

# Eine Bestimmung der Oberflächenqualität von Fahrradinfrastruktur durch Smartphone-Beschleunigungsdaten mithilfe des k-means++-Algorithmus

## *Determination of the Surface Quality of the Bicycle Infrastructure by Smartphone Acceleration Data Using the k-means++ Algorithm*

Stefan Kranzinger<sup>1</sup>, Sven Leitinger<sup>1</sup>

<sup>1</sup>Salzburg Research Forschungsgesellschaft mbH · stefan.kranzinger@salzburgresearch.at

**Zusammenfassung:** Die Bereitstellung einer gut instandgehaltenen Fahrradinfrastruktur wirkt sich positiv auf den Fahrkomfort sowie die Sicherheit im Straßenverkehr aus und schafft die Voraussetzung, dass dieses als Alternative zum motorisierten Individualverkehr oder zum öffentlichen Verkehr angenommen wird. In dieser Studie wird mit Daten aus Smartphone-Beschleunigungssensoren und mithilfe eines k-means++-Algorithmus ein Clusterverfahren angewandt, um die Qualität von Fahrbahnabschnitten festzustellen. Dieses ist einfach und kostengünstig auf großflächige Radinfrastrukturen anwendbar und zeigt in einer Fallstudie für die Stadt Salzburg, dass schlecht befahrbare Fahrbahnabschnitte gut lokalisiert und infolge gezielt instandgehalten werden können.

**Schlüsselwörter:** Fahrbahnoberflächenqualität, Smartphone-Beschleunigungssensoren, k-means++-Clustering

**Abstract:** *Providing a well-maintained cycling infrastructure has a positive impact on the comfort and safety of cycling in road traffic and creates the conditions for its acceptance as an alternative to private or public transport. This study uses data from smartphone accelerometers and a k-means++ algorithm to determine the quality of road sections. This is easy and cost-effective to apply to large-scale cycling networks and shows in a case study for the city of Salzburg that poorly passable lane sections can be well localised and subsequently maintained in a targeted manner.*

**Keywords:** *Road surface roughness, smartphone accelerometers, k-means++ clustering*

## 1 Motivation

Eine gut instandgehaltene Fahrradinfrastruktur ist die Voraussetzung, dass diese als Alternative zum motorisierten Individualverkehr oder öffentlichen Verkehr angenommen wird. Die Qualität der Fahrradinfrastruktur hängt im großen Maße vom Zustand der Fahrbahnoberfläche ab. Ist diese von guter Qualität wirkt sich das laut Vittorio et al. (2014) positiv auf die Sicherheit und den Komfort der Verkehrsteilnehmer\*innen aus.

Um eine gute Qualität der Fahrbahnoberfläche zu gewährleisten, ist es notwendig Fahrbahnschäden zu lokalisieren, damit diese infolge gezielt instandgehalten werden kann. Die nachfolgende Studie stellt daher eine Methode zur Qualitätsmessung der Fahrbahnoberfläche von Fahrradinfrastruktur vor. Hierfür werden in einer Fallstudie Daten von Smartphone-Beschleunigungssensoren für das Hauptradwegenetz der Stadt Salzburg gesammelt. Diese Daten werden anschließend mit einem k-means++-Algorithmus verarbeitet, um mithilfe eines Clusterverfahrens Fahrbahnabschnitte automatisiert in Gruppen mit guter und schlechter

Qualität einzuteilen. Ziel ist es eine leicht umsetzbare und kostengünstige Methode für die Identifikation von schlecht befahrbaren Fahrbahnabschnitten einer beliebigen Fahrradinfrastruktur vorzustellen. Somit soll die Instandhaltung erleichtert und die Attraktivität von Fahrradnetzen erhöht werden.

Nachfolgend wird in dieser Studie auf die relevante Forschung der behandelten Fragestellung eingegangen. Weiterhin werden die Datenaufzeichnung, das Testsetting und das verwendete Clusterverfahren vorgestellt. Danach folgen eine Darstellung der Fallstudie, die Beschreibung und die Diskussion der Ergebnisse.

## 2 Stand der Forschung

In der Literatur gibt es eine Reihe von Arbeiten, in welchen mithilfe von Daten aus Smartphone-Beschleunigungssensoren die Fahrbahnqualität der Fußgänger\*innen-, Fahrrad- sowie Personenkraftwagen (PKW)-Infrastruktur bestimmt wird. Die Verwendung von Smartphone-Beschleunigungssensoren sind laut Zang et al. (2018) besonders deshalb interessant, weil auf Fußgänger\*innen- und Fahrrad-Infrastruktur die ansonsten üblichen Technologien, wie fahrzeugmontiertes Laserscanning, oft nicht verwendet werden können. Weiters argumentieren Douangphachanh & Oneyama (2013), die kostengünstige und leicht umzusetzende Sammlung von Daten mithilfe von Smartphone-Beschleunigungssensoren. Sattar et al. (2018) geben in einem Literaturreview einen Überblick über Studien in Bezug zur Bewertung der Fahrbahnoberfläche durch Smartphone-Sensoren. In diesem Kapitel soll durch die Vorstellung ausgewählter Arbeiten zu diesem Thema der Stand der Forschung genauer beleuchtet werden.

