

## ENERGIEEFFIZIENZ DURCH BIG-DATA-ANALYSEN IN DER GEBÄUDEAUTOMATION

L. Lauss<sup>1</sup>, J. Mehnert<sup>2</sup>, J. Lebert<sup>2</sup>, T. Auer<sup>1</sup> und S. Plesser<sup>3</sup>

<sup>1</sup>Lehrstuhl für Gebäudetechnologie und klimagerechtes Bauen, TU München, Deutschland

<sup>2</sup>Institut für Gebäude- und Solartechnik, TU Braunschweig, Deutschland

<sup>3</sup>Steinbeis-Innovationszentrum energie+, Braunschweig, Deutschland

E-Mail: lukas.lauss@tum.de; j.mehnert@tu-braunschweig.de

### KURZFASSUNG

Eine große Anzahl von Bestandsgebäuden verfügt über Gebäudeautomationssysteme, die Sensor- und Aktordaten kontinuierlich erfassen. Diese Betriebsdaten spiegeln sowohl die Funktionsweise des Gebäudes als Gesamtes, als auch einzelner gebäudetechnischer Anlagen wider und fallen im Gebäudebetrieb massenhaft (Big Data) an. Im Rahmen dieses Forschungsvorhabens wurde ein algorithmischer Big-Data-Baukasten als informationstechnologisches Rahmenwerk entwickelt, mit dessen Hilfe sich verschiedene Fragestellungen rund um die Analyse von Betriebsdaten effizient beantworten lassen. Der innovative Schwerpunkt liegt in der Entwicklung von skalierbaren und automatisierten Lösungen zur Minimierung von Performance Gaps. Durch die experimentelle Anwendung der Big-Data-Analysemethoden konnte die Identifikation von Potenzialen zur energetischen Betriebsoptimierung im Gebäudebestand aufgezeigt werden.

### ABSTRACT

A large number of existing buildings have building automation and control systems that continuously record sensor and actuator data. These operating data reflect the functionality of the building as a whole as well as of individual building technology systems and are generated in large quantities (Big Data) during building operation. Within the framework of this research project, an algorithmic Big Data construction kit was developed as an information technology framework, with the help of which various questions concerning the analysis of operational data can be answered efficiently. The innovative focus lies in the development of scalable and automated solutions to minimize performance gaps. The experimental application of the Big Data analysis methods showed the identification of potentials for energy operation optimization in existing buildings.

### EINLEITUNG

Die ambitionierten Klimaziele der Bundesregierung lauten, bis 2050 einen nahezu klimaneutralen Gebäudebestand zu realisieren (Umweltbundesamt, 2016). Für diese Umsetzung müssen wirksame Strategien und Maßnahmen entwickelt werden. Die bauphysikalischen Optimierungsmöglichkeiten, welche die Transmissions- und Lüftungswärmeverluste betreffen sind größtenteils ausgeschöpft (Correal et al., 2014). Die Stärkung der Energieeffizienz gewinnt mehr an Bedeutung und die technische Gebäudeausrüstung (TGA) in Kombination mit der Gebäudeautomation (GA) wird priorisiert betrachtet. Resultierend gilt es Energieeinsparpotenziale in Gebäuden mit einer Gebäudeautomation ausfindig zu machen und zu nutzen. Dementsprechend bedeutend sind optimal ausgeführte und im Betrieb fehlerfrei funktionierende Anlagen, die dem planerischen Soll entsprechen und anschließend optimiert werden könnten. Folglich trägt eine energieeffiziente Gebäudebetriebsführung dazu bei, die Energieeffizienz zu steigern und die Treibhausgas-Emissionen zu senken.

In einem Großteil der Gebäude ist dies nicht der Fall, wodurch ein hohes Energieeinsparpotential gegeben ist. In diesem Zusammenhang tritt die Problematik auf, dass eine Leistungslücke zwischen dem vorhergesagten Energieverbrauch in der Entwurfsphase von Gebäuden und dem Energieverbrauch dieser sich in Betrieb befindlichen Gebäuden auftritt (Fisch et al., 2007, Réhault et al., 2014). Etabliert hat sich hierbei die Beschreibung des Performance Gaps, also der Abweichung zwischen dem technisch erzielbarem *Soll* und dem tatsächlich auftretenden *Ist*. Aus energetischer Sicht werden die Einbußen bezogen auf den Energiebedarf mit 5 % bis 30 % beziffert (Wilde, 2014). Die Ursachen für den erhöhten Energieverbrauch liegen meist bei Betriebsfehlern in der Gebäude- und Anlagentechnik sowie deren Automation (Fütterer et al., 2017).

Gleichzeitig sind Qualitätssicherungsprozesse im Gebäudebetrieb wenig etabliert und das Zusammenspiel zahlreicher Akteure wie z. B. Objektbetreibern, Dienstleistern und Nutzern erschwert diesen Prozess. Zusätzlich stellt die stetig zunehmende Komplexität der Gebäude und die Wechselwirkungen der Funktionen von Anlagen ein Hindernis für einen minimalen Energieverbrauch dar.

Ein Indikator hierfür ist die rapide wachsende Vernetzung im Zeitalter des Internet-of-Things (IoT) (Mauerer et al., 2018). Werden Anlagen- und Funktionsstörungen sowie Betriebsfehler im Lebenszyklus von Gebäuden nicht identifiziert und demnach nicht beseitigt, ist ein energieeffizienter Gebäude- und Anlagenbetrieb nicht gewährleistet. Das Ziel der Energieeffizienz im Gebäudebetrieb wird dann nur minderwertig und teilweise erfüllt.

