

Schlussbericht vom 14.08.2023

zu IGF-Vorhaben Nr. 21516 N

Thema

Semantische Interoperabilität von Beleuchtungssystemen, Analyse und Auswertung anhand von realen Daten zur Folgerung von Use Cases in unterschiedlichen realen Standorten

Berichtszeitraum

01.11.2020 - 28.02.2023

Forschungsvereinigung

Forschungsvereinigung Elektrotechnik beim ZVEI e.V.

Forschungseinrichtung(en)

Forschungseinrichtung 1: Technische Universität Darmstadt, Fachgebiet Adaptive lichttechnische Systeme und visuelle Verarbeitung

Forschungseinrichtung 2: Fachhochschule Dortmund, Institut für Kommunikationstechnik

Gefördert durch:

Inhaltsverzeichnis

Zusammenfassung.....	6
1. Einleitung.....	8
1.1. Technische und Wirtschaftliche Problemstellung.....	8
1.2. Wirtschaftliche Bedeutung der Forschungsergebnisse für KMU	9
1.3. Arbeitshypothese	14
1.4. Einfluss der Corona-Pandemie auf den Projektverlauf	14
1.5. Über das Projekt und die Projektpartner.....	15
2. Arbeitspakete	16
2.1. AP1 Vorbereitung der Auswertung	16
2.2. AP2 Datenmodelle und Cloudanbindung.....	18
2.3. AP3 Einheitliche Semantik	28
2.4. AP4 Use Cases.....	37
2.5. AP5 Betrieb und Wartung der Anlagen.....	64
3. Zusammenfassung und Ausblick.....	66
3.1. Anbindung an IoL Cloud.....	66
3.2. Semantik.....	66
3.3. Use Cases	66
3.4. Ausblick.....	67
4. Erläuterung zur Verwendung der Zuwendung	68
5. Ergebnistransfer in die Wirtschaft.....	68
5.1. Veröffentlichungen und Wissenstransfer	68
5.2. Durchgeführte Transfermaßnahmen (von Projektbeginn bis heute).....	69
5.3. Geplante Transfermaßnahmen (auch nach Projektende)	71
6. Literatur.....	73
7. Anhang.....	76

Abbildungsverzeichnis

Abbildung 1: Finale Struktur IoL-Plattform.	20
Abbildung 2: Komponenten des SBS.....	21
Abbildung 3 Beispiel unterschiedlicher Modellierungsstrategien.....	26
Abbildung 4 Auszug vereinheitlichtes Datenformat.....	27
Abbildung 5 Semantische Verknüpfung von Geräteinformationen im Kontext von WoT.	28
Abbildung 6: BoT der TuDa-Anlage.....	32
Abbildung 7 BoT-Beschreibung zu 3D-Rendering	33
Abbildung 8: Erweiterte BoT-Struktur	34
Abbildung 9: Übersicht der Docker-Dienste.....	36
Abbildung 10: Übersicht der Datenstruktur mit acht Standorten, die die Daten in die Datenbank schrieben, die zusätzlich erhobenen Wetter- und Sonnenaufgangsdaten, die zur Datenauswertung genutzt werden.	37
Abbildung 11: Automatisch generierte Darstellung eines Beispiel Standorts aus dem Projekt mit Fenstern, Türen, Lichtquellen, Sensoren und den Raumabmessungen.....	38
Abbildung 12: Trends und Saisonalitäten mit verschiedenen Zeitkonstanten.....	40
Abbildung 13: Pearson Korrelation von Außentemperatur, Außenluftfeuchtigkeit, Bewölkungsgrad, Temperatur, Luftfeuchtigkeit, Beleuchtungsstärke und Leistung an einem Leuchtenpunkt.....	40
Abbildung 14: Ersetzen fehlender Werte eines Temperatursensors durch Werte vom Vortag sowie einem gewichteten Mittelwert der umliegenden Sensoren.	41
Abbildung 15: Visualisierung der verwendeten LSTM-Modellarchitektur. Jeder LSTM-Block besitzt 64 Units, die mit einem Dense-Layer bestehen aus 32 Neuronen verbunden sind.	44
Abbildung 16: Hyperparameter-Suchraum für die untersuchten Vorhersage-Modelle: ARIMA, HMM, XGBoost und LSTM.	45
Abbildung 17: RMSE und MAPE-Score für die Vorhersage von Leistung, Beleuchtungsstärke und Temperatur an einem Beispiel-Standort für einen Vorhersagezeitraum von vier Stunden.	46
Abbildung 18: Modellvergleich für die Vorhersage der Beleuchtungsstärke mit einem Vorhersagezeitraum von vier Stunden.....	46
Abbildung 19: Punktuelle Anomalie eines Temperatursensors.	47
Abbildung 20: Kontextuelle Anomalie eines Helligkeitssensors.	47
Abbildung 21: Kollektive Anomalie der Leistung einer Leuchte, da diese leuchte in der ersten Woche sechs Tage eingeschaltet war und jede weitere Woche fünf Tage.....	48
Abbildung 22: Beispiel der Anomaliedetektion anhand eines Temperatursensors. Als Anomalien dienen 25 zufällig ausgewählte Zeitbereiche, deren gemessene Werte künstlich um 15 % erhöht wurden. Zur Optimierung wurde die Precision-Recall Kurve verwendet.....	49
Abbildung 23: Confusion-Matrix, ROC und Precision-Recall-Kurven für die Modelle: Simple, ARIMA, XGBoost, LSTM und HDBSCAN zur Evaluierung der vorhergesagten Anomalien aus Abbildung 20 für einen Temperatursensor.....	51
Abbildung 24: Beispielhafte Zerlegung von Sensorsignalen in Tageslicht- und Kunstlichtanteil. Zeitdaten (time) für 01.11.2021 bis 28.02.2022, sowie die Detailansicht für 01.02.2022 bis 03.02.2022 dargestellt.	52
Abbildung 25: Beleuchtungsstärke auf Sensor durch Beleuchtungsanlage. Qualität der Sensorsignale nicht ausreichend, um einen Trend zu identifizieren. Zeitraum: 14.09.2021 bis 04.02.2022.	53

Abbildung 26: Heatmap der an der Decke verbauten Helligkeitssensoren eines Standorts mit Einstrahlwinkel aus Richtung Raumboden.....	54
Abbildung 27: CS-Wert Abschätzung nach Truong et al. [31] mit den Eingangswerten Beleuchtungsstärke und der CCT.....	55
Abbildung 28: Abschätzung der Tageslicht-CCT mit der Annahme, dass die CCT des Tageslichts sich zwischen 6500 K und 12000K im Zeitraum von Sonnenaufgang und Sonnenuntergang bewegt.	55
Abbildung 29: Visualisierung kombinierte Raumparameter. Die linke Visualisierung zeigt den CS-Wert am 04.10.2021 um 14:00 Uhr orts aufgelöst. Die rechte Darstellung zeigt eine Heatmap der Helligkeit kombiniert mit der Temperatur.	57
Abbildung 30: Kerndichteschätzung von allen Helligkeitssensoren eines Standorts.	58
Abbildung 31: Darstellung der vier KDE-Cluster der Beleuchtungsstärkesensoren.....	59
Abbildung 32: Grundriss eines Standorts mit Zuordnung der Sensoren zu den vier Clustern 0 – 3 aus Abbildung 31. Cluster 3 beinhaltet nur einen Sensor auf Tischhöhe mit Einstrahlrichtung von oben. Alle weiteren Sensoren sind an der Decke befestigt mit Einstrahlrichtung von unten.....	60
Abbildung 33: Jupyter-Notebook Entwicklungsumgebung der Datenanalyse-Toolbox mit Beschreibungstext, Code und Parametrisierung.....	62
Abbildung 34: Beispiel der Anomaliedetektion anhand eines Temperatursensors. Als Anomalien dienen drei zufällig ausgewählte Stellen, die sich über einen Zeitraum von sechs Tagen strecken und deren gemessene Werte um 15% über das Tukey Fenster erhöht wurden . Zur Optimierung wird die Precision-Recall Kurve verwendet	76
Abbildung 35: Metriken für die Beurteilung der Anomaliedetektion aus Beispiel in Abbildung 34. Zur Optimierung wird die Precision-Recall Kurve verwendet.....	77
Abbildung 36: Beispiel der Anomaliedetektion anhand eines Temperatursensors. Als Anomalien dienen drei zufällig ausgewählte Stellen, die sich über einen Zeitraum von sechs Tagen strecken und deren gemessene Werte um 15% über das Tukey Fenster erhöht wurden . Zur Optimierung wird die ROC verwendet	78
Abbildung 37: Metriken für die Beurteilung der Anomaliedetektion aus Beispiel Abbildung 36. Zur Optimierung wird die ROC verwendet.	79

Tabellenverzeichnis

Tabelle 1: Erweiterte Informationen zu Datenparametern.....	16
Tabelle 2: Exemplarische Merkmale verschiedener Datenpunkte einer Testanlage	39
Tabelle 3: Aufbau einer Confusion-Matrix mit TN (true negative), FP (false positive), FN (false negative) und TP (true positive) Anzahl an Werten. Das Label beschreibt die korrekte Zuordnung zur den zwei Klassen „Anomalie“ und „keine Anomalie“	50
Tabelle 4: Untersuchung unterschiedlicher Abtastintervalle für die Vorhersage von Beleuchtungsstärke, Temperatur und Luftdruck. Die Zahlen repräsentieren den RMSE und die Reduzierung der Datenmenge wurde mittels Mittelwertbildung durchgeführt.....	61

Zusammenfassung

Projektname: Semantische Interoperabilität von Beleuchtungssystemen, Analyse und Auswertung anhand von realen Zustandsdaten zur Folgerung von Use Cases in unterschiedlichen realen Standorten (Internet of Light, IoL)“

Projektzeitraum: 01.11.2020 bis 28.02.2023

Forschungseinrichtungen: Fachgebiet für Adaptive Lichttechnische Systeme und Visuelle Verarbeitung der TU Darmstadt (FE 1), Institut für Kommunikationstechnik der FH Dortmund (FE 2)

Ziel des Projektes: Die hochauflösende Erfassung von Zustandsdaten realer Beleuchtungsanlagen und der Aufbau eines zugehörigen Digitalen Zwillings ist eine Voraussetzung für explorative Datenanalysen sowie die daraus resultierende Generierung neuartiger applikationsspezifischer Geschäftsfelder (Use Cases). Dies ist insbesondere dann von Bedeutung, wenn neben den physikalischen Größen der semantische Kontext der Beleuchtungsanlage erfasst werden soll, um auch die Möglichkeit zu eröffnen, wissensbasierte KI-Methoden einzusetzen. Dazu bilden maschinenlesbare Anlageninformationen einen essenziellen Grundstein. Schlussfolgernd eröffnen sich darüber hinaus für viele Unternehmen in der Lichtindustrie neue Geschäftsfelder im Hinblick auf digitale Dienstleistungen z.B. im Bereich situationsadaptiver, intelligenter, KI-gestützter Beleuchtungssysteme und Predictive Maintenance. Bei diesem Ziel ist die Verwendung einer semantischen Beschreibung z.B. mit der hier entwickelten WoT Beschreibung auf der Basis von SENSE-WoT sehr hilfreich. Durch die Diskussion innerhalb der Industrie beispielsweise auch im Rahmen des Fachverbandes Licht des Verbands der Elektro- und Digitalindustrie (ZVEI) gewinnt das Erfassen und Auswerten von Daten zunehmend an Bedeutung, da hier ein erhebliches Fortschrittspotential im Hinblick auf die internationale Wettbewerbsfähigkeit der Unternehmen gesehen wird. Die Datenerhebung und -speicherung unterschiedlicher und verteilter Anlagen bildet die Grundlage, um ein bisher ungenutztes Innovationspotenzial hinsichtlich künftiger Dienste-Orientierung in Beleuchtungsanlagen zu erschließen. Die erhobenen Daten müssen hierfür jedoch unter Verwendung einer einheitlichen Semantik beschrieben werden. Im Rahmen des Projektes IoL wurde, zunächst für Analysezwecke, in der Cloud ein herstellerübergreifendes System-Monitoring mit Fokus auf professionellen Beleuchtungsanlagen unter Reallabor-Bedingungen aufgebaut, um Fragestellungen aus den Bereichen Einheitlichkeit des Datenformates, Datenplausibilität, datengetriebene Use Cases, Anwendernutzen und Algorithmen-Auswahl zu bearbeiten.

Ergebnisse:

Zur Erreichung der vorgenannten Ziele wurden zunächst die Testanlagen hinsichtlich ihrer Datenstruktur sowie Erhebungsintervalle untersucht und anschließend individuelle Datenkonvertierungs-Konzepte und -Module umgesetzt, sowie eine Cloudinfrastruktur zur Zustandsdatenverwaltung (inkl. Historie) und zur semantischen Annotation aufgebaut (AP1-2). Zur einheitlichen Beschreibung wurden Ontologieuntersuchungen für die Anwendungsdomäne Licht durchgeführt und ein erweitertes Web-of-Things-Modell auf der Basis von SENSE-WoT entworfen (AP3). Ein inhaltsbasiertes Monitoring-System wurde entwickelt und implementiert, welches Fehler und Anomalien in den Datenfeldern detektiert. Des Weiteren wurden verschiedene Vorhersagemethoden für die Raumparameter untersucht, die es ermöglichen,

zukünftige interne und externe Zustände vorherzusagen. Ausgehend von den real vorhandenen Sensoren wurden die Möglichkeiten zur Reduzierung der Sensoranzahl bei gleicher Abbildung der Use Cases abgeleitet. Die optimale Raumposition abhängig von frei wählbaren Parametern wurde mittels automatisch erzeugter Heatmaps realisiert. Die entwickelten Analysemethoden, Prädiktionsalgorithmen und Use Case Demonstrationen wurden in einer umfassenden Toolbox zusammengefasst, die es ermöglicht, die Datensätze selektiv zu verarbeiten, Visualisierungen der Zwischenschritte zu schaffen und so schließlich Erkenntnisse zu gewinnen (AP4).

Im Rahmen des Projektes wurden dabei pandemiebedingt keine Vermessungen der Anlagen durchgeführt, sondern der benötigte Informationsgehalt der semantisch angereicherten Datensätze für eine Abbildung der Use Cases ohne vor-Ort-Begehung in den Fokus gestellt. Um die Konnektivität und Datenqualität der Anlagen sicherzustellen, wurden zyklische Prüfungen implementiert. Außerdem wurden regelmäßige Updates der Cloudanbindung vorgenommen, um sicherzustellen, dass die Anlagen nahtlos mit den Cloud-Systemen kommunizieren können (AP5).

Das Forschungsziel wurde erreicht.

1. Einleitung

Das folgende Kapitel gibt einen kurzen Überblick zur Zielstellung des Projektes und erläutert die Motivation für den gestellten Antrag sowie die durchgeführten Untersuchungen.

1.1. Technische und Wirtschaftliche Problemstellung

Datenerhebung und -speicherung unterschiedlicher und verteilter Anlagen bildet die Grundlage, um ein bisher ungenutztes Innovationspotenzial hinsichtlich **künftiger Dienste-Orientierung** und der Anwendung von KI-Systemen z.B. für eine intelligente Lichtsteuerung in Beleuchtungsanlagen zu erschließen. Beleuchtungsanlagen, einschließlich zugehöriger Sensorik, müssen hierfür jedoch in der Lage sein, Daten zu erfassen und diese mit einer **einheitlichen Semantik** in eine Cloud zu übertragen, dies ist insbesondere für zukünftige KI-Anwendungen in heterogenen Systemen erforderlich.

Im Rahmen des Projektes IoL wurde, zunächst für Analysezwecke, in der Cloud ein herstellerübergreifendes System-Monitoring mit Fokus auf professionellen Beleuchtungsanlagen unter Reallabor-Bedingungen aufgebaut.

Einheitliche Semantik

Zu Projektbeginn war die Interoperabilität von professionellen Beleuchtungsanlagen herstellerübergreifend nur eingeschränkt durch den Einsatz entsprechender Gateway-Lösungen gegeben. Dies hatte zur Folge, dass nur ein reduzierter Funktionsumfang in die jeweils andere angebundene Systemwelt übersetzt werden konnte. Somit waren die verfügbaren Funktionen für bestimmte Dienste im professionellen Einsatz nicht vollständig gekoppelt. Auch der Einsatz von Internet of Things (IoT)-Plattformen (z.B. AWS IoT, Azure IoT) zur Kopplung der Systeme war im Rahmen dieses Projektes nicht zielführend und kann Nachteile aufweisen (z.B. Vendor-Login, ggf. geschlossenes Datenformat). Zwar bieten diese Plattformen eine erhöhte Konnektivität, jedoch ist eine semantische Anreicherung der Zustandsdaten und die Beschreibung von Umgebungsbedingungen im Hinblick auf hersteller- und gewerkeübergreifende Mehrwertdienste aufgrund geschlossener Datenformate nur sehr eingeschränkt möglich.

Die semantische Beschreibung von Daten professioneller Beleuchtungsanlagen und deren räumlicher Kontexte, als Bestandteil intelligenter Gebäude der Zukunft, kann einen wertvollen Beitrag zur herstellerübergreifenden und flexiblen Umsetzung von Mehrwertdiensten leisten. In diesem Zusammenhang sind die Anforderungen von Building Information Modeling (BIM)-Projekten innerhalb des „Stufenplan Digitales Planen und Bauen“ [1] des BMVI zu nennen. Somit kann das Projekt IoL als ein zusätzlicher Anstoß zur schrittweisen Nutzung und Einführung des BIM und der Nutzung der Informationen aus professionellen Lichtplanungstools angesehen werden, indem deren semantische Informationen maschinenlesbar für die jeweilige Anlage im realen Betrieb zur Verfügung gestellt werden können.

Hierzu wurden in IoL semantisch angereicherte Daten generiert, welche auch die Entwicklung von hersteller- und gewerkeübergreifenden Services ermöglichen, ohne tiefere Kenntnisse der zugrundeliegenden Technologien auf Seiten der Dienstentwickler und Data Scientists zu erfordern. Auch Anwendungen der Künstlichen Intelligenz (KI) wie Maschinelles Lernen (ML)

werden auf der Basis von semantisch beschriebenen Daten umfangreiche Marktveränderungen bewirken.

Den Unternehmen kommt im Hinblick auf die Semantik für den KI/ML-Bereich und Automatisierung die Aufgabe der Beschreibung von lichttechnischen Eigenschaften auch mit Bezug zum zukünftigen Digitalen Zwilling eines Gebäudes zu, der den aktuellen Zustand eines Gebäudes und dessen entsprechende Historie abbildet. Das Projekt IoL lieferte hier wichtige Impulse und Beiträge, um die mittelständische Lichtindustrie für einen zunehmend durch das IoT beeinflussten Markt zu stärken und beispielsweise den Einsatz von KI/ML im Bereich Lichttechnik zum Beispiel in Form von selbstlernenden Beleuchtungssystemen zu ermöglichen.

Intelligente Beleuchtung

Die Integration von Sensoren und Aktoren in vernetzte Beleuchtungsanlagen ermöglicht diesen, den Zustand ihrer Umgebung zu erfassen. Die Anlage selbst kann durch eine entsprechend intelligente Applikationssoftware auf die erfassten Umgebungs- und Leuchtzustände reagieren und die Beleuchtung anpassen. Für eine bestmögliche Optimierung ist die Kenntnis von historischen wie auch aktuellen Zustandsdaten Voraussetzung. Diese werden daher im Kontext von IoL event-basiert oder in vorgebbaren Abtastintervallen in der IoL-Cloud abgelegt und verwaltet. Die persistente Datenablage und -verwaltung in der Cloud bietet Anlagenbetreibern wie auch Anlagennutzern gleichzeitig neue Potentiale. So ist beispielsweise die bedarfsorientierte Planung von Wartungszyklen ohne eine vorherige vor-Ort-Begehung möglich. Durch optimiertes Licht entsteht für den Anlagennutzer eine positive human-physiologische und kognitive Auswirkung. Beide hier exemplarisch genannten Anwendungsfälle führen zu einem ökonomischen Vorteil auf beiden Seiten. Dies begünstigt neben der Migrationsmöglichkeit für zusätzliche Komponenten oder die dienste-basierte Konfigurationsänderung wiederum die Investitionsbereitschaft in intelligente Beleuchtungsanlagen. Auch während des Betriebs kann durch die Schaffung neuer datenbasierter Applikationen und Dienste der Anlagenwert nachhaltig gesteigert werden. Im Rahmen des Projektes IoL wurden daher zukunftsweisende Anwendungsfälle anhand realer Zustandsdaten aus 8 verschiedenen, realen Standorten im Live-Betrieb abgeleitet. Dies erlaubt insbesondere den KMU einen direkten Einblick in die Erarbeitung neuer Anwendungsfälle und damit zusammenhängender dienste-basierter Geschäftsmodelle.

1.2. Wirtschaftliche Bedeutung der Forschungsergebnisse für KMU

Die positiven Entwicklungen der Beleuchtungstechnologie der letzten Jahre hat zu Leuchten-Lichtausbeuten von 180 - 200 lm/W geführt und damit ein großes Energieeinsparpotenzial erschlossen. Aktuelle europäische Ecodesign- und Labeling-Regulierungen fördern die Umsetzung dieser Potenziale. Allein durch eine weitere Optimierung der Lichtquellentechnologien sind jedoch in nächster Zeit keine signifikanten Verringerungen des Energiebedarfs von Beleuchtungssystemen mehr zu erwarten. Die weitere technologische und wirtschaftliche Entwicklung der Lichtbranche hängt immer weniger von einzelnen Komponenten und mehr von umfassenden Lösungen und Dienstleistungen in den einzelnen Anwendungen ab. „Smarte“ Beleuchtungsanlagen mit intelligenten Steuerungen und Sensoren können nicht nur zu einer weiteren Reduzierung des Energiebedarfs, sondern auch zu weiteren Vorteilen für Hersteller, Betreiber und Nutzer von Beleuchtungsanlagen führen. Die parallele dynamische

Entwicklung der Digitalisierung und des Internets der Dinge weist für die Lichtindustrie große Potentiale für einen substanziellen Auf- und Ausbau dienste-orientierter Geschäftsmodelle, sogenanntem „Light as a Service“ (LaaS), auf. So prognostizieren Marktanalysen für die nächsten Jahre einen Zuwachs beim Umsatz mit digitalen Lösungen von mindestens 10% jährlich weltweit allein im Bereich Licht [2] [3] [4]. Diese Entwicklung in Richtung Digitalisierung, Konnektivität und Datennutzung stellt derzeit noch eine Hürde für die überwiegend mittelständische Lichtindustrie in Deutschland dar [5]. Durch die Ergebnisse des Projekts IoL soll diese Hürde gesenkt werden. Dabei müssen aktuelle IT-Technologien wie das Internet der Dinge und Cloud-basierte Lösungen zur Lichtsteuerung beherrscht sowie die Umsetzung digitaler Dienste verstanden werden, um sich auch in Zukunft auf dem Weltmarkt behaupten zu können. Dazu ist eine einheitliche Art der Kommunikation im Bereich professioneller Beleuchtung eine grundlegende Voraussetzung. Beispielsweise haben sich führende internationale Technologie-Unternehmen zu einer Working-Group zusammengeschlossen [6], um in dem Projekt „Connected Home over IP“ (CHIP) [7] einen lizenzfreien Konnektivitäts-Standard zur Erhöhung der Kompatibilität zwischen Smart Home-Produkten zu entwickeln. Solch ein Standard fehlt im Bereich professioneller Beleuchtung derzeit gänzlich. Die Integration aktuell in der Entwicklung befindlicher Standards wie CHIP in professionelle Beleuchtungseinrichtungen ist aufgrund einer fehlenden einheitlichen Semantik und der verwendeten Feldbussysteme derzeit nicht zielführend. Insbesondere stellen die unterschiedlichen Anwendungsdomänen Smart Home und Smart Building unterschiedliche kommunikationstechnische Anforderungen beispielsweise hinsichtlich der Latenzzeiten.

Vor dem Hintergrund des zunehmenden Vordringens großer IT-Konzerne in immer mehr Applikationsbereiche ist es daher wichtig, dass die deutsche Lichtindustrie zeitnah durch eigene Bestrebungen, wie dieses Projektvorhaben, Entwicklungsimpulse zur Digitalisierung nach ihren eigenen Vorstellungen setzt und sich in die internationale Standardisierung kompetent einbringt.

Aus den zuvor genannten Gründen wurde die Idee für das Projekt IoL als gemeinschaftliche Projektarbeit zwischen der deutschen Lichtindustrie und den Forschungseinrichtungen erarbeitet.

Das Projekt IoL hatte die Diversität der aktuellen Marktsituation im Hinblick auf unterschiedliche Sensoren, Leuchten und deren Daten mit acht unterschiedlichen Standorten dargestellt. Die Komplexität des Themas wurde aufgezeigt und bietet den KMU einen anwendungsorientierten Einstieg in die Thematik Internet of Things (IoT). Die erarbeitete einheitliche Semantik bietet eine Lösung für die Erhebung und Beschreibung von Daten zur gesamtheitlichen Verarbeitung in einer Cloud. Die Notwendigkeit einer Building Topology Ontology (BoT) wurde aufgezeigt und sollte bei Planung und Installation von neuen Anlagen erstellt werden. Mithilfe der so angereicherten Daten kann der Markt für KMU mit ressourcengebundenen Geschäftsmodellen um digitale Dienste, wie „Light as a Service“ (LaaS), ergänzt werden. Die detaillierte Beschreibung des Datenflusses in diesem Projekt kann als Basis für den Einstieg in eine digitale Dienst-Umgebung genutzt werden und zeigt einen Weg, welcher erfolgreich mit acht unterschiedlichen Anlagen aufgebaut wurde und als Blaupause für weitere KMU dienen kann. Die einheitliche Semantik im Bereich Licht erleichtert ebenfalls weiteren Branchen die Anbindung deren Dienste an Lichtapplikationen. Weiterhin werden die Tools und Schritte innerhalb einer Cloud-Umgebung erläutert und können in die jeweiligen Systeme

eines KMU übernommen werden. Die Details, die hierzu nötig sind, wurden mit der aufgebauten IoL-Cloud für KMU beispielhaft beschrieben.

Drei exemplarische Use Cases in diesem Projekt stellen grundlegende Beispiele für eine Datenauswertung dar. Weiterhin evaluieren die Use Cases den Nutzen der Daten für die Unternehmen. Der Datenfluss, die Vorverarbeitungs- und Auswertungsmethoden können für weiterführende individuelle Use Cases abgewandelt werden. Die Methoden können von KMU besonders leicht in eigene Produkte oder Systeme integriert werden und verursachen deshalb vergleichsweise wenig zusätzlichen Entwicklungsaufwand für den Beginn mit digitalen Geschäftsmodellen.

Die erarbeitete Datenanalyse-Toolbox, die alle verwendeten Methoden in diesem Projekt zusammenfasst, wird den Projektpartnern und alle interessierten Firmen als Source-Code in der state-of-the-art Programmiersprache Python zur Verfügung gestellt und kann als Grundgerüst für den einfachen Eintritt in eine digitale Dienste Umgebung genutzt werden.

Von den Erkenntnissen und Anwendungsmöglichkeiten der Ergebnisse des Projektes IoL zum Aufbau einer Cloud, Übertragung der Daten von einzelnen Standorten mit unterschiedlichen Licht- und Sensorpunkten, sowie einer allgemeinen Datenauswertung können neben den allein im ZVEI-Fachverband Licht organisierten etwa 130 Unternehmen auch zahlreiche weitere – vor allem mittelständische – Branchenunternehmen sowie branchenübergreifend weitere Unternehmen z.B. aus den Bereichen Gesundheitswesen, Gebäudetechnologien und Energiewirtschaft profitieren. Die Projektergebnisse bieten grundsätzliche Herangehensweisen zur Verbesserung der semantischen Interoperabilität und zeigen exemplarische digitale Services zur Anwendung in verschiedenen Einsatzgebieten innerhalb der Lichtindustrie. Die Ergebnisse, etwa zur Interoperabilität, haben in generalisierter Form zudem Potenzial für eine Anschlussfähigkeit für Anwendungen auch in anderen Branchen.

Wissenschaftlich technischer Ansatz

Review-Paper wie [8] [9] [10] geben eine abstrahierte Übersicht zum Einsparpotential durch derzeitige, intelligente Beleuchtungssysteme in Abhängigkeit von verfügbaren Sensordaten, Sensorpositionen, Raumnutzungsarten sowie Steuerungs- und Regelungsalgorithmen. Als thematisch verwandte, bereits abgeschlossene Forschungsprojekte sind GreenerBuildings [11], EnLight [12] und das OpenAIS Projekt [13] [14] zu nennen.

Die Zielsetzung des OpenAIS Projektes war die Definition einer optimierten, offenen Systemarchitektur für intelligente Beleuchtungssysteme. In einer Pilotanlage wurde ein voll funktionsfähiges herstellerübergreifendes Lichtsteuerungssystem auf Basis der OpenAIS Architektur realisiert. Des Weiteren lag der Fokus beim Projekt OpenAIS auf der Steuerung eines Lichtsystems. Dies ist im beantragten Projekt IoL nicht der Fall. Dabei war das Ziel neben der Bereitstellung der o.g. semantischen Informationen für eine darüber angeordnete Dienstenumgebung eine realzeitnahe herstellerübergreifende Steuerung der Leuchten zu erreichen und damit eine gruppenbasierte Steuerung über IP-Kommunikation in der Dienstenumgebung zu ermöglichen. Im Zuge dessen ist ein dediziertes Datenmodell zur Realisierung eingesetzt worden, um diese Qualitätsmerkmale zu erfüllen. Hierfür wurde weitestgehend auf bestehende Ontologien zurückgegriffen und diese zur vollumfassenden

Beschreibung der Daten zusammengefügt und ergänzt. Das so erarbeitete Datenmodell kann zur Unterstützung von Normen und Standards genutzt werden. Dies ist besonders für KMU hilfreich, um die große Protokollvielfalt für die Lichtbranche klar zu selektieren und eine Richtung aufzuzeigen.

In keinem der bisherigen Projekte wurde bearbeitet, welcher Nutzen aus den Sensor- und Betriebsgerätedaten gewonnen werden kann, wenn diese nicht nur momentan zur dynamischen Lichtsteuerung, sondern im Langzeitkontext ausgewertet werden. Langzeitdaten wie diese sind dabei das Gold des 21. Jahrhunderts [15] und unterstützen die effektive Gestaltung von Prozessen und Weiterentwicklung von Produkten und Dienstleistungen. Der deutsche Mittelstand nutzt das Potential dieser Daten bisher kaum [16].

Im Rahmen des Projektes wird daher als datengetriebener und effizienzsteigernder Use Case einer intelligenten Beleuchtungsanlage beispielhaft ein Raummanagement nach Raumparametern entwickelt, um den Raum in verschiedenen Zonen einteilen und je nach Nutzungsszenario des Raums die optimale Zone für diese Nutzung identifizieren zu können. Wenn an den häufig benutzten Arbeitsplätzen im Raum das Lichtniveau und die biologische nicht-visuelle Wirksamkeit nicht ausreicht, sollen die Kunstlichtanlagen dementsprechend eingestellt werden (Lichtniveaueinstellung). Mittels einer einfachen Konfiguration können unterschiedliche Parameter(kombinationen) für die Identifikation herangezogen werden und KMU können diese Softwarebasis nutzen, um dies in bestehende oder neue Services zu integrieren. Die Auswertung eines „einfachen“ Beleuchtungsstärkesensors für die Abschätzung der nicht-visuellen Wirksamkeit mittels des circadian stimulus kann für KMU insbesondere einen großen Marketing Mehrwert bieten.

Über ein solches System wird die Arbeitsplatzbelegung auf Tageslichteinfall optimiert und so kann, bei gesteigerter circadianer Wirksamkeit und weniger Kunstlichteinsatz, energieeffizient gearbeitet werden. Dieser Anwendungsfall ist angelehnt an die besonders weit entwickelten Beleuchtungssysteme insbesondere im Pflegebereich. Des Weiteren wird im Rahmen dieses Projektes ein dynamisches Lebensdauer-Monitoring basierend auf den realen Betriebsdaten der Leuchten entwickelt. Mithilfe dieses Use Cases für eine prädiktive Anlagenwartung konnte gezeigt werden, dass mit einem einfachen Beleuchtungsstärkesensor die Degradation theoretisch abgeschätzt, sowie der circadian stimulus (CS) als Bewertung der biologischen nicht-visuellen Wirksamkeit an unterschiedlichen Orten im Raum festgestellt werden kann. Über Predictive Maintenance lassen sich in Zukunft Anfahrtswege für die Wartungstechniker optimieren und Probleme bzw. Auffälligkeiten präventiv untersuchen.

Sowohl im EnLight als auch im OpenAIS Projekt wird als offener Forschungsansatz der Standby-Energieverbrauch der Vielzahl von eingesetzten Sensoren genannt [5] [1]. In diesem Projekt sollte deshalb die Reduktion der aktiven Sensoranzahl in einem System bei gleicher Funktionalität durch die Einführung von virtuellen Sensoren untersucht werden. Die Methode der Kerndichteschätzung gepaart mit einem dichte-basierten Clustering-Algorithmus bietet für die Planung von neuen Anlagen eine Methode zur Optimierung der Anzahl an verbauten Sensoren mit mathematischer Modellierung für weitere Positionen im Raum. Die Optimierung der Sensoranzahl reduziert die Anlagenkosten, spart die Energie für die Sensorelektronik ohne die Informationsrelevanz und Komplexität der Anlagen zu reduzieren. KMU können somit besonders kostengünstige Lösungen anbieten ohne die Datenqualität für die Auswertung in Use Cases zu reduzieren.

