

TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Energiewirtschaft und Anwendungstechnik

Wärmeverbrauchsanalyse auf Basis einer raumbezogenen Zusammenführung von Gebäudedaten

Uwe Peter Böhme

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik
der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs (Dr.-Ing.)

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr.-Ing. Josef Kindersberger

Prüfer der Dissertation: 1. Univ.-Prof. Dr.-Ing. Ulrich Wagner
2. Univ.-Prof. Dr.-Ing. Klaus Diepold
3. Univ.-Prof. Dr. rer. nat. Thomas Hamacher

Die Dissertation wurde am 19.06.2013 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 10.12.2013 angenommen.

Zusammenfassung

In der vorliegenden Arbeit werden Datenmodelle und Algorithmen entworfen, um raumbezogene Informationen über den Gebäudebestand einer Stadt zu verarbeiten. Unter Verwendung einer Geodatenbank und von Text Mining-Methoden lassen sich erstmals Gebäudedaten aus vielen, nicht flächendeckenden Quellen unterschiedlicher geografischer Skalen und mit Angaben in natürlicher Sprache zusammenführen. Die Herkunft von Einzelangaben bleibt nachvollziehbar, Datenkonflikte werden anhand der Qualität der Quellen automatisiert gelöst. Darauf aufbauend werden Gebäudeparameter berechnet, deren Einfluss auf den Energieverbrauch für Raumheizung analysiert und Schätzfehler bei der Berechnung von Wärmekatastern quantifiziert. Mit diesem Ansatz lassen sich kosteneffizient Grundlagen für Wärmesystemanalysen schaffen, die auf städtischen Daten basieren.

Abstract

This thesis is about the development of data models and algorithms to handle spatially referenced information regarding the building stock of a municipality. Using a spatial database and text-mining-methods, building information from multiple sources is conflated while the source of the information remains transparent for every parameter and building. The sources possess various spatial resolutions and elements with unclassified natural language. Conflicts arising of contradictory data are automatically resolved by comparing the source quality. Based on this conflated information, building parameters are calculated and their influence on energy consumption for room heating is analyzed. Errors from the estimation of heat consumption values are quantified. This approach allows creating a basis for heat system analyses based on municipal data.

Inhaltsverzeichnis

Inhaltsverzeichnis	V
1. Einleitung	1
2. Methoden	5
2.1. Grundlagen und Begriffe relationaler Datenbanksysteme	5
2.2. Geoinformationssystem und Geodatenbank	7
2.3. Grundlagen und Begriffe des <i>Text Mining</i>	9
2.4. Wärmebedarfsbilanzierung und Klimakorrektur	10
2.4.1. Energieform	11
2.4.2. Klimatische Anpassung	11
2.4.3. Heizenergiebilanzierung	13
2.5. Regressionsanalyse	15
3. Konzept zur Verarbeitung der Gebäudeparameter	19
3.1. Generalisierte Struktur der Datenquellen und Datenaufbereitung	21
3.1.1. Attributtypen der Inputdaten	22
3.1.2. Güteindikator für Datenquellen	23
3.1.3. Mustererkennung in Textfeldern	24
3.2. Modellierung und Verwendung von Raumbezügen	27
3.2.1. Aufbau eines Datenmodells der Raumbezüge	27
3.2.2. Verwendung der Raumbezüge zum Import von Primärparametern	31
3.2.3. Verarbeitung der Adresse als Raumbezug externer Daten	33
3.3. Primärparameterverwaltung	37
3.3.1. Datenmodell der Primärparameterverwaltung	37
3.3.2. Zusammenführung der Primärparameter	40
3.3.3. Beispiele für die Verwendung der Primärparameterverwaltung . .	44
3.4. Gebäudeparametermodell	46
4. Quellen für Gebäudedaten in Oldenburg	49
4.1. Gebäudedatenquellen mit Gebäudebezug	50
4.1.1. Archiv (Bauordnung)	50
4.1.2. Bauanträge (Bauordnung)	51
4.1.3. Digitalisierte Gebäudekarteikarten (Bauordnung)	52
4.1.4. Automatisierte Liegenschaftskarte (ALK)	53
4.1.5. Laserscan-Überfliegung	53
4.1.6. Manuelle Luftbilddauswertung	54
4.1.7. Kaufpreissammlung des Gutachterausschusses für Grundstückswerte	55
4.1.8. Oldenburger Häuserbuch	56

4.1.9. Vertragszahlen des Stromnetzbetreibers	57
4.2. Gebäudedatenquellen mit Gebietsbezug	58
4.2.1. Rechtsverbindliche Bebauungspläne (B-Pläne)	58
4.2.2. Öffentliche Brandkasse	58
4.2.3. Baufertigstellungsstatistik des LSKN	59
4.2.4. Wohnraumzählung der Volkszählung 1987	59
5. Berechnung und Analyse von Gebäudeparametern in Oldenburg	61
5.1. Gebäudebaujahr und Baualtersklasse	62
5.1.1. Datenquellen	62
5.1.2. Klassifizierungen des Baujahres	64
5.1.3. Bestimmung der Datenqualität aus dem Vergleich der Quellen	65
5.1.4. Schätzung nicht bekannter Gebäudebaujahre	67
5.1.5. Analyse und Kontrolle der Gebäudebaujahresverteilungen	68
5.2. Gebäudeklasse	72
5.2.1. Bestimmung der Gebäudeklasse	72
5.2.2. Analyse der Gebäudeklassenverteilung	75
5.3. Gebäudetyp	76
5.4. Gebäudehüllflächen und Gebäudevolumen	77
5.4.1. Datengrundlage und Vervollständigung	77
5.4.2. Berechnung der Gebäudehüllflächen und des Volumens	78
5.4.3. Analyse und Literaturvergleich	81
5.5. Geschosshöhe und Grundflächen	83
5.5.1. Quellen bzw. Berechnung der Geschosshöhe aus der Gebäudehöhe	83
5.5.2. Berechnung und Umrechnung von Grundflächen	84
5.5.3. Auswertung des Verhältnisses zwischen Wohnfläche und Volumen	86
5.6. Gebäudenutzung	87
5.6.1. Transformation der Gebäudenutzungsdaten in Oldenburg	88
5.6.2. Berechnung der Gebäudezonengrößen für verschiedene Nutzungen	90
5.6.3. Analyse der Nutzungsarten und zugeordneten Nutzflächen im Stadtgebiet Oldenburg	91
6. Analyse der Wärmeverbräuche von Ein- und Zweifamilienhäusern in Oldenburg	95
6.1. Gebäudeauswahl und Berechnung des klimabereinigten Heizwärmeverbrauchs	95
6.1.1. Gebäudeauswahl	95
6.1.2. Gebäudedaten und anonymisierte Gasliefermengen	96
6.1.3. Berechnung klimakorrigierter Endenergieverbräuche für Raumwärme	98
6.1.4. Berechnung spezifischer Wärmeverbräuche	99
6.2. Mittlere Endenergieverbräuche und Literaturvergleich	100
6.2.1. Differenzierung nach Baualtersklasse und Gebäudebaujahr	100
6.2.2. Vergleich mit Literaturwerten	103
6.2.3. Differenzierung nach Baualtersklasse, Volumen und A/V-Verhältnis	104
6.3. Verteilung und Multiple Regression gruppierter Heizwärmeverbräuche	106
6.3.1. Kumulierte Verteilung des spezifischen Wärmeverbrauchs	107
6.3.2. Multiple Regressionsanalyse des Heizwärmeverbrauchs	109

6.4. Unsicherheit bei der Wärmeverbrauchsschätzung	111
6.4.1. Konfidenzintervall des statistischen Schätzfehlers	114
6.4.2. Konfidenzintervall des statistischen und systematischen Schätzfehlers	116
6.4.3. Übertragbarkeit der Ergebnisse auf Mehrfamilienhäuser	117
6.5. Vergleich von Wärmeverbrauch und Wärmebedarf	118
6.5.1. Wärmebedarfsberechnung und Literaturvergleich	118
6.5.2. Vergleich von Wärmeverbrauch und -bedarf	118
6.5.3. Effekt von Energieeffizienzmaßnahmen	121
7. Zusammenfassung und Ausblick	123
A. Anhang	127
A.1. Räumliche Darstellung der ermittelten Gebäudebaujahre in Oldenburg . .	127
A.2. Transformation der Stromvertragsklassen auf die GHD-HH-Klassen	136
Literaturverzeichnis	XV
Akronyme	XVII
Abbildungsverzeichnis	XX
Tabellenverzeichnis	XXI
Danksagung	XXIII

1. Einleitung

Es ändert sich das System nach den lokalen Verhältnissen und diese müssen deshalb in jedem Fall genau geprüft werden bevor man sich für die Wahl des Systems entscheidet.

Oskar von Miller (1891)

Energiewirtschaftliches Umfeld des Wärmeverbrauchs

Dem Wärmesektor kommt bei der Umstrukturierung und Dekarbonisierung des Energiesystems eine „herausragende Bedeutung“ [58, S.61] zu, da 58 % des Endenergieverbrauchs in Deutschland auf den Wärmeverbrauch zurückzuführen sind. Die Hälfte dieses Anteils wird für Raumwärme aufgewendet, die vor allem im Sektor Haushalte, gefolgt vom Sektor Gewerbe, Handel und Dienstleistung (GHD) anfällt [58].

Die Umstrukturierung wird aus zwei Gründen angestrebt. Die bei der Verbrennung fossiler Energieträger entstehenden CO₂-Emissionen in Deutschland sollen als weltweites Leuchtturmprojekt klimagerechteren Wirtschaftens gesenkt werden. Zugleich belasten die steigenden Preise der fossilen Energieträger die wirtschaftliche Entwicklung und das Lebensniveau der Bevölkerung [58, S.52]. Strategien sind Effizienzsteigerung, d. h. die Reduktion des Wärmeverbrauchs durch Gebäudesanierung und die Optimierung von Wärmeerzeugungsanlagen, Abwärmennutzung aus Kraft-Wärme-Kopplung (KWK) und der Einsatz von erneuerbaren Energien. Insbesondere durch die letzten beiden Punkte erfolgt eine immer engere Verzahnung des Wärme- und Stromsektors.

Durch die Komplexität der Umstrukturierung wird eine systemische Analyse, Planung und Optimierung des Wärmesektors notwendig. Diese adressiert vor allem technische und wirtschaftliche Aspekte.

Hierzu gehört der Vergleich zentraler Wärmeinfrastrukturen mit dezentralen Einzelanlagen. Bei einer Wärmeversorgung auf Basis erneuerbarer Energien ist sowohl zentral bereitgestellte Fernwärme als auch der Transport von Biomasse für dezentrale Feuerungsanlagen mit Transportverlusten behaftet [62]. Dies macht es notwendig, die räumliche Lage von Verbrauchern und Erzeugern bzw. Ressourcen zu berücksichtigen. Gewünscht ist eine möglichst feine räumliche Auflösung, z. B. die Unterscheidung einzelner Gebäude.

Die Treiber und Potentiale des Gesamtsystems müssen in der Modellierung korrekt abgebildet werden. Die aktuelle Debatte über die Wirtschaftlichkeit von energetischen Sanierungen im Gebäudebestand zeigt, dass der Wärmebedarf - eine theoretische Größe zur energetischen Beurteilung eines Gebäudes - in vielen Fällen mit dem tatsächlich messbarem Wärmeverbrauch verwechselt wird [75]. Letzterer beinhaltet den wichtigen Faktor

Nutzerverhalten und ist der eigentliche Treiber bei der Bestimmung der Energienachfrage und der Potentiale zur Energieeinsparung. Ein Effekt dieser Verwechslung sind Beschlüsse zu unrealistischen Emissions-Minderungsziele auf kommunaler Ebene.

Grundlage der geforderten Energiesystemmodellierung ist somit die Verfügbarkeit von Daten, um die lokalen Modell- und Randbedingungen korrekt abzubilden. Hierzu gehören raumbezogene Gebäudedaten für die Analyse und Ermittlung eines flächendeckenden Wärmebedarfs und vor allem Wärmeverbrauchs.

Aus einer Vielzahl von Gründen bieten sich kommunale Untersuchungsgebiete für hoch aufgelöste, raumbezogene Analysen des Wärmesektors an. Dies zeigen 893 Klimaschutzkonzepte bis Ende des Jahres 2011 [13]. Entsprechende Grundlagen sind durch den Leitspruch der Agenda 21 „*think global, act local*“ gelegt [6]. Beispielsweise existieren für die Stadt Oldenburg entsprechende Analysen seit dem Jahr 1992 [63, 59, 54].

Zu den Gründen gehört vor allem das Subsidiaritätsprinzip. Ein Siedlungsraum ist die angemessene Größe für die Planung von Wärmeinfrastrukturen und sollte daher einer Planung auf administrativ höheren Ebenen vorgezogen werden. Ein Vorbild bietet der Flächennutzungsplan, der - eingeschränkt durch die Regionalplanung - in kommunaler Hand liegt. Darüber hinaus besteht in einer Kommune die größte Nähe zwischen den Planenden und den Bürgern, was den gesellschaftlichen Diskurs fördert. Ein Großteil der existierenden Daten bezüglich des Gebäudebestandes liegt in der kommunalen Verwaltung in einer administrativen Hand. Sie ist daher der erste Ansprechpartner bei der Datenermittlung.

Eine Schwierigkeit kommunaler Daten ist, dass diese aus Datenbeständen unterschiedlicher Abteilungen zusammengeführt werden müssen, unstrukturierte Elemente wie beschreibenden Text enthalten und zum Teil nicht flächendeckend bzw. vollständig vorliegen. Die Einbindung der strukturreichen und umfangreichen Einzeldaten in Stadtenergieanalysen ist nicht Stand der Technik. Deshalb bauen Analysen auf einer kleinen Datenbasis auf, obwohl unter Umständen bessere Informationen vorhanden wären. Einer Verwendung umfangreicher Daten ist mit hohen Kosten durch die in den meisten Fällen manuell geprägte Datenverarbeitung verbunden.

Stand der Technik

Die Speicherung raumbezogener Gebäudedaten ist durch verschiedene Ansätze möglich. Hierzu gehören relationale Datenbanken [52, 44], die Verwendung von *Layer* - d. h. Karten mit entsprechenden Attributen (z. B. die Automatisierte Liegenschaftskarte) [4] - oder hierarchische XML-Strukturen [60]. Der Begriff Gebäudedatenbank wird nicht immer nur in Verbindung mit einer tatsächlichen Datenbank genutzt, sondern beschreibt generell die Speicherung von Gebäude- bzw. Siedlungsmerkmalen. Die zitierten Quellen haben gemeinsam, dass sie keine Konzepte für eine strukturierte Datenerfassung und -kombination, sondern nur Speicherformate für „fertig verfügbare“ Gebäudedaten bieten.

Die Erfassung von Gebäudedaten ist vor allem durch Fernerkundung [76], Analyse digitaler Informationen von Verwaltungen und Organisationen, Begehungen [37], Befragungen [30, 53] oder durch den Erwerb der Daten von Datenhändlern [43] möglich.

Im Bereich der Fernerkundung und automatischen Strukturerkennung gibt es verschiedene Ansätze einer flächendeckenden Gebäudedatenermittlung. Diese beinhalten die Erkennung der 3D-Struktur der Gebäude durch Laserscan oder Stereoluftbilder und die Analysen von Dachflächen und Siedlungsflächen durch Multispektralaufnahmen [76]. Beispiele für darauf basierende Wärmebedarfsanalysen sind Studien für die Stadt Berlin [45] und für ein Siedlungsgebiet in Stuttgart [29]. Daneben existieren Verfahren für die Erkennung von Siedlungsstrukturen und eine darauf basierende Abschätzung der Einwohnerzahlen bzw. des Wärmebedarfs mittels Kartendaten [56, 55]. Fernerkundungsverfahren analysieren hauptsächlich die Gebäudegeometrie und Oberflächen in unterschiedlicher Detailtiefe. Alle weiteren Informationen werden davon abgeleitet. Die Erfassung historischer Informationen wie beispielsweise des Baujahrs ist automatisiert nicht zuverlässig möglich, da historische Karten bzw. Luftbilder in Schwarz-Weiß vorliegen und die Qualität nicht ausreicht.

Die Verwendung bereits bestehender Datenbestände aus kommunalen und Landesverwaltungen bzw. Organisationen wie den Gutachterausschüssen für Grundstücksbewertung wird in dieser Arbeit verfolgt. Andere Arbeiten mit einem ähnlichen Ansatz sind [52, 37]. Erste Schritte bei der Ermittlung von Gebäudedaten und der Datenverwaltung mittels einer Geodatenbank für das Stadtgebiet Oldenburg sind in Vorbereitung auf diese Arbeit als studentisches Projekt durchgeführt wurden [44]. Viele Klimaschutzkonzepte und Energienutzungspläne stützen ihre Analysen ebenfalls auf kommunale Daten. Jedoch beinhalten die genannten Arbeiten und Studien keinen strukturierten und verallgemeinerten Umgang mit strukturreichen Daten verschiedener Quellen, sondern beschreiben die Datenquellen und eine beispielhafte Verarbeitung, die geprägt ist durch manuelle Datenaufbereitung und -zuordnung.

Die Daten des Zensus 2011 geben die Zahl der Gebäude mit Wohnraum und die Wohnungszahl differenziert nach Baualtersklasse, Gebäudetyp, Eigentumsform und Heizungsart an. Die Angaben sind auf kommunaler Ebene aggregiert und können in Zukunft wichtige Kontrollgrößen liefern [73].

Statistische Untersuchungen des Gebäudebestandes verwenden in den meisten Fällen manuell gepflegte Datenbestände und ergänzend dazu Befragungen und Begehungen [30, 20]. Für kommunale Untersuchungen ist dieser Ansatz ab einer bestimmten Größe der Kommune meist sehr kostenaufwendig [37].

Studien zu den Wärmeverbräuchen für Typgebäude werden überwiegend unter Beteiligung der Wohnungswirtschaft [30], Heizungsablesern wie Metrona oder Techem [77, 69] oder Befragungen [53] durchgeführt. Jede dieser Studien basiert auf einem Stichprobenumfang mit jeweils über 10.000 Gebäuden aus dem nationalen Gebäudebestand in Deutschland. Die Studien geben mittlere Kennwerte der Wärmeverbräuche an. Die Streuung der Verbräuche wird für kleine und große Mehrfamilienhäuser in [69, 33] untersucht. Forschungsarbeiten für die raumbezogene Analyse von kommunalen Wärmeverbräuchen existieren beispielsweise für München [31], Augsburg [38] oder Salzburg [10].

Ziel und Struktur der Arbeit

Das Ziel der Arbeit ist, Daten zum Gebäudebestand in einer Stadt oder Kommune aus vielen und zum Teil nicht flächendeckenden Quellen zusammenzuführen. Die Daten beziehen sich auf unterschiedliche geografische Skalen wie Adressen, Gebäude oder Gebiete und beinhalten unklassifizierte Angaben in natürlicher Sprache. Dies macht es möglich, mit begrenzten finanziellen sowie personellen Kapazitäten eine Gebäudedatenbank aufzubauen, die als Diskussions- und Arbeitsgrundlage für die Umstrukturierung des lokalen Wärmesektors genutzt werden kann. Die Datenverarbeitung erfolgt gebäudescharf. Darüber hinaus soll der Einfluss von Gebäudeparametern auf den Energieverbrauch für Raumheizung analysiert werden. Die mit der Prognose von Wärmeverbräuchen verbundenen Schätzfehler werden quantifiziert.

Kapitel 2 gibt eine Einführung in die Methoden, die in der Arbeit verwendet werden. Kapitel 3 erläutert das Konzept zur Verarbeitung der Gebäudeparameter. Hierzu gehört die Erfassung der Original- bzw. Primärparameter, die Zusammenführung und Konfliktbereinigung der Primärparameter und der Umgang mit Raumbezügen.

Kapitel 4 bis 6 beziehen sich auf die Umsetzung des Datenverarbeitungskonzepts für das Untersuchungsgebiet Oldenburg. Kapitel 4 gibt einen Überblick und eine kurze Beschreibung der Datenquellen. Kapitel 5 zeigt die Berechnung und Analyse der für verschiedene Verfahren der Wärmebedarfsberechnung bzw. für die Verbrauchsdatenanalyse notwendigen Gebäudeparameter auf Basis der in Oldenburg verfügbaren Daten. In Kapitel 6 werden die Gebäudeparameter mit Wärmeverbrauchsdaten kombiniert und die Art sowie die Stärke des Zusammenhangs für ca. 17.500 Ein- und Zweifamilienhäuser statistisch analysiert.

2. Methoden

In diesem Kapitel werden Begriffe und Methoden erklärt, die in der Arbeit zur Anwendung kommen.

2.1. Grundlagen und Begriffe relationaler Datenbanksysteme

Eine relationale Datenbank besteht aus miteinander verknüpften Tabellen. Sie ermöglicht die Speicherung großer Datenmengen und erfüllt folgende Anforderungen, die im Rahmen dieser Arbeit an die Datenverwaltung gestellt werden:

- Daten werden so gut wie nicht redundant gehalten. Die Einschränkung „so gut wie“ ermöglicht zum Teil eine einfachere Verarbeitung.
- Die Konsistenz der Daten ist gesichert.
- Die Verarbeitung der Daten erfolgt schnell, es gibt einen einfach handhabbaren und vielseitigen Zugriff auf die Daten.

Das Strukturierungsmodell einer relationalen Datenbank basiert auf der relationalen Algebra und den darin formulierten Mengenoperationen. Mit der weitgehend standardisierten Datenbanksprache *Structured Query Language* (SQL) ist die Manipulation der Daten mit nur wenigen Zeilen Code möglich [39, S.19ff].

Eine Tabelle repräsentiert ein Objekt, beispielsweise das Objekt „Gebäude“. Beschrieben werden kann das Gebäude anhand von Attributen. Jede Tabellenspalte beinhaltet ein Attribut, z. B. die Gebäudehöhe. Jede Tabellenzeile enthält die Daten für eine Objektinstanz, z. B. für ein bestimmtes Gebäude mit den zugehörigen Attributen. Anhand der Anzahl und der semantischen Eigenschaften der Attribute wird ein Objekt beschrieben. Diese Eigenschaften spiegeln nicht die gesamte Realität wieder, sondern stellen ein Modell der Realität dar. Das Modell umfasst die für die Anwendung relevanten Aspekte. Ein Beispiel bietet die Tabelle „Gebäude“ in Abbildung 2.1(a). Die Tabelle enthält die vier Attribute „GebäudeID“, „FlurstückID“, „Grundfläche“ und „Baujahr“. In diesem Fall interessiert die Anwendung nur, auf welchem Flurstück ein Gebäude steht, welche Fläche der Gebäudeumriss umfasst und wann das Gebäude gebaut wurde.

Die Struktur der Tabelle Gebäude wird in Abbildung 2.1(b) in Anlehnung an ein Entity-Relationship-Diagramm dargestellt. Dies ist eine grafische Methode zur Darstellung von komplexen Beziehungsgebilden und dient dem Entwurf von Datenmodellen oder Software allgemein [7]. Das Attribut „GebäudeID“ ist mit „PK“, d. h. Primärschlüssel bzw. *Primary*

2. Methoden

Key gekennzeichnet. Der Wert eines Primärschlüsselattributs identifiziert einen Datensatz eindeutig in einer Tabelle. Er darf weder leer sein noch mehrmals vorkommen und wird durch die Formatierung fett gekennzeichnet. Alle mit „ID“ endenden Attribute sind in dieser Arbeit Primärschlüssel.

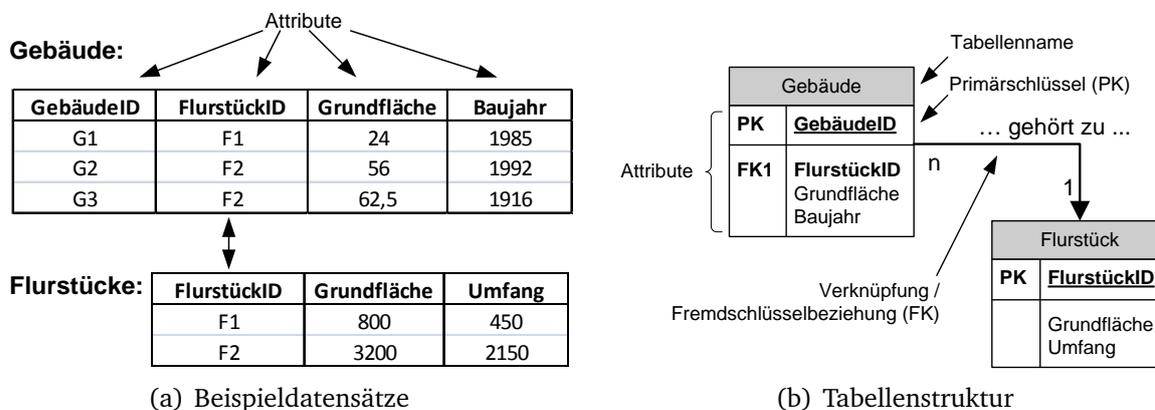


Abbildung 2.1.: Darstellung zweier in Relation zueinander stehender Tabellen

Ein wichtiger Aspekt relationaler Datenbanken ist die sogenannte Normalisierung der Daten. Hierdurch soll vermieden werden, immer gleiche Attributwerte in jeweils mehreren Datensätzen stehen. Anhand von Referenzen, sogenannten Fremdschlüsselbeziehungen, werden zusammengehörige Datensätze verschiedener Tabellen zueinander in Beziehung gesetzt [67]. In Abbildung 2.1 steht die Tabelle Gebäude in einer relationalen Beziehung zur Tabelle Flurstück. Die Verbindung zwischen einem Gebäudedatensatz und einem Flurstückdatensatz erfolgt aufgrund des Attributs und zugleich Fremdschlüssels (FK) „FlurstückID“ in beiden Tabellen. Da die FlurstückID Primärschlüssel in der Tabelle „Flurstück“ ist, verweist ein Datensatz der Tabelle „Gebäude“ eindeutig auf einen Datensatz in der Tabelle „Flurstück“. Die Kardinalität der Verbindung ist $n : 1$. Das heißt, es gibt keines, eines oder mehrere Gebäude ($n \in [0, 1, 2, 3, \dots]$), die einem Flurstück zugeordnet sind. Genau aufgrund dieser Möglichkeit, dass mehrere Gebäude auf dem selben Flurstück stehen, sind die Daten in zwei Tabellen getrennt, d. h. normalisiert. Bei nur einer Tabelle müssten für jedes Gebäude alle dazugehörigen Flurstückinformationen gespeichert werden. Dies führt zu Redundanzen, die beispielsweise aufgrund verschiedener Schreibweisen Inkonsistenzen in den Daten nach sich ziehen können. Durch die Fremdschlüsselbeziehungen werden zueinander konsistente Datensätze erzwungen. Im oberen Beispiel muss beim Einfügen von neuen Gebäuden das dazugehörige Flurstück in der Tabelle Flurstück existieren. Sonst verweigert das Datenbanksystem das Anlegen des neuen Datensatzes. Die Konsistenzbedingungen hängen von der Güte des Datenbankmodells ab und liegen in der Verantwortung des Datenbankdesigners.

Außer der dargestellten $n : 1$ -Kardinalität von Fremdschlüsselbeziehungen kommen in der Arbeit $n : m$ -Kardinalitäten vor, wobei n bzw. m Werte größer gleich Null ($n, m \geq 0$) annehmen können. Dies ist beispielsweise der Fall bei der Zuordnung von Gebäude und Adressen. Ein Gebäude kann keine bis mehrere Adressen, z. B. Gebäudeeingänge, haben und zugleich kann eine Adresse auf kein bis mehrere Gebäude verweisen. Dieser Fall wird mit Hilfe einer dritten Tabelle implementiert (siehe hierzu [67]).

Datenbanksysteme ermöglichen darüber hinaus eine Benutzerverwaltung und einen Zugriffsschutz für die Daten oder einzelne Bestandteile sowie den parallelen Zugriff mehrerer Benutzer. Details zum Aufbau relationaler Datenbanken auf Basis der Normalformenlehre und eine Einführung in SQL bieten [39, 67].

Die Attribute relationaler Datenbanken sind atomare Objekte, die nicht weiter sinnvoll unterteilt werden können. Hierzu gehört beispielsweise eine Zahl oder Zeichenkette. Objektrelationale Datenbanken erweitern dieses Konzept um die Möglichkeit, strukturierte Datentypen, sogenannte Objekte, als einzelnes Attribut einer Tabelle zu speichern und Objektmethoden zu definieren. Damit ist es möglich, Geodaten durch die Definition von Klassen und zugehörigen Methoden in die Datenverwaltung mit SQL zu integrieren [11].

2.2. Geoinformationssystem und Geodatenbank

Ein Geoinformationssystem (GIS) ist eine auf räumliche Daten spezialisierte Sonderform eines Informationssystems, welches „zur Unterstützung beliebiger technischer und organisatorischer Einrichtungen bei der Erfassung, Speicherung, Verarbeitung und Darstellung von Informationen“ dient [11, S.2].

Eine Geodatenbank beinhaltet vorwiegend die Speicherung von Raum- und Sachdaten und ist daher primär nur ein Bestandteil eines GIS [11]. Es existieren verschiedene Geodatenbanksysteme. In der Arbeit wird das objektrelationale Datenbanksystem „Postgres“ mit der Erweiterung „PostGIS“ verwendet [78, 64]. Neben der Fähigkeit zur Geodaten-speicherung enthält eine PostGIS/Postgres-Datenbank eine umfangreiche Sammlung von Funktionen für die Verarbeitung von Geodaten. Dadurch erfüllt sie die im Rahmen dieser Arbeit an ein GIS gestellten Hauptanforderungen: die Speicherung und Verarbeitung der Daten.

Für den Import von Daten in eine PostGIS/Postgres-Datenbank existieren mehrere Schnittstellen. Räumliche Daten liegen meist in dem weit verbreiteten Shapefile-Format vor. Mithilfe eines *shp2pgsql*-Konverters werden die Daten als Tabelle in die Datenbank kopiert [34]. Die Darstellung der Daten in Form von Karten geschieht mithilfe des Programms Quantum GIS [65], die Darstellung von Grafiken und Statistiken mithilfe der Software R [5]. Das PostGIS/Postgres-Datenbanksystem und alle genannten Programme sind *Open Source*.

Die Speicherung von Geodaten als Objekte in PostGIS basiert auf dem weit verbreiteten Standard des *Simple-Feature-Modells*. Es umfasst Klassen und Methoden für die Beschreibung und Bearbeitung der Geometrien. Die abstrakte Klasse *Geometry* ist als Oberklasse aller Geometrien die Schnittstelle zwischen der Geodatenobjekt-Klassenhierarchie und dem umgebenden Datenbanksystem. Dies zeigt Abbildung 2.2. Darin ist das Beispiel des vorherigen Kapitels um die Datentypen der Attribute und die Spalte „Geometrie“ erweitert. Auf diese Weise „verschwindet“ die komplexe Struktur der Geodaten hinter der abstrakten Oberklasse *Geometry*. Das Geometrie-Attribut kann bei der Manipulation der Daten mit SQL wie die anderen Attribute behandelt werden.

Tabelle Gebäude:

Attribute

Attributname	GebäudeID	FlurstückID	Grundfläche	Baujahr	Geometrie
Datentyp	character	character	double	integer	Geometry
Datensätze	G1	F1	24	1985	Polygone(...
	G2	F2	56	1992	Polygone(...
	G3	F2	62,5	1916	Polygone(...

Abbildung 2.2.: Tabelle mit Sach- und Geometrieattributen

Grundsätzlich werden im *Simple-Feature-Modell* die Geometrien in punktförmige, linienförmige und flächenhafte Objekte unterschieden [82]. Darüber hinaus gibt es sogenannte *GeometryCollection*, die mehrere Entitäten von jeweils einer Grundgeometrieart beinhalten. Ein Beispiel der drei Grundgeometriearten gibt Abbildung 2.3.

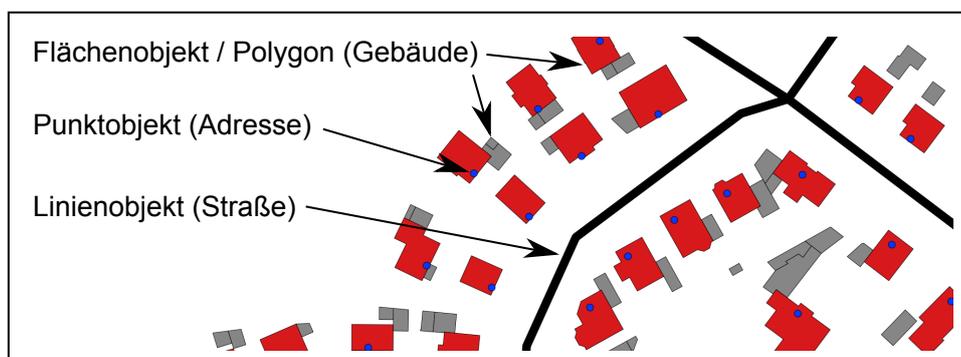


Abbildung 2.3.: Karte der drei Grundgeometriearten; Hauptgebäude in Rot, Nebengebäude in Grau

Linien- und Flächenobjekte können auf eine Menge von Punktobjekten zurückgeführt werden. Ein Gebäudepolygon ist definiert aufgrund der Lage seiner Eckpunkte und deren Reihenfolge. Zwei Eckpunkte werden jeweils mit einer geraden Strecke verbunden. Die Lage eines Punktes wird durch die angegebenen Koordinatenwerte immer relativ zu einem Koordinatenbezugssystem angegeben. Geodaten mit unterschiedlichen Koordinatenbezugssystemen müssen in ein gemeinsames System umgerechnet werden. Im Fall von Geodaten sind die Koordinatenbezugssysteme georeferenziert, d. h. sie beziehen sich auf Positionen der Erdoberfläche. Für georeferenzierte Koordinatenbezugssysteme existiert ein eindeutiger *European Petroleum Survey Group Geodesy* (EPSG)-Code. Anhand dieses EPSG-Codes wird das Koordinatenbezugssystem von Geodaten in PostGIS angegeben und kann mithilfe eines Funktionsaufrufs in andere Bezugssysteme umgerechnet werden [11].

In der Arbeit werden die Geodaten in dem projizierten Gauß-Krüger-Koordinatensystem mit dem EPSG-Code 31467 verwendet. In einem projizierten Koordinatensystem wird ein Teil der Oberfläche der dreidimensionalen Erdoberfläche auf eine zweidimensionale Ebene abgebildet. Das projizierte Koordinatenbezugssystem besitzt rechtwinklig zueinander

stehende Achsen und erleichtert die Darstellung auf einem Kartenblatt. Eine Erklärung unterschiedlicher Koordinatenbezugssysteme und deren Zusammenhang gibt [11].

2.3. Grundlagen und Begriffe des Text Mining

Texte repräsentieren Wissen und Informationen in nicht strukturierter Form [40]. Daher ist es nicht möglich, diese Informationen ohne vorherige Bearbeitung in relationalen Datenbanken zu speichern.

Text Mining umfasst statistische oder musterbasierte Verfahren, um Zusammenhänge und Informationen in Texten automatisiert oder semi-automatisiert zu erkennen. Dies ist von Nutzen bei der Verarbeitung großer Mengen von Text. Die erkannten Strukturen sollen in Analogie zum *Data Mining* neue Erkenntnisse generieren, die vorher unbekannt waren [40, 47].

In dieser Arbeit werden sowohl statistische als auch musterbasierte Verfahren eingesetzt. Statistische Verfahren dienen dem automatisierten Vergleich verschiedener und zum Teil falscher Adress-Schreibweisen, z. B. Hauptstraße, Hauptstr. oder Hptstraße (siehe Kapitel 3.2.3) für die richtige räumliche Zuordnung von Daten. Musterbasierte Verfahren ermöglichen eine semiautomatische Auswahl relevanter Datensätze beim Gebäudedatenimport. Kapitel 3.1.3 beschreibt, wie alle Vorkommen, die der Struktur „Neubau eines Gebäudes“ entsprechen, identifiziert werden.

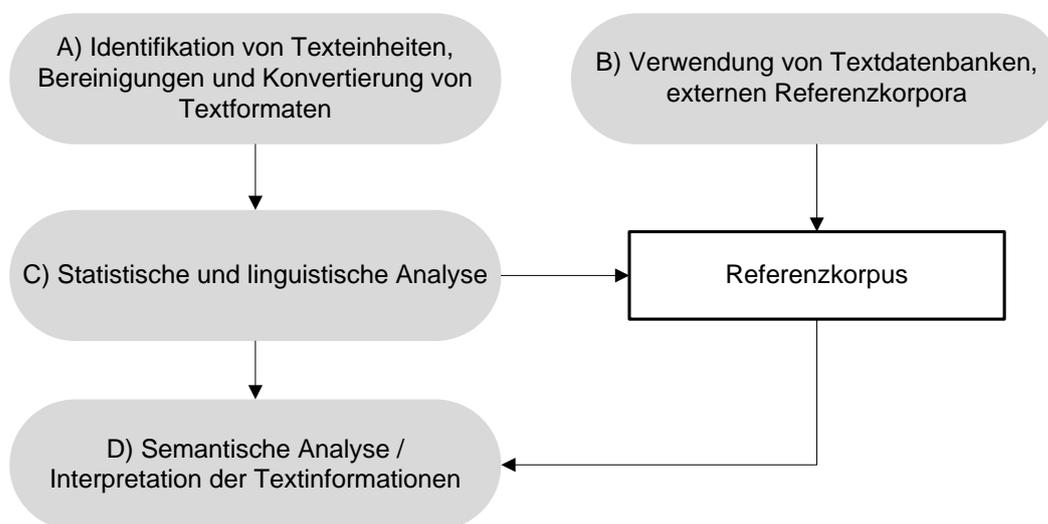


Abbildung 2.4.: Ablaufdiagramm von *Text Mining*-Analysen, eigene Abbildung in Anlehnung an [40]

Abbildung 2.4 skizziert die einzelnen Schritte der *Text Mining*-Verarbeitung. Der Schritt „A“ dient der Textvorbereitung. Einzelne Texteinheiten werden aus den sie umgebenden und zum Teil nicht klar abgegrenzten Strukturen extrahiert. Überflüssige Tabulatoren und Zeilenumbrüche werden entfernt. Vor allem für die richtige Verarbeitung von Umlauten

in der deutschen Sprache muss die Textkodierung erkannt und beim Einlesen der Texte berücksichtigt werden.

Grundlage vieler *Text Mining*-Verfahren ist, die Texte mit einem Referenzkorpus zu vergleichen. Ein Referenzkorpus ähnelt einem Wörterbuch. Er besteht aus einer Sammlung von Begriffen in richtiger Schreibweise, die unter Umständen mit zusätzlichen Informationen versehen sind. Das Hinzufügen zusätzlicher Informationen zu Wörtern nennt man „annotieren“. Annotationen können Informationen über die Art des Wortes enthalten, z. B. ob es ein Verb oder Nomen ist. Sie können zudem Fachinformationen oder semantische Verknüpfungen beinhalten, z. B. dass der Begriff „Einfamilienhaus“ ein Unterbegriff von „Gebäude“ ist. Referenzkorpora werden im Fachgebiet Computerlinguistik, zu der *Text Mining* gehört, in domänenspezifische und allgemeine linguistische Korpora eingeteilt. Domänenspezifische Korpora beinhalten vor allem Schlüsselwörter aus einer Wissensdomäne bzw. einem Fachbereich. Mit ihnen kann beispielsweise untersucht werden, ob ein zu testender Text zu einer bestimmten Wissensdomäne gehört [40].

Ein Referenzkorpus wird aus Stichwörtern oder Phrasen des zu untersuchenden Textes gebildet. Entsprechende Terme können durch eine statistische Analyse identifiziert und in den meisten Fällen manuell annotiert werden. Alternativ dazu stellen verschiedene Organisationen Textdatenbanken oder dateibasierte Referenzkorpora zur Verfügung (Schritt „B“). Beispiele hierfür sind das Wortschatzportal der Uni Leipzig [1], das Stammformenlexikon des Projekts Morphy [50] oder Informationen aus bereits durchgeführten Analysen [40].

Die statistische und linguistische Analyse von Texten kann beliebig komplex werden. Die Basis der Analyse ist nahezu immer die Zerlegung der Texte in einzelne Wörter. Stopwörter wie beispielsweise die Wörter „und“, „in“ oder „zum“ werden entfernt. Diese Wörter kommen sehr häufig vor, tragen jedoch oft nicht zum Inhalt bei.

Vor allem im Deutschen kommen häufig sogenannte Komposita, d. h. zusammengesetzte Wörter vor. Zum Beispiel ist „Einfamilienhaus“ ein zusammengesetzter Ausdruck für „Haus für eine Familie“. Im Englischen dagegen besteht der Term aus zwei Wörtern: „*detached house*“. Die Zerlegung von Komposita in die einzelnen Bestandteile und deren Normalisierung basiert oftmals auf Referenzkorpora [40, 15].

Der letzte Schritt in Abbildung 2.4, die semantische Analyse ist das eigentliche Herzstück des *Text Mining*. Mithilfe der Referenzkorpora werden anhand von statistischen Verfahren und der Erkennung syntaktischer Muster Informationen aus dem Text gewonnen. Die zwei Anwendungsbereiche der Arbeit werden in den Kapiteln 3.1.3 und 3.2.3 detailliert beschrieben.

2.4. Wärmebedarfsbilanzierung und Klimakorrektur

Der Wärmebedarf ist die Energiemenge, die unter Annahme eines standardisierten Nutzerverhaltens auf Basis der bauphysikalischen Eigenschaften eines Gebäudes berechnet wird. Er ist eine rein kalkulatorische Größe [80]. Die Berechnungsmethoden können in die

drei Kategorien Hochrechnungen, Bilanzverfahren und Simulationen eingeteilt werden. Hochrechnungen weisen die geringste Komplexität auf, Simulationen die höchste. Mit der Komplexität des Verfahrens werden immer mehr Gebäudeeigenschaften berücksichtigt, zugleich steigt jedoch auch der Umfang der für die Berechnung zu ermittelnden Daten. Nähere Informationen dazu bietet die im Rahmen dieser Arbeit entstandene Publikation „Verbesserte Aussagekraft kommunaler Wärmebedarfsanalysen“ [9].

Vom Wärmebedarf unterschieden werden muss der Wärmeverbrauch. Dieser bezeichnet die tatsächlich messbare Menge an Energie. Der Verbrauch hängt neben den Gebäudeparametern vom individuellen Nutzerverhalten und Effekten ab, die nicht oder nur mit einem hohen Datenaufwand zu erfassen sind. Der Wärmebedarf eines Gebäudes kann vom Wärmeverbrauch um den Faktor 2 bis 3 abweichen [9, 53].

2.4.1. Energieform

Die Angabe der Energieform ist wichtig, wenn Wärmeverbrauchsdaten aus unterschiedlichen Quellen oder Ergebnisse verschiedener Wärmeberechnungsverfahren miteinander verglichen werden [66, 46]. Hierfür werden in der Arbeit die in Abbildung 2.5 dargestellten Begriffe der Energieumwandlungskette verwendet.

- **Primärenergie** ist die geförderte Energie inklusive der für die Förderung und den Transport aufgewendeten Energie. Sie entspricht dem Öl bzw. Gas im Boden oder dem noch im Wald befindlichen Holz.
- **Endenergie** ist die am Gebäude angelieferte Energie in Form von Heizöl, Gas oder Holz. Bei der Umwandlung von Primärenergie entstehen im Falle fossiler Brennstoffe Förderverluste und ganz allgemein Transportverluste.
- **Heizwärme** ist die Energie des warmen Heizungswassers nach dem Verlassen der Feuerungsanlage. Bei der Verfeuerung entstehen Anlageverluste durch den Wirkungsgrad der Anlage und die nicht kontinuierliche Nutzung der Anlage.
- **Raumwärme** bezeichnet die Energie der warmen Raumluft im Gebäude. Bei der Berechnung sind jedoch Leitungsverluste des Gebäude-Heizsystems und Verteilverluste am Heizkörper abzuziehen.

2.4.2. Klimatische Anpassung

Klimarandbedingungen der Bedarfsberechnung Für die zeitliche und räumliche Vergleichbarkeit von Wärmebedarfsberechnungsverfahren und Berechnungskennwerten müssen die verwendeten klimatischen Bedingungen bekannt sein. Eine wichtige Angabe ist die Gradtagszahl (GTZ), die aus der Summe der Temperaturdifferenz zwischen einer definierten Rauminnentemperatur (z. B. 20°C) und der mittleren Außentemperatur eines Tages (z. B. 3°C) berechnet wird, wenn die Außentemperatur eine gewisse Heizgrenztemperatur (meist 15°/12°C) an dem Tag unterschreitet [80]. Bei Temperaturen unterhalb der Heizgrenztemperatur wird davon ausgegangen, dass ein Gebäude beheizt ist. Zur

2. Methoden

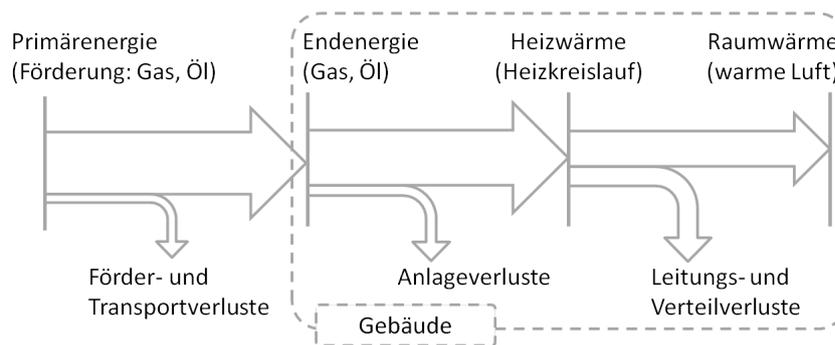


Abbildung 2.5.: Energieformen und Verluste entlang der Energieumwandlungskette mit Bezug zur Raumwärme im Gebäudebereich

Nachvollziehbarkeit der GTZ müssen daher die Innenraumtemperatur und die Heizgrenztemperatur mit angegeben werden. Nach obigem Beispiel lautet die Notation $GTZ(20/15)$ bzw. $GTZ(20/12)$.

Der Standort für das deutsche Referenzklima ist bis zum 31.12.2011 Würzburg, ab dem 01.01.2012 ist es Potsdam [69]. Die Angaben der Gradtagszahlen verschiedener Literaturquellen sind nicht einheitlich, z. B. werden für das langjährige Mittel in Deutschland Werte von $GTZ_D(19/15)=86,4 \text{ kWh/a}$ [24, Tab.A.2] und $GTZ_D(19/15)=82,296,4 \text{ kWh/a}$ [27] angegeben.

Die GTZ haben einen großen Einfluss auf die nach DIN-V-4108-6 [24] berechneten Wärmebedarfe. Anhand der GTZ werden die aufgrund der Gebäudeeigenschaften berechnete Verlustleistung je Kelvin Temperaturdifferenz hochgerechnet (siehe Kapitel 2.4.3). Der Wärmebedarf ist direkt proportional zur GTZ.

Klimabereinigung von Wärmeverbräuchen Klimakorrekturefaktoren dienen der Umrechnung von gemessenen Wärmeverbrauchswerten aus unterschiedlichen klimatischen Bedingungen. Der Klimakorrekturefaktor k ist bis 2009 aus dem Verhältnis der GTZ des mittleren deutschen Klimas $GTZ_D(20/15)$ mit 20°C Innen- und 15°C Heizgrenztemperatur und der entsprechenden lokalen $GTZ_{\text{Lokal}}(20/15)$ berechnet wurden [35].

Seit 2009 ist ein neues Verfahren durch das Bundesministerium für Verkehr, Bau und Stadtentwicklung (BmVBS) vorgeschrieben, welches mithilfe eines Messnetzes von ca. 500 Stationen in Deutschland und einer räumlichen Interpolation genauere Korrekturfaktoren für 8200 Postleitzahlbereiche in Deutschland liefert [35, 36, 14, 19]. Das neue Verfahren berücksichtigt die Unterschiede in den klimatischen Verhältnissen, da es neben der Temperatur mehrere Einflussgrößen wie beispielsweise Geländehöhe und Bebauungsdichte einbezieht. Es werden städtische und ländliche Gebiete klimatisch unterschieden. Zum Beispiel weisen die innenstadtnahen Postleitzahlbereiche 26121 bis 26125 in Oldenburg für alle angegebenen Jahre höhere Klimakorrekturefaktoren auf als die Außenbezirke (siehe Tabelle 2.1). Ein Gebäude in der Innenstadt verbraucht somit

klimabereinigt weniger Energie als das identische Gebäude im Außenbezirk. Die Abweichung von Klimakorrekturfaktoren zwischen dem neuen und alten Verfahren sind zum Teil erheblich [36].

Tabelle 2.1.: Klimafaktoren für Oldenburg im Überblick

Verfahren	Bezug	Jahr			
		2007	2008	2009	2010
Altes Verfahren [27]	Wetterstation Emden (Flughafen)	1,21	1,13	1,09	0,93
Neues Verfahren [19]	Durchschnitt Oldenburg	1,28	1,19	1,16	0,98
	PLZ 26121	1,31	1,21	1,18	0,99
	PLZ 26122	1,29	1,20	1,17	0,98
	PLZ 26123	1,30	1,20	1,18	0,98
	PLZ 26125	1,29	1,20	1,17	0,98
	PLZ 26127	1,28	1,19	1,16	0,98
	PLZ 26129	1,27	1,18	1,15	0,97
	PLZ 26131	1,24	1,16	1,13	0,96
	PLZ 26133	1,26	1,18	1,15	0,97
	PLZ 26135	1,26	1,18	1,15	0,97

2.4.3. Heizenergiebilanzierung

In der Arbeit werden die Wärmebedarfe für den Sektor Haushalte auf Basis des Heizperiodenbilanzverfahrens berechnet. Das Verfahren ist als Funktion der Gebäudedatenbank implementiert. Dadurch ist die automatisierte Berechnung des Wärmebedarfs für alle Gebäude möglich. Das Heizperiodenbilanzverfahren berechnet den Jahres-Raumwärmebedarf. Es werden für die gesamte Heizperiode konstante mittlere Randbedingungen angenommen [24]. Einflüsse von Anlagen- und Verteilverlusten werden anhand von Faktoren bei der Umrechnung des Raumwärmebedarfs in andere Energieformen berücksichtigt.

Der Raumwärmebedarf Q_h eines Gebäude ergibt sich aus vier bauphysikalischen Effekten, die bilanziell miteinander verrechnet werden (Gleichung 2.1). Dazu gehören Transmissionsverluste Q_T aufgrund der Wärmeleitung durch die thermische Hülle des Gebäudes (Wände, Fenster, Dach, Boden), Lüftungsverluste Q_V durch den Austausch von warmer Innen- gegen kalte Außenluft, innere Wärmegewinne Q_i durch die Abwärme von Menschen und elektrischen Geräten und solare Wärmegewinne Q_S aufgrund der Sonneneinstrahlung durch Fensterflächen. Die Transmissions- und Lüftungsverluste ergeben sich aus der Multiplikation der Transmissions- bzw. Lüftungsverlustleistung H_T bzw. H_V mit der GTZ.

$$Q_h = Q_T + Q_V - (Q_i + Q_S) = (H_T + H_V) \cdot \text{GTZ} - (Q_i + Q_S) \quad (2.1)$$

2. Methoden

Ein übersichtliche Darstellung der Berechnung bietet die Veröffentlichung „Energetische Bilanzierung von Wohngebäuden nach DIN V 18599“ [41, S.8].

In die Berechnung fließen die Bauteilflächen der Gebäude unter Berücksichtigung angrenzender Gebäude, Wärmeleitwerte (U-Wert) der Gebäudetypologie des Institut für Wohnen und Umwelt, Darmstadt (IWU) und die in Tabelle 2.2 angegebenen Randbedingungen. Bei der Umrechnung des Raumwärmebedarfs in Endenergie werden Leitungsverluste mit der Aufwandszahl $e_V = 1,16$ und Anlagenverluste mit der Aufwandszahl $e_p = 1,1$ angenommen. Der Raumwärmebedarf wird mit den Aufwandszahlen multipliziert.

Die Gebäudetypologie des IWU für Wohngebäude klassifiziert die Gebäude anhand von Baualtersklassen, in der diese errichtet wurden und der Größe bzw. Art des Gebäudes (z. B. Einfamilienhaus, Mehrfamilienhaus bzw. Hochhaus). Die Baualtersklassen orientieren sich in ihrer Einteilung an historischen Ereignissen bzw. Wohngebäude betreffende gesetzlichen Regulierungen (siehe Kapitel 5.1.2). Die Untersuchung erfasst Gebäude in ganz Deutschland und gibt für jede Gebäudeklasse einen Repräsentanten an, der die für die Klasse spezifische Bauweise im Mittel beschreibt. Die Wärmeleitwerte sind IST-Werte, bei denen Sanierungen anteilig berücksichtigt sind [53, 57].

Tabelle 2.2.: Berechnungsparameter für das Heizperiodenbilanzverfahren

Parameter	Wert	Bemerkung
Wärmebrücken- koeffizient	0,05	Aufschlag auf die Transmissionswärmeverluste durch nicht berücksichtigte Wärmebrücken
Luftaustauschfaktor	0,19	Anteil des Luftvolumens im Gebäude, das pro Stunde ausgetauscht wird
Solarer Faktor	0,567	Korrekturfaktor für solare Gewinne
Spezifische Innere Gewinne	22	Innere Gewinne in kWh/m ² a
Auslastungsgrad	0,95	
Gebäudetypologie	IWU 2011 [53]	Die Gebäudetypologie liefert die statistischen Angaben zu den U-Werten der Bauteile der thermischen Hülle
Klimadaten	81,192 kWh/a	Mittleres Klima des Referenzpunktes Würzburg

Die Monatsbilanz- und DIN-V-18599-Verfahren sind Wärmebedarfsberechnungsverfahren mit einer feineren zeitlichen Auflösung. Das Monatsbilanzverfahren bietet eine monatliche Auflösung des Raumwärmebedarfs [57, 61]. Das DIN-V-18599-Berechnungsverfahren modelliert Gebäude und Anlagentechnik gemeinsam. Darüber hinaus werden verschiedene Zonen im Gebäude mit unterschiedlichen Nutzungszeiten und -temperaturen berücksichtigt [22]. Die erhöhte zeitliche Auflösung bietet keinen Mehrwert im Rahmen dieser Arbeit, benötigen jedoch einen höheren Implementierungsaufwand. Ein weiteres Beispiel für den Einsatz des Heizperiodenbilanzverfahrens ist das Projekt Tabula des IWU [53].

2.5. Regressionsanalyse

Das Ziel der Regressionsanalyse ist es, die Wirkung einer unabhängigen Variablen auf eine abhängige Variable zu analysieren. Falls mehrere unabhängige Variablen berücksichtigt werden, spricht man von einer Multiplen Regressionsanalyse. Sie gehört zu den Multivariaten Analyseverfahren der Statistik [3]. Untersucht werden kann beispielsweise, inwiefern der Heizwärmeverbrauch (abhängige Variable) eines Gebäudes von dessen Volumen abhängt (unabhängige Variable).

Die abhängige Variable muss metrisch, die unabhängigen Variablen können metrisch und nominal sein [3]. Eine nominale Variable ist beispielsweise die Baualtersklasse. Den Baualterklassen kann keine Rangordnung zugewiesen werden, da nicht grundsätzlich eine Klasse besser oder höher ist als die andere.

Die lineare Regression zeigt den linearen Zusammenhang zwischen den Variablen, der die geringste Abweichung zwischen tatsächlichen und geschätzten Werten aufweist. Verdeutlichen soll dies Abbildung 2.6 anhand einer beispielhaften Stichprobe von 15 Gebäuden. Die Datenpunkte bezeichnen den Heizwärmeverbrauch für Gebäude unterschiedlicher Volumina. Die Regressionsgerade \hat{Q} zeigt zum Teil hohe Abweichungen von den tatsächlichen Datenpunkten. Ihr liegt die Regressionsfunktion

$$\hat{Q} = b_0 + b_1 V \quad (2.2)$$

zugrunde, wobei \hat{Q} der durch die Regression geschätzte Wärmeverbrauch und V das Gebäudevolumen ist. b_0 und b_1 sind die Koeffizienten, die durch die Minimierung des verbleibenden Fehlers zwischen dem tatsächlichen Verbrauch Q und dem geschätzten Verbrauch \hat{Q} bestimmt werden.

Der verbleibende Fehler wird Residualgröße bzw. Residuum ϵ genannt mit $\epsilon = Q - \hat{Q}$. Der gemessene Verbrauchswert kann ausgedrückt werden als Summe des geschätzten Verbrauchs und der Residualgröße (siehe Gleichung (2.3)).

$$Q_k = \hat{Q}_k + \epsilon_k = b_0 + b_1 V_k + \epsilon_k \quad (2.3)$$

Das Bestimmtheitsmaß R^2 ist ein Gütemaß zur Überprüfung, wie gut das Regressionsmodell die Realität abbildet. Es ist definiert als Quotient der erklärten Streuung zur Gesamtstreuung (siehe Gleichung (2.4)). Die Streuung entspricht hierbei der Varianz, d. h. der Quadratsumme der Abweichung zwischen dem entsprechenden Wert und dem Mittelwert.

$$R^2 = \frac{\sum_k (\hat{Q} - \bar{Q})^2}{\sum_k (Q - \bar{Q})^2} = \frac{\text{erklärte Streuung}}{\text{Gesamtstreuung}} \in [0, 1] \quad (2.4)$$

Beispielhaft wird dies in Abbildung 2.6 anhand des Datenpunkts 13 gezeigt. Der Mittelwert wird als horizontale gestrichelte Linie dargestellt. Die erklärte Streuung ist die Differenz zwischen dem Mittelwert $\bar{Q} = 8,9 \text{ MWh/a}$ und dem geschätzten Verbrauch $\hat{Q}_{13} = 11,0 \text{ MWh/a}$. Die Gesamtstreuung ist die Differenz zwischen Mittelwert und tatsächlichem Verbrauch $Q_{13} = 12,7 \text{ MWh/a}$.

2. Methoden

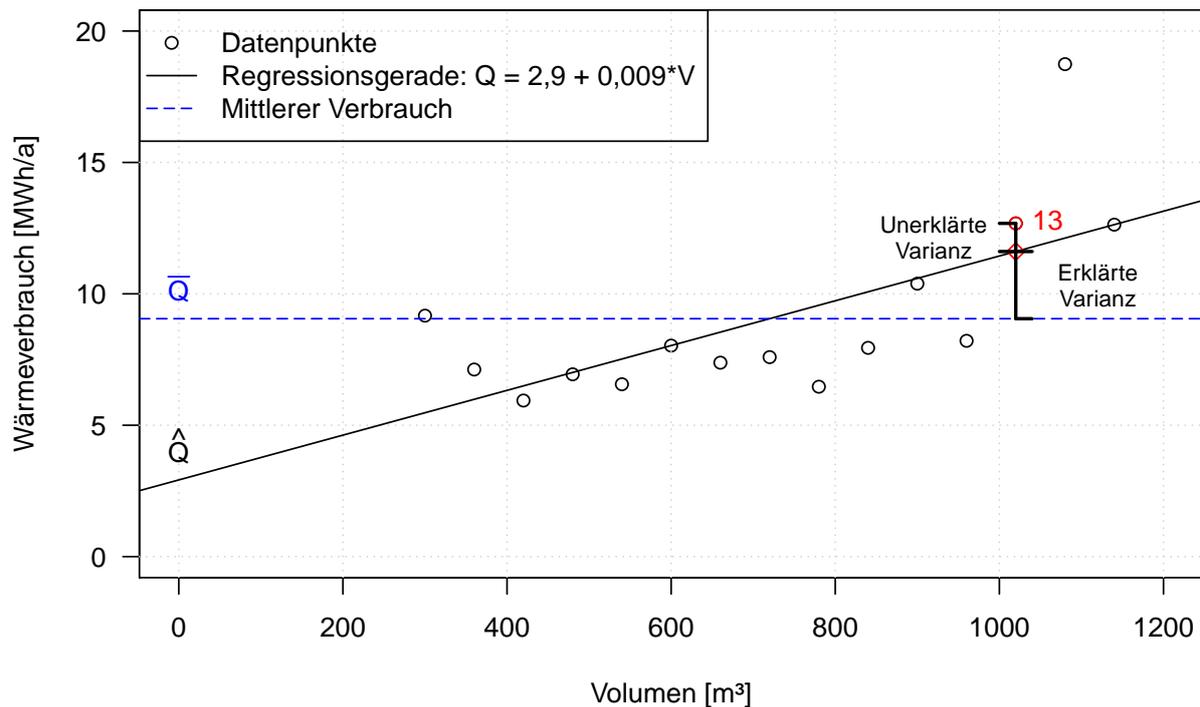


Abbildung 2.6.: Lineare Regression für Wärmeverbräuche von Gebäuden unterschiedlichen Volumens

Für alle 15 Datenpunkte ergibt sich für das Beispiel ein Bestimmtheitsmaß $R^2 = 0,36$. Das bedeutet aus Gleichung (2.4), dass 36 % der gesamten Varianz durch die Regression erklärt werden können. Für den Fall mehrerer unabhängiger Variablen wird das Bestimmtheitsmaß noch um die Zahl der Freiheitsgrade korrigiert und als „Korrigiertes Bestimmtheitsmaß“ R_{kor}^2 bezeichnet (siehe hierzu [3]).

Die 15 Datenpunkte sind eine Stichprobe aus der Grundgesamtheit von vielen Gebäuden. Die für diese Stichprobe ermittelten Regressionsparameter b_0 und b_1 sind Realisationen von Zufallsvariablen, da sie von den zufälligen Schwankungen der Datenpunkte abhängen [3, S.77]. Anhand eines statistischen F-Tests wird geprüft, ob der Einfluss der unabhängigen Variable Volumen auf die abhängige Variable Verbrauch wirklich existiert oder nur zufällig in dieser Stichprobe beobachtet wird. Das Ergebnis des Tests ist das sogenannte Signifikanzniveau (p-Wert). Der p-Wert deutet die Wahrscheinlichkeit an, dass die Stichprobe der 15 Beispielwerte existiert, wenn man zugleich annimmt, dass es keinen Zusammenhang zwischen Volumen und Wärmeverbrauch gibt. Er wird berechnet als bedingte Wahrscheinlichkeit. Der p-Wert für das Beispiel ist $p = 0,017$. Diese Angabe wird mit einer Vertrauenswahrscheinlichkeit verglichen. Bei einer üblichen Vertrauenswahrscheinlichkeit von $\alpha = 95\%$ ist das Referenz-Signifikanzniveau $1 - \alpha = 0,05$. Da für das Beispiel $p < 1 - \alpha$ gilt, wird von einem Einfluss des Volumens V auf den Wärmeverbrauch Q ausgegangen. Eine Vertiefung der hier nur knapp behandelten Thematik bieten [3, 42].

Der linearen Regression liegen Prämissen zugrunde, deren Einhaltung geprüft werden

muss [3].

- Hierzu gehört vor allem, dass die Residualgröße bzw. der Fehlerterm $\epsilon_k \sim N(0, \sigma^2)$ normalverteilt um den Mittelwert Null ist mit einer für alle k konstanten Varianz σ^2 , auch Homoskedastizität genannt.
- Es darf keine lineare Abhängigkeit zwischen den erklärenden Variablen geben und die Fehlerterme ϵ dürfen nicht miteinander korrelieren.