Douangphachanh & Oneyama (2013), verwenden Smartphone-Beschleunigungsdaten, welche mithilfe von PKW gesammelt wurden, um die Fahrbahnqualität zu messen. Die Ergebnisse der Autoren zeigen eine lineare Korrelation zwischen der gemessenen Beschleunigung und der Fahrbahnunebenheit, welche durch den *International Roughness Index* (IRI) (Sayers, 1995; Smith & Ram, 2016) repräsentiert wird. Die Korrelation hängt von der gefahrenen Geschwindigkeit ab und ist am signifikantesten bei Geschwindigkeiten mit weniger als 60 km/h. Sadjadi (2017) untersucht mithilfe von Daten aus Smartphone-Beschleunigungssensoren die Fahrbahnqualität von Autostraßen. Die Ergebnisse des Autors zeigen, dass die gemessene Fahrbahnunebenheit stark mit der subjektiven Einschätzung der fahrenden Person sowie mit dem IRI korreliert. Hari Krishnan & Gopi (2017) klassifizieren mithilfe von Smartphone-Beschleunigungssensoren Fahrbahnschäden wie Schlaglöcher oder Fahrbahnbuckel für Autostraßen. Die Ergebnisse der Autoren zeigen, dass bei einer Geschwindigkeit zwischen 15 und 20 km/h die genaueste Schätzung der Stärke solcher Fahrbahnschäden möglich ist. Li & Goldberg (2018) liefern ein weiteres Beispiel für die Detektion von Fahrbahnschäden und die Bewertung der Fahrbahnqualität von PKW-Infrastruktur durch Smartphone-Beschleunigungssensoren. Als größte Herausforderung zur akkuraten Bestimmung der Fahrbahnqualität aus unterschiedlichen Testfahrten, nennen die Autoren die Genauigkeit des Smartphone-GNSS-Empfängers. Hier kann es durch Abweichungen des GNSS-Empfängers zu Ungenauigkeiten in der Lokalisierung von Fahrbahnschäden kommen.

Zang et al. (2018) beschreiben eine Methode zur Bewertung der Oberflächenqualität von Fußgänger\*innen- und Fahrradinfrastruktur. Unter Verwendung des IRI, wurde von den Autoren ein Algorithmus zur Erkennung von Schlaglöchern bzw. Fahrbahnbuckel entwickelt.

Die Resultate des Algorithmus basieren auf GNSS- und Beschleunigungsdaten von Smartphones und zeigen eine gute Detektion und Lokalisierung von Schlaglöchern und Fahrbahn-buckel. Litzenberger et al. (2018) versuchen die Fahrbahnqualität einer Fahrradinfrastruktur vorauszusagen. Für die Datenaufzeichnung werden zwei verschiedene Beschleunigungssensoren sowie unterschiedliche Kalibrierungen (Geschwindigkeit und Reifendruck) verwendet. Die Ergebnisse zeigen, dass die Fahrbahnqualität mithilfe der Daten eines Smartphone-Beschleunigungssensors mit einer Genauigkeit von mehr als 97 % prognostiziert werden kann. Wijerathne et al. (2018) präsentieren einen Algorithmus mit welchem Fahrbahnunebenheiten einer Fahrradinfrastruktur, durch die Analyse von Daten aus Smartphone-Beschleunigungssensoren, erkannt werden können. Die Autoren verwenden eine Wavelet-Transformation, um anhand einer übermäßigen Rauschvarianz der Beschleunigungsdaten und des Lipschitz-Exponenten, Fahrbahnunebenheiten zu detektieren.

Die Literaturrecherche ergibt somit, dass es mit kostengünstigen Smartphone-Beschleunigungssensoren möglich ist, die Fahrbahnqualität darzustellen bzw. Fahrbahnunebenheiten zu detektieren. Es gibt Anwendungen in Bezug auf Fußgänger\*innen-, Fahrrad- sowie PKW-Infrastruktur. Ein Großteil der Studien verwendet den IRI, um die Fahrbahnoberflächenqualität zu bestimmen, wohingegen für die Klassifikation von Fahrbahnschäden komplexe Algorithmen wie Wavelet-Transformation verwendet werden.

Diese Studie versucht mithilfe von Smartphone-Beschleunigungssensoren die Qualität der Fahrradinfrastruktur darzustellen. Hierzu sollen, anstatt des in der Literatur viel verwendeten IRI, Features mithilfe von Beschleunigungsdaten berechnet werden, mit denen eine Klassifikation der Fahrbahnoberfläche vorgenommen werden kann. Somit soll eine Alternative Methode zum IRI erforscht werden, mit welcher die Fahrbahnoberfläche bestimmt werden kann. Als Methode für die Klassifizierung wird ein k-means++-Algorithmus verwendet, da dieser ein einfach zu implementierendes und leistungsstarkes Verfahren darstellt, um Fahrbahnabschnitte zu klassifizieren. Somit soll in dieser Studie folgende Forschungsfrage beantwortet werden:

*Können mithilfe von aus Smartphone-Beschleunigungsdaten berechneten Features und der Anwendung eines k-means++-Algorithmus, die Abschnitte einer Fahrradinfrastruktur gemäß ihrer Qualität bestimmt werden?*

Anschließend wird im Detail auf die verwendeten Daten und die angewendete Methode eingegangen.

### **3 Daten und Methodologie**

#### **3.1 Aufzeichnung und Weiterverarbeitung der Daten**

Die Datenaufzeichnung erfolgt mithilfe einer Smartphone-App, welche bereits in anderen Studien (Brunauer & Rehr, 2016; Brunauer et al., 2018) zur Aufzeichnung von Mobilitätsdaten verwendet wurde. Als Datenoutput werden GNSS-Koordinaten, Zeitstempel und für jeden identifizierten Ausschlag eine Beschleunigung in  $\text{m/s}^3$  erfasst. Diese Daten werden anschließend in einem von uns selbst entwickelten cloudbasierten Analysedienst zur Analyse von Trajektorien auf der Open Telekom Cloud (“Open Telekom Cloud,” n. d.) weiterverarbeitet. Durch dieses System werden die gesammelten Daten durch die Funktion der Ausrei-

ßer-Erkennung und -Entfernung, einen Kartenabgleich und der Generierung von abschnittsbasierten Befahrungsinformationen weiterverarbeitet (Brunauer et al., 2018).

Brunauer & Rehl (2016) liefern eine detaillierte Beschreibung dieser Verarbeitung, in welcher durch heuristische Regeln Beschleunigungen, Geschwindigkeiten und Positionsausreißer korrigiert werden. Die Autoren merken an, dass die gesammelten Daten auf einem dreidimensionalen Beschleunigungssignal ( $m/s^2$ ), welches mit 50 Hz aufgezeichnet wurde, sowie einer GNSS-Positionsmessung mit einer Abtastrate von 1 Hz basieren. Für die Synchronisation der gesammelten Daten wird der GNSS-Zeitstempel zum nächstgelegenen Zeitstempel des Beschleunigungsmessers zugeordnet. Um eine Beschleunigungsmessung zwischen zwei GNSS-Positionsmessungen auf einem Straßennetzgraphen zuzuordnen, wird laut Brunauer & Rehl (2016) ein abstands-basiertes lineares Referenzierungsverfahren angewendet. Mit diesem kann jede Beschleunigungsmessung einem linearen Referenzpunkt auf einem Segment, inklusive Fahrtrichtung und der relativen Position auf dem Segment in Form eines Offsets vom Segmentanfang, zugewiesen werden.