Zuzüglich zu den bisher genannten Use Cases ist die Extraktion von bisher unbekanntem Anwendungsfällen über die Analyse von Zusammenhängen in den Zustandsdaten ein Bestandteil des IoL-Projektes, um das genannte Daten-Potential zu erschließen.

Aus kommunikationstechnischer Sicht wurde im Projekt IoL, im Gegensatz zu den vorgenannten thematisch verwandten Projekten der Ansatz verfolgt, bestehende, mit unterschiedlichen Feldprotokollen kommunizierende Anlagen, welche als Ganzes über ein IP-Gateway oder Cloud2Cloud angesprochen werden können, in eine herstellerübergreifende Cloudlösung einzubinden und insbesondere die ausgetauschten Datenmodelle zu vereinigen und semantisch anzureichern. Hierbei wurde im Gegensatz zu OpenAIS ein öffentliches Modell aus der Web of Things (WoT) Architektur mit semantischen Erweiterungen verwendet. Das so in diesem Projekt erarbeitete Datenformat kann aufgrund seiner semantischen Interoperabilität mit geringem Integrationsaufwand in die lichttechnischen Systeme anderer Unternehmen integriert werden.

Voraussetzung für die Umsetzung der zuvor beschriebenen Potentiale und Use Cases ist eine systemübergreifende und damit einheitliche Modellierung der Daten. Dies wird im Projekt anhand der Konzepte des digitalen Zwillings (Digital Twin, DT, virtuelle Repräsentation eines Beleuchtungssystems im Kontext der Gegebenheiten eines Gebäudes) und der semantischen Interoperabilität umgesetzt. Das Konzept zur WoT-Beschreibung eines Smart Home/Smart Building wurde im BMWI-Projekt SENSE [6] unter Beteiligung des ZVEI, des DFKI Saarbrücken und dem IKT der FH Dortmund erarbeitet. Zu einer einheitlichen Semantik und weiterführend einer semantischen Interoperabilität von Systemen/Systemwelten gehören u.a. auch die Beschreibung der messtechnischen Eigenschaften wie physikalischer Größe und Einheit, Grenzwerte, Granularität der Amplitude, Abtast-Intervalle und Messverfahren. Die so ermittelten Datenstrukturen können von KMU genutzt werden und in eigene Use Cases integriert werden, um eine für zukünftige Konzepte gewappnete Darstellung und Speicherung der Daten zu gewährleisten.

Somit kann eine maschinenverstehbare Vernetzung hergestellt werden. Diese ermöglicht die Beschreibung und funktionale Einordnung eines Gerätes (Sensor und Aktor) im technischen Kontext sowie in dessen Anordnung im Gebäude/Raum und den vielfältigen Relationen zueinander.

Neben der Speicherung von Systemeigenschaften, Parametern und Zustandswerten, beinhaltet der DT auch die Modellierung der Zusammenhänge von Teilkomponenten des Gesamtmodells. Um die Eindeutigkeit der im DT enthaltenen Strukturen und Elemente sicherzustellen, eignet sich die Verwendung von Semantik-Web-Technologien. Basierend auf öffentlich zugänglichen Ontologien [1],[2],[3] besteht eine formale Beschreibung der modellierten Zusammenhänge. Datenformate wie RDF oder JSON-LD [4] sind für eine maschinenlesbare Weiterverarbeitung der Modelldaten geeignet. Sollen weitergehende Erkenntnisse aus dem Modell ermittelt werden, können implizite Zusammenhänge ggf. über ein geeignetes Reasoning (semantisches Auflösen von Zusammenhängen) ermittelt werden. Damit stehen die geplanten Arbeiten auch im direkten Bezug zum Arbeitskreis „Semantische Interoperabilität“ des ZVEI [5] und den vom BMWI geförderten Projekten SENSE [6], SUITE [7] und ForeSight [8]. Die miteinander vernetzten Projekte sind in den Bereichen Smart Home bzw. Smart Living und Ambient Assisted Living verortet. In IoL wurden die Ergebnisse aus SENSE als Basis für weitere Untersuchungen eingebracht. Aufgrund der Komplexität der Leuchten-Datenpunkte und Parameter sowie der

unterschiedlichen Projektziele konnte dies nur in Rahmen eines dedizierten Projektes realisiert werden.

Die in diesem Projekt erarbeiteten Grundlagen können KMU bei der Entstehung neuer Produkte und erweiterter Geschäftsmodelle sowie der Steigerung der Produktqualität gepaart mit einer schnellen Reaktionsgeschwindigkeit durch die digitale Vernetzung und den sich daraus entwickelnden neuen Märkten dienen.

1.3. Arbeitshypothese

In diesem Projekt werden die Daten aus Testanlagen, welche über 8 Standorte verteilt installiert sind, in einer gemeinsamen IoL-Cloud zusammengeführt. Ziel dieses Vorhabens ist die semantische Beschreibung der in den Testanlagen generierten Zustandsdaten bei gleichzeitiger Erprobung der Architektur für die Datenspeicherung und Datenübertragung. Hierauf aufbauend werden die zentral gespeicherten Daten analysiert und ausgewertet, sowie datengetriebene, lichttechnische Use Cases untersucht und bearbeitet. Die Betrachtung von realen Daten aus unterschiedlichen Standorten und Raumarten birgt großes Potential für eine allgemeine Typisierung und Generalisierung der Anwendungsfälle und einen direkten Transfer der praktischen Erfahrungen in die Industrie.

1.4. Einfluss der Corona-Pandemie auf den Projektverlauf

Die Projektlaufzeit war durch die Corona-Pandemie geprägt. Entgegen der Planung gemäß Projektantrag mussten vor-Ort-Analysen der anzubindenden Testanlagen weitgehend entfallen und über Remote-Zugriff und kontinuierlichen Austausch der Projektmitarbeiter mit den technischen Ansprechpartnern der Unternehmen kompensiert werden. Zu Beginn des Projektes wurde mit zwölf Standorten gerechnet, wobei Pandemiebedingt acht Standorte final im Projekt IoL eingebunden werden konnten. Während der Projektlaufzeit sind fünf Anlage ausgeschieden. Die Gründe hierfür waren, Standortwechsel des Unternehmens, Umnutzung der geplanten Räumlichkeiten als Corona-Testzentrum und unternehmensinterne Gründe, wodurch der Testraum nicht mehr zur Verfügung gestellt werden konnte. Infolgedessen wurde durch FE1 ein Besprechungsraum als weiterer Testraum und Standort zur Verfügung gestellt, welcher als erste Test- und Beispiel-Anlage diente.

Seitens der Forschungseinrichtungen wurde ebenfalls die Suche und Verfügbarkeit von qualifiziertem Personal, sowie die Einstellungs- und Einarbeitungsbedingungen durch die Pandemie erschwert. Weiterhin hatte COVID-19 zu einer starken Einschränkung der Zutritts- und Nutzungsmöglichkeiten der Büros und Labore bei den Forschungseinrichtungen geführt, wodurch ein anfänglich langsamerer Fortschritt im Projekt stattgefunden hatte, welcher zum Ende ausgeglichen werden konnte.

Zusammenfassend, konnten trotz aller genannten Einschränkungen die primären Projektziele vollumfänglich erreicht werden.

1.5. Über das Projekt und die Projektpartner

FE 1: Technische Universität Darmstadt, Koordinierende FE

Fachbereich Elektrotechnik und Informationstechnik, Fachgebiet Lichttechnik;
Hochschulstraße 4a, 64289 Darmstadt;
Prof. Dr.-Ing. habil. Tran Quoc Khanh

FE 2: Fachhochschule Dortmund

Fachbereich Informationstechnik, Institut für Kommunikationstechnik,
Fachgebiet Kommunikationstechnik und digitale Signalverarbeitung;
Sonnenstraße 96, 44139 Dortmund;
Prof. Dr.-Ing. Ingo Kunold

2. Arbeitspakete

Aufgrund der während der Projektlaufzeit aufgetretenen pandemischen Lage musste die geplante Vorgehensweise abgewandelt werden. Es wurden vorrangig Remote-Zugänge zu den Testanlagen und Laboren genutzt, sowie die vorgesehene Trennung zwischen Erhebungsphase und Implementierung durch eine agile sequenzielle Anbindung der Testanlagen ersetzt. Trotz der erforderlichen Anpassungen konnten die Projektziele wie geplant umfänglich erreicht werden. Im Nachfolgenden werden die Projektergebnisse je Arbeitspaket näher erläutert und den geplanten Zielen gegenübergestellt.

2.1. AP1 Vorbereitung der Auswertung

In der initialen Projektphase wurden verschiedene Maßnahmen zur Erhebung des IST-Zustandes der anzubindenden Testanlagen durchgeführt. Hierzu wurden zunächst anhand von Fragebögen die grundsätzlichen lichttechnischen (Tageslichteinfall, Sensortypen, Leuchtentypen und -positionen, erfassbare Betriebsdaten, Anwendungsszenario, etc.) und kommunikationstechnischen Eigenschaften (IP-fähig, Protokolle, Schnittstellen, etc.) der Anlagen erhoben. Die identifizierten Merkmale wurden als Ausgangsbasis an die nachfolgend beschriebenen Prozesse übergeben.

2.1.1. Konzeption der Datenkonvertierungsmodule

Die erhobenen technischen Rahmenbedingungen (Kommunikationsprotokolle, Schnittstellen, WAN-Anbindung, etc.) der Testanlagen zeigten eine hohe Diversität im Bereich der eingesetzten Protokolle und Datenformate. Jedoch konnten MQTT und REST als Schnittmenge der unterstützten Schnittstellen zum Datenabruf unter den untersuchten Anlagen identifiziert werden.

Anhand der erhobenen Parameter wurden ein initiales Clustering der Testanlagen durchgeführt, und anschließend in einer zweiten Iteration erweiterte Informationen zu den Datenparametern erhoben. Da im Projekt der Fokus auf die lichttechnische Daten-Erfassung lag, wurden hier primär die Datenpunkte der verfügbaren Sensoren (je Sensor 1..n Datenpunkte) erfasst und beschrieben (siehe Tabelle 1). Dieser Parameterkatalog wurde als Basis zur Konzeption eines homogenen Datenmodells in AP2 und zur Konzeption der Softwaremodule in AP3 eingebracht.

Tabelle 1: Erweiterte Informationen zu Datenparametern.

Parameter	Beschreibung
ID	ID in Datenbank
Parametername	Bezeichnung in Datenbank
Messbereich (min, max)	Abgedeckter Messbereich
Datenblatt	Link zu Datenblatt
Kurzbeschreibung	Kurze Beschreibung des Parameters

Im Rahmen der vorliegenden Untersuchung hat sich als zielführend erwiesen den Parameternamen so zu wählen, dass eine eindeutige Identifikation und Ableitung der Funktion ermöglicht wird.

Für die semantische Beschreibung der Anlagen durch das WoT-Modell sind Angaben zum Messbereich des jeweiligen Datenparameters von besonderer Bedeutung. Unterschiedliche Messgeräte und deren Interpretation durch die Anlagenbetreiber führen zu verschiedenen Varianten eines funktional gleichwertigen Messparameters wie beispielsweise der Temperatur (Lufttemperatur, Gehäusetemperatur, Temperatur Vorschaltgerät, Platinen-Temperatur), was eine anlagenspezifische Betrachtung und Beschreibung notwendig macht und in einen nachgelagerten Schritt in ein einheitliches Datenformat überführt werden muss.

Die semantische Beschreibung unterstützt dabei, die physikalischen Geräte und Parameter abzubilden, was sowohl eine detaillierte Beschreibung der Datenparameter als auch des Gerätes selbst einschließt. Das WoT-Modell beinhaltet zudem einen Hardware-Link, der auf das Online-Datenblatt des Gerätes verweist und es ermöglicht, bei Bedarf weitere Informationen zum Betrieb und zur Konfiguration abzurufen. Die Repräsentation der Daten wurde so weit wie möglich an die DALI-Spezifikationen angelehnt.

Angesichts der verschiedenen Varianten von Datenparametern mit ähnlicher Funktion erwies sich eine Kurzbeschreibung im WoT-Modell als hilfreich, um Unterschiede oder genaue Funktionen zu identifizieren. Darüber hinaus wurden die Beschreibungen so formuliert, dass alle relevanten Informationen komprimiert verfügbar sind.

2.1.2. Datenevaluation (Vollständigkeit und Redundanz)

Bedingt durch die in Kapitel 1.4 benannten Einflüsse der Corona-Pandemie konnte eine qualitative Kontrolle der Anlagendaten auf Relevanz, Erhebungsintervall und Datenmenge zunächst nicht wie geplant für alle Testanlagen in dieser frühen Phase des AP1 vorgenommen werden. Diese Untersuchungen wurden aber für jede Testanlage sukzessive nach erstmaliger Anbindung der Anlage und im Laufe des Langzeittests nachgeholt. Ab Beginn der Datenaufzeichnungen wurde für jede angebundene Testanlage ein Datenmonitoring etabliert, welches automatisiert die eingehenden Daten analysiert und eine vollständige Datenaufnahme unterstützt. Bei Nichterhalt von Datenpunkten wurden die Anlagenbetreiber informiert und die Ursache besprochen. Eine genaue Beschreibung des Monitorings ist in Kapitel 2.5.2 zu finden.

Bei der Anbindung der Anlagen wurden die Daten mit den Anlagenbetreibern besprochen und für dieses Projekt wichtige Datenpunkte identifiziert, sodass Datenpunkte eindeutige Eigenschaften und Zustände beschreiben, die nicht aus anderen Datenpunkten abgeleitet oder anderen Einheiten beschrieben werden können. Weiterhin wurde für die Vergleichbarkeit der Anlagen über die Use Cases darauf geachtet, dass ähnliche Datenpunkte aus den acht Anlagen gespeichert werden. Durch die iterative Anbindung der Anlagen wurden alle so bestimmten Datenpunkte gespeichert und nach der Anbindung aller Anlagen ausgewertet. Dies ermöglichte ein breites Spektrum an Daten aufzunehmen, die im Datenverarbeitungsschritt Verwendung finden konnten. Die explizit umgesetzten Maßnahmen, sowie Use Cases werden in Kapitel 2.5 näher erläutert.

2.1.3. Analyse Erhebungsintervalle

In Abhängigkeit des zu betrachtenden Use Cases variieren die erforderlichen Erhebungsintervalle der zugrundeliegenden Datenparameter. Ziel bei der Identifikation der optimalen

Erhebungsintervalle war es eine allgemeine Empfehlung für alle eingehenden Datenparameter zu erarbeiten. Die initialen Erhebungsintervalle wurden als Ausgangssituation durch FE2 dokumentiert und basierend hierauf zusammen mit FE1 ein Vorschlag erarbeitet. Anschließend wurde dieser mit den Anlagenbetreibern sofern technisch möglich umgesetzt. In Fällen, wo die jeweilige Anlagenkonfiguration keine Anpassung zuließ, wurde durch die verarbeitende Seite bei FE2 eine Interpolation bzw. ein Downscaling des jeweiligen Datenparameters vorgenommen. Eine detaillierte Betrachtung der Erhebungsintervalle und der Auswirkung bei Änderung wird in Kapitel 2.4.3 erläutert.

2.1.4. Zielabgleich

Die in AP1 laut Antrag definierten Ziele

- Module zur Datenkonvertierung (Umsetzung in AP3) konzipiert (FE 2),
- Daten auf Vollständigkeit und Redundanz geprüft (FE 1+2),
- optimierte Erhebungsintervalle identifiziert (FE 1+2)

wurden mit leichten zeitlichen Verschiebungen erreicht.

2.2. AP2 Datenmodelle und Cloudanbindung

2.1.1. Konzepte zur Zustandsdatenverwaltung

Basierend auf initialen Diskussionen im Projektbegleitenden Ausschuss hinsichtlich potenzieller künftiger bzw. exemplarischer Anwendungsfälle und den bisherigen Erfahrungen der Forschungseinrichtungen wurden verschiedene Eigenschaften für eine zuverlässige und performante Zustandsdatenverwaltung identifiziert und in ein Konzept überführt.

Zunächst ist es erforderlich, die Zustandsdaten mit einer **einheitlichen Beschreibung** zu versehen, welche auch ohne genauere Kenntnis der Datenquelle verständlich ist. Im Projekt wurde mit der Web of Things Thing-Description [9] ein etablierter Standard des W3C verfolgt. Eine solche Beschreibung sollte aus Performance-Gründen nicht dauerhaft mit den Zustandsdaten übertragen werden, sondern wird in einem entsprechenden Repository (z.B. Mongo-DB) hinterlegt. Die Verknüpfung von Zustandsdaten und Beschreibung erfolgt über eine entsprechende ID.

Neben der effizienten Archivierung ist auch ein performantes **Zeitreihen-Handling** inkl. der Verfügbarkeit komplexer Filtermechanismen zur Realisierung von realzeitnahen Anwendungen erforderlich. Hierzu wurde die Zeitreihendatenbank InfluxDB ausgewählt, welche zusätzliche zum effizienten Handling von Zeitreihen auch über eine Integration von Watchdog-Funktionalitäten sowie des Visualisierung-Tools Grafana verfügt.

Weiterhin ist es erforderlich eine flexible **Zugriffskontrolle** zu etablieren, welche einerseits den FE den zur explorativen Analyse erforderlichen Zugriff auf den Datenbestand ermöglicht, andererseits aber die Zugriffsmöglichkeiten auf die Daten der jeweiligen Testanlagen separat und selektiv konfigurierbar macht. Hierzu wurde eine API mit Token-basierten Methoden konzipiert.

2.2.1. Konzepte zur Daten-Übertragung und -Bereitstellung

Basierend auf den Informationen aus der initialen Datenerhebung zu den in den Testanlagen eingesetzt Protokollen und Schnittstellen konnten MQTT und REST als vorrangig verfügbare Ansätze identifiziert werden. Daher wurde ein duales Datenübertragungskonzept formuliert, welches beide Methoden unterstützt. Für einzelne Anlagen war es zudem erforderlich individuelle Ansätze (bspw. Bereitstellung dedizierter MQTT-Broker) abzubilden.

Die eingehenden Daten sind im nächsten Schritt unter Berücksichtigung der in Kapitel 2.1.1 genannten Aspekte strukturiert in eine (Zeitreihen-) Datenbank zu überführen. Nach Evaluation etablierter Parsing-Ansätze wurde für den Projektkontext ein Konzept auf Basis von Node-RED entworfen, welches eine skalierbare Lösung mit Script-basierter oder grafischer Handhabung darstellt.

Die Bereitstellung der aufgenommen Zustandsdaten wurde in Abstimmung zwischen dem Projektkonsortium und den FE über die in AP4 implementierte Datenanalyse-Toolbox umgesetzt. Hierdurch wird unter Einhaltung der definierten Zugriffsmechanismen und Schnittstellen innerhalb der IoL-Cloud eine komfortable Aufbereitung und Analyse der Zustandsdaten je Testanlage unterstützt.

2.2.2. Konzepte zur Cloud-Infrastruktur

Basierend auf den Ergebnissen der vorgelagerten Untersuchungen wurde die Cloud-Infrastruktur konzipiert und prototypisch implementiert (Abbildung 1). Die im Projekt bestehenden Rahmenbedingungen aus dezentralen Datenquellen, zentraler Aufbereitung/Datenhaltung und (perspektivisch) mehrerer nachgelagerter Systeme zur weiteren Datennutzung (Mehrwertdienste) favorisieren eine Cloud-Infrastruktur, welche in die Teilbereiche Datenquellen (links), Datenaufbereitung/-Archivierung(Mitte) und Auswertedienste (rechts) unterteilt ist.

Im Bereich Hersteller/Anlagenbetreiber sind die Testanlagen inkl. der jeweiligen bereitgestellten Gateways / Schnittstellen angeordnet, welche die Anlagendaten für das IoL-WoT-System zur Verfügung stellen. Die dortigen Komponenten werden von den jeweiligen Anlagenbetreibern überwacht und weisen weitgehend eine nicht durch das Projekt IoL beeinflusste (beeinflussbare) Konfiguration auf. Innerhalb des IoL-WoT-Systems sind u.a. die Software-Komponenten für das Payload-Mapping der eingehenden Datenströme auf das WoT-Modell, eine Zeitreihen-Datenbank zur Archivierung der Zustandsdaten, ein Repository für Thing-Descriptions, sowie eine API-Komponente zur Bereitstellung einer gesicherten Schnittstelle zu nachgelagerten Auswerte- und Mehrwert-Diensten angelegt.

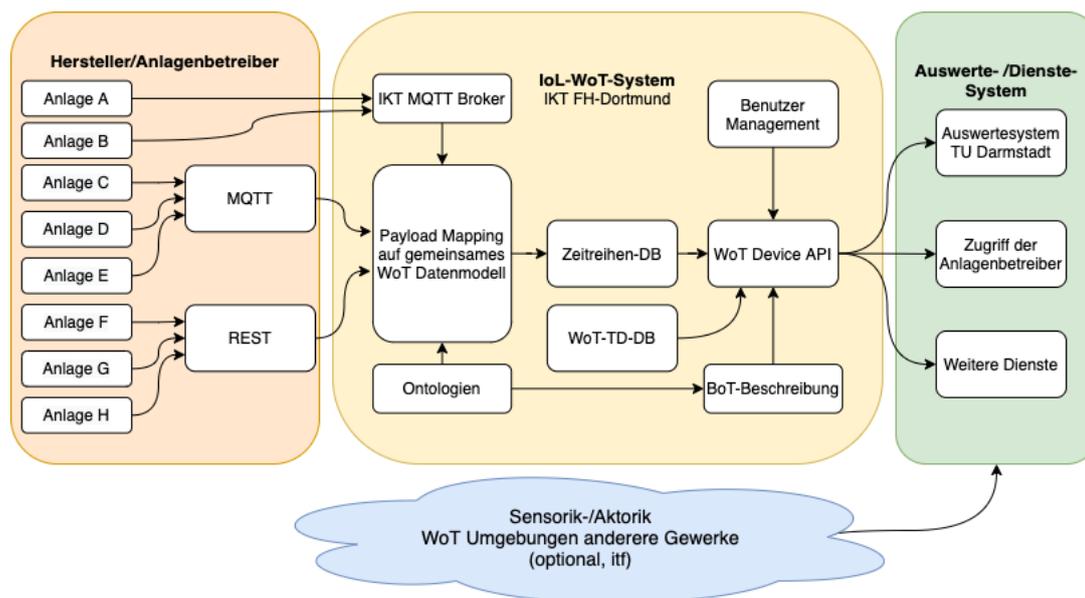


Abbildung 1: Finale Struktur IoL-Plattform.

Einige Anlagen und die TUDa stellen jeweils einen eigenen MQTT-Broker zur Verfügung. Dieser fungiert als zentrale Stelle für den Datenaustausch zwischen den Anlagen und der IoL-Cloud. Die Zustandsdaten werden in dedizierten Topics auf dem Broker verwaltet und mittels eines Publish/Subscribe-Mechanismus an die konfigurierten Endpunkte (hier Payload-Mapping Modul der IoL-Cloud) übermittelt.

Einige andere Anlagen wurden über separate MQTT-Broker innerhalb der IoL-Cloud angebunden und die Daten analog zu den anderen Anlagen über die MQTT-Schnittstelle des Payload-Mapping-Moduls verarbeitet.

Weitere Anlagen bieten eine REST-Schnittstelle über einen eigenen Dienst oder externen Provider an. Hier wurden die optimalen Abrufintervalle individuell für jede Anlage identifiziert und über dedizierte Pull-Mechanismen mit dem Payload-Mapping-Modul verknüpft.

Vorbereitend auf die Verarbeitung der Anlagendaten im IoL-WoT-System waren die WoT- und BoT-Beschreibungen für jede Anlage manuell zu erstellen und mit weiteren Ontologien semantisch anzureichern. Dieser Prozess, welcher im Wesentlichen in AP3 durchgeführt wurde, würde bei einer vollständigen Integration der semantischen Annotation in den Planungs- und Installationsprozess von professionellen Beleuchtungsanlagen, zukünftig durch ein automatisches Mapping ersetzt werden können.

Im Bereich Auswerte-/Dienste-System sind externe Dienste mit Zugriff auf das IoL-WoT-System verortet. Die Abfragen von Zustandsdaten aus der IoL-Cloud werden durch die WoT-Device API verarbeitet. Ein integriertes Benutzer- und Rollenmanagement stellt hierbei den berechtigten Zugriff sicher. Zusätzliche Mehrwertdienste (bspw. dedizierte Analyse-Tools einzelner Hersteller) können ebenfalls über diese Schnittstelle an die IoL-Cloud angebunden werden.

Der **Smart Building Server (SBS)** bildet die zentrale cloudbasierte Backend-Komponente des IoL-WoT-Systems und ist mit einer Microservice-Architektur implementiert. Ein SBS unterstützt eine Vielzahl von verteilten Instanzen (Testanlagen), die wiederum mehrere lokale Controller (Smart Device Controller, SDC) bündeln können. Für den Zugriff auf die Services des SBS stehen

mehrere RESTful APIs zur Verfügung, die über eine Open API Spezifikation konforme Beschreibung definiert sind und in den nächsten Abschnitten näher beschrieben werden.

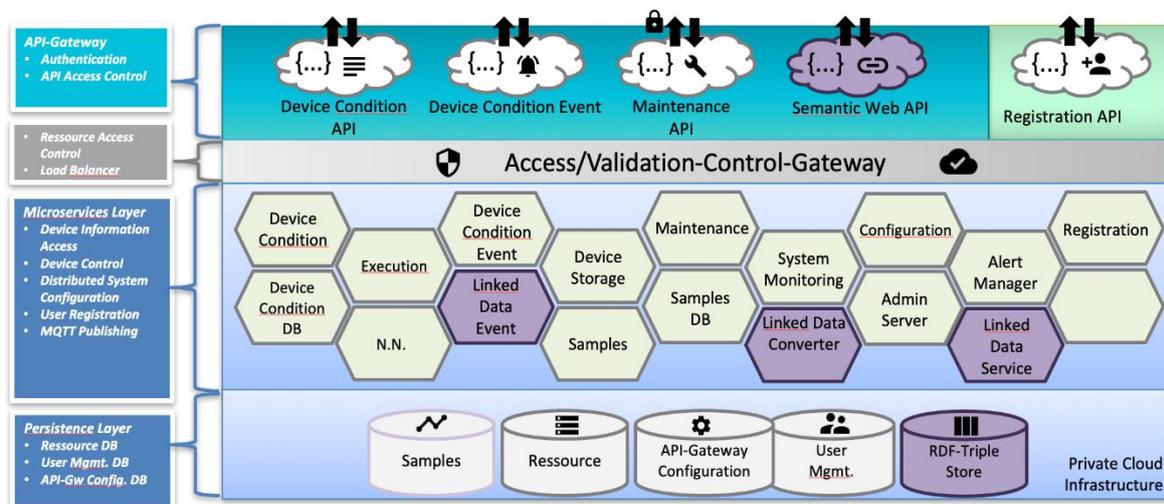


Abbildung 2: Komponenten des SBS.

API Gateway

Der zentrale Einstiegspunkt zur Nutzung der am SBS bereitgestellten APIs wird über ein API Gateway (Application Programming Interface Gateway) verwaltet. Das API Gateway regelt den benutzerorientierten Zugriff auf Private und Public APIs. Restriktionen und entsprechende Security-Mechanismen [10] können durch die Verwendung des API-Gateways an zentraler Stelle implementiert und verwaltet werden. Die nachgelagerten Microservices müssen keine spezielle Zugriffsberechtigung erhalten, da sie ausschließlich über das API-Gateway erreichbar sind. Somit können bspw. Sicherheitsupdates zentral und zeitnah eingespielt werden. Ein Beispiel für einen restriktiven API-Zugriff könnte die Nutzung der Maintenance API sein. Der Zugriff auf dieses API ist nur über eine Administrator Rolle möglich. Ein weiterer Vorteil des API Gateways ist die zentralisierte Überwachung der Systemzustände durch Funktionen wie Logging, Traffic Control oder IP-Restriction. Der IP-Security wird durch die Zertifikatsverwaltung und Absicherung der Kommunikation über TLS Rechnung getragen. Zur Bandbreitenbegrenzung wird für die Funktion des Traffic Controls die sogenannte Rate Limiting Methode eingesetzt. Eine weitere Funktion des API Gateways ist die Orchestrierung der verfügbaren Services.

Authentication API

Die Authentication API dient der eindeutigen Identifikation eines zuvor registrierten Benutzers. Ein Benutzer meldet sich unter Verwendung seiner Benutzername-Passwortkombination an der Login-Ressource an und erhält bei gültiger Kombination einen Access Token. Dieser Token wird bei der Nutzung aller anderen APIs zur Authentifizierung verwendet. Dieser Access Token identifiziert den jeweiligen Benutzer gleichermaßen wie die Benutzername-Passwortkombination, hat aber den Vorteil, dass der Benutzer seine Passwörter nicht für die Anbindung externer Plattformen benötigt. Diese Vorgehensweise ist an das in der Industrie etablierte OAuth-Verfahren angelehnt und verhindert die unnötige Verteilung sensibler Benutzerdaten wie bspw. das geheime Benutzerpasswort.

Der Access Token ist als geheim anzusehen und muss deshalb in jedem Fall sicher aufbewahrt bzw. gespeichert werden. Ist ein Access Token dennoch in unbefugte Hände geraten, kann dieses Token über die Authentication API zurückgezogen werden. In diesem Fall wird das kompromittierte Token an die logout-Ressource gesendet. Ein nachgelagerter Dienst entfernt daraufhin das Token aus der Liste der gültigen Access Token.

Da jedes Access Token nur für einen fest definierten Zeitraum gültig ist, existiert eine refresh-Ressource um das Token zu erneuern. Zur Authentifizierung wird das „noch“ gültige Access Token an die refresh-Ressource gesendet. Als Antwort erhält der Benutzer ein neu generiertes Access Token mit identischer Gültigkeitsdauer. Das alte Access Token ist ab diesem Zeitpunkt ungültig und kann nicht mehr verwendet werden.

Zum Generieren der Access Token wird der JSON¹ Web Token (JWT) Standard verwendet. Ein JWT bietet bereits die Möglichkeit grundlegende Informationen wie Aussteller, Ausstellungsdatum und Gültigkeitsdauer im Token zu speichern. Durch ein digitales Signaturverfahren kann eine unbefugte Manipulation des JWT weitgehend ausgeschlossen werden.

Device Condition API

Die Device Condition API (DCA) ermöglicht einen Zugriff auf die Geräteinformationen des jeweiligen Benutzers. Jeglicher Zugriff ist durch eine Benutzerauthentifizierung geschützt und kann deshalb nur im Benutzerkontext erfolgen. Die Schnittstellen dieses APIs sind in Form eines RESTful WebServices und damit ressourcen-orientiert umgesetzt. Die Repräsentation der Geräteinformationen erfolgt in Form des DeviceData-Modells. Die Device Condition API ist die primäre Webservice Schnittstelle für externe Plattformen und Dienste, die Geräte-basierte Services bereitstellen. Im Projekt IoL wird die DCA genutzt, um die allgemeinen Geräteinformationen und die aktuellen Zustandswerte der Sensoren weiter zu verarbeiten und Visualisierungen der Gebäudeinstallation zu realisieren. Weitergehend können in Kombination mit anderen Datenquellen (bspw. Wettervorhersage) intelligente Services (Smart Services) bereitgestellt werden, die sich an typischen Use Cases einer intelligenten Beleuchtungslösung orientieren.

Device Condition Event API

Um außerhalb des IoL-Backends eine Event-Stream-basierte Datenkommunikation zu ermöglichen, wird für nachgelagerte Systeme am SBS ein WebSocket-Endpunkt zur Verfügung gestellt. Verbindungen können nach einer Client-Authentifizierung über einen gültigen Access Token hergestellt werden. Anschließend werden auf dem TLS gesicherten Kanal alle Zustandsänderungen der jeweiligen Datenpunkte des angemeldeten Benutzers publiziert (Push-Prinzip). Diese Zustandsänderungen werden durch das Senden von Parameter-Value-Objekten mit JSON Encoding realisiert. Die Datenkommunikation auf dem WebSocket-Endpunkt findet ausschließlich unidirektional als Event-Bus vom SBS zum Client statt.

¹ JSON Web Token Standard (JWT) - RFC 7519

Maintenance API

Das Maintenance API ist eine administrative Webservice Schnittstelle, die interne Daten des IoT-Backends und dessen Teilsysteme über eine einheitliche und mit der Open API Spezifikation dokumentierte Schnittstelle verfügbar macht. Restriktive Zugriffsberechtigungen verhindern den unbefugten Zugriff auf das Maintenance API.

Über das API können unter anderem Log-Files aller Teilkomponenten abgefragt werden. Dazu muss eine Logging-Ressource unter Angabe der gewünschten Teilkomponente per HTTP-GET adressiert werden. Durch dynamische Anpassung des Log-Levels können je nach Bedarf Detailinformationen der einzelnen Komponenten ermittelt werden.