## MOTIVATION

Gebäude und Anlagen werden heute zunehmend komplexer. Die Vernetzung im Zeitalter des IoT nimmt stetig zu und die Anforderungen an Energieeffizienz und Nutzerkomfort steigen zunehmend (Waide et al., 2014). Die verschiedenen Einflüsse, wie zum Beispiel Witterung und Nutzerverhalten, machen den Einsatz zusätzlicher Mess-, Steuer- und Regelungstechnik (MSRT) zur Erhaltung eines effizienten Anlagenbetriebs unabdingbar. Aus diesem Grund verfügen moderne Nicht-Wohngebäude (NWG) in der Regel über eine Gebäudeleittechnik (GLT), welche die Sensor- und Aktordaten kontinuierlich erfasst. Diese GLT-Systeme generieren große Datenmengen, welche den Betrieb des Gebäudes widerspiegeln. Dadurch wird Aufschluss über die Funktionsweise des Gebäudes als Gesamtes, aber auch über einzelne Anlagen und Gewerke gegeben. Die heutigen Daten einer Leittechnik können Messwerte von Sensoren und Aktoren, z. B. Temperaturen, Volumenströme, Drücke oder Ventilstellungen in unterschiedlichen zeitlichen Auflösungen (häufig sekundlich, minütlich, viertelstündlich, etc.) erfassen und speichern. Zudem werden Fehlermeldungen von Steuergeräten und Aktoren, wie z. B. Pumpen oder Ventilatoren erfasst.

Diese Daten fallen heute im Gebäudebetrieb massenhaft an, jedoch mangelt es an Strategien diese automatisiert auszuwerten und somit Rückschlüsse ziehen zu können. Insbesondere die Datenmenge und -vielfalt („Big Data“) stellt hohe Anforderungen an eine nachgelagerte Auswertung großer Betriebszeiträume. Im Sinne einer umfassenden, energetischen und funktionstechnischen Überprüfung von technischen Anlagen müssen größere Zeiträume ausgewertet und viele der vorhandenen Messkanäle miteinander kombiniert werden. Dabei werden schnell Datenmengen erreicht, für die klassische Auswerteverfahren und konventionelle Infrastrukturen an ihre Grenzen stoßen. Des Weiteren wird sich die bereits heutzutage schon große Datenmenge, in Zukunft noch weiter erhöhen und verdichten.

Für die Analyse von Betriebs- und Automationsdaten eignen sich Big Data Technologien bzw. Methoden der künstlichen Intelligenz. Im Folgenden werden daher die Rahmenwerkzeuge und Prozesse, welche zur Umsetzung dieser Analysen, bezogen auf Betriebsdaten die mittels der Gebäudeautomation aufgezeichnet wurden, detailliert beschrieben.

Zu diesem Zweck stehen die Bereiche Datenaufbereitung, Datenklassifikation, Merkmalsanalyse und Fehlererkennung gebäudetechnischer Anlagen im Fokus der Untersuchungen. Das Ziel dieses Forschungsvorhabens ist somit die Entwicklung von skalierbaren und automatisierten Lösungen zur Unterstützung der Minimierung von Performance Gaps, die durch diverse Mängel während der Planung, Errichtung und dem Betrieb von Gebäuden entstehen. Dadurch soll bestmöglicher Nutzerkomfort bei gleichzeitig hoher Energieeffizienz gewährleistet und ein Beitrag zur energetischen Betriebsoptimierung des Gebäudebestandes geleistet werden.

## METHODIK VON BIG DATA ANALYSEN

Für die Analyse von Betriebs- und Automationsdaten werden Big Data Technologien ein großes Potential, insbesondere in der Zusammenarbeit mit unterschiedlichen Akteuren, welche für eine Sicherstellung der Betriebsperformance notwendig sind eingeräumt (Bitkom, 2012). Im Rahmen des Projektes wurde dafür ein algorithmischer Big Data Baukasten als informationstechnologisches Rahmenwerk entwickelt, mit dessen Hilfe sich verschiedene Fragestellungen rund um die Analyse von Betriebsdaten effizient beantworten lassen (Reiß et al., 2019). Um Prozesse handhabbar zu gestalten, wurde dieser Baukasten nach dem „Teile & Herrsche-Prinzip“ aufgebaut (siehe *Abbildung 1*). Für die Umsetzung wurden sowohl reale Betriebsdaten, die über die Gebäudeautomation von zwei großen NWG gewonnen wurden, als auch mittels dynamischer Gebäudesimulationsmodelle synthetisch erzeugte Betriebsdaten verwendet. Die Anwendung des Big Data Baukastens wurde im Rahmen der Forschungsarbeit in zahlreichen Experimenten aus den Bereichen Datenaufbereitung, Datenklassifikation, Merkmalsanalyse und Fehlererkennung der TGA demonstriert. Ziel des Forschungsvorhabens ist somit die Entwicklung von skalierbaren und automatisierten Lösungen zur energetischen Betriebsoptimierung des Gebäudebestandes.

