# Agentenbasierte Städtesimulation zur integrierten Modellierung menschlichen Verhaltens und urbaner Versorgungssysteme

Von der Fakultät für Informatik der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

David Neuroth, M.Eng.

Berichtende:
Univ.-Prof. Dr. Christina Büsing
Univ.-Prof. Dr. Heidi Heinrichs

Tag der mündlichen Prüfung: 10. April 2026

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek verfügbar.

**David Neuroth**

*Agentenbasierte Städtesimulation zur integrierten Modellierung menschlichen Verhaltens und urbaner Versorgungssysteme*
12. Dezember, 2025

✉ neuroth@combi.rwth-aachen.de

## Abstract

Human behavior is essential to a variety of research topics. For example, the design of urban supply systems depends on the way people interact with them. This includes power distribution systems, as well as guaranteed access to essential resources, such as medical care. To determine the optimal setup of these systems, a flexible model is required that can map both current and future behavior in order to analyze changes in behavioral patterns. To that end, interactions of people with their environment need to be considered. Interdependencies regarding transportation and non-residential buildings are particularly important. To accurately account for these interactions, people must be represented not in isolation, but as part of a whole modeled city.

To accomplish this, an existing behavior simulation was expanded and turned into a complete city simulation that considers every resident in a city individually. In order to efficiently manage the significant computational demands, a parallel simulation architecture was developed. This framework distributes the workload and frequently synchronizes information to maintain integrity and plausibility.

The city simulation realistically models people's everyday behavior and generates electricity load profiles, transportation profiles, as well as visit frequency for supermarkets, pharmacies, and similar institutions. This integrated simulation allows for complex investigations. For example, it can help to assess the influence of size, location, and availability of selected non-residential buildings on people's activities and, subsequently, power consumption.

## Kurzfassung

Bei einer Vielzahl wissenschaftlicher Fragestellungen spielt menschliches Verhalten eine fundamentale Rolle. Beispielsweise ist für Design und Dimensionierung von Versorgungssystemen entscheidend, wann und auf welche Art Menschen mit diesen Systemen interagieren. Derartige Systeme können Strom- und Wärmenetze sein, aber etwa auch der gesicherte Zugang zu medizinischer Versorgung. Um diese Systeme zu optimieren, ist ein flexibles Modell notwendig, mit dem gegenwärtiges und zukünftiges Verhalten abgebildet und mit dem der Einfluss von Verhaltensänderungen analysiert werden kann. Hierbei spielen auch Wechselwirkungen von Personen mit ihrer Umgebung eine große Rolle. Insbesondere Interaktionen in Bezug auf städtische Infrastruktur wie Transportwege und Nichtwohngebäude können erheblichen Einfluss haben. Um diese Interaktionen realistisch modellieren und Interventionen integriert betrachten zu können, dürfen Personen dabei nicht isoliert modelliert werden. Stattdessen muss es möglich sein, eine Stadt als Ganzes vollständig in einem Modell abzubilden.

Daher wurde ein existierendes Modell zur allgemeinen Verhaltenssimulation von Menschen weiterentwickelt und zu einer vollwertigen Städtesimulation ausgebaut, sodass jede einzelne Person in ganzen Städten korrekt repräsentiert werden kann, inklusive der Interaktionen zwischen mehreren Personen. Hierfür wurde eine synchrone Simulationsarchitektur aufgebaut, die eine parallele Verhaltenssimulation aller Bewohner mit regelmäßiger Synchronisation über einen Nachrichtenaustausch kombiniert. Dadurch kann die enorme Rechenlast einer Städtesimulation unter Ausnutzung moderner Hochleistungsrechner effizient bewältigt werden, ohne die Integrität und Plausibilität des Modells zu beeinträchtigen.

Die Städtesimulation bildet realistisches Alltagsverhalten ab und kann anhand dessen Stromlastprofile, Transportprofile, aber auch Besuchshäufigkeiten in Nichtwohngebäuden wie Apotheken, Arztpraxen oder Supermärkten generieren. Erst durch die integrierte Simulation einer ganzen Stadt kann eine Vielzahl komplexer Fragestellungen untersucht werden, etwa welchen Einfluss Standorte, Größe und Verfügbarkeit verschiedener Nichtwohngebäude auf das alltägliche Verhalten und damit auf den Stromverbrauch haben können.

## Danksagung

In den über drei Jahren, die ich mit dem Anfertigen dieser Arbeit verbracht habe, habe ich von vielen Seiten aus meinem Umfeld unglaublich viel Hilfe und Unterstützung erfahren.

Zuallererst möchte ich mich bei Christina für die hervorragende Betreuung bedanken. Noch vor drei Jahren hätte ich nicht geglaubt, dass eine externe Promotion so konstruktiv und unkompliziert ablaufen kann. Christina hat mich, wo immer möglich, einbezogen, mich fachlich wie persönlich unterstützt und mir jederzeit hilfreiches Feedback und interessante Anregungen gegeben. Ohne sie als Betreuerin wäre diese Arbeit überhaupt nicht möglich gewesen.

Genauso möchte ich Noah danken, der mich die gesamte Zeit über angeleitet und betreut hat. Noah hat mir den Freiraum gegeben, meine Ideen auf meine eigene Weise umzusetzen, mich aber gleichzeitig immer wieder in die richtige Richtung gelenkt, und mir bei etlichen Problemen schnell und pragmatisch weitergeholfen.

Daneben möchte ich mich auch bei den tollen Kolleginnen und Kollegen vom Team Regionale Energiesysteme bedanken, für die spannenden Projekte und die schöne Zusammenarbeit. Insbesondere Kristina, Julian und Kathi standen mir jederzeit mit Rat und Tat zur Seite, wenn ich nicht weiterwusste, und viele ihrer Beiträge bilden die Grundlage meiner Arbeit.

Mein ganz besonderer Dank gilt allen Mitgliedern des Combi-Teams. Seit unserem ersten Treffen bei der INOC 2022 wurde ich vom gesamten Team unglaublich herzlich und freundlich aufgenommen. Vor allem Timo und Corinna, die mich von sich aus direkt in die Gruppe integriert haben und dafür gesorgt haben, dass ich mich dort wirklich willkommen fühle, bin ich dafür sehr dankbar. Insbesondere möchte ich mich auch ausdrücklich bei Sophia bedanken, die mit mir zusammen die letzten Wochen bis zur Abgabe durchgestanden hat, und die trotz Stress und knapper Zeit noch ein gemeinsames Projekt mit mir begonnen hat, das nun den Abschluss dieser Arbeit bildet.

Abschließend bedanke ich mich aufrichtig bei meiner Familie dafür, dass sie mich in all der Zeit nach Kräften unterstützt und in schwierigen Phasen immer wieder aufgebaut und ermutigt hat.

Danke euch allen für mehr Unterstützung, als ich mir je hätte wünschen können!

*David Neuroth*
*Aachen, Dezember 2025*

# Inhaltsverzeichnis

**Abstract** i

**Kurzfassung** iii

**Danksagung** v

**1 Einleitung** 1

1.1 Motivation . . . 1
1.2 Zielsetzung . . . 2
1.3 Struktur der Arbeit . . . 2

**2 Wissensstand** 3

2.1 Agentenbasierte Modellierung . . . 4
2.2 Verhaltenssimulation . . . 4
2.2.1 Arten von Verhaltensmodellen . . . 5
2.2.2 Modellkomplexität . . . 7
2.2.3 Daten für die Verhaltenssimulation . . . 8
2.2.4 Markov-Ketten-basierte Ansätze . . . 8
2.2.5 Andere Modellierungsansätze . . . 10
2.2.6 Vergleich der Verhaltensmodelle . . . 13
2.3 Städtesimulation . . . 17
2.3.1 Urban Building Energy Models . . . 17
2.3.2 Urban System Energy Models . . . 18
2.3.3 Transportsimulationen . . . 19
2.3.4 Landnutzungsmodelle . . . 21
2.3.5 Geosimulationen . . . 23
2.3.6 Vergleich der Modelle . . . 24

**3 Methodik** 27

3.1 Anforderungen . . . 27
3.2 Fallbeispiel Jülich . . . 28
3.3 Ansatz und Modell . . . 28
3.3.1 Szenarioerstellung . . . 29
3.3.2 Parallele Massensimulation . . . 32
3.3.3 Struktur der Städtesimulation . . . 34
3.3.4 Außer-Haus-Aktivitäten . . . 38
3.3.5 Simulation der Stadt . . . 42

3.4 Implementierung . . . 44
3.4.1 Implementierung des CityScenarioGenerators . . . 45
3.4.2 Erweiterung des LoadProfileGenerators . . . 51

**4 Auswertung der Städtesimulation 55**
4.1 Skalierbarkeit . . . 55
4.2 Szenariovalidierung . . . 57
4.3 Verhaltensvalidierung . . . 61
4.3.1 Das ETHOS.ActivityAssure-Framework . . . 61
4.3.2 Aggregierte Verhaltensvalidierung . . . 62
4.3.3 Verhaltensvalidierung nach Personen- und Tageskategorien . . . 68
4.4 Auswertung des Transports . . . 76
4.5 Validierung und Auswertung der Lastprofile . . . 79
4.5.1 Stromverbrauch der Stadt . . . 79
4.5.2 Gleichzeitigkeit der Lastprofile . . . 83
4.5.3 Geografische Auswertung der Last . . . 85
4.5.4 Auswirkungen verschiedener Elektroauto-Ladeleistungen . . . 88
4.6 Auslastung der Gesundheitsversorgung . . . 89
4.6.1 Aktueller Stand in Jülich . . . 89
4.6.2 Alternative Szenarien . . . 92

**5 Diskussion 99**

**6 Zusammenfassung und Ausblick 103**

**A Anhänge 105**
A.1 Einbinden der MPI.NET-Bibliothek . . . 105
A.2 Abfrage zu Nichtwohngebäuden . . . 105
A.3 Modell für optimale Apothekenstandorte . . . 106

**Abbildungsverzeichnis 110**

**Tabellenverzeichnis 111**

**Literaturverzeichnis 112**

# 1 Einleitung

Das Leben im 21. Jahrhundert spielt sich hauptsächlich in Städten ab. Bereits heute lebt mehr als die Hälfte der Weltbevölkerung in Städten, in Europa sogar über 75% [1]. Die Tendenz ist steigend, für das Jahr 2050 wird weltweit ein Urbanisierungsgrad von etwa 70% erwartet [2].

Viele dieser Menschen verbringen einen großen Teil ihres Alltags in Städten, sie wohnen und arbeiten dort, und eine Vielzahl weiterer sonstiger Aktivitäten, seien es Erledigungen oder Freizeitbeschäftigungen, findet mehrheitlich oder ausschließlich dort statt. Diese Bereiche sind dabei nicht klar getrennt, sondern vielmehr eng miteinander verwoben. Kombinierte Aktivitäten, wie auf dem Weg von der Arbeit nach Hause einkaufen zu gehen, oder in Abhängigkeit vom Verkehr unterschiedliche Transportmittel oder auch Aufbruchszeiten zu wählen sind nur einige Beispiele für die vielfältigen Abhängigkeiten und Einflüsse der verschiedenen Aspekte aufeinander. Menschen bewegen sich nicht nur passiv durch ihre Umwelt, sondern interagieren mit ihr und den anderen Personen darin. Wenn jemand in einem Geschäft besonders lange bedient wird, hat das unmittelbare Folgen für den Tagesablauf der anderen Kunden, die noch warten. Diese wechselseitige Beeinflussung macht Abläufe in Städten komplex. Eine Änderung in der Verfügbarkeit von öffentlichem Nahverkehr kann Auswirkungen auf Arbeitszeiten, ausgeübte Hobbies oder regelmäßig besuchte Orte haben.

So vielseitig wie menschliche Aktivitäten sind auch die Bereiche, die davon betroffen sind. Der Energieverbrauch im Wohnbereich etwa wird maßgeblich dadurch bestimmt, wann Personen zuhause sind, und hängt somit auch von der städtischen Struktur und typischen Tagesabläufen ab. Auch die Auslastung von Verkehrsnetzen und Besucherzahlen in Geschäften werden hierdurch beeinflusst.

Viele dieser genannten Thematiken sind Gegenstand aktueller Diskussionen und haben unmittelbaren Bezug zu wichtigen zukünftigen Entwicklungen. Im Rahmen der Energiewende müssen lokale Energiesysteme massive Änderungen durchlaufen, mit gravierenden Konsequenzen für die angeschlossenen Haushalte. Der Umstieg hin zu erneuerbaren Energien muss auch den Transportsektor miteinbeziehen. Für beide Bereiche ist es unabdingbar, die davon betroffenen Personen zu berücksichtigen, um eine optimale Lösung zu finden.

Weitere offene Probleme gibt es etwa im Bereich der Gesundheitsversorgung. Um ein Mindestmaß an Versorgung sicherzustellen, müssen ausreichend viele Einrichtungen mit genügend Kapazität vorhanden sein. Auch in dieser Frage spielen die Tagesabläufe der Stadtbewohner und ihre Bewegungsmuster in Abhängigkeit von existierenden Standorten eine entscheidende Rolle.

Eine Vielzahl relevanter und aktueller Probleme betrifft somit das Leben in Städten auf eine Weise, dass eine Betrachtung nur im Kontext der Stadt Sinn ergibt. Nur so können umfassende und praktikable Lösungen gefunden werden, die sich an den gegebenen Strukturen sowie Verhaltensweisen und Bedürfnissen der Menschen orientieren.

## 1.1 Motivation

Um die beschriebenen Probleme gesamtheitlicher angehen zu können, wird eine Möglichkeit benötigt, die relevanten Vorgänge in Städten integriert abbilden zu können. Es existieren zahlreiche

individuelle Modelle für eingeschränkte Bereiche wie Energieverbrauch, Transport oder das Verhalten von Bewohnern in Gebäuden. Für komplexere Fragestellungen reichen diese Modelle jedoch nicht aus, da sie die jeweils anderen Aspekte weitgehend außer Acht lassen. Insbesondere genügt auch keine simple Kombination dieser Modelle, da Wechselwirkungen der einzelnen Komponenten auf diese Weise nicht berücksichtigt werden können. Stattdessen ist ein Modell erforderlich, dass die jeweils relevanten Komponenten gemeinsam betrachtet.

## 1.2 Zielsetzung

Zur Unterstützung integrierter Untersuchungen von Städten soll eine Simulation entwickelt werden, die in der Lage ist, die für verschiedene Fragestellungen relevanten Bestandteile gemeinsam zu modellieren. Ausgangspunkt soll dabei das Verhalten der Bewohner sein. Aus den simulierten Aktivitäten werden Aufenthaltsorte, Energiebedarfe und Einflüsse auf Einrichtungen und Elemente der Stadtinfrastruktur abgeleitet. Da das menschliche Verhalten somit den Kern der gesamten Simulation darstellt, muss es entsprechend mit hoher Genauigkeit modelliert werden.

Sämtliche Komponenten müssen räumlich und zeitlich präzise erfasst werden, um gültige Aussagen für alle Person und jede berücksichtigte Einrichtung treffen zu können. Insbesondere gegenseitige Beeinflussung von Personen und Auswirkungen auf ihr Verhalten und auf die Stadt sind von Interesse und sollen soweit möglich detailliert abgebildet werden.

## 1.3 Struktur der Arbeit

In den folgenden Kapiteln wird die Entwicklung einer vollständigen, integrierten Städtesimulation beschrieben, die die genannten Anforderungen erfüllt. Zunächst wird in Kapitel 2 der aktuelle Stand im Bereich der Verhaltenssimulation und der Modellierung von Städten wiedergegeben, um die Grundlage für die Städtesimulation zu schaffen und zu zeigen, dass bisher kein Modell existiert, das die beschriebenen Probleme lösen kann. Darauf aufbauend werden in Kapitel 3 der Modellierungsansatz sowie die technischen Details der Umsetzung erklärt. Dabei wird auf den vollständigen Ablauf der Modellierung eingegangen, begonnen bei der Sammlung von Daten für die Definition vollständiger Städteszenarien, bis hin zur effizienten Ausführung des Simulationsprogramms auf geeigneten Rechnerarchitekturen. Die Funktionalität der Städtesimulation wird anschließend in Kapitel 4 anhand eines beispielhaften Szenarios demonstriert. Dazu wird die Stadt Jülich als Ganzes simuliert und in Bezug auf menschliches Verhalten, Energieverbräuche sowie Auslastung bestimmter Orte untersucht. Hierbei werden auch alternative Szenarien modelliert und in Bezug auf ihre Auswirkungen analysiert. Abschließend wird die entwickelte Simulation in Kapitel 5 bewertet und hinsichtlich ihrer möglichen Anwendungen evaluiert.

# 2 Wissensstand

Die Simulation menschlichen Verhaltens ist eine sehr komplexe Aufgabe, die in der Vergangenheit bereits auf vielfältige Art und Weise angegangen wurde. Dabei gibt es erhebliche Unterschiede im gewählten Vorgehen und in der betrachteten Fragestellung. Von der Simulation großer Menschenmassen bis zum Versuch, die Gedankengänge einer einzelnen Person abzubilden, wurden unterschiedlichste Themengebiete detailliert bearbeitet. In den vergangenen Jahrzehnten wurden auch bereits mehrere Ansätze zur Simulation von menschlichem Verhalten in Städten vorgestellt, mit unterschiedlichen Methoden und Absichten. Hierzu zählen beispielsweise alltägliche Aktivitäten, die Nutzung von Haushaltsgeräten und das Reisen mit verschiedenen Transportmitteln. Aufgrund der Komplexität menschlichen Verhaltens ist insbesondere die agentenbasierte Modellierung in diesem Bereich weit verbreitet. Daher kommt dieser Technik im Folgenden eine besondere Bedeutung zu.

Abbildung 2.1 gibt einen Überblick über die wichtigsten Schwerpunkte dieser Arbeit. In diesem Kapitel werden die drei relevanten Teilbereiche individuell vorgestellt, inklusive eventueller Überschneidungen. Zunächst werden in Abschnitt 2.1 die Grundlagen der agentenbasierten Modellierung erklärt. Anschließend werden in Abschnitt 2.2 allgemein die Simulation menschlichen Verhaltens und dazu verwendete Methodiken beschrieben. Zuletzt wird in Abschnitt 2.3 auf existierende Modelle und Plattformen für die Simulation von Städten eingegangen.

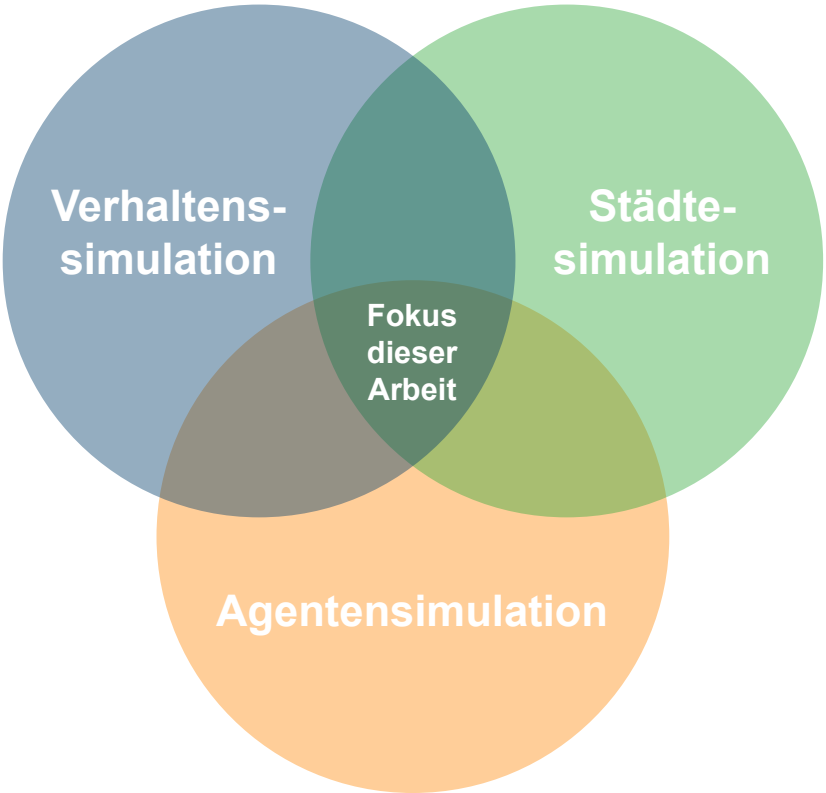



**Abbildung 2.1:** Überblick über die Schwerpunkte dieser Arbeit

## 2.1 Agentenbasierte Modellierung

Die agentenbasierte Modellierung bezeichnet eine Form von Bottom-Up-Modellen, bei denen die einzelnen Bestandteile eines Systems, die Agenten, individuell abgebildet werden. Was genau ein Agent darstellt, kann dabei sehr verschieden sein. So können beispielsweise Gebäude, Fahrzeuge, Unternehmen oder Menschen als Agenten fungieren. Wichtig ist lediglich, dass jeder Agent ein eigenständiges Objekt mit Eigenschaften und eigenem Verhalten bildet. Selbst einfaches Agentenverhalten kann dabei durch Interaktionen und Wechselwirkungen zu sehr komplexem Systemverhalten führen. Diese Eigenschaft wird als Emergenz bezeichnet und gilt als eine der wichtigsten Stärken von agentenbasierter Modellierung [3]. Des Weiteren lassen sich viele Systeme sehr intuitiv mit Agenten beschreiben, da sie natürlicherweise aus autonomen Einheiten bestehen, deren Verhalten oft leichter zu beschreiben ist als das des gesamten Systems. Ein wichtiger Bestandteil agentenbasierter Modellierung ist zudem die Umgebung, die die Agenten je nach Modell räumlich, zeitlich oder auf andere Weise einordnet.

Im Bereich der Simulation von menschlichem Verhalten werden bereits seit längerem Agentenmodelle eingesetzt. Die konkreten Umsetzungen variieren dabei von regelbasierten Ansätzen über probabilistische Modelle bis hin zu neuronalen Netzen [4]. Im Entscheidungsprozess der Agenten können mehrere Bereiche berücksichtigt werden. Neben kognitiven Überlegungen gehören dazu auch emotionale oder soziale Komponenten sowie eine Lernfähigkeit. Daraus wurden bereits einige zum Teil sehr ausführliche Entscheidungsarchitekturen entwickelt [5]. Aufgrund ihrer hohen Komplexität werden diese Architekturen im Gebäude- und Energiebereich bisher kaum eingesetzt, einige Modelle sind aber davon inspiriert.

## 2.2 Verhaltenssimulation

Menschliches Entscheidungsverhalten im Allgemeinen ist bereits seit langem Gegenstand wissenschaftlicher Fragestellungen. Auf der einen Seite ist es für etliche Prozesse auf verschiedenen räumlichen und zeitlichen Ebenen ausschlaggebend, etwa bei Verkehrsflüssen, Bewegungen von Menschenmassen oder der Verbreitung energieeffizienter Geräte. Auf der anderen Seite lässt es sich aufgrund der Vielzahl relevanter Einflussfaktoren grundsätzlich nur schwierig modellieren [6]. Daher wurden bereits viele verschiedene Ansätze entwickelt, die dieses Problem mit unterschiedlichen Methoden und Zielen angehen.

Allgemein kann Verhalten als "beobachtbare Aktion oder Reaktion einer Person auf externe oder interne Stimuli" [7] definiert werden. Modelle verwenden allerdings in der Regel spezifischere Definitionen von Verhalten und grenzen ein, was modelliert werden soll, um die Komplexität zu reduzieren. Beispielsweise wird häufig nur das Verhalten in einem bestimmten Kontext, wie zum Beispiel zuhause, im Büro oder unterwegs betrachtet. In dieser Arbeit steht energietechnisch relevantes Verhalten im Vordergrund, daher wird der Fokus auf Aktivitäten im Alltag gelegt, also die Nutzung von Geräten und Transportmitteln innerhalb einer Stadt. Dabei sind in erster Linie die Aufenthalte der Personen zuhause und in verschiedenen Nichtwohngebäuden sowie ihre Gerätenutzung von Interesse.

In verschiedenen Bereichen, die sich mit der Simulation menschlichen Verhaltens beschäftigen, wurden bereits etliche Modelle entwickelt. Im Folgenden werden zunächst allgemeine Eigenschaf-

ten und mögliche Kategorisierungen von Verhaltensmodellen beschrieben. Zudem werden die von diesen Modellen verwendeten Daten und die Bedeutung von Modellkomplexität bei der Verhaltenssimulation erklärt. Anschließend wird in Abschnitt 2.2.4 und Abschnitt 2.2.5 eine repräsentative Auswahl an Modellen vorgestellt, um einen Überblick über die verwendeten Methoden zu geben. Dazu wurden Modelle ausgesucht, die neue Konzepte implementieren oder bekannte Methoden neu kombinieren und so den aktuellen Stand der Technik in der Verhaltensmodellierung abbilden. Die Modelle werden dabei danach unterteilt, ob sie Markov-Ketten als grundlegende Kernkomponente einsetzen, oder ob sie einen anderen Ansatz verfolgen. Abbildung 2.2 gibt einen Überblick über die ausgewählten Verhaltensmodelle.

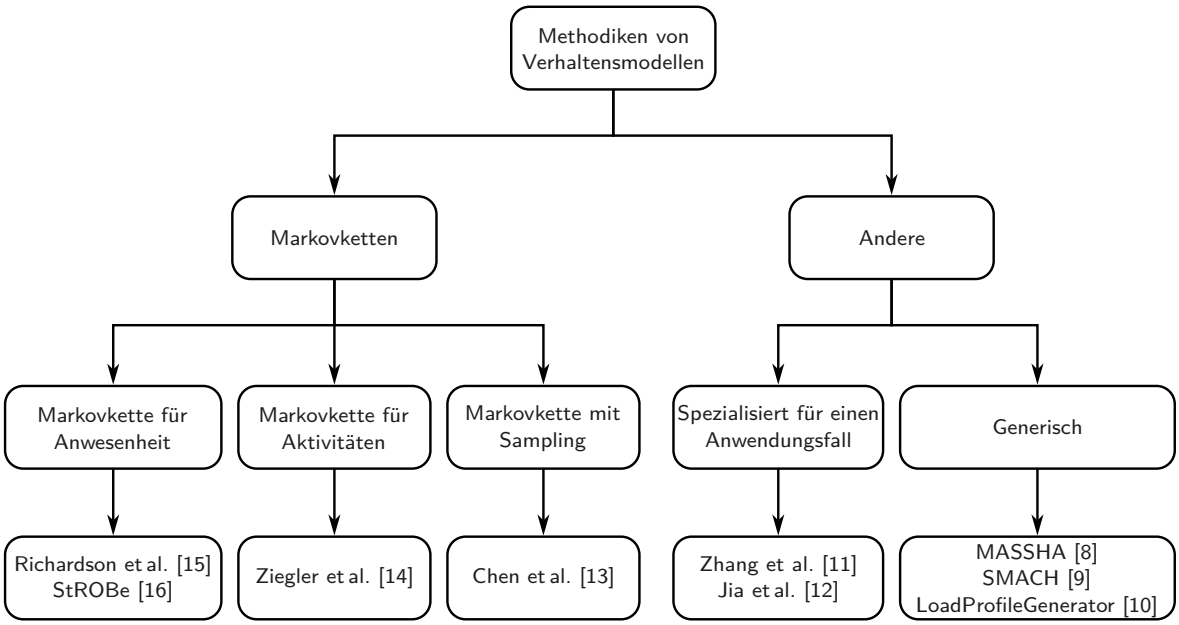



**Abbildung 2.2:** Die im Folgenden vorgestellten Verhaltensmodelle, sortiert nach ihrer Methodik

### 2.2.1 Arten von Verhaltensmodellen

Aufgrund der vielen betroffenen Anwendungsgebiete und der unterschiedlichen Anforderungen lassen sich Kategorien von Verhaltensmodellen nicht immer eindeutig voneinander abgrenzen. Stattdessen existieren verschiedene Arten von Modellen mit unterschiedlichen Eigenschaften, die teilweise ineinander übergehen.

Eine Art der Verhaltensmodellierung, die sehr aktiv behandelt wird, befasst sich mit *Occupant Behavior*, also dem Verhalten von Bewohnern. Damit wird vorrangig wärme-relevantes Verhalten in Gebäuden bezeichnet. Das schließt sämtliche Interaktionen mit der Gebäudehülle ein, wie beispielsweise das Öffnen von Fenstern, das Einstellen von Jalousien oder Thermostaten, aber auch die Nutzung elektrischer Geräte. Diese Formen von Verhalten sind typischerweise für Gebäudesimulationen relevant, die Wärmeabgaben in Gebäuden anhand physikalischer Modelle berechnen, um den jährlichen Heiz- und Kühlbedarf zu ermitteln. Es gibt mehrere Faktoren, die sich auf das Verhalten von Bewohnern auswirken können. Die Einstellung von Heizung, Klimatisierung oder beweglichen Teilen der Gebäudehülle gehört zum sogenannten adaptiven Verhalten. Dieses dient der Aufrechterhaltung oder Wiederherstellung des persönlichen Komforts und wird unter anderem durch die Raumtemperatur und die $CO_2$-Konzentration beeinflusst [17]. Nicht-adaptives

Verhalten, wie die Nutzung von Haushaltsgeräten, hängt dagegen stärker von der Uhrzeit und persönlichen Verhaltensmustern ab [18]. Im Bereich der Occupant-Behavior-Modelle wird auch der Frage nachgegangen, wie das Verhalten von Bewohnern mithilfe von Sensordaten ermittelt werden kann. Hierbei wird beispielsweise versucht, anhand der Daten von vorhandenen oder eigens angebrachten Sensoren auf die Anwesenheit und Aktivität von Personen in einem Raum zu schließen [19]. Eine derartige Datenerfassung ist allerdings sehr aufwendig und in großem Maßstab insbesondere in Bezug auf Datenschutz kaum realisierbar, daher kommt dieser Typ an Occupant-Behavior-Modellen für diese Arbeit nicht in Betracht.

Occupant-Behavior-Modelle lassen sich auf verschiedene Weise klassifizieren. Grundsätzlich kann zwischen deterministischen und stochastischen Ansätzen unterschieden werden [20]. Deterministische Modelle nutzen fest vorgegebene Zeitpläne für nicht-adaptive Aktivitäten und Regeln für adaptives Verhalten, abhängig vom Zustand der simulierten Umgebung, etwa der Raumtemperatur. Die Regeln können dabei manuell definiert oder beispielsweise über Regression bestimmt werden. Stochastische Modelle basieren dagegen auf Wahrscheinlichkeitsverteilungen und ermitteln so je nach Uhrzeit, vorheriger Beschäftigung und gegebenenfalls weiteren Faktoren mögliche Folgeaktivitäten. Eine verbreitete Methode hierfür sind Markov-Ketten verschiedener Art, häufig jedoch homogene oder inhomogene erster Ordnung [21]. Diese Verfahren werden in Abschnitt 2.2.4 genauer erklärt.

Zur Berechnung der notwendigen Koeffizienten werden passende Daten benötigt. Der hierbei üblicherweise verwendete Typ an Daten wird in Abschnitt 2.2.3 beschrieben.

Balvedi, Ghisi und Lamberts [21] und Li et al. [22] wählen eine vergleichbare Aufteilung und bezeichnen deterministische Modelle als Modelle mit festem Zeitplan. Zusätzlich werden hier datenbasierte und agentenbasierte Modelle als separate Kategorien betrachtet. Datenbasierte Modelle nutzen Techniken des maschinellen Lernens, wie Entscheidungsbäume oder Support Vector Machines, zur Bestimmung des Verhaltens anhand von Umgebungsparametern. Teilweise wird dies auch mit vorhergehendem Clustering zur Identifikation üblicher Zustände und Verhaltensmuster kombiniert. Dafür müssen diese Modelle mit großen Mengen passender Daten trainiert werden, was ihre Anwendbarkeit erschwert. Außerdem wird ihre Präzision und ihr mögliches Anwendungsgebiet durch Qualität und Art der Trainingsdaten bestimmt [23]. Agentenbasierte Modelle verfolgen einen feingranularen Ansatz, basierend auf der individuellen Simulation aller Bewohner mitsamt ihrem Entscheidungsverhalten. Li et al. [22] heben hervor, dass diese Kategorisierungen nicht immer eindeutig sind und es Überschneidungen geben kann, beispielsweise zwischen agentenbasierten und probabilistischen Ansätzen. Eine präzise Abbildung von menschlichem Verhalten ist in vielen Fällen essenziell, um gute Ergebnisse zu erhalten. Bei Gebäude- und Städtesimulationen gilt eine unzureichende Berücksichtigung vom Bewohnerverhalten als einer der Hauptgründe für die sogenannte Performance Gap, also die Unterschiede zwischen simulierten und tatsächlichen Energieverbräuchen [20, 22, 23].

Weitere Ansätze zur Verhaltenssimulation neben Occupant-Behavior-Modellen finden sich in Lastprofilgeneratoren oder im Bereich der Human Activity Recognition. Erstere versuchen, für definierte Gebäude oder Haushalte Lastprofile verschiedener Typen zu generieren. Dazu muss nicht notwendigerweise das Verhalten der Bewohner simuliert werden, Top-Down-Ansätze funktionieren beispielsweise über die Zerlegung von gemessenen Lastprofilen auf höherer Ebene. Bottom-Up-

Lastprofilgeneratoren modellieren hingegen meist, ähnlich wie Occupant-Behavior-Modelle, die Aktivitäten der Bewohner, um über damit zusammenhängende Geräteaktivierungen auf entstehende Lasten schließen zu können [24]. Da in dieser Arbeit auf die Generierung von Aktivitätsprofilen abgezielt wird, werden hier lediglich Bottom-Up-Lastprofilgeneratoren betrachtet.

Das Forschungsfeld Human Activity Recognition beschäftigt sich mit der Erkennung menschlichen Verhaltens in sogenannten intelligenten Umgebungen, also Gebäuden oder Arealen, die zu diesem Zweck mit entsprechenden Sensoren ausgestattet sind. Dabei werden unter anderem verschiedene Sensorkonzepte erprobt. Wie zuvor bereits erwähnt, sind experimentelle Datensammlungen in derartigen Szenarien aufwendig. Für intelligente Umgebungen gilt dies umso mehr, da nicht beliebig viele solcher Umgebungen zur Verfügung stehen, und Modifikationen an existierenden Umgebungen in der Regel mit hohem Aufwand verbunden sind [25, 8]. Daher wird auch hier auf Modelle zurückgegriffen, die definierte Umgebungen mitsamt den Menschen darin abbilden können. Neben den eigentlichen Aktivitäten der Menschen wird dabei vor allem Wert auf die simulierte Auslösung von Sensoren gelegt.

### 2.2.2 Modellkomplexität

Existierende Verhaltensmodelle unterscheiden sich stark in Komplexität, Auflösung und Detailgrad. Das hängt einerseits mit der Weiterentwicklung von Modellen und Methoden zusammen, andererseits aber auch mit unterschiedlichen Anforderungen durch die jeweiligen Verwendungszwecke. Bei neueren Modellen ist laut Gaetani, Hoes und Hensen [23] eine Tendenz zu höherer Komplexität erkennbar. Diese Entwicklung wird ambivalent betrachtet, da eine höhere Modellkomplexität zu erhöhten Abweichungen aufgrund falscher Annahmen führen kann. Ziel ist, ein Modell zu finden, dass komplex genug ist, um alle relevanten Sachverhalte ausreichend präzise darzustellen, ohne unnötige Komplexität einzuführen. Um die Wahl einer geeigneten Modellkomplexität zu erleichtern, haben Malik et al. [6] ein Level-of-Detail-Framework für Occupant-Behavior-Modelle entwickelt. Der Fokus wurde aufgrund ihrer allgemein hohen Komplexität und ihrer aktuellen Beliebtheit im Occupant-Behavior-Bereich auf agentenbasierte Modelle gelegt. Das Framework betrachtet zehn grundlegende Eigenschaften von Occupant-Behavior-Modellen und kategorisiert Methoden zur Lösung verschiedener Teilprobleme. Anhand dessen werden vier Detailstufen gebildet, die sich beispielsweise darin unterscheiden, wie Agenten dargestellt werden, ob sie lernfähig sind oder ob sie mit ihrer Umgebung oder anderen Agenten interagieren können. Stufe O-0 stellt nur durchschnittliches statisches Bewohnerverhalten dar und ist für aggregierte Betrachtungsweisen mit niedriger Auflösung geeignet. Bei den Stufen O-1 und O-2 steigt die Komplexität schrittweise weiter an, unter anderem durch die Modellierung dynamischen Verhaltens und unterschiedlicher Bewohnergruppen. Stufe O-3 berücksichtigt schließlich individuelles Agentenverhalten inklusive Interaktionen in hoher räumlicher und zeitlicher Auflösung. Allerdings sind auch hybride Modelle möglich, die je nach Bedarf nur einzelne Aspekte detaillierter oder rudimentärer betrachten und somit zwischen den definierten Stufen liegen. Das ist notwendig, da die Wahl eines geeigneten Detailgrades sehr kontextabhängig ist [6]. Das Framework stellt daher auch keinen eindeutigen Algorithmus zur Festlegung des Detailniveaus dar, sondern eine Orientierungshilfe.

### 2.2.3 Daten für die Verhaltenssimulation

Geeignete Daten spielen bei vielen Verhaltensmodellen eine zentrale Rolle. Datenbasierte Modelle sind in erster Linie von einem guten Datensatz abhängig, um Zusammenhänge zwischen Umgebungsparametern und Verhaltensmustern erlernen und wiedergeben zu können. Darüber hinaus betrifft die Datenqualität insbesondere probabilistische Modelle. Diese basieren auf der Generierung von Aktivitäten anhand geeigneter stochastischer Verteilungen. Um diese Verteilungen aufstellen zu können, werden Daten darüber, wann welche Aktivitäten ausgeübt werden, benötigt. Hierfür werden in der Regel Zeitverwendungsdaten herangezogen [26]. Dabei handelt es sich um Umfragen, bei denen eine ausgewählte Personengruppe über einen festgelegten Zeitraum von meist wenigen Tagen in festen Zeitintervallen von beispielsweise 10 Minuten ihre Tätigkeiten dokumentiert. Die entstehenden Tätigkeitsprofile werden anschließend aufbereitet, um etwa für jede Aktivität einen einheitlichen Namen zu verwenden. Meist werden die zu dokumentierenden Tage auf Wochentage und Wochenenden aufgeteilt, um repräsentative Aktivitätsprofile mit und ohne Arbeit zu erhalten. Damit die Profile statistisch analysiert und gruppiert werden können, beinhalten die Umfragen außerdem wichtige persönliche Merkmale und Angaben zur sozioökonomischen Situation der Befragten. Dazu gehören beispielsweise Alter, Geschlecht, Beruf und familiäre Situation. Zusätzlich können in einigen Zeitverwendungsumfragen auch sekundäre oder gemeinsam ausgeführte Aktivitäten erfasst werden.

Die so gesammelten Daten bieten eine gute Basis für die Modellierung menschlichen Verhaltens. Herausforderungen liegen insbesondere in mangelnden Kenntnissen über äußere Einflüsse sowie in der oft geringen Menge an Tagesprofilen pro Person, wodurch eine Einteilung in regelmäßige und außergewöhnliche Aktivitäten erschwert wird [9].

### 2.2.4 Markov-Ketten-basierte Ansätze

Eine gängige Methode zur Simulation menschlichen Verhaltens stützt sich auf Markov-Ketten. Dabei wird die Abfolge an Aktivitäten als stochastischer Prozess modelliert und es wird versucht, auf Basis der bisherigen Zustände auf den jeweils nächsten zu schließen. Die Ordnung der Markov-Kette bestimmt, wie viele der letzten Zustände bei der Auswahl des nächsten berücksichtigt werden, bei einer Markov-Kette erster Ordnung fließt also nur der aktuelle Zustand in die Entscheidung ein. Falls diese bedingten Wahrscheinlichkeiten in jedem Zeitschritt konstant sind, handelt es sich um eine homogene Markov-Kette, andernfalls um eine inhomogene Markov-Kette. Die konkreten Wahrscheinlichkeiten werden in der Regel anhand eines geeigneten Datensatzes ermittelt. Im Folgenden werden einige Modelle vorgestellt, die auf verschiedene Art und Weise Markov-Ketten einsetzen.

#### Verhaltensmodell von Richardson

Ein häufig zitierter und weiterverwendeter Ansatz zu Verhaltenssimulation stammt von Richardson et al. [15]. Die Grundlage dafür bildet ein reines Occupancy-Modell, also ein Modell, das lediglich die Anwesenheit der Bewohner simuliert [27]. Als Datenbasis wurde eine britische Zeitverwendungsumfrage aus dem Jahr 2000 in 10-Minuten-Auflösung genutzt. Anhand dieser Daten wurden Wahrscheinlichkeiten für eine Änderung der Anzahl anwesender Personen in einem Haushalt berechnet, modelliert als eine Markov-Kette. Diese Wahrscheinlichkeiten sind lediglich vom

vorherigen Zustand und von der Uhrzeit abhängig, es handelt sich also um eine inhomogene Markov-Kette erster Ordnung. Somit entsteht für jeden Zeitschritt eine Übergangsmatrix, anhand derer stochastisch der nächste Zustand berechnet werden kann. Für Wochentage und Wochenenden wurden jeweils getrennte Matrizen erzeugt, um das unterschiedliche Verhalten insbesondere berufstätiger Leute wiedergeben zu können. Anwesenheitsprofile werden über einen Monte-Carlo-Ansatz generiert. Ausgehend von einem Startzustand wird dabei der jeweils nächste Folgezustand mithilfe der Übergangsmatrizen probabilistisch ermittelt. Diese Methodik wird als Markov-Ketten-Monte-Carlo-Technik bezeichnet.

Das eigentliche Verhaltensmodell baut hierauf auf. Jedem Haushalt wird eine vorgegebene Menge elektrischer Geräte zugewiesen. Jedes dieser Geräte wird mit einer Aktivität verknüpft, für die aus demselben Zeitverwendungsdatensatz ein Wahrscheinlichkeitsprofil ermittelt wird. Dieses Profil gibt für jeden Zeitschritt die Nutzungswahrscheinlichkeit des Geräts an, in Abhängigkeit von der Anzahl anwesender Bewohner. Auf diese Weise werden sowohl gleichzeitig genutzte Geräte als auch die geteilte Nutzung durch mehrere Bewohner berücksichtigt.

#### StROBe von Baetens

Baetens und Saelens [16] nutzen Teile des Modells von Richardson et al. [15] für ihr Occupant-Behavior-Modell StROBe, eine Abkürzung für Stochastic Residential Occupant Behavior. In StROBe werden belgische Zeitverwendungsdaten eingesetzt, sowie eine weitere Umfrage zur Ausstattung der Haushalte. Letztere wird genutzt, um Wahrscheinlichkeiten für den Besitz von Geräten zu berechnen, anhand derer die Geräteinventare von Haushalten stochastisch bestimmt werden können. Aus den Zeitverwendungsdaten werden auch hier Übergangsmatrizen für eine inhomogene Markov-Kette erster Ordnung erstellt, allerdings werden die Daten vorher mittels eines Clustering-Verfahrens aufgeteilt. So ergeben sich verschiedene Bewohnertypen mit unterschiedlichen Anwesenheitsmustern, die bei der Zusammenstellung von Haushalten berücksichtigt werden können. Für jedes Mitglied eines Haushalts wird dabei, abhängig vom Arbeitsstatus, zufällig einer der Bewohnertypen aus den Clustern ausgewählt. Für die Aktivitätsmodellierung wird auf die in Abschnitt 2.2.4 beschriebene Methode von Richardson et al. [15] zurückgegriffen.

#### Modell von Ziegler

Ziegler et al. [14] nutzen für ihren Lastprofilgenerator ebenfalls inhomogene Markov-Ketten erster Ordnung. Anders als bei den zuvor beschriebenen Modellen werden hierbei die Aktivitäten direkt durch die Markov-Kette bestimmt, es wird also kein Umweg mehr über die Simulation der Anwesenheit gemacht. Dabei werden auch wieder mehrere Markov-Ketten eingesetzt. Wie schon bei vorherigen Ansätzen werden die Zeitverwendungsdaten, in diesem Fall aus Deutschland und von 2012, nach sozioökonomischen Kriterien geclustert. Die bekannte Aufteilung in Wochentage und Wochenenden wird leicht angepasst, um abweichenden Verhaltensmustern an Freitagen gerecht zu werden. Zusätzlich wird eine saisonale Aufteilung in Sommer, Winter und Übergangsmonate vorgenommen.

### Hybrider Ansatz von Chen

Wie an den bereits vorgestellten Ansätzen zu erkennen ist, sind Markov-Ketten im Bereich der probabilistischen Occupant-Behavior-Modelle sehr beliebt und werden auf unterschiedliche Weise eingesetzt. Daher haben Chen et al. [13] diese Methode mit dem Sampling, einer anderen beliebten Methode in der Verhaltenssimulation, verglichen. Sampling basiert auf dem Aufstellen stochastischer Verteilungen zu Häufigkeit, Dauer und Uhrzeit einer Aktivität, in der Regel auch aus Zeitverwendungsdaten. Bei diesem Verfahren wird für jeden Aktivitätstypen zunächst die Anzahl auszuübender Aktivitäten für den nächsten Tag stochastisch ermittelt. Anschließend wird für jede Ausführung der Aktivität eine Uhrzeit und Dauer bestimmt. Dieses Verfahren kann beispielsweise mit einer vorgelagerten Markov-Kette zur Ermittlung des Anwesenheitsstatus kombiniert werden.

Im Vergleich zeigte der Sampling-Ansatz zwar eine gute Übereinstimmung bezüglich der Aktivitätshäufigkeiten und -dauern, allerdings große Abweichungen bei den typischen Uhrzeiten. Ein reiner Markov-Ketten-Ansatz, bei dem in jedem Zeitschritt eine Übergangsmatrix zur Vorhersage des Zustands im darauffolgenden Zeitschritt eingesetzt wurde, wies eine allgemein bessere Übereinstimmung mit den Zeitverwendungsdaten auf, allerdings eine gewisse Ungenauigkeit im Hinblick auf die Aktivitätsdauern [13]. Für ihr eigenes Modell wählten Chen et al. daher einen hybriden Ansatz. Dabei wird zur Bestimmung der jeweils nächsten Aktivität eine Markov-Kette verwendet, die Dauer der Aktivität wird aber über eine entsprechende Wahrscheinlichkeitsfunktion festgelegt. Die damit erzeugten Aktivitätsprofile passen erkennbar besser zu den Eingangsdaten.

## 2.2.5 Andere Modellierungsansätze

Neben den typischen Markov-Ketten-basierten Modellen gibt es viele weitere, die eigene, oftmals sehr individuelle Methoden für die Verhaltenssimulation nutzen. Einige davon sind sehr stark an einen bestimmten Anwendungsfall angepasst und lassen sich nur schwer generalisieren, andere sind dagegen weitgehend generisch gestaltet. Im Folgenden werden Beispiele für einige solcher Modelle mit verschiedenen ausgeklügelten Mechanismen zur Aktivitätssimulation beschrieben.

### MASSHA von Kamara-Esteban

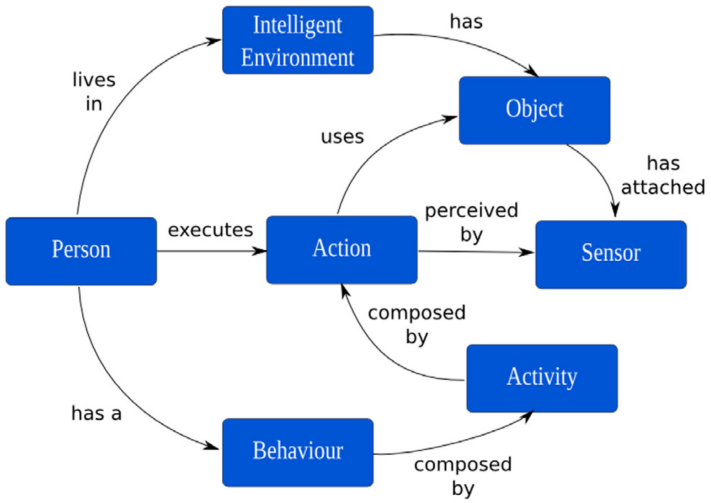



**Abbildung 2.3:** Struktur des Verhaltensmodells in MASSHA, von Kamara-Esteban et al. [8]

MASSHA [8] ist ein Modell aus dem Bereich Human Activity Recognition. Ziel dieses Modells ist es, eine intelligente, mit Sensoren versehene Umgebung nachzubilden und die Interaktionen von Menschen mit Sensoren zu simulieren. Das Modell arbeitet nach der in Abbildung 2.3 gezeigten Struktur. Relevant ist hiervon vor allem die hierarchische Aufteilung in Verhalten, Aktivitäten und Aktionen. Eine Aktion bezeichnet eine einzelne, spezifische Handlung, die von einem Sensor erfasst wird, beispielsweise das Öffnen eines Schranks. Jede Aktivität stellt eine zusammengehörige Handlungsfolge dar und setzt sich aus einer oder mehreren Aktionen zusammen, die parallel oder sequenziell ausgeführt werden. Das Verhalten bezeichnet die Gesamtheit aller Aktivitäten, die eine Person ausführt oder ausführen möchte. Als Eingabe erwartet MASSHA eine Liste zu erledigender Aktivitäten für jeden Bewohner. Diese enthält eine Reihe an Aktivitäten, die der Bewohner im Verlauf eines Tages ausführen soll, und gibt gegebenenfalls weitere Eigenschaften an, wie eine Priorität, abhängige Aktivitäten oder eine verpflichtende Startzeit. Während der Simulation versucht nun jeder Bewohner, möglichst viele seiner Aktivitäten zu absolvieren. Beendete Aktivitäten werden anhand einer Liste nachverfolgt, um Abhängigkeiten einhalten zu können, beispielsweise dass erst gegessen werden kann, nachdem gekocht wurde.

#### Modell von Zhang

Zhang, Siebers und Aickelin [11] haben ein sehr spezifisches, agentenbasiertes Verhaltensmodell entwickelt. Das mit der Simulationssoftware AnyLogic erstellte Modell bildet ein reales Bürogebäude inklusive aller Räume und der darin arbeitenden Personen akkurat nach. Der Fokus liegt auf flexiblen elektrischen Lasten durch Beleuchtung und die Nutzung von Computern. Licht und Computer wurden als passive Agenten realisiert, die lediglich auf Auslöser reagieren und den entsprechenden Strombedarf speichern. Mitarbeitende werden als aktive Agenten simuliert, die größtenteils in ihren Büros arbeiten und diese gelegentlich für kurze Zeit verlassen, etwa um einen anderen Raum aufzusuchen. Das Verhalten wird probabilistisch simuliert, mit festgelegten Übergangswahrscheinlichkeiten, abhängig vom aktuellen Zustand des Agenten. Die notwendigen Daten zur Berechnung dieser Wahrscheinlichkeiten sowie anderer relevanter Eigenschaften wurden über eine Umfrage unter allen Mitarbeitenden gesammelt. Auf diese Weise wurden auch Informationen dazu gewonnen, wie häufig Agenten ihre Computer eingeschaltet lassen oder ausschalten, wenn sie den Raum verlassen. Agenten können zudem mit anderen Agenten kommunizieren und sie zum Beispiel zu einem sparsameren Verhalten animieren. Dadurch konnten die Auswirkungen sozialer Netzwerke und Kontakte innerhalb des Büros auf das Verhalten und damit auf den Strombedarf untersucht werden. Allerdings ist das Modell durch seine Struktur und die Umfragedaten stark maßgeschneidert und in dieser Form nicht ohne Weiteres auf andere Szenarien anwendbar.

#### Modell von Jia

Beim agentenbasierten Ansatz von Jia et al. [12] wird die Software „Performance Moderator Function Server“ eingesetzt. Dabei handelt es sich um ein auf verschiedenen Submodellen basierendes Hilfsmittel zur Abbildung und Erklärung von menschlichem Verhalten. Darin gibt es drei grundlegende Aspekte, die das Verhalten einer Person beeinflussen. Dazu gehören die Wahrnehmung der Umgebung und möglicher Aktionen, ein Wertesystem für kurz- und langfristige Ziele und ein physiologisches System, das unter anderem das aktuelle Stressniveau speichert. In jedem Zeitschritt

werden vom Agenten wahrgenommene mögliche Aktionen regelbasiert gegeneinander abgewogen und diejenige, die dem Wertesystem zufolge am nützlichsten ist, wird ausgeführt. Auf diese Weise können Wechselwirkungen zwischen der Umgebung und dem Verhalten der Agenten abgebildet werden. Ähnlich wie beim Modell von Zhang, Siebers und Aickelin [11] in Abschnitt 2.2.5 wurde das Modell auf ein bestimmtes Bürogebäude abgestimmt. Grundlegende Parameter wie mögliche Interaktionen wurden in einer Befragung der Mitarbeitenden ermittelt. Simuliert wurde das Öffnen von Fenstern, Türen und Jalousien. Zu den wahrnehmbaren Umgebungsparametern gehören die Raumtemperatur, $CO_2$-Konzentration, Luftfeuchtigkeit und Helligkeit.

#### SMACH von Reynaud

In SMACH ("Simulation of Activity and Consumption in the Home") [9], einer Plattform für Verhaltenssimulation, wird agentenbasierte Modellierung mit dem Einsatz von Zeitverwendungsdaten kombiniert. Jeder Agent erhält zunächst einen eigenen Plan an Aktivitäten für den nächsten Tag und die nächste Woche. Diese Pläne werden aus Zeitverwendungsdaten generiert, aufgeteilt auf 27 Agententypen und zwei Tagestypen. Um regelmäßige Aktivitäten realistisch abbilden zu können, aber auch Freiraum für individuelle Abweichungen zu bieten, treffen die Autoren die vereinfachende Annahme, dass makroskopisch seltenes Verhalten auch auf mikroskopischer Ebene selten ist. Somit werden bei der Aufstellung der Pläne nur Aktivitäten und deren Zeitspannen berücksichtigt, die im gesamten Zeitverwendungsdatensatz häufig auftreten. Anschließend können die Pläne verändert werden, beispielsweise indem die Dauer einer Aktivität einmalig erhöht oder reduziert wird. Die so erstellten Pläne dienen jedoch lediglich als Vorgabe, wann die Ausführung einer Aktivität besonders wahrscheinlich ist. Aus dem Zeitverwendungsdatensatz werden alle ausführbaren Aktivitäten extrahiert und mit einer Dauer, einer Priorität, einem Ort und einem Gemeinschaftsattribut versehen. Über die Priorität wird nun in der Simulation gesteuert, wann ein Agent welche Aktivitäten ausführt. Während der im Plan vorgegebenen Zeiten für eine Aktivität wird deren Priorität stark erhöht, außerhalb davon wird sie verringert. Das Gemeinschaftsattribut sagt aus, ob eine Tätigkeit häufig gemeinsam mit anderen Haushaltsmitgliedern ausgeübt wird. Ist dieser Wert gesetzt und ein Agent führt eine solche Aktivität aus, so wird die Priorität für diese Aktivität bei allen weiteren Haushaltsmitgliedern ebenfalls erhöht. Auf diese Weise können neben gemeinsamen Beschäftigungen auch weitere Abhängigkeiten und Bedingungen realisiert werden, beispielsweise dass Geräte, die momentan von Anderen genutzt werden, vorübergehend nicht zur Verfügung stehen, oder andere Einflüsse aus der Umgebung wie variable Strompreise [28]. Dadurch ist dieser Ansatz im Vergleich zu vielen anderen Modellen sehr dynamisch und flexibel.

#### LoadProfileGenerator von Pflugradt

Im LoadProfileGenerator [10] wird ein gänzlich anderer agentenbasierter Ansatz verfolgt. Die Grundlage bildet ein Bedürfnismodell, das für die wesentlichen Entscheidungen der Agenten verantwortlich ist. Jeder Agent wird mit vielen verschiedenen Bedürfnissen versehen, darunter die Bedürfnisse zu arbeiten, zu essen oder zu schlafen, alle mit einer eigenen Gewichtung. Diese Bedürfnisse werden mit der Zeit stärker, und Agenten wählen ihre nächste Aktivität stets so, dass die gewichtete Summe der Bedürfniswerte optimiert wird. Abweichungen des Bedürfnisses vom Idealwert werden dabei in beide Richtungen negativ gewertet. Durch empirische Konfiguration der

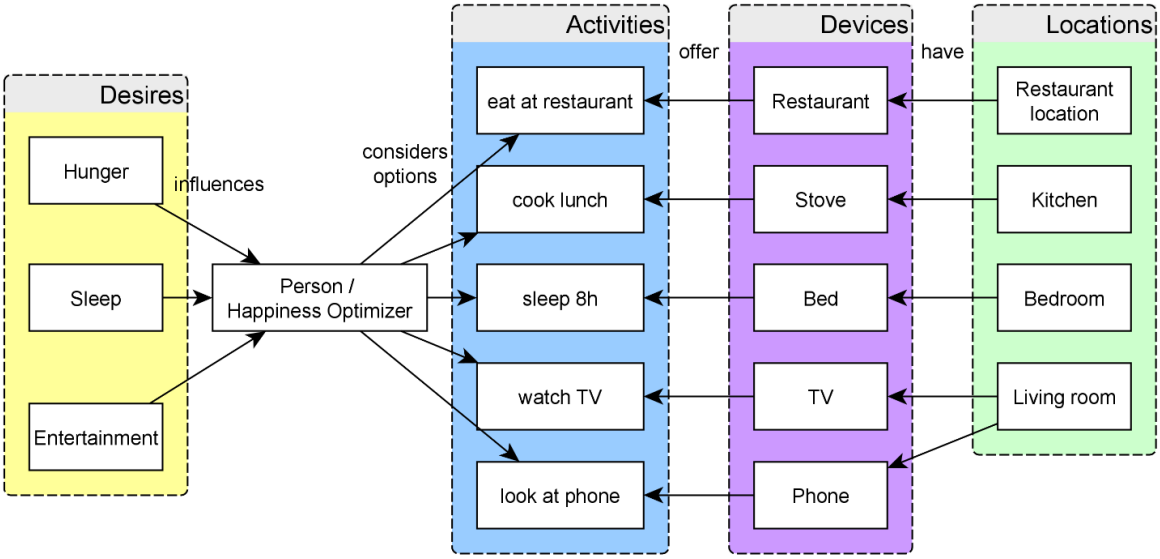



**Abbildung 2.4:** Überblick über den Kern des Verhaltensmodells im LoadProfileGenerator, aus [10]

Bedürfnisparameter können so plausible Häufigkeiten für alle alltäglichen Tätigkeiten erreicht werden. Um die Simulation realistischer zu gestalten, werden regelbasierte Einschränkungen getroffen, beispielsweise im Hinblick auf erlaubte Ausführungszeiten oder -bedingungen für eine Aktivität. So wird auch festgelegt, welche Aktivitäten unterbrochen oder mit anderen Agenten gemeinsam ausgeübt werden können.

