

FERNERKUNDUNG UND MASCHINELLES LERNEN ZUR ERFASSUNG VON URBANEM GRÜN

**Eine Analyse am Beispiel der
Verteilungsgerechtigkeit in
Deutschland**

Inaugural-Dissertation zur
Erlangung der Doktorwürde der
Philosophischen Fakultät der
Julius-Maximilians-Universität
Würzburg

vorgelegt von
Matthias Johann Weigand
aus Landsberg am Lech

Würzburg
2023



Erstgutachter: Prof. Dr. Stefan Dech
Zweitgutachter: Prof. Dr. Hannes Taubenböck
Drittgutachter: Prof. Dr. Jürgen Rauh
Tag des Kolloquiums: 13.12.2023

Diese Dissertation wurde im Rahmen des Promotionsstipendiumprogramms der Deutschen Bundesstiftung Umwelt (DBU) gefördert (Fördernummer 20016/468).

Gefördert durch das
Stipendienprogramm der



Deutsche
Bundesstiftung Umwelt

www.dbu.de

Diese Dissertation wurde angefertigt in der Abteilung "Georisiken und Zivile Sicherheit" am Deutschen Fernerkundungsdatenzentrum (DFD), Earth Observation Center (EOC) des Deutschen Zentrums für Luft- und Raumfahrt (DLR) in Oberpfaffenhofen.



DLR

Deutsches Zentrum
für Luft- und Raumfahrt

Danksagungen

Auf dem Weg zu dieser Arbeit habe ich von vielen Seiten Unterstützung erfahren. Ich möchte mich herzlich bei allen bedanken, die dazu beigetragen haben, diese Arbeit zu realisieren.

Insbesondere möchte ich mich bei Prof. Dr. Stefan Dech bedanken, dass er diese Arbeit von Beginn an unterstützt und diese am Deutschen Fernerkundungsdatenzentrum (DFD) des Deutschen Zentrums für Luft- und Raumfahrt (DLR) eingebunden hat. Dies ermöglichte mir eine vielseitige persönliche sowie fachliche Weiterentwicklung. Durch die vielfältigen Inspirationen und Ermutigungen von Prof. Dr. Hannes Taubenböck sowie dessen entgegengebrachtes Vertrauen als Teamleiter lernte ich immer wieder neue Perspektiven kennen. Von diesen Erfahrungen werde ich sicher noch lange profitieren und dafür möchte ich mich an dieser Stelle herzlich bedanken. Außerdem bedanke ich mich für die Zusammenarbeit mit Prof. Dr. Jürgen Rauh, die mich fachlich einen weiteren Schritt nach vorne gebracht hat.

Vor über 10 Jahren habe ich mein erstes Praktikum am DLR absolviert, dabei wurde mir meine Begeisterung für die Fernerkundung und Geoinformatik dank Dr. Michael Wurm bewusst. Er ist mein langjähriger Unterstützer, der mich auf dem Weg zu meinen Studienabschlüssen begleitet, der die Anbahnung dieses Promotionsvorhabens unterstützt hat und als Mentor für mich da war. Dafür danke ich ihm. Die zahlreichen Diskussionen, ob fachlich oder weit darüber hinaus, halfen mir dabei, mich in der Wissenschaftslandschaft zurechtzufinden und Fuß zu fassen.

Ich bedanke mich bei der Deutschen Bundesstiftung Umwelt (DBU) für die Förderung und Finanzierung meiner Promotion im Rahmen des Promotionsstipendienprogramms. Der interdisziplinäre Austausch im Kreis der Stipendiatinnen und Stipendiaten war für mich eine sehr wertvolle Erfahrung, um über den fachlichen Tellerrand zu blicken und zu verstehen, wie wichtig und gleichzeitig herausfordernd fachübergreifender wissenschaftlicher Austausch ist. Besonders möchte ich in diesem Zusammenhang Verena Exner und Dr. Hedda Schlegel-Starmann danken.

Vor und während dieser Dissertation habe ich ein unschätzbare Maß an Unterstützung von vielen Kolleginnen und Kollegen, Freundinnen und Freunden erfahren. Egal ob gute Gespräche oder aufbauende Worte, ob in der Teeküche, in Videokonferenzen oder beim Kolloquium an der Universität, ich möchte mich bei allen bedanken, die diese Erfahrung so einzigartig gemacht haben. Besonders möchte ich Ines Standfuß für die vielen Stunden danken, die wir uns verquatscht haben. Sie gab mir immer eine wertvolle Perspektive, wenn ich 'mal feststeckte und half mir so, viele knifflige Situationen zu meistern. Ich möchte Jeroen Staab danken, der für den nötigen Spritzer

Farbe im manchmal grauen Alltag aus Nullen und Einsen sorgte. In ihm habe ich einen Kollegen und Freund gefunden, der nie davor zurückscheute meine Ideen und Meinungen herauszufordern und mich so dazu anzuregen, die eigene Position zu hinterfragen und neue Wege zu beschreiten. Ariane Droin danke ich dafür, dass sie stets ein offenes Ohr hatte, wenn das Chaos überwältigend schien und sie so ein Ruhepol im Wirbelwind des Doktorandenalltags war. Der regelmäßige Austausch während der Pandemie half immer dabei neue Motivation zu finden. Ich danke Thomas Stark für die vielen Diskussionen zu den methodischen Fragen im Bereich des maschinellen Lernens. Diese halfen enorm dabei, mein eigenes neuronales Netz zu trainieren.

Allen Koautorinnen und Koautoren der Journalbeiträge und Publikationen, die im Rahmen dieser Dissertation entstanden sind, danke ich für das wertvolle Feedback und die konstruktive Kritik. Dadurch habe ich gelernt, meine eigene Arbeit besser einzuschätzen, den Fokus zu finden und beizubehalten. Außerdem möchte ich mich bei allen Kolleginnen und Kollegen in der Abteilung „Georisiken und Zivile Sicherheit“ am DFD bedanken. Prof. Dr. Günter Strunz danke ich für die Entfaltungsmöglichkeiten, die ich in dieser Abteilung erhalten habe. Ich danke Dr. Torsten Riedlinger, der mich insbesondere in der finalen Phase der Dissertation durch seinen Rat, das entgegengebrachte Vertrauen und die Freiheiten sehr unterstützt hat. Meine Kollegen Fabian Henkel, Ralph Kiefl und Martin Mühlbauer sorgten mit dem unkomplizierten und erfrischend pragmatischen Zugang zu Rechnerressourcen für die nötige Power, diese Arbeit fertigzustellen. Vielen Dank für diese Unterstützung. Des Weiteren bedanke ich mich bei Patrizia Hanika für ihre Arbeit und Unterstützung.

Ich danke allen Beitragenden der OpenStreetMap-Community, von deren gemeinschaftlichen Anstrengungen des Kartierens und dem gewährten offenen Datenzugang zu geographischen Daten diese Arbeit fundamental profitiert hat. Darüber hinaus wäre diese Arbeit ohne den Zugang zu freier und Open-Source-Software nicht möglich gewesen. Ich danke allen Autoren und Maintainern für ihr Engagement, das den freien Zugang zu wissenschaftlichem Rechnen ermöglicht.

Außerdem danke ich meinen Eltern Erika Weigand und Klaus Weigand, die mich stets auf meinem Bildungsweg gefördert haben, an mich geglaubt haben und immer eine Idee für einen Plan B hatten. Dieses ungebrochene Vertrauen, der Rückhalt und die Unterstützung über die Jahre hinweg ermöglichten mir diese Dissertation anzustreben und nun zu Ende zu bringen. Bei ihnen sowie meinen Schwestern möchte ich mich herzlich für das Verständnis über die gesamte Zeit, in der ich vorwiegend und bisweilen etwas zu sehr mit dieser Arbeit beschäftigt war, bedanken.

Zu guter Letzt geht mein tiefster Dank an Natalie Bienenstein. Sie hat mich über all die Zeit aufgemuntert, ungebrochenes Verständnis und unfassbare Geduld gezeigt und dabei selbst sehr oft zurückgesteckt. Ihr danke ich für den unendlichen Rückhalt, den sie mir auch in den anstrengendsten Zeiten gegeben hat. Ohne ihre beständige Unterstützung hätte ich diese Arbeit nicht vollenden können.

Zusammenfassung

Grünflächen stellen einen der wichtigsten Umwelteinflüsse in der Wohnumwelt der Menschen dar. Einerseits wirken sie sich positiv auf die physische und mentale Gesundheit der Menschen aus, andererseits können Grünflächen auch negative Wirkungen anderer Faktoren abmildern, wie beispielsweise die im Laufe des Klimawandels zunehmenden Hitzeereignisse. Dennoch sind Grünflächen nicht für die gesamte Bevölkerung gleichermaßen zugänglich. Bestehende Forschung im Kontext der Umweltgerechtigkeit (UG) konnte bereits aufzeigen, dass unterschiedliche sozio-ökonomische und demographische Gruppen der deutschen Bevölkerung unterschiedlichen Zugriff auf Grünflächen haben. An bestehenden Analysen von Umwelteinflüssen im Kontext der UG wird kritisiert, dass die Auswertung geographischer Daten häufig auf zu stark aggregiertem Level geschieht, wodurch lokal spezifische Expositionen nicht mehr genau abgebildet werden. Dies trifft insbesondere für großflächig angelegte Studien zu. So werden wichtige räumliche Informationen verloren. Doch moderne Erdbeobachtungs- und Geodaten sind so detailliert wie nie und Methoden des maschinellen Lernens ermöglichen die effiziente Verarbeitung zur Ableitung höherwertiger Informationen.

Das übergeordnete Ziel dieser Arbeit besteht darin, am Beispiel von Grünflächen in Deutschland methodische Schritte der systematischen Umwandlung umfassender Geodaten in relevante Geoinformationen für die großflächige und hochaufgelöste Analyse von Umwelteigenschaften aufzuzeigen und durchzuführen. An der Schnittstelle der Disziplinen Fernerkundung, Geoinformatik, Sozialgeographie und Umweltgerechtigkeitsforschung sollen Potenziale moderner Methoden für die Verbesserung der räumlichen und semantischen Auflösung von Geoinformationen erforscht werden. Hierfür werden Methoden des maschinellen Lernens eingesetzt, um Landbedeckung und -nutzung auf nationaler Ebene zu erfassen. Diese Entwicklungen sollen dazu beitragen bestehende Datenlücken zu schließen und Aufschluss über die Verteilungsgerechtigkeit von Grünflächen zu bieten.

Diese Dissertation gliedert sich in drei konzeptionelle Teilschritte. Im ersten Studienteil werden Erdbeobachtungsdaten der Sentinel-2 Satelliten zur deutschlandweiten Klassifikation von Landbedeckungsinformationen verwendet. In Kombination mit punktuellen Referenzdaten der europaweiten Erfassung für Landbedeckungs- und Landnutzungsinformationen des *Land Use and Coverage Area Frame Survey (LUCAS)* wird ein maschinelles Lernverfahren trainiert. In diesem Kontext werden verschiedene Vorverarbeitungsschritte der LUCAS-Daten und deren Einfluss auf die Klassifikationsgenauigkeit beleuchtet. Das Klassifikationsverfahren ist in der

Lage Landbedeckungsinformationen auch in komplexen urbanen Gebieten mit hoher Genauigkeit abzuleiten. Ein Ergebnis des Studienteils ist eine deutschlandweite Landbedeckungsklassifikation mit einer Gesamtgenauigkeit von 93,07 %, welche im weiteren Verlauf der Arbeit genutzt wird, um grüne Landbedeckung (GLC) räumlich zu quantifizieren.

Im zweiten konzeptionellen Teil der Arbeit steht die differenzierte Betrachtung von Grünflächen anhand des Beispiels öffentlicher Grünflächen (PGS), die häufig Gegenstand der UG-Forschung ist, im Vordergrund. Doch eine häufig verwendete Quelle für räumliche Daten zu öffentlichen Grünflächen, der *European Urban Atlas (EUA)*, wird bisher nicht flächendeckend für Deutschland erhoben. Dieser Studienteil verfolgt einen datengetriebenen Ansatz, die Verfügbarkeit von öffentlichem Grün auf der räumlichen Ebene von Nachbarschaften für ganz Deutschland zu ermitteln. Hierfür dienen bereits vom EUA erfasste Gebiete als Referenz. Mithilfe einer Kombination von Erdbeobachtungsdaten und Informationen aus dem OpenStreetMap-Projekt wird ein *Deep Learning*-basiertes Fusionsnetzwerk erstellt, welche die verfügbare Fläche von öffentlichem Grün quantifiziert. Das Ergebnis dieses Schrittes ist ein Modell, welches genutzt wird, um die Menge öffentlicher Grünflächen in der Nachbarschaft zu schätzen ($R^2 = 0.952$).

Der dritte Studienteil greift die Ergebnisse der ersten beiden Studienteile auf und betrachtet die Verteilung von Grünflächen in Deutschland unter Hinzunahme von georeferenzierten Bevölkerungsdaten. Diese exemplarische Analyse unterscheidet dabei Grünflächen nach zwei Typen: GLC und PGS. Zunächst wird mithilfe deskriptiver Statistiken die generelle Grünflächenverteilung in der Bevölkerung Deutschlands beleuchtet. Daraufhin wird die Verteilungsgerechtigkeit anhand gängiger Gerechtigkeitsmetriken bestimmt. Abschließend werden die Zusammenhänge zwischen der demographischen Komposition der Nachbarschaft und der verfügbaren Menge von Grünflächen anhand dreier exemplarischer soziodemographischer Gesellschaftsgruppen untersucht. Die Analyse zeigt starke Unterschiede der Verfügbarkeit von PGS zwischen städtischen und ländlichen Gebieten. Ein höherer Prozentsatz der Stadtbevölkerung hat Zugriff das Mindestmaß von PGS gemessen an der Vorgabe der Weltgesundheitsorganisation. Die Ergebnisse zeigen auch einen deutlichen Unterschied bezüglich der Verteilungsgerechtigkeit zwischen GLC und PGS und verdeutlichen die Relevanz der Unterscheidung von Grünflächentypen für derartige Untersuchungen. Die abschließende Betrachtung verschiedener Bevölkerungsgruppen arbeitet Unterschiede auf soziodemographischer Ebene auf.

In der Zusammenschau demonstriert diese Arbeit wie moderne Geodaten und Methoden des maschinellen Lernens genutzt werden können bisherige Limitierungen räumlicher Datensätze zu überwinden. Am Beispiel von Grünflächen in der Wohnumgebung der Bevölkerung Deutschlands wird gezeigt, dass landesweite Analysen zur Umweltgerechtigkeit durch hochaufgelöste und lokal feingliedrige geographische Informationen bereichert werden können. Diese Arbeit verdeutlicht, wie die Methoden der Erdbeobachtung und Geoinformatik einen wichtigen Beitrag leisten können, die Ungleichheit der Wohnumwelt der Menschen zu identifizieren und schlussendlich den nachhaltigen Siedlungsbau in Form von objektiven Informationen zu unterstützen und überwachen.

Abstract

Green spaces are one of the most important environmental factors for humans in the living environment. On the one hand they provide benefits to people's physical and mental health, on the other hand they allow for the mitigation of negative impacts of environmental stressors like heat waves which are increasing as a result of climate change. Yet, green spaces are not equally accessible to all people. Existing literature in the context of Environmental Justice (EJ) research has shown that the access to green space varies among different socio-economic and demographic groups in Germany. However, previous studies in the context of EJ were criticized for using strongly spatially aggregated data for their analyses resulting in a loss of spatial detail on local environmental exposure metrics. This is especially true for large-scale studies where important spatial information often get lost. In this context, modern earth observation and geospatial data are more detailed than ever, and machine learning methods enable efficient processing to derive higher value information for diverse applications.

The overall objective of this work is to demonstrate and implement methodological steps that allow for the transformation of vast geodata into relevant geoinformation for the large-scale and high-resolution analysis of environmental characteristics using the example of green spaces in Germany. By bridging the disciplines remote sensing, geoinformatics, social geography and environmental justice research, potentials of modern methods for the improvement of spatial and semantic resolution of geoinformation are explored. For this purpose, machine learning methods are used to map land cover and land use on a national scale. These developments will help to close existing data gaps and provide information on the distributional equity of green spaces.

This dissertation comprises three conceptual steps. In the first part of the study, earth observation data from the Sentinel-2 satellites are used to derive land cover information across Germany. In combination with point reference data on land cover and land use from the paneuropean *Land Use and Coverage Area Frame Survey (LUCAS)* a machine learning model is trained. Therein, different preprocessing steps of the LUCAS data and their influence on the classification accuracy are highlighted. The classification model derives land cover information with high accuracy even in complex urban areas. One result of the study is a Germany-wide land cover classification with an overall accuracy of 93.07 % which is used in the further course of the dissertation to spatially quantify green land cover (GLC).

The second conceptual part of this study focuses on the semantic differentiation of green spaces using the example of public green spaces (PGS), which is often the subject of EJ research. A frequently used source of spatial data on public green spaces, the *European Urban Atlas (EUA)*, however, is not available for all of Germany. This part of the study takes a data-driven approach to determine the availability of public green space at the spatial level of neighborhoods for all of Germany. For this purpose, areas already covered by the EUA serve as a reference. Using a combination of earth observation data and information from the OpenStreetMap project, a *Deep Learning*-based fusion network is created that quantifies the available area of public green space. The result of this step is a model that is utilized to estimate the amount of public green space in the neighborhood ($R^2 = 0.952$).

The third part of this dissertation builds upon the results of the first two parts and integrates georeferenced population data to study the socio-spatial distribution of green spaces in Germany. This exemplary analysis distinguishes green spaces according to two types: GLC and PGS. In this, first, descriptive statistics are used to examine the overall distribution of green spaces available to the German population. Then, the distributional equality is determined using established equality metrics. Finally, the relationships between the demographic composition of the neighborhood and the available amount of green space are examined using three exemplary sociodemographic groups. The analysis reveals strong differences in PGS availability between urban and rural areas. Compared to the rural population, a higher percentage of the urban population has access to the minimum level of PGS defined as a target by the World Health Organization (WHO). The results also show a clear deviation in terms of distributive equality between GLC and PGS, highlighting the relevance of distinguishing green space types for such studies. The final analysis of certain population groups addresses differences at the sociodemographic level.

In summary, this dissertation demonstrates how previous limitations of spatial datasets can be overcome through a combination of modern geospatial data and machine learning methods. Using the example of green spaces in the residential environment of the population in Germany, it is shown that nationwide analyses of environmental justice can be enriched by high-resolution and locally fine-grained geographic information. This study illustrates how earth observation and methods of geoinformatics can make an important contribution to identifying inequalities in people's living environment. Such objective information can ultimately be deployed to support and monitor sustainable urban development.

Inhaltsverzeichnis

Zusammenfassung	iii
Abstract	v
Abbildungsverzeichnis	xii
Tabellenverzeichnis	xiii
Abkürzungsverzeichnis	xv
1 Einleitung	1
1.1 Hintergrund	1
1.2 Forschungslücke	3
1.3 Ziele der Arbeit und Forschungsfragen	6
1.4 Aufbau der Arbeit	7
2 Grundlagen	11
2.1 Grünflächen als Ressource in der urbanen Umwelt	11
2.2 Gerechtigkeit der räumlichen Verteilung von Grünflächen	16
2.3 Fernerkundung zur Erfassung von Grünflächen	19
2.4 Maschinelles Lernen im Kontext der Fernerkundung	23
3 Konzeptioneller Rahmen	29
3.1 Beschreibung der Umwelt in Deutschland	29
3.2 Quantifizierung von öffentlichen Grünflächen in der Wohnumgebung	32
3.3 Gerechtigkeit der nachbarschaftlichen Grünflächen in Deutschland	35
4 Daten und Vorprozessierung	37
4.1 LUCAS als Referenzdaten für Landbedeckung	37
4.1.1 LUCAS Positionierung	39
4.1.2 LUCAS Vorprozessierung	41
4.2 Sentinel-2 Bilddaten	44
4.3 European Urban Atlas als Referenzdaten für öffentliche Grünflächen	49
4.4 OpenStreetMap Daten zu Grünflächentypen	51

4.5	Zensus 2011	55
4.6	Stadt-Land-Gradient	56
5	Methoden	57
5.1	Landbedeckungsklassifikation mit Sentinel-2 und LUCAS	57
5.1.1	Modellierung von Landbedeckung	57
5.1.2	Genauigkeitsabschätzung und Validierung	59
5.1.3	Erstellung einer deutschlandweiten Landbedeckungsklassifikation	60
5.1.4	Übertrag auf 2018	62
5.2	Deep Learning Datenfusion zur Erfassung öffentlichen Grüns	63
5.2.1	Vorbereitung des multimodalen Datensatzes	63
5.2.2	Modellierung mit Deep Learning	65
5.2.3	Genauigkeitsabschätzung	68
5.3	Analyse zur Verteilung von Grün in Deutschland	69
5.3.1	Datenvorbereitung	70
5.3.2	Deskriptive Statistiken	71
5.3.3	Verteilungsgerechtigkeit von Grünflächen	72
5.3.4	Grünflächenverfügbarkeit nach Bevölkerungsgruppen	74
6	Ergebnisse und Interpretation	77
6.1	Flächendeckende Landbedeckungsklassifikation mit Sentinel-2 und LUCAS	77
6.1.1	Evaluation der Stichprobenauswahl und Gesamtgenauigkeit	77
6.1.2	Großflächige Anwendung der Landbedeckungsklassifikation	78
6.1.3	Übertrag der Landbedeckungsklassifikation auf den Zeitraum 2018	80
6.2	Erfassung öffentlicher Grünflächen in Deutschland	82
6.2.1	Modellierungsgenauigkeit	83
6.2.2	Geographische Analyse der Modellierungsgenauigkeit	83
6.3	Gerechtigkeit der Verteilung von Grünflächen	88
6.3.1	Grünflächenverfügbarkeit in Deutschland	88
6.3.2	Gerechtigkeit von Grünflächenverteilung	91
6.3.3	Verfügbarkeit von Grünflächen und Bevölkerungszusammensetzung	95
7	Diskussion	101
7.1	Landbedeckungsklassifikation	101
7.1.1	Auswirkungen der Vorprozessierung von LUCAS-Daten	101
7.1.2	Positionierung von LUCAS-Stichproben	103
7.1.3	Bildklassifikation und Klassifikationsgenauigkeit	104
7.1.4	Übertrag auf 2018	105
7.2	Modellierung des öffentlichen Grünanteils mit Deep Learning	106
7.3	Gerechtigkeit der räumlichen Verteilung von Grünflächen	109
7.3.1	Grünflächenverfügbarkeit in Deutschland	109

7.3.2	Verteilungsgerechtigkeit und Einfluss der Demographie auf Grünflächen- verfügbarkeit	110
7.3.3	Limitierungen und weitere Forschung	112
8	Fazit und Ausblick	115
	Literatur	119
	Anhang	147
	Eidesstattliche Versicherung	157

Abbildungsverzeichnis

1.1	Raumkonzepte für die Abschätzung von Exposition gegenüber Umweltbelastungen	4
2.1	Gesundheitliche Wirkungspfade von Grünflächen	13
2.2	Spektralkanäle Sentinel-2, Landsat-7 und Landsat-8	21
3.1	Ablaufdiagramm der Landbedeckungsklassifikation	31
3.2	Ablaufdiagramm der Ableitung von öffentlichem Grünanteil	34
4.1	Überblick der LUCAS-Feldproben Punkte in Deutschland und Europa 2015	38
4.2	Abdeckung Deutschlands mit Sentinel-2 Bilddaten zwischen 2015 und 2017	45
4.3	Sentinel-2 Mosaik für Deutschland 2015–2017	47
4.4	Aggregate spektraler Indizes für Deutschland 2015–2017	48
4.5	European Urban Atlas in München	50
4.6	Abdeckung Deutschlands durch European Urban Atlas	51
4.7	Vergleich von OpenStreetMap Webansicht und Daten	53
5.1	Abdeckung Deutschlands mit Sentinel-2 Bilddaten in 2018	62
5.2	Häufigkeitsverteilung der öffentlichen Grünflächen je Nachbarschaft	64
5.3	Skizze des Deep Learning Fusionsnetzwerkes zur Ableitung öffentlicher Grünflächenanteile	66
5.4	Zweidimensionales Visualisierungsschema zur Einteilung der Grünflächenverfügbarkeit	71
5.5	Beispiel für eine Lorenzkurve	72
6.1	Gesamtgenauigkeit der Landbedeckungsklassifikation	78
6.2	Landbedeckungsklassifikation für ganz Deutschland	79
6.3	Vergleich der Landbedeckungsklassifikationen von 2015/2017 und 2018	82
6.4	Metriken des Trainingsverlaufs des Fusionmodells	84
6.5	Vergleich von Referenz und Prädiktion öffentlicher Grünflächen	85
6.6	Residuen des Fusionsmodells	85
6.7	Räumliches Beispiel der Modellierung öffentlichen Grüns	86
6.8	Geographische Analyse der Residuen des Fusionsmodells	87
6.9	Residuen des Fusionsmodells in Abhängigkeit zur Distanz zur Bebauung	87
6.10	Grünflächenverfügbarkeit in München	89

6.11 Verfügbarkeit von Grünflächen in Deutschland gesamt und nach Stadtgröße	89
6.12 Verfügbarkeit von Grünflächen in Deutschland nach Stadt-Land-Gradient und Gebäudetypen	90
6.13 Lorenzkurve der Verteilung von Grünflächen in Deutschland	91
6.14 Gerechtigkeitsmaße der Grünflächenverteilung nach Stadtgröße	93
6.15 Gerechtigkeitsmaße der Grünflächenverteilung nach Stadt-Land-Gradient	93
6.16 Gerechtigkeitsmaße der Grünflächenverteilung nach Gebäudetypen	94
6.17 Vergleich der Grünflächenverfügbarkeit für die 100 größten Städte Deutschlands	95

Tabellenverzeichnis

4.1	Hauptlandbedeckungsklassen des LUCAS-Datensatzes	42
4.2	Anzahl der LUCAS-Stichproben pro Experiment	42
4.3	Klassenhierarchie für die deutschlandweite Landbedeckungsklassifikation	44
4.4	Überblick über die Sentinel-2 Bilddaten	49
4.5	Aggregationsschema für Gebäudetypklassifikation im Zensus 2011	55
5.1	Hyperparameter für neuronale Regressionsnetzwerke zur Modellierung von öffentlichem Grün	68
6.1	Fehlermatrix für deutschlandweite Landbedeckungsklassifikation 2015–2017	80
6.2	Fehlermatrix für deutschlandweite Landbedeckungsklassifikation 2018	81
6.3	Gesamtstatistik zur Grünflächenverfügbarkeit in der Bevölkerung Deutschlands	89
6.4	Verteilungsgerechtigkeit von Grünflächen in Deutschland	92
6.5	Vergleichsergebnisse für Zusammenhang zwischen Grünflächenverfügbarkeit je Nachbarschaft und Demographie	96
6.6	Zusammenhang zwischen grüner Landbedeckung und Demographie	97
6.7	Zusammenhang zwischen öffentlichem Grün und Demographie	99
6.8	Zusammenhang zwischen grüner Landbedeckung und Demographie nach Urbanität	100
6.9	Zusammenhang zwischen öffentlichem Grün und Demographie nach Urbanität	100
A1	OpenStreetMap-Elemente zur Extraktion von Straßen und Schienen für Landbedeckungsklassifikation	147
A2	Gruppierungsschlüssel von OpenStreetmap Elementen für Fusionsnetzwerk	147

Abkürzungsverzeichnis

Abkürzung	Beschreibung
ANN	Artificial Neural Network
CNN	Convolution Neural Network
COVID-19	Corona Virus Disease 2019
DL	Deep Learning
ESA	Europäische Raumfahrtagentur (engl. <i>European Space Agency</i>)
EUA	European Urban Atlas
GIS	Geoinformationssystem
GLC	Grüne Landbedeckung (engl. <i>green land cover</i>)
KI	Künstliche Intelligenz
LB	Landbedeckung
LIDAR	Light Detection and Ranging
LN	Landnutzung
LUCAS	Land Use And Coverage Area Frame Survey
ML	Maschinelles Lernen
MODIS	Moderate Resolution Imaging Spectroradiometer
NASA	National Aeronautics and Space Association
NDBI	Normalized Difference Built-up Index
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
OSM	OpenStreetMap
PGS	Öffentliche Grünflächen (engl. <i>public green space</i>)
RADAR	Radio Detection and Ranging

Abkürzung	Beschreibung
RF	Random Forest
SAR	Synthetic Aperture RADAR
UG	Umweltgerechtigkeit
UHI	Urbaner Hitzeinsel Effekt
VGI	Von Freiwilligen zur Verfügung gestellte Geoinformationen (engl. <i>Volunteered Geographic Information</i>)
WHO	Weltgesundheitsorganisation (engl. <i>World Health Organization</i>)

Einleitung

1.1 Hintergrund

„Spazieren ist der neue Volkssport“ titelte 2021 die Sächsische Zeitung¹ zur Höhe der COVID-19-Pandemie während der weitreichenden Ausgangsbeschränkungen im Winter 2020/2021 in Deutschland und kontrastierte dabei leere Innenstädte mit belebten Fußwegen in Parks und Wäldern. Vermutlich waren Grünflächen nie zuvor stärker im kollektiven Bewusstsein der Bevölkerung als während und nach der pandemiebedingten Ausnahmesituation. Ausgangsbeschränkungen, wie im Frühjahr 2020 in Bayern (Bayerisches Ministerium für Gesundheit und Pflege, 2020), führten letztlich dazu, dass Menschen in Haushalten ohne eigenen privaten Garten einzig öffentliche Parks für die Erholung im Grünen nutzen konnten. Laut einer *forsa*-Umfrage² 2020 nutzten 28 % der Menschen in deutschen Großstädten Parks und Grünflächen häufiger als noch vor Beginn der Pandemie. Darüber hinaus gab beinahe die Hälfte der Befragten an, dass Grünflächen für sie nach Ausbruch der Pandemie an Bedeutung gewonnen haben. Das macht deutlich, wie wichtig ein ausreichendes Angebot an öffentlich zugänglichen Grünflächen im Stadtgebiet für die Gesundheit und das Wohlbefinden ist. Zwar handelte es sich bei den pandemiebedingten weitreichenden Ausgangsbeschränkungen um eine Ausnahmesituation, die zugrundeliegende Relevanz von Grünflächen im Stadtbild existiert jedoch auch außerhalb von derartigen strikten Maßnahmen.

Aus wissenschaftlicher Sicht ist es jedoch keine grundlegend neue Erkenntnis, dass Grünflächen zentrale Elemente einer lebenswerten (Stadt-) Umwelt sind. Im Gegenteil, Grünflächen spielen eine wichtige Rolle für die Gesundheit der Bevölkerung (siehe z. B. Hartig et al., 2014). Tagtäglich bewegen sich die Menschen zwischen Autoabgasen, industriellem Feinstaub, sind dem Straßenverkehrslärm ausgesetzt und müssen im Sommer Hitzewellen ertragen. Nur selten kann man sich diesen Umweltbelastungen entziehen und bei einem Spaziergang im Park oder im Wald durchatmen. Untersuchungen zur Beziehung zwischen der Gesundheit der Menschen

¹Sächsische Zeitung, 17.01.2021, „Spazieren ist der neue Volkssport“, zuletzt abgerufen am 27.07.2023: <https://www.saechsische.de/coronavirus/spazierengehen-neue-trend-raus-corona-lockdown-hobby-5358864-plus.html>

²Beschaffungsdienst Galabau, 13.07.2020, „Forsa-Studie zeigt: Parks und Grünflächen in Corona-Zeiten wichtiger denn je“, zuletzt abgerufen am 27.07.2023, <https://www.soll-galabau.de/aktuelle-news/ansicht-aktuelles/datum/2020/07/13/forsa-studie-zeigt-parks-und-gruenflaechen-in-corona-zeiten-wichtiger-denn-je.html>

und ihrer Wohnumgebung wurden verstärkt seit den 1980er Jahren durchgeführt (Diez Roux, 2001). Viele Studien untersuchen die positiven Auswirkungen lokaler grüner Infrastruktur auf das Klima (Middel et al., 2014), die Gesundheit (z. B. Maas et al., 2006; Villeneuve et al., 2012; Richardson et al., 2013; Honold et al., 2016), das Wohlbefinden und die soziale Integration (Lee und Maheswaran, 2011; Pope et al., 2015). Die letztgenannten Auswirkungen sind bei der städtischen Bevölkerung weltweit besonders deutlich geworden, da die COVID-19-Pandemie psychische Probleme wie Angst, Desorientiertheit, Stress und viele andere verstärkt hat (World Health Organization, 2020).

Es ist daher nicht verwunderlich, dass Grünflächen gemeinhin wichtige Faktoren für die Verbesserung der Lebensqualität sind (Sapena et al., 2021; Castelli et al., 2023) und zu den wichtigsten Faktoren zählen, die Menschen bei der Wahl des Wohnstandorts einbeziehen (Wurm et al., 2019c). Wie sehr die Menschen durch Umweltbelastungen betroffen sind oder von Umweltressourcen profitieren können, hängt stark von der täglichen Wohn- und Arbeitsumwelt ab, in der sie sich bewegen. Beispielsweise besteht ein Zusammenhang zwischen dem relativen Grünflächenanteil in der Nachbarschaft und der Selbsteinschätzung zur Beeinträchtigung der Lebensqualität durch fehlendes Grün (Wurm et al., 2010, S. 159).

Die Verfügbarkeit und Zugänglichkeit von Grünflächen im urbanen Raum sind dabei ein grundsätzlich räumliches Phänomen. Dies spiegelt sich beispielsweise in der Wahrnehmung der Bevölkerung zur Nutzung von Grünflächen wider. 32 % der Frauen und 29 % der Männer in einer repräsentativen Umfrage von Menschen in deutschen Großstädten gaben bereits vor den pandemiebedingten Ausgangsbeschränkungen an, dass eine bessere Erreichbarkeit von Parks und Grünanlagen ein Anreiz wäre, städtische Grünanlagen mehr zu nutzen³. Auch gibt es im Vergleich deutscher Großstädte signifikante Unterschiede in der Verfügbarkeit von Grünflächen im Stadtgebiet (Taubenböck et al., 2021), welche sich auch auf nachbarschaftlicher Ebene wiederfinden (Wüstemann et al., 2017a).

Lange ist bekannt, dass insbesondere in Großstädten öffentliche Grünflächen ein rares Gut und obendrein noch ungleich verteilt sind (z. B. Kabisch und Haase, 2014). Während der COVID-19-Pandemie wurden die existierenden Ungleichheiten lediglich plötzlich sehr deutlich. Diese ungleiche Verteilung führt langfristig zu ungleichen Folgen für die persönliche Entwicklung sowie der körperlichen und psychischen Gesundheit (Maschewsky, 2001, S. 18). Der ungleiche Zugang zu Grünflächen für unterschiedliche Gruppen der Gesellschaft ist immanent und wird als ein kritisches Anliegen der Forschung zu Umweltgerechtigkeit untersucht (UG, z. B. Jennings et al., 2012; Rigolon, 2016; Zuniga-Teran und Gerlak, 2019; Schüle et al., 2019; Zhuang et al., 2022). Das heißt, im Allgemeinen haben einige Stadtteile Zugang zu mehr oder höherwertigen Grünflächen als andere (z. B. Jünger, 2021), was die räumliche Verteilung von Grünflächen zu einer Frage der sozialen Gerechtigkeit macht. Allerdings bieten nicht alle Grünflächen in der

³Forsa-Umfrage, 05.06.2015, „Urbanes Grün und Gesundheit“, zuletzt abgerufen 27.07.2023, <https://die-gruene-stadt.de/wp-content/uploads/2015/08/2015-forsa-umfrage-gruen-und-gesundheit.pdf>

Stadtstruktur einen ähnlichen salutogenen (Antonovsky, 1997), d. h. gesundheitsfördernden, Nutzen (z. B. Wheeler et al., 2015; Akpinar et al., 2016; Mears et al., 2020). Eine wesentliche Art von städtischen Grünanlagen mit hoher Relevanz für die öffentliche Gesundheit und den sozialen Zusammenhalt in Städten sind öffentlich zugängliche Grünflächen (Gómez-Baggethun et al., 2013; Ludwig et al., 2021). Sie sind eine unverzichtbare Ressource für Menschen, die keinen eigenen Garten haben (Poortinga et al., 2021), bieten aber auch denjenigen, die Zugang zu privatem Grün haben, wertvolle Dienste (Lin et al., 2014; Berdejo-Espinola et al., 2021).

Diese Faktoren unterstreichen die Bedeutung von Grünflächen für die städtische Umwelt. Dies geht so weit, dass die Ziele für nachhaltige Entwicklung (engl. *sustainable development goals*, *SDG*) der Vereinten Nationen die Zielvorgabe 11.7 „bis 2030 wird der allgemeine Zugang zu sicheren, inklusiven und zugänglichen Grünflächen und öffentlichen Räumen gewährleistet sein“ als Teil des SDG 11 „Städte inklusiv, sicher, widerstandsfähig und nachhaltig machen“ enthalten (UN General Assembly, 2015). Und doch stehen urbane Grünflächen durch immer intensiveres Stadtwachstum und Nachverdichtung unter Druck (Xu et al., 2018a), obwohl sie nachweislich die Auswirkungen der städtischen Verdichtung auf die Lebenszufriedenheit verringern (He et al., 2022). Der Analyse von Grünflächen im Stadtgebiet und insbesondere der raumquantitativen Untersuchung des Zugangs zu Grünflächen kommt somit eine zentrale Rolle der geographischen Forschung zu.

1.2 Forschungslücke

Trotz der inhärenten räumlichen Natur der Wohnumwelt und deren gesundheitlichen Auswirkungen werden einige bisherige Untersuchungen dazu als „im Grunde nicht räumlich“ (engl. *essentially aspatial*) kritisiert (Auchincloss et al., 2012, S. 2). Diese Kritik bezieht sich auf die starke räumliche Aggregation von Umweltbedingungen in epidemiologischen Studien, welche in diskreten räumlichen Einheiten wie beispielsweise administrativen Flächen oder Postleitzahlgebieten vorliegen und so lokale Phänomene nicht gut abbilden. Abbildung 1.1 zeigt beispielhaft die unterschiedliche Abstraktion von Umweltbelastungen aus punktuellen Emissionsquellen sowie eine Anordnung von administrativen Einheiten. Die Kritik an „im Grunde nicht räumlichen“ Untersuchungen bezieht sich hierbei auf stark aggregierte (Umwelt-) Daten wie in Abbildung 1.1 b), die im lokalen Kontext die tatsächliche Exposition gegenüber Umweltbelastungen bzw. -ressourcen stark überprägt und zu Effekten wie sog. „ökologischen Fehlschlüssen“ (engl. *ecological fallacy*, Robinson, 1950; Openshaw, 1984) führen kann. Ökologische Fehlschlüsse entstehen, wenn Zusammenhänge auf übergeordneter Ebene mit individuellen Zusammenhängen gleichgesetzt werden. Diekmann (2007, S. 134) stellt klar, auch wenn „auf der Kollektivebene ein positiver Zusammenhang zwischen X und Y [besteht], dann kann auf der Individualebene zwischen X und Y (a) ebenfalls ein positiver Zusammenhang, (b) gar kein Zusammenhang oder (c) sogar ein negativer Zusammenhang existieren“. Die willkürliche Anordnung und Größe der häufig verwendeten administrativen Raumeinheiten stellt eine potenzielle Fehlerquelle in derartigen Studien dar, die bereits mit widersprüchlichen Ergebnissen in Zusammenhang gebracht wurde

(Bolte et al., 2011). Für die Untersuchung der Verfügbarkeit von Grünflächen lässt sich hieraus eine Relevanz von räumlich hoch aufgelösten Daten über die Verfügbarkeit von Grünflächen in der Wohnumgebung der Menschen ableiten.

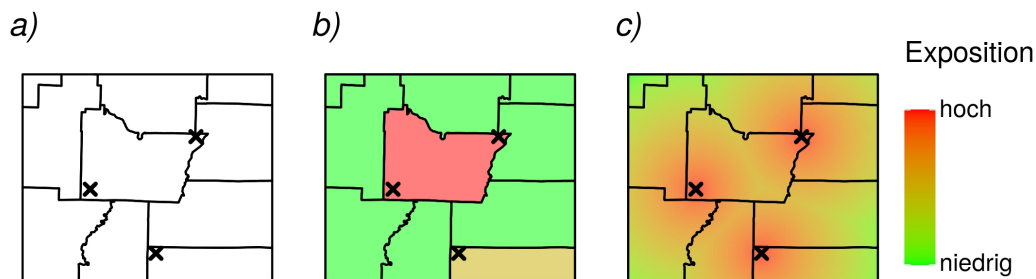


Abbildung 1.1: Unterschiedliche Raumkonzepte für die Abschätzung von Exposition gegenüber Umweltbelastungen (modifiziert nach Weigand et al., 2019). a) zeigt die Position von drei Emissionsquellen (x) von Umweltbelastungen im Raum sowie administrative räumliche Einheiten, z. B. Postleitzahlgebiete. Die Exposition der Menschen gegenüber Emissionsquellen beziehungsweise positiven Umweltressourcen sind in b) und c) auf unterschiedliche Arten modelliert. Während in b) die Exposition auf Basis der administrativen Einheit aggregiert wird, zeigt c) eine räumlich kontinuierliche Expositionsabschätzung.

Obwohl die Einflüsse räumlicher Aggregation auf die statistischen Modelle bereits lange unter dem Begriff des Problems der veränderbaren Gebietseinheit (engl. *modifiable areal unit problem*, *MAUP*) diskutiert werden (Openshaw, 1983; Fotheringham und Wong, 1991), ist die aggregierte Erfassung der räumlichen Verteilung von Umweltbelastungen bzw. Umweltressourcen in der Praxis häufig auf mangelnde Verfügbarkeit höher auflösender Daten zurückzuführen. Ebenfalls lässt sich ein gegensätzlicher Zusammenhang zwischen der Größe des Untersuchungsgebietes und der räumlichen Auflösung der Daten zu Umweltressourcen bzw. -belastungen feststellen (Weigand et al., 2019). Die Größe der Untersuchungsgebiete bisheriger Studien zur Verteilung von Grünflächen in Deutschland und international variiert dabei stark:

- Studiengebiete im Rahmen einzelner Städte (z. B. Heynen et al., 2006; Pham et al., 2012; Lakes et al., 2013; Kabisch und Haase, 2014; Dugord et al., 2014; Bertram und Rehdanz, 2015; Flacke et al., 2016; Apparicio et al., 2016; Pearsall, 2017; Banzhaf und Barrera, 2017; Voorde, 2017),
- wenige Städte (z. B. Villeneuve et al., 2012; Krekel et al., 2016; Sakieh et al., 2017; Wüstemann et al., 2017a) und
- auf Ebene von Nationen (z. B. Maas et al., 2006; Groenewegen et al., 2006; Mitchell und Popham, 2007, 2008; Berg et al., 2010).

Trotz der Relevanz wurde ein Mangel an geeigneten, zugänglichen, hochauflösenden, großräumigen Daten als zentraler limitierender Faktor für umfassendere und detailliertere Studien zu urbanem Grün identifiziert (Feltynowski et al., 2018). Frei zugängliche Geodaten von Erdbeobachtungssatelliten und aus anderen offenen Geodatenbanken haben jedoch das Potenzial, derartige Datenlücken zu schließen und die räumliche Abdeckung von Umweltgerechtigkeitsstudien zu

erweitern sowie deren räumliche Genauigkeit zu erhöhen. Sie ermöglichen es, Unterschiede in der Verfügbarkeit von Grünflächen über die städtischen Ballungsräume hinaus zu erfassen und somit Studien zur sozialen Gerechtigkeit auf größere Gebiete auszudehnen oder ihren räumlichen Detailgrad zu erhöhen. Viele Studien im Bereich des städtischen Grüns haben bereits Erdb Beobachtungsdaten verwendet, um Informationen über die Bodenbedeckung und -nutzung von urbanen Grünflächen abzuleiten. Daten zur vegetativen Landbedeckung und Vegetationsindizes etwa sind zugängliche Metriken zur Quantifizierung des Grünzustands eines Gebiets (z. B. Dewulf et al., 2016; Santos et al., 2016; Russette et al., 2021). Dabei tragen räumlich höher aufgelöste Bilder zu einer realistischeren Einschätzung des Grüns bei, insbesondere bei lokalen Anwendungen (Jimenez et al., 2022). Hier stellen die modernen, hochaufgelösten Satellitendaten des Sentinel-2 Sensors eine wichtige Datenquelle dar, welche sich von bisherigen großflächig frei verfügbaren Daten der Landsat Satelliten durch eine deutlich höhere räumliche Auflösung abheben. Jedoch steht eine Untersuchung der Eignung dieser Daten für den Zweck der Analyse von urbanen Grünflächen noch aus.

Neben der mangelnden Verfügbarkeit hochaufgelöster Daten zu Grünflächen ergibt sich aus der Literatur eine weitere Lücke im Kontext der Analyse städtischer Grünflächen. In den Untersuchungen zur Verfügbarkeit von Grünflächen wird häufig nicht zwischen verschiedenen Arten von Grünflächen unterschieden. In der umfangreichen Literatur gibt es große Unterschiede in der Art und Weise, wie städtisches Grün erfasst und analysiert wird (Kabisch, 2019). Studien, die Grünflächen analysieren, definieren diese häufig nicht ausreichend, was zu Inkompatibilitäten oder mitunter zu divergierenden Ergebnissen führen kann (Wolch et al., 2014; Taylor und Hochuli, 2017; Wüstemann et al., 2017b; Klompaker et al., 2018). Um dieses Problem zu lösen, können zusätzlich zu Satellitenbilddaten weitere Geodaten verwendet werden. Diese helfen detailliertere und semantisch differenzierte räumliche Daten von Grünflächen abzuleiten. Jüngste Studien haben beispielsweise den Einsatz von Datenfusion zwischen Satellitenbildern und anderen geographischen Informationen zur Ableitung detaillierter Landnutzungsinformationen demonstriert (Ludwig et al., 2021; Rosier et al., 2022). Hierfür kommen zunehmend Geodaten zum Einsatz, die von Freiwilligen zur Verfügung gestellt wurden (engl. *volunteered geographic information, VGI*, Goodchild, 2007). VGI liefern umfangreiche semantische Attribute der Landnutzung, die über die physische Beschreibung der Landbedeckung hinausgehen (Vargas-Munoz et al., 2021). Dies kann die Analysen von urbanen Grünflächen durch die Unterscheidung von Grüntypen wie Wäldern oder Parks von grünen landwirtschaftlichen Flächen erweitern.

Auch der kürzlich von Ludwig et al. (2021) entwickelte Ansatz zur semantisch detaillierten Beschreibung von urbanem Grün ist räumlich stark begrenzt. Jedoch zeigen städteübergreifende Analysen stark unterschiedliche Muster der Begrünung zwischen Städten (Kabisch et al., 2016; Wüstemann et al., 2017b; Zepp et al., 2020; Taubenböck et al., 2021). Viele Studien, die die gesundheitlichen Auswirkungen, die Verteilung und die Gerechtigkeit von Grünflächen untersuchen, konzentrieren sich darüber hinaus auf einzelne ausgewählte Städte und lassen die ländlichen Regionen um sie herum außer Acht. Dennoch leben in Deutschland etwa 31,9 % bis 50 % der

Bevölkerung außerhalb von urbanen Regionen, je nach Konzept der Urbanisierung, Datenlage, räumlichen Messeinheiten, Variablen oder verwendeten Schwellenwerten (Taubenböck et al., 2022). Angesichts der bekannten Varianz in der Verfügbarkeit von Grünflächen kann eine groß angelegte Ausweitung der Grünflächenbewertung auf ländliche Regionen eine breitere Perspektive für die Forschung zur sozialen Gerechtigkeit bieten. Eine großflächige Analyse von Methoden zur landesweiten Beschreibung der Verfügbarkeit von nachbarschaftlichen Grünflächen, unabhängig von arbiträren Grenzen, wie administrativen Stadtgrenzen, Städten bestimmter Größe und weiteren ist jedoch noch ausstehend.

1.3 Ziele der Arbeit und Forschungsfragen

Das übergeordnete Ziel der vorliegenden Dissertation besteht darin, methodische Schritte der systematischen Umwandlung umfassender Geodaten in relevante Geoinformationen für die großflächige Analyse von gesundheitsrelevanten Umwelteigenschaften aufzuzeigen und durchzuführen. Dies geschieht am Beispiel von Grünflächen in Deutschland. Diese Arbeit liegt damit an der Schnittstelle der Disziplinen Fernerkundung, Geoinformatik, Sozialgeographie und Umweltgerechtigkeitsforschung. Insbesondere sollen Potenziale großflächiger, moderner Geodaten für die Verbesserung der räumlichen und semantischen Auflösung von Geodaten erforscht werden sowie Analysen auf großflächige Untersuchungsgebiete ausgeweitet werden. Unter Zuhilfenahme modernster Methoden des maschinellen Lernens werden komplexe Merkmale der Landbedeckung und -nutzung erfasst und auf nationaler Ebene quantifiziert. Diese Entwicklungen sollen dazu beitragen bestehende Datenlücken zu schließen und Aufschluss über die Verteilungsgerechtigkeit einer wichtigen gesundheitlich relevanten Umweltressource zu bieten. Im Zentrum der Arbeit stehen drei Forschungsfragen:

1. Wie können moderne Satellitendaten von Sentinel-2 genutzt werden, um Landbedeckungsinformationen zu Grünflächen auch in komplexen Siedlungsstrukturen hochaufgelöst und flächendeckend abzuleiten?

Aufbauend auf bisherigen methodischen Ansätzen zur Landbedeckungsklassifikation stellt sich die Frage, welche Besonderheiten für die Klassifikation von hochaufgelösten Sentinel-2 Daten beachtet werden müssen. Die Beantwortung dieser Forschungsfrage zielt darauf ab, die Nutzbarkeit der Daten für die detaillierte und großflächige Beschreibung der Lebensumwelt der Menschen auch in komplexen Siedlungsstrukturen zu prüfen. Die entstehende Landbedeckungsklassifikation soll für die großräumige Analyse von urbanen Grünflächen auf Ebene der individuellen Nachbarschaft in allen bebauten Gebieten in Deutschland eingesetzt werden. Dabei sollen verschiedene Typen von Vegetationsklassen unterschieden werden. Es stehen vornehmlich räumliche und semantische Aspekte im Vordergrund, die die Qualität von Klassifikationen von Erdbeobachtungsdaten des Sentinel-2 Satelliten beeinflussen.

2. Ist es möglich, den Anteil öffentlich zugänglicher Grünflächen auf Nachbarschaftsebene mithilfe einer Kombination aus Satellitendaten und weiteren frei zugänglichen Geodaten hochgenau zu quantifizieren?

Im Kontext der Literatur zu positiven gesundheitlichen Einflüssen von Grünflächen im Siedlungsgebiet steht insbesondere der Grünflächenanteil in der Wohnumgebung der Menschen im Vordergrund. Darüber hinaus sind verschiedene Typen von urbanen Grünflächen unterschiedlich relevant für die Gesundheit. Hier stellt sich die Forschungsfrage, wie mithilfe großflächiger geographischer Daten aus diversen Quellen wie Satellitendaten und zusätzlichen frei zugänglichen Geodatenbanken relevante Geoinformationen abgeleitet werden können, um spezielle Typen von urbanem Grün großflächig zu beschreiben. Hierbei steht stets der räumlich lokale Bezug zur betroffenen Bevölkerung im Vordergrund, weswegen die Wohnumgebung in Form von Nachbarschaften zentrales Element zur Beantwortung der Forschungsfrage darstellt.

3. Wie gerecht sind grüne Landbedeckung sowie öffentliches Grün in der Wohnumgebung für die deutsche Bevölkerung verteilt?

Mit den im Rahmen von Forschungsfragen 1. und 2. erstellten Geodaten zur räumlichen Verteilung von Grünflächen und öffentlichen Grünanlagen werden im Rahmen dieser Forschungsfrage die Unterschiede der Verfügbarkeit von urbanem Grün in Abhängigkeit vom Typ untersucht. Genauer wird die Verteilungsungleichheit der Ressourcen „grüne Landbedeckung“ und „öffentliches Grün“ anhand von deskriptiven Statistiken und multivariaten Regressionsmodellen mit gemischten Effekten für die Beziehung zwischen nachbarschaftlichem Grünanteil und der demographischen Zusammensetzung quantifiziert.

1.4 Aufbau der Arbeit

Diese Dissertation behandelt die drei gestellten Forschungsfragen im Rahmen der folgenden Struktur:

In **Kapitel 2** werden die grundlegenden Konzepte eingeführt, auf denen diese Arbeit aufbaut, und es wird der Stand der Forschung präsentiert. Hierbei fungiert der Fokus der Arbeit – die Analyse der Grünflächenverteilung in Deutschland – stets als zentrales und verbindendes Element. Die Grundlagen setzen sich zusammen aus den Themenfeldern der Umweltgesundheits- und Umweltgerechtigkeitsforschung sowie den Inhalten der Erdbeobachtung und Geodatenverarbeitung. Darüber hinaus werden die notwendigen und relevanten Hintergrundinformationen für das Verständnis von aktuellen Methoden des maschinellen Lernens präsentiert, welche im Kontext der methodischen Ansätze dieser Dissertation angewandt werden.

In **Kapitel 3** werden die konzeptionellen Ansätze der Arbeit erläutert. Für jede der drei Forschungsfragen wird ein separater Versuchsaufbau konzeptualisiert, welcher mit Ablaufdiagrammen, grundlegenden Annahmen, sowie methodischer Ausrichtung detailliert wird. Zunächst wird die Analyse zur Eignung von Sentinel-2 Daten zur hochaufgelösten Landbedeckungsklassifikation

in Deutschland dargelegt. Es folgt die Entwicklung eines Modells zur Quantifizierung der Verfügbarkeit von öffentlichen Grünflächen. Dieses Kapitel gibt Aufschluss über die unterschiedlichen Anwendungen von diversen Geodaten und Methoden des maschinellen Lernens zur Beantwortung der ersten beiden Forschungsfragen. Im finalen Abschnitt wird schließlich die geographische Anwendung der beiden zuvor erstellten deutschlandweiten Datensätze dargelegt. Übergreifend stellt das Konzept der Dissertation somit eine umfangreiche Datenverarbeitungskette zum großflächigen Informationsgewinn dar.

In **Kapitel 4** werden die im Rahmen der Arbeit verwendeten Daten dargelegt sowie notwendige Vorprozessierungsschritte beschrieben. Unter anderem werden die genauen Schritte skizziert, mit denen die verwendeten Sentinel-2 Daten und daraus abgeleitete Bildmerkmale erstellt werden, die die Grundlage für die gesamte Arbeit darstellen. Darüber hinaus kommen weitere Datensätze zur detaillierten Beschreibung von semantischen Eigenschaften von Grünflächentypen aber auch zur Verortung der Bevölkerung zum Einsatz.

In **Kapitel 5** werden die Details der methodischen Vorgehensweisen erläutert. Hierbei stehen insbesondere die verwendeten Modelle des maschinellen Lernens im Vordergrund. Zusätzlich werden relevante Metriken präsentiert, welche zur Qualitätsevaluation der Ergebnisse herangezogen werden. Abschließend werden darüber hinaus die Methoden für die geographische Analyse der Verteilung von Grünflächen in der deutschen Bevölkerung dokumentiert.

In **Kapitel 6** werden die Ergebnisse der Arbeit dargelegt. Zunächst wird die Eignung von Sentinel-2 Bilddaten in Kombination mit modernen maschinellen Lernverfahren für die flächendeckende Landbedeckungsklassifikation in Deutschland mithilfe gängiger Genauigkeitsmaße evaluiert und eingeordnet. Darüber hinaus wird anhand eines Transfers die Übertragbarkeit der entwickelten Methodik auf einen späteren Zeitschritt demonstriert. Aufbauend auf den Ergebnissen der Landbedeckungsklassifikation werden anschließend das Resultat der Quantifizierung öffentlicher Grünflächen mittels Datenfusion aufgezeigt. Hierbei werden diverse Qualitätsmetriken verwendet, um die Güte der Modellschätzung zu verdeutlichen. Im letzten Schritt werden die diversen Ergebnisse zur Analyse der Verfügbarkeit von Grünflächen in Deutschland präsentiert. Hierbei werden verschiedene Aspekte beleuchtet, die zwischen flächen- und bevölkerungsgewichteten Analysen sowie in Bezug auf die demographische Komposition der Bevölkerung in Deutschland variieren.

In **Kapitel 7** werden die Ergebnisse aller Arbeitsschritte vor dem Hintergrund der anfangs definierten Forschungsfragen diskutiert. Sie werden in den Kontext der bestehenden Forschungsliteratur eingeordnet und vor dem Hintergrund der geographischen und gesellschaftlichen Implikationen erörtert.

Kapitel 8 bietet einen zusammenfassenden Gesamtüberblick über die drei thematischen Elemente der Arbeit. Hierin werden die vorgestellten Arbeiten abschließend in Bezug auf die Forschungsfragen und Ziele bewertet. Darüber hinaus werden Herausforderungen und mögliche Pfade für zukünftige Forschung skizziert.

Im Rahmen dieser Dissertation wurden drei Journalpublikationen, ein Konferenzbeitrag sowie eine Datenpublikation angefertigt. Inhalte dieser Dissertation basieren in Teilen auf diesen Veröffentlichungen:

- Weigand, M., Wurm, M., Dech, S., Taubenböck, H. (2019). „Remote Sensing in Environmental Justice Research – a Review“. In: *ISPRS International Journal of Geo-Information* 8.20, S. 28. <https://doi.org/10.3390/ijgi8010020>
- Weigand, M., Staab, J., Wurm, M., Taubenböck, H. (2020). „Spatial and semantic effects of LUCAS samples on fully automated land use/land cover classification in high-resolution Sentinel-2 data“. In: *International Journal of Applied Earth Observation and Geoinformation* 88, 102065. <https://doi.org/10.1016/j.jag.2020.102065>
- Weigand, M., Wurm, M., Droin, A., Stark, T., Staab, J., Rauh, J., Taubenböck, H. (2023). „Are public green spaces distributed fairly? A nationwide analysis based on remote sensing, OpenStreetMap and census data“. *Geocarto International*. <https://doi.org/10.1080/10106049.2023.2286305>
- Weigand, M., Staab, J., Wurm, M., Dech, S., Taubenböck, H. (2018). „Big Earth Data for Large Scale Mapping of Health Relevant Environmental Factors in Germany“. *EARSel 5th Joint Workshop „Urban Remote Sensing - Challenges & Solutions“*. Bochum, September 2018.
- Weigand, M., Staab, J., Wurm, M., Taubenböck, H. (2020). „Land Cover DE - Sentinel-2 - Germany, 2015“. *EOC Geoservice*. <https://doi.org/10.15489/1ccmlap3mn39>

Grundlagen

Im folgenden Kapitel werden die thematischen Grundlagen dargelegt, auf denen die Arbeiten dieser Dissertation aufbauen. Hierbei wird auf den aktuellen Stand der Forschung Bezug genommen. Zunächst wird die Relevanz von Grünflächen als wichtige Ressource in Siedlungsgebieten aufgezeigt sowie der Diskurs um die Gerechtigkeit der räumlichen Verteilung von Grünflächen vorgestellt. Darüber hinaus wird ausgeführt, wie moderne Erdbeobachtungsdaten und andere Geodaten genutzt werden können, um Grünflächen räumlich hochgenau zu beschreiben und zu kartieren. Anschließend werden Grundlagen und Methoden von maschinellen Lernverfahren erörtert, die in Kombination mit Erdbeobachtungsdaten großflächige Analysen von urbanem Grün ermöglichen.

2.1 Grünflächen als Ressource in der urbanen Umwelt

Grünflächen sind unbebaute und nicht verdichtete Flächen, die mit Vegetation bedeckt sind (Swanwick et al., 2003; Jorgensen und Gobster, 2010; Hunter und Luck, 2015). In Städten zählen dazu insbesondere Parks, Gärten, Wälder, Sportanlagen, aber auch Brachflächen oder freie Areale in Industriegebieten. Sie stellen einen wichtigen Teil der „urbanen Natur“ dar (Hartig et al., 2014). Außerhalb der Stadtgrenzen können Naturlandschaften, Wälder, Wiesen und dergleichen als natürliches Grün betrachtet werden. Je nach Art oder Zugänglichkeit dienen Grünflächen verschiedenen ökologischen Zwecken und erbringen unterschiedliche Ökosystemdienstleistungen (Jenerette et al., 2011; Gómez-Baggethun et al., 2013; Ensle und Kabisch, 2020).

Grünflächen wirken sich sowohl direkt als auch indirekt auf das Mikroklima in ihrer direkten Umgebung aus (Middel et al., 2014). Beispielsweise kühlt die Verdunstung von Wasser aus Pflanzen und Böden (Evapotranspiration) die Umgebungsluft und eignet sich damit gut, die zunehmenden Hitzeereignisse in dicht bebauten Siedlungsgebieten abzumildern. Der Effekt der Kühlung durch Evapotranspiration ist komplex und hängt stark von der allgemeinen Luftfeuchtigkeit und Temperatur ab (Oliveira et al., 2011). Im Vergleich zur bebauten Umgebung, ist die Luft in Parks jedoch im Durchschnitt ca. 1 °C kühler (Bowler et al., 2010). Hierbei ist der Anteil an Bäumen sowie großen zusammenhängenden Grünanlagen wie Parks oder Wäldern besonders wichtig (Bowler et al., 2010; Dugord et al., 2014). Zusammen mit nahegelegenen Wasserflächen haben Grünflächen so das Potenzial die hitzebedingte Mortalität zu verringern (Burkart et al., 2015). Neben der Auswirkung auf die Außentemperatur, konnte ebenfalls nachgewiesen werden,

dass sich Vegetationsflächen positiv auf die Innentemperatur in Gebäuden auswirken (Leichtle et al., 2023). Grünflächen werden daher auch als wirksames städteplanerisches Mittel zur Reduktion des sogenannten urbanen Hitzeinseleffekts (UHI, Oke, 1973) angesehen (z. B. Park et al., 2017; Targino et al., 2019).

Zusätzlich zum Einfluss auf die Lufttemperatur wirken sich Grünflächen positiv auf die Luftqualität aus. Die reinigende Wirkung bezieht sich hierbei vorrangig auf bestimmte Luftschadstoffe, wie z. B. Stickstoffdioxid (NO₂) und Feinstaub (Pugh et al., 2012). Grünflächen verringern die Schadstoffbelastung durch Streuung in der Luft sowie Ablagerung von Schadstoffen auf den Pflanzen (Hewitt et al., 2019). Vegetation in engen Straßenschluchten kann jedoch auch die Durchlüftung beeinträchtigen und somit zu lokalen Konzentrationen von Luftschadstoffen führen (Hartig et al., 2014).

Allgemein sind Grünflächen mit vielen Aspekten der physischen oder mentalen Gesundheit sowie des sozialen Wohlbefindens positiv assoziiert. Je höher der Grünanteil in der Wohnumgebung ist, desto besser ist die gesundheitliche Selbsteinschätzung der Menschen (Maas et al., 2006; Coppel und Wüstemann, 2017). Zudem ist die Morbidität der Anwohner geringer (Maas et al., 2009). Eine erhöhte Verfügbarkeit von Grünflächen in der Nachbarschaft steht in positivem Zusammenhang mit kardiovaskulärer Gesundheit, d. h. des Herz-Kreislauf-Systems (Villeneuve et al., 2012; Richardson et al., 2013). Grünflächen laden zu Bewegung und sportlicher Aktivität ein (Dewulf et al., 2016). Dabei wirkt sich sportliche Bewegung in Grünflächen stärker positiv auf die Erholung und den Stressabbau aus als weniger anstrengende Tätigkeiten wie Spaziergänge (Hansmann et al., 2007).

Eine erhöhte Verfügbarkeit und Qualität von Grünflächen in der Nachbarschaft steht in positivem Zusammenhang mit mentaler Gesundheit (Richardson et al., 2013) und geringeren Beschwerden durch psychische Belastungen (Pope et al., 2015). Stress und Kortisolausstoß sind signifikant verringert, je höher die Menge der Grünflächen in der Wohnumgebung ist (Thompson et al., 2012, 2014). Dieser Effekt ist ebenso zu für Wohnhäuser verzeichnen, die einen Blick auf mehr und diverseres Grün bieten (Honold et al., 2016). Eine kürzlich erschienene Studie von Helbich et al. (2020) legt nahe, dass die Verfügbarkeit von urbanem Grün die Suizidrate im späteren Leben reduziert. Städtische Grünflächen verbessern das Wohlbefinden, insbesondere in Gebieten mit geringer Grünausstattung, indem sie die präfrontale Aktivität bei der Verarbeitung negativer Emotionen reduzieren (Tost et al., 2019).

Bei Kindern wirken sich Grünflächen positiv auf das gesundheitliche und soziale Wohlbefinden aus (Dijk-Wesselius et al., 2018). Kinder, die in Nachbarschaften mit mehr Grünflächen aufwachsen, haben ein geringeres Risiko für psychiatrische Erkrankungen (Engemann et al., 2019). Grünflächen stehen zudem in Zusammenhang mit besseren schulischen Leistungen (Browning und Locke, 2020) und verringerten berufsbedingten Stressbelastungen junger Erwachsener (Herrera et al.,

2018). Grünflächen können auch die negativen gesundheitlichen Auswirkungen belastender Lebensereignisse abschwächen (Berg et al., 2010) und die Risiken von Depression in Jugendlichen reduzieren (Bray et al., 2022).

Der Einfluss des gesellschaftlichen Status auf die Sterblichkeit ist in Gebieten mit mehr Grünflächen signifikant niedriger als in Gebieten mit geringem Grünanteil (Mitchell und Popham, 2008). Je näher Menschen an Grünflächen wohnen, desto höher ist die Wahrscheinlichkeit, dass sie häufiger Grünflächen besuchen. In Anbetracht der diversen positiven gesundheitlichen Auswirkungen ist es nicht verwunderlich, dass Grünflächen einen wichtigen Faktor für die Wahl des Wohnortes darstellen (Wurm et al., 2019c). Hartig et al. (2014) identifizieren die Distanz, Erreichbarkeit und weitere Faktoren als wichtige Einflüsse auf die Wechselwirkung zwischen der natürlichen Umwelt und der Nutzungshäufigkeit oder -dauer.

Die Anzahl der Studien, die den Zusammenhang von Grünflächen und Gesundheit betrachten, nehmen bereits seit einigen Jahren zu (Hartig et al., 2014). Die genauen kausalen Wirkungswege jedoch sind komplex und direkte Rückschlüsse somit nicht oder nur bedingt zulässig (Lee und Maheswaran, 2011). Auch sind die Effektgrößen der Umwelteinflüsse auf die Gesundheit der Menschen trotz zunehmender Evidenz gering und häufig stark überlagert von anderen Einflussfaktoren, wie Einkommen, Arbeitsverhältnis, Bildung oder individuellem Verhalten (Hartig et al., 2014). Dennoch spielt die räumliche Komponente, d. h. der Verfügbarkeit und Zugänglichkeit von Grünflächen in der Lebensumwelt der Menschen, eine zentrale Rolle für die gesundheitliche Wirkung von Grünflächen (siehe Abbildung 2.1).

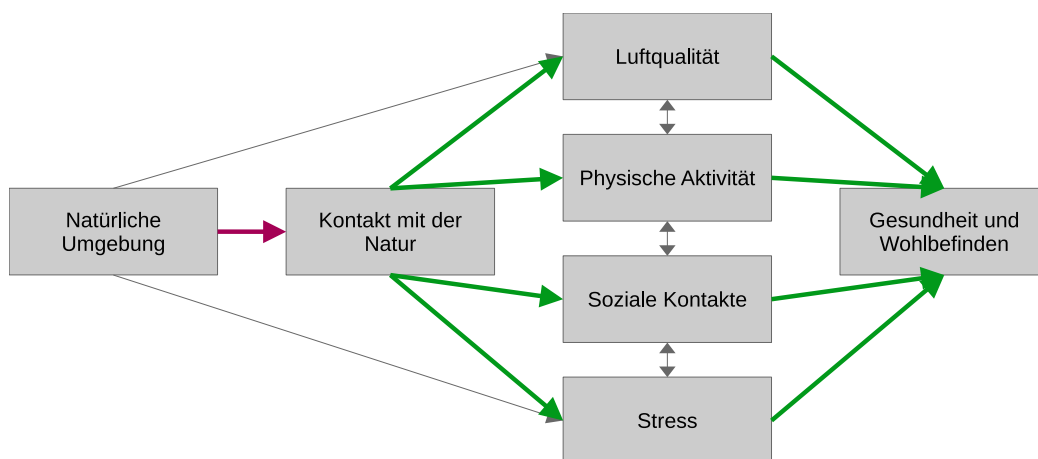


Abbildung 2.1: Wirkungswege von Grünflächen auf die Gesundheit nach Hartig et al. (2014). Die Autoren unterscheiden zwischen zwei Effektmodifikatoren. Erstens, die Effektmodifikation durch beispielsweise die Distanz und andere Erreichbarkeitsfaktoren, das Wetter, die wahrgenommene Sicherheit der Umgebung (lila). Zweitens die Effektmodifikation durch demographische bzw. gesellschaftliche Faktoren wie Geschlecht, Alter, sozioökonomischer Hintergrund und Gesellschaftskontext (grün). Da vier der Wirkungswege durch den Kontakt zur Natur beeinflusst werden, schreibt dieses Modell der räumlichen Effektmodifikation eine besondere Bedeutung zu.

