

55

André Müller

**Methodik zur datenbasierten
Typisierung von Quartieren
anhand baulicher Strukturen**

Geometrisch-topologische Grundlagen für
die energetische Bewertung des nationalen
Quartiersbestandes auf Basis von
Geoinformationsdaten

DISSERTATION

Methodik zur datenbasierten Typisierung von Quartieren anhand baulicher Strukturen

Geometrisch-topologische Grundlagen für die energetische Bewertung des nationalen Quartiersbestandes auf Basis von Geoinformationsdaten

Vom Fachbereich Bau- und Umweltingenieurwissenschaften
der Technischen Universität Darmstadt
zur Erlangung des akademischen Grades eines
Doktor-Ingenieurs (Dr.-Ing.)
genehmigte

DISSERTATION

von

André Müller M.Sc.

aus
Rüsselsheim

D17

Darmstadt 2023

Referent: Prof. Dr.-Ing. Carl-Alexander Graubner

Korreferent: Prof. Dr.-Ing. Hans-Joachim Linke

Tag der Einreichung: 17. Januar 2023

Tag der mündlichen Prüfung: 13. März 2023

Herausgeber:

Univ.-Prof. Dr.-Ing. Carl-Alexander Graubner

Anschrift:

Institut für Massivbau
Franziska-Braun-Straße 3
64287 Darmstadt

<http://www.massivbau.tu-darmstadt.de>

Müller, André:

Methodik zur datenbasierten Typisierung von Quartieren anhand baulicher Strukturen
Geometrisch-topologische Grundlagen für die energetische Bewertung des nationalen Quartiersbestandes
auf Basis von Geoinformationsdaten

1. Auflage Darmstadt

Dissertationsreihe des Instituts für Massivbau, Technische Universität Darmstadt; Heft 55

Online veröffentlicht bei TUprints der Universitäts- und Landesbibliothek Darmstadt (ULB)

URN: urn:nbn:de:tuda-tuprints-264174

URI: <https://tuprints.ulb.tu-darmstadt.de/id/eprint/26417>

Dieses Werk ist durch das Urheberrecht und/oder verwandte Schutzrechte geschützt. Sie sind berechtigt, das Objekt in jeder Form zu nutzen, die das Urheberrechtsgesetz und/oder einschlägige verwandte Schutzrechte gestatten. Für weitere Nutzungsarten benötigen Sie die Zustimmung der/des Rechteinhaber/s.

André Müller

Geboren 1984 in Rüsselsheim. Von 2004 bis 2009 kooperatives Ingenieurstudium „Systems Engineering“ an der Fachhochschule Wiesbaden zum Dipl.-Ing. Von 2009 bis 2012 Entwicklungsingenieur bei der Adam Opel AG. Von 2012 bis 2016 Studium zum M. Sc. in „Energy Science and Engineering“ an der Technischen Universität Darmstadt. Von Sep. 2013 bis Jan. 2014 Auslandsaufenthalt an der Universität Politècnica de Catalunya in Terrassa, Provinz Barcelona, Spanien. Seit 2016 Wissenschaftler am Institut Wohnen und Umwelt, Darmstadt, in den Forschungsfeldern „Energetische Gebäudebewertung und -optimierung“ sowie „Strategische Entwicklung des Gebäudebestandes“. Von 2017 bis 2023 externer Doktorand am Institut für Massivbau der Technischen Universität Darmstadt und assoziiertes Mitglied der Darmstädter Exzellenz-Graduiertenschule für Energiewissenschaft und Energietechnik der Deutschen Forschungsgemeinschaft DFG (GSC 1070).

VORWORT

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Institut Wohnen und Umwelt sowie als externer Doktorand am Institut für Massivbau der Technischen Universität Darmstadt. Meine Forschung war zudem eingebunden in das Forschungsvorhaben „EG2050:E4Q – Einbindung erneuerbarer Energieträger in die Energieversorgung vernetzter Quartiere“, welches durch das Bundesministerium für Wirtschaft und Klimaschutz gefördert (FKZ 03EGB0014B) und gemeinsam durch die genannten Institute bearbeitet wurde.

Meinem Doktorvater Herrn Prof. Dr.-Ing. Carl-Alexander Graubner danke ich außerordentlich für die stetige Unterstützung meiner Forschung, das mir entgegengebrachte Vertrauen und die Möglichkeit, auch als externer Doktorand, ein Teil seiner Forschungsgruppe gewesen sein zu dürfen. Ich konnte in der Zeit meiner Promotion am Institut für Massivbau viele Einblicke in mir neue Themenfelder und Methoden erlangen. Sein Anspruch, Ergebnisse höchster wissenschaftlicher Qualität und hoher Praxisrelevanz zu erzeugen, hat mich in meiner Arbeit inspiriert.

Herrn Prof. Dr.-Ing. Hans-Joachim Linke danke ich sehr herzlich für das Interesse an meinem Forschungsthema und für die Übernahme des Korreferates. Der Austausch über die Ergebnisse meiner Arbeit hat mich darin bestätigt, mit meiner Forschung an der Schnittstelle verschiedener Fachgebiete einen wichtigen Beitrag geleistet zu haben diese Fachgebiete besser zu verbinden. Ich hoffe, damit Impulse für nachfolgende Forschungsarbeiten – auch im Bereich der Geodäsie bzw. der Geoinformatik – geliefert zu haben.

Für das akribische und geduldige Korrektorat meiner Dissertationsschrift möchte ich Judith Heilmann, Johannes Koert, Patrick Wörner sowie meiner Mutter, meiner Schwiegermutter und meiner Frau danken.

Des Weiteren gilt ein besonderer Dank meinen Kolleginnen und Kollegen an beiden Instituten für deren Kollegialität und fachliche Begleitung. Ich empfinde es als besonderes Glück, an beiden Einrichtungen ein Arbeitsumfeld vorgefunden zu haben, in das ich mich fachlich und persönlich gut integrieren konnte und in dem ich mich im Rahmen meiner Promotion und durch die darüber hinausgehenden Aufgaben weiterentwickeln konnte. Insbesondere meinen Kolleginnen und Kollegen am Institut Wohnen und Umwelt möchte ich dafür danken, dass sie mir durch die Übernahme vieler Aufgaben innerhalb von Forschungsvorhaben sowie der Forschungsselbstverwaltung unseres Instituts die nötigen Freiräume für meine Dissertation ermöglicht haben.

Einen darüber hinausgehenden Dank richte ich zudem an Nikolaus Diefenbach, Marc Großklos und Holger Cischinsky, welche meine Arbeit in ihrer Rolle als Mentoren am Institut Wohnen und Umwelt kontinuierlich begleitet und stets kritisch hinterfragt haben. Sie lieferten damit wertvolle Impulse für meine Forschung. Monika Meyer möchte für die Unterstützung meiner Arbeit in ihrer Rolle als Institutsleitung danken.

Weiterhin möchte ich mich bei Saskia Falk, Judith Heilmann und Pascal Heit bedanken, welche mir im Rahmen ihrer Tätigkeit als studentische Hilfskräfte eine große Unterstützung bei der Bearbeitung des Projekts EG2050:E4Q waren.

Ein besonderer Dank gilt meiner Familie. Meinen Eltern, die mir meinen Lebensweg durch ihre Fürsorge während meiner Kindheit und Jugend erst ermöglicht haben und mich auch heute stets unterstützen. Meinen Eltern und Schwiegereltern für die Unterstützung meiner Familie in der Zeit, in der die Erarbeitung meiner Dissertation einen besonders großen zeitlichen Umfang einnahm und ich deshalb nur in geringem Umfang Zeit für den familiären Alltag daheim hatte.

Und schließlich danke ich von ganzem Herzen meiner Frau Anke und meinen Kindern Klara und Johann. Nur durch eure unendliche Geduld und fortwährende Nachsicht konnte ich die unzähligen Stunden an dieser Forschungsarbeit mit dem hier vorliegenden Buch auch zu einem erfolgreichen Abschluss bringen.

Die vielen weiteren Menschen, die mich auf meinem bisherigen Lebensweg begleitet haben möchte ich an dieser Stelle ebenfalls würdigen. Jede einzelne Erfahrung der vergangenen 38 Jahre im schulischen, beruflichen, sportlichen und privaten Umfeld hat zur Bildung und Entwicklung meiner Persönlichkeit und somit auch zur Entstehung dieser Dissertation beigetragen.

ZUSAMMENFASSUNG

Die deutsche Gesellschaft befindet sich in einem Transformationsprozess, um der Klimakrise zu begegnen. Vor dem Hintergrund der notwendigen Dekarbonisierung der Wärmeversorgung von Gebäuden haben sich Konzepte etabliert, welche für Quartiere, d. h. mehrere Gebäude in einem räumlichen Zusammenhang, die Reduktion von Treibhausgasemissionen durch eine Erhöhung der Gebäude- und Anlageneffizienz sowie durch die Einbindung erneuerbarer Energieträger erreichen. Während auf lokaler Handlungsebene eine Vielzahl an Projekten umgesetzt wird, fehlte die empirische Basis, um die Art und Anzahl von Quartieren in Deutschland zu erfassen und die Wirkung von Quartierskonzepten in Hoch- und Szenarienrechnungen zum Klimaschutz zu integrieren.

Die vorliegende Arbeit liefert eine Methodik zur Nutzung von Geobasisdaten für eine Analyse der Bebauungsstruktur und eine Einteilung dieser in geometrisch-topologisch ähnliche Gebiete mittels Clusteranalyse. Eine Vereinigung benachbarter Gebiete ähnlicher Bebauung führt im Anschluss zur Bildung von Quartieren, was schließlich eine statistische Analyse des Quartiersbestandes erlaubt. Daraus können Kenngrößen für die Beschreibung von Typvertretern für Quartiersgruppen, sogenannte Typquartiere, festgelegt werden. Die beispielhafte Anwendung der Methodik am Datensatz Berlins zeigt eine praxisgerechte Quartiersbildung sowie die Verwendbarkeit der Ergebnisse der Quartierstypisierung für vielfältige Forschungs- und Planungszwecke. Insbesondere eine Verwendung des nationalen Geodatenbestandes zur Quartierstypisierung eröffnet die Möglichkeit weiterführender Szenarioanalysen im Kontext nationaler Klimaschutzstrategien.

ABSTRACT

In the light of the climate crisis, German society is subject to an inevitable transformation process. Part of this is the need to reduce the heating-related carbon emissions of buildings by increasing the efficiency and the share of renewable energy sources. In order to reach that aim, concepts for building refurbishment and energy provision on a district level have been developed. However, a lack of empirical information on the number and delimitation of districts in Germany hinders the diffusion of successful concepts. In addition, due to the missing empirical data, the district level cannot be properly reflected in scenario analyses and climate change mitigation strategies.

This thesis presents a methodology which processes geospatial data to achieve a description of the urban morphology on a local level. By using a cluster analysis, areas of similar morphology are identified, and neighboring areas are combined to city districts. Thus, a statistical analysis of the district stock is possible, and archetype districts can be defined by the use of the statistical parameters. Accordingly, the empirical basis which can result from a study based on nationwide geospatial data will generate a benefit for research and planning purposes aiming at the decarbonization of the building stock on different spatial levels.

INHALTSVERZEICHNIS

Abkürzungsverzeichnis	IV
Parameter- und Variablenverzeichnis	VIII
1 Einleitung	1
1.1 Ausgangssituation und Motivation	1
1.2 Zielsetzung	4
1.3 Vorgehensweise	5
2 Betrachtungsraum „Quartier“	9
2.1 Einleitung und allgemeine Quartiersdefinitionen	9
2.2 Das Quartier in der Energieforschung	13
2.3 Definition eines kontextbezogenen Quartiersbegriffs	15
2.4 Existierende Ansätze zur Typisierung von Quartieren	19
2.4.1 Siedlungsstrukturtypen nach Roth (1980)	19
2.4.2 Weitere gebäudezentrierte Quartierstypologien	22
2.5 Softwarelösungen zur Bewertung energetischer Quartierskonzepte	28
2.6 Grenzen bestehender Typologien im Hinblick auf den Klimaschutz	32
2.6.1 Übertragung der Ergebnisse auf den deutschen Quartiersbestand	32
2.6.2 Anforderungen an die Datenbasis der Gebäudemerkmale und die Datenqualität	33
3 Methoden zur Gruppenbildung und Klassifizierung	35
3.1 Hinführende Erläuterungen	35
3.2 Methodische Grundlagen	37
3.3 Verfahren zur Clusteranalyse	43
3.3.1 Allgemeine Beschreibung der Clusteranalyse	43
3.3.2 Hierarchische Clusteranalyseverfahren	45
3.3.3 Partitionierende Clusteranalyseverfahren	47
3.3.4 Dichtebasierte Clusteranalyseverfahren	48
3.3.5 Gütekriterien der Gruppenfestlegung	50
3.4 Überwachtes Maschinelles Lernen zur Datenklassifizierung	52

3.5	Anwendungsbeispiele für Clusteranalysen im Kontext der energetischen Bewertung von Gebäuden und Quartieren	56
4	Geoinformationsdaten in Deutschland	61
4.1	Grundlegende Erläuterungen zu Geoinformationssystemen und –daten	61
4.2	Datenquellen und –formate von Geoinformationsdaten.....	64
4.2.1	Amtliches Liegenschaftskatasterinformationssystem.....	64
4.2.2	Digitales Basis-Landschaftsmodell	65
4.2.3	Verwaltungsgebiete	66
4.2.4	Amtliche Hauskoordinaten und Amtliche Hausumringe.....	67
4.2.5	Georeferenzierte Adressdaten.....	67
4.2.6	3D-Gebäudemodelle	68
4.3	Einsatzbereiche geoinformationstechnischer Daten	70
4.4	Einschränkungen bei der Nutzung von Geoinformationsdaten im Kontext der energetischen Bewertung von Gebäuden und Quartieren	73
5	Entwicklung einer Methodik zur geodatenbasierten Quartierstypisierung	77
5.1	Einführende Erläuterungen.....	77
5.2	Beschreibung verfügbarer GIS-Daten und verwendeter Software	79
5.3	Datenvorverarbeitung	85
5.4	Gruppenbildung mittels Clusteranalyse auf Ebene von Teilbauflächen	89
5.4.1	Allgemeine Erläuterungen zur durchgeführten Clusteranalyse.....	89
5.4.2	Clusteranalyse zur Bildung von Teilbauflächengruppen ähnlicher Bebauungsstruktur	93
5.4.3	Vorgelagerte Untersuchung zur Clusteranalyse von Einzelgebäuden.....	97
5.5	Bildung von Quartieren aus gruppierten Teilbauflächen	98
5.6	Statistische Analyse des Quartiersbestandes	103
5.7	Abschließende Anmerkungen zu den Randbedingungen der erarbeiteten Methodik.....	106
6	Anwendung der Methodik zur Typisierung von Quartieren am Beispiel Berlin.....	109
6.1	Einleitende Bemerkungen.....	109
6.2	Datenüberprüfung und Datenauswahl	110

6.2.1 Überprüfung der Gebäudemodellierung und Ansätze zur Datensatzreduzierung	110
6.2.2 Verarbeitung der zur Verfügung stehenden GIS-Daten	115
6.2.3 Auswahl der für die Methodenanwendung verwendeten Datenbasis.....	118
6.3 Datenvorverarbeitung des Datensatzes Berlin.....	119
6.4 Clusteranalyse auf Ebene von Teilbauflächen und Überprüfung der Gruppenbildung	122
6.4.1 Ausführung der Clusteranalyse zur Gruppenbildung	122
6.4.2 Vergleich der Varianten zur Festlegung einer geeigneten Gruppenanzahl	125
6.4.3 Plausibilisierung der Gruppenbildung für Teilbauflächen	127
6.5 Bildung von Quartieren aus gruppierten Teilbauflächen	132
6.6 Statistische Analyse des Berliner Quartiersbestandes und Festlegung von Kennwerten für Typquartiere	134
7 Möglichkeiten zur Quartierstypisierung des nationalen Geodatenbestandes	141
8 Resümee und Ausblick	145
8.1 Resümee	145
8.2 Ausblick.....	148
Literaturverzeichnis	151
Anhang.....	173

ABKÜRZUNGSVERZEICHNIS

°C	Grad Celcius
3D	dreidimensional
ADE	Application Domain Extension (engl.)
AdV	Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik
AFIS	Amtliches Festpunkt-Informationssystem
ALKIS	Amtliches Liegenschaftskatasterinformationssystem
ASW	average silhouette width (engl.), mittlere Silhouetten-Weite
ATKIS	Amtliches Topographisch-Kartographisches Informationssystem
Basis-DLM	Digitales Basis-Landschaftsmodell Deutschland
BGeoRG	Gesetzes über die geodätischen Referenzsysteme, -netze und geotopographischen Referenzdaten des Bundes
BKG	Bundesamt für Kartographie und Geodäsie
BMF	Bundesministerium für Finanzen
BMU	Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit bzw. Bundesministerium für Umwelt, Naturschutz, Bau und Reaktorsicherheit
BMWi	Bundesministerium für Wirtschaft und Technologie bzw. Bundesministerium für Wirtschaft und Energie
BMWK	Bundesministerium für Wirtschaft und Klimaschutz
CAD	computer-aided design (engl.), computergestütztes Konstruieren
CBD	central business district (engl.), Stadtzentrum bzw. Gewerbezentrum
CityGML	City Geographic Markup Language (engl.)
CLARA	Clustering Large Applications (engl.)
CLARANS	Clustering Large Applications based on Randomized Search (engl.)
DBSCAN	Density Based Spatial Clustering of Applications with Noise (engl.)
DH	Doppelhaus
District-ECA	District Energy Concept Adviser (engl.), Energiekonzeptberater für Stadtquartiere
EE	energetisches Einzelelement nach Hegger et al. (2012; 2015a)

EFH	Einfamilienhaus
EnEff:Stadt	Forschungsinitiative „Energieeffiziente Stadt“ des BMWi
EnEff2050	EnEff.Gebäude.2050 – Innovative Vorhaben für den nahezu klimaneutralen Gebäudebestand 2050
EST	Energetischer Stadtraumtyp nach Hegger et al. (2012; 2015a)
EU	Europäische Union
ExWoSt	Experimenteller Wohnungs- und Städtebau
GA	Georeferenzierte Adressdaten
GDI	Geodateninfrastruktur
GDI-DE	Geodateninfrastruktur Deutschland
GEG	Gebäudeenergiegesetz
GeoZG	Gesetz über den Zugang zu digitalen Geodaten
GFZ	Geschossflächenzahl
GIS	Geoinformationssystem, geoinformationstechnisch
GIS-Daten	Geoinformationsdaten
Gl.	Gleichung
GML	Geographic Markup Language (engl.)
GRZ	Grundflächenzahl
ha	Hektar
HK-DE	Amtliche Hauskoordinaten Deutschland
HU-DE	Amtliche Hausumringe Deutschland
IBZ	Bezeichnungsidentifikator der Art einer Verwaltungseinheit
IFEU	Institut für Energie- und Umweltforschung
inkl.	inklusive
INSPIRE	EU-Richtlinie zur Schaffung einer Geodateninfrastruktur in der Europäischen Gemeinschaft
IWU	Institut Wohnen und Umwelt GmbH
KBS	Koordinatenbezugssystem
KNN	Künstliches neuronales Netz

Abkürzungsverzeichnis

KSG	Klimaschutzgesetz
KT25	Kommunale Teilgebiete Deutschland im Maßstab 1:25.000
LoD	Level of Detail (engl.), Detaillierungsgrad
m	Meter
m ²	Quadratmeter
m ³	Kubikmeter
Max.	Maximum
MBN	Maß baulicher Nutzung
MFH	Mehrfamilienhaus
Min.	Minimum
MinPts	minimum number of points (engl.), Mindestzahl der dichteerreichbaren Nachbarn (dient als Dichtekriterium im Clusterverfahren DBSCAN)
MIT	Massachusetts Institute of Technology
ML	Machine Learning (engl.), Maschinelles Lernen
MW	Mittelwert
N	Anzahl, Grundgesamtheit
n	Anzahl, Stichprobenumfang
NAS	Normbasierte Austauschschnittstelle
NGIS	Nationale Geoinformations-Strategie
p _{0,25}	unteres Quartil
p _{0,50}	Median
p _{0,75}	oberes Quartil
PAM	Partitioning Around Medoids (engl.)
RH	Reihenhaus
RT	Raumtyp nach Roth (1980)
SD	Standardabweichung (engl. standard deviation)
SRT	energetischer Straßenraumtyp nach Hegger et al. (2012; 2015a)
ST	Siedlungs- bzw. Siedlungsstrukturtyp nach Roth (1980) oder Blesl (2002)
THG	Treibhausgas/-e

UMI	Urban Modeling Interface (engl.) des Sustainable Design Lab am MIT
UML	Unified Markup Language (engl.)
VG25	Verwaltungsgliederung Deutschland im Maßstab 1:25.000
XML	Extensible Markup Language (engl.)

PARAMETER- UND VARIABLENVERZEICHNIS

a	Fläche
C	Cluster bzw. Gruppe
d	Distanz bzw. Abstand
h	Höhe
i	Bezeichner für Laufvariable
I_{ASW}	Indikator der mittleren Silhouetten-Weite
j	Bezeichner für Laufvariable
k	Klassenanzahl, Gruppenanzahl
k_{opt}	geeignete Gruppenanzahl
m	Bezeichner für Laufvariable
MBN	Maß baulicher Nutzung
n	Bezeichner für Laufvariable
p, P	Wahrscheinlichkeit
\tilde{p}	Kumulierte Wahrscheinlichkeit
s	Ähnlichkeit
s	empirische Standardabweichung der Merkmalsausprägungen einer Variablen
sw	Silhouetten Weite (engl. silhouette width)
v	Merkmalsvektor eines Objektes
x	Merkmalsausprägung einer beliebigen Variablen
x	Objekt bzw. Merkmalsträger
\bar{x}	Mittelwert der Merkmalsausprägungen einer Variablen

1 EINLEITUNG

1.1 Ausgangssituation und Motivation

Die deutsche Politik, vielmehr noch die gesamte Gesellschaft der Bundesrepublik Deutschland, befindet sich in einem Transformationsprozess, um der vermutlich folgenschwersten globalen Krise seit dem Ende des Kalten Krieges adäquat zu begegnen. Heute bedrohen der Klimawandel und die damit einhergehenden Veränderungen der Umwelt die Lebensgrundlage eines Großteils der Weltbevölkerung. Der Klimawandel – also eine Verstärkung des natürlichen Treibhauseffekts, welcher durch im menschlichen Handeln begründete Treibhausgasemissionen induziert ist, – wird deshalb folgerichtig bereits als Klimakrise beschrieben. Dieser Krise wird durch stetig intensiviertere globale, europäische und nationale Bemühungen zur Erreichung einer maximalen mittleren Temperaturerhöhung der Erde von deutlich unter 2 °C, im Vergleich zum vorindustriellen Zeitalter, versucht entgegenzuwirken. (SRU 2020; BMF 2019; IPCC 2018; UN 2015)

Der öffentliche und häufig auch der bau- bzw. energiefachliche Diskurs beschränkt sich auf eine Diskussion der im Rahmen internationaler Verhandlungen festgelegten und in der deutschen Klimapolitik formulierten Zielwerte für bestimmte Stütz- und Zieljahre. Demgegenüber scheint die hinter den festgeschriebenen Treibhausgasreduktionszielen liegende Budgetbetrachtung – also das Grundprinzip, dass in Summe nur noch eine begrenzte Menge an Treibhausgasen emittiert werden kann, bis es zu unumkehrbaren, negativen Klimafolgewirkungen kommt – meist in der Auseinandersetzung mit dem Thema nicht ausreichend beachtet. Die mangelnde Berücksichtigung des Budgetansatzes führt jedoch zu einer verzögerten Reduktion der globalen und nationalen Treibhausgasemissionsmengen. Dadurch entsteht das Risiko, dass in der Zukunft deutlich stärkere Anstrengungen zur Reduktion von Emissionen erforderlich sein werden, oder im schlimmeren Fall, dass ein Aufholen versäumter Emissionsminderungen bis zu einem Zieldatum nicht mehr möglich sein könnte und das verfügbare Budget überschritten würde. (IPCC 2021; BMU 2019)

Das durch die deutsche Politik im Einklang mit den internationalen Vereinbarungen zum Klimaschutz formulierte Ziel der Transformation Deutschlands in eine klimaneutrale Gesellschaft erfordert die Berücksichtigung der bis zum Zieljahr 2045 zur Verfügung stehenden Restemissionsmengen in allen gesellschaftlichen Bereichen. Zu diesem Zweck wurden für die Sektoren Energiewirtschaft, Gebäude, Industrie, Verkehr sowie Landwirtschaft und Landnutzung Aktionspläne entwickelt, welche einen übergeordneten Rahmen für Klimaschutz formulieren und direkte Maßnahmen zur Reduktion der jeweiligen Treibhausgasemissionen beinhalten. Eine Überprüfung der Emissionsminderungen erfolgt seit dem Inkrafttreten des Klimaschutzgesetzes (2019) in regelmäßigen Abständen von wenigen Jahren und soll, insofern dies erforderlich ist, eine frühzeitige Korrektur der Maßnahmenpläne erlauben. (BMU 2019; BMF 2019; KSG, vom 18.08.2021)

Die jüngsten Überprüfungen der sektoralen Emissionsbilanzen im Gebäudebereich zeigen, dass die formulierten Emissionsminderungsziele nicht erreicht wurden und ein akuter Handlungsbedarf im Hinblick auf die Implementierung weiterführender Maßnahmen besteht (BMU 2021). Noch immer setzt sich der Gesamtbestand der Gebäude in Deutschland zum überwiegenden Teil aus Bauten zusammen, welche hinsichtlich der Höhe des Energieverbrauchs und der Art der Energieträger nicht mit der Zielvorgabe eines klimaneutralen Gebäudebestandes bis spätestens 2045 kompatibel sind. Wie bereits in repräsentativen Stichprobenerhebungen zu den Ausprägungen energetischer Gebäudemerkmale im Wohngebäudebestand für den Zeitraum von 2005 bis 2016 ermittelt (Cischinsky und Diefenbach 2018; Diefenbach et al. 2010; dena 2021) und in ähnlicher Art und Weise kürzlich auch für den Nichtwohngebäudebestand festgestellt (Hörner et al. 2021) zeigt sich eine zu geringe mittlere Rate der energetischen Modernisierung der Gebäudehülle innerhalb der vergangenen zwei Dekaden. Auch eine Umstellung der Gebäudewärmeversorgung hin zur (zumindest anteiligen) Bereitstellung von Wärme aus erneuerbaren Energieträgern ist zwar für den Bereich der Neubauten anvisiert. Für den deutlich größeren Anteil, den Bestandsgebäude am Gebäudesektor ausmachen, ist eine solche Transformation aber kurzfristig noch nicht erkennbar. (dena 2021; BDEW 2021; BWP 2021; Cischinsky und Diefenbach 2018)

Vor dem Hintergrund der unzureichenden Modernisierungsfortschritte im Gebäudesektor wurden deshalb zunehmend Modernisierungsansätze untersucht, welche mit dem „Quartiersmaßstab“ eine darüberliegende räumliche Ebene abdecken. Diese Ansätze fokussieren die Umsetzung eines integrierten Entwicklungskonzepts zur Verbesserung des Gebäudewärmeschutzes, zum Einsatz effizienter Wärme- und Kälteversorgungstechnologien sowie zur lokalen Erzeugung und Nutzung von Strom und Wärme aus regenerativen Energiequellen. (SRU 2020; Langenbrinck et al. 2017; Erhorn et al. 2010; Erhorn-Kluttig und Erhorn 2016)

Die zugrundeliegenden Überlegungen für die Etablierung von Ansätzen zur energetischen Gebäudemodernisierung auf Quartiersebene entspringen sowohl energetischen, ökonomischen als auch sozial-psychologischen Prinzipien bzw. Mechanismen (SRU 2020; Erhorn et al. 2010; Hentschel 2017). Quartierskonzepte stellen insofern einen Ansatzpunkt zur Gebäudemodernisierung und Implementierung effizienterer Versorgungstechnologien, unter Einbezug erneuerbarer Energien, dar.

Folglich haben in der Vergangenheit verschiedene theoretische Untersuchungen sowie geförderte Demonstrationsprojekte eine Vielzahl an Möglichkeiten aufgezeigt, wie durch (energetische) Modernisierung im Quartiersmaßstab unter Ausnutzung verschiedener Skalen- und Synergieeffekte eine Reduktion der gebäudeenergiebedarfsbedingten Treibhausgasemissionen erreicht werden kann. (Erhorn et al. 2010; Müller et al. 2012; Weißmann 2017; Stockinger 2019)

Trotz der positiven Beispiele scheint das Quartier noch immer nicht hinreichend als Betrachtungs- bzw. Optimierungsebene im Rahmen der städtebaulichen Planungspraxis für Neubauten, Modernisierungen sowie für die Erweiterung oder Transformation der Energieversorgungsinfrastruktur etabliert. Zwar verfolgen verschiedene Städte Strategien zur Erreichung einer meist individuell definierten Klimaneutralität, häufig werden hier jedoch die Handlungsspielräume und Bilanzgrenzen sehr eng gefasst und eine Steuerung der privatwirtschaftlichen Aktivitäten hin zu einem klimaneutralen Gebäudebestand wird nicht oder nur mittelbar berücksichtigt (Wissenschaftsstadt Darmstadt 2019; KSB Darmstadt 2020; Landeshauptstadt Mainz 2017). Eine Nutzung der erwarteten Potenziale des Quartiersmaßstabs erfolgt in der Praxis eher im Rahmen von Vorhaben, welche als „Energetische Stadt-sanierung“ gefördert werden (Langenbrinck et al. 2017). Eine Etablierung dieser Aktions-ebene im Alltag kommunaler und privatwirtschaftlicher Akteure mit Gebäudeenergiebezug fand jedoch in der Breite bisher nicht ausreichend statt.

Auch im wissenschaftlichen Diskurs wurden die Potenziale der Betrachtungsebene „Quar-tier“ bisher überwiegend im Hinblick auf die jeweiligen Einzelquartiere in Modellvorhaben oder theoretischen Untersuchungen behandelt. Zwar wurde eine Übertragbarkeit auf andere Quartiere im Rahmen der jeweiligen Projektberichte sowie in Publikationen der Begleit-forschung „Energiewendebauen“ für die gebäudebezogenen Inhalte des Energiefor-schungsprogramms der Bundesregierung gegeben (Begleitforschung EWB 2020; Müller et al. 2012; Beier et al. 2016), eine Verallgemeinerung und Übertragung auf den deutschen Gebäudebestand erfolgte jedoch nicht. Über die Modellvorhaben hinaus wurden zudem Softwarelösungen entwickelt, welche der Untersuchung abgegrenzter städtischer Teilge-biete dienen und die Erarbeitung von energetischen Quartierskonzepten erlauben sollen (vgl. Nouvel et al. 2015a; Budde et al. 2018; Reinhart et al. 2013). Eine Erweiterung des Untersuchungsraumes auf den gesamten Gebäude- bzw. Quartiersbestand in Deutschland und damit eine belastbare Verallgemeinerung der Beiträge von Quartierskonzepten zur Er-reichung der Klimaschutzziele im Gebäudesektor ist mit diesen Werkzeugen bisher jedoch ebenfalls nicht möglich.

Auch bereits in der Forschung erarbeitete sowie in der (Stadt-)Planungspraxis etablierte Typologien städtischer Bebauungsstrukturen bzw. städtischer Quartiere fanden bisher led-iglich im lokalen Kontext bzw. für thematisch oder räumlich stark abgegrenzte Untersu-chungen Anwendung (Everding und Lindner 2007; Wern et al. 2017). Dies lässt sich u. a. mit der zur Typisierung der raumbezogenen Bebauungsstrukturen nötigen Informations-dichte zu existierenden Einzelgebäuden begründen. Die notwendigen Informationen sind für größere Gebiete oder eine Vielzahl von Gebieten nur mit vergleichsweise hohem Auf-wand ermittelbar (vgl. Ali et al. 2021).

1.2 Zielsetzung

Bezugnehmend auf den zuvor umrissenen Forschungs- und Handlungsrahmen zu energetischen Modernisierungs- bzw. Versorgungskonzepten von Quartieren und unter Berücksichtigung der o. g. Hindernisse ergibt sich für die vorliegende Arbeit die Zielsetzung, eine Methodik zu entwickeln, welche es erlaubt, Typquartiere als Vertreter einer Gruppe von realen Quartieren zu definieren und deren bauliche Struktur sowie deren Häufigkeit im deutschen städtischen Gebäudebestand nachzuvollziehen. Die Methodik muss dabei in Abgrenzung zu bereits existierenden Arbeiten und Definitionen von Typquartieren und typischen Siedlungs- bzw. Bebauungsstrukturen eine großflächige oder gar flächendeckende statistische Analyse des Quartiersbestandes zulassen und sich deshalb elementar auf eine, den räumlichen Zusammenhang von Gebäuden abbildende, Datenbasis stützen. Gleichzeitig ergibt sich aus dem Anspruch einer möglichst hohen Verallgemeinerbarkeit der Ergebnisse die Notwendigkeit, dass diese Datenbasis über einen spezifischen, lokalen Kontext hinaus für einen größeren zusammenhängenden Ausschnitt des Gebäudebestandes vorliegt, wenn nicht sogar den deutschen Gebäudebestand vollständig in seinem räumlichen Zusammenhang abbildet. Mindestens sollte die Methodik jedoch so angelegt sein, dass deren Anwendung auf großräumige, zusammenhängende Ausschnitte des Gebäude- und damit Quartiersbestandes ohne Weiteres möglich ist.

Durch die Überführung von gebäudebezogenen Informationen des Vermessungs- und Katasterwesens in standardisierte, digitale Formate wurde innerhalb der vergangenen Dekade eine Vielzahl geoinformationstechnischer Daten (GIS-Daten) erzeugt, welche auf den verschiedenen Ebenen des föderal strukturierten Katasterwesens in Deutschland in verschiedenen Detaillierungsgraden vorliegen. Diese GIS-Datenbestände sollen folglich im Rahmen der vorliegenden Arbeit zunächst untersucht und bewertet werden, um eine geeignete Datenbasis auswählen bzw. sie aus den existierenden Datenbeständen aufbereiten zu können.

Die verfügbare Datenbasis bildet auch den Ausgangspunkt für die Entwicklung der Methodik zur datenbasierten Typisierung von Quartieren anhand baulicher Strukturen. Im Rahmen der Methodenentwicklung soll weiterhin die Auswahl eines oder mehrerer geeigneter statistischer Verfahren erfolgen, welche eine datengetriebene Einteilung des städtischen bebauten Raumes in Quartiere ermöglichen. Gleichzeitig soll zudem die praktische Relevanz und eine spätere Übertragbarkeit der Ergebnisse, d. h. der ermittelten Typquartiere und deren Häufigkeit, in den Kontext strategischer Untersuchungen zum Klimaschutz und der lokalen Quartiersplanung gewährleistet werden. Eine solche Übertragung der Ergebnisse der Typisierung und statistischen Beschreibung des deutschen Quartiersbestandes soll in dieser Arbeit abschließend beispielhaft dargestellt werden und den Mehrwert der durchgeführten Methodenentwicklung verdeutlichen.

Zur Erreichung der formulierten Ziele müssen die in den folgenden Kapiteln beschriebenen Untersuchungen Antworten auf die nachfolgenden Leitfragen geben:

- Wie ist die Betrachtungsebene „Quartier“ zu definieren und abzugrenzen, um die Zusammenführung von Gebäuden zu Quartieren zu ermöglichen und eine statistische Analyse sowie Typisierung zu erlauben?
- Welche Methoden zur Bildung von Gruppen von Gebäuden und anderen kleinräumigen Bezugsobjekten existieren und lassen sich vor dem Hintergrund der Zielsetzungen dieser Arbeit anwenden, um aus einer verfügbaren Datenbasis Typquartiere abzuleiten, die gleichzeitig eine breite praktische Anwendbarkeit besitzen?
- Welche flächendeckende Datenbasis ist für die Einteilung und statistische Analyse von Quartieren im Hinblick auf den räumlichen Zusammenhang von Gebäuden und die Erstellung energetischer Quartierskonzepte notwendig?
- Wie muss die Methode zur Typisierung und statistischen Analyse von Quartieren gestaltet werden und wie lassen sich die Ergebnisse im formulierten Kontext absichern?
- Wie können die Ergebnisse dieser Arbeit im Rahmen von Untersuchungen zu den Emissionsminderungspotenzialen im Gebäude- und Quartiersbestand genutzt werden und welche Impulse bzw. Empfehlungen liefert diese Anwendung für die Erarbeitung energetischer Quartierskonzepte?

Diese Forschungsfragen sollten aufgrund der Breite der in dieser Arbeit behandelten Sachverhalte nur als übergeordneter Rahmen verstanden werden. In den entsprechenden Kapiteln dieser Arbeit werden die jedem Themenkomplex zugrundeliegenden Fragestellungen und Herausforderungen vertiefend diskutiert und bearbeitet. Die Beantwortung der o. g. Leitfragen dieser Arbeit wird sowohl die Relevanz als auch den Mehrwert der hier durchgeführten Untersuchung verdeutlichen.

1.3 Vorgehensweise

Die Erreichung der im vorherigen Abschnitt formulierten Ziele dieser Arbeit sowie die Beantwortung der Forschungsfragen erfordern eine Verknüpfung verschiedener Themenkomplexe aus unterschiedlichen Forschungsdisziplinen. In *Kapitel 2* erfolgt zuerst eine umfassende Diskussion des Begriffs „Quartier“ und seiner Bedeutung in der gebäudebezogenen Energieforschung sowie bisheriger Ansätze zur Anwendung der räumlichen Ebene des Quartiers im Rahmen der Erstellung und Bewertung von Konzepten zur Wärme- und Stromversorgung. Für den Kontext der vorliegenden Untersuchung wird darauf aufbauend eine Definition des Quartiersbegriffs formuliert. Zudem werden existierende Typisierungskonzepte von Quartieren, Siedlungsstrukturen und Stadtraumtypen sowie Ansätze zur computergestützten Erarbeitung und Bewertung von energetischen Quartierskonzepten

vorgestellt. Den Abschluss des Kapitels bildet eine Darstellung der Grenzen der beschriebenen Ansätze. Hierbei werden die existierenden, eingangs bereits umrissenen Forschungslücken detailliert.

Kapitel 3 bietet einen breiten Überblick über die methodischen Grundlagen zu den im Rahmen dieser Arbeit untersuchten und angewandten Verfahren zur Gruppenbildung, der sog. Clusteranalyse. Hier erfolgt zunächst eine allgemeine Einordnung der clusteranalytischen Verfahren in die Methoden der statistischen und informationstechnischen Datenanalyse, welche durch die vergleichende Beschreibung einer Auswahl von Clustermethoden ergänzt wird. Ebenfalls werden sog. Klassifizierungsmethoden erläutert, welche die Zuordnung von Objekten eines Datensatzes in eine bestehende Gruppeneinteilung erlauben. Des Weiteren werden Forschungsarbeiten präsentiert, welche raum-, gebäude- oder energiebezogene Sachverhalte untersuchen und sich auf die Clusteranalyse als Untersuchungsmethode stützen. Abschließend erfolgt eine Diskussion über den Nutzen und die Grenzen der statistischen Verfahren zur Gruppenbildung von Datensätzen.

Einen Überblick zur generellen Verfügbarkeit von GIS-Daten, zu deren Erhebungshintergründen sowie zum Informationsgehalt verschiedener Datenangebote liefert *Kapitel 4*. Ein besonderer Fokus wird bei der Beschreibung der Daten auf die Datenbestände des Amtlichen Liegenschaftskatasterinformationssystems (ALKIS) sowie auf dreidimensionale (3D) Gebäudemodelle gelegt, welche innerhalb der vergangenen Jahre flächendeckend zur Bearbeitung von Aufgaben der öffentlichen Verwaltung und für Forschungszwecke zur Verfügung gestellt wurden. Das Kapitel schließt ebenfalls mit einer Vorstellung verschiedener Einsatzbereiche der GIS-Daten und zeigt die Chancen und Grenzen auf, die mit einer Verwendung der verfügbaren Daten vor dem Hintergrund der Zielsetzung dieser Arbeit verbunden sind.

Nach der Erarbeitung der spezifischen Hintergründe in den Themenkomplexen „Quartier“, „Clusteranalyse“, und „GIS-Daten“ folgt in *Kapitel 5* eine Zusammenführung im Rahmen der Entwicklung einer Methodik zur GIS-datenbasierten Quartierstypisierung. Hierfür werden zuerst die zur Verfügung stehenden GIS-Daten sowie die verwendete Software beschrieben. Im Anschluss daran wird eine mehrstufige Methodik erarbeitet, welche sich zusammensetzt aus ...

- 1.) einer statistischen Auswertung lokaler Gebäudepopulationen auf einer kleinskaligen räumlichen Ebene, den sog. Teilbauflächen, zur Beschreibung von kleinräumigen Bebauungsstrukturen (*Datenvorverarbeitung*),
- 2.) der Bildung von Gruppen hinsichtlich ihrer Bebauungsstruktur ähnlicher Teilbauflächen mittels *Clusteranalyse*,

- 3.) einer Vereinigung benachbarter Teilbauflächen ähnlicher Bebauungsstrukturen zu Quartieren (*Quartiersbildung*) und
- 4.) der *statistischen Analyse* des Quartiersbestandes sowie der *Festlegung von Kennwerten für Typquartiere* als Vertreter der Quartiersgruppen.

In *Kapitel 6* folgt die Anwendung der entwickelten Methodik auf einen Beispieldatensatz. Diese Untersuchungen haben das Ziel, die Ausführung der einzelnen Verfahrensschritte zu verdeutlichen und die Plausibilität sowie Praxistauglichkeit des entwickelten Verfahrens zu bewerten. Abschließend werden statistische Kennzahlen zur Häufigkeit der Typquartiere im untersuchten Beispieldatensatz ermittelt und Kenngrößen der bebauungsstrukturbeschreibenden Merkmale für Typquartiere, als Vertreter dieser Quartiersgruppen, festgelegt. Als Ergebnis der Anwendung der Methodik zur GIS-datenbasierten Typisierung von Quartieren stehen statistische Kennzahlen und Mengengerüste des Quartiersbestandes zur weiteren Verwendung zur Verfügung. Die Ergebnisse können z. B. als Hochrechnungsfaktoren für die Wirkung energetischer Quartierskonzepte in diesem Untersuchungsgebiet genutzt werden oder lokale Potenzialanalysen unterstützen.

Auf die Methodenentwicklung und -anwendung folgt in *Kapitel 7* die Beschreibung der Möglichkeiten zur Quartierstypisierung unter Verwendung des nationalen Datenbestandes. Hierbei werden zwei mögliche Ansätze vorgestellt und deren jeweilige Besonderheiten beschrieben. Es wird dabei aufgezeigt, dass die entwickelte Methodik ohne weitere Anpassungen über den räumlichen Kontext des Anwendungsbeispiels hinaus für die Typisierung anderer großflächiger Gebiete bis hin zum vollständigen nationalen Geodatensatz verwendbar ist.

Die vorliegende Forschungsarbeit schließt in *Kapitel 8* letztlich mit einem Resümee, in dem die gewonnen Erkenntnisse zusammengefasst werden und eine kurze übergeordnete Einordnung im Hinblick auf den aktuellen Stand der gebäude- und quartiersbezogenen Energieforschung in Deutschland erfahren. Ein Ausblick auf mögliche, anknüpfende Forschungsinhalte vervollständigt dieses Kapitel.

2 BETRACHTUNGSRAUM „QUARTIER“

2.1 Einleitung und allgemeine Quartiersdefinitionen

Im Rahmen von Kapitel 1 wurde die Notwendigkeit einer zeitnahen und deutlichen Reduktion des nutzungsbedingten Ausstoßes von Treibhausgasemissionen (THG-Emissionen) der Gebäude in Deutschland zur Erreichung eines klimaneutralen Gebäudebestandes bis zum Jahr 2045 hergeleitet. Soll auch die Erhöhung der mittleren Erdtemperatur auf deutlich unter 2 °C im Vergleich zum vorindustriellen Zeitalter begrenzt werden (UN 2015; BMUB 2016), muss die Einhaltung des bis dahin zur Verfügung stehenden sektoralen Emissionsbudgets ebenfalls sichergestellt werden. Hierfür sind – bezogen auf den Gebäudesektor – insbesondere eine Steigerung der energetischen Modernisierungsraten der Gebäudehüllbauteile von Bestandsgebäuden sowie eine Steigerung der Anteile erneuerbarer Energieträger an der Gebäudewärmeversorgung zwingend nötig (BCG 2021; Stiftung Klimaneutralität et al. 2021; Loga et al. 2018). Wie in der Einleitung dieser Arbeit bereits formuliert, wurde die Betrachtungsebene zur Steuerung und Umsetzung der Wärmewende in den vergangenen Jahrzehnten immer wieder über das Einzelgebäude hinaus in verschiedenen Kontexten auch auf Gebäudeensembles und Quartiere erweitert. Eine Diskussion existierender Definitionen und Anwendungen des Begriffs „Quartier“ sowie die Ableitung einer geeigneten Definition für die hier durchgeführten Untersuchungen liefern die nachfolgenden Abschnitte.

Im Umweltgutachten des Sachverständigenrates für Umweltfragen (SRU) wurde eine umfassende Analyse des Quartiers als „Raum für mehr Umwelt- und Klimaschutz“ dargestellt. Unter dem Begriff „Quartier“ wird dort „eine räumliche Einheit zwischen Gebäude- und Stadt(teil)ebene verstanden, deren Größe sich [kontextbezogen] nach der Eignung [...] richtet“ (SRU 2020, S. 406). Hinsichtlich des Forschungs- bzw. Handlungskontexts, aus welchem heraus das Quartier betrachtet werden kann, unterscheidet der SRU (2020) in seinem Gutachten vereinfacht zwischen einer sozialräumlichen und einer bau- bzw. infrastrukturellen Betrachtungsweise. Eine sozialräumliche Abgrenzung von städtischen Räumen zu Quartieren beruhe dabei in besonderem Maße auf identitätsstiftenden Merkmalen und der Wahrnehmbarkeit des eigenen Handelns für die im Quartier wohnenden bzw. agierenden Personen. Das Quartier wird unter diesen Gesichtspunkten auch als Kiez oder Stadtviertel benannt. Eine Abgrenzung von Quartieren aus sozialräumlicher Perspektive ergebe sich z. B. aus sich ändernden städtebaulichen Strukturen, wie der Bebauungsstruktur, oder natürlichen Barrieren im Stadtraum, wie Hauptstraßen oder Schienenwege. Die bau- und infrastrukturelle Betrachtungsebene wird im Hinblick auf Umwelt- und Klimaschutzaspekte beschrieben. Das Quartier ist demnach ein Verbund von Gebäuden und Infrastrukturen, dessen Abgrenzungskriterien und Größe im jeweiligen Handlungskontext definiert werden. Die unterschiedlichen Aufgaben der Stadt- und Infrastrukturplanung können es

deshalb erforderlich machen, jeweils verschiedene Teilräume als Quartier zu definieren, was zu einer nicht kongruenten Abgrenzung des Stadtraumes führt. „Die Quartiersebene [... vereint im bau- und infrastrukturellen Kontext] die strategische Suche nach zukunftsfähigen Lösungen, beispielsweise für die städtische Energieversorgung, mit der Umsetzung von konkreten und handhabbaren Maßnahmen“ (SRU 2020, S. 408). Eine übergeordnete Verallgemeinerbarkeit des Quartiersbegriffs ist deshalb nicht per se möglich und auch nicht notwendig. (Vgl. SRU 2020)

Die in der Analyse in SRU (2020) eher grobe Unterteilung der Perspektiven auf das Quartier als Handlungsraum soll zum besseren Verständnis der Komplexität des Quartiersbegriffs weiter verfeinert werden. Hierfür muss angemerkt werden, dass das Quartier neben seiner bisherigen Beschreibung als Handlungsraum, in welchem politische, sozialräumliche oder planerische Prozesse stattfinden, heute auch eine große Bedeutung in der Forschung hat. So finden sich unter dem Begriff der Quartiersforschung verschiedene Forschungsdisziplinen wieder, welche das Quartier selbst als Untersuchungsgegenstand definieren. Aber auch zur Beantwortung von ökonomischen, psychologischen, soziologischen, politikwissenschaftlichen oder natur- und ingenieurwissenschaftlichen Fragestellungen wird das Quartier häufig als räumliche oder nicht-räumliche Untersuchungseinheit bezeichnet. (Schnur 2008b)

Zwei wesentliche Beiträge zur raumbezogenen Analyse von sozioökonomischen und stadtplanerischen Fragestellungen, auf welche Untersuchungen der heutigen Quartiersforschung noch häufig verweisen, sind die Arbeiten von Jacobs (1976) und Castells (z. B. 1983). Sie verwenden in ihren jeweiligen (stadt-)soziologischen und sozioökonomischen Analysen die Begriffe „communities“, „communes“ und „neighborhoods“, um klassische, nicht raumbezogene Abgrenzungen von Teilen der Stadtgesellschaft vorzunehmen. Es erfolgt aber darüber hinaus auch die Betrachtung von Interdependenzen mit der Struktur des städtischen Raums und seiner Entwicklung, wofür dann die Begriffe „zones“ oder „districts“ verwendet werden.

Eine differenzierte Abhandlung über die Entstehung des Quartiers als Forschungsgegenstand und dessen Berücksichtigung in der wissenschaftlichen Modellbildung innerhalb unterschiedlicher soziologischer, ökonomischer und geographischer Forschungsdisziplinen kann z. B. in Schnur (2008a) und Cummings und Meisel (2013) gefunden werden. Schnur (2013) weist u. a. auf die Mehrdeutigkeit des Begriffs „Quartier“ im deutschen Sprachgebrauch hin, welcher sowohl die räumliche als auch die soziale bzw. sozioökonomische Abgrenzung (dann i. S. v. Nachbarschaften, einem Kiez, einer Gemeinschaft, aber auch einem Milieu) beschreiben kann (vgl. Schnur 2013, S. 26). Bezugnehmend auf eine Arbeit von Galster (1986) führt Schnur (2013) zehn raumbezogene Attribute an, anhand derer ein Quartier beschreibbar ist. Die Attribute benennt Schnur wie folgt:

1. *Bauliche Charakteristika (Bautypen, Modernisierungsgrad, Baudichte etc.)*
2. *Infrastrukturen (Straßen etc.)*
3. *Demographische Faktoren*
4. *Sozialer Status der Bevölkerung*
5. *Lokale Dienstleistungen (Quantität, Qualität)*
6. *Umweltfaktoren (Topographie, Emissionen etc.)*
7. *Erreichbarkeit (Topologie, Transportwesen etc.)*
8. *Politische Faktoren (politische Netzwerke, Partizipation etc.)*
9. *Soziale Interaktivität (soziale Netzwerke, Normen etc.)*
10. *Gefühlte Attribute (Ortsbindung, Quartiershistorie etc.)*

(Schnur 2013, S. 27–28)

Über diese Attribute hinaus diskutiert Galster (2001), ob eine universelle Abgrenzbarkeit des Quartiers als raumbezogene Einheit überhaupt möglich ist und leitet für sich ab, dass eine fallbezogene Abgrenzung von Quartieren notwendig ist:

Thus, my definition does not lead to the Holy Grail sought by much neighbourhood analysis of the 20th century: a means of unambiguously, meaningfully bounding urban neighbourhoods. [...] It suggests that the investigator would select a different parsing of urban space, depending on the particular neighbourhood attributes (or, equivalently, the neighbourhood typology) of interest.

(Galster 2001, S. 2113–2114)

Weiterhin liefert Galster (2001) mit dem Begriff „externality spaces“ – welcher im Folgenden als „Wirkungsbereich“ übersetzt wird – ein weiteres Konzept, welches das Quartier besser abgrenzbar machen soll. Der Wirkungsbereich ist demnach die räumliche Ausdehnung, innerhalb welcher eine Änderung von o. a. Attributen durch eine Person wahrgenommen wird und gleichzeitig Auswirkungen auf ihre Lebenssituation hat. Dies veranlasst Galster (2001) zur Formulierung der folgenden Merkmale für die Quartiersabgrenzung:

„Kongruenz“

der Grad, zu dem die Wirkungsbereiche mit anderen existierenden Grenzen übereinstimmen

„Allgemeingültigkeit“

der Grad, zu dem eine Überdeckung der Wirkungsbereiche verschiedener Einzelattribute vorliegt

„Übereinstimmung“

der Grad, zu dem eine Übereinstimmung der Wirkungsbereiche für verschiedene Personen in enger räumlicher Nähe vorliegt

(vgl. Galster 2001, S. 2114)

Laut Galster (2001) bildet die räumliche Ebene eines Baublocks oder einer geringen Anzahl von Baublöcken für einige der raumbezogenen Attribute die optimale räumliche Ausdehnung eines Quartiers. Die Merkmale „Kongruenz“ und „Übereinstimmung“ nehmen dann also jeweils ein hohes Maß an. Da dieser räumliche Zuschnitt allerdings hinsichtlich anderer Attribute bzw. Attributskombinationen weniger gute Ergebnisse bei den o. a. Bewertungsmerkmalen liefert, kann Galster (2001) keine weitere Eingrenzung seiner Quartiersdefinition liefern.

In seinen Ausführungen zu einer ‚Nachhaltigen Stadtplanung‘ formuliert Bott (2018) – beziehend auf die Arbeiten von Schnur (2008b), Vogelpohl (2008) und Lynch (2001) – das Quartier zudem als Überlagerung von drei Ebenen: einer städtebaulich-physischen, einer sozioökonomischen und einer symbolischen Ebene (Bott 2018, S. 21). Auch wenn er weiterhin eine überwiegend einheitliche Bautypologie als Merkmal eines Quartiers beschreibt, macht er deutlich, dass diese nicht als alleiniges Merkmal zur Quartiersbildung führt. Die baulich-räumliche Struktur konstituiert (aus einer stadtplanerischen) Sicht dennoch „das Grundgerüst, auf dem sich ein ausgeprägter Quartierscharakter entwickeln kann“ (Bott und Siedentrop 2018, S. 46). Die Stadt als Raumebene, welche eine Wahrnehmung als Ganzes erschwert, kann zudem als eine Collage begriffen werden, welche sich aus verschiedenen Quartieren zusammensetzt. Jedes einzelne Quartier steht dabei in Beziehung zu seinen Nachbarquartieren, bleibt aber dennoch identifizierbar und unterscheidbar (Bott und Siedentrop 2018, S. 46).

Wichtig scheint zudem der Hinweis, dass weder das Baugesetzbuch noch die verschiedenen städtebaulichen Förderprogramme klare Abgrenzungskriterien liefern, sondern die Forderung nach der Zweckmäßigkeit der Abgrenzung für den individuellen Anwendungsfall in den Vordergrund stellen (Bott 2018, S. 22–23). Da die Anwendungsfälle wiederum unterschiedliche Disziplinen der Quartiersforschung und verschiedene Dimensionen der Stadtplanung, des Verwaltungshandelns u. Ä. umfassen können, bleibt die bereits benannte Unschärfe des Begriffs „Quartier“ demnach bestehen. Auch hinsichtlich der Größenausdehnung bzw. der Anzahl der im Quartier befindlichen Gebäude können aus den angeführten Quellen keine Hinweise abgeleitet werden (vgl. auch Neitzel 2013). Aus diesem Grund soll vor der Formulierung eines eigenen Quartiersbegriffs für die Entwicklung einer Methodik zur Klassifizierung von Typquartieren die Berücksichtigung des Quartiers in der gebäudebezogenen Energieforschung genauer betrachtet werden. Damit soll der Kontext der erforderlichen Klimaschutzmaßnahmen im Gebäude- und Quartiersbestand eine hinreichende Berücksichtigung finden.

2.2 Das Quartier in der Energieforschung

Auch in der gebäudebezogenen Energieforschung verbindet das Quartier die räumlich-physische Ebene der gebauten Umwelt mit der sozioökonomischen und weiteren Forschungs- sowie Handlungsebenen. Wie die nachfolgende Diskussion verschiedener Quellen zeigen soll, liegt der Fokus hierbei jedoch stärker auf den materiellen Attributen, z. B. der Bebauungsstruktur, bauphysikalischen und gebäudetechnischen Eigenschaften des Quartiersgebäudebestandes sowie eventuell vorliegenden Energieinfrastrukturen. Motiviert durch verschiedene nationale und europäische Förderprogramme hat sich das Quartier in den vergangenen Dekaden in der gebäudebezogenen Energieforschung etabliert. Viele der Vorhaben untersuchten ingenieurtechnische Fragestellungen und erweiterten dadurch die seit Mitte des 20. Jahrhunderts stark sozioökonomisch geprägte Quartiersforschung. (Vgl. Cummings und Meisel 2013; Hentschel 2017; Erhorn et al. 2010)

So wurde im Forschungsprogramm „Experimenteller Wohnungs- und Städtebau“ (ExWoSt) des Bundesministeriums für Verkehr, Bau und Stadtentwicklung beispielsweise im Vorhaben „Anforderungen an energieeffiziente und klimaneutrale Quartiere“ (EQ) der Status quo der energetischen Stadterneuerung mit dem Quartier als Handlungsebene für die Erarbeitung und Umsetzung von Energiekonzepten untersucht. Das Forschungsdesign war sehr praxisorientiert gestaltet und erarbeitete eine Mess- und Bewertungsmethodik für energetische Analysen auf der Ebene des Quartiers (von Malotki et al. 2012). Im sich daran anschließenden Folgevorhaben EQ II wurden verschiedene Bilanzierungsverfahren und -werkzeuge gegenübergestellt, die vormals entwickelte Methodik erweitert und in einem Berechnungswerkzeug implementiert (von Malotki et al. 2016).

Ergänzend zur Quartiersforschung im Rahmen des Förderprogramms ExWoSt verstärkte auch das Bundesministerium für Wirtschaft und Technologie (BMWi) ab dem Jahr 2008 seine Energieforschung mit Quartiersbezug durch die Zusammenführung von Förderlinien des 5. Energieforschungsprogramms zum Forschungsschwerpunkt „Energieeffiziente Stadt“ (EnEff:Stadt). Im Rahmen dieses Förderkonzepts wurden Demonstrationsprojekte für eine Verbesserung der Energieeffizienz sowie für die Einbindung erneuerbarer Energieträger in die Versorgung ganzer Siedlungsgebiete bzw. Stadtteile bearbeitet. Zudem konsolidierte eine wissenschaftliche Begleitforschung das Wissen zu technischen und organisatorischen Aspekten der praktischen, quartiersbezogenen Energieforschung. Eine der Forschungsinitiative zugehörige Vernetzungsstrategie sollte darüber hinaus weitere Impulse für die quartiersbezogene Energieforschung liefern (vgl. Erhorn et al. 2010, S. 10–11). Die Vielzahl der untersuchten Siedlungsstrukturen sowie die berücksichtigten Technologien zur Effizienzsteigerung und/oder für den Einsatz erneuerbarer Energien in der Wärme- und Stromversorgung ist in Abbildung 2-1 (S. 14) dargestellt und veranschaulicht die inhaltliche Breite der quartiersbezogenen Energieforschung in den insgesamt drei Phasen des Förderkonzepts zwischen 2007 und 2015 (vgl. Beier et al. 2016).

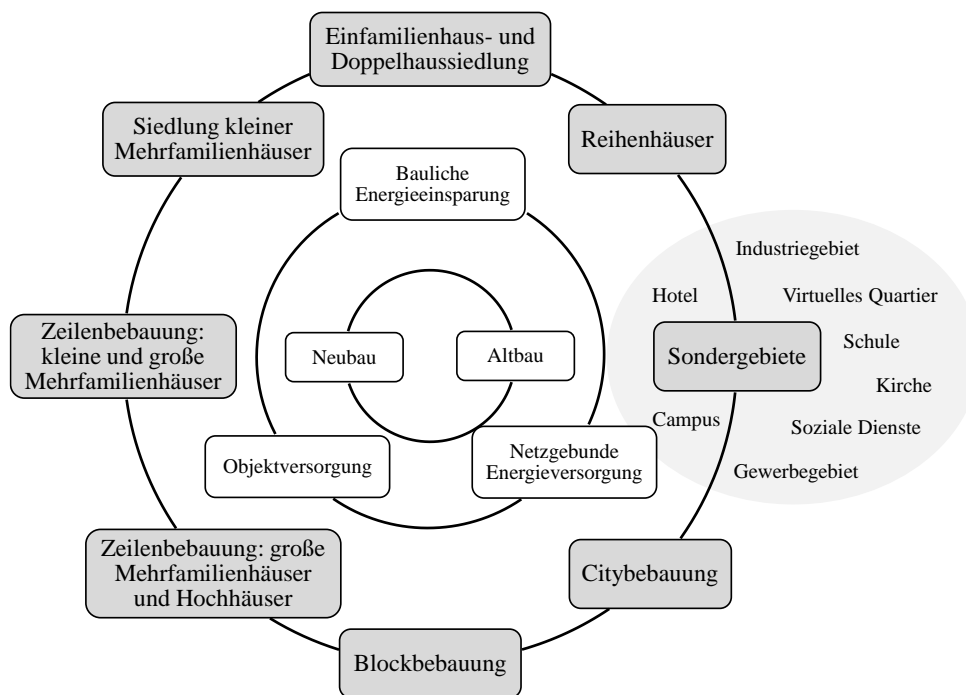


Abbildung 2-1: Übersicht der EnEff:Stadt-Demonstrationsvorhaben und der darin behandelten Siedlungstypen (äußere Ebene), der untersuchten Versorgungskonzepte (mittlere Ebene) sowie der Projekttypen (innere Ebene) [eigene Darstellung in Anlehnung an Beier et al. (2016, S. 30–31)]

Nach dem Ende der EnEff:Stadt-Förderung waren Quartierskonzepte weiterhin als Teil des 6. Energieforschungsprogramms Untersuchungsgegenstand und Kern vieler Demonstrationsvorhaben. Insbesondere die Förderinitiative „EnEff.Gebäude.2050 – Innovative Vorhaben für den nahezu klimaneutralen Gebäudebestand 2050“ (EG2050) und das Förderkonzept für „Reallabore der Energiewende“ erlauben heute die Erprobung innovativer technischer und organisatorischer Konzepte zur Wärme- und Stromversorgung unter Einbeziehung einer höheren räumlichen Ebene, z. B. dem Quartier oder einem gesamten Stadtgebiet, um Wege zur Erreichung der Ziele der Energiewende im Gebäudebereich¹ aufzuzeigen. (BMWi 2018, 2021)

Das Quartier erfährt auch im „Forschungsnetzwerk Energiewendebauen“ – einer Plattform zur besseren Vernetzung und zum koordinierten Wissenstransfer innerhalb der gebäudebezogenen Energieforschung – sowie in der aktuellen wissenschaftlichen Begleitforschung „Energiewendebauen“ des Bundesministeriums für Wirtschaft und Klimaschutz (BMWK) als eines von vier Modulen und eigenständiges Themenfeld der anwendungsorientierten

¹ Das Förderkonzept „Reallabore der Energiewende“ geht dabei über den üblichen Themenbereich Wärme- und Stromversorgung hinaus und fördert systemische und sektorenübergreifende Ansätze, welche auch die Themen Mobilität, Digitalisierung etc. beinhalten können (vgl. BMWi 2021).

Energieforschung besondere Beachtung. Die Gründe für eine Betrachtung des Quartiers als Forschungs- und Handlungsebene werden dabei wie folgt zusammengefasst:

Neben der energetischen Optimierung einzelner Gebäude birgt die ganzheitliche Betrachtung städtischer Siedlungsräume ein großes Potenzial zur Steigerung der Energieeffizienz. Städte [...] entscheiden über Versorgungsstrukturen, Bebauungspläne und energetische Standards in öffentlichen Gebäuden. Neue Technologien und moderne Planungsinstrumente können diesen Gestaltungsspielraum deutlich erweitern. Das Quartier ist die Bezugsebene, auf der alle städtischen Funktionen wie Wohnen, Dienstleistungen, Gewerbe, Infrastrukturen und Verkehr verortet sind. Hier werden technologische, sozialräumliche und ökonomische Prozesse sowie Energie- und Stoffströme sichtbar und steuerbar.

(Begleitforschung EWB 2022)²

Dass sich das Quartier über den Forschungskontext hinaus als Handlungsebene für Energieeffizienz und Klimaschutz etabliert hat, spiegelt sich in der Berücksichtigung dieser Ebene in den Landesgesetzgebungen, innerhalb des Gebäudeenergiegesetzes (GEG) sowie in den Klimaschutzaktivitäten der Europäischen Kommission wider. In einigen Bundesländern wird bereits die Berücksichtigung des Quartierszusammenhangs bei der Planung und Bewertung von städtischen Energiekonzepten und Infrastrukturen gefordert und/oder Planungshilfen für eine solche, über das Einzelgebäude hinaus erweiterte Betrachtung bereitgestellt (Köhler und Bürger 2022). Auf europäischer Ebene wird das Quartier als räumliche Ebene für die schnelle Steigerung der Gebäudemodernisierungsrate sowie einer deutlichen Erhöhung der Gesamteffizienz (d. h. Energiebedarfe und Energieerzeuger- und -verteilungsstrukturen) betrachtet und soll deshalb im Rahmen der „Renovierungswelle für Europa“ ebenfalls besondere Beachtung finden (EU Kommission 2020). Darüber hinaus sind die in den folgenden Abschnitten vorgestellten Studien ebenfalls als Ergebnisse der Energieforschung am Untersuchungsgegenstand „Quartier“ zu verstehen.

2.3 Definition eines kontextbezogenen Quartiersbegriffs

In Anlehnung an einige der bereits vorgestellten Arbeiten soll der Quartiersbegriff für die hier zu untersuchende Fragestellung definiert und dadurch eine Operationalisierung im Rahmen der Methodenentwicklung und Geodatenanalyse ermöglicht werden.

² Für eine übersichtliche Auswertung von Quartiersvorhaben wird auf die Veröffentlichung von Erhorn-Kluttig und Erhorn (2016) verwiesen.

Das Quartier wird in dieser Arbeit als eine abgrenzbare Raumebene verstanden, welche durch eine Ansammlung von Gebäuden des gleichen oder unterschiedlichen Bautypus in einem direkten räumlichen Zusammenhang gebildet wird. Es stellt eine zentrale Ebene dar, auf welcher sich das Handeln von Akteuren im Hinblick auf Klimaschutzmaßnahmen im Gebäudebestand erfahren und bewerten lässt. Der Fokus wird dabei auf die baulich-räumliche Ebene und die darin bestehenden „raumbevölkernden Objekte“, wie die Gebäude, Grundstücke und Baublöcke, gelegt.³ Durch deren Merkmale ist die baulich-räumliche Struktur des Quartiers mess- und beschreibbar und dadurch eine Erfassung des Quartiers möglich. Innerhalb des Quartiers sind die baulich-räumlichen Merkmale der Gebäude und Baublöcke zwar überwiegend homogen. Dies bedeutet jedoch nicht zwangsläufig, dass keine Gebäude anderen Bautyps oder unterschiedliche Baublockkonfigurationen vorliegen können. Eine Abgrenzung zu Nachbarquartieren erfolgt über den Vergleich der jeweiligen baulich-räumlichen Strukturmerkmale der die Quartiere bevölkernden Objekte oder aufgrund übergeordneter Grenzelemente, wie Hauptverkehrswege oder relevante Verwaltungsgrenzen.

Den kleinsten Maßstab eines Quartiers stellt somit die Baublockebene und die Anzahl der in diesem Baublock befindlichen Gebäude dar. Um eine deutliche Abgrenzung zur Ebene des Einzelgebäudes (welches z. B. alleine auf einem Baublock verortet sein kann) zu erreichen und um der Zielsetzung, die quartiersbezogenen Klimaschutzmaßnahmen in Deutschland besser bewertbar zu machen, gerecht zu werden, erscheint zudem eine Begrenzung der Mindestanzahl an Gebäuden in einem Quartier als sinnvoll. Auch hinsichtlich seiner Gesamtausdehnung ist ein Quartier beschränkt, um als intermediäre Einheit zwischen Einzelgebäuden und Stadtteilen oder z. B. vollständigen, großräumigen Versorgungsgebieten funktionieren zu können. Die beiden das Quartier begrenzenden Parameter „Mindestgebäudezahl“ und „maximale Flächenausdehnung“ lassen sich jedoch nicht aus der Literatur ableiten und sind stark kontextbezogen. Deshalb soll im Rahmen dieser Methodenentwicklung eine genauere Untersuchung dazu erfolgen.

Um die voranstehende Beschreibung des Quartiersbegriffs zu schärfen, sollen nachfolgend die als „raumbevölkernde Objekte“ beschriebenen Elemente „Grundstück“, „Baublock“ und „Gebäude“ definiert werden.

³ Über den Einfluss von sozioökonomischen Prozessen in Quartieren ist sich der Autor zwar bewusst, im Rahmen der geoinformationsdatenbasierten Untersuchung werden diese jedoch in dieser Arbeit nicht berücksichtigt.

Da der Begriff des Grundstücks im allgemeinen Sprachgebrauch eindeutig und nicht missverständlich ist, soll hierfür nur eine verkürzte Definition gegeben werden:

Das Grundstück ist ein räumlich abgegrenzter Teil der Erdoberfläche, der auf einem besonderen Blatt, dem Grundbuchblatt, für sich allein oder auf einem gemeinschaftlichen Grundbuchblatt unter einer eindeutigen Nummer des Bestandsverzeichnisses eingetragen ist (Grundstück im Rechtssinn).

(AdV 2018b, S. 193; siehe auch GBO, vom 05.10.2021)

Der Baublock ist gemäß der Definition des Verbandes Deutscher Städtestatistiker und des Deutschen Städtetags (1991) Teil eines „allgemeine(n), flächendeckende(n) Lokalisierungs- und Zuordnungssystem für statistisch-planerische und Verwaltungszwecke. [...] Blöcke werden in der Regel von Straßen und natürlichen oder baulichen Grenzen [...] oder Verwaltungsgrenzen höherer Ebene] von allen Seiten umschlossen“ (Deutscher Städtetag 1991, S. 16). Da neben den genannten Grenzen aber auch andere Kriterien (z. B. Nutzungsarten oder ordnungsrechtliche Gründe) in die Abgrenzung der Baublöcke einfließen können, ist auch ein direktes Angrenzen zweier Baublöcke generell möglich (vgl. Bömermann 2012, S. 82). Der Baublock ist somit eine weitestgehend klar abgegrenzte flächenhafte Einheit, welche „das gesamte bebaute und unbebaute Gemeindegebiet“ (Deutscher Städtetag 1991, S. 16) umfasst und für verschiedene statistische, planerische oder Verwaltungszwecke verwendet werden kann (vgl. Deutscher Städtetag 1991; Bömermann 2014, 2012). Diese Baublockbeschreibung soll auch in der vorliegenden Arbeit Verwendung finden – u. a., weil nahe liegt, dass der Baublock mit dieser Definition auch in den GIS-Daten der Katasterbehörden abgebildet ist.

Gebäude sind im baurechtlichen Sinne „selbstständig benutzbare, überdeckte bauliche Anlagen, die von Menschen betreten werden können und geeignet oder bestimmt sind, dem Schutz von Menschen, Tieren oder Sachen zu dienen“ (MBO, vom 21.09.2012, § 2 Abs. 2). Als bauliche Anlagen sind sie gemäß § 2 Abs. 2 S. 1 MBO mit dem Erdboden verbunden und aus Bauprodukten hergestellt. Gebäude lassen sich darüber hinaus anhand baulicher Parameter sowie der Nutzung in verschiedene Gebäudeklassen einteilen und in Regel- und Sonderbauten unterscheiden (MBO, vom 21.09.2012, § 2 Abs. 3). Das Gebäude kann zudem im Sinne der Erfassungskriterien für die Bautätigkeitsstatistik auch ein unterirdisches Bauwerk zu o. g. Zwecken sein und muss nicht zwangsläufig von Wänden umgeben sein (vgl. IT.NRW 2010, S. 4; Hecht 2014, S. 19–20).

[Als] einzelne Gebäude [gelten in diesem Kontext] freistehende Gebäude oder bei zusammenhängender Bebauung – z. B. Doppel- und Reihenhäuser – jedes Gebäude, das durch eine vom Dach bis zum Keller reichende Brandmauer von anderen Gebäuden getrennt ist. Ist keine Brandmauer vorhanden, so gelten die zusammenhängenden Gebäudeeinheiten als einzelne Gebäude, wenn sie ein eigenes Erschließungssystem (eigener Zugang und eigenes Treppenhaus) besitzen und für sich benutzbar sind.

(IT.NRW 2010, S. 4)

Zudem wird für statistische Zwecke sowie in den Geobasisdaten Deutschlands eine Unterteilung der Gebäude nach Nutzungsart in drei Hauptkategorien vorgenommen: Wohngebäude, Nichtwohngebäude und Gebäude von unmittelbarem öffentlichen Interesse (vgl. DESTATIS 2020, S. 5; AdV 2018a, 2021b). Neben der Unterteilung nach Nutzungen ist eine Abgrenzung zweier Gebäude auch über die Art der Erschließung möglich. Unter Einbezug einer eigenständigen Erschließung als notwendiges Gebäudemerkmal sind zwei nach den eingangs beschriebenen Kriterien eigenständige Bauwerke (d. h. betretbar, überdacht, trennende Brandwand ist vorhanden und es liegen z. B. unterschiedliche Nutzungen vor) nicht als einzelne Gebäude zu betrachten, wenn eine gemeinsame Erschließung vorliegt. In diesem Fall würde es sich vielmehr um ein Gebäude mit mehreren Gebäudeteilen handeln. Ob und wie das Merkmal einer eigenständigen Gebäudeerschließung im Rahmen dieser geodatenbasierten Untersuchung berücksichtigt werden kann, muss jedoch im weiteren Verlauf der Arbeit überprüft werden (siehe Kapitel 5.2 und Kapitel 6.2). Aus diesem Grund wird ein Gebäude im Rahmen dieser Arbeit nach den oben aufgeführten Kriterien, aber unabhängig von der Erschließungsart, definiert.

Die Definition des Quartiersbegriffs für diese Arbeit soll zudem im Rahmen der Entwicklung und Anwendung der Methodik zur Quartierstypisierung (siehe Kapitel 5 und Kapitel 6) weiter präzisiert werden, um bezüglich des Beitrags von Quartieren zum Klimaschutz im Gebäudebestand Aussagen mit einem möglichst hohen Nutzen für die Planungspraxis von Quartieren zu erreichen. In verschiedenen Forschungsvorhaben konnte gezeigt werden, dass eine quartiersbezogene Betrachtung energetischer Modernisierungsmaßnahmen und effizienter Wärme- und Stromversorgungskonzepte insbesondere für Quartiere mit einer vergleichsweise hohen Anzahl an Gebäuden je Einheit Siedlungsfläche zielführend ist (vgl. Beier et al. 2016; Erhorn et al. 2010). Auch für Energiekonzepte in ländlichen Bebauungsstrukturen wurde deren Mehrwert hinsichtlich der Reduktion von THG-Emissionen gezeigt (vgl. Ruppert et al. 2008; Pietruschka et al. 2015). Dennoch wird für diese Arbeit die These formuliert, dass bei einem Anteil der Wohneinheiten in Deutschland von rund 70 % in

kreisfreien Städten und städtischen Kreisen (DESTATIS et al. 2021, S. 262)⁴ die höchsten absoluten Einsparungspotenziale in Deutschland im verdichteten Gebäudebestand im urbanen Kontext zu erreichen sind. Der Untersuchungsgegenstand dieser Studie sind deshalb Quartiere mit städtischen oder stadähnlichen Bebauungsstrukturen, d. h. einer vergleichsweise hohen Dichte an Gebäuden je Einheit Siedlungsfläche.

Nach der Herleitung und Definition des Quartiersbegriffs an dieser Stelle muss im Rahmen der späteren Methodenentwicklung und Geodatenanalyse eine Operationalisierung erreicht werden. Hierfür ist zuerst eine weiterführende Untersuchung zu den in Deutschland verfügbaren GIS-Daten nötig (siehe Kapitel 4). Darauf muss ein Ablauf zur Datenaufbereitung und -analyse erarbeitet werden, welcher Gebäude, Teilbauflächen und Quartiere unter Berücksichtigung der formulierten Definitionen abgrenzen kann und sowohl die Typisierung als auch eine statistische Auswertung erlaubt (siehe Kapitel 5 und Kapitel 6).

2.4 Existierende Ansätze zur Typisierung von Quartieren

2.4.1 Siedlungsstrukturtypen nach Roth (1980)

Die Stadtplanung in Deutschland und somit auch die Entwicklung der baulichen Strukturen in Städten folgte in den vergangenen Jahrzehnten und Jahrhunderten verschiedenen Prinzipien und Leitbildern. Diese berücksichtigten in Teilen zwar auch energiebezogene Aspekte, fokussierten dabei jedoch hauptsächlich andere Grundprinzipien und epochenspezifische Vorstellungen von Städtebau und Stadtentwicklung (Albers und Wékel 2021). Keines der stadtplanerischen oder städtebaulichen Leitbilder bis Mitte bzw. Ende des 20. Jahrhunderts beinhaltete eine grundlegende Auseinandersetzung mit den Zusammenhängen von Stadtstruktur und Gebäudenutzwärmebedarfen sowie der Effizienz der Raumwärmebereitstellung (vgl. Everding und Kloos 2007). In den 1970er Jahren rückte als Folge der weltweiten Ölpreiskrise diese Thematik in den Vordergrund (Erhorn et al. 2010, S. 12–15; Müller 2021, S. 467–468).

In einer umfassenden Studie widmete sich Roth (1980) erstmals der Fragestellung, „in welchen typischen Siedlungsgebieten der Bundesrepublik Deutschland welche Wärmeversorgungssysteme in Zukunft am sinnvollsten einzusetzen sind [...]“. Die Untersuchung

⁴ In Behnisch (2008, S. 250–253) wird als Ergebnis einer raumbezogenen Datenanalyse geschätzt, dass sich etwa 70 % der Gebäude in Deutschland in den Ober-, Mittel- und Unterzentren nach dem Zentralen-Orte-Konzept der deutschen Raumplanungspraxis befinden (vgl. ROG, vom 22.12.2008). Eine ebenfalls von Behnisch durchgeführte Analyse unter Einbezug der raumstrukturellen Merkmale ergibt eine Abschätzung, dass rund 60 % des Gebäudebestandes in Gemeinden mit mehr als 10.000 Einwohnern vorliegt.

2 Betrachtungsraum „Quartier“

erfolgt(e) anhand von neun Siedlungstypen, deren Häufigkeit in der Bausubstanz der Bundesrepublik Deutschland mit statistischen Methoden festgestellt wurde“. (Roth 1980, S. 21)

Im Rahmen der Analyse fand eine Bewertung der Wirtschaftlichkeit von bautechnischen Maßnahmen (d. h. den Einsatz von Wärmedämmung) und Wärmeversorgungs Konzepten für die als Vertreter des Gebäudebestandes definierten neun „Siedlungstypen“ (siehe Abbildung 2-2) bei gleichzeitiger Differenzierung nach Altbauten und Neubauten als zu bewertende Gebäudeklassen statt. Da Roth seine Untersuchungen auf der Ebene der Siedlungsstrukturtypen (ST) durch- und in eine übergeordnete Bewertung der Raumstruktur überführte, entwickelte er ein „Raum-Siedlungsstrukturmodell“, welches im Folgenden kurz dargestellt wird. Details zu den Grundlagen und Annahmen des Modells werden zudem in Anhang A beschrieben.

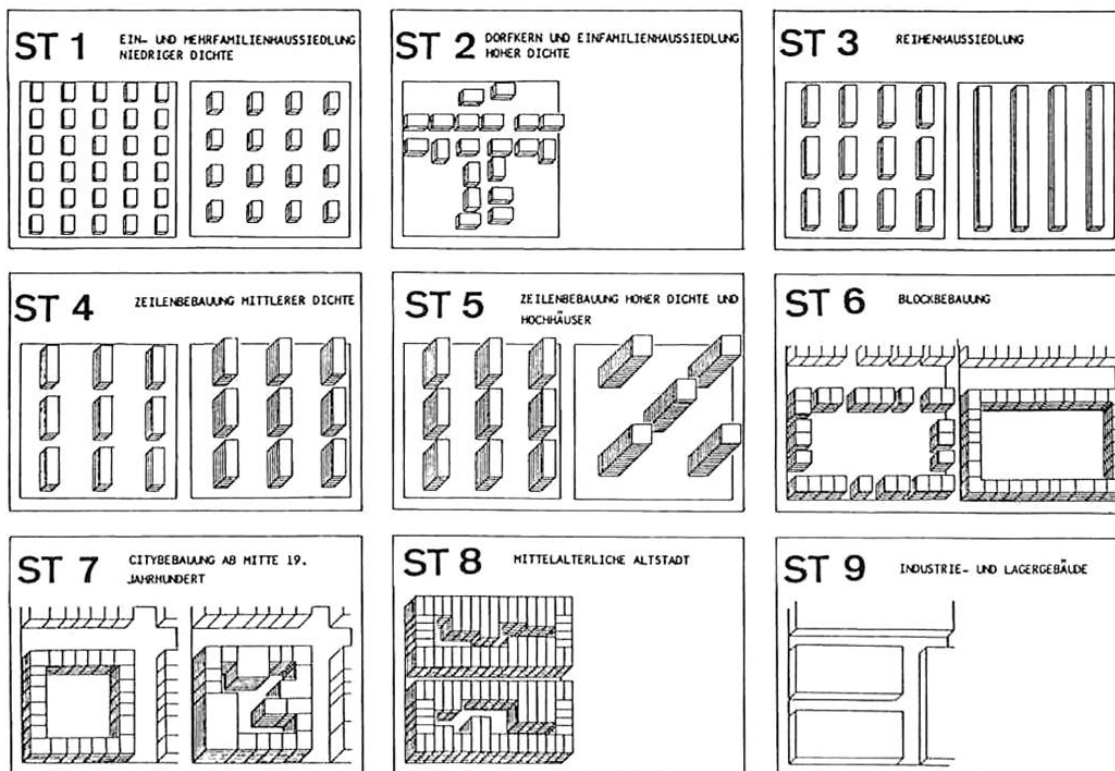


Abbildung 2-2: Siedlungsstrukturtypen nach Roth (1980) (siehe auch Tabelle Anhang A-1)[Roth und Häubi (1981, S. 972)]

Die Definition der neun Siedlungsstrukturtypen erfolgte durch Roth maßgeblich auf Basis einer ingenieurmäßigen und energiebezogenen Unterteilbarkeit der Bebauungsstrukturen in Deutschland und ist geprägt durch „direkt wärmeversorgungsrelevante bauliche und siedlungstechnische Merkmale wie Bauform, Baukonstruktion, Wärmedichte etc. Als Maßstab werden zusammenhängend bebaute Gebiete homogener Siedlungsstruktur (Zellen) gewählt [...]“ (Roth 1980, S. 49). Die Strukturtypen wurden dabei so abgegrenzt, dass sie auf topografischen Karten mit dem Maßstab 1:25.000 erkennbar sein sollten. Es erfolgte zudem

die Zuordnung eines Datenprofils mit strukturbeschreibenden Merkmalen zu jedem der Strukturtypen ST1 – ST9. Da es sich dabei um eine abstrakte Modellierung der gebauten Umwelt handelt, finden sich in den Datenprofilen meist Wertebereiche wieder, welche in der späteren Modellierung von Roth als Mittelwerte in die Berechnungsvorschriften eingehen. Die zur Differenzierung und Beschreibung der Siedlungsstruktur verwendeten Merkmale sind nachfolgend aufgeführt:

- Siedlungsspezifische Daten:
Geschosszahl, Geschossflächenzahl (GFZ) und Grundflächenzahl (GRZ), die Zahl der Gebäude je km², Baualter, Dachform, Beschattung durch Pflanzen
- Baulich-geometrische Daten:
Beheizte Nutzfläche, Oberflächen/Volumenziffer (F/V; heute als A/V-Verhältnis bezeichnet), Anteil Wand inkl. Fenster an der Gebäudeoberfläche, Anteil Fenster an der Außenwandfläche
- Baulich-technische Daten:
Wärmedurchgangskoeffizienten der Wand ohne Fenster, der Fenster, des Daches, der Bodenplatte bzw. der Kellerdecke
- Wärmetechnische Daten:
Mittlerer k-Wert (bei Altbauten; heute durch den U-Wert ersetzt), k_m -Sollwert⁵, Nutzwärmehöchstleistung, Anschlusswert nach DIN 4701, Höchstleistungswärmedichte (je km² Fläche)

(vgl. Roth 1980, S. 97–98)

Für die Überprüfung der Relevanz der festgelegten ST erfolgte zudem eine beispielhafte Auswertung topografischer Karten der Stadt Freiburg und eine manuelle Zuordnung der städtischen Bebauung zu den ST, welche im Anschluss anhand drei zufällig ausgewählter Baublöcke je Siedlungstyp überprüft wurde. Durch diese Überprüfung konnten zudem die Ausprägungen der raumbeschreibenden bau- und energietechnischen Merkmale mit den Modellannahmen der Datenprofile abgeglichen werden. (vgl. Roth 1980, S. 50)

Um eine Verallgemeinerung der Ergebnisse der Wirtschaftlichkeitsbetrachtungen von der Ebene der ST auf Deutschland zu ermöglichen, führte Roth zudem eine Untersuchung der

⁵ Dies „bezeichnet die Minimalanforderung der Wärmeschutzverordnung aufgrund des Oberflächen/Volumen-Verhältnisses“ (Roth 1980, S. 98) und ist in etwa vergleichbar mit dem heute im GEG als Anforderungsgröße definierten spezifischen, auf die wärmeübertragende Umfassungsfläche bezogenen Transmissionswärmeverlust (GEG, vom 08.08.2020).

Raumstruktur, als großräumliche Beziehung zwischen Siedlungen, durch und gliederte die Siedlungs- und Raumstruktur in vier Raumtypen (RT):

RT 1: Kernstädte von (großen) Stadtregionen

RT 2: Außenzonen von Stadtregionen

RT 3: Mittelstädte über 10.000 Einwohner außerhalb von Stadtregionen

RT 4: Dörfliche Siedlungen im ländlichen Raum

(vgl. Roth 1980, S. 41)

Aus den in den vier RT befindlichen Städten und Gemeinden wurde in einem zweiten Schritt eine Stichprobe von etwa 1 % der Nettosiedlungsfläche in Deutschland gezogen und für diese Städte und Gemeinden ebenfalls eine Auswertung von Bildkarten vorgenommen. Mit Hilfe von Korrelations- und Regressionsanalysen erfolgte anschließend die Untersuchung der raumstrukturellen Merkmale, für welche ein Zusammenhang mit der Siedlungsstruktur vermutet wurde. Die abgeleiteten Regressionsgleichungen erlaubten eine Schätzung der Nettosiedlungsfläche je SR im jeweiligen RT für Westdeutschland inkl. West-Berlin. Fehlerrechnungen zeigten laut Roth Standardfehler bei der Quantifizierung der ST je RT von 5 % bis 10 %.

Der Beschreibung der Siedlungs- und Raumstruktur durch ST und RT folgten eine Analyse der Möglichkeiten zur Wärmeeinsparung in Gebäuden durch bauliche Maßnahmen sowie ein Vergleich der verfügbaren Technologien und Kosten für die Gebäudewärmeversorgung. Eine betriebswirtschaftliche Optimierung von Wärmeschutz und Wärmeerzeugungsvarianten erlaubte die Evaluation wirtschaftlicher Wärmeversorgungssysteme je ST und eine Schätzung der Auswirkungen auf die damalige Struktur der Wärmeversorgung.