Wie Abbildung 2.4 zeigt, ist jede Aktivität mit einem Gerät an einem bestimmten Ort verknüpft. So können die Aufenthaltsorte der Agenten sowie entstehende Energie- und Ressourcenverbräuche nachverfolgt und damit schließlich Lastprofile berechnet werden. Dieses Verhaltensmodell ist sehr vielseitig und kommt vollkommen ohne Zeitverwendungsdaten aus, lediglich für die Geräte werden gemessene oder synthetische Lastkurven eingesetzt. Allerdings sind die Balancierung der Bedürfnisse und eine passende Einstellung aller Parameter für ein realistisches Verhalten kompliziert. Um plausible Simulationsergebnisse zu gewährleisten, wurden anhand statistischer Daten daher mehrere vordefinierte und validierte Modellhaushalte konfiguriert, die die Mehrheit der deutschen Haushalte ausreichend genau abdecken.

### 2.2.6 Vergleich der Verhaltensmodelle

Die gezeigten Modelle geben einen Überblick über die verschiedenen Ansätze zur Verhaltenssimulation. Jeder Modelltyp bringt dabei Vor- und Nachteile mit sich.

Deterministische oder statistische Ansätze sind verhältnismäßig unflexibel. Ihre Genauigkeit hängt stark von der Identifikation passender Einflussfaktoren aus der Umgebung ab, Gewohnheiten oder psychologische Faktoren können nicht gut abgebildet werden [22]. Insbesondere statische Verfahren ohne Berücksichtigung äußerer Einflüsse sind zudem nur in der Lage, beobachtetes Verhalten zu reproduzieren und können keine neuen Muster generieren [21].

Im Gegensatz dazu sind stochastische Modelle imstande, neue Verhaltensmuster zu erzeugen. Reine Markov-Ketten- oder Sampling-Ansätze haben, wie in Abschnitt 2.2.4 erklärt, jeweils be-

stimmte Schwächen, aber Chen et al. [13] konnten demonstrieren, dass ein hybrider Ansatz diese umgehen kann. Bei diesem Modelltyp kann insbesondere die Vielfältigkeit des Verhaltens leichter berücksichtigt werden. Durch die probabilistischen Elemente entstehen selbst bei mehrfacher Simulation desselben Haushalts unterschiedliche Aktivitätsprofile. Diese Variabilität erschwert jedoch gleichzeitig die Generierung konsistenten Verhaltens. Im aggregierten Maßstab, also bei wiederholter Ausführung eines stochastischen Modells für einen einzelnen oder mehrere Haushalte, ergeben sich häufig gut passende Verhaltensverteilungen, das gilt aber nicht unbedingt auch für individuelle Aktivitätsprofile [14]. Einer der relevantesten Nachteile von stochastischen Modellen besteht darin, dass sie viele Abhängigkeiten im menschlichen Verhalten nicht darstellen können. So wird der Anwesenheitsstatus oft getrennt von der Aktivität bestimmt [29]. Insbesondere die Einflüsse mehrerer Bewohner aufeinander, etwa durch gemeinsame Aktivitäten oder die Vermeidung einer Doppelnutzung von Geräten, werden häufig nicht oder nur in ausgewählten Einzelfällen berücksichtigt [21].

Datenbasierte Modelle können zwar die Zusammenhänge zwischen Umgebungsparametern und adaptivem Verhalten gut abbilden, benötigen im Vergleich zu anderen Modelltypen aber große Datenmengen [12]. Die zentrale Rolle des Datensatzes sorgt auch dafür, dass Modelle dieses Typs nicht in anderen Kontexten eingesetzt werden können, die von den Daten nicht abgedeckt werden [21].

Als letzter Modelltyp bleiben die agentenbasierten Ansätze, die im Bereich der Verhaltensmodelle in den letzten Jahren an Popularität gewonnen haben [30]. Diese Bezeichnung ist sehr generisch und kann viele Arten von spezifischen Modellen umfassen, die verschiedenen Ansätze haben jedoch gemein, dass sich durch die grundlegende Modellstruktur viele Möglichkeiten ergeben. Agentenmodelle betrachten jeden Bewohner mit individuellen Eigenschaften. Auf der einen Seite ermöglicht das, ähnlich wie stochastische Modelle, eine größere Vielfalt. Auf der anderen Seite können dadurch soziale und psychologische Aspekte in die Verhaltenssimulation einfließen [22]. Neben einem größeren Realismus beim Individualverhalten kann so auch Gruppenverhalten abgebildet werden [29]. Durch Emergenz ergeben sich aus einfachen Verhaltensregeln komplexe Muster auf aggregierter Ebene, die anderweitig schwierig zu modellieren wären [6]. Agentenmodelle sind im Allgemeinen flexibel und leicht erweiterbar, in Bezug auf die Verhaltensmodellierung, Interaktionen und äußere Einflüsse [22]. Beispielsweise ist es möglich, Veränderungen in der Umgebung oder der Demographie im Verlauf einer Simulation miteinzubeziehen. Zur Erstellung eines solchen Modells werden in der Regel viele Informationen benötigt, über die Art und Anzahl der Agenten, deren Beziehungen untereinander und ihre räumliche Einordnung in die Umgebung [23]. Durch die feingranulare Modellierung steigt zudem der Rechenaufwand bei größeren Szenarien rasch an [22]. Diese Faktoren können den Einsatz von agentenbasierten Simulationen erschweren. Es ist daher besonders wichtig, beim Entwurf die Zielkenngrößen der Simulation festzulegen und auf dieser Basis Auflösung, Eingangsdaten und Modellstruktur geeignet zu wählen. Dann ist der agentenbasierte Modelltyp vor allem bei hohen räumlichen und zeitlichen Auflösungen gut geeignet [30].

In Tabelle 2.1 werden die vorgestellten Modelle im Detail gegenübergestellt. Während bei den meisten Modellen Personen individuell simuliert werden, berücksichtigen viele davon keine dynamischen Einflüsse der Umgebung. Ein großer Teil der Modelle verfolgt stochastische Ansätze. Die

teilweise oder vollständig deterministischen Modelle berücksichtigen dagegen eine größere Anzahl äußerer und innerer Faktoren, wodurch ebenfalls die gewünschte Diversität in den Ergebnissen erzielt werden kann.

Viele der vorgestellten Verhaltensmodelle betrachten die Nutzung elektronischer Geräte, Anpassungen von Heizung und Klimaanlage sowie Interaktionen mit der Gebäudehülle wie das Öffnen von Fenstern. Die letzten beiden dieser Aktivitätstypen lassen sich nur sinnvoll im Rahmen einer Gebäudesimulation abbilden, da hierbei Wechselwirkungen mit dem physischen Zustand der aktuellen Umgebung bestehen. Da der Einsatz einer physikalischen Gebäudesimulation nicht Teil dieser Arbeit ist, werden diese Aspekte außer Acht gelassen. Stattdessen wird der Fokus bezüglich des Verhaltens auf Anwesenheit, Transport und Nutzung von Geräten gelegt. Für diese Komponenten sind physikalische Umgebungsparameter deutlich weniger relevant.

Aufgrund der hohen Auflösung und Detailstufe soll ein agentenbasierter Ansatz gewählt werden. Einige Modelle, wie die von Jia et al. oder Zhang, Siebers und Aickelin, sind stark auf eine spezifische Situation zugeschnitten. Sowohl SMACH als auch der LoadProfileGenerator sind detaillierte und umfangreiche Modelle, die bereits viele essenzielle Aspekte menschlichen Verhaltens berücksichtigen und Möglichkeiten zur Ergänzung weiterer Elemente bieten. SMACH benötigt zur Parametrierung passende Zeitverwendungsdaten. Für den LoadProfileGenerator sind diese Daten nicht unmittelbar erforderlich, dafür ist eine passende Konfiguration neuer Haushalte komplexer und für eine entsprechende Validierung werden ebenfalls Zeitverwendungsdaten eingesetzt. Im Gegensatz zu SMACH ist der LoadProfileGenerator jedoch quelloffen und damit frei nutzbar und erweiterbar. Daher wird in dieser Arbeit auf dieses Modell zurückgegriffen.

**Tabelle 2.1:** Vergleich der vorgestellten Verhaltensmodelle

| Model | Typ | Granularität | Umgebung | Daten | Primäres Ergebnis | Methodik | Besonderheit |
|---|---|---|---|---|---|---|---|
| Richardson 2010 [15] | stochastisch | Raum | statisch | TUS | Stromlastprofil | Anwesenheits-Markov-Kette | Gerätenutzungen abhängig von Anwesenheit |
| StROBe 2016 [16] | stochastisch | Person | statisch | TUS | Stromlastprofil, Wärme, DHW | Markov-Kette | Clustering |
| Ziegler 2020 [14] | stochastisch | Person | statisch | TUS | Stromlastprofil | Aktivitäts-Markov-Kette | Aufteilung nach Typtagen |
| Chen 2022 [13] | stochastisch | Person | statisch | TUS | Gebäudelastprofile | Markov-Kette & Sampling | Hybrider Ansatz |
| MASSHA 2017 [8] | stochastisch | Person | statisch | Smart Building Messdaten | Sensoraktivierungen | Priorisierter Aktivitätsplan | Modellierung auf Sensorebene |
| Zhang 2011 [11] | stochastisch | Person | statisch | Eigene Umfrage | Stromlastprofil | Zustandsbasiert | Kommunikation zwischen Agenten |
| Jia 2019 [12] | deterministisch | Person | dynamisch | Eigene Umfrage | Zustände von Fenster, Tür und Jalousie | Performance Moderator Function | Beurteilung von Optionen mittels eines Wertesystems |
| SMACH 2017 [9] | stochastisch | Person | dynamisch | TUS | Stromlastprofil | Aktivitätsstatistiken aus TUS-Daten | Dynamische Anpassung der Prioritäten |
| LoadProfile-Generator 2022 [10] | hybrid | Person | dynamisch | Gerätelastprofile | Lastprofile | Bedürfnis-basiert | Psychologisches Bedürfnismodell |

## 2.3 Städtesimulation

Im vorhergehenden Kapitel wurde die Simulation menschlichen Verhaltens behandelt. Dabei wurden stets einzelne Personen oder Haushalte betrachtet, gegebenenfalls unter Berücksichtigung von Abhängigkeiten der Mitglieder eines Haushalts untereinander. Derartige Abhängigkeiten existieren jedoch auch über die Grenzen eines Haushalts hinaus, wenngleich in geringerer Form. Eine Person, die sich gerade zuhause befindet, kann nicht gleichzeitig in einem Restaurant oder beim Arzt sein. Für eine realistische Modellierung der Auslastung von Unternehmen in einer Stadt ist daher eine konsistente Verhaltenssimulation der Bewohner einer Stadt notwendig. Eine Stadt besteht demnach nicht nur aus einer Menge unabhängiger Häuser, sondern aus einer Vielzahl an Personen mit möglichen Interaktionen und Einflüssen aufeinander. Messbar ist diese Beobachtung beispielsweise durch den maximalen Energieverbrauch einer Stadt, der deutlich unter der Summe der maximalen Energieverbräuche aller einzelnen Gebäude liegt. Die Vernachlässigung der relevanten Abhängigkeiten kann zu einer deutlichen Fehleinschätzung des maximalen Verbrauchs führen [20], da sich möglicherweise eine zu hohe oder zu niedrige Gleichzeitigkeit der Einzelverbräuche ergibt. Aus diesem Grund spielt die Entwicklung dedizierter Städtemodelle, die derartige Zusammenhänge berücksichtigen können, eine große Rolle.

In der Literatur finden sich verschiedene Typen von Modellen, die mit unterschiedlichen Methoden und Zielen versuchen, Städte abzubilden. Vor allem verschiedene Aspekte des Energiesystems werden hierbei häufig detailliert betrachtet. Dabei gibt es Ansätze, die sich auf einen Bereich fokussieren, ebenso wie Modelle, die mehrere Probleme mit einer integrierten Lösung angehen.

Um bisherige Herangehensweisen an vergleichbare Probleme zu zeigen, werden im Folgenden einige Typen von Städtemodellen vorgestellt, die sich mit zeitlichen und räumlichen Abhängigkeiten innerhalb einer Stadt beschäftigen und dabei insbesondere menschliches Verhalten berücksichtigen.

### 2.3.1 Urban Building Energy Models

Im Bereich der Städtemodellierung stellen insbesondere Urban Building Energy Models einen beliebten Ansatz dar. Dieser Modelltyp baut auf der physikalischen Gebäudesimulation auf. Es wird eine größere Anzahl an Gebäuden simuliert, wobei versucht wird, eine real existierende oder in Planung befindliche Stadt möglichst realitätsnah abzubilden. Dafür werden in erster Linie Daten von Geoinformationssystemen benötigt, die Lage und Geometrie jedes Gebäudes beschreiben. Je nach Datenquelle können das bloße Umrisse mit Höheninformationen oder detaillierte 3D-Modelle sein. Da meist kein vollständiges Gebäudemodell für jedes individuelle Haus erstellt werden kann, wird auf Gebäudearchetypen zurückgegriffen. Dabei wird eine feste Menge an Gebäudetypen definiert, für die jeweils mindestens ein Gebäudemodell erzeugt wird, und jedem Gebäude in der Stadt wird der passendste Archetyp zugewiesen.

Abhängig vom Fokus der Modellierung werden zusätzlich weitere energierelevante Aspekte miteinbezogen, beispielsweise die Schatten, die benachbarte Häuser aufeinander werfen, Mikroklimabereiche innerhalb der Stadt oder das Verhalten der Gebäudebewohner. Für Letzteres werden bisher meist einfachere Modelle als die in Abschnitt 2.2 vorgestellten verwendet, oft auch nur statische Aktivitätspläne [31]. Auch städtische Infrastruktur abseits des teilweise ebenfalls simulierten

Stromnetzes wird in der Regel nicht beachtet.

Abschließend werden alle aufgestellten Gebäudemodelle unter Berücksichtigung der zusätzlichen Faktoren simuliert. Durch die explizite Repräsentation jedes Gebäudes können beispielsweise die Auswirkungen von Renovierungsmaßnahmen gut modelliert werden. Während damit meist auf aggregierter Ebene realistische Ergebnisse generiert werden, da einige Fehler sich in der Summe ausgleichen, lassen sich bei einzelnen Gebäuden mitunter deutliche Abweichungen beobachten [32]. Um auch auf dieser Ebene genauere Resultate zu erzielen, wird insbesondere eine bessere, individuelle Abbildung einzelner Bewohner empfohlen. Im Folgenden wird ein Urban Building Energy Model genauer vorgestellt, das bereits eine verhältnismäßig komplexe Wärmesimulation bietet und für das starke Verbesserungen in Bezug auf Verhaltens- und Verkehrssimulation geplant sind.

#### CitySim (bis 2011)

CitySim [33] ist eine Plattform für Urban Building Energy Modelling, die aus dem Simulationstool SUNtool hervorgegangen ist. Der Schwerpunkt liegt hier auf einer präzisen Wärmesimulation jedes Gebäudes im betrachteten Gebiet. Dabei werden auch Einflüsse einzelner Gebäude aufeinander berücksichtigt, in Bezug auf Schatten und Wärmestrahlung. Für die Simulation des Bewohnerverhaltens stehen integrierte deterministische und stochastische Submodelle zur Verfügung. Diese arbeiten bisher nur auf Gebäudeebene, allerdings wurde bereits vorgeschlagen, CitySim mit der Transportsimulation MATSim auf Agentenebene zu koppeln [34]. Dadurch könnte auf die bisherigen Verhaltensmodelle verzichtet werden und stattdessen könnten von MATSim berechnete Ab- und Anreisezeiten genutzt werden. Mittels stochastischer Modelle auf Basis von Zeitverwendungsdaten könnten die entstehenden Anwesenheitszeiträume anschließend mit wärmerelevanten Aktivitäten gefüllt werden, um Verkehrsflüsse und Bewohnerverhalten in Gebäuden kohärent abzubilden.

### 2.3.2 Urban System Energy Models

Urban System Energy Models bilden einen verwandten Modelltyp. Statt komplexer Gebäudesimulationen konzentrieren sich diese Modelle in erster Linie auf die Modellierung der Energienetze einer Stadt [35]. Da hierfür nicht jedes Gebäude einzeln abgebildet werden muss, funktionieren diese Modelle in der Regel mit einer gröberen Auflösung. Aufgrund der ähnlichen Ansätze und sich überschneidender Anwendungsbereiche werden inzwischen auch Urban Building Energy Models mit Urban System Energy Models kombiniert. SynCity [36] stellt eine sehr umfassende Plattform für Urban Energy System Modelling dar, die verschiedene Teilbereiche berücksichtigt und miteinander verbindet. Im Folgenden wird hierauf genauer eingegangen.

#### SynCity (2010)

Keirstead, Samsatli und Shah [36] haben mit SynCity ein Toolkit für die Modellierung urbaner Energiesysteme entwickelt. Ihr Ziel ist die integrierte Simulation mehrerer relevanter Aspekte der Energienutzung in Städten. Dafür wurde ein modularer Ansatz verfolgt. SynCity besteht aus einem Simulationskern, mehreren Submodellen sowie einer Datenbank, die zusammen mit einer Ontologie die Schnittstelle zwischen dem Kern und den Submodellen bildet. In der Ontologie sind

Simulationsobjekte in den Kategorien Ressourcen, Prozesse, Technologien, Flächen und Agenten eindeutig definiert, um einen klaren und exakten Datenaustausch zu gewährleisten.

Zu den entwickelten Submodellen gehören ein Layout-Modell, ein Agent-Activity-Modell sowie ein Resource-Technology-Network-Modell. Das Layout-Modell dient der Optimierung von Gebäudestandorten für Greenfield- oder Brownfield-Szenarien und produziert einen günstigen Stadtplan. Darauf aufbauend können mithilfe des Agent-Activity-Modells Aktivitäten und Bewegungen aller Personen simuliert werden. Dafür wird ein vierstufiges Modell genutzt, wie es auch bei einigen Transportsimulationen üblich ist. Jede Person wird als eigenständiger Agent modelliert und wählt in jedem Zeitschritt eine Aktivität, einen Ort, an dem sie dieser Aktivität nachgehen kann, eine Form des Transports und eine Route zum gewählten Zielort. Für die Aktivitätenwahl werden Aktivitätsprofile für verschiedene Tageszeiten, Tagestypen (Wochentag oder Wochenende) und Jahreszeiten als Eingabe benötigt. Die Autoren nutzen hierfür das Modell TASHA [37]. Aus den Aktivitätsprofilen wird anhand sozioökonomischer Eigenschaften wie Alter und Einkommen ein eigener Aktivitätsplan für jede Person erstellt. Über die simulierten Aktivitäten werden anfallende Energie- und Rohstoffbedarfe ermittelt. Anhand der räumlichen und zeitlichen Verteilung dieser Bedarfe kann das Resource-Technology-Network-Modell abschließend eine optimale Energieinfrastruktur für die Stadt berechnen.

Erklärtes Ziel von SynCity ist unter anderem, die Modellparametrierung zu erleichtern und auf große Eingabedatensätze zu verzichten. Das wird durch die sequenzielle Ausführung der Modelle erreicht, bei der jedes Submodell durch die Ergebnisse der vorhergehenden Modelle parametriert wird. Das SynCity-Toolkit wurde auf eine geplante britische Stadt mit reduziertem $CO_2$-Ausstoß angewandt. Die aggregierten Modellergebnisse passten dabei gut zu den Planungsdaten der Stadt, die zeitlich und räumlich aufgelösten Daten konnten allerdings nicht verglichen werden. Ein detaillierteres Agentenmodell, dass das bisherige Agent-Activity-Modell ersetzen kann, soll laut den Autoren noch entwickelt werden, Genaueres wurde hierzu aber bisher nicht veröffentlicht [20].

### 2.3.3 Transportsimulationen

Eine andere Art von Modellen mit einem Fokus auf urbanen Umgebungen stellen Transportsimulationen dar. Diese zielen auf die Evaluation von Verkehrsinfrastrukturen im Rahmen einer Städteplanung ab. Zu diesem Zweck wird versucht, die Bewegungen aller Bewohner eines Gebiets passend wiederzugeben. Neben einer geeigneten Darstellung der räumlichen Umgebung, inklusive der nutzbaren Wege für verschiedene Fortbewegungsmittel, ist dafür die Simulation einiger menschlicher Entscheidungen auf unterschiedlichen Ebenen notwendig. Aufgrund ihrer guten Eignung für diese Form von Problemen, unter anderem durch die individuelle Abbildung einzelner Verkehrsteilnehmer mit eigenem Verhalten statt einer globalen Systembeschreibung, hat sich die Agentensimulation im Transportbereich etabliert. Dadurch ergeben sich verschiedene Vorteile, beispielsweise lassen sich viele intuitive Regeln und Einschränkungen leicht implementieren und durch die Nutzung heterogener Agenten kann eine realistische Diversität erreicht werden [38]. Im Vergleich zu Occupant Behavior finden Beeinflussungen der Agenten untereinander eher indirekt statt, beispielsweise durch Staus. Daher ist eine präzise, zeitlich und räumlich hoch aufgelöste Simulation von Bedeutung.

Transportsimulationen betrachten im Allgemeinen drei wesentliche Teilbereiche. Der erste Schritt

ist die Modellierung von Reisebedarf, also Informationen zu Zeit, Startpunkt und Ziel aller Reisen, die stattfinden sollen. Darauf folgen verkehrsbedingte Entscheidungen wie die Wahl des Transportmittels und der Route. Abschließend wird der Verkehrsfluss selbst simuliert.

Ein Nachteil der agentenbasierten Ansätze liegt im grundsätzlich hohen Rechenaufwand. Da in vielen Fällen große Szenarien, teils mit Millionen an Agenten, betrachtet werden und gleichzeitig eine hohe zeitliche Auflösung gefordert ist, wird häufig nur ein kurzer Zeitabschnitt simuliert, bei vielen Modellen ein einzelner Tag. Es handelt sich somit um Querschnittsanalysen, langfristige Veränderungen können nur über separate Szenarien untersucht werden [39]. Im Folgenden werden zwei der beliebtesten und umfassendsten Transportmodelle exemplarisch vorgestellt.

### MATSim

Ein sehr beliebtes und häufig eingesetztes Hilfsmittel bei Fragestellungen zur Mobilität in Städten ist MATSim [40]. Als aktivitätsbasiertes Framework für Multi-Agenten-Simulationen bietet es Module für die verschiedenen Schritte bei Transportsimulationen. Der Kern des Frameworks besteht aus einem sogenannten koevolutionären Ansatz. Die Grundlage für den Reisebedarf bilden Aktivitätspläne, die für alle Agenten über die Eingabedaten vorgegeben werden. Zu jeder Aktivität wird neben einer vorläufigen Startzeit ein Ausführungsort definiert. In einem iterativen Verfahren kann nun jeweils ein Teil der Agenten seine Aktivitätspläne mitsamt den dazugehörigen Entscheidungen bezüglich Routen, Abreisezeiten und Transportmittel variieren. Nach erneuter Simulation werden alle Pläne durch Scoring-Funktionen bewertet, um anzugeben, wie gut oder schlecht sie für den jeweiligen Agenten funktioniert haben. Neben langen Reisedauern und hohen Kosten können dabei auch Wartezeiten, zu späte Ankünfte, etwa beim Arbeitsplatz, oder zu kurze Dauern negativ gewertet werden. Über mehrere Iterationen hinweg werden für jeden Agenten mithilfe diskreter Entscheidungsmodelle aus allen verfügbaren Aktivitäts- und Reiseplänen die besten ausgewählt und weiterverwendet. So stellt sich nach einer ausreichenden Anzahl an Iterationen ein Gleichgewicht zwischen allen Agenten ein. Eine Aktivität bezeichnet in diesem Zusammenhang in erster Linie den Zeitraum zwischen zwei Reisen. Somit findet die Aktivitätsmodellierung bei MATSim in der Regel auf einem gröberen Niveau als bei typischen Occupant-Behavior-Modellen statt.

Die Verkehrswege im simulierten Gebiet werden über ein Netzwerk dargestellt, das über die Eingabe definiert wird. Sämtliche Straßen sind darin als Verbindungslinien enthalten, und Orte werden, automatisch oder manuell, einer Verbindungslinie zugeordnet, von der aus sie erreicht werden können. Da der Schwerpunkt bei MATSim auf großen Modellen und performanter Berechnung liegt, wird auf eine detaillierte physikalische Simulation mitsamt einem Car-Following-Modell verzichtet. Stattdessen wird jede Verbindung im Netzwerk als Warteschlange mit festgelegter Kapazität und Mindestdauer behandelt, die eingehende Verkehrsteilnehmer nacheinander am entgegengesetzten Ende hinauslässt, solange in den darauffolgenden Warteschlangen noch Platz frei ist.

### SUMO

Neben MATSim ist auch SUMO [41] eine beliebte Mikrosimulation im Transportbereich. Auch hier wird anhand von Straßennetzwerken gearbeitet, mit separaten Kanten für Fahrbahnen, Radwege, Gehwege etc. Die Bewegungen von Fahrzeugen und Fußgängern werden detailliert simuliert, inklusive eines Car-Following-Modells zur Simulation der Abstände zwischen aufeinanderfolgenden

Fahrzeugen. Für eine effiziente Berücksichtigung der Interaktionen von Fußgängern untereinander oder mit Fahrzeugen, beispielsweise an Fußgängerüberwegen, werden nur andere Verkehrsteilnehmer in geringem Abstand überprüft. Für eine derart feine Simulation werden präzise Daten zur Verkehrsinfrastruktur benötigt, es existieren allerdings Module, die diese Informationen, falls sie unvollständig sind, automatisch ergänzen können.

SUMO selbst ist nicht aktivitätsbasiert und benötigt als Eingabe eine fertige Liste an Routen mitsamt Startpunkt, Endpunkt und Uhrzeit. Es kann aber das Tool ACTIVITYGEN genutzt werden, um anhand von Bevölkerungsstatistiken über Aktivitätspläne Routen zu generieren. Dabei werden nur die Aktivitäten Arbeit, Schule und Freizeit unterstützt, also ein für Transportmodelle typisches, gröberes Abstraktionsniveau.

### 2.3.4 Landnutzungsmodelle

Neben Urban Building Energy Models, Urban System Energy Models und Transportsimulationen, alle mit kurzem bis mittelfristigem Zeithorizont, lassen sich in der Literatur auch Landnutzungsmodelle finden, die typischerweise deutlich längere Zeitspannen betrachten, meist mehrere Jahrzehnte. Ziel dieser Modelle ist es, die private und gewerbliche Nutzung von Grundstücken durch verschiedene Akteure zu modellieren und vorherzusagen. Dazu gehört, wann und wohin Haushalte, Geschäfte oder Firmen umziehen oder auf welche Weise ein Gebiet beispielsweise durch Grundstücksentwickler verändert wird. Dabei werden meist mehrere separate Submodelle für die verschiedenen Aspekte eingesetzt, zum Beispiel Markov-Ketten oder diskrete Entscheidungsmodelle unterschiedlicher Art [42].

Ein wichtiges Kriterium zur Bewertung eines Grundstücks stellt die Erreichbarkeit dar, das heißt wie gut die Verkehrsanbindung an verschiedene Ziele in der Umgebung ist. Diese hängt vom spezifischen Verkehrsfluss ab und kann durch die in Abschnitt 2.3.3 beschriebenen Transportmodelle ermittelt werden. Diese Modelle erfordern als Eingabedaten unter anderem Informationen zur Lage relevanter Orte und Gebäude in der Stadt, welche wiederum durch Landnutzungsmodelle erzeugt werden. Aus diesem Grund werden Landnutzungsmodelle mittlerweile meist in Kombination mit einem Transportmodell eingesetzt. Die Modelle werden dabei abwechselnd aufgerufen, um für ein spezielles Szenario zunächst die Flächennutzung und anschließend die dazu passenden Verkehrsflüsse zu bestimmen. Auf diese Weise können viele Modellaspekte realistischer dargestellt werden [43]. Diese gegenseitige Abhängigkeit wird als *Land-Use Transport Feedback Cycle* bezeichnet [44]. Diese Modellierungsstruktur und die damit verbundenen Datenflüsse sind in Abbildung 2.5 dargestellt. Aufgrund des hohen Rechenaufwands von Transportmodellen werden diese Modelle in der Regel seltener aufgerufen, sodass etwa nur für jedes fünfte Jahr ein Tag exemplarisch simuliert wird, während die Landnutzungsmodelle selbst meist für jedes Jahr im festgelegten Zeitraum ausgeführt werden [45].

Im Folgenden werden zwei komplexe und umfangreiche integrierte Landnutzungs- und Transportmodelle kurz vorgestellt.

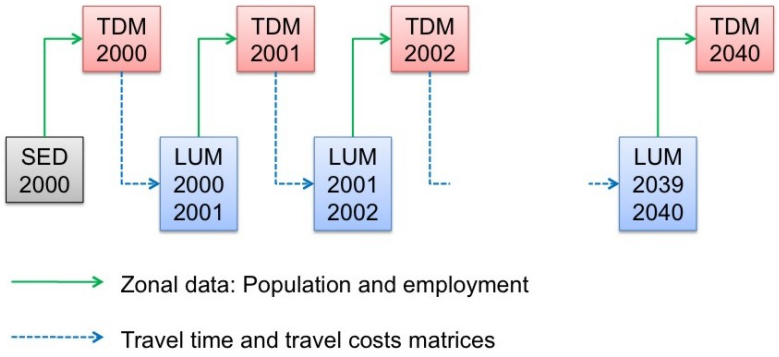



**Abbildung 2.5:** Kombination von Land-Use-Modellen mit Transportmodellen, von [43]

### UrbanSim (2002 – 2007)

Mit UrbanSim [45] wurde bereits seit Ende der 1990er Jahre eine Plattform zur Modellierung der langfristigen Entwicklung von Städten aufgebaut. Ursprünglich in Java programmiert, wurde UrbanSim im Verlauf der weiteren Entwicklung mehrfach neu entworfen und umstrukturiert. Die aktuelle Implementierung in Python wird als Teil des Urban Data Science Toolkit [46] zusammen mit verschiedenen Tools weiterentwickelt, die die Modellerstellung und die Ausführung der Kernkomponente von UrbanSim unterstützen können. Zu diesen Tools gehören unter anderem Orca zur Orchestrierung der Simulationsaufgaben und SynthPop zur Generierung von Modellpopulationen.

Mithilfe der integrierten Modellierung von Landnutzung und Transportsystemen untersucht UrbanSim, welche Auswirkungen politische Regulierungen, Investitionen in Transportinfrastruktur oder Änderungen in der Bevölkerung auf die Nutzung von Landflächen haben können. UrbanSim selbst bietet hierfür die Simulationsarchitektur, basierend auf einem Model Coordinator, der für Kommunikation und Datenverwaltung zuständig ist. Eingangsdaten und Zwischenergebnisse werden zentral in einer Datenbank gespeichert. Die eigentliche Modellierung wird auf mehrere unabhängige Modelle aufgeteilt. Üblicherweise wird mindestens jeweils ein Modell für die Standortwahl von Haushalten, von Geschäften und von Arbeitgebern sowie für die Entwicklung von Immobilien einbezogen. Aufgrund ihrer unterschiedlichen Zwecke nutzen diese Modelle verschiedene Methoden. Während die Demographie auf aggregierter Ebene modelliert wird, werden für Standortentscheidungen diskrete Wahlmodelle auf Haushalts- bzw. Arbeitgeberebene eingesetzt. Für die Transportsimulation können externe Modelle angebunden werden, um die Wechselwirkungen zwischen Flächennutzungen und Transportwegen zu berücksichtigen.

Eine verteilte Simulation ist mit UrbanSim nach derzeitigem Stand nicht möglich. Es wurden bereits Untersuchungen hinsichtlich einer möglichen parallelen Ausführung der Modelle angestellt, jedoch konnten aufgrund technischer Hindernisse von Python und der Architektur von UrbanSim nur eingeschränkt nutzbare Umsetzungsmöglichkeiten gefunden werden, die bisher noch nicht implementiert wurden [47].

### ILUMASS (2018)

Ein anderes integriertes Landnutzungs- und Transportmodell, ILUMASS [48], setzt ebenfalls auf einen modularen Ansatz. ILUMASS besteht aus mehreren individuellen Mikrosimulations-Modulen,

jedes mit einem anderen Ziel. Enthalten sind unter anderem Modelle zur demographischen Entwicklung, zur Erzeugung von Aktivitätsplänen und zur Simulation von Verkehrsflüssen. Die Ergebnisse der einzelnen Submodelle werden von einem zentralen Server gesammelt, gespeichert und an die jeweils nächsten Submodelle weitergeleitet. Dabei wird ein iterativer Ansatz verfolgt, bei dem Datenaustausch und Modellausführung wiederholt werden, bis ein Gleichgewicht erreicht wird und konsistente, stimmige Ergebnisse erzielt werden. Ein dediziertes Umwelt-Modul kann dadurch Indikatoren wie $CO_2$-Emissionen für verschiedene Teilbereiche der simulierten Stadt berechnen.

Der hohe Detailgrad und die Kopplung derart vieler Mikrosimulationen verursachen allerdings einen enormen Rechenaufwand, sodass das Modell in dieser Konfiguration nicht im benötigten Maßstab eingesetzt werden kann. Um Abhilfe zu schaffen, wurde eine Variante von ILUMASS mit reduzierter Komplexität in einigen Submodellen erstellt, die signifikant bessere Laufzeiten erreicht. Die Entwickler empfehlen daher, nicht sämtliche Fragestellungen in einem integrierten Modell mit Mikrosimulationen anzugehen, sondern einen Multi-Level-Ansatz zu verfolgen und jeden Teilbereich nur mit der dafür jeweils benötigten Granularität zu modellieren [49].

### 2.3.5 Geosimulationen

Einen weiteren Modelltyp im Bereich der Städtesimulationen bilden Geosimulationen. Diese Kombination aus Geoinformationssystemen und Simulationen stellt Objekte räumlich präzise dar, meist auf Haushalts- oder Personenebene [50]. Das bedeutet, dass jedes simulierte Objekt jederzeit einem Ort in der Simulationsumgebung zugeordnet werden kann. Auf diese Weise können Fragestellungen mit geographischem Schwerpunkt, etwa zur Infrastruktur einer Stadt, unter Berücksichtigung physikalischer, sozialer und demographischer Einflüsse betrachtet werden.

Ein umfangreiches und detailliertes Beispiel für eine Geosimulationsplattform stellt GeoWorldSim von Pijoan [51] dar. Von Grund auf in C++ und Qt entwickelt, nimmt die Performance hier einen hohen Stellenwert ein, um auch größere Modelle mit Millionen an Haushalten simulieren zu können. Dafür wurde eine ereignisgesteuerte Architektur für verteilte Simulation aufgebaut, die die parallele Berechnung durch Worker-Threads auf einem oder mehreren Rechenknoten ermöglicht. Anstatt einer Synchronisierung nach jedem Simulationszeitschritt verwendet jeder Worker eine eigene Event-Queue zur Abarbeitung der ihm zugewiesenen Agenten. Um zu verhindern, dass die Worker zeitlich zu stark divergieren, wird sichergestellt, dass alle Worker innerhalb eines festgelegten Zeitfensters bleiben.

GeoWorldSim ist modular aufgebaut. Personen, Haushalte, Geräte sowie Umgebungen werden als vollwertige Agenten implementiert. Für jeden davon können wiederverwendbare Fertigkeiten definiert werden, wie zum Beispiel bestimmte Abfragen der Umgebung, Kontaktierung eines anderen Agenten oder Ausgabe von Daten. So können mehrere Agenten ohne mehrfache, redundante Programmierung dieselben Funktionalitäten erhalten. Das Verhalten eines Agenten wird über eine Liste auszuführender Aktivitäten vorgegeben, die, je nach Angabe, sequenziell oder parallel abgearbeitet werden.

Eine Besonderheit von GeoWorldSim liegt in der Verwendung der Schema.org-Ontologie zur Definition von Agenten. Neben einigen vordefinierten Agentenklassen aus den wichtigsten Bereichen besteht für Anwender die Möglichkeit, mithilfe einer Object Factory Agenten aus Schema.org-

konformen Dateien im JSON-LD-Format zu generieren. Dadurch wird nicht nur die eindeutige und weit verbreitete Objekthierarchie von Schema.org genutzt, es erleichtert auch die Aufnahme neuer Agententypen in die Simulation.

Für die Simulationsumgebung wurde ein mehrstufiger Ansatz gewählt. Die geographischen Funktionen werden in Form mehrerer getrennter Umgebungs-Objekte umgesetzt, darunter eine physische Umgebung, die die Geometrien jedes Agenten überprüft und korrekt in übergeordnete Geometrien einordnet, sowie eine Netzwerkumgebung zur Darstellung von Verkehrsnetzen. Weitere Umgebungen übernehmen jeweils die Simulation sozialer Beziehungen, Übermittlung von Nachrichten und Steuerung der Ausführung.

Um die Vielseitigkeit von GeoWorldSim zu demonstrieren, wurden damit im Rahmen von Fallstudien mehrere Geosimulationen mit unterschiedlichen Zielen und Anwendungsbereichen erstellt, beispielsweise eine urbane Transportsimulation und eine Simulation zur sensorbasierten Erkennung menschlicher Aktivitäten in Smart Cities. Als Verhaltenssimulation ist MASSHA (siehe Abschnitt 2.2.5) in GeoWorldSim integriert. Gemeinsame Aktivitäten innerhalb eines Haushalts sind dabei nicht möglich, logische Abhängigkeiten zwischen Aktivitäten werden aber berücksichtigt [8].

### 2.3.6 Vergleich der Modelle

Wie aus dem hier gezeigten kurzen Überblick zur Städtemodellierung erkennbar ist, gibt es eine Vielzahl unterschiedlicher Ansätze und Methoden, um die dynamischen Prozesse auf verschiedenen Ebenen einer Stadt abzubilden. Keines der Modelle kann alle Aspekte einer Stadt berücksichtigen, daher werden jeweils unterschiedliche Schwerpunkte gesetzt.

Urban Building Energy Models konzentrieren sich auf physikalische Wärmesimulationen und beziehen nur speziell dafür relevantes Occupant Behavior mit ein, oftmals mit verhältnismäßig einfachen und ungenauen Methoden. Urban System Energy Models messen menschlichen Aktivitäten häufig ähnlich wenig Bedeutung bei. Bei beiden Modelltypen sind keine Verhaltensabhängigkeiten auf urbaner Ebene zu erkennen [20]. Obwohl zeitlich und räumlich präzise Verhaltensmodelle existieren, werden diese bei derartigen Modellen bisher nicht integriert [52].

Geosimulationen und Transportmodelle betrachten menschliches Verhalten in direkter Verbindung mit Bewegungen und Reisen, allerdings wird auch hier in der Regel ein verhältnismäßig hohes Abstraktionsniveau gewählt, mit nur wenigen möglichen Aktivitäten. Präzisere Verhaltensmodellierung beschränkt sich auf sehr eng abgegrenzte Teilbereiche, wie das Abstandhalten zum vorausfahrenden Fahrzeug.

Integrierte Landnutzungs- und Transportmodelle bilden, neben der eingebundenen Verkehrssimulation, längerfristige menschliche Entscheidungen ab. Aufgrund der langen Simulationszeiträume können sie keine präziseren, kurzfristigeren Verhaltensmodellierungen enthalten.

Bei allen vorgestellten Modellen werden Aktivitäten lediglich in einem einheitlichen Umfeld betrachtet, beispielsweise im Straßenverkehr, in privaten Haushalten oder in Büros. Eine integrierte Simulation über mehrere dieser Bereiche hinweg fehlt bislang [20].

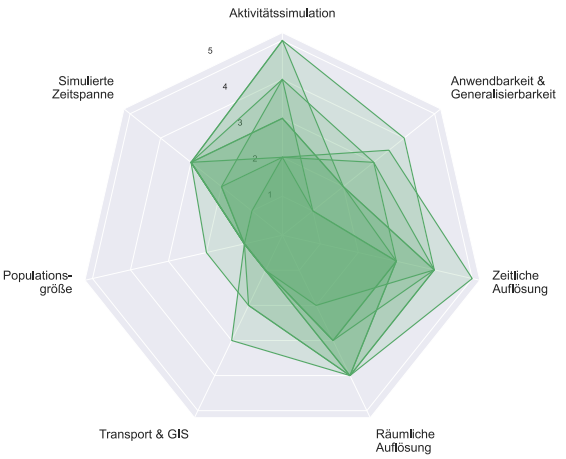
Insgesamt werden menschliche Aktivitäten bei der Berechnung von Ressourcen- und Energiever-

bräuchen auf Städteebene nur in geringem Umfang einbezogen. Da eine Korrelation dieser Bereiche nachgewiesen werden konnte, kann hier von einem großen Verbesserungspotenzial ausgegangen werden [52].

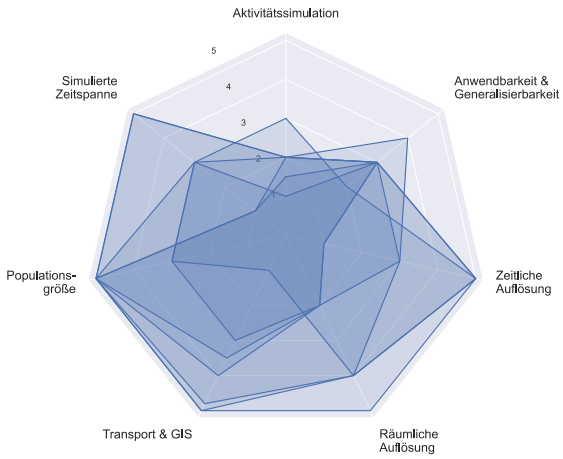
Allgemein wird für eine gründlichere Betrachtung der Vorgänge in Städten zukünftig verstärkt domänenübergreifende Modellierung eingesetzt [26]. Ein Trend in diese Richtung liegt in der Kopplung verschiedener Städtemodelle mit Transportsimulationen, um die Präzision angenommener An- und Abwesenheiten zu erhöhen.

In Abbildung 2.6 werden die unterschiedlichen Stärken und Schwächen aller gezeigten Modelle veranschaulicht. Die Bewertungen basieren auf Einschätzungen des Detailgrades, mit dem bestimmte Aspekte berücksichtigt werden. Sie sind als relative Abstufung zu verstehen und sollen einen Eindruck der Schwerpunkte eines Modells vermitteln, und stellen somit kein absolutes Maß für die Qualität der Ergebnisse dar. Gut zu erkennen sind darin die unterschiedlichen Prioritäten der in Abschnitt 2.2 vorgestellten Verhaltensmodelle im Vergleich zu den verschiedenen Städtemodellen. Zudem werden die Defizite der Städtemodelle im Bereich der Aktivitätssimulation vor dem Hintergrund der komplexen Verhaltensmodelle deutlich, und umgekehrt der Mangel an urbaner Modellierung bei Verhaltensmodellen.

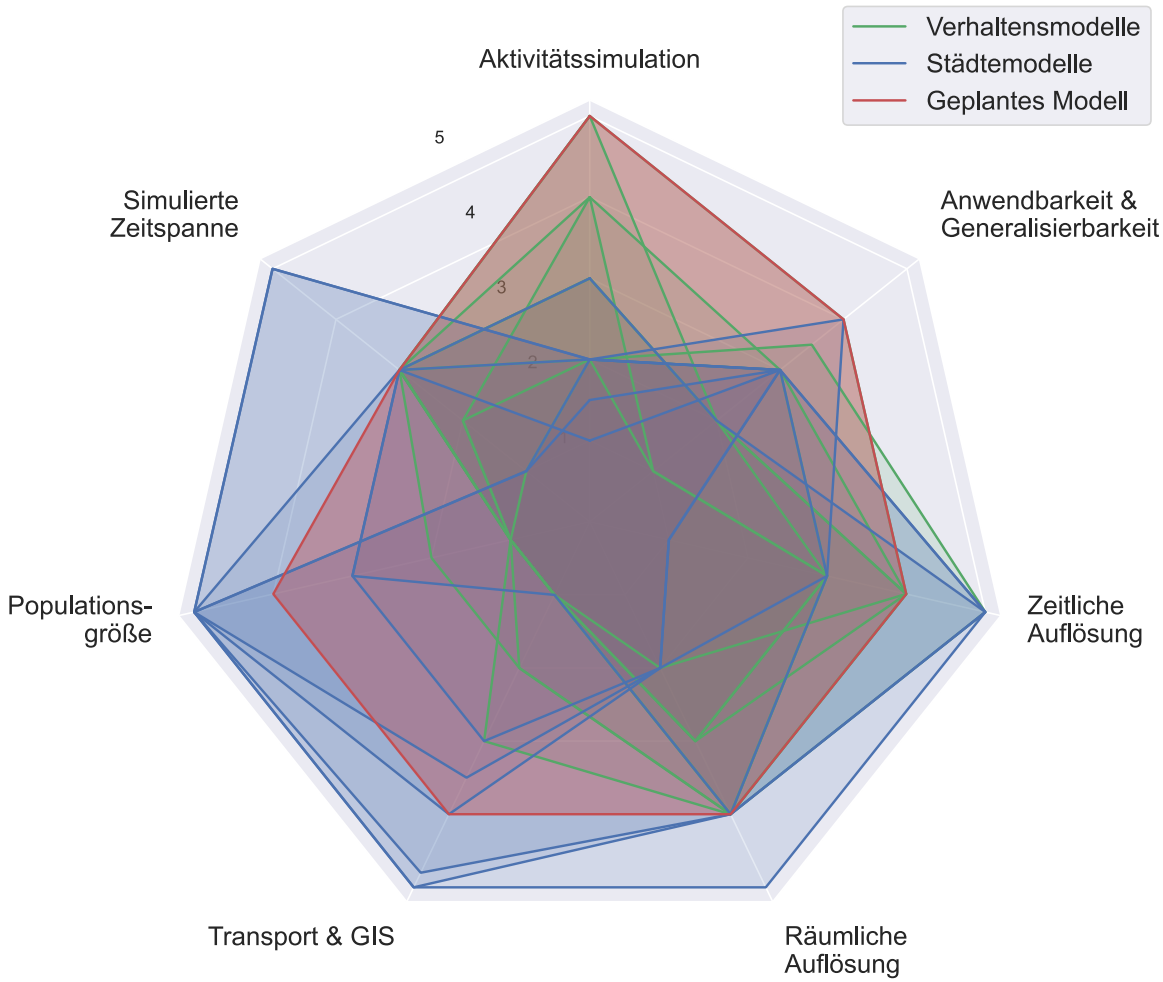
Um die aufgezeigten Lücken zu schließen, soll ein neues Modell entwickelt werden. In erster Linie soll dieses eine detaillierte Verhaltenssimulation mit der Modellierung auf Städteebene, inklusive Reisen und Aufenthalten in Unternehmen, kombinieren, um realistischere Aktivitäts- und Anwesenheitsprofile generieren zu können. Um den Rechenaufwand handhabbar zu halten, kann bei diesem Vorhaben nicht jeder Teilbereich in höchster Auflösung und Genauigkeit simuliert werden. Stattdessen soll jedes Gebiet mit angemessener Komplexität berücksichtigt werden, um mögliche Wechselwirkungen zu erhalten und gleichzeitig die Berechnung großer Szenarien zu erlauben. In Abbildung 2.6c werden die vorgesehenen Ausprägungen des in dieser Arbeit entwickelten Modells im Vergleich zu den existierenden Modellen dargestellt.




**(a)** Verhaltensmodelle




**(b)** Städtemodelle




**(c)** Geplantes Modell im Vergleich zu existierenden Modellen

**Abbildung 2.6:** Bewertung aller vorgestellten Verhaltens- und Städtemodelle anhand ausgewählter Kriterien auf einer Skala von 1 (grob) bis 5 (detailliert). Zum Vergleich wird in (c) die geplante Städtesimulation entsprechend eingeordnet.

# 3 Methodik

Um genauere und verlässlichere Aussagen über den zeitlichen Verlauf von Energieverbräuchen, Verkehr und Auslastungen von Ressourcen in ganzen Städten oder Stadtteilen treffen zu können, sind detaillierte Simulationen der relevanten menschlichen Aktivitäten erforderlich. Dabei müssen insbesondere mögliche Interaktionen und dadurch hervorgerufene Effekte berücksichtigt werden können, eine unabhängige Simulation aller einzelnen Einwohner oder einzelner Aspekte einer Stadt ist nicht ausreichend. Stattdessen wird ein integriertes Modell benötigt, das das Verhalten der kompletten Bevölkerung einer Stadt unter Berücksichtigung aller relevanten Einflüsse simuliert. Nur auf diese Weise können verschiedenste Szenarien mit kombinierten Interventionen beim Verhalten, Transport oder Energieverbrauch abgebildet werden.

Wie in Kapitel 2 gezeigt wurde, existiert bisher kein Modell, dass auf einem derartigen Detailgrad arbeitet und eine ausreichend komplexe Verhaltenssimulation enthält. Daher wurde in dieser Arbeit ein bestehendes Verhaltensmodell für Privathaushalte zu einer vollwertigen, integrierten Städtesimulation ausgebaut. Der Mehrwert dieser Simulation besteht insbesondere darin, dass die voneinander abhängigen Aspekte des menschlichen Verhaltens, des Transports und des Ressourcenverbrauchs gemeinsam betrachtet werden und so gegenseitige Einflüsse berücksichtigt werden können. Dadurch sollen entsprechende Fragestellungen beantwortet werden, zum Beispiel wie die Standorte von Geschäften alltägliches Verhalten, Transportwege und Lastkurven, aber auch die Auslastung der Geschäfte selbst beeinflussen.

Im Folgenden werden zunächst die Anforderungen definiert und ein beispielhafter Anwendungsfall vorgestellt. Anschließend wird das Konzept der entwickelten Städtesimulation erklärt und es wird auf grundlegende Designentscheidungen eingegangen. Schließlich wird die Implementierung aller erforderlichen Komponenten beschrieben.

## 3.1 Anforderungen

Das Ziel dieser Arbeit ist die Entwicklung einer eigenständigen Simulation zur Abbildung von Aktivitäten, Verkehr und Energieverbräuchen in Städten. Die Simulation soll die alltäglichen Aktivitäten jedes einzelnen Einwohners einer Stadt bestimmen, inklusive etwaiger Reisen zu anderen Orten innerhalb der Stadt. Daraus sollen Lastprofile für alle relevanten Lasttypen, insbesondere Elektrizität, berechnet werden. Zudem sollen Besuchszeiten und aktuelle Auslastungen aller Geschäfte einer Stadt ermittelt werden. Insgesamt soll die Städtesimulation somit in erster Linie Aktivitätsprofile, Lastprofile und Besuchsprofile erzeugen. Diese Daten sollen jeweils für jede betroffene Einheit in der Stadt angelegt werden. Es soll ein Aktivitätsprofil pro Person erstellt werden, ein Lastprofil pro Haushalt für jeden Lasttyp, sowie ein Besuchsprofil für jedes Geschäft.

Die Simulation muss die Bestandteile Transport, Aktivitäten und Ressourcenverbrauch integriert abbilden, also nicht einzeln durch Verkettung mehrerer unabhängiger Modelle. Stattdessen sollen die betroffenen Komponenten auf eine Weise modelliert werden, die es ermöglicht, Abhängigkeiten untereinander einzubeziehen. Hierfür soll eine offene und erweiterbare Simulationsinfrastruktur geschaffen werden, die bei Bedarf um für bestimmte Anwendungsfälle notwendige Einschränkungen erweitert werden kann. Explizit soll es möglich sein, Interaktionen verschiedener Bewohner einer Stadt miteinander einzubeziehen. In der einfachsten Form bezieht sich das beispielsweise auf das

gemeinsame Warten in einer Schlange, etwa vor einem Geschäft. Hierbei müssen sämtliche Abläufe zeitlich und logisch konsistent simuliert werden, um sicherzustellen, dass Wartezeiten plausibel abgebildet werden.

Eine derart detaillierte und umfassende Simulation einer hinreichenden Anzahl Personen ist sehr aufwendig. Um die Simulation sinnvoll einsetzen zu können, muss die Rechenzeit jedoch in einem handhabbaren Rahmen bleiben. Hierbei ist das Ziel, ganze Städte mit einer Größe von bis zu 100.000 Einwohnern mit einer Auflösung von einer Minute in akzeptabler Rechenzeit von maximal einigen Tagen simulieren zu können, eine geeignete Hochleistungsrechenplattform vorausgesetzt. Hierfür ist die effiziente Verwendung moderner paralleler Rechnerarchitekturen unumgänglich. Als üblicher Simulationszeitraum wird eine Spanne von einem Jahr angenommen, um jahreszeitliche Schwankungen zu berücksichtigen.

Um eine Stadt realitätsnah simulieren zu können, müssen alle wesentlichen Komponenten mit hinreichender Genauigkeit abgebildet werden. Dafür werden grundlegende Kenndaten wie die Anzahl an Wohn- und Nichtwohngebäuden benötigt. Für alle Wohngebäude müssen Informationen über die darin lebenden Bewohner bekannt sein, und für Nichtwohngebäude ist der Verwendungszweck relevant, also ob es sich beispielsweise um einen Supermarkt oder ein Schwimmbad handelt. Bei der anvisierten Städtegröße ist eine manuelle Dateneingabe unpraktikabel. Daher muss eine Möglichkeit implementiert werden, die Städtesimulation automatisch anhand einer geeigneten Datenbasis zu parametrieren.

## 3.2 Fallbeispiel Jülich

Als ein beispielhafter Anwendungsfall für die Städtesimulation wird Jülich betrachtet. Jülich ist eine mittlere Stadt im Kreis Düren in Nordrhein-Westfalen mit etwa 35.000 Einwohnern im Jahr 2025. Basierend auf der Auswertung der Gebäudedatenbank BUILDA [53] besteht die Stadt aus insgesamt 16.806 Gebäuden. Davon sind 9.108 reine Wohngebäude und 7.234 reine Nichtwohngebäude, die verbliebenen 464 Gebäude werden gemischt genutzt.

Dieses Fallbeispiel wird im Folgenden zur Veranschaulichung der Städtesimulation genutzt und in Kapitel 4 im Detail ausgewertet.

## 3.3 Ansatz und Modell

Das Konzept der entwickelten Städtesimulation besteht im Kern aus der gleichzeitigen Anwendung des gewählten Verhaltensmodells, des LoadProfileGenerators, für jeden einzelnen Haushalt im simulierten Gebiet. Aufgrund des damit verbundenen hohen Rechenaufwands wird eine parallele Rechenarchitektur mit mehreren Prozessen eingesetzt. Der LoadProfileGenerator (siehe Abschnitt 2.2.5) wurde als zugrundeliegendes Verhaltensmodell ausgewählt, da er im Hinblick auf Genauigkeit, Detailgrad und Erweiterbarkeit am besten zu den Anforderungen dieser Arbeit passt. Er bietet als einziges eine ausreichend komplexe Verhaltenssimulation jeder einzelnen Person und ist zudem als quelloffene Software frei erweiterbar und nutzbar.

Ein vollständiger Simulationsdurchlauf setzt sich aus den in Abbildung 3.1 dargestellten Schritten zusammen. Um die Erstellung von Städteszenarien von der eigentlichen Simulation zu entkoppeln und so Entwicklung und Anwendung des Modells zu erleichtern und zu beschleunigen, werden

für diese beiden Aufgaben separate Programme entwickelt. Das Simulationsprogramm verteilt die Wohn- und Nichtwohngebäude in einem Szenario gleichmäßig auf mehrere parallele Prozesse, die jeweils die entsprechenden Modelle vorbereiten. Anschließend startet der Hauptteil der Simulation, bei dem jeder Zeitschritt jedes Gebäudes einzeln simuliert wird. Zum Schluss werden die gespeicherten Daten ausgewertet und zusätzliche Ergebnisdateien generiert.

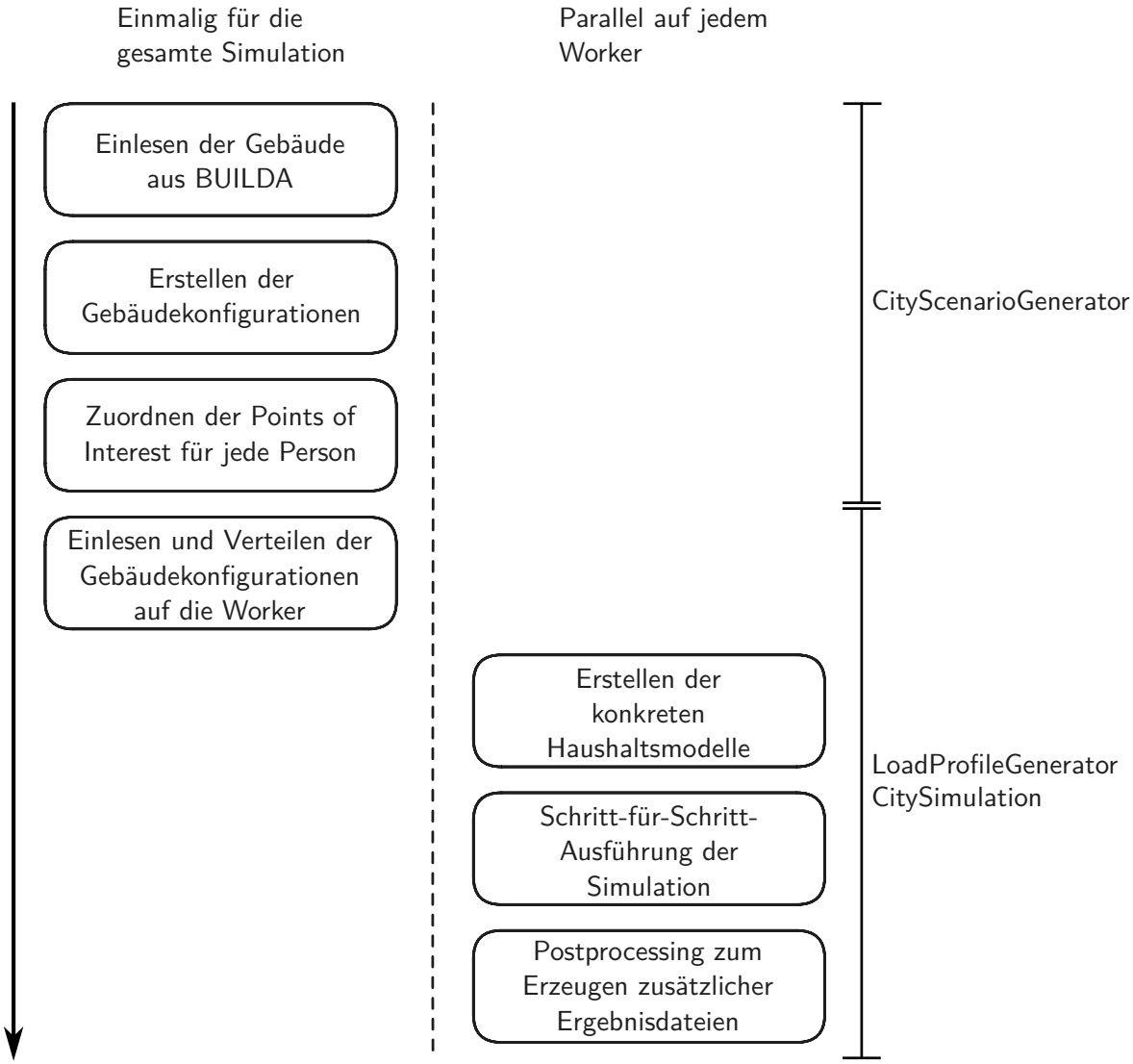



**Abbildung 3.1:** Wesentliche Schritte im Ablauf der Städtesimulation

Im Folgenden wird zunächst das Konzept des separaten Programms zur Szenario-Generierung vorgestellt. Anschließend wird in Abschnitt 3.3.2 der Aufbau der benötigten Simulationsinfrastruktur für eine synchrone, parallele Simulation aller Bewohner einer Stadt beschrieben, und in Abschnitt 3.3.3 die grundlegende Struktur der Städtesimulation erklärt. Schließlich wird in Abschnitt 3.3.4 genauer auf die Abbildung von Reisen und Aktivitäten innerhalb der Stadt eingegangen.