Die gesundheitlichen Vorteile werden mehr vom Typ der Grünfläche beeinflusst und nur schwach von der Menge an Grün alleine (Picavet et al., 2016). Je besser die wahrgenommene Zugänglichkeit von öffentlichen sowie die Verfügbarkeit von privaten Grünanlagen ist, desto höher auch die gesundheitliche Selbsteinschätzung der Bewohner (Poortinga et al., 2021). Nicht alle Typen von Grünflächen wirken sich jedoch gleichermaßen positiv auf die Gesundheit aus (z. B. Wheeler et al., 2015; Mears et al., 2020). Beispielsweise fanden Dennis et al. (2018) signifikante Unterschiede zwischen verschiedenen Grüntypen im Hinblick auf Gesundheitsindikatoren wie „potenziell verlorene Lebensjahre“ (engl. *years of potential life lost*). Akpinar et al. (2016) zeigten, dass urbane Grünflächen und Wälder sich positiv auf die Gesundheit auswirkten. Äcker, Weiden oder Brachflächen hingegen wirkten sich nicht signifikant positiv auf die Gesundheit der Menschen aus. In Siedlungen nehmen Grünflächen viele Formen an. Straßenbegleitgrün, bewachsene Brachflächen, Grünanlagen im Kontext von Gewerbe- und Industrieanlagen, stadtnahe Wälder, Auen und Wiesen, öffentliche Grün- und Parkanlagen oder private Gärten sind Beispiele für verschiedene Typen von Grünflächen im nahen Umfeld von Siedlungen und Städten. Öffentliches Grün stellt eine besonders wichtige Ressource in Städten dar. Die natürliche Umgebung wird im Stadtgebiet als besonders ästhetisch empfunden (Sang et al., 2016).

Die COVID-19-Pandemie hat verdeutlicht, dass öffentliches Grün wichtig ist. Insbesondere in Phasen strikter Lockdowns stiegen die Probleme psychischer Gesundheit wie beispielsweise Ängste, Verwirrung oder Stress deutlich an (World Health Organization, 2020). Kurz nach Beginn der Pandemie in Deutschland, wurden weitgreifende Einschränkungen der Bewegungsfreiheit für Menschen beschlossen^{4, 5, 6}. Öffentliche Grünanlagen wie Parks und nahegelegene Wälder waren daraufhin für viele Menschen, die nicht über einen privaten Garten verfügen können, zentrale Orte, um Zeit im Freien zu verbringen. Dies verdeutlicht die wichtige Rolle von öffentlichen Grünflächen im Alltag der Bevölkerung. Doch auch im Kontext der COVID-19-Pandemie ist die Nutzungshäufigkeit von Grünflächen stark beeinflusst von Umweltfaktoren wie Grünanteil in der Umgebung aber auch von sozioökonomischen Faktoren (Uchiyama und Kohnsaka, 2020).

Auch die Ausstattung von öffentlichen Grünflächen hat einen Einfluss auf deren gesundheitliche Relevanz. Generell zeigen Mitchell und Popham (2007) zwar den Zusammenhang auf, dass die Gesundheit der Menschen besser sei, je mehr Grünflächen zur Verfügung stehen, dieser ist jedoch abhängig zur Urbanität. Die Autoren führen als möglichen Erklärungsansatz an, dass sowohl Qualität als auch Quantität eine wichtige Rolle für den gesundheitlichen Zusammenhang spielen. Das Rahmenwerk von Lachowycz und Jones (2013) hebt hervor, dass die kausalen Pfade der menschlichen Gesundheit nur zum Teil durch die Exposition gegenüber Grünflächen erklärt werden.

⁴Pressemitteilung des Bayerisches Staatsministeriums für Wohnen, Bau und Verkehr vom 23.03.2020, zuletzt abgerufen am 27.07.2023: <https://www.stmb.bayern.de/med/aktuell/archiv/2020/ausgangsbeschränkung>

⁵Dritte Verordnung zur Bekämpfung des Corona-Virus vom 14. März 2020, Bundesland Hessen, zuletzt abgerufen 27.07.2023: <https://www.rv.hessenrecht.hessen.de/bshe/document/jlr-CoronaVVHE3V3P1>

⁶Pressemitteilung der Senatskanzlei Berlin: Eindämmung der Coronavirus – Berliner Senat beschließt weitgehende Kontaktbeschränkungen vom 22.03.2020, zuletzt abgerufen 27.07.2023: <https://www.berlin.de/rbmskzl/aktuelles/pressemitteilungen/pressemitteilung.910268.php>

Darüber hinaus stellen jedoch die Merkmale dieser Grünflächen einen potenziell moderierenden Faktor für diesen Pfad zu physischen und psychischen Gesundheitsvorteilen dar. Sie beeinflussen die Möglichkeit, Grünflächen zu nutzen, die potenzielle Motivation, sie zu nutzen, und die Einfachheit der Nutzung. In einer kürzlich veröffentlichten Studie zeigten Kabisch et al. (2021), dass ältere Parkanlagen das Risiko für Herz-Kreislaufkrankungen für Senioren deutlich senken. Dieser Effekt ist für jüngere, modernere Parkanlagen nicht signifikant. Ähnlich dazu fanden Vries et al. (2013), dass die Qualität von Grünflächen einen Einfluss auf deren gesundheitliche Wirkung hat. Ungenutzte Brachflächen im Stadtgebiet stehen, entgegen den positiven Effekten von gepflegten Grünflächen, im Zusammenhang mit schlechterem Wohlbefinden der Bewohner (Krekel et al., 2016). Das verdeutlicht die Relevanz der Landnutzung von mit Grün bedeckten Flächen im Stadtgebiet.

Im Gegensatz zu öffentlichem Grün stehen private Grünflächen nicht der gesamten Bevölkerung gleichermaßen zur Verfügung. Bei privaten Grünflächen handelt es sich um Privatgärten als Teil eines privaten Grundstücks, oder auch Grünanlagen, deren Zugang teilweise oder vollständig versperrt ist. Dazu zählen beispielsweise Kleingarten- oder Sportanlagen, welche nur für Mitglieder zugänglich sind. Poortinga et al. (2021) untersuchten die Vorteile von privaten und öffentlichen Grünflächen in Bezug auf die Auswirkungen der Ausgangsbeschränkungen aufgrund der COVID-19-Pandemie in Großbritannien. Sie kommen zu dem Schluss, dass die Verfügbarkeit sowohl von öffentlichen als auch privaten Grünflächen die negativen Effekte der einschränkenden Maßnahmen auf das Wohlbefinden der Menschen abschwächen. Hierbei kompensiert verfügbares öffentliches Grün ein Fehlen von (ausreichend) privatem Grün. Gleichzeitig kann umgekehrt privates Grün die positiven Vorteile von fehlenden öffentlichen Grünflächen für die mentale Gesundheit ausgleichen.

Grünflächen im urbanen Raum stellen erwartungsgemäß eine begrenzte Ressource dar. Darüber hinaus, variiert das Angebot an Grünflächen über verschiedene räumliche Skalen. Im Vergleich hunderter europäischer Städte identifizierten Wolff und Haase (2019) einen klaren Nord-Süd-Trend der Verfügbarkeit von urbanen Grünflächen. In nordeuropäischen Städten stehen deutlich mehr Grünflächen für die Bevölkerung zur Verfügung als in Südeuropa und der Balkanregion. Ähnliche Ergebnisse fanden Kabisch et al. (2016) in Bezug auf die Verfügbarkeit von öffentlichen Grünanlagen mit mehr als zwei Hektar in der Wohnumgebung. Die Autorinnen und Autoren weisen jedoch auch darauf hin, dass großräumige, für die Dimension von Analysen auf Ebene eines Kontinentes angelegte, Indikatoren lediglich ein breites Gesamtbild aufzeigen können, nicht aber lokale Phänomene und Auswirkungen beleuchten. Ähnlich finden auch Taubenböck et al. (2021) große Unterschiede der Grünanteile in deutschen Städten. Hierbei fällt auf, dass die ermittelten Werte je Stadt stark abhängig sind von der gewählten räumlichen Aggregationsmetrik. Im Vergleich acht europäischer Städte identifizierten Zepp et al. (2020) zwei zentrale Probleme in der Analyse des Grünflächenangebots. Einerseits stellen administrative Grenzen von Städten häufig willkürliche räumliche Einheiten dar, welche nicht mit den eigentlichen Siedlungsflächen übereinstimmen (vgl. auch Taubenböck et al., 2019). Andererseits existiert eine große Heterogenität in Bezug auf geographischen Daten zu Grünflächen.

Als Richtwert für die Mindestmenge an Grünflächen in der Wohnumgebung legt die Weltgesundheitsorganisation (engl. *World Health Organization, WHO*) 0,5 bis 1 Hektar in 300 Meter um den Wohnort fest (World Health Organization, 2017). Barrierefreier Zugang zu öffentlichen Grünflächen gilt als wichtiger Aspekt für eine seniorengerechte Stadtentwicklung (World Health Organization, 2007), denn Senioren, die die Verfügbarkeit von Grünflächen als ausreichend einschätzen und diese für zugänglich halten, suchen derartige Grünflächen auch häufiger auf (Enssle und Kabisch, 2020). Trotz der hohen Relevanz von Grünflächen und anderen salutogenen Ressourcen in der Wohnumgebung und im urbanen Raum, spielen weitere Faktoren, wie Umweltbelastungen, das soziale Umfeld durch die Familie oder Nachbarn und eine umfassende Versorgung mit medizinischen Einrichtungen eine große Rolle für die Gesundheit der Bevölkerung (Richardson und Mitchell, 2010; Claßen et al., 2012).

2.2 Gerechtigkeit der räumlichen Verteilung von Grünflächen

Vor dem Hintergrund der zuvor ausführlich beschriebenen positiven gesundheitlichen Auswirkungen von Grünflächen stellt sich die Frage, wie gerecht der Zugang zu Grünflächen in der Gesellschaft verteilt ist. Häufig sind Grünflächen räumlich ungleich verteilt und deshalb weniger zugänglich für bestimmte Gesellschaftsgruppen (z. B. Kabisch und Haase, 2014). Darüber hinaus haben Grünflächen eine besondere Relevanz für unterschiedliche Teile der Gesellschaft. Beispielsweise ist der Zusammenhang zwischen Grünflächenverfügbarkeit und Morbidität besonders ausgeprägt für Kinder und Menschen sozioökonomisch benachteiligter Gruppen (Maas et al., 2009). Generell ist die Frage der räumlichen Verteilung von Grünflächen Teil der größeren Umweltgerechtigkeitsdiskussion, die im Folgenden näher erläutert und in den historischen Kontext gesetzt wird.

Aus lokalen Sozial- und Bürgerrechtsbewegungen in den 1970er und 1980er Jahren in den USA entstand eine wissenschaftliche Diskussion zur ungerechten Verteilung von Umweltbelastungen und im erweiterten Sinn auch Umweltressourcen (Bolte et al., 2011; Rigolon et al., 2019). Das Thema der Umweltgerechtigkeit (UG, engl. *environmental justice*), manchmal auch umweltbezogene Gleichheit (engl. *environmental equity*) oder umweltbezogener Rassismus (engl. *environmental racism*), befasst sich in diesem Kontext mit der räumlichen und sozialen Verteilung von Umweltbelastungen auf soziale Gruppen oder Regionen (Maschewsky, 2001, S. 17). Dabei stehen verschiedene Aspekte im Zentrum der Umweltgerechtigkeitsforschung, darunter die räumliche und zeitliche Verteilung von Umweltbelastungen sowie die überproportionale Belastung von häufig sozial, ökonomisch oder politisch benachteiligten Personengruppen. Diese Ungleichheit wird dabei als ungerecht bewertet da sie die Forderung nach der Angleichung von Lebensverhältnissen verletzt, die z. B. in Artikel 3 Absatz 1 des Grundgesetzes der Bundesrepublik Deutschland „*Alle Menschen sind vor dem Gesetz gleich*“ verankert ist (Maschewsky, 2001, S. 17). Im Kontext der Umweltgerechtigkeit gibt es jedoch keine klare Definition was genau „gerecht“

ist. Vielmehr gibt es verschiedene Konzepte, z. B. eine gerechte Verteilung aller Umweltressourcen und Umweltbelastungen, eine Verteilung gemäß dem Beitrag jeder Person in einer Gesellschaft, sowie eine Verteilung gemäß der individuellen Bedürfnisse (Bolte et al., 2011).

Die vorliegende Arbeit steht im Kontext der Umweltgerechtigkeit, da die Verfügbarkeit von Grünflächen als räumliches Phänomen im Siedlungsgebiet inhärent ungleich verteilt sind. Dabei muss Ungleichheit jedoch nicht zwingend ungerecht sein. Ungleichheiten sind beispielsweise nicht ungerecht, wenn sie das Ergebnis aus freier Wahl sind. Ungerecht sind Ungleichheiten wiederum, wenn Menschen ohne Notwendigkeit benachteiligt werden, oder dies änderbar wäre (Resnik und Roman, 2007; Schröder-Bäck, 2012). Grundsätzlich lassen sich vier Alltags-theoretische Gerechtigkeitskonzepte unterscheiden, die Chancengerechtigkeit, Verteilungsgerechtigkeit, Vorsorgegerechtigkeit und Verfahrensgerechtigkeit. Maschewsky (2001, S. 42) definiert diese wie folgt:

- Chancengerechtigkeit beschreibt die „Gleichheit der Chancen und Risiken zu Beginn eines gesellschaftlichen Prozesses, z. B. bezüglich des Bildungszugangs vor der Erwerbsphase“,
- Verteilungsgerechtigkeit beschreibt die „Gleichheit der Vorteile und Nachteile am Ende eines gesellschaftlichen Prozesses, z. B. des Einkommens nach der Erwerbsphase“,
- Vorsorgegerechtigkeit beschreibt die „Vermeidung des Entstehens unnötiger Risiken durch Beendigung oder Veränderung bestimmter Handlungen, z. B. durch Verbot von Hochrisiko-Technologien“,
- Verfahrensgerechtigkeit beschreibt die „Gleichbehandlung verschiedener Personen/Gruppen bei der Einflussnahme auf soziale Prozesse, z. B. bei der juristischen Anfechtung politischer Entscheidungen“.

Weitere Literatur befasst sich im Rahmen der Gerechtigkeit des Zugangs zu Grünflächen beispielsweise mit Verteilungsgerechtigkeit, Verfahrensgerechtigkeit und interaktioneller Gerechtigkeit (Low, 2013; Köckler, 2017; Rigolon et al., 2019; Enssle und Kabisch, 2020). Letztere befasst sich mit der (Un-) Gerechtigkeit, die entsteht, wenn manche Gesellschaftsgruppen bestimmte Einrichtungen, z. B. öffentliche Toiletten oder Parkbänke, aufgrund bestimmter Einschränkungen nicht oder nur limitiert nutzen können (Alidoust und Bosman, 2015; Rigolon et al., 2019). Im Kontext geographischer Analysen steht vor allem die Verteilungsgerechtigkeit im Vordergrund (Claßen et al., 2012), das bedeutet der Zustand, dass Menschen unabhängig von ihrer sozio-ökonomischen Lage den gleichen räumlichen Zugriff auf bestimmte Umweltressourcen haben und Umweltbelastungen ebenfalls gleich verteilt sind (Rigolon et al., 2019). Diverse Aspekte der Umweltgerechtigkeitsforschung greifen auch in andere Forschungsfelder, beispielsweise zu Ökosystemdienstleistungen von urbanen Grünflächen (Calderón-Argelich et al., 2021).

Grundsätzlich nimmt Grün im Siedlungskörper ab, je dichter bebaut wird. Eine absolute Gleichverteilung des Zugangs von Grünflächen ist daher ausgeschlossen. Die räumliche Ungleichverteilung in europäischen Städten führt auch zu sozialer Ungleichheit in Bezug auf Grünflächenverfügbarkeit, wobei sozioökonomisch schwächer gestellte Gesellschaftsgruppen zusätzlich benachteiligt sind.

Diese Beobachtungen sind konsistent über mehrere Studien hinweg (siehe Schüle et al., 2019, für einen Forschungsüberblick). Ungerechtigkeit entsteht gemäß obengenannter Definitionen dann, wenn bestimmte soziale Gruppen systematisch besseren Zugang zu (ausreichend) Grünflächen haben als andere. Häufig sind dies sozial besser gestellte Gruppen (Claßen et al., 2012; Kabisch und Haase, 2014; Zuniga-Teran und Gerlak, 2019; Jünger, 2021). Darüber hinaus entsteht Ungerechtigkeit in Bezug auf Umwelteinflüsse einerseits vorab durch Diskriminierungseffekte gegenüber der lokalen Bevölkerung, etwa durch fehlende Einbindung in Standortentscheidungen, sowie Markteffekte, die Kosten-Nutzen-Vergleiche zwischen möglichen Standorten beeinflussen. Auch nachträglich kann Ungerechtigkeit in Form von Selektions- oder Drifteffekten entstehen, wenn durch die Verschlechterung der Wohnumgebung bestimmte Bevölkerungsgruppen weg- und soziale Randgruppen zuziehen (Maschewsky, 2001, S. 17).

Generell zeigt eine Vielzahl an Studien, dass die räumliche Verteilung von Grünflächen nicht gleichmäßig ist (z. B. Claßen et al., 2012; Jennings et al., 2012; Rigolon, 2016; Zuniga-Teran und Gerlak, 2019). Grünflächen sind eine der am häufigsten untersuchten Variablen in der Umweltgerechtigkeitsforschung (Zhuang et al., 2022). Zentraler Aspekt hierbei ist, dass manche Menschen Zugang zu weniger oder minderqualitativeren Grünflächen haben als andere (Jünger, 2021). Damit besteht auch das Potenzial für eine soziale Ungleichverteilung von Grünflächen. Häufig haben schlechter Verdienende und Minderheiten weniger Zugang zu Grünflächen wodurch soziale Ungerechtigkeit entsteht (Heynen et al., 2006; Mitchell und Popham, 2008; Pham et al., 2012; Lakes et al., 2013; Wüstemann et al., 2017a; Jünger, 2021). Weitere sozioökonomische Gruppen, die in diesem Kontext in der Literatur häufig betrachtet werden, sind beispielsweise Kinder unter 18 Jahren (z. B. Rehling et al., 2021) oder älteren Menschen ab 65 Jahren (z. B. Artmann et al., 2019; Kabisch und Haase, 2014).

Auf theoretischer Ebene gibt es verschiedene Modelle, die Umweltgerechtigkeit konzeptionell erfassen (Bolte et al., 2012b). Das Modell von Bolte et al. (2012a) schreibt lokalen Ressourcen in der Lebensumwelt, wie beispielsweise Grünflächen und öffentliche Freiräume, neben lokalen Umweltbelastungen zentrale Bedeutung im Zusammenhang zwischen sozialer Lage, Umwelt und Gesundheit zu. Auch andere Modelle beziehen den lokalen Kontext der Wohnumgebung oder Nachbarschaft als relevante Bezugsgröße für die Untersuchung der Umweltbelastungen und -ressourcen mit ein (z. B. Morello-Frosch und Shenassa, 2006; Voigtländer et al., 2011; Bolte et al., 2011). Der kleinräumigen Betrachtung der Umwelt kommt somit eine wichtige Rolle im Kontext der Umweltgerechtigkeitsforschung zu. Allen Modellen ist gemein, dass sie Umweltbelastung auf Ebene der Lebensumwelt des Individuums einbeziehen. Hieraus lässt sich für die geographische Anwendung eine Notwendigkeit nach der detaillierten räumlichen Beschreibung der individuellen Lebensumwelt ableiten, die im Widerspruch zu häufig gewählten räumlich stark aggregierten Datensätzen steht (Auchincloss et al., 2012). Die Analysen der hier vorliegenden Arbeit leisten einen Beitrag auf lokaler Ebene in Form der Qualität der Wohnumgebung zu bieten. Diese lokale Ebene befindet sich konzeptionell zwischen dem Makrolevel auf der einen Seite, welches umweltpolitische Dimensionen und Markteffekte umfasst und auf der anderen

Seit dem Mikrolevel, welches individuelle Einflüsse, wie die körperliche Prädisposition oder die sozioökonomische Situation der Menschen einbezieht (Soobader et al., 2006). Um die Verteilungsgerechtigkeit von Grünflächen zu quantifizieren, ist es notwendig, die Umweltressourcen und -belastungen zu quantifizieren, sowie detaillierte sozioökonomische Daten zur Bevölkerung zu kombinieren. Geoinformationen, wie sie beispielsweise aus Erdbeobachtungsdaten abgeleitet werden können, und Geoinformationssysteme (GIS) sind dabei hilfreiche Werkzeuge für die Evaluation von unterschiedlichen Umweltbedingungen sowie deren gesundheitliche Auswirkungen (Bolte et al., 2011).

2.3 Fernerkundung zur Erfassung von Grünflächen

Gerade vor dem Hintergrund der Komplexität und Dynamik von urbanen Räumen eignen sich Satellitendaten besonders gut für die Erfassung gesundheitsrelevanter Umweltbedingungen wie Grünflächen (Taubenböck et al., 2020). Satellitengestützte Erdbeobachtung mithilfe von Fernerkundungssensoren ermöglicht die flächendeckende Beschreibung der physischen Eigenschaften der Erdoberfläche (Egan, 2003, S. 1). Bei dieser Technologie werden elektromagnetische Strahlen gemessen, welche von den Objekten an der Erdoberfläche emittiert oder reflektiert werden. Der Vorteil dieses Verfahrens ist die Möglichkeit jederzeit unter objektiven Bedingungen wiederholbare quantitative Messungen zum Zustand der Erdoberfläche durchführen zu können (Taubenböck et al., 2010, S. 92). Man unterscheidet grundsätzlich zwischen zwei verschiedenen Aufnahmearten: aktive und passive Verfahren (Jones und Vaughan, 2010, S. 92).

Bei aktiven Fernerkundungsverfahren werden elektromagnetische Strahlen aktiv vom Sensor ausgesendet, welche dann nach der Interaktion mit den zu erfassenden Objekten, z. B. durch Reflexion, wieder gemessen werden können. Zu den aktiven Fernerkundungsverfahren im Kontext der Erdbeobachtung zählen vor allem RADAR (engl. *Radio detection and ranging*) und LIDAR (engl. *Light detection and ranging*).

Passive Verfahren nutzen hingegen elektromagnetische Strahlung aus anderen Quellen, beispielsweise Sonnenlicht (Egan, 2003, S. 44). Diese Verfahren messen das gestreute bzw. reflektierte Licht in verschiedenen Wellenlängen und erstellen so ein Abbild der Erde, ähnlich zu digitalen Kameras. Ein gängiges Verfahren passiver Fernerkundung ist die sogenannte multispektrale Fernerkundung. Hierbei wird die Intensität einzelner Bereiche des Lichtspektrums gemessen. Diese Abschnitte des Spektrums korrespondieren beispielsweise im sichtbaren Bereich des Lichts mit rotem, grünem und blauem Lichtanteilen. Jedoch können auch weitere, für Menschen unsichtbare Wellenlängenbereiche des Lichtes wie Infrarotstrahlung gemessen werden, welche wertvolle Informationen bieten, um unterschiedliche Eigenschaften der Erdoberfläche zu beschreiben.

Derzeit sind verschiedene satellitengestützte Sensoren zur zivilen Erdbeobachtung im Einsatz. Diese unterscheiden sich generell anhand ihrer Aufnahmeeigenschaften nach der räumlichen und spektralen Auflösung, der Wiederkehrrate sowie der Aufnahmegeometrie und Schwadbreite. Die

Vielzahl der Sensoren bietet eine breite Datenbasis für diverse Anwendungen. Dennoch eignen sich nicht alle Sensoren für jede Anwendung. Im Folgenden werden die wichtigsten Sensoren vorgestellt, welche im Kontext der Erfassung und Analyse von Grünflächen Relevanz haben.

MODIS⁷ (engl. *Moderate Resolution Imaging Spectroradiometer*) ist ein multispektraler Sensor an Bord der Terra und Aqua Satelliten der amerikanischen Luft- und Raumfahrtbehörde (engl. *National Aeronautics and Space Administration, NASA*). Der Sensor wurde konzipiert für eine hohe radiometrische Auflösung (12 Bit) in insgesamt 36 spektralen Bändern und ist im Jahr 1999 (Terra) bzw. 2002 (Aqua) gestartet. Die räumliche Auflösung beträgt 250 Meter bis 1 Kilometer bei einer Breite der Aufnahme (Schwadbreite) von 2330 km⁸. Durch die sehr niedrige Wiederkehrdauer von 1 bis 2 Tagen, werden auch kurzfristige Änderungen schnell erfasst (Kramer, 2002, S. 383 ff.). MODIS eignet sich daher besonders gut für großflächige oder globale Analysen von Landoberflächenveränderungen.

Eine weitere Serie von multispektralen Sensoren deren Daten häufig für die Erfassung von Grünflächen zum Einsatz kommen ist die Landsat-Flotte. Beginnend im Jahr 1972 mit Landsat-1 bis hin zu Landsat-9, der 2021 gestartet wurde, stellen die Satelliten der Landsat-Serie eine lange Zeitreihe von multispektralen Satellitenbildern in mittlerer bis hoher räumlicher Auflösung bereit (Landsat-1 bis -3 MSS 30 m - 80 m, Landsat-4 bis -5 TM 30 m - 120 m, Landsat-7 ETM+ 15 m - 60 m, Landsat-8 bis -9 OLI/TIRS 15 m - 100 m, Kramer, 2002, S. 440 ff.). Die verschiedenen spektralen Bänder der Landsat-Sensoren (siehe Abbildung 2.2 unten) erlauben Anwendungen der Bilddaten in diversen Bereichen, u. a. der Erfassung von Grünflächen (z. B. Chang et al., 2021). Besonders der freie Zugang zu Landsat-Daten seit 2009 stellt einen wichtigen Grund für den Erfolg dieser Daten für die Analyse von Landbedeckungsinformationen dar (Harris und Baumann, 2015; Shahtahmassebi et al., 2021).

Im Rahmen des Copernicus-Programms der europäischen Raumfahrtagentur (engl. *European Space Agency, ESA*) werden seit 2014 unterschiedliche Erdbeobachtungssatelliten betrieben. Dazu zählt Sentinel-2, welcher hochauflösende Aufnahmen im Bereich des optischen und infraroten Lichtspektrums bereitstellt⁹. Bilddaten der multispektralen Instrumente (engl. *multi-spectral instrument, MSI*) an Bord der Sentinel-2 Satelliten erreichen eine räumliche Auflösung von bis zu 10 m × 10 m und eine Aufnahmebreite von ca. 290 km sowie Wiederkehrrate von 5 Tagen. Abbildung 2.2 vergleicht die Spektralkanäle des Sentinel-2 MSI Sensors mit denen der Landsat-7 und Landsat-8 Missionen. Es wird deutlich, dass die Spektralkanäle im sichtbaren Licht (ca.

⁷NASA, ohne Datum, MODIS „About“, zuletzt abgerufen am 26.07.2023, <https://modis.gsfc.nasa.gov/about/>

⁸NASA, ohne Datum, MODIS „MODIS Design“, zuletzt abgerufen am 26.07.2023, <https://modis.gsfc.nasa.gov/about/design.php>

⁹ESA, 24.07.2015, Sentinel-2 Nutzerhandbuch, zuletzt abgerufen 26.07.2023: https://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook

450–700 nm, Sentinel-2 Bänder 2, 3, 4) sowie nahen Infrarot (Band 8) stark mit denen der Landsat-Missionen übereinstimmen, jedoch mit einer deutlich höheren räumlichen Auflösung von 10 m.

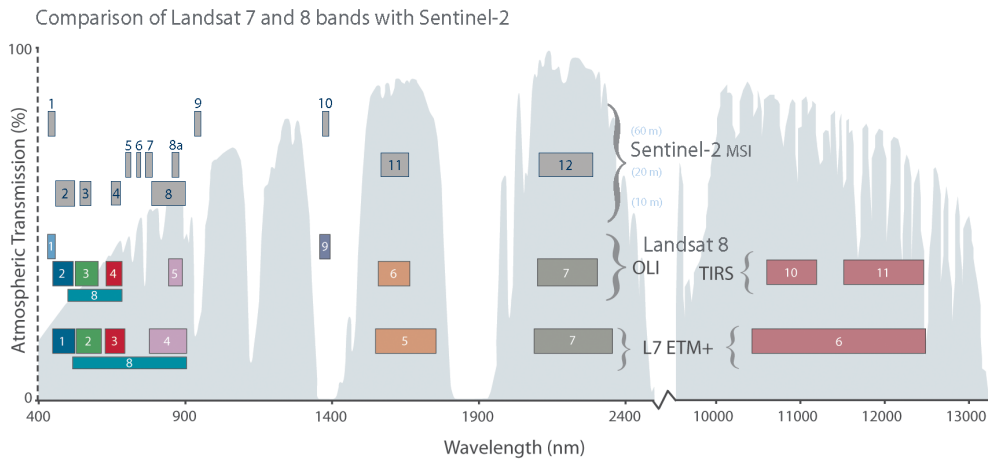


Abbildung 2.2: Vergleich der Spektralkanäle von Sentinel-2 (oben) und Landsat-8 (Mitte) und Landsat-7 (unten). Bänder des Sentinel-2 MSI Sensors sind horizontal gruppiert nach der räumlichen Auflösung der einzelnen Bänder: 10 m, Bänder 2, 3, 4 und 8; 20 m Bänder 5, 6, 7, 8a, 11 und 12; 60 m Bänder 1, 9 und 10. Die x-Achse beschreibt die Wellenlänge (engl. *wavelength*) des elektromagnetischen Spektrums, die graue Schattierung entlang der y-Achse zeigt die atmosphärische Durchlässigkeit des Lichts (engl. *atmospheric transmission*), welche durch die atmosphärische Absorption beeinflusst wird (Jones und Vaughan, 2010, S. 16 ff.). Bildrechte: USGS, public domain.

Im Kontext der Erfassung von Grünflächen bietet die Fernerkundung vielfältiges Potenzial. Chlorophyll und weitere bioaktive Substanzen sind im sichtbaren Licht (ca. 0,4 µm - 0,7 µm) und nahen Infrarot (ca. 0,7 µm - 1,4 µm) deutlich erkennbar (Solimini, 2016). Diese spektralen Eigenschaften werden häufig genutzt, um daraus spektrale Indizes abzuleiten (Jones und Vaughan, 2010, S. 145 f. und 164 ff., Huete, 2012). Einer der bekanntesten spektralen Indizes zur Beschreibung von Vegetation ist der sog. *Normalized Difference Vegetation Index* (NDVI, Rouse Jr. et al., 1974; Tucker, 1979). Dieser kombiniert die Rückstrahlung des roten Lichtes und nahen Infrarot zu einem normalisierten Index und skaliert zwischen -1 und 1. Wolken, Schnee und offene Wasserflächen zeichnen sich durch negative NDVI-Werte aus, während vitale Vegetation positive NDVI-Werte aufzeigt. Ziel des NDVI ist es durch die Kombination der beiden Spektren die Auswirkungen saisonaler Sonnenstandsunterschiede und atmosphärischer Einflüsse zu reduzieren (Rouse Jr. et al., 1974). Der NDVI ist definiert als die Differenz der Reflektanz im nahen Infrarot (*NIR*) und Roten Lichtspektrum (*RED*) geteilt durch deren Summe:

$$NDVI = \frac{NIR - RED}{NIR + RED} \tag{2.1}$$

Schon vor beinahe einem halben Jahrhundert ließen sich die räumliche Ausdehnung von Grünflächen sowie vegetative Eigenschaften, wie beispielsweise die Vitalität, Primärproduktion, oder der Ernteertrag mithilfe von fernerkundlichen Aufnahmen erfassen (Idso et al., 1977; Tucker und Sellers, 1986). Die stetigen Weiterentwicklungen von Sensoren und Methoden haben seither die Einsatzfähigkeit von multispektralen Fernerkundungsbildern deutlich gesteigert (Houborg et al., 2015). Im Kontext der Erfassung von urbanen Grünflächen kommen heute diverse Fernerkundungsdaten und -produkte häufig zum Einsatz (Shahtahmassebi et al., 2021). Studien zu urbanen Grünflächen nutzten Fernerkundungsdaten beispielsweise zur Quantifizierung des Grünanteils in der Umgebung (Dewulf et al., 2016; Santos et al., 2016; Russette et al., 2021).

Aufgrund der oben dokumentierten besonderen Eignung der satellitengestützten Erdbeobachtung für Beschreibung der Erdoberfläche, wurden in den vergangenen Jahren diverse Produkte erstellt, welche u. a. dabei helfen die räumliche Verteilung bestimmter Landbedeckungsklassen wie Vegetation besser zu verstehen. Diese Produkte sind häufig Ergebnis eines (automatisierten) Klassifikationsprozesses basierend auf Satellitendaten. Beispiele für derartige Landbedeckungsklassifikationen sind weltweite Datensätze wie ESA Worldcover¹⁰ oder Google Dynamic World¹¹ (Brown et al., 2022) oder regionale Datensätze wie etwa die japanische hochaufgelöste Landnutzungsklassifikation¹². Auch komplexere Konzepte von Landbedeckungsklassen, wie die Erfassung sog. lokaler Klimazonen (engl. *local climate zones*, LCZ, Stewart und Oke, 2012), wurden bereits großflächig aus Satellitendaten abgeleitet (Zhu et al., 2020; Sapena et al., 2021) und ermöglichen eine detaillierte Beschreibung der städtischen Umwelt. Bildmaterial mit höherer räumlicher Auflösung trägt zu einer realistischeren Einschätzung der Grünfläche bei, insbesondere bei lokalen Anwendungen (Jimenez et al., 2022).

Eine verbleibende Problematik mit satellitengestützten Erdbeobachtungsdaten ist jedoch, dass die Perspektive der Draufsicht anders ist als die der Menschen am Boden. Daher können Verzerrungen zwischen der Wahrnehmung von Grünflächen am Boden und dem in Fernerkundungsdaten gemessenen Grünflächenanteil auftreten (Jiang et al., 2017; Helbich et al., 2021). Manche Eigenschaften von Grünflächen können in der Folge nicht ausreichend erhoben werden (Feltynowski et al., 2018). Dennoch erlauben Fernerkundungsdaten Analysen über arbiträre Grenzen hinweg, wie beispielsweise Stadt- oder Nationalgrenzen.

Jüngste Fortschritte in der Technologie von Satelliten und Fernerkundungssensoren, den Bildklassifikationsverfahren und der Verarbeitungsinfrastruktur ermöglichen großflächige Datenanalysen, z. B. auf regionaler (Leinenkugel et al., 2019), nationaler (Mack et al., 2017), kontinentalen (Pflugmacher et al., 2019) oder sogar globalen Ebene (Chen et al., 2014; Brown

¹⁰ESA, ohne Datum, „Worldwide land cover mapping“, zuletzt abgerufen 26.07.2023, <https://esa-worldcover.org>

¹¹Google und World Resources Institute, ohne Datum „Dynamic world, A near realtime land cover dataset for our constantly changing planet“, zuletzt abgerufen 26.07.2023, <https://www.dynamicworld.app/>

¹²JAXA, ohne Datum, ALOS Research and Application Project, „High-Resolution Land Use and Land Cover Map Products“, zuletzt abgerufen 26.07.2023, https://www.eorc.jaxa.jp/ALOS/en/dataset/lulc_e.htm

et al., 2022). In multitemporalen Analysen werden Landbedeckungsveränderungen beobachtet, um die Dynamik der Landoberfläche zu erfassen, wie z. B. Veränderungen des globalen Waldbestandes (Hansen et al., 2013) und des städtischen Wachstums (Taubenböck et al., 2012). Im Vergleich zu schlechter aufgelösten MODIS und Landsat Bilddaten, versprechen die ebenfalls frei zugänglichen europäischen Sentinel-2 Daten mit einer geometrischen Auflösung von $10\text{ m} \times 10\text{ m}$ eine bessere räumliche Genauigkeit für Landbedeckungs- und Landnutzungskartierung (Close et al., 2018; Sánchez-Espinosa und Schröder, 2019). Insbesondere im Kontext der Erkennung von Landbedeckungsklassen im städtischen Umfeld ist eine bessere räumliche Auflösung von großer Bedeutung, da so auch kleine Grünflächen erkannt werden können. Helbich et al. (2021) skizzieren ein Spannungsfeld zwischen den bisherigen häufig verwendeten Bildern der Landsat-Satelliten und weit höher aufgelösten Orthophotos oder Straßenansichten (engl. *street view images*). Nachteil von Landsat ist dabei die räumliche Auflösung, welche kleine, isolierte Grünflächen schwer erkennen lässt. Orthophotos und Straßenansichten hingegen sind nur schlecht großflächig einsetzbar. Sentinel-2 stellt hier aufgrund der hohen Auflösung bei gleichzeitig großer Abdeckung und freiem Datenzugang einen vielversprechende Zwischenweg dar.

2.4 Maschinelles Lernen im Kontext der Fernerkundung

Das Feld des maschinellen Lernens (ML, engl. *machine learning*) ist sehr umfangreich und assoziierte Methoden finden zunehmend Verwendung in Aufgaben und Produkten des täglichen Lebens. Die disruptive Natur kürzlich vorgestellter Anwendungen, gemeinhin unter dem Sammelbegriff der *künstlichen Intelligenz* (KI, engl. *artificial intelligence, AI*) bekannt, führen zu einem besonderen Maß an Aufmerksamkeit im medialen Diskurs. Unter dem Begriff findet sich eine beinahe unermessliche Vielfalt von Anwendungen. Die folgenden Grundlagen zielen darauf ab, die Anwendung von ML im Kontext geographischer Fernerkundung mit einem Fokus auf Anwendungen im Kontext urbaner Grünflächen zusammenzufassen sowie die notwendigen Grundlagen zu präsentieren, die für das Verständnis der vorgelegten Arbeit notwendig sind.

Mitchell (1997, S. 2) definiert maschinelles Lernen wie folgt (a. d. engl.): „Ein Computerprogramm lernt aus einer Erfahrung E bezogen auf eine Aufgabe T und ein Leistungsmaß P , wenn die Performanz die Aufgabe T zu erfüllen, gemessen nach dem Leistungsmaß P , mit zunehmender Erfahrung E steigt.“

Als Aufgabe T kommen dabei diverse Anwendungen infrage (siehe Goodfellow et al., 2016, S. 97–100). Im Rahmen der hier vorliegenden Arbeit sind davon zwei von Interesse: Klassifikation und Regression. Ziel des Algorithmus der Klassifikation ist es zu erkennen, welcher diskreten Klasse eine Beobachtung zugehört. Im Bereich der geographischen Erdbeobachtung kommen derartige ML-Verfahren beispielsweise für die Kartierung von festgelegten Landbedeckungsklassen zum Einsatz. Ziel der Regression ist es hingegen eine Beobachtung einem Wert entlang einer kontinuierlichen Skala zuzuordnen, z. B. zur Schätzung von Bebauungsanteilen in einem Gebiet.

Abhängig zur Aufgabe T werden verschiedene Leistungsmaße P verwendet, wie beispielsweise die Fehlerrate, d. h. der Anteil der fehlerhaft klassifizierten Beobachtungen, oder der mittlere Fehler, d. h. die mittlere Abweichung zwischen der gemessenen und prädierten Zielvariable (Goodfellow et al., 2016, S. 100 f.).

Die Erfahrung E des maschinellen Lernverfahrens basiert auf der Möglichkeit, Strukturen und Muster aus einem Datensatz zu erlernen. Im Rahmen der geographischen Fernerkundung kommen häufig sogenannte überwachte ML-Verfahren zum Einsatz (Ma et al., 2017). Dabei handelt es sich um Lernverfahren, in denen dem Algorithmus sogenannte Referenzdaten oder (engl.) *ground truth*-Daten im Rahmen des Trainingsprozesses präsentiert werden. Diese stellen die Grundlage des Lernens der Zusammenhänge zwischen einem Merkmalsraum und der gesuchten Zielvariable (Klasse oder kontinuierlicher Wert) und somit die Erfahrung E dar (Goodfellow et al., 2016, S. 101, 136). Zentrale Eigenschaft der Algorithmen des maschinellen Lernens ist somit, dass der Lösungsweg nicht auf zuvor bestimmten, explizit definierten Entscheidungen oder Hypothesen beruht, sondern aus der Gesamtmenge aller möglicher Hypothesen die Beste ausgewählt wird, die basierend auf dem vorliegenden Datensatz zur Entscheidungsfindung führt (Mitchell, 1997, S. 14).

Methoden des maschinellen Lernens sind im Kontext der Erdbeobachtung inzwischen ein etabliertes Mittel für die großflächige Analyse von Satellitenbildern (Feng und Li, 2020) und werden insbesondere in hochdimensionalen Datensätzen und für komplexe Probleme eingesetzt (Maxwell et al., 2018). Maschinelles Lernen steht dabei im Kontrast zu parametrischen Methoden, wie z. B. Maximum Likelihood, manuellen Schwellwertverfahren oder physikalischen Modellen. Das Grundkonzept des maschinellen Lernens lässt sich in zwei Schritte unterteilen: erstens das Lernen des Modells aus bekannten Referenzdaten, das sogenannte Training, und zweitens die Vorhersage oder Prädiktion für alle Pixel des Satellitenbildes mithilfe des trainierten Modells. Das Ziel des Lernvorgangs ist dabei, Muster in den Satellitendaten sowie daraus abgeleiteten Bildmerkmalen zu identifizieren, welche dabei helfen, eine bestimmte Landbedeckungsklasse oder kontinuierliche Variablen zu identifizieren und von anderen Ausprägungen zu unterscheiden. Die genauen Zusammenhänge werden dabei nicht explizit festgelegt, sondern ergeben sich dynamisch aus den Daten (Lary et al., 2016). Die Qualität des klassifizierten Bildes hängt dabei unter anderem vom verwendeten ML-Algorithmus sowie der Qualität der Eingangsdaten ab (Maxwell et al., 2018).

Für überwachte Lernverfahren in der Fernerkundung sind zweierlei Daten notwendig: Einerseits Referenzdaten für die Zielvariable, z. B. Landbedeckungsklassen, andererseits die vom Erdbeobachtungssensor erfassten spektralen Bilddaten und optional daraus abgeleitete Bildmerkmale oder spektrale Indizes (Camps-Valls, 2009).

Referenzdaten dienen als Beispiele anhand derer der maschinelle Lernalgorithmus erfahren kann, welche Eigenschaften etwa eine bestimmte Landbedeckungsklasse ausmachen. Während des Trainings- und Testvorgangs wird eine Vielzahl an Beispielen benötigt, welche die Heterogenität

innerhalb und zwischen den Ausprägungen der Zielvariable abbilden. Im Kontext der fernerkundungsbezogenen Anwendung von ML-Verfahren ist es stets schwierig ausreichend Referenzdaten zu beziehen, da deren manuelle Erhebung sehr aufwändig und zeitintensiv ist (Congalton, 1991). Darüber hinaus ist es besonders wichtig, dass die Referenzdaten im Trainings- und Testdatensatz räumlich voneinander getrennt sind, um eine unabhängige und somit verlässliche Evaluation über den Erfolg des maschinellen Lernverfahrens durchzuführen (Karasiak et al., 2021).

Agrund vermehrter Anwendungen von maschinellem Lernen in der Erdbeobachtung steigt die Nachfrage nach hochqualitativen Referenzdaten. Daher wurden in den vergangenen Jahren im Rahmen unterschiedlicher Initiativen Referenzdatensätze erstellt. Zhu et al. (2020) erstellten beispielsweise einen Referenzdatensatz für die Klassifikation von Lokalen Klimazonen anhand von Daten der Sentinel-1 und Sentinel-2 Missionen. Die Initiative Geo-Wiki¹³ stellt diverse Referenzdaten für bebauten Gebiete sowie beispielsweise landwirtschaftliche Nutzflächen bereit, die bereits in diversen Applikationen zum Einsatz kommen¹⁴. Dieser globale Referenzdatensatz für Landbedeckungs- bzw. Landnutzungsdaten (LB/LN) wird dabei im Rahmen bürgerwissenschaftlicher (engl. *citizen science*) Initiativen von Freiwilligen durch visuelle Bildinterpretation gesammelt und beinhaltet ca. 150.000 Referenzpunkte (Fritz et al., 2017). Trotz der enormen Anstrengungen, die unternommen wurden, um einen konsistenten Referenzdatensatz von LB/LN-Informationen bereitzustellen, ist die Anzahl der verfügbaren Stichproben in einigen Regionen jedoch relativ gering. So enthält der Datensatz bis dato¹⁵ lediglich Referenzpunkte für 374 verschiedene Orte in Deutschland. Die Daten unterscheiden sich wiederum in ihrer räumlichen und semantischen Konsistenz. Eine ähnliche Initiative namens MapSwipe¹⁶ nutzt den Beitrag Freiwilliger für die manuelle Klassifikation und Kartierung einzelner Beispielbilder im Kontext humanitärer Anwendungen. Diese Daten spielen als Referenz eine wichtige Rolle für die Klassifikation von Erdbeobachtungsdaten (z. B. Chen et al., 2019). EuroSAT ist ein Referenzdatensatz, welcher speziell für die Landnutzungs- und Landbedeckungsklassifikation aus Sentinel-2 Daten entwickelt wurde. Dieser Datensatz stellt 27.000 Bildelemente in Europa bereit, welchen jeweils eine zugehörige LB/LN-Information zugeordnet ist (Helber et al., 2019). Das Konzept der freiwillig zur Verfügung gestellten geographischen Informationen (engl. *volunteered geographic information, VGI*) ist besonders bekannt aus dem OpenStreetMap-Projekt¹⁷ (OSM). Diese Daten wurden ebenfalls bereits erfolgreich als Referenz für Fernerkundungsbildklassifikationen eingesetzt (Schultz et al., 2017; Wan et al., 2017; Maggiori et al., 2017).

¹³Geo-Wiki, ohne Datum, „Earth Observation & Citizen Science“, zuletzt abgerufen 26.07.2023, <https://www.geo-wiki.org>

¹⁴Geo-Wiki, ohne Datum, „About Us“, zuletzt abgerufen 26.07.2023, <https://www.geo-wiki.org/page/about>

¹⁵Zenodo, 13.06.2017, „A global dataset of crowdsourced land cover and land use reference data (2011-2012)“, <https://zenodo.org/record/6572482>

¹⁶MapSwipe, ohne Datum, „Swipe and tap. Put communities on the map.“, zuletzt abgerufen 26.07.2023, <https://mapswipe.org>

¹⁷OpenStreetMap, ohne Datum, zuletzt abgerufen 26.07.2023, <https://openstreetmap.org>

Derartige Datensätze stellen die Grundlage vieler moderner geographischer Anwendungen von Erdbeobachtungsdaten dar. Durch die Kombination mit geeigneter Referenzdaten und fernerkundlicher Bilddaten kann ein maschinelles Lernverfahren trainiert werden, welches die Eigenschaften der Zielvariable den physischen Eigenschaften der Erdoberfläche modelliert. Hierbei ist es hilfreich, möglichst aussagekräftige beschreibende Merkmale bereitzustellen. Etwa im Falle einer Landbedeckungsklassifikation müssen die Bilddaten die entscheidenden Merkmale bieten, die Trennung der unterschiedlichen Klassen zu ermöglichen und gleichzeitig die Variation innerhalb jeder Klasse abzubilden. Dabei helfen diverse direkte und höherwertige Informationen aus den Bildern: Einerseits sind die spektralen Informationen selbst relevante Datenpunkte zur Unterscheidung einzelner Landbedeckungsklassen. Beispielsweise sind zur Unterscheidung von Vegetation zu bebauter Fläche die Reflektanzen des grünen Lichtes besonders relevant. Andererseits können spektrale Indizes, wie der NDVI, genutzt werden, um Zusammenhänge zwischen verschiedenen Bändern in einem Pixel zu abstrahieren. Auch räumlich-kontextuelle Merkmale, wie die Textur einer Nachbarschaft von Pixeln oder Objektsegmentierung stellt eine hilfreiche Information dar (Li et al., 2014). Weiteres Potenzial den Merkmalsraum der Bilddaten zu vergrößern liegt in der Verwendung multitemporaler Bilddaten, die dank des wiederkehrenden Überflugs von Satelliten entstehen. Hieraus können beispielsweise zeitliche Verläufe der phänologischen Zyklen abgeleitet werden (Liang und Schwartz, 2009). Je mehr Bildmerkmale zur Verfügung stehen, desto besser können Landbedeckungseigenschaften abstrahiert werden (Khatami et al., 2016).

Für das maschinelle Lernen im Bereich der Erdbeobachtung werden die Bilddaten räumlich mit den Referenzinformationen verknüpft werden. Jeder Referenzbeobachtung steht demnach ein Vektor von beschreibenden Bildmerkmalen gegenüber. Der Zusammenhang zwischen dem Zielwert y , repräsentiert durch den Referenzdatensatz, z. B. in Form einer Landbedeckungsklasse, soll in vorwärts gerichteten maschinellen Lernverfahren somit als Funktion der Bildmerkmale x und assoziierter Gewichte sowie Parameter θ modelliert werden (vgl. Goodfellow et al., 2016, S. 163), z. B.:

$$y = f(x; \theta) = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \times \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix} + b \quad (2.2)$$

Methoden des maschinellen Lernens wurde bereits in diversen Studien erfolgreich im Rahmen der Erdbeobachtung eingesetzt. Die Einsatzzwecke reichen von der Klassifikation von urbaner Landbedeckung mithilfe von Multispektraldaten (z. B. Huang et al., 2007) hin zu selbstlernenden Verfahren im Kontext komplexer Siedlungsstrukturen (z. B. Geiß et al., 2019). Auch im Kontext der die Unterscheidung verschiedener Siedlungstypen findet maschinelles Lernen Einsatz, etwa zur Klassifikation von SAR-Daten der TerraSAR-X Mission (z. B. Wurm et al., 2017a) oder

multispektralen Daten von Sentinel-2 (z. B. Wurm et al., 2017b). Großflächige Anwendungen von Landbedeckungsklassifikationen wurden ebenfalls bereits durchgeführt, beispielsweise wurden mithilfe von Daten der Landsat-Satelliten Landbedeckungsklassifikationen auf regionaler (Leinenkugel et al., 2019), nationaler (Mack et al., 2017) oder kontinentaler Ebene (Pflugmacher et al., 2019) durchgeführt. Die komplexen Daten von voll polarisierten SAR-Daten konnten anhand von ML zur Erkennung von Landbedeckung genutzt werden (z. B. Du et al., 2015). Die Klassen des *Local Climate Zones*-Konzeptes wurden mit maschinellem Lernen erfolgreich aus diversen fernerkundlichen Datenprodukten abgeleitet (z. B. Bechtel und Daneke, 2012; Danylo et al., 2016). Diverse Pflanzenarten und -vitalitätsfaktoren wurden durch ML im Kontext urbaner Gebiete (z. B. Tooke et al., 2009) oder küstennaher Gebiete untersucht (Dai et al., 2023) und zur Unterscheidung verschiedener Feldfrüchte genutzt (z. B. Skakun et al., 2016; Kussul et al., 2018). Mittels multisensoraler Daten und maschinellen Lernverfahren konnten einzelne Gebäudetypen unterschieden werden (z. B. Geiß et al., 2015). Diese Beispiele bilden nur einige der vielzähligen Anwendungsmöglichkeiten von „klassischen“ Methoden des maschinellen Lernens dar.

Seit ca. 2014 kommt in der Fernerkundung zunehmend eine Klasse von maschinellen Lernverfahren zum Einsatz, die auf sog. tiefen neuronalen Netzen (engl. *deep neural networks*) basiert und unter dem Begriff *Deep Learning* (DL) zusammenzufassen ist (Zhu et al., 2017; Ma et al., 2019). Als Resultat der zunehmenden Verfügbarkeit von leistungsstarken Prozessierungskapazitäten, größeren frei verfügbaren Datensätzen sowie zunehmenden Möglichkeiten und Techniken zur Verarbeitung tieferer Netzwerke, stieg die Popularität und Nutzbarkeit dieser Art von Algorithmen an (Goodfellow et al., 2016, S. 26) und sie finden zunehmend Anwendung in der geographischen Fernerkundung und Erdbeobachtung (Zhu et al., 2017). Die besondere Stärke von Deep Learning liegt darin, hochkomplexe Zusammenhänge zwischen einzelnen Merkmalsausprägungen in einem ML-Problem zu generalisieren, z. B. in der Objekterkennung (Goodfellow et al., 2016, S. 151). Ein bedeutender Unterschied zwischen „klassischen“ Methoden des maschinellen Lernens und Deep Learning ist die Fähigkeit von Deep Learning relevante Merkmale selbst zu erlernen. Die sogenannte (engl.) *backpropagation*, d. h. die rückwärtsgewandte Betrachtung des Klassifikationsvorgangs, ermöglicht die neuronalen Netzwerke entlang des optimalen Merkmals zur Entscheidungsfindung auszurichten (LeCun et al., 2015). Dadurch erübrigt sich der mitunter sehr aufwändige Arbeitsschritt klassischer ML-Ansätze, relevante Bildmerkmale *a priori* zu errechnen und für das Modell bereitzustellen. Deep Learning Modelle erlernen hingegen simultan die optimalen Bildmerkmale, die zur Erkennung der Zielvariablen relevant sind, sowie die Zielvariable selbst (Quinn et al., 2018). Dies resultiert häufig in besseren Ergebnissen als traditionellere Methoden erreichen könnten, erfordert jedoch mehr Rechenaufwand und besondere Rechenhardware.

Die Einsatzzwecke von Deep Learning in der Erdbeobachtung sind sehr umfangreich. Tiefe neuronale Netze eignen sich etwa dazu, einzelnen Bildausschnitten eine Landbedeckungsklasse zuzuordnen, ein Vorgang der häufig als Szenenklassifikation (engl. *scene classification*) bezeichnet wird (Cheng et al., 2020). Daneben wurden Methoden entwickelt, mit denen pixelbasierte

Klassifikationen von Bildinhalten durchgeführt werden können. Diese können unter dem Sammelbegriff der semantischen Segmentierung (engl. *semantic segmentation*) zusammengefasst werden (Kemker et al., 2018).

Die unterschiedlichen Anwendungsgebiete und Architekturen der neuronalen Netze können dabei stark variieren. Szenenklassifikation sogenannter *Convolutional Neural Networks* (*CNN* Fukushima, 1980; Waibel et al., 1989; LeCun et al., 1990) wurde eingesetzt, um verschiedene LB/LN-Klassen zu erkennen (z. B. Naushad et al., 2021; Ma et al., 2021). Auch die Klassifikation von *Local Climate Zones* konnte so durchgeführt werden (Qiu et al., 2020b; Zhu et al., 2022). Beispielsweise konnten CNNs bereits erfolgreich für die Kartierung und Unterscheidung komplexer Siedlungsmuster eingesetzt werden (Wurm et al., 2019b; Stark et al., 2020) oder zur Erkennung einzelner Gebäude (Maggiori et al., 2017; Quinn et al., 2018; Chan et al., 2022). Des Weiteren wurden mithilfe von CNNs Siedlungsmasken aus Sentinel-2 Bilddaten abgeleitet (Qiu et al., 2020a). Geiß et al. (2020) setzten Deep Learning-Methoden zur Verbesserung von Landbedeckungsklassifikationsergebnissen ein. Zhang et al. (2019) verwendeten komplexe Architekturen zur Klassifikation von Landnutzung und Landbedeckung. Unter Kombination heterogener Datensätze konnten mithilfe von Deep Learning auch Landnutzungen auf nationaler Ebene ausgewertet werden (Rosier et al., 2022). Eine besondere Herausforderung der DL-Methoden ist jedoch deren Anforderung von enorm großen Datenmengen, welche zum Training bereitgestellt werden müssen (Cheng et al., 2017; Quinn et al., 2018; Hua et al., 2022). Es ist daher insbesondere das Feld der Anwendungen von Deep Learning in der Fernerkundung, die von den obengenannten Initiativen zur Veröffentlichung von großangelegten Referenzdaten profitiert.

Sowohl klassische maschinelle Lernverfahren als auch Deep Learning-Methoden stellen einen integralen Teil der modernen Erdbeobachtung dar und ermöglichen die automatisierte und effiziente Inwertsetzung der zunehmenden Datenmengen. Die Auswahl des geeigneten Verfahrens geschieht stets abhängig vom individuellen Anwendungsfall und muss an die Voraussetzungen, wie die verfügbaren Bild- und Referenzdaten, sowie die anvisierte Aufgabenstellung angepasst werden.

Konzeptioneller Rahmen

Die Arbeiten dieser Dissertation teilen sich in drei getrennte Aufgaben. Im folgenden Kapitel werden die konzeptionellen Überlegungen dargelegt, welche als Grundlage für die gewählten Daten (Kapitel 4) und Methoden (Kapitel 5), die Struktur der Ergebnisdarstellung (Kapitel 6) sowie Diskussion (Kapitel 7) dienen.

3.1 Beschreibung der Umwelt in Deutschland

Um aus umfangreichen Satellitendaten höherwertige Informationen abzuleiten, die gewinnbringend für die Untersuchung der räumlichen Verteilung und Zugänglichkeit von Grünflächen genutzt werden können, ist es zunächst notwendig, diese Rohdaten in Geoinformationen zu überführen. Eine Methode zur Gewinnung von Informationen aus Satellitenbilddaten ist die räumliche Kartierung mithilfe der Klassifikation von Landbedeckung bzw. Landnutzung (LB/LN). Um im Rahmen der Dissertation die hochgenaue landesweite Analyse von urbanen Grünflächen zu ermöglichen, soll zunächst eine solche Klassifikation von Satellitendaten durchgeführt werden. Dabei stehen drei zentrale Aspekte im Vordergrund:

1. Die differenzierte Beschreibung unterschiedlicher Landbedeckungsklassen von Vegetation,
2. die hohe räumliche Auflösung der finalen Landbedeckungsklassifikation, um auch komplexe räumliche Muster und kleinräumige Grünflächen in Siedlungsgebieten zu erfassen,
3. die flächendeckende und automatisierte Anwendbarkeit, um eine Untersuchung von Grünflächen in ganz Deutschland zu ermöglichen und die Übertragbarkeit auf andere Zeiträume oder Einzugsgebiete zu gewährleisten.

Unter Berücksichtigung dieser Vorgaben stellen die Daten des Sentinel-2 Satelliten eine wertvolle Datenquelle dar. Im Vergleich zu Landsat bietet Sentinel-2 eine bis zu neunmal höhere räumliche Auflösung für multispektrale Aufnahmen im sichtbaren Licht und nahen Infrarot. Dies reduziert die Größe der minimalen Kartiereinheit (engl. *minimum mapping unit*, MMU), also die Größe der Objekte, die gerade groß genug sind, um erfasst werden zu können. Im Vergleich zu zuvor für derartige Klassifikationen verwendete Satellitendaten werden so auch kleinere Grünflächen im Siedlungsgebiet erkennbar.

Neben den Satellitendaten werden für eine erfolgreiche automatisierte Landbedeckungsklassifikation geeignete Referenzdaten benötigt. Die Erstellung dieser Daten für einen einzigen Zweck ist häufig sehr aufwändig und umfangreich (Congalton, 1991). So auch im hier vorliegenden Fall, welcher zum Ziel hat eine flächendeckende, landesweite Klassifikation zu erstellen. In ähnlichen großflächigen Ansätzen in der Literatur wird deshalb auf einen bestehenden Datensatz des sog. *Land Use and Coverage Area Frame Survey* (LUCAS) zurückgegriffen (z. B. Mack et al., 2017; Close et al., 2018; Pflugmacher et al., 2019; Leinenkugel et al., 2019; d'Andrimont et al., 2021a). Dieser beinhaltet wertvolle Informationen, welche für die Klassifikation von Landbedeckungsinformationen aus Erdbeobachtungsdaten gewinnbringend eingesetzt werden können. Der LUCAS-Datensatz wurde jedoch zuvor noch nicht in Kombination mit Sentinel-2 Daten verwendet, weswegen im Rahmen dieser Dissertation die Eignung dieses Datensatzes für die Klassifikation von Sentinel-2 untersucht wird. Auch eine Analyse der verschiedenen Vorprozessierungsschritte steht im Fokus.

Ziel dieser konzeptionellen Einheit ist es, eine umfassende Bewertung der räumlichen und semantischen Auswirkungen von LUCAS-Stichproben auf großräumige LB/LN-Klassifikationen aus Sentinel-2 Daten vorzunehmen. Hierbei werden vier verschiedene Ansätze zur Auswahl von LUCAS-Referenzdaten für die Verwendung in großflächigen, hochauflösenden Landbedeckungskartierungen verglichen. Einerseits werden die acht übergeordneten Landnutzungsklassen des LUCAS-Datensatzes mithilfe von Sentinel-2 Daten klassifiziert. Die Ansätze, die von Mack et al. (2017) und Pflugmacher et al. (2019) entwickelt wurden, um großflächige Bilddaten der Landsat-Satelliten auf Basis von LUCAS in-situ Stichproben zu klassifizieren, werden im Rahmen dieser Dissertation auf Sentinel-2 übertragen und so deren Anwendbarkeit in höher auflösenden Bilddaten getestet. Darüber hinaus wird ein Vorprozessierungsansatz von Weigand et al. (2020) präsentiert, welcher einen Schwerpunkt auf die im Kontext der Dissertation relevanten Grünflächen legt.

Die flächendeckende Ableitung von Landbedeckungsinformationen aus Satellitenbilddaten erfolgt mithilfe eines überwachten maschinellen Lernverfahren. Hierbei liegt die Annahme zugrunde, dass verschiedene Landbedeckungsklassen durch klar identifizierbare und abgrenzbare Bildeigenschaften voneinander zu trennen sind (Maxwell et al., 2018). Anhand einer im Vergleich zur Gesamtfläche Deutschlands relativ geringen Menge von Referenzdaten werden Modelle erstellt. Diese haben das Ziel die Merkmale zu identifizieren, mit deren Hilfe die Landbedeckungsklassen voneinander getrennt werden können. Hierfür werden die Referenzdaten anhand ihrer Geokoordinate räumlich mit den Bilddaten verknüpft werden. Daraus resultiert ein tabellarischer Datensatz, der als Eingangsdatensatz zum Training der maschinellen Lernmodelle dient. Der Datensatz beinhaltet die Zielklasse an jedem Referenzpunkt, die Spektraldaten aus Satellitenbildern sowie abgeleiteten Bildmerkmalen und zusätzliche Informationen aus sekundären Geodaten, die Bebauung und versiegelte Flächen ausweisen. In einem iterativen Verfahren werden alle Modelle mit variierenden

Sets von Trainings- und Testpunkten trainiert und mithilfe geeigneter Genauigkeitsmaße evaluiert. So wird die Verlässlichkeit der Ergebnisse in Relation zur Auswahl der Trainings- und Testpunkte geschätzt.

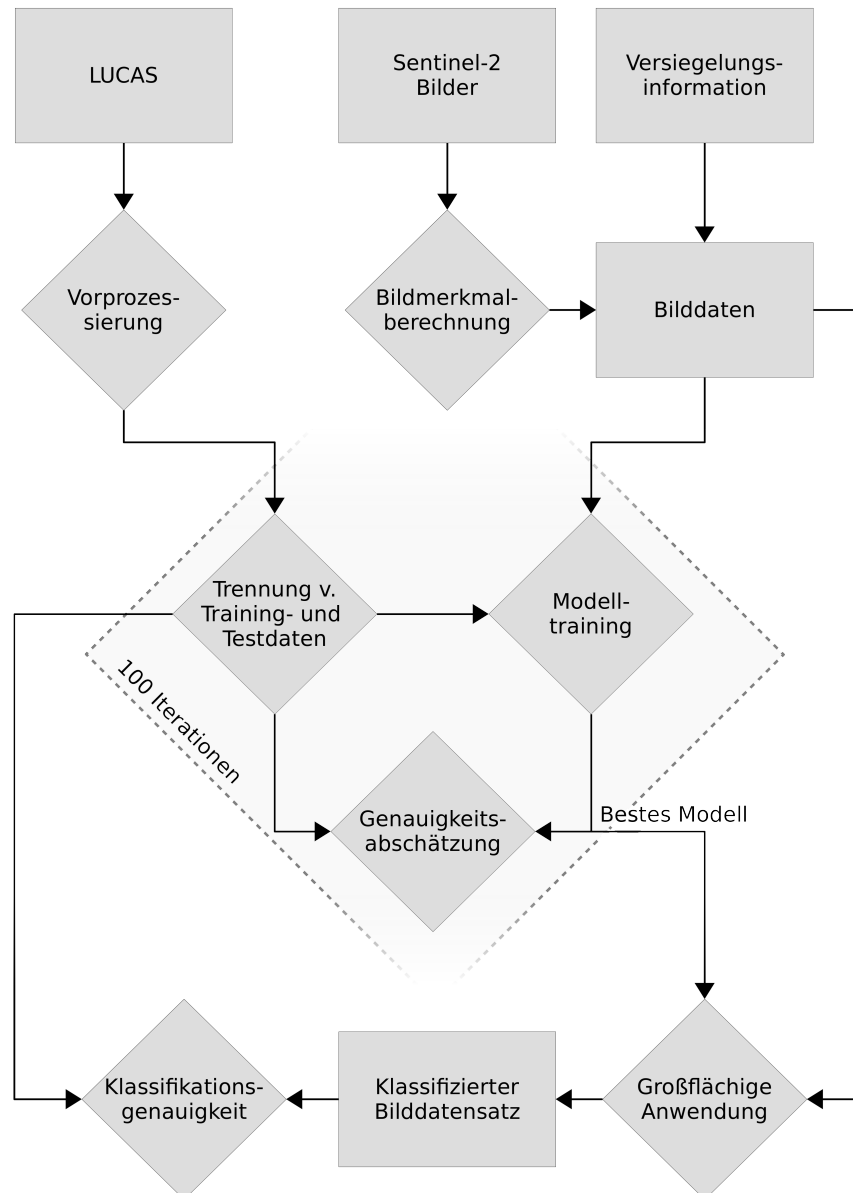


Abbildung 3.1: Ablaufdiagramm der Daten (□) und Methoden (◇) der Landbedeckungsklassifikation für Deutschland (verändert nach Weigand et al., 2020).

Auf Basis der gesammelten Erkenntnisse wird eine hochaufgelöste Landbedeckungsklassifikation für ganz Deutschland erstellt. Diese dient wiederum in den folgenden Kapiteln der Dissertation als Datengrundlage für die Analyse der räumlichen und sozialen Verteilung von Grünflächen in Deutschland. Abbildung 3.1 zeigt den Ablauf der Studie von den Eingangsdaten über die methodische Verarbeitung bis hin zum finalen Ergebnis der Landbedeckungsklassifikation.

Die Methode zur Vorprozessierung von LUCAS-Referenzdaten und automatisierten Klassifikation von Sentinel-2 Multispektraldaten wird zunächst entwickelt für den Zeitraum 2015–2017, zielt jedoch auch auf eine wiederholbare Anwendbarkeit ab. Durch die langfristige Verfügbarkeit der Sentinel-2 Bilddaten über diesen Zeitraum hinaus, sowie durch die dreijährige Erhebung des LUCAS-Surveys ist es möglich diese Klassifikation mehrfach durchzuführen. Im Rahmen dieser Arbeit wird die Methodik daher auf den Zeitraum des Jahres 2018 erneut angewendet, um die Übertragbarkeit zu demonstrieren.

Teile dieses konzeptionellen Schrittes wurden bereits in Weigand et al. (2020) veröffentlicht.

3.2 Quantifizierung von öffentlichen Grünflächen in der Wohnumgebung

Städte werben oft damit, die „grünste Stadt“ zu sein. Doch selbst auf der Makroebene hängt der Grünheitsgrad einer Stadt stark von der Objektivität der zur Analyse verwendeten Daten und der Definition des räumlichen Geltungsbereichs ab und kann daher je nach gewählter Metrik stark variieren (Taubenböck et al., 2021). Diese Abweichungen nehmen auf lokalen Skalen noch zu (Wüstemann et al., 2017a). Darüber hinaus ist nicht alles Grün aus gesundheitlicher Sicht gleichermaßen relevant (Mears et al., 2020). Straßenbegleitgrün mag sich positiv auf die Luftqualität auswirken (Pugh et al., 2012), dient aber nicht als Ort der körperlichen Betätigung oder des sozialen Miteinanders. Ebenso kann ein privater Garten oder eine Kleingartenanlage für manche Menschen ein Zufluchtsort vor der Hektik der Stadt sein, aber sie sind in der Regel nicht für die Öffentlichkeit zugänglich. Wenn man also über die Frage „*Wie grün ist eine Stadt?*“ hinausgeht und sich der Frage „*Wie viele öffentlich zugängliche Grünflächen, Parks und Gärten gibt es in einer Stadt?*“ zuwendet, ist es wichtig, zwischen verschiedenen Arten von Grün zu unterscheiden. Daher ist es notwendig, über die physische Darstellung von grünen Flecken in der (städtischen) Landschaft, d. h. die Bodenbedeckung, hinauszugehen und semantische Informationen über die Art des Grüns, d. h. die Bodennutzung, hinzuzufügen.

In dieser Studie werden *öffentliche* Grünflächen (engl. *public green spaces, PGS*) in Übereinstimmung mit Ludwig et al. (2021) definiert als bewachsene oder natürliche Flächen, von denen allgemein angenommen werden kann, dass sie für die Öffentlichkeit frei zugänglich sind. Dazu gehören Wälder, Wildnisgebiete, städtische Parks, Spielplätze und öffentliche Gärten. Im Gegensatz dazu unterscheiden sich *nicht-öffentliche* Grünflächen dadurch, dass sie sich in privatem Besitz befinden und von der freien öffentlichen Zugänglichkeit und Nutzung ausgeschlossen sein können. Dies ist in der Regel der Fall bei Privatgärten, Flächen mit begrenztem Zugang wie Sportanlagen oder Kleingärten, Industrie- oder Gewerbeflächen, landwirtschaftlichen Kulturen, Flugplätzen, militärischen Sperrgebieten oder anderweitig nicht nutzbaren begrüntem oder be-

wachsenen Flächen. Ziel ist es die Verfügbarkeit, d. h. die Menge an öffentlichen Grünflächen in Hektar zu quantifizieren, da dies eine robuste Messgröße für innerstädtische Ungleichheiten ist, beispielsweise im Vergleich zu entfernungsbasierten Metriken (Rigolon, 2016).