2.4.2 Weitere gebäudezentrierte Quartierstypologien

Seit der Erstellung der oben beschriebenen Studie von Roth (1980) gab es in den vergangenen vierzig Jahren immer wieder Untersuchungen, welche das Konzept der Strukturierung und Typisierung der Bebauungsstruktur im Kontext der jeweiligen Forschungsfrage aufgriffen. Die bedeutendsten Weiterentwicklungen werden nachfolgend beschrieben. Dabei kann die Bedeutsamkeit sowohl in einer bemerkenswerten inhaltlich-methodischen Weiterentwicklung liegen als auch in der Häufigkeit, in der die Arbeit in der aktuellen energiebezogenen Quartiersforschung Anwendung findet.

In seiner Forschung über die „räumlich hoch aufgelöste Modellierung leitungsgebundener Energieversorgungssysteme zur Deckung des Niedertemperaturwärmebedarfs“ differenziert Blesl (2002) die von Roth (1980) erarbeitete Typologie der Siedlungsstrukturen weiter aus. Das Ziel seiner Arbeit bestand in einer kleinräumigen Untersuchung von Wärmeversorgungskonzepten (z. B. dem Anschluss von städtischen Teilgebieten an eine leitungsgebundene Wärmeversorgung). Die computergestützte Analyse verfügbarer topografischer

Bildkarten ermöglichte eine Verfeinerung der Klassifizierung und erlaubte somit eine belastbarere Modellierung der wärme- und kostentechnischen Zusammenhänge. Die Analysen stützten sich dabei neben den Berechnungen mit von ihm definierten charakteristischen Merkmalen für die nun insgesamt 18 Siedlungstypen⁶ auch auf die Auswertung der Daten für eine von ihm festgelegte Modellgemeinde. (vgl. Blesl 2002)

Da die kleinräumige Betrachtung eine genauere Bestimmung der Leitungslängen einer netzgebundenen Wärmeversorgung innerhalb der Siedlungstypen erforderlich macht, liefert Blesls Typologie auch diesbezüglich weiterführende charakteristische Merkmale. Die grafische Darstellung der Siedlungstypen ähnelt den Abbildungen von Roth (1980), akzentuiert allerdings die stärkere Bedeutung der Straßenzuglängen für die leitungsgebundene Wärmeversorgung (siehe Abbildung 2-3)⁷.

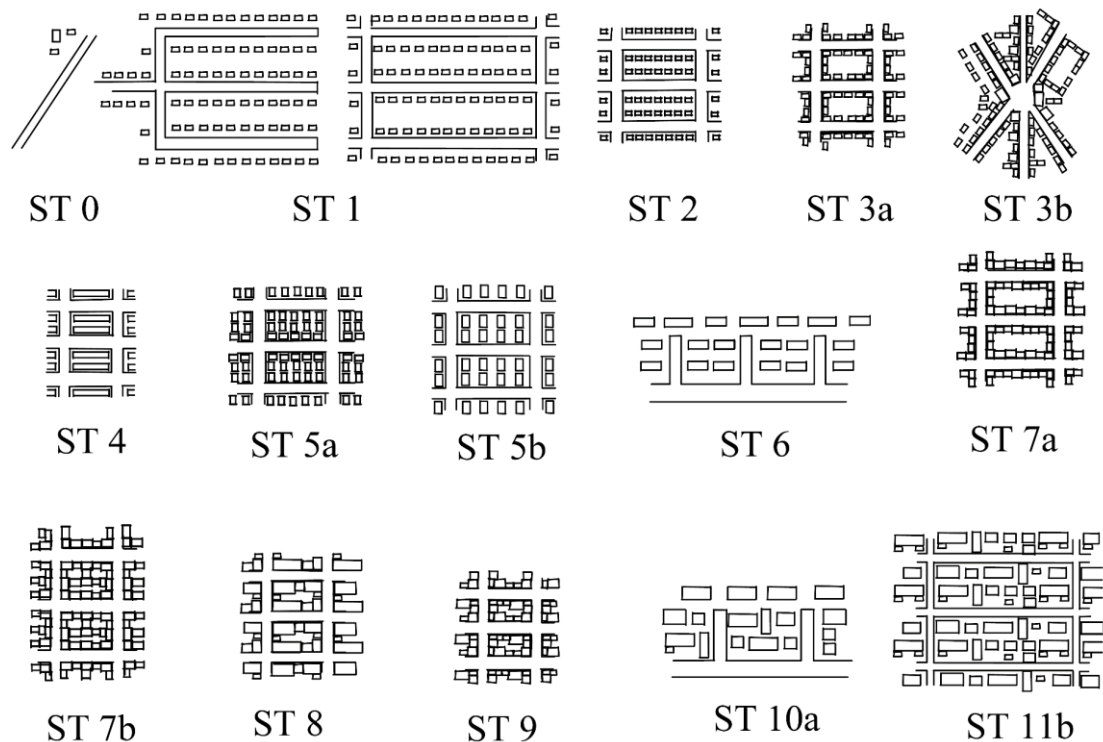


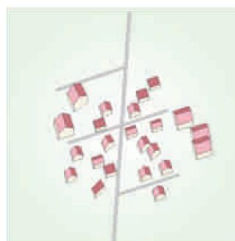
Abbildung 2-3: Siedlungstypen nach Blesl (2002) [Blesl (2002, S. 143)]

⁶ Eine Beschreibung sowie die Übersicht der charakteristischen Merkmale der Siedlungstypen findet sich in Anhang A.

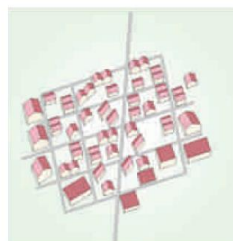
⁷ Für die Siedlungstypen 10b, 11a und 12 liegen in Blesl (2002) keine Darstellungen vor.

Weitere Arbeiten von Everding und Kloos (2007) sowie Everding und Lindner (2007) griffen das Werkzeug der Typisierung städtischer Bebauungsstrukturen auf und untersuchten anhand von 20 Stadtraumtypen die Potenziale einer passiven und aktiven Solarenergienutzung in Bestandsbauten. Die Methodenentwicklung und eine allgemeinere Untersuchung findet sich in Everding und Kloos (2007); in Everding und Lindner (2007) wird darüber hinaus die Erarbeitung eines „solaren Rahmenplans“ für die Stadt Berlin vorgestellt. Dieser Rahmenplan sollte die Potenziale für die solare Nutzung innerstädtischer Flächen abschätzen und Handlungsempfehlungen für eine nachhaltige Stadtentwicklung zur Erreichung lokaler Klimaschutzziele liefern.

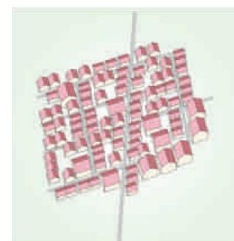
Eine weitere sehr detaillierte Auseinandersetzung mit dem Quartiersbegriff und der Typologisierung von Stadträumen aufgrund der Bebauungsstruktur liefern Erhorn et al. (2010). Unter dem Titel „Energetische Quartiersplanung in Deutschland“ werden die Ergebnisse der wissenschaftlichen Begleitforschung für die Forschungsinitiative „Energieeffiziente Stadt“ (EnEff:Stadt) zusammengefasst und eine breite Übersicht über den damaligen Forschungsstand gegeben. Über die reine Beschreibung der existierenden Forschungsergebnisse hinaus nehmen Erhorn et al. (2010) zudem eine eigene Typisierung der Siedlungsstrukturen in Deutschland vor (siehe Abbildung 2-4). In Anlehnung an die Arbeiten von Roth (1980) und Blesl (2002), aber mit dem Ziel einer besseren Berücksichtigung der Betrachtungsebene des Stadtquartiers, leiten die Autoren zwölf Siedlungstypen für Wohngebäude und drei Haupttypen für den Nichtwohngebäudebereich, mit jeweils einigen Untertypen, sowie einen Typ für bebaubare Freiflächen ab (Erhorn et al. 2010, S. 28–34).



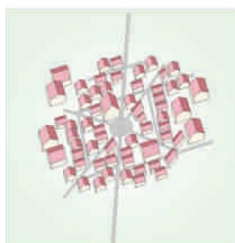
ST1: Lockere offene Bebauung



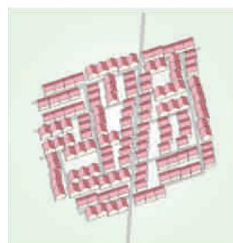
ST2: Einfamilienhäuser- und Doppelhäusersiedlung



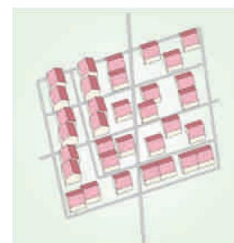
ST3a: Ehemaliger/städtischer Dorfkern



ST3b: Ländlicher Dorfkern



ST4: Reihenhaussiedlung

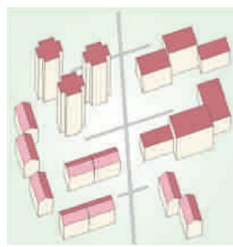


ST5a: Siedlung kleiner Mehrfamilienhäuser

Abbildung 2-4: Darstellung von 12 Siedlungstypen für den Wohngebäudebereich nach Erhorn et al. (2010) [Erhorn et al. (2010)] (Fortsetzung auf Folgeseite)



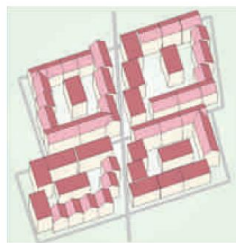
ST5b: Zeilenbebauung mit kleinen und größeren Mehrfamilienhäusern



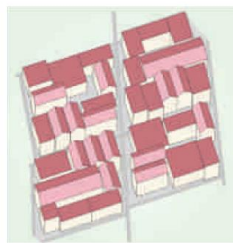
ST6: Zeilenbebauung mit großen hohen Mehrfamilienhäusern und Hochhäusern



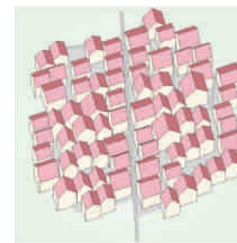
ST7a: Blockbebauung niedriger Dichte



ST7b: Blockbebauung hoher Dichte



ST8: Citybebauung



ST9: Historische Altstadt

Abbildung 2-4 Fortsetzung

Wie auch in den vorherigen Studien lieferten Erhorn et al. (2010) für ihre Siedlungstypen charakteristische Kennwerte bzgl. geometrisch-baulicher Merkmale sowie energetische Kennwerte und nahmen zudem eine grobe Zuordnung typischer Baujahresspannen zu den Siedlungstypen vor (Erhorn et al. 2010, S. 35–38). Da reale Quartiere auch aus baulichen Strukturen bestehen können, welche sich in mehrere Siedlungstypen aufteilen, wurde ein Verfahren vorgeschlagen, welches eine gewichtete Addition der Kennwerte der jeweiligen Siedlungstypen vorsieht, um beschreibende Kennwerte (z. B. die durchschnittliche beheizte Nutzfläche oder die mittlere Wärmeverbrauchsichte) für das reale Quartier als Untersuchungsgegenstand ermitteln zu können. Die erstellte Typologie der Siedlungsstrukturen nutzten Erhorn et al. zur Eingruppierung der Quartiere, welche im Rahmen der „EnEff:Stadt“-Vorhaben untersucht wurden (siehe auch Beier et al. 2016, S. 30). Eine weiterführende empirische Untersuchung mit dem Ziel einer Quantifizierung der Siedlungstypen oder Realquartiere innerhalb des deutschen Gebäudebestandes erfolgte im Rahmen der wissenschaftlichen Begleitforschung zu EnEff:Stadt jedoch nicht. (Vgl. Erhorn et al. 2010; Erhorn-Kluttig und Erhorn 2016; Beier et al. 2016)

Eine deutliche Weiterentwicklung des von Roth (1980) entwickelten Ansatzes erfolgte von 2009 bis 2015 im „EnEff:Stadt“-Forschungsvorhaben „UrbanReNet“. Hegger et al. (2012; 2015a) erweiterten in diesem Vorhaben die bisher (mit Ausnahme von Everding und Kloos (2007)) nur auf baulich-geometrische Parameter beschränkte und auf Gebäudewärmebedarfe ausgerichtete Typologie um eine Klassifizierung hinsichtlich der Potenziale für verschiedene erneuerbare Energieträger. Hierbei wurden sowohl gebäudeintegrierte bzw.

-nahe Potenziale als auch städtische Freiflächen berücksichtigt. So wurden in verschiedenen Parameterstudien und ergänzenden Einzeluntersuchungen die Potenziale zur Nutzung von Erdwärmesonden und solarer Strahlungsenergie sowie durch den Einsatz von lokal verfügbarer Biomasse für definierte Stadtstrukturen abgeschätzt. Auch hinsichtlich der Gebäudewärmebedarfe wird in Hegger et al. (2012) ein deutlich differenzierteres Vorgehen gewählt. Die Wärmebedarfe für die in einem energetischen Stadtraumtyp üblicherweise vorkommenden Gebäudetypen werden nach Baualtersklasse sowie energetischem Modernisierungszustand der Gebäude angegeben. Die von Hegger et al. verwendete Bezeichnung „energetische Stadtraumtypen“ (EST) unterstreicht den beschriebenen zusätzlichen energetischen Informationsgehalt. (Hegger et al. 2012; Dettmar et al. 2020)

Hegger et al. (2012) unterteilen die städtischen Strukturen in insgesamt 13 EST, wobei für drei dieser EST jeweils zwei Untergruppen vorliegen (siehe Tabelle 2-1)⁸. In Ergänzung zu den bereits vorgestellten Siedlungstypen für bauliche Strukturen anderer Autoren umfasst diese Typisierung auch Freiflächen, welche hinsichtlich der energetischen Potenziale eine bedeutende Rolle innerhalb größerer städtischer Teilgebiete spielen können. Dabei werden als Raumebenen mit Strukturmerkmalen das Gebäude, die Parzelle, die Sektion (als homogen bebauter Teil eines Baublocks), der Baublock und das Gebiet (als Zusammenschluss mehrerer Baublöcke) definiert (Hegger et al. 2015a, S. 70–74; siehe auch Anhang A).

Ähnlich, wie bereits von Erhorn et al. (2010) beschrieben, verfolgen auch Hegger et al. (2012; 2015a) den Ansatz, dass reale Untersuchungsräume, d. h. Stadtquartiere oder noch größere städtische Gebiete, als eine Kombination von EST gebildet werden können. Das Untersuchungsgebiet wird im Vorhaben „UrbanReNet“ dafür in Sektionen unterteilt, welche jeweils einem EST zugeordnet werden. Ergänzende Raumeinheiten, wie Wasserflächen, Straßenräume und energetische Einzelelemente, sind ebenfalls zu berücksichtigen und bilden mit den Sektionen das Quartier bzw. städtische Untersuchungsgebiet.

Eine ursprünglich als Projektbaustein ebenfalls vorgesehene Untersuchung zu den Häufigkeitsverteilungen der EST in Deutschland sollte „die größten energetischen Einsparpotenziale und eine Gewichtung bei der Entwicklung von Energiekonzepten [aufzeigen]. Auch Schwerpunkte bei der Ausrichtung zukünftiger Forschungsschwerpunkte sollten aus der quantitativen Verteilung der EST abgeleitet werden“ (Hegger et al. 2015a, S. 170). Dieses Vorhaben konnte jedoch nicht umgesetzt werden, was die Autoren mit einem besonders hohen Aufwand aufgrund des sehr heterogenen deutschen Siedlungsbestandes begründen. Eine Verallgemeinerungsfähigkeit von Untersuchungsergebnissen für Städte bzw.

⁸ Eine Darstellung der typischen stadtmorphologischen Strukturen der einzelnen EST ist in Anhang A dieser Arbeit zu finden.

Stadtgebiete aus einer Stichprobenerhebung wäre nur eingeschränkt gegeben (vgl. Hegger et al. 2015a, S. 170).

Tabelle 2-1: Beschreibung der energetischen Stadtraumtypen nach Hegger et al. (2012) [vgl. Hegger et al. (2012, S. 29)]

EST 1	Kleine, freistehende Wohnbebauung niedriger bis mittlerer Geschossigkeit A – freistehende Einfamilienhausbebauung B – freistehende Mehrfamilienhausbebauung	Überwiegende Wohnnutzung
EST 2	Reihenhausbebauung	
EST 3	Zeilenbebauung niedriger bis mittlerer Geschossigkeit	
EST 4	Großmaßstäbliche Wohnbebauung hoher Geschossigkeit A – großmaßstäbliche Zeilenbebauung B – freistehende Wohnhochhäuser	
EST 5	Blockrandbebauung	
EST 6	Dörfliche Bebauung	Überwiegende Mischnutzung
EST 7	Historische Altstadtbebauung	
EST 8	Innenstadtbebauung	
EST 9	Geschäfts-, Büro-, und Verwaltungsgebiet	Büro- und gewerbliche Nutzung
EST 10	Gewerbegebiet A – Gewerbegebiet, überwiegend Handel B – Gewerbegebiet, überwiegend Produktion	
EST 11	Öffentliche Parkanlagen	Freiraum
EST 12	Friedhofsanlagen	
EST 13	Kleingartenanlagen	
EE	Energetische Einzelemente	

Die beschriebenen (Weiter-)Entwicklungen der Typologien für Siedlungsstrukturen und Quartiere sollen im nachfolgenden Abschnitt um einen ausschnittshaften Überblick zu existierenden Softwarelösungen für die Bewertung energetischer Quartierskonzepte ergänzt werden. Im Anschluss daran erfolgt eine Erörterung der Grenzen der bestehenden Typologien sowie der Bewertungstools vor dem Hintergrund der Forschungsfragen dieser Arbeit.

2.5 Softwarelösungen zur Bewertung energetischer Quartierskonzepte

Ergänzend zu den verschiedenen Weiterentwicklungen des Typisierungsansatzes für Siedlungsstrukturen und Quartiere wurden in den vergangenen Jahren im Rahmen von Forschungsvorhaben auch verschiedene computergestützte Werkzeuge mit Energie- und Quartiersbezug entwickelt. Diese dienen u. a. zur Berechnung von Wärmebedarfen von Gebäuden im untersuchten Stadtgebiet und zur Erstellung sowie Bewertung von Energieversorgungskonzepten unter Beachtung des baulich-räumlichen Kontextes der Gebäude. Eine Auswahl der verfügbaren Softwarelösungen wird nachfolgend vorgestellt und deren Fähigkeiten sowie Einschränkungen diskutiert.

Im Rahmen von zwei Forschungsvorhaben zwischen 2011 und 2016, welche durch das Bundesinstitut für Bau-, Stadt- und Raumforschung gefördert wurden, entstand am Institut Wohnen und Umwelt (IWU) das Excel-basierte Berechnungswerkzeug „EQ-Tool“. Auf Grundlage einer breit angelegten Untersuchung zum damaligen Stand der energetischen Stadterneuerung und der Anforderungen an ein unterstützendes Bewertungsinstrument energetischer Quartierskonzepte wurde das EQ-Tool entwickelt (von Malotki et al. 2012; von Malotki et al. 2016). Es erlaubt eine systematisierte Erfassung von Daten zu einem lokalen Gebäudebestand hinsichtlich dessen Bebauungsstruktur, Gebäudeenergieverbräuchen und Wärmeschutzniveaus. Ökonomische und ökologische Rahmenbedingungen sowie empirische Ergebnisse zu energetischen Modernisierungsraten von Gebäudehüllbauteilen werden genutzt, um eine Projektion der Modernisierungsfortschritte⁹ ohne jegliche Intervention vorzunehmen. Darauf aufbauend können verschiedene Szenarien mit kontextbezogenen Schwerpunkten der energetischen Modernisierung des Quartiers vorgenommen werden und so THG-Minderungspotenziale ermittelt und eine Projektion des Verlaufs der energetischen Modernisierung des Quartiers erstellt werden. Eine genaue Abbildung des räumlichen Zusammenhangs der Gebäude im Quartier sowie der Auswirkungen dieser räumlichen Randbedingungen auf die Wärmeversorgungskonzepte erfolgt im EQ-Tool nicht. Das Bewertungswerkzeug wurde seit der Erstellung in einer Vielzahl von Vorhaben zur Bewertung von realen Quartiersmodernisierungskonzepten eingesetzt und im jeweiligen Vorhabenkontext weiterentwickelt (z. B. Bagherian et al. 2017). Die letzte Aktualisierung erfolgte im Jahr 2021 im Rahmen des Forschungsvorhabens „EG2050:E4Q“ (Koch 2021).

⁹ Indikatoren für den Modernisierungsfortschritt sind die mittlere Modernisierungsrate, end- und primärenergetische Bedarfs- bzw. Verbrauchsreduktion sowie die Reduktion von THG-Emissionen im untersuchten Quartier im Vergleich zum Startjahr der Betrachtung.

Ein weiteres Werkzeug zur Bilanzierung und Bewertung von energetischen Quartierskonzepten ist der „Energiekonzept-Berater für Quartiere“ (engl. District-Energy Concept Adviser; District-ECA). Er wurde vom Fraunhofer-Institut für Bauphysik im Rahmen der wissenschaftlichen Begleitforschung für die Förderinitiative EnEff:Stadt (vgl. Abschnitt 2.4.2) entwickelt und in einer ersten Version im Jahr 2012 veröffentlicht. Eine Aktualisierung inkl. verschiedener Erweiterungen erfolgte im Jahr 2018 im Rahmen eines durch die Europäische Union (EU) geförderten Vorhabens. (Erhorn-Kluttig et al. 2013; Budde et al. 2018)

District-ECA bietet dem Anwender eine grafische Oberfläche zur Erstellung eines energetischen Quartierskonzepts mit Hilfe einer Auswahl an Typgebäuden für Wohn- und Nichtwohnzwecke sowie einer Vielzahl kombinierbarer Wärme- und Stromversorgungsvarianten auf Gebäude- und Quartiersebene und darüber hinaus ein räumlich übergeordnetes Versorgungsnetz. Die Berechnung der Energiebedarfe erfolgt dabei gemäß DIN V 18599:2016 in einem Monatsbilanzverfahren für 15 mögliche Klimazonen in Deutschland. Die Typgebäude können in District-ECA in verschiedenen energetischen Ausgangs- und Modernisierungszuständen modelliert werden. Als energetische Bewertungsindikatoren finden der End- und Primärenergiebedarf sowie die CO₂-Emissionen für die Wärme- und Strombereitstellung im Quartier Anwendung (siehe Abbildung 2-5). Neben dem Programmmodul zur eigenen Modellierung eines energetischen Quartierskonzepts liefert District-ECA zudem ein Benchmarking-Modul zur Gegenüberstellung von Quartiersverbrauchswerten und den mittleren Energieverbräuchen von Gebäuden in Deutschland sowie einem Zielwert. Das Benchmarking erlaubt lediglich eine unscharfe Abbildung der Gebäude im Quartier. Zudem ist die Datenbasis dieses Benchmarking-Moduls von 2007 bzw. 2011, was ebenfalls nur zu einer eingeschränkten Aussagefähigkeit führen dürfte (vgl. Fraunhofer-Gesellschaft; Budde et al. 2018; Fraunhofer IBP 2018). Der räumliche Zusammenhang der Gebäude im Quartier wird wie auch im EQ-Tool fast vollständig außer Acht gelassen. Lediglich über eine Anpassung der Verlustfaktoren für die netzgebundene Wärmebereitstellung können Besonderheiten, wie eine sehr geringe Wärmebedarfsdichte, abgebildet werden (vgl. Budde et al. 2018, S. 36–38; Fraunhofer IBP 2018).

Während die beiden oben beschriebenen Softwarelösungen die Bewertung von energetischen Quartierskonzepten über die zusammengefasste Betrachtung von Einzelgebäuden vornehmen, basieren andere Entwicklungen direkt auf geometrischen Gebäudemodellen und bilden somit den räumlichen Zusammenhang der Gebäude eines Quartiers ab. Das Programm „Sim Stadt“ der Hochschule für Technik besteht aus verschiedenen Programmmodulen und -bibliotheken, welche eine vollständige energetische Bewertung eines eingegrenzten Stadtgebiets ermöglichen. Für dieses Stadtgebiet müssen entsprechende GIS-Daten und weiterführende Gebäudeinformationen vorliegen. Die Datenverarbeitung besteht dabei vereinfacht beschrieben aus:

1. der Vorverarbeitung und Analyse von dreidimensionalen Gebäudemodellen (welche z. B. im Format CityGML bereitgestellt werden; vgl. Kapitel 4),
2. der Anreicherung der Gebäudedaten um bauphysikalisch-energetische Parameter und Kennwerte,
3. einer Wärmebedarfsberechnung nach DIN V 18599 oder einer Zuweisung von Wärmebedarfswerten auf Basis einer Zuordnung der realen Gebäude zu Beispielgebäuden der IWU-Gebäudetypologie sowie einer Abschätzung der Strombedarfe,
4. einer weitestgehend automatischen Dimensionierung von Varianten der Quartierswärmeversorgung sowie einer Wärmenetzauslegung für vernetzte Wärmeversorgungskonzepte.

Das Ergebnis der Quartiersanalyse mit Sim Stadt ist ein „Vergleich der betrachteten Wärmeversorgungsvarianten anhand festgelegter Kennzahlen und Schlüsselindikatoren wie den Installations- und Betriebskosten sowie jährlichen Treibhausgasemissionen oder (dem) Primärenergiebedarf“ (Weiler et al. 2021, S. 43). Im Vergleich zu den einfacheren Quartiersbilanzierungswerkzeugen EQ-Tool und District-ECA erlaubt Sim Stadt eine deutlich detaillierte Abbildung der Bebauung eines Quartiers und Berücksichtigung dieser in den energetischen Berechnungen. Dieser Detaillierungsgrad hängt jedoch maßgeblich von der verfügbaren Datengrundlage ab. Nur wenn detaillierte Informationen über das Gebäudebaualter und den energetischen Zustand der jeweiligen Gebäude im Quartier vorliegen, können diese in der Bewertung berücksichtigt werden. Fehlen diese Informationen, werden die Lücken über entsprechende Werte aus den hinterlegten Gebäude- und Bauteilbibliotheken gefüllt. Die Berechnung der Wärmebedarfe unterscheidet sich dann aber nur unwesentlich von den anderen Bilanzierungswerkzeugen. Die Betrachtung des räumlichen Zusammenhangs im Rahmen der Bewertung und des Vergleichs von Wärmeversorgungskonzepten bleibt jedoch in jedem Fall als Mehrwert erhalten. (Vgl. Nouvel et al. 2015a; Eicker und Schumacher 2012; Weiler et al. 2018; Weiler et al. 2021; Coors 2018)

Eine weitere Modellierungsumgebung wurde in den vergangenen Jahren am Sustainable Design Lab des Massachusetts Institute of Technology (MIT) entwickelt. Diese nutzt dreidimensionale Gebäudemodelle, um Simulationen für Quartiere und Stadtgebiete zu erlauben. Das sog. Urban Modeling Interface (UMI) stellt dabei die Schnittstelle dar zwischen der CAD-Software Rhino3D und den Berechnungsmodulen für eine Analyse ...

- der Gebäudeenergiebedarfe in der Nutzungsphase,
- der „grauen“ Energie der Gebäude aus der Herstellungsphase,
- der Erreichbarkeit und Fußläufigkeit der Objekte im Quartier,
- der Tageslichtversorgung der Gebäude im Quartier,
- der Möglichkeiten für eine urbane Lebensmittelproduktion sowie
- von Quartiersversorgungskonzepten.

Die energiebezogenen Gebäudesimulationen werden dabei mit dem Simulationswerkzeug „Energy+“ des US-amerikanischen Energieministeriums (Department of Energy)¹⁰ durchgeführt und erzeugen stündliche Bedarfswerte aller betrachteten Gebäudemodelle im Quartier¹¹. Zur Bewertung des Quartiers werden jahresbezogene Kennwerte berechnet (siehe Abbildung 2-5). Bei fehlenden Informationen über den energetischen Zustand der Gebäude und über den gemessenen Wärmeverbrauch der Gebäude erfolgt auch im UMI eine Verwendung von Daten typischer Gebäude nach Baualterklasse und Nutzungsart. Falls bekannt, kann ebenfalls das Jahr der letzten Gebäudemodernisierung zur Abschätzung des energetischen Zustandes verwendet werden. (Reinhart et al. 2013; Cerezo Davila et al. 2016)

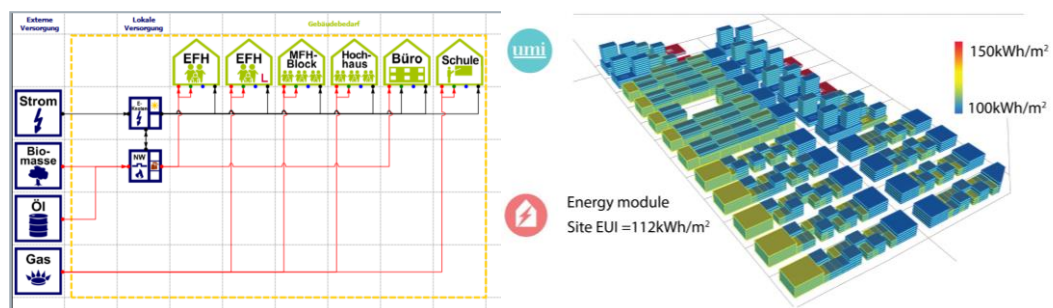


Abbildung 2-5: Modellierungsoberfläche im Programm District-ECA (links) und Darstellung des Gebäudeenergiebedarfs durch UMI (rechts) [links: eigene Darstellung; rechts: SDL (2022)]

In einer bemerkenswerten Arbeit von Nagpal (2020) wurde diese Softwareumgebung zudem genutzt, um den Energiebedarf des MIT-Campus nicht nur im Sinne einer deterministischen Untersuchung mit definierten bauphysikalisch-energietechnischen Parametern zu simulieren, sondern um eine probabilistische Berücksichtigung von möglichen Streuungen der nötigen Simulationseingangsdaten vorzunehmen. Durch dieses Verfahren konnte eine Erhöhung der Robustheit in der Optimierung des energetischen Quartierskonzepts erreicht werden. (Nagpal 2019, 2020)

Die in diesem Abschnitt dargestellten Softwarelösungen zur Erstellung und Bewertung von energetischen Modernisierungskonzepten für Quartiere zeigen unterschiedliche Detaillierungsniveaus, welche zu unterschiedlichen Stärken und Anwendungsfeldern der Programme führen. Eine kurze Diskussion der erkennbaren Begrenzungen der Methoden zur Typisierung von Stadtstrukturen und Quartieren sowie der Berechnungswerkzeuge schließt

¹⁰ Weiterführende Informationen zur Software „Energy+“ sind auf der Website des U.S. Department of Energy (<https://energyplus.net>) zu finden.

¹¹ Die Energiebedarfsberechnung erfolgt dabei unter der vereinfachten Darstellung der Gebäude als sog. „Shoobox“-Modelle. Alle Gebäude werden somit als vereinfachte Klötzchenmodelle abgebildet.

sich im nachfolgenden Abschnitt an. Diese Diskussion soll grundlegende Hinweise für die spätere, im Rahmen der Methodenentwicklung durchgeführte Auswahl der Datenbasis (vgl. Kapitel 5) liefern.

2.6 Grenzen bestehender Typologien im Hinblick auf den Klimaschutz

2.6.1 Übertragung der Ergebnisse auf den deutschen Quartiersbestand

Wie in den beiden vorangegangenen Abschnitten erläutert, wurden vor dem Hintergrund unterschiedlicher Fragestellungen Typologien von Siedlungsstrukturen und Quartieren erstellt sowie computergestützte Modellierungs- und Bewertungswerkzeuge entwickelt. Bei genauerer Betrachtung zeigen sich im Hinblick auf die im Rahmen der vorliegenden Arbeit formulierten Fragestellungen Grenzen der existierenden Arbeiten.

Als einzige Studie liefert Roth (1980) ein Konzept, welches die Verallgemeinerung seiner Ergebnisse auf den gesamtdeutschen Raum erlaubt. Diese Untersuchung stützt sich dabei auf empirische Daten zur Raumstruktur in Deutschland und einer Regressionsanalyse der Siedlungsstrukturen in einer im Vorhaben ausgewählten 1 %-Stichprobe der Kommunen in Deutschland. Die Schätzung der Häufigkeiten von Siedlungstypen mittels statistischer Regression ist jedoch zwangsläufig mit Unsicherheiten behaftet. Diese beruhen maßgeblich auf der Tatsache, dass die Raumebene der Kommunen (d. h. Städte und Gemeinden) als Merkmalsträger der Stichprobenuntersuchung und Regressionsanalyse dient. Merkmale sind ausschließlich soziodemografische und sozioökonomische Faktoren. Eine Berücksichtigung von Merkmalen in höherer räumlicher Auflösung (z. B. auf der Ebene von Ortsteilen oder statistischen Erhebungsbezirken) erfolgte nicht. Die Schätzung der Häufigkeiten von Siedlungstypen im deutschen Gesamtbestand nach Roth (1980) beruht folglich auf der Annahme, dass die Anteile der Typquartiere in zwei Kommunen ähnlicher Merkmalsausprägungen hinsichtlich Demografie und sozioökonomischer Situation ebenfalls ähnlich sind. Ein Beleg dieser Annahme erfolgte in der Studie und anderen zugehörigen Veröffentlichungen nicht. Deshalb und aufgrund der sehr individuellen städtebaulich-historischen Entwicklung der Städte in Deutschland innerhalb der vergangenen Jahrhunderte¹² wird diese Annahme in Zweifel gezogen.

Die weiteren, in den vorherigen Abschnitten dargestellten Ergebnisse liefern zudem keinen (direkten) Ansatz zur Übertragung von Quartierstypologien oder Ergebnissen der softwaregestützten Erstellung und Bewertung von Energiekonzepten für Quartiere auf einen

¹² Ausführliche Darstellungen der Stadtentwicklung und Stadtentwicklungsplanung in Deutschland können z. B. Albers und Wékel (2021), Streich (2011) und Bott (2018) entnommen werden.

nationalen Gebäude- bzw. Quartiersbestand. Zwar formulieren Erhorn et al. (2010) sowie Hegger et al. (2012; 2015b) Möglichkeiten zur Abbildung heterogen bebauter Siedlungsstrukturen durch eine Zusammenfassung mehrerer Siedlungstypen bzw. EST, dennoch ist eine Übertragung nicht ohne Weiteres möglich – auch wegen der Diskrepanz zwischen den zur Definition verwendeten Gebäudemerkmalen sowie flächendeckend verfügbaren Gebäudedaten in Deutschland¹³.

2.6.2 Anforderungen an die Datenbasis der Gebäudemerkmale und die Datenqualität

Alle vorgestellten Typologisierungen von Stadt- bzw. Siedlungsstrukturen oder Quartieren beruhen auf einer – wenn auch unterschiedlich stark differenziert – detaillierten Beschreibbarkeit energiebedarfsrelevanter Merkmale der Bebauung. Dabei leiten sich in allen Ansätzen die quartiers- oder siedlungstypbezogenen Energiekennwerte von definierten Gebäudemodellen und deren energetischer Bilanzierung ab. Auch den Softwarelösungen liegt dieses Konzept zugrunde, wobei die mit GIS-Daten arbeitenden Werkzeuge Sim Stadt und UMI zumindest hinsichtlich der Gebäudegeometrie die reale Situation der untersuchten Gebäude abbilden können. Im Falle fehlender Informationen zu bauphysikalischen bzw. energietechnischen Parametern verwenden beide Hilfsmittel jedoch ebenfalls Bilanzierungsparameter von Beispiel- oder Typgebäuden. Die Bilanzierungsergebnisse der Softwarelösungen sind somit immer mit der Unsicherheit behaftet, dass ein als unsaniert oder nur wenig modernisiert modelliertes Gebäude in der Realität bereits eine umfassende Modernisierung erfahren hat. Die Potenziale zur Reduktion der Energiebedarfe und THG-Emissionen in einem untersuchten Quartier werden in diesem Falle deutlich überschätzt. Im Hinblick auf eine Quantifizierung von Quartieren und deren Beiträge zum Klimaschutz zeigt sich, dass sowohl die flächendeckende Einordnung von Gebäuden, Quartieren und Stadtteilen in die bestehenden Typologien als auch eine flächendeckende Anwendung der Softwarelösungen eine detaillierte energetische Erfassung aller Gebäude in Deutschland erfordern würde. Mindestens jedoch müsste eine stichprobenbasierte Untersuchung durchgeführt werden, die mit einer hohen räumlichen Auflösung (z. B. Ebene der Baublöcke oder statistischen Bezirke) die baustrukturellen Merkmale und energetischen Merkmale der Einzelgebäude ermittelt und im Sinne einer Kalibrierung der Typologien oder Softwarelösungen eine Übertragung auf den Gesamtbestand hinreichend absichert. Im Rahmen der Methodenentwicklung in Kapitel 5 wird auf die Begrenzungen der bestehenden Methoden erneut eingegangen.

¹³ Zusätzlich zu den Ausführungen im nachfolgenden Abschnitt soll bzgl. der verfügbaren Datenbasis auf das Kapitel 4 dieser Arbeit verwiesen werden.

Insgesamt kann attestiert werden, dass die Betrachtungsebene des Quartiers für die Implementierung von Maßnahmen zur Erreichung der Klimaschutzziele im Gebäudebestand einen nicht vernachlässigbaren Maßstab bildet. Aus den Erfahrungen der vergangenen Forschung kann somit die Prämisse abgeleitet werden, dass energetische Modernisierungskonzepte im Quartiersmaßstab die Umsetzung der „sowieso“ notwendigen Modernisierungsschritte auf Einzelgebäudeebene unterstützen können und dabei zu optimierten Lösungen beitragen. Die sich daraus ergebende Bedeutung des Quartiers für den Klimaschutz im Gebäudesektor unterstreicht die Wichtigkeit eines verbesserten Verständnisses des deutschen Quartiersbestandes. Die Ergebnisse der hier vorliegenden Arbeit können mit der Beschreibung, der Quantifizierung sowie einer statistischen Beschreibung von Typquartieren einen Beitrag zur Verbesserung eben dieses Verständnisses leisten. Über die bereits diskutierten Quellen hinausgehende Arbeiten, auf welchen die voranstehend formulierte Einschätzung zum Nutzen der Quartiersebene ebenfalls beruhen, sind Freudenberg et al. (2019), Freudenu et al. (2021), Huttenloher (2020), Feldmann (2009), Reicher und Schmidt (2021), Langenbrinck et al. (2017) und Erhorn-Kluttig et al. (2011).

3 METHODEN ZUR GRUPPENBILDUNG UND KLASSIFIZIERUNG

3.1 Hinführende Erläuterungen

In einer Vielzahl wissenschaftlicher Disziplinen stellt sich bei der Analyse von Forschungsdaten die Frage, ob eine Einteilung der beobachteten Objekte in Gruppen möglich ist und nach welchen Kriterien diese im jeweiligen Forschungskontext sinnvoll vorgenommen werden kann. Beispielhaft können hier die Arbeiten in der Biologie zur systematisierten Klassifizierung von Lebewesen genannt werden (vgl. Behnisch 2008, S. 46). In den in Kapitel 2 vorgestellten Arbeiten zur Typologisierung von Siedlungsstrukturen und Quartieren erfolgte die Gruppenbildung auf Basis ingenieurmäßiger Festlegungen zur Einteilbarkeit von Gebäuden aufgrund ihrer baulichen und geometrischen Parameter. Die weiterführende Einteilung, z. B. der Siedlungstypen durch Roth (1980), wurde ebenfalls ingenieurmäßig unter Beachtung praxisrelevanter Rahmenbedingungen (nämlich der visuellen Abgrenzbarkeit auf topografischen Karten sowie einer gewünschten Unterteilung nach Nutzflächendichte) vorgenommen. Ein solches Vorgehen der begründeten Kriterienfestlegung erlaubt eine gewisse Feinjustierung und kann zur passgenauen Gruppeneinteilung für eine spezifische Fragestellung genutzt werden. Die festgelegten Einteilungskriterien können darüber hinaus verwendet werden, um eine Klassifizierung weiterer Datensätze vorzunehmen.

Neben dieser ebenfalls sehr anspruchsvollen Methode, Objekte in Gruppen einzuteilen, wurden im Laufe des vergangenen Jahrhunderts verschiedene mathematische und informationstechnische Methoden entwickelt, die eine Gruppierung (Clusteranalyse) und Klassifizierung von Daten erlauben. In der deutschen Sprache wird der Begriff „Klassifizierung“ oftmals synonym für eine Einteilung von Objekten eines Datensatzes in vorher nicht bekannte Gruppen als auch für eine Zuordnung von Objekten zu bestehenden, d. h. im Vorfeld der Analyse bekannten, Gruppen genutzt. Andere Arbeiten unterscheiden sprachlich zwischen strukturentdeckenden und strukturprüfenden Verfahren (vgl. Backhaus et al. 2018), wobei dort unter den strukturprüfenden Verfahren die Klassifizierungsverfahren zur Gruppenzuordnung verstanden werden. In englischsprachigen Arbeiten ist die sprachliche Differenzierung mit den beiden Begriffen „clustering“ (i. S. v. Gruppenfindung bzw. -bildung) und „classification“ (i. S. v. Gruppenzuordnung) klarer nachzuvollziehen. In der hier vorliegenden Arbeit wird sprachlich unterschieden zwischen „Clusteranalyse“, „dem Clustern von Daten“ oder „der Gruppenbildung“ als Prozess der Gruppen- bzw. Klassenfindung und der „Klassifizierung“ als Prozess der Gruppenzuordnung. (Vgl. Han et al. 2012; Backhaus et al. 2018; Kaufman und Rousseeuw 1990)

Die Clusteranalyseverfahren und Klassifizierungsmethoden sind zudem Teil einer größeren Menge von informationstechnischen Datenanalysemethoden, welche unter dem Oberbegriff des Maschinellen Lernens (engl. machine learning; ML) gegliedert werden. Hierbei

gehören die Clusteranalysemethoden per se zur Gruppe der nicht überwachten Lernmethoden (engl. unsupervised ML), da zunächst eine Gruppeneinteilung gefunden werden muss. Die Klassifizierungsmethoden hingegen sind den überwachten oder teilüberwachten Lernmethoden (engl. supervised bzw. semi-supervised ML) zuzuordnen. Hier besteht das Grundprinzip darin, dass für einen vorhandenen Datensatz eine Gruppenzuordnung existiert und in einem (meist größeren) Datensatz die Objekte anhand ihrer Merkmalsausprägungen in die bereits bestehenden Gruppen einsortiert werden sollen. Eine Einteilung einiger häufig genutzter Verfahren ist in Abbildung 3-1 dargestellt. (Vgl. Forsyth 2019; Jiang 2021; Han et al. 2012)

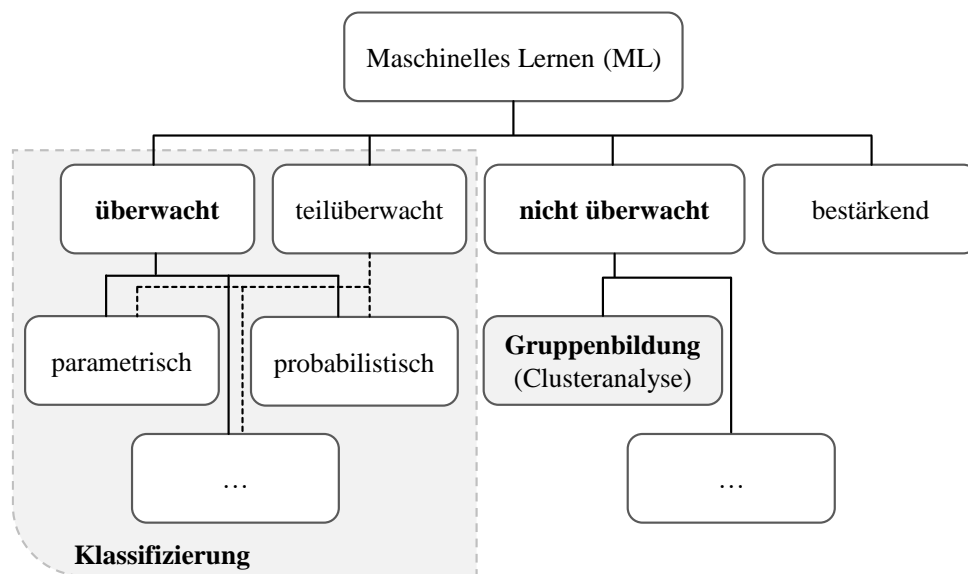


Abbildung 3-1: Einordnung von Clusteranalyseverfahren und Klassifizierungsmethoden als Teilbereiche des Maschinellen Lernens [eigene Darstellung in Anlehnung an Behnisch (2008), Mahrishi et al. (2021) und Plaue (2021)]

Eine Auswahl der in Abbildung 3-1 gezeigten Verfahren, welche im Kontext der Fragestellung dieser Arbeit und für die Methodenentwicklung in Kapitel 5 von besonderem Interesse sind, wird in den nächsten Abschnitten vorgestellt. Im nachfolgenden Abschnitt dieses Kapitels werden demnach zunächst die Grundlagen der Clusteranalyse und Klassifizierung vorgestellt. Die Beschreibung der mathematischen Methoden zur Datenanalyse erfolgt dabei sehr stark eingegrenzt und im Hinblick auf die im späteren Verlauf der Methodenentwicklung beschriebenen Daten und Zusammenhänge. Bei der Beschreibung der Clusteranalyseverfahren sowie der Klassifizierungsmethoden kann aufgrund der Vielzahl an Anwendungsgebieten und Verfahren nur eine Auswahl der existierenden Verfahren dargestellt werden. Die Auseinandersetzung mit den existierenden Möglichkeiten der Gruppierung und Klassifizierung von Daten i. A. und GIS-Daten im Besonderen bildet einen wichtigen Baustein für die in Kapitel 5 beschriebene Methodenentwicklung zur Typisierung des deutschen Quartiersbestandes.

3.2 Methodische Grundlagen

Die Clusteranalyse sowie die Klassifizierungsverfahren sind Methoden zur mathematisch-informationstechnischen Einteilung von Objekten anhand ihrer beobachteten Merkmale. Hierbei werden grundlegende Begrifflichkeiten und Einteilungssystematiken aus anderen Disziplinen, wie der deskriptiven Statistik, d. h. der beschreibenden Analyse quantitativer und qualitativer Daten, aber auch aus der induktiven Statistik als ein Aufgabengebiet zur Ermittlung der Beziehung zwischen Teilgesamtheiten und der Grundgesamtheit, zugrunde gelegt (vgl. Handl und Kuhlenkasper 2018, S. 3–4). Eine für die Darstellung der Clusteranalysemethoden sowie Klassifizierungsverfahren erforderliche Auseinandersetzung mit diesen Grundlagen wird im Folgenden gegeben. Auf die in den nachfolgenden Abschnitten vereinfacht beschriebenen Begriffe und Verfahren wird in Kapitel 5 jeweils in direktem Bezug zu den verwendeten Daten und Methoden vertiefend eingegangen.

Die Datenanalyse i. A. sowie die beiden oben benannten Aufgabengebiete der Statistik beschäftigen sich mit der Beschreibung von Objekten und deren Merkmalen. Untersuchungsobjekte können dabei z. B. Befragungsteilnehmer, aber auch andere physische oder nicht physische Einheiten sein, deren Merkmale quantitativ messbar oder zumindest qualitativ beschreibbar sind. Die untersuchten Objekte werden deshalb als Merkmalsträger bezeichnet. Da die Statistik das Ziel verfolgt, eine über ein einzelnes Objekt hinausgehende Beschreibung von Merkmalsausprägungen und deren Zusammenhängen zu liefern, werden mehrere Objekte zu einem Datensatz zusammengefasst und dessen Zusammensetzung im Hinblick auf die Ausprägungen der einzelnen Merkmale mittels verschiedener charakteristischer Werte beschrieben. (Vgl. Backhaus et al. 2018; Hedderich und Sachs 2020)

Die Merkmale – auch als Variablen bezeichnet – können dabei in Abhängigkeit ihrer Art bzw. Messbarkeit in unterschiedliche Merkmalstypen, die sog. Skalenniveaus, unterteilt werden (siehe Tabelle 3-1, S. 38). Nominale Variablen bilden qualitative Eigenschaften der Objekte und innerhalb der Merkmalsausprägungen keine Rangfolge ab. Auch binäre Variablen, also die Darstellung, ob ein Objekt eine Eigenschaft besitzt (wahr = 1) oder nicht (falsch = 0), gehören zu den nominalen Variablen. Ordinal skalierte Variablen sind ebenfalls als qualitative Merkmalsbeschreibungen zu verstehen, es liegt jedoch eine Rangfolge der Merkmalsausprägungen vor (Ausprägung A ist „besser“ als Ausprägung B; wobei „besser“ und „schlechter“ kontextbezogen zu definieren sind). Diese auch als nicht-metrisch bezeichneten Skalen lassen die Erfassung von Häufigkeiten der Variablenausprägungen im Datensatz und somit die Angabe eines Modalwertes¹⁴ zu. Zudem kann im Falle der

¹⁴ Der Modalwert oder auch Modus ist der Wert der am häufigsten beobachteten Merkmalsausprägung innerhalb der untersuchten Datenreihe (vgl. Eckle-Kohler und Kohler 2017).

ordinalen Variablen die Ermittlung des Medians¹⁵ vorgenommen werden. Metrische Variablen sind numerische Merkmale von Objekten, welche sich auf einer Skala mit gleich großen Abstandswerten zwischen den möglichen Ausprägungen innerhalb des Wertebereichs befinden.¹⁶ In dieser Gruppe der Skalen wird zudem zwischen intervallskalierten Variablen und ratio- oder verhältnisskalierten Variablen unterschieden. Für Erstere existiert dabei kein definierter Nullpunkt. Der Wertebereich spannt damit einen relativen Bewertungsrahmen zwischen den Merkmalsträgern innerhalb des Datensatzes auf. Für intervallskalierte Variablen lassen sich die arithmetischen Operationen Addition und Subtraktion ausführen und der arithmetische Mittelwert sowie die empirische Standardabweichung berechnen. Die Verhältnisskala besitzt über die Eigenschaften der Intervallskala hinaus einen definierten Nullpunkt, welcher i. d. R. dem Nichtvorhandensein der Ausprägung entspricht. Dieser Skalentyp erlaubt deshalb die Anwendung aller arithmetischer Operationen. (Vgl. Backhaus et al. 2018, S. 10–13; Hedderich und Sachs 2020)

Tabelle 3-1: Skalenniveaus, Merkmalsarten und mögliche Rechenoperationen [eigene Darstellung basierend auf Backhaus et al. (2018) und Hedderich und Sachs (2020)]

Skala	Merkmal	Rechenoperationen¹⁾
<i>nicht-metrisch (kategorial)</i>		
Nominalskala	Klassifizierung qualitativer Merkmalsausprägungen	Häufigkeiten
Ordinalskala	Rangwert mit Ordinalzahlen	Median, Quantile
<i>metrisch (kardinal)</i>		
Intervallskala	Skala mit gleichgroßen Abschnitten ohne natürlichen Nullpunkt	Addition, Subtraktion, Mittelwert
Ratioskala	Skala mit gleichgroßen Abschnitten mit natürlichem Nullpunkt	Multiplikation, Division

¹⁾ Zusätzlich sind die Rechenoperationen niedrigerer Skalenniveaus anwendbar.

¹⁵ Der Median ist definiert als der Wert einer Datenreihe, für den mindestens 50 % der beobachteten Werte kleiner oder gleich und mindestens 50 % der Werte größer oder gleich sind (Handl und Kuhlenkasper 2018, S. 83).

¹⁶ In verschiedenen Lehrbüchern wird die Einteilung der Variablen mit anderen Begriffen bezeichnet. So können anstelle der Bezeichnungen „nicht-metrisch“ und „metrisch“ auch die Termini „kategorial“ und „numerisch“ Anwendung finden (siehe z. B. Dümbgen (2016)). Die metrischen Variablen können zudem noch in diskrete und stetige Variablen unterschieden werden, wobei im Fall diskreter Variablen die Merkmale nur endlich viele Ausprägungen innerhalb des Wertebereichs annehmen können. Bei stetigen Variablen hingegen können beliebige Ausprägungen innerhalb des Wertebereichs der Variable vorliegen (Sibbertsen und Lehne 2021, S. 6).

Auf eine detailliertere Beschreibung der Berechnung von charakteristischen Werten statistischer Daten, wie z. B. den Median, den Mittelwert oder Quantilwerte, sowie Streuungsmaßen, wie der empirischen Standardabweichung, wird an dieser Stelle weitestgehend verzichtet und auf einschlägige Lehrbücher und die weiterführende Fachliteratur zum Thema verwiesen (z. B. Sibbertsen und Lehne 2021; Dümbgen 2016; Backhaus et al. 2018; Handl und Kuhlenkasper 2018; Plaue 2021).

Häufigkeitsverteilungen und Wahrscheinlichkeiten von Merkmalsausprägungen

Nach einer Einteilung der Merkmalsausprägungen in überschneidungsfreie Werteklassen können für nicht-metrische oder metrische Variablen Häufigkeitsverteilungen der Variablenwerte erzeugt und mittels Häufigkeitstabellen oder Histogrammen dargestellt werden. Aus der Häufigkeitsverteilung kann direkt auch die Wahrscheinlichkeitsverteilung der Variablen abgeleitet werden, indem je Klasse der Anteil der Merkmalswerte in einer Klasse an der untersuchten Gesamtheit ermittelt wird. Werden jeweils die relativen Häufigkeiten (also Wahrscheinlichkeiten des Vorkommens einer Merkmalsausprägung) einer Klasse mit den relativen Häufigkeiten der Klassen mit kleineren Merkmalswerten aufsummiert, so erhält man eine Verteilungsfunktion, welche eine Aussage erlaubt, wie wahrscheinlich das Auftreten eines Merkmalswertes ist, der kleiner oder gleich einem Zielwert ist (siehe Abbildung 3-2, links). Aus der empirischen Häufigkeitsverteilung stetiger Variablen lässt sich zudem die empirische Dichtefunktion ableiten, wenn die Klassenbreite zur Einteilung aller Objekte stetig verringert wird. Die Dichtefunktion ist dabei so definiert, dass die Fläche unterhalb der Dichtefunktion den Wert 1 ergibt (siehe Abbildung 3-2, rechts). Für diskrete Variablen können stetige Dichte- bzw. Verteilungsfunktionen durch Approximations- bzw. Schätzverfahren ermittelt werden. (Vgl. Sibbertsen und Lehne 2021; Handl und Kuhlenkasper 2018)

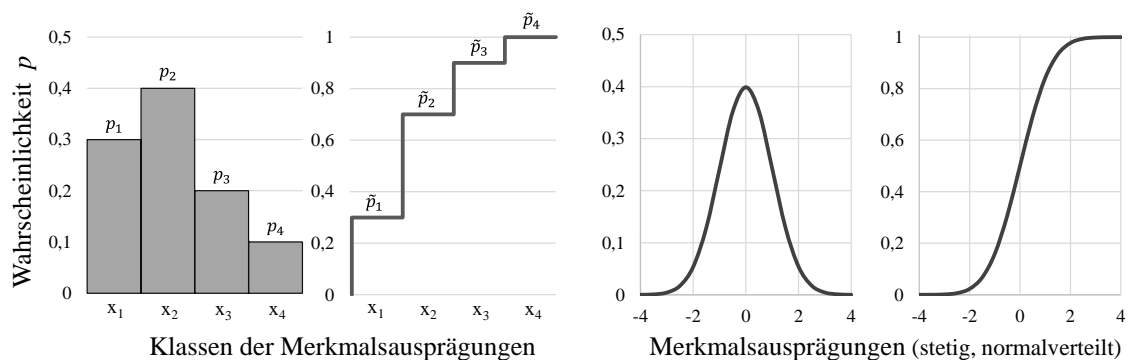


Abbildung 3-2: Wahrscheinlichkeiten einer in vier Werteklassen eingeteilten Variable x (links) und die daraus abgeleitete nicht stetige Verteilungsfunktion (2. v. links) sowie Dichtefunktion (2. v. rechts) und Verteilungsfunktion einer stetigen, normalverteilten Variable (rechts) [eigene Darstellung in Anlehnung an Wörner (2020), Sibbertsen und Lehne (2021) und Hedderich und Sachs (2020)]

Normalisierung und Standardisierung

Zur besseren Beschreibbarkeit der Variablen der Objekte eines Datensatzes und der Verteilungen von deren Merkmalsausprägungen sowie für einige der später vorgestellten Clusterverfahren ist eine Normalisierung oder Standardisierung der Daten erforderlich. Hierbei werden die Daten hinsichtlich ihrer Lage und relativen Beziehung zueinander transformiert, sodass die Verteilung der Merkmalsausprägungen gewünschte Eigenschaften annimmt. Im Rahmen der Normalisierung wird dabei i. d. R. darauf abgezielt, den Wertebereich einer Variablen auf ein Intervall von 0 bis 1 einzuschränken. Die Normierung des Objekts x zu x^s kann für die metrische Variable i gemäß der Berechnungsvorschrift in Gl. 3-1 durchgeführt werden. Bei einer Standardisierung hingegen erfolgt die Datentransformation zur Erreichung einer Nulllage des Mittelwertes \bar{x} und einer Standardabweichung s von 1. Diese Art der Datentransformation erlaubt einen Vergleich von Objekten aus unterschiedlichen Datensätzen (bzw. Stichproben) mittels ihrer Position innerhalb der standardisierten Verteilung der Merkmalsausprägungen. Das Verfahren der Standardisierung wird auch „Z-Transformation“ genannt. Hierbei wird der Mittelwert \bar{x} der Merkmalsausprägungen aller n Merkmalsträger der untersuchten Stichprobe bzw. der Gesamtheit vom Wert der Variablen x des jeweiligen Objekts i subtrahiert und durch die Standardabweichung s dieser Variablen dividiert (siehe Gl. 3-2). Andere Datentransformationen können zudem dazu verwendet werden, um Zusammenhänge bzw. Abhängigkeiten zwischen unterschiedlichen Merkmalen besser begreifbar, darstellbar und für weiterführende Rechenverfahren nutzbar zu machen. So wird z. B. das Logarithmieren von Merkmalswerten häufig dazu genutzt, um nicht-lineare Zusammenhänge zu linearisieren und somit entsprechende Rechenverfahren anwendbar zu machen. (Vgl. Hedderich und Sachs 2020)

$$x_i^s = \frac{x_i - x_{min}}{x_{max} - x_{min}} ; \text{mit } i \in [1, \dots, n] \quad \text{Gl. 3-1}$$

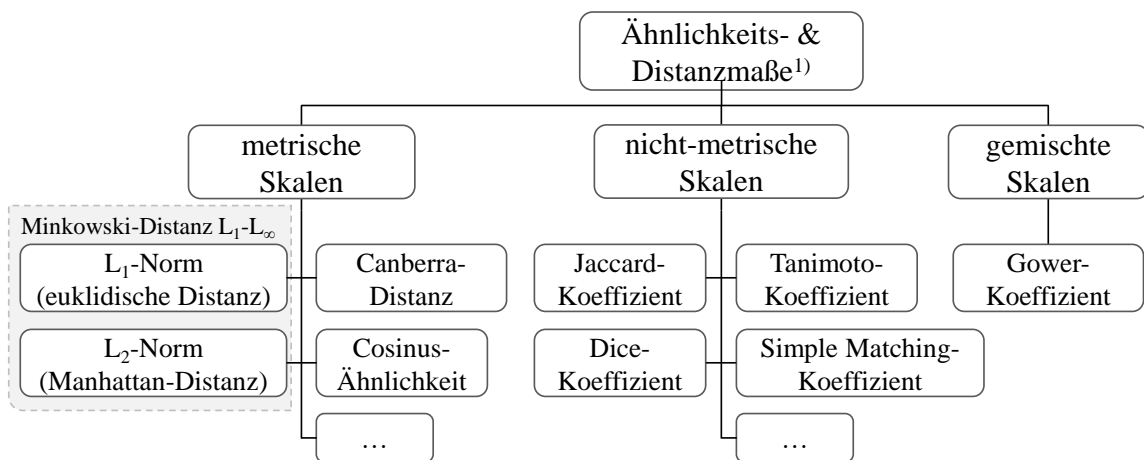
$$z_i = \frac{x_i - \bar{x}}{s} ; \text{mit } i \in [1, \dots, n] \quad \text{Gl. 3-2}$$

Ähnlichkeits- und Distanzmaße

Zur Analyse der Ähnlichkeit bzw. der Unterschiedlichkeit zweier Datenpaare, wie die Werte einer Variablen zweier Objekte, muss ein Abstand dieser Objekte anhand ihrer Merkmalsausprägungen ermittelt werden. Hierfür stehen verschiedene Verfahren zur Verfügung, welche vom Skalenniveau der Variablen abhängen. So ist bei binären oder kategorialen Variablen eine grundlegend andere Systematik zum Vergleich der Merkmalsausprägungen zu verfolgen, als dies bei metrischen Variablen möglich ist. Eine Übersicht über

die auch als Ähnlichkeitsmaße und Distanzmaße bezeichneten Verfahren zur Abstandsfeststellung ist in Abbildung 3-3¹⁷ gegeben.

Ein sehr häufig für Clusteranalysen verwendetes Distanzmaß ist die euklidische Norm oder euklidische Distanz, mit Hilfe derer für metrische Variablenwerte in einem zwei- bis n-dimensionalen euklidischen Raum der Abstand (oder auch Norm genannt) berechnet werden kann. Werden die zwei betrachteten Objekte und ihre Merkmalsausprägungen als Vektoren v_1 und v_2 bezeichnet, so ist die euklidische Vektornorm die Wurzel des Skalarprodukts der Vektoren (siehe Gl. 3-3). Im zwei- oder dreidimensionalen euklidischen Raum entspricht der Abstand dabei der Länge. (Vgl. Arens et al. 2018, S. 746–750; Westermann 2020, S. 50)



¹⁾Durch entsprechende Umrechnungsvorschriften kann aus einem Ähnlichkeitsmaß ein Distanzmaß ermittelt werden.

Abbildung 3-3: Ähnlichkeits- und Distanzmaße für den Vergleich von Variablen unterschiedlicher Skalenniveaus [eigene Darstellung basierend auf Kaufman und Rousseeuw (1990), Gower (1971) und Drost (2018)]

$$\|v\|_{\mathbb{R}^n} = \sqrt{(v_1 - v_2)^2} = \sqrt{\sum_{i=1}^n (v_{1,i} - v_{2,i})^2} ; \text{ mit } n \in \mathbb{N} \quad \text{Gl. 3-3}$$

Ein anderes, ebenfalls im Rahmen von Clusteranalysen verwendetes Distanzmaß ist die Manhattan-Distanz. Vereinfacht kann diese Berechnung im zweidimensionalen Raum als die Ermittlung der Länge des Abstandes zwischen zwei Punkten bei ausschließlich rechtwinkliger Wegführung beschrieben werden. Gl. 3-4 (S. 42) stellt die Berechnungsvorschrift für zwei Punkte x_1 und x_2 im n-dimensionalen Raum dar.

¹⁷ Die Übersicht ist wegen der Vielzahl an, z. T. für spezielle Anwendungsfälle definierten, Verfahren nur als beispielhafte Auswahl (sehr) häufig verwendeter Ähnlichkeitsmaße zu verstehen.

Abbildung 3-4 zeigt die Manhattan-Distanz zudem beispielhaft im zweidimensionalen Raum und stellt sie der euklidischen Distanz gegenüber. Die beiden Distanzmaße werden in einer verallgemeinerten Formulierung für unterschiedliche Vektorräume, der Minkowski-Distanz, zusammengefasst (vgl. Kaufman und Rousseeuw 1990, S. 10–13).

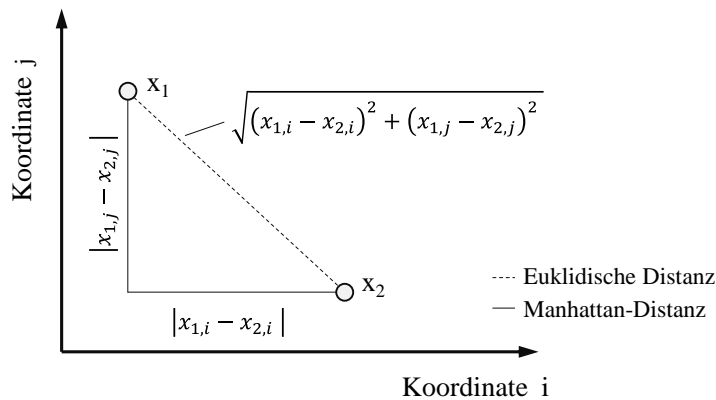


Abbildung 3-4: Euklidische Distanz (gestrichelte Linie) und Manhattan-Distanz (durchgezogene Linie) für zwei Punkte im zweidimensionalen Raum [eigene Darstellung basierend auf Kaufman und Rousseeuw (1990)]

$$d_{\mathbb{R}^n} = \sum_{i=1}^n (|x_{1,i} - x_{2,i}|) ; \text{ mit } n \in \mathbb{N} \quad \text{Gl. 3-4}$$

Als nicht-metrische Ähnlichkeitsfunktionen sollen an dieser Stelle der Jaccard-, der Dice-, der Simple Matching- sowie der Tanimoto-Koeffizient für binäre Daten vorgestellt werden. Das Grundprinzip der Ähnlichkeitsbewertung für binäre Daten kann über eine Kontingenztafel einfach veranschaulicht werden (siehe Tabelle 3-2), denn es geht bei der Festlegung eines solchen Koeffizienten um die Definition, bei welcher Kombination von Merkmalsausprägungen der zu vergleichenden Objekte eine Ähnlichkeit vorliegt. Für den Jaccard-Koeffizienten ist die Gleichheit eines binären Wertepaares nur für den Fall festgelegt, in dem beide Werte den Zustand „wahr“ (oder 1) annehmen (Fall *a* in Tabelle 3-2).

Der gleichen Festlegung folgt auch der Dice-Koeffizient, für den allerdings zusätzlich noch eine doppelte Gewichtung der Übereinstimmung definiert ist. Der Zustand „falsch“ beider Objekte wird hierbei als fehlende Information und damit als nicht interpretierbares Merkmal verstanden. Ergänzend zu dem Fall einer Übereinstimmung der Wahrheitswerte der beiden binären Variablen betrachten der Simple Matching- und der Tanimoto-Koeffizient auch das gemeinsame Vorliegen des Zustandes „falsch“ (oder 0) als Übereinstimmung, unterscheiden sich jedoch in der Gewichtung der Nichtübereinstimmungen. Die Gleichungen Gl. 3-5 bis Gl. 3-8 stellen die Koeffizienten in Bezug auf Tabelle 3-2 dar. Die als Ähnlichkeiten dargestellten Koeffizienten lassen sich zudem ebenfalls als Unähnlichkeitskoeffizienten darstellen, wenn der Zähler der jeweiligen Gleichung die als Nichtübereinstimmung

bewerteten Zustände enthält (vgl. Kaufman und Rousseeuw 1990, S. 23–27; Behnisch 2008, S. 33; Hennig 2015, S. 11–13). Eine Übertragung der Ähnlichkeitskoeffizienten von binären auf kategoriale Variablen ist generell möglich und Teil vieler softwareseitiger Implementationen.

Tabelle 3-2: Kontingenztabelle der möglichen Kombinationen von Zuständen zweier binärer Variablen [eigene Darstellung nach Kaufman und Rousseeuw (1990)]

		Objekt j	
		Zustand	
Objekt i	Wahr	<i>a</i>	<i>b</i>
	Falsch	<i>c</i>	<i>d</i>

$$s(i, j) = \frac{a}{a + b + c} \text{ Jaccard} \quad \text{Gl. 3-5}$$

$$s(i, j) = \frac{2a}{2a + b + c} \text{ Dice} \quad \text{Gl. 3-6}$$

$$s(i, j) = \frac{a + d}{a + b + c + d} \text{ Simple Matching} \quad \text{Gl. 3-7}$$

$$s(i, j) = \frac{a + d}{(a + d) + 2(b + c)} \text{ Tanimoto} \quad \text{Gl. 3-8}$$

Eine weiterführende Beschreibung der hier vorgestellten und einiger anderer Ähnlichkeits- bzw. Distanzmaße finden sich z. B. in Hennig et al. (2015), Kaufman und Rousseeuw (1990) oder in der vielfältigen Fachliteratur über die Erstellung und Anwendung von Clusteranalyseverfahren.

3.3 Verfahren zur Clusteranalyse

3.3.1 Allgemeine Beschreibung der Clusteranalyse

Die Clusteranalyse als Methode zur Verarbeitung und Auswertung statistischer Daten kann als ein Instrument verstanden werden, welches deutlich über die ausschließliche Anwendung von mathematisch-informationstechnischen Algorithmen¹⁸ hinausgeht. Vielmehr ist bereits die Formulierung einer Fragestellung bzw. die Festlegung des Untersuchungsziels der Datenanalyse der erste Schritt der Clusteranalyse. Darüber hinaus besteht eine

¹⁸ Ein Algorithmus ist dabei als eine endliche, geordnete Menge wohldefinierter Regeln zur Verarbeitung von Werten bzw. Daten mit dem Ziel der Problemlösung zu verstehen (vgl. ISO 2015 nach Plaue 2021, S. 189).

Clusteranalyse aus der Datenauswahl und -aufbereitung, der Festlegung geeigneter Distanz- bzw. Ähnlichkeitsmaße sowie der Auswahl eines geeigneten Clusteralgorithmus, der technischen Durchführung bzw. Ausführung des Algorithmus, der Analyse der Ergebnisse und einer abschließenden Interpretation. Hierbei kann es erforderlich sein, einzelne Schritte oder den gesamten Prozess iterativ zu durchlaufen. Die Clusteranalyse kann ebenso zur Anpassung der Fragestellung bzw. neuen Fragestellungen führen (siehe Abbildung 3-5). (Vgl. Behnisch 2008; Han et al. 2012)

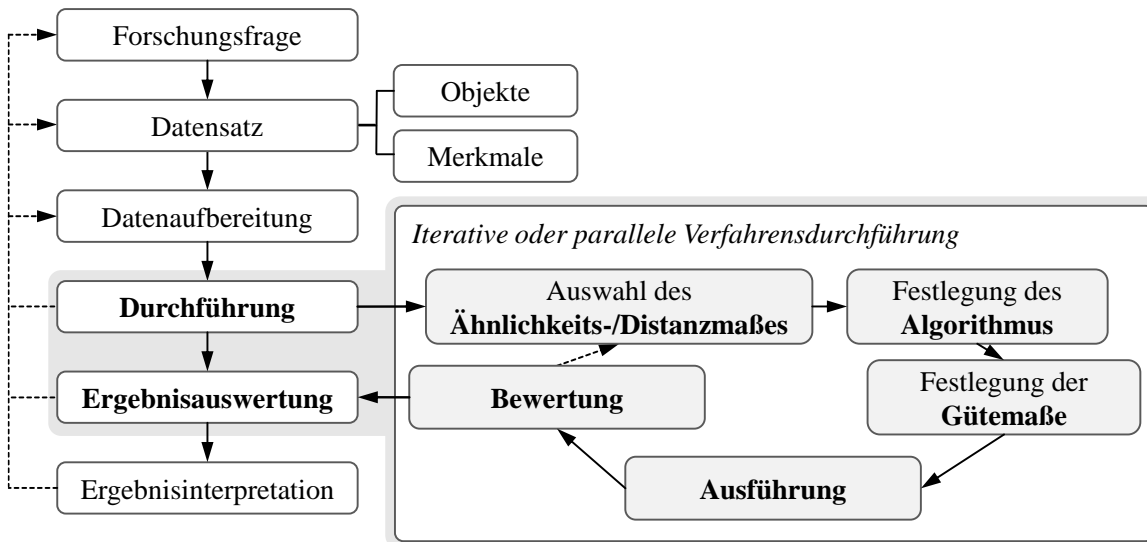


Abbildung 3-5: Prozessschritte einer Clusteranalyse [eigene Darstellung in Anlehnung an Behnisch (2008)]

Während Möglichkeiten zur Datenaufbereitung sowie der Ähnlichkeits- und Distanzmaße bereits im vorherigen Kapitel 3.2 aufgezeigt wurden, wird nachfolgend ein Überblick über die Vielzahl der Clusteranalyseverfahren gegeben und eine Auswahl von Clusteralgorithmen, also die informationstechnischen Rechenoperationen zur Gruppenbildung, beschrieben. Die Clusteralgorithmen werden i. d. R. nach der Art der Methodik zur Gruppenbildung in hierarchische, partitionierende, dichte-basierte und rasterbasierte Verfahren unterschieden (siehe Abbildung 3-6). Dabei ist das gemeinsame Ziel der Clusteranalyseverfahren, eine Gruppierung zu erreichen, welche eine möglichst große Homogenität von Objekten innerhalb derselben Gruppe und eine möglichst große Heterogenität der unterschiedlichen Gruppen zueinander ergibt. Eine der Hauptfragen, welche es bei der Durchführung einer Clusteranalyse zu beantworten gilt, ist die Festlegung einer geeigneten Anzahl an Gruppen (häufig auch Klassen genannt). Auch die für eine Bewertung der Gruppierungsgüte, also der inneren Homogenität und äußeren Heterogenität, nötigen Verfahren sowie die darauf aufbauenden Methoden zur Festlegung einer geeigneten Gruppenanzahl werden nachfolgend vorgestellt. (Vgl. Han et al. 2012; Hennig 2015; Behnisch 2008)

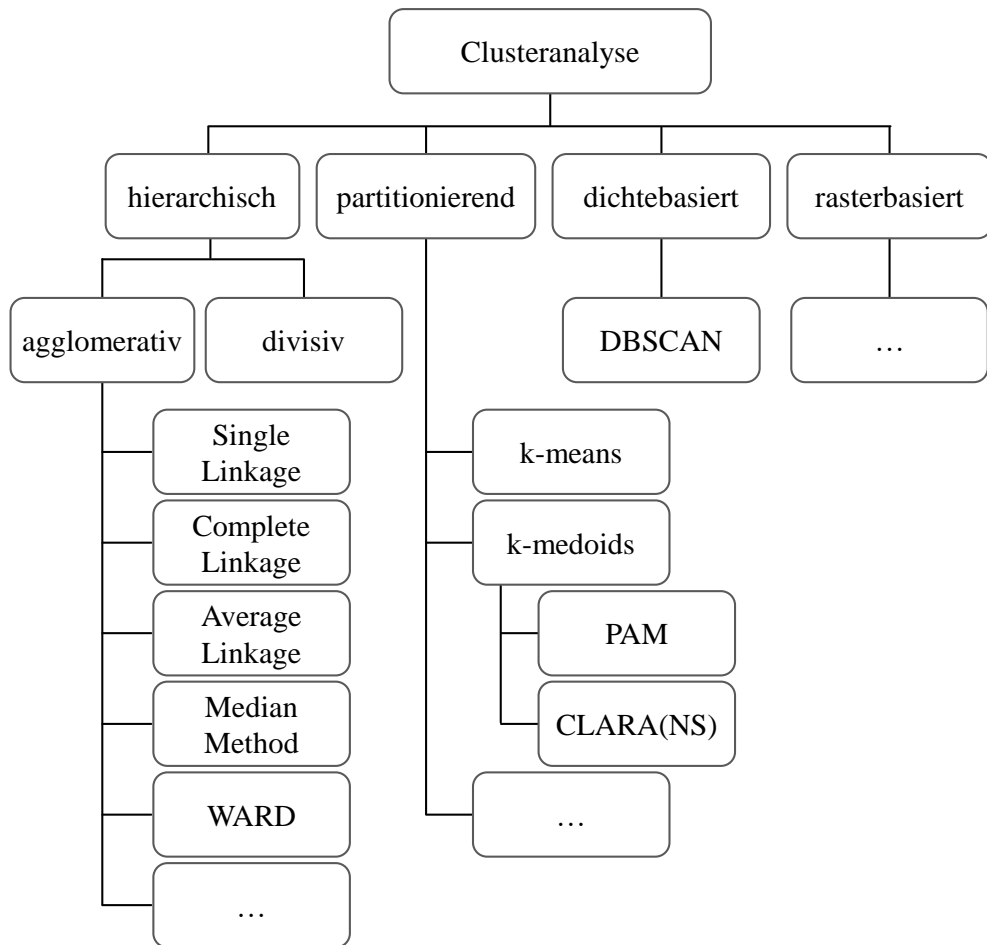


Abbildung 3-6: Gruppen und Algorithmen von Clusteranalyseverfahren [eigene Darstellung in Anlehnung an Backhaus et al. (2018, S. 457) und Behnisch (2008, S. 49)]

3.3.2 Hierarchische Clusteranalyseverfahren

Das Grundprinzip der Clusteranalyse mit hierarchischen Verfahren basiert auf der Erstellung einer Distanzmatrix für alle Objekte eines Datensatzes. Mittels eines festzulegenden Distanzmaßes wird die Unterschiedlichkeit aller Objekte zu allen anderen Objekten festgestellt. Ausgehend vom Urzustand, in dem jedes Objekt eine eigene Gruppe darstellt, folgt ein iterativer Prozess, in dessen Verlauf die Elemente zu Gruppen vereinigt werden. Im ersten Schritt der agglomerativen-hierarchischen Algorithmen werden die beiden Objekte des Datensatzes, welche insgesamt die geringste Distanz zueinander aufweisen – sich also am ähnlichsten sind – zu einer Gruppe vereinigt. Ein Datensatz von n Elementen wird somit nach der ersten Iteration reduziert auf $n-1$ Gruppen. In jeder weiteren Iteration werden nachfolgend wieder die Objekte der geringsten Distanz miteinander vereinigt, wobei die Objekte in diesen Schritten sowohl noch nicht vereinigte Einzelobjekte des Ursprungsdatensatzes als auch die bereits gebildeten Gruppen aus zwei oder mehr Elementen sind. Welche Elemente für die Distanzberechnung zwischen einer bereits gebildeten Gruppe und einem weiteren Einzelelement oder einer weiteren Gruppe herangezogen werden, ist Teil einer vor

der Anwendung des agglomerativen Clusteralgorithmus festzulegenden Verfahrensauswahl. Es können dabei die am nächsten beieinander liegenden (single linkage) oder die am weitest entfernten (complete linkage) Objekte der Gruppe(n) als distanzbestimmende Objekte definiert werden. Aber auch die Berücksichtigung von Distanzen zwischen den Mittelpunkten (im Sinne des empirischen Mittelwertes; average linkage) oder der Schwerpunkte (im Sinne eines Medians; median method) der Gruppen können für die Bestimmung der Zwischengruppendistanz herangezogen werden. Verfahren, die von einer einzelnen Gruppe, welche alle Objekte des Datensatzes enthält, beginnen und eine Unterteilung je Verfahrensschritt vornehmen, werden als *divisiv* bezeichnet. (Vgl. Han et al. 2012; Kaufman und Rousseeuw 1990; Backhaus et al. 2018)

Weitere Methoden, z. B. die Ward-Methode, stützen sich bei der Vereinigung der Objekte zu Gruppen bzw. bei der Vereinigung von bereits in einer vorherigen Iteration gebildeten Gruppen nicht auf die Distanz als Kriterium. Bei der Ward-Methode wird die bestehende Varianz innerhalb der existierenden Gruppen untersucht und die Zuordnung eines weiteren Objekts zu einer Gruppe nach dem Kriterium des minimalen Varianzzuwachses bei einer Vereinigung vorgenommen. Es werden also jene Objekte oder Gruppen vereinigt, die insgesamt zu einer geringstmöglichen Zunahme der Varianz innerhalb aller Gruppen führen. (Vgl. Legendre und Legendre 1998, S. 329–333)

Um eine abschließende Gruppeneinteilung zu erhalten, ist es erforderlich, den agglomerativen Verfahren eine Gruppenanzahl k vorzugeben. Somit wird der Vereinigungszustand für ebendiese Gruppenanzahl den Objekten des Datensatzes als jeweilige Gruppenzugehörigkeit zugewiesen. Die Darstellung der Hierarchiebeziehung der einzelnen Objekte sowie die Vereinigungsreihenfolge erfolgt i. d. R. in sog. Dendrogrammen (siehe Abbildung 3-7, S. 47). Alle Objekte des Ursprungsdatensatzes werden in diesem Diagramm auf der Abszisse abgebildet und deren Vereinigung durch ein rechtwinkliges Liniengeflecht dargestellt. An der Lage der Vereinigungspunkte auf der Ordinate lässt sich die Reihenfolge der Gruppenbildung und die Gruppeneinteilung für die festgelegte Gruppenanzahl erkennen. (Vgl. Han et al. 2012; Legendre und Legendre 1998)

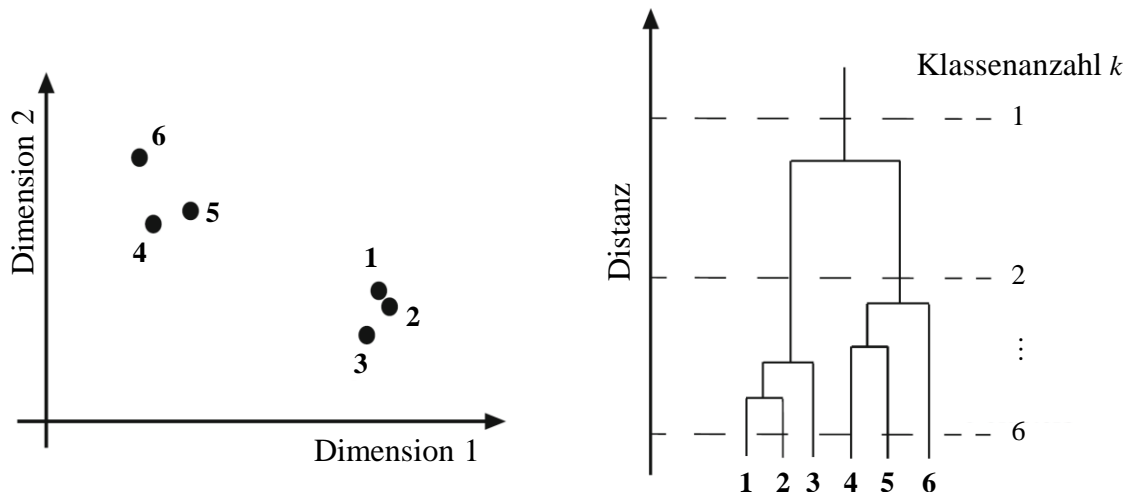


Abbildung 3-7: Datenpunkte (links) und zugehöriges Dendrogramm (rechts) als Beispiel der Hierarchiebeziehung und Vereinigungsreihenfolge im Rahmen der hierarchischen Gruppenbildung [Forsyth (2019), angepasste Darstellung]

3.3.3 Partitionierende Clusteranalyseverfahren

Die partitionierenden Clusteranalyseverfahren verfolgen ein von den hierarchischen Verfahren abweichendes Konzept zur Objekt- bzw. Gruppenvereinigung. Das methodische Grundprinzip der partitionierenden Verfahren besteht darin, dass die Clusterzugehörigkeit der Objekte von Iterationsschritt zu Iterationsschritt geändert werden kann, wenn dies zur Erfüllung der Zielfunktion beiträgt.¹⁹ Eines der bekanntesten und am häufigsten eingesetzten partitionierenden Verfahren stellt der „k-means“-Algorithmus dar. Die Zielfunktion des Algorithmus ist die Minimierung der quadrierten euklidischen Distanz, also die Maximierung der Homogenität innerhalb der Gruppen bei gleichzeitiger Maximierung des Abstandes zu anderen Gruppen. Zur Reduktion der benötigten Rechenzeit werden jedoch nicht alle beliebigen Zuordnungsmöglichkeiten für alle Objekte des Datensatzes und die vorgegebene Gruppenanzahl k berechnet. Vereinfachend wird für jede der k zu erzeugenden Gruppen der Mittelpunkt (engl. centroid) über eine Mittelwertbildung berechnet und als Bezugsobjekt für die weiteren Schritte des Algorithmus festgelegt. Im ersten Schritt der Ausführung werden zufällig k Elemente als Startpunkte für die Gruppenbildung gewählt. Danach folgt eine Berechnung der Distanz aller Objekte zu den Clustermittelpunkten (bzw. im ersten Schritt zu den Startpunkten) mittels euklidischer Distanz und eine Zuordnung der Objekte zur Gruppe des nächstgelegenen Mittelpunktes. Im nächsten Iterationsschritt

¹⁹ Im Unterschied zu dieser „globalen“ Anpassung der Gruppenzugehörigkeit bei den partitionierenden Clusteranalyseverfahren erfolgt bei den hierarchischen Methoden lediglich eine „lokale“ Gruppeneinteilung im jeweiligen Iterationsschritt. Bestehende Gruppen werden beim agglomerativen Vorgehen nur um neue Objekte erweitert, Objekte aber in einer späteren Iteration nicht von einer zur anderen Gruppe verschoben.

werden für die formierten Gruppen neue Mittelwerte berechnet und diese für eine erneute Distanzberechnung genutzt. Der Algorithmus ist damit in der Lage, lokale Minima zu finden und besitzt in vielen Softwareimplementierungen die Möglichkeit, eine maximale Anzahl an Iterationen als Abbruchkriterium sowie ein mehrfaches Ausführen mit unterschiedlichen Startpunkten²⁰ festzulegen. (Vgl. Han et al. 2012; Kaufman und Rousseeuw 1990; Zhou 2021)

Anstelle eines errechneten Gruppenmittelpunktes kann das partitionierende Verfahren auch eines der Objekte, welches dem Gruppenmittelpunkt am ähnlichsten ist, als Mittelpunkt bestimmen. Dieses Objekt wird dann als „Medoid“ bezeichnet und dient als Bezugsobjekt für die Feststellung der Ähnlichkeit mit den Objekten des Datensatzes. Das Verfahren wird deshalb „k-medoids“-Algorithmus bzw. „Partitioning Around Medoids“ (PAM) genannt. Neben der Festlegung von Mittelpunkten unterscheidet sich zudem die Zielfunktion: Es wird die Summe der Distanz aller Objekte des Datensatzes zu den jeweiligen Medoiden minimiert. Da die benötigten Rechenoperationen bei großen Datensätzen leicht über die Grenze verfügbarer Computerleistung hinausgehen, kann die Suche optimaler Medoide auch an einer (möglichst repräsentativen) Teilmenge des Datensatzes durchgeführt werden. Die Zuordnung aller Objekte des Datensatzes zu den ermittelten Gruppen erfolgt dann erst nach Bestimmung der Medoide am reduzierten Datensatz. Diese Erweiterung des Algorithmus wurde softwaretechnisch unter dem Namen „Clustering Large Applications“ (CLARA)²¹ umgesetzt. (Vgl. Kaufman und Rousseeuw 1990)

3.3.4 Dichtebasierte Clusteranalyseverfahren

Dichtebasierte Clusteranalyseverfahren, z. B. der Algorithmus „Density Based Spatial Clustering of Applications with Noise“ (DBSCAN)²², verfolgen ein unterschiedliches Konzept zur Bildung von Gruppen. Während die vorherigen Verfahren die Abstände aller Objekte zu allen anderen Objekten oder Gruppen bzw. Centroiden oder Medoiden für eine Gruppenzuordnung nutzen, wird für den Algorithmus DBSCAN der Suchradius ϵ definiert, der die zu untersuchende Umgebung eines jeden Objektes im Datensatz festlegt. Die

²⁰ Zwar erhöht sich durch die mehrfache Ausführung die Rechenzeit bis zur abschließenden Gruppeneinteilung, jedoch wird das Risiko verringert, ein ungünstiges lokales Minimum der Intraclustervarianz als finale Gruppeneinteilung zu verwenden.

