

Verschwundene Karten

Objektunterschiede und räumliche Verbreitung von
Löschen in OpenStreetMap mit Fokus auf
Konfliktregionen

Manuel Kraft

Matrikelnummer: 4064202
E-Mail: manuel.kraft@online.de

Masterarbeit
Master of Science

Geographisches Institut
Universität Heidelberg
Deutschland
10.04.2025

Erster Betreuer: Prof. Dr. Sven Lautenbach
Zweiter Betreuer: Dr. Michael Auer

Abstract

Löschenungen in OpenStreetMap bieten Einblicke in sozial-geographische Ereignisse aus der Vergangenheit und liefern Rückschlüsse auf die OSM-Datenqualität. Deshalb wurde die räumliche Verbreitung und die Objektunterschiede von OSM-Löschenungen weltweit herausgearbeitet. Lokale Auffälligkeiten wurden anschließend detaillierter analysiert, um Erkenntnisse zu den Hintergründen der Löschenungen zu sammeln. Des Weiteren wurden fünf Regionen untersucht, deren Territorium von mehreren Bevölkerungsgruppen beansprucht wird. Damit sollte beurteilt werden, ob sich in den Löschenungen territoriale Konflikte widerspiegeln.

Die zugrunde liegenden Daten stammen von der OSHDB. Sie wird von der Heigit gGmbH verwaltet und bietet Zugriff auf die vollständige Historie aller OSM-Elemente. Verwendet wurde dabei ein Datensatz, der alle OSM-Objekte beinhaltet, die im Juli 2024 auf der OSM-Karte sichtbar waren oder bis Juli 2024 von der OSM-Karte gelöscht wurden. Dieser Datensatz wurde im Anschluss anhand eigenständig geschriebener Programme in Python und R vorprozessiert und analysiert. Die globalen thematischen Karten wurden mit QGIS erstellt und der globale Raum dabei in rund 290km² große Hexagone zergliedert.

Die Ergebnisse der Objektunterschiede zeigen, dass administrative Grenzen am häufigsten gelöscht werden. Danach folgen touristische Ziele und Verkaufseinrichtungen. Selten gelöscht werden stehende Gewässer, Barrieren, Telekommunikationssysteme und Elektrizitäts-Einrichtungen. Kein auffälliges Löscherhalten gibt es bei militärischen Objekten. Insgesamt werden OSM-Objekte, die nicht nach den üblichen Konventionen der OSM-Community beschrieben wurden, überproportional häufig gelöscht.

Die Ergebnisse der räumlichen Verbreitung von Löschenungen zeigen, dass es keine klar erkennbaren großflächigen Unterschiede gibt. Stattdessen treten sichtbare räumliche Unterschiede überwiegend lokal auf. Die Ursachen von Löschenungen in Regionen mit besonders vielen Löschenungen können in vier verschiedene Typen unterteilt werden und hängen meistens mit umfangreichen Datenimports zusammen.

In den meisten untersuchten Konfliktregionen sind Löschenungen mit Bezug zu dem territorialen Konflikt zwar vorhanden, allerdings machen sie nur einen kleinen Teil der dortigen Löschenungen aus. Nur in Bergkarabach spiegelt sich der Konflikt deutlich in den Löschenungen wider. Insgesamt zeigen die Ergebnisse, dass OSM-Löschenungen mehr von Harmonie als von Streit geprägt sind.

Danksagungen

Ein Dank gilt dem Datenspeicherungs-Service SDS@hd, welches unterstützt wird von dem Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg (MWK) und der Deutschen Forschungsgemeinschaft (DFG) mit dem Zuschuss INST 35/1503-1 FUGG.

Ein weiterer Dank geht an Rafael Troilo, Forscher am *Heidelberg Institute for Geoinformation Technology* (Heigit gGmbH), der eine wichtige Hilfestellung für den Zugang und der Aufbereitung der gelöschten OSM-Daten war.

Inhaltsverzeichnis

1.	Einleitung.....	5
2.	Datengrundlage.....	9
3.	Methodik.....	11
3.1	Software.....	12
3.2	Hexagone als räumliche Gliederung.....	13
3.3	Datenkorrektur.....	14
3.4	Permutationstests zum Vergleich von zwei Regionen	15
3.5	Unterschiedene Objekte	17
3.5.1	Gegenständliche Objekte	17
3.5.2	Adressen	17
3.5.3	Objekte ohne Informationen	18
3.5.4	Sonstige Objekte	19
3.6	Ignorierte Löschungen	19
3.6.1	Hausdorff-Distanz als Maß für geometrische Ähnlichkeit	20
3.7	Untersuchte Konfliktregionen.....	22
4.	Ergebnisse	25
4.1	Globale Räumliche Verbreitung	25
4.2	Globale Objektunterschiede.....	28
4.3	Löschungsreiche Regionen	34
4.3.1	Dschibuti.....	34
4.3.2	Niederlande	36
4.3.3	Finnland und Lettland	38
4.3.4	Südafrika	42
4.3.5	Australien	44
4.4	Konfliktregionen	47
4.4.1	Westjordanland	47
4.4.2	Kaschmir	51
4.4.3	Karabach	57
4.4.4	Arunachal Pradesh	62
4.4.5	Westsahara	67
5.	Interpretation	72
5.1	Globale Räumliche Verbreitung	72
5.2	Globale Objektunterschiede.....	73
5.3	Löschungsreiche Regionen	75
5.4	Konfliktregionen	79
6.	Diskussion.....	84
7.	Fazit	87

1. Einleitung

OpenStreetMap (OSM) ist das erfolgreichste Crowdsource-Projekt geographischer Informationen (FOODY et al. 2017, in: MINGHINI & FRASSINELLI 2019, S.1). Die zugrundeliegende OSM-Datenbank wird unter der *Open Database License*¹ verbreitet. Damit kann die Datenbank von jedem frei genutzt und verändert werden. Die besonders leichte Erstellung neuer OSM Daten, auch für Nutzer ohne geografischen oder computertechnischen Hintergrund, macht das Beitreten attraktiv und führt mittlerweile zu knapp 10 Millionen registrierten Nutzer² (MINGHINI & FRASSINELLI 2019).

Während einige Beitragende als Einzelpersonen agieren, haben andere Nutzer Communities gegründet³. Die Communities haben unterschiedliche Größen und reichen von kleinen Gemeinden bis hin zu länderübergreifenden Regionen. OSM ist mit seinen Communities also auch ein soziales Projekt, bei dem verschiedene Sichtweisen aufeinandertreffen.

Dabei gilt normalerweise die sogenannte *Linus-Law* Annahme: Die Beiträge mehrerer Akteure konvergieren zu einem gemeinsamen Ergebnis und die Datenqualität steigt (McGOUGH et al., S. 174). Von Besonderheit sind dabei Löschungen. Also Änderungen, die keine Konvergenz anstreben, sondern einen radikalen Neuanfang.

Es gibt bisher keine Studien, in denen die Löschungen von OSM-Objekten analysiert werden. Dabei haben solche Informationen das Potential, einerseits sozial-geographische Ereignisse aus der Vergangenheit zu erkennen und andererseits ergänzende Aussagen über die OSM-Datenqualität zu liefern.

So können sich Meinungsverschiedenheiten und soziale Konflikte in den Löschungen widerspiegeln. Sofern solche Motive die OSM-Karte beeinflussen, können damit Rückschlüsse über den zugrundeliegenden Konflikt getroffen werden. Allerdings ist das auch ein Hinweis auf eine niedrigere OSM-Datenqualität, weil durch Meinungsverschiedenheiten die OSM-Karte in einer Art und Weise verändert wird, wie sie nicht der realen Welt entspricht.

¹ <https://www.openstreetmap.org/copyright/en>

² https://planet.openstreetmap.org/statistics/data_stats.html

³ <https://www.openstreetmap.org/communities>

Auch wenn OSM-Objekte besonders häufig oder besonders selten gelöscht werden, können sowohl sozial-geographische Ereignisse erkannt als auch Aussagen zur Datenqualität der Objekte getroffen werden. So können z.B. besonders viele Löschungen von militärischen Anlagen einen Hinweis darauf liefern, dass aus militär-strategischen Gründen die Öffentlichkeit von solchen Objekten nichts erfahren soll. Auch sind häufig gelöschte Objekte weniger vertrauenserweckend als Objekte, die selten gelöscht werden.

Weitere Aussagen über die Datenqualität und über sozial-geographische Ereignisse können anhand der räumlichen Verbreitung von Löschungen getroffen werden. So können auffallend viele Löschungen in einer Region z. B. das Ergebnis politischer Umgestaltungen sein. Sofern die Löschungen über einen langen Zeitraum auftreten, kann das auch auf bewusste Verschleierung von Informationen oder organisierten Vandalismus hinweisen, sodass Aussagen über die OSM-Datenqualität getroffen werden können.

In dieser Arbeit werden die weltweiten Löschungen explorativ untersucht. Das heißt, dass zunächst herausgearbeitet wird, wie die globale räumliche Verteilung der Löschungen ist und welche OSM-Objekte unter- und überproportional gelöscht werden.

Darauf aufbauend erfolgt eine genauere Untersuchung von Regionen mit besonders vielen Löschungen, sodass Hintergründe zu deren Ursachen herausgefunden werden können. Auch werden Konfliktregionen genauer untersucht, um zu beurteilen, ob sich Meinungsverschiedenheiten und politische Spannungen in den Löschungen widerspiegeln.

Die Forschungsfragen lauten dabei wie folgt:

1. Wie sieht die globale räumliche Verteilung von Löschungen aus und gibt es großflächige Unterschiede?
2. Welche OSM-Objekte werden unter- und welche überproportional häufig gelöscht?
3. Bezogen auf Regionen mit besonders vielen Löschungen: Was sind die Ursachen solcher Löschungen?
4. Haben territoriale Konflikte in der realen Welt einen Einfluss darauf, dass mehr Löschungen auftreten?

Zu der ersten und dritten Forschungsfrage gibt es keine Hypothesen, weil zu wenig Vorwissen darüber besteht. Die Herangehensweise für die Beantwortung dieser Forschungsfragen erfolgt deshalb explorativ.

Zu der zweiten Forschungsfrage gibt es folgende Hypothesen:

1. Objekte, die nicht dem informellen Standard der OSM-Community⁴ entsprechen, werden überproportional häufig gelöscht, weil sie der OSM-Karte keine ausreichenden Informationen liefern.
2. Administrative Grenzen werden überproportional häufig gelöscht, weil sie sich einerseits häufig verändern und andererseits Brennpunkt von Meinungsverschiedenheiten sind.
3. Militärisch bedeutsame Objekte werden überproportional häufig gelöscht, weil manche Akteure ein Interesse daran haben, dass solche Objekte nicht der Öffentlichkeit bekannt gemacht werden sollen.
4. Natürliche Objekte wie Flüsse und Seen werden kaum gelöscht, weil sie üblicherweise nicht umstritten sind und sich selten verändern.

Zu der vierten Forschungsfrage gibt es folgende Hypothese:

1. In Regionen mit territorialen Konflikten gibt es überproportional viele Löschungen. So sorgen politische Meinungsverschiedenheiten zwischen den Kartierern dafür, dass Objekte mehrmals neu erstellt und wieder gelöscht werden. Zum anderen werden Objekte aufgrund taktischer Überlegungen verschleiert, um sie vor anderen Konfliktparteien zu verbergen.

Um die Forschungsfrage zu beantworten, ob territoriale Konflikte einen Einfluss darauf haben, dass mehr Löschungen auftreten, werden mehrere Konfliktregionen untersucht. Gemeinsam ist diesen Regionen, dass ihr gesamtes Territorium seit über 15 Jahren von mindestens zwei Konfliktparteien beansprucht wird. So werden in dieser Arbeit alle OSM-Löschen von 2007 bis 2024 analysiert und nicht nur die der letzten drei Jahre. Deshalb werden neuere Konflikte, wie z.B. der seit 2022 andauernde Ukraine-Krieg, nicht behandelt.

Für die Arbeit wurde der weltweite, historische OSM-Datensatz ausgewertet. Die Daten stammen von der *OpenStreetMap History Database* (OSHDB) der Heigit gGmbH. Der Datensatz enthält alle Löschen von 2007 bis 2024 und alle Objekte, die 2024 auf der OSM-Karte sichtbar waren. Eine detaillierte Beschreibung dazu erfolgt in Kapitel 2.

In Kapitel 3 wird das methodische Vorgehen der Datenanalyse beschrieben. Dabei werden die verwendeten Hilfsmittel aufgelistet (3.1), die Hexagone als räumliche Einheit erklärt (3.2), Korrekturen ungenauer Datenwerte beschrieben (3.3) und die Anwendung statistischer Tests thematisiert (3.4). Auch werden die verschiedenen OSM-Objektgruppen präsentiert (3.5) und

⁴ https://wiki.openstreetmap.org/w/index.php?title=Map_features&oldid=2804789

erklärt, wie bestimmte Löschungen von der Analyse ausgeschlossen wurden (3.6). Zuletzt werden die ausgewählten Konfliktregionen präsentiert mitsamt ihrer zugrunde liegenden Konflikte (3.7).

In Kapitel 4 werden die Ergebnisse der Datenanalyse präsentiert. Dabei wird zunächst die globale räumliche Verbreitung thematisiert (4.1). Anschließend werden die Objektunterschiede behandelt (4.2) und Regionen detaillierter untersucht, in denen besonders viele Löschungen stattgefunden haben (4.3). Der letzte Ergebnisabschnitt behandelt die Konfliktregionen (4.4).

In Kapitel 5 werden die Ergebnisse interpretiert und mögliche Erklärungsversuche für das Löschverhalten angesprochen. In Kapitel 6 wird die Aussagekraft der Ergebnisse diskutiert und auf die Limitierungen dieser Arbeit eingegangen. Zum Abschluss werden die wichtigsten Ergebnisse zusammengefasst und Bilanz gezogen, inwiefern mit dieser Arbeit neue Erkenntnisse zur OSM-Datenqualität und zu sozialen Verhaltensweisen der OSM-Community gewonnen werden konnten.

2. Datengrundlage

Die primäre Datenquelle für die Arbeit ist die *OpenStreetMap History Database* (OSHDB). Sie wird von der Heigt gGmbH verwaltet und bietet Zugriff auf die vollständige Historie aller OSM-Elemente. Aufgrund der Größe der historischen OSM-Daten sind die Möglichkeiten, direkt damit zu arbeiten, limitiert. Außerdem gibt es wenig benutzerfreundliche Software-Analysetools. Das Ziel der OSHDB ist es, OSM-Daten für Forscher und andere Interessierte leichter zugänglich zu machen. Das wird u.a. durch effiziente Datenspeicherung und effiziente Datenabfragen ermöglicht. (RAIFER et al. 2024)

Aus der OSHDB wurde dabei ein Datensatz aus Parquet-Dateien abgeleitet, der die Historie aller Änderungen auf der OSM-Karte beinhaltet. Dabei repräsentiert jede Änderung ein Objekt mit Informationen zu vorherigem und nachfolgendem Zustand. Auch beinhaltet jede Änderung die Information, ob es sich um die letzte Änderung handelt und das Objekt damit auf der derzeitigen OSM-Karte sichtbar ist oder ob es sich um eine Löschung handelt. Betrachtet wurden für diese Analyse dabei nur Objekte, die bis zum 17.07.2024 gelöscht wurden oder am 17.07.2024 auf der OSM-Karte sichtbar waren. Insgesamt wurden damit rund 1,67 Milliarden Objekte analysiert, davon etwa 1,44 Milliarden Objekte aus der OSM-Karte und knapp 230 Millionen Objekte, die gelöscht wurden. Nicht dazu gehören Objekte, die als Relationen⁵ abgespeichert sind.

Der Zugriff auf den Datensatz erfolgt über das Rechencluster *bwForCluster Helix*. Helix ist Teil der bwHPC-Strategie (High Performance Computing) des Landes Baden-Württemberg, „über die unter anderem an fünf Standorten Hochleistungsrechencluster für verschiedene wissenschaftliche Fachrichtungen zur Verfügung gestellt werden“ (UNIVERSITÄTSRECHENZENTRUM HEIDELBERG 2025). „Ziel von bwHPC ist es, eine [...] Grundversorgung mit Rechen- und Datenhaltungsressourcen für die Forscher:innen im Land zu gewährleisten“ (ebd.).

Abgelegt ist der Datensatz auf dem sogenannten *Scientific Data Storage* bzw. *SDS@hd*. Das ist ein staatlicher Dienstleister des Landes Baden-Württembergs für die sichere Speicherung wissenschaftlicher Daten, auf die häufig zugegriffen wird.⁶ Mit Helix können die darauf abgespeicherten Daten schnell abgefragt und weiterverarbeitet werden.

⁵ <https://wiki.openstreetmap.org/wiki/Relation>

⁶ <https://wiki.bwhpc.de/e/SDS@hd>

Für einige Regionen wurden Daten zu dem zeitlichen Verlauf der sichtbaren OSM-Objekte untersucht, um sie mit dem zeitlichen Verlauf der gelöschten OSM-Objekte aus dem oben genannten Datensatz zu vergleichen. Diese Daten stammen aus dem *Ohsome Dashboard*⁷ der Heigit gGmbH .

Die räumliche Verbreitung der Löschungen wurde mit thematischen Karten visualisiert. Die zugrundeliegenden Staatsgrenzen und Küstenlinien wurden von einem Datensatz der UN übernommen.⁸

⁷ <https://dashboard.ohsome.org>

⁸ <https://public.opendatasoft.com/explore/dataset/world-administrative-boundaries/export/?flg=en-us>

3. Methodik

Die getroffenen Aussagen in dem Ergebniskapitel stützen sich ausschließlich auf eigene Datenanalysen. Dafür wurden mehrere Softwareprogramme verwendet, die in Kapitel 3.1 vorgestellt werden. Anschließend wird thematisiert, wie der globale Raum in Hexagone zergliedert (3.2) und ungenaue Werte präziser gemacht wurden (3.3). Danach wird der Permutationstest beschrieben, mit dem beurteilt wurde, ob in einer Region überproportional viel gelöscht wurde oder nicht (3.4). In Kapitel 3.5 werden die verschiedenen OSM-Objektgruppen vorgestellt und im Anschluss erklärt, wie bestimmte Löschungen von der Analyse ausgeschlossen wurden (3.6). Abschließend werden die untersuchten Konfliktregionen mit ihrem zugrunde liegenden Konflikt beschrieben (3.7).

3.1 Software

Python:

Python ist eine quelloffene und leicht verständliche Programmiersprache⁹. Besonders mächtig wird Python durch seine Bibliothek mit vielen erweiterbaren Funktionalitäten bzw. Paketen. Die Python-Programme für diese Arbeit stützen sich hauptsächlich auf Pakete für einfache Datenanalysen (*Pandas*), für Schnittstellen zu Datenbanken (*duckdb*), für den Umgang mit georeferenzierten Geometrien (*Shapely*, *PyProj*, *Geopandas*) und für einfache Grafiken (*matplotlib*).

R:

R ist eine quelloffene Programmiersprache, die hauptsächlich für statistische Auswertungen verwendet wird¹⁰. Mit den R-Programmen für diese Arbeit wurden die Hexagone für die räumliche Gliederung erstellt, statistische Tests durchgeführt und einfache Ergebnis-Grafiken visualisiert.

QGIS:

QGIS ist eine quelloffene Software für die Erfassung, Verarbeitung, Analyse und Visualisierung von geographischen Informationen¹¹. In dieser Arbeit wurde QGIS verwendet, um Karten zu erstellen, die die räumliche Verbreitung von Löschungen visualisieren.

Eine Auflistung aller Python- und R-Programme, inklusive der verwendeten Pakete, befindet sich auf dem folgenden öffentlichen Github-Repository:
https://github.com/ManuelKraft/master_thesis

⁹ <https://www.python.org/about/>

¹⁰ <https://www.r-project.org/about.html>

¹¹ <https://qgis.org/project/overview/>

3.2 Hexagone als räumliche Gliederung

Bei der Datenanalyse wurde der globale Raum in Hexagone zergliedert. Für jedes Hexagon wurde dabei gezählt, wie viele gelöschte und sichtbare OSM-Objekte sich mit ihm überschneiden.

Mit den Hexagonen wurden die globalen Karten erstellt, welche die räumliche Verbreitung von Löschungen visualisieren. Bei der Wahl der räumlichen Auflösung musste sinnvoll abgewogen werden:

Bei einer zu hohen Auflösung wären einige Hexagone nicht groß genug, um genug Daten zu erfassen. Außerdem erschwert ein zu hoher Rechenaufwand die Datenanalyse. Bei einer zu niedrigen Auflösung wären einige Hexagone zu groß, um relevante kleinräumige Verbreitungsmuster zu erfassen.

In dieser Arbeit wurde festgelegt, dass die Hexagone eine Fläche von rund 288 km² haben bzw. Level 11 der Auflösungsskala von Barnes & Sahr (Barnes & Sahr 2017). Das macht weltweit knapp 1,8 Millionen Geometrien.

Hexagone eignen sich besonders gut für die Darstellung globaler Karten. So gibt es bei der Umpunktierung der gekrümmten Erdoberfläche auf eine zweidimensionale Karte immer Verzerrungen. Diese nehmen mit abnehmendem Maßstab zu. Mit Hexagonen werden diese Verzerrungen minimiert. (BARNES & SAHR 2017; BIRCH et al. 2007, S. 350)

Das ist in dieser Arbeit vorteilhaft, weil viele globale Karten und großflächige regionale Karten gezeigt werden.

Außerdem verbessern die Hexagone die Datenanalyse dahingehend, dass die Ränder näher am Zentrum liegen als bei einem üblichen Rastergitter und dadurch Verzerrungen durch Randeffekte minimiert werden (ebd.). Damit werden räumliche Abfragen verbessert, wie z.B. bei dieser Arbeit die Überschneidung der Hexagone mit OSM-Objekten.

Die Hexagone wurden mit der Programmiersprache R und dem R-Paket *dggridR* erstellt (BARNES & SAHR 2017). Einige wenige Hexagone, deren Projektion sich als schwierig herausgestellt hat, wurden eigenständig mit Python erfasst und herausgefiltert. Um Darstellungsprobleme zu vermeiden, wurden Hexagone entfernt, die den Anti-Meridian schneiden (KRAFT 2025).

3.3 Datenkorrektur

Bei der Analyse der räumlichen Verbreitung von Löschungen wird für jedes Hexagon die Anzahl an gelöschten Objekten und die Anzahl an sichtbaren Objekten ermittelt, die sich mit ihm räumlich überschneiden. Diese Größen korrelieren stark miteinander (Abbildung 6). Die reine Anzahl an gelöschten Objekten reicht also nicht aus für die Bewertung eines auffälligen Löschverhaltens. Deshalb wird die Anzahl an gelöschten Objekten mit der Anzahl an sichtbaren Objekten ins Verhältnis gesetzt:

$$\text{Ratio} = \text{Anzahl gelöschter Objekte} / \text{Anzahl sichtbarer Objekte}$$

In dieser Arbeit ist *Ratio* die entscheidende Maßzahl für die Analyse von Löschungen. Dieser Wert hängt von der Anzahl an sichtbaren Objekten ab (dem Nenner) und wird umso ungenauer, je kleiner der Nenner wird. Auch wird der Wert ungenauer, wenn der Nenner besonders stark streut. Deshalb werden die *Ratio*-Werte schwierig zu vergleichen sein, wenn die Nenner nicht über alle Hexagone einen konstanten Wert haben oder nicht überall besonders groß sind. (ANSELIN et al. 2006, S. 14f.)

Eine Maßnahme, um die *Ratio*-Werte vergleichbarer zu machen, ist die Glättung der Werte mittels bayesianischer Statistik. Das Grundprinzip dabei ist, dass mit zuvor getroffenen Modellannahmen (*Prior*) unbekannte Werte präziser abgeschätzt werden können (*Posterior*) (ANSELIN et al. 2006, S. 40).

In dieser Arbeit wird der sogenannte *Global Linear Empirical Bayes Smoother* angewandt. Die zugrunde liegenden Parameter (*Priors*) werden dabei direkt aus den Daten berechnet. Die Werte nähern sich hierbei dem globalen Mittelwert an. Diese Näherung ist umso stärker, je größer die globale Varianz aller *Ratio*-Werte und je kleiner der Nenner des jeweiligen einzelnen *Ratio*-Wertes ist. (ANSELIN et al. 2006, S. 47f.)

Die Korrektur der Daten anhand dieser Methode erfolgte mit einem eigenständig geschriebenen Python-Programm (KRAFT 2025). Trotz der damit besseren Aussagekraft werden *Ratio*-Werte von Hexagonen mit weniger als 30 sichtbaren Objekten nicht berechnet, weil die Korrektur bei zu kleinen Nennern nur begrenzt hilfreich ist. Stattdessen erhalten diese Hexagone die Anmerkung, dass sie zu wenig Daten beinhalten (siehe u.a. graue Flächen in Abbildung 7).

3.4 Permutationstests zum Vergleich von zwei Regionen

In der Einleitung wurde die These aufgestellt, dass in Konfliktregionen überproportional viele OSM-Objekte gelöscht werden. Um diese These zu überprüfen, werden die Löschungen von Konfliktregionen mit den Löschungen zweier anderer Regionen verglichen:

1. Das direkte Umland, welches etwa viermal so groß ist
2. Die komplette weltweite Landfläche, welche nicht zu den ausgewählten Konfliktregionen gehört

Dafür wird ein statistischer Test durchgeführt, der die Konfliktregion und die jeweilige Vergleichsregion auf einen signifikanten Unterschied ihrer Löschungen testet. In dieser Arbeit wurde der Permutationstest verwendet, weil dieser keine Informationen der zugrundeliegenden Werteverteilung benötigt und die Daten nicht normalverteilt sein müssen (LAUTENBACH 2025).

Das Grundprinzip von Permutationstests ist, dass man die Daten unter der Annahme der Nullhypothese permutiert und die statistische Kennzahl, die von Interesse ist, auf Grundlage der Permutation neu berechnet. Dies macht man sehr oft und vergleicht die Verteilung der so zustande gekommenen statistischen Kennzahl mit dem Wert, der sich aus den beobachteten Daten ergibt. Daraus lässt sich abschätzen, wie wahrscheinlich das Auftreten des beobachteten Wertes unter der Nullhypothese ist. (ebd.)

Bezogen auf diese Arbeit ist die Nullhypothese, dass in der jeweiligen Konfliktregion und in der jeweiligen Vergleichsregion gleich viel gelöscht wird. Die untersuchte statistische Kennzahl ist dabei der Mittelwert.

Im vorherigen Kapitel wurde die Zergliederung des globalen Raumes in Hexagone beschrieben, welche u.a. die Anzahl an gegenständlichen Löschungen und die Anzahl an gegenständlichen sichtbaren Objekten als Wert beigefügt haben (siehe Kapitel 3.2). Aufgrund der starken Korrelation zwischen diesen Größen wird für die Datenanalyse das Verhältnis dieser Werte, kurz *Ratio*, betrachtet (siehe Kapitel 3.3).

Als Ausgangssituation des Permutationstests gibt es n Hexagone, die sich räumlich mit der jeweiligen Konfliktregion überschneiden und n Hexagone, die sich mit der Vergleichsregion überschneiden. Unter Annahme der Nullhypothese können die *Ratio*-Werte beliebig den beiden Verteilungen zugeordnet werden, ohne dass sich die untersuchte statistische

Kennzahl, in diesem Fall der Mittelwert, verändert. Um das zu testen, werden die Hexagone der Konfliktregion und die Hexagone der Vergleichsregion, umgangssprachlich formuliert, miteinander *vermischt* und zwei neuen, jeweils n großen Verteilungen zugewiesen. Von den Verteilungen werden die zwei Mittelwerte berechnet und deren Differenz ermittelt. Dieses Vorgehen wird mehrmals wiederholt, in dieser Arbeit 10.000 mal bei dem Vergleich mit dem Umland und 1.000.000 mal bei dem Vergleich mit der kompletten globalen Landfläche. Das Resultat ist eine Verteilung der simulierten Mittelwertdifferenzen, die mit der Differenz der tatsächlich beobachteten Mittelwerte verglichen werden kann (Abbildung 1).

In dieser Arbeit haben die Vergleichsregionen immer eine größere Fläche als die Konfliktregionen, sodass es dort stets mehr überschneidende Hexagone gibt. Deshalb werden vor jeder Permutation n zufällige Hexagone der Vergleichsregion ausgewählt.

Karabach vs. Umland

Mittelwert Karabach ≈ 0.139 , Mittelwert Umland ≈ 0.071
Differenz der beobachteten Mittelwerte ≈ 0.068

Permutationstest

Nullhypothese: Beobachteter Mittelwert für Karabach nur zufällig größer

Alternativhypothese: Beobachteter Mittelwert für Karabach nicht zufällig größer

p-Wert $< 1e-03 < 0.05 \rightarrow$ Verwerfung der Nullhypothese

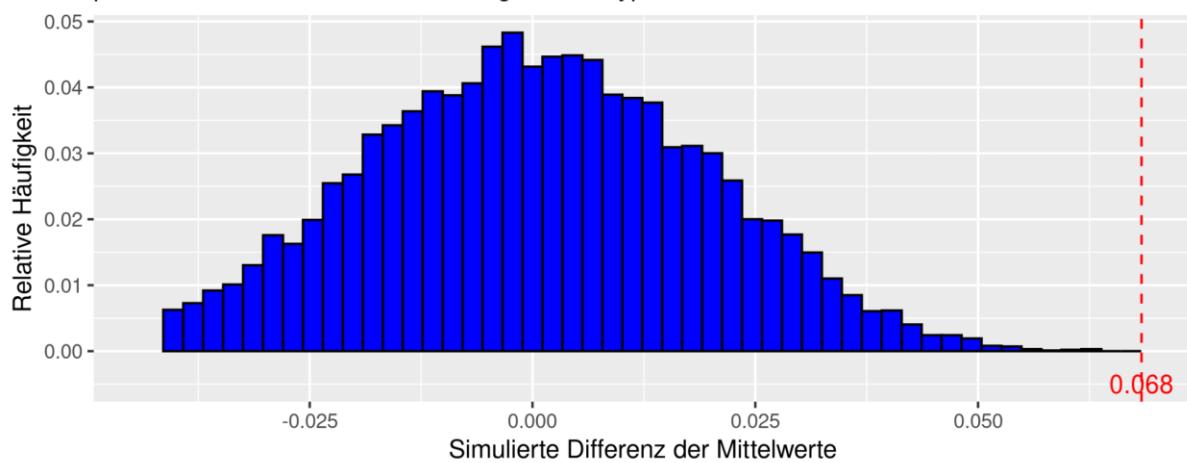


Abbildung 1: Fallbeispiel mit Anwendung des Permutationstests

3.5 Unterschiedene Objekte

Objekte werden in OpenStreetMap mithilfe von sogenannten *Tags* repräsentiert. Ein Tag besteht aus einem Schlüssel-Wert-Paar (engl.: *key, value*) und ist ein Attribut des Objektes¹². Es gibt dabei keine offizielle Reglementierung, wie viele Tags mit welchem Inhalt einem Objekt hinzugefügt werden. Dennoch hat sich die OSM-Community auf informelle Standards geeinigt, mit denen die meisten Objekte beschrieben werden können (ebd.). Diese Standards sind nicht statisch, sondern können sich ändern. So kommen häufig Vorschläge auf, welche neuen Tags man für welche Objekte benutzen könnte¹³. Die Unterscheidung der OSM-Objekte in dieser Arbeit basiert auf dem Stand von Januar 2025.

3.5.1 Gegenständliche Objekte

Die meisten Objekte in OpenStreetMap repräsentieren physisch sichtbare Gegenstände in der realen Welt. Diese Objekte werden innerhalb der deutschen OSM-Community als *gegenständliche* Objekte und in der englischen OSM-Community als *Map Features* bezeichnet.

Dabei gibt es eine Liste an Keys, welche üblicherweise für solche Objekte verwendet werden. Nicht immer werden die Objekte damit exakt beschrieben, sondern sie haben weitere Tags beigefügt. Dennoch beinhalten sie üblicherweise einen Tag mit einem der folgenden 29 Keys:

Aerialway, Aeroway, Amenity, Barrier, Boundary, Building, Craft, Emergency, Geological, Healthcare, Highway, Historic, Landuse, Leisure, Man Made, Military, Natural, Office, Place, Power, Public Transport, Railway, Route, Shop, Sport, Telecom, Tourism, Water, Waterway

In dieser Arbeit werden alle OSM-Objekte als *Gegenständlich* bezeichnet, die mindestens einen Tag mit einem dieser Keys haben.

3.5.2 Adressen

OpenStreetMap beinhaltet auch Objekte, die in der realen Welt nicht sichtbar sind. Dazu gehören Adressen, also Hausnamen, Straßennamen oder Hausnummern. Üblicherweise sind Adressinformationen einem gegenständlichen Objekt als zusätzliche Tags angefügt.

¹² https://wiki.openstreetmap.org/w/index.php?title=Map_features&oldid=2804789

¹³ <https://wiki.openstreetmap.org/w/index.php?title=Changelog&oldid=2778521>

Allerdings zählt es auch zum Standard, ein separates Objekt mit der Adressinformation anzulegen. Das gilt, wenn z.B. ein Gebäude mehrere Hausnummern hat oder die Geometrie des zugehörigen Objektes unbekannt ist. Dennoch ist in den meisten Fällen das Anlegen eines separaten Objektes nicht notwendig und einige Programme, welche die OSM-Datenqualität bewerten, zeichnen dafür einen Fehler auf. Verkompliziert wird die Sachlage dadurch, dass OSM-Communities einzelner Länder unterschiedliche Handlungsweisen vorgeben.¹⁴

Daher können solche Objekte, müssen aber nicht auf eine niedrige Datenqualität hinweisen und werden deshalb als separate Gruppe aufgeführt. In dieser Arbeit werden alle OSM-Objekte als Adressen angesehen, die nicht gegenständlich sind (3.2.1) und die mindestens einen Tag mit einem Key haben, der das Präfix ‘addr:’ beinhaltet.

3.5.3 Objekte ohne Informationen

Objekte, die keine Informationen beigefügt haben, können keinem Objekt in der realen Welt zugeordnet werden und gelten deshalb als Hinweis auf eine niedrige Datenqualität.

