

## Kurzbericht

für das im Rahmen des Förderprogramms „Ressourceneffiziente Abwasserbeseitigung NRW, Förderbereich 6: Forschungs- und Entwicklungsprojekte zur Abwasserbeseitigung“ geförderte F&E-Vorhaben:

# Wasserwirtschaftliche Ermittlungen und Anwendung von digitalen Versiegelungsdaten aus Amtlichen Liegenschaftskatasterdaten

## WaLKIS

Landesamt für Natur,  
Umwelt und Verbraucherschutz  
Nordrhein-Westfalen





Abschlussbericht zum Projekt

**Wasserwirtschaftliche Ermittlungen und Anwendung von digitalen Versiegelungsdaten aus  
Amtlichen Liegenschaftskatasterdaten (WaLKIS)**

Gefördert von

Landesamt für Natur,  
Umwelt und Verbraucherschutz  
Nordrhein-Westfalen



Aachen, April 2018  
FiW an der RWTH Aachen

Dr.-Ing. F.-W.- Bolle



## **Projektbearbeitung**

### **Institution**

Forschungsinstitut für Wasser- und Abfallwirtschaft  
an der RWTH Aachen (FiW) e.V.  
Kackertstraße 15-17  
52056 Aachen

### **Bearbeiter**

Dr.-Ing. Gesa Kutschera

Sajjad Tabatabaei, M.Sc.

Leon Indjike, M.Sc.

## Inhaltsverzeichnis

|        |                                                                                         |    |
|--------|-----------------------------------------------------------------------------------------|----|
| 1      | Einleitung .....                                                                        | 1  |
| 2      | Versiegelte Flächen .....                                                               | 2  |
| 3      | Datengrundlage .....                                                                    | 4  |
| 3.1    | ALKIS.....                                                                              | 4  |
| 3.2    | DGM .....                                                                               | 4  |
| 3.3    | Bevölkerungskataster .....                                                              | 4  |
| 3.4    | Referenzdaten von Emschergenossenschaft und Lippeverband.....                           | 5  |
| 4      | Methodik .....                                                                          | 8  |
| 4.1    | Vorbemerkung .....                                                                      | 8  |
| 4.2    | Bestimmung der Gebäude- und öffentlichen Verkehrsflächen.....                           | 8  |
| 4.3    | Bestimmung von Zuwegungsflächen .....                                                   | 8  |
| 5      | Ermittlung der Eingangsdaten für die statistische Abschätzung der Zuwegungsflächen..... | 10 |
| 5.1    | Erzeugung der Erklärungsvariablen .....                                                 | 10 |
| 5.2    | Referenzdaten .....                                                                     | 11 |
| 6      | Untersuchungsergebnisse.....                                                            | 12 |
| 6.1    | Gebäudeflächen aus ALKIS .....                                                          | 12 |
| 6.2    | Verkehrsflächen aus ALKIS.....                                                          | 20 |
| 6.3    | Abgeleitete Zuwegungsflächen aus ALKIS .....                                            | 25 |
| 6.3.1  | Klassifizierungsmodell .....                                                            | 25 |
| 6.3.2  | Clusteranalyse und Regressionsmodelle .....                                             | 28 |
| 6.3.3  | Validierungsergebnisse .....                                                            | 31 |
| 7      | Zusammenfassung .....                                                                   | 34 |
| 8      | Ausblick .....                                                                          | 36 |
| 9      | Literaturverzeichnis .....                                                              | 38 |
| Anhang | .....                                                                                   | 40 |

## Abbildungsverzeichnis

|                                                                                                                                                                                                                               |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Abbildung 1: Neuversiegelung und Anteil der versiegelten Flächen in der BRD .....                                                                                                                                             | 2  |
| Abbildung 2: Verbandsgebiet der Emschergenossenschaft und Lippeverband .....                                                                                                                                                  | 5  |
| Abbildung 3: Stand der Aktualisierung der EGLV-Daten im Untersuchungsgebiet der<br>Emschergenossenschaft und Lippeverband .....                                                                                               | 7  |
| Abbildung 4: Anteil der unterschiedlichen versiegelten Flächen im Untersuchungsgebiet der EGLV....                                                                                                                            | 7  |
| Abbildung 5: Übersicht der Vorgehensweise zur Abschätzung der Zuwegungsflächen.....                                                                                                                                           | 9  |
| Abbildung 6: Gegenüberstellung der summierten Gebäudeflächen auf einzelne Flurstücke aus ALKIS<br>und EGLV, links: Stärke des linearen Zusammenhangs; rechts: Grad der Übereinstimmung in<br>Prozent.....                     | 14 |
| Abbildung 7: Prozentuales Defizit der durch ALKIS erfassten Gebäudeflächen im Vergleich zu EGLV -<br>Gebäudeflächen .....                                                                                                     | 16 |
| Abbildung 8: Verteilung der Abweichung zwischen korrigierten ALKIS-Gebäudeflächen und EGLV-<br>Gebäudeflächen (links: Methode 1, rechts: Methode 2).....                                                                      | 17 |
| Abbildung 9: Verlauf der Abweichungen zwischen EGLV- und ALKIS Gebäudepolygonen bei<br>unterschiedliche Korrekturansätzen .....                                                                                               | 18 |
| Abbildung 10: Zusammenhang zwischen der Gebäudegröße und die Größe der Abweichungen zu<br>EGLV-Gebäudepolygone.....                                                                                                           | 18 |
| Abbildung 11: Vergleich der Übereinstimmungen von ALKIS Gebäudeflächen und korrigierten ALKIS<br>Gebäudeflächen mit EGLV Gebäudeflächen .....                                                                                 | 19 |
| Abbildung 12: Übereinstimmung der Flächensummen der korrigierten ALKIS-Gebäudeflächen mit<br>EGLV-Gebäudeflächen in den Kommunen von Emschergenossenschaft und Lippeverband.                                                  | 20 |
| Abbildung 13: Übereinstimmung der Objektart "Straßenverkehr" mit EGLV Daten bezogen auf<br>Kommunen im Einzugsgebiet der Emschergenossenschaft und Lippeverband .....                                                         | 21 |
| Abbildung 14: Erfassung von befestigten Verkehrsflächen durch EGLV und ALKIS.....                                                                                                                                             | 22 |
| Abbildung 15: Flächenüberschuss bzw. Flächendefizit mit der Objektart „Straßenverkehr“ (rot) und<br>erweiterter Auswahl (schwarz) im Vergleich zu EGLV – Verkehrsflächen.....                                                 | 23 |
| Abbildung 16: Erfassung von Verkehrsflächen durch ALKIS und EGLV (Stand 2008) in Bad<br>Sassendorf (Stand 2017).....                                                                                                          | 24 |
| Abbildung 17: Erfassungsgenauigkeit der Verkehrsflächen durch ALKIS - Daten (Stand Dortmund<br>2018, Dorsten 2018, Recklinghausen 2018, Essen 2015) (EGLV Stand: Dortmund,<br>Recklinghausen, Essen 2014, Dorsten 2010) ..... | 25 |
| Abbildung 18: Fehlerschätzung für den Generalisierungsfehler .....                                                                                                                                                            | 26 |

|                                                                                                                                                                 |    |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Abbildung 19: Ranking der Gewichtung der Regressoren .....                                                                                                      | 27 |
| Abbildung 20: Einzugsgebiet des Hüller Bachs im Untersuchungsgebiet der EGLV .....                                                                              | 28 |
| Abbildung 21: Zuordnung von Flurstücken zu Clustern mit unterschiedlicher Güte der<br>Regressionsmodelle .....                                                  | 30 |
| Abbildung 22: Ausgewählte Teileinzugsgebiete mit Abweichungen unter und über 20 % zwischen<br>EGLV-Zuwegungsflächen und prognostizierten Zuwegungsflächen ..... | 31 |
| Abbildung 23: Digitalisierte Flurstücke in der Städtereion für die Validierung der Regressionsmodelle<br>(Stand 2016) .....                                     | 32 |
| Abbildung 24: Prozentuale Übereinstimmung der Prognosewerte für Zuwegungsflächen mit manuell<br>digitalisierten Zuwegungsflächen aus der Städtereion.....       | 33 |

## Tabellenverzeichnis

|                                                                                                                                                                                |    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Tabelle 1: Unterteilung der befestigten Flächen anhand der „Leistungsbeschreibung für die ArcGIS-konforme Erfassung der befestigten Flächen“ der EGLV .....                    | 6  |
| Tabelle 2: Ableitung der unabhängigen Variablen (Regressoren) aus ALKIS .....                                                                                                  | 10 |
| Tabelle 3: Auswahl von Objektarten zur Bestimmung von Gebäudeflächen .....                                                                                                     | 12 |
| Tabelle 4: Zusammensetzung unterschiedlicher Fälle bei der Gegenüberstellung von Gebäudepolygonen aus ALKIS und EGLV auf einzelne Flurstücke im Untersuchungsgebiet EGLV ..... | 13 |
| Tabelle 5: Die Übereinstimmung der summierten Gebäudeflächen auf Flurstücke getrennt nach Siedlungen .....                                                                     | 15 |
| Tabelle 6: Auswahl an ALKIS – Objektarten, die Verkehrsflächen abbilden .....                                                                                                  | 22 |
| Tabelle 7: Prognose von Zuwegungen auf Flurstücke mit der Methode "Entscheidungsbäume" .....                                                                                   | 27 |
| Tabelle 8: Ergebnisse der statistischen Abschätzung von Zuwegungsflächen und durchschnittliche Übereinstimmungen mit Kalibrierungsdaten.....                                   | 29 |

## **Abkürzungsverzeichnis**

DWA-A Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall Arbeitsblatt  
EG/LV Emschergenossenschaft/Lippeverband  
GDI Geodateninformationsstruktur  
OS Objektschlüssel



## 1 Einleitung

Daten zu versiegelten Flächen sind aufgrund der Folgewirkungen für den Wasserhaushalt und in diesem Zusammenhang einer geringeren Grundwasserneubildung, eines größeren Oberflächenabflusses und sinkender Evapotranspiration eine relevante Basisinformation für diverse wasserwirtschaftliche Aufgabenstellungen (BLUME, 1992). Trotz zahlreicher integraler Funktionen existiert für das Land NRW weder eine einheitliche Methodik zur Erhebung der Versiegelungsdaten noch eine zentrale Datenbank, in der Versiegelungsinformationen für wasserwirtschaftliche Anwendungen vorgehalten werden. Die bisherige Abschätzung des Versiegelungsgrads von Siedlungsbereichen erfolgt dabei beispielsweise auf Basis der Analyse von Daten aus digitalisierten Luftbildern oder abgeschätzten Durchschnittswerten für Verkehrs- und Siedlungsflächen.

Diese Methoden sind jedoch für eine landesweit flächendeckende Erhebung von versiegelten Flächen in NRW im Hinblick auf verschiedene Gütekriterien wie beispielsweise Wiederholbarkeit und eine automatisierte Erhebung nicht geeignet, um mit einem vertretbaren Aufwand den unterschiedlichen Anforderungen aus der Praxis gerecht zu werden. In Abhängigkeit der Aufgabenstellung ist aktuell eine aufwendige Kartierung der Bodenversiegelung in unterschiedlicher räumlicher Auflösung notwendig. Mit der bundesweiten Einführung des "Amtlichen Liegenschaftskatasterinformationssystems" (ALKIS) existiert erstmals eine geeignete Datengrundlage, die eine für die flächendeckende Erfassung von versiegelten Flächen ermöglicht.

Gebäude-, Verkehrs- und Zuwegungsflächen bilden die drei Flächenkategorien, die den größten Anteil an versiegelten Flächen darstellen. Durch vorhandene Flächenpolygone sollen Hausumringen und Verkehrsflächen anhand einer objektorientierten Auswahl und Zuwegungsflächen mittels statistischen Verfahren ermittelt werden.

Ziel des Projektes ist es, den zuständigen Wasserbehörden sowie weiteren Anwendern von Daten zu versiegelten Flächen eine EDV-Anwendung zur Verfügung zu stellen, mit der befestigte Flächen für beliebig abgrenzbare Raumeinheiten in NRW mittels einer wissenschaftlich abgesicherten, dokumentierten und reproduzierbaren Methode abgeleitet werden können. Im Einzelnen soll überprüft werden, ob ALKIS-Daten für die flächendeckende Erhebung von Gebäudeflächen und von Verkehrsflächen geeignet sind und ob es möglich ist aus ALKIS-Daten befestigte Zuwegungsflächen abzuleiten.

## 2 Versiegelte Flächen

In Deutschland entstanden täglich im Durchschnitt 104 ha neue Siedlungs- und Verkehrsflächen (1992 - 2015). Zwischen 1992 und 2011 ist der Anteil der versiegelten Fläche insgesamt um 19,4 % angestiegen. In NRW liegt der Zuwachs bei 14,3 %. Von der Entwicklung und Zunahme der Siedlungs- und Verkehrsflächen abgesehen, haben die Bundesländer Bayern mit 17,1 % und Nordrhein-Westfalen mit 16,3 % (Abbildung 1) im bundesvergleich die höchsten Versiegelungsanteile (Stand 2011). Der zunehmende Flächenverbrauch hat Auswirkungen auf die Umwelt, bspw. begünstigt er Hochwasser und verändert den natürlichen Wasserhaushalt (STATISTISCHES BUNDESAMT, 2017).

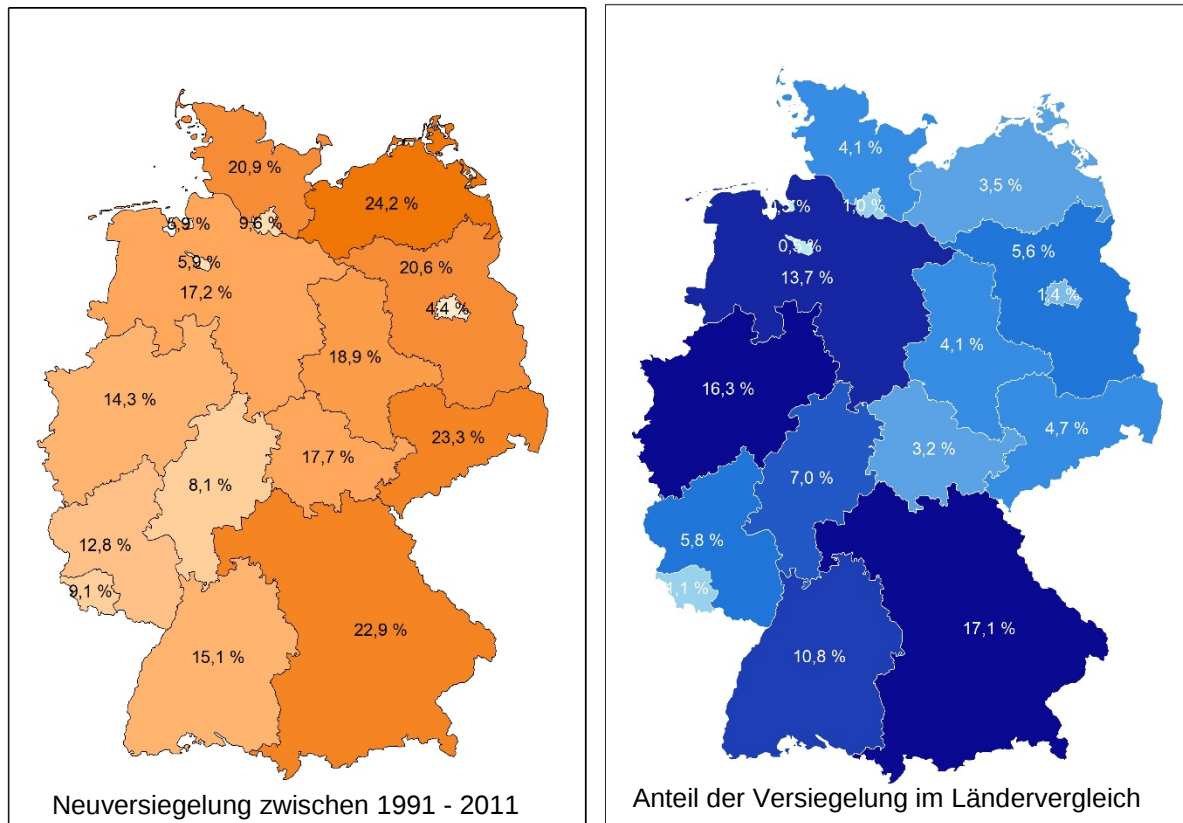


Abbildung 1: Neuversiegelung und Anteil der versiegelten Flächen in Deutschland

Informationen über versiegelte Flächen sind in Verbindung mit ihrer räumlichen Lage eine wichtige Basisinformation in der Wasserwirtschaft. Als versiegelte Flächen werden nach Arbeitsblatt DVWK-A 198) durchlässige und undurchlässige Flächen mit Unterbau bezeichnet. Versiegelungsdaten werden regelmäßig von den zuständigen Behörden sowie Fachplanern, die sich mit Themen des Umweltschutzes befassen, für die Stadt- und Landesplanung und für die Wasserwirtschaft genutzt. Die Forderung nach zeitlich hoch aufgelösten Versiegelungsdaten wird in der Politik immer lauter, um umweltpolitische und stadtplanerische Strategien nachhaltig sichtbar zu machen (SENATSWERWALTUNG FÜR STADTENTWICKLUNG UND WOHNEN, 2012).

Auch in der Fachwelt gewinnen Daten zu versiegelten Flächen zunehmend an Bedeutung, z.B. bei der Erarbeitung des neuen Arbeitsblattes DWA-A 102, welches sich noch im Gelbdruck befindet. Dieses Arbeitsblatt 102 befasst sich mit den „Grundsätzen zur Bewirtschaftung und Behandlung von

Regenwetterabflüssen zur Einleitung in Oberflächengewässer“. Es enthält emissions- und immissionsbezogene Grundsätze und Vorgaben zum Umgang mit niederschlagsbedingten Siedlungsabflüssen. Darin ist als wesentliche Grundlage der Betrachtungen zu Menge (Quantität) und Beschaffenheit (Qualität) des Oberflächenabflusses eine möglichst genaue Flächenermittlung mit den Abfluss beeinflussenden Merkmalen genannt (BECK ET AL., 2016).

### 3 Datengrundlage

Das Land NRW stellt aktuell über 2.700 Datensätze für unterschiedliche Bereiche des öffentlichen Lebens über sein Open Data Portal zu Verfügung, unter anderem Transport- und Verkehr, Umwelt und Klima, Gesundheit, Kultur und Freizeit sowie Geobasisdaten. Diese sind seit dem 01. Januar 2017 kostenlos auf der Internetplattform [www.opengeodata.nrw.de](http://www.opengeodata.nrw.de) verfügbar. Für die Nutzung der Daten wird hauptsächlich die Datenlizenz „Deutschland 2.0 Namensnennung“ verwendet. Sie erfordert lediglich die Nennung des Urhebers und einen Hinweis, ob diese verändert wurden. Die Daten können deshalb für beliebige Zwecke — auch für die kommerzielle Verwendung — ohne Genehmigung und Gebührenpflicht verwendet werden.

#### 3.1 ALKIS

Im Amtlichen Liegenschaftskataster-Informationssystem (im Folgenden ALKIS) werden alle Daten des Liegenschaftskatasters zusammengeführt. Hier werden im Wesentlichen Flurstücke und Gebäude (Liegenschaften) bundesweit dargestellt, beschrieben und nachgewiesen. Das Liegenschaftskatasterinformationssystem erfüllt unterschiedliche Aufgaben, u. a. die Lieferung von Basisdaten für die Geodateninformationsstruktur (GDI) in Deutschland und dient als amtliches Verzeichnis der Grundstücke für den Eigentumsnachweis im Grundbuch (AdV, 2018). Objektarten (Feature-Class). Die Aktualisierung der ALKIS-Daten wird bei Bedarf vorgenommen und stellt eine wichtige Anforderung für die Praxis dar. Der Aktualisierungsprozess dauert von der Entstehung der Veränderung bis zur Freigabe des fortgeführten Datenbestandes ca. sechs bis zwölf Wochen (AdV, 2008b).