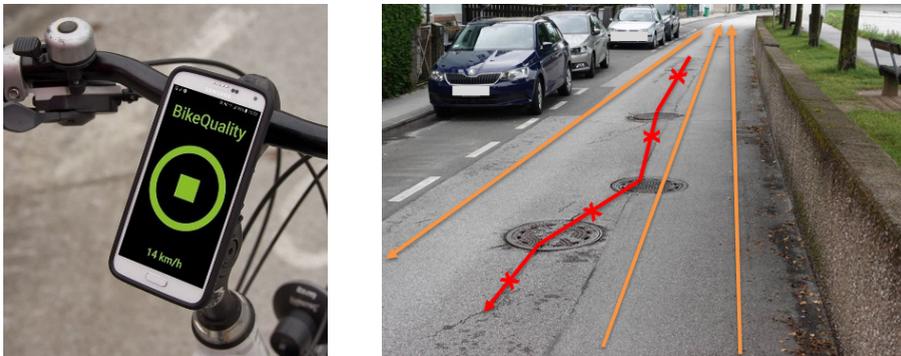
Weiterhin wird ein Map-Matching-Algorithmus von Rehl et al. (2018) verwendet, welcher auf einer Arbeit von Sauerwein (2013) beruht. Laut Rehl et al. (2018) verwendet der Map-Matching-Algorithmus mehrere geometrische und topologische Maße um Pfadkandidaten auszuwählen. Diese Pfadkandidaten werden durch Fahrtrajektorien repräsentiert, die sich aus zeitlich geordneten Trackpunkten zusammensetzen. Diese werden wiederum als Kombination einer WGS84-Koordinate und einem Zeitstempel dargestellt. Weiterhin wird laut Rehl et al. (2018) das digitale Straßennetz als ein Graph aus geographisch referenzierten Straßensegmenten modelliert. Diese sind topologisch so verbunden, dass dadurch mögliche Bewegungsoptionen innerhalb des Straßennetzes dargestellt werden. Am Ende des Map-Matching Verfahrens wird ein Pfad ausgegeben, welchem die jeweilige Entität höchstwahrscheinlich gefolgt ist (Rehl et al., 2018).

Durch die Anwendung des Map-Matching-Algorithmus werden die GNSS-Positionen dem digitalen Radnetzgraphen zugeordnet. Infolge ist es möglich, dass jede Beschleunigungsmessung inklusive Fahrtrichtung einem linearen Referenzpunkt auf dem Radnetzgraphen zugeteilt werden kann (Brunauer & Rehl, 2016). Anschließend kann der Radnetzgraph in beliebig lange Abschnitte unterteilt werden. In der diskreten Fahrzeugsimulation werden für die Analyse Zellen von  $7,5 \cdot 7,5$  m definiert, da davon ausgegangen wird, dass dieser Raum von einem Fahrzeug eingenommen wird (Nagel, 1998). In der Literatur wird diese Einteilung jedoch als zu grob angesehen, um eine Fahrzeugbewegung unter realen Bedingungen simulieren zu können (Bham & Benekohal, 2004; Lan & Chang, 2005). Die Einteilung in 7,5 Abschnitte ist zudem stark von Argumenten getrieben, die für die Fahrzeugsimulation sinnvoll sind (Nagel, 1998), jedoch für den in dieser Studie analysierten Anwendungsfall nicht umgelegt werden können. Daher wurde in dieser Studie die zur Analyse verwendete Abschnittslänge, durch eine Expert\*inneneinschätzung, empirisch festgelegt. Hierfür wurden mehrere Testsegmente mit starken Fahrbahnschäden und gut befahrbaren Fahrbahnoberflächen gewählt. Diese Segmente stellen einen Straßenabschnitt zwischen zwei Abbiegemöglichkeiten dar und wurden unter der Berücksichtigung der Fahrtrichtung in jeweils 5, 10 und 20 m lange Abschnitte segmentiert. Ein Vergleich der durchschnittlichen Beschleunigung zwischen diesen drei Abschnittslängen hat ergeben, dass bei einer Länge von 10 m der Zustand der Fahrbahnoberfläche am besten abgebildet werden konnte. Folgend wird für die weitere Untersuchung eine Abschnittslänge von 10 m herangezogen. Hierbei ist anzumerken, dass auch kürzere bzw. längere Abschnitte für das Clustering-Verfahren verwendet werden

könnten. Bei längeren Abschnitten ist zu erwarten, dass lokal auftretende Unebenheiten nicht so deutlich sichtbar werden, wenn diese an Abschnitte mit einer guten Fahrbahnoberfläche angrenzen, da starke Ausschläge im Durchschnitt an Gewicht verlieren würden. Bei kürzeren Abschnitten würden GNSS-Lokalisierungsfehler an Bedeutung zunehmen, da diese bei einer höheren Anzahl an Abschnitten öfter auftreten können.

Somit werden für jeden 10 m Abschnitt des Untersuchungsnetzes Features generiert, welche als Input für den Clusteralgorithmus verwendet werden. Hierbei ist anzumerken, dass es sich bei diesen 10 m Abschnitten um statische Abschnitte des Straßennetzes handelt. Daher ist sichergestellt, dass es sich auch bei unterschiedlichen Tracks immer um die gleichen 10 m Abschnitte handelt, welche für die Analyse herangezogen werden. Zudem ist darauf hinzuweisen, dass aufgrund der individuellen Länge eines Segmentes nicht jeder Abschnitt genau 10 m beträgt. Für Abschnitte mit weniger als 10 m, wurden die erfassten Beschleunigungsdaten auf die Länge von 10 m hochgerechnet, um eine Vergleichbarkeit zu gewährleisten. Hierbei ist anzumerken, dass sich die kürzeren Abschnitte immer am Ende eines Segmentes befinden, an welchem der/die Fahrradfahrer\*in dieses wieder verlässt.