Es gibt jedoch auch informationslose Objekte, die Teilmenge einer *Relation* sind. Eine Relation ist ein spezieller Datentyp, um logische Beziehungen zwischen Objekten zu modellieren und verweist dafür auf andere Objekte. Anzuwenden sind Relationen u.a. bei sogenannten *Multi-Polygonen*, also Flächen mit verknüpften “Exklaven” oder “Löchern”. Um z.B. eine Insel innerhalb eines Sees abzuspeichern, müssen zwei Objekte angelegt werden. Ein äußerer Ring (*way*) für den See und ein innerer Ring für die Insel. Eine Relation verweist nun auf den See und die Insel. Die Tags zu dem See sind dann meist in der Relation enthalten, aber nicht in dem See-Objekt selbst.¹⁵

Relationen werden in dem analysierten Datensatz nicht behandelt, dafür aber deren verwiesene Objekte. Diese gehören zum informellen Standard der OSM-Community, seien ihnen Informationen beigefügt oder nicht. Objekte, die weder Informationen beigefügt haben, noch Teil einer Relation sind, gehören dagegen nicht dazu. Daraus kann geschlussfolgert werden, dass Objekte ohne Informationen, ebenso wie Adressen, auf eine niedrige Datenqualität hinweisen können, aber nicht müssen.

¹⁴ <https://wiki.openstreetmap.org/w/index.php?title=Key:addr:&oldid=2795834>

¹⁵ <https://wiki.openstreetmap.org/w/index.php?title=Relation&oldid=2773710>

3.5.4 Sonstige Objekte

Objekte, die weder gegenständlich sind noch Adressen beschreiben, entsprechen im Regelfall nicht den Konventionen der OSM-Community und können deshalb nicht eindeutig einem Objekt in der realen Welt zugeordnet werden. Deshalb gelten sie in dieser Arbeit als Hinweis auf eine niedrige Datenqualität.

3.6 Ignorierte Löschungen

Nicht alle Löschungen in der OpenStreetMap-Datenbank sind von Dauer. Häufig löschen Nutzer ein Objekt, um es in leicht abgewandelter Form neu zu erstellen. Solche Objekte wurden demnach nicht wirklich gelöscht, sondern modifiziert. Modifizierungen sind zwar in OSM durchführbar, aber nicht alle Kartierer nutzen diese Möglichkeit.

In dieser Arbeit wurden alle Löschungen von der Analyse ausgeschlossen, für die in einem fast identischen Zeitraum ein neues Objekt erstellt wurde, das die gleichen Informationen (Tags) beigefügt hat und eine ähnliche Geometrie besitzt (Abbildung 2). Insgesamt wurden damit etwa 20,5 Millionen 'gelöschte' Objekte erfasst und von der weiteren Analyse ausgeschlossen. Das entspricht rund 9% aller Löschungen im Datensatz.

Wenn ein Nutzer eine Menge an Änderungen auf der OSM-Karte macht, werden diese Änderungen in einem sogenannten *Changeset* gruppiert abgespeichert. Ein gelöschtes und ein neu erstelltes Objekt gelten in dieser Arbeit als zeitlich nahezu identisch, wenn sie im gleichen *Changeset* gespeichert sind bzw. die gleiche Changeset-ID haben.

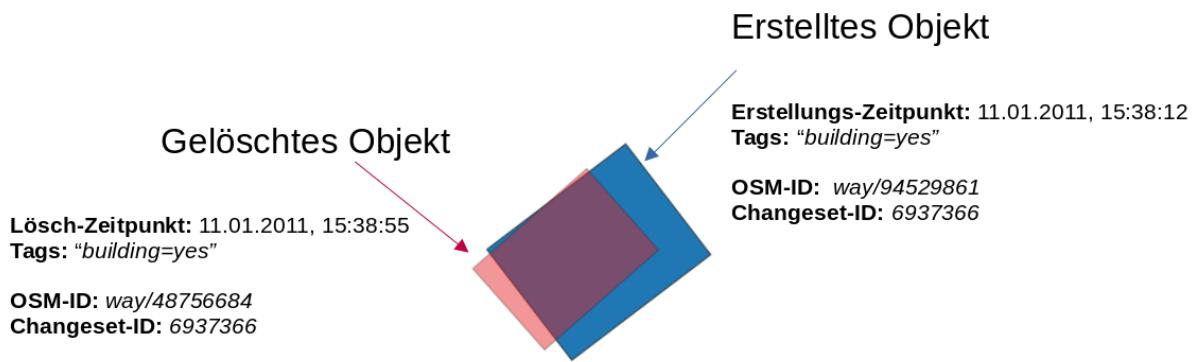


Abbildung 2: Reales Fallbeispiel eines gelöschten Objektes, bei dem zu einem fast identischen Zeitpunkt ein neues Objekt mit ähnlicher Geometrie und gleichen beigefügten Informationen (Tags) hinzugefügt wurde

3.6.1 Hausdorff-Distanz als Maß für geometrische Ähnlichkeit

Um zu bestimmen, ab wann ein gelöschtes Objekt eine ähnliche Geometrie zu einem neu erstellten Objekt hat, wurde die sogenannte Hausdorff-Distanz zwischen beiden Objekten abgemessen. Die Hausdorff-Distanz ist eine Maßeinheit, mit der die Ähnlichkeit zwischen Geometrien beschrieben werden kann.

Seien A und B zwei Geometrien. Dann wird von jedem Punkt a in A die Distanz zum kürzesten Punkt in B abgemessen. Die maximale Distanz aus der berechneten Menge ist die Hausdorff-Distanz.

$$H'(A, B) = \max_{a \in A} d(a, B)$$

(NADLER 1978, in: HUNG & YANG 2004)

Weil die Hausdorff-Distanz von B zu A anders sein kann, ist der größere der beiden Werte die finale Hausdorff-Distanz.

$$H(A, B) = \max\{H'(A, B), H'(B, A)\}$$

(ebd.)

Um zu entscheiden, was die maximale Hausdorff-Distanz ist, bis zu der zwei Geometrien als ähnlich gelten, muss sinnvoll abgewogen werden: Bei einem zu hohen Schwellenwert würden zu häufig zwei Geometrien als ähnlich betrachtet werden, die in Wirklichkeit nicht

ähnlich sind. Bei einem zu niedrigen Schwellenwert würden zu häufig zwei Geometrien unrechtmäßig als verschieden betrachtet werden. In dieser Arbeit gelten zwei Geometrien als ähnlich, wenn die Hausdorff-Distanz maximal 30 Meter beträgt. Die zwei OSM-Objekte aus dem vorherigen Fallbeispiel gelten demnach als ähnlich, weil deren Hausdorff-Distanz 3,5 Meter beträgt und damit unterhalb des Schwellenwertes von 30 Metern liegt (Abbildung 3).

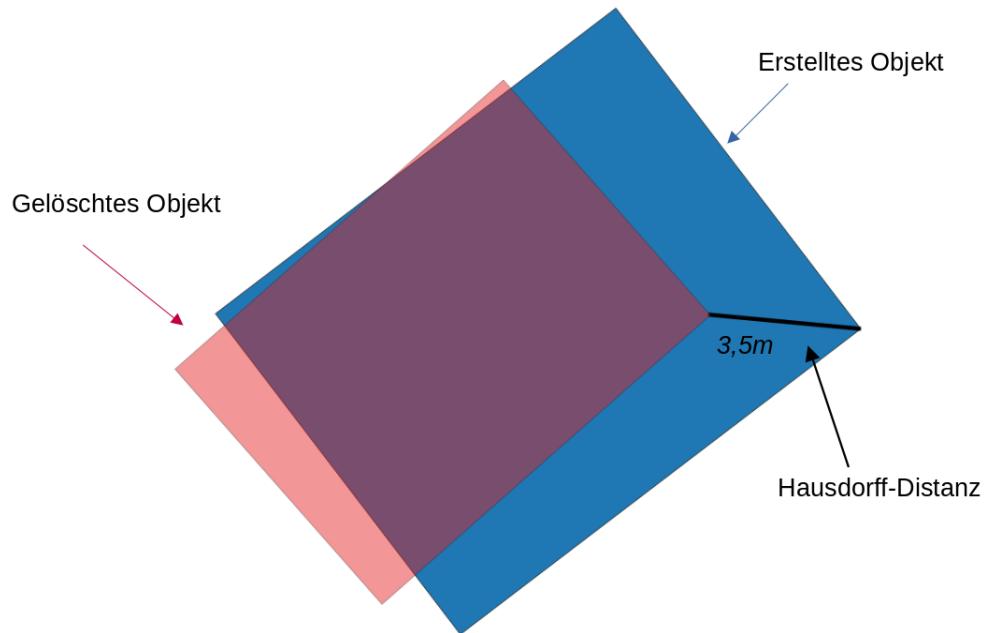


Abbildung 3: Anwendung der Hausdorff-Distanz

3.7 Untersuchte Konfliktregionen

In diesem Abschnitt werden alle untersuchten Konfliktregionen aufgelistet und deren zugrunde liegender Konflikt kurz beschrieben. Gemeinsam ist allen Regionen, dass ihr Territorium seit über 15 Jahren von mindestens zwei Mächten beansprucht wird.

Westjordanland:

Das Westjordanland liegt zwischen Israel und Jordanien. Von der UN-Vollversammlung wurde es im Teilungsplan von 1947 dem zu gründenden arabischen Staat Palästina zugesprochen. Seit dem Sechs-Tage-Krieg von 1967 wird es von Israel besetzt. Von den Palästinensern dagegen wird es als Teil eines eigenen zukünftigen Staates beansprucht. Mit den Osloer Verträgen von 1990 und 1995 sollte der gestaffelte Rückzug israelischer Streitkräfte aus dem Westjordanland erfolgen. Ziel war der schrittweise Weg einer friedlichen Zweistaaten-Lösung mit einem souveränen palästinensischen Staat. (RND 2023)

Mittlerweile hätte sich Israel aus dem Westjordanland längst zurückziehen müssen. Doch dazu ist es nie gekommen. Israel kontrolliert weiterhin mehr als 60% der Region. Die expansive Siedlungspolitik der letzten Jahre von Seiten der israelischen Regierung macht einen Rückzug noch unwahrscheinlicher. Dabei gibt es viele Zwischenfälle von radikalen Siedlern, die gewaltsam gegen Palästinenser vorgehen und ihre Landflächen einnehmen. Seit dem Terrorangriff der palästinensischen Terrororganisation Hamas auf Israel am 07.10.2023 treten die radikalen Siedler noch gewaltamer auf.

(TAGESSCHAU 2024)

Derzeit kommt es im Westjordanland regelmäßig zu Militäreinsätzen mit in der Folge massiven gewalttätigen Ausschreitungen (AUSWÄRTIGES AMT 2025).

Westsahara:

Die Westsahara ist ein Territorium ohne Selbstverwaltung, das von Marokko besetzt ist. Bis 1976 war die Westsahara offiziell eine spanische Kolonie. Gegen Ende der Kolonialzeit wurde die Volksfront zur "Befreiung" gegründet (Polisario-Front). Diese kämpft für eine unabhängige Demokratische Arabische Republik Sahara auf dem gesamten Territorium. Gleichzeitig erhoben Marokko und Mauretanien territoriale Ansprüche in der Region. 1975 drangen diese Länder in die Westsahara ein, beendeten die spanische Kolonialherrschaft und zwangen tausende Sahrauis zur Flucht. Die sahrauischen Flüchtlinge gründeten

daraufhin eine Exilregierung in Algerien. Mauretanien zog sich 1979 aus der Westsahara zurück, sodass Marokko auch den südlichen Teil besetzte.

Der Krieg zwischen der Polisario-Front und Marokko dauerte bis 1991, als ein Waffenstillstandsabkommen unterzeichnet wurde. Unter Leitung der Vereinten Nationen sollte ein Selbstbestimmungsreferendum durchgeführt werden, das über den Status der Westsahara entscheiden sollte. Streitigkeiten über die Wählerschaft verhinderten jedoch lange dessen Durchführung. Daraufhin schlugen die Vereinten Nationen 2003 einen Plan zur Volkszählung vor. Marokko lehnte diesen Vorschlag jedoch ab, sodass das Referendum über den Status der Westsahara seit über 30 Jahren aussteht.

Während des Krieges begann Marokko 1981 mit der Errichtung einer mittlerweile 2.700 Kilometer langen, befestigten Sand- und Steinmauer (Berm), die die Westsahara in zwei Gebiete teilt. Der westliche Teil wird von Marokko kontrolliert und der östliche und südliche Teil von der Polisario-Front.

(BUNDESZENTRALE FÜR POLITISCHE BILDUNG 2021)

Kaschmir:

Kaschmir ist eine Gebirgsregion im Nordwesten des Himalayas, die sowohl von Indien als auch von Pakistan beansprucht wird. Nach der Teilung Britisch-Indiens 1947 in das säkulare Indien und das muslimische Pakistan gab es Uneinigkeit über die Zugehörigkeit Kaschmirs. Das führte zu einem Krieg, der 1949 mit einem Waffenstillstandsabkommen und einer Waffenstillstandslinie endete. Diese Linie entspricht der heutigen de-facto-Grenze zwischen Indien und Pakistan. Der südöstliche Teil Kaschmirs ist unter Kontrolle Indiens und der nordwestliche Teil unter Kontrolle Pakistans.

Nach einer langsamen Normalisierung haben sich die Beziehungen seit 2014 wieder verschlechtert, weil der neue indische Premierminister Modi einen konfrontativen Kurs verfolgt und Indien als aufstrebende Großmacht ansieht.

Militante separatistische Gruppen innerhalb Kaschmirs und der indische Anspruch auf die von China besetzte angrenzende Region Aksai Chin machen Kaschmir zu einer der stärksten militarisierten Regionen der Welt.

(BUNDESZENTRALE FÜR POLITISCHE BILDUNG 2024)

Bergkarabach:

Bergkarabach ist eine gebirgige Region, die sowohl von Armenien als auch von Aserbaidschan beansprucht wird. Die 1991 erklärte Unabhängigkeit des überwiegend armenisch besiedelten Bergkarabach von Aserbaidschan wurde international nie anerkannt, militärisch jedoch durch Armenien abgesichert.

Die Ursprünge der Auseinandersetzung gehen bis ins 18. Jahrhundert zurück, eskalierten jedoch im Zuge des Zerfalls der Sowjetunion. Nach verlustreichen kriegerischen Auseinandersetzungen von 1992 bis 1994 gab es ein Waffenstillstandsabkommen. Damit endeten zwar die Kampfhandlungen, ein Friedensabkommen mit einer nachhaltigen Lösung des Konflikts gab es aber nicht.

Im September 2023 eroberte Aserbaidschan mit einer militärischen Großoffensive die Region. Seit dem 01.01.2024 ist Bergkarabach offiziell aserbaidschanisches Staatsgebiet. (LANDESZENTRALE FÜR POLITISCHE BILDUNG BADEN-WÜRTTEMBERG 2024)

Arunachal Pradesh:

Arunachal Pradesh liegt im westlichen Himalaya und wird von Indien kontrolliert. China erhebt jedoch auch Anspruch auf die Region. Der Konflikt geht mindestens bis ins Jahr 1914 zurück. Damals einigten sich Britisch-Indien und Tibet darauf, dass die Provinz Arunachal Pradesh Teil des Territoriums Britisch-Indiens wird. China zweifelt die völkerrechtliche Verbindlichkeit der damaligen Konvention an, da Tibet niemals Souveränität erlangt habe und somit keine Partei eines völkerrechtlichen Vertrags habe sein können. (KONRAD ADENAUER STIFTUNG 2023)

Immer wieder kommt es zu Auseinandersetzungen und Provokationen an der chinesischen Grenze. (IPG-JOURNAL 2023)

4. Ergebnisse

In diesem Kapitel werden die Ergebnisse der Datenanalyse zu den OSM-Löschen vorgestellt. Dabei werden zunächst die globalen Ergebnisse (4.1 und 4.2), anschließend die Ergebnisse zu den Regionen mit besonders vielen Löschen (4.3) und die Ergebnisse zu den Konfliktregionen thematisiert (4.4).

4.1 Globale Räumliche Verbreitung

Die großen Industriestaaten aus Nordamerika, Europa und Japan haben die meisten Löschen, gefolgt von besonders bevölkerungsreichen Regionen wie Indien und Südostasien (Abbildung 4).

Auffällig sind Regionen, die wenig bis gar nicht besiedelt sind und dennoch viele Löschen haben. Dazu zählt der Punkt mitten im Atlantischen Ozean, wo sich der 0. Längengrad (Meridian) mit dem 0. Breitengrad (Äquator) kreuzt.

Löschen gibt es besonders häufig dort, wo es auch viele derzeit sichtbare OSM-Objekte gibt (Vergleich Abbildung 4 und Abbildung 5). Der Korrelationskoeffizient nach Pearson beträgt rund 0.72 (Abbildung 6). Das gilt nach Cohen als besonders starker Zusammenhang (COHEN 1988, in: GIGNAC & SZODORAI 2016, S. 74).

In dieser Arbeit werden Löschen deshalb ins Verhältnis zu den derzeit existierenden OSM-Objekten gesetzt (Abbildung 7). Dabei zeigt sich, dass großflächige Unterschiede zwischen ganzen Kontinenten oder zwischen dem globalen Norden und dem globalen Süden nicht zu erkennen sind. Die USA, Japan, Türkei und Polen weisen als einzige großflächige Staaten überproportional viele Löschen auf, ansonsten sind keine großflächigen Verbreitungsmuster auszumachen. Stattdessen gibt es viele lokale Auffälligkeiten, auf die in Kapitel 4.3 genauer eingegangen wird.

Während Abbildung 7 die globale Verteilung aller Löschen visualisiert, zeigt Abbildung 8 die globale Verteilung aller Löschen von gegenständlichen Objekten. Beim Vergleich beider Abbildungen ist erkennbar, dass gegenständliche Objekte seltener gelöscht werden.

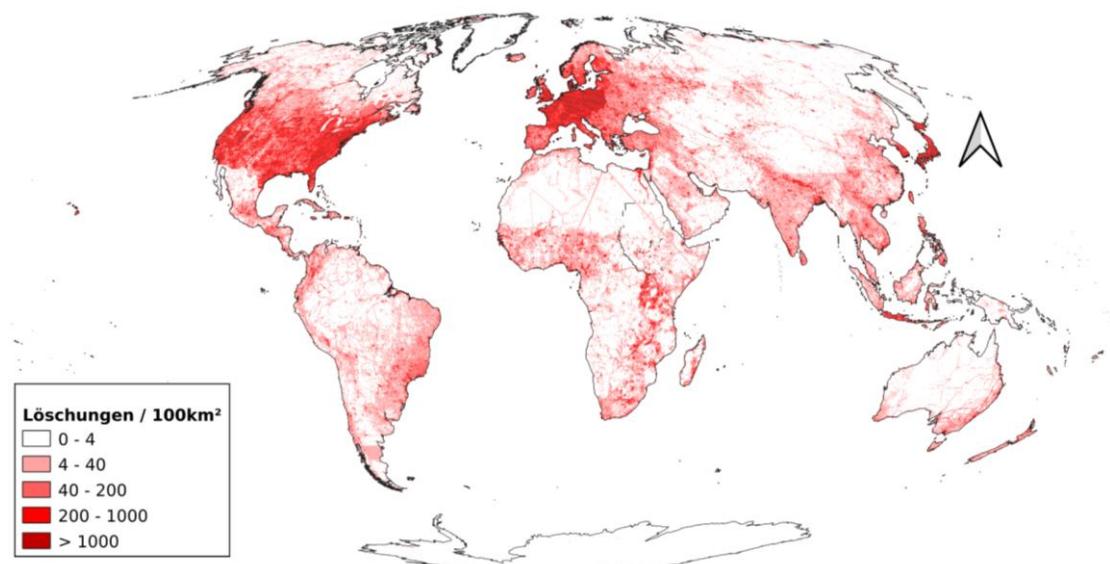


Abbildung 4: Globale Verteilung von gelöschten OSM-Objekten

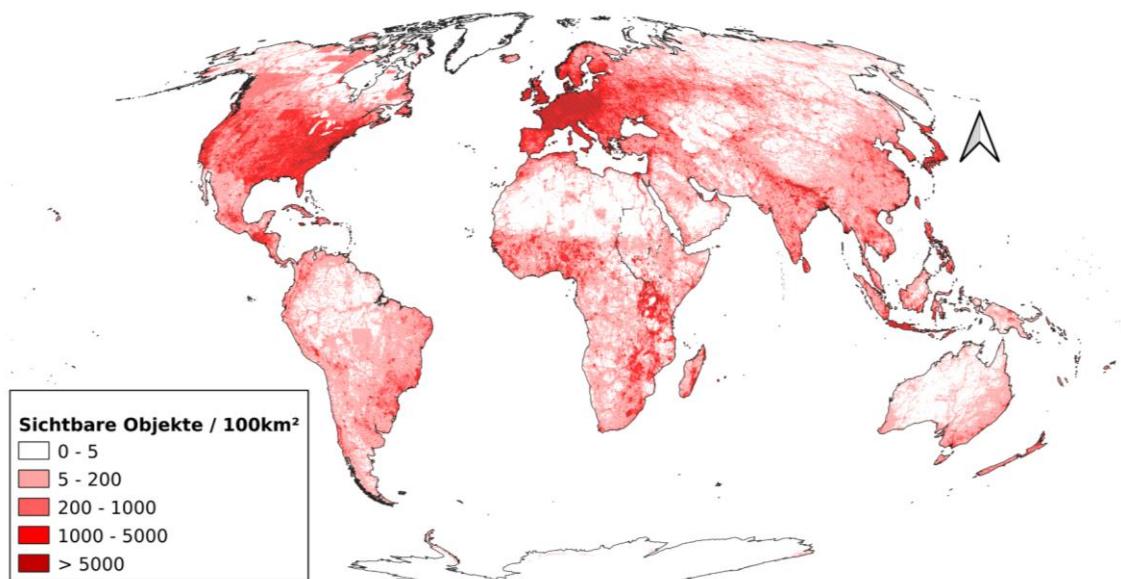


Abbildung 5: Globale Verteilung von sichtbaren OSM-Objekten

Korrelation zwischen Anzahl an Löschungen und Anzahl sichtbarer Objekte
Pearsons R ≈ 0.72

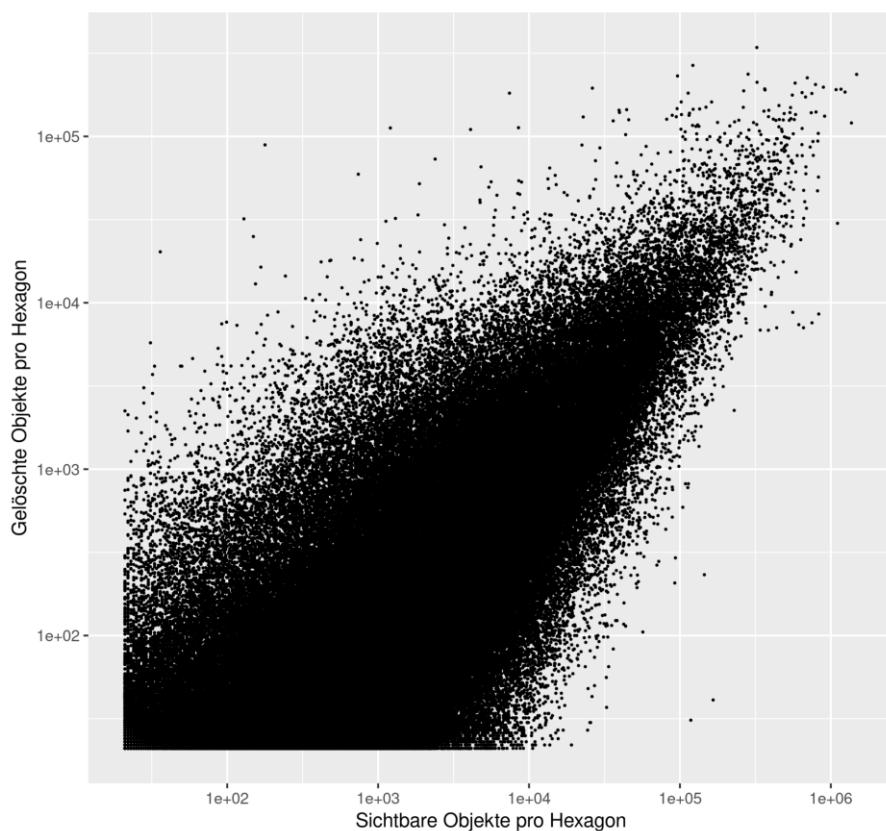


Abbildung 6: Korrelation zwischen Anzahl gelöschter- und Anzahl sichtbarer Objekte

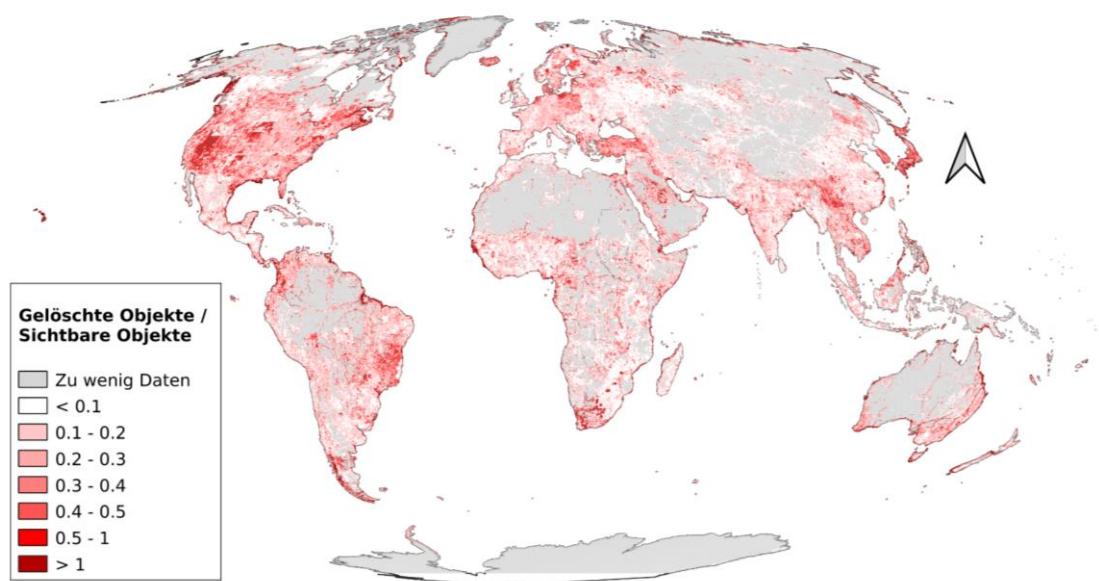


Abbildung 7: Globale Verteilung von Löschungen normiert durch die Anzahl sichtbarer Objekte

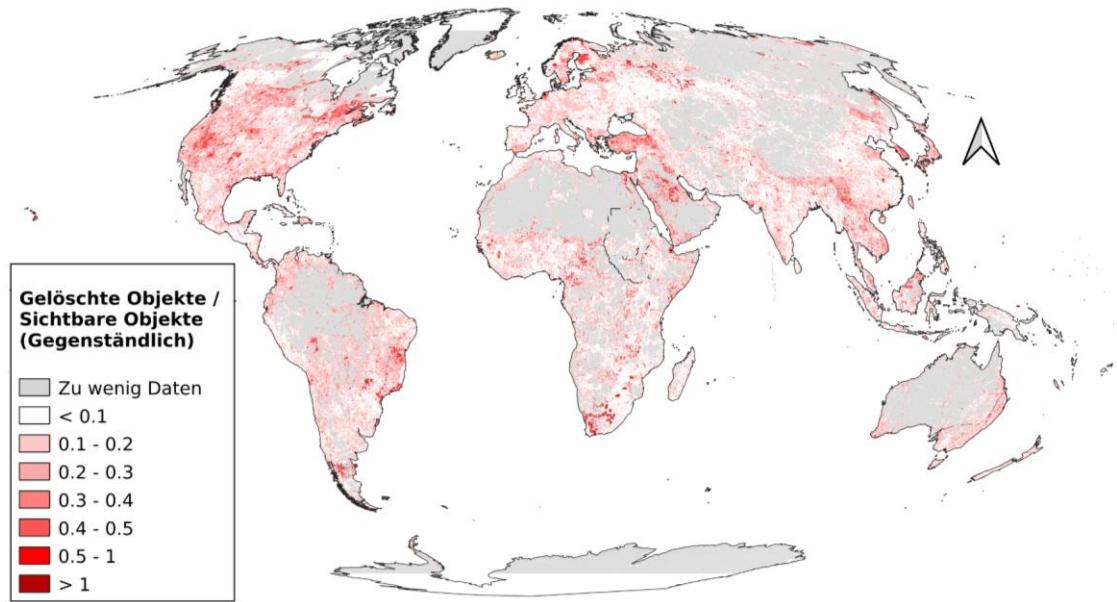


Abbildung 8: Globale Verteilung von Löschungen (nur gegenständliche Objekte) normiert durch die Anzahl sichtbarer Objekte (nur gegenständliche Objekte)

4.2 Globale Objektunterschiede

Gegenständliche Objekte (3.5.1) machen zwar den größten Teil aller Löschungen aus, werden aber im Verhältnis zu ihrer Häufigkeit auf der sichtbaren Karte selten gelöscht. Während ihr Anteil auf der OSM-Karte fast 80% beträgt, liegt ihr Anteil an allen Löschungen nur bei rund 51%. Adressen werden verhältnismäßig häufig gelöscht (Fallbeispiel Abbildung 10). Ihr Anteil an allen Löschungen beträgt etwa 8% und das sind drei Prozentpunkte mehr als ihr Anteil an der OSM-Karte. Verhältnismäßig häufiger gelöscht werden Objekten ohne Informationen (Fallbeispiel Abbildung 11). Ihr Anteil an der OSM-Karte liegt bei knapp 13%, während ihr Anteil an allen Löschungen etwa 22% beträgt. Am verhältnismäßig häufigsten gelöscht werden Objekte, die nicht nach den informellen Standards der OSM-Community getaggt wurden (Fallbeispiel Abbildung 12). Ihr Anteil an der OSM-Karte macht lediglich 2,5% aus, ihr Anteil an allen Löschungen dagegen knapp 19%. (Abbildung 9)

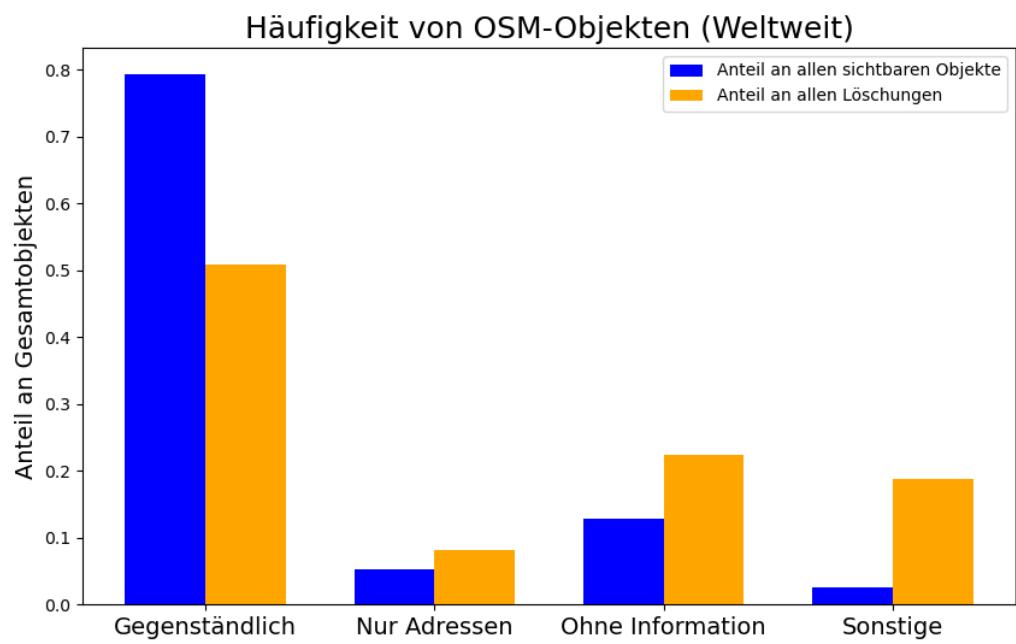


Abbildung 9: Häufigkeit der OSM-Objektgruppen weltweit

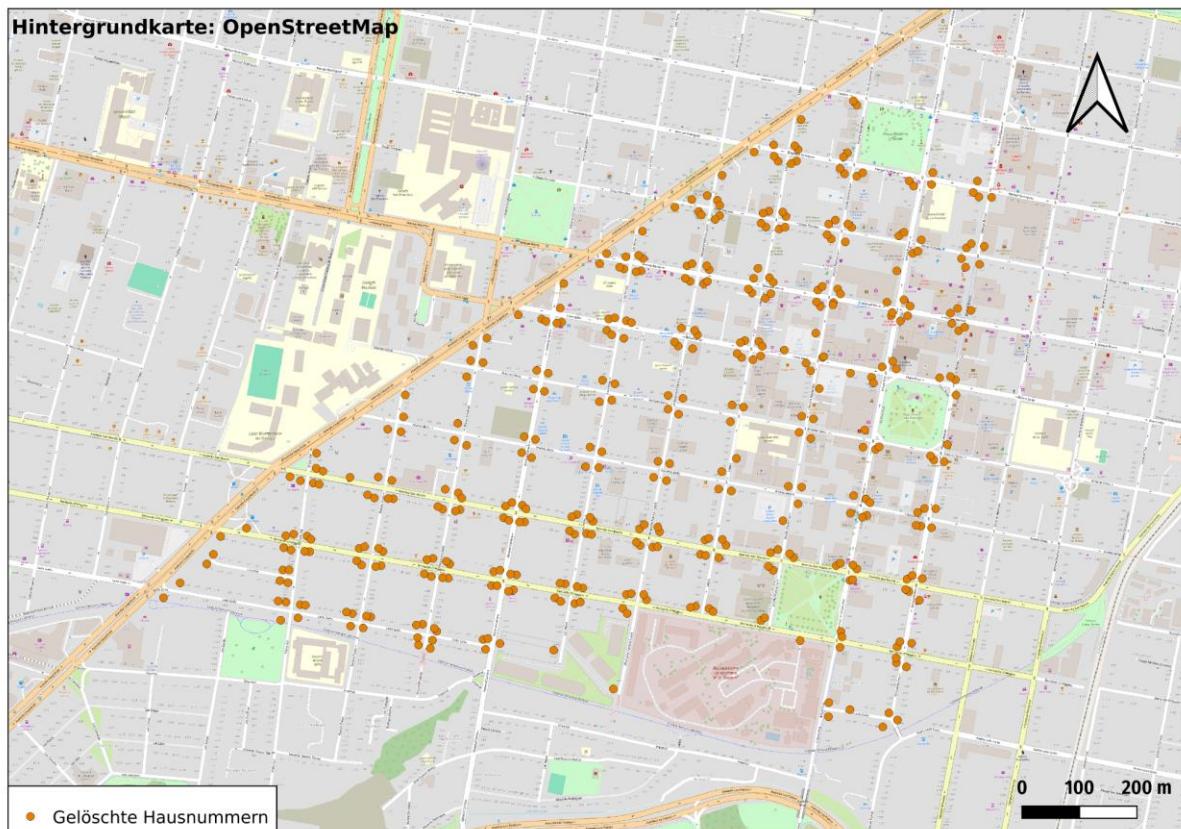


Abbildung 10: Fallbeispiel gelöschter Adressen in Temuco, Chile¹⁶

¹⁶ <https://osm.org/changeset/26258719>



Abbildung 11: Fallbeispiel eines gelöschten informationslosen Objektes in Norwegen¹⁷