Um die Hardwareauslastung des SBS und dessen Teilkomponenten zu ermitteln, existiert eine sogenannte Runtime-Statistic-Ressource. Diese Ressource liefert bei einer HTTP-GET Anfrage detaillierte Information in Bezug auf die Hardwareauslastung der Teilsysteme. Es können derzeit unter anderem die CPU-Auslastung, die Speicherauslastung und der verwendete Swap-Speicher abgefragt werden. Die Runtime-Statistic-Ressource bietet somit eine Möglichkeit Engpässe im System frühzeitig zu erkennen und entsprechende Gegenmaßnahmen einzuleiten.

Eine weitere Aufgabe des Maintenance API ist das Einspielen von Software-Updates in den verteilten CPP Systemkomponenten, um bspw. Sicherheitslücken in verwendeten Software-Modulen zu beheben.

Die Configuration-Ressource ermöglicht das Einspielen von System-Konfigurationen für alle Systemkomponenten der SBS.

Device Condition DB

Als Datenbanksystem für den SBS dient in der Implementierung eine SQL-fähige relationale Datenbank. Die Datenbank ist als Microservice implementiert und läuft in einem Docker-Container, der über eine Service-Registry adressiert werden kann. Dieser Container ist aufgrund der bestehenden Konfiguration nur innerhalb des Docker-Netzwerkes und damit für Microservices im selben Netzwerk erreichbar. Zugriffe auf die Device Condition DB sind nur nach einer vorhergehenden Authentifizierung möglich.

Die Device Condition DB speichert die DeviceData-Objektstruktur. Diese Struktur beinhaltet unter anderem Geräteinformation, Sensor-Zustände und Gebäudeinformation. Da diese Daten immer im Kontext des Benutzers zu sehen sind, wird zusätzlich eine eindeutig Benutzer-ID mit den zugehörigen Nutzdaten verknüpft. Neben den DeviceData-Strukturen werden Konfigurationen aller Teilsysteme in der Device Condition DB gespeichert. Der Zugriff auf die Konfigurationsdaten ist nur durch die Administrator-Rolle möglich.

Lesende oder schreibende Datenbankzugriffe von anderen Microservices erfolgen über fest definierte RESTful Webservice Schnittstellen zu einem vorgelagerten Datenbank-Service. Durch diese Vorgehensweise wird nur an einer Stelle im System eine direkte Datenbankverbindung aufgebaut. Die Verbindung ist durch die Verwendung von TLS gesichert. Der Handshake im TLS-Verbindungsaufbau erfordert eine beidseitige Authentifizierung in Form eines Zertifikataustausches.

Authorization DB (User/API Access)

Die Authorization DB speichert Benutzerprofile in einer relationalen Datenbank. Da diese Datenbank sensible Informationen archivierte, sind an dieser Stelle hohe Sicherheitsanforderungen bezüglich des Datenschutzes einzuhalten. Eine Sicherheitsmaßnahme ist die Trennung der Datenbank von den anderen Datenbank-Systemen. Weitere Sicherheitsmechanismen werden unter anderem in Bezug auf die Zugriffsrestriktionen und die Art der Datenspeicherung vorgenommen. Einige dieser Mechanismen werden im Folgenden exemplarisch dargestellt.

Ein Benutzer registriert sich initial über den sogenannten Registration Service am IoL-Backend. Bei der Registrierung wird ein Benutzerprofil angelegt, das mindestens die Grundinformationen (Email-Adresse, Alias und ein Passwort) enthält. Zusätzlich können optional weitere Angaben wie Vor- und Nachname und Adresse gemacht werden. Das Passwort des Benutzers wird als sogenannter Salted Hashwert gespeichert. Dadurch ist das Klartext-Passwort innerhalb des Backends nicht bekannt. Bei einer Anmeldung eines Benutzers wird lediglich der errechnete Hashwert mit dem Wert in der Datenbank verglichen. Jedem Benutzer wird eine eindeutige Customer Identity (ID) zugeordnet. Diese ID wird zur Verknüpfung zwischen einem Benutzer und seinen in der Device Condition DB gespeicherten Daten verwendet.

Die Authorization DB steht in Form eines Docker-Containers zur Verfügung und wird nur durch einen zugehörigen Authorization DB Service in die Microservice Umgebung eingebunden. Dieser Service authentifiziert sich durch eine Benutzername/Passwort-Kombination bei der Datenbank. Die Verbindung zwischen dem Service und der eigentlichen Datenbank ist durch TLS verschlüsselt.

Benötigt ein Microservice Daten der Authorization DB, müssen diese Daten über die RESTful Webservice Schnittstelle des Authorization Service abgerufen werden.

Authorization Service

Der Authorization Service ist für die Erzeugung von Access Token und API-Keys zuständig. Dieser Service bildet damit die Grundlage für die zugehörige Authentication API. Soll ein neues Access Token ausgestellt werden, muss zunächst die Identität des Benutzers per Benutzername/Passwort-Kombination sichergestellt werden. Nach Empfang der Login-Daten wird zunächst eine Verbindung zur Authorization DB aufgebaut. In diesem Schritt werden der Benutzername und der Hashwert des Passworts mit der Datenbank abgeglichen. Ist diese Kombination gültig, wird ein neues Access Token generiert und mit der Authorization DB synchronisiert. Zusätzlich muss das Access Token der Device Condition DB, welches für den Gerätezugriff im Nutzerkontext benötigt wird, durch das neue Access Token ersetzt werden. Anschließend wird das ungültige Access Token aus der Authorization DB entfernt.

Message Broker

Als Message Broker des SBS wird eine MQTT Plattform verwendet. Der Broker ist so konfiguriert, dass eine Authentifizierung für jeden Client notwendig ist. Zusätzlich sind die Kommunikationskanäle des Brokers per TLS gesichert. Der Broker bildet die Kernkomponente für eine Event-Stream basierte Datenkommunikation zwischen den Teilkomponenten des SBS. Der Message Broker ermöglicht eine Verteilung der Geräte-, Konfigurations- und

Gebäudezustandsdaten, die als Broker-Clients fungieren. Die Clients können Datenquellen (Publisher) oder Datensinken (Subscriber) sein. Ein typischer Publisher ist eine Testanlage, die Zustandsdaten von Sensoren und Aktoren als Parameter-Objekt auf das zugehörige Parameter-Topic sendet. Ein Topic ist ein Pfad für einen bestimmten Informationskanal, der von einem oder mehreren Subscribern adressiert werden kann. Möchte eine Teilkomponente alle Messages eines Topics erhalten, können Teile des Topic-Pfades mit sogenannten Wildcards versehen werden. So kann bspw. der Device Persistence Service alle Updates der DeviceData-Objekte empfangen und persistieren.

Ein weiteres Einsatzgebiet des Message Brokers sind Remote Procedure Calls (RPC), die vom SBS zum Schalten von Aktoren eingesetzt werden. Die RPC Technik bietet in diesem Zusammenhang die Möglichkeit synchrone Schalthandlungen an die verteilten SDCs zu senden. Jeder RPC Aufruf erhält eine synchrone Rückantwort. Somit kann zu jeder Zeit bestimmt werden, ob eine Schalthandlung zwischen zwei Endpunkten erfolgreich durchgeführt wurde oder in der Kommunikationskette ein Fehler aufgetreten ist.

2.2.3. Anbindung der Testanlagen an IoL-Cloud

Nachfolgend wird der Prozess zur Anbindung der von den Mitgliedern des Projektbegleitenden Ausschusses zur Verfügung gestellten Testanlagen beschrieben.

Initial wurden die in AP1 erstellten Fragebögen zur Erhebung der Anlagenparameter, Protokolle und Schnittstellen, welche durch die technischen Ansprechpartner bereitgestellt wurden, evaluiert. Anschließend wurden anhand der aus dem Fragebogen gewonnenen Informationen die API-Dokumentationen und das vom Anlagenbetreiber verwendete Datenmodell analysiert.

Da Pandemie-bedingt nicht alle Testanlagen kontinuierlich für eine Anbindung an die IoL-Cloud bereitstanden, wurden die verschiedenen Anlagen in einem sukzessiven Prozess nach Verfügbarkeit angebunden. Diese Abweichung gegenüber dem laut Antrag geplanten Vorgehen konnte durch parallele bilaterale Abstimmungen mit den technischen Ansprechpartnern auf Herstellerseite weitgehend kompensiert werden. So konnten die BoT-Beschreibung (basierend auf den Geometriedaten der Anlagen) sowie die WoT-Beschreibung initial ohne bestehende Kommunikation mit der jeweiligen Testanlage angelegt werden.

Mit Anbindung der entsprechenden Testanlage wurden die vorbereiteten WoT- und BoT-Beschreibungen gegen die Live-Daten der Anlage validiert und ggf. Anpassungen vorgenommen und sodann basierend auf den finalen Beschreibungen das individuelle Payload Mapping implementiert. Dieses Mapping berücksichtigt unter anderem Anlagenbesonderheiten bei der Datenübertragung, dem Zeitstempel und Datenformat. Mit vorliegendem Mapping wurden die eingehenden Zustandsdaten kontinuierlich in eine Zeitreihen-Datenbank persistiert, und mit automatisierten Prüfroutinen gekoppelt.

2.2.4. Erstellung und Implementierung eines einheitliches Datenformats

Im Verlauf des Projektes wurden die bereitgestellten Anlagen sukzessive angebunden. Somit konnte eine vollständige Übersicht über die Gesamtheit der in den Testanlagen verwendeten Datenmodelle entgegen der ursprünglichen Planung erst im letzten Drittel der Projektlaufzeit erstellt werden. Mit steigender Anzahl angebundener Anlagen zeigte sich jedoch bereits, dass auch bei vergleichbaren Parametern und gleicher Technologie große Unterschiede in der herstellerseitig gewählten (Daten-)Modellierung bestehen.

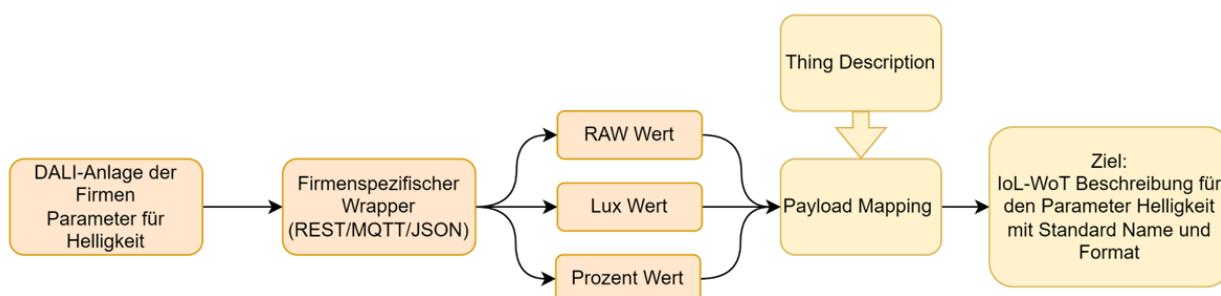


Abbildung 3 Beispiel unterschiedlicher Modellierungsstrategien.

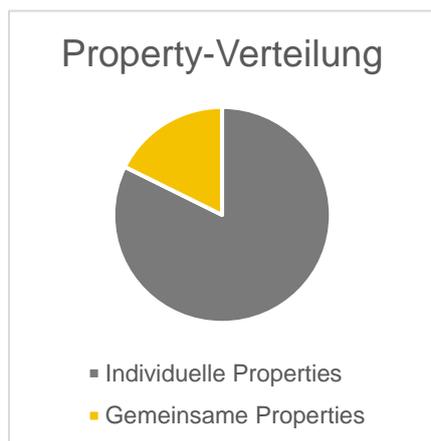
Um eine allgemeine Beschreibung in WoT zu realisieren, wurden die Properties aller angebotenen Anlagen hinsichtlich ihrer Funktionen, dem verwendeten Datentyp, der Beschreibung und soweit vorhanden der angegebene Minimal- und Maximal-Werte verglichen.

Wenn übereinstimmende Properties zwischen mindestens zwei Anlagen festgestellt wurden, wurde für die Property ein gemeinsamer Name gewählt, aus den verwendeten Datentypen ein gemeinsamer gebildet, eine gemeinsame Beschreibung geschaffen und wenn möglich gemeinsame Minima und Maxima gebildet. Abbildung 4 zeigt hierzu einen Auszug aus der Übereinstimmungsmatrix.

Property Manufacturer_A	dimmingLevel_current	operatingTime_value
Beschreibung Manufacturer_A	Computed dimming level in percent	Operating time in minutes
Property Manufacturer_B		
Beschreibung Manufacturer_B		
Property Manufacturer_C	led_brightness	operation_time_dev_total
Beschreibung Manufacturer_C	led brightness in %	total operation time in hours
Property Manufacturer_D		
Beschreibung Manufacturer_D		
Property Manufacturer_E	dimmer_value	timeLEDModul
Beschreibung Manufacturer_E	Current dim value	Operating time LED module
Property Manufacturer_F	lightLevel	lampOperationTime
Beschreibung Manufacturer_F	DALI control gear light level	
Schema	integer; integer; integer; integer	integer; number; integer; integer
min	0; 0; 0; 0	0; x; 0
max	100; 100; 100; 100	21247483647; x; 2147483647
Summe Ähnliche P.	4	4
Allgemeine Property	dimmingLevelPercent	operatingTime
Allgemeine Beschreibung	Dimming Level in Percent	Operating Time
Gemeinsames Schema	integer	number
min	0	0
max	100	21247483647

Abbildung 4 Auszug vereinheitlichtes Datenformat.

Die so erstellten Property Beschreibungen stellen die Basis für ein einheitliches Datenformat dar und können zusammen mit der IoL-WoT Beschreibung in künftigen Projekten Verwendung finden. Das Vorgehen, die gewählten Bezeichnungen und Beschreibungen sind ein Vorschlag für ein einheitliches Datenmodell und werden für eine weitere Diskussion im ZVEI-Fachverband Licht sowie für weitere Interessenten bereitgestellt.



Insgesamt wurden bei dieser Analyse 34 verschiedene Properties identifiziert, wovon jedoch lediglich 6 Properties von mindestens 2 Anlagen verwendet wurden. Die verbleibenden 28 Properties ließen sich aufgrund ihrer Diversität nicht weiter kombinieren. Zur Wahrung der Funktionalität eines einheitlichen Datenformates werden diese individuellen Properties vollständig im einheitlichen Datenmodell repliziert. Perspektivisch empfiehlt es sich, branchenweit beispielsweise einen Katalog von Basis-Properties (z. B. Dimmwert in Prozent, siehe Abbildung 4) zu definieren, um den Anteil übereinstimmender Properties zu steigern. Hiervon unberührt bleibt die grundsätzliche Vorgehensweise, individuelle Properties (bspw. zur Abbildung herstellerspezifischer Features) in einem hierfür vorgesehenen Bereich des Datenmodells einzugliedern.

2.2.5. Zielabgleich

Die in AP2 laut Antrag definierten Ziele

- Konzepte zur Zustandsdatenverwaltung (FE 2), Daten-Übertragung und -Bereitstellung (FE 1+2), sowie zur Cloud-Infrastruktur (FE 2) erstellt,
- Testanlagen an IoL-Cloud angebunden (FE 2)
- Einheitliches Datenformat erstellt und prototypisch implementiert (FE 1+2)

wurden erreicht.

2.3. AP3 Einheitliche Semantik

In diesem AP wurden die Untersuchungen zur Überführung der aktuell herstellerseitig verwendeten Datenmodelle in Richtung eines einheitlichen, semantisch angereicherten Datenmodells durchgeführt. Die in Abbildung 5 skizzierten Verknüpfungen zwischen Device, Function, Location und Hardware basieren auf Erkenntnissen der Projektes SENSE [6] und ermöglichen eine maschinenlesbare Abbildung von Zusammenhängen, welche u.a. für die Umsetzung von Mehrwertdiensten (z.B. Human Centric Lighting, Predictive Maintenance) erforderlich sind.

Die gezeigte Struktur ist das Ergebnis einer mehrstufigen Iteration und deckt in ihrer aktuellen Version die Anforderungen der in IoL betrachteten Use Cases vollständig ab. In den folgenden Abschnitten werden die in diesem Kontext durchgeführten Arbeiten näher beschrieben.

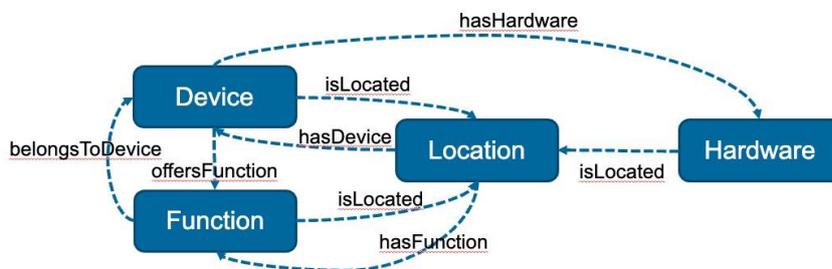


Abbildung 5 Semantische Verknüpfung von Geräteinformationen im Kontext von WoT.

2.3.1. Identifikation von Ontologien für die Anwendungsdomäne Licht

Die Basis für eine einheitlich semantische Beschreibung der Daten wird durch das bereits in Vorgängerprojekten verwendete WoT-Modell realisiert. Im Rahmen des BMWi-Projektes SENSE hat sich bereits gezeigt, dass dieses durch seine Flexibilität und Erweiterbarkeit um externe Ontologien für die Anwendung im Projekt IoL geeignet ist.

Der Ansatz im Projekt SENSE hat sich aufgrund der Thematik vorrangig auf Sensorik und Aktorik im Kontext der Gebäudeautomation konzentriert und war in der ursprünglichen Form nicht dafür ausgelegt Leuchten qualitativ und umfassend zu beschreiben. Das Ziel der initialen Recherche in IoL war es deshalb, eine Ontologie für die Domäne Licht zu identifizieren und deren Eignung in Kombination mit dem WoT-Ansatz für das Projekt IoL zu verifizieren. Im Rahmen der Ontologie-Recherche bestätigte sich jedoch die ursprüngliche Annahme, dass eine tragfähige Lösung zur Beschreibung von professionellen lichttechnischen Anlagen nur durch eine geeignete Kombination mehrerer Ontologien zu erreichen war. Nachfolgend werden die als für IoL geeignet identifizierten Ontologien aufgeführt und die wesentlichen Eigenschaften zusammengefasst. Unter Verwendung dieser Ontologien wurden in den nachgelagerten Arbeitspaketen die semantischen Annotationen der Parameter der jeweiligen Testanlagen umgesetzt.

Die Ontologie Schema.org [11] ist ein öffentlich gepflegtes Projekt mit dem Ziel Daten strukturiert im Internet auf Webseiten, in elektronischen Dokumenten und weiteren Bereichen anzureichern und zu verwalten. Die Begriffe der Ontologie können aufgrund der interoperablen Struktur in relationalen Strukturen wie RDF oder JSON-LD verwendet werden. Schema.org wird bereits in vielen Diensten von Google, Microsoft oder Pinterest zum Anreichern und zur Kategorisierung der Daten verwendet.

Iotschema.org [12] ist eine Erweiterung der Ontologie schema.org. Ziel der Ontologie ist es, Webanwendungen die Interaktion mit der physischen Welt auf der Grundlage maschinell interpretierbarer Informationen zu ermöglichen. Die semantische Interoperabilität für vernetzte Dinge ermöglicht es diese in verschiedenen IoT-Ökosystemen zu integrieren. Die Vokabeln für die Ontologie Iotschema.org werden in einem offenen Community-Prozess, unter Verwendung der W3C Community Group Schema Extensions For IoT sowie über GitHub weiterentwickelt.

OM oder Ontology of units of Measure ist eine Ontologie, welche Konzepte und Abhängigkeiten für wissenschaftliche Forschung modelliert [13]. Die Ontologie hat den Fokus auf die semantische Beschreibung von Einheiten, Mengen, Messungen und Dimensionen, welche für die genaue Abbildung von Forschungsergebnissen relevant ist. Ergänzend zu standardmäßig beschriebenen Einheiten in anderen Ontologien können über diese auch abgeleitete SI-Einheiten oder spezielle Energieeinheiten abgedeckt werden.

Die Smart Application REFerence Ontology (SAREF) [14] hat das Ziel Interoperabilität zwischen Lösungen mehrerer Anbieter in unterschiedlichen Bereichen des Internet Of Things (IoT) zu ermöglichen und somit zur Entwicklung eines globalen digitalen Marktes beizutragen. SAREF setzt auf eine Struktur rund um ein Device. Dies kann ein Sensor/Aktuator oder weiße Ware sein und besitzt Eigenschaften wie Funktionen oder Interaktionsmöglichkeiten in Form von Befehlen.

Saref4Ener [6] ist eine Erweiterung der Ontologie SAREF für die Domäne Energie. Es wurde in Kooperation mit Energy@Home [15], EEBus [16] und großen Industrieverbänden in Deutschland und Italien erarbeitet, um eine Verbindung der unterschiedlichen Datenmodelle für die einheitliche Verwendung von smarten Geräten wie Smart Grids zur Energieverwaltung im Haushalt zu ermöglichen.

Die Verwendung der genannten Ontologien war erforderlich, um die vorliegenden Parameter der Testanlagen in Summe eindeutig zu beschreiben. Im Einzelfall können Anlagen mit weniger komplexen Parameter-Strukturen / einfacheren Use-Cases auch unter Verwendung eines Subsets der oben genannten Ontologien hinreichend beschrieben werden.

2.3.2. Erweiterung des WoT-Modells

In AP2 wurde das WoT-Modell, welches auf Erkenntnissen aus dem BMWi-Projekt SENSE basiert, für die Anforderungen im Projekt IoL erweitert. Diese Erweiterung umfasst lichttechnische Parameter für die WoT-Thing-Description zur Beschreibung der Geräte und eine BoT-konforme Struktur zur Beschreibung der Liegenschaft. Nachfolgend werden diese Erweiterungen näher erläutert.

WoT-Thing-Description

Jedes WoT Thing wird durch einen Ressource Identifier in Form eines Uniform Resource Identifier (URI) identifiziert. Diese URI wird benutzt um einen Digitalen Zwilling [17] von Aktoren, Sensoren und Informationsdiensten zu identifizieren. Die Dinge (Things) werden so konsumierenden Anwendungen als lokale Softwareobjekte mit Eigenschaften, Aktionen und Ereignissen zur Verfügung gestellt. Dies geschieht unabhängig vom physischen Standort oder dem für den Zugriff auf sie verwendeten Protokollen. Die Beschreibung eines Gerätes oder Services per Thing Description liegt üblicherweise im Format JSON-LD vor, und umfasst neben der Beschreibung grundlegender Geräteeigenschaften auch das Verhalten, die Interaktionsmöglichkeiten, das Datenschema, die Sicherheitskonfiguration und die Protokollbindungen. Das Registrieren, Abrufen und Löschen von Thing Descriptions wird mittels eines WoT-Thing-Directory [18] über eine definierte REST-API realisiert.

Im Folgenden wird der allgemeine Aufbau einer WoT Thing Description anhand eines Beispiels beschrieben.

Das Feld mit dem Namen **context** bietet die Möglichkeit die Thing Description (TD) um zusätzliche Vokabular-Begriffe zu erweitern. Dies bildet die Grundlage für semantische Annotationen und Erweiterungen von Kernmechanismen wie Protokollbindungen, Datenschemata und Sicherheitsschemata.

Das nächste Feld **id** enthält die URI die das Thing eindeutig identifiziert.

Der **title** bietet einen lesbaren Titel für das Thing zum Beispiel für die Anzeige in einer Bedienoberfläche.

Die Felder **security** und **securityDefinitions** gehören zusammen. In den **securityDefinitions** werden unter Verwendung von Namen, hier `bearer_sc`, die Zugriffskontrollmechanismen spezifiziert. Erst durch das Feld **security**, durch Angabe des Namens, wird eine Security Definition aktiviert. Dies kann auf verschiedenen Ebenen des Things geschehen. Die Verwendung mindestens einer Security Definition auf der obersten Ebene ist Pflicht.

Das Feld **@type** kann auf unterschiedlichen Ebenen auftauchen und enthält semantische Schlagwörter (Tags), welche die aktuelle Ebene des Things beschreiben. Zum Beispiel auf der obersten Ebene kann beschrieben werden, ob das Thing ein Aktor oder ein Sensor ist.

Die **description** beinhaltet eine zusätzliche Information und kann auf unterschiedlichen Ebenen gefunden werden. Je nach Ebene kann die Information unterschiedliche Inhalte haben, zum Beispiel die Beschreibung der möglichen Befehle, die eine Action akzeptiert.

version kann Versionsinformationen über das Thing enthalten. Dies beschränkt sich nicht nur auf die Version der TD, sondern es ist auch möglich zusätzliche Informationen wie Versionsnummern für Hardware und Firmware des Things zu hinterlegen.

In den Feldern **created** und **modified** ist es möglich Datumsstempel abzulegen. In **created** den Zeitpunkt der Erstellung des Things und in **modified** das Datum der letzten Veränderung der Beschreibung.

Das Feld **base** enthält eine Basis-URI, die für alle relativen URI-Referenzen verwendet wird.

Die folgenden Felder werden in der Beschreibung von WoT als Interaction Affordances zusammengefasst, sie heißen **property**, **action** und **event**. Sie sind das Kernstück eines Things, da ohne sie keine Funktion gegeben ist. Ein Thing sollte mindestens eine der drei Interaction Affordances enthalten. Alle drei enthalten das Feld **forms** das einen Satz von Formular-Hypermedia-Steurelementen enthält, die beschreiben, wie eine Operation durchgeführt werden kann. Ebenfalls enthalten alle hier genannten Interaction Affordances ein Schema, welches semantisch die Funktion der jeweiligen Affordance beschreibt. Weitere mögliche Felder sind **title**, **description**, **readOnly** und **writeOnly**. Die letzteren beiden zeigen an, ob die jeweilige Interaction Affordance nur gelesen oder geschrieben werden kann oder ob beides möglich ist.

Die Property Affordances unter dem Schlagwort **property** beschreiben die Zustände des Things. Ein Thing kann mehrere Property Affordances haben. Zum Beispiel die unterschiedlichen Funktionen einer Leuchte: Farbtemperatur, Dimmlevel, Temperatur werden jeweils eine eigenständige Property Affordances geführt.

Unter **action** werden im Thing die Action Affordances beschrieben, die es erlauben eine Funktion des Things aufzurufen. Ein Beispiel hierfür ist das Abrufen von Historien-Daten.

Das letzte Feld in dieser Auflistung ist **hardware**. Hier kann die Hardware beschrieben werden die physisch die Funktionen, die im Thing beschrieben werden, zur Verfügung stellt. Im Projekt wurde dies durch Verlinkung auf die Produkt-Website des Herstellers realisiert.

BoT-Beschreibung

Für den Kontext der Betrachtung und der Erarbeitung von Use-Cases zum Thema Licht ist eine räumliche Zuordnung sowie die Kenntnis von natürlichen Lichtquellen im Raum erforderlich. Im WoT-Modell kann eine entsprechende räumliche Zuordnung ohne eine externe Ontologie jedoch nicht beschrieben werden.

In der Planungs- und Bauphase von Gebäuden gewinnt BIM (Building Information Modeling), eine Arbeitsmethode für die vernetzte Planung, Bau und Bewirtschaftung von Gebäuden, welche alle Gewerke mit einbezieht, zunehmend an Bedeutung. In BIM werden alle relevanten Bauwerksdaten digital modelliert und in virtuellen Modellen des Bauwerks unterschiedliche Aspekte visualisiert. In diesem Kontext stellt die Building Topology Ontology (BoT) eine geeignete, BIM-konforme Lösung zur Beschreibung der räumlichen Anordnung dar [19].

BoT ist eine semantische Beschreibung angelehnt an das Verfahren BIM und beschreibt die Basisstruktur des Bauwerks über eine hierarchische, maschinenlesbare Struktur. Dabei wird das Architekturmodell beschrieben, da es alle Informationen zum Gebäude und dessen Unterteilung enthält.

Das Konzept von BoT setzt auf Zonen in der virtuellen oder realen Welt mit dreidimensionaler Ausdehnung, welche alle Aspekte des Gebäudes beschreiben. Die Site beschreibt das Grundstück, auf dem ein oder mehrere Gebäude stehen. Das Building ist die unabhängige Einheit des Gebäudes mit individuellen Eigenschaften und Struktur. Ein Storey stellt eine Etage im Gebäude dar und fasst alle Elemente darauf zusammen. Der Space ist ein dreidimensionaler Raum mit Begrenzung zugehörig zu einer Etage im Gebäude. Im Folgenden wird die BoT-Beschreibung am Beispiel der Testanlage der TuDa (Abbildung 6) erläutert.

```

@context : [ 1 item ]
▼ 0 : {
  bot : https://w3id.org/bot#
  rdf : http://www.w3.org/1999/02/22-rdf-syntax-ns#
  wot : https://www.w3.org/2019/wot/td#
  product : https://w3id.org/product#
  saref : https://saref.etsi.org/core/
  schema : https://schema.org/
  geo : https://pur1.org/geojson/vocab#
}
]
@id : https://ikt-dortmund.de:8443/ict-gw/v1/sites/TU_Darmstadt
@type : [ 1 item ]
schema:name : TU_Darmstadt
bot:hasBuilding : [ 1 item ]
▼ 0 : {
  @id : https://ikt-dortmund.de:8443/ict-gw/v1/sites/TU_Darmstadt/buildings/TU_Darmstadt_building
  ▶ @type : [ 1 item ]
  schema:name : TU_Darmstadt_building
  ▼ bot:hasStorey : [ 1 item ]
  ▼ 0 : {
    @id : https://ikt-dortmund.de:8443/ict-gw/v1/sites/TU_Darmstadt/buildings/TU_Darmstadt_building/storeys/ground_floor
    ▶ @type : [ 1 item ]
    schema:name : ground_floor
    ▼ bot:hasSpace : [ 1 item ]
    ▼ 0 : {
      @id : https://ikt-dortmund.de:8443/ict-gw/v1/sites/TU_Darmstadt/buildings/TU_Darmstadt_building/storeys/ground_floor/spaces/test_room
      ▶ @type : [ 1 item ]
      schema:name : test_room
      ▶ bot:hasElement : [ 2 items ]
      ▶ bot:adjacentElement : [ 1 item ]
      ▶ geo:geometry : { 3 props }
      ▶ view3D : { 1 prop }
    }
  ]
  ▶ geo:geometry : { 3 props }
}
]
▶ bot:hasElement : [ 1 item ]
▶ geo:geometry : { 3 props }
}
]
geo:geometry : { 3 props }

```

Abbildung 6: BoT der TuDa-Anlage.

Jede Zone der BoT-Beschreibung weist neben der individuellen Unterstruktur eine @id, @type, schema:name und geo:geometry auf. Die @id gibt über den eindeutigen Pfad die Zuordnung auf dem Grundstück und im Gebäude an. Über den @type wird festgelegt, um welchen Typ von Zone es sich handelt. Mit schema:name wird die menschenlesbare Bezeichnung der Zone angegeben. In geo:geometry wird die Fläche der Zone als Polygon beschrieben und dazu verwendet, um die Dimensionen in 3D-Koordinaten modellierbar zu machen.

Der @context wird analog zum context-Attribut im Abschnitt zur WoT-Beschreibung verwendet und beschreibt, um welche Vokabular-Begriffe diese erweitert wird. Speziell für die BoT wird die Ontology Product eingeführt, da über diese Elemente wie Heizungen, Türen und Fenster korrekt beschrieben werden können.

Mit der @id wird auf den Speicherort der BoT-Beschreibung in der WoT-IoT-DB verwiesen. Auch wenn der Link offen angegeben wird, ist ein direkter Zugriff nur über das interne Netzwerk von FE2 oder extern über die WoT-Device API möglich. In schema:name wird der menschenlesbare Name TU_Darmstadt angegeben.

Über bot:hasBuilding wird das Gebäude mit Bezeichnung TU_Darmstadt_building als Unterelement der Site TU_Darmstadt beschrieben. Dies wird analog dazu in der @id sites/TU_Darmstadt/buildings/TU_Darmstadt_Buildings sichtbar. Da das Gebäude nicht so groß wie das Grundstück ist, variiert hier auch die Dimension beschrieben durch geo:geometry.

Über bot:hasStorey wird dem Gebäude TU_Darmstadt_building die Ebenenstruktur über Etagen hinzugefügt.

Die @id sites/TU-

Darmstadt/buildings/TU_Darmstadt_Buildings/storeys/ground_floor zeigt durch die Struktur, dass die Etage ground_floor im Gebäude verortet ist. Aufgrund der Gebäudegeometrie kann die Fläche für jede Etage vom Gebäude abweichen.

Über bot:hasSpace wird der Etage ground_floor die Zone test_room mit @id sites/TU-Darmstadt/buildings/TU_Darmstadt_Buildings/storeys/ground_floor/spaces/test_room hinzugefügt. So wie die Etagen weist jeder Raum unterschiedliche Dimensionen auf und benötigt die eigene Beschreibung durch geo:geometry. Zudem enthält der Raum Elemente, welche über hasElements strukturiert werden.

Anders als Raumelemente sind Fenster und Türen oftmals Teil von mehr als einem Raum, wenn diese eine Verbindung haben oder aneinander angrenzen. Deshalb werden im Konzept von BoT Fenster und Türen als anliegende oder angrenzende Elemente beschrieben. Im Fall von zwei Räumen mit Verbindung oder Sicht in den jeweils anderen Raum erhält jeder Raum die Wandelemente als adjacentElements.

