

## Forschungsdatenbank Nichtwohngebäude (ENOB: dataNWG): E.1.4.3 Zwischenbericht Gebäudemerkmale

Hartmann, André; Behnisch, Martin; Hecht, Robert; Meinel, Gotthard;  
Schorcht, Martin; Schwarz, Steffen

Veröffentlichungsversion / Published Version

Zwischenbericht / interim report

### Empfohlene Zitierung / Suggested Citation:

Hartmann, A., Behnisch, M., Hecht, R., Meinel, G., Schorcht, M., & Schwarz, S. (2020). *Forschungsdatenbank Nichtwohngebäude (ENOB: dataNWG): E.1.4.3 Zwischenbericht Gebäudemerkmale*. Dresden: Leibniz-Institut für ökologische Raumentwicklung e.V.. <https://doi.org/10.26084/a7sy-b025>

### Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:  
<https://creativecommons.org/licenses/by/4.0/deed.de>

### Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see:  
<https://creativecommons.org/licenses/by/4.0>



ENOB: dataNWG

## Forschungsdatenbank Nichtwohngebäude

Leibniz-Institut für ökologische Raumentwicklung

Weberplatz 1  
01217 Dresden  
Germany

Tel: +49 (0)351 / 4679-254  
Fax: +49 (0)351 / 4679-212

g.meinel@ioer.de  
www.ioer.de



BERGISCHE  
UNIVERSITÄT  
WUPPERTAL

E.1.4.3

## Zwischenbericht Gebäudemerkmale

Gefördert vom

**Bundesministerium für Wirtschaft und Energie**

im Förderbereich

**Energieoptimierte Gebäude und Quartiere**

der Fördermaßnahme *Anwendungsorientierte nichtnukleare FuE* des 6. Energieforschungsprogramms der Bundesregierung



Bundesministerium  
für Wirtschaft  
und Energie

## Impressum

Projekt	<b>Forschungsdatenbank Nichtwohngebäude. Repräsentative Primärdatenerhebung zur statistisch validen Erfassung und Auswertung der Struktur und der energetischen Qualität des Nichtwohngebäudebestands in Deutschland.</b>
Kurztitel	<b>ENOB:dataNWG</b>
Teilprojekt	Erhebung von Gebäudemerkmalen und Gebäudeklassifikation durch Geodatenanalyse
Gefördert mit Mitteln von	Bundesministerium für Wirtschaft und Energie (BMWi) im Förderschwerpunkt Energieoptimierte Gebäude und Quartiere des 6. Energieforschungsprogramms der Bundesregierung
Förderkennzeichen	03 ET1315 B
Fördermittelnehmer	Leibniz-Institut für ökologische Raumentwicklung
Projektpartner	Institut Wohnen und Umwelt (IWU)  Bergische Universität Wuppertal, Fachgebiet Ökonomie des Planens und Bauens (BUW-ÖPB)
Auftragnehmer	
Bericht	E1.4.3
Verfasser	André Hartmann
Mitarbeit	Martin Behnisch, Robert Hecht, Gotthard Meinel,  Martin Schorcht, Steffen Schwarz
Datum	16. Dezember 2020
ISBN	
Dokument	

## Inhalt

1	Einleitung – Geodaten und Vorgehen .....	1
2	Beschreibung der Eingangsdaten .....	1
2.1	Amtliche Hausumringe (HU-DE) .....	2
2.2	Georeferenzierte Adressdaten (GA) und amtliche Hauskoordinaten (HK-DE) .....	3
2.3	3D-Gebäudemodell LoD1 Deutschland (LoD1-DE) - Grundlegende Informationen.....	5
3	Workflow zur Berechnung der Gebäudemerkmale .....	8
3.1	Datenvorverarbeitung .....	8
3.2	Attributübertragung LoD1-HU.....	10
3.3	Erstellung des Merkmalssets .....	10
3.3.1	Formbezogene Merkmale .....	10
3.3.2	Umgebungsbezogene Merkmale .....	11
3.3.3	Kategorisierung der Gebäudefunktion .....	11
3.4	Berechnung der Merkmalssets.....	14
3.4.1	Skalenebenen .....	14
3.4.2	Programmatische Umsetzung .....	15
4	Ergebnis .....	15
5	Referenzen.....	18

## 1 Einleitung – Geodaten und Vorgehen

Im Projekt ENOB: dataNWG sollen mittels statistischer Hochrechnungen Strukturdaten zum Nichtwohngebäudebestand in Deutschland erhoben werden. Ein wichtiger Schritt dahin ist die Erstellung einer Auswahlgrundlage, aus der Objekte für Vor-Ort Begehungen gezogen werden können (siehe Bericht E.1.4.5 Stichprobe: Modellierung und Ziehung). Die weiteren Schritte – Errechnung von Relevanzwahrscheinlichkeiten mittels binär-logistischer Regression, Erstellung von Erhebungsbezirken, Konzeption des Stichprobenauswahlverfahrens und Erstellung von Begehungsunterlagen – werden in eigenen Berichten dargestellt. Gegenstand dieses Berichts ist die Aufbereitung der Geodaten und die darauffolgende Erstellung der Variablen, welche in das Regressionsverfahren eingehen sollen. Die Variablen sollen eine möglichst gute numerische Charakterisierung der sekundären Erhebungseinheiten liefern, den Hausumringen des amtlichen Liegenschaftskatasters. Angelegt werden diese Merkmale in der Attributtabelle einer Geodatenbank<sup>1</sup>.

Ausgangspunkt für die folgenden Prozessierungen sind die Geobasisdaten des Bundesamtes für Kartografie und Geodäsie (BKG), namentlich die Hausumringe (HU-DE), die Georeferenzierten Adressdaten (GA) und die LoD1 3D-Gebäudemodelle. Jeder dieser Datensätze enthält bereits zum Teil umfangreiche Attribute, von denen einige in das zu bildende Merkmalsset eingehen (siehe Tab. 1)

Notwendig ist daher zunächst die Zusammenführung der verschiedenen Eingangsdaten in einem Geodatensatz.

**Tabelle 1: Übersicht zu den wichtigsten Eingangsdaten.**

	Stichtag	Objektanzahl	Attributan-zahl	Veröffent-licht durch	Datenfor-mat
<b>HU-DE (Hausum-ringe)</b>	01.04.2015	Ca. 50 Mio.	2	BKG	ESRI Shapefiles
<b>GA (Georefe-renzierte Adressda-ten)</b>	01.04.2015	Ca. 20 Mio.	25	BKG	CSV-Tabellen
<b>LoD1 3D-Ge-bäude</b>	Verschieden	Ca. 57 Mio.	22	BKG	XML-Files

## 2 Beschreibung der Eingangsdaten

Im Folgenden werden die oben erwähnten Daten kurz beschrieben. Die Zusammenführung der Daten erfolgt auf Grundlage der Hausumringe, d.h. diese werden der maßgebliche Merkmalsträger für alle folgenden Untersuchungen. Obwohl das

<sup>1</sup> Bei der verwendeten GIS-Software handelte es sich um ArcGIS 10. Die Geodaten wurden in ESRI File Geodatabases gehalten.