Abbildung 12: Fallbeispiel eines gelöschten Objektes mit unausreichenden Informationen bei Kassel¹⁸

¹⁷ <https://osm.org/changeset/21240694>

¹⁸ <https://www.osm.org/changeset/13107839>

Der größte Anteil von gelöschten gegenständlichen Objekten entfällt mit rund 40% auf Gebäude (*building*). Allerdings repräsentieren sie auch mehr als jedes zweite gegenständliche Objekt auf der sichtbaren Karte und werden deshalb verhältnismäßig selten gelöscht. Den zweitgrößten Anteil an gelöschten gegenständlichen Objekten machen Straßen (*highway*) mit rund 28% aus. Mit deutlichem Abstand folgen natürliche Flächen (*natural*), Nutzungsflächen (*landuse*) und relevante Einrichtungen (*amenity*). (Abbildung 13)

Von den gelöschten gegenständlichen Objekten werden Grenzen (*boundary*) am verhältnismäßig häufigsten gelöscht. Das sind Objekte, die politische und andere administrative Flächen abgrenzen. Deren Anzahl an Löschungen entspricht etwa zwei Dritteln der Anzahl an Grenzen in der OSM-Karte. Auch Objekte zur Repräsentation von Siedlungen (*place*), touristische Einrichtungen (*tourism*) und Verkaufsorte (*shop*) werden überproportional häufig gelöscht. Am verhältnismäßig seltensten gelöscht werden Objekte für Telekommunikationssysteme (*telecom*) und Elektrizitäts-Einrichtungen (*power*). Deren Anzahl an Löschungen entspricht weniger als 4% von ihrer Anzahl in der OSM-Karte. Auch stehende Gewässer (*water*), Barrieren (*barrier*) wie Zäune oder Wände, Gesundheitseinrichtungen (*healthcare*), Geologische Ortsmarkierungen (*geological*), Gebäude allgemein (*building*) und handwerkliche Einrichtungen (*craft*) werden vergleichsweise selten gelöscht. Deren Anzahl an gelöschten Objekten ist stets weniger als 8% der Anzahl auf der OSM-Karte. Keinen besonderen Effekt gibt es u.a. bei kfz-befahrenen Straßen (*highway*) und militärischen Objekten (*military*). (Abbildung 15)

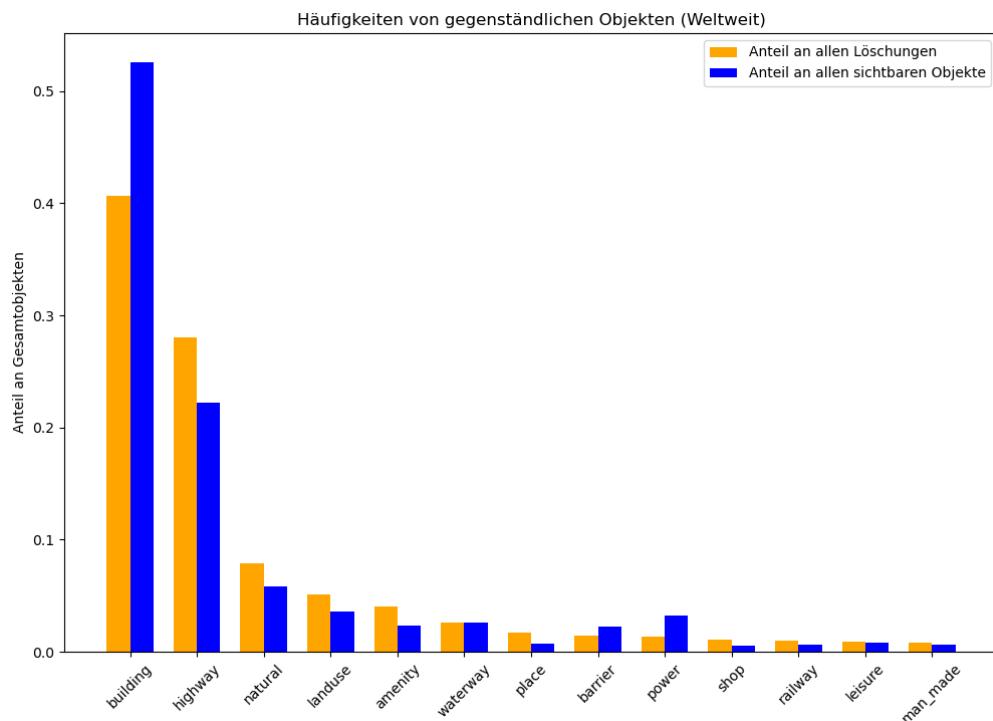


Abbildung 13: Häufigkeit oft vorkommender gegenständlicher Objekte

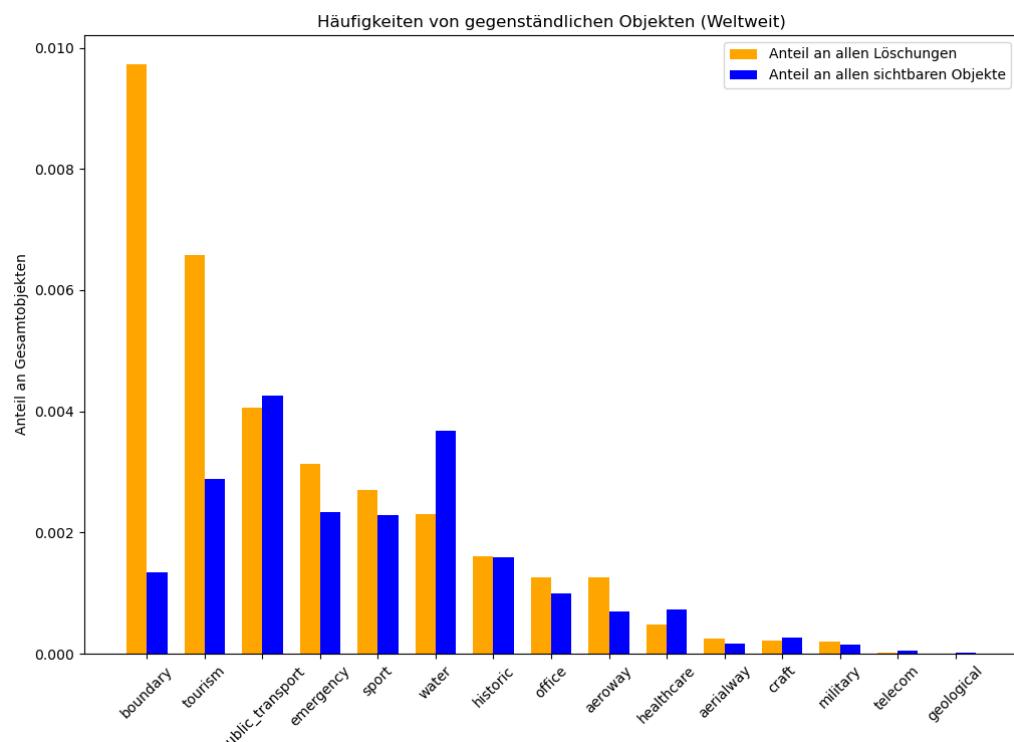


Abbildung 14: Häufigkeit selten vorkommender gegenständlicher Objekte

Löschen von gegenständlichen Objekten (Weltweit)

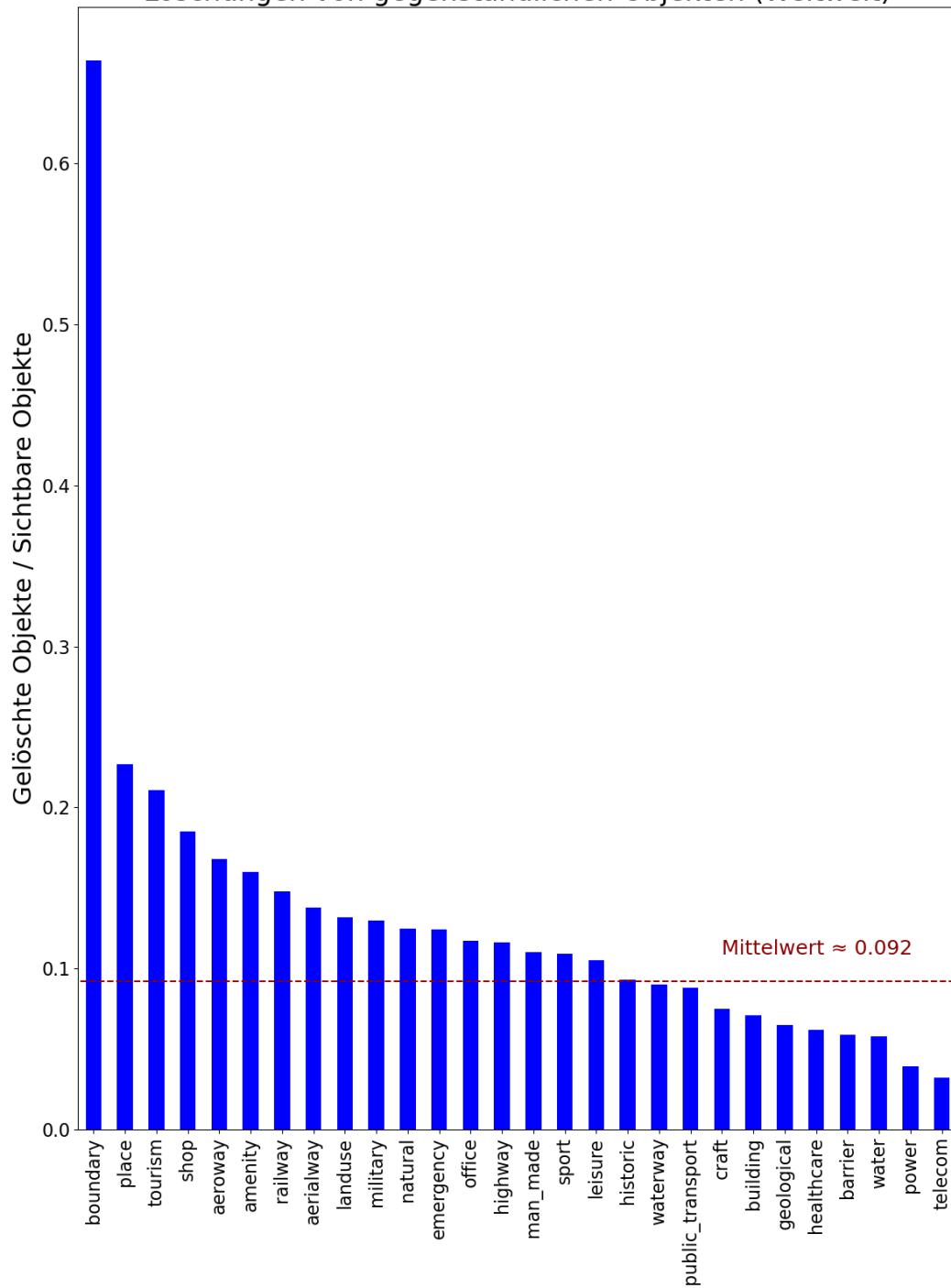


Abbildung 15: Verhältnis zwischen gelöschten und sichtbaren gegenständlichen Objekten

4.3 Löschungsreiche Regionen

Bei der Betrachtung der Weltkarte in Abbildung 7 fallen einige weitere Regionen auf, in denen deutlich mehr gelöscht wurde als im Umland. Die Hintergründe dazu werden in diesem Kapitel präsentiert.

4.3.1 Dschibuti

In Dschibuti übersteigt die Anzahl gelöschter Objekte größtenteils die Anzahl sichtbarer Objekte. Im direkten Umland fallen die Werte deutlich ab. (Abbildung 16)

Nahezu 100% aller Löschen sind Gebäude (Abbildung 17) und haben den folgenden Kommentar beigelegt: "*undiscussed and undocumented import*"¹⁹

Im zeitlichen Verlauf wird deutlich, dass die Löschen innerhalb eines kurzen Zeitraumes im Jahr 2022 passiert sind. Kurz zuvor stieg die Anzahl sichtbarer Gebäude sprunghaft an. Auch im Jahr 2020 gab es während des Anstieges der sichtbaren Objekte mehr Löschen als üblich, wenn auch in einem deutlich geringeren Ausmaß als bei dem Ereignis von 2022. (Abbildung 18)

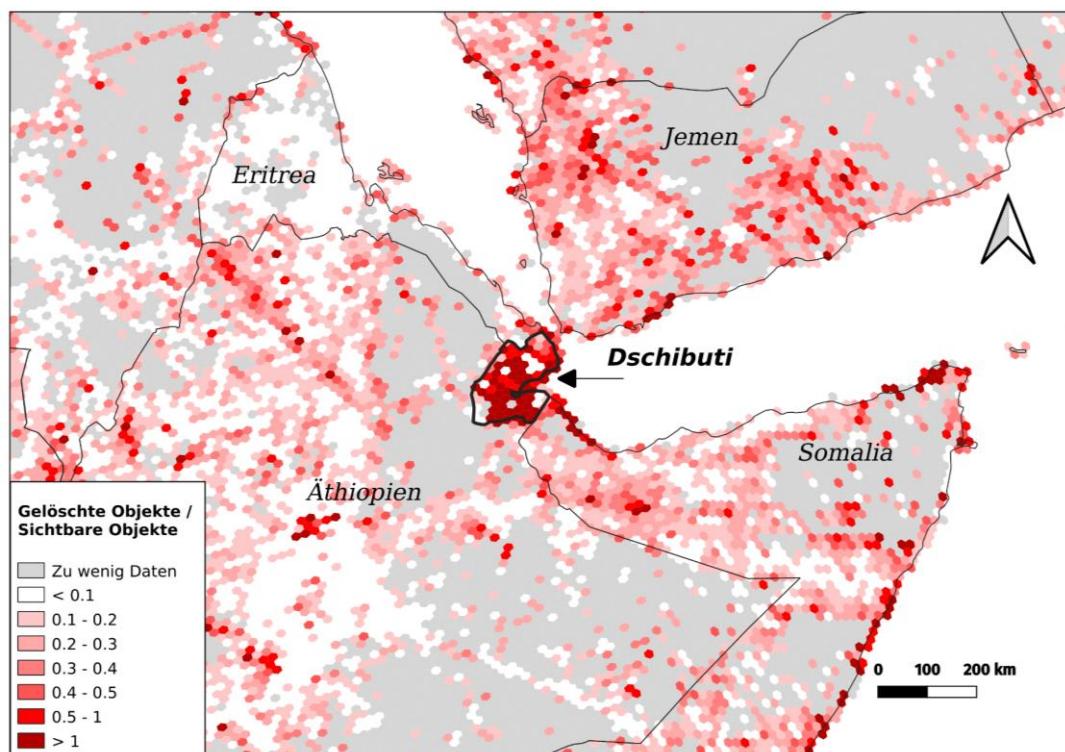


Abbildung 16: Löschen in der Region um Dschibuti

¹⁹ u.a. <https://www.osm.org/changeset/121965162>

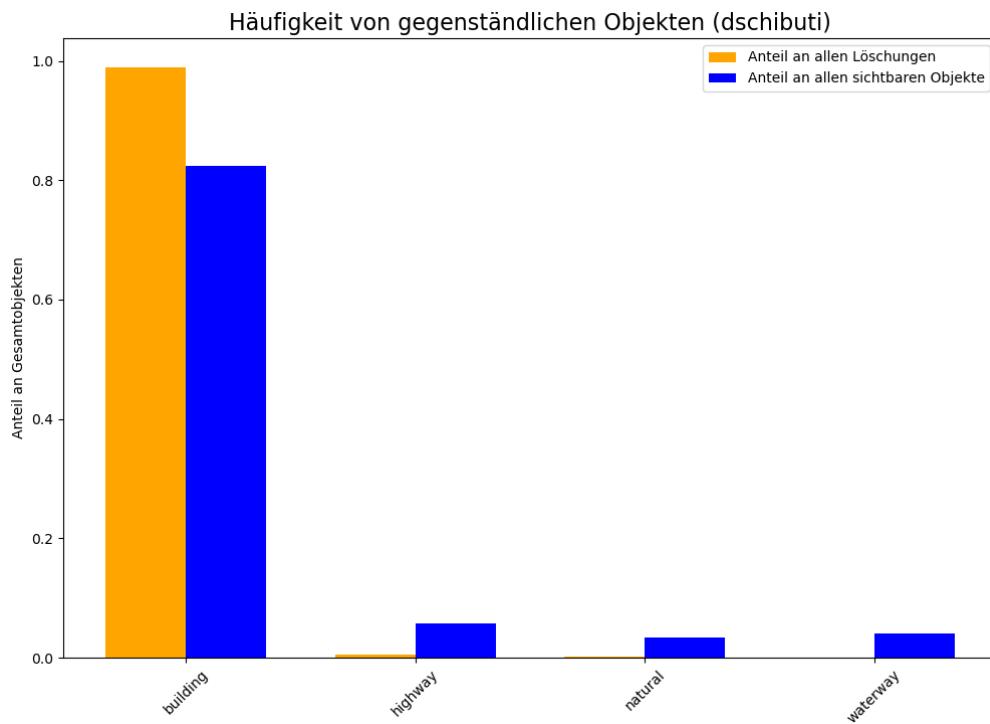


Abbildung 17: Häufigkeit gegenständlicher Objekte in Dschibuti

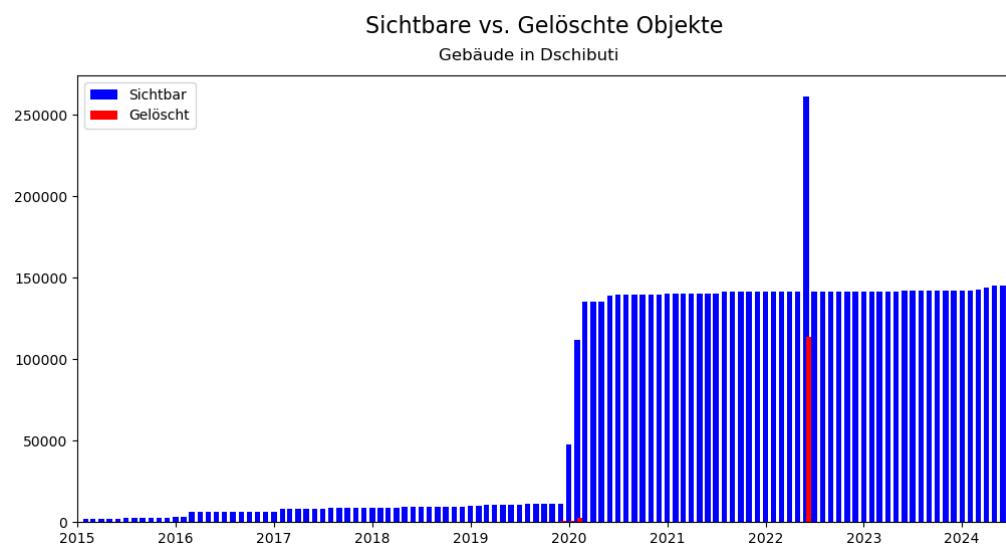


Abbildung 18: Zeitlicher OSM-Verlauf von gelöschten und sichtbaren Gebäuden in Dschibuti

4.3.2 Niederlande

In den Niederlanden wird über die gesamte Landesfläche hinweg mehr gelöscht als in den europäischen Nachbarländern. Auffällig ist dabei, wie markant die Werte an der Landesgrenze abfallen. (Abbildung 19)

Ursache dafür ist, dass Gebäude im Verhältnis zu ihrem Auftreten auf der sichtbaren Karte mit einem Wert von etwa 33% deutlich häufiger gelöscht werden als üblich.

Das alleine hat deshalb so starke Auswirkungen auf die Gesamtzahl an Löschungen, weil Gebäude den mit Abstand höchsten Anteil daran haben. So werden Grenzen und Bahninfrastrukturen verhältnismäßig häufiger gelöscht, machen gemeinsam aber weniger als 5% aller Löschungen aus. (Abbildung 20)

Die meisten dieser Gebäude wurden 2014 gelöscht. Gleichzeitig hat sich die Anzahl der sichtbaren Gebäude deutlich erhöht. (Abbildung 21)

Die Löschungen hatten den folgenden Kommentar beigefügt: "Import BAG"²⁰

BAG steht dabei für "Basic Registration Addresses and Buildings" und ist ein Datensatz der niederländischen Regierung zu Gebäuden und Adressen.²¹ Vermutlich gab es also im Jahr 2014 ein Projekt der niederländischen OSM-Community, bei dem Gebäude des BIG-Datensatzes in OSM importiert und dabei einige veraltete Gebäude gelöscht wurden.

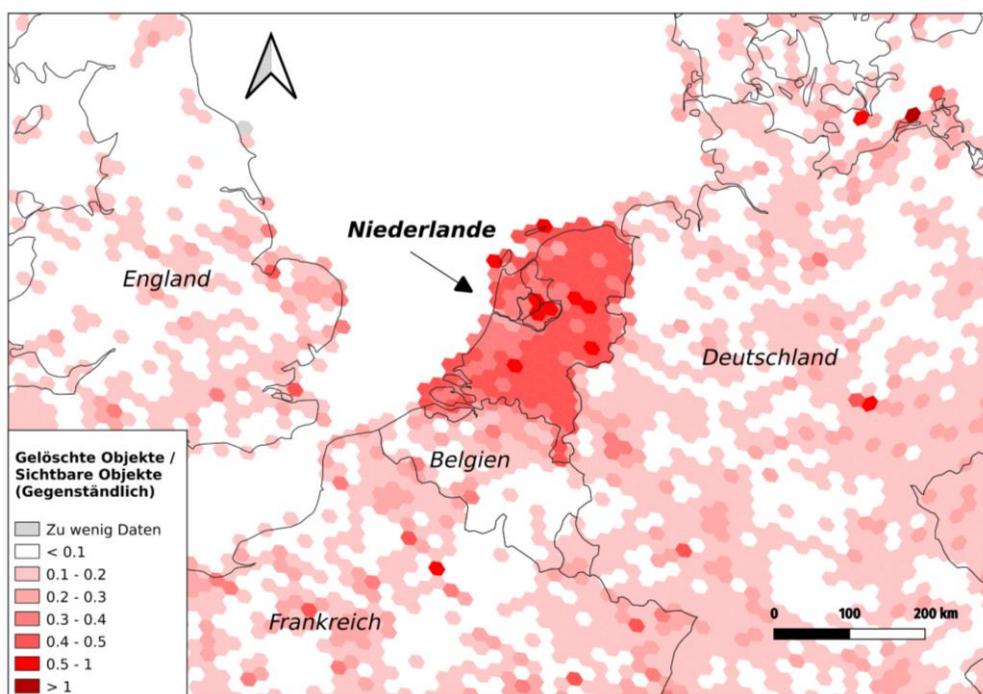


Abbildung 19: Löschungen in der Region rund um die Niederlande

²⁰ u.a. <https://www.osm.org/changeset/21375910>

²¹ <https://data.europa.eu/data/datasets/basisregistratie-adressen-en-gebouwen-bag-?locale=en>

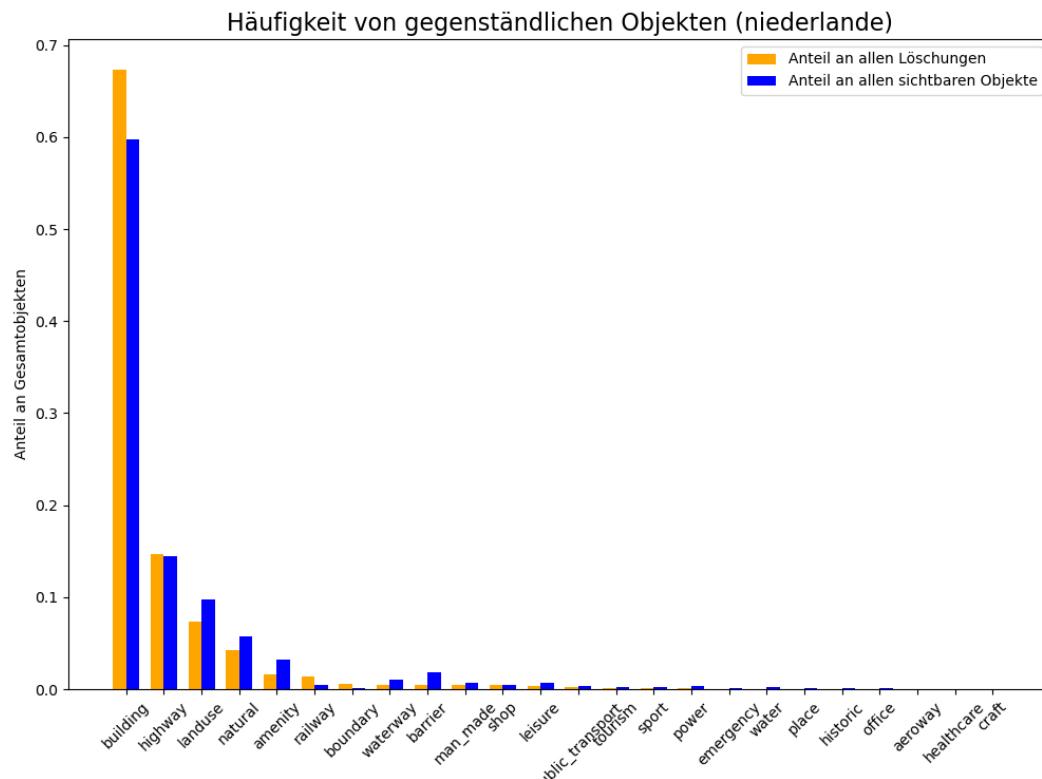


Abbildung 20: Häufigkeit gegenständlicher Objekte in den Niederlanden

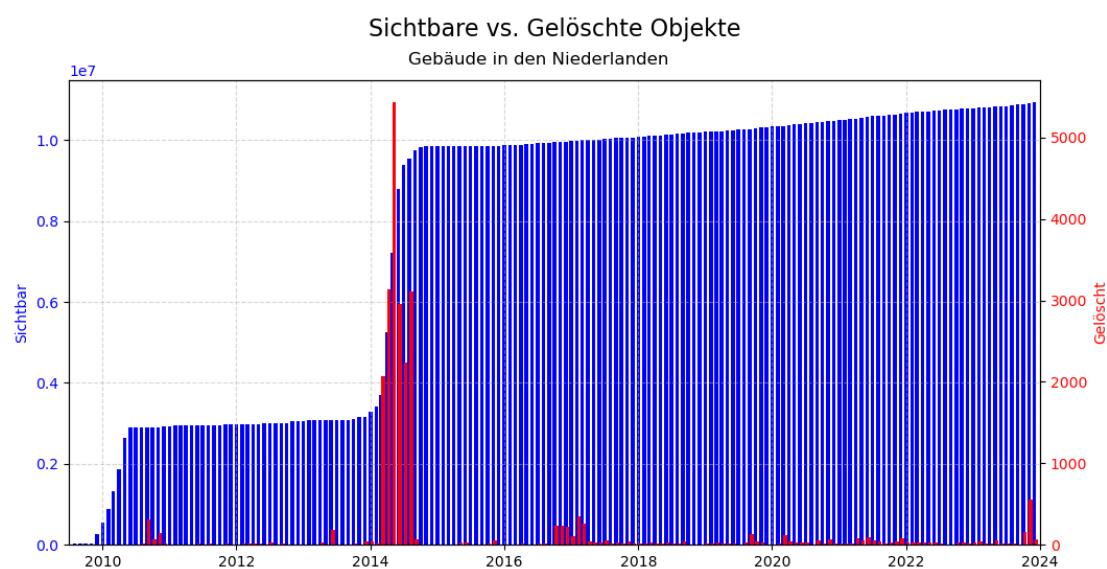


Abbildung 21: Zeitlicher OSM-Verlauf von gelöschten und sichtbaren Gebäuden in den Niederlanden

4.3.3 Finnland und Lettland

Auffällig viele Löschungen gibt es in Lettland und im Zentrum Finnlands (Abbildung 22). Grund für die vielen Löschungen in Lettland sind die verhältnismäßig vielen Löschungen von Grenzen (Abbildung 23). Im Zentrum Finnlands dagegen gibt es auffällig viele Löschungen von Sportinfrastruktur und Landnutzungsflächen (Abbildung 24).

Die Löschungen der lettischen Grenzen und der finnischen Sporteinrichtungen erfolgten kurz nachdem sich die Anzahl der sichtbaren Objekte erhöht hatte (Abbildung 25 und Abbildung 26). Sie waren damit mutmaßlich eine Reaktion auf einen unerwünschten Datenimport. Bestätigt wird das durch die beigefügten Kommentare.

So lautet der Kommentar für den Großteil gelöschter lettischen Grenzen folgendermaßen:

- “Delete buggy and undisussed import (boundary tags on nodes etc). [...]”²²

Die Kommentare für den Großteil der gelöschten finnischen Sportinfrastruktur lauten folgendermaßen:

- “Snowmobile: Oversized Removal”²³
- “Snowmobile: Fixing”²⁴
- “Snowmobile: Elimination of overlaps”²⁵

Die Löschungen der finnischen Landnutzungsflächen erfolgten innerhalb weniger Tage im Winter 2021. Sie waren ebenfalls die Reaktion auf einen unerwünschten Import, weil sich die Kommentare auf Rückabwicklungen (*reverting*) zu einem vorherigen Zustand beziehen:

- “Reverting 72911944 - ways”²⁶
- “Reverting 73004319 - ways”²⁷
- ...

Der Import passierte allerdings etwa ein Jahr vor den Löschungen. Damit waren vermutlich falsche Daten für etwa ein Jahr auf der OSM-Karte sichtbar, ehe die OSM-Community beschloss, sie zu entfernen.