#### 3.2 DGM

Das Digitale Geländemodell mit einer Gitterweite 25 m (DGM25) beschreibt die Geländeformen der Erdoberfläche und ist in Lage und Höhe georeferenziert (UTM Zone 32, DHHN92). Die Lagegenauigkeit beträgt 1 bis 3 m. Das DGM1 ist sowohl in der räumlichen Auflösung als auch Höhengenuigkeit deutlich genauer.

#### 3.3 Bevölkerungskataster

Beim Zensus 2011 wurden mehrere Millionen Datensätze erhoben, um aktuelle Zahlen zum Leben, Wohnen und Arbeiten in Deutschland zu gewinnen. Mit dem Zensus 2011 nahm Deutschland an einer EU-weiten Zensusrunde teil, die künftig alle zehn Jahre stattfinden wird. Die Ergebnisse liegen auf Bundes-, Länder-, Kreis- und Gemeindeebene vor. Außerdem sind einzelne Ergebnisse für Gitterzellen verfügbar. In diesem Projekt wurden 100-Meter-Gitter verwendet. Im Jahr 2021 wird der nächste Zensus stattfinden. Mit dem Zensus sollen in Deutschland die amtliche Einwohnerzahl sowie weitere tief gegliederte Daten zur Bevölkerung, Erwerbstätigkeit und zum Gebäude- und Wohnungsbestand gewonnen werden. Diese Daten sind flächendeckend verfügbar und stehen zur freien Verfügung.

### 3.4 Referenzdaten von Emschergenossenschaft und Lippeverband

Die Emschergenossenschaft und der Lippeverband sind öffentlich-rechtliche Wasserwirtschaftsverbände, die im Sinne der Gewässerunterhaltung der Flüsse Emscher und Lippe sowie ihrer Nebenflüsse aktiv sind. Die Emscher und die Lippe sind rechtsseitige Nebenflüsse des Rheins. Die in diesem Gebiet erzeugten Versiegelungsinformationen, die in einem manuellen Prozess anhand von Orthophotos digitalisiert werden, stehen als georeferenzierte Referenzdaten zur Verfügung (Abbildung 2).

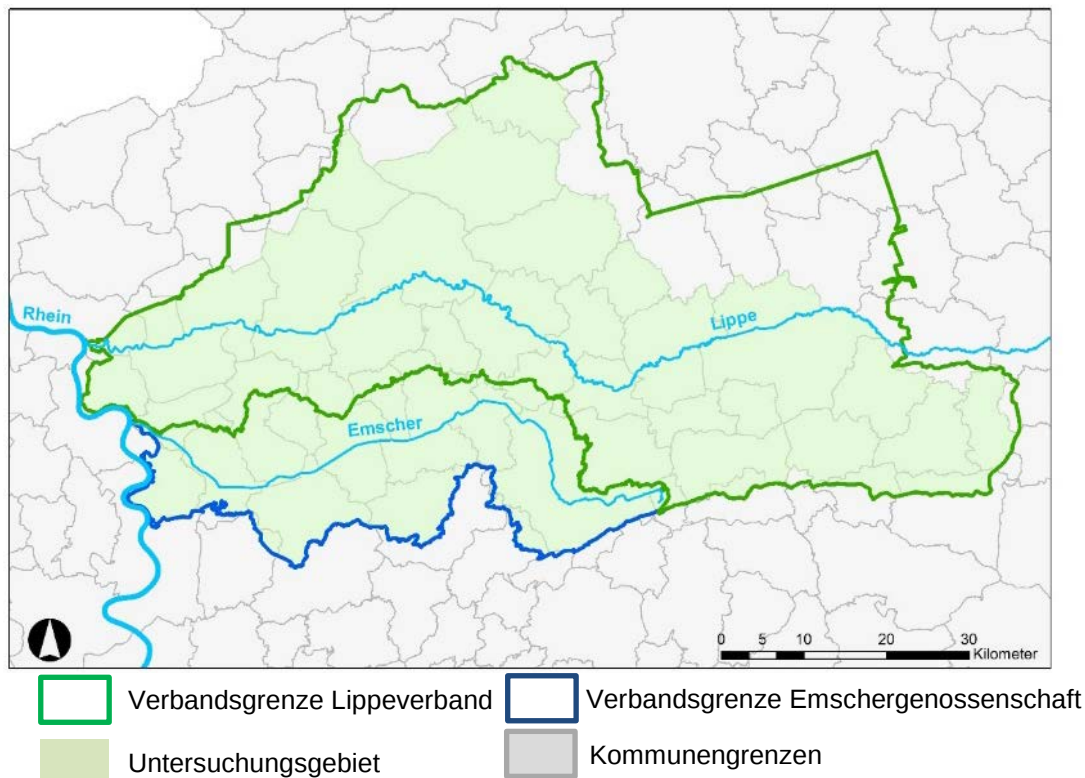


Abbildung 2: Verbandsgebiet von Emschergenossenschaft und Lippeverband

Die Bestimmung der Versiegelungsflächen erfolgt anhand der „Leistungsbeschreibung für die ArcGIS-konforme Erfassung der befestigten Flächen“ von EGLV. Die Flächenpolygone werden in sechs Befestigungsarten unterteilt, welche jeweils mit einem Objektschlüssel (OS) in der Attributtabelle kodiert sind (Tabelle 1). „Halden und Deponien“ beinhalten in Betrieb befindliche bzw. bereits rekultivierte Halde- und Deponieflächen. Dachflächen und Gebäude umfassen Wohn- und Verwaltungsgebäude, Nebengebäude und Gartenhäuser. Öffentliche Verkehrsflächen beschreiben befestigte wasserundurchlässige Flächen. Zu den öffentlichen Verkehrsflächen gehören Straßen, öffentlich genutzte Parkplätze, Autobahnrastplätze, Fuß- und Radwege entlang von Straßen, Einkaufsstrassen, Fußgängerzonen, Marktplätze sowie Bahnsteige. Der Schienenverkehr wird nicht als befestigte Fläche erfasst. Die Zuwegungsflächen werden in „Privat saubere“ (OS 2300) und „Privat schmutzige“ (OS 2400) Flächen unterteilt. „Privat saubere“ Flächenpolygone sind Flächen mit geringer Belastung, bspw. Zuwegungen an und zu Gebäuden, Terrassen an Wohnhäusern, Schulhöfe, Sport- und Tennisplätze sowie Freibäder und festinstallierte Schwimmbecken. „Privat schmutzige“ Flächen hingegen sind

Zuwegungen, die befahrbar sind und im Niederschlagsfall behandlungsdürftigen Abfluss verursachen. Unter „Privat schmutzige“ Flächen fallen beispielweise Garagenzufahrten und befahrbare Wege auf Privatgrundstücken, Park- und Lagerplätze, Fabrik- und Garagenhöfe, Tankstellen, sowie Flughäfen, Schlamm- und Klärbecken (EGLV, 2012).

Tabelle 1: Unterteilung der befestigten Flächen anhand der „Leistungsbeschreibung für die ArcGIS-konforme Erfassung der befestigten Flächen“ von EGLV

| Objekt-Schlüssel (OS) | Erläuterung                      |
|-----------------------|----------------------------------|
| 0060                  | Halden und Deponien in Betrieb   |
| 0070                  | Halden und Deponien rekultiviert |
| 2100                  | Dachflächen – Gebäude            |
| 2200                  | Öffentliche Verkehrsflächen      |
| 2300                  | Privat saubere Flächen           |
| 2400                  | Privat schmutzige Flächen        |

In Abbildung 4 ist der Anteil der versiegelten Gebäude-, Verkehrs und Zuwegungsflächen für das gesamte EGLV-Gebiet und für vier ausgewählte Kommunen dargestellt. Es wird deutlich, dass flächendeckend eine Verteilung von jeweils ca. ein Drittel zugrunde liegt. Deshalb ist die Ermittlung aller drei Bestandteile unabdingbar. Die zur Anwendung kommenden Referenzdaten (aus Orthophotos abgeleitete versiegelte Flächen) werden im Schnitt alle fünf Jahre aktualisiert. Die Aktualisierung erfolgt nicht flächendeckend, sondern gemeindeweise nach Bedarf. Für die hier durchgeführte Analyse liegen die Daten mit dem Bearbeitungsstand 2011 - 2017 zu Grunde (Siehe Abbildung 3). Der Stand der Daten ist bei der Beurteilung der Analyseergebnisse stets zu berücksichtigen.



## 4 Methodik

### 4.1 Vorbemerkung

Das FiW konnte in einer Vergleichsstudie, im Auftrag des Lippeverbands (LV) und Wupperverbands (WV) in Kooperation mit dem LANUV, eine hohe Übereinstimmung zwischen relevanten Objektarten aus ALKIS (Gebäude- und öffentliche Verkehrsflächen) und abgeleiteten Flächenpolygonen aus Luftbildern feststellen. Damit wurde die Grundannahme, dass eine (halb)automatisierbare Ableitung von befestigten Flächen auf Basis von landesweit verfügbaren, amtlichen Daten realisierbar ist, bestätigt. Weiterhin wurde festgestellt, dass Zuwegungsflächen zu Gebäuden, z.B. Garageneinfahrten oder Terrassen, anhand von ALKIS-Objektarten nicht selektiert werden können.

Im Folgenden werden in Kapitel 4.2 Methoden zur Einschätzung der Eignung von ALKIS-Daten für die flächendeckende Identifizierung von Gebäudeflächen und öffentlichen Verkehrsflächen und in Kapitel 4.3 statistische Verfahren zur Abschätzung von Zuwegungsflächen vorgestellt.

### 4.2 Bestimmung der Gebäude- und öffentlichen Verkehrsflächen

Zur Bestimmung der Gebäudeflächen und öffentlichen Verkehrsflächen aus ALKIS werden in einem ersten Schritt ALKIS-Objektarten anhand ihrer Semantik, Definition und Attributarten und anschließend visuell überprüft (qualitativ). Abschließend werden die ausgewählten Objekte mithilfe von aus Orthophotos abgeleiteten und manuell digitalisierten Flächenpolygonen auf ihre Güte überprüft (quantitativ).

### 4.3 Bestimmung von Zuwegungsflächen

Die Zuwegungsflächen können anhand der Selektion von ALKIS-Objektarten nicht unmittelbar erfasst werden. Vor diesem Hintergrund werden für die statistische Abschätzung von Zuwegungen Methoden eingesetzt, die eine möglichst präzise Ableitung von Zuwegungsflächen anhand von ALKIS-Daten ermöglichen. Im Allgemeinen kommen zwei Lerntechniken zum Einsatz: das überwachte Lernen (supervised learning) und das nicht überwachte Lernen (unsupervised learning). Für die Bestimmung der Zuwegungsflächen werden Algorithmen des überwachten Lernens angewendet. Denn zusätzlich zu der erklärenden Variablen  $\vec{x}$  ist ein Beobachtungswert  $y$  gegeben. Mit Hilfe einer bekannten Datenmenge und bekannten Ausprägungen werden Modelle trainiert, die fundierte Abschätzungen für den Ausgang von neuen und unbekannten Daten erzeugen können. Die aus Orthophotos abgeleiteten versiegelten Zuwegungsflächen, die als Beobachtungswerte dienen, können als qualitative Information zur Klassifikation von Bezugsflächen beitragen, um die Existenz von Zuwegungen vorherzusagen. Als quantitative Information dienen sie der Regressionsanalyse und der Abschätzung von Flächengrößen. Zum Einsatz kommen lineare und nicht lineare multiple Regressionsmodelle. Des Weiteren lassen sich große Datenmengen durch ein einziges Regressionsmodell oft nur unzureichend abbilden. Deshalb werden bei Bedarf Teilmodelle (Cluster) gebildet. Jedes Cluster wird anschließend separat betrachtet und mit einem Regressionsmodell nachgebildet. Clusteranalysen werden verwendet, um beispielsweise verborgene Gruppierungen in Daten zu identifizieren.

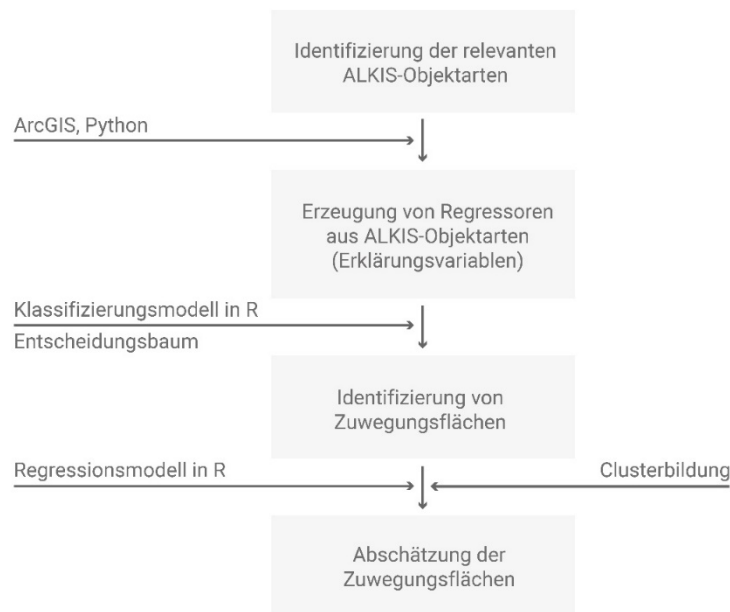


Abbildung 5: Übersicht der Vorgehensweise zur Abschätzung der Zuwegungsflächen

Abbildung 5 zeigt eine Übersicht über die Analyseschritte und die angewendeten Werkzeuge und Methoden. Zunächst werden relevante Objektarten aus ALKIS identifiziert. Aus den ausgewählten Objektarten werden Regressoren bestimmt, die eine statistische Abschätzung der Zuwegungsflächen ermöglichen. Da eine Vielzahl der Bezugsflächen keine Zuwegungen aufweisen, müssen diese erst identifiziert und aus den weiteren statistischen Analysen ausgeschlossen werden. Dazu dient ein Klassifizierungsmodell mit der Methode des Entscheidungsbaumes. Abschließend werden unter Anwendung von Transformationsverfahren und Clusteranalysen Regressionsmodelle erzeugt, um die Zuwegungsflächen abzuschätzen.

## 5 Ermittlung der Eingangsdaten für die statistische Abschätzung der Zuwegungsflächen

### 5.1 Erzeugung der Erklärungsvariablen

Bei Auswahl von Merkmalen, die als unabhängige Variablen zur statistischen Analyse dienen sollen, müssen

- Objektivität (keine Beeinflussung durch den Bearbeiter)
- Zuverlässigkeit (übereinstimmende Ergebnisse bei wiederholtem Erstellen der Variablen) und
- Gültigkeit (Sachverhalt entspricht der Realität)

der Daten berücksichtigt werden. Bei der Erzeugung dieser Merkmale ist darauf zu achten, dass sich keine systematischen Fehler (*bias*) einschleichen, die sich auf eine Vielzahl der Stichproben auswirken. Unvermeidbar hingegen sind zuverlässige Fehler, wobei ihre Größe für die Reproduzierbarkeit ausschlaggebend ist. Unter Beachtung dieser Kriterien wurde im Rahmen der Untersuchungen mit Hilfe der Skriptsprache Python ein Tool zur automatischen Bestimmung und Ausgabe von diversen unabhängigen Variablen aus dem ALKIS-Datensatz entwickelt und angewendet. Die unabhängigen Variablen resultieren aus ALKIS-Objektarten, dem digitalen Geländemodell (DGM 25) sowie dem Bevölkerungskataster. Die Auswahl der Regressoren aus ALKIS wurde anhand logischer Überlegungen und vermuteter Zusammenhänge zwischen den Regressoren und Zuwegungsflächen getroffen. Die Regressoren werden in metrische und kategoriale Regressoren unterteilt. Bei metrischen Merkmalen lassen sich die Ausprägungen durch Zahlen darstellen (quantitative Merkmale). Im Gegensatz dazu ermöglichen kategoriale Regressoren eine qualitative Unterscheidung der Zuwegungsflächen z.B. die Unterscheidung von Zuwegungsflächen, die sich in „Wohnbauflächen“ und in „Industrie- und Gewerbeflächen“ befinden.

Tabelle 2: Ableitung der unabhängigen Variablen (Regressoren) aus ALKIS

| Regressoren                                                  |
|--------------------------------------------------------------|
| Flurfläche                                                   |
| Anzahl der angrenzenden Flurstücke                           |
| Länge der angrenzenden Flurstücke                            |
| Länge der umlaufenden Verkehrsflächen                        |
| Gebäudefläche                                                |
| Gebäudeanzahl auf einem Flurstück                            |
| Bebauungsdichte (Verhältnis Gebäudefläche – Flurstückfläche) |
| Geländeneigung                                               |
| Kategorie Siedlungen                                         |
| Kategorie Maßgebliche Gebäudefunktion                        |
| Kategorie Gebäudeanzahl                                      |

|                              |
|------------------------------|
| Kategorie Bevölkerungsdichte |
| Kategorie Bebauungsdichte    |

In Tabelle 2 sind die erzeugten Regressoren tabellarisch aufgelistet. Die ermittelten Regressoren werden im Folgenden erläutert.

## 5.2 Referenzdaten

Die Zuwegungsflächen werden im EGLV-Datensatz als „Privat Saubere“ und „Privat Schmutzige“ Flächenpolygone geführt (vgl. Kapitel 3.4). Ziel der Untersuchungen ist es, die Flächengrößen von Zuwegungsflächen möglichst präzise abzuschätzen. Die Behandlungsbedürftigkeit dieser Flächen spielt dabei keine Rolle. Deshalb werden die „Privat Saubere“ und „Privat Schmutzige“ Flächenpolygone zusammengeführt und mit der Objektart Flurstück aus ALKIS verschnitten. Dabei wurde festgestellt, dass ca. 20 % der Flurstücke keine Überlagerung mit Zuwegungsflächen aufweisen. Hier werden im Allgemeinen zwei Fälle unterschieden: Zum einen existieren Flurstücke, die tatsächlich keine Zuwegungsflächen aufweisen und zum anderen Flurstücke, die zwar eine Zuwegung besitzen, die aber durch die EGLV-Daten nicht erfasst wurde. Eine Quantifizierung der Unvollständigkeit ist nicht möglich, da keine weiteren Vergleichsdaten vorliegen. Deshalb wird vereinfacht davon ausgegangen, dass relevante abflusswirksame Zuwegungsflächen, wie z.B. asphaltierte und betonierte Flächen, vollständig erfasst wurden. Vor diesem Hintergrund wird ein Klassifikationsmodell (Entscheidungsbäume) eingesetzt, mit dem Flurstücke ohne Zuwegungsflächen identifiziert und aus den Regressionsanalysen ausgeschlossen werden können.

## 6 Untersuchungsergebnisse

### 6.1 Gebäudeflächen aus ALKIS

Tabelle 3 zeigt die Auswahl der Gebäude aus ALKIS-Objektarten inkl. der Attributart, anhand derer die Objekte selektiert werden. Beispielsweise werden in der Objektart „Sonstige Bauwerke oder Sonstige Anlagen“ mit der Kennung 51009 die Attribute mit den Wertarten 1610 „Überdachung“ und 1611 „Carport“ selektiert. Die Gesamtheit der Auswahl beschreibt die Gebäudeflächen, die aus ALKIS-Daten selektiert werden können.