ATKIS Basis-DLM im Zuge der Datenaufbereitung zum Einsatz kommt, wird auf eine umfassende Beschreibung an dieser Stelle verzichtet. Das ATKIS Basis-DLM enthält alle topographisch relevanten Objekte. Zur weiteren Vertiefung wird auf die zum Basis-DLM umfangreich vorhandene Literatur verwiesen (AdV 2009, Schumacher 2009).

## 2.1 Amtliche Hausumringe (HU-DE)

Unter der Produktbezeichnung „Hausumringe Deutschland“ (HU-DE) werden seit 2009 mit diesen Daten jährlich die Grundrisspolygone der Gebäude des Amtlichen Liegenschaftskatasters für ganz Deutschland zentral bereitgestellt (Westenberg, 2013). Zusammen mit den Georeferenzierten Adressdaten (GA) sind sie der umfangreichste Geometriedatenbestand zu Gebäuden der Bundesrepublik. Vorteile dieser Datensätze gegenüber kommerziellen oder freien Alternativen sind: einheitliche Datenformate, flächendeckende Verfügbarkeit für Deutschland und eine bestehende gesetzliche Fortführungspflicht.

**Abbildung 1: Beispiel einer Visualisierung der Hausumringe (HU-DE © GeoBasis-DE / BKG 2015)**

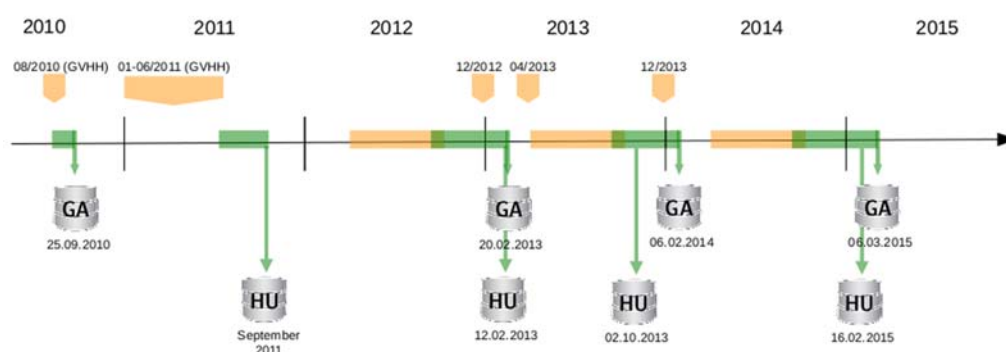


Die HU-DE enthalten flächenhaft modellierte Objekte aus dem Amtlichen Liegenschaftskataster Informationssystem (ALKIS). Diese Umringe stellen die Gebäudegrundfläche der Gebäude bzw. Gebäudeteile dar (siehe Abbildung 1) und enthalten nur die Amtlichen Gemeinde- und Regionalschlüssel (AGS, RS) als Attribute. Die Hausumringe werden durch die Länder jährlich erhoben und aktualisiert. Die Qualität der Geometrien der Hausumringe wurde in Hartmann et al. (2016) untersucht. Grundsätzlich kann von einer Erfassung von mehr als 95 % ausgegangen werden, wobei die fehlenden Umringe meist mit aktueller Bautätigkeit in Verbindung gebracht werden können. In Sachsen und Mecklenburg-Vorpommern gibt es im ländlichen Raum größere Erfassungslücken. Ein Abgleich mit Orthophotos bestätigt eine sehr gute Positionsgenauigkeit der Hausumringe. Im Zuge dieses Projektes wird mit den Hausumringen des Jahres 2015 gearbeitet.

## 2.2 Georeferenzierte Adresdaten (GA) und amtliche Hauskoordinaten (HK-DE)

Seit 2003 werden die Gebäudeadressen des Liegenschaftskatasters als Hauskoordinaten (HK) überregional, durch die Zentrale Stelle Hauskoordinaten und Hausumringe (ZSHH, ehemals GVHH), herausgegeben (Westenberg, 2013). Sie werden als Tabellen im CSV-Format geliefert. Die in dieser Tabelle hinterlegten Werte für Rechts- und Hochwert erlauben die visuelle Darstellung der Lage in einem GIS. Die Ableitung der amtlichen Hauskoordinaten erfolgt für alle Hauptgebäude mit Hausnummern aus ALKIS, d. h. Nebengebäude mit Hausnummer sind in dem Datenbestand nicht enthalten. Die Gebäudezuordnung ist in ALKIS als Attribut abgelegt (Haupt- oder Nebengebäude), wobei Gebäude mit Lagebezeichnung und Hausnummer stets Hauptgebäude, Gebäude mit Lagebezeichnung mit Pseudonummer und laufender Nummer stets Nebengebäude sind. Gebäude, die eine Lagebezeichnung mit Pseudonummer ohne laufende Nummer haben, sind nur in gekennzeichneten Gebäudenutzungsfällen Hauptgebäude. Die Hauskoordinaten sind umfangreich attribuiert, u.a. durch: eine Kennung des Datensatzes, eine eindeutige Nummer, Qualität der georeferenzierten Gebäudeadresse, Schlüssel für Land, Regierungsbezirk, Kreis/kreisfreie Stadt, Gemeinde, Orts- bzw. Gemeindeteile, Straße, Hausnummer, Adressierungszusatz, Koordinatenwerte, Straßename, Postleitzahl, postalischer Ortsname, Zusatz zum postalischen Ortsnamen und Postalischer Ortsteil.

**Abbildung 2: Bearbeitungszeiträume bis zur Veröffentlichung der Geodaten des BKG. Über der Zeitachse ist der Zeitpunkt des Datenbankauszuges angegeben (Hartmann 2015). An die Aufbereitung der HK bei der ZSHH (orange) schließt sich die Weiterbearbeitung durch das BKG (grün) bis zur Veröffentlichung an.**



Die Georeferenzierten Gebäudeadressen (GA) werden vom BKG angeboten. Sie werden aus den Amtlichen Hauskoordinaten (HK-DE) der ZSHH erstellt, deren Stichtag der 01. April eines Jahres ist. Mitte des Jahres erfolgt die Übermittlung der HK-Daten an das BKG, dort werden Gebietsstände vereinheitlicht und Adressen von Drittanbietern (nexiga GmbH, 2014 und Adressangaben der Deutschen Post Direkt GmbH) ergänzt. Diese Bearbeitungen nehmen ca. 4-5 Monate in Anspruch, zwischen dem Datenbankauszug aus ALKIS und der Veröffentlichung der GA vergehen also in der Regel etwa 10 Monate, was bei Verwendung der Daten zu berücksichtigen ist. Ein detaillierter Überblick über die zeitliche Abfolge der Bearbeitung der GA

ist in Abbildung 2 dargestellt. Im Zuge dieses Projektes wird mit den GA des Jahres 2015 gearbeitet.