²² u.a. <https://www.osm.org/changeset/68316594>

²³ u.a. <https://www.osm.org/changeset/32938436>

²⁴ u.a. <https://www.osm.org/changeset/32836012>

²⁵ u.a. <https://www.osm.org/changeset/31269538>

²⁶ u.a. <https://www.osm.org/changeset/111808786>

²⁷ u.a. <https://www.osm.org/changeset/111810935>

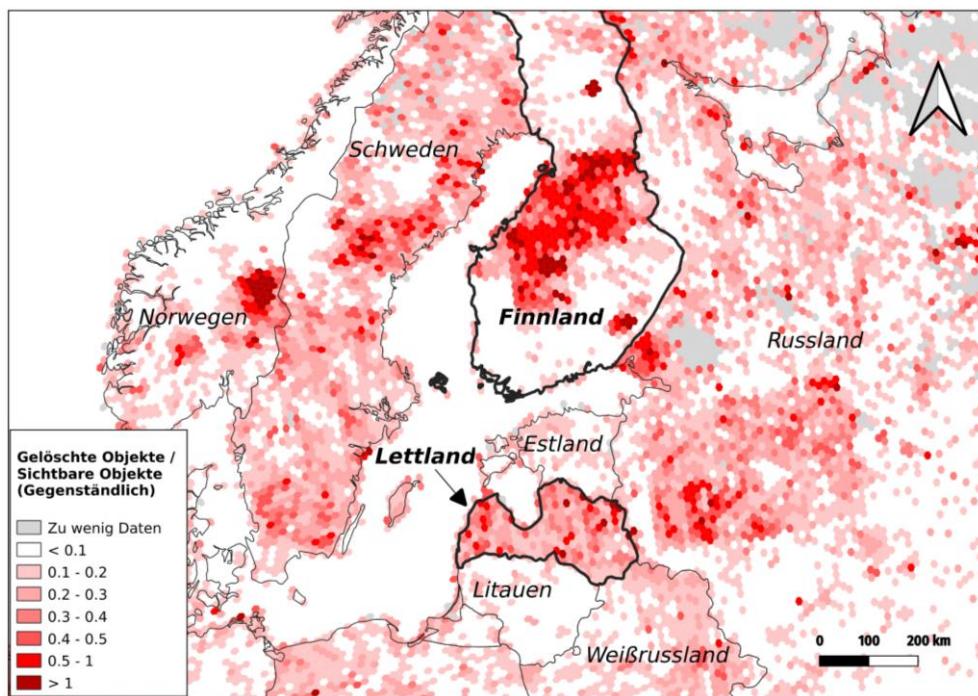


Abbildung 22: Löschungen in der Region rund um Finnland und Lettland

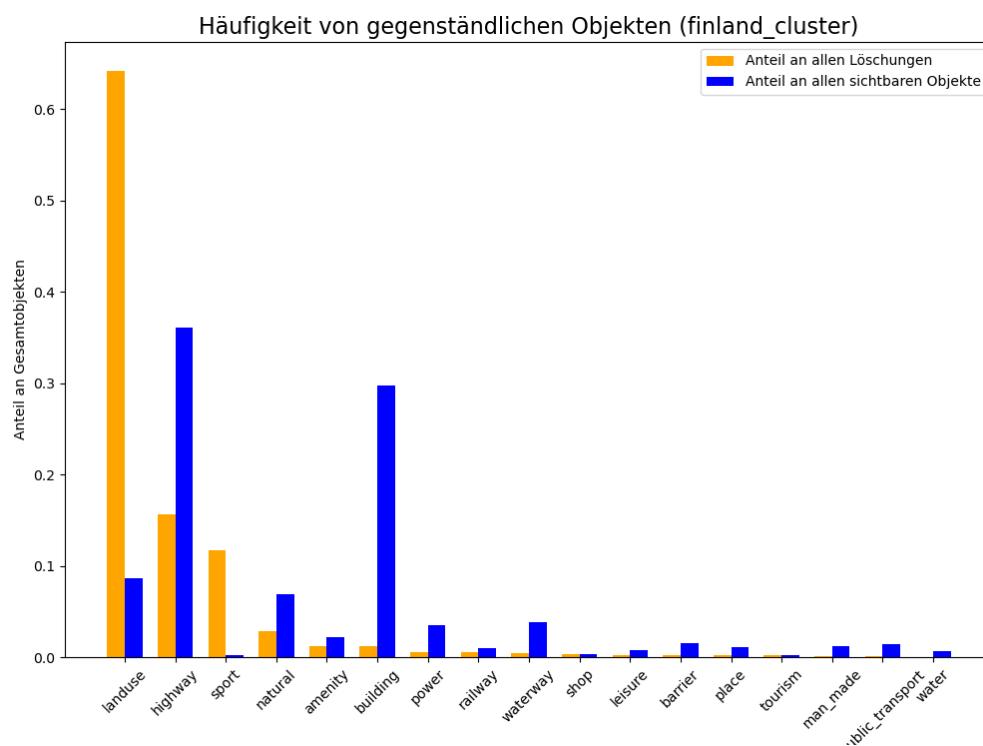


Abbildung 23: Häufigkeit gegenständlicher Objekte in finnischen Regionen mit vielen Löschungen

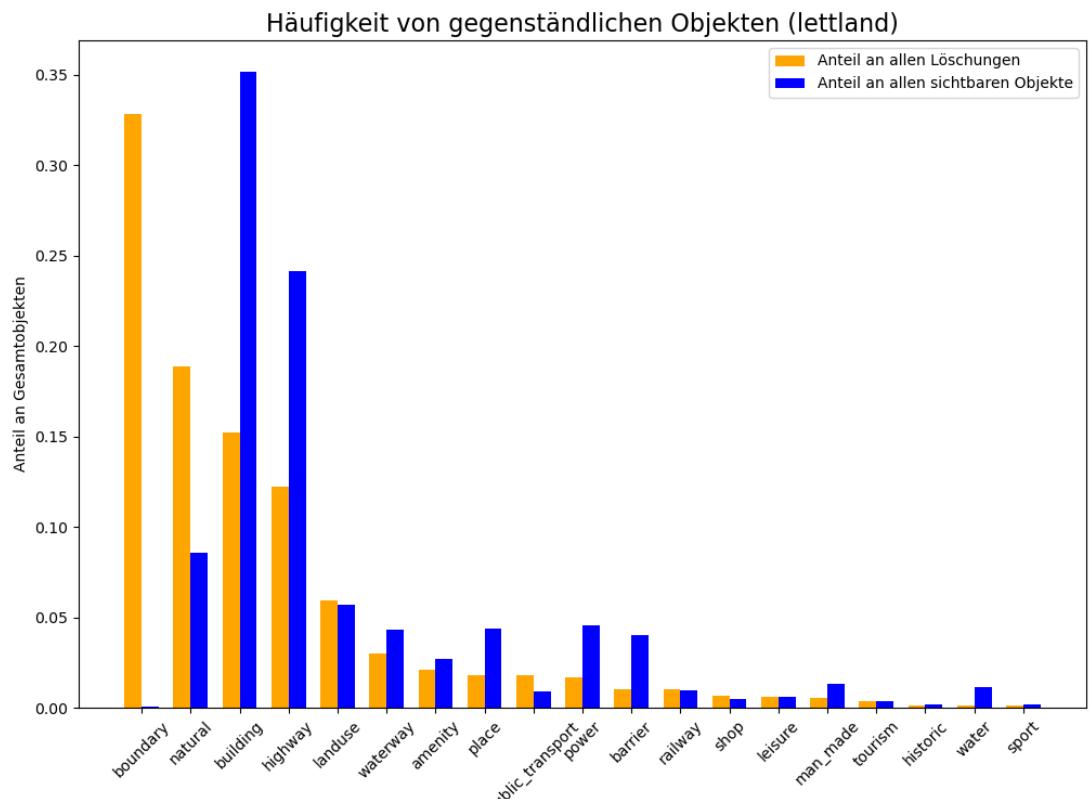


Abbildung 24: Häufigkeit gegenständlicher Objekte in Lettland

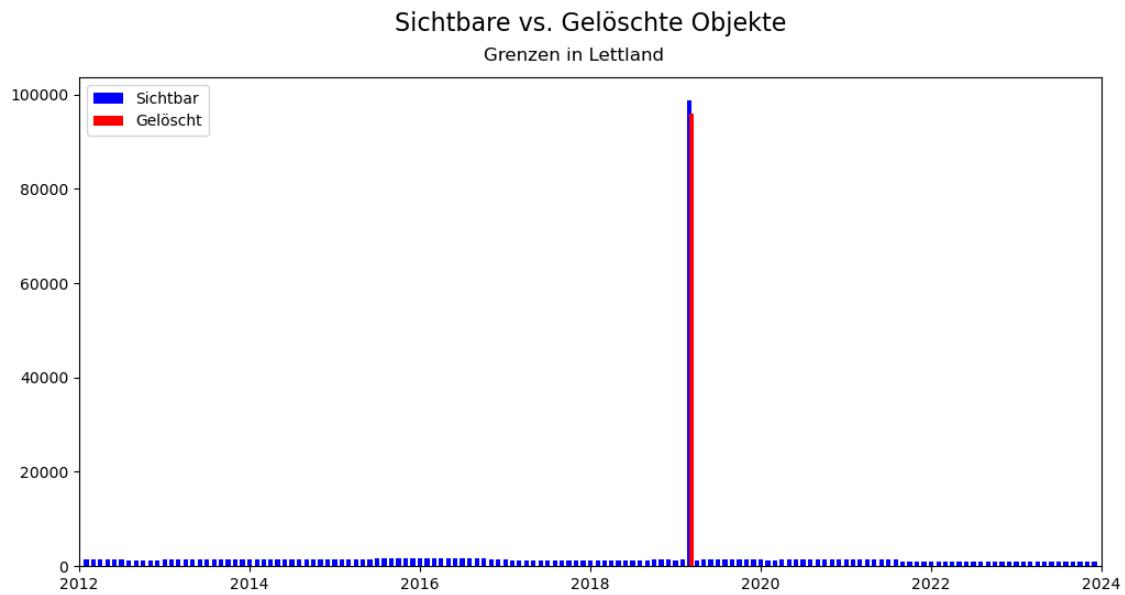


Abbildung 25: Zeitlicher OSM-Verlauf von gelöschten und sichtbaren Grenzen in Lettland

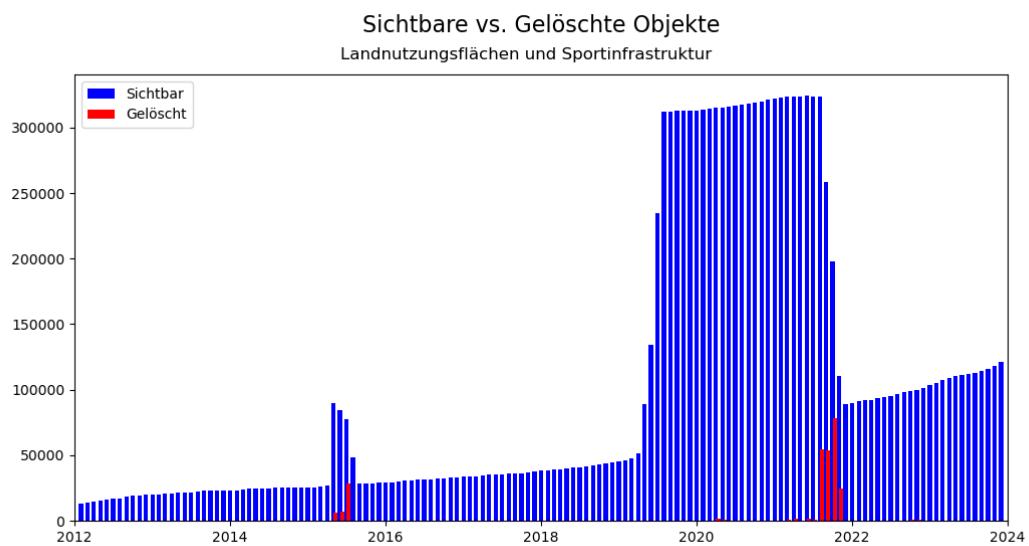


Abbildung 26: Zeitlicher OSM-Verlauf von gelöschten und sichtbaren Landnutzungsflächen und sportbezogener Infrastruktur in finnischer Region mit vielen Löschungen

4.3.4 Südafrika

In Südafrika gibt es einige räumliche Cluster mit sehr vielen Löschungen. Dabei übersteigt die Anzahl an Löschungen zumeist die Anzahl an sichtbaren Objekten. Außerhalb der Cluster gibt es deutlich weniger Löschungen. (Abbildung 27)

Unter den Löschungen sind fast ausschließlich Gebäude (Abbildung 28). Sie fanden innerhalb eines kurzen Zeitraumes im Jahr 2023 statt, kurz nachdem sich die Anzahl der sichtbaren Gebäude stark erhöht hat (Abbildung 29).

Fast alle Löschungen hatten folgenden Kommentar beigefügt:

- “remove undiscussed and buggy building import in South Africa”²⁸

Ursache für die vielen Löschungen in Südafrika ist also, dass ein unerwünschter Import rückgängig gemacht wurde.

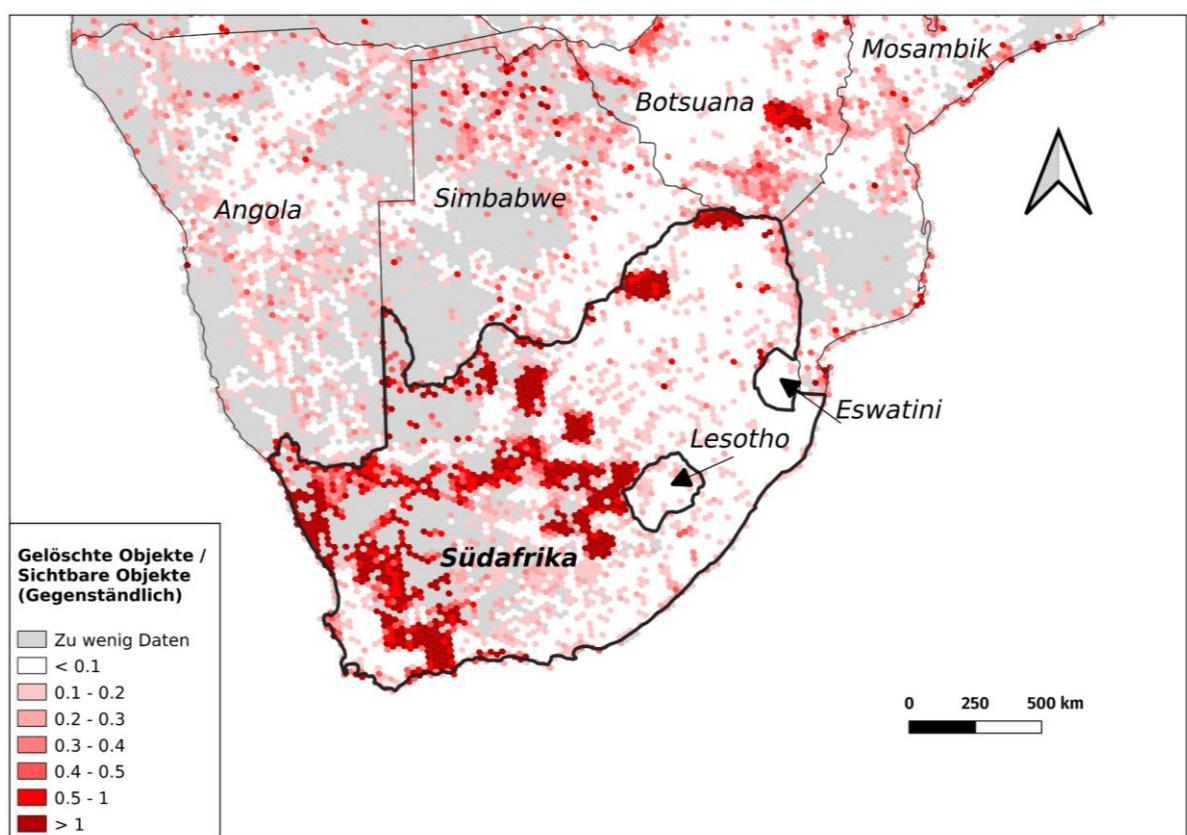


Abbildung 27: Löschungen in der Region rund um Südafrika

²⁸ u.a. <https://osm.org/changeset/140178652>

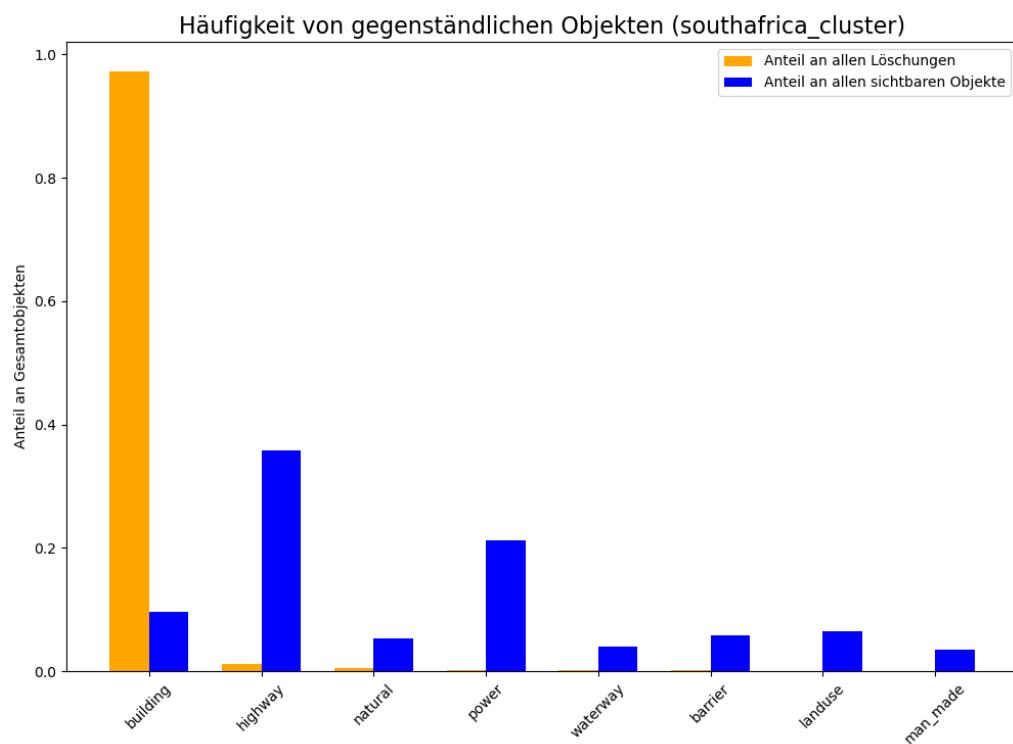


Abbildung 28: Häufigkeit gegenständlicher Objekte in südafrikanischen Regionen mit besonders vielen Löschungen

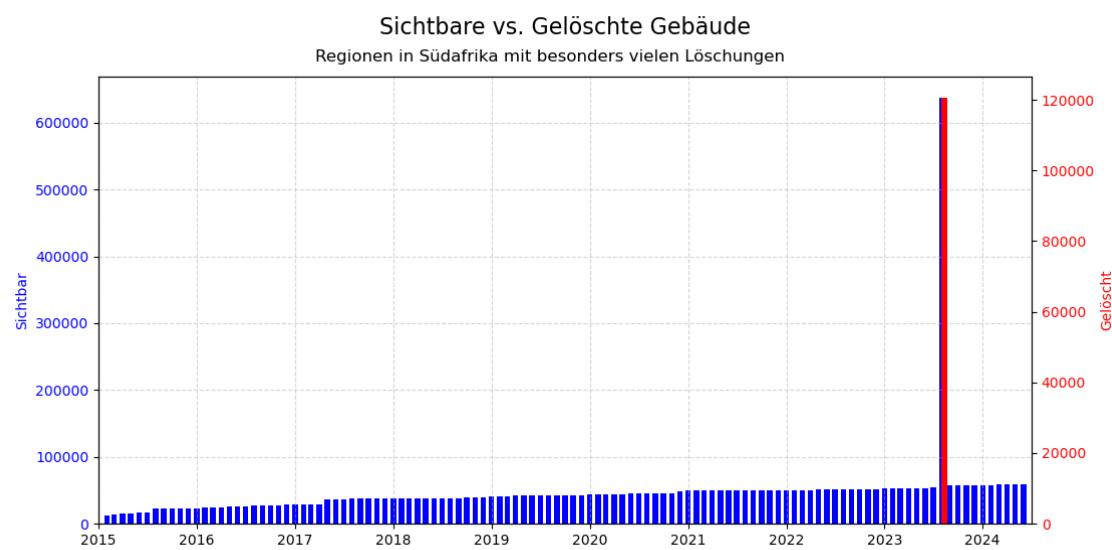


Abbildung 29: Zeitlicher OSM-Verlauf von gelöschten und sichtbaren Gebäuden in südafrikanischen Regionen mit besonders vielen Löschungen

4.3.5 Australien

Im östlichen Australien gibt es eine Linie mit besonders vielen Löschungen, die sich sichtbar vom Rest des Kontinents abhebt (Abbildung 30). Diese Linie verläuft entlang von mehreren zusammenhängenden Highways (Abbildung 31). In diesen Bereichen werden Highways auch mit Abstand am meisten gelöscht (Abbildung 32). Die meisten Löschungen passierten dabei innerhalb eines kurzen Zeitraumes Ende 2007 und Mitte 2012 (Abbildung 33).

Die Löschungen von 2007 geschahen kurz nachdem besonders viele Objekte der sichtbaren Karte hinzugefügt wurden. Vermutlich wurde nach einem umfangreichen Import ein Teil davon wieder gelöscht, weil die Datenqualität nicht ausreichte. Belegbar ist die Vermutung nicht, weil keine Kommentare zu den Löschungen hinterlegt wurden.

Die Löschungen von 2012 haben größtenteils den folgenden Kommentar hinterlegt:

- “*Updates based on the redaction process*”²⁹

Im Jahr 2012 wurde die Lizenz der OSM-Daten automatisiert umgeändert auf die Open Database License (ODbl), um sie sicherer und kompatibler für Datenbanken zu machen.³⁰ Bei der Umwandlung gingen etwa 1% aller OSM-Daten verloren.³¹ Damit ist das vermutlich der Grund für die Löschungen in diesem Zeitraum.

²⁹ u.a. <https://osm.org/changeset/12317984>

³⁰

https://osmfoundation.org/wiki/Licence/Historic/We_Are_Changing_The_License#Why_are_we_changing_the_license?

³¹ https://wiki.openstreetmap.org/wiki/Open_Database_License

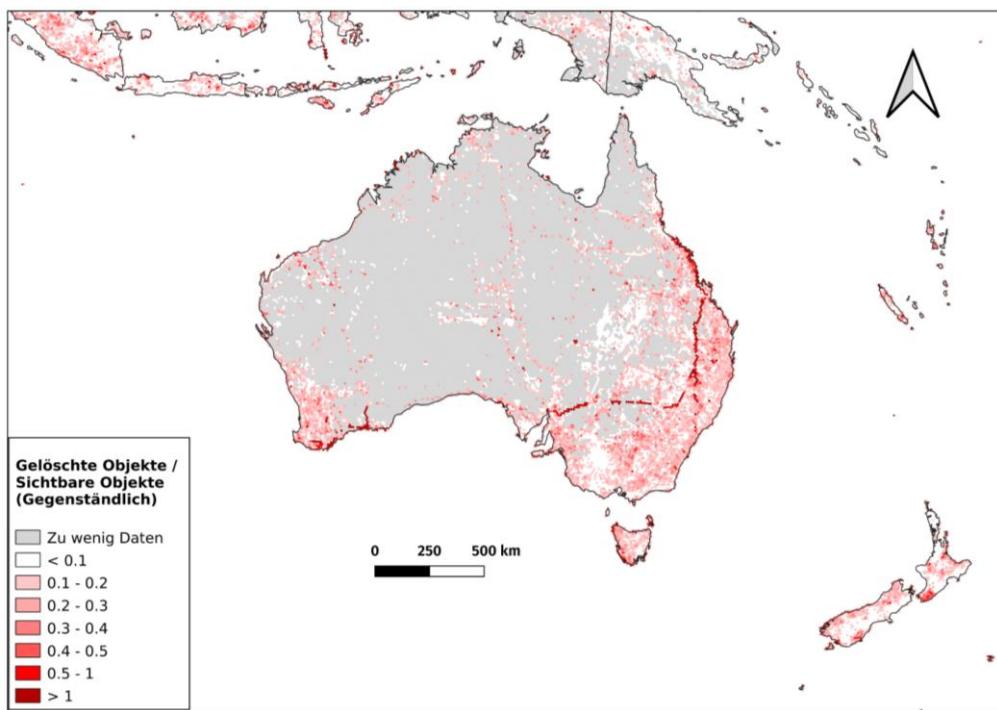


Abbildung 30: Löschungen in der Region rund um Australien

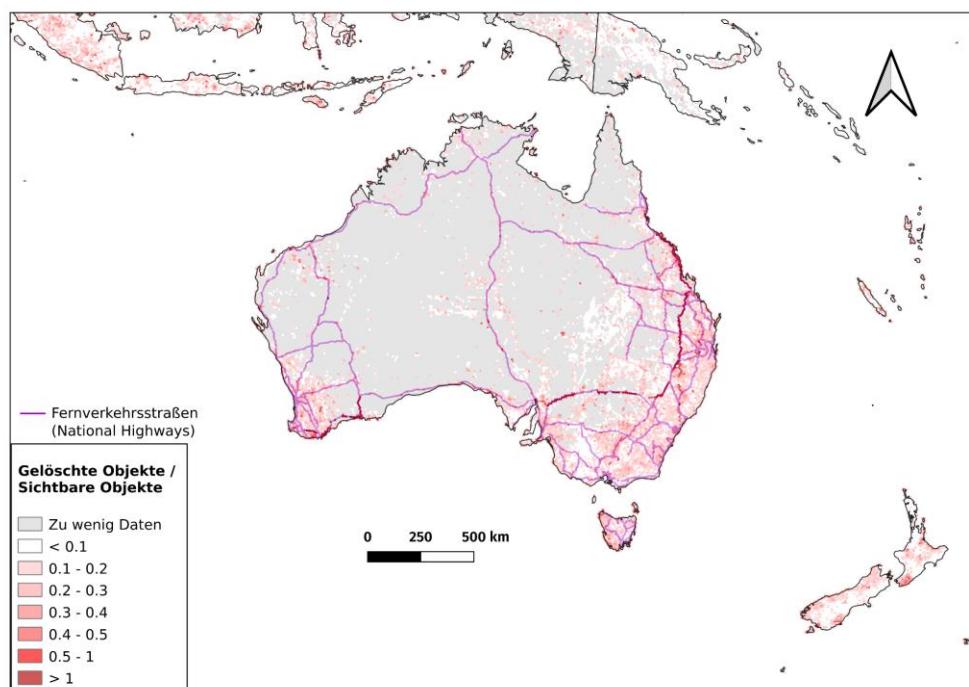


Abbildung 31: Löschungen in der Region rund um Australien mit Highways

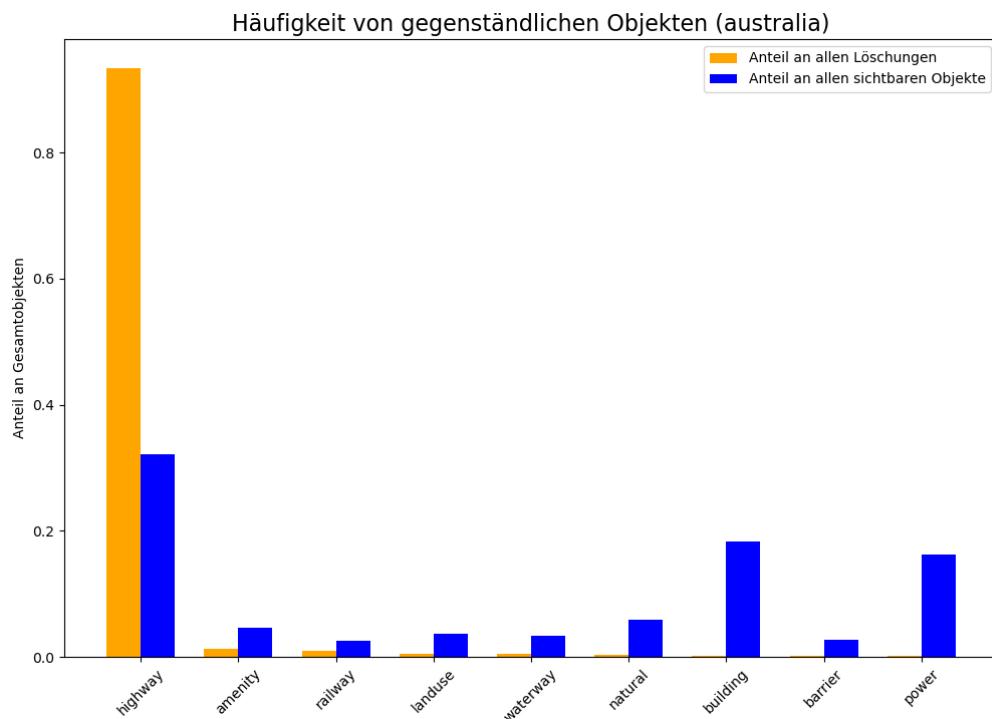


Abbildung 32: Häufigkeit gegenständlicher Objekte im Bereich der australischen Highways mit besonders vielen Löschungen

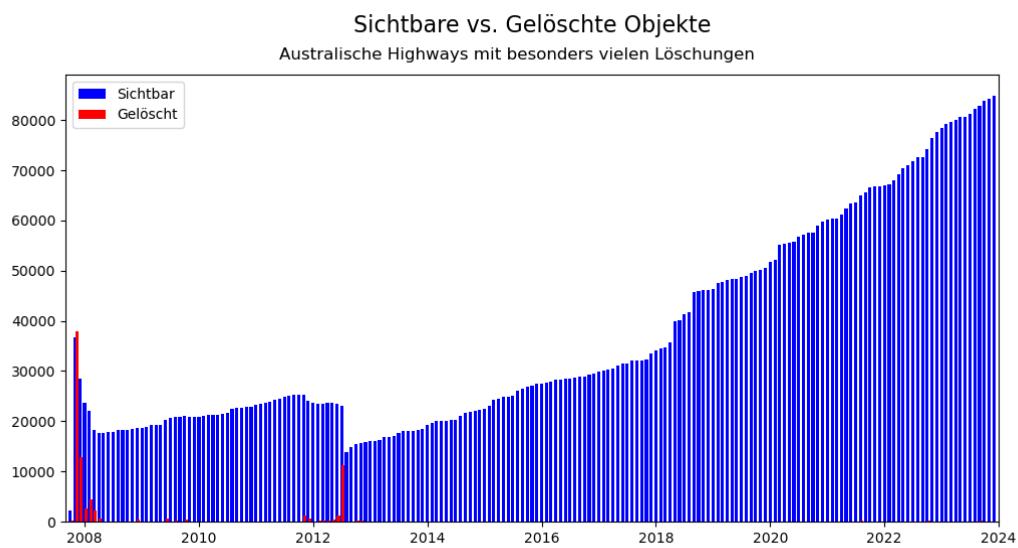


Abbildung 33: Zeitlicher OSM-Verlauf von gelöschten und sichtbaren Objekten im Bereich der australischen Highways mit besonders vielen Löschungen

4.4 Konfliktregionen

Bei der Untersuchung der Konfliktregionen liegt der Fokus auf den Löschungen gegenständlicher Objekte. So soll die anfangs gestellte These untersucht werden, ob in Konfliktregionen real existierende Objekte überproportional häufig gelöscht werden, z.B. weil sie umstritten sind oder aus strategischen Gründen verschleiert werden sollen.

4.4.1 Westjordanland

Für das Westjordanland wird direkt sichtbar, dass besonders viel gelöscht wurde. Im Süden und Norden der Region übersteigen die Werte größtenteils die 100%-Marke. Das heißt, dass dort mehr Objekte gelöscht wurden als derzeit auf der OSM-Karte existieren. Im Umland fallen die Werte, vor allem auf jordanischem Territorium, deutlich ab. (Abbildung 34)

Das Verhältnis der Anzahl gelöschter Objekte mit der Anzahl existierender Objekte liegt für alle Hexagone im Durchschnitt bei etwa 147,1%. Das ist ein deutlich höherer Wert als im Umland (8,6%) und weltweit (15%). Damit ist die Aussage, dass im Westjordanland nicht zufällig mehr gelöscht wird als im Umland und im globalen Mittel, statistisch signifikant. (Abbildung 35 und Abbildung 36)

Die Löschungen im Westjordanland betreffen ausschließlich gegenständliche Objekte (Abbildung 37). Davon entfallen die Allermeisten auf Gebäude (*building*) mit einem Anteil von rund 90% (Abbildung 38). Diese werden auch verhältnismäßig am häufigsten gelöscht (Abbildung 39). Auch Ortsbezeichnungen (*place*) und Barrieren (*barrier*) werden verhältnismäßig häufig gelöscht. Auffällig ist, wie selten Straßen (*highway*) gelöscht werden, trotz ihres hohen Anteils an sichtbaren Objekten.

Die gelöschten Gebäude wurden größtenteils im Frühjahr 2023 entfernt, nachdem die Anzahl sichtbarer Gebäude stark angestiegen war (Abbildung 40). Der dazugehörige Kommentar hieß:

- “revert buggy and undiscussed import of buildings in Israel”³²

Damit waren die vielen Löschungen im Westjordanland das Ergebnis einer umfangreichen Rückabwicklung eines unerwünschten Imports.