Die Verfügbarkeit von Grünflächen kann auf unterschiedlichen räumlichen Ebenen gemessen werden. Beispielsweise verwenden einige Studien die durchschnittliche Grünfläche in administrativen Raumeinheiten oder Postleitzahlengebieten mit deutlichen Unterschieden in Größe, Urbanität und Bevölkerung (z. B. Akpinar et al., 2016). Dies kann jedoch zu Verzerrungen führen, die durch das Problem der modifizierbaren Gebietseinheiten (engl. *modifiable areal units problem*, MAUP, Openshaw, 1983) verursacht werden. Dies bedeutet, dass das Ergebnis der Analyse durch die willkürliche Anordnung der Gebietseinheiten verzerrt werden kann. Hier wird versucht, die Verzerrung durch das MAUP zu verringern, indem mithilfe standardisierter räumlicher Einheiten für das gesamte Untersuchungsgebiet gearbeitet wird. Dabei steht die Verfügbarkeit von Grünflächen im Lebensumfeld der Menschen im Fokus, nämlich in der Nachbarschaft ihrer Wohnadresse. Dies geschieht da die umgebende Nachbarschaft die individuelle Wahrnehmung der städtischen Struktur (Wurm et al., 2019a) und somit die Gewohnheiten der Fußgänger stark beeinflusst (Droin et al., 2023). In der Literatur über städtisches Grün finden sich verschiedene Definitionen für die Nachbarschaft (für eine umfassende Übersicht siehe Kabisch, 2019). Oft werden diese um die Wohnadresse mit unterschiedlicher Entfernung abgegrenzt, z. B. 300 m (Xu et al., 2018b,a; Wolff und Haase, 2019; Barber et al., 2021), 400 m (Flacke et al., 2016), 500 m (Wüstemann et al., 2017a), 100–3000 m (Klomp maker et al., 2018). Im Rahmen dieser Arbeit wird die Entfernung, die Kabisch (2019) in den jüngsten Studien als die am häufigsten verwendete Entfernung identifiziert haben, verwendet: eine Nachbarschaft mit einem Radius von 500 m.

Flächendeckend für Deutschland werden die Grünflächenanteile für Landbedeckung mit Grünflächen aus der in Kapitel 3.1 beschriebenen Landbedeckungsklassifikation abgeleitet. Da diese jedoch keine Informationen über die semantischen Typen der Grünflächen enthalten, sollen in diesem methodischen Schritt frei verfügbare Geodaten verwendet werden, um detailliertere und semantisch differenzierte Geoinformationen über Grünflächen abzuleiten. Jüngste Studien haben durch den Einsatz von Datenfusion zwischen Satellitenbildern und zusätzlichen Geodaten, u. a. auf Basis von VGI, gezeigt, wie über die Landbedeckung hinaus detaillierte Landnutzungsinformationen erfasst werden können (Dennis et al., 2018; Ludwig et al., 2021; Rosier et al., 2022). VGI liefern umfangreiche semantische Attribute der Landnutzung, die über die physische Beschreibung der Landbedeckung hinausgehen (Vargas-Munoz et al., 2021). Dies erweitert die Analysen urbaner Grünflächen durch die Unterscheidung von Grüntypen wie Wäldern oder Parks gegenüber grüner landwirtschaftlicher Flächen und ermöglicht eine Beschreibung der Zugänglichkeit der Grünflächen. Im Rahmen der Dissertation wird zu diesem Zweck ein Datenfusionsansatz auf Deep Learning-Basis genutzt.

Der methodische Aufbau zur Beschreibung der Wohnumwelt der Bewohner Deutschlands ist in Abbildung 3.2 dargestellt. Aufbauend auf verschiedenartigen Datensätzen, nämlich Sentinel-2 Bildern, Aggregaten aus OpenStreetMap sowie dem European Urban Atlas als Referenzdatensatz wird Menge von öffentlichen Grünflächen in der Wohnumgebung der Menschen in Deutschland geschätzt. Um die heterogenen Datensätze zu vereinen, wird ein maschinelles Lernverfahren auf Basis von tiefen neuronalen Netzen angewandt. Genaugenommen werden zwei unterschiedliche Typen von neuronalen Netzen, nämlich ein ANN und ein CNN, synergetisch fusioniert, sodass die verschiedenartigen Eingangsdaten simultan zur Quantifizierung der öffentlichen Grünflächen auf Nachbarschaftsebene genutzt werden können.

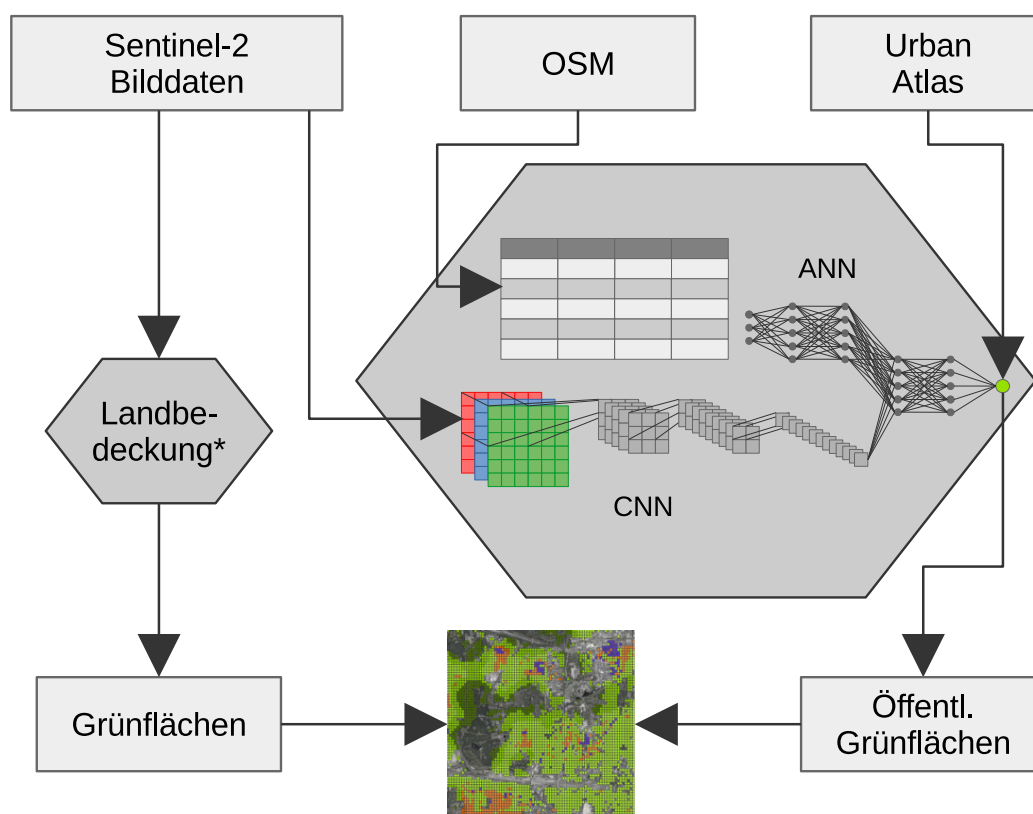


Abbildung 3.2: Ablaufdiagramm der Methode zur Ableitung von öffentlichem Grünanteil in der Wohnumgebung. Hierbei werden Daten (□) und Methoden (◇) unterschieden. Zum Einsatz kommt ein Fusionsnetzwerk bestehend aus einem (engl.) *Convolutional Neural Network* (CNN) und einem (engl.) *Artificial Neural Network* (ANN).

Teile dieses konzeptionellen Schrittes wurden bereits in Weigand et al. (2023b) veröffentlicht.

3.3 Gerechtigkeit der nachbarschaftlichen Grünflächen in Deutschland

Bisherige Untersuchungen zur Verteilung von Grünflächen sind räumlich meist stark begrenzt. In Fallstudien werden überwiegend einzelne Städte oder eine geringe Zahl von Großstädten betrachtet. Dies ist unter anderem auf den immensen Aufwand zur detaillierten Erfassung der Umwelt zurückzuführen, welcher eine detailliertere Betrachtung der Umweltbelastungen bzw. Umweltressourcen auf Individualebene erschwert.

Basierend auf den hochaufgelösten und deutschlandweit erfassten Geodaten zur räumlichen Verteilung von grüner Landbedeckung und öffentlichen Grünanlagen (wie in den konzeptionellen Schritten 3.1 und 3.2 vorgestellt), ist es möglich die sozialräumliche Verteilungsgerechtigkeit von Grünflächen deutschlandweit zu analysieren. Dabei stehen die Menschen im Mittelpunkt, das bedeutet die Analyse betrachtet die Verfügbarkeit von Grünflächen in Bezug auf die lokale Umwelt am Wohnort der Bewohnerinnen und Bewohner Deutschlands. Diese Art der Aggregation arbeitet direktional invers zu Konzepten der räumlichen Aggregation auf Basis vorgegebener Raumeinheiten anhand von administrativen Flächen, wie beispielsweise Stadtbezirken oder Postleitzahlgebieten. Zur Verortung der Bevölkerung in Deutschland wird ein ebenfalls landesweit verfügbarer Datensatz aus der nationalen Zensuserhebung herangezogen. Anhand der Kombination dieser einzigartigen Datensätze wird es erstmals möglich, die Gerechtigkeit der Verteilung von Grünflächen in Deutschland auf einem hohen räumlichen Detailgrad zu untersuchen.

Eine absolut gerechte Verteilung von Grünflächen für alle Bewohner eines Landes ist de facto unmöglich. Deshalb wird in dieser Arbeit, Gerechtigkeit anhand von drei Gesichtspunkten betrachtet:

Einerseits soll untersucht werden, welcher Teil der Bevölkerung gewisse Mengen von Grünflächen in der Wohnumwelt vorfindet. Basierend auf der Nachbarschaft im Radius von 500 Metern (siehe oben, Kapitel 3.2), werden die Verfügbarkeiten von grüner Landbedeckung sowie öffentlicher Grünflächen im Wohnumfeld jeweils in drei Klassen unterschieden. Die Verfügbarkeit von grüner Landbedeckung wird dabei in Terzile geteilt: niedrige, mittlere und hohe Verfügbarkeit von grüner Landbedeckung. Öffentliche Grünflächen, welche eine Teilmenge aller Grünflächen und somit der grünen Landbedeckung darstellen, werden entsprechend der von der Weltgesundheitsorganisation vorgegebenen Mengen eingeteilt (World Health Organization, 2017). Demnach stellen zwischen 1,8 und 3,6 Hektar ein Mindestmaß von öffentlich zugänglichen Grünflächen in einer 500 m Nachbarschaft dar. Analog zur grünen Landbedeckung werden diese Schwellwerte genutzt um den Anteil der Bevölkerung mit entsprechendem Zugang zu öffentlichem Grün zu quantifizieren.

Der zweite Ansatz zur Untersuchung der Gerechtigkeit der Verteilung bedient sich etablierter Maße zur Analyse von Verteilungsgerechtigkeit. Hiermit ist es möglich die Gerechtigkeit der Verteilung einer Ressource innerhalb einer Gesellschaft zu bestimmen.

Im dritten Schritt werden mithilfe differenzierter demographischer Daten des Zensus einzelne Gesellschaftsgruppen betrachtet. Im Detail wird untersucht in welchem Zusammenhang die Verfügbarkeit von Grünflächen mit der demographischen Zusammensetzung der Nachbarschaft in Bezug auf den Anteil von Kindern, Senioren und Menschen mit Migrationshintergrund steht.

Durch die Anwendung auf ganz Deutschland eröffnen sich Analyseschritte, welche in bisherigen Studien nicht berücksichtigt werden konnten. Durch die große Grundgesamtheit ist es somit erstmals möglich, Gerechtigkeitsmaße über alle Städte und Gemeinden in Deutschland hinweg konsistent zu betrachten. Darüber hinaus ist es möglich Städte unterschiedlicher Größe miteinander zu vergleichen. Diese Tatsache wird im Rahmen der Analysen genutzt, um die Gerechtigkeitsmaße entlang des Grades der Urbanisierung der Siedlungsgebiete in Deutschland zu diskutieren.

Teile dieses konzeptionellen Schrittes wurden bereits in Weigand et al. (2023b) veröffentlicht.

Daten und Vorprozessierung

Für die einzelnen konzeptionellen Schritte dieser Dissertation werden verschiedenartige Datensätze kombiniert. 1) Für die Landbedeckungsklassifikation mit maschinellen Lernverfahren werden zwei Datensätze kombiniert: flächendeckende Satellitendaten und punktuelle Referenzdaten. 2) Zur Quantifizierung von öffentlichem Grün auf Nachbarschaftsebene mit *Deep Learning*-Verfahren werden weitere Referenzdaten in Polygonform, Satellitenbilddaten sowie eine Datenbank frei zugänglicher Geodaten mit landesweiter Ausdehnung eingesetzt. 3) Für den dritten konzeptionellen Ansatz, der bevölkerungsbezogenen Analyse der Grünflächenverteilung in Deutschland werden zusätzlich Daten zur Bevölkerung sowie Urbanität hinzugezogen.

4.1 LUCAS als Referenzdaten für Landbedeckung

Im Kontext der erdbeobachtungsgestützten Analyse der Umwelt sind präzise, aktuelle und repräsentative Referenzdaten unerlässlich. Hochgenau lokalisierte in-situ-Informationen spielen eine fundamentale Rolle für die Entwicklung von verlässlichen Klassifikationsmodellen auf Basis von Algorithmen (d'Andrimont et al., 2020). Da die manuelle und projektspezifische Erhebung derartiger Referenzdaten, speziell in großflächigen Anwendungen wie hier angestrebt, extrem zeit- und kostenintensiv ist, wird im Rahmen dieser Studie auf existierende in-situ-Daten zurückgegriffen.

Das paneuropäische *Land Use and Coverage Area Frame Survey*¹⁸ (LUCAS) ist ein umfassendes System zur Erfassung von in-situ-Informationen, das 2001 von der Europäischen Union ins Leben gerufen wurde, um statistische Daten zur Landnutzung und Landbedeckung zu erheben (Gallego und Delincé, 2010). Ausgehend von einem systematischen Gitter von Referenzpunkten Punkten mit einem Abstand von 2 km × 2 km, das sich über ganz Europa erstreckt, sammelt geschultes Personal in einem dreijährigen Turnus detaillierte Informationen im Feld oder per Interpretation von Luftbilddaten. Diese räumlich stratifizierte Verteilung ist ein großer Vorteil von LUCAS, da hierdurch die Anfälligkeit gegenüber einer räumlichen Autokorrelation der Datenpunkte verringert wird, die wichtig für die Unabhängigkeit und Repräsentativität der Daten ist. Jeder der ca. 1,1 Millionen Punkt des regelmäßigen 2 km Gitters wurde initial durch Fotointerpretation

¹⁸Eurostat Statistics explained, ohne Datum, „LUCAS – Land use and land cover survey“, zuletzt abgerufen 26.07.2023, https://ec.europa.eu/eurostat/statistics-explained/index.php?title=LUCAS_-_Land_use_and_land_cover_survey

einer vordefinierten Klasse von Landbedeckung zugeordnet (EUROSTAT, 2022). Für jede der dreijährlichen Erfassungen werden auf Basis dieser Landbedeckungsklassen sogenannte Feldproben (engl. *field sample*) ausgewählt. Abbildung 4.1 zeigt die Verteilung der LUCAS-Feldprobe 2015 in Europa eingefärbt nach den acht Hauptlandbedeckungsklassen der Erfassung.

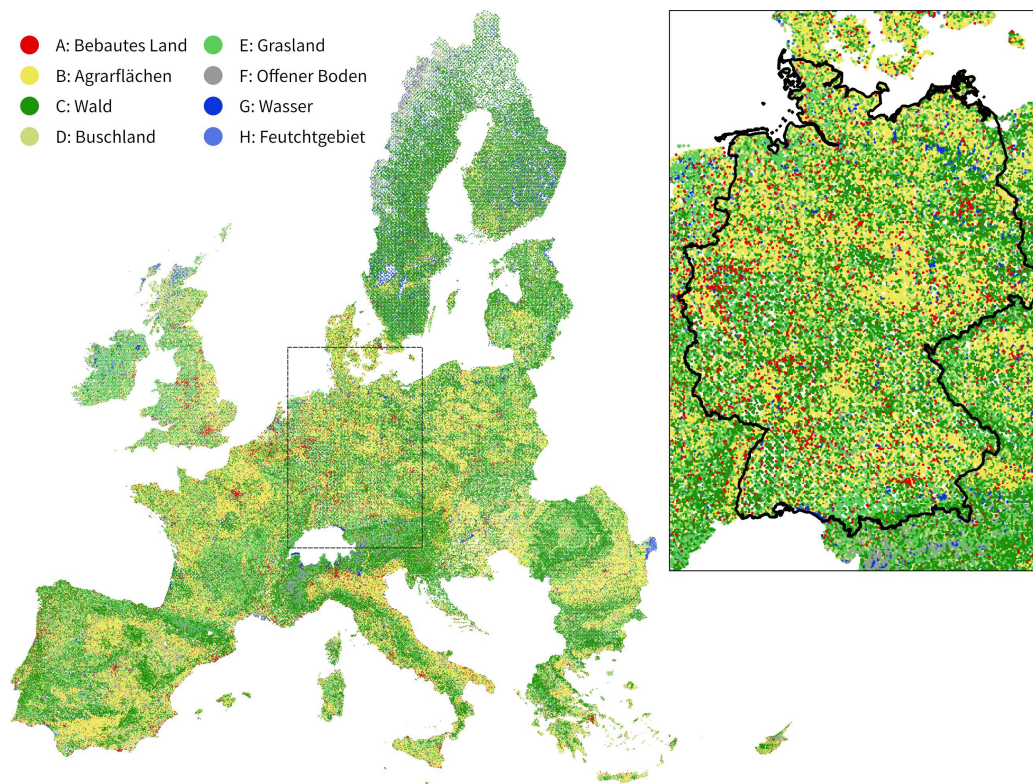


Abbildung 4.1: Überblick der LUCAS Feldproben-Punkte in Deutschland und den Ländern der Europäischen Union 2015. Insgesamt wurden 2015 338.359 Punkte untersucht, davon liegen 26.590 in Deutschland.

Im Feld werden Fotos von den einzelnen Punkten angefertigt und ökologische Parameter dokumentiert (Orgiazzi et al., 2018), darunter u. a. Landbedeckung, Landnutzung und Umwelteigenschaften. In Kombination mit detaillierten Metadaten über die Qualität der Beobachtung selbst, wie z. B. dem Erfassungsdatum oder dem GPS-Standort des Beobachters im Feld, werden die gesammelten in-situ-Informationen anschließend veröffentlicht. Des Weiteren werden die lokalen Gegebenheiten je Referenzpunkt mithilfe von 5 Fotos dokumentiert, eines vom zu erfassenden Punkt selbst, und vier in alle Himmelsrichtungen (Nord, Süd, West und Ost, d'Andrimont et al., 2020).

LUCAS in-situ-Daten wurden bereits als Trainings- und Validierungsdaten in Kombination mit Fernerkundungsdaten verwendet. Einige Studien konzentrierten sich auf die vegetative Bodenbedeckung, um Grünflächenkarten (Zillmann et al., 2014; Tassopoulou et al., 2019), landwirtschaftliche Bestandsaufnahmen (Conrad et al., 2010; Esch et al., 2014; Kussul et al., 2018) oder Vegetationsüberwachung (Khaliq et al., 2018) abzuleiten. Andere nutzten das breite Spektrum

der Landbedeckungsarten für semantisch umfassendere Ansätze der Landbedeckungsklassifikation (Mack et al., 2017; Close et al., 2018; Pflugmacher et al., 2019; Leinenkugel et al., 2019). In der Regel, insbesondere bei großräumigen Ansätzen, werden für solche Anwendungen Landsat-Bilder mit einer räumlichen Auflösung von 30 m verwendet. Im Gegensatz dazu wurde LUCAS in Kombination mit Sentinel-2, mit einer geometrischen Auflösung von bis zu 10 m, nur für lokale oder regionale Skalen (Khaliq et al., 2018; Close et al., 2018) verwendet. Eine großmaßstäbliche nationale Landbedeckungsklassifikation mit Sentinel-2 steht daher noch aus.

Aufgrund der hohen räumlichen Auflösung der Multispektraldaten von Sentinel-2 können Positionsfehler in den Referenzdaten die Ergebnissenauigkeit von Bildklassifikationsverfahren erheblich beeinträchtigen. Während der Aufnahme einzelner Messpunkte der LUCAS-Erhebung im Feld wird die Position, von der aus ein Probenpunkt vermessen wird, mit GPS-Geräten erfasst. Diese Methode ist anfällig für GPS-Positionsfehler, die eine Quelle von Unsicherheiten darstellen, die in früheren Forschungsarbeiten zu Fernerkundungsanwendungen diskutiert wurden (z. B. Pflugmacher et al., 2019; d'Andrimont et al., 2020, 2021b). Darüber hinaus beziehen sich die für jeden LUCAS-Punkt aufgezeichneten LB/LN-Informationen immer auf den ursprünglich vorgesehenen Standort, der auch als theoretischer Standort bezeichnet wird (EUROSTAT, 2015b). Dieser theoretische Punkt kann zum Zeitpunkt der Erhebung nicht immer erreicht werden, da er sich auf einem Gewässer oder Privatgrundstück (EUROSTAT, 2015a) befinden kann. Dies kann zu einem räumlichen Versatz zwischen dem GPS-Punkt und dem beobachteten theoretischen Standort führen. Wenn die GPS-Position in einer Fernerkundungsklassifikation zur Extraktion von Spektralinformationen für ein nachfolgendes maschinelles Lernverfahren verwendet wird, werden so ggf. Landbedeckungsinformationen falsch zugeordnet, was zu einer Verringerung der Klassifikationsgenauigkeit führen kann.

Die relevanten Informationen des LUCAS-Datensatzes bestehen aus der räumlichen Koordinate auf der einen Seite und den erfassten Landbedeckungs- bzw. Landnutzungsinformationen sowie weiteren Umweltparametern auf der anderen. Um ein optimales Klassifikationsprodukt aus diesen Daten abzuleiten, bedürfen Positionierung und Landbedeckungsinformationen gesonderter Aufbereitung. Im Folgenden werden die Eigenschaften, Probleme und nötigen Vorprozessierungsschritte für diese beiden Teilinformationen genauer beleuchtet.

4.1.1 LUCAS Positionierung

Jede LUCAS-Stichprobe befindet sich im Schnittpunkt eines regelmäßigen 2 km langen INSPIRE-Gitters¹⁹, in dem die LB/LN-Informationen erhoben werden. Bei der Erhebung des LUCAS-Datensatzes im Jahr 2015 war es den Beobachtern jedoch in 88,3 % der Fälle nicht möglich, den Stichprobenpunkt am genauen Standort zu erreichen, da diese sich auf Privatgrundstücken, in unzugänglichen (Feucht-)Gebieten oder in dicht bebauten städtischen Strukturen befanden.

¹⁹Europäische Kommission, ohne Datum, „INSPIRE Knowledge Base“, zuletzt abgerufen 26.07.2023: <https://inspire.ec.europa.eu/>

Schwer zugängliche Punkte können im Allgemeinen nur aus der Ferne beobachtet werden. In solchen Fällen wird der GPS-Standort aufgezeichnet, von dem aus der theoretische Punkt beobachtet wird, wie auch die Entfernung zwischen den Punkten. Das bedeutet, dass die aufgezeichnete GPS-Position, an der der Beobachter die Informationen über den Probepunkt sammelt, nicht unbedingt räumlich mit dem theoretischen Punkt übereinstimmt. So kann zum Beispiel ein LUCAS-Punkt, der im Meer liegt, vom Ufer aus zuverlässig bewertet werden.

Bis zur LUCAS-Erhebung 2015 enthielten die Daten lediglich diese GPS-Standorte als geographische Koordinaten. In der Konsequenz wurde die Beobachtungsdistanz als zusätzliches Qualitätskriterium für die Auswahl der LUCAS-Stichproben verwendet (z. B. Pflugmacher et al., 2019), was nicht ausschließen lässt, dass die GPS-Standorte für die Verortung der LUCAS-Stichproben verwendet wurden. Insbesondere bei hochauflösenden Bildern können jedoch sehr kleine räumliche Ungenauigkeiten Fehler in den Prozess der Landbedeckungsklassifikation einbringen. Beispielsweise würde ein Probenpunkt, der sich auf einem Gewässer befindet und vom Ufer aus beurteilt wird, als Folge der verschobenen Position eine falsche spektrale Signatur für nachfolgende maschinelle Lernprozesse liefern. Die Anwendung strenger Regeln für den Ausschluss von Proben hingegen, die aus der Entfernung bewertet werden, führt zum Verlust vieler Probenpunkte. Besonders in kleineren Untersuchungsgebieten kann dies zu sehr kleinen Stichproben führen. Die genauen Kriterien zur Auswahl der Punkte unterliegen so immer der Abwägung zwischen Probenanzahl und räumlich-semantischer Verlässlichkeit.

Der theoretische Standort eines LUCAS-Punktes (aus Gründen der Lesbarkeit im Folgenden als *GRID* bezeichnet), d. h. der Standort, für den die Landbedeckungsinformationen aufgezeichnet werden, ist bis 2015 lediglich in der Punkt-ID kodiert. Die Koordinate des *GRID*-Punktes kann damit rekonstruiert werden, z. B.:

- LUCAS-ID = 12345678
 - *GRID* Nordwert = 1.234.000
 - *GRID* Ostwert = 5.678.000
 - Bezugssystem: ETRS89-extended / LAEA Europe (EPSG: 3035).

Bei der Erhebung 2015 beträgt der durchschnittliche Abstand zwischen den Standorten von *GRID* und *GPS* 24,22 m (Median = 5 m, 1. Quartil = 2 m und 3. Quartil = 16 m). In der Literatur wurde bis dato noch nicht behandelt, welchen Einfluss diese räumliche Abweichung insbesondere für die Klassifikation von Sentinel-2 Bilddaten hat. Deshalb werden in der hier vorliegenden Arbeit stets beide Koordinatenpaare verwendet, um den systematischen Fehler zu ermitteln, der durch die Verwendung der *GPS*-Koordinaten entstehen kann.

4.1.2 LUCAS Vorprozessierung

Neben der Position der LUCAS-Punkte können auch die semantischen Landnutzungs- und Landbedeckungsinformationen sowie Metadaten, die für jeden LUCAS-Punkt gesammelt wurden, zur Auswahl geeigneter Referenzdaten genutzt werden. In verwandten Studien wurden verschiedene Ansätze für diesen Auswahlprozess verwendet. Diese Ansätze unterscheiden sich in Bezug auf den semantischen Detailgrad und die Qualitätskriterien, die zur Auswahl gültiger Stichproben angewendet werden. Während zum Beispiel Pflugmacher et al. (2019) insgesamt 12 LB/LN-Klassen verwenden, betrachten Close et al. (2018) nur fünf Zielklassen. Diese Unterschiede in der Anzahl der Zielklassen können sich auf die Genauigkeit des Klassifikationsmodells auswirken (Ma et al., 2017). Eine umfassende Bewertung dieser Ansätze zur Vorverarbeitung von LUCAS-Eingabedaten für ihre Verwendung in der Fernerkundungsklassifikation steht jedoch noch aus.

Die wichtigsten Informationen, die in der LUCAS-Datenbank enthalten sind, sind die detaillierten Informationen zu den an den Punkten vorherrschenden Landbedeckungs- und Landnutzungs-klassen. In Tabelle 4.1 sind die acht Hauptklassen (*A-H*) des LUCAS-Datensatzes aufgeführt, welche insgesamt 84 detaillierte Unterklassen zusammenfassen. Im Kontext der Klassifikation von Erdbeobachtungsdaten muss zwischen Landbedeckung und Landnutzung unterschieden werden. Fernerkundungsdaten aus satellitengestützten Sensoren eignen sich besonders gut zur Erfassung der physischen Erscheinung der Erdoberfläche (siehe Kapitel 2.3). Die Unterscheidung von Landbedeckungseigenschaften stehen daher im Vordergrund für bildgebende Klassifikationsverfahren. Landnutzungsarten haben dagegen hingegen semantischen Charakter und werden im Rahmen der Landbedeckungsklassifikation meist lediglich implizit verwendet. In der Konsequenz ist es in verwandten Studien üblich, die ursprüngliche Hierarchie von LUCAS Landbedeckungs- und Landnutzungsklassen zu reklassifizieren. Ein Beispiel hierfür sind Baumplantagen, wie z. B. Kirschbaumkulturen. In der ursprünglichen LUCAS-Hierarchie werden sie als Ackerland (*B*) eingestuft, obwohl sich ihr physisches Erscheinungsbild (z. B. die Vegetationshöhe) von den meisten anderen landwirtschaftlichen Flächen mit niedriger Vegetation unterscheidet, die in der Klasse *B* enthalten sind. Im Ablauf eines Bildklassifikationsvorgangs führt dies unweigerlich zu einer Heterogenität der Klassen und kann damit die Klassifikationsgenauigkeit beeinträchtigen.

Darüber hinaus enthalten die LUCAS-Beispiele eine Vielzahl qualitativer und quantitativer Metadaten. Bestehende Studien, die LUCAS als Referenzdaten verwenden, nutzen diese Metriken, um geeignete Punkte mit hoher Qualität für das Training und die Validierung einer Fernerkundungsbildklassifikation auszuwählen (z. B. Conrad et al., 2010; Mack et al., 2017; Pflugmacher et al., 2019). Diese unterschiedlichen Ansätze der Vorverarbeitung, im Folgenden als Vorprozessierungsschema (kurz Schema) bezeichnet und symbolisiert durch S , nutzt verschiedene Metadatenkriterien für die Auswahl der Punkte basierend auf der Relevanz für die zu untersuchenden Landbedeckungsklassen und der räumlichen Auflösung der Satelliten. Um die Auswirkungen der Neuklassifizierung und Auswahl der LUCAS-Stichproben zu bewerten, werden in dieser Arbeit vier verschiedene Schemata verglichen: S_0 als Orientierungswert, S_1 und S_2 , die aus aktuellen Studien übernommen werden, und S_3 als neu entwickeltes Schema für die Aufbereitung

Tabelle 4.1: Hauptlandbedeckungsklassen des LUCAS-Datensatzes. Für weitere Details siehe beispielsweise EUROSTAT (2015b).

Kürzel	Klassenbezeichnung
A	Bebautes Land
B	Agrarflächen
C	Waldflächen
D	Gebüsch
E	Grasland
F	Offener Boden, Moos und Flechten
G	Wasser
H	Feuchtgebiete

von LUCAS-Daten für hochaufgelöste Multispektraldaten der Sentinel-2 Satelliten. Die Tabelle 4.2 fasst die Anzahl der Referenzpunkte nach der Vorverarbeitung zusammen. Aufgrund der unterschiedlichen Auswahlkriterien unterscheidet sich die genaue Anzahl von Punkten zwischen den einzelnen Schemata, jedoch bei ähnlicher Größenordnung.

Tabelle 4.2: Die Anzahl der LUCAS-Stichproben in Deutschland für die in dieser Studie getesteten Vorprozessierungsschemata S . Der Positionierungsansatz wird durch *GRID* und *GPS* angegeben; die Unterschiede sind auf ungültige *GPS*-Punkte und auf Beobachtungsabstände zurückzuführen, die einen Schwellwert überschreiten.

S_0		S_1		S_2		S_3	
GRID	GPS	GRID	GPS	GRID	GPS	GRID	GPS
26745	24883	20879	19381	14568	14567	19291	17967

S_0 : Als Vergleichsschema werden keine LUCAS-Punkte anhand von Attributinformatoren aussortiert, sondern lediglich die LB-Klassen auf die ursprünglichen acht Hauptklassen (A-H) reduziert. Ungültige *GPS*-Punkte werden aus der Stichprobe ausgeschlossen.

S_1 : Das von Mack et al. (2017) entwickelte Schema behält sieben Landbedeckungsklassen aus der ursprünglichen LUCAS-Hierarchie vollständig bei und unterscheidet bei Waldflächen zwischen zwei Unterklassen: Laubwaldgebiete (engl. *broadleaved woodlands*, C1) und Nadelwaldgebiete (engl. *coniferous woodlands*, C2). Insgesamt umfasst es also neun Landbedeckungsklassen. Um reine Landsat-Pixel zu erhalten, schlossen die Autorinnen und Autoren Stichproben aus, die sich innerhalb eines homogenen Flecks von weniger als 1 ha befinden, wenn sie nicht als bebaute Flächen (engl. *artificial land*, A) klassifiziert sind. Diese Ausnahme war für den Klassifikationsprozess wichtig, da sonst fast keine Referenzpunkte für bebaute Gebiete erhalten geblieben wären. Manuelle

räumliche Korrekturen, wie sie in der ursprünglichen Arbeit von Mack et al. (2017) vorgenommen wurden, wurden in der hier vorliegenden Studie nicht durchgeführt, um Vergleichbarkeit und Reproduzierbarkeit mit anderen Schemata herzustellen.

S_2 : Im Rahmen einer europaweiten Anwendung unterschieden Pflugmacher et al. (2019) sowohl Wälder als auch landwirtschaftliche Flächen nach deren Saisonalität. Dieses Stichprobenschema unterscheidet neben Laub- und Nadelwäldern auch Mischwaldgebiete (engl. *mixed woodlands*, C3) und trennt permanente Schneeflächen bzw. Gletscher (engl. *permanent snow/glaciers*, G50) von anderen Wasserflächen (G). So entstehen insgesamt 12 Landbedeckungsklassen. Darüber hinaus nutzen die Autoren die detaillierten LUCAS-Attributinformationen, um bestimmte Klassen auszuschließen, wie z. B. nicht bebaute lineare Merkmale (engl. *non-built-up linear features*, A22), die auf Bildern mit niedriger Auflösung nicht zu erkennen sind, und temporäres Grasland (engl. *temporary grassland*, B55). Bezüglich der räumlichen Auflösung von Landsat-8 werden mehrere qualitative Filter angewendet. Dazu gehören eine Mindestgröße von 0,5 ha, ein Landbedeckungsgrad von mehr als 50 % und vor allem eine maximale Beobachtungsentfernung von 30 m. Sowohl für bebaute Flächen als auch für mehrjähriges Ackerland wurden von den Autoren weniger strenge Filterkriterien gewählt.

S_3 : Aufbauend auf früheren Ansätzen wurde von Weigand et al. (2020) ein weiteres Klassenschema für die hochauflösende Landbedeckungskartierung mit LUCAS und Sentinel-2 vorgestellt. Aufgrund der hohen räumlichen Auflösung von Sentinel-2 müssen alle Stichproben, die nicht der Klasse bebautes Land (engl. *artificial land*, A) zugehören, innerhalb einer homogenen Fläche von mehr als 0,5 ha liegen. Darüber hinaus wurden nur Beobachtungen berücksichtigt, die als „auf dem Punkt“ (engl. *on the point*) beschrieben sind. Dadurch wurden LUCAS-Referenzpunkte ausgeschlossen, die sich einer Grenze zwischen zwei LB/LN-Typen liegen, sei es beispielsweise an einer Feldgrenze oder an einem kleinen linearen Element, wie einem Pfad (vgl. EUROSTAT, 2015a, S. 18). Da mehrere verschiedene Landbedeckungsarten nebeneinander existieren können, z. B. Agroforstwirtschaft (vgl. EUROSTAT, 2015a, S. 26), werden alle Proben mit mehreren Landbedeckungsklassen, die sich innerhalb kleiner vegetativer Flecken mit einer Größe von weniger als 20 m × 20 m befinden (siehe EUROSTAT, 2015a, S. 25), ausgeschlossen. Wasserflächen und Feuchtgebiete, die zum Zeitpunkt der Begehung vorübergehend trocken oder überflutet sind (z. B. ein Flussbett EUROSTAT, 2015a, S. 99), werden ebenfalls ausgeschlossen, da davon auszugehen ist, dass es sich hierbei um eine zeitlich sehr heterogene Landbedeckungsklasse handelt. Das Schema S_3 unterscheidet grundsätzlich zwischen bebauten und natürlichen Flächen. Zu Letzteren zählen Vegetationsflächen, offener Boden und Gewässer. Vegetationsflächen werden aufgrund ihrer unterschiedlichen ökologischen und mikroklimatischen Auswirkungen nach Höhe und Saisonalität des Wachstums in vier Kategorien eingeteilt: Hohe mehrjährige Vegetation umfasst Nadelbaumbestände, wohingegen hohe saisonale Vegetation aus Laubbäumen besteht. Niedrige saisonale Vegetation ist durch starke saisonale Schwankungen im Jahresverlauf gekennzeichnet, wie beispielsweise Ackerland. Niedrige ganzjährige Vegetation ist beständiger, z. B. Grasland oder Sträucher. Es wird eine neue Nomenklatur eingeführt, welche in der Tabelle 4.3

zusammengefasst ist. Während einige Klassen am Boden für einen menschlichen Beobachter gut unterscheidbar sind, kann davon ausgegangen werden, dass diese Klassen in der Draufsicht der Erdbeobachtungssatelliten nicht oder nur schwer zu differenzieren sind, wie z. B. eine schmale, mit Bäumen überwachsene Straße (Conrad et al., 2010). Daher werden diese Klassen aus dem Schema grundsätzlich ausgeschlossen (vgl. Tabelle 4.3 rechte Spalte). Aus den ursprünglich acht LUCAS-Klassen werden während der Vorverarbeitung somit sieben neue LC-Klassen gebildet, die im Folgenden für die Bildklassifikation mit Sentinel-2 verwendet werden.

Tabelle 4.3: Neu entwickelte Klassenhierarchie der LUCAS-Punkte für die hochaufgelöste, deutschlandweite Landbedeckungskartierung mit Sentinel-2 nach dem Klassifikationsschema S_3 . Die rechte Spalte enthält alle in S_3 ausgeschlossenen LUCAS-Klassen. Tabelle modifiziert nach Weigand et al. (2020).

Klasse	Beschreibung	Beinhaltete LUCAS-Klassen	Grundsätzlich ausgeschlossene Klassen
A	Bebaute Areale	A	A13; A22 ausschließend derer mit den
S	Offener Boden	F10 – F30	LN Klassen: U21, U22, U31, U34,
T1	Hohe saison. Veg.	C10, B71 – B74	U36 or U37; A30; B55; B71 – B75; B83;
T2	Hohe mehrj. Veg.	C21 – C23	Bx1; Bx2; C3; D10; E10; E30; F40; G50;
V1	Niedr. saison. Veg.	B (exkl. B71 – B74, B84)	H21 – H23
V2	Niedr. mehrj. Veg.	B84, D20, E20, H11, H12	
W	Wasserflächen	G	

4.2 Sentinel-2 Bilddaten

Die Wahl der Satellitenbilddaten für die Erstellung der bundesweiten Landbedeckungsklassifikation zugunsten von Sentinel-2 wird von verschiedenen Aspekten beeinflusst. Einerseits ist es im Rahmen der Analyse von Grünflächen sinnvoll auf einen Sensor mit ausreichender spektraler Auflösung zurückzugreifen, um neben sichtbarem Licht auch z. B. nahes Infrarot, u. A. für spektrale Indizes nutzen zu können. Wie in Kapitel 2.3 ausgeführt, helfen diese Informationen dabei Vegetationsflächen gut zu erkennen. Darüber hinaus ist es für den hier geplanten großflächigen Ansatz hilfreich, auf eine große Anzahl von Bildaufnahmen mit gleichzeitig großer Aufnahmefläche zurückgreifen zu können, sodass ausreichend wolkenfreie Satellitenbildszenen zu einem Mosaik für das gesamte Untersuchungsgebiet zusammengefasst werden können. Auch die hohe räumliche Auflösung von $10\text{ m} \times 10\text{ m}$ ist besser geeignet als beispielsweise die der Landsat Satelliten, um heterogene Stadtgebiete möglichst genau beschreiben zu können. Für die Entwicklung der hier beschriebenen Methode stehen zunächst die Bilddaten des am 23.05.2015 gestarteten Sentinel-2 A Satelliten zur Verfügung.

Optische Erdbeobachtungsdaten wie die des Sentinel-2 sind abhängig von Wetter und anderen atmosphärischen Verhältnissen. Wolkenbedeckung, Dunst, Rauch oder andere atmosphärische Einflüsse können die Sicht auf den Erdboden in manchen Wellenlängenbereichen erschweren.

Dies wiederum resultiert in der Veränderung des gemessenen Lichtspektrums am betrachteten Bildpunkt im Vergleich zu einer Aufnahme unter optimalen Bedingungen bei klarem Himmel. Um diese Effekte zu entfernen, können atmosphärische Korrekturen mittels algorithmischer Methoden durchgeführt werden (z. B. ATCOR[®] Richter und Schläpfer, 2016a,b). Derartige Ansätze sind jedoch aufwändig und erfordern zusätzliche Daten. Um atmosphärische Einflüsse auch ohne komplexe atmosphärische Korrekturen möglichst gering zu halten, ist es ebenfalls möglich ein Aggregat aus einer großen Zahl von Satellitenbildern abzuleiten, welches Extremwerte, wie sie bei Wolken, Schatten und Dunst zu erwarten sind, minimiert (Gómez et al., 2016). So können beispielsweise Satellitenbilder anhand der Gesamtwolkenbedeckung, welche den Metadaten zu entnehmen ist, selektiert werden. Dies kann jedoch in Gebieten mit häufiger Wolkenbedeckung die Anzahl an verfügbaren Bilddaten stark reduzieren.

Im hier vorgelegten Fall werden zunächst alle Sentinel-2 (Level 1C, L1C) Szenen über Deutschland mit einer Wolkenbedeckung unter 60 % zur Erstellung eines landesweiten Mosaiks ausgewählt. Eine zusätzliche Wolkenmaskierung wird unter Verwendung des in den Sentinel-2 Bildern enthaltenen QA60-Bandes durchgeführt. Es werden nur Sentinel-2 Szenen im Zeitraum von Juni 2015 bis April 2017 ausgewählt, insgesamt 389 Szenen. Abbildung 4.2 zeigt die Anzahl der wolkenfreien Aufnahmen pro Pixel im Untersuchungsgebiet.

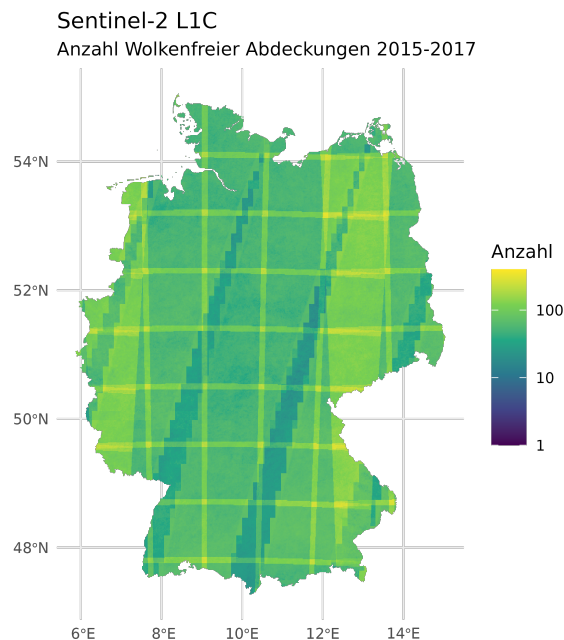


Abbildung 4.2: Abdeckung Deutschlands mit wolkenfreien Sentinel-2 L1C Bilddaten für den Zeitraum zwischen 2015 und 2017, eingefärbt in logarithmischer Skala. Die Anzahl der wolkenfreien Aufnahmen pro Pixel variiert dabei zwischen 23 (1. Perzentil, 5. Perzentil = 30) und 218 (99. Perzentil, 95. Perzentil = 145). Rechteckige Artefakte in Gelb zeigen die Überlagerung der einzelnen Bildausschnitte (engl. *granules*), die zur internen Datenhaltung in der Prozessierungsumgebung Google Earth Engine verwendet werden. Die diagonalen Streifen ergeben sich durch die Geometrie der Flugbahn der Sentinel-2 Satelliten.

Aus der Gesamtzahl aller Bildaufnahmen werden mehrere Mosaik abgeleitet. Für alle ausgewählten Sentinel-2 Szenen werden die Bänder 2, 3, 4, und 8 selektiert. Diese entsprechen den Reflektanzen im blauen (B), grünen (G) und roten (R) Bereich des sichtbaren Lichts sowie dem nahen Infrarot (NIR). Für jedes dieser Bänder werden alle Aufnahmen aus den Sommermonaten (Mai bis September) überlagert. Anschließend wird für jedes Pixel im Untersuchungsgebiet der Medianwert ermittelt. Das so entstandene flächendeckende Mosaik, welches den Einfluss von Wolken und Wolkenschatten minimiert (Gómez et al., 2016). Für die weitere Verarbeitung werden die Kanäle einzeln aus der Google Earth Engine (siehe Gorelick et al., 2017) exportiert und zu einem Komposit aus Medianmosaik der Spektralbänder 2, 3, 4 und 8 mit 10 m räumlicher Auflösung zusammengefügt. Abbildung 4.3 zeigt eine Echtfarbandarstellung des resultierenden Mosaiks.

Zusätzlich zu den Aggregaten der spektralen Bänder werden Mosaik für drei spektrale Indizes errechnet: *Normalized Difference Vegetation Index* (Siehe Kapitel 2.3 für Details, insbes. Formel 2.1), *Normalized Difference Water Index* (NDWI McFeeters, 1996), und *Normalized Difference Built-Up Index* (NDBI Zha et al., 2003). Letztere sind definiert wie folgt:

$$NDWI = \frac{GREEN - NIR}{GREEN + NIR} \quad (4.1)$$

$$NDBI = \frac{SWIR - NIR}{SWIR + NIR} \quad (4.2)$$

wobei *GREEN* die Reflektanz im grünen Spektralbereich darstellt (Band 3), *NIR* das nahe Infrarot (Band 8) und *SWIR* das kurzwellige Infrarot (Band 11).

Diese Indizes wurden bereits in verwandten Ansätzen zur LB/LN-Klassifikation (Griffiths et al., 2013; Pelletier et al., 2016; Leinenkugel et al., 2019; Tassopoulou et al., 2019) erfolgreich eingesetzt. In dieser Arbeit werden die Spektralindizes für alle Sentinel-2 Szenen im Studienzeitraum errechnet, jedoch werden die Indizes anders die optischen Daten, für den gesamten Zeitraum von Juni 2015 bis April 2017 errechnet, d. h. auch für die Wintermonate. Dies ermöglicht die Nachbildung der Saisonalität der Vegetationsflächen über das ganze Jahr hinweg und liefert nützliche Informationen für die Unterscheidung der vegetativen LC-Typen (Griffiths et al., 2014; Mack et al., 2017; Griffiths et al., 2019). Die Vielzahl an Messwerten pro Pixel wird wie auch die optischen Bilddaten mithilfe statistischer Aggregation reduziert. Für die Indizes werden drei Perzentile errechnet, das 25., das 50. (Median) und das 75. Perzentil. So entstehen insgesamt drei Mosaik für jeden Index. Abbildung 4.4 zeigt die Medianaggregate des NDVI, NDBI und NDWI in Deutschland. Aus der Kombination von vier Spektralbändern und insgesamt neun spektralen Indizes resultieren somit 13 deutschlandweite Bildebenen.

Um die Prädiktionsgenauigkeit des Klassifikationsverfahrens auf Basis von maschinellem Lernen zu erhöhen, werden aus diesen ersten Bildebenen zusätzliche Merkmale abgeleitet. Es wurde bereits nachgewiesen, dass Texturmerkmale dazu beitragen, die Klassifikationsgenauigkeit

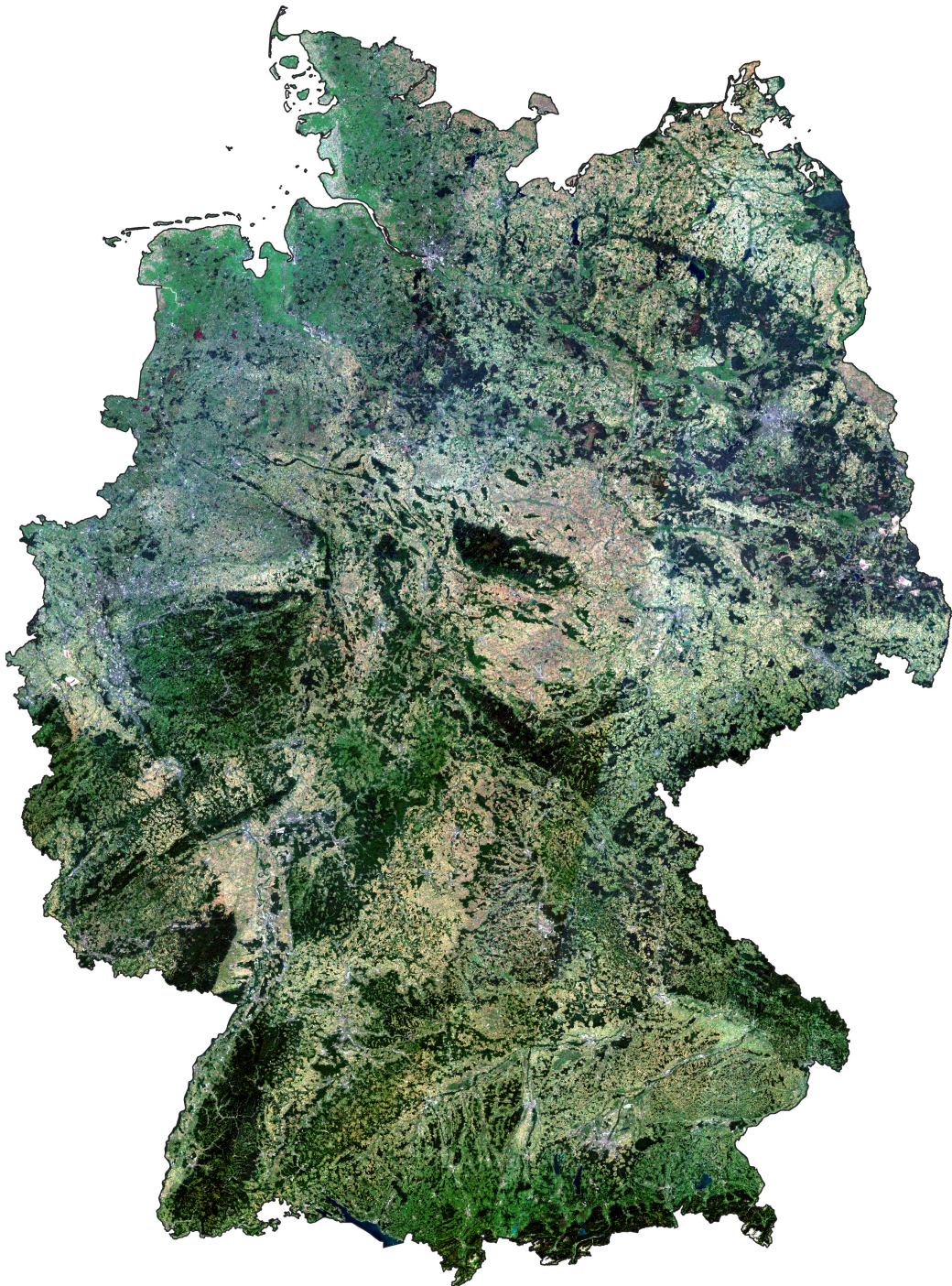


Abbildung 4.3: Landesweites Mosaik von Sentinel-2 Bilddaten über Deutschland. Es handelt sich um eine Echtfarbendarstellung der Sentinel-2 Spektralkanäle 4, 3, und 2 (rot, grün, blau). Die räumliche Auflösung des Datensatzes ist $10\text{ m} \times 10\text{ m}$.

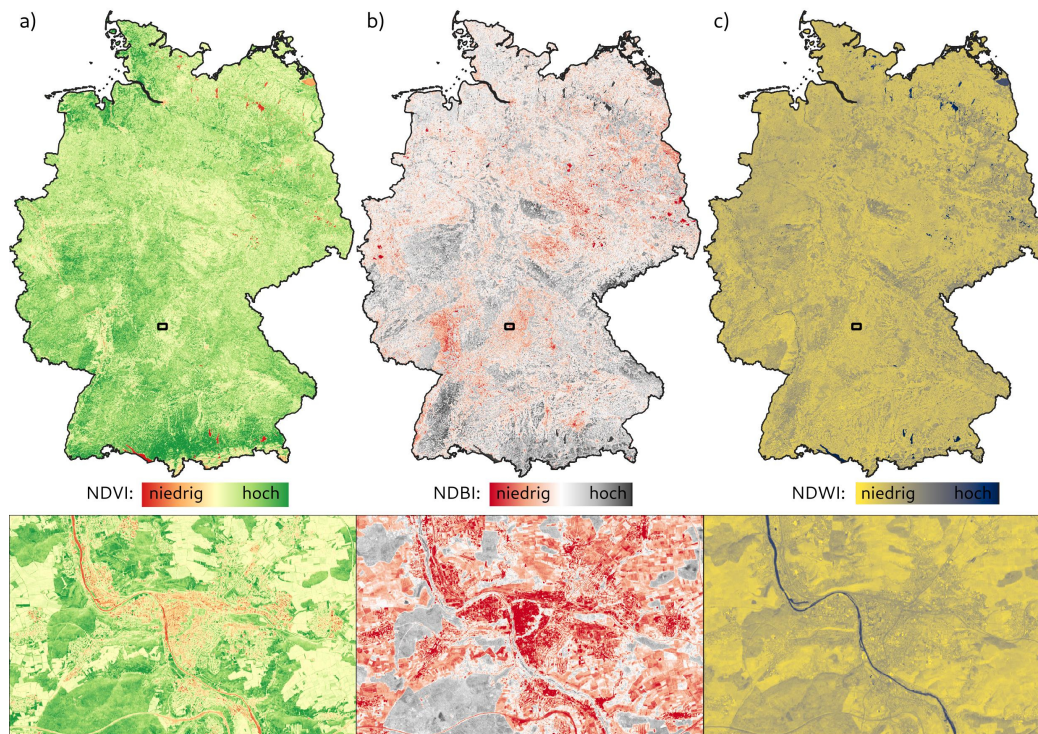


Abbildung 4.4: Landesweite Medianaggregate von NDVI (a), NDBI (b) und NDWI (c) aus Sentinel-2 Bilddaten über Deutschland (oben). Detailausschnitte von der Stadt Würzburg (unten). Indizes Die räumliche Auflösung der Datensätze ist $10\text{ m} \times 10\text{ m}$.

von Landbedeckungsdaten signifikant zu erhöhen (Khatami et al., 2016; Li et al., 2014). Bei einer räumlichen Auflösung von $10\text{ m} \times 10\text{ m}$ werden viele der Landbedeckungsklassen, wie z. B. urbane Umgebungen, durch ihre angrenzende Nachbarschaft von Pixeln definiert, zumal mehrere der in diesem Klassifikationsansatz verwendeten LUCAS-Klassen die lokale Landbedeckung innerhalb eines kreisförmigen Bereichs von bis zu 20 m (Karydas et al., 2015) beschreiben. Deshalb werden im Rahmen dieser Studie fokale statistische Texturmaße erster Ordnung (Median, Mittelwert und Standardabweichung) für drei Nachbarschaften (Fenstergrößen von 3×3 Pixel, 5×5 Pixel und 9×9 Pixel) für alle 13 Bildebenen errechnet. Darüber hinaus werden für die vier spektralen Mosaik (R, G, B, NIR) Texturen zweiter Ordnung auf Basis der Graustufen-Koinzidenzmatrix (engl. *grey level co-occurrence matrix*, *GLCM*, Haralick et al., 1973; Haralick, 1979) berechnet. Derartige lokale Bildstatistiken wurden bereits mehrfach erfolgreich im Kontext der Erdbeobachtungsklassifikation eingesetzt (Shanmugan et al., 1981; Du et al., 2015; Wurm et al., 2017a,b). Insgesamt werden so 226 Bildschichten aus den Sentinel-2 Bildern abgeleitet.

Zusätzliche Informationen, insbesondere zur Unterscheidung von sehr heterogenen Landbedeckungsflächen, werden durch landesweite Versiegelungsinformationen für drei Bebauungstypen bereitgestellt; diese sind Schienen, Straßen und Gebäude. Für Schienen und Straßen werden

Daten aus dem OpenStreetMap-Projekt²⁰ bundesweit extrahiert (siehe Tabelle A1 im Anhang für verwendete Elemente; für detaillierte Beschreibung der Datenbeschaffenheit siehe Kapitel 4.4). Gebäudegrundrisse werden aus Level-of-Detail-1-Gebäudemodellen des deutschen Bundesamts für Kartographie und Geodäsie abgeleitet. Für jedes 10 m-Pixel der deutschlandweiten Rasterdaten wird der relative Flächenanteil der Überlagerung der jeweiligen Polygondatensätze errechnet. Mit diesen drei zusätzlichen Versiegelungsebenen werden somit insgesamt 229 Schichten als Datenbasis für die Bildklassifizierung verwendet. Diese sind in Tabelle 4.4 übersichtlich zusammengefasst.

Tabelle 4.4: Überblick über die aus Sentinel-2 Mosaiken und Hilfsdaten berechneten Bildebenen für Deutschland. Fokale Bildstatistiken und die Graustufen-Koinzidenzmatrix (engl. *grey level co-occurrence matrix, GLCM*) werden in drei Fenstergrößen abgeleitet (3 × 3 Pixel, 5 × 5 Pixel und 9 × 9 Pixel).

	Mosaik	Fokale Statistiken			GLCM Texturmaße			Total
		3 × 3	5 × 5	9 × 9	3 × 3	5 × 5	9 × 9	
a) Spektrale Bänder	4	12	12	12	32	32	32	136
b) Spektrale Indizes	9	27	27	27				90
c) Zusatzdaten	3							3

$$\Sigma = 229$$

4.3 European Urban Atlas als Referenzdaten für öffentliche Grünflächen

Im Rahmen der Dissertation soll die Analyse über die rein physische Beschreibung des städtischen Grüns mittels grüner Landbedeckung hinaus erweitert werden. Hierfür bedarf es räumlich-quantitativer Daten zu öffentlichen Grünanlagen (engl. *public green spaces, PGS*). Aufbauend auf Ansätzen in früheren Studien (z. B. Kabisch und Haase, 2014; Krekel et al., 2016; Wüstemann et al., 2017b; Oehrlein et al., 2019) werden Daten aus dem europäischen Städteatlas 2018 (engl. *European Urban Atlas, EUA*)²¹ als Grundlage für die Quantifizierung der Verfügbarkeit öffentlicher Grünflächen verwendet. Zur Abgrenzung von öffentlichen Grünflächen werden die EUA-Klassen grüne städtische Gebiete (engl. *green urban areas*), Wälder (engl. *forests*) und krautige Vegetationsverbände (engl. *herbaceous vegetation associations*) in Übereinstimmung mit der bestehenden Literatur ausgewählt. Abbildung 4.5 zeigt die räumliche Verteilung und Ausdehnung der Landbedeckungs- und Landnutzungsklassen des EUA in München.

²⁰OpenStreetMap, zuletzt abgerufen am 27.07.2023: <https://openstreetmap.org>

²¹Copernicus, ohne Datum, „Urban Atlas“, zuletzt abgerufen 27.07.2023 <https://land.copernicus.eu/local/urban-atlas>

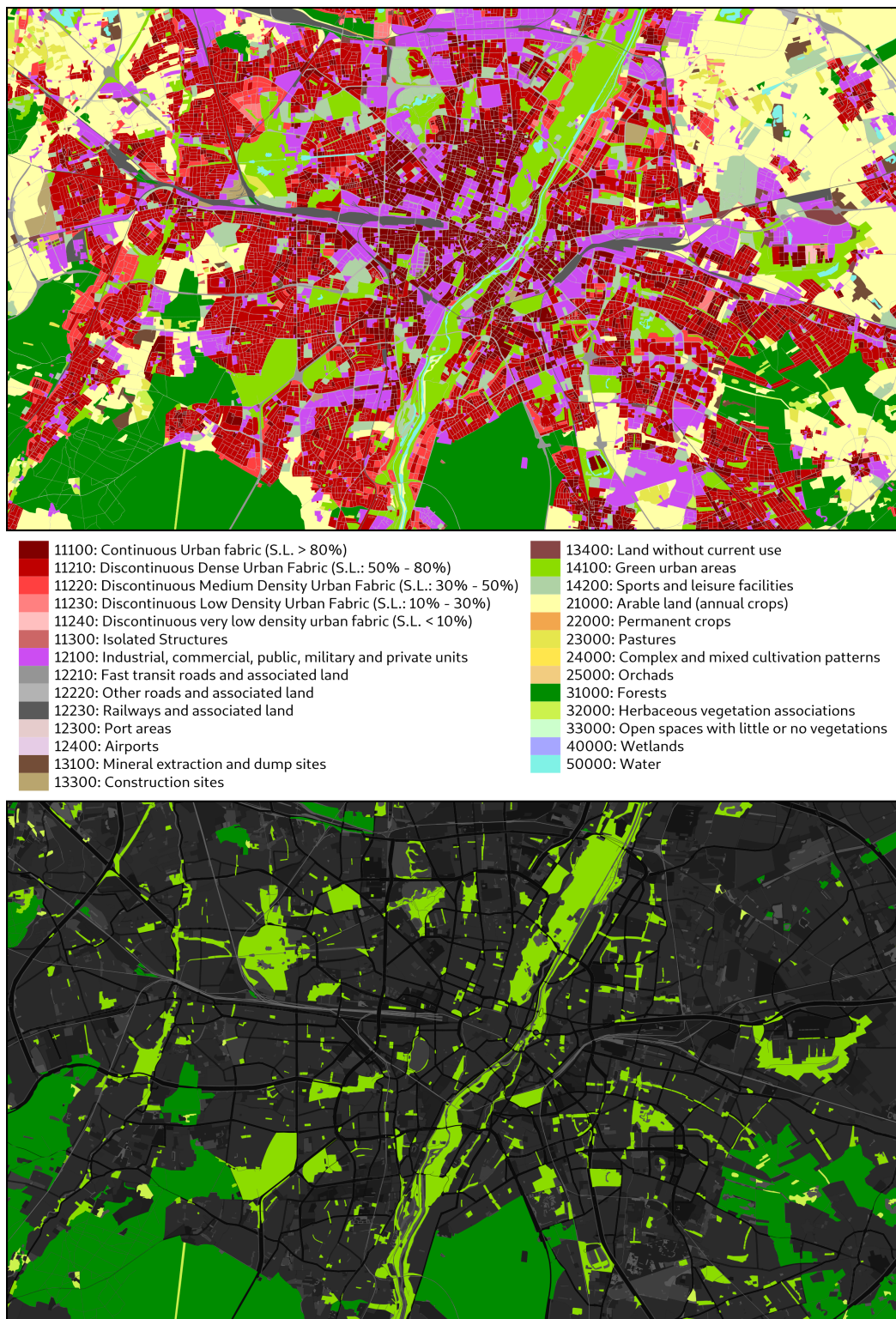


Abbildung 4.5: European Urban Atlas Datensatz in München (oben). Die Klassen (engl.) *Green Urban Areas*, *Forests* und *Herbaceous Vegetation Associations* (unten), die in dieser Studie als Referenzdatensatz dienen für das Training eines maschinellen Lernverfahrens zu Quantifizierung der Verfügbarkeit von öffentlichem Grün.

Der EUA-Datensatz deckt jedoch nicht die gesamte Fläche Deutschlands ab, sondern nur 96 räumlich-funktionale städtische Gebiete, in denen 75 % der deutschen Bevölkerung leben (vgl. Abbildung 4.6). Um die Analysen im Sinne der Grundmotivation der vorliegenden Arbeit über die Grenzen dieser Gebiete hinaus auszudehnen und somit ganz Deutschland abzudecken, wird die Verfügbarkeit öffentlicher Grünflächen mittels Datenfusion mehrerer räumlich umfassender Datensätze und eines maschinellen Lernansatzes extrapoliert. Ähnlich wie bei neueren Ansätzen, die Satellitenbilddaten und vektorbasierte geographische Daten kombinieren (Rosier et al., 2022; Georganos et al., 2022), werden Satellitenbilder und OpenStreetMap (OSM)-Daten kombiniert, um die Zielgröße zu modellieren: die Menge der verfügbaren öffentlichen Grünflächen in der Nachbarschaft. Hierfür dient der Datensatz des European Urban Atlas als Referenz. Die Geometrien der EUA-Daten werden in eine PostgreSQL / PostGIS Datenbank integriert, um im weiteren Verlauf der Datenaufbereitung effizienten Zugriff darauf zu ermöglichen.

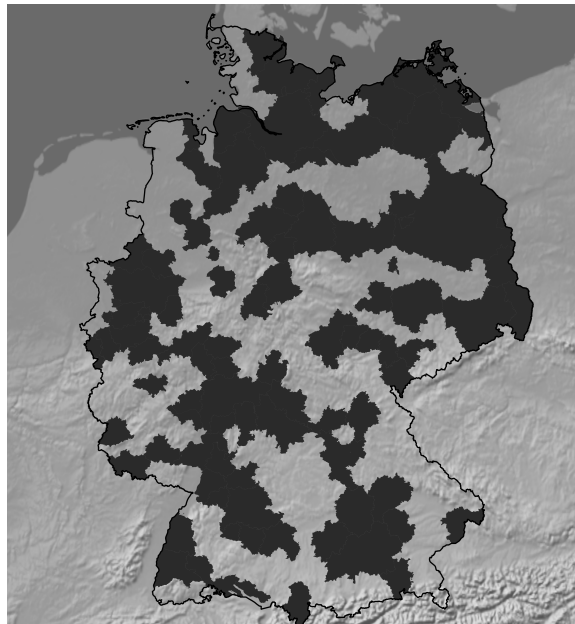


Abbildung 4.6: Abdeckung Deutschlands durch die 96 funktionalen urbanen Gebiete (engl. *functional urban areas*), für welche Daten des (engl.) *European Urban Atlas* zur Verfügung stehen.

4.4 OpenStreetMap Daten zu Grünflächentypen

OpenStreetMap (OSM) ist ein Projekt zur Schaffung und Verbreitung einer frei verfügbaren und editierbaren Datenbank von Geodaten weltweit²². Die Sammlung von räumlichen Vektordaten in Form von Punkten, Linien und Polygonen sowie detaillierten Attributdaten und sind für diverse Zwecke einsetzbar, z. B. Kartographien oder Streckennavigation. OpenStreetMap-Daten stellen

²²OpenStreetMap Wiki, ohne Datum, Hauptseite, zuletzt abgerufen 27.07.2023, https://wiki.openstreetmap.org/wiki/Main_Page

eine interessante Datenquelle dar, um in Kombination mit Erdbeobachtungsdaten hochdetailliert Landnutzungsinformationen abzuleiten (Vargas-Munoz et al., 2021). Beispielsweise verbesserte die Hinzunahme von Sehenswürdigkeiten oder anderen Landmarken (engl. *points of interes*, *POI*) und Straßennetzwerken zu Satellitenbilddaten die Klassifikationsgenauigkeit diverser Landnutzungsklassen (Rosier et al., 2022). OpenStreetMap-Daten zeigen große Ähnlichkeiten mit der anderen Landnutzungskartierungen in europäischen Großstädten (Arsanjani und Vaz, 2015). Abbildung 4.7 zeigt die Standardansicht der OSM-Kartendarstellung (oben) sowie die zugrundeliegenden Daten, wie sie im Webeditor verfügbar sind (unten). Die Daten der OSM-Datenbank unter der (engl.) *Open Database License (ODbL)*²³ verfügbar können jederzeit heruntergeladen werden²⁴. Beispielsweise stellt die Geofabrik GmbH wöchentliche Datenbankauszüge für einzelne Regionen bereit²⁵.

Zur Repräsentation räumlicher Informationen stehen in OSM drei Elemente zur Verfügung²⁶: Knoten (engl. *nodes*), Wege (engl. *ways*) und Relationen (engl. *relations*). Wege eignen sich zur Darstellung von Linienobjekten oder Flächen. Linien werden in Form offener Wege abgebildet, d. h. der erste und letzte Punkt sind nicht identisch. Flächen hingegen sind geschlossene Wege, deren erste und letzte Stützpunkte identisch sind. Zusätzlich zur räumlichen Information verfügt jedes Objekt in OSM über Attributinformationen. Anders als gängige tabellarische Geodatenformate, wie z. B. Geopackage²⁷, wird in der OSM-Datenbank ein System von Schlüsselwerten (engl. *tags*) im (engl.) *key=value* Format verwendet. Dies bietet den Vorteil, dass Objekte individuell und frei attribuiert werden können und nicht einem vorgegebenen Schema von Attributspalten müssen. Umfangreiche Dokumentation und Hilfestellung²⁸ für die freiwillig Beitragenden der OSM-Gemeinschaft stellen dabei die Grundlage für konsistente Attributierung der kartierten Objekte dar. Codeblock 4.1 zeigt beispielhaft die Notation eines OSM-Objektes anhand des Parks „Hirschgarten“ in München.

²³Open Knowledge Foundation, Open Data Commons, ohne Datum, „Open Data Commons Open Database License (ODbL)“, zuletzt abgerufen 27.07.2023, <https://www.opendatacommons.org/licenses/odbl>

²⁴An dieser Stelle möchte ich allen Mitwirkenden des OpenStreetMap-Projekts danken. Der freie Zugang zu OpenStreetMap-Daten unter den Bedingungen der ODbL ist fundamental für diese Arbeit.

²⁵Geofabrik Downloads, ohne Datum, „OpenStreetMap Data Extracts“, zuletzt abgerufen 27.07.2023, <https://download.geofabrik.de>

²⁶OpenStreetMap-Wiki, 12.05.2023, „Elements“, zuletzt abgerufen 27.07.2023, <https://wiki.openstreetmap.org/wiki/Elements>

²⁷Open Geospatial Consortium (OGC), ohne Datum, „GeoPackage“, zuletzt abgerufen 27.07.2023, <https://www.geopackage.org/>

²⁸OpenStreetMap-Wiki, 17.06.2023, „Map Features“, zuletzt abgerufen 27.07.2023, https://wiki.openstreetmap.org/wiki/Map_features



Abbildung 4.7: Beispiele der standardisierten OpenStreetMap Webkarte im „OSM Carto“-Stil (oben) und den zugrundeliegenden Daten im Webeditor „iD“ (unten) für den Park „Hirschgarten“ in München. Bildquelle: © OpenStreetMap contributors.

