

Daniel Schulz, David Knodt, Sebastian Konietzny  
Michael Mock, Robert Spindler

**MONALisa – Automatisiertes  
MONitoring, Alarming und  
VisuaLisieren von Sensordaten  
der technischen Gebäudeausrüstung  
zur Erschließung niedrig investiver  
Energieeinsparpotenziale**

Bei dieser Veröffentlichung handelt es sich um die Kopie des Abschlussberichtes einer vom Bundesministerium für Verkehr, Bau und Stadtentwicklung -BMVBS- im Rahmen der Forschungsinitiative »Zukunft Bau« geförderten Forschungsarbeit. Die in dieser Forschungsarbeit enthaltenen Darstellungen und Empfehlungen geben die fachlichen Auffassungen der Verfasser wieder. Diese werden hier unverändert wiedergegeben, sie geben nicht unbedingt die Meinung des Zuwendungsgebers oder des Herausgebers wieder.

Dieser Forschungsbericht wurde mit modernsten Hochleistungskopierern auf Einzelanfrage hergestellt.

Die Originalmanuskripte wurden reprotechnisch, jedoch nicht inhaltlich überarbeitet. Die Druckqualität hängt von der reprotechnischen Eignung des Originalmanuskriptes ab, das uns vom Autor bzw. von der Forschungsstelle zur Verfügung gestellt wurde.

© by Fraunhofer IRB Verlag

2017

ISBN 978-3-7388-0031-9

Vervielfältigung, auch auszugsweise,  
nur mit ausdrücklicher Zustimmung des Verlages.

**Fraunhofer IRB Verlag**

Fraunhofer-Informationszentrum Raum und Bau

Postfach 80 04 69  
70504 Stuttgart

Nobelstraße 12  
70569 Stuttgart

Telefon 07 11 9 70 - 25 00  
Telefax 07 11 9 70 - 25 08

E-Mail [irb@irb.fraunhofer.de](mailto:irb@irb.fraunhofer.de)

[www.baufachinformation.de](http://www.baufachinformation.de)

[www.irb.fraunhofer.de/bauforschung](http://www.irb.fraunhofer.de/bauforschung)

# Projekt MONALisa

Endbericht, April 2017

# **MONALisa**

## **Automatisiertes MONitoring, Alarming und VisuaLisieren von Sensordaten der technischen Gebäudeausrüstung zur Erschließung niedrig investiver Energieeinsparpotenziale**

Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS  
in Sankt Augustin

Antragsforschung

Daniel Schulz  
David Knodt  
Dr. Sebastian Konietzny  
Dr. Michael Mock  
Robert Spindler

Der Forschungsbericht wurde mit Mitteln der Forschungsinitiative Zukunft Bau des Bundesinstitutes für Bau-, Stadt- und Raumforschung gefördert.  
(Aktenzeichen: SWD-10.08.18.7-15.12)  
Die Verantwortung für den Inhalt des Berichtes liegt beim Autor.

<b>1</b>	<b>EINFÜHRUNG .....</b>	<b>4</b>
1.1	AUFBAU UND VORGEHENSWEISE IM PROJEKT .....	5
1.2	ENERGIESPARMAßNAHMEN .....	6
1.3	UNTERSUCHUNGSOBJEKTE .....	7
1.3.1	<i>Das Gebäude des Umweltbundesamtes.....</i>	<i>8</i>
<b>2</b>	<b>DATENAUFBEREITUNG .....</b>	<b>14</b>
2.1	DATENEXPORT .....	16
2.2	ARCHITEKTUR .....	16
2.3	DATENIMPORT .....	19
2.4	DATENSEMANTIK .....	20
2.5	DATENVALIDIERUNG .....	21
2.6	DATENVERARBEITUNG UND ANALYSE-TOOLS .....	24
2.6.1	<i>Intern.....</i>	<i>25</i>
2.6.2	<i>Extern .....</i>	<i>28</i>
<b>3</b>	<b>DATENEXPLORATION UND PLAUSIBILITÄT .....</b>	<b>30</b>
3.1	AUTOMATISIERTE ZUORDNUNGEN VON SOLL- UND IST-WERTEN .....	30
3.2	EXPLORATION DES FERNWÄRMEHEIZKREISES .....	31
3.2.1	<i>Leistung des Fernwärmeheizkreises .....</i>	<i>32</i>
3.2.2	<i>Steuerung einer Heizungsanlage.....</i>	<i>34</i>
<b>4</b>	<b>DATENANALYTIK .....</b>	<b>36</b>
4.1	EINFÜHRUNG .....	36
4.2	FRAGESTELLUNGEN .....	37
4.3	DATENANALYSEN .....	38
4.3.1	<i>Visuelle Datenanalyse zur Identifikation von Defekten.....</i>	<i>38</i>
4.3.2	<i>Ventilstellungen bei Heizung und Wärmerückgewinnung .....</i>	<i>38</i>
4.3.3	<i>Differenztemperatur zwischen ein- und ausfließendem Wasser der Fernwärmeübergabe .....</i>	<i>40</i>
4.3.4	<i>Detektion von großen Differenzen zwischen Soll- und Ist-Werten .....</i>	<i>44</i>
4.3.5	<i>Analyse der Ventilstellungen.....</i>	<i>47</i>
4.3.6	<i>Automatisierte Detektion von Schwingungen.....</i>	<i>50</i>
4.3.7	<i>Wetteranalyse.....</i>	<i>54</i>
<b>5</b>	<b>VISUALISIERUNG .....</b>	<b>72</b>
5.1	ÜBERSICHT ALLER STATISCHEN HEIZUNGSANLAGEN .....	72
5.2	INTERAKTIVES DASHBOARD .....	73
<b>6</b>	<b>ZUSAMMENFASSUNG .....</b>	<b>74</b>
6.1	WEITERE ERKENNTNISSE AUS DEM FORSCHUNGSPROJEKT .....	75
<b>7</b>	<b>KOMMUNIKATION .....</b>	<b>78</b>
7.1	MEETINGS / TELEFONKONFERENZEN (AUSWAHL) .....	78
7.2	AUßENDARSTELLUNG .....	78
<b>8</b>	<b>ABBILDUNGSVERZEICHNIS .....</b>	<b>79</b>
<b>9</b>	<b>LITERATURVERZEICHNIS .....</b>	<b>81</b>

Mit dem Begriff „Klimawandel“ werden im Allgemeinen langfristige Klimaschwankungen bezeichnet, wie sie sich schon immer in der Erdgeschichte seit Entstehung der Atmosphäre zugetragen haben. Diese natürlichen Schwankungen werden jedoch seit Beginn der industriellen Revolution durch anthropogen verursachte Veränderungen des Klimas überlagert. Aufgrund dessen wird der Begriff des Klimawandels seit Anfang der 1980er Jahre zunehmend für die vom Menschen verursachte Erwärmung der Atmosphäre verwendet. Damit verbunden sind deren katastrophale Folgen, wie Überschwemmungen, Dürreperioden und Hitzewellen.

Um dem Klimawandel aktiv entgegen zu treten, wurde durch die nationale und internationale Politik die Energiewende eingeleitet. Im Zentrum der Energiewende stehen dabei insbesondere zwei Maßnahmen. Dies sind:

- a) der Nutzungswechsel von fossilen Energieträgern hin zu erneuerbarer Energien, und
- b) ein nachhaltiges Energiemanagement.

Während bei a) vor allem der langfristige Nutzungswechsel von Energien wie z.B. von Erdöl und Erdgas hin zu Wind und Wasser im Zentrum steht, werden als Kernelemente eines nachhaltigen Energiemanagements die kurzfristige Steigerung der Energieeffizienz und das Einsparen unnötiger Verbräuche angesehen [1]. Eine besondere Rolle spielen in diesem Prozess vor allem Gebäude, da diese ca. 30-40 Prozent des gesamten Endenergieverbrauchs einer Volkswirtschaft auf sich vereinen [2]. Maßnahmen, die zur Energieeinsparung und Steigerung der Energieeffizienz von Gebäuden eingesetzt werden, leisten somit einen großen Beitrag bei der Umsetzung der Energiewende.

In diesem Kontext spielen insbesondere die Betriebs- und Energiekosten der technischen Gebäudeausrüstung eine besondere Rolle, da diese einen Großteil der Gesamtkosten innerhalb des Lebenszyklus eines Gebäudes ausmachen [3]. Als maßgeblicher Ansatzpunkt zur Anwendung von Energiesparmaßnahmen dient dabei vor allem die Gebäudeautomation. Hierunter ist die Gesamtheit aller Überwachungs-, Steuerungs-, Regel- und Optimierungseinrichtungen innerhalb eines Gebäudes zu verstehen [4].

Repräsentiert durch ein Netz aus verschiedenartigen Sensoren wird die Gebäudeautomation als Basis eines dynamischen Energie- und Gebäudemanagements eingesetzt. Darauf aufbauend wird eine Steuerung der Gebäudeautomation softwarebasiert durch die Gebäudeleittechnik (GLT) durchgeführt. Diese hat in der Regel die Aufgabe, die Funktionsfähigkeit der Gebäudeautomation zu überwachen und bestimmte Betriebsparameter optimal zu konfigurieren. Hierzu zählen z.B. alle Konditionen, die das Raumklima in einem Gebäude bestimmen. In Gänze ist dies keine triviale Aufgabe, sondern stellt hohe Anforderungen sowohl an die Software als auch an das Fachpersonal.

Aktuelle Ansätze im Rahmen der Anwendung der GLT schöpfen aufgrund der Komplexität des Gesamtsystems das Potential einer Optimierung der Gebäudeautomation jedoch nicht vollständig aus. Analytische Konzepte und Datenanalysen basierend auf historischen Daten sind häufig in den verwendeten Softwarekomponenten der GLT unterrepräsentiert. Zudem werden die Systeme häufig von Fachpersonal bedient, dem der technische Hintergrund fehlt, alle Prozesse vollständig zu verstehen und in ihrer Komplexität zu überblicken. Beides führt dazu, dass das Potential an Energiesparmaßnahmen nur unzureichend ausgeschöpft wird.