### 3.3.1 Szenarioerstellung

Die Erstellung von Szenarien für die Städtesimulation ist eine komplexe Aufgabe. Es müssen große Datenmengen aus verschiedenen Quellen verarbeitet, kombiniert und in geeigneter Form für die

Simulation bereitgestellt werden. Die einzelnen Elemente sind dabei bereits stark voneinander abhängig. So muss beispielsweise bekannt sein, wo in einer Stadt sich Supermärkte oder Arztpraxen befinden, um festlegen zu können, welche dieser Orte die Bewohner eines Haushalts aufsuchen. Für eine realistische Auswahl der Zielorte sind wiederum korrekte Distanzangaben von einem Haushalt zum potenziellen Ziel essenziell. Um ein konsistentes Szenario zu erhalten, müssen somit alle für den LoadProfileGenerator relevanten Eingangsdaten beschafft und in der richtigen Reihenfolge miteinander verknüpft werden. Der Umfang der Daten und die Komplexität der einzelnen Schritte machen dabei einen weitgehend automatischen Generierungsprozess erforderlich.

Das von Grund auf neu entwickelte Programm *CityScenarioGenerator* übernimmt diese Aufgabe. Ein vollständiges Städteszenario enthält alle benötigten Informationen über das Simulationsziel und kann genutzt werden, um den LoadProfileGenerator für jeden Haushalt im Gebiet passend zu parametrieren. Dazu werden alle zu berücksichtigenden Wohn- und Nichtwohngebäude mitsamt allen benötigten Eigenschaften aus der Gebäudedatenbank BUILDA [53] ausgelesen. Für jedes Wohngebäude wird eine passende Simulationskonfiguration für den LoadProfileGenerator erstellt, basierend auf den in der Datenbank verfügbaren Angaben. Alle Nichtwohngebäude werden klassifiziert und in die Menge der verfügbaren *Points of Interest* aufgenommen. Aus dieser Menge werden für jede Person in jedem Haushalt einige bestimmte Points of Interest ausgewählt. Die entstehenden Daten werden als Dateien in einer vorgegebenen Ordnerstruktur abgelegt und bilden damit ein vollständiges Städtesimulationsszenario. Dieses Szenario kann von der eigentlichen Städtesimulation eingelesen werden.

#### Zuordnung von Wohn- und Nichtwohngebäuden

Die für die Städtesimulation erforderlichen Eingabedaten lassen sich im Wesentlichen in zwei Typen von Objekten unterteilen, in Wohngebäude und Nichtwohngebäude. Erstere beherbergen die zu simulierenden Personen in ihren Haushalten, letztere dienen als Points of Interest für diverse Außer-Haus-Aktivitäten. Für beide werden die Daten aus BUILDA entnommen, da diese Datenbank Angaben zu sämtlichen Gebäuden in Deutschland enthält. Da alle Gebäude bereits als Wohn- oder Nichtwohngebäude klassifiziert sind, werden sie aufgrund ihrer unterschiedlichen Verwendung in der Städtesimulation getrennt verarbeitet. Bei gemischt genutzten Gebäuden werden die beiden Verwendungszwecke separat behandelt.

Sowohl Wohn- als auch Nichtwohngebäude werden einzeln in eine für die Simulation geeignete Form übertragen. Bei Wohngebäuden bedeutet das die passende Abbildung aller Haushalte im LoadProfileGenerator. Für Nichtwohngebäude muss ein Point-of-Interest-Typ gewählt werden, damit sie bei der Verhaltenssimulation durch die Agenten auf eine für ihren Verwendungszweck angemessene Weise besucht werden.

#### Zielort-Auswahl der Agenten

Eine wichtige Komponente der Städtesimulation bilden die Entscheidungen der Personen, welche Points of Interest sie aufsuchen wollen. Bei der zugrundeliegenden Modellierung dieses Aspekts gibt es im Wesentlichen zwei Möglichkeiten, die sich darin unterscheiden, ob die Auswahl relevanter Points of Interest für jede Person während oder vor der Simulation durchgeführt wird. Bei dem ersten Ansatz wählt jeder Agent wie bisher Aktivitäten aus seiner Aktivitätsliste aus,

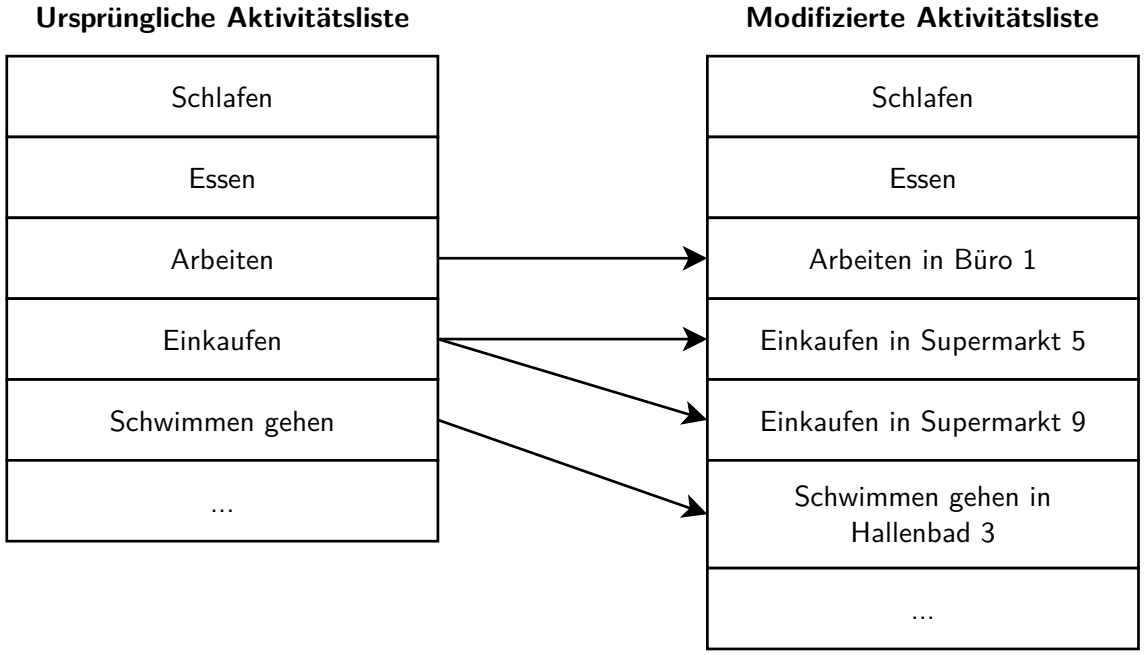



**Abbildung 3.2:** Alle Außer-Haus-Aktivitäten einer Person werden zu Beginn der Städtesimulation angepasst und durch Varianten für einen spezifischen Point of Interest ersetzt. Dabei sind auch mehrere gleiche Aktivitäten möglich, wenn diese an unterschiedlichen Orten ausgeübt werden.

und bei Außer-Haus-Aktivitäten wird dynamisch ein Zielort ermittelt. Diese Variante bringt jedoch einige Nachteile mit sich. Aktivitätsentscheidungen finden sehr häufig statt, und bei jeder dieser Entscheidungen alle verfügbaren Points of Interest des passenden Typs zu überprüfen wäre rechenaufwendig. Außerdem können verschiedene Points of Interest desselben Typs unterschiedliche Eigenschaften haben, wie zum Beispiel verschiedene Öffnungszeiten, die ebenfalls bei jeder Auswahl berücksichtigt werden müssten. Da Points of Interest im Rahmen der Parallelisierung auf mehrere Prozesse verteilt werden, ist dafür entweder eine duplizierte Datenstruktur mit Informationen über alle Points of Interest notwendig, die für jeden Prozess kopiert wird, oder es muss eine Vielzahl einzelner Nachrichten zur Informationsabfrage verschickt werden. Zusätzlich würde ohne weitere Einschränkung die ständige Auswahl neuer Zielorte gefördert. In der Realität besuchen Menschen jedoch meist nur eine kleine Menge gewohnter Zielorte regelmäßig [54]. Aus diesen Gründen ist der beschriebene Ansatz nicht praktikabel.

Der zweite Ansatz umgeht dieses Problem dadurch, dass für jede Person im Vorhinein relevante Points of Interest festgelegt werden. Die Liste der verfügbaren Aktivitäten jeder Person wird dann beim Simulationsstart anhand dieser Ziele, wie in Abbildung 3.2 gezeigt, modifiziert, sodass für jeden Point of Interest eine eigene Version sämtlicher dort angebotenen Aktivitäten enthalten ist. Auf diese Weise speichert jeder Agent nur Informationen zu für ihn relevanten Zielen. Zudem können die existierenden Prüfmechanismen, ob eine Aktivität gerade verfügbar ist, ausgenutzt werden, um etwa die Öffnungszeiten von Geschäften zu berücksichtigen, sofern diese bekannt sind. Dieser Ansatz ist weniger flexibel als der erste, da die gewählten Points of Interest für die gesamte Simulationsdauer festgesetzt sind und nicht verändert werden können. Da diese Modellierung jedoch für diesen Anwendungsfall realistisch genug ist und der Rechenaufwand so deutlich geringer ausfällt, wurde der zweite Ansatz gewählt.

#### Definition von Routen

Für die Simulation der Wege zwischen den verschiedenen Orten, die jeder Agent aufsucht, müssen Routen mit korrekten Distanzen und Reisedauern vorgegeben werden. Um die Erreichbarkeit aller wichtigen Zielorte zu gewährleisten und Personen im Hinblick auf direkte Reisen zwischen verschiedenen Orten nicht einzuschränken, muss ein vollständiges Routennetzwerk für alle Zielorte jeder einzelnen Person angelegt werden. Jede Route besteht aus zwei IDs, einer für den Startpunkt und einer für den Endpunkt. Für eine genauere Abbildung von Wegstrecken und Verkehr werden Routen gerichtet modelliert, es muss also zu jeder Route eine passende Rückroute geben. Alternativ kann zur Vereinfachung über ein Attribut definiert werden, dass alle Routen in beide Richtungen gelten sollen. Dann wird beim Import in der Städtesimulation jedes Routenobjekt zusätzlich mit vertauschtem Start und Ziel erzeugt.

Um die größtmögliche Flexibilität hinsichtlich Reisen und Aktivitätswahl zu bieten, werden alle Orte gesammelt, die von einem Agenten genutzt werden, und Routen von jedem dieser Orte zu jedem anderen erstellt. Somit kann jeder Agent jederzeit alle für ihn relevanten Orte aufsuchen, und muss nicht etwa zuerst nach Hause zurückkehren. Das ermöglicht insbesondere die für die Transportmodellierung wichtigen Wege-Ketten, die beispielsweise auftreten, wenn eine Person auf dem Rückweg von der Arbeit einkaufen geht. Bei der Erzeugung der Routen dürfen immer nur die Orte berücksichtigt werden, die für eine einzelne Person von Bedeutung sind, und nicht alle für einen Haushalt relevanten Orte auf einmal. Letzteres würde zwar ermöglichen, alle für den Haushalt benötigten Routen in einem Schritt zu erzeugen. Da die Anzahl an Routen in einem vollständig verknüpften Netzwerk aber quadratisch mit der Anzahl an Knoten steigt und viele der Orte nur für jeweils eines der Haushaltsmitglieder relevant sind, würde so eine Vielzahl ungenutzter Routen generiert. Das wird in Abbildung 3.3 anhand eines Minimalbeispiels illustriert. Bei einer größeren Anzahl an Orten kann der Anteil irrelevanter Routen in Mehrpersonenhaushalten erheblich sein.

### 3.3.2 Parallele Massensimulation

Die grundlegende Architektur für die Städtesimulation ist aus verschiedenen Gründen ein kritischer Punkt. Die Abbildung ganzer Städte mit bis zu 100.000 Einwohnern geht mit einem erheblichen Rechenaufwand einher. Der LoadProfileGenerator benötigt für die Simulation eines Haushalts für die Zeitspanne von einem Jahr mit minimaler Ergebnisausgabe etwa eine Minute auf einem üblichen Desktop PC. Für eine ganze Stadt ergibt sich damit eine Gesamtrechenzeit von fast 70 Tagen. Daher ist eine effiziente und gut skalierende Implementierung von äußerster Wichtigkeit. Um verfügbare Rechnerstrukturen bestmöglich ausnutzen zu können, ist eine flexible Parallelisierung der Simulation notwendig. Zudem muss die Architektur für Erweiterungen offen gestaltet sein, damit zukünftig weitere Untermodelle hinzugefügt und der Realismus der Simulation verbessert werden können.

#### Wahl eines geeigneten Frameworks

Für die Parallelisierung der Simulation musste ein geeignetes Framework gewählt werden. Dabei bestand die Möglichkeit, entweder ein generisches Parallelisierungsframework oder ein Framework zur agentenbasierten Modellierung einzusetzen. Da der LoadProfileGenerator in der Sprache C# entwickelt wurde und der Aufwand für eine Neuimplementierung in einer anderen Sprache un-

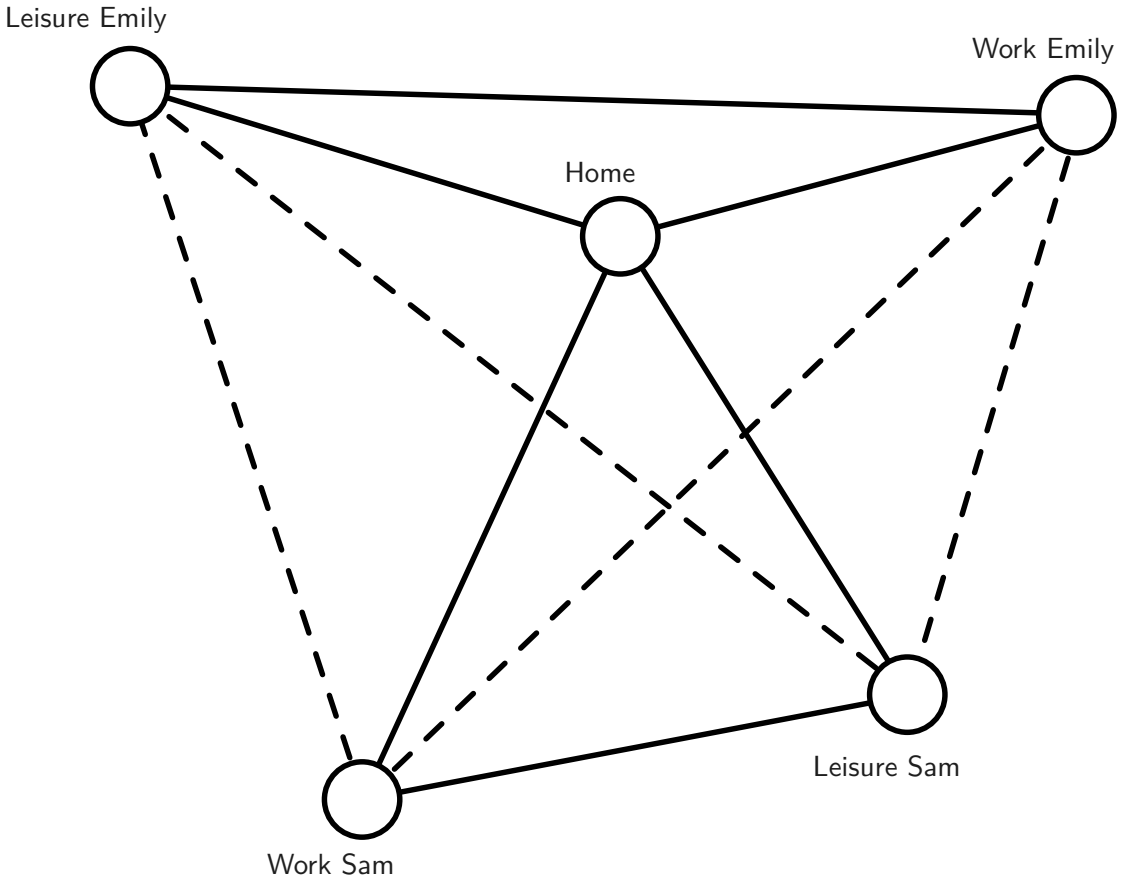



**Abbildung 3.3:** Beispiel für das Routennetzwerk eines Zwei-Personen-Haushalts. Beide Haushaltsmitglieder haben eigene Arbeitsplätze und Orte für Freizeitaktivitäten, die für die jeweils andere Person nicht relevant sind. Wenn für jede Person individuell die notwendigen Routen ermittelt werden anstatt für alle zusammen, können ungenutzte Routen (hier gestrichelt dargestellt) ausgelassen werden.

verhältnismäßig gewesen wäre, kamen nur Frameworks in Betracht, die eine performante und funktional umfangreiche Einbindung von C#-Programmen gestatten.

Zwar sind zahlreiche agentenbasierte Lösungen verfügbar, davon sind allerdings nur wenige leistungsfähig genug für die erforderliche Größenordnung. Von diesen Einschränkungen abgesehen besteht zudem grundsätzlich die Problematik, dass viele Frameworks Rahmenbedingungen für die Modellimplementierung vorgeben, die insbesondere die freie Einbindung eines existierenden Programms zur Abbildung des Agentenverhaltens erschweren oder sogar ausschließen. Aus diesem Grund wurde diese Möglichkeit verworfen.

Im Bereich der generischen Parallelisierungsframeworks ist das Message Passing Interface (MPI) ein etablierter und weit verbreiteter Standard. MPI umfasst dabei allgemein Interprozesskommunikation, also den Austausch von Nachrichten zwischen Prozessen auf verteilten Computersystemen. Es stellt Funktionen zum Senden und Empfangen von Daten bereit, aber darüber hinaus keine Hilfsmittel für die Koordination der Prozesse oder des Simulationsablaufs. Die Logik hierfür muss daher vollständig selbst implementiert werden. Dafür gibt es keinerlei Einschränkungen hinsichtlich der Programmstruktur, sodass der LoadProfileGenerator als Verhaltensmodell für die Bewohner problemlos eingebunden werden kann. Mit MPI.NET [55] existiert eine gut nutzbare und hin-

reichend vollständige Bibliothek für die Einbindung in C#-Projekte. Die Details der Einbindung von MPI.NET in den Bauprozess unter verschiedenen Systemen sind in Anhang A.1 beschrieben. Dadurch ist die Nutzung von MPI in einem Projekt mit dem LoadProfileGenerator ohne großen Aufwand möglich. Aus diesen Gründen stellte MPI die beste Option für die entwickelte Städtesimulation dar und wurde als zugrundeliegendes Framework verwendet.

#### Anwendung von MPI

Bei der Verwendung von MPI werden mehrere eigenständige Prozesse des gewünschten Programms gestartet, die alle denselben Programmcode ausführen. Diese Prozesse werden als Worker bezeichnet. Über Fallunterscheidungen und die Auswahl von Daten in Abhängigkeit von einer von MPI vergebenen Prozessnummer, des sogenannten Ranks, kann dafür gesorgt werden, dass alle Prozesse unterschiedliche Aufgaben erledigen. Für die Kommunikation zwischen zwei oder mehr Prozessen müssen alle Beteiligten die passende MPI-Funktion aufrufen. Standardmäßig blockieren alle MPI-Funktionen, bis alle beteiligten Prozesse den Funktionsaufruf erreicht haben, für einige Funktionen existieren jedoch auch nicht-blockierende Varianten, die unter Umständen effizientere Programmabläufe ermöglichen.

### 3.3.3 Struktur der Städtesimulation

Für die parallele Simulation einer Stadt werden mittels MPI mehrere Worker gestartet. Da alle Haushalte und Nichtwohngebäude individuell simuliert werden, ist eine feste, gleichmäßige Verteilung aller Gebäude auf die Worker die naheliegendste Lösung. Dadurch muss jedes Gebäude immer nur auf dem dafür zuständigen Worker initialisiert und gespeichert werden. Eine andere Möglichkeit besteht in der zentralen Speicherung aller Gebäude in einer Datenbank mit dynamischer Abarbeitung durch die Worker. Diese Variante wäre unter Umständen flexibler und würde automatisch eine gleichmäßige Lastverteilung und damit optimale Ressourcenausnutzung bieten. Im Gegenzug wären viele Speicherzugriffe zum Versenden aller Gebäudedaten notwendig, was sowohl Speicher als auch Netzwerk des Rechenclusters belasten würde. Da die Simulation eines Haushalts für einen einzelnen Zeitschritt nur wenige Millisekunden dauert, die dafür benötigten Daten aber bei Größen im Bereich einiger hundert Megabyte liegen und darüber hinaus auf komplexe Datenstrukturen verteilt sind, würde das bloße Laden der Daten in diesem Ansatz einen kaum zu umgehenden Flaschenhals für die gesamte Simulation darstellen. Selbst in einem optimistischen Szenario mit einer Übertragungsrate von 10 GBit/s und perfekter Parallelisierung der Übertragung ergibt sich bei 100 Workern eine Übertragungszeit von bis zu 40 Sekunden pro Zeitschritt, während die reine Simulationsdauer für einen Zeitschritt unter einer halben Sekunde liegt. Dieser Ansatz ist somit in Bezug auf die Rechenzeit sehr ineffizient. Daher wurde diese Option nicht weiterverfolgt und stattdessen die feste Verteilung der Gebäude umgesetzt.

#### Synchronisation der Simulation

Das Hauptziel der Städtesimulation ist die korrekte Abbildung von gegenseitigen Abhängigkeiten und Einflüssen der einzelnen Personen aufeinander, auch über die Grenzen eines Haushalts hinaus. Durch die Parallelisierung sind die Informationen auf alle Worker verteilt und unter Umständen nicht beliebig abruf- oder änderbar. Um die Auswirkungen des Verhaltens von Personen aus unter-

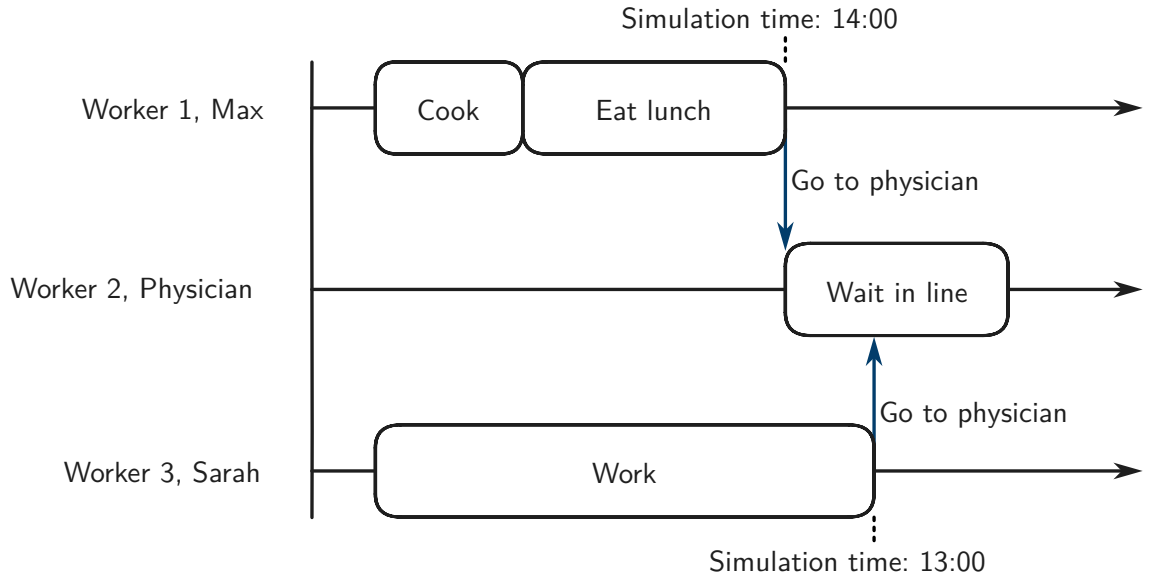



**Abbildung 3.4:** Probleme ohne Synchronisation der Worker. Wenn Worker nicht exakt gleich schnell simulieren, kann die Plausibilität der Simulation nicht sichergestellt werden. In diesem Beispiel wäre Max in der Warteschlange vor Sarah, obwohl Sarah zeitlich früher zum Arzt gegangen ist.

schiedlichen Haushalten insgesamt konsistent berücksichtigen zu können, ist jedoch wichtig, dass es nicht zu Widersprüchen kommt und nicht mit veralteten Daten gearbeitet wird. Andernfalls können bestimmte Situationen in der Simulation nicht korrekt behandelt werden, wie in Abbildung 3.4 anhand eines Beispiels gezeigt wird. Worker 1 und Worker 3 rechnen unterschiedlich schnell, daher bewegt sich die Simulationszeit auseinander, sodass Worker 1 bereits eine Stunde weiter fortgeschritten ist. Dadurch landet Max in der Warteschlange beim Arzt auf Worker 2 vor Sarah, obwohl Sarah innerhalb der Simulation bereits eine Stunde zuvor zum Arzt gegangen ist.

Um derartige Situationen zu vermeiden, muss die gesamte Simulation regelmäßig synchronisiert werden. Als eingebundenes Verhaltensmodell simuliert der LoadProfileGenerator immer einen Haushalt oder ein ganzes Haus auf einmal. Änderungen, die auch andere Worker betreffen, umfassen dabei in erster Linie Bewegungen von Agenten, beispielsweise den Aufbruch einer Person zur Arbeit. Diese Änderungen können in jedem Zeitschritt auftreten, daher muss nach jedem Zeitschritt eine Synchronisierung stattfinden. Die Städtesimulation wurde daher so aufgebaut, dass alle Haushalte im Zielgebiet stets nur für einen Zeitschritt auf einmal simuliert werden (siehe Abbildung 3.5). Danach werden Nachrichten ausgetauscht, um alle betroffenen Worker über eventuelle Änderungen zu informieren. Anschließend wird mit dem nächsten Zeitschritt fortgefahren. Diese Prozedur wird wiederholt, bis alle Zeitschritte im vorgegebenen Zeithorizont abgearbeitet wurden. Daraufhin ist der Hauptteil der Simulation abgeschlossen.

Abgesehen von der Synchronisierung in jedem Zeitschritt arbeiten alle Worker unabhängig voneinander. Insbesondere speichern sie ihre Ergebnisse selbstständig auf einem geteilten Speichersystem. Dadurch wird vermieden, dass große Ergebnisdatenmengen über MPI verschickt werden müssen, was aufgrund beschränkter Bandbreite viel Zeit in Anspruch nähme. Der LoadProfileGenerator verfügt über eine Vielzahl verschiedener Ergebnisoptionen, die individuell aktiviert oder deakti-

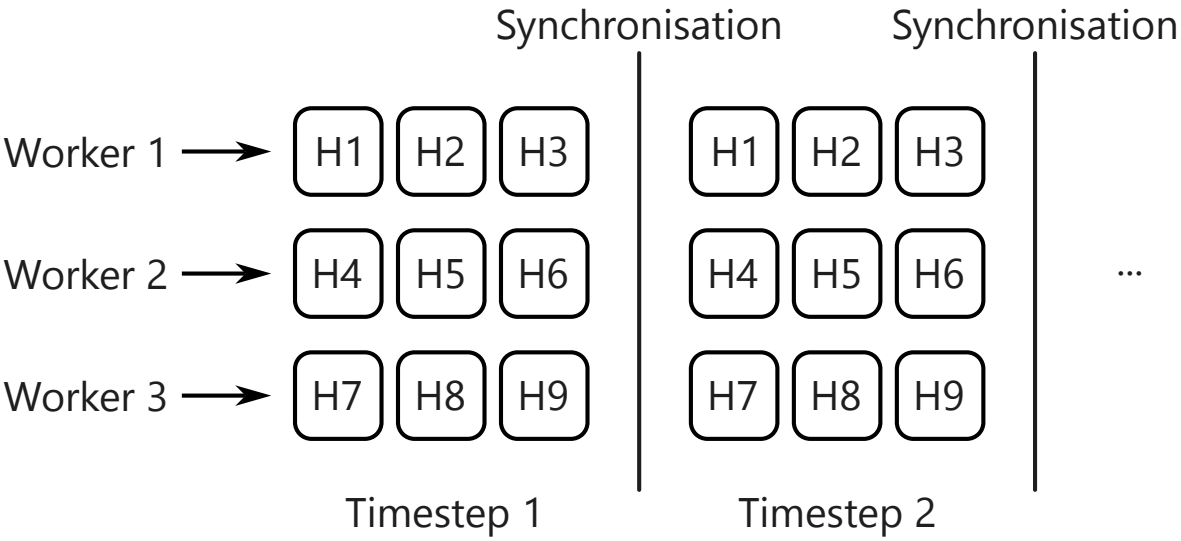



**Abbildung 3.5:** Verteilung der Gebäude und Synchronisierung während der Simulation

viert werden können. Für die Untersuchungen in dieser Arbeit wurden Aktivitätsprofile, Lastprofile auf Haus- und Haushaltsebene sowie der Aufenthaltsort jeder Person in jedem Zeitschritt als Ausgabe gewählt. Selbst in diesem reduzierten Umfang liegt die Gesamtgröße der Ergebnisdateien pro Haushalt bei durchschnittlich 100 MB. Bei der Simulation einer Stadt wie Jülich mit etwa 16.000 Haushalten werden somit über 1,5 TB an Ergebnisdaten produziert, daher ist eine effiziente Handhabung der Daten unumgänglich.

#### Verteilung der Arbeitslast

Der Vorteil der beschriebenen Synchronisierungsmethode liegt in der Sicherheit, dass alle Daten zu jeder Zeit konsistent bleiben und Agenten ihre Entscheidungen stets auf Basis aktueller Informationen treffen können. Nachteilig ist, dass schnellere Worker unter Umständen auf langsamere warten müssen, wodurch die Performance beeinträchtigt wird. Allerdings wird erwartet, dass für die Städtesimulation nicht viel mehr als 100 Worker eingesetzt werden, da jeder Worker mindestens einen eigenen Prozessorkern erfordert und die gesamte Simulation sonst zu viele Hardwareressourcen reservieren würde. Die Anzahl simulierter Gebäude liegt dagegen wesentlich höher, im Fallbeispiel Jülich etwa gibt es fast 10.000 Wohngebäude. Somit werden im Normalfall mindestens 100 Gebäude auf einem Worker simuliert, und es ist zu erwarten, dass sich Unterschiede bei der Dauer eines Simulationsschritts größtenteils ausgleichen. Das gilt allerdings nur für homogene Rechnerarchitekturen, also wenn alle Worker auf ähnlich leistungsfähigen Maschinen laufen. Zudem kann der Rechenaufwand für die Simulation einzelner Gebäude unter Umständen stark variieren, da der Aufwand in erster Linie von der Anzahl an Bewohnern abhängt. Ein Einfamilienhaus mit nur einem Bewohner kann somit um ein Vielfaches schneller simuliert werden als beispielsweise ein Wohngebäude mit 10 oder mehr Parteien. Um zu vermeiden, dass durch Zufall viele Häuser mit besonders vielen Bewohnern denselben Workern zugewiesen werden und sich merkliche Unterschiede in der Arbeitslast ergeben, wird eine in Bezug auf die Gesamtanzahl simulierter Personen möglichst faire Verteilung angestrebt. Da ein Haus nicht weiter aufgeteilt werden kann und immer als Ganzes einem Worker zugeordnet wird, existiert hierfür nicht immer eine perfekte Lösung. Daher werden die Häuser heuristisch mithilfe eines Greedy-Algorithmus verteilt, indem sie

absteigend nach Bewohnerzahl sortiert und dem Worker mit der bisher geringsten Personenzahl zugeteilt werden. In der Praxis gibt es in vielen Städten einen großen Anteil an Häusern mit nur einem Bewohner, im Anwendungsbeispiel Jülich macht diese Kategorie beispielsweise fast 18% aller Gebäude und 5% aller Bewohner aus. In solchen Fällen reicht das Greedy-Verfahren aus, um eine optimale Verteilung zu erhalten.

#### Initialisierung der Simulation

Wie in Abbildung 3.1 zu sehen ist, beginnt die Städtesimulation mit einem obligatorischen nichtparallelen Teil. Dieser umfasst das Einlesen der Konfiguration, Verteilen der Gebäude sowie einige notwendige Initialisierungen, die nicht parallelisiert werden können, und wird ausschließlich von einem Worker ausgeführt. Dazu gehört auch der Aufbau eines Registers, das speichert, auf welchem Worker jeder Point of Interest simuliert wird. Ohne dieses Register können Nachrichten, die einen bestimmten Point of Interest betreffen, später nicht an den richtigen Worker adressiert werden. Außerdem müssen die Zufallszahlengeneratoren auf jedem Worker korrekt initialisiert werden.

Zufallszahlengeneratoren werden an verschiedenen Stellen der Simulation für probabilistische Effekte eingesetzt. Um die gesamte Simulation reproduzierbar zu halten, sodass dieselbe Eingabe stets zu exakt denselben Ergebnissen führt, kann in der Konfiguration ein Seed angegeben werden. Ein Seed ist eine Zahl oder Zeichenfolge, die einen Zufallszahlengenerator konfiguriert, sodass er eine bestimmte Zahlenfolge ausgibt. Derselbe Seed resultiert dabei immer in exakt derselben Zufallszahlenfolge. Wenn also mehrere Simulationsdurchläufe mit demselben initialen Seed gestartet werden, wird an jeder probabilistischen Stelle in der Simulation stets derselbe Wert erzeugt und die gesamten Simulationsergebnisse sind identisch. Das funktioniert allerdings nur, solange in der gesamten Simulation ausschließlich auf einen einzigen Zufallszahlengenerator zugegriffen wird. In einer parallelen Architektur mit getrennten Prozessen ist das nicht möglich, hier muss jeder Prozess seinen eigenen Zufallszahlengenerator erzeugen. Dabei darf nicht einfach jeder Prozess den in der Konfiguration vorgegebenen Seed nutzen, da ansonsten die Zufallszahlengeneratoren in allen Prozessen die gleichen Zahlenfolgen erzeugen würden, was zu unerwünschten Nebeneffekten führen kann. Jeder Prozess braucht eine eigene Zufallszahlenfolge, die jedoch letztendlich von dem initialen Seed in der Eingangskonfiguration abhängt. Dieses Problem kann gelöst werden, indem ein einzelner Worker zu Beginn einen gesonderten Zufallszahlengenerator mit dem Seed aus der Eingabe parametriert, und diesen dann nutzt, um einen zufälligen Seed für jedes zu simulierende Gebäude zu generieren. Anschließend wird für jedes Gebäude ein eigener, mit dem entsprechenden Seed konfigurierter Zufallszahlengenerator verwendet. Somit erhält die Simulation jedes Gebäudes eine eigene Zahlenfolge, die jedoch vom initialen Seed abhängt, wodurch die Ergebnisse vollständig reproduzierbar sind. Dieses Verfahren ist auch nicht von der Anzahl an Workern abhängig, sodass Simulationen mit unterschiedlich vielen Workern dennoch identische Ergebnisse liefern.

#### Hauptteil der Simulation

Nach diesen Schritten beginnt der parallele Teil, den jeder Worker gleichermaßen durchläuft. Neben dem Initialisieren der benötigten Simulator-Klassen muss auch für jedes Wohngebäude ein erster Vorbereitungsschritt im LoadProfileGenerator durchlaufen werden. In der Simulationskonfiguration werden Haushalte über sogenannte *Templates* angegeben. Diese benutzt der Load-

ProfileGenerator, um Modelle ähnlicher, aber nicht identischer Haushalte eines bestimmten Typs zu erstellen und für die Simulation zu nutzen. Auf diese Weise wird vermieden, dass es zu Artefakten in den generierten Daten kommt, etwa weil zu viele Haushalte zu exakt gleicher Uhrzeit kochen. Diese Generierung der Haushaltsmodelle aus den Templates muss daher noch vor Beginn der eigentlichen Simulation geschehen.

Abbildung 3.6 zeigt die wesentlichen Abläufe im Hauptteil der Städtesimulation, die auf jedem Worker gleichzeitig stattfinden. In der Hauptschleife werden nacheinander alle Zeitschritte abgearbeitet. Jeder Zeitschritt beginnt mit der Simulation der Haushalte, der reisenden Agenten und der Points of Interest über jeweils dafür zuständige Simulator-Klassen. Diese liefern dabei gegebenenfalls Nachrichten an andere Simulatoren zurück. Um gleiche Konditionen zu schaffen, werden diese Nachrichten jedoch nicht direkt verarbeitet, auch wenn das möglich wäre, solange sich der adressierte Transport- oder Point-of-Interest-Simulator auf demselben Worker befindet. Stattdessen werden alle Nachrichten gesammelt und am Ende des Zeitschritts auf einmal ausgetauscht. Jeder Worker kann dabei potenziell Nachrichten für jeden anderen Worker produzieren. Daher wird dieser Austausch mit der MPI-Funktion *Alltoall* umgesetzt, die die Verteilung aller Nachrichten in einem Schritt erledigt. Diese Funktion muss aufgrund der Art der Kommunikation in jedem Worker blockieren, der schnellste Worker muss also immer warten, bis auch der letzte Worker den Funktionsaufruf erreicht hat. In Testsimulation von 1000 Häusern mit 12 Workern hat sich gezeigt, dass in etwa 84 % der Fälle ein Worker für einen bestimmten anderen Worker keine Nachrichten hat. Da im Voraus jedoch nicht bekannt ist, welcher Worker Nachrichten an welche anderen senden möchte, ist eine Reduzierung des Kommunikationsaufwands durch gezielten paarweisen Nachrichtenaustausch nicht möglich.

### 3.3.4 Außer-Haus-Aktivitäten

Der LoadProfileGenerator modelliert die Bedürfnisse jeder Person und bestimmt darüber, welche Aktivität diese jeweils als nächstes ausführt. Auf diese Weise entsteht nach und nach ein Aktivitätsprofil. Dabei werden innerhalb des Hauses alle Daten verwaltet, die für die Aufrechterhaltung der Rahmenbedingungen notwendig sind. So werden genutzte Haushaltsgeräte für die Dauer des Gebrauchs markiert, sodass kein anderer Haushaltsbewohner zur gleichen Zeit darauf zugreifen kann. Normalerweise reicht diese haushaltsinterne Überprüfung aus. Im Falle der Städtesimulation müssen allerdings alle Außer-Haus-Aktivitäten gesondert betrachtet werden, da hier äußere Einflüsse eine Rolle spielen können. Dazu zählen sowohl Aktivitäten an anderen Orten, wie zum Beispiel dem Arbeitsplatz, als auch die dazugehörigen Wege. Diese müssen bei der Verhaltenssimulation dementsprechend gesondert behandelt werden.

#### Simulation außerhalb des Haushalts

Um die neuen Randbedingungen zu implementieren, kommen zwei grundsätzlich verschiedene Optionen in Betracht. Ein Agent könnte bei jeder Bewegung logisch dem Worker, der für seinen neuen Standort zuständig ist, zugeordnet werden. Dieser Worker könnte dann zum Ende des Aufenthalts eine Folgeaktivität bestimmen. Alternativ verbleibt die Verantwortung für einen Agenten und dessen Entscheidungen durchgängig bei dem Worker, der für den Haushalt des Agenten zuständig ist. Die erste Variante hat den Vorteil, dass keine weitere Nachricht über das Ende jeder

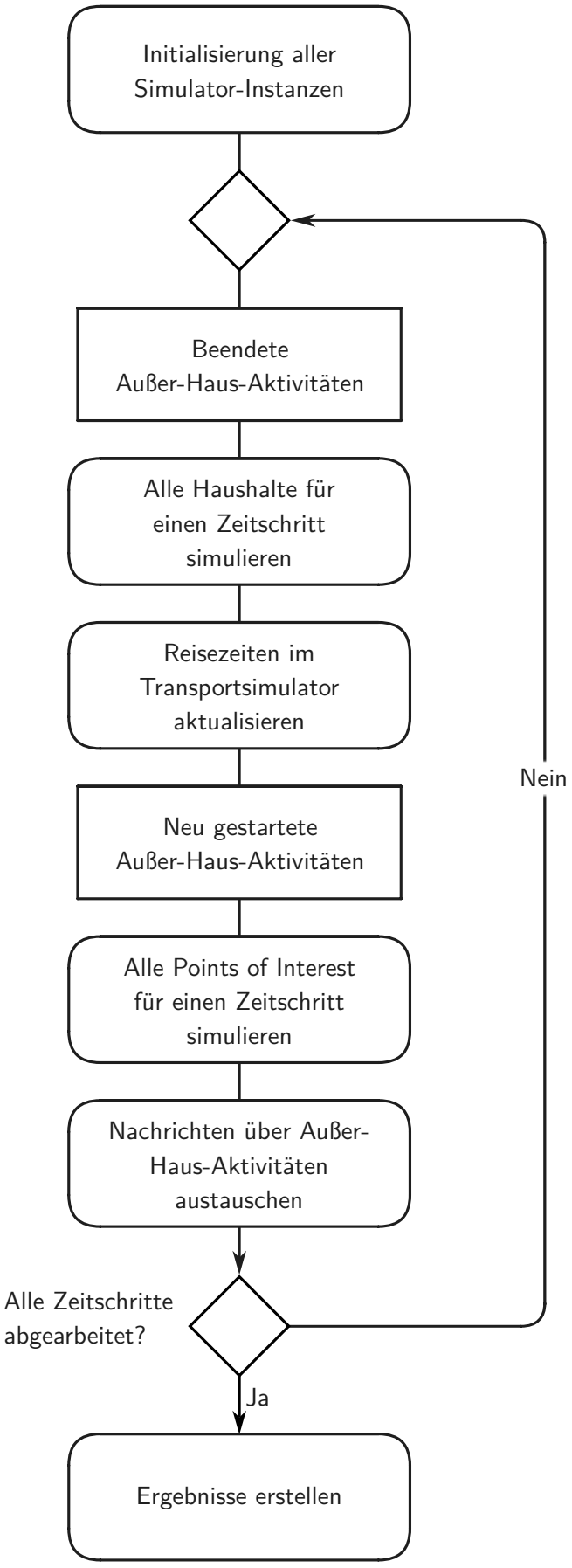



**Abbildung 3.6:** Wesentliche Schritte im Ablauf der Städtesimulation auf jedem Worker

Außer-Haus-Aktivität notwendig ist. Allerdings müssen sämtliche für die Aktivitätswahl notwendigen Daten stets an den aktuell zuständigen Worker geschickt werden. Das umfasst den Agenten selbst, aber auch die verfügbaren Geräte, Daten über andere Haushaltsmitglieder und allgemeine Bedingungen. Einerseits wird dadurch der Kommunikationsaufwand drastisch erhöht. Andererseits kann aber auch die Sicherung der Konsistenz nicht mehr gewährleistet werden. Es wäre beispielsweise denkbar, dass zwei Mitglieder desselben Haushalts sich beide gleichzeitig an unterschiedlichen Orten außer Haus befinden und ihre nächste Aktivität wählen müssen. Selbst wenn auf ihren jeweiligen Workern alle Informationen über nutzbare Geräte und verfügbare Aktivitäten bereitstehen, kann nicht verhindert werden, dass beide dieselbe Aktivität wählen und somit ein Gerät doppelt genutzt wird. Bei der lokalen Simulation mehrerer Agenten auf demselben Worker kann diese Situation ausgeschlossen werden, indem einer der Agenten zuerst wählt und damit Vorrang vor dem anderen hat. Da aber innerhalb eines Zeitschritts keine weitere Kommunikation zwischen zwei Workern stattfindet, ist diese Form der Kontrolle nicht möglich. Zur Bewältigung dieses Problems kann lediglich eine vorläufige Aktivitätswahl durchgeführt werden, die bei der Ankunft am Zielort erneut geprüft wird. Das sichert zwar die Konsistenz, führt aber unter Umständen zu unnötigen Reisen, da Agenten im Falle einer doch nicht verfügbaren Aktivität als Ersatz möglicherweise eine Aktivität an einem anderen Ort wählen. Zudem wird das Problem des hohen Kommunikationsaufwands nicht gelöst. Insgesamt ist dieser Ansatz daher ungeeignet.

Die zweite Variante ist von diesen Problemen nicht betroffen. Bei einer Außer-Haus-Aktivität wird nicht der gesamte Agent verschickt, sondern lediglich eine kurze Nachricht, um die betroffene Instanz darüber zu informieren. Im Falle einer Reise ist das der zuständige Transportsimulator, der die Berechnung der Reisedauer übernimmt. Bei Aktivitäten an anderen Orten erhält der Point-of-Interest-Simulator des jeweiligen Ortes die Nachricht, um den Aufenthalt der Person simulieren zu können. Nach Abschluss einer Außer-Haus-Aktivität muss dann der zuständige Simulator wieder den Haushaltssimulator des Agenten benachrichtigen, damit dieser eine neue Aktivität auswählen kann. Das betrifft sowohl Transportsimulatoren bei Reisen als auch Point-of-Interest-Simulatoren bei dortigen Aktivitäten. Diese Nachrichten müssen stets korrekt versendet und übermittelt werden, damit für jeden Agenten ein lückenloses Aktivitätsprofil erstellt werden kann. Der Vorteil dieser Variante besteht in den geringeren zu versendenden Datenmengen, sowie darin, dass die gesamte Modellstruktur für einen Agenten und seinen Haushalt nur auf einem Worker aufgebaut und verwaltet werden muss. Aus diesen Gründen wurde der zweite Ansatz gewählt.

Bei den meisten alltäglichen Aktivitäten wie Schlafen oder Essen fallen keine Nachrichten an, da sie ausschließlich innerhalb des eigenen Haushalts stattfinden. Wenn aber ein Agent beispielsweise arbeiten geht, müssen dafür mindestens vier Nachrichten versandt werden. In Abbildung 3.7 ist dieses Szenario exemplarisch dargestellt. Der Agent in diesem Beispiel wird auf Worker 1 simuliert. Eine Aktivitätswahl durch den LoadProfileGenerator wird dabei durch eine rote Raute symbolisiert. Nach einigen lokalen Aktivitäten entscheidet sich der Agent für die Aktivität „Arbeit“, muss dafür jedoch zunächst zu seinem Arbeitsplatz fahren. Daher wird die entsprechende Reise initiiert und eine Nachricht an den Transportsimulator auf demselben Worker geschickt. Dieser berechnet die Reisedauer und informiert den zuständigen Haushaltssimulator, sobald der Agent sein Ziel erreicht hat. Der Haushaltssimulator prüft nun die geplanten Aktivitäten und startet, wenn es zu keinen Abweichungen gekommen ist, die Aktivität „Arbeit“. Dafür wird eine

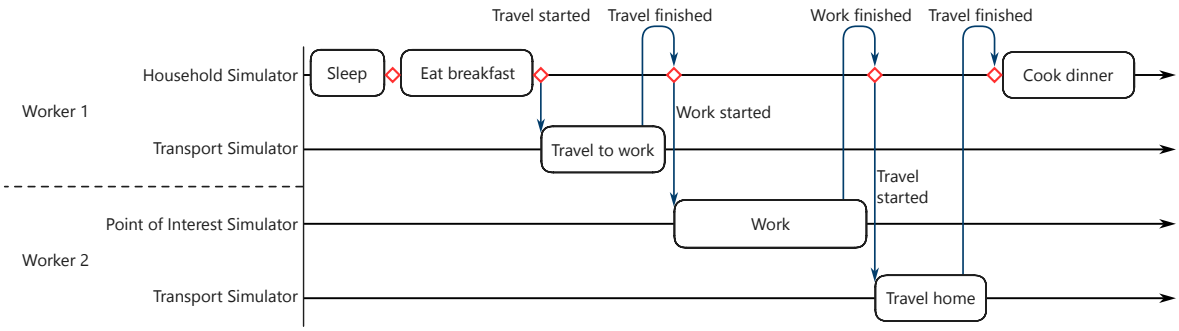



**Abbildung 3.7:** Nachrichten zum Austausch eines Agenten, der eine Außer-Haus-Aktivität ausführt. Jeder blaue Pfeil stellt eine Nachricht an den entsprechenden Simulator dar, jede rote Raute eine simulierte Aktivitätswahl des Agenten.

Nachricht an den zuständigen Point-of-Interest-Simulator versandt, der daraufhin die Aufenthaltsdauer des Agenten ermitteln kann. Der korrekte Zielworker wird dabei über das zuvor angelegte Point-of-Interest-Register ermittelt. Sobald die Aktivität beendet ist, wird wiederum eine Nachricht an den Haushaltssimulator gesendet, der nun die nächste Aktivität und eine gegebenenfalls dafür notwendige Reise bestimmt. Der Übersichtlichkeit halber stellt Abbildung 3.7 lediglich einen Ausschnitt dar. Für gewöhnlich ist in einer Städtesimulation jeder Worker für mehrere Haushalts- und Point-of-Interest-Simulatoren zuständig.

### Aktivitäten im LoadProfileGenerator

Eine grundlegende Eigenschaft von Außer-Haus-Aktivitäten ist ihre unbekannte Dauer. Im gesamten Verhaltensmodell des LoadProfileGenerators wurde grundsätzlich davon ausgegangen, dass die Dauer einer jeden Aktivität im Vorhinein bekannt ist. Zwar variiert die Dauer aller Aktivitäten entsprechend einer Normalverteilung mit festgelegten Parametern, die genaue Dauer für eine einzelne Ausübung einer Aktivität wird aber zu deren Beginn mithilfe des globalen Zufallszahlengenerators bestimmt. Somit kann direkt bei der Entscheidung für eine Aktivität festgelegt werden, wie lange die Person beschäftigt ist und wie lange die dazugehörigen Geräte für andere Nutzer blockiert sind.

Diese Annahme kann in der Städtesimulation nicht aufrechterhalten werden. Die Dauer von Außer-Haus-Aktivitäten wie Reisen oder Einkaufen kann nun durch externe Faktoren beeinflusst werden und wird dynamisch ermittelt, möglicherweise sogar in einem anderen Worker. Daher müssen für sämtliche damit zusammenhängende Abfragen und Kontrollen zur Plausibilität im Kern der Verhaltenssimulation neue Mechanismen etabliert werden. Diese basieren in erster Linie darauf, dass Agenten und Geräte zu Beginn einer Außer-Haus-Aktivität für unbestimmte Zeit als beschäftigt beziehungsweise blockiert markiert werden. Erst, wenn das Ende der Aktivität signalisiert wurde, können diese Markierungen zurückgesetzt und die Geräte somit wieder freigegeben werden. Zusätzlich können alle mit einer Aktivität verbundenen Programmschritte, die auf die Dauer zugreifen, wie beispielsweise das Aktivitäts-Logging, anders als zuvor erst beim Abschluss der Aktivität ausgeführt werden.

Einige der im Zusammenhang mit einer Aktivität zu erledigenden Aufgaben sind verhältnismäßig komplex. Daher wurde darauf verzichtet, sämtliche betroffenen Modellkomponenten an die neuen

Anforderungen anzupassen, sodass die Änderungen nur dort durchgeführt wurden, wo dies zu Abbildung aller relevanten Situationen erforderlich war. Zu den nicht betroffenen Bereichen gehört in erster Linie das Blockieren genutzter Geräte. Alle tatsächlich relevanten Geräte befinden sich im LoadProfileGenerator innerhalb des Haushalts. Für jegliche Außer-Haus-Aktivitäten, wie beispielsweise Einkaufen, wurden Platzhalter-Geräte genutzt. Bei diesen ist jedoch weder eine Reservierung sinnvoll, noch müssen entstehende Lasten nachverfolgt werden, da diese nicht im betrachteten Haushalt selbst anfallen. Daher wurden diese Geräte ersatzlos entfernt, ohne Einbußen in den Simulationsergebnissen. Somit konnte die Einschränkung gesetzt werden, dass lediglich bei Aktivitäten innerhalb des Haushalts Geräte reserviert und blockiert werden können. Für die gesondert implementierten Transportmittel kann diese Einschränkung nicht gelten, da diese ausschließlich für Außer-Haus-Aktivitäten eingesetzt werden.

### 3.3.5 Simulation der Stadt

Der LoadProfileGenerator als Verhaltensmodell bestimmt lediglich das Verhalten einzelner Personen unter Berücksichtigung aller anwesenden Haushaltsmitglieder. Aufenthalte außerhalb des eigenen Haushalts werden auch mit den durchgeführten Anpassungen und Erweiterungen lediglich statisch simuliert, also mit fest im Vorhinein ermittelter Dauer, ohne externe Einflüsse zu beachten. Um derartige Einflüsse in der Städtesimulation abzubilden, wurden Point-of-Interest-Simulatoren und Transportsimulatoren entwickelt. Diese eigenständigen Klassen sind für die Abbildung aller relevanten Vorgänge an jeweils einem Point of Interest beziehungsweise beim Transport zuständig. Ihre Aufgabe besteht in erster Linie darin, die Aufenthalts- oder Reisedauer jeder Person zu ermitteln und die zuständigen Simulatoren über das Ende von Reisen und Aktivitäten zu informieren.

#### Transportsimulation

Transportsimulation ist ein komplexes Themengebiet mit eigenen Methodiken und Anforderungen. Der LoadProfileGenerator verfügt zwar über eine einfache Simulation von Wegen mit konstanter Länge, allerdings ohne weitere Logik zur Verkehrsmittelwahl oder der Abhängigkeit der Reisedauer von Wochentag und Uhrzeit. Um diesen Bereich dennoch möglichst realistisch abdecken zu können, wird die Möglichkeit einer Modellkopplung in die Städtesimulation eingebaut. Diese Kopplung läuft wie in Abbildung 3.8 dargestellt ab. Zunächst erstellt der CityScenarioGenerator ein Städteszenario mitsamt allen Einwohnern, Gebäuden, sowie den Zielortpräferenzen der Personen. Diese Informationen können von einem Transportmodell ausgewertet und analysiert werden, um präzise Angaben zu Wegstrecken, variablen Reisedauern sowie Auswahlwahrscheinlichkeiten für verfügbare Verkehrsmodi zu liefern. Anhand dieser Daten können anschließend in der Städtesimulation die Parameter zur Abbildung des Reiseverhaltens gesetzt werden. Die im Laufe einer Simulation bestimmten Werte für Reisedauern können somit unverändert genutzt werden, sodass die Transportsimulator-Klasse in erster Linie das Logging der reisenden Personen übernehmen muss.

Durch die explizite Abbildung jeder Reise jeder einzelnen Person bietet die Städtesimulation selbst das Potenzial für eine sehr detaillierte agentenbasierte Mobilitätssimulation. So könnte beispielsweise jede Straße im Verkehrsnetz mit allen dort befindlichen Fahrzeugen individuell modelliert werden, was durch Emergenz gegebenenfalls auch die Simulation von Staus ermöglichen würde.

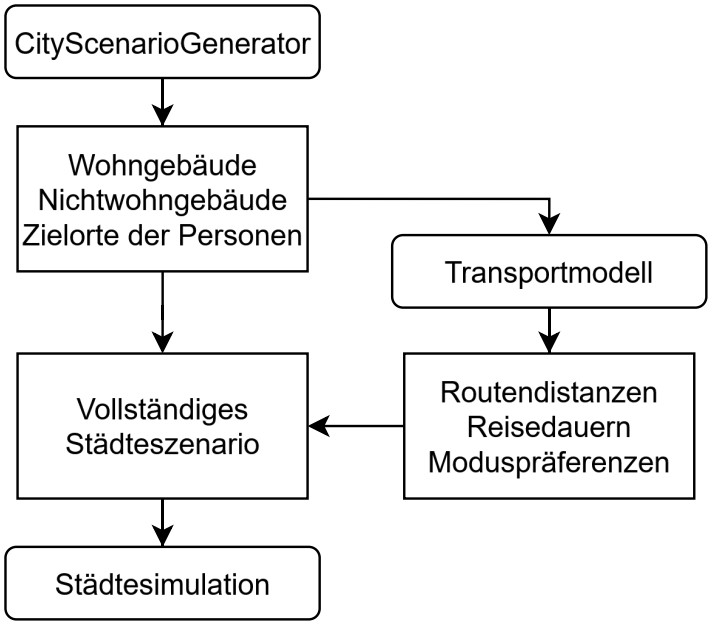



**Abbildung 3.8:** Darstellung der optionalen Kopplung mit einem zusätzlichen Transportmodell. Ohne Transportmodell wird in der Städtesimulation auf Distanzen in Luftlinie und vordefinierte Reisegeschwindigkeiten zurückgegriffen.

Die Architektur für derartige Erweiterungen ist durch die gekapselte Transportsimulation gegeben. Eine solch detaillierte Modellierung liegt jedoch außerhalb des Rahmens dieser Arbeit und stellt lediglich ein Beispiel dar, wozu die hier aufgebaute Plattform zur Städtesimulation in Zukunft genutzt werden könnte.