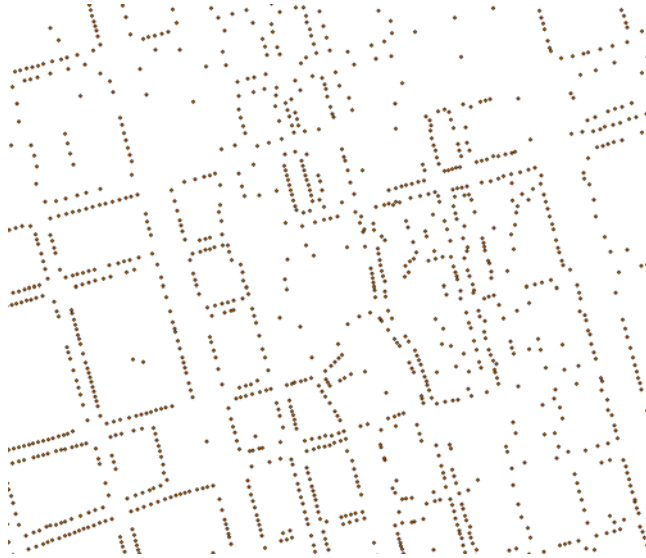
Die HK-DE der Bundesländer wurden in Hartmann (2015) hinsichtlich ihrer Qualität untersucht. Aufgrund ihrer umfangreichen Attribuierung müssen hinsichtlich der Qualitätseinschätzung die Lage und die thematische Genauigkeit betrachtet werden. Da die Hauskoordinaten schon für geplante Bauvorhaben in die Flurstücke gesetzt werden können, d. h. gegenüber den Hausumringen potentiell im Vorlauf sind, kann davon ausgegangen werden, dass es sich um einen sehr aktuellen und vollständigen Datensatz handelt. Ein Qualitätsattribut gibt Auskunft über Lage der Koordinate zum Umring bzw. dem Vorhandensein eines Gebäudes (BKG 2015):

- A – Koordinate im Umring
- B – Koordinate im Flurstück, Gebäude vorhanden
- R – Koordinate im Flurstück, Gebäude nicht sicher vorhanden
- P – Koordinate von Drittanbieter in Umring (nur in den GA des BKG)
- X – Koordinate von Drittanbieter nicht in Umring (nur in den GA des BKG)

Der Anteil der Koordinaten der Qualitätsstufe ‚A‘ kann herangezogen werden, um die Vollständigkeit der Hausumringe abzuschätzen. Dabei wurde berücksichtigt, dass es durch thematische Fehler Koordinaten geben kann, die als ‚A‘-Koordinaten geführt werden, obwohl sie außerhalb des Umrings liegen oder ihnen gar kein Umring zugeordnet werden kann. Untersuchungen von Hartmann (2015) zeigen, dass der Anteil dieser Koordinaten an allen ‚A‘-Koordinaten in den meisten Bundesländern bei unter 0,5 % liegt (Saarland: 2 % und Bremen: 1 %). Die spätere räumliche Verschneidung von HU-DE und GA wird also unter Verwendung einer abstandsbaasierten Verschneidungsregel ablaufen. Auf diese Weise werden auch Adressen zugeordnet, die etwas außerhalb ihres zugehörigen HU liegen.



**Abbildung 1: Beispiel Hauskoordinaten als Punkt Feature-Layer  
(HK-DE © GeoBasis-DE / BKG 2015)**



### **2.3 3D-Gebäudemodell LoD1 Deutschland (LoD1-DE) - Grundlegende Informationen**

3D-Gebäudemodelle sind deutschlandweit erstmals ab 2015 verfügbar und beinhalten neben der Umringgeometrie auch die Höheninformationen der Gebäude. Das ableitbare Gebäudevolumen und die Informationen zur Gebäudefunktion aus dem Liegenschaftskataster bieten in Kombination eine neue Möglichkeit für die Gebäudebestandsanalyse (Jahn et al. 2015). Die Daten können dezentral über die Landesvermessungsämter als auch zentral über die ZSHH und für Bundes- und Forschungseinrichtungen über das BKG bezogen werden. Ein Update erfolgt einmal jährlich, die Abgabe erfolgt bezogen auf Gebiets- und Verwaltungseinheiten im CityGML-Format. Da die Geometrien aus ALKIS abgeleitet sein sollten, kann von derselben Lagegenauigkeit wie für die HU-DE ausgegangen werden. Welche Objekte und Attribute in den Daten modelliert sind, hängt von den länderspezifischen Vermessungsverordnungen und Einmessungspflichten des Katasters ab. Nicht enthalten sind unterirdischen Gebäude.

Laut den Vorgaben der AdV (Arbeitsgemeinschaft der Vermessungsverwaltungen) werden aus ALKIS die folgenden Objektgruppen extrahiert und abgebildet (AdV 2015):

**Tabelle 1: Gebäude- und Bauwerksinformationen in ALKIS**

Kodierung	Bezeichnung
51001	AX_Turm
51002	AX_BauwerkOderAnlageFuerIndustrieUndGewerbe
51006	AX_BauwerkOderAnlageFuerSportFreizeitUndErholung
51007	AX_HistorischesBauwerkOderSonstigeEinrichtung
31001	AX_Gebaeude
51003	AX_VorratsbehaelterSpeicherbauwerk
51009	AX_SonstigesBauwerkOderSonstigeEinrichtung

Die derzeit verfügbare Detailstufe LoD1 (Level of Detail 1), auch als Klötzchenmodell bezeichnet, bildet die einfachste der vier Modellierungsstufen (Gröger et al. 2008). Die Gebäude werden als 3D-Körper (Solid) mit attributierter Höhe dargestellt. Die Dachform ist in dieser Detailstufe nicht modelliert.

Die Qualität der Höheninformation ist von besonderem Interesse, da sie die Grundlage für die Abschätzung von Gebäudevolumina, Geschossflächen und Wohnflächen bilden kann. Die Datenerstellung erfolgt nach Vorgaben der AdV, die in einer Profilbeschreibung (AdV 2016) festgehalten sind. In den Datenbeständen finden sich allerdings dennoch Länderspezifika, welche gegebenenfalls beachtet werden müssen. Die meisten Geometrien werden aus dem Liegenschaftskataster abgeleitet, nur Sachsen und Hessen weichen davon ab (AdV 2019). Die Datenquelle der Lage bestimmt direkt die Ausprägung der Geometrie und ist somit wichtig, um regionale Unterschiede in den Positionsgenauigkeiten zu erklären. Im Kontext von Gebäudevolumenanalysen und damit verbundenen Kennwerten sind die Erfassungsmethoden und die Verfahren zur Ermittlung der mittleren Dachhöhe der „Klötzchen“ relevant. Der Großteil der Geometrien besitzt eine Höheninformation, welche aus Laserscandaten oder automatischer Photogrammetrie stammt. Zu einem geringeren Anteil werden allerdings auch andere Methoden zur Höhenbestimmung verwendet. Die räumliche Verteilung dieser verschiedenen Methoden zeigt sich deutlich zwischen den Bundesländern, aber auch sehr kleinräumig z.B. in einem Stadtteil. Aufgrund der Simplifizierung der Dachform im LoD1-Klötzchenmodell werden Dachformen, von Flachdächern abgesehen, nicht völlig formgetreu und somit auch nicht völlig volumengetreu abgebildet. Die objektspezifischen Höhenwerte werden in den überwiegenden Fällen über einen Mittelwert oder einen Median der vorhandenen Informationspunkte im Objekt gebildet. Auch der höchste gemessene Punkt (First) oder der niedrigste (Traufe) können Verwendung finden. Die entsprechenden Informationen sind in den Metadaten hinterlegt und müssen in die Analysen entsprechend einfließen. Darüber hinaus wird die abgebildete Höheninformation von der Genauigkeit des zugrundeliegenden 3D-Geländemodells beeinflusst, wobei die verwendeten Datengrundlagen nach den Bundesländern stark variieren.