Eine ansatzweise Verletzung der Prämissen stellt nicht das gesamte Ergebnis in Frage, sollte jedoch bei der Bewertung desselben berücksichtigt werden [3].

3. Konzept zur Verarbeitung der Gebäudeparameter

Im Untersuchungsgebiet Oldenburg und in vielen anderen Gemeinden stehen Gebäudedaten für die Erstellung von Wärmekatastern bzw. Heizenergieanalysen nicht zentralisiert zur Verfügung, sondern müssen aus unterschiedlichen Quellen extrahiert und zusammengeführt werden. Bisherige Ansätze sind geprägt von einem hohen Zeitaufwand bei der Aufarbeitung der Datengrundlage [52, 37]. Dieser Zeitaufwand entsteht nicht nur durch den Datenbezug und die damit verbundenen Abstimmungsprozesse, sondern auch durch die technische Zuordnung der erhaltenen Informationen und die notwendige Weiterverarbeitung.

Anforderungen Es können somit folgende Anforderungen an das Datenverarbeitungskonzept identifiziert werden:

- Da die Beschaffung von Informationen und Daten sehr zeitaufwendig ist, liegt mit Beginn der Datenanalyse meist keine finale Datenmenge vor. Neu hinzukommende Daten sollen schnell und automatisiert in die Analyse aufgenommen werden können.
- Eine transparente Datenverarbeitung soll es ermöglichen, unklare Ergebnisse auf möglicherweise fehlerhafte Datenquellen zurückführen zu können. Hierzu gehört, dass auch nach der Verarbeitung der Daten die Informationsquelle für jede Angabe bekannt ist und mit in die Analyse einfließen kann.
- Trotz des Strukturreichtums der Quelldaten soll eine möglichst generische Lösung als Werkzeug zur Verfügung stehen, die auf verschiedene Gebäudeparameter anwendbar ist. Sowohl eine Erweiterung um weitere Gebäudeparameter als auch eine Übertragbarkeit auf andere Untersuchungsgebiete soll möglich sein.
- Die Berücksichtigung verschiedener Datenquellen in unterschiedlichen EDV-Formaten, Raumbezügen und Erfassungszeitpunkten führt dazu, dass sich Angaben überschneiden und gegebenenfalls widersprechen. Diese Konflikte sollen möglichst automatisiert und nach klaren, nachvollziehbaren Regeln gelöst werden, bevor eine energiewirtschaftliche Analyse durchgeführt wird.
- Die aufbereiteten Gebäudedaten und darauf aufbauende Ergebnisse sollen räumlich als auch thematisch aggregiert werden können, beispielsweise für Vergleiche mit statistischen Daten.
- Die in der Arbeit ausgewerteten Aspekte werden nicht als abschließend betrachtet. Das Konzept soll deshalb Flexibilität bei der Beantwortung neuer Fragestellungen zeigen.

3. Konzept zur Verarbeitung der Gebäudeparameter

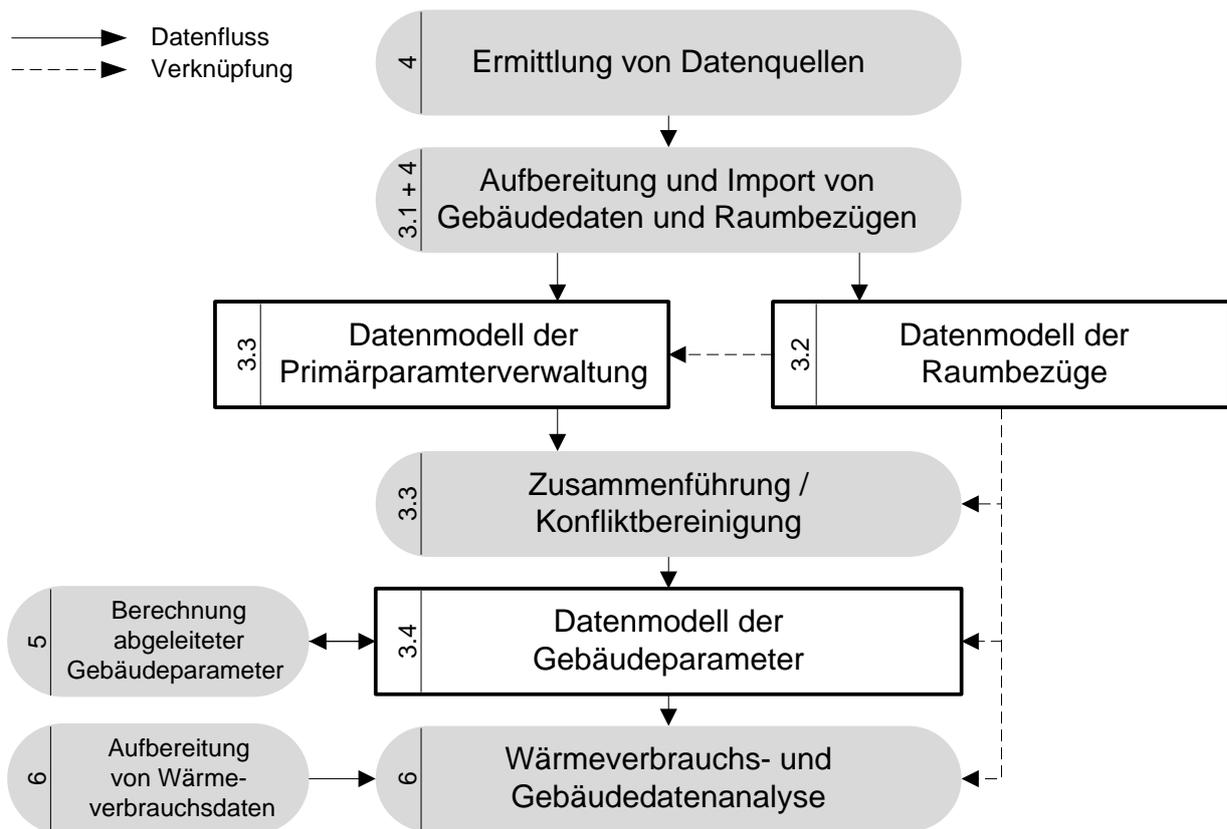


Abbildung 3.1.: Ablaufdiagramm der Gebäudedatenverarbeitung, Methoden (rund) und Datenmodelle (eckig) mit vertikalen Kapitelangaben

Die genannten Anforderungen treten beispielsweise bei den Angaben zu den Gebäudebaujahren in Oldenburg auf. Die Quellen reichen von historischen Aufarbeitungen bis zur digitalen Erfassung von Bauanträgen seit Mitte der 90er Jahre. Keine der Quellen bietet ein vollständiges Bild des Gebäudebestandes. Erst durch andauernde Recherchen konnten die Datenquellen identifiziert werden, teilweise erstreckte sich die Datenaufnahme weit in die Zeit der Datenanalyse hinein. Die Quellen weisen Adressbezüge, Gebäudebezüge und Gebietsbezüge auf. Angaben verschiedener Quellen bezüglich eines Gebäudes widersprechen sich.

Strukturierung der Datenverarbeitung Die Datenverarbeitung beginnt mit der Ermittlung und Integration unterschiedlicher Datenquellen und endet mit der Verfügbarkeit aufbereiteter Gebäudedaten für Wärmeverbrauchsanalysen (siehe Abbildung 3.1).

Da die Datenquellen unterschiedliche Formate und Strukturen aufweisen, wird die Aufbereitung und der Import an die ermittelten Quellen angepasst. Beim Datenimport werden Sach- und Raumdaten getrennt. Zugleich wird sichergestellt, dass Bezüge zwischen diesen korrekt gesetzt werden. Die separate Speicherung ermöglicht eine flexible und transparente Datenverarbeitung. Die für den Import notwendigen Datentypen und Analyseschritte werden in Kapitel 3.1 beschrieben.

Grundlage für die Zusammenführung und -analyse der Daten sind die Raumbezüge und deren Zugehörigkeit zueinander. Sie werden in dem Datenmodell der Raumbezüge verwaltet. Die Raumbezüge dienen der Zuordnung von Primärparametern und der räumliche Analyse, z. B. bei räumlicher Aggregation. Der Umgang mit Raumbezügen wird in Unterkapitel 3.2 beschrieben. Eine Voraussetzung für die erfolgreiche Datenverarbeitung ist, dass die Raumbezüge vollständig und korrekt bekannt sind.

Der Strukturreichtum der Informationen aus den Datenquellen wird durch das **Datenmodell der Primärparameterverwaltung** erfasst. Die Primärparameter sind die Informationen, die direkt aus den Datenquellen extrahiert werden. Sie stammen aus den verschiedenen Quellen, wurden zu verschiedenen Zeitpunkten aufgenommen und weisen unterschiedliche Raumbezüge auf. Um die Datenstruktur erweitern zu können, wird soweit möglich vermieden, Parameter, Quellen und Angaben fest in der Datenstruktur zu verankern (siehe Unterkapitel 3.3). Der Fokus liegt auf der möglichst originalgetreuen Erfassung der Quelldaten mit ihrem ursprünglichen Raumbezug. Die gespeicherten Daten hängen damit von den Datensätzen des Datenmodells der Raumbezüge ab.

Auf Basis dieses Konzepts ist eine weitgehend standardisierte und automatisierte **Zusammenführung** und **Konfliktbereinigung** der Daten möglich. Da sich Angaben unterschiedlicher Quellen widersprechen können, erzeugt die Zusammenführung einen Datensatz pro Gebäude mit den jeweils hochwertigsten verfügbaren Daten.

Das Datenmodell der **Gebäudeparameterverarbeitung** beinhaltet die für eine energie-wirtschaftliche Analyse notwendigen Parameter (siehe Unterkapitel 3.4). Komplexe bzw. nicht direkt in Quellen verfügbare Gebäudeparameter wie Oberfläche/Volumen-Verhältnis oder der Gebäudetyp werden aus den Primärparametern abgeleitet und berechnet.

Unterschiede zwischen der generalisierten Problemlösung und der Ausführung in Oldenburg Bei der Darstellung des Konzepts in diesem Kapitel werden zwei Herangehensweisen unterschieden. Einerseits wird in den Datenmodellen und Methoden allgemein die Struktur der Datenzuordnung und des Zusammenhangs zwischen verschiedenen Raumbezügen zu erfasst. Andererseits besteht diese Komplexität bei den vorhandenen Daten im Untersuchungsgebiet Oldenburg nicht immer. Zum Teil sind daher Vereinfachungen möglich, auf die im weiteren Text differenziert eingegangen wird.

3.1. Generalisierte Struktur der Datenquellen und Datenaufbereitung

Eine generalisierte Struktur von externen Datenquellen erlaubt es, die einzelnen Bestandteile einer konkreten Quelle zielgerichtet für den Import der Informationen zu analysieren. Es wird davon ausgegangen, dass die Daten tabellarisch vorliegen. Ein Datensatz kann somit in unterschiedliche Attribute unterteilt werden.

3.1.1. Attributtypen der Inputdaten

Datenquellen enthalten zum Teil Informationen über mehrere Gebäudeparameter und sind sehr unterschiedlich aufgebaut. Tabelle 3.1 gibt ein Beispiel für die tabellarische Struktur der Datenquellen. Für die bessere Gliederung der Datenverarbeitung werden Attributtypen definiert. Jedes Attribut, d. h. jede Spalte, kann einem Attributtyp zugeordnet werden.

Tabelle 3.1.: Bestimmung der Attributtypen für einen anonymisierten Beispieldatensatz

Akten- zeichen	Eingang	Grundstück	Vorhabenbeschreibung
12xxx- 1963	01.01.1963	Oldenburg Musterstra- ße 67	Neubau eines Einfamilienwohnhauses Hinw.: 12xxx-1983 Bauantr. v. 09.05.1983 Neubau v. Garagen
↓ ID	↓ Skalar	↓ Raumbezug	↓ Textfeld

Das Beispiel zeigt darüber hinaus, dass die Datenquellen Informationen für verschiedene Gebäudeparameter beinhalten können. Zum einen kann anhand der Spalte „Eingang“ das Baujahr der Gebäudes abgeleitet werden. Dies gilt jedoch nur, wenn der Datensatz einen „Neubau“ bezeichnet. Zum anderen enthält die „Vorhabenbeschreibung“ einen Hinweis auf die Gebäudeklasse „Einfamilienhaus“.

Sowohl der Raumbezug als auch die „Vorhabenbeschreibung“ sind Textfelder und werden mithilfe von *Text Mining*-Methoden analysiert (siehe Kapitel 3.2.3 und 3.1.3).

Die Attribute können in fünf verschiedene Attributtypen eingeteilt werden. Ein einzelner Datensatz kann keinen bis mehrere dieser Attributtypen beinhalten.

1. Das **Erstellungsdatum** oder kurz Datum des Datensatzes besagt, wann der Datensatz angelegt wurde und gibt Auskunft über die Aktualität der Daten. Bei historisch gewachsenen Quellen kommt es vor, dass innerhalb einer Quelle mehrere Datensätze mit dem gleichen Raumbezug existieren. Unter Umständen muss die Information noch in ein einheitliches Zeitformat „YYYY-MM-TT“ gebracht werden. Falls eine entsprechende Angabe - wie in oberem Beispiel - fehlt, muss sie geschätzt werden. Sie ist notwendig für die Zusammenführung der Primärparameter.
2. Die **ID** identifiziert den Datensatz eindeutig durch eine kodierte Bezeichnung, um die Ergebnisse verschiedener Bearbeitungsschritte wieder zueinander zu führen. Die Eindeutigkeit muss sichergestellt sein.
3. Der **Raumbezug** bringt die Gebäudeinformationen mit Positionen oder Orten auf der Erdoberfläche in Verbindung [7]. Er wird entweder direkt durch eine räumliche Geometrie angegeben oder verweist indirekt anhand einer kodierten Bezeichnung (ID) bzw. einer textlichen Lagebezeichnung, z. B. Adresse auf eine räumliche Geometrie. Die Bearbeitung wird in Kapitel 3.2 näher beschrieben.

4. Ein **Skalar** ist ein Feldtyp, dessen Daten entweder einen Zahlenwert (z. B. Baujahr „1956“) oder einen kodierten Bezeichner (z. B. Dachform „Flachdach“) beinhalten. Kodierte Bezeichner sind Zeichenketten, die bei gleicher Bedeutung die selbe Schreibweise aufweisen müssen.
5. Ein **Textfeld** enthält Beschreibungen als Fließtext oder besteht aus nicht normierten Bezeichnern. Zu letzterem gehören beispielsweise Angaben mit verschiedenen Schreibweisen. Aufgrund der Heterogenität und Komplexität von Textfeldern ist die Semantik nicht ohne weitere Bearbeitung automatisiert erschließbar. Es können auch mehrere Informationen in einem Feld enthalten sein, z. B. „Neubau von Einfam.-Haus mit Garage“. Gesuchte Merkmale müssen aus dem Text extrahiert und in eine einheitliche Schreibweise gebracht werden. Die Vorgehensweise ist näher in Kapitel 3.1.3 beschrieben.

3.1.2. Güteindikator für Datenquellen

Durch die Verwendung mehrerer Quellen kommt es vor, dass für einen Parameter eines Gebäudes mehrere sich widersprechende Angaben existieren. Aufgrund der Datenmenge ist eine manuelle Überprüfung der einzelnen Werte wirtschaftliche nicht zielführend. Die Datenquellen werden deshalb anhand eines sogenannten Güteindikators bezüglich ihrer Qualität bewertet. Bei der Zusammenführung der Quellen wird die Angabe der am besten bewerteten Quelle übernommen. Alle anderen Angaben bleiben ungenutzt.

Für die Bestimmung des Güteindikators werden folgende Sachverhalte berücksichtigt:

- Gründe für die Erfassung der Informationen: Ein gesetzlicher Auftrag ist in der Regel der stärkste Grund, da davon ausgegangen werden kann, dass die Daten nahezu immer richtig und sorgfältig erfasst worden sind.
- Weitere Hintergrundinformationen zu den Datenquellen wie beispielsweise die Einschätzung der mit der Datenaufnahme befassten Personen und Organisationen bzw. der Qualität.
- Verarbeitungsfähigkeit der Informationen: Beispielsweise Schwierigkeiten bei der Zuordnung der Daten zu den in der Datenbank gespeicherten Raumbezügen führen zu einer niedrigeren Bewertung.

Der Güteindikator gibt damit einen Anhaltspunkt, wie stark den Angaben einer Quelle vertraut werden kann. Darüber hinaus ist der relative Bezug zweier Güteindikatoren von verschiedenen Datenquellen wichtig, da bei der Zusammenführung der Angaben kein Mindestindikator vergeben wird, sondern die Angabe mit dem höchsten Güteindikator - so schlecht er auch unter Umständen ist - übernommen wird. Um dies zu verhindern, können Quellen bei der Zusammenführung ab- oder zugeschaltet werden. Ein Beispiel, bei dem eine Quelle aufgrund ihrer sich als unzutreffend herausgestellten Inhalte nicht berücksichtigt wird, ist in Kapitel 5.1.3 beschrieben. Tabelle 3.2 gibt die Skala der Indikatoren und die dazugehörige Entscheidungsgrundlage an.

3. Konzept zur Verarbeitung der Gebäudeparameter

Tabelle 3.2.: Güteindikator-Skala und Entscheidungsgrundlage für die Einordnung einer Quelle

Indikatorwert	Entscheidungsgrundlage
10	Gesetzlicher Auftrag zur Datenerhebung, Daten als Zahlenwerte oder Bezeichner
7-9	Hochqualitative Datenerhebung, z. T. mit Kontrollen ohne gesetzlichen Auftrag; fehlerfreie Datenverarbeitung im Rahmen dieser Arbeit
3-6	Fehlerbehaftete Datenerhebung bzw. fehlerbehaftete Datenverarbeitung nicht auszuschließen
0-2	Statistische Daten oder aggregierte Informationen ohne genauen Bezug zum Gebäude

3.1.3. Mustererkennung in Textfeldern

Für die Extraktion von Informationen und Zusammenhängen aus Textfeldern ist eine einfache Suche nach einzelnen Wörter meist nicht ausreichend. Der Inhalt von Textfeldern besteht aus natürlicher Sprache. Er weist fehlerhafte Schreibweisen oder unterschiedliche Wörter für einen gleichen bzw. ähnlichen Bedeutungszusammenhang oder syntaktische Variationen auf.

Die in dieser Arbeit untersuchten Texte in Textfeldern bestehen nicht aus ganzen Sätzen, sondern sind knapp gehaltene Angaben im Nominalstil, deren syntaktischer Aufbau relativ regelmäßig ist. Daher bieten sich Verfahren der Mustererkennung an [40]. Das Ziel der Mustererkennung soll an folgendem beispielhaften Textfeld des Datensatzes in Tabelle 3.1 verdeutlicht werden. Das Textfeld beinhaltet die Kurzbeschreibung eines Bauantrags aus den Gebäudekarteikarten der Stadtverwaltung Oldenburg. Das beschriebene Verfahren ist auf andere Muster und Datenquellen übertragbar.

„Neubau eines Einfamilienwohnhauses, Hinw.: 12346-1983 Bauantr. v. 09.05.1983
Neubau v. Garagen“

Für den menschlichen Leser erschließt sich der Zusammenhang schnell. Es handelt sich um den „Neubau eines Gebäudes, das beheizt wird“. Im hinteren Teil wird als Hinweis auf einen anderen Bauantrag verwiesen, der den Neubau von Garagen beinhaltet. Eine manuelle Prüfung dieses Bedeutungszusammenhangs für über 65.000 Datensätze ist jedoch ein unverhältnismäßig großer Aufwand im Rahmen einer kommunalen Wärmeanalyse.

Das Muster kann in drei Aspekte zerlegt werden. Die folgenden Angaben in spitzen Klammern sind sogenannte *Tags* bzw. Markierungen und drücken den Typ der Angabe aus [40]. Die *Tags* stehen zum Teil in einem hierarchischen Verhältnis. Das erste Wort des Musters muss zur Bedeutungsklasse <Neubau> gehören. Nach dem ersten Wort folgen ein oder mehrere Verbindungswörter wie „eines“, „von“ oder „von einem“. Die Mustersyntax variiert somit in einem gewissen Rahmen. Das letzte Wort des Musters muss zur Bedeutungsklasse <Beheiztes Gebäude> gehören. Zusammengesetzt ergibt sich das

Muster „<Neubau> ... <Beheiztes Gebäude>“, wobei die Anzahl der Verbindungswörter „...“ nicht beliebig ist. Es gibt beispielsweise auch Datensätze, deren Syntax folgendem Schema entspricht: „Neubau einer Garage, Umbau eines Mehrfamilienwohnhauses“. In diesem Fall wurde nicht ein Gebäude neu gebaut, sondern nur die Garage. Wenn man eine zu starke Variation der Mustersyntax zulässt, würde dieser Fall trotzdem zum Muster passen und fälschlicherweise erkannt.

Mustererkennung im Bereich *Text Mining* bedeutet, dass in den Texten nach einem spezifischen syntaktischen Ausdruck gesucht wird. Die Vorgehensweise erfolgt nach der verallgemeinert dargestellten Schrittfolge in Abbildung 2.4 aus Kapitel 2.3. Die notwendigen Referenzkorpora werden mithilfe einer manuellen, statistisch unterstützten Analyse auf Basis der zu untersuchenden Textfelder erstellt. Sie listen alle zu einem Oberbegriff bzw. zu einer Bedeutungsklasse gehörenden Wörter [40]. Die Muster werden mithilfe regulärer Ausdrücke und unter Berücksichtigung dieser Referenzkorpora erkannt. Dadurch kann die zu berücksichtigende Variation der Mustersyntax formuliert werden [25].

Erstellen eines Referenzkorpus Ein Referenzkorpus ist ein annotiertes Wörterbuch. Für die Erkennung des oben beschriebenen Beispiels werden die relevanten Wörter mit ihren Oberbegriffen bzw. Bedeutungsklassen <Neubau> und <Beheiztes Gebäude> gelistet.

Mögliche Quellen für einen angepassten Referenzkorpus sind Synonymsammlungen oder Textdatenbanken wie beispielsweise [1]. Diese Quellen berücksichtigen jedoch nicht abweichende Schreibweisen und zum Teil komplexe Komposita (z. B. „Einfamilien-doppelhäuser“). Eine Zerlegung der vorkommenden Komposita ist aufgrund von Falschschreibweisen und abbrechenden Wörtern (z. B. „Einfamilienh“) nur mit großer manueller Nacharbeit möglich. Daher ist die Verwendung externer Referenzkorpora für diesen Anwendungsfall ungeeignet.

Der Referenzkorpus wird aus den zu untersuchenden Texten selbst mithilfe des Zipf'schen Gesetzes erstellt [40]. Hierfür werden die Wörter der zu untersuchenden Texte absteigend nach ihrer Häufigkeit sortiert. Ein bestimmtes Wort hat dann die Position n . Laut dem Zipf'schen Gesetz ist die Wahrscheinlichkeit, dass dieses Wort im Text auftritt gleich $1/n$, d. h. der reziproke Wert seiner Position in der Häufigkeitsliste.

Mithilfe eines *Tokenizers* werden die Texte in einzelne Wörter zerlegt [15]. Ein Wort ist hierbei definiert als eine Gruppe von beieinander stehenden Buchstaben des deutschen Alphabets in Groß- und Kleinschreibung. Die Häufigkeitsverteilung der Wörter in den Gebäudekartekarten der Stadt Oldenburg zeigt Abbildung 3.2. Die fünf häufigsten Wörter sind in der Abbildung mit ihrer Position angegeben. Es kommen nur ca. 200 Wörter häufiger als 100 mal vor. Die Kurve des Zipf'schen Gesetzes ist gestrichelt eingezeichnet. Zwischen beiden Kurven erkennt man eine grundsätzliche Übereinstimmung.

Auf Basis dieser Wortliste ist es einfach, manuell alle relevanten Einträge mit einer Häufigkeit > 80 zu einer Bedeutungsklasse zuzuordnen. Wörter wie das dritthäufigste vorkommende Wort „Band“ können ignoriert werden, da sie für die Mustererkennung

3. Konzept zur Verarbeitung der Gebäudeparameter

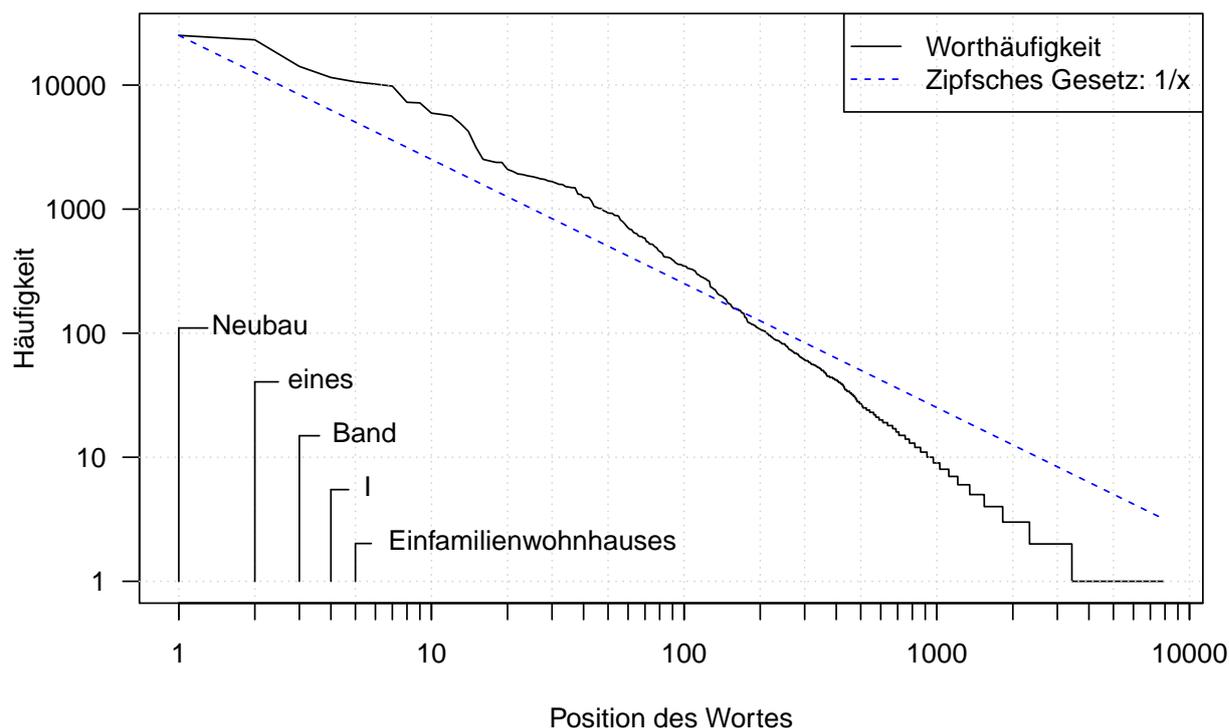


Abbildung 3.2.: Häufigkeitsverteilung der Wörter in den Textfeldern der Gebäudekartenkarten

von Neubauten nicht relevant sind. Nach der Markierung relevanter Einträge mit einer Häufigkeit > 80 wird die Wortliste in eine alphabetische Reihenfolge umsortiert. Falschschreibweisen und alternative Ausdrücke gruppieren sich dadurch rund um die häufig vorkommenden und bereits markierten Wörter. Beispielsweise liegt das nur 33 Mal vorkommende Wort „Betriebsgebäude“ direkt neben dem 403 mal vorkommenden Wort „Betriebsgebäudes“.

Der Zeitaufwand für den manuellen Aufbau eines Referenzkorpus, gestützt durch die Analyse der Häufigkeitsliste liegt bei maximal einer Stunde und erlaubt die Analyse von Tausenden von Datensätzen. Zum Oberbegriff <Neubau> werden 38 Begriffe (z. B. „Neub.“, „Nb.“, „Errichtung“, „Neuerstellung“, etc.) gefunden. Der Oberbegriff <Beheizte Gebäude> umfasst knapp 300 Begriffe.

Mustererkennung mit Regulären Ausdrücken Das Muster „<Neubau> ... <Beheiztes Gebäude>“ des beschriebenen Beispiels wird technisch durch den folgenden Regulären Ausdruck: „(<Neubau>)([^]*){0,2} (<Beheiztes Gebäude>)“ gebildet [25]. Der mittlere Teil „([^]*){0,2}“ beschreibt, dass maximal zwei („{0,2}“) jeweils durch ein Leerzeichen getrennte Zeichenketten („[^]*“) zugelassen sind. Die Tags werden durch die Einträge des Referenzkorpus und einem ODER-Verbindungszeichen ersetzt.

Der dadurch entstehende Reguläre Ausdruck ist klein genug, um Laufzeitschwierigkeiten zu vermeiden. Die Erkennung der Muster in ca. 65.000 Datensätzen dauert nur wenige

Sekunden. Es werden insgesamt 8730 Datensätze mit diesem Muster erkannt. Das Muster kommt in 390 Variationen vor. Die Variation „Neubau eines Einfamilienwohnhauses“ tritt am häufigsten auf. Es werden jedoch auch Einträge wie „Neubau DHH“ erkannt.

3.2. Modellierung und Verwendung von Raumbezügen

Die Bezeichnung „Raumbezug“ wird in dieser Arbeit im Sinne eines räumlichen Verweises verwendet. Informationen werden mit Orten auf der Erdoberfläche verknüpft [7]. Er wird entweder direkt durch eine räumliche Geometrie angegeben oder verweist indirekt anhand einer kodierten Bezeichnung (ID) oder einer textlichen Lagebezeichnung, z. B. Adresse auf eine räumliche Geometrie.

Das Datenmodell der Raumbezüge bietet eine feste Verknüpfung zwischen den drei Formen einer Raumbezugsangabe. Darüber hinaus verbindet es unterschiedliche Raumbezüge, die in einem logischen Zusammenhang zueinander stehen. Beispielsweise wird eine Adresse mit dem zugehörigen Gebäude durch eine Fremdschlüsselverbindung verknüpft.

Viele Gebäudedatenquellen beinhalten keine Koordinaten, sondern geben den Raumbezug anhand einer Lagebezeichnung an. Das Datenmodell der Raumbezüge und entsprechende Bearbeitungsroutinen ermöglichen den Vergleich dieser Lagebezeichnung mit den in der Datenbank gespeicherten Raumbezügen. Beim Import der Gebäudedaten wird die externe Lagebezeichnung mit dem passenden internen Raumbezug ersetzt. Dies ist die Grundlage, um die importierten Gebäudedaten einander zuzuordnen und miteinander zu vergleichen.

3.2.1. Aufbau eines Datenmodells der Raumbezüge

Im Datenmodell der Raumbezüge werden fünf Raumbezugsklassen unterschieden. Abbildung 3.3 visualisiert die Raumbezugsklassen anhand eines beispielhaften Kartenausschnitts. Die Gebäude befinden sich in Flurstücken, wobei sie unterschieden werden in Hauptgebäude, die einen oder mehrere Adresspunkte innerhalb des Gebäudegrundrisses aufweisen, und Nebengebäude ohne Adresspunkte. Fast alle Hauptgebäude, jedoch auch einige Nebengebäude haben Zugang zu einem oder mehreren Versorgungsanschlüssen. Die Gebiete werden angedeutet durch die lila Begrenzungslinien, die die „Blockseiten“ der kleinräumigen Gliederung der Stadt Oldenburg darstellen. Die Raumbezugsklassen sind folgendermaßen zu verstehen:

- **Gebiete** untergliedern das kommunale Gesamtgebiet. Die Summe der über die zu untersuchende Region verteilten Gebiete ergeben eine sogenannte Gliederungsebene und sind räumlich disjunkt. Gebiete verschiedener Gliederungsebenen können sich überlagern. Zu den Gliederungsebenen gehören z. B. Postleitzahlbereiche, Bebauungsplangebiete, Rasterzellen und statistische Bezirke der Stadtverwaltung.

3. Konzept zur Verarbeitung der Gebäudeparameter



Abbildung 3.3.: Kartenausschnitt zur Darstellung der fünf Raumbezugsklassen in Oldenburg

- **Flurstücke** sind aufgrund der zusätzlichen Attribute und ihrer Bedeutung bezüglich der Eigentumsverhältnisse eine Sonderform räumlicher Gliederungsebenen und werden gesondert behandelt.
- **Adresspunkte** geben die Lage einer Adresse an.
- **Versorgungsanschlüsse** für Strom, Gas oder Wärme geben die Lage des Übergabepunkts vom Versorgungsnetz in ein Gebäude an. Falls keine Georeferenz bekannt ist, wird die Lage mit Hilfe der angegebenen Adresse geschätzt.
- **Gebäude** beschreiben die Lage und Form der Gebäudegrundflächen. In dieser Arbeit ist ein Gebäude definiert anhand der von ihm bekannten Gebäudegrundfläche.

Modell der Raumbezüge Aus dem semantischen Zusammenhang der Raumbezugsklassen ergibt sich das in Abbildung 3.4 dargestellte Datenmodell der Raumbezüge. Jede Tabelle repräsentiert eine Raumbezugsklasse. Gemeinsam ist allen Klassen der grundlegende Aufbau. Jede Klasse verfügt über

- ein Primärschlüsselattribut („...ID“), das jeden Datensatz eindeutig datenbankintern identifiziert,
- ein oder mehrere Attribute für die kodierte Bezeichnung oder textliche Lagebezeichnung der Raumbezüge und
- ein Attribut, in dem die Raumgeometrie gespeichert ist.

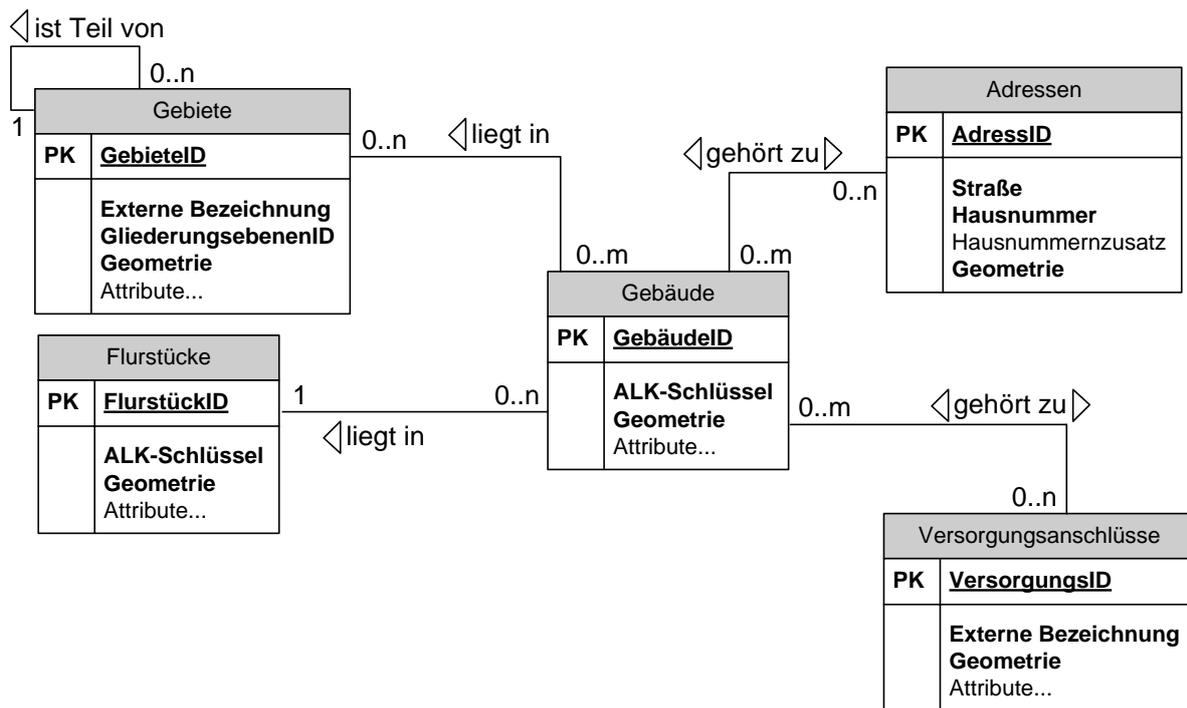


Abbildung 3.4.: Datenmodell der Raumbezugsklassen

Hinzu kommen weitere klassenspezifische Attribute. Die Form der externen Bezeichnung variiert. Bei den Adressen besteht sie aus Straße, Hausnummer und Hausnummernzusatz, bei den Versorgungsanschlüssen aus einer ID-Nummer. Die Gebäude und Flurstücke werden durch die jeweiligen ID-Nummern der Automatisierte Liegenschaftskarte (ALK) identifiziert. Die Eindeutigkeit der externen Bezeichnungen ist nicht durch die Datenbanksyntax sicher gestellt. Mögliche daraus entstehende Probleme bei der Datenzuordnung können nicht im Rahmen des Datenbankmodells geklärt werden, sondern bedürfen einer inhaltlichen Kontrolle der Raumbezüge.

Darüber hinaus gibt die Abbildung 3.4 die Art der Relation zwischen den Raumbezugsklassen an. Bei Versorgungsanschlüssen kann es vorkommen, dass entweder mehrere Gebäude sich einen Versorgungsanschluss teilen oder mehrere Versorgungsanschlüsse auf nur ein Gebäude verweisen. Dies wird dargestellt als $n : m$ -Verbindung, wobei n bzw. m Werte größer gleich Null ($n, m \geq 0$) annehmen. Analog gilt dies für Adressen. Bei den größeren Gebäuden in der Karte von Abbildung 3.3 ist sichtbar, dass einem Gebäudegrundriss mehrere Adressen zugeordnet sind. Für die Umsetzung von $n : m$ -Beziehungen zwischen Tabellen ist jeweils eine dritte Zuordnungstabelle notwendig. Diese sind im abgebildeten Raumbezugsmodell nicht dargestellt, um die wesentlichen Aspekte hervorzuheben.

Eine Besonderheit ergibt sich im Verhältnis zwischen Gebäuden und Gebieten. Ein Gebäude gehört zu genau einem Gebiet einer Gliederungsebene. In der Tabelle „Gebiete“ sind jedoch die Gebiete verschiedener Gliederungsebenen gespeichert. Dadurch steht ein Gebäude mit mehreren Gebieten unterschiedlicher Ebenen, d. h. mehreren Datensätzen der Tabelle „Gebiete“ im Zusammenhang. Zugleich gehören zu jedem Gebiet mehrere

3. Konzept zur Verarbeitung der Gebäudeparameter

Gebäude, womit sich eine $n : m$ -Verbindung zwischen den Tabellen ergibt. Die $1 : n$ -Verbindung der Tabelle „Gebiete“ mit sich selbst visualisiert die Zuordnung eines Gebiets zu dem Gebiet der nächst höheren Gliederungsebene bei hierarchisch strukturierten Ebenen. Ein Beispiel hierfür sind Ebenen der kleinräumigen Gliederung der städtischen Verwaltung. Deren unterste Ebene umfasst die in Abbildung 3.3 dargestellten Blockseiten. Mehrere Blockseiten werden zu einem Block zusammengefasst, etc.

Die feste Verknüpfung von Datensätzen ermöglicht manuelle Anpassungen von beispielsweise fehlerhaft oder unzureichend verknüpften Raumbezügen. Nach der erstmaligen Verknüpfung sind rechenintensive Operationen der räumlichen Verschneidung nicht mehr notwendig. Dies führt zu einer schnellen Datenverarbeitung bei der Analyse und Abfrage der Daten.

Aufbau der Raumbezüge für das Untersuchungsgebiet Oldenburg Die Raumbezüge für Oldenburg entstammen größtenteils der ALK. Die Pflege der ALK ist eine hoheitliche Aufgabe der Vermessungsämter. Die Daten umfassen den Gebäudebestand, die Flurstücke und die Adressen vollständig und korrekt. Damit ist die Voraussetzung von möglichst vollständigen Raumbezügen gegeben. Die Angaben zu den Gebieten werden durch das Amt für Geoinformation und Statistik der Stadt Oldenburg bereit gestellt.

Die Verbindung zwischen Raumbezügen unterschiedlicher Raumbezugsklassen wird für das Untersuchungsgebiet Oldenburg aufgrund der Lage der Geometrien berechnet. Das bedeutet, dass zwei sich räumlich überlagernde Geometrien einander zugeordnet werden. Es kann vorkommen, dass dadurch die Realität bei der Verbindung von Adressen und Gebäuden bzw. bei Versorgungsanschlüssen und Gebäuden nicht vollständig abgebildet wird. Im Falle von Adressen werden diese als Punktgeometrien repräsentiert und können dadurch nur innerhalb eines Gebäudegrundrisses liegen. Das entsprechende Gebäude wird in dem Fall als Hauptgebäude gekennzeichnet. Angrenzende oder benachbarte Gebäude ohne Adresse sind Nebengebäude. Gebäudeparameter, die mit einer Adresse versehen sind, werden beim Import nur auf das Hauptgebäude bezogen, nicht auf das Nebengebäude. Falls die Informationen sich auf beide Gebäude beziehen oder unter Umständen sogar nur auf das Nebengebäude, findet eine Falschzuordnung statt. Das kann aufgrund der verfügbaren Daten jedoch nicht verhindert werden. Diese Tatsache wird bereits in der Diplomarbeit von Christian Linder thematisiert [52]. Es führt außerdem dazu, dass im Untersuchungsgebiet die Verbindung zwischen den Adressen und Gebäuden keine $n : m$ -Kardinalität, sondern eine $1 : m$ -Kardinalität aufweist. Eine Adresse bezieht sich auf maximal ein Gebäude.

Die vom Netzbetreiber erhaltenen Versorgungsanschlussdaten weisen keine gültigen Geometrien auf. Daher werden Gebäude und Anschlüsse anhand der mitgelieferten Adresse einander zugeordnet. Der Umgang mit unterschiedlichen Adressschreibweisen ist in Kapitel 3.2.3 beschrieben. Aufgrund der vereinfachten Zuordnung von Gebäuden und Adressen im Untersuchungsgebiet besteht für die Verbindung zwischen Gebäuden und Versorgungsanschlüssen ebenfalls nur eine $1 : m$ -Kardinalität.

Gebäude, Gebiete und Flurstücke haben zweidimensionale Polygone als Geometrien. Da es vorkommen kann, dass ein Gebäudepolygon die Grenze zwischen zwei Gebiets- oder

Flurstückpolygonen schneidet, wird der Schwerpunkt des Gebäudepolygons berechnet und dieser nulldimensionale Punkt mit den Flurstückspolygonen verglichen [44].

Vor dem Import von Raumbezügen in das Datenmodell werden deren Geometrien überprüft. Dabei müssen folgende Bedingungen erfüllt sein:

- Die Geometrien der Raumbezüge dürfen sich nicht überlappen. Sollte dies doch der Fall sein, müssen die Daten vor dem Import bereinigt werden. Sonst würde das gleiche Gebäude mehrfach in den Daten geführt werden, woraus Fehler bei der Analyse resultieren. Die verwendeten ALK-Daten enthalten ca. 100 einander überlagernde Gebäude. Von diesen wird immer nur eines importiert.
- Die Dimension der Geometrie und das Koordinatenbezugssystem der zu importierenden Raumbezügen müssen übereinstimmen. Für Gebäude und Gebiet ist dies ein Polygon, für Adressen und Versorgungsanschlüsse ein Punkt.

Tabelle 3.3 gibt die Zahl der Datensätze für die drei wichtigsten Raumbezugsklassen Gebäude, Adresse und Versorgungsanschluss auf der Matrixdiagonalen an. Die Anzahl erfolgreicher Zuordnungen zwischen jeweils zwei unterschiedlichen Klassen ist aus den nicht-diagonalen Matrixelementen abzulesen. Circa 50 % bzw. 44.676 Gebäude sind Hauptgebäude und verfügen über eine Adresse. Nahezu alle Hauptgebäude sind folglich mit einer Adresse bzw. einem Versorgungsanschluss verbunden

Tabelle 3.3.: Raumbezüge und Zuordnungshäufigkeiten in Oldenburg

	Gebäude	Adresse	Versorgung
Gebäude	87.389	44.676	44.487
Adresse		44.678	40.716
Versorgung			46.421

3.2.2. Verwendung der Raumbezüge zum Import von Primärparametern

Beim Import von Primärparametern wird der ursprüngliche Raumbezug - im Folgenden als **externer Raumbezug** bezeichnet - durch einen **internen Raumbezug** ersetzt. Das heißt, jedem Datensatz werden zwei Attribute hinzugefügt: die Raumbezugsklasse und die ID des internen Raumbezugs, auf den sich der Datensatz bezieht.

Der Ablauf des Imports hängt von der Form der Raumbezugsangabe ab. Tabelle 3.4 gibt für jede in Oldenburg verwendete Datenquelle die Raumbezugsklasse und die Form des Raumbezugs an. Folgende Vorgehensweisen werden angewendet:

- Beim **Adressvergleich** werden die internen Adressangaben und externen Adressangaben schreibweisenunabhängig verglichen. Dies ist vertieft im folgenden Kapitel 3.2.3 beschrieben.

3. Konzept zur Verarbeitung der Gebäudeparameter

Tabelle 3.4.: Raumbezüge der Datenquellen in Oldenburg

Quelle	Raumbezugs- klasse	Form der Raum- bezugsangabe	Umgang mit dem Raumbezug
Archiv (Bauordnung)	Gebäude	ID	ID-Abgleich
Bauanträge (Bauordnung)	Adresse	Lagebezeichnung	Adressvergleich
Gebäudekartei- karten (Bauordnung)	Adresse	Lagebezeichnung	Adressvergleich
Automatisierte Liegenschaftskarte	Gebäude	Geometrie	Geometrie ist bereits importiert
Laserscan	Gebäude	Geometrie	Räumlicher Vergleich mit internen Gebäudepolygonen
Manuelle Luftbildererkennung	Gebäude	Geometrie	Räumlicher Vergleich mit internen Gebäudepolygonen
Kaufpreissammlung	Adresse	Lagebezeichnung	Adressvergleich
Oldenburger Häuserbuch	Adresse	Lagebezeichnung	Adressvergleich
Stromnetzbetreiber	Versorgungs- anschluss	Lagebezeichnung	Adressvergleich
Bebauungspläne	Gebiete	Geometrie	Import der Geometrie
Einwohner- meldeamt	Adresse	Lagebezeichnung	Adressvergleich

- Der **ID-Abgleich** wird durchgeführt zwischen den externen und internen kodierten Bezeichnungen (IDs) der entsprechenden Raumbezugsklassen. Beispielsweise ist jeder Datensatz der ALK mit einer siebenstelligen Hexadezimalzahl (z. B. 002J3N9) eindeutig bezeichnet, die eine direkte Zuordnung verschiedener Datensätze ermöglicht.
- Der **räumliche Vergleich mit internen Gebäudepolygonen** erfolgt analog zu der oben beschriebenen Zuordnung von zwei sich überlagernden Geometrien.
- Beim **Import der Geometrie** werden die Geometrien der externen Datenquelle in das Modell der Raumbezüge importiert und stehen dann als interne Raumbezüge zur Verfügung. Das ist beispielsweise der Fall bei den Bebauungsplan-Gebieten zu. Das Datenmodell der Raumbezüge ist so konzipiert, dass ein Import von neuen Gebieten möglich ist, ohne dass die Struktur des Modells geändert werden muss.

3.2.3. Verarbeitung der Adresse als Raumbezug externer Daten

Die Adresse ist ein besonderer Raumbezug, da die Angabe semantisch komplex und in der Alltagssprache verwurzelt ist. Dadurch entstehen Ungenauigkeiten und kleine Variationen, die zwar das Verständnis eines menschlichen Lesers nicht behindern, jedoch eine algorithmisierte Verarbeitung erschweren. Hierzu gehören folgende Abweichungen:

- Schreibfehler oder unterschiedliche Schreibweisen beim Straßennamen
- Abkürzungen von Straßennamen, z. B. „Str.“
- Angabe von Adressbereichen wie beispielsweise „Hauptstr. 1-5“
- Umfang der Adressangabe, wie der Name der Stadt, die Postleitzahl oder die Namen von berühmten Gebäuden, z. B. „Schloss“
- Zusatzangaben wie „Nebengebäude“

Zugleich ist die Adresse ein sehr wichtiger Raumbezug, da viele Datenquellen für diese Arbeit keine anderen Bezüge als die Adresse aufweisen. Hierzu gehören vor allem historische Informationen, die den Gebäudeparameter Baujahr betreffen.

Ein erster Ansatz für den Vergleich unterschiedlich geschriebener Adressen wurde im Rahmen der durch mich betreuten Studienarbeit „Energieeffiziente Stadt Oldenburg“ entwickelt [44]. Darin wird für Straßennamen eine Straßennamen-Zuordnung auf Basis eines Matching-Algorithmus beschrieben. Der Ansatz muss für jede Datenquelle manuell angepasst werden und erfordert einen hohen Aufwand, wenn Falschschreibweisen oder komplexere Veränderungen der Straßennamen berücksichtigt werden sollen.

Dieser erste Ansatz wurde im Rahmen dieser Arbeit weiterentwickelt und stärker automatisiert. Durch die Neuerungen ist er übertragbar auf neue Datenquellen oder Untersuchungsgebiete. Das entwickelte Verfahren ermöglicht es, dass Adressangaben externer Datensätze zu den in der Tabelle „Adressen“ gespeicherten internen Adressangaben automatisiert zugeordnet werden. Bei der Zuordnung ist es möglich, die Ergebnisse manuell

3. Konzept zur Verarbeitung der Gebäudeparameter

zu kontrollieren und unter Umständen zu korrigieren. Eine erfolgreiche Zuordnung der externen und internen Adressangabe erlaubt die Integration und Verarbeitung der mit der Adresse verknüpften Gebäudeparameter-Angaben auf Basis des vorgestellten Modells der Raumbezüge.

Die Struktur von Adressbereichen, die mehrere Einzeladressen umfassen, wird in Tabelle 3.5 angegeben. Angaben in spitzen Klammern sind sogenannte *Tags* bzw. Markierungen und drücken den Typ der Angabe aus [40]. Die *Tags* stehen in einem hierarchischen Verhältnis.

Tabelle 3.5.: Hierarchie der für die Adresszuordnung relevanten *Tags*

Tags	Beispiel
<Adressbereichsangabe>	Hauptstr. 5-9
- <Straßenname>	Hauptstr.
- <Straßenabschnitt>	5-9
<Adressangabe>	Hauptstraße 5a
- <Straßenname>	Hauptstraße
- <Hsnr>	5
- (<HsnrZs>)	a

Die „<Adressbereichsangabe>“ beinhaltet den „<Straßenname>“ und „<Straßenabschnitt>“. Da sich die Arbeit auf das Stadtgebiet Oldenburg beschränkt, werden der Ortsname oder Postleitzahlangaben nicht berücksichtigt. Der *Tag* „<Straßenabschnitt>“ beinhaltet sowohl eine einzelne Hausnummer (z. B. 3), zum Teil mit Hausnummernzusatz (z. B. 3a) als auch einen Bereich von Hausnummern mit oder ohne Hausnummernzusätzen (z. B. 3a-c). Eine einzelne <Adressangabe> besteht immer aus einem Straßennamen (<Straßenname>), einer Hausnummer (<Hsnr>) und optional einem Hausnummerzusatz (<HsnrZs>).

Die Verarbeitung der <Adressbereichsangabe> erfolgt in fünf Schritten:

1. Aus dem angegebenen Raumbezug wird die <Adressbereichsangabe> extrahiert und die Ortsangabe, Postleitzahl, etc. entfernt. Aus „Oldenburg, Hauptstr. 5-9“ wird „Hauptstr. 5-9“.
2. Daraufhin wird diese in <Straßenname> und <Straßenabschnitt> getrennt. Trennung von „Hauptstr. 5-9“ in „Hauptstr.“ und „5-9“.
3. Der <Straßenname> wird mit den in der Datenbank gespeicherten Angaben <Straßenname> verglichen, um den internen Namen mit der größten Übereinstimmung zu finden. Die größte Übereinstimmung mit „Hauptstr.“ hat in diesem Fall der interne Straßenname „Hauptstraße“.
4. Der <Straßenabschnitt> wird aufgetrennt in eine Liste von <Hsnr> mit optionalem <HsnrZs>. „5-9“ schließt im Fall von Oldenburg die Hausnummern 5, 7 und 9 ein.

5. Zuletzt wird geprüft, ob die aus der <Adressbereichsangabe> extrahierten und aufbereiteten Inhalte der <Adressangabe> in der Datenbank existieren. Es wird getestet, ob die aufgeteilten Adressangaben „Hauptstraße 5“, „Hauptstraße 7“ und „Hauptstraße 9“ als interne Adressen existieren.

Die Anforderungen an Schritt 1 variieren für unterschiedliche Datenquellen und müssen jeweils angepasst werden. Eine Herausforderung stellen die Schritte 3 und 4 dar. Sie werden im Folgenden näher beschrieben. Bei einer erfolgreichen Prüfung im letzten und 5. Schritt wird in der Datenbank ein Verweis zwischen der ID des externen Datensatzes, zu dem der Raumbezug gehört, und der internen Adress-ID gespeichert.

Schreibweisenunabhängiger Vergleich der Straßennamen Straßennamen können durch Schreibfehler, Abkürzungen oder das Weglassen von Bindestrichen verändert werden. Meist ist diese Veränderung nicht sehr umfangreich, führt jedoch dazu, dass ein direkter Vergleich nicht erfolgreich ist. Der Straßename ist eine Zeichenkette und besteht aus mindestens einem Wort.

Die Liste der in der Datenbank gespeicherten Straßennamen ist der Referenzkorpus für die Analyse (siehe Kapitel 2.3). Diese Straßennamen werden im Folgenden als **interne** Straßennamen I bezeichnet. Die Straßennamen der zu importierenden Daten werden im Folgenden als **externe** Straßennamen X bezeichnet.

Für den Vergleich interner und externer Straßennamen werden beide Zeichenketten in Abschnitte der Länge n aufgeteilt. Die Abschnitte werden als N-gramme bezeichnet. Aufeinanderfolgende N-gramme überlappen sich. Beispielsweise für $n = 3$, sogenannte Trigramme, wird das Wort „Energie“ in folgende N-gramme zerlegt: „_E, _En, ene, ner, erg, rgi, gie, ie_, e_“, wobei „_“ für ein Leerzeichen steht.

Der Quotient aus der Zahl der übereinstimmenden N-gramme, geteilt durch die gesamte Anzahl der N-gramme beider Zeichenketten X und I gibt ein Maß $\phi(X, I)$ für den Grad der Übereinstimmung an. Der Faktor 2 in Gleichung (3.1) dient der Normierung von $\phi(X, I)$ auf den Wertebereich $[0, 1]$.

$$\phi(X, I) = \frac{2 \cdot |\text{n-grams}(X) \cap \text{n-grams}(I)|}{|\text{n-grams}(X)| + |\text{n-grams}(I)|} \quad (3.1)$$

Eine genaue Übereinstimmung ergibt $\phi(X, I) = 1$, keine Übereinstimmung $\phi(X, I) = 0$ [48, 15].

Das Maß $\phi(X, I)$ wird für jede Kombination aus einem externen Straßennamen X und jedem in der Datenbank gespeicherten internen Straßennamen I berechnet. Als Ergebnis wird dem externen Straßennamen X derjenige interne I zugeordnet, für welchen $\phi(X, I)$ maximal ist. Ein unterer Schwellenwert für das Maß $\phi(X, I)$ ist nicht definiert, da sich ein solcher nicht sinnvoll festlegen lässt. Aufgrund der fehlenden Untergrenze findet der Algorithmus in jedem Fall einen internen Straßennamen I für X . Die Voraussetzung für eine erfolgreiche automatisierte Zuordnung ist daher, dass für ein Untersuchungsgebiet alle Straßennamen in der Datenbank gespeichert sind. Sonst wird die nach obigem Prinzip am ähnlichsten geschriebene Straßenbezeichnung I ausgewählt.

3. Konzept zur Verarbeitung der Gebäudeparameter

Um fehlerhafte Zuordnungen zu vermeiden, sind die folgenden manuellen Kontroll- und Korrekturmöglichkeiten in dem Verfahren implementiert:

- Nach der Zuordnung eines internen Straßennamens für alle externen Zeichenketten wird die Zuordnungshäufigkeit überprüft. Wenn unterschiedliche externe Zeichenketten X der gleichen internen Zeichenkette I zugeordnet werden, erfolgt eine Ausgabe dieser Fälle zur manuellen Kontrolle.
- Darüber hinaus wird die Zuordnung der Straßennamen alphabetisch geordnet in einer Textdatei ausgegeben. Diese Textdatei ist die Basis für den Vergleich der gesamten Adressen. Da jeder externe Straßename nur einmal angegeben wird, ist die Liste kürzer und übersichtlicher als die originalen Ortsbezüge und eine manuelle Korrektur einfacher möglich.

Trennung von Straßenabschnitten Straßenabschnitte beinhalten mindestens eine Hausnummer und sind sehr heterogen aufgebaut. Bei der Angabe von Hausnummern-Bereichen wie beispielsweise „5-9“ ist das Hausnummern-Nummerierungssystem der Kommune ausschlaggebend. In Oldenburg befinden sich alle ungeraden Hausnummern auf einer Straßenseite, alle geraden auf der gegenüberliegenden. Die Angabe „5-9“ umfasst daher nur die Hausnummern 5, 7 und 9, nicht jedoch die Nummern 6 und 8.

Tabelle 3.6 beschreibt die vorkommenden Kombinationen der Tags $\langle \text{Hsnr} \rangle$ und $\langle \text{HsnrZs} \rangle$, aus denen ein $\langle \text{Straßenabschnitt} \rangle$ aufgebaut sein kann. Für jede Kombination ist die Trennregel kurz formuliert und mit einem Beispiel verdeutlicht. „-, +/“ sind die vorkommenden Verbindungszeichen.

Tabelle 3.6.: Regeln für die Trennung von Straßenabschnitten

Tag-Kombination	Trennregel	Beispiel
$\langle \text{Hsnr} \rangle - \langle \text{Hsnr} \rangle$	Trennung des Straßenabschnitts, wobei nur ungerade oder gerade Nummern berücksichtigt werden	1-5 \rightarrow 1,3,5 bzw. 2-4 \rightarrow 2,4
$\langle \text{Hsnr} \rangle \langle \text{HsnrZs} \rangle - \langle \text{HsnrZs} \rangle$	Trennung der Hausnummernzusätze in alle dazwischenliegenden Buchstaben	1a-c \rightarrow 1a, 1b, 1c
$\langle \text{Hsnr} \rangle + / , \langle \text{Hsnr} \rangle$	Trennung des Abschnitts in die zwei angegebenen Hausnummern	1+3 bzw. 1/3 bzw. 1,3 \rightarrow 1, 3
$\langle \text{Hsnr} \rangle \langle \text{HsnrZs} \rangle + / , (\langle \text{Hsnr} \rangle) \langle \text{HsnrZs} \rangle$	Trennung des Abschnitts in die zwei angegebenen Hausnummerzusätze	1a + (1)c \rightarrow 1a, 1c
$\langle \text{Hsnr} \rangle + / , \langle \text{Hsnr} \rangle \langle \text{HsnrZs} \rangle$	Trennung des Abschnitts in die Hausnummer und Hausnummer mit Zusatz	1 + 1a \rightarrow 1, 1a
$\langle \text{Hsnr} \rangle , \langle \text{Hsnr} \rangle , \langle \text{Hsnr} \rangle , \dots$	Trennung der Aufzählung von Hausnummern	1,2,4,6 \rightarrow 1, 2, 4, 6

3.3. Primärparameterverwaltung

Das Ziel der Primärparameterverwaltung ist, den Strukturreichtum der aus Quellen erhaltenen Informationen für eine systematische Weiterverarbeitung zu erfassen. Die Daten variieren in vier Aspekten. Sie

- stammen aus verschiedenen Quellen,
- stehen im Zusammenhang mit unterschiedlichen Raumbezugsklassen,
- sind zu abweichenden Zeitpunkten aufgenommen und
- betreffen verschiedene Gebäudeparameter.

Diese Aspekte werden in Metadatenattributen gespeichert. Sie tragen keine inhaltlichen Informationen, sondern dienen der systeminternen Verwaltung [7]. Die inhaltlichen Informationen werden Sachdaten genannt und unten beschrieben. Beispielsweise gehören zu dem Sachdatum „1963“ die Metadaten „Gebäudekarteikarten“ als Quellenangabe, „Adresse: Nadorsterstr. 122 b“ als Angabe des Raumbezugs, „1992“ als Zeitpunkt der Datenaufnahme bzw. Datensatzerstellung und „Baujahr“ als Parametername. Die Logik des Datenmodells der Metadaten fordert, dass die Kombination der vier Aspekte bzw. Metadaten eindeutig ist. Diese Tatsache ist wichtig für die Strukturierung des Datenmodells und der darauf aufbauenden Zusammenführung der Daten.

Die Primärparameterverwaltung speichert die Gebäudeparameter mit dem originären Raumbezug. Somit muss der Datenimport nicht wiederholt werden, falls die Zuordnung der Raumbezüge korrigiert wird. Die Gebäudeparameter werden erst bei der Zusammenführung der Daten von ihrem originären Raumbezug auf das Gebäude übertragen. Der Import der Primärparameter und die Bearbeitung und Verknüpfung der Raumbezüge kann somit parallel geschehen. Darüber hinaus werden Analysen der Primärparameter erleichtert und Scheingenauigkeiten vermieden.

Neben den vier Metadatenattributen umfassen die verschiedenen Gebäudeparameter unterschiedlich viele Einzelangaben. Ein Beispiel für die auftretende Vielfalt und die damit verbundene Problematik bietet Abbildung 3.5(a). Alle fünf Datensätze beziehen sich auf eines oder beide Gebäude der Abbildung 3.5(b). Die Raumbezüge sind so dargestellt, wie sie von den Datenquellen bereit gestellt werden. Anschlüsse, Gebäude und Bebauungspläne werden jeweils über eine ID adressiert. Der Gebäudeparameter Nutzung beinhaltet drei Angaben, alle anderen jeweils eine Angabe. Die Metadaten für Parameter, Quelle und Raumbezug sind angegeben, der Zeitpunkt der Datenaufnahme fehlt.