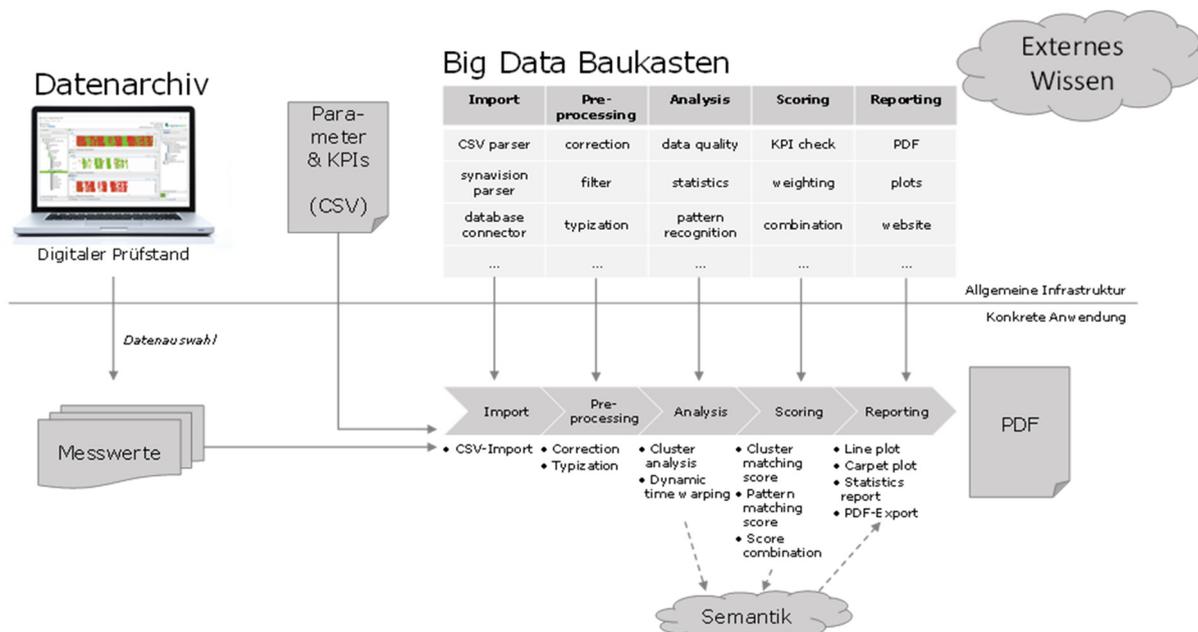


Abbildung 1: Überblick der Gesamtarchitektur im Projekt

Innerhalb des Baukastens sind Fragestellungen gezielt in algorithmische Lösungen zerlegt. Auf diese Weise wird sichergestellt, dass komplexe Prozesse, die darüber hinaus stets automatisierbar sein müssen, handhabbar werden. Der Baukasten setzt sich aus einzelnen Baukastenelementen zusammen, welche sich den Kategorien Daten-Konnektivität, Vorverarbeitung, Analyse, Bewertung und Reporting zuordnen lassen (Entwicklungsebene 1 in *Abbildung 2*). Kombiniert und orchestriert werden einzelne Elemente zu Anwendungsszenarien zusammengestellt und bilden somit wiederverwendbare Module (Ebene 2). Innerhalb der Anwendungsebene (Ebene 3) erfolgt eine automatisierte Wissenserkenntnis und Merkmalsextraktion, indem die Verfahren auf reale Betriebsdaten angewandt werden.

Die Entwicklung von Machine-Learning Modellen erfordert ein Training der Algorithmen mit einer großen Menge zu diesem Zwecke geeigneter Daten. Zu diesem Zweck wurden im Rahmen des Projektes neben realen Betriebsdaten, die über die Gebäudeautomation von zwei großen NWG gewonnen wurden, auch Simulationsdaten verwendet. Diese synthetischen Betriebsdaten wurden auf Basis von Gebäude und Anlagensimulationsmodellen erzeugt. Die detaillierte Modellierung von Gebäude- und Anlagentechnik inklusive Heizung, Lüftung, Kühlung und Beleuchtung mit allen relevanten Regelalgorithmen ermöglicht die realitätsnahe Abbildung des thermischen Verhaltens der Gebäude- und Anlagentechnik sowie deren Wechselwirkungen. Die Simulationsdaten repräsentieren synthetische Betriebsdaten sowohl zum korrekten als auch zum fehlerhaften Gebäude- und Anlagenbetrieb. Dies wird erreicht durch die gezielte Implementierung definierter Betriebsfehler.

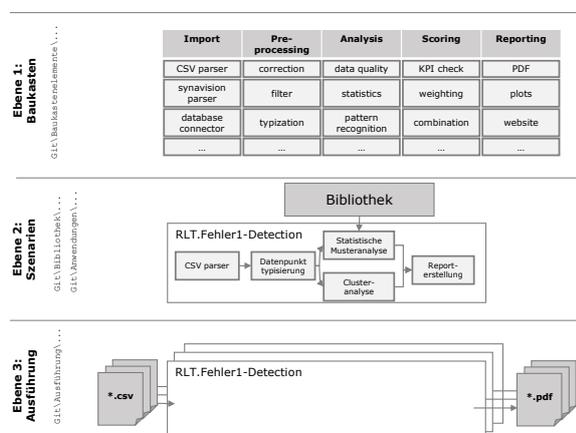


Abbildung 2: Entwicklungsebenen mithilfe des algorithmischen Big Data Baukastens

Im Vergleich zu realen Betriebsdaten bietet diese Methodik den großen Vorteil, dass hinsichtlich der Verfügbarkeit und Vollständigkeit von Automationsdaten keine Einschränkungen bestehen. Durch die Simulationen kann die große Bandbreite an möglichen Szenarien im Gebäude- und Anlagenbetrieb abgebildet sowie unterschiedliche Merkmale in die virtuellen Datenquellen implementiert werden. Im Kontext der Simulationen wurde auch untersucht, welche Effekte einzelne Parameter auf den Gebäude- und Anlagenbetrieb haben und inwieweit sich Betriebsfehler auf den Energiebedarf auswirken können.

## EXPERIMENTELLE ANWENDUNG VON BIG DATA ANALYSEN

Die Analysen wurden an Betriebsdaten aus realen Nicht-Wohngebäuden durchgeführt, welche durch die in der Gebäudeautomation integrierte Speicherfunktion abrufbar waren. Diese integrierte Speicherung der Betriebsdaten ist in den häufigsten NWG gegeben. Andernfalls sind Datenlogger als Software- oder Hardwarelösung zum Nachrüsten verfügbar, z. B. erreichbar über das BACnet Protokoll (Kranz, 2013). Da die Schnittstellen hier jedoch sehr heterogen und proprietär sind, wurde der *Digitale Prüfstand* der synavision GmbH eingesetzt, um die Verbindung unterschiedlicher Gebäude zu vereinheitlichen und zu vereinfachen (siehe *Abbildung 1*). Neben den realen Betriebsdaten wurde der Betrieb gebäudetechnischer Anlagen auch mittels dynamischer Simulationen modelliert, die den Vorteil bieten zielführend und schnell unterschiedliche (Fehler-)szenarien abzubilden.