Ein weiteres Qualitätskriterium ist die Aussagefähigkeit und Vollständigkeit der Nutzungsinformation (Gebäudefunktion). Wünschenswert wären möglichst einheitliche und detaillierte Bezeichnungen für die Gebäudefunktionen. In einer semantischen Auswertung der Gebäudefunktion konnten aber auch Klassen wie „Gebäude allgemein“, oder „nach Quellenlage nicht zu spezifizieren“ identifiziert werden. Die folgende Tabelle 3 zeigt das Ergebnis einer semantischen Auswertung aller Gebäudefunktionen in Form des Anteils der verwertbaren Gebäudefunktion nach Bundesländern. Länder wie Baden-Württemberg, Berlin, Hamburg, Hessen, Mecklenburg-Vorpommern, Niedersachsen, Nordrhein-Westfalen, Rheinland-Pfalz, Saarland, Sachsen-Anhalt, aber auch Brandenburg und Schleswig-Holstein weisen eine sehr gute bzw. noch gute Vollständigkeit der Nutzungsinformation auf.

**Tabelle 2: Anteil von 3D-Gebäudemodellen mit Nutzungsinformationen differenziert nach Bundesländern (Stand 2016)**

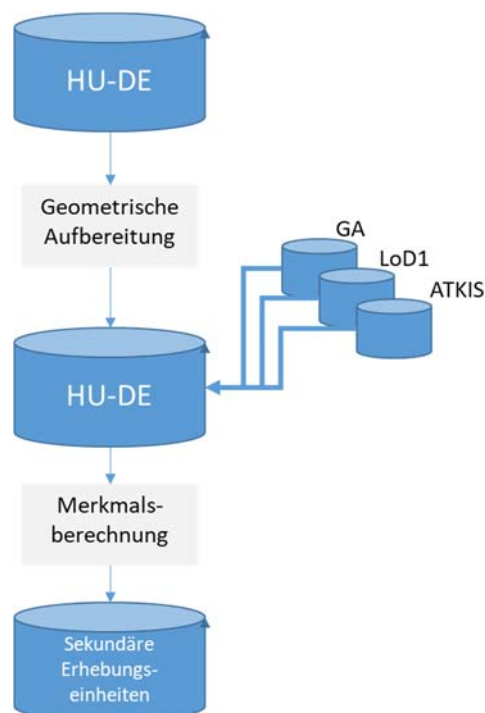
Bundesland	Prozentualer Anteil der Geometrien mit Nutzungsinformation [%]
BW	100,0
BY	52,0
BE	100,0
BB	89,0
HB	0,0
HH	99,9
HE	99,0
MV	99,8
NI	100,0
NW	96,8
RP	100,0
SL	99,9
SN	28,6
ST	99,9
SH	88,0
TH	57,0

In den Bundesländern Bayern, Bremen, Sachsen und Thüringen stellt sich die Datenlage dagegen erheblich schlechter dar. In Bayern sind zwar einige Nutzungsklassen vergeben, der überwiegende Anteil der Polygone trägt jedoch nur die Information Hauptgebäude oder Nebengebäude. In Bremen ist in den vorliegenden Daten nur die Information „Gebäude allgemein“ hinterlegt. Speziell in Sachsen, aber auch in Thüringen sind aktuell noch erhebliche Datenlücken bei der Nutzungsinformation zu verzeichnen. Es ist allerdings davon auszugehen, dass sich die Qualität in den kommenden Jahren steigern wird.

### 3 Workflow zur Berechnung der Gebäudemerkmale

Die Gebäudegeometrie ist sowohl mit den HU-DE als auch den LoD1-DE gegeben, wobei beide in verschiedenem Umfang attribuiert sind. Aus der Dokumentation der LoD1-Daten gingen für die Bundesländer uneinheitliche Datenstichtage hervor. Dies war einer der Gründe, die HU-DE mit ihrem festen Datenstand als geometrische Grundlage der Datenbank zu wählen. Da die HU-DE aber keine weiteren Attribute außer der Gemeindegemeinschaften enthalten, bestand der nächste Schritt in der Übertragung der Attributdaten der LoD1 auf die HU-DE. Dazu wurden die Polygone beider Datensätze räumlich verschnitten und Attribute wie u.a. Gebäudehöhe und Funktion übertragen. Daran anschließend erfolgt die Berechnung der weiteren Merkmale in der Datenbank. Es ergibt sich der nachfolgend dargestellte Workflow (Abbildung 5).

**Abbildung 5: Darstellung des gesamten Workflows. Am Ende der Merkmalsberechnung liegen die HU-DE als sekundäre Erhebungseinheiten vor.**



#### 3.1 Datenvorverarbeitung

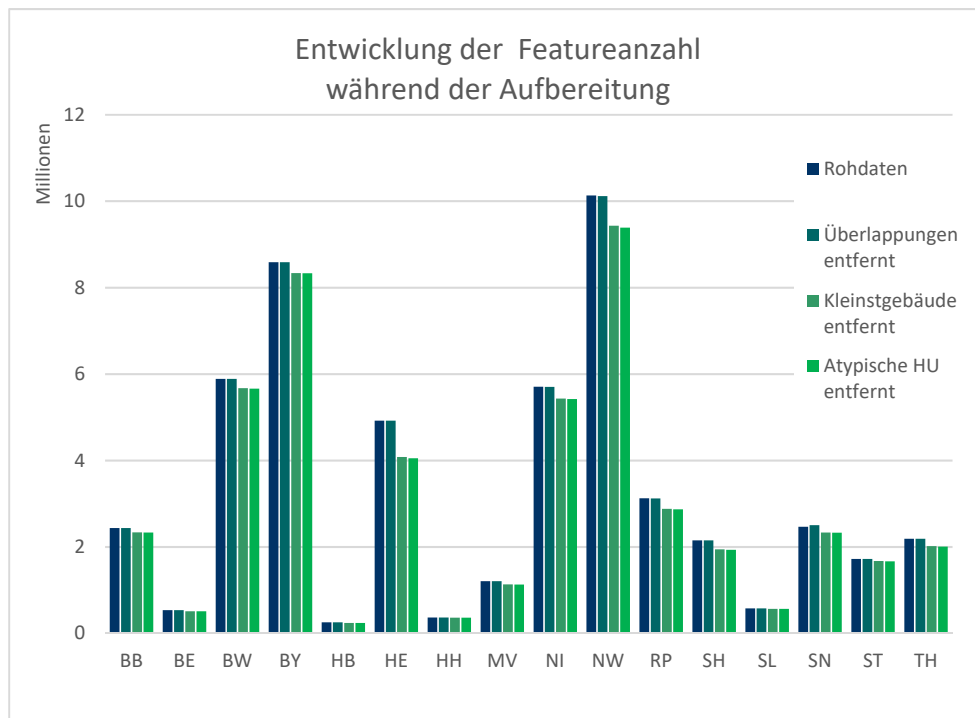
Der Datenverschneidung und Merkmalsberechnung vorgeschaltet ist eine Datenvorverarbeitung. Dies betrifft die Beseitigung von Überlappungen im Datensatz sowie von Umringen, die kleiner sind als der Einmessungsschwellwert von 10m<sup>2</sup>. Auch Umringe, die eine stark atypische Form und gleichzeitig eine Größe < 190 m<sup>2</sup> aufwiesen, wurden durch die Vereinigung mit benachbarten Umringen beseitigt (Hartmann et al. (2016)).