Tabelle 3: Auswahl von Objektarten zur Bestimmung von Gebäudeflächen

|                    | Objektart                                           | Kennung | Attributart zur Auswahl | Wertarten               |
|--------------------|-----------------------------------------------------|---------|-------------------------|-------------------------|
| <b>Dachflächen</b> | Gebäude                                             | 31001   | Lage zur Erdoberfläche  | 1200<br>ausgenommen     |
|                    | Turm                                                | 51001   |                         |                         |
|                    | Bauwerke oder Anlage für Industrie und Gewerbe      | 51002   | Bauwerksfunktion        | 1215,1220,<br>129, 1350 |
|                    | Vorratsbehälter<br>Speicherbauwerk                  | 51003   |                         |                         |
|                    | Bauwerk od. Anlage für Sport, Freizeit und Erholung | 51006   | Bauwerksfunktion        | 1431,1490               |
|                    | sonstige Bauwerke oder sonstige Einrichtung         | 51009   | Bauwerksfunktion        | 1610,1611               |

Bei der Gegenüberstellung der summierten Gebäude- und Dachflächen auf Flurstücken werden 1.021.570 Flurstücke untersucht. Die Gebäudepolygone aus beiden Datensätzen werden durch Verschneidung mit Flurstücken diesen zugeordnet. Sie befinden sich alle innerhalb der Objektgruppe Siedlung der „Tatsächlichen Nutzung“. Hierbei werden sämtliche Gebäude- und Dachflächen auf einem Flurstück zu einer Einheit verschmolzen und mit dem anderen Datensatz verglichen.

Aus der Verschmelzung der Dach- und Gebäudeflächen auf einem Flurstück resultieren verschiedene Kombinationen. Es existieren sowohl Flurstücke, die weder ALKIS-Gebäudepolygone noch EGLV-Gebäudepolygone beinhalten als auch Flurstücke, die entweder ALKIS-Gebäudepolygone oder EGLV-Gebäudepolygone aufweisen und schließlich Flurstücke, in denen Gebäudeflächen aus beiden Datensätzen verfügbar sind. Tabelle 4 ist zu entnehmen, dass bei 19,97 % der betrachteten Flurstücke keine Zuordnung zu Gebäudepolygonen möglich ist. Auf diesen Flurstücken befinden sich keine Gebäude. Bei ca. 3,6 % der Daten befinden sich nur EGLV-Gebäudepolygone auf den Flurstücken und machen insgesamt eine Fläche von 1,7 km<sup>2</sup> aus. Dieser Flächenanteil konnte durch ALKIS-Daten nicht erfasst werden. Bei ca. 2,2 % der Daten befinden sich ausschließlich ALKIS-Gebäudepolygone auf den

Flurstücken mit einer Gesamtfläche von 3,2 km<sup>2</sup>. Hierbei wurde ein Überschuss an Gebäudeflächen erfasst. Die beiden zuletzt genannten Fälle bestätigen die Aussage, dass beide Datensätze die Realität nicht vollständig darstellen. Gründe dafür sind unter anderem Neubauten, die aufgrund der Aktualisierungsperioden der Datensätze noch nicht erfasst wurden. Ein weiterer Grund ist die unvollständige Erfassung von Gebäuden durch Bearbeiter bzw. Katasterämter. Die restlichen Flurstücke (ca. 75 %) werden sowohl von ALKIS–Gebäudepolygonen als auch von EGLV–Gebäudepolygonen überlagert. Im gesamten Untersuchungsgebiet ist ein Defizit von ca. 16 km<sup>2</sup> bei den ALKIS–Daten zu beobachten. Im Folgenden ist dargestellt, wie diese Abweichung zustande kommt.

Tabelle 4: Zusammensetzung unterschiedlicher Fälle bei der Gegenüberstellung von Gebäudepolygonen aus ALKIS und EGLV auf einzelne Flurstücke im Untersuchungsgebiet von EGLV

|   | ALKIS        | EGLV         | Anzahl<br>Datensätze | Anteil [%] | Fläche [m <sup>2</sup> ]<br>ALKIS | Fläche [m <sup>2</sup> ]<br>EGLV |
|---|--------------|--------------|----------------------|------------|-----------------------------------|----------------------------------|
| 1 | Kein Gebäude | Kein Gebäude | 201.229              | 19,97      | 0                                 | 0                                |
| 2 | Kein Gebäude | Gebäude      | 37.022               | 3,62       | 0                                 | 1.720.311                        |
| 3 | Gebäude      | Kein Gebäude | 20.692               | 2,18       | 3.254.279                         | 0                                |
| 4 | Gebäude      | Gebäude      | 717.523              | 74,23      | 189.731.800                       | 205.816.378                      |

In Abbildung 6 werden die summierten Gebäudeflächen auf einzelnen Flurstücken aus ALKIS und dem EGLV–Datensatz in einem Punktdiagramm (links) dargestellt. Es wird deutlich, dass ein sehr starker linearer Zusammenhang zwischen den Datensätzen besteht. Der Korrelationskoeffizient beträgt  $r = 0,99$ . Daraus lässt sich schließen, dass in der Summe eine sehr hohe Übereinstimmung zwischen den erfassten Gebäudeflächen in beiden Datensätzen vorherrscht. In dem rechten Boxplot in Abbildung 6 ist diese Übereinstimmung visualisiert, unabhängig davon, ob durch ALKIS größere oder kleinere Gebäudeflächen erfasst werden. Die mittlere Übereinstimmung der durch beide Datensätze erfassten Gebäudeflächen beträgt im Durchschnitt 86,32 % (rote gestrichelte Linien im Boxplot). Dreiviertel (1. Quantil) der Daten zeigen eine höhere Übereinstimmung von 81,5 % (blaue gestrichelte Linie im Boxplot). Die Streuung der Beobachtungen um den Mittelwert liegt bei 12,7 %. Die Ursache für die Diskrepanz zwischen ALKIS und EGLV–Daten wird deutlich, indem die Übereinstimmung zwischen den einzelnen Siedlungen (Objektarten der Objektartengruppe „Siedlung“ der „Tatsächlichen Nutzung“) gegenübergestellt wird. In Tabelle 5 ist die Übereinstimmung der summierten Gebäudeflächen auf Flurstücke für einzelne Siedlungen dargestellt. Die Siedlungen der „Tatsächlichen Nutzung“ werden in der ersten Spalte aufgeführt.

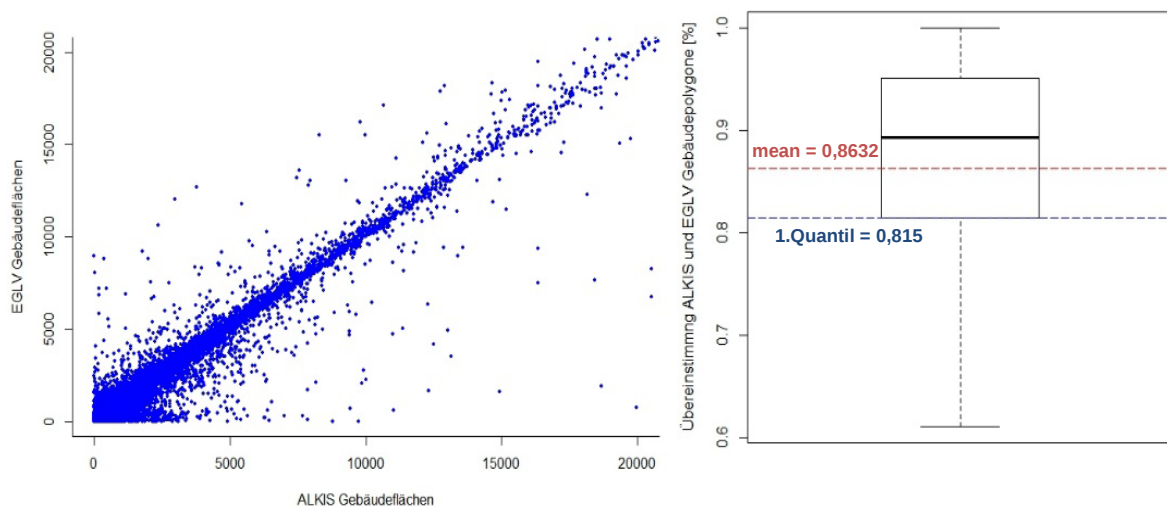


Abbildung 6: Gegenüberstellung der summierten Gebäudeflächen auf einzelne Flurstücke aus ALKIS und EGLV, links: Stärke des linearen Zusammenhangs; rechts: Grad der Übereinstimmung in Prozent

In der zweiten Spalte ist die Anzahl und in der dritten Spalte der Anteil der Daten pro Siedlungstyp dargestellt. Es wird deutlich, dass sich ein Großteil der Flurstücke auf „Wohnbauflächen“ (77,1 %) befindet. Weitere relevante Siedlungsgebiete sind „Industrie und Gewerbegebiete“, „Fläche gemischter Nutzung“, „Fläche besonderer funktionaler Prägung“ und „Sport-, Freizeit- und Erholungsflächen“. Die durchschnittliche Übereinstimmung zwischen ALKIS und EGLV-Daten (Spalte 4) ist in den relevanten Siedlungen sehr hoch, ausgenommen ist die Siedlung „Sport-, Freizeit- und Erholungsflächen“. Hier liegt die mittlere Übereinstimmung bei 67,7 %. In den folgenden Spalten 5 und 6 sind die durch ALKIS und EGLV-Daten erfassten Gebäudeflächen aufsummiert dargestellt. Der vorletzten Spalte „Delta“ ist das Defizit bzw. der Überschuss durch ALKIS zu entnehmen. Ausschließlich in den Siedlungen „Sport-, Freizeit- und Erholungsflächen“ wird durch ALKIS ein Überschuss an Gebäudeflächen erfasst. Das größte Defizit entsteht in den Siedlungen der Kategorie „Wohnbaufläche“. Aus diesen Beobachtungen lässt sich zunächst ableiten, dass ALKIS-Daten im Allgemeinen ein Defizit aufweisen und dieses flächendeckend (u.a. verteilt auf 77,1 % der Daten in Wohnbauflächen) vorliegt. Eine Überprüfung auf Extremwerte hat gezeigt, dass sich das Defizit bzw. der Überschuss zwischen ALKIS und EGLV-Gebäudeflächen für 98 % der Daten zwischen  $-44 \text{ m}^2$  und  $130 \text{ m}^2$  befinden. In diesem Bereich sind 87 % der Defizite bzw. der Überschüsse abgedeckt. Deshalb wird ausgeschlossen, dass das Defizit bzw. der Überschuss zwischen ALKIS und EGLV-Daten auf wenige Beobachtungen mit großer Gewichtung zurückzuführen ist. Die wichtigste Ursache für das flächendeckende Defizit an Gebäudeflächen, die durch ALKIS erfasst werden, ist die Erfassung der Grundfläche. Nicht berücksichtigt werden die Dachüberhänge, die bspw. bei Familienhäusern und Wohnanlagen

vorkommen. Diese These wird weiterhin durch die Erfassung der maßgeblichen Gebäudefunktion<sup>1</sup> auf Flurstücke bestätigt. Die Analyse hat gezeigt, dass 97 % des gesamten Flächendefizits auf Flurstücken mit der maßgeblichen Gebäudefunktion „Wohnhaus“ zu finden sind.

Tabelle 5: Die Übereinstimmung der summierten Gebäudeflächen auf Flurstücke getrennt nach Siedlungen

| Siedlungen                             | Anzahl [-] | Anteil [%] | Übereinstimmung [%] | EGLV [m <sup>2</sup> ] | ALKIS [m <sup>2</sup> ] | Delta [m <sup>2</sup> ] | Delta Anteil [%] |
|----------------------------------------|------------|------------|---------------------|------------------------|-------------------------|-------------------------|------------------|
| Wohnbaufläche                          | 584.879    | 77,1       | 85,5                | 105.711.335            | 93.204.378              | -12.506.957             | 77,8             |
| Industrie und Gewerbeflächen           | 40.299     | 5,3        | 87,9                | 52.077.179             | 50.846.029              | -1.231.150              | 7,7              |
| Halde                                  | 17         | 0,002      | 56,0                | 5.696                  | 4.025                   | -1.671                  | 0,01             |
| Bergbaubetrieb                         | 14         | 0,002      | 86,2                | 45.788                 | 42.223                  | -3.565                  | 0,02             |
| Tagebau Grube Steinbruch               | 6          | 0,0008     | 73,0                | 697                    | 805                     | 108                     | 0,0007           |
| Fläche Gemischter Nutzung              | 59.986     | 7,9        | 88,7                | 29.371.568             | 2.750.9347              | -1.862.221              | 11,6             |
| Fläche besonderer funktionaler Prägung | 65.241     | 8,6        | 92,2                | 15.232.479             | 14.668.639              | -563.840                | 3,5              |
| Sport-, Freizeit- und Erholungsfläche  | 7.334      | 1,0        | 67,7                | 3.198.151              | 3.298.149               | 99.998                  | 0,6              |
| Friedhof                               | 503        | 0,1        | 82,4                | 173.486                | 158.206                 | -15.280                 | 0,1              |

Dies zeigt sich auch bei der Betrachtung der gesamten erfassten Gebäudeflächen in den Kommunen. In Abbildung 7 sind die Kommunen sowie der prozentuale Anteil der fehlenden Gebäudeflächen visualisiert, die durch ALKIS nicht erfasst wurden. Diese bewegen sich zwischen 0 % und maximal 15 %. Die orange und gelb gefärbten Gebiete befinden sich mehrheitlich in weniger dicht besiedelten Kommunen. Die Kommunen mit einem größeren Defizit als 10 % machen ca. 18 % der gesamten Daten und 27 % des gesamten Flächendefizits im Untersuchungsgebiet aus, die durch ALKIS nicht erfasst werden konnten. Ein Drittel der Defizite verteilt sich somit auf ca. ein Viertel der Daten. Zusammenfassend lassen sich aus den Untersuchungen folgende Aussagen ableiten, wobei die Aktualität der EGLV – Daten berücksichtigt werden sollte.

- Die durch ALKIS erfassten summierten Gebäudeflächen auf Flurstücken zeigen eine hohe Übereinstimmung (Standardabweichung= 12,7 %) mit EGLV–Gebäudeflächen.
- Die Abweichung zwischen ALKIS- und EGLV–Gebäudeflächen ist flächendeckend vorhanden, insbesondere bei Wohnbauflächen.

<sup>1</sup> Die Funktion eines Gebäudes ist dann maßgeblich, wenn sie zu konkurrierenden Gebäuden die größte Fläche auf einem Flurstück aufweist. Sollte sich beispielsweise auf einem Flurstück eine Garage, ein Gartenhaus und ein Wohnhaus befinden, so wird die Gebäudefunktion des Wohnhauses als maßgeblich gewertet, weil das Wohnhaus im Vergleich zu den anderen Gebäuden die größte Fläche besitzt.

- Ursache hierfür ist die Erfassung der Grundflächen ohne Berücksichtigung der Dachüberhänge, dies kommt insb. bei Wohnhäusern vor.

Für eine Empfehlung zur Anpassung des vorhandenen Flächendefizits in ALKIS-Gebäudepolygonen wurden zwei unterschiedliche Ansätze für dieselben Flurstücke untersucht. Bei der ersten Methode wird die durchschnittliche Abweichung zwischen EGLV- und ALKIS-Gebäudeflächen auf Flurstücke ermittelt und anschließend zu einzelnen ALKIS-Gebäudeflächen summiert. Bei der zweiten Methode wird hingegen der Mittelwert der prozentualen Abweichung zwischen EGLV- und ALKIS-Gebäudeflächen auf Flurstücke ermittelt. Die resultierende Fläche, die sich aus der Multiplikation des erzeugten Faktors mit den ALKIS-Gebäudeflächen ergibt, wird zu ALKIS-Gebäudeflächen hinzugefügt.

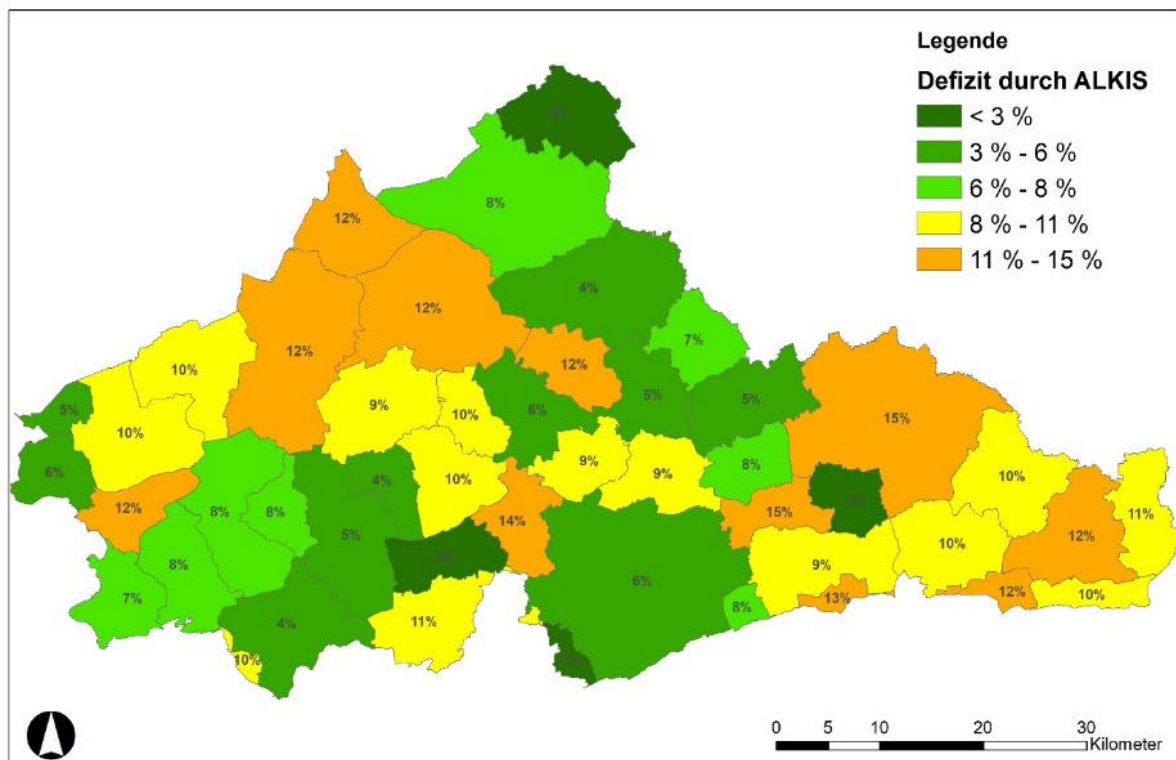


Abbildung 7: Prozentuales Defizit der durch ALKIS erfassten Gebäudeflächen im Vergleich zu EGLV - Gebäudeflächen

Hierzu werden beispielhaft Flurstücke mit der maßgeblichen Gebäudefunktion „Wohnhaus“ betrachtet. Für Methode 1 ergibt sich eine absolute Abweichung von ca. 20 m<sup>2</sup> pro Flurstück. Methode 2 ergibt eine mittlere prozentuale Abweichung von 14,5 %. Werden diese Flächen mit aktuellen ALKIS-Gebäudeflächen summiert, ergibt sich bei einer erneuten Untersuchung der Abweichungen zwischen EGLV-Gebäudeflächen und ALKIS-Gebäudeflächen über das gesamte Untersuchungsgebiet bei Methode 1 eine Abweichung von 0,5 m<sup>2</sup> und eine Abweichung von 2,7 % bei Methode 2. Bei der ersten Methode wird unabhängig von der vorhandenen Gebäudefläche pauschal 20 m<sup>2</sup> hinzuaddiert. Die Verteilung der Abweichungen der korrigierten ALKIS-Gebäudeflächen gegenüber EGLV-Gebäudeflächen ist in Abbildung 8 links dargestellt. Bei der zweiten Methode wird die vorhandene

Gebäudefläche immer um 14,5 % vergrößert. Bei Methode 2 führt die Anpassung der ALKIS-Gebäudeflächen zum Teil zu extrem großen Abweichungen der korrigierten Gebäudeflächen.