Codeblock 4.1: Beispiel für ein OpenStreetMap-Objekt in Form der XML-Notation. Zeilenumbrüche aus Formatierungsgründen sind durch ↵ gekennzeichnet. Interessant für diese Arbeit sind vor allem die *key=value*-Paare, z. B. `<tag k="leisure" v="park"/>`, sowie die Geometrien (hier wird auf die Geometrie eines anderen Objektes verwiesen `<member type="way" ref="148696657" role="outer"/>`).

```
<osm version="0.6" generator="CGImap 0.8.8 (2794161 spike-06.openstreetmap.
  ↵ org)" copyright="OpenStreetMap and contributors" attribution="<http://
  ↵ www.openstreetmap.org/copyright>" license="<http://opendatacommons.org
  ↵ /licenses/odbl/1-0/>">
<relation id="1396807" visible="true" version="11" changeset="100687449"
  ↵ timestamp="2021-03-09T07:39:49Z" user="map_per" uid="11758604">
  <member type="way" ref="148696657" role="outer"/>
  <tag k="leisure" v="park"/>
  <tag k="name" v="Hirschgarten"/>
  <tag k="type" v="multipolygon"/>
  <tag k="wikidata" v="Q261850"/>
  <tag k="wikipedia" v="de:Hirschgarten (München)"/>
</relation>
</osm>
```

Im Rahmen dieser Arbeit werden für die Quantifizierung der Präsenz, Absenz oder Zugänglichkeit von (öffentlichen) Grünflächen zunächst alle relevanten *key=value*-Paare identifiziert. In der Vielzahl von möglichen Attributinformationen sind jedoch viele ähnliche Objekte enthalten, weswegen alle für diese Studie relevanten OSM-Merkmale in semantisch konsistente Sammlungen verschiedener Schlüsselwertpaare gruppiert werden. Diese Merkmalsaufbereitung geschieht, um ähnliche Informationen aus verwandten Tags zu konsolidieren, von denen jedes für sich genommen auf nationaler Ebene potenziell räumlich sehr dispers verteilt ist und somit nur selten in den Nachbarschaften zu finden ist. Insgesamt werden 71 Gruppen gebildet. Dazu gehört z. B. die Gruppe „öffentliche Freizeiteinrichtungen“, die alle mit den Schlüsselwertpaaren `leisure=common`, `leisure=dog_park`, `leisure=fitness_station`, `leisure=nature_reserve`, `leisure=park`, `leisure=picnic_table`, `leisure=playground` und `leisure=slipway` gekennzeichneten Merkmale beinhaltet. Die vollständige Tabelle der OSM Gruppierung, der zugeordneten *key=value*-Paare sowie der jeweils relevanten Geometrielemente ist im Anhang zu finden, siehe Tabelle [A2](#).

Zur effizienten Abfrage der Objekte im Rahmen der deutschlandweiten Analyse werden die OpenStreetMap-Daten in Form einer PostgreSQL / PostGIS Datenbank abgespeichert. Hierbei werden die Knoten in Punktgeometrien und Wege sowie Relationen in Linien- und Polygoneometrien umgewandelt.

4.5 Zensus 2011

Für die Verortung der Bevölkerung in Deutschland werden die neuesten verfügbaren Zensusdaten aus dem Jahr 2011 verwendet. Das Statistische Bundesamt (destatis) stellt demographische Bevölkerungsdaten in räumlichen Rastern nach dem INSPIRE-Standard zur Verfügung²⁹ (Neutze, 2015). Im Rahmen dieser Arbeit werden drei verschiedene Datensätze aus dem Zensus 2011 abgeleitet: 1) Gesamtbevölkerung, 2) Bebauungstypen und 3) Demographie auf Nachbarschaftsebene.

Mit dem ersten Datensatz wird die räumliche Verteilung der Bevölkerung in der Bundesrepublik Deutschland erfasst. Die Gesamtzahl der lokalen Bevölkerung wird in einem INSPIRE-konformen Rasterdatensatz mit einer räumlichen Auflösung von 100 m × 100 m (1 ha) erfasst. Die Zuordnung der Bewohnerinnen und Bewohner zu einer Gitterzelle basiert auf der Wohnadresse. Insgesamt gibt es in Deutschland 3.177.723 bewohnte Gitterzellen im Zensus 2011.

Der zweite Datensatz beschreibt die räumliche Verteilung der Wohngebäudetypen in ganz Deutschland ebenfalls auf der Basis von 1 ha-Rasterzellen. Genauer handelt es sich um die Anzahl der Häuser nach Typ je Rasterzelle. In der vorliegenden Arbeit werden diese Informationen genutzt, um den Zusammenhang zwischen bestimmten Wohnungstypen und der Verfügbarkeit von Grünflächen zu analysieren. Auf der räumlichen Ebene von 1 ha aggregiert der Zensus die Anzahl der Gebäude in sieben Klassen, welche 4 Obergruppen zuzuordnen sind (vgl. Tabelle 4.5). Im Rahmen dieser Arbeit werden Bebauungstypen nach der übergeordneten Klasse bewertet. Für die Analyse der Grünflächenverfügbarkeit werden nur Zensus-Gitterzellen ausgewählt, die mindestens fünf Gebäude enthalten und nur Gebäude derselben aggregierten Klasse aufweisen. So wird der Einfluss von Zellen mit gemischten Bebauungstypen ausgeschlossen und es ist möglich die übergeordneten Tendenzen des Zusammenhangs zu identifizieren.

Tabelle 4.5: Aggregationsschema für Gebäudetypklassifikation im Zensus 2011. Die Anzahl der Zellen, welche mindestens fünf Gebäude der jeweiligen Klasse enthalten, ist ebenfalls aufgeführt.

Klassen	Anzahl Zellen	Zensus Gebäudetypklassifikation
Freistehendes Haus	716949	Freistehendes Einfamilienhaus
		Freistehendes Zweifamilienhaus
Doppelhaus	42066	Einfamilienhaus: Doppelhaushälfte
		Zweifamilienhaus: Doppelhaushälfte
Reihenhaus	34425	Einfamilienhaus: Reihenhaus
		Zweifamilienhaus: Reihenhaus
Mehrfamilienhaus	22214	Mehrfamilienhaus: 3 - 6 Wohnungen
		Mehrfamilienhaus: 7 - 12 Wohnungen
		Mehrfamilienhaus: 13 Wohnungen und mehr

²⁹Zensus 2011, ohne Datum, „Ergebniss des Zensus 2011 zum Download - erweitert“, zuletzt abgerufen 27.07.2023, <https://www.zensus2011.de/DE/Home/Aktuelles/DemografischeGrunddaten.html>

Der dritte Datensatz, der aus dem Zensus 2011 abgeleitet wird, enthält aggregierte demographische Informationen über die Zusammensetzung der Bevölkerung in der Nachbarschaft. Diese Daten werden in Gitterzellen mit einer räumlichen Auflösung von $1 \text{ km} \times 1 \text{ km}$ (1 km^2) bereitgestellt. Es werden drei Metriken zur demographischen Zusammensetzung verwendet, die Informationen über bestimmte breit definierte Gesellschaftsgruppen enthalten. Diese decken sich mit Bevölkerungsgruppen, die in der Literatur über Grünflächen regelmäßig zu finden sind: 1) der Anteil von Kindern unter 18 Jahren (Rehling et al., 2021), 2) älteren Menschen ab 65 Jahren (Artmann et al., 2019; Kabisch und Haase, 2014), im Folgenden auch Senioren genannt, 3) und Ausländern (Jünger, 2021), d. h. Personen mit ausschließlich nicht-deutscher Staatsangehörigkeit.

In den Rasterdaten des Zensus werden die Personen anhand der Geo-Koordinate der Wohnanschrift einer bestimmten Rasterzelle zugeordnet. Alle Standortdaten aus dem Zensus 2011 werden zur Gewährleistung der Privatsphäre Perturbationsmaßnahmen unterzogen. Jede Zelle wird durch ihren Mittelpunkt räumlich referenziert, um sie mit der Verfügbarkeit von Grünflächen in der Nachbarschaft zu kombinieren. Für die Analyse des Zusammenhangs zwischen Grünflächen und demographischer Zusammensetzung bei einer Auflösung von $1 \text{ km} \times 1 \text{ km}$ wird die Verfügbarkeit von grüner Landbedeckung und öffentlichem Grün als Durchschnitt aus allen zugrunde liegenden $100 \text{ m} \times 100 \text{ m}$ Zensus-Gitterzellen ermittelt.

4.6 Stadt-Land-Gradient

Es wird erwartet, dass die Verfügbarkeit von Grünflächen in ländlichen und städtischen Regionen unterschiedlich ist, da Grünflächen in dichter besiedelten Gebieten zugunsten anderer Landbedeckungstypen verdrängt sind, z. B. Bebauung. Daher werden alle Maße für die Verfügbarkeit von Grünflächen in der Nachbarschaft in Abhängigkeit von der Einwohnerzahl der Gemeinde analysiert, wozu Klassen entlang der Schwellenwerte von 5.000, 10.000, 50.000 und 100.000 Einwohnern gebildet werden. Basierend auf der Überschneidung der Zentralkoordinate und den administrativen Grenzen Deutschlands wird jede Rasterzelle des Zensus einer Gemeinde zugeordnet.

Zusätzlich zur Beschreibung des Stadt-Land-Gradienten anhand der Gemeindegröße wird eine zusätzliche feingliedrigere Differenzierung herangezogen. Die Klassifizierung des Stadt-Land-Gradienten von Taubenböck et al. (2022) erlaubt die Beziehung zwischen dem städtischen Kontext einer Nachbarschaft und der Verfügbarkeit von Grünflächen noch besser zu verstehen. Diese Einteilung von „ländlich“ bis „städtisch“ basiert auf einer Vielzahl von Abstraktionen des Urbanisierungsgrades. Daraus wird eine wahrscheinkeitsbasierte Metrik abgeleitet, um eine endgültige Klassifizierung des Stadt-Land-Gefälles zu erstellen. So wird auf der Grundlage von Urbanitätsquantilen jeder bewohnte Stadtteil in Deutschland einer von fünf Klassen von „ländlich“ bis „städtisch“ zugeordnet. Dieser Datensatz ist für die vorliegende Studie von Bedeutung, da er eine feinere Unterscheidung zwischen verschiedenen Urbanitätsgraden auf der Ebene von Stadtteilen bzw. Nachbarschaften ermöglicht. Dies erhöht die Interpretierbarkeit durch eine schrittweise Unterscheidung zwischen den Stadtzentren und den Randgebieten.

Methoden

Die methodischen Arbeitsschritte dieser Arbeit orientieren sich an den drei konzeptionellen Einheiten, die in Kapitel 3 vorgestellt wurden. Im Folgenden wird zunächst der methodische Aufbau zur pixelbasierten Landbedeckungsklassifikation erläutert, anschließend die Quantifizierung der Verfügbarkeit von öffentlichen Grünflächen mittels Datenfusion und maschinellem Lernen. Zuletzt wird die methodische Grundlage gelegt für die Analyse der Verteilungsgerechtigkeit von Grünflächen in Deutschland.

5.1 Landbedeckungsklassifikation mit Sentinel-2 und LUCAS

Der methodische Aufbau der Landbedeckungsklassifikation mithilfe von Sentinel-2 und LUCAS-Daten teilt sich in vier Bereiche. Im Folgenden wird zuerst die Modellierung mithilfe eines maschinellen Lernverfahrens skizziert. Anschließend wird erläutert, wie die Klassifikationsgenauigkeit der einzelnen Modelle validiert wird. Darauf folgt die Anwendung der Modellierungsergebnisse in Form einer flächendeckenden Landbedeckungsklassifikation für ganz Deutschland. Mithilfe eines weiteren Bilddatensatzes für das Jahr 2018 wird abschließend die Übertragbarkeit der Methode demonstriert.

5.1.1 Modellierung von Landbedeckung

Entscheidend für die Auswahl des geeigneten maschinellen Lernverfahrens zur pixelbasierten Modellierung der Landbedeckungsinformationen ist die Form, Anzahl und Beschaffenheit der Ausgangsdaten. Im hier vorliegenden Fall stehen flächendeckende Sentinel-2 Bilddaten in 10 Metern räumlicher Auflösung sowie punktuelle Referenzdaten der LUCAS-Erhebung zur Verfügung. Durch räumliche Verschneidung, d. h. der räumlichen Überlagerung von Bild- und Punktdaten und die Extraktion aller Pixelwerte am entsprechenden Referenzpunkt, entsteht ein tabellarischer Datensatz. Dieser beinhaltet einerseits die Zielvariable in Form der Landbedeckungsklasse entsprechend der LUCAS Vorprozessierungsschemata S_0 bis S_3 (siehe Abschnitt 4.1.2) und andererseits die Werte der 229 Bildmerkmale (siehe Abschnitt 4.2).

Ein geeignetes ML-Verfahren zur Klassifizierung derartiger Datensätze ist der sog. (engl.) *Random Forest* Klassifikator (RF, Breiman, 2001). Dieser Klassifikator hat sich bereits in der Vergangenheit bewährt und liefert selbst in hochdimensionalen Merkmalsräumen gute und ro-

buste Ergebnisse (Geiß et al., 2015; Khatami et al., 2016; Pelletier et al., 2016; Wurm et al., 2017a,b). Der Random Forest Algorithmus basiert dabei auf dem Prinzip der Klassifikations- und Regressionsbäume (engl. *classification and regression trees, CART*), einem hierarchischen System von schwellwertbasierten Entscheidungen zur Lösungsfindung (Quinlan, 1986). Die besondere Eigenschaft von Random Forests ist, dass sie aus einem Ensemble von vielen einzelnen CART zusammengesetzt sind, welche im Verfahren des (engl.) *bootstrap aggregating*, besser bekannt unter der Abkürzung *bagging*, entstehen (Breiman, 1996). Hierbei wird für die Erstellung der einzelnen Entscheidungsbäume stets nur eine Teilmenge aller verfügbaren Trainingsbeobachtungen per Zufallsauswahl (mit Zurücklegen) verwendet. Dies verhindert eine zu starke Anpassung an das Trainingsset (engl. *overfitting*, Hastie et al., 2003, S. 587 f.) und erhöht somit die Generalisierungsfähigkeit des gelernten Modells (z. B. Ham et al., 2005). Neben der zufälligen Auswahl von Trainingsbeobachtungen werden im Random Forest Algorithmus auch die beschreibenden Variablen zufällig ausgewählt, d. h. die Spalten des tabellarischen Datensatzes, welche zur Entscheidungsfindung genutzt werden (Prinzie und Poel, 2007). Dies wird auch als (engl.) *attribute bagging* bezeichnet. Während des Lernvorgangs wird so eine Vielzahl von Entscheidungsbäumen auf unterschiedlichen Teilen des Datensatzes trainiert, die in Form eines Ensembles von gleichberechtigten Klassifikationswegen gesammelt werden. Hierdurch entwickelt der Random Forest Ansatz eine hohe Stabilität gegenüber Variationen im Datensatz und führt zu robusten Klassifikationsergebnissen. Zum Zeitpunkt der Inferenz, d. h. in der Anwendung des Modells auf neue, bisher ungesehene Daten, werden alle Beobachtungen (im Fall der Fernerkundung Pixel) anhand aller Entscheidungsbäume einzeln klassifiziert. Die finale Klasse eines Pixels entsteht durch die Mehrheitsentscheidung, sprich welche Klasse pro Pixel am häufigsten als Ergebnis der Entscheidungsbäume prädiziert wird.

Im Vergleich gängiger maschineller Lernverfahren schneidet Random Forest mit höheren Genauigkeiten in einer Vielzahl von Klassifikationsproblemen vor anderen Verfahren ab, wie beispielsweise (engl.) *Support Vector Machines* (Fernández-Delgado et al., 2014). Random Forest Klassifikatoren zeigen hohe Klassifikationsgenauigkeit und niedrige Varianz der Ergebnisse (Sheykhmousa et al., 2020). Darüber hinaus eignen sich Random Forest Modelle gut für die effiziente Verarbeitung hochdimensionaler Datensätze, die gegebenenfalls hohe Multikollinearität aufweisen (Belgiu und Drăguț, 2016). Insbesondere letztere ist in dieser Studie anzunehmen, da Bildtexturen, die in verschiedenen Fenstergrößen gerechnet wurden, ähnlichen Trends folgen könnten.

In dieser Studie werden mehrere Random Forest Modelle mit je 500 Entscheidungsbäumen und einer Anzahl von 15 zufällig ausgewählten Variablen (d. h. Bildmerkmalen) pro Entscheidung trainiert. Um den Einfluss der zugrundeliegenden Randomisierung zu ermitteln, und um sicherzustellen, dass die Ergebnisse unabhängig von der Ziehung der Zufallszahlen akzeptable Ergebnisse liefern, werden je Vorprozessierungsschema ($S_0 - S_3$) und Positionierungsansatz (*GPS*

und *GRID*) 100 Modelle mit variierenden Zufallszahlen trainiert und evaluiert. Die Ziehung der Zufallszahlen wird durch die Verwendung unterschiedlicher Startwerte für Pseudozufallszahlen (engl. *seed*) beeinflusst.

5.1.2 Genauigkeitsabschätzung und Validierung

Komplexere maschinelle Lernverfahren, wie Random Forest, büßen im Gegensatz zu einfacheren Verfahren an Interpretierbarkeit ein. Da die genaue Entscheidungsfindung nur schwer nachzuvollziehen ist, werden sie auch mitunter als (engl.) „*black box*“ bezeichnet. In der Praxis erzielen derartige Verfahren jedoch häufig höhere Genauigkeiten als andere Klassifikationsverfahren (Sheykhmousa et al., 2020). Um die Güte der Entscheidungsfindung dennoch quantifizieren zu können, wird eine Genauigkeitsabschätzung (engl. *accuracy assessment*) durchgeführt (Congalton, 1991). Hierbei wird die tatsächliche Landbedeckungsklasse eines Punktes mit derjenigen verglichen, die der Random Forest Klassifikator dem Punkt anhand der extrahierten Bilddaten zuweist.

Um eine verlässliche Genauigkeitsabschätzung zu erhalten, wird die Gesamtheit der Referenzbeobachtungen in zwei unabhängige, stratifizierte Stichprobengruppen mit einem Verhältnis von 80 % für das Training und 20 % für die Validierung aufgeteilt. Die Genauigkeit des Klassifikationsverfahrens kann bestimmt werden, indem die Validierungspunkte nach Abschluss des Trainingsvorgangs durch das Modell klassifiziert und anschließend die Klassifikationsergebnisse und die tatsächliche Landbedeckungsklasse der Validierungspunkte verglichen werden. Hierzu wird eine Fehlermatrix (engl. *error matrix*) erstellt, welche die Referenzklasse und die prädierte Klasse gegenüberstellt. Basierend auf dieser Fehlermatrix können diverse Genauigkeitsmaße abgeleitet werden, z. B. die Gesamtgenauigkeit (engl. *overall accuracy*, O) oder klassenbasierte Genauigkeitsmaße (Congalton, 1991). Die Gesamtgenauigkeit O beschreibt den Anteil aller richtig klassifizierten Referenzpunkte an der Gesamtheit der Referenzpunkte.

In einem Fall mit n Landbedeckungsklassen entsteht so eine $n \times n$ Fehlermatrix, welche die Anzahl v prädierten Klassen in Zeilen (i) und Referenzklassen in Spalten (j) notiert, aus der sich die Gesamtgenauigkeit O wie folgt errechnet:

$$O = \frac{\sum_{i=j=1}^n v_{ij}}{\sum_{i=1}^n \sum_{j=1}^n v_{ij}} \quad (5.1)$$

Darüber hinaus können die klassenspezifischen Genauigkeiten – Nutzergenauigkeit (engl. *user's accuracy*, U_i) sowie Produzentengenauigkeit (engl. *producer's accuracy*, P_j) – für jede Klasse i bzw. j abgeleitet werden:

$$U_i = \frac{v_{ij}}{\sum_{j=1}^n v_{ij}} \quad (5.2)$$

$$P_j = \frac{p_{ii}}{\sum_{i=1}^n v_{ij}} \quad (5.3)$$

Die Nutzergenauigkeit U_i beschreibt dabei den Anteil der korrekt identifizierten Referenzbeobachtungen einer Klasse in Relation zur gesamten Zahl der Referenzbeobachtungen dieser Klasse. Sie stellt somit ein Maß für Unterklassifikation dar. Die Nutzergenauigkeit ist außerhalb des Fachbereichs der Fernerkundung auch unter (engl.) *positive predictive value* oder (engl.) *precision* bekannt (Altman und Bland, 1994a). Die Produzentengenauigkeit P_j hingegen misst den Anteil an richtig klassifizierten Beobachtungen einer Klasse in Relation zu allen Beobachtungen, die als die jeweilige Klasse prädiziert werden. Sie ist somit ein Maß für Überklassifikation und ist auch als Sensitivität (engl. *sensitivity*) oder (engl.) *recall* bekannt (Altman und Bland, 1994b).

Da sowohl die Teilung der Stichproben in Trainings- und Validierungsbeobachtungen als auch die Generierung der einzelnen Entscheidungsbäume des Random Forest Algorithmus auf Zufallsstichproben beruhen, können die Auswirkungen der Zufallsauswahl nicht ausgeschlossen werden. Unter Berücksichtigung dieser Tatsache werden für jedes Experiment 100 Modelle basierend auf unterschiedlichen Ausgangswerten für (Pseudo-) Zufallszahlen (engl. *seeds*) erstellt. Anhand der Standardabweichung der Gesamtgenauigkeit über die 100 Iterationen je Experiment lässt sich so die Robustheit der Modelle zusammenfassen. Aus insgesamt 4 Vorprozessierungsschemata (S_0 - S_3), 2 Positionierungsvarianten (GPS vs GRID), und 100 Iterationen ergeben sich insgesamt 800 trainierte Modelle. Diese Modelle werden zunächst nicht auf den gesamten Bilddatensatz angewandt, sondern dienen lediglich zum Vergleich der Performanz der Modelle in Bezug auf die Vorprozessierungsschemata sowie die Auswahl des besten Modells für die anschließende Erstellung des landesweiten Klassifikationsproduktes.

5.1.3 Erstellung einer deutschlandweiten Landbedeckungsklassifikation

Auf der Grundlage der Gesamtgenauigkeit wird aus den 800 durchgeführten Klassifikationsexperimenten leistungsstärkste Modell ausgewählt. Dieses wird anschließend zur Klassifizierung der gesamten Bilddaten eingesetzt. Daraus entsteht ein großflächiges Landbedeckungsklassifikationsprodukt für Deutschland mit einer räumlichen Auflösung von 10 m × 10 m. Als letzter Schritt werden die Ergebnisse der Karte mithilfe der sog. marginalen proportionalen Kartengenauigkeitsschätzung (engl. *marginal proportional map accuracy estimation*, Card, 1982; Olofsson et al., 2014; Stehman und Foody, 2019) bewertet. Diese bietet im Vergleich zur klassischen Fehlermatrix den Vorteil, dass die Genauigkeitsbewertung stärker Bezug nimmt auf die Verteilung der Landbedeckungsklassen im finalen Klassifikationsprodukt und sie somit stärker von Einflüssen

des Designs der Stichprobe im Referenzdatensatz entkoppelt ist (Stehman und Foody, 2019). Im Vergleich zur klassischen Genauigkeitsabschätzung wird die Fehlermatrix der stratifizierten Zufallsstichproben dafür mit der Vorkommenswahrscheinlichkeit der einzelnen Klassen (\hat{p}_{ij}) in der endgültigen Karte gewichtet (Olofsson et al., 2013, 2014; Stehman et al., 2018):

$$\hat{p}_{ij} = W_i \frac{v_{ij}}{v_i} \quad (5.4)$$

wobei W_i der Fläche einer Klasse ($A_{m,i}$) in Relation zur gesamten kartierten Fläche (A_{tot}) entspricht:

$$W_i = \frac{A_{m,i}}{A_{tot}} \quad (5.5)$$

Die mithilfe der tatsächlich kartierten Verteilung der Landbedeckungsklassen gewichtete Fehlermatrix wird im Anschluss verwendet, um Schätzungen für die Gesamtgenauigkeit (\hat{O}), Nutzergenauigkeit (\hat{U}) und Produzentengenauigkeit (\hat{P}) zu berechnen (Stehman und Foody, 2019):

$$\hat{O} = \sum_{i=1}^n \hat{p}_{ii} \quad (5.6)$$

$$\hat{U}_i = \frac{\hat{p}_{ii}}{\hat{p}_{i.}} \quad (5.7)$$

$$\hat{P}_i = \frac{\hat{p}_{ii}}{\hat{p}_{.i}} \quad (5.8)$$

Die Standardfehler der geschätzten Genauigkeitsmaße ergeben sich aus der Quadratwurzel der Varianz \hat{V} der Gesamtgenauigkeit und der klassenbasierten Nutzer- und Produzentengenauigkeit:

$$\hat{V}(\hat{O}) = \sum_{i=1}^c W_i^2 \hat{U}_i \frac{1 - \hat{U}_i}{n_{i+} - 1} \quad (5.9)$$

$$\hat{V}(\hat{U}_i) = \hat{U}_i \frac{1 - \hat{U}_i}{n_{i+} - 1} \quad (5.10)$$

$$\hat{V}(\hat{P}_j) = \frac{1}{\hat{N}_{+j}^2} \left[\frac{N_{j+}^2 (1 - \hat{P}_j)^2 \hat{U}_j (1 - \hat{U}_j)}{n_{j+} - 1} + \hat{P}_j^2 \sum_{i \neq j}^c N_{i+}^2 \frac{\frac{n_{ij} (1 - \frac{n_{ij}}{n_{i+}})}{n_{i+} - 1}}{n_{i+} - 1} \right] \quad (5.11)$$

Diese Maße zur Schätzung der Genauigkeit der landesweiten Landbedeckungsklassifikation von Sentinel-2 Daten in Deutschland werden für die Interpretation der Ergebnisse sowie Einschätzung der Stärken und Schwächen des Landbedeckungsprodukts herangezogen.

5.1.4 Übertrag auf 2018

Zur Demonstration der Übertragbarkeit der entwickelten Methodik wird diese auf einen späteren Zeitpunkt angewendet. Zu diesem Zweck wird mithilfe von Sentinel-2 Bilddaten aus dem Jahr 2018 sowie der LUCAS-Daten aus demselben Jahr erneut eine Klassifikation angefertigt. Im Unterschied zu der initialen Entwicklung für den Zeitraum von 2015 bis 2017 werden hier lediglich Bilddaten aus einem Jahr verwendet, was dadurch möglich ist, dass durch den Start von Sentinel-2 B im Jahr 2017 deutlich mehr Bilddaten innerhalb des kürzeren Zeitraums zur Verfügung stehen. Abbildung 5.1 zeigt die Häufigkeit der Abdeckung des Untersuchungsgebietes Deutschland mit Bilddaten von Sentinel-2 A und B im Zeitraum von Mai bis September 2018. Ähnlich wie für den zuvor verwendeten Zeitraum im ersten Klassifikationsaufbau stehen auch für diesen Zeitraum über ganz Deutschland ausreichend wolkenfreie Sentinel-2 Aufnahmen zur Verfügung, aus denen Mosaik abgeleitet werden können. Durch die erneute Ableitung der Landbedeckungsklassifikation lassen sich zeitliche Veränderungen an der Erdoberfläche einfach erkennen.

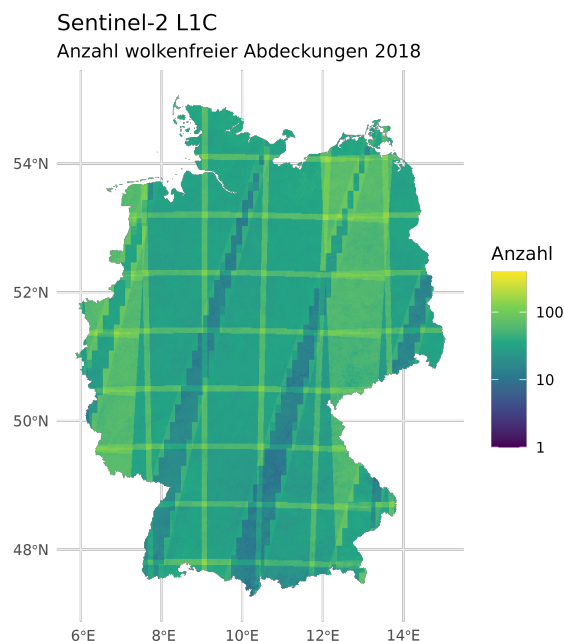


Abbildung 5.1: Abdeckung Deutschlands mit wolkenfreien Sentinel-2 L1C Bilddaten für das Jahr 2018 eingefärbt in logarithmischer Skala. Aus allen verfügbaren Aufnahmen werden lokale Aggregate abgeleitet, die Anzahl der wolkenfreien Aufnahmen pro Pixel variiert dabei von 12 (1. Perzentil, 5. Perzentil = 15) bis zu 117 (99. Perzentil, 95. Perzentil = 82). Rechteckige Artefakte in Gelb zeigen die Überlagerung der einzelnen Bildausschnitte (engl. *granules*), die zur internen Datenhaltung in der Prozessierungsumgebung Google Earth Engine verwendet werden. Die diagonalen Streifen ergeben sich durch die Geometrie der Flugbahn der Sentinel-2 Satelliten.

Für diesen Vergleich werden das Vorprozessierungsschema und der Positionierungsansatz analog zur Erstellung der deutschlandweiten Landbedeckungsklassifikation für 2015/2017 verwendet. Da die Stichprobe für jede der dreijährigen Erhebungen im Rahmen des LUCAS-Surveys variiert

wird, ist es jedoch nicht möglich die identischen Referenzpunkte wie zuvor für die Modellierung zu verwenden. Hierdurch werden möglicherweise Schwankungen in den Modellierungsergebnissen verursacht. Es ist zu erwarten, dass ein Änderungsvergleich zwischen dem Klassifikationsprodukt von 2015/2017 und 2018 nicht ausschließlich durch die tatsächliche Veränderung am Boden beeinflusst wird, sondern auch durch die unterschiedlichen Referenzpunkten.

5.2 Deep Learning Datenfusion zur Erfassung öffentlichen Grüns

Zusätzlich zu Landbedeckungsinformationen soll im zweiten konzeptionellen Schritt dieser Dissertation auch räumliche Informationen zur Landnutzung von Grünflächen im Kontext von Siedlungsgebieten beschrieben werden. Aufbauend auf den Satellitendaten und Landbedeckungsinformationen aus dem vorherigen Abschnitt werden hierfür weitere Daten hinzugezogen, um Grünflächen stärker semantisch zu differenzieren. Das Ziel ist hierbei die Fläche öffentlichen Grüns in einer Nachbarschaft von 500 m Radius um einen beliebigen Punkt zu ermitteln.

Hierfür sollen verschiedenartige Daten kombiniert werden, anhand derer die Zusammenhänge von Landbedeckung und anderen räumlichen Merkmalen mithilfe von maschinellem Lernen modelliert werden. Diese haben das Ziel eine Quantifizierung der Ressource öffentliches Grün auf Nachbarschaftsebene zu ermöglichen. Der methodische Ansatz der hier vorliegenden Arbeit baut dafür auf die Hinzunahme von Geodaten des OpenStreetMap-Projekts, welche sehr detaillierte semantische Informationen über diverse räumliche Objekte bereitstellen. Durch die Kombination aus Satellitendaten und OSM sollen mögliche Datenlücken in einem der beiden Datensätze ausgeglichen werden. Die methodische Herausforderung besteht darin, die pixelbasierten Fernerkundungsdaten in Form von Bildern mit den vektorbasierten Daten aus OpenStreetMap zu kombinieren.

Im Folgenden wird zunächst die Datenaufbereitung vorgestellt, die die Grundlage für die Modellierung zur Quantifizierung von öffentlichem Grün stellt. Darauf folgt die Beschreibung der auf Deep Learning aufbauenden Modellierung von öffentlichen Grünflächenanteilen auf Nachbarschaftsebene. Abschließend wird erläutert, wie die Genauigkeit des entstehenden Datenproduktes bewertet werden kann.

5.2.1 Vorbereitung des multimodalen Datensatzes

Für die Modellierung von öffentlichem Grün werden drei Datensätze kombiniert: Satellitenbilddaten, Aggregate aus OpenStreetMap, sowie European Urban Atlas-Daten (EUA). Letztere dienen dabei als Referenz für die Verfügbarkeit von öffentlichem Grün. Um die Daten für die synergetische Verarbeitung vorzubereiten, werden alle Datensätze auf Ebene der einheitlichen 500 m Nachbarschaft aufbereitet. Ausgehend von einem Radius von 500 m um zentralen Punkt ergibt sich eine Nachbarschaft mit einem Durchmesser von 1000 m. Beim hier genutzten Datensatz rasterbasierter Satellitendaten mit einer Auflösung von 10 m pro Sentinel-2 Pixel wird dies in einen Ausschnitt von 101×101 Pixeln um den Zentrumsunkt, entsprechend 1010 m Kantenlänge

bzw. 102,01 Hektar Gesamtfläche, überführt. Diese werden im Folgenden als „Kacheln“ oder „Nachbarschaft“ bezeichnet. Basierend auf dieser Definition wird ein regelmäßiges Gitter über alle Flächen Deutschlands gelegt, an denen EUA-Daten zur Verfügung stehen (vgl. Abbildung 4.6). Ausschnitte, die nicht vollständig mit EUA-Daten abgedeckt sind, werden entfernt, um fehlerhafte Quantitäten von öffentlichem Grün zu vermeiden. Darüber hinaus wird der Datensatz auf Kacheln beschränkt, welche maximal 1000 m von bebautem Gebiet entfernt liegen, um den Datensatz auf die im Rahmen der Studie relevanten Siedlungsgebiete zu fokussieren. Die Gesamtanzahl aller Kacheln für Deutschland, die diese Kriterien erfüllen, beträgt 115.220.

Für jede dieser Kacheln wird aus dem EUA die Menge an verfügbaren öffentlichen Grünflächen pro Nachbarschaft in Hektar errechnet. Dies stellt die Zielvariable der folgenden Modellierung dar. Abbildung 5.2 zeigt, dass der überwiegende Teil der Kacheln eine geringere Anzahl an öffentlichem Grün aufweist. Dies ist erwartbar, da die Kacheln in bebautem Gebiet liegen.

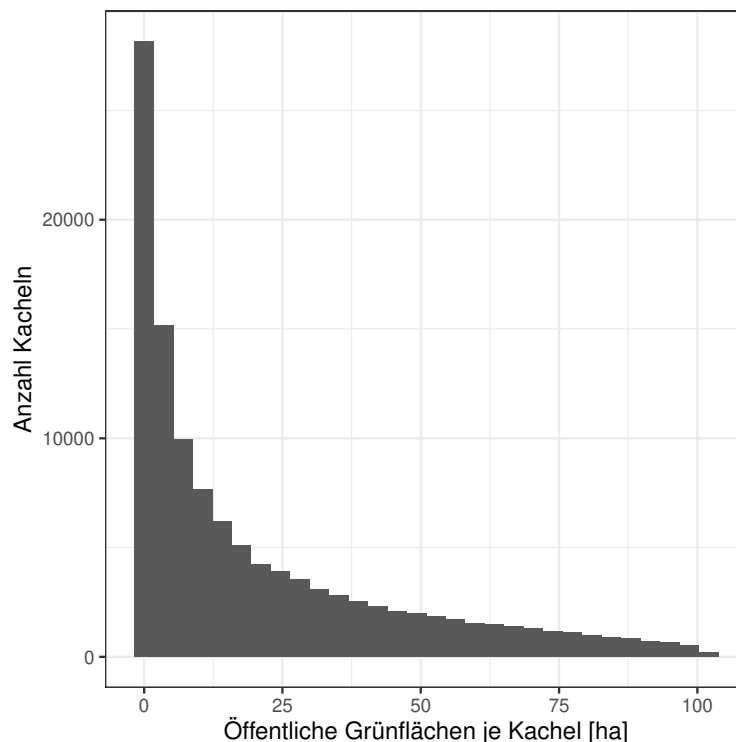


Abbildung 5.2: Häufigkeitsverteilung der Fläche öffentlichen Grüns in den 115.220 Kacheln im Eingangsdatensatz.

Neben der Zielvariable – der verfügbaren Fläche öffentlichen Grüns – wird für jede Kachel ein standardisierter multimodaler Eingangsdatensatz von beschreibenden Variablen erzeugt. Dieser umfasst zwei Teile: multispektrale Bilddaten im Rasterformat auf der einen Seite sowie tabellarische Daten über OpenStreetMap-Elemente in der jeweiligen Zelle auf der anderen Seite. Die multispektralen Bilddaten bauen auf dem deutschlandweiten Sentinel-2 Mosaik von 2018 auf (siehe Abschnitt 4.2). Für jede Kachel beinhalten diese Mosaik Bildinformationen des sichtbaren

Lichtes (Rot, Grün, Blau), des Median-Aggregats des NDVI sowie der Landbedeckungsklassifikation von 2018 (siehe Abschnitt 5.1.4). Für jede Kachel enthält der Bilddatensatz somit 5 Bildebenen mit 101×101 Pixeln.

Zusätzlich zu den Bilddaten werden die OpenStreetMap-Elemente pro Kachel aggregiert. Hierfür werden alle OpenStreetMap-Elemente mit der Ausdehnung der einzelnen Kacheln verschnitten. Aufbauend auf der Gruppierung von OSM-Elementen in insgesamt 71 Gruppen (siehe Abschnitt 4.4) werden relevante Informationen der 3 Geometrieformen (Punkt, Linie, Polygon) aggregiert. Dies umfasst pro Kachel die Anzahl aller Punkte pro Gruppe, die kumulative Länge aller Linienobjekte sowie die kumulative Fläche aller Polygone je Gruppe. Tabelle A2 (siehe Anhang) fasst die möglichen Geometrieformen je Gruppe zusammen. Es entsteht ein Merkmalsvektor mit insgesamt 94 Dimensionen pro Kachel, da manche der 71 Gruppen von mehreren Geometrieformen besetzt sind.

Der beschreibende Datensatz, der als Eingangsinformation für die Abschätzung der Fläche an öffentlichem Grün je Kachel dient, besteht somit aus drei Elementen: 1) Bilddaten und 2) OpenStreetMap-Aggregat als erklärende Variablen sowie 3) der Fläche öffentlicher Grünanlagen in Hektar als abhängige Variable. Für die Verarbeitung in neuronalen Netzen werden alle Bilddatensätze beziehungsweise Merkmals- und Zielvektoren zwischen Minimum und Maximum auf linear einen Zahlenraum von 0 bis 1 skaliert (engl. *min-max normalization*), wie es in bildverarbeitenden Anwendungen üblich ist (Goodfellow et al., 2016, S. 441).

5.2.2 Modellierung mit Deep Learning

Aufbauend auf dem hochdimensionalen, heterogenen Datensatz zur Beschreibung der Landbedeckung und semantischer Informationen zur Landnutzung wird ein Deep Learning Regressionsmodell entwickelt, welches heterogene Daten simultan verarbeiten kann. Das Ziel der Modellierung ist, die Informationen aus beiden verschiedenartigen Eingangsdatensätzen gleichzeitig zu nutzen, um die Fläche an öffentlichem Grün pro Kachel, d. h. in dieser Nachbarschaft, abzuschätzen. Aufgrund ihrer Vielseitigkeit bieten neuronale Netze in diesem Anwendungsszenario besonders einzigartige Voraussetzungen. Einerseits haben bildverarbeitende Netzwerke auf Basis der (engl.) *Convolutional Neural Networks*, (*CNN*) eindrucksvoll demonstriert, wie auch komplexe Bildinformationen im Rahmen von Erdbeobachtungsanalysen erfasst und quantifiziert werden können (LeCun et al., 2015). Darüber hinaus bieten klassischere Ansätze der künstlichen neuronalen Netzwerke (engl. *artificial neural networks*, *ANN*) die Möglichkeit, Merkmale aus tabellarischen Daten zu abstrahieren, um eine Zielvariable beschreiben zu können.

Beiden, CNN und ANN, liegt die gleiche fundamentale Idee zugrunde. Inspiriert von der Funktion eines neuronalen Netzes im menschlichen Gehirn bilden neuronale Netze den Prozess des Lernens nach, indem sie ein Eingangssignal mit einem Ergebnissignal verknüpfen (Goodfellow et al., 2016, S. 13 ff.). Die Verknüpfung geschieht in Form von mathematischen Funktionen und Gewichten, welche im Laufe des Trainings des Netzwerkes angepasst werden können, sodass ein

Eingangssignal zum gewünschten Ergebnis führt. Moderne neuronale Netzwerke – das schließt die hier genannten CNN und ANN ein – besitzen eine Vielzahl von Ebenen zwischen den Eingangs- und Ausgangsebenen, welche miteinander mit vielen Funktionen und Gewichten verknüpft sind. Dies ermöglicht komplexeste Zusammenhänge abzubilden (LeCun et al., 2015). Diese werden daher als tiefe neuronale Netze (engl. *deep neural networks*) bezeichnet. Durch die im Grunde gleiche Funktionsweise ist es daher möglich, die separaten Netzwerke zur Verarbeitung von Bild- und tabellarischen Daten zu einem Netzwerk zu fusionieren.

Diese Dissertation baut daher auf bestehenden Arbeiten zu bildverarbeitenden Netzwerken auf und entwirft ein Fusionsnetzwerk, welches simultan Bilddaten sowie OpenStreetMap-Aggregate verarbeiten kann. Abbildung 5.3 skizziert die Anordnung des Netzwerkes. Die Referenz für die Zielvariable – Fläche von öffentlichem Grün pro Nachbarschaft – wird durch die Daten des European Urban Atlas bereitgestellt.

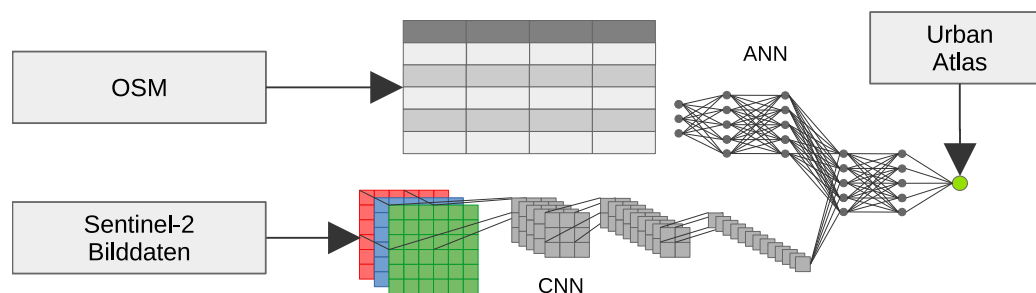


Abbildung 5.3: Skizze des Deep Learning Fusionsnetzwerkes zur Ableitung öffentlicher Grünflächenanteile auf Nachbarschaftsebene. Der grüne Punkt rechts markiert die Zielvariable des Netzwerkes, welche im Fall der Inferenz dem Ausgabewert entspricht.

Die Bilddaten dienen als Input für ein CNN aus welchem höherwertige Informationen über die Verfügbarkeit von Grünflächen bzw. öffentlichem Grün erlernt werden. Im Rahmen dieser Dissertation wird auf das (engl.) *Multilevel Feature Fusion*-basierte Netzwerk aufgebaut (*Feature Fusion* ist hier zu unterscheiden vom Fusionsansatz von ANN und CNN), welches von Qiu et al. (2020b) speziell für die szenenbasierte Klassifikation von Landbedeckung, im speziellen *Local Climate Zones*, entwickelt wurde. Es besteht aus vier Blöcken aus einer Kombination von Faltungsebenen (engl. *convolution layers*) gefolgt von Batch-Normalisierung und Poolingebenen, mit der Besonderheit, dass die einzelnen Blöcke mit einer letzten Poolingebene verbunden sind (sog. engl. *skip connections*, für Details siehe Qiu et al., 2020b). Dieses Netzwerk zeichnet sich im Gegensatz zu etablierten, nicht fernerkundungsspezifischen Netzwerken, wie beispielsweise *MobileNet-V2* (Sandler et al., 2018), *VGG16* (Simonyan und Zisserman, 2014), *ResNet50* (He et al., 2016) oder *ResNext* (Xie et al., 2016) durch eine deutlich niedrigere Komplexität und geringere Anzahl an Parametern aus. Trotz der reduzierten Komplexität erzielte dieses Modell dennoch ähnliche oder bessere Ergebnisse für die Klassifikation von Erdbeobachtungsdaten (Qiu et al., 2020b). Die geringe Komplexität stellt im hier angestrebten Vorhaben einen entscheidenden Vorteil dar, da durch die Fusion mit weiteren Datensätzen zusätzliche Komplexität entsteht. Ein

weniger komplexes CNN verringert unnötige Gesamtkomplexität des Fusionsmodells. Dies wirkt sich positiv auf die Rechenzeit aus und verringert die Tendenz des Netzwerks zu Überanpassung (engl. *overfitting* Géron, 2019, S. 457).

Die OpenStreetMap-Daten werden in ein ANN gespeist. Der Unterschied dieses Netzwerkes zum CNN liegt darin, dass anstatt von Faltungsebenen sogenannte vollverbundene Ebenen (engl. *fully connected layer*, oder *densely connected layer*) verwendet werden. Durch die zunehmende Verringerung der Knoten der inneren Ebenen zwischen dem Eingangsebenen und der letzten Ebene, wird eine Abstraktion der hochdimensionalen Eingangsdaten hin zu der singulären kontinuierlichen Zielvariable, öffentliches Grün, erreicht. Das Netzwerk weist entsprechend der Anzahl von OSM-Aggregaten 94 Eingangsknoten auf gefolgt von versteckten Ebenen mit 20, 10 und 5 Knoten. Alle Knoten sind gefolgt von Batch-Normalisierung und aktiviert mit der (engl.) *Rectified Linear Unit* (Fukushima, 1969; Goodfellow et al., 2016, S. 187, ReLU). Die Batch-Normalisierung (Ioffe und Szegedy, 2015) dient zur Vermeidung von Extremwerten innerhalb des Netzwerkes und ist deshalb ein effektives Mittel Optimierung des Netzwerkes (Goodfellow et al., 2016, S. 309 ff.).

Das Fusionsnetz erweitert das ANN und das CNN, indem es deren letzte vollverbundene Ebenen zu einer gemeinsamen FusionsEbene verkettet. Darauf folgen drei vollverbundene Ebenen mit 20, 15, 10 und 5 ReLU-aktivierten Knoten. Während des Lernprozesses wird der Huber Verlust (Huber, 1964) genutzt, um die Regressionsleistung des Modells ermitteln. In der rückwärtsgewandten Optimierung wird der Verlust dazu genutzt das Modell zu optimieren, d. h. die verfügbare Fläche von öffentlichem Grün in einer Nachbarschaft zunehmend besser zu quantifizieren. Der Huber Verlust setzt sich aus zwei Komponenten zusammen: aus einer linearen Komponente bei größeren Fehlerraten und einer quadratischen Komponente bei niedrigeren Fehlerraten. Die lineare Komponente reduziert die Anfälligkeit für Ausreißer und die quadratische Komponente hilft der schnelleren und besseren Konvergenz zum optimalen Ergebnis im Vergleich zu anderen Verlustmaßen, wie beispielsweise dem mittleren absoluten Fehler (Géron, 2019, S. 293). Die Architekturen der Netzwerke sind im Anhang in Codeblock A1 abgedruckt.

Wie für eine Regression mit neuronalen Netzwerken empfohlen, weisen alle Modelle eine lineare Aktivierungsfunktion (manchmal auch als „ohne Aktivierung“ bezeichnet) in der letzten Ebene auf (Géron, 2019, S. 293). Vor der finalen Ebene wird eine (engl.) *Dropout*-Ebene eingezogen. Diese Ebene blockiert während dem Trainingsvorgang des Netzwerkes zufällig einzelne Verbindungen, um Überanpassung zu verringern (Srivastava et al., 2014; Bejani und Ghatee, 2021) und die Generalisierungsfähigkeit des Netzwerkes zu verbessern. Die Hyperparameter, die für das Training der Netzwerke verwendet werden sind in Tabelle 5.1 aufgeführt und werden im weiteren Verlauf näher erläutert.

Vor dem eigentlichen Trainingsvorgang werden die mit zufälligen Gewichten und Parametern initialisierten Netzwerke mit einer extrem niedrigen Lernrate, der sog. Aufwärmlehnrate (engl. *warmup learning rate*) trainiert, was dabei hilft, die Knoten des Netzwerks entlang des Verlustgradienten auszurichten, entlang dessen das Modell optimiert wird. Der eigentliche Trainingsvorgang

Tabelle 5.1: Hyperparameter für neuronale Regressionsnetzwerke zur Modellierung von öffentlichem Grün.

Hyperparameter	Wert
Batch Size	512
Dropout	0,1
Learning Rate	1×10^{-4}
Reduce Learning Rate on Plateau Patience	20 Epochen, Faktor 0,2
Early Stopping	40 Epochen
Warmup Epochen	3
Warmup Learning Rate	1×10^{-15}

erfolgt zunächst mit einer einheitlichen Lernrate von anfänglich 1×10^{-4} . Nach Abschluss jeder Epoche, d. h. einem vollständigen Durchlauf aller zur Verfügung stehenden Trainingsdaten, werden die Validierungsdaten genutzt, um die zwischenzeitige Genauigkeit des Modells zu ermitteln. Sollte die Genauigkeit für 20 Epochen nicht gesunken sein, wird die Lernrate im Faktor 0,2 reduziert. Im Trainingsvorgang werden zunächst die beiden Teil-Netzwerke CNN und ANN separat trainiert. Nachdem die Teil-Netzwerke getrennt optimiert sind, werden anschließend deren vortrainierte Parameter für das Training des Fusionsnetzwerkes verwendet. Empirisch zeigte sich, dass daraufhin das Training des Fusionsnetzwerkes schneller und besser zur optimalen Lösung konvergierte.

Der Trainingsvorgang endet, nachdem 40 Epochen lang keine weitere Verbesserung des Netzwerkes verzeichnet werden kann. Das Ergebnis des Trainingsvorgangs ist das Modell, welches nach Evaluierung der Validierungsdaten den niedrigsten Wert des Huber Verlusts aufweist. Der beste Stand des Fusionsnetzwerkes wird im Folgenden verwendet, um die verfügbare Fläche von öffentlichem Grün auf Nachbarschaftsebene zu ermitteln.

5.2.3 Genauigkeitsabschätzung

Das Training, die Validierung sowie die Evaluation der neuronalen Netzwerke beruht auf den insgesamt 115.220 Kacheln bzw. Nachbarschaften in ganz Deutschland. Um die Unabhängigkeit zwischen den Trainings- und Evaluierungsdaten des Modells zu gewährleisten, werden alle Kacheln in separate Trainings-, Validierungs- und Testdatensätze aufgeteilt, die respektive 70 %, 15 % und 15 % der Daten enthalten. Nach Abschluss des Trainings wird das beste Modell auf alle Kacheln des Testdatensatzes angewandt, um die Modellgüte bestimmen zu können. Diese wird aus der Differenz zwischen Referenzdaten und Prädiktion bestimmt. Hierfür stehen verschiedene Metriken zur Verfügung.

Einerseits wird die Genauigkeit des Regressionsansatzes mithilfe des empirischen Bestimmtheitsmaßes R^2 (engl. *coefficient of determination*) ausgewertet (Bahrenberg et al., 2010, S. 198 ff.). Dieses misst das Verhältnis der Variation der Modellergebnisse (systematische Variation) zur Gesamtvariation der Referenzdaten.

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (5.12)$$

wobei y_i der Referenzwert der Beobachtung i aus der Gesamtheit der Beobachtungen n darstellt, \hat{y}_i den Schätzwert des Modells und \bar{y} den mittleren Referenzwert der Zielvariable. Das Bestimmtheitsmaß misst die Güte einer Modellschätzung in Form einer einheitslosen Zahl zwischen 0 und 1, wobei höhere Werte eine gute Übereinstimmung zwischen dem Referenzwert und der Modellschätzung anzeigen.

Zusätzlich zum Bestimmtheitsmaß ist es üblich die Magnitude des Fehlers zu bestimmen. Hierzu eignet sich der mittlere absolute Fehler (engl. *mean absolute error, MAE*), der die durchschnittliche Abweichung zwischen Referenz und Modellergebnis in der Einheit der Zielvariable abbildet.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (5.13)$$

Ähnlich zum *MAE* misst die Wurzel aus dem mittleren quadratischen Fehler (engl. *root mean squared error, RMSE*) die Abweichung zwischen Prädiktion und Referenz in der Einheit der Zielvariable. Im Gegensatz zum *MAE* ist der *RMSE* sensitiver gegenüber starken Über- bzw. Unterschätzungen, sogenannten Ausreißern (engl. *outlier*). Der *RMSE* berechnet sich wie folgt:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (5.14)$$

Mithilfe dieser Metriken werden Ergebnisse der Prädiktion von Grünflächenanteilen auf Nachbarschaftsebene evaluiert, um die Effektivität der Methode evaluieren zu können.

5.3 Analyse zur Verteilung von Grün in Deutschland

Zur Untersuchung der Verteilung von Grün in der Wohnumgebung der Bevölkerung Deutschlands werden die Ergebnisse aus den vorherigen methodischen Schritten kombiniert (siehe auch Weigand et al., 2023b). Die Ausstattung der Nachbarschaft mit Grünflächen wird anhand von zwei verschiedenen Metriken analysiert: die Verfügbarkeit von grüner Landbedeckung (engl. *green land*

cover, GLC) und von öffentlichen Grünflächen (engl. *public green spaces*, PGS). Nachfolgend wird zunächst die Kombination dieser Datensätze mit Bevölkerungsdaten beschrieben, die diese Analysen ermöglicht. Anschließend werden die methodischen Grundlagen für die deskriptive Statistik zur Beschreibung der Verteilung von Grünflächen für die Bevölkerung Deutschlands, die Analyse der Verteilungsgerechtigkeit sowie die Analyse des Zusammenhangs zwischen Demographie und Grünflächenverfügbarkeit auf Stadtteilebene dargelegt. Der zentrale Beitrag dieser Studie ist, dass eine derartige Analyse von hochaufgelösten Umweltdaten erstmals auf nationaler Ebene durchgeführt wird.

5.3.1 Datenvorbereitung

Den konzeptionellen Ausgangspunkt für die Datenaufbereitung dieser Analyse stellt der Zensus-Datensatz (siehe Abschnitt 4.5) dar. Für die Zentrumsunkte aller ca. 3,1 Mio. räumlichen $100\text{ m} \times 100\text{ m}$ Gitterzellen der Zensusdaten wird zunächst die rechteckige Nachbarschaft um einen Radius von 500 m gezogen. Mithilfe dieser Nachbarschaften werden für jeden Zensuspunkt die verfügbaren Grünflächen individuell lokal erfasst und zugeordnet. Hierbei wird zwischen grüner Landbedeckung (GLC) und öffentlichem Grün (PGS) unterschieden.

Zur Quantifizierung der Gesamtfläche grüner Landbedeckung je Nachbarschaft wird die Landbedeckungsklassifikation für das Jahr 2018 (siehe Abschnitt 5.1.4) genutzt. Es wird die Gesamtfläche von Baumbewuchs und Grünland aufsummiert. Namentlich werden die Klassen *niedrige mehrjährige Vegetation*, *hohe mehrjährige Vegetation* und *hohe saisonale Vegetation* verwendet. Die Menge verfügbarer öffentlicher Grünflächen wird mithilfe des zuvor erstellten Modells zur Quantifizierung öffentlichen Grüns auf Nachbarschaftsebene (siehe Abschnitt 5.2) ermittelt. Hierfür werden die Eingangsdaten der Nachbarschaft jeder einzelnen Zensus-Gitterzelle extrahiert. Mithilfe des Fusionsmodells wird daraus letztlich der Anteil öffentlichen Grüns prädiziert. So steht schließlich die Menge grüner Landbedeckung und öffentlicher Grünflächen für die Nachbarschaft um jeden Zensus-Gitterpunkt zur Verfügung.

Aus den soziodemographischen Informationen des Zensus in Kombination mit den Grünflächendaten ergibt sich ein tabellarischer Datensatz, welcher im Folgenden Eingang findet in die statistische Untersuchung der Verteilung von Grünflächen in Deutschland. Die direkte Gegenüberstellung der unterschiedlichen Erfassung von Grünflächen (GLC und PGS) ermöglicht dabei einen Vergleich der beiden Werte.

Für die Analyse des Zusammenhangs zwischen der demographischen Zusammensetzung der Bevölkerung und den verfügbaren Grünflächen liegen Zensusdaten lediglich für Gitterzellen mit einer räumlichen Auflösung von $1\text{ km} \times 1\text{ km}$ vor. Die verfügbare Menge von Grünflächen für diese Gitterzellen wird anhand räumlicher Aggregation des Mittelwertes aller darunterliegenden $100\text{ m} \times 100\text{ m}$ Gitterzellen ermittelt.

5.3.2 Deskriptive Statistiken

Als erster Aspekt der Grünflächenverteilung wird die Verfügbarkeit von Grünflächen in der Wohnumgebung der Bevölkerung in Deutschland mithilfe deskriptiver Statistiken ermittelt. Als Ausgangsbasis für die Analyse der sozialen Gerechtigkeit der Grünflächenverfügbarkeit wird verglichen, welcher Anteil der Bevölkerung in Deutschland Zugang zu ausreichenden Mengen an öffentlichem Grün hat. Basierend auf den Zielvorgaben der World Health Organization (2017) sollten die Einwohner mindestens 1,8 ha bis 3,6 ha öffentliches Grün innerhalb der hier verwendeten 500 m Nachbarschaft vorfinden, welche die räumliche Grundlage dieser Studie darstellt. Dabei wird davon ausgegangen, dass Gerechtigkeit bei der Versorgung mit öffentlichem Grün in Deutschland gegeben ist, wenn die gesamte Bevölkerung 3,6 ha oder mehr öffentliches Grün zur Verfügung in der Wohnumgebung vorfindet.

Zusätzlich zu öffentlichem Grün wird die Verfügbarkeit von grüner Landbedeckung in allen Wohngebieten in Deutschland quantifiziert. Da grüne Landbedeckung eine Obermenge von Grünflächen ist, d. h. sie erfasst neben öffentlichen Grünflächen viele weitere Grünflächentypen, werden andere Schwellenwerte festgelegt. Die Nachbarschaften werden anhand der bevölkerungsgewichteten Verteilung in Deutschland in Terzile nach GLC (niedrig, mittel, hoch) eingeteilt. Daraus ergeben sich Schwellenwerte für GLC von 25,4 ha und 44,2 ha.

Um die Komplexität zu verringern und die Lesbarkeit zu verbessern, wird ein zweidimensionales Klassifizierungsschema unter Verwendung der GLC- und PGS-Schwellenwerte gewählt (siehe Abbildung 5.4). Dies geschieht, um die Zusammensetzung der Grünflächen auf Nachbarschaftsebene durch das gemeinsame Auftreten von GLC- und PGS-Werten besser zu verstehen. Das Klassifizierungsschema kommt zum Einsatz, um die kumulativen Anteile der Bevölkerung für

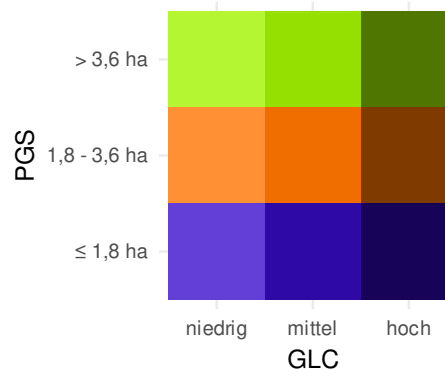


Abbildung 5.4: Zweidimensionales Visualisierungsschema zur Einteilung der Grünflächenverfügbarkeit auf Nachbarschaftsebene. Dieses Schema dient zur visuellen Interpretation der Ergebnisgrafiken der Analyse. Entlang der Horizontalen werden zunehmende Anteile grüner Landbedeckung (GLC) eingeteilt. Diese sind durch zunehmende Farbtiefen verdeutlicht. Entlang der Vertikalen werden zunehmende Anteile öffentlicher Grünflächen (PGS) angetragen, erkennbar durch verschiedene Farben.

jede Klasse der Verfügbarkeit von Grünflächen zu berechnen. Dies ermöglicht die Identifizierung von Disparitäten in der Bevölkerung und die Quantifizierung des Anteils der Bevölkerung mit einer unter dem Zielwert liegenden PGS-Verfügbarkeit. Dabei wird die räumliche Beschaffenheit des Zensus-Gitters genutzt, um verschiedene Gruppierungsstrategien anzuwenden: 1) gesamtes Deutschland, 2) nach Gemeindegröße, 3) feingranulares Land-Stadt-Gefälle und 4) vorherrschender Wohnungstyp.

5.3.3 Verteilungsgerechtigkeit von Grünflächen

Über die reine Beschreibung der Grünflächenverteilung hinaus ist es möglich die Verteilungsgerechtigkeit innerhalb der Bevölkerung oder einzelner Untergruppen zu bestimmen. Hierfür stehen diverse Gerechtigkeitsmaße zur Verfügung, welche die Verteilungsungleichheit einer Variable in einem Index aggregieren (McGregor et al., 2019). Visuell kann die Gerechtigkeit einer Verteilung mithilfe der Lorenzkurve dargestellt werden (siehe Abbildung 5.5). Sie stellt die Ungleichverteilung einer Ressource in einer Gesellschaft dar durch die Gegenüberstellung des Anteils der Bevölkerung und des relativen Anteils dessen an der Gesamtsumme der Ressource. Im Kontext dieser Dissertation und im Kontext der Umweltgerechtigkeit stellt die Ressource beispielsweise die Menge von Grünflächen dar.

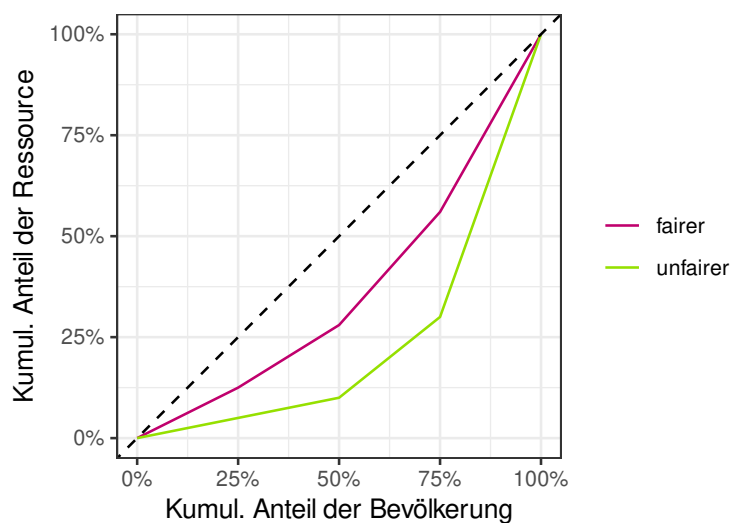


Abbildung 5.5: Beispiel für eine Lorenzkurve zeigt zwei Verteilungen. Die Abbildung zeigt, eine fairere Verteilung (pink) der Ressource und eine unfairere (grün). Eine gerechte Verteilung der Ressource wäre erreicht, wenn die Linie entlang der diagonale (gestrichelte Linie) zu liegen käme.

Zur Bewertung der sozialen Gerechtigkeit der räumlichen Grünflächenverteilung in einer Maßzahl wird häufig der sog. Gini-Koeffizient eingesetzt (z. B. Kabisch und Haase, 2014; Xu et al., 2018b). Dieser Koeffizient ist ein Maß für die Ungleichheit der Verteilung oder Konzentration. Er kann verwendet werden, um die Verteilungsgerechtigkeit einer Ressource innerhalb einer Bevölkerung zu quantifizieren. In dieser Analyse wird der Gini-Koeffizient verwendet, um die

Verteilungsgerechtigkeit der Menge von grüner Landbedeckung und öffentlicher Grünflächen in der Wohnumgebung der Bevölkerung Deutschlands zu bewerten. Der Gini-Koeffizient G ist definiert als

$$G = \frac{\sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|}{2n \sum_{j=1}^n x_j} \quad (5.15)$$

wobei x_i die verfügbare Grünfläche für ein Individuum i der Bevölkerung von n Personen ist. Der Gini-Koeffizient nimmt Werte zwischen 0 und 1 an. Während 0 auf eine völlig gleichmäßige Verteilung der Ressource hindeutet, zeigt 1 die größtmögliche Ungleichverteilung an, d. h. die Anhäufung der verfügbaren Ressource für nur ein Individuum. Zusammen mit dem Gini-Koeffizienten ermöglicht die Lorenzkurve die Visualisierung der Ungleichheit entlang der beiden kumulativen Dimensionen der Begünstigten und der Ressource.

Trotz der weitverbreiteten Nutzung des Gini-Koeffizienten, auch im Rahmen der Umweltgerechtigkeitsforschung, wird dieser bisweilen kritisiert, da zwei unterschiedliche Verteilungen gleiche Gini Werte hervorbringen können (Liu und Gastwirth, 2020). Darüber hinaus impliziert der Gini-Koeffizient eine Wertung darüber welche Teile der Verteilung am wichtigsten sind (Levy et al., 2006). Deshalb wird empfohlen zusätzlich zum Gini-Koeffizienten weitere Maße der Verteilungsgerechtigkeit einzusetzen.

Das Atkinson Ungleichverteilungsmaß (Atkinson, 1970) beispielsweise adressiert einige Kritikpunkte des Gini-Koeffizienten. Dieser Index wird von Marshall et al. (2014) und Sheriff und Maguire (2020) für Umweltgerechtigkeitsanalysen verwendet und ist besonders geeignet für die Analyse von gesundheitlicher Ungleichheit (Levy et al., 2006).

$$A(\epsilon) = 1 - \frac{1}{\bar{y}} \left(\sum_{i=1}^N y_i^{1-\epsilon} \right)^{\frac{1}{1-\epsilon}} \quad \text{für } \epsilon \neq 1 \quad (5.16)$$

$$A(\epsilon) = 1 - \frac{1}{\bar{y}} \left(\prod_{i=1}^N y_i \right)^{\frac{1}{N}} \quad \text{für } \epsilon = 1 \quad (5.17)$$

wobei der Faktor ϵ es ermöglicht, eine Gewichtung zwischen unteren und oberen Bereichen der Verteilung vorzunehmen. Je höher ϵ , desto stärker werden die unteren Bereiche der Verteilung gewichtet (Atkinson, 1970; Cowell, 2011, S. 52; McGregor et al., 2019). Für ϵ werden typischerweise Werte zwischen 0,2 und 2 verwendet (Levy et al., 2006). In dieser Untersuchung werden in Anlehnung an andere Studien die Werte $\epsilon = 0,5$ und $\epsilon = 0,75$ parallel betrachtet

(Kawachi und Kennedy, 1997; Marshall et al., 2014; De Nicolò et al., 2022, 2023). Dies ermöglicht eine differenzierte Betrachtung der Verteilungsgerechtigkeit zusätzlich zum Gini-Koeffizienten. Durch den Unterschied der beiden Werte ϵ können Ungerechtigkeiten für Menschen mit weniger Grünflächen stärker betont werden.

Ähnlich wie bei der deskriptiven Statistik werden verschiedene Gruppierungsstrategien angewandt, d. h. 1) gesamtes Deutschland, 2) nach Gemeindegröße, 3) nach dem Land-Stadt-Gefälle und 4) nach dem vorherrschenden Wohnungstyp, um die Verteilung der Grünflächenressourcen in jeder Bevölkerung zu analysieren.

5.3.4 Grünflächenverfügbarkeit nach Bevölkerungsgruppen

Zuletzt wird der Zusammenhang zwischen der demographischen Zusammensetzungen von Siedlungsgebieten und die Verfügbarkeit von Grünflächen in der Nachbarschaft untersucht. Damit soll festgestellt werden, ob eine der ausgewählten Bevölkerungsgruppen systematisch Nachteile oder Vorteile in Bezug auf Grünflächen erfährt, und wie sich diese zwischen GLC und PGS unterscheiden. Hierfür werden die Informationen über die demographische Zusammensetzung auf der Skala von $1 \text{ km} \times 1 \text{ km}$ Zensus-Gitterzellen (vgl. Abschnitt 4.5) verwendet.

Mithilfe einer Regressionsanalyse wird die Beziehung zwischen der Grünflächenverfügbarkeit je Gitterzelle als Zielgröße entlang einer kontinuierlichen Achse (in ha) in Abhängigkeit zur Bevölkerungsdichte und der demographischen Zusammensetzung modelliert. Zugrunde liegt die Annahme, dass der systematische Zusammenhang zwischen der abhängigen und den unabhängigen Variablen mithilfe der Regressionsterme quantifiziert werden kann.

Die Erstellung eines einzelnen Regressionsmodells über das gesamte Untersuchungsgebiet Deutschland würde jegliche lokalen Besonderheiten maskieren und aggregieren. Es ist jedoch davon auszugehen, dass lokale Einflüsse unterschiedlicher historischer, geographischer und sozio-demographischer Gegebenheiten eine starke Rolle für den Zusammenhang zwischen Grünflächen und demographischer Zusammensetzung spielen. Um die zugrundeliegenden allgemeinen Trends zu ermitteln, werden lineare Regressionsmodelle mit gemischten Effekten (engl. *linear mixed effects models*, *LMM*) verwendet. In diesen Modellen wird die Beziehung zwischen der Menge der verfügbaren Grünflächen (GLC/PGS) und der multivariaten demographischen Zusammensetzung des Stadtteils modelliert. Um lokale Einflüsse auf die zugrundeliegenden Beziehungen zu berücksichtigen, werden in die LMM zufällige Effekte (engl. *random effects*, *RE*) auf verschiedenen Ebenen (der Ebene der Gemeinde, der Ebene der Stadtgröße und des Stadt-Land-Gradienten) einbezogen. Die Regressionsterme der Grünflächenvariablen, d. h. die Steigungskoeffizienten der Regressionsgleichung, werden als fixierte Effekte (engl. *fixed effects*, *FE*) geschätzt, welche allgemeine Trends der Beziehung zwischen GS und demographischer Zusammensetzung in Deutschland beschreiben. Alle LMMs werden für die Bevölkerungszahl (im Logarithmus) kontrolliert, da die

Bevölkerungsdichte als wichtiger Proxy für die Bebauungsdichte angesehen wird und somit einen inhärenten Einfluss auf den möglichen Umfang der verfügbaren Grünflächen hat. Das Modell wird formalisiert als

$$y_{g,h,i} = \beta_0 + \beta_{pop} \log(pop) + \beta_1 X_1 + \dots + \beta_n X_n + u_g + u_h + u_i + \epsilon \quad (5.18)$$

wobei $y_{g,h,i}$ die Zielvariable ist, d. h. die Menge an Grünfläche (entweder GLC oder PGS) in Hektar in Relation zu den Zufallseffekten u_g , u_h und u_i , entsprechend der Zufallseffekte auf Ebene der Gemeinde, der Stadtgröße und des Stadt-Land-Gradienten. $\beta_{1\dots n}$ sind die fixierten Effekte der n unabhängigen Variablen $X_{1\dots n}$ und ϵ ist ein Vektor der zufälliger Fehlerterme.