³² u.a. <https://osm.org/changeset/132641930>

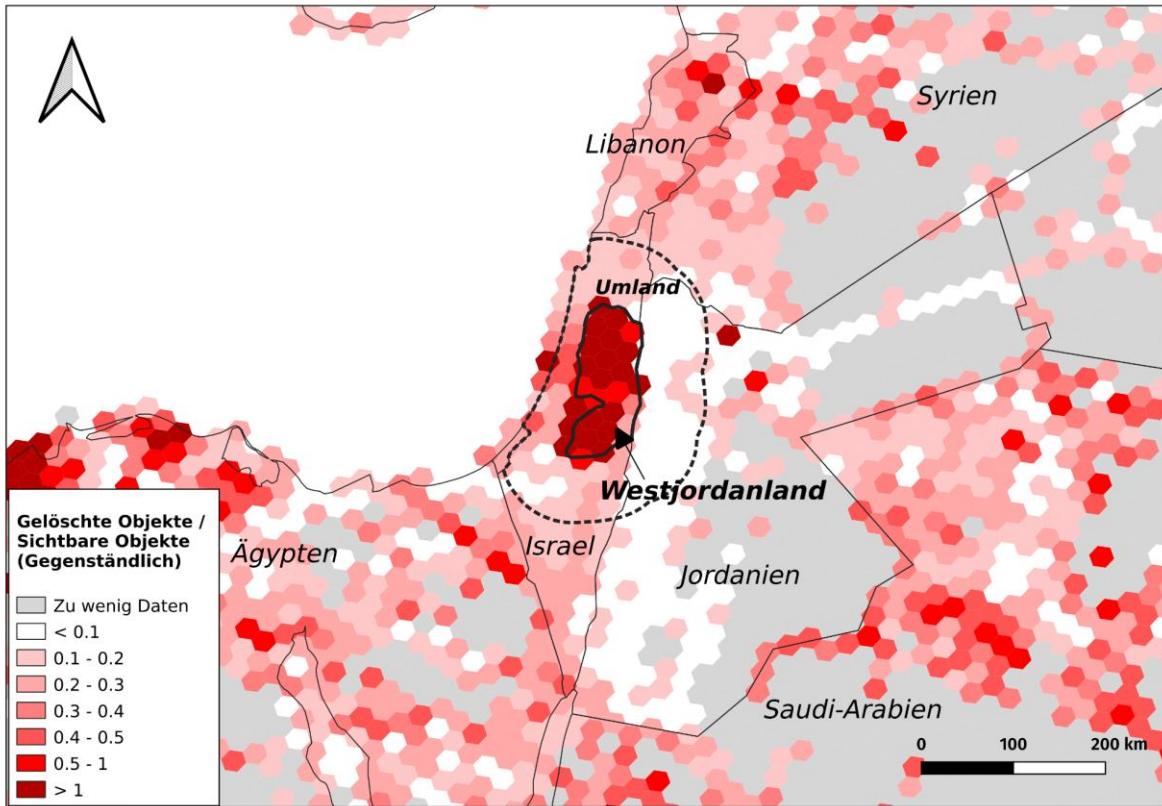


Abbildung 34: Löschungen in der Region rund um das Westjordanland

Westjordanland vs. Umland

Mittelwert Westjordanland ≈ 1.471 , Mittelwert Umland ≈ 0.086
 Differenz der beobachteten Mittelwerte ≈ 1.385

Permutationstest

Nullhypothese: Beobachteter Mittelwert für Westjordanland nur zufällig größer
 Alternativhypothese: Beobachteter Mittelwert für Westjordanland nicht zufällig größer

p-Wert $< 1e-03 < 0.05 \rightarrow$ Verwerfung der Nullhypothese

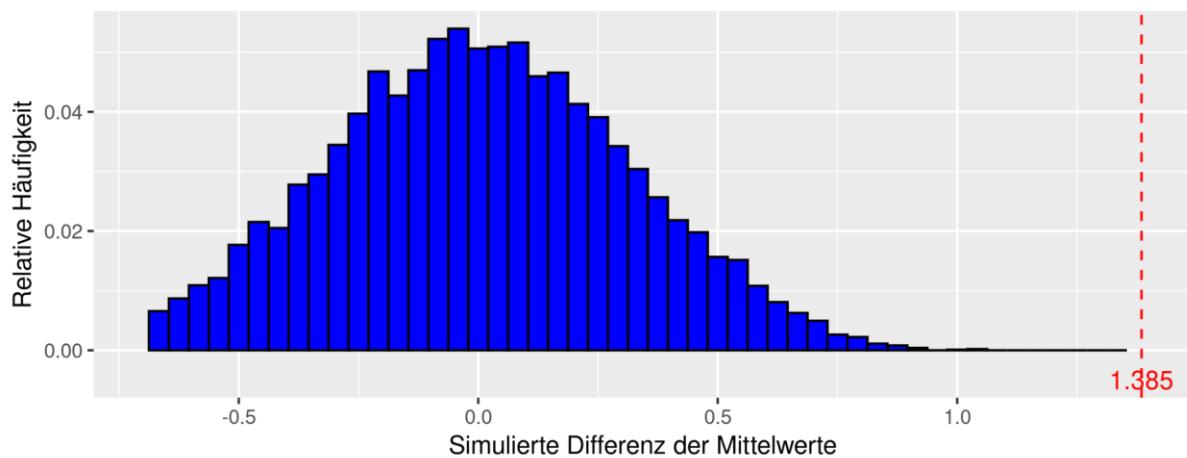


Abbildung 35: Permutationstest zur Überprüfung, ob sich die Menge an Löschungen im Westjordanland signifikant von dessen Umland unterscheidet

Westjordanland vs. Global

Mittelwert Westjordanland ≈ 1.471 , Mittelwert Global ≈ 0.15

Differenz der beobachteten Mittelwerte ≈ 1.321

Permutationstest

Nullhypothese: Beobachteter Mittelwert für Westjordanland nur zufällig größer

Alternativhypothese: Beobachteter Mittelwert für Westjordanland nicht zufällig größer

p-Wert $\approx 0.001 < 0.05 \rightarrow$ Verwerfung der Nullhypothese

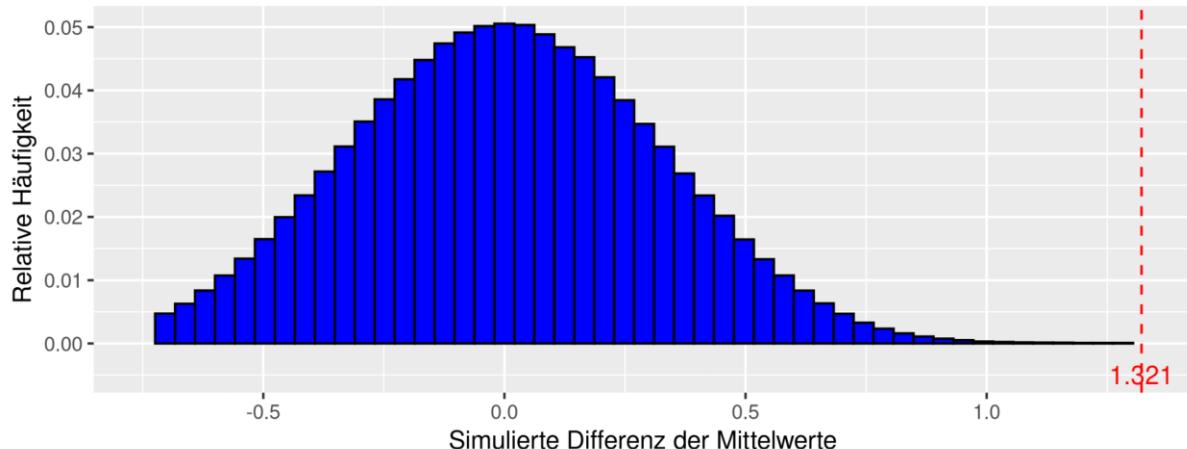


Abbildung 36: Permutationstest zur Überprüfung, ob sich die Anzahl an Löschungen im Westjordanland signifikant von dem globalen Raum unterscheidet

Häufigkeit von OSM-Objekten (Westjordanland)

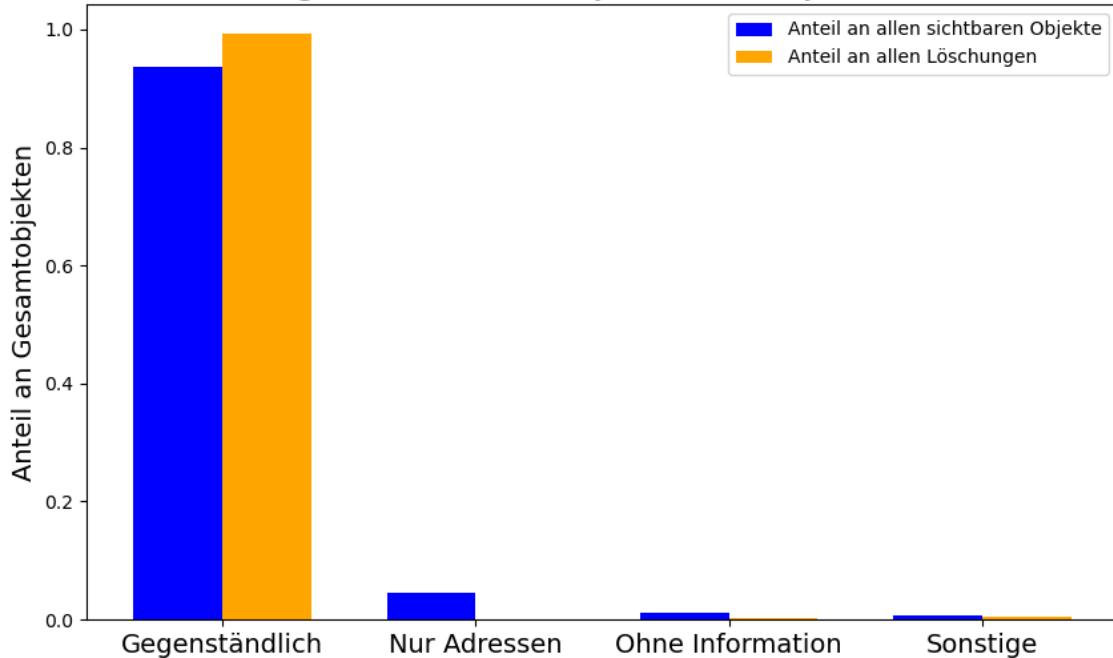


Abbildung 37: Häufigkeit der OSM-Objektgruppen im Westjordanland

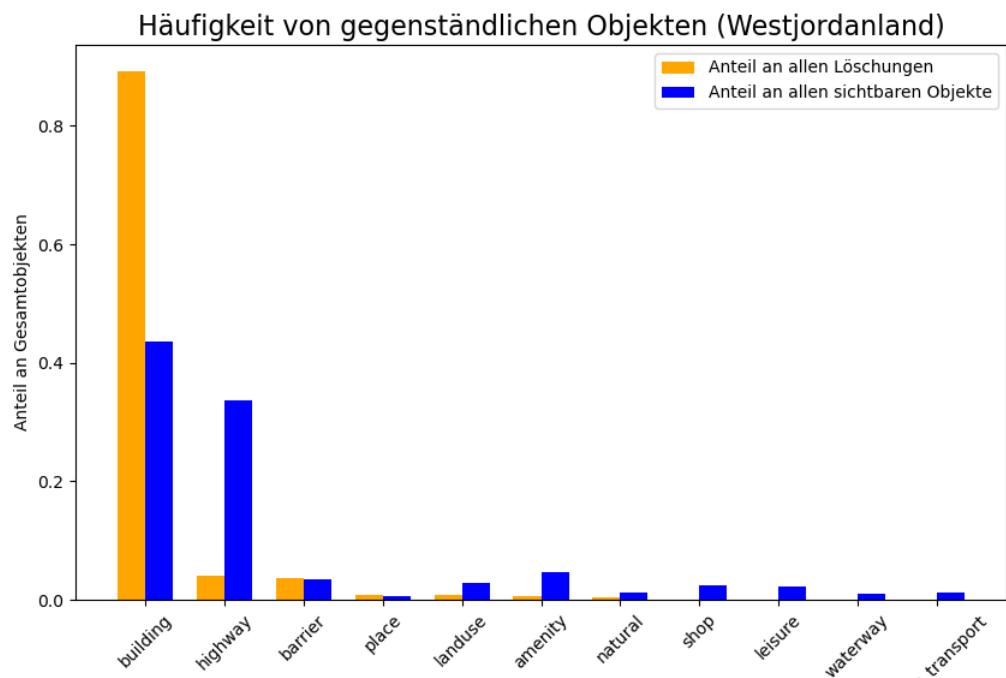


Abbildung 38: Häufigkeit gegenständlicher Objekte im Westjordanland

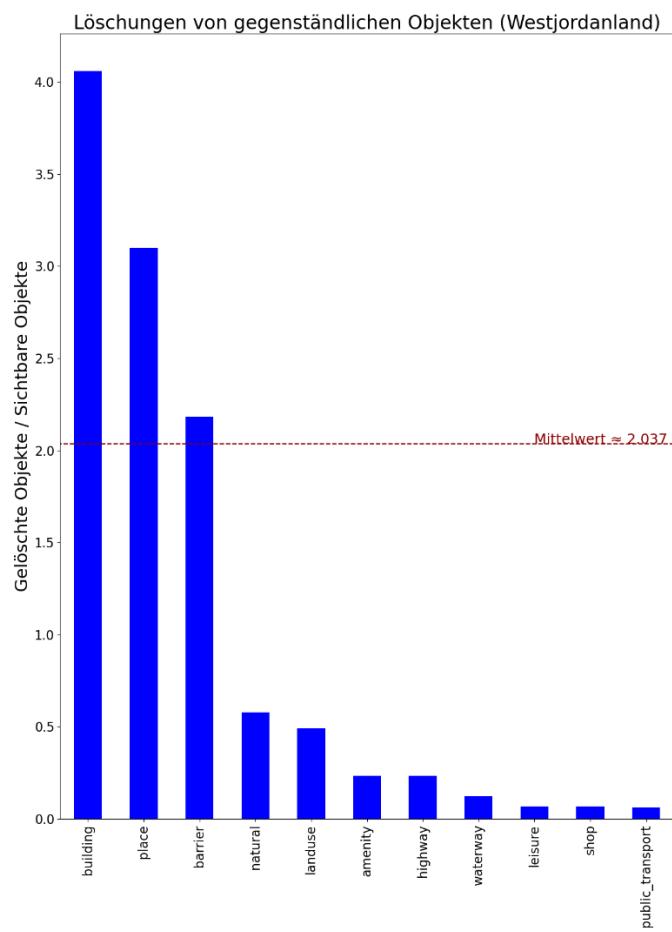


Abbildung 39: Verhältnis zwischen gelöschten und sichtbaren gegenständlichen Objekten im Westjordanland

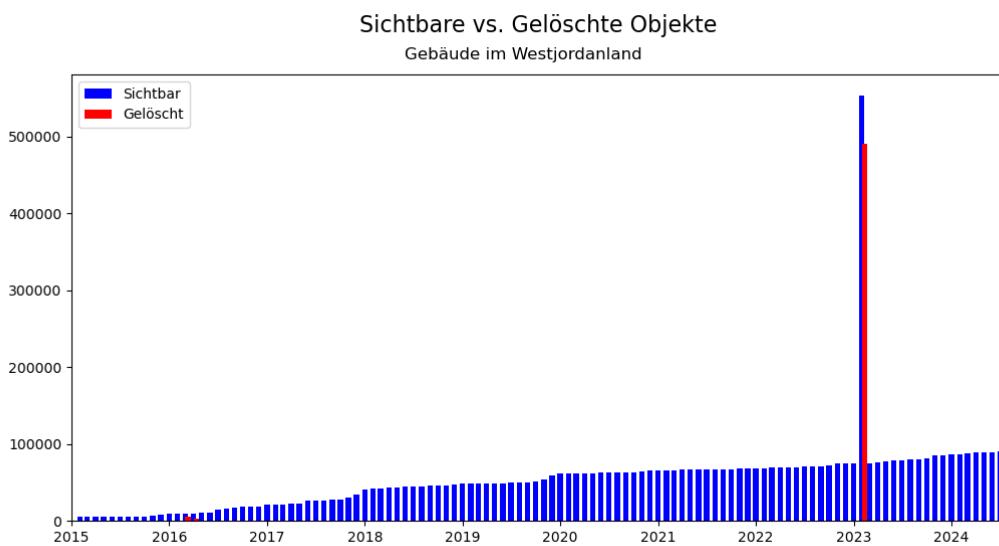


Abbildung 40: Zeitlicher OSM-Verlauf von gelöschten und sichtbaren Gebäuden im Westjordanland

4.4.2 Kaschmir

In der Kaschmir-Region gibt es vereinzelte Gebiete im zentralen und südlichen Teil mit vielen Löschungen. Im angrenzenden pakistanschen Umland fallen die Werte deutlich ab. Im indischen, afghanischen und tadschikistanischen Umland dagegen bleiben die Werte auf ähnlichem Niveau. Im äußersten Nordosten und Nordwesten von Kaschmir und dem chinesischen Umland gibt es dagegen zu wenig Daten, um verlässliche Aussagen zu treffen. (Abbildung 41)

Das Verhältnis der Anzahl gelöschter Objekte mit der Anzahl existierender Objekte liegt für alle Hexagone bei durchschnittlich rund 11,8%. Damit wird etwas mehr gelöscht als im Umland (11,2%). Der Unterschied von 0,2% ist allerdings so klein, dass er auch aus Zufall hätte entstehen können (Abbildung 42). Im Vergleich zu allen Regionen weltweit (15%) wird weniger gelöscht (-3,2%). Auch dieser Unterschied ist nicht groß genug, um einen statistisch signifikanten Unterschied nachzuweisen und damit zu behaupten, dass in Kaschmir tatsächlich weniger gelöscht wird als global (Abbildung 43).

Unter den Löschungen machen gegenständliche Objekte mit etwa der Hälfte aller Objekte die größte Gruppe aus. Allerdings ist deren Anteil geringer als weltweit. Objekte, die nicht nach dem informellen Standard der OSM-Community beschrieben wurden, werden dagegen häufiger gelöscht als weltweit. (Abbildung 44)

Unter den gelöschten gegenständlichen Objekten kommen Straßen (35%), Repräsentationen von Orten (23%), Gebäude (16%) und nutzbare Einrichtungen (10%) am häufigsten vor. (Abbildung 45)

Am verhältnismäßig häufigsten gelöscht werden Objekte zur Repräsentation von Orten (*place*). Die Anzahl an Löschungen entspricht dabei rund 53% der Anzahl von Objekten auf der sichtbaren Karte. Am zweithäufigsten gelöscht werden nutzbare Einrichtungen (*amenity*) mit einem Wert von knapp 40%. Auch touristisch relevante Orte (*tourism*) werden mit einem Wert von knapp 15% noch überproportional häufig gelöscht. Am verhältnismäßig seltensten gelöscht werden mit deutlichem Abstand Elektrizitäts-Einrichtungen (*power*). (Abbildung 46)

Die meisten Löschungen passierten im Sommer 2012 und Herbst 2013 (Abbildung 47). Die Löschungen aus dem Sommer 2012 haben überwiegend folgenden Kommentar beigefügt:

- “*Updates based on the redaction process*”³³

Für die Hintergründe zu diesem Kommentar siehe Kapitel 4.3.5.

Die Löschungen aus dem Herbst 2013 erfolgten kurz nachdem sich die Anzahl sichtbarer Objekte erhöht hatte. Sie haben überwiegend folgenden Kommentar beigefügt:

- “*remove un-discussed and faulty import of India place names*”³⁴

Ein Nutzer wollte demnach Ortsbezeichnungen in indischer Sprache ergänzen, auch auf pakistanisch-kontrolliertem Gebiet. Ob dieser Import politisch motiviert ist und mit dem territorialen Konflikt zwischen Indien und Pakistan zusammenhängt, kann damit allerdings nicht belegt werden. Weitere häufig auftretende Kommentare, die auf den Konflikt hindeuten, gibt es keine.

³³ u.a. <https://osm.org/changeset/12382331>

³⁴ u.a. <https://osm.org/changeset/17612894>

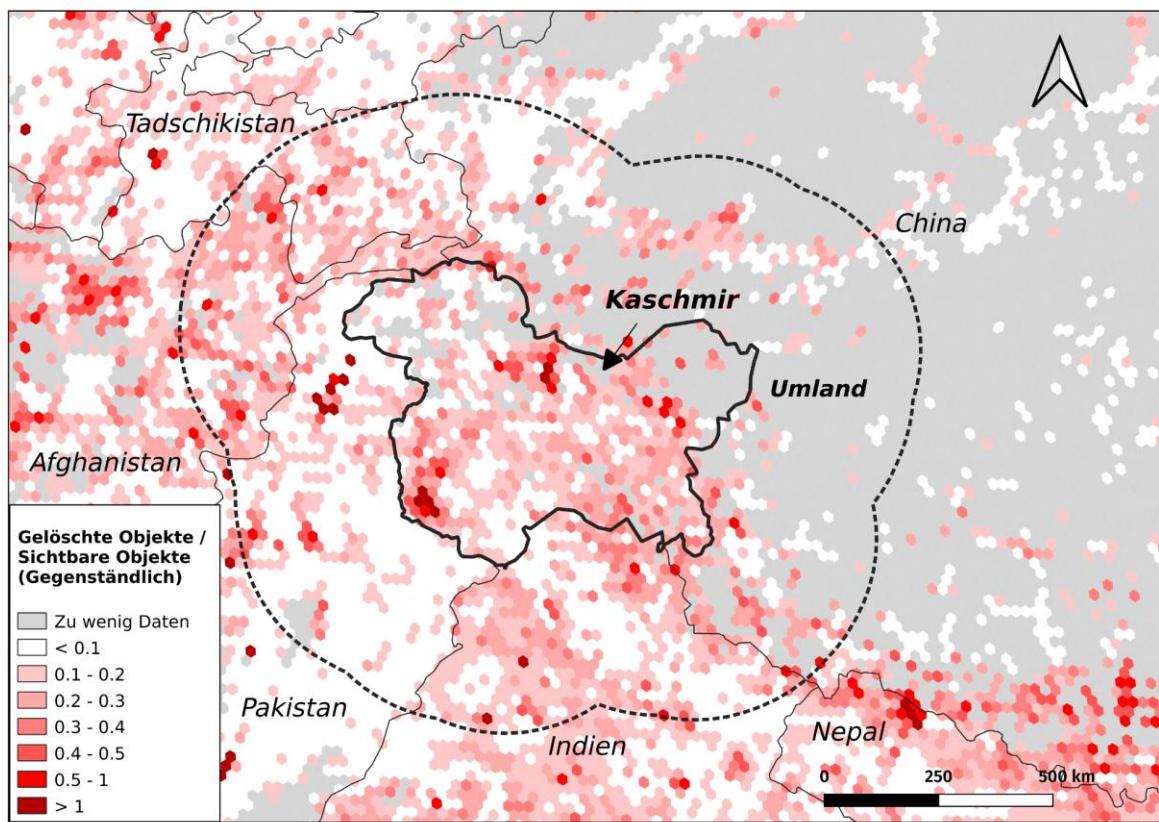


Abbildung 41: Löschungen in der Region rund um Kaschmir

Kaschmir vs. Umland

Mittelwert Kaschmir ≈ 0.118 , Mittelwert Umland ≈ 0.112
 Differenz der beobachteten Mittelwerte ≈ 0.005

Permutationstest

Nullhypothese: Beobachteter Mittelwert für Kaschmir nur zufällig größer

Alternativhypothese: Beobachteter Mittelwert für Kaschmir nicht zufällig größer

p-Wert $\approx 0.402 > 0.05 \rightarrow$ Beibehaltung der Nullhypothese

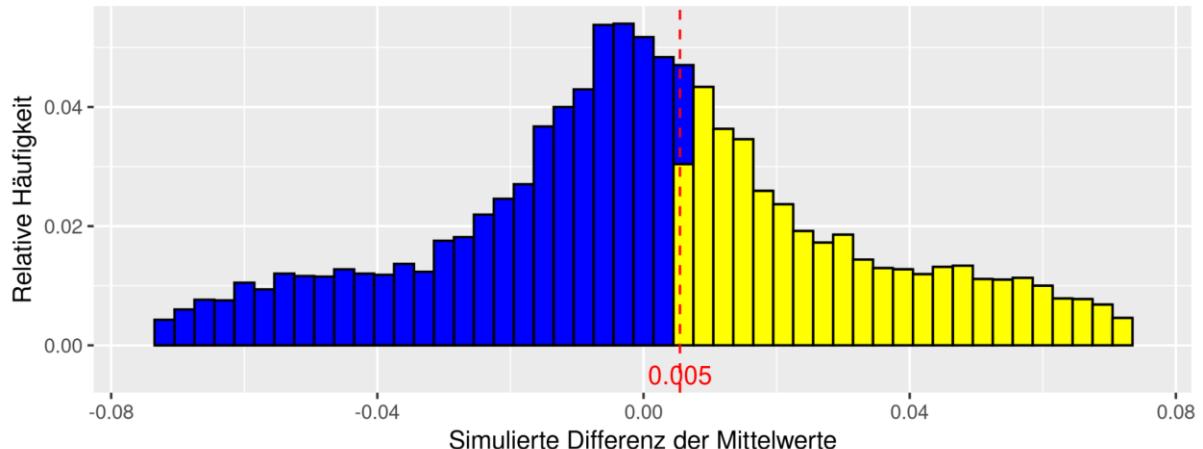


Abbildung 42: Permutationstest zur Überprüfung, ob sich die Menge an Löschungen in Kaschmir signifikant von dessen Umland unterscheidet; Der p-Wert entspricht dem Anteil der gelben Fläche an der Gesamtfläche der Balken

Kaschmir vs. Global

Mittelwert Kaschmir ≈ 0.118 , Mittelwert Global ≈ 0.15
 Differenz der beobachteten Mittelwerte ≈ -0.032

Permutationstest

Nullhypothese: Beobachteter Mittelwert für Kaschmir nur zufällig kleiner

Alternativhypothese: Beobachteter Mittelwert für Kaschmir nicht zufällig kleiner

p-Wert $\approx 0.094 > 0.05 \rightarrow$ Beibehaltung der Nullhypothese

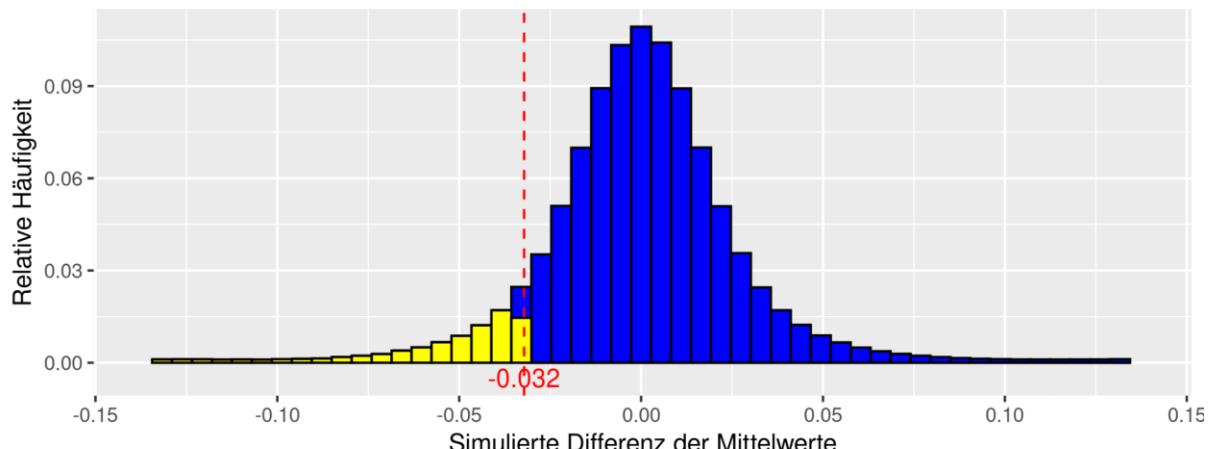


Abbildung 43: Permutationstest zur Überprüfung, ob sich die Anzahl an Löschungen in Kaschmir signifikant von dem globalen Raum unterscheidet; Der p-Wert entspricht dem Anteil der gelben Fläche an der Gesamtfläche der Balken

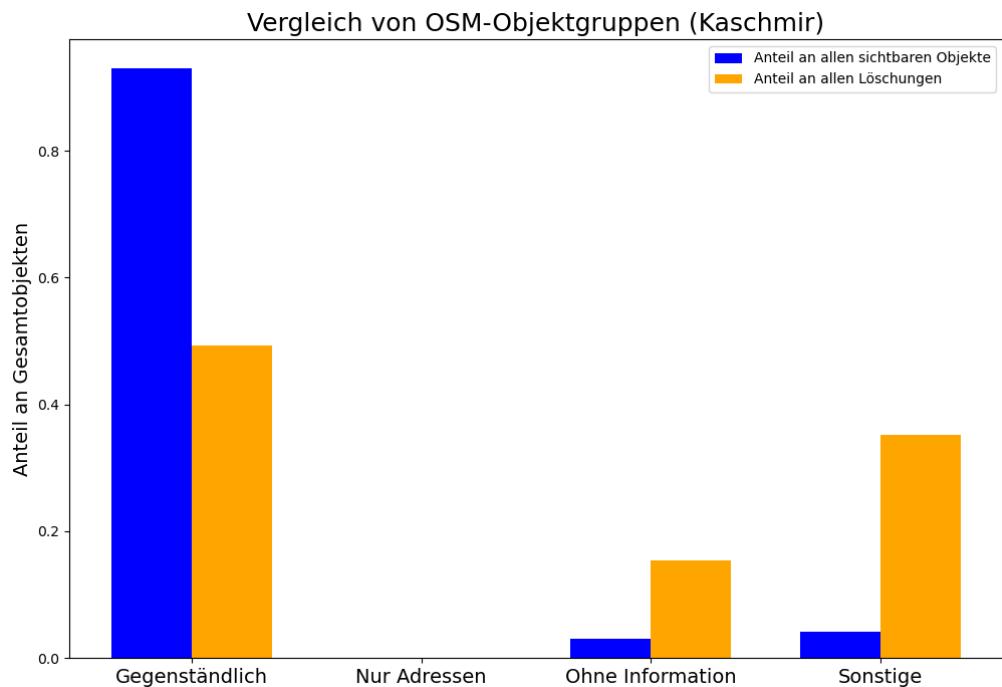


Abbildung 44: Häufigkeit der OSM-Objektgruppen in Kaschmir

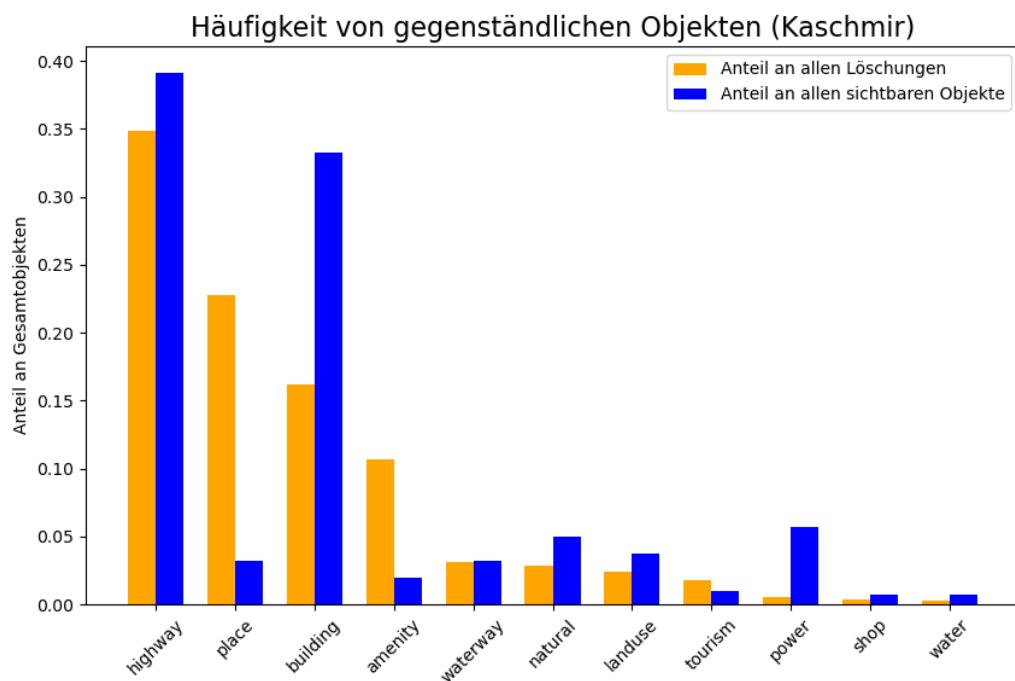


Abbildung 45: Häufigkeit der gegenständlichen Objekte in Kaschmir

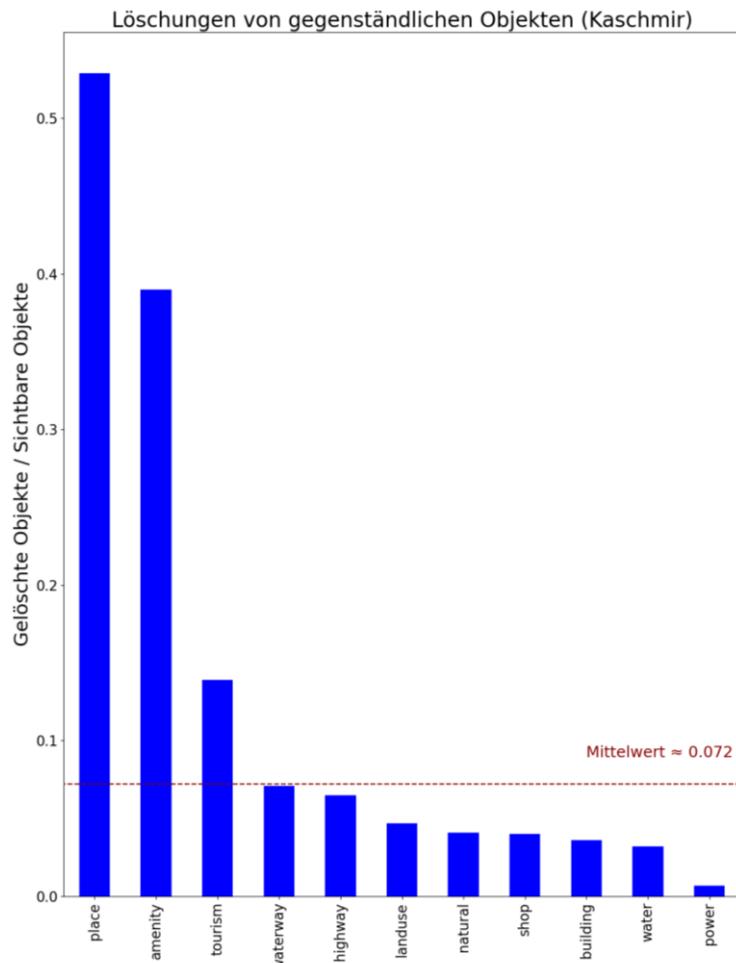


Abbildung 46: Verhältnis zwischen gelöschten und sichtbaren gegenständlichen Objekten in Kaschmir

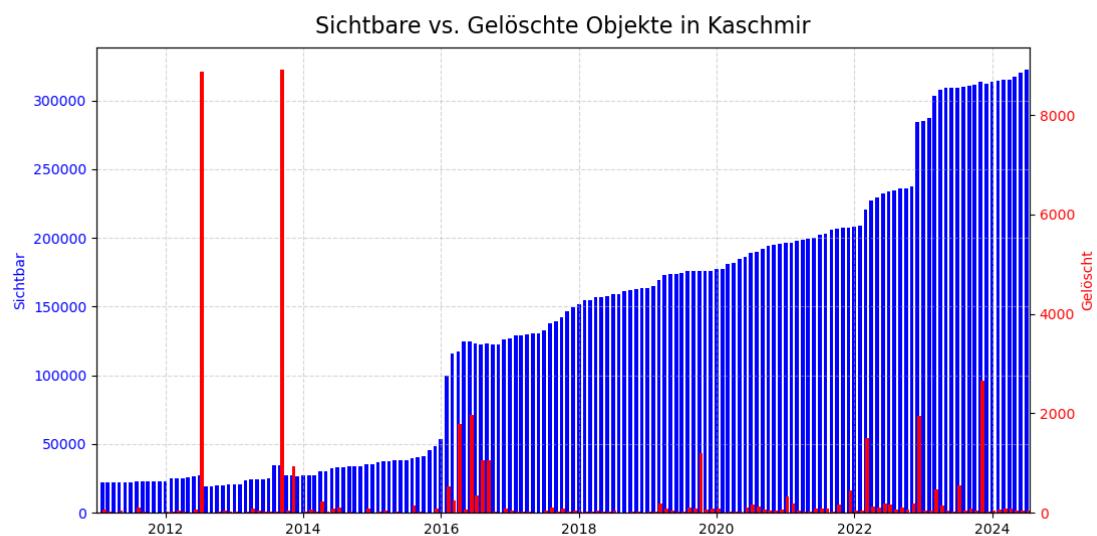


Abbildung 47: Zeitlicher OSM-Verlauf von gelöschten und sichtbaren Objekten in Kaschmir

4.4.3 Karabach

In Karabach gibt es an der nördlichen Grenze zu Aserbaidschan und an der südlichen Grenze zum Iran etwas mehr Löschungen als in den anderen Regionen. Im Umland von Karabach gibt es etwas mehr Löschungen an der Armenisch-Aserbaidschanischen Grenze. Ansonsten sind keine klaren räumlichen Muster zu erkennen. (Abbildung 48)

Das Verhältnis der Anzahl gelöschter Objekte mit der Anzahl existierender Objekte liegt für alle Hexagone bei durchschnittlich rund 13,9%.

Das ist ein fast doppelt so hoher Wert wie für das Umland von Karabach (7,1%). Damit ist die Aussage, dass der Mittelwert für Karabach nicht zufällig größer ist als der Mittelwert des Umlandes, statistisch signifikant (Abbildung 49).

Im Vergleich zu allen Löschungen weltweit (15%) wird in Karabach etwas weniger gelöscht (-1,1%). Allerdings ist der Unterschied so klein, dass er auch aus Zufall hätte entstehen können (Abbildung 50).

Unter den Löschungen machen gegenständliche Objekte mit etwa 85% den größten Anteil aus (Abbildung 51). Darunter fallen größtenteils Straßen mit einem Anteil von rund 35% und Gebäude mit einem Anteil von knapp 23% (Abbildung 52). Alle weiteren Objekte haben einen Anteil von je unter 5%.