²¹ Eine Erweiterung des CLARA-Algorithmus ist unter der Bezeichnung CLARANS von Ng und Han (2002) veröffentlicht worden.

²² Das Attribut „Spatial“ bezieht sich dabei nicht auf eine raumbezogene Analyse im Sinne der Geoinformationstechnik, sondern auf die Untersuchung einer hinsichtlich des Abstandes vom Objekt begrenzten Umgebung im n-dimensionalen Merkmalsraum (wobei n der Anzahl der Objektvariablen entspricht).

Objekte des Datensatzes werden dabei in einer zufälligen Reihenfolge ausgewählt und überprüft. Liegen andere Objekte innerhalb des Suchradius des ausgewählten Objektes, werden sie als „Dichte erreichbar“ definiert und es wird eine Objektkette gebildet. Für die jeweils nächsten Elemente der Kette wird die zu untersuchende Umgebung wiederum auf weitere Objekte überprüft. Neben dem Radius muss zudem eine vom Anwender festgelegte Mindestobjektanzahl (minimum number of points, *MinPts*) innerhalb der Objektumgebung vorhanden sein, damit ein Objekt für die Bildung bzw. Fortführung einer Objektkette genutzt wird. Regionen im n-dimensionalen Raum, in denen Objekte dicht beieinander liegen, werden somit zu einer Gruppe formiert. Hat sich aus einer Objektkette eine Gruppe bzw. ein Cluster konstituiert, so wird zufällig ein anderes, noch keiner Gruppe zugeordnetes Objekt ausgewählt. Nicht erreichbare Objekte werden als potenzielle Ausreißer markiert. Im Gegensatz zu den oben beschriebenen Verfahren, welche auf einem globalen Distanzkriterium beruhen und somit zu einer bevorzugt radialen Gestalt²³ der Gruppen führen, können dichte-basierte Clusteranalyseverfahren andere Strukturen innerhalb des Datensatzes besser nachvollziehen. Herausforderungen bei der Anwendung von DBSCAN bestehen in der Festlegung der beiden Verfahrensparameter ϵ und *MinPts* und einer hohen Sensitivität der Ergebnisse gegenüber der festgelegten Parameterwerte. (Vgl. Han et al. 2012; Sander 2016; Khan et al. 2018)

Die charakteristischen Eigenschaften der drei grundlegend verschiedenen Arten von Clusteranalyseverfahren werden in Tabelle 3-3²⁴ (S. 50) zusammengefasst. Zu den jeweiligen Algorithmen existieren weiterhin Anpassungen, z. B. zur Reduktion der Rechenzeit oder zur besseren Berücksichtigung von im Kontext spezifischer Fragestellungen vorliegenden Besonderheiten. Andere wichtige Weiterentwicklungen der Clusteranalyseverfahren sind Methoden zur Behandlung von Datensätzen mit gemischtskalierten Merkmalen. Als weiterführende Literatur für ergänzende Algorithmen bzw. Anpassungen sowie für Vergleichsstudien von Clusteranalyseverfahren wird auf Hennig (2015), Han et al. (2012), Giordani et al. (2020) und die in Tabelle Anhang B-1 genannten Quellen verwiesen.

²³ Im zweidimensionalen Raum entspricht dies einer kreisförmigen, im dreidimensionalen Raum einer sphärischen Form der Gruppen.

²⁴ Raster- bzw. gitterbasierte Methoden kommen bisher nur sehr selten zur Anwendung und eignen sich nicht für die Untersuchung in dieser Arbeit, weshalb sie in der zusammenfassenden Tabelle nicht aufgeführt sind.

3 Methoden zur Gruppenbildung und Klassifizierung

Tabelle 3-3: Charakteristische Eigenschaften der Clusteranalyseverfahren nach Gruppen [eigene Darstellung basierend auf Han et al. (2012)]

Clusteranalyseverfahren	Charakteristische Eigenschaften
Hierarchische Methoden	<ul style="list-style-type: none"> • Die Gruppeneinteilung folgt einem hierarchischen Ordnungsprinzip. • Gebildete bzw. aufgelöste Gruppen können nicht noch einmal verändert werden. • Es existieren verschiedene Ordnungsprinzipien. • Eine vollständige Gruppierung mit sich gegenseitig ausschließender Gruppenzugehörigkeit wird erzeugt.
Partitionierende Methoden	<ul style="list-style-type: none"> • Eine abstandsbaasierte Gruppeneinteilung erfolgt und führt zu einer „sphärischen“ Form der Gruppen. • Eine vollständige Gruppierung mit sich gegenseitig ausschließender Gruppenzugehörigkeit wird erzeugt. • Eine Nutzung von Bezugsobjekten (z. B. Mittelwert oder Median) zur Rechenzeitverkürzung ist möglich.
Dichtebasierte Methoden	<ul style="list-style-type: none"> • Es erfolgt die Bildung von Gruppen in Bereichen mit hoher Objektdichte, welche durch Bereiche mit geringer Objektdichte abgegrenzt sind. • Die Form der Gruppen ist frei und entspricht der Form des räumlichen Dichteverlaufs. • Die Festlegung einer Mindestanzahl von Nachbarn sowie eine Erkennung von Ausreißern ist möglich. • Ausreißer können als nicht gruppierte Elemente behandelt werden, in diesem Fall liegt keine vollständige Gruppierung vor.

3.3.5 Gütekriterien der Gruppenfestlegung

Zur Beurteilung der Güte der Gruppeneinteilung durch einen Clusteranalysealgorithmus bzw. für eine getroffene Auswahl von Verfahrensparametern wurden ebenfalls Kriterien und Methoden entwickelt. Die am häufigsten verwendeten Verfahren sollen hier skizziert werden.

Die „mittlere Silhouetten-Weite“ (engl. average silhouette width; ASW; auch Silhouettenkoeffizient) wurde von Kaufman und Rousseeuw (1990) als Kriterium zur Beurteilung einer geringen Intracusterhomogenität bei möglichst hoher Interclusterheterogenität entwickelt. Die Silhouetten-Weite sw wird dabei definiert als die Differenz des geringsten Abstandes (d) eines Objekts x im Cluster C_i mit $i = (1, \dots, k)$ zum ähnlichsten Objekt einer benachbarten Gruppe C_j mit $j = (1, \dots, k)$ sowie $j \neq i$ und dem mittleren Abstand zu allen Objekten der eigenen Gruppe dividiert durch das Maximum dieser beiden Abstandswerte (Gl. 3-9 bis Gl. 3-11). Die ASW entspricht dann dem Mittelwert der Silhouetten-Weite über alle Objekte 1 bis n eines Datensatzes (Gl. 3-12, S. 51). In einer vergleichenden Studie von Arbelaitz et al. (2013) wurde dieses Bewertungskriterium zur Festlegung einer

zweckmäßigen Gruppenanzahl als besonders gut geeignet bewertet. (Vgl. Kaufman und Rousseeuw 1990; Akhanli und Hennig 2020; Arbelaitz et al. 2013; Han et al. 2012).

$$sw(x) = \frac{b(x) - a(x)}{\max\{a(x), b(x)\}} \quad \text{Gl. 3-9}$$

mit

$$a(x) = \frac{\sum_{x' \in C_i, x' \neq x} d(x, x')}{|C_i| - 1} \quad \text{Gl. 3-10}$$

$$b(x) = \min_{j \in [1, k], j \neq i} \left\{ \frac{\sum_{x' \in C_j} d(x, x')}{|C_j|} \right\} \quad \text{Gl. 3-11}$$

$$I_{ASW} = \frac{1}{n} \sum_{x=1}^n sw(x) \quad \text{Gl. 3-12}$$

Wird der Indikator I_{ASW} iterativ für verschiedene Gruppenanzahlen k berechnet, kann die Gruppenanzahl mit der größten, d. h. besten mittleren Silhouetten-Weite ermittelt und für die finale Gruppeneinteilung der Objekte des Datensatzes in Gruppen genutzt werden (vgl. Charrad et al. 2014). Ein Beispiel für die grafische Ermittlung der optimalen Gruppenanzahl mittels ASW wird in Abbildung 3-8 (links, S. 52) gegeben. Im gezeigten Beispiel liegt ein globales und ein lokales Optimum vor. In solchen Fällen muss der Anwender entscheiden, welche Gruppenanzahl und welches Gruppierungsergebnis der Beantwortung der Forschungsfrage am dienlichsten sind.

Neben der mittleren Silhouetten-Weite gibt es eine Vielzahl weiterer Verfahren. In Studien findet häufig das sog. „Elbow“-Kriterium Anwendung. Diesem Kriterium liegt eine Untersuchung der Intraclustervarianz zugrunde, welche der Summe des quadratischen Abstandes aller in einem Cluster liegenden Objekte zum Mittelpunkt des Clusters entspricht (engl. within sum of squares; WSS). Der Grundgedanke der Elbow-Methode besteht darin, dass bei steigender Gruppenanzahl die Reduktion der WSS abnimmt, weitere Gruppen also die Homogenität innerhalb der Gruppen bei steigender Fragmentierung des Datensatzes nur gering verbessern. Als optimale Gruppenanzahl k_{opt} wird dabei der Wert angenommen, an welchem der Gradient der WSS von k_i zu k_{i+1} am größten ist.²⁵ Meist wird die Ermittlung der optimalen Gruppenanzahl k_{opt} in Studien in einem grafischen Verfahren durch den Anwender selbst vorgenommen. (Vgl. Han et al. 2012; Akhanli und Hennig 2020; Ullmann et al. 2021)

²⁵ Diese Stelle ist in der grafischen Festlegung der Klassenanzahl meist als Knick (Elbow) erkennbar.

In Abbildung 3-8 (rechts) wurde die Visualisierung des Gradienten der WSS durch die unter der Kurve liegenden Winkel α in Grau für jede Gruppenanzahl vorgenommen. Der größte Winkel zeigt die optimale Gruppenanzahl k_{opt} an (dunkelgrau).

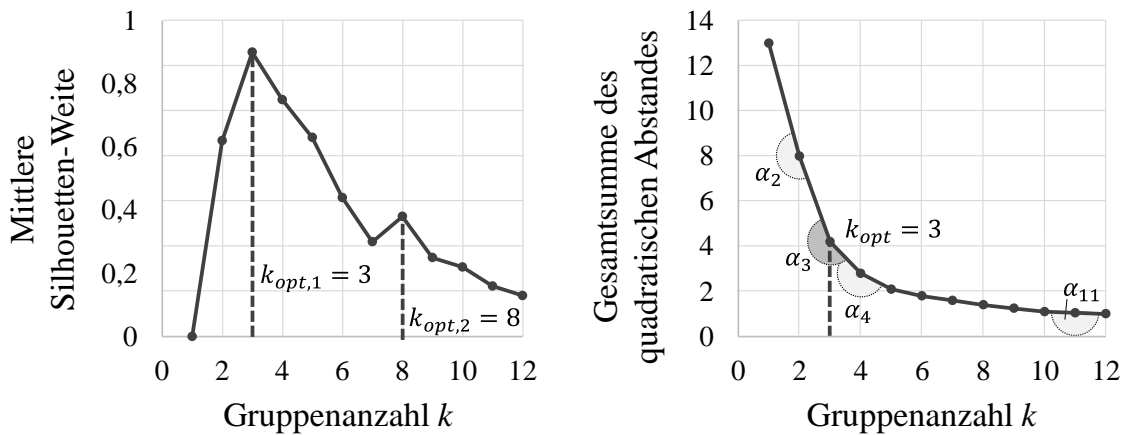


Abbildung 3-8: Festlegung einer geeigneten Gruppenanzahl durch die Auswertung der mittleren Silhouetten-Weite (links) und mittels der Elbow-Methode (rechts) [eigene Darstellung]

Für eine Übersicht weiterer Indizes zur Bewertung der Güte von Gruppierungen wird insbesondere auf Charrad et al. (2014), Akhanli und Hennig (2020) und Hennig (2022) verwiesen. Die iterative Untersuchung von Indexwerten zur Festlegung geeigneter Gruppenanzahlen kann zudem auch mittels anderer Indizes vorgenommen werden (vgl. Giordani et al. 2020; Charrad et al. 2014). Weiterhin diskutieren einige der in Tabelle Anhang B-1 benannten Studien ebenfalls diese Thematik.

3.4 Überwachtes Maschinelles Lernen zur Datenklassifizierung

Die in Kapitel 3.3 dargestellten Methoden dienen der Analyse von Objekten eines Datensatzes und der Bestimmung von Gruppen. Soll der Datensatz um zusätzliche Objekte erweitert werden oder ein anderer Datensatz mit gleichen Merkmalen und einer ähnlichen Verteilung der Eigenschaften in diese bereits ermittelten Gruppen „einsortiert“ werden, so bieten sich verschiedene Verfahren an. Neben den klassischen statistischen Verfahren, wie der Regression oder generalisierten linearisierten Modellen, können verschiedene überwachte Methoden des Maschinellen Lernens (ML) eingesetzt werden. Auch in diesem Forschungs- bzw. Arbeitsfeld wurde in den vergangenen Jahrzehnten eine Vielzahl an verschiedenen Methoden und Weiterentwicklungen erarbeitet, welche heute für unterschiedliche Forschungskontexte angewendet werden (vgl. Forsyth 2019; Zhou 2021). In den nachfolgenden Abschnitten wird deshalb nur eine begrenzte Einführung gegeben und eine Auswahl der Verfahren, welche für die Aufgabenstellung besonders geeignet sind (siehe Kapitel 5), im Detail vorgestellt.

Überwachtes ML kann als ein Prozess der Datenverarbeitung beschrieben werden, bei dem eine bekannte Datenstruktur genutzt wird, um ein möglichst genaues Vorhersagemodell zu entwickeln, welches bei begrenztem Bedarf an informationstechnischen Ressourcen eine robuste Einteilung oder Bewertung von neu hinzugekommenen Daten der gleichen Datenstruktur erlaubt. Die grundlegenden Prinzipien, nach denen das Vorhersagemodell eines überwachten ML-Algorithmus funktioniert, können sich dabei jedoch stark unterscheiden und deshalb in die in Abbildung 3-9 dargestellten Arten eingeteilt werden. Die Anwendungsgebiete von überwachten ML-Verfahren sind dabei heute vielfältig und beinhalten z. B. die Text-, Bild- und Spracherkennung sowie die Erkennung biometrischer Daten, die technische, logistische und/oder betriebswirtschaftliche Entscheidungsunterstützung sowie die Fehlerüberwachung und Wartungsprognose technischer Systeme. (Vgl. Forsyth 2019; Bell 2014; Zhou 2021; Kant Hiran et al. 2021; Yang et al. 2018)

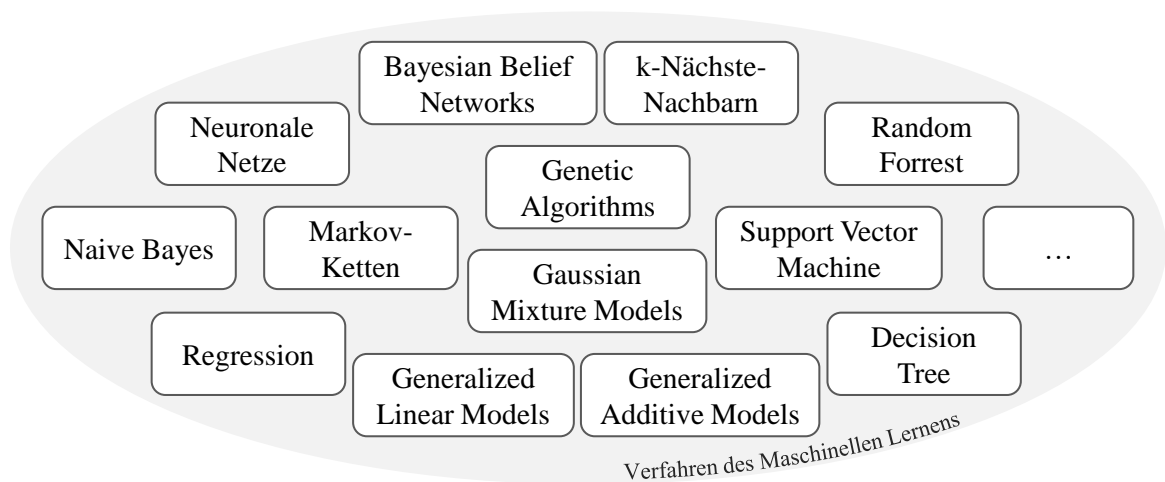


Abbildung 3-9: Häufig verwendete überwachte Verfahren des Maschinellen Lernens [eigene Darstellung]

Wie in Abbildung 3-9 dargestellt kann bei den überwachten ML-Verfahren zudem eine Unterteilung in regelbasierte und verteilungsbasierte Algorithmen vorgenommen werden. Zu den bekanntesten Vertretern der regelbasierten Algorithmen gehören die Verfahren „Decision Tree“ (dt. Entscheidungsbaum) und „Random Forrest“. Beiden liegt zugrunde, dass durch den Algorithmus Unterscheidungskriterien definiert und für die verschiedenen Merkmale festgelegt werden, welche nacheinander ausgeführt zu einer möglichst hohen Übereinstimmung der ursprünglichen Eingruppierung der Daten und der durch den Algorithmus vorgenommenen Eingruppierung führen. Während der „Entscheidungsbaum“-Algorithmus für einen Testdatensatz einen einzelnen Entscheidungsbaum erstellt, variiert der „Random Forrest“-Algorithmus in mehreren Durchläufen zufällig die Teilmenge des Datensatzes, die zur Erstellung eines Entscheidungsbaumes verwendet wird. Für jeden der Durchläufe wird ein Entscheidungsbaum erzeugt und im Anschluss ermittelt, welche Merkmale und welche Unterscheidungsregeln für die einzelnen Merkmale am stärksten zu einer hohen Übereinstimmung der Gruppierungen geführt haben. Mit diesem Klassifizierungs-

modell können dann weitere Daten klassifiziert werden. (Vgl. Forsyth 2019; Maheswari Viswanadhula et al. 2021)

Der „k-Nächste-Nachbarn“-Algorithmus wird als einer der einfachsten überwachten ML-Algorithmen beschrieben (vgl. Maheswari Viswanadhula et al. 2021). Das Grundprinzip erfordert nur wenige Rechenschritte zur Bestimmung der Distanz eines zu klassifizierenden Objekts zu den Objekten eines bereits klassifizierten Trainingsdatensatzes. Als Eingabeparameter des Algorithmus wird eine Anzahl k an Nachbarn des zu klassifizierenden Objekts festgelegt. Für jedes dieser k Nachbarobjekte wird nun deren Klasseneinteilung abgefragt und die Klasse mit der höchsten Häufigkeit als Klasse für das zu klassifizierende Objekt verwendet (siehe Abbildung 3-10). Das zu verwendende Distanzmaß (siehe auch Kapitel 3.2) kann dabei ebenso festgelegt werden wie die Anzahl der zu berücksichtigenden Nachbarn. Bei der Anwendung dieses ML-Verfahrens muss jedoch beachtet werden, dass in höherdimensionalen Räumen, d. h. für Objekte mit einer Vielzahl an Variablen, die Aussagekraft der Nachbarschaftsbeziehung abnehmen kann und somit der Fehler bei der Klassifizierung i. d. R. zunimmt. (Vgl. Forsyth 2019; Zhou 2021)

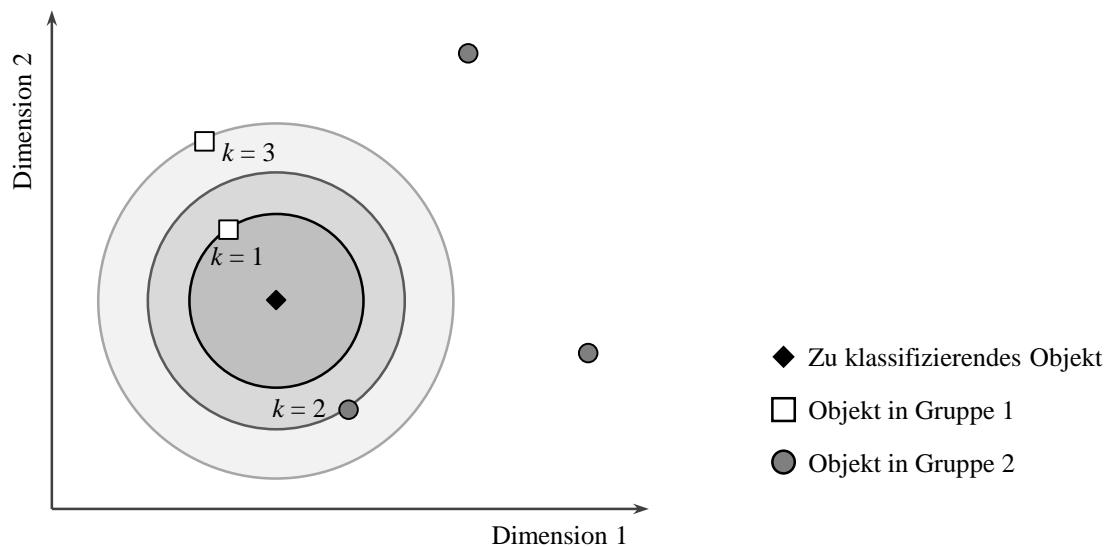


Abbildung 3-10: k -Nächste-Nachbarn-Zuordnung mit Äquidistanzflächen des euklidischen Abstandes im zweidimensionalen Merkmalsraum (für $k=3$ erfolgt die Zuordnung zu Gruppe 1) [eigene Darstellung in Anlehnung an Zhou (2021)]

Ein weiteres häufig verwendetes Verfahren aus der Gruppe des überwachten ML ist der „Naive Bayes Klassifizierer“. Das Grundprinzip beruht auf dem Bayes Theorem, d. h. auf der Annahme von bedingten Wahrscheinlichkeiten: Liegen (Vor-)Informationen zu einem der Merkmale eines Objektes vor, so lässt sich darüber die Wahrscheinlichkeit abschätzen, dass ein anderes Merkmal einen Wert innerhalb eines bestimmten Wertebereichs annimmt. Zudem wird für die Nutzung des Naive Bayes Klassifizierers angenommen, dass alle Variablen unabhängig sind.

Auf Basis dieser Grundannahmen kann die Wahrscheinlichkeit bestimmt werden, mit der ein zu klassifizierendes Objekt einer bereits bekannten Klasse angehört. Die Zielfunktion (hier Kostenfunktion genannt) ist dabei die Maximierung der Zugehörigkeitswahrscheinlichkeit, also die Zuordnung eines Objekts zu der Klasse, für welche die Zugehörigkeitswahrscheinlichkeit am höchsten ist. Die Wahrscheinlichkeit $P(C_i|X)$, dass ein Objekt mit den Merkmalsausprägungen $X = (x_1, \dots, x_n)$ zur Klasse i mit $i = (1, \dots, k)$ gehört, wird A-posteriori-Wahrscheinlichkeit genannt und ergibt sich aus dem im Folgenden dargestellten Zusammenhang (siehe Gl. 3-13, vgl. Zhou 2021, S. 159–160; Hedderich und Sachs 2020, S. 184–186). Die sog. A-priori-Wahrscheinlichkeit $P(C_i)$ gibt die Wahrscheinlichkeit an, mit der ein Objekt zur Klasse i gehört. Diese Wahrscheinlichkeit kann im Falle des überwachten Lernens z. B. aus der Klassenbesetzung eines bereits klassifizierten Datensatzes abgeleitet werden. Multipliziert wird diese A-Priori-Wahrscheinlichkeit mit der bedingten Wahrscheinlichkeit²⁶ $P(X|C_i)$ für die Zugehörigkeit der Merkmalsausprägungen in X zur Klasse C_i . Dieses Produkt wird weiterhin durch die Summe der Wahrscheinlichkeiten einer Zugehörigkeit zu den k feststehenden Klassen $P(X)$ dividiert. Der Klasse C_i mit der höchsten A-posteriori-Wahrscheinlichkeit wird das Objekt dann zugeordnet. (Vgl. Hedderich und Sachs 2020; Zhou 2021; Boehmke 2022; Zhang 2004; Berthold et al. 2020; Kant Hiran et al. 2021; Backhaus et al. 2018)

$$P(C_i|X) = \frac{P(X|C_i) \cdot P(C_i)}{\sum_{i=1}^k P(X|C_i) \cdot P(C_i)} = \frac{P(X|C_i) \cdot P(C_i)}{P(X)} = \frac{P(C_i)}{P(X)} \prod_{j=1}^n P(x_j|C_i)$$

Gl. 3-13

Der Naive Bayes Klassifizierer erlaubt und erfordert es, für eine hohe Güte der Klassifizierung auch für die unterschiedlichen Variablen eigene A-Priori-Wahrscheinlichkeitsverteilungen zugrunde zu legen. Hierbei können für verschiedene Skalenniveaus angepasste Verteilungstypen verwendet werden. Ein weiterer Vorteil des Naive Bayes Klassifizierers ist, dass er mit einer geringfügigen Anpassung auch Merkmalsausprägungen berücksichtigen kann, die im vorherigen Trainingsdatensatz nicht vorhanden waren (x_j') und somit eine Zugehörigkeitswahrscheinlichkeit $P(x_j'|C_i)$ zur Klasse C_i von 0 aufweisen würden (vgl. Zhou 2021). Für diesen Fall kann der als „additive Glättung“ oder „Laplace’sche Glättung“ bezeichnete Ansatz angewendet werden. Hierbei wird für die unbekannte Merkmalsausprägung ein ein- oder mehrmaliges Auftreten im Trainingsdatensatz „vorgetäuscht“ und somit

²⁶ Für die Beschreibung dieser bedingten Wahrscheinlichkeit wird auch im Deutschen häufig der englischsprachige Begriff „Likelihood“ verwendet, um eine klarere Unterscheidung zwischen den unterschiedlichen Konzepten einer Wahrscheinlichkeit der Realisierung eines Merkmalswertes und der bedingten Wahrscheinlichkeit einer Klassenzugehörigkeit zu erreichen.

die Zugehörigkeitswahrscheinlichkeit dieser Merkmalsausprägung zu jeder der k Klassen größer als 0. Bei einer relativ großen Menge an Objekten und Variablen führt dieser additive Term insgesamt aber aufgrund des geringen Einflusses auf die totale Wahrscheinlichkeit $P(X)$ nur zu einer äußerst geringen Veränderung der A-Posteriori-Wahrscheinlichkeiten für alle anderen Objekte. Ein Wechsel von einer Klasse i zu einer Klasse $h \neq i$ würde erst dann erfolgen, wenn die A-Posteriori-Wahrscheinlichkeit der Klasse h nach der Erweiterung mit dem additiven Term größer würde als die Wahrscheinlichkeit der Zugehörigkeit zur Klasse i . (Zhou 2021; Boehmke 2022; Mijaka 2020)

Bei vollständigem Wissen über die A-Priori-Verteilungen und die bedingten Wahrscheinlichkeiten ist der Naive Bayes Klassifizierer ideal. Demzufolge wird diese optimale Klassifikation auch als obere Schranke und Referenz für die Güte anderer überwachter ML-Verfahren genutzt. Dieses Wissen ist allerdings in der Realität nicht mit endlichem Aufwand erreichbar, weshalb in der Praxis eingesetzte Naive Bayes Klassifizierer hinsichtlich der Güte der Klassifizierung unterhalb dieser Schranke bleiben. Jedoch zeigt das Verfahren auch bei Datensätzen, bei denen Variablen nicht durchweg unabhängig sind, noch sehr gute Klassifizierungsergebnisse und wird deshalb in vielen Forschungskontexten eingesetzt. (Zhang 2004; Forsyth 2019; Zhou 2021; Mijaka 2020)

Ausführliche Beschreibungen zu den vorgestellten überwachten ML-Verfahren sowie anderer Methoden können der Fachliteratur entnommen werden (z. B. Forsyth 2019; Han et al. 2012; Zhou 2021; Maheswari Viswanadhula et al. 2021; Jiang 2021).

3.5 Anwendungsbeispiele für Clusteranalysen im Kontext der energetischen Bewertung von Gebäuden und Quartieren

Nach der Einführung in die Grundlagen der Methoden zur Gruppenbildung und zur Klassifizierung statistischer Daten sowie der Vorstellung einer Auswahl der wichtigsten Algorithmen sollen an dieser Stelle einige Studien zusammenfassend dargestellt werden, welche diese Methoden zur Beantwortung wissenschaftlicher Fragestellungen einsetzen. Aufgrund der vielfältigen Forschungskontexte wird dabei eine Auswahl getroffen, bei welcher der Fokus auf Themenfeldern liegt, die der Forschungsfrage dieser Arbeit naheliegen – d. h. einen Bezug zum Klimaschutz in Gebäuden und Quartieren haben und/oder Geodaten zur Analyse räumlicher Strukturen der gebauten Umwelt nutzen.

In der bereits zitierten Studie von Behnisch (2008) werden verschiedene Klassifizierungsverfahren genutzt, um einen Datensatz mit sozioökonomischen Daten sowie Geoinformationsdaten für Gemeinden in Deutschland zu analysieren und eine Zuordnung zu zuvor definierten Klassen bzw. Gruppen vorzunehmen. In einem ersten Anwendungsfall erzeugt Behnisch eine Einteilung von Gemeinden hinsichtlich des Merkmals „Beschäftigungsschwerpunkte nach Wirtschaftssektoren“ in sieben definierte Gruppen. Die Basis der Gruppenzuordnung bilden raumstrukturelle Merkmale auf Gemeindeebene und Merkmale des

Bevölkerungs- und Beschäftigungswachstums. Als Methode wendet Behnisch ein Gaußsches Mixtur-Modell zur Klassifizierung an, welches mit einer Zielfunktion zur Maximierung der Zugehörigkeitswahrscheinlichkeit zu einer der vorgegebenen Gruppen die Gemeinden den verschiedenen Gemeindetypen (z. B. Gemeinde mit Beschäftigungsschwerpunkt im Sektor Gewerbe, Dienstleistung und Handel) zuordnet und hierfür die jeweiligen Unterscheidungsgrenzen der Variablen nach mehrfachen Wiederholungen des Verfahrens optimal festlegt. Für die Klassifizierung hinsichtlich der Variablen zum Wachstum von Gemeinden nutzt Behnisch ein regelbasiertes künstliches neuronales Netz (KNN). Die Klassenzuordnung mittels KNN wird zudem mit einem dichte-basierten Clusteralgorithmus validiert (vgl. Behnisch 2008). In einer abschließenden Untersuchung bedient sich Behnisch zudem eines logarithmischen Regressionsverfahrens, um den deutschen Gebäudebestand anhand von verschiedenen nationalen, aber auch kleinräumigen statistischen Informationen sowie Geobasisdaten zu schätzen. Seine Arbeit schließt mit einer Gegenüberstellung und Diskussion der Ergebnisse. Dabei wird die Verteilung der Gemeinden auf die verschiedenen Gemeindetypen im Hinblick auf die Merkmale „Beschäftigungsschwerpunkt“, „Raumstruktur“, „Bevölkerungswachstum“ bzw. „Beschäftigungswachstum“ und „Gebäudezahl“ verglichen.

Hecht (2014) nutzt in seiner Studie Geobasisdaten und geotopografische Karten, um „Methoden zur automatischen Klassifizierung von Gebäudegrundrissen (zu) erarbeiten, (zu) analysieren und (zu) bewerten mit dem Ziel, diese für die gebäudebasierte Analyse der Siedlungsstruktur zu nutzen“ (Hecht 2014, S. V). Die „Verfahrensentwicklung [... erfolgt] vor dem Hintergrund des Einsatzes in der kleinräumigen Analyse der Siedlungsstruktur auf Grundlage von Geobasisdaten in Deutschland“ (Hecht 2014, S. 143). Hecht untersucht dabei unter anderem, aus welchen Datenquellen welche geometrischen, topologischen²⁷ und statistischen sowie semantischen Informationen zu Gebäuden, Gebäuderegionen²⁸, Baublöcken, Gebäudeumgebungen und Siedlungskörpern ermittelt werden können. Hierbei verwendet er die betrachteten und im späteren Verlauf der Arbeit im Rahmen der Klassifizierung angewendeten Datenbestände topografischer Karten, Hausumringe und Vektordaten des Amtlichen Topografisch-Kartografischen Informationssystems (ATKIS) sowie

²⁷ In diesem Kontext und im weiteren Verlauf der Arbeit beschreiben topologische Eigenschaften von Objekten Informationen zu deren räumlicher Lage in Bezug zu anderen Objekten desselben Datensatzes (z. B. die Anzahl von Gebäuden mit einem Abstand geringer oder gleich eines festgelegten (Such-)Radius).

²⁸ Der Begriff „Gebäuderegion“ beschreibt in der Studie von Hecht (2014) eine Vereinigung von Gebäuden und Gebäudeketten, welche direkt benachbart sind, d. h. deren Wände sich berühren.

ALKIS-Flächenpolygone für Gebäude und 3D-Gebäudemodelle.²⁹ Kategoriale Merkmale des Datensatzes werden bei Hecht für die spätere Anwendung der Klassifizierungsalgorithmen Dummy-codiert und fehlende Werte über ein k-Nächste-Nachbarn-Imputationsverfahren³⁰ ergänzt. Für einen Datensatz der festgelegten Datenstruktur wird eine softwaretechnische Routine festgelegt, die mit 16 verschiedenen Klassifizierungsalgorithmen – welche z. T. eine jeweils sehr spezielle Optimierung von Parametern erfordert oder für welche eine Berücksichtigung von Parameterintervallen vorgenommen wird – den Gebäudedatensatz im Hinblick auf eine vorher durch Hecht definierte Gebäudetypologie klassifiziert. Dieses Verfahren wird abschließend an sieben Gemeinden unterschiedlicher Flächenausdehnung, Einwohnerzahl sowie raumstruktureller Zugehörigkeit³¹ erprobt und ein Vergleich der Klassifizierungsalgorithmen hinsichtlich ihrer Klassifizierungsgüte sowie der benötigten Zeit für das Training und das Klassifizieren der untersuchten Daten vorgenommen. Anhand verschiedener Beurteilungs- bzw. Vergleichskriterien benennt Hecht (2014) den „Random Forrest“-Algorithmus für die durch ihn untersuchten Datensätze und Merkmalsauswahlen als das Verfahren mit der höchsten Klassifizierungsgüte. Aus weiterführenden Untersuchungen leitet er zudem ab, dass nur ein begrenzter Teil der ihm zur Verfügung stehenden Merkmale zu einer hohen Güte des Klassifikationsergebnisses beiträgt. Diese sind Merkmale zum Vorliegen einer amtlichen Hauskoordinate, zur Gebäudeanzahl in einer Gebäuderegion, zur Bebauungsart (freistehende oder geschlossene Bebauung), zur Nutzungskategorie, zur Gebäudefläche bzw. Fläche der Gebäuderegion sowie das Verhältnis des Umfangs eines Gebäudegrundrisses zur Grundrissfläche. (Vgl. Hecht 2014)

Die Forschungsarbeit Hechts kann als ein nützliches Beispiel für die Anwendung von Klassifizierungsalgorithmen zur Nachvollziehung bekannter sowie für einen Trainingsdatensatz vorliegender Gruppeneinteilungen verstanden werden. Im Hinblick auf die Ergebnisse der Arbeit von Hecht können jedoch keine Aussagen dazu getroffen werden, ob die verwendeten Trainingsdaten eine ausreichende Repräsentativität für eine Anwendung des dort erarbeiteten „Random Forrest“-Klassifizierungsmodells auf einen deutlich größeren, z. B. gesamtdeutschen Datensatz erlauben. Auch eine weiterführende Untersuchung von Quartiersstrukturen war nicht Teil von Hechts Arbeiten. Zwar zeigt er an den analysierten Gemeinden beispielhaft, wie sich aus der Anzahl bzw. Flächendichte der Gebäude innerhalb von

²⁹ Zum besseren Verständnis der hier benannten Datenquellen wird auf die Ausführungen in Kapitel 4 der vorliegenden Arbeit verwiesen.

³⁰ Bei Hecht (2014) wird der Datensatz im Falle fehlender Merkmalsausprägungen durch den Mittelwert der Merkmalsausprägungen der zehn ähnlichsten Objekte vervollständigt. Die ähnlichsten Objekte werden dabei über die Berechnung der euklidischen Distanz im n -dimensionalen Variablenraum ermittelt.

³¹ Es handelt sich dabei um die Kategorien „städtisch“, „kleinstädtisch“ und „ländlich“.

Klassen Bebauungsstrukturtypen, ähnlich z. B. denen von Roth, ableiten lassen. Eine Untersuchung der Durchmischung von Gebäuden und Bebauungsstrukturen sowie Abgrenzung von Quartieren ist jedoch nicht Teil der Studie.

Andere besonders nennenswerte Arbeiten sind die Studien von Jochem et al. (2020) und Jochem und Tatem (2021), in denen die Autoren Gebäudegrundrissepolygone hinsichtlich geometrischer und topologischer Eigenschaften analysieren und mit Hilfe eines Gaußschen Finite-Mixtur-Modells geografische Rasterzellen der untersuchten Städte bzgl. der Bebauungsstruktur in Gruppen einteilen (vgl. Jochem et al. 2020). Der bei der Ermittlung der topologischen Gebäudeeigenschaften zu berücksichtigende Radius um ein Gebäude sowie die Rasterauflösung können dabei beliebig angepasst werden (vgl. Jochem und Tatem 2021). Weiler et al. (2018), Weinand et al. (2019) und Cao et al. (2013) setzen ebenfalls Clusteranalyseverfahren für die Untersuchung und Gruppierungen von Gebäuden bzw. Gemeinden ein. Weiler et al. (2018) nutzen für ihre Untersuchung zur Bestimmung energetisch zusammenhängender Gebäudegruppen 3D-Geoinformationsdaten und leiten bzw. schätzen für diese aus ihrer Modellierungsumgebung „SimStadt“ auf Basis einer gebäudetypologischen Datenbank energierelevante Gebäudeeigenschaften ab. Diese dienen als Eingangsdaten für eine Faktorenanalyse, welche wiederum eine agglomerative-hierarchische Clusteranalyse speist. Die Festlegung der Gruppenanzahl wird dabei durch eine Auswertung von 25 Indizes vorgenommen. Die Gruppenbesetzung wird anschließend mit der Einteilung der Gebäude in der deutschen Gebäudetypologie des IWU nach Loga et al. (2015) verglichen. Dabei wird der Schluss gezogen, dass eine Typisierung der Gebäude mittels Clusteranalyse und auf Basis eines, das Gebäude hinreichend beschreibenden Datensatzes der bloßen intuitiven Zuordnung zu den Typgebäudekategorien vorzuziehen ist.

Der in Weiler et al. (2018) angewendete Verfahrensablauf wurde ursprünglich durch Weinand et al. (2019) im Kontext einer Clusteranalyse der deutschen Gemeinden entwickelt. Die Verfahrensschritte sind deshalb die gleichen. Die Arbeiten unterscheiden sich aber hinsichtlich der Variablenanzahl des Eingangsdatensatzes deutlich. Weinand et al. verwenden einen 59 Variablen umfassenden Datensatz, welcher sowohl sozioökonomische als auch gebäudeenergetische und andere energierelevante Eigenschaften der Gemeinden beschreibt. Die durchgeführte Faktorenanalyse leistet somit einen deutlich größeren Beitrag zur Reduktion der benötigten Rechenzeit bei der Berechnung der Objekthierarchie mittels Ward-Algorithmus als in der vorherigen Studie von Weiler et al. Das Ergebnis der Studie ist eine Einteilung aller Gemeinden Deutschlands in Gruppen, welche hinsichtlich sozioökonomischer und energetischer Merkmale ähnlich sind. Eine Übertragung von erfolgreichen Klimaschutzstrategien einer Gemeinde auf Gemeinden der gleichen Gemeindeklasse könnte somit zu einer schnelleren Umsetzung von Klimaschutzmaßnahmen führen.

Auch Cao et al. (2013) versuchen, eine Gruppierung von Gebäudegrundrissen zur Analyse städtischer Strukturen anhand von Clusteranalyseverfahren vorzunehmen. Der Fokus liegt

hierbei insbesondere auf der Einteilung des untersuchten Stadtraumes in verschiedene Gruppen – dies können Stadtteile oder Quartiere sein. In der Arbeit diskutieren Cao et al., dass in einer Clusteranalyse unter gleichzeitiger Beachtung der Gebäudekoordinaten und der Gebäudemerkmale Gruppen entstehen können, die zwar hinsichtlich der Gebäudeeigenschaften ähnlich sind, jedoch aufgrund ihrer räumlichen Entfernung vom Algorithmus in unterschiedliche Gruppen eingeteilt werden. Um diese Beschränkung der klassischen Clusteranalyseverfahren zu vermeiden, wenden sie ein mehrstufiges Verfahren an. Im ersten Schritt werden die Gebäude hinsichtlich ihrer Eigenschaften untersucht und diese charakteristischen Gebäudeeigenschaften auf der Ebene einer beliebigen Anzahl von zufällig gewählten, den Stadtraum vollständig und überschneidungsfrei abbildenden Teilräumen³² aggregiert. Diese Bebauungsstruktur wird als Signatur des Teilraumes bezeichnet. Alle Teilräume werden anschließend mit einem „k-means“-Algorithmus klassifiziert, dessen Klassenanzahl nach einer Parametervariation für den untersuchten Datensatz auf neun festgelegt wurde. Die eingeteilten Signaturklassen können als Siedlungsstrukturen verstanden werden, weisen aber im Vergleich zu den in Kapitel 2.4 beschriebenen Typisierungen eine größere Heterogenität auf. Im abschließenden Verfahrensschritt erfolgt der Einsatz eines Algorithmus aus der Gruppe der k-medoids-Verfahren in Kombination mit einer Optimierungsfunktion. Die zufälligen Teilräume werden mit Teilräumen, die eine ähnliche Signatur besitzen und sich in räumlicher Nähe befinden, vereinigt. Die daraus entstandenen Gruppen werden anschließend den ähnlichsten Signaturgruppen, d. h. Siedlungsstrukturen zugeordnet. Cao et al. liefern somit einen Verfahrensvorschlag zur raumbezogenen Clusteranalyse, unter Auflösung der sonst vorliegenden Beschränkungen im Falle einer Gruppierung mit Gebäudekoordinaten und anderen Gebäudeeigenschaften als Eingangsdaten.

Diese Auswahl an Studien zur Anwendung von Clusteranalyseverfahren sowie Algorithmen des überwachten ML zeigt den möglichen Mehrwert dieser Methoden im Hinblick auf Klimaschutzmaßnahmen im Gebäudebereich und liefert hilfreiche Ansätze, die bei der Entwicklung einer geoinformationsdatenbasierten Methodik zur Klassifizierung von Typquartieren anhand baulicher Strukturen in Kapitel 5 berücksichtigt werden.

³² Eine detaillierte Beschreibung dieser räumlichen Bezugsebene findet sich in Cao et al. (2013) nicht. Ebenfalls bleibt der Mechanismus zur Bildung dieser Teilgebiete in der Studie unbeleuchtet.

4 GEOINFORMATIONSDATEN IN DEUTSCHLAND

4.1 Grundlegende Erläuterungen zu Geoinformationssystemen und –daten

Die Motivation für die in dieser Arbeit formulierte Fragestellung leitet sich aus dem Wissensdefizit über das reale Vorhandensein von typischen Siedlungsstrukturen und Quartieren ab. Gleichzeitig besteht ein hohes Interesse seitens der Energieforschung am Quartier als Handlungsebene für Klimaschutzmaßnahmen im Gebäudesektor (siehe auch Kapitel 1). Während in den früheren Arbeiten zu Siedlungs- und Quartierstypen, welche in Kapitel 2 vorgestellt wurden, keine Geoinformationsdaten für eine flächendeckende Analyse der städtischen Bebauungsstrukturen in Deutschland zur Verfügung standen bzw. die flächendeckende Auswertung von geotopografischen Karten schlicht wegen der damals unzureichenden Rechnerleistung nicht möglich war, kann im Rahmen dieser Arbeit auf eine strukturierte Geodatenbasis für Deutschland zurückgegriffen werden. Dennoch existieren bisher keine Geoinformationsdaten (GIS-Daten), welche die Raumebene des Quartiers über einen themenspezifischen und lokalen Untersuchungsrahmen hinaus flächendeckend abbilden würden. Bevor in den nachfolgenden Abschnitten detailliert auf die in Deutschland verfügbaren GIS-Daten eingegangen wird, gilt es, einige grundlegende Begriffe und Prinzipien im Kontext der Geodatenverarbeitung zu erläutern und einen Überblick über die Entwicklung der Geodateninfrastruktur in Deutschland in den vergangenen Jahrzehnten zu geben. Ein Verständnis von wesentlichen (Daten-)Konzepten der Geoinformatik und deren Anwendung sind zwar für die Datenverarbeitung zur Entwicklung der Typisierungsmethodik in Kapitel 5 und Kapitel 6 erforderlich, jedoch kein Forschungsgegenstand dieser Arbeit und sollen deshalb an dieser Stelle nicht detailliert diskutiert werden.³³

Als Geoinformationen werden Objekte der realen Welt, die sog. Geoobjekte, verstanden, wobei als wesentliche Elemente ein Raumbezug und eine eindeutige Beschreibbarkeit vorliegen müssen. Diese Geoobjekte werden unter Zuhilfenahme eines konzeptionellen Modells, d. h. auf Basis festgelegter Vorschriften zur Datenstruktur, digital abgebildet. Dabei können Geoobjekte definierten Objektklassen zugeordnet werden und eine quantitative oder qualitative Beschreibung von Eigenschaften vorliegen. Die Merkmale eines Geoobjekts können dabei unterteilt werden in geometrische, topologische, thematische und temporale Eigenschaften sowie in ergänzende Metadaten (vgl. Bill 2018, S. 790). Die Verarbeitung der Geoinformationsdaten erfolgt in einem sog. Geoinformationssystem (GIS). Das GIS besteht aus den GIS-Daten sowie aus Hard- und Software, die für deren Verarbeitung

³³ Für weitere Informationen zu den Konzepten der Geoinformationsdatenverarbeitung siehe z. B. Bill (2010), Ehlers und Schiewe (2012), Sester (2019), Coors (2005), Coors et al. (2016) oder Kutzner (2016).

benötigt werden, und ist somit ein computergestütztes Hilfsmittel, mit dem ein Anwender raumbezogene Problemstellungen in unterschiedlichsten Themenfeldern bearbeiten, neue Informationen aus den Geoinformationen ableiten und andere, nicht raumbezogene Informationen mit den GIS-Daten verknüpfen kann (vgl. Bill 2018, S. 788–789; Kutzner 2016, S. 5–8).

Da das GIS grundlegend als ein lokales System für die Verarbeitung lokal gewonnener Daten betrachtet werden kann, ist für die Generierung eines bedeutenden Mehrwertes von GIS-Daten für politische, gesellschaftliche und wissenschaftliche Zwecke ein übergeordneter Rahmen zu definieren, welcher die Erhebung, Verarbeitung, Speicherung und Nutzung von GIS-Daten regelt. Dieser Rahmen wird durch internationale und nationale Standards zur Struktur der Geoobjekte, Objektinformationen und Beziehungen zu anderen Geoobjekten sowie zu Softwareschnittstellen gegeben und ermöglicht damit neben der „syntaktischen Interoperabilität (etwa Austauschformate oder Diensteschnittstellen) aber auch [...] (eine) semantische Interoperabilität (gemeinsame Vokabulare oder semantische Referenzsysteme) für [...] Informationsinfrastrukturen“ (Bernard et al. 2019, S. 92). Dies ermöglicht den Aufbau einer Geodateninfrastruktur (GDI) als Werkzeug und Organisationsstruktur für *„verteilte Geodaten und -dienste, Netzwerke, Standards und Richtlinien zum verantwortungsvollen Austausch und Umgang mit den zur Verfügung stehenden Geoinformationen sowie [mit den] institutionellen, organisatorischen, technologischen und wirtschaftlichen Ressourcen zu Entwicklung, Betrieb und Pflege der Geodateninfrastruktur“* (Bernard et al. 2019, S. 92–93; Hervorh. im Original). Ein weiteres Merkmal einer GDI ist zudem die direkte Nutzbarkeit der in der GDI bereitgestellten Ressourcen, d. h. der Zugriff und die Verwendung von Geodaten und -diensten ohne weitere technische oder organisatorische Aufwände zu eigenen Datenverarbeitungszwecken (vgl. Bernard et al. 2019, S. 93)³⁴.

Die Etablierung einer nationalen „Geodateninfrastruktur Deutschland“ (GDI-DE) wurde bereits durch den im Jahre 1998 gegründeten Interministeriellen Ausschuss für Geoinformationswesen vorbereitet und durch einen Beschluss des Kanzleramtschefs und der Chefs der Staats- und Senatskanzleien der Länder im Jahre 2003 in die Wege geleitet (vgl. BMI 2005, S. 1; GDI-DE 2015, S. 5). Durch die von Bund, Ländern und Kommunen getragene GDI-DE werden die vorhandenen Geoinformationen vernetzt, veröffentlicht und nutzbar

³⁴ Auf eine detaillierte Auseinandersetzung mit den Anforderungen an eine GDI und deren erfolgreichen Aufbau sowie Betrieb soll an dieser Stelle verzichtet werden, da sie für die Beantwortung der Forschungsfrage dieser Arbeit nur einen geringen Mehrwert liefert. Für weiterführende Informationen soll jedoch auf entsprechende Fachliteratur wie Bernard (2005), Bill (2010), Linke und Hickel (2015) oder Ehlers und Schiewe (2012) verwiesen werden.

sowie über standardisierte Schnittstellen verknüpfbar gemacht (vgl. GDI-DE 2015, S. 5). Zu Beginn musste dabei eine Vielzahl von Herausforderungen gelöst werden, welche sich u.a. aus der föderalen Verwaltungsstruktur in Deutschland ergaben. So wurde für die standardisierte GIS-Datenerfassung und -verarbeitung in den katasterführenden bzw. Vermessungsverwaltungen der Länder ein einheitliches Datenreferenzmodell durch die Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik (AdV) erarbeitet. Das sog. AAA-Referenzmodell³⁵ erlaubte eine Angleichung der Datenstrukturen. Weitere zu verbessernde Aspekte waren zudem eine bis dahin wenig vereinheitlichte Beschreibung der Metadaten (also z. B. der Qualität und Aktualität von GIS-Daten) sowie eine zwischen den einzelnen Bundesländern stark unterschiedliche Entgelt- und Nutzungspolitik. (Vgl. BMI 2005, S. 3; AdV 2008, S. 8–9)

Parallel zur Erarbeitung der GDI-DE und darin integrierter Geodatendienste wurde auf Ebene der EU eine „Richtlinie zur Schaffung einer Geodateninfrastruktur in der Europäischen Gemeinschaft“ (INSPIRE) erarbeitet und 2007 verabschiedet. INSPIRE, wie auch der GDI-DE, liegen die Ziele zugrunde, Geodaten für eine Vielzahl von politischen und gesellschaftlichen Wirkungs- und Entscheidungsbereichen zugänglich zu machen. Hierdurch soll eine transparente und interoperable Daten-, Dienste- und Verwaltungsstruktur ermöglicht werden, die neben der allgemeinen Zugänglichkeit eine einfache Nutzbarkeit der in der EU bzw. in Deutschland verfügbaren GIS-Daten erlaubt. (Vgl. GDI-DE 2015; INSPIRE, vom 14.03.2007)

In Deutschland wurde die INSPIRE-Richtlinie durch das „Gesetz über den Zugang zu digitalen Geodaten“ (GeoZG) in der nationalen Gesetzgebung verankert. Das GeoZG schafft den rechtlichen Rahmen für „den Zugang zu Geodaten, Geodatendiensten und Metadaten von geodatenhaltenden Stellen sowie [...] für] die Nutzung dieser Daten und Dienste, insbesondere für Maßnahmen, die Auswirkungen auf die Umwelt haben können“ (GeoZG, vom 25.02.2021, § 1). Darüber hinaus wurde die GDI-DE stetig weiter ausgebaut und mit der sog. „GeoInfoDok“ eine „Dokumentation zur Modellierung der Geoinformationen des amtlichen Vermessungswesens“ geschaffen, welche die kontinuierlichen technischen Neuerungen in der Geodatenverarbeitung, wie die verstärkte Modellierung von dreidimensionalen GIS-Daten, ebenfalls abdeckt (vgl. AdV 2008, S. 4–7). Auch durch das Inkrafttreten des „Gesetzes über die geodätischen Referenzsysteme, -netze und geotopographischen Referenzdaten des Bundes“ (BGeoRG) sowie mit der Formulierung einer „Nationalen

³⁵ Die Abkürzung „AAA“ bezeichnet dabei die Begriffe AFIS (Amtliches Festpunkt-Informationssystem), ALKIS (Amtliches Liegenschaftskatasterinformationssystem) und ATKIS (Amtliches Topographisch-Kartographisches Informationssystem), welche den grundlegenden Rahmen zur räumlichen Vermessung und modellhaften Abbildung der Erdoberfläche in Deutschland bilden. (Vgl. BMI 2005, S. 3)

Geoinformations-Strategie“ (NGIS) wurden die Aufgaben der Akteure, denen die Verantwortung für den Betrieb der GDI-DE übertragen worden sind, klarer formuliert und Grundprinzipien für die Sicherstellung des gesamtgesellschaftlichen Mehrwertes der GDI-DE verankert. Die „Open Data und Public Sector Information-Richtlinie“ der EU³⁶ sowie die „Open-Data“-Strategie der deutschen Bundesregierung tragen dazu bei, dass sich die weitreichenden Nutzungspotenziale einfach und frei zugänglicher GIS-Daten in Politik und Gesellschaft entfalten können. (Vgl. GDI-DE 2015; Open Data- und PSI-Richtlinie; BMI 2021b, 2021a, S. 22–38)

Die für Deutschland als Teil der GDI-DE verfügbare Geodatenbasis³⁷ sowie die im Rahmen der vorliegenden Arbeit für die Typisierung, Quantifizierung und statistische Beschreibung von Quartieren in der gebauten Umwelt verfügbaren Daten werden im Nachfolgenden beschrieben.

4.2 Datenquellen und –formate von Geoinformationsdaten

4.2.1 Amtliches Liegenschaftskatasterinformationssystem

Das Amtliche Liegenschaftskatasterinformationssystem (ALKIS) ist die Fortführung der Automatisierten Liegenschaftskarte und des Automatisierten Liegenschaftsbuchs der katasterführenden Behörden der Länder. Es dient der vollständigen und flächenhaften Abbildung aller Grund- und Flurstücke und ermöglicht somit die Erfassung beschreibender Daten zu Grundstücken und Gebäuden. Damit ist es das amtliche Verzeichnis aller Grundstücke im Sinne der Grundbuchordnung. Die Daten des ALKIS werden in der Datenmodellierungssprache Unified Modeling Language (UML) geführt, welche eine „vollständige und unzweifelhaft interpretierbare, formale Beschreibung von Inhalt und Struktur von Datenbeständen“ (AdV 2008, S. 12) erlaubt. Die Datenausgabe erfolgt textbasiert in der Extensible Markup Language (XML), unter Berücksichtigung der Kodierungsregeln der Geography Markup Language (GML) und den weiterführenden Spezifikationen der sog. „Normbasierten Austauschchnittstelle“ (NAS). Die mittels NAS zur Verfügung stehenden ALKIS-Daten können zur vereinfachten Anwendung auch in andere Datenformate übertragen werden (z. B. in das sog. Shape-Datenformat des Softwareanbieters ESRI). (Vgl. AdV 2008; Senatsverwaltung für Stadtentwicklung, Bauen und Wohnen Berlin; AdV 2019)

³⁶ Diese wurde durch eine Änderung des E-Government-Gesetzes und durch die Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors im nationalen Recht umgesetzt.

³⁷ Es wird lediglich eine Auswahl der Geobasisdaten vorgestellt, welche im Kontext der wissenschaftlichen Fragestellung dieser Arbeit zielführend erscheint.

Die Inhalte des ALKIS sind nach den Vorgaben des AAA-Referenzmodells, welche in der GeoInfoDok dokumentiert sind, strukturiert und werden durch den sog. ALKIS-Objektdatenkatalog beschrieben (siehe AdV 2018b). Das ALKIS enthält die in Tabelle Anhang C-1 dargestellten Objektbereiche und Objektartengruppen, welche wiederum in verschiedene Objektarten und Merkmale, hier Attribute genannt, unterteilt sind. Für jedes Objekt im ALKIS stehen geometrische, topologische, thematische und temporale Attributwerte sowie ergänzende Metadaten zur Verfügung. Ein Beispiel dieser Datenstruktur ist für den Objektbereich „Gebäude“ in Tabelle Anhang C-1 dargestellt. Das ALKIS als zentrale Datenstruktur für Grundstücks- und Gebäudeinformationen ist in Deutschland auf Ebene der Landeskatasterbehörden verfügbar, jedoch nicht flächendeckend für den Abruf von nicht personenbezogenen Daten kostenfrei. Es bildet allerdings die Basis für die Erstellung von deutschlandweiten Datensätzen durch das Bundesamt für Kartographie und Geodäsie (BKG) und für eine Verschneidung mit den Daten des Amtlichen Topographisch-Kartographischen Informationssystems (ATKIS). In den nachfolgenden Abschnitten werden deshalb diejenigen GIS-Daten beschrieben, welche flächendeckend für das Bundesgebiet vorliegen.

4.2.2 Digitales Basis-Landschaftsmodell

Einen Datensatz, welcher die Erdoberfläche in Deutschland vollständig beschreibt, stellt das Digitale Basis-Landschaftsmodell (Basis-DLM) dar. Es bildet die topografischen Objekte der Landschaft im Vektorformat ab, wobei die Geoobjekte einer bestimmten Objektart zugeordnet werden. Die Geoobjekte werden im Datensatz durch ihre räumliche Lage, ihren geometrischen Typ, beschreibende Attribute und Beziehungen zu anderen Objekten (Relationen) sowie durch einen deutschlandweit eindeutigen Objektidentifikator definiert. Der Umfang der beinhalteten Attribute orientiert sich an der Topographischen Karte im Maßstab 1:25.000, jedoch hat die Abbildung im Basis-DLM eine höhere Lagegenauigkeit der punkt- und linienförmigen Objekte. (Vgl. BKG 2022, S. 4)

Die Objektebenen des Basis-DLM sind dabei in einem dem ALKIS ähnlichen Muster strukturiert, umfassen jedoch überwiegend die Geoobjekte einer höheren räumlichen Ebene. Eine beispielhafte Darstellung dieser Struktur zeigt Tabelle 4-1 (S. 66). Es wird deutlich, dass die Datenstruktur des Basis-DLM eine vollständige Abbildung der gebäudebezogenen Informationen aus dem ALKIS (Objektartengruppe „Gebäude“ mit ID 31001) erlaubt. Auf die im Datensatz des BKG verfügbaren Informationen wird in Kapitel 5 detailliert eingegangen.

4 Geoinformationsdaten in Deutschland

Tabelle 4-1: Gruppierung der Geoobjekte im Basis-DLM nach Themen, Ebenen und Objektarten [eigene Darstellung basierend auf AdV (2018a)]

Thema	Ebene (Inhalt)	Objektarten (ID)
Siedlung		
	Ortslage (Sie01)	
		52001 AX_Ortslage
	Baulich geprägte Flächen (Sie02)	
		41001 AX_Wohnbauflaeche
		41002 AX_IndustrieUndGewerbeflaeche
		41003 AX_Halde
		41004 AX_Bergbaubetrieb
		41005 AX_TagebauGrubeSteinbruch
		41006 AX_FlaecheGemischterNutzung
		...
	Bauwerke und sonstige Einrichtungen (Sie03)	
		...
	Besondere Anlagen auf Siedlungsflächen (Sie04)	
		...
	Gebäude (Sie05)	
		31001 AX_Gebaeude
		31002 AX_Bauteil
	51001 AX_Turm	
	Inkl. Attribute aus 12002 AX_LagebezeichnungMitHausnummer 12003 AX_LagebezeichnungMitPseudonummer	
Verkehr		
...		

4.2.3 Verwaltungsgebiete

Der Datensatz der Verwaltungsgliederung im Maßstab 1:25.000 (VG25) „umfasst die administrativen Verwaltungseinheiten der Bundesrepublik Deutschland von der Staats- bis zur Gemeindeebene“ (BKG 2021b, S. 5). Mit Ausnahme der Regierungsbezirke werden die Verwaltungsebenen flächendeckend für Deutschland aufgebaut und sind in ihrer Hierarchie über den sog. „Amtlichen Regionalschlüssel“ strukturiert. Die jeweilige Art der Verwaltungseinheit kann durch einen im Datensatz verfügbaren Bezeichnungsidentifikator (IBZ) nachvollzogen werden.³⁸ (Vgl. BKG 2021b, 2020)

³⁸ Bei Betrachtung der untersten durch den Datensatz VG25 abgebildeten administrativen Ebene der Gemeinde beschreibt der IBZ die Unterscheidung in Städte und Gemeinde unter Beachtung einer ggf. vorliegenden Kreisfreiheit oder Gemeinschaftsangehörigkeit der Kommune sowie ggf. in gemeindefreies Gebiet (IBZ 60–65).

Zur weiteren Untergliederung der Verwaltungsstrukturen in deutschen Großstädten ab 250.000 Einwohnern ist zudem der Datensatz der „Kommunalen Teilgebiete“ verfügbar, welcher ebenfalls im Maßstab 1:25.000 (KT25) die kommunalen Teilgebiete erster Ordnung abbildet (vgl. BKG 2018). Die kommunalen Teilgebiete entsprechen der stadtspesifischen Verwaltungsgliederung und tragen in den jeweiligen Städten unterschiedliche Bezeichnungen. Sie werden im Rahmen der vorliegenden Arbeit vereinfacht als Bezirke oder Stadtbezirke beschrieben. Dabei gilt es zu berücksichtigen, dass mit diesem Begriff nicht die Ebene der Regierungsbezirke, welche in einigen Bundesländern existiert, bezeichnet werden soll.

4.2.4 Amtliche Hauskoordinaten und Amtliche Hausumringe

Ebenfalls abgeleitet aus den Vermessungs- bzw. Katasterinformationen der Länder stehen die beiden Datensätze Amtliche Hauskoordinaten (HK-DE) und Amtliche Hausumringe (HU-DE) zur Verfügung. Die HK-DE stellen die Gebäude in Deutschland als Punktelemente dar, während die HU-DE die Gebäudegrundrisse als Flächenpolygone abbilden. Die Objekte im Datensatz HU-DE sind dabei nicht nur Objekte der Objektartengruppe „Gebäude“ (ID 31001), sondern auch Geoobjekte der Objektartengruppe „Bauwerke und Einrichtungen in Siedlungsflächen“ (ID 5100X, vgl. Tabelle 4-1 und Tabelle Anhang C-1). Weiterführende Attribute, d. h. Eigenschaften der Gebäude, wie für das ALKIS beschrieben, sind in diesem Datensatz nicht enthalten. Die geometrischen Abbildungen der Gebäude können jedoch über einen eindeutigen Objektidentifikator oder über räumliche Abfragen mit Objekten bzw. Attributen anderer Datenquellen vereinigt werden. (Vgl. BKG 2021c, 2021d; Meinel und Burckhardt 2013)³⁹

4.2.5 Georeferenzierte Adressdaten

Die Georeferenzierten Adressdaten für Deutschland (GA) sind, wie auch die HK-DE, Punktkoordinaten, welche mit dem Merkmal „Hausadresse“ und weiteren Informationen zur Verwaltungszugehörigkeit einen Datensatz bilden. Im Unterschied zu den HK-DE beruhen die GA jedoch auf einer weiter gefassten Datenbasis. Neben den HK-DE liegt den GA zudem ein Adressdatensatz der Deutschen Post zugrunde. Dieser wird zur Überprüfung, insbesondere aber auch zur Ergänzung von (noch) in den HK-DE fehlenden Adressen verwendet. Der Datensatz enthält ca. 22,8 Millionen Adressdaten von Deutschland. (Vgl. BKG 2021a)

³⁹ In Meinel und Burckhardt (2013) werden die HK-DE fälschlicherweise mit den Georeferenzierten Adressdaten gleichgesetzt. Dieser Datensatz wird jedoch noch aus weiteren Datenquellen gespeist und ist somit als eigenständig anzusehen.

4.2.6 3D-Gebäudemodelle

Neben der zweidimensionalen, d. h. punkt-, linien- oder flächenhaften Abbildung der Gebäude in den oben beschriebenen Geodatenquellen sind in den vergangenen Jahren durch die katasterführenden Stellen der Länder auf Basis der Vorgaben des AAA-Referenzmodells aus den verfügbaren gebäudebezogenen Geobasisdaten dreidimensionale Gebäudeabbildungen erstellt worden. Dies erfolgte durch eine Kombination der ALKIS-Daten und der Daten des Digitalen Geländemodells oder anderer Datenquellen, wie digitalen Orthophotos. (Vgl. AdV 2008)

In Deutschland liegen die 3D-Gebäudemodelle im Format der City Geographic Markup Language (CityGML-Format) vor, generell sind aber auch andere Datenformate für 3D-Gebäudemodelle verfügbar (z. B. CityJSON oder Kollada). Das CityGML-Format ist eine textbasierte, semantische Beschreibung der 3D-Gebäudedaten. Es folgt den Festlegungen der Extensible Markup Language (XML), als Beschreibungsformat für die Abbildung von Daten und deren Beziehungen in Textform. Die inhaltliche Strukturierung für die 3D-Gebäudemodelle wird durch ein konzeptionelles Datenmodell in der Unified Markup Language (UML) und den darin definierten Objektklassen und Schemata vorgegeben (vgl. Kutzner 2016; Gröger et al. 2012; AdV 2008). Für die Anwendung im Rahmen der GDI-DE wurde durch die AdV ein AdV-CityGML-Profil definiert. Die geometrische und inhaltliche Beschreibung der 3D-Gebäudemodelle besteht aus den bekannten gebäudebezogenen Informationen des ALKIS, die Gebäude sind hinsichtlich ihrer Geometrie jedoch als sog. „Solids“, d. h. Volumenobjekte modelliert, welche sich wiederum aus verschiedenen Flächenelementen für jeden Teil der Gebäudeoberfläche zusammensetzen. Die Detaillierung der 3D-Gebäudemodelle (engl. level of detail; LoD) kann sich dabei jedoch unterscheiden. Definiert ist der LoD1 als Abbildung der Gebäude in Form von Klötzchenmodellen ohne Berücksichtigung der realen Dachform. Im LoD2 werden die Dachformen, bis auf wenige Sonderformen, detailliert abgebildet. In den LoDs mit höherer Indexierung werden die Gebäudemodelle weiter detailliert. So weisen die Oberflächen von LoD3-Modellen auch eine explizite Modellierung von Fenstern und Türen auf und im LoD4 sind Innenausstattungen von Gebäuden abbildbar (vgl. Abbildung 4-1). Der Datensatz des BKG enthält die in Tabelle 4-2 dargestellten Inhalte. (Vgl. Coors 2005; Coors et al. 2016; AdV 2008, 2021a; Gruber und Recklinghausen 2020; LDBV Bayern 2021)

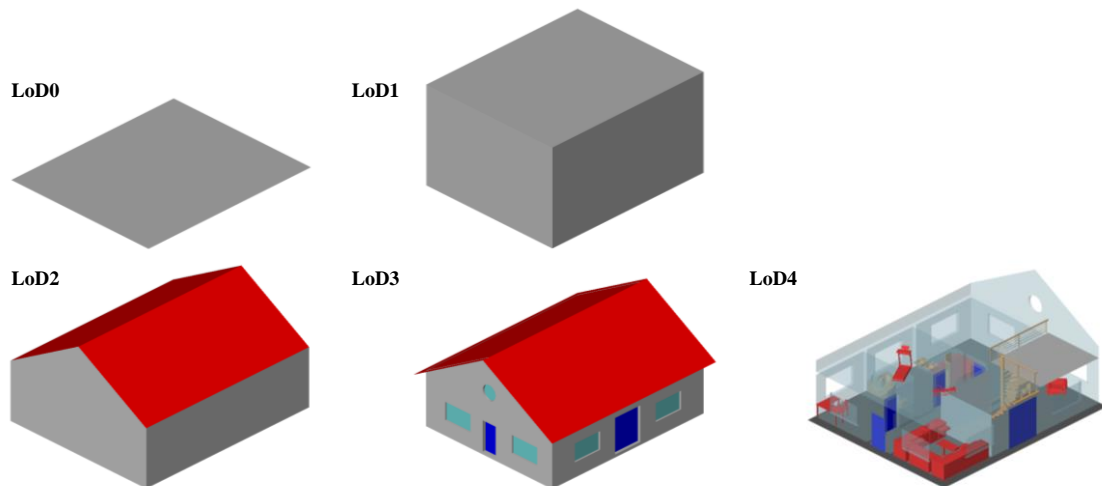


Abbildung 4-1: Repräsentation eines Beispielgebäudes im LoD0 bis LoD4 (einige Objektschichten wurden für die Darstellung des LoD4 ausgeblendet) [eigene Darstellung; Daten: Institut für Automation und Angewandte Informatik, Karlsruher Institut für Technologie]

Tabelle 4-2: Art der Abbildung von Gebäuden im 3D-Gebäudemodell LoD2 sowie Gebäudeattribute [eigene Darstellung basierend auf LDBV Bayern (2021)]

Geometrische Abbildung der Gebäude im Modell
<ul style="list-style-type: none"> Die Gebäudedarstellung erfolgt mit standardisierten Dachformen.
<ul style="list-style-type: none"> Die Geometrie des Körpers wird durch Volumenelemente (Solids) und deren Flächenaggregate (Multisurface) beschrieben.
<ul style="list-style-type: none"> Gemeinsam genutzte Geometrien werden redundant geführt.
<ul style="list-style-type: none"> Die Höhengenaugigkeit beträgt größtenteils 1 m; größere Abweichungen sind in Einzelfällen bei komplexen Dachformen möglich.
Attribute der Gebäudemodelle
<ul style="list-style-type: none"> Objektidentifikator
<ul style="list-style-type: none"> Gebäudefunktion
<ul style="list-style-type: none"> Qualitätsangaben (Metadaten)
<ul style="list-style-type: none"> Höhe des Gebäudes als Differenz in Metern zwischen dem höchsten Bezugspunkt und dem tiefsten Bezugspunkt des Gebäudes
<ul style="list-style-type: none"> Amtlicher Gemeindeschlüssel
<ul style="list-style-type: none"> Generalisierte Dachform
<ul style="list-style-type: none"> Name (falls geführt)

Durch die föderale Struktur des Vermessungswesens in Deutschland und einige Freiheitsgrade bei der Erstellung der 3D-Gebäudemodelle aus den o. g. Basisdaten können zwischen den Modellen der Bundesländer systematische Unterschiede bestehen. Diese betreffen insbesondere die Datengrundlagen für die Erfassung von Gebäudehöhen und Dachformen. Aber auch die Gebäudehöhe eines LoD1-Modells kann unterschiedlichen Definitionen

folgen.⁴⁰ Diese Besonderheiten sind über die Metadaten des 3D-Gebäudedatensatzes nachvollziehbar und in Abhängigkeit der Anwendungszwecke zu berücksichtigen. Die Modellierung der Gebäude als Gebäudeteile unterscheidet sich ebenfalls von Bundesland zu Bundesland, was je nach Anwendungszweck zu berücksichtigen ist (siehe auch Tabelle 5-2 in Kapitel 5.2). (Vgl. Schwarz et al. 2021)

4.3 Einsatzbereiche geoinformationstechnischer Daten

Die Einsatzbereiche geoinformationstechnischer Daten auf den verschiedenen Verwaltungsebenen des föderalen Systems in Deutschland sowie in anderen gesellschaftlichen Bereichen sind vielfältig. So können als zentrale Arbeitsfelder Planungsaufgaben der kommunalen Verwaltungen sowie Landesverwaltungen betrachtet werden, in deren Arbeitsprozessen schon seit Dekaden die analogen Plandaten durch digitalisierte Datenformate ergänzt und heute überwiegend ersetzt wurden. Auch die Themenfelder der raumbezogenen Forschung sind vielfältig. So finden sich heute in allen klassischen Teilgebieten der Forschung auch Forschungsfelder mit Raumbezug. Einen sehr guten Überblick über die raumbezogene Nachhaltigkeitsforschung, welche selbst als ein Querschnittsthema zu verstehen ist, gibt z. B. die Schriftenreihe „Flächennutzungsmonitoring“ des Leibniz-Institut für ökologische Raumentwicklung (z. B. Meinel et al. 2021; Meinel et al. 2013). Die Schaffung von flächendeckenden GIS-Datensätzen im Rahmen der GDI-DE erlaubt jedoch weit über die öffentlichen Aufgaben und wissenschaftlichen Fragestellungen hinausgehende Handlungs- und Geschäftsfelder. Ein Beispiel für ein solches GIS-Daten getriebenes Geschäftsfeld ist das Geomarketing, welches für die Methoden des klassischen Marketings, wie der Kundensegmentierung, eine Verknüpfung raumbezogener Attribute und eine Anreicherung mit Gebäudeinformationen aus den Geobasisdaten beinhaltet. Eine mikrogeographische Marktsegmentierung lässt so z. B. das Abschätzen von Modernisierungswahrscheinlichkeiten in Abhängigkeit von der Gebäudeart und -lage sowie den sozioökonomischen und soziodemografischen Merkmalen der Eigentümer zu (vgl. Kroth und Wawrzyniak 2022; Herter und Mühlbauer 2018). Aber auch im Kontext der gebäude- und quartiersbezogenen Energieforschung nimmt, wie bereits in Kapitel 2 und Kapitel 3 gezeigt, die Nutzung von GIS-Daten zu. Einige Beispiele, welche im Hinblick auf die in dieser Arbeit zu beantwortenden Fragestellungen nützliche Erkenntnisse liefern, werden nachfolgend vorgestellt und diskutiert.

⁴⁰ Die Höhe eines LoD1-Gebäudemodells kann z. B. dem Mittelwert der Höhe aller Punkte einer das Dach beschreibenden Punktwolke entsprechen, aber auch die Höhe der Firstlinie oder der Traufe sein (vgl. Schwarz et al. (2021, S. 200–201)).

Ein in den vergangenen Dekaden stark in den Fokus der gebäudebezogenen Energieforschung gerücktes Thema ist die automatisierte energetische Bewertung von Gebäuden sowie Siedlungsbereichen unterschiedlicher räumlicher Ausdehnung. Beispiele für Softwarelösungen wurden bereits in Kapitel 2.5 erläutert. Diese Werkzeuge und andere Forschungsarbeiten im selben Kontext nutzen zur Modellierung der Energiebedarfe auf Gebäude- und Quartiersebene GIS-Daten in unterschiedlicher Art und räumlicher Auflösung sowie mit unterschiedlichen ergänzenden Metadaten. Für eine Modellierung von Gebäudeenergiebedarfen leiten z. B. Nouvel et al. (2015a), Weiler et al. (2018), Zirak et al. (2020) und Dochev et al. (2020) aus 3D-CityGML-Daten für das Baualter und die Gebäudeart typische bauphysikalische Parameter ab und speisen damit Softwaremodule zur Simulation von Gebäudeenergiebedarfen und Wärmenetzen. Ähnliche Modellierungsansätze, allerdings in einem kanadischen bzw. US-amerikanischen Kontext, verfolgen Hosseini Haghighi et al. (2021), Chen et al. (2017), Rezaei et al. (2021), Cerezo Davila et al. (2016) und Abdolhosseini Qomi et al. (2016). Diese Arbeiten zeigen, dass eine energetische Bewertung von Gebäuden ohne die detaillierte Verfügbarkeit von bauphysikalisch oder energietechnisch relevanten Gebäudeinformationen eine Vielzahl an Modellannahmen zu den Gebäuden und deren energetischen Zustand erfordert. In der Studie von Cerezo Davila et al. (2016) führt der Einbezug ergänzender, qualitätsgesicherter Datenquellen mit gebäudebeschreibenden Merkmalen folglich zu einer insgesamt höheren Güte der Modellergebnisse. Abdolhosseini Qomi et al. (2016) nutzen ein probabilistisches Modell zur Bestimmung des Wärmeverbrauchs, um die Modellunsicherheiten zu reduzieren und berücksichtigen die Unsicherheiten für einen Großteil der insgesamt 74 Modelleingangsparameter. Mutani und Todeschi (2021) nutzen weiterhin eine Gebäudeenergieausweisdatenbank zur semantischen Anreicherung der verfügbaren GIS-Daten. Basierend auf dieser Datengrundlage erarbeiten die Autoren einen Wärmeatlas für die Stadt Turin, welcher es z. B. ermöglicht, die Stadtregionen zu identifizieren, die eine besonders niedrige energetische Modernisierungsrate im Gebäudebestand aufweisen und somit das Ziel einer aktiven Intervention werden sollten. Aus den Studien kann abgeleitet werden, dass die Verfügbarkeit von gesicherten Gebäudeinformationen – insbesondere aus der energetischen Gebäude- bzw. Modernisierungsplanung – einen deutlichen Mehrwert für die energetisch-geoinformationstechnische Analyse von Gebäuden bietet.

Über die eher kleinräumigen Analysen der o. g. Studien hinaus wurden an der Schnittstelle von Verwaltungseinrichtungen für Stadt- und Infrastrukturplanung und Versorgungsunternehmen – z. T. mit Unterstützung von wissenschaftlichen Instituten der gebäudebezogenen Energieforschung – sog. Wärmeatlanten oder Wärmekataster entwickelt. Das voranstehend bereits vorgestellte Prinzip der Nutzung und ggf. die Anreicherung von GIS-Daten und

deren Weiterverarbeitung zur Ermittlung von Gebäudeenergiebedarfen⁴¹ bleibt dabei weitestgehend gleich. Für die großräumige Analyse ganzer Städte und Regionen werden meist jedoch keine dynamischen Gebäudesimulationen durchgeführt, sondern eine Einordnung der Gebäude in der untersuchten Region in Gruppen einer bestehenden Gebäudetypologie vorgenommen und bereits verfügbare (z. B. nutzflächenspezifische) Energiekennwerte der Typgebäude auf die Gebäude des Untersuchungsgebietes übertragen. Als Beispiele sollen an dieser Stelle die Wärmekataster des Saarlands (Wern et al. 2017) und der Hansestadt Hamburg (BUKEA 2019; Dochev et al. 2018; Dochev et al. 2019; Dochev et al. 2020) sowie der Energieatlas des Landes Baden-Württemberg (LUBW 2022) und der (gebäudescharfe) Wärmeetlas für Deutschland des Konsortiums aus dem Institut für Energie- und Umweltforschung, GEF Ingenieur AG und geomer GmbH (Blömer 2021) angeführt werden. Die Datenverarbeitung, die Modellierung der Wärmebedarfe und damit auch der Nutzen der Wärmekataster muss jedoch kritisch hinterfragt werden, da eine Anwendung der Ergebnisse nur mit entsprechenden Einschränkungen sinnvoll erscheint. So übertragen die o. g. Modelle Daten zum Gebäudebaualter aus einem Datensatz des Statistischen Bundesamtes mit einer räumlichen Auflösung von 100 m x 100 m auf alle Gebäude, die sich in dieser Rasterzelle befinden. Auch liegen in den Gebäudedaten keine Details über den realen energetischen Zustand, also das aktuelle Modernisierungsniveau und aktuelle bauphysikalische Parameter vor. Weiterhin ist die Übertragung der Energiekennzahlen von Typologiegebäuden auf die Modellabbildungen der realen Gebäude ein weiterer Prozessschritt, welcher mit hohen Modellunsicherheiten einhergeht. Denn den Typologiegebäuden liegen bestimmte Modellannahmen zur Gebäudegeometrie, den materiellen und bauphysikalischen Eigenschaften sowie ggf. zu baulichen Besonderheiten zugrunde, welche sich i. d. R. nicht mit dem Ist-Zustand der realen Gebäude decken. Dennoch kann die Nutzung dieser Wärmekataster z. B. für eine raumbezogene Wärmeleitplanung hilfreich sein, weil hier i. A. keine gebäudescharfen Aussagen nötig sind und die Analysen eine räumliche Ausdehnung von Versorgungsgebieten bzw. Netzabschnitten umfassen. Die verfügbaren Geobasisdaten liefern also auch in diesem Kontext einen wichtigen Beitrag für die Erarbeitung von Klimaschutzmaßnahmen im Gebäudesektor.

Im Forschungsvorhaben EnOB:dataNWG wurde ein anderer Nutzen aus dem Vorliegen deutschlandweiter Geobasisdaten gezogen. Zur ersten repräsentativen Stichprobenerhebung über die Anzahl und Merkmale von Nichtwohngebäuden in Deutschland wurden die Geobasisdaten HU-DE, GA und 3D-Gebäudemodelle des LoD1 verwendet, um einen

⁴¹ In den verschiedenen Wärmekatastern werden als Ergebniskennwerte Energiebedarfe oder verbrauchskalibrierte Endenergiebedarfe angegeben. Auf die Differenzierung zwischen diesen Arten der Ergebnisdarstellung wird im weiteren Text mit dem Ziel einer besseren Lesbarkeit verzichtet.

Datensatz von Nichtwohngebäuden anhand geometrischer, topologischer und semantischer Informationen zu erstellen. Aus diesem Datensatz ließen sich anschließend, unter Berücksichtigung verschiedener stichprobentheoretischer Anforderungen, Erhebungsbezirke abgrenzen, welche eine repräsentative Stichprobenerhebung des Nichtwohngebäudebestandes erlaubten. Die Gebäudemerkmale der Stichprobenobjekte wurden im Vorhaben EnOB:dataNWG in verschiedenen Detaillierungsstufen erfasst und die Informationen zu den Gebäuden aus einer Detaillierungsstufe jeweils bei der Erhebung von Daten auf der nächsten Detailstufe auf ihre Richtigkeit hin überprüft. Die GIS-Daten dienten hier also nicht als Grundlage einer Modellierung, sondern als Grundgesamtheit der Gebäude in Deutschland, aus welcher eine Nichtwohngebäudestichprobe gezogen wurde. (Vgl. Hörner et al. 2021; Hartmann et al. 2020b; Hartmann et al. 2020a; Busch und Müller 2020; Cischinsky et al. 2021)

Die vorgestellten Arbeiten unterstreichen den Wert von GIS-Daten für die gebäude- und quartiersbezogene Energieforschung. Dennoch lassen sich aus diesen Studien auch Herausforderungen bei der Auswahl und Verarbeitung von GIS-Daten ableiten, die im nachfolgenden Kapitel zusammenfassend dargestellt werden.

4.4 Einschränkungen bei der Nutzung von Geoinformationsdaten im Kontext der energetischen Bewertung von Gebäuden und Quartieren

Die Voraussetzungen zur Erzielung belastbarer Ergebnisse im Rahmen der GIS-Daten-basierten, energetischen Bewertung von Gebäuden, Quartieren und größeren Siedlungsbereichen werden aus den im vorigen Kapitel vorgestellten Studien abgeleitet. Sie umfassen einige wenige Aspekte, welche die Prozessschritte GIS-Datenauswahl und -verarbeitung, energetische Modellierung sowie Ergebnisdarstellung betreffen. Einige der Formulierungen können dabei als Beschreibung eines für die gebäude- und quartiersbezogene Energieforschung idealen GIS-Datensatzes verstanden werden.

Eine möglichst detaillierte Abbildung von energierelevanten Gebäudeeigenschaften in den GIS-Daten erlaubt eine deutliche Verbesserung der Modellierung von Energiebedarfen auf Gebäude- und Quartiersebene. Zwar sind in den vergangenen Jahren aus geoinformatischer Sicht die Möglichkeiten für eine Anreicherung von GIS-Daten (insbesondere den 3D-Gebäudemodellen) z. B. mit der CityGML EnergyADE⁴² geschaffen worden (vgl. Aguiaro

⁴² Eine Application Domain Extension (ADE) ist eine ergänzende Schemabeschreibung, zu einer definierten semantischen Modellierung, hier CityGML. Die EnergyADE liefert somit eine Struktur, welche es erlaubt einen CityGML-Datensatz um energierelevante Gebäudeinformationen in einer einheitlichen und automatisiert verarbeitbaren Form zu erweitern.

et al. 2018; Nouvel et al. 2015b). Vorhanden sind solche detaillierten Gebäudeinformationen in der Breite jedoch nicht. Auch wenn diese Daten für einen (Groß-)Teil des Gebäudebestandes verfügbar gemacht werden könnten, wäre eine Nutzung erst vor dem Hintergrund der Gesetzgebung zum Schutz personenbezogener Daten auf europäischer und Bundesebene zu prüfen.

Die verschiedenen Geobasisdaten in Deutschland erlauben bisher keine einheitliche sowie eindeutige Abbildung und energetische Modellierung von Gebäuden. Dieser Zustand ist darin begründet, dass die Abbildung von Gebäudekörpern, Gebäudeadressen und anderen gebäudezugehörigen Bauteilen keine vollständige semantische Zuordnung zulassen und der Anwender unter individuell festzulegenden Modellannahmen ein Zuordnungs- bzw. Ausschlusschema definieren muss. So mussten die GIS-Daten im Vorhaben EnOB:dataNGW z. B. erst durch einen Bereinigungsalgorithmus vorverarbeitet werden, um untypische Grundrisspolygone auszuschließen. Zudem wurde eine Anreicherung der Daten der HU-DE mit den GA vorgenommen. Zwischen den Objekten der beiden Datensätze besteht jedoch nicht immer eine 1:1- oder 1:n-Zuordnung. Dies bedeutet, dass einzelne Adressdaten keinem Grundriss zugeordnet werden können und umgekehrt für viele Gebäudegrundrisse keine Adressdaten vorliegen, obwohl diese Grundrisspolygone als eigenständige Gebäude in den Datensätzen HU-DE oder 3D-Gebäudemodelle enthalten sind (vgl. Hartmann et al. 2020a; Hartmann et al. 2016; Hecht 2014, S. 173–176). Wegen nicht verfügbarer semantischer Informationen⁴³ zu diesen Grundrissen bleibt die Handhabung der Gebäudepolygone ohne Adresspunkt somit ebenfalls dem Anwender überlassen (vgl. auch Kapitel 6.2). Für eine Anwendung in der gebäude- und quartiersbezogenen Energieforschung sind deshalb zwei Prämissen zu formulieren: Erstens sollten alle vorliegenden Informationen über eine 1:1- bzw. 1:n-Beziehung den GIS-Gebäudeabbildungen zugeordnet werden können. Zweitens sollten Gebäudegrundrisse oder Umringe semantische Informationen besitzen, die eine Zuordnung bzw. Abgrenzung zu anderen Gebäuden sowie eine Einteilung in energiebilanzrelevante und -irrelevante Polygone⁴⁴ ermöglichen.

Ergänzend zu den Freiheitsgraden, welche sich bei der Modellierung von GIS-Daten ohne ein vollständiges Schema zur Zuordnung von Daten unterschiedlicher Quellen ergeben,

⁴³ Diese Informationen könnte z. B. die eindeutige Zuordnung eines Hinterhauses, welches als eigenständiges Gebäude im GIS-Datensatz abgebildet ist, zu einem Haupthaus sein.