Weitergehende Anwendungen

Bei der Entwicklung der BoT-Strukturen wurden auch Ansätze für weitergehende Anwendungen auf Basis der semantischen Beschreibungen erarbeitet. Einer dieser Ansätze war es, das WoT-Modell mit BoT so zu erweitern, dass eine Visualisierung in einer 3D-Engine ermöglicht wird. Dazu mussten erweiterte Beschreibungsmöglichkeiten für Elemente in Raum und Wänden eingeführt werden. Hierzu wurde eine ergänzende Datenstruktur eingeführt, welche weiterhin eine valide Beschreibung nach dem BoT-Standard darstellt.

Mit der erweiterten Dateistruktur kann die geographische Ausrichtung sowie die Dimensionen der Elemente beschrieben werden. Auf dieser Basis wurde ein Script in der 3D-Engine Unity erstellt, welches das beschriebene Gebäude mit Fenstern und Türen darstellt (Abbildung 7).



Abbildung 7 BoT-Beschreibung zu 3D-Rendering

Die erweiterte BoT-Beschreibung wurde im standardisierten und maschinenlesbaren JSON-Format (ISO/IEC 21778:2017) umgesetzt. Die Struktur folgt einem gängigen Vorgehen bei der Verwendung von Unity und ist in Abbildung 8 dargestellt. Das Beispiel zeigt Elemente in der Laborumgebung der FH Dortmund.

```

{
  @id : https://ikt-dortmund.de:8443/ict-gw/v1/sites/FH Do
  elements : [ 0 items ]
  0 : {
    @id : https://ikt-dortmund.de:8443/ict-gw/v1/sites/FH Do/Buildings/FH Do building/elements/door
    name : door
    position : [ 3 items ]
      0 : 2.97
      1 : 0
      2 : 0
    rotation : [ 3 items ]
      0 : 0
      1 : 0
      2 : 0
    scaling : [ 3 items ]
      0 : 1
      1 : 1
      2 : 1
    shape : {
      type : simple
      format : custom
      data : [ 3 items ]
        0 : 0.98
        1 : 0.2
        2 : 2.1
    }
  }
}

```

Abbildung 8: Erweiterte BoT-Struktur

Die übergeordnete @id referenziert die vorliegende BoT-Beschreibung in der WoT-IoT-DB und legt den Basispfad für die zu beschreibenden Elemente fest. Zu beschreibende Elemente werden als JsonObject im JSONArray elements organisiert.

Jedes JsonObject hat eine eindeutige @id, welche auf das Element in der BoT-Beschreibung zeigt. Dies ist notwendig, da eine eindeutige Zuordnung nur über die enthaltene URI möglich ist. Der name enthält den menschenlesbaren Namen des Objekts und dient der Übersicht, falls das Dokument manuell überprüft wird. Geometriedaten werden separiert in den vier Unterelementen position, rotation, scaling und shape gespeichert.

Für die BoT-Beschreibung wurde festgelegt, dass jegliche Betrachtung von Koordinaten im Raum von der Ecke links unten ausgeht. Das JsonObject position beschreibt diese Koordinate als 3D-Vektor.

Da für die Betrachtung und Planung von Licht neben dem künstlichen Anteil auch das natürliche Licht wichtig ist, wird für jedes Objekt die geographische Ausrichtung über das JsonObject rotation beschrieben. Mit dem Ziel diese anhand bereits existierender Ansätze zu modellieren wird der Kompass mit 0-360° als Referenz verwendet.

Das JsonObject scaling ist von der Objektstruktur in Unity inspiriert und gibt an, in welchem Verhältnis zur Originalgröße dieses im 3D-Modell in Unity dargestellt werden soll. Aktuell findet diese Information noch keine Verwendung, soll aber für spätere Erweiterungen genutzt werden.

Das JsonObject shape weist eine komplexere Struktur auf, da es die eigentlichen Daten zur Dimension vom Objekt enthält. Hier wurde der IFC-Standard (Industry Foundation Classes) aufgrund der offenen Struktur und der Möglichkeit zum standardisierten Datenaustausch gewählt. Shape hat drei Unterparameter type, format und data, welche eine automatisierte Verarbeitung ermöglichen. Die Parameter type und format beschreiben welche Daten wie aufbereitet in data

gespeichert sind. Abhängig davon wird in data das passende Format mit den eigentlichen Daten gespeichert.

2.3.3. Validierung der semantischen Beschreibung gegen IoL-Use Cases

Die prototypische Umsetzung der im Projekt erarbeiteten Use Cases stellte besondere Anforderungen an die semantische Annotation der Zustandsdaten (z.B. räumliche Verortung, Adaption der Messeinheiten, etc.). Die beteiligten Forschungseinrichtungen haben diese Anforderungen bereits zu Projektbeginn diskutiert und eine Vorgehensweise entworfen, um die Herausforderungen seitens der semantischen Annotation mit den Use Case-getriebenen Anforderungen abzugleichen. Es wurde ein iterativer Prozess etabliert, welcher eine initiale Anforderungsanalyse auf Use Cases-Ebene mit agilen Entwicklungsphasen und Feedbackrunden umfasste. Im Laufe dieses Prozesses wurden erforderliche Modifikationen an den Thing-Descriptions der Testanlagen vorgenommen und wo erforderlich weitere Ontologien einbezogen.

Aufgrund der pandemie-bedingten, sukzessiven Anbindung der Testanlagen wurde dieser Prozess entgegen der ursprünglichen Planung mehrfach durchlaufen, was sich jedoch aufgrund der heterogenen Datenstrukturen als unverzichtbar erwies.

2.3.4. Implementierung der Basis-Services

Die IoL-Cloud wird durch FE2 unter Verwendung einer Docker-Infrastruktur innerhalb eines zertifizierten Rechenzentrums gehostet. Docker ist eine Softwareumgebung, mit der die Entwicklung und Verwaltung von Anwendungen vereinfacht wird. Die Umgebung kann lokal wie auch in der Cloud betrieben werden und ist mit vielen aktuellen Standardtechnologien und Protokollen kompatibel. Dienste in Docker werden in sogenannten Containern abgelegt und können unkompliziert verwaltet und überwacht werden. Der Betrieb der IoL-Cloud erfordert verschiedene Basis-Services, welche im Folgenden beschrieben werden.

<input type="checkbox"/> Name	State  Filter 
<input type="checkbox"/> node-red-2	healthy
<input type="checkbox"/> kong	healthy
<input type="checkbox"/> node-red	healthy
<input type="checkbox"/> kong-postgres	healthy
<input type="checkbox"/> mosquito-	running
<input type="checkbox"/> wot-device-api	running
<input type="checkbox"/> mosquito-	running
<input type="checkbox"/> mosquito-	running
<input type="checkbox"/> mosquito	running
<input type="checkbox"/> timeseries-database	running
<input type="checkbox"/> thing-directory	running
<input type="checkbox"/> portainer	running
<input type="checkbox"/> eureka-service	running
<input type="checkbox"/> konga	running
<input type="checkbox"/> rdf4j-server	running
<input type="checkbox"/> docker-registry	running

Abbildung 9: Übersicht der Docker-Dienste.

Der Dienst **Node-RED** wird in den Containern *node-red* und *node-red-2* für die Verarbeitung und Speicherung der eingehenden Daten aller Anlagen verwendet. Node-RED bietet eine browserbasierte Benutzeroberfläche an, über welche mit intuitiver Bedienung, verschiedene Protokolle und Funktionen realisiert werden können. Node-RED benötigt nur wenig Ressourcen und ist somit auch auf Mikrocontrollern lauffähig. Insgesamt laufen zwei Instanzen des Dienstes Node-RED um bei hohem Datenaufkommen ein Loadbalancing zu erreichen.

Der Dienst **Mosquitto** in den Containern wird verwendet, wenn Anlagenbetreiber keinen eigenen MQTT-Broker für die Datenübertragung zur Verfügung stellen konnten.

Die **“wot-device-api”** wird für die Abfrage der Zustandsdaten über eine NDA-konforme Schnittstelle der Anlagenbetreiber und Forschungseinrichtungen FE1+2 verwendet.

RDF4J ist eine auf Java basierte Datenbank für semantisch aufbereitete Informationen wie RDF und Linked-Data mit Möglichkeiten diese wie in einer Datenbank zu filtern und zu verwalten. Die WoT- und BoT-Beschreibungen der Anlagen werden hier vorgehalten und können über die WoT-Device-API abgerufen und ausgewertet werden.

Der Docker-Container **“docker-registry”** stellt den serverseitigen gleichnamigen Dienst Docker Registry zur Verfügung. Mit diesem Dienst werden alle Docker-Images für die Basisdienste zentral und kontrolliert gespeichert, sowie von dort im System verteilt.

Im Docker-Container **“timeseries-database”** wird eine InfluxDB als Cloud-Datenbank betrieben. InfluxDB ist eine Zeitreihen-Datenbank für IoT-Entwickler und Cloud-Anwendungen und bietet eine leistungsstarke Engine optimiert für die Verarbeitung von Echtzeitdaten. Neben der Unterstützung der Programmiersprachen JavaScript, Python, Go, PHP und C# für clientseitige

Abfragen gibt es viele Möglichkeit den Funktionsumfang durch kostenfreie Plugins zu erweitern. Die Zeitreihen-Datenbank ist kompatibel zu allen aktuellen Technologien und wird im Projektrahmen direkt durch Node-RED mit Daten beschrieben. InfluxDB bietet eine integrierte Deadman-Check Funktion, um Anomalien bei der Datenaufzeichnung automatisiert zu detektieren.

Die Dienste **kong**, **konga** und **kong-postgres** realisieren die externe Erreichbarkeit für autorisierte Zugriffe und bilden das Benutzer- und Rollenmanagement ab.

2.3.5. Zielabgleich

Die in AP3 gemäß Antrag definierten Ziele

- Ontologien für die Anwendungsdomäne Licht identifiziert (FE 1+2),
- erweitertes WoT-Modell liegt vor (FE 2),
- semantische Beschreibung gegen Use Cases in AP4 validiert (FE 1+2),
- Basis-Services implementiert (FE 2)

wurden erfolgreich erreicht.

2.4. AP4 Use Cases

Im folgenden Kapitel werden die drei implementierten Use Cases zum Monitoring der Betriebszuverlässigkeit, die Vorhersage von Raumparametern, sowie die Bestimmung eines minimalen Sets an benötigten Sensoren vorgestellt. Um die Privacy der teilnehmenden Unternehmen zu wahren, werden die acht Standorte mit den Laufnummern 1 bis 8 kodiert und die Evaluation exemplarisch durchgeführt und präsentiert. In Kapitel 2.4.4 wird eine Toolbox vorgestellt, mit dieser die Unternehmen für ihre eigenen Standorte und Daten die gesamte Analyse durchführen können.

Für die weitere Verarbeitung der Daten wird die Struktur aus Abbildung 10 der Datenverarbeitung genutzt. Für die Evaluation und bandbreitenunbeschränkte Verarbeitung der Daten werden diese in der ersten Phase aus der Datenbank von der IoL-Cloud heruntergeladen und lokal gespeichert.

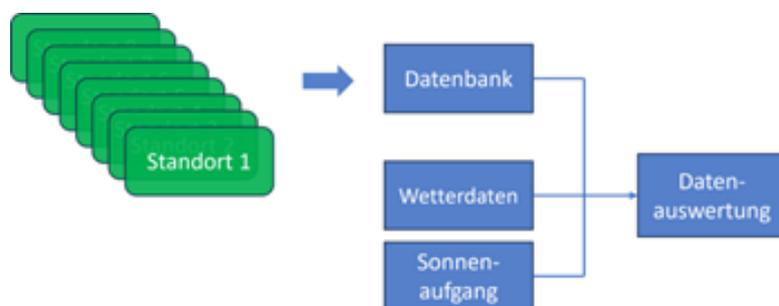


Abbildung 10: Übersicht der Datenstruktur mit acht Standorten, die die Daten in die Datenbank schreiben, die zusätzlich erhobenen Wetter- und Sonnenaufgangsdaten, die zur Datenauswertung genutzt werden.

Zur Begrenzung der Datenmenge können Standorte, Start- und Endzeit, sowie einzelne Sensoren gewählt werden. Hierfür werden Anfragen an die IoL-Cloud gestellt und die Daten im in Kapitel 2.3 (Einheitliche Semantik) vorgestellten Format empfangen. Für eine einfache Weiterverarbeitung werden die Datenwerte der Sensoren in einer Tabelle gespeichert, wobei jede Spalte einen Sensortyp darstellt und jede Zeile einen zeitlichen Eintrag. Die Metainformationen werden in einer weiteren Tabelle gespeichert. Mithilfe der übersichtlichen tabellarischen Darstellung werden fehlende und redundante Daten identifiziert. Redundante Dateneinträge werden in dieser Phase aus der Tabelle entfernt. Fehlende Daten werden linear interpoliert und markiert um diese in den nachfolgend vorgestellten Anwendungsfällen mit komplexeren Ansätzen zu ersetzen, falls nötig.

Für die Vorhersage von Raumparametern werden Wetterdaten von der Openweathermap.org API [20], Sonnenaufgangs- und Untergangsdaten von <https://sunrise-sunset.org> [21] bezogen.

Zur Visualisierung eines Standortes mit seinen Sensoren und Aktoren wurde eine Methode zur Generierung einer 2D-Darstellung aus der BoT-Beschreibung entwickelt. Ein Beispiel ist in Abbildung 11 zu sehen mit Positionen von Sensoren, Lichtquellen, Fenstern, Türen, sowie den Raummaßen.

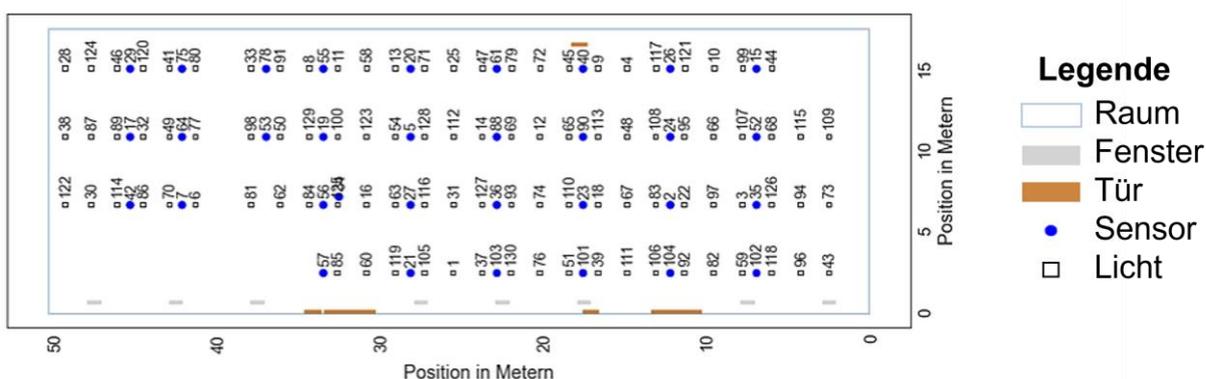


Abbildung 11: Automatisch generierte Darstellung eines Beispiel Standorts aus dem Projekt mit Fenstern, Türen, Lichtquellen, Sensoren und den Raumabmessungen.

Datenvorverarbeitung

Für eine erste Evaluation wurde eine deskriptive Statistik auf die Sensordaten angewandt, deren Schritte nachfolgend erläutert werden.

Zunächst wurden im Rahmen einer explorativen Datenanalyse grundsätzliche Merkmale der erfassten Daten untersucht, beispielsweise der Datentyp, Wertebereich, Trends und Saisonalitäten, Korrelationen und statistische Verteilung.

Tabelle 2 zeigt einen Auszug der bestimmten Merkmale zu verschiedenen Datenpunkten einer Testanlage.

Tabelle 2: Exemplarische Merkmale verschiedener Datenpunkte einer Testanlage.

Datenpunkt	Beleuchtungsstärke	Temperatur	Luftfeuchtigkeit	Leistung
<i>Datentyp</i>	kontinuierlich	kontinuierlich	kontinuierlich	kontinuierlich
<i>Fehlende Daten</i>	Ja	Ja	Ja	Ja
<i>Trend</i>	Nein	Nein	Nein	Nein
<i>Saisonalität</i>	Tages/Jahres	Tages	Tages/Jahres	Tages/Wochen
<i>Minimum</i>	0 lx	27,95 °C	8,43 %	0 W
<i>Maximum</i>	1144,58 lx	51,56 °C	30,95 %	56,29 W
<i>Mittelwert</i>	85,76 lx	36,84 °C	17,95 %	7,49 W

Basierend auf den gewonnenen Erkenntnissen erfolgte die Konzeption eines Vorverarbeitungsablaufs mit Möglichkeiten zum temporalen Resampling (für spätere Untersuchungen zu optimalen Erhebungsintervallen), dem Ersetzen (Interpolieren) fehlender Werte sowie dem Skalieren der Daten (Angleichen verschiedener Größenordnungen zur Vermeidung numerischer Probleme in der Weiterverarbeitung durch Vorhersagemodelle o.ä.).

Hinsichtlich Trends und Saisonalitäten lassen sich unterschiedliche Zeithorizonte feststellen. Abbildung 12 zeigt Temperatur, Luftfeuchtigkeit, Beleuchtungsstärke und Leistung einer Leuchte über zwei Beispiel-Zeiträume. Auffällig ist hier eine Lücke zwischen Februar und Juni 2022. Diese resultiert aus einem Fehler in der Datenübertragung und muss für die weitere Verarbeitung gesondert behandelt werden. Im Vergleich zwischen Innen- und Außentemperatur ist erkennbar, dass es sich um einen klimatisierten Raum handelt und die Innentemperatur so gut wie keine Jahres-Saisonalität aufweist. Eine Ausnahme stellen hier Zeitpunkte mit besonders hohen Außentemperaturen dar, die vermutlich dazu führen, dass ein Halten der gewünschten Temperatur nicht mehr möglich ist. Bei der Luftfeuchtigkeit und der Beleuchtungsstärke hingegen sind klare Abhängigkeiten von der Jahreszeit erkennbar. Die Leistung der Leuchte wiederum weist keine Jahres-Saisonalität auf, zeigt jedoch bei der Betrachtung einer Woche, dass die Beleuchtung nur an Werktagen verwendet wurde. Zudem ist bei allen Datenpunkten eine Tages-Saisonalität feststellbar. Die saisonalen Komponenten für Tages-, Wochentage und Jahreszeit als somit wichtige Einflussgrößen werden über eine Sinusfunktion, deren Wert sich aus dem Verhältnis zwischen dem Zeitpunkt der Messwertaufnahme und dem betrachteten Merkmal (Tag: 24 h, Woche: 7 d, Jahr: 365 d) gemäß der Funktion:

$$x = \sin(\text{Zeitpunkt} \cdot 2\pi \cdot \text{Periodenlänge})$$

ergibt, in die Datenverarbeitung integriert.

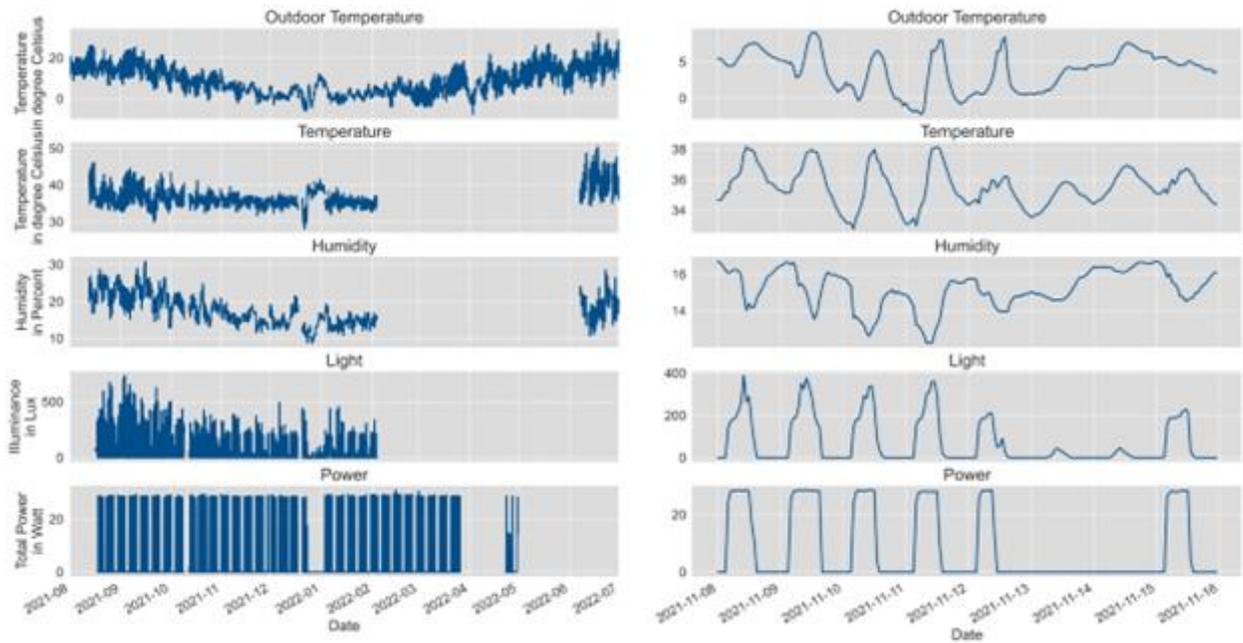


Abbildung 12: Trends und Saisonalitäten mit verschiedenen Zeitkonstanten.

Abbildung 13 zeigt beispielhaft die Korrelation zwischen den in Abbildung 12 betrachteten Merkmalen. Während bei der hohen Korrelation zwischen der Leistung einer Leuchte und der gemessenen Beleuchtungsstärke auch eine Kausalität angenommen werden kann, ist dies bei der negativen Korrelation zwischen Luftfeuchtigkeit im Außenbereich und Beleuchtungsstärke nicht der Fall, der Zusammenhang besteht hier indirekt über die Tageszeit.

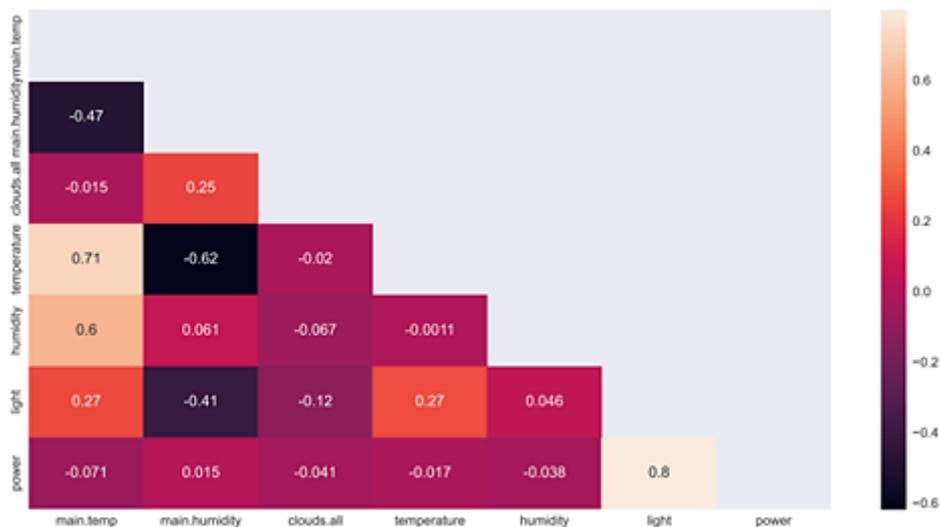


Abbildung 13: Pearson Korrelation von Außentemperatur, Außenluftfeuchtigkeit, Bewölkungsgrad, Temperatur, Luftfeuchtigkeit, Beleuchtungsstärke und Leistung an einem Leuchtenpunkt.

Im nächsten Schritt werden fehlende Werte ersetzt, um eine Verwendung von Vorhersagemodellen zu ermöglichen, da diese häufig lückenhaften Daten nicht verarbeiten können. Bei der Ersetzung fehlender Datenpunkte wird zwischen kurzfristigen und langfristigen Ausfällen unterschieden. Bei kurzfristigen Ausfällen fehlen nur einzelne Werte bis zu einem Zeitraum von wenigen Stunden. Diese können durch Wiederholung des letzten bekannten Wertes oder durch lineare Interpolation ersetzt werden. Kommt es zu einem längerfristigen Ausfall von

mehreren Stunden oder sogar Tagen, ist eine einfache Interpolation oder Wertwiederholung nicht mehr sinnvoll. Eine naheliegende Möglichkeit zum Ersetzen fehlender Werte ist die Verwendung des Wertes vom Vortag zur selben Uhrzeit, unter Berücksichtigung aller den Datenpunkt betreffenden Saisonalitäten. Befinden sich an einem Standort mehrere Sensoren des gleichen Typs, kann auch der gewichtete Mittelwert der umliegenden Sensoren zum Ersetzen der fehlenden Daten genutzt werden. Hierzu sind Informationen über die Positionen der Sensoren erforderlich, da die Werte von näherliegenden Sensoren stärker gewichtet werden sollten. Ein Beispiel für den Umgang mit fehlenden Datenpunkten ist in Abbildung 14 dargestellt.

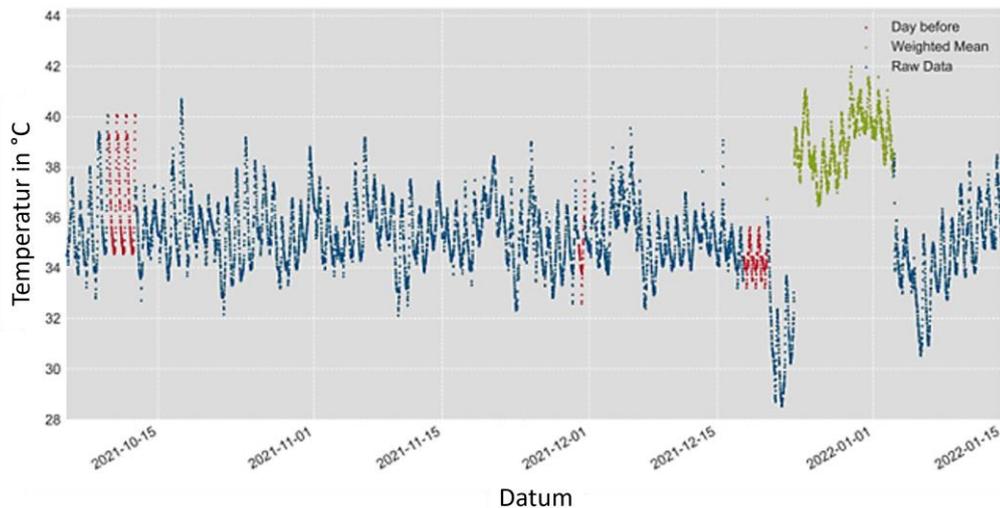


Abbildung 14: Ersetzen fehlender Werte eines Temperatursensors durch Werte vom Vortag sowie einem gewichteten Mittelwert der umliegenden Sensoren.

Im letzten Schritt der Vorverarbeitung erfolgt die Skalierung der Daten. Diese ist insbesondere bei der multivariaten Analyse mit unterschiedlichen Größenordnungen und unterschiedlichen Wertebereichen erforderlich. Häufig basieren Vorhersage-Modelle auf der Betrachtung von Abständen zwischen den Werten. Besitzt ein Feature jedoch einen größeren Wertebereich als die übrigen, sind auch seine Abstände größer, was möglicherweise zu einer stärkeren Gewichtung führt. Bei der Skalierung von Daten wird zwischen Normalisierung und Standardisierung unterschieden. Normalisierung transformiert alle Daten in einen bestimmten Wertebereich, während Standardisierung dazu führt, dass alle Daten einen Mittelwert von Null und eine Standardabweichung von 1 besitzen. Beide Verfahren werden als Vorverarbeitungsschritt in diesem Projekt angewendet.

Die vorverarbeiteten Daten werden im Folgenden für die Erprobung verschiedener Use Cases aus dem Projekt verwendet.

2.4.1. Monitoring zur Betriebszuverlässigkeit

Während der Erfassungszeit waren keine Schadensereignisse in den Beleuchtungssystemen zu beobachten, eine Absehbarkeit solcher Ereignisse aus den erfassten Leuchten-Daten (beispielsweise Temperatur, Leistungsaufnahme) und externer Sensorik (beispielsweise Helligkeitssensoren) konnte daher nicht untersucht werden.

Infolgedessen werden zwei Themen für die Betriebszuverlässigkeit betrachtet: In der Anlage auftretende Zustandsanomalien sowie die Identifikation des graduellen Lichtstromrückgangs basierend auf den erhobenen Leuchtendaten. Für eine exakte Anomalie-Detektion müssen fehlende Werte präzise interpoliert werden, wozu datengetriebene Methoden zur Zeitreihenvorhersage verwendet wurden.

Zeitreihenvorhersage

Da es sich bei den in dieser Arbeit vorliegenden Daten um Zeitreihen handelt, die sich dem Bereich des Internet of Things (IoT) zuordnen lassen, wurde die Suche nach dem Stand der Wissenschaft auf diesen Bereich beschränkt. Beispielsweise lassen sich durch Recurrent Neural Networks (RNN) bzw. Long Short-Term Memories (LSTM) sequentielle Daten, wie Bewegungsabläufe oder Stromverbräuche, über einen Machine Learning-Ansatz vorhersagen [22].

Einen ausführlichen Vergleich verschiedener Modelle sowie wichtige Voraussetzungen zur Analyse von IoT Daten liefern Saleem und Chishti [23]. Es werden etwa Lineare Regression zur Vorhersage von Energieverbräuchen, Entscheidungsbäume zur Bewegungserkennung oder eine Support Vector Machine (SVM) zur Überwachung von akustischen Signalen genannt. Als Deep Learning Modelle werden LSTM-Modelle für die Vorhersage der Luftqualität und LSTM bzw. CNN-Modelle zur Aktivitätserkennung vorgestellt. Im Bereich der Zeitreihenvorhersage sind es unterschiedliche Modelle wie Auto-Regressive Integrated Moving Average (ARIMA), Hidden Markov Model (HMM) und RNN für verschiedenste Szenarien, wie z. B. der Wetter- (ARIMA, RNN) und Energieverbrauchsvorhersage (ARIMA), der Anomalie oder Bewegungserkennung (HMM), oder der Aktivitätserkennung durch eine Kombination von Sensoren (RNN). [23]

Ein Vergleich verschiedener Machine Learning Methoden für die Vorhersage der Sonneneinstrahlung wurde durch Voyant et al. [24] durchgeführt. Es wurde festgestellt, dass die verschiedenen Modelle je nach gegebener Voraussetzung unterschiedlich gut geeignet waren, wobei Support Vector Machines (SVM), Entscheidungsbäume und Random Forests vielversprechende Resultate lieferten - Deep Learning Modelle wurden nicht berücksichtigt.

Für die schnell und sprunghaft veränderliche Leistung und den Stromverbrauch zeigt die Literatur zur Vorhersage folgende Modelle auf: Leistung wurde erfolgreich mittels ARIMA [25] und XGBoost [26] für stündliche und wöchentliche Vorhersagen modelliert. Stromverbrauch konnte mit einem 15 Minuten Zeithorizont ebenfalls mit XGBoost und LSTMs erfolgreich vorhergesagt werden [27]. Die jeweiligen Untersuchungen zeigten einen geringen Fehler mit den jeweils verwendeten Modellen.

Anhand dieser Auswahl an Vorarbeiten ist gut zu erkennen, dass viele verschiedene Modelle existieren, die alle ihre eigenen Vor- und Nachteile besitzen und für unterschiedliche Situationen unterschiedlich gut geeignet sein könnten. Daher wurde ein klassisches statistisches Modell der Zeitreihenvorhersage wie das ARIMA Modell, leicht verständliche und erklärbare Modelle wie Entscheidungsbäume (XGBoost) und Markov Ketten (HMM) sowie ein Deep Learning Modell (LSTM) zu Vergleichszwecken in diesem Projekt parallel untersucht.

Der wesentliche Parameter bei der Verwendung des ARIMA Modells ist seine Ordnung, welche sich mit Hilfe des „pmdarima“ Pakets anhand der vorliegenden Daten bestimmen lässt [28], [29]. Hierbei können dann die jeweiligen maximalen Ordnungen sowie Startordnungen definiert

werden. Das ARIMA-Modell muss iterativ ausgewertet werden. Dies bedeutet, dass für den ersten Zeitpunkt in den Testdaten der gewünschte Zeitraum (in diesem Fall vier Stunden) vorhergesagt wird. Für den nächsten Zeitpunkt wird der wahre Wert des vorherigen Zeitpunkts zu den Trainingsdaten hinzugefügt, und das Modell wird neu trainiert. Dieser Vorgang wird so lange wiederholt, bis alle Zeitpunkte in den Testdaten vorhergesagt wurden.