Am verhältnismäßig häufigsten gelöscht werden Landnutzungsflächen (*landuse*). Die Anzahl an Löschungen entspricht dabei rund 23% der Anzahl von Objekten auf der sichtbaren Karte. Knapp dahinter folgen Straßen (*highway*) mit einem Wert von etwa 22%. Auch natürliche Flächen (*natural*) werden mit einem Wert von rund 18% noch überproportional häufig gelöscht. Am verhältnismäßig seltensten gelöscht werden mit deutlichem Abstand Elektrizitäts-Einrichtungen (*power*). (Abbildung 53)

Unter den Löschungen befinden sich eine Menge beigefügter Kommentare, die auf den Konflikt hindeuten:

- “*No any political additional on Azerbaijan map please. Please OSM team doesnt allow anybody to do any changes here. I removed all additions in accordance with world regulations*”³⁵

³⁵ <https://www.osm.org/changeset/81651580>

- “Nagorno-Karabakh Republic does not exist. It is not recognized by any country in the world. This is the de jure part of Azerbaijan. By showing this name you are violating international law”³⁶
- “There is not country called Republic of Nagorny Karabagh...Its Azerbaijan territory recognised by all countries in the world, its just some separatists always changing the name too impress their terrorist organization.”³⁷

Die meisten Löschungen passierten im Herbst 2023 (Abbildung 54). Zu dieser Zeit gab es die militärische Großoffensive Aserbaidschans mit anschließender Wiedereingliederung von Karabach in das aserbaidschanische Staatsgebiet. Die Löschungen währenddessen hatten überwiegend folgende Kommentare beigelegt:

- “The city came under Azerbaijani control on October 15, and all non-functional objects were deleted from the map. and fake street address names in Armenian were deleted”³⁸
- “added azerbaijani names and deleted dont working objects”³⁹
- “Artsakh army raised the white flag on September 21. Armenian armed forces were removed from Karabakh. The state of Artsakh was abolished. Karabakh was officially recognized as Azerbaijan”⁴⁰
- Übersetzt aus dem Russischen: “The separatist regime "Nagorno-Karabakh" ceased to exist”⁴¹
- “controled by azerbaijan army”⁴²
- “By 29 September 2023, police of the Ministry of Internal Affairs of Artsakh left all their weapons in Khankendi and completely abandoned the region. Azerbaijani police vehicles began patrolling the city on 29 September.”⁴³
- “Reverting a number of mass removals of Armenian names”⁴⁴

³⁶ <https://osm.org/changeset/41552116>

³⁷ <https://osm.org/changeset/77296883>

³⁸ <https://osm.org/changeset/143248924>

³⁹ <https://osm.org/changeset/145995195>

⁴⁰ u.a. <https://osm.org/changeset/141652578>

⁴¹ u.a. <https://osm.org/changeset/141864426>

⁴² u.a. <https://osm.org/changeset/141931109>

⁴³ u.a. <https://osm.org/changeset/141944377>

⁴⁴ <https://osm.org/changeset/141978430>

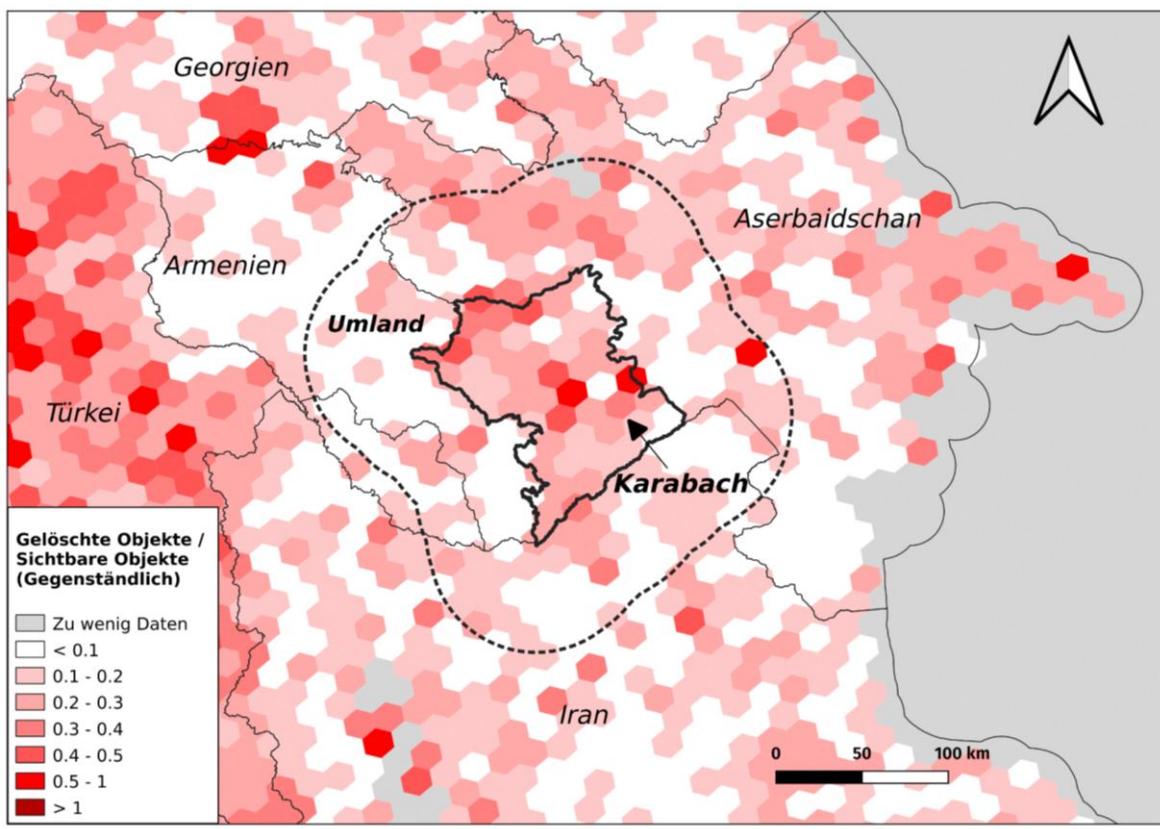


Abbildung 48: Löschungen in der Region um rund Karabach

Karabach vs. Umland

Mittelwert Karabach ≈ 0.139 , Mittelwert Umland ≈ 0.071
Differenz der beobachteten Mittelwerte ≈ 0.068

Permutationstest

Nullhypothese: Beobachteter Mittelwert für Karabach nur zufällig größer

Alternativhypothese: Beobachteter Mittelwert für Karabach nicht zufällig größer

p-Wert $< 1e-03 < 0.05 \rightarrow$ Verwerfung der Nullhypothese

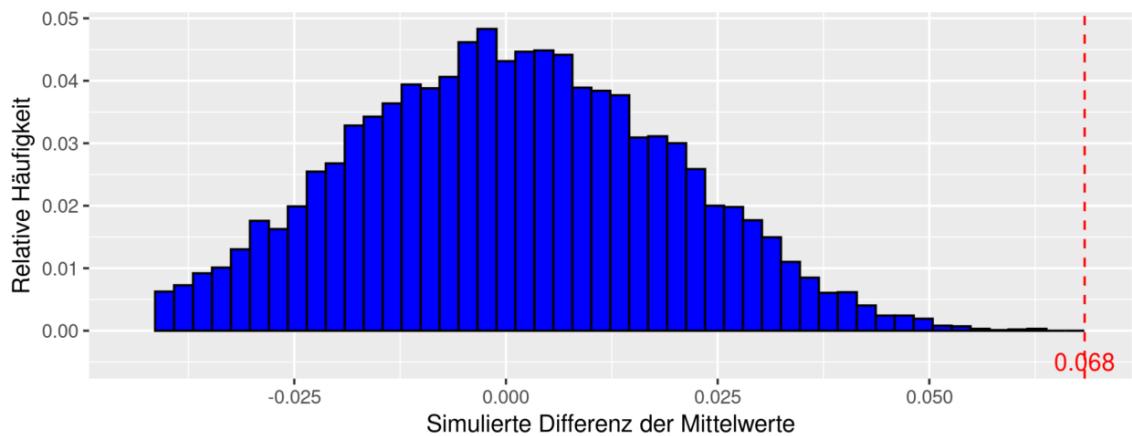


Abbildung 49: Permutationstest zur Überprüfung, ob sich die Menge an Löschungen in Karabach signifikant von dessen Umland unterscheidet

Karabach vs. Global

Mittelwert Karabach ≈ 0.139 , Mittelwert Global ≈ 0.15

Differenz der beobachteten Mittelwerte ≈ -0.01

Permutationstest

Nullhypothese: Beobachteter Mittelwert für Karabach nur zufällig kleiner

Alternativhypothese: Beobachteter Mittelwert für Karabach nicht zufällig kleiner

p-Wert $\approx 0.386 > 0.05 \rightarrow$ Beibehaltung der Nullhypothese

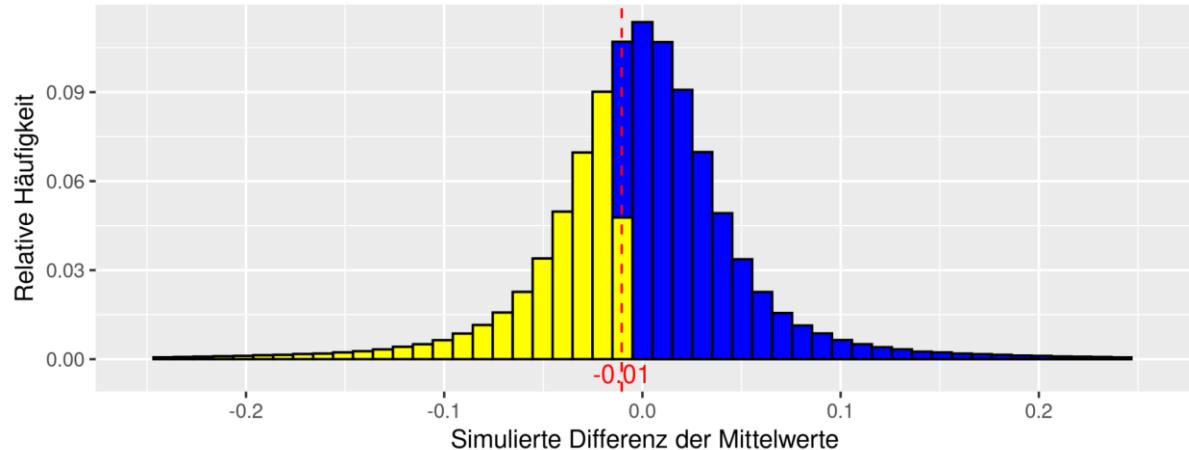


Abbildung 50: Permutationstest zur Überprüfung, ob sich die Menge an Löschungen in Karabach signifikant von dem globalen Raum unterscheidet

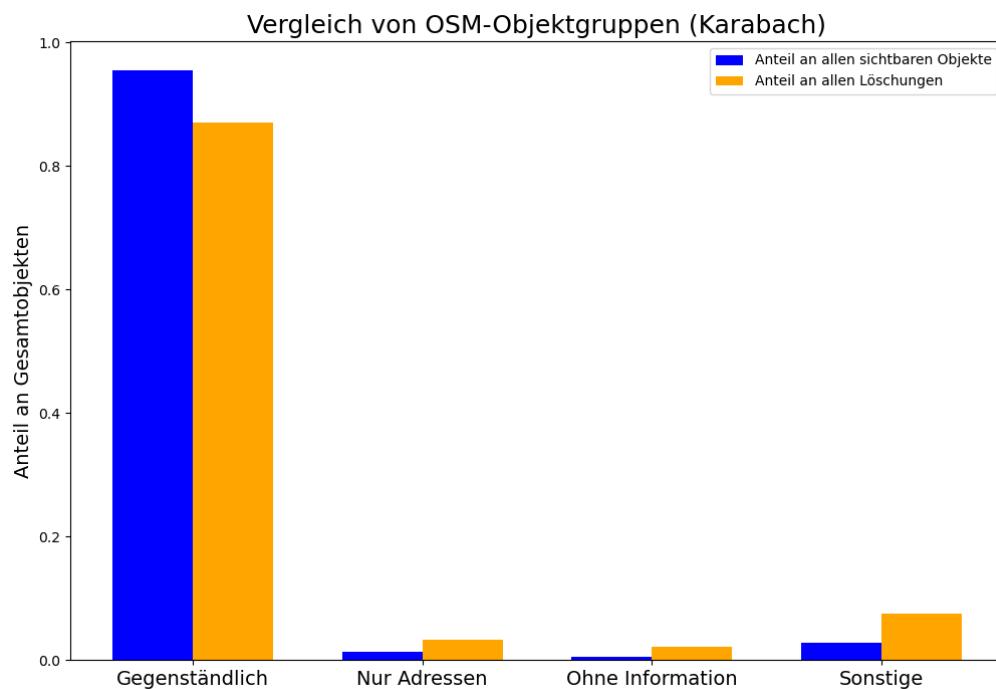


Abbildung 51: Häufigkeit der OSM-Objektgruppen in Karabach

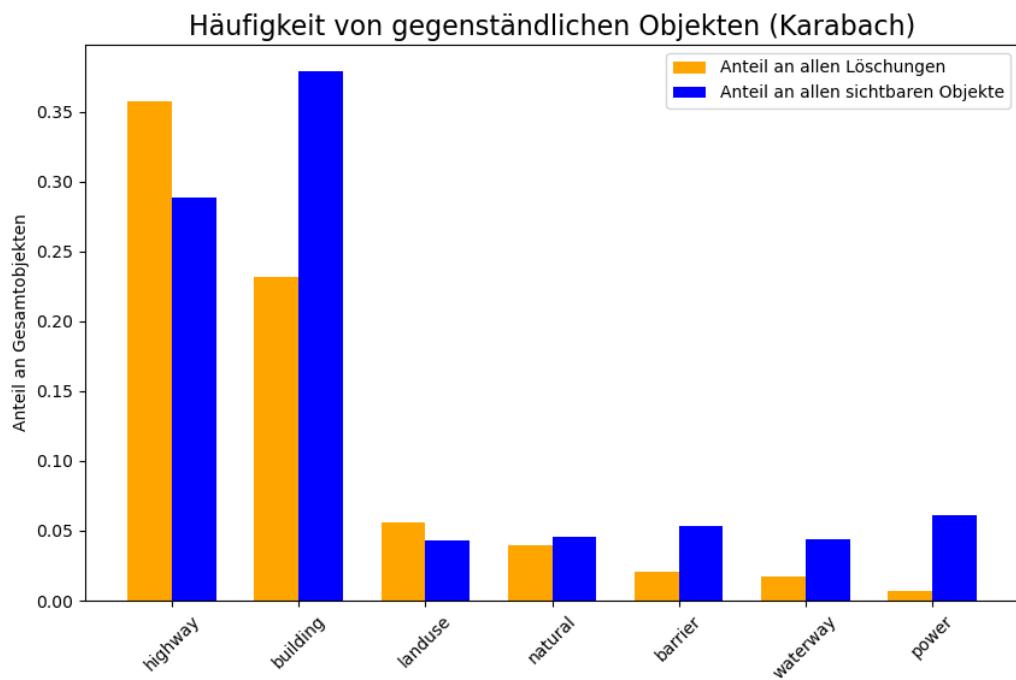


Abbildung 52: Häufigkeit der gegenständlichen Objekte in Karabach

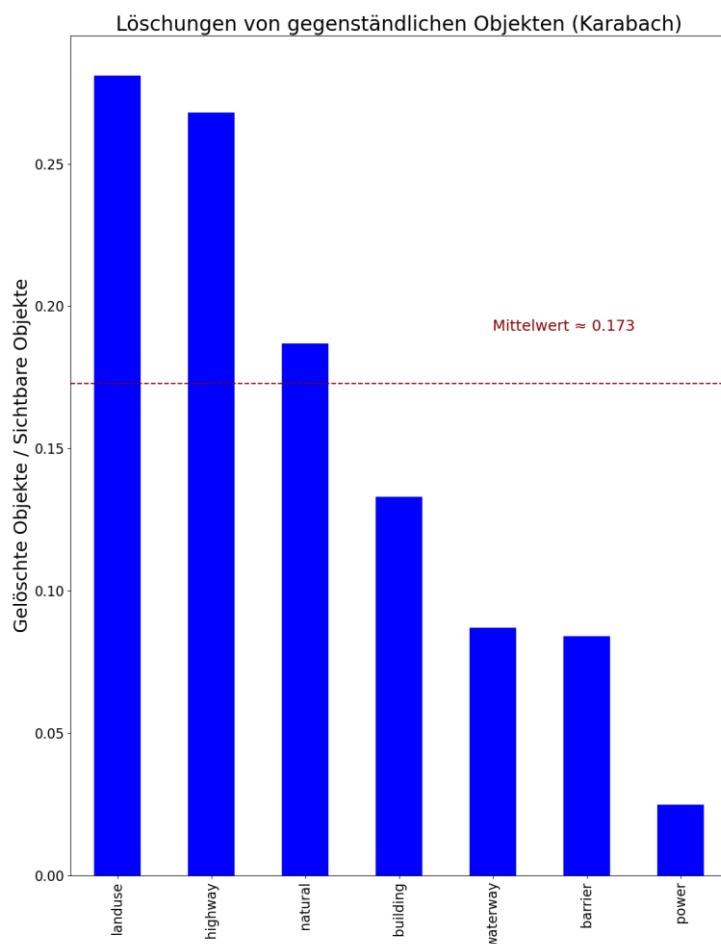


Abbildung 53: Verhältnis zwischen gelöschten und sichtbaren gegenständlichen Objekten in Karabach

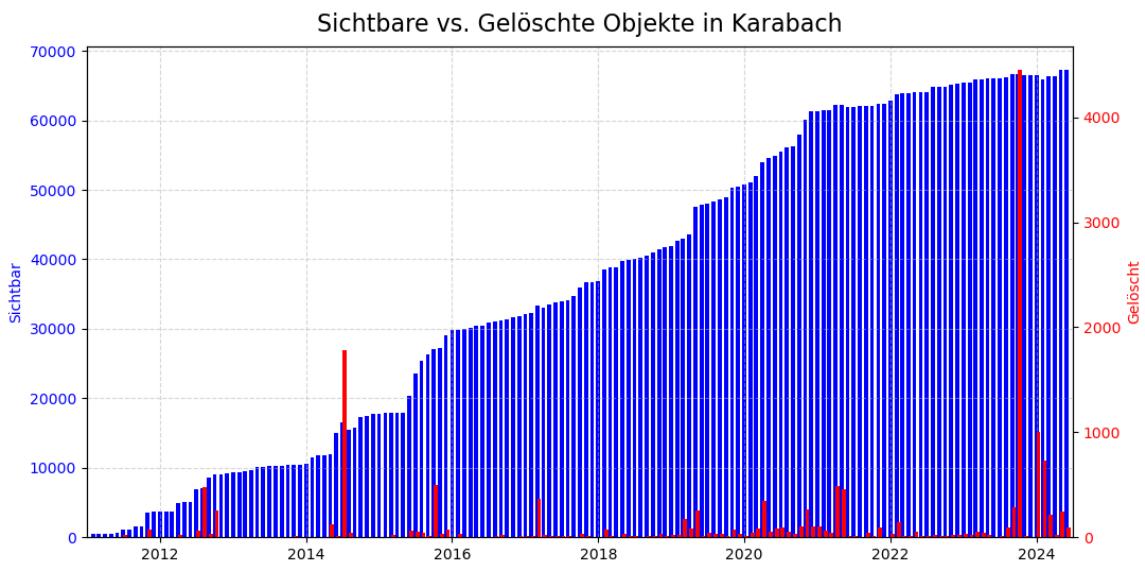


Abbildung 54: Zeitlicher OSM-Verlauf von sichtbaren und gelöschten Objekten in Karabach

4.4.4 Arunachal Pradesh

In Arunachal Pradesh gibt es im östlichen Teil der Region vereinzelt mehr Löschungen. Ansonsten liefern einige Teile des Landes zu wenig Daten, um ein klares räumliches Muster sichtbar zu machen. Im Umland gibt es besonders viele Löschungen im äußersten Nordosten Indiens und südlich von Bhutan. (Abbildung 55)

Das Verhältnis der Anzahl gelöschter Objekte mit der Anzahl existierender Objekte liegt für alle Hexagone im Durchschnitt bei etwa 17,6%. Damit wird in Arunachal Pradesh etwas mehr gelöscht als im Umland (15,8%) und weltweit (15%). Diese Unterschiede sind aber so klein, dass sie auch aus Zufall hätten entstehen können. (Abbildung 56 und Abbildung 57)

Unter den Löschungen machen gegenständliche Objekte mit einem Anteil von knapp 80% die größte Gruppe aus. Die anderen Objekte sind ausschließlich solche, die nicht nach den informellen Standards der OSM-Community beschrieben wurden. Deren Anteil von etwa 20% ist ein verhältnismäßig hoher Wert. (Abbildung 58)

Unter den gelöschten gegenständlichen Objekten fällt knapp die Hälfte auf Objekte zur Repräsentation von Orten (*place*). Danach folgen Straßen (*highway*) mit einem Anteil von

knapp 25%. Auffällig ist, wie gering mit 5% der Anteil an Gebäuden ist, obwohl sie am häufigsten auf der sichtbaren OSM-Karte vorkommen. (Abbildung 59)

Am verhältnismäßig häufigsten gelöscht werden mit Abstand Repräsentationen benannter Orte. Die Anzahl an Löschungen entspricht dabei mehr als 70% der Anzahl von Objekten auf der sichtbaren Karte. Ansonsten gibt es keine weiteren Objekte, die überproportional häufig gelöscht werden. Am verhältnismäßig seltensten gelöscht werden mit Abstand Gebäude. (Abbildung 60)

Die meisten Löschungen passierten im Frühjahr 2014 (Abbildung 61) und hatten folgenden Kommentar beigelegt: “*Revert undisussed imports from unknown source*”. Demnach wurde ein unerwünschter Datenimport wieder gelöscht. Diese Löschungen beinhalteten hauptsächlich Repräsentationen benannter Orte (*place*).

Vereinzelt lassen die beigelegten Kommentare die politischen Spannungen zwischen China und Indien erahnen:

- “*update the border line of China, Because Southern Tibet is owned and administrated by China.*”⁴⁵ (Anm: China bezeichnet Arunachal Pradesh als ‘Südtibet’)
- “*Erroneous International India-China Border*”⁴⁶
- “*this is not china*”⁴⁷
- “*revert erroneous deletions and border edits between China and India*”⁴⁸
- “*moving Chinese language name from `name` tag to `name:zh`. Please do not use the `name` tag to make political statements. This area is administered by India, and should follow Indian naming conventions*”⁴⁹
- “*Moving more Chinese names to `name:zh` tags*”⁵⁰
- “*Fixing name tags on waterways. This area is a defacto part of India and should follow Indian naming conventions, Chinese language should not be in the name tag.*”⁵¹
- “*The path is controversial*”⁵² (Übersetzt aus dem Chinesischen)

⁴⁵ <https://osm.org/changeset/8098467>

⁴⁶ <https://osm.org/changeset/8701684>

⁴⁷ <https://osm.org/changeset/29224549>

⁴⁸ <https://osm.org/changeset/41673567>

⁴⁹ <https://osm.org/changeset/50184932>

⁵⁰ <https://osm.org/changeset/50185046>

⁵¹ <https://osm.org/changeset/50272255>

⁵² <https://osm.org/changeset/100943243>

Allerdings machen diese Löschungen weniger als 5% aller Löschungen in Arunachal Pradesh aus. Demnach lässt sich festhalten, dass sich der Konflikt in den Löschungen kaum widerspiegelt.

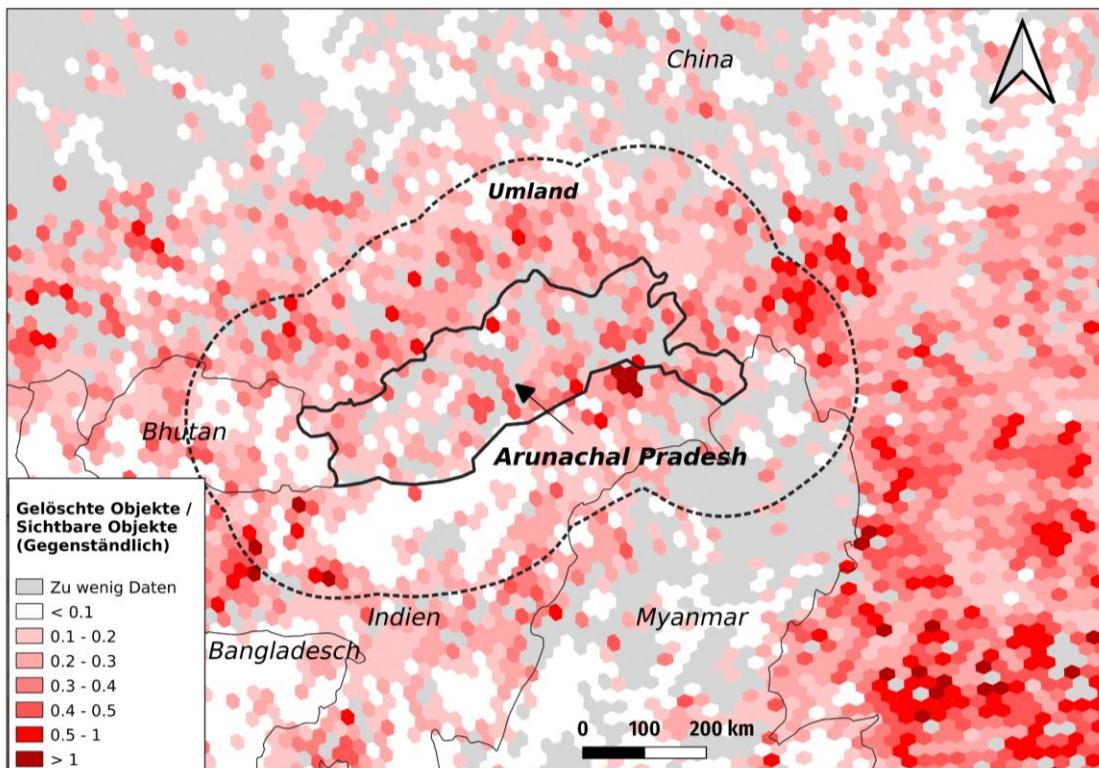


Abbildung 55: Löschungen in der Region rund um Arunachal Pradesh

Arunachal Pradesh vs. Umland

Mittelwert Arunachal Pradesh ≈ 0.176 , Mittelwert Umland ≈ 0.158
Differenz der beobachteten Mittelwerte ≈ 0.018

Permutationstest

Nullhypothese: Beobachteter Mittelwert für Arunachal Pradesh nur zufällig größer
Alternativhypothese: Beobachteter Mittelwert für Arunachal Pradesh nicht zufällig größer

p-Wert $\approx 0.372 > 0.05 \rightarrow$ Beibehaltung der Nullhypothese

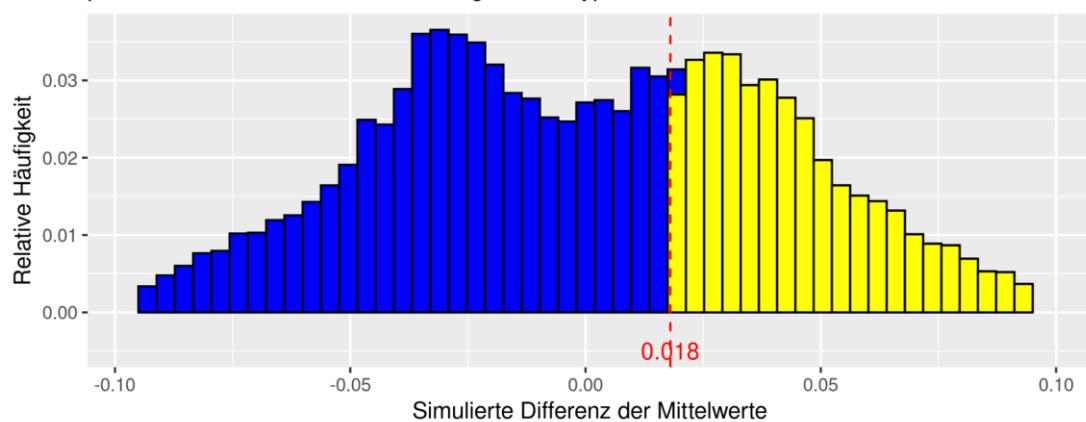


Abbildung 56: Permutationstest zur Überprüfung, ob sich die Menge an Löschungen in Arunachal Pradesh signifikant von dessen Umland unterscheidet; Der p-Wert entspricht dem Anteil der gelben Fläche an der Gesamtfläche der Balken

Arunachal Pradesh vs. Global

Mittelwert Arunachal Pradesh ≈ 0.176 , Mittelwert Global ≈ 0.15

Differenz der beobachteten Mittelwerte ≈ 0.026

Permutationstest

Nullhypothese: Beobachteter Mittelwert für Arunachal Pradesh nur zufällig größer

Alternativhypothese: Beobachteter Mittelwert für Arunachal Pradesh nicht zufällig größer

p-Wert $\approx 0.318 > 0.05 \rightarrow$ Beibehaltung der Nullhypothese

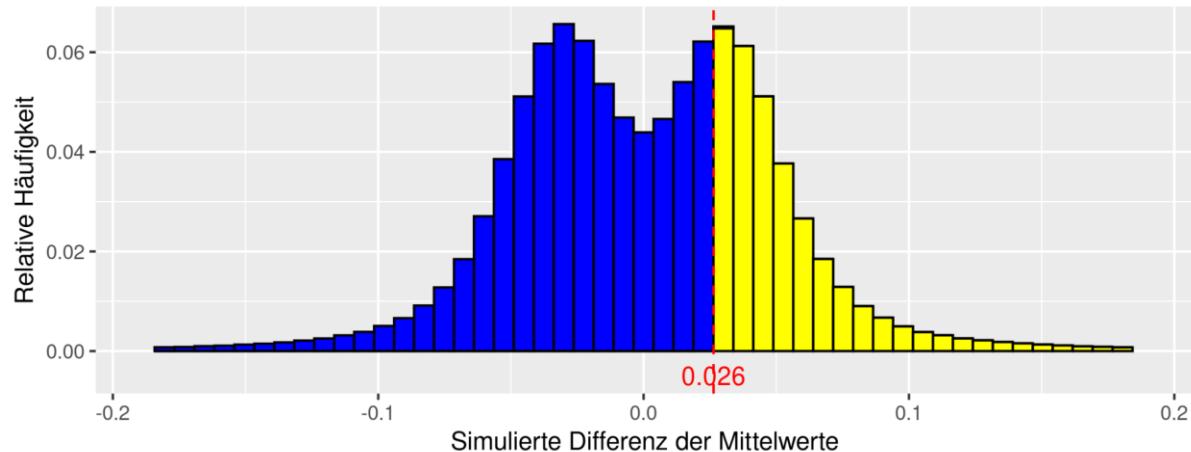


Abbildung 57: Permutationstest zur Überprüfung, ob sich die Menge an Löschungen in Arunachal Pradesh signifikant von dem globalen Raum unterscheidet; Der p-Wert entspricht dem Anteil der gelben Fläche an der Gesamtfläche der Balken

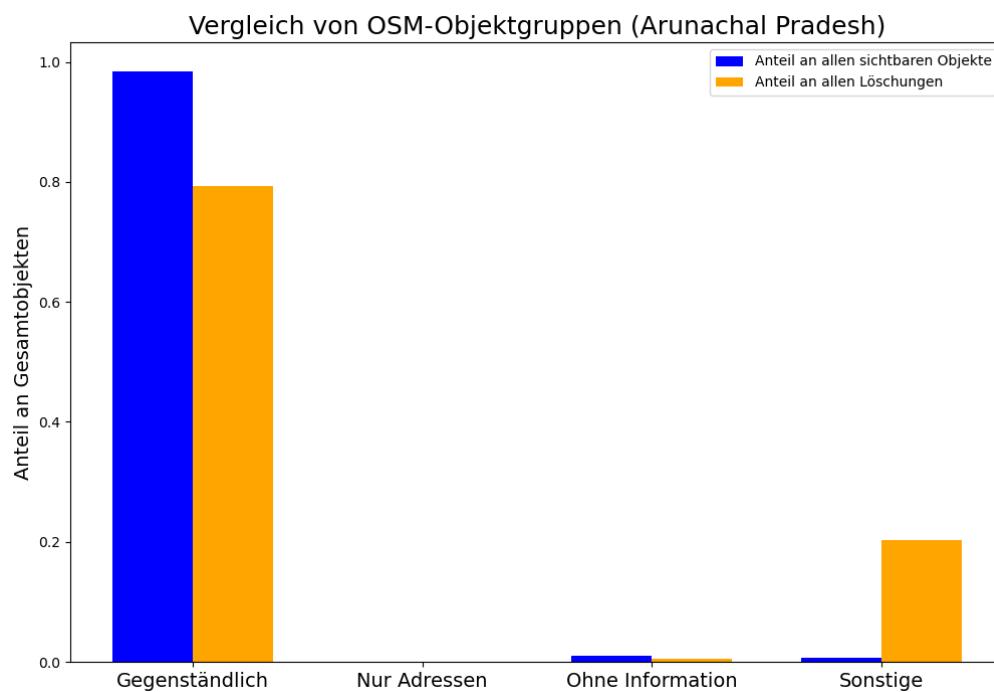


Abbildung 58: Häufigkeit der OSM-Objektgruppen in Arunachal Pradesh

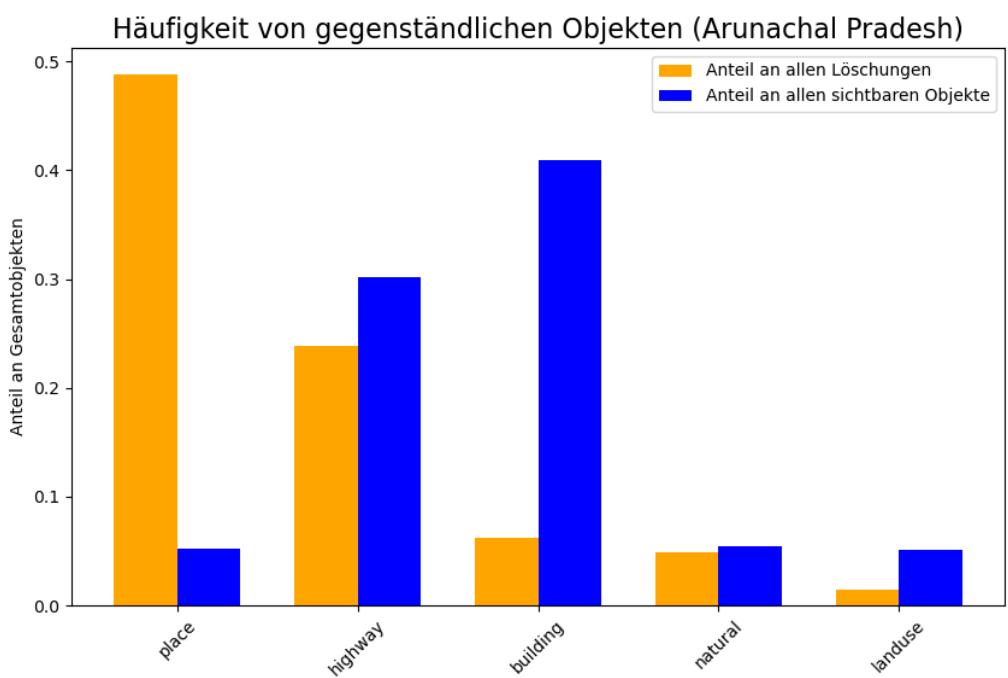


Abbildung 59: Häufigkeit der gegenständlichen Objekte in Arunachal Pradesh

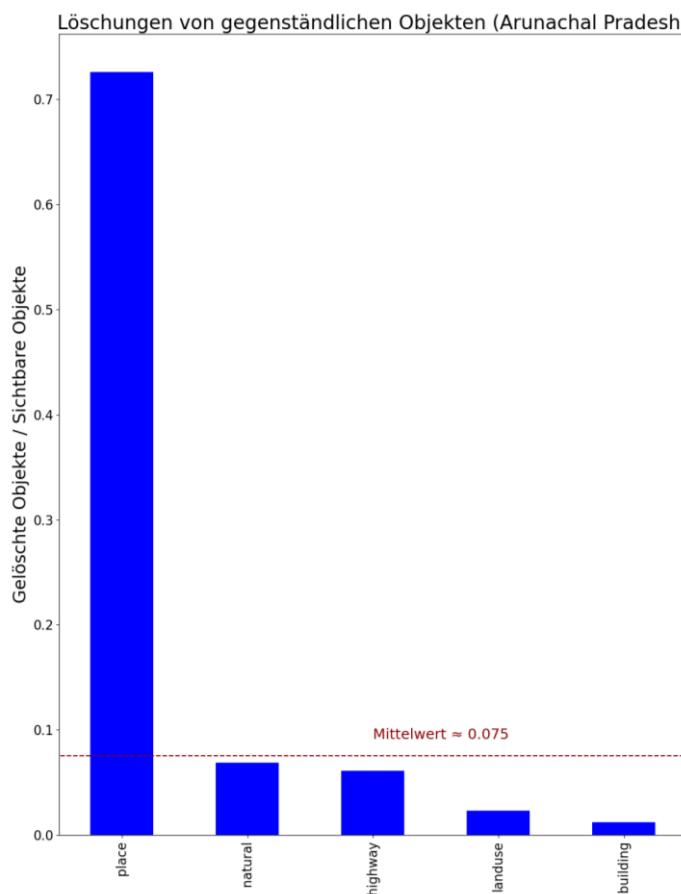


Abbildung 60: Verhältnis zwischen gelöschten und sichtbaren gegenständlichen Objekten in Arunachal Pradesh