⁴⁴ Für einige der Geoobjekte der HU-DE und 3D-Gebäudemodelle kann eine solche Einteilung über die Gebäudefunktion erfolgen. Eine Vielzahl der Polygone weist jedoch ein bilanzrelevantes Funktionsattribut auf, obwohl keine vordergründige Zugehörigkeit zum energiebilanzrelevanten Baukörper besteht. Beispiele sind Gebäudevordächer, Balkone und andere Anbauten, welche die Nutz- bzw. Energiebezugsfläche der Gebäude nicht erhöhen (siehe hierzu auch die Ausführungen in Kapitel 6.2).

entstehen durch fehlende energetische bzw. bauphysikalische Informationen Unsicherheiten. Eine energetische Modellierung sollte deshalb auf probabilistischen Prinzipien beruhen – insbesondere im Kontext der geodatenbasierten Modellierung, durch welche gebäudescharfe Ergebnisse visualisiert werden können. Energiebedarfe sind damit nicht als einzelner gebäudebezogener Kennwert zu errechnen, sondern als Wertebereiche oder eine Auswahl möglicher Bedarfswerte für unterschiedliche Konfigurationen für Bilanz- bzw. Simulationseingangsparameter⁴⁵.

Hinsichtlich der Ergebnisdarstellung ist als Anforderung an energetisch-geoinformationstechnische Analysen zu formulieren, dass eine Darstellung auf der größten, d. h. räumlich höchsten Ebene der Eingangsdaten erfolgen sollte, falls mit deterministischen Methoden gearbeitet wird. Andernfalls sollten die Unsicherheiten aufgrund der Übertragung von gebäuderelevanten Daten zwischen Objekten unterschiedlicher räumlicher Auflösung (z. B. von statistischen Daten der Ebene des Erhebungsbezirks auf Einzelgebäudegrundrisse) auch in grafischen Darstellungen kenntlich gemacht werden und eine Fehlerabschätzung innerhalb der Studie erfolgen.

Die vorgestellten Einschränkungen und Erkenntnisse aus der bisherigen Anwendung von GIS-Daten in der gebäude- und quartiersbezogenen Energieforschung werden im nachfolgenden Kapitel mit dem Ziel der Entwicklung einer geoinformationsdatenbasierten Methodik zur Klassifizierung von Typquartieren herangezogen. Dafür sind die vorliegenden Datenquellen zu benennen und deren Handhabung vor dem Hintergrund der beschriebenen Anforderungen an die Datenbasis, Datenverarbeitung und Ergebnisdarstellung zu erläutern.

⁴⁵ In den in Kapitel 4.3 vorgestellten Arbeiten wurde über den Mittelwert der Baujahre der Gebäude in einer Rasterzelle das Gebäudealter bestimmt und aus typologischen Kennwerten ein Satz an Bilanz- bzw. Simulationseingangsdaten abgeleitet. Eine Verbesserung der Ergebnisse kann z. B. durch die Berücksichtigung mehrerer Eingangsparametersätze für unterschiedliche energetische Gebäudeniveaus von „unsaniert“ bis „energetisch umfassend saniert“, erzielt werden.

5 ENTWICKLUNG EINER METHODIK ZUR GEODATENBASIERTEN QUARTIERSTYPISIERUNG

5.1 Einführende Erläuterungen

Die im Rahmen dieser Arbeit entwickelte Methode zur geodatenbasierten Quartierstypisierung leitet sich aus den in Kapitel 1 und Kapitel 2 skizzierten klimapolitischen und auf die Quartiers- und Energieforschung bezogenen Hintergründen ab. Diese können wie folgt vereinfacht zusammengefasst werden:

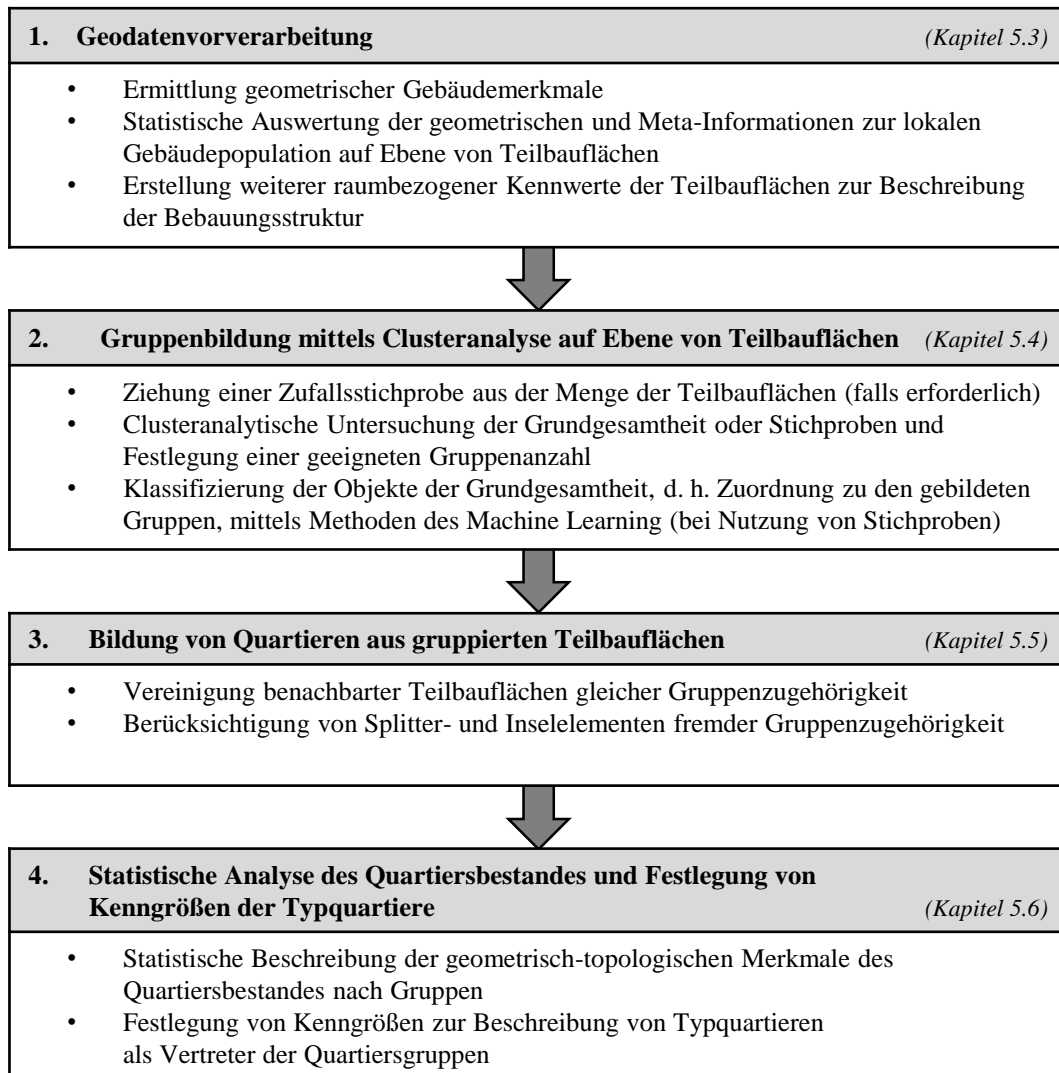
- Zur Erreichung der Klimaschutzziele im Gebäudebestand bedarf es immenser Anstrengungen bezüglich der energetischen Modernisierung von Gebäuden sowie einer weitestgehenden Umstellung der Art der Wärmeversorgung hin zum breiten Einsatz erneuerbarer Energieträger (vgl. Kapitel 1.1).
- Das Quartier, als dem Gebäude übergeordnete Raumebene, wird in Forschung und Praxis als ein Forschungsgegenstand bzw. Aktionsraum betrachtet, der es erlaubt, über die ohnehin notwendigen Maßnahmen auf Gebäudeebene hinaus sowohl eine Steigerung der Gesamteffizienz als auch eine beschleunigte Umsetzung von Modernisierungsmaßnahmen zu erreichen (vgl. Kapitel 2.2).
- Über einen spezifischen Kontext, d. h. über lokale Quartiersvorhaben, hinaus, existiert keine Möglichkeit zur Verallgemeinerung von Erkenntnissen aus Modellvorhaben bei gleichzeitiger quantitativer Abschätzung der Wirkungen von quartiersbezogenen Maßnahmen (vgl. Kapitel 2.6).

Als weitere Grundlage der Methodenentwicklung wurden zudem die mathematisch-informationstechnischen Zusammenhänge zur Clusteranalyse und der Klassifizierung in Kapitel 3 dargestellt sowie in Kapitel 4 ein Überblick über Geoinformationssysteme und der in Deutschland geführten Geodatenbasis gegeben. Für die Beantwortung der eingangs formulierten Forschungsfragen werden in den nachfolgenden Abschnitten demnach die verfügbare Datenbasis vorgestellt und eine Datenauswahl vorgenommen, bevor die mehrstufige Methode zur geodatenbasierten Typisierung von Quartieren anhand baulicher Strukturen erarbeitet wird. Diese beinhaltet, wie in Abbildung 5-1 (S. 80) skizziert, die Erfassung der Bebauungsstruktur auf einer kleinskaligen Raumebene, den sog. Teilbauflächen (Kapitel 5.3) und die Einteilung von Teilbauflächen in Gruppen ähnlicher Bebauungsstruktur (Kapitel 5.4). Wie später noch detaillierter erläutert wird, ist im Rahmen dieser Einteilung von Teilbauflächen in Gruppen zudem wegen informationstechnischer Beschränkungen eine differenzierte Datenverarbeitung für besonders umfangreiche Datensätze nötig. Daher erfolgt die Gruppenbildung in einem solchen Fall an einer Stichprobe und im Anschluss werden alle Objekte der Grundgesamtheit der Teilbauflächen durch ein geeignetes Klassifizierungsverfahren den gebildeten Gruppen zugeordnet.

5 Entwicklung einer Methodik zur geodatenbasierten Quartierstypisierung

Datenüberprüfung und -auswahl

Methodik zur GIS-datenbasierten Quartierstypisierung



Verwendung von Typquartieren und Mengengerüsten

Abbildung 5-1: Verfahrensschritte der Methodik zur GIS-datenbasierten Typisierung von Quartieren

Weiterhin umfasst die in diesem Kapitel entwickelte Methodik neben der Zusammenführung von Teilbauflächen zu Quartieren (Kapitel 5.5) auch deren Quantifizierung und statistische Beschreibung sowie die Festlegung von Kennwerten für Typquartiere (Kapitel 5.6). Im Rahmen der statistischen Beschreibung der Quartiere ist das Bewusstsein darüber wichtig, dass in diese Analyse die Einschränkungen der zugrundeliegenden Datenbasis eingehen. Dies betrifft sowohl die Vollständigkeit der Abbildung aller Gebäude und Teilbauflächen in Deutschland als auch die für die Gebäude vorliegenden Merkmalsausprägungen. Die Datenbasis wird damit als Bezugsdatensatz und Referenz betrachtet,

welche eine detaillierte Analyse und die Ableitung von wissenschaftlichen Erkenntnissen erlaubt. Wo Abweichungen zwischen dieser Referenz und der Realität aus anderen Studien bekannt sind, werden diese Diskrepanzen diskutiert und über modellinterne Anpassungen entschieden. Das Ziel der Arbeit besteht jedoch, wie in Kapitel 1 dargelegt, nicht in einer z. B. stichprobenbasierten Überprüfung der Repräsentativität der verfügbaren Geobasisdaten hinsichtlich der Zahl der Gebäude und der für die Untersuchung verwendeten Merkmale. Vielmehr soll die Methodik die kontinuierliche Fortschreibung der Basisdaten berücksichtigen und somit eine stetige Verbesserung ermöglicht werden.

Im nachfolgenden Kapitel 5.2 werden die zur Entwicklung und zur späteren Anwendung der Methodik verfügbaren GIS-Daten sowie die verwendete Software beschrieben, bevor die einzelnen Verfahrensschritte detailliert erläutert werden. Die Anwendung der Methodik am Beispiel eines auf die Stadt Berlin eingegrenzten Datensatzes erfolgt in Kapitel 6. Anhand der Anwendung der Methodik und der daraus gewonnenen Erkenntnisse werden die Ergebnisse der Methodik plausibilisiert und ihre Praxistauglichkeit verdeutlicht.

5.2 Beschreibung verfügbarer GIS-Daten und verwendeter Software

Für die Methodenentwicklung und deren Anwendung im Rahmen dieser Arbeit konnten die Geobasisdaten Digitales Basis-Landschaftsmodell (Basis-DLM), Georeferenzierte Adressdaten (GA), Verwaltungsgebiete (VG25), Kommunale Teilgebiete (KT25) und 3D-Gebäudemodelle im LoD2 kostenfrei vom Bundesamt für Kartographie und Geodäsie (BKG) bezogen werden (siehe Tabelle 5-1, S. 80).⁴⁶ Die ebenfalls zur Verfügung stehenden 3D-Gebäudemodelle im LoD1 werden für die Methodenentwicklung und -anwendung nicht verwendet, da alle Gebäudeeigenschaften daraus auch über den Datensatz im LoD2 verfügbar sind. Die Amtlichen Hausumringe (HU-DE) und Amtlichen Hauskoordinaten (HK-DE), welche in anderen raum- und gebäudebezogenen Untersuchungen Anwendung finden (vgl. Kapitel 4.3), liefern für diese Arbeit ebenfalls keine über die geometrischen und semantischen Daten der 3D-Gebäudemodelle und Adresspunkte der GA hinausgehenden Informationen. Es ist zudem anzumerken, dass der Datensatz Basis-DLM keine Gebäudeobjekte des ALKIS enthält, obwohl die Datenstruktur des Basis-DLM eine Abbildung dieser Daten ermöglichen würde (vgl. Tabelle 4-1, S. 66).

⁴⁶ Teile dieser Arbeit sind im Rahmen des Forschungsvorhabens „EG2050:E4Q – Einbindung erneuerbarer Energieträger in die Energieversorgung vernetzter Quartiere“ bearbeitet worden. Die Vergabe der Nutzungslizenz erfolgte für das Teilvorhaben „Entwicklung einer geoinformationsdatenbasierten Typisierung städtischer Quartiere“ am Institut Wohnen und Umwelt, Darmstadt (FKZ 03EGB0014B).

Alle verfügbaren Daten liegen flächendeckend für das deutsche Bundesgebiet vor und wurden im Mai bzw. Juli 2021 vom BKG bereitgestellt.⁴⁷ Der Datensatz bildet damit Aktualisierungen der in ALKIS verfügbaren Informationen bis einschließlich 2020 ab. Die verwendeten GIS-Daten sind in Abbildung 5-2 visualisiert. Die Abstufung der räumlichen Ebenen von Siedlungen (VG25) zu Bezirken (KT25), Ortslagen und Teilbauflächen (Basis-DLM) sowie Gebäudegrundrissen (3D-Gebäudemodell LoD2) ist dabei farblich gekennzeichnet. Die braun gekennzeichneten Geoobjekte entsprechen den Elementen des jeweiligen Datensatzes. Falls der nachfolgende Datenausschnitt in einer 1:n-Beziehung abgebildet werden kann, ist diese durch eine rote Kennzeichnung markiert. Die vorherige Ebene wird jeweils durch Grautöne gekennzeichnet.

Tabelle 5-1: Im Rahmen dieser Untersuchung zur Verfügung stehende GIS-Daten für Deutschland

Datensatz	Beschreibung	Kürzel	Datenformat
Verwaltungsgebiete	Städte und Gemeinden	VG25_GEM	Shape
Kommunale Teilgebiete	Stadtbezirke	KT25	Shape
Basis-DLM	Ortslage	Basis-DLM (sie01_f)	Shape
	Baulich geprägte Flächen	Basis-DLM (sie02_f)	Shape
	Straßen	Basis-DLM (ver01_l)	Shape
3D-Gebäudemodelle	Gebäude und Gebäudeteile	LoD2	CityGML
Georeferenzierte Adressdaten	Adresspunkte	GA	CSV

⁴⁷ Die 3D-Gebäudedaten im LoD2 standen ab Juli 2021 für die Nutzung zur Verfügung. Alle anderen angegebenen Geobasisdaten wurden im Mai 2021 an den Lizenznehmer überstellt. Der Datensatz für die Freie Hansestadt Bremen (inkl. Bremerhaven) wurde als gegenüber dem BKG-Datensatz von 2020 korrigierte Version für das Jahr 2021 direkt über das Landesamt GeoInformation Bremen bezogen.



Abbildung 5-2: Räumliche Ebenen der GIS-Daten zur Methodenentwicklung (am Beispiel Berlin)

(1): Städte und Gemeinden (VG25), Markierung: Berlin;

(2): Kommunale Teilgebiete (KT25), Markierung: Charlottenburg-Wilmersdorf;

(3): Ortslagen (Basis-DLM sie01_f); Markierung: Charlottenburg;

(4): Teilbauflächen (Basis-DLM sie02_f), Ausschnitt: Ortslage Charlottenburg;

(5): Vereinigte Gebäudegrundrisse (LoD2)

(6): Georeferenzierte Adressdaten (GA)

[eigene Darstellung; Daten: GeoBasis-DE / BKG (2021) bzw. BKG, Deutsche Post Direkt GmbH, Statistisches Bundesamt, Wiesbaden (2021)]

Die 3D-Gebäudemodelle im LoD2 bilden die geometrisch-topologische Grundlage der raumbezogenen Analyse dieser Arbeit und liefern zudem wichtige Informationen zur Beschreibung der Gebäude. Die Datenstruktur des CityGML-Formats⁴⁸ bildet die im Datensatz befindlichen Gebäude dabei in Abhängigkeit der Komplexität ihrer Geometrie als einzelnes Objekt oder als mehrere Gebäudeteile ab. Geometrische Volumenkörper und deren begrenzende Flächenelemente sind wiederum als Bestandteile der Einzelgebäude bzw. der Gebäudeteile definiert. Während die Datenstruktur des CityGML-Standards und des AAA-Referenzmodells eine Vielzahl von beschreibenden Attributen für die semantische Anreicherung eines Datensatzes für 3D-Gebäudemodelle erlaubt, ist in dem für Deutschland flächendeckend bereitgestellten Datensatz nur ein geringer Teil davon verfügbar (siehe Anhang C und Tabelle Anhang C-3).

Die modellhafte Abbildung der Gebäude geht zudem mit einigen Herausforderungen einher. Es handelt sich dabei z. B. um die Abbildung von Gebäuden, welche in der menschlichen Wahrnehmung als ein Gebäudekomplex aus Haupt- und Nebengebäude(n) mit funktionalem Zusammenhang aufgefasst werden, aber im Datensatz eigenständige Objekte darstellen, d. h. verschiedene Gebäudeidentifikationsnummern besitzen. Auch die Unterteilung von Gebäuden in mehrere Gebäudeteile lässt mitunter nicht immer eindeutige Rückschlüsse darauf zu, welche der Gebäudeteile in einer gebäudebezogenen Datenanalyse als relevant⁴⁹ berücksichtigt werden sollten. So werden im Datensatz z. B. auch Überdachungen von Hauseingängen und ähnliche bauliche Strukturen als Gebäudeteile modelliert, welche aus energetischer Sicht nicht als Bezugsfläche für die Gruppierung und Bewertung von Gebäuden herangezogen werden sollten. Auf diese Aspekte der Datenmodellierung wird im Rahmen der Datenaufbereitung und -verarbeitung bei der Anwendung der Methodik in Kapitel 6.2 (S. 110) ein besonderes Augenmerk gerichtet.

Auch die GA des BKG stehen für die Datenauswertung und Methodenentwicklung zur Verfügung. Da die GA jedoch, wie noch detailliert wird, keinen besonderen Mehrwert für die Analyse der vorliegenden Daten besitzen, werden diese in der Methodenentwicklung nicht weiter berücksichtigt (vgl. Kapitel 6.2). Die Anzahl der im Datensatz der 3D-Gebäudemodelle abgebildeten Gebäude und Gebäudeteile sowie der Adressdatenpunkte der GA ist in Tabelle 5.2 dargestellt. Die Gegenüberstellung legt den Schluss nah, dass in der Art

⁴⁸ Die Daten des BKG basieren auf der Version 1.0 des CityGML-Standards.

⁴⁹ Die Relevanz eines Gebäudeteils ergibt sich im Kontext dieser Untersuchung einerseits aus der energetischen Wirkung einer Berücksichtigung bzw. eines Außer-Acht-Lassens von Gebäudeteilen bei der energetischen Bewertung von Gebäuden und Quartieren sowie deren Versorgungskonzepten. Zum anderen wird die Relevanz bestimmt durch die Auswirkungen auf die zu bewertenden baulich-räumlichen Strukturen und dadurch auf die Einteilung des Stadtraumes in Quartiere.

der modellhaften Abbildung der Gebäude zwischen den Bundesländern z. T. erhebliche Unterschiede bestehen, insbesondere bei der Modellierung von Gebäuden aus unterschiedlich vielen Gebäudeteilpolygonen. Vor der Verwendung der Daten zur Anwendung der Methode zur GIS-datenbasierten Quartierstypisierung ist deshalb in besonderem Maße Wert auf eine weiterführende Datenprüfung und -auswahl zu legen, welche auch die unterschiedlich detaillierte Modellierung von Gebäuden aus Gebäudeteilpolygonen berücksichtigt (siehe auch Kapitel 6.2).

Tabelle 5-2: Gebäude- und Adressbestand in 3D-Gebäudemolldaten (LoD2) und Georeferenzierten Adressdaten (GA) des BKG⁵⁰ [eigene Datenauswertung; Daten: GeoBasis-DE / BKG (2021), BKG, Deutsche Post Direkt GmbH, Statistisches Bundesamt, Wiesbaden (2021), Landesamt GeoInformation Bremen (2021)]

Bundesland	3D-Gebäudedatensatz (2020)			GA	
	Grundflächenpolygone	Gebäude	Verhältnis	Adresspunkte	Verhältnis
Baden-Württemberg	6.384.129	6.035.996	1,06	3.132.671	1,93
Bayern	8.641.577	8.622.392	1,00	3.689.111	2,34
Berlin	969.861	640.397	1,51	405.384	1,58
Brandenburg	2.922.346	2.318.876	1,26	864.023	2,68
Bremen	313.496	313.222	1,01	176.774	1,80
Hamburg	724.446	379.610	1,91	287.736	1,32
Hessen	6.831.016	4.918.568	1,39	1.617.275	3,04
Mecklenburg-Vorpommern	826.551	1.125.669	0,73	508.863	2,21
Niedersachsen	10.343.517	6.796.528	1,52	2.611.958	2,60
Nordrhein-Westfalen	27.354.882	10.763.451	2,54	4.474.148	2,41
Rheinland-Pfalz	4.025.553	3.127.691	1,29	1.424.446	2,20
Saarland	837.973	677.954	1,24	342.245	1,98
Sachsen	2.684.619	2.070.421	1,30	989.778	2,09
Sachsen-Anhalt	3.269.066	1.746.674	1,87	671.955	2,60
Schleswig-Holstein	2.776.686	2.292.795	1,21	956.423	2,40
Thüringen	2.646.206	2.241.990	1,18	628.729	3,57
Gesamt	81.551.898	54.077.164		22.781.519	

⁵⁰ In dieser Tabelle entsprechen die Gebäude einer geodaten-technischen Abbildung mit Gebäudepolygonen bzw. Polygonen für Gebäudeteile und eigener Root-ID. Dieser Gebäudebegriff unterscheidet sich von anderen statistischen Abgrenzungen, die z. B. anhand der Adresszugehörigkeit vorgenommen werden.

Zur Verarbeitung der GIS-Daten werden verschiedene kostenfreie Softwarelösungen benutzt und mit deren Hilfe der Datenimport, die Datenverarbeitung und die Visualisierung von GIS-Daten durchgeführt. Da die 3D-Gebäudemodelle für Deutschland in dem bereits komprimierten Datenformat des BKG etwa 800 GB Speicherplatz in Anspruch nehmen, wird für die Verarbeitung dieser und der anderen GIS-Daten eine „PostgreSQL“-Datenbank (Version 13) erstellt. Durch die Datenbankerweiterung „PostGIS“ (Version 3.1) wird diese Datenbank auch in die Lage versetzt, GIS-Daten zu verarbeiten und räumliche Operationen durchzuführen.⁵¹ Der Zugriff auf die Datenbank erfolgt über die PostgreSQL-integrierte und kommandozeilenbasierte Schnittstelle „psql“ oder mit der Anwendung „pgAdmin 4“ (Version 6). Zur Visualisierung der Geoinformationsdaten wird die Software „QGIS“ (Version 3.20) verwendet, welche mit verschiedenen Modulen zur Geodatenverarbeitung zur Verfügung steht. Die Verarbeitung der im CityGML-Format vorliegenden 3D-Gebäudemodelle erfordert darüber hinaus den Einsatz weiterer Softwarelösungen. Zur allgemeinen Visualisierung im Rahmen der Datenprüfung wird der „FKZ-Viewer“ (Version 6.2) des Karlsruher Institut für Technologie eingesetzt. Dieses Werkzeug ermöglicht eine direkte, dreidimensionale Ansicht der 3D-Gebäudemodelldaten im CityGML-Format, was nicht zum Funktionsumfang von QGIS gehört.

Für den Import der 3D-Gebäudemodelle in die PostGIS-Datenbank kann auf die Lösung „3DCityDB-Database“⁵² (Version 5.0) zurückgegriffen werden. Die 3DCityDB stellt dabei einen Übersetzer zwischen der objekt-relationalen PostGIS-Datenbank sowie den textbasierten und in der UML-Struktur aufgebauten CityGML-Datenmodellen dar. Das hierarchisch-strukturierte Beziehungsschema im CityGML-Format wird durch die 3DCityDB somit in eine datenbankgerechte Abbildung überführt.⁵³ Neben einer grafischen Oberfläche bietet der zugehörige 3DCityDB-Importer-Exporter die Möglichkeit der Ausführung der Java-basierten Importroutinen auch über die Kommandozeile des Microsoft Betriebssystems Windows und somit eine vergleichsweise einfache Ausführung von Stapelverarbeitungsdateien oder weiterführenden Automatisierungsprozessen. (Yao et al. 2018; Kunde 2013; Nagel et al. 2020)

⁵¹ Im Folgenden wird der Begriff PostGIS-Datenbank verwendet, um die Fähigkeit der GIS-Datenverarbeitung in der PostgreSQL-Datenbank zu unterstreichen.

⁵² Die PostGIS-Erweiterung 3DCityDB-Database wird nachfolgend lediglich mit 3DCityDB bezeichnet.

⁵³ Auf eine ausführliche Beschreibung der informationstechnischen Hintergründe und Anwendungsmöglichkeiten der 3DCityDB soll hier verzichtet werden, da diese Ausführungen nur wenig Mehrwert für das Verständnis der weiteren Datenverarbeitungsprozesse und Methodenentwicklung liefern würden. Für weitere Informationen zur 3DCityDB siehe z. B. Yao et al. (2018), Kunde (2013) und Nagel et al. (2020).

Für die statistisch-informationstechnische Datenanalyse wird die Statistik-Software „R“ (Version 4.1.1) in der Anwendungsumgebung „R Studio“ (Version 1.4.1717) eingesetzt, da in PostgreSQL sowohl die Clusteranalyseverfahren als auch die Machine Learning-Algorithmen (vgl. Kapitel 3) nur rudimentär abgebildet sind. Neben der kostenfreien Verfügbarkeit zeichnet sich die Software R insbesondere durch ein Anwender- und Entwicklernetzwerk aus, welches in einer Vielzahl von Themenfeldern sowie mit einer Vielzahl von Methoden arbeitet und entwickelte Programmcodes für eine Verwendung durch andere Nutzer kostenfrei zur Verfügung stellt. Für die in dieser Arbeit wichtigen Clusteranalyse- und Machine Learning-Verfahren existieren viele sog. „Packages“, d. h. Programmcodes mit definierten Datenverarbeitungsroutinen. Eine Übersicht der wichtigsten für diese Studie verwendeten Packages ist Tabelle Anhang C-4 zu entnehmen. Durch den Einsatz verschiedener weiterer Funktionspakete in R kann die Software in der vorliegenden Arbeit als zentraler Kern der Datenverarbeitung genutzt werden, über den ein direkter Datenaustausch mit der PostGIS-Datenbank stattfindet und auch Steuerbefehle für die Windows-Kommandozeile formuliert und ausgeführt werden.

5.3 Datenvorverarbeitung

Ziel der Datenvorverarbeitung, als dem ersten Verfahrensschritt innerhalb der Methodik zur GIS-datenbasierten Quartierstypisierung, ist es, die Informationen zur Bebauungsstruktur auf einer kleinskaligen Raumebene, den Teilbauflächen, auszuwerten und diese Informationen so aufzubereiten, dass sie in der anschließenden Clusteranalyse die Bildung von Teilbauflächengruppen ähnlicher Bebauungsstruktur erlauben. Im Kontext der vorliegenden Arbeit sind hierfür eine Auswahl relevanter und verfügbarer Merkmale von Gebäuden zu treffen sowie eine Form zur Beschreibung der topologischen Charakteristik der Gebäude innerhalb der jeweiligen Teilbaufläche festzulegen. Beides folgt der Zielsetzung, das Quartier als etablierte Betrachtungs- und Handlungsebene in der gebäudebezogenen Energieforschung durch eine datenbasierte Typisierung besser beschreibbar und für unterschiedliche Fragestellungen operationalisierbar zu machen. Wie in Abbildung 5-1 (S. 78) dargestellt folgt die Vorverarbeitung der Geoinformationsdaten einer vorgelagerten Überprüfung, Verarbeitung und Auswahl der für die Analyse und Typisierung zur Verfügung stehenden GIS-Daten. Während die Vorverarbeitungsschritte und die dafür nötigen Festlegungen in diesem Kapitel in verallgemeinerter Form dargelegt werden, kann eine beispielhafte Anwendung Kapitel 6.2 entnommen werden.

Für eine Beschreibung und Unterscheidung der energetischen Qualität der Gebäudeobjekte stehen in den GIS-Daten nur wenige Informationen zur Verfügung, welche im Rahmen von Energiebilanzberechnungen zur Ermittlung des Gebäudeenergiebedarfs benötigt werden. Die Eingangsdaten einschlägiger Energiebilanzverfahren können dabei vereinfacht in fünf Gruppen unterteilt werden: geometrische, bauphysikalische, anlagentechnische,

klimatische sowie nutzungsbedingte (vgl. DIN V 4108-6:2003-06; DIN V 18599-1:2018-09; Loga und Diefenbach 2013). Neben dem Standort der Gebäude, welcher in den Bilanzverfahren spezifische klimatische Eingangsparameter bedingt, sind ausschließlich die geometrischen Gebäudeeigenschaften in den flächendeckenden GIS-Daten verfügbar. Aus der Gruppe der geometrischen Bilanzeingangsparameter kann das Verhältnis der wärmeübertragenden Gebäudeaußenfläche zum Gebäudevolumen (A/V-Verhältnis), welches eine abgeleitete Größe der jeweiligen Gebäudegeometrie ist, als maßgebliche, den Energiebedarf bestimmende Kenngröße formuliert werden.⁵⁴ Im Hinblick auf den verfügbaren Datenbestand kann diese Kenngröße durch die Verwendung der Gebäudegrundfläche sowie der Gebäudehöhe näherungsweise abgebildet werden (siehe Tabelle 5-3, S. 88)⁵⁵.

Neben der energetischen Qualität der Gebäude ist im Rahmen von Untersuchungen möglicher Versorgungskonzepte auf Quartiersebene zudem die Wärmeabnahmedichte innerhalb eines Versorgungsgebietes ein wichtiges Kriterium. Die Wärmeabnahmedichte erlaubt die Abschätzung der Effizienz von leitungsgebundenen Wärmeversorgungssystemen (siehe auch Kapitel 2.4.1 und Anhang A), denn sie bestimmt das Verhältnis der Wärmeverluste im Leitungsnetz zu der für die Deckung des Wärmebedarfs des Versorgungsgebietes abgegebenen Wärmemenge⁵⁶ und ergibt sich aus dem Quotienten der Summe der in einem Netz zu versorgenden Nutz-, Wohn- bzw. Energiebezugsfläche und der Fläche des Versorgungsgebietes. Die Wärmeabnahmedichte ermöglicht somit eine Abschätzung der Effizienz von Wärmeversorgungsnetzen als Ergebnis einer vorliegenden Bebauungsstruktur. Die zur Verfügung stehenden GIS-Daten erlauben jedoch keine direkte Bestimmung der Wärmeabnahmedichte, da eine Ermittlung der Gebäudewärmebedarfe nicht möglich ist. Aus

⁵⁴ In einer vereinfachten Betrachtung sollen die bauphysikalischen, anlagentechnischen, klimatischen und nutzungsbedingten Bilanzparameter als konstant betrachtet werden. Der absolute Gebäudeheizwärmebedarf wird somit maßgeblich durch die Fläche der wärmeübertragenden Außenbauteile bestimmt. Die in der Energiebilanz ermittelten spezifischen Energiebedarfskennwerte ergeben sich wiederum aus einem Bezug auf die Nutz-, Wohn- bzw. Energiebezugsfläche eines Gebäudes. Aufgrund der in der Baupraxis notwendigen Mindestgeschosshöhen sowie aus Wirtschaftlichkeitsgründen typischen Geschosshöhen steht die Bezugsfläche in direktem Zusammenhang mit dem sich ergebenden Gebäudevolumen und das A/V-Verhältnis kann als Indikator einer energieeffizienten Gebäudegeometrie verwendet werden. (Vgl. DIN V 4108-6:2003-06; DIN V 18599-1:2018-09; Loga und Diefenbach 2013)

⁵⁵ Die ebenfalls zur Verfügung stehenden Informationen zur Gebäudedachform werden hier nicht noch einmal explizit berücksichtigt, da die Modellierungsvorschriften der 3D-Gebäudemodelle im LoD2 bereits eine Abbildung bei der Berechnung der mittleren Gebäudehöhe vorsehen und sie somit implizit enthalten sind.

⁵⁶ Auch für diese vereinfachende Betrachtung sind die bauphysikalischen Parameter der Bauteile des Wärmenetzes sowie anlagentechnische Parameter, Betriebsstrategien und Netztopologien als konstant anzunehmen.

diesem Grund kann auch für diesen Kennwert lediglich eine geometrisch-topologische Abbildung der Zusammenhänge berücksichtigt werden. Für die spätere Gruppenbildung von Teilbauflächen ähnlicher Bebauungsstrukturen werden deshalb zwei Maße der baulichen Nutzung definiert. Während das *Maß baulicher Nutzung I* (MBN I) die Summe der Gebäudegrundfläche je Teilbaufläche ins Verhältnis zur Fläche der Teilbaufläche setzt, ist das *Maß baulicher Nutzung II* (MBN II) als Quotient aus der Summe der Gebäudevolumina und der Fläche der Teilbaufläche formuliert (siehe Gl. 5-1 und Gl. 5-2).

$$\text{MBN I}_i = \frac{\sum a_j}{a_i} ; \text{mit } i \in [1, \dots, n] \text{ und } j \in [1, \dots, m] \quad \text{Gl. 5-1}$$

$$\text{MBN II}_i = \frac{\sum(a_j * h_j)}{a_i} ; \text{mit } i \in [1, \dots, n] \text{ und } j \in [1, \dots, m] \quad \text{Gl. 5-2}$$

Mit	a_j	=	Grundfläche des Gebäudes j
	a_i	=	Grundfläche der Teilbaufläche i
	h_j	=	Mittlere Höhe des Gebäudes j
	MBN I _{i}	=	Maß baulicher Nutzung I der Teilbaufläche i
	MBN II _{i}	=	Maß baulicher Nutzung II der Teilbaufläche i
	n	=	Anzahl Teilbauflächen
	m	=	Anzahl Gebäude auf jeweiliger Teilbaufläche

Neben den bereits formulierten Kenngrößen zur Beschreibung der Bebauungsstrukturen auf der Ebene von Teilbauflächen werden noch weitere Kennzahlen festgelegt, welche sowohl gebäudeenergetische Zusammenhänge als auch städtebauliche Aspekte beschreiben und somit eine Relevanz für den praktischen Nutzen der zu erzeugenden Typquartiere besitzen. Neben den Mittelwerten der geometrischen Gebäudegeometrieparameter wird zudem deren Streuung in Form der Standardabweichung s angegeben⁵⁷, um die Homogenität der Bebauungsstruktur ebenfalls darzustellen und im Rahmen der späteren Gruppierung berücksichtigen zu können. Weiterhin werden Kenngrößen definiert, welche es erlauben, die städtebaulich wahrnehmbare Bebauungsdichte zu berücksichtigen. So wird die flächenspezifische Anzahl der Gebäude der Teilbaufläche ebenso als Kenngröße festgelegt wie die flächenspezifische Anzahl der Gebäuderegionen der Teilbaufläche. Eine Gebäuderegion wird in diesem Zusammenhang in Anlehnung an Hecht (2014) als Vereinigung sich berührender Gebäudeobjekte innerhalb einer Teilbaufläche verstanden. Die Gebäuderegion ist zum einen aus gebäudeenergetischen Gründen von Bedeutung, denn sie liefert als zusätzliches Merkmal die Information, dass die wärmeübertragende Außenfläche eines Gebäudes

⁵⁷ Die für die Beschreibung der Bebauungsstruktur auf Ebene von Teilbauflächen verwendete Standardabweichung bezieht sich auf die auf dieser Teilbaufläche befindliche Grundgesamtheit der Gebäude.

nicht vollständig durch die vereinfachte Darstellung über Gebäudegrundfläche und -höhe zu ermitteln ist, da angebaute Gebäude die wärmeübertragende Außenfläche entsprechend verringern.⁵⁸ Durch die Verwendung des Mittelwertes und der Standardabweichung der Regionengrundfläche sowie des Mittelwertes der Gebäudeanzahl je Gebäuderegion wird zudem der städtebaulichen Charakteristik eines Siedlungsgebietes Rechnung getragen und eine möglichst praxisnahe Gruppierung ermöglicht. Tabelle 5-3 stellt die für die Gruppierung der Teilbauflächen im Rahmen der Clusteranalyse genutzten Parameter dar.

Tabelle 5-3: *Energetisch-städtebaulich relevante Merkmale zur Beschreibung der Bebauungsstruktur einer Teilbaufläche*

Name	Beschreibung	Einheit
<i>Mittelwert der Gebäudegrundfläche</i>	Mittelwert der Grundfläche aller Gebäude innerhalb der Teilbaufläche	m ²
<i>Standardabweichung der Gebäudegrundfläche</i>	Standardabweichung der Grundfläche aller Gebäude innerhalb der Teilbaufläche	m ²
<i>Mittelwert der mittleren Gebäudehöhe</i>	Mittelwert der mittleren Gebäudehöhe aller Gebäude innerhalb der Teilbaufläche	m
<i>Standardabweichung der mittleren Gebäudehöhe</i>	Standardabweichung der mittleren Gebäudehöhe aller Gebäude innerhalb der Teilbaufläche	m
<i>Spezifische Anzahl Gebäude</i>	Anzahl der Gebäude innerhalb der Teilbaufläche bezogen auf die Fläche der Teilbaufläche	1/m ²
<i>Spezifische Anzahl Gebäuderegionen</i>	Anzahl der Gebäuderegionen innerhalb der Teilbaufläche bezogen auf die Fläche der Teilbaufläche	1/m ²
<i>Mittelwert der Regionengrundfläche</i>	Mittelwert der Grundfläche aller Gebäuderegionen innerhalb der Teilbaufläche	m ²
<i>Standardabweichung der Regionengrundfläche</i>	Standardabweichung der Grundfläche aller Gebäuderegionen innerhalb der Teilbaufläche	m ²
<i>Mittelwert der Gebäudeanzahl je Region</i>	Mittelwert der Anzahl an Gebäuden, die eine gemeinsame Gebäuderegion bilden	-
<i>Maß baulicher Nutzung I (MBN I)</i>	Quotient aus der Summe aller Gebäudegrundflächen und der Grundfläche der Teilbaufläche (vergleichbar mit der Grundflächenzahl)	m ² /m ²
<i>Maß baulicher Nutzung II (MBN II)</i>	Quotient aus der Summe aller Bruttorauminhalte der Gebäude und der Grundfläche der Teilbaufläche	m ³ /m ²

⁵⁸ Der Begriff der Gebäuderegion soll hier an wenigen Beispielen veranschaulicht werden: Zwei Doppelhäuser bzw. Doppelhaushälften bilden eine gemeinsame Gebäuderegion. Mindestens drei in Reihenbauweise aneinandergebaute Baukörper einzelner Gebäude bilden ebenfalls eine Gebäuderegion.

Im Rahmen der Anwendung der Methodik können über die hier festgelegten Kenngrößen noch weiterführende Kennzahlen ausgewiesen werden, die dann einen informativen Charakter besitzen und keine Eingangsgröße der im nachfolgenden Abschnitt beschriebenen Gruppenbildung durch Clusteranalyse sind (vgl. Kapitel 6.2 und Anhang C-6).

5.4 Gruppenbildung mittels Clusteranalyse auf Ebene von Teilbauflächen

5.4.1 Allgemeine Erläuterungen zur durchgeführten Clusteranalyse

Dem Verfahrensschritt der Datenvorverarbeitung zur Beschreibung der Bebauungsstruktur auf Ebene von Teilbauflächen folgt die Durchführung einer Clusteranalyse zur Gruppierung von Teilbauflächen. Durch die im Rahmen der Clusteranalyse ausgeführte Gruppenbildung für Teilbauflächen ähnlicher Bebauungsstrukturen wird eine datenbasierte Einteilung der Bebauungsstrukturen des untersuchten Gebietes erreicht. Wie bereits in Kapitel 3.1 erläutert steht der Gruppenbildung mittels Clusteranalyse eine ingenieurmäßige Einteilung der Objekte innerhalb des Datensatzes nach definierten Kriterien, z. B. gemäß bestehenden Typologien, gegenüber.⁵⁹ Sowohl die clusteranalytische Bildung von Teilbauflächengruppen als auch die Verwendung bestehender Typologien für Siedlungsstrukturen bzw. Stadtraumtypen scheinen i. A. anwendbar, um eine empirische Fundierung mittels GIS-Daten zu erarbeiten. In der hier beschriebenen Methodenentwicklung soll jedoch unabhängig von vorherigen Typologien, welche sich aus räumlich, zeitlich oder thematisch stark abgegrenzten Fragestellungen ergaben, ein Verfahren für eine möglichst allgemeingültige Gruppeneinteilung erarbeitet werden. Dieses datenbasierte Vorgehen liefert somit eine objektive, nachvollziehbare und zukünftig fortschreibbare Beschreibung der Bebauungsstrukturen des Gebäudebestandes auf einer kleinskaligen räumlichen Ebene.

Wie bereits in Kapitel 3.3 erläutert sind für die Durchführung einer Clusteranalyse nach der Auswahl der zu berücksichtigenden Attribute ein geeignetes Distanz- oder Ähnlichkeitsmaß festzulegen und eines der verfügbaren Gruppenbildungsverfahren auf den Datensatz anzuwenden. Weiterhin sind die Kriterien zur Beurteilung der Güte einer Gruppierung zu definieren, eine Beurteilung von Gruppierungen mit verschiedenen Gruppenanzahlen vorzunehmen und schlussendlich eine im Kontext des Untersuchungszwecks als geeignet betrachtete Gruppenanzahl festzulegen. Die Festlegung einer für den Untersuchungszweck

⁵⁹ Ein Exkurs über die Möglichkeiten und Beschränkungen einer ingenieurmäßigen Einteilung der Gebäudeabbildungen in den flächendeckend zur Verfügung stehenden GIS-Daten wird zudem in Anhang C-5 gegeben.

geeigneten Gruppenanzahl ist dabei abhängig vom gewählten Verfahren und notwendig für z. B. die hierarchischen sowie für das partitionierende Verfahren PAM.⁶⁰

Im Rahmen der Untersuchungen hat sich gezeigt, dass die Objektanzahl des Datensatzes bereits für Teildatensätze der flächendeckend vorliegenden Geodaten die Grenzen des durch die Software R nutzbaren Arbeitsspeichers bei Weitem überschreiten. Da eine Distanzmatrix im Verhältnis $(n)^2$ zur Länge n des Vektors, d. h. im vorliegenden Fall zur Anzahl der in Gruppen einzuteilenden Teilbauflächen, steigt, überschreitet die im Arbeitsspeicher gehaltene Datenmenge schnell die durch Soft- und Hardware vorgegebenen Grenzen. Für die Durchführung der Untersuchung wird somit auch die Möglichkeit gegeben, eine Kombination aus Clusteranalyse zur Gruppenbildung basierend auf einer Stichprobe des Datensatzes und eine Klassifizierung zur Gruppenzuordnung verbleibender Elemente aufzubauen. Der durch die Clusteranalyse in Gruppen eingeteilte Datensatzausschnitt wird im Rahmen der Klassifizierung genutzt, um einen der überwachten ML-Algorithmen anzulernen und ein Klassifizierungsmodell zu erstellen. Daran anschließend wird dieses Modell verwendet um eine Klassifizierung, d. h. eine Zuordnung der verbliebenen Objekte des Datensatzes zu den im Rahmen der Clusteranalyse festgelegten Gruppen, vorzunehmen (siehe Abbildung 5-3).

Um die Grenzen der Speicherverwaltung der Software auch im Prozessschritt der Klassifizierung nicht zu überschreiten, kann – falls erforderlich – eine Aufspaltung der Menge der Restobjekte des Datensatzes in mehrere Teilmengen erfolgen, welche voneinander unabhängig klassifiziert werden können.⁶¹ Generell lässt sich an dieser Stelle jedoch feststellen, dass weiterführende informationstechnische Entwicklungspotenziale für die direkte Verarbeitung besonders großer Datensätze mit Clusteranalyseverfahren in der Softwareumgebung R identifiziert wurden.

⁶⁰ In dem in Kapitel 3.3 vorgestellten dichte-basierten Verfahren DBSCAN ergibt sich die Gruppenanzahl jedoch über die Wahl der Eingangsparameter des Verfahrens. Dort gilt es, die Anwendbarkeit einer Gruppeneinteilung für die formulierte Forschungsfrage im Anschluss (meist subjektiv) zu beurteilen.

⁶¹ Für die Software R existiert die Möglichkeit, Algorithmen parallelisiert, d. h. in mehreren Instanzen des Programms auf demselben Computer auszuführen. Jede der initiierten Instanzen kann dabei den für die Ausführung verfügbaren Speicher vollständig ausnutzen, sofern die Ressourcen des zur Verfügung stehenden Computers dies erlauben. Von dieser Möglichkeit wird durch die Anwendung des „future.apply“-Packages in dieser Arbeit Gebrauch gemacht (vgl. Bengtsson 2021).

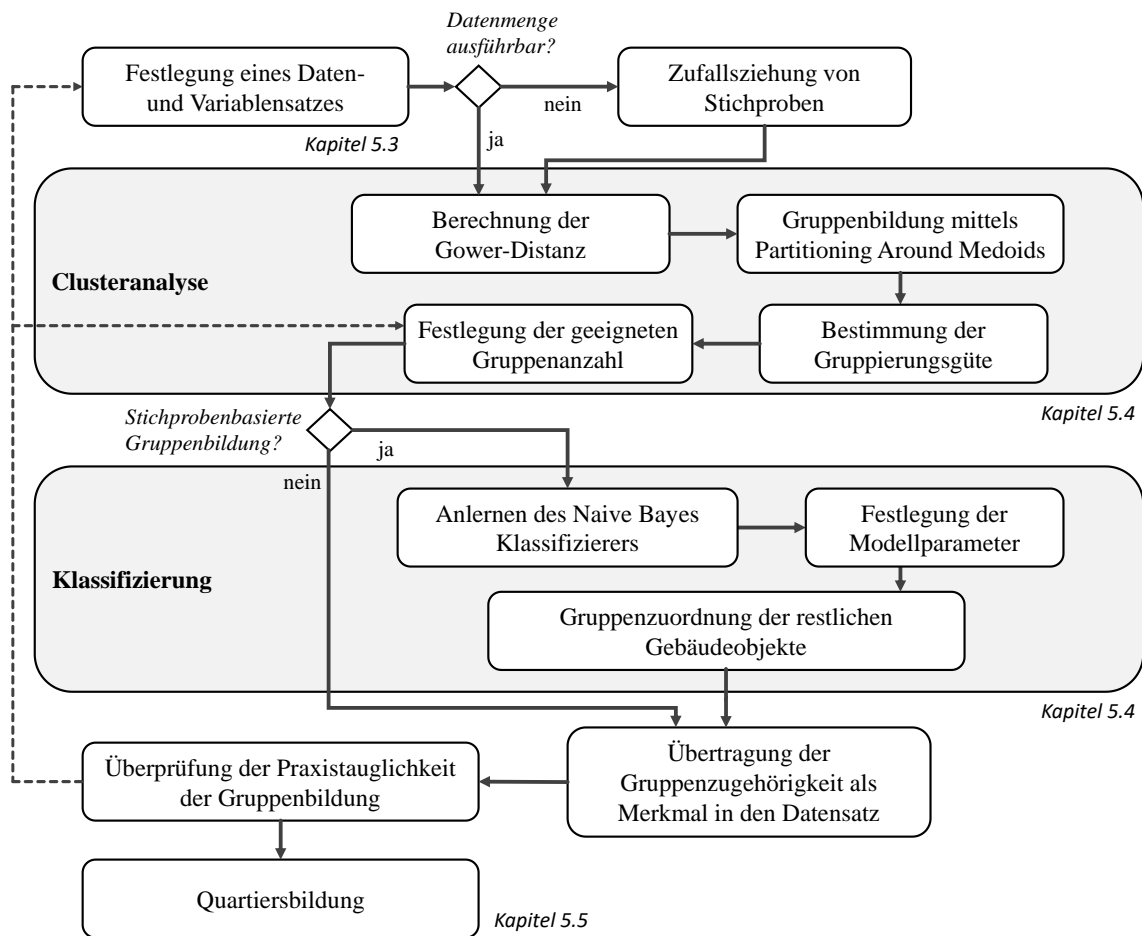


Abbildung 5-3: Ablauf zur Gruppenbildung aus einer Stichprobe der Teilbauflächen (Clusteranalyse) und Gruppenzuordnung übriger Objekte des Datensatzes (Klassifizierung) [eigene Darstellung]

Die Erfassung des räumlichen Zusammenhangs von Einzelgebäuden und eine Beschreibung der Bebauungsstruktur auf einer höheren räumlichen Ebene erfolgten bereits im Rahmen der Datenvorverarbeitung (Kapitel 5.3). Aus den Merkmalen auf Ebene der Teilbauflächen sollen zunächst Teilbauflächengruppen ähnlicher Bebauungsstrukturen und nachfolgend Quartiere gebildet werden. Dies erlaubt die spätere Quantifizierung der abgegrenzten Quartiere und deren statistische Analyse. In einem einfachen Gedankenspiel zur Gruppierung von Einzelgebäuden soll zudem gezeigt werden, dass eine Quartierstypisierung nicht erreicht werden kann durch die einmalige Anwendung der Gruppenbildung durch Clusteranalyse, wenn gleichzeitig die räumliche Zusammengehörigkeit und die geometrisch-inhaltlichen Attribute der Geoobjekte berücksichtigt werden sollen. Zwar legen Beschreibungen einiger existierender Clusteranalyseverfahren, wie DBSCAN, die Anwendbarkeit für solche raumbezogenen Fragestellungen nahe, die genauere Betrachtung zeigt jedoch ein gegenteiliges Bild: Die Clusteranalyse eines Datensatzes, in dem sowohl die Position von Objekten im Raum bzw. deren Abstand von Nachbarobjekten als auch die Ähnlichkeit der objektbeschreibenden Merkmale berücksichtigt wird, führt zu einer Einteilung in Gruppen, die abhängig von der Gewichtung der beiden Attributkategorien ist. Eine

starke Gewichtung der räumlichen Nähe führt zu Gruppen, die durch den Abstand der Gebäude zueinander bestimmt werden, welche jedoch eine starke Heterogenität hinsichtlich der Ausprägung geometrisch-inhaltlicher Merkmale aufweisen können. Wird die Gewichtung der geometrisch-inhaltlichen Ähnlichkeit der Objekte in den Vordergrund gestellt, führt dies dazu, dass sich die Objekte der Gruppen räumlich stark durchmischen (siehe Abbildung 5-4).



Abbildung 5-4: Vergleich der Gruppenbildung bei Schwerpunktsetzung auf die räumliche Nähe von Gebäuden (links) bzw. auf die geometrisch-inhaltlichen Gebäudeattribute (rechts) für fünf Gruppen [eigene Darstellung; Daten – GeoBasis-DE / BKG (2021)]

Des Weiteren werden bei dem Versuch einer direkten Typisierung von Quartieren aus Einzelgebäuden bei der Wahl eines hohen Gewichtes auf die räumliche Nähe und der Festlegung einer hohen Gruppenanzahl Quartiere gleicher Bebauungsstrukturen aber entsprechender Distanz zueinander in unterschiedliche Gruppen (bzw. Typen) eingeteilt. Die Überlegungen dieses gebäudebezogenen Gedankenexperiments lassen sich ohne Einschränkungen auf die Bildung von Gruppen ähnlicher Teilbauflächen übertragen, da auch dort die Zweiteilung in Attribute der räumlichen Position und Merkmale zur Beschreibung der Bebauungsstruktur existiert. Das Ergebnis einer einstufigen Einteilung von Quartieren widerspricht damit der Zielsetzung der Ableitung von Typquartieren als Vertreter einer Gruppe von Quartieren gleicher oder ähnlicher Bebauung und macht somit die im Rahmen der Arbeit entwickelte und aus mehreren Verfahrensschritten bestehende Methodik erforderlich.

5.4.2 Clusteranalyse zur Bildung von Teilbauflächengruppen ähnlicher Bebauungsstruktur

Die Gruppeneinteilung der Teilbauflächen mittels Clusteranalyse erfolgt basierend auf den in Kapitel 5.3 erarbeiteten Merkmalen zur Beschreibung der Bebauungsstruktur der Teilbauflächen (siehe Tabelle 5-3, S. 88) sowie auf dem in Abbildung 5-3 (S. 91) dargestellten Ablauf zur Gruppenbildung mittels Clusteranalyse einer Stichprobe des verfügbaren Datensatzes und nachgelagerter Klassifizierung. Die nachgelagerte Klassifizierung hat, wie voranstehend bereits erörtert, den Zweck, die restlichen Objekte außerhalb der Stichprobe den Teilbauflächengruppen zuzuordnen, welche im Rahmen der Clusteranalyse gebildet wurden.

Für die Durchführung der Clusteranalyse muss eine Festlegung des Distanzmaßes und des anzuwendenden Algorithmus zur Gruppeneinteilung getroffen werden. Um eine Distanzberechnung für einen Datensatz mit gemischt skalierten Variablen zu ermöglichen, wird das Verfahren nach Gower (1971) in der R-Implementierung des Funktionspakets *cluster* ausgeführt (vgl. Maechler et al. 2021). Dieses Verfahren erzeugt eine Distanzmatrix aus der Gewichtung der Unähnlichkeiten der metrischen und kategorialen Variablen. Für die metrischen Variablen wird hierbei nach einer Normierung der Merkmalswerte die euklidische Distanz ermittelt. Für die kategorialen oder binären Variablen erfolgt die Bestimmung der Ähnlichkeit s der Merkmalsausprägungen der jeweiligen Objekte nach dem Prinzip des Simple-Matching-Koeffizienten, d. h. gleiche Merkmalsausprägungen werden als ähnlich betrachtet und erhalten den Ähnlichkeitswert 1 (vgl. Kapitel 3.2). Die Distanz d der kategorialen oder binären Objekte ergibt sich aus dem Zusammenhang in Gl. 5-3.

$$d_{\text{Simple-Matching}} = 1 - s_{\text{Simple-Matching}} ; \text{mit } s \in [0,1] \quad \text{Gl. 5-3}$$

Wird vom Anwender keine abweichende Gewichtung vorgenommen, so werden alle Variablen im Rahmen der Distanzermittlung mittels Gower-Algorithmus gleich gewichtet. Enthält ein Datensatz Objekte, für welche eine Merkmalsausprägung nicht existiert, so wird diese Variable für die Berechnung der Gower-Distanz zwischen diesem und den anderen Objekten des Datensatzes außer Acht gelassen. Es ist zu beachten, dass für Datensätze mit einer hohen Anzahl an kategorialen Variablen, in Abhängigkeit der Charakteristik der Werteverteilung metrischer Attribute, eine überproportionale Gewichtung der kategorialen Merkmale eintreten kann. Für den vorliegenden Datensatz, welcher sich ausschließlich aus metrischen Variablen zusammensetzt, besteht eine solche Situation jedoch nicht. Im Rahmen einer Fortschreibung der Methodik, z. B. aufgrund einer zukünftigen Anreicherung der flächendeckend verfügbaren GIS-Daten mit weiterführenden, aus energetischer Sicht relevanten Gebäudemerkmalen, muss dies aber ggf. berücksichtigt werden.

Für die Clusteranalyse des Datensatzes wurde das bereits in Kapitel 3.3 vorgestellte Verfahren Partitioning Around Medoids (PAM) als geeignet identifiziert. Das Verfahren zur

Gruppenbildung nutzt direkt die Objektdistanzen, welche durch die Anwendung des Gower-Algorithmus in Form einer Distanzmatrix berechnet wurden. Es findet keine weitere euklidische Distanzberechnung oder die Bildung von Gruppenmittelwerten – welche für kategoriale Variablen nicht definiert sind – innerhalb des PAM-Algorithmus statt. Das ausgewählte Verfahren erlaubt somit die Gruppenbildung ohne spezifische, systematische Fehler aufgrund der Berechnung der Distanz für nicht-metrische Datensatzvariablen.

Als Bewertungskriterium für die Güte der Gruppeneinteilung wird die mittlere Silhouetten-Weite (ASW) festgelegt. Diese Festlegung ist ebenfalls vor dem Hintergrund einer möglichen Berücksichtigung gemischtskalierter Variablen notwendig. Die Silhouetten-Weite beruht auf der Auswertung der paarweisen Objektdistanzen und kann deshalb die in der Distanzmatrix verfügbaren Abstände der Gebäudeobjekte verwenden. Andere Verfahren, z. B. das Elbow-Kriterium auf Basis der Summe des quadratischen Abstandes aller in einem Cluster liegenden Objekte zum Mittelpunkt des Clusters (engl. within sum of squares; WSS), stützen sich auf eine Varianzanalyse bezogen auf den Gruppenmittelpunkt unter Anwendung einer euklidischen Metrik, welche für kategoriale Attribute eines Datensatzes keine korrekte Abbildung der Objekte erlaubt (siehe auch Kapitel 3.3.5).

Da der ausgewählte PAM-Algorithmus zur Gruppenbildung eine exogene Festlegung hinsichtlich der Gruppenanzahl erfordert, wird ein Intervall an möglichen Gruppenanzahlen festgelegt und eine iterative Ausführung dafür vorgenommen. Basierend auf der Recherche zu bereits bestehenden Siedlungsstruktur- und Stadtraumtypen in Kapitel 2.4 sowie eigenen Voruntersuchungen zeigte sich ein Intervall der Gruppenanzahl $k \in [2, \dots, 20]$ als zielführende Rahmensetzung. Für jede dieser Gruppenanzahlen wird das Kriterium zur Gruppierungsgüte, die ASW, berechnet und anhand lokaler Maxima die geeignete Gruppenanzahl festgelegt (siehe Abbildung 5-5 und Abbildung 5-6).

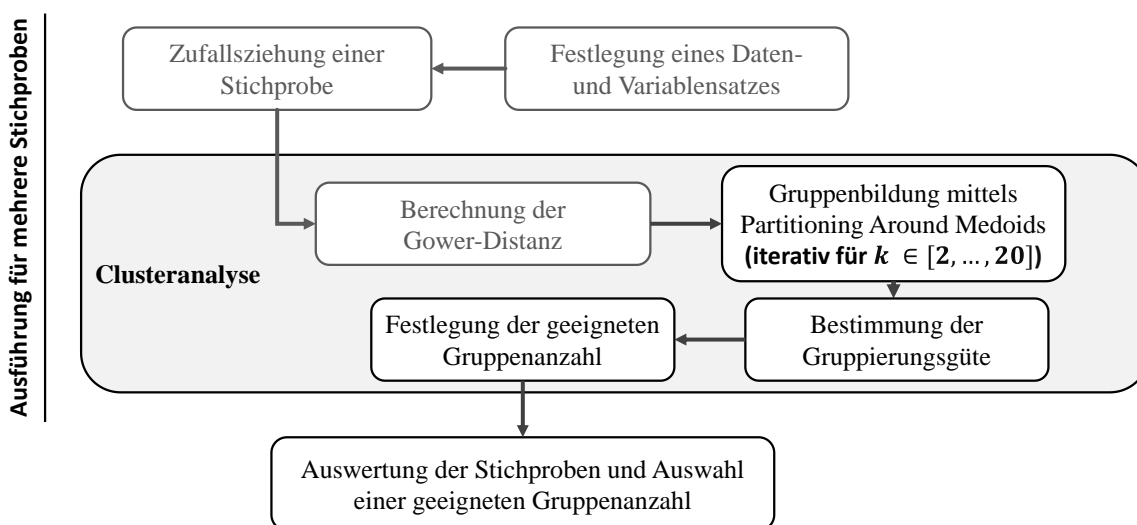


Abbildung 5-5: Iterative Bestandteile der Festlegung einer geeigneten Gruppenanzahl und der clusteranalytischen Gruppenbildung für Teilbauflächen

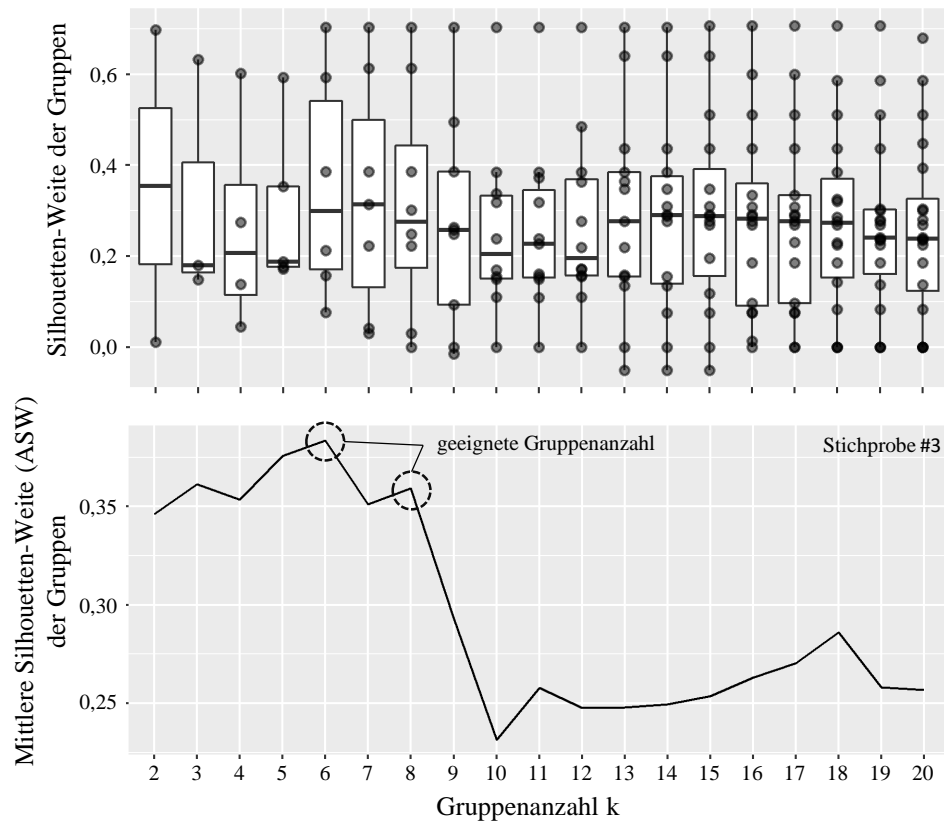


Abbildung 5-6: Beispielhafte Auswertung der mittleren Silhouetten-Weite für die mittels PAM-Algorithmus gebildeten Teilbauflächen für Gruppenanzahlen $k \in [2, \dots, 20]$
 (Hohe Werte der Silhouetten-Weite je Gruppe (oben) bzw. mittleren Silhouetten-Weite (unten) entsprechen einer hohen Homogenität der Gruppe bei guter Abgrenzung zur nächstgelegenen Nachbargruppe.)

Neben dem numerischen Wert der ASW wird bei der Festlegung einer geeigneten Gruppenanzahl zudem berücksichtigt, wie sich die Silhouetten-Weite der einzelnen Gruppen darstellt. Bei ähnlicher ASW wird die Gruppenanzahl bevorzugt, welche weniger Gruppen mit niedriger Silhouetten-Weite besitzt. Durch diese Festlegung wird für eine möglichst hohe Anzahl an Gruppen eine möglichst scharfe Abgrenzung zu anderen Gruppen erreicht.

Bei der endgültigen Festlegung der Gruppenanzahl wird die ASW jedoch nicht als das ausschließliche Kriterium angewandt, da häufig eine vergleichsweise geringe Gruppenanzahl die höchsten ASW-Werte ergibt. Dieser Umstand deckt sich jedoch nicht mit den ingenieurmäßigen Einteilungssystematiken für Siedlungsstruktur- bzw. Stadtraumtypen in den bereits vorgestellten Studien von z. B. Roth (1980) oder Hegger et al. (2012; 2015a) und gefährdet die Zielsetzung dieser Arbeit, einen praxisnahen Mehrwert für die Abschätzung von Emissionsminderungspotenzialen durch Quartierskonzepte zu liefern. Deshalb wird die Festlegung getroffen, dass auch im Falle von ASW-Maxima bei einer Gruppenanzahl von zwei bis vier Gruppen eine höhere Gruppenanzahl ausgewählt werden soll und zu prüfen ist, ob diese im Hinblick auf die Anwendbarkeit der Ergebnisse zu bevorzugen ist.

Wie in den allgemeinen Erläuterungen zum Aufbau der clusteranalytischen Untersuchung (Kapitel 5.1) bereits ausgeführt muss angesichts der großen Anzahl an Objekten in einem flächendeckenden Datensatz und der existierenden soft- und hardwareseitigen Beschränkung auf eine Kombination von Gruppierung und Klassifizierung zurückgegriffen werden. Wegen der Notwendigkeit des zweistufigen Prozesses aus Gruppenbildung und Klassifizierung wird für die Auswahl einer geeigneten Anzahl von Gruppen für die Einteilung der Teilbauflächen des Datensatzes, ergänzend zu der in Abbildung 5-3 (S. 91) dargestellten Clusteranalyse, ein iteratives Vorgehen gewählt (siehe Abbildung 5-5, S. 94). Die Ziehung von mehreren Zufallsstichproben aus dem Datensatz soll Zufallseffekte, z. B. durch eine hohe Anzahl „untypischer“ Teilbauflächen in einer Stichprobe, minimieren und somit die Robustheit der Festlegung der Gruppenanzahl steigern. Für jede der Stichproben erfolgen die Clusteranalyse mit Gruppenbildung im Intervall $k \in [2, \dots, 20]$, eine Bewertung der Gruppierungsgüte und die Festlegung einer geeigneten Gruppenanzahl. Aus einer Gegenüberstellung der Ergebnisse der Gruppenbildung je Stichprobe wird dann jene Gruppenanzahl gewählt, die im Sinne des Modus am häufigsten als geeignet betrachtet wurde. Eine Verdeutlichung des Vorgehens wird durch die Anwendung der Methodik in Kapitel 6.4 erreicht.

Für die Klassifizierung, d. h. die Zuordnung der restlichen Teilbauflächen zu den anhand der Stichprobe gebildeten Teilbauflächengruppen ähnlicher Bebauungsstruktur, wird der Naive Bayes Klassifizierer in der Implementierung des Packages „klaR“ in der Software R verwendet. Er kann, wie bereits in Kapitel 3.4 ausgeführt, mit unterschiedlichen Skalenniveaus arbeiten und durch eine ergänzte Kerndichteschätzung auch nicht normalverteilte Merkmalsausprägungen berücksichtigen. Zudem wird ein Laplace'scher Glättungswert verwendet, welcher die Berücksichtigung von A-priori-Wahrscheinlichkeiten nicht im Trainingsdatensatz vorhandener Merkmalsausprägungen erlaubt. Der Naive Bayes Klassifizierer ist dadurch ein robuster Algorithmus, der die Zugehörigkeitswahrscheinlichkeit von bisher nicht zu Gruppen zugeordneten Objekten zu ebendiesen Gruppen ermittelt und die Objekte der Gruppe mit der höchsten Zugehörigkeitswahrscheinlichkeit zuteilt.

Die als Ergebnis von Gruppenbildung und Klassifizierung vorliegende Zuordnung aller Teilbauflächen zu einer der festgelegten Teilbauflächengruppen wird zur weiteren Anwendbarkeit der Gruppeneinteilung im Datensatz der Teilbauflächen als neues Attribut „Teilbauflächengruppe“ aufgenommen. Die Gruppenzugehörigkeit wird zudem von der räumlich höheren Ebene der Teilbauflächen an die Gebäude des Gebäudedatensatzes vererbt, um eine statistische Beschreibung der Kenngrößen von Teilbauflächen und der

Gebäudepopulationen⁶² der Teilbauflächen vornehmen zu können. Durch die Einteilung der Teilbauflächen in Gruppen sowie durch die Erweiterung des Datensatzes um das Attribut der Teilbauflächengruppe können in den nächsten Verfahrensschritten die Bildung von Quartieren aus Teilbauflächen gleicher Gruppenzugehörigkeit sowie die statistische Beschreibung inkl. einer Festlegung von Kennzahlen für Typquartiere als Gruppenvertreter des Quartiersbestandes vorgenommen werden.

5.4.3 Vorgelagerte Untersuchung zur Clusteranalyse von Einzelgebäuden

Es wurden der Methodenentwicklung vorgelagerte Untersuchungen durchgeführt, in denen eine clusteranalytische Einteilung der Einzelgebäude aus den verfügbaren GIS-Daten vorgenommen wurde. Ziel war es, eine Gruppeneinteilung der Gebäude zu erzeugen, welche im weitesten Sinne existierenden Gebäudetypologien bzw. den grundlegenden Systematiken der Siedlungsstruktur- bzw. Stadtraumtypen entspricht. Die Zugehörigkeit der Gebäude zu Gebäudegruppen sollte als weiteres Merkmal in der Analyse der Bebauungsstruktur auf Ebene von Teilbauflächen verwendet werden. Die Clusteranalyse der Einzelgebäude wurde dabei mehrfach für Datensatzvarianten mit unterschiedlichem Umfang an Gebäudemerkmalen angewendet (Tabelle 5-4, S. 98) und unterschiedliche Gruppenanzahlen untersucht, die gemäß dem Gütekriterium der ASW als geeignet erschienen. Die Merkmale wurden im Rahmen der Gruppeneinteilung mittels PAM-Algorithmus dabei jeweils gleichgewichtet. Die Gruppeneinteilung der Einzelgebäude wurde an ausgewählten Gebieten des Datensatzes vorgenommen, in denen eine in der subjektiven Wahrnehmung den bestehenden Gebäudetypen entsprechende, d. h. hinsichtlich der Gebäudetypen homogene Bebauungssituation (z. B. Reihenhaussiedlung) vorliegt.

Als Ergebnis dieser Clusteranalysen für Einzelgebäude konnte keine Gruppeneinteilung gefunden werden, die für einen überwiegenden Teil der betrachteten Bebauungssituationen eine geeignete Gruppeneinteilung ergab. Es muss deshalb attestiert werden, dass ein Nachvollziehen bestehender Typologien für Einzelgebäude ein stark verändertes Vorgehen erfordert, welches nicht mit der Zielsetzung der datenbasierten Gruppeneinteilung in Einklang zu bringen ist und nachfolgend umrissen wird:

⁶² Der Begriff der Gebäudepopulation wird an dieser Stelle eingeführt, um eine Verwechslung von lokalen, kleinräumigen Gebäudebeständen auf Teilbauflächen oder in Quartieren mit Gebäudebeständen höherer Verwaltungsebenen, z. B. Städten, Bundesländern oder dem Bund, zu vermeiden.

- Die nachvollziehende Gruppenbildung für Einzelgebäude erfordert eine merkmalsbasierte Formulierung der Gruppeneinteilung sowie die (manuelle) Zuordnung einer Teilmenge eines Gebäudedatensatzes zu den Gruppen.
- Diese Teilmenge kann als Referenz zur Auswahl und Optimierung eines Klassifizierungsalgorithmus genutzt werden. Mit dem Klassifizierungsmodell kann anschließend eine Einteilung aller Gebäude in die bestehenden Gruppen der verwendeten Typologie vorgenommen werden.
- Die manuelle Zuordnung der Teilmenge der Gebäude kann ebenfalls zu einer gezielten Optimierung der Eingangsvariablen, der Variablen Gewichte sowie zur Festlegung der Gruppenanzahl einer Clusteranalyse dienen.

Beide formulierten Ansätze (ausschließliche Klassifizierung bzw. Optimierung der Verfahrensparameter der Clusteranalyse) basieren dabei auf einer vorab verfügbaren oder durch einen Anwender in einem spezifischen Forschungskontext selbst festgelegten Gebäudetypologie. Auch die Vorabenteilung der Gebäudeobjekte in die Gruppen der verwendeten Typologien entspricht einer subjektiven Zuordnung. Dieses Vorgehen weicht deshalb von dem im Rahmen dieser Arbeit verfolgten datengetriebenen Vorgehen mit dem Ziel einer möglichst allgemeingültigen Typisierung von Quartieren ab.

Tabelle 5-4: Variablensätze zur Clusteranalyse der Einzelgebäude mit dem Ziel einer Einteilung in Gruppen ähnlich bestehender Typologien

Variablensätze				
Gering	Gering+	Mittel	Mittel+	Hoch
<ul style="list-style-type: none"> • Grundfläche • mittlere Gebäudehöhe • Funktionsklasse 	<i>Gering</i> zzgl. <ul style="list-style-type: none"> • Anzahl direkt angrenzender Gebäude 	<i>Gering+</i> zzgl. <ul style="list-style-type: none"> • Dachform 	<i>Mittel</i> zzgl. <ul style="list-style-type: none"> • Roeck-Wert • Perimeterlänge • Kompaktheitszahl • Länge-zu-Breite-Verhältnis 	<i>Mittel</i> zzgl. <ul style="list-style-type: none"> • Merkmale der Gebäuderegion (Fläche, Länge-zu-Breite-Verhältnis, Anzahl Gebäude)

5.5 Bildung von Quartieren aus gruppierten Teilbauflächen

Im dritten Schritt der Methodik zur GIS-datenbasierten Typisierung von Quartieren werden die Teilbauflächen, die eher kleinräumige Objekte darstellen, zu Quartieren zusammengeführt (vgl. Abbildung 5-1, S. 78). Hierbei soll eine möglichst praxisnahe Abgrenzung von Quartieren erreicht werden, um eine Anwendung in der gebäudebezogenen Energieforschung zu ermöglichen. Es soll dabei berücksichtigt werden, dass der Quartiersbegriff in der städtebaulichen und energiebezogenen Praxis in unterschiedlichen Kontexten auch hinsichtlich der Flächenausdehnung von Bebauungssituationen unterschiedlich formuliert

wird. So werden z. B. für Projekte des Wärmecontractings meist kleinräumige Quartiere mit nur wenigen Gebäuden untersucht, während in Stadterneuerungsprozessen kommunaler Verwaltungen großräumige Quartiere und Stadtteile behandelt werden (HoE et al. 2021). Ob eine Möglichkeit besteht, die Quartiersbildung sowohl für Akteure, welche in kleinmaßstäblichen Quartieren von wenigen Gebäuden arbeiten, als auch für großskalige, z. B. stadtplanerische Aufgabenstellungen, dienlich zu gestalten, wird in den nachfolgenden Abschnitten sowie im Rahmen der Anwendung der Methodik in Kapitel 6.5 untersucht.

Um eine praxisgerechte Quartiersbildung aus den Teilbauflächen zu ermöglichen, können bauliche, verwaltungstechnische sowie Infrastrukturelemente, welche eine zerschneidende Wirkung haben können bzw. sollen, als Quartiersgrenze berücksichtigt werden. Zu diesem Zweck werden die bisherigen flächenhaften Datensätze der PostGIS-Datenbank um die Geodatenabbildung des deutschen Straßennetzes ergänzt. Die als Teil des Basis-DLM verfügbaren Straßenachsen lassen sich durch Filtersetzungen soweit einschränken, dass daraus ein Netz übergeordneter Straßen entsteht (siehe Abbildung 5-7, S. 100, und Tabelle 5-5, S. 100)⁶³.

Im Weiteren sind die Liniengeometrien des Straßennetzes durch einen Puffer zu Flächenelementen zu erweitern und die Bereiche der Überschneidung mit angrenzenden Teilbauflächen von diesen Teilbauflächen zu subtrahieren. Somit wird eine Kopie der Teilbauflächengeometrien erzeugt, welche einen definierten Abstand zu Straßenachsen besitzt und damit in der weiterführend beschriebenen Quartiersbildung als Grenze dient. Weitere zerschneidende Verkehrsanlagen, z. B. oberirdische Bahntrassen oder Binnengewässer müssen nicht gesondert berücksichtigt werden, da der Datensatz der Teilbauflächen diese Elemente bereits durch einen entsprechenden Abstand zwischen den Teilbauflächen abbildet. Im Rahmen der nachfolgenden Beschreibung der Systematik zur Zusammenführung von Teilbauflächen zu Quartieren bilden die Geoobjekte der Verkehrsinfrastruktur Grenzlinien des Quartierswachstums.

⁶³ In einer vorgelagerten Untersuchung zeigte sich die Geodatenabbildung des übergeordneten Berliner Straßennetzes, welche durch die Berliner Senatsverwaltung bereitgestellt wird, als besonders geeignet, um daraus infrastrukturbedingte Grenzen von Quartieren abzuleiten (Senatsverwaltung für Umwelt, Verkehr und Klimaschutz Berlin 2017). Die dort als übergeordnete Straßen der Kategorien 0, I und II attribuierten Liniengeometrien führen zu einem weitmaschigen Straßennetz, welches auch größere, aus stadtplanerischer Sicht interessante Quartierszuschnitte erlaubt. Das Schlüsselmerkmal „Straßenkategorie“ ist im flächendeckend für Deutschland verfügbaren Datensatz jedoch nicht vorhanden, weshalb durch die Filtersetzung ein engmaschigeres Straßennetz gebildet wird, was im Falle seiner Anwendung als Grenze folglich auch die maximale Flächenausdehnung von Quartieren begrenzt.

5 Entwicklung einer Methodik zur geodatenbasierten Quartierstypisierung



Abbildung 5-7: Beispiel des übergeordneten Straßennetzes als Zerschneidungselement für die Bildung von Quartieren aus Teilbauflächen (Beispiel Berlin, schwarz = Straßenachsen) [Daten: GeoBasis-DE / BKG (2021)]

Tabelle 5-5: Filterkriterien zur Erzeugung eines übergeordneten Straßennetzes als Grenzelement der Quartiersbildung aus Teilbauflächen

ATKIS-Attributart	Kennung	Wert(e)	Bezeichnung
Widmung	WDM	1301	Bundesautobahn
		1303	Bundesstraße
		1305	Landesstraße, Staatsstraße
		1306	Kreisstraße
Verkehrsbedeutung innerörtlich	BDI	1000	Durchgangsverkehr
Verkehrsbedeutung überörtlich	BDU	1000	Überörtlicher Verkehr
		1001	Fernverkehr
		1002	Regionalverkehr
„Hat direkt unten“ existiert	HDU_x ¹⁾	1	Existiert

¹⁾ Das Attribut HDU_x beschreibt ein direkt unter der Fahrbahn liegendes Verkehrsbauwerk, z. B. eine Brücke oder einen Tunnel.

Die Methodik ermöglicht zudem weitere Grenzen der Quartiersbildung zu berücksichtigen, um den verschiedenen möglichen Anwendungsfeldern besser gerecht zu werden. So können beispielsweise die Grenzen unterschiedlicher Objekte aus den verfügbaren Geobasisdaten Basis-DLM und VG25 (vgl. Kapitel 4.2) definiert werden und somit eine Quartiersbildung z. B. nur innerhalb einer Ortslage, eines Stadtbezirks, einer Gemeinde bzw. Stadt, oder eines Bundeslandes erlaubt werden. Findet keine Verwendung dieser ergänzenden grenzbildenden Elemente statt, so können z. B. auch über verwaltungstechnische Wirkungsbereiche hinweg Quartiere entstehen.⁶⁴

Das Grundprinzip des Quartierswachstums folgt in dieser Arbeit der Bestrebung, vergleichsweise homogene Bebauungsstrukturen zu einem Bezugsraum, hier Quartier, zusammenzufassen, um damit eine Basis für stadträumlich-energetische Analysen zu liefern. Es baut deshalb auf der Einteilung der Teilbauflächen in Gruppen ähnlicher Bebauungsstruktur auf. Entgegen der Zielsetzung einer möglichst homogenen Bebauung innerhalb eines Quartieres kann in begründeten Fällen eine Zuordnung von Teilbauflächen anderer Gruppenzugehörigkeit zu einem Quartier ebenfalls sinnvoll und praxisgerecht sein, wenn dadurch z. B. kleinflächige Inselquartiere innerhalb eines sonst homogenen Gebietes vermieden oder untypische Fragmente zu einem sonst homogenen Quartier zugeordnet werden. Die Berechnung von zusammenhängenden Flächenkonstrukten kann mit denselben raumbezogenen Vereinigungsalgorithmen durchgeführt werden, die auch für die Erzeugung und Auswertung der Gebäuderegionen genutzt wurden. Mittels raumbezogenem DBSCAN-Algorithmus werden Ketten bzw. Aggregationen sich berührender Teilbauflächenpolygone erzeugt. Die bestehenden Programmcodes müssen neben der Änderung des Bezuges auf die Teilbauflächen zudem noch dahingehend angepasst werden, dass eine Beschränkung der Zusammenführung auf Polygone gleicher Gruppenzugehörigkeit erfolgt. Dies wird durch eine iterative Ausführung für jede Teilbauflächengruppe mit einem entsprechend gefilterten Teildatensatz ermöglicht (siehe Abbildung 5-8, S. 102).

⁶⁴ Die Möglichkeit zur Berücksichtigung dieser Grenzelemente für die Quartiersbildung kann verdeutlicht werden, wenn die Perspektiven unterschiedlicher Akteure der gebäudebezogenen Energieforschung und -planung skizziert werden. Während ein kommunales Stadtplanungsamt meist ausschließlich die sich in seinem Verwaltungsgebiet befindlichen Quartiere erfassen und bearbeiten wird, können in übergeordneten nationalen Potenzialstudien oder privatwirtschaftlichen Unternehmungen auch Quartiere und Quartierskonzepte von Interesse sein, welche sich über eine Verwaltungsgrenze hinweg erstrecken und somit zwei Städten zugehörig sind.

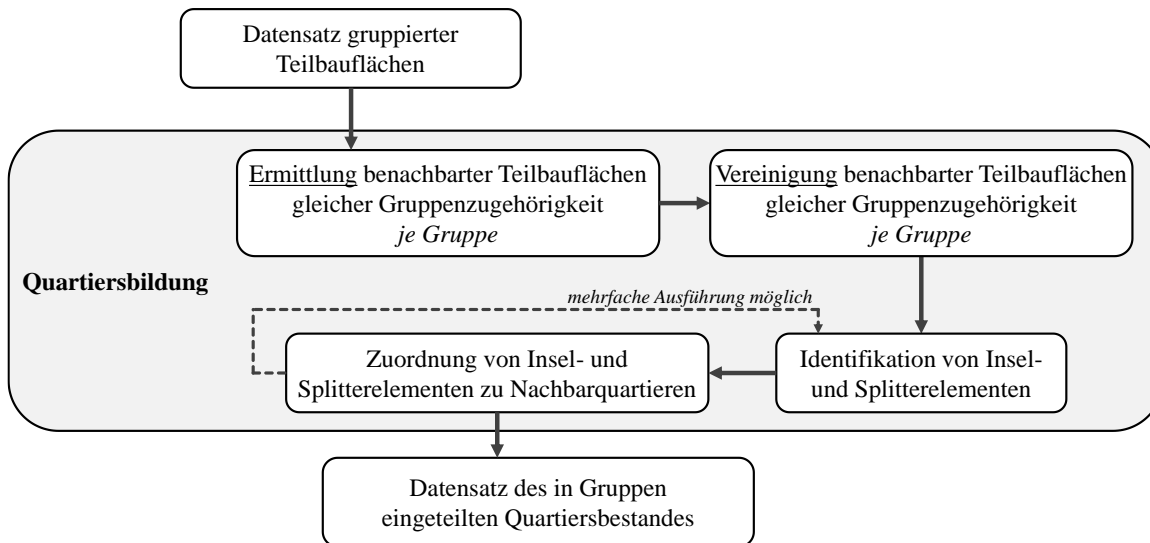


Abbildung 5-8: Schematischer Ablauf der Quartiersbildung aus einem Datensatz gruppierter Teilbauflächen

Die Ergebnisse dieses ersten Verfahrensteils des Quartierswachstums werden weiterhin auf das Vorhandensein splitter- bzw. inselhafter Polygone untersucht, die als Quartiersfragmente in der Praxis i. d. R. nicht als eigenständige Untersuchungsgebiete herangezogen werden. Es zeigt sich, dass hinsichtlich der flächenmäßigen Ausdehnung und der Anzahl von Gebäuden eine Vielzahl kleiner Inselgebiete entstanden ist. Diese können meist sowohl aus stadtstruktureller als auch aus energietechnischer Sicht den benachbarten Flächen zugeschlagen werden, ohne eine übermäßige Durchmischung der Bebauungsstruktur zu erzeugen. Grenzt ein solches kleinflächiges Gebiet an mehrere andere Flächenkonstrukte, so wird im Rahmen der Methodenentwicklung die Festlegung getroffen, dass dieses Quartiersfragment dem größten der benachbarten Gebiete zugeordnet wird. Das Grenzkriterium der flächenmäßigen Relation von absorbierendem und absorbiertem Gebiet ist dabei flexibel und kann durch eine Parametervariation im Rahmen der Anwendung festgelegt werden (siehe Abbildung 5-9). Die mehrfache Wiederholung dieses Wachstumsprozesses ist ebenfalls möglich und kann zu einer weiteren Reduktion kleinflächiger Insel- bzw. Splitterpolygone führen. Der Prozess der Quartiersbildung führt somit zur Entstehung unterschiedlich großer Quartiere, deren jeweilige Bebauungsstruktur weitestgehend ähnlich und der Struktur der Teilbauflächengruppen entsprechend ist. Für diese Quartiere kann nachfolgend eine statistische Beschreibung erfolgen und aus dem Bestand je Quartiersgruppe ein Typquartier abgeleitet werden.