### **Entwicklung synthetischer Betriebsdaten**

Im Rahmen dieses Projektes werden virtuelle Daten durch die simulationstechnische Abbildung von typischen Bürogebäuden und der dazugehörigen Anlagentechnik sowie Automation erzeugt. Durch die dynamisch-thermischen Gebäude- und Anlagensimulationen in *IDA Indoor Climate and Energy (ICE)* werden unterschiedliche Varianten zur Energieerzeugung, -verteilung, und -übergabe erstellt. Typische Anwendungsbereiche des Modellierungs- und Simulationswerkzeuges *IDA-ICE* sind die Simulation von Innenraumklima und den dafür benötigten Energieströmen in Gebäuden, basierend auf der Modellierung und Simulation der Gebäudesubstanz, als auch der Anlagentechnik. Zusätzlich zu dem thermischen Verhalten der Gebäude und deren Räume werden auch die MSRT einschließlich der Hydraulik für Heizung, Lüftung, Kühlung sowie die Beleuchtung durch Kunst- und Tageslicht simuliert. Ein parameterbasierter Modellaufbau ermöglicht dabei eine breite Variation verschiedener Einflussfaktoren auf den Gebäude- und Anlagenbetrieb und erlaubt Untersuchungen, die im realen Gebäudekontext schwer oder nicht darstellbar sind. Dementsprechend ergeben sich durch den Einsatz von synthetischen Betriebsdaten aufgrund der möglichen Bandbreite an Parametervariationen Vorteile für das Training von Machine-Learning Algorithmen gegenüber realen Betriebsdaten. Durch diesen simulationsbasierten Prozess werden synthetische Betriebsdaten für den korrekten und fehlerfreien Betrieb sowie mittels implementierter Betriebsfehler für Fehlerfälle generiert.

Beim Aufbau des Gebäudemodells wurde ein typisches Bürogebäude, welches den Durchschnitt dieser NWG-Kategorie in Deutschland abbildet, modelliert. Das Referenz-Gebäudemodell basiert auf den normativen Grundlagen eines Referenz-

Büroraums, der durch 20 Zonen (à 20.25 m<sup>2</sup> für ein 2 Personenbüro mit hoher Belegungsdichte nach DIN V 18599-10) und unterschiedlicher Orientierung auf Gebäudemastab (405.0 m<sup>2</sup> NF) skaliert wird. Der weitere Aufbau des Modells orientiert sich ebenfalls an den Randbedingungen des Nutzungsprofils für Gruppenbüros nach DIN V 18599-10. Im Fokus der gesamtheitlichen Untersuchungen stehen die Anlagentechnik und deren Automation, weshalb Bürogebäude mit einem entsprechenden technischen Ausrüstungsgrad gewählt werden. Qualitätssicherungsmaßnahmen für die Energieeffizienz von Gebäuden wie technisches Monitoring oder umfangreiche Funktionen der GA werden, erst ab einer entsprechenden Baualtersklasse eingesetzt. Deshalb orientiert sich der Energiestandard an einer entsprechenden Baualtersklasse und für das Gebäudemodell werden die bauphysikalischen Kennwerte der EnEV 2009 für die typischen Konstruktionsaufbauten zu Grunde gelegt. Die Bereitstellung der Heizenergie zur Raumwärmeversorgung und zur Konditionierung der Zuluft in den Heizregistern der Raumluftheizung (RLT) Anlage erfolgt mittels Gasbrennwertkessel. Sowohl die Wärmeerzeuger als auch die Wärmeabgabesysteme in Form von statischen Heizkörpern wurden auf Basis der Heizlastberechnung nach DIN EN 12831 dimensioniert. Die mechanische Lüftung ist als Teilklimaanlage (TKA) modelliert und ist mit den thermodynamischen Luftbehandlungsfunktionen *heizen, kühlen, entfeuchten* für die Konditionierung zur Zu- und Abluftversorgung der Büroräume verantwortlich. Die modellierte RLT-Anlage orientiert sich an der TGA im Bereich Raumluftheizung des VDI-Mustergebäudes nach VDI 6009-1 und ist daher als TKA ausgeführt. Nach (Werner, 2008) werden in Büro- und Verwaltungsgebäuden Teilklimaanlagen dreimal so häufig wie Vollklimaanlagen eingesetzt.

Die fehlerhaften Datensätze werden anhand konkreter Betriebsfehler (siehe Abschnitt „Fehlerkatalog“) wie z. B. Abweichung von definierter Soll-Zulufttemperatur, defektes Regelventil, Über-/unterschreitung der Betriebszeiten, Offset Zuluft-Temperaturfühler generiert. Unter fehlerhaft wird hierbei, eine Abweichung vom definierten Normbetrieb, verstanden. Diese Abweichungen können in unterschiedlicher Charakteristik (z. B. Zeitpunkt, Intensität, Dauer, etc.) auftreten, sodass diese dynamische Variabilität auch bei der Modellierung berücksichtigt ist. Durch diese Variation unsicherer Randbedingungen entsteht eine Vielzahl an Simulationen mit Datensätzen, die den fehlerhaften Betrieb repräsentieren. Unter anderem werden die synthetischen Daten zum Entwickeln und Trainieren von Machine-Learning Algorithmen für den Big Data Baukasten verwendet. In diesem Kontext wurden die simulationstechnisch

abzubildenden Datenpunkte an die in der Praxis verfügbaren Datenpunkte angepasst. Diese Anpassung vermeidet ein unerwünschtes Trainieren der Algorithmen auf Abhängigkeiten zwischen virtuellen Datenpunkten aus den Simulationen und realen Datenpunkten aus der GLT.

### Anwendungsbeispiel: Klassifikation und semantische Anreicherung

Das Wissen um die Semantik eines Datenpunktes ist die Eingangsvoraussetzung für weitere automatisierte Verfahrensschritte und somit der erste elementare Baustein in der Vorverarbeitung.