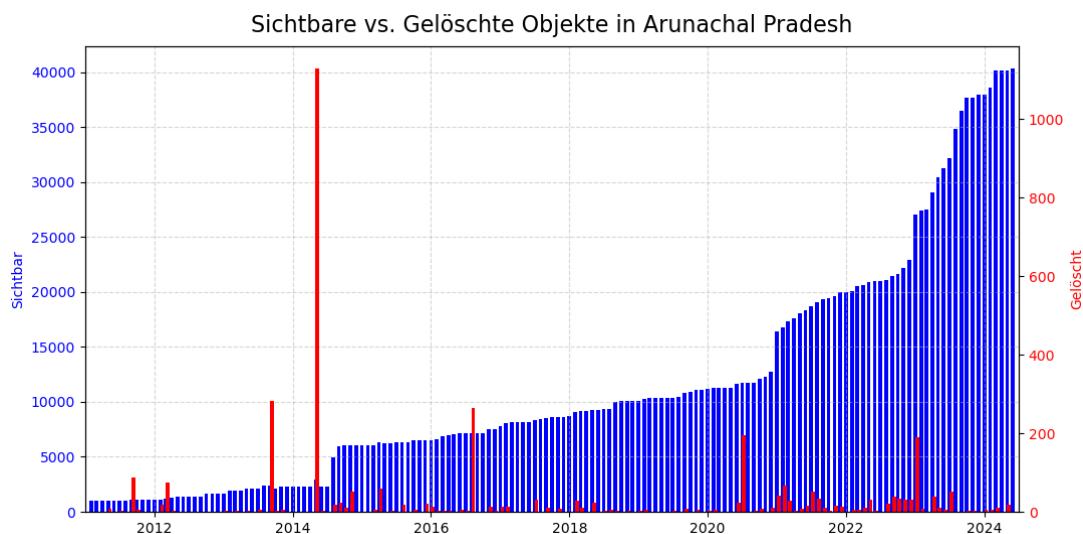


Abbildung 61: Zeitlicher OSM-Verlauf von gelöschten und sichtbaren Objekten in Arunachal Pradesh

4.4.5 Westsahara

Die Westsahara und dessen Umland liegen größtenteils in der Wüste. Deshalb reicht die Datengrundlage nicht aus, um anhand statistischer Tests zu untersuchen, ob im Vergleich zum Umland und weltweit überproportional viel gelöscht wird.

Zumindest in den Küstengebieten gibt es viele Löschungen. Auch ist die inoffizielle Grenze zwischen Marokko und der Polisario-Front zu erkennen, an der ebenfalls ein vergleichsweise hohes Löscherhalten erkennbar ist. Diese von Marokko errichtete Sand- und Steinmauer, auch Berm genannt, verläuft mitten durch die Region und trennt beide Konfliktparteien. (Abbildung 62 und Abbildung 63)

Unter den Löschungen befinden sich überwiegend Objekte, die nicht nach den informellen Standards der OSM-Community beschrieben wurden. Ursache dafür ist, dass in den Küstenregionen die Küstenlinie korrigiert wurde und dafür Objekte gelöscht wurden, die in diese Kategorie fallen. Die Westsahara ist damit die einzige untersuchte Region in dieser Arbeit, bei der gegenständliche Objekte nicht die Mehrheit der Löschungen ausmachen. (Abbildung 64)

Unter den gelöschten gegenständlichen Objekten machen Straßen (22%), Landnutzungsflächen (20%) und Barrieren (18%) den größten Anteil aus. Barrieren werden

damit häufiger gelöscht als Gebäude (11%), obwohl letztgenannte den mit Abstand größten Anteil an sichtbaren Objekten ausmachen. (Abbildung 65)

Unter den gegenständlichen Objekten werden Landnutzungsflächen am verhältnismäßig häufigsten gelöscht. Ihre Anzahl an Löschungen entspricht etwa 33% der Anzahl an Landnutzungs-Objekten auf der sichtbaren Karte. Auf Platz zwei folgen Barrieren mit einem Anteil von rund 12%. Am verhältnismäßig seltensten gelöscht werden dagegen Gebäude. (Abbildung 66)

Die Löschungen fanden zu verschiedenen Zeitpunkten statt (Abbildung 67).

Bei den besonders vielen Löschungen um das Jahr 2012 und Anfang 2014 und Ende 2017 wird anhand der beigefügten Kommentare deutlich, dass es sich überwiegend um konfliktlose Datenkorrekturen handelt:

- “fix unwayed segments”⁵³
- “remapping (Bing)”⁵⁴
- “Western Sahara, align coastline”⁵⁵

Die Löschungen an der Grenze zwischen Marokko und der Polisario-Front sind dagegen auch Ergebnis umstrittener Kartenveränderungen:

- “restored Sahrawi Democratic Republic, see previous discussions in the forum, on the ground rule”⁵⁶
- “reverted sockpuppet edits”⁵⁷
- “it is a dispute region”⁵⁸

Der Begriff *sockpuppet* steht hierbei für anonyme Accounts, die nur angelegt werden, um irreführende Informationen zu verbreiten (ELSNER 2013). Insgesamt sind allerdings die Kommentare, aus denen der territoriale Konflikt deutlich wird, in der Minderheit.

⁵³ u.a. <https://osm.org/changeset/3263956>

⁵⁴ <https://osm.org/changeset/10530250>

⁵⁵ u.a. <https://osm.org/changeset/52428443>

⁵⁶ u.a. <https://osm.org/changeset/38424655>

⁵⁷ u.a. <https://osm.org/changeset/38480265>

⁵⁸ <https://osm.org/changeset/141638677>

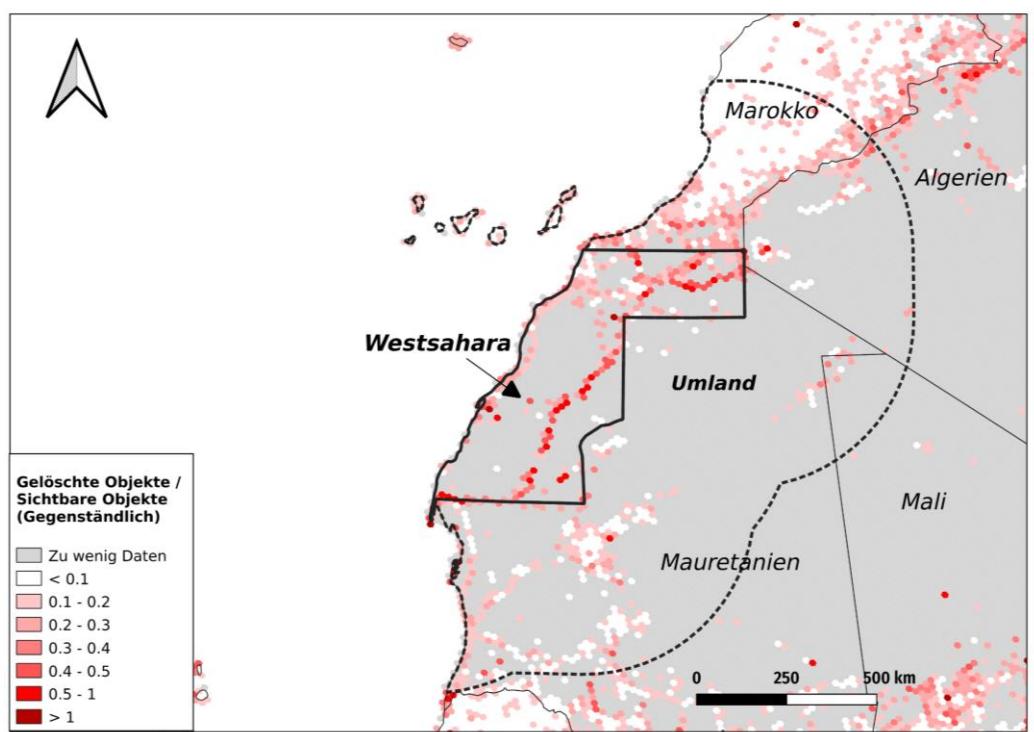


Abbildung 62: Löschungen in der Region rund um die Westsahara

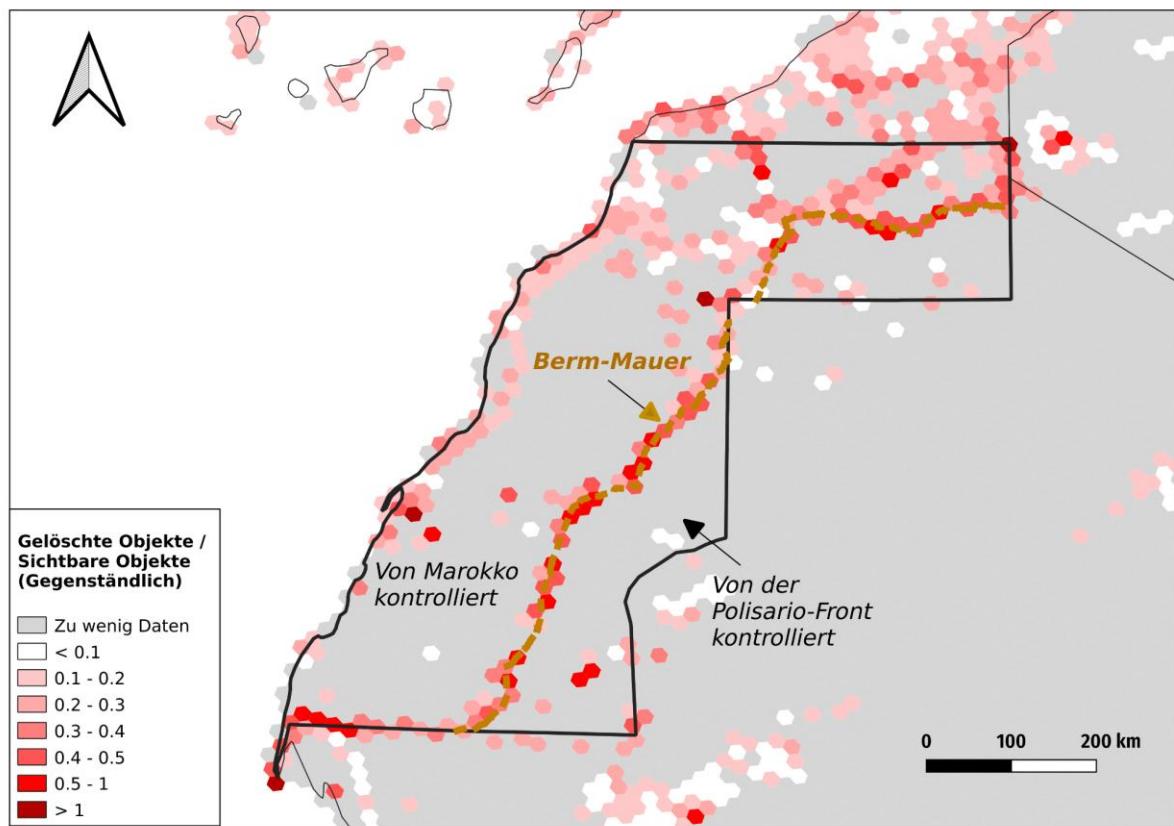


Abbildung 63: Löschungen in der Region rund um die Westsahara mit Berm-Mauer

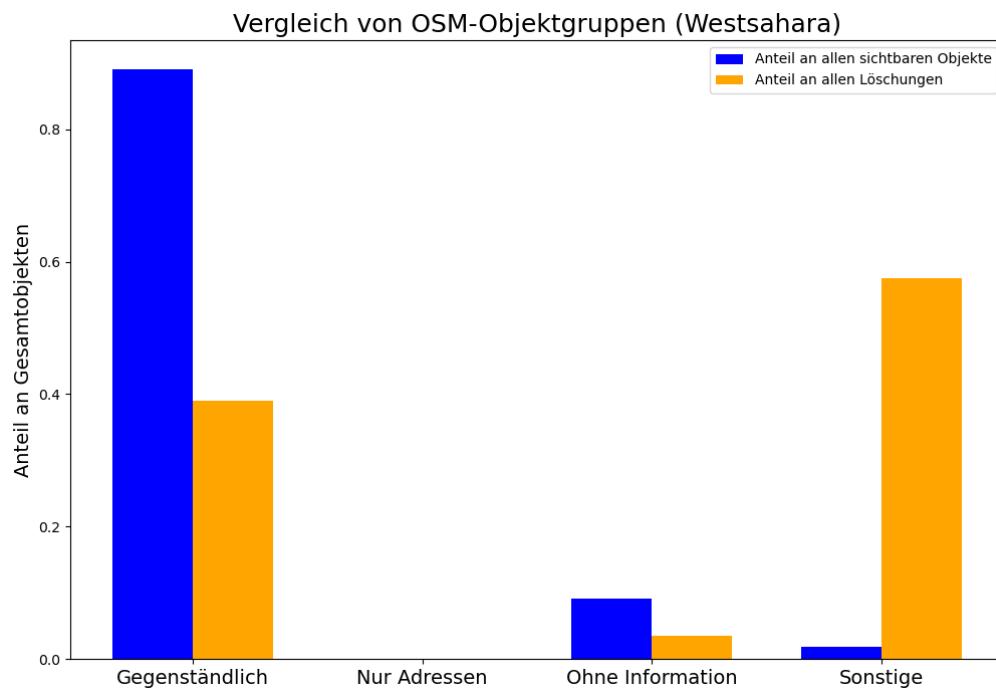


Abbildung 64: Häufigkeit der OSM-Objektgruppen in der Westsahara

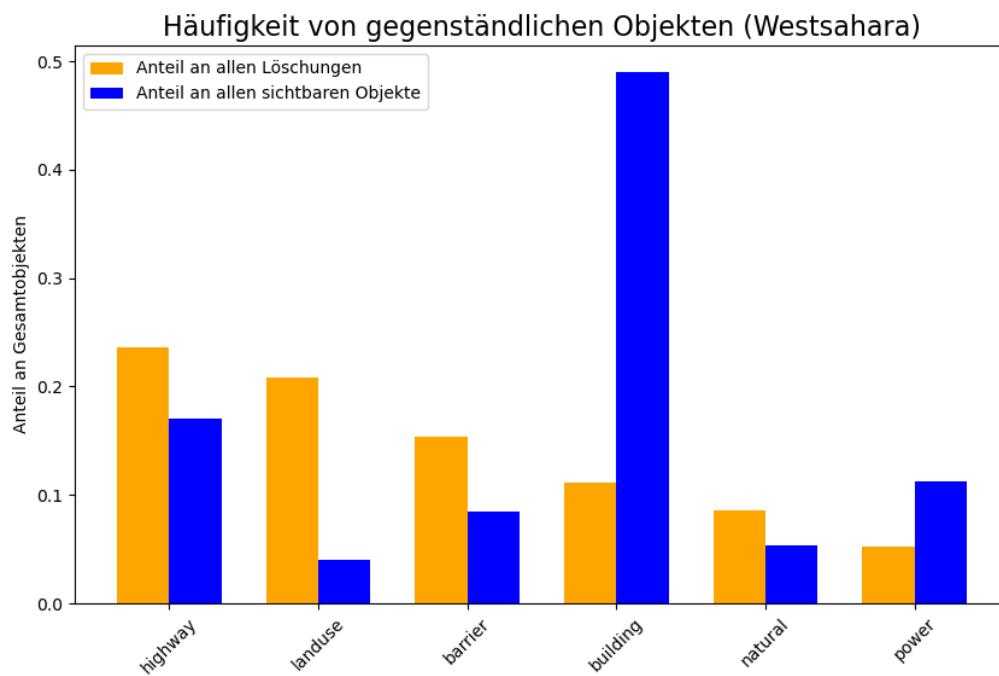


Abbildung 65: Häufigkeit gegenständlicher Objekte in der Westsahara

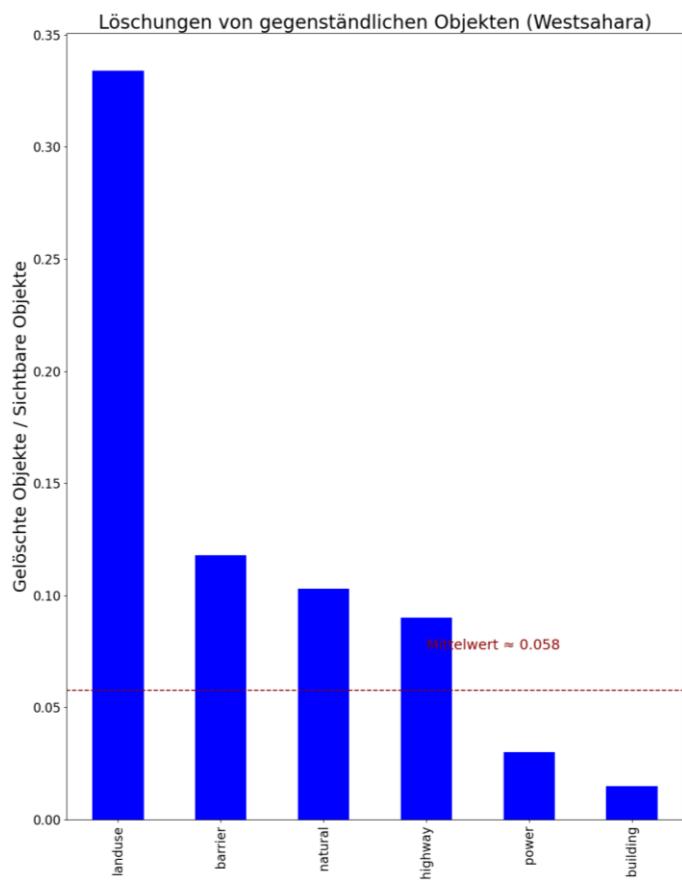


Abbildung 66: Verhältnis zwischen gelöschten und sichtbaren gegenständlichen Objekten in der Westsahara

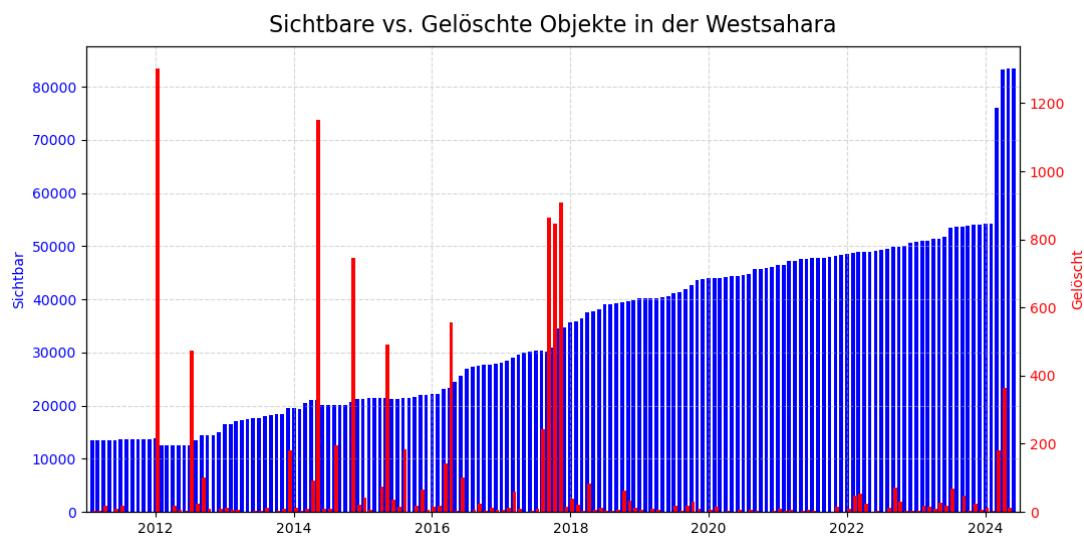


Abbildung 67: Zeitlicher OSM-Verlauf von gelöschten und sichtbaren Objekten in der Westsahara

5. Interpretation

Zunächst werden die globalen Ergebnisse, anschließend die Ergebnisse zu den Regionen mit besonders vielen Löschungen und zuletzt die Ergebnisse zu den Konfliktregionen interpretiert und eingeordnet.

5.1 Globale Räumliche Verbreitung

Bei der ersten Forschungsfrage aus der Einleitung geht es um die globale räumliche Verbreitung von Löschungen und ob großflächige Unterschiede zu erkennen sind. Grundsätzlich gibt es vor allem dort viele Löschungen, wo auch viele sichtbare Objekte auf der OSM-Karte sind. Dieser positive Zusammenhang konnte auch statistisch nachgewiesen werden (Abbildung 6).

Eine der wenigen Ausnahmen ist der Ort des 0. Längen- und Breitengrades. Dieser Ort wird in sozialen Medien und wissenschaftlichen Artikeln auch als *Null Island* bezeichnet und ist das Ergebnis einer fehlerhaften digitalen Verarbeitung räumlicher Informationen (JUHÁSZ & MOONEY 2022). In der realen Welt gibt es dort nichts anderes als eine Boje zur Wetterbeobachtung mitten im Ozean (ebd.). In der digitalen Welt dagegen entstehen dort viele Geoobjekte mit fehlenden oder fehlerhaften Koordinaten (ebd.). Es werden an diesem Ort also deshalb so viele OSM-Objekte gelöscht, weil dort viele Objekte versehentlich erstellt wurden.

Bei der Normierung der Löschungen durch die Anzahl an sichtbaren Objekten zeigt sich, dass großflächige Unterschiede zwischen ganzen Kontinenten oder zwischen dem globalen Norden und dem globalen Süden nicht zu erkennen sind. Die USA, Japan, Türkei und Polen weisen als einzige großflächige Staaten überproportional viele Löschungen auf. Ansonsten treten sichtbare Unterschiede überwiegend lokal auf.

5.2 Globale Objektunterschiede

Bei der zweiten Forschungsfrage aus der Einleitung geht es darum, welche Objekte unter- und welche Objekte überproportional häufig gelöscht werden. Dabei wurden vier Hypothesen aufgestellt, die in diesem Kapitel teilweise bestätigt und teilweise widerlegt werden.

Die Ergebnisse zeigen, dass Gegenständliche Objekte verhältnismäßig selten gelöscht werden. Das liegt vermutlich daran, dass sie definitiv dem informellen Standard der OSM-Community entsprechen (3.2.1). Dadurch fällt der Löschungsgrund für eine nicht standardmäßige Kartierung weg. Adressen und Objekte ohne Informationen werden verhältnismäßig häufig gelöscht. Diese Objekte haben das Merkmal, dass sie teilweise den Standards der OSM-Community entsprechen, teilweise aber auch nicht (3.2.2 und 3.2.3). Damit gibt es vermutlich häufiger den Löschungsgrund für eine nicht standardmäßige Kartierung. Sonstige Objekte, denen nicht die standardmäßigen Informationen beigelegt sind, entsprechen üblicherweise nicht den Vorstellungen der OSM-Community. Das ist vermutlich der Hauptgrund dafür, dass diese Objektgruppe mit Abstand am verhältnismäßig häufigsten gelöscht wird (19% aller Löschungen vs. 2,5% auf der OSM-Karte). Insgesamt bestätigen die Ergebnisse damit die zu Beginn aufgestellte Hypothese, wonach Objekte, die zu wenig nachvollziehbare Informationen beigelegt haben, überproportional häufig gelöscht werden.

Grenzen, größtenteils für administrative Flächen aber auch für geschützte Naturflächen, werden mit Abstand am häufigsten gelöscht. Ein Grund dafür ist vermutlich, dass sie sich durch politische Entscheidungen häufig verändern und deshalb alte Grenzen gelöscht werden. Auch können zwischen OSM-Nutzern politische Meinungsverschiedenheiten bestehen, was zu wechselhaften Erstellungen und Löschungen führt. Touristische Ziele werden ebenfalls überproportional häufig gelöscht. Eine Vermutung dafür ist, dass manche Bewohner an einem Ort mit touristischem Potential nicht möchten, dass ihr Gebiet öffentlichen Andrang ausgesetzt wird, sodass sie touristische Objekte aus der OSM-Karte löschen. Auch könnte Uneinigkeit darüber bestehen, ob ein Objekt touristisch relevant ist oder nicht. Des Weiteren werden Verkaufsorte besonders häufig gelöscht. Solche Objekte unterliegen dynamischen Wirtschaftsprozessen und werden vermutlich häufiger gelöscht, weil sie z.B. im Standortkampf unterliegen und umziehen müssen. Üblicherweise nutzen Verkaufsläden Einrichtungen, die auch nach Verschwinden des Geschäfts bestehen und nicht abgerissen werden. Weshalb solche Objekte als Ganzes gelöscht werden, anstatt nur die beigelegten Informationen zu ändern, bleibt unklar. Objekte für Telekommunikationssysteme und

Elektrizitäts-Einrichtungen werden am seltensten gelöscht. Vermutlich weil es selten Gründe gibt, solche beständigen Einrichtungen, wenn einmal errichtet, abzureißen und woanders aufzubauen. Die gleiche Begründung kann man für Gebäude allgemein und Barrieren wie Wände oder Absperrungen heranziehen, die ebenfalls selten gelöscht werden. Auch geologisch interessante Gebiete und stehende Gewässer werden wohl deshalb selten gelöscht, weil sie sich so gut wie nie wandeln. Noch keine Erklärung gibt es dafür, dass Löschungen selten bei gesundheitlichen und handwerklichen Einrichtungen auftreten.

Insgesamt haben sich zwei zu Beginn aufgestellte Hypothesen bestätigt: Administrative Grenzen werden besonders häufig und natürliche Flächen wie stehende Gewässer oder geologisch interessante Gebiete besonders selten gelöscht. Nicht bestätigt werden konnte dagegen die Hypothese, dass militärische Anlagen besonders häufig gelöscht werden.

5.3 Löschungsreiche Regionen

Bei der dritten Forschungsfrage aus der Einleitung geht es darum, was die Ursachen von Löschungen in Regionen mit besonders vielen Löschungen sind. Aus den Ergebnissen können vier Typen von Löschungen abgeleitet werden (Tabelle 1).

	<i>Typ C (Classic)</i>	<i>Typ RI (Revert Import)</i>	<i>Typ PI (Process Import)</i>	<i>Typ D⁵⁹ (Dispute)</i>
Beschreibung	Objekte werden dauerhaft aus der Karte gelöscht	Rückabwicklung eines Imports	Löschen im Kontext von Nachbearbeitungen eines Imports	Neben vielen Löschungen auch viele Neuerstellungen, vermutlich aufgrund von Meinungsverschiedenheiten
Formale Kriterien	<p>Die Anzahl an sichtbaren Objekten sinkt in einem Monat um mindestens 20% im Vergleich zu der durchschnittlichen Anzahl sichtbarer Objekte der vorherigen fünf Monate</p> <p>In den folgenden zwei Monaten steigt die Anzahl der sichtbaren Objekte um weniger als 20%</p>	<p>Die Anzahl an sichtbaren Objekten steigt innerhalb von ein bis drei Monaten um mindestens 10% an</p> <p>In demselben Zeitraum ist die Anzahl gelöschter Objekte dreimal so hoch wie in den drei vorangegangenen Monaten</p> <p>Innerhalb der nächsten zwölf Monate gibt es einen Zeitraum von ein bis drei Monaten, bei dem die Anzahl an sichtbaren Objekten um fast die gleiche Menge abnimmt, wie sie in Bedingung 1 zugenommen hat (Betrag des Rückgangs ist zwischen 90% und 110% des Anstieges)</p>	<p>In einem Zeitraum von ein bis drei Monaten steigt die Anzahl sichtbarer Objekte um mindestens 10% der vorherigen drei Monate an</p> <p>In demselben Zeitraum ist die Anzahl gelöschter Objekte dreimal so hoch wie in den drei vorangegangenen Monaten</p>	<p>In einem Zeitraum von ein bis drei Monaten macht die Summe der Löschungen mindestens 5% der durchschnittlichen Anzahl an sichtbaren Objekten aus</p>

Tabelle 1: Die vier verschiedenen Löschungstypen in der Übersicht. Für andere Regionen wurden sie noch nicht getestet. Das wäre eine sinnvolle Ergänzung zu dieser Arbeit

Zum einen den klassischen Fall, bei dem Objekte dauerhaft aus der OSM-Karte gelöscht werden und die Anzahl an sichtbaren OSM-Objekten auf der Karte weniger wird (Abbildung 68). Dieser Löschungstyp wird in dieser Arbeit als Typ C (“Classic”) bezeichnet.

⁵⁹ Typ D wird zwar erst in Kapitel 5.4 eingeführt, aber hier der Übersicht halber bereits erwähnt; Für diesen Typ sollten idealerweise semantische Kriterien ergänzt werden, d.h. dass die Kommentare zu den Löschungen betrachtet und nach Hinweisen auf den territorialen Konflikt untersucht werden

Formal ist Typ C unter den folgenden Bedingungen vorhanden:

1. Die Anzahl an sichtbaren Objekten sinkt in einem Monat um mindestens 20% im Vergleich zu der durchschnittlichen Anzahl sichtbarer Objekte der vorherigen fünf Monate
2. In den folgenden zwei Monaten steigt die Anzahl der sichtbaren Objekte um weniger als 20%

In den untersuchten Regionen trat Typ C vor allem bei der automatisierten Lizenzänderung der OSM-Daten im Jahr 2012 auf. Bei dieser Umwandlung gingen etwa 1% aller OSM-Daten verloren. Ansonsten trat dieser Typ in den untersuchten Regionen kaum in Erscheinung.

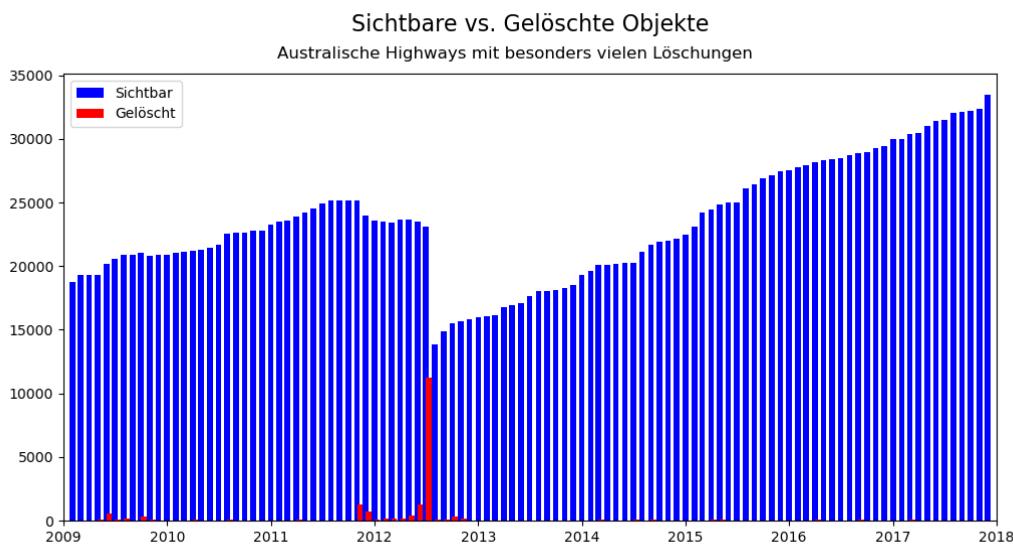


Abbildung 68: Löschungstyp C, Fallbeispiel Australien

Des Weiteren gibt es Löschungen, die eine Reaktion auf einen umfangreichen Datenimport sind. Dabei wird der Import rückgängig gemacht und der vorherige Zustand auf der OSM-Karte wiederhergestellt (Abbildung 69). Dieser Löschungstyp wird in dieser Arbeit deshalb als Typ RI (“Revert Import”) bezeichnet.