**Abbildung 6: Der Datensatz der HU-DE vor (links) und nach den Aufbereitungsschritten (rechts).**



Die Beseitigung der Überlappungen erfolgte durch eine *Union*-Funktion, d.h. die Bildung aller Schnittflächen zwischen den Geometrien. Im Überlagerungsfall entstehen auf diese Weise identische Schnittflächen, von denen alle bis auf eine entfernt werden. Diese verbliebene Schnittfläche wird mit dem Nachbarpolygon mit der längsten gemeinsamen Kante zusammengefasst. Nachdem diese geometrischen Änderungen vorgenommen sind, liegt ein überlappungsfreier Datensatz für die Weiterverarbeitung vor. Wie sich diese Vorbereitung auf die Daten auswirkt, ist beispielhaft in der Abbildung 6 gezeigt.

**Abbildung 7: Entwicklung der Feature-Anzahl nach den einzelnen Aufbereitungsschritten.**



Die Anzahl der Features in den HU wird auf diese Weise erwartungsgemäß etwas reduziert (Abbildung 7).

## 3.2 Attributübertragung LoD1-HU

Nach den vorbereitenden Arbeiten kann die Zusammenführung der Geodaten beginnen. Dabei muss beachtet werden, dass die Bestände unterschiedliche Stichtage haben. Es kann nicht davon ausgegangen werden, dass sich in jedem Fall eine volle Überlagerung und Verschneidung der Datensätze ergibt.

Bezogen auf die Vollständigkeit der Grundrissdaten wurden die HU-DE als Referenz betrachtet. Mit dieser Annahme wurden die HU-DE über die *SpatialJoin\_analysis*-Funktion in ArcGIS mit den LoD1 überlagert und die Merkmale an erstere übertragen. Die räumlichen Beziehungen dieser Überlagerung können durchaus komplex sein. Zu den zeitlich bedingten Unterschieden der Vollständigkeit kommen solche, die auf verschiedene Modellierungsarten zurückgehen. Generell wurde so verfahren, dass einem HU-Objekt nur ein Merkmal zugewiesen wird, wenn es von mindestens einem LoD1-Gebäude überlagert wird (mindestens also 1:1-Beziehung). Des Weiteren wird ein LoD1-Objekt ohne HU-Entsprechung nicht einfach in den Datensatz übernommen, da die Grundriss-Referenz der HU-DE nicht geändert werden sollte. Die Höhe wurde im Fall einer 1:n-Zuordnung zwischen HU und LoD1 über eine *MergeRule*-Einstellung als Median übergeben, die Einstellung für semantische Attribute in diesem Fall war der Modalwert. Die auf diese Weise angefügten Attribute stellen die Grundlage für das im Folgenden beschriebene Merkmalsset dar.

Über eine räumliche Verschneidung wurden auch die Anzahlen der Adressen pro Hausumring als Attribut angefügt sowie die Nutzungsart (d.h. Wohnnutzung, Mischnutzung, Industrie- und Gewerbe oder Verkehr) des Baublockes aus dem ATKIS-Basis-DLM, in welchem sich die Hausumringe befinden (Abbildung 7).

## 3.3 Erstellung des Merkmalssets

Nachdem die HU um die Attribute aus LoD1, GA und ATKIS angereichert wurden, erfolgt die Berechnung der Merkmale. Im Folgenden werden die Merkmale kurz beschrieben sowie die Annahmen zu ihrer Aussagekraft erläutert. In Anlehnung an Hecht (2014) können die Merkmale eingeteilt werden in solche, die

- Form und Größe (Morphologie)
- Umgebung/Nachbarschaft (Kontext) oder die
- Nutzungsinformationen (Semantik)

des Umrings beschreiben. Dies kann wiederum verschiedene Maßstabsebenen betreffen. So ist die Form eher auf den Umring selbst und evtl. eine geometrisch zusammenhängende Region von Umringen (Umringregion) bezogen. Die Nachbarschaft bzw. Umgebung des Umrings kann den Umring, die Umringregion bis hin zum Baublock umfassen. Derartige Merkmalssets sind bereits seit mehreren Jahren Gegenstand der Geoinformatik. Eine sehr ausführliche Darstellung zu Merkmalssets in der Gebäudeanalyse findet sich in Hecht (2014).

### 3.3.1 Formbezogene Merkmale

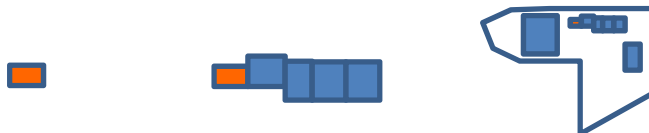
Unter die formbeschreibenden Merkmale fallen u.a. jene, die eine Beziehung zwischen den geometrischen Parametern Umfang und Fläche des Umringspolygons herstellen. Ein Wert für die Fläche allein lässt keine Aussage über die Form einer geometrischen Figur zu. Durch die Verbindung mit anderen geometrischen Parametern

in einem Index lässt sich aber z.B. einschätzen, wie sehr eine Form einem Kreis, einer Ellipse o.ä. ähnelt. Diese Indizes dienen als Kompaktheitsmaße und erlauben mittelbar eine Aussage über die Komplexität des Grundrisses. Es wird vermutet, dass im Bereich der Wohngebäude eher einfache Formen vorherrschen und sich komplexere, verwinkelte Grundrisse eher im Bereich der Nichtwohngebäude finden lassen. In der Literatur (u.a. Maceachren 1985, Steiniger et al. 2008, Wurm et al. 2016) werden eine Reihe Kompaktheitsmaße für verschiedene Anwendungsbereiche diskutiert.

### 3.3.2 Umgebungsbezogene Merkmale

Wie aus Abbildung 8 ersichtlich ist, lassen sich auf den gezeigten Skalenebenen verschiedene Ausprägungen der Merkmale erwarten. Die Betrachtung der über den einzelnen HU hinausgehenden Nachbarschaft (als Umringregion oder als weiteres Gefüge im Baublock) erweitert die rein formbezogenen Merkmale um solche, die Aussagen zu Homogenität und Varianzen gestatten. So ist zu erwarten, dass innerhalb eines geschlossenen Wohnblocks die darin enthaltenen Hausumringe eher ähnliche geometrische Merkmale haben. Auch über den Baublock betrachtet sollte das Gefüge der Gebäude in diesem Fall eher homogen sein. Bei ausgeprägter industrieller oder gewerblicher Nutzung kann hingegen eine Vielzahl an speziell genutzten Objekten (z.B. Heizhaus, Werkshalle, Bürotrakt) mit weit streuenden geometrischen Merkmalen angenommen werden.