Ein vielversprechendes Verfahren zur Klassifikation unbekannter Sensor- und Aktordaten ist das *Deep Learning* (Goodfellow et al., 2016). Die Idee hierbei ist es, ein Modell auf Basis bekannter und bereits klassifizierter Daten zu trainieren, um anschließend das trainierte Modell für die Klassifikation unbekannter Sensor- und Aktordaten zu nutzen. Deep Learning basiert dabei auf *künstlichen neuronalen Netzen* (Kruse et al., 2015), die wiederum aus sogenannten *Neuronen* und deren Verbindungen zueinander bestehen. Sie bilden ähnlich dem menschlichen Gehirn, je nach erlerntem Kontext, gewichtete Verbindungen zwischen den Neuronen aus. *Abbildung 3* zeigt ein neuronales Netz, das die viertelstündigen Messwerte eines Tagesausschnitts einer Zeitreihe verarbeitet und ausgabeseitig eine Klassifikation in die vier untersuchten physikalischen Einheiten vornimmt. Hierbei handelt es sich um die Demonstration einer Grobklassifikation. Die Gesamtheit der Datenpunkte, die für die Analyse der Performance notwendig sind, wird in eine Klassifikationsstruktur eingeteilt. Mithilfe dieser Taxonomie können Datenpunkte exakt beschrieben werden. Dafür wurden die Klassen *Einheit*, *Medium*, *System*, *Subsystem*, *Komponente* und *Ort* eingeführt. Neben einer Grobklassifikation wurde auch eine Feinklassifikation durchgeführt, die zusätzlich zur Einheitenklassifikation eine Unterscheidung nach dem Ort (Einbauort) vorsieht. Am Beispiel des Datenpunkttyps Temperatur bedeutet dies, dass das neuronale Netz trainiert wird, um Vorlauf- von Rücklauftemperaturen oder Luft- von Wassertemperaturen zu unterscheiden. Im Rahmen des Forschungsdemonstrators wurden verschiedene neuronale Netze zur Klassifikation erprobt. Dabei wurden drei verschiedene Netzwerkarchitekturen des Deep Learnings implementiert.

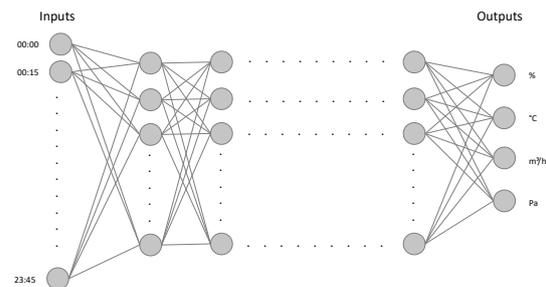


Abbildung 3: Schematische Darstellung eines neuronalen Netzes zur Klassifikation von Datenpunkten nach Einheiten

Das beschriebene Beispiel zur Grobklassifikation wurde mit ca. 48.000 wöchentlichen Samples aus RLT-Anlagen des realen Gebäudebetriebs trainiert (Lebert, 2019). Die Probe und Klassifikation erfolgte mit jeweils ca. 4.000 wöchentlichen Samples aus Realdatensätzen, die jedoch nicht Teil des Trainings waren.

Die Deep Learning Verfahren erzielten während der durchgeführten Experimente bereits in einfachen Konfigurationen gute bis sehr gute Ergebnisse (siehe Abschnitt „Ergebnisse und Diskussion“). Um die Eignung abschließend zu beurteilen, müssten jedoch weitere Experimente mit deutlich größeren Datenmengen durchgeführt sowie darüber hinaus Optimierungen der Verfahren implementiert werden, um eine Überanpassung der Modelle durch zu hohen oder zu geringen Variantenreichtum zu verhindern. Größere Datenmengen können beispielsweise über synthetische Daten durch eine dynamische Simulation erreicht werden. Der Einsatz der synthetisch erzeugten Zeitreihendaten zum Training von maschinellen Lernalgorithmen wurde zudem anhand des Random Forest Algorithmus geprobt und evaluiert.

Die Zeitreihenverläufe, aus denen nach dem Abruf aus der Gebäudeautomation zu Anfang lediglich Werte und Zeitstempel als Informationen verfügbar waren liegen nun typisiert und mit semantischen Information angereichert vor. Diese sind, nach einer für Gebäudebetriebsanalysen geeigneten Taxonomie, klassifiziert und somit für eine Weiterverarbeitung mittels computergestützter Methoden verarbeitbar. Weitere Anwendungsbeispiele, die im Zuge des Projektes an realen Anlagen praxisnah erprobt wurden, werden im Folgenden kurz erläutert.

#### Fehlerkatalog:

Dieses kategorische Bewertungssystem dient als Erfassungsmatrix. Dabei werden Fehler nach Kategorien der DIN EN 15232 gewerkegetreu zugeordnet. Zudem können den katalogisierten Fehlern energetische Auswirkungen und die Häufigkeit des Auftretens zugeordnet werden. Mithilfe dessen kann ein Prioritätsindex abgeleitet werden. Weiterhin werden Fehlern entsprechende Handlungsempfehlungen zugeordnet, die wiederum

die Grundlage für einen entsprechenden Feedback-Mechanismus darstellt (Wussler, 2018).

#### Merkmalsidentifikation:

Dabei werden typische Merkmale zum Beispiel als Aktivitätsprofil für eine Schnittmenge an Daten durch eine Normwoche beschrieben. In diesem Zuge werden viele Wochen überlagert und für jeden Zeitpunkt einer Normwoche eine Aktivitätsmaßzahl abgeleitet.

#### Automatisierte Fehlererkennung:

Mittels der Komposition eines sogenannten *150%-Modells* einer RLT-Anlage können die Betriebsmodi eines Systems und deren Zulässigkeit bewertet werden. Dabei wird das System RLT in seine  $n$  Subsysteme (Vorerhitzer, Kühler, Befeuchter, etc.) zerlegt. Ein Subsystem kann dabei einen Zustand "An" oder "Aus" annehmen. Das System (z. B. RLT) kann demnach  $2^n$  Systemzustände annehmen, die jeweils nur "Zulässig" oder "Unzulässig" sein müssen. Zusammen mit den Baukastenelementen Datenpunktktyperkennung und Merkmalsidentifikation können die Systemzustände und somit ein fehlerhafter Betrieb für jeden Zeitpunkt automatisiert hergeleitet werden.