Das im Folgenden beschriebene Forschungsprojekt leistet einen wesentlichen Beitrag zur Optimierung des Energie- und Gebäudemanagements mit dem Ziel, die Energie- und Betriebskosten eines Gebäudes nachhaltig zu senken. Dies geschieht durch die datengetriebene Erschließung von Energiesparmaßnahmen. Für die Erschließung wird prototypisch ein Prozess aufgezeigt, der auf der Basis innovativer Algorithmen der Statistik, des Data Mining und der Zeitreihenanalyse historische Sensordaten der GLT analytisch auswertet. Allgemeines Ziel ist es den Prozess in seiner Breite von den Rohdaten bis hin zu den Analyseergebnissen aufzuzeigen. Eine Quantifizierung der Energieeinsparpotentiale findet dabei nicht statt. Im Zentrum des Prozesses stehen dabei Anwendungsbeispiele. Als Datenbasis dienen Sensordaten eines Bundesgebäudes.

## 1.1 Aufbau und Vorgehensweise im Projekt

Für die Durchführung des Projektes wurde der Projektverlauf in mehrere Phasen unterteilt. Diese spiegeln sich in den einzelnen Abschnitten dieses Berichtes wieder, welche zur Orientierung hier kurz beschrieben werden. Zu Beginn des Projektes galt es zunächst das Domänenwissen über die GLT und die Untersuchungsobjekte aufzubauen (siehe Kapitel 1). Hierzu zählt z.B. das Wissen über Prozesse innerhalb der GLT, Funktionsweisen einzelner Sensoren aber auch historische Informationen zum Untersuchungsobjekt. Letzte sind z.B. unerlässlich um Analyseergebnisse zielführend interpretieren zu können. Den Kern des Projektes stellt die eigentliche Datenanalyse dar. Allen Prozessen, die hiermit in Verbindung stehen, sind einzelne Kapitel gewidmet. Diese sind in Abbildung 1 dargestellt. Dies beginnt mit der Aufbereitung und Exploration der Daten, gefolgt von der Analyse der Daten auf der Basis unterschiedlicher Analysemethoden und schließt mit der Visualisierung der Ergebnisse der Analysen ab. Im anschließenden Kapitel 6 erfolgt die Zusammenfassung der Ergebnisse.



Abbildung 1: Aufbau des Berichtes

Dieses Vorgehen orientiert sich am Cross - Industry Standard Process für Data Mining (CRISP-DM, [5]). Dieser hilft bei der Planung und Verwaltung von Analyseprojekten und wurde 1996 im Rahmen eines von der EU geförderten Projektes entwickelt (vgl. Abbildung 2). Seit seiner Entwicklung wird dieser Prozess in Analyseprojekten angewendet. An dieser Stelle sei auf eine Eigenschaft bzw. eine Erfahrung aus der Anwendung des Prozesses hingewiesen:

1) Wie in Abbildung 2 zu sehen, ist dieser Prozess nicht unidirektional. Er sieht Rückschritte zu vorherigen Teilprozessen vor. Dieses ist explizit gewünscht, um den gesamten Prozess und seinen Fortschritt immer wieder zu hinterfragen. Im Kontext dieses Projektes hat sich insbesondere die Rückbesinnung zum Verständnis der Daten und des Business als enorm wichtig herausgestellt. Da das Wissen über die Daten und die Prozesse ausschließlich extern vorlag, war dies eine zeitliche Herausforderung.

2) Viele Erfahrungen in der Anwendung des Prozesses haben im Lauf der Zeit gezeigt, dass die Datenaufbereitung ein nicht zu unterschätzender Teil hinsichtlich der Zeitplanung ist. Es gibt Erfahrungen die einen Großteil des zeitlichen Aufwandes in der Datenaufbereitung sehen. Dies ist vor allem darin begründet, dass die Daten nicht für die Analysen erfasst und entsprechend an die Algorithmen angepasst werden müssen. Auch dies kann für dieses Projekt nachvollzogen werden, was in den folgenden Kapiteln beschrieben wird.

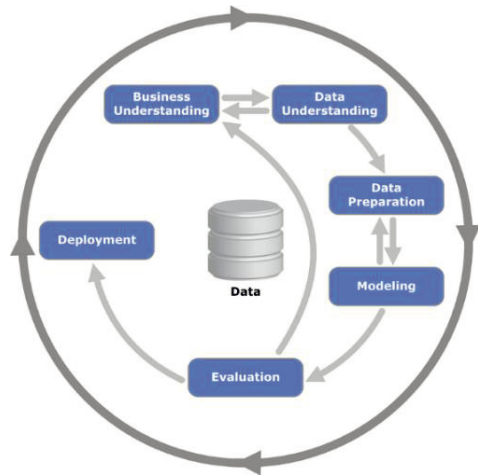


Abbildung 2: CRISP-DM Modell  
(Quelle: Wikipedia [5])

## 1.2 Energiesparmaßnahmen

Im Allgemeinen existieren vielfältige Möglichkeiten die für den Betrieb eines Gebäudes notwendigen Energiemengen zu optimieren und ggf. zu reduzieren. Die folgende Abbildung gibt einen Überblick über mögliche Energiesparmaßnahmen und bewertet diese hinsichtlich Ihres Einsparpotentials und der dafür notwendigen Aufwände in Form einer relativen Kostenübersicht:

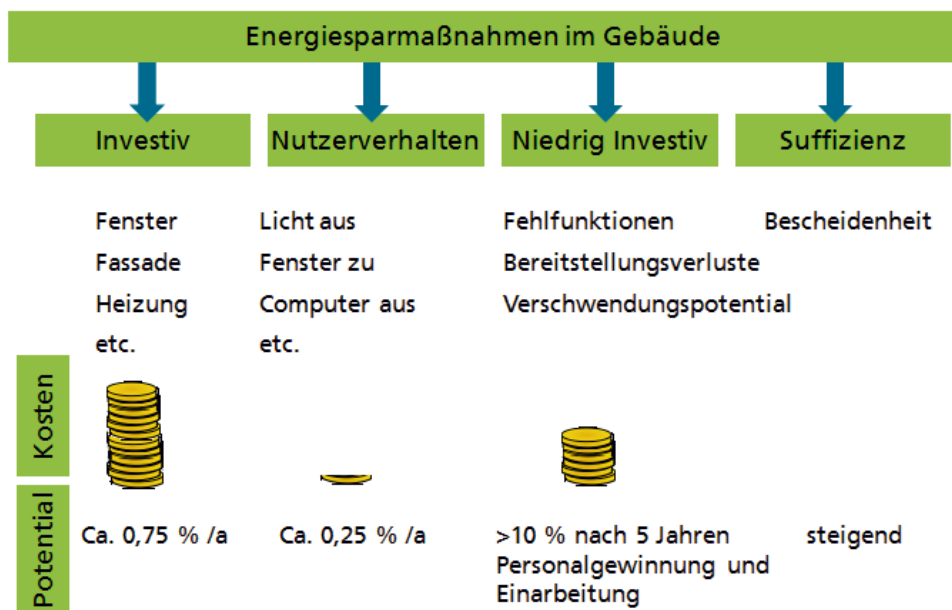


Abbildung 3: Übersicht Energiesparmaßnahmen für Gebäude

Die in diesem Forschungsprojekt eingesetzten Methoden zur Erschließung von Energiesparpotentialen sind der Kategorie der niedrig investiven Maßnahmen zuzuordnen. Hierzu zählen allen Maßnahmen, die mit verhältnismäßig geringem finanziellen Aufwand kurzfristig Einsparpotentiale im regulären Gebäudebetrieb aufdecken und erschließen können. Auf lange Sicht können mit diesen Maßnahmen >10% der Energie eingespart werden. Die Maßnahmen lassen sich wiederum in drei Kategorien einteilen:



- In die Kategorie der Fehlfunktionen fallen alle Maßnahmen, die Elemente innerhalb der GLT aufdecken, die abseits des Normbetriebs betrieben werden. Hierzu zählen z.B. Pumpen, die im laufenden Betrieb ausfallen oder Lüfter, welche nicht mehr genügend Leistung aufbringen.
- In die Kategorie der Bereitstellungsverluste fallen alle Maßnahmen, die Elemente innerhalb der GLT aufdecken, bei denen dauerhaft mehr Energie zur Verfügung gestellt, als an der jeweiligen Position zu einem Zeitpunkt benötigt wird. Hierzu zählen z.B. Klappen und Ventile, die aufgrund Ihrer Stellung mehr Luft oder Wasser durchlassen als angefordert wird.
- In die Kategorie der Verschwendungspotentiale fallen alle Maßnahmen, die Elemente innerhalb der GLT aufdecken, bei denen dauerhaft mehr Energie verbraucht als benötigt wird. Hierzu zählen z.B. zu hoch eingestellte Sollwerte für die Vorlauftemperatur der Heizungsanlage.

Abseits der Erschließung niedrig investiver Energiesparpotentiale, existieren noch weitere Maßnahmen, um Energiesparpotentiale zu erschließen. Diese seien der Vollständigkeit halber hier kurz beschrieben.