**Abbildung 8: Maßstabsebenen für Merkmale: Umring, Umringregion, Baublock (v.l.n.r.)**



### 3.3.3 Kategorisierung der Gebädefunktion

Die aufbereiteten HU-DE tragen die aus den LoD1-Daten überführte Nutzungsinformation pro Umring. Diese Nutzungsinformation ist aus dem Liegenschaftskatasterinformationssystem (ALKIS) abgeleitet. Mit dem ALKIS-Objektartenkatalog in Version 6.0 der AdV (AdV 2008) liegt eine einheitliche Definition dieser Nutzungskategorien vor. Aufgrund der Neuheit des Produktes LoD1 und der Umstellung auf ALKIS in den vergangenen Jahren finden sich dennoch einige Länderspezifika mit

Auswirkungen auf die Nutzungsinformation. Der Katalog enthält mehrere hundert Gebäudenutzungstypen.

**Abbildung 9: Gebäudeklassifikation mit energetischer Einordnung (BBSR 2013)**



Es ergeben sich für die Identifikation von Nichtwohngebäuden relevante und irrelevante Nutzungstypen. Abbildung 9 gibt eine grundlegende Übersicht über die Nutzungseinstufung (BBSR 2013). Hinsichtlich der Gebäudeklassifikation ist ebenfalls entscheidend, ob die Nichtwohngebäude im Rahmen der Energieeinsparverordnung (EnEV) relevant sind. EnEV-relevante Nichtwohngebäude (NWG) sind in den Kategorien Bildung, Büro und Verwaltung, Gewerbe und Industrie, Heilbehandlung, Handel und Dienstleistung, Sport, Kultur, Beherbergung und Gastronomie zu finden. Nicht von Bedeutung sind in diesem Zusammenhang u.a. Gebäude für Land- und Forstwirtschaft, Gebäude der technischen Erschließung/Infrastruktur, Gebäude der verkehrlichen Erschließung/Infrastruktur, des Weiteren sonstige NWG wie Garagen sowie Wohngebäude, also Gebäude mit überwiegender wohnräumlicher Nutzung.

Um für die binär-logistische Regression eine verwertbare Datengrundlage zu schaffen, mussten die Gebäudeklassen einzeln eingestuft werden.<sup>2</sup> Damit alle Nutzungsklassen beachtet werden können, wurden alle in den LoD1-Daten vorkommenden Nutzungsinformationen redundanzfrei ausgelesen. Anschließend erfolgte die typologische Einstufung der Nutzungsklassen durch das gesamte Forschungskonsortium. Für die Einstufung wurden die folgenden Kategorien gebildet.

<sup>2</sup> Das mit dieser Kategorienbildung verfolgte Ziel wird im Bericht „E.1.4.5 Stichprobe: Modellierung und Ziehung“ erläutert.



**Tabelle 4: Numerische Kodierung der Oberkategorien für die Nutzungsarten**

1	zu EnEV-relevantem Nichtwohngebäude gehörend
2	zu EnEV-irrelevantem Nichtwohngebäude gehörend (inkl. Garagen)
3	Zu Nichtwohngebäude gehörend, aber unklar, ob EnEV-relevant
0	Zu Wohngebäude gehörend
-1	Zu keinem Gebäude gehörend (z.B. Bauwerke)
-2 <sup>3</sup>	aus konzeptionellen Gründen gestrichen
9	Zu Gebäude gehörend, aber unklar, ob NWG, Keine Einstufung möglich

Die Kategorien 1, 2 und 3 klassifizieren jeweils NWG, wobei Kategorie 2 eindeutig nicht EnEV-relevante Gebäude beinhalten soll, welche für die Untersuchung nahezu keine Bedeutung haben. Eindeutig EnEV-relevante Gebäude wie Schulen, Polizeiämter, Krankenhäuser oder Gerichte finden sich in Kategorie 1, hier lassen die Geodaten eine vergleichsweise sichere Einordnung zu. Gebäude in Kategorie 3 wie kirchliche Einrichtungen (ohne Sakralbauten), Werkstätten oder Lagerhallen können nicht ohne weiteres als EnEV-relevant identifiziert oder abgelehnt werden.

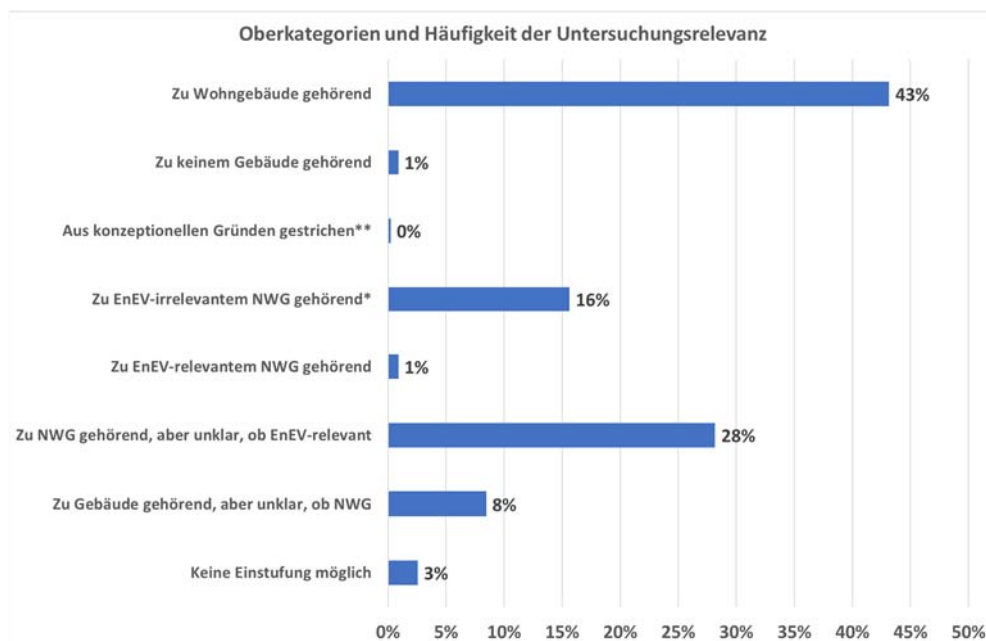
Wenn der überwiegende Teil (>50%) des Gebäudes der Wohnnutzung dient, so ist es kein NWG, sondern ein Wohngebäude. Kategorie 0 kennzeichnet Wohngebäude, zu denen Einzelhäuser, aber auch Reihenhäuser gehören. Bauwerke wie Wasserbehälter, Öltanks oder Funkmasten (-1) erfüllen dagegen noch nicht einmal die Gebäudeeigenschaft. Ebenfalls finden u.a. Kioske, Wochenendhäuser oder Schutzhütten aus konzeptionellen Gründen keine Beachtung. Objekte der Kategorie -2 wurden nur für die Pilotphasen-Bundesländer Nordrhein-Westfalen und Thüringen gebildet.

In den Geodaten sind auch Nutzungsinformationen wie Gebäude (allgemein), Kinderheim oder Forsthaus hinterlegt. Hier ist nicht eindeutig erkennbar, ob eine Nichtwohngebäudenutzung dominant ist. Diese Objekte befinden sich in Kategorie 9.