Aufgrund des iterativen Ansatzes gibt es für jeden Zeitpunkt mehrere Vorhersagen. Daher wird nur der erste Vorhersage-Wert zur Bestimmung des Scores verwendet. Außerdem wird für den Modellvergleich und die Parameterauswahl hier nur die univariate Analyse betrachtet, da die multivariate Modellierung keine signifikante Verbesserung der Vorhersagen ergab. Zusätzlich zu den Vorverarbeitungsparametern können die Anzahl der Bäume und die verwendeten Merkmale für das XGBoost-Modell variiert werden. Ein vorzeitiger Abbruch, der das Training beendet, wenn das Hinzufügen weiterer Bäume die Vorhersagen nicht verbessert, ist ebenfalls möglich. Die Anzahl der Runden ohne Verbesserung kann festgelegt werden. Die Auswertung erfolgt anhand der Validierungsdaten. Da die Variation der Anzahl der Bäume auch bei vorzeitigem Abbruch keine signifikante Veränderung der Vorhersagen zur Folge hatte, wurde der Wert auf 50 Bäume festgelegt. Daher beschränkt sich die Abstimmung der Hyperparameter für das XGBoost-Modell auf die Vorverarbeitungsparameter und die Merkmalsauswahl. Zusätzlich zu den Merkmalen, die dem Modell übergeben werden, werden der aktuelle Monat, der Wochentag und die Stunde als numerische Werte einbezogen. Um einen Vorhersagezeitraum von vier Stunden zu ermöglichen, wird das Training mit um vier Stunden nach hinten verschobenen Werten durchgeführt. Dadurch kann jeweils nur ein Wert vorhergesagt werden, und frühere Vorhersagen werden verwendet, um den gesamten Zeitraum zu füllen. Die Auswahl der Merkmale basiert auf einer vom XGBoost-Algorithmus ermittelten Punktzahl, welcher die Relevanz des Features widerspiegelt.

Das LSTM-Modell hat zusätzliche Parameter wie die Anzahl der Trainingsepochen und die Angabe der vergangenen Stunden als zusätzliche Merkmale. Für jeden Zeitschritt dient ein Array der Größe (Anzahl der vergangenen Zeitschritte, Anzahl der Merkmale) als Input für das Modell. Dieses Array wird dann in die LSTM-Schicht mit 64 Einheiten (units) eingespeist. Der Ausgang der LSTM-Schicht ist mit einem Dense-Layer aus 32 Neuronen verbunden, die ihrerseits mit einem Dense-Layer mit einem einzelnen Neuron als Ausgang verbunden ist. Der Wert dieses einzelnen Neurons entspricht dem vorhergesagten Wert. Der Aufbau der LSTM-Modellarchitektur ist in Abbildung 15 dargestellt. Die Verwendung von tieferen Netzen mit mehreren LSTM-Schichten und zusätzlichen Dense- oder Dropout-Schichten (Layern) führte nicht zu einer signifikanten Verbesserung. Es wurden 50 Trainingsepochen verwendet, und das Training wurde vorzeitig abgebrochen, wenn in 10 aufeinanderfolgenden Epochen keine Verbesserung bei den Validierungsdaten zu verzeichnen war.

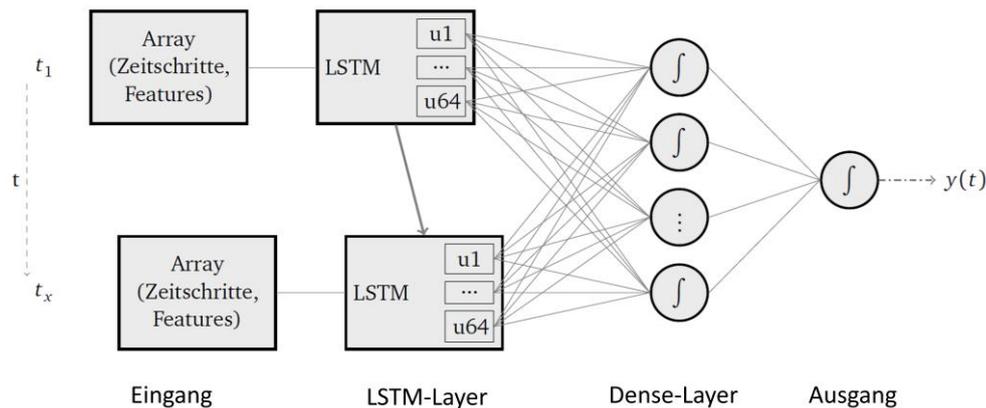


Abbildung 15: Visualisierung der verwendeten LSTM-Modellarchitektur. Jeder LSTM-Block besitzt 64 Units, die mit einem Dense-Layer bestehen aus 32 Neuronen verbunden sind.

Bei dem HMM-Modell können neben den Vorverarbeitungsparametern und der Auswahl der Merkmale auch die Anzahl der zugrundeliegenden Zustände, die Anzahl der vergangenen Stunden und die Verwendung einer Zufallsvariable bei der Vorhersage als Hyperparameter definiert werden. Im Gegensatz zu den anderen Modellen wird jedoch nicht die vorhergesagte Variable selbst als Merkmal verwendet, sondern ihre Veränderung. Zunächst wird der Wert zum aktuellen Zeitschritt erfasst. Zusätzlich werden der Wert des vorangegangenen Zeitschrittes sowie die Maximal- und Minimalwerte innerhalb des definierten Zeitfensters (Anzahl der vergangenen Stunden) gespeichert. So kann die relative Veränderung zum vorherigen Wert sowie zum Maximum und Minimum innerhalb des Zeitfensters als Merkmal verwendet werden. Zusätzlich kann ein kurzfristiger Trend berücksichtigt werden, der beschreibt, ob die beobachtete Variable innerhalb des Zeitfensters überwiegend zunimmt, abnimmt oder konstant bleibt. Dieser Trend wird als überwiegend abnehmend (-1) bezeichnet, wenn mehr als 65 % der Differenzen innerhalb des Zeitfensters negativ sind. Er wird als überwiegend ansteigend (1) beschrieben, wenn weniger als 35 % der Unterschiede negativ sind, und als konstant (0) für einen Anteil negativer Unterschiede zwischen 35 % und 65 %. Da das HMM-Modell jeweils nur einen Zeitschritt vorhersagen kann, werden längere Vorhersagen iterativ ermittelt. Dies führt zu mehreren Vorhersagen für jeden Zeitpunkt in längeren Zeiträumen. Für den Modellvergleich wird der erste vorhergesagte Wert verwendet.

Startpunkt der Arbeit mit jeglichen Prädiktionsmodellen ist die Bestimmung der Hyperparameter. Abbildung 16 zeigt den Hyperparameter-Suchraum für die genannten Modell-Varianten.

Vorverarbeitung	Scaler	ohne, minmax, standard
	Füllen mit letztem Wert	1H, 4H
	Füllmethode	day before, mean, interpolate (6H), (weighted mean)
ARIMA	Featureauswahl	univariat
HMM	Featureauswahl	Verwendung aller Features
	Zustände	3, 6
	zurückliegende Stunden	1, 2, 4
	kurzfristiger Trend	Ja, Nein
	Zufallsanteil	Ja, Nein
XGBoost	Featureauswahl	automatische Featureauswahl
LSTM	Featureauswahl	automatische Featureauswahl
	zurückliegende Stunden	36, 48

Abbildung 16: Hyperparameter-Suchraum für die untersuchten Vorhersage-Modelle: ARIMA, HMM, XGBoost und LSTM.

Die optimalen Merkmale variieren je nach Datentyp und dem verwendeten Modell. Insgesamt erwiesen sich jedoch in Bezug auf die Vorverarbeitung die Verwendung des Standard-Scalers, das Auffüllen fehlender Werte mit dem letzten beobachteten Wert für eine Stunde und die Verwendung des Mittelwerts für "nan"-Werte (not a number (nan) zur Identifikation fehlender Werte) als am besten geeignet. In Bezug auf die Hyperparameter der Modelle wurde festgestellt, dass für das HMM-Modell die Verwendung von sechs Zuständen, die Berücksichtigung nur der letzten Stunde, die Einbeziehung des kurzfristigen Trends und das Weglassen der Zufallskomponente bei der Vorhersage die besten Modellierungsergebnisse erbrachten. Bei der Vorhersage mit dem LSTM-Modell erwies sich die Berücksichtigung der 36 vorangegangenen Stunden als am effektivsten. Werden die jeweils besten Parameter sowie Features verwendet, kann ein Vergleich zwischen den Modellen und gegenüber Baseline-Modellen (letzten Wert fortsetzen oder Wert zum gleichen Zeitpunkt des Vortages) erfolgen.

Im direkten Vergleich wird der RMSE (root mean squared error) zum Bestimmen der Hyperparameter und Auswählen des geeignetsten Modells je Feature gewählt. Ein Vorhersagezeitraum von vier Stunden wurde betrachtet, da eine Veränderung der Features innerhalb dieser Zeit zu erwarten und damit eine Bewertung der Vorhersage möglich ist. Eine längere Vorhersage von mehreren Tagen hingegen wurde als nicht notwendig angesehen, da auf Erkenntnisse aus der Vorhersage schnell genug reagiert werden könnte.

Bei der Betrachtung der Metriken in Abbildung 17 wird deutlich, dass der Einsatz von Prognosemodellen sinnvoll ist und zu besseren Ergebnissen führt als bei den Baseline-Modellen. Als zugrunde gelegte Basismodelle der Vorhersage wurde einerseits der letzte Wert (Baseline: persistence) oder der Wert der gleichen Uhrzeit des Tages zuvor (Baseline: day before) genutzt. Der Vergleich der Modelle zeigt, dass XGBoost und HMM insgesamt die besten Ergebnisse liefern, mit den niedrigsten RMSE-Werten. Der MAPE (mean absolute percentage error) Score konnte nicht für alle Datenpunkte angegeben werden, da bei der Berechnung eine Division durch den Messwert erfolgt. Bei tatsächlichen Werten nahe Null ergibt sich jedoch ein sehr großer Wert, der keine Vergleichbarkeit zulässt und daher nicht dargestellt wird. Ein Vergleich über die verschiedenen Datenpunkte zeigt, dass alle Modelle bei sich langsam ändernden Größen, wie

bei der Temperatur, bessere Ergebnisse liefern als bei Verläufen mit größeren Sprüngen (Leistung, Beleuchtungsstärke).

Metrik	Datenpunkt	baseline persistence	baseline day before	ARIMA	XGBoost	LSTM	HMM
RMSE	Leistung in W	9,41	8,64	7,69	2,64	3,80	3,94
	Beleuchtungsstärke in lx	30,25	18,49	19,77	7,33	13,36	10,80
	Temperatur in °C	6,40	5,82	5,84	1,98	5,79	1,83
MAPE	Leistung	-	-	-	-	-	-
	Beleuchtungsstärke	1,49	0,49	2,26	0,21	0,60	0,22
	Temperatur	0,15	0,15	0,17	0,06	0,15	0,04

Abbildung 17: RMSE und MAPE-Score für die Vorhersage von Leistung, Beleuchtungsstärke und Temperatur an einem Beispiel-Standort für einen Vorhersagezeitraum von vier Stunden.

In Abbildung 18 sind die Vorhersageergebnisse nach dem Hyperparametertuning der Modelle für einen Beleuchtungsstärkesensor dargestellt. Der sprunghafte Anstieg der Beleuchtungsstärke weist den größten Fehler auf.

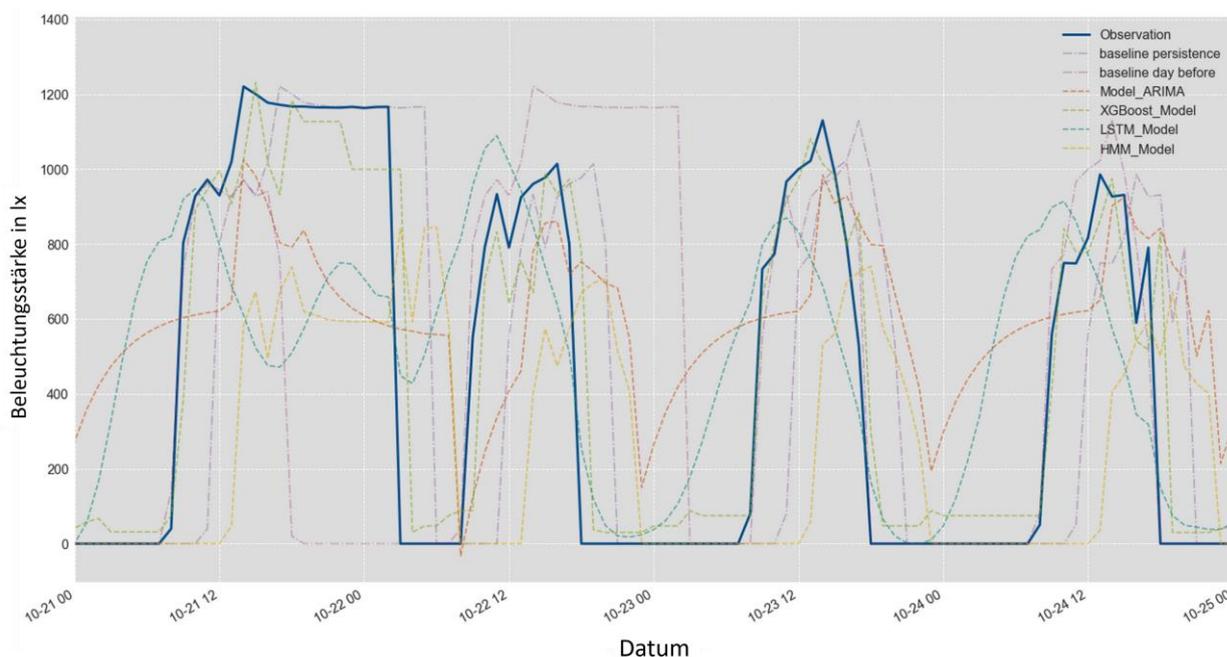


Abbildung 18: Modellvergleich für die Vorhersage der Beleuchtungsstärke mit einem Vorhersagezeitraum von vier Stunden.

Anomaliedetektion

Es wurden keine Informationen zu Fehlerzuständen seitens der Anlagenbetreiber erfasst, daher wurde zunächst simulativ eine Anomaliedetektion erarbeitet. Für die Simulation wurden die Sensordaten mit 25 künstlichen Fehlern beaufschlagt, die in der Realität im gleichen Maße auftreten könnten. Grundsätzlich können Anomalien in drei Typen eingruppiert werden: Punktuell, Kontextuell und Kollektiv.

Punktuelle Anomalien beschreiben einzelnen Punkte die als Ausreißer gedeutet werden können, da diese unterschiedliche Werte als die Gesamtheit der Datenaufweisen [30]. In Abbildung 19 ist dies beispielhaft für einen Temperatursensor und einem ausreißenden Wert auf 40 °C dargestellt.

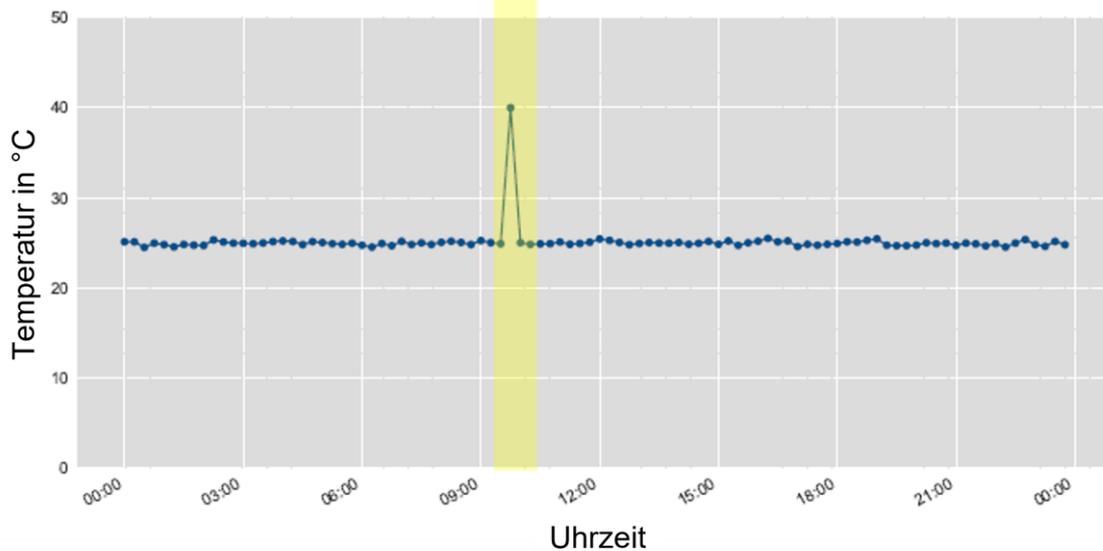


Abbildung 19: Punktuelle Anomalie eines Temperatursensors.

Kontextuelle Anomalien beschreiben einzelne Datenpunkte, die sich im Kontext anormal verhalten [30]. Ein Beispiel ist in Abbildung 20 für den Helligkeitsverlauf eines Sensors dargestellt. Dieser Sensor weist einen zyklischen Verlauf über den Tages- und Nachtrhythmus auf. Ein Wert ist um ca. 13 Uhr 0 lx und ist anormal, da dies nicht dem Tagesverlauf folgt, jedoch normal wäre, wenn dies in der Nacht auftreten würde, somit stellt dieser Wert bezogen auf den Tageskontext eine Anomalie dar.

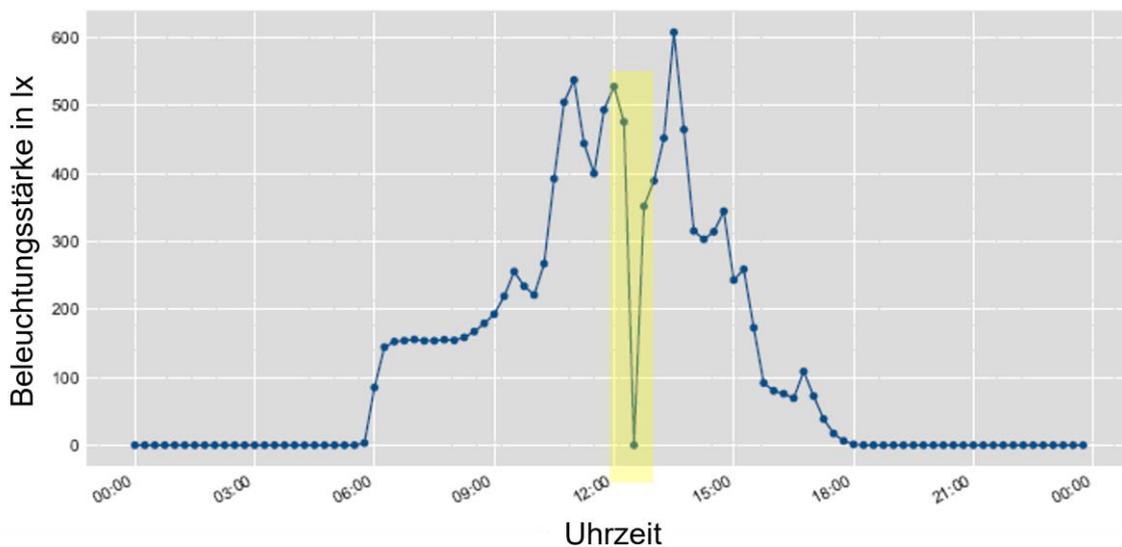


Abbildung 20: Kontextuelle Anomalie eines Helligkeitssensors.

Den dritten Typ von Anomalien stellen kollektive Ausreißer dar [30]. Abbildung 21 zeigt den Leistungsverlauf einer Leuchte über fünf Wochen. In der ersten Woche wird die Leuchte sechs Tage verwendet, wohingegen in den darauffolgenden Wochen die Leuchte nur fünf Tage verwendet wird. Daraus folgt, dass der sechste Tag in der ersten Woche eine Anomalie darstellt.

Anomalien müssen somit nicht immer negativ konnotiert sein, sondern stellen zunächst nur eine Abweichung vom erwarteten Verhalten dar. Erst durch Domänenwissen und ggf. Hintergrundinformationen können festgestellte Anomalien bewertet werden.

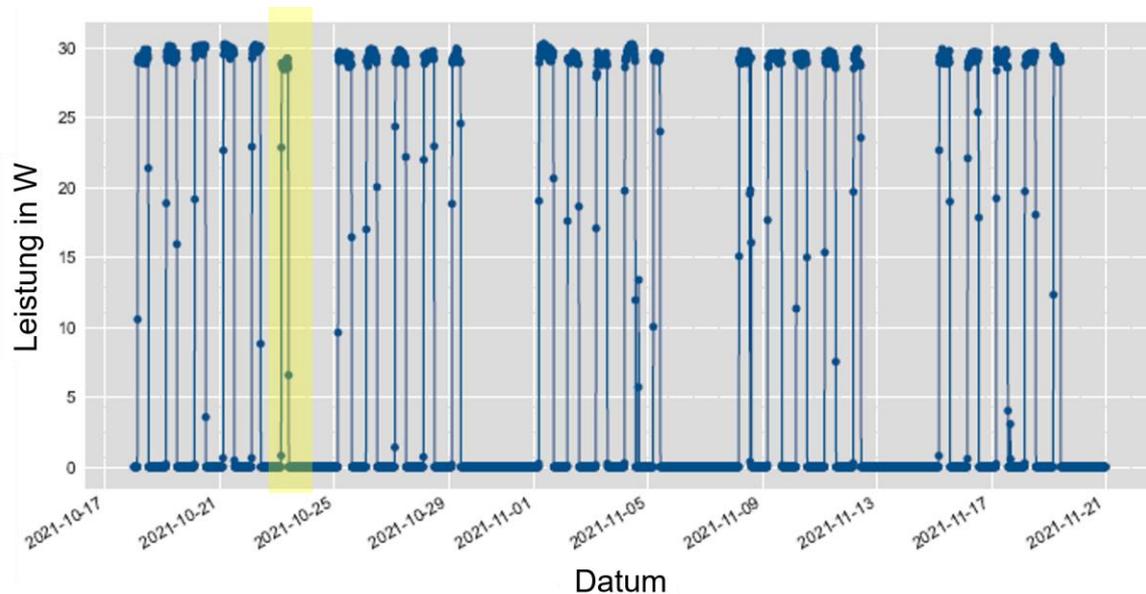


Abbildung 21: Kollektive Anomalie der Leistung einer Leuchte, da diese leuchte in der ersten Woche sechs Tage eingeschaltet war und jede weitere Woche fünf Tage.

Zur Überprüfung der Detektionsmethodik für diese drei Anomalie-Typen wurden zwei Varianten von Ausreißern eingefügt: Einzelne um 15 % erhöhte Werte, welche zufällig aus den Daten ausgewählt wurden. Dies simuliert kurzzeitige Fehlverhalten in der Anlage und entsprechen punktuellen und kontextuellen Anomalien. Außerdem wurden über einen Zeitraum von sechs Tagen alle Daten um 15 % erhöht, um Sensor- und Anlagen-Fehlverhalten zu simulieren und die Identifikation von kollektiven Anomalien zu untersuchen. Um einen „Sprung“ zu vermeiden, erfolgt die Manipulation der Daten über ein Tukey Fenster, welches für einen langsamen Anstieg bzw. Abfall der erhöhten Daten sorgt. Der Vorteil der Simulation der Fehler zur Bewertung der Erkennungsmethoden ist dabei, dass der Start- und Endzeitpunkt der Anomalien exakt bekannt ist. Insgesamt wurden drei Arten von Anomalie Erkennungsmethoden untersucht:

1. Gleitender Mittelwert (simple)
2. ARIMA, XGBoost, HMM und LSTM-Modelle
3. Dichtebasiertes Clustering: HDBSCAN

Um eine schnelle Verarbeitung der Daten zu gewährleisten, wurden ein Resampling aller Zeitseriendaten auf Intervalle von einer Stunde durchgeführt (Mittelwert der vorliegenden Daten während der jeweiligen Stunde). Als Referenz zur Anomaliedetektion wurde der gleitende Mittelwert, sowie die gleitende Standardabweichung genutzt. Werte die über der Standardabweichung liegen wurden als Anomalie gekennzeichnet. Bei den datengetriebenen Ansätzen ARIMA, XGBoost, HMM und LSTM wird der RMSE als tolerierbare Distanz genutzt. Weiterhin wurden 80 % der Daten für das Training der Modelle herangezogen und 20 % für die Evaluation. Als dritte Art der Anomalie Erkennung wird das dichte-basierte Clustering-Verfahren HDBSCAN genutzt, um Werte, die sich stärker von anderen Werten unterscheiden, zu markieren.

In Abbildung 22 sind beispielhaft die Daten eines Temperatursensors, sowie die jeweiligen Vorhersagen des gleitenden Mittelwert Modells (simple), der vier Modelle XGBoost, ARIMA, LSTM und HMM, sowie das HDBSCAN Clustering dargestellt. Es wird zwischen Cluster, Global und Local Anomalien unterschieden. Hierbei beschreibt eine Lokalanomalie Ausreißer, die nur an einem Sensor auftreten, Globalanomalien betreffen Werte, die bei allen Sensoren zum gleichen Zeitpunkt erhöht auftreten. Clusteranomalien sind jene, die bei mehreren umliegenden Sensoren detektiert werden können, jedoch nicht bei allen Sensoren des gleichen Typs in einem Standort.

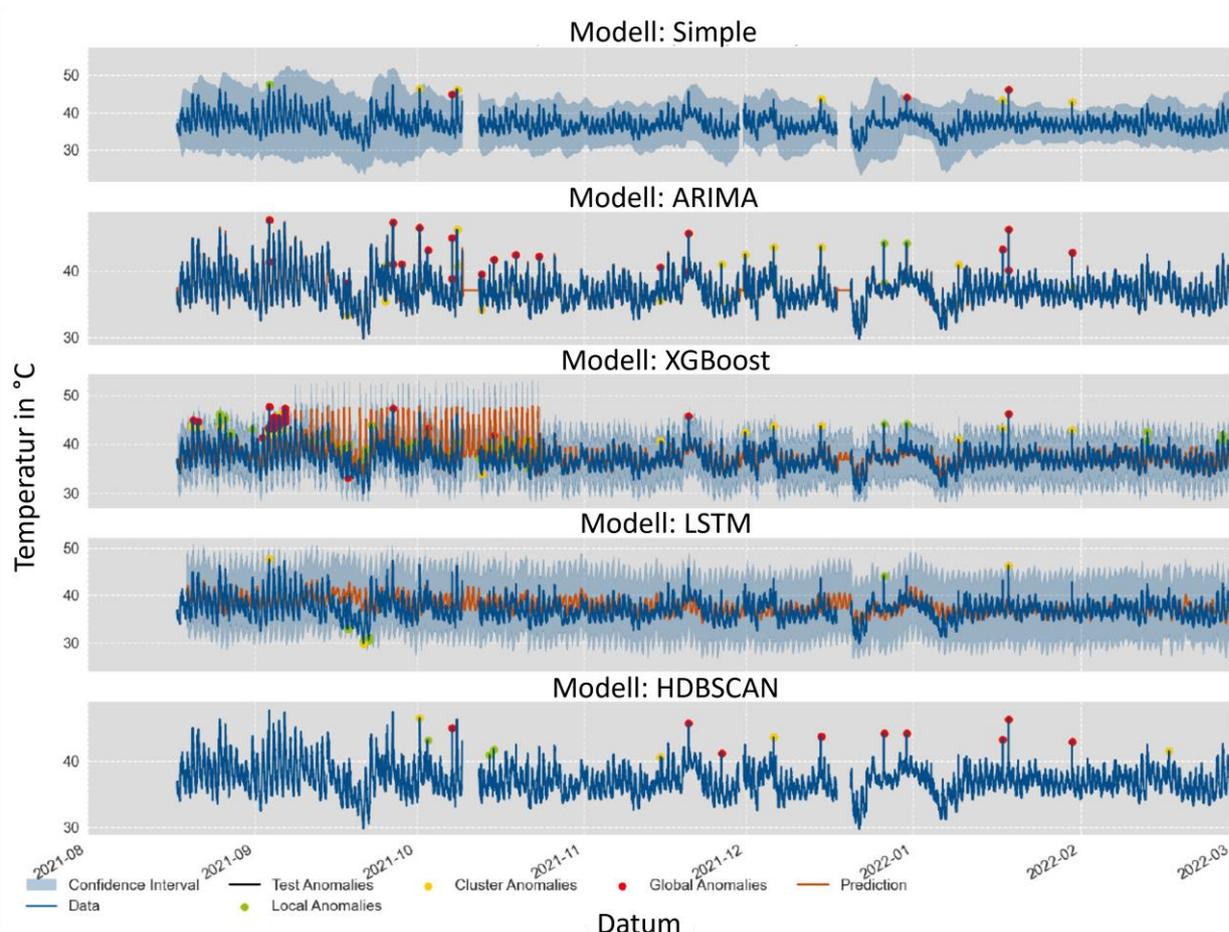


Abbildung 22: Beispiel der Anomaliedetektion anhand eines Temperatursensors. Als Anomalien dienen 25 zufällig ausgewählte Zeitbereiche, deren gemessene Werte künstlich um 15 % erhöht wurden. Zur Optimierung wurde die Precision-Recall Kurve verwendet.

Die Evaluierung der gefundenen Anomalien wird auf Basis der Confusion-Matrix, Receiver Operating Characteristic (ROC) und Precision-Recall Kurven durchgeführt. Hierfür werden die erkannten mit den tatsächlich, synthetisch eingebrachten Anomalien verglichen. Die Confusion-Matrix gibt die Anzahl korrekt als Anomalie gekennzeichnete (true positive - TP), falsch als Anomalie gekennzeichnete (false positive - FP), korrekt als keine Anomalie gekennzeichnete (true negative - TN) und falsch als keine Anomalie gekennzeichnete (false negative - FN) Datenpunkte an. Der schematische Aufbau einer Confusion-Matrix ist in Tabelle 3 gegeben.

Tabelle 3: Aufbau einer Confusion-Matrix mit TN (true negative), FP (false positive), FN (false negative) und TP (true positive) Anzahl an Werten. Das Label beschreibt die korrekte Zuordnung zur den zwei Klassen „Anomalie“ und „keine Anomalie“.

	Prediction	
Label	TN	FP
	FN	TP

Bei den gewählten Algorithmen Simple, XGBoost, LSTM, HMM, ARIMA und HDBSCAN kann der Grenzwert, ab dem ein Datenpunkt einer bestimmten Klasse zugeordnet wird, frei gewählt werden. Daher existieren verschiedene Möglichkeiten zur Bewertung und der Bestimmung des optimalen Grenzwerts. Eine gängige Möglichkeit bietet die Receiver Operating Characteristic (ROC). Sie gibt den Kompromiss zwischen true positive TP rate und false positive FP rate an. Je nachdem, welches Verhältnis zwischen TP und FN als akzeptabel gewählt wird, ergibt sich ein unterschiedlicher Punkt auf der Kurve, welcher diesen Grenzwert einhält. Eine vollständig zufällige Zuordnung zwischen TP und FN würde eine Diagonale als ROC erzeugen. Bei Daten, die sehr ungleich verteilt sind zwischen den zwei Klassen „Anomalie“ und „keine „Anomalie“, ist jedoch die ROC nicht optimal als Entscheidungskriterium. Dies ist der Fall für die vorliegenden Daten, da grundsätzlich davon ausgegangen werden kann, dass keine Anomalie vorliegt. Eine geeignete Alternative zur ROC stellt hier die Precision-Recall Kurve dar. Sie zeigt den Kompromiss zwischen Precision ($TP/(TP+FP)$) und Recall ($TPR = TP/(TP+FN)$). Die Bestimmung des optimalen Grenzwertes zur Entscheidung über eine Anomalie erfolgt dazu wie bei der ROC entweder über eine subjektive Entscheidung (akzeptables Verhältnis zwischen TP und FN) oder über die Maximierung des F1 Scores = $2TP/(2TP + FP + FN)$. Eine vollständig zufällige Zuordnung zwischen TP und FN würde eine horizontale Linie mit dem Verhältnis von positiven zu Gesamtfällen als Wert aufweisen. In Abbildung 23 sind die Confusion-Matrix, ROC und die Precision-Recall Kurve für die Anomalien des Temperatursensors aus Abbildung 20 dargestellt. Zur Optimierung wurde die Precision-Recall Kurve verwendet. Die über den F1 Score gefundenen Grenzwerte sind als rote Punkte gekennzeichnet. Anhand dieses Grenzwertes werden die Daten als Anomalie bzw. keine Anomalie klassifiziert und in der jeweiligen Confusion-Matrix links dargestellt.