## ERGEBNISSE & DISKUSSION

### **Datenpunktktyperkennung durch Algorithmen des maschinellen Lernens**

Aus den Untersuchungen zur Datenpunktktyperkennung durch Deep Learning Verfahren ist hervorgegangen, dass die Effektivität der Klassifikation signifikant von der Anzahl der zur Verfügung stehenden Samples je Klasse abhängig ist. Je mehr Datenpunkte für das Training zur Verfügung stehen desto bessere Ergebnisse liefern die Algorithmen. Des Weiteren wurde festgestellt, dass die Ausgeglichenheit des Trainingsdatensatzes von großer Bedeutung ist. Dies lässt sich daran festmachen, dass die Raten korrekt klassifizierter Klassen bei häufig auftretenden Datenpunkten höher sind als die der selten vorkommenden Klassen. Insgesamt wurden die häufig vertretenen Datenpunkte der Klassifikationsebene „Einheit“ mit einer True-Positive-Rate (TPR) von 70-97 % erkannt, während Klassen mit geringerer Datengrundlage zum Trainieren eine TPR von 40 % definitiv richtig klassifizierten Datentypen erreicht haben.

Das Training der neuronalen Netze zur Unterscheidung von Temperaturen, die an verschiedenen Orten innerhalb von Anlagen gemessen werden, wie z. B. Zuluft- und Ablufttemperaturen, erzielte korrekt vorhergesagte Klassifikationsraten von maximal 69 %. Daher werden für die Erkennung der Datenpunkte in dieser feineren Aufschlüsselung empfohlen die Verfahren weiter zu optimieren, um die Ergebnisse zu steigern. Potential zur Optimierung bietet beispielsweise die semantische Dekomposition der vorhandenen Informationen. Durch diese Dekomposition, beispielsweise des Zeitstempels in

Wochentag, Monat, Jahreszeit oder etwa des Sonnenauf- und untergangs wird der semantisch nutzbare Kontext für die Analyse zusätzlich angereichert. Die natürliche Struktur der Deep Learning Verfahren eignet sich hierfür in besonderem Maße, da sich die Klassifikation von abstrakten Merkmalen hin zu konkreten Sensor- und Aktortypen verfeinern lässt. Im Rahmen des Forschungsprojektes wurde die grundsätzliche Eignung der Verfahren demonstriert.

### **Energetische Auswirkungen von Betriebsfehlern**

Neben der Bereitstellung der virtuellen Datenquellen werden die Simulationen auch zur Analyse von Stellschrauben im Gebäudebetrieb und zur Bewertung der Auswirkungen von Betriebsfehlern auf den Energiebedarf verwendet. Hierbei soll der Einfluss einzelner Betriebsparameter auf den Gebäude- und Anlagenbetrieb untersucht werden. Als Zielgrößen werden die jeweiligen Komponenten der RLT-Anlagen bzw. deren in Wechselwirkung stehenden Energieströme gewählt. Dazu zählen der durch die Wärmerückgewinnung (WRG) rückgewonnene Anteil an Wärme ( $Q_{WRG}$ ), dem Lufterhitzer zugeführte Wärme ( $Q_{VE}$ ), dem Luftkühler abgeführte Wärme ( $Q_{KUE}$ ), dem Luftnacherhitzer zugeführte Wärme ( $Q_{NE}$ ), die aufzuwendende Ventilatorarbeit ( $W_{Vent}$ ), der raumintegrierten Heizung zugeführte Wärme ( $Q_{ZH}$ ) für die Raumwärmeversorgung und die gesamte eingebrachte Wärme ( $Q_H$ ). Anhand dieser Zielgrößen sollen einerseits die direkten Auswirkungen auf Bauteilebene und andererseits die indirekten Auswirkungen auf sekundäre Größen beurteilt werden.

Wie bereits in den vorangegangenen Beschreibungen erläutert, kann der korrekte Betrieb eine große Bandbreite an Betriebszuständen annehmen. In diesem Zusammenhang soll der Einfluss von unterschiedlichen Zulufttemperaturen auf die definierten Zielgrößen untersucht werden. Die untenstehende *Abbildung 4* zeigt daher beispielhaft die unterschiedlichen Nutzenergiebedarfe für den korrekten Betrieb der Referenz-Teilklimaanlage in Kombination mit der Variation der Zulufttemperatur in 0.5 K Schritten zwischen 18.0 °C und 22.0 °C. Diese Analyse kann aber auch als Energieeinsparpotential aufgrund von fehlerhaften Zulufttemperaturen angesehen werden und steht somit in Bezug zu dem Betriebsfehler „Unter-/Überschreitung der Soll-Zulufttemperatur“.

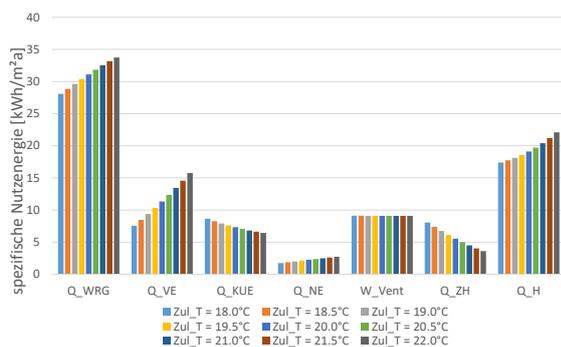


Abbildung 4: Einfluss der Zulufttemperatur auf den Nutzenergiebedarf

Erwartungsgemäß zeigt sich bei den definierten Randbedingungen der Simulationen die charakteristische Ausprägung aller Zielgrößen, mit Ausnahme der gleichbleibenden Ventilatorarbeit, eine Veränderung durch die Variation der Zulufttemperatur. Es lässt sich feststellen, dass mit steigender Zulufttemperatur der durch die WRG rückgewonnene Anteil an Wärme, die dem Luftvor- und Luftnacherhitzer zugeführte Wärme sowie die gesamt eingebrachte Wärme ansteigt. Im Gegensatz dazu sinkt sowohl die vom Luftkühler abgeführte Wärme als auch die durch die dezentrale Heizung zugeführte Wärme für die Raumwärmeversorgung. Das bedeutet, umso höher das von der RLT-Anlage bereitgestellte Temperaturniveau, desto weniger Energie müssen die Wärmeabgabesysteme in den Räumen aufwenden. Im Hinblick auf die gesamte Wärmemenge kann festgehalten werden, dass bei einer Erhöhung bzw. Senkung der Zulufttemperatur pro Kelvin im Bereich 18.0 – 22.0 °C eine Einsparung bzw. Mehrverbrauch von ca. 20 % bzw. 4.7 kWh/m²a Nutzenergie vorliegt.