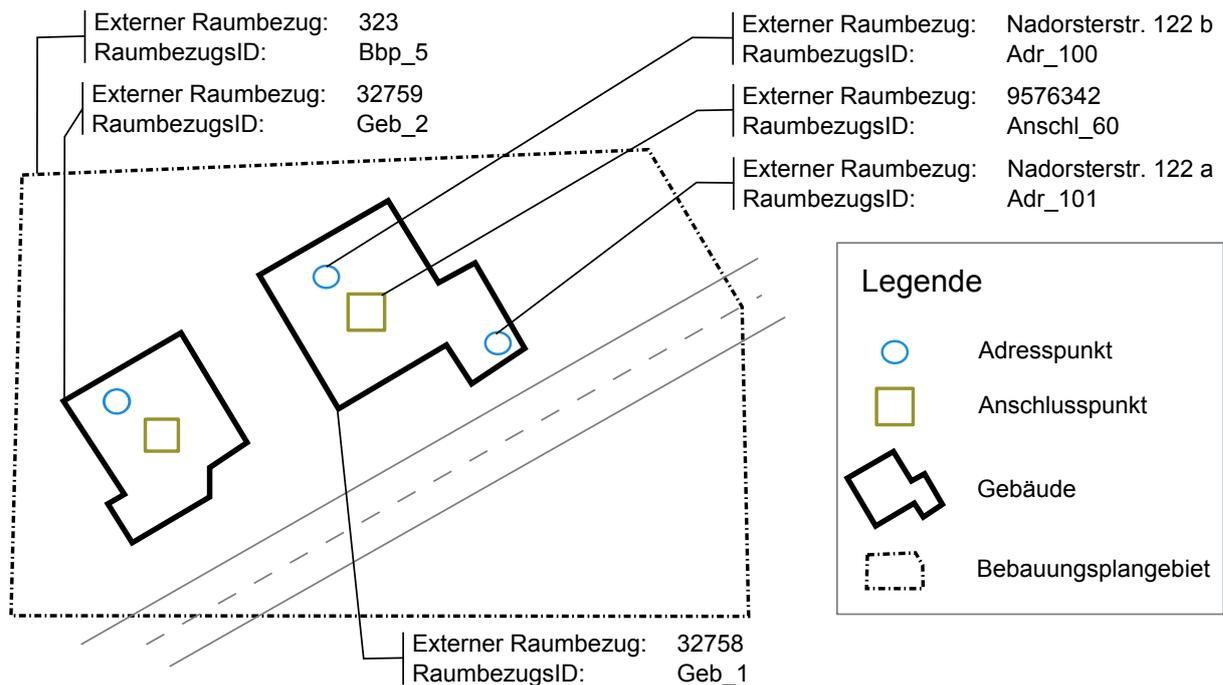
3.3.1. Datenmodell der Primärparameterverwaltung

Beim Start eines Datenerfassungsprojekts ist meist nicht vollständig bekannt, welche Informationen zur Verfügung stehen werden. An das Datenmodell wird daher die Anforderung gestellt, dass für jede der vier Metadatenattribute möglichst einfach neue Werte

3. Konzept zur Verarbeitung der Gebäudeparameter

Datensatz	Externer Raumbezug	Parameter	Angabe	Quelle
1	Adresse: Nadorsterstr. 122 b	Baujahr	1963	Gebäudekarteikarten
2	Adresse: Nadorsterstr. 122 a	Baujahr	1965	Gebäudekarteikarten
3	Anschluss: 9576342	Nutzung	3 Einheiten Wohnen 1 Einheit Einzelhandel 1 Einheit Gewerbe	EWE Netz Anschlüsse
4	Gebäude: 32758	Dachform	"Flachdach"	Gebäudekarteikarten
5	Bebauungsplan: 323	Baujahr	1980	Bebauungspläne

(a) Gebäudeparameter unterschiedlicher Quellen



(b) Karte mit zwei Gebäuden zur Visualisierung der Gebäudeparameter-Raumbezüge

Abbildung 3.5.: Beispiel für den Strukturreichtum von Primärparameterangaben verschiedener Quellen für zwei Gebäude

hinzugefügt werden können. Es könnte z. B. sein, dass vorher nicht behandelte Daten über den Gebäudeparameter „Alter des Heizsystems“ berücksichtigt werden sollen. Das Datenmodell ist so konzipiert, dass die Einbindung neuer Informationen keine strukturellen Änderungen erzwingen, sondern in der vorgegebenen Datenstruktur möglich bleibt.

Metadaten Abbildung 3.6 zeigt das Datenmodell der Primärparameterverwaltung als Entity-Relationship-Modell. Die Tabellen können in die zwei Bereiche Meta- und Sachdaten eingeteilt werden. Die zentrale Tabelle ist die Tabelle Metadaten. Sie enthält das eindeutige Primärschlüsselattribut „MetadatenID“ und die Metadatenattribute. Der Raumbezug wird auf zwei Attribute RaumbezugsID und RaumbezugID aufgeteilt. Die Kombination der fünf Metadatenattribute ist ebenso eindeutig wie der Primärschlüssel

MetadatenID. Dies wird durch Bedingungen sicher gestellt, die der Tabelle zugewiesen werden.

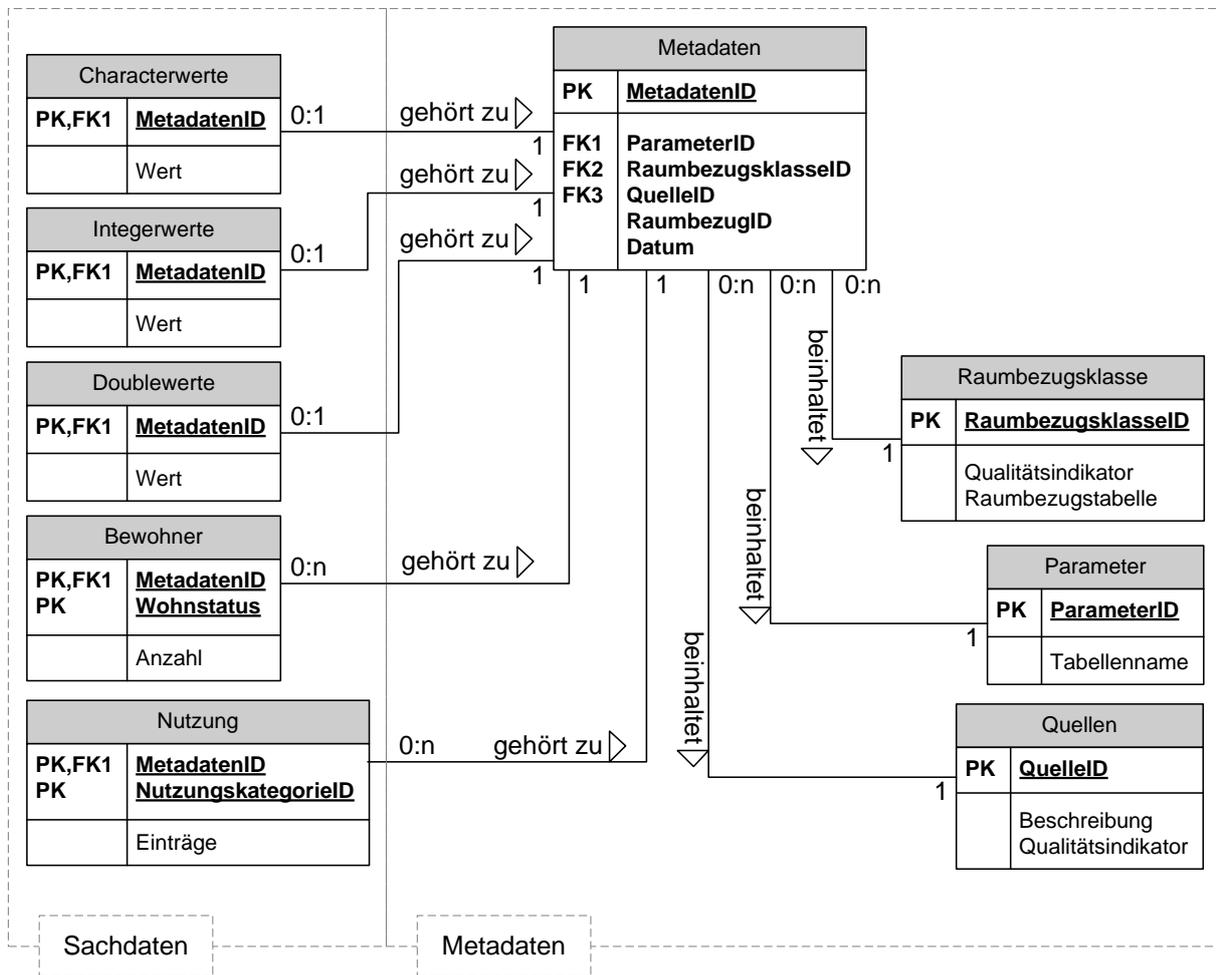


Abbildung 3.6.: Datenmodell der Primärparameterverwaltung

Die Fremdschlüsselbeziehungen der zentralen Tabelle „Metadaten“ mit den Tabellen „Raumbezugsklassen“, „Parameter“ und „Quellen“ erzwingen eine konsistente Bezeichnung der Metadatenattribute. Es wird geprüft, ob die jeweiligen Attributwerte der Tabelle Metadaten in den Tabellen vorhanden sind. Nur in diesem Fall ist das Einfügen des neuen Datensatzes in die Tabelle „Metadaten“ erfolgreich. Dieser Sachverhalt und die automatische Prüfung durch das Datenbanksystem wird „referentielle Integrität“ genannt. Die referentielle Integrität verhindert Fehler, beispielsweise durch falsche Schreibweisen und sichert damit die Konsistenz der Datenhaltung.

Ein Datensatz der Tabelle „Metadaten“ bezeichnet damit eine Einheit von Angaben für einen Gebäudeparameter und für ein Gebäude. Zu dieser Einheit können ein einziger oder auch mehrere Sachdatensätze gehören. Dies verdeutlicht das Beispiel von Abbildung 3.7. Das Beispiel gibt die Angaben aus Abbildung 3.5(a) nach dem Import wieder. Datensätze mit gleicher MetadatenID gehören jeweils zueinander. Den Datensätze 1, 3 und 4 der Tabelle Metadaten stehen jeweils ein Datensatz in den Tabellen Integerwerte

3. Konzept zur Verarbeitung der Gebäudeparameter

und Characterwerte gegenüber. Der Datensatz 3 beschreibt die Nutzung eines Gebäudes durch drei Einträge in der Tabelle Nutzung. Diese Beschreibung umfasst jedoch mehrere Sachdatensätze. Die ursprünglichen Raumbezugsangaben wurden ersetzt durch interne Raumbezugs-IDs. Es gibt zwei sich widersprechende Datensätze für das Baujahr aus unterschiedlichen Quellen.

Die Trennung von Sach- und Metadaten ermöglicht, dass trotz der verschiedenen Struktur unterschiedlicher Gebäudeparameterangaben die Informationen standardisiert verarbeitet werden können.

Tabelle Metadaten

MetadatenID	ParameterID	QuelleID	RaumbezugsklasseID	RaumbezugsID	Datensatz-Erstellung
1	Baujahr	Gebäudekarteikarten	Adresse	Adr_100	1992
2	Baujahr	Gebäudekarteikarten	Adresse	Adr_101	1992
3	Nutzung	EWE Netz Anschlüsse	Anschluss	Anschl_60	2009
4	Dachform	Gebäudekarteikarten	Gebäude	Geb_1	1992
5	Baujahr	Bebauungspläne	Bebauungsplan	Bbp_5	1980

Tabelle Integerwerte

MetadatenID	Wert
1	1963
2	1965
5	1980

Tabelle Characterwerte

MetadatenID	Wert
4	Flachdach

Tabelle Nutzung

MetadatenID	NutzungskategorieID	Wert
3	Wohnen	3
3	Einzelhandel	1
3	Gewerbe	1

Abbildung 3.7.: Beispiel für die Primärparameterverwaltung mit Meta- und Sachdatentabellen; Fortführung von Abbildung 3.5

Sachdaten Die Sachdaten unterscheiden sich bezüglich der Datentypen und der Struktur aus einem oder mehreren dazugehörigen Attributen. Aus diesem Grund gibt es insgesamt fünf Sachdatentabellen (siehe das Datenmodell in Abbildung 3.6). Drei sind für die Speicherung von Informationen der Basisdatentypen Integer, Character und Double vorgesehen. Die zwei weiteren Tabellen beziehen sich jeweils auf einen Gebäudeparameter.

Die Beispiele in Abbildung 3.7 sollen die unterschiedlichen Kardinalitäten zwischen den Sachdatentabellen und der Metadatatabelle veranschaulichen. Die Sachdatentabellen „Integerwerte“ und „Characterwerte“ weisen eine 1:1-Kardinalität zur „Metadaten“-Tabelle auf. Es gibt immer nur einen Datensatz je Datensatz in der Tabelle Metadaten. Ebenso gilt dies für die nicht dargestellte Tabelle „Doublewerte“. Da die Tabelle „Nutzungen“ noch Bedingungen beinhaltet, ist die Kardinalität zur „Metadatatabelle“ $n : 1$. Das heißt, ein Datensatz in der „Metadatatabelle“ ist mit mehreren Sachdaten-Datensätzen verbunden. Dies gilt ebenso für die nicht dargestellte Tabelle „Bewohner“.

3.3.2. Zusammenführung der Primärparameter

Die Zusammenführung der Primärparameter beschreibt den Vorgang, dass aus den möglicherweise unterschiedlichen Primärparameterangaben verschiedener Quellen eine Angabe

als „die richtige“ für einen Gebäudeparameter eines Gebäude ausgewählt und für die weitere Bearbeitung und Analyse zur Verfügung gestellt wird.

Metadatenebene Aufgrund des oben beschriebenen Datenmodells sind alle die Metadatenebene betreffenden Schritte für jeden Parameter gleich und somit standardisiert durchführbar. Für jedes Gebäude und jeden Parameter werden automatisiert die Datensätze verglichen und die besten daraus selektiert.

Abbildung 3.8 soll die Vorgehensweise für den Parameter Baujahr veranschaulichen. Der Algorithmus der Zusammenführung der Daten wird für jeweils einen Gebäudeparameter ausgeführt, den der Nutzer auswählt. Im Beispiel ist es das Baujahr. Hierfür existieren die drei Datensätze 1, 2 und 5 in der Tabelle Metadaten. Das Attribut „Datensatzerstellung“ gibt den Zeitpunkt der Datensatzerstellung an.

Der zweite Schritt, die Transformation des Raumbezugs, basiert auf den in Kapitel 3.2 dargestellten Beziehungen zwischen den Raumbezügen. Dabei wird anhand der Metadaten Raumbezugsklasse und Raumbezug jedem Datensatz der Tabelle „Metadaten“ eine oder mehrere GebäudeIDs zugeordnet. Im Beispiel beziehen sich die Adressen der Datensätze 1 und 2 auf das Gebäude Geb_1, der Bebauungsplan des Datensatzes 5 auf die Gebäude Geb_1 und Geb_2 (siehe zum Vergleich die Karte in Abbildung 3.5(b)).

Um verschiedene Quellen und Raumbezüge miteinander zu vergleichen, wird für beide Kategorien in Kapitel 3.1.2 ein Güteindikator eingeführt. Der Güteindikator ist eine numerische Bewertung der Quellen bzw. der Raumbezüge und erlaubt die Aufstellung einer Rangordnung. Die Schritte 3 und 4 fragen für jede Quelle und jede Raumbezugsklasse die Güteindikatorwerte ab und stellen diese für den nächsten Schritt zur Verfügung.

In Schritt 5 werden für jedes Gebäude die sich darauf beziehenden Datensätze sortiert. Die Platzierung gibt die Hierarchie des Datensatzes in der erhaltenen Reihenfolge an. Die Sortierung basiert auf den drei Attributen „Güteindikator Quelle“, „Datensatzerstellung“ und „Güteindikator Raumbezugsklasse“, wobei die drei Bedingungen gestaffelt sind:

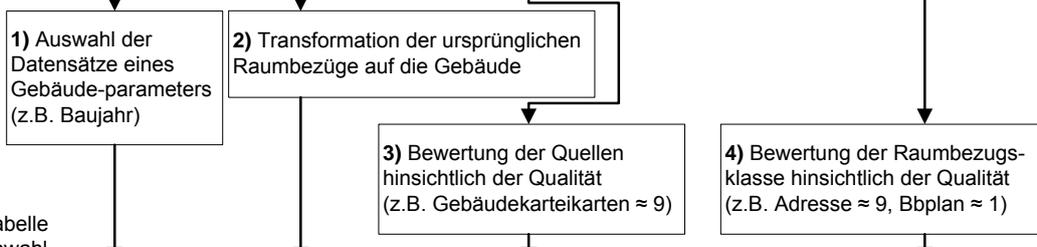
1. Die erste und wichtigste Ordnungsbedingung ist der Güteindikator der Datenquelle, wobei die Sortierung absteigend erfolgt.
2. Die zweite Ordnungsbedingung ist das Datum des Datensatzes und damit seine Aktualität. Die Ordnung erfolgt ebenfalls absteigend.
3. Die dritte und unterste Ordnungsbedingung ist der Güteindikator für die Raumbezugsklasse. Er dient der Priorisierung von räumlich fein gegliederten Angaben (z. B. Adresse) gegenüber weniger fein gegliederten Angaben (z. B. Postleitzahlgebiet).

Es werden nur die Datensätze weiter verwendet, die die Platzierung 1 aufweisen. Das Beispiel von Abbildung 3.8 zeigt jedoch auch anhand der Datensätze 1 und 2, dass für ein einzelnes Gebäude (Geb_1) mehr als ein Datensatz erstplatziert sein kann. Dies liegt daran, dass mehrere Adressen sich auf ein Gebäude beziehen und für jede Adresse eine Angabe existiert. Wenn die Adressen aus der gleichen Quelle stammen und das gleiche Erstelldatum aufweisen, stehen für eine Hierarchisierung dieser Daten keine weiteren

3. Konzept zur Verarbeitung der Gebäudeparameter

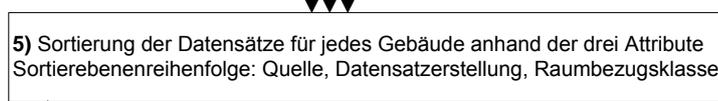
Tabelle Metadaten

Meta-datenID	ParameterID	RaumbezugsID	QuelleID	Datensatz-Erstellung	Raumbezugs-klasseID
1	Baujahr	Adr_100	Gebäudekartekarten	1992	Adresse
2	Baujahr	Adr_101	Gebäudekartekarten	1992	Adresse
3	Nutzung	Anschl_60	EWE Netz Anschlüsse	2009	Anschluss
4	Dachform	Geb_1	Gebäudekartekarten	1992	Gebäude
5	Baujahr	Bbp_5	Bebauungspläne	1980	Bebauungsplan



Temporäre Tabelle Metadatenauswahl

Meta-datenID	ParameterID	GebäudeID	Güteindikator Quelle	Datensatz-Erstellung	Güteindikator Raumbezugs-klasse
1	Baujahr	Geb_1	9	1992	10
2	Baujahr	Geb_1	9	1992	10
5	Baujahr	Geb_1	1	1980	1
5	Baujahr	Geb_2	1	1980	1



Temporäre Tabelle Ausgewählte_Datensätze

Meta-datenID	GebäudeID	Platzierung	# Datensätze je Gebäude	# Gebäude je Datensatz
1	Geb_1	1	2	1
2	Geb_1	1	2	1
5	Geb_1	2		
5	Geb_2	1	1	2

Wenn >1 besteht ein Datenkonflikt, der aufgrund der semantischen Zusammenhänge gelöst werden muss

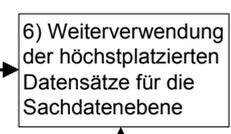


Abbildung 3.8.: Beispiel für die Zusammenführung der Primärparameter auf Metadatenebene

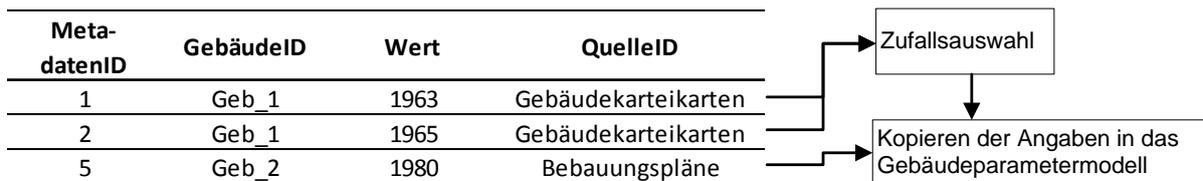


Abbildung 3.9.: Beispiel für die Zusammenführung der Primärparameter auf Sachdatenebene

Metadaten mehr zur Verfügung. Wie in diesem Fall weiter zu verfahren ist, hängt von der Semantik der Daten ab und wird daher auf der Sachdatenebene behandelt.

Bei der Datenübergabe werden die Zahl der betroffenen Gebäude je Datensatz und die betroffenen Datensätze je Gebäude gezählt. Als Ergebnis der Metadatenebene ergeben sich damit die Tabelle „Ausgewählte_Datensätze“ aus Abbildung 3.8. Mit diesen Informationen wird einerseits der beschriebene Fall gekennzeichnet, dass trotz der Zusammenführung mehrere Datensätze auf ein Gebäude verweisen. Andererseits ist es für die Weiterverarbeitung von quantitativen Parametern wichtig, ob ein Datensatz Informationen für mehrere Gebäude beinhaltet.

Bei den Daten im Untersuchungsgebiet Oldenburg haben alle Datensätze einer Quelle die gleiche Raumbezugsklasse. Damit wäre die dritte Ordnungsbedingung nicht notwendig. Sie ist jedoch der Vollständigkeit halber implementiert. Ein denkbarer Anwendungsfall wäre beispielsweise, Bebauungspläne unterschiedlicher Gebietsgröße zu unterscheiden und dadurch Angaben von sich überlappenden Bebauungsplänen feiner zu differenzieren.

Sachdatenebene Die Selektion der Daten ergibt für einen Parameter eines Gebäudes auf Basis der Metadaten keinen, einen oder mehrere Datensätze ($m \geq 0$). Im Falle mehrerer Datensätze gibt es keine Metadaten-Merkmale, anhand derer die Datensätze miteinander verglichen werden können. Wie im Beispiel von Abbildung 3.8 in der untersten Tabelle gezeigt, kann es auch vorkommen, dass sich die Angaben auf mehr als ein Gebäude beziehen ($m \geq 1$). Die verschiedenen Gebäudeparameter weisen unterschiedliche Strukturen auf und sind somit nicht mehr einheitlich behandelbar. Für deren Typisierung sind die folgenden Kriterien zu beachten:

- Ein Parameter kann aufgrund seiner Natur eine additive oder nicht additive Angabe sein. Beispielsweise ist das Baujahr eines Gebäudes nicht additiv, da mehrere Angaben nicht miteinander addiert werden. Die Einwohnerzahl eines Gebäudes ist eine additive Größe. Die Angaben zweier Datensätze, die sich beispielsweise auf die zwei Hauseingänge eines Gebäudes beziehen, können miteinander addiert werden.
- Falls ein Parameter eine additive Angabe ist, kann man diese in die Kategorien Absolutwert oder Verhältniswert unterteilen. Ein Absolutwert ist beispielsweise die absolute Bruttogrundfläche für eine bestimmte Nutzung. Deren Anteil an der gesamten Fläche ist ein Verhältniswert. Bei Verhältniswerten muss darauf geachtet werden, dass nach Umrechnungen die Summe der Werte 100 % ergibt.
- Die Frage, ob ein Parameter eine oder mehrere Nebenbedingungen enthält, führt zu einem dritten Unterscheidungsmerkmal. Die Gebäudenutzung z. B. hat als Nebenbedingung die Art der Nutzung. Da ein Sach-Datensatz nur eine Nutzung und deren Anteil am Gebäude enthält, gehören zur vollständigen Information immer mehrere Datensätze.

Diese Typisierung hilft, die Behandlung der Parameter teilweise zu standardisieren. Tabelle 3.7 gibt eine Übersicht der dadurch entstehenden Fälle und der vorzunehmenden Anpassungen der Daten.

3. Konzept zur Verarbeitung der Gebäudeparameter

Tabelle 3.7.: Parametertypen und Operationen für die Datenverarbeitung auf Sachebene

Parametertyp	n Datensätze pro Gebäude	m Gebäude pro Datensatz
Nicht additiv	(1) Mittelwertbildung / Auswahl	(2) Zuweisung zu jedem Gebäude
Additiv	(3) Auswahl oder Summe über n	(4) Aufteilung unter Berücksichtigung weiterer Informationen

Bei nicht additiven Parametern wie beispielsweise dem Baujahr oder der Höhe wird abhängig vom Kontext aus den Angaben ein Mittelwert gebildet oder eine der Angaben ausgewählt (Tabelle 3.7, Fall (1)). Abbildung 3.9 führt das bisher entwickelte Beispiel für die Verarbeitung des Parameters Baujahr fort. Da hierbei keine Informationen für die Begründung der Auswahl zur Verfügung stehen, wird einer der beiden Datensätze für Geb_1 zufällig ausgewählt. Eine Mittelwertbildung ist angemessen, wenn beispielsweise mehrere Höhenangaben für ein Gebäude existieren. Wenn es eine nicht additive Angabe für mehrere Gebäude gibt, kann diese ohne Anpassung auf mehrere Gebäude übertragen werden (Tabelle 3.7, Fall (2)).

Für additive Parameter hängt es vom Zusammenhang ab, ob ein Datensatz aus möglicherweise mehreren ausgewählt wird oder ob die Angaben der Datensätze summiert werden (Tabelle 3.7, Fall (3)). Eine Auswahl würde aufgrund fehlender weiterer Informationen zufällig erfolgen. Falls eine additive Angabe sich auf mehrere Gebäude bezieht, muss eine Aufteilung stattfinden. Im einfachsten Fall kann die Gesamtmenge zu gleichen Teilen auf die Gebäude verteilt werden. Alternativ ist die Berechnung der Anteile auf Basis weiterer Gebäudeparameter möglich (Tabelle 3.7, Fall (4)). Ein Beispiel hierfür ist die Anzahl der Bewohner zweier Gebäude. Dabei ist eine Gleichverteilung oder eine Verteilung abhängig von der Gebäudegröße denkbar.

Aufgrund der vorhandenen Daten und der daraus folgenden vereinfachten Verknüpfung der Raumbezüge tritt Fall (4) in Oldenburg nicht auf. Tabelle 3.8 listet die in Oldenburg bekannten Primärparameter, deren Parametertyp und die Lösung der vorkommenden Fälle auf. Die Höhenangaben in Oldenburg haben bereits in der Datenquelle einen direkten Gebäudebezug. Ein wichtiger, hier nicht gelisteter Primärparameter ist die Gebäudegrundfläche und -geometrie. Sie gehört zu den Raumbezügen, aus denen sie direkt berechnet werden kann.

3.3.3. Beispiele für die Verwendung der Primärparameterverwaltung

Übersicht der gespeicherten Daten Abbildung 3.10 gibt eine Übersicht der Daten für das Untersuchungsgebiet Oldenburg. Auf den Achsen sind die Primärparameter und die Datenquellen angetragen. Die Zahlen in den Kreisen geben die Anzahl der Gebäude an, für die ein entsprechender Primärparameter aus einer Quelle bekannt ist. Beispielsweise

sind aus der Quelle „Bauanträge“ Angaben zum Baujahr für 4.770 Gebäude gespeichert. Die Farben der Kreise verdeutlichen die Zahl der Gebäude. Die Skala reicht von Blau (sehr wenige Datensätze) bis Rot (ca. 80.000 Datensätze).

Ersetzen von Standardwerten für einzelne Gebäude Die Primärparameterverwaltung bietet die Möglichkeit, für verschiedene Gebäudeparameter mittlere Standardwerte zu definieren. Die Standardwerte erhalten einen niedrigen Güteindikatorwert und einen Raumbezug, der alle Gebäude einschließt. Falls keine anderen Daten zur Verfügung stehen, werden automatisch die Standardwerte bei der Zusammenführung der Gebäudeparameter herangezogen. Dadurch ist eine schnelle erste Auswertung für ein Untersuchungsgebiet möglich. Falls für kleine Gebiete oder einzelne Gebäude genauere Informationen zur Verfügung stehen, werden diese anstatt der Standardwerte gewählt. Die Primärparameterverwaltung bietet hierbei folgende Vorteile:

- Die Standardwerte werden für ein beliebiges Gebiet angegeben und behalten diesen originalen Gebietsbezug in der Primärparameterverwaltung. Bei der Änderung eines Wertes ist nur ein Datensatz betroffen.
- Der Nutzer muss sich nicht darum kümmern, welche Daten für welches Gebäude die genaueste Aussage ermöglichen. Allein durch die Vergabe eines Güteindikatorwertes für die zu berücksichtigenden Datenquellen wird automatisch für jedes Gebäude die Quelle mit dem höchsten Güteindikatorwert gewählt.
- Angaben aus Datenquellen mit unterschiedlichen räumlichen Bezugsgrößen (Gebäude, Stadtteil, Region, Land) werden automatisch und für alle Gebäude in einem Gebiet miteinander verglichen.

Tabelle 3.8.: Primärparametertypen der Daten in Oldenburg

Primärparameter	Parametertyp	Auftretende Fälle (Tab. 3.7)	Lösung
Baujahr	nicht additiv	(1), (2)	Zufällige Auswahl von Datensätzen -
Geschosszahl	Nicht additiv	(1)	Zufällige Auswahl von Datensätzen
First-, Traufhöhe	Nicht additiv		
Bewohner	Additiv	(3)	Summe der Bewohner je Wohnstatus
Stromanschlüsse	Additiv	(3)	Summe der Anschlüsse je Nutzungsform

3. Konzept zur Verarbeitung der Gebäudeparameter

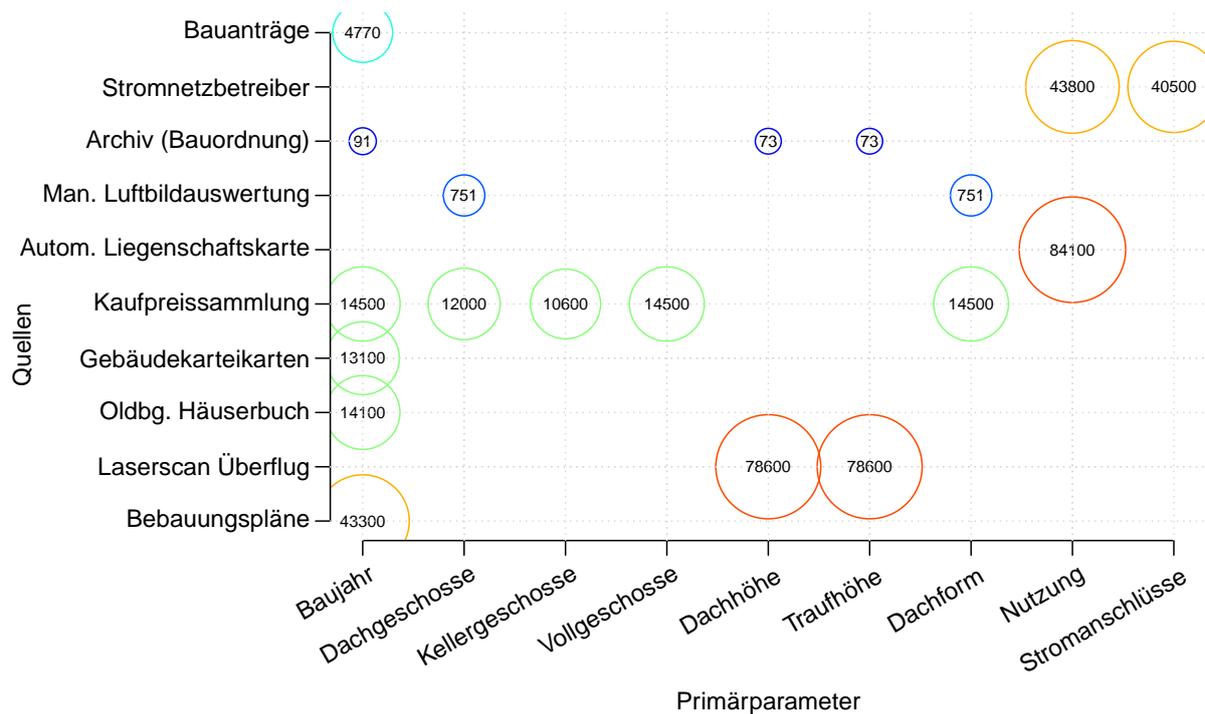


Abbildung 3.10.: Statistik der Primärparameterverwaltung für Parameter und Datenquellen; die Farben der Kreise verdeutlichen die Menge der Datensätze von Blau (sehr wenige) bis Rot (ca. 80.000).

- Es können mehrere Szenarien von Gebäudeparameterkombinationen erstellt werden, indem verschiedene Quellen für die Zusammenführung ausgewählt oder nicht ausgewählt werden.
- Falls zu einem späteren Zeitpunkt neue Daten hinzukommen, kann die Zusammenführung schnell erneut durchgeführt werden.

3.4. Gebäudeparametermodell

In diesem Datenmodell sind alle Gebäudeparameter enthalten, die für die Berechnung des Wärmebedarfs oder für die Charakterisierung der Gebäude bei der Wärmeverbrauchsanalyse notwendig sind. Es speichert sowohl die zusammengeführten Primärparameter als auch daraus abgeleitete Gebäudeparameter, wodurch ein einfacher Zugriff auf die Daten möglich ist. Nach der oben beschriebenen Zusammenführung werden die Primärparameter in das Gebäudedatenmodell kopiert. Wie die einzelnen abgeleiteten Gebäudeparameter berechnet werden, ist in Kapitel 5 näher beschrieben. Die Berechnung hängt von der lokal vorliegenden Datenlage ab und erfordert daher für das Untersuchungsgebiet angepasste Verfahren.

Das Gebäudeparametermodell umfasst die Angaben, die mit vertretbarem Aufwand für

eine stadtweite bzw. kommunale Analyse ermittelt werden können. Diese Gebäudeparameter werden Basisparameter genannt. Hierzu gehört beispielsweise die Größe von Bauteilen der thermischen Gebäudehülle. Diese Angabe kann aus der Höhe und dem Grundflächenumfang abgeschätzt werden. Angaben zum Wandaufbau oder dem U-Wert einzelner Gebäude sind Detailparameter, die für Tausende von Einzelgebäuden nicht ermittelt werden können. Da diese Detailparameter für die Berechnung des Wärmebedarfs notwendig sind, werden mittlere Werte aus statistischen Gebäudetypologien verwendet.

Eine Erweiterung des Modells um zusätzliche Attribute bzw. Basisparameter erfordert eine strukturelle Änderung der Tabellen. Die Änderungen sind jedoch aufgrund des geringen Strukturreichtums einfach durchführbar und beeinflussen nicht die anderen Datenmodelle. Es ist möglich, mehrere Instanzen des Gebäudedatenmodells zu erzeugen, um unterschiedliche Datenszenarien miteinander zu vergleichen.

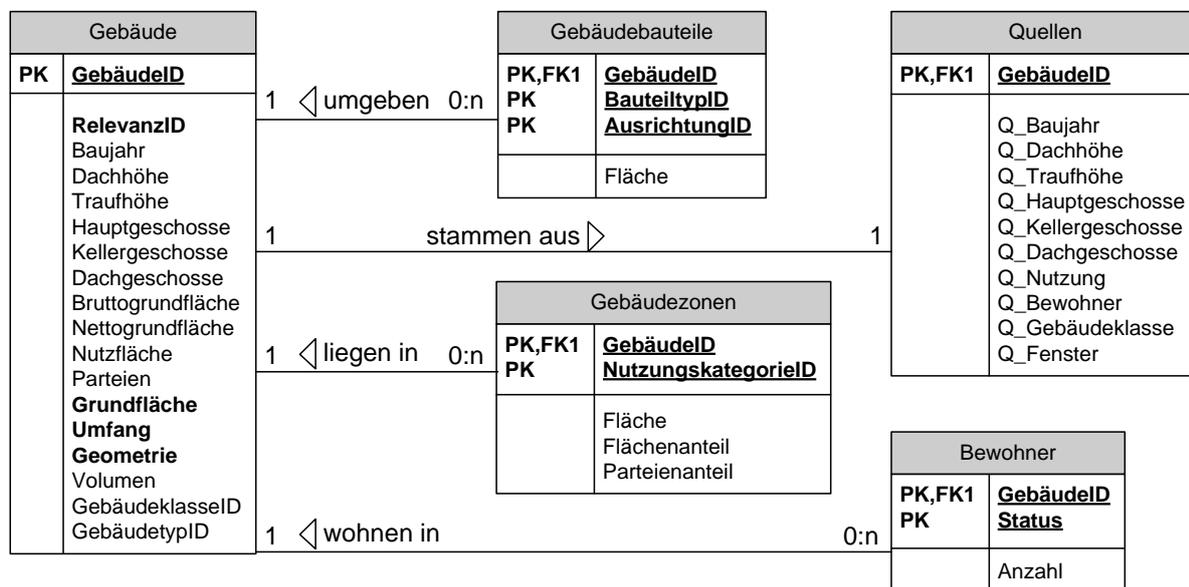


Abbildung 3.11.: Datenmodell der Gebäudeparameter

In Abbildung 3.11 sind die zum Gebäudeparametermodell gehörigen Tabellen mit den Attributen und der Kardinalität der Tabellenverknüpfungen gezeigt. Die Tabelle „Gebäude“ enthält alle Parameter, die nur einmal pro Gebäude auftreten. Ein Gebäude in der Realität entspricht genau einem Datensatz der Tabelle „Gebäude“ und ist eindeutig definiert durch den Primärschlüssel „GebäudeID“.

Für eine einfachere Darstellung der Gebäudeparameter in Karten besitzt jede Instanz des Gebäudeparametermodells ein eigenes Geometrieattribut „Geometrie“. Es muss daher nicht immer auf die Raumbezüge verwiesen werden. Anhand der „RelevanzID“ werden die Gebäude in Haupt- und Nebengebäude gegliedert. Nebengebäude sind größtenteils Garagen und kleinere Gebäude, die bei der Wärmeberechnung nicht berücksichtigt werden. Die weiteren Attribute der Tabelle Gebäude werden im Kapitel 5 näher beschrieben.

Zu einem Gebäude gehören mehrere Kategorien von Bauteilen, Zonen und Bewohnern. Es kann jeweils null bis mehrere Datensätze je Gebäude geben. Daher werden diese

3. Konzept zur Verarbeitung der Gebäudeparameter

in jeweils eigenen Tabellen gespeichert und anhand der GebäudeID mit dem Gebäude verbunden („PK“ bedeutet „Primärschlüssel“, „FK“ Fremdschlüssel). Die Tabellen sind folgendermaßen aufgebaut:

- Gebäudebauteile beschreiben die Teile der äußeren Gebäudehülle und sind gegliedert nach den Bauteiltypen „Fassade“, „Wand“, „Fenster“, „Dach“ und „Boden“. Darüber hinaus werden die fünf Ausrichtungen „Süd“, „Ost“, „West“, „Nord“ und „Horizontal“ unterschieden. Für jedes Bauteil, Ausrichtung und Gebäude wird die Fläche gespeichert.
- Die Bewohner werden nach dem Wohnstatus „Hauptwohnsitz“, „Nebenwohnsitz“ und „Alleiniger Wohnsitz“ unterschieden.
- Die Gebäudezonen beschreiben unterschiedliche Bereiche der Gebäudenutzung. Die Nutzung ist definiert anhand einer NutzungskategorieID. Unterschieden werden die Fläche in m², der Anteil der Fläche an der Gesamtfläche des Gebäudes und der Anteil der Nutzung an der Zahl aller sich im Gebäude befindlichen Parteien (Wohn- oder Gewerbeinheiten).

Die Tabelle „Quellen“ enthält für jeden Datensatz in der Tabelle „Gebäude“ genau einen Datensatz. Die Attribute „Q_...“ geben die bereits in der Primärparameterverwaltung verwendete „QuellenID“ an. Damit ist für die einzelnen Gebäudeparameter die Herkunft der Angaben bei der Analyse nachvollziehbar.

Im Untersuchungsgebiet Oldenburg liegen Informationen über die Gebäudesanierung und die Heizanlagen nicht vor. Deshalb sind hierfür keine Attribute im Gebäudeparametermodell vorgesehen. Beide Arten von Informationen würden jedoch durch jeweils eine zusätzliche Tabelle abgebildet werden, die anhand der GebäudeID mit dem jeweiligen Eintrag der Tabelle Gebäude verknüpft wird. Eine jeweils zusätzliche Tabelle ist empfehlenswert, da unter Umständen mehr als ein Heizungssystem bzw. mehr als eine Sanierungsmaßnahme pro Gebäude abgebildet werden soll.

4. Quellen für Gebäudedaten in Oldenburg

Die Analyse des Gebäudebestandes in Oldenburg basiert auf neun Datenquellen mit einem direkten Gebäudebezug und vier Quellen mit Gebietsbezug. Tabelle 4.1 listet alle Quellen und deren Güteindikator-Wert, dem Unterkapitel für weitere Beschreibungen und dem Zeitraum, in welchem die Daten erhoben wurden. Der Datenumfang verschiedener Quellen variiert zum Teil stark. Die umfangreichste Quelle ist die Automatische Liegenschaftskarte mit ca. 84.000 Datensätzen. Die manuelle Ermittlung von Baujahren aus dem Archiv des Bauamtes umfasst nur knapp 200 Datensätze, da die Ermittlung sehr zeitaufwendig ist.

Der Erhebungszeitraum für einen Großteil der Informationen liegt um das Jahr 2009. Quellen mit abweichendem Erhebungszeitraum beinhalten meist Angaben zum Baujahr und sind historisch gewachsen. Für jede Quelle wird ein Güteindikatorwert im Rahmen der Arbeit vergeben. Eine Beschreibung des Güteindikators enthält Kapitel 3.1.2.

Tabelle 4.1.: Gebäudedatenquellen in Oldenburg

Quelle	Güte-indikator	Kapitel	Aktualität der Angaben	Erfasste Gebäude
- Archiv (Bauordnung)	9	4.1.1	2012	188
- Bauanträge (Bauordnung)	10	4.1.2	1997 - 2010	4.700
- Gebäudekarteikarten (Bauordnung)	6	4.1.3	1930 - 1995	8.000
- Automatisierte Liegenschaftskarte (Geodaten)	10/6	4.1.4	2009	84.000
- Laserscan (Geodaten)	4	4.1.5	2009	78.000
- Manuelle Luftbilderkennung (Geodaten)	8	4.1.6	2011	831
- Kaufpreissammlung	8	4.1.7	1980 - 2009	13.000
- Oldenburger Häuserbuch	5	4.1.8	1995 - 2007	6.000
- Vertragszahlen des Stromnetzbetreibers	9	4.1.9	2009	40.000
- Bebauungspläne	1	4.2.1	1950 - 2000	
- Brandkasse	0	4.2.2	2011	
- Baufertigstellungsstatistik	0	4.2.3	1980 - 2011	
- Wohnraumzählung 1987	0	4.2.4	1987	

Die Gebäudedaten der Stadtverwaltung stammen aus unterschiedlichen Fachbereichen. Hierzu gehören die Stadtplanung bzw. Stadtentwicklung, das Bauamt, das Statistikamt

und das Amt für Geoinformation. Der Informationsaustausch, welche Daten verfügbar sind, und die Kommunikation, wofür die Daten verwendet werden sollen, nimmt in meist Monate in Anspruch. Ergänzend zu den Beschreibungen der Datenquellen in diesem Kapitel beinhalten die Diplomarbeit [52] und die Studie [37] weitere Angaben.

4.1. Gebäudedatenquellen mit Gebäudebezug

Die Datenquellen in diesem Kapitel beinhalten Angaben mit Bezug auf ein oder wenige Gebäude. Die Raumbezüge der Daten sind Gebäude oder Adressangaben. Die Beispiele in den Tabellen in diesem Kapitel sind anonymisiert, indem Adressen verändert und eindeutige Kennzeichnungen mit „xxx“ ersetzt sind.

4.1.1. Archiv (Bauordnung)

Das Archiv des Fachdiensts für Bauordnung beinhaltet die Bauakten der Gebäude in Oldenburg weitgehend vollständig. Die Akten sind nicht digitalisiert und bestehen größtenteils aus handschriftlichen Plänen und Anträgen. Informationen müssen manuell aus den Akten ausgelesen werden.

Tabelle 4.2.: Datensätze des Archivs (Bauordnung); die Datensätze werden jeweils in einer Spalte dargestellt.

Parameter	Datensatz 1	Datensatz 2
GebäudeID	555776	559193
Adresse	Hauptstr. 32	Hauptstr. 43 D
Baujahr (des Hauptgebäudes)	1970	vor 1987
Traufhöhe	2,75	2,57
Firsthöhe	8,85	8,3
Dachform	Walmdach	Walmdach
Hauptgeschosse	1	1
Kellergeschosse	1	0
Dachgeschosse	1	1
unbewohnte Dachgesch. (Abstellkammer)	1	1
Raumhöhe Hauptgeschoss (inkl. Decke)	2,75	2,87
Status	OK	OK
Bemerkung	bis 1997 Nebenstr. 32	Flachdachaufstockung 1987

Im Rahmen dieser Arbeit wird eine testweise Ermittlung von Gebäudedaten für ausgewählte Bezirke in Oldenburg durchgeführt. Die Ermittlung konzentriert sich auf Gebäude, für die aus anderen Datenquellen kein Baujahr verfügbar ist. Die hierfür involvierten Mitarbeiter werden mit angepassten und vorbereiteten Excel-Listen ausgestattet, um die benötigten Informationen im Archiv vor Ort digital aufzunehmen. Für die Ermittlung von ca. 100 Gebäudedatensätzen ist ein zeitlicher Aufwand von ca. 22 Personenstunden notwendig. Die Datenermittlung umfasst neben dem Baujahr die First- und Traufhöhen sowie die Geschosshöhen der Gebäude. In Summe werden für 188 Gebäude ein Baujahr und für 174 Gebäude Höhenangaben ermittelt. Bei einigen Bauakten fehlen die Bauzeichnungen für die Ermittlung von Höhenangaben.

Die manuelle Datenaufnahme unter der Annahme einer gewissenhaften Arbeitsweise begründet den Güteindikator-Wert 9 für die erhaltenen Angaben. Tabelle 4.2 gibt ein Beispiel für zwei Datensätze und den erhaltenen Angaben. Die GebäudeID und Adresse dienen als Raumbezug der Informationen, einerseits bei der Suche der Bauakte im Archiv, andererseits bei der Rückübermittlung der Daten.

4.1.2. Bauanträge (Bauordnung)

Der Fachdienst für Bauordnung bearbeitet bzw. archiviert alle Bauanträge seit 1996 mit der Software ProBauG. Die Software ist ein Frontend für eine Access-Datenbank und wird von ca. 50% der Gemeinden in Deutschland für die Bauordnung genutzt. Der Datenexport ist anhand einer Excel- bzw. Textdatei möglich. Die Bauordnung ist eine hoheitliche Aufgabe der Kommunen. Hieraus resultieren klare gesetzliche Vorgaben zur Bearbeitung und es kann davon ausgegangen werden, dass die Daten zu 100% korrekt sind. Daher wird dieser Quelle im Rahmen der Arbeit der höchste Güteindikator-Wert 10 gegeben. Die Daten der Stadt Oldenburg ergeben knapp 4.700 Angaben über Neubauten bis zum Jahr 2009.

Tabelle 4.3.: Datensätze der Bauanträge (Bauordnung)

Eingang	Grundstück (Raumbezug)	Vorhabenbeschreibung	Datum Fertigst.
12.12.1996	Musterstraße 1 Gemarkung: Oldenburg, Flur: 1 Flurstück: xxx	Neubau eines Einfamilienwohn- hauses (Veränderte Ausführung zu xxx)	01.07.1997
25.10.1996	Musterstraße 69 Gemarkung: Oldenburg, Flur: 2 Flurstück: xxx	Neubau eines Mehrfamilien- wohnhauses mit 5 WE Neubau einer Stellplatzanlage	18.09.1998

In Tabelle 4.3 sind die relevanten Spalten von zwei Datensätzen beispielhaft dargestellt. Das Datum „Eingang“ gibt das Antragsdatum, das Datum „Datum Fertigst.“ die Fertigstel-

lung an. Da für die Bauvorschriften das Datum der Antragstellung wichtig ist, wird dieses als Baujahr übernommen. Dadurch können sich jedoch im Vergleich mit der Baufertigstellungsstatistik des Landesamt für Statistik und Kommunikation Niedersachsen (LSKN) Unterschiede ergeben. Der Raumbezug der Daten ist immer eine Adressangabe, die wie in Kapitel 3.2.3 beschrieben verarbeitet wird. Die Spalte Vorhabenbeschreibung ist vom Typ Textfeld. Die den Neubau eines Gebäudes betreffenden Datensätze werden mithilfe des Textmusters „<Neubau> ... <Beheiztes Gebäude>“ erkannt (siehe Kapitel 3.1.3).

4.1.3. Digitalisierte Gebäudekartekarten (Bauordnung)

Gebäudekartekarten sind ein analoger Verweis auf eine Bauakte und beinhalten die knappe Beschreibung eines Bauantrags. Sie sind vergleichbar mit den digitalen ProBauG-Einträgen. In den 90er Jahren wurde durch die Stadtverwaltung Oldenburg begonnen, die analogen Gebäudekartekarten zu digitalisieren und in das Verwaltungssystem ProBauG zu integrieren. In der Vergangenheit waren deutlich mehr Veränderungen am oder im Gebäude meldepflichtig als seit den Entbürokratisierungsmaßnahmen Ende der 90er Jahre. Die Gebäudekartekarten beinhalten neben tatsächlichen Neubauanträgen auch Anträge für die Erneuerung des Heizungssystems oder die Umnutzung von Gebäuden.

Die älteste digitalisierte Gebäudekartekarte ist mit dem Baudatum 1889 versehen. Ab dem Jahr 1920 existiert eine nennenswerte Anzahl von Gebäudekartekarten. Aus diesen digitalisierten Informationen werden für diese Arbeit ca. 65.500 Datensätze zur Verfügung gestellt. Den Daten liegt ein Verwaltungsvorgang zugrunde, weshalb eine hohe Genauigkeit unterstellt werden kann. Bei der Digitalisierung wurden jedoch für ca. 5000 Einträge anstatt des wirklichen Datums des Verwaltungsvorgangs ein Standarddatum 1950 oder 1970 eingetragen. Alle Einträge mit diesen Angaben bleiben bei der Analyse unberücksichtigt. Da weitere Fehler nicht vollständig ausgeschlossen werden können, wird ein Güteindikatorwert 6 vergeben. Es kann nicht davon ausgegangen werden, dass die Digitalisierung in anderen Kommunen ebenfalls durchgeführt wurde.

Tabelle 4.4.: Datensätze der digitalisierten Gebäudekartekarten

Aktenzeichen	Eingang	Grundstück	Vorhaben
xxx-1963	01.01.1963	Oldenburg Musterstraße 14	Neubau eines Einfamilienwohnhauses Hinw.: 34381-1983 Bauantr. v. 09.05.1983 Neubau v. Garagen
xxx-1969	01.01.1969	Oldenburg Muster-Str. 23b	Neubau 2 GASHEIZUNGEN INSGES.35.000 WE
xxx-1995	01.01.1950	Oldenburg Musterstr. 40	Änderung eines Wohn- und Bürogebäudes (Nutzungsänderung: Wohnraum in Büro - Umbau Dachgeschoß); Neubau ...

In Tabelle 4.4 sind drei Datensätze beispielhaft dargestellt. Das Beispiel soll einerseits die Ähnlichkeit zu den ProBauG-Einträgen in Tabelle 4.3, andererseits die größere Vielfalt bei der Vorhabenbeschreibung zeigen. Eine Besonderheit ist, dass die ID, das sogenannten Aktenzeichen die Jahresangabe für einen Eintrag in den letzten vier Ziffern beinhaltet (z. B. 11952-1969 gibt das Jahr 1969). Der Ortsbezug wird über die im Feld Grundstück angegebene Adresse mit der in Kapitel 3.2.3 beschriebenen Methode des Adressvergleichs hergestellt. Die Spalte Vorhaben wird anhand der in Kapitel 3.1.3 beschriebenen Musterrerkennung von Textfeldern analysiert, um alle Datensätze zu extrahieren, die das Muster „<Neubau> ... <Beheiztes Gebäude>“ enthalten.

4.1.4. Automatisierte Liegenschaftskarte (ALK)

Die ALK ist ein raumbezogenes Informationssystem, in welchem die Geodaten für Flurstücke, Gebäude und Adressen verwaltet werden. Die Führung der ALK ist in Deutschland die Aufgabe der Vermessungsämter der Länder. Die Gemeinden beziehen die ALK für das jeweilige Gemeindegebiet von den Ländern. Nähere Informationen bietet [52]. Die Pflege der ALK ist gesetzliche Hoheitsaufgabe und die Angaben mit sehr hoher Wahrscheinlichkeit korrekt. Daher erhalten alle Raumbezüge der Quelle den höchsten Güteindikator-Wert 10. Die ALK beinhaltet einen Datensatz für jedes Gebäude. Dies sind in Oldenburg ca. 82.100 Gebäude.

Die Geometrien der ALK bilden das Grundgerüst der Gebäudedatenbank, da mit diesen das Datenmodell der Raumbezüge gefüllt wird. Vor dem Import der Geometrien werden unterirdische Gebäude und sich überschneidende Gebäudegrundrisse herausgefiltert. Bei den Flurstücken und Adressen werden nur die Geometrien und Lagebezeichnungen als Daten in das Modell der Raumbezüge übernommen. Für die Gebäude liefert die ALK neben den Geometrien und einer eindeutigen Gebäudeobjekt-ID bzw. Lagebezeichnung noch einen sogenannten Objektschlüssel und die dazugehörige Bezeichnung. Der Objektschlüssel lässt eine grobe Abschätzung der Gebäudenutzung zu. Tabelle 4.5 zeigt drei Beispieldatensätze.

Die Geometrie besteht aus einem Polygon (siehe Kapitel 2.2), von welchem jeweils die ersten zwei Eckpunkte angegeben sind. Die ALK umfasst insgesamt 105 verschiedene Objektschlüssel für Gebäude, wovon in Oldenburg 86 vorkommen. Fast 40.000 Datensätze tragen den Objektschlüssel mit der Bezeichnung „Wohngebäude“, fast ebenso viele die Bezeichnung „Gewerbegebäude“. Hierbei muss jedoch beachtet werden, dass unter Gewerbegebäude alle sonstigen Gebäude wie Garagen, Schuppen, etc. fallen. Öffentliche Gebäude wie „Hochschulgebäude“, „Parkhaus“ oder „Museum“ sind fein untergliedert und kommen nur vereinzelt vor.

4.1.5. Laserscan-Überfliegung

Im Rahmen einer Feinstaub-Emissionsstudie wurde für die Stadt Oldenburg eine Laser-scanning-Überfliegung mit vier Messpunkten pro Quadratmeter Erdoberfläche durch-

4. Quellen für Gebäudedaten in Oldenburg

Tabelle 4.5.: Datensätze der ALK mit Objektart und den jeweils ersten zwei Koordinaten der Gebäudegeometrie

Objekt ID	Objektart- schlüssel	Objektartbezeichnung	Geometrie
O0H9xxx	1101	Öffentl. Gebäude	POLYGON(((3447754.7289 5887755.78395,3447760.1339 5887768.71195, ...
O060xxx	931	Wohngebäude	POLYGON(((3446714.2051 5891740.079,3446724.8619 5891737.139, ...
O0ANxxx	932	Gewerbegebäude	POLYGON(((3449447.07505 5887782.967,3449448.66105 5887780.006, ...

geführt. Die Überfliegung fand zu Beginn des Jahres 2009 statt. Aus den Messpunkten wurde durch eine räumliche Überlagerung mit den Gebäudegrundflächen der ALK vier Höhen berechnet: die maximale Höhe, die Firsthöhe (oberer Dachabschluss), die als Traufhöhe bezeichnete Kante des unteren Dachabschlusses und die maximale Höhe über Normal Null (NN). Die für die Arbeit wichtigen Angaben Trauf- und Firsthöhe sind in Abbildung 5.10 auf Seite 78 dargestellt.

Nach Auskünften des Fachdienstes für Statistik und Geoinformation wurden die einzelnen Messpunkte nicht interpoliert, um die Gebäudekanten und daraus wiederum die Höhenangaben zu berechnen. Es fand nur eine statistische Analyse der Messpunkte statt. Einzelne Stichproben ergaben eine zum Teil sehr schlechte Qualität der Ergebnisse, da z. B. Vordächer bzw. tiefere Gebäudeteile als unterer Dachabschluss erkannt wurden. Bei einfachen Gebäudegeometrien sind die Ergebnisse jedoch weitgehend korrekt. Für die Quelle wird der Güteindikatorwert 4 vergeben.

Die Quelle umfasst ca. 78.500 Gebäude und stellt damit flächendeckend für das gesamte Stadtgebiet Höhenangaben bereit. Der Raumbezug ist die Gebäudegeometrie, wodurch Fehler bei der Zuordnung unterschiedlicher Raumbezugsklassen ausgeschlossen werden können.

4.1.6. Manuelle Luftbildauswertung

Bei der Analyse der Laserscan-Daten sind Flachdächer nicht ausreichend berücksichtigt. Im Rahmen dieser Arbeit werden testweise Flachdächern für die Postleitzahlgebiet 26121 und 26127 in Oldenburg auf Basis hochauflösender, georeferenzierter Luftbildaufnahmen vom April 2011 ermittelt. Hierfür werden die Flachdächer durch einen Mitarbeiter des Fachdienstes Statistik und Geoinformation mit einem georeferenzierten Punkt innerhalb des zugehörigen Gebäudepolygons gekennzeichnet. Die räumliche Überlagerung der

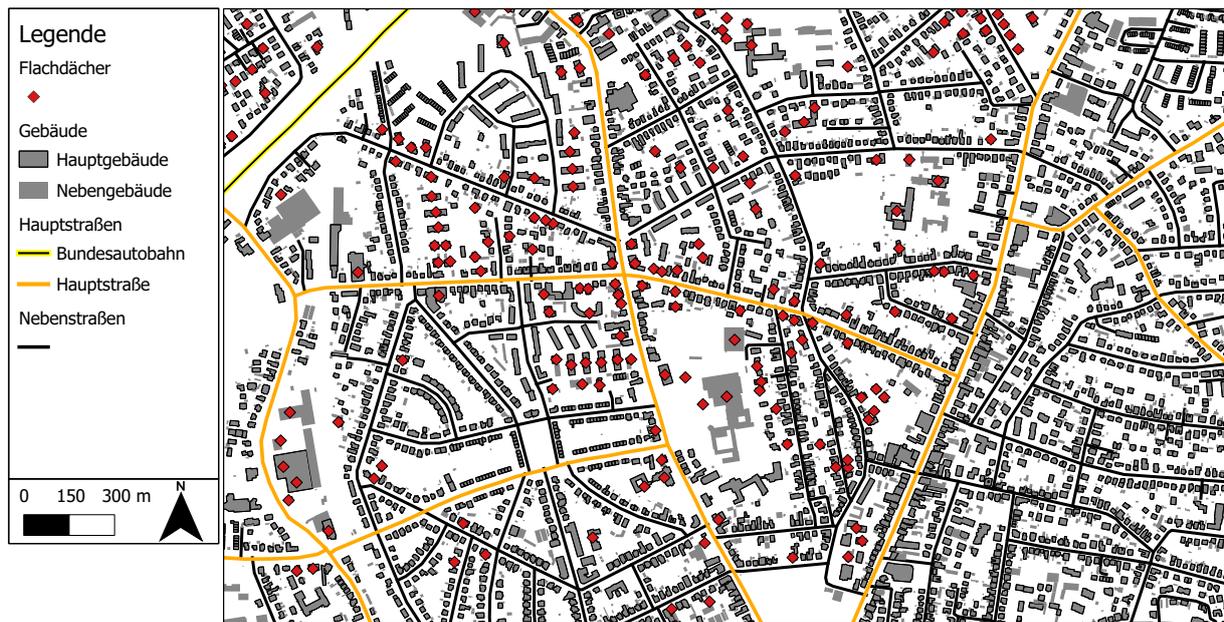


Abbildung 4.1.: Kartenausschnitt mit manuell erkannten Flachdächern in Oldenburg

Punkte mit den Gebäudepolygonen der Datenbank erlaubt die Zuordnung der aufgenommenen Daten. Dächer, für die mithilfe der orthogonalen Luftbildperspektive keine sichere Entscheidung getroffen werden kann, werden durch Schrägluftaufnahmen verifiziert. Für die Schrägluftaufnahmen dient die Vogelperspektive des Kartendienstes Bing Maps.

Problematisch ist die Einstufung komplexer Dachformen innerhalb eines Gebäudegrundrisses. In diesem Fall wird die flächenmäßig überwiegende Dachform gewählt. Für die Erfassung eines Gebiets von knapp 14 km² durch das GIS System und die Nacharbeitung mit Bing Maps werden ca. 12 Personenstunden benötigt. In Abbildung 4.1 sind die Flachdächer als rote Raute gekennzeichnet.

Unter der Voraussetzung einer gewissenhaften Arbeit sind die Angaben korrekt und gebäudescharf bearbeitet. Unsicherheiten in den Ergebnissen entstehen durch die vereinfachte Klassifizierung komplexer Dachformen. Daher wird der Quelle der Güte-Indikator 8 zugewiesen.

4.1.7. Kaufpreissammlung des Gutachterausschusses für Grundstückswerte (GAG) des Landesamts für Geoinformation und Landentwicklung (LGLN)

Der Gutachterausschuss ist ein selbstständiges Gremium, das anhand von Wertermittlungs- und Verkehrswertgutachten Informationen über Grundstückswerte und Marktzusammenhänge für Bürger, Behörden und Unternehmen zur Verfügung stellt. Die Kaufpreissammlung ist die Grundlage für genannte Gutachten und erfasst alle Immobilienverkäufe.

4. Quellen für Gebäudedaten in Oldenburg

Tabelle 4.6.: Relevante Attribute der Kaufpreissammlung

Attribut	Anteil am Hauptgebäudebestand	Datensätze	Anzahl verschiedener Attributewerte
Baujahr	48 %	19.550	211
Dachform	36 %	14.800	8
Geschosszahl	37 %	14.900	20

Die Verkäufe werden anhand des Kaufvertrags, der Bauakte, Fragebögen und unter Umständen Ortsbesichtigungen ausgewertet und gespeichert [49]. In Oldenburg werden die Daten seit den 80er Jahren digital gespeichert. Die Gebäudeinformationen werden durch die Gutachter soweit möglich plausibilisiert. Für die Arbeit werden ca. 20.700 Gebäudedatensätze im Stadtgebiet Oldenburg mit einer Vielzahl von gebäudebezogenen Attributen zur Verfügung gestellt. Informationen zu Eigentümern, Werten und Preisen sind nicht beinhaltet. Für die Datenquelle wird ein Güteindikatorwert 9 vergeben.

Der Raumbezug der Datensätze ist durch eine Adressangabe gegeben. Nicht jeder Datensatz enthält Einträge zu allen Attributen. Die Kaufpreissammlung ist in dieser Arbeit vor allem eine Quelle für die Ermittlung des Baujahrs. Tabelle 4.6 listet die relevanten Attribute auf. Die Anzahl verschiedener Attributewerte zeigt, dass insgesamt Gebäude aus 211 Baujahren erfasst sind und acht Dachformen sowie 20 Geschosszahlen unterschieden werden.

4.1.8. Oldenburger Häuserbuch

Das Oldenburger Häuserbuch ist die geschichtliche Aufarbeitung von ca. 10.000 Gebäuden in Oldenburg mit Informationen zu deren Geschichte, Abriss, Neubauten und Eigentümern. Das Buch entstand aus der jahrzehntelangen Archivarbeit von Herrn Wachtendorf. Das Speicherformat ist Fließtext. Im Rahmen des Masterprojekts „Energieeffiziente Stadt Oldenburg“ wurde die Datei auf Basis der Textformatierung in jeweils einen Datensatz pro Gebäude aufgeteilt und die Adressangabe extrahiert. Die Identifikation eines Baujahrs basiert auf einer Mustererkennung. Das Muster ist zusammengesetzt aus den Schlüsselwörtern „neu gebaut“, „neu errichtet“, „Erbauungsjahr“ und einer Jahreszahl. Es können knapp 6.000 Einträge mit einer Baujahresangabe identifiziert werden [44].