Damit einheitliche Fahrbedingungen ermöglicht werden, wurden weitere einheitliche Vorgaben zur Datenaufzeichnung festgelegt: Für die Datenaufzeichnung werden baugleiche Smartphones und Fahrräder verwendet. Die Fahrräder müssen eine starre Gabel besitzen und der Reifendruck wurde abhängig vom verwendeten Reifen definiert. Für die Aufzeichnung wurde das Smartphone auf allen Fahrrädern an der gleichen Position befestigt (s. Abb. 1). Den Proband\*innen wurde in einer Einschulung die Vorgaben zur Fahrweise vermittelt. Weiterhin wurde den Proband\*innen eine Fahrgeschwindigkeit zwischen 15 und 20 km/h vorgegeben. Wie in Abbildung 1 dargestellt, wurde eine geradlinige Fahrweise vorgeschrieben ohne dass Fahrbahnbuckel, Schlaglöcher oder Kanaldeckel gezielt angefahren werden mussten. Somit sollte eine gleichmäßige Verteilung der Befahrungen und infolge eine einheitliche Qualitätsbewertung ermöglicht werden.



**Abb. 1:** Vorgaben zur Datenerfassung: Montage des Smartphones am Fahrradlenker (links) und vorgegebene Fahrweise (rechts)

### 3.2 Clusterverfahren

Als Clusterverfahren wurde der k-means++-Algorithmus verwendet, welcher von Arthur & Vassilvitskii (2006) eingeführt wurde. Dieser Algorithmus erweitert den weit verbreiteten k-means-Algorithmus von Lloyd (1982) durch eine randomisierte Seeding-Technik. Der von Lloyd (1982) entwickelte k-means-Algorithmus wählt zuerst zufällig  $k$  Zentren aus. Die zu clusternden Datenpunkte werden anschließend dem sich am nächsten befindlichen Zentrum zugewiesen. Danach wird das Zentrum neu berechnet und der Vorgang wiederholt sich bis sich der Prozess stabilisiert (Arthur & Vassilvitskii, 2006; Lloyd, 1982). Der k-means++-Algorithmus von Arthur & Vassilvitskii (2006) modifiziert den Initialisierungsprozess zur Auswahl der Clusterzentren. In diesem wird für jeden Datenpunkt der Abstand zum nächstgelegenen Clusterzentrum quadriert  $D(x)^2$ . Anschließend werden zufällig neue Clusterzentren ausgewählt, wobei die Wahrscheinlichkeit, dass ein Datenpunkt als neues Clusterzentrum ausgewählt wird, proportional zu  $D(x)^2$  definiert wird. Wenn durch die Wiederholung dieses Vorgangs  $k$  Clusterzentren festgelegt wurden, wird der k-means-Algorithmus von Lloyd (1982) verwendet (Arthur & Vassilvitskii, 2006). Laut Ergebnissen von Arthur & Vassilvitskii (2006) liefert der k-means++-Algorithmus durch diesen modifizierten Initialisierungsprozess oft schnellere und genauere Ergebnisse als der ursprüngliche k-means-Algorithmus. Als Nachteil der Verwendung des k-means++-Algorithmus soll erwähnt werden, dass dieser sehr sensitiv auf die Auswahl des ersten Zentrums reagiert (Öztürk, Cavusoglu, & Zengin, 2015). Weiterhin argumentieren Bachem et al. (2016), dass der k-means++-Algorithmus nicht auf große Datensätze skaliert werden kann. Dies ergibt sich daraus, dass bei jedem Seeding-Schritt und jeder Iteration die Berechnung der Abstände zwischen den Clusterzentren und den Datenpunkten neu durchgeführt werden muss (Bachem et al., 2016).

Zusammenfassend handelt es sich beim k-means++-Algorithmus, um ein einfach anzuwendendes und schnelles Clustering-Verfahren mit einer starken empirischen Performance (Arthur & Vassilvitskii, 2006; Bachem et al., 2016). Aufgrund seiner einfachen Implementierbarkeit treten geringe methodische Barrieren für eine breite Anwendung zur Bewertung von Fahrradinfrastruktur auf. Weiterhin kann das Verfahren leicht für neue und erweiterte Datensätze angepasst werden, was einen flexiblen Einsatz ermöglicht. Zudem können im Vergleich zur Verwendung des IRI vielfältige Features für die Qualitätsbewertung der Fahrbahnabschnitte verwendet werden. Daher stellt die Verwendung des k-means++-Algorithmus eine interessante Alternative zu herkömmlichen Methoden wie dem IRI dar, um Fahrbahnoberflächen bewerten zu können.