Garagen wären zwar prinzipiell als NWG zu betrachten, sind aber im Rahmen der Studie aufgrund ihrer Größe und ihres Energieverbrauchs nicht relevant. In einigen Bundesländern (z.B. Baden-Württemberg) ist die Attribuierung dieser Nutzungsinformation überdurchschnittlich genau, in anderen hingegen nicht. Garagen weisen jedoch grundsätzlich charakteristische Formfaktoren auf. Zur besseren Nachvollziehbarkeit und Validierung der Geodaten wurden diese Objekte zunächst in die Kategorie 98 eingeordnet, in Abbildung 10 aber der Kategorie 2 zugeschlagen. Im Zuge der Auswertung bleiben somit Rückschlüsse hinsichtlich der Attribuierungsgenauigkeit für Garagen in den Geodaten erwartet.

<sup>3</sup> Nur für die Pilotphasen-Bundesländer Nordrhein-Westfalen und Thüringen gebildet.

**Abbildung 10: Oberkategorien und Häufigkeitsverteilung der Untersuchungsrelevanz in den LoD1-Objekten.**



Anhand der LoD1-Geometrien wurde zunächst ein Verteilungsmuster (Gesamtdeutschland) erstellt, siehe Abbildung 10. Die relativen Ausprägungen in den Klassen lassen sich auch auf den schlussendlichen HU-DE basierten Datensatz übertragen.

Den größten Anteil bilden mit ca. 43% Wohngebäude. Die weiteren nicht relevanten Klassen 2, -1, -2 und 98 kommen zusammengefasst auf einen Anteil von ca. 16%. Die (unter Umständen relevanten) Klassen 1, 3 und 9 kommen zusammen auf ca. 40% Gesamtanteil. Insgesamt konnten nur sehr wenige Umringe (Klasse 1) als eindeutig zu einem EnEV-relevanten NWG gehörend identifiziert werden, ihr Anteil liegt bei nur ca. 1%.

### 3.4 Berechnung der Merkmalssets

Die Errechnung der Merkmale erfolgt nach Abschluss der geometrischen Aufbereitung und Datenanreicherung der Umringspolygone. Zum einen soll im Anschluss ein aussagekräftiges Set an Variablen für die binär-logistische Regression vorliegen, zum anderen sollte die Berechnung der Merkmale, in Anbetracht der Objektanzahl, in einem vertretbaren zeitlichen Rahmen erfolgen können.

#### 3.4.1 Skalenebenen

Wie bereits erwähnt, beziehen sich einige der Merkmale nicht nur direkt auf ein Umringobjekt, sondern auf eine höhere morphologische Skalenebene, der das Objekt zugehört (vgl. Abb. 8). So werden auch Merkmale für die Umringregion oder den Baublock, in dem der Umring liegt, gebildet. Die entsprechenden Skalenebenen müssen ebenfalls als Geodaten vorliegen bzw. zum Teil aus den Umringspolygonen erstellt werden.

Die Umringregionen werden durch eine *Dissolve*-Funktion als geometrische Verschmelzung direkt aneinandergrenzender Umringe gebildet. Aus einzelnen gereihten Hausumringen eines Straßenblockes entsteht also ein Umring für den ganzen Block. Die innerhalb dieses Blockes liegenden Umringe erhalten die ID des Blockes, was die statistische Auswertung von Merkmalen für die Umringregion ermöglicht.

Die Baublöcke sind (siehe 3.2) dem ATKIS-Basis-DLM entnommen, da schon bei der vorbereitenden Datenfusion die IDs dieser Blöcke zusammen mit der Nutzungsart aus ATKIS an die HU-DE übergeben wurden.

### 3.4.2 Umsetzung

Die Berechnung der meisten Merkmale erfolgte direkt aus den Features in der Geodatenbank. Die Prozessierung lief pro Bundesland und über eine teilweise parallelisierte Ausführung von Python-Skripten ab. In ArcGIS kann nicht parallel auf eine Geodatenbank zugegriffen werden, d.h. für die Nutzung paralleler Prozesse müssen die Daten auf verschiedene Geodatenbanken aufgeteilt werden. Die Aufteilung nach Bundesländern erwies sich in diesem Zusammenhang als praktikabel. Alle Merkmalsberechnungen konnten über die ArcGIS-Python-Schnittstelle realisiert werden.

## 4 Ergebnis

Das Ergebnis der oben beschriebenen Verfahren ist ein um ca. 40 numerische Merkmale angereicherter Geodatenbestand, der ca. 49 Mio. Objekte enthält. Die Datenmenge in der Geodatenbank umfasst etwa 19 GB. Die folgende Tabelle gibt einen Überblick über die errechneten Merkmale und ihre Definition.

**Tabelle 5: Errechnete Merkmale für die HU-DE. Merkmalstypen sind Morphologie (M), Kontext (K), Semantik (S).**

Attribut	Definition	Merkmalstyp
Shape_Area	Flächeninhalt eines einzelnen Umrings [m <sup>2</sup> ]	M
Shape_Length	Umfang eines einzelnen Umrings [m]	M
N_NODES	Anzahl Stützpunkte eines Umrings	M
AREAREG	Flächeninhalt zusammenhängender Umringe (Umringregion) [m <sup>2</sup> ]	K
PERIREG	Umfang Umringregion [m]	K
AREARECT	Fläche des schmalsten umschreibenden Rechtecks (Minimum Bounding Rectangle - MBR) [m <sup>2</sup> ]	M
MINOCIRC	Durchmesser des kleinsten umschreibenden Kreises [m]	M
RECTLENG	Länge des MBR [m]	M

<b>RECTWID</b>	Breite des MBR [m]	M
<b>AREABLK</b>	Flächeninhalt des Baublockes [m <sup>2</sup> ]	K
<b>PERIBLK</b>	Umfang des Baublockes [m]	K
<b>MINBLDIST</b>	Minimaler Abstand Umring zu Baublockgrenze [m]	K
<b>MINDISTREG</b>	Minimaler Abstand Umring zu Umring [m]	K
<b>UFE</b>	Umfang-Flächen-Verhältnis des Umrings	M
<b>RATIOAREA</b>	Verhältnis der Umringfläche zu Fläche des MBR	M
<b>Shape_Idx</b>	Kreisähnlichkeit (Circularity) [0...1]	M
<b>RECTRATIO</b>	Längen-/Breiten-Verhältnis des MBR	M
<b>SCHUMM</b>	Verhältnis der Durchmesser des flächengleichen Kreises und des kleinsten umschreibenden Kreises	M
<b>SHAPER</b>	Shapeindex der Umringregion [0...1]	K
<b>PATIO</b>	Anzahl der Löcher in Umringregion	K
<b>N_NBR</b>	Anzahl der benachbarten Umringe	K
<b>ADDNUMB</b>	Anzahl der Hauskoordinaten des Umrings	S
<b>ORIENT</b>	Orientierung der längsten Seite des MBR	M
<b>VOLGEB</b>	Volumen des Umrings [m <sup>3</sup> ]	M
<b>VOLREG</b>	Volumen der Umringregion [m <sup>3</sup> ]	K
<b>MIN_AREA</b>	Minimale Fläche innerhalb der Umringregion [m <sup>2</sup> ]	K
<b>MAX_AREA</b>	Maximale Fläche innerhalb der Umringregion [m <sup>2</sup> ]	K
<b>MEAN_AREA</b>	Mittelwert der Flächen der Umringregion [m <sup>2</sup> ]	K
<b>STD_AREA</b>	Standardabweichung der Flächen der Umringregion [m <sup>2</sup> ]	K
<b>MIN_VOL</b>	Minimales Volumen innerhalb der Umringregion [m <sup>3</sup> ]	K