Allen anderen Maßnahmen voran stehen investive Maßnahmen. Bei diesen werden die Energiesparpotentiale hauptsächlich auf der Basis von kostspieligen Maßnahmen in Form des Einsatzes von neuen Materialien und Techniken ausgeschöpft. Hierzu zählen z.B. ein optimierter Wärmeschutz durch den Einsatz einer neuen Dämmung und der Einbau effizienterer Heizungssysteme. Ihr Einsparpotential ist in diesen Fällen eher langfristig zu betrachten. Ergänzend dazu existieren Maßnahmen, welche die Nutzer des jeweiligen Gebäudes betreffen. Hierzu zählen insbesondere Maßnahmen, die das Nutzerverhalten hin zu einem energetisch sorgsamem Umgang prägen. In diesem Kontext sind ebenfalls Maßnahmen zu nennen, welche sich mit der menschlichen Lebenseinstellung im Umgang mit Energie beschäftigen (Suffizienz).

### 1.3 Untersuchungsobjekte

Im unmittelbaren Vorfeld des Forschungsprojektes wurden als Untersuchungsobjekte zwei Gebäude des Bundes in Zusammenarbeit mit dem BBSR bestimmt (Abbildung 4: Untersuchungsobjekte). Dies sind der Gebäudekomplex des Umweltbundesamtes (UBA) mit Hauptsitz in Dessau-Roßlau und der Gebäudekomplex des Bundesministeriums für Gesundheit (BMG) mit Hauptsitz in Bonn. Beide Gebäude sind mit ihren Baujahren 2005 (UBA) und 2007 (BMG), gemessen an anderen verfügbaren Datenquellen, relativ jung und verfügen deshalb über moderne Technik und Ausstattung für den Betrieb einer GLT. Gleichzeitig konnte über ihre Betriebsjahre eine historische Datenbank aller Sensordaten der Gebäudeleittechnik aufgebaut werden. Der Inhalt dieser Datenbanken bildet den zentralen Teil des Projektes und es gilt diesen verfügbar zu machen und anschließend nach zu bestimmenden analytischen Kriterien auszuwerten.

Im Rahmen des Projektverlaufs wurde entschieden, sich auf die Daten des Umweltbundesamtes zu konzentrieren (vgl. Kapitel 2.3 Datenimport). Für das Verständnis der in den nächsten Kapiteln folgenden Auswertungen und Analysen ist ein breites Hintergrundwissen zum Gebäude des UBA notwendig. Aufgrund dessen erfolgen im folgenden Kapitel eine kurze Beschreibung der Historie des UBA und die Erläuterung einer Auswahl an Anlagen, welche maßgeblich über die GLT gesteuert werden. Dabei wird sich auf die wesentlichen Punkte, die zum weiteren Verständnis der folgenden Kapitel notwendig sind, konzentriert.

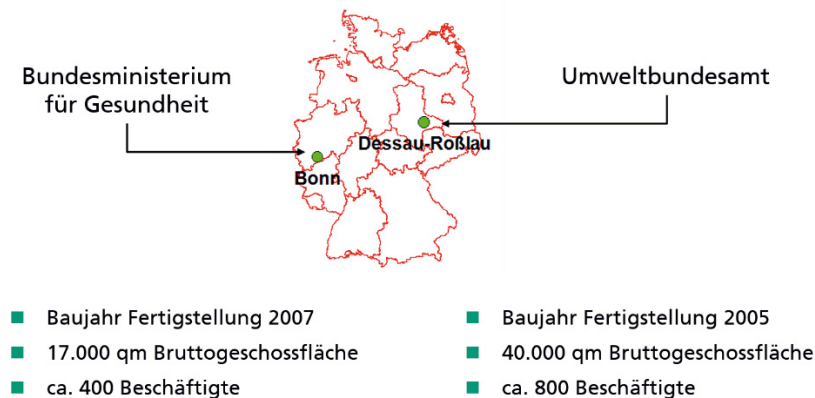


Abbildung 4: Untersuchungsobjekte

### 1.3.1 Das Gebäude des Umweltbundesamtes

Im Jahre 2005 wurde der Hauptsitz des Umweltbundesamtes von Berlin in das heutige Dessau-Roßlau (damals nur Dessau) in Sachsen-Anhalt verlegt. Vor Ort entstand hierfür ein neuer Gebäudekomplex. Dessen Grundsteinlegung fand 2002 statt. Offiziell eröffnet und in Betrieb genommen wurde das Gebäude am 11. Mai 2005. Seit dieser Zeit findet die Aufzeichnung der GLT-Daten statt. Als Standort für den Gebäudekomplex wurde eine ehemalige Industriefläche im „Gasviertel“ gewählt. Dieses befindet sich in unmittelbarer Nähe zum Hauptbahnhof und zur Innenstadt Dessaus. Hierdurch wurde dem Wunsch eines möglichst bürgernahen Konzeptes Rechnung getragen.

Der Gebäudekomplex des UBA setzt sich in Gänze aus mehreren Gebäuden zusammen. Dies sind das Hauptgebäude mit Hörsaal, angegliederter Bibliothek, die Kantine und der „Wörlitzer Bahnhof“. Letzter ist das denkmalgeschützte Gebäude eines ehemaligen Bahnhofs, welches heute als Bürogebäude vom UBA genutzt wird. Ergänzend befinden sich ein Parkhaus und ein Abstellplatz für Zweiräder auf dem Gelände. Die Anordnung der Gebäude vor Ort ist in nebenstehender Abbildung zu sehen.

Abbildung 5: Lageplan Gebäude des Umweltbundesamtes (eigene Aufnahme)



Das Hauptgebäude besitzt die markante Form einer Schleife. Diese verfügt im südwestlichen Teil über ein Forum, welches den Eingangsbereich darstellt. Daran nördlich anschließend bildet die Schleife einen Innenteil. Dieser wurde ursprünglich als nach oben offener Innenhof geplant. Man entschied sich jedoch dazu, diesen Innenteil nachträglich mit einem Glasdach zu versehen, so dass dieser den Charakter eines Atriums bzw. Wintergartens mit mediterranem Mikroklima erlangte. Das Glasdach wurde dabei als Sheddach konzipiert (Abbildung 6: Sheddach des UBA von innen (links) und außen (rechts)). In diesem sind teilweise Solarzellen, ein innen liegender Sonnenschutz und Kippfenster integriert. Letzte dienen ganzjährig sowohl der Lüftung als auch der Regulierung der Luftfeuchte und Temperatur. Die Steuerung ist automatisiert. Die Regelung ist jeweils in die GLT integriert.



Abbildung 6: Sheddach des UBA von innen (links) und außen (rechts)

Innerhalb der Schleife liegen die Büros der ca. 800 Mitarbeiter und Mitarbeiterinnen. Diese sind hauptsächlich als Einzelbüros konzipiert und verteilen sich mit ihrer jeweiligen Größe von ca. 11qm über 4 Stockwerke. Dabei existieren jeweils ein innerer und ein äußerer Teil der Schleife. Während im inneren Teil der Schleife die Bürofenster jeweils zum Atrium hinaus zeigen, besitzen die Büros des äußeren Teils eine Fensterfront als Teil der äußeren Fassade. Der innere und der äußere Teil der Schleife sind durch einen innenliegenden Flur voneinander getrennt (Abbildung 7: Innenflur der Schleife (links), Atrium (Mitte) und Beispiel-Büro des UBA). Die Fassade selbst setzt sich hauptsächlich aus den Materialien Holz, Stahl und Glas zusammen. Die Decken des Hauptgebäudes sind aus Beton gefertigt.

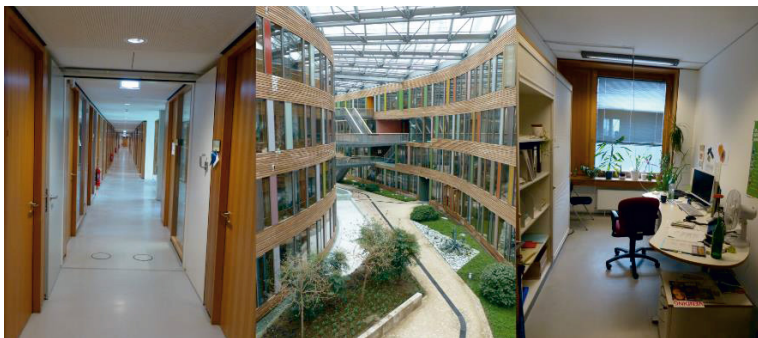


Abbildung 7: Innenflur der Schleife (links), Atrium (Mitte) und Beispiel-Büro des UBA (eigene Aufnahmen)

Auf dem Dach des UBA befinden sich verschiedene Sensoren zur Erfassung und Bewertung der lokalen Wetterbedingungen. Hierzu zählt z.B. auch die Außentemperatur, welche eine zentrale Rolle bei der Steuerung der Heizungs- und Lüftungsanlagen spielt. Historisch betrachtet wurden bei der Fertigstellung des UBA zunächst zwei Sensoren zur Messung der Außentemperatur am Gebäudekomplex angebracht. Dies waren zu einem ein Sensor an der Westseite der Fassade des Gebäudes und zum anderen ein Sensor an der Wetterstation. Letzte wurde entlang der Achse 3 auf dem Dach im südöstlichen Teil des Gebäudekomplexes angebracht (siehe Abbildung 8: Achsen des UBA). Dabei dienten die Messwerte des Temperatursensors an der Fassade zunächst als Richtwert für die Soll-Temperaturen der Heizungsanlage.



Abbildung 8: Achsen des UBA

Bereits kurz nach Inbetriebnahme des UBA stellte sich jedoch heraus, dass die Exponiertheit des Fassadensensors die Temperaturmessung zur Steuerung der Heizungsanlage negativ beeinflusst. Dies ist in der westlichen Exposition des Sensors begründet, was vor allem durch die erhöhte globale Sonnenstrahlung an manchen Nachmittagen zu einer Abschaltung der Heizungsanlage führte. Aufgrund dessen wurde im Frühjahr 2005 auf 2006 zur Steuerung der Heizungsanlage auf die Temperatursensorik der Wetterstation entlang der Achse 3 umgestellt.