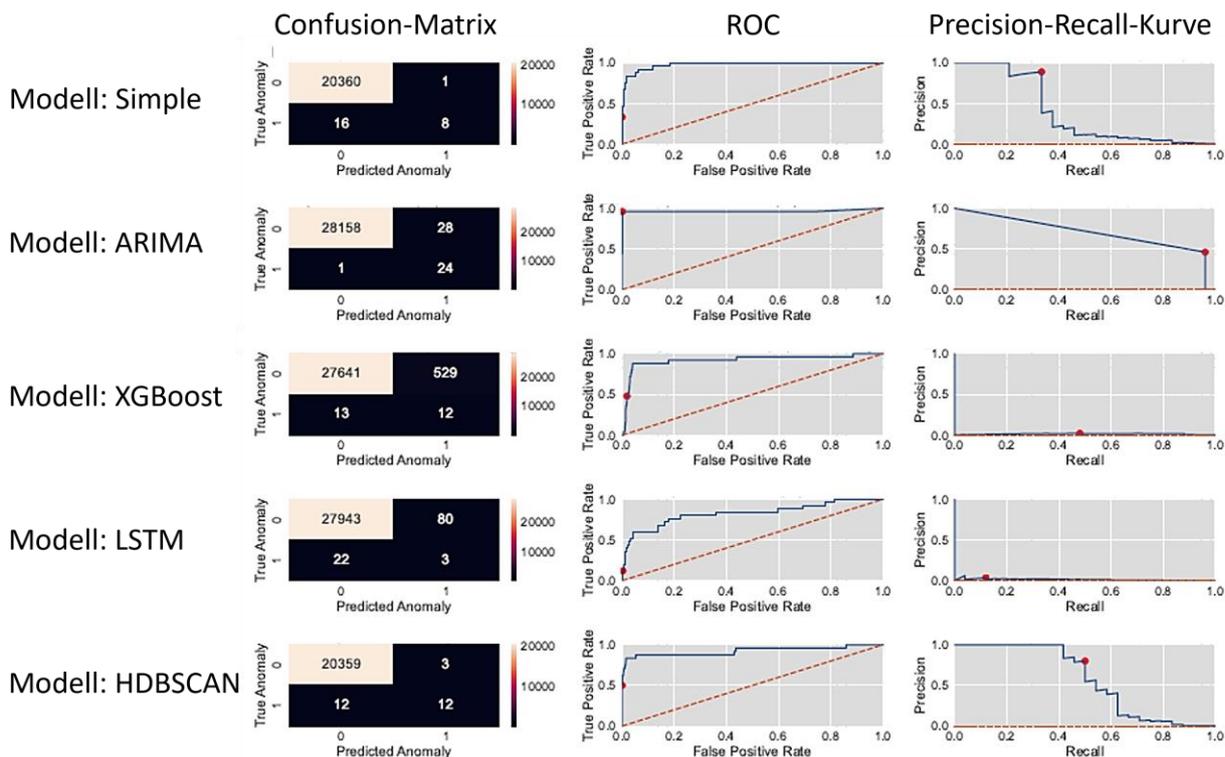


Abbildung 23: Confusion-Matrix, ROC und Precision-Recall-Kurven für die Modelle: Simple, ARIMA, XGBoost, LSTM und HDBSCAN zur Evaluierung der vorhergesagten Anomalien aus Abbildung 20 für einen Temperatursensor.

Im Anhang befinden sich die Abbildungen (Abbildung 34 - Abbildung 37) für die sechs Tage erhöhten Anomalien. Bei der Gegenüberstellung der Algorithmen weist der Simple Ansatz (Nutzung des Mittelwerts und der Standardabweichung als Toleranz), sowie das ARIMA Modell die besten Ergebnisse für punktuelle und kontextuelle Anomalien auf. Für kollektive Anomalien mit und ohne sanfte Anstiege/Rückgänge zeigt sich, dass alle Methoden deutlich schlechtere Resultate liefern, wobei der XGBoost und LSTM-Algorithmus noch am besten geeignet sind. Dies kann dadurch erklärt werden, dass die drei übrigen Methoden dem Verlauf der Daten sehr stark folgen und daher eine schleichende Änderung des Systems nicht registrieren, während XGBoost und LSTM deutlich universeller, allerdings weniger genau, Daten vorhersagen. Durch diese Generalisierung können diese Modelle längerfristige Anomalien besser detektieren. Bezüglich der insgesamt schlechteren Detektion müssen die erkannten Anomalien (siehe Abbildung 34 bis Abbildung 37 im Anhang) genauer betrachtet werden. Hierbei fällt auf, dass insbesondere die XGBoost und LSTM-Modelle die erzeugten Anomalien grundsätzlich korrekt erkannt haben, jedoch nicht jeder Datenpunkt innerhalb des manipulierten Zeitraums markiert wurde. Durch die Verwendung des Tukey Fensters und dem damit verbundenen Anstieg bzw. Abfall der Werte, werden gerade am Anfang oder Ende eines Anomalie-Fensters die Veränderungen durch die Detektionsalgorithmen noch nicht erkannt. Dadurch ergibt sich ein rechnerisch höherer Fehler-Score, obwohl der Hauptzeitraum der Anomalien sehr zuverlässig detektiert wurde. Zusätzlich wird angenommen, dass alle nicht manipulierten Daten frei von Anomalien sind. Da dies jedoch nicht sichergestellt werden kann, ergibt sich auch hier eine Verschlechterung der rechnerischen Fehlerkennzahlen.

Lichtstromdegradation

Bezüglich der Auswertung von Trends in dem realisierten Lichtstrom der Anlagen (potenzielle Degradation) stellte sich bei den geeigneten Anlagen (mindestens Helligkeitssensorik für die Eigenabstrahlung vorhanden) zunächst die Aufgabe, eine Zerlegung der Sensorsignale in Fremdlicht (Tageslicht oder nicht in das System eingebundene Beleuchtung) und Kunstlichtanteil vorzunehmen. Zu diesem Zweck wurde eine Methodik entwickelt, welche die Einschaltzeitpunkte der künstlichen Beleuchtung und die Sonnenaufgangszeit am Standort analysiert und aus der Sprunghöhe der Sensorsignale den statischen Kunstlichtanteil (Eigenabstrahlung des Systems) für den folgenden Tag ermittelt. So konnte für jeden Zeitpunkt der Tageslichtanteil und der Kunstlichtanteil aus einem einzelnen Sensor abgeschätzt werden, siehe Abbildung 24. Der bestimmte Tageslichtanteil wird für die Demonstration des HCL-Raummanagement Use Cases im nächsten Abschnitt verwendet.

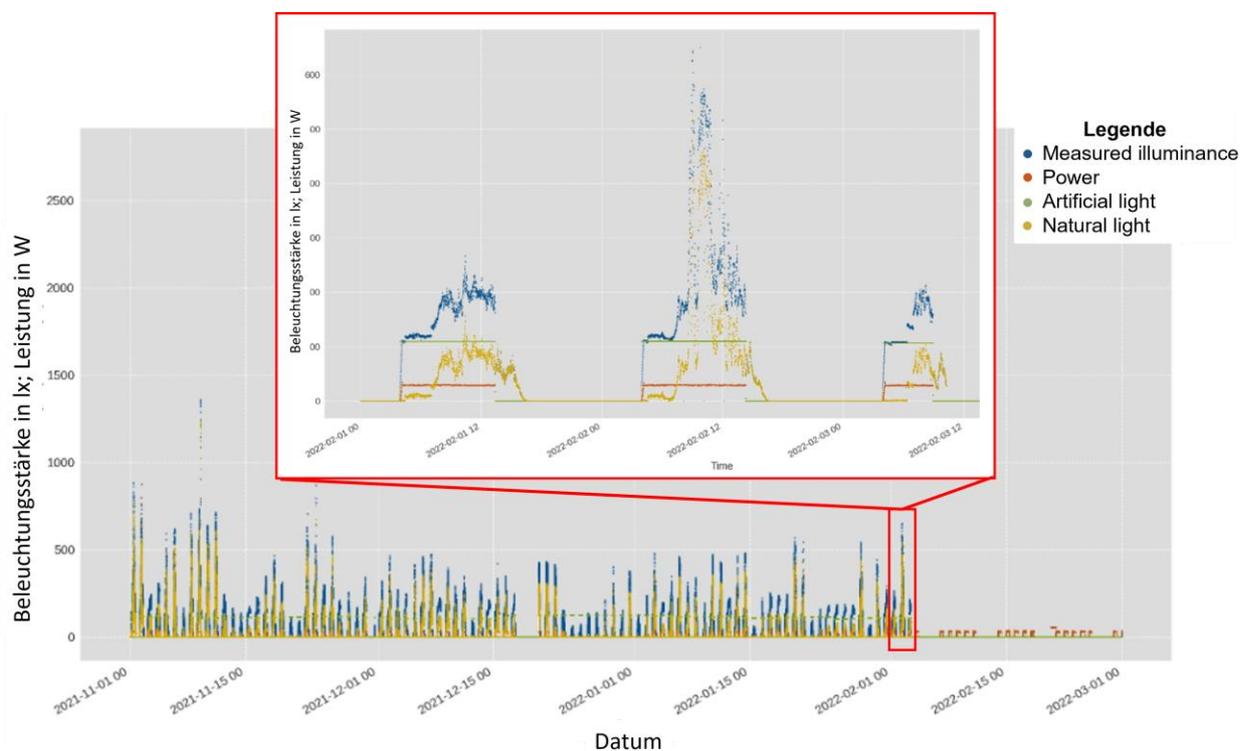


Abbildung 24: Beispielhafte Zerlegung von Sensorsignalen in Tageslicht- und Kunstlichtanteil. Zeitdaten (time) für 01.11.2021 bis 28.02.2022, sowie die Detailansicht für 01.02.2022 bis 03.02.2022 dargestellt.

Zur Bestimmung der künstlichen Beleuchtung werden die Leistungsdaten der Leuchten herangezogen. Anhand der Leistung können Zeiträume identifiziert werden, in denen die Leuchten ausgeschaltet waren und wann diese eingeschaltet wurden. Die Sonnenaufgangsdaten werden genutzt, um zu bestimmen, ob kurz vor dem Zeitpunkt des Einschaltens (ein Sensorwert vor dem Einschalten) die Sonne den Raum nicht beleuchtet hat und dies eine valide Referenz ist. Zur Bestimmung der künstlichen Beleuchtungsstärke wurde ein Zeitraum von 15 Minuten betrachtet aus dem die Median-Beleuchtungsstärke berechnet wird. Der Zeitraum von 15 Minuten nach dem Einschalten darf daher ebenfalls nicht durch Störlicht, wie Sonnenlicht beeinflusst sein. Die Störlichtbeleuchtungsstärke (natural light) kann durch Subtraktion von gemessenem Beleuchtungsstärkewert (measured illuminance) des Sensors und der künstlichen Beleuchtung (artificial light) bestimmt werden.

Mithilfe der so ermittelten künstlichen Beleuchtungsstärke pro Tag kann über die Analyse längerer Zeiträume der realisierte Lichtstrom der Leuchten auf Trends untersucht werden. Abbildung 23 zeigt beispielhaft die Entwicklung der durch die künstliche Beleuchtung realisierten Beleuchtungsstärken auf einem Sensor über einen Zeitraum von 4,5 Monaten.

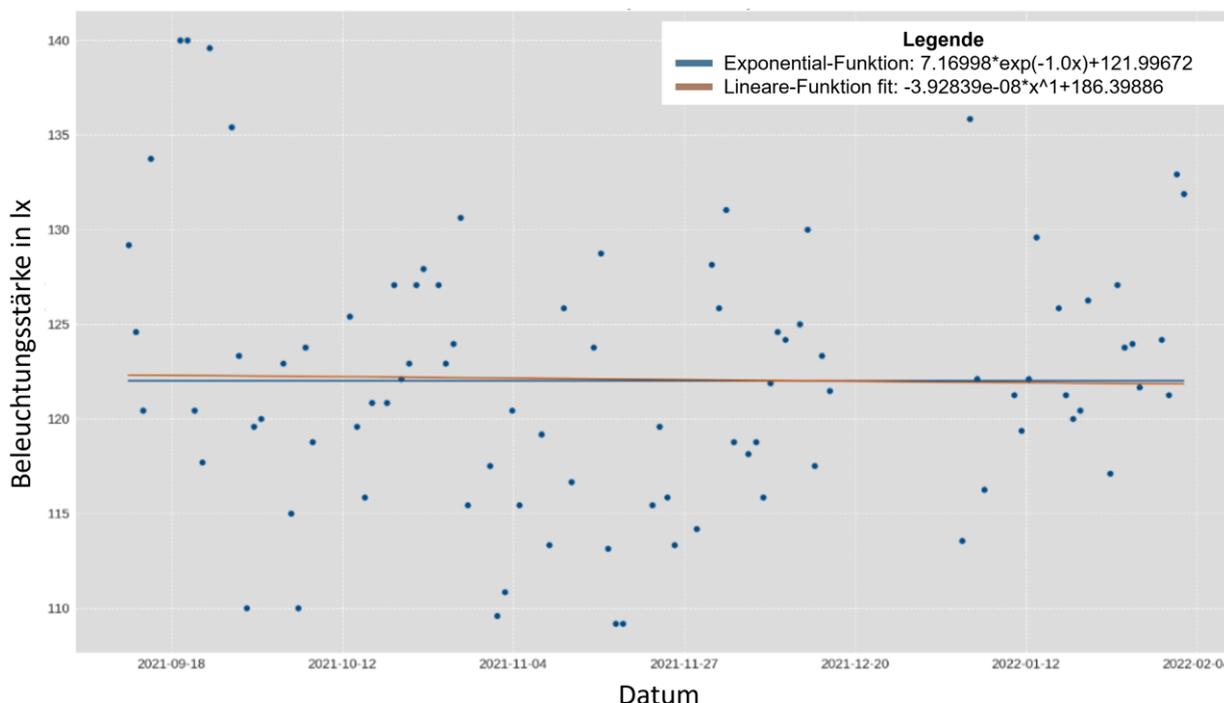


Abbildung 25: Beleuchtungsstärke auf Sensor durch Beleuchtungsanlage. Qualität der Sensorsignale nicht ausreichend, um einen Trend zu identifizieren. Zeitraum: 14.09.2021 bis 04.02.2022.

Exemplarisch wurden die künstlichen Beleuchtungsstärkewerte mit einer Polynom- und einer Exponentialfunktion gefittet. In Abbildung 25 zeigt sich mit beiden Funktionen kein klarer Trend der erfassten Sensor-Beleuchtungsstärke und somit kein Hinweis auf eine feststellbare Leuchtendegradation (für alle Leuchten-Sensor-Systeme/möglichen Standorte in diesem Projekt). Dies ist auf mehrere Ursachen zurückzuführen. So lässt der betrachtete Zeitraum von einem 4,5 Monaten für ein Leuchtensystem, bei herstellenspezifizierte Betriebsbedingungen, keine signifikante Leuchtenalterung erwarten. Zusätzlich waren die Leuchten nicht täglich oder nur für einige Stunden eingeschaltet. Weiterhin ist der Zeitpunkt der Installation der Leuchtmittel nicht bekannt. Weiterhin trägt die Messgenauigkeit des Sensors im betrachteten Messbereich zusätzlich zu einer Verschleierung des ohnehin minimal ausgeprägten Trends bei. Um hier weitere Aussagen treffen zu können, wäre eine tiefer gehende Untersuchung mit neu installierten Leuchtmitteln über einen längeren Zeitraum mit entsprechenden Referenzmessungen mit entsprechend kalibrierten Beleuchtungsstärke-Messgeräten erforderlich. Voraussetzung für diese beiden Untersuchungen ist die Möglichkeit, die künstliche von der natürlichen Beleuchtung zu unterscheiden. Dies konnte lediglich an Standorten erfolgen, an denen regelmäßig die künstliche Beleuchtung zu Zeitpunkten verwendet wird, an denen keine natürliche Beleuchtung vorhanden ist.

Als Fazit kann somit festgehalten werden, dass für eine Überwachung von Lichtstromdegradation zusätzliche Bedingungen erfüllt werden müssen: Idealerweise wird für diesen Use Case Sensorik

direkt in die Leuchte integriert, um unabhängig der (unkontrollierbaren) Außenbedingungen die Emission der Leuchten erfassen zu können. Ein Degradationsmonitoring ist theoretisch auch mit suboptimaler (externer) Sensorik möglich, wenn die Leuchten auch zu Zeitpunkten verwendet werden (eingeschaltet werden können), an denen kein Fremdlicht vorhanden ist. Im Ablauf dieses Projekts waren die zusammenhängenden Beobachtungszeiträume aller Anlagen (Betriebsdauer im Beobachtungszeitraum) klein und die verwendeten Leuchten teilweise bereits seit längerem im Betrieb (potenziell bereits in der asymptotischen Phase der Degradation), sodass keine signifikante (feststellbare) Degradation zu nachzuweisen ist. Die Streuung der Daten in Abbildung 25 ist auf die notwendige (suboptimale) mathematische Abschätzung der künstlichen Beleuchtung sowie die Messungsgenauigkeit der verwendeten Sensorik verbunden mit dem entsprechenden Applikationsszenario zurückzuführen.

2.4.2. Vorhersage von Raumparametern

Bei der Analyse von Sensordaten ist die Modellierung und damit verbunden die Möglichkeit der Vorhersage von kommenden Zuständen ein interessanter Bestandteil. Dazu existieren Klassen unterschiedlicher Verfahren, die abhängig vom Anwendungsfall unterschiedlich gut geeignet sind (siehe Kapitel Zeitreihenvorhersage). In diesem Projekt war das erklärte Ziel dieser Auswertung eine umfassende, für das Wohlbefinden der Nutzenden relevante Eigenschaften des Raumes möglichst orts aufgelöst berechnen und vorhersagen zu können. Die konkrete Eigenschaft ist hierbei abhängig von der Nutzung des Raumes ist muss von den Anlagenbetreibern definiert werden. In Abbildung 26 ist beispielsweise der Verlauf der geschätzten Tageslichtverteilung zu drei Uhrzeiten 05:00, 09:00 und 10:00 Uhr am 04.10.2021 eines Standorts gezeigt. Hierbei können Mitarbeiter in diesem Raum Plätze im Raum bevorzugen, die eine höhere Helligkeit aufweisen und dies aus Abbildung 26 ableiten.

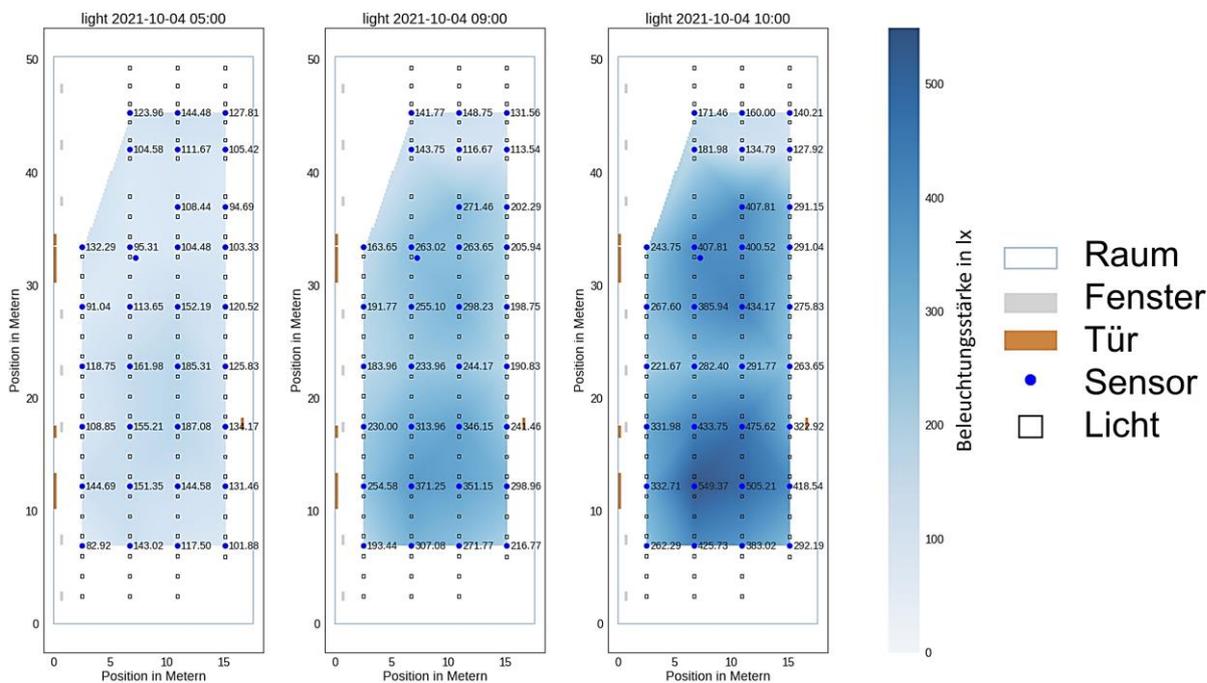


Abbildung 26: Heatmap der an der Decke verbauten Helligkeitssensoren eines Standorts mit Einstrahlwinkel aus Richtung Raumboden.

Die erreichte Vorhersage von über die Sensoren/Leuchtvorschaltgeräte erfassbaren Parametern kann mit wenigen Annahmen noch für eine Aussage zur biologischen Wirksamkeit (Circadian Stimulus, CS Version 2018) der Beleuchtung erweitert werden. Mit einer Annäherung von Truong et al. [31] kann über die Beleuchtungsstärke und die Farbtemperatur (CCT) von Licht der circadiane Stimulus (CS-Wert) abgeschätzt werden. Die Gleichung zur Abschätzung des CS-Wertes ist in Abbildung 27 dargestellt.

$$CS(E_v, CCT_{ANSI}) = \begin{cases} 0.7 - \frac{0.7}{1 + \frac{3.65}{10^6} \left(\frac{(CCT_{ANSI}/[K])^{1.404} E_v}{1000 [lx]} \right)^{1.1500}} & \text{if } (CCT_{ANSI} > 3710 \text{ K}) \ \& \ (R_a > 80) \\ 0.7 - \frac{0.7}{1 + \frac{3.40}{1000} \left(\frac{(CCT_{ANSI}/[K])^{0.783} E_v}{1000 [lx]} \right)^{1.1012}} & \text{if } (CCT_{ANSI} \leq 3710 \text{ K}) \ \& \ (R_a > 80) \end{cases}$$

Abbildung 27: CS-Wert Abschätzung nach Truong et al. [31] mit den Eingangswerten Beleuchtungsstärke und der CCT.

Der CS-Wert beschreibt die circadiane Wirksamkeit des Lichts auf einen Menschen und ab einem Wert über 0.35 gilt das Licht als wirksam [32]. Neben der bereits vorgestellten Zerlegung in Tageslicht- und Kunstlichtanteil der sensorisch erfassten Beleuchtungsstärke wird also die Kenntnis der CCT des Tageslichts und des Kunstlichts vorausgesetzt. Der Tagesverlauf des Sonnenlichts hinsichtlich der CCT kann dabei über Abbildung 28 vereinfacht modelliert werden.

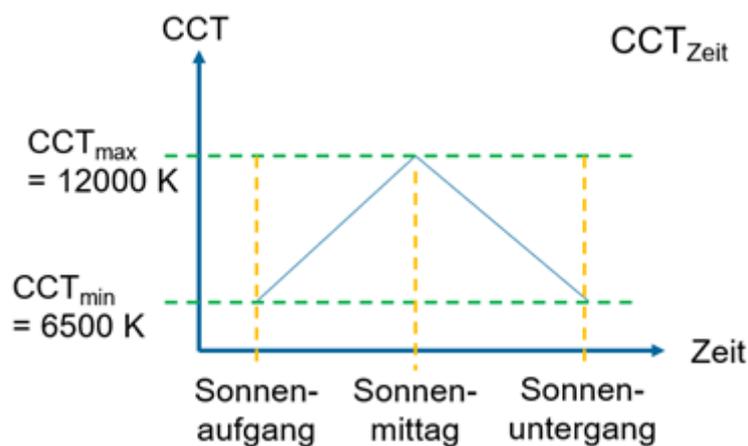


Abbildung 28: Abschätzung der Tageslicht-CCT mit der Annahme, dass die CCT des Tageslichts sich zwischen 6500 K und 12000K im Zeitraum von Sonnenaufgang und Sonnenuntergang bewegt.

Die Bewölkung kann dabei über den Bewölkungsgrad und die folgende Formel eingebracht werden.

$$CCT_{Wolken} = Wolken[\%] * (CCT_{max} - CCT_{min})/100$$

Der zeitaufgelöste Bewölkungsgrad wurde in diesem Projekt standortspezifisch aus der Online-Wetterdatenbank von openweathermap.org abgerufen. Die tageszeitabhängige CCT des Tageslichts ergibt sich dann über

$$CCT_{Tageslicht} = xy2CCT(f_{Zeit} * (CCT2xy(CCT_{Zeit})) + f_{Wolken} * (CCT2xy(CCT_{Wolken}))),$$

wobei $f_{\text{Zeit}} = 0.5$ mit $\{0 \leq f_{\text{Zeit}} \leq 1\}$, $f_{\text{Wolken}} = 1 - f_{\text{Zeit}}$ das Verhältnis zwischen Zeiteinfluss und Wolkeneinfluss auf die CCT des Tageslichts beschreibt. Das Verhältnis von 1:1 wurde empirisch ermittelt. Die Bezeichnung $xy2CCT$ und $CCT2xy$ beschreibt eine Funktion zur Umrechnung von xy -Normfarbwertanteile in die korrelierte Farbtemperatur (CCT), sowie die Umrechnung der CCT in xy -Normfarbwertanteile. Daraus folgt, dass final das bekannte Verhältnis aus Kunstlicht und Tageslicht verwendet werden kann, um die CCT des Mischlichts zu bestimmen:

$$CCT_{\text{gesamt}} = xy2CCT(w_1 * CCT2xy(CCT_{\text{Kunstlicht}}) + w_2 * CCT2xy(CCT_{\text{Tageslicht}})),$$

wobei $w_1 = E_{\text{Kunstlicht}} / E_{\text{sensor}}$ und $w_2 = 1 - w_1$. Über dieses Vorgehen ließe sich der CS-Wert des Lichts approximieren, welches auf den Beleuchtungsstärkesensor fällt. Dies entspricht in den Projektanlagen allerdings nicht der Beleuchtung im Arbeitsbereich der Raumnutzenden, da die Sensoren bis auf eine Ausnahme an den Leuchten verortet waren. Somit ist ein weiterer Schritt nötig, der die gemessene Beleuchtungsstärke in der Leuchtenebene auf eine Beleuchtungsstärkeschätzung für die Arbeitsebene umrechnet:

$$E_v = E_{\text{sensor}} * E_{\text{reference}} / E_{\text{Kunstlicht}},$$

wobei das Verhältnis zwischen Referenzebene und Sensorebene in der jeweiligen Anlage und für den jeweiligen Sensor bestimmt werden müsste. Eine Abschätzung kann mit der Annahme erfolgen, dass auf der Referenzebene eine Beleuchtungsstärke von $E_{\text{reference}} = 500 \text{ lx}$ vorherrscht, was dem Minimum aus der Norm DIN EN 12464-1 für Büroarbeiten entspricht. Dabei wird davon ausgegangen, dass durch die Sensorpositionierung und den Sensorblickwinkel in den betrachteten Anlagen (nach unten in den Arbeitsraum) keine direkte Einstrahlung von Sonnenlicht auf den Sensor zu erwarten ist, und daher der Einfluss der (über den Tagverlauf potenziell veränderlichen) Beleuchtungsgeometrie (Tageslicht zu Kunstlicht) durch das räumliche Integral des Sensorblickfeldes gering ist.

Insgesamt lassen sich die Vorhersagen und abgeleiteten Größen dann kombinieren, um beispielsweise zu Darstellungen wie in Abbildung 29 zu kommen, woraus sich Aussagen für ein anwendungszentriertes Raummanagement ortsaufgelöst ableiten lassen.

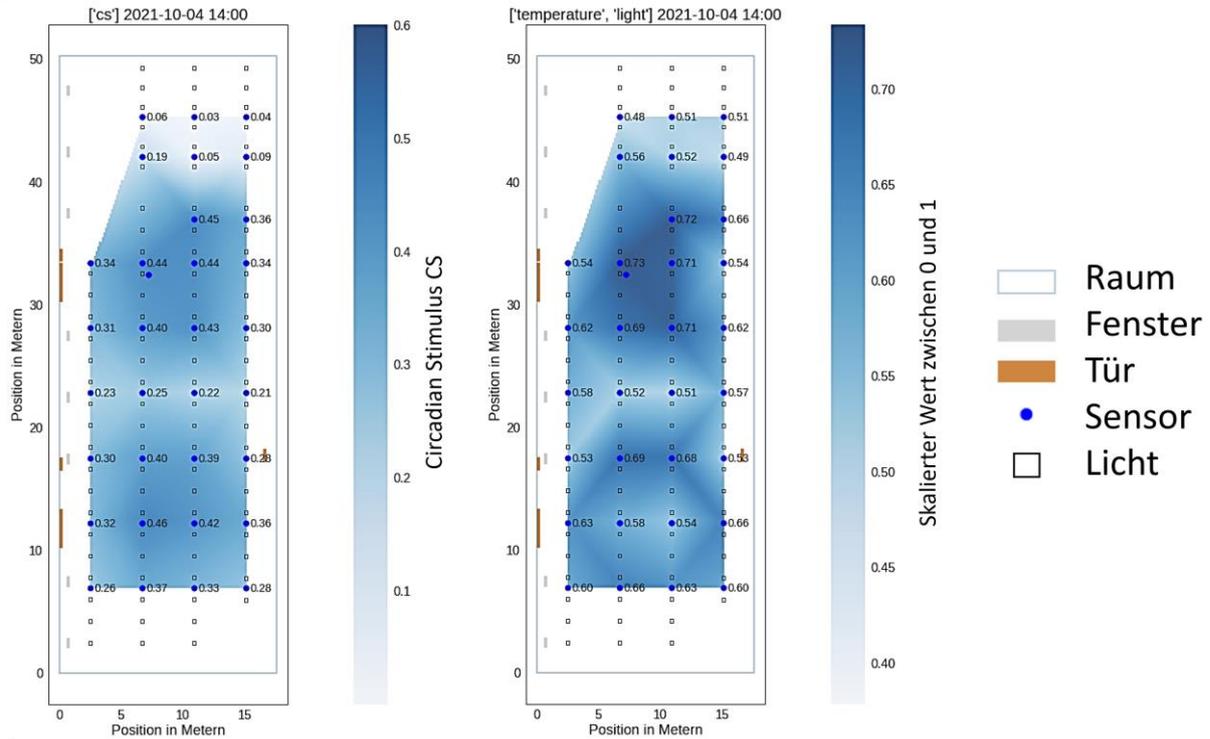


Abbildung 29: Visualisierung kombinierte Raumparameter. Die linke Visualisierung zeigt den CS-Wert am 04.10.2021 um 14:00 Uhr ortsaufgelöst. Die rechte Darstellung zeigt eine Heatmap der Helligkeit kombiniert mit der Temperatur.

Im Beispiel in Abbildung 29 sind zwei Heatmaps zu sehen, die mit unterschiedlichen Parametern erzeugt wurden. Auf der linken Seite ist eine Heatmap auf Basis des CS-Wertes aufgetragen, wobei ein höherer Wert eine höhere biologische Wirksamkeit bedeutet. Für die Wahl des Arbeitsplatzes sollte daher ein Platz mit einer dunkelblauen Einfärbung gewählt werden. Der dargestellte CS-Wert wurde mithilfe der zuvor vorgestellten Abschätzung berechnet. Auf der rechten Seite der Abbildung 29 ist die Temperatur und die Helligkeit einbezogen und ebenfalls stellen dunkelblau eingefärbte Stellen eine bessere Wahl dar. Hierbei ist zu beachten, dass eine niedrigere Temperatur als besser erachtet wurde, daher wird eine niedrigere Temperatur mit einer höheren Helligkeit als gut beachtet. Hierbei können parameterspezifische Optimalwerte (z. B. Temperatur um 22°) und die Polarität der Größen (z. B. größere CS-Werte sind positiv) für die Bewertung vorgegeben werden.

Das Ziel dieses Use Cases, Bereiche in einem Raum zu identifizieren in welchen die besten Arbeitsbedingungen gegeben sind, wurde erreicht. Die Heatmap-Darstellung kann für jeden Standort bestimmt werden und die Korrelationsparameter für die Eignungsbewertung können frei gewählt und an die Aufgabe der Standorte angepasst werden. Die Berechnung des CS-Wertes ergänzt beispielhaft die Bewertung der ortsaufgelösten Raumparameter um eine biologische nicht-visuelle Metrik.

2.4.3. Minimales Set an Sensoren

Die Minimierung des aktiven Sensorsets wurde in zwei Dimensionen betrachtet. Einerseits wurde (bei entsprechender Anzahl) der Frage nachgegangen, wie viel Information sich zwischen Sensoren gleichen Typs in einem Raum überschneiden, wie viele Sensoren gleichen Typs also ohne Informationsverlust abgeschaltet werden könnten. Andererseits ist es von wissenschaftlichem Interesse, welche Zusammenhänge zwischen Sensordaten verschiedenen Typs existieren, die sich zu einer verbesserten Modellierung der Raumparameter(vorhersage) durch einen multivariaten Ansatz nutzen lassen. Zur Beantwortung dieser Fragen wurde eine Kerndichteschätzung (engl. Kernel density estimation, KDE) durchgeführt. Diese Methode kann für alle beliebigen Sensoren angewendet werden und zeigt die Verteilung der Sensorwerte mittels einer Wahrscheinlichkeitsverteilung. In Abbildung 30 sind für einen Standort die vorhandenen 27 Helligkeitssensoren und deren KDE's dargestellt. Zur Vermeidung eines hohen Anteils an 0 lx Werten wurde in der Abbildung 30 nur die Werte zwischen Sonnenaufgang und Sonnenuntergang betrachtet. Ein klarer Unterschied im Verhalten von verschiedenen Sensoren ist an den verschiedenen Kurvenverläufen erkennbar. Einige Sensoren weisen zwei Peaks auf, weitere bleiben über einen längeren Bereich annähernd konstant und ein Sensor erfasst nur sehr geringe Beleuchtungsstärken.