#### Aufenthaltssimulation für Points of Interest

Die Basisimplementierung der Point-of-Interest-Simulation ist wie die Transportsimulation sehr simpel gehalten. Sie verwaltet lediglich die aktuell anwesenden Personen, loggt entsprechende Informationen und kommuniziert das Ende jeder Aktivität. Als Aufenthaltsdauer wird die vom LoadProfileGenerator vorgegebene Aktivitätsdauer unverändert übernommen. Für viele Point-of-Interest-Typen ist diese Modellierung bereits ausreichend. Bei Sportaktivitäten im Freien ist beispielsweise keine größere Beeinflussung durch andere Personen zu erwarten.

Für einige Point-of-Interest-Typen ist dagegen eine spezifischere Modellierung notwendig, abhängig von der Art und Weise, wie Besuche entsprechender Einrichtungen üblicherweise ablaufen. Der Fokus wird hier auf Points of Interest mit einer Warteschlange gelegt, also auf Orte, an denen Kunden oder Besucher nacheinander abgefertigt werden. Dieses Konzept ist sehr verbreitet und kommt beispielsweise in Supermärkten, Apotheken, Arztpraxen oder Tankstellen vor. Der Einfluss der Warteschlange auf die gesamte Aufenthaltsdauer hängt dabei von verschiedenen Faktoren ab, es wird also eine flexible Modellierung benötigt.

Hierfür wird eine separate Simulator-Klasse erstellt. Diese bildet die Kundenwarteschlange explizit ab, jeder neu ankommende Besucher wird also zunächst in einer Queue-Datenstruktur gespeichert. Zusätzlich gibt es eine konstante, aber pro Instanz frei einstellbare Anzahl an Service-Slots. Diese repräsentieren die Bereiche, in denen Kunden nach dem Warten bedient werden, also beispielsweise

die Kassen in einem Supermarkt. Jeder Slot kann nur einen Kunden gleichzeitig aufnehmen. Sobald ein Kunde fertig bedient wurde und der entsprechende Slot frei wird, rückt der Kunde an vorderster Position der Warteschlange nach. Erst dann startet für diesen Kunden die eigentliche Aktivität, im Falle einer Arztpraxis also beispielsweise die Behandlung, mit der im Vorhinein vom LoadProfileGenerator festgelegten Dauer.

Auf diese Weise werden Wartezeiten korrekt abgebildet, allerdings könnte dies bei Überlastung dazu führen, dass die Warteschlange immer weiter anwächst. Um das zu verhindern, wird ein Abbruchkriterium für neue Besucher eingeführt. Bei der Ankunft wird anhand der Länge der Warteschlange und der Anzahl verfügbarer Service-Slots die voraussichtliche Wartezeit abgeschätzt. Übersteigt diese eine festgelegte Mindestwartedauer und ein bestimmtes Vielfaches der Aktivitätsdauer des Kunden, so bricht der Kunde seinen Besuch ab und verlässt unmittelbar den Point of Interest. Die Aktivität wird dabei zur Nachverfolgung als gescheitert markiert. Durch diesen Mechanismus werden unnatürlich lange Wartezeiten und Geschäfte, die niemals schließen können, verhindert. Gleichzeitig können anhand der Anzahl an gescheiterten Aktivitäten die Auswirkungen begrenzter Service-Kapazitäten in einem Gebiet untersucht werden.

Basierend auf Erfahrungswerten wurden eine Mindestwartedauer von 10 Minuten sowie das Vierfache der Aktivitätsdauer als Standardparameter gewählt. Zur Übersicht ist der vollständige Ablauf der Vorgänge in einem Point of Interest mit Warteschlange in Abbildung 3.9 dargestellt.

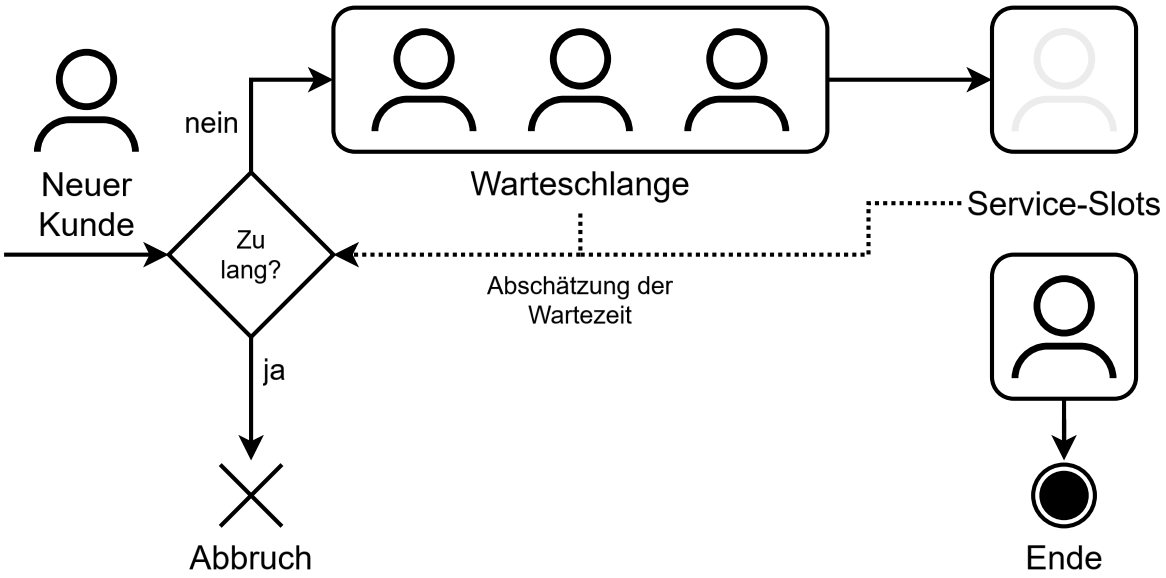



**Abbildung 3.9:** Simulation der Abläufe in Points of Interest mit Warteschlange

## 3.4 Implementierung

Die Städtesimulation wurde, dem in Abschnitt 3.3 vorgestellten Konzept entsprechend, im Rahmen eines C#-Programms als Erweiterung des LoadProfileGenerators implementiert. Da der Quellcode des LoadProfileGenerators bereits öffentlich zugänglich ist [56], wurden die Anpassungen und Ergänzungen im selben Repository vorgenommen. Der Großteil der Erweiterungen für die Simulation von Städten ist dabei im neu hinzugefügten Projektverzeichnis *CitySimulation* enthalten.

Abschnitt 3.4.1 beschreibt im Folgenden zunächst die Implementierung der Szenarioerstellung.

Abschnitt 3.4.2 geht anschließend auf die dafür notwendigen Änderungen am LoadProfileGenerator selbst ein.

### 3.4.1 Implementierung des CityScenarioGenerators

Die aus dem LoadProfileGenerator entwickelte Städtesimulation erfordert eine große Menge detaillierter Eingabedaten, um alle Komponenten des Modells korrekt zu parametrieren. Um diese Daten bereitzustellen, wurde, wie in Abschnitt 3.3.1 skizziert, eigens das Programm CityScenarioGenerator entwickelt. Es hat die Aufgabe, geeignete Daten zu sammeln, aufzubereiten und daraus die gewünschten Szenarien zusammenzustellen. Das Programm wurde in Python implementiert, da eine existierende Client-Bibliothek für die verwendete Datenbank den Datenabruf erleichtert und die Performance aufgrund des vergleichsweise geringen Rechenaufwands keine große Rolle spielt.

Als primäre Eingabe erfordert der CityScenarioGenerator eine Ortsangabe, die direkt an BUILDA weitergereicht wird, um die entsprechenden Gebäude abzurufen. Nach der Verarbeitung der Gebäudedaten werden dazu passende Konfigurationen für die Simulation erstellt. Die genaue Vorgehensweise hierbei wird im Folgenden getrennt für Wohn- und Nichtwohngebäude beschrieben. Anschließend werden einige Präferenzen der Bewohner in Bezug auf besuchte Orte und Verkehrswege festgelegt. Schließlich werden alle erstellten Konfigurationen in Form von JSON-Dateien gespeichert. Abschnitt 3.4.1 zeigt die Struktur, in der diese Dateien abgelegt werden, sodass die gesamte Szenariodefinition leicht verschoben und wieder eingelesen werden kann.

```
minimal_scenario
├── calcspec.json
├── city.json
├── houses
│   ├── DEA_DENW40AL6V00002U.json
│   ├── DEA_DENW40AL6V00003L.json
│   ├── DEA_DENW40AL7u00000t.json
│   └── DEA_DENW40AL85000073.json
├── routes
│   └── det_mx_All_0to86400.json
└── statistics
    ├── households_per_house.json
    ├── household_types.json
    ├── poi_types_all.json
    └── poi_types_relevant.json
```

**Abbildung 3.10:** Verzeichnisstruktur und Dateien eines Städteszenarios mit vier Häusern

#### Erstellen der Haushaltskonfigurationen

Die Datenbank BUILDA liefert zu jedem Wohngebäude eine basierend auf Zensusdaten geschätzte Anzahl an Haushalten. Zu jedem Haushalt werden die Gesamtanzahl an Bewohnern, die Anzahl an Autos, sowie jeweils die Anzahl an Frauen, Senioren und Voll- oder Teilzeitbeschäftigten unter

den Bewohnern angegeben. Anhand der Anteile der jeweiligen Personenkategorien im Haushalt wird nun ein möglichst gut passendes Haushalts-Template unter den im LoadProfileGenerator vordefinierten Templates ausgewählt. Das geschieht mittels der in Gleichung (3.1) gezeigten Bewertungsfunktion. Diese gibt die Abweichung $D$ eines Templates zu den gegebenen Spezifikationen eines Haushalts an.

$$\begin{aligned} D &= 1000 \cdot N + 100 \cdot R_w + 10 \cdot R_f + R_s \\ N &\in \mathbb{N} \\ R_w, R_f, R_s &\in [0, 1] \end{aligned} \tag{3.1}$$

$N$ bezeichnet dabei die Anzahl Bewohner im Haushalt, $R_w$ den Anteil berufstätiger Bewohner, $R_f$ den Anteil Frauen und $R_s$ den Anteil Senioren. Die so definierte Abweichung $D$ wird für alle 65 Templates berechnet. Falls die minimale Abweichung eindeutig ist, wird das entsprechende Template übernommen, andernfalls wird aus den Templates mit minimaler Abweichung zufällig eines gewählt. Durch die Gewichtung in der Funktion wird dabei der korrekten Bewohnerzahl die größte Bedeutung zugemessen, und den Anteilen bestimmter Untergruppen dementsprechend ein geringerer Einfluss.

Im LoadProfileGenerator kann jedem Haushalt zusätzlich ein *TransportationDeviceSet* zugewiesen werden, das die Menge der verfügbaren Transportmittel definiert. Daher wird an dieser Stelle für jeden Haushalt ein passendes TransportationDeviceSet mit der korrekten Anzahl an Autos sowie grundsätzlich Zugang zu öffentlichen Verkehrsmitteln und Fahrrädern bestimmt. Die gesammelten Daten werden, zusammen mit der Gebäude-ID und den dazugehörigen Koordinaten, in einer eigenen JSON-Datei im Unterverzeichnis „houses“ gespeichert (siehe Abschnitt 3.4.1).

#### Erstellen der Point-of-Interest-Konfigurationen

Bei den Nichtwohngebäuden wird ähnlich vorgegangen, nur werden keine Haushaltsdefinitionen, sondern Points of Interest für die Simulation erstellt. Basierend auf einer Kategorisierung des Amtlichen Liegenschaftskatasterinformationssystem (ALKIS) wird der Typ jedes Gebäudes bestimmt und es können dazu passende *Locations* gewählt werden. Locations stellen im LoadProfileGenerator Orte dar, an denen Personen sich aufhalten und Tätigkeiten nachgehen können. Diese Locations können sich innerhalb des eigenen Haushalts befinden, wie beispielsweise das Schlafzimmer, oder außerhalb, wie etwa der Arbeitsplatz. Insgesamt gibt es 69 Locations, 46 davon außer Haus. Die Zuordnung, welche Aktivitäten in welcher Location zur Verfügung stehen, erfolgt über eine gesonderte Abstraktionsebene, die *Household Traits*. Diese sind für alle Locations fest vordefiniert, um typische Orte im menschlichen Alltag abzubilden. Über das Zuweisen von Locations können Points of Interest somit in die Aktivitätssimulation eingebunden werden. Diese Zuweisung geschieht über eine fest definierte Zuordnung. Dabei muss allerdings beachtet werden, dass eine Location der Verwendung im LoadProfileGenerator nach lediglich ein Ort ist, an dem bestimmte Aktivitäten ausgeübt werden können. Das umfasst sowohl Nichtwohngebäude, wie Schulen oder Supermärkte, als auch Wohngebäude, etwa die Wohnung eines Freundes. Damit die Zuordnung sinnvoll umgesetzt werden kann, müssen alle Außer-Haus-Locations des LoadProfileGenerators entsprechend klassifiziert werden.

Wie in Tabelle 3.1 dargestellt, wird dabei zwischen vier Arten von Locations unterschieden. Am häufigsten sind Locations, die Nichtwohngebäude repräsentieren. Ob ein Nichtwohngebäude einer dieser Locations entspricht, kann über ein manuell definiertes Mapping anhand der ALKIS-Gebäudekategorie entschieden werden. Bei einigen Kategorien ist eine 1-zu-1-Zuordnung möglich, so wird etwa die Kategorie „Laden“ stets durch die Location „Supermarket“ abgebildet. Es gibt aber auch Fälle, in denen die ALKIS-Kategorie keine eindeutige Zuordnung zulässt. Beispielsweise

**Tabelle 3.1:** Klassifizierung aller Locations im LoadProfileGenerator für eine geeignete Zuordnung zu Gebäuden. Die Kategorie gibt an, auf welche Art von Gebäude sich die jeweilige Location bezieht. Für die Locations werden die englischen Originalbezeichnungen verwendet. Locations, die sich innerhalb des eigenen Haushalts befinden, wie zum Beispiel das Schlafzimmer, werden nicht zugewiesen und sind daher nicht aufgeführt.

| Nichtwohngebäude | Arbeitsplatz | Wohngebäude | Kein Gebäude |
|---|---|---|---|
| Shopping Mall | Office Work-place | House of friends or relatives | Hiking Location |
| Dance Club | Workplace (shift worker) | Home[1] | Walking Location |
| University | Volunteer Workplace | | Mountain Road |
| Summer Camp | | | Running Path |
| Cinema Location | | | Bicycle Route |
| Soccer Practice Location | | | Fishing Location |
| Cafe | | | Hunting Forest |
| School | | | Festival Location |
| Museum | | | Flea Market Location |
| Kindergarden | | | |
| Dance Studio | | | |
| Fitness Studio | | | |
| Concert House | | | |
| Community College | | | |
| Singing School | | | |
| Doctors Office | | | |
| Outdoor Swimming Pool | | | |
| Indoor Swimming Pool | | | |
| Bow Range | | | |
| Bar Location | | | |
| Yoga Studio | | | |
| Car Wash | | | |
| Local Recreational Area | | | |
| Opera | | | |
| Musical Society | | | |
| Supermarket | | | |
| Food Market | | | |
| Theater | | | |
| Horse Stable | | | |
| Golf Club | | | |
| Generic Errand Location | | | |

[1] Die Location „Home“ bezieht sich auf den Arbeitsplatz einer Haushaltshilfe und liegt daher ebenfalls in einem Wohngebäude.

gibt es mehrere Locations, die für ein „Gebäude für Sportzwecke“ infrage kommen. Bei solchen Kategorien wird für jedes Gebäude einzeln eine der Locations zufällig ausgewählt. Darüber hinaus gibt es einige Gebäudekategorien, die für die alltägliche Aktivitätssimulation keinerlei Bedeutung haben. Dazu gehören etwa Schornsteine oder Funkmasten. Diese Gebäude werden verworfen, ohne einen Point of Interest daraus zu erstellen.

Die zweite Kategorie von Locations bezeichnet einen Arbeitsplatz. Personen arbeiten für gewöhnlich ebenfalls in Nichtwohngebäuden, allerdings ist hier ein separates Mapping erforderlich, das angibt, welche Art von Arbeit in einem bestimmten Gebäudetyp möglich ist. So gibt es in Krankenhäusern sowohl Stellen mit Schichtdienst als auch mit gewöhnlichen Büroarbeitszeiten. Bei anderen Gebäudetypen, wie beispielsweise einem Gemeindehaus, können dagegen auch ehrenamtliche Tätigkeiten angenommen werden. Um alle verschiedenen Möglichkeiten abbilden zu können, wird an dieser Stelle nicht nur eine Arbeitsplatz-Location ausgewählt. Stattdessen werden für jedes Nichtwohngebäude mehrere Points of Interest erstellt, eine für jede Arbeitsvariante. Viele Nichtwohngebäude stellen einen Service für Kunden bereit und können außerdem als Arbeitsplatz dienen, daher werden die Zuordnungen dieser beiden Location-Typen unabhängig voneinander durchgeführt. Falls mehrere Points of Interest aus einem Nichtwohngebäude erstellt werden, werden diese in der Simulation vollständig getrennt voneinander behandelt. Das schränkt die Möglichkeiten für die Modellierung dieser Gebäude geringfügig ein, da beispielsweise keine gemeinsame Berücksichtigung der Anzahl anwesender Mitarbeiter und Kunden eines Geschäfts möglich ist. Derartige Effekte lassen sich in der verwendeten Aktivitätsmodellierung allerdings auch nicht ohne Weiteres sinnvoll einbeziehen und werden für diese Städtesimulation als vernachlässigbar angesehen. Durch die beschriebene Zuordnung kann es für ein Gebäude niemals mehrere Points of Interest desselben Location-Typs geben. Eine Kombination aus Gebäude-ID und Location-Name ist somit stets eindeutig und kann als ID für den Point of Interest genutzt werden.

Die dritte Location-Kategorie besteht aus Orten, die in anderen Wohngebäuden liegen, beispielsweise die Wohnungen von Bekannten. Solche Locations sind für soziale Aktivitäten wie Besuche von Freunden oder Familie relevant. Die entsprechenden Points of Interest werden für jede Person daher im Rahmen der Festlegung der Zielpräferenzen zufällig aus allen Wohngebäuden gewählt.

Übrig bleiben Locations, die nicht an ein Gebäude gebunden sind und somit nicht anhand der Gebäudedatenbank generiert werden können. Entsprechende Points of Interest für diese Locations müssen auf anderem Wege erstellt werden. Da diese Locations größtenteils individuelle, unabhängige Tätigkeiten definieren, wie beispielsweise einen Spaziergang im Wald, kommt es hierbei nicht auf eine exakte Ortsangabe oder korrekte Gruppierung von Personen innerhalb eines Points of Interest an. Randomisiert erzeugte Instanzen sind somit an dieser Stelle vorerst ausreichend.

Insgesamt werden also für jede ALKIS-Gebäudekategorie zwei Mengen an passenden Locations festgelegt, eine für die Rolle des Gebäudes als Arbeitsplatz und eine für die Nutzung als Kunde. Ein beispielhafter Auszug aus dem erstellten ALKIS-Kategorie-Location-Mapping ist in Abbildung 3.11 dargestellt. Das vollständige Mapping wurde als Teil des CityScenarioGenerators [57] veröffentlicht und ist im Repository unter „data/alkis_mapping“ einsehbar.

Alle generierten Points of Interest werden aufgrund der geringen Größe zusammen in der Datei

**Abbildung 3.11:** Beispiel der Zuordnung passender Locations zu ALKIS-Gebäudekategorien. Für ein Gebäude werden stets alle Einträge aus „work_locations“ und ein zufälliger Eintrag aus „non_work_locations“ übernommen.

```
"Gaststätte, Restaurant": {
        "non_work_locations": [
                "Bar Location",
                "Cafe"
        ],
        "work_locations": [
                "Workplace (shift worker)"
        ]
},
"Fabrik": {
        "non_work_locations": [],
        "work_locations": [
                "Workplace (shift worker)",
                "Office Workplace"
        ]
},
"Theater, Oper": {
        "non_work_locations": [
                "Opera",
                "Theater"
        ],
        "work_locations": [
                "Workplace (shift worker)"
        ]
}
```

„city.json“ gespeichert (siehe Abschnitt 3.4.1). Jede Point-of-Interest-Definition enthält dabei den entsprechenden Location-Typ, die Gebäudekoordinaten sowie optional eine Angabe zu den Öffnungszeiten.

### Import zusätzlicher Nichtwohngebäude

Wie sich nach anfänglichen Tests gezeigt hat, sind die ALKIS-Daten zu Nichtwohngebäuden teilweise sehr unvollständig. Insbesondere kleinere Unternehmen wie Arztpraxen fehlen häufig. So wird beispielsweise im gesamten Gebiet der Stadt Jülich kein einziger Arzt angegeben. Um dieses Problem zu lösen, wird OpenStreetMap[1] als zusätzliche Datenquelle für einige ausgewählte Point-of-Interest-Typen verwendet. OpenStreetMap bietet als freies Projekt zur Sammlung von Geodaten wesentlich mehr und detailliertere Daten als ALKIS.

Für den Gebäudeimport aus OpenStreetMap werden mithilfe des Filterungs-Werkzeugs overpass turbo[2] alle Gebäude im Zielgebiet abgefragt, die eine bestimmte Schlüssel-Wert-Kombination enthalten, beispielsweise die Kombination „Amenity“ → „doctors“. Die vollständige Abfrage ist in Anhang A.2 angegeben.

Ergänzend ist auch der Import von Gebäuden aus weiteren Datenquellen, beispielsweise aus Adress-

[1] https://www.openstreetmap.org/
[2] https://overpass-turbo.eu/index.html

büchern, über manuell erstellte Eingabedateien möglich. Um bei dieser Vielzahl an Quellen Dopplungen zu vermeiden, wird ein Spatial Join durchgeführt. Dabei wird anhand der Koordinaten der zu importierenden Gebäude überprüft, welche davon sich so nah beieinander befinden, dass sie sich auf dasselbe Gebäude beziehen. In diesen Fällen werden die verfügbaren Informationen zum Verwendungszweck des Gebäudes aus allen Quellen zusammengeführt.

Jede der untersuchten Datenquellen für sich war in Tests unvollständig, kombiniert ergibt sich aber eine gute Abdeckung. Für Jülich konnten so beispielsweise insgesamt 57 Arztpraxen identifiziert werden.

#### Umsetzung der Point-of-Interest-Wahl

Die Point-of-Interest-Wahl wird als Teil der Szenarioerstellung im CityScenarioGenerator durchgeführt. Dabei wird für jede Person unabhängig eine feste Anzahl an Zielen pro Typ bestimmt. Die Auswahl erfolgt zufällig und ist abhängig von der Entfernung in Luftlinie vom Wohnort der Person. Dazu werden von Strobel und Pruckner [58] empirisch berechnete Deterrence-Funktionen eingesetzt. Diese Funktionen weisen jedem möglichen Zielort anhand der Entfernung und des Typs ein passendes Gewicht zu. Das ist notwendig, da Personen bei bestimmten Point-of-Interest-Typen, wie beispielsweise dem Arbeitsplatz, längere Wege in Kauf nehmen als bei anderen.

Um zu gewährleisten, dass jede Person alle ihre Aktivitäten ausführen kann, wird mindestens eine Instanz von jedem Point-of-Interest-Typen zugewiesen. Viele davon werden später nicht benötigt, eine Person ohne das Hobby Fußballspielen nutzt beispielsweise nie die Location „Soccer Practice Location“. In diesen Fällen wird der entsprechende Point of Interest bei der Initialisierung der Städtesimulation für diese Person übergangen, der dadurch entstehende Zusatzaufwand ist somit minimal.

Die festgelegten Point-of-Interest-Präferenzen werden für jede Person als Teil der Haushaltskonfiguration gespeichert, also innerhalb der JSON-Datei für das entsprechende Haus. Zur Optimierung werden außerdem in jeder Hauskonfiguration die gesammelten für die Haushalte des Hauses relevanten Points of Interest aufgelistet. Das reduziert den Rechenaufwand, da die Simulation einzelner Häuser im Hinblick auf die zugrundeliegenden Datenstrukturen weitestgehend getrennt erfolgt. So muss in den Strukturen zur Simulation eines Hauses nur ein Bruchteil der gesamten Orte überhaupt berücksichtigt werden. Die vollständige Liste an Points of Interest in der separaten „city.json“-Datei wird nur zu Beginn der Städtesimulation benötigt, um die Point-of-Interest-Simulatoren zu initialisieren und auf die Worker zu verteilen.

#### Generierung von Routen

Nachdem feststeht, welche Orte jede Person aufsuchen will, müssen Routen für die entsprechenden Wege angelegt werden. Die Routen können neben den Ortspräferenzen als Teil der Datei „city.json“ gespeichert werden. Um die Routenerstellung jedoch entkoppeln zu können, besteht alternativ die Möglichkeit, die Routendefinitionen separat in der Datei „routes.json“ abzulegen. Diese Datei kann dann im Rahmen der in Abschnitt 3.3.5 beschriebenen Modellkopplung bei Bedarf von einem externen Transportmodell unter Verwendung der Gebäudedefinitionen und Ortspräferenzen generiert werden.

Eine Routendefinition fungiert als Sammelobjekt für eine Reise, die mit unterschiedlichen Transportmodi erfolgen kann. Daher müssen Distanzen und Reisedauern für jeden relevanten Modus angegeben werden. Im LoadProfileGenerator sind grundsätzlich die Transportmittel Auto, Fahrrad und Bus (stellvertretend für jeglichen öffentlichen Personennahverkehr) implementiert, außerdem können Agenten zu Fuß zu ihrem Ziel gehen. Die Moduswahl erfolgt über feste Gewichtungen, die in der Routendefinition anzugeben sind. Um individuelle Unterschiede bei den Transportvorlieben berücksichtigen zu können, werden mehrere Gewichtungsangaben erwartet. Von Interesse ist dabei, ob ein Haushalt über ein Auto oder über ein Abonnement für öffentlichen Nahverkehr verfügt. Damit alle relevanten Situationen ideal abgebildet werden können, ergeben sich die in Tabelle 3.2 dargestellten sieben Fälle.

**Tabelle 3.2:** Für die Verkehrsmittelwahl ergeben sich im Hinblick auf die verfügbaren Informationen neun mögliche Fälle. Sieben davon werden berücksichtigt, sodass Haushalte mit und ohne Auto sowie mit und ohne ÖPNV-Abonnement unterschieden werden können. Für die Nutzung des ÖPNV ist ein Abonnement nicht erforderlich, es erhöht aber die Wahrscheinlichkeit. Da unter Umständen nicht bekannt ist, ob ein Haushalt über ein ÖPNV-Abonnement verfügt, gibt es für diesen Fall eine zusammengefasste Kategorie. Außerdem gibt es eine weitere Kategorie für den Fall, dass zu beiden Kriterien keine Informationen vorliegen.

| | ÖPNV-Abonnement verfügbar | | |
|---|---|---|---|
| Auto verfügbar | Ja | Nein | Unbekannt |
| Ja | X | X | X |
| Nein | X | X | X |
| Unbekannt | | | X |

Aufgrund der Option, ein detailliertes Transportmodell für die Routengenerierung einzusetzen, ist im CityScenarioGenerator lediglich eine simple Ersatzimplementierung enthalten. Diese erzeugt ausschließlich Auto- oder ÖPNV-Routen und verwendet dabei die Gebäudedistanz in Luftlinie.

### 3.4.2 Erweiterung des LoadProfileGenerators

Zur Integration in die Städtesimulation mussten einige grundlegende Anpassungen am LoadProfileGenerator vorgenommen werden. Um die bisherige Verwendungsmöglichkeit als Lastprofilgenerator für einzelne Haushalte nicht zu beeinträchtigen, wurden die wesentlichen internen Änderungen so vorgenommen, dass sie für den Nutzer nicht sichtbar sind. Spezielle Erweiterungen für die Städtesimulation sind nur in einem gesonderten Modus aktiv. Allgemeine Änderungen, von denen auch der Modus ohne Städtesimulation profitiert, wurden dagegen dauerhaft übernommen.

#### Unterschiedliche Aktivitätstypen

Die tiefgreifendste Änderung, die für die Integration des LoadProfileGenerators in die Städtesimulation notwendig war, betrifft die Modellierung von Aktivitäten unbekannter Dauer. Da, wie in Abschnitt 3.3.4 beschrieben, nicht die gesamte Funktionalität der bisher im LoadProfileGenerator implementierten Aktivitäten auch für Außer-Haus-Aktivitäten benötigt wird, werden die beiden Aktivitätstypen voneinander getrennt. In der Städtesimulation muss also zwischen lokalen Aktivitäten mit im Vorhinein bekannter Dauer und Außer-Haus-Aktivitäten mit unbekannter Dauer

unterschieden werden. Hierzu wurde zur Abbildung geplanter und bereits ausgeführter Aktivitäten eine zusätzliche Abstraktionsebene hinzugefügt. Die dazugehörigen Elemente sind in Abbildung 3.12 dargestellt. Das Interface IActivity und die abstrakte Basisklasse Activity repräsentieren dabei allgemein eine Aktivität, die gerade ausgeführt wird oder für die eine zukünftige Ausführung geplant ist. Implementiert wird dieses Interface von den Klassen StaticActivity und DynamicActivity. Die Klasse StaticActivity steht dabei für Aktivitäten mit im Vorhinein bekannter Länge, und DynamicActivity entsprechend für Aktivitäten unbekannter Länge. Für beide gibt es jeweils eine Spezialisierung als normale Aktivität und als Reise-Aktivität. Im normalen Modus zur Simulation eines Haushalts werden alle Aktivitäten über die Klassen LocalActivity und StaticTravelActivity abgebildet. Bei Verwendung der Städtesimulation werden nur lokale Aktivitäten im Haushalt durch LocalActivity-Objekte repräsentiert. Für Außer-Haus-Aktivitäten werden stattdessen die Klassen RemoteActivity beziehungsweise DynamicTravelActivity für Reisen eingesetzt. StaticTravelActivity wird in der Städtesimulation nicht verwendet, da jede Reise hier grundsätzlich eine unbekannte Dauer hat.

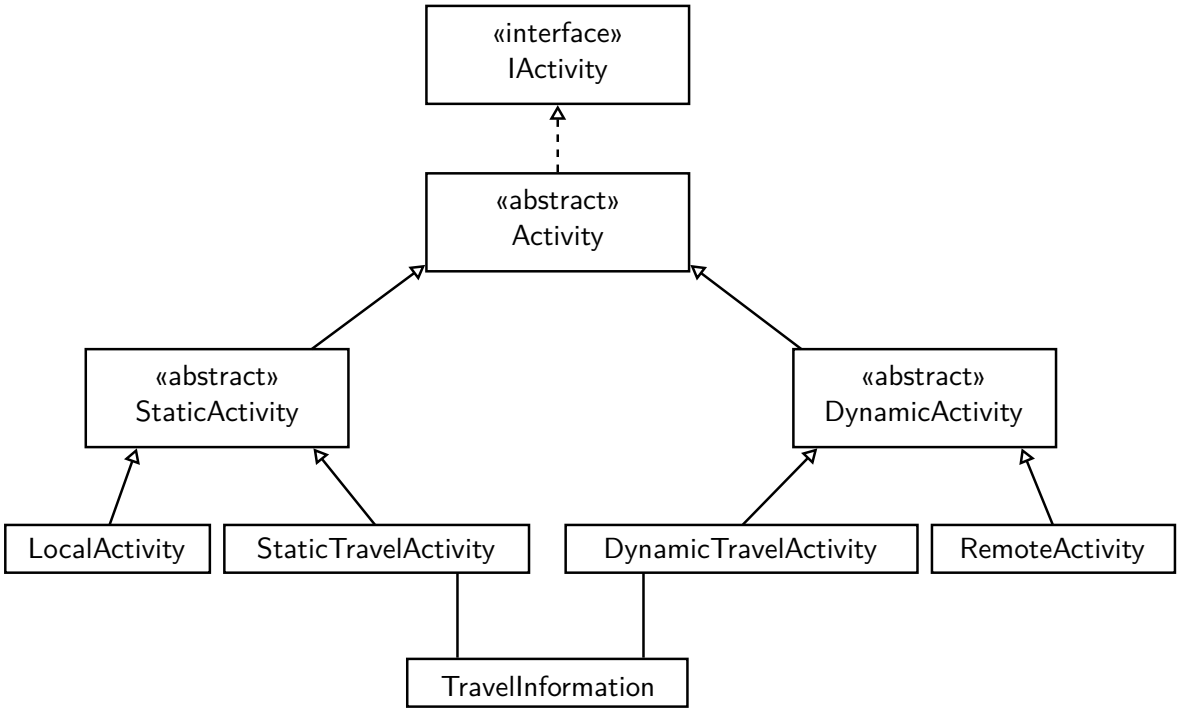



**Abbildung 3.12:** UML-Klassendiagramm der zusätzlichen Interfaces und Klassen zur Darstellung der verschiedenen Aktivitätsvarianten

Die Aktivitätsobjekte werden genutzt, um alle relevanten Daten zu einer Aktivität so lange zu speichern, wie sie benötigt werden. Dazu gehören insbesondere die Startzeit, die ausführende Person und der Standort. Das Interface wurde dabei so einfach wie möglich gehalten. Im Wesentlichen gibt es eine Start-Methode, um die Ausführung einer Aktivität zu beginnen, eine IsFinished-Methode, die regelmäßig aufgerufen wird, um den Status der Aktivität zu prüfen, sowie eine Finish-Methode, um die Aktivität abzuschließen, sobald sie beendet ist. Nachdem ein Aktivitätsobjekt einmal angelegt wurde, kann die gesamte notwendige Logik mithilfe dieser Methoden erfolgen. Die Aktivitätsobjekte übernehmen dann, entsprechend der Variante, das Logging aller relevanten Informationen, Anpassen von Variablen und Freigeben von Geräten. Dadurch sind

außerhalb dieser Klassen kaum Fallunterscheidungen für den Simulationsmodus und die Art der Aktivität notwendig. Eine beliebige Aktivität und die dazugehörige Reise zum Ausführungsort werden als zwei getrennte Aktivitäten modelliert. Da die Reise aber überhaupt erst wegen der Aktivität angetreten wird, werden die beiden dazugehörigen Aktivitätsobjekte gleichzeitig angelegt. Eine Person kann allerdings zu jedem Zeitpunkt nur eine Aktivität ausüben. Daher werden die Objekte in einer Aktivitätswarteschlange verwaltet, sodass nach Abschluss der Reise nahtlos mit der Zielaktivität fortgefahren werden kann. Diese Architektur bietet zukünftig auch die Möglichkeit, komplexere Algorithmen zur Vorausplanung von Aktivitäten einzusetzen.

#### Nachrichten zu Außer-Haus-Aktivitäten

Abbildung 3.13 zeigt die Nachrichten-Klassen, die angelegt werden, um Start oder Ende einer Außer-Haus-Aktivität in der Städtesimulation zu melden. Diese Klassen haben keine eigene Funktionalität, sondern speichern lediglich alle benötigten Informationen zum jeweiligen Ereignis in einer Form, die leicht serialisiert werden kann. Der Aspekt der Serialisierbarkeit ist an dieser Stelle besonders wichtig, da diese Nachrichten mittels MPI an andere Worker versandt werden müssen. Als Serialisierungsformat wurde dazu JSON gewählt, da die Implementierung durch in .NET verfügbare Standardkomponenten sehr leicht und für die verhältnismäßig geringen Datenmengen ausreichend effizient ist.

Um Rechenzeit und Bandbreite zu sparen, wurde der Umfang der Nachrichten auf das notwendige Minimum reduziert. Beide Klassen setzen sich daher aus so wenigen Attributen wie möglich zusammen. Enthalten sind in erster Linie eindeutige Bezeichner, um die betroffene Person und die relevanten Orte identifizieren zu können. Anhand der Orts-ID wird über das Point-of-Interest-Register außerdem der zuständige Worker ermittelt. Nachrichten zum Start einer neuen Außer-Haus-Aktivität liefern zusätzlich relevante Daten über die gestartete Aktivität selbst mit. Neben dem Namen der Aktivität wird die erwartete Dauer angegeben, und ob es sich um eine Reiseaktivität handelt. Die Dauer bezieht sich dabei auf die vom LoadProfileGenerator probabilistisch bestimmte Zeitspanne für diese Aktivität. Anders als bei lokalen Aktivitäten, bei denen die Dauer direkt übernommen wird, dient sie bei Außer-Haus-Aktivitäten lediglich als Information für den Point-of-Interest-Simulator. Dieser kann die tatsächliche Dauer der Aktivität, abhängig von Faktoren wie der Anzahl gleichzeitig anwesender Personen, frei anpassen.

Eine einzelne Nachricht hat eine Größe von etwa 200 Bytes, der genaue Wert hängt von den verwendeten Namen und IDs ab. Im Verlauf einer ganzjährigen Simulation von Jülich mit 50 Workern werden insgesamt etwa $10^8$ Nachrichten versandt, es ergibt sich also eine durchschnittliche Gesamtdatenmenge von 18 GB. Da sich diese Nachrichten sehr gleichmäßig auf alle Zeitschritte der Simulation verteilen, spielt diese Menge bei üblichen Datenübertragungsraten in modernen verteilten Rechnersystemen von mindestens 1 GBit/s keine limitierende Rolle.

#### *Sites* und *Locations*

Für die Modellierung von Reisen werden Locations in sogenannten *Sites* gruppiert. Zwischen den Sites werden dann paarweise Distanzen und Transportmodalitäten fest definiert. Der Einfachheit halber wurden bisher lediglich die Sites „Home“, „Workplace“, „School“, „Shopping“ und „Event Location“ verwendet, wobei insbesondere letztgenannte genutzt wurde, um verschiedene Locations

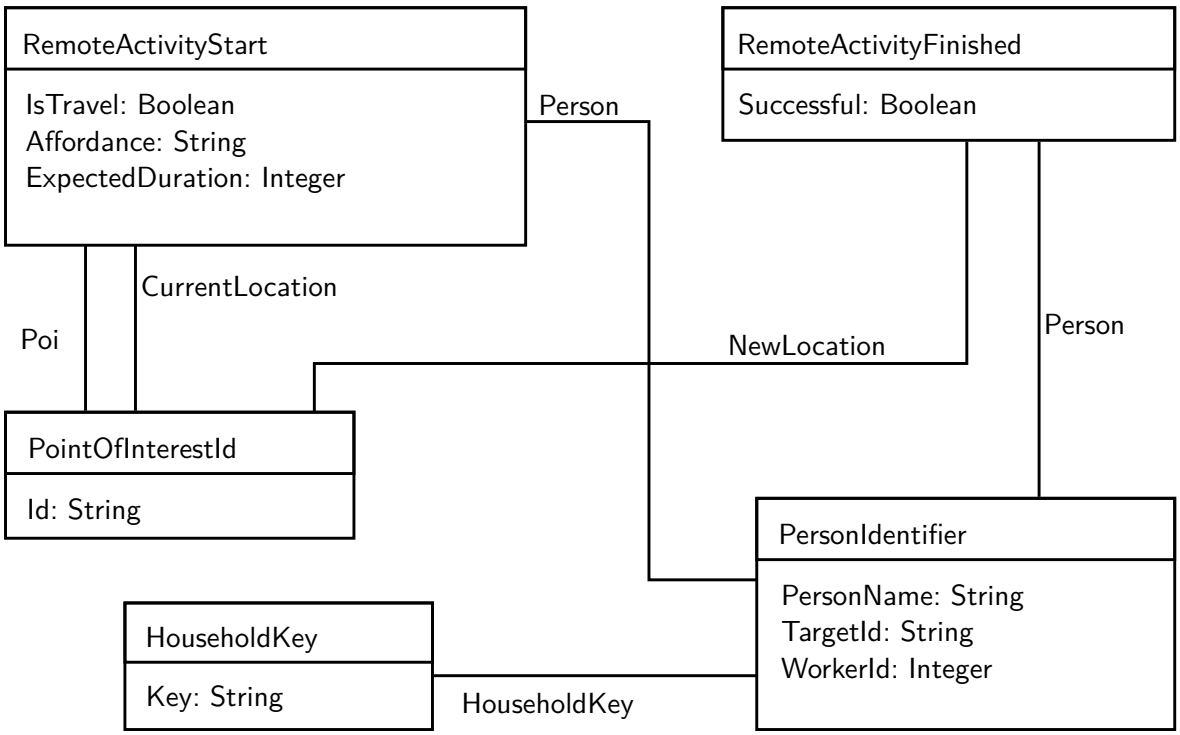



**Abbildung 3.13:** UML-Klassendiagramm der Nachrichten- und ID-Klassen, die genutzt werden, um über Ereignisse in der Städtesimulation zu informieren

zu sammeln, die in keine der anderen Kategorien fallen. Für die Abbildung ganzer Städte musste diese feste Verteilung um eine dynamischere Lösung ergänzt werden. Von den vordefinierten Sites wird lediglich „Home“ beibehalten. Für jeden Point of Interest der Stadt wird anschließend ein eigener Site angelegt. Routen verbinden dann stets zwei Sites miteinander.

Selbst bei optimiertem Vorgehen können bei größeren Szenarien sehr viele Routen-Objekte erzeugt werden, im Beispiel Jülich sind es etwa 100 Millionen Routen. Allein das Speichern dieser großen Menge an Objekten in der zentralen Datenbank des LoadProfileGenerators nimmt viel Zeit in Anspruch. Durch die Implementierung einer gesonderten Speicherfunktion, die alle Routen-Objekte innerhalb einer einzigen Datenbank-Session speichert, konnte diese Zeit um die Hälfte reduziert werden.

# 4 Auswertung der Städtesimulation

Die Städtesimulation wurde entwickelt, um das Verhalten aller Bewohner einer Stadt integriert und unter Berücksichtigung von Interaktionen und Wechselwirkungen abzubilden. In diesem Kapitel wird die Funktionalität der Simulation demonstriert, getestet und validiert, um zu zeigen, dass sie die gestellten Anforderungen erfüllt, und um mögliche Anwendungen aufzuzeigen.

Da die Performance bei derart komplexen Simulationen einer großen Anzahl an Objekten von besonderer Bedeutung ist, wird in Abschnitt 4.1 zunächst die Leistungsfähigkeit und Skalierbarkeit der Simulation getestet und bewertet. Um die Gültigkeit der Simulationsergebnisse zu überprüfen, werden anschließend verschiedene Szenarien für die Stadt Jülich als Anwendungsfall definiert und simuliert. Die dabei generierten Daten werden dann im Hinblick auf das Verhalten der Einwohner, die unternommenen Reisen, Energieverbräuche sowie Besuche bei Points of Interest ausgewertet und validiert.

## 4.1 Skalierbarkeit

Die Simulation einer ganzen Stadt ist ein aufwendiges, komplexes und rechenintensives Unterfangen. Die Skalierbarkeit und Performance der Städtesimulation waren daher in der Entwicklung besonders relevant. Durch das anspruchsvolle Ziel, Städte im Bereich von 10.000 bis 100.000 Einwohnern trotz detaillierter Verhaltensmodellierung in kurzer Zeit simulieren zu können, musste das Programm entsprechend stark optimiert werden. Die dafür zwingend notwendige Ausnutzung hochgradig paralleler Rechnerarchitekturen stellte darüber hinaus spezielle Anforderungen an die Implementierung.

Zwecks Untersuchung der Performance ist es sinnvoll, die bereits in Abschnitt 3.3 erwähnten drei Abschnitte der Städtesimulation getrennt zu betrachten. Ausschlaggebender Faktor für diese Trennung ist die Synchronisation der Worker. Im Bereich der Parallelisierung ist die Kommunikation zwischen den Prozessen ein wesentlicher Punkt, der die Effizienz der parallelen Software maßgeblich bestimmen kann. Die Initialisierung der Haushaltsmodelle sowie das Postprocessing enthalten nur zu Beginn einen einzigen Kommunikationsaufruf zur Synchronisierung und Verteilung der Arbeitslast. Ab dann arbeitet jeder Prozess seine Arbeitspakete ab, ungebremst von weiterer Kommunikation. Aus diesem Grund skalieren diese beiden Abschnitte nahezu ideal mit der Anzahl an Workern. Eine Verdopplung der Workeranzahl bedeutet die halbe Arbeitslast für jeden Worker und somit auch in etwa eine Halbierung der Rechenzeit.

Gemäß Amdahls Gesetz [59] ist die Beschleunigung durch Parallelisierung eines vorgegebenen Problems durch den seriellen Teil des Programms begrenzt. Bei der Städtesimulation kommt ergänzend hinzu, dass die Arbeitslast nicht beliebig fein verteilt werden kann. Sobald jeder Worker nur noch ein einziges Haus bearbeitet, ist keine weitere Aufteilung mehr möglich.

Diese Betrachtungsweise gilt für vorgegebene Probleme fester Größe. Die Städtesimulation kann auch danach bewertet werden, bis zu welcher maximalen Größe Städte modelliert werden können. In diesem Fall findet Gustafsons Gesetz [60] Anwendung. Demnach wächst mit zunehmender Parallelisierung die Größe der in einer vorgegebenen Zeit simulierbaren Städte linear an. Bessere Hardware eröffnet somit neue Möglichkeiten der Modellierung, und aufgrund des linearen

Zusammenhangs ist diese Verbesserung effizient und theoretisch beliebig skalierbar.

Dies gilt jedoch nur für Initialisierung und Postprocessing. Die Kommunikation in der Hauptsimulation beeinträchtigt die Effizienz der Parallelisierung merklich. Zur Veranschaulichung der unterschiedlichen Skalierbarkeit der einzelnen Schritte der Städtesimulation wurde ein Test durchgeführt. Dazu wurde ein kleines Szenario mit 100 Einfamilienhäusern generiert und mit variierenden Workeranzahlen zwischen 1 und 100 simuliert. Die Rechenzeiten der Simulationsläufe sind in Abbildung 4.1 dargestellt. Darin ist klar zu erkennen, dass die Hauptsimulation schlechter skaliert

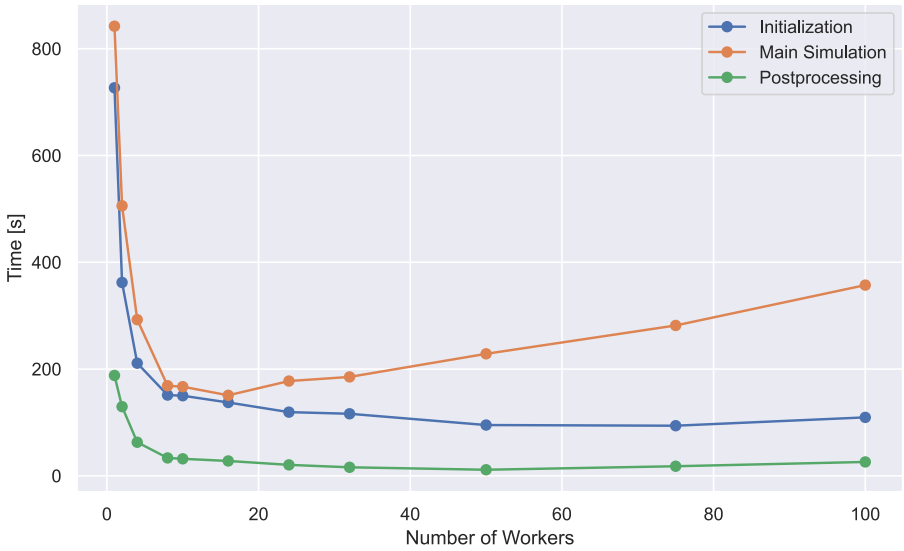
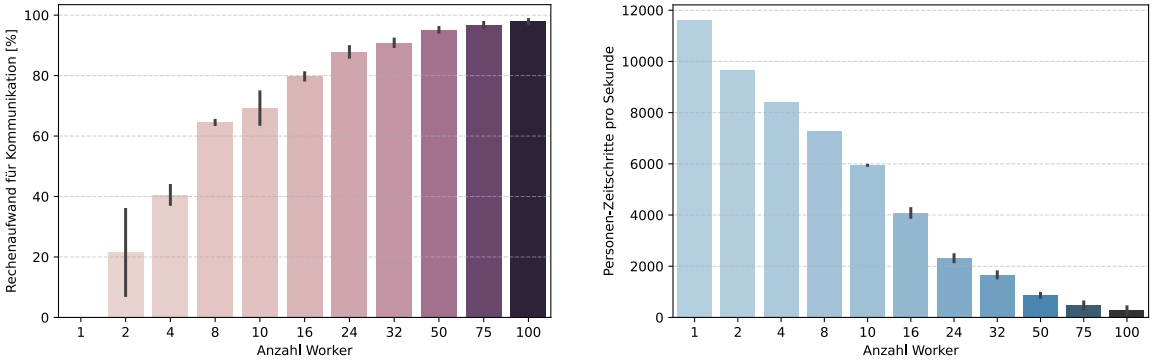



**Abbildung 4.1:** Veränderung der Rechenzeiten für Initialisierung, Hauptsimulation und Postprocessing bei steigender Anzahl paralleler Worker

und ab einem bestimmten Punkt, in diesem Fall ab 24 Workern, der zunehmende Kommunikationsaufwand die Performance der Simulation wieder verschlechtert. Der Zusatzaufwand besteht dabei zum einen im Senden und Empfangen von Nachrichten, und zum anderen im Warten auf andere Worker, da jede Kommunikation stets alle Worker miteinbezieht.

Wie erheblich der Kommunikationsaufwand werden kann, ist in Abbildung 4.2 zu sehen. Mit steigender Anzahl an Workern steigt der Anteil der Rechenzeit in der Hauptsimulation, der für MPI-Funktionsaufrufe verwendet wird. Im Extremfall, wenn jeder Worker nur ein Haus simuliert, wird fast die gesamte Rechenleistung nur für Kommunikation aufgewendet. Gleichzeitig sinkt die Simulationsgeschwindigkeit jedes einzelnen Workers. Um die Rechengeschwindigkeit eines Workers bei unterschiedlichen Anzahlen simulierter Häuser vergleichen zu können, wird hier die Metrik Personen-Zeitschritte pro Sekunde verwendet. Diese gibt an, wie viele einzelne Zeitschritte wie vieler Personen ein Worker innerhalb einer Sekunde simuliert hat. Abbildung 4.2b zeigt die durchschnittliche Geschwindigkeit der Worker anhand dieser Metrik an. Hier ist deutlich zu sehen, dass jede Erhöhung der Workeranzahl die Effizienz jedes einzelnen Workers verringert. Solange dieser Abfall der Rechengeschwindigkeit durch die größere Anzahl an Workern überkompensiert wird, ist die Parallelisierung sinnvoll und beschleunigt insgesamt die Städtesimulation. Die Anzahl an Workern, bei der dieser Punkt überschritten wird, hängt dabei von der Größe des Szenarios ab. Wird die Anzahl zu simulierender Häuser bei gleicher Workeranzahl größer, steigt der Rechenauf-




**(a)** Kommunikationsaufwand **(b)** Simulationsgeschwindigkeit

**Abbildung 4.2:** Auswirkungen des Kommunikationsaufwands bei einer hohen Anzahl an Workern auf den Rechenaufwand für MPI-Funktionen (a) und die Simulationsgeschwindigkeit (b)

wand pro Worker, der Kommunikationsaufwand bleibt aber nahezu gleich. In Tests erreichte eine Simulation von Jülich mit über 9.000 Häusern selbst auf 64 Workern noch etwa 3.500 Personen-Zeitschritte pro Sekunde. Je größer das Problem ist, desto mehr Worker können somit effizient eingesetzt werden.

Die geringen Fehlerbalken in Abbildung 4.2 zeigen außerdem, dass die Schwankungen bezüglich Geschwindigkeit zwischen den Workern größtenteils sehr gering ausfallen. Das bedeutet, dass die Gesamtlast, wie in Abschnitt 3.3.3 angenommen, ausreichend gut verteilt wird und ein aufwendiger Mechanismus zum Lastausgleich nicht notwendig ist. Es ist lediglich zu beachten, dass diese Aussage nur für homogene Rechnerarchitekturen gilt. Bei starken Unterschieden in der Rechenleistung bremsen langsamere Prozessorkerne die schnelleren aus. Für die Zukunft besteht eine Möglichkeit, auch auf heterogenen Architekturen eine größtmögliche Effizienz zu erreichen, darin, zunächst die Rechenleistung jedes Workers zu messen und die Häuser anschließend entsprechend zu verteilen, sodass leistungsfähigere Worker eine größere Arbeitslast erhalten.

## 4.2 Szenariovalidierung

Als Anwendungsfall und Test für die Städtesimulation wird, wie in Abschnitt 3.2 erklärt, die Stadt Jülich verwendet. Jülich ist hierfür aufgrund seiner Größe sehr gut geeignet, da es mit über 30.000 Einwohnern aufwendig zu simulieren ist und somit einen sinnvollen Test für die Skalierbarkeit der Städtesimulation darstellt. Gleichzeitig enthält es eine Vielzahl verschiedener Nichtwohngebäude und ermöglicht so, die neu entwickelte Funktionalität anhand eines realen Beispiels zu erproben.

Die in der Städtesimulation produzierten Ausgaben wie Last und Aktivitäten hängen maßgeblich von den Eigenschaften der Bevölkerung ab. Für gute Simulationsergebnisse ist somit ein gutes Szenario mit einer passenden Bevölkerungsstruktur wichtig. Daher werden zunächst die mit dem CityScenarioGenerator erstellten Szenariodaten dahingehend überprüft und anhand offizieller Bevölkerungsdaten validiert.

Abweichungen in der generierten Bevölkerung können grundsätzlich zwei Ursachen haben. Zum einen wurden die Haushaltsdaten in BUILDA in Teilen heuristisch aus den Zensusdaten von 2011 bestimmt. Zensus-Daten liegen in 100m-Gitterzellen vor und mussten für die Datenbank auf Haus-

haltsebene disaggregiert werden, wobei sich in der Summe leichte Diskrepanzen ergeben können. Zum anderen kann es sein, dass im LoadProfileGenerator kein perfekt passendes Haushaltstemplate für einen Haushalt existiert. Wie in Abschnitt 3.4.1 beschrieben, wurde stets das am besten geeignete Template gewählt, basierend auf einer Bewertungsfunktion. Da diese Funktion die Geschlechter und Altersklassen weniger stark gewichtet, kann es insbesondere bei diesen Statistiken zu Abweichungen kommen. BUILDA definiert für Jülich 54 Haushalte mit mehr als 6 Mitgliedern, die im LoadProfileGenerator bisher so nicht abgebildet sind. Dazu gehört beispielsweise auch ein Seniorenheim mit 144 über-80-jährigen Bewohnern. Sämtliche derartigen Sonderfälle abzubilden ist nicht realistisch. Für 72% der insgesamt 15.910 Haushalte konnte jedoch ein in allen Kategorien passendes Template gefunden werden, ein großer Teil der Bevölkerung wird also bereits sehr gut repräsentiert.

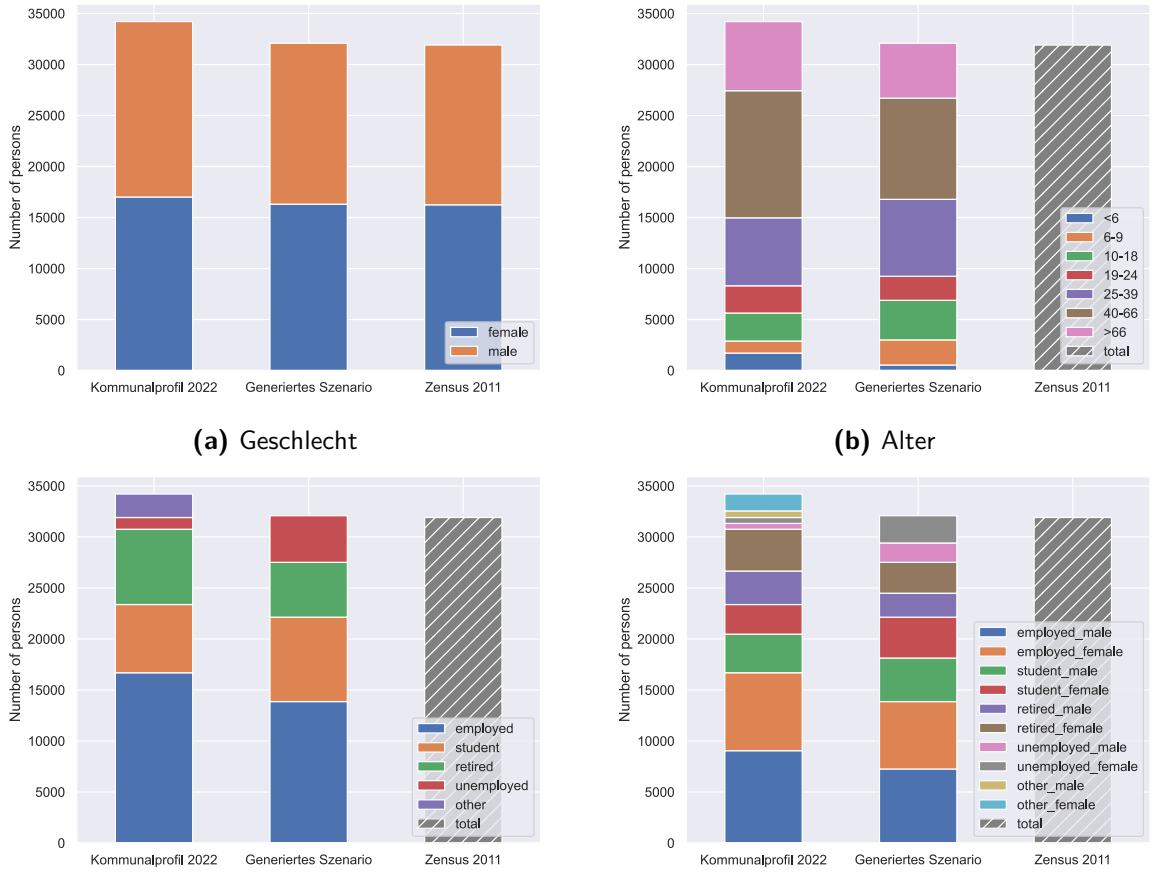
Tabelle 4.1 stellt die Einwohnerzahlen aus der offiziellen Bevölkerungsstatistik vom Statistischen Bundesamt, die Haushaltsdaten aus BUILDA und die Statistiken des daraus erzeugten Szenarios gegenüber. Die Abweichung bei der Gesamteinwohnerzahl liegt bei etwa 0.5%, und auch der Anteil an Frauen passt mit 50,8% gut zum Vergleichswert von 50,9%.