Im Jahre 2008 wurde eine fehlerhafte Steuerung der Kippfenster im Sheddach diagnostiziert. Trotz Bedarf blieben die Kippfenster im Sommer geschlossen, was zu einer Überhitzung des Atriums und somit zu Unbehagen unter der Belegschaft des UBA führte. Als Ursache wurde ein fehlerhafter Regensensor der Wetterstation entlang der Achse 3 identifiziert. Zeitgleich mit dem Tausch des Sensors wurde als Äquivalent eine zweite Wetterstation in Betrieb genommen. Diese befindet sich auf dem Dach entlang der Achse 97 im nordöstlichen Teil des UBA. Die hierdurch erzeugte Redundanz an Wettersensoren wird ausgenutzt, um eine Validierung der erfassten Wetterdaten durchzuführen. Nur wenn beide Wetterstationen identische Wetterphänomene detektieren, wird über die GLT eine Steuerung der relevanten Anlagen veranlasst. Eine Übersicht über die beiden Wetteranlagen auf dem Dach des UBA zeigt die folgende Abbildung:

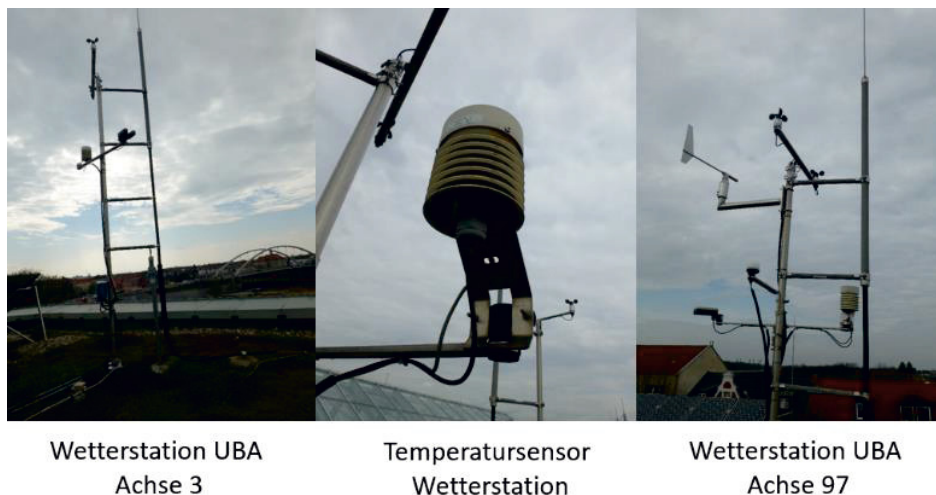


Abbildung 9: Wetterstationen und Außentemperatursensor des UBA (eigene Aufnahmen)

Mit der Inbetriebnahme der Wetterstation entlang der Achse 97 wurde deren Temperatursensor für die Steuerung der Heizungsanlage herangezogen. Gleichzeitig wurde der Temperatursensor für die Fassade auf die Nordseite des Gebäudes verlegt. In diesem Kontext bleibt bei der Verwendung der Daten zu beachten, dass die verschiedenen Anlagen des UBA für den Außentemperatursensor zwar unterschiedliche Identifikationsnummern (AKS, vgl. folgendes Kapitel) verwenden, diese aber jeweils auf den identischen Sensor innerhalb der Wetterstation entlang der Achse 97 zugreifen.

Die Messung der Außentemperatur an der Wetterstation entlang der Achse 97 hat maßgeblichen Einfluss auf die Beheizung des UBA. Für die Betrachtung dieses Berichtes lässt sich diese grundsätzlich in zwei Bereiche einteilen. Dies sind die statische Heizung und die dynamische Heizung, welche auch als Lüftung bezeichnet wird. Über beide Techniken wird Wärme im UBA verteilt und verbraucht. Insgesamt besitzt das UBA sechs Heizkreise getrennt nach den folgenden Bereichen (vgl. Abbildung 10: Innengebäude, Außengebäude, Hörsaal, Bibliothek, raumlufttechnische Anlage (RLT) bzw. Lüftung und Absorptionskältemaschine.





Abbildung 10: Heizkreise des UBA (eigene Aufnahmen)

Bis auf wenige Räume existieren keine Klimaanlage. Diese sind lediglich für den Serverraum Druckerei, Schulungsräume, Pforte und den Hörsaal vorhanden. Die Erzeugung von Wärme geschieht hauptsächlich über zwei Wärmetauscher, die an das Fernwärmenetz gekoppelt sind in Kombination mit Maßnahmen der Wärmerückgewinnung. Ergänzt wird dieser Prozess mit der Verwendung regenerativer Energie. Die Abhängigkeiten für die Steuerung dieser Komponenten werden im Folgenden kurz erläutert:

Über die Außentemperatur und eine definierte Heizkurve (vgl. Abbildung 11) wird der Sollwert der Vorlauftemperatur der statischen Heizung (Innengebäude und Außengebäude) berechnet. Die Lüftung orientiert sich dagegen auf manuell festgelegten Sollwerten, die über die langjährige Erfahrung beim Betrieb der GLT entstanden und angepasst sind. Die Lüftung ist nur dann aktiv, wenn auch die statische Heizung aktiviert ist. Die Grenztemperatur für die Heizung liegt bei einer Außentemperatur von  $15^{\circ}\text{C}$ . Liegt die Außentemperatur über diesem Wert, werden sowohl Heizung als auch Lüftung ausgeschaltet. Um ein ständiges An- und Ausschalten im Grenzbereich zu vermeiden, liegt der Steuerung eine Hysterese zugrunde. Ergänzend hierzu sind in der Steuerung definierte Betriebszeiten (Tag / Nacht, Wochenende, Feiertage) hinterlegt, die eine entsprechenden Verschiebung der Heizkurze und der Ein-/ Ausschaltzeitpunkte der Anlagen vorsehen.

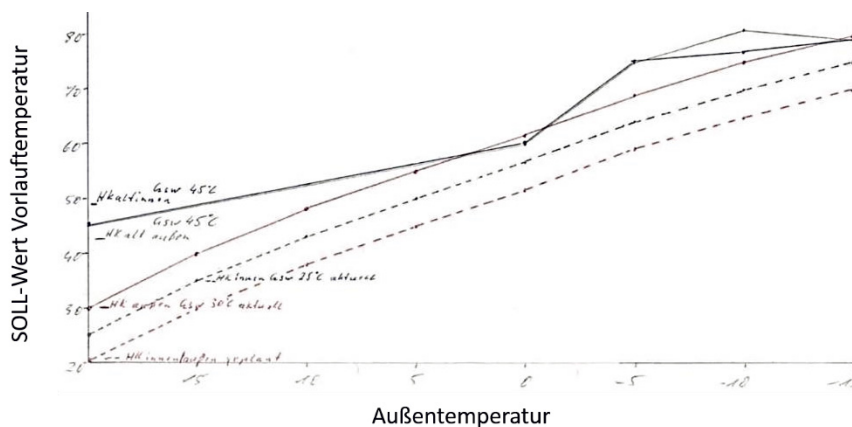


Abbildung 11: Zeichnung der Heizkurve des UBA

Über die Differenz zwischen der Soll- und Ist-Temperatur werden die jeweiligen Ventilstellungen der statischen Heizung und Lüftung gesteuert (vgl. Abbildung 12 und Abbildung 13). Falls die Ist-Temperatur unter den korrespondierenden Sollwerten liegt, werden die entsprechenden Ventile geöffnet, so dass wärmeres Wasser bzw. Zuluft eingeleitet werden kann (Differenz in a) steuert Ventilstellung in b)). Liegen die umgekehrten Verhältnisse vor, werden die Ventile geschlossen. Bei der Lüftung ist ergänzend die Wärmerückgewinnung zu beachten. Deren Leistung wird zuerst ausgeschöpft bevor die Leistung der Heizung hinzugeschaltet wird (Differenz in c) steuert die Ventilstellung d) & e) [zuerst d) bis zur Erschöpfung, dann erst zusätzlich e)). Die Dimensionierung und die jeweiligen prozentualen Öffnungsgrade der Ventile in Abhängigkeit der Anforderung wurden vom Erbauer der Anlage festgelegt und können nicht über die GLT gesteuert werden.