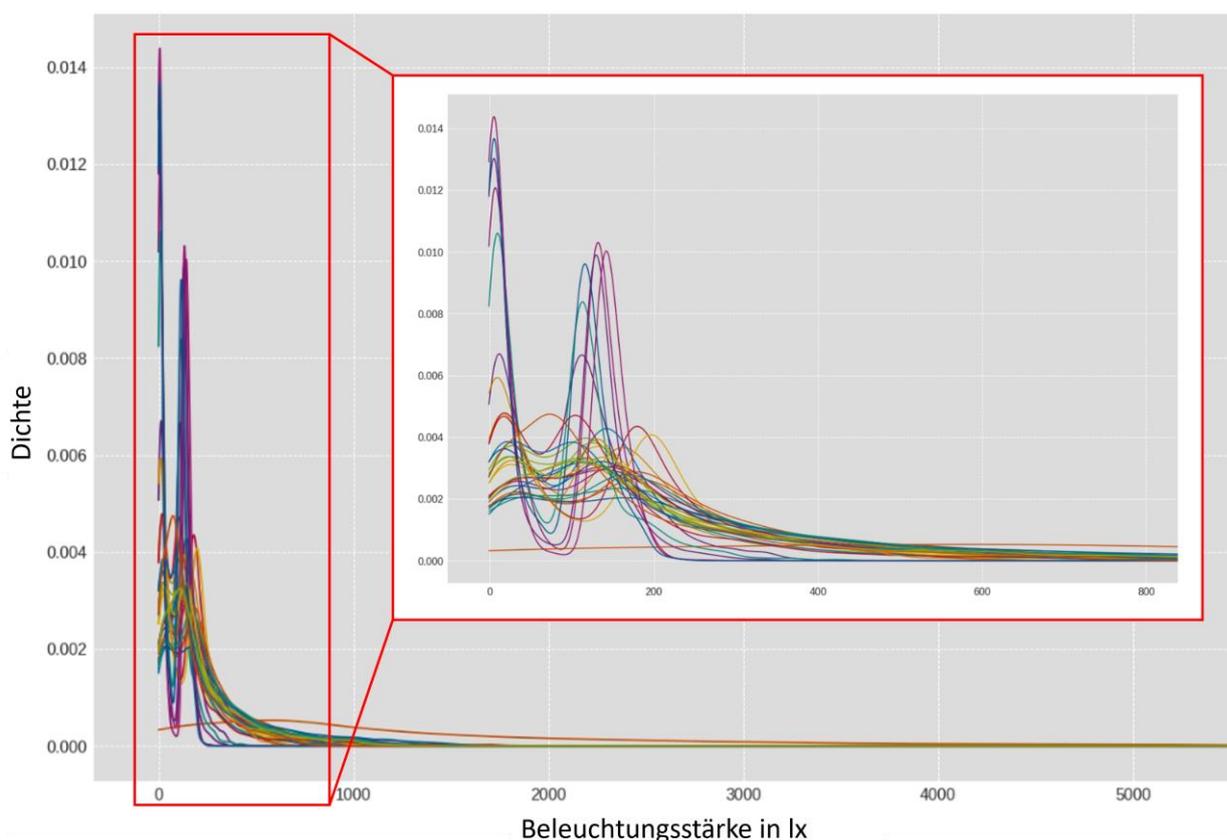


Abbildung 30: Kerndichteschätzung von allen Helligkeitssensoren eines Standorts.

Für eine genauere Betrachtung wurden die KDE's mit einem Clustering-Algorithmus in vier Klassen eingeteilt. Abbildung 31 zeigt die jeweiligen KDE-Verläufe in Grau, zugeordnet in die jeweiligen Cluster 0 - 3. In rot ist der Mittelwert pro Klasse dargestellt.

Die erste Frage für diesen Use Case zur Minimierung der Sensoranzahl, welche die Überlappung an Informationen zwischen den Sensoren untersucht, kann hiermit beantwortet werden. Mittels einer KDE kann für jeden Sensortyp in einem Standort mittels Clustering sich gleich verhaltende Sensoren identifiziert werden. Dies hat zur Folge, dass ebenfalls im Minimalfall nur ein Sensor pro Cluster benötigt wird, der das Verhalten für alle Sensoren im Cluster stellvertretend empfängt. Im Beispiel aus Abbildung 31 wären somit vier Beleuchtungssensoren ausreichend.

Dies schließt ein, dass im Vorfeld Messungen von Sensoren an verschiedenen Positionen im Raum vorhanden sind. Es ist ausreichend, wenn diese Informationen von zusätzlichen Sensoren von wenigen Messtagen vorliegen, um eine Korrelationsanalyse mittels KDE durchzuführen und mathematische Zusammenhänge zwischen den Positionen und dem minimalen Set an Sensoren zu ermitteln. Durch den mathematischen Bezug zwischen realen Sensoren und beliebiger Position kann im Anschluss für eine beliebige Anzahl an Punkten im Raum das Sensorverhalten abgeschätzt werden.

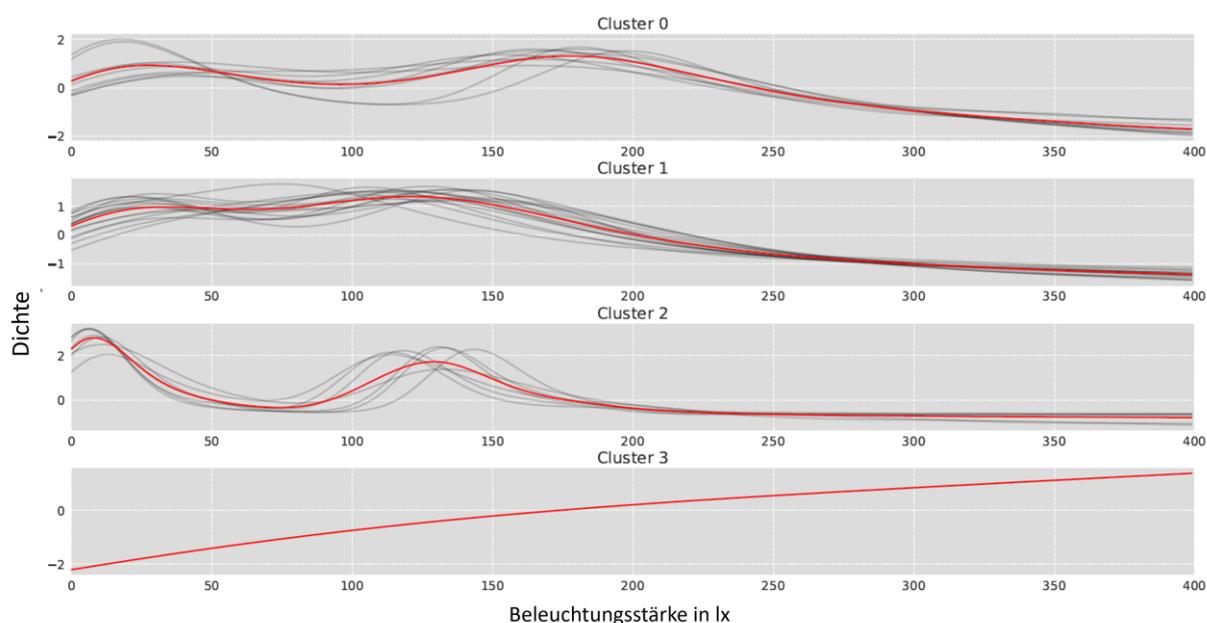


Abbildung 31: Darstellung der vier KDE-Cluster der Beleuchtungsstärkesensoren.

Mittels der BoT Beschreibung des Standorts wurde automatisiert der Grundriss in Abbildung 32 erstellt und die Sensorpositionen mit den jeweiligen Cluster-Nummern eingetragen. Aus dieser Darstellung ist ersichtlich, dass Sensoren in Cluster 2 sich Großteils rechts in einem separaten Bereich befinden. Verläufe in Cluster 1 beschreiben fensternahe Orte im Raum. Cluster 0 spezifiziert Positionen in der Mitte des Raums. Der Sensor in Cluster 3 stellt eine Besonderheit dar und wurde aus diesem Grunde in einem Cluster individuell betrachtet, da dieser auf Tischhöhe angebracht ist, mit Einstrahlrichtung von oben. Alle anderen Sensoren sind an der Decke befestigt und das Licht strahlt von unten ein.

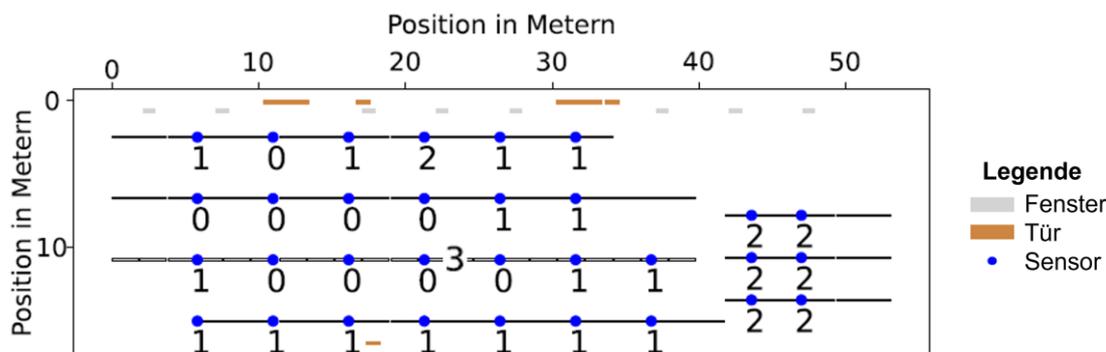


Abbildung 32: Grundriss eines Standorts mit Zuordnung der Sensoren zu den vier Clustern 0 – 3 aus Abbildung 31. Cluster 3 beinhaltet nur einen Sensor auf Tischhöhe mit Einstrahlrichtung von oben. Alle weiteren Sensoren sind an der Decke befestigt mit Einstrahlrichtung von unten.

Mithilfe der Darstellung der Sensorcluster im Grundriss (siehe Abbildung 32) können verschiedene Beleuchtungszonen identifiziert werden und eventuell unerwünschte Unterschiede durch Anpassung der Beleuchtungsanlage behoben werden.

Mithilfe dieses allgemeinen Vorgehens können beliebige Sensortypen in unterschiedlichen Räumen untersucht und vollautomatisiert bearbeitet werden.

Zusätzlich zur Bestimmung des minimalen Sets an Sensoren wurde das Abtastintervall untersucht. Hierbei besteht der Kompromiss zwischen einem möglichst geringen Informationsverlust durch große Intervalle und einer Reduzierung der Datenmenge und damit des Rechen- und Zeitaufwands. Für die Bestimmung des optimalen Abtastintervalls wurden diese variiert und die Zeitreihenvorhersage mittels ARIMA, XGBoost, LSTM und HMM-Algorithmen erneut angestoßen. Bei der Simulation der Abtastintervalle ergaben sich zwei Möglichkeiten:

- direkte Simulation längerer Intervalle durch die Verwendung des ersten Wertes je Intervall und verwerfen der übrigen Werte. ("first")
- Simulation durch Mittelwertbildung über das zusammengefasste Intervall ("mean")

Die Modelle wurden anschließend mit den vier unterschiedlichen Abtastintervallen 10, 20, 30 und 60 Minuten und den zwei Möglichen Datenreduktionsmethoden („mean“, „first“) neu trainiert. Die Tests erfolgten mit den Sensorwerten für die Beleuchtungsstärke, Temperatur und Luftdruck. Die Daten dieser Sensortypen lagen in kürzeren Zeitintervallen von zehn Minuten vor und konnten daher für größere Abtastintervalle zusammengefasst werden.

Die Evaluation wurde anhand des RMSE als Metrik durchgeführt. Um eine Vergleichbarkeit zu schaffen, wurde ebenfalls das Hyperparameter Tuning erneut durchgeführt, um die besten Parameter für die jeweiligen Intervalle und Modelle zu finden. In Tabelle 4 ist das Ergebnis zu sehen. Für diese drei Datentypen weist XGBoost in den meisten Fällen den geringsten RMSE auf. Weiterhin zeigt sich ein unterschiedlich starker Einfluss des Abtastintervalls auf die Vorhersage bei den verschiedenen Datenpunkten. Bei der Beleuchtungsstärke entstehen bei größeren Abtastintervallen bessere Scores. Hingegen ist dies für die Temperatur und den Luftdruck nicht der Fall. Dieses Verhalten lässt sich durch die Art der jeweiligen Zeitreihen erklären. So besitzt die Beleuchtungsstärke einen Wertebereich zwischen 0 lx und ca. 1800 lx und weist durch Ein- und Ausschalten der künstlichen Beleuchtung Sprünge von etwa 800 lx auf.

Sprünge in Daten weisen häufig einen erhöhten Fehler auf, kontinuierliche Verläufe mit geringerer Steigung sind weniger anspruchsvoll für die Vorhersage. Hinzu kommen starke Schwankungen durch die natürliche Beleuchtung, die den kontinuierlichen Verlauf sprunghaft verändern. Die Verwendung größerer Abtastintervalle führt dabei vor allem aufgrund der Mittelwertbildung zu besseren Vorhersagen (Sprünge werden durch die Mittelung abgeflacht). Bei Sensortypen mit geringeren Schwankungen wie beispielsweise der Temperatur oder noch längerfristigen Veränderungen wie dem Luftdruck hat das Abtastintervall nur einen geringen Einfluss oder eine schnellere Abtastung führt zu geringfügig besseren Ergebnissen. Dies kann auf eine größere Anzahl an verfügbaren Trainingsdaten, aufgrund des geringeren Abtastintervalls, zurückgeführt werden. Als Fazit zeigt sich, dass die Wahl des Abtastintervalls bei langsam veränderlichen Größen nur einen geringen Einfluss hat und bei zeitlich schnell veränderlichen Größen (vor allem aufgrund der Mittelwertbildung) mit größeren Abtastintervallen bessere Ergebnisse erzielt werden können.

Tabelle 4: Untersuchung unterschiedlicher Abtastintervalle für die Vorhersage von Beleuchtungsstärke, Temperatur und Luftdruck. Die Zahlen repräsentieren den RMSE und die Reduzierung der Datenmenge wurde mittels Mittelwertbildung durchgeführt.

Datenpunkt	Modell	10 min	20 min	30 min	60 min
<i>Beleuchtungsstärke in lx</i>	ARIMA	527,65	478,47	-	478,92
	XGBoost	227,49	208,04	197,62	172,68
	LSTM	336,74	333,21	341,55	312,86
	HMM	19947,4	328,41	258,18	259,82
<i>Temperatur in °C</i>	ARIMA	0,50	0,51	0,52	0,54
	XGBoost	0,14	0,14	0,14	0,14
	LSTM	0,60	0,36	0,34	0,46
	HMM	0,57	0,48	0,51	0,42
<i>Luftdruck in hPa</i>	ARIMA	1,22	1,26	1,30	1,42
	XGBoost	0,57	0,59	0,63	0,63
	LSTM	2,11	1,97	2,14	2,27
	HMM	0,63	0,72	0,47	0,63

2.4.4. Datenanalyse-Toolbox

Die vorgestellten Methoden in den Use Cases wurden im Hinblick auf die Heterogenität der acht Projekt-Anlagen konzipiert und auf eine allgemeine Herangehensweise geachtet. Die Vereinheitlichung der Datenstrukturen in der IoL-Cloud ermöglichte diese anlagenübergreifende Betrachtung. Aufgrund der hohen Diversität der untersuchten Standorte wurde beispielhaft die Auswertung in diesem Kapitel für verschiedene Standorte dargestellt. Für eine weitere Evaluation der Sensordaten und der genauen Betrachtung der implementierten Methoden wird zusätzlich zu diesem Bericht eine Datenanalyse-Toolbox den acht Standorten zur Verfügung gestellt.

Alle verwendeten Methoden und Algorithmen wurden vollständig in der Programmiersprache Python in Version 3.9 implementiert. Die Datenanalyse-Toolbox wurde in einem Jupyter-Notebook realisiert. Ein Jupyter-Notebook ist eine Web-basierte Entwicklungsplattform, die ermöglicht Code Fragmente in einer über den Web-Browser gesteuerten Umgebung auszuführen und diese gleichzeitig in formatierten Text (Dokumentation mit Konfigurations- und Anwendungshinweisen) einzubinden. Ein weiterer Vorteil ist, dass die Ergebnis-Ausgabe der

Code-Fragmente direkt im Jupyter-Nodebook unter dem Code erfolgt. Die Datenanalyse-Toolbox fußt damit auf einer einfach einzurichtenden und schnell zu benutzenden Plattform.

Für die Einrichtung der Datenanalyse-Toolbox wurde eine Setup-Batch-Datei erzeugt, die die gesamte Installation unter Windows übernimmt. Hierfür ist Windows 10 oder neuer nötig. Mit dem Beenden der Installation steht eine weitere Batch-Datei zur Verfügung, die das Jupyter-Notebook im Browser startet. Abbildung 33 zeigt die Ansicht mit Beschreibungstext, Code und Parametrisierung.

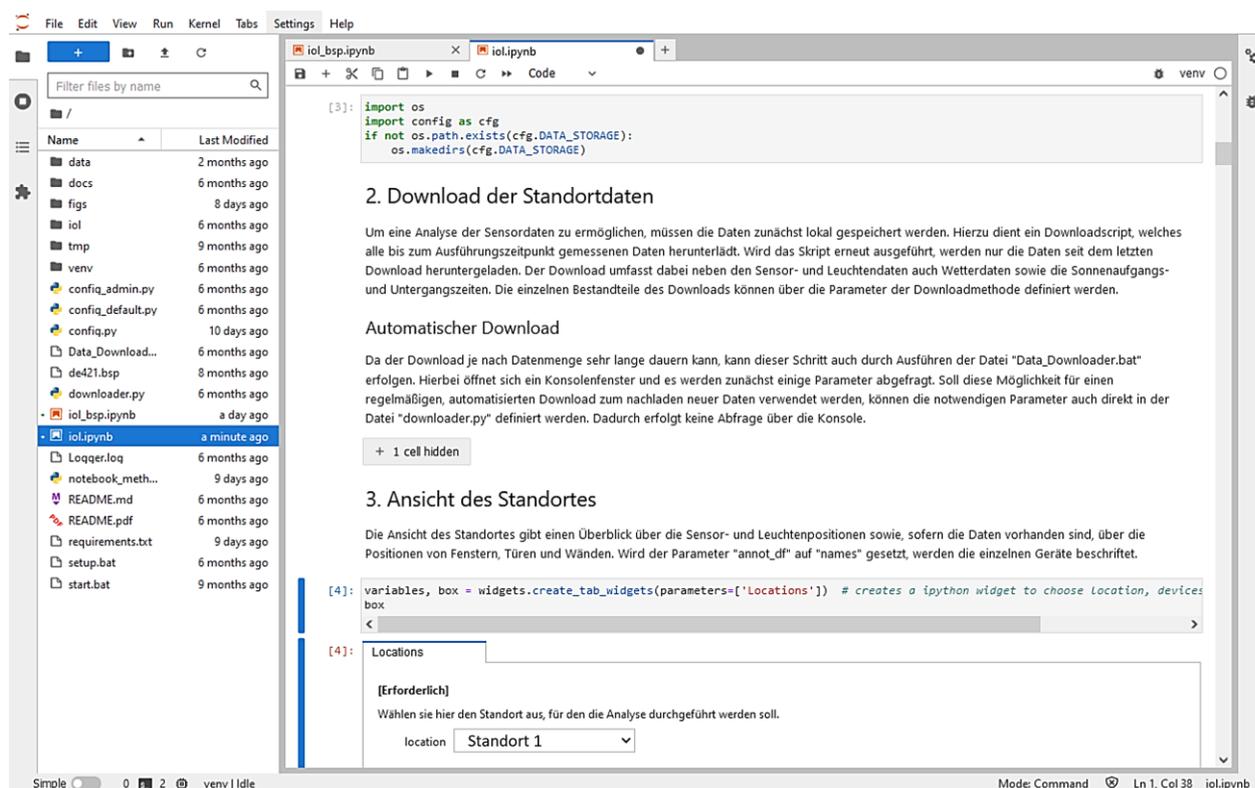


Abbildung 33: Jupyter-Notebook Entwicklungsumgebung der Datenanalyse-Toolbox mit Beschreibungstext, Code und Parametrisierung.

Das Notebook dient der Anwendung verschiedener Analysemethoden auf vorhandene Sensordaten. Es ist dabei in mehrere Bereiche unterteilt, die im Folgenden erläutert werden.

1. **Erstellung einer Konfigurationsdatei:** Um die Analyse an den jeweiligen Standort anzupassen, muss eine Konfiguration erstellt werden. Diese beinhaltet den Speicherort der Daten sowie die Koordinaten der Anlage, falls auch Wetterdaten und Sonnenaufgangs- und Untergangszeiten berücksichtigt werden sollen. Außerdem müssen die jeweiligen Autorisierungstokens angegeben werden. Die Konfigurationsdatei ermöglicht zudem eine Anpassung der Log-level, sowie eine Anpassung des Farbschemas der Visualisierungen.
2. **Download der Standortdaten:** Für die Durchführung der Analyse müssen die Anlagendaten lokal gespeichert sein. Ein Download der Daten kann dazu mithilfe eines beigefügten Skriptes angestoßen werden.

3. **Ansicht des Standortes:** Die Ansicht des Standortes gibt einen Überblick über die Sensor- und Leuchtenpositionen sowie, sofern die Daten vorhanden sind, über die Positionen von Fenstern, Türen und Wänden.
4. **Explorative Datenanalyse:** Die explorative Datenanalyse dient dazu einen ersten Überblick über die vorhandenen Daten zu erhalten. Hierbei ist eine Ansicht der Rohdaten, die Erzeugung einer statistischen Übersicht sowie verschiedene Visualisierungen möglich. Diese beinhalten einen Scatterplot der alle Daten anzeigt, einen Pairplot und eine Übersicht der Korrelationen zum Erkennen gegenseitiger Abhängigkeiten.
5. **Anwendung der Zeitreihenvorhersage:** In diesem Abschnitt kann eine Vorhersage ausgehend von den gemessenen Daten mit Hilfe vier verschiedener Machine Learning Algorithmen (ARIMA, XGBoost, LSTM, HMM) erfolgen. Dabei können die Modell- und Vorverarbeitungsparameter entweder manuell festgelegt oder mittels Hyperparameterertuning bestimmt werden. Ein Vergleich der Modelle, auch mit zwei zusätzlichen Baseline-Modellen, ist ebenfalls möglich.
6. **Detektion von Anomalien:** Die Detektion von Anomalien erfolgt über verschiedene Methoden: Verwendung von gleitendem Mittelwert und Standardabweichung, Verwendung eines Clusteralgorithmus (HDBSCAN) und Verwendung der Modellierungen von bereits trainierten Modellen (ARIMA, XGBoost, LSTM, HMM).
7. **Analyse der Beleuchtung:** In diesem Kapitel wird die Beleuchtung genauer betrachtet. Dies beinhaltet eine Trennung von künstlichem und natürlichen Anteil der gemessenen Beleuchtungsstärke, eine Darstellung der Alterung der Leuchten, eine Kerndichteschätzung sowie eine Betrachtung verschiedener Beleuchtungszonen.
8. **Use Case: Sensoreinsparung:** Der Use Case Sensoreinsparung beschäftigt sich mit der Frage nach der Anzahl und Art der benötigten Sensoren. Hierbei ist das Ziel, eine ausreichende Erfassung der Umgebung mit einer möglichst geringen Anzahl an Sensoren zu erreichen. Dazu sollten die Sensorwerte möglichst frei von Korrelation zu anderen Sensoren, die Anzahl der Sensoren je Sensortyp dem Raum angepasst und die Abtastung dem Sensortyp angepasst sein.
9. **Use Case: Dynamisches Monitoring der Leuchten-Betriebszuverlässigkeit:** Die Verwendung von Sensoren und Leuchtendaten über einen längeren Zeitraum ermöglicht die Erkennung von Unregelmäßigkeiten über eine Anomaliedetektion. Da diese Unregelmäßigkeiten, wenn sie nicht absichtlich hervorgerufen wurden, auf Fehler in der Anlage bzw. den Leuchten hindeuten, ist so ein dynamisches Monitoring der Leuchten-Betriebszuverlässigkeit möglich.
10. **Use Case: Raummanagement:** Der Use Case Raummanagement beschäftigt sich mit der räumlichen Vorhersage von Unterschieden der Datenpunkte. Die Messwerte der Sensoren und abgeleitete Größen wie der CS-Wert können (ggf. mathematisch verknüpft) zu einem bestimmten Zeitpunkt im Grundriss dargestellt und darüber Bewertungen von bestimmten Raumpositionen unter standortspezifischen Kriterien durchgeführt werden.

Zur zusätzlichen Beschreibung dient eine Readme.pdf Datei, welche weitere detaillierte Informationen zur Verwendung und der Ausgabe in den einzelnen Bereichen des Jupyter-Notebooks gibt. Die drei Informationsquellen, dieser Bericht, die Readme-Datei und das Notebook selbst beschreiben den Ablauf von der Installation bis zur Auswertung, sodass grundlegende Programmierkenntnisse für einen Umgang mit der Toolbox ausreichen.

Die gesamte Auswertung für diesen Bericht basiert auf der Datenanalyse-Toolbox und kann somit pro Standort individuell durchgeführt werden.

2.4.5. Zielabgleich

Die zu AP4 definierten Ziele

- Monitoring zur Betriebszuverlässigkeit fertiggestellt (FE 1+2)
- Vorhersage von Raumparametern erstellt (FE 1+2)
- minimales Set an Sensoren zur Use Cases Abbildung bestimmt (FE 1)
- Datenanalyse-Toolbox zusammengestellt (FE 1)

wurden vollständig erreicht.

2.5. AP5 Betrieb und Wartung der Anlagen

In AP5 wurden die Aufgaben um den Laufzeit-Betrieb der IoL-Cloud und der angebundenen Anlagen gebündelt.

2.5.1. Vermessung der Anlagen

Von einer Umsetzung der für Projektbeginn geplanten Vermessung der Anlagen musste aufgrund der pandemischen Lage abgesehen werden. Zweck der Vermessungen wäre neben der Erfassung der geometrischen Verhältnisse auch eine lichttechnische Charakterisierung der Anlagen gewesen, um eine messtechnische Referenz zu den Sensordaten zu besitzen. Die geometrischen Daten wurden schließlich über iterative Rückfragen bei den Anlagenbetreibern gewonnen und durch die FE2 direkt in die erarbeitete Datenstruktur integriert. Die Datenauswertung wurde so konzipiert, dass an entsprechenden Stellen getroffene Annahmen über die lichttechnischen Parameter der Anlagen durch reale Messwerte ersetzt werden können. Die aus dem Stand der Technik und Wissenschaft abgeleiteten Annahmen schränken dabei den Erkenntnisgewinn im Hinblick der Arbeitshypothesen nicht ein.

2.5.2. Zyklische Prüfung der Konnektivität und Datenqualität der Anlagen

Für jede Anlage wurden innerhalb der IoL-Cloud sogenannte Deadman-Checks implementiert. Diese automatisierten Überprüfungen überwachen die in der Datenbank eingehenden Daten und signalisieren, wenn von einer der Anlagen über einen individuellen Zeitraum keine Daten mehr in der Datenbank eingehen. Da die Übertragungsintervalle der Anlagen zwischen „einmal täglich“ und „sekündlich“ variierten, wurden die Deadman-Checks Anlagen-individuell eingerichtet, um die unterschiedlichen Sende- bzw. Abruffrequenzen der Anlagen zu berücksichtigen und eine möglichst zeitnahe Benachrichtigung bei Fehlern zu ermöglichen.

Ausbleibende Daten sind auf zwei primäre Gründe zurückzuführen:

1. Anlagenseitige Updates mit Einfluss auf die Datenstruktur: Während der Projektlaufzeit kam es vereinzelt zu unangekündigten Änderungen im Bereich der Hardware oder Software der Testanlagen. Diese Abweichungen gegenüber der abgebildeten (Daten-)Struktur verhinderten ein vollständiges Payload-Mapping, was

zu einem entsprechenden Fehler-Status führte. Diese Ausfälle konnten zumeist kurzfristig nach Rücksprache mit den technischen Ansprechpartnern der Testanlage durch individuelle Mapping-Updates behoben werden.

2. Gateway-Ausfälle:

Die andere Ursache für Alarmierungen durch die Deadman-Checks waren anlagenseitige, kommunikationstechnische Ausfälle (defekte Komponenten, unterbrochene WAN-Anbindung, etc.). In diesem Fall wurden die technischen Ansprechpartner durch FE2 über den Anlagenstatus informiert.

Ergänzend zu den automatisierten Prüfungen wurde die Datenqualität regelmäßig manuell über vordefinierte Dashboard-Visualisierungen überwacht.

2.5.3. Updates der Cloudanbindung

Während der Projekt-Laufzeit wurde die IoL-Cloud durch FE2 zunächst aufgebaut und anschließend administriert und funktional erweitert. Teil dieser Arbeiten war die agile Weiterentwicklung der Anbindungen unter Berücksichtigung der Anforderungen der jeweiligen Anlagen. In diesem Zuge wurde zum Beispiel in der Docker-Umgebung zusätzliche MQTT-Instanzen hinzugefügt, um den Anlagenbetreibern, die diesen Weg zur Datenübermittlung gewählt haben, eine Lösung zu bieten, welche die Anforderungen an den Datenschutz gewährleistet.

Weiterhin wurden kontinuierlich Optimierungen (z.B. Load-Balancing-Maßnahmen, Komprimierung der Mapping-Strategien, etc.) und sicherheitsrelevante Aktualisierungen durchgeführt.

2.5.4. Zielabgleich

Die in AP5 gemäß Antrag definierten Ziele

- Vermessung der Anlagen durchgeführt (FE 1)
- Zyklische Prüfung der Konnektivität und Datenqualität der Anlagen (FE 2)
- Updates der Cloudanbindung (FE 2)

wurden zum Teil erreicht.

3. Zusammenfassung und Ausblick

3.1. Anbindung an IoL Cloud

Im Rahmen des Projektes wurden acht Testanlagen im Hinblick auf die verwendeten Datenformate, -modelle und Kommunikationsprotokolle analysiert. In einer initialen Erhebungsphase wurden anhand von Fragebögen licht- und kommunikationstechnische Eigenschaften ermittelt. Basierend auf diesen Erkenntnissen wurden Konzepte erarbeitet, um die Daten der verschiedenen Testanlagen an eine zentrale IoL Cloud zu übertragen.

Bereits in dieser initialen Erhebungsphase zeigte sich, dass auch bei gleichartigen Funktionen (z.B. Dimmlevel auslesen) in den Testanlagen sehr heterogene Datenformate zur Anwendung kamen.

Aufgrund der vorherrschenden Corona-Bedingungen während der Projektlaufzeit wurde in Abstimmung mit den teilnehmenden Unternehmen und dem Projektausschuss eine gegenüber dem Antrag abweichende Vorgehensweise zur Anlagenanbindung beschlossen und die Testanlagen sukzessive nach ihrer jeweiligen Verfügbarkeit an die IoL Cloud angebunden. Die initialen Informationen zu den Testanlagen wurde somit kontinuierlich um Realdaten erweitert.

3.2. Semantik

Um im Rahmen der Projektlaufzeit eine maximale Datenbasis für die in AP4 durchzuführenden Arbeiten aufzuzeichnen, wurden die jeweiligen Anlagendaten unmittelbar nach Erteilung der Zugriffsberechtigung anhand der ausgewählten Ontologien unter Anwendung des WoT semantisch annotiert und in einer Zeitreihendatenbank aufgezeichnet.

Mit steigender Anzahl angebundener Anlagen wurden die Anforderungen an ein gemeinsames, Hersteller-übergreifendes Datenformat deutlicher und die Konzeption eines einheitlichen Datenformates wurde in eine prototypische Implementierung überführt. Hierbei erfolgte ein kontinuierlicher Abgleich mit den Anforderungen aus AP4.

3.3. Use Cases

Die drei Use Cases zum Monitoring der Betriebszuverlässigkeit, der Vorhersage von Raumparametern, sowie die Untersuchung des Sensoreinsparpotenzials wurden im Detail erläutert. Die Betriebszuverlässigkeit wurde anhand einer Anomalie-Erkennung implementiert und die Funktionsweise konnte mit simulierten Anomalien gezeigt werden. Die Untersuchung der Lichtstromdegradation wies keinen Abfall in der Helligkeit auf. Im zweiten Use Case wurden voll automatisiert Heatmap-Darstellungen mithilfe des Grundrisses aus der BoT Beschreibung erzeugt. Eine beliebige Anzahl und Art an Sensortypen kann von den Standorten genutzt werden, um dabei für einen Satz von Ziel-Parametern die optimale Position im Raum zu identifizieren. In diesem Zug wurde eine prototypische Abschätzung des circadianen Stimulus anhand der Beleuchtungsstärke und des Bewölkungsgrads vorgestellt. Im dritten Use Case konnte mittels Kerndichteschätzung sich ähnlich verhaltende Sensoren erkannt und gruppiert werden. Um die Datenmenge der Sensoren zu reduzieren und gleichzeitig den Informationsgehalt hochzuhalten, wurden auch die Abtastintervalle untersucht. Als Resultat sollten schnell veränderliche Werte alle

30 Sekunden bis 10 Minuten abgetastet werden und sich langsam verändernde Werte im Bereich von 60 Minuten. Die vollständige Implementierung wurde in einer Datenanalyse-Toolbox (Jupyter-Notebook mit Dokumentation, Konfigurations- und Anwendungshinweisen) zusammengefasst und ermöglicht die vollständige individuelle Auswertung für jeden Standort durch die Unternehmen selbst.