Die Informationen basieren auf alten Quellen und der Fokus der Untersuchung von Herrn Wachtendorf ist der historische Werdegang der Gebäude. Es kann mit hoher Wahrscheinlichkeit davon ausgegangen werden, dass die Informationen korrekt sind. Unsicherheiten entstehen durch das Alter der Quellen und mögliche Fehler bei der automatischen Analyse der Textinformationen. Daher erhält die Quelle den Güteindikator 5.

4.1.9. Vertragszahlen des Stromnetzbetreibers

Der Netzbetreiber stellt die Stromnetzinfrastruktur zur Verfügung und ist für das Ablesen der Strommengen zuständig. Diese teilt er den Lieferanten zum Zweck der Abrechnung mit. Der Stromnetzbetreiber erfasst somit alle Anschlüsse, unabhängig von welchem Energielieferanten ein Haushalt oder Unternehmen den Strom bezieht. Ein Anschluss entspricht damit einem Stromvertrag. In Oldenburg gibt es ca. 102.000 Anschlüsse bzw. Stromverträge. Diese sind verteilt auf ca. 54.000 Anschlussobjekte. Ein Anschlussobjekt entspricht in den meisten Fällen einem Hausanschluss, zu dem ein bis mehrere Zähler für die einzelnen Verträge gehören.

Jeder Vertrag hat eine Branchenangabe. Die Branchen sind klassifiziert in knapp 280 unterschiedliche Nutzungskategorien, die wiederum den fünf Branchentypen Haushalt, Landwirtschaft, Produzierendes Gewerbe, Dienstleistungen und Verteilerwerke/Kraftwerke zugeordnet sind. Die Branchenschlüssel entstammen einer Gewerbetypologie, die historisch für die Verwaltung der Anschlüsse der EWE Netz entstanden ist. Ziel der Typologie ist vor allem die Planung von Lastkurven, jedoch spielen auch Marketing und statistische Gesichtspunkte eine Rolle. Seitens des Stromnetzbetreibers werden somit die Zahl der Stromverträge, gegliedert nach der Branche und den Anschlussobjekten zur Verfügung gestellt. Die Raumbezugsangaben der Anschlussobjekte sind Adressen (siehe Kapitel 3.2.1). Tabelle 4.7 gibt ein Beispiel für die Struktur der Daten.

Die Daten entstammen dem Abrechnungssystem der EWE Netz und haben daher eine hohe Genauigkeit. Die Zuordnung der Anschlussobjekte und Gebäude kann in geringem Umfang unklar sein, wenn Hinter- oder Nebengebäude durch ein Anschlussobjekt mit versorgt werden. Daher erhält diese Quelle den Güteindikator 9.

Tabelle 4.7.: Datensätze zu Stromanschlüssen

Anschlussobjekt	Straße	Hausnr.	Verträge	Branche
180xxx	Musterstr.	30	6	Haushalte
			1	Übrige Dienstleistung
180xxx	Musterstraße	48	1	Gaststätten
			1	Anderer Einzelhandel

4.2. Gebäudedatenquellen mit Gebietsbezug

Gebietsbezogene Datenquellen umfassen Angaben, die nicht direkt auf ein Gebäude verweisen, sondern auf das gesamte Untersuchungsgebiet oder Teile davon.

4.2.1. Rechtsverbindliche Bebauungspläne (B-Pläne)

Ein Bebauungsplan (B-Plan) wird von der Kommune für ein kommunales Teilgebiet beschlossen und bestimmt die Art der Bebauung wie beispielsweise die Gebäudetypen, Nutzungsvorschriften für die Gebäude und möglicherweise auch Dachform und Aussehen der Gebäude. Beschlossene B-Pläne enthalten ein Datum der Rechtsverbindlichkeit, ab welchem sich die Bebauung nach den Vorgaben richten muss. Die Stadt Oldenburg hatte bis zur Fertigstellung der Arbeit ca. 50% der rechtsverbindlichen B-Pläne digitalisiert. Die Gebietsangaben sind als Polygon-Geometrien für einen Teil von Oldenburg enthalten. Diesen sind Sachdaten zugeordnet. Für die Arbeit interessant ist das Datum der Rechtsverbindlichkeit, weitere Lagebeschreibungen und die Art des Bebauungsplans mit Unterscheidung zwischen allgemeinem und vorhabenbezogenem Bebauungsplan. Im Fall letzterer ist davon auszugehen, dass die Gebäude innerhalb des Gebiets kurz nach der Rechtsverbindlichkeit des Bebauungsplanes errichtet wurden.

Die Übertragbarkeit des Datums der Rechtsverbindlichkeit auf die Baujahre der im B-Plan-Gebiet stehenden Gebäude ist in vielen Fällen problematisch. Im Planungsgebiet können sich beispielsweise bereits vor dem Beschluss des B-Plans Gebäude befunden haben oder es wurden erst mit großer zeitlicher Verzögerung Neubauten realisiert. Darüber hinaus geht aus den digitalisierten B-Plänen nicht der Grund des Bebauungsplans hervor. Hinzu kommt, dass sich einzelne B-Pläne zum Teil überlappen und aus den verfügbaren Daten keine Anhaltspunkte entnommen werden können, welchem Bebauungsplan man in diesem Fall den Vorzug geben muss. Daher wird den Bebauungsplänen der Güteindikator 1 zugewiesen.

4.2.2. Öffentliche Brandkasse

Die Öffentliche Landesbrandkasse ist ein Versicherungsunternehmen in Oldenburg. Bis ca. 1990 war das Unternehmen Monopolversicherer für die Pflichtversicherung Gebäude-Brandschutz in der Region Oldenburg. Ein Aspekt bei der Wertbestimmung eines Gebäudes im Rahmen der Brandschutzversicherung ist das Baujahr. Die Öffentliche Landesbrandkasse verfügt über dieses Attribut, gibt jedoch aus Datenschutzgründen keine Versicherungsdaten - wozu das Baujahr gehört - in nicht-anonymisierter Form, d.h. mit einer Adresse versehen heraus. Es werden jedoch zur Validierung der Daten die den Versicherungsverträgen hinterlegten Baujahre mit der Angabe der Postleitzahl zur Verfügung gestellt. Die Anzahl der Baujahresdaten deckt einen Großteil der Hauptgebäude in Oldenburg ab.

Die Daten dienen der Validierung der ermittelten und geschätzten Baujahre (siehe Kapitel 5.1.5). Da die Angaben nur statistisch vorliegen, erhalten sie den Güteindikatorwert Null.

4.2.3. Baufertigstellungsstatistik des Landesamt für Statistik und Kommunikation Niedersachsen (LSKN)

Die Baufertigstellungsstatistik der LSKN erfasst die abgeschlossenen Neubau-Bauanträge. Die Daten für die Stadt Oldenburg werden jährlich vom Bauamt der Stadt gemeldet. Bei der Übermittlung gab es laut Auskunft des Bauamtes Oldenburg in den 90er Jahren Übermittlungsfehler, weshalb einige Neubauten erst mit einem zeitlichen Versatz von einem oder zwei Jahren gemeldet wurden. Die Zeitreihe beinhaltet somit Fehler, die nicht mehr vollständig nachvollzogen werden können. Da die Angaben nur statistisch vorliegen, erhalten sie den Güteindikatorwert Null.

4.2.4. Wohnraumzählung der Volkszählung 1987

Die Volkszählung von 1987 beinhaltete vier Fragen zum Gebäude. Eine Frage davon betraf das Baujahr des Gebäudes [74]. Die Einteilung umfasste die Baualtersklassen bis 1900, 1900 - 1918, 1919 - 1948, 1949 - 1957, 1958 - 1962 und jahresscharfe Angaben ab 1963. Eine weitere die Gebäudeart, wobei zwischen Wohngebäude, sonstiges Gebäude mit Wohnraum und bewohnter Unterkunft unterschieden wurde. Da die Angaben nur statistisch vorliegen, erhalten sie den Güteindikatorwert Null.

5. Berechnung und Analyse von Gebäudeparametern in Oldenburg

Gebäudeparameter sind die Angaben, die in die Wärmebedarfs- bzw. Wärmeverbrauchsberechnung eingehen. Sie sind in den meisten Fällen nicht direkt in den Datenquellen enthalten, sondern werden aus den Primärparametern abgeleitet. Die dafür verwendeten Verfahren hängen davon ab, welche Quellen und Primärparameter für ein Untersuchungsgebiet zur Verfügung stehen. In Oldenburg sind dies die in Abbildung 5.1 in der linken Spalte dargestellten Parameter. Darauf aufbauend werden die Gebäudeparameter berechnet. Die Verfahren sind als Funktionen der PostgreSQL/PostGIS-Geodatenbank implementiert.

Für die Wärmebedarfsberechnung werden im Stadtenergiebereich vor allem Hochrechnungs- und Bilanzverfahren angewendet. Für Hochrechnungsverfahren werden die Gebäude anhand des Gebäudetyps und der Nutzung differenziert. Hochrechnungsverfahren benötigen für jede der berücksichtigten Typ-Nutzungs-Klassen einen spezifischen Bedarfswert, der mit einer Bezugsfläche multipliziert wird. Bezugsflächen sind z. B. die Energienutzflächen oder Wohnflächen. Bilanzverfahren benötigen Informationen über die Flächen der thermischen Hülle, das Gebäudevolumen und die Gebäudenutzung (siehe Kapitel 2.4.3). Darüber hinaus werden Gebäudeparameter, die aufgrund des damit verbundenen Aufwands nicht erhoben werden - z. B. die U-Werte (Wärmeleitwerte) - anhand des Gebäudetyps aus der Literatur zugeordnet.

Die Abbildung 5.1 gezeigten Informationsflüsse gelten für jedes Gebäude. Jede Box repräsentiert einen Parameter. Die verbindenden Pfeile geben an, auf Basis welcher Primärparameter die Gebäudeparameter berechnet werden. Falls für ein Gebäude keine Angaben vorhanden sind, werden diese entweder auf Basis der Angaben benachbarter Gebäude oder aus anderen, bekannten Gebäudeparametern abgeleitet. Zur besseren Übersicht sind diese „Ausnahmen“ nicht dargestellt. Sie werden jedoch in den entsprechenden Unterkapiteln beschrieben. Die Gebäudeparameter „Therm. Hüllflächen“ und „Grundflächen“ sind Oberbegriffe für mehrere Gebäudeparameter.

Die dick umrandeten Parameter sind Eingangsgrößen für die in Kapitel 6 beschriebene Verbrauchsdatenanalyse. Graue Flächen geben einen Verweis auf die Unterkapitel, in dem die jeweiligen Schritte verdeutlicht sind. Zum Teil werden darin die Primärparameter ausführlich analysiert, wie beispielsweise die Angaben zum Baujahr. Es findet ein Vergleich der errechneten Gebäudeparameter mit statistischen Daten statt, soweit diese zur Verfügung stehen.

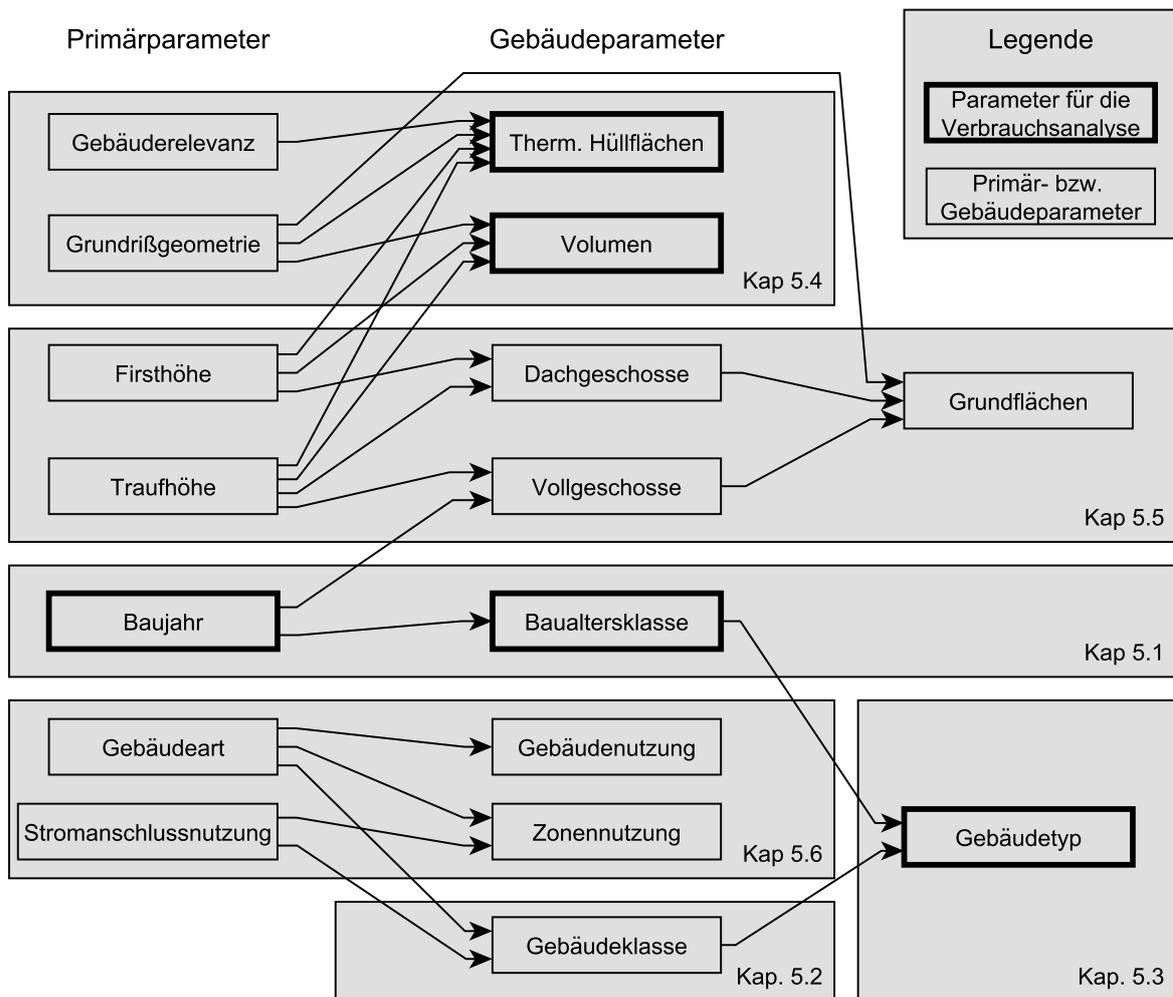


Abbildung 5.1.: Informationsfluss zwischen Primär- und Gebäudeparametern mit Angabe der jeweiligen Unterkapitel

5.1. Gebäudebaujahr und Baualtersklasse

Die Änderungen der gesetzlichen Rahmenbedingungen für den Wärmeschutz in Wohngebäuden seit dem Zweiten Weltkrieg führen über ihren Regulierungsmechanismus dazu, dass das Baujahr eines Gebäudes ein wichtiger Einflussfaktor auf den Heizwärmeverbrauch ist. Die geänderten Rahmenbedingungen wirken sich beispielsweise in den U-Werten der Gebäudetypologien des Wohngebäudebereichs [53, 71] aus.

5.1.1. Datenquellen

In Oldenburg waren vor Beginn dieser Arbeit keine Angaben zu den Gebäudebaujahren zentral und digitalisiert verfügbar. Generell existieren vollständige Angaben zu den Ge-

bäudebaujahren in einer Kommune meist nur in den Archiven der Bauordnungsämter [37]. Die Aufarbeitung und Digitalisierung dieser Daten erfordert jedoch viel manuelle Arbeit und ist dementsprechend zeit- und kostenaufwändig. Im Rahmen dieser Arbeit werden daher sechs zumeist digital verfügbare Datenquellen für das Stadtgebiet Oldenburg identifiziert, die Nutzungserlaubnis geklärt und die Daten für die weitere Verwendung aufbereitet. Tabelle 5.1 gibt eine Übersicht. Der Überprüfung der Angaben dienen drei unterschiedliche Statistiken.

Tabelle 5.1.: Datenquellen für den Parameter Baujahr

Datenquelle	Güteindikator
<i>Quellen für die Parameterbestimmung</i>	
- Bauanträge des Fachdienstes für Bauordnung (Kapitel 4.1.2)	10
- Gebäudekarteikarten des Fachdienstes für Bauordnung (Kapitel 4.1.3)	6
- Kaufpreissammlung des Gutachterausschusses für Grundstücksbewertung (Kapitel 4.1.7)	8
- Historische Aufarbeitung durch das Oldenburger Häuserbuch (Kapitel 4.1.8)	5
- Bebauungspläne (B-Pläne) des Fachdienstes für Statistik (Kapitel 4.2.1)	1
- Archiv des Fachdienstes für Bauordnung (Kapitel 4.1.1)	9
<i>Quellen für den statistischen Vergleich</i>	
- Gebäudebaujahre aus den Brandversicherungsverträgen des Versicherers Öffentliche Landesbrandkasse , differenziert nach Postleitzahlen (Kapitel 4.2.2)	
- Gebäude- und Wohnraumstatistik der Wohnraumzählung 1987 (Kapitel 4.2.4)	
- Baufertigstellungsstatistik des (Kapitel 4.2.3)	

Für jede Datenquelle sind das Datenformat und der Güteindikator angegeben, welcher die Vertrauenswürdigkeit der Quellen quantifiziert (siehe Kapitel 3.1.2). Die Wahl des Güteindikators ist in den angegebenen Unterkapiteln von Kapitel 4 begründet, in denen die Quellen ausführlicher beschrieben sind:

- Die Bauanträge dokumentieren alle seit Mitte der 90er Jahre gemeldeten und abgeschlossenen Bauvorhaben von Gebäuden. Aufgrund der hoheitlichen Aufgabe der Verwaltung kann davon ausgegangen werden, dass diese Datenquelle alle in dieser Zeit entstandenen Gebäude umfasst.
- Die Gebäudekarteikarten sind eine Kurzdarstellung von bearbeiteten Anträgen im Bauordnungsamt vor der digitalisierten Verarbeitung. Sie lagen ursprünglich in Papierform vor und wurden im Untersuchungsgebiet seit den 90er Jahren nachträglich digitalisiert.

- Das Oldenburger Häuserbuch besteht aus der manuellen Aufarbeitung der Geschichte von über 10.000 Gebäuden in Oldenburg. Die Arbeit wird seit über zehn Jahren durch eine interessierte Privatperson durchgeführt.
- Die Bebauungspläne liegen für das Untersuchungsgebiet als georeferenzierte Polygone mit dem Datum der Rechtsverbindlichkeit vor. Der Inhalt eines Bebauungsplans ist im Detail nicht bekannt.
- Zur testweisen Ergänzung von fehlenden Baujahresdaten fanden Recherchen im Archiv des Bauordnungsamtes statt. Hierzu wurden auf Basis der Adressen die Bauakten nach entsprechenden Angaben durch Praktikanten der Stadt Oldenburg untersucht.
- Die Kaufpreissammlung des Gutachterausschusses für Grundstücksbewertung beinhaltet detaillierte Informationen zu verkauften und bewerteten Grundstücken und Gebäuden. Im Untersuchungsgebiet liegen diese Akten digital und tabellarisch vor.
- Zur statistischen Überprüfung der ermittelten Baujahre wird die fortgeschriebene Gebäude- und Wohnraumstatistik der Volkszählung 1987 des LSKN verwendet. Einen jährlichen Vergleich bietet die Baufertigstellungstatistik seit ca. 1980. Die Daten der öffentlichen Landesbrandkasse sind differenziert nach Postleitzahlen und machen einen räumlich feineren Vergleich möglich.

5.1.2. Klassifizierungen des Baujahres

Es gibt für den Wohnungsbau in Deutschland seit 1918 verschiedene Perioden, in denen die Baustruktur als weitgehend homogen angenommen wird. Die Perioden werden als Baualtersklassen bezeichnet. Ihre Abgrenzung begründet sich mit den zur jeweiligen Zeit herrschenden Rahmenbedingungen aufgrund von Normen oder Gesetzen wie beispielsweise den Novellen der Wärmeschutzverordnungen. Diese führten dazu, dass sich Baumaterialien und Bauweisen und damit auch die Qualität der thermischen Gebäudehüllen änderten (siehe Tabelle 5.2).

Das IWU differenziert den Gebäudebestand in 10 Baualtersklassen, wobei sich die Klassen A und B auf den Zeitraum vor 1918 beziehen [53]. Grundlage für die Einteilung der Baualtersklassen nach 1979 sind die Änderungen der Wärmeschutzverordnungen, die für den Wohnungsbau gesetzliche Mindeststandards angeben [57, S. 23]. Bei der Analyse von Verbräuchen für die Raumwärme im Geschosswohnungsbau wird eine vereinfachte Klassifizierung verwendet (siehe Tabelle 5.2, Spalte 4) [70, 33, 77]. Im weiteren Verlauf der Arbeit bezieht sich der Begriff „Baualtersklassen“ auf die differenziertere IWU-Typisierung, soweit nicht anders vermerkt. Gebäude ab 2009 werden nicht zum Gebäudebestand gerechnet, sondern als Neubau bezeichnet.

Tabelle 5.2.: Baualtersklassen von Wohngebäuden nach IWU [26] (erste Spalte) und verschiedener Wärmeverbrauchsstudien [70, 33, 77] (vierte Spalte)

IWU-Bez.	Zeitraum	Ereignisse zu Beginn der Baualtersklasse	Vereinf. Klassifiz.
A / B	- 1918		
C	1919 - 1948	Gründung der Weimarer Republik	vor WSV082
D	1949 - 1956	Gründung der Bundesrepublik Deutschland	
E	1957 - 1968	Einführung der DIN 4108	
F	1969 - 1978	Ergänzung der DIN 4108	
G	1979 - 1983	Inkrafttreten der ersten Wärmeschutzverordnung (01.11.1977)	
H	1984 - 1994	Inkrafttreten der zweiten Wärmeschutzverordnung (01.01.1984)	WSV082
I	1995 - 2001	Inkrafttreten der dritten Wärmeschutzverordnung (01.01.1995)	WSV095
J	2002 - 2006	Inkrafttreten der Energieeinsparverordnung (EnEV 2002)	EnEV
	2007 - 2008	Inkrafttreten der Novellierung der Energieeinsparverordnung (EnEV 2007)	
	2009 -	Inkrafttreten der Novellierung der Energieeinsparverordnung (EnEV 2009)	

5.1.3. Bestimmung der Datenqualität aus dem Vergleich der Quellen

Da keine ausführliche Referenzquelle existiert, deren Korrektheit für die verschiedenen Baujahre verifiziert ist, wird die Qualität der Datenquellen nur durch deren Gegenüberstellung eingeschätzt. Die einzige Quelle mit manuell kontrollierten Angaben ist das Archiv des Bauamts. Dabei handelt es sich jedoch nur um ca. 100 Angaben.

Die Gegenüberstellung aller anderen Quellen findet jeweils paarweise statt. In Abbildung 5.2 entspricht jeder Datenpunkt einem Gebäude, für das eine Baujahresangabe aus den beiden an den Achsen vermerkten Quellen existiert. Es sind nicht alle Kombinationen berücksichtigt, sondern nur der Vergleich gegenüber den Daten der Kaufpreissammlung des Gutachterausschusses. Die Datensätze dieser Quelle sind zeitlich relativ gleichmäßig verteilt (siehe Kapitel 4.1.7) und aufgrund der Art der Datenerhebung kann von einer hinreichenden Korrektheit der Daten ausgegangen werden.

Die Darstellung der Datenpunkte ist leicht transparent, um eine Häufung von übereinanderliegenden Datenpunkten sichtbar zu machen. Bei allen vier Unterabbildungen ist die Häufung der Datenpunkte entlang der $x=y$ -Geraden deutlich sichtbar. Für die spätere Un-

5. Berechnung und Analyse von Gebäudeparametern in Oldenburg

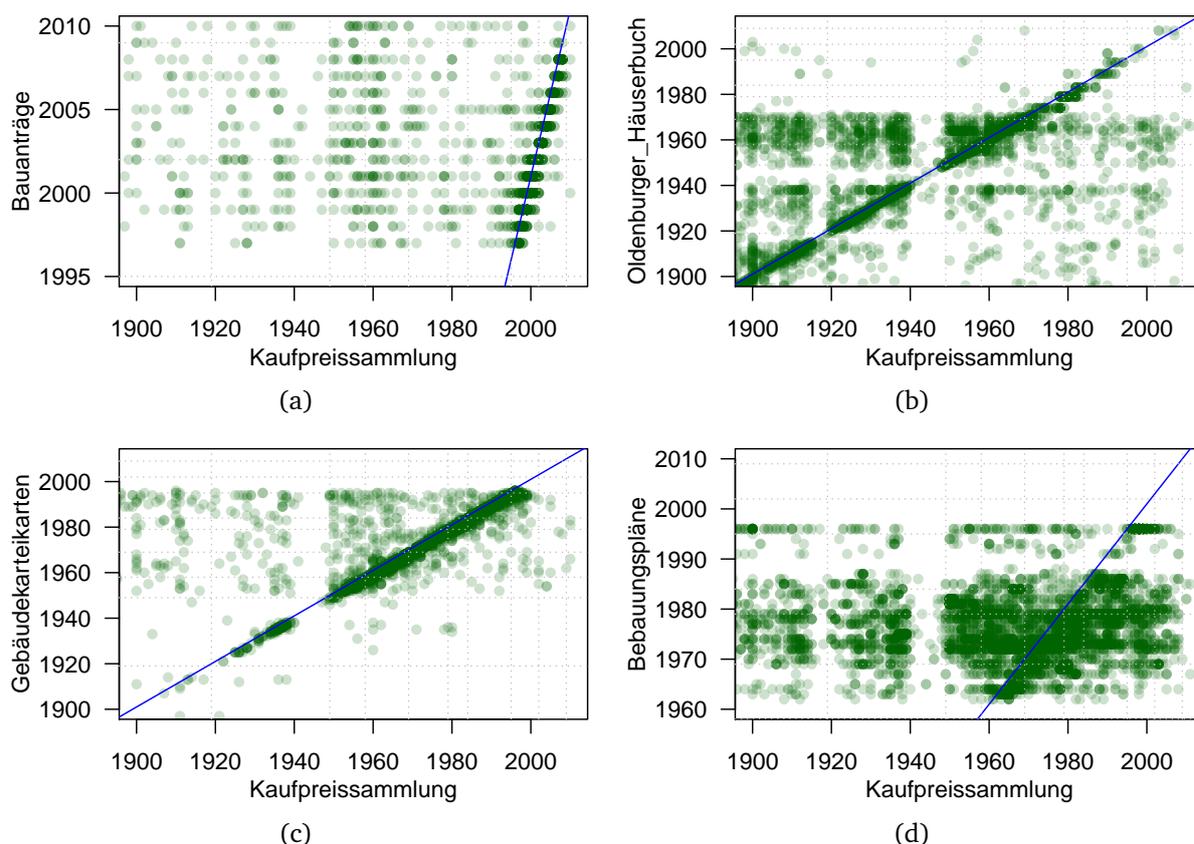


Abbildung 5.2.: Paarweiser Vergleich der Quellenangaben zum Gebäudebaujahr verschiedener Quellen mit Angabe der Winkelhalbierenden in blau und den IWU-Baualtersklassen als Gitterstruktur

tersuchung ist jedoch nicht das genaue Baujahr entscheidend, sondern die Baualtersklasse. Die Baualtersklassen werden durch die grau gepunkteten Gitter gekennzeichnet.

Auffallend viele Gebäude mit Baujahren vor 1970 in beiden Quellen in Abbildung 5.2(b) werden im Oldenburger Häuserbuch mit einem späteren Baujahr geführt, als in den Daten des Gutachterausschusses. Da die Daten der Kaufpreissammlung jedoch erst ab 1980 aufgenommen wurden, wird in der Arbeit davon ausgegangen, dass beim Auslesen der Daten des Häuserbuchs Erweiterungen oder Veränderungen der Gebäude als Neubau verwechselt wurden und die Angaben damit unzuverlässiger sind. Für die Angaben aus Bebauungsplänen und der Kaufpreissammlung ist eine schlechte Übereinstimmung zu erkennen (siehe Abbildung 5.2(d)). Dies wird nachfolgend näher diskutiert.

Die von der Winkelhalbierenden abweichenden Datenpunkte beim Vergleich der Datensätze der Bauanträge mit denen der Kaufpreissammlung repräsentieren Gebäude, die nach dem Zeitpunkt ihres Verkaufs abgerissen wurden (siehe Abbildung 5.2(a)). Wenn auf dem Grundstück mit der gleichen Adresse ein Neubau stattfand, wurde dieser durch die Bauanträge erfasst.

Für alle möglichen paarweisen Kombinationen der fünf Datenquellen gibt die Tabelle 5.3

Tabelle 5.3.: Anzahl und prozentualer Anteil der paarweise übereinstimmenden Baualtersklassen

absolut \ %	Bauanträge	Gebäude- karteikarten	Bebauungs- pläne	Oldbg. Häuser- buch	Kaufpreis- sammlung
Bauanträge		4 %	19 %	6 %	59 %
Gebäudekarteikarten	11		37 %	62 %	75 %
Bebauungspläne	345	2.003		18 %	25 %
Oldbg. Häuserbuch	32	1.814	847		76 %
Kaufpreissammlung	831	3.082	1.456	3.313	

die absolute Zahl und den prozentualen Anteil der paarweise übereinstimmenden Baualtersklassen an. Dies entspricht den Punkten aus Abbildung 5.2, die in den Gitterzellen unter der $x=y$ -Geraden liegen. Die Übereinstimmung zwischen miteinander vergleichbaren Datenquellen liegt bei ca. 60 %-76 %. Allein der Vergleich mit den Bebauungsplänen fällt deutlich schlechter aus.

Die geringste Übereinstimmung ergibt sich zwischen den Bauanträgen und den Gebäudekarteikarten mit 4 % und dem Oldenburger Häuserbuch mit 6 %. Dies hat jedoch systemische Gründe, da diese Quellen faktisch disjunkt sind. Das heißt, nur für 11 bzw. 32 Gebäude beinhalten jeweils beide Quellen eine Angabe. Wie auch in Abbildung 5.2 sichtbar, beginnen Angaben zu den Bauanträgen erst nach 1996. Ab diesem Zeitpunkt werden sie digital erfasst. Dagegen ist das Ziel der historischen Aufarbeitung, möglichst alte Gebäude zu erfassen. Die meisten Angaben der Gebäudekarteikarten stammen ebenfalls aus dem Zeitraum vor 1995.

Baujahre aus Bebauungsplänen Beim Vergleich der Bebauungspläne mit den anderen vier Datenquellen in Tabelle 5.3 fallen die durchweg niedrigen Anteile übereinstimmender Baualtersklassen (zwischen 18 % und 37 %) auf.

Eine etwas genauere Sichtweise bietet die Tabelle 5.4. Sie zeigt, bei wie vielen Bebauungsplänen ein gewisser Prozentsatz an übereinstimmenden Angaben je Bebauungsplan erreicht wird. Bei 145 Bebauungsplänen ist der Anteil der übereinstimmenden Angaben unter 10 %. Bei 35 Bebauungsplänen ist der Anteil zwischen 90 % und 100 %. Bei näherer Analyse erkennt man, dass die Bebauungspläne mit hoher Übereinstimmung (>70,%) nur sehr kleine Gebiete umfassen. Aufgrund dieser schlechten Übereinstimmung der Bebauungspläne mit den anderen Datenquellen werden die Informationen der Bebauungspläne in der Arbeit nicht genutzt.

5.1.4. Schätzung nicht bekannter Gebäudebaujahre

Aus allen Datenquellen gemeinsam wird für ca. 76 % der Gebäude ein Baujahr ermittelt. Für die Berechnung des Wärmebedarfs ist jedoch die Angabe einer Baualtersklasse bei

Tabelle 5.4.: Häufigkeitsverteilung der Bebauungspläne je Übereinstimmungswahrscheinlichkeit der Gebäudebaujahre von Bebauungsplänen und gebäudescharfen Datenquellen

Klasse (bis):	10 %	20 %	30 %	40 %	50 %	60 %	70 %	80 %	90 %	100 %
Häufigkeit:	145	32	16	24	27	5	4	8	5	35

allen Gebäuden notwendig. Das Baujahr korreliert nicht mit anderen Gebäudeparametern, weshalb eine Klassifizierung auf Basis anderer nicht erfolgreich ist. Die Gebäude ohne Baujahr sind relativ gleichmäßig im Untersuchungsgebiet verteilt und somit nahezu immer von anderen Gebäude mit einer Angabe umgeben.

Für das Schätzverfahren wird davon ausgegangen, dass die Verteilung der Baualtersklassen von Gebäuden in einem lokalen Gebiet rund um das „Zielgebäude“, das Gebäude ohne Angabe, repräsentativ für das jeweilige Gebiet ist. Die Gebäude mit einer Angabe werden im Folgenden „Referenzgebäude“ genannt. Als lokales Gebiet werden statistische Gliederungsbezirke der Blockebene gewählt. Dies ist die zweitkleinste Gliederungsebene und beinhaltet im Mittel 80 Gebäude pro Block. Alternativ ist die Möglichkeit implementiert, die Gebäude in einem zu bestimmenden Radius um das Zielgebäude für die Bestimmung der lokalen Verteilung zu wählen.

Aus den Baualtersklassen der Referenzgebäude wird eine kumulierte Wahrscheinlichkeitsverteilung der Klassen im Intervall zwischen Null und Eins berechnet. Hieraus wird durch das Ziehen einer Zufallszahl in diesem Intervall eine Klasse ausgewählt. Jede Baualtersklasse wird durch ein Baujahr repräsentiert. Dieses Repräsentativ-Baujahr der zufällig ausgewählten Baualtersklasse wird dem Zielgebäude zugewiesen.

Für die Berechnung der Wahrscheinlichkeitsverteilung werden nur Angaben von Datenquellen mit einem Güteindikator größer fünf gewählt, damit die Schätzung auf möglichst fehlerfreien Angaben basiert. Beim wiederholten Ziehen einer Zufallszahl wird die aus den Referenzgebäuden abgeleitete Wahrscheinlichkeitsverteilung nachgebildet. Das Verfahren wird für die ca. 10.000 Gebäude ohne Angaben durchgeführt.

Die Baualtersklassen I und J, d. h. Neubauten ab 1995 werden bei der Schätzung nicht berücksichtigt. Es kann aufgrund der Datenquelle Bauanträge und der Angaben aus den anderen Quellen davon ausgegangen werden, dass die Baujahre aller Gebäude aus diesen Zeitperioden bekannt sind. Diese Annahme bestätigt ebenfalls der Vergleich mit der Neubaustatistik der Landesamt für Statistik und Kommunikation Niedersachsen (LSKN) (siehe Abbildung 5.5).

5.1.5. Analyse und Kontrolle der Gebäudebaujahresverteilungen

Aus den Angaben der oben beschriebenen und als verwendbar definierten Datenquellen wird jedem Gebäude ein Baujahr zugewiesen. Für ca. 25 % der Gebäude existieren

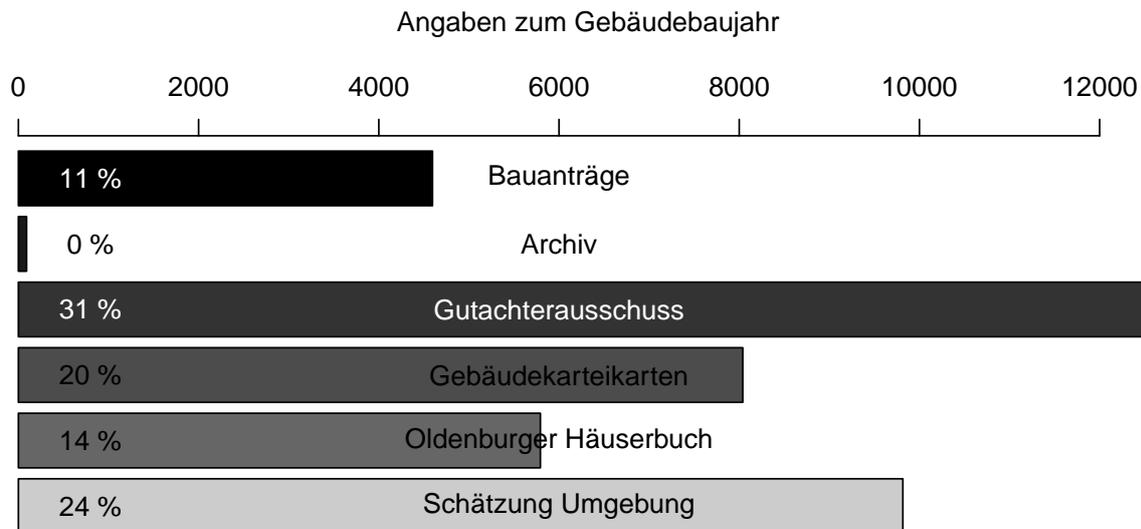


Abbildung 5.3.: Häufigkeitsverteilung der konfliktbereinigten Angaben zum Gebäudebaujahr je Datenquelle

mehrere Angaben pro Gebäude. Aus diesen wird mit der in Kapitel 3.1.2 beschriebenen Methode eine Angabe ausgewählt.

Abbildung 5.3 gibt die Zahl der Baujahresangaben je Datenquelle und deren prozentualen Anteil am gesamten Hauptgebäudebestand an. Die Datenquellen sind anhand des Güteindicators von oben nach unten sortiert. Die Quelle „Bauanträge“, die nahezu alle Gebäude seit 1996 abdeckt, enthält ein Baujahr für 11 % der Gebäude. Ca. 100 Angaben entstammen der manuellen Recherche im Archiv des Bauordnungsamtes. Das sind weniger als 0,5 %. Die Kaufpreissammlung des Gutachterausschusses war die ergiebigste Quelle mit über 12.000 Baujahren bzw. einem Anteil von 31 %.

Insgesamt liefern städtische Daten, bestehend aus den Bauanträgen und den Gebäudekarteikarten, Informationen für 30 %, das sind ca. 10.000 Gebäude. Die Bebauungspläne liefern keinen Beitrag, da sie wegen der oben beschriebenen schlechten Qualität nicht berücksichtigt sind. Insgesamt können für ca. 76 % der Hauptgebäude, dies entspricht 31.294 Gebäuden, ein Baujahr aus einer der Datenquellen bestimmt werden. Daher muss bei 24 % der Gebäude ein Baujahr geschätzt werden.

Für die Darstellung in Abbildung 5.4 werden die Baujahre in die oben beschriebenen neun Baualtersklassen nach der IWU-Klassifizierung zusammengefasst. Die Höhe der Säulen gibt die absolute Anzahl der Gebäude an, die im entsprechenden Zeitraum neu errichtet und bis 2008 nicht wieder abgerissen wurden. Die Grau-Schattierung zeigt, aus welchen Quellen die Angaben für eine Baualtersklasse stammen. Die rote und blaue Linie repräsentiert die statistischen Daten des Zensus 1987 und der Öffentlichen Brandkasse.

Der höchste Zuwachs im Gebäudebestand fand in den Baualtersklassen E und F, also zwischen 1958 und 1978, sowie Anfang der 90er Jahre statt. Bei einem Vergleich der

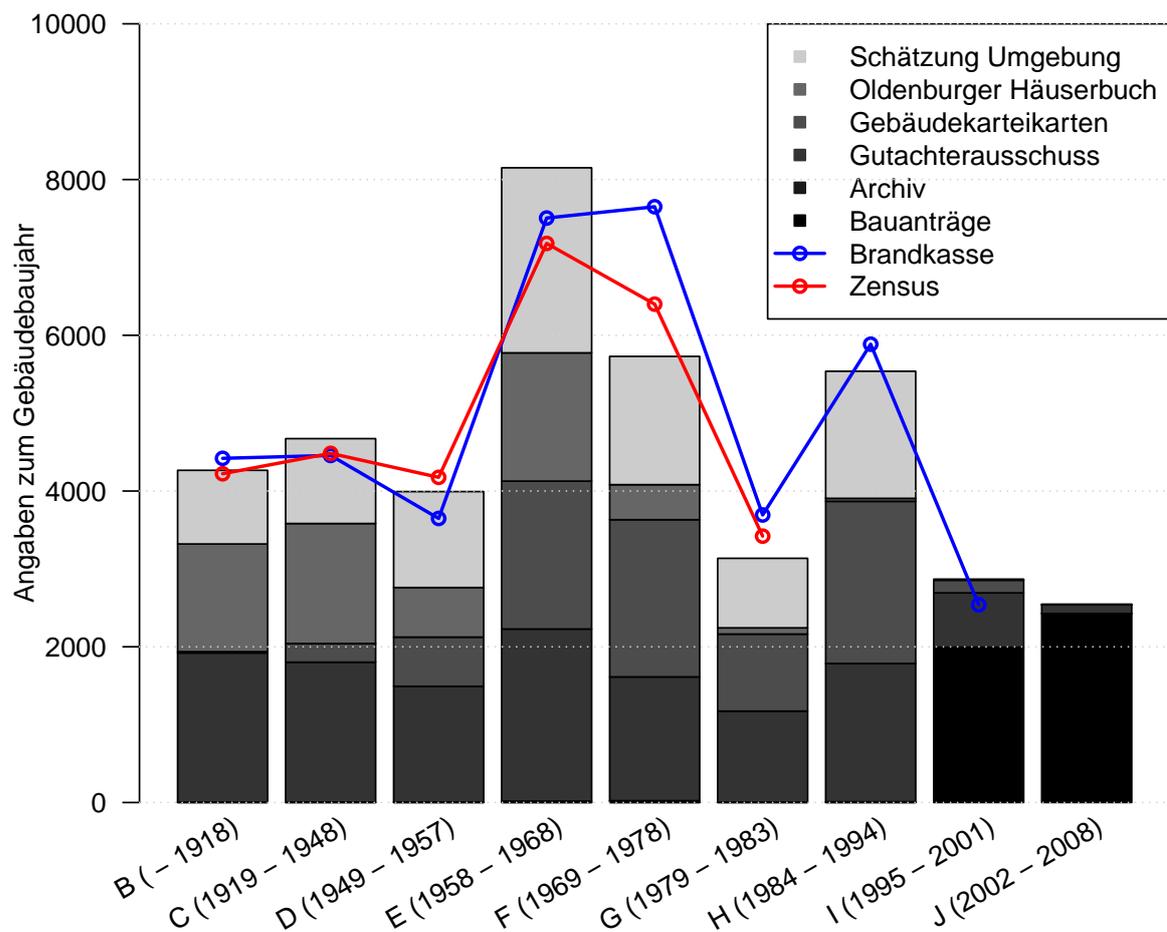


Abbildung 5.4.: Häufigkeitsverteilung der konfliktbereinigten Angaben zum Gebäudebaujahr je Baualterklasse und Datenquelle

Balken muss jedoch beachtet werden, dass die Baualterklassen keine äquidistanten Zeiträume bilden. In jedem Balken sind die Angaben der Quellen nach dem Güteindikator von unten nach oben sortiert. Hierbei ist auch die zeitliche Verteilung der Angaben verschiedener Quellen erkennbar. Angaben der Quelle Bauanträge betreffen nur die Jahre nach 1996, d. h. in den Baualterklassen I und J. Angaben des Oldenburger Häuserbuchs betreffen überwiegend ältere Gebäude bis 1968. Die Angaben des Gutachterausschusses sind zeitlich relativ gleich verteilt. Bei I und J wird dies deshalb nicht sichtbar, da die Bauanträge einen höheren Güteindikator haben und daher als Baujahresangaben ausgewählt werden. Die Zahl der Angaben der Umgebungsschätzung ist proportional zur jeweiligen Anzahl der Angaben aus den Datenquellen.

Insbesondere für die Baualterklassen B bis D stimmen die Angaben der Datenbank und die statistischen Daten sehr gut überein. In der Baualterklasse E wird die Zahl der Gebäude bezogen auf die Zensus-Angaben um ca. 13 % überschätzt, dagegen in den Klassen F um ca. 11 % und G um ca. 9 % unterschätzt. Dies kann damit begründet werden, dass diese Baualterklassen in den verfügbaren Datenquellen ebenfalls über- bzw.

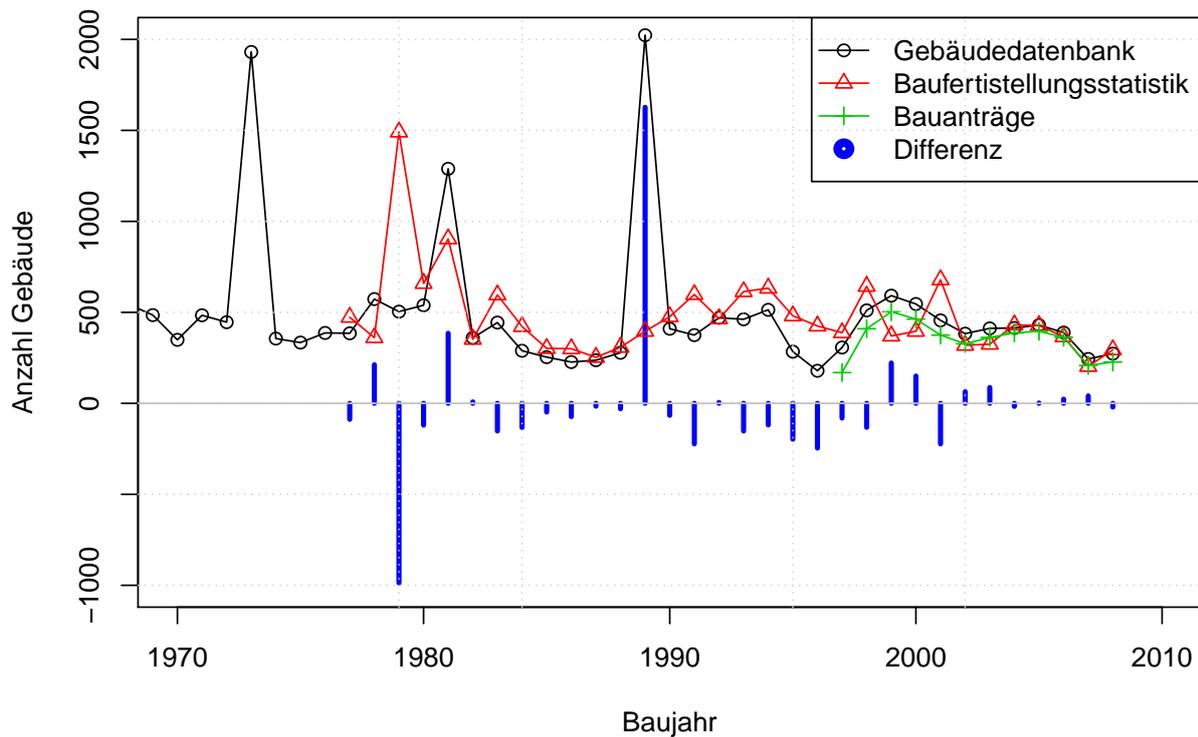


Abbildung 5.5.: Zeitreihe der Neubauten mit jährlicher Auflösung seit 1970 und Literaturvergleich

unterrepräsentiert sind, die Schätzung jedoch durch die bekannten Baujahre geprägt ist.

Ab der Baualterklasse H sind keine Angaben aus dem Zensus mehr verfügbar, da dieser bereits 1987 statt fand. Die Originaldaten der öffentlichen Brandkasse enthalten zwar nicht alle Gebäude, dafür jedoch eine sehr hohe und statistisch aussagefähige Anzahl. Diese wird auf den gesamten Hauptgebäudebestand hochgerechnet. Der prozentuale Anteil der Gebäude je Baualtersklasse bleibt hierbei gleich.

Darstellung der historischen Stadtentwicklung Da die Gebäude einen Raumbezug besitzen, kann die beschriebene Entwicklung in Karten dargestellt werden. Dies macht es möglich, anhand des Baujahres die historische Entwicklung der Stadt Oldenburg zu visualisieren. In den Abbildung A.1 bis A.8 im Anhang ab Seite 128 ist der Oldenburger Gebäudebestand im jeweils letzten Jahr der IWU-Baualtersklassen als Karte abgebildet. Die Karten enthalten auch die geschätzten Baujahre. Bei Gebäuden, die in der Vergangenheit abgerissen und neu gebaut wurden, wird nur der Neubau dargestellt.

Baufertigungsstatistik Ab dem Jahr 1977 können die Angaben der Gebäudedatenbank mit den jährlichen Angaben der Baufertigstellungsstatistik des LSKN abgeglichen werden, dargestellt in Abbildung 5.5. Die drei Peaks bei der Anzahl der Gebäude der Datenbank entstehen aufgrund der Schätzverfahren für Gebäude ohne Quellenangaben

(siehe Kapitel 5.1.4). Zum Teil erklären sich die Schwankungen in der Statistik durch verspätete Meldungen an das Landesamt. Der Peak in 1978 kann mit der für 1979 in Kraft getretenen Energieeinsparverordnung erklärt werden. Es ist zu vermuten, dass viele Bauwillige einen Bauantrag noch vor deren Inkrafttreten gestellt haben. Der Vergleich der ermittelten Baujahre mit der Statistik zeigt trotz hoher Abweichungen in einzelnen Jahren eine relativ gute Übereinstimmung bei der zeitlichen Mittelung. Summiert ergeben sich für den Vergleichszeitraum von 1978 bis 2010 eine Abweichung von 210 Gebäuden beim Vergleich der LSKN-Statistik mit den Daten der Gebäudedatenbank. Das entspricht einem Anteil von 1,3 %.

5.2. Gebäudeklasse

Die Gebäudeklassen unterscheiden den Gebäudebestand anhand der Zahl der Parteien und der Hauptnutzungsform je Gebäude. Für Gebäude mit ausschließlich oder überwiegend privater Nutzung entsprechen die Klassen der Klassifikation des IWU [53]. Hinzu kommt die Kategorie Nichtwohngebäude (NWG) für Gebäude, die keine Wohnnutzung oder eine überwiegend gewerbliche Nutzung aufweisen. Eine Übersicht bietet Tabelle 5.5.

Tabelle 5.5.: Gebäudeklassen und Klassenbeschreibung

Klassenname	Beschreibung
EFH/ZFH	Ein- und Zweifamilienhäuser
RH	Reihenhäuser, in Reihe gebaute (mind. 3 Gebäude) und jeweils angrenzende Ein- und Zweifamilienhäuser
MFHkl	Kleine Mehrfamilienhäuser und Gebäude mit Mischnutzung (Wohnen und GHD) mit bis zu neun Parteien
MFHgr	Große Mehrfamilienhäuser und Gebäude mit Mischnutzung ab zehn Parteien
NWG	Nichtwohngebäude mit einer beliebigen Anzahl an Parteien

5.2.1. Bestimmung der Gebäudeklasse

Es erfolgt eine Einteilung des Gebäudebestandes in fünf Gebäudeklassen mithilfe eines regelbasierten Verfahrens. Für die Zuordnung dient die Gebäudenutzung der ALK, die Stromvertragszahlen und -nutzungsarten des Stromnetzbetreibers und die Lage der Gebäude. Bei Gebäuden ohne Stromvertragsinformationen, wird die Unterscheidung auf Basis der Gebäudevolumina getroffen. Abbildung 5.6 gibt die Regeln und Vorgehensweise an. Die regelbasierte Klassifikation ist gut implementierbar, leicht modifizierbar und jederzeit nachvollziehbar. Dies sind Vorteile gegenüber Regressionsansätzen oder Fuzzy-Klassifizierungen [56].

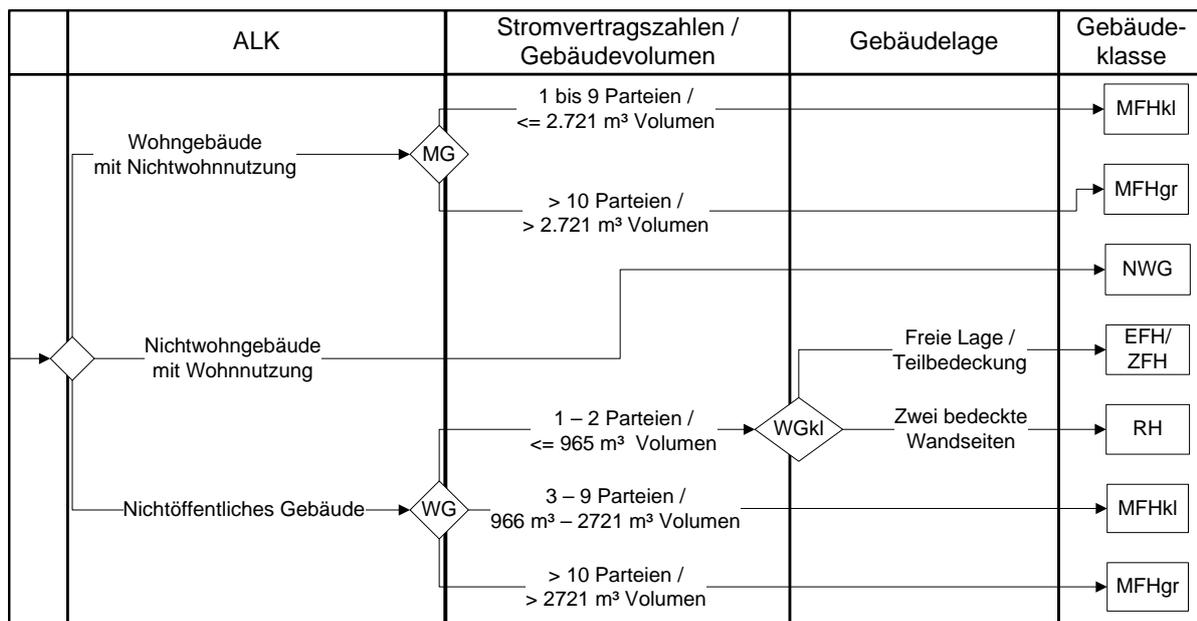


Abbildung 5.6.: Entscheidungsbaum zur Bestimmung der Gebäudeklasse auf Basis der ALK-Nutzungsschlüssel, der EWE-Netz-Stromvertragskategorisierungen und der relativen Gebäuelage

Der erste Klassifikationsschritt unterscheidet die Gebäude in Wohngebäude (WG), mischgenutzte Gebäude (MG) und NWG. Die ALK in Niedersachsen beinhaltet für jeden Gebäudegrundriß eine Objektartangabe (siehe Kapitel 4.1.4). Öffentliche Gebäude werden relativ fein unterschieden, private und gewerbliche Nutzungen nur sehr grob. Die Differenzierung der Objektartnutzungen ist für eine energiewirtschaftliche Analyse nicht ausreichend. Für eine Klassifikation in WG, MG und NWG sind die Angaben jedoch sehr hilfreich, da die Objektarten eine entsprechende Einteilung bereits beinhalten. In der Arbeit werden die Objektarten mit der Teilbezeichnung „nichtöffentliche Gebäude“ (Objektartschlüssel 0931 und 0932) als WG klassifiziert. MG sind „Wohngebäude mit ...“ einer Nichtwohnnutzung, die „Öffentlich“, „Handel und Dienstleistung“, oder „Gewerbe und Industrie“ (Objektartschlüssel 2111, 2121, 2131) sein kann. NWG sind dagegen die drei Nichtwohnnutzungen „... mit Wohnen“ (Objektartschlüssel 2141, 2151, 2161) und alle restlichen Typen. Eine Übersicht der Objektartschlüssel gibt Tabelle 5.8 im Kapitel 5.6.

Für die Klassifikation von Stromkunden in verschiedene Kundentypen besitzt jeder Stromvertrag des Stromnetzbetreibers eine Angabe zur Verbraucherklasse. Diese Klassifizierung ist genauer als die Objektarten der ALK und unterscheidet vor allem verschiedene Gewerbebereiche. Die Zahl der Stromverträge erlaubt darüber hinaus eine Abschätzung, wie viele Parteien in einem Gebäude sind. Dabei wird angenommen, dass je Partei ein Stromvertrag existiert. In Mehrfamilienhäusern (MFH) gibt es zum Teil einen Gemeinschaftsanschluss (z. B. für die Beleuchtung gemeinsam genutzter Flächen). Dieser Anschluss hat jedoch eine eigene Nutzungsklasse und wird bei der Bestimmung der Parteien nicht berücksichtigt. WG werden auf dieser Basis in kleine Wohngebäude (WGkl) mit maximal zwei Parteien, kleine Mehrfamilienhäuser (MFHkl) mit bis zu neun Parteien

5. Berechnung und Analyse von Gebäudeparametern in Oldenburg

und große Mehrfamilienhäuser (MFHgr) mit mindestens zehn Parteien eingeteilt. MG werden nur in MFHkl und MFHgr anhand der Grenze von neun Parteien unterschieden. Bei NWG findet keine weitere Unterscheidung statt, da die Gebäudetypen vor allem für Analysen im Sektor Haushalte dienen. Der Sektor GHD ist sehr heterogen, wodurch eine Gewerbegebäudetypisierung erschwert wird [68]. Unterscheidungen hier werden anhand der spezifischen Gebäude- bzw. Zonennutzungen (siehe Kapitel 5.6) getroffen.

Für knapp 1.000 Gebäude sind keine Stromvertragsdaten verfügbar. Deshalb werden diese anhand des Gebäudevolumens klassifiziert. Die Klassengrenzen werden auf Basis der Gebäudevolumina der bereits klassifizierten Gebäude mit Stromvertragsdaten bestimmt. Für jede Gebäudeklasse wird der Median der Gebäudevolumina berechnet. Die Mitte zwischen den Medianen zweier Klassen wird als Grenzwert festgelegt. Dies ergibt 965 m^3 für die Unterscheidung von WGkl und MFHkl sowie 2721 m^3 für die Unterscheidung von MFHkl und MFHgr. Die Verteilung der Gebäudevolumina je Gebäudeklasse ist in Abbildung 5.8(a) gezeigt.

Die WGkl werden darüber hinaus in Reihenhäuser (RH) und Ein- oder Zweifamilienhäuser (EFH/ZFH) je nach Nachbarschaftsverhältnis der Gebäude unterschieden. Auf Basis der ALK-Gebäude-Grundflächenpolygone wird mithilfe der GIS-Funktionalität der Datenbank analysiert, ob zwei Gebäude direkt aneinander grenzen. Falls ein Gebäude, das zuvor als WGkl klassifiziert wurde, zwei direkt angrenzende andere WGkl aufweist, wird das Gebäude der Klasse RH zugeordnet; alle weiteren WGkl, die ein RH berühren, ebenfalls. Falls die Bedingungen nicht zutreffen, gehört es dem Typ EFH/ZFH an. Zum Teil sind die Grundflächenpolygone der ALK fehlerhaft, d. h. zwei Grundflächen berühren sich nicht, sondern haben einen Abstand von ca. $0,00005 \text{ m}$. Dieser Abstand ist in der Realität jedoch nicht existent. Diese Fälle werden mit Hilfe eines räumlichen Puffers ausgeglichen.

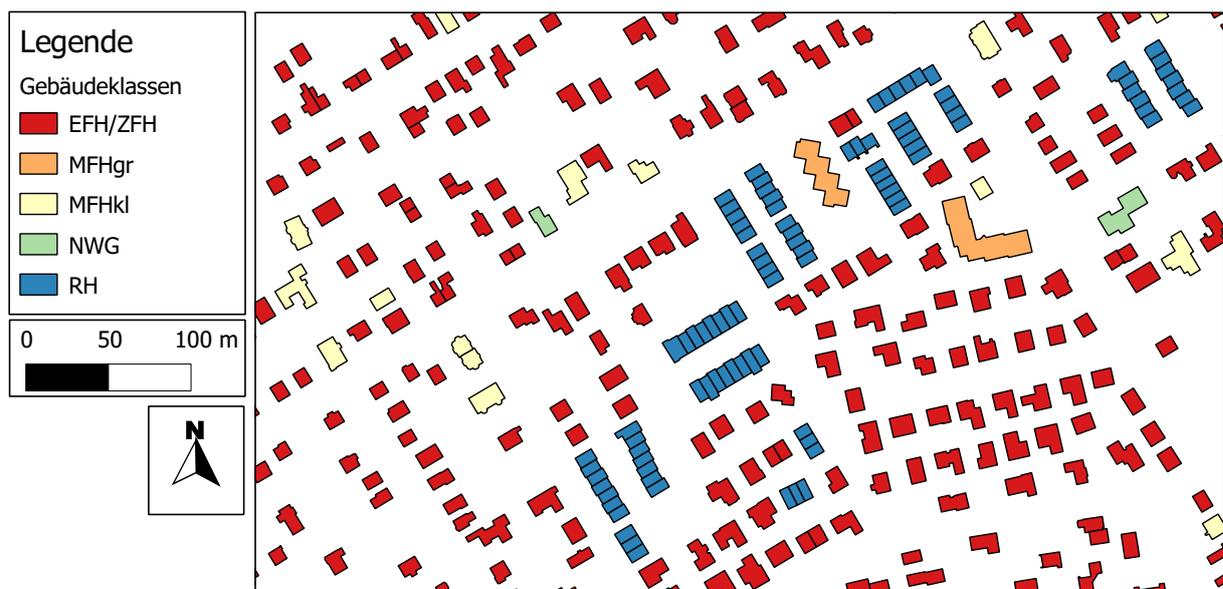


Abbildung 5.7.: Kartenausschnitt mit farblich klassifizierten Gebäuden in Oldenburg

5.2.2. Analyse der Gebäudeklassenverteilung

EFH/ZFH stellen mit über 28.000 Gebäude den größten Anteil der ca. 41.000 Hauptgebäude in Oldenburg dar. Hinzu kommen ca. 3.800 Reihenhäuser. Es gibt 5.763 MFHkl und 1.262 MFHgr. Beide Typen bestehen aus Gebäuden mit reiner Wohnnutzung und mit Mischnutzung. Überwiegend gewerblich genutzt werden ca. 2.500 Gebäude. Abbildung 5.7 veranschaulicht die Gebäudeklassifikation durch die farbliche Darstellung der verschiedenen Klassen. Die Karte zeigt einen überwiegend mit Ein- und Zweifamilienhäusern bebauten nordöstlichen Ausschnitt von Oldenburg. Die Reihenhäuser sind in blau gekennzeichnet.