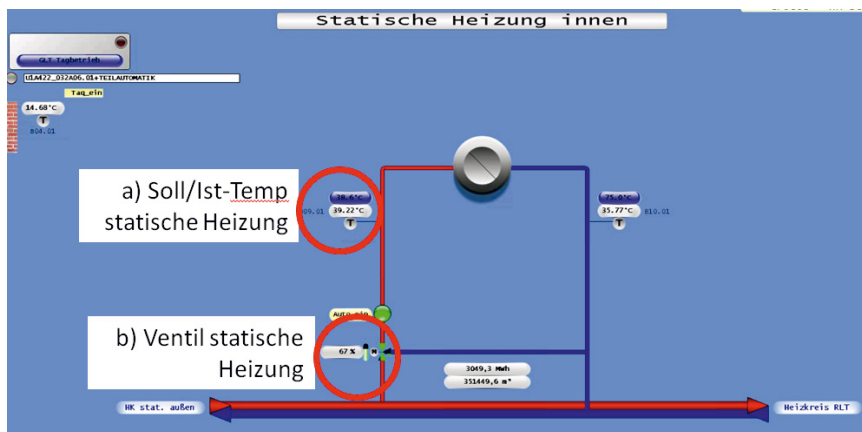


Abbildung 12: GLT-Abbild statische Heizung innen

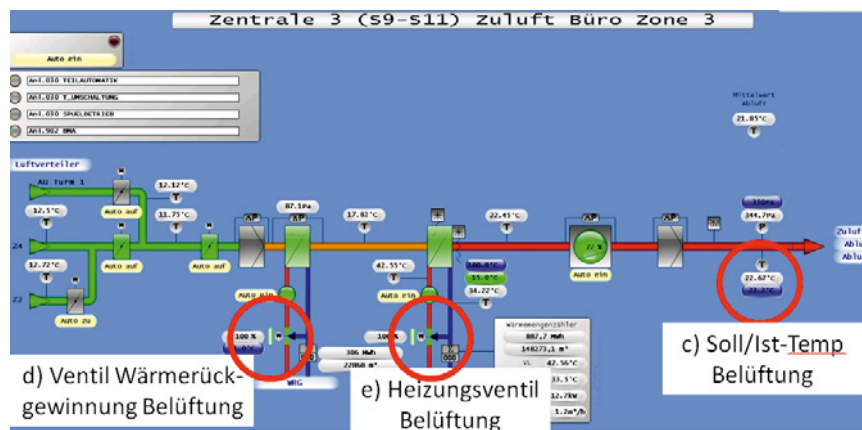


Abbildung 13: GLT-Abbild RLT-Zentrale 3

Die Ventilstellungen in der statischen Heizung und der Lüftung steuern wiederum die Soll-Temperaturwerte des Vorlaufs des sekundären Heizkreislafs in der Fernwärmeübergabe (vgl. Abbildung 14). Die Abhängigkeiten in der GLT setzen sich somit bis zur primären Energiequelle durch. Falls mindestens ein Ventil eine 100%ige Auslastung aufweist, wird die genannte Solltemperatur in der Fernwärmeübergabe erhöht. Besteht diese Anforderung nicht, schließen sich die Ventile der Fernwärme wieder, so dass die Solltemperatur sinkt (die Ventilstellung in b) und e) [aber nicht d)] steuert die Solltemperatur in f, die Differenz in f) steuert wiederum die Ventilstellung in g)). Im Regelfall sind für diesen Prozess die Sensorwerte der Lüftung ausschlaggebend, da diese die höchste Anforderung an Wärme besitzt.



Im folgenden Kapitel müssen die Daten für die im weiteren Verlauf erfolgenden Analysen zunächst grundlegend aufbereitet werden. Hierzu wird der Prozess von der Datenbeschaffung über das Aufsetzen einer Architektur bis hin zu ersten Statistiken zur Datenvalidierung aufgezeigt.

Bei der Beantragung des Projektantrages bestand die Annahme, dass die historischen Daten aus jeweils beiden Untersuchungsobjekten über einen Fernzugriff für die Analyse beim Fraunhofer IAIS verfügbar gemacht werden können. Diese Annahme beruht auf Erfahrungen, die im Vorfeld mit der Sensordatenbank eines weiteren Gebäudes des Bundes gemacht werden konnten, dessen GLT auf der identischen Technik wie der beiden Untersuchungsobjekte beruht. Bei dem Gebäude handelt es sich um das Energieeffizienzhaus Plus in Berlin (Fasanenstraße). Dessen Sensordatenbank beruht auf einem von der Systemplattform Wonderware verwalteten Microsoft SQL Server, der über einen geeigneten Client über einen Fernzugriff angesteuert werden kann. Einmal über einen Fernzugriff verbunden, können die Daten der jeweiligen Datenpunkte aus der Datenbank über den Client mittels SQL-Skripte abgefragt und extern verfügbar gemacht werden.

Nach der Kontaktaufnahme mit den jeweils technisch Verantwortlichen der beiden Gebäude, stellte sich jedoch in den ersten Gesprächen heraus, dass für beide Untersuchungsobjekte kein Fernzugriff auf die historischen Sensordatenbanken der GLT für das Projekt in Aussicht gestellt werden konnte. Die Gründe hierfür unterscheiden sich bei den beiden Bundeseinrichtungen:

UBA: Eine Fernabfrage der historischen Sensordaten der GLT ist analog zum Energieeffizienzhaus Plus über den Microsoft SQL Server möglich, kann jedoch auf der Basis der aktuellen Software-Version erst ab dem Datenbestand ab Ende Februar 2015 durchgeführt werden. Dies ist durch eine ab Ende Februar 2015 eingespielte neue Software Version der Systemplattform Wonderware bedingt, deren Lizenzen keine Abfragen der älteren historischen Sensordaten erlaubt. Um auf die historischen Sensordaten vor Februar 2015 zugreifen zu können, müssten neue Lizenzgebühren vom UBA an den Hersteller der Systemplattform abgeführt werden. Diese Beträge waren jedoch aus dem Budget des Projektes nicht tragbar.

BMG: Die installierte Systemplattform besitzt aus technischer Sicht über die IT-Infrastruktur des BMG keine Möglichkeit des Fernzugriffs auf den Microsoft SQL Server. Die Daten können lediglich vor Ort aus der historischen Datenbank mittels SQL-Skripten geschrieben werden.

Aufgrund dieses Hindernisses musste ein Workaround gefunden werden, der die Daten für das Projekt verfügbar macht (Abbildung 16: Workaround für die Bereitstellung der Daten). Dieser bestand in der Möglichkeit, die Daten nicht über den Microsoft SQL Server vor Ort in den Bundesministerien im Fernzugriff zu exportieren, sondern eine Kopie der Microsoft SQL Datenbank anzufertigen und den Export anschließend intern beim Fraunhofer IAIS durchzuführen. Diese Möglichkeit war sowohl durch die technische Infrastruktur beim BMG, als auch über die Lizenzen des UBA gegeben.



**Bundeseinrichtungen****(UBA, BMG)****Fraunhofer IAIS**

Datenaufbereitung

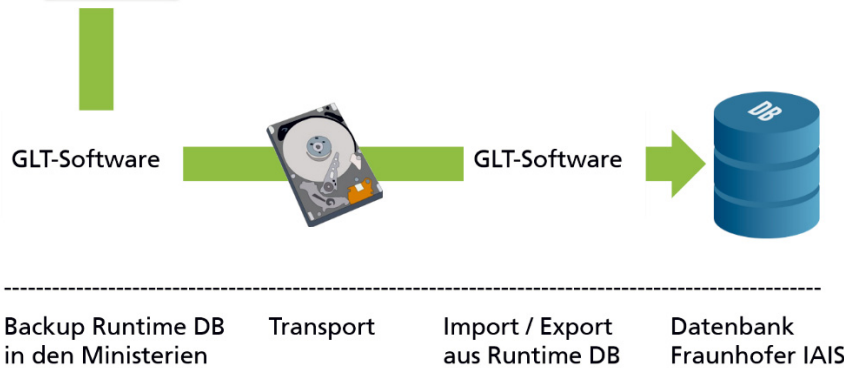


Abbildung 16: Workaround für die Bereitstellung der Daten

Die Anfertigung der Kopie der Datenbank wurde mit Unterstützung der jeweiligen technischen Betreuer der GLT vor Ort in den Bundeseinrichtungen und den jeweils externen Systembetreuern durchgeführt. Dies sind Herr Bretschneider und Herr Walther (FM-Energie) für das UBA, und Herr Dussing und Herr Schmitt (FMSbase) für das BMG. Für den Transport der angelegten Kopie der Datenbank wurde aufgrund von Sicherheitsbedenken der jeweiligen IT ein ftp-Server ausgeschlossen. Aufgrund dessen kam für die jeweilige Kopie der Datenbank eine externe Festplatte zum Einsatz, welche mit einem TrueCrypt-Container für den gesicherten Datentransport ausgestattet war. Die Kopie der Datenbank wurde auf die Festplatte kopiert und anschließend an das Fraunhofer IAIS auf dem Postweg (UBA) und bei einem Vor-Ort-Termin (BMG) übergeben. Ergänzend hierzu wurde jeweils eine Datenpunktliste übergeben, welche eine Auflistung aller Sensoren (GLT und Systemparameter) enthält.

Nach Empfang der Datenträger lag eine Datenmenge von ca. 180 GB für die Daten des UBA und ca. 5 GB für die Datenmenge des BMG vor. Diese verteilen sich auf ca. 485.000 Dateien in 3.650 Ordnern für das UBA und ca. 40.000 Dateien in 2.551 Ordner für das BMG. Die Menge an Datenpunkten liegt beim UBA 3.768, während diese beim BMG bei 1.471 liegt. Da von der GLT-Software in der Regel 1 Ordner pro Tag angelegt wird (bei einem Neustart des Systems wird zusätzlich ein weiterer Ordner erstellt), ist bei den Daten des UBA von einer ca. 10-jährigen und beim BMG von einer ca. 7-jährigen Datenhistorie auszugehen. Diese Zeiträume decken sich in etwa mit den Baujahren der jeweiligen Gebäude.

Für die Weiterverarbeitung der Sensordaten musste beim Fraunhofer IAIS jeweils die identische Systemplattform analog zu den Systemplattformen der beiden Bundesministerien aufgesetzt werden. Dies war notwendig, da die Kopien der jeweiligen Datenbanken in einem proprietären Format vorliegen, die mit keiner anderen Software gelesen und verarbeitet werden können. Die Verwendung der identischen Systemplattform war somit eine notwendige Bedingung, ohne deren Sicherstellung keine Weiterverarbeitung der Sensordaten hätte erfolgen können.

Für das Aufsetzen der Systemplattform wurde dem Fraunhofer IAIS eine kostenlose Forschungslizenz der Systemplattform Wonderware vom Hersteller Schneider Electric zur Verfügung gestellt. Diese besitzt jeweils eine 30-tägige Laufzeit. Während der Exportphase der Sensordaten aus den Datenbanken musste die Forschungslizenz mehrmals verlängert werden. Da die beiden Systemplattformen beim UBA und beim BMG unterschiedliche Versionsnummern und Datenbankversionen trugen, musste Wonderware individuell auf zwei getrennten Rechnern installiert und für das Aufsetzen der Datenbank und den Export der Sensordaten vorbereitet werden.