3.4. Ausblick

In regelmäßigen Diskussionen mit dem Projektbegleitenden Ausschuss wurden weitergehende vorwettbewerbliche Forschungsfelder thematisiert, welche über die Ansätze eines einheitlichen IoL-Datenmodells hinausgehen. So wurde u.a. eine Erweiterung des in IoL verfolgten Ansatzes um die zusätzliche Beschreibung von **Aktorik** als sinnvoll bewertet. Dies würde auch im Hinblick auf neue Use Cases eine breitere Anwendung ermöglichen. Die Forschungseinrichtungen haben hierzu bereits in der Abschlussphase des Projektes Planungsaktivitäten aufgenommen und stehen diesbezüglich auch im Austausch mit dem Fachverband Licht des ZVEI bzw. mit der Forschungsvereinigung Elektrotechnik.

4. Erläuterung zur Verwendung der Zuwendung

Im Folgenden werden die geleisteten Personenmonate (Einzelansatz A.1, wiss.-techn. Personal) für die jeweilige Forschungseinrichtung aufgeführt.

Jahr	FE 1	FE 2
2020	2,0	-
2021	13,35	8,0
2022	10,5	20,0
2023	1,0	-
<i>Gesamt</i>	<i>26,85</i>	<i>28,0</i>

Die geleistete Arbeit entspricht in vollem Umfang dem begutachteten und bewilligten Antrag und war damit notwendig und angemessen.

Es wurden keine gewerblichen Schutzrechte erworben oder angemeldet.

5. Ergebnistransfer in die Wirtschaft

Nachfolgend werden die bereits durchgeführten und die noch geplanten Transfermaßnahmen erläutert.

5.1. Veröffentlichungen und Wissenstransfer

Die Inhalte und entwickelten Ergebnisse des Projektes IoL werden in den fortlaufend genannten Maßnahmen an die Lichtbranche, sowie Branchenübergreifend veröffentlicht, um den Transfer in die Wirtschaft zu unterstützen. Die Resultate und Feststellungen aus dem Projekt werden in einem Artikel in einer branchenspezifischen Fachzeitschrift, wie beispielsweise der „Licht“ publiziert und tragen zum Ergebnistransfer an alle Unternehmen in dieser Branche bei. Der Artikel wird nach Abschluss des Projektes mit der Verknüpfung von allen erarbeiteten Schlussfolgerungen verfasst. Weiterhin hat das Projekt eine große Bandbreite an aktuell relevanten Themen abgedeckt und die Teilgebiete der Farbsensorik, intelligenten Leuchten-Ansteuerung und LED-Alterung wird durch FE1 im Detail weiter untersucht und in Dissertationen zu diesen Themen veröffentlicht werden. Zusätzliche Veröffentlichungen in Fachjournals zu diesen Dissertationsthemen sind geplant. Die Inhalte wurden außerdem in Form von insgesamt vier Abschlussarbeiten (Bachelor-/Masterarbeit) an den Forschungseinrichtungen näher untersucht. Während und nach der Projektlaufzeit werden in den Vorlesungen Smart Lighting I/II und KI2 bei den Forschungseinrichtungen die Themen von Internet of Light an Studierende weitergegeben, um das Wissen langfristig in die Wirtschaft mit einzubringen.

5.2. Durchgeführte Transfermaßnahmen (von Projektbeginn bis heute)

Maßnahme	Ziel	Ort/Rahmen	Status
Maßnahme A Treffen mit PA	Ergebnisse aus Datenerhebung, Analysen; gegenseitiger Wissenstransfer	A1: FE1 Workshop Datenanalyse	A1: Erste Erkenntnisse: 22.11.2021 A2: 29.04.2021 A3: 22.01.2021, 29.04.2021, 17.06.2021, 19.08.2021, 22.11.2021 A4: Erste Erkenntnisse: 17.06.2021, A4: Detailgespräch: 19.08.2021, 22.11.2021, 27.04.2022, 22.09.2022 A4: Erläuterung aller Findings: 01.02.2023
		A2: FE2 Workshop IoT Cloud API	
		A3: FE1 Meilenstein-Treffen	
		A4: FE2 Workshop Semantik (Domäne Licht)	
Maßnahme B Fachtagungen	Austausch mit Wissenschaft und Anwendern	B1: Swiss Lighting Forum	Pandemiebedingt ausgefallen
		B2: Licht 2021 in Bamberg	Pandemiebedingt ausgefallen
		B3: Lux Junior Ilmenau	Pandemiebedingt ausgefallen
		B4: Smart Energy and Systems Conference	Pandemiebedingt ausgefallen
		B5: Light & Building	B5: Das Projekt wurde auf der Light & Building 2022 durch FE1 vorgestellt
		B6: Tagung LED Professionals Bregenz	Pandemiebedingt ausgefallen
Maßnahme C Akadem. Lehre	Vermittlung der Ergebnisse	C1: FE 1 Vorlesung Lichttechnik 2	C1, C2: Vorlesung 14./21.07.2021 Smart Lighting VL I /II, Praktikum zu Lichtsensorik konzipiert und wegen der COVID-19 Pandemie digital durchgeführt (21.06. – 12.07.2021) C3: Vorlesung 17.01.2023, Vorstellung der Projektansätze und -
		C2: FE 1 Praktikum Lichttechnik 2	
		C3: FE 2 Vorlesung Verteilte Energieinformationssysteme und Anwendungen	

		<p>C4: FE 2 Vorlesung Einführung in maschinelles Lernen und Künstliche Intelligenz</p>	<p>ergebnisse</p> <p>C4: Überführt in Vorlesung KI2 im Sommersemester 2022 (2023 wg. Forschungsfreiemester entfallen)</p>
		<p>C5: FE 2 Praktikum Digitale Technologien Fachpraktikum 1 + 2</p>	<p>C5: Praktikum Sommersemester 2022 und 2023 „Aufbau einer Datenübertragung für verteilte Systeme“</p>
		<p>C6: Studentische Abschlussarbeiten (Bachelor-/Masterarbeit)</p>	<p>C6: Während der Projektlaufzeit wurden Themengebiete des Projektes in zwei Abschlussarbeiten bei FE1 und in zwei Abschlussarbeiten bei FE2 eingebracht</p>
Maßnahme D Gremienarbeit	Vermittlung für Normung und Anwendung	<p>D1: FE 1 Beratende Tätigkeit z.B. für die Gremien des ZVEI Fachverbands Licht</p>	<p>D1, D2: Erkenntnisse werden fortlaufend in die Gremienarbeit eingebracht: Projekttreffen IoL und Projekttreffen des Arbeitskreises semantische Interoperabilität</p> <p>Vermittlung des Projekts in der AAT LTT Sitzung vom ZVEI-Fachverband Licht am 03.05.2022</p> <p>Highlights aus dem Projekt IoL wurden in der HAT LTT Sitzung vom ZVEI-Fachverband Licht am 24.01.2023 vermittelt</p>
		<p>D2: FE2 Beratende Tätigkeit für den ZVEI: Arbeitskreis Semantische Interoperabilität</p>	
Maßnahme E Öffentlichkeit	Bekanntmachung Vorhaben und Ziele	<p>E1: Webpräsenz, -news</p>	<p>E1: Präsentation des Projekts auf Webseiten der FE1 und FE2 und innerhalb des FV Licht des ZVEI</p> <p>E2: Das Projekt wurde auf der Light & Building 2022 durch FE1 vorgestellt</p>
		<p>E2: Projektvorstellung auf Light & Building 2020</p>	

Maßnahme F Anwender	Erfahrungsaustausch; Hilfe bei Problemlösung	F1: Beratung zu explorativer Datenanalyse	F1, F2: Fortlaufend in den Projekttreffen: 22.01.2021, 29.04.2021, 17.06.2021, 19.08.2021, 22.11.2021, 27.04.2022, 22.09.2022, 01.02.2023 F3, F4: Fortlaufende bilaterale Gespräche zwischen den beiden FEs und einzelner Mitglieder des PA
		F2: Beratung zu Implementation neuer Use Cases	
		F3: Beratung im Bereich Cloudanbindung	
		F4: Beratung zum Thema Semantik	
Maßnahme G	Austausch mit Forschungsprojekten	G1: Projekttreffen u.a. SENSE, SUITE und ForeSight zum Austausch der Entwicklungsergebnisse z.B. zu WoT, Onotolgien	G1: Fortlaufende Synchronisation der Arbeiten in den Projekt-Entwicklergruppen in FE2

5.3. Geplante Transfermaßnahmen (auch nach Projektende)

Maßnahme	Ziel	Ort/Rahmen	Status
Maßnahme H Transfer in die Wirtschaft	Nutzbarmachung der wesentlichen Ergebnisse	H1: Beiträge in Fachzeitschriften (Licht, Lighting Research & Technology)	H1: Artikel in der Fachzeitschrift Licht in Planung
		H2: Konferenzen/Workshops (Licht 2025, Light & Building, Smart Energy and Systems Conference, IEEE Cando EPE)	H2: Durch die Relevanz des Projektes für KMU wird die Thematik auf den Konferenzen und Workshops eingereicht und fokussiert
		H3: Individuelle Beratung Industrie/öffentl. Einrichtungen	H3: Die Inhalte des Projekts werden bei Bedarf der Industrie und öffentlichen Einrichtungen auch nach Projektende in
		H4: Input in ForeSight Plattform und SENSE Labor	individuellen Beratungsgesprächen näher erläutert
			H4: Kontinuierlich über Maßnahme G1 erfolgt

Maßnahme I Folgeprojekte	Aufbau/Weiterführung der Forschungsaktivitäten	I1: Analyse anlagenspezifischer Use Cases	I1: Hilfe bei der Anlagenspezifischen Umsetzung der Datenanalyse-Toolbox wird angeboten
		I2: Kooperationsprojekte mit Industrie	I2: Gespräche über ein Folgeprojekt „IoL2“ in den Gremiensitzungen besprochen und die Gespräche werden weitergeführt
		I3: Kooperationsprojekt im RuhrValley Verbund	I3: Eine Weiterführung der Ergebnisse für die Konnektivität wird im RuhrValley Verbund besprochen
Maßnahme J Akadem. Lehre	Vermittlung der Ergebnisse	J1: FE 1 Vorlesung (VL) Lichttechnik 2	J1: In Vorlesung Smart Lighting VL I /II bei FE1 inkludiert und wird weiterhin ein Bestandteil bleiben
		J2: FE 1 Praktikum Lichttechnik 2	J2: Praktikum zu Lichtsensorik konzipiert und wird in folgenden Praktika weitergeführt
		J3: FE 2 Vorlesung Verteilte Energieinformationssysteme und Anwendungen	J3: In Vorlesung inkludiert und wird weiterhin ein Bestandteil bleiben
		J4: FE 2 Vorlesung Einführung in maschinelles Lernen und Künstliche Intelligenz	J4: Wird im Rahmen der Vorlesung KI2 weiterhin präsentiert.
		J5: FE 2 Praktikum Digitale Technologien Fachpraktikum 1 + 2	J5: Themenkomplex Datenübertragung verteilter Systeme bleibt weiter Bestandteil der FP 1+2
		J6: Dissertationen	J6: Themen und Ergebnisse des Projekts fließen in Dissertationen bei FE1 mit ein. Teile der Thematik, wie Farbsensoren, intelligente Lichtsteuerung oder die LED-Alterung wird separat vertieft

<p>Maßnahme K Standardisierung</p>	<p>Standardisieren der Kommunikation</p>	<p>K1: Vorträge/Infoveranstaltungen bei Normungsgremien, zuarbeiten Verbands-gremien etc.</p>	<p>K1: Die Entwicklungsergebnisse z.B. zu WoT, Ontologien werden in Projekttreffen von SENSE, SUITE , ForeSight, Applied Excellence Department, VDI Arbeitsgruppen sowie künftigen Forschungsvorhaben mit eingebracht</p>
--	--	---	---

6. Literatur

[1] „KNX IoT 1.0“. KNX Association. [Online]. Verfügbar unter: https://www.knx.org/wAssets/docs/downloads/Marketing/Flyers/KNX-IoT/KNX-IoT_en.pdf

[2] ETSI.org, „SAREF extension for Building“. <https://saref.etsi.org/saref4bldg/v1.1.2/>

[3] „SAREF4ENER“. <https://saref.etsi.org/saref4ener/v1.1.2/>

[4] „JSON for Linking Data“. <https://json-ld.org>

[5] A. Hüneburg, H. Hoffmann, und I. Kunold, „ZVEI Studie Semantische Interoperabilität“. 2017.

[6] I. Kunold, „Projekt SENSE, IKT FH Dortmund“. <https://www.fh-dortmund.de/projekte/sense.php>

[7] BMWK, „SUITE -Intelligente, adaptive Assistenzservices auf Basis von Smart Meter und Gebäudeautomationsdaten“, *digitale-technologien.de*. https://www.digitale-technologien.de/DT/Redaktion/DE/Standardartikel/Smarte-Datenwirtschaft-Projekte/SDW_suite.html

[8] BMWK, „ForeSight - Plattform für kontextsensitive, intelligente und vorausschauende Smart Living Services“, *digitale-technologien.de*. https://www.digitale-technologien.de/DT/Redaktion/DE/Standardartikel/KuenstlicheIntelligenzProjekte/KuenstlicheIntelligenz_ErsterFoerderaufuf/ki-projekt_foresight.html

[9] Web of Things Interest Group, „Web of Things Thing Description“. W3C. [Online]. Verfügbar unter: <https://www.w3.org/TR/wot-thing-description/>

[10] U. Großmann, I. Kunold, und C. Engels, Hrsg., *Smart Energy 2016: Digitalisierung der Energieversorgung - Treiber und Getriebene*. in E-Business. Glückstadt: vwh, Verlag Werner Hülsbusch, Fachverlag für Medientechnik und -wirtschaft, 2016.

[11] „Schema.org Data Model“. <https://schema.org/docs/datamodel.html>

- [12] „Iotschema“. <https://schema.org/docs/iot-gettingstarted.html>
- [13] H. Rijgersberg, M. Van Assem, und J. Top, „Ontology of units of measure and related concepts“, *Semantic Web*, Bd. 4, Nr. 1, S. 3–13, 2013, doi: 10.3233/SW-2012-0069.
- [14] ETSI, „SAREF: the Smart Applications REFerence ontology“. <https://saref.etsi.org/core/v3.1.1/>
- [15] „Energy@home association“. <http://www.energy-home.it>
- [16] „EEBus Initiative e.V.“ <https://www.eebus.org>
- [17] M. Kuller, F. Kohlmorgen, N. Karaoglan, M. Niemeyer, I. Kunold, und H. Wohrle, „Conceptual design of a digital twin based on semantic web technologies in the smart home context“, in *2020 IEEE 3rd International Conference and Workshop in Óbuda on Electrical and Power Engineering (CANDO-EPE)*, Budapest, Hungary: IEEE, Nov. 2020, S. 000167–000172. doi: 10.1109/CANDO-EPE51100.2020.9337749.
- [18] S. Sir, F. Kohlmorgen, M. Biskup, M. Kuller, I. Kunold, und H. Wohrle, „Discovery and synchronization of thing descriptions in a distributed architecture“, in *2022 IEEE 5th International Conference and Workshop Óbuda on Electrical and Power Engineering (CANDO-EPE)*, Budapest, Hungary: IEEE, Nov. 2022, S. 000141–000146. doi: 10.1109/CANDO-EPE57516.2022.10046394.
- [19] M. H. Rasmussen, M. Lefrançois, G. F. Schneider, und P. Pauwels, „BOT: The building topology ontology of the W3C linked building data group“, *Semantic Web*, Bd. 12, Nr. 1, S. 143–161, Nov. 2020, doi: 10.3233/SW-200385.
- [20] „Openweathermap: Weather forecasts, nowcasts and history in a fast and elegant way“. <https://openweathermap.org/>
- [21] „Sunset and sunrise times“. <https://sunrise-sunset.org>
- [22] R. Krishnamurthi, A. Kumar, D. Gopinathan, A. Nayyar, und B. Qureshi, „An Overview of IoT Sensor Data Processing, Fusion, and Analysis Techniques“, *Sensors*, Bd. 20, Nr. 21, S. 6076, Okt. 2020, doi: 10/gjhjk5.
- [23] T. J. Saleem und M. A. Chishti, „Data Analytics in the Internet of Things: A Survey“, *Scalable Comput. Pract. Exp.*, Bd. 20, Nr. 4, S. 607–630, Dez. 2019, doi: 10/gsjn4q.
- [24] C. Voyant *u. a.*, „Machine learning methods for solar radiation forecasting: A review“, *Renew. Energy*, Bd. 105, S. 569–582, Mai 2017, doi: 10/gdq9px.
- [25] G. R. Newsham und B. J. Birt, „Building-level occupancy data to improve ARIMA-based electricity use forecasts“, in *Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building*, Zurich Switzerland: ACM, Nov. 2010, S. 13–18. doi: 10/bzc4d3.
- [26] H. Zheng und Y. Wu, „A XGBoost Model with Weather Similarity Analysis and Feature Engineering for Short-Term Wind Power Forecasting“, *Appl. Sci.*, Bd. 9, Nr. 15, S. 3019, Juli 2019, doi: 10/gsjn4z.

- [27] C. Li u. a., „Power Load Forecasting Based on the Combined Model of LSTM and XGBoost“, in *Proceedings of the 2019 the International Conference on Pattern Recognition and Artificial Intelligence*, Wenzhou China: ACM, Aug. 2019, S. 46–51. doi: 10/gsjn42.
- [28] T. C. Nokeri, „Forecasting Using ARIMA, SARIMA, and the Additive Model“, in *Implementing Machine Learning for Finance*, Berkeley, CA: Apress, 2021, S. 21–50. doi: 10.1007/978-1-4842-7110-0_2.
- [29] „pmdarima: ARIMA estimators for Python“. <http://alkaline-ml.com/pmdarima/>
- [30] M. Fahim und A. Sillitti, „Anomaly Detection, Analysis and Prediction Techniques in IoT Environment: A Systematic Literature Review“, *IEEE Access*, Bd. 7, S. 81664–81681, 2019, doi: 10/gp668h.
- [31] W. Truong, B. Zandi, Vt. Drl-ng, und T. K. Dr-Ing habil, „Circadian metric – Computation of circadian stimulus using illuminance, correlated colour temperature and colour rendering index“, *Build. Environ.*, Bd. 184, Nr. August, S. 107146, 2020, doi: 10.1016/j.buildenv.2020.107146.
- [32] M. S. Rea, R. Nagare, und M. G. Figueiro, „Modeling Circadian Phototransduction: Quantitative Predictions of Psychophysical Data“, *Front. Neurosci.*, Bd. 15, S. 615322, Feb. 2021, doi: 10/gsjn45.

7. Anhang

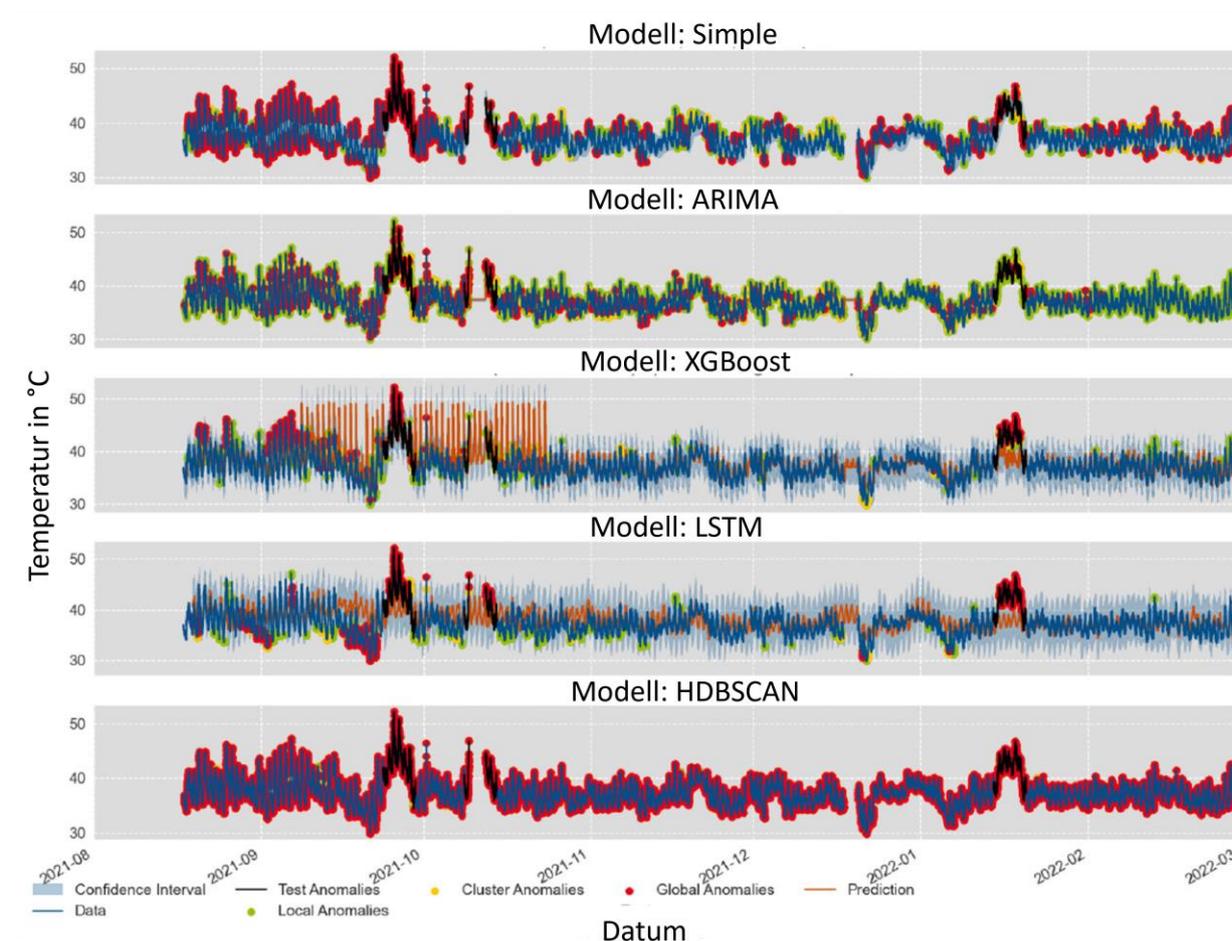


Abbildung 34: Beispiel der Anomaliedetektion anhand eines Temperatursensors. Als Anomalien dienen drei zufällig ausgewählte Stellen, die sich über einen Zeitraum von sechs Tagen strecken und deren gemessene Werte um 15% über das Tukey Fenster erhöht wurden. Zur Optimierung wird die Precision-Recall Kurve verwendet.

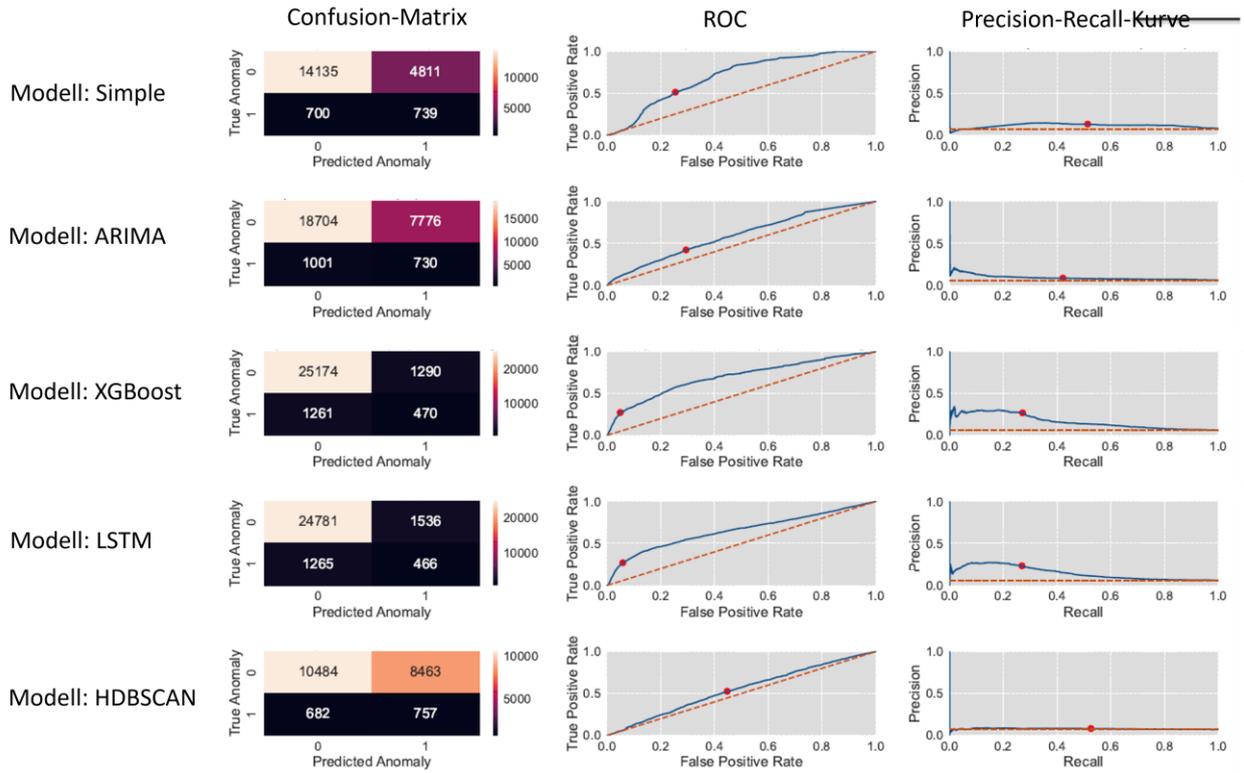


Abbildung 35: Metriken für die Beurteilung der Anomaliedetektion aus Beispiel in Abbildung 34. Zur Optimierung wird die **Precision-Recall Kurve** verwendet.

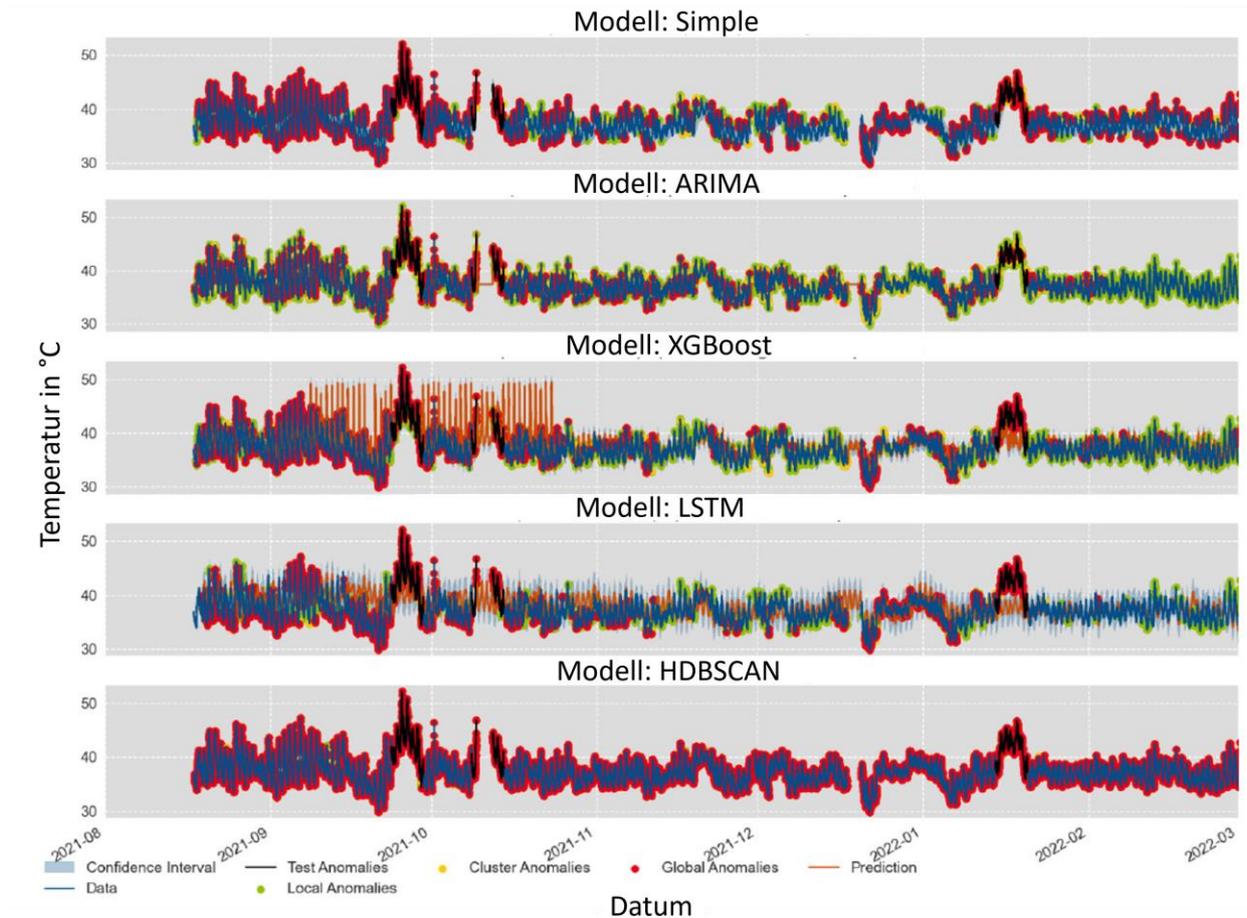


Abbildung 36: Beispiel der Anomaliedetektion anhand eines Temperatursensors. Als Anomalien dienen drei zufällig ausgewählte Stellen, die sich über einen Zeitraum von sechs Tagen strecken und deren gemessene Werte um 15% über das Tukey Fenster erhöht wurden. Zur Optimierung wird die ROC verwendet.

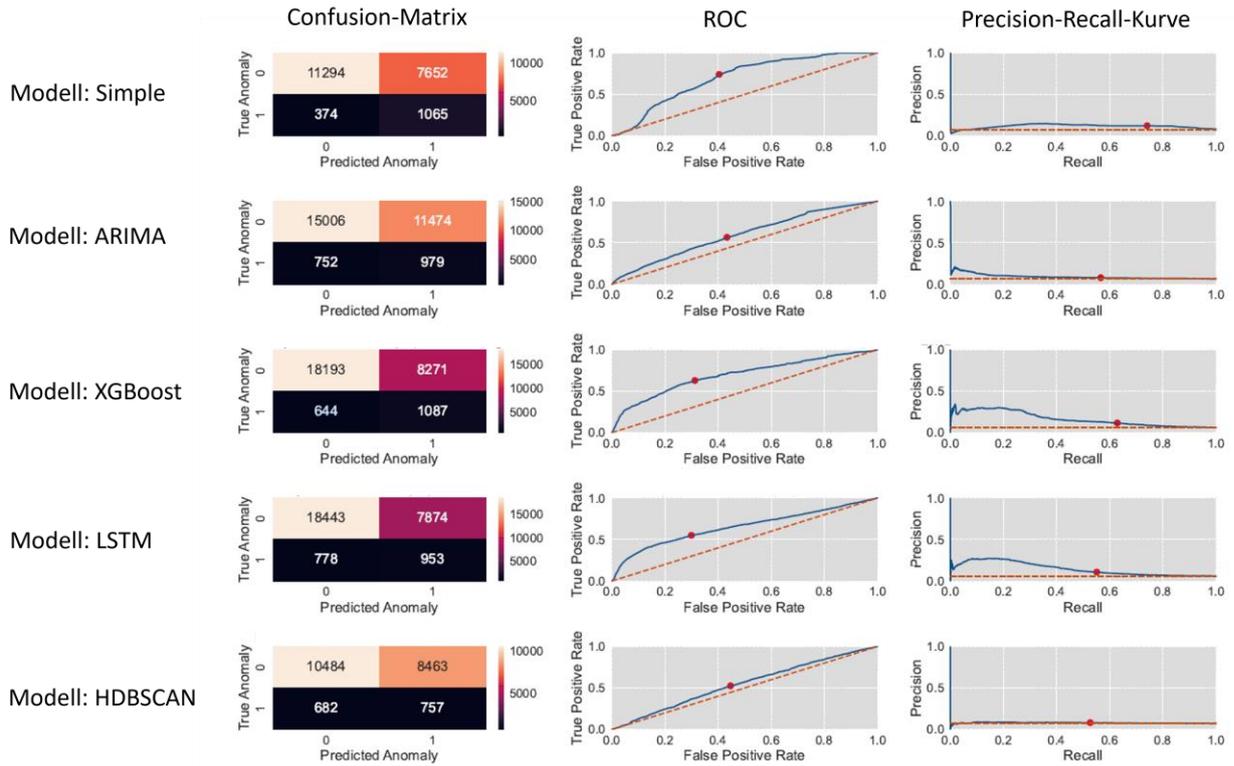


Abbildung 37: Metriken für die Beurteilung der Anomaliedetektion aus Beispiel Abbildung 36. Zur Optimierung wird die **ROC** verwendet.