<b>MAX_VOL</b>	Maximales Volumen innerhalb der Umringregion [m <sup>3</sup> ]	K
<b>MEAN_VOL</b>	Mittelwert der Volumina innerhalb der Umringregion [m <sup>3</sup> ]	K
<b>STD_VOL</b>	Standardabweichung der Volumen der Umringregion [m <sup>3</sup> ]	K
<b>MIN_Z</b>	Minimale Höhe innerhalb der Umringregion [m]	K
<b>MAX_Z</b>	Maximale Höhe innerhalb der Umringregion [m]	K
<b>MEAN_Z</b>	Mittelwert der Höhen innerhalb der Umringregion [m]	K
<b>STD_Z</b>	Standardabweichung der Höhe der Umringregion [m]	K
<b>SUMABUBL</b>	Summe der Umringfläche im Baublock [m <sup>2</sup> ]	K
<b>MEAN_AREAH</b>	Mittlerer Flächeninhalt der Umringe im Baublock [m <sup>2</sup> ]	K
<b>STD_AREAH</b>	Standardabweichung der Umringflächen im Baublock [m <sup>2</sup> ]	K
<b>POPDENS</b>	Bevölkerung im 100m Raster (Zensus 2011, DESTATIS)	S

## 5 Referenzen

AdV (2008): Dokumentation zur Modellierung der Geoinformation des amtlichen Vermessungswesens (GeoInfodok) – ALKIS-Objektartenkatalog. Version 6.0. Stand: 11.04.2008. Online: <http://www.adv-online.de/GeoInfoDok/GeoInfoDok-6.0/broker.jsp?uMen=d3b70780-c5f2-bc61-f27f-31c403b36c4c> (02.11.20)

AdV (2009): Dokumentation zur Modellierung der Geoinformation des amtlichen Vermessungswesens (GeoInfodok)- Hauptdokument. Version 6.0.1. Stand 01.07.2009. Online: <http://www.adv-online.de/GeoInfoDok/GeoInfoDok-6.0/binarywriterservlet?imgUid=8f830072-8de8-9221-d5ad-8f138a438ad1&uBasVariant=11111111-1111-1111-1111-111111111111> (02.11.20)

AdV (2015): Datenformatbeschreibung 3D-Gebäudemodell LoD1 Deutschland (LoD1-DE). GeoBasis-DE – Hauskoordinaten und Hausumringe der deutschen Liegenschaftenkatasters. Bezirksregierung Köln.

AdV (2016): AdV-CityGML-Profile für 3D-Gebäudemodelle. Online: <http://www.adv-online.de/AdV-Produkte/Weitere-Produkte/3D-Gebaeudemodelle-LoD/binarywriterservlet?imgUid=f760754e-b81c-c514-611d-1565d1cbf297&uBasVariant=11111111-1111-1111-1111-111111111111> (02.11.2020)

AdV (2019): Produktblatt 3D-Gebäudemodelle LoD1. Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland. Online: <http://www.adv-online.de/AdV-Produkte/Standards-und-Produktblaetter/Produktblaetter/binarywriterservlet?imgUid=a432cda7-7ad9-f61f-e302-c303b36c4c2e&uBasVariant=11111111-1111-1111-1111-111111111111> (02.11.2020)

BBSR (2013): Deilmann, C.; Behnisch, M.; Dirlich, S.; Gruhler, K.; Hagemann, U.; Peterreit, R.; Kunz, C.; Peterreit, K.: Systematische Datenanalyse im Bereich der Nichtwohngebäude - Erfassung und Quantifizierung von Energieeinspar- und CO<sub>2</sub>-Minderungspotenzialen. Berlin: BMVBS, 2013 (BMVBS-Online-Publikation; 27/2013).

BKG (2015): Georeferenzierte Adressdaten – GA. Dokumentation zur Lieferung. Bundesamt für Kartographie und Geodäsie. Stand 25.11.2015.

Gröger, G.; Kolbe, T. H.; Czerwinski, A.; Nagel, C. (2008): OpenGIS city geography markup language (CityGML) encoding standard. Open Geospatial Consortium Inc.(2008).

Hartmann, A. (2015): Analyse der Gebäudebestandsentwicklung in Deutschland aus Grundriss-bezogenen amtlichen Geobasisdaten, Diplomarbeit, TU-Dresden.

Hartmann, A.; Meinel, G.; Hecht, R.; Behnisch, M. (2016): A Workflow for Automatic Quantification of Structure and Dynamic of the German Building Stock Using Official Spatial Data. In: ISPRS International Journal of Geo-Information 5(8), 142.

Hecht, R. (2014): Automatische Klassifizierung von Gebäudegrundrissen - Ein Beitrag zur kleinräumigen Beschreibung der Siedlungsstruktur. Berlin: Rhombos-Verlag, IÖR-Schriften 63.

Jahn, A.; Hecht, R.; Meinel, G. (2015): 3D-Gebäudemodelle - Grundlage siedlungsstruktureller Analysen am Beispiel Sachsens. In: Meinel, G.; Schumacher, U.;

Behnisch, M.; Krüger, T. (Hrsg.): Flächennutzungsmonitoring VII. Boden – Flächenmanagement – Analysen und Szenarien. Berlin: Rhombos-Verlag, 2015, IÖR-Schriften 67, S. 233-243.

Maceachren, A. (1985). Compactness of geographic shape: comparison and evaluation of measures. *Geografiska Annaler, Series B*, 67(1), S. 53-67.

Meinel, G.; Hecht, R.; Herold, H. (2009): Analyzing building stock using topographic maps and GIS. In: *Building Research & Information* 37 (5-6), S. 468–482.

Meinel, G.; Hecht, R.; Herold, H.; Schiller, G. (2008): Automatische Ableitung von stadtstrukturellen Grundlagendaten und Integration in einem Geographischen Informationssystem. Bonn: BBR, Forschungen / BBR 134.

Schumacher, U. (2009): ATKIS-ALK(IS)-Orthobild – Vergleich von Datengrundlagen eines Flächenmonitorings. In: Meinel, G.; Schumacher, U. (Hrsg.): Flächennutzungsmonitoring – Konzepte, Indikatoren, Statistik. Aachen. S. 47-67.

Steiniger, S.; Lange, T.; Burghardt, D.; Weibel, R. (2008): An Approach for the Classification of Urban Building Structures Based on Discriminant Analysis Techniques. *Transactions in GIS*, 12 (1), S. 31-59.

Westenberg, G.; Will, K. (2013): Geometrieinformationen zum Gebäudebestand – die Produkte Hauskoordinaten, Hausumringe und 3D-Gebäudemodelle. In: Meinel, G.; Schumacher, U.; Behnisch, M. (Hrsg.): Flächennutzungsmonitoring V. Methodik – Analyseergebnisse – Flächenmanagement. Berlin: Rhombos – Verlag, 2013, IÖR Schriften 61, S. 147-154.