Ul Haq, I. (2020). Traffic congestion prediction based on estimated time of arrival. *PloS (Public Library of Science) one*, 15(12), e0238200.
6. Thakuriah, P., & Tilahun, N. (2013). Incorporating weather information into real-time speed estimates: Comparison of alternative models. *Journal of Transportation Engineering*, 139(4), 379–389. [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000506](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000506).
7. Maze, T. H., Agarwai, M., & Burchett, G. (2006). Whether weather matters to traffic demand, traffic safety, and traffic operations and flow. *Transportation research record 1948, Transportation Research Board, Washington, DC*.
8. Saberi, M., & Bertini, R. L. (2010). Empirical analysis of the effects of rain on measured freeway traffic parameters. *89th annual meeting of the transportation research board, Transportation Research Board, Washington, DC*.
9. Huang, S.-H., & Ran, B. (2003). An application of neural network on traffic speed prediction under adverse weather condition. *82nd annual meeting of the transportation research board, Transportation Research Board, Washington, DC* (S. 1–21). [http://www.researchgate.net/profile/Bin\\_Ran/publication/265318230\\_An\\_Application\\_of\\_Neural\\_Network\\_on\\_Traffic\\_Speed\\_Prediction\\_Under\\_Adverse\\_Weather\\_Condition/links/54999d8e0cf22a83139625a2.pdf%5Cnhttp://www.ltrc.lsu.edu/TRB\\_82/TRB2003-000915](http://www.researchgate.net/profile/Bin_Ran/publication/265318230_An_Application_of_Neural_Network_on_Traffic_Speed_Prediction_Under_Adverse_Weather_Condition/links/54999d8e0cf22a83139625a2.pdf%5Cnhttp://www.ltrc.lsu.edu/TRB_82/TRB2003-000915).
10. Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE). *Geoscientific Model Development Discussions*, 7(1), 1525–1534.
11. Bundesministerium für Digitales und Verkehr. (2019). *Optimierung der Einsatz- und Routenplanung von Speditions- und Fernbusunternehmen unter der Berücksichtigung vorhergesagter Schlechtwetterbedingungen und Parkplatzverfügbarkeiten – MeteoValue live*. <https://www.bmvi.de/SharedDocs/DE/Artikel/DG/mfund-projekte/meteo-value-life.html>.

Die Kapitel 3 und 4 werden unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht. Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation in den Kapiteln.