Als Orientierung zur Auswirkung der lokalen Effekte auf die Gesamtschätzung der LMM werden darüber hinaus klassische multivariate lineare Regressionsmodelle ohne fixierte und zufällige Effekte geschätzt. Diese haben die Form:

$$y = \beta_0 + \beta_{pop} \log(pop) + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \quad (5.19)$$

wobei y die Zielvariable ist, deren Wert mithilfe der linearen Steigungskoeffizienten β_i und der Bevölkerungsanteile X_i geschätzt wird. X_1 ist der Anteil von Ausländern, X_2 der Anteil von Kindern und X_3 der Anteil von Senioren.

Es werden getrennte Modelle für GLC und PGS erstellt, um mögliche Unterschiede und Ähnlichkeiten zwischen den beiden Grünflächenmetriken zu ermitteln. In verschiedenen Modellen werden zudem abwechselnd Variablen mit zufälligen Effekten und unabhängige Variablen verwendet, um festzustellen, ob eine Kombination von Variablen zusammenwirkende oder unterschiedliche Wirkungen hervorruft.

Ergebnisse und Interpretation

Im Folgenden werden die Ergebnisse der drei konzeptionellen Schritte gezeigt. Die raumquantitativen oder statistischen Ergebnisse werden dabei begleitet von einer initialen Einordnung.

6.1 Flächendeckende Landbedeckungsklassifikation mit Sentinel-2 und LUCAS

Die flächendeckende Landbedeckungsklassifikation für Deutschland aus Sentinel-2 Daten wird im Folgenden in drei Einzelschritten betrachtet. Zunächst werden verschiedene Vorprozessierungsschemata sowie Positionierungsansätze von LUCAS-Daten hinsichtlich ihrer Gesamtgenauigkeit evaluiert. Anschließend wird ein Modell auf das ganze Untersuchungsgebiet der Landesfläche Deutschlands angewandt und eine detaillierte Genauigkeitsabschätzung durchgeführt. Schließlich wird der Übertrag der Methode, welche basierend auf Satellitendaten aus dem Zeitraum von 2015 bis 2017 entwickelt, wurde auf Daten des Jahres 2018 angewandt.

6.1.1 Evaluation der Stichprobenauswahl und Gesamtgenauigkeit

Zu Beginn der Evaluation der Landbedeckungsklassifikation aus Sentinel-2 und LUCAS steht der Vergleich zwischen verschiedenen Vorprozessierungsansätzen von LUCAS-Daten. Die Ergebnisse der Gesamtgenauigkeit der Modelle geben Aufschluss über die Auswirkungen der verschiedenen Vorverarbeitungsschemata ($S_0 - S_3$), die auf die LUCAS-Stichproben und die Positionierung der Stichprobenpunkte (*GPS* und *GRID*) angewendet wurden (vgl. auch Weigand et al., 2020). Alle Ergebnisse sind in Abbildung 6.1 zusammengefasst. Das Balkendiagramm stellt die durchschnittliche Gesamtgenauigkeit für die Experimente dar und ist nach Vorprozessierungsschema und Positionierungsansatz gruppiert. Die Fehlerspannen, die durch die Erstellung von 100 Modellen mit wechselnden Seeds für Pseudozufallszahlen entstehen, werden durch die Fehlerbalken angezeigt.

Beim Blick auf das Balkendiagramm zeigt sich ein deutlicher Unterschied zwischen den präsentierten Vorprozessierungsschemata der LUCAS-Beispieldaten. Die Gesamtgenauigkeiten für alle Experimente reichen von 78,1 % bis 93,6 %. Es ist festzustellen, dass kein direkter Zusammenhang zwischen den Ergebnissen der Gesamtgenauigkeit und der Anzahl der Klassen besteht. Schemata S_1 (9 Klassen) und S_3 (7 Klassen) erzielten höhere Gesamtgenauigkeiten als

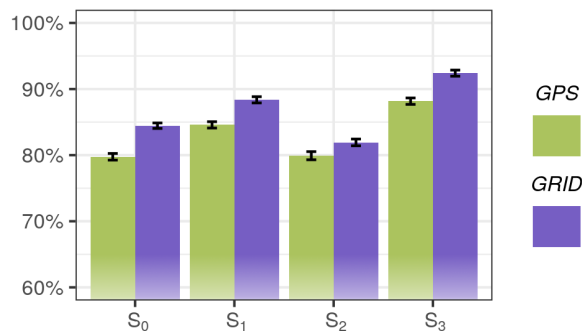


Abbildung 6.1: Gesamtklassifikationsgenauigkeit für alle Experimente (verändert nach Weigand et al., 2020). Die Farben zeigen die verschiedenen Positionierungsansätze an, die einzelnen Verfahren zur Vorverarbeitung der Proben, S_x , sind entlang der x-Achse ausgerichtet. Fehlerbalken zeigen die Standardabweichung vom Mittelwert der 100 Experimente mit unterschiedlichen Zufallszahlen an.

Schemata S_0 (8 Klassen) und S_2 (13 Klassen). Dennoch wird die höchste Genauigkeit von dem Schema mit der geringsten Klassenanzahl erreicht (S_3). Die Fehlerspanne für alle Experimente beträgt im Durchschnitt 0,57 % (max. 0,76 %), was auf robuste Modellierungsergebnisse hinweist.

Vergleicht man die Positionierungsansätze *GPS* und *GRID*, so zeigt sich, dass letzterer bei allen Experimenten zu einer durchgängig höheren Gesamtgenauigkeit führt. Die Steigerungen der Gesamtgenauigkeit variieren je nach Vorprozessierungsschema von 2,0 % für S_2 bis 4,7 % für S_0 . Eine Varianzanalyse, in der die Gesamtgenauigkeiten der einzelnen Modelldurchläufe je Positionierungsansatz verglichen werden, sowie ein daran anschließender Tukey-Duckworth-Test quantifizieren den Unterschied der Gesamtgenauigkeit zwischen den Positionierungsansätzen auf 3,7 %.

6.1.2 Großflächige Anwendung der Landbedeckungsklassifikation

Nach der Bewertung der Gesamtgenauigkeit aller 800 Experimente wird eine deutschlandweite pixelbasierte Klassifikation durchgeführt. Basierend auf den vorherigen Ergebnissen wird das beste Modell ausgewählt. Dabei handelt es sich um das Vorprozessierungsschema S_3 und die *GRID*-Standorte. Dieses Modell wird auf die Bilddaten angewandt, um eine flächendeckende Klassifikation für das deutsche Bundesgebiet durchzuführen.

Abbildung 6.2 zeigt das vollständige Klassifikationsergebnis. Die Karte veranschaulicht die Heterogenität der Landschaften in Deutschland auf verschiedenen Skalen. Auf nationaler Ebene ist ein Trend zu niedriger Vegetation im Norden und höherer Vegetation im Süden des Landes zu erkennen. Landmarken, wie die Mittelgebirgslandschaft, sind durch dichte Baumvegetation klar erkennbar. Detaillierte Ansichten des Großraums Berlin in Abbildung 6.2 (oben rechts) zeigen ein komplexes Netz aus miteinander verbundenen städtischen Gebieten, Städten und Dörfern mit dazwischen liegenden Ackerflächen und Wiesen. In der Detailvergrößerung lassen sich zudem feine

Strukturen wie einzelne Straßen und unterschiedliche Vegetationsdecken (z. B. Anordnungen von Laub- und Nadelwäldern) unterscheiden. Selbst in dicht bebauten Gebieten erkennt der Klassifikator in den hochaufgelösten Bildern kleine Flächen von Stadtgrün.

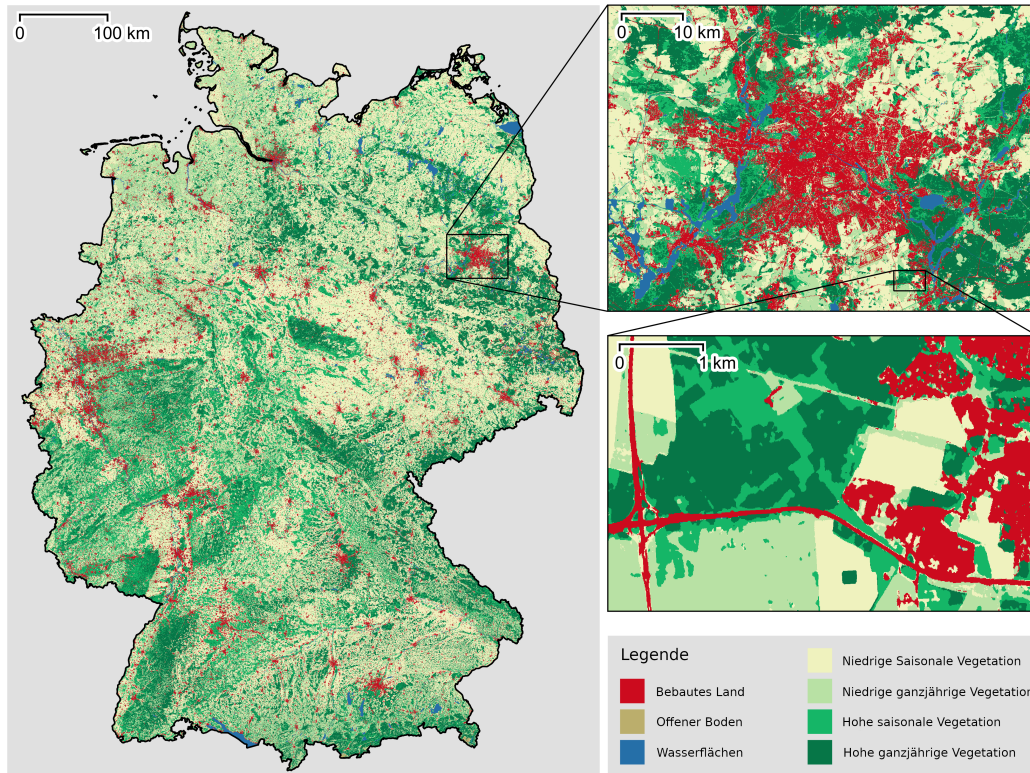


Abbildung 6.2: Überblick über die Landbedeckungsklassifikation aus Sentinel-2 Bildern und LUCAS-Referenzdaten, gemäß dem Vorprozessierungsschema S_3 (verändert nach Weigand et al., 2020). Die Zoom-Fenster zeigen den Detaillierungsgrad von der Landschaftsansicht über die städtische Struktur bis hin zu feinkörnigen Landbedeckungsmustern.

Tabelle 6.1 zeigt die detaillierte Fehlermatrix gewichtet nach dem geschätzten Modell basierend auf den Flächenanteilen einzelner Klassen am Gesamtergebnis (W_i). Insgesamt erreicht die resultierende Klassifikation eine geschätzte Genauigkeit (\hat{O}) von 93,07 % (Standardfehler = 0,43). Die geschätzte Gesamtgenauigkeit \hat{O} bezieht im Vergleich zur Gesamtgenauigkeit O die Klassenhäufigkeit im gesamten Klassifikationsprodukt ein und weicht daher leicht von der Gesamtgenauigkeit (93.6 %) ab.

Die klassenbasierte Genauigkeit betreffend erreichen alle Klassen eine hohe oder sehr hohe geschätzte Nutzergenauigkeit ($\hat{U} > 84$ %). Die niedrigste Nutzergenauigkeit und gleichzeitig der höchste Standardfehler wird von der am geringsten vertretenen Klasse, *offener Boden* (S), erreicht, was auf die generell geringe Anzahl von Referenzpunkten in dieser Klasse zurückzuführen ist. Dies deutet auf höhere Überklassifikation dieser Klasse hin, die aufgrund ihres geringen Vorkommens im Gesamtprodukt ($W_i = 0,002$) insgesamt einen geringen Einfluss auf das Gesamtergebnis ausübt. Die geschätzte Produzentengenauigkeit erreicht ebenfalls für die meisten Klassen sehr hohe Werte ($\hat{P} > 90$ %). Die Landbedeckungsklassen *offener Boden* (S) und *Wasser* (W) erreichen

Tabelle 6.1: Fehlermatrix für die deutschlandweite Landbedeckungsklassifikation 2015–2017 mit Schätzungen für die Gesamtgenauigkeit (\hat{O}), die Nutzergenauigkeit (\hat{U}) und die Produzentengenauigkeit (\hat{P}) sowie Standardfehler (SF, siehe Stehman und Foody (2019) für Details). Referenzwerte werden in den Spalten verortet, Prädiktionen werden in Zeilen geführt. W_i quantifiziert den Anteil der prädizierten Klassen am Gesamtergebnis. Tabelle verändert nach Weigand et al. (2020).

Präd./ Ref.	A	S	T1	T2	V1	V2	W	\hat{U} [%]	(SF)	W_i
A	6,90	0,15	0,15	0,05	0	0,20	0,15	90,73	(2,37)	0,076
S	0	0,16	0	0	0	0,03	0	84,62	(10,42)	0,002
T1	0,13	0	14,86	0,89	0,06	0,64	0,29	88,11	(1,41)	0,169
T2	0,03	0	0,42	15,84	0,06	0,22	0	95,59	(0,85)	0,166
V1	0,02	0,02	0,06	0	33,69	1,11	0	96,52	(0,45)	0,349
V2	0,08	0,05	0,74	0,19	0,93	20,18	0,21	90,17	(1,03)	0,223
W	0	0,02	0	0	0	0,03	1,44	96,84	(1,8)	0,015
\hat{P} [%]	96,42	39,78	91,51	93,36	96,98	90,05	68,94	$\hat{O} = 93,07$	(0,43)	
(SF)	(1,16)	(10,18)	(1,13)	(1,02)	(0,46)	(0,98)	(4,91)			

jedoch eine deutlich geringere geschätzte Produzentengenauigkeit von 39,78 % (SE = 10,18 %) bzw. 68,94 % (SE = 4,91 %), was darauf hindeutet, dass bei diesen Landbedeckungsklassen höhere Auslassungsfehler bzw. stärkere Unterklassifikationen auftreten. Während die Streuung zwischen den Klassen hauptsächlich zwischen den Vegetationsklassen, insbesondere zwischen hoher Vegetation (T1 und T2) und zwischen niedriger Vegetation (V1 und V2), besteht, lässt sich bei den bebauten Flächen (A) eine leichte Überschneidung mit Referenzpunkten der anderen Landbedeckungsklassen feststellen. Dies indiziert eine leichte Überschätzung von bebauten Flächen, was ebenfalls vom Verhältnis der Nutzergenauigkeit (\hat{U}) zur Produzentengenauigkeit (\hat{P}) bestätigt wird.

Diese Landbedeckungsklassifikation steht im EOC Geoservice des Deutschen Fernerkundungsdatenzentrums (DFD) des Deutschen Zentrums für Luft- und Raumfahrt (DLR) unter folgender DOI frei zum Download zur Verfügung: <https://doi.org/10.15489/1ccmlap3mn39>.

6.1.3 Übertrag der Landbedeckungsklassifikation auf den Zeitraum 2018

Nachdem die Methode zunächst für Sentinel-2 Bilddaten aus einem mehrjährigen Zeitraum entwickelt wurde, soll auch die Anwendbarkeit der Methode für Bilddaten aus nur einem Jahr getestet werden. Ein kürzerer Aufnahmezeitraum der Bilddaten hat den Vorteil, dass dieser zeitlich näher am Erfassungszeitraum der LUCAS-Daten liegt und zudem zeitlich stark variierende Flächen und Dynamiken besser raumzeitlich abgegrenzt werden können. Analog zur Bildklassifikation für den vorigen Zeitraum wird hierfür das Vorprozessierungsschema S_3 und die GRID-Positionierung

Tabelle 6.2: Fehlermatrix für die landesweite Landbedeckungsklassifikation 2018 mit Schätzungen für die Gesamtgenauigkeit (\hat{O}), die Nutzergenauigkeit (\hat{U}) und die Produzentengenauigkeit (\hat{P}) sowie Standardfehler (SF, siehe Stehman und Foody (2019) für Details). Referenzwerte werden in den Spalten verortet, Prädiktionen werden in Zeilen geführt. W_i quantifiziert den Anteil der prädizierten Klassen am Gesamtergebnis.

Präd./ Ref.	A	S	T1	T2	V1	V2	W	\hat{U} [%]	(SF)	W_i
A	6,49	0,11	0,04	0	0,08	0,38	0	91,40	(2,06)	0,071
S	0	0,08	0	0	0	0	0	100,00	(0,00)	0,001
T1	0,03	0	13,80	1,20	0,20	0,79	0,06	85,84	(1,49)	0,161
T2	0	0,03	0,81	15,64	0,06	0,27	0,09	92,55	(1,11)	0,169
V1	0,04	0,06	0,02	0,06	32,20	1,06	0,02	96,24	(0,46)	0,335
V2	0,31	0	0,95	0,58	1,75	21,41	0,08	85,37	(1,18)	0,251
W	0	0	0	0	0	0	1,31	100,00	(0)	0,013
\hat{P} [%]	94,55	27,82	88,39	89,48	93,9	89,54	83,9	$\hat{O} = 90,93$	(0,47)	
(SF)	(1,38)	(0,00)	(1,28)	(1,14)	(0,65)	(0,96)	(4,52)			

der LUCAS-Punkte verwendet. Die Ergebnisse der Klassifikation von Bilddaten zusammen mit LUCAS-Referenzdaten aus dem Jahr 2018 sind in Form der Fehlermatrix zusammengefasst (siehe Tabelle 6.2).

Wie auch für den Zeitraum von 2015 bis 2017, zeigt die Fehlermatrix der Landbedeckungsklassifikation 2018 eine hohe Gesamtgenauigkeit ($\hat{O} = 90,93 \%$), wengleich sie auch wenige Prozent unter dem vorigen Produkt liegt. Die erneut sehr hohen Nutzergenauigkeiten (\hat{U}) deuten an, dass nur geringe Überklassifikationen zu verzeichnen sind. Die stärksten Fehlklassifikationen bezogen auf die Verteilung der Landbedeckungsklassen in Deutschland konzentrieren sich wie zuvor auf die Vegetationsklassen. Diese umfassen besonders Fehlklassifikationen zwischen den Landbedeckungsklassen hoher Vegetation ($T1$ und $T2$) einerseits und den Klassen der niedrigen Vegetation ($V1$ und $V2$) andererseits. Auch die Werte der Produzentengenauigkeit zeigen ähnliche Muster wie im vorhergehenden Zeitraum. Die meisten Klassen weisen lediglich geringe Auslassungsfehler auf. Eine Ausnahme stellt die am geringsten vertretene Klasse *offener Boden* (S) dar. Referenzpunkte dieser Klasse sind besonders selten im Referenzdatensatz enthalten und werden überwiegend fehlklassifiziert. Wie zuvor fällt eine besondere Überlappung der Klasse *offener Boden* (S) mit der Klasse *bebautes Land* (A) auf.

Abbildung 6.3 zeigt einen Vergleich zwischen den Landbedeckungsklassifikationen von 2015/2017 und 2018 im Bereich der Stadt Würzburg. Die direkte visuelle Gegenüberstellung der Klassifikationsergebnisse zeigt, dass Dynamiken, wie beispielsweise die Baumaßnahmen³⁰ am Hubland Würzburg (vgl. Abbildung 6.3 mittig in den beiden unteren Bildausschnitten) in der ersten Klassifikation (2015/2017, linkes Bildbeispiel) klar als *offener Boden* (S) zu erkennen sind

³⁰Stadt Würzburg, ohne Datum, Hubland Luftbilder 2016, zuletzt abgerufen 27.07.2023, <https://www.wuerzburg-hubland.de/bildergalerie/luftbilder/520949.2016.html>

und 2018 abgeschlossen waren³¹. Ein weiterer Unterschied zwischen den beiden Klassifikationsergebnissen zeigt sich in der feingliedrigeren Erfassung von Grünflächen im Siedlungskörper im Klassifikationsergebnis von 2018. Auch innerhalb von Vegetationsflächen zeigen sich deutliche Unterschiede zwischen den Klassifikationsergebnissen, wie beispielsweise im oberen Bildausschnitt der Zoom-Fenster (siehe Abb. 6.3 unten) oder im Ringpark der Stadt Würzburg im linken Bereich der Zoom-Fenster. Während 2015/2017 die Klasse *hohe saisonale Vegetation* (T1) überwiegt, wird dort für das Jahr 2018 zunehmend *niedrige ganzjährige Vegetation* (V2) klassifiziert. Die Fehlermatrix bietet einen Hinweis auf derartige Klassifikationsunterschiede in Form prozentualer Fehler.

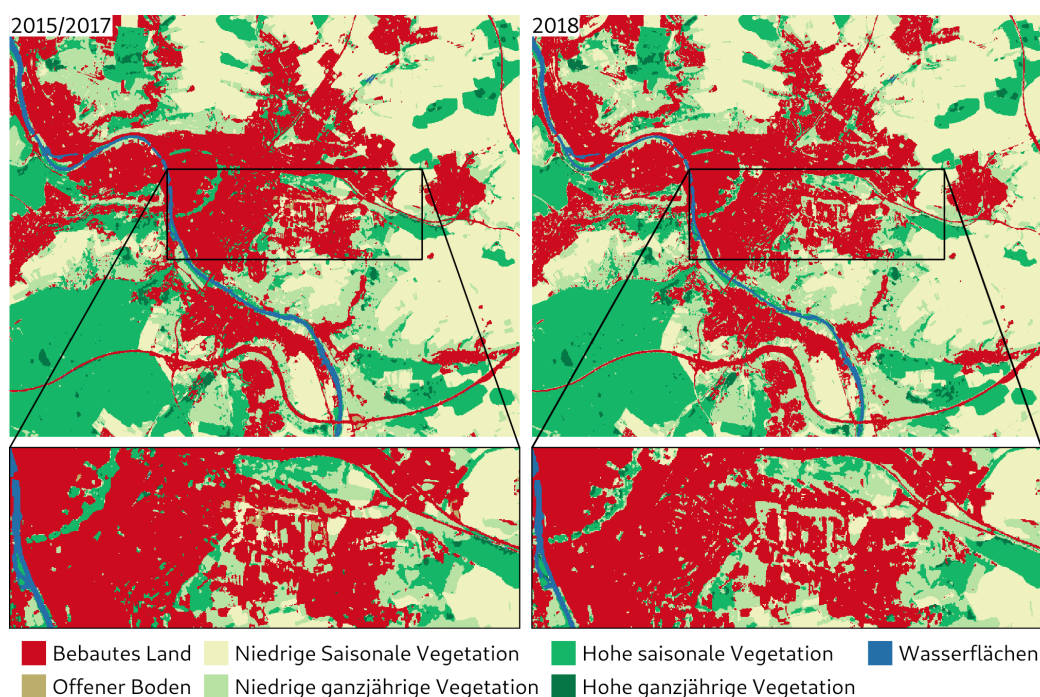


Abbildung 6.3: Vergleich der Landbedeckungsklassifikationen von 2015/2017 (links) und 2018 (rechts) am Beispiel der Stadt Würzburg.

6.2 Erfassung öffentlicher Grünflächen in Deutschland

Die Ergebnisse des zweiten konzeptionellen Ansatzes dieser Dissertation zur Schätzung der verfügbaren Fläche von öffentlichem Grün je Nachbarschaft werden im Folgenden aufgezeigt. Zunächst wird die Deep Learning-basierte Modellierung anhand des Trainingsverlaufes sowie der Genauigkeitsmetriken evaluiert. Anschließend werden anhand geographischer Beispiele die Ergebnisse räumlich verdeutlicht.

³¹Stadt Würzburg, ohne Datum, Hubland Luftbilder 2018, zuletzt abgerufen 27.07.2023, <https://www.wuerzburg-hubland.de/bildergalerie/luftbilder/520936.2018.html>

6.2.1 Modellierungsgenauigkeit

Abbildung 6.4 zeigt den Verlauf des Modelltrainings des Fusionsmodells. Es werden verschiedene Metriken für Trainings- und Validierungsschritte unterschieden. Die Verläufe zeigen ein asymptotisches Verhalten des Huber Verlusts wie auch der Fehlermetriken hin zu einem Minimalwert. Nach 3 Aufwärmepochen, welche dazu dienen, das Netzwerk nach randomisierter Initialisierung mit sehr geringer Lernrate (1×10^{-15}) auf die gradientbasierte Optimierung vorzubereiten, sinkt der Fehler insbesondere in den ersten ca. 32 Epochen. Der sichtbare Unterschied zwischen *MAE* und *RMSE* legt nahe, dass in diesen ersten Epochen vornehmlich die starken Ausreißer reduziert werden, welche deutlich stärker durch den *RMSE* überbetont werden. Nach ca. 35 Trainingsepochen werden nur noch geringe Performanzzugewinne erzielt. Sowohl Trainings- als auch Validierungsmetriken sinken stetig asymptotisch ab, was zeigt, dass keine Überanpassung zu verzeichnen ist.

Durch die Evaluation des trainierten Netzwerkes mithilfe einer unabhängigen Stichprobe von Referenznachbarschaften wird die Modellgüte bestimmt. Insgesamt erzielt das Fusionsmodell ein Bestimmtheitsmaß von $R^2 = 0,952$ mit einem mittleren absoluten Fehler von $MAE = 3,31$ ha und einem *RMSE* von 5,50 ha. In Abbildung 6.5 sind die Werte der Prädiktion und die Referenz der Flächen an öffentlichem Grün für alle Nachbarschaften des Testdatensatzes gegenübergestellt. Hier zeigt sich, dass das Modell den Trend der Grünflächenanteile gut abbilden kann. Im optimalen Fall würde das Modell die Referenzwerte perfekt nachbilden können. Alle Punkte würden dann auf der Diagonale verortet, welche durch die gestrichelte Linie verdeutlicht wird. Hier zeigt sich eine deutliche Konvergenz der Prädiktion entlang dieses Equilibriums. Die seitlichen Histogramme (oben und rechts) der Häufigkeitsverteilung der Flächenwerte zeigen die Übereinstimmung der Streuungen zwischen der Referenz und der Prädiktion. Dennoch gibt es einzelne, teils deutliche Abweichungen des Modells.

Die Abweichungen zwischen Prädiktion und Referenz, auch Residuen genannt, werden in Abbildung 6.6 separat aufgetragen. Hier wird sichtbar, dass das Modell speziell für Gebiete, welche in der Referenz geringe Mengen an öffentlichem Grün aufweisen diese leicht überschätzt werden. Gebiete mit mehr öffentlicher Grünfläche werden hingegen vom Modell tendenziell etwas unterschätzt. Diese Tendenz lässt sich aus dem geschätzten linearen Trend (grüne Linie) ableiten. Die Häufigkeitsverteilung der Residuen (abzulesen am Histogramm am rechten Rand) zeigt, dass die Mehrzahl der Residuen um den Wert 0 verteilt sind.

6.2.2 Geographische Analyse der Modellierungsgenauigkeit

Über die rein statistische Betrachtung der Genauigkeit hinaus spielen geographische Aspekte eine wichtige Rolle in der Evaluation eines derartigen Modells. Abbildung 6.7 zeigt die räumliche Anordnung einiger Referenznachbarschaften am Beispiel Münchens. Es ist deutlich zu sehen,

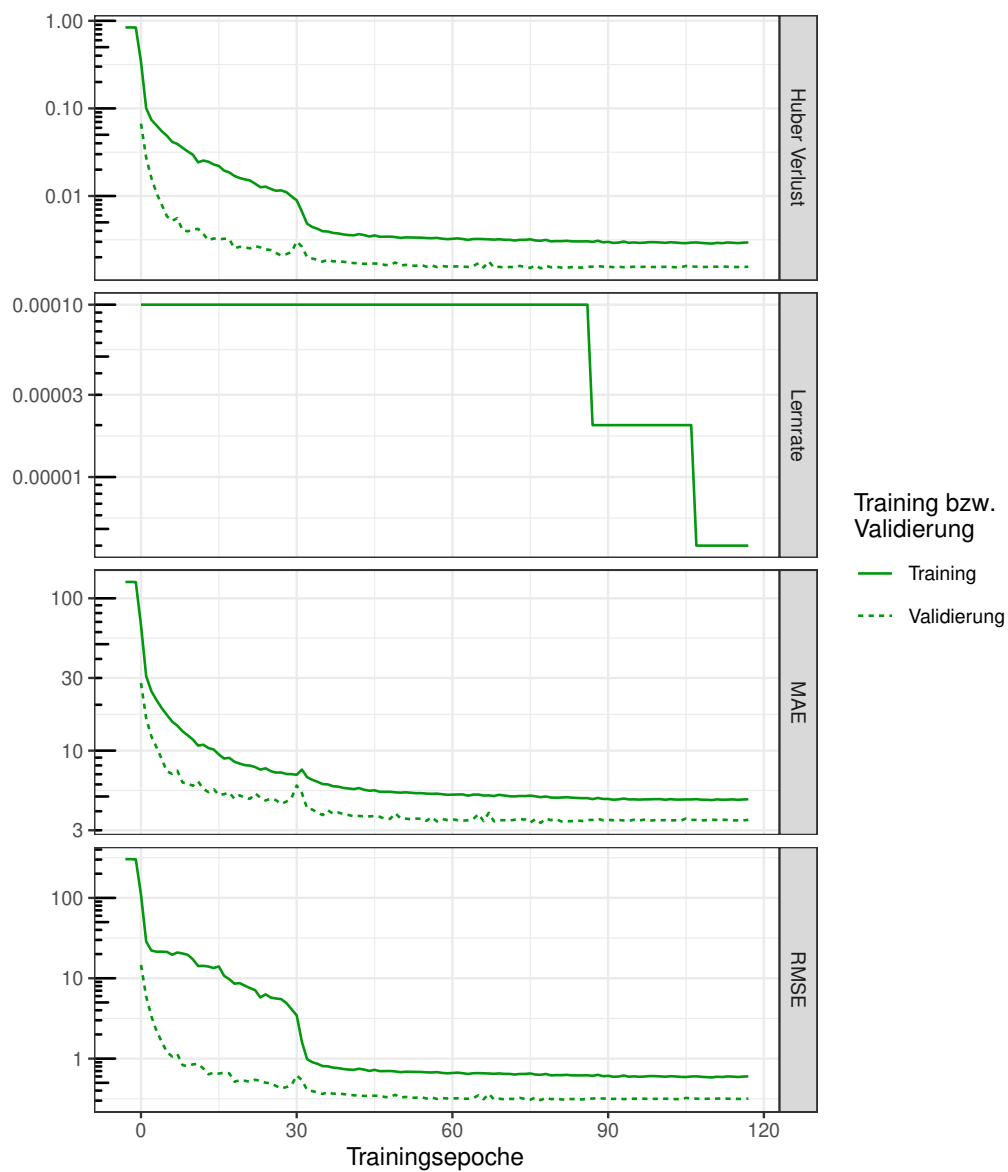


Abbildung 6.4: Metriken des Trainingsverlaufs des Fusionmodells. Der Verlauf wird entlang der Trainingsepochen aufgetragen (x-Achse). Im Lauf des Trainings kann der Huber Verlust, welcher zur Optimierung des Regressionsmodells genutzt wird, gesenkt werden. Die Senkung der Lernrate erfolgt, wenn 20 nach Epochen keine weitere Verbesserung des Validierungsverlusts mehr erzielt werden konnte. Der mittlere absolute Fehler (MAE) und die Wurzel aus dem mittleren quadratischen Fehler (RMSE) dienen zur qualitativen Bewertung des Lernerfolges.

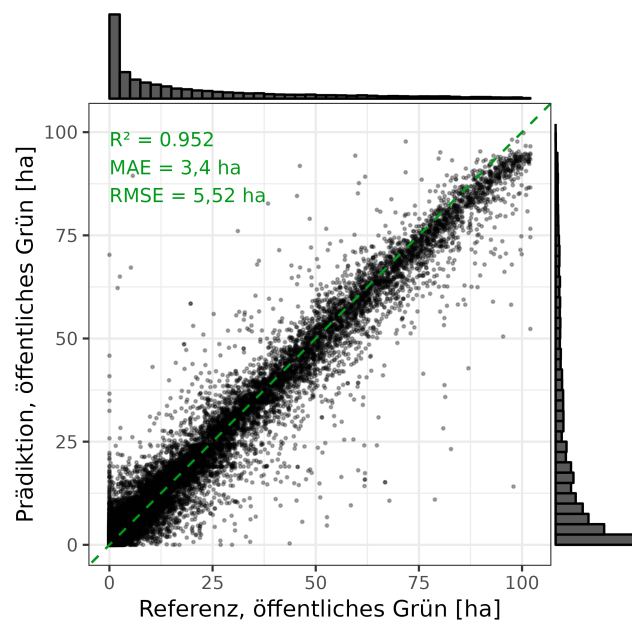


Abbildung 6.5: Vergleich von Referenz (x-Achse) und Prädiktion (y-Achse) öffentlicher Grünflächen des Fusionsmodells auf Ebene von Nachbarschaften. Die optimale Schätzung ist durch die gestrichelte Linie auf der Diagonalen repräsentiert.

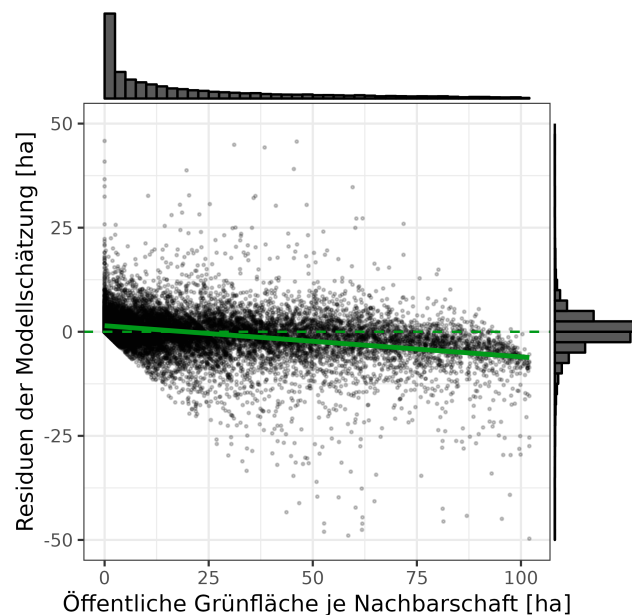


Abbildung 6.6: Gegenüberstellung der Referenz (x-Achse) mit den Residuen des Fusionsmodells, ebenfalls in Hektar (y-Achse). Ein optimales Modell würde für alle Referenznachbarschaften Residuen von 0 schätzen, gekennzeichnet durch die gestrichelte grüne Linie. Die durchgezogene grüne Linie zeigt den Trend der Residuen in Form eines linearen Modells an.

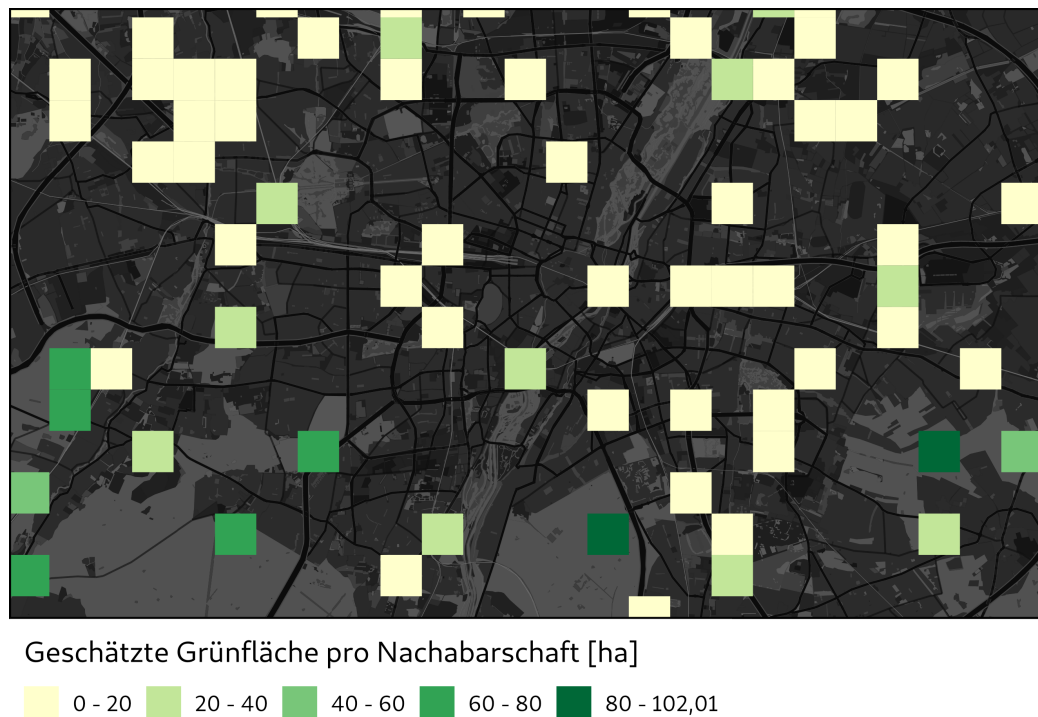


Abbildung 6.7: Räumliches Beispiel der Modellierung der verfügbaren Fläche öffentlichen Grüns je für ausgewählte Referenznachbarschaften im Münchener Stadtgebiets.

wie Gebiete in der Innenstadt vorwiegend niedrige Werte an öffentlichem Grün aufweisen während Nachbarschaften am Stadtrand, z. B. im Südwesten und Südosten von größeren Anteilen öffentlicher Grünflächen geprägt sind.

In [Abbildung 6.8](#) sind die Abweichungen des Modells vom Referenzwert aufgetragen. Hier wird deutlich, dass die PGS Fläche der meisten Nachbarschaften durch das Modell gut geschätzt wird (Kacheln in hellem Gelb). Dennoch treten, wie den obigen Ergebnissen zu Residuen zu entnehmen ist, mitunter deutliche Abweichungen zwischen Referenz- und Schätzwert auf. Es sind hier keine deutlichen räumlichen Muster erkennbar und Über- wie Unterschätzungen (in Blau- bzw. Rottönen) scheinen unabhängig vom Bebauungsgrad.

Diese Unabhängigkeit der Schätzwerte und Residuen bestätigt sich weiter, wenn man die Residuen der Modellierung von öffentlichem Grün in Relation der Distanz zur Bebauung betrachtet (siehe [Abbildung 6.9](#)) lässt sich kein bestimmter Trend ausmachen. Daraus lässt sich schließen, dass die Performanz des Fusionsmodells unabhängig vom Grad der Bebauungsdichte und der Urbanität ist.

Die Experimente mit dem Fusionsnetzwerk zeigen exemplarisch, dass durch die Kombination von Satelliten- und OSM-Daten Zusammenhänge gelernt werden können, um Landnutzungsinformationen zu lernen (hier am Beispiel öffentlicher Grünflächen). Das resultierende Modell

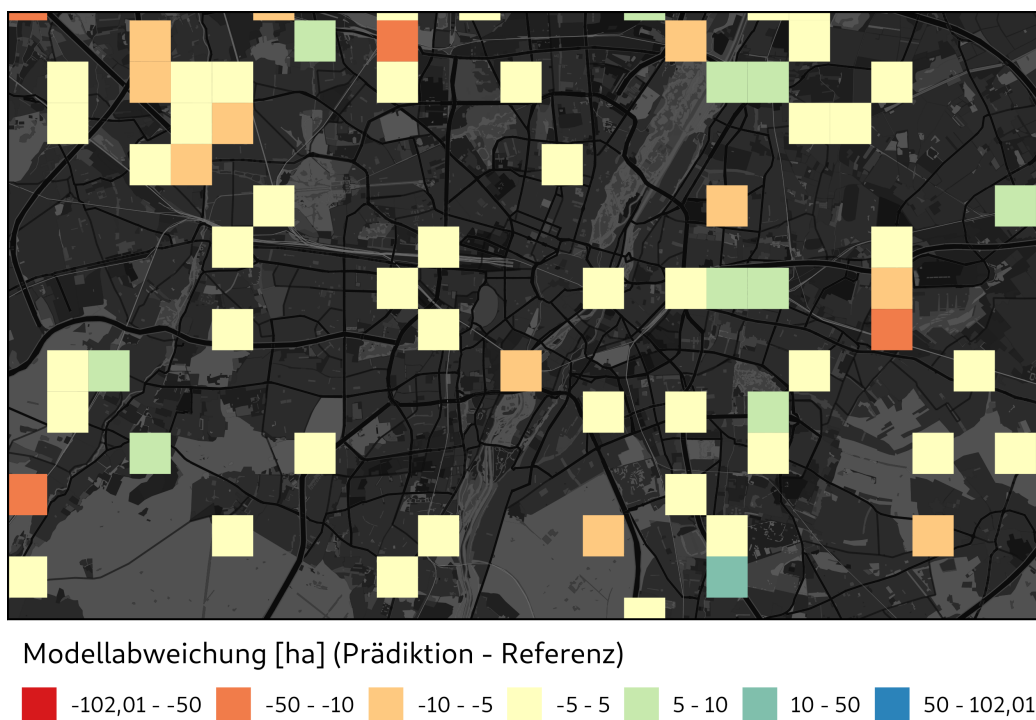


Abbildung 6.8: Geographische Visualisierung der Residuen des Fusionsmodells. Rötliche Farbtöne kennzeichnen Nachbarschaften, in welchen das Modell den Referenzwert für öffentliche Grünflächen unterschätzt, bläuliche Farbtöne zeigen Nachbarschaften mit Überschätzung.

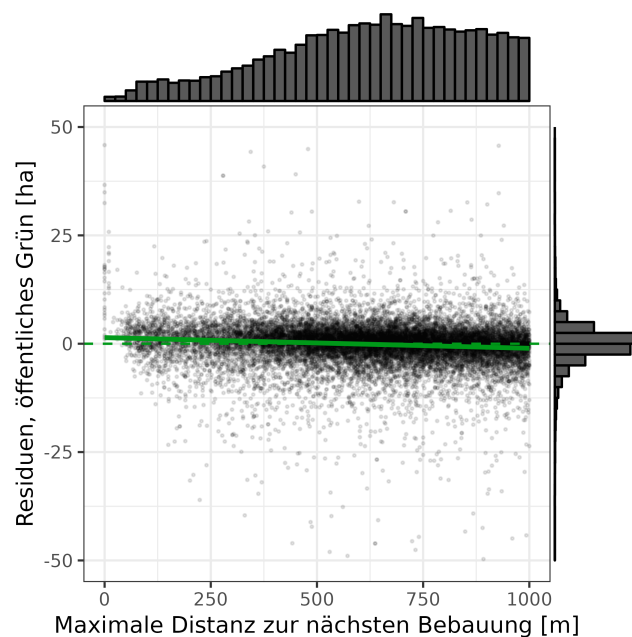


Abbildung 6.9: Residuen des Fusionsmodells (y-Achse) in Abhängigkeit zur Distanz der Nachbarschaft zu Bebauung (x-Achse). Die gestrichelte grüne Linie zeigt die Schätzung des optimalen Modells entlang der Nulllinie der y-Achse an. Der Trend aller Residuen ist mithilfe eines linearen Modells geschätzt und weicht nur geringfügig von dieser Nulllinie ab (durchgezogene grüne Linie).

ermöglicht die Informationen zur Verfügbarkeit von öffentlichem Grün aus dem European Urban Atlas, welcher im Original nicht flächendeckend für ganz Deutschland vorliegt, auf das gesamte Bundesgebiet zu extrapolieren. Diese dienen im Folgenden als Grundlage für die raumquantitative Verteilungsanalyse von Grün in Deutschland.

6.3 Gerechtigkeit der Verteilung von Grünflächen

Die Ergebnisse des dritten konzeptionellen Schwerpunktes dieser Dissertation werden im Folgenden betrachtet. Anhand der in vorherigen Arbeitsschritten erstellten Landbedeckungsklassifikation und der Quantifizierung von öffentlichem Grün wird die Verteilung von Grünflächen in Deutschland auf Nachbarschaftsebene zunächst mithilfe von deskriptiven Statistiken analysiert. Im Zentrum stehen hierbei stets die Menschen in Deutschland. Die Wohnumgebung derer dient als Ausgangspunkt für die folgenden Analysen und wird lokalisiert mithilfe von geocodierten Zensusdaten. Auf die deskriptiven Statistiken folgt die Analyse der Verteilungsgerechtigkeit dieser gesundheitsrelevanten Ressource in der Wohnumwelt der Menschen. Abschließend wird der Zusammenhang zwischen der demographischen Bevölkerungszusammensetzung und den Grünflächenanteilen untersucht und beschrieben.

6.3.1 Grünflächenverfügbarkeit in Deutschland

Die Grünflächenanteile für alle Zensus-Nachbarschaften in München sind in Abbildung 6.10 kartiert. Es wird deutlich, wie die Grünflächenverfügbarkeit auf Nachbarschaftsebene im Stadtgebiet variiert. Während der vorwiegende Teil der Nachbarschaften mehr als 3,6 ha öffentliches Grün beinhalten (grüne Farbtöne), stechen einzelne Gebiete mit geringerem PGS-Anteil hervor (orange und blaue Farbtöne), beispielsweise Forstenried im Südwesten der Stadt. Dieses Stadtviertel ist geprägt von vielen Einfamilienhäusern mit Privatgärten.

Die Verteilung von Grünflächen in der Wohnumgebung der Menschen in Deutschland ist in Tabelle 6.3 zusammengefasst. Im Durchschnitt finden die Menschen in Deutschland 37,4 ha an grüner Landbedeckung in ihrer Nachbarschaft vor (Median 34,0 ha). Der große Unterschied zwischen dem ersten und dem dritten Quartil sowie die hohe Standardabweichung von 21,2 ha zeigt an, dass es große Schwankungen innerhalb der Bevölkerung gibt.

Im Vergleich zur Menge an grüner Landbedeckung beträgt die durchschnittliche Fläche an öffentlichem Grün mit 12,3 ha (Median 8,2 ha) nur ein Drittel. Wie zuvor für die GLC weist die hohe Standardabweichung 12,1 ha auf eine große Streuung der Verteilung hin. Die relativ geringere Abweichung zwischen Median und Mittelwert bei der Verteilung grüner Landbedeckung bedeutet, dass diese weniger von Ausreißern geprägt ist als die Verteilung von öffentlichem Grün.

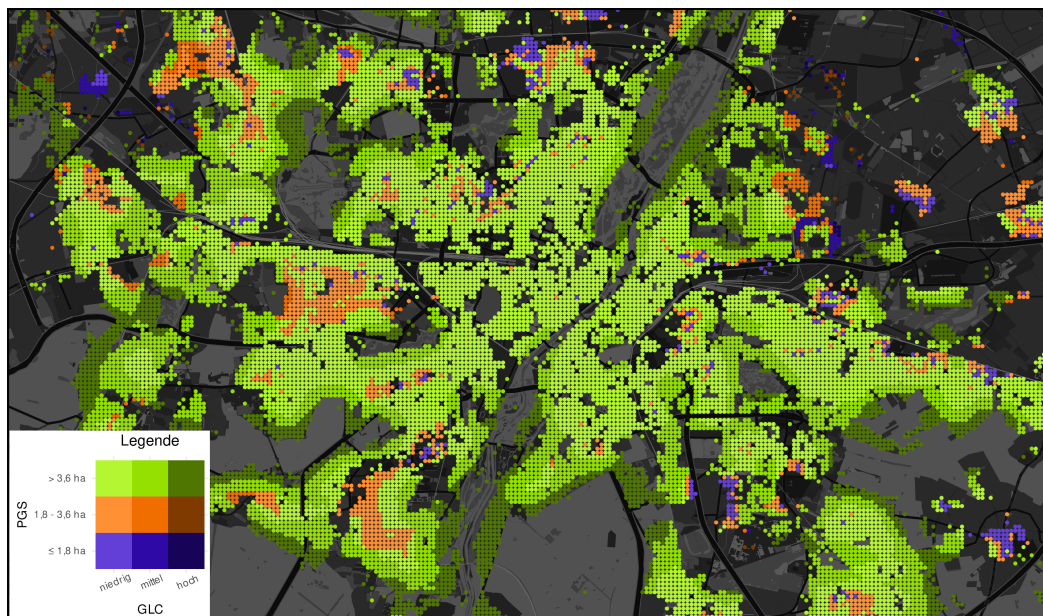


Abbildung 6.10: Grünflächenverfügbarkeit in der Nachbarschaft aller Zensus Gitterpunkte in München klassifiziert nach öffentlichem Grün (PGS, gekennzeichnet durch verschiedene Farben) und grüner Landbedeckung (GLC, gekennzeichnet durch verschiedene Farbtöne).

Tabelle 6.3: Gesamtstatistik zur Grünflächenverfügbarkeit in der Bevölkerung Deutschlands in Hektar.

Typ	Mittelwert	Std. Abw.	1. Quartil	Median	3. Quartil
Grüne Landbedeckung (GLC)	37,4	21,2	21,2	34,0	50,6
Öffentliche Grünflächen (PGS)	12,3	12,1	4,5	8,2	15,9

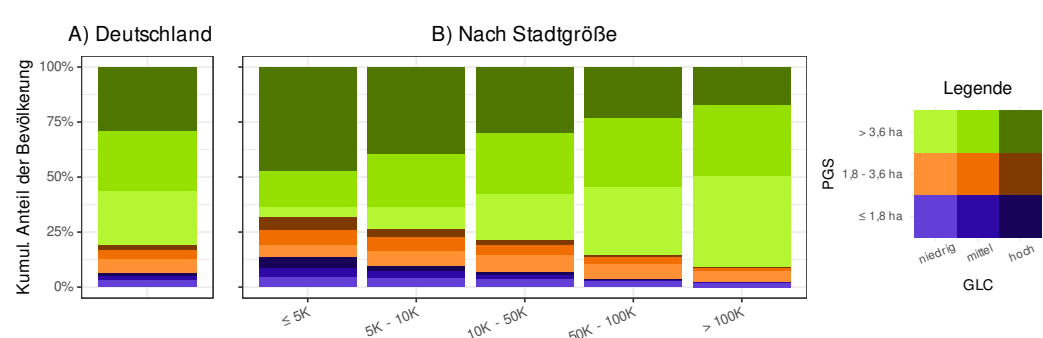


Abbildung 6.11: Verfügbarkeit von grüner Landbedeckung (GLC) und öffentlichen Grünflächen (PGS) auf Nachbarschaftsebene in Deutschland nach kumulierten Anteilen an der Bevölkerung. Die einzelnen Diagramme verwenden unterschiedliche Gruppierungsvariablen: A) Gesamtdeutschland, B) nach Gemeindegröße in Tausend (K) Einwohnern. Abbildung verändert nach Weigand et al. (2023b).

Abbildung 6.11 A) zeigt die Verfügbarkeit von GLC und PGS für die deutsche Bevölkerung als kumulative Anteile. Dabei haben 80,8 % der Menschen Zugang zu 3,6 ha oder mehr PGS in ihrer Nachbarschaft, 12,6 % haben zwischen 1,8 ha und 3,6 ha, weitere 6,6 % haben 1,8 ha oder weniger PGS zur Verfügung. Das bedeutet, 19,2 % der Bevölkerung in Deutschland sind im Hinblick auf die von der WHO gesetzten Ziele unterversorgt.

Betrachtet man die Ergebnisse nach Gemeindegröße in Abbildung 6.11 B), wird deutlich, dass kleinere Gemeinden im Allgemeinen einen höheren Anteil an grüner Landbedeckung aufweisen, was durch den größeren Anteil an dunkleren Farbtönen zum Ausdruck kommt. Umgekehrt steigt der Anteil der Bevölkerung, der Zugang zu mehr öffentlichen Grünflächen hat, mit der Stadtgröße. Während in kleinen Gemeinden mit weniger als 5.000 Einwohnern nur zwei Drittel der Bevölkerung Zugang zu 3,6 ha öffentlichen Grünanlagen oder mehr haben, steigt dieser Anteil in Städten mit mehr als 100.000 Einwohnern auf 91 %. Dies verdeutlicht, dass die Bedeutung von öffentlichen Grünflächen als wichtigste Ressource für Grünflächen mit der Gemeindegröße zunimmt während in kleineren Gemeinden der Anteil von hohen Mengen grüner Landbedeckung sehr deutlich dominiert.

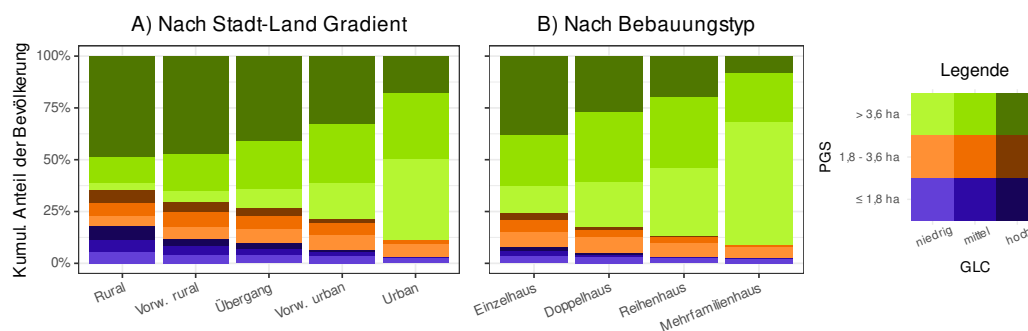


Abbildung 6.12: Verfügbarkeit von grüner Landbedeckung (GLC) und öffentlichen Grünflächen (PGS) auf Nachbarschaftsebene in Deutschland nach kumulierten Anteilen an der Bevölkerung. Die einzelnen Diagramme verwenden unterschiedliche Gruppierungsvariablen: A) nach dem 5-Klassen-Stadt-Land-Gefälle von (Taubenböck et al., 2022) und B) in Stadtvierteln mit bestimmten Gebäudetypen. Abbildung verändert nach Weigand et al. (2023b).

Die feinere lokale Auflösung des Land-Stadt-Gefälles durch Taubenböck et al. (2022), welches anders als die Stadtgröße den Urbanitätsgrad auf Ebene von einzelnen Nachbarschaften erfasst, unterstreicht diese Ergebnisse noch (vgl. Abbildung 6.12 A)). In ländlichen Gebieten gibt es eine Mehrheit von Nachbarschaften mit mittleren und hohen GLC-Werten, während 35,5 % der Bevölkerung nicht genügend PGS in der Nachbarschaft zur Verfügung haben. Im Gegensatz dazu haben nur 11,7 % der Bevölkerung in städtischen Zentren Zugang zu weniger als 3,6 ha PGS.

Die Verfügbarkeit von öffentlichem Grün und grüner Landbedeckung in Stadtvierteln mit unterschiedlichen Wohnungstypen in Abbildung 6.12 B) zeigt ebenfalls auffällige Trends. Mehr als die Hälfte (61,5 %) der Bevölkerung in Nachbarschaften mit freistehenden Einfamilien-, Zweifamilien- oder Reihenhäusern findet mehr als 25 ha GLC in ihrem Viertel vor (mittlere

und dunkle Farbtöne). Im Gegensatz dazu finden 24,5 % der Bevölkerung in Gebieten mit freistehenden Einfamilienhäusern eine unterdurchschnittliche Verfügbarkeit von GLC vor. Im Vergleich zur Gesamtbevölkerung Deutschlands ist dieser Anteil 5,3 % höher. Die Ergebnisse für Gebiete mit Doppelhaushälften, Reihenhäusern und Mehrfamilienhäusern zeigen, dass diese Gebiete zunehmend weniger GLC und insgesamt mehr PGS-Verfügbarkeit aufweisen.

6.3.2 Gerechtigkeit von Grünflächenverteilung

Die Gerechtigkeit der Verteilung der gesundheitsrelevanten Umweltressource Grünflächen wird ebenfalls anhand der zwei Typen von Grünflächen, grüne Landbedeckung und öffentliches Grün, betrachtet. Die Unterschiede der Verteilung unter der deutschen Bevölkerung werden in Abbildung 6.13 in Form der Lorenzkurve verdeutlicht. Es ist erkennbar, dass die Gesamtmenge grüner Landbedeckung gerechter in der Bevölkerung verteilt ist, als öffentliche Grünflächen. Während die Hälfte der Gesamtfläche von grüner Landbedeckung unter ca. 70 % der Bevölkerung verteilt sind, ist die Hälfte der Menge von öffentlichem Grün auf ca. 80 % der Bevölkerung verteilt. Das bedeutet die restlichen 50 % der grünen Landbedeckung bzw. des öffentlichen Grüns entfallen auf 30 % beziehungsweise 20 % der Bevölkerung. Hierbei ist zu berücksichtigen, dass die beiden Lorenzkurven relativ zur Gesamtmenge der jeweiligen Ressource fungieren, welche deutlich divergieren. Dennoch zeigt der Vergleich der beiden Lorenzkurven eindrucksvoll, dass öffentliches Grün einer weniger ausgeglichenen Verteilung in der Bevölkerung unterliegt.

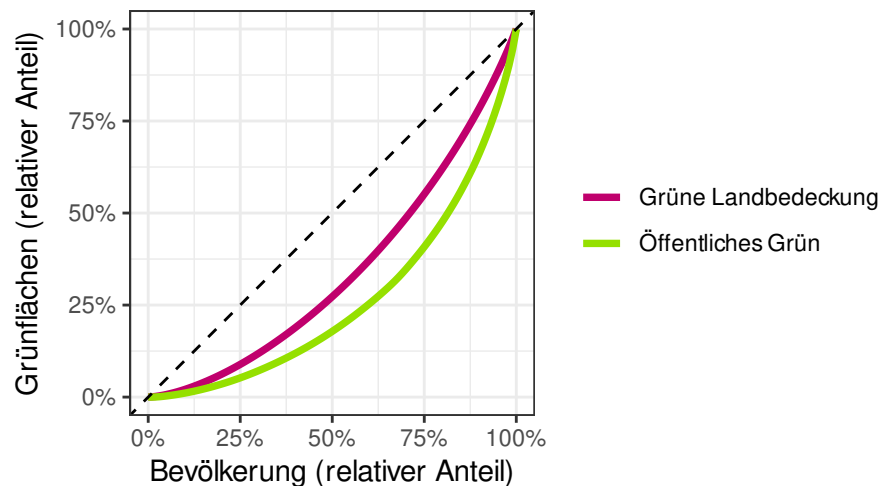


Abbildung 6.13: Lorenzkurve der Verteilung von Grünflächen in Deutschland. Die gestrichelte Diagonale zeigt den Zustand absoluter Gleichverteilung der Ressource an. Abbildung verändert nach Weigand et al. (2023b).

Neben der visuellen Einordnung der Verteilungsgerechtigkeit der Grünflächenverfügbarkeit in Deutschland durch Lorenzkurven werden im Folgenden Gini- und Atkinson-Koeffizienten zur Bewertung der Verteilungsgerechtigkeit hinzugezogen. In Tabelle 6.4 sind die Ergebnisse

der Koeffizienten für ganz Deutschland aufgeführt. Insgesamt ist die Ungleichheit der GLC-Verfügbarkeit deutlich geringer (Gini = 0,32), als die Ungleichheit der PGS-Verfügbarkeit (Gini = 0,48). Die Atkinson-Koeffizienten bestätigen diese Differenz, doch fügen sie eine weitere interessante Komponente hinzu. Bei stärkerer Gewichtung der Bereiche der Ressourcenverteilung mit geringerem Grünanteil – durch höheres ϵ – steigt die Ungleichheit der Verteilung für öffentliches Grün deutlich stärker an als die der grünen Landbedeckung. Wie auch zuvor die Lorenzkurve deutet dies darauf hin, dass öffentliches Grün weit mehr konzentriert ist als grüne Landbedeckung.

Tabelle 6.4: Verteilungsgerechtigkeit von Grünflächen in Deutschland nach Typ.

Typ	Gini	Atkinson ($\epsilon = 0,5$)	Atkinson ($\epsilon = 0,75$)
Grüne Landbedeckung (GLC)	0,318	0,086	0,132
Öffentliche Grünflächen (PGS)	0,477	0,188	0,285

Vergleicht man die Gini-Indizes für GLC und PGS nach der Größe der Gemeinde (vgl. Abb. 6.14), so wird deutlich, dass die Verteilungsgerechtigkeit von grüner Landbedeckung und öffentlichen Grünflächen mit der Stadtröße variiert. In der Tat besteht eine umgekehrte Beziehung zwischen den Gini- bzw. Atkinson-Indizes von GLC und PGS. Im Falle von Großstädten mit mehr als 100.000 Einwohnern erreichen die Gerechtigkeitsmetriken für GLC ($G = 0,34$, $A(0,5) = 0,10$, $A(0,75) = 0,16$) und PGS ($G = 0,39$, $A(0,5) = 0,12$, $s(0,75) = 0,18$) vergleichbare Werte, während sie für die Bevölkerung in Gemeinden mit weniger als 5.000 Einwohnern drastisch divergieren (GLC: $G = 0,24$; PGS: $G = 0,56$). Auffällig ist darüber hinaus, dass die Differenzen zwischen Atkinson-Koeffizienten mit unterschiedlichem ϵ bei zunehmender Stadtgröße inverse Trends aufweisen. Für öffentliches Grün nähern sich die Gerechtigkeitsmetriken in größeren Städten zunehmend einander an, während sie für grüne Landbedeckung zunehmend divergieren, wenn auch in geringerem Abstand. Daraus kann gefolgert werden, dass Verfügbarkeit von öffentlichem Grün in Großstädten deutlich fairer verteilt ist als in kleineren Gemeinden und die Gerechtigkeit insbesondere für die Menschen mit besonders geringem Zugang zu öffentlichem Grün zunimmt. Gegensätzlich ist die Fläche von grüner Landbedeckung in kleineren Gemeinden regelmäßiger verteilt.

In Erweiterung der Analyse nach Gemeindegröße verdeutlichen die Ergebnisse der Indizes entlang des Land-Stadt-Gefälles, dass die Ungleichheit zwischen GLC und PGS auch bei einer stärkeren lokalen Differenzierung der Urbanität bestehen bleibt (vgl. Abbildung 6.15). Die Verteilungsgerechtigkeit von öffentlichen Grünflächen sinkt entlang des Gradienten von ruralen Gebieten zu urbanen Zentren stetig an, gleichzeitig sinkt die Gleichmäßigkeit des Zugangs zu grüner Landbedeckung erst am Übergang zu vorwiegend urbanen Gebieten.

Auch der Zusammenhang des Gebäudetyps mit der Ungleichheit der Grünflächen zeigt klare Trends (vgl. Abbildung 6.16). Die Ungleichheit der GLC nimmt von Gebieten mit Einfamilienhäusern zu Gebieten mit Mehrfamilienhäusern zu. Dieser Trend stimmt mit den vorherigen Ergebnissen überein, da Mehrfamilienhäuser in dichteren städtischen Zentren häufiger anzutreffen

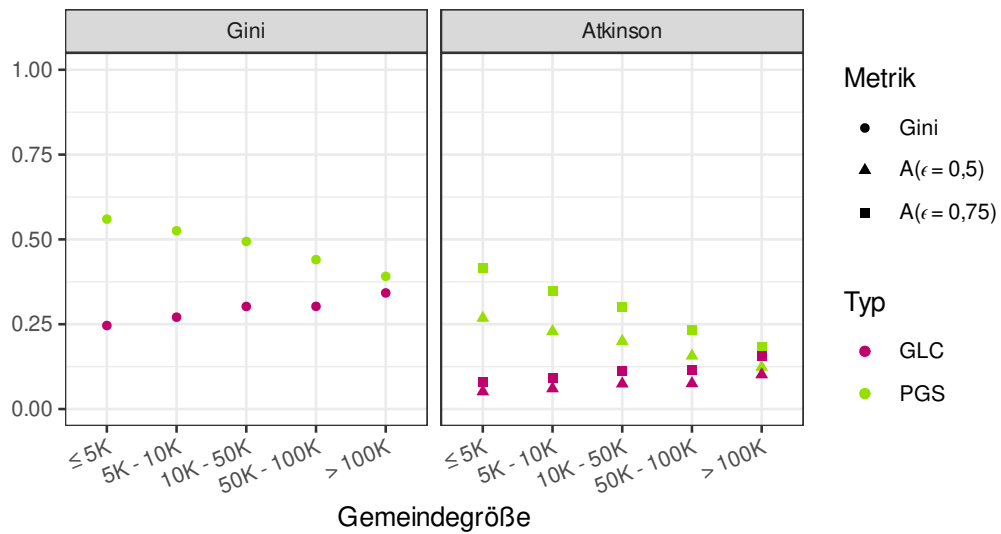


Abbildung 6.14: Gini- (links) und Atkinson-Koeffizienten (rechts) als Gerechtigkeitsmaße der Grünflächenverteilung in Deutschland für grüne Landbedeckung (GLC, pink) und öffentliche Grünflächen (PGS, grün) nach der Gemeindegröße, angegeben in Klassen von Tausend (K) Einwohnern pro Gemeinde. Abbildung verändert nach Weigand et al. (2023b).

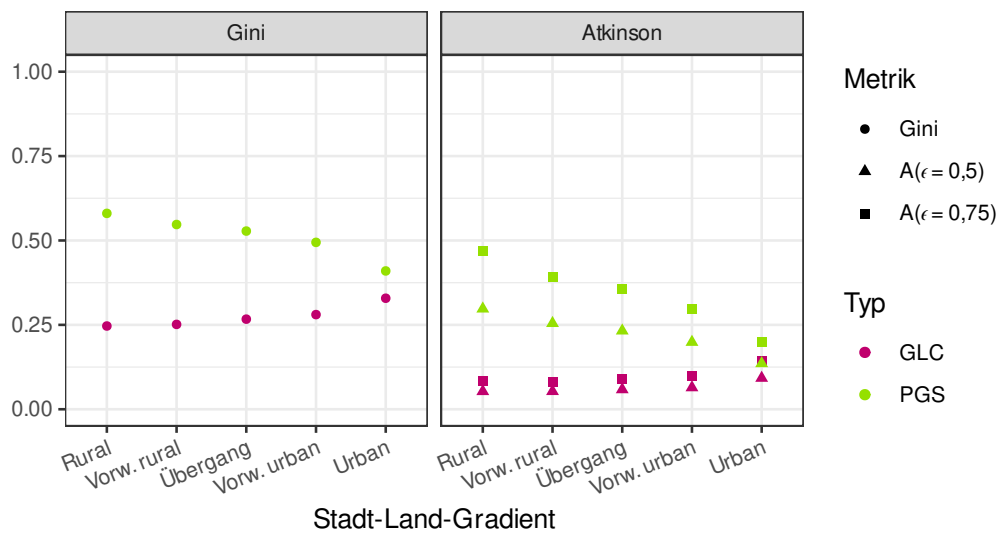


Abbildung 6.15: Gini- (links) und Atkinson-Koeffizienten (rechts) als Gerechtigkeitsmaße der Grünflächenverteilung in Deutschland für grüne Landbedeckung (GLC, pink) und öffentliche Grünflächen (PGS, grün) nach der Klassifizierung der Stadtteile entlang des Stadt-Land-Gradienten. Abbildung verändert nach Weigand et al. (2023b).

sind. Bemerkenswert ist, dass die Ungleichheit der Ressource GLC (Gini = 0,38) die Ungleichheit der PGS (Gini = 0,35) in Mehrfamilienhausgebieten übersteigt. Gleiches gilt für beide Atkinson-Koeffizienten. Das bedeutet, dass in all diesen Vierteln die verfügbare Gesamtmenge an grüner Landbedeckung weniger gleichmäßig auf die Bevölkerung in diesen Gebieten Deutschlands verteilt ist als die verfügbare öffentliche Grünfläche. Diese Ergebnisse zeigen, dass die Ungleichheit bei der Verfügbarkeit von GLC in der Nachbarschaft mit der Stadtgröße, der Urbanität und dichteren Wohnformen zunimmt. Umgekehrt nimmt entlang der gleichen Achsen die Ungleichheit bei der PGS-Verfügbarkeit ab.

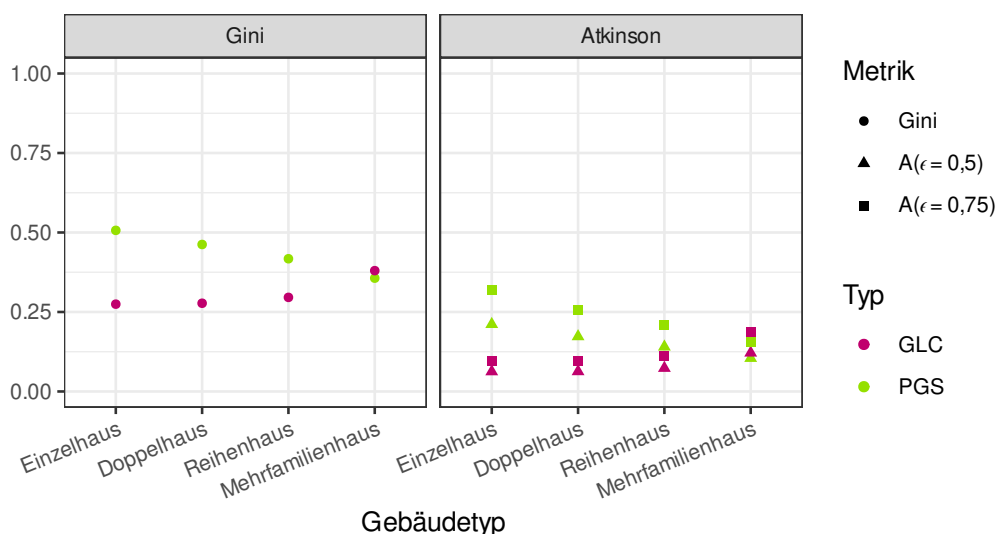


Abbildung 6.16: Gini- (links) und Atkinson-Koeffizienten (rechts) als Gerechtigkeitsmaße der Grünflächenverteilung in Deutschland für grüne Landbedeckung (GLC, pink) und öffentliche Grünflächen (PGS, grün) nach Stadtteilen mit unterschiedlichen Gebäudetypen. Abbildung verändert nach Weigand et al. (2023b).