**Tabelle 4.1:** Vergleich der offiziellen Bevölkerungsstatistiken für Jülich aus dem Zensus 2011 mit den Daten in BUILDA und dem daraus generierten Szenario

| Kategorie | Zensus 2011 [61] | BUILDA [53] | Generiertes Szenario |
|---|---|---|---|
| Einwohner | 31.909 | 32.297 | 32.081 |
| Weiblich | 16.246 | 17.022 | 16.309 |
| Männlich | 15.663 | 15.275 | 15.772 |

Für das Jahr 2022 sind zusätzlich nach Altersklassen und Beschäftigung aufgeschlüsselte Statistiken verfügbar. Diese werden in Abbildung 4.3 mit den Daten des generierten Szenarios verglichen, um die Anteile der jeweiligen Gruppen zu überprüfen. Da die Einwohnerzahl zwischen 2011 und 2022 um über 2.000 Personen auf 34.208 angewachsen ist, sind die Werte hier im Vergleich zu den Szenariodaten durchgängig leicht erhöht. Soweit verfügbar werden zusätzlich die Statistiken aus den Zensusdaten von 2011 dargestellt, aus Datenmangel aber teils ohne Klassifizierung. Die Grafiken zeigen, dass die Verteilung insbesondere in Bezug auf Geschlecht und Alter im Szenario gut getroffen wird. Leichte Ungenauigkeiten finden sich beispielsweise beim Anteil an Kleinkindern unter 6 Jahren und dem Anteil an Schulkindern zwischen 6 und 9 Jahren. Insgesamt stimmt die Menge an Kindern unter 10 Jahren aber sehr gut, sodass die Auswirkungen für die Verhaltenssimulation gering ausfallen.

Wie in Abbildung 4.3c zu sehen ist, liegt die Menge an Arbeitslosen im Szenario vergleichsweise hoch. Dabei muss jedoch berücksichtigt werden, dass es im LoadProfileGenerator lediglich die in der Grafik gezeigten Beschäftigungsstatus gibt. Alle Personen, die nicht berufstätig oder im Ruhestand sind und nicht studieren, fallen somit automatisch in die Kategorie arbeitslos, wodurch diese Gruppe scheinbar überrepräsentiert wird. Zusammenfassend lässt sich sagen, dass in Bezug auf die betrachteten Bevölkerungsgruppen zwar teilweise Abweichungen vorkommen, mitunter bedingt durch unpräzise oder veraltete Eingangsdaten. Diese sind jedoch im Verhältnis gering und jede Gruppe wird auf angemessene Weise in der Szenariopopulation repräsentiert.




**(a)** Geschlecht

**(b)** Alter

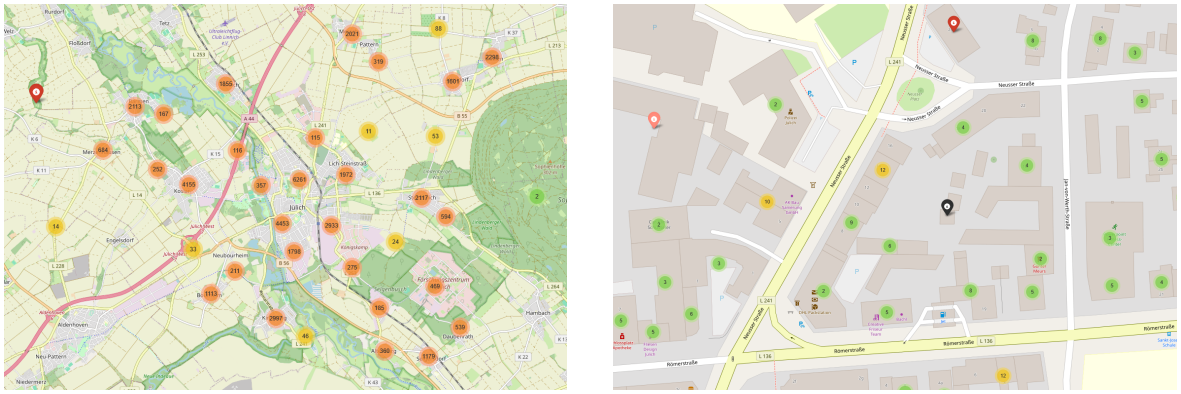
**(c)** Beschäftigungsstatus

**(d)** Beschäftigungsstatus & Geschlecht

**Abbildung 4.3:** Vergleich der Bevölkerungsstruktur innerhalb des generierten Jülich-Szenarios mit Daten aus dem Regionalprofil für die Stadt Jülich und dem Zensus 2011, anhand von Geschlecht (a), Alter (b), Beschäftigungsstatus (c) und Beschäftigungsstatus kombiniert mit Geschlecht (d). Die Vergleichsdaten vom Zensus 2011 enthalten keine Angaben zum Alter oder Beschäftigungsstatus, daher wird in (b), (c) und (d) nur die Gesamtanzahl dargestellt.

Neben der Bevölkerung ist die Menge an Points of Interest, die für das Szenario generiert werden, entscheidend. Anhand der Karten in Abbildung 4.4 ist zu erkennen, dass sich die enthaltenen Nichtwohngebäude über das gesamte Stadtgebiet erstrecken. Eine nähere Betrachtung wie in Abbildung 4.4b zeigt, dass meist mehrere Points of Interest pro Gebäude angelegt wurden. Für Innenstädte mit Mehrparteienhäusern, in denen mehrere Geschäfte enthalten sein können, entspricht das den Erwartungen. Die Anhäufung der Points of Interest wird dadurch verstärkt, dass, wie in Abschnitt 3.4.1 beschrieben, in der Regel zwei Point-of-Interest-Instanzen pro Geschäft angelegt werden, je eine für die Nutzung als Kunde und als Angestellter. Dazu kommen Points of Interest, die in Wohngebäuden liegen (siehe Tabelle 3.1), sodass sich insbesondere bei den in Innenstädten häufig vorkommenden gemischt genutzten Gebäuden teils 10 oder mehr Points of Interest ansammeln können.

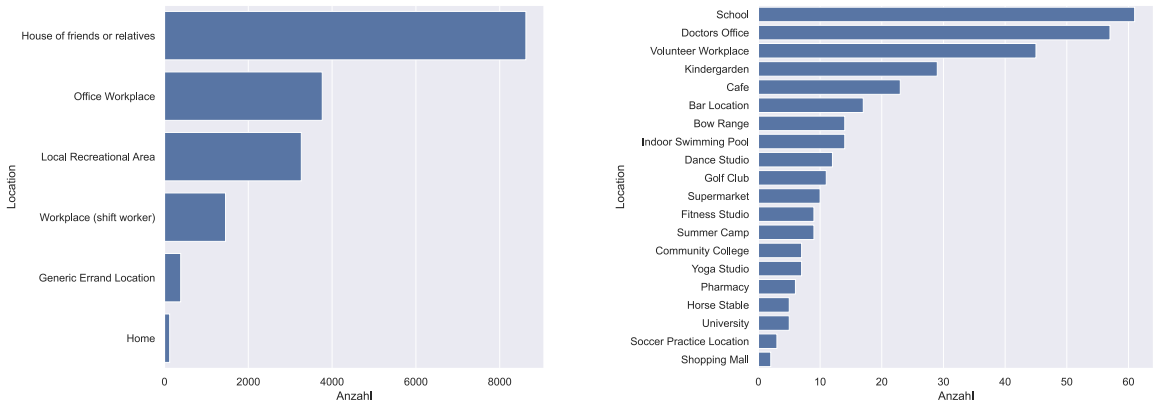
Insgesamt enthält das Szenario 17.869 Points of Interest, davon beziehen sich allerdings 8.625 auf Wohngebäude. Weitere 5.267 entfallen auf Arbeitsplätze. Ein großer Teil der übrigen Points of Interest repräsentiert nicht näher spezifizierte Orte für Freizeitgestaltung. Es verbleiben 703 Points of Interest für Geschäfte, Arztpraxen und Ähnliches. Die genaue Verteilung aller Points of Interest



**(a)** Stadt Jülich **(b)** Ausschnitt der Innenstadt

**Abbildung 4.4:** Übersicht über die im Szenario enthaltenen Nichtwohngebäude in ganz Jülich (a) und in einem Ausschnitt der Jülicher Innenstadt (b). Jeder Kreis steht dabei für eine Gruppe an Points of Interest, wobei die Zahl im Kreis die Anzahl angibt.

wird in Abbildung 4.5 gezeigt.




**(a)** Häufige Points of Interest **(b)** Seltenere Points of Interest

**Abbildung 4.5:** Anzahl der verschiedenen Point-of-Interest-Typen, sortiert nach Häufigkeit

Die Anzahl an Supermärkten, Arztpraxen oder ähnlichen Institutionen in Jülich lässt sich mangels geeigneter Daten nicht exakt validieren. Als Anhaltspunkt können die Gewerbeanmeldungen in Jülich aus dem Jahr 2015 verwendet werden [62]. Dieser Datensatz enthält sämtliche in Jülich gemeldeten Unternehmen, allerdings lediglich versehen mit nicht standardisierten Freitextbeschreibungen. Eine geeignete Zuordnung zu den Point-of-Interest-Typen der Städtesimulation ist damit schwierig. Aus diesem Grund wird lediglich überprüft, ob die Größenordnung der Werte für einige ausgewählte Kategorien plausibel ist.

Im Jahr 2015 waren insgesamt 2.140 Unternehmen in Jülich angemeldet. Da jedes Unternehmen als potenzieller Arbeitsplatz in Betracht kommt, kann diese Angabe mit der Gesamtanzahl an Arbeitsplatz-Points-of-Interest verglichen werden. Letztere ist mit 5.267 etwa doppelt so hoch. Verursacht wird diese Diskrepanz größtenteils durch nahe beieinander liegende Nichtwohngebäude, die vom CityScenarioGenerator als unterschiedliche Points of Interest gewertet werden. Die Aufenthaltsdauer am Arbeitsplatz wird jedoch typischerweise nicht durch die Anzahl anwesender Personen beeinflusst. Somit ist eine Überschätzung an dieser Stelle nicht problematisch, da die

Ortsangaben und damit die Wege der Personen auch in diesen Fällen korrekt sind.

Zum Überprüfen der Einkaufsmöglichkeiten werden alle Unternehmen gesammelt, die das Stichwort „Einzelhandel“ enthalten. In Jülich gibt es davon insgesamt 487. Alle Points of Interest mit einem entsprechenden Typ, also unter anderem Supermärkte, Apotheken und Einkaufszentren, ergeben zusammengenommen 448 Orte. Das entspricht 92% der gemeldeten Unternehmen. Damit wird dieser Bereich im Szenario gut abgebildet. Zuletzt kann exemplarisch die Anzahl an Apotheken überprüft werden, da diese in den Gewerbedaten eindeutig benannt sind. Sowohl das Szenario als auch die Vergleichsdaten enthalten 6 Apotheken, dieser Unternehmenstyp wird somit im Szenario optimal wiedergegeben.

## 4.3 Verhaltensvalidierung

Das menschliche Verhalten ist Kernelement der entwickelten Städtesimulation. Es wird mit hohem Detailgrad simuliert und ist unmittelbar für fast alle weiteren Ergebnisse der Simulation, einschließlich der Energieverbräuche, Reisen und Besuche in Geschäften maßgeblich verantwortlich. Aufgrund dieser zentralen Rolle wird das Verhalten gründlich anhand geeigneter Daten validiert. Hierfür wird das eigens zu diesem Zweck entwickelte Validierungsframework ETHOS.ActivityAssure [63] eingesetzt, das in Abschnitt 4.3.1 vorgestellt wird.

Dabei werden zunächst in Abschnitt 4.3.2 aggregierte Statistiken zu allen ausgeübten Aktivitäten über die gesamte Szenariopopulation und Simulationsdauer untersucht. Diese vermitteln einen ersten Eindruck von der Plausibilität der Ergebnisse. Darauf aufbauend werden in Abschnitt 4.3.3 einzelne Aktivitätsprofilkategorien genauer betrachtet.

### 4.3.1 Das ETHOS.ActivityAssure-Framework

ETHOS.ActivityAssure ist ein Framework, das bei der Validierung von Verhaltensmodellen unterstützen soll. Hierzu wurde die Veröffentlichung „ETHOS.ActivityAssure—An Open-Source Validation Framework for Synthetic European Activity Profiles“ [64] verfasst, die Entwicklung und Anwendung des Frameworks im Detail beschreibt. Die wichtigsten Inhalte werden im Folgenden kurz zusammengefasst.

ETHOS.ActivityAssure ermöglicht das Kategorisieren von Aktivitätsprofilen nach Tages- und Personentyp und die anschließende Berechnung von Verhaltensstatistiken für jede Kategorie. Damit können sowohl durch ein Modell wie die Städtesimulation generierte Daten als auch im Rahmen einer Umfrage gesammelte Daten aggregiert werden. Das ist insbesondere für Zeitverwendungsstudien nützlich, die für die Validierung von Verhaltensmodellen sehr hilfreich sind, aber aus Datenschutzgründen oft nicht öffentlich zur Verfügung gestellt werden. ETHOS.ActivityAssure ermöglicht es, aus Zeitverwendungsdaten aggregierte Statistiken zu erstellen, die ohne Probleme veröffentlicht werden und somit leicht zur Validierung herangezogen werden können. Dies wurde mit der europäischen Zeitverwendungsstudie HETUS von 2010 [65] gemacht, der entstandene Statistikdatensatz ist zusammen mit ETHOS.ActivityAssure frei verfügbar [66].

Für die eigentliche Validierung bietet ETHOS.ActivityAssure Funktionen zum Vergleichen zweier Statistikdatensätze anhand von Grafiken und Indikatoren. Beides kann auf aggregierter Ebene oder einzeln pro Aktivität und Personenkategorie ausgewertet werden. Um den Vergleich zu erleichtern,

werden alle Aktivitäten dabei in 15 Kategorien eingeteilt.

Für die Validierung der Städtesimulation werden alle 32.081 erzeugten Aktivitätsprofile kategorisiert und anschließend dazu genutzt, Verhaltensstatistiken zu berechnen. Diese werden dem aus HETUS erstellten Validierungsdatensatz gegenübergestellt, um etwaige Abweichungen zu finden und die betroffenen Personenkategorien zu identifizieren.

### 4.3.2 Aggregierte Verhaltensvalidierung

Für eine erste Überprüfung der Aktivitätsdaten werden die mit ETHOS.ActivityAssure erstellten Statistiken über alle Kategorien aggregiert. Damit können Aussagen darüber getroffen werden, wie gut die Städtesimulation im Allgemeinen Aktivitätsmuster abbildet, und ob es bei bestimmten Aktivitäten grundsätzlich größere Abweichungen gibt. Bei der Aggregation muss jede Kategorie geeignet gewichtet werden, um die betrachtete Bevölkerung angemessen zu repräsentieren. Der HETUS-Datensatz enthält bereits Gewichte, diese sind allerdings für eine Abbildung der gesamten deutschen Bevölkerung gedacht. Stattdessen werden zunächst die in der Städtesimulation generierten Aktivitätsprofile verarbeitet und kategorisiert. Die Anzahl an Aktivitätsprofilen pro Kategorie wird dann als Gewicht verwendet, sowohl für die Statistiken der Städtesimulation als auch für die Zeitverwendungsdaten. Auf diese Weise berücksichtigen beide so erzeugten aggregierten Statistiken die demographische Verteilung im generierten Jülich-Szenario.

#### Zeitverbrauch der Aktivitäten

Abbildung 4.6 zeigt anhand dieser Statistiken, wie viel Zeit im Schnitt pro Tag für verschiedene Aktivitäten aufgewendet wird. Dabei ist gut zu erkennen, dass die wesentliche Aufteilung aus den Zeitverwendungsdaten von der Städtesimulation wiedergegeben wird. Die mit Abstand zeitintensivste Aktivität ist Schlafen mit ungefähr 8 Stunden täglich. Die insgesamt geringe Arbeitszeit kommt zustande, da für die gezeigten Werte der Durchschnitt aus Arbeits- und Ruhetagen sowie arbeitenden und nicht berufstätigen Personen gebildet wird.

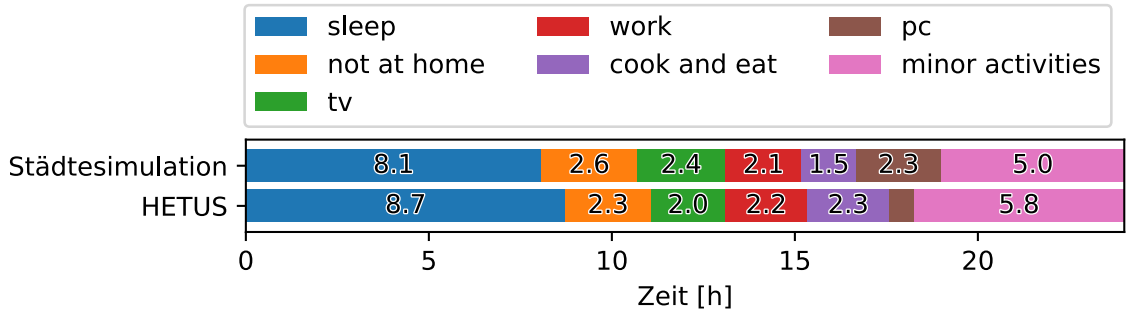



**Abbildung 4.6:** Durchschnittliche Zeit in Stunden, die pro Tag mit bestimmten Aktivitäten zugebracht wird, für die HETUS-Zeitverwendungsdaten und die Städtesimulation im Vergleich. Aktivitäten mit geringerem zeitlichem Umfang wurden zusammengefasst.

Die Zeit für Schlaf und Arbeit wird von der Städtesimulation leicht unterschätzt, die mit Fernsehen und sonstigen Aktivitäten außer Haus zugebrachte Zeit wird dagegen leicht überschätzt, aber die Unterschiede sind lediglich gering. Deutlichere Abweichungen gibt es beim Kochen und Essen

sowie bei der PC-Nutzung. Letztere ist im LoadProfileGenerator als eines von mehreren Hobbys enthalten. Da in einem Modell nicht die gesamte Vielfalt möglicher Freizeitbeschäftigungen abgebildet werden kann, wurde eine exemplarische Auswahl getroffen und implementiert. Diese Hobbys sind damit naturgemäß überrepräsentiert, da sie stellvertretend für alle ähnliche Aktivitäten stehen. Die an dieser Stelle sehr spezifische Aktivitätskategorisierung sorgt dafür, dass das Beispiel der PC-Nutzung lediglich besonders heraussticht. Für das sonstige Verhalten, inklusive Außer-Haus-Aktivitäten und Reisen, spielt die genaue Art der Freizeitbeschäftigungen zuhause keine große Rolle.

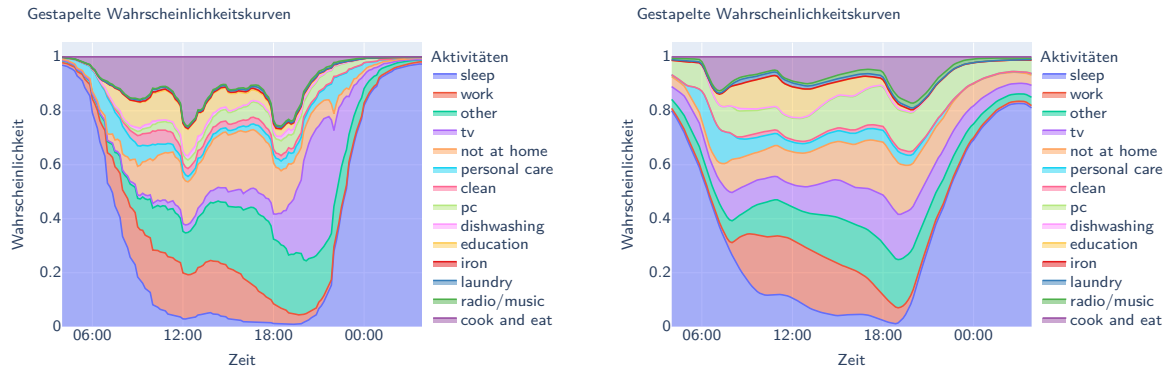
Dasselbe gilt auch für Kochen und Essen. Eine Schwierigkeit besteht hierbei darin, dass der LoadProfileGenerator Kochen und Essen stets als kombinierte Aktivität modelliert. Dadurch sind diese beiden Teilbereiche im Nachhinein nicht mehr unterscheidbar. Ein zu hoher Anteil an Kochen würde aber aufgrund des hohen Energiebedarfs bei der Lastvalidierung auffallen. Insgesamt wird ein Anteil von 87% der Zeit korrekt auf die Aktivitätskategorien verteilt. Die Zeitanteile stimmen somit gut überein, in einigen Bereichen wäre aber eine leichte Verbesserung möglich. Um nähere Informationen über mögliche Ursachen der Abweichungen zu erlangen, wird im Folgenden untersucht, wie die Abweichungen im Tagesverlauf schwanken.

#### Zeitlicher Verlauf des Verhaltens

Bei der Simulation von menschlichem Verhalten kommt es nicht nur auf eine passende Verteilung der verfügbaren Zeit an, sondern auch auf die Veränderungen im Verlauf des Tages. Beispielsweise schläft die Mehrheit der Bevölkerung im Regelfall nachts und arbeitet tagsüber. Diese Muster müssen durch die Verhaltenssimulation wiedergegeben werden, und werden daher im Folgenden anhand von Aktivitätswahrscheinlichkeitskurven untersucht. Eine solche Wahrscheinlichkeitskurve gibt dabei an, mit welcher Wahrscheinlichkeit eine Aktivität zu verschiedenen Uhrzeiten ausgeübt wird.

Abbildung 4.7 zeigt die Wahrscheinlichkeitskurven für alle Aktivitäten gestapelt an, für die Validierungsdaten und für die Simulationsergebnisse. Hier ist zu sehen, dass die relevantesten Muster sich in beiden Datensätzen stark ähneln. Insbesondere das typische Muster für Schlaf ist deutlich erkennbar, sowie die Hauptarbeitszeit. Auch die drei klassischen Tageszeiten für Mahlzeiten sind sichtbar, allerdings leicht anders ausgeprägt. So findet im LoadProfileGenerator das Frühstück im Schnitt etwa eine Stunde früher statt. Zudem ist die Wahrscheinlichkeit für das Mittagessen deutlich weniger erhöht als in den Validierungsdaten. Das liegt daran, dass der LoadProfileGenerator Arbeit als eine einzelne ununterbrochene Tätigkeit modelliert, ohne jegliche Mittagspausen. Auf diese Weise lässt sich auch der zu geringe Anteil von Kochen und Essen insgesamt in der Städtesimulation erklären, der in Abbildung 4.6 sichtbar wurde.

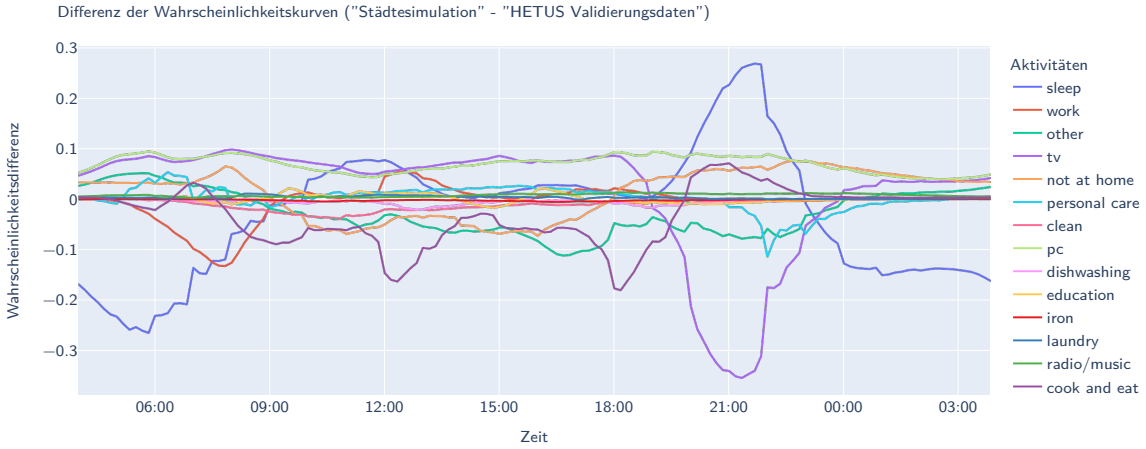
Die Kurven in Abbildung 4.7c zeigen jeweils die Differenz zwischen den beiden Wahrscheinlichkeitskurven für eine Aktivität aus den Ergebnissen der Städtesimulation und den Validierungsdaten. Positive Werte geben somit an, dass die entsprechende Aktivität zu dieser Uhrzeit in der Städtesimulation häufiger stattfindet als in den Validierungsdaten, bei negativen Werten ist es umgekehrt. Auf diese Weise sind zeitliche Unterschiede klarer erkennbar. So fällt anhand der Ausschläge der blauen Kurve für Schlaf auf, dass die durchschnittliche Schlafenszeit im LoadProfileGenerator leicht nach vorne verschoben ist. Zu erkennen ist das an der negativen Differenz um 06:00 Uhr




**(a)** HETUS




**(b)** Städtesimulation




**(c)** Differenz

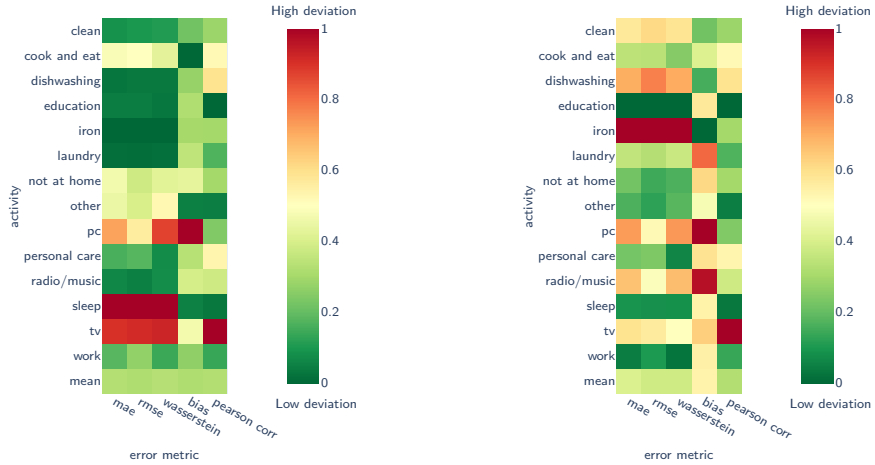
**Abbildung 4.7:** Gestapelte Wahrscheinlichkeitskurven für die HETUS-Daten (a) und die Städtesimulation (b), sowie die Differenz dazwischen (c)

und der positiven Differenz etwa gleicher Höhe um 21:00 Uhr. Morgens schlafen in der Städtesimulation im Durchschnitt also weniger Personen, am Abend dafür mehr. Wie zuvor bereits festgestellt wurde, stimmt die Schlafdauer insgesamt dennoch gut mit den Validierungsdaten überein.

Um einzelne Aktivitäten mit besonders hoher Abweichung zu identifizieren, werden die Vergleichsindikatoren von ETHOS.ActivityAssure genutzt. Dazu gehören der Mean Absolute Error (MAE), der Root Mean Square Error (RMSE), die Wasserstein-Distanz, der Bias und der Pearson-Korrelationskoeffizient. Formeln für alle Indikatoren sind in der Veröffentlichung zu ETHOS.ActivityAssure angegeben [64]. Alle fünf Indikatoren beziehen sich auf die Differenz der Wahrscheinlichkeitskurven der jeweiligen Aktivität. Jeder der Indikatoren hebt bestimmte Aspekte der Kurven stärker hervor. Während MAE, RMSE und Wasserstein-Distanz eine größere Differenz auch entsprechend als stärkere Abweichung werten, betrachtet der Pearson-Korrelationskoeffizient lediglich die Veränderung der Wahrscheinlichkeiten und nicht deren absolute Werte. Beim Bias gleichen sich positive und negative Abweichungen aus, dieser Indikator gibt also an, wie stark sich die Gesamtzeit unterscheidet, die mit einer Aktivität zugebracht wird.

Abbildung 4.8 stellt die Werte der fünf Indikatoren für alle Aktivitäten in den aggregierten Statistikdatensätzen dar. Den gesamten Satz an Indikatoren gibt es zusätzlich in einer skalierten

Version. Dabei werden die Indikatorwerte antiproportional zum Gesamtanteil der jeweiligen Aktivität skaliert. Auf diese Weise werden Abweichungen bei selteneren Aktivitäten wie Putzen stärker berücksichtigt. Die unskalierten Indikatoren werden von den häufigsten Aktivitäten wie Schlaf dominiert. Dahingegen gehören die skalierten Werte für Schlaf zu den besten Indikatorwerten. Die hohen Werte für Schlaf hängen somit in erster Linie damit zusammen, dass die Wahrscheinlichkeitskurve für diese Aktivität über einen großen Teil des Tages sehr hoch liegt, bei nahezu 100%. Eine vergleichsweise kleine zeitliche Abweichung, wie das zu frühe Aufstehen im LoadProfileGenerator, schlägt sich dadurch stark in den unskalierten Indikatoren nieder.




**(a)** Standard **(b)** Skaliert

**Abbildung 4.8:** Heatmaps mit den Indikatoren aus ETHOS.ActivityAssure zum Vergleich der HETUS-Daten mit den Simulationsergebnissen, in normaler (a) und skalierter (b) Variante. Zur besseren Lesbarkeit wurden alle Indikatoren auf den einheitlichen Wertebereich $[0, 1]$ transformiert, mit exakter Übereinstimmung beim Wert $0$.

Die Indikatoren für Fernsehen deuten in beiden Versionen auf eine hohe Abweichung hin, skaliert und unskaliert. Der Pearson-Korrelationskoeffizient liegt dabei besonders hoch. Das ist ein deutliches Indiz dafür, dass die Form der Wahrscheinlichkeitskurven im LoadProfileGenerator nicht zu der der Validierungsdaten passt, dass also das zeitliche Muster nicht übereinstimmt. Um das zu bestätigen, werden in Abbildung 4.9 beide Wahrscheinlichkeitskurven übereinander gezeigt. Hier ist in den HETUS-Daten ein markantes Muster zu erkennen, mit dem Großteil der Fernsehaktivitäten zwischen 18 und 0 Uhr. Dieses fehlt in den Simulationsergebnissen völlig, hier findet Fernsehen nahezu gleichmäßig verteilt zu jeder Tageszeit statt. An dieser Stelle könnte der LoadProfileGenerator daher zukünftig verbessert werden, indem ein Teil der Fernsehaktivitäten auf dieses Zeitfenster beschränkt wird.

Zu den am besten vom LoadProfileGenerator abgebildeten Aktivitäten gehören Bildung und persönliche Pflege, die bei sämtlichen Indikatoren gute Werte erzielen. Die Kurven in Abbildung 4.10 stimmen, bis auf einige kleinere Abweichungen wie einen fehlenden abendlichen Anstieg bei Pflege, sehr gut überein.

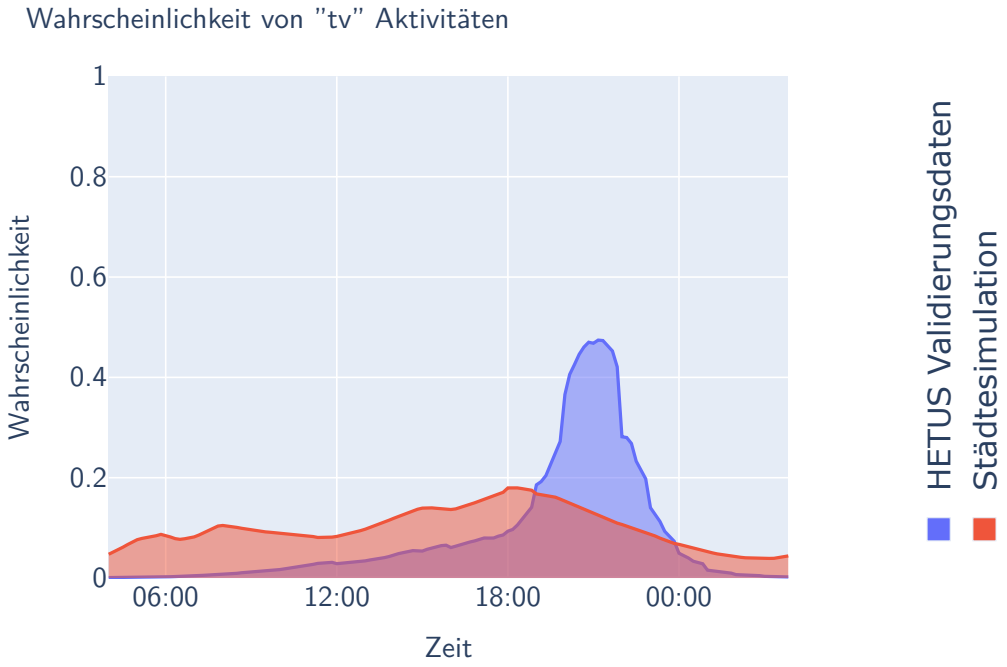
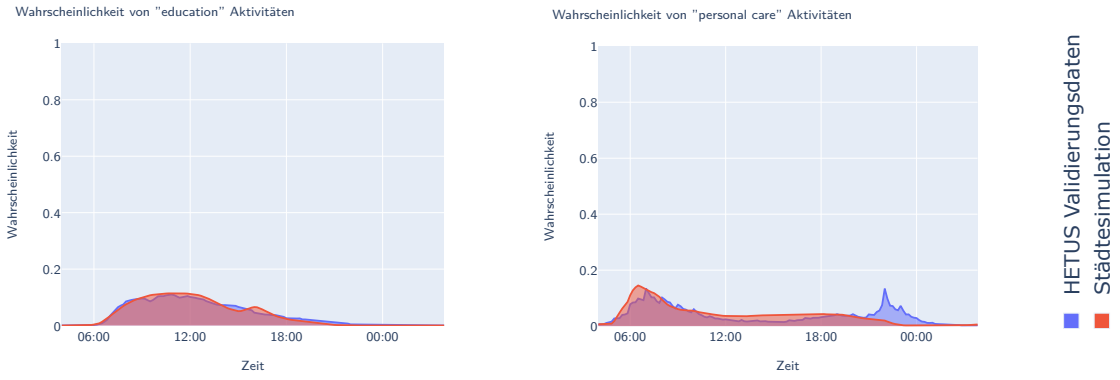



**Abbildung 4.9:** Wahrscheinlichkeitskurve für die Aktivität Fernsehen, in den Validierungsdaten und den Ergebnissen der Städtesimulation




**(a)** **(b)**

**Abbildung 4.10:** Wahrscheinlichkeitskurven für die Aktivitäten Bildung (a) und persönliche Pflege oder Körperpflege (b), in den Validierungsdaten und den Ergebnissen der Städtesimulation

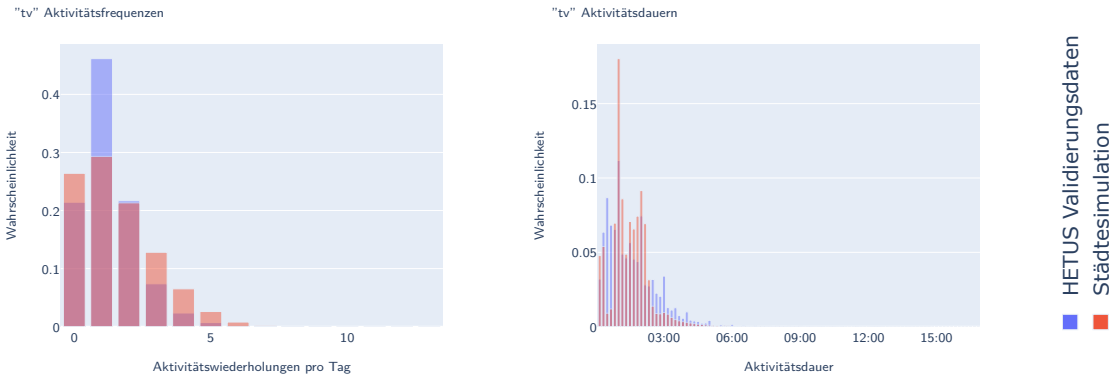
### Aktivitätsdauer und -frequenz

Wahrscheinlichkeitskurven geben zwar Auskunft über die Uhrzeiten, zu denen Aktivitäten ausgeübt werden, daraus lässt sich jedoch nicht ablesen, um wie viele separate Wiederholungen derselben Aktivitäten es sich handelt, und wie lange diese andauern. Je nach Aktivität kann es hierbei große Unterschiede geben. Diese Eigenschaften beeinflussen vor allem die Regelmäßigkeit des Verhaltens, die wiederum große Bedeutung für verschiedene haushaltsbezogene Fragestellungen hat, wie etwa die Rentabilität von Heimenergiesystemen. Daher werden diese Aspekte gesondert untersucht.

Bei der Aktivität Fernsehen zeigt sich in Abbildung 4.11, dass die Verteilungen im Allgemeinen

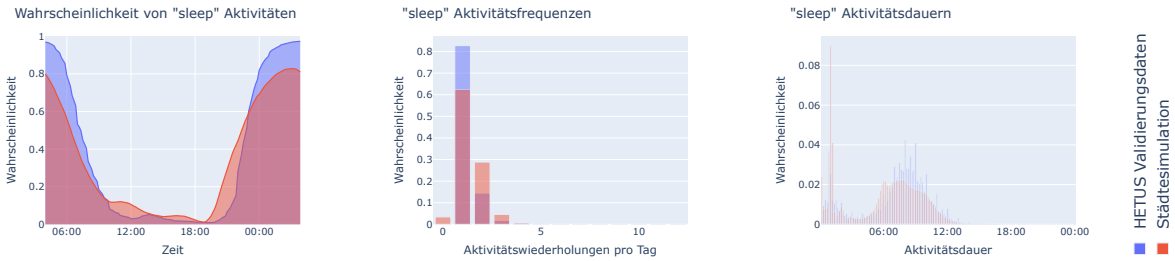
beide gut passen. In den Validierungsdaten ist die Häufigkeit für exakt eine Fernsehaktivität pro Tag noch deutlich höher, aber in Bezug auf die ungefähre Verteilung der Frequenz und auch der Aktivitätsdauer stimmen die Ergebnisse der Städtesimulation gut mit dem HETUS-Datensatz überein. Das bedeutet, dass die Modellierung dieser Aktivitäten im LoadProfileGenerator grundsätzlich gut funktioniert. Lediglich der zeitliche Fokus in den Abendstunden könnte, wie zuvor erwähnt, stärker betont werden.

Aufgrund ihrer großen Bedeutung für den Tagesablauf werden die Aktivitäten Schlafen und Arbeiten im Detail untersucht. Wie in Abbildung 4.12 anhand der Häufigkeiten und Dauern zu sehen ist, wird im LoadProfileGenerator der Anteil kurzer Schlafaktivitäten von etwa einer Stunde zu hoch eingeschätzt, in den Validierungsdaten ist die Wahrscheinlichkeit für nur einen einzigen Schlafzeitraum pro Tag höher. In Kombination mit der leicht erhöhten Wahrscheinlichkeitskurve gegen 12 Uhr deutet das darauf hin, dass der klassische Mittagsschlaf in der Städtesimulation übermäßig häufig vorkommt. Neben der bereits erwähnten Verschiebung der Schlafenszeit nach vorne fällt auch ein leichter Unterschied in der Steigung auf, die Wahrscheinlichkeitskurve der Städtesimulation fällt morgens langsamer ab, und auch der abendliche Anstieg ist flacher. Die Varianz der Schlafenszeit ist im LoadProfileGenerator somit im Vergleich zu den HETUS-Daten etwas zu hoch. Diese Abweichungen betreffen jedoch allesamt eher geringfügige Details. Die Verteilung der Dauern langer Schlaf-Aktivitäten stimmt gut überein, und insgesamt wird vor allem




**(a)** **(b)**

**Abbildung 4.11:** Verteilungen der Wiederholhäufigkeit (a) und Dauer (b) von Fernsehaktivitäten

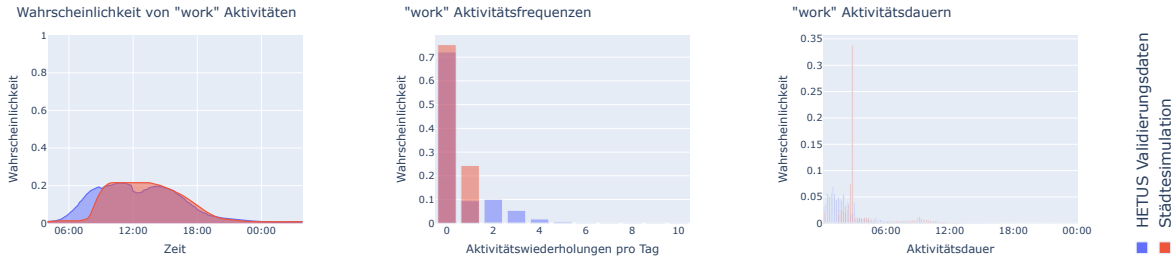



**(a)** **(b)** **(c)**

**Abbildung 4.12:** Wahrscheinlichkeitskurve (a), Häufigkeit (b) und Dauer (c) von Schlaf-Aktivitäten pro Tag, Vergleich der Validierungsdaten (blau) mit den Ergebnissen der Städtesimulation (rot)

das Schlafverhalten in der Nacht gut abgebildet.

Bei den Statistiken der Arbeitsaktivitäten in Abbildung 4.13 fallen in allen Aspekten leichte Unterschiede auf. In den HETUS-Daten gibt es ein merkliches Tal in der Wahrscheinlichkeitskurve gegen Mittag, Arbeit findet häufiger pro Tag statt, ist im Schnitt dafür aber kürzerer. Alle diese Unterschiede lassen sich auf die bereits beschriebene fehlende Modellierung einer Mittagspause zurückführen. In den HETUS-Daten wird ein Vollzeit-Arbeitstag durch die Pause geteilt, sodass sich zwei kurze Arbeitsabschnitte ergeben, in der Städtesimulation ist das dagegen nicht so. Diese Unterschiede sind somit lediglich Modellierungsartefakte, die ignoriert werden können, solange der Mittagspause bei der Arbeit für den Simulationszweck keine besondere Relevanz zukommt. Auffällig ist dagegen der tendenziell spätere Arbeitsbeginn im LoadProfileGenerator, besonders im Hinblick auf das frühere Aufstehen.




**(a)** **(b)** **(c)**

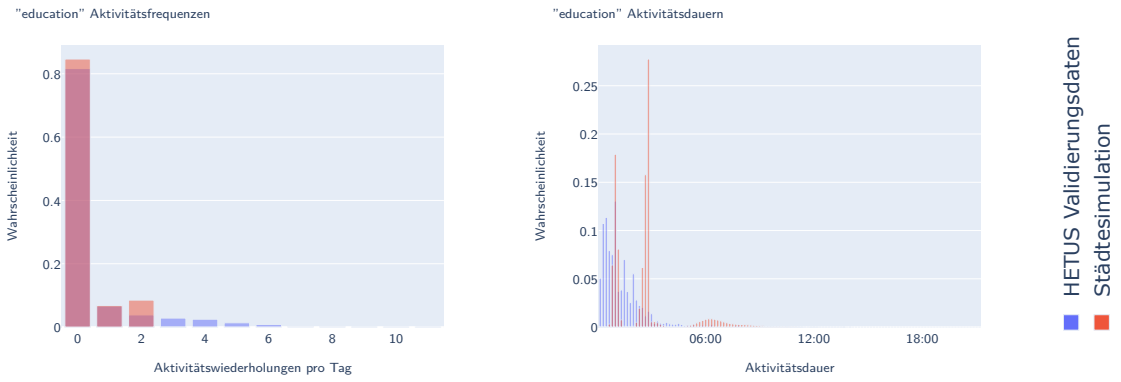
**Abbildung 4.13:** Wahrscheinlichkeitskurve (a), Häufigkeit (b) und Dauer (c) von Arbeitsaktivitäten pro Tag, Vergleich der Validierungsdaten (blau) mit den Ergebnissen der Städtesimulation (rot)

Die wesentlichen Eigenschaften der Arbeit werden ausreichend abgedeckt. Die hier gezeigten Daten sind jedoch Durchschnittswerte der gesamten Simulationspopulation. Wichtiger ist an dieser Stelle eine passende Abbildung aller Untergruppen, also der arbeitenden sowie der arbeitslosen Teile der Bevölkerung. Daher werden in Abschnitt 4.3.3 einige Aktivitätsprofil-Kategorien differenzierter betrachtet.

Beim Abgleich der Aktivitätsdauern bei Bildung in Abbildung 4.14 fallen deutliche Unterschiede auf, die im Kontrast zu den gut passenden Frequenzen und Wahrscheinlichkeitskurven stehen. Die zwei herausstechenden Spitzen in den Daten der Städtesimulation sind dabei eine Folge der gewählten Methode zur Verhaltenssimulation. Im LoadProfileGenerator sind nur verhältnismäßig wenige Aktivitäten dieser Kategorie definiert. Diese haben größtenteils eine hohe Dauer, da, wie bei der Arbeit, auch hierbei keine Pausen modelliert werden. Die Dauer einzelner Aktivitäten wird zwar mithilfe einer Normalverteilung variiert, durch die geringe Anzahl an Varianten ergeben sich aber dennoch einzelne, klar unterscheidbare Bereiche für mögliche Dauern. Ähnlich wie bei der Arbeit gilt jedoch auch hier, dass diese Ungenauigkeit des Modells für die Analysen in dieser Arbeit keine Rolle spielt

### 4.3.3 Verhaltensvalidierung nach Personen- und Tageskategorien

Die Untersuchung im vorherigen Kapitel hat gezeigt, dass die Städtesimulation auf aggregierter Ebene die Verhaltensmuster der HETUS-Daten im Wesentlichen gut abbilden kann. Allerdings




**(a)** **(b)**

**Abbildung 4.14:** Verteilungen der Wiederholhäufigkeit (a) und Dauer (b) von Bildungsaktivitäten

unterscheidet sich das Verhalten innerhalb der Bevölkerung mitunter stark. Beispielsweise hat der nicht-arbeitende Teil im Vergleich zu Vollzeitbeschäftigten jeden Tag etwa 8 bis 9 Stunden zusätzlich zur Verfügung, die für andere Aktivitäten genutzt werden. Um zu überprüfen, ob die Städtesimulation auch diese einzelnen Gruppen für sich genommen plausibel modelliert, muss eine getrennte Validierung erfolgen. Dafür kann die Kategorisierung in ETHOS.ActivityAssure genutzt werden. Diese klassifiziert jedes 24h-Aktivitätsprofil nach Land, Geschlecht, Beschäftigungsstatus sowie danach, ob es sich um einen Arbeitstag oder einen freien Tag handelt. Durch die Zusammenlegung der Tagestypen bei Beschäftigungsstatus, die davon nicht betroffen sind, ergeben sich insgesamt 16 Kategorien pro Land. Der LoadProfileGenerator enthält zum aktuellen Zeitpunkt keine teilzeitbeschäftigten Männer, somit verbleiben noch 14 Kategorien. Da diese Arbeit sich auf die Anwendung der Städtesimulation auf deutsche Szenarien beschränkt, wird kein internationaler Vergleich durchgeführt.

### Verhaltensunterschiede zwischen Profiltypen

Abbildung 4.15 bietet eine Übersicht über die Zeit, die in den 14 Kategorien insgesamt mit verschiedenen Aktivitäten zugebracht wird. Die jeweils passenden Kategorien der Validierungsdaten und der Simulationsergebnisse sind dabei direkt untereinander dargestellt. Diese Übersicht zeigt deutlich, wie verschieden das Verhalten der Personengruppen sein kann. Am deutlichsten stechen die Unterschiede in der Arbeitszeit heraus. Die Zeiteinteilung entspricht hier vollständig den Erwartungen, mit im Durchschnitt über 8 Stunden pro Tag für Vollzeitbeschäftigte und etwa 6 Stunden pro Tag für Teilzeitbeschäftigte, jeweils an Arbeitstagen. An freien Tagen sowie in Profilen der anderen Kategorien findet kaum Arbeit statt. Eine Sonderrolle nehmen hierbei Studierende ein, da ihre Hauptbeschäftigung in die Kategorie Bildung fällt. In Abbildung 4.15 ist Bildung aufgrund des insgesamt kleinen Anteils in den sonstigen Aktivitäten enthalten, daher ist bei Studierenden an Arbeitstagen ein deutlich erhöhter Anteil dieser sonstigen Aktivitäten zu sehen. Im Schnitt beträgt die an diesen Tagen von Studierenden für Bildung aufgebrachte Zeit etwa 6,5 Stunden.

Die Unterschiede in der Arbeitszeit sind allerdings durch die Aktivitätsprofilkategorisierung vorgegeben. Bemerkenswerter sind die weiteren Verhaltensunterschiede bei Aktivitätstypen, die nicht in die Kategorisierung einfließen und dennoch innerhalb der Profilkategorien stark variieren. Diese

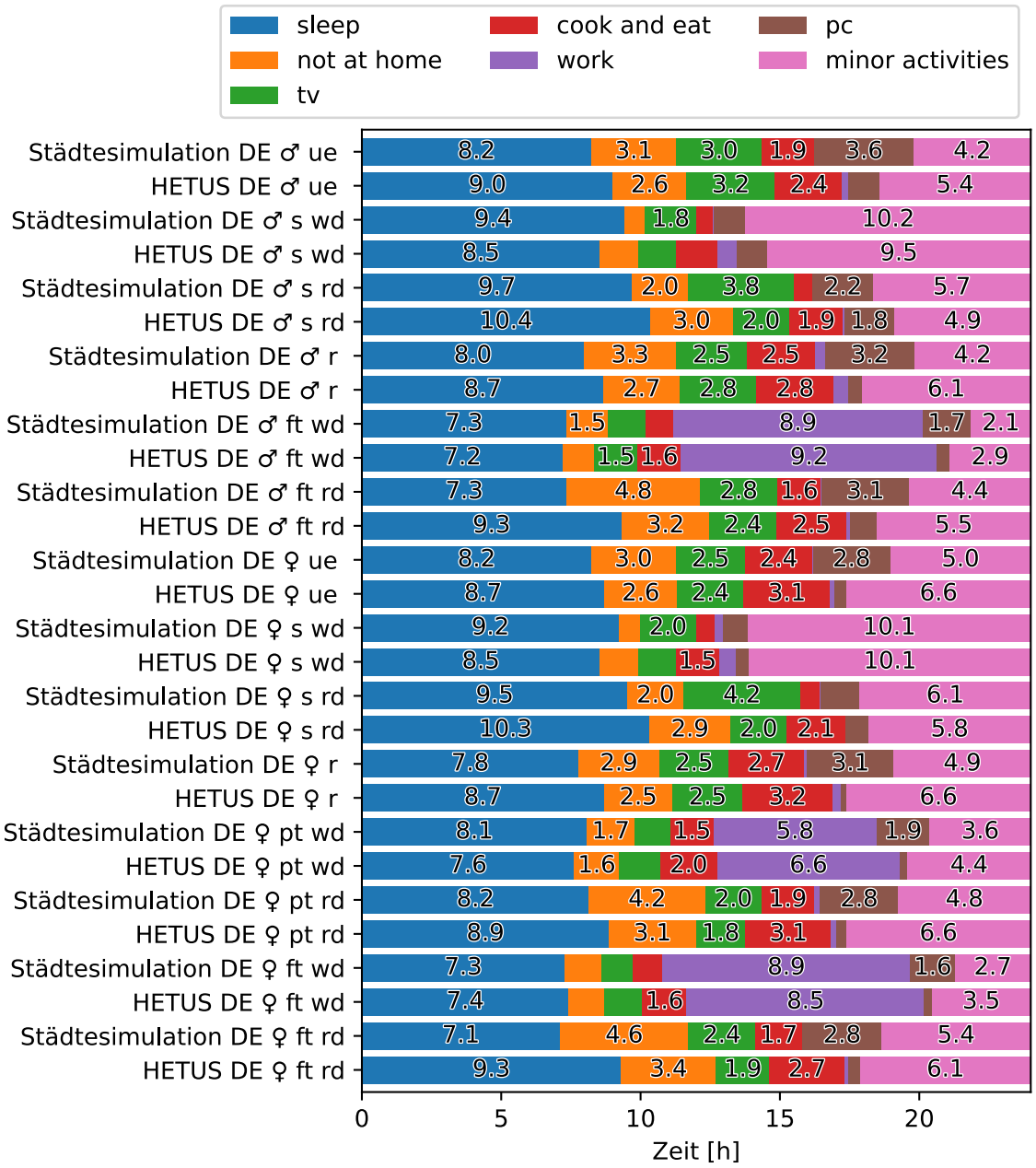



**Abbildung 4.15:** Durchschnittliche Zeit in Stunden, die pro Tag mit bestimmten Aktivitäten zugebracht wird, einzeln für verschiedene Personen- und Tageskategorien. Jede Kategorie wird jeweils für die Zeitverwendungsdaten und die Städtesimulation gezeigt. Bedeutung der Abkürzungen: ♀=weiblich, ♂=männlich, ue=arbeitslos, s=studierend, r=Ruhestand, ft=Vollzeit, pt=Teilzeit, wd=Arbeitstag, rd=freier Tag

Unterschiede gehen deutlich über die Haupttätigkeit der Person hinaus und zeigen damit, welchen Einfluss die Merkmale der Kategorisierung auf den üblichen Tagesablauf haben. Sehr gut zu erkennen sind beispielsweise die signifikanten Schwankungen in der Schlafenszeit. Grundsätzlich wird an freien Tagen mehr geschlafen als an Arbeitstagen, in den meisten Kategorien beträgt der Unterschied knapp 2 Stunden. Studierende schlafen an beiden Tagestypen im Durchschnitt am längsten und Rentner am kürzesten. Die mit Fernsehen verbrachte Zeit ist vor allem unter den Arbeitslosen und Rentnern verhältnismäßig hoch und in der arbeitenden Bevölkerung niedrig,

insbesondere an Arbeitstagen. Im Hinblick auf PC-Nutzung lässt sich eine klare Tendenz in Bezug auf die Geschlechter ausmachen. Bei allen weiblichen Profiltypen liegt die tägliche PC-Nutzung nur etwa halb so hoch wie beim männlichen Gegenstück. Bei persönlicher Pflege ist es umgekehrt, wobei die Differenzen geringer ausfallen. Anderweitige Aktivitäten außer Haus, zu denen etwa Einkäufe oder verschiedene Freizeitbeschäftigungen zählen, finden bei der arbeitenden Bevölkerung sehr viel häufiger an freien Tagen statt als an Arbeitstagen, mit einem teils um fast 200% erhöhten Zeitanteil.

Diese deutlichen Variationen im Verhalten zeigen, dass es unerlässlich ist, alle Bevölkerungsgruppen individuell abzubilden und ihr Verhalten entsprechend angepasst zu simulieren, um repräsentative Ergebnisse zu erhalten. Die Städtesimulation ist dazu in der Lage und reproduziert die beobachteten Muster größtenteils sehr gut. Insbesondere die geschlechts- und altersabhängigen Unterschiede in der aufgewendeten Zeit für Fernsehen, persönliche Pflege und Aktivitäten außer Haus finden sich in den Simulationsergebnissen wieder. Auch die verschiedenen Schlafmengen der Personenkategorien lassen sich erkennen, lediglich die Differenz zwischen freien Tagen und Arbeitstagen ist zu gering. Studierende werden in diesem Bezug teilweise nicht passend eingeordnet, so wird in dieser Gruppe die Zeit für Außer-Haus-Aktivitäten im Verhältnis zur restlichen Bevölkerung unterschätzt und die Zeit für Fernsehen an freien Tagen überschätzt. Zudem wird die geschlechtsspezifische PC-Nutzung nicht berücksichtigt, da diese Aktivität wie bereits beschrieben im LoadProfileGenerator als ein Beispiel-Hobby umfassend eingesetzt wird.

Die wesentlichen Aussagen, die aus Abbildung 4.15 hervorgehen, sind somit zum einen, dass das Verhalten abhängig von der Personengruppe stark variiert, und zum anderen, dass diese Variationen von der Städtesimulation gut wiedergegeben werden. Um die Qualität der Simulationsergebnisse pro Personenkategorie zu quantifizieren, werden in Abbildung 4.16 die jeweiligen korrekt auf die Aktivitätstypen verteilten Anteile der Zeit dargestellt. Die Werte sind absteigend sortiert, um hervorzuheben, in welchen Bereichen die Simulation bessere oder weniger gute Ergebnisse liefert. Auffällig ist hierbei, dass Arbeitstage, insbesondere bei Vollzeitbeschäftigten oder Studierenden, am besten repräsentiert werden, wohingegen die Abweichungen an freien Tagen etwas höher liegen. Mit einem Minimum von 80% Übereinstimmung wird jedoch auch die schlechteste Kategorie, vollzeitbeschäftigte Frauen an freien Tagen, ausreichend gut wiedergegeben. Die Städtesimulation ist somit grundsätzlich in der Lage, das spezifische Verhalten unterschiedlicher Bevölkerungsgruppen realistisch zu modellieren.

### Zeitanteile nach Kategorien

Um genau zu überprüfen, wie gut die Städtesimulation verschiedene Aktivitäten für die einzelnen Profiltypen abbildet, werden erneut Heatmaps mit Vergleichsindikatoren zu Hilfe genommen. Anhand der Bias-Heatmap in Abbildung 4.17 können Unterschiede im Zeitanteil der Aktivitäten leicht identifiziert werden. Hier ist zu beachten, dass der Bias an dieser Stelle nicht in einen Absolutwert transformiert wurde und vorzeichenbehaftet ist. Positive Werte, in der Grafik in gelb, zeigen einen im Vergleich zu HETUS zu hohen Zeitanteil der Aktivität in der Städtesimulation an, negative Werte in blau entsprechend einen zu geringen Anteil. Eine gute Übereinstimmung entspricht somit Werten nahe 0. Die skalierten Bias-Werte in Abbildung 4.17b setzen dabei die Abweichung in Relation zum Zeitanteil einer Aktivität.