## 2.1 Datenexport

Im Folgenden bestand der Hauptteil der Arbeit darin die Datenbanken aus den beiden Bundeseinrichtungen zunächst zu reproduzieren und anschließend die Daten aus den Datenbanken zu exportieren. Für die Durchführung dieser Aufgabe galt es im ersten Schritt die Wonderware Software und deren Möglichkeiten des Anlegens einer Datenbank aus einer Kopie zu studieren und umzusetzen. Nach Abschluss dieser Arbeiten lagen die Daten analog zu den Systemplattformen in den jeweiligen Bundeseinrichtungen auf den Rechnern des Fraunhofer IAIS vor, so dass der Export der Sensordaten vorbereitet werden konnte. Hierzu wurde ein Export-Skript für den jeweiligen SQL-Server geschrieben, das die vollständigen Daten der jeweiligen Datenbank in ein Text-Format schreibt. Die Vorgabe, welche Sensoren aus der jeweiligen Datenbank gezogen werden, wurde durch die Datenpunktlisten der beiden Ministerien vorgegeben. Diese enthalten jeweils sowohl die Datenpunkte der Sensoren aus der GLT als auch die Datenpunkte des technischen Systems, welches im Hintergrund die GLT betreibt.

Der Export der Daten aus der Datenbank selbst kann grundsätzlich in zwei Varianten durchgeführt werden: In der ersten Variante werden die Daten mit vorab global festzulegenden festen Zeitabständen exportiert. Dies bedeutet, dass die Datenpunkte nach dem Export aus der Datenbank eine exakt vorab definierte Zeitdifferenz (z.B. 1 Sekunde, 5 Minuten etc.) im Text-Format aufweisen. Sofern keine Änderung der jeweiligen Sensorwerte vorliegt, enthält jeder Datensatz den identischen Wert seines jeweiligen Vorgängers. In der zweiten Variante wird der Export als Delta der Datenpunkte formuliert. Dies bedeutet, dass immer dann ein neuer Eintrag aus der Datenbank in das Text-Format exportiert wird, wenn sich laut Aufzeichnung der Sensorwert gegenüber seinem jeweiligen Vorgänger geändert hat. Die Zeitdifferenz zwischen den Einträgen ist somit bei der zweiten Variante variabel.

Bei der Verarbeitung der Daten stellen Algorithmen bestimmte Anforderungen an die Daten und deren Eigenschaften bezüglich der zeitlichen Auflösung. Nur wenn diese erfüllt sind, kann das Maximum an analytischen Möglichkeiten aus einem Algorithmus herausgeholt werden. Da zum Zeitpunkt des Exports die Art der einzusetzenden Algorithmen noch unbekannt war, mussten die Daten ohne zeitliche Artefakte aus der Datenbank exportiert werden. Zudem waren die individuellen Eigenschaften jedes einzelnen Datenpunktes unbekannt. Dies können ggf. konstante Zeitabstände aufweisen oder aber individuell in Abhängigkeit des GLT-Betriebs gesteuert sein.

Als Methodik für den Export der Sensordaten aus den GLT-Datenbanken der beiden Bundeseinrichtungen wurde sich für die zweite Variante entschieden. Dies stellte zudem sicher, dass das Datenvolumen nicht künstlich in die Höhe getrieben wurde. Nach der Festlegung der Methodik konnte somit der Prozess des Exports angestoßen werden, der die Daten zunächst in das Text-Format schreibt. Hierzu wurde für die jeweiligen Rechner ein SQL-Export-Skript geschrieben und getestet, welches diesen Prozess nach einem manuellen Anstoß autonom steuert. Für das Aufsetzen des Systems, das Schreiben des Skripts und die Laufzeit des Exports kann eine gesamte Laufzeit von ca. 5 Wochen festgehalten werden.

## 2.2 Architektur

Für die weitere Verarbeitung der Daten musste zu diesem Zeitpunkt eine Entscheidung bezüglich der einzusetzenden Architektur getroffen werden. Hierzu stehen z.B. klassische relationale Datenbanken oder neuartige Big Data Systeme zur Wahl. Zur Durchführung der konkreten Aufgaben im Projekt ist jedoch außer den in Kapitel 2.6 beschriebenen Systemen und Tools der Einsatz von speziellen Big-Data Technologien

nicht erforderlich gewesen. Bei dem zu analysierenden Datenvolumen von 185 GB (UBA + BMG) an Rohdaten sind herkömmliche, datenbank-basierte Technologien vollkommen ausreichend. Dies liegt letztendlich darin begründet, dass im Projekt ausschließlich Sensordaten aus zwei Gebäuden zur Verfügung standen und verarbeitet worden sind. Diese Daten konnten, wie in Abschnitt 2.1 beschrieben, auf herkömmlichen externen Festplatten transportiert werden.

Im Kontext von Industrie 4.0 stellt sich jedoch durchaus die allgemeine Frage, welche Architektur-Szenarien für die Erfassung und Auswertung von Sensordaten aus vielen Gebäuden denkbar und sinnvoll sind. Aufgrund dessen wird im Folgenden ein Ansatz beschrieben, der in der Lage ist, die dabei anfallenden Datenmengen zu bewerkstelligen. Dabei wird teilweise auf Kapitel Bezug genommen, die hierfür eine Relevanz besitzen. Zur Anwendung kommen diese für das Projekt aufgrund der überschaubaren Datenmenge nicht. Bei den Architektur-Szenarien stellt sich insbesondere die Frage nach einem Architekturkonzept, das es erlaubt:

- Daten aus einer skalierbaren Menge von Gebäuden zu erfassen und zu speichern,
- die gespeicherten Daten analog zu den in den Kapiteln 2 bis 4 beschriebenen Schritten zu vorbereiten, zu validieren und zu analysieren,
- und Ergebnisse der Analyse in geeigneter Form in die GLT einzubringen.

Grundvoraussetzung für den Aufbau einer passenden Architektur ist es, dass die auszuwertenden GLT-Installationen über geeignete Netzwerkzugänge mit einem Rechencluster verbunden sind, in dem die Daten gespeichert und ausgewertet werden können. Dieses Rechencluster wird typischerweise in der „Cloud“ realisiert, wobei hier noch unterschieden wird, ob es sich dabei um eine öffentliche oder eine private Cloud-Lösung handelt. In beiden Fällen sollten über standardisierte Verfahren der Verschlüsselung und Zugriffskontrolle die Vertraulichkeit und Authentizität der Daten gesichert werden. Gegebenenfalls können auch Verfahren zur Anonymisierung der Daten eingesetzt werden, um die Privatsphäre von GLT-Betreibern und den in den Gebäuden arbeitenden Personen zu schützen.

Unter den genannten Voraussetzungen lässt sich im Rechencluster eine Big-Data Architektur nach dem etablierten Grundmuster der Lambda-Architektur aufbauen, welches in Abbildung 17: Grundaufbau der Lambda-Architektur veranschaulicht ist.

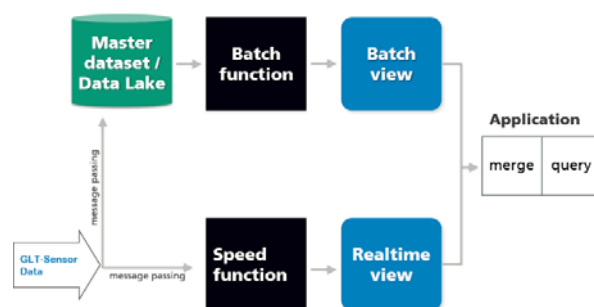


Abbildung 17: Grundaufbau der Lambda-Architektur

Die Lambda-Architektur ist ein bewährter Ansatz zur Analyse, Konstruktion und Erläuterung von Big-Data-Anwendung [6]. Die in der Architektur vorgesehene Modularisierung spiegelt typische Anforderungen an Big-Data-Anwendungen wider und systematisiert diese. Der Architektur-Ansatz hilft, Systemanalysen durchzuführen, die technische und nicht-funktionale Anforderungen an neue Anwendungen aufdecken und beurteilen -- unabhängig davon, in welcher Form und welchem Umfang die Module als technische Komponenten der Anwendung realisiert werden. Ein mögliches Ergebnis

einer solchen Systemanalyse kann auch sein, dass bestimmte Teile der Lambda-Architektur in einer konkreten Anwendung nicht benötigt werden.

Die folgenden Überlegungen nutzen die Lambda-Architektur, um die Schritte zur Speicherung, Verarbeitung und Analyse von Sensordaten sowie die Bereitstellung von Analyseergebnissen abzubilden und zu verorten.

Eine der Grundaufgaben einer Big-Data Architektur zur Verarbeitung von GLT-Daten ist es, große Volumen an Sensor-Daten speichern zu können. Dabei ist es oftmals sogar unerheblich, genau definieren zu können, zu welchen Zweck die Daten aktuell genau eingesetzt werden sollen. Allein das Vorhanden der Daten birgt schon das Potential, für neue Ideen der Analyse und Auswertung vorbereitet zu sein. Daher ist es ein wichtiges Prinzip von Big-Data Architekturen, Daten in ihrer Roh-Form möglichst unverändert aufnehmen und für weitere Analysen bereitstellen zu können. Diese Vorgehensweise wird oft auch als das Anlegen eines sogenannten „Data Lake“ bezeichnet, im Kontext der Lambda Architektur wird hier vom sogenannten „Master Dataset“ gesprochen.