Abbildung 6.17 zeigt einen Vergleich der Verfügbarkeit von Grünflächen in den 100 bevölkerungsreichsten Städten in Deutschland. Zwischen einzelnen Städten sind starke Unterschiede zu verzeichnen. Der Anteil der Bevölkerung, der Zugang zu mehr öffentlichen Grünanlagen als das WHO-Ziel hat, schwankt von fast 100 % der Bevölkerung in Bergisch Gladbach bis zu 55,3 % in Minden. Im Kontrast zur deutlichen Zunahme von Bevölkerungsteilen, welche weniger als 3,6 Hektar öffentliche Grünflächen in der Wohnumgebung vorfinden, spiegeln die Gini- und Atkinson-Koeffizienten für die Verteilung des Zugangs zu dieser Ressource in diesen Städten keinen derart drastischen Trend wider. Dies verdeutlicht, dass derartige Koeffizienten zwar die Verteilungsgleichheit der Grünflächenanteile quantifizieren, aber nicht in der Lage sind, wiederzugeben, welcher Anteil der Bevölkerung Zugang zu ausreichenden Grünflächenanteilen hat. Deshalb sind beide Informationen komplementär zur Beurteilung der Umweltgerechtigkeit von Grünflächen in deutschen Städten notwendig. Abbildung 6.17 zeigt auch, dass PGS-Anteile und GLC-Terzile nicht stark korrelieren. Entsprechend ist ein Rückschluss von der Verfügbarkeit grüner Landbedeckung auf die Zugänglichkeit von öffentlichem Grün nicht direkt möglich.

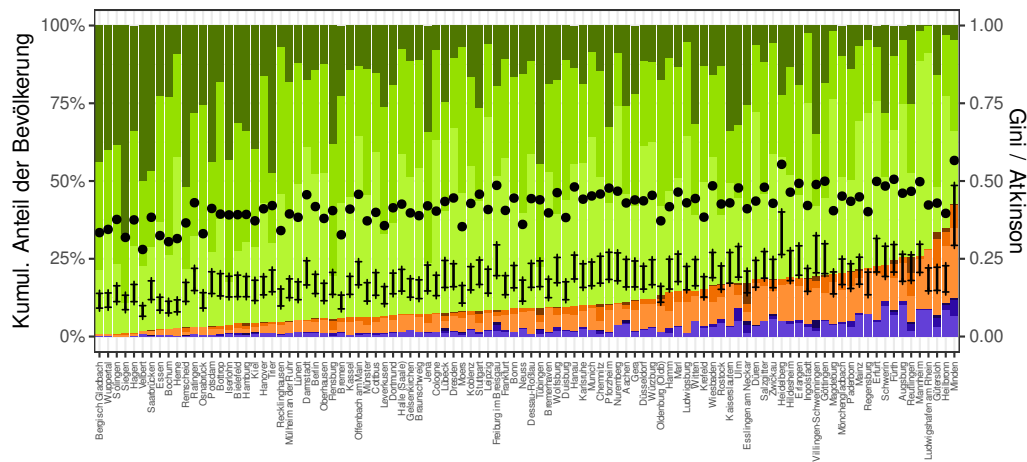


Abbildung 6.17: Vergleich der Grünflächenverfügbarkeit nach kumuliertem Bevölkerungsanteil (farbige Balken), PGS-Gini-Koeffizient (schwarze Punkte) sowie Atkinson-Indizes ('+' verbunden durch schwarze Linien, $\epsilon = 0,5$ oben, $\epsilon = 0,75$ unten) für die 100 bevölkerungsreichsten Städte in Deutschland. Die Farben folgen der Legende in der Abbildung 5.4. Die Städte sind nach dem Anteil der Bevölkerung geordnet, der mehr als 3,6 Hektar PGS in der Nachbarschaft hat, was dem WHO-Ziel entspricht (grün gefärbte Balken). Zu den wichtigsten Ergebnissen gehören: Die Bandbreite der Bevölkerung mit PGS-Werten über dem Zielwert variiert erheblich zwischen den größten Städten. Der PGS-Gini-Koeffizient korreliert nur schwach mit den PGS-Bevölkerungsanteilen über dem Zielwert ($r = -0,52$) mit deutlichen Ausreißern wie Darmstadt oder Heilbronn. Abbildung verändert nach Weigand et al. (2023b).

6.3.3 Verfügbarkeit von Grünflächen und Bevölkerungszusammensetzung

Die Beziehung zwischen der Verfügbarkeit von Grünflächen und der Bevölkerungszusammensetzung in Deutschland wird mithilfe von multivariaten linearen Modelle mit gemischten Effekten (LMM) untersucht. Diese ermöglichen globale Muster zu erkennen, ohne dass diese von lokalen Effekten, z. B. auf Stadtebene überlagert werden. Dies wird im Folgenden genutzt, um systemische Unterschiede der Verfügbarkeit von Grünflächen in Abhängigkeit zur demographischen Bevölkerungszusammensetzung auf Nachbarschaftsebene zu quantifizieren.

Als Vergleichsgrundlage dienen zunächst zwei reguläre lineare Modelle, welche ohne gemischte Effekte erstellt werden. In Tabelle 6.5 sind die Funktionsterme der beiden Modelle zusammengefasst, sowohl für grüne Landbedeckung (GLC, Spalte 2) als auch für öffentliches Grün (PGS, Spalte 3). Es wird deutlich, dass zwar statistische signifikante Zusammenhänge für alle Terme zu verzeichnen sind, die Modelle jedoch nur sehr geringe Bestimmtheitsmaße (R^2) zeigen. Diese Divergenz weist bereits darauf hin, dass die Modelle nicht alle Aspekte des Zusammenhangs zwischen Demographie und Grünflächenverfügbarkeit abbilden können. Als Erweiterung zu den obigen Basisexperimenten, werden im Folgenden LMMs verwendet, um lokale Effekte abbilden zu können.

Tabelle 6.5: Vergleichsergebnisse zweier multivariater linearer Regressionsmodelle ohne fixierte oder zufällige Effekte für Zusammenhang zwischen grüner Landbedeckung (GLC) bzw. öffentlichem Grün (PGS) und Demographie. Angegeben sind der y-Achsen Schnittpunkt (engl. *y-intercept*) sowie die Steigungskoeffizienten aller Variablen in Hektar. Die Ergebnisse dieser Modelle dienen als Basiswerte für die Interpretation der Ergebnisse der LMMs im vorliegenden Abschnitt. Signifikanzniveau ***p-Wert < 0,001.

	GLC	PGS
y-Schnittpunkt	61,765 ***	29,190 ***
log(Gesamtbev.)	-2,902 ***	-2,358 ***
Anteil Ausländer	-0,082 ***	0,178 ***
Anteil Kinder	0,115 ***	-0,123 ***
Anteil Senioren	0,096 ***	0,033 ***
Modell R^2	0,046	0,064
Anzahl Gitterzellen	174.126	174.121

In Tabelle 6.6 sind die verschiedenen Modelle zusammengefasst, die einerseits unterschiedliche Variablen mit fixiertem Effekt und andererseits Variablen mit zufälligem Effekt verwenden. Entlang der Spalten sind verschiedene Kombinationen von Variablen bzw. fixierten Effekten (FE) und zufälligen Effekten (RE) geführt, welche dazu dienen, die Magnitude der verschiedenen Einflüsse und Wechselwirkungen zu quantifizieren.

Erwartungsgemäß zeigen alle Modelle, dass die GLC mit zunehmender Bevölkerungszahl am stärksten abnimmt, was die Bedeutung der Kontrolle der Bevölkerungsdichte unterstreicht, die ein starker Proxy für die Bebauungsdichte ist. Ein Anstieg des Logarithmus zur Basis 10 der Bevölkerung um 1, d. h. ein absoluter Anstieg um das Zehnfache, ist mit einem durchschnittlichen Rückgang der GLC um etwa 2,5 ha verbunden (Standardfehler 0,025 ha).

In Bezug auf die demographischen Merkmale zeigen die fixierten Effekte der Modelle durchweg eine geringere Menge an GLC bei höherem Anteil von ausländischer Bevölkerung in einer Nachbarschaft, nämlich -0,195 ha bis -0,226 ha GLC pro 1 % Anstieg. Umgekehrt sind sowohl ein höherer Anteil an Kindern als auch ein höherer Anteil an älteren Menschen in der Nachbarschaft mit einem höheren Anteil an GLC verbunden, wenn auch in geringerem Umfang. Variieren die einbezogenen unabhängigen Variablen (im Vergleich der Modelle 1 - 4), ist festzustellen, dass die fixierten Effekte in allen Modellen ähnlich sind. Dies zeigt, dass die verschiedenen Variablen in der Tat disjunkte Effekte haben, was auch durch die allgemein niedrigen Korrelationen zwischen den demographischen Merkmalen der Nachbarschaft unterstützt wird.

Die Auswirkungen der verschiedenen Variablen mit Zufallseffekten verdeutlichen die Bedeutung der lokalen Geographie und ihren Einfluss auf die GLC-Exposition verschiedener Bevölkerungsgruppen. Dies wird durch den Unterschied der fixierten Effekte für die demographische

Tabelle 6.6: Zusammenhang von grüner Landbedeckung (GLC, in ha) und geschätzten fixierten Effekten (FE) demographischer Variablen, der Bevölkerung (log.), dem Anteil von Ausländern, Kindern und älteren Menschen auf Nachbarschaftsebene in Deutschland abhängig zu verschiedenen zufälligen Effekten (RE) je Gemeinde, Stadtgröße und Stadt-Land-Gradient. Signifikanzniveaue *** p-Wert < 0,001, ** p-Wert < 0,01. Tabelle verändert nach Weigand et al. (2023b).

	GLC (1)	GLC (2)	GLC (3)	GLC (4)	GLC (5)	GLC (6)	GLC (7)	GLC (8)	GLC (9)
y-Schnittpunkt FE	64,854 ***	65,266 ***	65,014 ***	64,708 ***	63,050 ***	62,989 ***	62,624 ***	59,470 ***	59,213 ***
log(Gesamtbev.) FE	-2,544 ***	-2,766 ***	-2,815 ***	-2,754 ***	-2,750 ***	-2,687 ***	-2,688 ***	-2,663 ***	-2,530 ***
Anteil Ausländer FE	-0,195 ***			-0,226 ***	-0,224 ***	-0,208 ***	-0,208 ***	-0,028 **	-0,003
Anteil Kinder FE		0,018 ***		0,033 ***	0,033 ***	0,030 ***	0,030 ***	0,102 ***	0,092 ***
Anteil Senioren FE			0,042 ***	0,041 ***	0,041 ***	0,040 ***	0,040 ***	0,069 ***	0,066 ***
Gemeinde RE	✓	✓	✓	✓	✓	✓	✓		
Stadtgrößenklasse RE					✓		✓	✓	
Stadt-Land-Gradient RE						✓	✓		✓
Modell R ²	0,618	0,621	0,621	0,624	0,624	0,626	0,626	0,050	0,053
Anzahl Gitterzellen	211.271	204.091	198.740	174.126	174.126	174.091	174.091	174.126	174.091

Zusammensetzung zwischen Modellen mit Zufallseffekten nur nach Stadtgrößenklasse oder Land-Stadt-Gefälle (Modelle 8 und 9) und Modellen mit Zufallseffekten für Gemeinden (Modelle 1 - 7) deutlich. Die fixierten Effekte der Demographie variieren weniger zwischen Modellen mit verschiedenen Variablen für die demographische Zusammensetzung als bei der Kontrolle für verschiedene Variablen mit Zufallseffekten. Interessanterweise weisen die Modelle 8 und 9 deutlich niedrigere Bestimmtheitsmaße auf. Darüber hinaus werden nur marginale Verbesserungen des R^2 -Wertes der Modelle festgestellt, wenn zusätzlich zu den fixierten Effekten der Gemeinden auch die zufälligen Effekte der Stadtgrößenklasse und des Land-Stadt-Gefälles kontrolliert werden (Modelle 5 - 7). Dies deutet erneut an, dass lokale Effekte auf Gemeindeebene sehr stark sind.

Analog zu GLC zeigt die Analyse des Zusammenhangs zwischen öffentlichem Grünanteil (PGS) und demographischen Merkmalen der Nachbarschaft (siehe Tabelle 6.7), dass öffentliches Grün am deutlichsten mit höherer Bevölkerungsdichte ($\log(\text{Gesamtbev.})$ FE) abnimmt. Im Gegensatz zur grünen Landbedeckung ist der y -Schnittpunkt für PGS etwa halb so groß. Dies ist intuitiv, da es im Allgemeinen weniger Hektar PGS als GLC gibt und erstere eine Untermenge der grünen Landbedeckung ist. Im Gegensatz zu GLC wird ein höherer Anteil an Kindern mit einer geringeren Menge an öffentlichem Grün assoziiert (-0,074 bis -0,0114), und Stadtteile mit einem höheren Anteil an Ausländern weisen eine höhere Verfügbarkeit von PGS auf (0,036 bis 0,048). Letzteres fällt auf, da es den signifikantesten und unerwartetsten Unterschied zwischen den GLC- und PGS-Modellen darstellt. Die fixierten Effekte für alle Variablen sind in allen Modellen ähnlich.

Auch hier spielen die lokalen Bedingungen in den einzelnen Städten und Gemeinden eine wichtige Rolle bei den Beziehungen zwischen PGS und den unabhängigen Variablen. Ähnlich wie bei GLC nimmt das Bestimmtheitsmaß R^2 ab, wenn die Zufallseffekte der Gemeinden in den Modellen 8 und 9 ausgeschlossen werden. Diese Modelle haben auch deutlich andere fixierte Effekte als die Modelle mit fixierten Effekten auf Gemeindeebene. Insgesamt zeigen die Modelle geringe Korrelationen zwischen den Variablen des Bevölkerungsanteils und die meisten FE und RE weisen eine hohe Signifikanz auf (p -Wert $< 0,001$ (***)).

Die Erstellung individueller Modelle für die Klassen des Land-Stadt-Gefälles mit Zufallseffekten pro Gemeinde (vgl. Tabellen 6.8 und 6.9) zeigt bemerkenswerte Unterschiede in der Größenordnung der fixierten Effekte zwischen den Land-Stadt-Klassen. Bei der GLC nimmt der fixierte Effekt der Bevölkerungsdichte vom ländlichen zum städtischen Raum um das Vierfache zu. Dies ist bei PGS nicht der Fall. Der Ausländeranteil steht in ländlichen Gebieten in einem positiven Zusammenhang mit der GLC (0,102 ***), in städtischen Gebieten jedoch in einem negativen (-0,327 ***). Von einigen Ausnahmen abgesehen, nehmen die absoluten Effektgrößen der demographischen Zusammensetzungen für GLC tendenziell in Richtung städtischerer Gebiete zu. Dieser Trend ist für PGS nicht so stark ausgeprägt. Gleichzeitig nimmt das R^2 der Modelle mit zunehmender Urbanität deutlich ab, und die statistische Signifikanz, gemessen an den zugehörigen p -Werten, ist im Vergleich zu den globalen Modellen teilweise geringer. Dies deutet darauf hin, dass die zugrundeliegenden systemischen Beziehungen zwischen den verschiedenen Urbanitätsgraden variieren.

Tabelle 6.7: Zusammenhang von öffentlichen Grünflächen (PGS, in ha) und geschätzten fixierten Effekten (FE) demographischer Variablen, der Bevölkerung (log.), dem Anteil von Ausländern, Kindern und älteren Menschen auf Nachbarschaftsebene in Deutschland abhängig zu verschiedenen zufälligen Effekten (RE) je Gemeinde, Stadtgröße und Stadt-Land-Gradient. Signifikanzniveaus ***p-Wert < 0,001. Tabelle verändert nach Weigand et al. (2023b).

	PGS (1)	PGS (2)	PGS (3)	PGS (4)	PGS (5)	PGS (6)	PGS (7)	PGS (8)	PGS (9)
y-Schnittpunkt FE	30,998 ***	31,533 ***	30,833 ***	31,127 ***	32,978 ***	31,835 ***	32,809 ***	30,775 ***	30,520 ***
log(Gesamtbev.) FE	-3,066 ***	-2,888 ***	-3,072 ***	-2,877 ***	-2,885 ***	-2,896 ***	-2,895 ***	-2,519 ***	-2,575 ***
Anteil Ausländer FE	0,048 ***			0,042 ***	0,039 ***	0,037 ***	0,036 ***	0,141 ***	0,128 ***
Anteil Kinder FE		-0,074 ***		-0,082 ***	-0,082 ***	-0,082 ***	-0,082 ***	-0,113 ***	-0,114 ***
Anteil Senioren FE			0,021 ***	0,018 ***	0,017 ***	0,018 ***	0,017 ***	0,035 ***	0,034 ***
Gemeinde RE	✓	✓	✓	✓	✓	✓	✓		
Stadtgrößenklasse RE					✓		✓	✓	
Stadt-Land-Gradient RE						✓	✓		✓
Modell R ²	0,456	0,459	0,457	0,455	0,455	0,455	0,455	0,067	0,069
Anzahl Gitterzellen	211.271	204.091	198.740	174.126	174.126	174.091	174.091	174.126	174.091

Tabelle 6.8: Zusammenhang von grüner Landbedeckung (GLC, in ha) und geschätzten fixierten Effekten (FE) demographischer Variablen getrennt nach Urbanität entlang der 5 Stadt-Land Klassen (vgl. Taubenböck et al., 2022). Signifikanzniveaus ***p-Wert < 0.001, ** p-Wert < 0.01, * p-Wert < 0.05. Tabelle verändert nach Weigand et al. (2023b).

	GLC (1)	GLC (2)	GLC (3)	GLC (4)	GLC (5)
y-Schnittpunkt FE	59,726 ***	63,079 ***	62,946 ***	64,989 ***	63,731 ***
log(Gesamtbev.) FE	-1,176 ***	-2,109 ***	-2,646 ***	-3,525 ***	-4,783 ***
Anteil Ausländer FE	0,102 ***	-0,054 ***	-0,137 ***	-0,185 ***	-0,327 ***
Anteil Kinder FE	-0,011	-0,022 *	0,011	0,029 *	0,206 ***
Anteil Senioren FE	0,017 **	0,015 **	0,052 ***	0,065 ***	0,255 ***
Stadt-Land Klasse	Rural	Vorw. rural	Übergang	Vorw. urban	Urban
Gemeinde RE	✓	✓	✓	✓	✓
Modell R^2	0,730	0,661	0,625	0,586	0,497
Anzahl Gitterzellen	28.768	64.168	41.665	22.836	16.654

Tabelle 6.9: Zusammenhang von öffentlichen Grünflächen (PGS, in ha) und geschätzten fixierten Effekten (FE) demographischer Variablen getrennt nach Urbanität entlang der 5 Stadt-Land Klassen (vgl. Taubenböck et al., 2022). Signifikanzniveaus ***p-Wert < 0,001, ** p-Wert < 0,01, * p-Wert < 0,05, · p-Wert < 0,1. Tabelle verändert nach Weigand et al. (2023b).

	PGS (1)	PGS (2)	PGS (3)	PGS (4)	PGS (5)
y-Schnittpunkt FE	29,683 ***	33,789 ***	30,997 ***	32,449 ***	30,577 ***
log(Gesamtbev.) FE	-2,958 ***	-3,336 ***	-2,640 ***	-2,538 ***	-2,474 ***
Anteil Ausländer FE	0,121 ***	0,086 ***	0,081 ***	0,000	-0,016
Anteil Kinder FE	-0,032 ***	-0,076 ***	-0,077 ***	-0,099 ***	-0,067 ***
Anteil Senioren FE	0,007	0,008 ·	0,038 ***	0,025 **	0,108 ***
Stadt-Land Klasse	Rural	Vorw. rural	Übergang	Vorw. urban	Urban
Gemeinde RE	✓	✓	✓	✓	✓
Modell R^2	0,622	0,483	0,433	0,463	0,361
Anzahl Gitterzellen	28.768	64.168	41.665	22.836	16.654

Diskussion

In diesem Kapitel werden die Ergebnisse der einzelnen konzeptionellen Schritte diskutiert und in den Kontext anderer wissenschaftlichen Publikationen eingeordnet.

7.1 Landbedeckungsklassifikation

Der im Rahmen dieser Dissertation gewählte Ansatz zur Klassifikation von hochaufgelösten, multispektralen Sentinel-2 Bilddaten in Kombination mit Referenzdaten, die aus dem LUCAS-Datensatz abgeleitet wurden, bietet verschiedene Erkenntnisse. Im Folgenden werden diese entlang der vier Teilaspekte der LUCAS Vorprozessierung und Positionierung, der Bildklassifikation und dem Übertrag auf einen weiteren Zeitraum diskutiert.

7.1.1 Auswirkungen der Vorprozessierung von LUCAS-Daten

Die Experimente des ersten konzeptionellen Schrittes dieser Dissertation untersuchen den Einsatz von Sentinel-2 Bilddaten, LUCAS-Referenzpunkten und maschinellem Lernen zur automatisierten Landbedeckungsklassifikation. Die Ergebnisse zeigen, dass die semantische Vorverarbeitung der LUCAS-Referenzdaten einen starken Einfluss auf die Genauigkeit der Klassifikationsergebnisse hat. Dennoch ist ein direkter Vergleich zwischen den Vorprozessierungsschemata ($S_0 - S_3$) nicht zulässig. Alle Schemata wurden für einen bestimmten Zweck der europaweiten oder nationalen LC-Kartierung entwickelt. Sie unterscheiden sich in der Auswahl der Stichprobe und der semantischen Klassenhierarchie und setzen so unterschiedliche Schwerpunkte für die Betrachtung des Raumes und der Landbedeckung bzw. Landnutzung.

Die Genauigkeit der einzelnen Schemata hängt dabei u. a. von der Anzahl der Klassen und Referenzstichproben ab, die zur Bewertung des Klassifikationsmodells verwendet werden. Im Allgemeinen erzielt ein Landbedeckungsprodukt mit weniger Zielklassen eine höhere Genauigkeit, eine Tendenz, die auch in anderen Studien beobachtet wurde (Ma et al., 2017). Dieser Effekt ist darauf zurückzuführen, dass die spektralen Signaturen in einem Experiment mit weniger Klassen klarer abgrenzbar sind. Klassifizierungssysteme mit mehr Zielklassen sind im Umkehrschluss anfälliger für Fehlklassifizierungen. Dennoch lässt sich in der Gesamtschau der 800 durchgeführten Klassifikationsexperimente sagen, dass der gewählte Random Forest Klassifikator für alle Vorpro-

zessierungsschemata hohe Genauigkeiten von 80 % und mehr erzielen konnte. Dies deckt sich mit umfassenden Ergebnissen zur Performanz dieses Klassifikators von z. B. Fernández-Delgado et al. (2014).

Von den drei selektiven Vorverarbeitungsschemata ($S_0 - S_3$) zeigen die Ergebnisse von S_2 die höchste Rate an Fehlklassifikationen. Dies ist unter anderem auf die große Anzahl von LC-Klassen (12) und die relativ wenigen verbleibenden Proben (ca. 14.600, vgl. Tab. 4.2) zurückzuführen. Dieses Schema wurde ursprünglich von Pflugmacher et al. (2019) für paneuropäische Klassifizierungszwecke entwickelt. In einer kleineren Region, im hier vorliegenden Fall Deutschland, sind wesentlich weniger Referenzpunkte für das Training des maschinellen Lernmodells verfügbar. Eine Einschränkung bei der Verwendung von LUCAS-Daten als Trainingsdaten für die Klassifizierung von Fernerkundungsbildern ist daher, dass sie nur auf regionaler oder nationaler Ebene erfolgreich eingesetzt werden können (Weigand et al., 2020). Je kleiner das zu untersuchende Gebiet ist, desto unwahrscheinlicher ist es, dass die LUCAS-Daten für die Erstellung eines stabilen Modells geeignet sind (siehe Leinenkugel et al., 2019). Daher wurden LUCAS-Daten in einigen Studien mit anderen Daten kombiniert (z. B. Griffiths et al., 2019), während in anderen Studien Proben mittels (engl.) *Crowd-Sourcing*, d. h. dem gemeinschaftlichen Erstellen von Referenzdaten, gesammelt wurden (Bayas et al., 2016).

Spannend wäre auch eine Ausweitung der Trainingsdaten auf den gesamten LUCAS-Datensatz. Moderne Technologien wie (engl.) *Spatio-Temporal Assets Cataloge (STAC)*³² in Kombination mit (engl.) *Cloud Optimized Geotiffs (COG)*³³ oder cloudbasierte Prozessierungsplattformen wie die Google Earth Engine (Gorelick et al., 2017) ermöglichen zunehmend den gezielteren Datenzugriff auf kleinste Ausschnitte in der im Gesamten beinahe unerschöpfliche Menge von Erdbeobachtungsdaten. Mithilfe dieser Technologien könnten umfangreiche Klassifikationsmodelle trainiert werden, auch wenn das eigentliche Studiengebiet eine zu geringe Anzahl an LUCAS-Punkten aufweist. Eine derartige Anwendung kann die mitunter prohibitiven Ressourcenanforderungen, die notwendig wären um Bilddaten für die gesamte EU vorzuhalten, zu umgehen. Unabhängig vom Einsatz von LUCAS-Daten werden diese technologischen Fortschritte bei der automatisierten Datensuche das Auffinden von geeigneten Bilddaten für Landbedeckungsklassifikationen zunehmend erleichtern.

Insgesamt liefern alle getesteten Vorprozessierungsschemata Landbedeckungsklassifikationen mit hoher semantischer Genauigkeit. Die Anwendung von S_1 und S_2 , die ursprünglich für Landsat Bilder mit geringerer Auflösung entwickelt wurden, zeigen ähnliche Genauigkeiten im Vergleich zu den jeweiligen Originalarbeiten (Mack et al., 2017; Pflugmacher et al., 2019). Dies unterstreicht die Übertragbarkeit der bestehenden Ansätze für die Klassifikation von Sentinel-2 Bildern. Mit S_3

³²STAC, ohne Datum, The STAC Specification, zuletzt abgerufen am 27.07.2023, <https://stacspec.org/en/about/stac-spec/>

³³GDAL, ohne Datum, „COG – Cloud Optimized GeoTIFF generator“, zuletzt abgerufen am 27.07.2023, <https://gdal.org/drivers/raster/cog.html>

haben Weigand et al. (2020) eine neue strukturierte Klassenhierarchie vorgestellt, die Landbedeckungsklassen beschreibt, die mit hoher semantischer Genauigkeit klassifiziert werden können und sich gut eignen, um Grünflächen auch im Siedlungskörper zu erkennen.

7.1.2 Positionierung von LUCAS-Stichproben

Die Ergebnisse der Gegenüberstellung verschiedener Positionierungsansätze unterstreichen die Tatsache, dass die Lokalisierung von LUCAS-Punkten für Fernerkundungsanalysen weder intuitiv noch trivial ist. Vor der Veröffentlichung des Datensatzes aus dem Jahr 2018 waren jedoch lediglich die GPS-Koordinaten im veröffentlichten LUCAS-Datensatz enthalten. *GRID*-Koordinaten mussten daher separat erhoben oder aus der LUCAS-ID rekonstruiert werden. *GPS*-Koordinaten, von denen aus die LUCAS-Punkte im Feld begutachtet wurden, hingegen werden zusammen mit den Daten bereitgestellt und erscheinen daher für den Benutzer leichter zugänglich. Auch wenn der Abstand zwischen den beiden Positionierungsansätzen *GPS* und *GRID* in den meisten Fällen nur wenige Meter beträgt, hat die gewählte Verortungsmethode bei allen Experimenten einen großen Einfluss auf die Genauigkeit des Klassifikationsergebnisses (siehe Abbildung 6.1). Es ist offensichtlich, dass die Wahl der theoretischen Position der Punkte (*GRID*) gegenüber der angegebenen *GPS*-Position die Klassifizierungsgenauigkeit unabhängig von S verbessert (Weigand et al., 2020). Diese ungerichtete Verschiebung führt dennoch zu messbaren Klassifikationsfehlern, ein Fakt der insbesondere bei hochauflösenden Bildern zum Tragen kommt. Die Stärke dieses Effekts hängt selbst nach Anwendung eines maximalen Beobachtungsabstands für *GPS*-Standorte von dem gewählten Vorverarbeitungsschema ab. Beispielsweise ist der Effekt zwischen *GPS* und *GRID* bei S_2 am geringsten, da hier bereits strenge Regeln für den Abstand zwischen dem Beobachter und dem beobachteten Punkt gelten. Der Fallstrick bei der Positionierung tritt am deutlichsten bei Stichproben der Klasse *Wasserflächen* (G) auf, bei denen sich die *GPS*-Koordinaten typischerweise am Ufer und nicht im Wasser befinden. Mit der ursprünglichen *GPS*-Position wird der spektrale Merkmalsraum beispielsweise durch Bewuchs am Ufer verzerrt. Der Ausschluss von Informationen über diese Gebiete auf der Grundlage der Beobachtungsdistanz (d. h. der Distanz zwischen dem *GPS*-Punkt und dem ursprünglichen *GRID*-Punkt) ignoriert jedoch die Tatsache, dass die Informationen verlässlich vom Ufer aus registriert werden können, auch wenn die Distanz besonders groß sein kann. Da der Positionierungsansatz einen großen Einfluss auf die Klassifizierungsgenauigkeit bei hochauflösenden Fernerkundungsanwendungen hat, ist zu empfehlen stets die *GRID*-Koordinaten der LUCAS-Erhebungen zu rekonstruieren. Seit 2018 sind diese auch in den Rohdaten enthalten als sogenannte „theoretische Koordinaten“ (engl. *theoretical point*). Auf diese Weise muss keine Auswahl von Stichproben auf der Grundlage einer maximalen Beobachtungsentfernung vorgenommen werden, und es können mehr Referenzpunkte für den Klassifizierungsprozess zur Verfügung gestellt werden.

7.1.3 Bildklassifikation und Klassifikationsgenauigkeit

Die Kombination von in-situ Daten aus der LUCAS-Datenbank mit Satellitenbildern kann zu mehreren häufigen Fehlern führen. So werden beispielsweise Effekte wie schattige Bereiche, die im Bild dargestellt sind, in der LUCAS-Klassenhierarchie nicht antizipiert. Diese können, insbesondere bei zunehmender räumlicher Bildauflösung, städtische Gebiete mit vielen hohen Gebäuden und Straßenschluchten aufgrund ihrer spektralen Ähnlichkeit fälschlicherweise als Wasserflächen eingestuft werden (Wieland et al., 2019). Darüber hinaus sind schmale, von Bäumen überwachsene Straßen, die für einen menschlichen Beobachter am Boden deutlich erkennbar sind, aus der Vogelperspektive nicht als solche zu erkennen. Diese werden daher mit hoher Wahrscheinlichkeit fälschlicherweise einer Vegetationsklasse, wie beispielsweise *hohe saisonale Vegetation* (*T1*), also Laubbäumen, zugeordnet.

Klassen, die in geringerer Zahl in der Stichprobe vertreten sind, wie z. B. *offener Boden* (*S*), sind anfällig für eine größere Anzahl von Klassifikationsfehlern. Die Fehlermatrizen zeigen, dass diese spezielle Klasse deutlich häufiger falsch klassifiziert wurde als die anderen Landbedeckungsklassen in Experimenten. Dies könnte neben einer geringeren Stichprobengröße auf eine spektrale Ähnlichkeit mit anderen Klassen zurückzuführen sein, wie z. B. Baumaterialien, die der Klasse *bebautes Land* (*A*) zugeordnet sind.

In diesem methodischen Aufbau wurden insgesamt 229 Spektral-, Textur- und zusätzliche Bildebene für die Bildklassifizierung verwendet. Eigentlich sollten spektrale Indizes aus bodennahen Reflektanzen (engl. *at-surface reflectances*) abgeleitet werden anstatt aus sensornahen Reflektanzen (engl. *at-sensor reflectances* Jones und Vaughan, 2010, S. 167). Im Fall von Sentinel-2 entspricht das Level-2A Produkten, welche die Oberflächenreflektanz abbilden³⁴. Diese Daten standen zum Zeitpunkt der Datenerhebung noch nicht flächendeckend im nötigen Umfang zur Verfügung, sollten aber in zukünftigen Anwendungen verwendet werden, um die Datenqualität weiter zu steigern.

Im Vergleich zu vorherigen, Landsat-basierten Studien (z. B. Mack et al., 2017; Pflugmacher et al., 2019; Leinenkugel et al., 2019) wurden mehr Bildmerkmale gerechnet. Dies war möglich, da durch die höhere räumliche Auflösung von Sentinel-2 von 10 m × 10 m so auch kleinräumigere Texturen erfasst werden konnten. Insgesamt wurde dadurch ein Bilddatensatz mit einem komprimierten Datenvolumen von 2,5 Terabyte erstellt. Obwohl diese Ergebnisse vielversprechend sind, sind sie mit einem hohen Rechenaufwand verbunden und erfordern geoinformatische Fachkenntnisse in der Analyse großer geographischer Daten. In zukünftigen Anwendungen könnten daher Techniken zur gezielten Auswahl der relevantesten Bildmerkmale angewandt werden, um den Rechenaufwand und notwendigen Speicherplatz zu verringern.

³⁴Sentinel Online, Level-2A User Guide, zuletzt aufgerufen am 01.05.2023: <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/product-types/level-2a>

Die Anwendung der verschiedenen Vorverarbeitungsschemata für die nationale Landbedeckungsklassifikation im Rahmen dieser Arbeit zeigt, dass die LUCAS-Stichprobendaten eine geeignete Datenbasis für hochauflösende Landbedeckungskartierungen mit Sentinel-2 Daten darstellen. Maschinelle Lernverfahren wie Random Forest sind in der Lage die komplexen Zusammenhänge zwischen Bildmerkmalen und verschiedenen Landbedeckungsklassen gut zu erfassen und diese voneinander zu unterscheiden. Landschaftsmerkmale können so auf lokaler bis nationaler Ebene mit hoher Genauigkeit identifiziert werden. Kleinräumigere Analysen, wie beispielsweise die Abgrenzung individueller Bäume im urbanen Kontext erfordern jedoch weiterhin sehr hoch aufgelöste Bilddaten (z. B. Leichtle et al., 2021), wenngleich bekannt ist, dass auch der städtische Baumbestand ungleich verteilt sein kann (Baró et al., 2019).

Die endgültige Landbedeckungskarte bietet eine beispiellose Grundlage für künftige Anwendungen, die von der hohen räumlichen Auflösung profitieren, wie z. B. interdisziplinäre Analysen menschlicher Lebensräume für Umweltgerechtigkeit (vgl. Weigand et al., 2019). Bis jetzt wurde dieser Datensatz bereits für verschiedenste Studien eingesetzt. Beispielsweise untersuchten (Tennikl et al., 2019) und Taubenböck et al. (2021) Grünflächen in deutschen Großstädten. Staab et al. (2022a,b) nutzten die Landbedeckungseigenschaften zur Modellierung der Lärm-Umwelt-Interaktion in einer sogenannten Landnutzungsregression (engl. *land use regression*). Rutschmann (2023) nutzte die Klassifikation zur Beschreibung von Landschaftseigenschaften im Kontext der Analyse vom Verhalten von Honigbienen. Dies zeigt ein weites Anwendungsspektrum der Daten über die Beschreibung von Grünflächen im Rahmen dieser Dissertation hinaus.

7.1.4 Übertrag auf 2018

Die Anwendung des entwickelten Ansatzes aus der Vorprozessierung der Referenzdaten sowie der Übertrag auf einen weiteren Zeitraum zeigt verschiedene Aspekte eindrucksvoll. Einerseits wird deutlich, dass die Auswahl der Bilddaten aus einem Jahreszeitraum – und damit einer Vegetationsperiode – ausreicht, um aussagekräftige Bildmerkmale zu aggregieren und eine Landbedeckungsklassifikation erfolgreich durchzuführen. Sentinel-2 bietet, zumindest für das hier gewählte Untersuchungsgebiet, ausreichend Bilddaten, aus denen Quantilsaggregate für die Spektralen Indizes NDVI, NDBI und NDWI extrahiert werden können, welche zu einer hochgenauen Klassifikation beitragen.

Übergeordnete Dynamiken im Landschaftsbild, wie am Beispiel von großräumigen Baumaßnahmen in Würzburg demonstriert werden konnte, werden durch den Vergleich der beiden Klassifikationsergebnisse gut erfasst. Trotz ähnlich hoher Genauigkeiten der beiden Klassifikationsergebnisse zeigen, gestaltet sich eine pixelbasierte Veränderungsanalyse (engl. *change detection*) jedoch schwierig. Die Modelle der beiden Klassifikationsprodukte betonen verschiedene Aspekte der Umwelt auf unterschiedliche Weise. Dafür gibt es zwei mögliche Gründe: Einerseits wirkt sich die Zeitspanne, in der die Bilddaten aufgenommen wurden auf die Bildmerkmale aus. Im ersten Fall (2015-2017) wurden mehrere Vegetationsperioden einbezogen, im zweiten Modell (2018) lediglich eine. Andererseits basieren die beiden Modelle auf unterschiedlichen Stichproben aus der

LUCAS-Erhebung, da für jede der dreijährigen Erhebungen eine andere Zufallsstichprobe aus der Grundgesamtheit von ca. 1,1 Million Punkten in Europa ausgewählt wird. Die Lage dieser Punkte bestimmt schlussendlich, welche Landbedeckungseigenschaften im Modell abstrahiert und gelernt werden können. Um die Vergleichbarkeit der Produkte zu erhöhen, müssten weitere Experimente mit synchronisierten Stichprobensets durchgeführt werden. Hierbei könnte der zuvor genannte STAC/COG-basierte Datenbezug helfen, um effizient Bilddaten für synchronisierte Stichproben von LUCAS zu erhalten.

7.2 Modellierung des öffentlichen Grünanteils mit Deep Learning

Urbanes Grün ist aufgrund der vielfachen Vorteile für die Menschen wichtiger Gegenstand moderner Umweltgerechtigkeits- bzw. Umweltgesundheitsforschung. Häufig werden zur Bestimmung des Grünanteils in der Wohnumgebung der Menschen alle Typen von Vegetationsflächen betrachtet und beispielsweise unter dem Begriff der „Grünheit der Nachbarschaft“ (engl. *neighborhood greenness*) zusammengefasst (z. B. Reid et al., 2018). Zunehmend mehr geographische Datenprodukte aus Satellitendaten, entweder aus Vegetationsindizes (z. B. NDVI) oder Landbedeckungskartierungen, ermöglichen derartige Grünflächenverfügbarkeit großflächig und hoch aufgelöst zu bestimmen. Eine zu allgemeine Betrachtung von Grünflächen in Form von beispielsweise Landbedeckungsinformationen erlauben jedoch keine separate Beurteilung der Verteilung verschiedener Typen urbaner Grünflächen oder auch deren gesundheitlichen Einflusses.

Vor diesem Hintergrund war es das Ziel des zweiten konzeptionellen Schrittes dieser Dissertation, moderne maschinelle Lernverfahren zu nutzen, um die Verfügbarkeit von öffentlichen Grünanlagen, als ein Beispiel eines bestimmten Typs von Grünflächen, auf Nachbarschaftsebene flächendeckend für Deutschland bestimmen zu können. Hierfür wurde aufgebaut auf bestehende Datensätze des European Urban Atlas, die auch bisher im Rahmen von Analysen der Umweltgerechtigkeit von urbanem Grün eingesetzt wurden, jedoch nur Teile des Untersuchungsgebietes abdecken. Das hier vorgestellte Konzept nutzt moderne Methoden des maschinellen Lernens, um derartige Informationen über den nachbarschaftlichen öffentlichen Grünanteil für das gesamte Untersuchungsgebiet zu quantifizieren. Dies erschließt die Aufnahme von knapp 20 Millionen Einwohnerinnen und Einwohnern Deutschlands, knapp 25 % der Gesamtbevölkerung, welche zuvor nicht für derartige Untersuchungen erfasst waren.

Für die Modellierung der Regression der öffentlichen Grünfläche auf Nachbarschaftsebene wird im Rahmen dieser Arbeit ein Fusionsansatz aus Satellitenbild- und OpenStreetMap-Daten gewählt, welche in Form eines neuronalen Netzwerks simultan verarbeitet werden. Die gewählte Datenfusionsstrategie zielte darauf ab, den heterogenen Charakter von OSM-Daten mit Bilddaten in einem modernen datengetriebenen Verfahren zu vereinen. Die entwickelte Methodik unterscheidet sich daher von anderen Arbeiten, die beispielsweise einen rein CNN-basierten Ansatz wählten und OSM-Daten rasterisieren (Audebert et al., 2017; Henry et al., 2021; Doda et al., 2022). Im Gegensatz zu ähnlichen Ansätzen in der Literatur, welche ebenfalls Daten aus

Satellitenbildern und OSM nutzen (z. B. Rosier et al., 2022; Ludwig et al., 2021), wurden die Vektordaten im hier gewählten methodischen Aufbau aus OSM zunächst auf Nachbarschaftsebene aggregiert und in Form eines Merkmalsvektors einem ANN zugeführt. Ähnliche Aggregate wurden zuvor von Tuia et al. (2017) bereits in einem Random Forest-basierten Klassifikationsverfahren zur Landnutzungsklassifikation eingesetzt. Dieser Entscheidung liegt der Versuch zugrunde, die Gesamtkomplexität des Fusionsnetzwerkes möglichst zu reduzieren, um eine großflächige, deutschlandweite Anwendung zu einem akzeptablen Rechenaufwand zu ermöglichen.

Die hier präsentierten exemplarischen Ergebnisse des Modelltrainings sowie der Performanzmetriken zeigen, dass das Fusionsmodell in der Lage ist, die nötigen Zusammenhänge aus den Datensätzen zu extrahieren und die Fläche öffentlichen Grüns auf Nachbarschaftsebene mit hohem Bestimmtheitsmaß zu schätzen. Die Ergebnisse zeigen eine klare Konvergenz des Modells zum Optimalwert. Das hohe Bestimmtheitsmaß zeigt jedoch nur einen Teil der Modellperformanz. Die Betrachtung der Residuen zeigt, dass selten auch sehr hohe Abweichungen zwischen dem Schätzwert und der Referenz auftreten. Fehlschätzungen des Modells werden zu ähnlichen Teilen als Über- wie Unterschätzungen ohne erkennbare Muster getroffen. Die Analyse der Residuen zeigt ebenfalls keine Systematik in Abhängigkeit zur Distanz zu Bebauung, was eine höhere Zahl von Fehleinschätzungen in dicht besiedelten Gebieten ausschließen lässt. Die räumliche Verteilung der Residuen des Netzwerkes zeigt, dass keine räumlichen Muster der Über- und Unterschätzungen erkennbar sind.

Bezogen auf den methodischen Ansatz der tiefen neuronalen Netzwerke bietet sich nach diesem Machbarkeitsbeweis noch Optimierungspotenzial. Beispielsweise wäre denkbar eine andere Aktivierungsfunktion als die lineare Aktivierung für die Ausgabebene zu wählen. Denkbar wäre beispielsweise eine ReLU-Aktivierung, da für die Zielvariable im hier vorliegenden Fall nur positive Werte geschätzt werden sollen (Géron, 2019, S. 293). Im Rahmen der Datenarbeit wurden hier die Hyperparameter des Modells manuell ermittelt und unterliegen so dem Einfluss der Auswahl eines Experten. Die automatisierte Suche nach den optimalen Hyperparametern könnte hier helfen die statistisch optimalen Parameter zu ermitteln. Neben der manuellen Suche nach Hyperparametern bieten sich verschiedene automatisierte Verfahren an (Géron, 2019, S. 320; Goodfellow et al., 2016, S. 420 ff.): systematische Suche (engl. *grid search*), randomisierte Suche (engl. *randomized search*) oder modellbasierte Optimierung. Automatisierte Hyperparameteroptimierung erfordert jedoch beträchtliche Rechnerressourcen und es ist daher ratsam diese nur durchzuführen, wenn die finale Architektur des Modells feststeht. Die im Kontext der Trainingsdatenvorbereitung von neuronalen Netzwerken bisweilen verwendete Augmentierung (Lecun et al., 1998; Krizhevsky et al., 2012; Goodfellow et al., 2016, S. 445), d. h. Transformation und Variation von z. B. Bilddaten durch leichte Variation von Farben oder Rotation, wurde hier nicht angewendet, da bereits sehr viele Referenznachbarschaften zur Verfügung standen. Dennoch könnte diese Art der Datenaufbereitung genutzt werden, insbesondere wenn kleinere Datensätze zur Verfügung stehen, um die teilweise prohibitiven Kosten von ausreichend Referenzdaten zu senken (Ma et al., 2019).

Das im Rahmen dieser Arbeit vorgestellte Fusionsmodell zur Quantifizierung von öffentlichen Grünanlagen kann als Machbarkeitsbeweis angesehen werden, dass die Fusion von heterogenen Datensätzen aus Satelliten- und OSM-Daten mittels eines CNN- und ANN-basierten Fusionsnetzwerkes in der Lage ist stabile Schätzwerte zu liefern. Speziell vor dem Hintergrund der geringen Margen, welche im Rahmen der WHO-Zielvorgaben für öffentliches Grün in der Wohnumgebung herrschen (World Health Organization, 2017), sollten zukünftige Modellansätze auf die Minimierung der mittleren Fehler vornehmlich in Gebieten mit sehr geringer Grünflächenverfügbarkeit geachtet werden.

Eine derartige Verwendung von neuronalen Netzen im Kontext der Quantifizierung von urbanen Grünflächen bisher nicht bekannt. Nach dem hier geführten Machbarkeitsnachweis des Fusionsansatzes zur Ableitung von Landnutzungseigenschaften sind Erweiterungen denkbar. Die Ausweitung auf weitere oder mehrere Typen von Grün ist denkbar und müsste in nachfolgenden Studien untersucht werden. Auch die Ausweitung zur pixelbasierten Kartierung von öffentlichen Grünflächen mithilfe von CNN-basierter semantischer Segmentierung ist eine interessante Fortführung des hier vorgestellten Ansatzes. Diese bietet den Vorteil, dass die über die reine Verfügbarkeitsmetrik hinaus die Erreichbarkeit mithilfe von Routing innerhalb der individuellen Nachbarschaft ermittelt werden könnte, zum Beispiel in fußläufiger Reichweite (Wüstemann et al., 2017a; Droin et al., 2023). Darüber hinaus könnten gegebenenfalls weitere Aspekte von Grünflächen erfasst werden, da bekannt ist, dass Ausstattung der Grünflächen hat einen Einfluss auf die Akzeptanz bzw. die Zufriedenheit der Menschen über die Grünflächen (Wang et al., 2021). Die hier vorgestellte Methodik zur Quantifizierung von öffentlichem Grün kann perspektivisch genutzt werden, um Nachbarschaften mit Defiziten an öffentlichem Grün zu identifizieren. Bezogen auf Planungen zur Begrünung der Städte sollten insbesondere stark exponierte Flächen priorisiert werden, da die Effektivität hinsichtlich der kühlenden Wirkung besonders hoch sind (Massaro et al., 2023).

Die entwickelte Methodik baut fundamental auf den Daten des OpenStreetMap-Projekts auf. Diese Daten, deren Potenzial für geographische Forschung schon früh erkannt wurde (Goodchild, 2007; Flanagin und Metzger, 2008), stellen eine wertvolle Quelle von hochaufgelösten und semantisch breit gefächerten räumlichen Informationen dar. Dennoch haben derartige von Freiwilligen gesammelten Geodaten ganz besondere Eigenheiten, welche die Verlässlichkeit der Ergebnisse beeinflussen können. Zwar kann eine hohe Genauigkeit gemessen werden, doch die Vollständigkeit der Landbedeckungs- und Landnutzungsinformationen in der OSM-Datenbank variiert auf globaler Ebene stark (Zhou et al., 2022). Am Beispiel von Grünflächen konnte bereits gezeigt werden, dass private Grünflächen weniger häufig kartiert werden als öffentliche Grünanlagen (Texier et al., 2018; Ludwig et al., 2021). Um unerkannte Datenlücken zu kompensieren wurde im hier gewählten methodischen Aufbau die Datenfusion mit Satellitendaten gewählt. Auch gewisse Inkonsistenzen von Attributdaten im Bezug auf die semantische Markierung von Grünflächenobjekten in OSM (Ludwig et al., 2019) werden im vorgestellten Ansatz durch die Gruppierung ähnlicher Elemente (vgl. Tabelle A2) adressiert. Es wäre ebenfalls vorstellbar, Grünflächen weiter zu differenzieren

durch die Deep Learning-basierte Fusion von multimodalen Bilddaten aus Satellitenbildern und Bodenansichten von Google Street View zur Differenzierung urbaner Landnutzung (Srivastava et al., 2019).

7.3 Gerechtigkeit der räumlichen Verteilung von Grünflächen

Im dritten konzeptionellen Schritt dieser Dissertation wurden die zuvor erstellten Datenprodukte zu Landbedeckung und öffentliche Grünflächen kombiniert, um diese in eine Analyse zur Verteilung von Grünflächen unter der Bevölkerung Deutschlands zu überführen. Diese stellt die erste Anwendung auf nationalem Level dar bei gleichzeitig hochaufgelöstem räumlichen Detailgrad, welcher die Wohnumgebung aller Einwohnerinnen und Einwohner beinhaltet.

7.3.1 Grünflächenverfügbarkeit in Deutschland

In dieser Arbeit wurde die Verfügbarkeit von Grünflächen und öffentlich zugänglichen Grünflächen in den Wohngebieten der Menschen in Deutschland quantifiziert. In Übereinstimmung mit der bestehenden Literatur (z. B. Voorde, 2017) unterstreicht eine schwache Korrelation zwischen PGS und GLC ($r = 0,617$) die Bedeutung der Analyse von verschiedenen Grüntypen. Frühere Studien mit hoher räumlicher Auflösung waren auf einzelne Städte oder EUA-Gebiete beschränkt (z. B. Kabisch et al., 2016; Zepp et al., 2020); in dieser Studie wurde der Umfang erweitert, indem die Verfügbarkeit von öffentlichem Grün mithilfe eines Ansatzes aus Datenfusion mit Deep Learning modelliert wurde. Die in dieser Dissertation entwickelte Methode zur Datenerfassung half, die räumliche Begrenzung der Daten des European Urban Atlas zu umgehen, und ermöglichte eine Analyse der öffentlichen Grünflächen unabhängig von willkürlichen Datengrenzen.

Die bundesweite Analyse zeigte, dass die Mehrheit der deutschen Bevölkerung mehr öffentliche Grünflächen in der Wohnumgebung vorfindet als das von der WHO definierte Minimum. Vor allem in ländlichen Regionen sinkt der Anteil der Menschen, die diese Ziele erreichen, jedoch deutlich. Auch zwischen verschiedenen Städten fanden sich große Unterschiede. Dies ergänzt die Ergebnisse von Taubenböck et al. (2021), in denen betont wird, dass die Vielfalt der lokalen Landschaft eine wichtige Rolle für den Umfang der Grünflächen in einer Stadt spielt. Die vorliegende Arbeit fügt jedoch zwei Perspektiven hinzu: Erstens bewertet sie verschiedene Arten von Grün, insbesondere öffentlich zugängliche Grünflächen. Und zweitens führt sie eine hochauflösende, nationale Analyse der verfügbaren Grünflächen auf Nachbarschaftsebene durch, die die gesamte Bevölkerung in Deutschland einschließt.

Die Analyse ergibt signifikante Unterschiede zwischen der Verfügbarkeit von Grünflächen in städtischen und ländlichen Regionen, sowohl in Bezug auf die verfügbare Menge als auch auf die Verteilungsgerechtigkeit unter der Bevölkerung. Die Ergebnisse zeigen, dass öffentliches Grün vor allem in dicht besiedelten Regionen eine wichtige Ressource darstellt. Aber auch in Städten ist ein hoher Anteil an GLC nicht unbedingt mit einem hohen Anteil an PGS gleichzusetzen.

In Stadtvierteln mit überwiegend Ein- und Zweifamilienhäusern ist die Verfügbarkeit von GLC zwar höher, die Verfügbarkeit von PGS ist jedoch immer noch geringer als in dichter besiedelten Gebieten mit Mehrfamilienhäusern.

Ebenso ist es in ländlichen Regionen wahrscheinlicher, dass die Verfügbarkeit von öffentlichen Grünflächen unter dem Zielwert liegt. Ähnliche Unterschiede sind auch bei anderen gesundheitsrelevanten Einrichtungen zu erwarten (Pearce et al., 2006). Darüber hinaus handelt es sich bei öffentlichen Grünflächen in ländlichen Regionen eher um natürliche Umgebungen, die nachweislich einen größeren positiven Einfluss auf die Gesundheit haben (Wheeler et al., 2015). Welche Grünflächen in der Nachbarschaft zu finden sind, hängt stark von der individuellen Wohnumgebung ab. In ländlichen Gebieten wohnen die Menschen womöglich nicht in der Nähe eines Parks, können aber auch in den Feldern und nahe gelegenen Wäldern spazieren gehen. Die genaue Zusammensetzung von Grünflächen, z. B. auch landwirtschaftlich genutzter Flächen, beeinflusst den gesundheitlichen Einfluss auf unterschiedliche Weise. Es ist daher notwendig, die Wege in Abhängigkeit von der Urbanität zu analysieren und die möglicherweise unterschiedlichen gesundheitlichen Auswirkungen salutogener Ressourcen in verschiedenen Regimen zu bewerten. Schließlich haben frühere Studien gezeigt, dass der Zugang zu und die Exposition gegenüber Grünflächen nur schwach korreliert sind (Jarvis et al., 2020). Die Studie bewertet nicht die objektive und subjektive Bedeutung von öffentlich zugänglichen Grünflächen, sondern liefert eine Datengrundlage und einen methodischen Ansatz zur Unterstützung solcher Studien; künftige Forschungen sollten untersuchen, wie die ländliche Bevölkerung Grünflächen im Allgemeinen und PGS im Speziellen nutzt und bewertet und wie sich dies von Stadtbewohnern unterscheidet.

7.3.2 Verteilungsgerechtigkeit und Einfluss der Demographie auf Grünflächenverfügbarkeit

Das wichtigste Ergebnis dieser Studie ist, dass es einen signifikanten Unterschied in der Verteilungsgerechtigkeit gibt, je nachdem, welche Art von Grünflächen berücksichtigt wird. Bei allen Experimenten mit Ausnahme von Gebieten mit Mehrfamilienhäusern war die Verfügbarkeit von GLC gleichmäßiger über die Bevölkerung verteilt als die von PGS. Dies deutet darauf hin, dass die Ressource öffentliches Grün in der deutschen Bevölkerung konzentrierter und damit unverhältnismäßiger verteilt ist als GLC. Wie von Voorde (2017) erörtert, unterstreicht dies die Relevanz der Einbeziehung der Art und Zugänglichkeit von Grünflächen in die Stadtplanung, da GLC als Referenzmaß für die Grünbedeckung die Verfügbarkeit von nutzbaren Grünflächen überbewerten und den unverhältnismäßigen Zugang für bestimmte Bevölkerungsgruppen unterschätzen könnte.

Darüber hinaus unterstreichen die divergierenden Trends der Gini-Koeffizienten zwischen den beiden Grünflächenarten die starken Unterschiede zwischen ländlichen und städtischen Gebieten. Dies verdeutlicht die Vielfalt der Grünflächenverfügbarkeit und die Notwendigkeit einer lokalisierten Bewertung der Grünflächenressourcen. Regionale Unterschiede sind stark von den lokalen geographischen Gegebenheiten wie dem Grad der Verstädterung abhängig. Die Studie unterstreicht auch, dass die Ungleichheit bei den öffentlichen Grünflächen in ländlichen Regionen

besonders groß ist. Dort wird die Ungleichheit der Verfügbarkeit von öffentlichem Grün zum einen durch private Gärten und landwirtschaftliche Flächen und zum anderen durch Häuser in der Nähe größerer Wälder verstärkt.

Der Vergleich von Gini- bzw. Atkinson-Koeffizienten zwischen verschiedenen Bevölkerungsgruppen muss jedoch mit Vorsicht interpretiert werden. Diese Gerechtigkeitsmaße beziehen sich auf unterschiedliche Gesamtmengen der verfügbaren Ressourcen. Sie beschreiben lediglich die Verteilung der verfügbaren Menge. Der Vergleich für die 100 bevölkerungsreichsten Städte in Deutschland hat gezeigt, dass die Gini- und Atkinson-Koeffizienten beispielsweise keine direkten Rückschlüsse darauf erlaubt, welcher Anteil der Bevölkerung mit öffentlichem Grün gemessen am WHO-Schwellwert unterversorgt sein könnte. Auch außerhalb der städtischen Gebiete nimmt die Bedeutung des öffentlichen Grüns als Hauptquelle für Grünflächen in der Umgebung abnehmen. Während der Park in der Stadt die wichtigste Quelle für öffentliches Grün sein könnte, sind in ländlichen Gebieten andere und vielfältigere Formen von Grün, wie Wiesen oder Felder, für die Menschen relevanter. Die getroffenen Annahmen im Datenmodell, beispielsweise durch die Auswahl der EUA-Klassen *Green urban areas*, *Forests* und *Herbaceous vegetation associations*, sind nicht in der Lage diese subjektiven Unterschiede in der Präferenz der Anwohnerinnen und Anwohner widerzuspiegeln.

Die höhere Verfügbarkeit von ausreichend PGS in größeren Städten kann auch als Zeugnis einer gezielten und erfolgreichen Stadtplanung gesehen werden. Dort, wo aufgrund zunehmender Bebauungsdichte von vornherein weniger grüne Landbedeckung vorhanden ist, ist die Bedeutung der Bereitstellung von öffentlichem Grün für die Stadtbevölkerung besonders hoch und gleichsam wichtig. Dort, wo nicht jede Anwohnerin und jeder Anwohner Platz für einen eigenen Garten hat, ist der Anteil der Menschen, die Zugang zu einem ausreichenden PGS haben, höher.

Die Beziehung zwischen der demographischen Zusammensetzung und der Verfügbarkeit von Grünflächen in der Nachbarschaft befasst sich mit Ungleichheiten in Bezug auf verschiedene demographische Merkmale. Die Bevölkerungsdichte hatte insgesamt den größten Einfluss auf die Verfügbarkeit von GLC und PGS, was die Bedeutung der Kontrolle dieser Variable unterstreicht. Während ein zunehmender Anteil ausländischer Bevölkerung in der Nachbarschaft mit geringeren GLC-Flächen verbunden ist, ist dieser Trend bei PGS umgekehrt. Dies steht im Einklang mit der Tatsache, dass die Mehrheit der ausländischen Bevölkerung in Städten lebt. Ein ähnlicher Zusammenhang wurde von Barbosa et al. (2007) für das Vereinigte Königreich festgestellt. Der Unterschied in der Größenordnung zwischen den beiden Effekten zeigt jedoch, dass die Verringerung der GLC nicht vollständig durch einen gleichwertigen Anstieg der PGS kompensiert wird. Demnach sind Stadtteile mit einem höheren Anteil an Zuwanderern in Bezug auf die Vorteile von GLC, z. B. Umgebungskühlung (Park et al., 2017), generell stärker benachteiligt. Außerdem wird deutlich, dass PGS in diesen Vierteln eine besonders wichtige Quelle für Grünflächen ist.

Im Gegensatz dazu waren die Beziehungen zwischen der Verfügbarkeit von GLC und Stadtteilen mit einem höheren Anteil an Kindern sowie einem höheren Anteil an älteren Menschen weit weniger ausgeprägt. Dies kann als Hinweis darauf gedeutet werden, dass die systemische Ungleichheit in Stadtteilen mit einem höheren Anteil ausländischer Bevölkerung stärker ausgeprägt ist. Die geringe positive Assoziation zwischen dem Anteil der Kinder und der GLC und die geringe negative Assoziation mit der PGS kann ein Hinweis auf die Suburbanisierungsbewegung junger Familien sein, die in die Vororte mit einem höheren Anteil an privatem Grün im Gegensatz zu öffentlichen Parks ziehen.

Die Magnituden der fixierten Effekte sind über die Experimente unterschiedlicher Zusammensetzungen stabil, was auf unabhängige Effekte der einzelnen Bevölkerungsanteile hindeutet. Das Bestimmtheitsmaß R^2 ist am höchsten, wenn für zufällige Effekte auf Gemeindeebene kontrolliert wird. Dies deutet darauf hin, dass lokale Phänomene tatsächlich einen starken Einfluss auf die Beziehung zwischen der Verfügbarkeit von Grünflächen und der Bevölkerungszusammensetzung haben. Die Achsen und Effekte der Bevölkerung blieben über verschiedene Variablenkombinationen hinweg stabil, was auf eine starke Unabhängigkeit von Bevölkerungsanteilen und Variablen mit zufälligen Effekten hindeutet.

7.3.3 Limitierungen und weitere Forschung

Bestimmt durch den Input in das PGS-Modell werden hier verfügbare Grünflächen geschätzt für rechteckige Nachbarschaften, die einen Puffer von 500 m umschließen. Hier wäre es besser, realistischere räumliche Konzepte zur Definition der Nachbarschaft zu wählen. Die PGS-Verfügbarkeit könnte beispielsweise mit Konzepten wie auf der Laufdistanz basierenden individuellen Nachbarschaften (Droin et al., 2023) kombiniert werden. So wird die Analyse der realistischen Erfahrung von Fußgängern gerechter, wobei möglicherweise unterschiedliche Gehgeschwindigkeiten berücksichtigt werden müssen. Die methodische Herausforderung hierbei liegt in der irregulären Form derartiger Nachbarschaften, was eine Anpassung der hier demonstrierten Methodik erfordern würde. Zudem haben Schindler et al. (2022) gezeigt, dass die Bereitschaft auch größere Strecken zurückzulegen, um Grünflächen zu besuchen, abhängig von sozioökonomischen Faktoren, stark variiert. Eine Ausweitung der Analysen auf die Umgebung des Arbeitsplatzes der Personen (z. B. Rauch et al., 2021) würde ein differenzierteres Bild der Verteilung der Erreichbarkeit von Grünflächen ermöglichen. Außerdem hat sich gezeigt, dass die gesundheitlichen Auswirkungen von Grünflächen von der Größe der Nachbarschaft abhängen (Reid et al., 2018; Su et al., 2019), was derartige Studien anfällig für das Problem der modifizierbaren Flächeneinheiten (MAUP) macht. In dieser Arbeit wurden die Dimensionen auf Basis der Literatur abgeleitet. Um mögliche Wirkungen von PGS auf die Gesundheit zu ermitteln, müssten mehrere Nachbarschaftsgrößen, möglicherweise in Kombination mit der Begehrbarkeit der Nachbarschaft, berücksichtigt werden. Dies rechtfertigt die Ausweitung künftiger Studien, um die Auswirkungen von Stadtteilen unterschiedlicher Größe zu untersuchen.

Der Grünflächentyp PGS wurde auf Basis von EUA-Daten ermittelt. Wie von Kabisch et al. (2016) erörtert, weisen diese Daten einige inhärente Einschränkungen auf, da deren Landbedeckungs- bzw. Landnutzungsklassen mitunter sehr breit definiert sind. Am Beispiel von *Green urban areas* kann man sehen, dass sie mehrere verschiedene Klassen von städtischem Grün umfassen. Außerdem kann die minimale Kartiereinheit von 0,25 ha pro Objekt kleinere Grünflächenmerkmale in komplexen städtischen Umgebungen kaschieren. Um eine größere Vielfalt an Grünflächentypen zu berücksichtigen, könnte der in dieser Studie verwendete Datenfusionsansatz auf explizitere und komplexere Grünflächenontologien erweitert werden (z. B. Ismayilova und Timpf, 2022). Auch das Klassifizierungsschema der Sentinel-2-basierten Landbedeckungsklassifikation, wie es von Weigand et al. (2020) vorgeschlagen und in Abschnitt 5.1.4 flächendeckend für Deutschland kartiert wurde, kann zu Verzerrungen führen. Hier ist darauf hinzuweisen, dass in der jüngeren Vergangenheit verschiedene großmaßstäbliche Landbedeckungsprodukte veröffentlicht wurden, die für ähnliche Analysen verwendet werden können, z. B. Dynamic World (Brown et al., 2022) oder ESA World Cover³⁵.

Im Rahmen dieser Arbeit werden die letzten verfügbaren Zensusdaten in Deutschland aus dem Jahr 2011 verwendet. Aktuellere Daten waren aufgrund von Verzögerungen durch die COVID-19-Pandemie³⁶ nicht verfügbar. Darüber hinaus bringt das räumliche Rasterformat einige Einschränkungen für diese Studie mit sich. Erstens erlaubt es keine detaillierte Beschreibung von Personen anhand mehrerer sozioökonomischer Variablen, d. h. Migrationshintergrund *und* Bildung *und* Einkommen *und* Alter. Daher war es in dieser Studie nicht möglich, die sozioökonomischen Variablen zu kontrollieren, die die Ergebnisse verfälschen. Dies wäre mit detaillierteren Paneldaten möglich (z. B. Krekel et al., 2016; Wüstemann et al., 2017a). Zweitens führen Randomisierungseffekte zu Verzerrungen in dünn besiedelten Zellen. Diese Zellen sind naturgemäß eher in weniger besiedelten Gebieten zu finden, was in ländlichen Gebieten oder am Rande von Städten der Fall sein dürfte. Dies kann sich auf diese Studie auswirken, da die Tendenz besteht, den Anteil der Menschen in Regionen mit geringerer Bevölkerungsdichte und möglicherweise hoher GLC/PGS-Verfügbarkeit zu unterschätzen.

Auch wenn Verwaltungsgrenzen zur Unterscheidung verschiedener Städte und ihrer Bevölkerung verwendet wurden, spiegeln sie nicht unbedingt die Realität in Bezug auf die Stadtregionen wider (Taubenböck et al., 2019). Es wäre daher interessant, den Einfluss der morphologischen Stadtregionen auf die Verfügbarkeit von Grünflächen zu untersuchen.

Trotz der zuvor genannten Limitierungen dieser Studie zeigt die hier entwickelte Methodik zur großflächigen Analyse von Grünflächenverteilung hohes Potenzial, die Gerechtigkeit der Umweltbedingungen in der Wohnumgebung auf einem neuen Level zu betrachten. Durch die

³⁵ESA, ohne Datum, „Worldwide land cover mapping“, zuletzt abgerufen 26.07.2023, <https://esa-worldcover.org>

³⁶Zensus 2022, 02.09.2020, „Bundeskabinett beschließt Verschiebung des Zensus 2021 um ein Jahr“, zuletzt abgerufen 26.07.2023, https://www.zensus2022.de/DE/Aktuelles/verschiebung_beschluss.html

objektive Erfassung der Erdoberfläche ermöglicht die Fernerkundung in diesem Kontext die Vergleichbarkeit von gesundheitsrelevanten Umweltfaktoren über Raum und Zeit herzustellen. Dies kann in zukünftigen Untersuchungen dazu dienen unabhängiger von bestehenden, mitunter räumlich oder semantisch begrenzten Datensätzen zu werden. Besonders für den internationalen Vergleich bedeutet das, Aspekte der Gerechtigkeit der Verfügbarkeit von beispielsweise Grünflächen konsistenter abbilden zu können. Allgemein ermöglicht der Einsatz von maschinellen Lernverfahren die zunehmende Masse an verfügbaren Geodaten in relevante Geoinformationen zu transformieren und diese so auch für Umweltanalysen zugänglich zu machen. Erst dadurch wird es möglich, Schwerpunkte zu setzen, die in bestehenden Geodaten nicht (explizit) abgebildet sind und so neue Fragen der gesellschaftlichen Transformation zu beantworten.



Fazit und Ausblick

Vor dem Hintergrund der zunehmend spürbaren Auswirkungen des Klimawandels, welche auch im Bericht des (engl.) *Intergovernmental Panel on Climate Change (IPCC)* von 2022 dokumentiert sind (Lwasa et al., 2022), spielen Grünflächen eine zentrale Rolle für die nachhaltige Stadtentwicklung. Sie sind darüber hinaus auch in den Zielvorgaben der *Sustainable Development Goals* der Vereinten Nationen verankert. Dennoch ist eine der großen Herausforderungen der Zukunft, die Entwicklung und den Erhalt der Grünflächen in den Städten auch sozial nachhaltig zu gestalten (Haase et al., 2017), denn die nachträgliche Begrünung von Nachbarschaften kann unter Umständen weitere Ungerechtigkeit hervorrufen (Rigolon und Németh, 2018). Die Möglichkeit die Verteilung von Umweltfaktoren wie Grünflächen hoch detailliert zu erfassen und Entwicklungen zu quantifizieren ist essenziell, um nachhaltige Stadtentwicklung auch objektiv messbar zu machen. Organisationen wie beispielsweise die Weltgesundheitsorganisation monitoren daher zunehmend die Verteilung von gesundheitsrelevanten Umwelteinflüssen (Braubach, 2012). Um die genaue sozialräumliche Verteilung von Grünflächen zu beschreiben, sind hoch aufgelöste, semantisch angemessene und flächendeckende Daten notwendig. Nur so können gesetzte Ziele auch eingehalten und verifiziert werden.

Siedlungsgebiete als menschliches Habitat sind jedoch komplex und divers (vgl. Taubenböck et al., 2018; Zhu et al., 2022; Standfuß et al., 2023a,b; Weigand et al., 2023a; Droin et al., in Begutachtung) und räumliche Daten zur Verteilung von Umwelteinflüssen stehen häufig nur auf stark aggregierter Ebene bereit. Die Untersuchungsgebiete in Umweltgerechtigkeitsstudien sind darüber hinaus häufig auf kleine Gebiete wie einzelne Städte beschränkt. Vor diesem Hintergrund verfolgte diese Dissertation die Potenziale moderner Erdbeobachtungs- bzw. Geodaten zu ergründen, bisherige Datenlücken mithilfe von maschinellem Lernen zu schließen und die räumliche Auflösung der erfassten Umwelteigenschaften zu erhöhen.

Die Arbeit war dabei geleitet von drei übergeordneten Forschungsfragen, die im Folgenden einzeln beantwortet werden sollen. Der erste konzeptionelle Teil dieser Arbeit befasste sich mit der Erfassung hochaufgelöster Landbedeckungsinformationen:

Forschungsfrage 1: *Wie können moderne Satellitendaten von Sentinel-2 genutzt werden, um Landbedeckungsinformationen zu Grünflächen auch in komplexen Siedlungsstrukturen hochaufgelöst flächendeckend abzuleiten?*

Die Ergebnisse dieser Dissertation zeigen, dass die Kombination aus LUCAS-Referenzpunkten und Sentinel-2 erfolgreich zur Ableitung von Landbedeckungsinformationen eingesetzt werden können. Mit einem Random Forest Klassifikationsalgorithmus konnten hierfür hohe bis sehr hohe Genauigkeiten erzielt werden ($\hat{O} > 90\%$). Dies gilt nicht nur für die von Weigand et al. (2020) speziell entwickelte Klassenhierarchie, sondern schließt bisherige Vorprozessierungsschemata mit ein, die ursprünglich für die Klassifikation von Landsat-Daten entwickelt wurden. Auch diese konnten erfolgreich auf Sentinel-2 angewendet werden. Mit der hohen räumlichen Auflösung von $10\text{ m} \times 10\text{ m}$ bietet diese Datenbasis die Möglichkeit, auch kleine Bereiche von Grünflächen in komplexen Siedlungsstrukturen zu identifizieren.