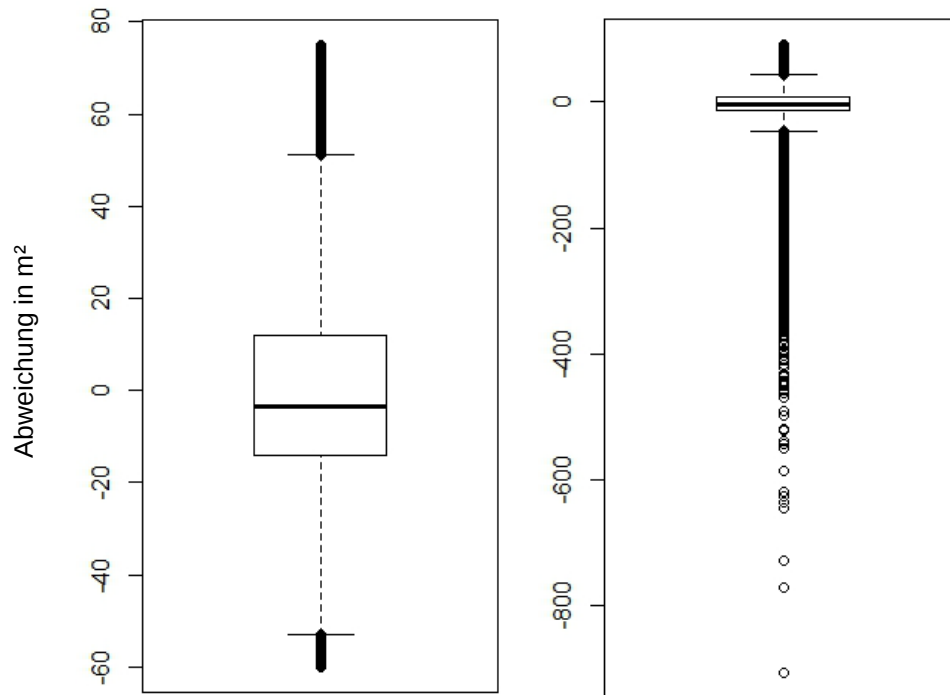


Abbildung 8: Verteilung der Abweichung zwischen korrigierten ALKIS-Gebäudeflächen und EGLV-Gebäudeflächen (links: Methode 1, rechts: Methode 2)

Es wird eine große Anzahl an Extremwerten erzeugt, die in der Summe aber auch für kleinräumige Untersuchungen die Anpassung negativ beeinflusst. Dies ist insbesondere bei ALKIS-Gebäudeflächen der Fall, die bereits eine größere Fläche als EGLV-Gebäude abbilden. Abbildung 9 zeigt die Verläufe der Abweichungen von korrigierten ALKIS-Gebäudepolygonen zu EGLV-Gebäudepolygonen für beide Methoden. Es wird deutlich, dass bei der Anpassung der ALKIS-Gebäudeflächen durch die mittlere prozentuale Abweichung (roter Verlauf) ein sehr großer Sprung im negativen Bereich entstehen kann. Anhand dieser Untersuchung wurden für die Flurstücke mit verschiedenen maßgeblichen Gebäudefunktionen Anpassungswerte erstellt (siehe Anhang 1). Damit eine ausreichende Repräsentativität gegeben ist, wurden ausschließlich maßgebliche Gebäudefunktionen betrachtet, die mindestens 1.000 Mal im Untersuchungsgebiet vorkommen.

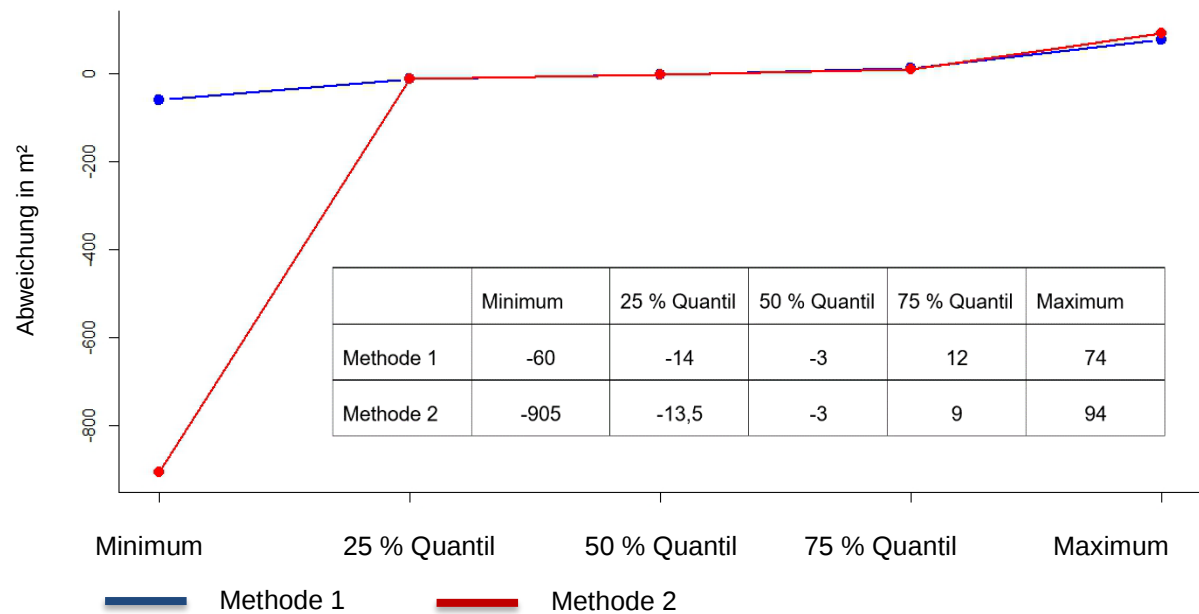


Abbildung 9: Verlauf der Abweichungen zwischen EGLV- und ALKIS Gebäudepolygonen bei unterschiedlichen Korrekturansätzen

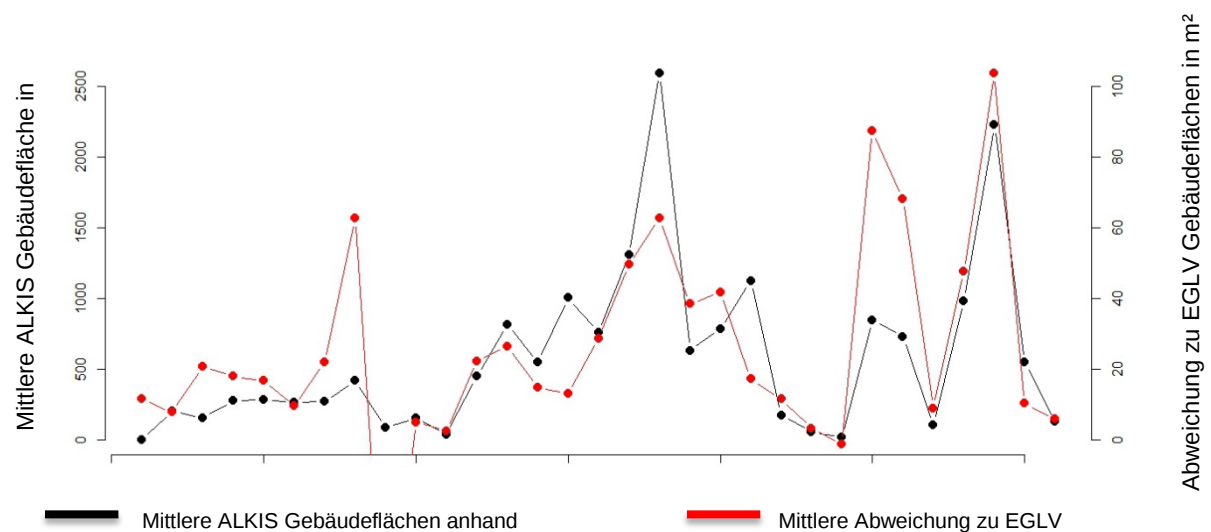


Abbildung 10: Zusammenhang zwischen der Gebäudegröße und die Größe der Abweichungen zu EGLV-Gebäudepolygonen

Die ermittelten Anpassungswerte korrelieren stark mit den Gebäudeflächen. D.h. je größer die Gebäudeflächen im Mittel ausfallen, desto größere absolute Abweichungen zu EGLV-Gebäudeflächen sind zu erwarten. Diese Erkenntnis ist plausibel, denn je größer die Gebäude werden, desto größer fallen die Dachüberhänge aus. Beispielsweise haben „Wohngebäude“ eine mittlere Flächengröße von 153 m<sup>2</sup>, das durchschnittliche Defizit zu EGLV-Gebäudeflächen und somit der Anpassungswert beträgt 20 m<sup>2</sup>. Garagen hingegen haben eine durchschnittliche Größe von 54 m<sup>2</sup> und einen Anpassungswert von 3,1 m<sup>2</sup>. Land- und forstwirtschaftliche Betriebsgebäude sind mit 845 m<sup>2</sup> weitaus größer, der durchschnittliche Anpassungswert dazu beträgt 86,1 m<sup>2</sup>. Abbildung 10 zeigt diesen Zusammenhang.

Jeder schwarze Punkt zeigt die mittlere Größe der ALKIS-Gebäude für eine bestimmte maßgebliche Gebädefunktion auf einem Flurstück (linke Y-Achse). Die korrespondierenden roten Punkte zeigen die Abweichungen zu EGLV-Gebäudeflächen (rechte Y-Achse).

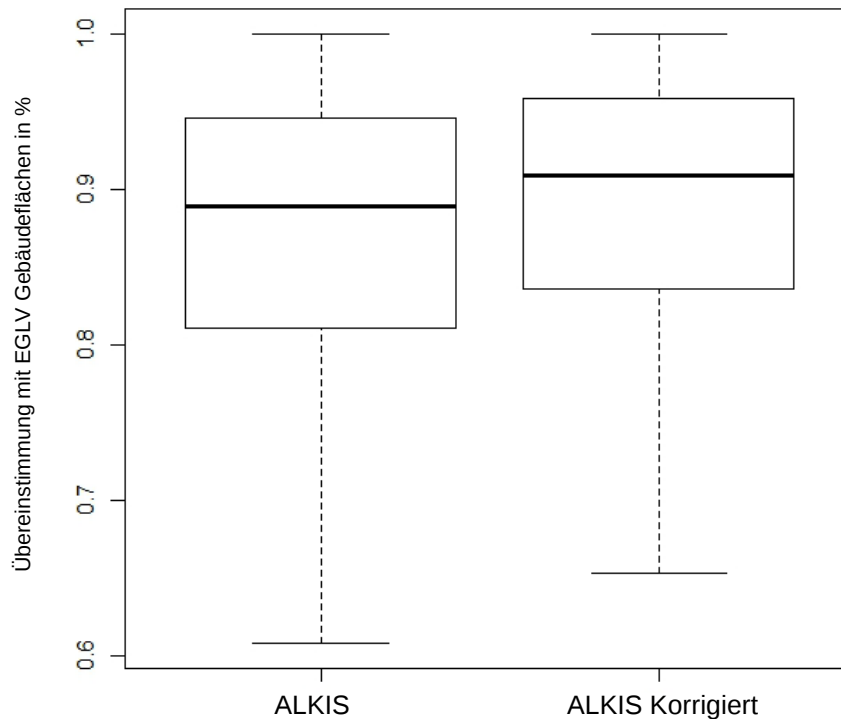


Abbildung 11: Vergleich der Übereinstimmungen von ALKIS Gebäudeflächen und korrigierten ALKIS-Gebäudeflächen mit EGLV-Gebäudeflächen

Bei Anwendung der Korrekturen auf die ALKIS-Gebäudeflächen kommen deutlich höhere Übereinstimmungen mit EGLV-Gebäudeflächen zustande. Abbildung 11 zeigt die Verbesserung der Übereinstimmungen durch die Korrektur der ALKIS-Gebäudeflächen. Die Box setzt deutlich höher bei ca. 85 % an und die mittlere Übereinstimmung steigt von 86 % auf 87,5 % an. Diese Optimierung zeigt sich vor allem bei der Summierung der Gebäudeflächen auf größere räumliche Einheiten, z.B. der Kommunen. Abbildung 12 zeigt die Übereinstimmung der korrigierten ALKIS Gebäudeflächen im Vergleich zu EGLV-Gebäudeflächen. Die Defizite werden mit maximal 6 % beziffert, in einigen Kommunen kommt es auch zu minimalen Flächenüberschüssen.

Dieses Ergebnis zeigt, dass anhand der Korrekturwerte eine sehr gute Annäherung der ALKIS-Gebäudeflächen an EGLV-Gebäudeflächen möglich ist.

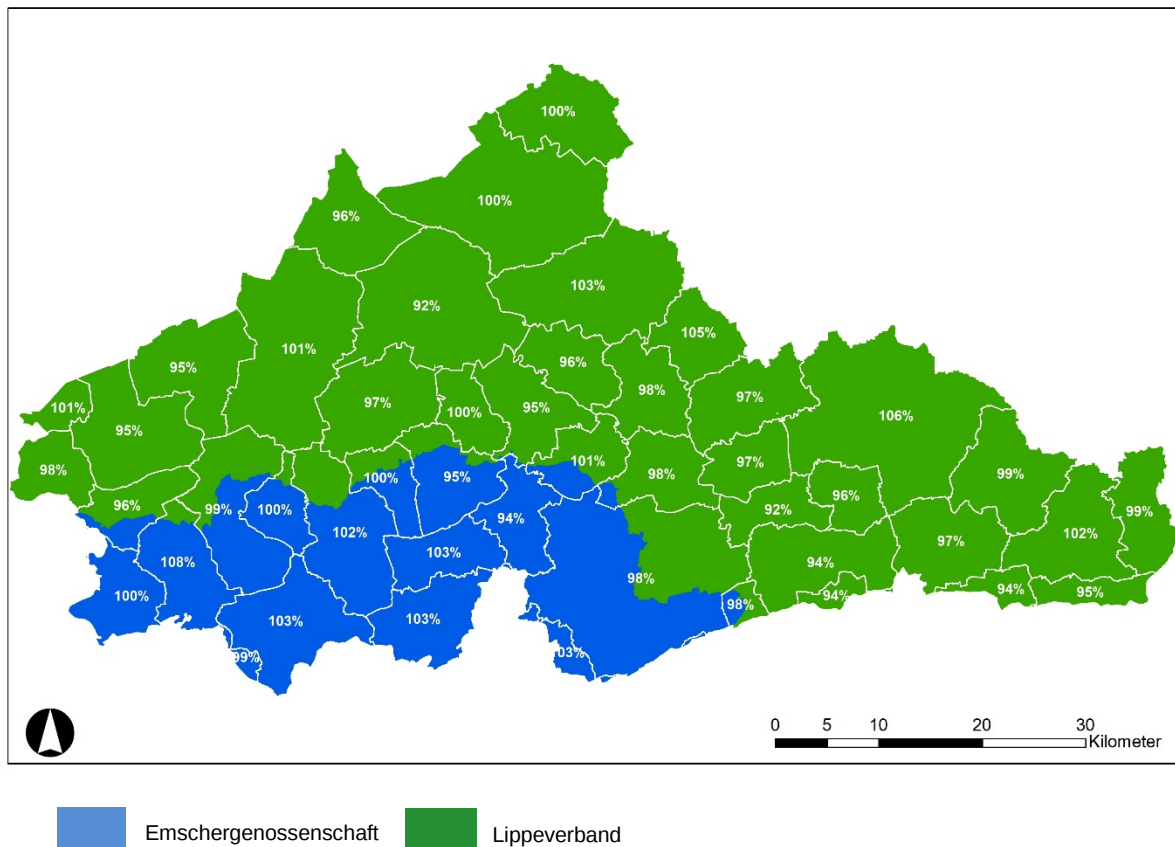


Abbildung 12: Übereinstimmung der Flächensummen der korrigierten ALKIS-Gebäudeflächen mit EGLV-Gebäudeflächen in den Kommunen von Emschergenossenschaft und Lippeverband

## 6.2 Verkehrsflächen aus ALKIS

Die identifizierten Objektarten, die Verkehrsflächen abbilden, werden in einem Lastenheft detailliert beschrieben. Die Selektion der durch die EGLV-Daten erfassten Verkehrsflächen ist anhand von ALKIS-Objektarten möglich. Wird ausschließlich die Objektart „Straßenverkehr“ (Kennung 42001) ausgewählt, entstehen für die einzelnen Kommunen, die in Abbildung 13 dargestellten Übereinstimmungen (Summe der Verkehrsflächen). Es ist festzustellen, dass im Verbandsgebiet der Emschergenossenschaft relativ hohe Übereinstimmungen mit Toleranzen von 10 % erreicht werden. Tendenziell nimmt diese Übereinstimmung jedoch in Richtung Norden ab wobei die Aktualität der EGLV-Daten immer berücksichtigt werden sollte. In Richtung Norden nimmt die Aktualität der Daten deutlich ab.

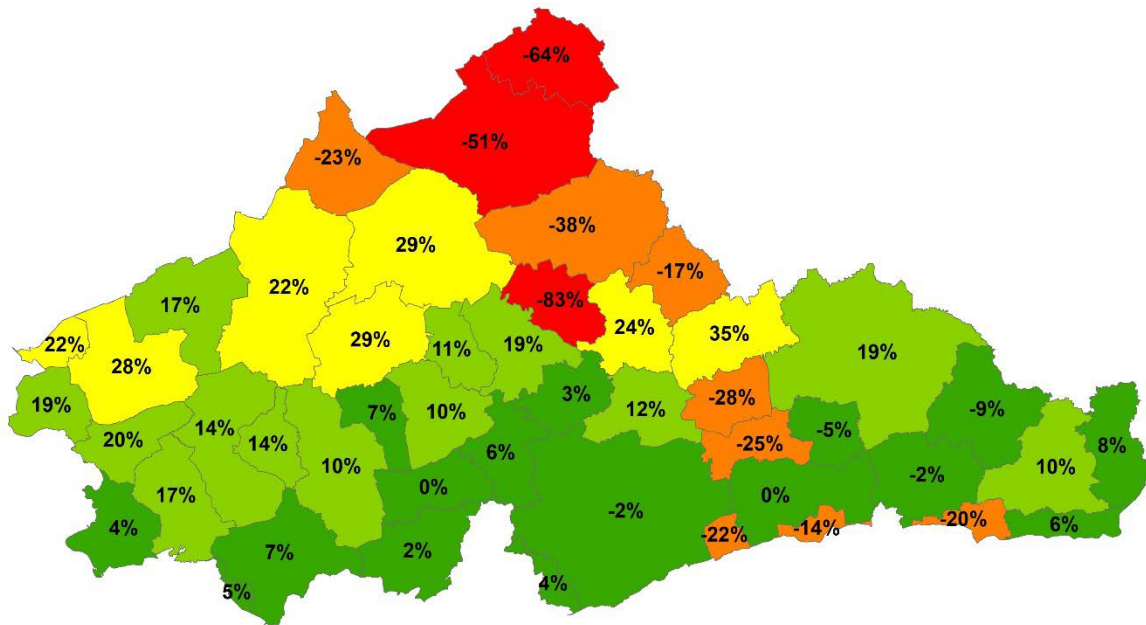


Abbildung 13: Übereinstimmung der Objektart "Straßenverkehr" mit EGLV Daten bezogen auf Kommunen im Einzugsgebiet der Emscher genossenschaft und Lippeverband

Abbildung 14 zeigt die Kommune Nottuln im Norden des Untersuchungsgebietes. In der linken Darstellung sind die durch die EGLV erhobenen befestigten Verkehrsflächen in blau zu erkennen. Im Vergleich dazu sind in der mittleren Abbildung Verkehrsflächen abgebildet, die durch die Objektart Straßenverkehr erfasst werden. Das Straßennetz wird mit einem Flächendefizit von 64 % erfasst. Über die Objektart „Straßenverkehr“ hinaus können zahlreiche weitere Objektarten (bspw. „Wege und Parkplätze“) aus ALKIS selektiert werden. Die Objektarten gehören den beiden Objektartengruppen „Verkehr“ (Objektartenbereich „Tatsächliche Nutzung“) und „Bauwerke, Anlagen und Einrichtungen für den Verkehr“ (Objektartenbereich „Bauwerke, Einrichtungen und sonstige Angaben“) an. Die Objektartengruppe mit der Bezeichnung „Verkehr“ und der Kennung 42000 enthält die bebauten und nicht bebauten Flächen, die dem Verkehr zur Verfügung stehen. Die Objektartengruppe mit der Bezeichnung „Bauwerke, Anlagen und Einrichtungen für den Verkehr“ und der Kennung 53000 enthält hingegen flächen- oder linienförmige Anlagen, die dem Verkehr dienen. In der rechten Abbildung ist das Ergebnis der erweiterten Auswahl in grüner Farbe zu sehen. Das Straßennetz wird deutlich dichter erfasst. Der ALKIS-Datensatz ermöglicht die Erfassung eines Flächenüberschusses von 43 % gegenüber dem EGLV-Datensatz. Der Stand der Erfassung der EGLV-Daten ist dabei zu berücksichtigen.