Grundprinzipien der Aufnahme und Speicherung von Daten in das Master Dataset sind:

- Daten aus den verschiedensten Quellen werden in das System aufgenommen und in Rohform, schemafrei abgelegt. Dabei werden einkommende Daten pro Datenelement mit einem oder mehreren Zeitstempeln versehen, die angeben, wann die Daten entstanden sind und wann sie in das System aufgenommen worden sind.
- Die Daten sind „immutable“, sie werden im Master Dataset nicht verändert. Wenn nötig, werden sie für eine Anwendung in eine neue Repräsentation überführt, ohne jedoch die Rohdaten zu verändern oder löschen. Wenn gelöscht wird, dann nur unter Einhaltung eines besonders dafür vorgesehenen Prozesses.
- Durch Aufbewahrung der mit einem Zeitstempel versehenen Rohdaten können neue Anwendungen stets auf den größtmöglichen Datenbestand zugreifen, es können historische Analysen und Modelle unter Berücksichtigen der Historie erstellt werden, und der Anwendungszustand kann zu einem beliebigen früheren Zeitpunkt passend rekonstruiert werden.

Neben den Roh-Daten werden in dem Data-Lake aufbereitete Daten gespeichert und für spätere Analysen zur Verfügung gestellt. Zur Aufbereitung der Daten werden „Batch-Funktionen“ im Big-Data System programmiert (siehe Abbildung 17: Grundaufbau der Lambda-Architektur), die dann die in Abschnitt 2.5 beschriebenen Schritte der Datenvalidierung durchführen. Auch die in Abschnitt 2.6.1.2 beschriebene „fit\_tags“ Funktion lässt sich als Batch-Funktion realisieren, die dann als „Batch-View“ einen Ausschnitt der Sensordaten der durchzuführenden Analyse mit jeweils passendem Zeitraster bereitstellt.

Die eigentlichen Analysen der Sensordaten erfolgen meist explorativ und interaktiv nur auf einem für die spezielle Fragestellung relevanten Teil der Sensoren und Zeitraum. Dies bedeutet im Sinne der Lambda-Architektur, dass für eine Analyse eine spezielle „Batch-View“ erzeugt und in einem für die Analyse geeigneten Tool materialisiert wird. So kann z.B. für eine explorative Analyse über eine vom Big-Data System bereitgestellte SQL-Schnittstelle der gewünschte Teil der validierten Daten in ein Visualisierungs-Tool geladen werden.

Die Ergebnisse einer Analyse können unterschiedlicher Art sein. Sie können zum einen allgemeiner Natur sein (vgl. z.B. Abschnitt 4.3.1 und 4.3.7). Im Resultat geben diese Hinweise auf eine Reparatur oder Konfiguration des GLT-Systems. Diese Art von Ergebnissen erfordert eine Bewertung durch „menschliche“ Experten und resultieren in vom Menschen durchzuführende Aktionen (wie das Anbringen und Konfigurieren eines zusätzlichen Wärmesensors). Diese Art von Ergebnissen und der dazugehörigen Reaktion erfordert keine weitere besondere Unterstützung seitens der Architektur.

Andere Analyse-Ergebnisse können aber auch durchaus helfen, Verbesserungen zu „automatisieren“ und in den Regelbetrieb einzuführen. So zeigt etwa die Analyse aus 4.3.4.2, dass bei einigen Regelkreisen des betrachteten GLT Systems unerwünschte „Schwingungen“ auftreten. Neben der Erkenntnis, dass in diesen Regelkreisen die Dimensionierung und Regelung überprüft werden sollte, zeigt sich aber auch, dass es sinnvoll und machbar ist, derartige Muster automatisiert zu erkennen. Eine solche Muster-Erkennung in Echtzeit würde in der Lambda-Architektur typischerweise in der „Speed Function“ realisiert werden. Obwohl dies technisch machbar ist, wäre im vorliegenden Anwendungsfall eine Realisierung innerhalb der jeweiligen GLT vorzuziehen. Dies bedeutet, dass die Daten-Analyse auf den historischen Daten im Data-Lake genutzt wird, um kritische Verhaltensmuster zu identifizieren. Die Überprüfung, ob und welche Sensoren und Regelkreise solche Muster aufweisen, könnte jedoch in der GLT in Echtzeit realisiert werden, ohne dass die Sensordaten dazu unmittelbar an das Rechencluster geschickt werden müssen.

## 2.3 Datenimport

Nach der Durchführung des Exports der Daten wurden diese in eine Oracle-Datenbank im räumlich getrennten, internen Bereich der IT des Fraunhofer IAIS importiert. Hierzu wurden die Daten zunächst von den jeweiligen lokalen Rechnern auf den Rechner der Datenbank per mobilen Datenträger verschoben. Dies beschleunigt den Import, da man beim Schreiben der Daten in die Datenbank nicht den Limitierungen des Netzes, wie z.B. Schwankungen der Netzkapazität und Netzausfällen, unterlegen ist.

Nach Abschluss des Kopierens konnte der Import der Daten aus den jeweiligen Dateien in die Datenbank angestoßen werden. Hierzu wurde zunächst eine leere Tabelle mit der Meta-Struktur der Quelltable in der neuen Datenbank angelegt. Hierdurch wird den Daten eine Vorgabe gemacht, wie diese in die Datenbank zu importieren sind. Im Anschluss wurde ein Import-Skript auf SQL-Basis geschrieben und ausgeführt, welches die Daten in die zuvor angelegte Tabelle schreibt. Im Unterschied zu den Originaldaten, in denen die Daten in einzelnen Dateien über mehrere Ordner verteilt vorliegen, werden im Ergebnis die Daten in eine einzelne Tabelle geschrieben.

Die folgenden Tabellen geben einen abschließenden Überblick über die in den Einzelschritten angefallenen Datenmengen:

Backup Runtime Datenbank	TEXT-Dateien	Zieldatenbank (Oracle)
ca. 180 GB Dateigröße 485.000 Dateien 3.650 Ordner	ca. 1.25 TB Dateigröße 479.00 Dateien 131 Ordner	Ca. 7 Mrd. Datenzeilen

Tabelle 1: Eckdaten der Stufen der Datenaufbereitung (UBA)

Backup Runtime Datenbank	TEXT-Dateien	Zieldatenbank (Oracle)
ca. 5 GB Dateigröße 40.000 Dateien 2.551 Ordner	ca. 9.4 GB Dateigröße ca. 36.000 Dateien 104 Ordner	

Tabelle 2: Eckdaten der Stufen der Datenaufbereitung (BMG)

Zu diesem Zeitpunkt des Projektes musste festgestellt werden, dass die Datenaufbereitung aus den vorab genannten Gründen wesentlich mehr Zeit in Anspruch genommen hatte als ursprünglich geplant. Dieser zeitliche Mehraufwand war in den kalkulierten Zeitplan nicht eingeplant, da zu Beginn von barrierefreien Zugangsbedingungen ausgegangen worden war. Um den Faktor Zeit für die Datenaufbereitung nicht noch weiter zu strapazieren, wurde beschlossen, den Datenraum für die folgenden Schritte der Datenaufbereitung und Analyse einzuschränken. Aufgrund dessen wurde sich auf die Datenbasis der GLT des UBA konzentriert, da diese im Vergleich zur GLT des BMG weitaus mehr Datenpunkte besitzt. Zudem ist in der Datengrundlage eine längere Datenhistorie ausgewiesen.

---

Datenaufbereitung

---

## 2.4 Datensemantik

Jeder Datenpunkt ist mit einem eindeutigen Identifier (ID) und einer Beschreibung im Klartext versehen. Diese ID befindet sich sowohl in den Daten in der Datenbank, als auch an vielen Sensoren im Gebäude. Über die Beschreibung wird im Allgemeinen der Inhalt des jeweiligen Datenpunktes entschlüsselt. Somit wird für den Nutzer verständlich, was sich hinter dem jeweiligen Datenpunkt verbirgt. Diese Beschreibung ist für die verwendete Datenbasis in großen Teilen jedoch nicht selbsterklärend, da sie z.T. mit Abkürzungen versehen ist, die sich von sich aus nicht dem Nutzer erschließen.



Abbildung 18: AKS Vorlauf-Temperatur Sek. Kreislauf

Eine semantische Beschreibung der Datenpunkte im Klartext ist jedoch für das Verständnis der Datenpunkte unabdingbar. Um die Beschreibung hinter den einzelnen Datenpunkten aufzuschlüsseln wurde deshalb auf die Eigenschaften der ID zurückgegriffen. Die ID ist nach den Vorgaben des Anlagenkennzeichnungsschlüssels (AKS) aufgebaut, welche in einer DIN-Norm festgehalten sind. Demnach handelt es sich bei dem AKS um eine hierarchisch aufgebaute ID, welche speziell für Anlagenteile der GLT vom Bund entwickelt wurde. Durch die Verknüpfung des AKS in Daten mit den entsprechenden AKS in der DIN konnten somit alle Datenpunkte im Klartext aufgelöst und dargestellt werden.

Diese Verknüpfung ist jedoch anfänglich fehlgeschlagen. Der Grund liegt in einer abweichenden Darstellung des AKS in den Datenpunkten im Vergleich zur DIN-Norm. Während in der DIN-Norm der AKS über 30 Stellen verfügt, weisen die Datenpunkte lediglich eine Länge von 29 Stellen auf. Eine Verknüpfung schlug folglich zunächst fehl. Erst mit der Identifizierung der fehlenden Stelle in der Betriebsmittelkennzeichnung des AKS konnte das Problem behoben werden. Die Lösung bestand in der Ergänzung der Betriebsmittelkennzeichnung um eine 0 an zweiter Stelle (vergleiche Abbildung 19: Anlagenkennzeichnungsschlüssel).