Durch die erneute Anwendung der entwickelten Klassifikationsmethodik auf einen zweiten Zeitraum konnte obendrein die Reproduzierbarkeit der Methode demonstriert werden. Die beiden Klassifikationsprodukte ermöglichen es, Landschaftsdynamiken auf übergeordneter Ebene zu beschreiben. Eine weitergehende, genauere Untersuchung des Einflusses der ausgewählten LUCAS-Punkte ist jedoch erforderlich, um eine höhere Vergleichbarkeit der beiden Klassifikationsprodukte auf Pixelebene zu ermöglichen. Hierbei werden neue Technologien wie STAC und COG sowie cloudbasierte Prozessierungsumgebungen helfen, indem sie großflächige Prozessierung von Erdbeobachtungsdaten zunehmend zugänglicher machen. Derartige Anwendungen werden auch über die Klassifikation mit LUCAS-Referenzdaten hinaus die Zukunft fernerkundlicher Applikationen nachhaltig mitbestimmen.

Die Zunahme an hochaufgelösten und frei zugänglichen Erdbeobachtungsdaten in Kombination mit Methoden des maschinellen Lernens und insbesondere Deep Learning sind von zentraler Bedeutung für die Forschung rund um die Klassifikation von Landbedeckung (Feng und Li, 2020). Zudem sind zunehmend umfassende Referenzdatensätze verfügbar, welche die spezifischen Anforderungen fernerkundlicher Bildverarbeitung adressieren, z. B. EuroSAT (Helber et al., 2019) oder Datensätze aus dem Kontext des So2Sat-Projekts (Zhu et al., 2020; Doda et al., 2022). In den kommenden Jahren werden diese Faktoren weiter dazu beitragen, Geodaten in großflächige und semantisch umfangreiche Geoinformationen zur Landbedeckung zu transformieren, diese zugänglich zu machen und so die Nutzbarkeit von Erdbeobachtungstechnologien für viele interdisziplinäre Einsatzzwecke zu erhöhen.

Aufbauend auf den Ergebnissen der Landbedeckungsklassifikation wurde das Ziel verfolgt, verschiedene Typen von Grünflächen weiter zu differenzieren:

Forschungsfrage 2: Ist es möglich, den Anteil öffentlich zugänglicher Grünflächen auf Nachbarschaftsebene mithilfe einer Kombination aus Satellitendaten und weiteren frei zugänglichen Geodaten hochgenau zu quantifizieren?

Zur semantischen Anreicherung der Grünflächen in Deutschland wurden zunächst Vektordaten des OpenStreetMap-Projekts aufbereitet. Mit dem Ziel Satellitenbilder und OSM-Aggregate simultan zu verarbeiten, wurde ein *Deep Learning*-basiertes maschinelles Lernverfahren entwickelt, welches die Fläche öffentlichen Grüns auf Nachbarschaftsebene ableitet. Der wissenschaftliche Beitrag

dieser Arbeit liegt in dem hier erfolgreich geführten Machbarkeitsbeweis der derartigen Fusion von Bild- und aggregierten Vektordaten zur Quantifizierung von Landnutzungstypen am Beispiel öffentlichen Grüns. Zuvor lagen detaillierte Informationen zur Verfügbarkeit von öffentlichem Grün in der Wohnumgebung lediglich für drei Viertel der Bevölkerung Deutschlands vor. Mithilfe des hier angewandten Deep Learning Modells konnte die Abdeckung dieser auf die gesamte Bevölkerung Deutschlands erweitert werden.

Die Methode der Datenfusion von OSM und Bilddaten lässt sich auf viele Anwendungsbereiche übertragen. Sie ermöglicht die Erfassung semantischer, nicht zwingend physischer Merkmale zur Unterscheidung von Objekten. Im Kontext urbaner Grünflächen wäre beispielsweise denkbar, weitere Typen von Grün zu unterscheiden, wie beispielsweise Straßenbegleitgrün, Schrebergärten, Sportanlagen und viele mehr. Dennoch ergibt sich aus der entwickelten Methode eine Abhängigkeit von den Beiträgen Freiwilliger zur Datenbank der OpenStreetMap. Obgleich durch die Fusion eventuelle Lücken in Satellitenbildern oder OSM-Daten ausgeglichen werden können, muss in weiteren Untersuchungen die genaue Wechselwirkung derartiger Inkonsistenzen genauer beleuchtet werden. In der vorliegenden Arbeit handelte es sich bei der Zielvariable des maschinellen Lernenverfahrens (öffentliche Grünfläche in der Nachbarschaft) um eine kontinuierliche Variable, weshalb ein Regressionsverfahren eingesetzt wurde. Es bleibt zu erkunden, ob für diese spezielle Anwendung, auch vor dem Hintergrund der anschließenden Analyse, ein Klassifikationsnetzwerk bessere Genauigkeiten erzielen könnte, z. B. in Klassen $PGS < 3,6$ ha und $PGS \geq 3,6$ ha.

Mit den Datensätzen der vorhergehenden Arbeitsschritte wurde es möglich, die Grünflächen aus dem Blickwinkel der Umweltgerechtigkeit deutschlandweit zu betrachten und deren Verteilung anhand zweier unterschiedlicher Definitionen zu untersuchen: grüne Landbedeckung (GLC) und öffentliche Grünflächen (PGS). Die Kombination mit den geo-lokalisierten Zensusdaten erlaubte es, die Verteilung der Grünflächenverfügbarkeit auf die Bevölkerung und ihre Beziehung zur gesellschaftlichen Zusammensetzung zu quantifizieren. Dabei stand die folgende Frage im Mittelpunkt:

Forschungsfrage 3: Wie gerecht sind grüne Landbedeckung sowie öffentliches Grün in der Wohnumwelt für die deutsche Bevölkerung verteilt?

Diese Frage wurde mithilfe deskriptiver Statistiken und (engl.) *Mixed Effects Models* bearbeitet. Insgesamt haben mehr als 80 % der Bevölkerung Deutschlands Zugang zu mehr öffentlichem Grün, als von der Weltgesundheitsorganisation als Zielvorgabe definiert wurde. Sowohl die Verfügbarkeit von grüner Landbedeckung als auch von öffentlichen Grünflächen variiert signifikant mit dem Grad der Urbanität und mit der Gebäudestruktur in der Nachbarschaft. Diese beiden Ressourcen, GLC und PGS, sind jedoch sehr unterschiedlich auf die Bevölkerung verteilt und zeigen entlang des Stadt-Land-Gradienten sogar gegenläufige Trends. In Stadtteilen mit einem höheren Anteil ausländischer Bewohner ist die Verfügbarkeit von GLC geringer; ein Trend, der sich für PGS in Deutschland nicht feststellen lässt.

Diese Studie baut auf den landesweiten Bevölkerungsdaten des Zensus 2011 auf. Dieser erlaubt jedoch nur eine begrenzte Unterscheidung einzelner sozioökonomischer und soziodemographischer Gruppen der Gesellschaft. Zukünftig können die hier entwickelten landesweiten Geoinformationen zu Grünflächen in der Nachbarschaft mit detaillierteren sozioökonomischen Daten kombiniert werden. Dies kann dabei helfen, weitere soziale Ungleichheiten in Bezug auf Grünflächen in Deutschland zu identifizieren. Für eine derartige Untersuchung bieten sich in Deutschland beispielsweise Daten des Sozioökonomischen Panels (SOEP, Wagner et al., 2007) an. Auf diese Weise könnten auch unterschiedliche Präferenzen und Anforderungen einzelner Gesellschaftsgruppen an Grünflächen (Swanwick, 2009; Calderón-Argelich et al., 2023) differenziert berücksichtigt werden. In Verbindung mit Gesundheitsdaten kann die Fusion von derartigen Geoinformationen genutzt werden, um die salutogenen Wirkungspfade und Effekte des Zugangs zu verschiedenen Grünflächentypen in der Bevölkerung besser zu verstehen.

Die Umweltgerechtigkeit steht im Kontext vieler interdisziplinärer Betrachtungsweisen. Dieses viele Disziplinen umspannende Fachgebiet lädt dazu ein, dass viele unterschiedliche Herangehensweisen an dessen fachlichen Schwerpunkten zusammentreffen (Knable und Yu, 2023). Wie die Literaturrecherche von Weigand et al. (2019) gezeigt hat, werden im Bezug auf räumliche Analysen jedoch selten flächendeckende und hoch aufgelöste Ansätze gewählt. Vor diesem Hintergrund widmete sich die vorliegende Arbeit dem Feld aus einer quantitativen, raumwissenschaftlichen und datengetriebenen Perspektive, um diese Datenlücke zu schließen und zu demonstrieren, dass landesweite Analysen mit existierenden Daten möglich sind und neue Erkenntnisse liefern. Davon werden zukünftige Untersuchungen zu gesundheitlichen Einflüssen der Umwelt beziehungsweise Analysen der Umweltgerechtigkeit profitieren. Das gewonnene Wissen kann als Grundlage und Entscheidungsunterstützung für die nachhaltige Stadtentwicklung dienen.

Literatur

- Akpinar, A., Barbosa-Leiker, C. und Brooks, K. R. (2016). „Does green space matter? Exploring relationships between green space type and health indicators“. In: *Urban Forestry & Urban Greening* 20, S. 407–418. DOI: [10.1016/j.ufug.2016.10.013](https://doi.org/10.1016/j.ufug.2016.10.013).
- Alidoust, S. und Bosman, C. (2015). „Planning for an ageing population: links between social health, neighbourhood environment and the elderly“. In: *Australian Planner* 52.3, S. 177–186. DOI: [10.1080/07293682.2015.1034145](https://doi.org/10.1080/07293682.2015.1034145).
- Altman, D. G. und Bland, J. M. (1994a). „Statistics Notes: Diagnostic tests 2: predictive values“. In: *BMJ* 309.6947, S. 102–102. DOI: [10.1136/bmj.309.6947.102](https://doi.org/10.1136/bmj.309.6947.102).
- Altman, D. G. und Bland, J. M. (1994b). „Statistics Notes: Diagnostic tests 1: sensitivity and specificity“. In: *BMJ* 308.6943, S. 1552–1552. DOI: [10.1136/bmj.308.6943.1552](https://doi.org/10.1136/bmj.308.6943.1552).
- Antonovsky, A. (1997). *Salutogenese: Zur Entmystifizierung der Gesundheit*. Bd. 36. Forum für Verhaltenstherapie und psychosoziale Praxis. Tübingen: dgvt-Verlag. 222 S.
- Apparicio, P., Séguin, A.-M., Dubé, J. et al. (2016). „Spatial distribution of vegetation in and around city blocks on the Island of Montreal: A double environmental inequity?“ In: *Applied Geography* 76, S. 128–136. DOI: [10.1016/j.apgeog.2016.09.023](https://doi.org/10.1016/j.apgeog.2016.09.023).
- Arsanjani, J. J. und Vaz, E. (2015). „An assessment of a collaborative mapping approach for exploring land use patterns for several European metropolises“. In: *International Journal of Applied Earth Observation and Geoinformation* 35, S. 329–337. DOI: [10.1016/j.jag.2014.09.009](https://doi.org/10.1016/j.jag.2014.09.009).
- Artmann, M., Mueller, C., Goetzlich, L. und Hof, A. (2019). „Supply and Demand Concerning Urban Green Spaces for Recreation by Elderlies Living in Care Facilities: The Role of Accessibility in an Explorative Case Study in Austria“. In: *Frontiers in Environmental Science* 7, S. 12. DOI: [10.3389/fenvs.2019.00136](https://doi.org/10.3389/fenvs.2019.00136).
- Atkinson, A. B. (1970). „On the measurement of inequality“. In: *Journal of Economic Theory* 2.3, S. 244–263. DOI: [10.1016/0022-0531\(70\)90039-6](https://doi.org/10.1016/0022-0531(70)90039-6).
- Auchincloss, A. H., Gebreab, S. Y., Mair, C. und Diez Roux, A. V. (2012). „A review of spatial methods in epidemiology, 2000–2010“. In: *Annual review of public health* 33, S. 107–122. DOI: [10.1146/annurev-publhealth-031811-124655](https://doi.org/10.1146/annurev-publhealth-031811-124655).
- Audebert, N., Le Saux, B. und Lefèvre, S. (2017). „Joint Learning from Earth Observation and OpenStreetMap Data to Get Faster Better Semantic Maps“. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. ISSN: 2160-7516, S. 1552–1560. DOI: [10.1109/CVPRW.2017.199](https://doi.org/10.1109/CVPRW.2017.199).

- Bahrenberg, G., Giese, E., Mevenkamp, N. und Nipper, J. (2010). *Statistische Methoden der Geographie – Band 1 Univariate und bivariate Statistik*. Hrsg. von H. Gebhardt, R. Baumhauer, J. Bendix und P. Reuber. 5. Aufl. Studienbücher der Geographie. Stuttgart: Gebr. Borntraeger.
- Banzhaf, E. und Barrera, F. de la (2017). „Evaluating public green spaces for the quality of life in cities by integrating RS mapping tools and social science techniques“. In: *2017 Joint Urban Remote Sensing Event (JURSE)*. IEEE, S. 4. DOI: [10.1109/jurse.2017.7924559](https://doi.org/10.1109/jurse.2017.7924559).
- Barbosa, O., Tratalos, J. A., Armsworth, P. R., Davies, R. G., Fuller, R. A., Johnson, P. und Gaston, K. J. (2007). „Who benefits from access to green space? A case study from Sheffield, UK“. In: *Landscape and Urban Planning* 83.2-3, S. 187–195. DOI: [10.1016/j.landurbplan.2007.04.004](https://doi.org/10.1016/j.landurbplan.2007.04.004).
- Baró, F., Calderón-Argelich, A., Langemeyer, J. und Connolly, J. J. (2019). „Under one canopy? Assessing the distributional environmental justice implications of street tree benefits in Barcelona“. In: *Environmental Science & Policy* 102, S. 54–64. DOI: [10.1016/j.envsci.2019.08.016](https://doi.org/10.1016/j.envsci.2019.08.016).
- Barber, A., Haase, D. und Wolff, M. (2021). „Permeability of the city – Physical barriers of and in urban green spaces in the city of Halle, Germany“. In: *Ecological Indicators* 125.107555, S. 12. DOI: [10.1016/j.ecolind.2021.107555](https://doi.org/10.1016/j.ecolind.2021.107555).
- Bayas, J. L., See, L., Fritz, S., Sturn, T., Perger, C., Dürauer, M., Karner, M., Moorthy, I., Schepaschenko, D., Domian, D. und McCallum, I. (2016). „Crowdsourcing In-Situ Data on Land Cover and Land Use Using Gamification and Mobile Technology“. In: *Remote Sensing* 8.905, S. 18. DOI: [10.3390/rs8110905](https://doi.org/10.3390/rs8110905).
- Bayerisches Ministerium für Gesundheit und Pflege (2020). *Vollzug des Infektionsschutzgesetzes - Vorläufige Ausgangsbeschränkung anlässlich der Corona-Pandemie*. Bekanntmachung vom 20.03.2020, Az. Z6a-G8000-2020/122-98.
- Bechtel, B. und Daneke, C. (2012). „Classification of Local Climate Zones Based on Multiple Earth Observation Data“. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 5.4, S. 1191–1202. DOI: [10.1109/jstars.2012.2189873](https://doi.org/10.1109/jstars.2012.2189873).
- Bejani, M. M. und Ghatge, M. (2021). „A systematic review on overfitting control in shallow and deep neural networks“. In: *Artificial Intelligence Review* 54.8, S. 6391–6438. DOI: [10.1007/s10462-021-09975-1](https://doi.org/10.1007/s10462-021-09975-1).
- Belgiu, M. und Drăguț, L. (2016). „Random forest in remote sensing: A review of applications and future directions“. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 114, S. 24–31. DOI: [10.1016/j.isprsjprs.2016.01.011](https://doi.org/10.1016/j.isprsjprs.2016.01.011).
- Berg, A. E. van den, Maas, J., Verheij, R. A. und Groenewegen, P. P. (2010). „Green space as a buffer between stressful life events and health“. In: *Social science & medicine* 70.8, S. 1203–1210. DOI: [10.1016/j.socscimed.2010.01.002](https://doi.org/10.1016/j.socscimed.2010.01.002).
- Bertram, C. und Rehdanz, K. (2015). „The role of urban green space for human well-being“. In: *Ecological Economics* 120, S. 139–152. DOI: [10.1016/j.ecolecon.2015.10.013](https://doi.org/10.1016/j.ecolecon.2015.10.013).

- Berdejo-Espinola, V., Suárez-Castro, A. F., Amano, T., Fielding, K. S., Oh, R. R. Y. und Fuller, R. A. (2021). „Urban green space use during a time of stress: A case study during the COVID-19 pandemic in Brisbane, Australia“. In: *People and Nature* 3.3, S. 597–609. DOI: [10.1002/pan3.10218](https://doi.org/10.1002/pan3.10218).
- Bolte, G., Pauli, A. und Hornberg, C. (2011). „Environmental justice: Social disparities in environmental exposures and health: Overview“. In: *Encyclopedia of Environmental Health*. Hrsg. von J. O. Nriagu. Bd. 2. Elsevier, S. 459–470. DOI: [10.1016/B978-0-444-52272-6.00685-1](https://doi.org/10.1016/B978-0-444-52272-6.00685-1).
- Bolte, G., Bunge, C., Hornberg, C., Köckler, H. und Mielck, A. (2012a). „Umweltgerechtigkeit durch Chancengleichheit bei Umwelt und Gesundheit. Eine Einführung in die Thematik und Zielsetzung dieses Buches“. In: *Umweltgerechtigkeit – Chancengleichheit bei Umwelt und Gesundheit: Konzepte, Datenlage und Handlungsperspektiven*. Hrsg. von G. Bolte, C. Bunge, C. Hornberg und A. Mielck. 1. Aufl., S. 15–37.
- Bolte, G., Voigtländer, S., Razum, O. und Mielck, A. (2012b). „Modelle zur Erklärung des Zusammenhangs zwischen sozialer Lage, Umwelt und Gesundheit“. In: *Umweltgerechtigkeit – Chancengleichheit bei Umwelt und Gesundheit: Konzepte, Datenlage und Handlungsperspektiven*. Hrsg. von G. Bolte, C. Bunge, C. Hornberg, H. Köckler und A. Mielck. Bern: Hans Huber, S. 39–50.
- Bowler, D. E., Buyung-Ali, L., Knight, T. M. und Pullin, A. S. (2010). „Urban greening to cool towns and cities: A systematic review of the empirical evidence“. In: *Landscape and urban planning* 97.3, S. 147–155. DOI: [10.1016/j.landurbplan.2010.05.006](https://doi.org/10.1016/j.landurbplan.2010.05.006).
- Bray, I., Reece, R., Sinnett, D., Martin, F. und Hayward, R. (2022). „Exploring the role of exposure to green and blue spaces in preventing anxiety and depression among young people aged 14–24 years living in urban settings: A systematic review and conceptual framework“. In: *Environmental Research* 214, S. 114081. DOI: [10.1016/j.envres.2022.114081](https://doi.org/10.1016/j.envres.2022.114081).
- Braubach, M. (2012). „Gesundheitsrelevante Umweltgerechtigkeit in Europa: die Arbeit der Weltgesundheitsorganisation (WHO)“. In: *Umweltgerechtigkeit – Chancengleichheit bei Umwelt und Gesundheit: Konzepte, Datenlage und Handlungsperspektiven*. Hrsg. von G. Bolte, C. Bunge, C. Hornberg, H. Köckler und A. Mielck. Bern: Hans Huber, S. 151–158.
- Breiman, L. (2001). „Random Forests“. In: *Machine Learning* 45.1, S. 5–32. DOI: [10.1023/a:1010933404324](https://doi.org/10.1023/a:1010933404324).
- Breiman, L. (1996). „Bagging predictors“. In: *Machine Learning* 24.2, S. 123–140. DOI: [10.1007/bf00058655](https://doi.org/10.1007/bf00058655).
- Browning, M. H. und Locke, D. H. (2020). „The greenspace-academic performance link varies by remote sensing measure and urbanicity around Maryland public schools“. In: *Landscape and Urban Planning* 195, S. 103706. DOI: [10.1016/j.landurbplan.2019.103706](https://doi.org/10.1016/j.landurbplan.2019.103706).

- Brown, C. F., Brumby, S. P., Guzder-Williams, B., Birch, T., Hyde, S. B., Mazzariello, J., Czerwinski, W., Pasquarella, V. J., Haertel, R., Ilyushchenko, S., Schwehr, K., Weisse, M., Stolle, F., Hanson, C., Guinan, O., Moore, R. und Tait, A. M. (2022). „Dynamic World, Near real-time global 10 m land use land cover mapping“. In: *Scientific Data* 9.1, S. 17. DOI: [10.1038/s41597-022-01307-4](https://doi.org/10.1038/s41597-022-01307-4).
- Burkart, K., Meier, F., Schneider, A., Breitner, S., Canário, P., Alcoforado, M. J., Scherer, D. und Endlicher, W. (2015). „Modification of Heat-Related Mortality in an Elderly Urban Population by Vegetation (Urban Green) and Proximity to Water (Urban Blue): Evidence from Lisbon, Portugal“. In: *Environmental Health Perspectives* 124.7, S. 927–934. DOI: [10.1289/ehp.1409529](https://doi.org/10.1289/ehp.1409529).
- Calderón-Argelich, A., Benetti, S., Anguelovski, I., Connolly, J. J., Langemeyer, J. und Baró, F. (2021). „Tracing and building up environmental justice considerations in the urban ecosystem service literature: A systematic review“. In: *Landscape and Urban Planning* 214, S. 104130. DOI: [10.1016/j.landurbplan.2021.104130](https://doi.org/10.1016/j.landurbplan.2021.104130).
- Calderón-Argelich, A., Anguelovski, I., Connolly, J. J. und Baró, F. (2023). „Greening plans as (re)presentation of the city: Toward an inclusive and gender-sensitive approach to urban greenspaces“. In: *Urban Forestry & Urban Greening* 86, S. 127984. DOI: [10.1016/j.ufug.2023.127984](https://doi.org/10.1016/j.ufug.2023.127984).
- Camps-Valls, G. (2009). „Machine learning in remote sensing data processing“. In: *2009 IEEE International Workshop on Machine Learning for Signal Processing*. IEEE. DOI: [10.1109/mlsp.2009.5306233](https://doi.org/10.1109/mlsp.2009.5306233).
- Card, D. H. (1982). „Using Known Map Category Marginal Frequencies to Improve Estimates of Thematic Map Accuracy“. In: *Photogrammetric Engineering and Remote Sensing* 48.3, S. 431–439.
- Castelli, C., d’Hombres, B., Dominicis, L. de, Dijkstra, L., Montalto, V. und Pontarollo, N. (2023). „What makes cities happy? Factors contributing to life satisfaction in European cities“. In: *European Urban and Regional Studies*, S. 096977642311553. DOI: [10.1177/09697764231155335](https://doi.org/10.1177/09697764231155335).
- Chang, M.-E., Zhao, Z.-Q., Chang, H.-T. und Shu, B. (2021). „Urban green infrastructure health assessment, based on landsat 8 remote sensing and entropy landscape metrics“. In: *European Journal of Remote Sensing* 54.1, S. 417–430. DOI: [10.1080/22797254.2021.1948357](https://doi.org/10.1080/22797254.2021.1948357).
- Chan, C. Y.-C., Weigand, M., Alnajar, E. A. und Taubenböck, H. (2022). „Investigating the capability of UAV imagery for AI-assisted mapping of Refugee Camps in East Africa“. In: *Proceedings of the Academic Track at State of the Map 2022*. Hrsg. von M. Minghini, P. Liu, H. Li, A. Y. Grindberger und L. Juhász. Florence, Italy: Zenodo, S. 45–48. DOI: [10.5281/ZENODO.7004576](https://doi.org/10.5281/ZENODO.7004576).
- Chen, J., Ban, Y. und Li, S. (2014). „Open access to Earth land-cover map“. In: *Nature* 514.7523, S. 434–434. DOI: [10.1038/514434c](https://doi.org/10.1038/514434c).

- Cheng, G., Han, J. und Lu, X. (2017). „Remote Sensing Image Scene Classification: Benchmark and State of the Art“. In: *Proceedings of the IEEE* 105.10, S. 1865–1883. DOI: [10.1109/jproc.2017.2675998](https://doi.org/10.1109/jproc.2017.2675998).
- Chen, J., Zhou, Y., Zipf, A. und Fan, H. (2019). „Deep Learning From Multiple Crowds: A Case Study of Humanitarian Mapping“. In: *IEEE Transactions on Geoscience and Remote Sensing* 57.3, S. 1713–1722. DOI: [10.1109/tgrs.2018.2868748](https://doi.org/10.1109/tgrs.2018.2868748).
- Cheng, G., Xie, X., Han, J., Guo, L. und Xia, G.-S. (2020). „Remote Sensing Image Scene Classification Meets Deep Learning: Challenges, Methods, Benchmarks, and Opportunities“. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 13, S. 3735–3756. DOI: [10.1109/jstars.2020.3005403](https://doi.org/10.1109/jstars.2020.3005403).
- Claßen, T., Heiler, A. und Brei, B. (2012). „Urbane Grünräume und gesundheitliche Chancengleichheit – längst nicht alles im “grünen Bereich”“. In: *Umweltgerechtigkeit durch Chancengleichheit bei Umwelt und Gesundheit-Konzepte, Datenlage und Handlungsperspektiven*. Hrsg. von G. Bolte, C. Bunge, C. Hornberg, H. Köckler und A. Mielck. Bern: Huber, S. 113–123.
- Close, O., Benjamin, B., Petit, S., Fripiat, X. und Hallot, E. (2018). „Use of Sentinel-2 and LUCAS Database for the Inventory of Land Use, Land Use Change, and Forestry in Wallonia, Belgium“. In: *Land* 7.4, S. 154. DOI: [10.3390/land7040154](https://doi.org/10.3390/land7040154).
- Conrad, C., Goessl, A., Lex, S., Metz, A., Esch, T., Konrad, C., Goettlicher, G. und Dech, S. (2010). „Mapping crop distribution in administrative districts of southwest Germany using multi-sensor remote sensing data“. In: *Remote Sensing for Agriculture, Ecosystems, and Hydrology XII*. Hrsg. von C. M. U. Neale und A. Maltese. SPIE, S. 1–9. DOI: [10.1117/12.865113](https://doi.org/10.1117/12.865113).
- Congalton, R. G. (1991). „A review of assessing the accuracy of classifications of remotely sensed data“. In: *Remote Sensing of Environment* 37.1, S. 35–46. DOI: [10.1016/0034-4257\(91\)90048-b](https://doi.org/10.1016/0034-4257(91)90048-b).
- Coppel, G. und Wüstemann, H. (2017). „The impact of urban green space on health in Berlin, Germany: Empirical findings and implications for urban planning“. In: *Landscape and Urban Planning* 167, S. 410–418. DOI: [10.1016/j.landurbplan.2017.06.015](https://doi.org/10.1016/j.landurbplan.2017.06.015).
- Cowell, F. (2011). *Measuring Inequality*. Oxford University Press. DOI: [10.1093/acprof:osobl/9780199594030.001.0001](https://doi.org/10.1093/acprof:osobl/9780199594030.001.0001).
- Dai, J., Luo, F., Chen, Z., Zhou, G. und Zeng, J. (2023). „Pixel-based unique phenological feature composite method: tackle the challenges of Spartina alterniflora identification and salt marsh classification in mixed vegetations area“. In: *Journal of Applied Remote Sensing* 17.02, S. 22. DOI: [10.1117/1.jrs.17.024510](https://doi.org/10.1117/1.jrs.17.024510).
- Danylo, O., See, L., Bechtel, B., Schepaschenko, D. und Fritz, S. (2016). „Contributing to WUDAPT: A Local Climate Zone Classification of Two Cities in Ukraine“. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 9.5, S. 1841–1853. DOI: [10.1109/jstars.2016.2539977](https://doi.org/10.1109/jstars.2016.2539977).

- d'Andrimont, R., Yordanov, M., Martinez-Sanchez, L., Eiselt, B., Palmieri, A., Dominici, P., Gallego, J., Reuter, H. I., Joebges, C., Lemoine, G. und Velde, M. van der (2020). „Harmonised LUCAS in-situ land cover and use database for field surveys from 2006 to 2018 in the European Union“. In: *Scientific Data* 7.1. DOI: [10.1038/s41597-020-00675-z](https://doi.org/10.1038/s41597-020-00675-z).
- d'Andrimont, R., Verhegghen, A., Lemoine, G., Kempeneers, P., Meroni, M. und Velde, M. van der (2021a). „From parcel to continental scale – A first European crop type map based on Sentinel-1 and LUCAS Copernicus in-situ observations“. In: *Remote Sensing of Environment* 266, S. 112708. DOI: [10.1016/j.rse.2021.112708](https://doi.org/10.1016/j.rse.2021.112708).
- d'Andrimont, R., Verhegghen, A., Meroni, M., Lemoine, G., Strobl, P., Eiselt, B., Yordanov, M., Martinez-Sanchez, L. und Velde, M. van der (2021b). „LUCAS Copernicus 2018: Earth-observation-relevant in situ data on land cover and use throughout the European Union“. In: *Earth System Science Data* 13.3, S. 1119–1133. DOI: [10.5194/essd-13-1119-2021](https://doi.org/10.5194/essd-13-1119-2021).
- De Nicolò, S., Ferrante, M. R. und Pacei, S. (2022). „Small Area Estimation of Inequality Measures using Mixtures of Betas“. In: S. 30. arXiv: [2209.01985](https://arxiv.org/abs/2209.01985).
- De Nicolò, S., Ferrante, M. R. und Pacei, S. (2023). „Mind the Income Gap: Bias Correction of Inequality Estimators in Small-Sized Samples“. In: S. 21. arXiv: [2107.08950](https://arxiv.org/abs/2107.08950).
- Dennis, M., Barlow, D., Cavan, G., Cook, P., Gilchrist, A., Handley, J., James, P., Thompson, J., Tzoulas, K., Wheeler, C. P. und Lindley, S. (2018). „Mapping Urban Green Infrastructure: A Novel Landscape-Based Approach to Incorporating Land Use and Land Cover in the Mapping of Human-Dominated Systems“. In: *Land* 7.1, S. 17. DOI: [10.3390/land7010017](https://doi.org/10.3390/land7010017).
- Dewulf, B., Neutens, T., Van Dyck, D., De Bourdeaudhuij, I., Broekx, S., Beckx, C. und Van de Weghe, N. (2016). „Associations between time spent in green areas and physical activity among late middle-aged adults“. In: *Geospatial health* 11.3, S. 225–232. DOI: [10.4081/gh.2016.411](https://doi.org/10.4081/gh.2016.411).
- Diez Roux, A. V. (2001). „Investigating neighborhood and area effects on health“. In: *American journal of public health* 91.11, S. 1783–1789. DOI: [10.2105/ajph.91.11.1783](https://doi.org/10.2105/ajph.91.11.1783).
- Diekmann, A. (2007). *Empirische Sozialforschung*. Hrsg. von B. König. 10. Aufl. Reinbeck bei Hamburg: Rowohlt Taschenbuch. 784 S.
- Dijk-Wesselius, J. E. van, Maas, J., Hovinga, D., Vugt, M. van und Berg, A. E. van den (2018). „The impact of greening schoolyards on the appreciation, and physical, cognitive and social-emotional well-being of schoolchildren: A prospective intervention study“. In: *Landscape and Urban Planning* 180, S. 15–26. DOI: [10.1016/j.landurbplan.2018.08.003](https://doi.org/10.1016/j.landurbplan.2018.08.003).
- Doda, S., Wang, Y., Kahl, M., Hoffmann, E. J., Ouan, K., Taubenböck, H. und Zhu, X. X. (2022). „So2Sat POP - A Curated Benchmark Data Set for Population Estimation from Space on a Continental Scale“. In: *Scientific Data* 9.1, S. 1–11. DOI: [10.1038/s41597-022-01780-x](https://doi.org/10.1038/s41597-022-01780-x).
- Drain, A., Wurm, M. und Taubenböck, H. (2023). „The Individual Walkable Neighborhood - Evaluating people-centered spatial units focusing on urban density“. In: *Computers, Environment and Urban Systems* 99, S. 101893. DOI: [10.1016/j.compenvurbsys.2022.101893](https://doi.org/10.1016/j.compenvurbsys.2022.101893).
- Drain, A., Wurm, M., Weigand, M., Gawlas, C. und Taubenböck, H. (in Begutachtung). „How does pedestrian permeability vary across cities? A fine-grained assessment for all large cities in Germany“. In: *Computers, Environment and Urban Systems*.

- Du, P., Samat, A., Waske, B., Liu, S. und Li, Z. (2015). „Random Forest and Rotation Forest for fully polarized SAR image classification using polarimetric and spatial features“. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 105, S. 38–53. DOI: [10.1016/j.isprsjprs.2015.03.002](https://doi.org/10.1016/j.isprsjprs.2015.03.002).
- Dugord, P.-A., Lauf, S., Schuster, C. und Kleinschmit, B. (2014). „Land use patterns, temperature distribution, and potential heat stress risk—The case study Berlin, Germany“. In: *Computers, Environment and Urban Systems* 48, S. 86–98. DOI: [10.1016/j.compenvurbsys.2014.07.005](https://doi.org/10.1016/j.compenvurbsys.2014.07.005).
- Egan, W. G., Hrsg. (2003). *Optical Remote Sensing – Science and Technology. Science and Technology*. 1. Aufl. Boca Raton: CRC Press, S. 528.
- Engemann, K., Pedersen, C. B., Arge, L., Tsiogiannis, C., Mortensen, P. B. und Svenning, J.-C. (2019). „Residential green space in childhood is associated with lower risk of psychiatric disorders from adolescence into adulthood“. In: *Proceedings of the National Academy of Sciences* 116.11, S. 5188–5193. DOI: [10.1073/pnas.1807504116](https://doi.org/10.1073/pnas.1807504116).
- Enssle, F. und Kabisch, N. (2020). „Urban green spaces for the social interaction, health and well-being of older people— An integrated view of urban ecosystem services and socio-environmental justice“. In: *Environmental Science & Policy* 109, S. 36–44. DOI: [10.1016/j.envsci.2020.04.008](https://doi.org/10.1016/j.envsci.2020.04.008).
- Esch, T., Metz, A., Marconcini, M. und Keil, M. (2014). „Combined use of multi-seasonal high and medium resolution satellite imagery for parcel-related mapping of cropland and grassland“. In: *International Journal of Applied Earth Observation and Geoinformation* 28, S. 230–237. DOI: [10.1016/j.jag.2013.12.007](https://doi.org/10.1016/j.jag.2013.12.007).
- EUROSTAT (2015a). *LUCAS 2015 (Land Use / Cover Area Frame Survey) – Technical reference document C1: Instructions for Surveyors*. Techn. Ber. Eurostat, Regional statistics und Geographic Information, S. 1–140.
- EUROSTAT (2015b). *LUCAS 2015 (Land Use /Cover Area Frame Survey) – Technical Reference Document: C3 Classification (Land Use and Land Cover)*. Techn. Ber. European Commission, Eurostat.
- EUROSTAT (2022). *New LUCAS 2022 sample and subsamples design: criticalities and solutions – 2022 edition*. Hrsg. von European Commission. Statistical Office of the European Union. Luxembourg: Publications Office of the European Union. DOI: [10.2785/957524](https://doi.org/10.2785/957524).
- Feltynowski, M., Kronenberg, J., Bergier, T., Kabisch, N., Łaskiewicz, E. und Strohbach, M. W. (2018). „Challenges of urban green space management in the face of using inadequate data“. In: *Urban Forestry & Urban Greening* 31, S. 56–66. DOI: [10.1016/j.ufug.2017.12.003](https://doi.org/10.1016/j.ufug.2017.12.003).
- Feng, M. und Li, X. (2020). „Land cover mapping toward finer scales“. In: *Science Bulletin* 65.19, S. 1604–1606. DOI: [10.1016/j.scib.2020.06.014](https://doi.org/10.1016/j.scib.2020.06.014).
- Fernández-Delgado, M., Cernadas, E., Barro, S. und Amorim, D. (2014). „Do we Need Hundreds of Classifiers to Solve Real World Classification Problems?“ In: *Journal of Machine Learning Research* 15, S. 3133–3181.

- Flanagin, A. J. und Metzger, M. J. (2008). „The credibility of volunteered geographic information“. In: *GeoJournal* 72.3-4, S. 137–148. DOI: [10.1007/s10708-008-9188-y](https://doi.org/10.1007/s10708-008-9188-y).
- Flacke, J., Schüle, S. A., Köckler, H. und Bolte, G. (2016). „Mapping environmental inequalities relevant for health for informing urban planning interventions – A case study in the city of Dortmund, Germany“. In: *International journal of environmental research and public health* 13.7, S. 711. DOI: [10.3390/ijerph13070711](https://doi.org/10.3390/ijerph13070711).
- Fotheringham, A. S. und Wong, D. W. S. (1991). „The Modifiable Areal Unit Problem in Multivariate Statistical Analysis“. In: *Environment and Planning A* 23.7, S. 1025–1044. DOI: [10.1068/a231025](https://doi.org/10.1068/a231025).
- Fritz, S., See, L., Perger, C., McCallum, I., Schill, C., Schepaschenko, D., Duerauer, M., Karner, M., Dresel, C., Laso-Bayas, J.-C., Lesiv, M., Moorthy, I., Salk, C. F., Danylo, O., Sturn, T., Albrecht, F., You, L., Kraxner, F. und Obersteiner, M. (2017). „A global dataset of crowdsourced land cover and land use reference data“. In: *Scientific Data* 4, S. 170075. DOI: [10.1038/sdata.2017.75](https://doi.org/10.1038/sdata.2017.75).
- Fukushima, K. (1969). „Visual Feature Extraction by a Multilayered Network of Analog Threshold Elements“. In: *IEEE Transactions on Systems Science and Cybernetics* 5.4, S. 322–333. DOI: [10.1109/tssc.1969.300225](https://doi.org/10.1109/tssc.1969.300225).
- Fukushima, K. (1980). „Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position“. In: *Biological Cybernetics* 36.4, S. 193–202. DOI: [10.1007/bf00344251](https://doi.org/10.1007/bf00344251).
- Gallego, J. und Delincé, J. (2010). „The European Land Use and Cover Area-Frame Statistical Survey“. In: *Agricultural Survey Methods*. Hrsg. von R. Benedetti, M. Bee, G. Espa und F. Piersimoni. Chichester: John Wiley & Sons, Ltd. Kap. 10, S. 149–168. DOI: [10.1002/9780470665480.ch10](https://doi.org/10.1002/9780470665480.ch10).
- Geiß, C., Pelizari, P. A., Marconcini, M., Sengara, W., Edwards, M., Lakes, T. und Taubenböck, H. (2015). „Estimation of seismic building structural types using multi-sensor remote sensing and machine learning techniques“. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 104, S. 175–188. DOI: [10.1016/j.isprsjprs.2014.07.016](https://doi.org/10.1016/j.isprsjprs.2014.07.016).
- Geiß, C., Aravena Pelizari, P., Blickensdörfer, L. und Taubenböck, H. (2019). „Virtual Support Vector Machines with self-learning strategy for classification of multispectral remote sensing imagery“. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 151, S. 42–58. DOI: [10.1016/j.isprsjprs.2019.03.001](https://doi.org/10.1016/j.isprsjprs.2019.03.001).
- Geiß, C., Zhu, Y., Qiu, C., Mou, L., Zhu, X. X. und Taubenböck, H. (2020). „Deep Relearning in the Geospatial Domain for Semantic Remote Sensing Image Segmentation“. In: *IEEE Geoscience and Remote Sensing Letters*, S. 1–5. DOI: [10.1109/lgrs.2020.3031339](https://doi.org/10.1109/lgrs.2020.3031339).
- Georganos, S., Hafner, S., Kuffer, M., Linard, C. und Ban, Y. (2022). „A census from heaven: Unraveling the potential of deep learning and Earth Observation for intra-urban population mapping in data scarce environments“. In: *International Journal of Applied Earth Observation and Geoinformation* 114, S. 103013. DOI: [10.1016/j.jag.2022.103013](https://doi.org/10.1016/j.jag.2022.103013).

- Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow Concepts, Tools, and Techniques to Build Intelligent Systems. Concepts, Tools, and Techniques to Build Intelligent Systems*. 2. Aufl. Sebastopol, CA: O'Reilly Media. 856 S.
- Gómez-Baggethun, E., Gren, Å., Barton, D. N., Langemeyer, J., McPhearson, T., O'Farrell, P., Andersson, E., Hamstead, Z. und Kremer, P. (2013). „Urban Ecosystem Services“. In: *Urbanization, Biodiversity and Ecosystem Services: Challenges and Opportunities*. Hrsg. von T. Elmqvist, M. Fragkakis, J. Goodness, B. Güneralp, P. J. Marcotullio, R. I. McDonald, S. Parnell, M. Schewenius, M. Sendstad, K. C. Seto und C. Wilkinson. Dordrecht: Springer, S. 175–251. DOI: [10.1007/978-94-007-7088-1_11](https://doi.org/10.1007/978-94-007-7088-1_11).
- Gómez, C., White, J. C. und Wulder, M. A. (2016). „Optical remotely sensed time series data for land cover classification: A review“. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 116, S. 55–72. DOI: [10.1016/j.isprsjprs.2016.03.008](https://doi.org/10.1016/j.isprsjprs.2016.03.008).
- Goodfellow, I., Bengio, Y. und Courville, A. (2016). *Deep Learning*. Cambridge, MA: MIT Press. 775 S.
- Goodchild, M. F. (2007). „Citizens as sensors: the world of volunteered geography“. In: *GeoJournal* 69.4, S. 211–221. DOI: [10.1007/s10708-007-9111-y](https://doi.org/10.1007/s10708-007-9111-y).
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. und Moore, R. (2017). „Google Earth Engine: Planetary-scale geospatial analysis for everyone“. In: *Remote Sensing of Environment* 202, S. 18–27. DOI: [10.1016/j.rse.2017.06.031](https://doi.org/10.1016/j.rse.2017.06.031).
- Griffiths, P., Müller, D., Kuemmerle, T. und Hostert, P. (2013). „Agricultural land change in the Carpathian ecoregion after the breakdown of socialism and expansion of the European Union“. In: *Environmental Research Letters* 8.4, S. 045024. DOI: [10.1088/1748-9326/8/4/045024](https://doi.org/10.1088/1748-9326/8/4/045024).
- Griffiths, P., Kuemmerle, T., Baumann, M., Radeloff, V. C., Abrudan, I. V., Lieskovsky, J., Munteanu, C., Ostapowicz, K. und Hostert, P. (2014). „Forest disturbances, forest recovery, and changes in forest types across the Carpathian ecoregion from 1985 to 2010 based on Landsat image composites“. In: *Remote Sensing of Environment* 151, S. 72–88. DOI: [10.1016/j.rse.2013.04.022](https://doi.org/10.1016/j.rse.2013.04.022).
- Griffiths, P., Nendel, C. und Hostert, P. (2019). „Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping“. In: *Remote Sensing of Environment* 220, S. 135–151. DOI: [10.1016/j.rse.2018.10.031](https://doi.org/10.1016/j.rse.2018.10.031).
- Groenewegen, P. P., Van den Berg, A. E., De Vries, S. und Verheij, R. A. (2006). „Vitamin G: effects of green space on health, well-being, and social safety“. In: *BMC public health* 6.1, S. 49. DOI: [10.1186/1471-2458-6-149](https://doi.org/10.1186/1471-2458-6-149).
- Haase, D., Kabisch, S., Haase, A., Andersson, E., Banzhaf, E., Baró, F., Brenck, M., Fischer, L., Frantzeskaki, N., Kabisch, N., Krellenberg, K., Kremer, P., Kronenberg, J., Larondelle, N., Mathey, J., Pauleit, S., Ring, I., Rink, D., Schwarz, N. und Wolff, M. (2017). „Greening cities – To be socially inclusive? About the alleged paradox of society and ecology in cities“. In: *Habitat International* 64, S. 41–48. DOI: [10.1016/j.habitatint.2017.04.005](https://doi.org/10.1016/j.habitatint.2017.04.005).

- Ham, J., Chen, Y., Crawford, M. und Ghosh, J. (2005). „Investigation of the random forest framework for classification of hyperspectral data“. In: *IEEE Transactions on Geoscience and Remote Sensing* 43.3, S. 492–501. DOI: [10.1109/tgrs.2004.842481](https://doi.org/10.1109/tgrs.2004.842481).
- Hansmann, R., Hug, S.-M. und Seeland, K. (2007). „Restoration and stress relief through physical activities in forests and parks“. In: *Urban Forestry & Urban Greening* 6.4, S. 213–225. DOI: [10.1016/j.ufug.2007.08.004](https://doi.org/10.1016/j.ufug.2007.08.004).
- Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O. und Townshend, J. R. G. (2013). „High-Resolution Global Maps of 21st-Century Forest Cover Change“. In: *Science* 342.6160, S. 850–853. DOI: [10.1126/science.1244693](https://doi.org/10.1126/science.1244693).
- Hartig, T., Mitchell, R., Vries, S. de und Frumkin, H. (2014). „Nature and Health“. In: *Annual Review of Public Health* 35.1, S. 207–228. DOI: [10.1146/annurev-publhealth-032013-182443](https://doi.org/10.1146/annurev-publhealth-032013-182443).
- Harris, R. und Baumann, I. (2015). „Open data policies and satellite Earth observation“. In: *Space Policy* 32, S. 44–53. DOI: [10.1016/j.spacepol.2015.01.001](https://doi.org/10.1016/j.spacepol.2015.01.001).
- Haralick, R. M., Shanmugam, K. und Dinstein, I. (1973). „Textural Features for Image Classification“. In: *IEEE Transactions on Systems, Man, and Cybernetics* SMC-3.6, S. 610–621. DOI: [10.1109/tsmc.1973.4309314](https://doi.org/10.1109/tsmc.1973.4309314).
- Haralick, R. (1979). „Statistical and structural approaches to texture“. In: *Proceedings of the IEEE* 67.5, S. 786–804. DOI: [10.1109/proc.1979.11328](https://doi.org/10.1109/proc.1979.11328).
- Hastie, T., Tibshirani, R. und Friedman, J. H. (2003). *The Elements of Statistical Learning*. Springer, S. 549.
- He, K., Zhang, X., Ren, S. und Sun, J. (2016). „Deep residual learning for image recognition“. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, S. 770–778. DOI: [10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90).
- He, D., Miao, J., Lu, Y., Song, Y., Chen, L. und Liu, Y. (2022). „Urban greenery mitigates the negative effect of urban density on older adults’ life satisfaction: Evidence from Shanghai, China“. In: *Cities* 124, S. 103607. DOI: [10.1016/j.cities.2022.103607](https://doi.org/10.1016/j.cities.2022.103607).
- Helber, P., Bischke, B., Dengel, A. und Borth, D. (2019). „EuroSAT: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification“. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 12.7, S. 2217–2226. DOI: [10.1109/jstars.2019.2918242](https://doi.org/10.1109/jstars.2019.2918242).
- Helbich, M., O’Connor, R. C., Nieuwenhuijsen, M. und Hagedoorn, P. (2020). „Greenery exposure and suicide mortality later in life: A longitudinal register-based case-control study“. In: *Environment International* 143, S. 105982. DOI: [10.1016/j.envint.2020.105982](https://doi.org/10.1016/j.envint.2020.105982).
- Helbich, M., Poppe, R., Oberski, D., Emmichoven, M. Z. van und Schram, R. (2021). „Can’t see the wood for the trees? An assessment of street view- and satellite-derived greenness measures in relation to mental health“. In: *Landscape and Urban Planning* 214, S. 104181. DOI: [10.1016/j.landurbplan.2021.104181](https://doi.org/10.1016/j.landurbplan.2021.104181).

- Henry, C., Hellekes, J., Merkle, N., Azimi, S. M. und Kurz, F. (2021). „Citywide Estimation of Parking Space Using Aerial Imagery and OSM Data Fusion with Deep Learning and Fine-grained Annotation“. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLIII-B2-2021, S. 479–485. DOI: [10.5194/isprs-archives-xliii-b2-2021-479-2021](https://doi.org/10.5194/isprs-archives-xliii-b2-2021-479-2021).
- Herrera, R., Markevych, I., Berger, U., Genuneit, J., Gerlich, J., Nowak, D., Schlotz, W., Vogelberg, C., Mutius, E. von, Weinmayr, G., Windstetter, D., Weigl, M., Heinrich, J. und Radon, K. (2018). „Greenness and job-related chronic stress in young adults: a prospective cohort study in Germany“. In: *BMJ Open* 8.6, e021599. DOI: [10.1136/bmjopen-2018-021599](https://doi.org/10.1136/bmjopen-2018-021599).
- Hewitt, C. N., Ashworth, K. und MacKenzie, A. R. (2019). „Using green infrastructure to improve urban air quality (GI4AQ)“. In: *Ambio* 49.1, S. 62–73. DOI: [10.1007/s13280-019-01164-3](https://doi.org/10.1007/s13280-019-01164-3).
- Heynen, N., Perkins, H. A. und Roy, P. (2006). „The political ecology of uneven urban green space: the impact of political economy on race and ethnicity in producing environmental inequality in Milwaukee“. In: *Urban Affairs Review* 42.1, S. 3–25. DOI: [10.1177/1078087406290729](https://doi.org/10.1177/1078087406290729).
- Honold, J., Lakes, T., Beyer, R. und Meer, E. van der (2016). „Restoration in Urban Spaces: Nature Views From Home, Greenways, and Public Parks“. In: *Environment and Behavior* 48.6, S. 796–825. DOI: [10.1177/0013916514568556](https://doi.org/10.1177/0013916514568556).
- Houborg, R., Fisher, J. B. und Skidmore, A. K. (2015). „Advances in remote sensing of vegetation function and traits“. In: *International Journal of Applied Earth Observation and Geoinformation* 43, S. 1–6. DOI: [10.1016/j.jag.2015.06.001](https://doi.org/10.1016/j.jag.2015.06.001).
- Huang, X., Zhang, L. und Li, P. (2007). „Classification and Extraction of Spatial Features in Urban Areas Using High-Resolution Multispectral Imagery“. In: *IEEE Geoscience and Remote Sensing Letters* 4.2, S. 260–264. DOI: [10.1109/lgrs.2006.890540](https://doi.org/10.1109/lgrs.2006.890540).
- Hua, Y., Marcos, D., Mou, L., Zhu, X. X. und Tuia, D. (2022). „Semantic Segmentation of Remote Sensing Images With Sparse Annotations“. In: *IEEE Geoscience and Remote Sensing Letters* 19, S. 1–5. DOI: [10.1109/lgrs.2021.3051053](https://doi.org/10.1109/lgrs.2021.3051053).
- Huber, P. J. (1964). „Robust Estimation of a Location Parameter“. In: *The Annals of Mathematical Statistics* 35.1, S. 73–101. DOI: [10.1214/aoms/1177703732](https://doi.org/10.1214/aoms/1177703732).
- Huete, A. R. (2012). „Vegetation Indices, Remote Sensing and Forest Monitoring“. In: *Geography Compass* 6.9, S. 513–532. DOI: [10.1111/j.1749-8198.2012.00507.x](https://doi.org/10.1111/j.1749-8198.2012.00507.x).
- Hunter, A. J. und Luck, G. W. (2015). „Defining and measuring the social-ecological quality of urban greenspace: a semi-systematic review“. In: *Urban Ecosystems* 18.4, S. 1139–1163. DOI: [10.1007/s11252-015-0456-6](https://doi.org/10.1007/s11252-015-0456-6).
- Idso, S. B., Jackson, R. D. und Reginato, R. J. (1977). „Remote-Sensing of Crop Yields“. In: *Science* 196.4285, S. 19–25. DOI: [10.1126/science.196.4285.19](https://doi.org/10.1126/science.196.4285.19).
- Ioffe, S. und Szegedy, C. (2015). „Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift“. In: *Proceedings of the 32nd International Conference on Machine Learning*. PMLR, S. 448–456.
- Ismayilova, I. und Timpf, S. (2022). „Towards an Ontology of Urban Green Spaces“. In: *GI_Forum 2022* 2, S. 47–57. DOI: [10.1553/giscience2022_02_s47](https://doi.org/10.1553/giscience2022_02_s47).

- Jarvis, I., Gergel, S., Koehoorn, M. und Bosch, M. van den (2020). „Greenspace access does not correspond to nature exposure: Measures of urban natural space with implications for health research“. In: *Landscape and Urban Planning* 194, S. 103686. DOI: [10.1016/j.landurbplan.2019.103686](https://doi.org/10.1016/j.landurbplan.2019.103686).
- Jenerette, G. D., Harlan, S. L., Stefanov, W. L. und Martin, C. A. (2011). „Ecosystem services and urban heat riskscape moderation: water, green spaces, and social inequality in Phoenix, USA“. In: *Ecological Applications* 21.7, S. 2637–2651. DOI: [10.1890/10-1493.1](https://doi.org/10.1890/10-1493.1).
- Jennings, V., Gaither, C. J. und Gragg, R. S. (2012). „Promoting Environmental Justice Through Urban Green Space Access: A Synopsis“. In: *Environmental Justice* 5.1, S. 1–7. DOI: [10.1089/env.2011.0007](https://doi.org/10.1089/env.2011.0007).
- Jiang, B., Deal, B., Pan, H., Larsen, L., Hsieh, C.-H., Chang, C.-Y. und Sullivan, W. C. (2017). „Remotely-sensed imagery vs. eye-level photography: Evaluating associations among measurements of tree cover density“. In: *Landscape and Urban Planning* 157, S. 270–281. DOI: [10.1016/j.landurbplan.2016.07.010](https://doi.org/10.1016/j.landurbplan.2016.07.010).
- Jimenez, R. B., Lane, K. J., Hutyra, L. R. und Fabian, M. P. (2022). „Spatial resolution of Normalized Difference Vegetation Index and greenness exposure misclassification in an urban cohort“. In: *Journal of Exposure Science & Environmental Epidemiology* 32.2, S. 213–222. DOI: [10.1038/s41370-022-00409-w](https://doi.org/10.1038/s41370-022-00409-w).
- Jones, H. G. und Vaughan, R. A. (2010). *Remote Sensing of Vegetation: Principles, Techniques, and Applications. Principles, Techniques, and Applications*. 1. Aufl. Oxford: Oxford University Press, S. 353.
- Jorgensen, A. und Gobster, P. H. (2010). „Shades of Green: Measuring the Ecology of Urban Green Space in the Context of Human Health and Well-Being“. In: *Nature and Culture* 5.3, S. 338–363. DOI: [10.3167/nc.2010.050307](https://doi.org/10.3167/nc.2010.050307).
- Jünger, S. (2021). „Land use disadvantages in Germany: A matter of ethnic income inequalities?“ In: *Urban Studies* 59.9, S. 1819–1836. DOI: [10.1177/00420980211023206](https://doi.org/10.1177/00420980211023206).
- Kabisch, N. und Haase, D. (2014). „Green justice or just green? Provision of urban green spaces in Berlin, Germany“. In: *Landscape and Urban Planning* 122, S. 129–139. DOI: [10.1016/j.landurbplan.2013.11.016](https://doi.org/10.1016/j.landurbplan.2013.11.016).
- Kabisch, N., Strohbach, M., Haase, D. und Kronenberg, J. (2016). „Urban green space availability in European cities“. In: *Ecological Indicators* 70, S. 586–596. DOI: [10.1016/j.ecolind.2016.02.029](https://doi.org/10.1016/j.ecolind.2016.02.029).
- Kabisch, N., Püffel, C., Masztalerz, O., Hemmerling, J. und Kraemer, R. (2021). „Physiological and psychological effects of visits to different urban green and street environments in older people: A field experiment in a dense inner-city area“. In: *Landscape and Urban Planning* 207, S. 103998. DOI: [10.1016/j.landurbplan.2020.103998](https://doi.org/10.1016/j.landurbplan.2020.103998).
- Kabisch, N. (2019). „The Influence of Socio-economic and Socio-demographic Factors in the Association Between Urban Green Space and Health“. In: *Biodiversity and Health in the Face of Climate Change*. Hrsg. von M. R. Marselle, J. Stadler, H. Korn, K. N. Irvine und A. Bonn. Cham: Springer International Publishing, S. 91–119. DOI: [10.1007/978-3-030-02318-8_5](https://doi.org/10.1007/978-3-030-02318-8_5).

- Karydas, C., Gitas, I., Kuntz, S. und Minakou, C. (2015). „Use of LUCAS LC Point Database for Validating Country-Scale Land Cover Maps“. In: *Remote Sensing* 7.5, S. 5012–5041. DOI: [10.3390/rs70505012](https://doi.org/10.3390/rs70505012).
- Karasiak, N., Dejoux, J.-F., Monteil, C. und Sheeren, D. (2021). „Spatial dependence between training and test sets: another pitfall of classification accuracy assessment in remote sensing“. In: *Machine Learning*. DOI: [10.1007/s10994-021-05972-1](https://doi.org/10.1007/s10994-021-05972-1).
- Kawachi, I. und Kennedy, B. P. (1997). „The relationship of income inequality to mortality: Does the choice of indicator matter?“ In: *Social Science & Medicine* 45.7, S. 1121–1127. DOI: [10.1016/s0277-9536\(97\)00044-0](https://doi.org/10.1016/s0277-9536(97)00044-0).
- Kemker, R., Salvaggio, C. und Kanan, C. (2018). „Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning“. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 145, S. 60–77. DOI: [10.1016/j.isprsjprs.2018.04.014](https://doi.org/10.1016/j.isprsjprs.2018.04.014).
- Khatami, R., Mountrakis, G. und Stehman, S. V. (2016). „A meta-analysis of remote sensing research on supervised pixel-based land-cover image classification processes: General guidelines for practitioners and future research“. In: *Remote Sensing of Environment* 177, S. 89–100. DOI: [10.1016/j.rse.2016.02.028](https://doi.org/10.1016/j.rse.2016.02.028).
- Khaliq, A., Musci, M. A. und Chiaberge, M. (2018). „Understanding effects of atmospheric variables on spectral vegetation indices derived from satellite based time series of multispectral images“. In: *2018 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*. IEEE. DOI: [10.1109/aipr.2018.8707430](https://doi.org/10.1109/aipr.2018.8707430).
- Klompaker, J. O., Hoek, G., Bloemsa, L. D., Gehring, U., Strak, M., Wijga, A. H., Brink, C. van den, Brunekreef, B., Lebret, E. und Janssen, N. A. (2018). „Green space definition affects associations of green space with overweight and physical activity“. In: *Environmental Research* 160, S. 531–540. DOI: [10.1016/j.envres.2017.10.027](https://doi.org/10.1016/j.envres.2017.10.027).
- Knoble, C. und Yu, D. (2023). „Environmental justice: An evolving concept in a dynamic era“. In: *Sustainable Development*, S. 1–18. DOI: [10.1002/sd.2519](https://doi.org/10.1002/sd.2519).
- Köckler, H. (2017). *Umweltbezogene Gerechtigkeit – Anforderungen an eine zukunftsweisende Stadtplanung*. Hrsg. von U. Altröck und H. Kegler. Frankfurt a. M.: PL Academic Research. 248 S.
- Kramer, H. J. (2002). *Observation of the Earth and Its Environment. Survey of Missions and Sensors*. 4. Aufl. Berlin, Heidelberg: Springer. 1509 S. DOI: [10.1007/978-3-642-56294-5](https://doi.org/10.1007/978-3-642-56294-5).
- Krekel, C., Kolbe, J. und Wüstemann, H. (2016). „The greener, the happier? The effect of urban land use on residential well-being“. In: *Ecological Economics* 121, S. 117–127. DOI: [10.1016/j.ecolecon.2015.11.005](https://doi.org/10.1016/j.ecolecon.2015.11.005).
- Krizhevsky, A., Sutskever, I. und Hinton, G. E. (2012). „ImageNet classification with deep convolutional neural networks“. In: *NIPS'12: Proceedings of the 25th International Conference on Neural Information Processing Systems*. Hrsg. von F. Pereira, C. J. C. Burges, L. Bottou und K. Q. Weinberger. Bd. 1. Red Hook, NY: Curran Associates Inc., S. 1097–1105. DOI: [10.1145/3065386](https://doi.org/10.1145/3065386).

- Kussul, N., Mykola, L., Shelestov, A. und Skakun, S. (2018). „Crop inventory at regional scale in Ukraine: developing in season and end of season crop maps with multi-temporal optical and SAR satellite imagery“. In: *European Journal of Remote Sensing* 51.1, S. 627–636. DOI: [10.1080/22797254.2018.1454265](https://doi.org/10.1080/22797254.2018.1454265).
- Lachowycz, K. und Jones, A. P. (2013). „Towards a better understanding of the relationship between greenspace and health: Development of a theoretical framework“. In: *Landscape and Urban Planning* 118, S. 62–69. DOI: [10.1016/j.landurbplan.2012.10.012](https://doi.org/10.1016/j.landurbplan.2012.10.012).
- Lakes, T., Brückner, M. und Krämer, A. (2013). „Development of an environmental justice index to determine socio-economic disparities of noise pollution and green space in residential areas in Berlin“. In: *Journal of Environmental Planning and Management* 57.4, S. 538–556. DOI: [10.1080/09640568.2012.755461](https://doi.org/10.1080/09640568.2012.755461).
- Lary, D. J., Alavi, A. H., Gandomi, A. H. und Walker, A. L. (2016). „Machine learning in geosciences and remote sensing“. In: *Geoscience Frontiers* 7.1, S. 3–10. DOI: [10.1016/j.gsf.2015.07.003](https://doi.org/10.1016/j.gsf.2015.07.003).
- LeCun, Y., Bengio, Y. und Hinton, G. (2015). „Deep learning“. In: *Nature* 521.7553, S. 436–444. DOI: [10.1038/nature14539](https://doi.org/10.1038/nature14539).
- LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., Hubbard, W. und Jackel, L. (1990). „Handwritten Digit Recognition with a Back-Propagation Network“. In: *Advances in Neural Information Processing Systems 2*. Hrsg. von D. Touretzky. Bd. 2. NIPS 1989. San Francisco, CA: Morgan Kaufmann Publishers, S. 396–404. DOI: [10.5555/109230.109279](https://doi.org/10.5555/109230.109279).
- Lecun, Y., Bottou, L., Bengio, Y. und Haffner, P. (1998). „Gradient-based learning applied to document recognition“. In: *Proceedings of the IEEE* 86.11, S. 2278–2324. DOI: [10.1109/5.726791](https://doi.org/10.1109/5.726791).
- Lee, A. C. K. und Maheswaran, R. (2011). „The health benefits of urban green spaces: a review of the evidence“. In: *Journal of Public Health* 33.2, S. 212–222. DOI: [10.1093/pubmed/fdq068](https://doi.org/10.1093/pubmed/fdq068).
- Leinenkugel, P., Deck, R., Huth, J., Ottinger, M. und Mack, B. (2019). „The Potential of Open Geodata for Automated Large-Scale Land Use and Land Cover Classification“. In: *Remote Sensing* 11.19, S. 2249. DOI: [10.3390/rs11192249](https://doi.org/10.3390/rs11192249).
- Leichtle, T., Zehner, M., Kühnl, M., Martin, K. und Taubenböck, H. (2021). „Urban Trees – Detection, Delineation, Quantification, and Characterisation based on VHR Remote Sensing“. In: *CITIES 20.50 – Creating Habitats for the 3rd Millennium: Smart – Sustainable – Climate Neutral. Proceedings of REAL CORP 2021, 26th International Conference on Urban Development, Regional Planning and Information Society*. Hrsg. von M. Schrenk, V. V. Popovich, P. Ziele, P. Elisei, C. Beyer, J. Ryser und G. Stöglehner, S. 1029–1039.
- Leichtle, T., Kühnl, M., Droin, A., Beck, C., Hiete, M. und Taubenböck, H. (2023). „Quantifying urban heat exposure at fine scale - modeling outdoor and indoor temperatures using citizen science and VHR remote sensing“. In: *Urban Climate* 49, S. 101522. DOI: [10.1016/j.uclim.2023.101522](https://doi.org/10.1016/j.uclim.2023.101522).

- Levy, J. I., Chemerynski, S. M. und Tuchmann, J. L. (2006). „Incorporating concepts of inequality and inequity into health benefits analysis“. In: *International Journal for Equity in Health* 5.1, S. 19. DOI: [10.1186/1475-9276-5-2](https://doi.org/10.1186/1475-9276-5-2).
- Li, M., Zang, S., Zhang, B., Li, S. und Wu, C. (2014). „A Review of Remote Sensing Image Classification Techniques: the Role of Spatio-contextual Information“. In: *European Journal of Remote Sensing* 47.1, S. 389–411. DOI: [10.5721/eujrs20144723](https://doi.org/10.5721/eujrs20144723).
- Liang, L. und Schwartz, M. D. (2009). „Landscape phenology: an integrative approach to seasonal vegetation dynamics“. In: *Landscape Ecology* 24.4, S. 465–472. DOI: [10.1007/s10980-009-9328-x](https://doi.org/10.1007/s10980-009-9328-x).
- Lin, B. B., Fuller, R. A., Bush, R., Gaston, K. J. und Shanahan, D. F. (2014). „Opportunity or Orientation? Who Uses Urban Parks and Why“. In: *PLoS ONE* 9.1. Hrsg. von A. M. Merenlender, e87422. DOI: [10.1371/journal.pone.0087422](https://doi.org/10.1371/journal.pone.0087422).
- Liu, Y. und Gastwirth, J. L. (2020). „On the capacity of the Gini index to represent income distributions“. In: *METRON* 78.1, S. 61–69. DOI: [10.1007/s40300-020-00164-8](https://doi.org/10.1007/s40300-020-00164-8).
- Low, S. (2013). „The Routledge Research Companion to Planning and Culture“. In: Hrsg. von G. Young und D. Severson. 1. Aufl. London: Routledge. Kap. Public Space and Diversity: Distributive, Procedural and Interactional Justice for Parks, S. 295–310.
- Ludwig, C., Hecht, R., Lautenbach, S., Schorcht, M. und Zipf, A. (2019). „Assessing the Completeness of Urban Green Spaces in OpenStreetMap“. In: *Proceedings of the Academic Track, State of the Map 2019, 21-22, Heidelberg*. Hrsg. von M. Minghini, A. Y. Grinberger, L. Juhász, G. Yeboah und P. Mooney. Zenodo, S. 21–22. DOI: [10.5281/ZENODO.3387701](https://doi.org/10.5281/ZENODO.3387701).
- Ludwig, C., Hecht, R., Lautenbach, S., Schorcht, M. und Zipf, A. (2021). „Mapping Public Urban Green Spaces Based on OpenStreetMap and Sentinel-2 Imagery Using Belief Functions“. In: *ISPRS International Journal of Geo-Information* 10.251, S. 1–25. DOI: [10.3390/ijgi10040251](https://doi.org/10.3390/ijgi10040251).
- Lwasa, S., Seto, K. C., Bai, X., Blanco, H., Gurney, K. R., Kilkış, Ş., Lucon, O., Murakami, J., Pan, J., Sharifi, A. und Yamagata, Y. (2022). „Urban Systems and Other Settlements“. In: *IPCC 2022, Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Hrsg. von P. Shukla, J. Skea, R. Slade, A. A. Khoualdjia, R. van Diemen, D. McCollum, M. Pathak, S. Some, P. Vyas, R. Fradera, M. Belkacemi, A. Hasija, G. Lisboa, S. Luz und J. Malley. Cambridge, UK und New York, NY, USA: Cambridge University Press. Kap. 8, S. 861–952. DOI: [10.1017/9781009157926.010](https://doi.org/10.1017/9781009157926.010).
- Ma, L., Li, M., Ma, X., Cheng, L., Du, P. und Liu, Y. (2017). „A review of supervised object-based land-cover image classification“. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 130, S. 277–293. DOI: [10.1016/j.isprsjprs.2017.06.001](https://doi.org/10.1016/j.isprsjprs.2017.06.001).
- Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G. und Johnson, B. A. (2019). „Deep learning in remote sensing applications: A meta-analysis and review“. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 152, S. 166–177. DOI: [10.1016/j.isprsjprs.2019.04.015](https://doi.org/10.1016/j.isprsjprs.2019.04.015).

- Ma, A., Wan, Y., Zhong, Y., Wang, J. und Zhang, L. (2021). „SceneNet: Remote sensing scene classification deep learning network using multi-objective neural evolution architecture search“. In: 172, S. 171–188. DOI: [10.1016/j.isprsjprs.2020.11.025](https://doi.org/10.1016/j.isprsjprs.2020.11.025).
- Maas, J., Verheij, R. A., Groenewegen, P. P., De Vries, S. und Spreeuwenberg, P. (2006). „Green space, urbanity, and health: how strong is the relation?“ In: *Journal of Epidemiology & Community Health* 60.7, S. 587–592. DOI: [10.1136/jech.2005.043125](https://doi.org/10.1136/jech.2005.043125).
- Maas, J., Verheij, R. A., Vries, S. de, Spreeuwenberg, P., Schellevis, F. G. und Groenewegen, P. P. (2009). „Morbidity is related to a green living environment“. In: *Journal of Epidemiology & Community Health* 63.12, S. 967–973. DOI: [10.1136/jech.2008.079038](https://doi.org/10.1136/jech.2008.079038).
- Mack, B., Leinenkugel, P., Kuenzer, C. und Dech, S. (2017). „A semi-automated approach for the generation of a new land use and land cover product for Germany based on Landsat time-series and Lucas in-situ data“. In: *Remote Sensing Letters* 8.3, S. 244–253. DOI: [10.1080/2150704x.2016.1249299](https://doi.org/10.1080/2150704x.2016.1249299).
- Maggiori, E., Tarabalka, Y., Charpiat, G. und Alliez, P. (2017). „Convolutional Neural Networks for Large-Scale Remote-Sensing Image Classification“. In: *IEEE Transactions on Geoscience and Remote Sensing* 55.2, S. 645–657. DOI: [10.1109/tgrs.2016.2612821](https://doi.org/10.1109/tgrs.2016.2612821).
- Marshall, J. D., Swor, K. R. und Nguyen, N. P. (2014). „Prioritizing Environmental Justice and Equality: Diesel Emissions in Southern California“. In: *Environmental Science & Technology* 48.7, S. 4063–4068. DOI: [10.1021/es405167f](https://doi.org/10.1021/es405167f).
- Massaro, E., Schifanella, R., Piccardo, M., Caporaso, L., Taubenböck, H., Cescatti, A. und Duveiller, G. (2023). „Spatially-optimized urban greening for reduction of population exposure to land surface temperature extremes“. In: *Nature Communications* 14.2903, S. 1–10. DOI: [10.1038/s41467-023-38596-1](https://doi.org/10.1038/s41467-023-38596-1).
- Maschewsky, W. (2001). *Umweltgerechtigkeit, Public Health und soziale Stadt*. Frankfurt a. Main: VAS. 238 S.
- Maxwell, A. E., Warner, T. A. und Fang, F. (2018). „Implementation of machine-learning classification in remote sensing: an applied review“. In: *International Journal of Remote Sensing* 39.9, S. 2784–2817. DOI: [10.1080/01431161.2018.1433343](https://doi.org/10.1080/01431161.2018.1433343).
- McFeeters, S. K. (1996). „The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features“. In: *International Journal of Remote Sensing* 17.7, S. 1425–1432. DOI: [10.1080/01431169608948714](https://doi.org/10.1080/01431169608948714).
- McGregor, T., Smith, B. und Wills, S. (2019). „Measuring inequality“. In: *Oxford Review of Economic Policy* 35.3, S. 368–395. DOI: [10.1093/oxrep/grz015](https://doi.org/10.1093/oxrep/grz015).
- Mears, M., Brindley, P., Jorgensen, A. und Maheswaran, R. (2020). „Population-level linkages between urban greenspace and health inequality: The case for using multiple indicators of neighbourhood greenspace“. In: *Health & Place* 62.102284, S. 1–11. DOI: [10.1016/j.healthplace.2020.102284](https://doi.org/10.1016/j.healthplace.2020.102284).