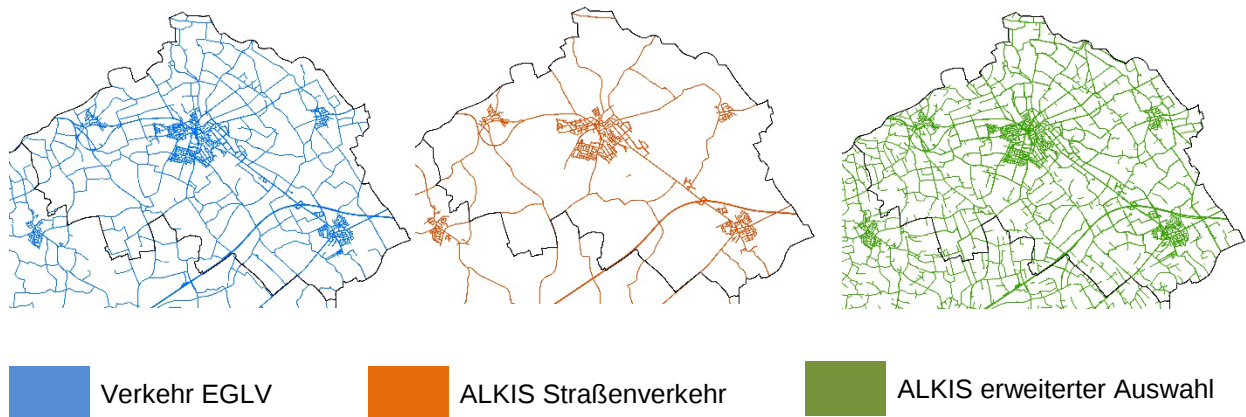


Abbildung 14: Erfassung von befestigten Verkehrsflächen durch EGLV und ALKIS

Weiterhin sei erwähnt, dass die zur Verfügung gestellten Referenzdaten von EGLV für den Bereich Verkehr keine Bahnkörper und Gleise abbilden. Gleise und Verkehrsbegleitflächen für den Bahnverkehr können durchaus anhand von ALKIS-Objektarten erfasst werden. Sie werden für die Gegenüberstellung der Datensätze vernachlässigt, um die Vergleichbarkeit zu gewährleisten. Die Bahnkörper werden durch die beiden Objektarten „Bahnverkehr“ und „Bahnverkehrsanlagen“ abgebildet. In Tabelle 6 ist die Auswahl der Objektarten aufgeführt, die Verkehrsflächen charakterisieren.

Tabelle 6: Auswahl an ALKIS – Objektarten, die Verkehrsflächen abbilden

| Objektartengruppen                                         | Objektart                   | Kennung | Attributart zur Auswahl | Wertarten                                                                          |
|------------------------------------------------------------|-----------------------------|---------|-------------------------|------------------------------------------------------------------------------------|
| <b>Verkehr</b>                                             | Straßenverkehr              | 42001   | Nutzungsfunktion        | 5130, NULL                                                                         |
|                                                            | Weg                         | 42006   | Nutzungsfunktion        | 5210, 5211, 5212, 5220, 5230, 5240, 5250, 5260, 9999                               |
|                                                            | Platz                       | 42009   | Nutzungsfunktion        | 5130, 5310, 5320, 5330, 5340, 5350                                                 |
| <b>Bauwerke, Anlagen und Einrichtungen für den Verkehr</b> | Bauwerke im Verkehrsbereich | 53001   | Bauwerksfunktion        | 1800, 1801, 1802, 1803, 1804, 1805, 1807, 1808, 1809, 1810, 1820, 1830, 1840, 1845 |
|                                                            | Straßenverkehrsanlagen      | 53002   | Verkehrsanlagenart      | 1000                                                                               |
|                                                            | Weg, Pfad, Steig            | 53003   | Nutzungsart             | 1106, 1111                                                                         |

|  |                               |       |                    |                           |
|--|-------------------------------|-------|--------------------|---------------------------|
|  | Bahnverkehrs-<br>anlage       | 53004 | Bahnhofskategorie  | 1010, 1020, 1030          |
|  | Flugverkehrs-<br>anlage       | 53007 | Verkehrsanlagenart | 1310, 1320, 1330,<br>5531 |
|  | Bauwerk im<br>Gewässerbereich | 53009 | Bauwerksfunktion   | 2030, 2020, 2050          |

Wird die erweiterte Auswahl erneut mit EGLV-Verkehrsflächen verglichen, ist festzustellen, dass in allen Kommunen Flächenüberschüsse zu beobachten sind. Diese reichen von 7 % bis über 60 % (siehe Abbildung 15). Die Aktualität der EGLV-Daten spielt bei dieser Beobachtung eine wichtige Rolle. Der Stand der letzten Aktualisierung der EGLV-Daten ist aus Abbildung 3 zu entnehmen. Weiterhin wird ersichtlich, dass diejenigen Kommunen, die im ersten Fall die größten Flächendefizite gezeigt haben (rote Prozentzahlen in Abbildung 15), mit der erweiterten Auswahl die größeren Flächenüberschüsse aufweisen. Daraus lässt sich im Allgemeinen die Aussage ableiten, dass die Verkehrsflächen in den Kommunen unterschiedlich vollständig durch die Objektart „Straßenverkehr“ dargestellt werden können. Durch die Auswahl weiterer ALKIS-Objektarten können die EGLV-Verkehrsflächen vollständig abgebildet werden. Darüber hinaus werden zusätzliche Verkehrsflächen erfasst, die im EGLV-Datensatz nicht vollständig enthalten sind.

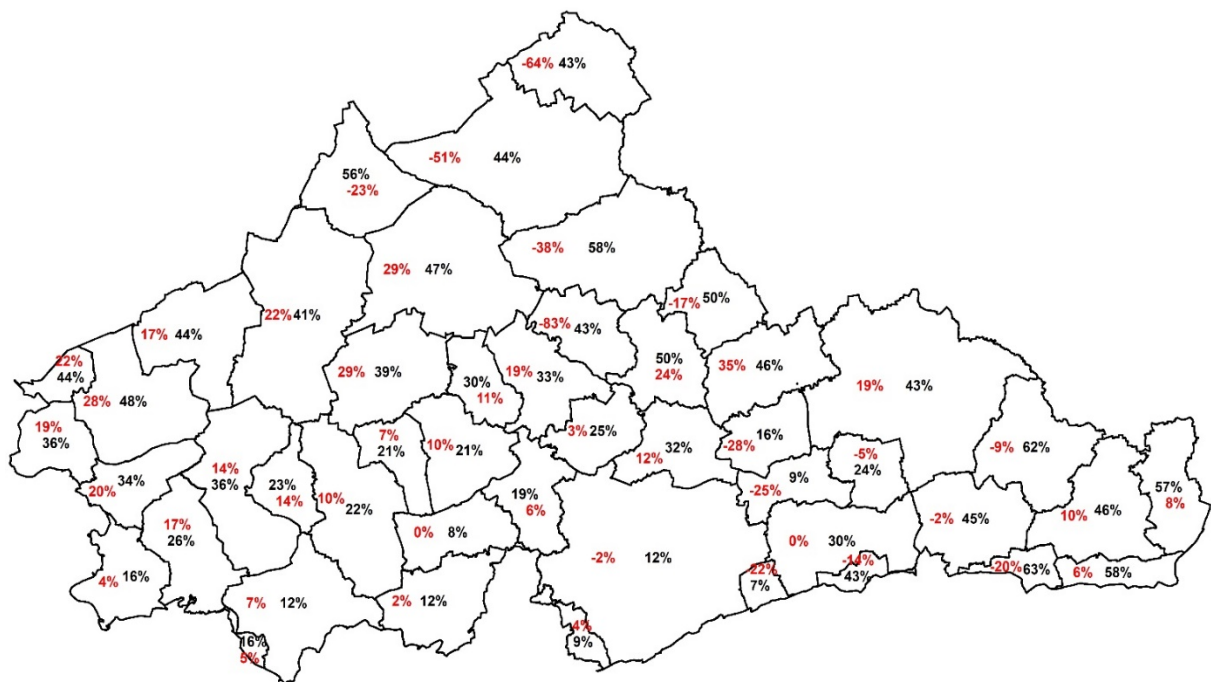




Abbildung 16: Erfassung von Verkehrsflächen durch ALKIS und EGLV (Stand 2008) in Bad Sassendorf (Stand 2017)

Bei einer detaillierteren Betrachtung der Verkehrsflächen muss jedoch eine Relativierung dieser allgemeinen Aussage getroffen werden. Einerseits können ALKIS-Daten in vielen Fällen tatsächlich zusätzliche Flächen erfassen, wie in Abbildung 16 zu sehen ist.

Andererseits ist zu berücksichtigen, dass ein Teil der durch ALKIS erfassten Verkehrsflächen im EGLV-Datensatz unter Zuwegungsflächen, insbesondere „Privat Schmutzige Flächen“ fallen (Parkplätze), erfasst wurde. Des Weiteren ist die Erfassungsgenauigkeit der ALKIS-Daten in den verschiedenen Kommunen sehr unterschiedlich. In Abbildung 17 sind Autobahnkreuze in unterschiedlichen Kommunen zu sehen. Durch die Unterscheidung der Wertearten z.B. in „Fahrbahn“ und „Begleitflächen“ ist die Trennung der befestigten Verkehrsflächen theoretisch gewährleistet (Abbildung 17 links unten: Dortmund). In der Praxis wird jedoch deutlich, dass die Handhabung bei der Erzeugung der ALKIS-Daten sehr unterschiedlich sein kann, wie in den weiteren Beispielen in Abbildung 17 zu sehen ist. Dabei wurden in diesen Fällen die Fahrbahnen sowie die Begleitflächen als „Fahrbahn“ kategorisiert. Diese Beispiele zeigen, dass Flächenüberschüsse im ALKIS-Datensatz nicht zwingend als zusätzlich erfasste Verkehrsflächen angenommen werden müssen, sondern auch durch eine grobe Erfassung verursacht werden können.



Abbildung 17: Erfassungsgenauigkeit der Verkehrsflächen durch ALKIS - Daten (Stand Dortmund 2018, Dorsten 2018, Recklinghausen 2018, Essen 2015) (EGLV Stand: Dortmund, Recklinghausen, Essen 2014, Dorsten 2010)

Zusammenfassend wird für die Verkehrsflächen festgestellt, dass anhand von ALKIS-Daten versiegelte Verkehrsflächen, die durch die EGLV erfasst werden, teilweise vollständig abgeleitet werden können. Mit ALKIS-Daten können zusätzliche Verkehrsflächen abgebildet werden. Diese sind abhängig von der Vollständigkeit bzw. Erfassungsgenauigkeit der Datengrundlage, die wiederum von der Handhabung der Katasterämter bei der Fortführung der ALKIS-Daten abhängen.

## 6.3 Abgeleitete Zuwegungsflächen aus ALKIS

### 6.3.1 Klassifizierungsmodell

Mit dem Ziel der Identifizierung von Zuwegungsflächen auf Flurstücken wurde mit Hilfe der Methode des Entscheidungsbaumes ein Prognosemodell erstellt, das aus 200 „Entscheidungsbäumen“ besteht. Hierbei handelt es sich um ein Klassifizierungsmodell, mit dessen Hilfe Flurstücke mit und ohne Zuwegungsflächen voneinander unterschieden werden sollen. Flurstücke werden in den Klassen „Zuwegung Wahr“ (TRUE) und „Zuwegung Falsch“ (FALSE) unterteilt. Für diese Betrachtung kommen ausschließlich metrische Regressoren zum Einsatz. In Abbildung 18 ist der Verlauf der Fehlerschätzung für den Generalisierungsfehler zu sehen. Dieser Wert wird auch als „OOB-Wert“ („out of bag“) bezeichnet. Die Schätzung für den Generalisierungsfehler ist eine Abschätzung für eine falsche

Klassifizierung von unbekannten Daten (z.B. Kalibrierungsdaten) unter Verwendung des trainierten Modells. Mit jedem weiteren „Entscheidungsbaum“ sinkt die Fehlerschätzung, bis sie sich auf dem Niveau von ca. 11 % stabilisiert. Für die Anwendung bedeutet der OOB-Wert, dass bei Klassifizierung von neuen unbekannten Daten ca. 11 % Flurstücke falsch klassifiziert werden. Im Umkehrschluss bedeutet dieser Wert, dass in 89 % der Fälle die Existenz einer Zuwegungsfläche auf einem Flurstück korrekt prognostiziert wird.

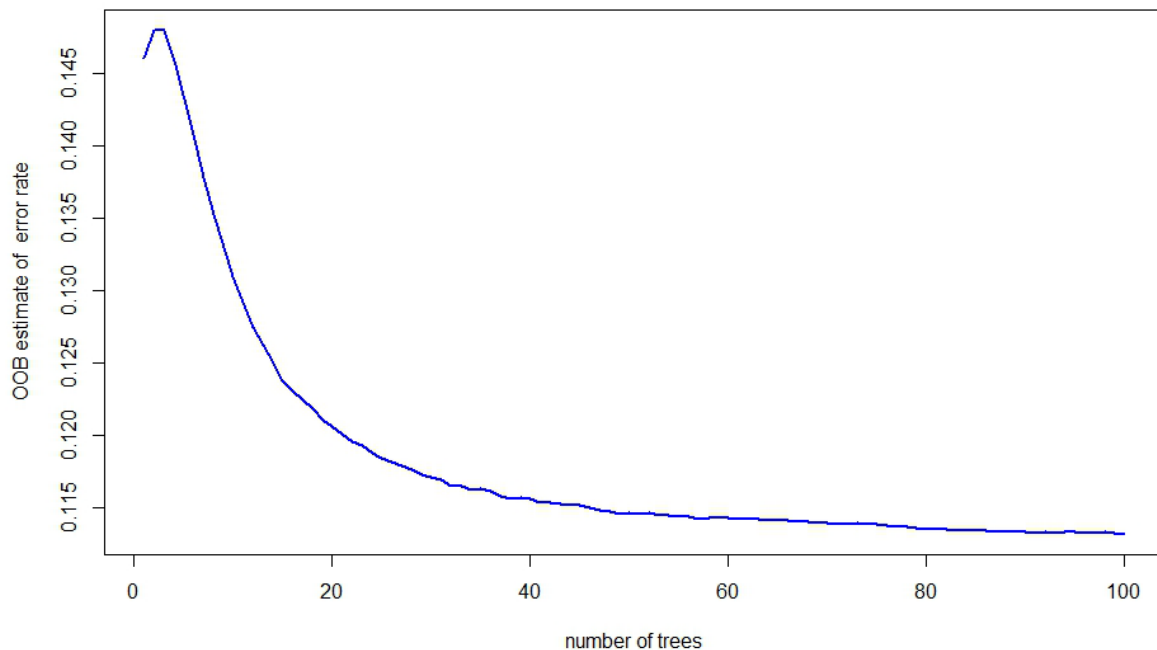


Abbildung 18: Fehlerschätzung für den Generalisierungsfehler

In Abbildung 19 ist ein Ranking der Regressoren nach ihrer Relevanz zu sehen, anhand derer die Entscheidung für die Klassifizierung gefällt wird. Dabei wird ermittelt, wie sehr ein Regressor den Gini-Index<sup>2</sup> über die gesamten Entscheidungspfade reduziert. D.h. es wird überprüft, inwieweit ein Regressor die Unreinheit der Entscheidungen reduziert und immer wieder die gleiche Entscheidung trifft. Die größten Einflüsse werden durch die Regressoren „Verhältnis der Gebäudefläche zur Flurstückfläche“ und „Gebäudefläche“ realisiert. Dieses Ergebnis zeigt, dass Gebäude eine besonders wichtige Rolle bei der Abschätzung von Zuwegungsflächen spielen. Die geringste Bedeutung für das Modell haben hingegen die Geländehöhe und die Anzahl der angrenzenden Flurstücke.

<sup>2</sup> Der Gini-Index ist ein Maß für die Ungleichverteilung (KOLANOSKI, 2008).

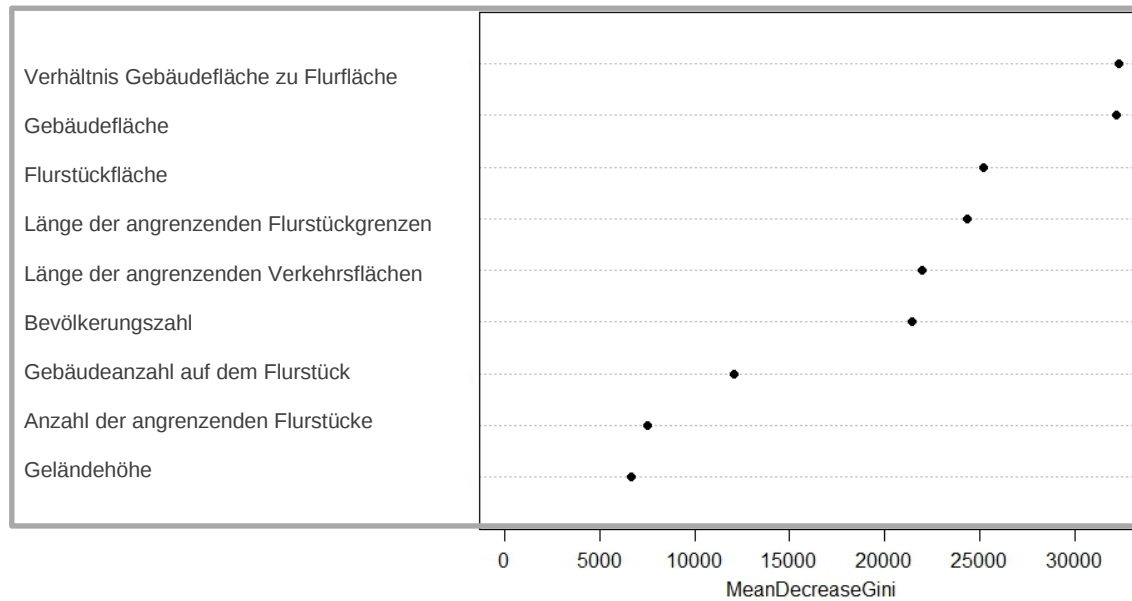


Abbildung 19: Ranking der Gewichtung der Regressoren

In Tabelle 7 sind die Ergebnisse der Prognose des erstellten Klassifizierungsmodells zu sehen. „FALSE“ steht für Flurstücke, die keine Zuwegung aufweisen und „TRUE“ steht für Flurstücke mit Zuwegungsflächen. Aus der ersten Zeile der Tabelle ist zu entnehmen, dass 3.723 Flurstücke, die tatsächlich keine Zuwegung aufweisen, auch als solche klassifiziert wurden. Bei 2.278 Flurstücken wurde fälschlicherweise eine Zuwegung ermittelt. In der zweiten Zeile werden bei 4.685 Flurstücken keine Zuwegungsflächen prognostiziert, obwohl sie in Wahrheit eine Zuwegungsfläche besitzen. Die restlichen 48.686 Flurstücke werden korrekt klassifiziert.