Formal ist Typ RI unter den folgenden Bedingungen vorhanden:

1. Die Anzahl an sichtbaren Objekten steigt innerhalb von ein bis drei Monaten um mindestens 10% an
2. In demselben Zeitraum ist die Anzahl gelöschter Objekte dreimal so hoch wie in den drei vorangegangenen Monaten

3. Innerhalb der nächsten zwölf Monate gibt es einen Zeitraum von ein bis drei Monaten, bei dem die Anzahl an sichtbaren Objekten um fast die gleiche Menge abnimmt, wie sie in Bedingung 1 zugenommen hat (Betrag des Rückgangs ist zwischen 90% und 110% des Anstieges)

Ursache für Typ *RI* ist, dass die Datenqualität des Imports mangelhaft oder die OSM-Richtlinien⁶⁰ für Imports nicht beachtet wurden. Das heißt, dass z.B. keine nachvollziehbare Dokumentation erstellt wurde, wie genau die Integration der externen Daten in den bestehenden Datensatz abläuft.

Nicht immer wird der Import sofort rückgängig gemacht: In Finnland dauerte es knapp ein Jahr, bis die Objekte wieder gelöscht wurden (Abbildung 69). Insgesamt kam Löschungstyp *RI* bei den untersuchten Regionen am häufigsten auf (Dschibuti, Finnland, Lettland, Südafrika)

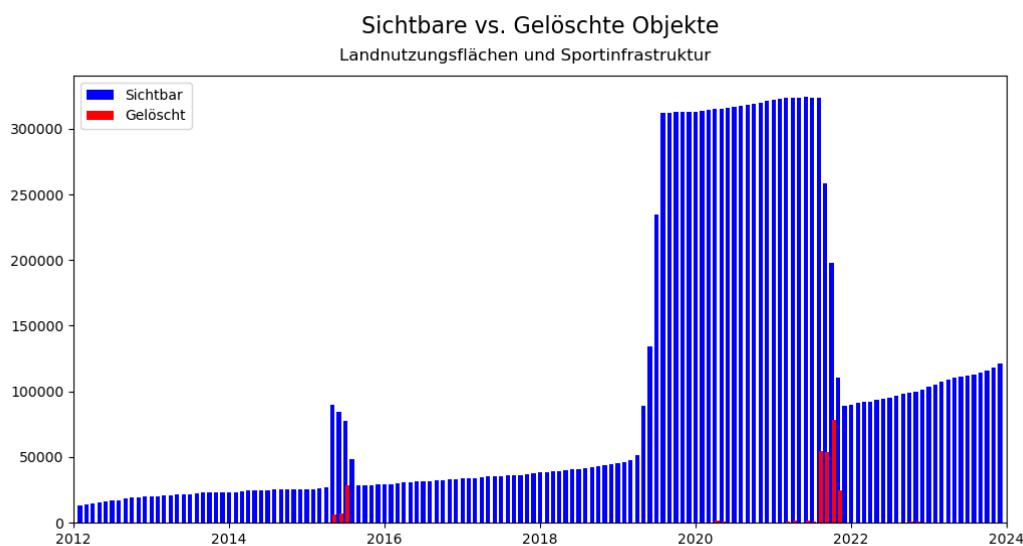


Abbildung 69: Löschungstyp *RI*, Fallbeispiel Finnland

Bei einem weiteren Löschungstyp erfolgen die Löschungen im Kontext eines umfangreichen Imports. Das heißt, dass entweder nur Teile des Imports wieder entfernt oder zeitgleich andere Objekte aus der OSM-Karte herausgelöscht werden. Trotz der auftretenden Löschungen steigt demnach die Anzahl sichtbarer Objekte auf der OSM-Karte an (Abbildung 70). Aufgrund der Nachbearbeitung des Imports wird dieser Löschungstyp als Typ *PI* ("Process Import") bezeichnet.

⁶⁰ <https://wiki.openstreetmap.org/wiki/Import/Guidelines>

Formal ist Typ *P*/*I* unter den folgenden Bedingungen vorhanden:

1. In einem Zeitraum von ein bis drei Monaten steigt die Anzahl sichtbarer Objekte um mindestens 10% der vorherigen drei Monate an
2. In demselben Zeitraum ist die Anzahl gelöschter Objekte dreimal so hoch wie in den drei vorangegangenen Monaten

Auslöser für Typ *P*/*I* kann z.B. sein, dass Nutzer einen externen Datensatz importieren und anschließend die Objekte davon wieder entfernen, die entweder ausreichend gut von anderen Objekten repräsentiert sind oder die nicht den erwarteten Qualitätsstandards entsprechen. Ein weiterer Auslöser kann sein, dass durch den Import vorherige Objekte überholt und deshalb aus der OSM-Karte gelöscht werden.

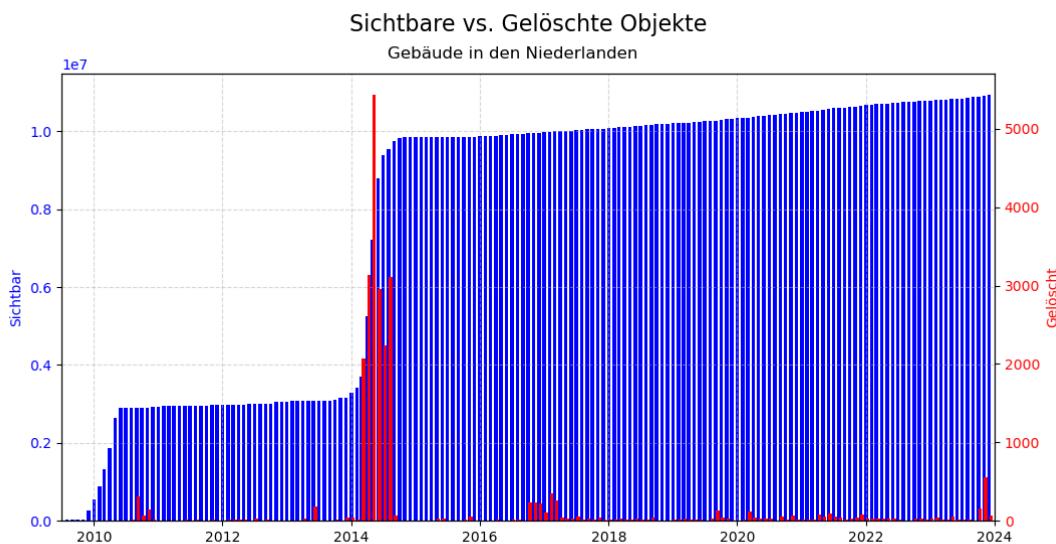


Abbildung 70: Löschungstyp M, Fallbeispiel Niederlande

Insgesamt betrachtet passieren in einer Region besonders dann viele Löschen, wenn ihnen ein umfangreicher Import vorangegangen ist. Die Löschen kommen dann dadurch zustande, dass dieser Import entweder komplett rückgängig gemacht (Typ *R*/*I*) oder nachbearbeitet wird (Typ *P*/*I*). Der klassische Fall, bei dem Objekte dauerhaft aus der Karte gelöscht werden (Typ *C*), kommt seltener vor.

5.4 Konfliktregionen

Bei der vierten Forschungsfrage aus der Einleitung geht es darum, ob territoriale Konflikte in der realen Welt einen Einfluss darauf haben, dass mehr Löschungen auftreten. Die Auswertungen für die untersuchten Konfliktregionen zeigen dabei unterschiedliche Ergebnisse.

Das Westjordanland hat zwar deutlich mehr Löschungen als das Umland vorzuweisen, allerdings beziehen sich diese Löschungen überwiegend auf die Rückabwicklung eines unerwünschten Datenimports, der nichts mit dem Konflikt zu tun hat.

In Kaschmir unterscheidet sich die Anzahl der Löschungen nicht signifikant vom Umland und die meisten Löschungen haben ebenfalls nichts mit dem Konflikt zwischen Indien und Pakistan zu tun. In Arunachal Pradesh ist die Anzahl der Löschungen auch nicht signifikant größer. Allerdings gibt es hier vereinzelte Löschungen mit klarem Bezug zu dem Konflikt zwischen Indien und China. Sowohl in Kaschmir als auch in Arunachal Pradesh gab es einen umfangreichen Import von Ortsbezeichnungen in indischer Sprache, der anschließend rückgängig gemacht wurde. Unklar ist dabei, ob dieser Import politisch motiviert war oder nicht. In der Westsahara machen unpolitische Datenkorrekturen den größten Teil der Löschungen aus. Lediglich an der inoffiziellen Grenze zwischen Marokko und der Polisario-Front gibt es Hinweise darauf, dass das Territorium umstritten ist. Solche Löschungen sind aber die Ausnahme und nicht der Regelfall.

In Karabach dagegen spiegelt sich der Konflikt zwischen Armenien und Aserbaidschan eindeutig in den Löschungen wider. So stieg die Anzahl der Löschungen stark an, als Aserbaidschan im September 2023 eine militärische Großoffensive begann und innerhalb weniger Tage die Region unter seine Kontrolle brachte. Aus den beigefügten Kommentaren wird deutlich, dass ein Teil der Löschungen gemacht wurde, um "nicht-funktionsfähige" Objekte zu entfernen und um aserbaidschanische Bezeichnungen hinzuzufügen. Was unter nicht-funktionsfähigen Objekten zu verstehen ist und weshalb sie gelöscht wurden, wird aus den Kommentaren nicht deutlich. Da die gelöschten Objekte dabei überwiegend zivile Einrichtungen wie Verkaufsläden und Friseursalons sind, handelt es sich vermutlich um verlassene Gebäude von armenischen Flüchtlingen, die für nichts mehr zu gebrauchen waren.

Repräsentationen benannter Orte (*place*) werden in Kaschmir und Arunachal Pradesh am verhältnismäßig häufigsten gelöscht. Grund dafür ist der umfangreiche Import mit indischen Ortsbezeichnungen, der nach kurzer Zeit wieder gelöscht wurde. Auch der hohe Anteil an

gelöschten Gebäuden im Westjordanland lässt sich auf einen Import dieser Art zurückführen. In Karabach und der Westsahara werden Landnutzungsflächen (*landuse*) auffällig häufig gelöscht. Eine Erklärung gibt es dafür bislang nicht. Auch ist unklar, weshalb in Karabach Straßen (*highway*) besonders häufig gelöscht wurden.

Bei der Betrachtung des zeitlichen Verlaufs von Löschungen und sichtbaren Objekten auf der Karte ist im Westjordanland und in Arunachal Pradesh Löschungstyp *Rl* (siehe 5.2) dominant (Abbildung 71 und Abbildung 72). Damit spiegeln sich in den Löschungen hauptsächlich Rückabwicklungen falscher Imports wider.

In Kaschmir und in der Westsahara gibt es eine Kombination mehrerer Löschungstypen (Abbildung 73 und Abbildung 74). Das deutet darauf hin, dass es in diesen Regionen nicht nur eine, sondern mehrere Ursachen für Löschungen gibt.

In Karabach passierten besonders viele Löschungen, als sich der territoriale Konflikt zwischen Armenien und Aserbaidschan zuspitzte. Die Anzahl der sichtbaren Objekte sank währenddessen kaum. Auch gab es keinen vorangegangenen Import wie bei Löschungstyp *Rl*. Das deutet darauf hin, dass neben den Löschungen gleichzeitig auch viele neue Objekte der Karte hinzugefügt wurden. Das könnte an Meinungsverschiedenheiten zwischen Kartierern liegen, bei denen OSM-Objekte im Wechsel gelöscht und wieder neu erstellt werden. Deshalb wird neben den drei vorgestellten Löschungstypen in Kapitel 5.3 ein weiterer Typ *D* ("*Dispute*") in dieser Arbeit eingeführt. (Abbildung 75)

Formal ist Typ D unter den folgenden Bedingungen vorhanden:

1. In einem Zeitraum von ein bis drei Monaten macht die Summe der Löschungen mindestens 5% der durchschnittlichen Anzahl an sichtbaren Objekten aus
2. Gleichzeitig gibt es keine Verringerung der Anzahl sichtbarer Objekte um mehr als 1%

Insgesamt lässt sich anhand der untersuchten Regionen festhalten, dass Löschungen mit Bezug zu einem territorialen Konflikt existieren, meistens aber nur einen kleinen Teil der Löschungen ausmachen. Nur in Karabach spiegelte sich der territoriale Konflikt deutlich in den Löschungen wider. Stattdessen kommen vor allem solche Löschungen vor, denen ein umfangreicher Datenimport vorangegangen ist. Dieser Import wird dann entweder komplett rückgängig gemacht (Typ *Rl*) oder nachbearbeitet (Typ *Pt*).

Die Hypothese aus der Einleitung, wonach sich territoriale Konflikte eindeutig in den Löschungen widerspiegeln, kann demnach weder bestätigt noch widerlegt werden.

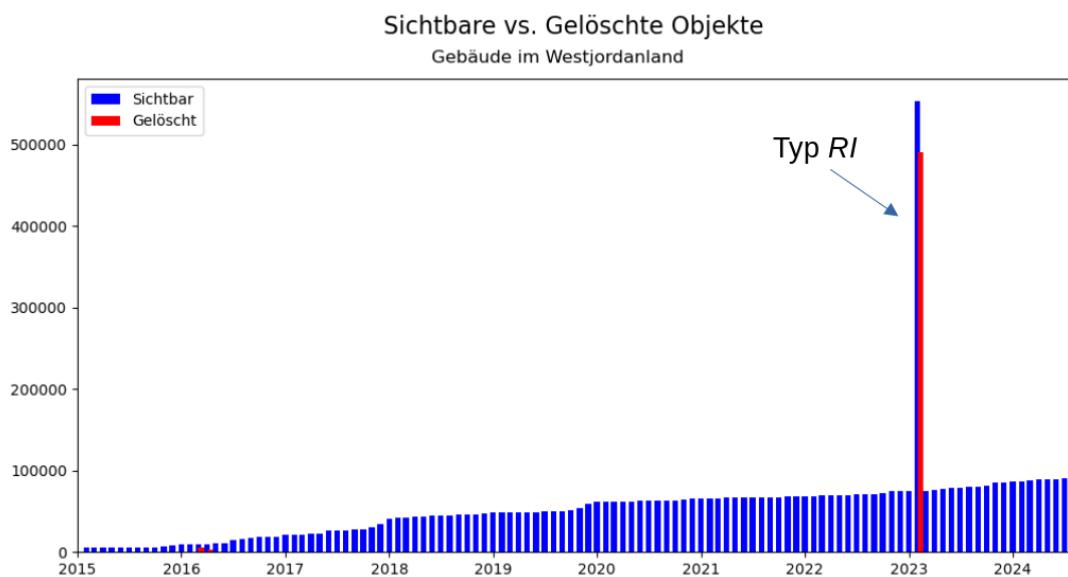


Abbildung 71: Zeitlicher OSM-Verlauf von gelöschten und sichtbaren Gebäuden im Westjordanland mit Typ-Beschriftung

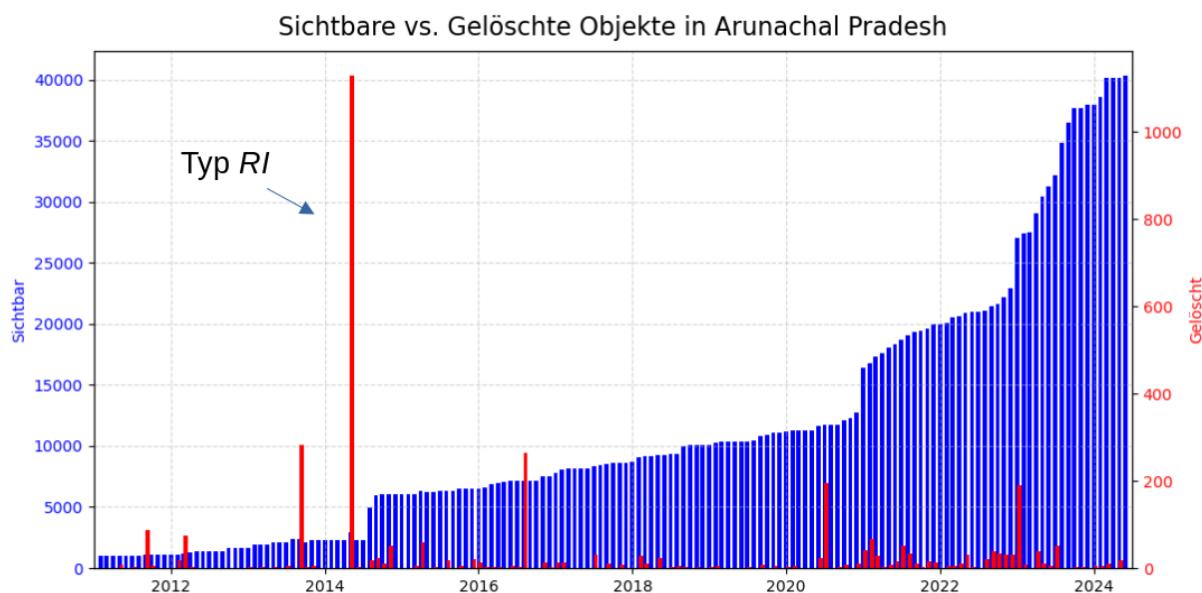


Abbildung 72: Zeitlicher OSM-Verlauf von gelöschten und sichtbaren Objekten in Arunachal Pradesh mit Typ-Beschriftung

Sichtbare vs. Gelöschte Objekte in Kaschmir

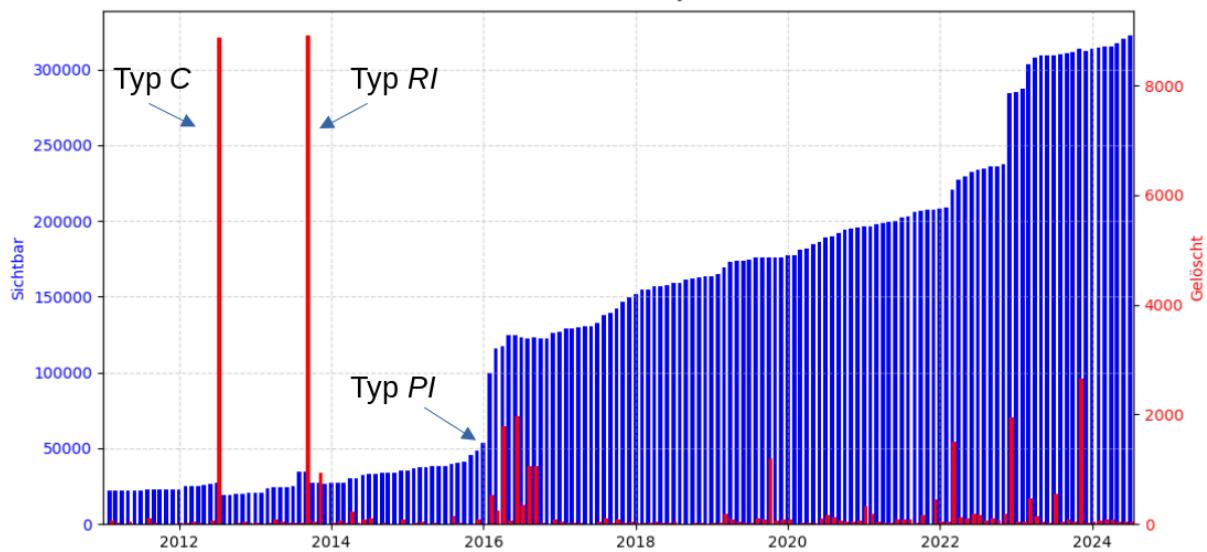


Abbildung 73: Zeitlicher OSM-Verlauf von gelöschten und sichtbaren Objekten in Kaschmir mit Typ-Beschriftung

Sichtbare vs. Gelöschte Objekte in der Westsahara

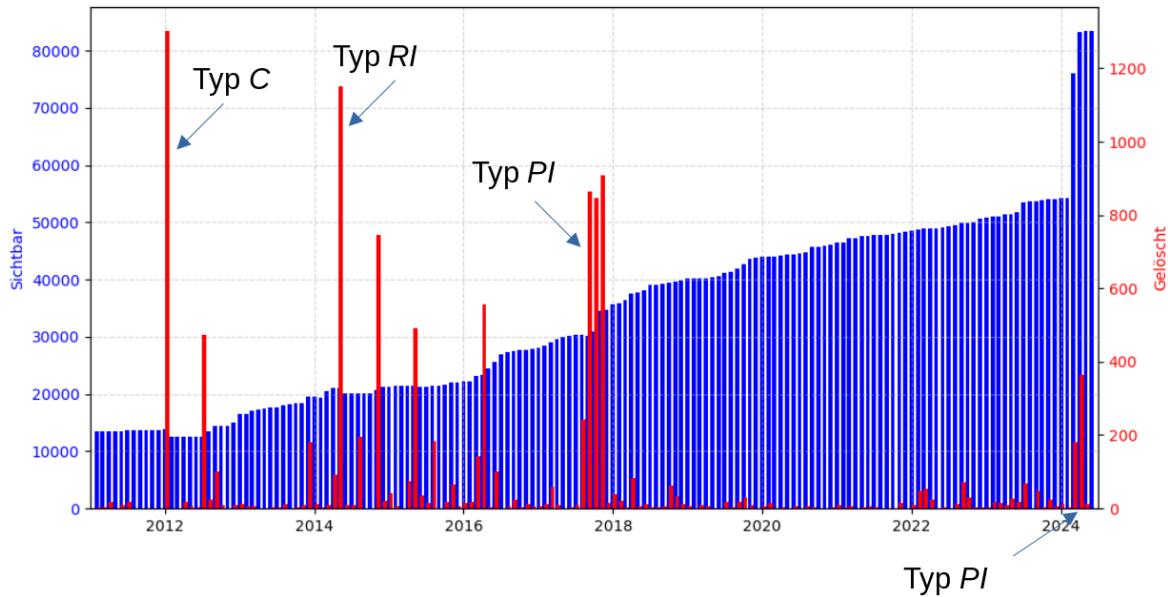


Abbildung 74: Zeitlicher OSM-Verlauf von gelöschten und sichtbaren Objekten in der Westsahara mit Typ-Beschriftung

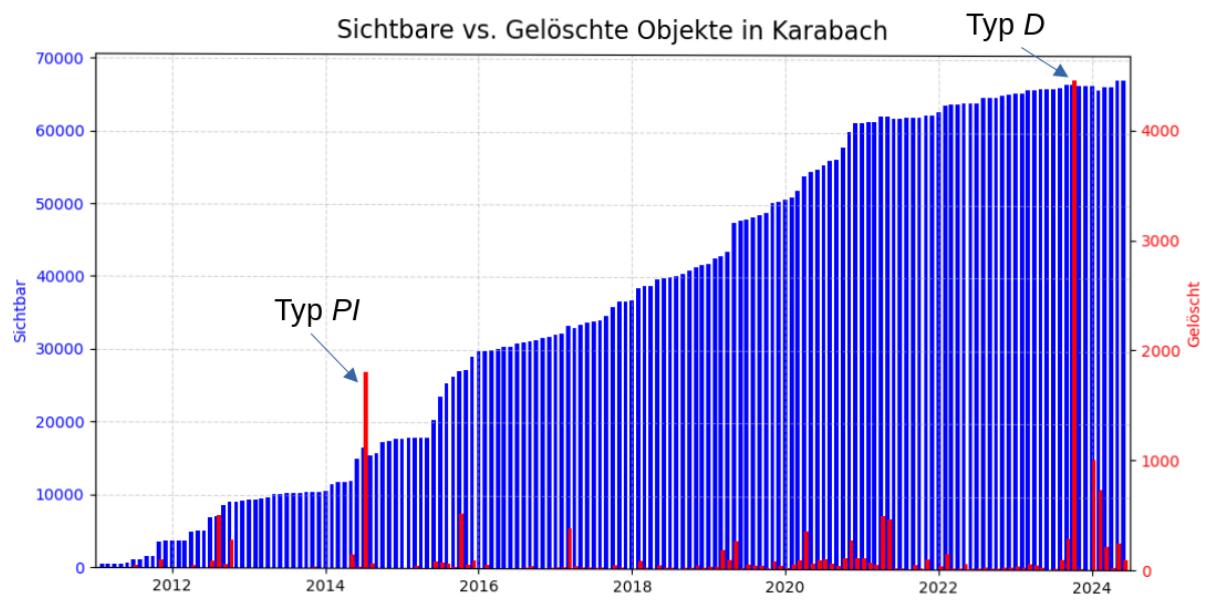


Abbildung 75: Zeitlicher OSM-Verlauf von gelöschten und sichtbaren Objekten in Karabach mit Typ-Beschriftung

6. Diskussion

Die Datenanalyse für die Herausarbeitung grundlegender Charakteristika von OSM-Löschen ist einigen Limitierungen unterworfen, auf die in diesem Kapitel eingegangen wird.

Bei der Analyse der Objektunterschiede wurde u.a. die Gruppe der gegenständlichen Objekte eingeführt. Diese haben das Merkmal, dass sie physisch sichtbare Objekte aus der realen Welt repräsentieren. Eine Ausnahme sind dabei Grenzen (*boundary*), welche häufig nicht physisch sichtbar sind. Sie wurden dennoch dieser Gruppe zugewiesen, weil sie wie alle anderen gegenständlichen Objekte die Eigenschaft haben, dass sie definitiv zum informellen Standard der OSM-Community gehören. Die Gruppenbezeichnung *Gegenständlich* ist deshalb teilweise irreführend und könnte z.B. durch *primäre Objekte* ersetzt werden. Darauf wurde verzichtet, weil dieser Begriff weniger aussagekräftig wäre.

Spannend wäre es, zu erfahren, was die Hauptmotive für das Löschen administrativer Grenzen sind. Damit kann besser beurteilt werden, ob Löschen primär aus Meinungsverschiedenheiten oder aus korrekten Kartenaktualisierungen erfolgen. Demnach wäre das eine wichtige Fragestellung für zukünftige Forschungsarbeiten.

Anhand der Regionen mit besonders vielen Löschen wurden vier verschiedene Löschungstypen abgeleitet. Es wurde allerdings nicht getestet, ob diese Typen für andere Regionen anwendbar sind. Um die eingeführten Typen demnach zu validieren, wäre auch hier eine nachfolgende Forschungsarbeit dazu sinnvoll.

Manche Objekte sind in OpenStreetMap als Relationen gespeichert, die in der Datenanalyse nicht berücksichtigt wurden. Das ist grundsätzlich nicht problematisch, weil solche Objekte zum einen weniger als 1% aller analysierten Daten ausmachen und zum anderen auf Objekte verweisen, die in der Analyse erfasst sind. Bedeutsam wird das jedoch bei der Analyse einzelner gegenständlicher Objekte, welche überproportional häufig als Relation gespeichert sind. Dazu zählen Routing-Objekte (*route*) und Grenzen (*boundary*). Routing-Objekte werden zu über 90% als Relation gespeichert⁶¹ und wurden deshalb bei der Präsentation der Ergebnisse ignoriert, weil das Ergebnis von weniger als 10% solcher Objekte nicht repräsentativ ist. Bei Grenzen ist etwa jedes dritte Objekt als Relation

⁶¹ <https://taginfo.openstreetmap.org/keys/route>

gespeichert.⁶² Dennoch wurden sie in den Ergebnissen aufgeführt, weil zwei Drittel der erfassten Objekte als ausreichend repräsentativ betrachtet werden. Außerdem haben die Ergebnisse zu den Grenzen hohe Aussagekraft, weil diese Objekte am verhältnismäßig häufigsten gelöscht werden. Selbst wenn keine einzige als Relation abgespeicherte Grenze gelöscht worden wäre, betrüge ihr Anteil an Löschungen von der Anzahl an Objekten in der OSM-Karte mehr als 40%. Das wäre damit immer noch der mit Abstand höchste Wert im Vergleich zu allen anderen Objekten.

Einige gelöschte Objekte des zugrundeliegenden Datensatzes wurden bereits vor der Datenanalyse ausgeschlossen, weil sie keine wirklichen Löschungen, sondern Modifizierungen sind (3.6). Eines der Kriterien für den Ausschluss dabei war, dass ein neu erstelltes Objekt mit ähnlicher Geometrie existiert. Um zu bestimmen, ob zwei Geometrien eine ähnliche Geometrie haben, wurde die Hausdorff-Distanz verwendet (3.6.1). Dabei werden die kürzesten Distanzen aller Teilpunkte von Geometrie A zum kürzesten Punkt in Geometrie B abgemessen, dasselbe umgekehrt von Geometrie B zu Geometrie A. Der maximale Wert davon ist definiert als die Hausdorff-Distanz (siehe Formel 1, 2 und 3). In der mathematischen Theorie hat eine Geometrie beliebig viele Teilpunkte. Bei der praktischen Anwendung in der Geoinformatik begrenzen sich die Teilpunkte einer Fläche üblicherweise auf die Stützpunkte der Begrenzungslinie. So wurde in dieser Arbeit die Hausdorff-Distanz mit dem Python-Paket *shapely* berechnet, welche nur die Stützpunkte als Teilpunkte erfasst, was die Berechnung unpräziser macht⁶³. Shapely bietet dabei eine Möglichkeit, die Anzahl an Teilpunkten zu erhöhen, sodass die Berechnung der Hausdorff-Distanz präziser wird (ebd.). Weil damit aber auch der Rechenaufwand steigt, wurde darauf vorerst verzichtet. Für zukünftige Forschungsarbeiten zu OSM-Löschungen wäre das eine sinnvolle Ergänzung.

Ein weiteres Kriterium für den Ausschluss eines gelöschten Objektes von der Analyse war, dass die beigefügten Informationen bzw. Tags *identisch* mit den Tags des neu erstellten Objektes sein müssen (3.6). Der Grundgedanke dabei ist, dass wenn z.B. das gelöschte Objekt eine Bibliothek ist und das neu erstellte Objekt ein Restaurant, das gelöschte Objekt als tatsächliche Löschung erfasst wird, auch wenn sich die Geometrien beider Objekte ähneln sollten. Problematisch ist das, wenn z.B. sowohl das gelöschte Objekt, als auch das neu erstellte Objekt Bibliotheken sind, die neu erstellte Bibliothek aber noch eine Zusatzinformation hat (z.B. den Tag *wheelchair=yes*). In so einem Fall ist das neu erstellte

⁶² <https://taginfo.openstreetmap.org/keys/boundary>

⁶³ https://shapely.readthedocs.io/en/stable/reference/shapely.hausdorff_distance.html

Objekt so ähnlich zu dem gelöschten Objekt, dass man es eher als eine Modifizierung des alten Objektes interpretieren kann. Diese Löschungen sollten idealerweise erfasst und ignoriert werden, was in dieser Arbeit nicht passiert.

Lösbar wäre das Problem, indem gelöschte Objekte erst dann als potentielle Kandidaten für den Ausschluss der Datenanalyse in Betracht kommen, wenn deren beigefügte Tags nicht identisch, sondern *ähnlich* zu den Tags des neu erstellten Objekts sind. Dafür könnte z.B. überprüft werden, ob sowohl Lösung als auch Erstellung dem gleichen gegenständlichen Objekt zugeordnet werden können. Das heißt, dass einer der 29 Keys für gegenständliche Objekte (aufgelistet in 3.5.1) in beiden Objekten vorkommt. Als fortgeschrittene Methode könnte die Levenshtein-Distanz in Betracht gezogen werden, welche zwei Zeichenketten auf ihre Ähnlichkeit überprüft (Po 2020). In dieser Arbeit wurde auf solche Methoden verzichtet, weil sowohl die technische Implementierung als auch die maschinelle Berechnung deutlich aufwändiger wäre. Für zukünftige Forschungsarbeiten zu OSM-Löschen wäre auch das eine sinnvolle Ergänzung.

Würden solche fortgeschrittenen Methoden angewandt und somit auch Löschungen ignoriert werden, deren Informationen sich mit neu erstellten Objekten nur ähneln, statt sich zu gleichen, dann wären vermutlich deutlich mehr Löschungen ignoriert worden als in dieser Arbeit. Damit kann zumindest festgehalten werden, dass mit der hier angewandten Methodik kaum Löschungen ignoriert wurden, die nicht ignoriert werden sollten.

In Südafrika wurden im Jahr 2023 nach den Ergebnissen dieser Arbeit mehr Gebäude gelöscht als auf der sichtbaren OSM-Karte existieren. Der Grund dafür konnte auch nach umfangreicher Nachprüfung nicht herausgefunden werden.

In dieser Arbeit wurden Regionen mit besonders vielen Löschungen detaillierter untersucht, um mehr über die Hintergründe der Löschungen zu erfahren. Die globale Karte zur räumlichen Verteilung der Löschungen zeigt, dass es noch viele weitere Regionen gibt, in denen auffällig viel gelöscht wird (Abbildung 7). Beispielsweise in der Türkei oder im Südwesten der USA. Aus Zeitgründen wurde darauf verzichtet. Dabei wäre es interessant zu erfahren, ob die Löschungen dort ebenfalls überwiegend auf vorangegangene Datenimports zurückzuführen sind und damit die Ergebnisse dieser Arbeit bestätigen. Für zukünftige Arbeiten lohnt sich deshalb eine detaillierte Analyse dieser Regionen.

7. Fazit

Insgesamt bestätigen die Ergebnisse, dass sowohl die Art des Objektes als auch der Ort einen Einfluss auf die Menge an Löschungen haben. Damit wurden ergänzende Hinweise für die OSM-Datenqualität geliefert. Am Beispiel des zeitlichen Zusammenhangs von Löschungen und der militärischen Großoffensive in Karabach wurde offengelegt, dass die OSM-Karte auch Austragungsort politischer Spannungen ist. Die globalen thematischen Karten zur räumlichen Verbreitung von Löschungen ermöglichen weiterführende Forschungsarbeiten, bei denen lokale Auffälligkeiten genauer untersucht werden können, um mehr über die Hintergründe der Löschungen herauszufinden.

Die Kategorisierung der Löschungen in vier verschiedene Löschungstypen liefert einen Überblick über die relevantesten Ursachen von Löschungen und erleichtert es Forschern, ihre eigenen Ergebnisse zu diesem Thema sinnvoll zu interpretieren und einzuordnen.

Die Vergleiche der Löschungstypen machen deutlich, dass Löschungen meistens auf umfangreiche Datenimports folgen und seltener auf Meinungsverschiedenheiten zurückzuführen sind. Deshalb sollten Regionen mit vielen Löschungen nicht automatisch als "Kriegszonen" abgestempelt werden. Schlussendlich kann also festgehalten werden, dass Löschungen mehr von Harmonie als von Streit geprägt sind.