- Middel, A., Häb, K., Brazel, A. J., Martin, C. A. und Guhathakurta, S. (2014). „Impact of urban form and design on mid-afternoon microclimate in Phoenix Local Climate Zones“. In: *Landscape and Urban Planning* 122, S. 16–28. DOI: [10.1016/j.landurbplan.2013.11.004](https://doi.org/10.1016/j.landurbplan.2013.11.004).
- Mitchell, R. und Popham, F. (2007). „Greenspace, urbanity and health: relationships in England“. In: *Journal of Epidemiology & Community Health* 61.8, S. 681–683. DOI: [10.1136/jech.2006.053553](https://doi.org/10.1136/jech.2006.053553).
- Mitchell, R. und Popham, F. (2008). „Effect of exposure to natural environment on health inequalities: an observational population study“. In: *The Lancet* 372.9650, S. 1655–1660. DOI: [10.1016/S0140-6736\(08\)61689-X](https://doi.org/10.1016/S0140-6736(08)61689-X).
- Mitchell, T. M. (1997). *Machine Learning*. New York: McGraw-Hill.
- Morello-Frosch, R. und Shenassa, E. D. (2006). „The Environmental “Riskscape” and Social Inequality: Implications for Explaining Maternal and Child Health Disparities“. In: *Environmental Health Perspectives* 114.8, S. 1150–1153. DOI: [10.1289/ehp.8930](https://doi.org/10.1289/ehp.8930).
- Naushad, R., Kaur, T. und Ghaderpour, E. (2021). „Deep Transfer Learning for Land Use and Land Cover Classification: A Comparative Study“. In: *Sensors* 21.23, S. 8083. DOI: [10.3390/s21238083](https://doi.org/10.3390/s21238083).
- Neutze, M. (2015). „Gitterbasierte Auswertungen des Zensus 2011“. In: *Statistik & Informationsmanagement* 2, S. 64–67.
- Oehrlein, J., Niedermann, B. und Haurert, J.-H. (2019). „Analyzing the Supply and Detecting Spatial Patterns of Urban Green Spaces via Optimization“. In: *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, S. 137–158. DOI: [10.1007/s41064-019-00081-0](https://doi.org/10.1007/s41064-019-00081-0).
- Oke, T. R. (1973). „City size and the urban heat island“. In: *Atmospheric Environment (1967)* 7.8, S. 769–779. DOI: [10.1016/0004-6981\(73\)90140-6](https://doi.org/10.1016/0004-6981(73)90140-6).
- Oliveira, S., Andrade, H. und Vaz, T. (2011). „The cooling effect of green spaces as a contribution to the mitigation of urban heat: A case study in Lisbon“. In: *Building and Environment* 46.11, S. 2186–2194. DOI: [10.1016/j.buildenv.2011.04.034](https://doi.org/10.1016/j.buildenv.2011.04.034).
- Olofsson, P., Foody, G. M., Stehman, S. V. und Woodcock, C. E. (2013). „Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation“. In: *Remote Sensing of Environment* 129, S. 122–131. DOI: [10.1016/j.rse.2012.10.031](https://doi.org/10.1016/j.rse.2012.10.031).
- Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E. und Wulder, M. A. (2014). „Good practices for estimating area and assessing accuracy of land change“. In: *Remote Sensing of Environment* 148, S. 42–57. DOI: [10.1016/j.rse.2014.02.015](https://doi.org/10.1016/j.rse.2014.02.015).
- Openshaw, S. (1983). *The modifiable areal unit problem*. Concepts and techniques in modern geography 38. Norwich, England: Geo Books.
- Openshaw, S. (1984). „Ecological Fallacies and the Analysis of Areal Census Data“. In: *Environment and Planning A* 16.1, S. 17–31. DOI: [10.1068/a160017](https://doi.org/10.1068/a160017).

- Orgiazzi, A., Ballabio, C., Panagos, P., Jones, A. und Fernández-Ugalde, O. (2018). „LUCAS Soil, the largest expandable soil dataset for Europe: a review“. In: *European Journal of Soil Science* 69.1, S. 140–153. DOI: [10.1111/ejss.12499](https://doi.org/10.1111/ejss.12499).
- Park, J., Kim, J.-H., Lee, D. K., Park, C. Y. und Jeong, S. G. (2017). „The influence of small green space type and structure at the street level on urban heat island mitigation“. In: *Urban Forestry & Urban Greening* 21, S. 203–212. DOI: [10.1016/j.ufug.2016.12.005](https://doi.org/10.1016/j.ufug.2016.12.005).
- Pearce, J., Witten, K. und Bartie, P. (2006). „Neighbourhoods and health: a GIS approach to measuring community resource accessibility“. In: *Journal of epidemiology and community health* 60.5, S. 389–395. DOI: [10.1136/jech.2005.043281](https://doi.org/10.1136/jech.2005.043281).
- Pearsall, H. (2017). „Staying cool in the compact city: Vacant land and urban heating in Philadelphia, Pennsylvania“. In: *Applied Geography* 79, S. 84–92. DOI: [10.1016/j.apgeog.2016.12.010](https://doi.org/10.1016/j.apgeog.2016.12.010).
- Pelletier, C., Valero, S., Inglada, J., Champion, N. und Dedieu, G. (2016). „Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas“. In: *Remote Sensing of Environment* 187, S. 156–168. DOI: [10.1016/j.rse.2016.10.010](https://doi.org/10.1016/j.rse.2016.10.010).
- Pflugmacher, D., Rabe, A., Peters, M. und Hostert, P. (2019). „Mapping pan-European land cover using Landsat spectral-temporal metrics and the European LUCAS survey“. In: *Remote Sensing of Environment* 221, S. 583–595. DOI: [10.1016/j.rse.2018.12.001](https://doi.org/10.1016/j.rse.2018.12.001).
- Pham, T.-T.-H., Apparicio, P., Séguin, A.-M., Landry, S. und Gagnon, M. (2012). „Spatial distribution of vegetation in Montreal: An uneven distribution or environmental inequity?“ In: *Landscape and urban planning* 107, S. 214–224. DOI: [10.1016/j.landurbplan.2012.06.002](https://doi.org/10.1016/j.landurbplan.2012.06.002).
- Picavet, H. S. J., Milder, I., Kruize, H., Vries, S. de, Hermans, T. und Wendel-Vos, W. (2016). „Greener living environment healthier people?: Exploring green space, physical activity and health in the Doetinchem Cohort Study“. In: *Preventive Medicine* 89, S. 7–14. DOI: [10.1016/j.ypmed.2016.04.021](https://doi.org/10.1016/j.ypmed.2016.04.021).
- Poortinga, W., Bird, N., Hallingberg, B., Phillips, R. und Williams, D. (2021). „The role of perceived public and private green space in subjective health and wellbeing during and after the first peak of the COVID-19 outbreak“. In: *Landscape and Urban Planning* 211, S. 104092. DOI: [10.1016/j.landurbplan.2021.104092](https://doi.org/10.1016/j.landurbplan.2021.104092).
- Pope, D., Tisdall, R., Middleton, J., Verma, A., Ameijden, E. van, Birt, C., Macherianakis, A. und Bruce, N. (2015). „Quality of and access to green space in relation to psychological distress: results from a population-based cross-sectional study as part of the EURO-URHIS 2 project“. In: *The European Journal of Public Health*, S. 35–38. DOI: [10.1093/eurpub/ckv094](https://doi.org/10.1093/eurpub/ckv094).
- Prinzie, A. und Poel, D. V. den (2007). „Random Multiclass Classification: Generalizing Random Forests to Random MNL and Random NB“. In: *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, S. 349–358. DOI: [10.1007/978-3-540-74469-6_35](https://doi.org/10.1007/978-3-540-74469-6_35).

- Pugh, T. A. M., MacKenzie, A. R., Whyatt, J. D. und Hewitt, C. N. (2012). „Effectiveness of Green Infrastructure for Improvement of Air Quality in Urban Street Canyons“. In: *Environmental Science & Technology* 46.14, S. 7692–7699. DOI: [10.1021/es300826w](https://doi.org/10.1021/es300826w).
- Qiu, C., Schmitt, M., Geiß, C., Chen, T.-H. K. und Zhu, X. X. (2020a). „A framework for large-scale mapping of human settlement extent from Sentinel-2 images via fully convolutional neural networks“. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 163, S. 152–170. DOI: [10.1016/j.isprsjprs.2020.01.028](https://doi.org/10.1016/j.isprsjprs.2020.01.028).
- Qiu, C., Tong, X., Schmitt, M., Bechtel, B. und Zhu, X. X. (2020b). „Multilevel Feature Fusion-Based CNN for Local Climate Zone Classification From Sentinel-2 Images: Benchmark Results on the So2Sat LCZ42 Dataset“. In: *IEEE Journal of Selected Topics In Applied Earth Observation and Remote Sensing* 13, S. 2793–2806. DOI: [10.1109/jstars.2020.2995711](https://doi.org/10.1109/jstars.2020.2995711).
- Quinn, J. A., Nyhan, M. M., Navarro, C., Coluccia, D., Bromley, L. und Luengo-Oroz, M. (2018). „Humanitarian applications of machine learning with remote-sensing data: review and case study in refugee settlement mapping“. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376.2128, S. 20170363. DOI: [10.1098/rsta.2017.0363](https://doi.org/10.1098/rsta.2017.0363).
- Quinlan, J. R. (1986). „Induction of decision trees“. In: *Machine Learning* 1.1, S. 81–106. DOI: [10.1007/bf00116251](https://doi.org/10.1007/bf00116251).
- Rauch, S., Taubenböck, H., Knopp, C. und Rauh, J. (2021). „Risk and space: modelling the accessibility of stroke centers using day- & nighttime population distribution and different transportation scenarios“. In: *International Journal of Health Geographics* 20.1. DOI: [10.1186/s12942-021-00284-y](https://doi.org/10.1186/s12942-021-00284-y).
- Rehling, J., Bunge, C., Waldhauer, J. und Conrad, A. (2021). „Socioeconomic Differences in Walking Time of Children and Adolescents to Public Green Spaces in Urban Areas—Results of the German Environmental Survey (2014–2017)“. In: *International Journal of Environmental Research and Public Health* 18.5, S. 2326. DOI: [10.3390/ijerph18052326](https://doi.org/10.3390/ijerph18052326).
- Reid, C. E., Kubzansky, L. D., Li, J., Shmool, J. L. und Clougherty, J. E. (2018). „It’s not easy assessing greenness: A comparison of NDVI datasets and neighborhood types and their associations with self-rated health in New York City“. In: *Health & Place* 54, S. 92–101. DOI: [10.1016/j.healthplace.2018.09.005](https://doi.org/10.1016/j.healthplace.2018.09.005).
- Resnik, D. B. und Roman, G. (2007). „Health, Justice, and the Environment“. In: *Bioethics* 21.4, S. 230–241. DOI: [10.1111/j.1467-8519.2007.00547.x](https://doi.org/10.1111/j.1467-8519.2007.00547.x).
- Richardson, E. A. und Mitchell, R. (2010). „Gender differences in relationships between urban green space and health in the United Kingdom“. In: *Social Science & Medicine* 71.3, S. 568–575. DOI: [10.1016/j.socscimed.2010.04.015](https://doi.org/10.1016/j.socscimed.2010.04.015).
- Richardson, E., Pearce, J., Mitchell, R. und Kingham, S. (2013). „Role of physical activity in the relationship between urban green space and health“. In: *Public Health* 127.4, S. 318–324. DOI: [10.1016/j.puhe.2013.01.004](https://doi.org/10.1016/j.puhe.2013.01.004).
- Richter, R. und Schläpfer, D. (2016a). *ATCOR-2/3 User Guide, Version 9.0.2*. Techn. Ber. Wil, S. 1–263.

- Richter, R. und Schläpfer, D. (2016b). *Atmospheric and Topographic Correction - (ATCOR Theoretical Background Document)*. Techn. Ber. Wil, S. 1–146.
- Rigolon, A. und Németh, J. (2018). „We’re not in the business of housing:” Environmental gentrification and the nonprofitization of green infrastructure projects“. In: *Cities* 81, S. 71–80. DOI: [10.1016/j.cities.2018.03.016](https://doi.org/10.1016/j.cities.2018.03.016).
- Rigolon, A., Fernandez, M., Harris, B. und Stewart, W. (2019). „An Ecological Model of Environmental Justice for Recreation“. In: *Leisure Sciences* 44.6, S. 655–676. DOI: [10.1080/01490400.2019.1655686](https://doi.org/10.1080/01490400.2019.1655686).
- Rigolon, A. (2016). „A complex landscape of inequity in access to urban parks: A literature review“. In: *Landscape and Urban Planning* 153, S. 160–169. DOI: [10.1016/j.landurbplan.2016.05.017](https://doi.org/10.1016/j.landurbplan.2016.05.017).
- Robinson, W. S. (1950). „Ecological Correlations and the Behavior of Individuals“. In: *American Sociological Review* 15.3, S. 351–357. DOI: [10.1093/ije/dyn357](https://doi.org/10.1093/ije/dyn357).
- Rosier, J. F., Taubenböck, H., Verburg, P. H. und Vliet, J. van (2022). „Fusing Earth observation and socioeconomic data to increase the transferability of large-scale urban land use classification“. In: *Remote Sensing of Environment* 278, S. 113076. DOI: [10.1016/j.rse.2022.113076](https://doi.org/10.1016/j.rse.2022.113076).
- Rouse Jr., J. W., Haas, R. H., Schell, J. A. und Deering, D. W. (1974). „Monitoring vegetation systems in the Great Plains with ERTS“. In: *NASA. Goddard Space Flight Center 3d ERTS-1 Symp., Vol. 1, Sect. A*, S. 309–317.
- Russette, H., Graham, J., Holden, Z., Semmens, E. O., Williams, E. und Landguth, E. L. (2021). „Greenspace exposure and COVID-19 mortality in the United States: January–July 2020“. In: *Environmental Research* 198.111195, S. 1–7. DOI: [10.1016/j.envres.2021.111195](https://doi.org/10.1016/j.envres.2021.111195).
- Rutschmann, B. (2023). „Occurrence and population density of wild-living honey bees in Europe and the impact of different habitat types on their foraging and overwintering success“. doctoralthesis. Universität Würzburg. DOI: [10.25972/OPUS-28673](https://doi.org/10.25972/OPUS-28673).
- Sakieh, Y., Jaafari, S., Ahmadi, M. und Danekar, A. (2017). „Green and calm: Modeling the relationships between noise pollution propagation and spatial patterns of urban structures and green covers“. In: *Urban Forestry & Urban Greening* 24, S. 195–211. DOI: [10.1016/j.ufug.2017.04.008](https://doi.org/10.1016/j.ufug.2017.04.008).
- Sang, Å. O., Knez, I., Gunnarsson, B. und Hedblom, M. (2016). „The effects of naturalness, gender, and age on how urban green space is perceived and used“. In: *Urban Forestry & Urban Greening* 18, S. 268–276. DOI: [10.1016/j.ufug.2016.06.008](https://doi.org/10.1016/j.ufug.2016.06.008).
- Santos, T., Tenedório, J. und Gonçalves, J. (2016). „Quantifying the City’s Green Area Potential Gain Using Remote Sensing Data“. In: *Sustainability* 8.1247, S. 1–16. DOI: [10.3390/su8121247](https://doi.org/10.3390/su8121247).
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A. und Chen, L.-C. (2018). „MobileNetV2: Inverted Residuals and Linear Bottlenecks“. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

- Sánchez-Espinosa, A. und Schröder, C. (2019). „Land use and land cover mapping in wetlands one step closer to the ground: Sentinel-2 versus landsat 8“. In: *Journal of Environmental Management* 247, S. 484–498. DOI: [10.1016/j.jenvman.2019.06.084](https://doi.org/10.1016/j.jenvman.2019.06.084).
- Sapena, M., Wurm, M., Taubenböck, H., Tuia, D. und Ruiz, L. A. (2021). „Estimating quality of life dimensions from urban spatial pattern metrics“. In: *Computers, Environment and Urban Systems* 85.101549, S. 1–11. DOI: [10.1016/j.compenvurbsys.2020.101549](https://doi.org/10.1016/j.compenvurbsys.2020.101549).
- Schultz, M., Voss, J., Auer, M., Carter, S. und Zipf, A. (2017). „Open land cover from OpenStreetMap and remote sensing“. In: *International Journal of Applied Earth Observation and Geoinformation* 63, S. 206–213. DOI: [10.1016/j.jag.2017.07.014](https://doi.org/10.1016/j.jag.2017.07.014).
- Schüle, S. A., Hiliz, L. K., Dreger, S. und Bolte, G. (2019). „Social Inequalities in Environmental Resources of Green and Blue Spaces: A Review of Evidence in the WHO European Region“. In: *International Journal of Environmental Research and Public Health* 16.1216, S. 1–15. DOI: [10.3390/ijerph16071216](https://doi.org/10.3390/ijerph16071216).
- Schindler, M., Le Texier, M. und Caruso, G. (2022). „How far do people travel to use urban green space? A comparison of three European cities“. In: *Applied Geography* 141.102673, S. 1–11. DOI: [10.1016/j.apgeog.2022.102673](https://doi.org/10.1016/j.apgeog.2022.102673).
- Schröder-Bäck, P. (2012). „Ethische Kriterien der Gerechtigkeit für den Zusammenhang von Umwelt und Gesundheit“. In: *Umweltgerechtigkeit – Chancengleichheit bei Umwelt und Gesundheit: Konzepte, Datenlage und Handlungsperspektiven*. Hrsg. von G. Bolte, C. Bunge, C. Hornberg, H. Köckler und A. Mielck. Bern: Hans Huber, S. 61–72.
- Shahtahmassebi, A. R., Li, C., Fan, Y., Wu, Y., Lin, Y., Gan, M., Wang, K., Malik, A. und Blackburn, G. A. (2021). „Remote sensing of urban green spaces: A review“. In: *Urban Forestry & Urban Greening* 57, S. 126946. DOI: [10.1016/j.ufug.2020.126946](https://doi.org/10.1016/j.ufug.2020.126946).
- Shanmugan, K. S., Narayanan, V., Frost, V. S., Stiles, J. A. und Holtzman, J. C. (1981). „Textural Features for Radar Image Analysis“. In: *IEEE Transactions on Geoscience and Remote Sensing* GE-19.3, S. 153–156. DOI: [10.1109/tgrs.1981.350344](https://doi.org/10.1109/tgrs.1981.350344).
- Sheriff, G. und Maguire, K. B. (2020). „Health Risk, Inequality Indexes, and Environmental Justice“. In: *Risk Analysis* 40.12, S. 2661–2674. DOI: [10.1111/risa.13562](https://doi.org/10.1111/risa.13562).
- Sheykhmousa, M., Mahdianpari, M., Ghanbari, H., Mohammadimanesh, F., Ghamisi, P. und Homayouni, S. (2020). „Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review“. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 13, S. 6308–6325. DOI: [10.1109/jstars.2020.3026724](https://doi.org/10.1109/jstars.2020.3026724).
- Simonyan, K. und Zisserman, A. (2014). „Very Deep Convolutional Networks for Large-Scale Image Recognition“. In: S. 1–14. arXiv: [1409.1556](https://arxiv.org/abs/1409.1556) [cs.CV].
- Skakun, S., Kussul, N., Shelestov, A. Y., Lavreniuk, M. und Kussul, O. (2016). „Efficiency Assessment of Multitemporal C-Band Radarsat-2 Intensity and Landsat-8 Surface Reflectance Satellite Imagery for Crop Classification in Ukraine“. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 9.8, S. 3712–3719. DOI: [10.1109/jstars.2015.2454297](https://doi.org/10.1109/jstars.2015.2454297).

- Solimini, D. (2016). *Understanding Earth Observation*. Cham: Springer International Publishing. DOI: [10.1007/978-3-319-25633-7](https://doi.org/10.1007/978-3-319-25633-7).
- Soobader, M., Cubbin, C., Gee, G. C., Rosenbaum, A. und Laurenson, J. (2006). „Levels of analysis for the study of environmental health disparities“. In: *Environmental research* 102.2, S. 172–180. DOI: [10.1016/j.envres.2006.05.001](https://doi.org/10.1016/j.envres.2006.05.001).
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. und Salakhutdinov, R. (2014). „Dropout: A Simple Way to Prevent Neural Networks from Overfitting“. In: *Journal of Machine Learning Research* 15, S. 1929–1958. DOI: [10.5555/2627435.2670313](https://doi.org/10.5555/2627435.2670313).
- Srivastava, S., Vargas-Muñoz, J. E. und Tuia, D. (2019). „Understanding urban landuse from the above and ground perspectives: A deep learning, multimodal solution“. In: *Remote Sensing of Environment* 228, S. 129–143. DOI: [10.1016/j.rse.2019.04.014](https://doi.org/10.1016/j.rse.2019.04.014).
- Stark, T., Wurm, M., Zhu, X. X. und Taubenböck, H. (2020). „Satellite-Based Mapping of Urban Poverty With Transfer-Learned Slum Morphologies“. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 13, S. 5251–5263. DOI: [10.1109/jstars.2020.3018862](https://doi.org/10.1109/jstars.2020.3018862).
- Staab, J., Schady, A., Weigand, M., Lakes, T. und Taubenböck, H. (2022a). „Predicting traffic noise using land-use regression—a scalable approach“. In: *Journal of Exposure Science & Environmental Epidemiology*, S. 1–12. DOI: [10.1038/s41370-021-00355-z](https://doi.org/10.1038/s41370-021-00355-z).
- Staab, J., Schady, A., Wolf, K., Behzadi, S., Dallavalle, M., Weigand, M., Lakes, T. und Taubenböck, H. (2022b). „2000 land-use regressions for road traffic noise predictions – how sample selection affects extrapolation weights“. In: *PROCEEDINGS of the 24th International Congress on Acoustics*. ICA. Gyeongju, Korea, S. 1–13.
- Standfuß, I., Geiß, C., Kühnl, M., Droin, A., Mast, J., Wurm, M., Siedentop, S., Heider, B. und Taubenböck, H. (2023a). „In the tension between large-scale analysis and accuracy - Identifying and analysing intra-urban (sub-)centre structures comparing official 3D-building models and TanDEM-X nDSMs“. In: *Computers, Environment and Urban Systems* 102, S. 101953. DOI: [10.1016/j.compenvurbsys.2023.101953](https://doi.org/10.1016/j.compenvurbsys.2023.101953).
- Standfuß, I., Weigand, M., Gähler, M., Milbert, A., Dosch, F., Eichfuss, S., Sander, N., Geiß, C. und Taubenböck, H. (2023b). „Categorizing urban areas into “core”, “fringe”, and “periphery” based on the built-up morphology“. In: *2023 Joint Urban Remote Sensing Event (JURSE)*. IEEE. DOI: [10.1109/jurse57346.2023.10144133](https://doi.org/10.1109/jurse57346.2023.10144133).
- Stewart, I. D. und Oke, T. R. (2012). „Local Climate Zones for Urban Temperature Studies“. In: *Bulletin of the American Meteorological Society* 93.12, S. 1879–1900. DOI: [10.1175/bams-d-11-00019.1](https://doi.org/10.1175/bams-d-11-00019.1).
- Stehman, S. V., Fonte, C. C., Foody, G. M. und See, L. (2018). „Using volunteered geographic information (VGI) in design-based statistical inference for area estimation and accuracy assessment of land cover“. In: *Remote Sensing of Environment* 212, S. 47–59. DOI: [10.1016/j.rse.2018.04.014](https://doi.org/10.1016/j.rse.2018.04.014).

- Stehman, S. V. und Foody, G. M. (2019). „Key issues in rigorous accuracy assessment of land cover products“. In: *Remote Sensing of Environment* 231.111199, S. 1–23. DOI: [10.1016/j.rse.2019.05.018](https://doi.org/10.1016/j.rse.2019.05.018).
- Su, J. G., Dadvand, P., Nieuwenhuijsen, M. J., Bartoll, X. und Jerrett, M. (2019). „Associations of green space metrics with health and behavior outcomes at different buffer sizes and remote sensing sensor resolutions“. In: *Environment International* 126, S. 162–170. DOI: [10.1016/j.envint.2019.02.008](https://doi.org/10.1016/j.envint.2019.02.008).
- Swanwick, C., Dunnett, N. und Woolley, H. (2003). „Nature, Role and Value of Green Space in Towns and Cities: An Overview“. In: *Built Environment* 29.2, S. 94–106. DOI: [10.2148/benv.29.2.94.54467](https://doi.org/10.2148/benv.29.2.94.54467).
- Swanwick, C. (2009). „Society’s attitudes to and preferences for land and landscape“. In: *Land Use Policy* 26, S62–S75. DOI: [10.1016/j.landusepol.2009.08.025](https://doi.org/10.1016/j.landusepol.2009.08.025).
- Targino, A. C., Coraiola, G. C. und Krecl, P. (2019). „Green or blue spaces? Assessment of the effectiveness and costs to mitigate the urban heat island in a Latin American city“. In: *Theoretical and Applied Climatology* 136.3-4, S. 971–984. DOI: [10.1007/s00704-018-2534-1](https://doi.org/10.1007/s00704-018-2534-1).
- Tassopoulou, M., Verde, N., Mallinis, G., Georgiadis, C., Kaimaris, D. und Patias, P. (2019). „Demonstrating the potential of remote sensing to support sustainable development goals implementation: case studies over Greece“. In: *Seventh International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2019)*. Hrsg. von G. Papadavid, K. Themistocleous, S. Michaelides, V. Ambrosia und D. G. Hadjimitsis. Bd. 11174. SPIE, S. 1–13. DOI: [10.1117/12.2533634](https://doi.org/10.1117/12.2533634).
- Taubenböck, H., Heldens, W., Heiden, U. und Wurm, M. (2010). „Physische Indikatoren für die Stadtplanung“. In: *Fernerkundung im urbanen Raum*. Hrsg. von H. Taubenböck und S. Dech. Darmstadt: Wissenschaftliche Buchgesellschaft. Kap. 5.4, S. 86–93.
- Taubenböck, H., Esch, T., Felbier, A., Wiesner, M., Roth, A. und Dech, S. (2012). „Monitoring urbanization in mega cities from space“. In: *Remote Sensing of Environment* 117, S. 162–176. DOI: [10.1016/j.rse.2011.09.015](https://doi.org/10.1016/j.rse.2011.09.015).
- Taubenböck, H., Kraff, N. J. und Wurm, M. (2018). „The morphology of the Arrival City - A global categorization based on literature surveys and remotely sensed data“. In: *Applied Geography* 92, S. 150–167. DOI: [10.1016/j.apgeog.2018.02.002](https://doi.org/10.1016/j.apgeog.2018.02.002).
- Taubenböck, H., Weigand, M., Esch, T., Staab, J., Wurm, M., Mast, J. und Dech, S. (2019). „A new ranking of the world’s largest cities—Do administrative units obscure morphological realities?“ In: *Remote Sensing of Environment* 232.111353, S. 1–14. DOI: [10.1016/j.rse.2019.111353](https://doi.org/10.1016/j.rse.2019.111353).
- Taubenböck, H., Schmich, P., Erbertseder, T., Müller, I., Tenikl, J., Weigand, M., Staab, J. und Wurm, M. (2020). „Satellitendaten zur Erfassung gesundheitsrelevanter Umweltbedingungen: Beispiele und interdisziplinäre Potenziale“. In: *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz* 63.8, S. 936–944. DOI: [10.1007/s00103-020-03177-w](https://doi.org/10.1007/s00103-020-03177-w).

- Taubenböck, H., Reiter, M., Dosch, F., Leichtle, T., Weigand, M. und Wurm, M. (2021). „Which city is the greenest? A multi-dimensional deconstruction of city rankings“. In: *Computers, Environment and Urban Systems* 89, S. 101687. DOI: [10.1016/j.compenvurbsys.2021.101687](https://doi.org/10.1016/j.compenvurbsys.2021.101687).
- Taubenböck, H., Droin, A., Standfuß, I., Dosch, F., Sander, N., Milbert, A., Eichfuss, S. und Wurm, M. (2022). „To be, or not to be ‘urban’? A multi-modal method for the differentiated measurement of the degree of urbanization“. In: *Computers, Environment and Urban Systems* 95.101830, S. 1–12. DOI: [10.1016/j.compenvurbsys.2022.101830](https://doi.org/10.1016/j.compenvurbsys.2022.101830).
- Taylor, L. und Hochuli, D. F. (2017). „Defining greenspace: Multiple uses across multiple disciplines“. In: *Landscape and Urban Planning* 158, S. 25–38. DOI: [10.1016/j.landurbplan.2016.09.024](https://doi.org/10.1016/j.landurbplan.2016.09.024).
- Tenikl, J., Wurm, M., Weigand, M., Staab, J., Müller, I. und Taubenböck, H. (2019). „Satellitengestützte Vermessung von städtischem Grün in deutschen Städten“. In: *Flächenutzungsmonitoring XI*. Hrsg. von G. Meinel, U. Schumacher, M. Behnisch und T. Krüger. Bd. 77. IÖR Schriften. Berlin: Rhombos, S. 221–228.
- Texier, M. L., Schiel, K. und Caruso, G. (2018). „The provision of urban green space and its accessibility: Spatial data effects in Brussels“. In: *PLOS ONE* 13.10. Hrsg. von E. Arcaute, e0204684. DOI: [10.1371/journal.pone.0204684](https://doi.org/10.1371/journal.pone.0204684).
- Thompson, C. W., Roe, J., Aspinall, P., Mitchell, R., Clow, A. und Miller, D. (2012). „More green space is linked to less stress in deprived communities: Evidence from salivary cortisol patterns“. In: *Landscape and Urban Planning* 105.3, S. 221–229. DOI: [10.1016/j.landurbplan.2011.12.015](https://doi.org/10.1016/j.landurbplan.2011.12.015).
- Thompson, C. W., Aspinall, P. und Roe, J. (2014). „Access to Green Space in Disadvantaged Urban Communities: Evidence of Salutogenic Effects Based on Biomarker and Self-report Measures of Wellbeing“. In: *Procedia - Social and Behavioral Sciences* 153, S. 10–22. DOI: [10.1016/j.sbspro.2014.10.036](https://doi.org/10.1016/j.sbspro.2014.10.036).
- Tooke, T. R., Coops, N. C., Goodwin, N. R. und Voogt, J. A. (2009). „Extracting urban vegetation characteristics using spectral mixture analysis and decision tree classifications“. In: *Remote Sensing of Environment* 113.2, S. 398–407. DOI: [10.1016/j.rse.2008.10.005](https://doi.org/10.1016/j.rse.2008.10.005).
- Tost, H., Reichert, M., Braun, U., Reinhard, I., Peters, R., Lautenbach, S., Hoell, A., Schwarz, E., Ebner-Priemer, U., Zipf, A. und Meyer-Lindenberg, A. (2019). „Neural correlates of individual differences in affective benefit of real-life urban green space exposure“. In: *Nature Neuroscience* 22.9, S. 1389–1393. DOI: [10.1038/s41593-019-0451-y](https://doi.org/10.1038/s41593-019-0451-y).
- Tucker, C. J. und Sellers, P. J. (1986). „Satellite remote sensing of primary production“. In: *International Journal of Remote Sensing* 7.11, S. 1395–1416. DOI: [10.1080/01431168608948944](https://doi.org/10.1080/01431168608948944).
- Tucker, C. J. (1979). „Red and photographic infrared linear combinations for monitoring vegetation“. In: *Remote Sensing of Environment* 8.2, S. 127–150. DOI: [10.1016/0034-4257\(79\)90013-0](https://doi.org/10.1016/0034-4257(79)90013-0).

- Tuia, D., Moser, G., Wurm, M. und Taubenböck, H. (2017). „Land use modeling in North Rhine-Westphalia with interaction and scaling laws“. In: *2017 Joint Urban Remote Sensing Event (JURSE)*. IEEE, S. 1–4. DOI: [10.1109/jurse.2017.7924542](https://doi.org/10.1109/jurse.2017.7924542).
- Uchiyama, Y. und Kohsaka, R. (2020). „Access and Use of Green Areas during the COVID-19 Pandemic: Green Infrastructure Management in the “New Normal”“. In: *Sustainability* 12.23, S. 9842. DOI: [10.3390/su12239842](https://doi.org/10.3390/su12239842).
- UN General Assembly (2015). *Transforming our world: The 2030 agenda for sustainable development*. Techn. Ber. A/RES/70/1, 21 October 2015.
- Vargas-Munoz, J. E., Srivastava, S., Tuia, D. und Falcao, A. X. (2021). „OpenStreetMap: Challenges and Opportunities in Machine Learning and Remote Sensing“. In: *IEEE Geoscience and Remote Sensing Magazine* 9.1, S. 184–199. DOI: [10.1109/mgrs.2020.2994107](https://doi.org/10.1109/mgrs.2020.2994107).
- Villeneuve, P. J., Jerrett, M., Su, J. G., Burnett, R. T., Chen, H., Wheeler, A. J. und Goldberg, M. S. (2012). „A cohort study relating urban green space with mortality in Ontario, Canada“. In: *Environmental research* 115, S. 51–58. DOI: [10.1016/j.envres.2012.03.003](https://doi.org/10.1016/j.envres.2012.03.003).
- Voigtländer, S., Mielck, A. und Razum, O. (2011). „Die Bedeutung des kleinräumigen Kontexts für Gesundheit: Entwurf eines Erklärungsmodells“. In: *Das Gesundheitswesen* 74.11, S. 702–709. DOI: [10.1055/s-0031-1285863](https://doi.org/10.1055/s-0031-1285863).
- Voorde, T. V. de (2017). „Spatially explicit urban green indicators for characterizing vegetation cover and public green space proximity: a case study on Brussels, Belgium“. In: *International Journal of Digital Earth* 10.8, S. 798–813. DOI: [10.1080/17538947.2016.1252434](https://doi.org/10.1080/17538947.2016.1252434).
- Vries, S. de, Dillen, S. M. van, Groenewegen, P. P. und Spreeuwenberg, P. (2013). „Streetscape greenery and health: Stress, social cohesion and physical activity as mediators“. In: *Social Science & Medicine* 94, S. 26–33. DOI: [10.1016/j.socscimed.2013.06.030](https://doi.org/10.1016/j.socscimed.2013.06.030).
- Wagner, G. G., Frick, J. R. und Schupp, J. (2007). „The German Socio-Economic Panel Study (SOEP) - Evolution, Scope and Enhancements“. In: *SSRN Electronic Journal*. DOI: [10.2139/ssrn.1028709](https://doi.org/10.2139/ssrn.1028709).
- Waibel, A., Hanazawa, T., Hinton, G., Shikano, K. und Lang, K. (1989). „Phoneme recognition using time-delay neural networks“. In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 37.3, S. 328–339. DOI: [10.1109/29.21701](https://doi.org/10.1109/29.21701).
- Wan, T., Lu, H., Lu, Q. und Luo, N. (2017). „Classification of High-Resolution Remote-Sensing Image Using OpenStreetMap Information“. In: *IEEE Geoscience and Remote Sensing Letters* 14.12, S. 2305–2309. DOI: [10.1109/lgrs.2017.2762466](https://doi.org/10.1109/lgrs.2017.2762466).
- Wang, Z., Zhu, Z., Xu, M. und Qureshi, S. (2021). „Fine-grained assessment of greenspace satisfaction at regional scale using content analysis of social media and machine learning“. In: *Science of The Total Environment* 776.145908, S. 1–13. DOI: [10.1016/j.scitotenv.2021.145908](https://doi.org/10.1016/j.scitotenv.2021.145908).
- Weigand, M., Wurm, M., Dech, S. und Taubenböck, H. (2019). „Remote Sensing in Environmental Justice Research – a Review“. In: *ISPRS International Journal of Geo-Information* 8.20, S. 1–28. DOI: [10.3390/ijgi8010020](https://doi.org/10.3390/ijgi8010020).

- Weigand, M., Staab, J., Wurm, M. und Taubenböck, H. (2020). „Spatial and semantic effects of LUCAS samples on fully automated land use/land cover classification in high-resolution Sentinel-2 data“. In: *International Journal of Applied Earth Observation and Geoinformation* 88.102065, S. 1–9. DOI: [10.1016/j.jag.2020.102065](https://doi.org/10.1016/j.jag.2020.102065).
- Weigand, M., Worbis, S., Sapena, M. und Taubenböck, H. (2023a). „A structural catalogue of the settlement morphology in refugee and IDP camps“. In: *International Journal of Geographical Information Science*, S. 1–27. DOI: [10.1080/13658816.2023.2189724](https://doi.org/10.1080/13658816.2023.2189724).
- Weigand, M., Wurm, M., Droin, A., Stark, T., Staab, J., Rauh, J. und Taubenböck, H. (2023b). „Are public green spaces distributed fairly? A nationwide analysis based on remote sensing, OpenStreetMap and census data“. In: *Geocarto International* 38.1, S. 2286305. DOI: [10.1080/10106049.2023.2286305](https://doi.org/10.1080/10106049.2023.2286305).
- Wheeler, B. W., Lovell, R., Higgins, S. L., White, M. P., Alcock, I., Osborne, N. J., Husk, K., Sabel, C. E. und Depledge, M. H. (2015). „Beyond greenspace: an ecological study of population general health and indicators of natural environment type and quality“. In: *International Journal of Health Geographics* 14.17, S. 1–17. DOI: [10.1186/s12942-015-0009-5](https://doi.org/10.1186/s12942-015-0009-5).
- Wieland, M., Li, Y. und Martinis, S. (2019). „Multi-sensor cloud and cloud shadow segmentation with a convolutional neural network“. In: *Remote Sensing of Environment* 230, S. 111203. DOI: [10.1016/j.rse.2019.05.022](https://doi.org/10.1016/j.rse.2019.05.022).
- Wolch, J. R., Byrne, J. und Newell, J. P. (2014). „Urban green space, public health, and environmental justice: The challenge of making cities 'just green enough'“. In: *Landscape and Urban Planning* 125, S. 234–244. DOI: [10.1016/j.landurbplan.2014.01.017](https://doi.org/10.1016/j.landurbplan.2014.01.017).
- Wolff, M. und Haase, D. (2019). „Mediating Sustainability and Liveability—Turning Points of Green Space Supply in European Cities“. In: *Frontiers in Environmental Science* 7. DOI: [10.3389/fenvs.2019.00061](https://doi.org/10.3389/fenvs.2019.00061).
- World Health Organization (2007). *Global age-friendly cities: A guide*. Geneva.
- World Health Organization (2017). *Urban green spaces: a brief for action*. Copenhagen, Denmark: World Health Organization. Regional Office for Europe.
- World Health Organization (2020). *Mental health and psychosocial considerations during the COVID-19 outbreak*. <https://www.who.int/docs/default-source/coronaviruse/mental-health-considerations.pdf>. World Health Organization.
- Wurm, M., Goebel, J. und Wagner, G. G. (2010). „Integration raumrelevanter Indikatoren in sozial- und verhaltenswissenschaftliche Analysen“. In: *Fernerkundung im urbanen Raum*. Hrsg. von H. Taubenböck und S. Dech. Darmstadt: Wissenschaftliche Buchgesellschaft. Kap. 7.3, S. 153–162.
- Wurm, M., Taubenböck, H., Weigand, M. und Schmitt, A. (2017a). „Slum mapping in polarimetric SAR data using spatial features“. In: *Remote Sensing of Environment* 194, S. 190–204. DOI: [10.1016/j.rse.2017.03.030](https://doi.org/10.1016/j.rse.2017.03.030).





- Wurm, M., Weigand, M., Schmitt, A., Geiß, C. und Taubenbock, H. (2017b). „Exploitation of textural and morphological image features in Sentinel-2A data for slum mapping“. In: *2017 Joint Urban Remote Sensing Event (JURSE)*. IEEE, S. 1–4. DOI: [10.1109/jurse.2017.7924586](https://doi.org/10.1109/jurse.2017.7924586).
- Wurm, M., Goebel, J., Wagner, G. G., Weigand, M., Dech, S. und Taubenböck, H. (2019a). „Inferring floor area ratio thresholds for the delineation of city centers based on cognitive perception“. In: *Environment and Planning B: Urban Analytics and City Science*, S. 1–15. DOI: [10.1177/2399808319869341](https://doi.org/10.1177/2399808319869341).
- Wurm, M., Stark, T., Zhu, X. X., Weigand, M. und Taubenböck, H. (2019b). „Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks“. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 150, S. 59–69. DOI: [10.1016/j.isprsjprs.2019.02.006](https://doi.org/10.1016/j.isprsjprs.2019.02.006).
- Wurm, M., Weigand, M., Stark, T., Goebel, J., Wagner, G. G. und Taubenböck, H. (2019c). „Modelling the impact of the urban spatial structure on the choice of residential location using ‘big earth data’ and machine learning“. In: *2019 Joint Urban Remote Sensing Event (JURSE)*. IEEE. DOI: [10.1109/jurse.2019.8808942](https://doi.org/10.1109/jurse.2019.8808942).
- Wüstemann, H., Kalisch, D. und Kolbe, J. (2017a). „Access to urban green space and environmental inequalities in Germany“. In: *Landscape and Urban Planning* 164, S. 124–131. DOI: [10.1016/j.landurbplan.2017.04.002](https://doi.org/10.1016/j.landurbplan.2017.04.002).
- Wüstemann, H., Kalisch, D. und Kolbe, J. (2017b). „Accessibility of urban blue in German major cities“. In: *Ecological Indicators* 78, S. 125–130. DOI: [10.1016/j.ecolind.2017.02.035](https://doi.org/10.1016/j.ecolind.2017.02.035).
- Xie, S., Girshick, R., Dollár, P., Tu, Z. und He, K. (2016). „Aggregated Residual Transformations for Deep Neural Networks“. In: S. 1–10. arXiv: [1611.05431 \[cs.CV\]](https://arxiv.org/abs/1611.05431).
- Xu, C., Haase, D. und Pauleit, S. (2018a). „The impact of different urban dynamics on green space availability: A multiple scenario modeling approach for the region of Munich, Germany“. In: *Ecological Indicators* 93, S. 1–12. DOI: [10.1016/j.ecolind.2018.04.058](https://doi.org/10.1016/j.ecolind.2018.04.058).
- Xu, C., Haase, D., Pribadi, D. O. und Pauleit, S. (2018b). „Spatial variation of green space equity and its relation with urban dynamics: A case study in the region of Munich“. In: *Ecological Indicators* 93, S. 512–523. DOI: [10.1016/j.ecolind.2018.05.024](https://doi.org/10.1016/j.ecolind.2018.05.024).
- Zepp, H., Groß, L. und Inostroza, L. (2020). „And the winner is? Comparing urban green space provision and accessibility in eight European metropolitan areas using a spatially explicit approach“. In: *Urban Forestry & Urban Greening* 49.126603. DOI: [10.1016/j.ufug.2020.126603](https://doi.org/10.1016/j.ufug.2020.126603).
- Zha, Y., Gao, J. und Ni, S. (2003). „Use of normalized difference built-up index in automatically mapping urban areas from TM imagery“. In: *International Journal of Remote Sensing* 24.3, S. 583–594. DOI: [10.1080/01431160304987](https://doi.org/10.1080/01431160304987).
- Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J. und Atkinson, P. M. (2019). „Joint Deep Learning for land cover and land use classification“. In: *Remote Sensing of Environment* 221, S. 173–187. DOI: [10.1016/j.rse.2018.11.014](https://doi.org/10.1016/j.rse.2018.11.014).

- Zhou, Q., Wang, S. und Liu, Y. (2022). „Exploring the accuracy and completeness patterns of global land-cover/land-use data in OpenStreetMap“. In: *Applied Geography* 145, S. 102742. DOI: [10.1016/j.apgeog.2022.102742](https://doi.org/10.1016/j.apgeog.2022.102742).
- Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F. und Fraundorfer, F. (2017). „Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources“. In: *IEEE Geoscience and Remote Sensing Magazine* 5.4, S. 8–36. DOI: [10.1109/mgrs.2017.2762307](https://doi.org/10.1109/mgrs.2017.2762307).
- Zhu, X. X., Hu, J., Qiu, C., Shi, Y., Kang, J., Mou, L., Bagheri, H., Haberle, M., Hua, Y., Huang, R., Hughes, L., Li, H., Sun, Y., Zhang, G., Han, S., Schmitt, M. und Wang, Y. (2020). „So2Sat LCZ42: A Benchmark Data Set for the Classification of Global Local Climate Zones [Software and Data Sets]“. In: *IEEE Geoscience and Remote Sensing Magazine* 8.3, S. 76–89. DOI: [10.1109/mgrs.2020.2964708](https://doi.org/10.1109/mgrs.2020.2964708).
- Zhu, X. X., Qiu, C., Hu, J., Shi, Y., Wang, Y., Schmitt, M. und Taubenböck, H. (2022). „The urban morphology on our planet – Global perspectives from space“. In: *Remote Sensing of Environment* 269.112794, S. 1–11. DOI: [10.1016/j.rse.2021.112794](https://doi.org/10.1016/j.rse.2021.112794).
- Zhuang, S., Bolte, G. und Lakes, T. (2022). „Exploring Environmental Health Inequalities: A Scientometric Analysis of Global Research Trends (1970–2020)“. In: *International Journal of Environmental Research and Public Health* 19.7394, S. 1–25. DOI: [10.3390/ijerph19127394](https://doi.org/10.3390/ijerph19127394).
- Zillmann, E., Gonzalez, A., Herrero, E. J. M., Wolvelaer, J. van, Esch, T., Keil, M., Weichelt, H. und Garzon, A. M. (2014). „Pan-European Grassland Mapping Using Seasonal Statistics From Multisensor Image Time Series“. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7.8, S. 3461–3472. DOI: [10.1109/jstars.2014.2321432](https://doi.org/10.1109/jstars.2014.2321432).
- Zuniga-Teran, A. und Gerlak, A. (2019). „A Multidisciplinary Approach to Analyzing Questions of Justice Issues in Urban Greenspace“. In: *Sustainability* 11.11, S. 3055. DOI: [10.3390/su11113055](https://doi.org/10.3390/su11113055).

Anhang

Tabelle A1: OpenStreetMap-Elemente zur Extraktion von Straßen und Schienen für die deutschlandweite Landbedeckungsklassifikation. Detaillierte Informationen zu den einzelnen Elementen kann sind im OpenStreetMap-Wiki unter *Map Features* zu finden: https://wiki.openstreetmap.org/wiki/Map_features (zuletzt abgerufen 27.07.2023).

Key	Value
highway	motorway, motorway_link, trunk, trunk_link, primary, primary_link, secondary, secondary_link, tertiary, tertiary_link, unclassified, residential, service, living_street, pedestrian
railway	rail, light_rail, subway, tram
tunnel	yes (Ausschlusskriterium)

Tabelle A2: Gruppierungsschlüssel von OpenStreetMap-Elementen nach `key=value`-Paaren für Fusionsnetzwerk sowie jeweils relevante Geometrietypen: Punkte (engl. *nodes*) = , Pfade (engl. *ways*) = , Flächen (engl. *areas*) = , Relationen (engl. *relations*) = . Die Gruppierung ist Grundlage für die Aggregation von OSM-Elementen auf Nachbarschaftsebene. Alle Elemente einer Gruppe wurde zusammengefasst um je Gruppe und Geometrie ein Aggregat abzuleiten: die kumulative Anzahl von Punkten, die kumulative Länge von Linien sowie die kumulative Fläche von Flächenelementen.






Gruppe	Assoziierte <code>Key=Value</code> Paare und Geometrietypen
access_agricultural	access=agricultural 
access_forestry	access=forestry 
access_no	access=no 
access_permissive	access=permissive 
access_private	access=private 

Tabelle A2: (fortgesetzt)

Gruppe	Assoziierte Key=Value Paare und Geometrietypen
access_yes	access=yes 
aeroways	aeroway=aerodrome   , aeroway=apron  , aeroway=gate  , aeroway=hangar  , aeroway=helipad   , aeroway=heliport   , aeroway=navigationaid  , aeroway=runway  , aeroway=spaceport   , aeroway=taxiway  , aeroway=terminal   , aeroway=windsock  , building=hangar 
agricultural	building=barn  , building=farm_auxiliary  , building=greenhouse  , building=stable 
amenity	amenity=bicycle_parking   , amenity=fountain   , amenity=bbq  , amenity=bench  , amenity=drinking_water  , amenity=shelter   , amenity=toilets   , amenity=waste_basket  , amenity=clock  , amenity=funeral_hall   , amenity=grave_yard   , amenity=hunting_stand   , amenity=kneipp_water_cure    , amenity=lounger  , amenity=place_of_worship   , leisure=firepit   , man_made=water_tap 
barrier	barrier=ditch  , barrier=fence  , barrier=wall   , barrier=yes   
bicycle_no	bicycle=no 
bicycle_yes	bicycle=yes 
building_yes	building=yes 
educational	amenity=college   , amenity=kindergarten   , amenity=language_school   , amenity=library   , amenity=music_school   , amenity=school   , amenity=university   , building=kindergarten  , building=school  , building=university  , building=college 
foot_no	foot=no 
foot_permissive	foot=permissive 
foot_private	foot=private 
foot_yes	foot=yes 

Tabelle A2: (fortgesetzt)

Gruppe	Assoziierte Key=Value Paare und Geometrietypen
gate	barrier=gate  
hedge	barrier=hedge  
historic_site	historic=aircraft   , historic=aqueduct   , historic=archaeological_site    , historic=battlefield   , historic=building   , historic=castle   , historic=castle_wall   , historic=charcoal_pile   , historic=church   , historic=city_gate   , historic=farm   , historic=fort   , historic=manor   , historic=memorial   , historic=monument   , historic=ruins    , historic=tomb   , historic=tower   , historic=wayside_cross  , historic=wayside_shrine  
landuse_allotments	landuse=allotments 
landuse_basin	landuse=basin 
landuse_brownfield	landuse=brownfield 
landuse_cemetery	landuse=cemetery 
landuse_commercial	landuse=commercial  
landuse_construction	landuse=construction 
landuse_farmland	landuse=farmland 
landuse_farmyard	landuse=farmyard 
landuse_forest	landuse=forest 
landuse_grass	landuse=grass 
landuse_greenfield	landuse=greenfield 
landuse_industrial	landuse=industrial 
landuse_landfill	landuse=landfill 
landuse_meadow	landuse=meadow 
landuse_military	landuse=military 
landuse_orchard	landuse=orchard 





















































Tabelle A2: (fortgesetzt)

Gruppe	Assoziierte Key=Value Paare und Geometrietypen
landuse_plantnursery	landuse=plant_nursery ☐ ☐
landuse_quarry	landuse=quarry ☐ ☐
landuse_railway	landuse=railway ☐ ☐
landuse_recreationground	landuse=recreation_ground ☐ ☐
landuse_religious	landuse=religious ☐ ☐
landuse_residential	landuse=residential ☐ ☐
landuse_retail	landuse=retail ☐ ☐
landuse_villagegreen	landuse=village_green ☐ ☐
landuse_vineyard	landuse=vineyard ☐ ☐
leisure_public	leisure=common ☐ ☐ ☐ ☐, leisure=dog_park ☐ ☐ ☐ ☐, leisure=fitness_station ☐ ☐ ☐ ☐, leisure=nature_reserve ☐ ☐ ☐ ☐, leisure=park ☐ ☐ ☐ ☐, leisure=picnic_table ☐ ☐ ☐ ☐, leisure=playground ☐ ☐ ☐ ☐, leisure=slipway ☐ ☐ ☐ ☐
leisure_restrictive	leisure=dance ☐ ☐ ☐ ☐, leisure=disc_golf_course ☐ ☐ ☐ ☐, leisure=escape_game ☐ ☐ ☐ ☐, leisure=marina ☐ ☐ ☐ ☐, leisure=miniature_golf ☐ ☐ ☐ ☐, tourism=camp_site ☐ ☐ ☐ ☐, tourism=caravan_site ☐ ☐ ☐ ☐
leisure_sport	leisure=pitch ☐ ☐ ☐ ☐, leisure=sports_centre ☐ ☐ ☐ ☐, leisure=stadium ☐ ☐ ☐ ☐, leisure=swimming_area ☐ ☐ ☐ ☐, leisure=swimming_pool ☐ ☐ ☐ ☐, leisure=track ☐ ☐ ☐ ☐, leisure=water_park ☐ ☐ ☐ ☐
medical_building	amenity=hospital ☐ ☐ ☐ ☐, amenity=nursing_home ☐ ☐ ☐ ☐, building=hospital ☐ ☐ ☐ ☐
military	building=military ☐ ☐ ☐ ☐, building=bunker ☐ ☐ ☐ ☐, military=airfield ☐ ☐ ☐ ☐, military=bunker ☐ ☐ ☐ ☐, military=barracks ☐ ☐ ☐ ☐, military=checkpoint ☐ ☐ ☐ ☐, military=danger_area ☐ ☐ ☐ ☐, military=nuclear_explosion_site ☐ ☐ ☐ ☐, military=obstacle_course ☐ ☐ ☐ ☐, military=office ☐ ☐ ☐ ☐, military=range ☐ ☐ ☐ ☐, military=training_area ☐ ☐ ☐ ☐
natural_beach	natural=beach ☐ ☐ ☐ ☐

Tabelle A2: (fortgesetzt)

Gruppe	Assoziierte Key=Value Paare und Geometrietypen
natural_grassland	natural=grassland ☐☒
natural_heath	natural=heath ☐☒
natural_sand	natural=sand ☐☒
natural_scrub	natural=scrub ☐☒
natural_spring	natural=spring ☐☐
natural_water	natural=water ☐☒
natural_wetland	natural=wetland ☐☒
natural_wood	natural=wood ☐☒, natural=tree_row ☐☒, natural=tree ☐☐
park	boundary=national_park ☐☒, boundary=protected_area ☐☒
public_building	amenity=community_centre ☐☒☒, amenity=theatre ☐☒☒, amenity=courthouse ☐☒☒, amenity=townhall ☐☒☒, building=cathedral ☐☒, building=chapel ☐☒, building=church ☐☒, building=mosque ☐☒, building=religious ☐☒, building=synagogue ☐☒, building=temple ☐☒, building=civic ☐☒, building=government ☐☒, building=public ☐☒, building=toilets ☐☒, building=train_station ☐☒, building=transportation ☐☒, building=conservatory ☐☒
restaurants	amenity=bar ☐☒☒, amenity=biergarten ☐☒☒, amenity=cafe ☐☒☒, amenity=fast_food ☐☒☒, amenity=food_court ☐☒☒, amenity=ice_cream ☐☒☒, amenity=pub ☐☒☒, amenity=restaurant ☐☒☒
route_bicycle	route=bicycle ☐☒☒, route=mtb ☐☒☒
route_foot	route=foot ☐☒☒, route=running ☐☒☒
route_hiking	route=hiking ☐☒☒
sports_building	building=grandstand ☐☒, building=pavilion ☐☒, building=riding_hall ☐☒, building=sports_hall ☐☒, building=stadium ☐☒

Tabelle A2: (fortgesetzt)

Gruppe	Assoziierte Key=Value Paare und Geometrietypen
tourist_attraction	tourism=attraction  , tourism=gallery  , tourism=information  , tourism=museum  , tourism=picnic_site  , tourism=theme_park  , tourism=viewpoint  , tourism=wilderness_hut  , tourism=zoo 
traffic_foot_cycle	highway=pedestrian  , highway=track  , highway=footway   , highway=bridleway  , highway=steps  , highway=path  , highway=cycleway 
traffic_motorized	highway=motorway  , highway=trunk  , highway=primary  , highway=secondary  , highway=tertiary  , highway=unclassified   , highway=residential  , highway=motorway_link  , highway=trunk_link  , highway=primary_link  , highway=secondary_link  , highway=tertiary_link  , highway=living_street  , highway=service  , highway=bus_guideway  , highway=escape  , highway=raceway  , highway=road  , highway=busway 
water	water=river  , water=oxbow  , water=canal  , water=lake  , water=reservoir  , water=pond  , water=basin  , water=lagoon  , waterway=river   , waterway=riverbank  , waterway=stream   , waterway=canal  

Codeblock A1: Netzwerkarchitekturen der CNN-, ANN- und Fusion-Regressionsnetzwerke zur Quantifizierung von öffentlichen Grünflächen auf Nachbarschaftsebene. Die verwendeten Softwareversionen lauten: Python v. 3.9.6, Tensorflow v. 2.8.

```
-----CNN-----
Model: "model"
-----
Layer (type)                 Output Shape                 Param #
-----
input_img (InputLayer)       [(None, 101, 101, 5)]      0

greennet_v2_skipcon (Function) (None, 30)                  122942
onal)

| input_1 (InputLayer)       [(None, 101, 101, 5)]      0
| conv_1_1 (Conv2D)          (None, 101, 101, 16)      6496
| bn_1_1 (BatchNormalization) (None, 101, 101, 16)      64
| cnn_relu_1_1 (Activation)  (None, 101, 101, 16)      0
| conv_1_2 (Conv2D)          (None, 101, 101, 16)      20752
| bn_1_2 (BatchNormalization) (None, 101, 101, 16)      64
| cnn_relu_1_2 (Activation)  (None, 101, 101, 16)      0
| maxpool_1 (MaxPooling2D)  (None, 50, 50, 16)        0
| conv_2_1 (Conv2D)          (None, 50, 50, 32)        12832
| bn_2_1 (BatchNormalization) (None, 50, 50, 32)        128
| cnn_relu_2_1 (Activation)  (None, 50, 50, 32)        0
| conv_2_2 (Conv2D)          (None, 50, 50, 32)        25632
| bn_2_2 (BatchNormalization) (None, 50, 50, 32)        128
| cnn_relu_2_2 (Activation)  (None, 50, 50, 32)        0
| maxpool_2 (MaxPooling2D)  (None, 25, 25, 32)        0
| conv_3_1 (Conv2D)          (None, 25, 25, 64)        18496
| bn_3_1 (BatchNormalization) (None, 25, 25, 64)        256
| cnn_relu_3_1 (Activation)  (None, 25, 25, 64)        0
| conv_3_2 (Conv2D)          (None, 25, 25, 64)        36928
| bn_3_2 (BatchNormalization) (None, 25, 25, 64)        256
| cnn_relu_3_2 (Activation)  (None, 25, 25, 64)        0
| maxpool_3 (MaxPooling2D)  (None, 12, 12, 64)        0
| globalavgpool2d_1 (GlobalAveragePooling2D) (None, 16)                0
| globalavgpool2d_2 (GlobalAveragePooling2D) (None, 32)                0
| globalavgpool2d_3 (GlobalAveragePooling2D) (None, 64)                0
| greennet_dense_1 (Dense)   (None, 15)                 255
| greennet_dense_2 (Dense)   (None, 10)                 330
| greennet_dense_3 (Dense)   (None, 5)                  325
| cnn_bn1 (Dropout)          (None, 15)                 0
| cnn_bn2 (Dropout)          (None, 10)                 0
```

```

|
| cnn_bn3 (Dropout)          (None, 5)          0
|
| cnn_relu_act1 (Activation) (None, 15)         0
|
| cnn_relu_act2 (Activation) (None, 10)         0
|
| cnn_relu_act3 (Activation) (None, 5)          0
|
| final_greenet_concat (Conc (None, 30)         0
| atenate)
|-----|
| final_big_dropout_dropout ( (None, 30)         0
| Dropout)
|
| CNNRegressionLayer (Dense) (None, 1)          31
|-----|
|
| Total params: 122,973
| Trainable params: 122,525
| Non-trainable params: 448
|-----|

```

```

-----ANN-----
Model: "model_1"

```

```

-----|-----|-----|-----|
| Layer (type)          | Output Shape | Param # |
|-----|-----|-----|-----|
| input_osm (InputLayer) | [(None, 94)] | 0
|
| ann_dense_0 (Dense)    | (None, 20)   | 1900
|
| ann_batchnorm_0 (BatchNorma | (None, 20)   | 80
| lization)
|
| ann_relu_0 (Activation)   | (None, 20)   | 0
|
| ann_dense_1 (Dense)      | (None, 10)   | 210
|
| ann_batchnorm_1 (BatchNorma | (None, 10)   | 40
| lization)
|
| ann_relu_1 (Activation)   | (None, 10)   | 0
|
| ann_dense_2 (Dense)      | (None, 5)    | 55
|
| ann_batchnorm_2 (BatchNorma | (None, 5)    | 20
| lization)
|
| ann_relu_2 (Activation)   | (None, 5)    | 0
|
| final_big_dropout_dropout ( (None, 5)         0
| Dropout)
|
| ANNRegressionLayer (Dense) | (None, 1)    | 6
|-----|-----|-----|-----|

```

```

|
| Total params: 2,311
| Trainable params: 2,241
| Non-trainable params: 70
|-----|

```

```

-----Fusion-----
Model: "model"

```

```

-----|-----|-----|-----|-----|
| Layer (type)          | Output Shape | Param # | Connected to |
|-----|-----|-----|-----|-----|
| input_osm (InputLayer) | [(None, 94)] | 0       | []
|
| ann_dense_0 (Dense)    | (None, 20)   | 1900    | ['input_osm[0][0]']
|
| ann_batchnorm_0 (BatchNormaliz | (None, 20)   | 80      | ['ann_dense_0[0][0]']
| ation)
|
| ann_relu_0 (Activation)   | (None, 20)   | 0       | ['ann_batchnorm_0[0][0]']
|
| ann_dense_1 (Dense)      | (None, 10)   | 210     | ['ann_relu_0[0][0]']
|-----|-----|-----|-----|-----|

```

```

ann_batchnorm_1 (BatchNormaliz (None, 10)      40    ['ann_dense_1[0][0]']
ation)

ann_relu_1 (Activation)          (None, 10)      0    ['ann_batchnorm_1[0][0]']

ann_dense_2 (Dense)              (None, 5)       55   ['ann_relu_1[0][0]']

ann_batchnorm_2 (BatchNormaliz (None, 5)       20   ['ann_dense_2[0][0]']
ation)

input_img (InputLayer)          [(None, 101, 101, 5 0  []
)]

ann_relu_2 (Activation)          (None, 5)       0    ['ann_batchnorm_2[0][0]']

greennet_v2_skipcon (Functiona (None, 30)      122942 ['input_img[0][0]']
1)
-----|
| input_1 (InputLayer)          [(None, 101, 101, 5 0  []
| )]
|
| conv_1_1 (Conv2D)             (None, 101, 101, 16 6496  []
| )
|
| bn_1_1 (BatchNormalization)   (None, 101, 101, 16 64   []
| )
|
| cnn_relu_1_1 (Activation)      (None, 101, 101, 16 0    []
| )
|
| conv_1_2 (Conv2D)             (None, 101, 101, 16 20752 []
| )
|
| bn_1_2 (BatchNormalization)   (None, 101, 101, 16 64   []
| )
|
| cnn_relu_1_2 (Activation)      (None, 101, 101, 16 0    []
| )
|
| maxpool_1 (MaxPooling2D)      (None, 50, 50, 16) 0    []
|
| conv_2_1 (Conv2D)             (None, 50, 50, 32) 12832 []
|
| bn_2_1 (BatchNormalization)   (None, 50, 50, 32) 128   []
|
| cnn_relu_2_1 (Activation)      (None, 50, 50, 32) 0    []
|
| conv_2_2 (Conv2D)             (None, 50, 50, 32) 25632 []
|
| bn_2_2 (BatchNormalization)   (None, 50, 50, 32) 128   []
|
| cnn_relu_2_2 (Activation)      (None, 50, 50, 32) 0    []
|
| maxpool_2 (MaxPooling2D)      (None, 25, 25, 32) 0    []
|
| conv_3_1 (Conv2D)             (None, 25, 25, 64) 18496 []
|
| bn_3_1 (BatchNormalization)   (None, 25, 25, 64) 256   []
|
| cnn_relu_3_1 (Activation)      (None, 25, 25, 64) 0    []
|
| conv_3_2 (Conv2D)             (None, 25, 25, 64) 36928 []
|
| bn_3_2 (BatchNormalization)   (None, 25, 25, 64) 256   []
|
| cnn_relu_3_2 (Activation)      (None, 25, 25, 64) 0    []
|
| maxpool_3 (MaxPooling2D)      (None, 12, 12, 64) 0    []
|
| globavgpool2d_1 (GlobalAverage (None, 16)      0    []
| Pooling2D)
|
| globavgpool2d_2 (GlobalAverage (None, 32)      0    []
| Pooling2D)
|
| globavgpool2d_3 (GlobalAverage (None, 64)      0    []

```

```

| Pooling2D)
|
| greennet_dense_1 (Dense)      (None, 15)      255      []
| greennet_dense_2 (Dense)      (None, 10)      330      []
| greennet_dense_3 (Dense)      (None, 5)       325      []
| cnn_bn1 (Dropout)            (None, 15)      0        []
| cnn_bn2 (Dropout)            (None, 10)      0        []
| cnn_bn3 (Dropout)            (None, 5)       0        []
| cnn_relu_act1 (Activation)    (None, 15)      0        []
| cnn_relu_act2 (Activation)    (None, 10)      0        []
| cnn_relu_act3 (Activation)    (None, 5)       0        []
|
| final_greennet_concat (Concate (None, 30)      0        []
| nate)
-----
FinalANNLayer (Dense)          (None, 5)       30      ['ann_relu_2[0][0]']
FinalCNNLayer (Dense)          (None, 5)       155     ['greennet_v2_skipcon[0][0]']
concatenate (Concatenate)      (None, 10)      0       ['FinalANNLayer[0][0]',
      'FinalCNNLayer[0][0]']
fusion_dense_0 (Dense)          (None, 20)      220     ['concatenate[0][0]']
fusion_batchnorm_0 (BatchNorma (None, 20)      80      ['fusion_dense_0[0][0]']
lization)
fusion_relu_0 (Activation)      (None, 20)      0       ['fusion_batchnorm_0[0][0]']
fusion_dense_1 (Dense)          (None, 15)      315     ['fusion_relu_0[0][0]']
fusion_batchnorm_1 (BatchNorma (None, 15)      60      ['fusion_dense_1[0][0]']
lization)
fusion_relu_1 (Activation)      (None, 15)      0       ['fusion_batchnorm_1[0][0]']
fusion_dense_2 (Dense)          (None, 10)      160     ['fusion_relu_1[0][0]']
fusion_batchnorm_2 (BatchNorma (None, 10)      40      ['fusion_dense_2[0][0]']
lization)
fusion_relu_2 (Activation)      (None, 10)      0       ['fusion_batchnorm_2[0][0]']
fusion_dense_3 (Dense)          (None, 5)       55      ['fusion_relu_2[0][0]']
fusion_batchnorm_3 (BatchNorma (None, 5)       20      ['fusion_dense_3[0][0]']
lization)
fusion_relu_3 (Activation)      (None, 5)       0       ['fusion_batchnorm_3[0][0]']
final_big_dropout_dropout (Dro (None, 5)      0       ['fusion_relu_3[0][0]']
pout)
fusion_FusionRegressionLayer ( (None, 1)      6       ['final_big_dropout_dropout[0][0]']
Dense)
=====
Total params: 126,388
Trainable params: 125,770
Non-trainable params: 618
-----

```


Versicherung an Eides Statt

Ich, Matthias Johann Weigand, versichere an Eides Statt durch meine Unterschrift, dass ich die Dissertation "*Fernerkundung und Maschinelles Lernen zur Erfassung von urbanem Grün - Eine Analyse am Beispiel der Verteilungsgerechtigkeit in Deutschland*" selbständig und ohne fremde Hilfe angefertigt, alle Stellen, die ich wörtlich oder dem Sinne nach aus Veröffentlichungen entnommen habe, als solche kenntlich gemacht und ich auch keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt habe.

Ich versichere an Eides Statt durch meine Unterschrift, dass ich die Regeln der Universität Würzburg über gute wissenschaftliche Praxis eingehalten habe, insbesondere, dass ich die Gelegenheit zum Promotionsvorhaben nicht kommerziell vermittelt bekommen und insbesondere nicht eine Person oder Organisation eingeschaltet habe, die gegen Entgelt Betreuer bzw. Betreuerinnen für die Anfertigung von Dissertationen sucht.

Ich versichere an Eides Statt, dass ich die vorgenannten Angaben nach bestem Wissen und Gewissen gemacht habe und dass die Angaben der Wahrheit entsprechen und ich nichts verschwiegen habe. Die Strafbarkeit einer falschen eidesstattlichen Versicherung ist mir bekannt, namentlich die Strafandrohung gemäß § 156 StGB bis zu drei Jahren Freiheitsstrafe oder Geldstrafe bei vorsätzlicher Begehung der Tat bzw. gemäß § 161 Abs. 1 StGB bis zu einem Jahr Freiheitsstrafe oder Geldstrafe bei fahrlässiger Begehung.

Matthias Johann Weigand
Landsberg am Lech, 31. August 2023