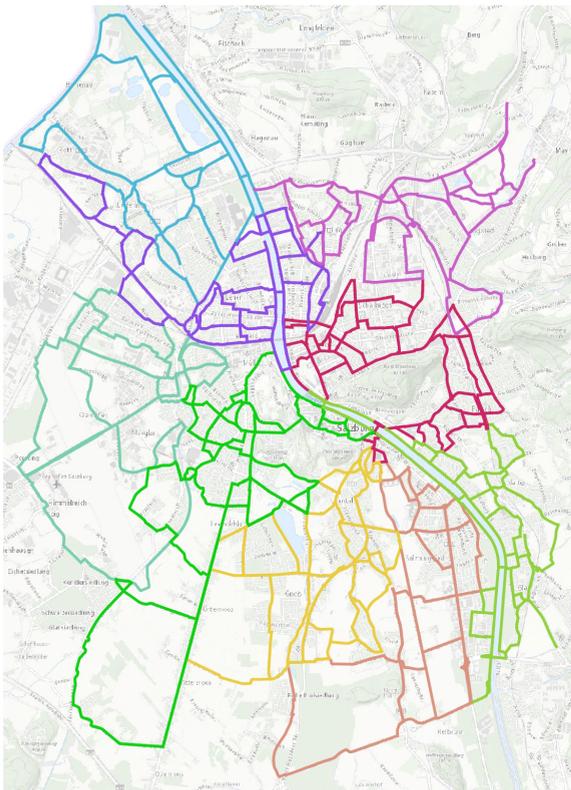
Für diese Studie werden, wie zuvor beschrieben, pro 10 m Abschnitt eines Segments Features berechnet. Nach Berechnung der Features werden Feature-Kombinationen bestimmt, welche anschließend mithilfe des k-means++-Algorithmus für die Clusterbildung verwendet wurden. Als Output wird jeder 10-m-Abschnitt einem Cluster zugeordnet. Diese Cluster werden weiters in Qualitätsstufen eingeteilt, womit eine Bewertung der Fahrbahnoberfläche möglich wird. Anschließend wird untersucht, welche Feature-Kombination gute bzw. schlechte Straßenabschnitte jeweils einem eigenen Cluster zugeordnet haben. Somit kann infolge die am besten geeignete Feature-Kombination für die Auswertung bestimmt werden. Um die Qualität des Outputs zu testen, wurde die Fahrbahnqualität von ausgewählten Straßenabschnitten, wie in Abbildung 2 dargestellt, durch eine Expert\*inneneinschätzung manuell bestimmt und mit den Ergebnissen des k-means++-Algorithmus verglichen. Die Feature-Kombination mit der höchsten Übereinstimmung zwischen Expert\*inneneinschätzung und Clusterverfahren wird für die weitere Analyse verwendet.



**Abb. 2:** Gute Fahrbahnoberflächenqualität (links) und (schlechte Fahrbahnoberflächenqualität (rechts)

## 4 Fallstudie: Stadt Salzburg

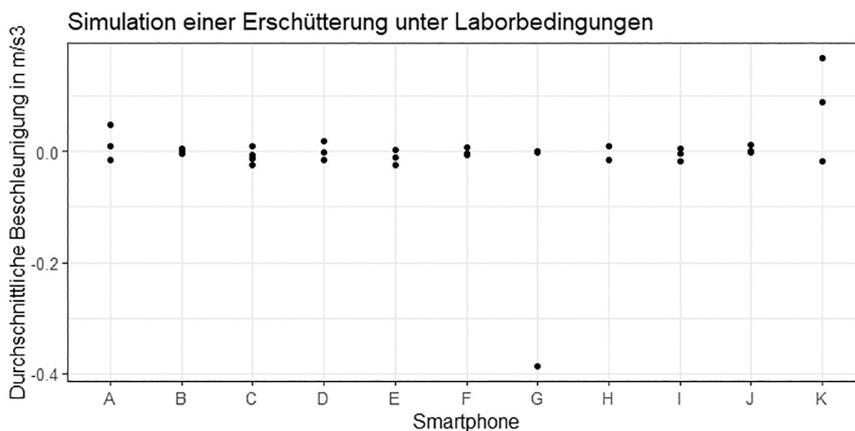
Die Datenaufzeichnung wurde im Zeitraum zwischen dem 18.05. – 05.07.2020 in der Stadt Salzburg entlang des Hauptradwegenetzes durchgeführt. Um eine gezielte Datenerfassung zu ermöglichen, wurde das Untersuchungsnetz in 9 Befahrungsrouten eingeteilt (siehe Abb. 3).



**Abb. 3:** Untersuchungsnetz in der Stadt Salzburg

Das Hauptradwegenetz musste in beiden Fahrtrichtungen befahren werden, somit ergab sich eine gesamte Befahrungslänge von 436,6 km. Jede Befahrungsrouten musste in beiden Richtungen 10 Mal befahren werden. Die Datenaufzeichnung wurde von 10 Proband\*innen durchgeführt. Für die Datenaufzeichnung wurden den Proband\*innen baugleiche Smartphones „Samsung S5neo“ und Fahrräder der Marke „Specialized Alibi Sport EQ“ mit starrer Gabel und Vollgummireifen zur Verfügung gestellt.

Um sicher zu stellen, dass die verwendeten Smartphones auch ähnliche Beschleunigungswerte bei gleichen Erschütterungen ausgeben, wurden Stöße unter Laborbedingungen simuliert. Hierfür wurden 11 Smartphones auf einer Platte abgelegt. Auf diese Platte wurde in mehreren Durchläufen ein Gewicht von einer standardisierten Höhe fallen gelassen. Dadurch sollte eine Erschütterung durch eine Fahrbahnebenheit simuliert werden. Die Beschleunigungswerte wurden aufgezeichnet und in Abbildung 4 dargestellt. Abbildung 4 zeigt, dass ähnliche Mittelwerte in der Beschleunigung zwischen den Smartphones erkennbar werden. Die Punkte veranschaulichen die durchschnittliche Beschleunigung pro Versuchsdurchgang für das jeweilige Smartphone. Hier sind bei den Smartphones G und K Ausreißer zu erkennen, welche aufgrund ihrer Größe vernachlässigt werden können. Dies lässt darauf schließen, dass die verwendeten Smartphones nicht gegeneinander abgeglichen werden müssen.



**Abb. 4:** Simulation einer Erschütterung unter Laborbedingungen

Nach Bereinigung der GNSS-Ausreißer können von den 337 Fahrradtracks noch insgesamt 5.166 der 5.642 km für die weitere Analyse verwendet werden. Diese wurden anschließend auf das Untersuchungsnetz von 436,6 km übertragen. Es wurde festgelegt, dass alle Befahrungen mit einer Geschwindigkeit von unter 9 km/h gefiltert werden. Bei diesen Befahrungen wird davon ausgegangen, dass es während der Fahrt zu einem Stopp gekommen ist und diese Daten daher nicht für die Analyse verwendet werden können. Nach dieser Filterung wurde eine Fehlermaßzahl berechnet, die den Anteil der Strecke angibt, welcher mit einer Durchschnittsgeschwindigkeit von weniger als 15 und mehr als 20 km/h befahren wurde. Diese zeigt, dass rund 47,9 % der gesamten Strecke nicht im vorgegebenen Geschwindigkeitskorridor zwischen 15 und 20 km/h befahren wurde. Dies war aufgrund von Steigungen und abfallenden Straßenverläufen jedoch zu erwarten. Erweitert man die Geschwindigkeitsgrenzen nur geringfügig um jeweils 3 km/h (12-23 km/h) sind nur rund 12,2 % der befahrenen Strecke

außerhalb des Korridors befahren worden. Bei einem Korridor von 10-25 km/h weichen nur rund 4 % der befahrenen Strecke ab.