Tabelle 7: Prognose von Zuwegungen auf Flurstücke mit der Methode "Entscheidungsbäume"

|                    | Prognose<br>Referenz | FALSE | TRUE   | Vorhersagegenauigkeit [%] |
|--------------------|----------------------|-------|--------|---------------------------|
|                    |                      |       |        |                           |
| Kalibrierungsdaten | FALSE                | 3.723 | 2.278  | 88,27                     |
|                    | TRUE                 | 4.685 | 48.686 |                           |

Mit dem Modell konnte somit die Existenz von Zuwegungsflächen in 88,27 % der Fälle richtig ermittelt werden. Mit dem Ergebnis dieses Modells wird die Existenz von Zuwegungsflächen auf Flurstücken mit einer, aus Sicht der Statistik, hohen Sicherheit abgeschätzt. Bei dem vorliegenden Modell wurde ein einziges Modell für das gesamte Untersuchungsgebiet entwickelt. Wie bereits beschrieben, können durch die Identifizierung von Clustern homogene Flurstücke gruppiert werden, die sowohl bei der Klassifizierung als auch bei den Regressionsanalysen zu besseren Ergebnissen führen.

### 6.3.2 Clusteranalyse und Regressionsmodelle

Im Zuge der statistischen Analyse wurde das Einzugsgebiet des Hüller Bachs als Kalibrierungsgebiet ausgewählt. Das Einzugsgebiet erstreckt sich über Teile von Bochum, Herne und Gelsenkirchen (EGLV-Daten Stand 2014-2016). Dieses Gebiet kam für die Entwicklung der Regressionsmodelle nicht zum Einsatz. Somit handelt es sich hierbei um unbekannte Daten für die erstellten Modelle. Die Regressionsmodelle wurden anhand der Daten im restlichen Einzugsgebiet der EGLV erstellt. Die digitalisierten Zuwegungsflächen aus dem EGLV-Datensatz kommen als Kalibrierungsdaten zum Einsatz.

Anhand einer manuellen Analyse wurden insgesamt zehn Cluster identifiziert. Die unterschiedlichen Regressionsmodelle haben zum Ziel, für möglichst viele Flurstücke die Zuwegungsflächen präzise abzuschätzen. Als Gütekriterium ist zum einen der Determinationskoeffizient  $R^2$  des Regressionsmodells relevant. Der Determinationskoeffizient ist ein Maß für die Güte der Prognosewerte, die angibt, welcher Anteil der beobachteten Varianz durch das Regressionsmodell erklärt werden kann. Zum anderen ist die Übereinstimmung der Prognosewerte mit Kalibrierungsdaten für die Auswahl der Cluster und die Güte der resultierenden Regressionsmodelle maßgebend. Bei der Clusterbildung wurde darauf Wert gelegt, dass sowohl die Maßzahl  $R^2$  als auch die Übereinstimmung der Prognosedaten mit Referenzdaten einen maximalen Wert annehmen.

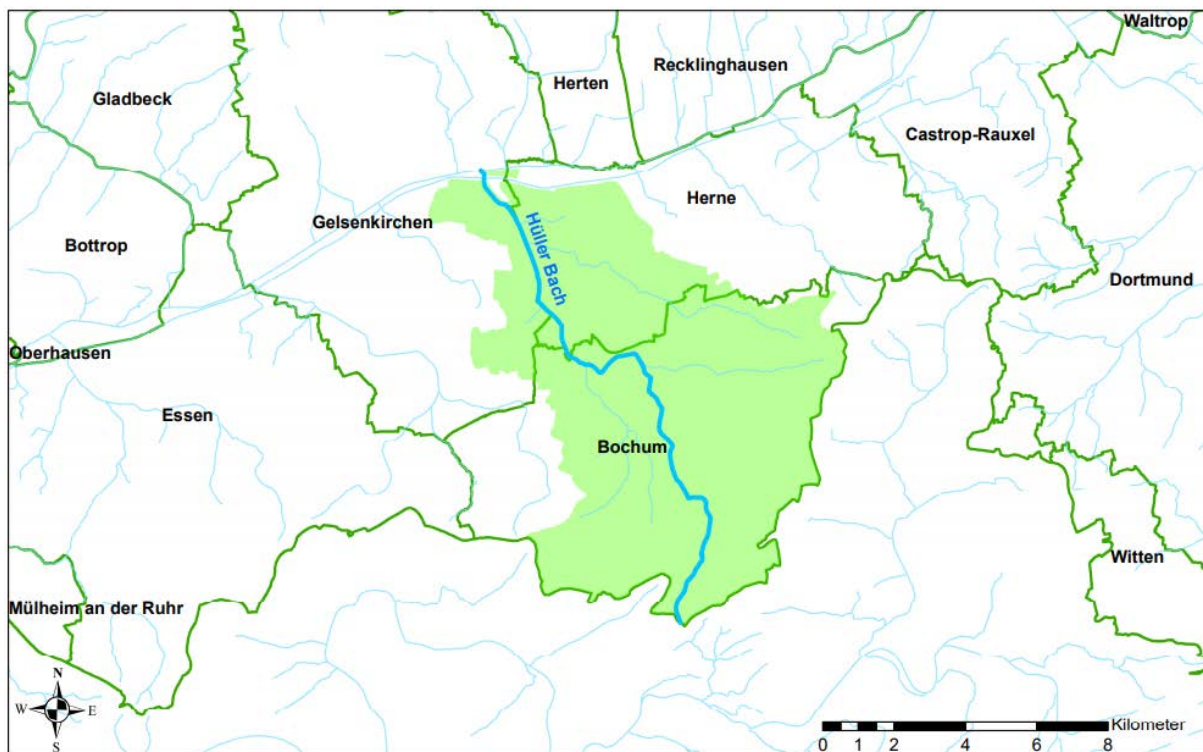


Abbildung 20: Einzugsgebiet des Hüller Bachs im Untersuchungsgebiet der EGLV

Tabelle 8: Ergebnisse der statistischen Abschätzung von Zuwegungsflächen und durchschnittliche Übereinstimmungen mit Kalibrierungsdaten

| Cluster                                                                                                                                                                                                       | Korrekte Klassifizierung Existenz von Zuwegungsflächen | Bestimmtheitsmaß des Regressionsmodells ( $R^2$ ) | Mittlere Übereinstimmung der realen Zuwegungsflächen mit Prognose |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------|---------------------------------------------------|-------------------------------------------------------------------|
| Flurstücke in Wohnbauflächen mit mindestens zwei Gebäuden                                                                                                                                                     | 0,97                                                   | 0,70                                              | 0,66                                                              |
| Flurstücke mit der maßgeblichen Gebäudefunktion „Garagen“ in Flächen besonderer funktionaler Prägung                                                                                                          | 0,91                                                   | 0,83                                              | 0,68                                                              |
| Flurstücke in Flächen Gemischter Nutzung und mindestens 3 Gebäuden                                                                                                                                            | 0,96                                                   | 0,73                                              | 0,63                                                              |
| Flurstücke in Flächen besonderer funktionaler Prägung mit mindestens 2 Gebäuden                                                                                                                               | 0,97                                                   | 0,81                                              | 0,59                                                              |
| Flurstücke in Industrie- und Gewerbeflächen mit mindestens 2 Gebäuden                                                                                                                                         | 0,96                                                   | 0,81                                              | 0,69                                                              |
| Flurstücke mit mindestens 1 Gebäude und einem Gebäude-Flurstück-Verhältnis zwischen 27% und 38%                                                                                                               | 0,95                                                   | 0,87                                              | 0,56                                                              |
| Flurstücke in Flächen Gemischter Nutzung mit 2 Gebäuden und mittlere bis hohe Bevölkerungsdichte (Flurstück befindet sich in einem Raster [100x100 m] mit einer Bevölkerungsanzahl größer 28 Personen )       | 0,93                                                   | 0,67                                              | 0,58                                                              |
| Flurstücke mit einer hohen Bevölkerungsdichte (Flurstück befindet sich in einem Raster [100x100 m] mit einer Bevölkerungsanzahl zwischen 47 - 70 Personen ) und einem Gebäude-Flurstück-Verhältnis größer 38% | 0,92                                                   | 0,70                                              | 0,58                                                              |
| Flurstücke mit einem und einem Gebäude-Flurstück-Verhältnis größer 38% und einer Flurstückfläche größer 90 m <sup>2</sup>                                                                                     | 0,90                                                   | 0,85                                              | 0,52                                                              |

In Tabelle 8 sind die Cluster und die Ergebnisse der Klassifizierungsmodelle bzw. Regressionsmodelle dargestellt (vollständige Tabelle in Anhang 2). Für ca. 70 % der Flurstücke im Kalibrierungsgebiet konnten Regressionsmodelle mit mittlerer und hoher Modellgüte erstellt werden. In dunkelgrüner Farbe sind fünf Cluster dargestellt, bei denen die Prognose über die Existenz von Zuwegungsflächen zwischen 91 % und 97 % liegt. Das Bestimmtheitsmaß dieser Cluster ist mit über 0,7 hoch, d.h. es ist eine hohe Varianzerklärung gegeben. Die mittlere Übereinstimmung der abgeschätzten Zuwegungsflächen und Zuwegungsflächen aus dem Einzugsgebiet des Hüller Bachs liegt im Mittel zwischen ca. 59 % bis 69 %. Die hellgrün hinterlegten Cluster zeigen analog zu den dunkelgrün hinterlegten Cluster hohe Bestimmtheitsmaße ( $R^2$  zwischen 0,67 und 87). Die Übereinstimmung der abgeschätzten Zuwegungsflächen mit den Referenzdaten fallen jedoch im Mittel geringer (52 % bis 58 %) aus. Für die restlichen Flurstücke ohne Gebäude (20 %) und Flurstücke mit sehr inhomogenen Strukturen (10 %) konnten keine Regressionsmodelle erzeugt werden. Die Zuteilung der Flurstücke im Einzugsgebiet des Hüller Bachs zu den Clustern mit mittlerer und hoher Modellgüte bzw. Flurstücke ohne Regressionsmodelle ist in

Abbildung 21 dargestellt. Die grünen und gelben Flurstücke zeigen in der linken Grafik Flurstücke, die einen der neun Cluster aus zugeordnet werden können. Diese machen ca. 70 % der gesamten Flurstücke in diesem Gebiet aus. Dabei sehen die Flächen in der linken und rechten Abbildung visuell gleich aus. Das liegt daran, dass Flurstücke ohne Regressionsmodelle (rot) meist sehr große Flächen darstellen, unter anderem „Sport-, Freizeit und Erholungsflächen“ oder „Friedhöfe“. Anhand der Definition der Cluster kann auf diese Weise vor der Erzeugung der Prognosewerte bereits bestimmt werden, für welche Flurstücke eine statistische Aussage möglich ist.

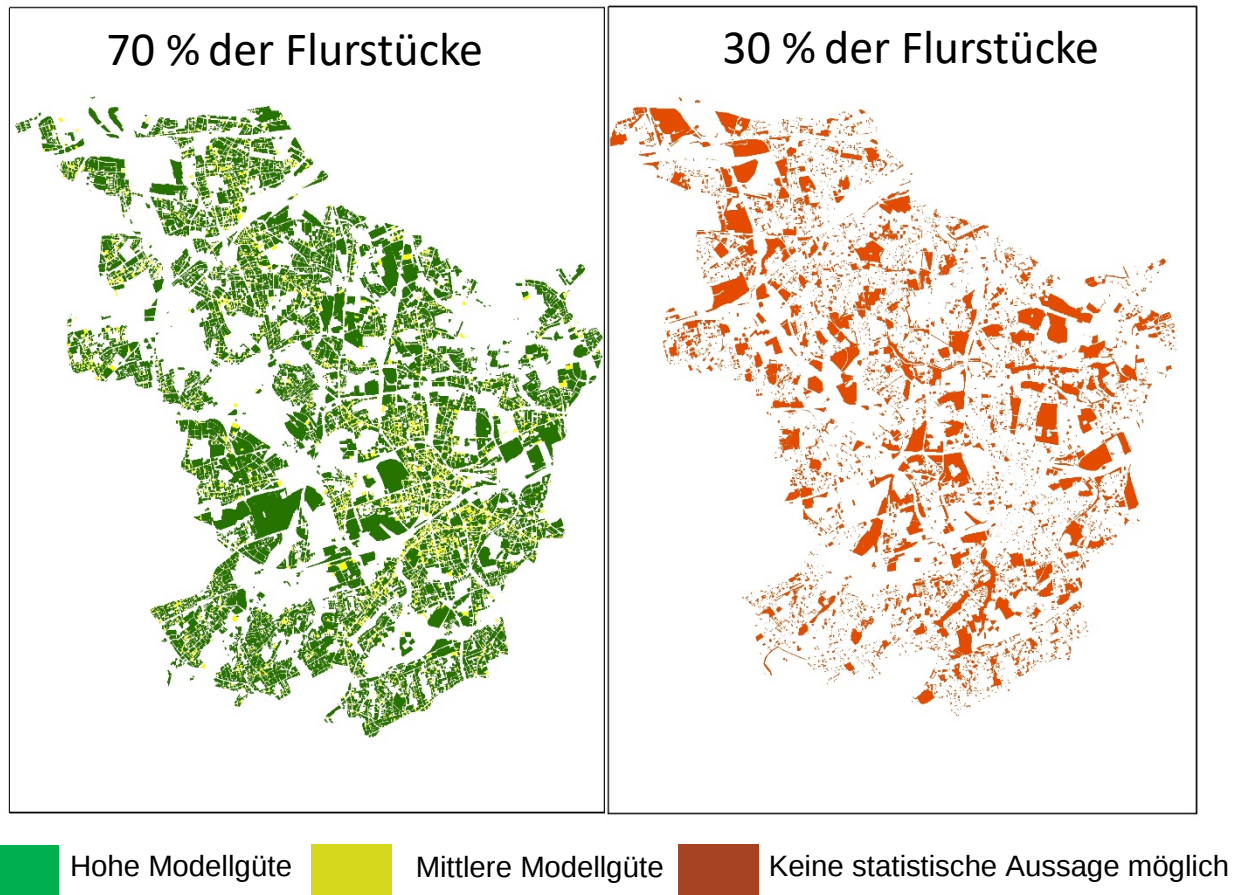


Abbildung 21: Zuordnung von Flurstücken zu Clustern mit unterschiedlicher Güte der Regressionsmodelle

Prinzipiell können die folgenden Flurstücke bei der statistischen Ermittlung von Zuwegungsflächen ausgeschlossen werden:

- Flurstücke kleiner als 15 m<sup>2</sup>
- Flurstücke mit einer größeren Gebäudeüberdeckung von 95 %
- Flurstücke, die sich in folgenden Flächenkategorien befinden: „Sport- Freizeit und Erholungsflächen“, „Friedhöfen“, „Halden“, „Bergbaubetriebe“ sowie „Tagebau, Grube und Steinbruch“ und
- Flurstücke ohne Gebäude

Alle weiteren Flurstücke werden erst ausgeschlossen, wenn sie keinem anderen Cluster zugeordnet werden können. Um die Güte der Prognosemodelle zu überprüfen, wurde das Einzugsgebiet Hüller Bach in 163 Teilgebiete unterteilt, die zu 20 Mischwasserbehandlungsanlagen gehören (Abbildung 22). Für die 163 (Teil-)Gebiete und die aggregierten 20 Einzugsgebiete von Mischwasserbehandlungsanlagen wurden die Zuwegungsflächen aus dem EGLV – Datensatz und den Regressionsmodellen ermittelt und verglichen. Rot hinterlegte Einzugsgebiete zeigen dabei eine

größere Abweichung als 20 %. Es ist festzustellen, dass je kleinteiliger die Einzugsgebiete werden, umso öfter nehmen die Abweichungen zu. Die Unterschiede gleichen sich bereits aus, wenn die Flächen auf die Ebene der 20 Mischwasserbehandlungsanlagen zusammengefasst werden.

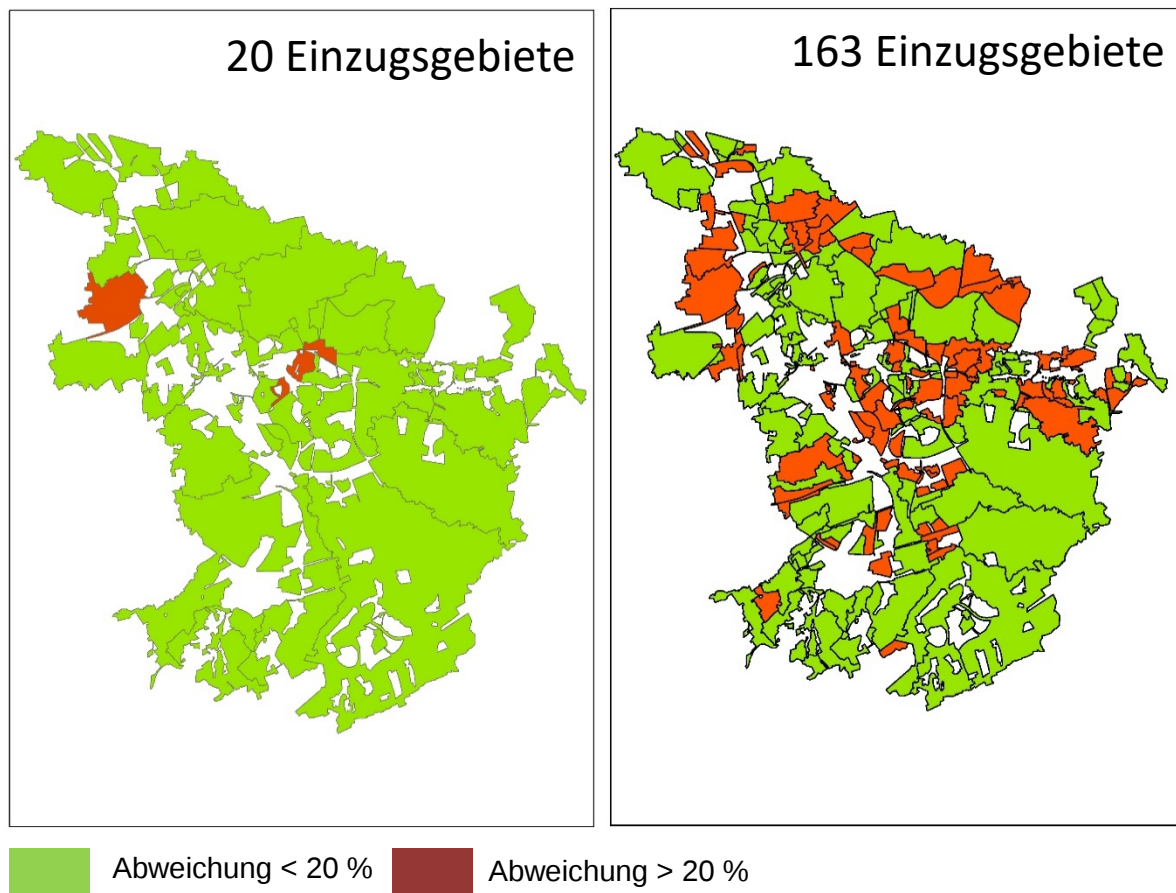


Abbildung 22: Ausgewählte Teileinzugsgebiete mit Abweichungen unter und über 20 % zwischen EGLV-Zuwegungsflächen und prognostizierten Zuwegungsflächen

Damit geht eine flächenmäßige Größenordnung für die Anwendung der Zuwegungsflächen einher. Das Einzugsgebiet des Hüller Bachs diente als Kalibrierungsdatensatz für die Anpassung der Cluster und die Erzeugung der Regressionsmodelle.