In Abbildung 5.8 wird die gefundene Zuordnung der Gebäude in Gebäudeklassen mit weiteren Gebäudeparametern verglichen, die für eine Einteilung in Frage kommen wür-

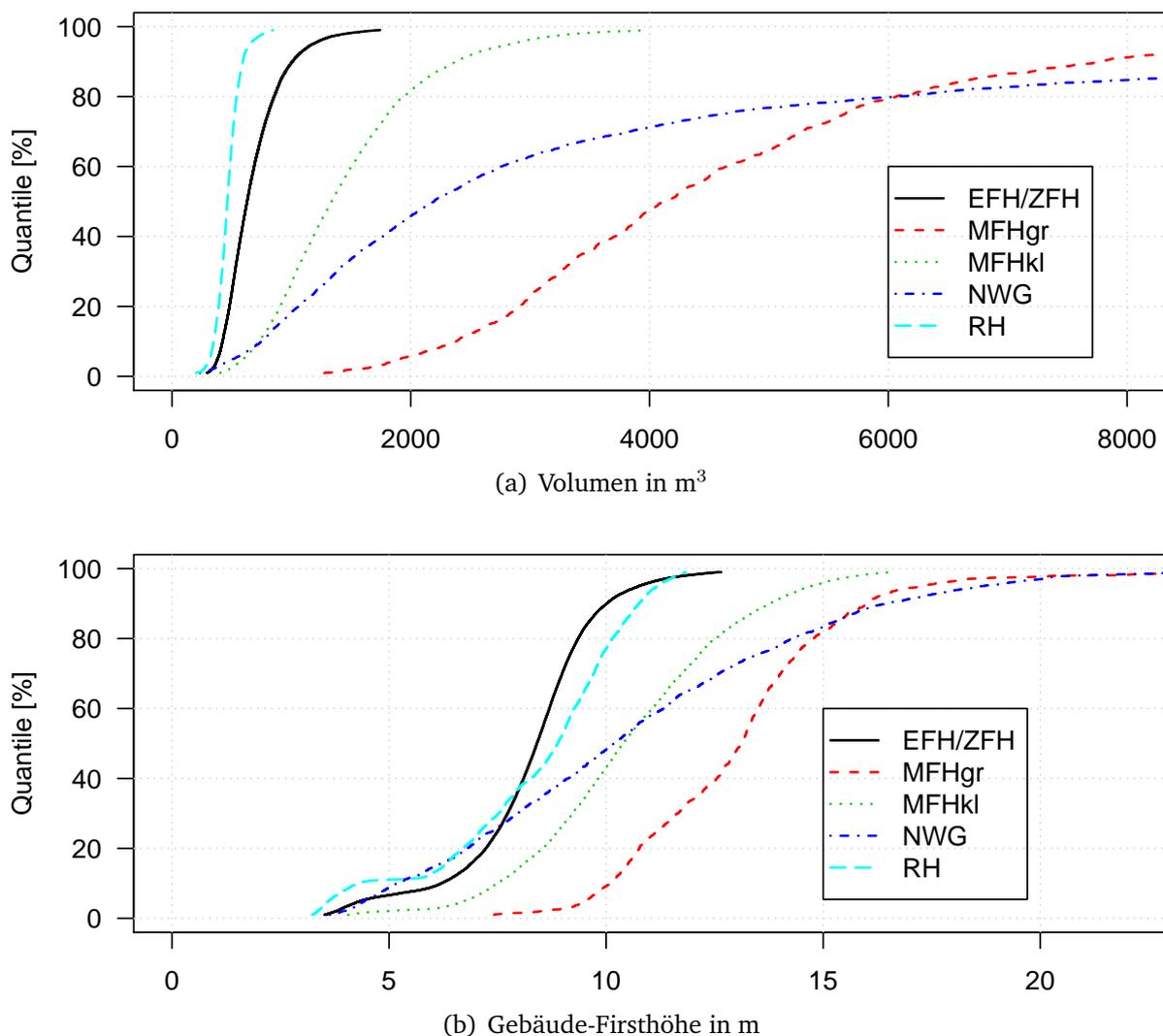


Abbildung 5.8.: Kumulierte Häufigkeitsverteilung der (a) Gebäudevolumina und (b) Gebäudehöhe je Gebäudeklasse

den. Als Darstellung wird ein sogenannter Quantilplot gewählt. Auf der X-Achse ist der Wertebereich angetragen, den der jeweilige Gebäudeparameter annimmt. Die Y-Achse zeigt die kumulierte Wahrscheinlichkeit zwischen 0 % und 100 % als Funktion der Gebäudeparameter. Dadurch lassen sich Quantile, wie beispielsweise der Median direkt ablesen. Je steiler eine Kurve verläuft, desto mehr Datenpunkte existieren bei einem entsprechenden Parameterwert. Die Kurven können gut miteinander verglichen werden, indem die entsprechenden Parameterwerte zu einem festen Quantilwert, z. B. 50 %, verglichen werden. Eine s-förmige Kurve lässt auf normalverteilte Werte schließen.

Die Verteilungskurven der Volumina je Gebäudeklasse in Abbildung 5.8(a) sind klar voneinander getrennt. Da sich RH und EFH/ZFH nur aufgrund der Gebäudelage unterscheiden, liegen die Verteilungen nah beieinander. Für MFHkl und MFHgr sind die Volumenverteilungen zu höheren Werten hin verschoben. Ein Vergleich der Kurven für ein bestimmtes Gebäudevolumen zeigt jedoch die Überlappung der Verteilungen. Beispielsweise weisen die obersten 20 % der MFHkl ebenso ein Volumen zwischen 2.000 und 4.000 m³ auf wie die unteren 50 % der MFHgr. Eine Klassifizierung auf Basis der Volumina wäre somit nicht eindeutig.

Die Verteilung der Firsthöhen je Gebäudeklasse EFH/ZFH, MFHkl und MFHgr zeigen starke Überlappungen in Abbildung 5.8(b). Insbesondere bei der Unterscheidung von MFHkl und MFHgr lässt die Höhe keine Differenzierung zu. Der Grund hierfür ist, dass sich Mehrfamilienhäuser (MFH) in Oldenburg nicht durch die Höhe, sondern eher durch die Grundflächengröße des Gebäudekomplexes unterscheiden.

NWG werden nicht anhand der Vertragszahlen unterschieden und beinhalten damit Gebäude aller Größenklassen, was aus den Abbildungen 5.8 (a) und (b) deutlich wird. Verteilung der NWG erstreckt sich über alle Volumina und Höhen.

Zusammenfassend gilt, dass anhand der Volumina die gefundene Klassifikation bestätigt werden kann. Aufgrund der Überlappungen der Verteilungen gibt es keine klare - im gewissen Sinne - natürliche Abgrenzung der Gebäudeklassen. In der Literatur werden die Gebäude zum Teil stark vereinfacht anhand ihrer Nutzfläche klassifiziert [2].

5.3. Gebäudetyp

Die in Kapitel 5.1 und 5.2 gebildeten neun Baualters- und fünf Gebäudeklassen werden zu 45 Gebäudetypen kombiniert. Deren Anteil am Gebäudebestand in Oldenburg zeigt Abbildung 5.9 anhand eines Mosaik-Plots. Die Flächen sind proportional zu den Anteilen. Die Prozentzahlen über den Spalten der Gebäudeklassen geben den kumulierten Anteil der jeweiligen Gebäudeklasse an.

Fast 70 % der Gebäude gehören zur Klassen der Ein- und Zweifamilienhäuser. An den Flächen deutlich sichtbar ist der Zuwachs von Gebäuden aller Klassen in den Baualtersklassen E und F (50er und 60er Jahre). Nahezu 80 % der MFHkl und MFHgr sind Gebäude vor 1979, dem Beginn von Baualtersklasse G und der Einführung der ersten Wärmeschutzverordnung, errichtet. Anfang der 90er Jahre wurden vor allem EFH/ZFH neu gebaut.

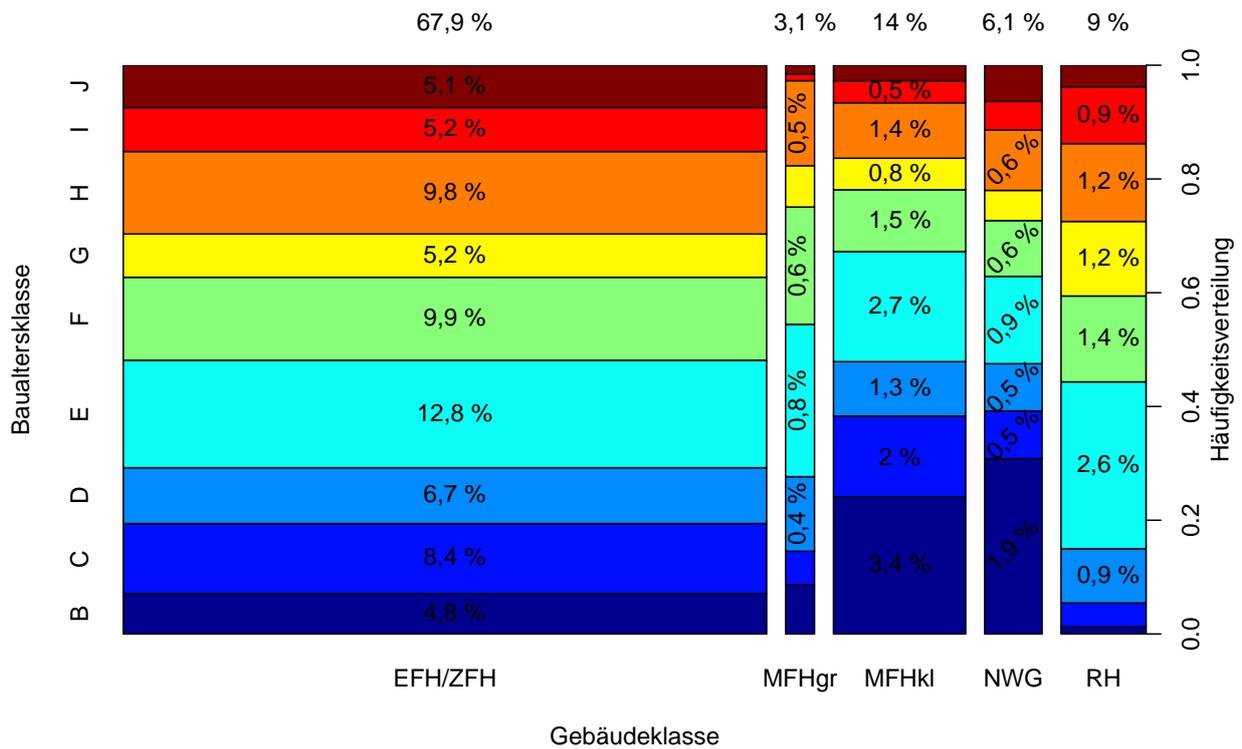


Abbildung 5.9.: Mosaikplot der Baualters- und Gebäudeklassen

5.4. Gebäudehüllflächen und Gebäudevolumen

Die Bauteilgrößen der thermischen Gebäudehülle sind wichtige Parameter bei der Berechnung des Wärmebedarfs durch Bilanzverfahren. Abbildung 5.10 soll die Bauteilbezeichnungen verdeutlichen. Die thermische Gebäudehüllfläche beinhaltet die Grundfläche, Wand-, Fenster- und Dachflächen mit Kontakt zur Außenluft. Türflächen werden vernachlässigt. Wandflächen und Fensterflächen werden gemeinsam als Fassadenflächen bezeichnet. In der Abbildung nicht gezeigt werden Fassadenflächen, die durch angrenzende Gebäude abgedeckt sind, wodurch die thermische Hüllfläche mit Außenluftkontakt verkleinert wird.

5.4.1. Datengrundlage und Vervollständigung

Die Berechnung der Gebäudehüllflächen verwendet die Grundflächengeometrie der Gebäudepolygone und die Angaben zur First- und Traufhöhe der Gebäude. Die Gebäudepolygone sind Bestandteil der ALK. Da nur Gebäude in der Arbeit berücksichtigt werden, die als Polygon in der ALK existieren, ist diese Datenquelle per Definition vollständig. Die Höhenangaben entstammen größtenteils aus Laserscan-Überfliegungen. Durch Erkennungsfehler oder fehlerhafte Daten treten Gebäude ohne Angaben auf. In diesem Fall werden Höhendaten auf Basis der Informationen der Nachbargebäude geschätzt. Das

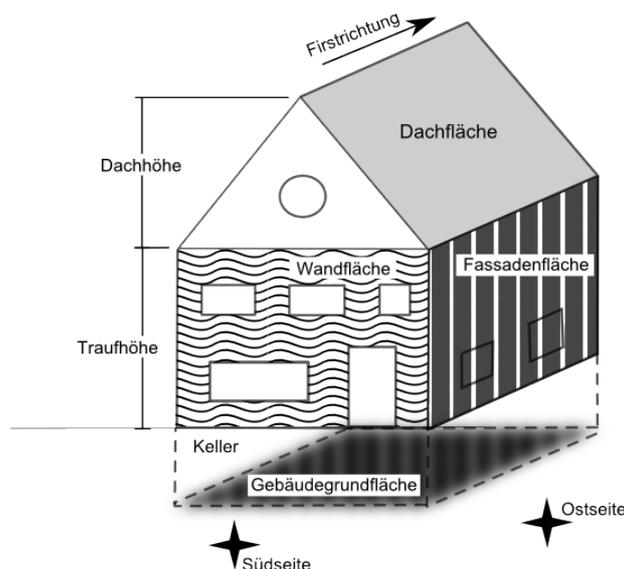


Abbildung 5.10.: Visualisierung der Gebäudehüllen-Parameter

Verfahren ist analog zu der in Kapitel 5.1.4 beschriebenen Methodik. Als Referenzgebäude berücksichtigt werden alle Gebäude der gleichen Relevanz wie das Zielgebäude in einem Umkreis von 80 m um dieses. Aus den Angaben der Trauf- und Firsthöhe der Referenzgebäude wird jeweils für beide Parameter ein gewichteter Mittelwert berechnet. Der Gewichtungsfaktor ist das Verhältnis der Grundflächen zwischen dem Zielgebäude ohne originäre Höhenangabe und dem jeweiligen Referenzgebäuden (siehe Gleichung (5.1)). Der Berechnung liegt die Annahme zugrunde, dass lokal, d. h. in einer relativ kleinräumigen Umgebung gleiche Grundflächen auf ein ähnliches Gebäude und folglich auf ähnliche Gebäudehöhen schließen lassen. Insgesamt werden für ca. 1200 Gebäude die Höhenangaben geschätzt. Falls beim ersten Durchgang mit Radius 80 m keine Referenzgebäude gefunden wird, erfolgt ein zweiter Durchgang mit Radius 300 m.

$$h_Z = \frac{\sum_R g(R) h_R}{\sum_R g(R)} \quad \text{mit} \quad g(R) = \begin{cases} \frac{G_Z}{G_R}, & \text{falls } G_Z \leq G_R \\ \frac{G_R}{G_Z}, & \text{falls } G_Z > G_R \end{cases} \quad (5.1)$$

Hierbei bezeichnet h^p den Parameter Höhe (Trauf- oder Firsthöhe), Z das Zielgebäude, R die Referenzgebäude, $g(R)$ den Gewichtungsfaktor und G die Grundfläche des jeweiligen Gebäudes. Die Höhenangaben beziehen sich auf ein Gebäudemodell mit einem Satteldach. Insbesondere bei Gebäuden mit komplizierten Dachformen, vor allem bei Gebäuden mit großen Grundflächen können die Höhen an einzelnen Stellen des Gebäudes deutlich von der angegebenen abweichen.

5.4.2. Berechnung der Gebäudehüllflächen und des Volumens

Die Fläche der thermischen Gebäudehülle A_{th} ist die Summe der im Folgenden näher beschriebenen Einzelflächen (siehe Tabelle 5.6). Summiert werden der untere Gebäu-

Tabelle 5.6.: Parameter der Gebäudehülle

Parameter	Symbol	Untergliederung
Wandflächen	A_W^d	vier horizontale Himmelsrichtungen d
Fensterflächen	A_w^d	vier horizontale Himmelsrichtungen d
Dachflächen	A_R^d	zwei horizontale Himmelsrichtungen d senkrecht zur Firstrichtung d_R
Bedeckte Fassadenflächen	A_{Cx}^d	beheiztes ($x = 1$) oder unbeheiztes ($x = 2$) Nachbargebäude, vier horizontale Richtungen
Grundfläche	G	

deabschluss G , die Wandflächen A_W^d , Fensterflächen A_w^d und Dachflächen A_R^d der vier Himmelsrichtungen d :

$$A_{\text{th}} = \sum_{d=1}^4 (G + A_W^d + A_w^d + A_R^d) \quad (5.2)$$

Da für die Mehrheit der Gebäude keine Angaben zur Unterkellerung der Gebäude existieren, wird bei der Berechnung der thermischen Hüllen davon ausgegangen, dass ein gegebenenfalls existierender Keller in der Regel unbeheizt ist. Der untere Abschluss der thermischen Hüllfläche ist somit die Kellerdecke mit der Größe der Gebäudegrundfläche.

Ein wichtiger Faktor für die Berechnung der Gebäudehüllflächen ist die Form des Gebäudegrundrisses. Diese wird rechnerisch durch den Grundrissumfang je Himmelsrichtung erfasst. Hierbei werden die vier Haupthimmelsrichtungen d unterschieden. Abweichungen von bis zu 45° zur Hauptachse einer Himmelsrichtung werden der entsprechenden Haupthimmelsrichtung zugeordnet. Der Umring des Gebäudegrundrisses wird jeweils zwischen zwei Eckpunkten in einzelne gerade Strecken der Länge u_i geteilt und abhängig von deren Azimut-Winkel Ψ_i die Himmelsrichtung bestimmt. Der Grundrissumfang u^d ergibt sich aus der himmelsrichtungsabhängigen Summe der einzelnen Streckenlängen (Gleichung (5.3)).

$$u^d = \sum_i f(d(u_i), d) \cdot u_i \quad \text{mit} \quad f(d(u_i), d) = \begin{cases} 1, & \text{falls } d(u_i) = d \\ 0, & \text{falls } d(u_i) \neq d \end{cases} \quad (5.3)$$

Firstrichtung Die Richtung des Dachfirsts d_R kann zwei Werte annehmen, Ost-West- sowie Süd-Nord-Richtung. Da keine Informationen über die Dachfirstausrichtung aus Daten ermittelbar ist, wird der Ansatz gebildet, dass der Dachfirst parallel zur längeren Gebäudekante verläuft. Die längere Gebäudekante wird aus dem Vergleich der Summe der Nord- und Südumringabschnitte bzw. Ost- und Westumringabschnitte bestimmt. Vernachlässigt werden hierbei komplexe Dachformen mit verschiedenen Firstrichtungen. Für Gebäude mit einfach strukturiertem Dach kann der Ansatz durch Stichproben mit Luftbildern sehr gut bestätigt werden.

Fassadenfläche je Himmelsrichtung Die Fassadenflächen A_F^d für jede der vier Himmelsrichtungen d werden aus der Länge des Gebäudeumrisses zur jeweiligen Himmelsrichtung, der Traufhöhe h_t und abhängig von der Ausrichtung und Form des Dachs aus den Giebelflächen berechnet. Die Giebelfläche ist die senkrechte Fassadenfläche zwischen Trauf- und Firsthöhe.

$$A_F^d = u^d h_T + 1/2 u^d h_R f(d, d^R) \quad \text{mit} \quad f(d, d^R) = \begin{cases} 0, & \text{falls } d = d^R \vee d^{-1} = d^R \\ 1, & \text{falls } d \neq d^R \vee d^{-1} \neq d^R \end{cases} \quad (5.4)$$

Hierbei bezeichnet d^{-1} die Gegenrichtung, d. h. wenn d Süden ist, ist d^{-1} Norden. Die Traufhöhe bezeichnet den unteren Dachabschluss. Bis zu dieser Höhe, kann die Wandfläche als Rechteck aus der Grundlänge mit entsprechender Himmelsrichtung und der Traufhöhe berechnet werden. Wie bei der Unterkellerung der Gebäude ist für die meisten Gebäude in Oldenburg die Dachform nicht bekannt. Daher wird für alle Gebäude ein Satteldach angenommen [79, S. IV-21]. Die Fallunterscheidung in Gleichung (5.4) führt dazu, dass für die senkrecht zur Richtung des Dachs stehenden Wandseiten die Firstfläche als Dreieck aus Grundlänge und Dachhöhe berechnet und zu der bereits berechneten Rechteckfläche hinzu addiert werden.

Durch angrenzende Gebäude bedeckte Fassadenflächen Eine ähnliche Vorgehensweise ist für die Berechnung von Fassadenflächen A_{Cx}^d implementiert, die durch angrenzende Gebäude abgedeckt sind. Hierfür werden die Strecken u_{Cx}^d der Gebäudeumrisse berechnet, an denen sich die Gebäude berühren. x steht hierbei für den Gebäuderelevanz-Wert der angrenzenden Gebäude. Ein Relevanz-Wert von 1 bezeichnet ein beheiztes Hauptgebäude, der Relevanz-Wert 2 ein unbeheiztes Nebengebäude. Durch die Multiplikation der Strecken U_{Cx}^d mit dem Minimum der Traufhöhen des Ziel- bzw. Referenzgebäudes $h_{t,Z}$ bzw. $h_{t,R}$ ergibt sich die bedeckte Fassadenfläche A_{Cx}^d .

$$A_{Cx}^d = u_{Cx}^d \cdot \min(h_{t,Z}, h_{t,R}) \quad (5.5)$$

Eine entsprechende Dachausrichtung und die möglicherweise Überdeckung von Fassadenflächen durch angrenzende Giebelwände wird nicht berücksichtigt.

Fensterflächen Für Oldenburg liegen keine Informationen über die Fensterflächen der Gebäude vor. Es müssen daher auf Basis der Gebäudetypologie entsprechende Werte geschätzt werden. Die Gebäudetypologie des IWU gibt die Fensterflächen für ein den Gebäudetyp repräsentierendes Typgebäude und für jede Himmelsrichtung an. Die Ausrichtung des Typgebäudes, beispielsweise bei einem Ost-West-ausgerichteten Reihenhaus beeinflusst hierbei die Angaben, da dieses Gebäude keine Fenster auf der Nord- und Südseite hat. Zudem ist die gesamte Fassadenfläche mit Außenluftkontakt angegeben.

Für die Typgebäude wird ein für alle Himmelsrichtungen gemitteltes Verhältnis von Fenster- und Fassadenfläche mit Außenluftkontakt berechnet. Dieses Verhältnis $w_{typ} = \frac{A_{w,typ}}{A_{F,typ}}$ wird als spezifische mittlere Fensterflächen-Kennzahl für die Gebäude in Oldenburg angenommen. $A_{w,typ}$ ist die über alle Himmelsrichtungen summierte Fensterfläche der

Typgebäudes. $A_{F,typ}$ die Fassadenfläche mit Außenluftkontakt. Damit ergibt sich für die Fensterflächen der Gebäude in Oldenburg unabhängig von der Himmelsrichtung d die Gleichung (5.6).

$$A_w^d = A_F^d \cdot w_{typ} = A_F^d \cdot \frac{A_{w,typ}}{A_{F,typ}} \quad (5.6)$$

Der Ansatz skaliert die Fensterfläche anhand der Größe der Fassadenfläche und garantiert außerdem, dass die Fensterfläche immer kleiner als die Fassadenfläche ist.

Wandflächen Die Wandfläche A_W^d ist die Summe der opaquen Bauteile mit Kontakt zur Außenluft und wird für die vier Haupthimmelsrichtungen unterschieden. Sie errechnet sich aus der Fassadenfläche A_F^d minus der bedeckten Fassadenflächen $A_{C_x}^d$ und Fensterflächen A_w^d .

$$A_W^d = A_F^d - A_{C_1}^d - A_{C_2}^d - A_w^d \quad (5.7)$$

Dachfläche je Himmelsrichtung Für die Berechnung der Dachfläche wird die Gebäudeform zu einem Quader vereinfacht, wobei Grundfläche und Umfang gleich den Werten der tatsächlichen Geometrie sind. Die Firstrichtung wird wie oben beschrieben parallel zur längeren Gebäudeseite der tatsächlichen Geometrie bestimmt. Unter der Annahme eines Satteldachs werden die beiden Dachflächen A_R nach Gleichung (5.8) berechnet.

$$A_R = \sqrt{h_R^2 + \left(\frac{1}{2}\min(a, b)\right)^2} \cdot \max(a, b) \quad (5.8)$$

Hierbei bezeichnen a und b die Seitenlängen des Quaders. Die Dachfläche wird so berechnet, dass der Dachfirst parallel zur längeren Seite des Quaders verläuft.

Bruttovolumen Bei der Berechnung wird ebenfalls ein Satteldach angenommen. Damit ergibt sich das durch die äußere Gebäudehülle umschlossene Bruttovolumen V_e durch die Gleichung (5.9).

$$V_e = G \cdot \left(h_f + \frac{1}{2}h_r\right) \quad (5.9)$$

5.4.3. Analyse und Literaturvergleich

Die Abbildung 5.11 gibt ein Beispiel für die berechneten Gebäudehüllflächen einer Doppelhaushälfte. Das Gebäude grenzt im Westen an die andere Doppelhaushälfte und im Osten an eine Garage an. Die um etwas 30° verdrehte Ausrichtung gehört zur Ost-West-Achse der Haupthimmelsrichtungen. Die Traufhöhe ist mit 2,70 m, die Dachhöhe mit 1,24 m angegeben. Die Fassadenflächen ergeben sich aus der Geometrie des Gebäudegrundrisses und den Höhen. Die Westfassade ist nahezu vollständig abgedeckt durch das angrenzende Hauptgebäude. Für das benachbarte Hauptgebäude geben die Laserscan-Daten eine kleinere Höhe an, weshalb die abgedeckte Fassadenfläche nicht mit der Gesamt-Fassadenfläche übereinstimmt. Auf Basis des Fensterflächenanteils der Gebäudetypologie

5. Berechnung und Analyse von Gebäudeparametern in Oldenburg

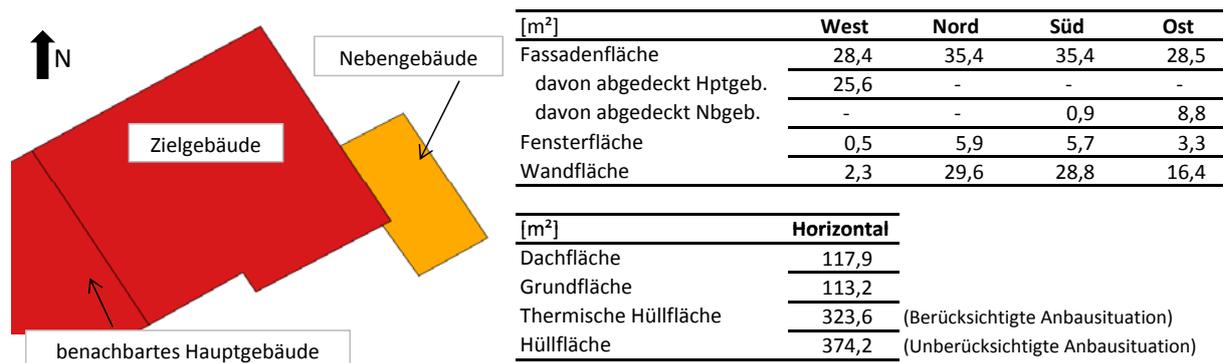


Abbildung 5.11.: Beispiel der thermischen Hüllflächen für eine Doppelhaushälfte mit angrenzender Garage

für den Gebäudetyp (EFH/ZFH der Baualtersklasse F) werden die Fensterflächen berechnet. Die Wandfläche berechnet sich dann als Differenz der Fassadenfläche minus der anderen Größen. Die Dachfläche ist aufgrund der geringen Dachhöhe nur geringfügig unterschiedlich zur Grundfläche. Die thermische Hüllfläche beinhaltet alle nicht abgedeckten Oberflächen und die Grundfläche.

Dieses Beispiel zeigt einige Verallgemeinerungen. Es ist unwahrscheinlich, dass in der Realität ein Fenster zu dem benachbarten Hauptgebäude nach Westen hinaus geht. Allerdings ist diese Einzelfallbetrachtung für den gesamten Gebäudebestand nicht durchführbar bzw. führt zu sehr vielen Fallunterscheidungen und Grenzfällen. Darüber hinaus wirkt die Dachhöhe mit 1,24 m sehr klein. Eine manuelle Überprüfung ergibt, dass das Gebäude im Bungalow-Stil errichtet ist.

Zur Überprüfung der berechneten Hüllflächen und Gebäudevolumen werden die mittleren A/V-Verhältnisse der Gebäude in Oldenburg mit den Angaben der Typgebäude des IWU in Abbildung 5.12 verglichen [53]. Jeder Datenpunkt steht für einen Gebäudetyp, wobei die Gebäudeklassen farblich unterschieden sind. Das A/V-Verhältnis ist der Quotient zwischen der thermischen Hüllfläche mit Kontakt zur Außenluft und dem Bruttovolumen des Gebäudes. Insgesamt ist erkennbar, dass die mittleren Angaben der Datenbank je Gebäudetyp nicht so stark voneinander abweichen, wie die Angaben des IWU, da gleichfarbige Punkte horizontal sehr ähnlich liegen. Die Angaben der Datenbank sind gemittelt aus jeweils über 1.000 Gebäuden, die Angaben des IWU entstammen einem Gebäude, das den jeweiligen Gebäudetyp repräsentiert. Die meisten Gebäudetypen weisen eine Abweichung kleiner als 25 % auf. Für einige Gebäudetypen der Klassen RH und MFHkl gibt das IWU deutlich kleinere A/V-Verhältnisse als die Berechnungen in Oldenburg an. Bei MFHkl ist der Grund, dass diese in Oldenburg freistehend, d. h. ohne angrenzende Gebäude gebaut sind. Die IWU-Typologie geht jedoch von einer Reihenbauweise aus. Die thermische Hülle beim IWU ist nicht immer die äußere Gebäudehülle. Zum Teil ist bei unbeheizten Dachgeschossen die thermische Hülle die obere Geschossdecke. Diese Informationen sind jedoch aus den verfügbaren Gebäudedaten in Oldenburg nicht extrahierbar.

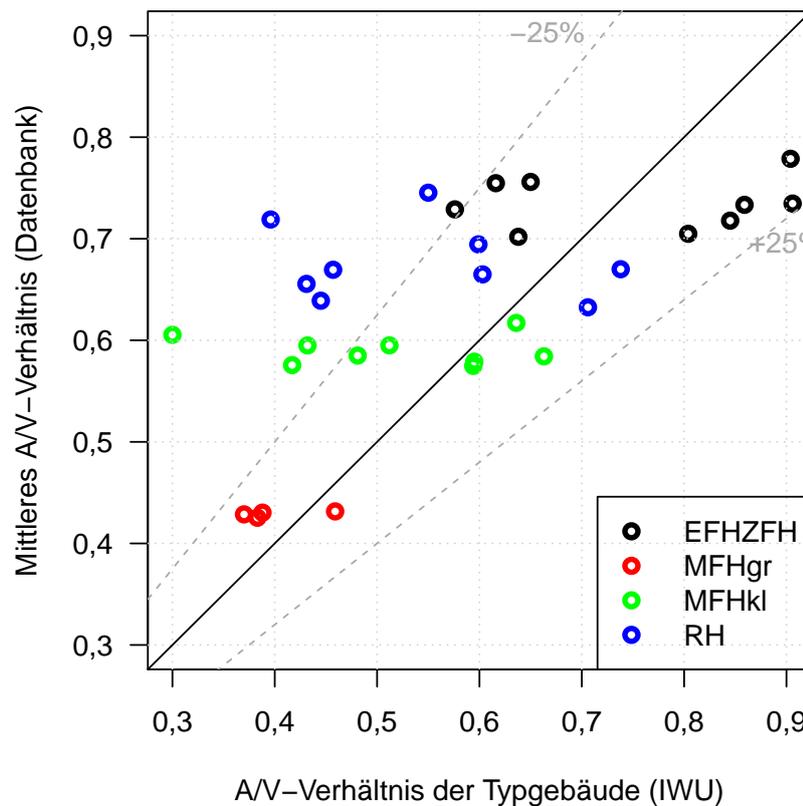


Abbildung 5.12.: A/V-Verhältnisse der Gebäudetypen und Literaturvergleich [53]

5.5. Geschosshöhe und Grundflächen

Die Geschosshöhe ist eine Voraussetzung für die Berechnung der Grundflächen eines Gebäudes. Die Grundflächen sind als Bezugsgrößen wichtig für die Wärmebedarfsberechnung mit Hochrechnungsverfahren.

5.5.1. Quellen bzw. Berechnung der Geschosshöhe aus der Gebäudehöhe

Für einen Teil der Gebäude können die Geschosshöhen aus den Daten der Kaufpreissammlung bezogen werden. Für alle anderen Gebäude werden die Geschosshöhen aus den Gebäudehöhen errechnet. Dabei werden Voll- und Dachgeschosse unterschieden. Die Zahl der Vollgeschosse wird auf Basis der Traufhöhen berechnet, soweit die Traufhöhenangaben plausibel erscheinen, d. h. größer gleich 2,8 m sind. Anhand der Dachhöhe wird das Vorhandensein eines Dachgeschosses bestimmt. Bei der Vollgeschosshöheberechnung ist die Baualtersklasse der Gebäude ein Kriterium für die Annahme der Geschosshöhen. Für Gebäude mit Baujahr vor 1918 (Typ B) wird eine Geschosshöhe von 3,7 m, für Gebäude bis 1945 (Typ C) von 3 m und für alle später erbauten Gebäude eine Höhe von 2,75 m angenommen (siehe Abbildung 5.13). Die Mindestdachhöhe für das Vorhandensein eines

Dachgeschosses ist 3,2 m. Dachhöhen größer als 10 m werden als fehlerhafte Angabe eingestuft. Die Höhenangaben werden in dem Fall für die Geschoszberechnung nicht verwendet. Der Entscheidungsbaum ist Ergebnis eines Forschungsprojekts mit der Jade Hochschule in Oldenburg, das im Rahmen dieser Arbeit betreut wurde [44].

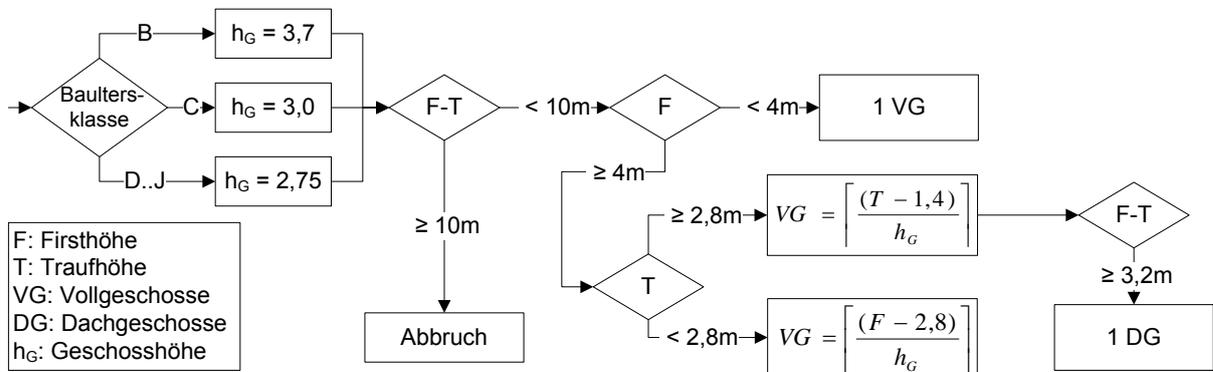


Abbildung 5.13.: Entscheidungsbaum für die Bestimmung der Geschoszberechnung anhand der Gebäudehöhe, Quelle: [44], eigene Darstellung

5.5.2. Berechnung und Umrechnung von Grundflächen

Eines der wichtigsten Kriterien bei der Angabe spezifischer Wärmeverbräuche oder -bedarfe ist die Bezugsgröße. In den meisten Fällen ist diese die Wohn- bzw. Nutzfläche oder die Energienutzfläche. Für die Anwendung von Hochrechnungsverfahren bei der Berechnung des Wärmebedarfs bzw. -verbrauchs ist die Kenntnis der Grundflächen Voraussetzung.

Abbildung 5.14 gibt eine Übersicht der unterschiedlichen Grundflächen. In vier Spalten werden die in der jeweiligen DIN vorkommenden und relevanten Flächen angegeben. Gleiche Grundflächen unterschiedlicher Normen liegen horizontal nebeneinander.

In der DIN 277 sind Grundflächen definiert und stehen in einem hierarchischen Verhältnis. Die **Bruttogrundfläche (BGF)** ist die Summe der Grundflächen aller Grundrissebenen eines Bauwerkes mit den Kategorien Nutzflächen, Funktionsflächen und Verkehrsflächen plus der Konstruktions-Grundfläche [21, Teil 1]. In der Arbeit wird die BGF als Summe der Grundfläche aller Geschosse berechnet. Die Begriffe **Bruttogeschossfläche** und **Nettogeschossfläche** sind nicht offiziell definiert, finden jedoch oftmals Verwendung für die Bruttogrundfläche bzw. Nettogrundfläche. Die **Nettogrundfläche (NGF)** gliedert sich in Nutzfläche, Technische Funktionsfläche und Verkehrsfläche [21, Teil 1]. Sie ist die Energiebezugsfläche nach DIN V 18599 [23].

Die **Nutzfläche (NF)** ist die „Summe der Grundflächen mit Nutzungen nach DIN 277-2:2005-02, Tabelle 1, Nr 1 bis Nr 7“ [21, Teil 1]. In reinen Wohngebäuden entspricht die Nutzfläche der Wohnfläche. Allgemein ist die NGF nach DIN 283 unterteilt in die Wohnfläche und die nutzbare Grundfläche von Wirtschaftsräumen und gewerblichen

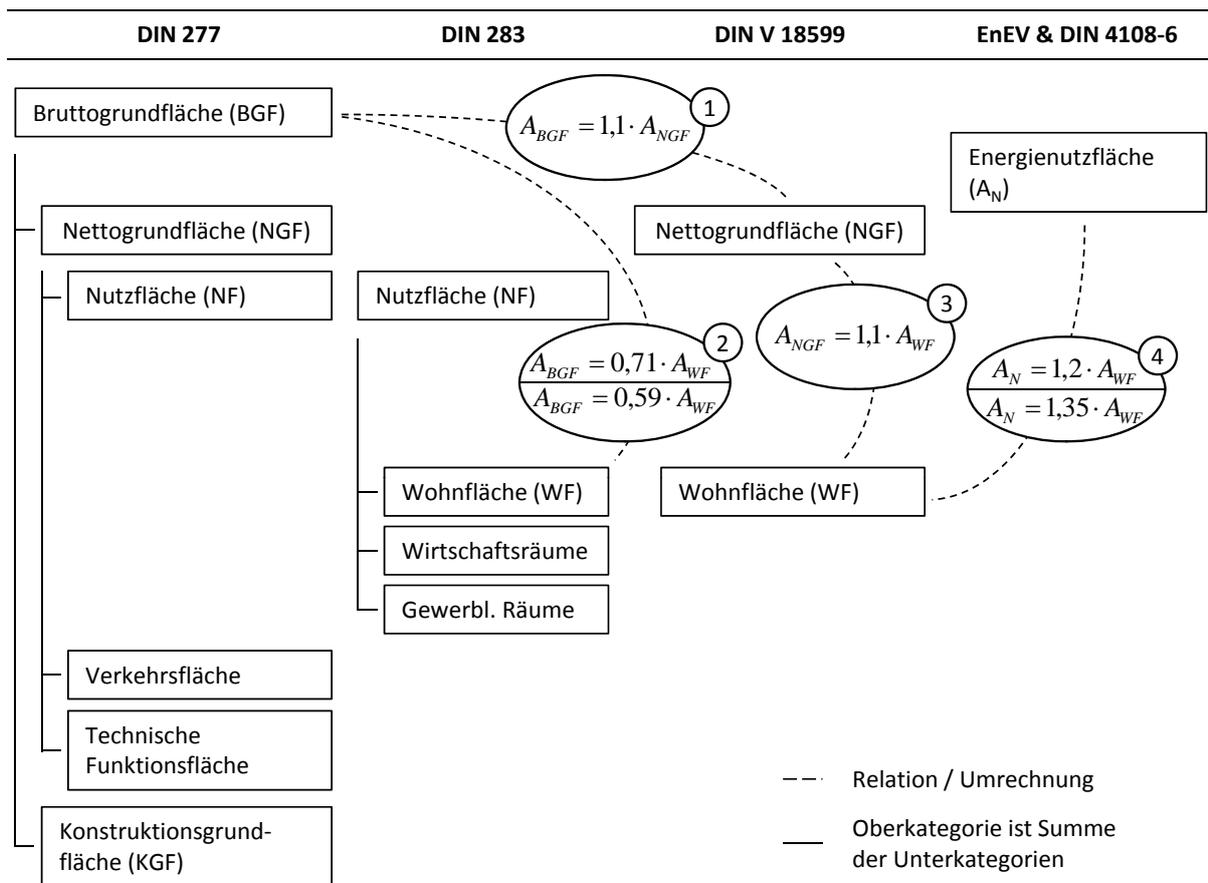


Abbildung 5.14.: Grundflächen-Begriffe in Normen und Verordnungen

Räumen. Die DIN 283 ist zwar nicht mehr gültig, wird jedoch weiter angewendet und hat sich zu einer anerkannten Regel der Technik entwickelt.

Die **Wohnfläche (WF)** ist die Summe der Grundflächen, die ausschließlich einer Wohnung oder einem Wohnheim zugeordnet werden [18].

Die Nutzfläche ist zu unterscheiden von der **beheizten Energienutzfläche (A_N)**. Letztere ist die Energiebezugsfläche der EnEV bzw. DIN 4108-6. Sie wird direkt aus dem Brutto-Gebäudevolumen V_e abgeleitet. Die Berechnungen unterscheiden sich je nach Gültigkeit der EnEV. Laut EnEV 2009 [17] gilt $A_N = 0,32V_e$. Eine Ausnahme bilden Bestandsgebäude mit einer Geschosshöhe $h_G < 2,5\text{ m}$ bzw. $h_G > 3\text{ m}$. Bei diesen gilt $A_N = 1/h_G - 0,04$.

In der Literatur existieren darüber hinaus verschiedene und sich zum Teil widersprechende pauschale Umrechnungen zwischen zwei Grundflächen. Die Aufzählungsnummern korrespondieren zu den Angaben in Abbildung 5.14:

1. Nach DIN 283 kann die Nettogrundfläche aus der Bruttogrundfläche durch den Faktor 1,1 abgeschätzt werden.
2. Wohnfläche von Ein-/Zweifamilienhäusern wird durch der VDI-Norm 3807 Blatt1 als 71 % der Bruttogrundfläche, die Wohnfläche im Geschosswohnbau als 59 % der

Bruttogrundfläche angegeben [80]. Hierzu gibt es jedoch unterschiedliche Angaben. Eine Machbarkeitsstudie für die Stadt Oldenburg gibt als Umrechnungsfaktoren 0,91 für Ein- und Zweifamilienhäuser bzw. 0,81 für Mehrfamilienhäuser an [12, S. 109].

3. Die DIN V 18599 erlaubt die Berechnung der Nettogrundfläche als das 1,1-fache der beheizten Wohnfläche.
4. Die Umrechnung zwischen der Gebäudenutzfläche und der Wohnfläche in in der EnEV mit dem Faktor 1,2 bzw. bei Ein- und Zweifamilienhäusern mit beheiztem Keller mit dem Faktor 1,35 angegeben. Das IWU geht jedoch von einem Umrechnungsfaktor 1,1 aus [53].

Im Untersuchungsgebiet werden die BGF F_B berechnet und daraus alle weiteren Flächen durch die in Abbildung 5.14 angegebenen Faktoren abgeschätzt. Die BGF setzt sich aus der Regelgeschossfläche F_R und der Dachgeschossfläche F_D nach Gleichung (5.10) zusammen. Die Regelgeschossfläche F_R ist die Zahl der Geschosse g , multipliziert mit der Gebäudegrundfläche G . Die Dachgeschossfläche wird aus der Regelgeschossfläche durch einen Faktor $f_{DG} = 0,8$ verkleinert, der die Verlustflächen aufgrund der Dachschräge pauschal berücksichtigt.

$$F_B = F_R + F_D = G \cdot g + G \cdot f_{DG} \quad (5.10)$$

5.5.3. Auswertung des Verhältnisses zwischen Wohnfläche und Volumen

In Abbildung 5.15 ist das Verhältnis zwischen Wohnfläche und Volumen für jeden Gebäudetyp als ein Vergleich zwischen den gemittelten Daten der Datenbank und Literaturwerten des IWU angegeben [53]. Für Datenpunkte auf der Winkelhalbierenden sind beide Angaben gleich. Dieser Indikator wird gewählt, um die berechneten Wohnflächen, die von der Berechnung der Stockwerke und den oben zitierten Korrekturfaktoren abhängen, mit dem Volumen als einem sehr viel leichter zu berechnenden Gebäudeparameter in ein Verhältnis zu setzen. Jeder Datenpunkt steht für einen Gebäudetyp. Die Daten des IWU zeigen eine höhere Streuung, da jede Angabe nur auf den Werten eines Typgebäudes basiert. Der Großteil der Datenpunkte spiegelt eine gute Übereinstimmung wieder. Die Datenpunkte für EFH/ZFH zeigen, dass in der Datenbank größere Flächen pro Volumen als bei den Typgebäuden berechnet werden. Aufgrund der pauschalisierten Rechnung in der Datenbank wird das Dachgeschoss bei der Berechnung der Wohnfläche mit einbezogen. Das ist jedoch nicht bei den Typgebäuden der Fall, die durch Befragung oder Begehung aufgenommen und genauer vermessen sind.

Ein Vergleich der berechneten Wohnflächen mit statistischen Daten der Stadt Oldenburg wird in Kapitel 5.6.2 durchgeführt.

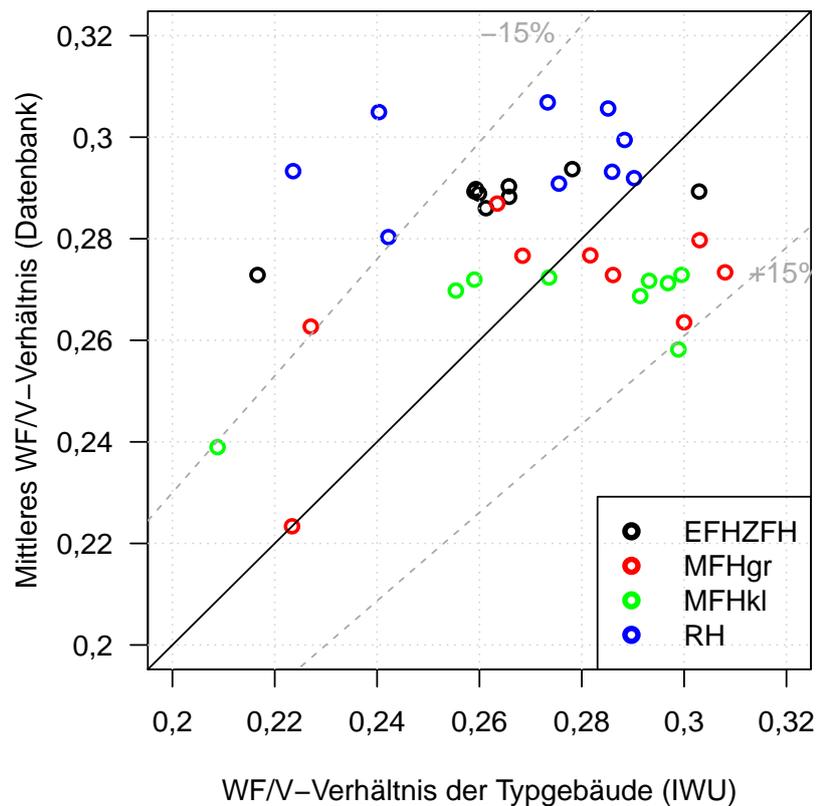


Abbildung 5.15.: Wohnfläche-Volumen-Verhältnisse der Gebäudetypen und Literaturvergleich

5.6. Gebäudenutzung

Die Gebäudenutzung ist eine kategorielle Größe. Angaben beziehen sich immer auf eine Klassifizierung. Insbesondere beim Vergleich von Angaben unterschiedlicher Quellen ist es notwendig, dass beide Quellen die selbe Klassifizierung aufweisen oder die Angaben vor dem eigentlichen Vergleich in eine gemeinsame Klassifizierung transformiert werden.

Es existieren eine Vielzahl von Klassifizierungen gewerblicher Nutzungen. Beispiele hierfür sind die hierarchische Wirtschaftszweigklassifizierung des Statistischen Bundesamtes [16], die Klassifizierung von Betrieben im Sektor GHD [68] oder Kundengruppenschlüssel zur Zuordnung von Lastprofilen [81]. Eine Klassifizierung teilt einzelne Objekte bezüglich ihrer wichtigen Charakteristika in Kategorien ein. Welche Charakteristika als wichtig erachtet werden, ist hierbei dem Ziel der Klassifizierung unterworfen.

Eine automatisierte Verarbeitung von Nutzungsangaben setzt daher Abbildungsregeln der verschiedenen originalen Klassifizierungen in eine einheitliche Zielklassifizierung voraus. Falls diese Abbildungsregeln eindeutig sind, d. h. als Funktionen ausgedrückt werden können, ist eine automatische Transformation unproblematisch. Falls nicht, müssen Annahmen für die automatisierte Abbildung getroffen werden oder manuelle Korrekturen erfolgen.

5.6.1. Transformation der Gebäudenutzungsdaten in Oldenburg

Die Gebäudenutzung wird in dieser Arbeit aus der Nutzungsklassifizierung der Stromverträge des Stromnetzbetreibers und der Gebäude-Objektart der ALK bestimmt (siehe Kapitel 4.1.9 und 4.1.4). Die Klassifizierung der Stromverträge ist detaillierter als die der ALK, weshalb die Stromverträge in erster Linie herangezogen werden. Falls ein Gebäude über keine Angaben zu den Stromverträgen verfügt, wird die Klassifizierung der ALK ergänzend verwendet. Es wird angenommen, dass jede Partei (Wohn- oder Gewerbeeinheit) eines Gebäudes einen Stromvertrag hat. Aus den Stromverträgen kann damit auf die Zahl der Parteien pro Gebäude und die Nutzung des Gebäudes geschlossen werden. Darüber hinaus werden gemischt genutzte Gebäude identifiziert und in verschiedene Nutzungszonen aufgeteilt.

Als Zielklassifizierung wird die Gewerbeklassifizierung der Studie Energieverbrauch des Sektors Gewerbe, Handel, Dienstleistung (GHD) verwendet [68]. Diese wird um die zwei Klassen „Haushalte EFH/ZFH“ und „Haushalte MFH“ erweitert, um den Sektor Haushalte ebenfalls abzubilden. Die Gewerbeklassifizierung ist das Ergebnis mehrerer mehrjähriger Analysen der Energieverbräuche im Sektor GHD. Dieser und der Sektor Haushalte stehen im Fokus von kommunalen Wärmeanalysen. Darüber hinaus haben die Sektoren Handel und Dienstleistung einen Anteil von 87,2 % an der Bruttowertschöpfung in Oldenburg [72, Tab. 0300]. Das produzierende Gewerbe hat einen geringen Anteil von 12,8 %. Die gebildete **GHD-HH-Klassifizierung** unterscheidet den GHD-Sektor vor allem anhand seiner Energieverbrauchsstruktur und umfasst insgesamt 30 Nutzungsklassen, die in 13 Branchen zusammengefasst sind. Branchen und Klassen sind in Tabelle 5.7 aufgelistet.

Die Transformation der Nutzungsklassifizierung umfasst die folgenden Aspekte:

- Die Nutzungsklassifizierung der Stromverträge ist mit 105 Nutzungsklassen relativ fein. Der Sinn dieser Klassifizierung ist die Zuordnung von Lastprofilen. 50 % der Klassen werden dem Sektor Industrie zugeordnet. Die Tabelle A.1 im Anhang zeigt die Zuordnungsregeln, die im Rahmen dieser Arbeit entwickelt wurden.
- Die Zuordnung von Betrieben des produzierenden Gewerbes zum Sektor GHD bzw. Industrie unterscheidet sich aufgrund der Zahl der Erwerbstätigen pro Betrieb [68]. Bei weniger als 20 Erwerbstätigen geht man vom Sektor GHD aus. Da Erwerbstätigenzahlen im Rahmen dieser Arbeit nicht verfügbar sind, wird eine pauschalisierte Zuordnung auf Basis der Klassenbezeichnungen durchgeführt.
- Die Objektartenklassifizierung der ALK beinhaltet nur 8 Klassen im Bereich der nicht-öffentlichen Nutzungsarten, die für eine Umwandlung in die GHD-HH-Klassifizierung interessant sind. Bei der Transformation dieser groben Klassifizierung auf die feinere Zielklassifizierung werden Annahmen über die prozentuale Aufteilung getroffen. Tabelle 5.8 zeigt die Klassen und deren Aufteilung. Die Kategorien werden hierbei nur auf die Zielklassen „Haushalte“ und „Sonstige betriebliche Dienstleistungen“ aufgeteilt.
- Die Branche Haushalte wird in EFH/ZFH und MFH anhand des Gebäudetyps des betroffenen Gebäudes aufgeteilt.

Tabelle 5.7.: Branchen, Nutzungsklassen und Anzahl der Gebäudeparteien in Oldenburg

Branche	Nutzungskategorie	Anz. Parteien	
Haushalte	Haushalte EFH/ZFH	39.664	
	Haushalte MFH	47.616	
Gartenbau	Gartenbau u. Gärtnereien	24	
Bekleidung, Leder, Spedition	Bekleidung, Leder, Textil	26	
	Spedit., Lagerei, Verkehrsvermittlung	75	
Baugewerbe	Baugewerbe	281	
	Kreditinstitute u. Versicherungen	255	
	Verlagsgewerbe	18	
	Sonst. betr. Dienstleistungen	2.143	
	Büroähnliche Betriebe	Gebietskörpersch. u. Sozialversich.	164
		Deutsche Bundespost / Postdienste	211
		Telekommunikation	20
		Deutsche Bahn AG	5
Herstellungsbetriebe	Metallgewerbe	316	
	Holzgewerbe	38	
	Papier- u. Druckgewerbe	44	
Handel	Einzelhandel	1.508	
	Großhandel	131	
	Handelsvermittlungen	4	
Krankenhäuser, Schulen, Bäder	Krankenhäuser	20	
	Schulen	207	
Beherbergung, Gaststätten, Heime	Beherbergungsgewerbe	50	
	Gaststättengewerbe	445	
	Org. ohne Erwerbszweck und Heime	468	
Nahrungsmittelgewerbe	Backgewerbe	87	
	Fleischereien	25	
	Restl. Nahrungsmittelgewerbe	9	
Wäschereien	Wäschereien u. (chem.) Reinigungen	33	
Landwirtschaft	Landwirtschaft	114	
Industrie	Industrie	618	

Tabelle 5.8.: Häufigste ALK-Objektarten mit den Gewichtungen für die Transformation der Nutzungsklassen

Schlüssel	Objektart-Beschreibung	Gewichtung Haushalte	Gew. Sonst. betr. Dienstl.
931	Nichtöffentliches Gebäude mit Schraffur unter einem Winkel von 50 gon zur längsten Gebäudeseite	100 %	0 %
932	Nichtöffentliches Gebäude mit Schraffur senkrecht zur längsten Gebäudeseite	0 %	100 %
2111	Wohngebäude mit Öffentlich	50 %	50 %
2121	Wohngebäude mit Handel und Dienstleistungen	50 %	50 %
2131	Wohngebäude mit Gewerbe und Industrie	50 %	50 %
2141	Öffentliches Gebäude mit Wohnen	10 %	90 %
2151	Gebäude für Handel und Dienstleistungen mit Wohnen	10 %	90 %
2161	Gebäude für Gewerbe und Industrie mit Wohnen	10 %	90 %

5.6.2. Berechnung der Gebäudezonengrößen für verschiedene Nutzungen

Die Zahl der Stromverträge einer bestimmten Nutzung geteilt durch die Summe der Stromverträge je Gebäude ergibt eine Gewichtung dieser Nutzung in dem Gebäude. Falls für ein Gebäude keine Stromverträge vorliegen, wird auf Basis der ALK-Gebäudeart eine pauschalisierte Gewichtung angenommen. Diese ist in Tabelle 5.8 für alle relevanten Objektarten prozentual angegeben.

Bei der Transformation von Nutzungsklassen in eine neue Klassifizierung verändert sich die Klassengewichtung. Falls beispielsweise zwei Klassen der Ursprungsklassifikation auf eine Klasse der Zielklassifikation abgebildet werden, werden die Einzelgewichtungen summiert. Gleichung (5.11) gibt die Berechnung verallgemeinert an, wobei $g_{u',u}$ der Transformationsfaktor von der Nutzungsklasse $u \in U$ der Ursprungsklassifizierung U zur Nutzungsklasse $u' \in U'$ der Zielklassifizierung U' ist. $n_{u',i}$ ist die Gewichtung der Nutzungsklasse u' im Gebäude i .

$$n_{u',i} = \sum_{u,i} n_u g_{u',u} \quad (5.11)$$

Der Transformationsfaktor g liegt in einem Wertebereich zwischen Null und Eins und es muss gelten, dass die Summe von g über alle Zielklassen $u' \sum_{u'} g_{u',u} = 1$ ist. Damit gibt der Transformationsfaktor an, welcher Teil der Gewichtung von der Klasse u auf die Klasse u' übergeht. Beispielsweise für den Objektartenschlüssel 931 in Tabelle 5.8 ist $g_{\text{Wohngebäude},931} = 1$, $g_{\text{Sonstiges Gewerbe},931} = 0$. Für den Objektartenschlüssel 2111 sind beide g -Werte jeweils 0,5, da die Fläche auf beide Zielklassen zu gleichen Teilen aufgeteilt wird.

Verschiedene Nutzungen nehmen in der Regel unterschiedlich große Zonen in einem Gebäude ein. In einem Wohngebäude mit Geschäften ist davon auszugehen, dass ein Geschäft in der Regel eine größere Fläche als eine Wohnung einnimmt. Bei einer Aufteilung der Grundfläche in unterschiedliche Gebäudezonen werden nutzungsspezifische Gewichtungsfaktoren f_u berücksichtigt. Nutzungsklassen des Sektors GHD werden in der Arbeit doppelt gewichtet, Klassen des Sektors Haushalte einfach. Der Annahme liegt der Gedanke zugrunde, dass in einem Gebäude durchschnittlich zwei Wohnungen pro Stockwerk liegen. Eine gewerbliche Einheit, z.B. ein Laden im Erdgeschoss nimmt jedoch die gesamte Stockwerksfläche ein und hat damit das Gewicht von zwei Wohnungen. Der gewichtete Anteil $n_{u,i}^g$ einer Zone mit Nutzung u an der gesamten Fläche eines Gebäudes i wird mit Gleichung(5.12) berechnet.

$$n_{u,i}^g = \frac{n_{u,i} f_u}{\sum_u (n_{u,i} f_u)} \quad (5.12)$$

5.6.3. Analyse der Nutzungsarten und zugeordneten Nutzflächen im Stadtgebiet Oldenburg

Neben der Darstellung der Nutzungsklassen und -branchen gibt Tabelle 5.7 die Anzahl der Parteien mit entsprechender Nutzung an. Es gibt ca. 39.000 Parteien bzw. Wohneinheiten in EFH/ZFH und RH sowie ca. 47.500 in MFH. Dies passt gut zusammen mit den laut Statistischem Jahrbuch ca. 84.500 Wohneinheiten in Oldenburg in 2008 [72, Tab. 0402]. Die am stärksten vertretenen Kategorien im Sektor GHD sind die Bereiche Handel und büroähnliche Betriebe.

Wohnflächen in Oldenburg Die Wohnflächenstatistik berücksichtigt nicht die Dachgeschosse von EFH/ZFH, RH und MFHkl, da diese in den meisten Fällen nicht als Wohnfläche gelten [44]. Für den Vergleich mit der Statistik werden die Wohnflächen nur auf Basis der

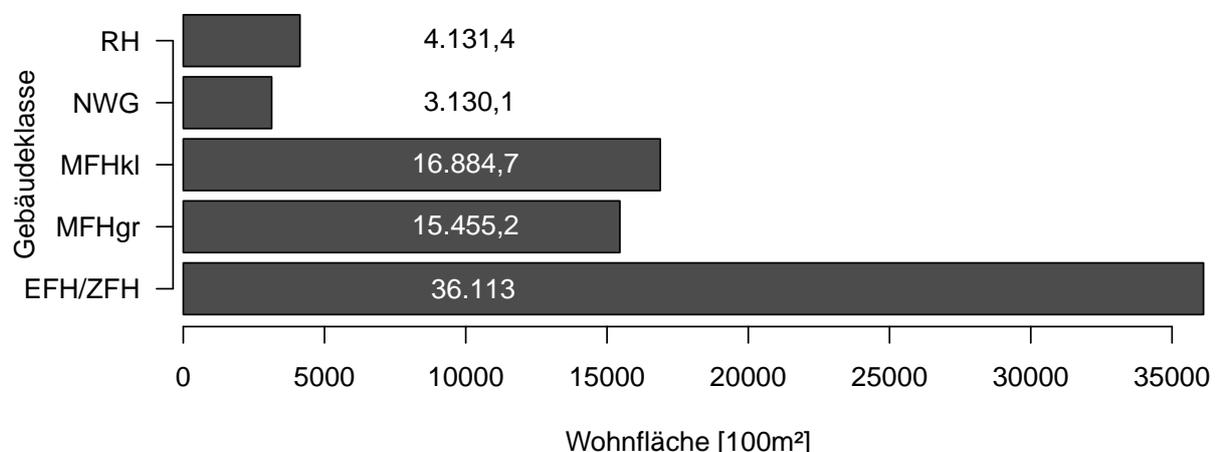


Abbildung 5.16.: Wohnfläche je Gebäudeklasse in Oldenburg

5. Berechnung und Analyse von Gebäudeparametern in Oldenburg

Hauptgeschosse berechnet. Dies bildet unter Umständen nicht die korrekte Wohnfläche ab, ermöglicht jedoch einen Vergleich der Angaben beider Quellen. Damit ergibt sich als Wohnfläche in Wohngebäuden 72.584 100m², in Nichtwohngebäuden 3.130 100m². Im Vergleich dazu gibt das Statistische Jahrbuch für Ende 2008 70.156 100m² Wohnfläche in Wohngebäuden und 1.122 100m² in Nichtwohngebäuden an [72, Tabelle 0403]. Eine Übersicht der Wohnflächen je Gebäudeklasse gibt Abbildung 5.16.

Die Wohnfläche in Nichtwohngebäuden wird in der Arbeit zu groß berechnet. Dies liegt daran, dass in Nichtwohngebäuden die Wohnanteile auf Basis der Stromverträge einen zu hohen Anteil an der gesamten Nutzfläche erhalten. Die dadurch erhaltene Verzerrung hat jedoch maximal einen Anteil von 2-3% an der gesamten Wohnfläche und kann vernachlässigt werden.

Gewerbeflächen in Oldenburg Gewerbeflächen nehmen in Oldenburg ca. 41.000 100m² Nutzfläche ein. Büroähnliche Betriebe und Handel sind mit einem Anteil von 60% repräsentiert (siehe Abbildung 5.17). In die Kategorie büroähnliche Betriebe fallen auch Gebietskörperschaften und Sozialversicherungen. Aufgrund ihres hohen Prozesswärmeverbrauchs werden Wäschereien als eigene Kategorie geführt. Landwirtschaft und Gartenbau sind in Oldenburg wenig vorhanden. Die dargestellten Gewerbeflächen verteilen sich auf insgesamt ca. 1.400 MFH, 370 EFH/ZFH und RH und 2.323 NWG.