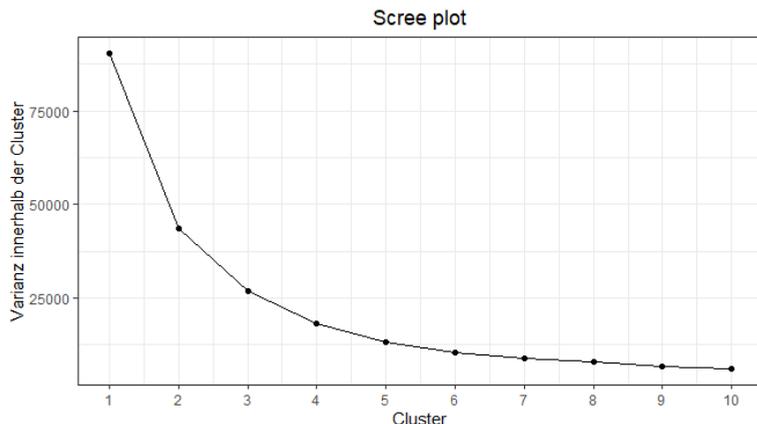
Bevor mit der Analyse des Untersuchungsnetzes begonnen wird, wurde eine Auswahl der zu verwendeten Features durchgeführt. Für die Klassifizierung wurden in einem ersten Schritt die Features in Tabelle 1 pro Fahrbahnabschnitt für jede einzelne Befahrung berechnet. Anschließend wird der Durchschnitt des jeweiligen Features über alle Befahrungen eines Fahrbahnabschnitts ermittelt. Hierbei ist anzumerken, dass ein Signal mit einer Beschleunigung größer gleich  $1.200 \text{ m/s}^3$  als starker Schlag klassifiziert wurde. Dieser Wert ist durch die Analyse von gezielten Befahrungen von Fahrbahnschäden durch eine Expert\*inneneinschätzung festgelegt worden.

**Tabelle 1:** Verwendete Features

#	Features
1	Durchschnittliche Beschleunigung
2	Mediane Beschleunigung
3	Durchschnittliche Standardabweichung
4	Minimale Beschleunigung
5	Maximale Beschleunigung
6	Durchschnittliche Anzahl an starken Schlägen
7	Durchschnittliche Beschleunigung der starken Schläge

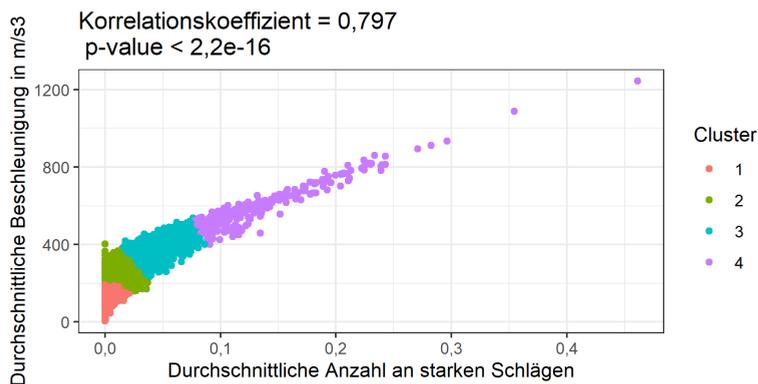
Für die Auswahl der finalen Kalibrierung des k-means++-Algorithmus wurden unterschiedliche Feature-Kombinationen, mit einem  $k$  von 3, 4 und 5 definiert. Die Analyse hat gezeigt, dass sich mithilfe der (1) *durchschnittlichen Anzahl an starken Schlägen* sowie der (2) *durchschnittlichen Beschleunigung über alle Schläge pro Abschnitt* die höchste Übereinstimmung zwischen den Ergebnissen des Clusterverfahrens und der Expert\*inneneinschätzung ergeben hat. Hierbei ist anzumerken, dass bewusst keine Unterscheidung zwischen Straßenbelägen bei der Bewertung durch die Expert\*innen vorgenommen wurde. Dies führt dazu, dass unter anderem Kopfsteinpflaster, aufgrund der starken Vibration durch den k-means++-Algorithmus und der Einschätzung der Expert\*innen, in die schlechteste Qualitätsstufe fallen und als sanierungsbedürftig eingestuft werden, obwohl hier kein unmittelbarer Sanierungsbedarf besteht. Trotzdem ist es laut unserer Sicht notwendig, Kopfsteinpflaster in der Analyse zu berücksichtigen, da diese für Fahrradfahrer\*innen ein unangenehmes Fahrgefühl durch starke Vibrationen auslösen.

Anschließend wurde die Anzahl der Cluster  $k$  mithilfe eines *scree plots* für die finale Auswertung bestimmt. Abbildung 5 zeigt auf der x-Achse die Anzahl der verwendeten Cluster und auf der y-Achse die Varianz innerhalb des jeweiligen Clusters. Man kann erkennen, dass bei einer Anzahl von 4 Clustern die Varianz nur mehr geringfügig abnimmt. Das Setting des k-means++-Algorithmus lautet daher wie folgt:  $k = 4$ ; Features (1) durchschnittliche Anzahl an starken Schlägen sowie (2) durchschnittliche Beschleunigung über alle Schläge pro Abschnitt.



**Abb. 5:** Scree plot zur Bestimmung der geeigneten Clusteranzahl

Abbildung 6 zeigt die Verteilung der Cluster, welche sich durch die Anwendung des k-means++-Algorithmus ergeben. Auf der x-Achse ist die *durchschnittliche Anzahl an starken Schlägen* und auf der y-Achse die *durchschnittliche Beschleunigung über alle Schläge* eines Fahrbahnabschnittes ersichtlich. Zwischen beiden Variablen ist mit einem Korrelationskoeffizienten von 0,79 eine deutliche signifikante Korrelation erkennbar.