Auf Basis dieser Betrachtung und den simulierten Energiebedarfen, welche grundsätzlich den korrekten Gebäude- und Anlagenbetrieb darstellen, kann bereits abgeleitet werden, wie groß der Einfluss auf den Performance Gap auch ohne Betriebsfehler sein kann. Darauf aufbauend rücken nun die Auswirkungen von Betriebsfehlern in den Fokus der Untersuchungen. Als exemplarischer Betriebsfehler wurde der im Realbetrieb identifizierte Betriebsfehler „gleichzeitiges heizen und kühlen“ modelliert und mit der unsicheren Randbedingung der Fehlerauftretsdauer simuliert. Bei dieser Betrachtung wurde als Referenz die Zulufttemperatur von 18.0 °C nach EnEV 2009 gewählt, um den Einfluss der Fehlerzeit untersuchen zu können. *Abbildung 5* zeigt die Ergebnisse der Zielgrößen für diesen Fehlerfall in Abhängigkeit der zeitlichen Ausprägungen für die in *Tabelle 1* aufgelisteten Fehlerzeiten und -dauern.

Tabelle 1: Zeitpunkte und Dauer der simulierten Betriebsfehler

FHL	BF-DAUER [UHRZEIT]	FHL	BF-DAUER [UHRZEIT]
_00	Di 10-12	_06	Werktags 12-13
_01	Werktags 10-12	_07	Fr. 16-18
_02	Di 10-12 & 15-17	_08	Mo, Fr 11-12 & 16-18
_03	Mo, Mi, Fr 10-12	_09	Di, Mi, Do 14-16
_04	Do 13-15	_10	Do, Fr 5-11
_05	Mi 10-17	_11	Täglich 5-18

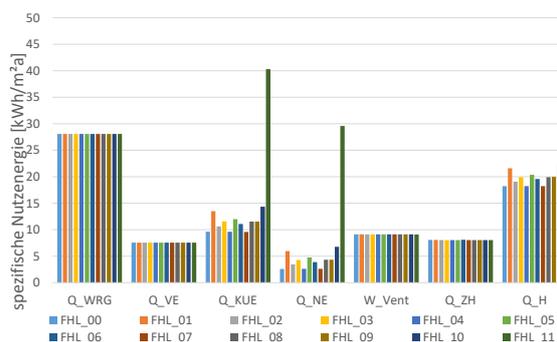


Abbildung 5: Einfluss des Betriebsfehlers „gleichzeitiges heizen und kühlen“ in Abhängigkeit der Fehlerauftretsdauer auf den Nutzenergiebedarf

Im Hinblick auf den energetischen Mehrverbrauch ist in diesem Szenario der dominierende Aspekt der Zeitraum in dem der Fehler auftritt. Grundsätzlich ergeben sich bei den Varianten an denen das gleichzeitige heizen und kühlen nur über ein paar wenige Stunden pro Tag oder Woche (alle bis auf FHL\_11) auftritt, ähnliche Ausprägungen der Zielgrößen. Wie zu erwarten, sind auf Bauteilebene bei Wärmerückgewinnung, Luftvorerhitzer, Ventilator und dezentraler Heizung trotz auftretendem Fehler keine Auswirkungen erkennbar. Im Gegensatz dazu schwanken der Energiebedarf für Luftkühler und Luftnacherhitzer sowie die resultierende gesamte eingebrachte Wärmemenge in Abhängigkeit der jeweiligen Fehlerdauer. Vergleicht man bei der Zielgröße Q\_H die beiden Varianten mit den kürzesten (FHL\_00/ FHL\_04/FHL\_07) und den längsten (FHL\_01/FHL\_10) Fehlerzeiten so ergeben sich Abweichungen von ca. 5.0 kWh/m²a Nutzenergie bzw. 22 %. Daraus kann abgeleitet werden, dass bei den definierten Randbedingungen der Simulationen ein energetischer Mehrverbrauch von ca. 1.0 kWh/m²a bzw. 4.4 % pro Stunde in denen der Fehler auftritt, verursacht wird. Grundsätzlich können die Auftretswahrscheinlichkeit und der Zeitraum von Betriebsfehlern stark variieren, da Fehler teilweise an vereinzelt Tagen, mehrmals im Monat oder

lediglich einmal im Jahr auftreten können. In diesem Zusammenhang gibt es nur sehr begrenzt belastbare Literatur und keine zufriedenstellende Datengrundlage. Jedoch ist es als ein sehr realistisches Szenario anzusehen, dass im Fehlerfall der Betriebsfehler dauerhaft vorhanden ist. Aus diesem Grund wird im Rahmen dieser Arbeit auch die Variante FHL\_11 eingeführt. Hierbei wird davon ausgegangen, dass sobald die RLT-Anlage in Betrieb ist, der Betriebsfehler auftritt. Diese Betrachtung resultiert in einem extrem stark ansteigenden Energiebedarf bei Kühlregister, Nacherhitzer und gesamter eingebrachter Wärmemenge im Vergleich zum korrekten Betrieb der Basisvariante mit Zulufttemperatur von 18.0 °C. Wird die Betriebszeit mit der Fehlerzeit gleichgesetzt, so ergibt sich bei Kühlbedarf ein energetischer Mehrverbrauch von ca. Faktor 3, bei Heizwärmebedarf für das Nachheizregister ca. Faktor 5 und bei Heizwärmebedarf gesamt ungefähr Faktor 2. Diese Analysen zeigen eindeutig den hohen Einfluss des *Technical Gap* auf den gesamtheitlichen *Performance Gap* und untermauern die Notwendigkeit nach energetischen Qualitätssicherungsmaßnahmen im Gebäude- und Anlagenbetrieb.