Der Raumbezug erlaubt eine Darstellung ausgewählter Daten als Karte. Beispielhaft werden in Abbildung 5.18 die Nutzfläche der Branche Handel dargestellt. Die Daten sind auf quadratische Raster mit einer Seitelänge von 250 m aggregiert. In Kartenabschnitten ohne Raster gibt es keine Handelsflächen. Dadurch wird die spinnenartige Struktur der Branche im Stadtgebiet Oldenburg sichtbar. Die Handelsflächen befinden sich in der

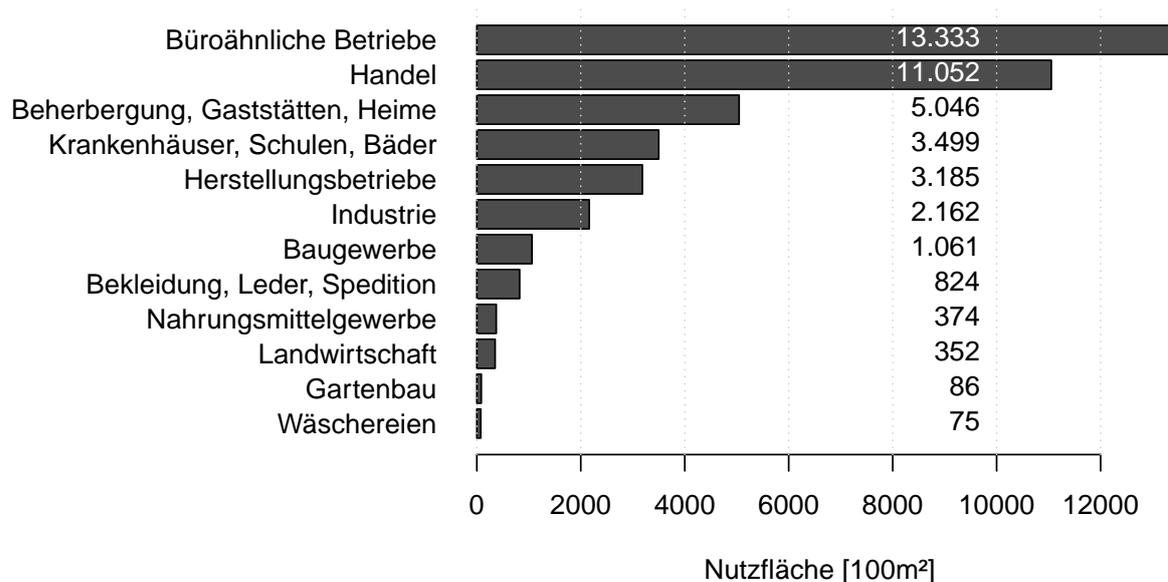


Abbildung 5.17.: Nutzfläche je Branche in Oldenburg

Innenstadt (grün-roter Bereich im Zentrum der Karte) und entlang der Ausfallstraßen nach Norden, Süden und Westen. Die zwei dunkelroten Raster im Nordwesten bzw. Osten geben außerhalb der Stadt liegende Handels-Gewerbegebiete an.

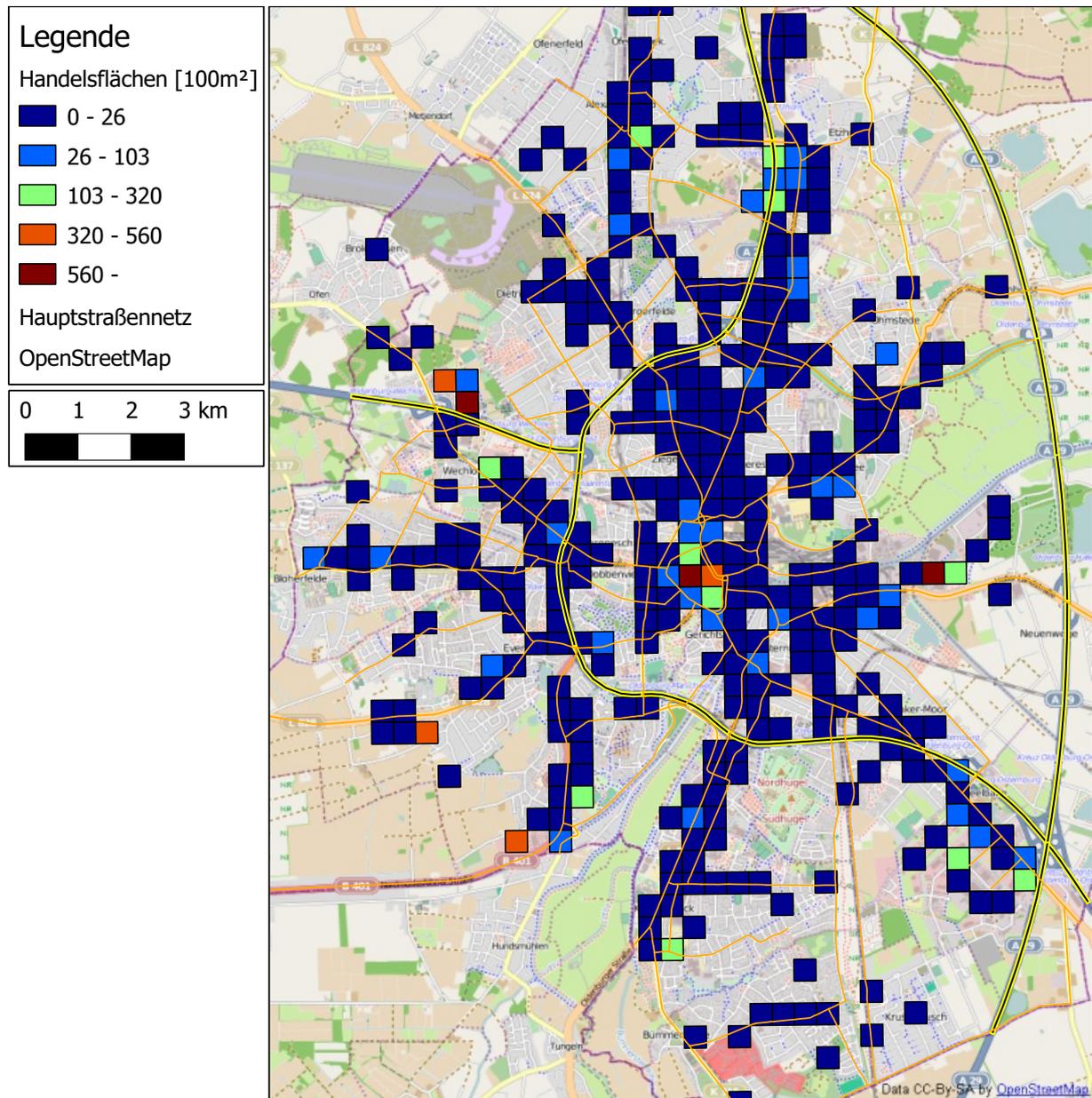


Abbildung 5.18.: Karte der berechneten Nutzflächen der Branche Handel in Oldenburg, Hintergrundgrafik: OpenStreetMap

6. Analyse der Wärmeverbräuche von Ein- und Zweifamilienhäusern in Oldenburg

Basierend auf den im vorherigen Kapitel abgeleiteten Gebäudeparametern werden in diesem Kapitel gemessene und klimabereinigte Wärmeverbräuche für ca. 17.500 Ein- und Zweifamilienhäuser analysiert. Der erste Schritt ist die Aufbereitung der gemessenen Verbrauchswerte (Kapitel 6.1). Es werden mittlere spezifische Wärmeverbräuche für verschiedene Kombinationen von Gebäudeparametern berechnet. Sie ergeben Kennwerte für den Vergleich der Daten mit Literaturangaben und zeigen, inwiefern Gebäudeparameter und Verbrauchswerte zusammenhängen (Kapitel 6.2). Kapitel 6.3 beschreibt mithilfe einer multiplen linearen Regression, inwiefern die Varianz der Verbrauchswerte durch die bekannten Gebäudeparameter erklärt werden kann. Bei einer Schätzung von Wärmeverbräuchen führt die nicht erklärbare Varianz zu Schätzfehlern. Für Szenarien mit einer variierenden Verfügbarkeit von Gebäudeparametern und unterschiedlich große Gruppen von Gebäuden werden die Schätzfehler quantifiziert (Kapitel 6.4). Im letzten Kapitel werden die gemessenen Wärmeverbräuche mit berechneten Wärmebedarfen verglichen (Kapitel 6.5).

6.1. Gebäudeauswahl und Berechnung des klimabereinigten Heizwärmeverbrauchs

6.1.1. Gebäudeauswahl

Die untersuchten 17.500 Ein- und Zweifamilienhäuser sind eine Auswahl aus dem gesamten Oldenburger Hauptgebäudebestand von ca. 40.000 Gebäuden. Die Auswahl erfolgt nach den folgenden Argumenten:

- In die Auswahl gehen nur Gebäude ein, die ausschließlich privat genutzt werden. Ziel ist eine Analyse im Sektor Haushalte.
- Die Stichprobe wird außerdem auf die Gebäudeklasse Ein- und Zweifamilienhäuser (EFH/ZFH) beschränkt. Die Schwankung der Verbräuche durch unterschiedliches Nutzerverhalten ist bei Gebäuden dieser Klasse sehr viel höher als bei Mehrfamilienhäusern. Gebäude mit mehr als zwei Wohneinheiten zeigen zum Teil deutliche Ausgleichseffekte des Nutzerverhaltens innerhalb eines Gebäudes und weisen somit eine geringere Streuung der Endenergieverbräuche auf.

- Die Gasliefermengen liegen aggregiert für ein Anschlussobjekt vor. Dies entspricht einem Hausanschluss. Zum Teil kommt es vor, dass mehrere Gebäude sich einen Anschluss teilen bzw. einem Gebäude mehrere Anschlussobjekte zugeordnet sind. Es werden jedoch in diesem Kapitel nur Gebäude berücksichtigt, denen genau ein Anschlussobjekt zugeordnet werden kann, um fehlerhafte Zuordnungen von Gasliefermengen und Gebäuden auszuschließen.
- Darüber hinaus finden nur Gebäude Eingang, bei denen kein Nebengebäude auf dem gleichen Flurstück steht, das als Wohngebäude klassifiziert ist. Diese Bedingung schließt ebenfalls Fehler bei der Zuordnung von Gasliefermengen und den Wohn- bzw. Bezugsflächen aus.
- Es werden nur Gebäude in die Auswahl aufgenommen, für die ein konkretes Baujahr aus einer der Quellen ermittelt werden kann. Hierbei berücksichtigt werden nur Quellen, die adressscharfe Angaben beinhalten und mit einem Güteindikator größer fünf bewertet sind.
- Bei der Analyse werden nur Gebäude mit einem zugeordneten Endenergieverbrauch für Raumwärme größer Null berücksichtigt.

Durch die genannten Bedingungen werden räumlich klar voneinander getrennte Gebäude mit jeweils eigenen Flurstücken in die Auswahl aufgenommen. Innenstadtbebauung mit gemischter Gebäudenutzung und Nebengebäuden auf dem gleichen Flurstück werden herausgefiltert.

6.1.2. Gebäudedaten und anonymisierte Gasliefermengen

Gasliefermengen sind die durch den Energieversorger abgerechneten Gasmengen, die zu jeweils einem Anschluss gehören. Falls sie einzelnen Personen oder Wohnungen zugeordnet werden können, unterliegen sie als personenbezogene Daten besonderen Datenschutzbestimmungen. Aufgrund ihrer wenigen Zahl von Wohnungen betrifft dies insbesondere Ein- und Zweifamilienhäuser. Der Datenschutz fordert die Zustimmung der betroffenen Personen zur Verarbeitung ihrer Daten. Der EWE NETZ GmbH ist es aufgrund dieser Bestimmungen und möglicher Irritationen von Kunden nicht möglich, personenbezogene Gasverbrauchsdaten ohne ein direktes eigenes geschäftliches Interesse für wissenschaftliche Zwecke zur Verfügung zu stellen.

Aus diesem Grund findet eine Anonymisierung der Verbräuche statt, indem auf Basis der Gebäudeparameter Cluster mit jeweils mindestens 15 Gebäuden pro Cluster gebildet werden. In einem Cluster befinden sich Gebäude mit ähnlichen Angaben zu den Gebäudeparametern Volumen, A/V-Verhältnis und Baualtersklasse. Diese Gebäudeparameter sind Einflussfaktoren bei der Wärmebedarfsrechnung (siehe Kapitel 2.4). Daher soll ihr Einfluss auf die tatsächlichen Wärmeverbräuche untersucht werden. Die Baualtersklassen umfassen neun Klassen. Das A/V-Verhältnis und Volumen sind metrische Variablen, die gleich große Intervalle eingeteilt werden. Die Intervallbreite des Volumens ist 60 m^3 , das A/V-Verhältnis wird in Schritte von $0,1 \text{ m}^2/\text{m}^3$ eingeteilt. Abbildung 6.1 gibt separat für alle drei Gebäudeparameter die Häufigkeitsverteilung der Intervalle an. Die Balkenbreite

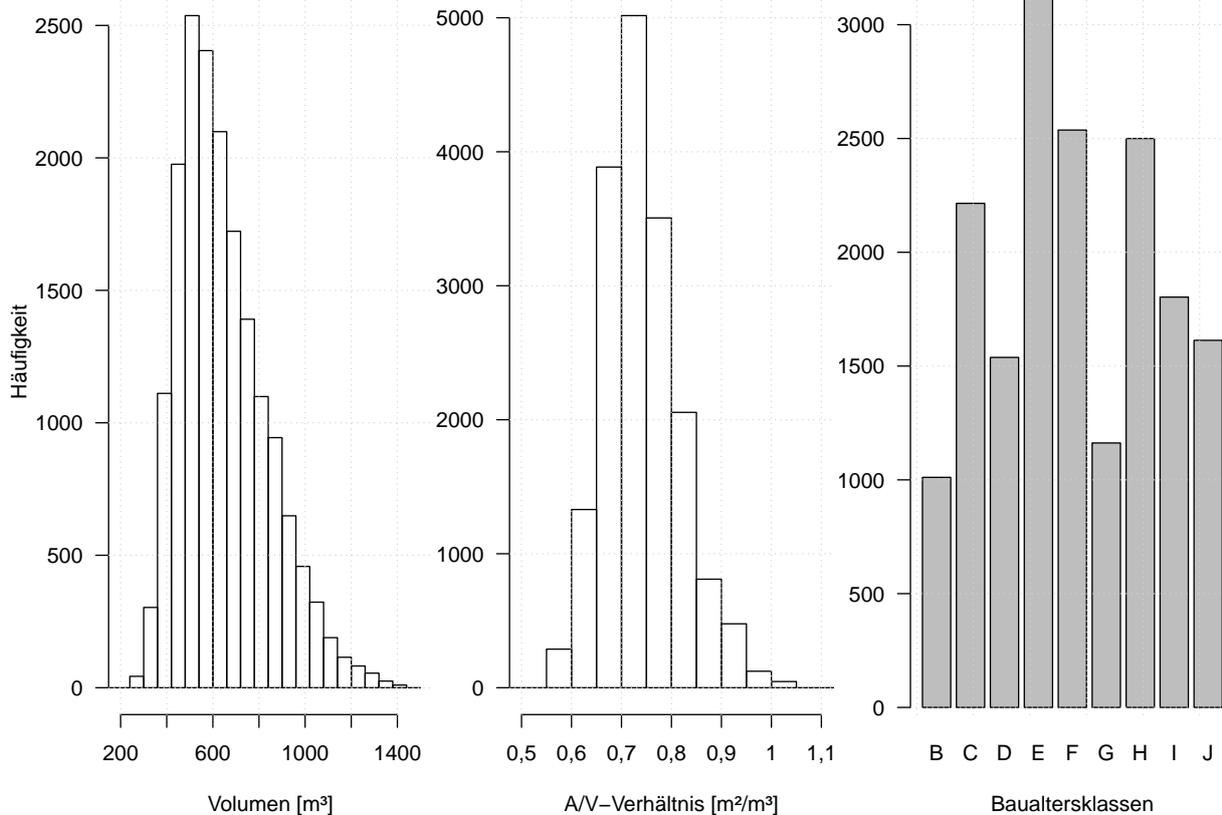


Abbildung 6.1.: Häufigkeitsverteilung der Gebäudeparameter Volumen, A/V-Verhältnis und Baualtersklasse

der Histogramme entspricht den Schrittweiten. 17.537 Gebäude werden durch die Kombination der Intervalle in insgesamt 280 Cluster eingeteilt. Jedes Cluster besitzt mindestens 15 Gebäude zum Zwecke der Anonymisierung. Zum Teil besitzen die Cluster jedoch mehr als 100 Gebäude. Das mittlere Gebäudevolumen ist 645 m^3 , das mittlere A/V-Verhältnis $0,74 \text{ m}^2/\text{m}^3$.

Die Übermittlung, welche Gebäude zu dem jeweiligen Cluster gehören, basiert auf den Anschlussobjekt-IDs der Gasanschlüsse. Seitens der EWE NETZ werden die anschlussobjektscharfen Verbrauchsdaten den Clustern zugeordnet und in eine zufällige Reihenfolge gebracht. Zurück übermittelt werden Einzeldatensätze, die jedoch nur noch mit dem entsprechenden Cluster verbunden sind. Ein Rückschluss auf einzelne Gebäude ist daher nicht möglich. Der Vorteil dieser Vorgehensweise ist die Verarbeitung nicht-personenbezogener Daten, die trotzdem als Einzeldatensätze für die notwendigen statistischen Untersuchungen vorliegen. Jedoch ist nicht mehr das genaue Volumen bzw. das genaue A/V-Verhältnis der Gebäude bekannt, sondern nur der Clusterwert. Daraus resultierende Toleranzen werden in Kapitel 6.1.4 behandelt.

Für alle folgenden Analyseschritte beinhaltet ein Gebäudedatensatz die in Tabelle 6.1 beispielhaft gelisteten Angaben.

Tabelle 6.1.: Beispiel eines Gebäudedatensatzes für die Verbrauchsauswertung

Angabe	Beispiel	Unsicherheit
Volumen-Klasse	660 m ³	± 30 m ³
A/V-Verhältnis-Klasse	0,8 m ² /m ³	± 0,05 m ³
Baualtersklasse	D (1945 - 1957)	
Verbrauch	32.465 kWh/m ² a	

6.1.3. Berechnung klimakorrigierter Endenergieverbräuche für Raumwärme

Alle angegebenen und berechneten Wärmeverbräuche stellen Endenergiewerte dar. Anlagenerluste und Leitungsverluste im Gebäude sind in den angegebenen Energiemengen enthalten, da diese aufgrund fehlender Daten über die Heizsysteme der Gebäude nicht herausgerechnet werden können. Für die Berechnung der für Raumwärme aufgewendeten Endenergie wird die Endenergie für Warmwasser von den Gasliefermengen abgezogen. Der Rechnung liegt die Annahme zugrunde, dass bei einem Gasanschluss Warmwasser ebenfalls mit Gas erwärmt wird. Der mittlere Endenergieverbrauch Q_{WW}^- für Warmwasser wird auf Basis des gesamten Endenergieverbrauchs, das entspricht den Gasliefermengen $Q_{Gas,t}$ abgeschätzt. Die Gasliefermengen hängen vom Jahr t der Angabe ab. Die Berechnung basiert auf der Studie „Vergleichswerte für Verbrauch bei Wohngebäuden“ [2] und auf der Publikationsreihe „Universelle Energiekennzahlen für Deutschland“ [32].

Die Publikationen berechnen den prozentualen Anteil der Endenergie für die Warmwasseraufbereitung an der gesamten verbrauchten Endenergie für ca. 150.000 Messwerte. Die Höhe des Anteils hängt von der Endenergie ab und beträgt für niedrige spezifische Endenergieverbrauchswerte von 50 kWh/m²a ca. 34 %, bei einem spezifischen Verbrauch von 200 kWh/m²a nur noch 16 %. Um für die in dieser Arbeit vorliegenden kontinuierlichen Gasliefermengen den Warmwasseranteil berechnen zu können, werden die Angaben der spezifischen Warmwasser-Verbrauchswerte q_{WW} pro m² Wohnfläche (WF) [2] in Abhängigkeit der spezifischen Gasliefermengen q_{Gas} linear interpoliert und ergeben Gleichung (6.1).

$$q_{WW}^- = 0,1025 \cdot q_{Gas} + 11,7 \frac{\text{kWh}}{\text{m}_{WF}^2 \text{a}} \quad (6.1)$$

Bezugsgröße der spezifischen Werte ist die Nutzfläche bzw. Wohnfläche F_{WF} eines Gebäudes. Wenn diese über den Zwischenschritt der beheizten Gebäudenutzfläche F_{hN} aus dem Gebäudevolumen geschätzt wird (Umrechnungswerte siehe Abbildung 5.14), erhält man Gleichung (6.2) für die Schätzung des Endenergieverbrauchs Q_{WW}^- für die Warmwasseraufbereitung. Er hängt vom Volumen und dem Gesamtenergieverbrauch ab.

Bei dieser Art der Bestimmung des Warmwasserverbrauchs wird das Nutzerverhalten, das sich im Gesamtenergieverbrauch widerspiegelt, auf den Warmwasserverbrauch übertragen. Je geringer der Gesamtenergieverbrauch, desto geringer ist auch der Warmwasserverbrauch. Zugleich wird berücksichtigt, dass sich der relative Anteil des Warmwassers

am Gesamtenergieverbrauch mit sinkendem Verbrauch vergrößert.

$$\begin{aligned}
 Q_{WW}^- &= q_{WW}^- \cdot F_{WF} = 0,1025 \cdot Q_{\text{Gas}} + 11,7 \frac{\text{kWh}}{\text{m}_{\text{WF}}^2 \text{a}} \cdot F_{WF} \\
 &= 0,1025 \cdot Q_{\text{Gas}} + 11,7 \frac{\text{kWh}}{\text{m}_{\text{WF}}^2 \text{a}} \cdot \frac{F_{hN}}{V_e} \cdot \frac{F_{WF}}{F_{hN}} \cdot V_e \\
 &= 0,1025 \cdot Q_{\text{Gas}} + 11,7 \frac{\text{kWh}}{\text{m}_{\text{WF}}^2 \text{a}} \cdot 0,32 \frac{\text{m}^2}{\text{m}^3} \cdot 0,8 \cdot V_e \quad (6.2)
 \end{aligned}$$

Um klimatische Schwankungen einzelner Jahre auszugleichen, werden die Endenergiewerte für Raumwärme mit Klimafaktoren des Deutscher Wetterdienst (DWD) auf Basis des seit 2011 gültigen Klimakorrekturenverfahrens korrigiert. Das Verfahren und die Korrekturwerte für Oldenburg sind in Kapitel 2.4.2 angegeben. Der klimabereinigte Endenergieverbrauch für Heizwärme Q_H wird mit den gebiets- und jahresabhängigen Klimafaktoren und unter Abzug des Warmwasserverbrauchs durch Gleichung (6.3) berechnet.

$$Q_H = (Q_{\text{Gas},t} - Q_{WW}^-) \cdot f_{K,t} \quad (6.3)$$

6.1.4. Berechnung spezifischer Wärmeverbräuche

Der spezifische Wärmeverbrauch q_H wird in dieser Arbeit auf die Energienutzfläche bezogen und ergibt sich nach Gleichung (6.4) [17, Anlage 1].

$$q_H = Q_H / (V_e \cdot 0,32) \quad (6.4)$$

Durch die Anonymisierung bei der Übertragung der Gasliefermengen ist für das Volumen aller Gebäude nur der Mittelwert des Volumenintervalls bekannt, zu dem jedes Gebäude gehört. Die Abweichungen zwischen tatsächlichem Volumen und Intervall-Mittelwert erzeugen Fehler bei der Berechnung der spezifischen Wärmeverbräuche. Die maximale Abweichung δV ist die halbe Intervallbreite, d. h. 30 m^3 . Das minimale Volumen von Gebäuden liegt bei 300 m^3 . Es wird davon ausgegangen, dass der Fehler im Vergleich zum Volumen klein ist und somit die lineare Fehlerfortpflanzungsformel angewendet werden kann [8].

$$\Delta q_H = \frac{\delta q_H}{\delta V} \Delta V = \frac{-Q}{0,32 \cdot V^2} \Delta V = \frac{-0,32 \cdot V q_H}{0,32 \cdot V^2} \Delta V = -q_H \frac{\Delta V}{V} \leq -0,108 q_H \quad (6.5)$$

Somit ergibt sich in Gleichung (6.5) mit Einsetzen von Gleichung (6.4) eine Abweichung des spezifischen Verbrauchs in Abhängigkeit des Volumens und des spezifischen Verbrauchs. Durch Einsetzen des kleinsten vorkommenden Volumens ($\min(V) = 267 \text{ m}^3$) und der maximalen Abweichung erhält man eine maximale Abweichung von 10,8% des spezifischen Verbrauchs, die mit $1/V$ abnimmt.

Die Abweichung bezieht sich jedoch nicht auf die Angaben in Kapitel 6.2.1, da diese Angaben bereits vor der anonymisierten Datenübergabe mit genauen gebäudespezifischen Volumina in spezifische Verbräuche nach Gleichung 6.4 umgerechnet werden. Für die

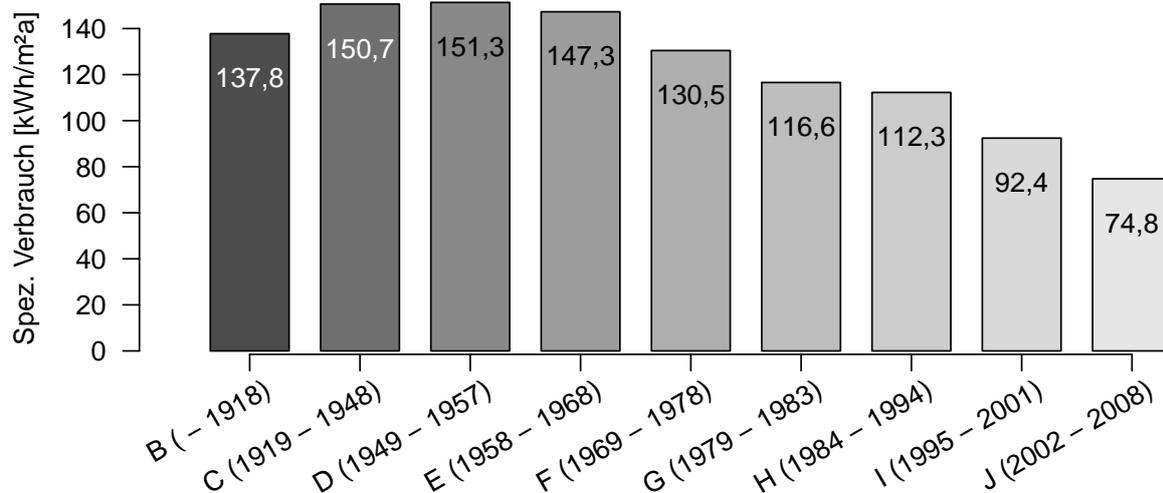


Abbildung 6.2.: Mittlerer spezifischer Endenergieverbrauch je m² Energienutzfläche

Auswertungen in Kapitel 6.3 und 6.4 kann dies nicht mit allen Angaben durchgeführt werden.

6.2. Mittlere Endenergieverbräuche und Literaturvergleich

Mittlere spezifische Wärmeverbräuche für verschiedene Kombinationen von Gebäudeparametern ergeben Kennwerte für den Vergleich der Daten mit Literaturangaben und zeigen, inwiefern Gebäudeparameter und Verbrauchswerte zusammenhängen. Durch die Mittelung der Verbräuche werden nutzerspezifische Schwankungen des Verbrauchs eliminiert.

6.2.1. Differenzierung nach Baualtersklasse und Gebäudebaujahr

In einem ersten Schritt werden die spezifischen mittleren Verbräuche der Ein- und Zweifamilienhäuser getrennt nach Baualtersklassen bzw. Baujahr analysiert. Abbildung 6.2 zeigt die spezifischen mittleren Endenergieverbräuche der jeweiligen Baualtersklasse. Bezugsgröße ist die Energienutzfläche des Gebäudes. Die Baualtersklassen, bezeichnet mit den Buchstaben B bis J, sind in Tabelle 5.2 auf Seite 65 näher beschrieben. Den höchsten spezifischen Verbrauch zeigen Gebäude der Baualtersklasse D. In diesem Zeitraum wurde die vorherrschende Wohnungsnot nach dem zweiten Weltkrieg zügig gelindert. Die Bauqualität rückte dabei in den Hintergrund. Auch in der darauf folgenden Baualtersklasse E änderte sich daran nicht viel. Für Gebäude ab der Baualterklasse F sind kontinuierlich sinkende spezifische Verbräuche zu beobachten.

Eine feinere zeitliche Auflösung bietet Abbildung 6.3. Sie gibt den mittleren spezifischen und klimabereinigten Endenergieverbrauch des Jahres 2009 von Ein- und Zweifamilienhäuser mit einem entsprechenden Baujahr seit 1900 an. Die grauen Balken im unteren Teil der Grafik bezeichnen die Zahl der Gebäude pro Baujahr, die in die Mittelwertbildung eingegangen sind. Deutlich sichtbar sind Lücken bei der Bebauung, hervorgerufen durch die zwei Weltkriege mit einer zeitlichen Verschiebung von ein bis zwei Jahre.

Auf der - bezüglich der Grafik - oberen Zeitachse und durch die vertikalen gepunkteten Linien sind die IWU-Baualtersklassen angetragen. Es ist jeweils das erste Jahr jeder neuen Klasse und der die Klasse repräsentierende Buchstabe angegeben. Die Einteilung der letzten vier Klassen ab 1979 richtet sich nach dem Inkrafttreten der Wärmeschutzverordnung (WSVO) bzw. den Novellierungen dieser und der Einführung der EnEV ab 2002 (siehe auch Tabelle 5.2).

Die waagerechten blauen Striche und Angaben repräsentieren die mittleren Endenergieverbräuche je Baualtersklasse in Abbildung 6.2. Bis Ende der 60er Jahre ist die Schwankungsbreite der jährlichen mittleren spezifischen Wärmeverbräuche groß. Die mittleren Verbräuche je Baualtersklasse unterscheiden sich jedoch kaum für die Baualtersklassen C bis E. Bereits für Gebäude mit Baujahr vor der ersten Wärmeschutzverordnung (Inkrafttreten 1979) gehen die Schwankungen der spezifischen mittleren Verbräuche aufeinanderfolgender Jahre deutlich zurück. Es kann vermutet werden, dass durch die Einführung von Normen in der Baupraxis und noch stärker durch die Einführung bindender gesetzlicher Rahmenbedingungen eine Angleichung energetischer Standards stattfand.

Seit Anfang der 70er Jahre ist ein klarer Abwärtstrend zu beobachten. Eine Stufenform des Wärmeverbrauchs, die aufgrund der regulatorischen Maßnahmen zu erwarten ist, zeigt sich ab der 3. Wärmeschutzverordnung. Zugleich ist aus der Abbildung jedoch ersichtlich, dass auch große Unterschiede innerhalb einer Baualtersklasse bestehen (z. B. Baualtersklasse F oder H), die durch die im Laufe der Jahre kontinuierlich verbesserte Baupraxis begründet werden. Die Daten belegen, dass Novellierungen der Wärmeschutzverordnungen antizipiert werden [70, S. 46].

Nicht erklärt werden kann der erhöhte spezifische Wärmeverbrauch für Gebäude mit Baujahr um 1990. Er könnte an einer nur unzureichenden Umsetzung der Wärmeschutzverordnungen liegen.

Vereinfachte Baujahresklassifizierung Die Baualtersklassen C, D und E weisen sehr ähnliche mittlere spezifische Verbräuche auf. Sie können daher nahezu ohne Informationsverlust in einer Klasse zusammengefasst werden. In verschiedenen Wärmeverbrauchsstudien wird ebenfalls eine vereinfachte Baualtersklassifizierung von nur vier Klassen verwendet (siehe auch Tabelle 5.2) [77, 33, 70].

Für die Einführung einer vereinfachten Klassifizierung auf Basis der jahresscharfen mittleren spezifischen Wärmeverbräuche wird ein hierarchisches Average-Linkage-Clusterverfahren verwendet [42]. Die Datenpunkte werden hierbei iteriert in hierarchisch angeordnete Cluster einsortiert. Durch die Wahl der Hierarchieebene kann damit die Zahl der

6. Analyse der Wärmeverbräuche von Ein- und Zweifamilienhäusern in Oldenburg

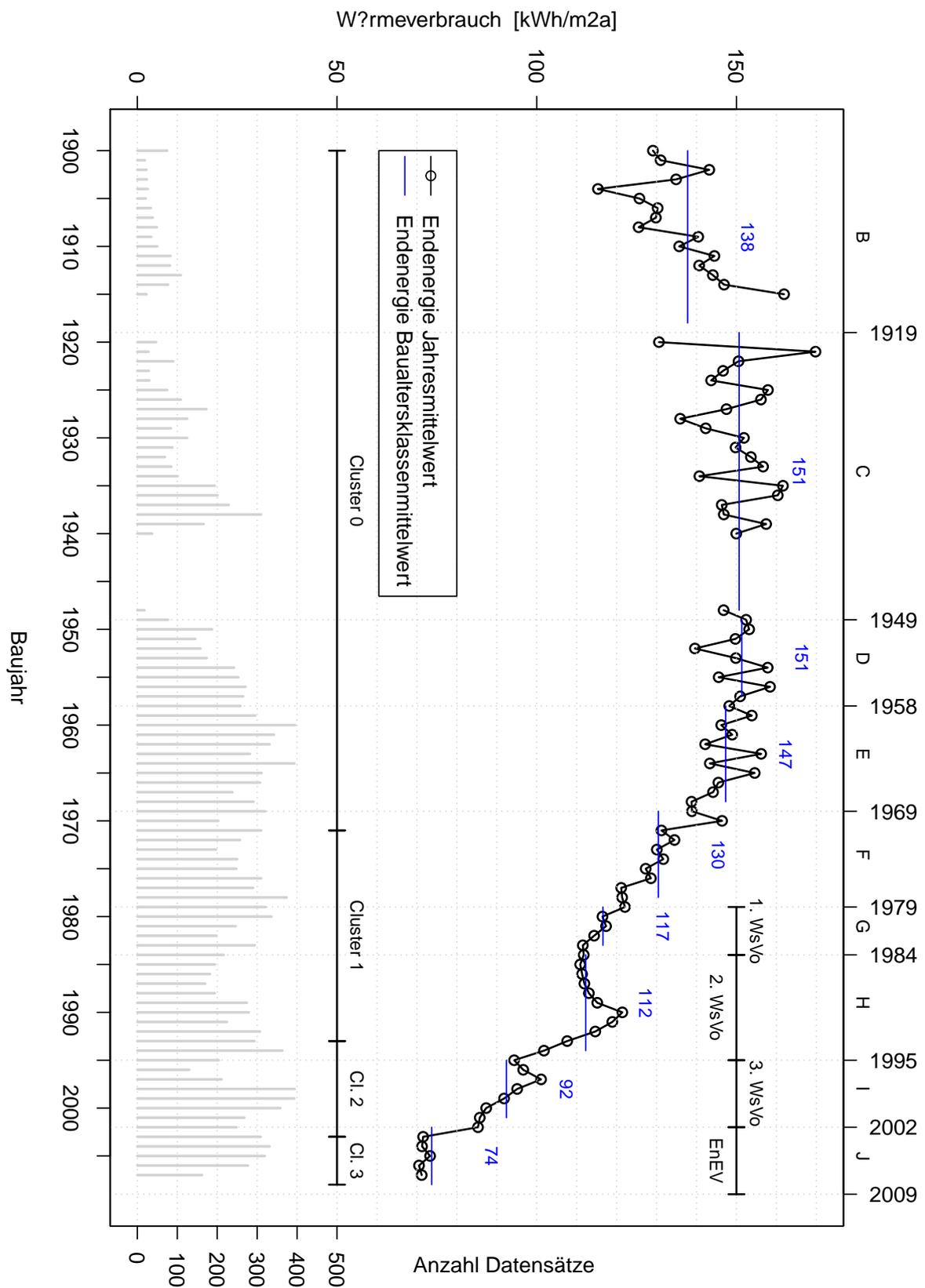


Abbildung 6.3.: Mittlere spezifische Endenergie je Baujahr, Mittelwert je IWU-Baualtersklasse, Anzahl der zugrundeliegenden Datensätze und Klassen einer hierarchischen Average-Linkage-Clustering

Cluster beeinflusst werden. In Anlehnung an oben zitierte Studien wird in Abbildung 6.3 das Ergebnis mit vier Clustern dargestellt. Die Cluster sind nummeriert von 0 bis 3.

Cluster 0 beschreibt den gesamten Zeitraum ab 1900 bis 1971. Der Zeitraum der ersten und zweiten Wärmeschutzverordnung wird zu Cluster 1 zusammen gefasst. Dabei machen sich die vermuteten antizipierten Mindeststandards und die Wirkungen der Ölpreiskrisen bemerkbar. Sie bewirken, dass Cluster 1 bereits 1971 beginnt anstatt 1979, in dem die erste Wärmeschutzverordnung in Kraft trat. Der Zeitraum von Cluster 2 beginnt zwei Jahre vor der dritten Wärmeschutzverordnung 1995. Cluster 3 beinhaltet die Jahre ab 2003 und entspricht damit nahezu der Einführung der EnEV.

Schlussfolgerung und vereinfachte Datenaufnahme des Baujahres Die Clusterung führt dazu, dass die Baualtersklassen nicht exakt mit den Jahren des Inkrafttretens der Wärmeschutzverordnungen definiert werden, wie dies verschiedene Wärmeverbrauchsstudien tun (siehe Tabelle 5.2). Die Ergebnisse der baujahresscharfen Analyse der Endenergieverbräuche stützen die These der Antizipation von Wärmeschutzmaßnahmen vor dem Inkrafttreten der Verordnungen [70].

Die vereinfachte Baualtersklassifizierung kann die Bestimmung der Baujahre für die Gebäude von Siedlungsgebieten enorm erleichtern. Unter der Annahme, dass eine Kommune seit ca. Mitte der 90er Jahre die Bauanträge digital bearbeitet, liegen für nahezu alle Gebäude, die im Zeitraum von Cluster 2 und 3 gebaut sind, ein Datensatz mit Datum der Baufertigstellung vor. Beispielsweise in Oldenburg werden Bauanträge seit 1996 digital bearbeitet. Für alle anderen Gebäude ist dann nur noch eine Einteilung in Cluster 0 und 1, den Zeitraum vor 1971 und ab 1971 notwendig. Diese Einteilung könnte manuell durch Luftbilder oder die Analyse historischer Karten geschehen, da nur Luftbilder eines einzigen Aufnahmezeitpunktes um das Jahr 1971 ausgewertet werden müssten. Damit ließen sich alle Gebäude identifizieren, die bis dahin gebaut wurden und zu Cluster 0 gehören. Gebäude, die nicht zu Cluster 2 und 3 gehören und nicht in den Luftbildern auftauchen, sind dann Cluster 1 zuzuordnen. Seit 1972 gibt es kontinuierliche Luftbilder auf Basis von Daten des Landsat-Satelliten, die zur Bestimmung von Neubauten heran gezogen werden können [29]. Validiert werden können die Daten durch die Baufertigstellungsstatistik, die beispielsweise für Oldenburg ab Ende der 70er Jahre jahresscharfe Daten liefert.

6.2.2. Vergleich mit Literaturwerten

Es existieren nicht viele Literaturangaben zum Heizwärmeverbrauch von Ein- und Zweifamilienhäusern. Vergleichswerte für die Nutzenergie Heizen bietet das IWU durch die im Rahmen der Tabula-Studie erarbeiteten verbrauchskorrigierten Wärmebedarfe. Hierbei wurde auf der Grundlage von ca. 1.700 Gebäuden in Deutschland ein wärmebedarfsabhängiger Faktor gebildet, um aus dem anhand der Standardrahmenbedingungen berechneten Wärmebedarf den Verbrauch abzuschätzen [53]. Für den Vergleich mit den gemittelten spezifischen Endenergieverbräuchen in Oldenburg werden die Heizwärmeangaben des IWU mit einem Anlagenfaktor $\eta_H = 1,1$ und einem Verteilfaktor $\eta_V = 1,16$ multipliziert

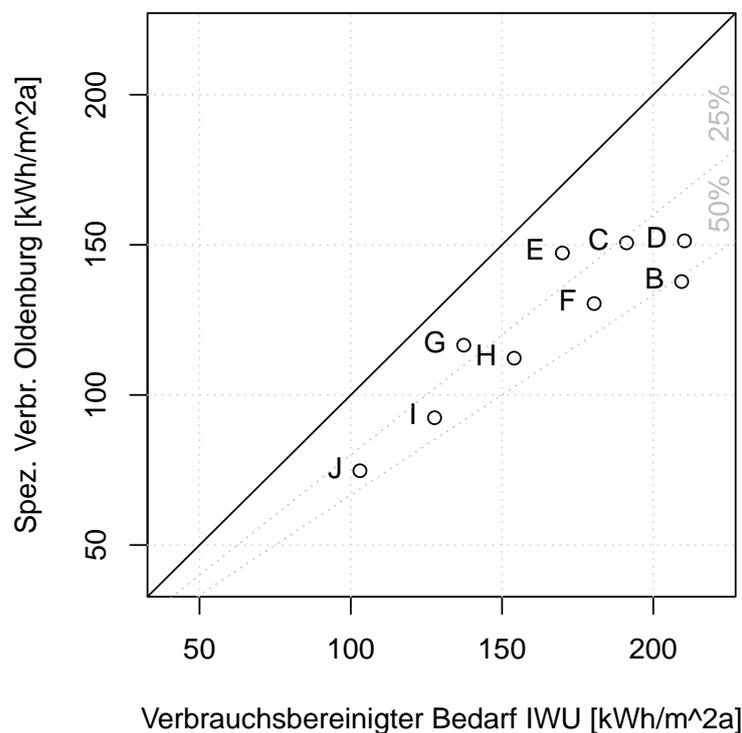


Abbildung 6.4.: Mittlerer spezifischer Endenergieverbrauch pro m² Energienutzfläche und Literaturvergleich

[51]. Der Vergleich in Abbildung 6.4 zeigt, dass die meisten IWU-Werte zwischen 25 % und 50 % über den Oldenburger spezifischen Verbräuchen liegen.

Im Rahmen der Arbeit wurde die Bachelorarbeit „Ermittlung von Energiestandards nach regionalen Gebäudetypologien der Stadt Oldenburg“ an der Jade Hochschule Oldenburg initiiert und betreut. Darin werden auf Basis von Energieberatungsprotokollen, Expertengesprächen und Literaturrecherchen die U-Werte (Wärmeleitwerte) von Ein- und Zweifamilienhäusern in Oldenburg mit der Gebäudetypologie des IWU verglichen [26]. Der Vergleich beschränkt sich auf die Baujahre zwischen 1949 und 1994 und bezieht sich auf den Urzustand der Gebäude. Die Ergebnisse der Arbeit lassen vermuten, dass die Abweichungen nicht die in Abbildung 6.4 gezeigten Unterschiede begründen. Die für Oldenburg ermittelten U-Werte weisen für Bodenplatte, Kellerdecke und Dachschräge sogar höhere Angaben als die des IWU aus.

6.2.3. Differenzierung nach Baualtersklasse, Volumen und A/V-Verhältnis

Die oben beschriebene Auswertung der spezifischen Verbräuche basiert nur auf einer Differenzierung der Ein- und Zweifamilienhäuser anhand ihres Baujahres bzw. der Baualtersklassen. Darauf aufbauend wird der baualtersklassenabhängige Einfluss der Parameter Volumen und A/V-Verhältnis untersucht. Abbildung 6.5 gibt den mittleren spezifischen Verbrauch in Abhängigkeit aller drei Gebäudeparameter an. Jede Unterabbildung bezieht

6.2. Mittlere Endenergieverbräuche und Literaturvergleich

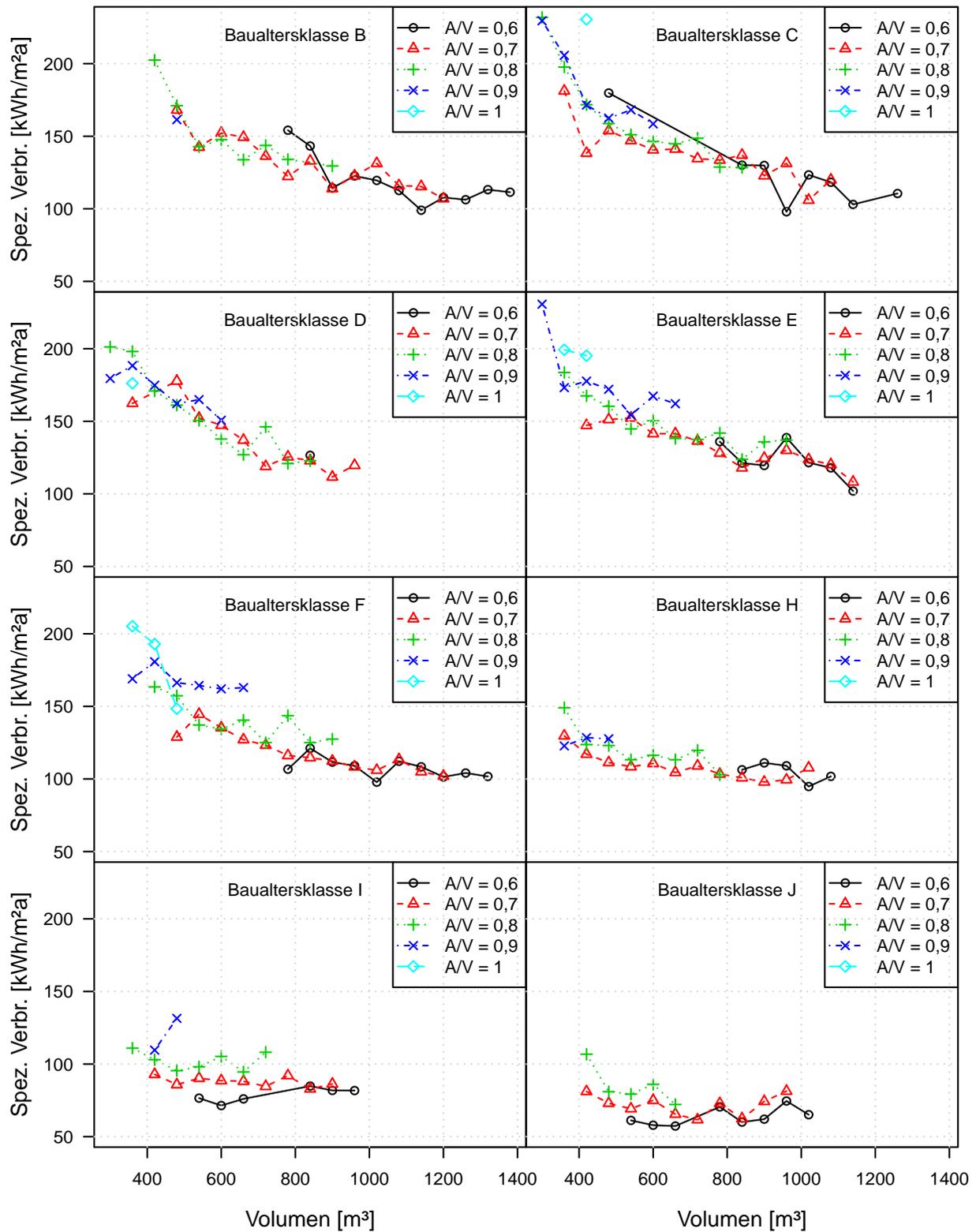


Abbildung 6.5.: Mittlerer spezifischer Endenergieverbrauch pro m² Energienutzungsfläche je Baualterklasse, Volumen und A/V-Verhältnis

sich auf eine Baualtersklasse. Darin ist der mittlere spezifische Verbrauch als Funktion der Volumina angetragen. Die verschiedenen Kurvenfarben und -strichtypen unterscheiden die Intervalle des A/V-Verhältnisses zwischen $0,6 \frac{\text{m}^2}{\text{m}^3}$ und $1 \frac{\text{m}^2}{\text{m}^3}$.

Die Korrelation zwischen Volumen und A/V-Verhältnis für die Baualtersklassen B bis H liegt zwischen -0,6 und -0,71. Bei großen Volumina zeigen die Kurven ausschließlich kleine A/V-Verhältnisse von $0,6 \frac{\text{m}^2}{\text{m}^3}$ und $0,7 \frac{\text{m}^2}{\text{m}^3}$. Für die Baualtersklasse I liegt der Korrelationskoeffizient bei -0,5, für die Baualtersklasse J bei -0,2. Der Zusammenhang zwischen A/V-Verhältnis und Volumen ist bei Gebäuden neueren Baujahres deutlich weniger ausgeprägt.

Die spezifischen Verbräuche sinken mit steigendem Gebäudevolumen in den Baualtersklassen B bis F stärker als in den jüngeren Baualtersklassen H bis J. Eine lineare Regression gibt ein Steigungsverhältnis m der Ausgleichsgerade von $m = -0,12 \frac{\text{kWh}/\text{m}^2\text{a}}{\frac{\text{m}^3}{\text{m}^3}}$ für die Baualtersklasse D. Bei der Baualtersklasse J ist es dagegen $m = -0,02 \frac{\text{kWh}/\text{m}^2\text{a}}{\frac{\text{m}^3}{\text{m}^3}}$.

Aus Abbildung 6.5 kann außerdem abgelesen werden, dass sich für ältere Baualtersklassen B bis F die mittleren spezifischen Verbräuche unterschiedlicher A/V-Verhältnisse bei gleichem Volumen in vielen Fällen nicht voneinander unterscheiden. Für die Baualtersklassen I und J zeigt der spezifische Verbrauch jedoch eine Abhängigkeit vom A/V-Verhältnis bei gleichem Volumen.

Eine Interpretation der Ergebnisse ist, dass neue Gebäude dichter sind. Bei diesen ist daher die Volumenabhängigkeit des Energieverbrauchs nicht so groß, da es nahezu keinen ungewollten Luftaustausch gibt. Bei älteren Gebäuden ist der ungewollte Luftaustausch größer und hat bei kleineren Gebäuden aufgrund der verkleinerten Energiebezugsfläche einen entsprechend größeren Einfluss auf den Heizenergieverbrauch.

Bei jüngeren Gebäuden legt der Einfluss des A/V-Verhältnisses auf den Energieverbrauch nahe, dass die Form, repräsentiert durch das A/V-Verhältnis einen Einfluss auf den Heizenergieverbrauch hat. Gebäude mit einer minimalen thermischen Hüllfläche, d. h. einem möglichst kleinen A/V-Verhältnis liefern daher bessere Energieverbrauchskennzahlen. Diese Interpretation basiert auf den Erfahrungen der Energiebedarfsrechnung (siehe Kapitel 2.4).

6.3. Verteilung und Multiple Regression gruppierter Heizwärmeverbräuche

Bei der Bildung mittlerer spezifischer Wärmeverbräuche im vorherigen Kapitel wird die große Varianz der Daten nicht berücksichtigt. Dadurch kann dargestellt werden, auf welche Art Gebäudeparameter und Verbrauch zusammenhängen. Die Stärke der Abhängigkeit wird jedoch in diesem Kapitel durch die Berücksichtigung der Einzelwerte und ihrer Verteilung berechnet.

6.3.1. Kumulierte Verteilung des spezifischen Wärmeverbrauchs

Die Endenergieverbräuche für Raumwärme weisen eine hohe Streuung zwischen knapp 40 kWh/m²a und über 300 kWh/m²a auf. Die Darstellung der Verteilung als Histogramme - wie beispielsweise in Abbildung 6.1 - erlaubt nicht den grafischen Vergleich verschiedener Verteilungen. Kumulierte Verteilungen machen dies möglich. Dabei werden die Wärmeverbräuche ihrer Größe nach sortiert dargestellt.

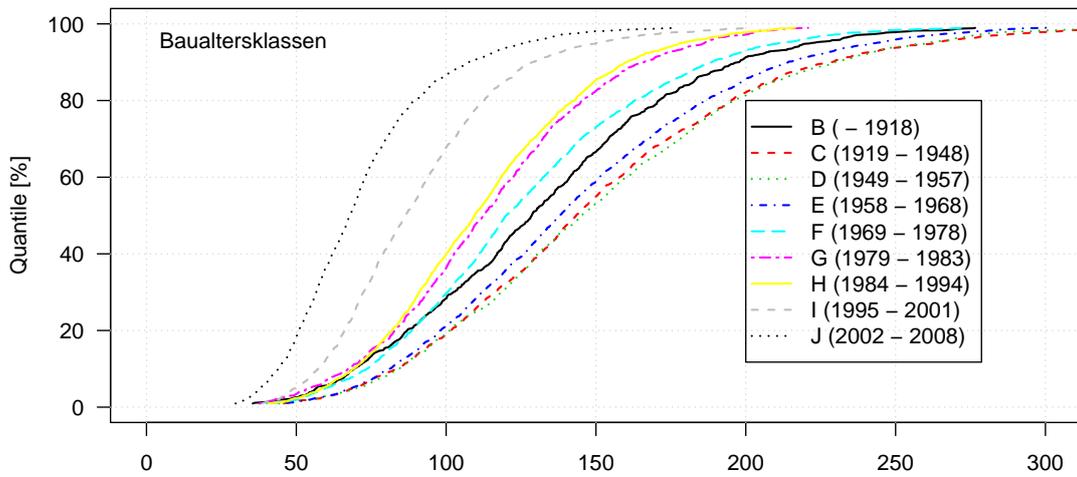
Abbildung 6.6 zeigt die kumulierten Verteilungen der spezifischen Wärmeverbräuche. Jede Kurve stellt eine Verteilung dar und beinhaltet die Verbräuche aller Gebäude mit dem entsprechenden Parameterwert. Auf der X-Achse ist der spezifische Endenergieverbrauch, auf der Y-Achse die Quantile als Funktion der spezifischen Verbräuche angetragen. Es entstehen damit Kurven in S-Form. Je größer die Steigung einer Kurve ist, desto mehr Gebäude mit einem entsprechenden Verbrauch kommen in den Daten vor. Je weiter rechts eine Verteilung liegt, desto höhere spezifische Verbräuche weist sie auf. Je größer der horizontale Abstand zweier Kurven ist, desto deutlicher unterscheiden sich die Verteilungen voneinander. Statistische Kenngrößen wie beispielsweise der Median und Quantile können aus der Darstellung direkt abgelesen werden.

In Abbildung 6.6(a) ist der Unterschied der Verteilung der spezifischen Wärmeverbräuche zwischen den jüngeren Baualtersklassen H bis J deutlich sichtbar. Die Mediane (50 %-Quantil) der Verteilungen haben einen Abstand von bis zu 20 kWh/m²a. Die oberen 15 % der Gebäude dieser jüngeren Baualtersklassen verbrauchen über 100 kWh/m²a. Die Verteilungen zeigen eine große Überlappung, d. h. die unterschiedlichen Kurven liegen vertikal übereinander. Alle Verteilungen haben Werte im Intervall zwischen 50 kWh/m²a und 170 kWh/m²a. Anhand der Baualtersklasse ist somit keine scharfe Aussage eines spezifischen Wärmeverbrauchs möglich. Eine große Einflussgröße für die Streuung der Verbräuche ist das Nutzerverhalten wie beispielsweise die Anwesenheitszeiten von Personen im Gebäude oder die durchschnittlichen Innenraumtemperaturen. Die hohe Streuung der Verbrauchsdaten tritt auch in veröffentlichten Wärmeverbrauchsstudien deutlich hervor [69, 33].

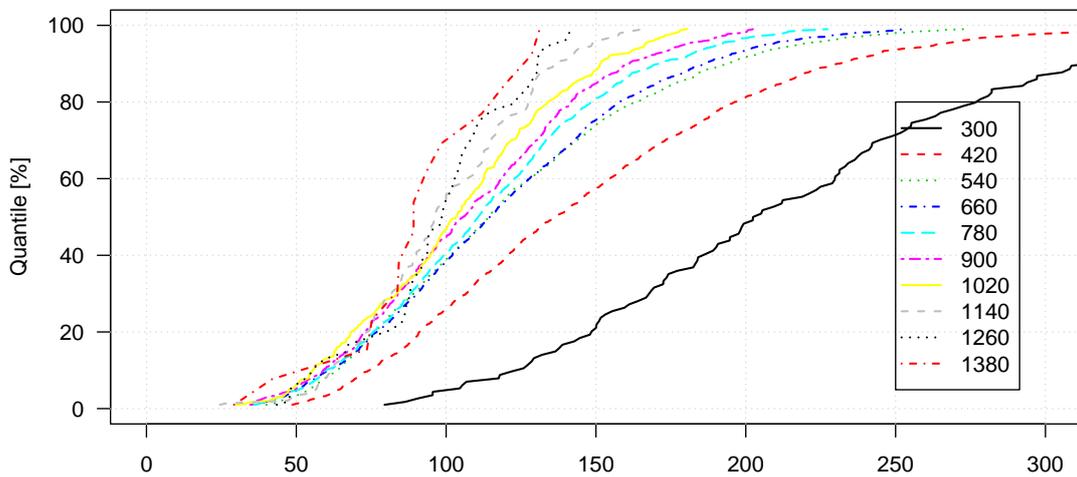
Für Abbildung 6.6(b) werden die Gebäude anhand der Volumina in Gruppen zwischen 300 m³ und 1380 m³ unterschieden. Gebäude mit kleinen Volumina unter 450 m³ zeigen einen sehr hohen und zu den anderen Kurven deutlich abgesetzte Verteilungen. Die Unebenheit der Kurve mit 300 m³ resultiert daraus, dass die Datenbasis nicht so viele Gebäude umfasst. Gebäude mit kleinen Volumina unter 450 m³ und zeigen eine sehr hohe Streuung der spezifischen Wärmeverbräuche. Dies ist zum Teil auch darauf zurück zu führen, wie der spezifische Verbrauch in Kapitel 6.1.4 berechnet wird. Die Schwankungsbreite der Verbräuche von Gebäuden mit großen Volumina über 1000 kWh/m³ ist sehr gering. Hinter den Verteilungen stehen jedoch nur sehr wenige Gebäude, was deutlich an dem kurvigen Verlauf der Verteilung sichtbar ist. Die Verteilung von nahezu allen Volumenklassen liegen sehr eng beieinander für spezifische Verbräuche unter 140 kWh/m²a.

Abbildung 6.6(c) zeigt, dass Gebäude mit einem kleineren A/V-Verhältnis geringere spezifische Verbräuche aufweisen. Zu der optisch klaren Abgrenzung der einzelnen Kurven

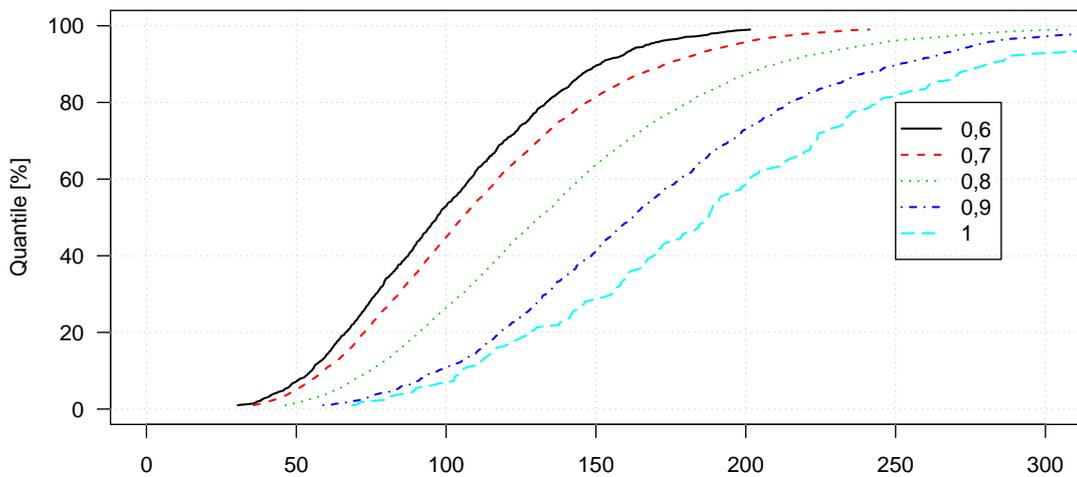
6. Analyse der Wärmeverbräuche von Ein- und Zweifamilienhäusern in Oldenburg



(a) Baualterklassen: Spez. Wärmeverbrauch in kWh/m²a



(b) Volumen: Spez. Wärmeverbrauch in kWh/m²a



(c) A/V-Verhältnis: Spez. Wärmeverbrauch in kWh/m²a

Abbildung 6.6.: Kumulierte Verteilung der spezifischen Endenergieverbräuche

führt die Tatsache, dass nur fünf Klassen unterschieden werden und die Intervallbreite des A/V-Verhältnisses mit $0,1 \text{ m}^2/\text{m}^3$ relativ grob ist. Wie bei Abbildung 6.5 beschrieben, korrelieren kleine A/V-Verhältnisse mit größeren Gebäudevolumina. Je größer das Gebäudevolumen, desto geringer ist der spezifische Verbrauch (siehe Abbildung 6.6(b)).

Möglicher Hinweis auf Sanierung Die Verteilungen jüngerer Baualtersklassen weisen in Abbildung 6.6(a) eine stärkere Steigung, d. h. eine geringere Varianz der spezifischen Heizwärmeverbräuche auf als die Verteilung älterer Baualtersklassen. Das 5 %-Quantil für die Daten der Baualtersklasse J liegt bei $38,2 \text{ kWh}/\text{m}^2$, das 95 %-Quantil bei $126,4 \text{ kWh}/\text{m}^2$. Somit liegen 90 % der Verbräuche innerhalb der Differenz beider Werte von $88,1 \text{ kWh}/\text{m}^2$. Bei der Baualtersklasse B beträgt das 5 %-Quantil $56,9 \text{ kWh}/\text{m}^2$, das 95 %-Quantil $221,9 \text{ kWh}/\text{m}^2$, der Unterschied zwischen beiden ist $165 \text{ kWh}/\text{m}^2$.

Diese höhere Streuung kann einerseits darauf zurückzuführen sein, dass das Nutzerverhalten in älteren Gebäuden deutlich heterogener als in jüngeren Gebäuden ist. Es gibt z. B. Personen, die aufgrund des schlechten energetischen Standards sehr sparsam mit Heizwärme umgehen. Weiterhin ist denkbar, dass bei älteren Gebäuden zum Teil eine Sanierung durchgeführt wurde und diese daher geringere Heizwärmeverbräuche aufweisen. Ein genauer Nachweis würde jedoch Informationen über tatsächlich durchgeführte Sanierungen benötigen, die im Rahmen dieser Arbeit nicht ermittelbar waren.

6.3.2. Multiple Regressionsanalyse des Heizwärmeverbrauchs

Ziel der Regressionsanalyse in diesem Kapitel ist eine Ursachenanalyse, die klären soll, wie stark der Einfluss der unabhängigen Variablen Baualtersklasse, Volumen und A/V-Verhältnis - die Gebäudeparameter - auf die abhängige Variable Wärmeverbrauch ist. Der kausale Zusammenhang zwischen den unabhängigen und der abhängigen Variablen ergibt sich durch das Formelwerk der Wärmebedarfsbilanzierung (siehe Kapitel 2.4). Die Stärke des Zusammenhang soll durch die Regressionsanalyse der empirischen Daten ermittelt werden. Eine Einführung in die Regressionsanalyse bietet das Kapitel 2.5.