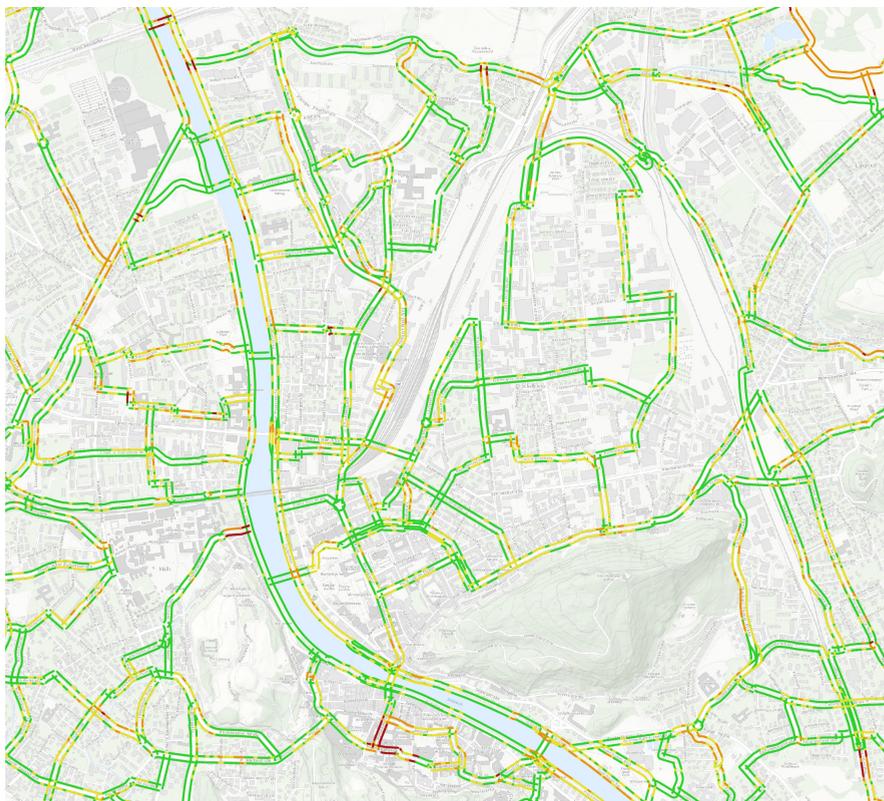
**Abb. 6:** Verteilung der Cluster

Aus der Visualisierung in Abbildung 6 kann abgeleitet werden, dass in Cluster 1 jene Abschnitte mit der besten und in Cluster 4 jene mit der schlechtesten Fahrbahnoberfläche zugeordnet wurden. Die absolute und relative Verteilung der Clustergruppen in Bezug auf Fahrbahnabschnitte und Kilometer ist in Tabelle 2 ersichtlich. Hier ergeben sich leichte Unterschiede, da aufgrund der Länge eines Straßensegments nicht jeder Fahrbahnabschnitt in genau 10 m unterteilt werden konnte. Die Clustergruppen wurden durch eine Expert\*inneneinschätzung auch mit qualitativen Informationen zu ihrer Befahrbarkeit ergänzt. Der Großteil der Fahrbahnabschnitte wird mit 58,4 % dem ersten und mit 31,7 % dem zweiten Cluster zugeordnet. Diese Clustergruppen wurden durch eine Expert\*inneneinschätzung als komfortabel bzw. fahrtauglich klassifiziert. Die restlichen Abschnitte verteilen sich mit 9,1 % auf

den dritten und mit 0,7 % auf den vierten Cluster. Diese beiden Clustergruppen stellen laut den Ergebnissen der verwendeten Methode, Abschnitte mit schlechter bzw. sehr schlechter Befahrbarkeit dar und sollten in weiterer Folge gezielt durch Wartungsarbeiten instandgehalten werden. Zudem wird ersichtlich, dass aufgrund der Datenbereinigung für ca. 8,4 km des Untersuchungsnetzes keine Clusterzuteilung möglich ist.

**Tabelle 2:** Verteilung der Clustergruppen

Cluster	Befahrbarkeit	Anzahl der Fahr- bahnabschnitte	Relativer Anteil der Fahrbahn- abschnitt in %	km	Relativer Anteil der km in %
1	Komfortabel	26.520	58,4	250,2	57,3
2	Fahrtauglich	14.410	31,7	135,9	31,1
3	Schlecht	4.150	9,1	39,1	9,0
4	Sehr schlecht	323	0,7	3,0	0,7
	Nicht ausgewertet			8,4	1,9
	<b>Gesamt</b>	<b>45.403</b>		<b>436,6</b>	



**Abb. 7:** Exemplarische visuelle Darstellung der Oberflächenqualität. Einteilung der Oberflächenqualität nach Farbe: ■ Komfortable, ■ Fahrtaugliche, ■ Schlechte und ■ Sehr schlechte Oberflächenqualität

Abbildung 7 zeigt eine visuelle Darstellung der Ergebnisse auf einem ausgewählten Kartenabschnitt. Die Oberflächenqualität der Fahrbahnabschnitte wird je nach Ergebnis des Clusteringverfahrens farblich dargestellt. Grüne und gelbe Abschnitte zeigen komfortable bzw. fahrtaugliche Fahrbahnabschnitte, wohingegen die orangen und roten Abschnitte schlechte bzw. sehr schlechte Fahrbahnabschnitte anzeigen. So können lokale hot-spots mit erhöhtem Wartungsbedarf gut visualisiert und lokalisiert werden, was für die Instandhaltungsarbeiten von großer Bedeutung ist. Die Ergebnisse zeigen somit, dass anhand der Anwendung des  $k$ -means++-Algorithmus, schlecht befahrbare Abschnitte der Fahrradinfrastruktur im Untersuchungsnetz der Stadt Salzburg detektiert werden können.