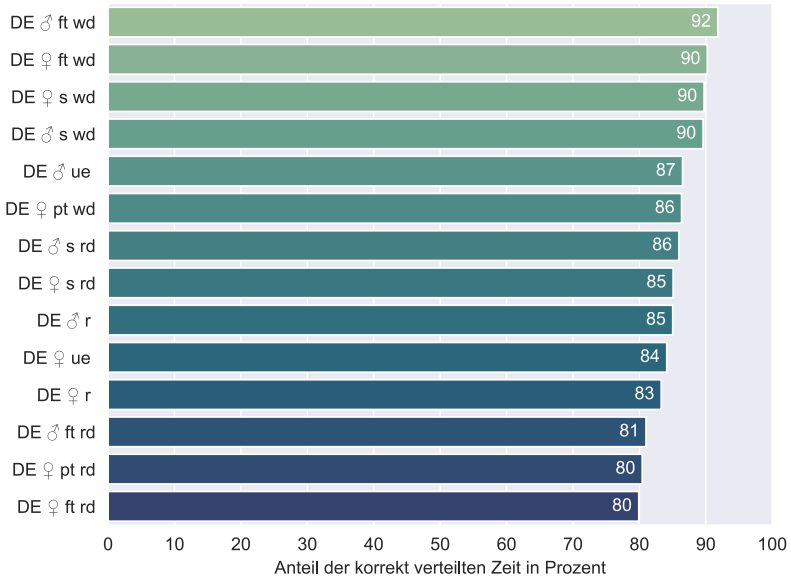



**Abbildung 4.16:** Anteil der Zeit, die korrekt auf die Aktivitätskategorien verteilt wird, für jeden Profiltyp. Bedeutung der Abkürzungen: ♀=weiblich, ♂=männlich, ue=arbeitslos, s=studierend, r=Ruhestand, ft=Vollzeit, pt=Teilzeit, wd=Arbeitstag, rd=freier Tag

Bei den unskalierten Werten fallen zunächst die hohen Abweichungen bei der PC-Nutzung auf, die zu den vorhergehenden Beobachtungen passen. Hier ist jedoch auch anzumerken, dass besonders hohe Abweichungen sich auf einige Profiltypen beschränken, wohingegen die Zeitanteile bei Studierenden gut übereinstimmen. Für eine Verbesserung müssten daher nur die anderen Personengruppen angepasst werden.

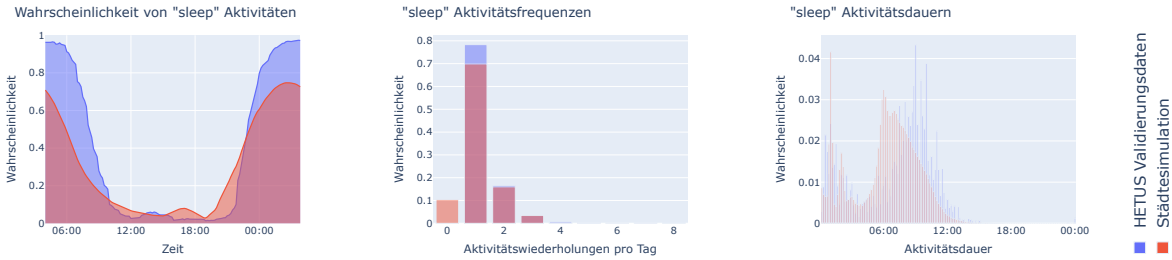
Die Menge an Schlaf ist bei Vollzeitbeschäftigten an freien Tagen zu gering. In Abbildung 4.18 werden exemplarisch die Daten für die Frauen dieser Kategorie untersucht. Dabei lassen sich mehrere Gründe für die geringe Schlafmenge ausmachen. Zum einen hat die Verteilung der Schlafdauern im längeren Bereich, also für den Nachtschlaf, einen merklich geringeren Mittelwert. Das äußert sich in der Wahrscheinlichkeitsverteilung in einem zu frühen Abfall der Schlafwahrscheinlichkeit




**(a)** Standard **(b)** Skaliert

**Abbildung 4.17:** Normaler (a) und skalierter (b) Bias für jede Aktivität in jeder Aktivitätsprofilkategorie

in den Morgenstunden. Zum anderen findet in der Städtesimulation an etwa 10% der Tage keine einzige Schlafaktivität statt, was in den Validierungsdaten so nie vorkommt. Bei den Männern dieser Kategorie bietet sich ein ähnliches Bild, allerdings weniger stark ausgeprägt. In beiden Fällen ist eine Korrektur im LoadProfileGenerator sinnvoll.




**(a)** Wahrscheinlichkeit **(b)** Häufigkeiten **(c)** Dauern

**Abbildung 4.18:** Wahrscheinlichkeitskurve (a), Häufigkeit (b) und Dauer (c) von Schlaf für vollzeitbeschäftigte Frauen an freien Tagen

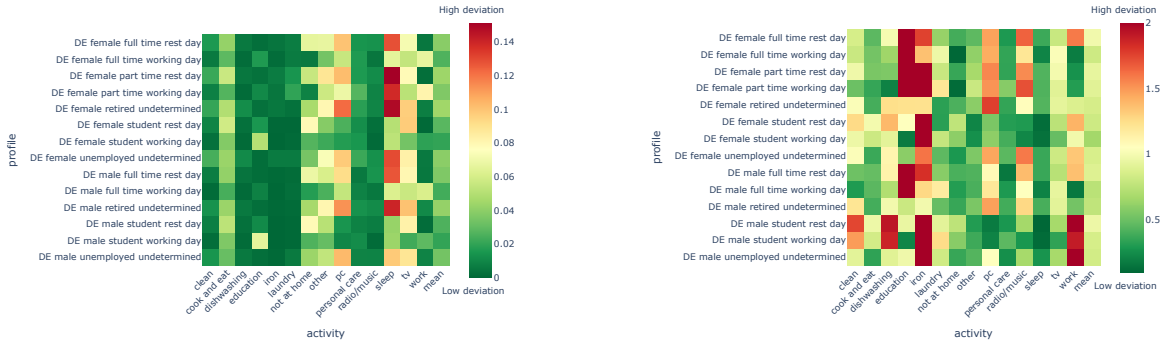
Weitere geringere Abweichungen betreffen bei Studierenden die Aktivitäten Kochen und Essen, die in ihrem Zeitanteil leicht unterschätzt werden. Die skalierten Bias-Werte zeigen außerdem eine in Relation zum Zeitaufwand große Abweichung für Bügeln sowie für Bildung in einigen Kategorien. Das hängt damit zusammen, dass Bügeln mit insgesamt 2% grundsätzlich eine sehr seltene Tätigkeit ist und im LoadProfileGenerator von vielen Personengruppen gar nicht ausgeübt wird. Bei Profiltypen mit höherem Bügelanteil, wie beispielsweise Rentnerinnen, wird Bügeln dagegen berücksichtigt und die Abweichung ist dementsprechend geringer. Daher fällt diese Ungenauigkeit nicht ins Gewicht.

### Zeitliche Verläufe nach Kategorien

Wie bereits in Abschnitt 4.3.2 für die aggregierten Daten gezeigt wurde, kann es trotz passender Zeitanteile Unstimmigkeiten in Bezug auf die Uhrzeiten geben, zu denen Aktivitäten ausgeübt werden. Dabei sind auch je nach Personen- oder Tagestyp unterschiedliche zeitliche Verläufe möglich. Daher muss dieser Aspekt gesondert für die einzelnen Profiltypen geprüft werden. Der Bias ist als Indikator hierfür nicht geeignet. Stattdessen wird in Abbildung 4.19 die Wasserstein-Distanz gezeigt. Diese nimmt nur positive Werte an, wobei $0$ für exakte Übereinstimmung und $1$ für maximal mögliche Distanz steht. Die Skalierung hebt allerdings die obere Grenze auf, sodass sich hier theoretisch der Wertebereich $[0, \infty)$ ergibt.

Durch ihre spezielle Berechnungsweise wertet die Wasserstein-Distanz leichte zeitliche Verschiebungen nur als geringfügige Abweichung. Die Indikatoren MAE und RMSE bewerten dagegen bloße Differenzen der y-Werte an jeder Stelle einer Kurve und neigen in solchen Fällen daher zu übermäßig hohen Distanzen. Ein Beispiel ist in Abbildung 4.20 abgebildet. Das Maximum der Kurve aus den Simulationsergebnissen ist hier nur leicht nach vorne verschoben und dementsprechend nahe an den Validierungsdaten.

Grundsätzlich lässt sich beobachten, dass hohe Abweichungen des Bias vom Idealwert $0$ auch zu hohen Wasserstein-Distanzen führen, die Umkehrung gilt aber nicht, da die Wasserstein-Distanz auch zeitliche Verschiebungen als Abweichung wertet. Ein deutliches Beispiel hierfür findet sich




**(a)** Standard **(b)** Skaliert

**Abbildung 4.19:** Normale (a) und skalierte (b) Wasserstein-Distanz für jede Aktivität in jeder Aktivitätsprofilkategorie

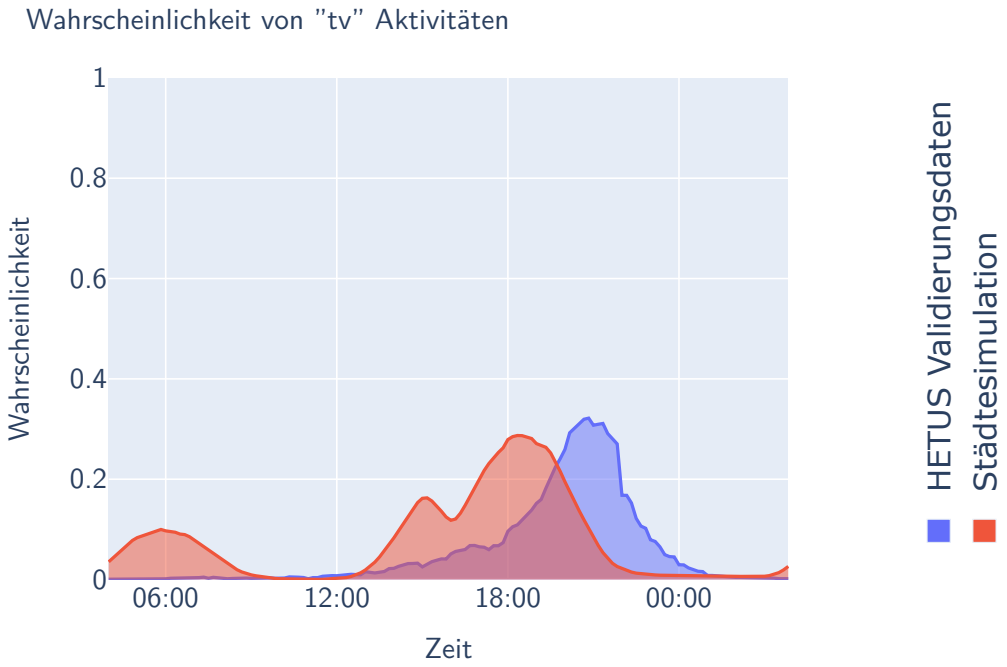
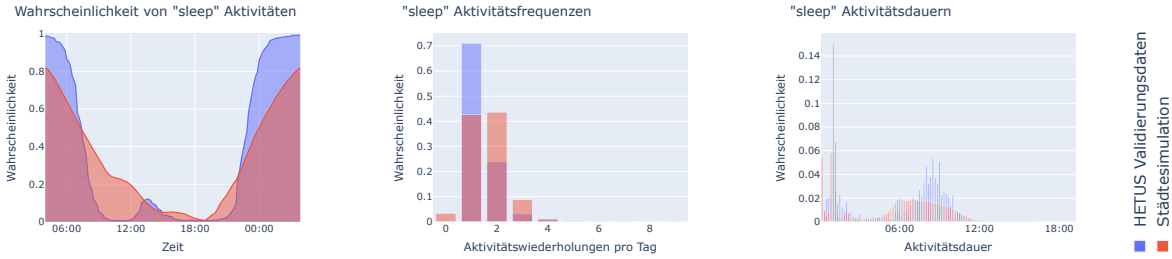



**Abbildung 4.20:** Wahrscheinlichkeitskurve für die Aktivität Fernsehen bei männlichen Studenten an Arbeitstagen

beim Schlaf. Während der Bias hier lediglich bei Vollzeitbeschäftigten an freien Tagen hohe Abweichungen anzeigt, liegt die Wasserstein-Distanz zusätzlich bei Rentnern und Teilzeitbeschäftigten verhältnismäßig hoch. In Abbildung 4.21 sind die entsprechenden Aktivitätseigenschaften für Rentner dargestellt. Aus den Wahrscheinlichkeitskurven wird die Ursache für die Diskrepanz der Indikatoren ersichtlich. Die Zeiten für den Nachtschlaf stimmen zwar ungefähr, variieren aber viel zu stark, sodass die Kurve der Städtesimulation zu flach ansteigt und abfällt. Ein Mittagsschlaf kommt zwar vor, allerdings leicht verfrüht, sodass die entsprechende Spitze in der Kurve mit dem späten Ende des Nachtschlafs verschmilzt. Ausgelöst wird dies durch zu große Varianz in sowohl Beginn als auch Dauer des Nachtschlafs. So entstehen die in Abbildung 4.19 angezeigten Abweichungen, obwohl die insgesamt mit Schlafen verbrachte Zeit gut übereinstimmt.

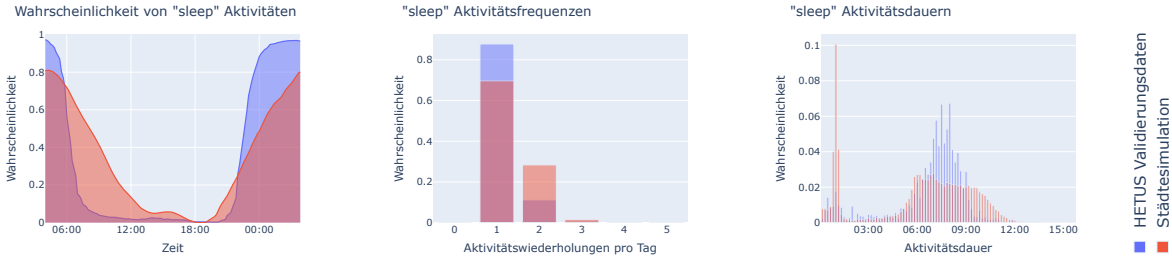



**(a)** Wahrscheinlichkeit **(b)** Häufigkeiten **(c)** Dauern

**Abbildung 4.21:** Wahrscheinlichkeitskurve (a), Häufigkeit (b) und Dauer (c) von Schlaf bei Rentnern

Bei teilzeitbeschäftigten Frauen sehen die Daten in Abbildung 4.22 ähnlich aus. Auch hier ist die Varianz in der Dauer zu groß, kombiniert mit einer Überrepräsentation des Mittagsschlafs.

Diese Beispiele belegen die Relevanz einer präzisen Validierung der zeitlichen Aktivitätsverläufe auch für einzelne Bevölkerungsgruppen. Nur auf diese Weise können derartige Unstimmigkeiten erkannt werden.

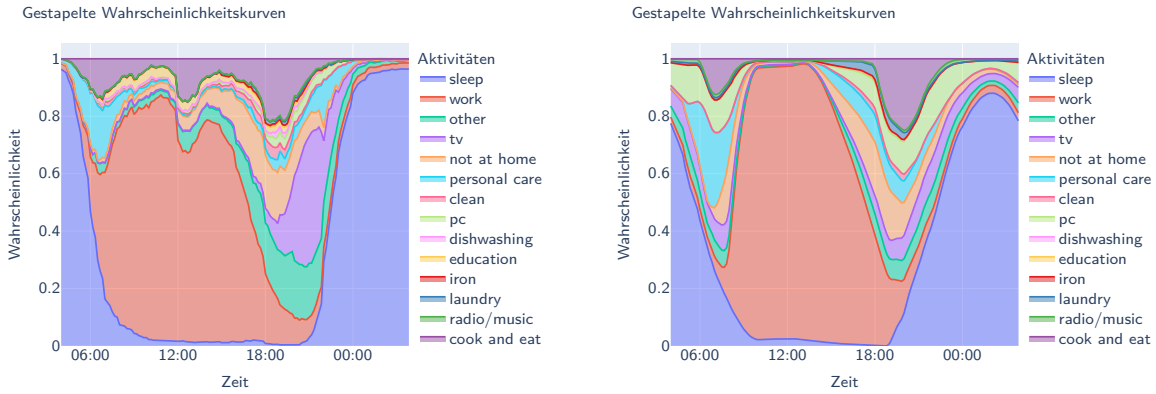



**(a)** Wahrscheinlichkeit **(b)** Häufigkeiten **(c)** Dauern

**Abbildung 4.22:** Wahrscheinlichkeitskurven (a) und Verteilungen der Dauer (b) von Schlaf bei teilzeitbeschäftigten Frauen

Für die Städtesimulation stellen die gezeigten Abweichungen aber bereits die Extremfälle dar. Die meisten Personen- und Tagestypen sind passender abgebildet und stimmen ausreichend gut mit den HETUS-Daten überein. Insbesondere Studierende und Vollzeitbeschäftigte werden sehr gut repräsentiert. Die beiden gestapelten Wahrscheinlichkeitskurven in Abbildung 4.23 für vollzeitbeschäftigte Frauen an Arbeitstagen weisen beispielsweise allgemein eine große Ähnlichkeit auf. Abgesehen von einigen spezifischen Ungenauigkeiten wie dem nicht modellierten Mittagessen und den zu stark verteilten Fernsehaktivitäten werden die Verhaltensmuster aus den HETUS-Daten passend wiedergegeben.

Insgesamt zeigt die Validierung mit ETHOS.ActivityAssure somit, dass im LoadProfileGenerator noch kleinere Unstimmigkeiten und Ungenauigkeiten enthalten sind, diese betreffen aber größtenteils Details und fallen in der Gesamtbetrachtung der Stadt nicht ins Gewicht. Die Städtesimulation ist in der Lage, Verhaltensmuster auf aggregierter Ebene sowie speziell für verschiedene Bevölkerungsgruppen angemessen zu berücksichtigen und plausible Aktivitätsprofile zu generieren.




**(a)** HETUS **(b)** Städtesimulation

**Abbildung 4.23:** Gestapelte Wahrscheinlichkeitskurven für die HETUS-Daten (a) und die Städtesimulation (b) für vollzeitbeschäftigte Frauen an Arbeitstagen

## 4.4 Auswertung des Transports

Transportmodellierung ist nicht der Hauptschwerpunkt der Städtesimulation. Stattdessen wird dieser Bereich im Detail von einem entsprechenden Modell abgebildet, dass zu diesem Zweck mit der Städtesimulation gekoppelt wurde, wie in Abschnitt 3.3.5 beschrieben. Aus diesem Grund wird das Reiseverhalten nicht vollumfänglich validiert. Da Wegzeiten aber auch einen Einfluss auf den Tagesablauf und ausgeübte Aktivitäten haben, wird an dieser Stelle ein kurzer Überblick gegeben.

Abbildung 4.24 zeigt die Häufigkeit von Reisen zu verschiedenen Zielorten hin oder von diesen Orten nach Hause. Das zu erkennende Muster entspricht dabei den Erwartungen, das häufigste Ziel macht der Arbeitsplatz aus, gefolgt von Supermärkten und weiteren Orten zur Erledigung von Einkäufen, sowie Schulen. Besuche bei Freunden und Verwandten kommen seltener, aber noch verhältnismäßig häufig vor. Universitäten und Arbeitsplätze mit Schichtdienst werden aufgrund des kleineren Personenkreises, für den diese Orte relevant sind, weniger häufig aufgesucht. Zuletzt folgen verschiedene Orte für Freizeitaktivitäten. Diese Häufigkeitsverteilung ist plausibel und passt zu den validierten Aktivitätsprofilen.

Neben der Häufigkeit der Reisen muss auch deren Distanz geprüft werden. In Abbildung 4.25 ist daher die Verteilung der Reisen auf unterschiedliche Distanzen dargestellt. Der Wertebereich passt dabei gut zur Stadt Jülich, mit Entfernungen von größtenteils weniger als 15 km und einem Schwerpunkt im Bereich sehr kurzer Strecken unterhalb von 5 km. Da Jülich insbesondere in der Stadtmitte viele verschiedene Points of Interest enthält, sind für die meisten Außer-Haus-Aktivitäten keine langen Reisen erforderlich.

Abbildung 4.26 liefert hierzu eine detailliertere Ansicht, indem die Verteilung der Distanzen separat für jede Point-of-Interest-Kategorie gezeigt wird. Viele typische Orte, die häufig in der Innenstadt zu finden sind, wie etwa Supermärkte, Apotheken oder Arztpraxen, haben dabei eine ähnliche Verteilung, mit vielen kurzen Strecken und wenigen längeren. Die Entfernung ist für die Auswahl von Geschäften und anderen Einrichtungen entscheidend, daher ist dieses Muster plausibel. Bei Reisen zu Häusern von Freunden und Verwandten lässt sich eine deutlich größere Streuung beobachten. Das liegt daran, dass diese Zielorte nicht in Abhängigkeit von der Distanz gewählt

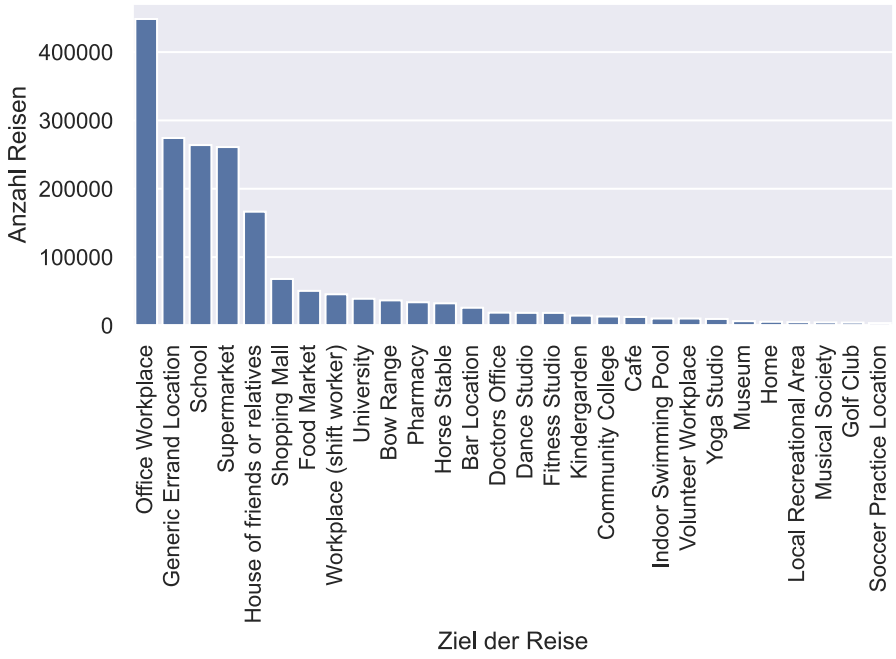



**Abbildung 4.24:** Verteilung der Anzahl an Reisen zu verschiedenen Zielorten in der Städtesimulation

werden, da für den Aufbau sozialer Beziehungen in der Regel gänzlich andere Faktoren eine Rolle spielen. Zudem stechen einige Orte mit auffälliger Verteilung hervor, in diesem Fall vor allem „Horse Stable“ und „Soccer Practice Location“. Bei beiden handelt es sich um eher seltene Points of Interest, die nur wenige Male innerhalb einer Stadt vorkommen. Dadurch kommt es hierbei zu einer stärker konzentrierten Verteilung der Distanzen. Die Verteilungen der Reisedistanzen passen somit zu den jeweiligen Zielorten und bilden die stärkeren Schwankungen und durchschnittlich höheren Strecken bei Reisen zu weniger häufigen Points of Interest realistisch ab.

Ein auffälliger Aspekt bei allen Distanzen ist die geringe Obergrenze von etwa 30 km. Diese kommt dadurch zustande, dass bei der Szenariogenerierung lediglich alle Gebäude ausgewählt werden, die laut ihrer offiziellen Adresse zur betrachteten Stadt gehören. Das Umland der Stadt wird in der aktuellen Implementierung ignoriert. Somit werden keine Reisen über die Stadtgrenzen hinaus abgebildet. Diese Einschränkung betrifft primär das Pendeln zum Arbeitsplatz, da dieser in vielen Fällen nicht im Wohnort der Person liegt. Der Rest der Simulation ist davon nur in geringem Maße durch im Mittel leicht kürzere Reisedauern betroffen, aber im Hinblick auf die simulierten Wege ist diese Einschränkung von Bedeutung.

Eine Möglichkeit, diese Limitierung zu umgehen, besteht darin, einen Puffer-Radius zu definieren und auch Nichtwohngebäude innerhalb dieser vorgegebenen Entfernung zur Stadt miteinzubeziehen. Da der Fokus dieser Arbeit jedoch nicht auf der genauen Modellierung von Wegstrecken liegt, wird diese Option für zukünftige Forschung offengelassen.

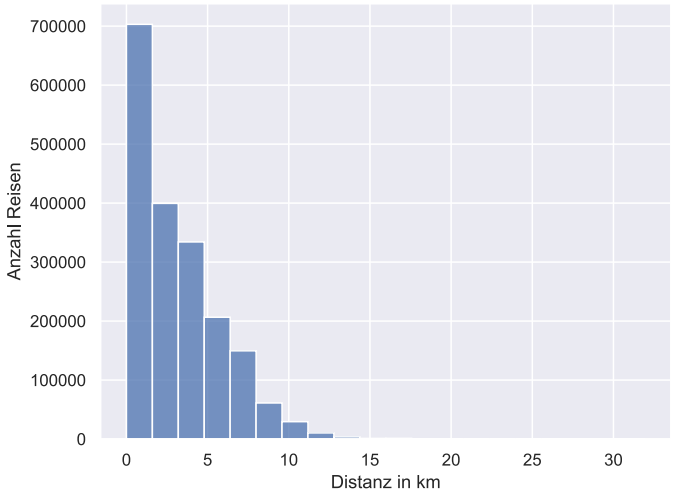



**Abbildung 4.25:** Häufigkeitsverteilung der Reisedistanzen nach Zielort-Kategorie

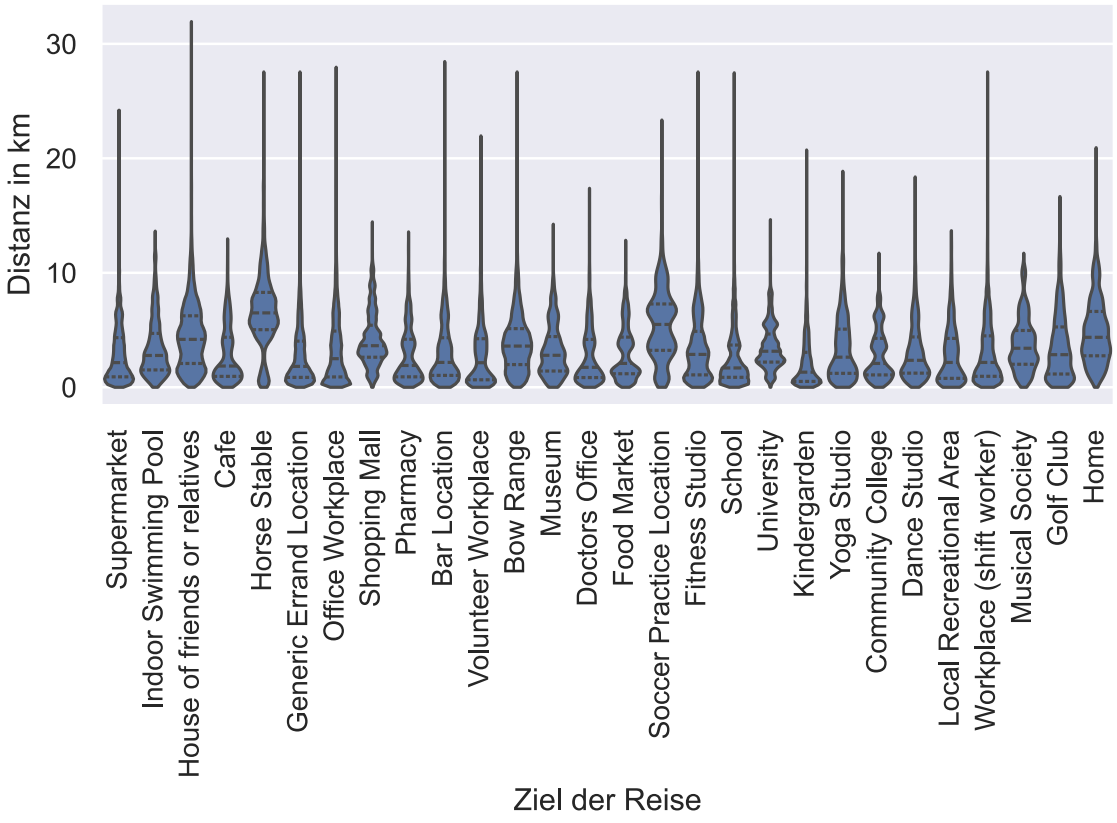



**Abbildung 4.26:** Häufigkeitsverteilung der Reisedistanzen für alle Bewohner von Jülich

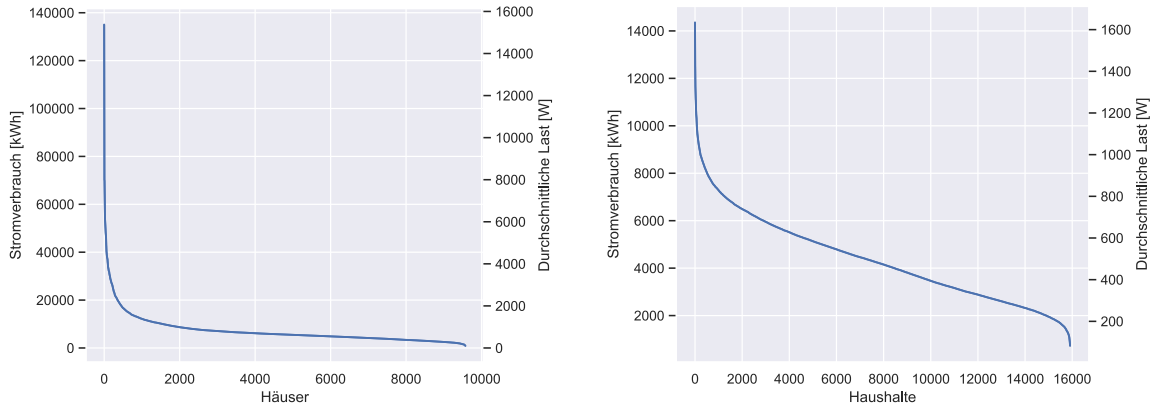
## 4.5 Validierung und Auswertung der Lastprofile

Ein wesentlicher Bestandteil der Städtesimulation ist die Abbildung von Energie- und Ressourcenverbräuchen in Haushalten. Der LoadProfileGenerator, auf dem die Städtesimulation basiert, erzeugt für jeden simulierten Haushalt Lastprofile aller relevanten Typen, also unter anderem Strom, Wasser und Gas. Im Folgenden wird der Fokus auf die Stromlastprofile gelegt.

Bei der Simulation der Stadt Jülich wurden insgesamt 15.910 Haushaltslastprofile allein für den Lasttyp Strom generiert. Diese Menge an Profilen kann nicht sinnvoll individuell überprüft werden. Einzelne Lastprofile auf Haus- und Haushaltsebene wurden zudem bereits von Pflugradt et al. [10] validiert. In dieser Arbeit liegt der Schwerpunkt dagegen auf der Simulation einer ganzen Stadt, daher werden die Profile mit verschiedenen Methoden auf aggregierter Ebene untersucht und ausgewertet.

### 4.5.1 Stromverbrauch der Stadt

Um die erzeugten Lastprofile zu prüfen, wird zunächst der gesamte Jahresverbrauch aller Häuser betrachtet. Abbildung 4.27a zeigt den Verbrauch der Häuser absteigend sortiert. Dabei ergibt sich eine typische Verteilung mit vielen Lasten in einem passenden Bereich für kleinere Haushalte. Der mittlere Jahresverbrauch liegt bei 7.220 kWh, und 83% aller Häuser haben einen Verbrauch unter 10.000 kWh. Lasten in diesem Bereich kommen deshalb so häufig vor, weil einer der größten Faktoren für den Energieverbrauch in Wohngebäuden die Anzahl der dort lebenden Personen ist. 6.613 der 9.572 Häuser in Jülich sind Einfamilienhäuser, und in 41% aller Haushalte lebt nur eine Person. Dadurch wohnen in 84% der Häuser insgesamt nur vier Personen oder weniger, sodass das für diese Größe typische Verbrauchsniveau stark repräsentiert ist.




**(a)** Verbrauch pro Haus

**(b)** Verbrauch pro Haushalt

**Abbildung 4.27:** Gesamtverbrauch und entsprechende Durchschnittslast jedes einzelnen Hauses (a) und Haushalts (b) im Simulationszeitraum, absteigend sortiert

Mit der Personenzahl hängen auch die Ausreißer in der Kurve zusammen. In dem Haus, das mit über 135.000 kWh den höchsten Verbrauch in Jülich hat, leben 85 Personen, und auch alle übrigen Häuser mit einem Verbrauch von 40.000 kWh oder mehr haben mindesten 10 Einwohner. Von derart großen Mehrfamilienhäusern gibt es allerdings nur wenige in der Stadt.

Analog hierzu zeigt Abbildung 4.27b den Verbrauch jedes Haushalts in Jülich. Die Form der Kurve

ist sehr ähnlich zu den Verbräuchen der Häuser, allerdings mit weniger extremen Maximalwerten. Das hängt damit zusammen, dass die maximale Haushaltsgröße, die im LoadProfileGenerator implementiert wurde, bei sechs Personen liegt. Die Spanne zwischen minimaler und maximaler Personenzahl ist daher bei Haushalten deutlich geringer als bei Häusern. Ein Großteil der Haushalte weist einen Jahresverbrauch im üblichen Bereich zwischen 2.000 kWh und 8.000 kWh auf, die Kurve ist somit plausibel.

Im Vergleich mit durchschnittlichen Haushaltsverbräuchen in Deutschland [67] fällt jedoch auf, dass die Simulation im Mittel um 20–30% höhere Verbräuche generiert. Um diese Abweichung genauer zu analysieren, wird in Abbildung 4.28 der durchschnittliche Jahresverbrauch aller 63 Haushaltstypen gezeigt, die in Jülich vorkommen. Anhand der Farbe ist zu erkennen, dass kleinere Haushalte erwartungsgemäß meist einen geringeren Energieverbrauch haben. Einige kleine Haushalte, unter anderem CHR51 und CHR54, fallen allerdings durch verhältnismäßig hohe Verbräuche auf. Diese können in Zukunft gesondert untersucht werden, um die Ursachen der hohen Lasten zu finden und, falls notwendig, zu beheben. Das allgemein erhöhte Niveau der Verbräuche im Vergleich zu den Validierungsdaten lässt sich durch veraltete und daher ineffiziente Haushaltsgeräte sowie übermäßig viele energieaufwendige Freizeitaktivitäten, wie die in Abschnitt 4.3 identifizierte hohe PC-Nutzung, erklären. Die Konfiguration modernerer Geräte und sparsamerer Aktivitäten kann in Zukunft dazu beitragen, den zu hohen Verbrauch abzusenken.

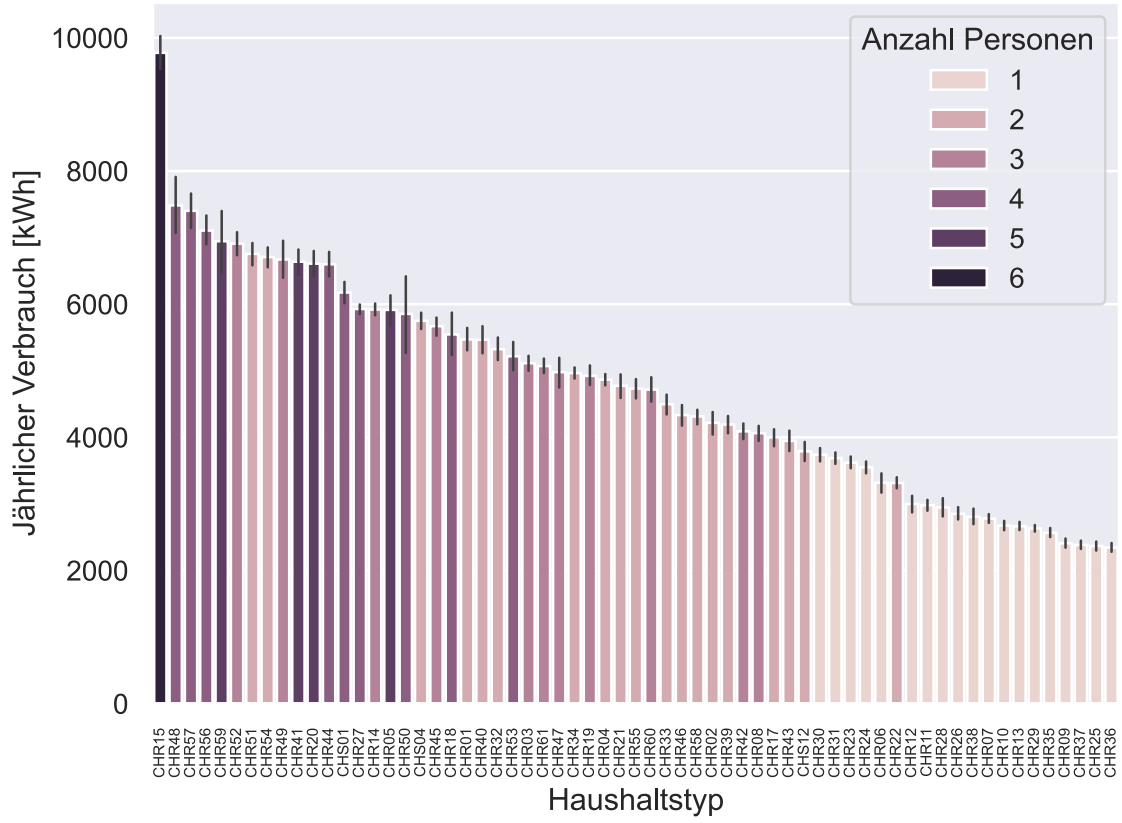



**Abbildung 4.28:** Durchschnittliche Jahresverbräuche nach Haushaltstyp. Die Farbe zeigt die Anzahl Personen im entsprechenden Haushaltstyp an.

Neben der Verteilung des Energieverbrauchs auf die einzelnen Häuser und Haushalte ist auch ein

realistischer zeitlicher Verlauf entscheidend. Um diesen betrachten zu können, zeigt das Rasterdiagramm in Abbildung 4.29 die aufsummierte Last aller Häuser der Stadt in jedem Simulationszeitschritt. In diesem Diagramm wird der Lastverlauf an einem Tag durch eine Spalte dargestellt, und

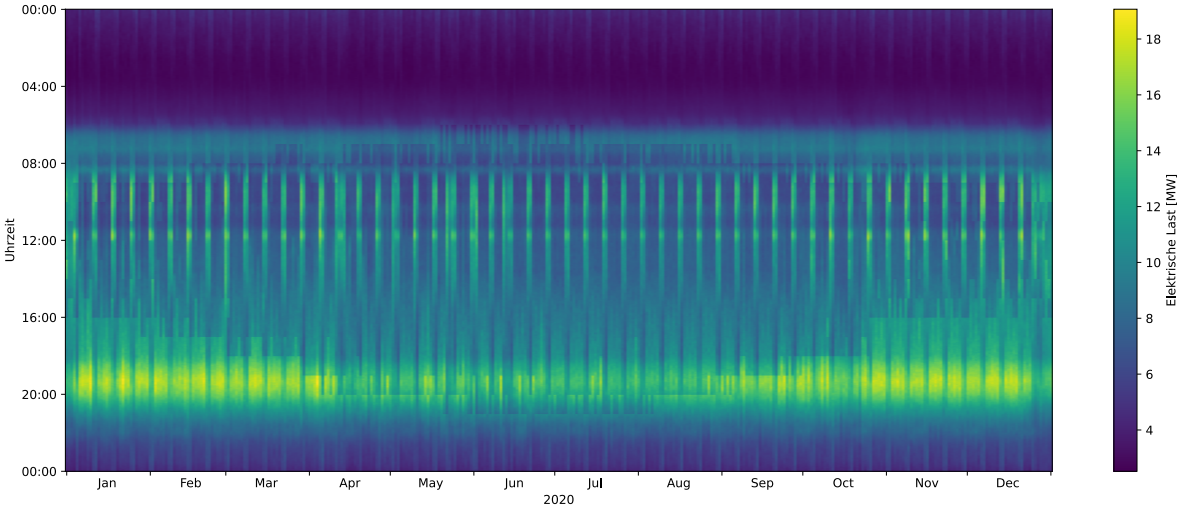



**Abbildung 4.29:** Generiertes elektrisches Lastprofil für Jülich für ein gesamtes Jahr

alle Tage bilden nebeneinander die Last im gesamten Jahr. Gut zu erkennen sind darin die Nachtruhe zwischen etwa 22 Uhr und 6 Uhr sowie die erhöhte Last nach dem üblichen Arbeitsende ab etwa 18 Uhr. Das Muster gleichmäßiger vertikaler Streifen entsteht durch Wochenenden. Da die meisten Personen wochentags einen großen Teil des Tages nicht zuhause verbringen, ist die Last in dieser Zeit geringer. Außerdem lässt sich ein ovales Muster erkennen, dass in den Wintermonaten zwischen 8 und 16 Uhr zu leicht verringerter Last führt und sich ausdehnt, bis es sich im Sommer von 6 bis 21 Uhr erstreckt. Dieses Muster wird in erster Linie durch die wechselnde Menge an Tageslicht hervorgerufen, die festlegt, wann Beleuchtung eingeschaltet wird. Das Summenprofil der Stadt weist somit alle regelmäßigen zeitlichen Schwankungen und Verläufe eines typischen Lastprofils auf.

Im Folgenden soll die Form der generierten Lastprofile genauer geprüft werden. Als Referenz werden hierfür die Standardlastprofile des BDEW [68] verwendet. Diese Sammlung an Standardlastprofilen wurde auf Basis von Messdaten erstellt und gibt typische Lastprofilverläufe für verschiedene Anwendungszwecke wieder, beispielsweise für Gewerbe oder Landwirtschaftsbetriebe. Das Haushaltsprofil H25 repräsentiert dabei den Strombedarf von Haushalten und enthält den Lastverlauf in 15-Minuten-Auflösung für einen einzelnen Tag. Dabei wird zwischen drei Tagestypen und den 12 Monaten unterschieden, sodass insgesamt 36 verschiedene Profile bereitgestellt werden. Mithilfe einer Dynamisierungsfunktion muss das jeweilige Profil anschließend für den gewünschten Zieltag angepasst werden. Alle Standardlastprofile sind grundsätzlich normiert und müssen vor der Verwendung auf den benötigten Gesamtenergieverbrauch skaliert werden. Auf diese Weise kann für jeden Tag eines Jahres ein passendes Standardprofil erzeugt werden, das als Vergleich für die generierten Daten genutzt werden kann.

Um die Verteilung der Lastwerte im erzeugten Städtelastprofil zu validieren, werden zunächst Jahresdauerlinien verwendet. Diese entstehen durch Sortierung aus den ursprünglichen Lastprofilen und geben einen Überblick darüber, wie lange welches Lastniveau anliegt. In Abbildung 4.30

wird die Jahresdauerlinie des Gesamtlastprofils von Jülich der aus den H25-Profilen errechneten Jahresdauerlinie gegenübergestellt. Wie dort zu sehen ist, passt die generierte Jahresdauerlinie der gesamten Stadt gut zur Form des H25-Standardlastprofils. Dabei fällt auf, dass in den Simulationsergebnissen häufiger besonders hohe oder niedrige Lastniveaus auftreten und dabei extremere Werte erreichen als das Standardlastprofil. Das liegt daran, dass das Standardlastprofil im Wesentlichen einem gemittelten Lastprofil entspricht, angepasst an den gewünschten Verbrauch. Das Profil hat dabei stets die gleiche Form und deckt übliche Lastverläufe auf aggregierter Ebene gut ab. Seltenere Fälle, in denen viele Häuser gleichzeitig eine außergewöhnlich hohe oder niedrige Last aufweisen, können so jedoch nicht abgebildet werden. Mit der Städtesimulation ist das möglich, da jedes Haus einzeln betrachtet wird und sich somit automatisch entsprechende Situationen per Zufall ergeben. Auf diese Weise kommen die zu erkennenden Unterschiede zustande.

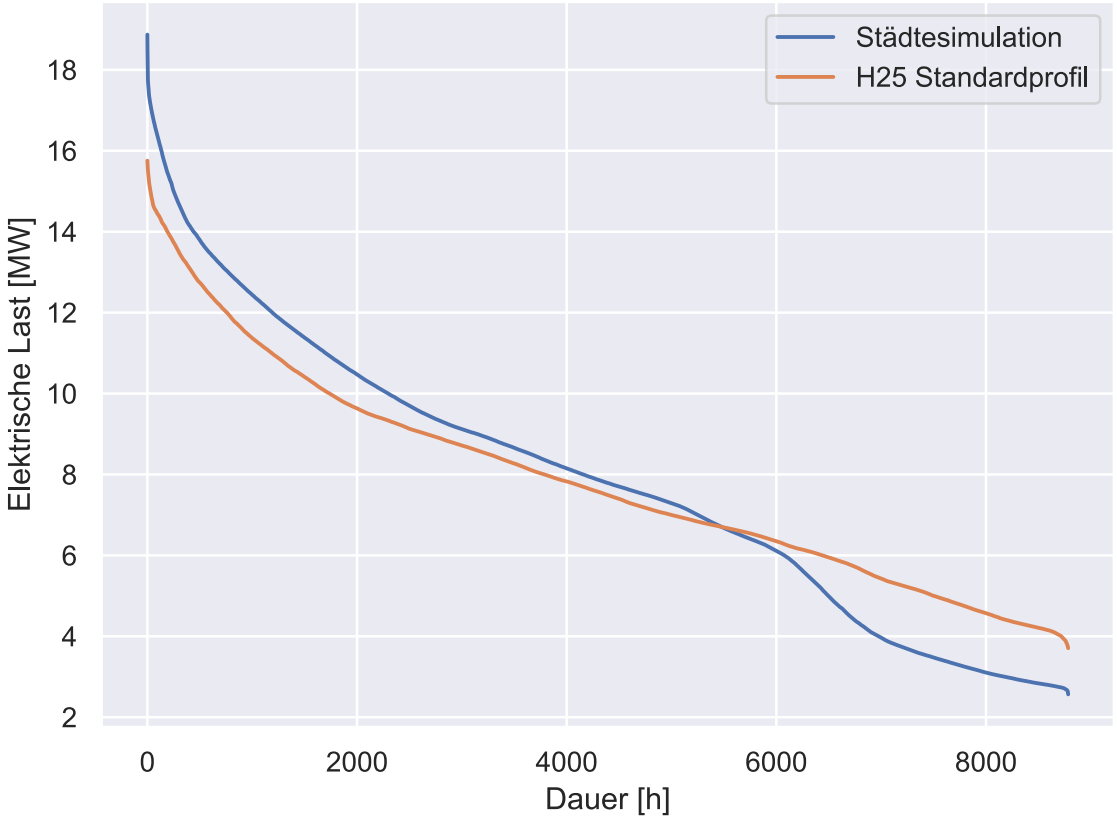



**Abbildung 4.30:** Die Last der gesamten Stadt als Jahresdauerlinie, im Vergleich zu einem entsprechend skalierten H25-Standardlastprofil. Jede Kurve gibt an, für wie lange welches Lastniveau insgesamt bestanden hat.

Zum Abgleich des täglichen Lastverlaufs werden nun aus allen Haushaltslastprofilen Statistik-Profile gebildet. Dabei wird jedes generierte Haushaltslastprofil in 366 24h-Profile aufgeteilt. Aus den 24h-Profilen aller Haushalte werden anschließend für jeden Zeitschritt Mittelwert, Quartile sowie Minimum und Maximum berechnet. Die so gesammelten Statistik-Werte sind in Abbildung 4.31 jeweils als eigenes Profil dargestellt, wieder im Vergleich zum H25-Standardlastprofil. Ähnlich wie im Rasterdiagramm in Abbildung 4.29 sind auch hier die Nachtruhe und die Feierabendzeit klar erkennbar. Das Minimum liegt fast konstant bei etwa 4 W. Diese Last kommt

durch den dauerhaften Stromverbrauch im Standby-Modus einiger Geräte, beispielsweise von Fernsehern, zustande und liegt damit durchgehend an. Beim Maximum sind dagegen Schwankungen zu beobachten, da hohe Werte hier entstehen, wenn zufällig mehrere größere Lasten zeitgleich anfallen. Die Wahrscheinlichkeit dafür ist nachmittags am größten, sodass der mit über 25 kW größte Maximalwert um 13:40 Uhr auftritt. Durch Personen in Schichtarbeit oder andere mit besonders früher oder später Schlafenszeit kann es aber auch nachts zu hohen Lastwerten kommen. Der kleinste der Maximalwerte liegt daher bei etwa 13 kW um 3:40 Uhr.

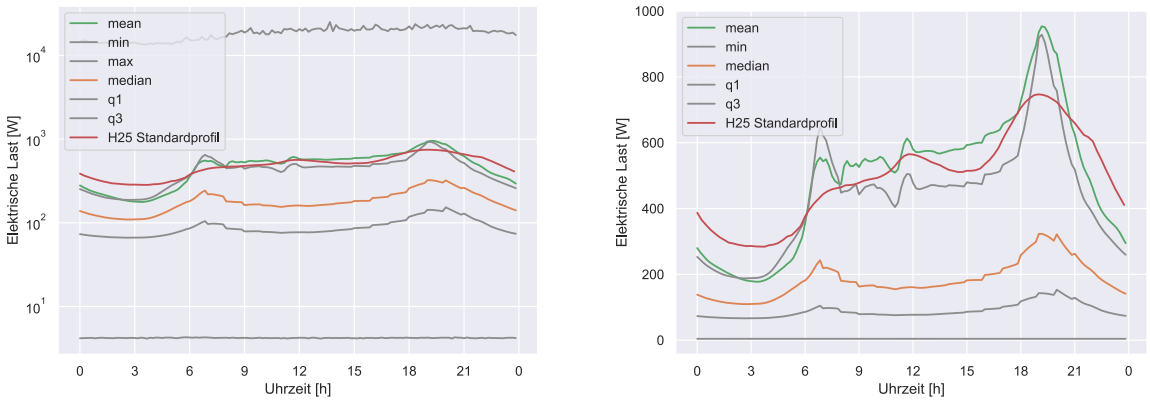



**Abbildung 4.31:** Statistik-Profile aller simulierten Häuser mit (a) und ohne (b) Maximum zur besseren Lesbarkeit. Zusätzlich wird jeweils das H25-Standardlastprofil dargestellt.

Die starke Abweichung zwischen Mittelwert und Median zeigt, dass der häufigste Lastbereich der meisten Haushalte deutlich niedriger liegt und der Durchschnitt durch einseitige Ausreißer nach oben verschoben wird. Diese Beobachtung ist konsistent mit der asymmetrischen Jahresdauerlinie der Stadt. Da das H25-Standardlastprofil konzipiert wurde, um aggregierte Lastgänge in Gruppen von Haushalten abzubilden, liegt es im Bereich des Mittelwertprofils. Es passt sowohl vom Lastniveau als auch vom zeitlichen Verlauf her gut, die Haushaltslasten werden also im Mittel gut modelliert.

### 4.5.2 Gleichzeitigkeit der Lastprofile

Anhand der gezeigten Minimal- und Maximalprofile wird die Bedeutung von Gleichzeitigkeit deutlich. Dieser Begriff bezieht sich darauf, ob bestimmte Muster in den Lastkurven einzelner Häuser oder Haushalte zeitlich synchron oder versetzt ablaufen. Bei einem synchronen Ablauf akkumulieren sich hohe oder niedrige Lasten zu Extremwerten auf Quartiers- oder Städteebene. Gibt es jedoch einen zeitlichen Versatz, gleichen sich die Lasten zu Durchschnittswerten aus. Aus diesem Grund ist die Gleichzeitigkeit von Lastprofilen ein wichtiger Aspekt im Zusammenhang mit lokalen Energiesystemen. Durch sie wird bestimmt, wie extrem Lastspitzen in bestimmten Gebieten ausfallen können und wie die dazugehörige Technik, etwa Transformatoren, ausgelegt werden muss.

Tabelle 4.2 veranschaulicht die Auswirkungen unterschiedlicher Gleichzeitigkeit anhand der Funk-

tion

$$F = \sum_{i=1}^{N} Y_i, \tag{4.1}$$

wobei $Y_i$ Lastwerte sind, die gleichverteilt aus der Menge $X_i$ der Werte im Lastprofil von Haushalt $i$ gezogen werden. Somit entspricht $F$ der Summe aus $N$ Werten, einem aus jedem Haushaltslastprofil. Damit gibt $F$ den Wertebereich an, den das Stadtlastprofil theoretisch annehmen könnte, wenn alle Haushaltslastprofile die Werte in passender Reihenfolge enthielten.

**Tabelle 4.2:** Statistiken des Stadtlastprofils im Vergleich zur Funktion $F$ aus Gleichung (4.1). Alle Werte sind in Megawatt angegeben. Dabei ist $E(X)$ der Mittelwert und $\sigma(X)$ die Standardabweichung von $X$.

| Statistik-Maß | Berechnung der Statistik für $F$ | $F$ | Stadtlastprofil |
|---|---|---|---|
| Mittelwert | $\sum_{i=1}^{N} E(X_i)$ | 7,86 | 7,86 |
| Minimum | $\sum_{i=1}^{N} \min X_i$ | 1,02 | 2,57 |
| Maximum | $\sum_{i=1}^{N} \max X_i$ | 215,65 | 18,87 |
| Standardabweichung | $\sqrt{\sum_{i=1}^{N} \sigma(X_i)^2}$ | 0,14 | 3,49 |

Mittelwert, Minimum, Maximum und Standardabweichung von $F$ wurden anhand der Formeln in Tabelle 4.2 aus allen $X_i$ berechnet. Die Mittelwerte stimmen überein, da der Mittelwert der Funktion $F$ unter der gegebenen Randbedingung, dass alle $X_i$ gleich viele Werte enthalten, dem Mittelwert des Stadtprofils entspricht. Bei den übrigen Kennzahlen sind aber deutliche Unterschiede zu beobachten. Minimum und Maximum des Stadtlastprofils sind nicht so extrem wie die entsprechenden Werte von $F$. Insbesondere das Maximum ist um ein Vielfaches kleiner, da, um den Höchstwert von $F$ anzunehmen, jedes Haushaltslastprofil sein Maximum zur gleichen Zeit erreichen müsste. Bei über 15.000 Haushalten ist das kein realistisches Szenario. Die Standardabweichung fällt dagegen beträchtlich kleiner aus. Das hängt damit zusammen, dass sich bei rein zufälligem Ziehen einer derart großen Anzahl an Werten Ausreißer im Schnitt ausgleichen, sodass der Mittelwert einer Stichprobe meist nahe dem Erwartungswert liegt. Die Standardabweichung des Stadtlastprofils ist höher, da die Lastprofile nicht unabhängig sind, sondern durch Faktoren wie Tageszeit, Temperatur oder Jahreszeit beeinflusst werden.

Um die Gleichzeitigkeit der Haushaltslastprofile zu untersuchen, kann der Gleichzeitigkeitsfaktor genutzt werden. Dieser kann für zwei oder mehr Lastprofile berechnet werden und gibt im Wertebereich $(0, 1]$ an, wie synchron die Profile in Bezug auf ihre Maximalwerte sind. Wenn alle Maximalwerte zum selben Zeitpunkt auftreten, hat der Gleichzeitigkeitsfaktor den Wert 1, da das summierte Profil in diesem Fall den größtmöglichen Maximalwert annimmt. Abbildung 4.32 zeigt daraus erstellte Gleichzeitigkeitskurven. Dazu wird, angefangen mit einem einzelnen Haushaltslastprofil, sukzessive ein weiteres Profil der Stadt hinzugenommen und der Gleichzeitigkeitsfaktor ermittelt. Da die so erstellte Kurve von der Reihenfolge der Profile abhängt, sind drei Kurven mit jeweils unterschiedlicher, zufälliger Reihenfolge der Profile abgebildet. Die Kurven folgen dabei einem hierfür üblichen Verlauf. Die Werte fallen schnell ab und stabilisieren sich schließlich bei einem Gleichzeitigkeitsfaktor von etwa 0.08. Nach etwa 1.000 Haushalten finden kaum noch Veränderungen statt, ab diesem Punkt haben weitere Haushalte nur noch geringfügigen Einfluss auf

den Maximalwert des aufsummierten Profils.

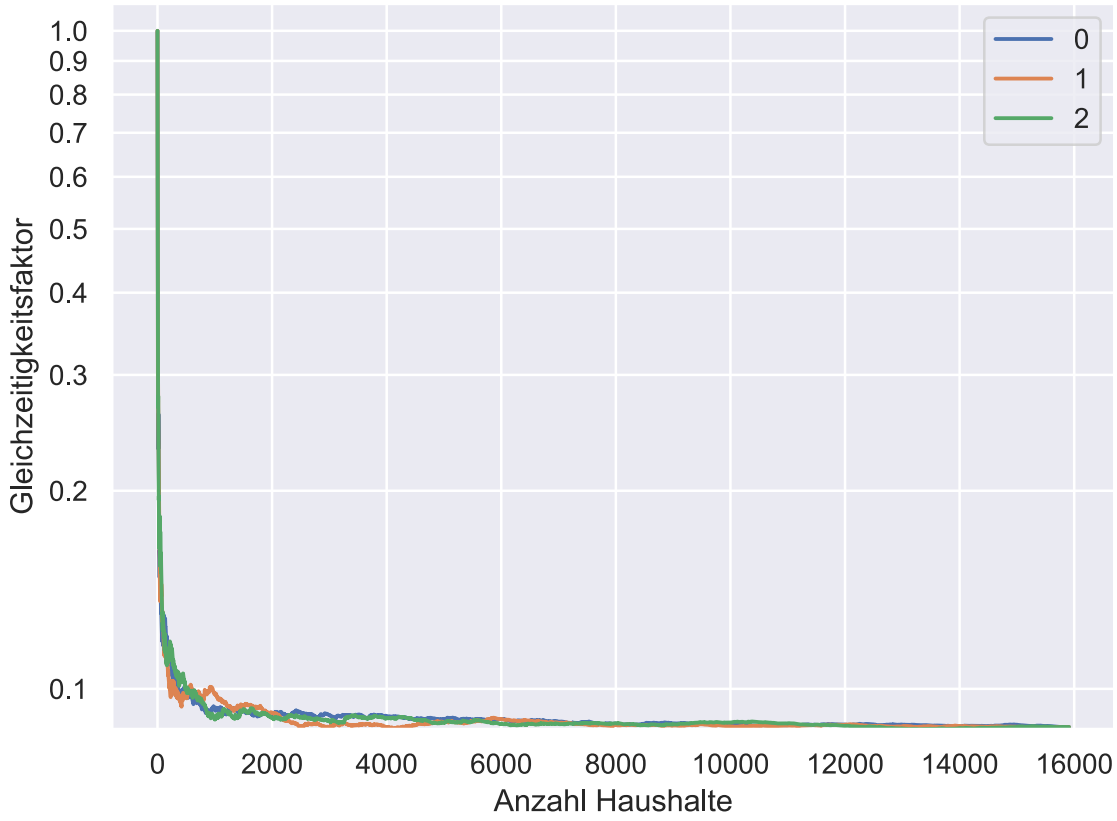



**Abbildung 4.32:** Drei Gleichzeitigkeitskurven mit zufällig gewählter Reihenfolge der Haushalte

### 4.5.3 Geografische Auswertung der Last

Eine wesentliche Besonderheit der Städtesimulation ist ihre Abbildung einer ganzen Stadt basierend auf geografischen Daten. Jedes Haus wird anhand seiner Merkmale modelliert und simuliert, im korrekten lokalen Zusammenhang zu anderen Wohn- und Nichtwohngebäuden. Während der Zweck eines üblichen Lastprofilgenerators in der Synthese zeitlich aufgelöster Lastdaten besteht, kann die Städtesimulation zeitlich und räumlich aufgelöste Daten generieren. Diese Daten können helfen, die Lastverteilung innerhalb einer Stadt im Detail zu untersuchen und komplexere Fragestellungen zu beantworten.