### 6.3.3 Validierungsergebnisse

Für die Validierung der Ergebnisse wurde ein Gebiet, bestehend aus 350 Flurstücken, in der Städteregion Aachen manuell digitalisiert. Hierbei wurde die Erfassung der Zuwegungsflächen anhand der Leistungsbeschreibung der EGLV durchgeführt, um die Vergleichbarkeit zu ermöglichen. Abbildung 23 zeigt den digitalisierten Bereich. Rechts sind mehrheitlich grün gefärbte Flurstücke aber auch einige farblose Flurstücke abgebildet. Bei den grün dargestellten Flurstücken handelt es sich um Flurstücke, die Clustern mit einer hohen Modellgüte zugeordnet werden. Diese sind hauptsächlich „Wohnbauflächen“ aber auch „Flächen gemischter Nutzung“ und „Flächen besonderer funktionaler Prägung“ sowie „Industrie- und Gewerbeflächen“. Die farblosen Flurstücke sind diejenigen, die aufgrund

von Plausibilisierungsmaßnahmen nicht betrachtet werden, z.B. Flurstücke, die vollständig von Gebäuden überlagert werden oder Clustern mit einer mittleren Güte bzw. keinem Cluster zugeordnet werden können. Die Prognose der Zuwegungsflächen wird von zusätzlichen Randbedingungen unterstützt, um möglichst realistische Prognosewerte zu generieren.



Abbildung 23: Digitalisierte Flurstücke in der Städtereion zur Validierung der Regressionsmodelle (Stand 2016)

Werden beispielsweise Zuwegungsflächen generiert, die größer sind als die effektiven Flächen (Flurstückfläche abzüglich der Gebäudeflächen) auf einem Flurstück, wird der Wert verworfen und der mittlere prozentuale Anteil an Zuwegungsflächen der Trainingsdaten übernommen. Dazu folgendes Beispiel: Für das erste Cluster, d.h. Flurstücke die sich auf „Wohnbauflächen“ befinden und die maßgebliche Gebäudefunktion „Wohnhaus“ haben, wird der durchschnittliche Anteil der Zuwegungsflächen berechnet. Dieser beträgt 26,9 % der effektiven Fläche (vgl. Anhang 2). Wird für ein Flurstück innerhalb dieses Clusters eine Zuwegungsfläche berechnet, die größer ist als die zur Verfügung stehende Fläche, wird die effektive Fläche mit dem durchschnittlichen Wert von 26,9 % multipliziert und mit dem Prognosewert ersetzt. Auf diese Weise wird sichergestellt, dass die Summe der Gebäudeflächen und Zuwegungsflächen maximal die Größe der Flurstückflächen einnehmen kann. In Abbildung 24 sind die Ergebnisse der Modellrechnungen im Vergleich zu den manuell digitalisierten Zuwegungsflächen dargestellt. Die roten Flurstücke mit 0 % Übereinstimmung zeigen die Fälle einer falschen Klassifizierung durch das „random forest“ Modell. Hierbei handelt es sich um Flurstücke, bei denen die Existenz von Zuwegungsflächen falsch ermittelt wurde. Dies ist bei ca. 10 % der 350 Flurstücke der Fall. Dieses Ergebnis bestätigt die in Kapitel 6.3.1 ermittelte Fehlereinschätzung für das Klassifikationsmodell. Für die restlichen Flurstücke wurden anhand der Regressionsmodelle bezüglich der Eigenschaften der Flurstücke statistisch Zuwegungsflächen ermittelt. Bei einer genaueren Betrachtung der Flurstücke wird festgestellt, dass bei homogenen Strukturen (blau gestreifte Umrandung) besonders hohe Übereinstimmungen (80 % - 100 %) zustande kommen.

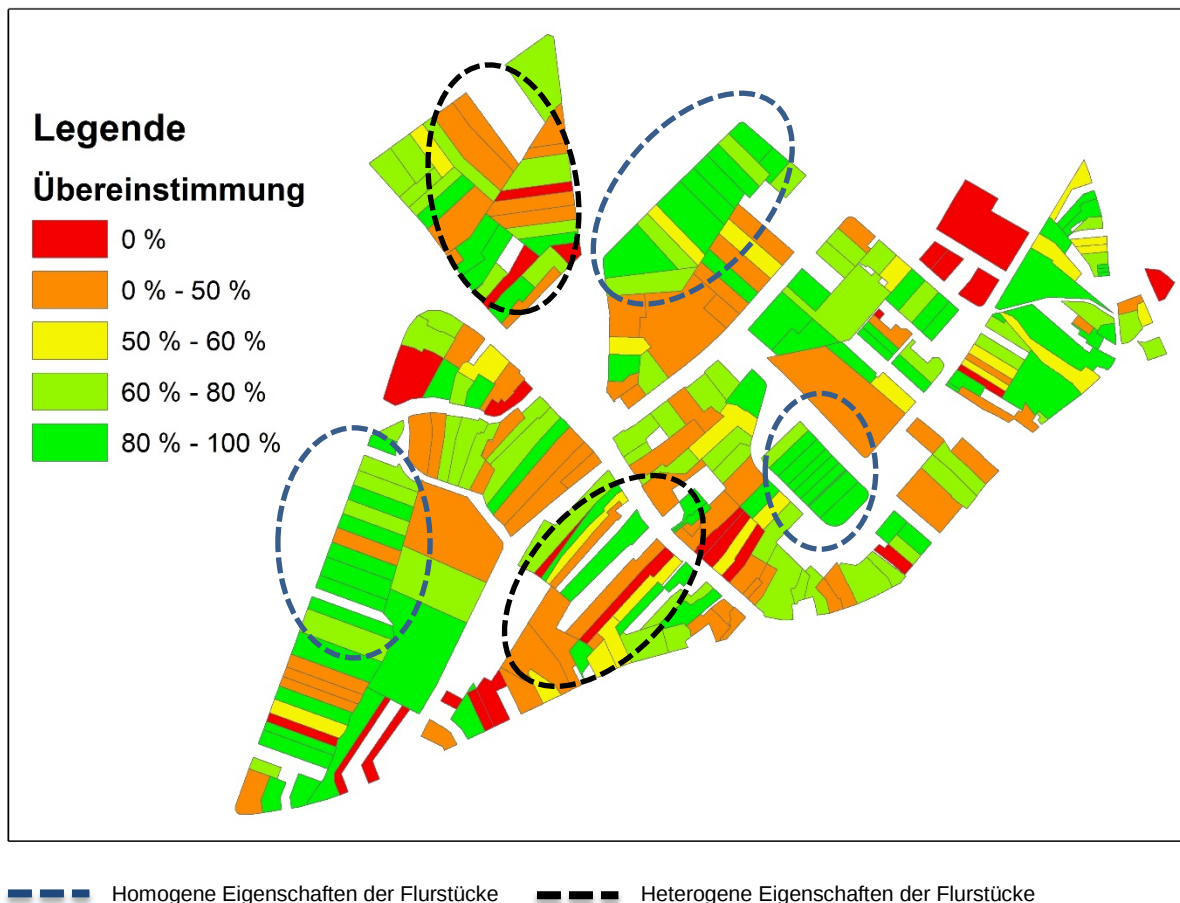


Abbildung 24: Prozentuale Übereinstimmung der Prognosewerte für Zuwegungsflächen mit manuell digitalisierten Zuwegungsflächen aus der Städteregion

Im Gegensatz dazu nehmen Flurstücke mit heterogenen Strukturen (schwarz gestreifte Umrandung) geringe Übereinstimmungswerte an. Die durchschnittliche Übereinstimmung zwischen Prognosewerten und manuell digitalisierten Zuwegungsflächen beträgt 60 % und befindet sich somit im Erwartungsbereich, welcher bei der Gegenüberstellung der Prognoseergebnisse im Einzugsgebiet (EZ) des Hüller Bachs festgestellt wurde. Damit ist die Übertragbarkeit der Prognosemodelle, die anhand der Trainingsdaten erzeugt und anhand von Kalibrierungsdaten aus dem Einzugsgebiet des Hüller Bachs optimiert wurde, auf beliebige Raumeinheiten in NRW vorläufig sichergestellt. Nichtsdestotrotz sind weitergehende Analysen anhand weiterer Validierungsdaten notwendig, um die Modelle weiter zu optimieren und höhere Übereinstimmungswerte zu erreichen. Für eine fundierte Beurteilung der statistisch ermittelten Zuwegungsflächen müssen Erfahrungen mit den Daten gesammelt werden. Hierfür werden Prognosedaten für das gesamte EGLV-Gebiet zur Verfügung gestellt, die innerhalb einer Evaluierungsphase von ca. einem Jahr durch Emschergenossenschaft und Lippeverband in sechs Schmutzfrachtmodellen verschiedener Regionen getestet werden, um die Auswirkungen der Abweichungen für unterschiedliche räumliche Einheiten unter Berücksichtigung von Niederschlag und Morphologie zu überprüfen.

## 7 Zusammenfassung

Befestigte Flächen werden in vielen für den Umweltschutz, für die Stadt- und Landschaftsplanung und die Wasserwirtschaft zuständigen Behörden sowie von entsprechenden Fachplanern regelmäßig genutzt. Wesentlicher Hintergrund in diesem Zusammenhang ist, dass befestigte Flächen immer eine Reduzierung der Versickerungsfläche für Niederschläge und somit eine Veränderung des Wasserhaushalts durch ein beschleunigtes Abflussverhalten dieser Flächen zur Folge haben. Daher sind sie eine wichtige wasserwirtschaftliche Basisinformation. Zudem ist der von den befestigten und unterschiedlich genutzten Flächen in die Kanalisation abfließende Niederschlag durch Reifenabrieb, Feinstaub und andere Mikro- und Markopartikel verunreinigt und belastet die Gewässer.

Im Allgemeinen können befestigte Flächen in Gebäude- und Dachflächen, Verkehrsflächen sowie Zuwegungsflächen unterteilt werden, die jeweils ca. 1/3 der befestigten Flächen ausmachen. Im Projekt „Wasserwirtschaftliche Ermittlungen und Anwendung von digitalen Versiegelungsdaten aus Amtlichen Liegenschaftskatasterdaten“ (WaLKIS) wurden die theoretischen Grundlagen und Analysemethoden zur Bestimmung von befestigten Gebäude- und Verkehrsflächen aus dem Liegenschaftskataster-Modell ALKIS sowie eine flächendeckende Analyse zur statistischen Abschätzung der befestigten Zuwegungsflächen aus den ALKIS-Daten und die Verifikation der Ergebnisse erarbeitet.

Zusammenfassend wird festgestellt, dass Gebäudeflächen anhand von ALKIS-Daten erfasst werden können und mit Hilfe der Korrekturwerte aus Kapitel „Korrekturwerte für ALKIS Gebäudepolygone“ eine gute bis sehr gute Übereinstimmung mit vorhandenen Referenzdaten auf Grundlage von digitalisierten Befliegungsdaten aufzeigen. Referenzdaten für die Analysen wurden von der EGLV zu Verfügung gestellt. Gebäudepolygone wurden in diesem Zusammenhang zunächst objektorientiert betrachtet (vgl. Kapitel „Objektorientierte Gegenüberstellung von Gebäudepolygone aus ALKIS und Referenzdaten“), um eine Aussage über die Lagegenauigkeit sowie Flächengleichheit im Vergleich zu den Referenzdaten zu ermöglichen. Diese Untersuchung ergab, dass bei der Betrachtung der einzelnen Gebäudepaare in ca. dreiviertel aller Fälle eine höhere Übereinstimmung als 80 % zwischen ALKIS-Daten und aus Befliegungsdaten digitalisierte Referenzdaten der EGLV existiert. Etwaige Abweichungen sind insbesondere durch Dachüberhänge, die in ALKIS nicht erfasst werden, aber auch aufgrund der unterschiedlichen Stände der Daten, zu erklären. Mit einer zweiten Methode zur Untersuchung der Gebäudedaten wurden Gebäudepolygone auf Flurstücke aggregiert. Eine Besonderheit ist hierbei die große Anzahl an Flurstücken, die beinahe vollständig von Gebäuden überlagert werden. Deshalb existiert eine sehr hohe Korrelation ( $r = 0,99$ ) zwischen ALKIS-Daten und aus Befliegungsdaten abgeleitete Referenzdaten der EGLV. Die Übereinstimmung der Gebäudeflächen fällt hingegen in Summe ähnlich wie bei der objektorientierten Methode aus und erreicht in ca. 75 % der Fälle eine Übereinstimmung von mehr als 80 % zwischen ALKIS-Daten und aus Befliegungsdaten ermittelte Gebäudeflächen. Durch die Anpassung der Gebäudeflächen mithilfe von Korrekturwerten werden deutlich höhere Übereinstimmungswerte erreicht. Mit dieser im Projekt entwickelten Methodik können mittels objektorientierter Auswahl aus ALKIS-Daten für ganz NRW flächendeckend Gebäudeflächen ermittelt werden.

Gleiches gilt für die Verkehrsflächen, wobei die Erfassung der Verkehrsflächen in ALKIS durch Kommunen sehr unterschiedlich gehandhabt wird. Insbesondere bei der Erfassung der Verkehrsbegleitflächen sowie der Objekte, die Verkehrsflächen in Siedlungen darstellen, weicht die Vollständigkeit und Genauigkeit der Erfassung stark voneinander ab. Zudem wurde der Bahnverkehr nur theoretisch betrachtet, da keine Referenzdaten für diesen Verkehrsbereich existieren.

Zuwegungsflächen können durch eine gezielte Auswahl von ALKIS-Objektarten nicht erfasst werden. Hierbei handelt es sich beispielsweise um Garagenzufahrten, befahrbare Wege auf Privatgrundstücken, Park- und Lagerplätze sowie Terrassen an Wohnhäusern. Zur Identifizierung von Zuwegungen auf Flurstücken wurde mit Hilfe eines Klassifizierungsmodells basierend auf der Methode des Entscheidungsbaumes ein Prognosemodell erstellt (siehe Kap. 7.3.1). Hierbei werden Flurstücke mit und ohne Zuwegungsflächen differenziert. Für dieses Modell wurde für das Validierungsgebiet in der Städteregion Aachen eine Übereinstimmungsquote der Flächenanteile von 90 % erreicht. Anschließend wurden neun Cluster (Gruppierungen) identifiziert und für jedes Cluster Regressionsmodelle generiert, die eine Abschätzung der Zuwegungsflächen ermöglichen. Für ca. 75 % der Daten konnten Regressionsmodelle mit mittlerer und hoher Modellgüte entwickelt werden. Dabei wurden nach der Kalibrierung für Cluster mit mittlerer Modellgüte durchschnittliche Übereinstimmungen zwischen 50 % und 60 % erreicht. Bei Clustern mit hoher Modellgüte schwankte dieser Wert zwischen 60 % und 70 %. Die Anwendung der Regressionsmodelle auf Validierungsdaten der Städteregion Aachen bestätigen die Ergebnisse aus der Kalibrierungsstufe. Im Rahmen dieses Projektes wurde erstmals gezeigt, dass anhand mathematischer Modelle eine statistische Abschätzung von Zuwegungsflächen mittels ALKIS-Daten möglich ist. Die Ergebnisse der statistischen Analyse sind für eine flächendeckende Anwendung für Raumeinheiten, beispielsweise Einzugsgebiete mittelgroßer Regenbecken oder vergleichbarer technischer Anlagen, geeignet. Die entwickelten Methoden und gewonnenen Erkenntnisse stellen eine bedeutsame, wissenschaftlich überprüfbare und reproduzierbare Grundlage für weiterführende Analysen und Anwendungen in der Wasserwirtschaft dar.

Die Ergebnisse sind im zuständigen Lenkungskreis, bestehend aus Wasserverbänden, unteren Wasserbehörden sowie Vertretern von Kommunen und Kreisen sowie des LANUV, auf große Zustimmung gestoßen und in ersten Praxisanwendungen überprüft worden. Im Rahmen der regelmäßigen Abstimmungsgespräche wurden weitere Optimierungsmöglichkeiten als Ausblick für eine mögliche zweite Phase des Projektes identifiziert.

## 8 Ausblick

Die im Rahmen des Projektes geschaffene wissenschaftlich abgesicherte, dokumentierte und reproduzierbare Methode zur Erfassung von versiegelten Flächen aus dem Amtlichen Liegenschaftskatastersystem liefert leider keinerlei Informationen über die Flächenbeschaffenheit und die Flächenabflusswirksamkeit. Sie weist daher noch keine weitreichenden Anwendungsmöglichkeiten auf. Aus diesem Grunde sind weiterführende Analysen angestrebt, die unter Anwendung weiterer Datenquellen sowie der Entwicklung weiterer GIS-Algorithmen Aussagen über die Belastung der Niederschlagsabflüsse sowie ihre Abflusswirksamkeit ermöglichen. Die Algorithmen dienen als Basis für die detaillierte Beschreibung der Methodik. Die entsprechenden Quellcodes werden in einem Begleitprozess an IT.NRW übergeben.

Die Ziele einer weitergehenden Analyse sind unter anderem die vollständige Automatisierung der Erfassungsmethode aus Phase I sowie die Identifizierung und Integration landesweiter zur Verfügung stehender Daten. Außerdem sollen die Prognosemodelle zur statistischen Abschätzung der Zuwegungsflächen optimiert werden, dies gilt der besseren Beurteilung der Beschaffenheit der Niederschlagsabflüsse aber auch der Ermittlung der Belastung der Niederschlagsabflüsse. Des Weiteren soll es möglich werden, die Abflusswirksamkeit der identifizierten befestigten Flächen zu beurteilen.

### **Landesweite, automatisierte Clusterbildung**

Mit der Generierung von Clustern (Gruppierung der Daten mit vergleichbaren statistischen Eigenschaften) wurde eine bessere Anpassung der Regressionsmodelle für die Abschätzung der Zuwegungsflächen ermöglicht. Diese Clusterbildung soll nun automatisiert werden. Hierfür wird ein Algorithmus entwickelt, der in der Lage ist, in Abhängigkeit der Modellgüte ( $R^2$ ) und Kalibrierungsergebnisse für die größtmögliche Menge der Daten die bestmögliche Annäherung der Regressionsmodelle durch die Wahl von geeigneten Clustern zu ermitteln.

### **Identifizierung und Integration von flächendeckend verfügbaren Daten aus NRW**

Weitere flächendeckend verfügbare Daten sollen identifiziert, auf ihre Güte überprüft und in Untersuchungen integriert werden. Diese Daten dienen beispielsweise der Verbesserung der Regressionsmodelle, der Erweiterung der Cluster und der besseren Abschätzung der Belastungskategorien.

### **Differenzierung der befestigten Flächen in Belastungskategorien nach DWA - A 102**

Die ermittelten befestigten Flächen aus diesem Projekt liefern eine geeignete Grundlage für die Differenzierung dieser Flächen nach den festgelegten drei Belastungskategorien (gering, mäßig, stark belastet). Eine weitere Grundlage zur Abschätzung der Abflussbelastung bilden die von EGLV zur Verfügung gestellten digitalisierten Flächenpolygone, die anhand von Luftbildern generiert wurden. In diesem Datensatz werden die Zuwegungsflächen in „Privat Saubere“ (verursachen nichtbehandlungsbedürftigen Niederschlagsabfluss) und „Privat Schmutzige“ (verursachen behandlungsbedürftigen Niederschlagsabfluss) Flächen unterteilt. Anhand von Bewertungsindikatoren

aus dem Arbeitsblatt A 102 (Teil A) und statistischen Analysen werden die ermittelten befestigten Flächen in die oben erwähnten drei Belastungskategorien unterteilt.

### **Abschätzung der Abflusswirksamkeit der ermittelten befestigten Flächen aus Phase I**

Die ermittelten befestigten Flächen geben weder einen Hinweis auf die Abflusswirksamkeit noch auf die Lage dieser (Zuwegungs-)Flächen. Insbesondere bei den statistisch ermittelten Zuwegungsflächen handelt es sich um den Anteil der befestigten Flächen auf Flurstücken. Mit Hilfe der Eigenschaften der Flurstücke, aber auch die Erschließung durch Verkehrsflächen, ist eine Beurteilung über die Abflusswirksamkeit der befestigten Flächen möglich. Durch Ortsbegehungen können die ermittelten Zustände validiert werden. Zudem können unter Verwendung von Geländemodellen mögliche Fließwege und die Kumulierung von Niederschlagsabflüssen ermittelt werden. Anhand der ermittelten Belastungskategorien erfolgt anschließend eine flächendeckende Abschätzung der zu erwartenden belasteten Niederschlagsabflüsse.