## 5 Diskussion

Diese Arbeit zeigt anhand einer Fallstudie für das Hauptradwegenetz der Stadt Salzburg ein leicht umsetzbares Verfahren zur Bestimmung der Fahrbahnoberflächenqualität einer Fahrradinfrastruktur. Anhand von Daten aus Smartphone-Beschleunigungssensoren, wurden mithilfe eines  $k$ -means++-Clusteralgorithmus schlecht befahrbare Fahrbahnabschnitte lokalisiert. Dies ermöglicht eine gezielte Instandhaltung und Verbesserung des Radnetzes und kann den Fahrkomfort und die Sicherheit für Fahrradfahrer\*innen erhöhen. Die beschriebene Vorgehensweise zur Bestimmung der Fahrbahnoberflächenqualität, kann auf jedes beliebige Radnetz umgelegt werden und stellt eine einfache und kostengünstige Methode dar, um die Attraktivität einer Fahrradinfrastruktur zu erhöhen. Als zukünftige Forschungsthemen könnten weitere Clusteralgorithmen untersucht und miteinander verglichen werden. Zudem ist anzumerken, dass die Länge der zur Analyse verwendeten Abschnitte als Modellparameter zu interpretieren ist und daher für neue Testgebiete erneut validiert werden sollte. Ein weiteres Forschungsthema würde somit die Messung des Einflusses der Länge der analysierten Abschnitte auf die Ergebnisse des Clusterverfahrens darstellen. Weiterhin könnten mithilfe von manuell gelabelten Abschnitten überwachte Lernverfahren verwendet werden, um schlecht befahrbare Abschnitte zu klassifizieren. Zudem könnte die in dieser Studie verwendete Methodologie auf weitere Radnetze angewendet werden, damit mehr Erkenntnisse auf ihre Funktionalität erzielt werden können. Eine mögliche Erweiterung würde zudem die Verwendung von zusätzlichen Features darstellen.

## Literatur

- Arthur, D., & Vassilvitskii, S. (2006). *k*-means++: *The advantages of careful seeding*. Stanford.
- Bachem, O., Lucic, M., Hassani, S. H., & Krause, A. (2016). Fast and Provably Good Seedings for  $k$ -Means. *Nips* (pp. 55–63).
- Bham, G. H., & Benekohal, R. F. (2004). A high fidelity traffic simulation model based on cellular automata and car-following concepts. *Transportation Research Part C: Emerging Technologies*, 12(1), 1–32.
- Brunauer, R., & Rehr, K. (2016). Supporting Road Maintenance with In-Vehicle Data: Results from a Field Trial on Road Surface Condition Monitoring. In *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*.  
<https://doi.org/10.1109/ITSC.2016.7795917>.

- Brunauer, R., Rehl, K., Kremser, W., & Moser, R. (2018). CarSense: Evaluation of the Potential of In-Vehicle Sensor Data for Road Operators. In *Proceedings of 7th Transport Research Arena TRA 2018*, April 16-1, Vienna, Austria.
- Douangphachanh, V., & Oneyama, H. (2013). A Study on the Use of Smartphones for Road Roughness Condition Estimation. In: *Proceedings of the Eastern Asia Society for Transportation Studies* (Vol. 9).
- Harikrishnan, P. M., & Gopi, V. P. (2017). Vehicle vibration signal processing for road surface monitoring. *IEEE Sensors Journal*, 17(16), 5192–5197.
- Lan, L. W., & Chang, C.-W. (2005). Inhomogeneous cellular automata modeling for mixed traffic with cars and motorcycles. *Journal of Advanced Transportation*, 39(3), 323–349.
- Li, X., & Goldberg, D. W. (2018). Toward a mobile crowdsensing system for road surface assessment. *Computers, Environment and Urban Systems*, 69, 51–62.
- Litzenberger, S., Christensen, T., Hofstätter, O., & Sabo, A. (2018). Prediction of road surface quality during cycling using smartphone accelerometer data. In *Multidisciplinary Digital Publishing Institute Proceedings* (Vol. 2, p. 217).
- Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2), 129–137.
- Nagel, K. (1998). From particle hopping models to traffic flow theory. *Transportation Research Record*, 1644(1), 1–9.
- Open Telekom Cloud. (n. d.). Retrieved from <https://open-telekom-cloud.com/de>.
- Öztürk, M. M., Cavusoglu, U., & Zengin, A. (2015). A novel defect prediction method for web pages using k-means++. *Expert Systems with Applications*, 42(19), 6496–6506.
- Rehl, K., Gröchenig, S., & Wimmer, M. (2018). Optimization and Evaluation of a High Performance Open-Source Map-Matching Implementation. In *Geospatial Technologies for All. Selected Papers of the 21th AGILE Conference on Geographic Information Science*. Lund, Sweden.
- Sadjadi, S. Y. (2017). Investigating for Road Roughness using Smartphone Sensors. *International Journal of Computers and Communication*, 11, 56–63.
- Sattar, S., Li, S., & Chapman, M. (2018). Road surface monitoring using smartphone sensors: A review. *Sensors*, 18(11), 3845.
- Sauerwein, T. (2013). *Optimization and evaluation of an Online Map-Matching Algorithm for Mid-Range Sampling Rates*. Universität Marburg.
- Sayers, M. W. (1995). On the calculation of international roughness index from longitudinal road profile. *Transportation Research Record*, (1501).
- Smith, K., & Ram, P. (2016). *Measures and Specifying Pavement Smoothness*. FHWA: Washington, DC, USA.
- Vittorio, A., Rosolino, V., Teresa, I., Vittoria, C. M., Vincenzo, P. G. et al. (2014). Automated sensing system for monitoring of road surface quality by mobile devices. *Procedia-Social and Behavioral Sciences*, 111, 242–251.
- Wijerathne, N., Viswanath, S. K., Hasala, M. S., Beltran, V., Yuen, C., & Lim, H. B. (2018). Towards comfortable cycling: A practical approach to monitor the conditions in cycling paths. In *2018 IEEE 4th World Forum on Internet of Things (WF-IoT)* (pp. 778–783).
- Zang, K., Shen, J., Huang, H., Wan, M., & Shi, J. (2018). Assessing and mapping of road surface roughness based on GPS and accelerometer sensors on bicycle-mounted smartphones. *Sensors*, 18(3), 914.