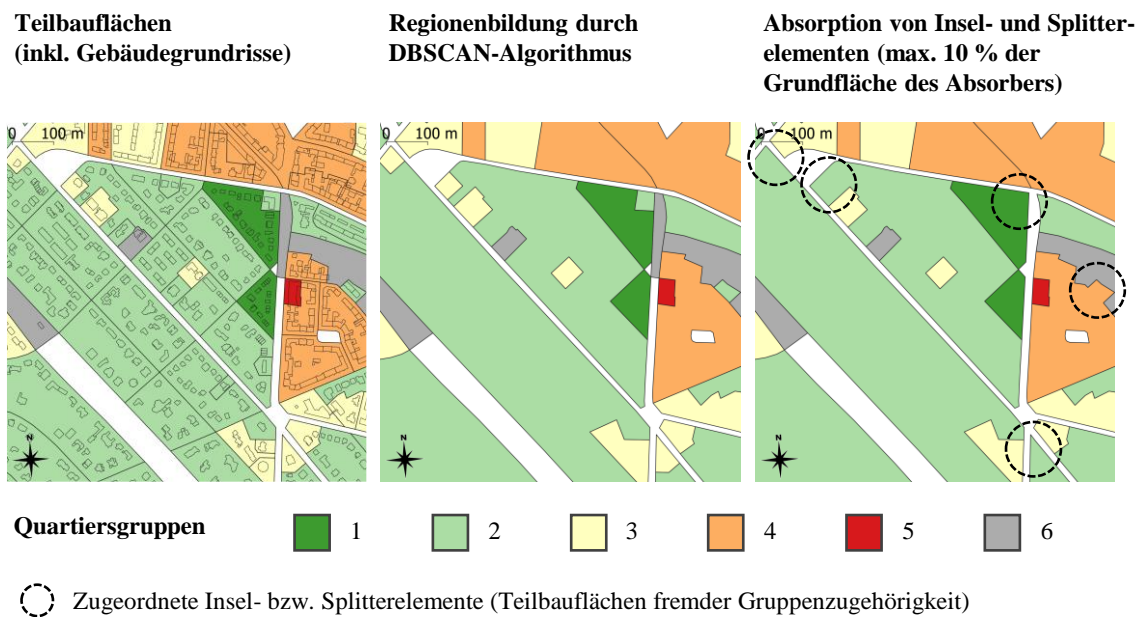


Abbildung 5-9: Beispiel der Bildung von Quartieren aus gruppierten Teilbauflächen (links) durch Vereinigung benachbarter Teilbauflächen gleicher Gruppenzugehörigkeit (Mitte) und Zuordnung von Insel- und Splitterelementen zur größten benachbarten Vereinigung von Teilbauflächen (rechts)

5.6 Statistische Analyse des Quartiersbestandes

Der Quartiersbildung folgen als Teil des letzten Verfahrensschrittes der Methodik zur GIS-datenbasierten Typisierung von Quartieren eine statistische Analyse des Quartiersbestandes sowie die Festlegung der beschreibenden Kennwerte für Typquartiere (siehe auch Abbildung 5-8). Diese Typquartiere, als Stellvertreter der aus den Teilbauflächen gebildeten Quartiersgruppen, ermöglichen die einfache Verwendung der Ergebnisse der Methoden-anwendung für weiterführende Studien in unterschiedlichen räumlichen und inhaltlichen Kontexten.

Die statistische Beschreibung der gebildeten Quartiere bzw. des erzeugten Quartiersbestandes erfolgt zweigeteilt. Zum einen sind in der gebäudebezogenen Energieforschung Kennzahlen bezüglich der Einzelgebäude und deren Merkmalsausprägungen innerhalb der Quartiersgruppen von Interesse. Diese erlauben einen Abgleich und ggf. eine Kalibrierung mit anderen Studien und empirischen Datenbasen und können zudem als Basisinformation für Potenzial- oder Szenarioanalysen auf verschiedenen räumlichen Ebenen dienen. Zum anderen wird auch das Quartier in seiner Aufteilung auf die erarbeiteten Quartiersgruppen und hinsichtlich seiner Merkmalsausprägungen statistisch ausgewertet. Hierbei liegt das Augenmerk auf der Erfassung und Beschreibung der Homogenität bzw. Heterogenität des Quartiersbestandes innerhalb der gebildeten Gruppen, z. B. bezogen auf die absolute Flächenausdehnung der Quartiere einer Gruppe oder hinsichtlich der Varianz der

flächenspezifischen Anzahl von Gebäuden. Diese Informationen sind für nachgelagerte Untersuchungen ebenfalls wichtig, um z. B. Unsicherheiten abschätzen zu können, wenn Potenzialanalysen ausschließlich auf Basis von Typquartieren durchgeführt wurden, oder um die Übertragbarkeit von lokalen Quartierskonzepten auf den Quartiersbestand bewerten zu können.

Für die statistische Beschreibung des Quartiersbestandes sowie der den Quartiersgruppen jeweils zugeordneten Gebäude stehen die in Tabelle 5-6 dargestellten Kenngrößen zur Verfügung. Zudem können für die Beschreibung einer höheren Raumebene als der Gebäudeebene jeweils auch statistische Kenngrößen der vorherigen Objekte herangezogen werden und als statistische Kennzahlen dienen (z. B. Mittlere Gebäudegrundfläche der Gebäude innerhalb einer Gebäuderegion). Da in unterschiedlichen Kontexten bei der Weiterverarbeitung der Ergebnisse der Methodenanwendung verschiedene Kennzahlen von Interesse sein können, wird für die allgemeine Anwendung und mit Bezug zur Forschungsfrage eine Auswahl an Merkmalen getroffen, die mindestens auszuwerten sind, damit die Ergebnisse einen Mehrwert für die gebäude- und quartiersbezogene Energieforschung sowie die Beurteilung von energetischen Quartierskonzepten besitzen. Die Merkmale leiten sich aus den in Kapitel 5.3 formulierten Überlegungen zu Gebäudeenergiebedarfen und netzgebundenen Wärmeversorgungsstrukturen ab und sind in Tabelle 5-6 entsprechend gekennzeichnet.

Der statistischen Auswertung und Beschreibung des in Quartiersgruppen eingeteilten Quartiersbestandes folgt die Festlegung von Kenngrößen für Typquartiere. Jede Quartiersgruppe soll dabei durch mindestens einen Typvertreter repräsentiert werden. Während i. d. R. für normalverteilte Variablen die Wahl des Mittelwertes eine geeignete Beschreibung eines synthetischen⁶⁵ Typvertreters erlaubt, ist im Hinblick auf die gebäude- und quartiersbezogenen Merkmale zu berücksichtigen, dass es sich z. T. um nicht normalverteilte Variablen handelt. Einige der Merkmalsausprägungen können z. B. nicht kleiner als Null bzw. eine definierte untere Grenze sein und es liegen rechtsschiefe Verteilungen vor (z. B. Grundfläche oder Gebäudehöhe). Aus diesem Grund wird festgelegt, dass für die Ableitung der synthetischen Typquartiere, als Vertreter der Quartiersgruppen, die Medianwerte der gruppenbezogenen Statistik verwendet werden. Diese berücksichtigen die Besonderheiten der Merkmalsverteilungen und erhöhen die Aussagekraft von weiterführenden Untersuchungen unter Verwendung der ermittelten Typquartiere.

⁶⁵ Der Begriff „synthetisch“ wird gewählt, um zu verdeutlichen, dass die erarbeiteten Typquartiere aus statistischen Kenngrößen abgeleitet werden. Demgegenüber können auch Beispielquartiere aus der Menge der Quartiere je Quartiersgruppe als Typvertreter herangezogen werden.

Tabelle 5-6: Für die statistische Beschreibung des Gebäude- und Quartiersbestandes auswertbare Merkmale (Für die gebäude- und quartiersbezogene Energieforschung und praxisnahe Ergebnisverwendung erforderliche Merkmale sind unterstrichen.)

Merkmalsgruppe	Merkmale [Einheit]
Objektpopulation	<ul style="list-style-type: none"> • <u>Anzahl Gebäude</u> [-] • Anzahl Gebäuderegionen [-] • Anzahl Teilbauflächen [-] • <u>Anzahl Quartiere</u> [-]
Gebäudemerkmale	<ul style="list-style-type: none"> • <u>Gebäudegrundfläche</u> [m²] • <u>Mittlere Gebäudehöhe</u> [m] • Gebäudevolumen [m³] • Dachform (Anteil) [%] • Funktionsklasse (Anteil) [%] • Roeck-Wert [-] • Kompaktheitszahl [-] • Perimeterlänge [m]
Merkmale der Gebäuderegion	<ul style="list-style-type: none"> • Anzahl Gebäude je Gebäuderegion [-] • Fläche der Gebäuderegion [m²] • Roeck-Wert [-] • Kompaktheitszahl [-] • Perimeterlänge [m]
Merkmale der Teilbaufläche	<ul style="list-style-type: none"> • Anzahl Gebäude je Teilbaufläche [-] • Anzahl Regionen je Teilbaufläche [-] • Fläche der Teilbaufläche [m²] • Maß baulicher Nutzung I [m²/m²] • Maß baulicher Nutzung II [m³/m²] • Spez. Anzahl Gebäude [1/ha] • Spez. Anzahl Regionen [1/ha]
Merkmale des Quartiers	<ul style="list-style-type: none"> • Anzahl Gebäude je Quartier [-] • Anzahl Regionen je Quartier [-] • Anzahl Teilbauflächen je Quartier [-] • <u>Fläche des Quartiers</u> [m²] • <u>Maß baulicher Nutzung I</u> [m²/m²] • <u>Maß baulicher Nutzung II</u> [m³/m²] • <u>Spez. Anzahl Gebäude</u> [1/ha] • Spez. Anzahl Regionen [1/ha] • Spez. Anzahl Teilbauflächen [1/ha]

Für einige der möglichen weiterführenden Studien unter Einbezug der Quartierstypen und Mengengerüste des Quartiersbestandes – insbesondere bei der Untersuchung netzgebundener Wärmeversorgungs-lösungen – scheint es zudem zielführend, eine weitere Differenzierung der festgelegten Typquartiere als Vertreter der Quartiersgruppen vorzunehmen. Dies ergibt sich insbesondere aus den erzielbaren energetischen und ökonomischen Synergieeffekten, die mit einem größeren⁶⁶ Versorgungsgebiet einhergehen. Bei der Anwendung der Methodik zur GIS-datenbasierten Quartierstypisierung ist deshalb zu entscheiden, ob neben der Festlegung von Kennwerten für Typquartiere auf Basis von Medianwerten der Merkmalsverteilungen der Quartiersgruppen zudem auch Typvertreter für Teilmengen kleinerer

⁶⁶ Der Begriff „größer“ kann sich dabei auf unterschiedliche Variablen beziehen, da sowohl höhere Wärmebedarfsdichten, höhere Anzahlen an Anschlussnehmern oder Wärmeerzeugungsanlagen größerer Spitzenleistungen jeweils zu unterschiedlichen Effizienzsteigerungen führen (z. B. durch Wirkungsgradsteigerungen oder ein verbessertes Verhältnis von Fixkosten zu variablen Kostenbestandteilen für den Betrieb und die Verwaltung einer netzgebundenen Wärmeversorgung).

oder größerer Quartiere innerhalb einer Quartiersgruppe erstellt werden sollen. Als zielführender Ansatz erscheint z. B. eine Aufteilung in drei Größenklassen, welche sich aus definierten Teilmengen zusammensetzen. Um bei der Unterteilung weiterhin aussagekräftige, d. h. als weitestgehend typisch zu betrachtende Quartierskennwerte zu erzielen, wird die Einteilung in drei Teilmengen anhand der Quartilsgrenzen vorgeschlagen. Das beschriebene Vorgehen zeigt Abbildung 5-10.

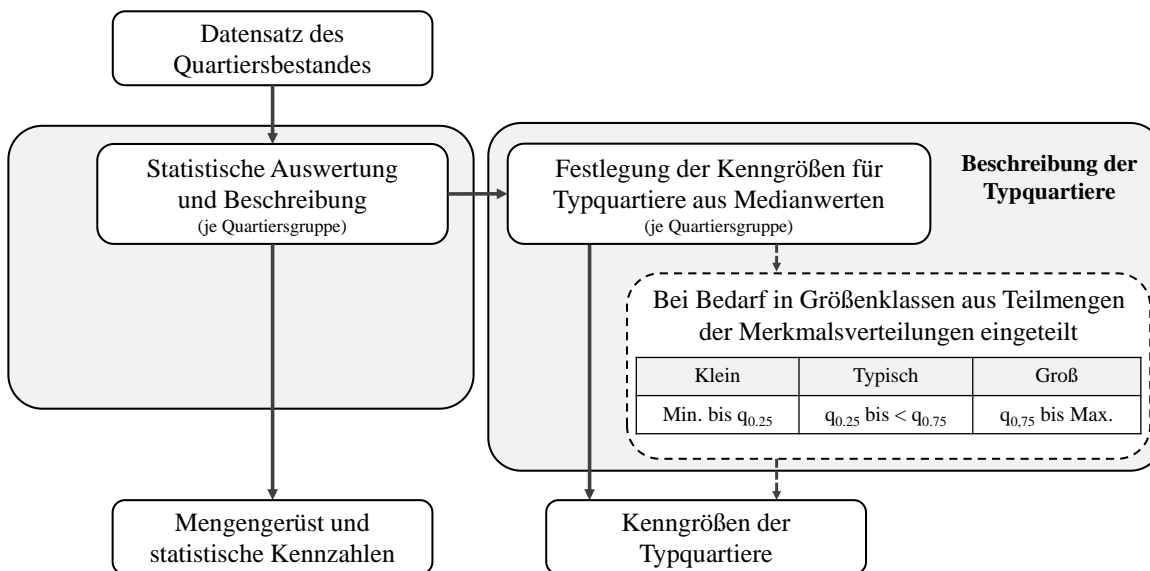


Abbildung 5-10: Statistische Analyse des Quartiersbestandes und Festlegung der Kennwerte für Typquartiere basierend auf Medianwerten der Merkmalsverteilungen (mit optionaler Aufteilung in Größenklassen)

5.7 Abschließende Anmerkungen zu den Randbedingungen der erarbeiteten Methodik

Nach der Erarbeitung und Beschreibung der Methodik zur GIS-datenbasierten Quartierstypisierung sollen, vor einer Anwendung der einzelnen Verfahrensschritte an einem Datensatz, noch einige ergänzende Anmerkungen formuliert werden. An dieser Stelle wird noch einmal auf die in Kapitel 4 vorgestellten und in Kapitel 5.2 weiter beschriebenen Geobasisdaten eingegangen. Wie bereits erläutert kann die – zum Zweck einer detaillierten energetischen Gebäudebewertung unvollständige – Datenbasis nur die geometrischen Einflüsse auf den Gebäudeenergiebedarf abbilden. Der sich aus den bauphysikalischen Eigenschaften oder der Gebäudenutzung ergebende Einfluss auf den Energiebedarf kann aus den flächendeckend verfügbaren Daten nicht ermittelt werden. Da die Methode und die Ergebnisse der Methoden-anwendung für vielfältige Forschungs- und Planungszwecke nutzbar sein sollen, wurde entschieden, keine Modellannahmen über die energetische Qualität der Gebäude zu treffen. Dieser Umstand ermöglicht und erfordert es gleichzeitig, für die Verwendung der Mengengerüste des Quartiersbestandes bzw. der abgeleiteten Typquartiere in nachgelagerten Vorhaben, eigene Modellannahmen über eine mittlere energetische Qualität der

Gebäude innerhalb der Quartiere oder eine weitere Unterteilung in Typquartiervarianten unterschiedlicher Effizienzlevels vorzunehmen. Diese offensichtliche Einschränkung der Methodik sowie der in Kapitel 6 erzeugten Ergebnisse führt aber ebenfalls zu einer breiten Nutzbarkeit von Methodik und Anwendungsergebnissen, denn eine Zurückrechnung empirischer Ergebnisse, welche unter Zuhilfenahme von Modellannahmen verarbeitet wurden, ist i. d. R. nicht oder nur unter Inkaufnahme systematischer Fehler möglich. Für den möglichen Anwendungsfall einer lokalen Potenzialabschätzung für eine Stadt oder einen Stadtteil können z. B. lokal vorliegende Informationen zu Modernisierungsniveaus der dortigen Gebäudepopulation verwendet werden, um die Ergebnisse dieser Arbeit zu ergänzen.

Weiterhin soll an dieser Stelle noch einmal auf die Betrachtungsebene „Quartier“ und dessen Erfassung eingegangen werden. Wie u. a. in Kapitel 2 ausgeführt wurde, ist die Betrachtungsebene des Quartiers in unterschiedlicher Weise für die gebäudebezogene Energieforschung von Bedeutung. Dabei hängt die Erfassung und Verarbeitung von Merkmalen, welche den Energiebedarf und die Energieversorgung beeinflussen, stark von der jeweiligen Fragestellung und der verfügbaren Informationslage ab. Als ideale Ausgangssituation ist hierbei die vollständige Abbildung aller Gebäude hinsichtlich aller energiebedarfsbedingenden Merkmale zu betrachten, welche folglich zu einer vollständigen Beschreibbarkeit der energiebezogenen Kennwerte eines Quartiers führen. In der Praxis ist ein solches Detaillierungsniveau jedoch i. d. R. nicht einmal für die Untersuchung lokaler, d. h. sehr kleinräumiger Gebiete vorzufinden. Eine Quartiersbetrachtung erfolgt häufig gerade auch deshalb, weil eine Aggregation auf dieser übergeordneten Raumebene dazu führt, dass der eigentliche Nachteil der stärkeren Unschärfe bei vollständiger Datenverfügbarkeit (z. B. durch Mittelwertbildung) für eine unvollständige Informationslage dazu führt, dass sich Unsicherheiten hinsichtlich der einzelnen Gebäudemerkmale ausgleichen und somit zwar unscharfe, aber robustere Ergebnisse erarbeitet werden können.⁶⁷

Darüber hinaus ist abschließend festzuhalten, dass die erarbeitete Methodik das Ziel verfolgt, Quartiere ähnlicher Bebauungsstrukturen zu bilden. Aufgrund der realen Bebauungssituationen führt die datenbasierte Quartiersbildung jedoch nur selten zu einer vollständigen Homogenität der gebildeten Quartiere. Dies ist als Unterschied zu den bestehenden,

⁶⁷ Für eine allgemeine Ausführung über die Zusammenhänge und Herausforderungen, die sich aus den Unsicherheiten der Eingangsparameter von Gebäudeenergiebilanzen ergeben, wird auf die beiden Studien Loga et al. (2019; 2021) verwiesen. Generell kann festgehalten werden, dass der Aspekt der Erfassung und Berücksichtigung von Unsicherheiten im Kontext der Gebäudeenergiebilanzierung erst seit Kurzem in den Fokus der Forschung gerückt ist. Die sich für die Quartiersebene darüber hinaus ergebenden Zusammenhänge, im Sinne der Fortpflanzung der gebäudespezifischen Unsicherheiten bei räumlicher Aggregation, sind ebenfalls bisher noch nicht hinreichend untersucht und beschrieben worden.

aus einer Theorie der energetischen Bewertung und Energieversorgung abgeleiteten, Typologien von Siedlungsstrukturen bzw. Stadtraumtypen festzustellen. Es ist weiterhin festzuhalten, dass in dieser Arbeit eine weitergefasste Quartiersabgrenzung, welche auch eine z. B. regelbasierte Verschmelzung von mehreren Quartieren mit innerhalb des Quartiers vergleichsweise homogenen, aber zwischen den Quartieren deutlich unterschiedlichen Bebauungsstrukturen nicht betrachtet wurde. Auch wenn diese als „Mischquartiere“ zu bezeichnenden Konstrukte gerade in der räumlich übergeordneten (Wärme-)Versorgungsplanung von Bedeutung sind, so ließen sich aus der Literaturstudie keinerlei Randbedingungen für einen verallgemeinerungsfähigen Algorithmus zur Bildung bzw. Abgrenzung von Mischquartieren ableiten. Die Entwicklung möglicher Forschungsansätze zur ergänzenden Abgrenzung von Mischquartieren muss an dieser Stelle deshalb als weitere offene, dieser Arbeit nachgelagerte Forschungsfrage formuliert werden.

6 ANWENDUNG DER METHODIK ZUR TYPISIERUNG VON QUARTIEREN AM BEISPIEL BERLIN

6.1 Einleitende Bemerkungen

Nach der in Kapitel 5 erfolgten Entwicklung der Methodik zur GIS-datenbasierten Typisierung von Quartieren anhand baulicher Strukturen, sollen die beschriebenen Verfahrensschritte nun angewendet werden. Für einen beispielhaften Datensatz werden zu diesem Zweck eine Datenvorverarbeitung erfolgen, eine Einteilung von Teilbauflächen in Gruppen in ähnlicher Bebauungsstruktur vorgenommen und aus diesen Quartiere gebildet sowie anschließend für die gebildeten Quartiersgruppen aus einer statistischen Auswertung Typquartiere als Typvertreter abgeleitet.

Für das Anwendungsbeispiel wird ein aus den flächendeckend verfügbaren GIS-Daten (vgl. Kapitel 5.2) auf die Stadt Berlin eingeschränkter Teildatensatz erstellt. Hintergrund der Festlegung auf das Untersuchungsgebiet Berlin ist, neben der Weiterführung einer vor der Bereitstellung der flächendeckenden GIS-Daten bereits durchgeführten Vorstudie, insbesondere die Vielfalt an Bebauungsstrukturen, welche in Berlin vorzufinden ist. Neben hochverdichtet bebauten innerstädtischen Gebieten mit überwiegender Wohn- oder Nicht-Wohnnutzung sind alle der in bestehenden Siedlungsstruktur- bzw. Stadtraumtypologien definierten Typen im Stadtgebiet vorzufinden (vgl. Kapitel 2.4). Die Anwendung der Methodik untergliedert sich dabei in eine vorgelagerte Datenüberprüfung und -auswahl sowie in die vier Verfahrensschritte der Methodik zur GIS-datenbasierten Typisierung von Quartieren:

1. Datenvorverarbeitung
2. Gruppenbildung mittels Clusteranalyse auf Ebene von Teilbauflächen
3. Bildung von Quartieren und Quartiersgruppen aus gruppierten Teilbauflächen
4. statistische Beschreibung des Quartiersbestandes und Festlegung von Kennwerten für Typquartiere

Die Anwendung der Methodik und die gewonnenen Erkenntnisse daraus werden in den nachfolgenden Kapiteln vorgestellt.

6.2 Datenüberprüfung und Datenauswahl

6.2.1 Überprüfung der Gebäudemodellierung und Ansätze zur Datensatzreduzierung

Vor der Verarbeitung der Daten wird eine Überprüfung der vorliegenden Datenbasis, im Hinblick auf die verfügbaren Datenquellen, auf die in der Literaturstudie⁶⁸ identifizierten Herausforderungen sowie auf offene Fragen bezüglich des Nutzens bestimmter Datenquellen, durchgeführt.⁶⁹ Für eine Übersicht der verfügbaren Daten sowie Datenquellen und -formate wird auf Tabelle 5-1 auf S. 80 sowie auf Abbildung 5-2 auf S. 81 verwiesen.

Im ersten Schritt wird überprüft, ob die Identifikation von Gebäudeteilen, welche sich außerhalb der energetischen Gebäudehülle befinden und zudem keine Auswirkungen auf die Bebauungsstruktur eines Quartiers haben, möglich ist. Hintergrund dieser Überprüfung sind die Erkenntnisse aus dem Forschungsvorhaben „EnOB:dataNWG“ des Institut Wohnen und Umwelt (Hörner et al. 2022; Hörner 2021), in dem der Aufbau einer Basis für die repräsentative Stichprobenerhebung des deutschen Nichtwohngebäudebestandes ebenfalls geodatenbasiert erfolgte. Im Vorhaben musste eine besondere Aufmerksamkeit auf die Identifikation von Gebäudevordächern und energetisch irrelevanten Gebäudeanbauten bei der Verarbeitung der als Datenbasis verwendeten Hausumringe für Deutschland (HU-DE) gelegt werden. Die Untersuchung einer Auswahl der LoD2-Gebäudemodelle innerhalb des Berliner Stadtgebietes in Anhang C-4 verdeutlicht, dass ohne eine breite empirische Untersuchung keine Regel abgeleitet werden kann, die eine eindeutige Abgrenzung von energetisch nicht relevanten Gebäudeteilen (z. B. Gebäudevordächer oder zu Wintergärten ausgebaute Terrassen) erlaubt – ohne dabei auch Geoobjekte zu erfassen, welche als relevant angenommen werden (z. B. Anbauten zu Wohn- und Arbeitszwecken). Vor diesem Hintergrund wird entschieden, die Gebäudeteile anhand des gemeinsamen Gebäudeidentifikators („Root-ID“) zu Gebäuden zu vereinigen, ohne dabei bestimmte Gebäudeteile unberücksichtigt zu lassen. Im Falle von als Gebäudeteile erfassten Vordächern führt dies zu einer geringfügigen systematischen Erhöhung der angenommenen Nutzfläche, jedoch zu einem geringen Einfluss auf das betrachtete Gebäudevolumen (siehe Anhang C). Insgesamt kann ein für die Zwecke dieser Studie akzeptabler systematischer Fehler im Hinblick auf energetische und baustrukturelle Parameter angenommen werden. Eine weitere empirische

⁶⁸ Siehe Kapitel 3.5, Kapitel 4.3 sowie Kapitel 4.4.

⁶⁹ Teile der GIS-Datenverarbeitung, z. B. der Import der Daten oder die Änderung von Koordinatenbezugsystemen bzw. eine Koordinatentransformation gehen als Arbeitsschritte der eigentlichen Datenprüfung voraus, werden mit dem Ziel einer nachvollziehbareren Darstellung der Verarbeitungsroutinen jedoch erst im Anschluss an die Erläuterungen zur Datenprüfung näher ausgeführt.

Absicherung wird jedoch empfohlen, wenn die Geodatenbasis z. B. als Grundlage für eine detaillierte Abbildung des Gebäudebestandes in einem Energiesystemmodell verwendet werden soll.

Im Rahmen der Datenüberprüfung wird zudem ein Augenmerk auf die Attribute „Gebäudegrundfläche“ und „Gebäudehöhe“ sowie auf die „Gebäudfunktion“ der Modellabbildungen gelegt. Hartmann et al. (2016) definieren in ihrem Verarbeitungsprozess für Geodaten eine untere Schwelle von 10 m² Grundfläche als Ausschlusskriterium von Gebäudegrundrissen der HU-DE. Diese Grenze wird von den Autoren aus der Landesbauordnung des Landes Mecklenburg-Vorpommern (LBauO M-V) abgeleitet und soll Baukörper abgrenzen, die nicht als Gebäude zu betrachten sind. Hierzu ist festzustellen, dass § 61 der LBauO M-V eine Aufzählung von Bauvorhaben ohne Verfahrenserfordernis beinhaltet, aber die Legaldefinition des Gebäudebegriffs dadurch nicht eingegrenzt wird (vgl. LBauO M-V, vom 15.10.2015). Eine Untersuchung am Beispiel von Berlin zeigt, dass die Festlegung von 10 m² als untere Grenze zu folgenden Ergebnissen führt: Viele Einzelgebäude, welche im Vergleich mit der realen Bebauungssituation eindeutig als nicht funktionale Gebäude im Kontext der Quartierstypisierung erkannt werden können, sind weiterhin Bestandteil des Datensatzes (z. B. Garagen oder Gartenhäuser). Umgekehrt nimmt bei einer Verschiebung der unteren Grenzen zu höheren Mindestgrundflächen die Zahl jener Gebäude zu, bei denen keine sichere Einschätzung über den Nutzen einer Elimination aus dem Datensatz getroffen werden kann (siehe Erläuterungen in Anhang C-4 sowie Abbildung Anhang C-2 und Abbildung Anhang C-3). Weil sich keine klare Festlegung aus der Literatur und der Überprüfung der Daten am Beispiel Berlin ableiten lässt, wird als Grenzwert eine pragmatische, die Funktionalität von Räumen in den Fokus rückende Festlegung von mehr als 16 m² für diese Arbeit getroffen. Unter Berücksichtigung üblicher Wandstärken von (Bestands-)Gebäuden und der Annahme eines Grundrisses mit geringem bis mittlerem Länge-zu-Breite-Verhältnis führt dies zu Nutzflächen von ca. 12 m² bis 15 m² und somit zu einem für verschiedene Anwendungszwecke nutzbaren Raum.⁷⁰

Über den Ansatz von Hartmann et al. (2016) hinaus wurde für die vorliegende Arbeit auch eine untere Grenze der Gebäudehöhe untersucht. Die Ausführungen und Beispiele in Anhang C-4 zeigen, dass zwar kein für alle Bebauungssituationen optimales Kriterium definiert werden konnte, aber dennoch eine Mindesthöhe von mehr als 2,5 m als ein geeignetes Kriterium festgelegt werden kann. Dieser Grenzwert reduziert den Datensatz um wenig

⁷⁰ Diese überschlägige Abschätzung der Nutzfläche beruht auf Wandstärken von 8 cm bis 24 cm je Außenwand sowie einem Länge-zu-Breite-Verhältnis von größer als 0,4.

oder nicht relevante Geoobjekte und geht dabei mit einem nur geringen negativen Einfluss auf die Abbildung der Bebauungs- und Siedlungsstruktur im Datensatz einher.

Hartmann et al. (2016) versuchen neben der Gebäudegrundfläche zudem ein weiteres Kriterium abzuleiten, welches in deren Studie genutzt wird, um untypisch langgestreckte Grundrisse bei räumlichen Analysen zum Gebäudebestand zu eliminieren. Es handelt sich dabei um eine Kombination der Kriterien:

- a) Das Breite-zu-Länge-Verhältnis ist kleiner als 0,3 und
- b) die Grundfläche ist größer als 190 m².

Die Festlegung dieser Grenzwerte lässt sich jedoch in Zweifel ziehen, da die aus Verteilungsfunktionen abgeleiteten Grenzwerte keine Bewertung hinsichtlich der realen Funktionalität betroffener Gebäudegrundrisse erfahren haben. Eine Überprüfung der unter das Kriterium fallenden Gebäude des LoD2-Datensatzes für die Stadt Berlin zeigt, dass die Abgrenzungskriterien auch typische, funktionale Geometrien von Wohngebäuden erfassen (siehe Abbildung Anhang C-4 und Abbildung Anhang C-7). Da bei der vorliegenden raumbezogenen Analyse eine möglichst vollständige Abbildung der in einem Quartier vorhandenen Gebäude erreicht werden soll, wird die Kombination der Kriterien nicht angewendet.

Auch das Attribut der Gebäudefunktion wird dahingehend überprüft, ob es zum Ausschluss von Gebäuden bzw. Gebäudegrundrissepolygonen ohne bzw. mit untergeordneter Relevanz für die hier vorliegende Studie genutzt werden kann, um dadurch den Datensatz einzugrenzen sowie die Typisierung von Quartieren zu vereinfachen. Zu diesem Zweck werden aus der Codeliste des AdV-CityGML-Standards (AdV 2021b) Gebäudefunktionen identifiziert, welche bei der raumbezogenen Analyse und Typisierung von Quartieren unberücksichtigt bleiben sollen, weil sie entweder wenig Relevanz in Bezug auf die Erarbeitung und Bewertung von allgemeinen, notwendigen Maßnahmen zum Klimaschutz im Gebäudebestand besitzen oder für die Gruppierung von Teilbauflächen in Quartiere anhand von baulichen Strukturen als wenig bedeutend betrachtet werden können. Letzteres geht damit einher, dass ein Vorhandensein dieser Gebäude keine gebietsprägende Wirkung hat, diese Gebäude also eher den Charakter von Einzelementen innerhalb einer Siedlungsstruktur besitzen, oder eine räumliche Agglomeration dieser Gebäude zu einer eigenen Raumstruktur führt, welche als Sonderfall zu betrachten ist und damit außerhalb einer Typologie liegt.⁷¹ Auf Basis dieser Überlegungen wird die in Tabelle 6-1 (S. 114) dargestellte Auswahl an

⁷¹ Als ein Beispiel für eine solche Siedlungsstruktur außerhalb der typischen, klimaschutzrelevanten Siedlungs- oder Quartierstypen können Zoos genannt werden, deren Flächenausdehnung und Gebäudeanzahl i. d. R. ausreichend groß sind, um als Quartier wahrgenommen zu werden. Für Analysen zur Dekarbonisierung der Wärmeversorgung des Gebäudebestandes sind diese jedoch von nachrangiger Bedeutung.

Filterkriterien für die Reduktion des Basisdatensatzes der 3D-Gebäudemodelle getroffen und im Rahmen der Datenverarbeitung implementiert.⁷²

Als weiterer Sachverhalt wird der mögliche Nutzen der Georeferenzierten Adressdaten (GA) zur erweiterten Vorverarbeitung der 3D-Gebäudedaten genauer betrachtet. Es gilt zu bewerten, ob die hohe Anzahl an 3D-Gebäudemodellen unter Einbezug der GA durch eine Vernachlässigung einzelner Geobjekte oder durch eine weitere Zusammenführung von Modellgebäuden bzw. deren Grundrissen verringert (z. B. indem Haupt- und Nebengebäude zusammengefasst werden) und gleichzeitig eine bessere Abbildung der Bebauungsstrukturen erreicht werden kann. Die Grundrisse der 3D-Gebäudemodelle wurden hierzu auf ihre räumliche Beziehung zu den GA hin überprüft. Eine beispielhafte Darstellung für eine innerstädtische Bebauungssituation (hier Berlin Kreuzberg) liefert Abbildung 6-1 (S. 115). Im linken Teil der Abbildung sind die Gebäude mit und ohne Adresspunkt im Gebäudegrundrisspolygon dargestellt. Dem gegenübergestellt wird im rechten Teil der Abbildung eine Visualisierung der realen Bebauungssituation mit Blockrand- und Blockinnenbauten überwiegender Wohnnutzung.

Der Vergleich zeigt, dass nur ein Teil der Gebäude im 3D-Gebäudedatensatz in komplexen Bebauungssituationen einen eigenen Adresspunkt besitzt. Eine Einschränkung der Geobjekte ausschließlich auf Grundrisse mit zugeordnetem Adresspunkt ist damit nicht zielführend für eine Reduktion des 3D-Gebäudemodelldatensatzes. Auch für eine weiterführende Vereinigung von Gebäuden ohne und mit Adresspunkt ist die notwendige Datenbasis nicht gegeben. Grundstückspolygone, welche als sinnvolles Zuordnungskriterium genutzt werden könnten, sind in keiner der flächendeckend verfügbaren Datenquellen der Geodatenbasis enthalten. Eine automatisierte Zuordnung von Gebäuden ohne Adresspunkt zu Gebäuden mit Adresspunkt auf Basis von anderen einfachen Zuordnungsregeln, z. B. anhand der längsten gemeinsamen Gebäudelinie, ist zwar einfach durchführbar, liefert jedoch für die Vielzahl an Bebauungssituationen häufig Ergebnisse, die sich nicht mit der wahrgenommenen Bebauungssituation decken. Die Geobjekte der 3D-Gebäudemodelle werden deshalb nicht weiter vereinigt oder auf Grundrisse mit vorliegendem Adresspunkt eingegrenzt.

⁷² Einige der aufgeführten Funktionsklassen werden z. B. auch im Umweltatlas der Stadt Berlin zur Abgrenzung von Nebengebäuden bei der Auswertung von städtebaulichen Kennzahlen verwendet (vgl. Gerstenberg 2019). Für weitere Ausführungen zur Gebäudefunktionsklasse als Abgrenzungskriterium siehe Anhang C.

6 Anwendung der Methodik zur Typisierung von Quartieren am Beispiel Berlin

Durch die in Tabelle 6-1 aufgeführten Kriterien wird eine Vielzahl von Kleinstgrundrissen aus dem Datensatz entfernt. Die betroffenen Objekte werden dabei hinsichtlich der Geometrie oder der Gebäudefunktion als von geringer oder fehlender Relevanz für die Zielsetzung dieser Arbeit bewertet. Die im Datensatz verbliebenen Objekte besitzen hingegen eine relevante Geometrie und Gebäudefunktion (z. B. Wohnnutzung). Bei Betrachtung der realen Bebauungssituationen erscheint die Richtigkeit der im Datensatz angegebenen Funktionszuordnung z. B. zur Funktionsklasse „Wohnen“ in den Geobasisdaten jedoch in einigen Fällen als fraglich (z. B. wenn Gartenhäuser vorliegen). Auch diesbezüglich ergibt sich somit ein Bedarf für empirische Absicherungen über die vorliegende Arbeit hinaus.

Tabelle 6-1: Filterkriterien zum Ausschluss von Objekten der 3D-Gebäudemodelldaten und Auswertung der Anzahl betroffener Objekte für Berlin (Gesamtanzahl Gebäudeobjekte = 640.397)

Filterkriterium	Beschreibung	Anzahl (Anteil) im LoD2-Datensatz für Berlin
Gebäudefunktion		
5100x	Andere (Infrastruktur-)Bauwerke	106.405 (16,6 %)
31001_1312	Wochenendhaus	13.728 (2,1 %)
31001_1313	Gartenhaus	4.152 (0,6 %)
31001_2074	Campinggebäude	69 (0,01 %)
31001_24xx	Gebäude für Verkehrszwecke (inkl. Garage, Parkhaus etc.)	83.665 (13,1 %)
31001_25xx	Gebäude zur Versorgung	5.488 (0,9 %)
31001_26xx	Gebäude zur Entsorgung	1.539 (0,2 %)
31001_27xx	Gebäude für Landwirtschaft (z. B. Schuppen)	25.040 (3,9 %)
31001_304x	Gebäude für religiöse Zwecke	1.387 (0,2 %)
31001_307x	Gebäude für Sicherheit und Ordnung	903 (0,14 %)
31001_308x	Friedhofsgebäude	280 (0,04 %)
31001_309x	Empfangsgebäude (Bahnhof, Flughafen etc.)	493 (0,08 %)
31001_326x	Gebäude im Zoo	333 (0,05 %)
31001_327x	Gebäude in botanischen Gärten	122 (0,02 %)
31001_3281	Schutzhütte (Touristik)	19 (< 0,01 %)
Summe Gebäudefunktion		243.604 (38,0 %)
Gebäudegrundfläche	≤ 16 m ²	132.436 (20,7 %)
Gebäudehöhe	≤ 2,5 m	86.397 (13,5 %)
Kombinierte Anwendung der Filterkriterien (Geobjekte können von mehreren Kriterien betroffen sein.)		287.673 (44,9 %)



Abbildung 6-1: Beispielhafter Datenausschnitt aus Berlin Kreuzberg nach der Überprüfung auf das Vorhandensein von GA innerhalb der Gebäudegrundrisse (links: grün - Gebäude hat Adresspunkt, violette - Gebäude hat keinen Adresspunkt; rot - Adresspunkt) und Vergleich mit realer Bebauungssituation (rechts)

[links: eigene Darstellung; Daten – GeoBasis-DE / BKG (2021) bzw. BKG, Deutsche Post Direkt GmbH, Statistisches Bundesamt, Wiesbaden (2021); rechts: Bilder – Google (2022); Kartendaten – GeoBasis-DE / BKG (2022)]

6.2.2 Verarbeitung der zur Verfügung stehenden GIS-Daten

Die Anwendung der voranstehend beschriebenen Ausschlusskriterien schließt sich im Zuge der Datenvorverarbeitung an den Import der GIS-Daten an. Das Vorgehen der Datenverarbeitung vor der eigentlichen Clusteranalyse und Klassifizierung der Teilbauflächen sowie der Bildung der Quartiere zum Zweck der Typisierung, wird in Abbildung 6-2 (S. 116) und Abbildung 6-3 (S. 117) dargestellt. Nach der Erstellung einer 3DCityDB-Instanz in der PostGIS-Datenbank werden die GIS-Daten in Abhängigkeit ihres Formats in das allgemeine Datenbankschema *public* oder das zur Abbildung von 3D-CityGML-Daten angelegte Schema *citydb* importiert. Hierfür wird ein Skript in R erstellt, welches die nötigen Importroutinen über die Windows Kommandozeile oder über eine direkte Datenbankschnittstelle anstößt und auch – falls erforderlich – eine Koordinatentransformation für die 3D-Gebäudemodelldaten durchführt.⁷³ Der Import der einzelnen Datenbestände ist dabei nicht an eine festgelegte Reihenfolge gebunden.

⁷³ Die 3D-Gebäudemodelle werden in Teildatensätzen je Bundesland bereitgestellt. Das Koordinatenbezugssystem (KBS) entspricht in Abhängigkeit der geografischen Lage des Bundeslandes entweder EPSG:25832 oder EPSG:25833. Für eine einheitliche Abbildung der Geoobjekte in der PostGIS-Datenbank werden die Geoobjekte in EPSG:28533 vor dem Import in die Hauptdatenbank mit dem 3DCityDB-Importer-Exporter transformiert, sodass die Abbildung dem KBS EPSG:25832 entspricht.

Die im CSV-Format vorliegenden GA werden über die Funktionen des verfügbaren R-Packages *sf* eingelesen und in die PostGIS-Datenbank übertragen. Nach dem vollständigen Import der 3D-Gebäudemodelle werden über das R-Skript zudem PostgreSQL-Anweisungen ausgeführt, welche die Geometrien der Gebäudeteile des 3D-Gebäudedatensatzes auf das Grundrisspolygon reduzieren und anhand des gemeinsamen Gebäudeidentifikators (Root-ID) vereinigen. Im Zuge dieser Vereinigung müssen einige der Objektmerkmale der Gebäudeteile für die Objektebene des Gebäudes aggregiert werden. Während die Flächen der Grundrisspolygone aller Gebäudeteile eines Gebäudes addiert werden, wird die Gebäudehöhe als flächengewichteter Mittelwert der Gebäudeteile auf das Gebäudeobjekt übertragen. Die in Form einer kategorialen Variable verfügbare Information zur Dachform eines Gebäudeteils wird bei der Vereinigung der Grundrisse in vier stetige, intervallskalierte Variablen transformiert. Jede dieser Variablen gibt den Anteil einer Dachformklasse nach Tabelle 6-2 (S. 117) an (vgl. Bayerische Vermessungsverwaltung 2018). Die anderen für die Untersuchung relevanten Merkmale der Gebäudeteile, z. B. die Gebäudefunktion, müssen nicht aggregiert werden, da sie von der Gebäudeebene an die Gebäudeteile vererbt wurden und somit für alle Gebäudeteile einer Root-ID identisch sind.

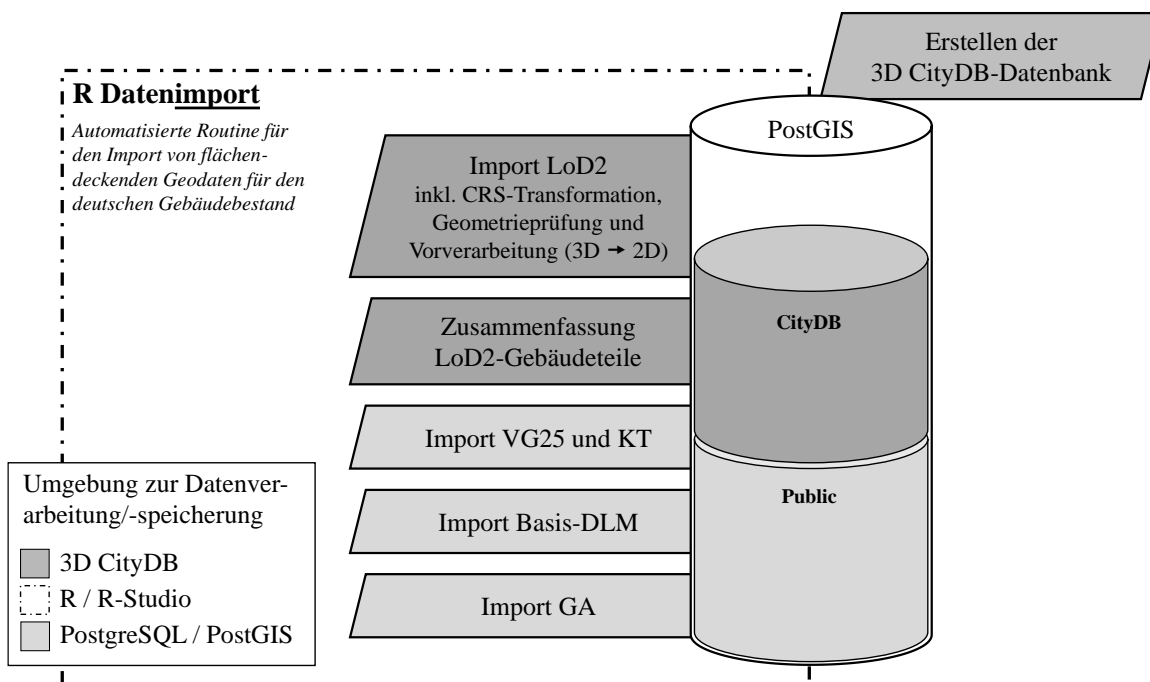


Abbildung 6-2: Schema des Geodatenimports und der Verarbeitung der LoD2-Gebäudedaten

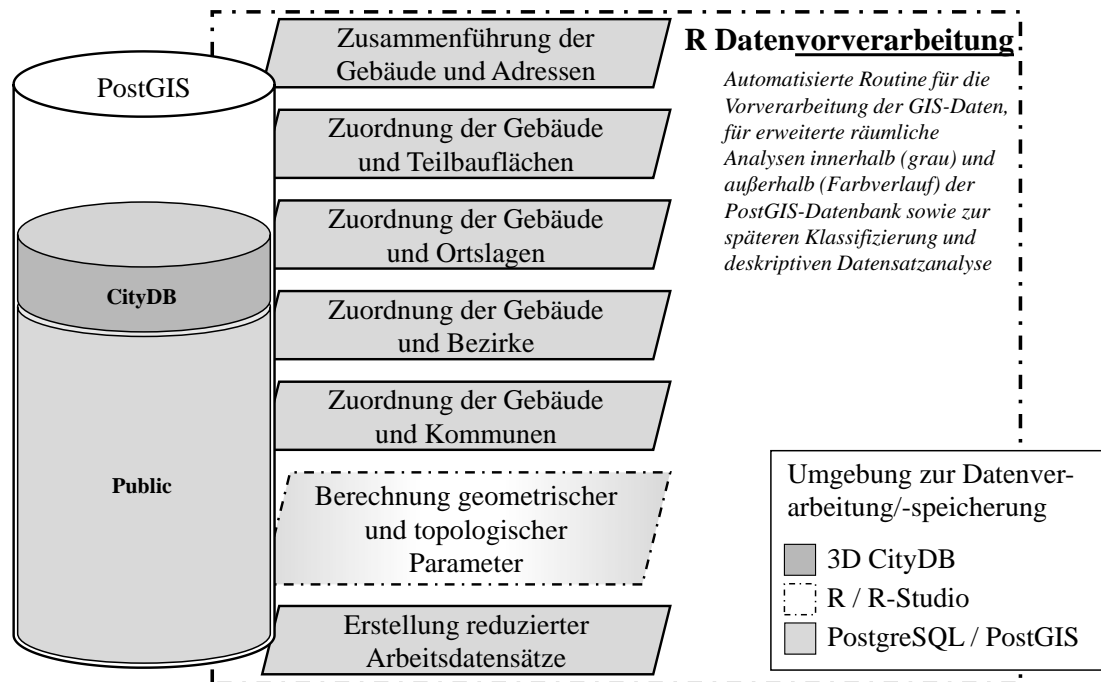


Abbildung 6-3: Schema der Geodatenverarbeitung zur Erstellung eines Basisdatensatzes und der Extraktion eines Arbeitsdatensatzes

Tabelle 6-2: Modellierung des Attributs „Dachform“ für Gebäudeteile und vereinigte Gebäudeobjekte

Modellebene	Abbildung der Dachformen			
Gebäudeteil	<ul style="list-style-type: none"> • Flachdach 	<ul style="list-style-type: none"> • Pultdach • versetztes Pultdach • Satteldach • Walmdach • Krüppelwalmdach • Mansardendach • Zeltdach 	<ul style="list-style-type: none"> • Mischform 	<ul style="list-style-type: none"> • Kegeldach • Kuppeldach • Bogendach • Turmdach • Sonstiges
Gebäude	Anteil mit Flachdach	Anteil mit Standarddachform	Anteil mit Mischform	Anteil mit sonstiger Dachform

Im Anschluss an die Vorverarbeitung der einzelnen Geobasisdaten erfolgt die Erstellung eines Datensatzes, in dem die Attribute der verschiedenen Datensätze auf die Gebäude als Merkmalsträger der untersten räumlichen Ebene übertragen werden. Dies erfolgt über die Abfrage der räumlichen Beziehungen der Geoobjekte innerhalb der PostGIS-Datenbank. Der auf diese Weise erstellte und mit Metainformationen angereicherte Gebäudedatensatz wird zur Beschreibung der Bebauungsstruktur auf Ebene von Teilbauflächen verwendet und dient somit als Grundlage für die weiteren Arbeitsschritte der Methodik. Für die in den

Kapiteln 5.3 bis 5.6 erläuterten Verfahrensschritte werden ausschließlich Teilbauflächen verwendet, welche nach Ausführung der Abgrenzungskriterien mit Gebäuden bebaut sind.

6.2.3 Auswahl der für die Methodenentwicklung verwendeten Datenbasis

Der verwendete Arbeitsdatensatz und der Umfang der dafür verwendeten Geobasisdaten sind in Tabelle 6-3 beschrieben. Es handelt sich dabei ausschließlich um vom BKG bereitgestellte flächendeckend verfügbare Datenquellen, welche für dieses Anwendungsbeispiel räumlich eingeschränkt wurden. Eine Übersicht der im Datensatz vorhandenen bzw. im Rahmen der Datenaufbereitung berechneten Gebäudemerkmale zeigt Tabelle 6-4.

Tabelle 6-3: Gliederung des für die Methodenentwicklung und für eine beispielhafte Anwendung der Quartierstypisierung erstellten Arbeitsdatensatzes

Räumliche Eingrenzung	Berlin
Beschreibung des Datenumfangs	Vorverarbeiteter Datensatz aus den Inhalten der Geobasisdaten: 3D-Gebäudemodell (LoD2), Basis-DLM (Ortslagen, Teilbauflächen, Straßennetz), Kommunale Teilgebiete (KT25), Verwaltungsgebiete (VG25) sowie Georeferenzierte Adressdaten (GA)
Abgrenzungskriterien	Gebäudeobjekte nach Anwendung der Abgrenzungskriterien (Grundfläche, Gebäudehöhe, Gebäudefunktion); Teilbauflächen, nur mit Gebäudepopulation ⁷⁴

Der gewählte Datensatz für Berlin zeigt sich als gut geeignet, um die Analyse von Gebäuden und Bebauungsstrukturen auf Ebene von Teilbauflächen zu ermöglichen und ist gleichzeitig informationstechnisch handhabbar. Tabelle 6-5 gibt die Anzahl der im Datensatz enthaltenen Teilbauflächen, Gebäuderegionen und Gebäude an. In den nachfolgenden Abschnitten werden die Anwendung der Methodik und die dabei durchgeführten Analysen sowie Festlegungen genauer beschrieben.

⁷⁴ Der Begriff der Gebäudepopulation wird in dieser Arbeit eingeführt, um eine Verwechslung von lokalen, kleinräumigen Gebäudebeständen mit Gebäudebeständen höherer Verwaltungsebenen, z. B. Städten, Bundesländern oder dem Bund, zu vermeiden.

Tabelle 6-4: Gebäudemerkmale des ausgewählten Datensatzes für die Beschreibung der Bebauungsstruktur auf Ebene von Teilbauflächen

Merkmale der Gebäude	
Gebäudefunktionsklasse ... 1: Wohnen ... 2: Wirtschaft und Gewerbe ... 3: öffentliche Zwecke ... 4: nicht spezifiziert	Grundfläche [m ²]
	Umfang des Gebäudegrundrisses [m]
	Mittlere Gebäudehöhe [m]
	Verhältnis von Länge zu Breite [-]
Anteil Dachfläche [%] ... Flachdach ... Standarddach ... Mischform ... sonstige Dachform	Verhältnis von Länge zu äquivalenter Breite [-]
	Kompaktheitsindex [-]
	Roeck-Wert [-]

Tabelle 6-5: Anzahl im Datensatz für die Methodenanwendung enthaltener Geobjekte und Gebäudeanzahl je Gebäudefunktionsklasse

Geobjekte		Anzahl	(Anteil)
Teilbauflächen		24.729	
Gebäuderegionen		211.459	
Gebäude		352.720	
Gebäudefunktionsklasse	1: Wohnen	290.206	(82,3 %)
	2: Wirtschaft und Gewerbe	52.740	(15,0 %)
	3: öffentliche Zwecke	9.331	(2,6 %)
	4: nicht spezifiziert	443	(0,1 %)

6.3 Datenvorverarbeitung des Datensatzes Berlin

Auf die Datenauswahl folgt die Datenvorverarbeitung als erster Verfahrensschritt gemäß dem in Abbildung 5-1 (S. 78) gezeigten Vorgehens zur GIS-datenbasierten Typisierung von Quartieren aus Teilbauflächen. Für den Beispieldatensatz der 3D-Gebäudemodelle in Berlin werden deshalb entsprechende Programmroutinen ausgeführt, welche die in Tabelle 5-3 (S. 88) dargestellten Variablen zur Beschreibung der Bebauungsstruktur auf Ebene von Teilbauflächen aus dem Gebäudedatensatz und einer raumbezogenen Abfrage bezüglich der Zugehörigkeit der Gebäude zu einer Teilbaufläche berechnen. Die statistischen Kennzahlen der Bebauungsstrukturvariablen der im Datensatz enthaltenen Teilbauflächen zeigt die Übersicht in Tabelle 6-6 (S. 120). In diesem Zusammenhang wird angemerkt, dass sich die Zeilenbezeichnungen „MW“ und „SD“ der Tabelle auf die Merkmals-träger „Gebäude“, „Gebäuderegion“ bzw. „Teilbaufläche“ (nur für Gebäudeanzahl, MBN I

6 Anwendung der Methodik zur Typisierung von Quartieren am Beispiel Berlin

und MBN II) innerhalb der jeweiligen Teilbauflächen beziehen. Die Spaltenbezeichnungen „MW“ und „SD“ zeigen hingegen die Auswertung der Mittelwerte bzw. Standardabweichungen der Gesamtheit der Teilbauflächen.

Tabelle 6-6: Statistische Kennzahlen der bebauungsstrukturbeschreibenden Variablen auf Ebene der Teilbauflächen (N = 24.729)

Merkmal	MW	SD	Min.	p_{0,25} p_{0,50} p_{0,75}	Max.
MW der Gebäudegrundfläche [m ²]	503,4	971,2	16,0	92,9 248,2 551,2	35.706,5
SD der Gebäudegrundfläche [m ²]	238,9	595,5	-	31,3 79,9 216,0	18.868,0
MW der mittleren Gebäudehöhe [m]	12,0	6,8	2,5	7,1 9,7 16,5	78,6
SD der mittleren Gebäudehöhe [m]	0,8	1,0	-	0,2 0,5 0,9	22,5
Spez. Anzahl Gebäude [1/ha]	9,3	7,5	0,0	3,1 7,8 14,3	105,4
Spez. Anzahl Gebäuderegionen [1/ha]	5,3	4,7	0,0	1,6 3,4 8,5	42,9
MW der Regionen- grundfläche [m ²]	2.415,0	4.908,8	16,0	169,4 595,4 2.485,3	219.541,5
SD der Regionen- grundfläche [m ²]	547,2	1.702,9	-	- 72,2 380,7	109.736,9
MW der Gebäude- anzahl je Region [-]	9,7	14,5	1,0	1,0 3,0 12,0	154,0
MBN I [m ² /m ²]	0,21	0,13	-	0,13 0,18 0,29	1,26
MBN II [m ³ /m ²]	0,89	1,91	-	0,05 0,24 0,84	27,30
Legende: MW = Mittelwert, SD = empirische Standardabweichung (der Grundgesamtheit), Min. = Minimum, p _{0,25} = Unteres Quartil, p _{0,50} = Median, p _{0,75} = Oberes Quartil, Max. = Maximum, MBN I = Maß baulicher Nutzung I, MBN II = Maß baulicher Nutzung II					

Tabelle 6-6 wird zudem ergänzt durch Abbildung 6-4 (S. 122), in welcher ein Histogramm der relativen Häufigkeit und die Dichtefunktion für die Merkmalsausprägungen der „mittleren Gebäudegrundfläche“ sowie die Dichtefunktion für die Merkmalsausprägungen der Variable „flächengewichtete mittlere Gebäudehöhe“ der Teilbauflächen des Datensatzes Berlin dargestellt werden. Hierbei ist zu beachten, dass die Dichtefunktionen eine Approximation der empirischen, nicht stetigen Merkmalsverteilungen der Grundgesamtheit der Teilbauflächen ($N = 24.729$) darstellen und sich in der Anwendbarkeit von den Wahrscheinlichkeitsdichtefunktionen stetig verteilter Zufallsvariablen unterscheiden. Wegen der im Vergleich zu Histogrammen besseren Lesbarkeit wird auch in den folgenden Teilen der Arbeit die Darstellung der Dichtefunktion bevorzugt.

Aus den Daten und Abbildungen ist die Bandbreite der Bebauungsstrukturen innerhalb der Teilbauflächen erkennbar. Durch die nachfolgend durchgeführte Clusteranalyse zur Gruppenbildung werden die Bandbreite der Bebauungsstruktur auf Ebene von Teilbauflächen analysiert, die Unähnlichkeit der Teilbauflächen in Form einer Distanzmatrix beschrieben und eine Einteilung der Teilbauflächen in Gruppen ähnlicher Bebauungsstruktur mittels PAM-Algorithmus vorgenommen.

Wie aus den gezeigten Dichteverläufen erkennbar, handelt es sich bei einigen der Variablen nicht um symmetrische, einer Normalverteilung ähnliche Merkmalsverteilungen. Eine vollständige Beschreibung der Verteilung der Merkmalsausprägungen erfordert deshalb neben dem Mittelwert und der Standardabweichung auch weitere Streuungsmaße, die sog. höheren Momente „Schiefe“ und „Kurtosis“ (auch „Wölbung“ genannt). Die Schiefe kann dabei als Maß der Asymmetrie betrachtet werden. Der Wert der Kurtosis beschreibt hingegen, ob eine Dichtefunktion eine eher spitze oder plateauartige Ausprägung besitzt. Diese Werte sind in Ergänzung zu den Angaben in Tabelle 6-6 und gemeinsam mit den Dichtefunktionen weiterer Merkmale in Anhang D-2 gegeben. Auch der empirische Variationskoeffizient als relative Formulierung der Streuung um den Mittelwert der Merkmalsverteilung wird dort zum Zweck der besseren Vergleichbarkeit der Varianz der betrachteten Variablen angegeben.

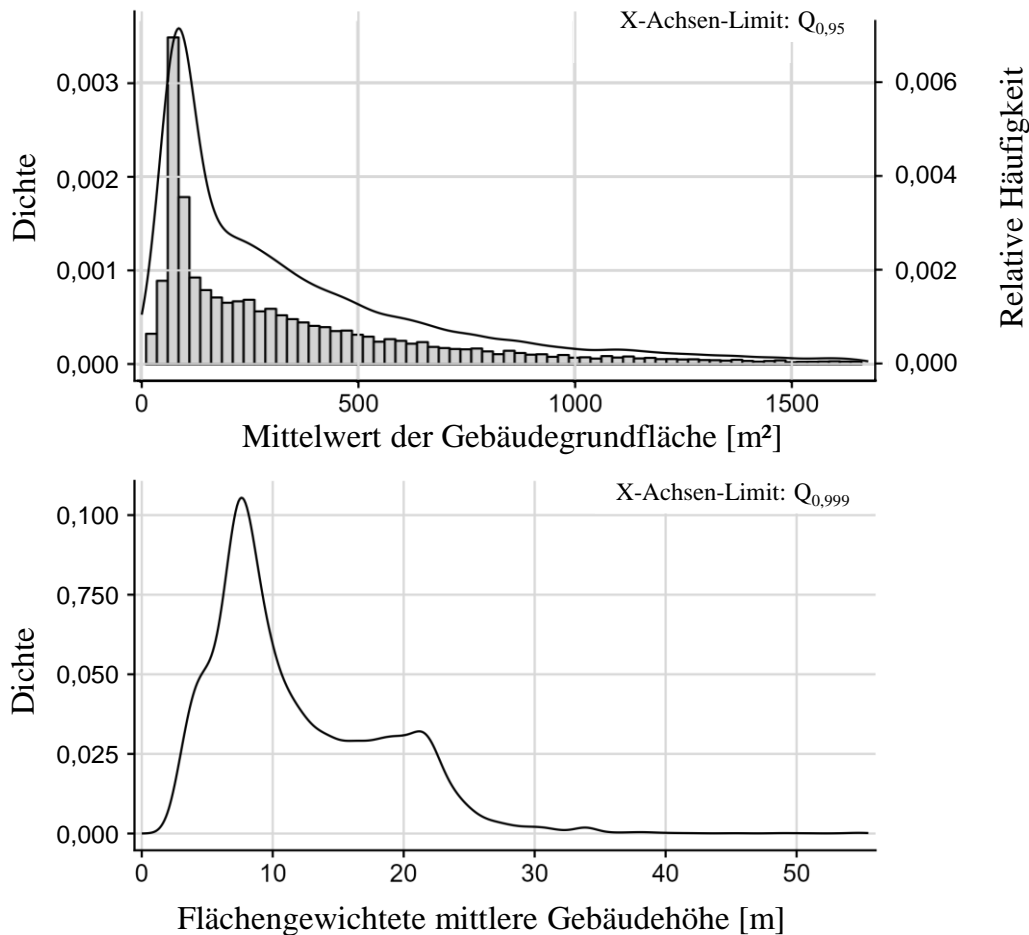


Abbildung 6-4: Dichtefunktionen der Bebauungsstrukturvariablen „Mittelwert der Gebäudegrundfläche“ und „Flächengewichtete mittlere Gebäudehöhe“ je Teilbaufläche (Darstellung entlang X-Achse abgeschnitten bei entsprechend angegebenem Quantilwert)

6.4 Clusteranalyse auf Ebene von Teilbauflächen und Überprüfung der Gruppenbildung

6.4.1 Ausführung der Clusteranalyse zur Gruppenbildung

Der im Rahmen der Datenvorverarbeitung erstellte Datensatz der Teilbauflächen mit den Bebauungsstrukturvariablen wird im Folgenden für die Clusteranalyse verwendet. Wie bereits in Abbildung 5-3 (S. 91) sowie in Kapitel 5.4.2 (S. 93) dargelegt, wird für die Clusteranalyse unter Verwendung der Gower-Metrik zuerst eine Distanzmatrix erstellt. Diese bildet in diesem Anwendungsbeispiel für Berlin die Distanzen der 24.729 Teilbauflächen des Datensatzes ab. Die Werte der Distanzmatrix beschreiben also die Unähnlichkeit der Teilbauflächen untereinander, wobei die Grundlage der Distanzberechnung die elf in Tabelle 5-3 (S. 88) bzw. Tabelle 6-6 (S. 120) gezeigten energetisch-städtebaulich relevanten Merkmale der Bebauungsstruktur sind, welche mittels Gower-Metrik gleichgewichtet in ein Unähnlichkeitsmaß übertragen werden. Der Distanzberechnung schließt sich die

Ausführung des PAM-Algorithmus an, welcher für eine iterativ vorgegebene Gruppenanzahl k eine Gruppeneinteilung vornimmt. Diese Einteilung minimiert dabei die Distanz der Objekte des Datensatzes zu den jeweiligen Medoiden der Gruppen.

Für den Teilbauflächendatensatz konnte die Gruppenbildung an den Grenzen der informationstechnischen Ressourcen, unter Inkaufnahme einer sehr hohen Ausführungsdauer, noch für alle Objekte durchgeführt werden. Das für die Beurteilung der Gruppeneinteilung festgelegte Gütemaß der Silhouetten-Weite bzw. mittleren Silhouetten-Weite (ASW) wird in Abbildung 6-5 (S. 124) für die einzelnen Gruppen (oben) bzw. die Gruppeneinteilung insgesamt (unten) dargestellt. Während bei einer Gruppenanzahl von drei für diesen Datensatz ein scheinbar geeignetes Ergebnis vorliegt, wird eine solche Festlegung der Gruppenanzahl jedoch aufgrund der fehlenden praktischen Relevanz nicht getroffen. Eine Unterteilung der Teilbauflächen und der darin befindlichen Bebauungsstrukturen ermöglicht keine Untersuchungen, die im Hinblick auf eine Bewertung von energetischen Quartierskonzepten sinnvolle Ergebnisse versprechen würden und weicht zudem stark von den bisher etablierten Siedlungsstruktur- und Stadtraumtypen ab. Es wird deshalb entschieden, das folgende lokale Maximum der ASW ($k_{opt} = 6$) als geeignet zu betrachten. Für die weiteren Schritte der beispielhaften Anwendung der Methodik zur GIS-datenbasierten Quartierstypisierung wird eine geeignete Gruppenanzahl von $k_{opt} = 6$ festgelegt.

Eine ergänzende Klassifizierung mittels Naive Bayes Klassifizierers ist in dieser beispielhaften Anwendung nicht nötig, da der PAM-Algorithmus auf alle Teilbauflächen angewendet wurde und somit die Grundgesamtheit der Teilbauflächen vollständig gruppiert wurde. Die Bildung von Quartieren aus den in Gruppen eingeteilten Teilbauflächen kann damit entsprechend den Beschreibungen in Kapitel 5.5 vorgenommen werden.

Vor der Ausführung der Quartiersbildung soll jedoch eine Auswahl der Merkmale der Bebauungsstruktur für die Gesamtheit der Teilbauflächen zu informativen Zwecken noch einmal in einer gruppenbezogenen Darstellung erarbeitet werden. Hierfür werden die statistischen Kenngrößen der Bebauungsstruktur für jede der durch eine Gruppe aufgespannten Teilmengen der Teilbauflächen ermittelt und ebenfalls tabellarisch dargestellt (Tabelle 6-7, S. 125).

Es zeigt sich insgesamt eine deutliche Differenzierung hinsichtlich der mittleren Gebäudegrundflächen und Gebäudehöhen der Gebäudepopulationen der Teilbauflächen sowie der anderen Bebauungsstrukturvariablen. Gleichzeitig lässt sich aus den statistischen Kennzahlen (Mittelwert, Standardabweichung sowie Schiefe und Kurtosis) ableiten, dass die zugrundeliegenden Verteilungen der Merkmalsausprägungen weiterhin eine hohe Varianz aufweisen und nicht normalverteilt sind. Ergänzend zur tabellarischen Darstellung finden sich in Anhang D-2 auch für die Teilmengen der Teilbauflächen je Teilbauflächengruppe ergänzende statistische Kennzahlen und Dichtefunktionen der Merkmale. Eine darüber hinausgehende Beschreibung der Gruppen erfolgt in Kapitel 6.4.3 im Rahmen eines Vergleiches der Einteilung mit ausgewählten Siedlungsgebieten in Berlin.

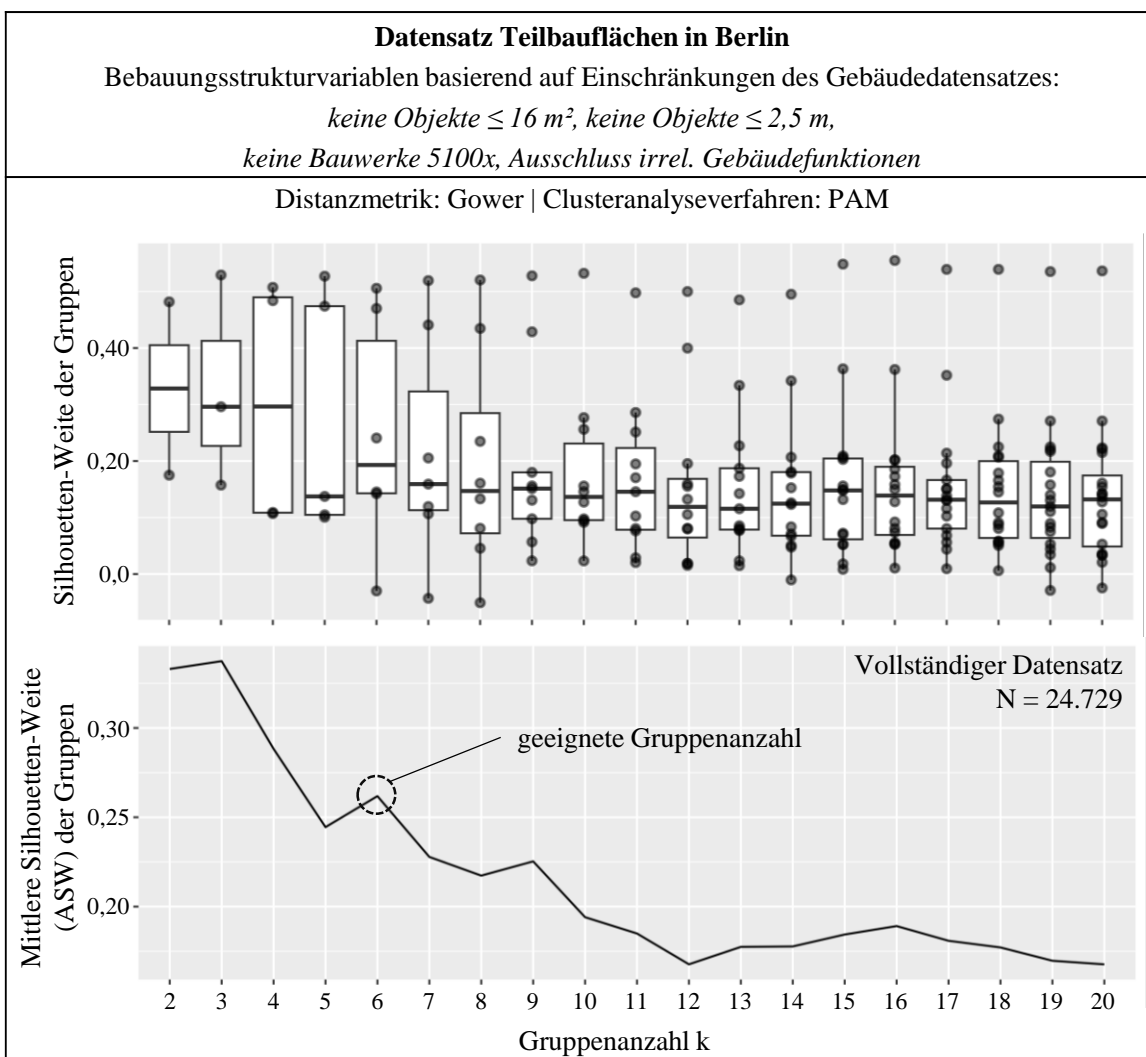


Abbildung 6-5: Festlegung einer geeigneten Gruppenanzahl k_{opt} aus der iterativen Gruppenbildung im Intervall $k \in [2, \dots, 20]$ für den vollständigen Teilbauflächendatensatz
 Oben: Silhouetten-Weite der einzelnen Gruppen je vorgegebener Gruppenanzahl k
 Unten: Mittlere Silhouetten-Weite (ASW) der Gruppeneinteilung

6 Anwendung der Methodik zur Typisierung von Quartieren am Beispiel Berlin

Tabelle 6-7: Statistische Kennzahlen der bebauungsstrukturbeschreibenden Variablen für die festgelegten Teilbauflächengruppen (Unterstreichung = Eingangsgröße für die Clusteranalyse)

Merkmal der Bebauungsstruktur	Teilbauflächengruppen					
	1	2	3	4	5	6
Anzahl Gebäude	195.679	90.778	49.516	63.608	4.310	12.052
Anzahl Teilbauflächen	5.335	4.471	5.895	3.782	2.032	3.214
Summe der Gebäude- grundfläche [Mio. m ²]	16,1	12,1	31,0	21,4	7,3	3,0
<u>MW der mittleren</u> <u>Gebäudegrundfläche</u> [m ²]	267,9	814,7	1.993,0	395,2	174,6	86,4
<u>SD der mittleren</u> <u>Gebäudegrundfläche</u> [m ²]	106,8	485,7	494,6	255,9	106,2	47,6
<u>MW der mittleren</u> <u>Gebäudehöhe</u> [m]	6,1	13,6	21,1	19,2	9,0	7,6
<u>SD der mittleren</u> <u>Gebäudehöhe</u> [m]	0,4	1,0	0,9	1,0	0,7	0,5
MW mittlere <u>Gebäude-</u> <u>anzahl je Gebäuderegion</u>	2,5	4,1	11,0	25,0	3,7	2,4
SD mittlere <u>Gebäudeanzahl</u> <u>je Gebäuderegion</u>	0,5	1,5	0,4	4,2	1,6	1,3
MW <u>spezifische Anzahl</u> <u>Gebäude</u> [1/ha]	2,3	3,9	3,8	11,7	12,0	17,8
MW des <u>MBN I</u> [m ² /m ²]	0,05	0,21	0,41	0,37	0,18	0,15
MW des <u>MBN II</u> [m ³ /m ²]	0,17	0,86	5,60	0,82	0,32	0,09

MW = Mittelwert; SD = empirische Standardabweichung; MBN = Maß baulicher Nutzung

6.4.2 Vergleich der Varianten zur Festlegung einer geeigneten Gruppenanzahl

Neben der Auswahl der im Folgenden weiterverwendeten, geeigneten Gruppenanzahl k_{opt} , als Ergebnis der Gruppenbildung für den vollständigen Datensatz der Teilbauflächen, wurde im Hinblick auf die Sicherstellung der Übertragbarkeit auf größere Datensätze, bis hin zur Erfassung des gesamtdeutschen Teilbauflächenbestandes, auch das Vorgehen der iterativen Gruppenbildung für unterschiedliche Zufallsstichproben aus der Gesamtheit der Teilbauflächen ausgeführt (vgl. Abbildung 5-5, S. 94). Aus dem Teilbauflächendatensatz wurden für diese weiterführende Untersuchung deshalb zehn Zufallsstichproben à 11.000 Teilbauflächenobjekte gezogen und als reduzierte Eingangsdatensätze für die Gruppenbildung verwendet. Aus den ermittelten Verläufen der Werte der ASW für die zehn Stichproben wurden anschließend die lokalen Maxima ausgezählt und der Modus dieser Maxima als geeignete Gruppenanzahl identifiziert (siehe Abbildung 6-6, S. 126, und Tabelle 6-8, S. 127).

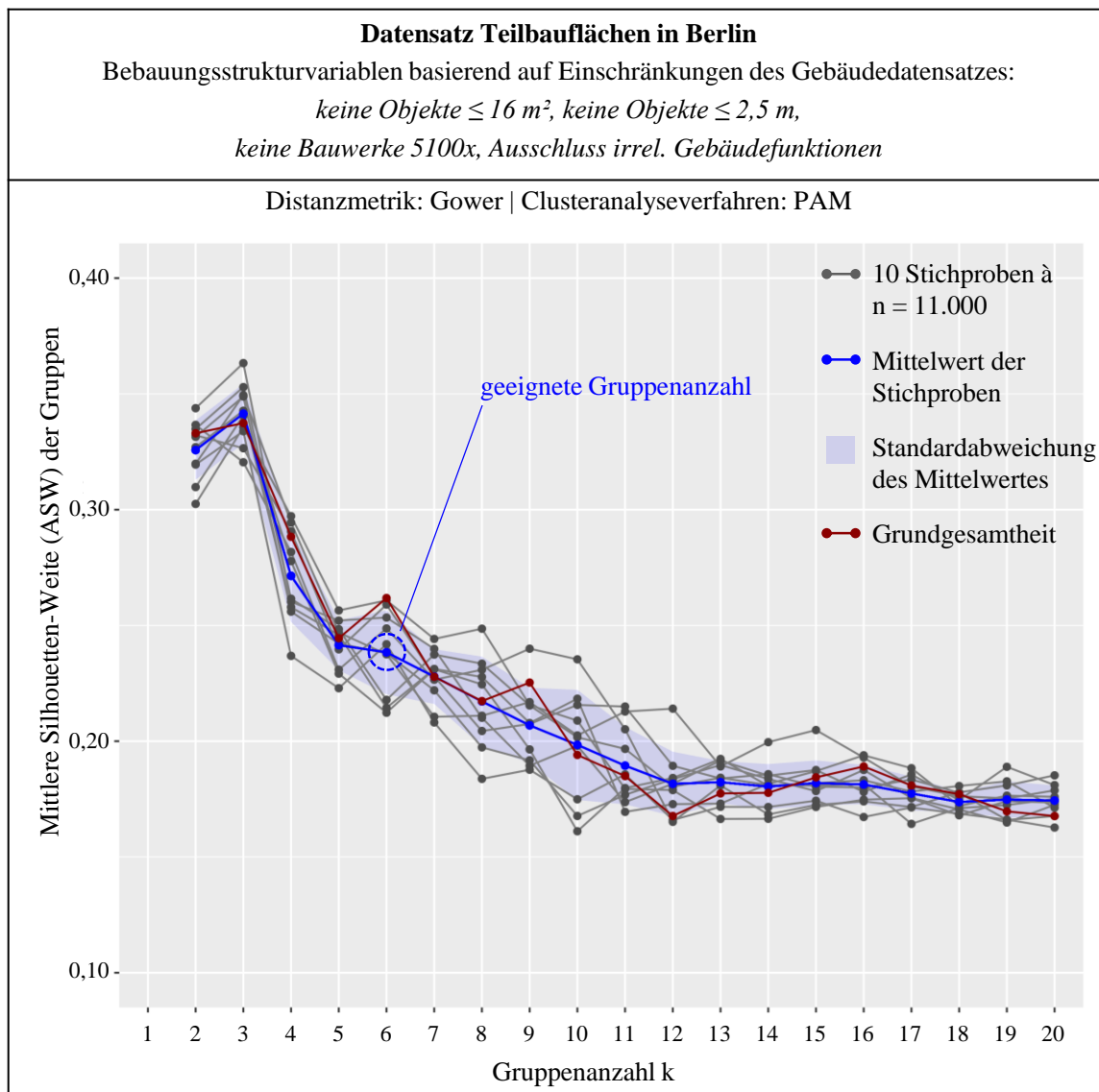


Abbildung 6-6: Ergebnisse der iterativen Berechnung der mittleren Silhouetten-Weite (ASW) und Auswahl der weiter zu berücksichtigenden Gruppenanzahl k_{opt}

Legende: ASW-Graphen der zehn Gebäudestichproben (grau); über alle Stichproben gemittelte ASW (blau); Standardabweichung vom Mittelwert der über alle Stichproben gemittelten ASW (hellblau)

Es zeigt sich bei der auf Stichproben basierenden Festlegung der geeigneten Gruppenanzahl eine Varianz der Ergebnisse zwischen den Stichproben. Der Verlauf des Mittelwertes der ASW aller Stichproben lässt vermuten, dass neben der Festlegung auf eine geeignete Gruppenanzahl $k_{opt} = 6$ auch die Gruppenanzahl $k_{opt} = 5$ oder 7 hätte gewählt werden können, weil ein klares lokales Maximum bei Betrachtung des ASW-Mittelwertes der zehn Stichproben fehlt. Bei der Anwendung der Methodik auf großflächige Datensätze sollte deshalb ein Augenmerk auf die Untersuchung und Festlegung der geeigneten Gruppenanzahl mittels PAM-Algorithmus gerichtet werden. Zudem sollten im Rahmen der vorliegenden Untersuchung identifizierte informationstechnische Optimierungsmöglichkeiten der Daten-

verarbeitung mittels R in weiterführenden Vorhaben durch programmiertechnische Maßnahmen für die clusteranalytische Analyse sehr großer Datensätze erschlossen werden.

Tabelle 6-8 Auswertung der Verläufe der mittleren Silhouetten-Weiten (ASW) für zehn Zufallsstichproben der Teilbauflächen in Berlin und Festlegung des Modus der je Stichprobe geeigneten Gruppenanzahl als insgesamt geeignete Gruppenanzahl k_{opt}

#	Gruppenanzahl k																			
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
1		x			x															
2		x				x														
3		x			x			x												
4		x																		
5		x			x		x													
6		x			x					x										
7		x																		
8		x				x			x											
9		x			x				x											
10		x				x		x												
Σ		10			5	3	1	1	2	1										

Hinweis: Der für alle Stichproben höchste ASW-Wert liegt bei einer Gruppenanzahl von 3. Eine Überprüfung der Geodaten zeigt jedoch, dass diese Einteilung nur einen geringen Mehrwert für die Energieforschung im Quartier hat. Es wurden die nächstbesten lokalen Maxima verwendet.

6.4.3 Plausibilisierung der Gruppenbildung für Teilbauflächen

Die mittels Clusteranalyse gebildeten Gruppen für Teilbauflächen sollen nachfolgend hinsichtlich der Ähnlichkeit der Bebauungsstruktur sowie ihrer Anwendbarkeit in der Energieforschung und für die praktische Modernisierungsplanung überprüft werden. Hierfür wurden aus dem Stadtgebiet Berlin insgesamt neun Siedlungsgebiete ausgewählt, welche eine subjektiv als überwiegend homogen wahrgenommene Bebauungsstruktur aufweisen und in Bezug auf existierende Typologien für Siedlungsstrukturen zu Typen bzw. Gruppen zugeordnet werden können (siehe Abbildung 6-7, S. 129)⁷⁵:

⁷⁵ Für eine Beschreibung der existierenden Typologien vgl. Kapitel 2.4 und Anhang A. Die Auswahl beruht auf einem Vergleich von GIS-Daten mit Satelliten- und Schrägbildaufnahmen (siehe auch Anhang D-1).

- Einfamilienhäuser (EFH), freistehend:
Freistehende, meist ein- bis dreigeschossige Baukörper geringer Grundfläche und mittlerer bis großer zugeordneter unbebauter Grundstücksfläche
- Doppelhäuser (DH):
Wohngebäude geringer Grundfläche und Geschossigkeit, bei denen jeweils zwei Baukörper eine Gebäuderegion bilden
- Reihenhäuser (RH):
Überwiegend Wohngebäude, die aus mehreren aneinandergebauten Gebäuden eine Gebäuderegion bilden; Grundflächen variieren von gering (Reiheneinfamilienhäuser) bis mittel (kleines Mehrfamilienhaus in Reihenbebauung); die Geschossanzahl liegt meist zwischen zwei und vier Geschossen
- Mehrfamilienhäuser (MFH), freistehend:
Freistehende Gebäude geringer bis mittlerer Grundfläche und meist drei bis vier Vollgeschossen sowie vergleichsweise großzügigen Freiflächen
- Mehrfamilienhäuser, Zeilenbebauung:
Mehrfamilienhäuser mittlerer bis höherer Grundfläche, die als ein langgestreckter Baukörper wahrgenommen werden und meist aus mehreren Eingängen bzw. Gebäudeteilen bestehen; in der Regel drei- bis sechsgeschossig
- Wohnhochhäuser:
Gebäude in Wohnanlagen mit hoher Grundfläche und hoher bis sehr hoher Geschossigkeit (mindestens sechs Vollgeschosse); Baukörper können sowohl in freistehender als auch in Zeilen- oder Reihenbauweise vorliegen
- Blockrandbebauung:
Einzelgebäude mittlerer Grundfläche und i. d. R. drei bis sechs Vollgeschossen, welche entlang der Ränder eines Baublockes in Reihenbauweise angeordnet sind; innerhalb eines Baublockes besteht ein vergleichsweise hoher Freiflächenanteil
- Blockrandbebauung, hohe Dichte:
Gebäudekonfiguration des Blockrandes wird durch zusätzliche Flügelgebäude oder Hinterhofbebauung stark verdichtet
- Industrie und Gewerbe:
Großflächige Industrie und Gewerbehallen sowie mittelgroße Gewerbegebäude

Wird die Einteilung der Teilbauflächen innerhalb der ausgewählten Siedlungsgebiete in die mittels Clusteranalyse erzeugten Gruppen betrachtet, so zeigt sich, dass die städtebauliche

Charakteristik der Gebiete durch die Gruppenbildung sehr gut erfasst und gruppiert wird (siehe Abbildung 6-8, S. 131).⁷⁶

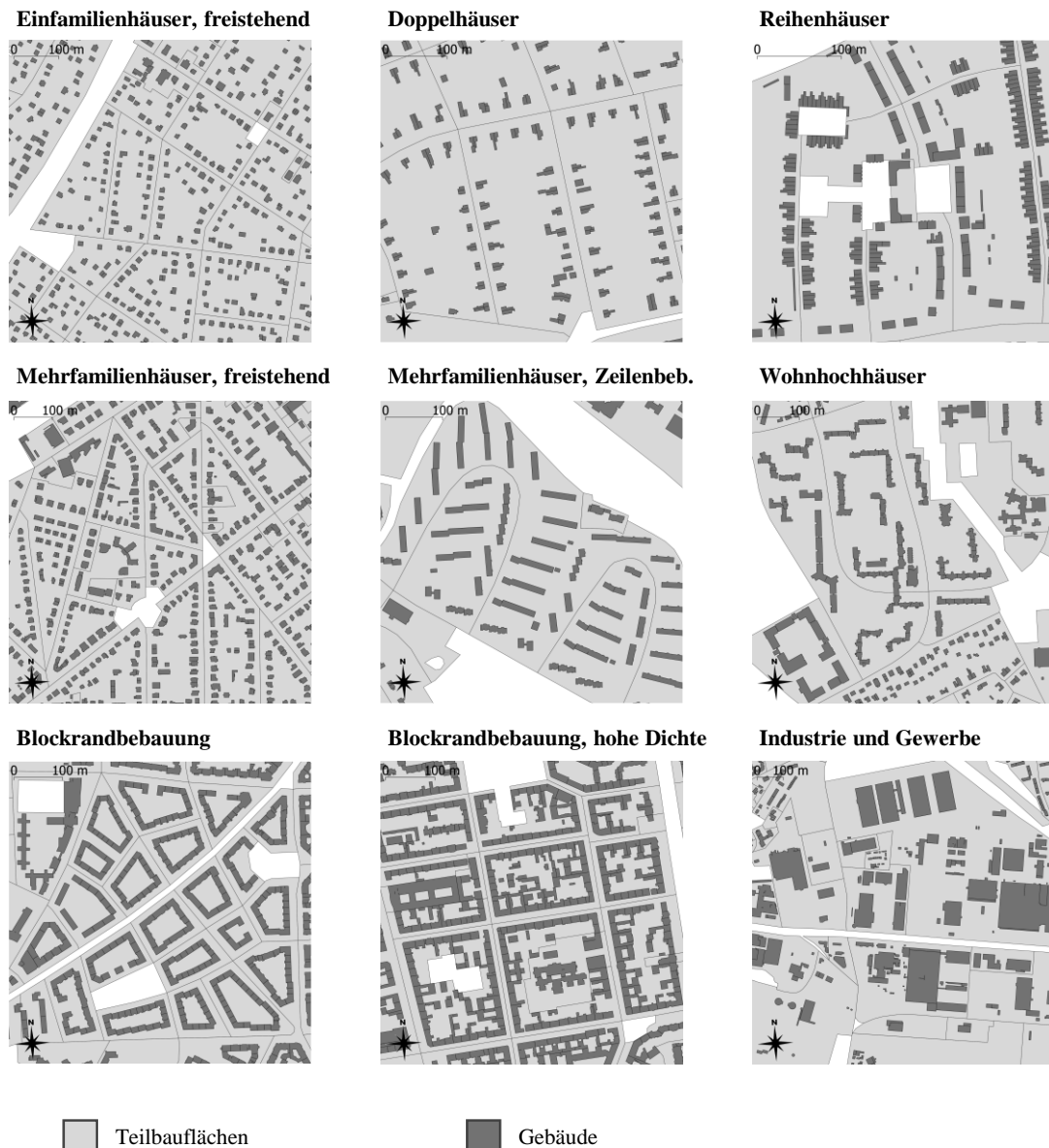
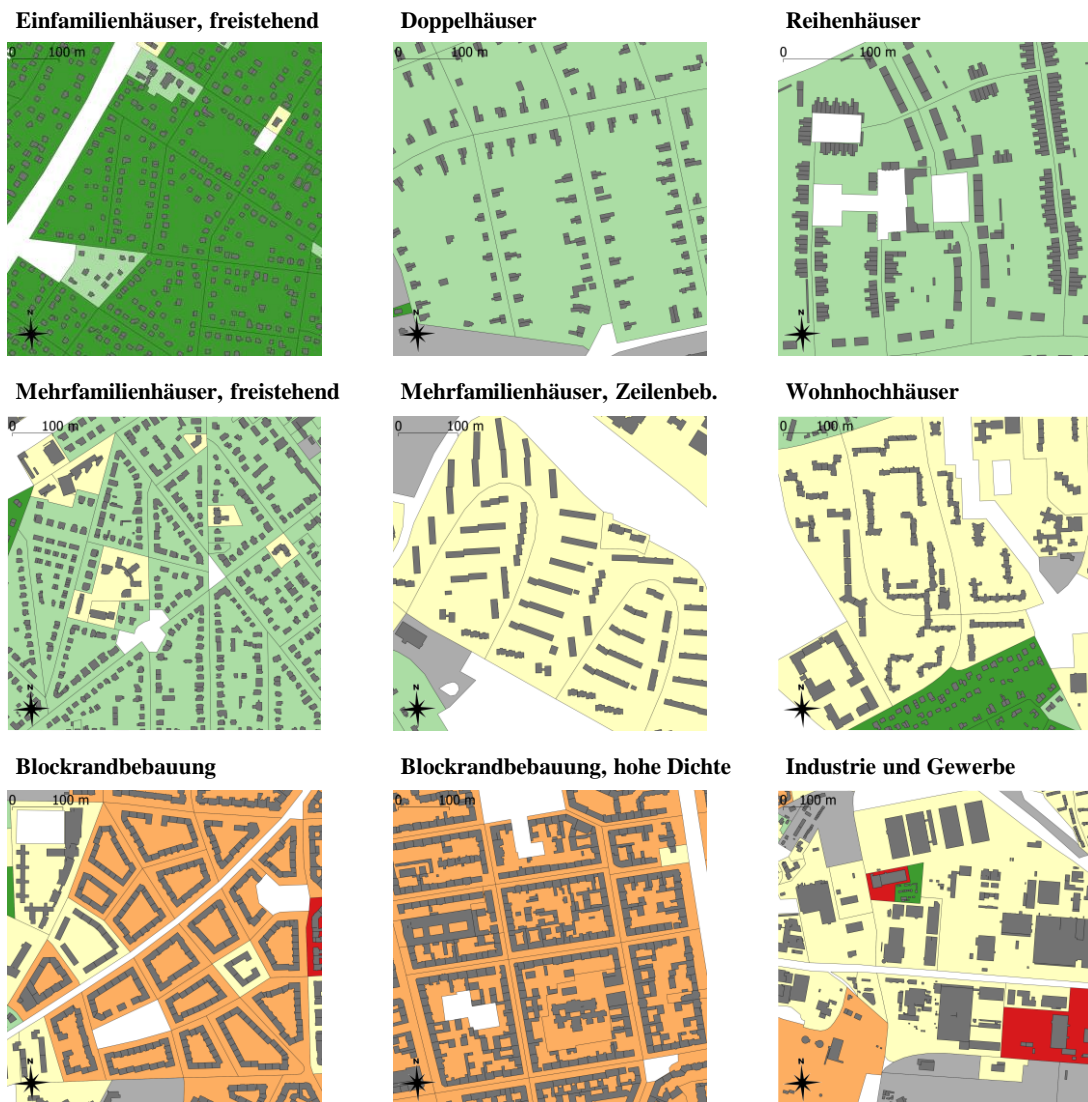


Abbildung 6-7: Ausgewählte Siedlungsgebiete Berlins für eine Überprüfung der Gruppenbildung für Teilbauflächen im Hinblick auf Plausibilität und Anwendbarkeit

⁷⁶ Die Benennung der Gruppen erfolgt im Rahmen der computergestützten Gruppenbildung zufällig. Für die einfachere Verwendung in dieser Arbeit wurden die Gruppen nach aufsteigendem MBN II sortiert benannt. Einzig Gruppe 6 wurde als „Sonderfall“ dieser Reihenfolge hinten angestellt. Für eine Begründung wird auf die nachfolgenden Absätze der Plausibilisierung der Gruppierungsergebnisse verwiesen.