Die Baualtersklasse der Gebäude ist eine nominale Variable und wird sowohl in der Klassifizierung der neun IWU-Baualtersklassen als auch in der in Kapitel 6.2.1 herausgearbeiteten vereinfachten Klassifizierung untersucht. Das Gebäudevolumen und A/V-Verhältnis sind metrische Variablen. Wegen der in Kapitel 6.1.2 beschriebenen Anonymisierung werden sie durch die Clusterwerte repräsentiert. Der Heizwärmeverbrauch als abhängige Variable ist eine metrische Variable und steht ungerundet zur Verfügung.

Die Gleichungen (6.6) und (6.7) beschreiben zwei Regressionsmodelle mit den Variablen tatsächlicher Heizwärmeverbrauch Q , Gebäudevolumen V , Gebäudebaualtersklasse A , A/V-Verhältnis S und dem Residuum ϵ für ein Gebäude k .

$$Q_k = b_0 + b_V V_k + b_A A_k + b_S S_k + \epsilon_k \quad (6.6)$$

$$\ln(Q_k) = b_0 + b_V V_k + b_A A_k + b_S S_k + \epsilon_k \quad (6.7)$$

Sie unterscheiden sich, da bei ersterer als abhängige Variable des Modells direkt der Heizwärmeverbrauch, bei letzterer der Logarithmus des Heizwärmeverbrauchs verwendet wird. Die Analyse der Residuen beider Modelle zeigt, dass das Logarithmus-Modell von Gleichung (6.7) eher die für die lineare Regression geforderte Prämisse $\epsilon_k \sim N(0, \sigma^2)$ erfüllt und daher besser geeignet ist. Insbesondere die Homogenität der Varianz ist hierbei besser erfüllt. Gleichung (6.7) wird daher für die Untersuchung verwendet.

Die Regressionsanalyse teilt die Varianz des Heizwärmeverbrauchs in einen „zufälligen“, d. h. nicht durch die Parameter erklärbaren Teil und in einen „systematischen“, erklärbaren Teil auf [3]. Die Stärke des Einflusses der unabhängigen Variablen auf die abhängige Variable wird aus dem Quotienten zwischen der systematischen, d. h. erklärbaren Varianz und der gesamten Varianz berechnet und als Bestimmtheitsmaß R^2 bezeichnet. Im folgenden verwendet wird das korrigierte Bestimmtheitsmaß R_{kor}^2 . Das ist das um die Zahl der Freiheitsgrade korrigierte R^2 [3, S. 76]. Die Freiheitsgrade sind die Differenz zwischen der Zahl der unabhängigen Variablen und der Zahl der Gleichungen.

Tabelle 6.2 gibt den Anteil der erklärbaren Varianz an der gesamten Varianz der Wärmeverbräuche für verschiedene Szenarien des Modells (Gleichung (6.7)) an. Die Szenarien unterscheiden sich darin, dass einzelne Summanden des Modells weggelassen werden und damit nicht in die Berechnung des Bestimmtheitsmaß eingehen. Aufgrund der hohen Datenmenge liegen die p-Werte aller Szenarien im Bereich von 10^{-16} . Der Einfluss der unabhängigen Variablen auf die abhängige Variable ist damit hoch signifikant und kann nicht zurückgewiesen werden.

Tabelle 6.2.: Anteil der erklärbaren Varianz an der gesamten Varianz der Endenergieverbräuche

Modellszenario	R_{kor}^2
Baualtersklasse nach IWU (9 Klassen) $\ln(Q_k) = b_0 + b_A A_k + \epsilon_k$	25,6 %
Vereinfachte Baualtersklasse (4 Klassen) $\ln(Q_k) = b_0 + b_A A_k + \epsilon_k$	23,6 %
Volumen $\ln(Q_k) = b_0 + b_V V_k + \epsilon_k$	21,2 %
A/V-Verhältnis $\ln(Q_k) = b_0 + b_S S_k + \epsilon_k$	0,3 %
Baualtersklasse, Volumen $\ln(Q_k) = b_0 + b_V V_k + b_A A_k + \epsilon_k$	39,8 %
Baualtersklasse, Volumen, A/V-Verhältnis $\ln(Q_k) = b_0 + b_V V_k + b_A A_k + b_S S_k + \epsilon_k$	40,0 %

Wenn für die Erklärung der Varianz der Wärmeverbräuche nur die Baualtersklasse der Gebäude herangezogen wird, können 25,6 % der Gesamtvarianz erklärt werden. Die

Regression mit der vereinfachten Baualtersklassifizierung bestätigt die Aussage in Kapitel 6.2.1, dass durch die vereinfachte Klassifizierung nahezu keine Informationen verloren gehen. Mithilfe der vereinfachten Klassifizierung werden 23,6 % der Varianz erklärt. Dies sind nur 2 Prozentpunkte weniger als bei dem Szenario mit 9 Baualtersklassen. In beiden Szenarien spielen die Parameter Volumen und A/V-Verhältnis keine Rolle.

Durch den Parameter Gebäudevolumen sind 21,2 % der Gesamtvarianz erklärt. Dagegen ist der alleinige Einfluss des A/V-Verhältnisses auf die Varianz der Wärmeverbräuche sehr gering mit nur 0,3 %. Dies drückt sich ebenfalls aus im Vergleich der unteren beiden Szenarien in Tabelle 6.2. Auf Basis aller drei verwendeten Gebäudeparameter sind 40,0 % der gesamten Varianz erklärbar, ohne das A/V-Verhältnis immer noch 39,8 %.

Daraus wird jedoch auch deutlich, dass 60 % der Varianz der Wärmeverbräuche somit entweder von Parametern, die im Rahmen dieser Arbeit nicht bekannt sind, oder vom Nutzerverhalten abhängen. Zu den unbekanntem Parametern gehören individuelle Abweichungen in der energetischen Qualität der Gebäude, das Heizungssystem und möglicherweise durchgeführte energetische Sanierungen. Im Rahmen der Arbeit war es nicht möglich, Informationen über diese Gebäudeparameter zu erhalten. Der Einfluss des Kesselalters beispielsweise hat einen Einfluss von bis zu 10 % auf den spezifischen Verbrauch [28].

6.4. Unsicherheit bei der Wärmeverbrauchsschätzung

Durch die bekannten Gebäudeparameter können maximal 40 % der Varianz der Wärmeverbräuche erklärt werden. Aufgrund des nicht erklärbaren Anteils der Varianz sind Abweichungen zwischen dem geschätzten und tatsächlichen Verbräuchen unvermeidlich. Dies wird durch die Angabe eines Unsicherheitsbereiches bzw. sogenannten Konfidenzintervalls des geschätzten Verbrauchs berücksichtigt. Bei einer Verbrauchsschätzung für eine Gruppe von Gebäuden verkleinert sich der Unsicherheitsbereich, da sich die Abweichungen gegenseitig ausgleichen.

Ziel der Analyse ist es, abhängig von der Größe der Gebäudegruppen und den für die Schätzung verwendeten Gebäudeparametern ein Konfidenzintervall mit einer Vertrauenswahrscheinlichkeit $\alpha = 95\%$ für den geschätzten Verbrauch anzugeben. Dies bedeutet, dass in 95 % der Fälle der tatsächliche Verbrauch innerhalb des geschätzten Konfidenzintervalls liegt.

Bildung zufälliger Gruppen von Gebäuden In räumlich differenzierten Untersuchungen zum Wärmeverbrauch von Gebäuden werden die Ergebnisse nicht für jedes Einzelgebäude angegeben, sondern für mehrere benachbarte Gebäude zusammengefasst. Die dabei entstehenden räumlichen Gebiete bilden eine sogenannte Aggregationsebene. Ein Beispiel für eine Aggregationsebene sind gleich große Rasterzellen. Alle Gebäude innerhalb einer Rasterzelle bilden eine Einheit.

Verallgemeinert ausgedrückt entspricht dies einer Gruppierung von Gebäuden auf Basis ihrer räumlichen Lage. In diesem Fall nimmt aufgrund der festen Gesamtzahl von 17.537 Gebäuden bei zunehmender Anzahl von Gebäuden je Gruppe die Zahl der Gruppen ab. Für Gruppen mit mehr als 100 Gebäuden ist eine statistisch valide Aussage nicht mehr möglich, da zu wenige Gruppen gebildet werden können.

Daher werden aus den 17.537 Datensätzen ohne Berücksichtigung der eigentlichen Lage der Gebäude per Zufallsverfahren Gebäude gezogen und zu Gruppen zusammengefasst. Es werden Gruppengrößen zwischen einem und 1.000 Gebäuden pro Gruppe unterschieden. Für jede Gruppengröße werden 10.000 Gruppen zufällig gezogen. Eine Gruppe entspricht somit einer Stichprobe aus der Grundgesamtheit aller untersuchter Ein- und Zweifamilienhäuser. 10.000 Stichproben pro Gruppengröße sind eine valide statistische Grundlage für die Angabe des Schätzfehlers mit einer Vertrauenswahrscheinlichkeit von 95 %. Dies bedeutet, dass für 9.500 Stichproben jeder Größe der tatsächliche Verbrauch innerhalb des angegebenen Konfidenzintervalls liegt.

Szenarien unterschiedlich feiner Gebäudeklassifizierung Die Konfidenzintervalle sollen nicht nur für verschiedene Gruppengrößen, sondern auch für unterschiedliche Szenarien angegeben werden. Die Szenarien unterscheiden sich aufgrund der berücksichtigten Gebäudeparameter. Sie entsprechen den Modellszenarien bei der multiplen linearen Regression im vorherigen Kapitel.

Tabelle 6.3.: Szenarien für Schätzfehleranalyse

Szenario	Berücksichtigte Parameter	Clusterzahl	Herkunft der Schätzwerte
ONE	- keine -	1	
AGE	Baualtersklassen	9	Mittelwerte der gemessenen Verbräuche
VOL	Volumen	20	
AVOL	Baualtersklassen und Volumen	126	
ALL	Baualtersklassen, A/V-Verhältnis und Volumen	280	
IWU	Baualtersklassen	9	
IWUkorr	Baualtersklassen	9	Durch Messwerte korrigierte Angaben des IWU

Tabelle 6.3 gibt eine Übersicht über die Szenarien. Für jedes Szenario sind die berücksichtigten Parameter und die Zahl der daraus resultierenden Cluster angegeben. Im Szenario ONE werden die Ein- und Zweifamilienhäuser nicht weiter unterschieden. Für alle Gebäude wird ein Schätzwert, der Mittelwert aller gemessenen Verbräuche angenommen. Bei den anderen Szenarien existiert für jedes Cluster ein Schätzwert. In den Szenarien ONE bis ALL werden die Schätzwerte der Untersuchung aus den gemittelten gemessenen Verbräuchen aller Gebäude in der jeweiligen Klasse berechnet. Für die Szenarien IWU und

IWUkorr entstammen die Schätzwerte Literaturangaben bzw. korrigierten Literaturangaben. Formal ausgedrückt besteht das Gebäudecluster k aus der Menge K von Gebäuden mit den zum Cluster passenden Eigenschaften.

Definition des statistischen und systematischen Schätzfehlers Die Differenz bzw. Abweichung zwischen dem Schätzwert eines Clusters und dem tatsächlichen Verbrauch eines Gebäudes in diesem Cluster wird als Schätzfehler bezeichnet. Er kann sowohl positiv als auch negativ sein. Für eine Gruppe von Gebäuden werden die Schätzfehler der Einzelgebäude addiert. Durch die unterschiedlichen Vorzeichen entstehen Ausgleichseffekte.

Der Schätzfehler ΔQ_i für ein Gebäude i ist definiert als die Differenz zwischen dem geschätzten Wärmeverbrauch \widetilde{S}_k und dem tatsächlichen Wärmeverbrauch Q_i eines Gebäudes (siehe Gleichung (6.8)). Der **relative Schätzfehler** Δq_i ist der Quotient aus dem Schätzfehler ΔQ und dem Schätzwert \widetilde{S} (siehe Gleichung (6.9)).

$$\Delta Q_i = Q_i - \widetilde{S}_k \quad \text{mit} \quad i \in K \quad (6.8)$$

$$\Delta q_i = \frac{\Delta Q_i}{\widetilde{S}_k} = \frac{Q_i - \widetilde{S}_k}{\widetilde{S}_k} \quad \text{mit} \quad i \in K \quad (6.9)$$

Der relative Schätzfehler wird auf den Schätzwert bezogen, um die Weiterverwendung der Ergebnisse zu erleichtern. Bei einer Schätzung des Wärmeverbrauchs interessiert es den Anwender, in welchem Intervall um den Schätzwert der tatsächliche Verbrauch liegt. Durch die Definition des relativen Schätzfehlers ist es möglich, dies aus den Ergebnissen dieser Arbeit abzulesen.

Der Schätzfehler wird in den statistischen und systematischen Schätzfehler unterschieden. Der statistische Schätzfehler ΔQ_i^{Stat} ist definiert als die Differenz zwischen dem Mittelwert \bar{Q}_k der tatsächlichen Verbräuche eines Clusters und den einzelnen Verbräuchen Q_i mit $i \in K$. Bei der Berücksichtigung mehrerer Gebäude eines Clusters gleichen sich die Abweichungen der Einzelgebäude vom Klassenmittelwert teilweise aus. Dadurch sinkt der statistische Schätzfehler und geht bei der Berücksichtigung aller Gebäude eines Clusters gegen Null.

Der systematische Schätzfehler wird berechnet aus der Differenz zwischen dem Schätzwert \widetilde{S}_k und dem Mittelwert \bar{Q}_k einer Gebäudeklasse k . Bei der Berücksichtigung mehrerer Gebäude der Klasse k gibt es keine Ausgleichseffekte, der systematische Schätzfehler wächst mit der Zahl der Gebäude pro Gruppe.

Bildung eines Konfidenzintervalls Der Schätzfehler wird für jede Stichprobe berechnet. Für jede Gruppengröße und für jedes Szenario erhält man somit eine Verteilung von 10.000 Werten. Die oben angegebene Vertrauenswahrscheinlichkeit $\alpha = 95\%$ bedeutet, dass die mittleren 95% der Verteilung berücksichtigt werden. Die äußeren 2,5% am unteren bzw. oberen Ende der Verteilung werden nicht berücksichtigt. Der minimale Schätzfehler entspricht somit dem 2,5%-Quantil dieser Verteilung, der maximale dem

97,5 %-Quantil. Beide zusammen ergeben das Konfidenzintervall der Schätzung mit der Vertrauenswahrscheinlichkeit von 95 %.

Die Angaben der nächsten beiden Unterkapitel drücken immer das Konfidenzintervall bzw. dessen Grenzen aus.

6.4.1. Konfidenzintervall des statistischen Schätzfehlers

Der erste Schritt der Analyse konzentriert sich auf den statistischen Schätzfehler, indem als Schätzwert \widehat{S}_k der Mittelwert \bar{Q}_k der jeweiligen Gebäudeklassen verwendet wird. Der systematische Schätzfehler ist Null. Die Berechnung wird für unterschiedliche Gruppengrößen und für die oben genannten Szenarien durchgeführt.

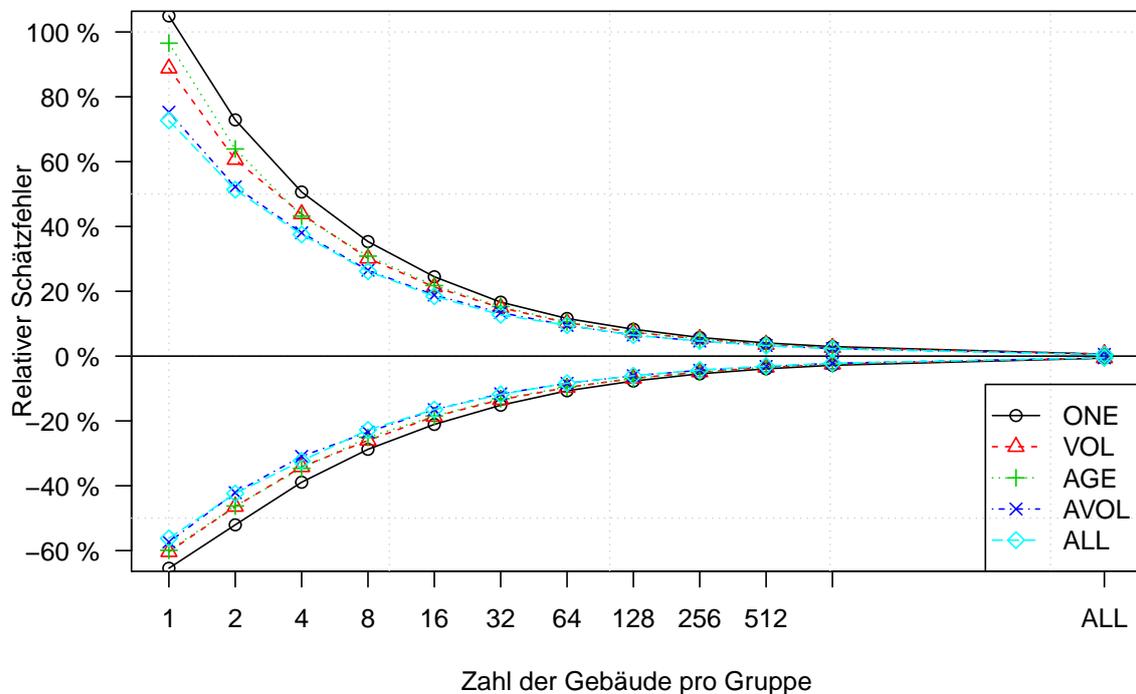


Abbildung 6.7.: 95 %-Quantil des relativen statistischen Schätzfehlers in % bezogen auf den geschätzten Heizwärmeverbrauch der Gruppe mit logarithmischer X-Achse

Die Kurven in Abbildung 6.7 geben die obere und untere Grenze des Konfidenzintervalls an. Die X-Achse ist logarithmisch. „ALL“ bezeichnet eine Kontrollgruppe, in der alle 17.537 Gebäude sind. Dadurch gleichen sich alle Abweichungen vollständig aus und der Schätzfehler bzw. die Breite des Konfidenzintervalls ist Null. Die Y-Achse gibt das Konfidenzintervall des relativen Schätzfehlers in Prozent des Schätzwerts an. Die Angaben sind dadurch von den absoluten Zahlen des Untersuchungsgebiets entkoppelt und können als Anhaltspunkt für die Wärmeverbrauchsschätzung in anderen Gebieten dienen. Eine Angabe von 50 % bzw. -50 % bedeutet, dass mit der Vertrauenswahrscheinlichkeit von 95 % der tatsächliche Wärmeverbrauch maximal 50 % größer bzw. kleiner ist als der

geschätzte. Da der tatsächliche Verbrauch unbegrenzt größer sein kann, aber seine natürliche untere Grenze bei Null hat, liegt das Konfidenzintervall nicht symmetrisch um Null.

Für Einzelgebäude, d. h. einer Gruppengröße von Eins, liegt der statistische Schätzfehler zwischen -60 % und +104 %. Das Szenario ONE zeigt den größten Schätzfehler für alle Gruppengrößen. Dies ist zu erwarten, da in diesem Szenario keine Informationen über die Gebäudeparameter berücksichtigt werden. Im Gegensatz dazu zeigt das Szenario ALL den geringsten Schätzfehler. Diese Tatsache kann in Zusammenhang mit der multiplen linearen Regression des vorherigen Kapitels gesetzt werden. Im Szenario ALL werden alle verfügbaren Informationen über die Gebäude für die Schätzung verwendet. 40 % der Heizverbrauchs-Varianz kann durch die Gebäudeparameter erklärt werden (siehe Tabelle 6.2). Da der statistische Schätzfehler vom Anteil der unerklärbaren Varianz abhängt, ist er in diesem Szenario am geringsten.

Die Verringerung des statistischen Schätzfehlers bei einer Verdopplung der Gebäude pro Gruppe für das Szenario ONE ist ähnlich groß wie der Unterschied zwischen den Szenarien ONE und ALL bei konstanter Gruppengröße. Das heißt, wenn keine Gebäudeparameter außer der Gebäudeklasse EFH/ZFH für die Schätzung zur Verfügung stehen, kann der Schätzfehler durch die Verdopplung der Gruppengröße auf ein ähnliches Niveau gesenkt werden wie bei voller Informationsverfügbarkeit. Dies liegt daran, dass durch den hohen nicht erklärbaren Anteil der Varianz der Heizwärmeverbräuche die Ausgleichseffekte bei der gemeinsamen Schätzung für mehrere Gebäude stärker sind als die Verkleinerung der nicht erklärbaren Varianz durch die Gebäudeparameter.

Für Gruppen von 10 Gebäuden variiert der tatsächliche Verbrauch nur noch zwischen -22 % und +26 % des geschätzten Verbrauchs, für Gruppen um die 100 Gebäude nur noch um $\pm 7\%$.

Mit wachsender Gruppengröße verkleinert sich auch die Differenz der spezifischen Schätzfehler unterschiedlicher Szenarien. Bei Einzelgebäuden wird durch die Verwendung von Volumen und Baualtersklasse (Szenario AVOL) der Schätzfehler um 23 Prozentpunkte gegenüber dem Szenario ONE verkleinert. Für Gruppen von 10 Gebäuden liegt der Unterschied nur noch bei 7 Prozentpunkten. Je größer die Gebäudegruppen für die Wärmeverbrauchsschätzung gewählt werden, desto weniger wird der Schätzfehler durch die Berücksichtigung zusätzlicher Gebäudeparameter bei der Schätzung verbessert.

Diskussion der Ergebnisse Für die Schätzung von Einzelgebäuden liegen die Konfidenzintervall-Grenzen des statistischen Schätzfehlers im genauesten Szenario außerhalb von -58 % bzw. +75 % des Schätzwerts. Die Aussagekraft eines Schätzwertes muss daher hinterfragt werden. Die vorgestellten Ergebnisse basieren auf Ein- und Zweifamilienhäuser mit rein privater Nutzung. Das bedeutet, die Grundgesamtheit der Daten ist bereits sehr homogen ausgewählt. Ausgleichseffekte bei der Gruppierung von Gebäuden führen dazu, dass die Varianz der Gruppen und damit der statistische Schätzfehler bereits für Gruppen von zwei bis fünf Gebäuden sehr stark zurück geht. Wenn man die Gruppen als räumliche Aggregation deutet, führt somit eine nur geringfügig geringere räumliche Auflösung zu deutlich präziseren Aussagen bezüglich des Wärmeverbrauchs.

Der sich sehr schnell verringernde Abstand zwischen den Szenarien unterschiedlich berücksichtigter Gebäudeparameter führt bei großräumigen Wärmeverbrauchsschätzungen dazu, dass die zum Teil aufwendig ermittelten Gebäudeparameter nur einen geringen Unterschied in der Genauigkeit der Aussage erzeugen.

6.4.2. Konfidenzintervall des statistischen und systematischen Schätzfehlers

Normalerweise sind für ein Untersuchungsgebiet keine Wärmeverbrauchsdaten bekannt, weshalb externe Schätzwerte aus der Literatur verwendet werden müssen. Diese zeigen meist eine Abweichung zu den Mittelwerten der tatsächlichen, jedoch unbekanntem Verbräuche. Ein Beispiel für eine derartige Abweichung ist der Vergleich der spezifischen Verbräuche in dieser Arbeit und der Angaben des IWU in Abbildung 6.4. Für die kombinierte Analyse des statistischen und systematischen Schätzfehlers werden diese IWU-Angaben herangezogen.

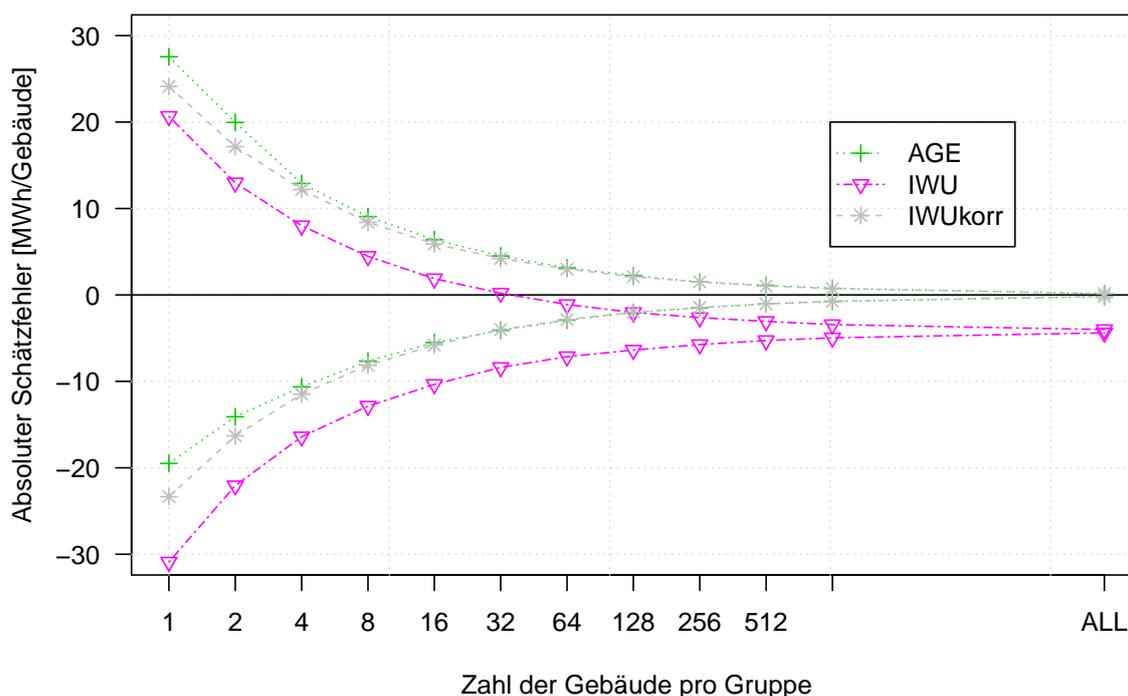


Abbildung 6.8.: 95 %-Quantile des absoluten Schätzfehlers in MWh/Gebäude mit logarithmischer X-Achse

Abbildung 6.8 zeigt für jede untersuchte Gruppengröße den absoluten Schätzfehler pro Gebäude, d. h. die Differenz zwischen Verbrauch und Schätzung (siehe Gleichung (6.8)). Die Verwendung des relativen Schätzfehlers würde einen Vergleich der unterschiedlichen Szenarien verzerren. Das Konfidenzintervall bzw. der Vertrauensbereich der Aussage liegt wiederum zwischen 2,5 % und 97,5 %. Verglichen werden die auf den IWU basierende Schätzung des Szenarios IWU mit dem bereits oben untersuchten Szenario AGE. Für die

Gruppengröße „ALL“ gibt die Abweichung des Szenarios IWU von Null den systematischen Schätzfehler pro Gebäude an. Dieser beträgt $-4,9 \text{ MWh/a}$ bzw. $-12,2 \%$ des mittleren tatsächlichen Verbrauchs. Die Abweichung ist negativ, da die Schätzwerte des IWUs zu groß sind. Der durchschnittliche Verbrauch eines Ein- bzw. Zweifamilienhauses im Untersuchungsgebiet ist ca. 26 MWh/a . Ein Schätzfehler von knapp -30 MWh/a pro Gebäude ergibt eine relative Abweichung von über 100% .

Eine Anpassung der Literatur-Schätzwerte an die realen Bedingungen verkleinert den systematischen Schätzfehler. Für das Szenario IWU_{korr} werden die IWU-Schätzwerte um den oben berechneten systematischen Schätzfehler $12,2 \%$ verkleinert. Die Anpassung an reale Bedingungen benötigt keine gemessenen einzelne Verbräuche, sondern kann mit aggregierten Liefermengen erfolgen. Diese sind für Untersuchungsgebiete meist leichter zu erhalten. Zum Teil werden in Gebieten mit leitungsgebundener Versorgung die aggregierten Liefermengen vom Energieversorger veröffentlicht.

Bei Szenario IWU_{korr} in Abbildung 6.4 werden die neun Schätzwerte der neun verwendeten Klassen mit nur einer aggregierten Verbrauchsmenge so korrigiert, dass die Abweichung in Summe Null ist. Dabei ist sehr unwahrscheinlich, dass alle neun Schätzwerte individuell richtig korrigiert werden. Das heißt, die einzelnen Schätzwerte sind weiterhin zu hoch oder zu niedrig. Nichtsdestotrotz ist davon auszugehen, dass die Korrektur für kleinere Gruppen von Gebäuden realistischere Ergebnisse liefert als nicht korrigierte Werte. Dies zeigt sich in den Konfidenzintervall-Grenzen des Szenarios IWU_{korr} für Gruppen mit mehr als vier Gebäuden. Sie liegen sehr nah an den Grenzen des Szenarios AGE.

Diskussion der Ergebnisse Ein Vergleich der Konfidenzintervallgrenzen in Abbildung 6.8 zeigt, dass durch die Korrektur der Schätzwerte die Konfidenzintervalle in Richtung der Nulllinie des Schätzfehlers verschoben werden. Bei der Verwendung mehrerer Klassen bzw. Schätzwerte sind statistische und systematische Schätzfehler voneinander abhängig. Das bedeutet, dass eine Verschiebung der Konfidenzintervalle die Breite derselben beeinflusst.

Für die Korrektur der Schätzwerte sind unterschiedliche Verfahren vorstellbar. Möglich ist die Verwendung aggregierter veröffentlichter Liefermengen. Alternativ könnten die Schätzwerte mit Stichproben gebäudescharfer Verbräuche korrigiert werden. Hierbei ist jedoch wieder auf eine Stichprobe in ausreichender Größe zu achten, um repräsentative Ergebnisse für Aussagen über den Gebäudebestand sicherzustellen.

6.4.3. Übertragbarkeit der Ergebnisse auf Mehrfamilienhäuser

Die Untersuchung der Schätzfehler für Mehrfamilienhäuser wird in dieser Arbeit aufgrund der eingeschränkten Datenlage nicht durchgeführt. Mehrfamilienhäuser beinhalten mehr Wohneinheiten als Ein- und Zweifamilienhäuser. Unterschiedliche Nutzerverhalten, die zu einem bedeutenden Teil für die Varianz des Wärmeverbrauchs verantwortlich gemacht werden können, gleichen sich in Mehrfamilienhäusern innerhalb des Gebäudes zum Teil

aus. Das bedeutet, dass die Varianz der Wärmeverbräuche von Mehrfamilienhäusern kleiner ist als die von Ein- und Zweifamilienhäusern.

Dadurch ist anzunehmen, dass die Konfidenzintervalle bei der Schätzung des Wärmeverbrauchs ebenfalls kleiner sind. Informationen zur Verteilung spezifischer Wärmeverbräuche in Mehrfamilienhäusern bieten die Publikationen [69] und [33].

6.5. Vergleich von Wärmeverbrauch und Wärmebedarf

Der Wärmebedarf ist eine Größe, die die Qualität der thermischen Gebäudehülle bewertet [75]. Für die Berechnung wird als Randbedingung ein standardisiertes Nutzerverhalten angenommen. Bei der Bewertung von Einzelgebäuden, beispielsweise im Rahmen des Energiebedarfsausweises sind die Randbedingungen durch den Gesetzgeber festgelegt, um die Vergleichbarkeit der Ergebnisse zu ermöglichen [17]. Für kommunale Energiebedarfsrechnungen existieren keine verbindlichen Vorgaben.

Der Wärmeverbrauch ist die tatsächlich verbrauchte und zumindest theoretisch messbare Menge an Wärme. Er ist damit die Grundlage für die Berechnung von Klimagasemissionen oder der tatsächlichen Einspareffekte der Gebäudesanierung.

6.5.1. Wärmebedarfsberechnung und Literaturvergleich

Die Berechnung des Wärmebedarfs in dieser Arbeit erfolgt anhand des Heizperiodenbilanzverfahrens. Annahmen und Randbedingungen sind in Kapitel 2.4 beschrieben. Das Verfahren ist als Algorithmus in der Datenbank implementiert. Zur Überprüfung der Ergebnisse werden die Wärmebedarfswerte des IWU herangezogen [53]. Diese beinhalten für jeden Gebäudetyp einen Wert. Abbildung 6.9 vergleicht die im Untersuchungsgebiet Oldenburg berechneten und gemittelten Raumwärmebedarfe mit den Angaben des IWU. Die Werte des IWU basieren auf nur einem Gebäude, das den Gebäudetyp repräsentiert. Die Buchstaben neben den Datenpunkten bezeichnen die Baualtersklasse.

Die Abweichung der Datenpunkte von der Winkelhalbierenden zeigt den Unterschied zwischen beiden Angaben. Dieser lässt sich vor allem auf die unterschiedlichen Gebäudehüllen zurückführen. Die Gebäude in Oldenburg weisen ein höheres A/V-Verhältnis auf als die IWU-Typgebäude, wie Abbildung 5.12 auf Seite 83 zeigt. Die Geraden mit der prozentualen Abweichung beziehen sich auf den Raumwärmebedarf in Oldenburg. Die Ergebnisse zeigen jedoch größtenteils eine maximale Abweichung von 25 % und sind damit eine akzeptable Grundlage für den Vergleich zwischen Wärmebedarf und Wärmeverbrauch.

6.5.2. Vergleich von Wärmeverbrauch und -bedarf

Für den Vergleich von Wärmeverbrauch und Wärmebedarf der Ein- und Zweifamilienhäuser in Oldenburg werden die Verbrauchs- bzw. Bedarfswerte je Baualtersklassen und

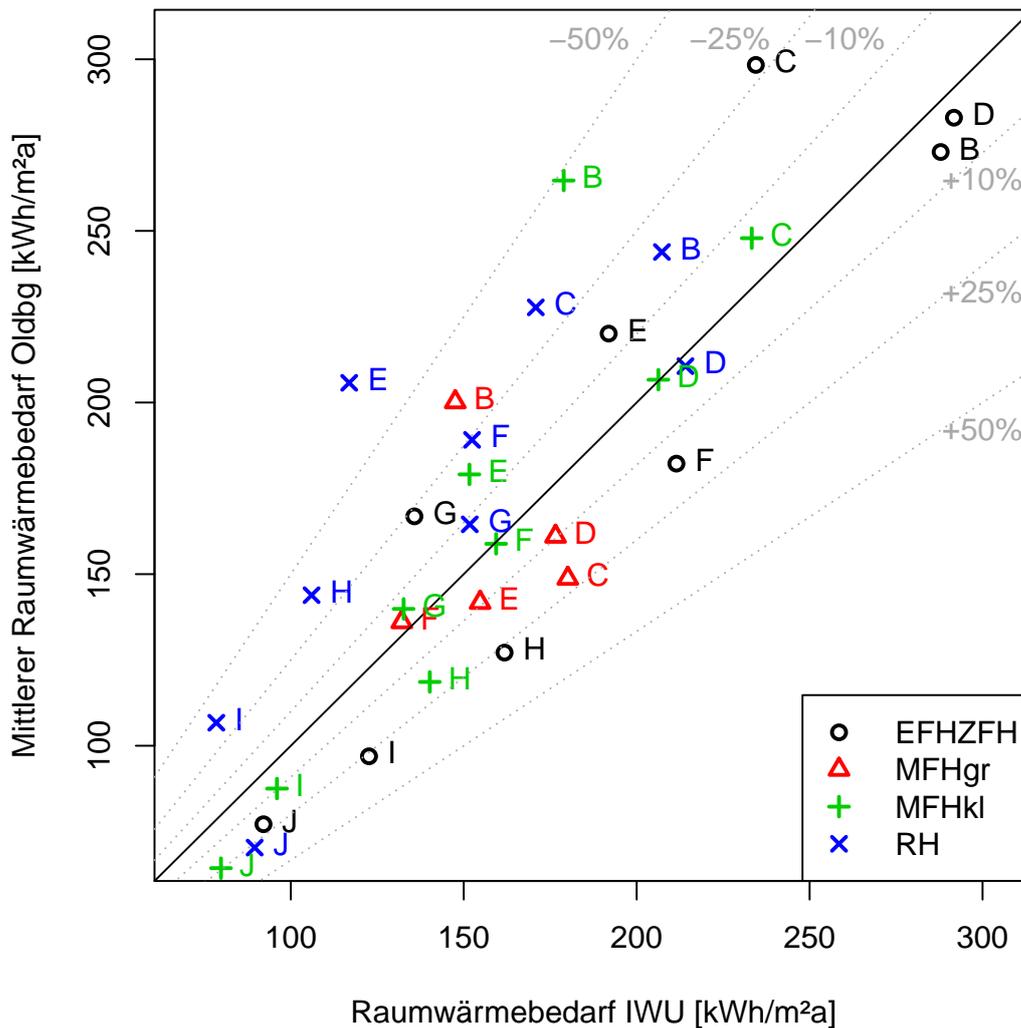


Abbildung 6.9.: Mittlerer Raumwärmebedarf je Gebäudetyp in Oldenburg und Literaturvergleich [53]

Volumenklassen (Abbildung 6.10) gemittelt. Es sind jeweils die spezifischen Verbräuche pro m^2 Gebäudenutzfläche abgebildet. Für die Umrechnung zwischen dem Raumwärmebedarf und der Endenergie wird der Anlagenfaktor mit 1,1 und der Verteilungsfaktor mit 1,16 angenommen [51]. Die Endenergieverbräuche sind klimabereinigt (siehe Kapitel 2.4), die Bedarfe sind mit der klimaangepassten Gradtagszahl von 81,192 kWh/a gerechnet.

In Abbildung 6.10 fällt auf, dass die Abweichungen der Datenpunkte von der 1:1-Linie, der Winkelhalbierenden, hauptsächlich von der Baualtersklasse der Gebäude abhängen. Für relativ junge Gebäude der Klassen H, I und J stehen Verbrauch und Bedarf in einem konstanten Verhältnis, wobei die Bedarfswerte ca. 40 % höher sind als die Verbrauchswerte. Bei älteren Gebäuden erhöht sich der Bedarf aufgrund der schlechter angenommenen U-Werte (Wärmeleitwerte) [53]. Insbesondere für Gebäude mit großen Volumina hingegen ist der spezifische Verbrauch nahezu gleich für alle Baualtersklassen.

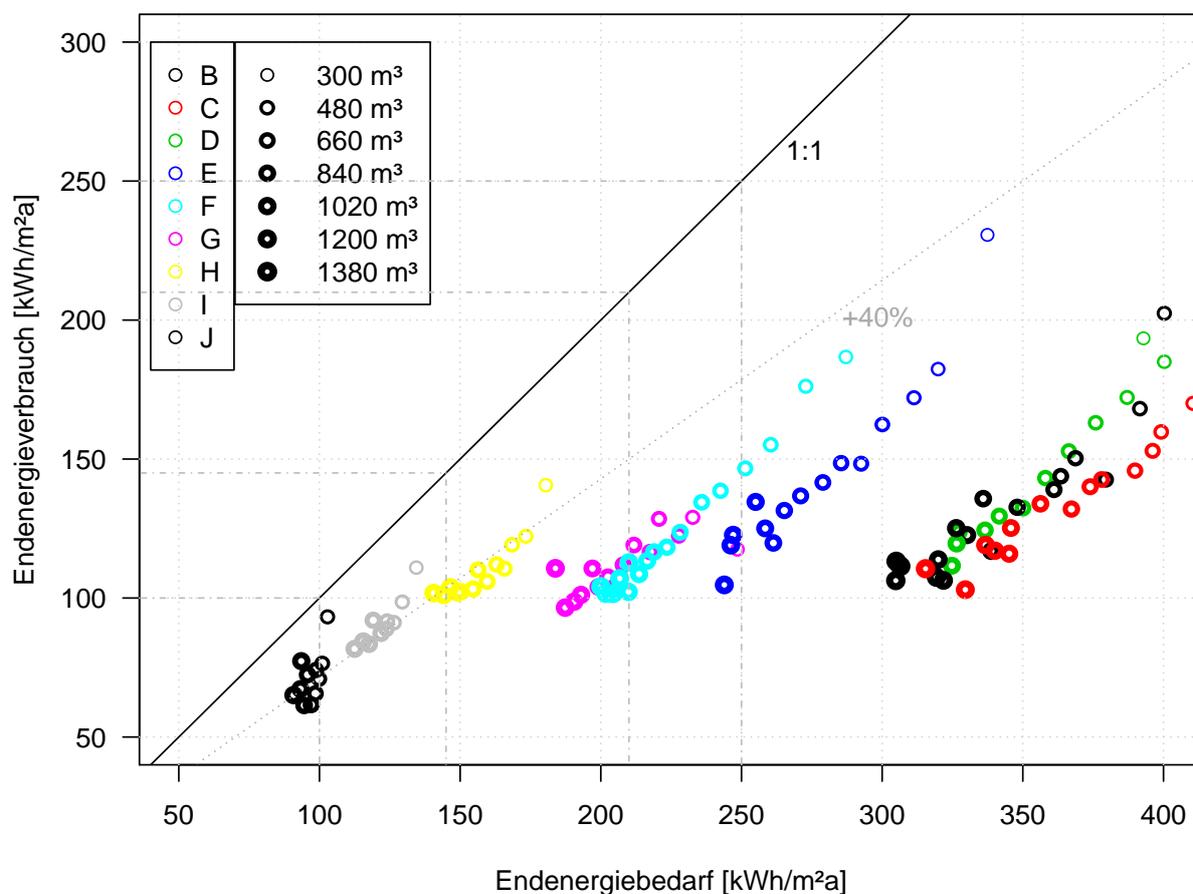


Abbildung 6.10.: Spezifischer Endenergieverbrauch und -bedarf je Baualters- und Volumenklasse für Ein- und Zweifamilienhäuser

Eine Begründung für die altersabhängige Abweichung von Bedarf und Verbrauch läge darin, dass die für die Berechnung verwendeten U-Werte der deutschen Gebäudetypologie nicht auf die Gebäude in Oldenburg zutreffen. Da jedoch die Wärmebedarfsberechnungen des IWU und in Oldenburg (siehe obige Abbildung 6.9) weitaus geringere Abweichungen aufweisen, ist dies unwahrscheinlich.

Für die Abweichung kommt vor allem ein baualterabhängiges Nutzerverhalten in Frage. Darauf lassen auch die Daten des IWU schließen [75, 53]. Dies wirkt sich durch verschiedene Faktoren aus. Insbesondere in älteren Gebäuden wird wahrscheinlich nicht gesamte Wohnfläche beheizt. Das bedeutet, dass die Bezugsfläche bei der Berechnung der spezifischen Bedarfe zu groß ist. Im standardisierten Nutzerverhalten wird angenommen, dass die mittlere Innenraumtemperatur 20°C beträgt, was bei alten Gebäuden oftmals nicht der Fall ist. Der Wärmebedarf ist jedoch von der Innenraumtemperatur abhängig. Hinzu kommt, dass Abwesenheitszeiten bei der Bedarfsberechnung nicht berücksichtigt, sondern während der gesamten Heizperiode von einer konstanten Innenraumtemperatur ausgegangen wird. Es ist davon auszugehen, dass alle drei Effekte zu dem Unterschied zwischen Bedarf und Verbrauch beitragen. Da keine Informationen über die Nutzer des Gebäudes vorliegen, kann dies im Rahmen dieser Arbeit nicht weiter untersucht werden.

Da die Mehrzahl der Gebäude in Oldenburg nicht umfassend saniert wurde, zeigt die Analyse sehr gut, wie das Nutzerverhalten von der Baualtersklasse und damit von der thermischen Qualität des Gebäudes abhängt.

Die Abhängigkeit des Energiebedarfs als auch des Energieverbrauchs vom Volumen der Gebäude für eine bestimmte Baualtersklasse ist ähnlich, insbesondere für ältere Gebäude der Baualtersklassen B bis H (-1994). Dies wird deutlich aus den parallel zur 1:1-Achse liegenden Datenpunkten je Baualtersklasse. Wie zu erwarten steigt der spezifische Bedarf und spezifische Verbrauch mit kleiner werdenden Volumina.

6.5.3. Effekt von Energieeffizienzmaßnahmen

Zu Energieeffizienzmaßnahmen gehören die energetische Sanierung der Gebäudehülle und der Einbau effizienterer Heiztechnologien. Dies führt bei gleich bleibendem Nutzerverhalten zu einer Verringerung der für Heizwärme verbrauchten Endenergie. In der öffentlichen Diskussion wird die Einsparung oftmals auf Basis der Wärmebedarfe quantifiziert, was zu überzogenen Erwartungen an die Gebäudesanierung führt. Tatsächlich jedoch müsste die Einsparung auf Basis der Wärmeverbräuche genannt werden.

Als Beispiel wird im Folgenden vereinfacht angenommen, dass Gebäude der Baualtersklasse B, C oder D durch Sanierungsmaßnahmen die energetische Qualität von Gebäuden der Baualtersklasse I erhalten. Der Unterschied von Gebäuden vor und nach der Sanierung entspricht somit dem Unterschied zwischen den Datenpunkten der Baualtersklasse I und der Baualtersklassen B-D in Abbildung 6.10. Der horizontale Abstand gibt die Differenz der Endenergiebedarfe, der vertikale Abstand die Differenz der Endenergieverbräuche an.

Für Gebäude mit einem Volumen über 1.000 m^3 erhält man eine Bedarfs-Differenz von ca. $200 \text{ kWh/m}^2\text{a}$. Der Bedarf der Baualtersklassen B-D beträgt ca. $320 \text{ kWh/m}^2\text{a}$, der der Baualtersklasse I ca. $120 \text{ kWh/m}^2\text{a}$. Die Verbrauchs-Differenz beträgt dagegen ca. $20 \text{ kWh/m}^2\text{a}$, das sind nur 10 % der Bedarfs-Differenz.

Die Bedarfsdifferenz gibt eine theoretische Bedarfsverbesserung an. Sie ist eine Kennzahl für den Vergleich von Sanierungsmaßnahmen verschiedener Gebäude. Trotz ihrer Einheit „ $\text{kWh/m}^2\text{a}$ “ darf sie jedoch nicht mit der Verringerung des realen Wärmeverbrauchs verwechselt werden. Die Verbrauchs-Differenz ist deutlich geringer, da einerseits der Energieverbrauch im unsanierten Zustand der Gebäude deutlich kleiner ist als der Bedarf [75]. Andererseits bewirkt die energetische Sanierung nicht nur eine reine Effizienzsteigerung bei gleichem Nutzerverhalten, sondern steigert ebenfalls den Wohnkomfort. Dies jedoch wirkt der Effizienzsteigerung entgegen und wird auch als *Rebound Effect* bezeichnet.

7. Zusammenfassung und Ausblick

Der Wärmesektor hat eine herausragende Bedeutung für das Gelingen der Energiewende in Deutschland. 31 % des gesamten Endenergieverbrauchs können auf die Erzeugung von Raumwärme zurück geführt werden. Die Komplexität der notwendigen Veränderungen und die hohen damit verbundenen Investitionskosten machen eine Energiesystemanalyse notwendig. Grundlage dieser Analyse sind Informationen über die Verbraucher und Erzeuger sowie deren räumliche Lage. Eine große Herausforderung stellt hierbei die Datenerhebung im Gebäudebestand dar.

Die kommunale Verwaltung ist die administrative Ebene mit den meisten Informationen zum Gebäudebestand im jeweiligen Gebiet. Deren Nutzung ist jedoch erschwert, da sich die Daten auf unterschiedliche räumliche Skalen beziehen, Angaben als beschreibende Elemente in natürlicher Sprache enthalten, nicht vollständig bzw. flächendeckend verfügbar sind und sich zum Teil widersprechen.

Datenverarbeitungskonzept

Das im Rahmen dieser Arbeit entwickelte Konzept ermöglicht die Zusammenführung und Berechnung von Gebäudeparametern hinsichtlich der Analyse der Wärmeverbräuche und -bedarfe (siehe Abbildung 7.1). Die verwendeten Daten stammen aus Quellen unterschiedlicher Vollständigkeit, Datenqualität, geografischer Skalen und Datenformate. Sie werden beim Import für eine automatisierte Verarbeitung aufbereitet und erhalten einen eindeutigen Raumbezug. Angaben aus Elementen in natürlicher Sprache werden durch *Text-Mining*-Methoden extrahiert. Die Methoden umfassen den schreibweisereisistenten Vergleich von Adressangaben und die Erkennung von Gebäudeinformationen durch Mustererkennungsverfahren.

Eine *Open Source*-Geodatenbank speichert alle Informationen und die zugehörigen Raumbezüge. Das dafür entworfene Datenmodell ist so konzipiert, dass neue Gebäudeparameter, Datenquellen und Raumbezüge ohne Änderung des Modells hinzugefügt werden können. Auf Basis miteinander verbundener Raumbezüge (z. B. Adresse und Gebäude) werden die Informationen einem oder mehreren Gebäuden zugeordnet. Konflikte aufgrund von widersprüchlichen Angaben für ein Gebäude werden anhand der Qualität der Quellen, der Aktualität der Daten und der geografischen Auflösung automatisiert gelöst.

Es sind Methoden implementiert, die Gebäudeparameter wie Gebäudegröße, Gebäudetyp, Gebäudenutzung und die Fassadenflächen unter Berücksichtigung angrenzender Gebäude berechnen. Für jedes Gebäude steht somit ein für die Untersuchung des Wärmeverbrauchs bzw. -bedarfs relevanter Datenbestand - vorstellbar als Gebäudesteckbrief - für Analysen zur Verfügung.

7. Zusammenfassung und Ausblick

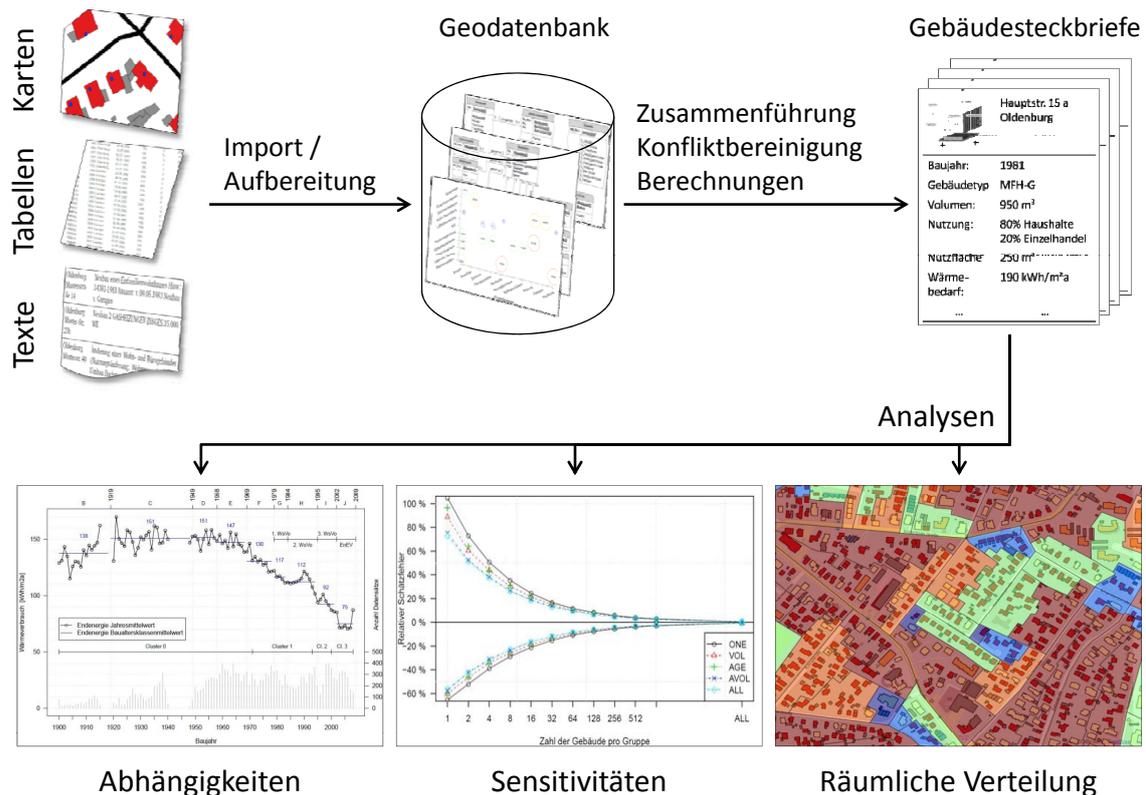


Abbildung 7.1.: Gebäudedatenverarbeitung und Typen der Analyse von Gebäude- und Verbrauchsdaten

Die Datenmodelle und Routinen gewährleisten, dass nach der Konfliktbereinigung und Berechnung für jede Angabe der Gebäudesteckbriefe die Datenquelle bzw. Berechnungsmethode gespeichert und nachvollziehbar ist. Das hilft bei der kritischen Auseinandersetzung mit den darauf aufbauenden Analysen und Schlussfolgerungen. Die automatisierte Konfliktbereinigung ermöglicht beispielsweise, dass Standardwerte beim Vorhandensein genauerer Informationen für einzelne Gebäude ersetzt werden. Ein mit der Zeit wachsender Datenbestand, z. B. aufgrund einer fortlaufenden Datenrecherche, kann ohne manuelle Bearbeitung nachträglich in die Gebäudesteckbriefe integriert werden.

Analyse von Gebäude- und Wärmeverbrauchsdaten

Auf Basis der zusammengeführten und berechneten Gebäudeparameter für das Untersuchungsgebiet Oldenburg umfasst die Analyse Gebäudeparameter und Wärmeverbräuche für ca. 17.500 Ein- und Zweifamilienhäuser. Sie berücksichtigt die für die Wärmebedarfsbilanzierung wichtigen Parameter Alter der Gebäude, Volumen sowie Oberfläche/Volumen-Verhältnis. Der generelle Zusammenhang zwischen Gebäudeparametern und Verbräuchen wird anhand gemittelter Wärmeverbräuche von Gebäuden ähnlicher Parameterkonfiguration bestimmt. Lineare multiple Regressionsanalysen zeigen, inwieweit die Streuung der Wärmeverbräuche durch die Gebäudeparameter erklärbar ist. Davon ausgehend werden

Konfidenzintervalle für eine Vorhersage von Wärmeverbräuchen auf Grundlage der Gebäudeparameter berechnet. Es findet eine Unterscheidung für Szenarien unterschiedlicher Parameterverfügbarkeit und für verschieden große Gruppen von Gebäuden statt.

Der Verlauf der mittleren Wärmeverbräuche in Abhängigkeit vom Baujahr der Gebäude seit dem Jahr 1900 zeigt die Antizipation der gesetzlichen Wärmeschutzverordnungen um jeweils einige Jahre. Die Ergebnisse belegen, dass für die Analyse der Verbräuche vier statt neun Baualtersklassen ausreichen. Die Baualtersklassen vor 1984 können zu einer Klasse zusammengefasst werden.

Der Vergleich theoretisch berechneter Wärmebedarfe und tatsächlich gemessener Wärmeverbräuche zeigt ein vom Gebäudebaujahr abhängiges Nutzerverhalten. Bei älteren Gebäuden übersteigt der Wärmebedarf den -verbrauch um 100 % bis 200 %; bei jüngeren um ca. 40 %. Diese Tatsache sollte bei der Berechnung von Sanierungspotentialen deutlicher Berücksichtigung finden.

Auf Basis der im Untersuchungsgebiet Oldenburg bekannten Gebäudeparameter sind 40 % der Varianz der Wärmeverbräuche erklärbar. 60 % der Varianz unterliegt dem Einfluss nicht ermittelter Parameter und dem individuellen Nutzerverhalten. Im Umkehrschluss bedeutet dies eine hohe Unsicherheit bei der Vorhersage von Wärmeverbräuchen. Das 95 %-Konfidenzintervall bei der Schätzung des Verbrauchs für einzelne Ein- und Zweifamilienhäuser reicht von -60 % bis +100 % des mittleren Verbrauchs. Bei einer Aggregation der Schätzwerte über mehrere Gebäude halbiert sich die Breite des Konfidenzintervalls mit Vervierfachung der Gruppengröße. Zugleich nimmt der Einfluss zusätzlicher Gebäudeparameter auf die Intervallbreite ab. Die ermittelten Gebäudeparameter Volumen und Baualter führen ab 30 aggregierten Gebäuden zu keiner signifikanten Verkleinerung der Konfidenzintervalle.

Ausblick

Das Datenverarbeitungskonzept könnte für eine periodisch fortgeschriebene Gebäudebestandsbilanzierung weiter entwickelt werden. Die dafür notwendige Modellierung von Lebenszyklen benötigt Gebäudedatensätze und -bauteile mit einem zusätzlichen Zeitparameter. Darüber hinaus wäre es insbesondere für die Prognose von Wärmeverbräuchen wünschenswert, weitere, in dieser Arbeit nicht verfügbare Einflussgrößen auf den Verbrauch zu identifizieren. Hierzu gehören beispielsweise die Art der Feuerungsanlagen und sozioökonomische Angaben zu den Gebäudebewohnern. Eine aufgrund der Datenlagen nicht durchführbare Verbrauchsdaten-Analyse von Mehrfamilienhäusern würde die vorgestellten Ergebnisse ergänzen und eine flächendeckende Fehlerabschätzung bei Wärmekatastern im Sektor Haushalte ermöglichen.

Das Datenerfassungskonzept ist ein Werkzeug, um Gebäudedaten aus verschiedenen Quellen zusammenzuführen und für einen Analyse aufzubereiten. Die Nachvollziehbarkeit der Herkunft der Informationen soll helfen, die erhaltenen Ergebnisse als Grundlage gesellschaftlicher und politischer Diskussionen zu etablieren.