### **Begleitung der Umsetzung bei IT.NRW**

Die Umsetzung der entwickelten Methodik soll durch IT.NRW auf der Basis des F&E-Berichtes und des Programm-Codes erfolgen. Die Kriterien zur Auswahl einer geeigneten EDV-Lösung sind bereits durch die Beteiligung von IT.NRW im Verlauf des Projektes berücksichtigt worden und sollen gemeinsam mit dem FiW differenziert und festgelegt werden. Schon heute zu nennende Eigenschaften sind die Performance und langfristige Unterhaltung des Systems (Versionskompatibilität). Bei der Entwicklung der EDV-Lösung müssen die Anforderungen der potenziellen Nutzer der Versiegelungsdaten Berücksichtigung finden. Die Daten sollen sowohl tabellarisch als auch georeferenziert für unterschiedliche räumliche Einheiten abgerufen werden können. Ein weiterer Aspekt ist die Bereitstellung der Daten für den Nutzer. Eine Möglichkeit stellt das Fachinformationssystem ELWAS dar.

## 9 Literaturverzeichnis

ADAMS, J. B.; MILTON, O. S. AND GILLESPIE, A. R. (1989): Simple Models for Complex Natural Surfaces: A Strategy for the Hyperspectral Era of Remote Sensing. pages 16–21.

ADV (2008A): ALKIS Objektartenkatalog. In: Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland. page 516.

ADV (2008B): Hauptdokument Version 6.0. In: Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland. page 516.

ADV (2018): Amtliches Liegenschaftskatasterinformationssystem (ALKIS®) In: Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland. Available at: <http://www.adv-online.de/AAA-Modell/ALKIS/>.

BACHOFER, F.; ESCH, T. AND KLEIN, D. (2009): Ableitung von versiegelungsgraden basierend auf hochaufgelösten fernerkundungsdaten mittels support vector machines (Derivation of soil sealing information from high resolution remote sensing data by application of support vector machines). *Angewandte Geo.*

BECK, R.; BECKER, M.; BORCHARDT, D.; BÜRGEL, B.; FUCHS, S.; HALLER, B.; JOSWIG, K.; MERTSCH, V.; PODRAZA, P.; REMMLER, F.; SCHMITT, T. G.; SCHWINGER, H.; UHL, M.; WWEIß, G. AND WELKER, A. (2016): DWA- A 102. In: DWA-Regelwerk.

BENN, O.; GEIBEL, P.; HUDIK, M.; KNEBEL, T. AND WYSOTZKI, F. (2005): Lernen von Entscheidungsbäumen bei objektabhängigen Fehlklassifikationskosten. In: Technische Universität Berlin.

BEZIRKSREGIERUNG KÖLN GEOBASIS NRW: WMS NW DOP-Übersicht – Metadaten DOP. URL: <https://www.tim-online.nrw.de/tim-online2/uebersicht.html?thema=dop> [Stand: 2019] (abgerufen am 27.03.2019).

BLUME, H. (1992): Handbuch des Bodenschutzes. In: ecomed Verlagsgesellschaft mbH.

BORCHARDT, D.; BÜRGEL, B.; DURCHSCHLAG, A.; GROTTKER, M.; HALLE, M.; PODRAZA, P.; SCHITTHELM, D. AND UHL, M. (2007): Merkblatt Merkblatt 3 / BWK. In: Fraunhofer IRB Verlag. Issue: November.

EG/LV (2012): Leistungsbeschreibung für die ArcGIS konforme Erfassung der befestigten Flächen  
Leistungsbeschreibung für die ArcGIS konforme Erfassung der befestigten Flächen. pages 1–8.

EMSCHERGENOSSENSCHAFT (2017): Daten und Fakten - Geschäftsbrief Emschergenossenschaft 2016/17. Issue: 201. pages 24–25. Available at: [http://www.eglv.de/fileadmin/user\\_upload/Daten\\_EG\\_dt\\_Oktober\\_2016\\_web.pdf](http://www.eglv.de/fileadmin/user_upload/Daten_EG_dt_Oktober_2016_web.pdf).

ESCH, T.; SCHORCHT, G. AND THIEL, M. (2017): Satellitengestützte Erfassung der Bodenversiegelung in Bayern. In: Bayerisches Landesamt für Umwelt (LfU).

FISCHER-STABEL, P.; SCHÄFER, K. AND WANNEMACHER, S. (2004): Satellitenbasierte Erfassung von Versiegelungsflächen – Entwicklung eines operationellen Auswerteverfahrens. In: Ministerium für Umwelt und Forsten Rheinland-Pfalz.

HOLTMANN, D. (2010): Grundlegende multivariate Modelle der sozialwissenschaftlichen Datenanalyse. - 3., veränd. Aufl.. Available at: <https://publishup.uni-potsdam.de/opus4-ubp/frontdoor/index/index/docId/4412>.

KOLANOSKI, H. (2008): Statistische Methoden der Datenanalyse. In: Humboldt-Universität zu Berlin. page 165.

KRESSLER, F. AND STEINNOCHER, K. (1995): Einsatz der Fernerkundung zur Überprüfung der Realnutzung im Raum Wien mittels Spectral Mixture Analysis. In: AGIT.

LIPPEVERBAND (2017): Daten und Fakten - Geschäftsbrief Lippeverband 2016/17.. Issue: 201. Available at: [http://www.eglv.de/fileadmin/user\\_upload/Daten\\_LV\\_dt\\_2016\\_Oktober\\_2016\\_web.pdf](http://www.eglv.de/fileadmin/user_upload/Daten_LV_dt_2016_Oktober_2016_web.pdf).

MATHWORKS (NO DATE): Machine Learning - Drei Dinge, die Sie wissen sollten. Available at: <https://de.mathworks.com/discovery/machine-learning.html> (Accessed: 20 March 2018).

MEINEL, G. AND HERNIG, A. (2005): Inwertsetzung von Geobasisdaten für Fachanwendungen am

Beispiel der Bodenversiegelungserhebung auf Grundlage des ATKIS Basis-DLM. In: Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation, . page 14,25.

MINISTERIUM FÜR WIRTSCHAFT, INNOVATION, DIGITALISIERUNG UND ENERGIE DES LANDES NORDRHEIN-WESTFALEN: OpenGeodaten.NRW. Digitale Orthophotos – Einzelkacheln. URL: <https://www.opengeodata.nrw.de/produkte/geobasis/dop/dop/> [Stand: 2017 - 2019] (abgerufen am 12.02.2019).

MINISTERIUM FÜR UMWELT, LANDWIRTSCHAFT, NATUR- UND VERBRAUCHERSCHUTZ DES LANDES NORDRHEIN-WESTFALEN: ELWAS-WEB. Elektronisches wasserwirtschaftliches Verbundsystem für die Wasserwirtschaftsverwaltung in NRW. URL: <https://www.elwasweb.nrw.de/elwas-web/index.jsf> [Stand: 09.04.2019] (abgerufen am 06.03.2019).

NETZBAND, M. (1998): Möglichkeiten und Grenzen der Fernerkundung zur Versiegelungskartierung in Siedlungsräumen. In: Sächsische Druck- und Verlagshaus GmbH.

SENATSWERWALTUNG FÜR STADTENTWICKLUNG UND WOHNEN (2012): Versiegelung In: Stadtentwicklung Berlin. Available at: [http://www.stadtentwicklung.berlin.de/umwelt/umweltatlas/dc102\\_01.htm#A1](http://www.stadtentwicklung.berlin.de/umwelt/umweltatlas/dc102_01.htm#A1) (Accessed: 20 February 2018).

STATISTISCHES BUNDESAMT (2017): Siedlungs- und Verkehrsfläche In: Umweltbundesamt. Available at: [https://www.umweltbundesamt.de/sites/default/files/medien/384/bilder/dateien/2\\_tab\\_suv\\_2017-04-07.pdf](https://www.umweltbundesamt.de/sites/default/files/medien/384/bilder/dateien/2_tab_suv_2017-04-07.pdf) (Accessed: 16 February 2018).

WOLF, C. AND HENNING, B. (2010): Handbuch der sozialwissenschaftlichen Datenanalyse. DOI: 10.1007/978-3-531-92038-2.

WOLLSCHLÄGER, D. (2017): Statistik und ihre Anwendungen Grundlagen der Datenanalyse mit R. Springer Spektrum. DOI: 10.1007/978-3-662-53670-4.

WOOLDRIDGE, J. M. (2008): Introductory Econometrics. A Modern Approach In: Cengage Learning Emea.

## Anhang

### Anhang 1: Korrekturwerte für ALKIS-Gebäudeflächen auf Flurstücke anhand der maßgeblichen Gebäudefunktion

| Flurstück mit der Maßgeblichen Gebäudefunktion | Wertart Gebäudefunktion | Anzahl Flurstücke | Korrekturwert [m²] | Mittlere Gebäudefläche [m²] |
|------------------------------------------------|-------------------------|-------------------|--------------------|-----------------------------|
| Wohngebäude                                    | 1000                    | 1520              | + 7,6              | 202,9                       |
| Wohnhaus                                       | 1010                    | 539850            | + 20,3             | 153                         |
| Zoll                                           | 1100                    | 6761              | + 17,6             | 277                         |
| Museum                                         | 1120                    | 8397              | + 16,5             | 281,9                       |
| Wohn- und Geschäftsgebäude                     | 1123                    | 1151              | + 9,5              | 263,9                       |
| Wohngebäude mit Gewerbe und Industrie          | 1130                    | 7303              | + 21,6             | 270,9                       |
| Land- und forstwirtschaftliches Wohngebäude    | 1210                    | 1126              | + 61,6             | 417                         |
| Überdachung                                    | 1610                    | 2492              | + 4,9              | 155,3                       |
| Carport                                        | 1611                    | 1167              | + 2,4              | 36,8                        |
| Gebäude für Wirtschaft oder Gewerbe            | 2000                    | 4393              | + 21,9             | 448,8                       |
| Gebäude für Handel und Dienstleistungen        | 2010                    | 4263              | + 26               | 816,2                       |
| Bürogebäude                                    | 2020                    | 1836              | + 14,6             | 547,9                       |
| Geschäftsgebäude                               | 2050                    | 969               | + 12,8             | 1003,2                      |
| Laden                                          | 2054                    | 1324              | + 28,2             | 760,7                       |
| Gebäude für Gewerbe und Industrie              | 2100                    | 8918              | + 48,7             | 1305,2                      |
| Fabrik                                         | 2111                    | 1530              | + 61,6             | 2592,7                      |
| Werkstatt                                      | 2120                    | 3846              | + 37,9             | 629,4                       |
| Gebäude für Vorratshaltung                     | 2140                    | 1101              | + 41,1             | 784,9                       |
| Lagerhalle, Lagerschuppen, Lagerhaus           | 2143                    | 6066              | + 16,9             | 1123,7                      |
| Betriebsgebäude zu Verkehrsanlagen (allgemein) | 2400                    | 1081              | + 11,4             | 173,1                       |
| Garage                                         | 2463                    | 59085             | + 3,1              | 54,6                        |

|                                                 |      |        |         |        |
|-------------------------------------------------|------|--------|---------|--------|
| Umformer                                        | 2523 | 5581   | + 1,3   | 19     |
| Land- und forstwirtschaftliches Betriebsgebäude | 2720 | 2922   | + 86,1  | 846,9  |
| Scheune                                         | 2721 | 1532   | + 67,1  | 728,7  |
| Schuppen                                        | 2723 | 5349   | + 8,7   | 105,6  |
| Stall                                           | 2724 | 3746   | + 46,9  | 983,2  |
| Allgemein bildende Schule                       | 3021 | 1222   | + 102,1 | 2227,4 |
| Kinderkrippe, Kindergarten, Kindertagesstätte   | 3065 | 1630   | + 10    | 546,2  |
| Sonstige                                        | -    | 257310 | + 4,8   | 192    |

## Anhang 2: Ergebnisse der Clusteranalyse und die Übereinstimmung mit Kalibrierungsdaten

| Cluster | Anzahl Trainingsdaten | Anzahl Kalibrierungsdaten (EZ Hüller Bach) | Summe der Flurstück-flächen | Summe der Zuwegungs-Flächen | Summe der Zuwegung-Flächen Prognose | Differenz | „Random Forest“ Klassifikation | Polynom des Regressions-modells | Bestimmtheitsmaß $R^2$ | Mittlere Übereinstimmung mit Prognose | durchschnittliche Zuwegungsanteil der effektiven Fläche [%] |
|---------|-----------------------|--------------------------------------------|-----------------------------|-----------------------------|-------------------------------------|-----------|--------------------------------|---------------------------------|------------------------|---------------------------------------|-------------------------------------------------------------|
| 1       | 399344                | 27394                                      | 18905189                    | 3086807                     | 3200626                             | 104%      | 0.9691                         | 8                               | 0.6976715              | 0.6559329                             | 0.269                                                       |
| 2       | 23852                 | 1557                                       | 284979                      | 113357                      | 113752                              | 100%      | 0.9094                         | 3                               | 0.825414               | 0.6802688                             | 0.7847028                                                   |
| 3       | 28149                 | 2152                                       | 2914455                     | 692931                      | 692545                              | 100%      | 0.9633                         | 8                               | 0.7334108              | 0.6333488                             | 0.4681425                                                   |
| 4       | 7012                  | 684                                        | 3383848                     | 872451                      | 905673                              | 104%      | 0.9678                         | 1                               | 0.8128634              | 0.5902227                             | 0.4187451                                                   |
| 5       | 20636                 | 1880                                       | 9108552                     | 3451763                     | 3353866                             | 97%       | 0.9596                         | 2                               | 0.8129894              | 0.6910591                             | 0.6408915                                                   |
| 6       | 56373                 | 4077                                       | 1921531                     | 369911                      | 401300                              | 108%      | 0.9453                         | 7                               | 0.8685742              | 0.5629012                             | 0.2726397                                                   |
| 7       | 5979                  | 707                                        | 275199                      | 60177                       | 61144                               | 102%      | 0.9264                         | 5                               | 0.6684832              | 0.5844017                             | 0.5741785                                                   |
| 8       | 12634                 | 1021                                       | 243844                      | 47809                       | 46428                               | 97%       | 0.9226                         | 7                               | 0.6980159              | 0.5827553                             | 0.3972396                                                   |
| 9       | 41146                 | 3707                                       | 1429633                     | 270641                      | 286512                              | 106%      | 0.8972                         | 8                               | 0.849509               | 0.5225935                             | 0.3869873                                                   |

## **Lastenheft**

# **ALKIS Flächenpolygone**

Wasserwirtschaftliche Ermittlungen und  
Anwendung von digitalen Versiegelungsdaten aus amtlichen  
Liegenschaftskatasterdaten - WaLKI

## **Projektbearbeitung**

### **Institution**

Forschungsinstitut für Wasser- und Abfallwirtschaft  
an der RWTH Aachen (FiW) e.V.  
Kackertstraße 15-17  
52056 Aachen

### **Bearbeiter**

Dr.-Ing. Gesa Kutschera

Sajjad Tabatabaei, M.Sc.

Leon Indjike, M.Sc.

## Tabellarische Zusammenfassung

### Objektartenauswahl

| Annahme         | Objektart                                           | Kennung | Attributart            | Wertarten                                            |
|-----------------|-----------------------------------------------------|---------|------------------------|------------------------------------------------------|
| Dachflächen     | Gebäude                                             | 31001   | Lage zur Erdoberfläche | 1200 ausgenommen                                     |
|                 | Turm                                                | 51001   |                        |                                                      |
|                 | Bauwerke od. Anlage für Industrie und Gewerbe       | 51002   | Bauwerksfunktion       | 1215,1220,1290, 1350                                 |
|                 | Vorratsbehälter Speicherbauwerk                     | 51003   |                        |                                                      |
|                 | Bauwerk od. Anlage für Sport, Freizeit und Erholung | 51006   | Bauwerksfunktion       | 1431,1490                                            |
|                 | sonstige Bauwerke od. sonstige Einrichtung          | 51009   | Bauwerksfunktion       | 1610,1611                                            |
| Verkehrsflächen | Straßenverkehr                                      | 42001   | Nutzungsfunktion       | 5130, NULL                                           |
|                 | Weg                                                 | 42006   | Nutzungsfunktion       | 5210, 5211, 5212, 5220, 5230, 5240, 5250, 5260, 9999 |
|                 | Platz                                               | 42009   | Nutzungsfunktion       | 5130, 5310, 5320, 5330, 5340, 5350                   |
|                 | Bahnverkehr                                         | 42010   | Nutzungsfunktion       | 2321, 2322, NULL                                     |
|                 | Flugverkehr                                         | 42015   | Verkehrsfunktion       | NULL, 5501                                           |

|                                                        |                                                     |       |                    |                                                                                |
|--------------------------------------------------------|-----------------------------------------------------|-------|--------------------|--------------------------------------------------------------------------------|
| Bauwerke, Anlagen und Einrichtungen im Verkehrsbereich | Bauwerke im Verkehrsbereich                         | 53001 | Bauwerksfunktion   | 1800, 1801, 1802, 1803, 1804, 1805, 1807, 1808, 1810, 1820, 1830, 1840, 1845   |
|                                                        | Straßenverkehrsanlagen                              | 53002 | Verkehrsanlagenart | 1000                                                                           |
|                                                        | Weg, Pfad, Steig                                    | 53003 | Nutzungsart        | 1106, 1111                                                                     |
|                                                        | Bahnverkehrsanlage                                  | 53004 | Bahnhofskategorie  | 1010, 1020, 1030                                                               |
|                                                        | Flugverkehrsanlage                                  | 53007 | Verkehrsanlagenart | 1310, 1320, 1330, 5531                                                         |
|                                                        | Bauwerk im Gewässerbereich                          | 53009 | Bauwerksfunktion   | 2030, 2050                                                                     |
| versiegelt, nicht behandlungsbedürftig                 | Bauwerke od. Anlage für Industrie und Gewerbe       | 51002 | Bauwerksfunktion   | 1230,1280,1390, 1400                                                           |
|                                                        | Bauwerk od. Anlage für Sport, Freizeit und Erholung | 51006 | Bauwerksfunktion   | 1430,1432,1450, 1470,                                                          |
|                                                        | sonstige Bauwerke od. sonstige Einrichtung          | 51009 | Bauwerksfunktion   | 1620,1621,1622, 1630,1640,1650, 1670,1700,1701, 1702,1703,1720, 1721,1722,1723 |
| versiegelt, behandlungsbedürftig                       | Bauwerke od. Anlage für Industrie und Gewerbe       | 51002 | Bauwerksfunktion   | 1210,1330,1331, 1332, 1333,1340                                                |
| Teilversiegelte Flächen                                | Bauwerk od. Anlage für Sport, Freizeit und Erholung | 51006 | Bauwerksfunktion   | 1410, 1411, 1420, 9999                                                         |

|                  |                                                     |       |                  |                                                                             |
|------------------|-----------------------------------------------------|-------|------------------|-----------------------------------------------------------------------------|
|                  | Weg, Pfad, Steig                                    | 53003 | Nutzungsart      | 1103, 1107, 1110, NULL                                                      |
|                  | Gleis                                               | 53006 | Ausführungsart   | 1200                                                                        |
| nicht versiegelt | Bauwerke od. Anlage für Industrie und Gewerbe       | 51002 | Bauwerksfunktion | 2140,1250,1260, 1270, 1310                                                  |
|                  | Bauwerk od. Anlage für Sport, Freizeit und Erholung | 51006 | Bauwerksfunktion | 1412,1440,1460, 1480, 1510                                                  |
|                  | sonstige Bauwerke od. sonstige Einrichtung          | 51009 | Bauwerksfunktion | 1740,1750,1760, 1761, 1762,1763,1770, 1780, 1781,1782,1783, 1790, 1791,9999 |
|                  | Weg, Pfad, Steig                                    | 53003 | Nutzungsart      | 1105                                                                        |
|                  | Vegetationsmerkmal                                  | 54001 |                  |                                                                             |
|                  | Untergeordnetes Gewässer                            | 55002 |                  |                                                                             |
|                  | Böschungsfäche                                      | 61002 |                  |                                                                             |