Die detaillierten, gebäudetypfokussierten Einteilungen der existierenden Typologien können durch die erarbeitete Gruppeneinteilung für Teilbauflächen mit sechs Gruppen nicht vollständig nachvollzogen werden. Dennoch kann sehr deutlich erkannt werden, dass die Teilbauflächen hinsichtlich der in Kapitel 5.3 als maßgeblich für energetische Betrachtungen auf Quartiersebene identifizierten Kenngrößen der baulichen Nutzung gut separiert sind. Die aus der Clusteranalyse gewonnene Einteilung der Teilbauflächen in Gruppen ähnlicher Bebauungsstruktur kann dabei wie folgt beschrieben werden:

- **Gruppe 1** besteht aus Gebieten mit einer Bebauung geringer städtebaulicher Dichte. Darin befinden sich überwiegend freistehende Einfamilienhäuser sowie als Doppelhäuser zu bezeichnende Gebäude mit kleinen Grundflächen und geringer Gebäudehöhe (ein bis drei Geschosse). Es existieren vergleichsweise große unbebaute Freiflächen in direkter Umgebung der Einzelgebäude.
- **Gruppe 2** umfasst ebenfalls Teilbauflächen mit geringer Bebauungsdichte. Die Gebiete sind jedoch durch andere Baukörper und ein höheres flächen- und volumenmäßiges Nutzungsmaß gekennzeichnet. Es befinden sich überwiegend Doppelhäuser und Reihenhäuser geringer Gebäudegrundflächen und Gebäudehöhen sowie kleinere, freistehende Mehrfamilienhäuser innerhalb der Teilbauflächen dieser Gruppe.
- **Gruppe 3** ist geprägt durch eine mittlere Bebauungsdichte und Gebäude mit mittlerer Geschossigkeit. Es handelt sich dabei sowohl um Zeilenbauten mit drei bis sechs Vollgeschossen als auch um höhergeschossige Wohnanlagen bis hin zu Wohnhochhäusern. Trotz der vergleichsweise hohen Nutzfläche je Gebäude ergibt sich aufgrund von umgebenden Freiflächen eine insgesamt mittlere flächen- und volumenmäßige Bebauungsdichte.
- **Gruppe 4** wird aus Teilbauflächen hoher bis sehr hoher städtebaulicher Dichte gebildet. Es handelt sich dabei zum Großteil um mehrgeschossige Wohn- und Nichtwohngebäude, welche in einer Blockrandkonfiguration gemeinsam eine oder nur wenige Gebäuderegionen auf einer Teilbaufläche bilden.
- **Gruppe 5** ist geprägt durch Teilbauflächen, auf denen sich einzelne oder mehrere Gebäudekomplexe mit großer Grundfläche und einem meist großen Länge-zu-Breite-Verhältnis befinden. Die Bebauungsstruktur besitzt nur sehr geringe Anteile nicht überbaute Fläche und ist üblicherweise in Stadtzentren (central business districts, CBD) oder peripheren städtischen Lagen mit guter Verkehrsanbindung vorzufinden.
- **Gruppe 6** besteht aus meist weitläufigen Teilbauflächen mit einer äußerst geringen Anzahl an Gebäuden. Es handelt sich dabei häufig um Sondernutzungsgebiete, in denen stark zweckgebundene Nichtwohnbauten vorzufinden sind (z. B. Parks).



Quartiersgruppen

- 1 Gebiete mit freistehender Bebauung geringer Dichte (überwiegend EFH)
 - 2 Gebiete geringer Bebauungsdichte (DH, RH, freistehende kleine MFH)
 - 3 Gebiete mittlerer Bebauungsdichte (MFH in Zeilenbebauung, Wohnhochhäuser)
 - 4 Stark verdichtete (Innen-)Stadtgebiete (meist Blockrandbebauung)
 - 5 Stark verdichtete Gebiete mit kompakter Bebauung (z. B. Gewerbe- oder Wohnkomplexe)
 - 6 Sondergebiet (meist großflächige Teilbauflächen mit wenigen Nichtwohngebäuden)
 - Gebäude
- geringe Dichte hohe Dichte

Abbildung 6-8: Vergleich der Gruppenbildung für Teilbauflächen mittels Clusteranalyse mit Siedlungsgebieten überwiegend homogener Bebauungsstruktur sowie Beschreibung der gebildeten Teilbauflächengruppen

Der Vergleich der mittels Clusteranalyse gebildeten Teilbauflächengruppen mit realen Bebauungssituationen im Untersuchungsraum Berlin hat gezeigt, dass die Gruppenbildung eine sinnvolle Einteilung hinsichtlich der Bebauungsstruktur erreicht hat. Diese ist aufgrund der festgelegten Anzahl von sechs Gruppen zwar im Vergleich zu bestehenden Gebäude- und Siedlungs- bzw. Raumstrukturtypologien als weniger differenziert zu betrachten, im Hinblick auf die mögliche Verwendung in Untersuchungen zu den Beiträgen von Energieeffizienz- und Wärmeversorgungs-lösungen im Quartiersmaßstab für die Erreichung der Emissionsminderungsziele im Gebäudebestand scheinen die maßgebenden Zusammenhänge, d. h. die geometrischen Gebäudeeigenschaften und die Raumbeziehungen der Gebäude, jedoch sehr klar herausgearbeitet. Daraus lässt sich ein hohes Maß der Anwendbarkeit der Ergebnisse der Gruppenbildung ableiten. Zudem besteht die Möglichkeit, in spezifischen Forschungs- bzw. Planungskontexten durch eine veränderte Datenauswahl oder eine andere Gruppenanzahl, eine vorhabenspezifische Gruppenbildung zu erzeugen.

Der erfolgreichen Einteilung von Teilbauflächen in Gruppen ähnlicher Bebauungsstruktur kann nun die Bildung von Quartieren aus ebendiesen Teilbauflächen folgen. Die oben beschriebene Gruppeneinteilung und die darin abgebildeten Bebauungsstrukturen werden dabei als grundlegende Einteilungssystematik der späteren Quartiere dienen.

6.5 Bildung von Quartieren aus gruppierten Teilbauflächen

Die Festlegung der geeigneten Gruppenanzahl $k_{opt} = 6$ führte als Ergebnis des PAM-Algorithmus zur Einteilung der Teilbauflächenobjekte in sechs Gruppen. Der Teilbauflächendatensatz mit dem ergänzten Merkmal der Gruppenzugehörigkeit wird nun verwendet, um den in Kapitel 5.5 beschriebenen Ablauf zur Bildung von (im vorliegenden Fall sechs) Quartiersgruppen aus Teilbauflächen durchzuführen (siehe auch Abbildung 5-8, S. 102). Hierfür ist, neben der Vereinigung direkt aneinander angrenzender Teilbauflächen gleicher Gruppenzugehörigkeit, zudem eine Untersuchung bzgl. der Zuordnung von Insel- bzw. Splitterteilbauflächen notwendig, um die Anzahl der Kleinstquartiere mit geringer praktischer Relevanz zu reduzieren.

Es werden im Rahmen der Quartiersbildung verschiedene Programmroutinen ausgeführt, welche innerhalb der R-Programmierungsumgebung sowie der PostGIS-Datenbank ...

- das Straßennetz als Grenze der Quartiersbildung anwendbar machen,
- Teilbauflächen identifizieren, die derselben Gruppe angehören und benachbart sind,
- diese Teilbauflächen zu Gebieten zusammenführen und
- eine Identifikation und Zuordnung von Insel- und Splitterelementen zu benachbarten Gebieten vornehmen.

Als Grenzelemente für die Quartiersbildung werden in diesem Anwendungsbeispiel ausschließlich Objekte der Verkehrsinfrastruktur berücksichtigt. Ein weiterführender Einbezug von z. B. Verwaltungsgrenzen der Stadtbezirke wird nicht durchgeführt, da das Ziel der beispielhaften Anwendung keine gebietsspezifische, d. h. lokale Beschreibung eines Quartiersbestandes ist. Eine Berücksichtigung weiterführender Grenzobjekte im Rahmen der Geometrieverarbeitungsroutinen ist jedoch durch geringfügige Anpassungen der Programmcodes möglich.

Für die Identifikation und Zuordnung von Inselementen wird im Zuge der Methoden-anwendung zudem eine vergleichende Untersuchung unterschiedlicher Grenzwerte für das festzulegende Flächenverhältnis von Inselement und aufnehmendem Gebiet durchgeführt. Dabei gilt es, ein geeignetes Flächenverhältnis zu identifizieren, bei dem der Einfluss der Zuordnung einer gruppenfremden Teilbaufläche die Ähnlichkeit der Bebauungsstruktur innerhalb der gebildeten Gebiete aus mehreren Teilbauflächen noch gering ist und dennoch eine praxisingerechte Reduktion der Anzahl von Kleinstquartieren erreicht wird. Für die Quartiersbildung am Beispieldatensatz Berlin zeigt sich eine Fläche des Inselementes von höchstens 10 % der Fläche des absorbierenden Gebietes als geeignet. Durch diese Zuordnung wurde die Anzahl von Kleinstquartieren verringert, bei einem als akzeptabel bewerteten Einfluss auf die Bebauungsstruktur der aufnehmenden Gebiete. Da die Zuordnung von Inselementen in diesem Beispiel der Methoden-anwendung nur einmalig ausgeführt wird, liegen auch nach der Quartiersbildung weiterhin Kleinstquartiere im Quartiersdatensatz vor. In weiterführenden Methoden-anwendungen ist in Abhängigkeit der jeweiligen Zielsetzungen durch den Anwender zu überprüfen, ob das Vorliegen einer begrenzten Anzahl von Kleinstquartieren akzeptabel oder ob eine mehrfache Ausführung der Zuordnung von Insel- und Splitterelementen notwendig ist.

Als Ergebnis der Quartiersbildung entsteht der in Abbildung 6-9 (S. 134) dargestellte Quartiersbestand in Berlin. Die Bebauungsstrukturen der Quartiersgruppen entsprechen dabei weitestgehend der aus der Gruppeneinteilung der Teilbauflächen gewonnenen Differenzierung. Durch die Zusammenführung der Teilbauflächen entlang der definierten Grenzelemente der Quartiersbildung und durch die Zuordnung einzelner, nicht gruppenzugehöriger Inselemente ergeben sich jedoch insgesamt andere Gebietszuschnitte, deren statistische Auswertung im letzten Teil der Methode zur GIS-datenbasierten Quartierstypisierung erfolgt. Diese Auswertung erlaubt schlussendlich auch die Ableitung der Typquartiere.

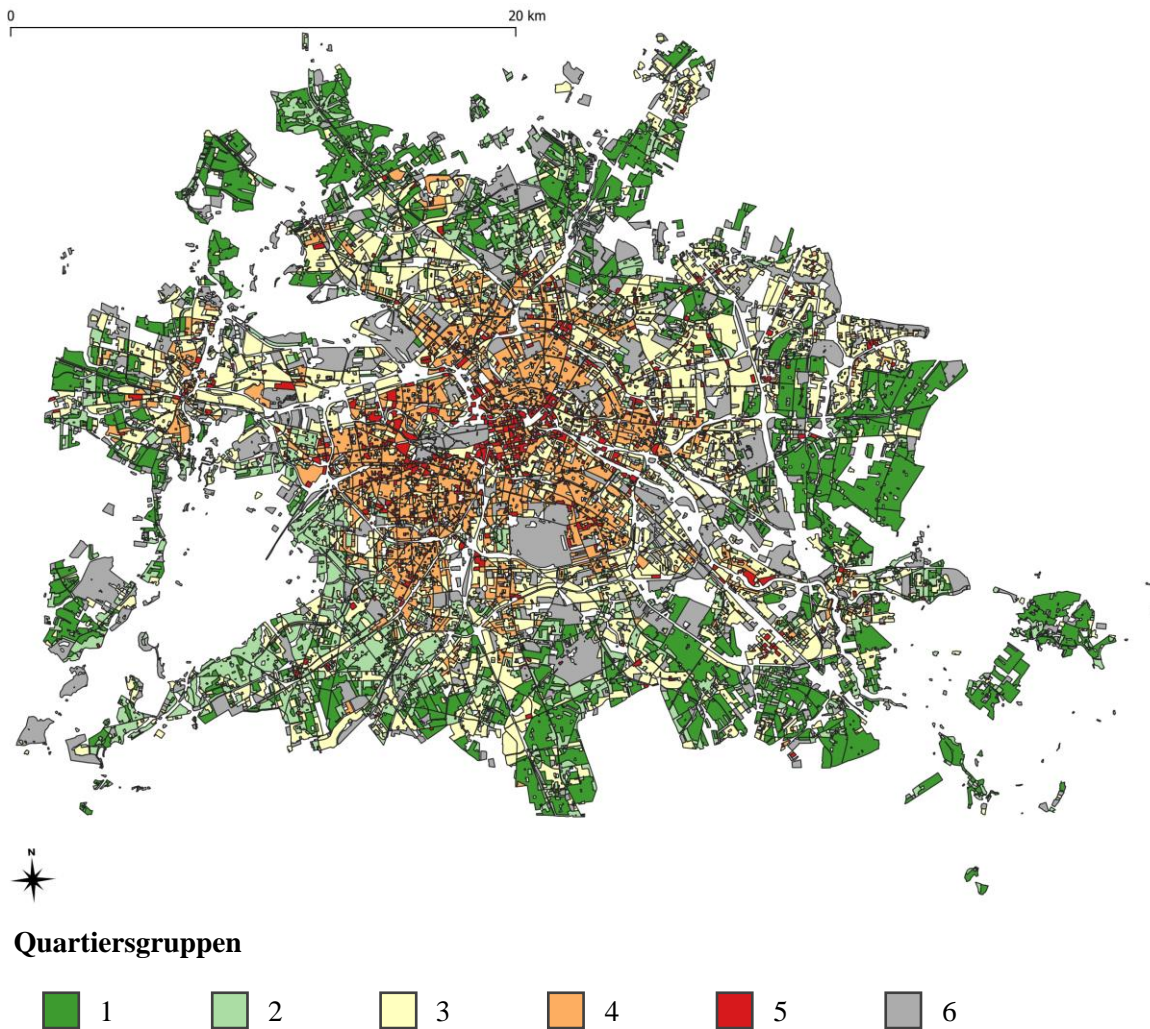


Abbildung 6-9: Quartiersbestand in Berlin basierend auf einer Einteilung der Teilbauflächen in 6 Gruppen ähnlicher Bebauungsstruktur, der Vereinigung benachbarter Teilbauflächen gleicher Gruppenzugehörigkeit sowie der einmaligen Ausführung der Zuordnung von Insel- und Splitterelementen (mit Flächenkriterium $A_{\text{Insel}} \leq 10\% A_{\text{Zielgebiet}}$)

6.6 Statistische Analyse des Berliner Quartiersbestandes und Festlegung von Kennwerten für Typquartiere

Die statistische Analyse des Berliner Quartiersbestandes erfolgt gemäß den Ausführungen in Kapitel 5.6 mit dem Ziel, die energetisch-städtebaulichen Merkmale der gebildeten Quartiere und Quartiersgruppen zu beschreiben und den Quartiersbestand zu quantifizieren. Im Rahmen der Untersuchung wird zum einen eine Auswertung der Quartiere bzw. der sie beschreibenden Merkmale je Quartiersgruppe vorgenommen. Darüber hinaus erfolgt zudem eine Analyse der Gebäudebestände innerhalb der Quartiersgruppen, sodass etwaige Kennwerte, welche im Rahmen der ausschließlichen Quartiersanalyse nur als statistische Kennzahlen angegeben werden, auch für einen Abgleich mit gebäudezentrierten Analysen

bzw. Datenbasen zur Verfügung stehen.⁷⁷ Die Auswertungen liefern als Ergebnis ein Mengengerüst, welches den Quartiersbestand bzw. den für Maßnahmen im Quartiersmaßstab als relevant betrachteten Gebäudebestand innerhalb des gewählten Untersuchungsgebietes Berlin in seiner Gesamtheit beschreibt und somit dessen Quantifizierung darstellt. Darüber hinaus ergibt sich aus der statistischen Analyse eine nach Quartiersgruppen differenzierte Beschreibung der energetisch relevanten Bebauungsstrukturvariablen. Wie bereits in den Kapiteln 5.3 und 5.7 ausgeführt bezieht sich die Beschreibung der energetischen Merkmale dabei wegen des Mangels an Daten zur energetischen Qualität der Gebäude ausschließlich auf die geometrischen Einflussparameter des Gebäudeenergiebedarfes.

Tabelle 6-9 (S. 136) stellt das Ergebnis der statistischen Auswertung des Berliner Quartiersbestandes dar. Neben der „Anzahl der Quartiere und Gebäude“ in den sechs Quartiersgruppen lassen sich die energiebedarfsrelevanten Merkmale der Quartiere (insbesondere die „mittlere Gebäudegrundfläche“, die „mittlere Gebäudehöhe“ oder die „mittlere Gebäudeanzahl je Gebäuderegion“ der Gebäudepopulationen je Quartier) sowie die energieverorgungsrelevanten Attribute, wie die „spezifische Anzahl von Gebäuden“ bzw. „spezifische Anzahl von Gebäuderegionen“ und die beiden Maße der baulichen Nutzung ablesen. Beim Vergleich der statistischen Kennwerte zeigt sich eine klare Differenzierung der sechs Gruppen hinsichtlich der genannten energierelevanten Merkmale. So kann beispielsweise festgestellt werden, dass der Mittelwert der Gebäudegrundfläche aller Gebäude eines Quartiers im Mittel in der Quartiersgruppe 1 mit einem Wert von 107,8 m² am geringsten ist, wohingegen Gruppe 5 die im Mittel größte mittlere Gebäudegrundfläche je Quartier besitzt (Zeile „MW mittlere Gebäudegrundfläche je Quartier [m²]“ in Tabelle 6-9). Bei Betrachtung der Standardabweichung der mittleren Gebäudegrundfläche je Quartier zeigt sich insbesondere für die Gruppen 3 und 5 jedoch auch, dass hier eine besonders hohe Varianz der Grundfläche existiert (Zeile „SD mittlere Gebäudegrundfläche je Quartier [m²]“). Es gibt also Quartiere innerhalb dieser Klassen mit stark unterschiedlichen mittleren Gebäudegrundflächen. In Quartiersgruppe 1 ist dieser Wert deutlich geringer, was eine stärkere Homogenität der Quartiere dieser Gruppe im Hinblick auf die mittlere Gebäudegrundfläche im Quartier bedeutet. Diese Feststellung kann ergänzt werden um die Bewertung der Standardabweichung der mittleren Gebäudegrundfläche innerhalb der einzelnen Quartiere.

⁷⁷ Die quartiersbezogene Auswertung liefert die gebäudebezogenen Merkmale nur als statistische Parameter innerhalb des Quartiers. So liegt dort die Gebäudegrundfläche z. B. nur als Variable „Mittelwert der Gebäudegrundfläche“ des Quartiers bzw. „Standardabweichung der Gebäudegrundfläche“ des Quartiers vor. Da die quartiersgruppenbezogene Statistik für diese Variable keine weitere Gewichtung vornimmt, ist der für eine Quartiersgruppe errechnete Mittelwert der „Mittelwerte der Gebäudegrundfläche je Quartier“ ungleich zum Mittelwert der Gebäudegrundfläche für alle Gebäude innerhalb der Quartiersgruppe.

6 Anwendung der Methodik zur Typisierung von Quartieren am Beispiel Berlin

Tabelle 6-9: Statistische Auswertung des Quartiersbestandes innerhalb der gebildeten Quartiersgruppen in Berlin

Variablen	Quartiersgruppen					
	1	2	3	4	5	6
Anzahl Quartiere	1.009	1.692	2.302	890	1.096	1.682
Anzahl Gebäude	157.989	74.447	45.526	60.198	3.774	9.968
MW Quartiersfläche [ha]	11,36	4,32	6,10	6,78	1,54	6,37
SD Quartiersfläche [ha]	22,23	8,47	10,11	9,35	2,09	18,04
MW Gesamtgebäudefläche je Quartier [1.000 m ²]	14,81	6,58	12,96	23,53	6,21	1,59
SD Gesamtgebäudefläche je Quartier [1.000 m ²]	28,06	13,17	23,43	34,83	9,11	3,06
MW mittlere Gebäudegrundfläche je Quartier [m ²]	107,8	191,3	810,5	415,8	2.212,8	343,0
SD mittlere Gebäudegrundfläche je Quartier [m ²]	45,5	122,9	789,1	293,9	2.703,8	363,7
MW der SDs der mittleren Gebäudegrundfläche je Quartier [m ²]	53,4	120,8	586,7	317,9	780,4	150,7
MW flächengew. mittlere Gebäudehöhe je Quartier [m]	8,2	9,2	13,4	18,9	20,8	7,0
SD flächengew. mittlere Gebäudehöhe je Quartier [m]	2,1	3,1	4,9	4,1	7,9	4,3
MW flächengew. SDs der mittleren Gebäudehöhe je Quartier [m]	0,4	0,6	1,1	0,7	1,0	0,6
MW mittlere Grundfläche der Gebäuderegion je Quartier [m ²]	298,2	558,2	2.016,6	7.260,0	6.799,6	774,8
MW der SDs mittlere Grundfläche der Gebäuderegion je Quartier [m ²]	173,9	320,1	1.223,6	2.471,3	1.093,9	273,2
MW mittlere Gebäudeanzahl je Gebäuderegion und Quartier	2,84	3,92	4,68	22,05	11,07	2,65
MW der SDs mittlere Gebäudeanzahl je Gebäuderegion und Quartier	1,69	2,00	2,20	6,97	1,28	0,72
MW Maß baulicher Nutzung I (MBN I) [m ² /m ²]	0,15	0,16	0,21	0,34	0,40	0,05
SD Maß baulicher Nutzung I (MBN I) [m ² /m ²]	0,06	0,08	0,07	0,08	0,13	0,04
MW Maß baulicher Nutzung II (MBN II) [m ³ /m ²]	1,27	1,53	2,67	6,50	7,95	0,34
SD Maß baulicher Nutzung II (MBN II) [m ³ /m ²]	0,73	0,96	1,14	2,19	3,51	0,32

MW = Mittelwert; SD = empirische Standardabweichung; MBN = Maß baulicher Nutzung

Während vorher die Varianz zwischen den Mittelwerten der Quartiere betrachtet wurde, ist diese dritte Kennzahl ein Maß für die Varianz der Gebäudegrundfläche innerhalb der Gebäudepopulationen der jeweiligen Quartiere (Zeile „MW der SDs der mittleren Gebäudegrundfläche je Quartier [m²]“). Hier zeigt sich, dass die Streuung der Gebäudegrundfläche innerhalb der einzelnen Quartiere im Mittel insbesondere für Gruppe 5 deutlich geringer ist. Die Grundflächen der Gebäude innerhalb eines Quartiers sind somit ähnlich, auch wenn die Quartiere der Gruppe sich hinsichtlich der mittleren Gebäudegrundfläche z. T. deutlich unterscheiden.

Neben der tabellarischen Darstellung der Ergebnisse wird für das Attribut „MBN I“ zudem in Abbildung 6-10 (S. 138) die Dichtefunktion je Gruppe gegeben. Diese Darstellung unterstützt die in Tabelle 6-9 erkennbare Differenzierung der Quartiere hinsichtlich der flächenmäßigen Bebauungsdichte mit geringer Bebauungsdichte in den Gruppen 1 und 2 und hoher Bebauungsdichte in den Gruppen 4 und 5. Die Dichtefunktionen anderer Merkmale sind in Anhang D-3 gegeben und bestätigen ebenfalls die meist gute Unterscheidung bei Betrachtung der Gruppenmittelwerte bzw. Mediane. Dennoch lässt sich attestieren, dass insbesondere die Quartiersgruppen 4 und 5 für einige der Merkmale eine hohe Varianz aufweisen und somit keine scharfe Gruppenabgrenzung in Bezug auf diese Merkmale erzielt wurde. Um eine vollständige Beschreibung der z. T. nichtsymmetrischen Merkmalsverteilungen zu liefern, werden in Anhang D-3 neben den Dichtefunktionen ebenfalls die Schiefe und die Kurtosis der Merkmalsverteilungen angegeben. Diese Größen geben die Abweichung der Verteilung der Merkmalswerte von einer Normalverteilung an.

Nach der Quantifizierung der Quartiere durch Beschreibung ihrer Anzahl sowie nach der Analyse statistischer Kennzahlen der verschiedenen Quartiersmerkmale in den sechs Quartiersgruppen kann eine abschließende Festlegung von Kennwerten für sechs „synthetische“ Typquartiere erfolgen. Da auch in diesem Schritt der Anwendung der Methodik eine auf den empirischen Kennzahlen beruhende Festlegung von Typquartieren erfolgt, sollen diese durch die Wahl des Begriffs „synthetisch“ explizit von anderen Möglichkeiten der Festlegung von Typvertretern abgegrenzt werden, z. B. realen Beispielquartieren. Wegen der häufig nicht normalverteilten Merkmalsverteilungen der bebauungsstrukturbeschreibenden Variablen wurde im Rahmen der Methodenentwicklung die Festlegung getroffen, dass als statistische Maßzahl für die Festlegung von Kenngrößen für Typquartiere Medianwerte der Merkmalsverteilungen verwendet werden.

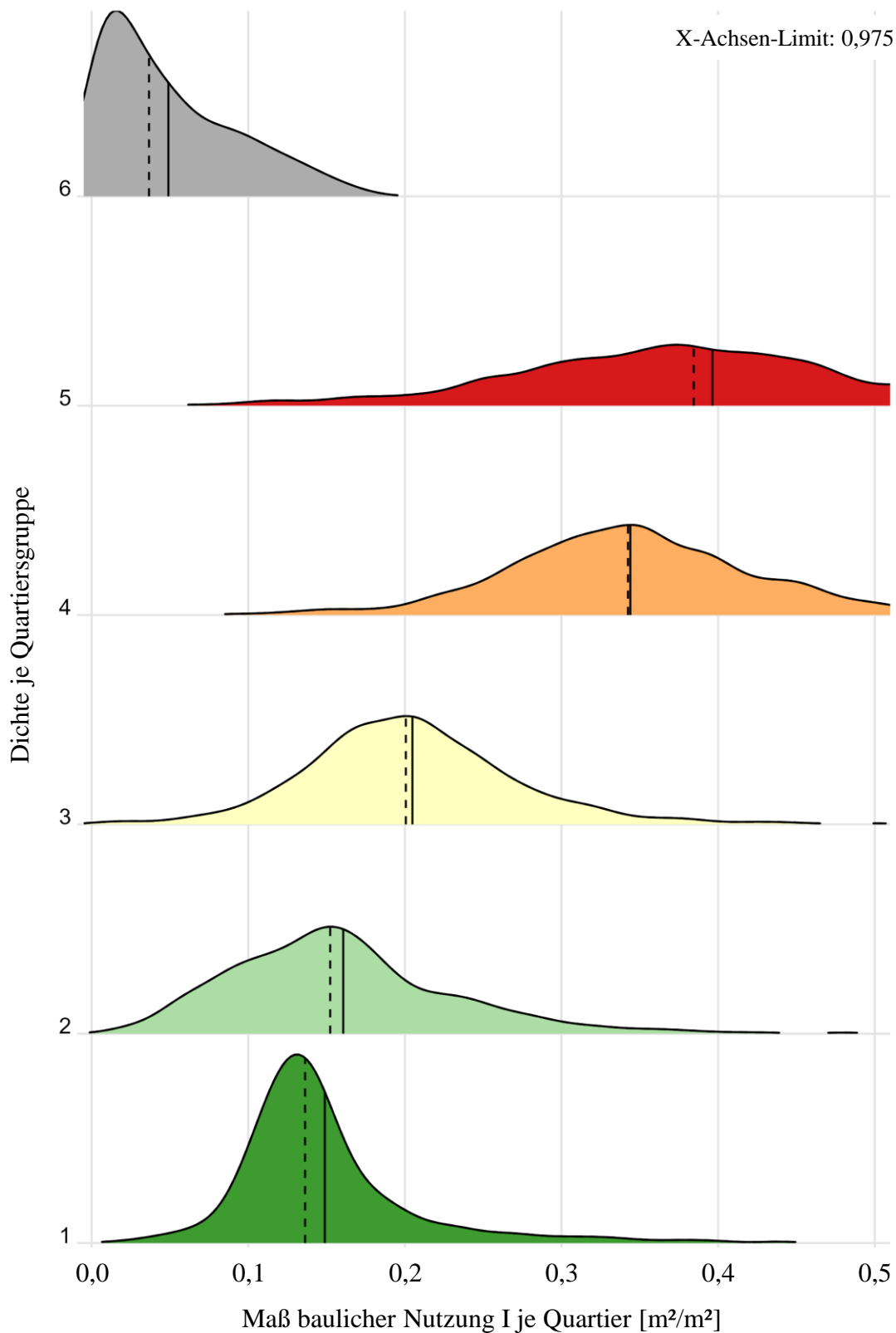


Abbildung 6-10: Dichtefunktion der Merkmalsausprägungen der Variable „Maß baulicher Nutzung I“ der Quartiersgruppen (Mittelwert = schwarze Linie; Median = gestrichelte Linie; Darstellung der X-Achse schneidet bei Quantilswert $Q_{0,95}$ ab)

Die Beschreibung der Bebauungsstruktur der synthetischen Typquartiere mittels Medianwerten für die hinsichtlich energiebedarfs- sowie energieverorgungsbezogener Fragestellungen relevanten Merkmale sind in Tabelle 6-10 dargelegt. Es ist erkennbar, dass die im Rahmen der Anwendung der Methodik festgelegten Parameter (z. B. geeignete Gruppenanzahl oder nur einmalige Iteration der Zuordnung von Inselgebieten) für diese beispielhafte Anwendung zu Typvertretern mit geringen bis mittleren Quartiersflächen und im stadtplanerischen Kontext eher zu geringen Gebäudezahlen führen. Aus den Kennwerten der Gebäudegeometrie sowie der Bebauungszusammenhänge lassen sich Modelle von Quartieren erstellen, welche sich in strategischen Untersuchungen oder Potenzialabschätzungen nutzen lassen.

Tabelle 6-10: Merkmale der aus der statistischen Auswertung des Quartiersbestandes abgeleiteten synthetischen Typquartiere für Berlin

Variablen	Typquartier					
	1	2	3	4	5	6
Anzahl Gebäude	48	16	7	27	2	2
Anzahl Gebäuderegionen	36	9	4	3	1	2
Quartiersfläche [ha]	3,55	1,73	2,43	3,33	0,85	1,88
Gebäudegrundfläche [m ²]	98,7	160,6	627,2	346,8	1.399,7	230,5
Mittlere Gebäudehöhe [m]	8,1	9,0	13,1	19,2	21,0	6,1
Maß baulicher Nutzung I (MBN I) [m ² /m ²]	0,14	0,15	0,20	0,34	0,38	0,04
Maß baulicher Nutzung II (MBN II) [m ³ /m ²]	1,13	1,36	2,58	6,23	7,52	0,24
Beschreibung der Bebauungsstrukturen gemäß Kapitel 6.4.3:						
1 Gebiete mit freistehender Bebauung geringer Dichte (überwiegend EFH) 2 Gebiete mit Bebauung geringer Dichte (DH, RH, freistehende MFH) 3 Gebiete mit mittlerer Bebauungsdichte (MFH in Zeilenbebauung, Wohnhochhäuser) 4 Stark verdichtete (Innen-)Stadtgebiete (meist Blockrandbebauung) 5 Stark verdichtete Gebiete mit kompakter Bebauung (z. B. Gewerbe- oder Wohnkomplexe) 6 Sondergebiet (meist großflächige Teilbauflächen mit wenigen Nichtwohngebäuden)						

Die festgelegten Kennwerte liefern insgesamt die Möglichkeit einer Verwendung für strategische Modelluntersuchungen auf einer hohen räumlichen Ebene, d. h. für dieses Anwendungsbeispiel für die Stadt Berlin. Weiterhin können die Typquartiere, als Vertreter der Quartiere der jeweiligen Gruppen, z. B. auch für Voruntersuchungen in einem lokalen, kleinräumigen Kontext eingesetzt werden. Dies ist insbesondere dann zielführend, wenn in besonders frühen Planungsphasen eine noch unzureichende Datenlage für ein Untersuchungsgebiet vorliegt. Die synthetischen Typquartiere können in diesem Kontext für Voruntersuchungen verwendet werden, welche zur Eingrenzung sinnvoller Versorgungs-

konzepte für die Zielsetzungen eines Vorhabens dienen. Eine weitere Detaillierung ist dann jedoch nach der vollständigen Grundlagenermittlung zwingend notwendig.

Während für eine Studie mit Fokus auf großflächige leitungsgebundene Wärmenetze mehrere Iterationen zu empfehlen sind, können die erzeugten Ergebnisse jedoch z. B. bei der Untersuchung von Möglichkeiten lokaler Nahwärmeversorgungs- und Contracting-Lösungen hilfreich sein. Es ist anzunehmen, dass die Gebäudeanzahl und flächenmäßige Ausdehnung sowie die sich aus den Gebäudegeometrien und Bebauungsdichten (in Abhängigkeit von der energetischen Qualität der Gebäude) ergebenden Wärme- und Leistungsbedarfe in Bereichen liegen, in denen privatwirtschaftliche Unternehmen ihre Geschäftsfelder sehen und unter Zuhilfenahme der Typquartiere und Mengengerüste den Ausbau effizienter Wärmeversorgungs-lösungen vorantreiben können.

Die Anwendung der Methodik hat somit gezeigt, dass die Einteilung von kleinräumigen Gebieten basierend auf den dort vorliegenden Bebauungsstrukturen möglich ist und die daraus ermittelten Quartiere eine sinnvolle Abgrenzung besitzen im Hinblick auf die Verwendung in weiterführenden energiebezogenen Forschungs- und Planungskontexten. In Kapitel 7 werden nun die Möglichkeiten aufgezeigt, welche die Anwendung der entwickelten Methodik zur Quartierstypisierung unter Verwendung des nationalen Geodatenbestandes eröffnen.

7 MÖGLICHKEITEN ZUR QUARTIERSTYPISIERUNG DES NATIONALEN GEODATENBESTANDES

Die erarbeitete und am Beispiel der Stadt Berlin angewendete Methodik zur Einteilung und statistischen Analyse von Quartieren sowie zur Festlegung von Kenngrößen für Typquartiere nutzt flächendeckend verfügbare GIS-Daten und ist mit dem Ziel der Anwendbarkeit für die Untersuchung des nationalen Quartiersbestandes entwickelt worden. Den vom Bundesamt für Kartographie und Geodäsie zur Verfügung gestellten Daten liegt dabei eine einheitliche Datenstruktur zugrunde, sodass ohne Anpassung der Verfahrensschritte der Methodik zwei Ansätze zur Anwendung der Methodik auf den nationalen Datenbestand möglich sind. Diese beiden Vorgehensweisen zur Quartierstypisierung unter Verwendung des nationalen Datenbestandes werden nachfolgend skizziert und deren Vor- und Nachteile diskutiert. Hierfür wird zur besseren Veranschaulichung des Ablaufs auf die Ergebnisse aus Kapitel 6 Bezug genommen. Das Vorgehen ist jedoch unabhängig vom Beispiel allgemein, d. h. auch für andere Gruppen von Teilbauflächen, anwendbar.⁷⁸

Auf der Basis der Datenvorverarbeitung in Kapitel 6.3 erfolgte mittels Clusteranalyse eine Gruppeneinteilung für alle Teilbauflächen in Berlin. Diese für die Grundgesamtheit⁷⁹ eines Teilbauflächenbestandes ermittelte Gruppeneinteilung kann, wie in Kapitel 5.4.2 für das Vorgehen unter Anwendung einer Zufallsziehung von mehreren Teilbauflächenstichproben beschrieben (vgl. Abbildung 5-3 auf S. 91), für die Erstellung eines Datenmodells für den Naive Bayes Klassifizierer verwendet werden. Für dieses ML-Verfahren bilden die mittels Clusteranalyse erzeugten Gruppenzuordnungen für die Teilbauflächen in Berlin somit eine Referenz. Der auf Basis dieser Referenzobjekte trainierte Naive Bayes Klassifizierer kann nun auch für andere Teilbauflächendatensätze gleicher Datenstruktur angewendet werden und die Wahrscheinlichkeit der Zugehörigkeit zu einer der Teilbauflächengruppen für jedes Objekt des neuen Datensatzes bestimmen. Die bisher nicht gruppierten Teilbauflächen werden im Rahmen dieser sog. Klassifizierung der Gruppe mit der höchsten Zugehörigkeitswahrscheinlichkeit zugeordnet. Zur Erstellung eines Teilbauflächendatensatzes gleicher Datenstruktur muss die Datenvorverarbeitung, die für Berlin vorgenommen wurde, auch

⁷⁸ Es kann sich hierbei sowohl um Teilbauflächen eines anderen Untersuchungsgebietes, welche mit der entwickelten Methodik eingeteilt wurden, als auch eine völlig andere Gruppeneinteilung handeln (z. B. die Verwendung von Teilbauflächen, welche manuell in ingenieurmäßig formulierte Typologien eingeteilt wurde).

⁷⁹ Im Falle der Methodikanwendung unter Zuhilfenahme einer Stichprobenziehung aus dem Teilbauflächendatensatz liegt bereits ein Datenmodell des Naive Bayes Klassifizierers vor, welches – nach der Vorverarbeitung des Gebäudedatenbestandes gemäß Schritt 2 der Methodik (vgl. Kapitel 5.3) – zur Gruppenzuordnung der Teilbauflächen des nationalen Datenbestandes angewendet werden kann.

für den nationalen Datensatz erfolgen. Dies bedeutet, dass die gleichen Ausschlusskriterien für den Datensatz der 3D-Gebäudemodelle angewendet werden müssen, bevor die Gebäude hinsichtlich ihrer geometrisch-topologischen Merkmale analysiert und die daraus errechneten Kennwerte der lokalen Gebäudepopulationen den Teilbauflächen zugeordnet werden. Die Anwendung des trainierten Naive Bayes Klassifizierers auf diesen neu erstellten, den nationalen Gebäude- und Teilbauflächenbestand abdeckenden Datensatz führt im Ergebnis zu einer Einteilung aller Teilbauflächen in die für den Datensatz Berlin erzeugten Gruppen (siehe Abbildung 7-1). Der so erzeugte nationale Teilbauflächendatensatz gruppierter Teilbauflächen kann nun für die Quartiersbildung verwendet werden. Hierbei ist insbesondere darauf zu achten, welche Grenzelemente bei der Vereinigung der Teilbauflächen zu Quartieren zu beachten sind (z. B. Verwaltungsgrenzen zwischen Kommunen).

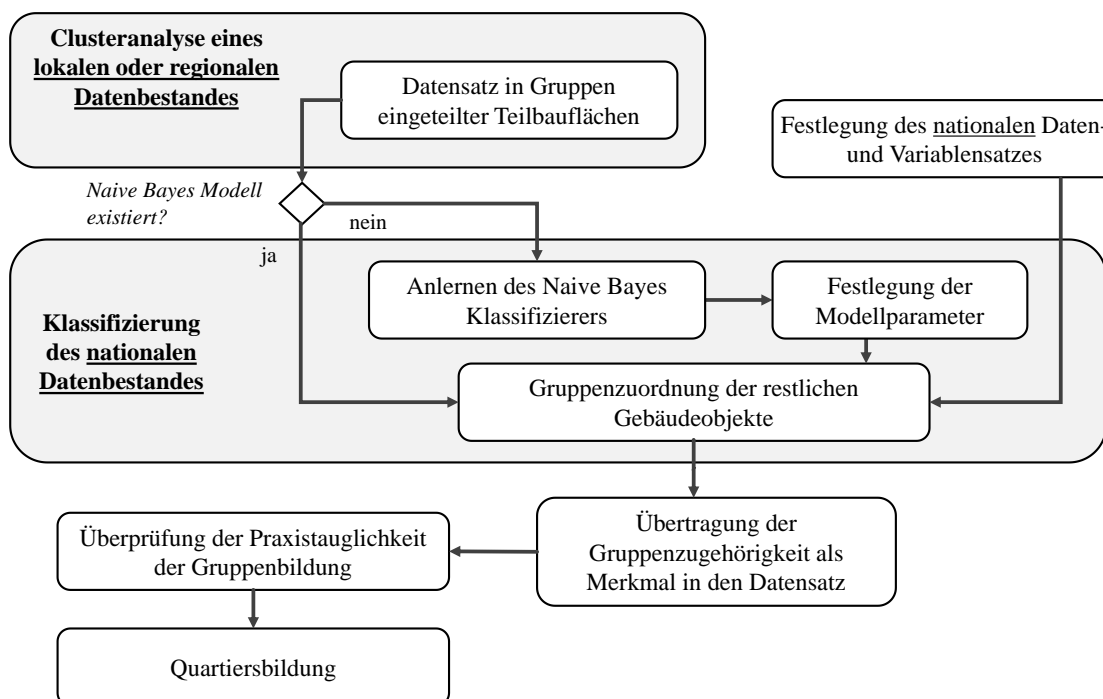


Abbildung 7-1: Verwendung einer mittels Clusteranalyse erzeugten Gruppeneinteilung von Teilbauflächen ähnlicher Bauungsstruktur zur Klassifizierung der Teilbauflächen des nationalen Gebäudbestandes

Das skizzierte Verfahren unter Verwendung der bestehenden Gruppeneinteilung der Teilbauflächen hat den Vorteil, dass die rechenzeitintensiven Operationen der Distanzmatrixbestimmung und der Gruppenbildung mittels PAM-Algorithmus nicht noch einmal durchgeführt werden müssen. Eine Anwendung geht deshalb mit vergleichsweise geringem Verarbeitungsaufwand einher. Gleichzeitig ist bei diesem Ansatz zur Quartierstypisierung unter Verwendung des nationalen Geodatenbestandes jedoch zu beachten, dass der als Referenz verwendete Datensatz (hier die Teilbauflächen in Berlin und deren bebauungsstrukturbeschreibende Variablen) Besonderheiten aufweisen kann, welche die Verallgemeinerungsfähigkeit der Gruppeneinteilung schwächen können. Aufgrund der in Berlin

vorgefundenen Vielfalt an Bebauungsstrukturen und der Historie als „geteilte Stadt“ kann davon ausgegangen werden, dass die ermittelte Gruppeneinteilung als Referenz für eine nationale Quartierstypisierung verwendet werden kann. Dennoch ist anzumerken, dass zu diesem Thema allgemein sowie im Speziellen im Vergleich von Groß- mit Klein- und Mittelstädten weitere Untersuchungen einen zusätzlichen Erkenntnisgewinn liefern können.

Die zweite Möglichkeit der Typisierung von Quartieren unter Verwendung des nationalen Datenbestandes ist die Ausführung aller Schritte der Methodik auf einen aus den Geobasisdaten erzeugten nationalen Datensatz. Auch hier gilt es, die Datenvorverarbeitung auszuführen, bevor die Clusteranalyse, wie in Abschnitt 5.4 beschrieben, für den vollständigen nationalen Untersuchungsraum durchgeführt wird. Aufgrund der enormen Datenmenge (vgl. Kapitel 5.2) ist zurzeit nur die Ausführung der Clusteranalyse unter Verwendung einer Stichprobenziehung umsetzbar. Mittels Clusteranalyse erfolgt somit erneut die iterative Gruppenbildung für unterschiedliche Gruppenanzahlen und nachgelagert die Festlegung einer geeigneten Gruppenanzahl. Nach der Ausführung des Algorithmus auf den nationalen Datensatz der gruppierten Teilbauflächen können abschließend die statistischen Analysen durchgeführt werden, um den nationalen Quartiersbestand zahlenmäßig und hinsichtlich der Ausprägung von bebauungsstrukturbeschreibenden Merkmalen zu erfassen und die Kenngrößen der Typquartiere als Vertreter der Quartiere des deutschen Quartiersbestandes festzulegen.

Im Hinblick auf die Ausführbarkeit der voran formulierten Arbeitsschritte kann die in der Arbeit entwickelte Methodik ohne Änderungen ausgeführt werden. Wegen der hohen Datenmengen ist jedoch insbesondere bei der Klassifizierung der Teilbauflächen, d. h. bei der Zuordnung nicht gruppierter Objekte zu den mittels Clusteranalyse gebildeten Teilbauflächengruppen, mit sehr hohen Ausführungszeiten zu rechnen. Bei diesem Ansatz der Quartierstypisierung unter Verwendung des nationalen Geodatenbestandes wird zudem empfohlen, weiterführende Untersuchungen zum Einfluss der Gruppenbildung auf Basis einer Stichprobe durchzuführen.

Insgesamt kann attestiert werden, dass eine datenbasierte Typisierung von Quartieren anhand baulicher Strukturen für den nationalen Gesamtbestand unter Anwendung der entwickelten Methodik möglich ist und die für Berlin beispielhaft erarbeiteten empirischen Ergebnisse ebenfalls für den nationalen Quartiersbestand ermittelt werden können.

8 RESÜMEE UND AUSBLICK

8.1 Resümee

Die Zielsetzung der vorliegenden Arbeit bestand darin, eine Methodik zu entwickeln, welche die Bildung von Typquartieren als Vertreter einer Gruppe realer Quartiere ermöglicht und deren bauliche Struktur sowie deren Häufigkeit im deutschen städtischen Gebäudebestand nachvollzieht. Diese Methodik sollte eine großflächige oder gar flächendeckende Quantifizierung des Quartiersbestandes erlauben und sich deshalb auf eine flächendeckend verfügbare, den räumlichen Zusammenhang von Gebäuden abbildende Datenbasis stützen.

In *Kapitel 2* wurde zunächst eine umfassende Literaturrecherche vorgestellt und der Begriff des „Quartiers“ und seine Bedeutung in unterschiedlichen Forschungsrichtungen diskutiert, insbesondere im Kontext der Gebäudeenergieforschung. Basierend auf diesen Erkenntnissen und mit Bezug zu den Zielsetzungen dieser Arbeit wurde im Anschluss eine kontextbezogene Definition des Quartiersbegriffs vorgenommen. Darüber hinaus wurden bestehende Typologien zur Einteilung von Siedlungs- bzw. Raumstrukturen erläutert und deren methodische Konzepte beschrieben. Nach einer Darstellung verschiedener Softwarelösungen zur energetischen Bewertung von Quartierskonzepten wurden die Grenzen bestehender Typologien und Bewertungshilfen aufgezeigt. Diese Beschränkungen sollten bei der Methodenentwicklung im Rahmen der vorliegenden Arbeit weiterhin Beachtung finden.

In *Kapitel 3* erfolgte die Erarbeitung der methodischen Grundlagen im Hinblick auf Verfahren zur Gruppenbildung und Klassifizierung. Es wurden die Bandbreite existierender Clusteranalyseverfahren sowie deren Anwendungsfelder und Beschränkungen vorgestellt, sodass eine spätere Verfahrensauswahl begründet werden konnte. Neben der Clusteranalyse als Prozess der Gruppenbildung wurden ebenfalls Verfahren zur Klassifizierung von Objekten, sog. „Machine Learning“-Verfahren, gegenübergestellt. Der Fokus lag dabei auf Algorithmen, welche die im Rahmen der Methodenentwicklung zur Verfügung stehenden Daten verarbeiten können. Die Diskussion verschiedener Anwendungsbeispiele für Clusteranalyse- und Klassifizierungsverfahren im Kontext der Gebäudeenergie- bzw. Quartiersforschung diente ebenfalls einer Absicherung der Verfahrensauswahl sowie Parameterfestlegung im Rahmen der Entwicklung einer Methodik zur Quartierstypisierung.

In *Kapitel 4* wurden grundlegende Erläuterungen zu Geoinformationssystemen und -daten (GIS bzw. GIS-Daten) gegeben und eine Auswahl für die Arbeit relevanter, flächendeckender Datenangebote in Deutschland beschrieben. Weiterhin wurden verschiedene Einsatzbereiche der benannten GIS-Daten aufgezeigt und deren bisherige Nutzung in der gebäudebezogenen Energieforschung diskutiert. Im Hinblick auf die Verwendung von GIS-Daten als Datengrundlage der Quartierstypisierung und Quantifizierung wurde zudem auf verschiedene Einschränkungen verwiesen, welche sich aus der Verfügbarkeit gebäude-

bezogener Informationen und aus existierenden Freiheitsgraden bei der modellhaften Abbildung der Umwelt in GIS ergeben.

In *Kapitel 5* erfolgte zuerst die Beschreibung der für diese Arbeit vorliegenden Geodaten-sätze sowie der zur Datenverarbeitung eingesetzten Softwarelösungen. Es wurde verdeutlicht, dass die Handhabung der Geodaten und die zur Typisierung nötigen statistisch-infor-mationstechnischen Verfahren nicht durch eine einzelne Softwarelösung darstellbar sind und für die Datenanalyse und Auswertung eine Vielzahl an Algorithmen zu entwickeln war. Ebenfalls erfolgte eine Diskussion über die Notwendigkeit einer der Anwendung der Methodik zur Quartierstypisierung vorgelagerten Datenüberprüfung und -auswahl.

Diesen Ausführungen schloss sich die Entwicklung der Methodik zur datenbasierten Quar-tierstypisierung aufgrund baulicher Strukturen an. Es wurden vier Verfahrensschritte fest-gelegt, welche erforderlich waren, um die formulierten Ziele der Arbeit zu erreichen. Der erste Verfahrensschritt war die Analyse der geometrisch-topologischen Beziehung der Ge-bäudepopulationen auf einer kleinskaligen Raumebene, den Teilbauflächen. Zu diesem Zweck wurde ein Variablensatz definiert, welcher die in den verfügbaren Datensätzen ent-haltenen Informationen auf gebäudeenergiebedarfsrelevante sowie für die leitungsgebun-dene Wärmeversorgung relevante Merkmale einschränkt. Diese Analyse lieferte damit eine Beschreibung der lokalen Bebauungsstruktur und ermöglichte es die Teilbauflächen als Merkmalsträger für die Bildung von Gruppen ähnlicher Bebauungsstruktur zu verwenden. Im zweiten Verfahrensschritt erfolgte die Durchführung einer Clusteranalyse mit der Ziel-setzung, den Teilbauflächenbestand in eine innerhalb der Anwendung der Methodik fest-zustellende geeignete Gruppenanzahl einzuteilen. Es wurde dabei ein Vorgehen festgelegt, welches sowohl eine direkte Gruppenbildung für kleine bis mittelgroße Datensätze vorneh-men kann als auch bestehende informationstechnische Ressourcenbeschränkungen bei der Gruppeneinteilung besonders großer Datensätze berücksichtigt. Für letztgenannte Aus-gangssituation wurde die Gruppenbildung durch Clusteranalyse um eine nachgelagerte Klassifizierung ergänzt und so eine vollständige Gruppierung des Teilbauflächendatensat-zes ermöglicht.

Der dritte Verfahrensschritt der Methodik zur Quartierstypisierung beinhaltet die Vereini-gung von benachbarten Teilbauflächen gleicher Gruppenzugehörigkeit, d. h. ähnlicher Be-bauungsstruktur, sowie die Zuordnung von kleinflächigen Insel- bzw. Splittergebieten an-derer Gruppenzugehörigkeit zu Quartieren. Hierfür wurde ein Algorithmus vorgestellt, der die kleinräumigen Teilbauflächen zu größeren zusammenhängenden Gebieten vereinigt und dabei sowohl Hauptverkehrswege als auch andere Grenzelemente berücksichtigen kann. Die durch Anwendung des Algorithmus erzeugten Quartiere lagen in einer Gruppie-rung nach Bebauungsstruktur vor, welche der vorher erzeugten Gruppeneinteilung für Teil-bauflächen entsprach. Den vierten Verfahrensschritt bildeten die statistische Analyse des in Quartiersgruppen eingeteilten Quartiersbestandes und die statistische Beschreibung der

Quartiere innerhalb der Quartiersgruppen. Basierend auf diesen Auswertungen werden abschließend geometrisch-topologische Kenngrößen (z. B. die Gebäudegrundfläche oder ein gebäudegrundflächenbezogenes Maß der baulichen Nutzung der Quartiersfläche) festgelegt, welche je Quartiersgruppe einen Typvertreter – das Typquartier – beschreiben.

In *Kapitel 6* wurde die entwickelte Methodik beispielhaft auf den Geodatenatz der Stadt Berlin angewendet. Im Vorfeld dieser Methodenanwendung wurden zuerst eine Überprüfung der für die Arbeit verfügbaren Geodaten vorgenommen und Kriterien festgelegt, die eine Auswahl der Daten erlaubten und sich an den Zielsetzungen der Arbeit orientierten. Dieser Datenauswahl schloss sich die Anwendung der Methodik an. Nach der Datenvorverarbeitung zum Zweck der Erfassung der lokalen Bebauungsstrukturen erfolgte die Durchführung der Clusteranalyse. Anhand eines Gütekriteriums für die Einteilung von Gruppen, der mittleren Silhouetten-Weite, konnten sechs Teilbauflächengruppen als eine geeignete Anzahl bestimmt werden und somit eine Einteilung der Teilbauflächen in Berlin vorgenommen werden. Zur Sicherstellung der Anwendbarkeit für den nationalen Datensatz wurde ebenfalls das kombinierte Verfahren aus Clusteranalyse für Stichproben und eine nachgelagerte Gruppenzuordnung der Restobjekte (Klassifizierung) ausgeführt. Für zehn Stichproben wurde jeweils mittels mittlerer Silhouetten-Weite eine geeignete Gruppenanzahl bestimmt. Diese geeigneten Gruppenanzahlen sowie die Verläufe der mittleren Silhouetten-Weite wurden nun den auf der Clusteranalyse der Grundgesamtheit der Teilbauflächen in Berlin basierenden Ergebnissen gegenübergestellt. Es konnte dadurch gezeigt werden, dass trotz einer gewissen Varianz der stichprobenbasierten Ergebnisse eine gute Anwendbarkeit dieses Verfahrens vorliegt.

Der Gruppeneinteilung der Teilbauflächen in Berlin folgten deren Vereinigung zu Quartieren und die statistische Analyse, welche im Ergebnis ein Mengengerüst der Gebäude und Quartiere in Berlin sowie die Beschreibung der relevanten geometrisch-topologischen Kenngrößen für sechs Gruppen ähnlicher Bebauungsstruktur liefert. Basierend auf der statistischen Auswertung des Quartiersbestandes in Berlin wurden zudem Kenngrößen festgelegt, welche Typquartiere als Gruppenvertreter beschreiben und in weiterführenden Untersuchungen im Kontext der Energieforschung oder Planungspraxis verwendet werden können.

In *Kapitel 7* wurden die Möglichkeiten aufgezeigt, basierend auf den Ergebnissen der vorherigen Kapitel, eine Quartierstypisierung für den nationalen Geodatenbestand durchzuführen. Es konnte dargelegt werden, dass eine Studie unter Einbezug des nationalen Geodatenbestandes zwei möglichen Ansätzen folgen kann. So wurde verdeutlicht, dass zum einen die für den Untersuchungsraum Berlin ermittelte bebauungsstrukturbezogene Gruppierung von Teilbauflächen verwendet werden kann, um die Teilbauflächen eines gesamtdeutschen Datensatzes mittels Klassifizierungsalgorithmen in ebendiese Gruppen einzusortieren. Zum anderen besteht die Möglichkeit einer Ausführung der Methodik unter

Anwendung der stichprobenbasierten Clusteranalyse und der nachgelagerten Klassifizierung auf den nationalen Datenbestand. Somit würde einer neuen Gruppeneinteilung der Teilbauflächen erzeugt. Die Ausführungen in Kapitel 7 zeigen, dass durch beide beschriebenen Ansätze Mengengerüste, statistische Beschreibungen und Kenngrößen für Typquartiere erzeugt werden können, welche den nationalen Quartiersbestand umfassen und in vielfältiger Weise in Forschung und Praxis weiterverwendet werden können.

Der Zusammenfassung der vorliegenden Arbeit wird an dieser Stelle eine kurze Einordnung der hier vorgestellten Inhalte und Ergebnisse auf einer übergeordneten Ebene und bezogen auf den aktuellen Stand der gebäude- und quartiersbezogenen Energieforschung nachgestellt. Wie in den einleitenden Kapiteln der Arbeit erläutert, fand in den vergangenen Dekaden eine kontinuierliche Auseinandersetzung der Forschung und Planungspraxis mit dem Quartier als Untersuchungsgegenstand bzw. Handlungsraum statt. Zur Realisierung einer sowohl energetisch, ökonomisch als auch ökologisch effizienten Wärme- und Stromversorgung des deutschen Gebäudebestandes ist eine integrale Betrachtung verschiedener Systembestandteile von der Bauteil- bzw. Anlagenebene bis hin zur nationalen und internationalen Ebene notwendig. Das Quartier kann hierzu als passende Aggregationsebene dienen, die lokale Zusammenhänge abbildet und dabei im Vergleich zum Einzelgebäude hinreichend verallgemeinerungsfähig ist, da in der Quartiersbetrachtung auch netzgebundene Versorgungsstrukturen bewertet werden können. Die Bemühungen zur stärkeren Nutzung dieser intermediären Ebene spiegeln sich in den entsprechenden Förderformaten u. a. des Bundesministeriums für Wirtschaft und Klimaschutz (BMWK) wider. Es zeigt sich jedoch, dass eine Übertragung der Ergebnisse einzelner Vorhaben bisher nur begrenzt (von Quartier zu Quartier) erreicht wurde. Ein Ansatz zur übergeordneten Operationalisierung der Ergebnisse dieses Forschungszweiges wurde mit der Datenbasierung der hier beschriebenen Methodik erstmals formuliert. Die Verbreitung der Ergebnisse innerhalb des „Forschungsnetzwerks Energie“ des BMWK sowie im Rahmen der Formate der wissenschaftlichen Begleitforschung „Energiewendebauen“ wird daher angestrebt, um die Bewertbarkeit von Klimaschutzmaßnahmen im Quartier zu verbessern und deren Umsetzung zu beschleunigen.

8.2 Ausblick

Nachfolgend sollen als Ausblick zudem eine Reihe möglicher Anknüpfungspunkte für weiterführende Untersuchungen formuliert werden. Es handelt sich hierbei um Themen, welche den unterschiedlichen Forschungsdisziplinen entspringen, die in der Arbeit abgedeckt wurden, und von der Ebene der Grundlagenforschung bis hin zur praxisnahen, angewandten Forschung reichen.

Die im Rahmen dieser Arbeit erstmals gelieferte geodatenbasierte Empirie des Quartiersbestandes könnte zukünftig zusätzlich abgesichert werden, indem eine Verknüpfung mit bestehenden Erhebungsinstrumenten zum Gebäudebestand erfolgt. Hierbei sind insbesondere die am Institut Wohnen und Umwelt durchgeführten repräsentativen Stichprobenerhebungen zum Wohn- und Nichtwohngebäudebestand in Betracht zu ziehen. Im Rahmen zukünftiger Erhebungen könnte über den gebäudeenergiebezogenen Fokus der bisherigen Erhebungen hinaus zudem eine Erfassung der räumlich-topologischen Einordnung der Gebäude erfolgen (z. B. durch Abfrage, welchem Quartierstyp die Bebauungssituation entspricht, in der sich das Gebäude befindet). Hierdurch könnten die stichprobenbasierten Gebäudedaten und die geodatenbasierte Quartiersempirie besser kombiniert werden. Eine Nutzung von energetischen Merkmalen zur Beschreibung des Gebäudebestandes in Szenarioanalysen zum Quartiersbestand wäre folglich mit weniger Unsicherheiten behaftet. Des Weiteren könnte die Verknüpfung von energetischen Merkmalen der Objekte einer Gebäudestichprobe mit den verfügbaren Geodaten weiterführende Analysen zu regionalen und lokalen Dynamiken der Gebäudemodernisierung erlauben. In Verbindung mit ergänzenden sozio-ökonomischen Daten würde so eine Untersuchung von Zusammenhängen außerhalb der rein energetisch-baulichen Dimension der energetischen Modernisierung und Energieversorgung von Gebäuden und Quartieren möglich. Auch eine Anpassung des bisherigen Stichprobendesigns ist denkbar, um neben der Gebäudeebene auch das Quartier als Träger energierelevanter Merkmale zu berücksichtigen. Eine solche Anpassung wäre z. B. die Stichprobenziehung unter Einbezug der mit der Methodik zur Quartierstypisierung gebildeten Teilbauflächen oder Quartiere als Erhebungsbezirke. Für jede dieser lokalen Gebäudedeformationen sollte dann möglichst eine Vollerhebung angestrebt werden, sodass beide Merkmalsebenen (Gebäude und Quartier) vollständig erfasst und beschreibbar sind.

Über das Handlungsfeld der empirischen Fundierung hinaus ist die Anwendung der Methodik auf den nationalen Geodatenbestand mittels Durchführung einer erneuten Clusteranalyse für die Gruppenbildung als nächster Schritt anzusehen, welcher im Nachgang der hier vorgestellten Grundlagenarbeit erfolgen wird. Die so gewonnene Quartierseinteilung kann im Anschluss daran mit den Typquartierskennzahlen für das Anwendungsbeispiel Berlin verglichen werden. Die nationale Quartierstypisierung ermöglicht auch einen kommunalen bzw. regionalen Vergleich von Gebäude- und Quartiersbeständen. Dieser kann hilfreich sein, um die Verteilung von (ökonomischen) Lasten, die sich aus der Notwendigkeit einer Dekarbonisierung der Gebäudewärmeversorgung ergeben, besser zu verstehen und ggf. weiterführende Steuerungsinstrumente für die Wärmewende im Gebäude und Quartier zu entwickeln.

Zudem können die Ergebnisse der Quartierstypisierung auch für die in den vergangenen Monaten und Jahren in den Fokus gerückte kommunale Wärmeplanung, als Instrument der raumbezogenen Energieversorgungsplanung, einen Mehrwert liefern. Zum einen sind mit

den Ergebnissen der vorliegenden Arbeit bzw. einer erneuten Quartierstypisierung für lokale Untersuchungsräume Szenarioanalysen und Potenzialstudien möglich, welche die Erstellung von kommunalen Wärmeplänen unterstützen können. Zum anderen bietet sich die Verwendung von Typquartieren, welche sich aus flächendeckenden oder nationalen Datenbanken ableiten, an, um eine begleitende Qualitätskontrolle in der Wärmeplanung zu etablieren. So könnten energetische, ökologische und ökonomische Kennzahlen von Energieversorgungskonzepten für Typquartiere als Benchmarks der individuellen Planungen von Kommunen dienen und eine Identifikation von Energiekonzepten ermöglichen, die im Vergleich mit anderen Versorgungskonzepten als i. d. R nicht hinreichend (kosten)effizient zu betrachten sind.

Die zusammenfassende Beschreibung der hier vorliegenden Arbeit sowie der Ausblick auf mögliche Erweiterungen, ergänzende Untersuchungen und weitere Anwendungsfälle verdeutlicht, dass diese Arbeit einen neuartigen Beitrag zur energetischen Bewertung von Gebäuden in ihrem räumlichen Zusammenhang liefert: Aus der Entwicklung der Methodik resultiert dabei ein flexibles Verfahren, das im Hinblick auf unterschiedliche Raumebenen sowie Forschungsfragen aus unterschiedlichen Handlungs- bzw. Forschungsfeldern eingesetzt werden kann. Die erarbeitete Einteilung von Quartieren am Beispiel Berlin und deren Quantifizierung sowie statistische Analyse liefern zudem ein Zahlengerüst für die Verknüpfung der Gebäude-, Quartiers- und Stadtebene unter Berücksichtigung des räumlichen Zusammenhangs des lokalen Gebäudebestandes und ermöglichen unter Zuhilfenahme vorhabenspezifischer Modellannahmen die Bewertung des Beitrags von Quartierskonzepten zur Reduktion gebäudebezogener Treibhausgasemissionen.

LITERATURVERZEICHNIS

- Abdolhosseini Qomi, Mohammad Javad; Noshadravan, Arash; Sobstyl, Jake M.; Toole, Jameson; Ferreira, Joseph; Pellenq, Roland J-M et al. (2016):** Data analytics for simplifying thermal efficiency planning in cities. In: *Journal of the Royal Society, Interface* 13 (117). DOI: 10.1098/rsif.2015.0971.
- AdV (2008):** Dokumentation zur Modellierung der Geoinformationen des amtlichen Vermessungswesens (GeoInfoDok). Hauptdokument. Version 6.0. Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV).
- AdV (2018a):** AdV-Produktspezifikation ATKIS-Basis-DLM-Shape. Version 1.1. Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV).
- AdV (2018b):** ALKIS-Objektartenkatalog DLKM. Version 7.1 rc.1. Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV).
- AdV (2019):** AdV-Produktspezifikation ALKIS-WFS und Ausgabeformate (Shape, CSV). Version 2.0.0. Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV) (AdV-Produktspezifikation ALKIS-WFS).
- AdV (2021a):** Produkt- und Qualitätsstandard für 3D-Gebäudemodelle. Version 2.2. Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV).
- AdV (2021b):** CityGML Codeliste des Attributs „BuildingFunctionType“. Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV).
- Agugiaro, Giorgio; Benner, Joachim; Cipriano, Piergiorgio; Nouvel, Romain (2018):** The Energy Application Domain Extension for CityGML: enhancing interoperability for urban energy simulations. In: *Open Geospatial Data, Software and Standards* 3 (1). DOI: 10.1186/s40965-018-0042-y.
- Akhanli, Serhat Emre; Hennig, Christian (2020):** Comparing clusterings and numbers of clusters by aggregation of calibrated clustering validity indexes. In: *Statistics and Computing* 30, S. 1523–1544. DOI: 10.1007/s11222-020-09958-2.
- Albers, Gerd; Wékel, Julian (2021):** Stadtplanung. Eine illustrierte Einführung. 4. Auflage. Darmstadt: wbg Academic.

Ali, Usman; Shamsi, Mohammad Haris; Hoare, Cathal; Mangina, Eleni; O'Donnell, James (2021): Review of urban building energy modeling (UBEM) approaches, methods and tools using qualitative and quantitative analysis. In: *Energy and Buildings* 246, S. 111073. DOI: 10.1016/j.enbuild.2021.111073.

Ankerst, Mihael; Breunig, Markus M.; Kriegel, Hans-Peter; Sander, Jörg (1999): OPTICS: Ordering Points To Identify the Clustering Structure. In: *Proc. ACM SIGMOD'99 Int. Conf. on Management of Data, Philadelphia PA, 1999*.

Arbelaitz, Olatz; Gurrutxaga, Ibai; Muguerza, Javier; Pérez, Jesús M.; Perona, Iñigo (2013): An extensive comparative study of cluster validity indices. In: *Pattern Recognition* 46 (1), S. 243–256. DOI: 10.1016/j.patcog.2012.07.021.

Arens, Tilo; Hettlich, Frank; Karpfinger, Christian; Kockelkorn, Ulrich; Lichtenegger, Klaus; Stachel, Hellmuth (2018): Mathematik. 4. Aufl. 2018. Berlin, Heidelberg: Springer Berlin Heidelberg.

Backhaus, Klaus; Erichson, Bernd; Plinke, Wulff; Weiber, Rolf (2018): Multivariate Analysemethoden. Eine anwendungsorientierte Einführung. 15., vollständig überarbeitete Auflage. Berlin, Heidelberg: Springer Gabler.

Bagherian, Behrooz; Koch, Thilo; Werner, Peter; Schmidt, Manfred; Biehl, Michael (2017): Energetisches Quartierskonzept Alte Schmelz - St. Ingbert. Unter Mitarbeit von Navid Sanaie, Eugen Rejek und Wolfgang Rejek. Institut Wohnen und Umwelt. Darmstadt.

Basille, Mathieu; Bucklin, David (2019): Package ‚rpostgis‘. R Interface to a ‘PostGIS’ Database.

Bayerische Vermessungsverwaltung (2018): Kundeninformation LoD2 Gebäudemodelle. Bayerische Vermessungsverwaltung.

BCG (2021): Klimapfade 2.0. Ein Wirtschaftsprogramm für Klima und Zukunft. Gutachten für den BDI. Boston Consulting (Group).

BDEW (2021): Foliensatz zur BDEW-Publikation Entwicklung des Wärmeverbrauchs in Deutschland. Basisdaten und Einflussfaktoren. Bundesverband der Energie- und Wasserwirtschaft (BDEW).

Begleitforschung EWB (2020): Energiewendebauen. Forschungserkenntnisse von der Komponente bis zum Quartier. Stuttgart: Fraunhofer IRB Verlag.

Begleitforschung EWB (2022): Quartier und Stadt. Wissenschaftliche Begleitforschung Energiewendebauen. Online verfügbar unter <https://www.energiewendebauen.de/themen/quartier-stadt>, zuletzt geprüft am 09.03.2022.

- Behnisch, Martin (2008):** Urban Data Mining. Operationalisierung der Strukturerkennung und Strukturbildung von Ähnlichkeitsmustern über die gebaute Umwelt. Dissertation. Universität Karlsruhe, Karlsruhe. Fakultät für Architektur.
- Beier, Carsten; Wrobel, Patrick; Schnier, Matthias; Kanngießer, Annedore; Erhorn, Hans; Erhorn-Kluttig, Heike et al. (2016):** EnEff:Stadt. Wissenschaftliche Begleitung der Förderinitiative „Energieeffiziente Stadt“ Phase 3. Schlussbericht. Fraunhofer-Institut für Umwelt-, Sicherheits- und Energietechnik UMSICHT; Fraunhofer-Institut für Bauphysik; Jank, R.; pro:21; Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung. Oberhausen.
- Bell, Jason (2014):** Machine Learning. Indianapolis, IN, USA: John Wiley & Sons, Inc.
- Bengtsson, Henrik (2021):** A Unifying Framework for Parallel and Distributed Processing in R using Futures. In: *The R Journal* 13 (2), S. 208. DOI: 10.32614/RJ-2021-048.
- Bernard, Lars (Hg.) (2005):** Geodateninfrastruktur. Grundlagen und Anwendungen. Heidelberg: Wichmann.
- Bernard, Lars; Brauner, Johannes; Mäs, Stephan; Wiemann, Stefan (2019):** Geodateninfrastrukturen. In: Sester, Monika (Hg.): Geoinformatik. Handbuch der Geodäsie, herausgegeben von Willi Freeden und Reiner Rummel. Berlin, Heidelberg: Springer Spektrum (Springer Reference Naturwissenschaften), S. 91–122.
- Berthold, Michael R.; Borgelt, Christian; Höppner, Frank; Klawonn, Frank; Silipo, Rosaria (2020):** Guide to Intelligent Data Science. Cham: Springer International Publishing.
- Bill, Ralf (2010):** Grundlagen der Geo-Informationssysteme. 5., völlig neu bearb. Aufl. Berlin, Offenbach: Wichmann.
- Bill, Ralf (2018):** Geoinformation / Geoinformationssysteme (GIS). In: ARL (Hg.): Handwörterbuch der Stadt- und Raumentwicklung. Ausgabe 2018. Hannover: Akademie für Raumforschung und Landesplanung.
- Bivand, Roger; Keitt, Tim; Rowlingson, Barry (2022):** Package ‘rgdal’. Bindings for the ‘Geospatial’ Data Abstraction Library.
- Bivand, Roger; Pebesma, Edzer J.; Gómez-Rubio, Virgilio (2013):** Applied Spatial Data Analysis with R. 2nd ed. 2013. New York, NY: Springer New York (Use R! Ser, v.10).
- BKG (2018):** Kommunale Teilgebiete 1:25 000. KT25. Dokumentation. Bundesamt für Kartographie und Geodäsie (BKG). Leipzig.

- BKG (2020):** Hinweise zur Verwaltungsgliederung. Anhang zur Spezifikation der VG-Produkte. Gültig ab 31.12.2020. Bundesamt für Kartographie und Geodäsie (BKG).
- BKG (2021a):** Georeferenzierte Adressdaten. GA. Dokumentation. Bundesamt für Kartographie und Geodäsie (BKG). Leipzig.
- BKG (2021b):** Verwaltungsgebiete 1:25 000. VG25. Dokumentation. Bundesamt für Kartographie und Geodäsie (BKG). Leipzig.
- BKG (2021c):** Hauskoordinaten Deutschland. HK-DE. Bundesamt für Kartographie und Geodäsie (BKG). Leipzig.
- BKG (2021d):** Hausumringe Deutschland. HU-DE. Dokumentation. Bundesamt für Kartographie und Geodäsie (BKG). Leipzig.
- BKG (2022):** Digitales Basis-Landschaftsmodell. Basis-DLM. Dokumentation. Bundesamt für Kartographie und Geodäsie (BKG). Leipzig.
- Blesl, Markus (2002):** Räumlich hoch aufgelöste Modellierung leitungsgebundener Energieversorgungssysteme zur Deckung des Niedertemperaturwärmebedarfs. Dissertation. Stuttgart: IER Bibliothek (Forschungsbericht, Bd. 92).
- Blömer, Sebastian (2021):** Wärmetlas 2.0. Modellbeschreibung. Institut für Energie- und Umweltforschung; GEF Ingenieur AG; geomer GmbH. Heidelberg.
- BMF (2019):** Klimaschutzprogramm 2030 der Bundesregierung zur Umsetzung des Klimaschutzplans 2050. Bundesministerium für Finanzen. Berlin.
- BMI (2005):** Bericht der Bundesregierung über die Fortschritte zur Entwicklung der verschiedenen Felder des Geoinformationswesens im nationalen, europäischen und internationalen Kontext. 1. Geo-Fortschrittsbericht der Bundesregierung. Bundesministerium des Innern (BMI). Berlin.
- BMI (2021a):** Räumliche Informationen im Alltag. Fünfter Bericht der Bundesregierung über die Fortschritte zur Entwicklung der verschiedenen Felder des Geoinformationswesens im nationalen, europäischen und internationalen Kontext. 5. Geo-Fortschrittsbericht der Bundesregierung. Bundesministerium des Innern, für Bau und Heimat (BMI). Berlin.
- BMI (2021b):** Open-Data-Strategie der Bundesregierung. Bundesministerium des Innern, für Bau und Heimat (BMI). Berlin.
- BMU (2019):** Klimaschutzplan 2050. Klimaschutzpolitische Grundsätze und Ziele der Bundesregierung. Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit (BMU). Berlin.
- BMU (2021):** Klimaschutzbericht 2021. Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit (BMU). Berlin.

- BMUB (2016):** Übereinkommen von Paris. Ins Deutsche übersetzte Textfassung des Klimaabkommens von Paris zum UNFCCC. Bundesministerium für Umwelt, Naturschutz, Bau und Reaktorsicherheit (BMUB). Berlin.
- BMWi (2018):** 7. Energieforschungsprogramm. Innovationen für die Energiewende. Bundesministerium für Wirtschaft und Energie. Berlin.
- BMWi (2021):** Förderkonzept „Reallabore der Energiewende“. Bundesministerium für Wirtschaft und Energie. Berlin.
- Boehmke, Bradley (2022):** Naïve Bayes Classifier. UC Business Analytics R Programming Guide. University of Cincinnati, Center for Business Analytics. Online verfügbar unter https://uc-r.github.io/naive_bayes, zuletzt aktualisiert am 04.03.2020, zuletzt geprüft am 24.02.2022.
- Bömermann, Hartmut (2012):** Stadtgebiet und Gliederungen. In: *Zeitschrift für amtliche Statistik Berlin Brandenburg* (1+2), S. 76–87.
- Bömermann, Hartmut (2014):** Berlin kleinräumig. In: *Zeitschrift für amtliche Statistik Berlin Brandenburg* (2), S. 20–22.
- Borah, B.; Bhattacharyya, D. K. (2004):** An improved sampling-based DBSCAN for large spatial databases. In: Proceedings of International Conference on Intelligent Sensing and Information Processing 2004. International Conference on Intelligent Sensing and Information Processing, 2004. Chennai, India, 04.–07. Januar 2004: IEEE, S. 92–96.
- Bott, Helmut (2018):** Das Quartier. In: Bott, Helmut; Grassl, Gregor und Anders, Stephan (Hg.): Nachhaltige Stadtplanung: DETAIL, S. 21–24.
- Bott, Helmut; Siedentrop, Stefan (2018):** Handlungsfeld Regional-, Stadt- und Quartierentwicklung. In: Bott, Helmut; Grassl, Gregor und Anders, Stephan (Hg.): Nachhaltige Stadtplanung: DETAIL.
- Budde, Eike; Ecker-Brinkmann, Astrid; Erhorn, Hans; Erhorn-Kluttig, Heike; Koutsomarkos, Konstantinos; Lyslow, Linda; Wössner, Simon (2018):** District Energy Concept Adviser. Edition MODER 2018. User Manual. Fraunhofer-Institut für Bauphysik. Stuttgart.
- BUKEA (2019):** Wärmekataster Handbuch. Behörde für Umwelt, Klima, Energie und Agrarwirtschaft (BUKEA), Hansestadt Hamburg. Hamburg.
- KSG, vom 18.08.2021:** Bundes-Klimaschutzgesetz (KSG).
- Busch, Roland; Müller, Ann-Katrin (2020):** Erhebungsmerkmale und Merkmalsausprägungen im Screening. Projektinformation. DataNWG-Projektinfo 3. Bergische Universität Wuppertal. Wuppertal.

- BWP (2021):** Branchenstudie 2021. Marktanalyse, Szenarien, Handlungsempfehlungen. Bundesverband Wärmepumpe (BWP). Berlin.
- Cao, Zechun; Wang, Sujing; Forestier, Germain; Puissant, Anne; Eick, Christoph F. (2013):** Analyzing the composition of cities using spatial clustering. In: Zheng, Yu.; Koonin, Steven E. und Wolfson, Ouri E. (Hg.): UrbComp '13. Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing. 2nd ACM SIGKDD International Workshop. Chicago, Illinois, 11.08.2013. New York, New York, USA: ACM Press.
- Castells, Manuel (1983):** The city and the grassroots. A cross-cultural theory of urban social movements. London: Arnold.
- Cerezo Davila, Carlos; Reinhart, Christoph F.; Bemis, Jamie L. (2016):** Modeling Boston: A workflow for the efficient generation and maintenance of urban building energy models from existing geospatial datasets. In: *Energy* 117, S. 237–250. DOI: 10.1016/j.energy.2016.10.057.
- Charrad, Malika; Ghazzali, Nadia; Boiteau, Véronique; Niknafs, Azam (2014):** NbClust : An R Package for Determining the Relevant Number of Clusters in a Data Set. In: *J. Stat. Soft.* 61 (6). DOI: 10.18637/jss.v061.i06.
- Chen, Yixing; Hong, Tianzhen; Piette, Mary Ann (2017):** Automatic generation and simulation of urban building energy models based on city datasets for city-scale building retrofit analysis. In: *Applied Energy* 205, S. 323–335. DOI: 10.1016/j.apenergy.2017.07.128.
- Cischinsky, Holger; Diefenbach, Nikolaus (2018):** Datenerhebung Wohngebäudebestand 2016. Datenerhebung zu den energetischen Merkmalen und Modernisierungsraten im deutschen und hessischen Wohngebäudebestand. 1. Aufl. Hg. v. Institut Wohnen und Umwelt. Institut Wohnen und Umwelt. Darmstadt.
- Cischinsky, Holger; Hörner, Michael; Rodenfels, Markus (2021):** Stichprobe: Modellierung und Ziehung. ENOB: dataNWG - Forschungsdatenbank Nichtwohngebäude. Zwischenbericht. Institut Wohnen und Umwelt. Darmstadt (E.1.4.5).
- Conway, Joe; Eddelbuettel, Dirk; Nishiyama, Tomoaki; Kumar Prayaga, Sameer; Tiffin, Neil (2022):** Package ‘RPostgreSQL’. R Interface to the ‘PostgreSQL’ Database System.
- Coors, Volker (Hg.) (2005):** 3D-Geoinformationssysteme. Grundlagen und Anwendungen. Heidelberg: Wichmann.
- Coors, Volker (2018):** SIMSTADT 2.0. 3D Simulation urbaner Energiesysteme unter Nutzung von Sensordaten. Foliensatz vom 15. Sächsischen GIS-Forum, Dresden, 30.01.2018. Dresden.

- Coors, Volker; Andrae, Christine; Böhm, Karl-Heinz (2016):** 3D-Stadtmodelle. Konzepte und Anwendungen mit CityGML. Berlin, Offenbach: Wichmann.
- Cummings, Veronika; Meisel, Uli (Hg.) (2013):** StadtQuartiere. Sozialwissenschaftliche, ökonomische und städtebaulich-architektonische Perspektiven. 1. Aufl. Essen, Ruhr: Klartext-Verlag.
- dena (2021):** dena Gebäudereport 2022. Zahlen, Daten, Fakten. Unter Mitarbeit von Simon Becker, Jonas Hagen, Rico Krüger, Heike Marcinek, Christina Stahl und Christian Stolte. Hg. v. Deutsche Energie-Agentur (dena). Berlin.
- DESTATIS (2020):** Statistik der Baugenehmigungen. Qualitätsbericht. Statistisches Bundesamt (DESTATIS). Wiesbaden.
- DESTATIS; WZB; BiB (2021):** Datenreport 2021. Statistisches Bundesamt (DESTATIS); Wissenschaftszentrum Berlin für Sozialforschung (WZB); Bundesinstitut für Bevölkerungsforschung (BiB). Bonn.
- Dettmar, Jörg; Drebes, Christoph; Sieber, Sandra (2020):** Energetische Stadtraumtypen. Strukturelle und energetische Kennwerte von Stadträumen. 2., überarbeitete und erweiterte Auflage. Stuttgart: Fraunhofer IRB Verlag.
- Deutscher Städtetag (1991):** Kommunale Gebietsgliederung. Empfehlungen zur Ordnung des Straßen-/Hausnummernsystems und zur Gliederung des Gemeindegebiets nach Gemeindeteilen, Blöcken und Blockseiten sowie DV-Organisation. Köln: Dt. Städtetag (Deutscher Städtetag Reihe H, DST-Beiträge zur Informationsgesellschaft und Stadtforschung. Reihe H, DST-Beiträge zur Statistik und Stadtforschung, 39).
- Diefenbach, Nikolaus; Cischinsky, Holger; Rodenfels, Markus (2010):** Datenbasis Gebäudebestand. Datenerhebung zur energetischen Qualität und zu den Modernisierungstrends im deutschen Wohngebäudebestand. 1. Aufl. Hg. v. Institut Wohnen und Umwelt. Institut Wohnen und Umwelt. Darmstadt.
- DIN V 18599-1:2018-09:** Energetische Bewertung von Gebäuden – Berechnung des Nutz-, End- und Primärenergiebedarfs für Heizung, Kühlung, Lüftung, Trinkwarmwasser und Beleuchtung. DIN Deutsches Institut für Normung e.V. Berlin.
- DIN V 4108-6:2003-06:** Wärmeschutz und Energie-Einsparung in Gebäuden. Teil 6: Berechnung des Jahresheizwärme- und des Jahresheizenergiebedarfs. DIN Deutsches Institut für Normung e.V. Berlin.
- Dochev, Ivan; Gorzalka, Philip; Weiler, Verena; Estevam Schmiedt, Jacob; Linkiewicz, Magdalena; Eicker, Ursula et al. (2020):** Calculating urban heat demands. An analysis of two modelling approaches and remote sensing for input data and validation. In: *Energy and Buildings* 226, S. 110378. DOI: 10.1016/j.enbuild.2020.110378.