A. Anhang

A.1. Räumliche Darstellung der ermittelten Gebäudebaujahre in Oldenburg

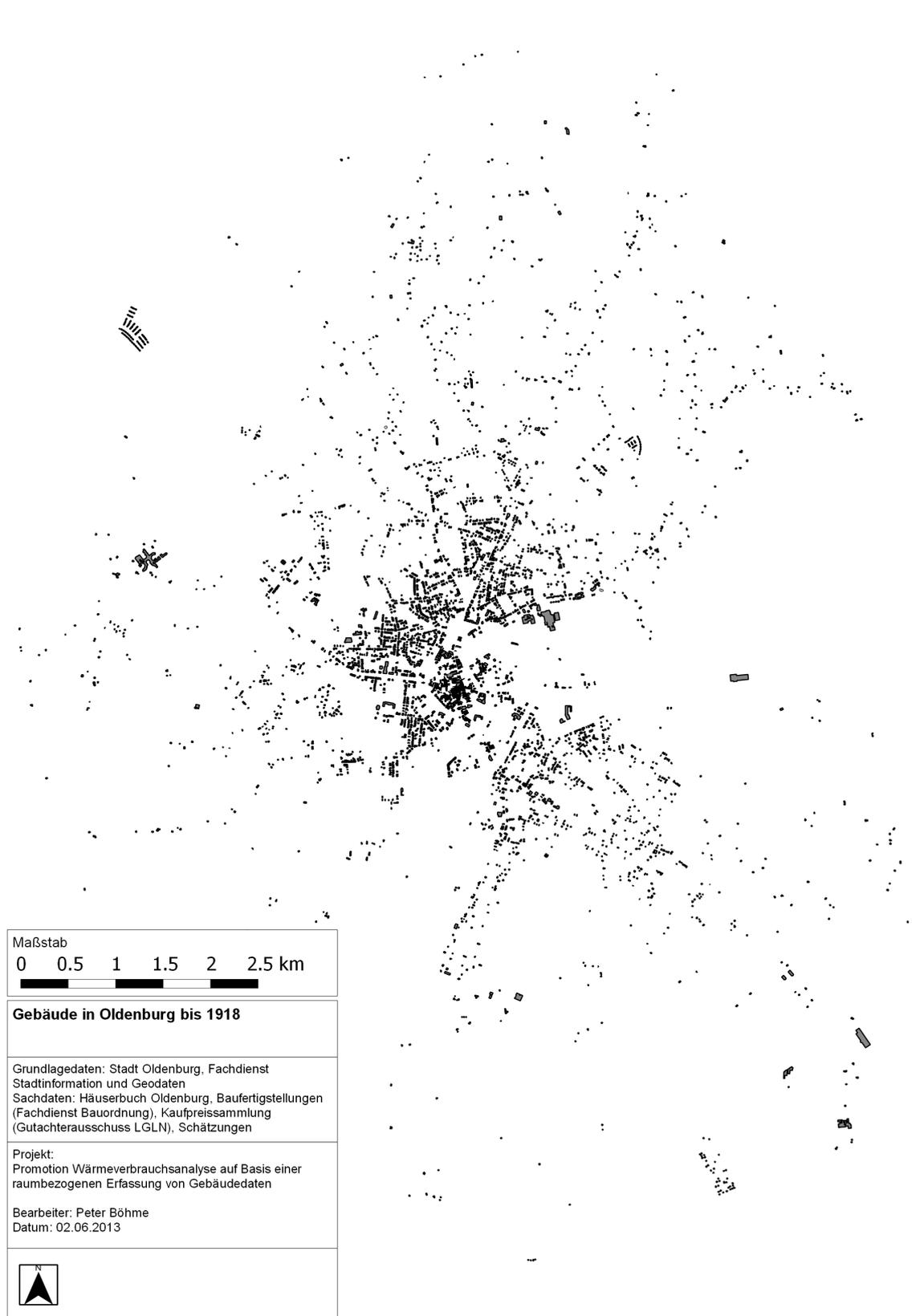


Abbildung A.1.: Karte der Gebäude mit Baujahr bis 1918

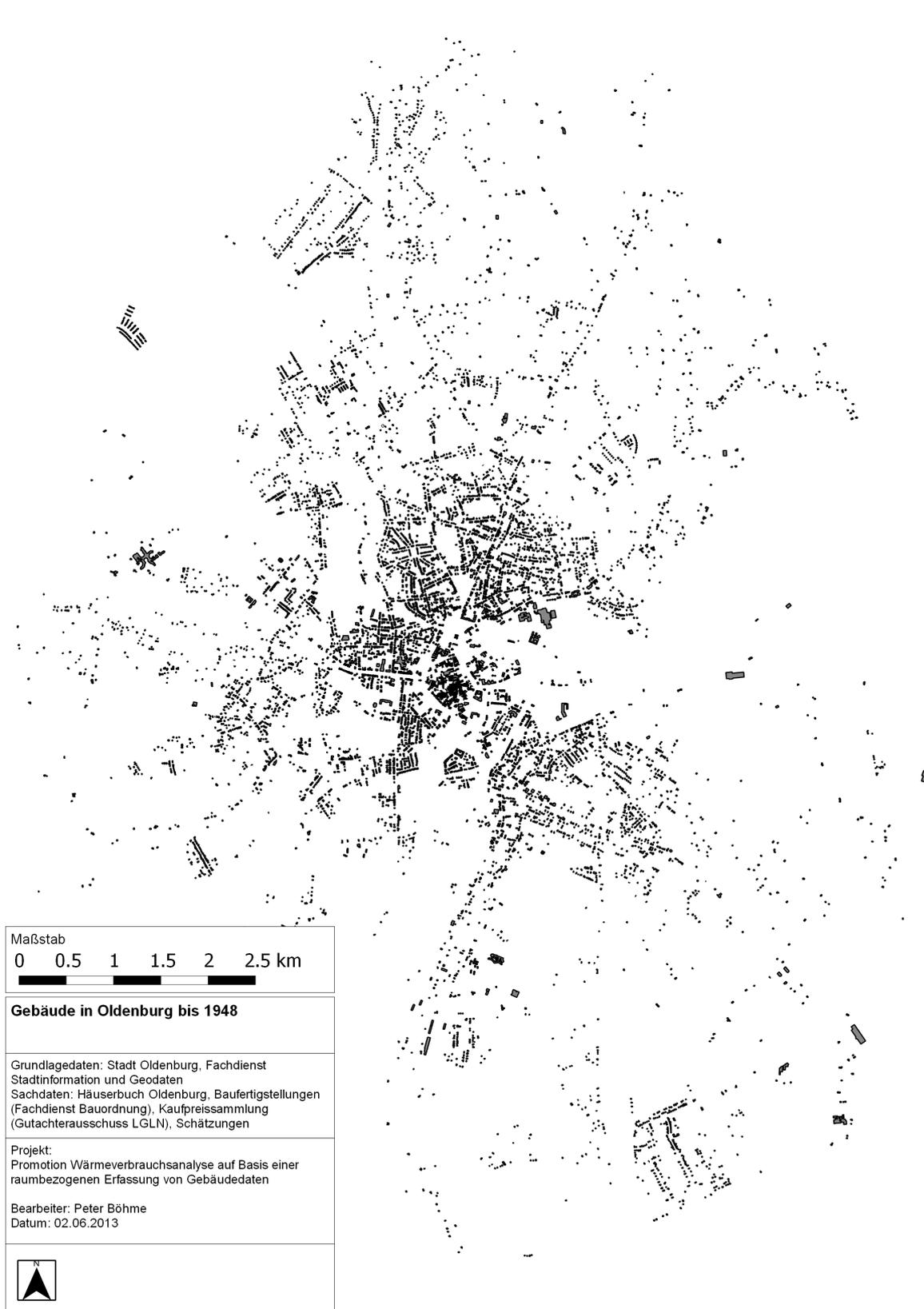


Abbildung A.2.: Karte der Gebäude mit Baujahr bis 1948

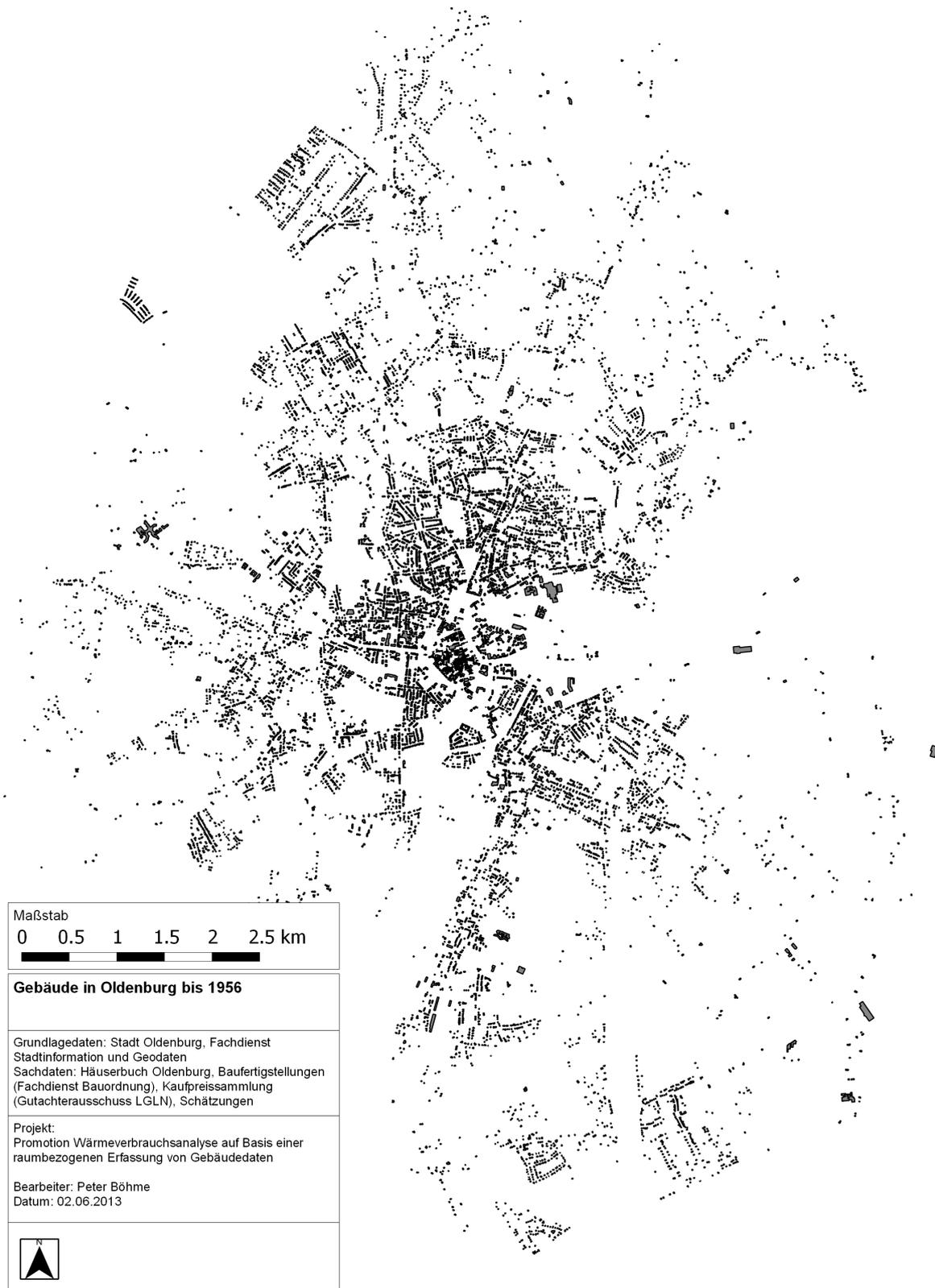


Abbildung A.3.: Karte der Gebäude mit Baujahr bis 1956

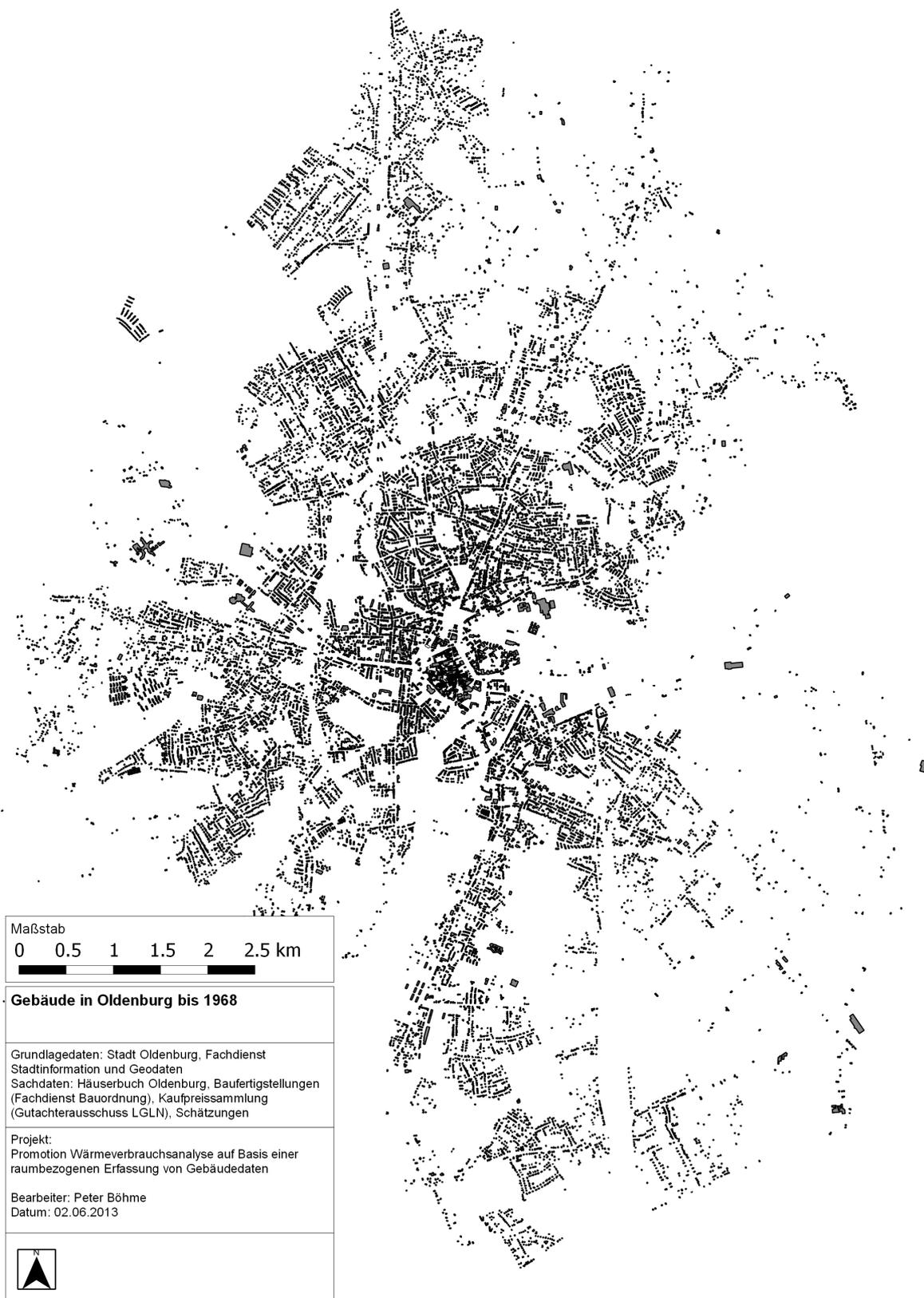


Abbildung A.4.: Karte der Gebäude mit Baujahr bis 1968

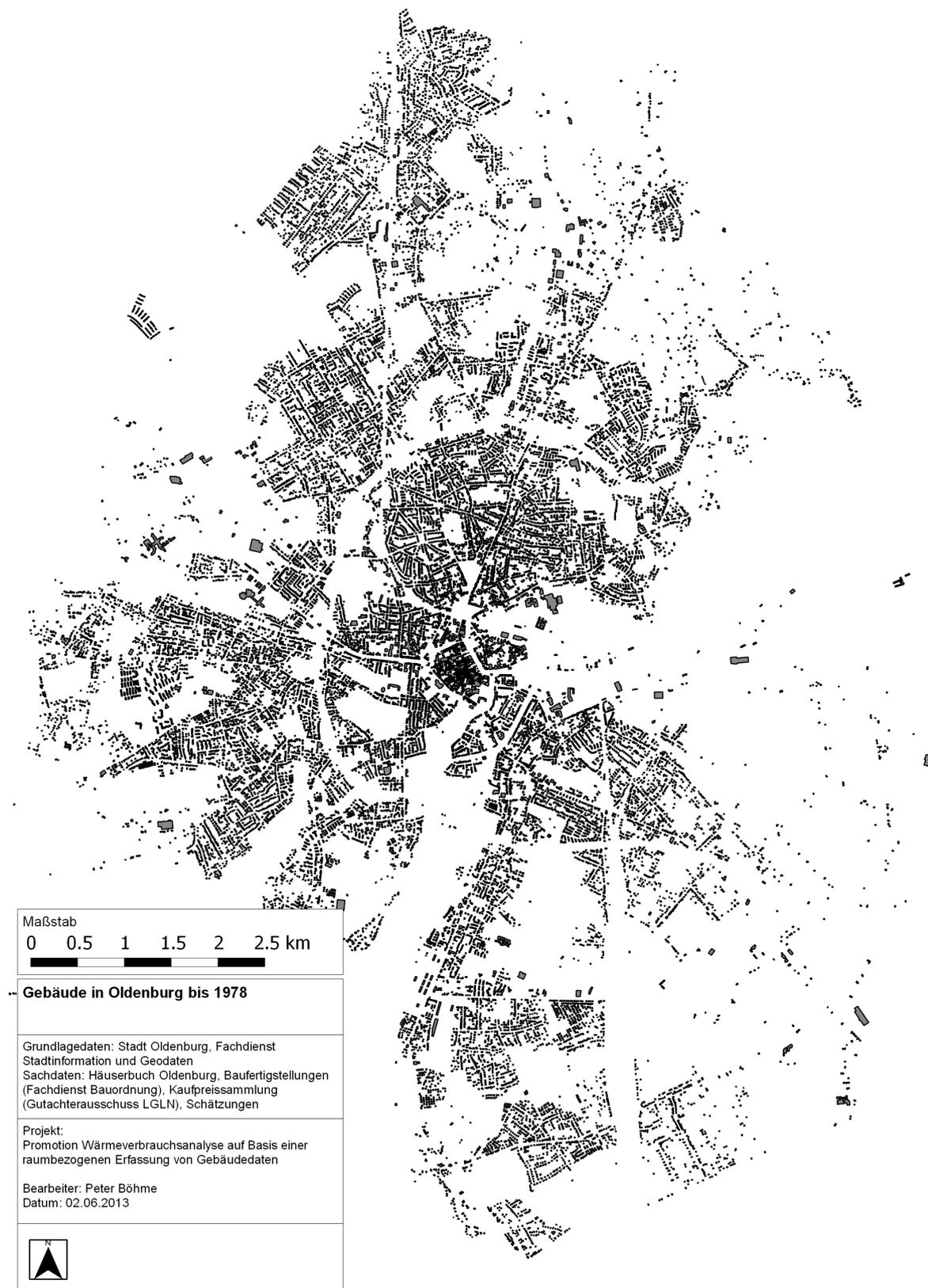


Abbildung A.5.: Karte der Gebäude mit Baujahr bis 1978

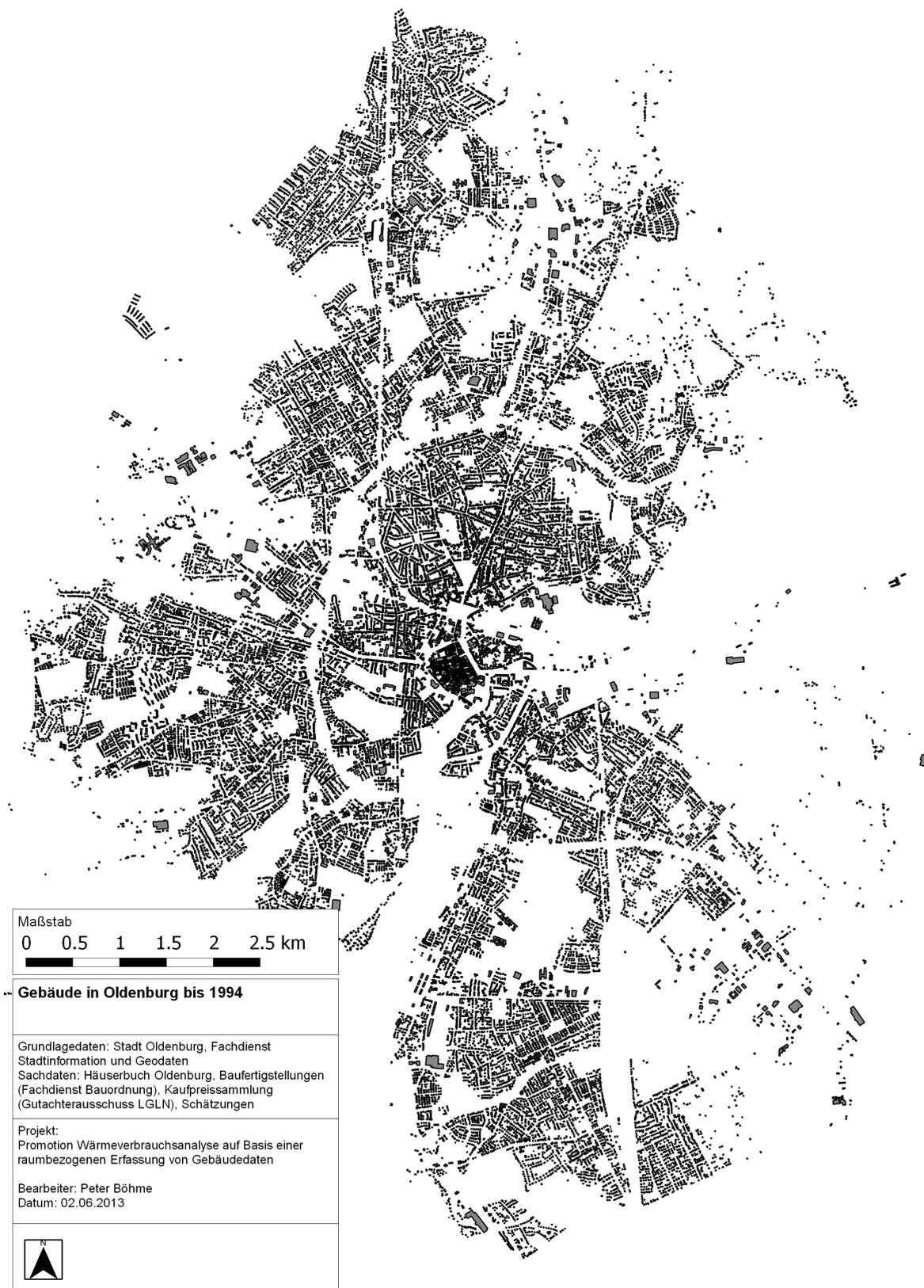


Abbildung A.6.: Karte der Gebäude mit Baujahr bis 1994

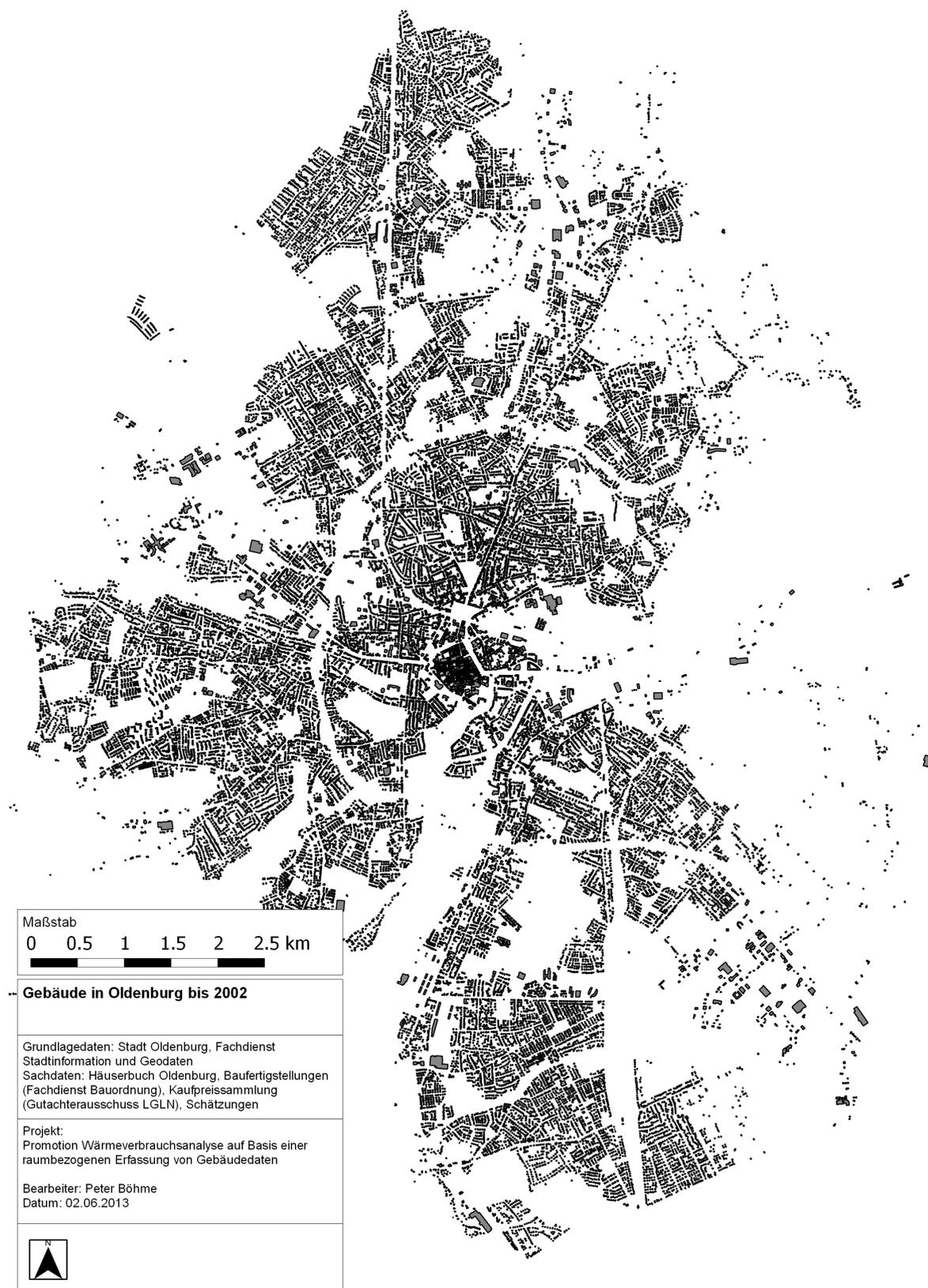


Abbildung A.7.: Karte der Gebäude mit Baujahr bis 2002

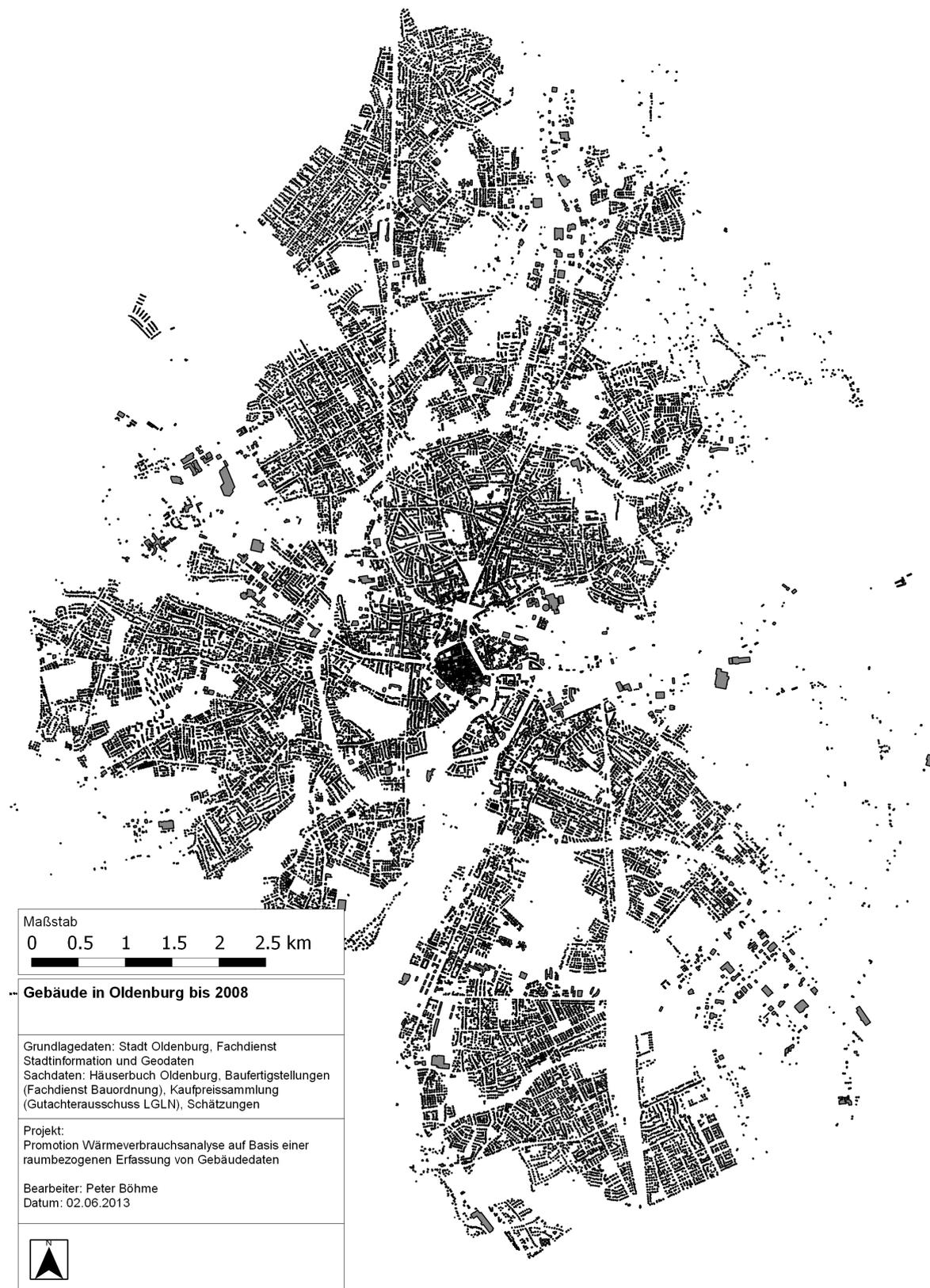


Abbildung A.8.: Karte der Gebäude mit Baujahr bis 2008

A.2. Transformation der Stromvertragsklassen auf die GHD-HH-Klassen

ID	Beschreibung	ID	Beschreibung
100	Haushalte	0-1	Haushalte EFH
110	Haushalte - Gemeinschaftsanlagen	0-2	Haushalte MFH
111	Haushalte - Heizungen für Mehrfamilienhäuser	0-2	Haushalte MFH
120	Haushalte - NATO-Angehörige	0-2	Haushalte MFH
200	Land-/Forstwirtschaft	9-10	Landwirtschaft
201	Gärtnereien, Baumschulen	10-11	Gartenbau u. Gärtnereien
202	Milchwirtschaft	9-10	Landwirtschaft
203	übrige Landwirtschaft	9-10	Landwirtschaft
400	Steinkohlebergbau	IN-	Industrie
401	Braunkohlebergbau	IN-	Industrie
402	Erdöl-und Erdgasgewinnung	IN-	Industrie
403	Torfgewinnung und -veredelung	IN-	Industrie
404	Übriger Bergbau	IN-	Industrie
405	Stein-, Sand- und Tongewinnung	IN-	Industrie
410	Zementherstellung	IN-	Industrie
411	Übrige Verarbeitung von Steinen und Erden	IN-	Industrie
412	Feinkeramik	3-1	Metallgewerbe
413	Glasherstellung und -verarbeitung	3-1	Metallgewerbe
420	Metallherst. eisenschaffende Betriebe	3-1	Metallgewerbe
421	Metallherst. Eisen-,Stahl- u.Tempergießereien	IN-	Industrie
422	Metallherst. Eisen-Ziehereien u.Kaltwalzwerke	3-1	Metallgewerbe
425	Metallherst. NE-Metallerzine-Metallhalbzeugw.	IN-	Industrie
426	Metallherst. NE-Metallgießereien	IN-	Industrie
427	Metallherst. NE-Metallziehereien	IN-	Industrie
430	Metallverarb. Elektronische Betriebe	3-1	Metallgewerbe
431	Metallverarb. Schiffbau	3-1	Metallgewerbe
432	Metallverarb. Strassenfahrzeugbau	3-1	Metallgewerbe
433	Metallverarb. Luftfahrzeug- u.Raumfahrzeugbau	3-1	Metallgewerbe
434	Metallverarb. Maschinenbau	3-1	Metallgewerbe
435	Metallverarb. Stahl-,Leichtmetallbau,Mechanik	3-1	Metallgewerbe
436	Metallverarb. Eisen-, Blech- und Metallwaren	3-1	Metallgewerbe
437	Metallverarb. Büromasch. DV-Geräte,-einricht.	3-1	Metallgewerbe
440	Chemie (im engeren Sinne)	IN-	Industrie
441	Kunststoffe Herstellung und Verarbeitung	3-1	Metallgewerbe
442	Gummi	3-1	Metallgewerbe

Fortsetzung auf der nächsten Seite

A.2. Transformation der Stromvertragsklassen auf die GHD-HH-Klassen

ID	Beschreibung	ID	Beschreibung
443	Spalt- und Brutstoffe Herstellung und Verarb.	IN-	Industrie
446	Mineralöle	IN-	Industrie
450	Holzbearbeitung	3-3	Holzgewerbe
451	Holzverarbeitung	3-3	Holzgewerbe
452	Zellstoff-,Holzschliff-,Papier- u.Pappeerzeug	IN-	Industrie
453	Papier- und Pappeverarbeitung	IN-	Industrie
454	Druckerei- und Vervielfältigungsgewerbe	3-4	Papier- u. Druckgewerbe
460	Textilherstellung	12-28	Bekleidung, Leder, Textil
461	Textilverarbeitung	12-28	Bekleidung, Leder, Textil
462	Ledererzeugung und -verarbeitung	12-28	Bekleidung, Leder, Textil
470	Mühlen	7-7	Restl. Nahrungsmittelge- werbe
471	Fisch- und Fleischverarbeitung	7-6	Fleischereien
472	Milchverarbeitung	3-1	Metallgewerbe
473	Nährmittel,Kartoffelerz,Obst- u.Gemüseverarb.	3-1	Metallgewerbe
474	Backwaren	7-5	Backgewerbe
475	Getränkeherstellung	IN-	Industrie
476	Übrige Nahrungs- und Genußmittel	7-7	Restl. Nahrungsmittelge- werbe
477	Futtermittel	7-7	Restl. Nahrungsmittelge- werbe
478	Tierische Veredelung	7-6	Fleischereien
480	Baugewerbe	1-9	Baugewerbe
481	Baustellen	1-9	Baugewerbe
482	Zuchtbetriebe	9-10	Landwirtschaft
489	Übrige Herstellung und Verarbeitung	3-1	Metallgewerbe
490	Energie- und Wasserversorgungsunternehmen	IN-	Industrie
491	Betriebsverbrauch	IN-	Industrie
492	EWE-Kavernen - Unterhaltung	IN-	Industrie
493	EWE-Kavernen - Ausspülung	IN-	Industrie
494	EWE-Verbrauch zur Energieerzeugung	IN-	Industrie
499	Netzbetreiber	IN-	Industrie
600	Bundesbahn	2-32	Deutsche Bahn AG
601	Andere Eisenbahnen	2-32	Deutsche Bahn AG
602	Straßenbahn- und Omnibusverkehr	3-1	Metallgewerbe
604	Schifffahrtsanlagen	12-34	Spedit., Lagerei, Verkehrs- vermittlung
606	Transport in Rohrleitungen	IN-	Industrie
607	Sonst Verkehrs-, Reise- u. Transportvermittl.	2-18	Sonst. betr. Dienstleistun- gen

Fortsetzung auf der nächsten Seite

A. Anhang

ID	Beschreibung	ID	Beschreibung
608	Post und Telekommunikation	2-30	Deutsche Bundespost / Postdienste
609	Mobilfunkanlagen	2-31	Telekommunikation
620	Behörden und öffentliche Verwaltungen	2-20	Gebietskörpersch. u. Sozialversich.
621	Städtische Abnahmestellen	IN-	Industrie
630	Schulen, Kindergärten	5-22	Schulen
631	Hochschulen, selbst. Forschungseinrichtungen	5-22	Schulen
632	Sportstätten	3-1	Metallgewerbe
633	Rundfunk- und Fernseheinrichtungen	2-31	Telekommunikation
634	Verlage	2-17	Verlagsgewerbe
635	Sonst. Kultur- und Unterhaltungseinrichtungen	6-19	Org. ohne Erwerbszweck und Heime
640	Einricht. für stationäre medizin. Behandlung	5-21	Krankenhäuser
641	Einricht. für ambulante medizin. Behandlung	2-18	Sonst. betr. Dienstleistungen
642	Einrichtungen für Tiermedizin	2-18	Sonst. betr. Dienstleistungen
650	Wohnheime und ähnliche Einrichtungen	6-19	Org. ohne Erwerbszweck und Heime
651	Kirchen und deren Verwaltungen	6-19	Org. ohne Erwerbszweck und Heime
652	Berufsorganisationen, Vereine, Parteien	6-19	Org. ohne Erwerbszweck und Heime
653	Einrichtungen der Abfallbeseitigung	IN-	Industrie
654	Einrichtungen der Abwasserbeseitigung	2-20	Gebietskörpersch. u. Sozialversich.
656	Friedhöfe	10-11	Gartenbau u. Gärtnereien
660	Großhandel, Außenhandel	4-27	Großhandel - nonfood
661	Großfl.Einzelhan., Warenhäuser, Einkaufszentren	4-25	Einzelhandel - nonfood
662	Anderer Einzelhandel	4-25	Einzelhandel - nonfood
663	Versandhandel	4-29	Handelsvermittlungen
664	Kühlhäuser und Trocknungsanlagen	IN-	Industrie
665	Lagerhaltungen	12-34	Spedit., Lagerei, Verkehrsvermittlung
670	Beherbergungsbetriebe mit u. ohne Gaststätten	6-15	Beherbergungsgewerbe
671	Gaststätten	6-16	Gaststättengewerbe
672	Campingplätze	6-15	Beherbergungsgewerbe
680	Heizzentralen, Heizwerke, Blockheizungen	IN-	Industrie
690	Körperpflege	IN-	Industrie

Fortsetzung auf der nächsten Seite

A.2. Transformation der Stromvertragsklassen auf die GHD-HH-Klassen

ID	Beschreibung	ID	Beschreibung
691	Reinigungen, Färbereien	8-8	Wäschereien u. (chem.) Reinigungen
692	Kreditinstitute und Versicherungsgewerbe	2-14	Kreditinstitute u. Versiche- rungen
693	Messewesen	2-18	Sonst. betr. Dienstleistun- gen
699	Übrige Dienstleistungen	2-18	Sonst. betr. Dienstleistun- gen
800	E-Verteiler	IN-	Industrie

Tabelle A.1.: Transformation der Nutzungsklassen der Stromverträge auf die GHD-HH-Klassen

Literaturverzeichnis

- [1] Abteilung Automatische Sprachverarbeitung. Wortschatzportal der Universität Leipzig. <http://wortschatz.uni-leipzig.de/>. [online; accessed 02.05.2013].
- [2] Lars Altendorf, Lars Kühl, Thomas Wilken, Norbert Fisch, Thomas Gawron, and Edmund Brandt. Vergleichswerte für Verbrauch bei Wohngebäuden. Berlin, 2012.
- [3] Klaus Backhaus. *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung*. Springer, Berlin [u.a.], 13 edition, 2011.
- [4] Norbert Bartelme. *Geoinformatik: Modelle, Strukturen, Funktionen*. Springer, Berlin, 4 edition, 2005.
- [5] Douglas Bates, John Chamber, Peter Dalgaard, and Robert Gentleman. The R Project for Statistical Computing. <http://www.r-project.org/>. [online; accessed 09.04.2013].
- [6] Bayerisches Landesamt für Umwelt (LfU). Agenda 21 - nachhaltige Entwicklung für das 21. Jahrhundert, 2004.
- [7] Ralf Bill and Marco L. Zehner. *Lexikon der Geoinformatik*. Wichmann, Heidelberg, 2001.
- [8] Peter Blüm. Einführung zur Fehlerrechnung im Praktikum. <http://www-ekp.physik.uni-karlsruhe.de/~bluem/Fehleranalyse.pdf>, 2002. [online; accessed 16.02.2013].
- [9] P. Böhme and T. Hamacher. Verbesserte Aussagekraft kommunaler Wärmebedarfsanalysen. In Andreas Koch, Ralf Bill, and Andreas Donabauer, editors, *Geoinformationssysteme 2013*, pages 7–16. Wichmann, 2013.
- [10] F. Botzenhart and T. Hamacher. Optimierung urbaner Energiesysteme: Analysen zur Struktur der Wärmeversorgung in Salzburg. In Andreas Koch, T. Kutzner, and Tobias Eder, editors, *Geoinformationssysteme*, pages 305–313. Wichmann, Berlin, 2012.
- [11] Thomas Brinkhoff. *Geodatenbanksysteme in Theorie und Praxis: Einführung in objektrelationale Geodatenbanken unter besonderer Berücksichtigung von Oracle Spatial*. Wichmann, Heidelberg, 3 edition, 2012.
- [12] BulwienGesa AG and Pesch und Partner. Machbarkeitsstudie Konversionsflächen: im Kontext des Immobilienmarktes der Stadt Oldenburg. Oldenburg, 2009.

- [13] Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit (BMU). Die Nationale Klimaschutzinitiative des Bundesumweltministeriums. http://www.bmu-klimaschutzinitiative.de/de/ziele_und_bilanz, 2012. [online; accessed 29.05.2013].
- [14] Bundesministerium für Verkehr, Bau und Stadtentwicklung (BMVBS). Bekanntmachung der Regeln für Energieverbrauchskennwerte im Wohngebäudebestand, 2009.
- [15] Bob Carpenter and Breck Baldwin. Text Analysis with LingPipe 4. New York, 2011.
- [16] Destatis. Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008). <https://www.destatis.de/DE/Methoden/Klassifikationen/GueterWirtschaftsklassifikationen/Content75/KlassifikationWZ08.html>, 2008. [online; accessed 13.02.2013].
- [17] Deutscher Bundestag. Verordnung über energiesparenden Wärmeschutz und energiesparende Anlagentechnik bei Gebäuden (Energieeinsparverordnung – EnEV), Fassung vom 30.09.2009: EnEV. [online; accessed 14.02.2011].
- [18] Deutscher Bundestag. Verordnung zur Berechnung der Wohnfläche, Fassung vom 25.11.2003: WoFlV.
- [19] Deutscher Wetterdienst. Klimafaktoren für Energieverbrauchsausweise. <http://www.dwd.de/klimafaktoren>, 2012.
- [20] N. Diefenbach, H. Cischinsky, M. Rodenfels, and K.-D Clausnitzer. Datenbasis Gebäudebestand: Datenerhebung zur energetischen Qualität und zu den Modernisierungstrends im deutschen Wohngebäudebestand. Darmstadt, 2010.
- [21] DIN Deutsches Institut für Normung e. V. Grundflächen und Rauminhalte (DIN 277), 02.2005.
- [22] DIN Deutsches Institut für Normung e. V. Energetische Bewertung von Gebäuden – Berechnung des Nutz-, End- und Primärenergiebedarfs für Heizung, Kühlung, Lüftung, Trinkwarmwasser und Beleuchtung (DIN V 18599), 02.2007.
- [23] DIN Deutsches Institut für Normung e. V. Energetische Bewertung von Gebäuden - Berechnung des Nutz-, End- und Primärenergiebedarfs für Heizung, Kühlung, Lüftung, Trinkwarmwasser und Beleuchtung - Teil 10: Nutzungsrandbedingungen, Klimadaten, 02.2007.
- [24] DIN Deutsches Institut für Normung e. V. Wärmeschutz und Energie-Einsparung in Gebäuden - Teil 6: Berechnung des Jahresheizwärme- und des Jahresheizenergiebedarfs (DIN V 4108-6), 06.2003.
- [25] Jeffrey E. F. Friedl. *Mastering regular expressions*. O'Reilly, Sebastapol and Calif, 3rd edition, 2006.
- [26] Institut für Wohnen und Umwelt (IWU). Deutsche Gebäudetypologie: Systematik und Datensätze. Darmstadt, 2003.

- [27] Institut für Wohnen und Umwelt (IWU). Gradtagszahlen für Deutschland. <http://www.iwu.de/downloads/fachinfos/energiebilanzen/#c203>, 2011. [online; accessed 13.09.2012].
- [28] Bernd Geiger and M. Hardi. Energieanstoßberatung des Kaminkehrerhandwerks in Bayern. München, 2002.
- [29] Christian Geiß, Hannes Taubenböck, Michael Wurm, Thomas Esch, Michael Nast, Christoph Schillings, and Thomas Blaschke. Remote Sensing-Based Characterization of Settlement Structures for Assessing Local Potential of District Heat. *Remote Sensing*, 3(7):1447–1471, 2011.
- [30] Timo Gniechwitz, Astrid Holz, Thorsten Schulze, and Dietmar Walberg. *Wohnungsbau in Deutschland - 2011 - Modernisierung oder Bestandsersatz: [Studie zum Zustand und der Zukunftsfähigkeit des deutschen "Kleinen Wohnungsbaus"]*, volume 59 of *Bauforschungsbericht*. Arbeitsgemeinschaft für Zeitgemäßes Bauen, Kiel, 2011.
- [31] T. Gobmaier, R. Corradini, D. Kraus, and U. Wagner. Strukturoptimierung in Ballungsgebieten: Energiebedarfsprognose für die Stadt München. München, 2007.
- [32] Martin Greller, Franz Schröder, Tobias Bögelein, Volker Hundt, and Peter Güttler. Universelle Energiekennzahlen für Deutschland - Teil 3: Spezifischer Energieverbrauch für zentrale Warmwasserbereitung und Relation zum Heizenergieverbrauch. *Bauphysik*, 32(3):154–159, 2010.
- [33] Martin Greller, Franz Schröder, Volker Hundt, Bernhard Mundry, and Olaf Papert. Universelle Energiekennzahlen für Deutschland - Teil 2: Verbrauchskennzahlenentwicklung nach Baualtersklassen. *Bauphysik*, 32(1):1–6, 2010.
- [34] The PostGIS Development Group. Postgis Manual Using PostGIS: Chapter 4: Data Management and Queries. <http://postgis.refractor.net/documentation/manual-1.5/ch04.html>. [online; accessed 09.04.2013].
- [35] Guido Halbig and Franz Schröder. Ortsgenaue Klimafaktoren für Deutschland - Teil I: Meteorologisches Modell und Bezug zu regulären klimakorrekturenkonventionen. *HLH Heizung, Lüftung/Klima, Haustechnik*, 10(62):50–55, 2011.
- [36] Guido Halbig and Franz Schröder. Ortsgenaue Klimafaktoren für Deutschland - Teil II: Empirische Vergleichsstudie über mit herkömmlichen und neuen Klimafaktoren generierte Energiekennwerte. *HLH Heizung, Lüftung/Klima, Haustechnik*, 11(62):37–42, 2011.
- [37] Gerhard Hausladen and Thomas Hamacher. Leitfaden Energienutzungsplan. München, 2010.
- [38] Joachim Herrmann. *Urbane Energieversorgung in Deutschland bis 2050: am Beispiel der Stadt Augsburg*. PhD thesis, Universität Augsburg, Augsburg, 2011.
- [39] Andreas Heuer, Gunter Saake, and Kai-Uwe Sattler. *Datenbanken kompakt*. mitp, Bonn, 1. edition, 2001.

- [40] Gerhard Heyer, Uwe Quasthoff, and Thomas Wittig. *Text mining: Wissensrohstoff Text : Konzepte, Algorithmen, Ergebnisse*. W3L-Verl., Herdecke, 2006.
- [41] Stefan Himburg. Energetische Bilanzierung von Wohngebäuden nach DIN V 18599. *Bauphysik*, 33(2):99–110, 2011.
- [42] Kurt Hornik, Reinhold Hatzinger, and Herbert Nagel. *R: Einführung durch angewandte Statistik*. Pearson Studium ein Imprint der Pearson Education, München, 1. edition, 2011.
- [43] infas geodaten. Datenkatalog Marktinformationen: Neu-Auflage 2012. Bonn, 2012.
- [44] T. Janßen, P. Lorkowski, S. Miotk, T. Schröder, and L. Sylvester. Energieeffiziente Stadt Oldenburg: Masterprojekt zum Konzept. Oldenburg, 2010.
- [45] Robert Kaden, Michael Prytula, Andreas Krüger, and Thomas Kolbe. Energieatlas Berlin: Vom Gebäude zur Stadt: Am Beispiel zur Abschätzung der Wärmeenergiebedarfe von Gebäuden. In Andreas Koch, Ralf Bill, and Andreas Donabauer, editors, *Geoinformationssysteme 2013*. Wichmann, 2013.
- [46] J. Knissel and R. Alles. Ökologischer Mietspiegel: Empirische Untersuchung über den möglichen Zusammenhang zwischen der Höhe der Vergleichsmiete und der wärmetechnischen Beschaffenheit des Gebäudes. Darmstadt, 2003.
- [47] M. Konchady. *Text mining application programming*. Charles River Media, Hingham and MA, 2006.
- [48] Greg Kondrak. N-gram similarity and distance. In Mariano Consens and Gonzalo Navarro, editors, *String processing and information retrieval*, pages 115–126. Springer, 2005.
- [49] Landesamt für Geoinformation und Landentwicklung Niedersachsen (LGLN). Gutachterausschüsse für Grundstückswerte in Niedersachsen. <http://www.gag.niedersachsen.de/>. [online; accessed 04.05.2013].
- [50] Wolfgang Lezius. Morphy. <http://www.wolfganglezius.de/doku.php?id=cl:morphy>. [online; accessed 02.05.2013].
- [51] Markus Lichtmeß and Karsten Voss. EnerCalC. <http://www.enob.info/?id=EnerCalC>, 2012. [online; accessed 25.01.2013].
- [52] Christian Linder. Konzept zur Erstellung kommunaler Gebäudedatenbanken in Bayern: Datenlagegrundlagen und Umsetzung am Beispiel Augsburgs, Diplomarbeit. http://www.staedtestatistik.de/fileadmin/kosis/AGK/DIPLOMARBEIT_Geb_DB_Linder.pdf, 2010. [online; accessed 25.01.2013].
- [53] T. Loga, N. Diefenbach, and R. Born. Deutsche Gebäudetypologie: Beispielhafte Maßnahmen zur Verbesserung der Energieeffizienz von typischen Wohngebäuden, 2011.
- [54] Heidi Ludewig, Karin Jahn, Bernd Eikmeier, Klaus Schäfer-Breede, Markus Otten, Kornelia Gerwien-Siegel, and Bernd Langer. Integriertes Energie- und Klimaschutzkonzept für die Stadt Oldenburg (InEKK): Schlussbericht, 2011.

- [55] G. Meinel, R. Hecht, and H. Herold. Verfahren zur Erhebung, Analyse und Visualisierung von Gebäudebestands- und Siedlungsentwicklungen auf Grundlage Topographischer Kartenreihen: Land Use-Discussion Paper, 2009.
- [56] G. Meinel, R. Hecht, H. Herold, and G. Schiller. Automatische Ableitung von stadtstrukturellen Grundlagendaten und Integration in einem Geographischen Informationssystem: Abschlussbericht. Bonn, 2008.
- [57] Isabell Nemeth. *Methodenentwicklung zur Bestimmung von Potenzialen der Energieeffizienzsteigerung im Haushalts- und GHD-Sektor: - Am Beispiel von drei Landkreisen in Bayern -*. PhD thesis, Technische Universität München, München, 2011.
- [58] Joachim Nitsch, Thomas Pregger, Tobias Naegler, Dominik Heide, Diego Luca de Tena, Franz Trieb, Yvonne Scholz, Kristina Nienhaus, Norman Gerhardt, Michael Sterner, Tobias Trost, Amany von Oehsen, Rainer Schwinn, Carsten Pape, Henning Hahn, Manuel Wickert, and Bernd Wenzel. Langfristszenarien und Strategien für den Ausbau der erneuerbaren Energien in Deutschland bei Berücksichtigung der Entwicklung in Europa und global, 2012.
- [59] Klaus Nottebaum, Frank Schöler, Robert Steinberger-Wilckens, and Hubert Westkämper. Klimaschutzbericht 1997 für die Stadt Oldenburg, 1997.
- [60] Open Geospatial Consortium (OGC). CityGML. <http://www.opengeospatial.org/standards/citygml>, 2013. [online; accessed 01.06.2013].
- [61] Bert Oschatz. Zwei Rechenverfahren - ein Ergebnis? DIN V 4108-6 / DIN V 4701-10 und DIN V 18599 bei Wohngebäuden. http://www.enbw-energiegemeinschaft.de/~upload/enbw-eg/veranstaltungen2009/cep_2009/vortragcep2009oschatz.pdf, 29.01.2009. [online; accessed 10.06.2013].
- [62] Angelika Paar, Florian Herbert, Martin Pehnt, Susanne Ochse, Stephan Richter, Stefanie Maier, Magalie Kley, Heiko Huther, Jens Kühne, and Ingo Weidlich. Transformationsstrategien von fossiler zentraler Fernwärmeversorgung zu Netzen mit höheren Anteilen erneuerbarer Energien: Endbericht: Gefördert vom Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit. Heidelberg and Leimen and Frankfurt am Main, 2013.
- [63] W. Pfaffenberger and W. Ströbele. Energiegutachten für die Stadt Oldenburg. Oldenburg, 1992.
- [64] PostGIS. PostGIS - Spatial and Geographic objects for PostgreSQL. <http://postgis.net/>. [online; accessed 09.04.2013].
- [65] QGIS Development Team. Quantum GIS. <http://www.qgis.org/>. [online; accessed 09.04.2013].
- [66] D. Reiter, F. Botzenhart, P. Mühlich, T. Hamacher, and A. Reuter. Developing a Roadmap for the Future Energy Infrastructure. In IAEE, editor, *10th IAEE European Conference*, pages 1361–1366, 2009.

- [67] Edwin Schicker. *Datenbanken und SQL: Eine praxisorientierte Einführung mit Hinweisen zu Oracle und MS-Access*. Teubner, Stuttgart and Leipzig and Wiesbaden, 3. edition, 2000.
- [68] Barbara Schlomann, Edelgard Gruber, Bernd Geiger, Heinrich Kleeberger, Urs Wehmhörner, Till Herzog, and Daria-Maria Konopka. *Energieverbrauch des Sektors Gewerbe, Handel, Dienstleistungen (GHD) für die Jahre 2004 - 2006*, 2009.
- [69] Franz Schröder, Martin Greller, Volker Hundt, Bernhard Mundry, and Olaf Papert. *Universelle Energiekennzahlen für Deutschland - Teil 1: Differenzierte Kennzahlverteilungen nach Energieträger und wärmetechnischem Sanierungsstand*. *Bauphysik*, 31(6):393–402, 2009.
- [70] Shell Deutschland Oil GmbH, HWWI, and ifeu. *Shell Hauswärme-Studie: Nachhaltige Wärmeerzeugung für Wohngebäude: Fakten, Trends und Perspektiven*, 2011.
- [71] B. Siepe, W. Schulz, Th Tech, and K.-H Schlusche. *Einsparmöglichkeiten beim Raumwärmebedarf des Wohngebäudebestandes in Hannover und Langenhagen*. Hannover, 1992.
- [72] Stadt Oldenburg. *Statistische Daten Online*. <http://www.oldenburg.de/startseite/politik/verwaltung/statistik/statistische-daten-online/inhaltsverzeichnis.html>. [online; accessed 20.01.2013].
- [73] Statistisches Bundesamt. *Merkmalsübersicht: Gebäude und Wohnungen*. https://www.zensus2011.de/SharedDocs/Downloads/DE/Merkmale/Merkmale_GWZ.pdf?__blob=publicationFile&v=6, 2013. [online; accessed 07.06.2013].
- [74] Statistisches Bundesamt. *Zensus 2011 - Geschichte: Die Volkszählung 1987*. <https://www.zensus2011.de/>, 2013. [online; accessed 02.05.2013].
- [75] Minna Sunikka-Blank and Ray Galvin. *Introducing the prebound effect: the gap between performance and actual energy consumption*. *Building Research & Information*, 40(3):260–273, 2012.
- [76] Hannes Taubenböck and S. Dech, editors. *Fernerkundung im urbanen Raum: Erdbeobachtung auf dem Weg zur Planungspraxis*. WBG, Darmstadt, 2010.
- [77] Techem. *Energiekennwerte 2011*, 2011.
- [78] The PostgreSQL Global Development Group. *PostgreSQL*. <http://www.postgresql.org/>. [online; accessed 06.02.2012].
- [79] Rainer Vallentin. *Energieeffizienter Städtebau mit Passivhäusern: Begründung belastbarer Klimaschutzstandards im Wohnungsbau*. Cuvillier, Göttingen, 1 edition, 2011.
- [80] Verein deutscher Ingenieure (VDI). *Energie- und Wasserverbrauchskennwerte für Gebäude - Grundlagen (VDI 3807 Blatt 1)*, 03.2007.

- [81] Vereinigung Deutscher Elektrizitätswerke (VDEW). Zuordnung der VDEW-Lastprofile zum Kundengruppenschlüssel. Frankfurt am Main, 2000.
- [82] Albert Zimmermann. *Basismodelle der Geoinformatik: Strukturen, Algorithmen und Programmierbeispiele in Java*. Hanser, Carl, München, 2011.

Akronyme

A_N Energienutzfläche 85

ALK Automatisierte Liegenschaftskarte 29, 30, 33, 53, 54, 72–74, 77, 88, 90

BGF Bruttogrundfläche 84, 86

BmVBS Bundesministerium für Verkehr, Bau und Stadtentwicklung 12

DWD Deutscher Wetterdienst 99

EFH/ZFH Ein- oder Zweifamilienhäuser 74–76, 86, 88, 91, 92

EPSG *European Petroleum Survey Group Geodesy* 8

GHD Gewerbe, Handel und Dienstleistung 1, 72, 74, 87, 88, 91

GIS Geoinformationssystem 7

GTZ Gradtagszahl 11–13

IWU Institut für Wohnen und Umwelt, Darmstadt 14, 64, 71, 72, 80, 82, 86, 103, 104, 118, 120

KWK Kraft-Wärme-Kopplung 1

LSKN Landesamt für Statistik und Kommunikation Niedersachsen 68, 71

LSKN Landesamt für Statistik und Kommunikation Niedersachsen 52, 59

MFH Mehrfamilienhäuser 76, 88, 91, 92

MFHgr große Mehrfamilienhäuser 74–76

MFHkl kleine Mehrfamilienhäuser 73–76, 82, 91

MG mischgenutzte Gebäude 73, 74

NF Nutzfläche 84

NGF Nettogrundfläche 84

NWG Nichtwohngebäude 72–74, 76, 92

RH Reihenhäuser 74, 76, 82, 91, 92

SQL *Structured Query Language* 5, 7

WF Wohnfläche 85

WG Wohngebäude 73

WGkl kleine Wohngebäude 73, 74

Abbildungsverzeichnis

2.1. Darstellung zweier in Relation zueinander stehender Tabellen	6
2.2. Tabelle mit Sach- und Geometrieattributen	8
2.3. Karte der drei Grundgeometriearten	8
2.4. Ablaufdiagramm von <i>Text Mining</i> -Analysen	9
2.5. Energieumwandlungskette mit Bezug zur Raumwärme	12
2.6. Lineare Regression für Wärmeverbräuche von Gebäuden unterschiedlichen Volumens	16
3.1. Ablaufdiagramm der Gebäudedatenverarbeitung	20
3.2. Häufigkeitsverteilung der Wörter in den Textfeldern der Gebäudekarteikarten	26
3.3. Kartenausschnitt zur Darstellung der fünf Raumbezugsklassen in Oldenburg	28
3.4. Datenmodell der Raumbezugsklassen	29
3.5. Beispiel für den Strukturreichtum von Primärparameterangaben	38
3.6. Datenmodell der Primärparameterverwaltung	39
3.7. Beispiel für die Primärparameterverwaltung	40
3.8. Beispiel für die Zusammenführung der Primärparameter auf Metadatenebene	42
3.9. Beispiel für die Zusammenführung der Primärparameter auf Sachdatenebene	42
3.10. Statistik der Primärparameterverwaltung für Parameter und Datenquellen	46
3.11. Datenmodell der Gebäudeparameter	47
4.1. Kartenausschnitt mit manuell erkannten Flachdächern in Oldenburg . . .	55
5.1. Informationsfluss zwischen Primär- und Gebäudeparametern	62
5.2. Paarweiser Vergleich der Quellenangaben zum Parameter Baujahr	66
5.3. Häufigkeitsverteilung der konfliktbereinigten Angaben zum Gebäudebaujahr je Datenquelle	69
5.4. Häufigkeitsverteilung der konfliktbereinigten Angaben zum Gebäudebaujahr je Baualtersklasse und Datenquelle	70
5.5. Zeitreihe der Neubauten mit jährlicher Auflösung seit 1970 und Literaturvergleich	71
5.6. Entscheidungsbaum zur Bestimmung der Gebäudeklasse	73
5.7. Kartenausschnitt mit farblich klassifizierten Gebäuden in Oldenburg . . .	74
5.8. Kumulierte Häufigkeitsverteilung der Gebäudevolumina und -höhen je Gebäudeklasse	75
5.9. Mosaikplot der Baualters- und Gebäudeklassen	77
5.10. Visualisierung der Gebäudehüllen-Parameter	78
5.11. Beispiel der thermischen Hüllflächen für eine Doppelhaushälfte	82
5.12. A/V-Verhältnisse der Gebäudetypen und Literaturvergleich	83
5.13. Entscheidungsbaum für die Bestimmung der Geschosszahl	84

5.14. Grundflächen-Begriffe in Normen und Verordnungen	85
5.15. Wohnfläche-Volumen-Verhältnisse der Gebäudetypen und Literaturvergleich	87
5.16. Wohnfläche je Gebäudeklasse in Oldenburg	91
5.17. Nutzfläche je Branche in Oldenburg	92
5.18. Karte der berechneten Nutzflächen der Branche Handel in Oldenburg, Hintergrundgrafik: OpenStreetMap	93
6.1. Häufigkeitsverteilung der Gebäudeparameter Volumen, A/V-Verhältnis und Baualterklasse	97
6.2. Mittlerer spezifischer Endenergieverbrauch je m ² Energienutzfläche	100
6.3. Mittlerer spezifischer Endenergieverbrauch je Baujahr und Zusatzinforma- tionen	102
6.4. Mittlerer spezifischer Endenergieverbrauch pro m ² Energienutzfläche und Literaturvergleich	104
6.5. Mittlerer spezifischer Endenergieverbrauch pro m ² Energienutzungsfläche je Baualterklasse, Volumen und A/V-Verhältnis	105
6.6. Kumulierte Verteilung der spezifischen Endenergieverbräuche	108
6.7. 95 %-Quantil des relativen statistischen Schätzfehlers in % bezogen auf den geschätzten Heizwärmeverbrauch	114
6.8. 95 %-Quantile des absoluten Schätzfehlers	116
6.9. Mittlerer Raumwärmebedarf je Gebäudetyp in Oldenburg und Literatur- vergleich	119
6.10. Spezifischer Endenergieverbrauch und -bedarf je Baualter- und Volumen- klasse für Ein- und Zweifamilienhäuser	120
7.1. Gebäudedatenverarbeitung und Typen der Analyse von Gebäude- und Verbrauchsdaten	124
A.1. Karte der Gebäude mit Baujahr bis 1918	128
A.2. Karte der Gebäude mit Baujahr bis 1948	129
A.3. Karte der Gebäude mit Baujahr bis 1956	130
A.4. Karte der Gebäude mit Baujahr bis 1968	131
A.5. Karte der Gebäude mit Baujahr bis 1978	132
A.6. Karte der Gebäude mit Baujahr bis 1994	133
A.7. Karte der Gebäude mit Baujahr bis 2002	134
A.8. Karte der Gebäude mit Baujahr bis 2008	135

Tabellenverzeichnis

2.1. Klimafaktoren für Oldenburg im Überblick	13
2.2. Berechnungsparameter für das Heizperiodenbilanzverfahren	14
3.1. Bestimmung der Attributtypen für einen anonymisierten Beispieldatensatz	22
3.2. Güteindikator-Skala und Entscheidungsgrundlage für die Einordnung einer Quelle	24
3.3. Raumbezüge und Zuordnungshäufigkeiten in Oldenburg	31
3.4. Raumbezüge der Datenquellen in Oldenburg	32
3.5. Hierarchie der für die Adresszuordnung relevanten <i>Tags</i>	34
3.6. Regeln für die Trennung von Straßenabschnitten	36
3.7. Parametertypen und Operationen für die Datenverarbeitung auf Sachebene	44
3.8. Primärparametertypen der Daten in Oldenburg	45
4.1. Gebäudedatenquellen in Oldenburg	49
4.2. Datensätze des Archivs (Bauordnung)	50
4.3. Datensätze der Bauanträge (Bauordnung)	51
4.4. Datensätze der digitalisierten Gebäudekarteikarten	52
4.5. Datensätze der ALK	54
4.6. Relevante Attribute der Kaufpreissammlung	56
4.7. Datensätze zu Stromanschlüssen	57
5.1. Datenquellen für den Parameter Baujahr	63
5.2. Baualtersklassen von Wohngebäuden	65
5.3. Anzahl und prozentualer Anteil der paarweise übereinstimmenden Baual- tersklassen	67
5.4. Häufigkeitsverteilung der Bebauungspläne	68
5.5. Gebäudeklassen und Klassenbeschreibung	72
5.6. Parameter der Gebäudehülle	79
5.7. Branchen, Nutzungsklassen und Anzahl der Gebäudeparteien in Oldenburg	89
5.8. Häufigste ALK-Objektarten mit den Gewichtungen für die Transformation der Nutzungsklassen	90
6.1. Beispiel eines Gebäudedatensatzes für die Verbrauchsauswertung	98
6.2. Anteil der erklärbaren Varianz an der gesamten Varianz der Endenergie- verbräuche	110
6.3. Szenarien für Schätzfehleranalyse	112
A.1. Transformation der Nutzungsklassen der Stromverträge auf die GHD-HH- Klassen	139

Danksagung

Die vorliegende Dissertation entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Max-Planck-Institut für Plasmaphysik und wurde gefördert durch die Konrad-Adenauer-Stiftung. Einen guten Teil meiner Arbeitszeit verbrachte ich als Gast am Lehrstuhl für Energiewirtschaft und Anwendungstechnik der TU München. Für die freundliche Aufnahme am Lehrstuhl und die nicht selbstverständliche Unterstützung durch die Mitarbeiterinnen und Mitarbeiter möchte ich mich herzlich bedanken.

Mein Dank gilt Prof. Ulrich Wagner für die Übernahme der Erstkorrektur, die Unterstützung meiner Arbeit und die Anregungen zum kritischen Hinterfragen in Ausrichtung und Struktur der Arbeit.

Ich danke Prof. Klaus Diepold für die Übernahme der Zweitkorrektur der Arbeit, sowie Prof. Josef Kindersberger für die Übernahme des Prüfungsvorsitzes.

Mein besonderer Dank gebührt Prof. Thomas Hamacher für die jahrelange Begleitung meiner Arbeit, die damit verbundenen inspirierenden Gespräche und die Motivation während den harten Zeiten einer Doktorarbeit.

Eine große Bereicherung waren für mich die fachlichen Gespräche mit meinen Kolleginnen und Kollegen am Max-Planck-Institut und am Lehrstuhl. Sie haben direkt und indirekt zum Gelingen dieser Arbeit beigetragen. Dafür möchte ich mich bei allen Kolleginnen und Kollegen und insbesondere bei Johannes Dorfner, Karl Janker, Katrin Schaber, Philipp Kuhn, Tobias Eder und Tobias Hartmann bedanken.

Herrn Michael Arndt und Herrn Henrik Tietje drücke ich meinen Dank aus für ihre außerordentliche Bereitschaft, mich in meiner Arbeit mit Daten und Kontakten zu weiteren Ansprechpartnern in Oldenburg zu unterstützen. In gleicher Weise möchte ich mich bei Prof. Manfred Weisensee bedanken, der mir einen für meine Promotion wichtigen Forschungsaufenthalt an der Jade Hochschule Oldenburg ermöglicht hat.

Für das sorgfältige Lektorat der Arbeit danke ich Johannes Dorfner, Uwe Böhme und Sarah Schöbel.

Bei meinen Eltern Karin und Uwe Böhme bedanke ich mich für Ihre ununterbrochene Unterstützung und Hilfsbereitschaft; nicht nur während der Zeit der Doktorarbeit, sondern auch auf dem Weg dahin.