In Abbildung 4.33 ist der Gesamtverbrauch eines Jahres für jedes Haus in Jülich abgebildet. Hier ist insbesondere der Stadtkern gut zu erkennen, der mit größeren Mehrfamilienhäusern eine überdurchschnittliche Last aufweist. Zum Abgleich zeigt Abbildung 4.34a die Anzahl an Personen, die in jedem Gebäude wohnen. Das grundsätzliche Muster weist in beiden Grafiken starke Ähnlichkeit auf.

Bei Betrachtung der durchschnittlichen Last pro Person in jedem Haus in Abbildung 4.34b wird allerdings auch deutlich, dass die Personenzahl nicht das einzige Kriterium für die Höhe des Energieverbrauchs ist. Stattdessen gibt es auch bei den relativen Werten eine breite Spanne an Lasten. Dabei lässt sich kein klares Muster erkennen, energieintensive und sparsame Häuser sind

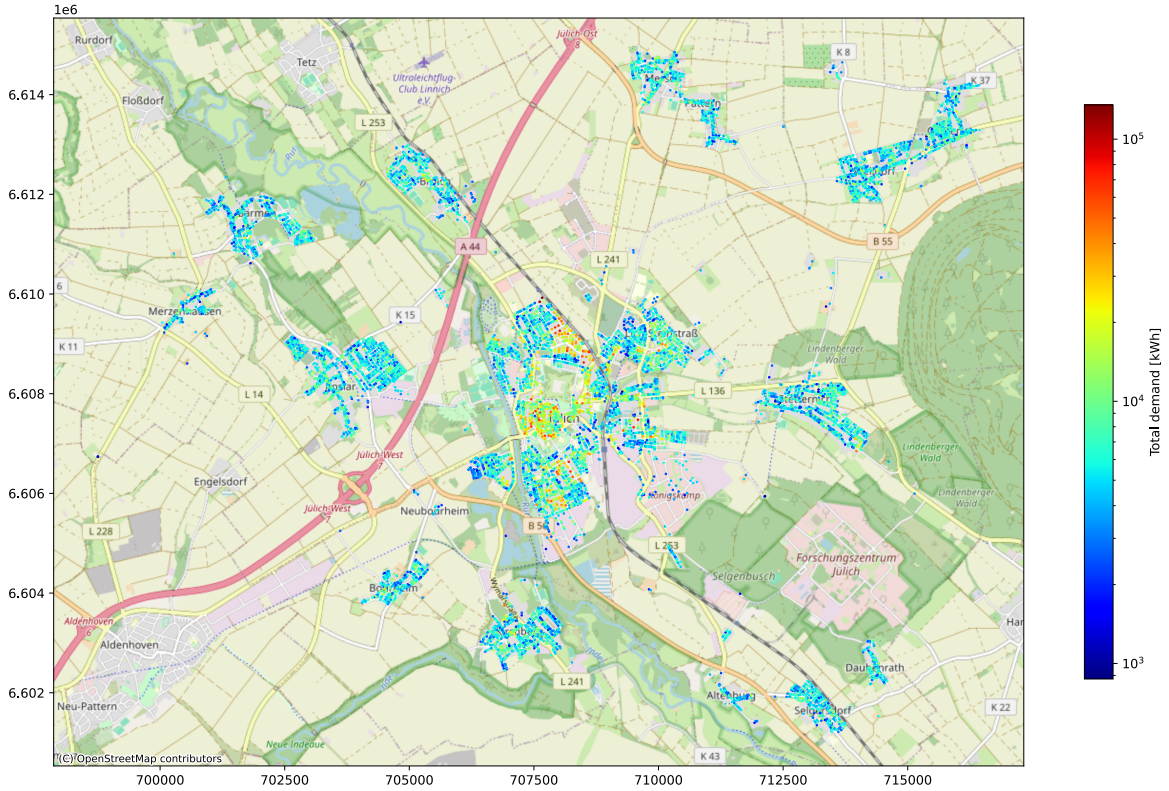
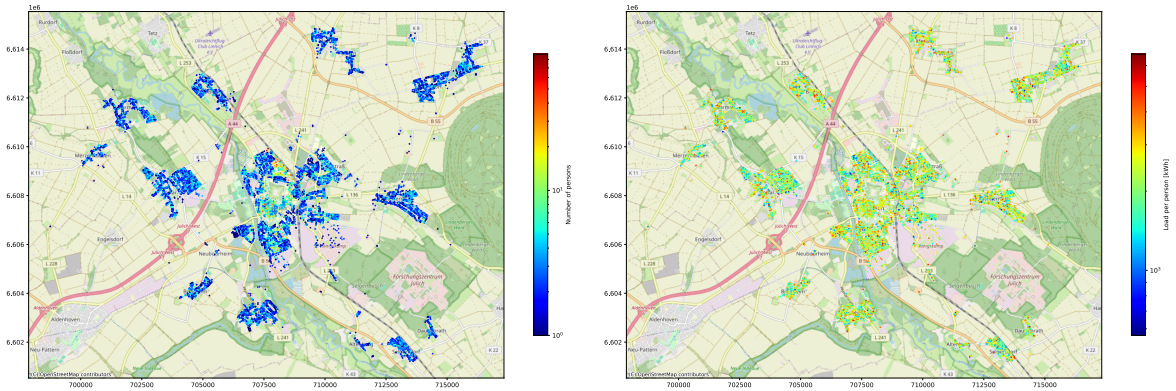



**Abbildung 4.33:** Simulierter Stromverbrauch aller Häuser in Jülich in einem Jahr

gleichermaßen über die gesamte Stadt verteilt.

Die Standardabweichung der Hauslastprofile in Abbildung 4.35 gibt Auskunft über typische Lastschwankungen. Auch hier sticht vor allem der Stadtkern heraus, allerdings weniger stark. Bei Häusern mit mehreren Haushalten ist der Wertebereich der Last größer, aber der Ausgleich durch geringe Gleichzeitigkeit der Haushaltsprofile sorgt dafür, dass die Standardabweichung nur unterproportional zur Anzahl Haushalte anwächst.

Da die Städtesimulation räumlich und zeitlich sehr präzise arbeitet und vollständige Lastdaten ohne vorherige Aggregation speichert, können beliebige Zeitpunkte aus dem Simulationsverlauf genauer untersucht werden. Abbildung 4.36 zeigt als ein mögliches Beispiel den Moment, in dem die Gesamtlast der Stadt ihr Maximum erreicht. So kann analysiert werden, welche Häuser besonders zu dieser hohen Gesamtlast beigetragen haben. Dabei fällt besonders auf, dass keine einzelnen Häuser herausstechen, sondern dass ein sehr gleichmäßiges Lastniveau vorliegt. Häuser mit niedriger Last bilden als vereinzelte blaue Punkte die Ausnahme. Im inneren Bereich der Stadt ist die Last durchgehend sehr hoch. Diese Beobachtung ist im Hinblick auf Datum und Uhrzeit des abgebildeten Zeitschritts schlüssig. Es handelt sich um einen Wochentag nach Ende der üblichen Arbeitszeit, genau zur Zeit des täglichen Lastmaximums, das bereits in Abschnitt 4.5.1 festgestellt wurde. Die Maximallast der Stadt wird also nicht durch einzelne Ausreißer entscheidend bestimmt, sondern durch gleichzeitige hohe Last in einem großen Teil der Häuser.




**(a)** Anzahl Personen **(b)** Energieverbrauch pro Person

**Abbildung 4.34:** Anzahl Personen (a) und der durchschnittliche Energieverbrauch pro Person (b) für jedes Haus in Jülich

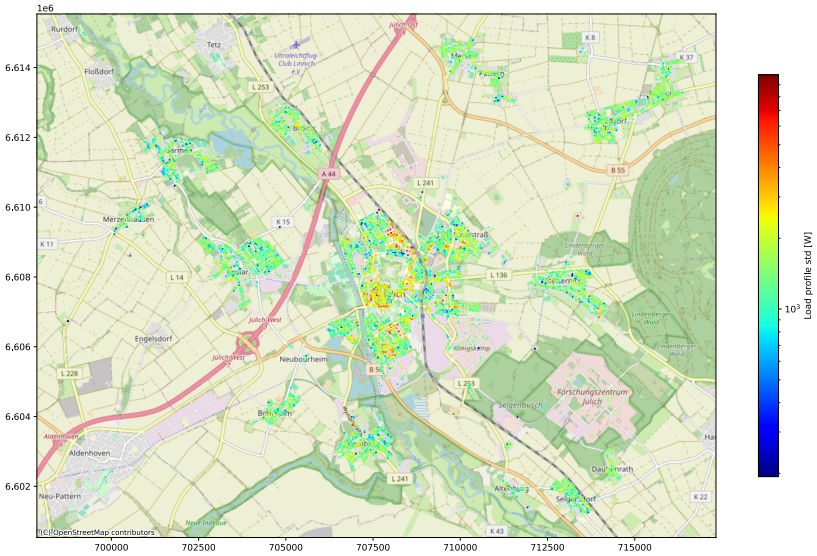



**Abbildung 4.35:** Standardabweichung der Lastprofile aller Häuser in Jülich

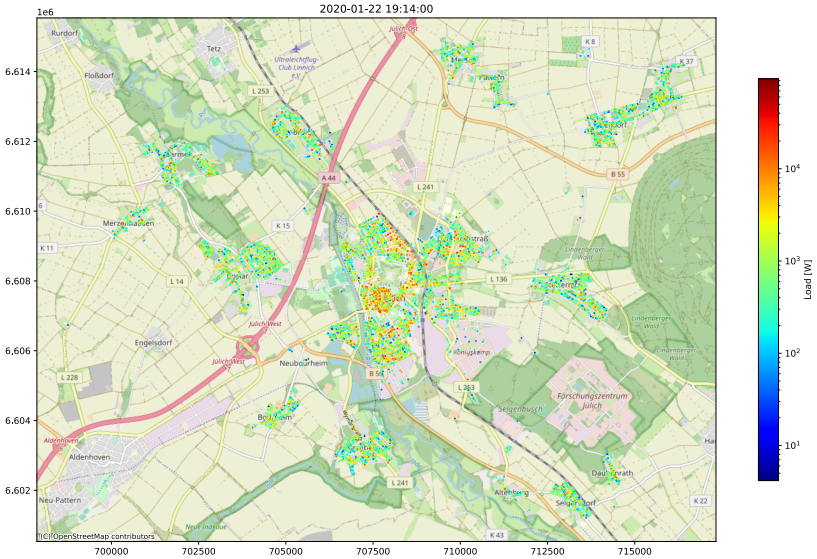
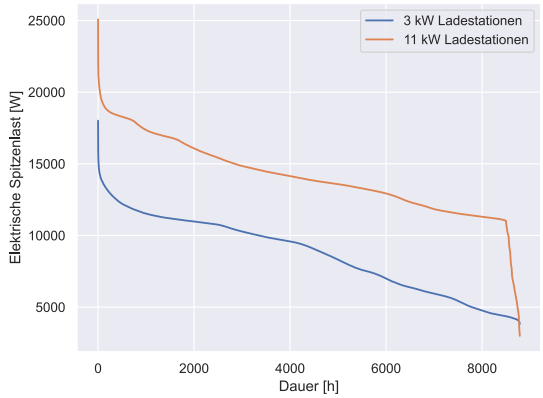



**Abbildung 4.36:** Alle Häuserlasten in Jülich vom Mittwoch, 22.01.2020, um 19:14 Uhr, dem Zeitpunkt der maximalen Gesamtstädtelast

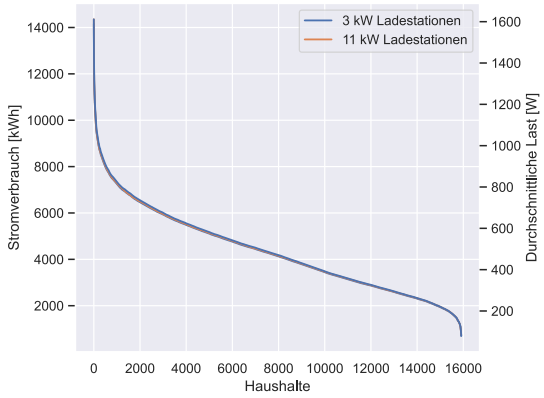
### 4.5.4 Auswirkungen verschiedener Elektroauto-Ladeleistungen

Die Städtesimulation ermöglicht durch die detaillierte Modellierung der einzelnen Komponenten einer Stadt die Betrachtung von Szenarien mit unterschiedlichen Parametern, um die Auswirkungen verschiedener Aspekte bewerten zu können. Im Folgenden werden beispielhaft die Folgen unterschiedlicher Ladeinfrastruktur für Elektrofahrzeuge auf den Stromverbrauch einer Stadt untersucht. Dazu wird das Jülich-Szenario in zwei verschiedenen Varianten simuliert, mit unterschiedlichen Ladeleistungen von 3 kW und 11 kW. In beiden Varianten wird zur Verdeutlichung der möglichen Auswirkungen in der Zukunft von einem vollständig elektrisch betriebenen Fahrzeugbestand ausgegangen.

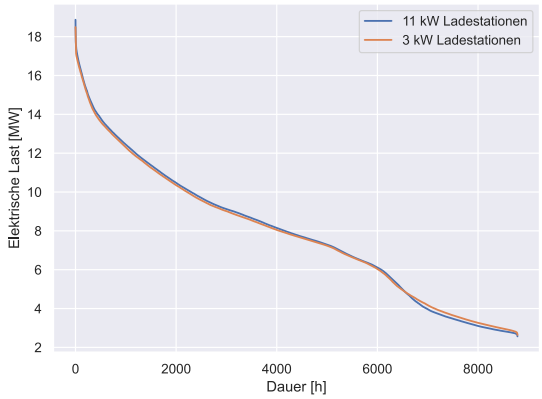
In Abbildung 4.37 werden die Auswirkungen der verschiedenen Ladeleistungen auf den Stromverbrauch der Stadt dargestellt. Die Unterschiede sind deutlich an den maximalen Haushaltslasten zu erkennen. Bei niedrigerer Ladeleistung treten Lasten unter 3 kW seltener auf, da die Ladezeiten länger dauern und somit häufiger mindestens die Ladelast von 3 kW anliegt. Lasten über 11 kW kommen dagegen kaum vor. Bei 11 kW Ladeleistung treten im Gegenzug häufiger deutlich höhere Lasten auf, insbesondere auch oberhalb von 11 kW. Die höchste Spitzenlast liegt erwartungsgemäß um etwa 8 kW höher als bei 3 kW Ladeleistung. Wie in Abbildung 4.37b zu sehen ist, bleiben die Gesamtverbräuche aller Haushalte aber unverändert, da in beiden Szenarien dieselbe Menge




**(a)** Maximale Last unter allen Haushalten




**(b)** Gesamtverbrauch aller Haushalte




**(c)** Aufsummierte Last der gesamten Stadt

**Abbildung 4.37:** Vergleich der Lastkurven in Szenarien mit 3 kW und 11 kW Ladeleistung für Elektroautos

an Energie zum Laden benötigt wird. Lediglich der zeitliche Verlauf ändert sich. Abbildung 4.37c zeigt die Auswirkungen auf das resultierende Lastprofil der gesamten Stadt. Bei 11 kW treten im Vergleich etwas häufiger besonders niedrige Lasten auf, dafür liegt die Spitzenlast der Stadt um 380 kW höher, wenngleich dieses Lastniveau nur selten erreicht wird. Dass die Spitzenlast auch im Szenario mit 11 kW nur um so wenig höher liegt, trotz einer Anzahl von 14.070 Autos, die durch die höhere Ladeleistung theoretisch eine Mehrlast von bis zu 112 MW verursachen könnten, liegt an der geringen Gleichzeitigkeit der Aufladungen. Da das Laden eines Autos mit 11 kW nach den meisten Fahrten nur einige Minuten dauert, finden stets nur verhältnismäßig wenige Aufladungen zur gleichen Zeit statt. Derartige Informationen sind insbesondere zur Auslegung der Verteilnetze einer Stadt relevant, da dazugehörige Anlagen und Leitungen entsprechend dimensioniert werden müssen.

## 4.6 Auslastung der Gesundheitsversorgung

Um die vielseitigen Einsatzmöglichkeiten der entwickelten Städtesimulation zu demonstrieren, wird mit dem Gesundheitssystems als Nächstes ein Bereich mit gänzlich anderen Anforderungen und Bewertungsmaßstäben betrachtet. Als Schwerpunkt wurden dabei örtliche Apotheken gewählt. Diese eignen sich besonders gut, da das übliche Besuchsmuster einer Einkaufsaktivität ähnelt und somit gut durch den LoadProfileGenerator simuliert werden kann. Durch die Synchronisierung der Simulation in jedem Zeitschritt und die explizite Abbildung von gegenseitigen Einflüssen der Personen bei ihrer Anwesenheit in Points of Interest können dabei die Auslastung der Geschäfte sowie Wartezeiten genau erfasst werden. Das ermöglicht eine Bewertung des untersuchten Systems im Hinblick auf Versorgung und Rentabilität.

### 4.6.1 Aktueller Stand in Jülich

Über OpenStreetMap sind in Jülich sechs Apotheken zu finden[1]. Für das Basisszenario wurde angenommen, dass jede dieser Apotheken über zwei Kassen verfügt, also zu jeder Zeit zwei Kunden gleichzeitig bedienen kann.

Zur Modellierung der Apothekenbesuche wurde im LoadProfileGenerator eine entsprechende Aktivität angelegt. Da alle Apotheken in Jülich sehr ähnliche Öffnungszeiten haben, wurde die Zeit von Montag bis Samstag zwischen 8 und 18 Uhr allgemein als zulässiger Zeitraum für diese Aktivität festgelegt. Die Angabe eines individuellen Zeitlimits pro Point of Interest wäre seitens der Simulation grundsätzlich möglich, für einen automatischen Import aus OpenStreetMap müssten dazu jedoch die Angaben zu Öffnungszeiten, falls verfügbar, eingelesen und in ein passendes Format gebracht werden. Aufgrund der geringen Schwankungen im gegebenen Szenario war das nicht notwendig.

Die Häufigkeit von Apothekenbesuchen variiert mitunter stark. Einige Personen suchen mehrmals pro Woche eine Apotheke auf, andere weniger als einmal im Jahr. Um diese Bandbreite möglichst gut abzudecken, wurden im LoadProfileGenerator fünf verschiedene Besuchsmuster definiert, von denen jeder Person zufällig eines zugewiesen wird. Um die Häufigkeiten insgesamt realistisch wiederzugeben, wurde dabei auf Statistiken und Studien zu Apothekenbesuchen zurückgegriffen

[1] Stand des OpenStreetMap-Datenabrufs: 24.10.2025

[69, 70, 71].

Auf der Karte in Abbildung 4.38 sind die sechs Apotheken in Jülich dargestellt. Die Farbskala zeigt die unterschiedliche Anzahl an Kundenbesuchen im Simulationszeitraum an. Hierbei ist zu erkennen, dass zentral gelegene Apotheken stärker besucht werden. Insbesondere der abgelegene Standort im Stadtteil Koslar (in der Abbildung links) wird um fast ein Drittel weniger häufig besucht als die Standorte in der Innenstadt. Diese Eigenschaft wurde nicht explizit modelliert, sondern ergibt sich aus der Präferenz der Personen für näher gelegene Points of Interest in Kombination mit der räumlichen Verteilung der Bevölkerung. Dieser Effekt ist somit ein Beispiel für die in Abschnitt 2.1 beschriebene Emergenz, eine große Stärke von Agentensimulationen.

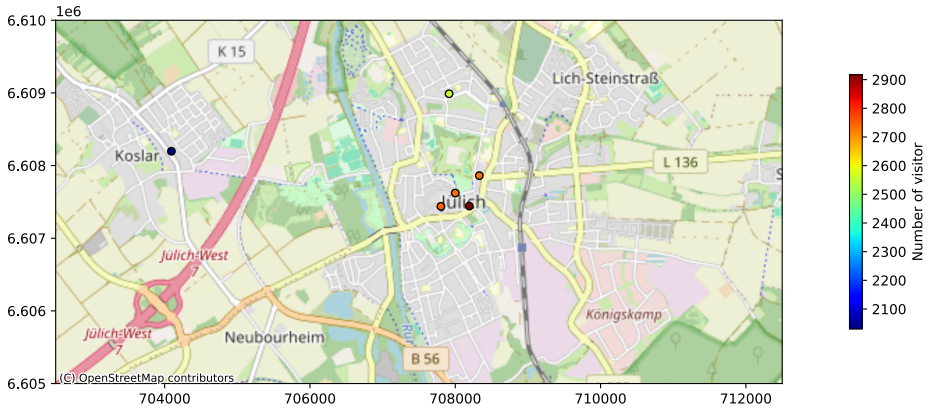
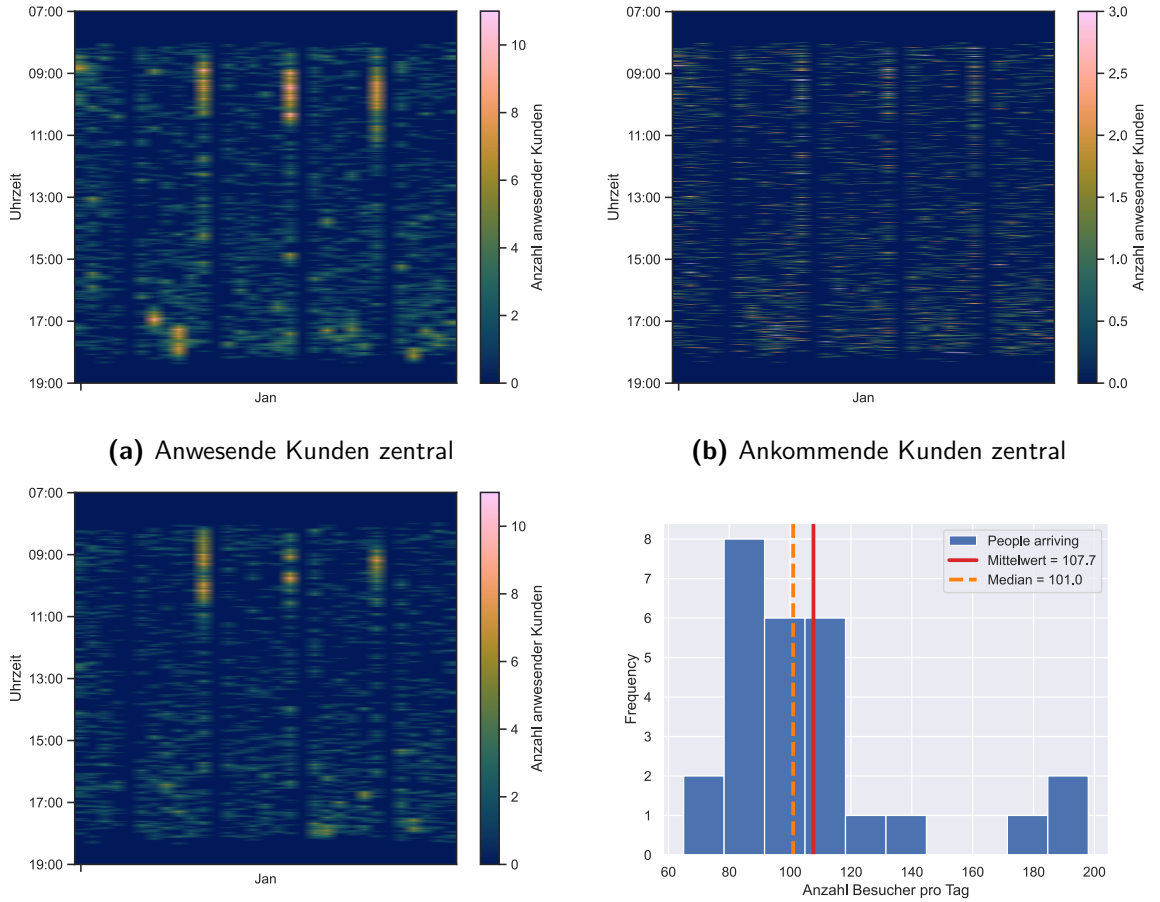



**Abbildung 4.38:** Existierende Apothekenstandorte in Jülich. Die Farbe zeigt die Anzahl der Kundenbesuche im Simulationszeitraum an.

Abbildung 4.39 zeigt den Betrieb in den Apotheken während der Öffnungszeiten. Zunächst ist zu sehen, dass dabei die Öffnungszeiten grundsätzlich eingehalten werden, außerhalb der vorgegebenen Zeiten und an Sonntagen gibt es keine Apothekenbesuche. Das Diagramm in Abbildung 4.39a zeigt die Anzahl anwesender Kunden in einer der Apotheken in der Jülicher Innenstadt für jeden Zeitschritt der Simulation. Hierbei stechen die Samstage mit besonders hohem Kundenaufkommen deutlich hervor. Das liegt daran, dass die Personen in der Simulation, der Realität entsprechend, an arbeitsfreien Tagen grundsätzlich mehr Zeit für sonstige Aktivitäten haben. Aktivitäten mit geringer Wiederholrate, die etwa nur wöchentlich oder noch seltener ausgeübt werden, finden daher bevorzugt an Wochenenden statt. Dabei ist zudem zu beachten, dass eine hohe Anzahl anwesender Kunden zu höheren Warte- und Aufenthaltszeiten führt und somit die Anwesenheit noch weiter steigert. Bei Betrachtung der bloßen Kundenankünfte pro Zeitschritt in Abbildung 4.39b fallen die Unterschiede daher deutlich weniger extrem aus.

Im Vergleich von Abbildung 4.39a und Abbildung 4.39c ist die Spanne der Besucherzahlen bei abgelegenen und zentralen Standorten zu erkennen. Sowohl der Wertebereich der Anzahl gleichzeitig anwesender Personen als auch die durchschnittliche tägliche Kundenzahl sind plausibel und passen zu Schätzungen des mittleren Kundenaufkommens in Apotheken.

Durch die explizite Simulation jeder Apotheke mitsamt der Warteschlange kann genau untersucht werden, wie lange jeder einzelne Kunde warten muss. Abbildung 4.40 zeigt die Verteilung der Wartezeiten zu unterschiedlichen Uhrzeiten für zwei Apotheken, zusammen mit der durchschnitt-




**(a)** Anwesende Kunden zentral

**(b)** Ankommende Kunden zentral

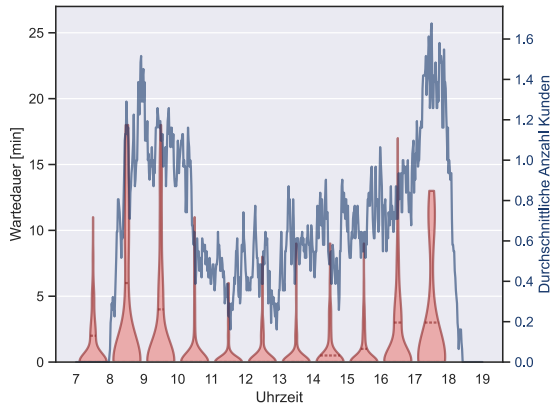
**(c)** Anwesende Kunden Koslar

**(d)** Tägliche Kundenzahl zentral

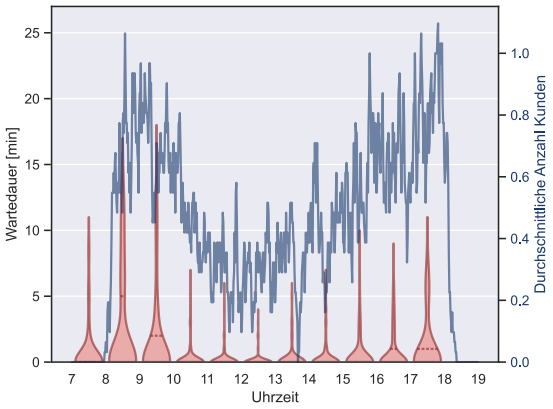
**Abbildung 4.39:** Rasterdiagramme zur Anzahl anwesender (a) und ankommender (b) Kunden in einer der zentral gelegenen Apotheken, sowie ankommende Kunden in der abgelegenen Apotheke in Koslar (c) und die Verteilung der täglichen Besuchszahlen in einer zentralen Apotheke (d).

lichen Anzahl anwesender Kunden als Referenz. Zusätzlich sind die Verteilungen der Wartezeit aus dem gesamten Simulationszeitraum dargestellt. In beiden Apotheken müssen die meisten Leute nicht warten und kommen sofort an die Reihe. Zu den Stoßzeiten morgens und abends kommt es zu kurzen Wartezeiten, größtenteils im Bereich weniger Minuten. Insbesondere abends, kurz vor Geschäftsschluss, sind die Wartezeiten etwas höher, weil zu dieser Zeit hauptsächlich noch Kunden anwesend sind, die bereits lange gewartet haben, und keine neuen Kunden mehr nachkommen. In der abgelegenen Filiale in Koslar sind die Wartezeiten zudem durchgehend geringer, da bei gleicher Verfügbarkeit und Kapazität ein geringerer Bedarf besteht.

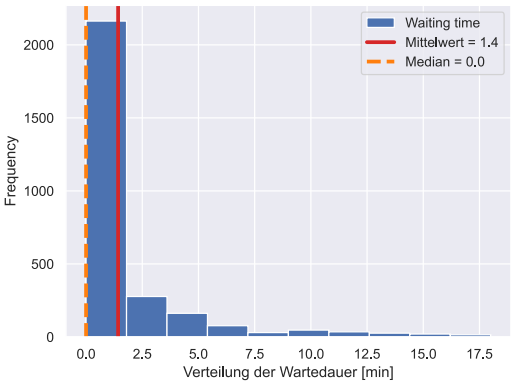
Für die Simulation von Points of Interest mit Warteschlangen wurde definiert, dass Kunden bei einer zu langen geschätzten Wartezeit ihren Besuch abbrechen. Diese Abbrüche stehen somit für Kunden, die nicht bedient werden konnten, und sind ein Maß für die Qualität der Versorgung. Die Anzahl nicht bedienter Kunden pro Tag ist in Abbildung 4.41 abgebildet. Darin ist zu erkennen, dass an einem Großteil der Tage sämtliche Kunden bedient werden können. An einigen Tagen kommt es zu wenigen Aktivitätsabbrüchen, möglicherweise verursacht durch Kunden mit besonders kurzer eigener Aktivitätsdauer und einer dementsprechend kurzen tolerierten Wartedauer. An vereinzelten Tagen liegt die Anzahl abgewiesener Kunden dagegen deutlich höher. Das sind die



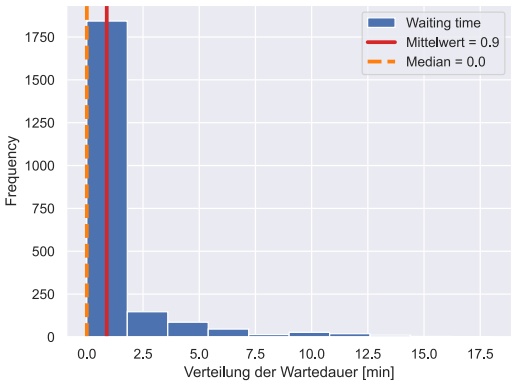

**(a)** Zentral




**(b)** Koslar




**(c)** Zentral



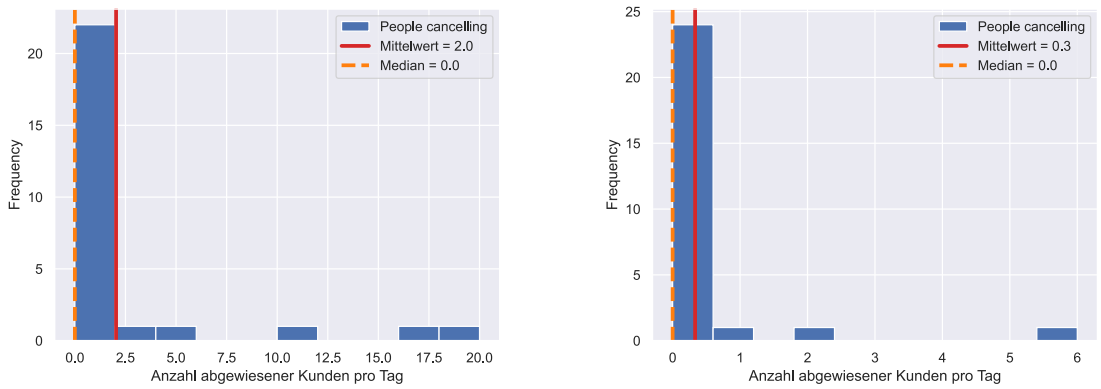

**(d)** Koslar

**Abbildung 4.40:** Verteilung der Wartezeit zu verschiedenen Uhrzeiten, für eine zentrale (a, c) und eine abgelegene Apotheke (b, d)

auch in den Rasterdiagrammen hervorstechenden Abschnitte mit besonders hoher Anwesenheit, in denen die jeweilige Apotheke durch den Kundenandrang überlastet ist und daher eine größere Anzahl an Kunden nicht versorgen kann. Diese Tage stellen Schwachpunkte im System der Gesundheitsversorgung dar, die reduziert oder im Idealfall vollständig vermieden werden sollten.

### 4.6.2 Alternative Szenarien

Eine der essenziellen Qualitäten eines Modells wie der Städtesimulation besteht in der Möglichkeit, nicht nur den aktuellen Stand abzubilden, sondern auch neue, hypothetische Szenarien betrachten und mit dem Status quo vergleichen zu können. So können in diesem Fall unterschiedliche Varianten bezüglich der Apotheken in Jülich simuliert werden, um herauszufinden, wie sich diese auf die Qualität des Gesundheitssystems auswirken. Um dies zu demonstrieren, werden im Folgenden drei Szenarien untersucht, die sich bezüglich der Apothekenstandorte leicht unterscheiden, ansonsten aber dem Basisszenario gleichen. Aus Geschwindigkeitsgründen wird jeweils nur ein Monat simuliert, da vorherige Simulationen gezeigt haben, dass die betrachteten Prozesse keinen signifikanten jahreszeitlichen Schwankungen unterliegen.



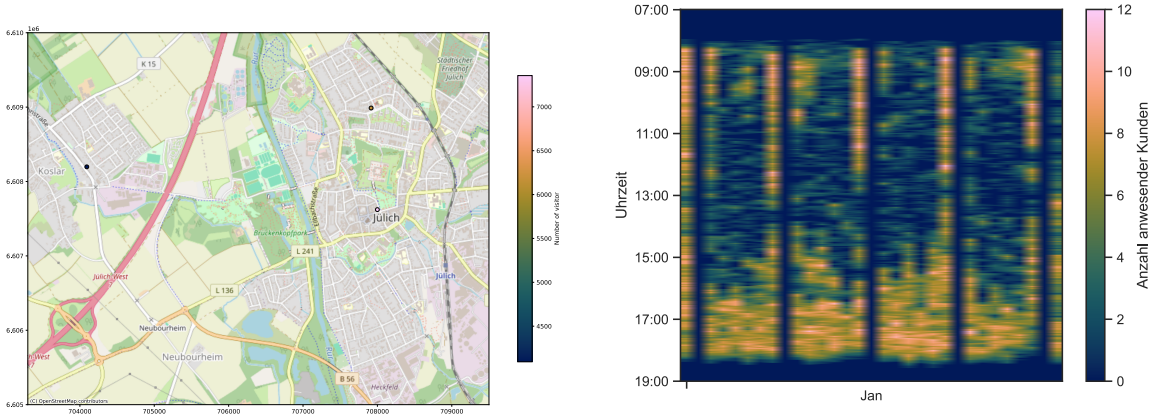

**(a)** Zentral **(b)** Koslar

**Abbildung 4.41:** Anzahl nicht bedienter Kunden pro Tag für eine zentrale (a) und eine abgelegene Apotheke (b)

### Szenario mit drei Apotheken

Als Erstes werden die Auswirkungen für den Fall untersucht, dass Apotheken ausfallen. Dazu wird ein identisches Städteszenario aufgebaut, noch vor der Zuordnung der Point-of-Interest-Präferenzen werden allerdings drei der Apothekenstandorte gelöscht. Die Bewohner verteilen sich somit, entsprechend den Deterrence-Funktionen im CityScenarioGenerator, auf die drei verbliebenen Standorte.

Wie in Abbildung 4.42a zu sehen ist, sind davon insbesondere die beiden zentraler gelegenen Standorte betroffen, da sie einen Großteil der zusätzlichen Kunden aufnehmen müssen. Das ist auch deutlich an der Anwesenheit in Abbildung 4.42b zu sehen, die nun häufiger an ihr Maximum heranreicht, insbesondere abends.

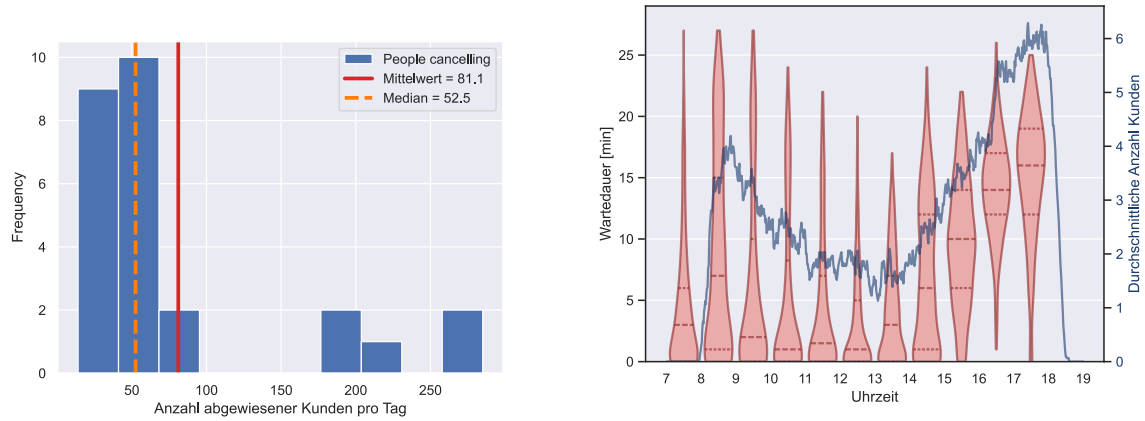



**(a)** Verbliebene Apotheken **(b)** Anwesenheit zentral

**Abbildung 4.42:** Die verbliebenen drei Apotheken im Szenario mit reduzierten Standorten (a) und die Anwesenheitszahl in der zentral gelegenen Apotheke (b)

Dies schlägt sich auch deutlich in der Anzahl Kunden nieder, die nicht bedient werden konnten. Die entsprechende Verteilung ist in Abbildung 4.43 abgebildet. Während dies zuvor eher ein Ausnahmefall war, werden jetzt jeden Tag etliche Kunden abgewiesen, im Durchschnitt über

80 täglich. Auch die Wartezeiten sind, vor allem in der Abendzeit, stark erhöht. Hier fällt auf, dass nicht nur die maximale Wartedauer steigt, sondern dass sich die gesamte Verteilung nach oben verschiebt. Zu den Stoßzeiten muss somit jeder Kunde im Mittel 15 Minuten warten.




**(a)** Abgewiesene Kunden **(b)** Wartezeit

**Abbildung 4.43:** Anzahl nicht bedienter Kunden pro Tag (a) und Wartezeit (b) für die zentrale Apotheke im Szenario mit reduzierten Standorten

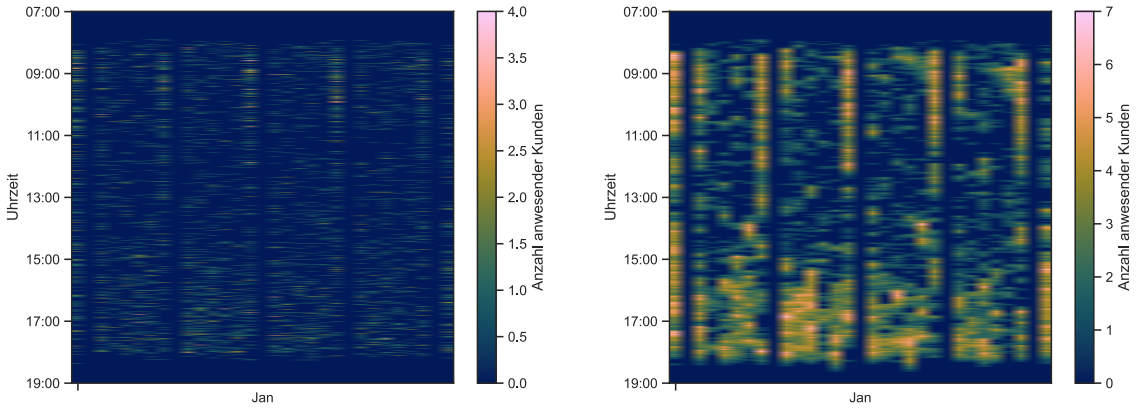
Dieses Szenario zeigt deutlich, dass eine derart starke Verringerung der Anzahl an Apotheken in Jülich zu einer Überlastung und dadurch zu einer Verschlechterung der Versorgung führen würde. Anhand der Besucherzahlen kann auch genau eingeschätzt werden, welche der verbliebenen Standorte hiervon wie stark betroffen wären.

### Szenario mit reduzierter Belegschaft

Um die Folgen eines Mangels an Fachkräften in der Pharmazie zu simulieren, wird die Anzahl an Kassen und somit gleichzeitig bedienbarer Kunden in jeder Apotheke von 2 auf 1 reduziert. Es ist somit keine parallele Abarbeitung der Kunden mehr möglich, und sobald ein Kunde bedient wird, müssen alle weiteren, die neu ankommen, zwangsläufig warten.

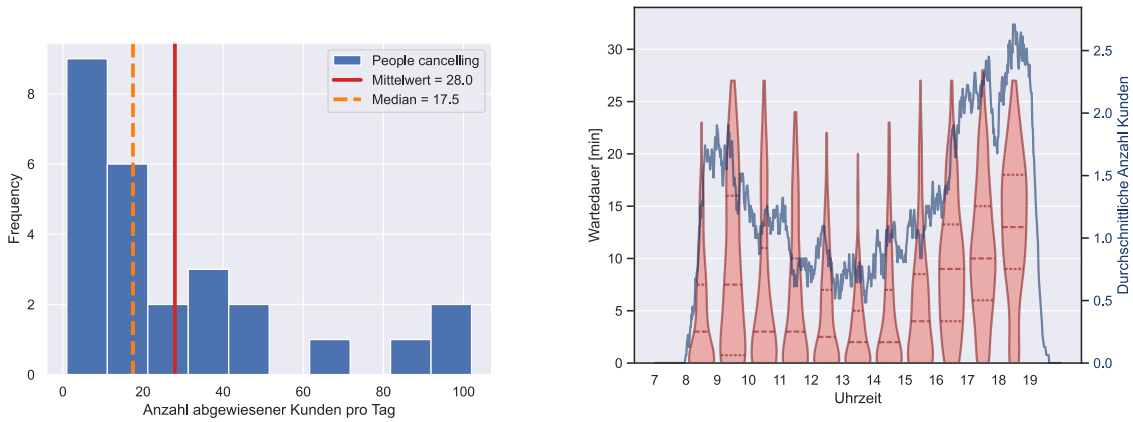
Durch diese Einschränkung wird die Kundenmenge pro Apotheke nicht beeinflusst, die Anzahl ankommender Kunden in jedem Zeitschritt in Abbildung 4.44 ist im Vergleich zum Basisszenario kaum verändert. Das Rasterdiagramm für die Anzahl anwesender Kunden zeigt dagegen eine deutliche Erhöhung, hervorgerufen durch die längeren Aufenthalte der Kunden in den Apotheken. An dieser Stelle ist zu beachten, dass die höchste Anzahl merklich unter dem Maximum im vorherigen Szenario mit reduzierter Apothekenzahl liegt.

Das hat auch Auswirkungen auf die Wartezeit und Anzahl abgewiesener Kunden, dargestellt in Abbildung 4.45. Im Vergleich zum vorherigen Szenario werden zwar pro Apotheke weniger Kunden abgewiesen, da jedoch jeder der Standorte nun deutlich weniger Kunden bewältigt, ergibt sich insgesamt ein ähnliches Bild. Auch hier ist zu beobachten, dass die Wartezeit durchgehend merklich ansteigt, insbesondere jedoch ab 17 Uhr. Sowohl Median als auch mittlere Wartedauer liegen dabei etwas höher. Beide Szenarien wirken sich also wie erwartet sehr negativ auf die Qualität der Versorgung aus, allerdings auf leicht unterschiedliche Weise.




**(a)** Ankommende Kunden **(b)** Anwesende Kunden

**Abbildung 4.44:** Anzahl ankommender (a) und insgesamt anwesender Kunden (b) für eine zentrale Apotheke im Szenario ohne parallele Bedienung




**(a)** Abgewiesene Kunden **(b)** Wartezeit

**Abbildung 4.45:** Anzahl nicht bedienter Kunden pro Tag (a) und Wartezeit (b) für eine zentrale Apotheke im Szenario ohne parallele Bedienung

### Szenario mit optimierter Standortplanung

Als Gegenentwurf zu den bereits gezeigten schlechteren Szenarien soll abschließend ein Szenario mit gezielter Standortplanung simuliert werden. Dabei werden mithilfe des Optimierungsmodells von Büsing, Leitner und Wrede [72] unter einer Menge vorgegebener Standortkandidaten die Standorte ausgewählt, die die gewählte Zielfunktion minimieren. Auf diese Weise soll untersucht werden, wie sich die Auslastung der Apotheken verändert, wenn lediglich die Position einzelner Apotheken variiert wird.

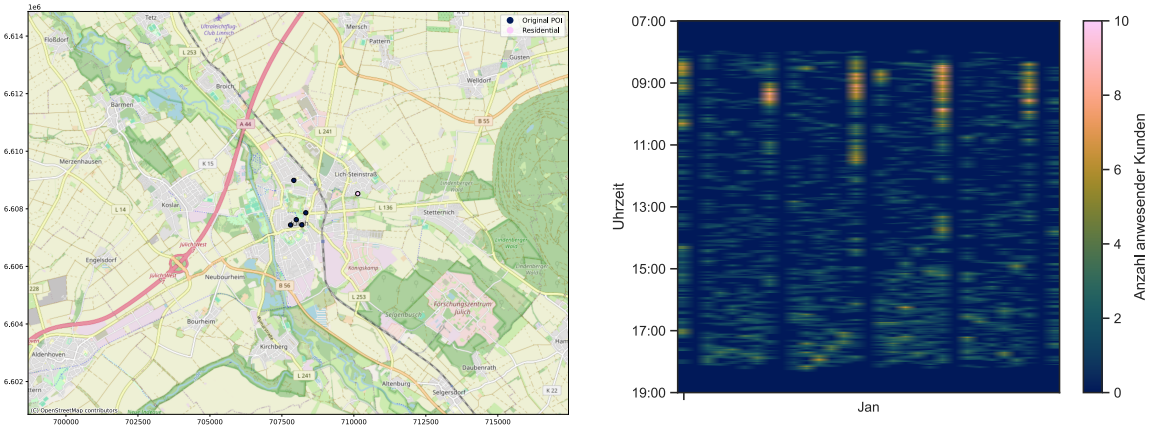
Um ein vergleichbares Szenario zu erhalten, wurde die Anzahl an Standorten, die eröffnet werden dürfen, auf sechs festgelegt. Da für die Untersuchung vor allem mögliche Standorte in der Nähe der Kunden von Interesse waren, wurden aus allen Wohngebäuden in Jülich 10 zufällig ausgewählt und als Standortkandidaten bestimmt. So kam eine über ganz Jülich verteilte Standortmenge zustande, mit einem Schwerpunkt auf dem dichter besiedelten Kerngebiet. Zusätzlich wurden die sechs tatsächlich existierenden Apothekenstandorte hinzugefügt. Aufgrund der Komplexität des Modells konnten Instanzen mit einer größeren Anzahl an Kandidatenstandorten vorerst

nicht berechnet werden. Weitere Modelloptimierungen oder effizientere Algorithmen würden die Optimierung solcher Instanzen aber aller Voraussicht nach ermöglichen.

Als Nebenbedingung wird im Modell gefordert, dass jede Person ihre jeweils nächstgelegene eröffnete Apotheke besucht. Damit wird dem Umstand Rechnung getragen, dass die Bewohner einer Stadt auch im Falle einer koordinierten Planung von Apothekenstandorten selbst auswählen, welche Filiale sie besuchen, und die Entfernung dabei eine entscheidende Rolle spielt. Um geringfügige Unterschiede in der Distanz nicht überzubewerten, werden alle Distanzen auf 10 Meter gerundet und Indifferenz zwischen gleich weit entfernten Standorten zugelassen.

Für die Kostenfunktion kommen verschiedene Varianten in Betracht, je nachdem, ob beispielsweise möglichst kurze Wege für die Personen zu ihrer jeweiligen Apotheke oder eine möglichst gleichmäßige Verteilung der Kunden auf die Standorte erzielt werden soll. Hier wurde letzterer Ansatz gewählt, in der Annahme, dass eine gleichmäßige Verteilung Überlastung reduziert und damit Ausfälle und Wartezeiten verringert, möglicherweise auf Kosten leicht erhöhter Wegstrecken. Um das zu erreichen, wird die Differenz der oberen und unteren Schranke für die Anzahl Kunden pro Apotheke minimiert. Die genaue Formulierung des Modells ist in Anhang A.3 zu finden.

Die vom Modell ausgewählten Standorte sind in Abbildung 4.46 dargestellt. Dabei fällt auf, dass fünf der existierenden Standorte übernommen wurden, lediglich die abgelegene Apotheke in Koslar wurde zugunsten eines zentraleren Standorts verworfen.




**(a)** Ausgewählte Standorte

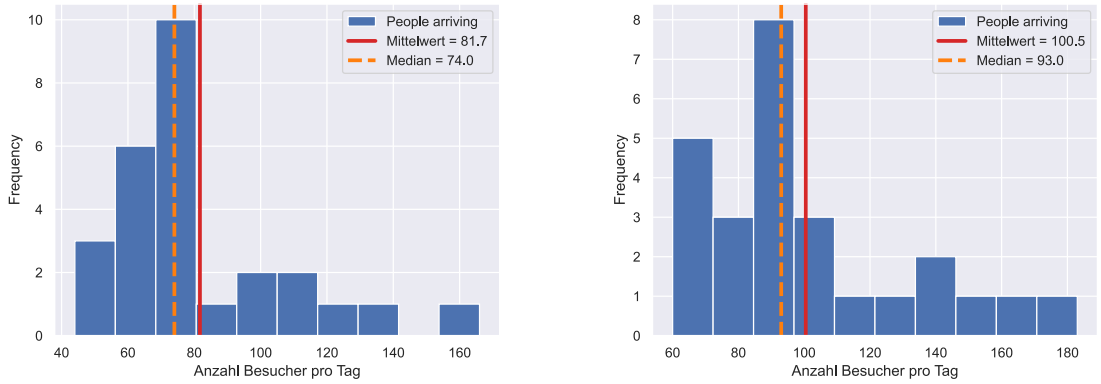
**(b)** Anwesende Kunden

**Abbildung 4.46:** Die vom Standortplanungsmodell gewählten Apothekenstandorte (a) und die Anzahl anwesender Kunden im neu eröffneten Apothekenstandort (b)

Abbildung 4.47 zeigt, dass der neue Standort stärker besucht ist als die ursprüngliche Filiale in Koslar, mit im Schnitt 19 Kundenbesuchen mehr pro Tag. Insgesamt ist dieses Szenario somit ausgeglichener.

### Auswertung der Szenarien

Um die Qualität des simulierten Gesundheitssystems zu bewerten, sind der Anteil nicht bedienter Kunden und die durchschnittliche Wartezeit geeignete Indikatoren. Beides ist in Tabelle 4.3 für jedes der vier Szenarien angegeben. Hier fallen insbesondere die deutlich schlechteren Werte der Szenarien mit weniger Apotheken und weniger Personal auf. Das optimierte Szenario liegt in bei-




**(a)** Standort in Koslar **(b)** Neu gewählter Standort

**Abbildung 4.47:** Anzahl täglicher Kundenbesuche in der Filiale in Koslar im Basisszenario (a) und des neu gewählten Standorts im optimierten Szenario (b)

**Tabelle 4.3:** Anteil nicht bedienter Kunden und durchschnittliche Wartezeit in jedem der vorgestellten Szenarien

| Szenario | Anteil abgewiesener Kunden | Mittlere Wartedauer in Minuten |
|---|---|---|
| Basisszenario | 2,6% | 1,6 |
| Drei Apotheken | 20,8% | 5,6 |
| Reduzierte Belegschaft | 20,6% | 6,4 |
| Optimierte Standorte | 2,7% | 1,6 |

den Werten sehr nah bei denen des Basisszenarios. Das lässt mehrere Schlussfolgerungen zu. Zum einen ist die Ausgangssituation im Basisszenario offenbar bereits verhältnismäßig gut, sodass das Modell in seiner bisherigen Form keine bessere Lösung finden konnte. Das ist plausibel, denn eine sehr unausgewogene Verteilung würde in der Realität nicht gut funktionieren und hätte vermutlich keinen Bestand. Zum anderen ist das Modell noch nicht ausgereift genug, um bessere Lösungen zu finden. Insbesondere konnten größere Lösungsräume mit mehr Kandidaten noch nicht optimal verarbeitet werden, sodass nur eine sehr eingeschränkte Menge an Standortkandidaten zur Auswahl stand. Es steht zu vermuten, dass mit einer größeren Auswahl bessere Lösungen gefunden werden können. Zuletzt ist aber die gewählte Kostenfunktion nur ein Ansatz, um den Bedarf gleichmäßig zu verteilen. Tatsächliche Überlastung tritt jedoch, wie in den Rasterdiagrammen zu sehen ist, stets in sehr begrenzten Zeiträumen auf und ist demnach stark von den Tagesabläufen der entsprechenden Bewohner abhängig. Die Gesamtanzahl an Kunden ist somit nur eine vereinfachende Abschätzung des Bedarfs, den eine Apotheke zu bewältigen hat. Eine detailliertere Kostenfunktion mit einem anderen Ansatz könnte den zeitlichen Aspekt der Apothekenbesuche möglicherweise besser abbilden und zu einem ausgewogeneren Ergebnis führen.

# 5 Diskussion

Im Rahmen dieser Arbeit wurde eine umfassende, integrierte Städtesimulation entwickelt. Das Ergebnis ist eine Simulationssoftware, die das alltägliche Verhalten aller Einwohner einer gesamten Stadt abbilden kann. Die Simulation generiert realistische Tagesabläufe, die angeben, wann eine Person beispielsweise schläft, arbeitet oder einem Hobby nachgeht. Dabei werden alle Objekte der Stadt räumlich und zeitlich präzise erfasst und über den Lauf der Simulation durchgängig nachverfolgt. Die Aufenthaltsorte aller Bewohner sind somit zu jedem Zeitpunkt bekannt. Abhängigkeiten zwischen Personen oder Einrichtungen werden in grundlegender Weise modelliert, und es wurden die Voraussetzungen geschaffen, um weitere Effekte und Einwirkungen zu implementieren. So wird zum Beispiel die Wartedauer in einer Schlange, etwa beim Einkaufen, explizit modelliert, in Abhängigkeit von den anwesenden Kunden und individuell für jeden Supermarkt in der betrachteten Stadt. Damit die Simulation in der Praxis sinnvoll und nutzbringend eingesetzt werden kann, wurde zum einen mit dem CityScenarioGenerator ein Softwarepaket zur automatischen Erstellung von Städteszenarien entwickelt, und zum anderen eine effiziente, parallelisierbare Simulationsarchitektur geschaffen, die Städte mit mehreren 10.000 Einwohnern in wenigen Stunden bis Tagen vollständig simulieren kann. Insgesamt wurde somit eine sehr umfangreiche und detaillierte Simulationsplattform aufgebaut, die effiziente Verhaltenssimulationen im Städtekontext für verschiedene Anwendungsszenarien ermöglicht.

In der Literatur existiert bereits eine Reihe an Modellen, die sich Prozessen in Städten widmen, beispielsweise die in Kapitel 2 vorgestellten Urban Building Energy Models oder Transportsimulationen. Diese Modelle stützen sich jedoch durchweg auf stark vereinfachte Abbildungen menschlichen Verhaltens. Größtenteils werden statische Aktivitätspläne genutzt, die teils sogar von einer externen Quelle vorgegeben werden müssen. Zudem ist die Menge möglicher Aktivitäten in der Regel eingeschränkt. Im Gegensatz dazu bietet die entwickelte Städtesimulation eine sehr detaillierte, erweiterbare und zeitlich wie räumlich hochaufgelöste Aktivitätsmodellierung. Durch das zugrundeliegende Bedürfnismodell werden die wichtigsten Aspekte menschlichen Verhaltens berücksichtigt und anschließend durch die Vielzahl verfügbarer Aktivitäten präzise abgebildet.

Der Detailgrad der Städtesimulation in Bezug auf menschliches Verhalten wird bisher nur von Modellen auf Haushaltsebene erreicht. Diese bilden allerdings weder städtische Infrastruktur noch Auswirkungen über einen einzelnen Haushalt hinaus ab. Aktivitätsprofile werden nur für jeden Haushalt individuell generiert, ohne gegenseitige Einflussnahme oder Wechselwirkungen durch Begegnungen und gleichzeitige Aufenthalte am selben Ort. Die Städtesimulation ist das erste Modell, das eine detaillierte Verhaltenssimulation mit der Simulation einer Stadt kombiniert. Ermöglicht wird das durch den Ansatz der synchronisierten Simulation, ohne die Änderungen nicht rechtzeitig ausgetauscht und berücksichtigt werden könnten. Nur so werden wechselseitige Einflüsse der Personen und anderer Modellkomponenten aufeinander erfasst, was bei bisherigen Modellen in dieser Form nicht möglich ist.