- Dochev, Ivan; Munoz Hidalgo, M. E.; Seller, Hannes; Peters, Irene (2018):** Assigning energetic archetypes to a digital cadastre and estimating building heat demand. An example from Hamburg, Germany. Working Papers. Arbeitsgebiet Infrastrukturplanung und Stadttechnik, HafenCity Universität Hamburg.
- Dochev, Ivan; Seller, Hannes; Peters, Irene (2019):** Spatial aggregation and visualisation of urban heat demand using graph theory. In: *International Journal of Sustainable Energy Planning and Management* 24, S. 115–124.
- Drost, Hajk-Georg (2018):** Philentropy: Information Theory and Distance Quantification with R. In: *JOSS* 3 (26), S. 765. DOI: 10.21105/joss.00765.
- Dümbgen, Lutz (2016):** Einführung in die Statistik. Basel: Springer Basel (Mathematik kompakt).
- Eckle-Köhler, Judith; Köhler, Michael (2017):** Eine Einführung in die Statistik und ihre Anwendungen. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Ehlers, Manfred; Schiewe, Jochen (2012):** Geoinformatik. Darmstadt: WBG (Geowissen kompakt).
- Eicker, Ursula; Schumacher, Jürgen (2012):** Simulation erneuerbarer Energiesysteme mit INSEL 8. In: *Sonnenenergie* (3), S. 50–51.
- Erhorn, Hans; Erhorn-Kluttig, Heike; Schrade, Johannes; Schmidt, Dietrich; Sager, Christina; Beier, Carsten et al. (2010):** Energetische Quartiersplanung in Deutschland. Wissenschaftliche Begleitung der Förderaktivität „Energieeffiziente Stadt“ - Schlussbericht Phase 1. Fraunhofer-Institut für Bauphysik; Fraunhofer-Institut für Umwelt-, Sicherheits- und Energietechnik UMSICHT; Volkswohnungen GmbH; pro:21. Stuttgart, Oberhausen, Bonn, Karlsruhe.
- Erhorn-Kluttig, Heike; Erhorn, Hans (2016):** Energetische Bilanzierung von Quartieren. Ergebnisse und Benchmarks aus Pilotprojekten - Forschung zur energieeffizienten Stadt. Stuttgart: Fraunhofer IRB Verlag (Schriftenreihe EnEff).
- Erhorn-Kluttig, Heike; Erhorn, Hans; Weber, Juri; Wössner, Simon; Budde, Eike (2013):** The District Energy Concept Adviser. A software tool to support urban decision makers in planning district energy supply schemes. In: ECEEE (Hg.): Rethink, renew, restart. Eceee summer study proceedings ; 3–8 June 2013, Belambra Les Criques, Toulon-Hyères, France. Stockholm, S. 721–727.
- Erhorn-Kluttig, Heike; Jank, Reinhard; Schrempf, Ludger; Dütz, Armand (2011):** Energetische Quartiersplanung. Methoden – Technologien – Praxisbeispiele. Stuttgart: Fraunhofer-IRB-Verl.

EU Kommission (2020): A Renovation Wave for Europe. Greening our buildings, creating jobs, improving lives. Communication from the Commission to the European Parliament, the Council, THE European Economic and Social Committee and the Committee of the Regions. Europäische Kommission. Brüssel.

Everding, Dagmar; Kloos, Michael (Hg.) (2007): Solarer Städtebau. Vom Pilotprojekt zum planerischen Leitbild. Stuttgart: Kohlhammer.

Everding, Dagmar; Lindner, Sigrid (2007): Solarer Rahmenplan Berlin. Beispiel der umweltpädagogischen Visualisierung einer PV-Anlage. In: *Bundesbaublatt* (10), S. 38–42.

Feldmann, Philipp (2009): Die strategische Entwicklung neuer Stadtquartiere unter besonderer Berücksichtigung innenstadtnaher oder innerstädtischer, brachgefallener Industrieareale. Dissertation. Universität Regensburg, Regensburg. International Real Estate Business School.

Forsyth, David (2019): Applied Machine Learning. 1st ed. 2019. Cham: Springer International Publishing (Springer eBook Collection).

Foss, Alex; Markatou, Marianthi; Ray, Bonnie; Heching, Aliza (2016): A semiparametric method for clustering mixed data. In: *Mach Learn* 105 (3), S. 419–458. DOI: 10.1007/s10994-016-5575-7.

Fraunhofer IBP (2018): District Energy Concept Adviser. Edition MODER 2018. Version 1.0.0.35. Stuttgart.

Fraunhofer-Gesellschaft: DistrictECA. Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V. Online verfügbar unter <https://www.district-eca.de/index.php?lang=de>, zuletzt geprüft am 07.03.2022.

Freudenau, Henrik; Siebert, Sebastian; Bußkamp, Maik; Bosch-Lewandowski, Simone; Ganser, Robin; Krug, Henning; Runge, Marten (2021): Neue Stadtquartiere - Konzepte und gebaute Realität. Bonn: Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR) im Bundesamt für Bauwesen und Raumforschung (BBR) (BBSR-Online-Publikation, 4/2021).

Freudenberg, Jens; Meyer, Hauke; Bäumer, Thomas; Huber, Stephanie; Popovic, Tobias; Schneider Gräfin zu Lynar, Uta et al. (2019): Das Quartier als Schlüssel zur Steigerung der Sanierungsrate. Erkenntnisse aus dem ‚Drei Prozent Projekt - energieeffizienter Sanierungsfahrplan für kommunale Quartiere 2050‘. Deutscher Verband für Wohnungswesen, Städtebau und Raumordnung; Hochschule für Technik Stuttgart; Beratungs- und Service-Gesellschaft Umwelt. Berlin.

Galster, George (1986): What is a Neighborhood? An Externality Space Approach. In: *International Journal of Urban and Regional Research* (10), S. 243–261.

- Galster, George (2001):** On the Nature of Neighbourhood. In: *Urban Studies* 38 (12), S. 2111–2124.
- GDI-DE (2015):** Nationale Geoinformations-Strategie. Die Welt mit Geoinformationen im Jahr 2025. Version 1.0. Arbeitsgruppe NGIS des Lenkungsgremium Geodateninfrastruktur Deutschland (GDI-DE).
- Gerstenberg, Josef (2019):** Fortschreibung des Umweltatlas. Daten und Karten zur „Städtebaulichen Dichte“. Technische Dokumentation. Senatsverwaltung für Stadtentwicklung und Wohnen Berlin. Berlin.
- GeoZG, vom 25.02.2021:** Gesetz über den Zugang zu digitalen Geodaten (GeoZG).
- GEG, vom 08.08.2020:** Gesetz zur Einsparung von Energie und zur Nutzung erneuerbarer Energien zur Wärme- und Kälteerzeugung in Gebäuden (Gebäudeenergiegesetz - GEG).
- Giordani, Paolo; Ferraro, Maria Brigida; Martella, Francesca (2020):** An Introduction to Clustering with R. Singapore: Springer Singapore (1).
- Gower, John C. (1971):** A General Coefficient of Similarity and Some of Its Properties. In: *Biometrics* 27 (4), S. 857. DOI: 10.2307/2528823.
- Gröger, Gerhard; Kolbe, Thomas H.; Nagel, Claus; Häfele, Karl-Heinz (Hg.) (2012):** OGC City Geography Markup Language (CityGML) Encoding Standard. OGC 12-019. Version 2.0.0. Open Geospatial Consortium (OGC).
- Gruber, Ulrich; Recklinghausen, Kreis (2020):** Official 3D Spatial Base Data in Germany. In: *PFG* 88 (1), S. 25–31. DOI: 10.1007/s41064-020-00097-x.
- GBO, vom 05.10.2021 (26.05.1994):** Grundbuchordnung.
- Han, Jiawei; Kamber, Micheline; Pei, Jian (2012):** Data mining. Concepts and techniques. 3rd ed. Amsterdam, Boston: Elsevier/Morgan Kaufmann (The Morgan Kaufmann series in data management systems).
- Handl, Andreas; Kuhlenkasper, Torben (2018):** Einführung in die Statistik. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Hartmann, André; Behnisch, Martin; Hecht, Robert; Meinel, Gotthard; Schorcht, Martin; Schwarz, Steffen (2020a):** Zwischenbericht Gebäudemerkmale. Teilbericht E1.4.3 im Projekt ENOB:dataNWG - Forschungsdatenbank Nichtwohngebäude. Leibniz-Institut für ökologische Raumentwicklung. Dresden.
- Hartmann, André; Hörner, Michael; Rodenfels, Markus (2020b):** Berechnung der Hüllflächen aller Gebäude in der Stichprobe. Projektinformation. DataNWG-Projektinfo 6. Leibniz-Institut für ökologische Raumentwicklung. Dresden.

- Hartmann, André; Meinel, Gotthard; Hecht, Robert; Behnisch, Martin (2016):** A Workflow for Automatic Quantification of Structure and Dynamic of the German Building Stock Using Official Spatial Data. In: *IJGI* 5 (8), S. 142. DOI: 10.3390/ijgi5080142.
- Hecht, Robert (2014):** Automatische Klassifizierung von Gebäudegrundrissen. Ein Beitrag zur kleinräumigen Beschreibung der Siedlungsstruktur. Dissertation. Berlin: Rhombos-Verl. (IÖR-Schriften, 63).
- Hedderich, Jürgen; Sachs, Lothar (2020):** Angewandte Statistik. Methodensammlung mit R. 17., überarbeitete und ergänzte Auflage. Berlin, Heidelberg: Springer Berlin Heidelberg (Springer eBook Collection).
- Hegger, Manfred; Dettmar, Jörg; Martin, Alexander; Meinberg, Thomas; Drebes, Christoph; Hesse, Ulrike et al. (2012):** UrbanReNet. Schlussbericht. Fachgebiet EEB, FB Architektur, Technische Universität Darmstadt. Darmstadt.
- Hegger, Manfred; Dettmar, Jörg; Meinberg, Thomas; Drebes, Christoph; Schulze, Joachim; Sieber, Sandra; Sylla, Ousmane (2015a):** UrbanReNet Phase II. Schlussbericht. Fachgebiet EEB, FB Architektur, Technische Universität Darmstadt. Darmstadt.
- Hegger, Manfred; Dettmar, Jörg; Meinberg, Thomas; Drebes, Christoph; Schulze, Joachim; Sieber, Sandra; Sylla, Ousmane (2015b):** UrbanReNet Phase II. Schlussbericht Anlage III. Guidebook zum Softwaretool. Fachgebiet EEB, FB Architektur, Technische Universität Darmstadt. Darmstadt.
- Hennig, Christian (2015):** Clustering strategy and method selection. In: Hennig, Christian; Meila, Marina; Murtagh, Fionn und Rocci, Roberto (Hg.): *Handbook of Cluster Analysis*. Boca Raton, London, New York: Chapman & Hall; CRC Press (Handbooks of Modern Statistical Methods).
- Hennig, Christian (2020):** Package ‘fpc’. Flexible Procedures for Clustering.
- Hennig, Christian (2022):** An empirical comparison and characterisation of nine popular clustering methods. In: *Adv Data Anal Classif* 16 (1), S. 201–229. DOI: 10.1007/s11634-021-00478-z.
- Hennig, Christian; Meila, Marina; Murtagh, Fionn; Rocci, Roberto (Hg.) (2015):** *Handbook of Cluster Analysis*. Boca Raton, London, New York: Chapman & Hall; CRC Press (Handbooks of Modern Statistical Methods).
- Hentschel, Sabine (2017):** Nachhaltige Quartiere im Bestand. Entwicklung von Instrumenten zur Bewertung und Optimierung von bestehenden Stadtquartieren im Hinblick auf Nachhaltigkeit am Beispiel der Stadt Essen. Dissertation. Universität Duisburg-Essen, Duisburg, Essen.

- Herter, Michael; Mühlbauer, Karl-Heinz (Hg.) (2018):** Handbuch Geomarketing. Märkte und Zielgruppen verstehen. Lokal. Global. Digital. Herbert-Wichmann-Verlag. 2., neu bearbeitete und erweiterte Auflage. Berlin: VDE VERLAG GMBH; Wichmann Verlag.
- HoE; IFM; IWU (2021):** E4Q - Einbindung erneuerbarer Energieträger in die Energieversorgung vernetzter Quartiere. Protokoll zum 1. Projekttreffen am 20.01.2020. House of Energy; Institut für Massivbau; Institut Wohnen und Umwelt. Darmstadt, Kassel.
- Hörner, Michael (2021):** Herausforderungen bei der Verwendung von Hausumringen zur Erzeugung einer Modellabbildung des deutschen Gebäudebestandes aus GIS-Daten. Persönliche Besprechungen. Darmstadt, 2021.
- Hörner, Michael; Cischinsky, Holger; Bischof, Julian; Schwarz, Steffen; Behnisch, Martin; Meinel, Gotthard et al. (2022):** EnOB:dataNWG. Forschungsdatenbank Nicht-WohnGebäude. Repräsentative Primärdatenerhebung zur statistisch validen Erfassung und Auswertung der Struktur und der energetischen Qualität des Nichtwohngebäudebestands in Deutschland. Schlussbericht. Hg. v. Institut Wohnen und Umwelt. Institut Wohnen und Umwelt; Leibniz-Institut für ökologische Raumentwicklung; Bergische Universität Wuppertal. Darmstadt.
- Hörner, Michael; Cischinsky, Holger; Diefenbach, Nikolaus (2021):** Stand und Dynamik der energetischen Modernisierung der Gebäudehülle im Bestand der Nichtwohngebäude. Teilbericht Strukturdaten (E 4.2-1) im Projekt ENOB:dataNWG. Forschungsdatenbank Nichtwohngebäude. Hg. v. Institut Wohnen und Umwelt. Institut Wohnen und Umwelt. Darmstadt.
- Hosseini Haghghi, Seyede Rabeeh; Uribarri, Pilar Monsalvete Álvarez de; Padsala, Rushikesh; Eicker, Ursula (2021):** Characterizing and structuring urban GIS data for housing stock energy modelling and retrofitting. In: *Energy and Buildings*, S. 111706. DOI: 10.1016/j.enbuild.2021.111706.
- Huang, Zhexue (1998):** Extensions to the k-Means Algorithm for Clustering Large Data Sets with Categorical Values. In: *Data Mining and Knowledge Discovery* 2 (3), S. 283–304. DOI: 10.1023/A:1009769707641.
- Huang, Zhexue (2017):** Clustering Large Data Sets With Mixed Numeric And Categorical Values. CSIRO Mathematical and Information Sciences. Canberra.
- Huttenloher, Christian (2020):** Das Quartier als Handlungsebene für Klimaschutz im Gebäudebestand. In: *Forum Wohnen und Stadtentwicklung (FWS)* (6), S. 287–290.

- IPCC (2018):** Zusammenfassung für politische Entscheidungsträger. In: Masson-Delmotte, Valérie; Zhai, Panmao; Pörtner, Hans-Otto; Roberts, Debra; Skea, Jim; Shukla, Priyadarshi R. et al. (Hg.): 1,5 °C globale Erwärmung. Ein IPCC-Sonderbericht über die Folgen einer globalen Erwärmung um 1,5 °C gegenüber vorindustriellem Niveau und die damit verbundenen globalen Treibhausgasemissionspfade im Zusammenhang mit einer Stärkung der weltweiten Reaktion auf die Bedrohung durch den Klimawandel, nachhaltiger Entwicklung und Anstrengungen zur Beseitigung von Armut. Deutsche Übersetzung auf Basis der Version vom 14.11.2018. Genf, Schweiz.
- IPCC (2021):** Zusammenfassung für die politische Entscheidungsfindung. In: Masson-Delmotte, Valérie; Zhai, Panmao; Pirani, Anna; Connors, Sarah; Péan, Clotilde; Berger, Sophie et al. (Hg.): Naturwissenschaftliche Grundlagen. Beitrag von Arbeitsgruppe I zum Sechsten Sachstandsbericht des Zwischenstaatlichen Ausschusses für Klimaänderungen. Deutsche Übersetzung auf Basis der Druckvorlage, Oktober 2021. Genf, Schweiz: Deutsche IPCC-Koordinierungsstelle, Bonn; Akademie der Naturwissenschaften ProClim, Bern; Bundesministerium für Klimaschutz, Umwelt, Energie, Mobilität, Innovation, Wien.
- ISO (2015):** Information technology – Vocabulary. Standard ISO/IEC 2382:2015. International Organization for Standardization (ISO). Genf, Schweiz.
- IT.NRW (2010):** Rahmenanleitung für die Erläuterungen zur Statistik der Bautätigkeit im Hochbau. Landesbetrieb Information und Technik Nordrhein-Westfalen.
- Jacobs, Jane (1976):** Tod und Leben großer amerikanischer Städte. 2. Aufl. Wiesbaden: Vieweg (Bauwelt Fundamente. 4: Urbanistik, Städtebau, Baupolitik).
- Jiang, Hui (2021):** Machine Learning Fundamentals: Cambridge University Press.
- Jochem, Warren C.; Pannell, Oliver; Chamberlain, Heather R.; Jones, Patricia; Tatem, Andrew J. (2020):** Classifying settlement types from multi-scale spatial patterns of building footprints. In: *Environment and Planning B: Urban Analytics and City Science* 48 (5), S. 1161–1179.
- Jochem, Warren C.; Tatem, Andrew J. (2021):** Tools for mapping multi-scale settlement patterns of building footprints: An introduction to the R package foot. In: *PloS one* 16 (2), e0247535. DOI: 10.1371/journal.pone.0247535.
- Kant Hiran, Kamal; Khazanchi, Deepak; Kumar Vyas, Ajay; Padmanaban, Sanjeevikumar (Hg.) (2021):** Machine Learning for Sustainable Development: De Gruyter.
- Karami, Amin; Johansson, Ronnie (2014):** Choosing DBSCAN Parameters Automatically using Differential Evolution. In: *IJCA* 91 (7), S. 1–11. DOI: 10.5120/15890-5059.
- Kaufman, Leonard; Rousseeuw, Peter (1990):** Finding Groups in Data. An Introduction to Cluster Analysis: John Wiley & Sons Inc.

Khan, Mohammad Mahmudur Rahman; Siddique, Md. Abu Bakr; Arif, Rezoana Bente; Oishe, Mahjabin Rahman (2018): ADBSCAN: Adaptive Density-Based Spatial Clustering of Applications with Noise for Identifying Clusters with Varying Densities. In: IEEE (Hg.): Proceedings of the 4th International Conference on Electrical Engineering and Information & Communication Technology (iCEEiCT 2018). 4th International Conference on Electrical Engineering and Information & Communication Technology (iCEEiCT 2018). Dhaka, Bangladesh, 13.–15.09.2018: IEEE, S. 107–111.

Koch, Thilo (2021): Aktualisierung des EQ-Tools im Rahmen des Forschungsvorhabens EG2050:E4Q. Persönliche Besprechungen. Darmstadt, 2021.

Köhler, Benjamin; Bürger, Veit (2022): Status quo der Wärmeplanung in Deutschland. Wärmewende. Die Energiewende im Wärmebereich. Öko-Institut e.V. Freiburg.

Kroth, Julia; Wawrzyniak, Barbara (2022): Infas360 Gebäudedaten. Bonn, Darmstadt, 02.03.2022. E-Mail an André Müller.

KSB Darmstadt (2020): Ergebnisdokumentation der 1. Sitzung des neu konstituierten Klimaschutzbeirats. Klimaschutzbeirat der Stadt Darmstadt.

Kunde, Felix (2013): CityGML in PostGIS. Portierung, Anwendung und Performanz-Analyse am Beispiel der 3D City Database von Berlin. Masterthesis. Universität Potsdam, Potsdam.

Kutzner, Tatjana (2016): Geospatial Data Modelling and Model-driven Transformation of Geospatial Data based on UML Profiles. Dissertation. Technische Universität München, München.

LBauO M-V, vom 15.10.2015: Landesbauordnung Mecklenburg-Vorpommern (LBauO M-V).

Landeshauptstadt Mainz (2017): Masterplan 100% Klimaschutz für die Landeshauptstadt Mainz. Unter Mitarbeit von Annerose Hörter, Julia Brandt, Andrea Graf, Johannes Trunzer, Ilka Erfurt, Johannes Gansler et al. Hg. v. Grün- und Umweltamt, Landeshauptstadt Mainz. 4K Kommunikation für Klimaschutz; Leipziger Institut für Energie GmbH. Mainz, Hannover, Leipzig.

Langenbrinck, Gregor; Rensing, Larissa; Wüllner, Lutz; Habermann-Nieße, Klaus; Klehn, Kirsten; Rosenau, Lena et al. (2017): KfW-Programm 432 „Energetische Stadtsanierung – Zuschüsse für integrierte Quartierskonzepte u. Sanierungsmanager“. Ergebnisse der Begleitforschung. In: *BBSR-Online-Publikationen 2017* (15).

LDBV Bayern (2021): Datenformatbeschreibung 3D-Gebäudemodell Deutschland in den Ausprägungen Level of Detail 1 und 2. Version 2.1. Landesamt für Digitalisierung, Breitband und Vermessung Bayern (LDBV Bayern).

Legendre, Pierre; Legendre, Louis (1998): Developments in Environmental Modelling. Numerical Ecology. 2. englische Auflage. 31 Bände. Amsterdam: Elsevier (20).

Linke, Hans-Joachim; Hickel, Christian (2015): Geoinformationssysteme für raumbezogene Umweltanalysen. In: Kaltschmitt, Martin und Schebek, Liselotte (Hg.): Umweltbewertung für Ingenieure. Methoden und Verfahren. Berlin, Heidelberg: Springer Berlin Heidelberg, S. 63–77.

Loga, Tobias; Diefenbach, Nikolaus (2013): TABULA Calculation Method – Energy Use for Heating and Domestic Hot Water. Hg. v. Institut Wohnen und Umwelt. Darmstadt, zuletzt geprüft am 27.12.2022.

Loga, Tobias; Großklos, Marc; Müller, André; Swiderek, Stefan; Behem, Guillaume (2021): Realbilanzierung für den Verbrauch-Bedarf-Vergleich. Realistische Bilanzierung und Quantifizierung von Unsicherheiten als Grundlage für den Soll-Ist-Vergleich beim Energieverbrauchscontrolling. MOBASY-Teilbericht. Institut Wohnen und Umwelt. Darmstadt.

Loga, Tobias; Müller, André; Hörner, Michael (2018): Wärmewende jetzt. Der Weg zu einer drastischen Senkung der CO₂-Emissionen im Gebäudesektor. Institut Wohnen und Umwelt (Schlaglicht, 02/2018).

Loga, Tobias; Stein, Britta; Diefenbach, Nikolaus; Born, Rolf (2015): Deutsche Wohngebäudetypologie. Beispielhafte Maßnahmen zur Verbesserung der Energieeffizienz von typischen Wohngebäuden. 2. erweiterte Auflage. Hg. v. Institut Wohnen und Umwelt. Institut Wohnen und Umwelt. Darmstadt.

Loga, Tobias; Stein, Britta; Hacke, Ulrike; Müller, André; Großklos, Marc; Born, Rolf et al. (2019): Berücksichtigung des Nutzerverhaltens bei energetischen Verbesserungen. BBSR-Online-Publikation Nr. 04/2019. Bonn: Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR) im Bundesamt für Bauwesen und Raumforschung (BBR) (BBSR-Online-Publikation, 04/2019).

LUBW (2022): Energieatlas Baden-Württemberg. Landesanstalt für Umwelt Baden-Württemberg (LUBW). Karlsruhe. Online verfügbar unter <https://www.energieatlas-bw.de/energieatlas>, zuletzt geprüft am 02.04.2022.

Lynch, Kevin (2001): Das Bild der Stadt. [Zitiert nach Bott (2018): Nachhaltige Stadtplanung]. Basel.

Maechler, Martin; Rousseeuw, Peter; Struyf, Anja; Hubert, Mia; Hornik, Kurt; Studer, Matthias et al. (2021): Package ‘cluster’. “Finding Groups in Data”: Cluster Analysis Extended Rousseeuw et al.

Maheswari Viswanadhula, Uma; Aluvalu, Rajanikanth; Keerthi Chennam, Krishna (2021): Application of machine learning algorithms for facial expression analysis.

In: Kant Hiran, Kamal; Khazanchi, Deepak; Kumar Vyas, Ajay und Padmanaban, Sanjeevikumar (Hg.): Machine Learning for Sustainable Development: De Gruyter, 77-95.

Mahrishi, Mehul; Sharma, Girish; Morwal, Sudha; Jain, Vipin (2021): Data model recommendations for real-time machine learning applications: a suggestive approach.

In: Kant Hiran, Kamal; Khazanchi, Deepak; Kumar Vyas, Ajay und Padmanaban, Sanjeevikumar (Hg.): Machine Learning for Sustainable Development: De Gruyter, S. 115–128.

Meinel, Gotthard; Burckhardt, Manuel (2013): Die Geobasisprodukte Hausumringe und Hauskoordinaten. Charakterisierung und Aufbereitung für Gebäudebestandsanalysen.

In: *pfg* 2013 (6), S. 575–588. DOI: 10.1127/1432-8364/2013/0201.

Meinel, Gotthard; Krüger, Tobias; Ehrhardt, Denise; Behnisch, Martin (Hg.)

(2021): Flächennutzungsmonitoring XIII. Flächenpolitik – Konzepte – Analysen – Tools. Leibniz-Institut für ökologische Raumentwicklung e. V. Berlin: Rhombos-Verlag (IÖR-Schriften, 79).

Meinel, Gotthard; Schumacher, Ulrich; Behnisch, Martin (Hg.) (2013): Flächennutzungsmonitoring V. Methodik, Analyseergebnisse, Flächenmanagement. Beiträge des 5. Dresdner Flächennutzungssymposium 2013. Leibniz-Institut für ökologische Raumentwicklung e. V. Berlin: Rhombos-Verl. (IÖR-Schriften, 61).

Mijaka, Michal (2020): Introduction to naivebayes package. Online verfügbar unter https://cran.r-project.org/web/packages/naivebayes/vignettes/intro_naivebayes.pdf, zuletzt geprüft am 11.03.2022.

Müller, André (2021): Energy performance evaluation of internal insulation as a measure for the modernization of existing buildings. In: Stahl, Thomas und Wakili, Karim (Hg.): Energy-Efficient Retrofit of Buildings by Interior Insulation. 1. Aufl.: Butterworth-Heinemann, S. 467–490.

Müller, Dirk; Osterhage, Tanja; Calí, Davide (2012): Quartierskonzept Energieeffizientes Rintheim - wissenschaftliche Begleitung. Abschlussbericht. Hg. v. RWTH Aachen. E.ON Energieforschungszentrum, RWTH Aachen. Aachen.

MBO, vom 21.09.2012: Musterbauordnung (MBO), zuletzt geprüft am 11.03.2022.

Mutani, Guglielmina; Todeschi, Valeria (2021): GIS-based urban energy modelling and energy efficiency scenarios using the energy performance certificate database.

In: *Energy Efficiency* 14 (5). DOI: 10.1007/s12053-021-09962-z.

Nagel, Claus; Yao, Zhihang; Kunde, Felix; Hudra, György; Kolbe, Thomas H. (2020): 3DCityDB Handbuch. Version 4.2.3.

- Nagpal, Shreshth (2019):** Auto-Calibrated Urban Building Energy Models as Continuous Planning Tools for Greenhouse Gas Emissions Management. PhD-Thesis. Massachusetts Institute of Technology, Boston. Department of Architecture.
- Nagpal, Shreshth (2020):** Urban Energy Analytics for Cost-Effective GHG Emissions Reduction. Vortrag am IBPSA Boston June 2020 Treffen. International Building Performance Association (IBPSA), Boston Chapter. Online, 23.06.2020.
- Neitzel, Michael (2013):** Gebaute Quartiere. Beziehungen zwischen wohnungswirtschaftlichen und städtebaulichen Zugängen. In: Cummings, Veronika und Meisel, Ulli (Hg.): StadtQuartiere. Sozialwissenschaftliche, ökonomische und städtebaulich-architektonische Perspektiven. 1. Aufl. Essen, Ruhr: Klartext-Verlag, S. 179–196.
- Ng, Raymond T.; Han, Jiawei (2002):** CLARANS: a method for clustering objects for spatial data mining. In: *IEEE Trans. Knowl. Data Eng.* 14 (5), S. 1003–1016. DOI: 10.1109/TKDE.2002.1033770.
- Nouvel, Romain; Brassel, Kai-Holger; Bruse, Marcel; Duminil, Eric; Coors, Volker; Eicker, Ursula; Robinson, Darren (2015a):** SimStadt, a new workflow-driven urban energy simulation platform for CityGML city models. In: Proceedings of CISBAT 2015 International Conference on Future Buildings and Districts. Sustainability from Nano to Urban Scale - Vol. I, 09.–11. September 2015. Lausanne, EPFL Solar Energy and Building Physics Laboratory. Lausanne, S. 889–894.
- Nouvel, Romain; Kaden, Robert; Bahu, Jean-Marie; Kaempf, Jerome; Cipriano, Piergiorgio; Lauster, Moritz et al. (2015b):** Genesis of the CityGML Energy ADE. In: Proceedings of CISBAT 2015 International Conference on Future Buildings and Districts. Sustainability from Nano to Urban Scale - Vol. I, 09.–11. September 2015. Lausanne, EPFL Solar Energy and Building Physics Laboratory. Lausanne, S. 931–936.
- Pebesma, Edzer (2018):** Simple Features for R: Standardized Support for Spatial Vector Data. In: *The R Journal* 10 (1), S. 439. DOI: 10.32614/RJ-2018-009.
- Pebesma, Edzer; Bivand, Roger (2005):** Classes and methods for spatial data in R. In: *R News* 5 (2), S. 9–13.
- Pebesma, Edzer; Rundel, Colin; Teucher, Andy (2022):** Package ‘lwgeom’. Bindings to Selected ‘liblwgeom’ Functions for Simple Features.
- Pietruschka, Dirk; Pietzsch, Ursula; Monien, Dirk (Hg.) (2015):** Vision 2020 die Plusenergiegemeinde Wüstenrot. Begleitforschung EnEff:Stadt. Stuttgart: Fraunhofer IRB Verl. (Schriftenreihe EnEff).
- Plaue, Matthias (2021):** Data Science. Grundlagen, Statistik und maschinelles Lernen. Berlin: Springer Spektrum (Lehrbuch).

- R Core Team (2021):** R. A language and environment for statistical computing. Wien: R Foundation for Statistical Computing. Online verfügbar unter <https://www.R-project.org/>, zuletzt geprüft am 21.08.2023.
- ROG**, vom 22.12.2008: Raumordnungsgesetz (ROG).
- Reicher, Christa; Schmidt, Anke (Hg.) (2021):** Handbuch Energieeffizienz im Quartier. Clever versorgen, umbauen, aktivieren. Wiesbaden: Springer Fachmedien Wiesbaden (Springer eBook Collection).
- Reinhart, Christoph F.; Dogan, Timur; Jakubiec, J. Alstan; Rakha, Tarek; Sang, Andrew (2013):** UMI - An Urban Simulation Environment for Building Energy Use, Daylighting and Walkability. In: IBPSA (Hg.): Proceedings of BS2013: 13th Conference of International Building Performance Simulation Association, Chambéry, France, August 26–28. 13th Conference of International Building Performance Simulation Association. Chambéry, France, 26.–28. August 2013. International Building Performance Association (IBPSA).
- Rezaei, Abolfazl; Samadzadegan, Bahador; Rasoulian, Hadise; Ranjbar, Saeed; Samareh Abolhassani, Soroush; Sanei, Azin; Eicker, Ursula (2021):** A New Modeling Approach for Low-Carbon District Energy System Planning. In: *Energies* 14 (5), S. 1383. DOI: 10.3390/en14051383.
- Open Data- und PSI-Richtlinie (20.06.2019):** Richtlinie (EU) 2019/1024 des Europäischen Parlaments und des Rates vom 20. Juni 2019 über offene Daten und die Weiterverwendung von Informationen des öffentlichen Sektors.
- INSPIRE**, vom 14.03.2007 (14.03.2007): Richtlinie 2007/2/EG des Europäischen Parlaments und des Rates zur Schaffung einer Geodateninfrastruktur in der Europäischen Gemeinschaft.
- Roever, Christian; Raabe, Nils; Luebke, Karsten; Ligges, Uwe; Szepannek, Gero; Zentgraf, Marc; Meyer, David (2022):** Package ‘klaR’. Classification and Visualization.
- Roth, Ueli (1980):** Wechselwirkungen zwischen der Siedlungsstruktur und Wärmeversorgungssystemen. Forschungsprojekt BMBau RS II 4-704102-77.10 (1980). Unter Mitarbeit von Fritz Häubi, Joachim Albrecht, Mathias Bischoff, Annemarie Deucher, Lucius Harder, Beatrice Langraf und Gisela Pape. Bonn (Schriftenreihe des Bundesministers für Raumordnung, Bauwesen und Städtebau 06, 044).
- Roth, Ueli; Häubi, Fritz (1981):** Wechselwirkung zwischen Siedlungsstruktur und Wärmeversorgungssystemen. In: *Schweizer Ingenieur und Architekt* 98 (44), S. 970–983. DOI: 10.5169/seals-74579.
- R-SIG-DB; Wickham, Hadley; Müller, Kirill (2022):** Package ‘DBI’. R Database Interface. R Special Interest Group on Databases (R-SIG-DB).

- Ruppert, Hans; Eigner-Thiel, Swantje; Girschner, Walter; Karpenstein-Machan, Marianne; Roland, Folker; Ruwisch, Volker et al. (Hg.) (2008):** Wege zum Bioenergie-dorf. Leitfaden für eine eigenständige Wärme- und Stromversorgung auf Basis von Biomasse im ländlichen Raum. Fachagentur Nachwachsende Rohstoffe. 1. Aufl. Gülzow: Fachagentur Nachwachsende Rohstoffe.
- Sander, Joerg (2016):** Density-Based Clustering. In: Sammut, Claude und Webb, Geoffrey I. (Hg.): Encyclopedia of Machine Learning and Data Mining. Boston, MA: Springer (Springer eBook Collection), S. 1–5.
- Schnur, Olaf (Hg.) (2008a):** Quartiersforschung. Zwischen Theorie und Praxis. 1. Aufl. Wiesbaden: VS Verlag für Sozialwissenschaften (VS Research Quartiersforschung).
- Schnur, Olaf (2008b):** Quartiersforschung im Überblick. Konzepte, Definitionen und aktuelle Perspektiven. In: Schnur, Olaf (Hg.): Quartiersforschung. Zwischen Theorie und Praxis. 1. Aufl. Wiesbaden: VS Verlag für Sozialwissenschaften (VS Research Quartiersforschung), S. 19–51.
- Schnur, Olaf (2013):** Quartiersforschung revisited. Konzepte und Perspektiven eines stadtgeographischen Forschungsfeldes. In: Cummings, Veronika und Meisel, Ulli (Hg.): StadtQuartiere. Sozialwissenschaftliche, ökonomische und städtebaulich-architektonische Perspektiven. 1. Aufl. Essen, Ruhr: Klartext-Verlag, S. 17–40.
- Schwarz, Steffen; Hartmann, André; Hecht, Robert; Schorcht, Martin; Meinel, Gotthard; Benisch, Martin (2021):** Bestandsaufnahme Amtliche 3D-Gebäudemodelle im LoD1. Eine Metadatenanalyse. In: *Zeitschrift für Geodäsie, Geoinformation und Landmanagement* 146 (3), S. 198–206. DOI: 10.12902/zfv-0352-2021.
- SDL (2022):** UrbanModellingInterface 3.0. Sustainable Design Lab, Massachusetts Institute of Technology. Online verfügbar unter <http://web.mit.edu/sustainabledesignlab/projects/umi/index.html>, zuletzt geprüft am 07.03.2022.
- Senatsverwaltung für Stadtentwicklung, Bauen und Wohnen Berlin:** ALKIS Berlin. Senatsverwaltung für Stadtentwicklung, Bauen und Wohnen Berlin. Online verfügbar unter https://fbinter.stadt-berlin.de/fb_daten/beschreibung/alkis_karte.html, zuletzt geprüft am 13.03.2022.
- Sester, Monika (Hg.) (2019):** Geoinformatik. Handbuch der Geodäsie, herausgegeben von Willi Freedon und Reiner Rummel. Springer-Verlag GmbH. Berlin, Heidelberg: Springer Spektrum (Springer Reference Naturwissenschaften).
- Sibbertsen, Philipp; Lehne, Hartmut (2021):** Statistik. Berlin, Heidelberg: Springer Berlin Heidelberg.

SRU (2020): Umweltgutachten 2020. Für eine entschlossene Umweltpolitik in Deutschland und Europa. Umweltgutachten 2020. Hg. v. Deutscher Bundestag (19. Wahlperiode). Sachverständigenrat für Umweltfragen (19/20590).

Stiftung Klimaneutralität; Agora Energiewende; Agora Verkehrswende (2021): Politikinstrumente für ein klimaneutrales Deutschland. 50 Empfehlungen für die 20. Legislaturperiode (2021–2025). Berlin.

Stockinger, Volker (2019): Wärme, Kälte und Strom im Quartier. In: *ÖW* 33 (3), S. 25. DOI: 10.14512/OEW340325.

Streich, Bernd (2011): Stadtplanung in der Wissensgesellschaft. Ein Handbuch. 2. Aufl. Wiesbaden: VS-Verl.

Ullmann, Theresa; Hennig, Christian; Boulesteix, Anne-Laure (2021): Validation of cluster analysis results on validation data: A systematic framework. In: *WIREs Data Min & Knowl.* DOI: 10.1002/widm.1444.

Ulrich, Tim (2014): Heuristiken für kombinierte Standort und Gebietsplanung mit vorgegebenen und zusätzlichen, frei wählbaren Standorten. Diplomarbeit. Karlsruher Institut für Technologie (KIT), Karlsruhe. Institut für Theoretische Informatik.

UN (2015): Paris Agreement to the United Nations Framework Convention on Climate Change (UNFCCC). United Nations. Paris.

Vogelpohl, Anne (2008): Stadt der Quartiere? Das Place-Konzept und die Idee von urbanen Dörfern. In: Schnur, Olaf (Hg.): *Quartiersforschung. Zwischen Theorie und Praxis.* 1. Aufl. Wiesbaden: VS Verlag für Sozialwissenschaften (Research Quartiersforschung).

von Malottki, Christian; Koch, Thilo; Blees, Volker (2016): EQII. Erweiterte Bilanzierung von Energieverbrauch und CO₂-Emissionen auf Quartiersebene. ExWoSt-Studie. Bonn: Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR) im Bundesamt für Bauwesen und Raumforschung (BBR) (ExWoSt-Informationen, 48/1).

von Malottki, Christian; Koch, Thilo; Brandis, Christof; Stete, Gisela; Blees, Volker (2012): EQ. Anforderungen an energieeffiziente und klimaneutrale Quartiere. Bonn: Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR) im Bundesamt für Bauwesen und Raumforschung (BBR) (ExWoSt-Informationen, 42/1).

Weiler, Verena; Duminil, Eric; Schröter, Bastian; Coors, Volker; Brüggemann, Thilo; Balbach, Bobo et al. (2021): Automatisierte Modellierung von Quartierswärmebedarfen auf Basis von 3D-Gebäudemodellen. In: *EUROHEAT&POWER* 50 (4-5), S. 40–45.

- Weiler, Verena; Weinand, Jann; Köhler, Sally; McKenna, Russel; Eicker, Ursula; Fichtner, Wolf (2018):** Bestimmung energetisch zusammenhängender urbaner Gebäudegruppen als Grundlage für die Quartiersplanung. In: IEE (Hg.): EnInnov2018. Neue Energie für unser bewegtes Europa. 15. Symposium Energieinnovation. Graz, 14.–16. Februar 2018. Institut für Elektrizitätswirtschaft und Energieinnovation (IEE), Technische Universität Graz. Graz.
- Weinand, Jann; McKenna, Russel; Fichtner, Wolf (2019):** Developing a municipality typology for modelling decentralised energy systems. In: *Utilities Policy* 57, S. 75–96.
- Weißmann, Claudia (2017):** Effizienter Einsatz erneuerbarer Energieträger in vernetzten Wohnquartieren. Entwicklung eines simulationsbasierten Verfahrens zur energetischen, ökologischen und ökonomischen Bewertung. Dissertation. Technische Universität Darmstadt, Darmstadt. Institut für Massivbau.
- Wern, Bernhard; Noll, Florian; Trapp, Manuel (2017):** Wärmekataster Saarland. Endbericht im Auftrag des Ministerium für Wirtschaft, Arbeit, Energie und Verkehr des Saarlandes. Ministerium für Wirtschaft, Arbeit, Energie und Verkehr des Saarlands: Institut für ZukunftsEnergie- und Stoffstromsysteme (izes).
- Westermann, Thomas (2020):** Mathematik für Ingenieure. Ein anwendungsorientiertes Lehrbuch. 8. Aufl. 2020. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Wickham, Hadley; Ooms, Jeroen; Müller, Kirill (2022):** Package ‘RPostgres’. Rcpp Interface to PostgreSQL.
- Wissenschaftsstadt Darmstadt (2019):** Höchste Priorität für Klimaschutz – Weltklima in Not – Darmstadt handelt. (SV-2019/0043). Unter Mitarbeit von Stadtverordnetenversammlung der Wissenschaftsstadt Darmstadt. Darmstadt, 16.08.2019.
- WPGP (2021):** Package ‘foot’. An R package for processing building footprint morphometrics. Version 0.8: WorldPop Research Group (WPGP), University of Southampton.
- Yang, Jucheng; Park, Dong Sun; Yoon, Sook; Chen, Yarui; Zhang, Chuanlei (2018):** Machine Learning and Biometrics. London: InTech.
- Yao, Zhihang; Nagel, Claus; Kunde, Felix; Hudra, György; Willkomm, Philipp; Donaubauer, Andreas et al. (2018):** 3DCityDB - a 3D geodatabase solution for the management, analysis, and visualization of semantic 3D city models based on CityGML. In: *Open geospatial data, softw. stand.* 3 (1). DOI: 10.1186/s40965-018-0046-7.
- Young, H. P. (1988):** Measuring the Compactness of Legislative Districts. In: *Legislative Studies Quarterly* 13 (1), S. 105–115.

Zhang, Harry (2004): The Optimality of Naive Bayes. In: Florida AI Research Society (Hg.): Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference. Miami Beach, Florida, May 17–19, 2004. Menlo Park, California: AAAI Press.

Zhou, Zhi-Hua (2021): Machine Learning. Singapore: Springer Singapore.

Zirak, Maryam; Weiler, Verena; Hein, Martin; Eicker, Ursula (2020): Urban models enrichment for energy applications: Challenges in energy simulation using different data sources for building age information. In: *Energy* 190, S. 116292. DOI: 10.1016/j.energy.2019.116292.

ANHANG

Anhang A	Ergänzende Ausführungen zu Siedlungs- und Quartierstypologien ...	174
Anhang B	Ergänzende Darstellung von clusteranalysebasierten Studien	184
Anhang C	Weiterführende Informationen zu Geoinformationsdaten.....	185
Anhang D	Ergebnisse weiterführender statistischer Auswertungen auf verschiedenen Raumebenen.....	196

ANHANG A ERGÄNZENDE AUSFÜHRUNGEN ZU SIEDLUNGS- UND QUARTIERSTYPOLOGIEN

Anhang A-1 Beschreibungen der Siedlungsstrukturtypen nach Roth (1980)

Das in Abbildung Anhang A-1 dargestellte Raum-Siedlungsstrukturmodell von Roth (1980) wurde verwendet, um auf Basis von statistischen Erhebungen und der Auswertung von Orthofotos eine Einteilung der Siedlungsstrukturtypen vorzunehmen sowie deren Bestand auf definierte Raumtypen in Deutschland hochzurechnen. Die Siedlungsstrukturtypen wurden dabei durch verschiedene charakteristische baulich-städtebauliche Kennzahlen beschrieben, welche aus unterschiedlichen Datenquellen, Studien oder Expertenannahmen abgeleitet wurden (siehe auch Abbildung Anhang A-2, S. 178).

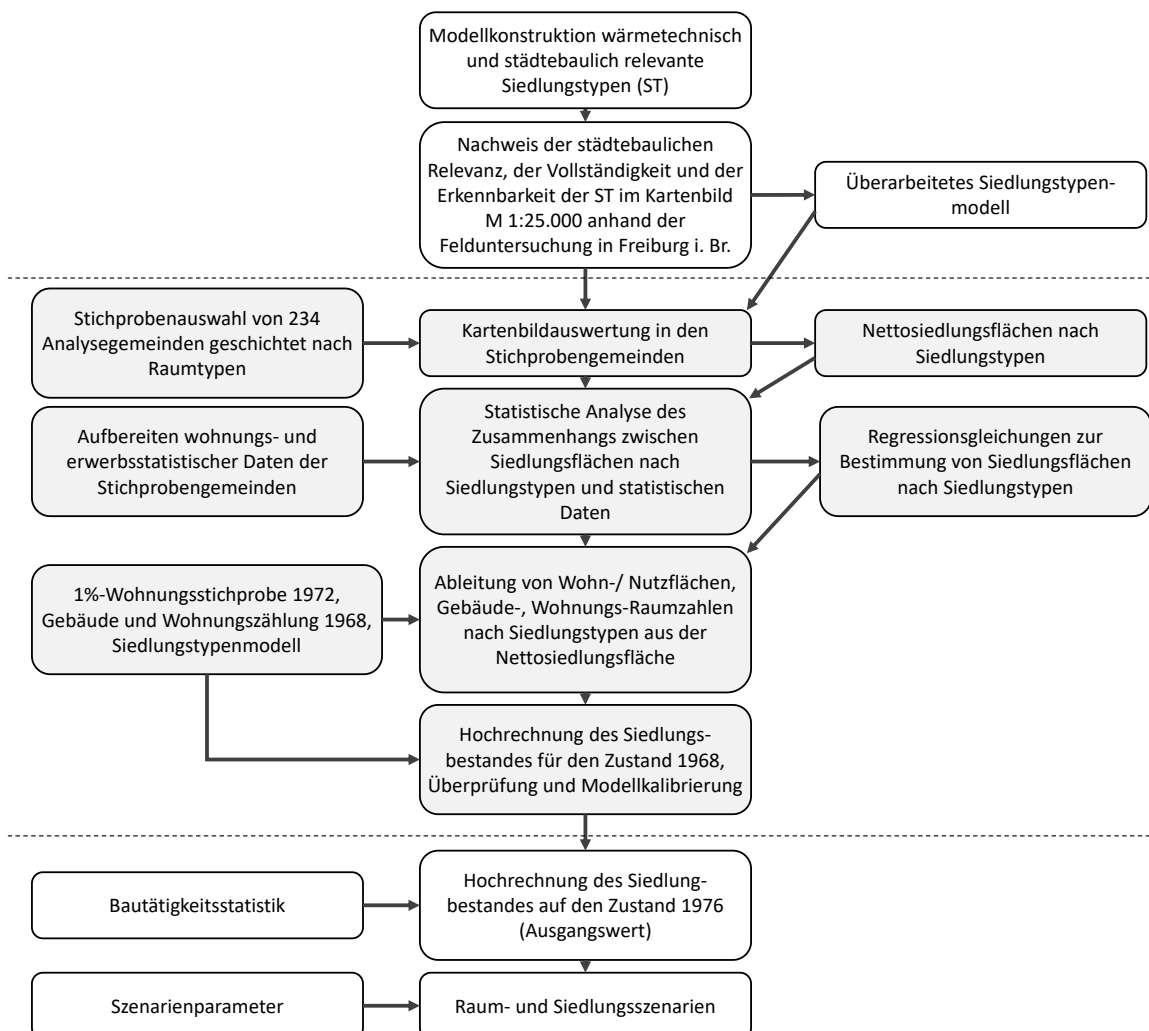


Abbildung Anhang A-1: Darstellung des in Roth (1980) verwendeten Raum-Strukturmodells zur Definition von Siedlungsstrukturtypen und einer Hochrechnung der Häufigkeiten für festgelegte Raumtypen [eigene Darstellung nach Roth (1980, S. 49)]

Die Kennwerte[, auf welchen die Beschreibung der Siedlungsstrukturtypen beruht,] stammen aus unterschiedlichen Untersuchungen:

- Die gebäude- und siedlungsspezifischen Werte (Geschoßzahl, Geschoßflächenziffer und Grundflächenziffer) sind die eigentlichen Erzeugungsparameter. Sie wurden aufgrund allgemeiner Kenntnisse über die Siedlungsstruktur so gewählt, daß sie den gesamten Siedlungsbestand identifizieren lassen.
- Baualter, Dachform, Beschattung durch Pflanzen und Fensterflächenanteil sind Primärmerkmale. Sie sind das Resultat einer eigenen Felduntersuchung.
- Sämtliche baulich-geometrischen Daten lassen sich ableiten, sei es durch Messungen aus den physischen Siedlungstypenmodellen, durch Berechnung oder durch Literaturlauswertung.
- Die baulich-technischen und die wärmetechnischen Werte stammen aus Normen oder der Literatur.
- Die k -Werte der einzelnen Bauteile entstammen teilweise einer Untersuchung des Battelle-Instituts [...] und einschlägigen Normen.
- Die mittleren k -Werte werden nach Kap. 2.4 berechnet und mit Literaturwerten verglichen. (Roth 1980, S. 49–50)

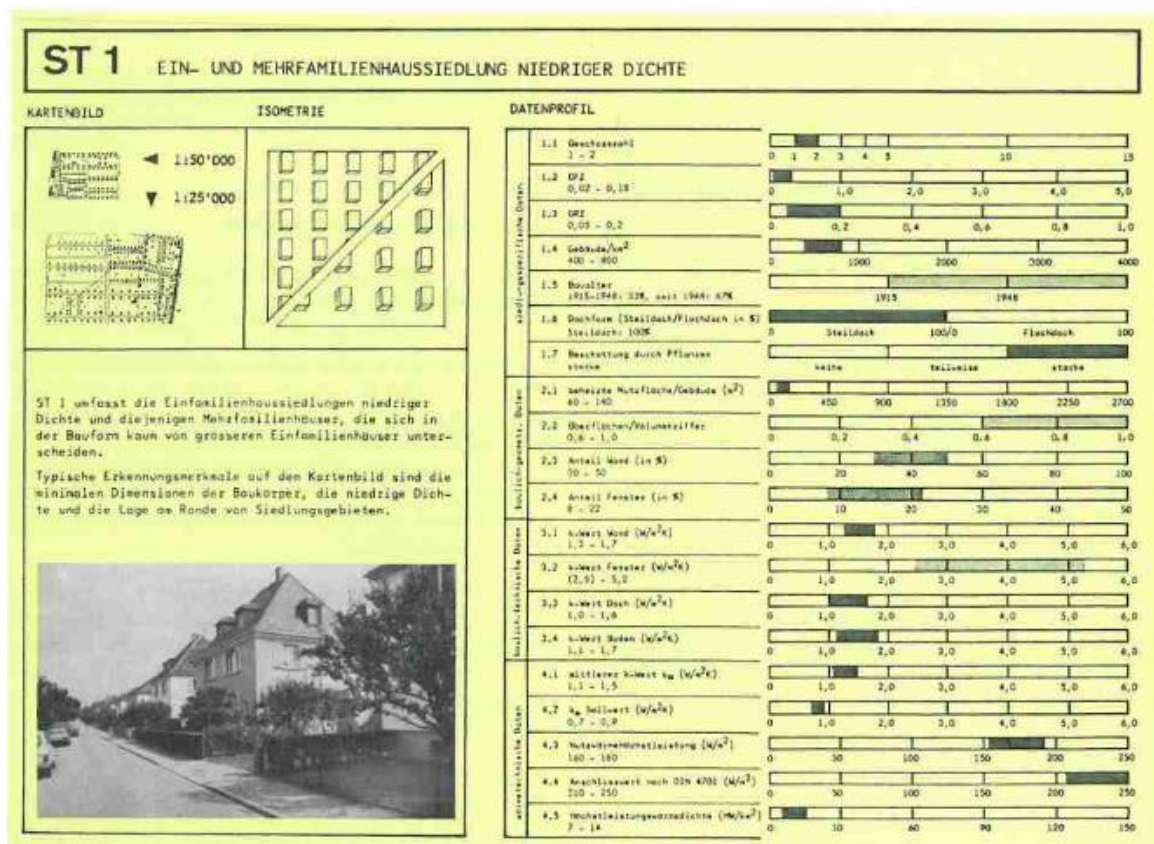


Abbildung Anhang A-2: Beschreibung charakteristischer Kennzahlen der Siedlungsstrukturtypen am Beispiel einer Ein- und Mehrfamilienhaussiedlung niedriger Dichte (Siedlungsstrukturtyp ST1) [Roth (1980)]

Für alle der definierten ST liefert Roth neben den charakteristischen Kennzahlen zudem eine kurze textliche Beschreibung, welche eine Nachvollziehbarkeit bzw. Erkennbarkeit von Siedlungsstrukturen im Kontext realer Stadtstrukturen erleichtern soll. Die textlichen Beschreibungen der ST von Roth sind in Tabelle Anhang A-1 zusammengefasst.

Tabelle Anhang A-1: *Textliche Beschreibung der Siedlungsstrukturtypen nach Roth (1980)*
 [Roth (1980, S. 99–103)]

ST 1	Ein- und Mehrfamilienhaussiedlung niedriger Dichte
<p>ST 1 umfasst die Einfamilienhaussiedlungen niedriger Dichte und diejenigen Mehrfamilienhäuser, die sich in der Bauform kaum von grösseren Einfamilienhäusern unterscheiden. Typische Erkennungsmerkmale auf dem Kartenbild sind die minimalen Dimensionen der Baukörper, die niedrige Dichte und die Lage am Rande von Siedlungsgebieten.</p>	
ST 2	Dorfkern und Einfamilienhaussiedlung hoher Dichte
<p>ST 2 repräsentiert die historischen Dorfkerne sowie dichte Einfamilienhaussiedlungen. Die Dorfkerne sind gekennzeichnet durch eine hohe Dichte und ein Strassennetz, das mehr von topographischen Bedingungen als von geometrischer Vorstellung geprägt ist. Die Einfamilienhaussiedlungen unterscheiden sich vom ST 1 vor allem durch ihre höhere Dichte.</p>	
ST 3	Reihenhaussiedlung
<p>ST 3 stellt Reiheneinfamilienhäuser dar. Auf dem Kartenbild sind sie durch eine dichte, parallele Anordnung und ein kleinmaschiges, rechtwinkliges Erschliessungsnetz zu erkennen. Dank dieser Merkmale können die Reiheneinfamilienhäuser trotz ähnlicher Grundrissform mit den Zeilenbauten des ST's 4 nicht verwechselt werden.</p>	
ST 4	Zeilenbebauung mittlerer Dichte
<p>ST 4 umfasst Zwei-, Drei- und Vierspänner mit drei bis fünf Geschossen. Diesem Siedlungstyp lässt sich ein grosser Teil der Nachkriegsbauten zuordnen. Entsprechend ihrem Baualter finden sich diese Gebäude vorwiegend am Stadtrand.</p>	
ST 5	Zeilenbebauung hoher Dichte und Hochhäuser
<p>Dass sowohl hohe Zeilenbauten wie auch Hochhäuser zu einem einzigen Siedlungstyp, ST 5, zusammengefasst werden, ergibt sich aus der Tatsache, dass beide Bauweisen nebeneinander im Stadtgefüge vorkommen. Auf dem Kartenbild ist ST 5 an der weitläufigen Verteilung grosser Gebäude erkennbar. Diese kommen praktisch nur in äusserster Randlage oder in neuen Stadtzentren vor.</p>	
ST 6	Blockbebauung
<p>Zu ST 6 gehören die städtischen Mehrfamilienhäuser der Jahrhundertwende, die zusammen ganze Strassenblöcke bilden. Typische Merkmale der Blockbebauung sind die darin liegenden grossen, zentralen Freiräume und die Hauptorientierung der Wohnungen auf die Strasse. Entsprechend ihrem Baualter schliesst diese Blockbauweise direkt an den Stadtkern an.</p>	
ST 7	Citybebauung ab Mitte 19. Jahrhundert
<p>ST 7 umfasst die städtischen Bauten, die Ende 19. Jahrhundert oder nach dem 2. Weltkrieg, während einer Phase wirtschaftlichen Aufschwungs, errichtet wurden. In dichten Baublöcken zusammengedrängt, enthalten diese Gebäude vorwiegend Büros und Geschäfte. In den meisten Städten befindet sich der ST 7 zwischen der mittelalterlichen Altstadt und dem damals neugebauten Bahnhof.</p>	

Tabelle Anhang A-1: Fortsetzung

ST 8	Mittelalterliche Altstadt
ST 8 repräsentiert die mittelalterliche Altstadt. Zentrale Lage, geschlossene Bauweisen und verwinkelte Erschliessung sind eindeutige Merkmale dieses Siedlungstyps, wie dies die Kartenbilder deutlich zeigen. Ebenfalls typisch sind die hohe Bebauungsdichte und die verhältnismässig hohen Gebäude.	
ST 9	Industrie und Lagergebäude
ST 9 umfasst sämtliche Industrie-, Lager- und Gewerbebauten. Diese liegen meistens am Stadtrand und haben oft den Anschluss an das Bahnnetz. Diese beiden Merkmale gestatten eine eindeutige Identifikation auf dem Kartenbild.	

Die Abgrenzung der Siedlungsstrukturtypen wird in Roth (1980, S. 50)

„[...] in einem zweiten Schritt [...durch eine] Felduntersuchung in Freiburg i. Br. geprüft. Anhand der topographischen Karte 1 :25 000 werden Gebiete (Zellen) gleicher Siedlungsstruktur ausgeschieden und Siedlungstypen zugeordnet [...].

Anschließend werden nach dem Zufallsprinzip je Siedlungstyp drei Baublöcke, wenn möglich in drei verschiedenen Zellen mit insgesamt rund 500 Gebäuden, zur Aufnahme in die Felduntersuchung bestimmt.“

Dem Autor zufolge konnte durch die Feldüberprüfung die Methodik zur Bildkartenanalyse bestätigt und die Ausprägungen der siedlungsstrukturellen Merkmale in den ST auf deren Validität hin geprüft werden. Anhand einer kleineren Anzahl von Gemeinden wurde zudem der Anteil verschiedener ST am Gesamtgemeindegebiet ermittelt und unter Zuhilfenahme verfügbarer raumbezogener Strukturdaten sowie mittels Korrelations- und Regressionsanalysen für eine Hochrechnung auf den Gemeindegesamtbestand verwendet. (Vgl. Roth 1980, S. 50–53)

Anhang A-2 Beschreibung und charakteristische Merkmale der Siedlungstypen nach Blesl (2002)

Mit Hilfe der Siedlungstypmethode werden in Blesl (2002) Stadtgebiete oder allgemein bebaute Flächen in Siedlungstypen unterteilt. Diese werden durch den Autor wie in Tabelle Anhang A-2 dargestellt beschrieben.

Tabelle Anhang A-2: *Textliche Beschreibung der Siedlungstypen nach Blesl (2002) [nach Blesl (2002, S. 137–140)]*

ST 0	Freistehende Einzelgebäude
<p>Bei diesem Siedlungstyp handelt es sich um meist freistehende Einzelgebäude, abgelegen von weiterer Bebauung.</p> <p>Die Gebäudevielfalt ist bei diesem Siedlungstyp groß, da Einfamilienhäuser, kleine Mehrfamilienhäuser und kleine Sonderbauten (z. B. Gaststätten oder Aussiedlerhöfe) beobachtet werden.</p>	
ST 1	Lockere offene Bebauung
<p>Die lockere offene Bebauung, die auch als „Streusiedlung“ bezeichnet werden kann, ist in Vorortgemeinden meist als Ausläufer von Einfamilienhaussiedlungen (ST 2) zu finden. Die unregelmäßige Bebauung wurde in Großstädten teilweise durch die Teilung der Grundstücke und den nachträglichen Bau von Einfamilienhäusern verdichtet.</p> <p>Die Wohngebäude sind fast vornehmlich 2-geschossig und größer als in den durchschnittlichen Einfamilienhaussiedlungen. Sie sind meistens freistehend bzw. einseitig angebaut an Garagen.</p>	
ST 2	Einfamilienhäuser- u. Doppelhäuser-Siedlung
<p>Die Einfamilienhaussiedlung repräsentiert überwiegend Wohngebiete, die an Stadträndern und in Vorortgemeinden gelegen ist. Der Siedlungstyp ist mit einem dichten regelmäßigen Straßennetz erschlossen.</p> <p>Die freistehenden Einfamilienhäuser sind mit 1 ½ bis 2 ½ Geschossen gebaut. Die Bandweite der Gebäudegröße ist aufgrund der unterschiedlichen Bauzeiten relativ groß. Sie reicht von sehr kleinen Arbeiterwohnhäusern der 30er Jahre bis zu großvolumigen Villen und kleinen Mehrfamilienhäusern.</p>	
ST 3a	Ehemaliger Dorfkern
<p>Der Siedlungstyp „ehemaliger Dorfkern“ befindet sich vornehmlich in Groß- und Mittelstädten. Hierbei handelt es sich um dicht bebaute Mischgebiete.</p> <p>Im Gegensatz zu Siedlungstyp 3b fehlt der eigentliche Dorfkern und der Straßenverlauf ist regelmäßig rechteckig. Die Gebäude sind direkt an den Gehweg gebaut. Knapp die Hälfte der Gebäude ist aneinander gebaut.</p>	
ST 3b	Zeilenbebauung hoher Dichte und Hochhäuser
<p>Der Siedlungstyp „ländlicher Dorfkern“ befindet sich vornehmlich in kleinen Eingemeindungen von Großstädten und im ländlichen Raum. Er ist durch einen zentralen Platz gekennzeichnet von dem aus sternförmig die Straßen weglaufen. Das Straßennetz ist mehr von topographischen, wirtschaftlichen und gesellschaftlichen Bedingungen als von geometrischen Verläufen geprägt.</p> <p>Der Siedlungstyp weist eine dichte Bebauung auf. Die Gebäude sind meistens aneinander und direkt an den Gehweg gebaut. Sie bestehen teilweise aus Fachwerk- und aus Mauerwerkbauten. Die Wohngebäude sind 1 ½- bis 2-geschossig und weisen zu ca. 75 % ein ausgebautes Dachgeschoß auf.</p>	

Tabelle Anhang A-2: Fortsetzung

ST 4	Reihenhäuser
<p>Der Siedlungstyp „Reihenhäuser“ befindet sich in Vororten und an Stadträndern meist in Verbindung mit ST 2. Im Siedlungsgebiet befinden sich vornehmlich Reihenhäuser.</p> <p>Das Gebiet ist engmaschig geometrisch durch Straßen erschlossen. Die Häuserzeilen sind dicht und überwiegend parallel angeordnet.</p>	
ST 5a	Siedlung kleiner Mehrfamilienhäuser
<p>Der Siedlungstyp ist erst ab Anfang der 80er Jahre entstanden. Er besteht beinahe überwiegend aus kleinen Mehrfamilienhäusern, die mit einem engen geometrischen Netz von Straßen erschlossen sind.</p> <p>Die Wohngebäude sind fast zum Großteil 2-geschossig, das Dachgeschoß ist vollständig ausgebaut. Die Gebäude sind freistehend bzw. einseitig angebaut an Garagen.</p>	
ST 5b	Zeilenbebauung mit kleinen und größeren Mehrfamilienhäusern
<p>Die Zeilenbebauung (bis max. 115 m Gesamtlänge) ist zum großen Teil in der Nachkriegszeit gebaut worden. Die Wohnblocks sind grobmaschig erschlossen. Sie befinden sich vorwiegend am Stadtrand von Groß- und Mittelstädten.</p> <p>Die großen Mehrfamilienhäuser sind 3- bis 5-geschossige Wohnbauten, die als 2-, 3- und 4-Spänner angeordnet sind.</p>	
ST 6	Zeilenbebauung mit großen hohen Mehrfamilienhäusern oder Hochhäusern
<p>Typisch für den Siedlungstyp ist die Randlage von Städten („Trabantenstädte“). In Abhängigkeit des Baualters sind Hochhäuser in Form von Scheibenhäuser [sic!] oder Wohncluster [sic!] zu beobachten, die im großen Gebäudeabstand voneinander stehen.</p> <p>Der Siedlungstyp wird verkehrstechnisch in erster Linie über Sackgassen erschlossen, die von einer Umgehungsstraße um das Siedlungsgebiet abzweigen.</p>	
ST 7a	Blockbebauung niedriger Dichte
<p>Der Siedlungstyp 7a befindet sich in Innenstadtbereichen von Großstädten. Die Straßenführung ist beinahe quadratisch. Typisches Merkmal sind die innen liegenden großen Freiflächen und die Hauptorientierung der Wohnungen auf die Straße.</p> <p>Zu diesem Siedlungstyp gehören die großen Mehrfamilienhäuser und Villen der Jahrhundertwende, die zusammen ganze Straßenblöcke bilden. Die Einzelgebäude sind meistens 3-geschossig. Die Gebäude sind teilweise aneinander gebaut.</p>	
ST 7b	Blockbebauung hoher Dichte
<p>Siedlungstyp 7b unterscheidet sich von Siedlungstyp 7a dahingehend, daß die gewerbliche Nutzung der unteren Etagen der großen Mehrfamilienhäuser verbreitet ist und die innen liegenden großen Freiflächen mit Gebäuden bebaut sind. Diese können nur durch Zufahrten in den Innenhof erreicht werden.</p> <p>Die Einzelgebäude sind meistens 3-geschossig und aneinander gebaut.</p>	
ST 8	Citybebauung
<p>Siedlungstyp 8 kennzeichnet die Citybebauung sehr hoher Dichte. Sie erstreckt sich in den meisten Städten vom Bahnhof bis zum Zentrum. Sie grenzt häufig an die mittelalterliche Altstadt an.</p> <p>Große kompakte Gebäude kennzeichnen das Gebiet. Teilweise sind bei der Citybebauung ganze Blöcke (vgl. Siedlungstyp 7a und 7b) durch wenige Gebäude vollständig überbaut. Die Gebäudenutzung variiert je nach Lage zwischen ausschließlichen Geschäfts-, Dienstleistungs-, Verwaltungs- und reinen Wohngebäuden.</p>	
ST 9	Historische Altstadt
<p>Siedlungstyp 9 entspricht dem mittelalterlichen Stadtkern. Diese [sic!] ist durch die geschlossene Bebauung mit verhältnismäßig hohen 3 ½-geschossigen Gebäuden gekennzeichnet.</p> <p>Typisch ist ein verwinkeltes Erschließungsnetz von Straßen und Gassen mit öffentlichen Plätzen.</p>	

Tabelle Anhang A-2: Fortsetzung

ST 10a	Öffentliche Sonderbauten (groß)
<p>Bei dem Siedlungstyp 10a öffentliche Sonderbauten (groß) kann zwischen einzelnen großen Sonderbauten (Schulzentrum oder Krankenhäuser) und Gebieten mit großen Sonderbauten (z. B. Universitätsgelände oder Kasernen) unterschieden werden.</p> <p>Gebiete mit großen Sonderbauten werden in erster Linie über Sackgassen erschlossen, die von einer umlaufenden Ringstraße abzweigen.</p>	
ST 10b	Öffentliche Sonderbauten (klein)
<p>Bei dem Siedlungstyp 10b öffentliche Sonderbauten (klein) existieren keine größeren Ansammlungen von kleinen Sonderbauten (z. B. Schulen, Kindergarten u. a.). An die Siedlungstypen 2, 5a, 5b, 6, 7 [sic!] schließen sich oft ohne städtebauliche Integration kleine öffentliche Sonderbauten an, die meistens von kleinen Grünanlagen bzw. Freiflächen umgeben sind.</p>	
ST 11a	Industriebauten
<p>Siedlungstyp 11a beinhaltet große Industriebauten mit meist unregelmäßigem Grundriß. Oft werden mehrere Gebäude unterschiedlichster Nutzung (Produktions-, Lager und Bürogebäude) zu einem „Werksgelände“ zusammengefaßt. Der Prozeßwärmebedarf variiert stark in Abhängigkeit der produzierten Produkte. Zwischen den einzelnen Gebäudekomplexen befinden sich große meist versiegelte Flächen.</p> <p>Der Siedlungstyp befinden [sic!] sich sehr häufig in Stadtrandlage mit guter Verkehrsanbindung (z. T. mit Bahnanschluß).</p>	
ST 11b	Gewerbliche Sonderbauten / Dienstleistungsbauten
<p>Dieser Siedlungstyp beinhaltet große von der Straße zurückversetzte Gewerbe-, Dienstleistungs- und Lagerbauten mit unregelmäßigen Grundrissen. Reine Gewerbe- und Industriegebiete sind vereinzelt ab ca. 1948, vermehrt ab den 70er Jahren entstanden.</p> <p>Der Siedlungstyp liegt meist in verkehrsgünstiger Stadtrandlage. Das Gebiet ist mit einem großmaßstäbigen orthogonalem Straßennetz erschlossen. Oftmals führen Ringstraßen um den Siedlungstyp herum.</p>	
ST 12	Sonstige Versorgungsgebiete und -gelände
<p>Siedlungstyp 12 beinhaltet die bisher nicht berücksichtigten Versorgungsgebäude und -gelände. Hierbei handelt es sich meistens um Sportanlagen wie Tennishallen, Tennisplätze, Freibäder, Sportplätze teilweise mit dazugehörigen Nebengebäuden sowie um Gärtnereien mit Gewächshäusern.</p>	

Die Typisierung von Siedlungsgebieten verfolgte insbesondere das Ziel, Wärmebedarfe bzw. -verbräuche und mögliche Strategien zur Verlegung einer netzgebundenen Wärmeversorgung innerhalb von Siedlungsgebieten zu bestimmen. Daher wurde ähnlich dem Vorgehen von Roth (1980) eine Charakterisierung der Siedlungstypen mittels energieverbrauchsbestimmender Parameter vorgenommen:

- Gebäudetyp
- Baualter
- Bandbreite der Gebäudegrundfläche sowie durchschnittliche Grundfläche im Siedlungstyp
- Anzahl der Gebäude
- Bandbreite des Abstandes zwischen Gebäude und Straße sowie durchschnittlicher Abstand zwischen Gebäude und Straße



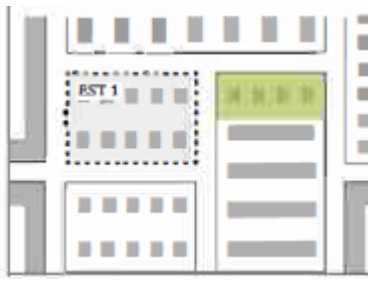
- Bandbreite der Anzahl einzelner Blöcke bezogen auf einen km², durchschnittliche Grundfläche eines Blockes sowie durchschnittliche Anzahl der Gebäude pro Block
- Abstand zwischen parallelen Straßen bzw. durchschnittlicher Abstand zwischen parallelen Straßen sowie Bandbreite des Abstandes zwischen Kreuzungen und durchschnittlicher Abstand zwischen Kreuzungen
- Gesamtlänge der Straßen bezogen auf einen km²
- Anzahl vom Zentrum abgehender Straßen
- Abstand zwischen Straßenabzweigungen bzw. durchschnittlicher Abstand zwischen Straßenabzweigungen

(vgl. Blesl 2002, S. 140–141)

Anhang A-3 Ergänzende Beschreibung der Einteilungssystematik in energetische Stadtraumtypen (EST) nach Hegger et al. (2012; 2015a)


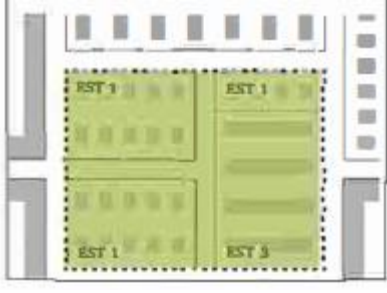
Zur Einteilung der energetischen Stadtraumtypen erarbeiteten Hegger et al. (2012; 2015a) eine detailliert beschriebene Systematik zur Übersetzung der auf Gebäudeebene vorgefundenen Strukturen in stadträumliche Gruppen bzw. Typen. Diese Systematik ist in Tabelle Anhang A-3 dargestellt.

Tabelle Anhang A-3: Raumebenen, als Träger von baulich-geometrischen Strukturmerkmalen zur Bestimmung der energetischen Stadtraumtypen in Hegger et al. (2012) [Quelle: Hegger et al. (2012, S. 71); angepasste Darstellung]⁸⁰

		
<p>Ein Gebäude ist immer einer Parzelle zugeordnet. Grundsätzlich wird nach der Anzahl Wohneinheiten in Ein- und Mehrfamilienhäuser und nach der Bauweise in freistehende und gekuppelte Gebäude unterschieden.</p>	<p>Die Parzelle bildet die kleinste städtebauliche Einheit. In der Regel verfügt jede Parzelle über einen Zugang zum öffentlichen Verkehrsnetz.</p>	<p>Die Sektion ist die kleinste räumlich zusammenhängende Einheit eines EST, EE oder SRT und besteht aus einer oder mehreren Parzellen.</p>

⁸⁰ Die Beschreibungen in Tabelle Anhang A-3 enthalten Korrekturen von in der Quelle vorhandenen orthografischen Fehlern. Diese sind der einfacheren Lesbarkeit an dieser Stelle nicht gesondert kenntlich gemacht.

Tabelle Anhang A-3: Fortsetzung

	
<p>Ein Baublock setzt sich aus einer oder mehreren Sektionen zusammen. Während einer Sektion immer nur ein EST zugeordnet werden kann, kann ein Baublock auch aus mehreren Sektionen und somit aus unterschiedlichen EST bestehen.</p>	<p>Das Gebiet bildet die Bilanzgrenze. In der Regel setzt sich dieses [Gebiet] aus mehreren Baublöcken verschiedener (EST) zusammen. [...]</p>

Neben Kartendaten lieferten Hegger et al. (2012; 2015a) auch Beispielbilder für die definierten EST (siehe Abbildung Anhang A-3). Diese Fotoaufnahmen und die sog. Schwarzpläne können den Ergebnissen dieser Arbeit (z. B. in Kapitel 6, Anhang C und Anhang D) gegenübergestellt werden.

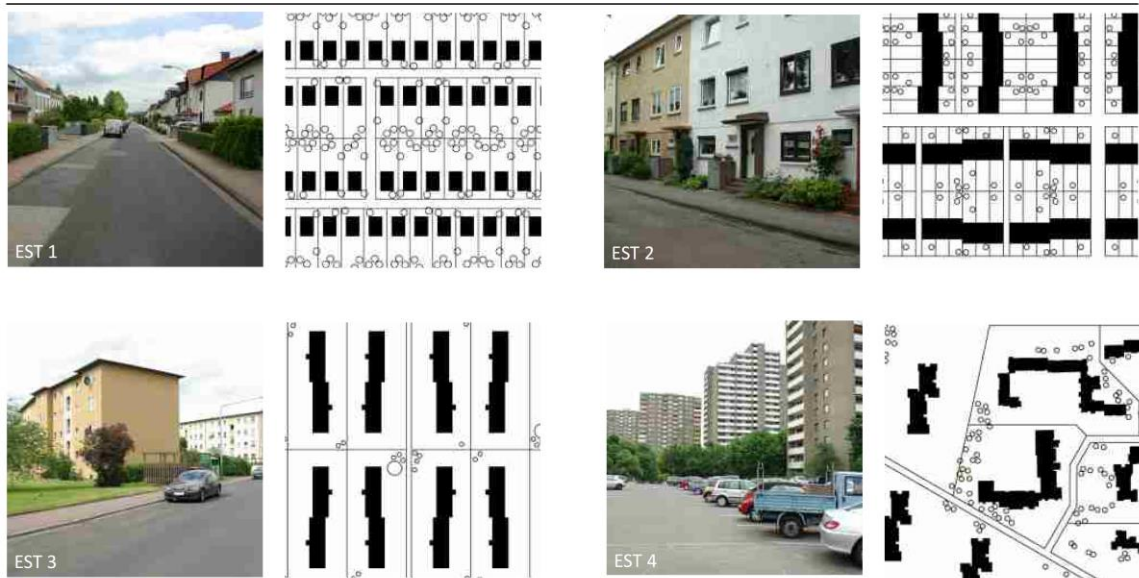


Abb. 6: Energetische Stadtraumtypen mit überwiegend Wohnnutzung (EST1 bis EST4)

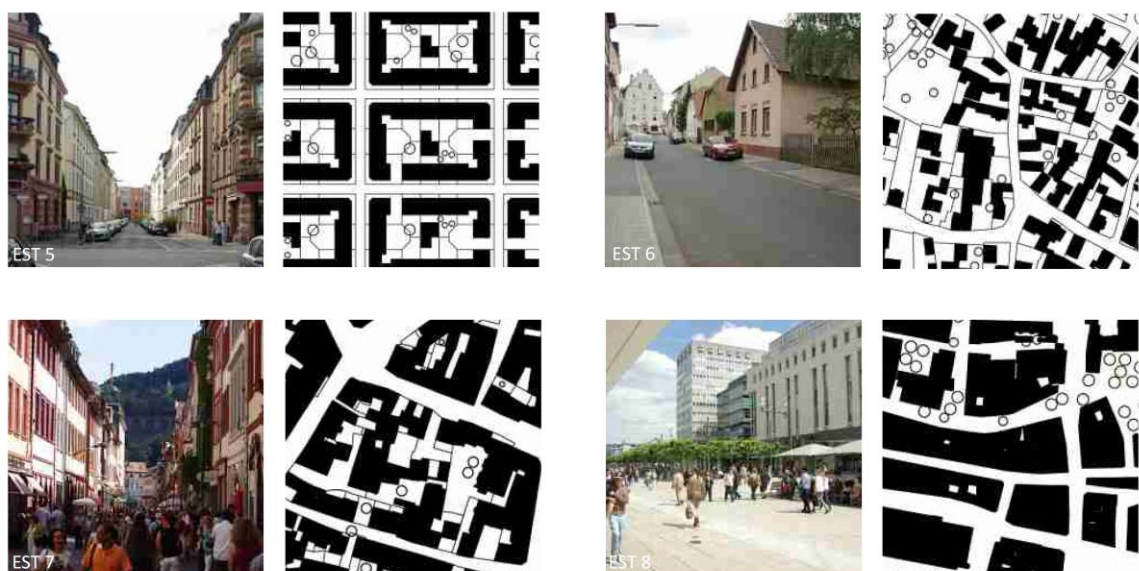


Abb. 7: Energetische Stadtraumtypen mit überwiegend Mischnutzung (EST5 bis EST8)

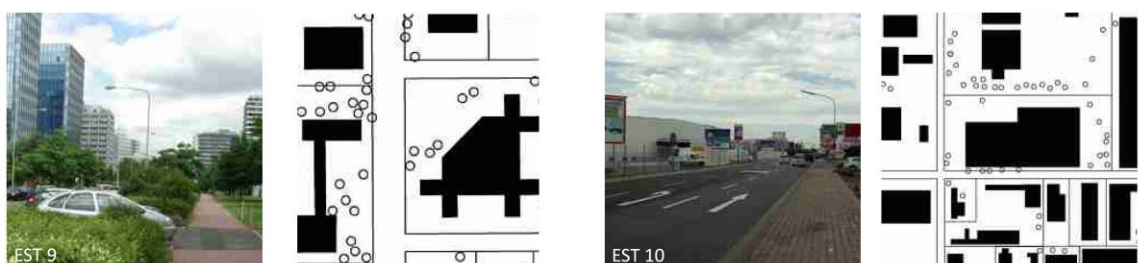


Abb. 8: Energetische Stadtraumtypen mit überwiegend Büro- bzw. gewerblicher Nutzung (EST9 bis EST10)

Abbildung Anhang A-3: Reale Beispiele und typische stadtmorphologische Strukturen der EST 1–10 nach Hegger et al. 2012; 2015a [Quelle: Hegger et al. (2015b, S. 6)]

ANHANG B ERGÄNZENDE DARSTELLUNG VON CLUSTER-ANALYSEBASIERTEN STUDIEN

Anhang B-1 Studien über die Entwicklungen von Clusteranalyseverfahren oder Untersuchungen unter Zuhilfenahme von Clusteranalysen

Tabelle Anhang B-1: Weiterführende Quellen über die Entwicklung, den Vergleich oder die Anwendung von Clusteranalyseverfahren [eigene Zusammenstellung]

Verfahrensart	Quelle	Inhalt
Partitionierend	Huang (1998) Huang (2017)	k-means ähnlicher Algorithmus unter Berücksichtigung von Prototypen (gleich den Medoiden) sowie angepasste Distanzberechnung für gemischtskalierte Datensätze; Zielfunktion, wie in k-means
	Foss et al. (2016)	Semiparametrische k-means-Clusteranalyse für gemischtskalierte Datensätze
Dichtebasierend	Ankerst et al. (1999)	Algorithmus OPTICS zur Parametervariation und Ermittlung einer übergeordneten Dichte-Clusterstruktur
	Borah und Bhattacharyya (2004)	Stichprobenbasierter DBSCAN-Algorithmus für große Datensätze
	Karami und Johansson (2014)	Automatische Festlegung der Verfahrensparameter ϵ und <i>MinPts</i>
	Khan et al. (2018)	Adaptiver DBSCAN-Algorithmus für Datensätze mit variierender Objektdichte

ANHANG C WEITERFÜHRENDE INFORMATIONEN ZU GEOINFORMATIONSDATEN

Anhang C-1 Objektartenbereiche und -gruppen des ALKIS

Tabelle Anhang C-1: Objektartenbereiche und Gruppen des ALKIS [eigene Darstellung nach AdV (2018b)]

Bereich	Gruppe
Flurstücke, Lage, Punkte	
	Angaben zum Flurstück
	Angaben zur Lage
	Angaben zum Netzkpunkt
	Angaben zum Punktort
	Fortführungsnachweis
	Angaben zur Reservierung
	Angaben zur Historie
Eigentümer	
	Personen- und Bestandsdaten
Gebäude	
	Angaben zum Gebäude
Tatsächliche Nutzung	
	Siedlung
	Verkehr
	Vegetation
	Gewässer
Bauwerke, Einrichtungen und sonstige Angaben	
	Bauwerke und Einrichtungen in Siedlungsflächen
	Bauwerke, Anlagen und Einrichtungen für den Verkehr
	Besondere Angaben zum Gewässer
	Besondere Angaben zum Verkehr
	Besondere Anlagen auf Siedlungsflächen
	Besondere Eigenschaften von Gewässern
	Besondere Vegetationsmerkmale
Relief	
	Digitales Höhenmodell
	Messdaten 3D
	Reliefformen
Gesetzliche Festlegungen, Gebietseinheiten, Kataloge	
	Administrative Gebietseinheiten
	Bodenschätzung, Bewertung
	Geographische Gebietseinheiten
	Kataloge
	Öffentlich-rechtliche und sonstige Festlegungen
Nutzerprofile	
	Angaben zu Nutzerprofilen
...	

Anhang

Tabelle Anhang C-2: Auszug aus dem ALKIS-Objektdatenkatalog für die Objektart „Gebäude“ und das Attribut „Bauweise“ [AdV (2018b)]

Attributart / Wertarten	Beschreibung	Wert
Bauweise	„Bauweise“ ist die Beschreibung der Art der Bauweise	
Freistehendes Einzelgebäude	'Freistehendes Einzelgebäude' ist ein freistehendes Wohngebäude mit in der Regel bis zu 2½ Geschossen (auch Villa, Landhaus, Bungalow).	1100
Freistehender Gebäudeblock	'Freistehender Gebäudeblock' ist ein freistehendes Wohngebäude (Mehrfamilienhaus), in der Regel 3 [sic!] - 8-geschossig.	1200
Einzelgarage	'Einzelgarage' ist eine einzeln stehende oder angebaute einzelne Garage als Abstellmöglichkeit für ein Fahrzeug, meist einen Pkw (auch eine einzelne Garage in Garagenhöfen).	1300
Doppelgarage	'Doppelgarage' ist eine einzeln stehende oder angebaute Garage als Abstellmöglichkeit für zwei Fahrzeuge, meist zwei Pkw.	1400
Sammelgarage	'Sammelgarage' ist eine Garage mit Abstellmöglichkeit für mehr als zwei Fahrzeuge, meist Pkw.	1500
Doppelhaushälfte	'Doppelhaushälfte' ist ein dreiseitig freistehendes Wohnhaus mit in der Regel bis zu 2½ Geschossen, an dem ein im allgemeinen gleichartiges Wohnhaus angebaut ist (etwa gleicher Baustil und ungefähr gleiche Baumaße).	2100
Reihenhaus	'Reihenhaus' ist eines von mehr als 2 gleichartig aneinanderggebauten Wohnhäusern mit in der Regel bis zu 2½ Geschossen in einer geschlossenen Häuserzeile (etwa gleicher Baustil und ungefähr gleiche Baumaße).	2200
Haus in Reihe	'Haus in Reihe' ist eines von mehr als zwei ungleichartigen, aneinanderggebauten Wohnhäusern, in der Regel mit bis zu 2 ½ Geschossen (z. B. in geschlossener Bauweise errichtete Wohngebäude in alten Ortskernen).	2300
Gruppenhaus	'Gruppenhaus' ist eines von mehr als 2 gleichartigen, aneinanderggebauten Wohnhäusern mit in der Regel bis zu 2½ Geschossen, die so gegeneinander verschoben sind, dass keine gemeinsame Achse gegeben ist.	2400
Gebäudeblock in geschlossener Bauweise	'Gebäudeblock in geschlossener Bauweise' ist eines von mehreren aneinanderggebauten Wohngebäuden (Mehrfamilienhäuser), in der Regel 3-8 geschossig; z. B. in Stadtkernen.	2500
Offene Halle	Unter einer "Offenen Halle" ist eine Halle zu verstehen, bei der alle vier Seiten offen sind. Hallen, [sic!] bei denen eine, zwei oder drei Seiten geschlossen sind, werden nur mit der entsprechenden Gebädefunktion ohne Belegung der Bauweise erfasst. Die offenen Gebäudeseiten sind mit "Besondere Gebäudelinie" Beschaffenheit "Offene Gebäudelinie" zu belegen.	4000
Sonstiges	'Sonstiges' bedeutet, dass das Gebäude eine Bauweise aufweist, die bekannt, aber nicht in der Liste der Wertarten aufgeführt ist.	9999

Anhang C-2 Vergleich der in ALKIS und im AdV-CityGML-Referenzmodell berücksichtigten mit den im Datensatz der 3D-Gebäudemodelle (LoD2) verfügbaren Gebäudeattribute und Metainformationen

Wie bereits in Kapitel 5.2 erläutert erlaubt das AdV-Referenzmodell die Übertragung einer begrenzten Zahl von Gebäudeattributen aus den Datenquellen des ALKIS auf die 3D-Gebäudemodelle. Für den vom BKG zur Verfügung gestellten Datensatz im LoD2 liegen jedoch nur wenige dieser Daten vor. Eine Gegenüberstellung der im AdV-Modell berücksichtigten Attribute und der verfügbaren Merkmale im für die Methodenentwicklung genutzten Datensatz für Berlin wird in der nachfolgenden Tabelle Anhang C-3 vorgenommen. Aus dieser Gegenüberstellung wird ersichtlich, dass einige, für diese energetisch-stadt-räumliche Untersuchung sicherlich nützlichen Gebäudemerkmale, z. B. die Bauweise oder das Baujahr, in dem flächendeckend verfügbaren Datensatz der 3D-Gebäudemodellen nicht vorhanden sind. Auch in den anderen vom BKG zur Verfügung gestellten Daten sind diese Attribute nicht vorhanden.