### ZUSAMMENFASSUNG & AUSBLICK

Mit der Baukastenstruktur steht ein geeignetes Rahmenwerkzeug bereit, um die Performance von Gebäuden auf kleinteilige Fragestellungen nach dem Teile & Herrsche-Prinzip herunterzubrechen. Mittels des Baukastens wurden Methoden und Algorithmen aus dem Bereich Big Data im Kontext des Gebäudebetriebs in konkrete Anwendungsszenarien überführt und praxisnah ausgeführt. Automatisierbare Methoden sind demnach ein großer Mehrwert, um ein technisches Monitoring digitalisiert zu unterstützen und den Mangel an ExpertInnen aufzufangen. Dafür bedarf es allerdings einen hohen Aufwand in der Implementierung, um auf viele technische Anlagen in Gebäuden generisch gültig zu sein.

Ein weiterer Schritt ist es neben Aussagen zur reinen Betriebsqualität, also die Auswirkungen auf Komfort und Energie, auch quantitative Bewertungen, also die Häufigkeit des Auftretens zu berücksichtigen.

Der Einsatz von Simulationsmodellen dient als virtuelle Datenbasis für die Generierung synthetischer Betriebsdaten für das Training von Verfahren des Maschinellen Lernens sowie deren Validationszwecken. Zusätzlich wurden durch die physikalischen und technischen Wirkzusammenhänge die Potentiale zur Energieeinsparung bzw. energetischen Mehrverbrauch durch die Variation unterschiedlicher korrekter Betriebszustände und Betriebsfehler quantifiziert.

Für weiterführende Forschungsaktivitäten in diesem Kontext sollte die Verwendung von stochastischen Gebäudemodellen unter Berücksichtigung von unsicheren Randbedingungen in Betracht gezogen werden. Hierbei könnte mit Monte-Carlo-Simulationen die große Bandbreite und dynamische Variabilität der Input-Parameter optimiert abgebildet werden. Dieser Prozess sollte sich wiederum positiv auf die Qualität der synthetischen Betriebsdaten, für das Entwickeln und Trainieren von Machine-Learning Algorithmen, auswirken.

### DANKSAGUNG

Dieses Vorhaben wurde mit Mitteln des Bundesministeriums für Wirtschaft und Energie (BMWi) unter dem Förderkennzeichen 03ET1426 gemeinsam im Konsortium mit dem Lehrstuhl für Gebäudetechnologie und klimagerechtes Bauen der Technischen Universität München, Lehrstuhl für Software Engineering der RWTH Aachen Lehrstuhl, WILO SE, synavision GmbH und dem siz energie+ durchgeführt.

### LITERATUR

- Bitkom Bundesverband Informationswirtschaft 2012. Telekommunikation und neue Medien e.V., Leitfaden: Management von Big-Data-Projekten, Bitkom.
- Correal, D., Gröschel, J., Möller, T., Neumann, S., Ruwinski, R. 2014. Energiemanagement für Bürogebäude. Leitfaden zur praktischen Einführung des Energiemanagements für Bürogebäude. Bericht Climate-KIC.
- Fisch, N., Plesser, S., et al. 2007. EVA - Evaluierung von Energiekonzepten für Bürogebäude. Schlussbericht, Forschungsprojekt, Institut für Gebäude und Solartechnik, Technische Universität Braunschweig
- Fütterer, J., Schild, T., Müller, D. 2017, Gebäudeautomationssysteme in der Praxis, Whitepaper ; RWTH-EBC 2017-001, Aachen, 2017, <http://dx.doi.org/10.18154/RWTH-2017-05671>
- Goodfellow I., Bengio Y., Courville A. 2016. Deep Learning, The MIT Press.
- Kranz H. R. 2013. BACnet Gebäudeautomation 1.12: Grundlagen in deutscher Sprache, Cci-Buch Beraten + Planen, 3. Aufl. cci Dialog.
- Kruse, R., Borgelt, C., Braune, C., Klawonn, F., Moewes, C., Steinbrecher M. 2015. Computational Intelligence: Eine methodische Einführung in Künstliche Neuronale Netze, Evolutionäre Algorithmen, Fuzzy-Systeme und Bayes-Netze, Computational Intelligence Springer Vieweg.

- Lebert J. 2019. Implementierung eines Algorithmus zur Klassifikation von Gebäudeautomationsdaten, Masterarbeit, TU Braunschweig.
- Mauerer J. et al. 2018. Studie Internet of Things 2019, IDG Business Media GmbH.
- Reiß, D., Mehnert, J., Plesser, S., Hannen, M. 2019. An algorithmic module toolkit to support quality management for building performance, In: E3S Web Conf. 111, S. 5002. DOI: 10.1051/e3sconf/201911105002.
- Réhault, N., Lichtenberg, G., Schmidt, F. et al. 2014. ModQS – Modellbasierte Qualitätssicherung des energetischen Gebäudebetriebs, Forschungsbericht, Fraunhofer-Institut für Solare Energiesysteme (ISE), Freiburg
- Umweltbundesamt (Hg.). 2016. Klimaneutraler Gebäudebestand 2050. Online verfügbar unter <http://www.umweltbundesamt.de/publikationen/klimaneutraler-gebäudebestand-2050>.
- Waide, P., Ure, J., Smith, G., Nordass, B. 2014. The scope for energy and CO<sub>2</sub>, savings in the EU through the use of building automation, Research Report, Waide Strategic Efficiency Ltd.
- Werner, P., Chmella-Emrich, E., Vilz, A. 2008. Folgen des Klimawandels: Gebäude und Baupraxis in Deutschland: BBR-Online-Publikation, Nr. 10/2008.
- Wilde P. de 2014. The gap between predicted and measured energy performance of buildings: A framework for investigation, *Autom. Constr.*, vol. 41, pp. 40–49.
- Wussler M. 2018. Entwicklung eines kategorischen Bewertungssystems zur technisch-wirtschaftlichen Evaluation von Betriebsfehlern gebäudetechnischer Anlagen. Masterarbeit, TU Braunschweig.