Durch die Vereinigung dieser beiden grundlegenden Aspekte in einem Modell können Städte erst ausreichend umfassend und integriert betrachtet werden, um komplexere Fragestellungen zu beantworten. So kann mit der Städtesimulation analysiert werden, wie sich das Verhalten der Einwohner auf das Besucheraufkommen in Geschäften oder die Auslastung der Verkehrsnetze auswirkt,

und wie diese Größen in Abhängigkeit von Standorten oder Öffnungszeiten variieren. Insbesondere können genauere Untersuchungen unter Berücksichtigung der Entscheidungen der Bewohner einer Stadt angestellt werden. Zum Beispiel kann auf diese Weise der Einfluss geänderter Ladenöffnungszeiten oder besserer Verkehrswege auf die Tagesabläufe der Bewohner bewertet werden. Zu derartigen Fragestellungen lassen sich mithilfe der Städtesimulation fundierte Aussagen treffen, die durch realistische Aktivitätssequenzen, Ortspräferenzen und dazu konsistente Reisezeiten, Auslastungen von Einrichtungen, Energieverbräuche und weitere Daten gestützt werden.

In Kapitel 4 wurde gezeigt, dass menschliches Verhalten und die daraus resultierenden Tagesabläufe sehr individuell sind und sich insbesondere zwischen verschiedenen Personentypen stark unterscheiden. Die detaillierte Abbildung einzelner Personen mit spezifischen Merkmalen ist somit für realistische Ergebnisse unerlässlich. Voraussetzung dafür ist eine angemessene Repräsentation der zu modellierenden Bevölkerung mit korrekter Demographie. Selbst für die existierenden Verhaltensmodelle mit hoher Genauigkeit existiert keine Methodik, um die Modelle für eine derart große Personenanzahl passend zu konfigurieren und eine entsprechend realitätsgetreue Simulation durchzuführen. Durch die umfangreiche Datenaufbereitung und Szenarioerstellung im CityScenarioGenerator wird eine präzise Modellierung mit der Städtesimulation erst ermöglicht. Beide Komponenten zusammengenommen ist die entwickelte Simulationsplattform in der Lage, individuell auf die Bevölkerung abgestimmte Tagesabläufe einer ganzen Stadt mit deutlich höherem Detailgrad zu generieren als bisherige Modelle.

Durch die explizite Modellierung verschiedener Points of Interest können die Auswirkungen von Standort, Dimensionierung und weiteren Eigenschaften diverser Einrichtungen direkt untersucht werden, sowohl im Hinblick auf die Tagesabläufe der Personen, die diese Einrichtungen besuchen, als auch auf die Einrichtungen selbst. Dabei können etwa die täglichen Besucherzahlen und ihre zeitliche Verteilung über den Tag analysiert werden, um Hinweise auf die Versorgung der Bevölkerung einerseits und die Rentabilität von Geschäften andererseits zu erhalten. Ein Vergleich mehrerer Szenarien könnte beispielsweise Auskunft über gut geeignete Standorte für eine neue Arztpraxis geben, die bestehende Praxen in der Umgebung entlasten kann. Derartige Daten sind bisher kaum verfügbar und können ohne eine synchronisierte und zeitlich und räumlich hochaufgelöste Verhaltenssimulation nicht in dieser Form gewonnen werden.

Die Modellierung von Transport und Reisezeiten erfolgt in der Städtesimulation selbst nur in sehr simpler Form, da es bereits viele Modelle gibt, die sich auf diesen Bereich spezialisieren und ihn sehr detailliert abdecken. Die in Kapitel 3 demonstrierte Option der Kopplung mit einem solchen Transportmodell bietet dennoch die Möglichkeit, passendes Transportverhalten bei den Tagesabläufen der Stadtbewohner miteinzubeziehen.

Der vollständig agentenbasierte Ansatz ist sehr rechenintensiv, garantiert dafür aber größtmögliche Flexibilität. Durch geringfügige Modifikation können etwa neue Gebäudetypen hinzugefügt und in die Simulation integriert werden. So können mit minimalem Aufwand verschiedenste Szenarien untersucht und entsprechende Daten generiert werden. Dass die effiziente Implementierung auf parallelen Rechnerarchitekturen trotzdem eine ausreichend gute Performance bietet, bestätigt, dass dieser Ansatz für diese Art von Problemen gut geeignet ist.

Die erstellte Simulation eröffnet durch ihren großen Umfang und gleichzeitig hohen Detailgrad

neue Möglichkeiten im Bereich der Modellierung von Städten. Ein wesentlicher Mehrwert besteht in der integrierten und räumlich sowie zeitlich hoch aufgelösten Modellierung. Jegliche Ergebnisdaten, ob Aktivitätsprofile, Lastkurven verschiedener Typen oder Besuchsprofile, werden für jede Instanz der simulierten Stadt für jede Minute innerhalb der betrachteten Zeitspanne generiert. Diese Daten lassen sich verwenden, um präzisere Aussagen im Hinblick auf mögliche Änderungen an der Städteinfrastruktur zu treffen. Dadurch können beispielsweise neue Varianten für die lokale Energieversorgung mit Lastprofilen für jedes einzelne Haus genau überprüft und so auf ihre Praxistauglichkeit hin bewertet werden. Auch die erzeugten Aktivitätsprofile ermöglichen tiefergehende Analysen, etwa in Bezug auf Gleichzeitigkeit verschiedener Verhaltensmuster und sich daraus ergebende Konsequenzen für verschiedene städtische Versorgungssysteme. Zwar wird das Verhalten während der Reisen von der Städtesimulation selbst nicht im Detail modelliert, Zeitpunkt und Ziel jeder Reise sind aber hilfreiche Informationen für detailliertere Untersuchungen der Verkehrsnetze.

Insbesondere die Daten zu verschiedenen Einrichtungen können auf vielfältige Weise eingesetzt werden. Neben der Ermittlung potenzieller Kundenzahlen für verschiedene Unternehmensstandorte können so Pläne für eine möglichst gute Abdeckung mit bestimmten Dienstleistungen erarbeitet werden. Die Simulation mehrerer Szenarien erlaubt den direkten Vergleich verschiedener Konzepte mit unmittelbarer Validierung, um Erreichbarkeit und Verfügbarkeit zu optimieren. In Kombination mit Modellen für die Bestimmung geeigneter Standortkandidaten oder weiterer Szenarioparameter kann so eine bestmögliche Versorgung bei minimalem Aufwand sichergestellt werden. Die Simulation kann also sowohl bei der Wahl möglichst profitabler Standorte für etwa Restaurants oder Supermarktfilialen helfen als auch eine möglichst geeignete Verteilung von Institutionen im medizinischen Bereich unterstützen, damit der Bedarf jedes Einwohners einer Stadt gedeckt werden kann.

Die in Abschnitt 4.6.2 beschriebene Kopplung mit einem Optimierungsmodell zur Planung von Apothekenstandorten demonstriert einen nützlichen Anwendungsfall für die Städtesimulation. Optimierte Pläne und Konzepte berücksichtigen aus Komplexitätsgründen in der Regel nicht alle Aspekte oder Dimensionen eines Problems, und es kann schwierig sein, zu entscheiden, ob bestimmte Vereinfachungen zulässig sind oder in der Realität zu großen Diskrepanzen führen. Die Städtesimulation kann in solchen Fällen genutzt werden, um die generierten Szenarien zu simulieren und die Auswirkungen der geplanten Veränderungen zu überprüfen und zu validieren.

Aufgrund des enormen Umfangs der Vorgänge selbst in kleinen Städten konnten einige Aspekte nur in vereinfachter Form berücksichtigt werden. Eine der relevantesten Einschränkungen betrifft die Auswahl der Gebäude im Rahmen der Szenariogenerierung. Im aktuellen Stand beschränkt sich die Simulation auf eine einzelne Stadt. Es werden daher lediglich alle Gebäude gesammelt, die offiziell zu der zu simulierenden Stadt gehören. Für Wohngebäude ist das eine sinnvolle Abgrenzung, und für viele Nichtwohngebäude wie Supermärkte oder Apotheken ist diese Einschränkung ebenso vollkommen ausreichend, da ein Großteil der alltäglichen Aktivitäten wie Einkaufen typischerweise in der Nähe des eigenen Zuhauses stattfindet. Einige Typen von Points of Interest verfügen jedoch über ein deutlich größeres Einzugsgebiet, sodass Personen sich hierfür über die Stadtgrenzen hinausbewegen. Insbesondere die Arbeitsplätze vieler Personen befinden sich nicht in deren Wohnort. Durch die fehlende Abbildung dieser längeren Reisen sind Pendeldistanzen im Bewegungsmuster der Bewohner im Vergleich zur Realität reduziert. Abgesehen von etwas längeren Reisedauern

sind die Auswirkungen davon auf die Ergebnisse der Städtesimulation allerdings minimal. Eine mögliche Lösung besteht darin, einen Pufferradius um die modellierte Stadt zu setzen. Innerhalb dieses Bereichs könnten dann zusätzlich alle Nichtwohngebäude als potenzielle Points of Interest zur Simulation hinzugefügt werden. Damit würde jedoch auch die Anzahl an Objekten in der Simulation deutlich steigen, sodass die Laufzeit gegebenenfalls beeinträchtigt wäre.

Eine weitere Einschränkung betrifft die Anzahl der Zielorte pro Person. In den durchgeführten Beispielrechnungen wurde aus Effizienzgründen nur ein Point of Interest jeder Kategorie für jede Person ausgewählt. Für einige Kategorien, beispielsweise für Supermärkte oder Orte zum Spazierengehen, wären mehrere Zielorte pro Person realistischer. Da die Gesamtanzahl an Besuchen in den betroffenen Orten davon nicht beeinflusst wird, sind die Auswirkungen gering. Eine gute Lösung wäre eine dynamische Bestimmung der Anzahl an Zielorten für jede Person, abhängig von der Kategorie. Für viele Fragestellungen ist die performantere Variante in der aktuellen Implementierung aber bereits ausreichend.

Familiäre und freundschaftliche Beziehungen werden in der Städtesimulation bislang nicht berücksichtigt. Häuser von Freunden und Verwandten werden lediglich wie alle Zielorte probabilistisch ausgewählt. Eine genauere Abbildung sozialer Netzwerke könnte für bestimmte Fragestellungen relevant sein, für die hier betrachteten Untersuchungen spielen diese Aspekte jedoch keine ausschlaggebende Rolle.

In Bezug auf die Performance der Städtesimulation wurde das gesetzte Ziel erreicht. Szenarien mittlerer Größe können in wenigen Tagen vollständig berechnet werden. Dadurch ist es möglich, in kurzer Zeit verschiedene Szenarien auszuwerten und zu vergleichen. Für aufwendige Studien, die eine sehr viel größere Anzahl an Simulationen erfordern, oder die, etwa im Rahmen eines iterativen Ansatzes, mehrere Simulationen nacheinander durchführen müssen, könnte die Simulationsdauer jedoch zu einem hemmenden Faktor werden. Für derartige Zwecke wären gegebenenfalls weitere Optimierungen oder Vereinfachungen notwendig.

Insgesamt wurde eine sehr umfangreiche, detaillierte und vielseitig einsetzbare Simulationssoftware erstellt. Die noch vorhandenen Einschränkungen können weitestgehend ohne großen Aufwand adressiert werden. Insbesondere der Einbezug weiterer Points of Interest aus dem Umland einer Stadt könnte die Ergebnisse für einige Zielort-Kategorien verbessern. Im Vergleich zu bisherigen Verhaltensmodellen konnte die Städtesimulation durch die Abdeckung ganzer Städte eine neue Modellierungsebene erreichen. Die Simulation weiter auszubauen, um das betrachtete Gebiet auszudehnen und zusätzliche Effekte und Wechselwirkungen über die Grenzen einer Stadt hinaus zu berücksichtigen, ist daher ein naheliegender und sinnvoller nächster Schritt für zukünftige Forschung. Des Weiteren besteht viel Potenzial für eine Anwendung der Städtesimulation, zum Beispiel im Energiebereich oder für die Verbesserung des Gesundheitssystems. In diesen und anderen Anwendungsgebieten können die Möglichkeiten der Städtesimulation genutzt werden, um existierende Systeme besser zu modellieren und neue Konzepte für effizientere, leistungsfähigere und menschenfreundlichere Städteinfrastrukturen zu finden.

# 6 Zusammenfassung und Ausblick

In dieser Arbeit wird die Entwicklung und Anwendung einer Simulationssoftware zur Modellierung von Städten beschrieben. Zunächst wird die Notwendigkeit eines integrierten Städtemodells mit detaillierter Verhaltensmodellierung begründet. Strukturen und Vorgänge in Städten sind sehr komplex und in hohem Maße von den beteiligten Menschen und deren Verhalten abhängig. Für eine sinnvolle Abbildung müssen daher die Tagesabläufe dieser Menschen berücksichtigt werden.

Bereits existierende Methoden zur Verhaltens- und Städtemodellierung werden in einem Überblick zum aktuellen Stand des Themas vorgestellt und verglichen. Dabei wird hervorgehoben, dass diese Modelle vor allem in Bezug auf menschliche Aktivitäten verhältnismäßig grob und ungenau sind, und sich meist auf einen stark eingegrenzten Bereich fokussieren. Eine vollständige, integrierte Simulation der wesentlichen Aspekte in einer Stadt, inklusive der vollständigen Tagesabläufe der Personen und ihrer Reisen und Aufenthalte an verschiedenen Orten, bietet keiner der Ansätze.

Daher wird im Anschluss die Entwicklung einer auf einem existierenden Verhaltensmodell basierenden Agentensimulation beschrieben. Die Grundlage bildet dabei das Message Passing Interface (MPI), das zur Parallelisierung der rechenintensiven Simulation eingesetzt wird. Darauf aufbauend wird die Struktur der Städtesimulation erläutert, bestehend aus der gleichmäßigen Verteilung aller Stadtbewohner auf mehrere Worker und einer Synchronisation in jedem Zeitschritt zur Aufrechterhaltung der Konsistenz. Zusätzlich wird der CityScenarioGenerator vorgestellt, ein separates Softwarepaket zur Erstellung realitätsgetreuer Städteszenarien unter Verwendung detaillierter Gebäude- und Unternehmensdaten.

Schließlich wird die Funktionalität der implementierten Städtesimulation anhand der Stadt Jülich als Anwendungsbeispiel gezeigt. Dabei werden die Aktivitäten der Bewohner im Detail ausgewertet und mit Zeitverwendungsdaten validiert. Anschließend werden der Haushaltsenergiebedarf der simulierten Stadt und die Auslastung verschiedener Einrichtungen analysiert. Durch den Vergleich der Simulationsergebnisse für verschiedene Städteszenarien wird der Nutzen der Städtesimulation zur Bewertung neuer Konzepte für städtische Infrastruktur demonstriert.

Abschließend wird die Städtesimulation als Ergebnis dieser Arbeit evaluiert und diskutiert. Der hohe Detailgrad und die integrierte Abbildung verschiedener Aspekte ermöglichen neue, tiefgehende Untersuchungen von Städten und Plänen für potenzielle Veränderungen. Einschränkungen bestehen im Wesentlichen in der Berücksichtigung der Umgebung einer Stadt. Diese Limitierungen können im Rahmen weiterer Forschung angegangen werden, um die Möglichkeiten, die die Städtesimulation bietet, weiter auszubauen und zu verbessern.

# A Anhänge

## A.1 Einbinden der MPI.NET-Bibliothek

Für die Einbindung unter Windows steht MPI.NET als nuget-Package zur Verfügung. Dieses Package wurde allerdings nur für die Verwendung mit der Windows-exklusiven MPI-Implementierung Microsoft MPI kompiliert und kann mit anderen MPI-Implementierungen nicht genutzt werden. Der Einsatz unter Linux-Systemen ist somit auf diesem Weg nicht möglich. Da mit dieser Arbeit aufgrund des hohen Rechenaufwands aber explizit auf Hochleistungsrechnerarchitekturen abgezielt wird, die häufig Linux einsetzen, muss hierfür eine Alternative gewählt werden. Unter Linux kann MPI.NET auch genutzt werden, es muss allerdings zunächst selbst gegen die gewünschte MPI-Implementierung, zum Beispiel OpenMPI, kompiliert werden. Die erstellten Bibliotheken müssen anschließend lediglich für das Zielprojekt verfügbar gemacht werden. Um die Verwendung unter beiden Systemen zu unterstützen, wurde die MPI.NET-Referenz in der Städtesimulation in Abhängigkeit vom Zielbetriebssystem formuliert, sodass sowohl unter Linux als auch unter Windows die korrekte Variante gewählt wird. Dadurch wird sichergestellt, dass die entwickelte Software ohne Anpassungen auf beiden Systemen kompiliert und ausgeführt werden kann.

## A.2 Abfrage zu Nichtwohngebäuden

Die folgende Abfrage wird genutzt, um über overpass turbo sämtliche relevanten Nichtwohngebäude für ein Zielgebiet, in diesem Fall Jülich, aus OpenStreetMap zu erhalten.

```
area[name="Jülich"]->.searchArea;
(
node["amenity"~"^(dentist|doctors|clinic|pub|biergarten|school|cine
↪ ma|nightclub|dancing_school|marketplace|cafe|bar|ice_cream|musi
↪ c_school|university|theatre|pharmacy|kindergarten)$"](area.sear
↪ chArea);
node["building"~"^(supermarket|school|riding_hall|stable|university
↪ |museum|kindergarten)$"](area.searchArea);
node["healthcare"~"^(audiologist|dentist|occupational_therapist|pod
↪ iatrist|doctor|physiotherapist|psychotherapist|optometrist|phar
↪ macy|speech_therapist)$"](area.searchArea);
node["health_person:type"~"^(physician)$"](area.searchArea);
node["leisure"~"^(horse_riding|swimming_pool)$"](area.searchArea);
node["office"~"^(physician)$"](area.searchArea);
node["shop"~"^(mall|supermarket|general)$"](area.searchArea);
node["tourism"~"^(museum)$"](area.searchArea);
); out;
```

## A.3 Modell für optimale Apothekenstandorte

Das folgende Modell wurde für die Standortplanung der Apotheken verwendet.

$$\min \quad \bar{z} - \underline{z} \tag{A.1a}$$

$$\text{s.t.} \quad \sum_{j \in J} x_{ij} = 1 \qquad \forall i \in I \tag{A.1b}$$

$$\sum_{i \in I} d_i x_{ij} \leq Q_j y_j \qquad \forall j \in J \tag{A.1c}$$

$$\sum_{k \in J^{>} ij} xik + y_j \leq 1 \qquad \forall i \in I,\ j \in J \tag{A.1d}$$

$$x_{ij} \leq y_j \qquad \forall i \in I,\ j \in J \tag{A.1e}$$

$$\sum_{j \in J} y_j = 6 \tag{A.1f}$$

$$\underline{z} \leq \sum_{i \in I} d_i x_{ij} \leq \bar{z} \qquad \forall j \in J \tag{A.1g}$$

$$x_{ij} \in 0, 1 \qquad \forall i \in I,\ j \in J \tag{A.1h}$$

$$y_j \in 0, 1 \qquad \forall j \in J \tag{A.1i}$$

$$\underline{z}, \bar{z} \in \mathbb{R}_{\geq 0} \tag{A.1j}$$

Die Kostenfunktion (A.1a) minimiert dabei den Unterschied im Bedarf, den die einzelnen Apotheken erfüllen. Bedingung (A.1b) sorgt dafür, dass alle Kunden nur jeweils eine Apotheke aufsuchen. Bedingung (A.1c) garantiert die Einhaltung der Kapazität jeder Apotheke. Die Kapazität wurde in diesem Fall anhand der in den vorgegebenen Öffnungszeiten maximal bedienbaren Kunden abgeschätzt. Bedingung (A.1d) sorgt dafür, dass alle Kunden ihren jeweils nächstgelegenen geöffneten Apothekenstandort wählen. Die für das hier beschriebene Szenario festgelegte Anzahl eröffneter Apotheken wird durch (A.1f) eingehalten. Bedingung (A.1g) legt die obere und untere Schranke des Kundenbedarfs fest, den jede Apotheke bedient. Diese Schranken werden in der Kostenfunktion benötigt. Alle übrigen Bedingungen sind redundant und nur zur Vereinfachung des Optimierungsproblems enthalten.

Das grundlegende Modell stammt von Büsing, Leitner und Wrede [72] und wurde für die Anwendung in dieser Arbeit leicht abgewandelt.

# Abbildungsverzeichnis


2.1 Überblick über die Schwerpunkte dieser Arbeit . . . . . 3

2.2 Die im Folgenden vorgestellten Verhaltensmodelle, sortiert nach ihrer Methodik . 5

2.3 Struktur des Verhaltensmodells in MASSHA, von Kamara-Esteban et al. [8] . . . 10

2.4 Überblick über den Kern des Verhaltensmodells im LoadProfileGenerator, aus [10] 13

2.5 Kombination von Land-Use-Modellen mit Transportmodellen, von [43] . . . . . 22

2.6 Bewertung aller vorgestellten Verhaltens- und Städtemodelle anhand ausgewählter Kriterien auf einer Skala von 1 (grob) bis 5 (detailliert). Zum Vergleich wird in (c) die geplante Städtesimulation entsprechend eingeordnet. . . . . 26

3.1 Wesentliche Schritte im Ablauf der Städtesimulation . . . . . 29

3.2 Alle Außer-Haus-Aktivitäten einer Person werden zu Beginn der Städtesimulation angepasst und durch Varianten für einen spezifischen Point of Interest ersetzt. Dabei sind auch mehrere gleiche Aktivitäten möglich, wenn diese an unterschiedlichen Orten ausgeübt werden. . . . . 31

3.3 Beispiel für das Routennetzwerk eines Zwei-Personen-Haushalts. Beide Haushaltsmitglieder haben eigene Arbeitsplätze und Orte für Freizeitaktivitäten, die für die jeweils andere Person nicht relevant sind. Wenn für jede Person individuell die notwendigen Routen ermittelt werden anstatt für alle zusammen, können ungenutzte Routen (hier gestrichelt dargestellt) ausgelassen werden. . . . . 33

3.4 Probleme ohne Synchronisation der Worker. Wenn Worker nicht exakt gleich schnell simulieren, kann die Plausibilität der Simulation nicht sichergestellt werden. In diesem Beispiel wäre Max in der Warteschlange vor Sarah, obwohl Sarah zeitlich früher zum Arzt gegangen ist. . . . . 35

3.5 Verteilung der Gebäude und Synchronisierung während der Simulation . . . . . 36

3.6 Wesentliche Schritte im Ablauf der Städtesimulation auf jedem Worker . . . . . 39

3.7 Nachrichten zum Austausch eines Agenten, der eine Außer-Haus-Aktivität ausführt. Jeder blaue Pfeil stellt eine Nachricht an den entsprechenden Simulator dar, jede rote Raute eine simulierte Aktivitätswahl des Agenten. . . . . 41

3.8 Darstellung der optionalen Kopplung mit einem zusätzlichen Transportmodell. Ohne Transportmodell wird in der Städtesimulation auf Distanzen in Luftlinie und vordefinierte Reisegeschwindigkeiten zurückgegriffen. . . . . 43

3.9 Simulation der Abläufe in Points of Interest mit Warteschlange . . . . . 44

3.10 Verzeichnisstruktur und Dateien eines Städteszenarios mit vier Häusern . . . . . 45

3.11 Beispiel der Zuordnung passender Locations zu ALKIS-Gebäudekategorien. Für ein Gebäude werden stets alle Einträge aus „work_locations“ und ein zufälliger Eintrag aus „non_work_locations“ übernommen. . . . . 49

3.12 UML-Klassendiagramm der zusätzlichen Interfaces und Klassen zur Darstellung der verschiedenen Aktivitätsvarianten . . . . . 52

3.13 UML-Klassendiagramm der Nachrichten- und ID-Klassen, die genutzt werden, um über Ereignisse in der Städtesimulation zu informieren . . . . . 54

4.1 Veränderung der Rechenzeiten für Initialisierung, Hauptsimulation und Postprocessing bei steigender Anzahl paralleler Worker . . . . . 56

4.2 Auswirkungen des Kommunikationsaufwands bei einer hohen Anzahl an Workern auf den Rechenaufwand für MPI-Funktionen (a) und die Simulationsgeschwindigkeit (b) . . . . . 57

4.3 Vergleich der Bevölkerungsstruktur innerhalb des generierten Jülich-Szenarios mit Daten aus dem Regionalprofil für die Stadt Jülich und dem Zensus 2011, anhand von Geschlecht (a), Alter (b), Beschäftigungsstatus (c) und Beschäftigungsstatus kombiniert mit Geschlecht (d). Die Vergleichsdaten vom Zensus 2011 enthalten keine Angaben zum Alter oder Beschäftigungsstatus, daher wird in (b), (c) und (d) nur die Gesamtanzahl dargestellt. . . . . 59

4.4 Übersicht über die im Szenario enthaltenen Nichtwohngebäude in ganz Jülich (a) und in einem Ausschnitt der Jülicher Innenstadt (b). Jeder Kreis steht dabei für eine Gruppe an Points of Interest, wobei die Zahl im Kreis die Anzahl angibt. . . 60

4.5 Anzahl der verschiedenen Point-of-Interest-Typen, sortiert nach Häufigkeit . . . . 60

4.6 Durchschnittliche Zeit in Stunden, die pro Tag mit bestimmten Aktivitäten zugebracht wird, für die HETUS-Zeitverwendungsdaten und die Städtesimulation im Vergleich. Aktivitäten mit geringerem zeitlichem Umfang wurden zusammengefasst. 62

4.7 Gestapelte Wahrscheinlichkeitskurven für die HETUS-Daten (a) und die Städtesimulation (b), sowie die Differenz dazwischen (c) . . . . . 64

4.8 Heatmaps mit den Indikatoren aus ETHOS.ActivityAssure zum Vergleich der HETUS-Daten mit den Simulationsergebnissen, in normaler (a) und skalierter (b) Variante. Zur besseren Lesbarkeit wurden alle Indikatoren auf den einheitlichen Wertebereich $[0, 1]$ transformiert, mit exakter Übereinstimmung beim Wert $0$. . . 65

4.9 Wahrscheinlichkeitskurve für die Aktivität Fernsehen, in den Validierungsdaten und den Ergebnissen der Städtesimulation . . . . . 66

4.10 Wahrscheinlichkeitskurven für die Aktivitäten Bildung (a) und persönliche Pflege oder Körperpflege (b), in den Validierungsdaten und den Ergebnissen der Städtesimulation . . . . . 66

4.11 Verteilungen der Wiederholhäufigkeit (a) und Dauer (b) von Fernsehaktivitäten . 67

4.12 Wahrscheinlichkeitskurve (a), Häufigkeit (b) und Dauer (c) von Schlaf-Aktivitäten pro Tag, Vergleich der Validierungsdaten (blau) mit den Ergebnissen der Städtesimulation (rot) . . . . . 67

4.13 Wahrscheinlichkeitskurve (a), Häufigkeit (b) und Dauer (c) von Arbeitsaktivitäten pro Tag, Vergleich der Validierungsdaten (blau) mit den Ergebnissen der Städtesimulation (rot) . . . . . 68

4.14 Verteilungen der Wiederholhäufigkeit (a) und Dauer (b) von Bildungsaktivitäten . 69

4.15 Durchschnittliche Zeit in Stunden, die pro Tag mit bestimmten Aktivitäten zugebracht wird, einzeln für verschiedene Personen- und Tageskategorien. Jede Kategorie wird jeweils für die Zeitverwendungsdaten und die Städtesimulation gezeigt. Bedeutung der Abkürzungen: ♀=weiblich, ♂=männlich, ue=arbeitslos, s=studierend, r=Ruhestand, ft=Vollzeit, pt=Teilzeit, wd=Arbeitstag, rd=freier Tag . . . 70

4.16 Anteil der Zeit, die korrekt auf die Aktivitätskategorien verteilt wird, für jeden Profiltyp. Bedeutung der Abkürzungen: ♀=weiblich, ♂=männlich, ue=arbeitslos, s=studierend, r=Ruhestand, ft=Vollzeit, pt=Teilzeit, wd=Arbeitstag, rd=freier Tag 72

4.17 Normaler (a) und skalierter (b) Bias für jede Aktivität in jeder Aktivitätsprofilkategorie . . . . . 72

4.18 Wahrscheinlichkeitskurve (a), Häufigkeit (b) und Dauer (c) von Schlaf für vollzeitbeschäftigte Frauen an freien Tagen . . . . . 73

4.19 Normale (a) und skalierte (b) Wasserstein-Distanz für jede Aktivität in jeder Aktivitätsprofilkategorie . . . . . 74

4.20 Wahrscheinlichkeitskurve für die Aktivität Fernsehen bei männlichen Studenten an Arbeitstagen . . . . . 74

4.21 Wahrscheinlichkeitskurve (a), Häufigkeit (b) und Dauer (c) von Schlaf bei Rentnern 75

4.22 Wahrscheinlichkeitskurven (a) und Verteilungen der Dauer (b) von Schlaf bei teilzeitbeschäftigten Frauen . . . . . 75

4.23 Gestapelte Wahrscheinlichkeitskurven für die HETUS-Daten (a) und die Städtesimulation (b) für vollzeitbeschäftigte Frauen an Arbeitstagen . . . . . 76

4.24 Verteilung der Anzahl an Reisen zu verschiedenen Zielorten in der Städtesimulation 77

4.25 Häufigkeitsverteilung der Reisedistanzen nach Zielort-Kategorie . . . . . 78

4.26 Häufigkeitsverteilung der Reisedistanzen für alle Bewohner von Jülich . . . . . 78

4.27 Gesamtverbrauch und entsprechende Durchschnittslast jedes einzelnen Hauses (a) und Haushalts (b) im Simulationszeitraum, absteigend sortiert . . . . . 79

4.28 Durchschnittliche Jahresverbräuche nach Haushaltstyp. Die Farbe zeigt die Anzahl Personen im entsprechenden Haushaltstyp an. . . . . 80

4.29 Generiertes elektrisches Lastprofil für Jülich für ein gesamtes Jahr . . . . . 81

4.30 Die Last der gesamten Stadt als Jahresdauerlinie, im Vergleich zu einem entsprechend skalierten H25-Standardlastprofil. Jede Kurve gibt an, für wie lange welches Lastniveau insgesamt bestanden hat. . . . . 82

4.31 Statistik-Profile aller simulierten Häuser mit (a) und ohne (b) Maximum zur besseren Lesbarkeit. Zusätzlich wird jeweils das H25-Standardlastprofil dargestellt. . 83

4.32 Drei Gleichzeitigkeitskurven mit zufällig gewählter Reihenfolge der Haushalte . . 85

4.33 Simulierter Stromverbrauch aller Häuser in Jülich in einem Jahr . . . . . 86

4.34 Anzahl Personen (a) und der durchschnittliche Energieverbrauch pro Person (b) für jedes Haus in Jülich . . . . . 87

4.35 Standardabweichung der Lastprofile aller Häuser in Jülich . . . . . 87

4.36 Alle Häuserlasten in Jülich vom Mittwoch, 22.01.2020, um 19:14 Uhr, dem Zeitpunkt der maximalen Gesamtstädtelast . . . . . 87
4.37 Vergleich der Lastkurven in Szenarien mit 3 kW und 11 kW Ladeleistung für Elektroautos . . . . . 88
4.38 Existierende Apothekenstandorte in Jülich. Die Farbe zeigt die Anzahl der Kundenbesuche im Simulationszeitraum an. . . . . 90
4.39 Rasterdiagramme zur Anzahl anwesender (a) und ankommender (b) Kunden in einer der zentral gelegenen Apotheken, sowie ankommende Kunden in der abgelegenen Apotheke in Koslar (c) und die Verteilung der täglichen Besuchszahlen in einer zentralen Apotheke (d). . . . . 91
4.40 Verteilung der Wartezeit zu verschiedenen Uhrzeiten, für eine zentrale (a, c) und eine abgelegene Apotheke (b, d) . . . . . 92
4.41 Anzahl nicht bedienter Kunden pro Tag für eine zentrale (a) und eine abgelegene Apotheke (b) . . . . . 93
4.42 Die verbliebenen drei Apotheken im Szenario mit reduzierten Standorten (a) und die Anwesenheitszahl in der zentral gelegenen Apotheke (b) . . . . . 93
4.43 Anzahl nicht bedienter Kunden pro Tag (a) und Wartezeit (b) für die zentrale Apotheke im Szenario mit reduzierten Standorten . . . . . 94
4.44 Anzahl ankommender (a) und insgesamt anwesender Kunden (b) für eine zentrale Apotheke im Szenario ohne parallele Bedienung . . . . . 95
4.45 Anzahl nicht bedienter Kunden pro Tag (a) und Wartezeit (b) für eine zentrale Apotheke im Szenario ohne parallele Bedienung . . . . . 95
4.46 Die vom Standortplanungsmodell gewählten Apothekenstandorte (a) und die Anzahl anwesender Kunden im neu eröffneten Apothekenstandort (b) . . . . . 96
4.47 Anzahl täglicher Kundenbesuche in der Filiale in Koslar im Basisszenario (a) und des neu gewählten Standorts im optimierten Szenario (b) . . . . . 97

# Tabellenverzeichnis

2.1 Vergleich der vorgestellten Verhaltensmodelle . . . . . 16

3.1 Klassifizierung aller Locations im LoadProfileGenerator für eine geeignete Zuordnung zu Gebäuden. Die Kategorie gibt an, auf welche Art von Gebäude sich die jeweilige Location bezieht. Für die Locations werden die englischen Originalbezeichnungen verwendet. Locations, die sich innerhalb des eigenen Haushalts befinden, wie zum Beispiel das Schlafzimmer, werden nicht zugewiesen und sind daher nicht aufgeführt. . . . . 47

3.2 Für die Verkehrsmittelwahl ergeben sich im Hinblick auf die verfügbaren Informationen neun mögliche Fälle. Sieben davon werden berücksichtigt, sodass Haushalte mit und ohne Auto sowie mit und ohne ÖPNV-Abonnement unterschieden werden können. Für die Nutzung des ÖPNV ist ein Abonnement nicht erforderlich, es erhöht aber die Wahrscheinlichkeit. Da unter Umständen nicht bekannt ist, ob ein Haushalt über ein ÖPNV-Abonnement verfügt, gibt es für diesen Fall eine zusammengefasste Kategorie. Außerdem gibt es eine weitere Kategorie für den Fall, dass zu beiden Kriterien keine Informationen vorliegen. . . . . 51

4.1 Vergleich der offiziellen Bevölkerungsstatistiken für Jülich aus dem Zensus 2011 mit den Daten in BUILDA und dem daraus generierten Szenario . . . . . 58

4.2 Statistiken des Stadtlastprofils im Vergleich zur Funktion $F$ aus Gleichung (4.1). Alle Werte sind in Megawatt angegeben. Dabei ist $E(X)$ der Mittelwert und $\sigma(X)$ die Standardabweichung von $X$. . . . . 84

4.3 Anteil nicht bedienter Kunden und durchschnittliche Wartezeit in jedem der vorgestellten Szenarien . . . . . 97

## Literaturverzeichnis


[1] *Urbanisierung und Städte weltweit*. did-206126-1. Statista, 2025. URL: https://de.statista.com/statistik/studie/id/206126/dokument/urbanisierung-und-verstaedterung/ (besucht am 28. 10. 2025).

[2] Ranjan K. Bose, Energy Sector Management Assistance Programme und World Bank, Hrsg. *Energy Efficient Cities: Assessment Tools and Benchmarking Practices*. Washington, D.C: World Bank, 2010. 227 S. ISBN: 978-0-8213-8104-5 978-0-8213-8309-4.

[3] Eric Bonabeau. „Agent-Based Modeling: Methods and Techniques for Simulating Human Systems". In: *Proceedings of the National Academy of Sciences of the United States of America* 99 (Suppl 3 14. Mai 2002), S. 7280–7287. ISSN: 0027-8424. DOI: 10.1073/pnas.082080899. PMID: 12011407. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC128598/ (besucht am 17. 06. 2022).

[4] Charles Macal und Michael North. „Introductory Tutorial: Agent-based Modeling and Simulation". In: *Proceedings of the Winter Simulation Conference 2014*. Proceedings of the Winter Simulation Conference 2014. Dez. 2014, S. 6–20. DOI: 10.1109/WSC.2014.7019874.

[5] Tina Balke und Nigel Gilbert. „How Do Agents Make Decisions? A Survey". In: *Journal of Artificial Societies and Social Simulation* 17.4 (2014), S. 13. ISSN: 1460-7425. DOI: 10.18564/jasss.2687. URL: https://www.jasss.org/17/4/13.html.

[6] Jeetika Malik et al. „A Level-of-Details Framework for Representing Occupant Behavior in Agent-Based Models". In: *Automation in Construction* 139 (Juli 2022), S. 104290. ISSN: 09265805. DOI: 10.1016/j.autcon.2022.104290. URL: https://linkinghub.elsevier.com/retrieve/pii/S0926580522001637 (besucht am 08. 02. 2023).

[7] Shuqin Chen et al. „Definition of Occupant Behavior in Residential Buildings and Its Application to Behavior Analysis in Case Studies". In: *Energy and Buildings* 104 (1. Okt. 2015), S. 1–13. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2015.06.075. URL: https://www.sciencedirect.com/science/article/pii/S0378778815301146 (besucht am 30. 10. 2022).

[8] O. Kamara-Esteban et al. „MASSHA: An Agent-Based Approach for Human Activity Simulation in Intelligent Environments". In: *Pervasive and Mobile Computing* 40 (2017), S. 279–300. ISSN: 1574-1192. DOI: 10.1016/j.pmcj.2017.07.007.

[9] Quentin Reynaud et al. „Using Time Use Surveys in Multi Agent Based Simulations of Human Activity." In: *ICAART (1)*. 2017, S. 67–77.

[10] Noah Pflugradt et al. „LoadProfileGenerator: An Agent-Based Behavior Simulation for Generating Residential Load Profiles". In: *Journal of Open Source Software* 7.71 (25. März 2022), S. 3574. ISSN: 2475-9066. DOI: 10.21105/joss.03574. URL: https://joss.theoj.org/papers/10.21105/joss.03574 (besucht am 04. 07. 2022).

[11] Tao Zhang, Peer-Olaf Siebers und Uwe Aickelin. „Modelling Electricity Consumption in Office Buildings: An Agent Based Approach". In: *Energy and Buildings* 43.10 (1. Okt. 2011), S. 2882–2892. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2011.07.007. URL: https://www.sciencedirect.com/science/article/pii/S0378778811003094 (besucht am 09. 11. 2022).

[12] Mengda Jia et al. „A Systematic Development and Validation Approach to a Novel Agent-Based Modeling of Occupant Behaviors in Commercial Buildings“. In: *Energy and Buildings* 199 (2019), S. 352–367.

[13] Jianli Chen et al. „Stochastic Simulation of Occupant-Driven Energy Use in a Bottom-up Residential Building Stock Model“. In: *Applied Energy* 325 (1. Nov. 2022), S. 119890. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2022.119890. URL: https://www.sciencedirect.com/science/article/pii/S0306261922011540 (besucht am 14. 10. 2022).

[14] Felix Ziegler et al. „A Probabilistic Modelling Approach for Residential Load Profiles“. In: *Zenodo* (27. Feb. 2020). DOI: 10.5281/ZENODO.3689339. URL: https://zenodo.org/record/3689339 (besucht am 26. 05. 2022).

[15] Ian Richardson et al. „Domestic Electricity Use: A High-Resolution Energy Demand Model“. In: *Energy and Buildings* 42.10 (1. Okt. 2010), S. 1878–1887. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2010.05.023. URL: https://www.sciencedirect.com/science/article/pii/S0378778810001854 (besucht am 14. 10. 2022).

[16] Ruben Baetens und Dirk Saelens. „Modelling Uncertainty in District Energy Simulations by Stochastic Residential Occupant Behaviour“. In: *Journal of Building Performance Simulation* 9.4 (2016), S. 431–447.

[17] Tianzhen Hong et al. „Ten Questions Concerning Occupant Behavior in Buildings: The Big Picture“. In: *Building and Environment* 114 (1. März 2017), S. 518–530. ISSN: 0360-1323. DOI: 10.1016/j.buildenv.2016.12.006. URL: https://www.sciencedirect.com/science/article/pii/S0360132316304851 (besucht am 26. 05. 2022).

[18] Yuan Jin et al. „Appliance Use Behavior Modelling and Evaluation in Residential Buildings: A Case Study of Television Energy Use“. In: *Building Simulation* 13.4 (1. Aug. 2020), S. 787–801. ISSN: 1996-8744. DOI: 10.1007/s12273-020-0648-8. URL: https://doi.org/10.1007/s12273-020-0648-8 (besucht am 26. 05. 2022).

[19] Salvatore Carlucci et al. „Modeling Occupant Behavior in Buildings“. In: *Building and Environment* 174 (1. Mai 2020), S. 106768. ISSN: 0360-1323. DOI: 10.1016/j.buildenv.2020.106768. URL: https://www.sciencedirect.com/science/article/pii/S0360132320301268 (besucht am 27. 05. 2022).

[20] Gabriel Happle, Jimeno A. Fonseca und Arno Schlueter. „A Review on Occupant Behavior in Urban Building Energy Models“. In: *Energy and Buildings* 174 (1. Sep. 2018), S. 276–292. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2018.06.030. URL: https://www.sciencedirect.com/science/article/pii/S0378778817333583 (besucht am 26. 05. 2022).

[21] Bruna Faitão Balvedi, Enedir Ghisi und Roberto Lamberts. „A Review of Occupant Behaviour in Residential Buildings“. In: *Energy and Buildings* 174 (1. Sep. 2018), S. 495–505. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2018.06.049. URL: https://www.sciencedirect.com/science/article/pii/S0378778817336447 (besucht am 26. 05. 2022).

[22] Jun Li et al. „Development and Improvement of Occupant Behavior Models towards Realistic Building Performance Simulation: A Review“. In: *Sustainable Cities and Society* 50 (1. Okt. 2019), S. 101685. ISSN: 2210-6707. DOI: 10.1016/j.scs.2019.101685. URL: https:

//www.sciencedirect.com/science/article/pii/S2210670719303774 (besucht am 01. 11. 2022).

[23] Isabella Gaetani, Pieter-Jan Hoes und Jan L. M. Hensen. „Occupant Behavior in Building Energy Simulation: Towards a Fit-for-Purpose Modeling Strategy". In: *Energy and Buildings* 121 (1. Juni 2016), S. 188–204. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2016.03.038. URL: https://www.sciencedirect.com/science/article/pii/S0378778816301918 (besucht am 05. 04. 2023).

[24] Lukas G. Swan und V. Ismet Ugursal. „Modeling of End-Use Energy Consumption in the Residential Sector: A Review of Modeling Techniques". In: *Renewable and Sustainable Energy Reviews* 13.8 (1. Okt. 2009), S. 1819–1835. ISSN: 1364-0321. DOI: 10.1016/j.rser.2008.09.033. URL: https://www.sciencedirect.com/science/article/pii/S1364032108001949 (besucht am 21. 10. 2022).

[25] S. Helal et al. „Persim - Simulator for Human Activities in Pervasive Spaces". In: Proceedings - 2011 7th International Conference on Intelligent Environments, IE 2011. 2011, S. 192–199. ISBN: 978-0-7695-4452-6. DOI: 10.1109/IE.2011.34.

[26] Flora D. Salim et al. „Modelling Urban-Scale Occupant Behaviour, Mobility, and Energy in Buildings: A Survey". In: *Building and Environment* 183 (1. Okt. 2020), S. 106964. ISSN: 0360-1323. DOI: 10.1016/j.buildenv.2020.106964. URL: https://www.sciencedirect.com/science/article/pii/S0360132320303231 (besucht am 01. 11. 2022).

[27] Ian Richardson, Murray Thomson und David Infield. „A High-Resolution Domestic Building Occupancy Model for Energy Demand Simulations". In: *Energy and Buildings* 40.8 (1. Jan. 2008), S. 1560–1566. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2008.02.006. URL: https://www.sciencedirect.com/science/article/pii/S0378778808000467 (besucht am 14. 10. 2022).

[28] Q. Reynaud et al. „Multi-Agent Simulation of Human Activity for Residential Electrical Load and Demand Forecasting and Its Application to Electric Mobility". In: *CIRED 2021 - The 26th International Conference and Exhibition on Electricity Distribution*. CIRED 2021 - The 26th International Conference and Exhibition on Electricity Distribution. Bd. 2021. Sep. 2021, S. 2808–2812. DOI: 10.1049/icp.2021.1512.

[29] Mathieu Schumann et al. „Multi-Agent Based Simulation of Human Activity for Building and Urban Scale Assessment of Residential Load Curves and Energy Use". In: *Building Simulation 2021 Conference*. 2021.

[30] Jeetika Malik et al. „Ten Questions Concerning Agent-Based Modeling of Occupant Behavior for Energy and Environmental Performance of Buildings". In: *Building and Environment* 217 (1. Juni 2022), S. 109016. ISSN: 0360-1323. DOI: 10.1016/j.buildenv.2022.109016. URL: https://www.sciencedirect.com/science/article/pii/S036013232200258X (besucht am 26. 04. 2023).

[31] Christoph F. Reinhart und Carlos Cerezo Davila. „Urban Building Energy Modeling – A Review of a Nascent Field". In: *Building and Environment* 97 (15. Feb. 2016), S. 196–202. ISSN: 0360-1323. DOI: 10.1016/j.buildenv.2015.12.001. URL: https://www.sciencedirect.com/science/article/pii/S0360132315003248 (besucht am 06. 02. 2023).

[32] C. Cerezo Davila, C.F. Reinhart und J.L. Bemis. „Modeling Boston: A Workflow for the Efficient Generation and Maintenance of Urban Building Energy Models from Existing Geospatial Datasets“. In: *Energy* 117 (2016), S. 237–250. ISSN: 0360-5442. DOI: 10.1016/j.energy.2016.10.057.

[33] Darren Robinson et al. „CITYSIM: Comprehensive Micro-Simulation of Resource Flows for Sustainable Urban Planning“. In: *IBPSA 2009 - International Building Performance Simulation Association 2009*. 27. Juli 2009.

[34] Darren Robinson, Urs Wilke und Frédéric Haldi. „Multi Agent Simulation of Occupants' Presence and Behaviour“. In: 16. Nov. 2011, S. 9.

[35] M. Ferrando et al. „Urban Building Energy Modeling (UBEM) Tools: A State-of-the-Art Review of Bottom-up Physics-Based Approaches“. In: *Sustainable Cities and Society* 62 (2020). ISSN: 2210-6707. DOI: 10.1016/j.scs.2020.102408.

[36] J. Keirstead, Nouri J. Samsatli und N. Shah. „Syncity: An Integrated Tool Kit for Urban Energy Systems Modelling“. In: 2009. URL: https://www.semanticscholar.org/paper/SYNCITY%3A-AN-INTEGRATED-TOOL-KIT-FOR-URBAN-ENERGY-Keirstead-Samsatli/6ac033651f959cbe935c5037278bd324367a7105 (besucht am 03. 02. 2023).

[37] James Keirstead und Aruna Sivakumar. „Using Activity-Based Modeling to Simulate Urban Resource Demands at High Spatial and Temporal Resolutions“. In: *Journal of Industrial Ecology* 16.6 (2012), S. 889–900. ISSN: 1530-9290. DOI: 10.1111/j.1530-9290.2012.00486.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1530-9290.2012.00486.x (besucht am 08. 02. 2023).

[38] Ana L. C. Bazzan und Franziska Klügl. „A Review on Agent-Based Technology for Traffic and Transportation“. In: *The Knowledge Engineering Review* 29.3 (Juni 2014), S. 375–403. ISSN: 0269-8889, 1469-8005. DOI: 10.1017/S0269888913000118. URL: https://www.cambridge.org/core/journals/knowledge-engineering-review/article/abs/review-on-agentbased-technology-for-traffic-and-transportation/00947E41817CA7336E0A130CBF353D16# (besucht am 05. 04. 2023).

[39] Andreas Horni, David Charypar und Kay W. Axhausen. „Variability in Transport Microsimulations Investigated With the Multi-Agent Transport Simulation MATSim“. In: *Arbeitsberichte Verkehrs-und Raumplanung* (Aug. 2011), 20 p. DOI: 10.3929/ETHZ-A-006686329. HDL: 20.500.11850/38291. URL: http://hdl.handle.net/20.500.11850/38291 (besucht am 31. 01. 2023).

[40] Kay W. Axhausen, Andreas Horni und Kai Nagel, Hrsg. *The Multi-Agent Transport Simulation MATSim*. Ubiquity Press, 2016. DOI: 10.5334/baw. URL: https://library.oapen.org/handle/20.500.12657/32162 (besucht am 27. 05. 2022).

[41] Pablo Alvarez Lopez et al. „Microscopic Traffic Simulation Using SUMO“. In: *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. 2018 21st International Conference on Intelligent Transportation Systems (ITSC). Nov. 2018, S. 2575–2582. DOI: 10.1109/ITSC.2018.8569938.

[42] James Keirstead, Mark Jennings und Aruna Sivakumar. „A Review of Urban Energy System Models: Approaches, Challenges and Opportunities“. In: *Renewable and Sustainable Energy Reviews* 16.6 (1. Aug. 2012), S. 3847–3866. ISSN: 1364-0321. DOI: 10.1016/j.rser.

2012.02.047. URL: https://www.sciencedirect.com/science/article/pii/S1364032112001414 (besucht am 08. 02. 2023).

[43] *Travel Forecasting Resource*. URL: https://tfresource.org (besucht am 02. 06. 2023).

[44] Michael Wegener. „Overview of Land Use Transport Models“. In: *Handbook of Transport Geography and Spatial Systems*. Bd. 5. Emerald Group Publishing Limited, 2004, S. 127–146.

[45] Paul Waddell. „UrbanSim: Modeling Urban Development for Land Use, Transportation, and Environmental Planning“. In: *Journal of the American planning association* 68.3 (2002), S. 297–314.

[46] *Urban Data Science Toolkit*. UrbanSim. URL: https://github.com/UDST (besucht am 21. 08. 2025).

[47] Aneesa Awaludin und David Chen. „UrbanSim Parallel Programming Capstone Paper“. In: *Computer Science and Engineering 481E course report* (2007).

[48] Dirk Strauch et al. „Linking Transport and Land Use Planning: The Microscopic Dynamic Simulation Model ILUMASS“. In: *Geodynamics*. CRC Press, 2004, S. 319–336.

[49] K. Spiekermann und M. Wegener. „Multi-Level Urban Models: Integration across Space, Time and Policies“. In: *Journal of Transport and Land Use* 11.1 (2018), S. 67–81. ISSN: 1938-7849. DOI: 10.5198/jtlu.2018.1185.

[50] Itzhak Benenson und Paul M. Torrens. *Geosimulation: Automata-Based Modeling of Urban Phenomena*. Hoboken, NJ: John Wiley & Sons, 2004. 287 S. ISBN: 978-0-470-84349-9.

[51] Ander Pijoan. „Geoworldsim: A Time-Asynchronous, Distributed and Intelligent Environment Based Geosimulation Platform“. http://purl.org/dc/dcmitype/Text. Universidad de Deusto, 2018. URL: https://documat.unirioja.es/servlet/tesis?codigo=250117 (besucht am 20. 02. 2023).

[52] F. Johari et al. „Urban Building Energy Modeling: State of the Art and Future Prospects“. In: *Renewable and Sustainable Energy Reviews* 128 (2020). ISSN: 1364-0321. DOI: 10.1016/j.rser.2020.109902.

[53] Kristina Dabrock et al. „Leveraging Machine Learning to Generate a Unified and Complete Building Height Dataset for Germany“. In: *Energy and AI* 17 (1. Sep. 2024), S. 100408. ISSN: 2666-5468. DOI: 10.1016/j.egyai.2024.100408. URL: https://www.sciencedirect.com/science/article/pii/S2666546824000740 (besucht am 10. 03. 2025).

[54] Laura Alessandretti et al. „Evidence for a Conserved Quantity in Human Mobility“. In: *Nature Human Behaviour* 2.7 (Juli 2018), S. 485–491. ISSN: 2397-3374. DOI: 10.1038/s41562-018-0364-x. URL: https://www.nature.com/articles/s41562-018-0364-x (besucht am 17. 12. 2024).

[55] Douglas Gregor, Andrew Lumsdaine und Ben Martin. *MPI.NET*. Version 1.4.0. Microsoft, 23. März 2019. URL: https://github.com/Microsoft/MPI.NET (besucht am 10. 12. 2025).

[56] Noah Pflugradt und David Neuroth. *LoadProfileGenerator*. Forschungszentrum Jülich - Jülich Systems Analysis (ICE-2), 2021. URL: https://github.com/FZJ-IEK3-VSA/LoadProfileGenerator (besucht am 10. 12. 2025).

[57] David Neuroth. *ETHOS.CityScenarioGenerator*. Forschungszentrum Jülich - Jülich Systems Analysis (ICE-2), 1. Dez. 2025. URL: https://github.com/FZJ-IEK3-VSA/CityScenarioGenerator (besucht am 02. 12. 2025).

[58] Leo Strobel und Marco Pruckner. „OMOD: An Open-Source Tool for Creating Disaggregated Mobility Demand Based on OpenStreetMap". In: *Computers, Environment and Urban Systems* 106 (1. Dez. 2023), S. 102029. ISSN: 0198-9715. DOI: 10.1016/j.compenvurbsys.2023.102029. URL: https://www.sciencedirect.com/science/article/pii/S0198971523000923 (besucht am 19. 09. 2023).

[59] Gene M. Amdahl. „Validity of the Single Processor Approach to Achieving Large Scale Computing Capabilities". In: *Proceedings of the April 18-20, 1967, Spring Joint Computer Conference*. AFIPS '67 (Spring). New York, NY, USA: Association for Computing Machinery, 18. Apr. 1967, S. 483–485. ISBN: 978-1-4503-7895-6. DOI: 10.1145/1465482.1465560. URL: https://dl.acm.org/doi/10.1145/1465482.1465560 (besucht am 24. 10. 2025).

[60] John L. Gustafson. „Reevaluating Amdahl's Law". In: *Commun. ACM* 31.5 (1. Mai 1988), S. 532–533. ISSN: 0001-0782. DOI: 10.1145/42411.42415. URL: https://dl.acm.org/doi/10.1145/42411.42415 (besucht am 24. 10. 2025).

[61] Statistisches Bundesamt. *Gemeindeverzeichnis-Sonderveröffentlichung*. destatis.de, Mai 2013. URL: https://www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/Administrativ/zensus-gemeinden.html (besucht am 10. 10. 2025).

[62] *Gewerbeanmeldungen Jülich*. 3. Nov. 2016. URL: https://offenedaten.kdvz.nrw/dataset/d16-gewerbeanmeldungen-j%C3%BClich (besucht am 25. 11. 2025).

[63] David Neuroth et al. *ETHOS.ActivityAssure*. Forschungszentrum Jülich, Jülich Systems Analysis (IEK-3), 16. Apr. 2024. URL: https://github.com/FZJ-IEK3-VSA/ActivityAssure (besucht am 16. 04. 2024).

[64] David Neuroth et al. „ETHOS.ActivityAssure—An Open-Source Validation Framework for Synthetic European Activity Profiles". In: *Energy and Buildings* 326 (1. Jan. 2025), S. 115036. ISSN: 0378-7788. DOI: 10.1016/j.enbuild.2024.115036. URL: https://www.sciencedirect.com/science/article/pii/S0378778824011526 (besucht am 18. 11. 2024).

[65] Eurostat. *Harmonised European Time Use Survey*. Microdata. Version 2010. S-CIRCABC, 2010. URL: https://ec.europa.eu/eurostat/web/microdata/harmonised-european-time-use-surveys (besucht am 14. 03. 2024).

[66] David Neuroth et al. *ETHOS.ActivityAssure Data Set*. Version 1.0.0. Zenodo, 22. März 2024. DOI: 10.5281/zenodo.10835209. URL: https://zenodo.org/records/10835209 (besucht am 16. 04. 2024).

[67] destatis. *Energy Consumption*. 4. Sep. 2020. URL: https://www.destatis.de/EN/Themes/Society-Environment/Environment/Material-Energy-Flows/Tables/electricity-consumption-households.html (besucht am 05. 12. 2025).

[68] BDEW. *Standardlastprofile Strom*. URL: https://www.bdew.de/energie/standardlastprofile-strom/ (besucht am 24. 09. 2025).

[69] Ifd Allensbach. *Bevölkerung in Deutschland nach Häufigkeit des Besuchens von Apotheken in den Jahren 2022 bis 2024*. https://de.statista.com/statistik/daten/stu-

die/1052863/umfrage/umfrage-zu-der-haeufigkeit-des-besuchs-von-apotheken/, 19. Juni 2024. URL: https://de.statista.com/statistik/daten/studie/1052863/umfrage/umfrage-zu-der-haeufigkeit-des-besuchs-von-apotheken/ (besucht am 02. 11. 2025).

[70] Lucas A. Berenbrok et al. „Evaluation of Frequency of Encounters With Primary Care Physicians vs Visits to Community Pharmacies Among Medicare Beneficiaries". In: *JAMA Network Open* 3.7 (15. Juli 2020), e209132. ISSN: 2574-3805. DOI: 10.1001/jamanetworkopen.2020.9132. URL: https://doi.org/10.1001/jamanetworkopen.2020.9132 (besucht am 22. 10. 2025).

[71] Samantha N Valliant et al. „Pharmacists as Accessible Health Care Providers: Quantifying the Opportunity". In: *Journal of Managed Care & Specialty Pharmacy* 28.1 (Jan. 2022), 10.18553/jmcp.2022.28.1.85. ISSN: 2376-0540. DOI: 10.18553/jmcp.2022.28.1.85. PMID: 34949110. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC8890748/ (besucht am 22. 10. 2025).

[72] Christina Büsing, Markus Leitner und Sophia Wrede. „Cover-based inequalities for the single-source capacitated facility location problem with customer preferences". In: *Computers & Operations Research* 182 (1. Okt. 2025), S. 107082. ISSN: 0305-0548. DOI: 10.1016/j.cor.2025.107082. URL: https://www.sciencedirect.com/science/article/pii/S0305054825001108 (besucht am 04. 11. 2025).