Tabelle Anhang C-3: Vergleich der in ALKIS verfügbaren Gebäudemerkmale mit den im AdV-CityGML-Modell angelegten Attributen und den im verfügbaren Datensatz des BKG vorhandenen Gebäudemerkmalen⁸¹

Gebäudeattribute gemäß ALKIS-Objektdatenkatalog ⁸²	Vorgesehen in AdV-CityGML-Modell	Verfügbar in Datensatz 3D-Gebäudemodelle (LoD2) für Berlin
Baujahr	Ja	-
Bauweise	-	-
Dachgeschossausbau	-	-
Externer Verweis/externe ID	Ja	Ja
Gebäudefunktion	Ja	Ja
Gebäudehöhe (gemessene bzw. modellierte)	Ja	Ja
Gebäudekennzeichen	-	-
gehört (Besitzer; nicht im öffentlichen ALKIS-Datensatz enthalten)	-	-
gehört zu/hängt zusammen mit (Verweis auf Gebäude, die Zusammenhang bilden)	-	-
Gemeindeschlüssel	Ja	Ja

⁸¹ Tabelle Anhang C-3 stellt lediglich die Merkmale der ALKIS-Objektebene AX_Gebäude dar. Weiterführende, das Gebäude betreffende Attribute können zudem in der Objektebene AX_Bauteil vorliegen, sofern die Gebäudeabbildung im Datensatz mehrere Bauteile umfasst.

⁸² In dieser Tabelle wurde zum besseren Verständnis nicht immer der im ALKIS-Objektdatenkatalog angegebene Merkmalsname verwendet.

Tabelle Anhang C-3: Fortsetzung

Gebäudeattribute gemäß ALKIS-Objektdatenkatalog ⁸³	Vorgesehen in AdV-CityGML- Modell	Verfügbar in Datensatz 3D-Gebäudemodelle (LoD2) für Berlin
Geschossfläche	-	-
Grundfläche	(indirekt)	(indirekt)
hängt zusammen mit	-	-
hat (Gebäuelagebezeichnung)	-	-
Hochhaus	-	-
Jahr des Abrisses	Ja	-
Name (<i>optionales Attribut</i>)	Ja	(Ja)
Nutzung	Ja	-
Objektidentifikator	Ja	Ja
Regionalsprache	-	-
Stockwerke oberhalb der Erdoberfläche	Ja	-
Stockwerke unterhalb der Erdoberfläche	Ja	-
verschiedene Qualitätsangaben/ Metainformationen	Ja	Ja
weitere Gebäudefunktion	-	-
zeigt auf/zeigt auf ohne (Verweis auf Lagebezeichnung mit/ohne Hausnummer)	-	-
Zustand	-	-
Zweitname	-	-
Gebäudeklasse	-	-

Anhang C-3 Für die Methodenentwicklung eingesetzte und als in besonderer Weise wichtig bewertete R-Packages

Ergänzend zu den Beschreibungen der softwareseitigen Ressourcen in Kapitel 5.2 wird in nachfolgender Tabelle eine Übersicht gegeben, welche R-Packages für die Methodenentwicklung einen besonderen Stellenwert besitzen.

⁸³ In dieser Tabelle wurde zum besseren Verständnis nicht immer der im ALKIS-Objektdatenkatalog angegebene Merkmalsname verwendet.

Tabelle Anhang C-4 Für die Methodenentwicklung verwendeten R-Packages

Name	Autor(en)
Clusteranalyse/ML	
<i>cluster</i> : “Finding Groups in Data”: Cluster Analysis Extended Rousseeuw et al.	Maechler et al. (2021)
<i>fpc</i> : Flexible Procedures for Clustering	Hennig (2020)
<i>klaR</i> : Classification and Visualization	Roever et al. (2022)
<i>stats</i> : The R Stats Package	R Core Team (2021)
Datenbankschnittstellen	
<i>DBI</i> : R Database Interface	R-SIG-DB et al. (2022)
<i>RPostgres</i> : Rcpp Interface to PostgreSQL	Wickham et al. (2022)
<i>RPostgreSQL</i> : R Interface to the ‘PostgreSQL’ Database System	Conway et al. (2022)
Geodatenverarbeitung	
<i>foot</i> : An R package for processing building footprints	WPGP (2021)
<i>lwgeom</i> : Bindings to Selected ‘liblwgeom’ Functions for Simple Features	Pebesma et al. (2022)
<i>rgdal</i> : Bindings for the ‘Geospatial’ Data Abstraction Library	Bivand et al. (2022)
<i>Rpostgis</i> : R Interface to a ‘PostGIS’ Database	Basille und Bucklin (2019)
<i>sf</i> : Simple Features for R	Pebesma (2018)
<i>sp</i> : Classes and Methods for Spatial Data	Pebesma und Bivand (2005); Bivand et al. (2013)

Anhang C-4 Ergänzende Ausführungen zur Überprüfung der für diese Arbeit vorliegenden GIS-Daten

Im Datensatz der 3D-Gebäudemodelle werden die Gebäude in Abhängigkeit von der Dachform und der Gebäudehöhe als unterschiedliche Gebäudeteile abgebildet. Hierbei lässt das AAA-Referenzmodell den katasterführenden Stellen gewisse Freiheiten in der Modellierung, was sich in den flächendeckenden Basisdaten des BKG widerspiegelt (siehe auch Tabelle 5-2, S. 83). In der LoD2-Modellierung des BKG wird den Gebäudeteilen ein gemeinsamer Gebäudeidentifikator zugeordnet, um den Gebäudezusammenhang der Gebäudeteile jederzeit nachvollziehen zu können.

Nachfolgend wird untersucht, ob für die Beantwortung der Fragestellung dieser Arbeit irrelevante Gebäudeteile über definierte Kriterien und Regeln identifiziert sowie aus dem Datensatz eliminiert werden können. Die Modellabbildungen werden zu diesem Zweck genauer betrachtet und die Zusammensetzung von Gebäuden aus Gebäudeteilen beispielhaft für unterschiedliche Bebauungssituationen überprüft sowie der realen Bebauungssituation gegenübergestellt. Hintergrund sind, neben eigenen Untersuchungen zu den

Modellierungsunterschieden der Geodatenquellen ALKIS und LoD2, auch Erkenntnisse aus Schwarz et al. (2021), Hartmann et al. (2016) und Hörner (2021).

Die Untersuchung des ALKIS- sowie des LoD2-Datensatzes für die Großstadt Berlin zeigt, dass bei der Modellierung von Gebäuden Anbauten von Gebäuden, wie Balkone oder Vordächer, meist als Gebäudeteile mit gleichem Merkmalswert für die Gebäudefunktion, wie das Hauptgebäude modelliert wurden. Nur ein geringer Teil der Anbauten ist als eigenständiges Objekt im LoD2-Geodatensatz als Bauwerk gekennzeichnet. Eine robuste Regel, zur Unterscheidung anhand festgelegter Differenzen der Gebäudehöhe sowie der Grundflächen zwischen zwei Gebäudeteilen, konnte durch die Untersuchung von zufälligen Beispielgebäuden nicht abgeleitet werden. Die Erstellung eines robusten Verfahrens zur Elimination von Anbauten erfordert vielmehr eine detaillierte Untersuchung und manuelle Klassifizierung der Gebäudeteile einer (am besten sogar repräsentativen) Gebäudestichprobe. Der hierfür notwendige Arbeitsumfang konnte im Rahmen der Methodenentwicklung zur Typisierung des deutschen Quartiersbestandes nicht geleistet werden. Abbildung Anhang C-1 stellt jedoch eines der untersuchten Beispiele dar und macht deutlich, dass sich die Abweichungen der geometrischen Gebäudekennzahlen aufgrund der Berücksichtigung energetisch irrelevanter Gebäudeteile in Grenzen halten. Für das gezeigte Mehrfamilienhaus liegen die Flächen- und Volumenabweichungen aufgrund der Berücksichtigung von Vordächern bei der Vereinigung von Gebäudeteilen zu einem Gebäude bei etwa 10 % respektive 1 %.

Mehrfamilienhaus (Zeilenbebauung)

Grundfläche	Gebäude:	611 m ²
	Vordach-Süd:	25 m ²
	Vordach-Nord:	25 m ²

Erhöhung Grundfläche ca. 10 %

Volumen	Gebäude:	18.330 m ³
	Vordach-Süd:	98 m ³
	Vordach-Nord:	98 m ³

Erhöhung Gebäudevolumen ca. 1 %

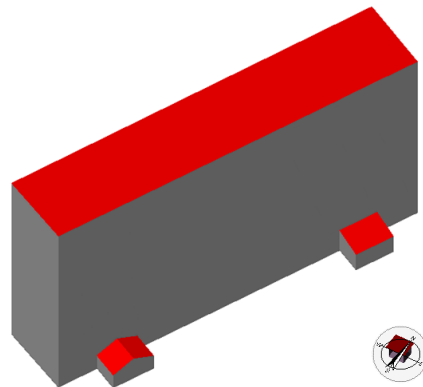


Abbildung Anhang C-1: Beispielhafte Überprüfung energetisch irrelevanter Gebäudeteile (hier: Gebäudeteile zur Erschließung)

Weiterhin wurde überprüft, ob die Festlegung einer minimal zulässigen Grundfläche als Kriterium zur Berücksichtigung von Gebäuden möglich ist. Diese Untersuchung leitet sich aus dem in Hartmann et al. (2016) formulierten Kriterium der Mindestgröße der Gebäudegrundfläche von 10 m² ab. Für den Datensatz wurden die Gebäude deshalb in zehn

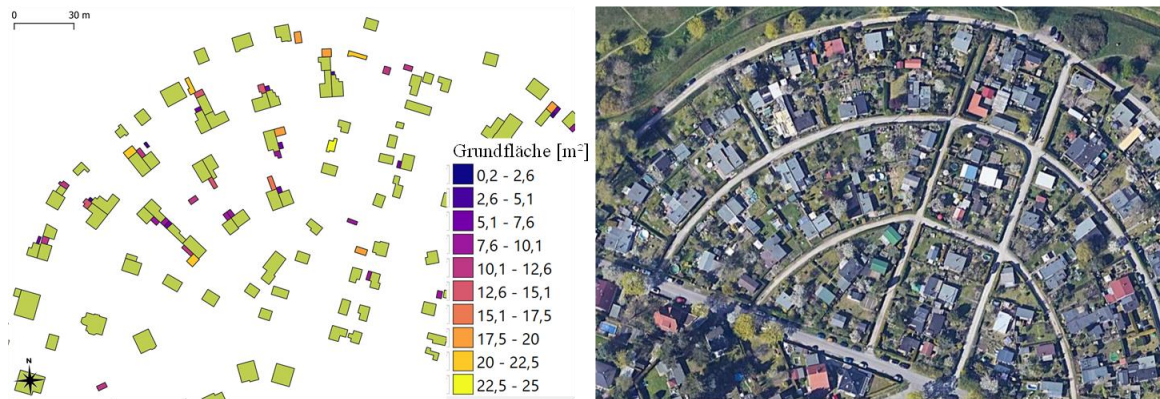


Abbildung Anhang C-2: Überprüfung des Abgrenzungskriteriums Mindestgrundfläche zum Ausschluss von Gebäudeobjekten aus dem GIS-Datensatz – Beispiel Doppelhäuser (links: GIS-Datenauswertung [eigene Darstellung, Daten: Geo-Basis DE / BKG (2021)]; rechts: Reale Bebauungssituation (Berlin Kaulsdorf) [Bild: Google (2022), Kartendaten – GeoBasis-DE / BKG (2022), GeoContent, Maxar Technologies])

Grundflächenklassen von 0 m² bis 25 m² eingeteilt und einzelne Ausschnitte des Geodaten-satzes mit der realen Bebauungssituation verglichen. Es zeigt sich, dass die Festlegung einer Mindestgrundfläche – insbesondere in weniger dicht besiedelten Siedlungsstrukturen – freistehende Gartenhäuser, Garagen, Schuppen und Gebäudeanbauten, z. B. Wintergärten, erfasst. Bei steigender Mindestfläche wird eine Abgrenzung und Einschätzung der Bebauungssituation in diesen Siedlungsbereichen jedoch zunehmend unsicherer und es ist anzunehmen, dass eine höhere Anzahl von Gebäuden aus dem Datensatz entfernt wird, die hinsichtlich der Bebauungsstruktur oder im Hinblick auf eine energetische Bewertung von Quartieren für die raumbezogene Analyse eine Bedeutung haben könnten (Abbildung Anhang C-2). Für höher verdichtete Siedlungsbereiche (siehe Abbildung Anhang C-3, S. 188) fallen nur wenige Gebäudeobjekte unter die höchste untersuchte untere Grenze von 25 m² Grundfläche.

Abbildung Anhang C-4 (S. 189) stellt zudem beispielhaft eine Bebauungssituation dar, in welcher das von Hartmann et al. (2016) ebenfalls vorgeschlagene Abgrenzungskriterium für untypische Grundrissformen dazu führt, dass Gebäude mit üblichen, in der Praxis funktionalen Grundrissen – hier Doppelhaushälften – aus dem Datensatz entfernt würden. Ein weiteres Beispiel stellt eine Bebauungssituation mit Reihenhäusern dar, in welcher die Abgrenzungskriterien ebenfalls zu einem fehlerhaften Ausschluss von Gebäudeobjekten führen (Abbildung Anhang C-7, S. 195).

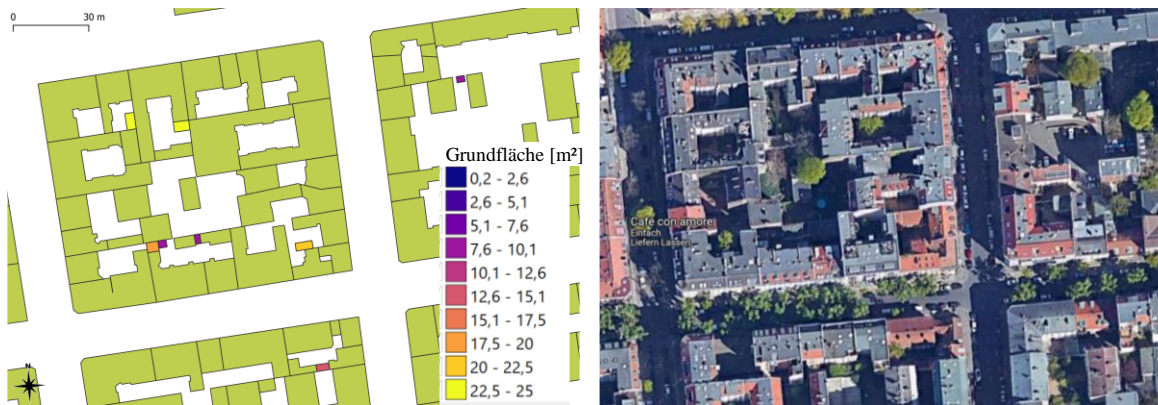


Abbildung Anhang C-3: Überprüfung des Abgrenzungskriteriums Mindestgrundfläche zum Ausschluss von Gebäudeobjekten aus dem GIS-Datensatz – Beispiel Blockrandbebauung (links: GIS-Datenauswertung [eigene Darstellung, Daten: Geo-Basis DE/BKG (2021)]; rechts: Reale Bebauungssituation (Berlin Charlottenburg) [Bild: Google (2022), Kartendaten – GeoBasis-DE/BKG (2022), GeoContent, Maxar Technologies])

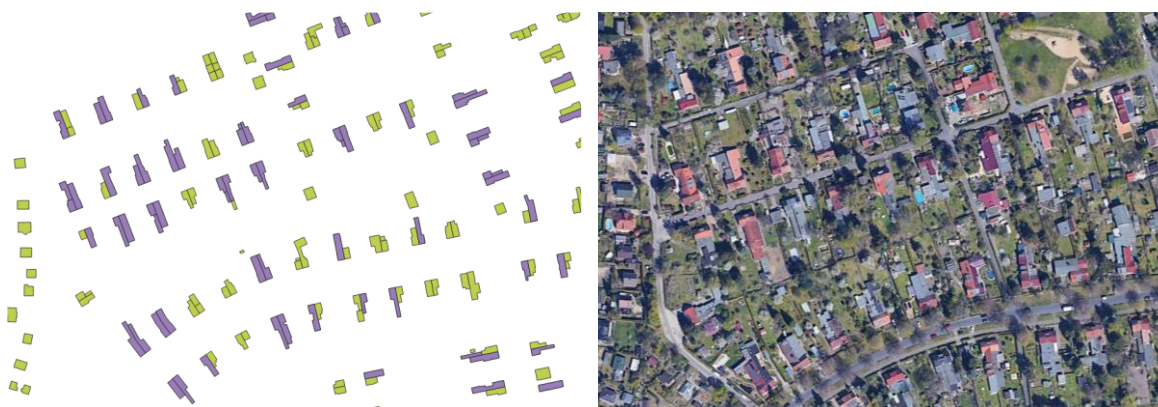


Abbildung Anhang C-4: Beispiel eines fehlerhaften Ausschlusses von Doppelhäusern durch die Abgrenzungskriterien Breite-zu-Länge-Verhältnis $\geq 0,3$ und Grundfläche ≥ 190 m (violett: Kriterien nicht erfüllt; links: GIS-Datenauswertung [eigene Darstellung, Daten: Geo-Basis DE/BKG (2021)]; rechts: Reale Bebauungssituation (Berlin Stadtrandsiedlung Malchow) [Bild: Google (2022), Kartendaten: GeoBasis-DE/BKG (2022)])

Auch für das mögliche Abgrenzungskriterium der Gebäudehöhe sind Beispiele der Überprüfung des Datensatzes in Abbildung Anhang C-5 und Abbildung Anhang C-6 gegeben. Für die Siedlungsstruktur mit freistehenden Häusern oder Doppelhäusern, welche zudem einen großen Teil an Bungalowbauten aufweist, fallen bei einer unteren Grenze der Gebäudehöhe von > 3 m bereits viele der Wohngebäude unter das Ausschlusskriterium (Abbildung Anhang C-5). Gleichfalls wäre in einem anderen Baubauungskontext auch ein höherer Grenzwert von bis zu 3,5 m möglich, bei dem weiterhin nur Gartenhäuser und überdachte Terrassen bzw. Wintergärten aus dem Datensatz entfernt würden, ohne fälschlicherweise auch Wohngebäude zu erfassen (Abbildung Anhang C-6).

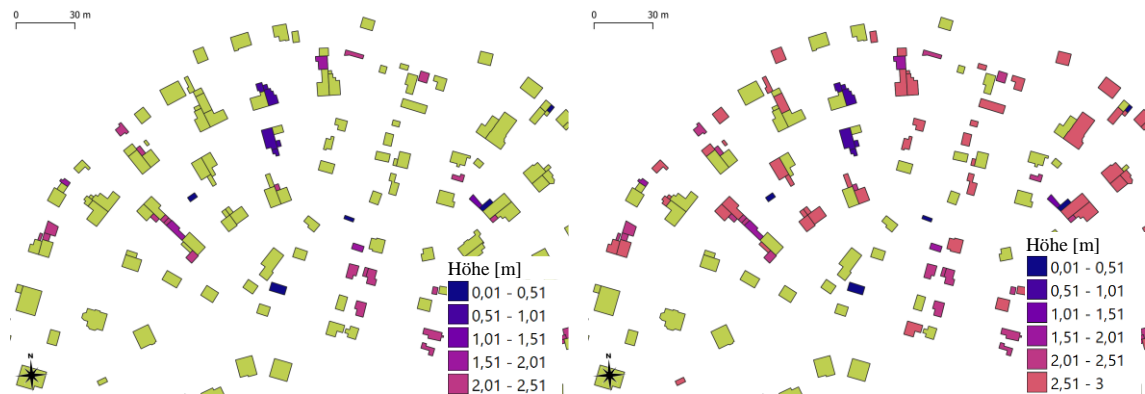


Abbildung Anhang C-5: Überprüfung der Festlegung des Abgrenzungskriteriums Gebäudehöhe zum Ausschluss von Gebäudeobjekten aus dem GIS-Datensatz – Beispiel Doppelhäuser (Berlin Kaulsdorf; links $< 2,5\text{ m}$; rechts $< 3\text{ m}$) [eigene Darstellung, Daten: Geo-Basis DE/BKG (2021)]

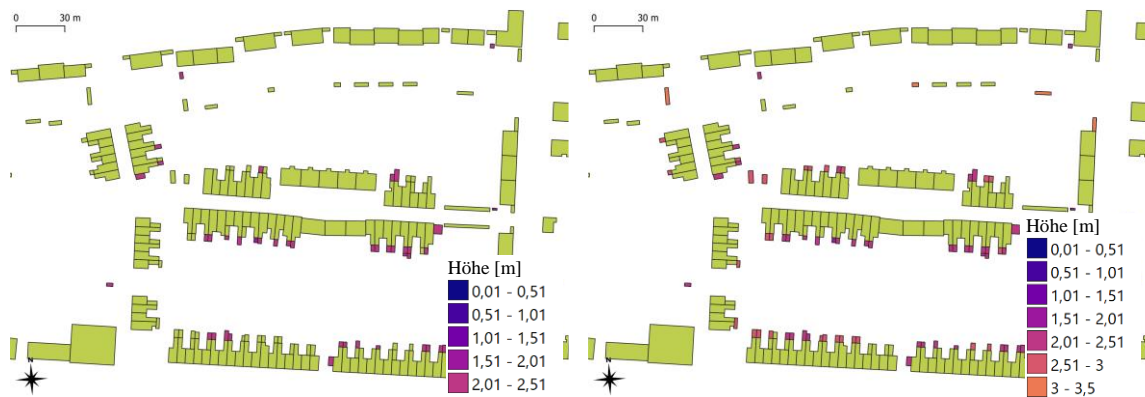


Abbildung Anhang C-6: Überprüfung der Festlegung des Abgrenzungskriteriums Gebäudehöhe zum Ausschluss von Gebäudeobjekten aus dem GIS-Datensatz – Beispiel Reihenhäuser (Berlin Staaken; links $< 2,5\text{ m}$; rechts $< 3,5\text{ m}$) [eigene Darstellung, Daten: Geo-Basis DE/BKG (2021)]

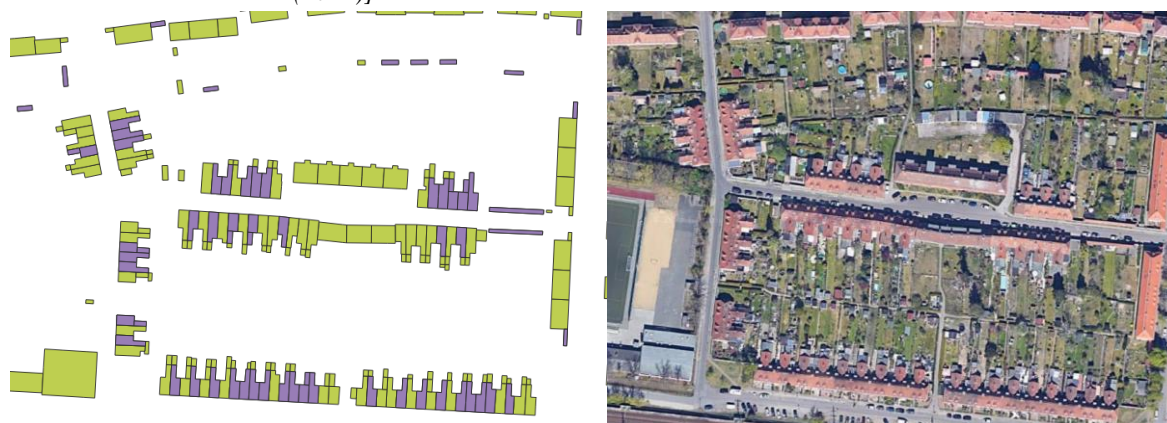


Abbildung Anhang C-7: Beispiel eines fehlerhaften Ausschlusses von Reihenhäusern durch die Abgrenzungskriterien Breite-zu-Länge-Verhältnis $\geq 0,3$ und Grundfläche $\geq 190\text{ m}^2$ (violett: Kriterien nicht erfüllt; links: GIS-Datenauswertung [eigene Darstellung, Daten: Geo-Basis DE/BKG (2021)]; rechts: Reale Bebauungssituation (Berlin Staaken) [Bild: Google (2022), Kartendaten: GeoBasis-DE/BKG (2022)])

Anhang C-5 Exkurs über die Möglichkeiten und Grenzen einer ingenieurmäßigen Typisierung der Gebäudeabbildungen

Im Rahmen dieser Arbeit wurde ebenfalls untersucht, inwieweit bestehende Ansätze zur Gebäudetypisierung und deren Anwendung im Kontext einer Quartierstypisierung möglich sind. Eine solche Bezugnahme auf Gebäudetypologien erfolgte z. B. in den Studien von Roth (1980) und Hegger et al. (2012; 2015a). Ein charakteristisches Gebäudemerkmal der dort berücksichtigten Gebäudetypen ist eine Bebauung als freistehende Ein- und Zweifamilienhäuser oder als Doppel-, Reihenhaus- oder Blockrandbebauung. Weitere Merkmale sind das Gebäudealter bzw. die Baualtersklasse und einhergehende (typische) Wärmeschutzniveaus sowie die Geschossigkeit und die Dachform. In den Datensätzen der Geodatenbasis für Deutschland liegt das Attribut der Bebauungsart auf Gebäudeebene nicht vor und kann ohne weiterführende empirische Analysen des realen Gebäudebestandes nicht aus den GIS-Daten abgeleitet werden. Das Vorhandensein von angrenzenden Gebäuden kann zwar ohne Weiteres durch raumbezogene Abfragen in der PostGIS-Datenbank ermittelt werden. Ursächlich für die Notwendigkeit eines auf Empirie gestützten Abgleichs der Geodaten mit der Realität ist jedoch die unscharfe Zuordnung von Gebäudefunktionen innerhalb des 3D-Gebäudemodelldatensatzes. Die unklare Datenlage bzgl. der realen Gebäudefunktion lässt keine sichere Unterscheidung der Nachbarbebauung eines Gebäudes zu, d. h. eine Bewertung, ob angrenzende Geobjekte Gebäude zur dauerhaften Nutzung oder z. B. Gartenhäuser oder Gewächshäuser sind. Über die Relevanz dieser Anbauten für die in dieser Arbeit formulierte Fragestellung kann somit nicht abschließend entschieden werden. Auch ein freistehendes Einfamilienhaus mit angrenzendem Gartenhaus wird ohne empirisch abgesicherte Modellierungsvorschriften als ein Gebäude mit einem oder mehreren Nachbargebäuden erfasst, sofern das Gartenhaus ebenfalls mit dem Attribut „Wohnfunktion“ im Geodatensatz abgebildet ist. Ohne empirische Begleitstudie wäre auch der systematische Fehler eines solchen Vorgehens nicht abschätzbar.

In den Geobasisdaten fehlen darüber hinaus jegliche Informationen zum Baualter bzw. der Baualtersklasse der Gebäude. Auch liegen keinerlei Informationen zu wärmeschutzrelevanten Gebäudeparametern vor, weshalb diese Attribute nicht zur Klassifizierung der Gebäude herangezogen werden können. Die Geschossigkeit der Gebäude kann aus dem im Geodatensatz der 3D-Gebäudemodelle vorliegenden Merkmal der gemessenen Gebäudehöhe abgeschätzt werden. Eine Modellierung ist für dieses Merkmal durch die Annahme einer typischen Geschosshöhe möglich. Auch die Dachform ist durch die detaillierte Abbildung standardisierter Dachformen der Gebäudemodelle im LoD2 vorhanden. Hierbei muss jedoch mit in der Realität sehr viel komplexeren Gebäudestrukturen umgegangen, d. h. das Vorliegen unterschiedlicher Dachformen mehrerer Gebäudeteile eines Gebäudes berücksichtigt werden. Für eine ingenieurmäßige Einteilung von Gebäuden als Basis einer Analyse der Siedlungsstrukturen höherer räumlicher Ebene, welche sich an die verwendeten

Gebäudetypen nach Roth (1980), Hegger et al. (2015a) oder an die deutsche Wohngebäudetypologie des IWU (Loga et al. 2015) anlehnen, stehen damit nur wenige Merkmale zur Verfügung.

Anhang C-6 Ermittlung weiterführender Gebäudemerkmale aus den Geometriedaten der GIS-Abbildung

Zur Berechnung weiterführender geometrischer Kennzahlen der Gebäudegrundrisse sowie Gebäuderegionen werden Standardfunktionen der PostGIS-Datenbankerweiterung sowie Funktionen der R-Packages „lwgeom“ (Pebesma et al. 2022) und „foot“ verwendet (WPGP 2021). Eine vereinfachte Beschreibung der Kennzahlen soll nachfolgend gegeben werden, um deren Bedeutung für die energetisch-stadtstrukturelle Untersuchung dieser Arbeit im Anschluss herleiten zu können.

Tabelle Anhang C-5: Beschreibung der Berechnung weiterführender Gebäudemerkmale

Kennzahl	Software Funktion	Beschreibung
Anzahl direkt angrenzender Gebäude	PostGIS <i>st_intersects()</i>	Raumbezogene Abfrage, wie viele Polygone des Datensatzes das jeweilige Grundrisspolygon berühren oder schneiden
Roeck-Wert	PostGIS <i>st_area()</i> <i>st_minimumboundingcircle()</i>	Quotient aus der Grundrissfläche und der Fläche des kleinsten den Grundriss umfassenden Kreises
Perimeterlänge	R-Package foot <i>fs_perimeter()</i>	Umfang des Gebäudegrundrisses
Kompaktheitszahl	R-Package foot <i>fs_compact()</i>	Quotient aus Perimeterlänge und Fläche eines Kreises mit gleichem Umfang (Kompaktheitsmaß nach Schwartzberg; vgl. Ulrich (2014; Young))
Länge-zu-Breite-Verhältnis	R-Package foot <i>fs_lwratio()</i>	Verhältnis von Länge zu Breite der Seiten des kleinsten, rotierten umfassenden Rechtecks eines Grundrisses (mit Länge = längere Kante; Breite = kürzere Kante)

Die Berechnung von Informationen zur Gebäuderegion eines Gebäudeobjekts im Geodatenatz erfolgt nach der Vereinigung der einzelnen Gebäude zu Gebäuderegionen. Eine Gebäuderegion entspricht dabei Grundrissen, welche in direktem Kontakt mit einem Nachbarn sind und über diesen die Bildung einer Kette von Grundrissen mit indirekt benachbarten Gebäuden ermöglichen. Hierfür werden in der PostGIS-Datenbank eine entsprechende Abfrage durchgeführt und neben der Anzahl der Gebäudeobjekte innerhalb einer Gebäuderegion auch geometrische Kennzahlen für die Gebäuderegionen ermittelt.

**ANHANG D ERGEBNISSE WEITERFÜHRENDER
STATISTISCHER AUSWERTUNGEN AUF VER-
SCHIEDENEN RAUMBEBENEN**

Anhang D-1 Vergleich der zweidimensionalen Visualisierung der Geodatenabbildung ausgewählter Siedlungsgebiete in Berlin mit Satelliten- und Schrägbildaufnahmen



Abbildung Anhang D-1: Zur Plausibilisierung der mittels Clusteranalyse erreichten Gruppeneinteilung herangezogene Siedlungsstrukturen in Berlin [Quelle: Google 2022, GeoBasis-DE/BKG, GeoContent, Maxar Technologies 2022]

Anhang D-2 Ergänzende bebauungsstrukturbeschreibende Kennzahlen und Abbildungen der Teilbauflächen in Berlin

Die statistischen Auswertungen der Teilbauflächen in Berlin (Kapitel 6.3) werden nachfolgend durch die Angabe statistischer Maßzahlen und weiterführende Abbildungen ergänzt. Es ist anzumerken, dass die Kennzahlen auf der Anwendung der beschriebenen Filter- bzw. Abgrenzungskriterien für energetisch-städtebaulich nicht relevante Gebäude und somit einem im Vergleich zum Datenbestand des BKG reduzierten Datensatz fußen. Die nachfolgenden Darstellungen besitzen überwiegend informativen Charakter, um dem Leser die Möglichkeit zu geben, sich ein besseres Verständnis über die energetisch relevanten Merkmale der Quartiere zu erarbeiten.

Tabelle Anhang D- 1: Ergänzende Maßzahlen zur vollständigen Beschreibung des Verlaufs der Dichtefunktionen für bebauungsstrukturbeschreibende Merkmale auf Ebene von Teilbauflächen

Merkmal	MW	SD	Variationskoeffizient	Schiefe	Kurtosis
MW der Gebäudegrundfläche [m ²]	503,4	971,2	1,9	10,1	189,6
SD der Gebäudegrundfläche [m ²]	238,9	595,5	2,5	9,3	145,8
MW der mittleren Gebäudehöhe [m]	12,0	6,8	0,6	8,3	6,5
SD der mittleren Gebäudehöhe [m]	0,8	1,0	1,0	4,9	52,1
Spez. Anzahl Gebäude [1/ha]	9,3	7,5	0,8	1,3	3,3
Spez. Anzahl Gebäuderegionen [1/ha]	5,3	4,7	0,9	1,1	0,6
MW der Regionengrundfläche [m ²]	2.415,0	4.908,8	2,03	11,1	366,3
SD der Regionengrundfläche [m ²]	547,2	1.702,9	3,11	22,8	1.193,2
MW der Gebäudeanzahl je Region [-]	9,7	14,5	1,70	2,8	11,4
MBN I [m ² /m ²]	0,21	0,13	0,63	0,97	1,3
MBN II [m ³ /m ²]	0,89	1,91	2,15	5,0	34,8
Legende: MW = Mittelwert, SD = empirische Standardabweichung (der Grundgesamtheit), MBN I = Maß baulicher Nutzung I, MBN II = Maß baulicher Nutzung II					

Die den Tabellenwerten zugehörigen Dichteverläufe können den nachfolgenden Grafiken entnommen werden.

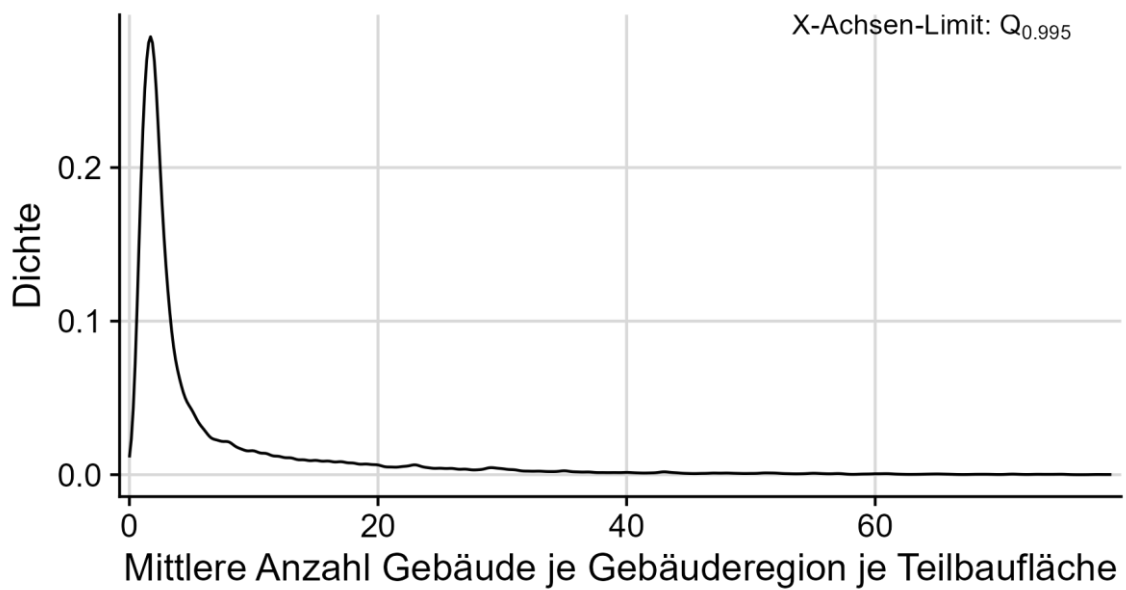


Abbildung Anhang D- 2: Dichteverlauf der Merkmalsausprägung „Mittlere Anzahl Gebäude je Gebäuderegion“ auf Ebene der Teilbauflächen für das Anwendungsbeispiel Berlin

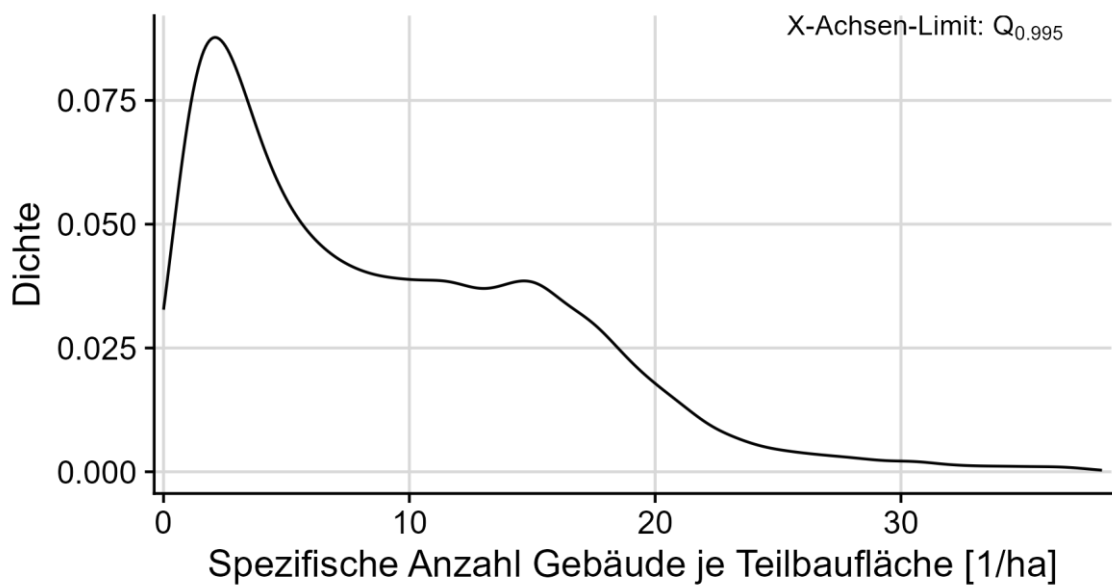


Abbildung Anhang D- 3: Dichteverlauf der Merkmalsausprägung „Spezifische Anzahl Gebäude“ auf Ebene der Teilbauflächen für das Anwendungsbeispiel Berlin

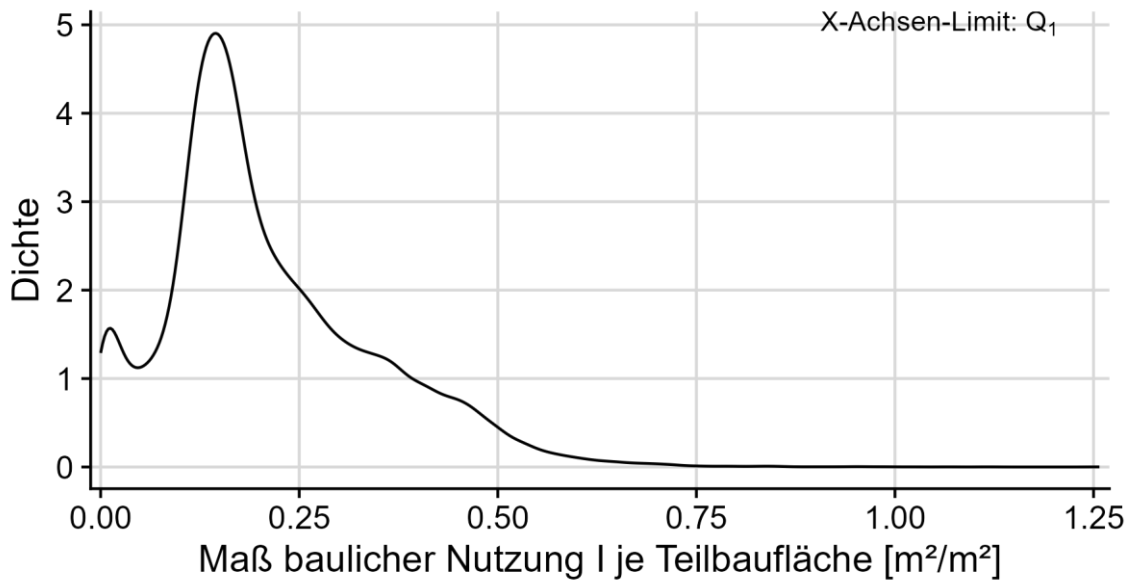


Abbildung Anhang D- 4: Dichteverlauf der Merkmalsausprägung „Maß baulicher Nutzung I“ auf Ebene der Teilbauflächen für das Anwendungsbeispiel Berlin

Anhang D-3 Ergänzende bebauungsstrukturbeschreibende Kennzahlen und Abbildungen der Quartiere in Berlin

Die statistischen Auswertungen des Quartiersbestandes in Berlin (Kapitel 6.6) werden nachfolgend durch weiterführende statistische Kennzahlen und Abbildungen ergänzt. Diesen Kennzahlen liegt die Einteilung der Teilbauflächen in Berlin in 6 Gruppen durch die Anwendung des clusteranalytischen Algorithmus PAM basierend auf einer mittels Gower-Metrik berechneten Distanzmatrix zugrunde. Die nachfolgenden Darstellungen besitzen überwiegend informativen Charakter, um dem Leser die Möglichkeit zu geben, sich ein besseres Verständnis über die energetisch relevanten Merkmale der Quartiere zu erarbeiten.

Tabelle Anhang D- 2: Ergänzende Maßzahlen zur vollständigen Beschreibung des Verlaufs der Dichtefunktionen für bebauungsstrukturbeschreibende Merkmale auf Ebene von Quartieren und unterteilt nach Quartiersgruppen

Variablen	Quartiersgruppen					
	1	2	3	4	5	6
MW Quartiersfläche [ha]	11,36	4,32	6,10	6,78	1,54	6,37
SD Quartiersfläche [ha]	22,23	8,47	10,11	9,35	2,09	18,04
Schiefe Quartiersfläche [-]	4,8	6,3	4,2	2,7	4,1	10,6
Kurtosis Quartiersfläche [-]	34,5	56,9	26,3	8,2	22,6	164,4
MW Gesamtgebäudefläche je Quartier [1.000 m ²]	14,81	6,58	12,96	23,53	6,21	1,59
SD Gesamtgebäudefläche je Quartier [1.000 m ²]	28,06	13,17	23,43	34,83	9,11	3,06
Schiefe Gesamtgebäudefläche je Quartier [1.000 m ²]	4,6	6,5	4,5	3,0	4,1	6,2
Kurtosis Gesamtgebäudefläche je Quartier [1.000 m ²]	30,0	62,5	28,2	10,1	23,2	60,8
MW mittlere Gebäudegrundfläche je Quartier [m ²]	107,8	191,3	810,5	415,8	2.212,8	343,0
SD mittlere Gebäudegrundfläche je Quartier [m ²]	45,5	122,9	789,1	293,9	2.703,8	363,7
Schiefe mittlere Gebäudegrundfläche je Quartier [m ²]	3,9	2,7	7,8	4,0	4,1	3,1
Kurtosis mittlere Gebäudegrundfläche je Quartier [m ²]	28,0	13,0	112,2	23,8	21,9	15,5
MW flächengew. mittlere Gebäudehöhe je Quartier [m]	8,2	9,2	13,4	18,9	20,8	7,0
SD flächengew. mittlere Gebäudehöhe je Quartier [m]	2,1	3,1	4,9	4,1	7,9	4,3
Legende: MW = Mittelwert, SD = empirische Standardabweichung (der Grundgesamtheit)						

Anhang

Tabelle Anhang D- 2: Fortsetzung

Schiefe flächengew. mittlere Gebäudehöhe je Quartier [m]	0,7	0,6	1,1	-0,3	1,7	4,9
Kurtosis flächengew. mittlere Gebäudehöhe je Quartier [m]	2,6	1,2	4,1	2,6	8,5	41,5
MW Maß baulicher Nutzung I (MBN I) [m ² /m ²]	0,15	0,16	0,21	0,34	0,40	0,05
SD Maß baulicher Nutzung I (MBN I) [m ² /m ²]	0,06	0,08	0,07	0,08	0,13	0,04
Schiefe Maß baulicher Nutzung I (MBN I) [m ² /m ²]	2,57	1,29	0,65	-0,02	0,96	0,89
Kurtosis Maß baulicher Nutzung I (MBN I) [m ² /m ²]	12,01	4,11	1,94	0,85	3,22	-0,15
MW Maß baulicher Nutzung II (MBN II) [m ³ /m ²]	1,27	1,53	2,67	6,50	7,95	0,34
SD Maß baulicher Nutzung II (MBN II) [m ³ /m ²]	0,73	0,96	1,14	2,19	3,51	0,32
Schiefe Maß baulicher Nutzung II (MBN I) [m ² /m ²]	2,8	1,4	0,5	0,6	1,4	1,3
Kurtosis Maß baulicher Nutzung II (MBN I) [m ² /m ²]	13,0	3,5	0,1	0,7	4,5	3,5
Legende: MW = Mittelwert, SD = empirische Standardabweichung (der Grundgesamtheit), MBN I = Maß baulicher Nutzung I, MBN II = Maß baulicher Nutzung II						

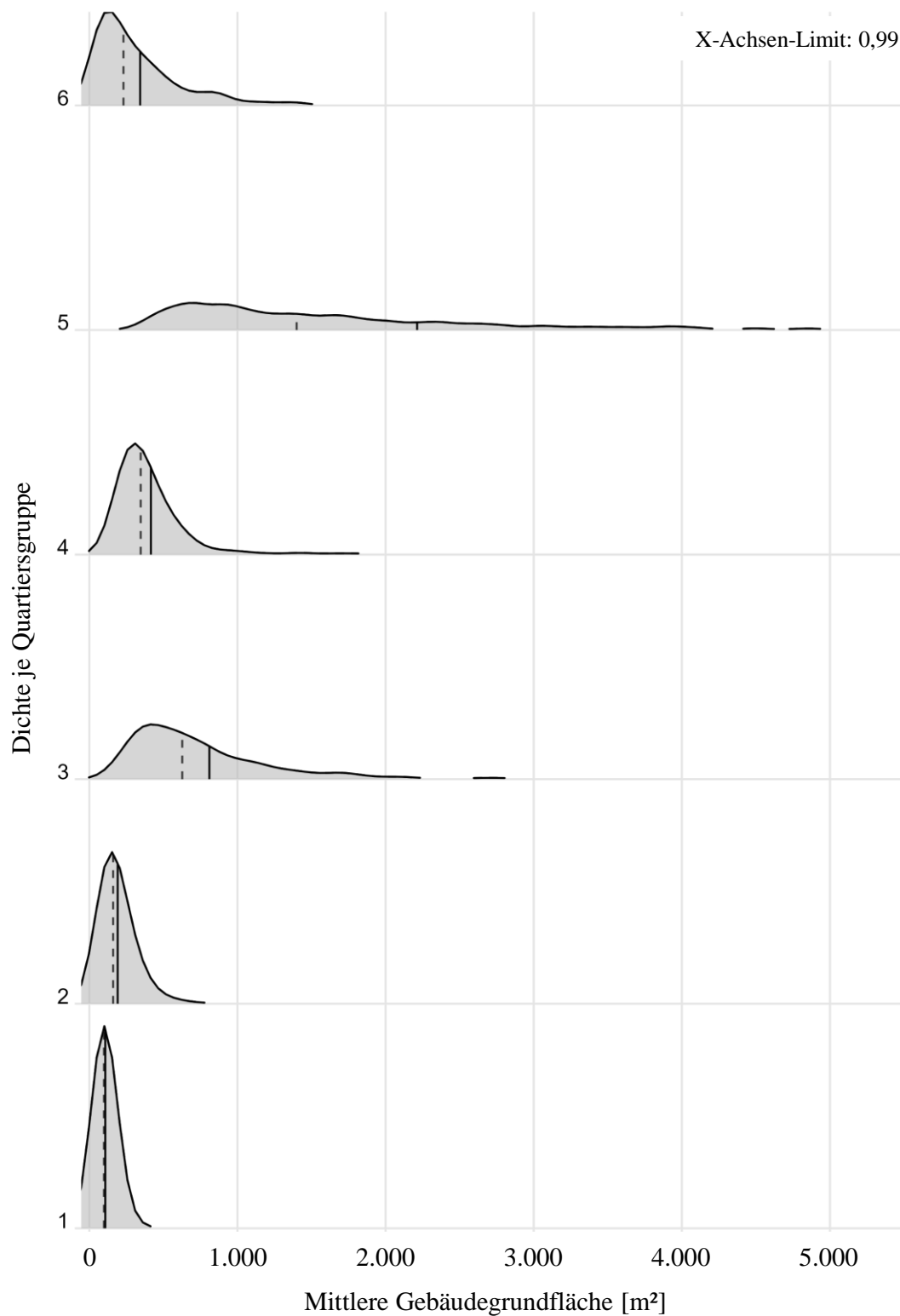


Abbildung Anhang D- 5: Dichteverlauf der Merkmalsausprägung „Gebäudegrundfläche“ auf Ebene der Quartiere für das Anwendungsbeispiel Berlin (Mittelwert = schwarze Linie, Median = gestrichelte Linie)

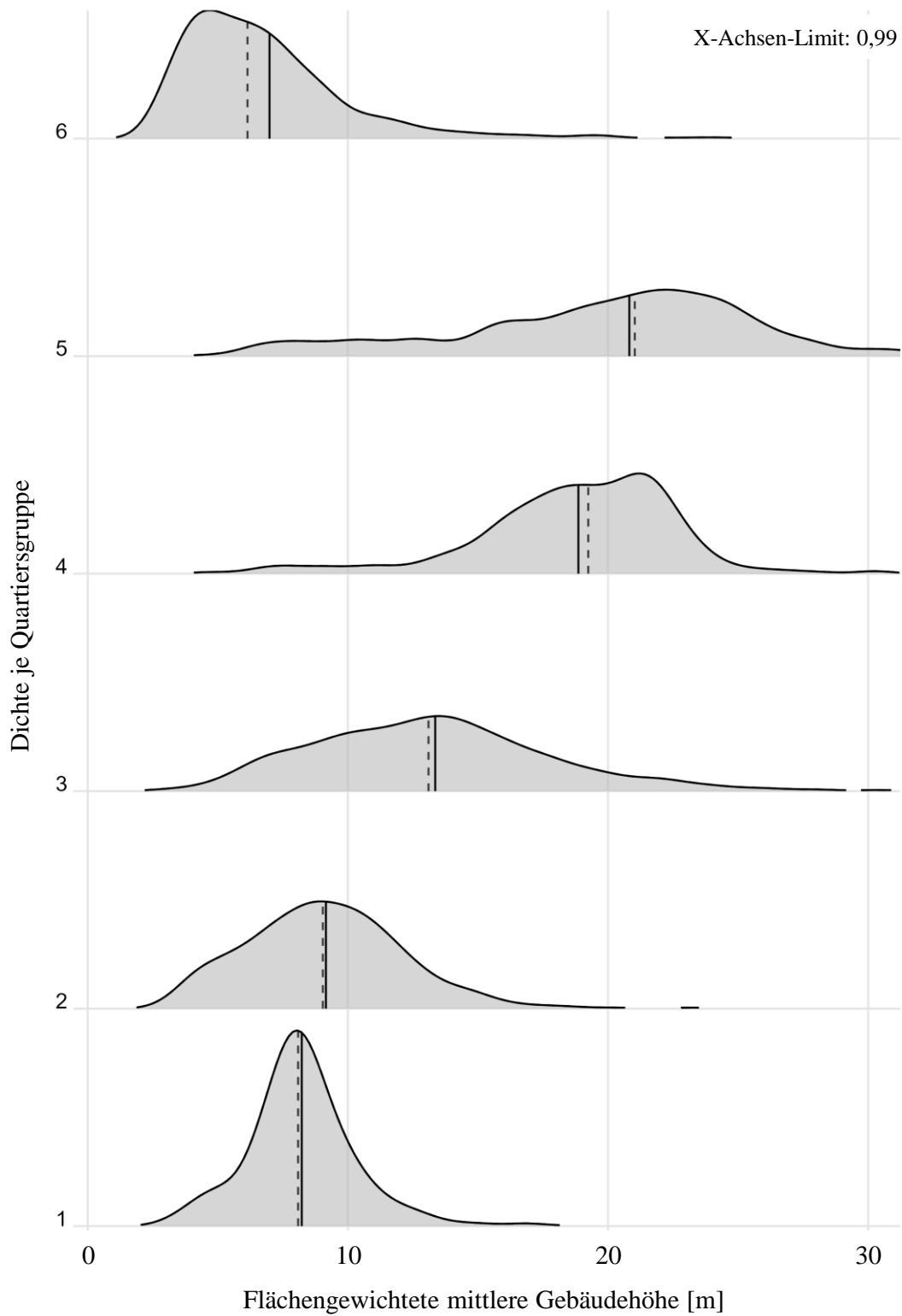


Abbildung Anhang D- 6: Dichteverlauf der Merkmalsausprägung „flächengewichtete mittlere Gebäudehöhe“ auf Ebene der Quartiere für das Anwendungsbeispiel Berlin (Mittelwert = schwarze Linie, Median = gestrichelte Linie)

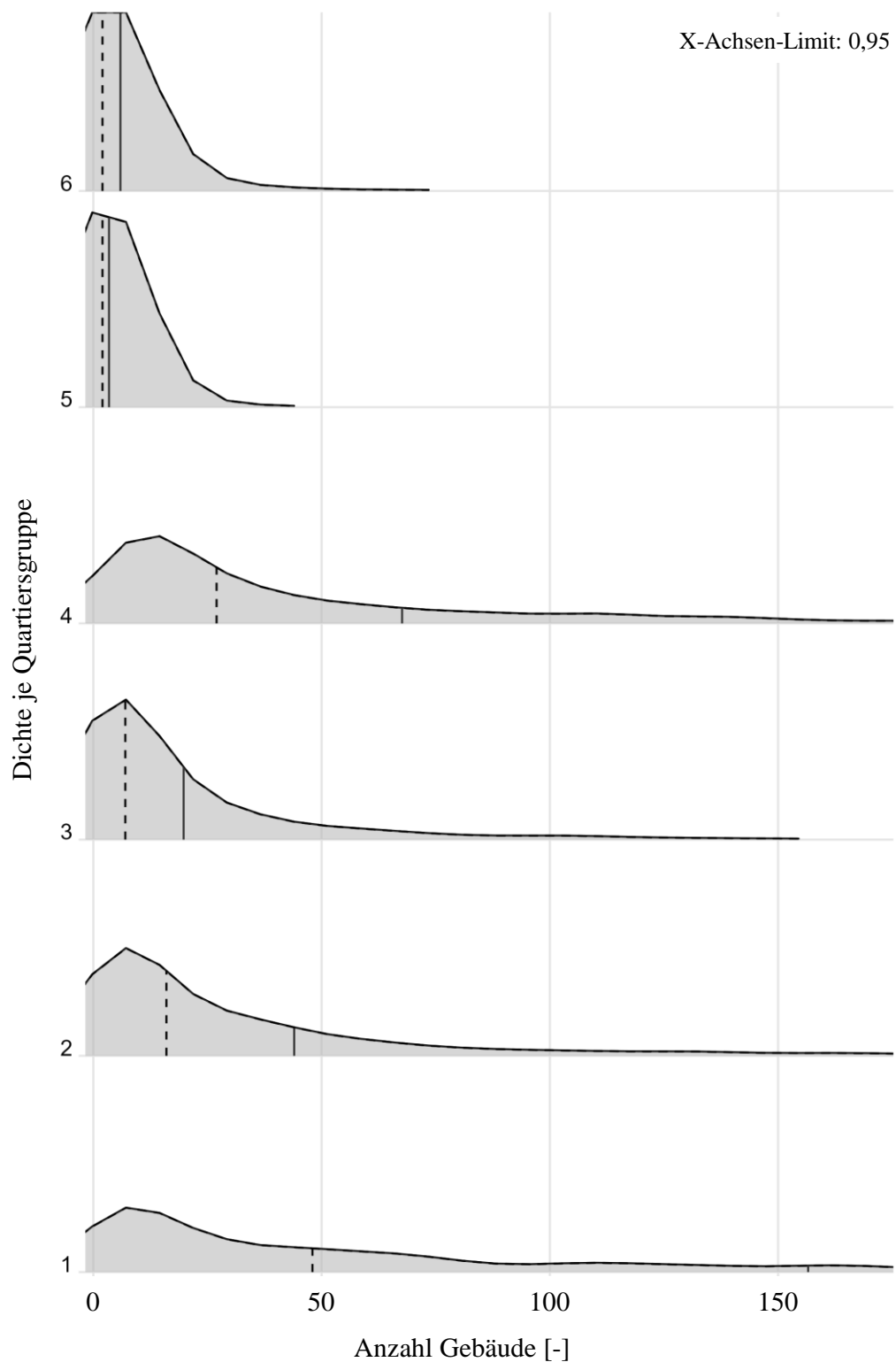


Abbildung Anhang D- 7: Dichteverlauf der Merkmalsausprägung „Anzahl Gebäude“ auf Ebene der Quartiere für das Anwendungsbeispiel Berlin (Mittelwert = schwarze Linie, Median = gestrichelte Linie)

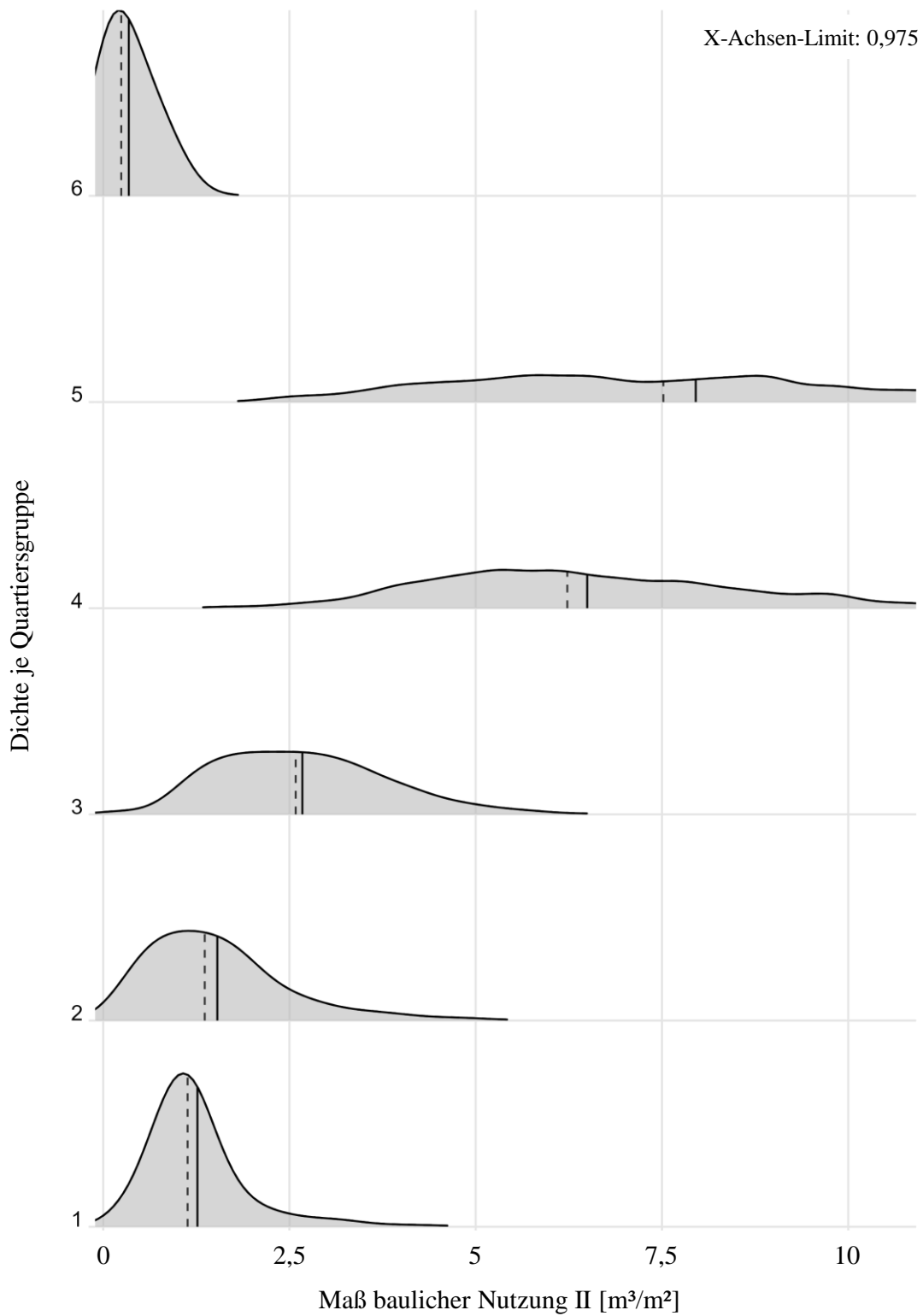


Abbildung Anhang D-8: Dichteverlauf der Merkmalsausprägung „Maß baulicher Nutzung II“ auf Ebene der Quartiere für das Anwendungsbeispiel Berlin (Mittelwert = schwarze Linie, Median = gestrichelte Linie)

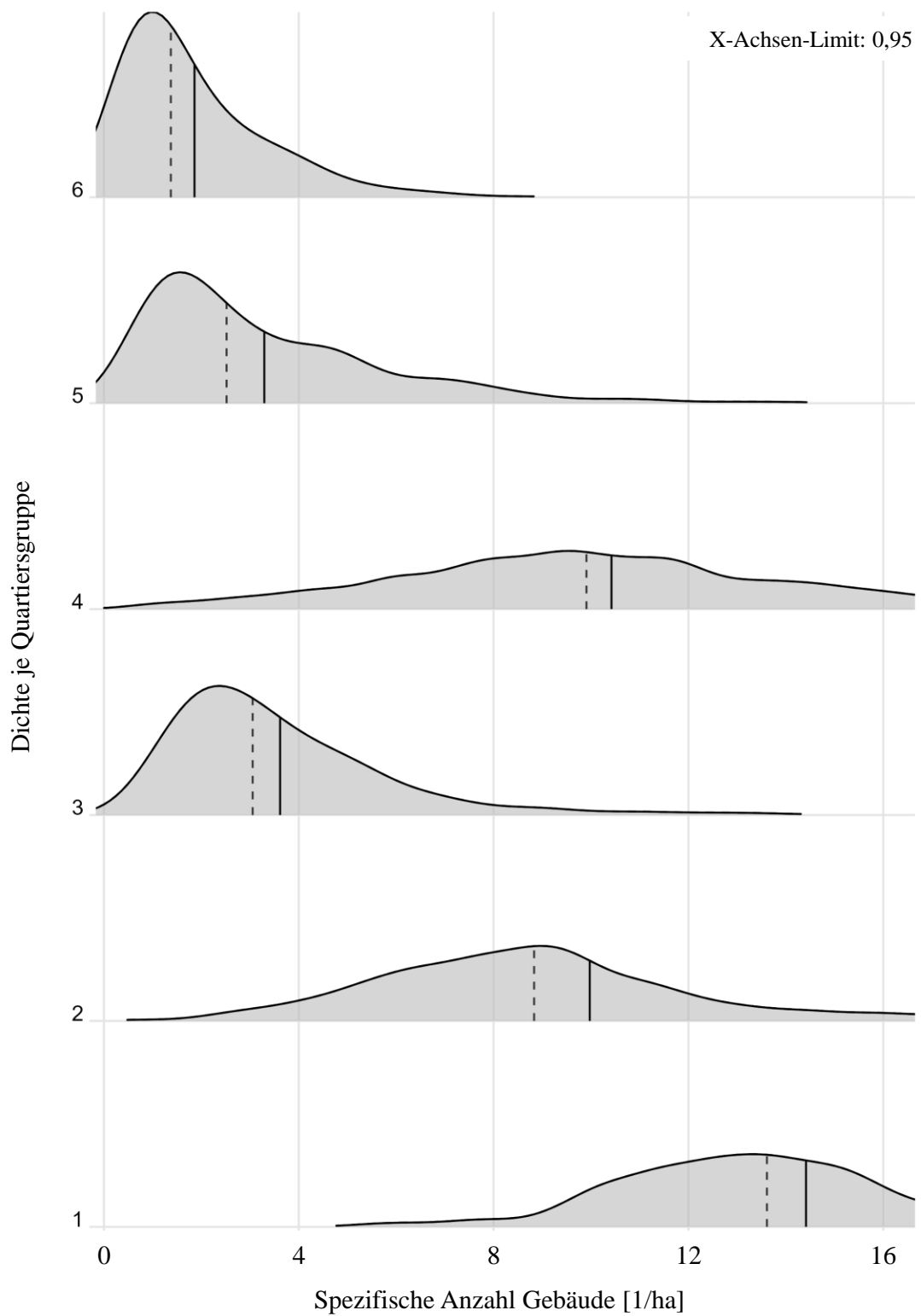


Abbildung Anhang D-9: Dichteverlauf der Merkmalsausprägung „Spezifische Anzahl Gebäude“ auf Ebene der Quartiere für das Anwendungsbeispiel Berlin (Mittelwert = schwarze Linie, Median = gestrichelte Linie)

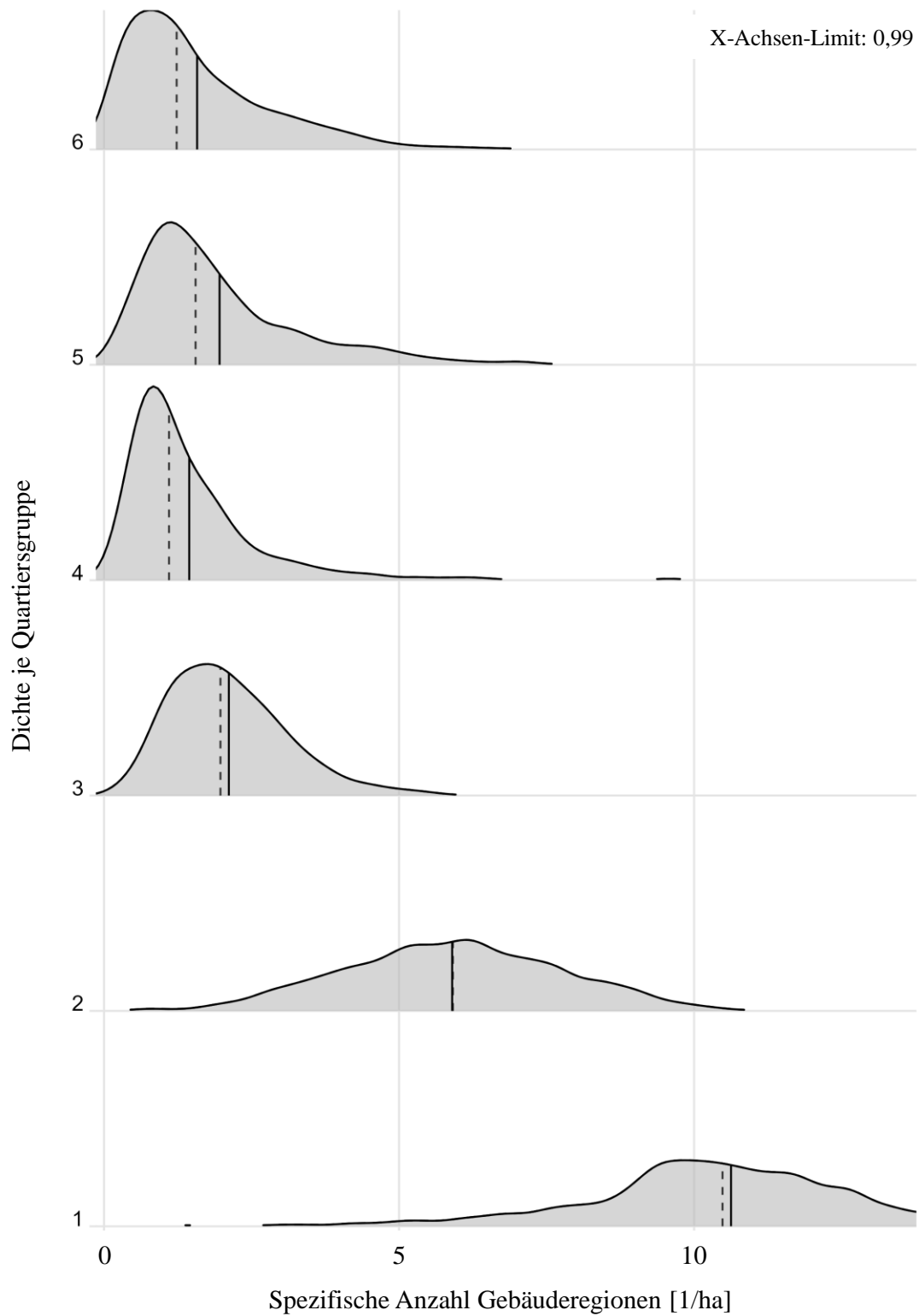


Abbildung Anhang D-10: Dichteverlauf der Merkmalsausprägung „Spezifische Anzahl Gebäuderegionen“ auf Ebene der Quartiere für das Anwendungsbeispiel Berlin (Mittelwert = schwarze Linie, Median = gestrichelte Linie)

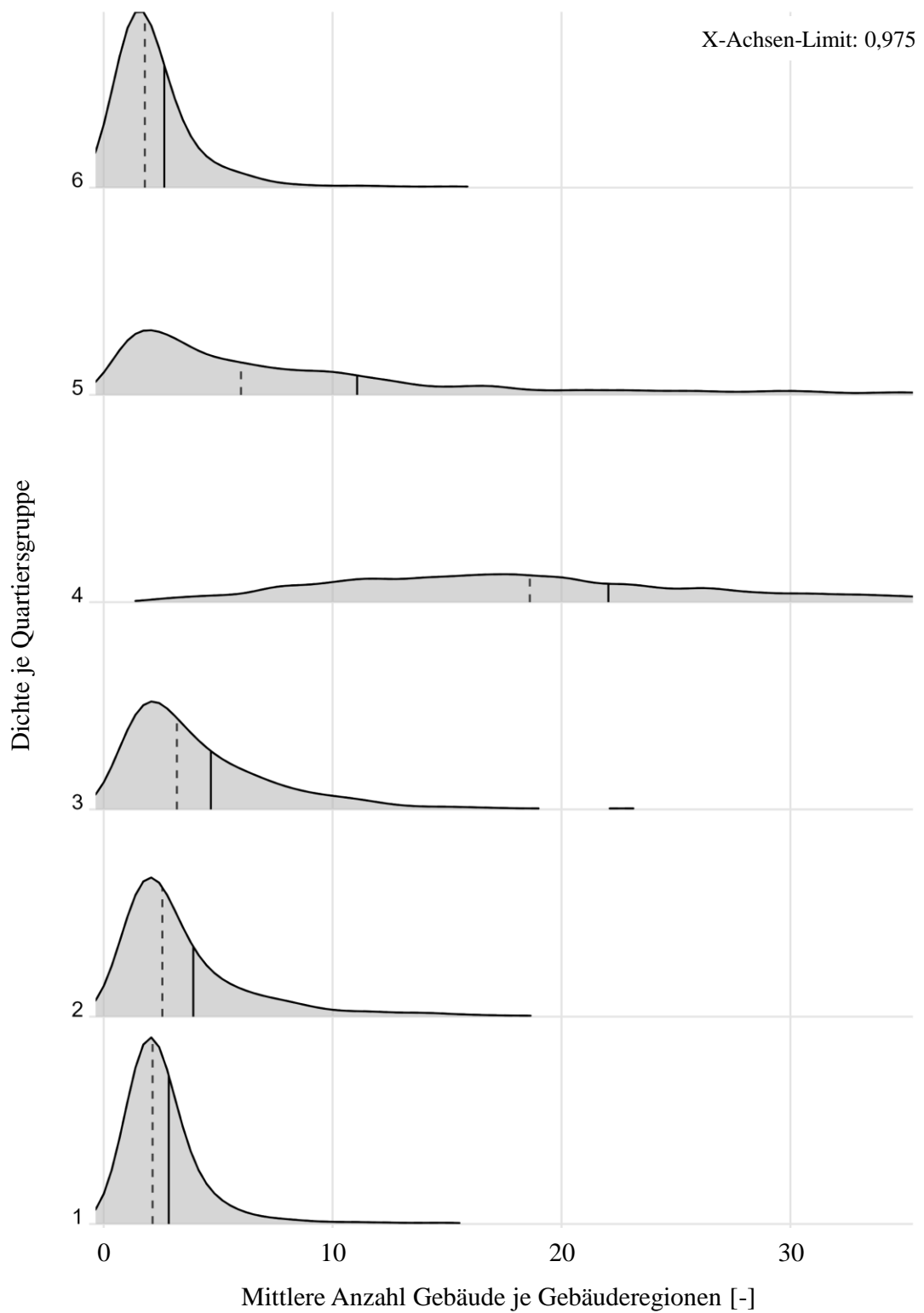


Abbildung Anhang D-11: Dichteverlauf der Merkmalsausprägung „Mittlere Anzahl Gebäude je Gebäuderegion“ auf Ebene der Quartiere für das Anwendungsbeispiel Berlin (Mittelwert = schwarze Linie, Median = gestrichelte Linie)

Dissertationsreihe

Univ.-Prof. Dr.-Ing. Carl-Alexander Graubner

Institut für Massivbau
Technische Universität Darmstadt

- Heft 1: **Stefan Kempf (2001)**
Technische und wirtschaftliche Bewertung der Mindestbewehrungsregeln für Stahlbetonbauteile
- Heft 2: **Katja Reiche (2001)**
Nachhaltigkeitsanalyse demontagegerechter Baukonstruktionen – Entwicklung eines Analysemodells für den Entwurf von Gebäuden
- Heft 3: **Michael Six (2001)**
Sicherheitskonzept für nichtlineare Traglastverfahren im Betonbau
- Heft 4: **Eric Simon (2002)**
Schubtragverhalten von Mauerwerk aus großformatigen Steinen
- Heft 5: **Holger Schmidt (2003)**
Versagenswahrscheinlichkeit unbewehrter Wand-Decken-Verbindungen bei Gasexplosionen im Fertigteilbau
- Heft 6: **Andreas Bachmann (2003)**
Ein wirklichkeitsnaher Ansatz der böenerregten Windlasten auf Hochhäuser in Frankfurt/Main
- Heft 7: **Duy Tien Nguyen (2004)**
Rotationskapazität von biegebeanspruchten Stahlbetonbauteilen mit Schubrissebildung
- Heft 8: **Gert Wolfgang Riegel (2004)**
Ein softwaregestütztes Berechnungsverfahren zur Prognose und Beurteilung der Nutzungskosten von Bürogebäuden
- Heft 9: **Christian Glock (2004)**
Traglast unbewehrter Beton- und Mauerwerkswände – Nichtlineares Berechnungsmodell und konsistentes Bemessungskonzept für schlanke Wände unter Druckbeanspruchung
- Heft 10: **Kati Herzog (2005)**
Lebenszykluskosten von Baukonstruktionen – Entwicklung eines Modells und einer Softwarekomponente zur ökonomischen Analyse und Nachhaltigkeitsbeurteilung von Gebäuden

- Heft 11: **Andreas Garg (2006)**
*Spannungszustände in Fahrbahnplatten weit gespannter Stahlverbundbrücken
Empfehlungen für die Herstellung der Ortbeton-Fahrbahnplatte von Talbrücken
mit der Schalwagenmethode*
- Heft 12: **Tilo Proske (2007)**
*Frischbetondruck bei Verwendung von Selbstverdichtendem Beton – Ein wirk-
lichkeitsnahes Modell zur Bestimmung der Einwirkungen auf Schalung und Rüs-
tung*
- Heft 13: **Simon Glowienka (2007)**
*Zuverlässigkeit von großformatigem Mauerwerk – Probabilistische Analyse von
großformatigem Mauerwerk aus Kalksandstein und Porenbeton mit Dünnbett-
vermörtelung*
- Heft 14: **Alexander Renner (2007)**
*Energie- und Ökoeffizienz von Wohngebäuden – Entwicklung eines Verfahrens
zur lebenszyklusorientierten Bewertung der Umweltwirkungen unter besonderer
Berücksichtigung der Nutzungsphase*
- Heft 15: **Guido Hausmann (2007)**
*Verformungsvorhersage vorgespannter Flachdecken unter Berücksichtigung der
stochastischen Eigenschaften*
- Heft 16: **Stefan Daus (2007)**
*Zuverlässigkeit des Klebeverbundes von nachträglich verstärkten Betonbautei-
len – Sicherheitskonzept für den Verbundnachweis von oberflächlich geklebter Be-
wehrung*
- Heft 17: **Thomas Kranzler (2008)**
*Tragfähigkeit überwiegend horizontal beanspruchter Aussteifungsscheiben aus
unbewehrtem Mauerwerk*
- Heft 18: **Lars Richter (2009)**
*Tragfähigkeit nichttragender Wände aus Mauerwerk – Ein nichtlineares Berech-
nungsmodell und Bemessungsverfahren für biegebeanspruchte Innen- und Au-
ßenwände*
- Heft 19: **Markus Spengler (2010)**
*Dynamik von Eisenbahnbrücken unter Hochgeschwindigkeitsverkehr –
Entwicklung eines Antwortspektrums zur Erfassung der dynamischen Trag-
werksreaktion*
- Heft 20: **Linh Ngoc Tran (2011)**
*Berechnungsmodell zur vereinfachten Abschätzung des Ermüdungsverhaltens
von Federplatten bei Fertigträgerbrücken*

- Heft 21: **Carmen Schneider (2011)**
Steuerung der Nachhaltigkeit im Planungs- und Realisierungsprozess von Büro- und Verwaltungsgebäuden – Entwicklung eines Instrumentes zur Vorbewertung und Optimierung der Nachhaltigkeitsqualität
- Heft 22: **Frank Ritter (2011)**
Lebensdauer von Bauteilen und Bauelementen – Modellierung und praxisnahe Prognose
- Heft 23: **Benjamin von Wolf-Zdekauer (2011)**
Energieeffizienz von Anlagensystemen zur Gebäudekühlung – Ein nutzenbezogener Bewertungsansatz
- Heft 24: **Eric Brehm (2011)**
Reliability of Unreinforced Masonry Bracing Walls – Probabilistic Approach and Optimized Target Values
- Heft 25: **Carolin Roth (2011)**
Lebenszyklusanalyse von Baukonstruktionen unter Nachhaltigkeitsgesichtspunkten – Ein Beitrag zur Beurteilung der Nachhaltigkeit von Gebäuden bei ungewissem Lebensweg
- Heft 26: **Kay-Uwe Thorn (2013)**
Neuentwicklung eines Ladungsrückhaltesystems aus textilbewehrten Betonfertigteilen
- Heft 27: **Andreas Greck (2013)**
Straßenbrücken mit einteiligem Verbundquerschnitt – Optimierung des Betonierablaufs der Fahrbahnplatte und Entwicklung eines Vorschlags zur vereinfachten globalen Systemberechnung beim Einsatz mehrerer Schalwagen
- Heft 28: **Martin Heimann (2013)**
Tragwerkszuverlässigkeit hochbeanspruchter Druckglieder aus ultrahochfestem Beton
- Heft 29: **Torsten Mielecke (2013)**
Nachhaltigkeitsbewertung von Einfamilienhäusern – Ein Beitrag zur Beurteilung und Optimierung der Nachhaltigkeitsqualität von kleinen Wohngebäuden
- Heft 30: **Sebastian Pohl (2013)**
Nachhaltigkeit im Gebäudebetrieb – Ein Vorschlag zur Diversifizierung des bau- und immobilienwirtschaftlichen Zertifizierungsregimes in Deutschland
- Heft 31: **Stefan Hainer (2015)**
Karbonatisierungsverhalten von Betonen unter Einbeziehung klinkerreduzierter Zusammensetzungen – Ein Modell zur Abschätzung des Karbonatisierungsfortschritts

- Heft 32: **Achim Knauff (2016)**
Heizwärmeerzeugung mit Solarthermie, Geothermie und Photovoltaik für Büro- und Verwaltungsgebäude
- Heft 33: **Ulf Grziwa (2017)**
Zuverlässigkeit schlanker UHPC-Druckglieder mit räumlich streuenden Materialeigenschaften
- Heft 34: **Björn Freund (2017)**
Frischbetondruck lotrechter, geneigter und gekrümmter Betonbauteile bei Verwendung von Betonen mit hoher Fließfähigkeit
- Heft 35: **Gökhan Uysal (2017)**
Kostenrisiken von Industriebauten mit flexiblen Nutzungsszenarien – Ein Berechnungsmodell für die Investitionsentscheidung unter Verwendung probabilistischer Methoden
- Heft 36: **Moien Rezvani (2017)**
Shrinkage model for concrete made of limestone-rich cements – An approach from cement paste to concrete
- Heft 37: **Claudia Weißmann (2017)**
Effizienter Einsatz erneuerbarer Energieträger in vernetzten Wohnquartieren – Entwicklung eines simulationsbasierten Verfahrens zur energetischen, ökologischen und ökonomischen Bewertung
- Heft 38: **Markus Blatt (2017)**
Beitrag zum Trag- und Verformungsverhalten von Stahlbetondeckenknoten
- Heft 39: **Michael Schmitt (2018)**
Tragfähigkeit ausfachender Mauerwerkswände unter Berücksichtigung der verformungsbasierten Membranwirkung
- Heft 40: **Valentin Förster (2018)**
Tragfähigkeit unbewehrter Beton- und Mauerwerksdruckglieder bei zweiachsig exzentrischer Beanspruchung
- Heft 41: **Ngoc Linh Tran (2018)**
Shear strength of slender reinforced concrete members without shear reinforcement
- Heft 42: **Jaroslav Kohoutek (2018)**
Zuverlässigkeit integraler Straßenbrücken in Massivbauweise
- Heft 43: **Jochen Zeier (2019)**
Thermisch getrennte Stützen-Decken-Anschlüsse im Stahlbetonbau – Ein Modell zum Tragverhalten bei großen Fugendicken
- Heft 44: **Katharina Fritz (2019)**
Qualifizierte Beurteilung der sektoralen Ressourceninanspruchnahme im deutschen Bauwesen

- Heft 45: **Patrick Wörner (2020)**
Einfluss des Nutzerverhaltens auf den Stromverbrauch in Wohngebäuden – Entwicklung eines komplexen Simulationsmodells für energetische Analysen
- Heft 46: **Martina Lohmeier (2020)**
Nachhaltigkeitsvergleich von Projektvarianten für Außerortsstraßen – Ein Vorschlag für den Abwägungs- und Entscheidungsprozess in der Vorplanung
- Heft 47: **Larissa Krieger (2020)**
Tragverhalten textilbewehrter Fertigteileplatten aus hochfestem Carbonbeton – Einfluss der Ausbreitung konzentrierter Einzellasten auf die Biegetragfähigkeit
- Heft 48: **Sarah Steiner (2020)**
Carbonation of concrete made of limestone-rich cement – CO₂ diffusivity and alteration of the hydrate phases
- Heft 49: **Redouan El Ghadioui (2020)**
Bemessung carbonbewehrter Betonbauteile unter besonderer Betrachtung des Gebrauchszustandes – Entwicklung eines Modells zur wirklichkeitsnahen Verformungsberechnung
- Heft 50: **Sebastian Hofmann (2021)**
Rissentwicklung in Betonbauteilen mit Basaltfaserkunststoffbewehrung bei wirklichkeitsnaher Betrachtung des Verbundverhaltens – Ein Modell zur Berechnung der Rissbreite im Grenzzustand der Gebrauchstauglichkeit
- Heft 51: **Dominik Müller (2022)**
Probabilistic Assessment of Existing Masonry Structures – The Influence of Spatially Variable Material Properties and a Bayesian Method for Determining Structure-Specific Partial Factors
- Heft 52: **Christian Herget (2022)**
Kalksteinmehl als Betonzusatzstoff für umweltfreundliche Betone – Bestimmung eines Anrechenbarkeitsfaktors zur Sicherstellung gleichwertiger Festbetoneigenschaften
- Heft 53: **Maximilian Brinkmann (2023)**
Tragfähigkeit druckbeanspruchten Lehm-mauerwerks unter Berücksichtigung nichtlinearer Feuchteprofile
- Heft 54: **Dominik Hiesch (2023)**
Bemessung von Betonfertigteilen mit vorgespannter Bewehrung aus Faserverbundkunststoff – Entwicklung eines Bemessungskonzeptes unter Berücksichtigung wirklichkeitsnaher Bauteilverformungen
- Heft 55: **André Müller (2023)**
Methodik zur datenbasierten Typisierung von Quartieren anhand baulicher Strukturen – Geometrisch-topologische Grundlagen für die energetische Bewertung des nationalen Quartiersbestandes auf Basis von Geoinformationsdaten

André Müller

Methodik zur datenbasierten Typisierung von Quartieren anhand baulicher Strukturen

Die deutsche Gesellschaft befindet sich in einem Transformationsprozess, um der Klimakrise zu begegnen. Vor dem Hintergrund der notwendigen Dekarbonisierung der Wärmeversorgung von Gebäuden haben sich Konzepte etabliert, welche für Quartiere, d.h. mehrere Gebäude in einem räumlichen Zusammenhang, die Reduktion von Treibhausgasemissionen durch eine Erhöhung der Gebäude- und Anlageneffizienz sowie durch die Einbindung erneuerbarer Energieträger erreichen. Während auf lokaler Handlungsebene eine Vielzahl an Projekten umgesetzt wird, fehlt die empirische Basis, um die Art und Anzahl von Quartieren in Deutschland zu erfassen und die Wirkung von Quartierskonzepten in Hoch- und Szenarienrechnungen zum Klimaschutz zu integrieren.

Die vorliegende Arbeit liefert eine Methodik zur Nutzung von Geobasisdaten für eine Analyse der Bebauungsstruktur und eine Einteilung dieser in geometrisch-topologisch ähnliche Gebiete mittels Clusteranalyse. Eine Vereinigung benachbarter Gebiete ähnlicher Bebauung führt im Anschluss zur Bildung von Quartieren, was schließlich eine statistische Analyse des Quartiersbestandes erlaubt. Daraus können Kenngrößen für die Beschreibung von Typvertretern für Quartiersgruppen, sogenannte Typquartiere, festgelegt werden. Die beispielhafte Anwendung der Methodik am Datensatz Berlins zeigt eine praxisgerechte Quartiersbildung sowie die Verwendbarkeit der Ergebnisse der Quartierstypisierung für vielfältige Forschungs- und Planungszwecke. Insbesondere eine Verwendung des nationalen Geodatenbestandes zur Quartierstypisierung eröffnet die Möglichkeit weiterführender Szenarioanalysen im Kontext nationaler Klimaschutzstrategien.