	Länderkennung	Ortskennung	Gebäude	Geschoss	Dokumentationsstand	Kostengruppe	Anlagennummer	Betriebsmittel	lfd. Nr.	Funktion	lfd. Nr.
Ebene	1	2	3	4	5	6	7	8	9	10	11
Stelle	1	2-6	7-9	10-11	12	13-16	17-19	20-23	24-26	27-28	29-30
Beispiel	D	06844	UBA	U1	A	431_	010	B04	_01	TX	01

Datenaufbereitung

Abbildung 19: Anlagenkennzeichnungsschlüssel

## 2.5 Datenvalidierung

Nachdem die Daten in die vorab angelegte Datenstruktur in der Datenbank importiert und aufbereitet wurden, erfolgte im nächsten Schritt eine Prüfung der Rahmenbedingungen hinsichtlich der Verwendbarkeit der Daten für tiefergehende Analysen. Analysen, die eine gewisse Datenhistorie voraussetzen, sind z.B. darauf angewiesen, dass die Daten über die zu untersuchenden Zeiträume verlässlich Datenpunkte aufweisen und keine großen Lücken in der Datenhistorie zu finden sind. Aufgrund dessen wurden auf den Daten Validierungen durchgeführt, die im Folgenden beschrieben werden. Diese sind grundsätzliche allgemeiner Natur und lassen sich auch auf Daten anderen GLT-Datenbanken anwenden.

1) In der ersten Validierung wurde auf Basis der erstellten Datenhistorie analysiert, ab welchem Zeitpunkt der jeweilige Datenpunkt in der Datenbank erscheint. Nur wenn der Großteil der Datenpunkte frühzeitig in der Datenbasis auftaucht ist davon auszugehen, dass eine lange Historie der jeweiligen Sensoren für weitere Analysen vorliegt. Die Ergebnisse zeigt die folgende Abbildung 20: Sensoren - Beginn der Aufzeichnung. Die Balken zeigen jeweils die Anzahl an Sensoren, die zu einem bestimmten Zeitpunkt in der Datenbank zum ersten Mal registriert wurden. Hier ist zu erkennen, dass der Großteil der Datenpunkte bereits seit Beginn der Aufzeichnung der Daten in 2006 vorliegt. Zudem ist ein weiterer Peak im Jahre 2014 in den Daten zu erkennen.

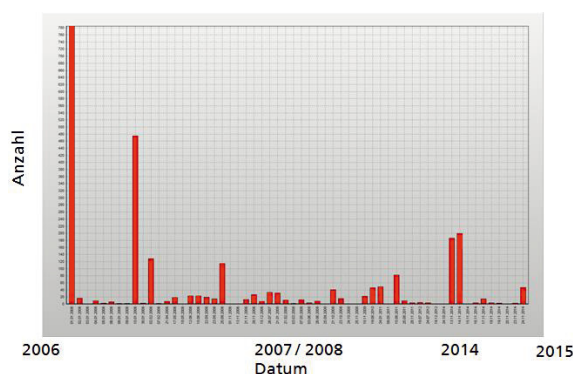


Abbildung 20: Sensoren - Beginn der Aufzeichnung

Die GLT ist ständig Veränderungen in Form von Optimierungen und Anpassungen unterlegen. Der zweite größere Peak in 2014 ist auf eine solche größere Veränderung der GLT zurückzuführen. Bei dieser wurde in 2014 die Umstellung auf Archestra - Technologie als Teil von Wonderware realisiert. Zudem wurde die Außenstelle Marienfelde an das System angebunden. Dies führte in der Datenbank sowohl zu neuen als auch veränderten Datenpunkten, die sich in Form dieses Peaks widerspiegeln. In den folgenden Kapiteln wird diese Umstellung und die Auswirkungen auf einen Teilbereich der AKS insbesondere bei der Betrachtung der SOLL-IST-Werte noch eine Rolle spielen.

2) In der zweiten Validierung wurde auf Basis der erstellten Datenhistorie analysiert, bis zu welchem Zeitpunkt der jeweilige Datenpunkt in der Datenbank erscheint. Nur wenn der Großteil der Datenpunkte möglichst spät letzte Einträge in der Datenbank aufweisen, kann von einer langen Datenhistorie für tiefer gehende Analysen ausgegangen werden. Das Ergebnis zeigt, dass alle zu finden Datenpunkte ihren letzten Sensorwert mit einem Zeitstempel aus 2015 liefern. Somit liefern alle seit Beginn der Aufzeichnung installierten Sensoren auch 2015 weiterhin Datenpunkte.

Datenaufbereitung

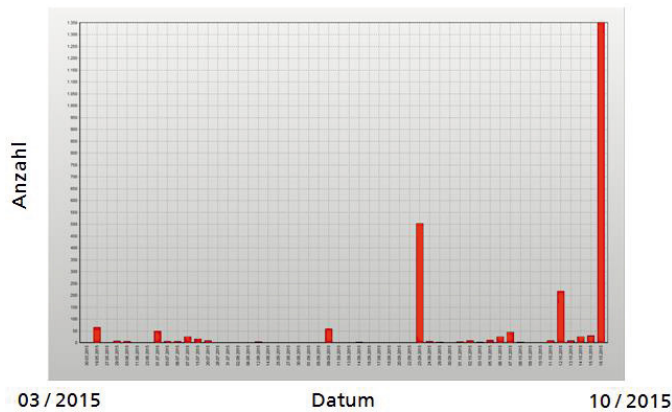


Abbildung 21: Sensoren – Datum Ende der Aufzeichnung

3) In der dritten Validierung wurde auf Basis der erstellten Datenhistorie analysiert, wie konstant die Sensorik der Datenpunkte über die gesamte Betriebszeit der GLT Daten liefern. Hierzu wurde im Detail analysiert, wie viele Datenpunkte pro Monat über die gesamte Datenhistorie mindestens 1 Datensatz pro Monat in der Datenbank der GLT liefern. Im Ergebnis kann auf eine konstante Menge an Datenpunkten der installierten Sensoren zurückgegriffen werden (vergleiche Abbildung 22: Anzahl Sensoren pro Monat). Zu Beginn der Aufzeichnung ist ein Anstieg der Anzahl an Sensoren zu erkennen. Dies ist auf die Installationsphase der GLT zurückzuführen, in der einzelne Anlagenteile sukzessive über die Zeit in Betrieb genommen wurden. Vereinzelte Lücken in den Daten sind dagegen auf Wartungsfenster oder kleinere Defekte zurückzuführen. Hier ist ebenfalls wieder deutlich der Anstieg der Anzahl der Sensoren im Jahre 2014 zu erkennen, was insbesondere im November 2014 zu Tage tritt und auf die zuvor beschriebene Umstellung zurückzuführen ist.

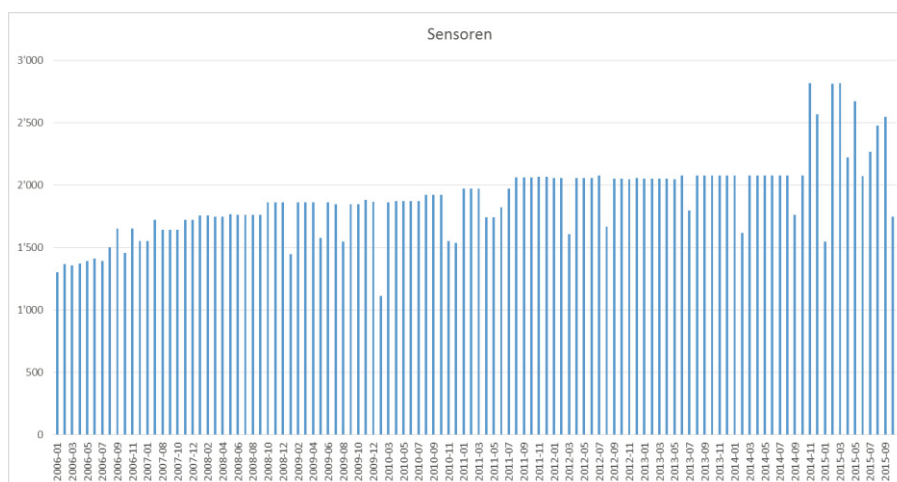


Abbildung 22: Anzahl Sensoren pro Monat



Auf der Basis dieser einzelnen Validierungsschritte kann auf eine langfristige und relativ konstante Datenbasis geschlossen werden, die für das Analysevorhaben zur Verfügung steht. Dieses Bild ergibt sich aus der Validierung der Gesamtmenge an Datenpunkten. Einzelne Datenpunkte können davon abweichen, was sich im weiteren Verlauf der Analyse noch als relevant herausgestellt hat. Aufgrund der großen Menge an Datenpunkten, kann dies nicht für jeden einzelnen Datenpunkt untersucht werden. Diese Analyse auf Einzelbasis wird individuell bei der Verwendung der jeweiligen Datenpunkte in Analysen durchgeführt.

4) Eine weitere Validierung betrifft die Werteausprägung der einzelnen Datenpunkte. Während bisher lediglich die Menge an Datenpunkten für eine Validierung herangezogen wurde, wird in der folgenden Validierung der jeweilige Wertebereich der Datenpunkte überprüft. Als Basis dienen hierzu die Vorgaben aus der zur Verfügung gestellten Datenpunktliste, welche jeweils die minimal und maximal zulässigen Wertebereiche pro verwendeter Einheit des Sensors vorgeben. Diese werden auch als *min-raw* und *max-raw* bezeichnet. Diese Vorgaben wurden mit den jeweiligen Minimal- und Maximalwerten aus der vorliegenden Datenmenge abgeglichen.

Das Ergebnis zeigt die folgende Übersicht (vergleiche Abbildung 23: minimale und maximale Werte pro Einheit). Dabei werden Sensorwerte außerhalb der zulässigen Wertebereiche in orange hinterlegt dargestellt. Bei den hinterlegten weißen Bereichen handelte es sich um Einheiten, deren jeweiliger Wertebereich eingehalten wurde. Die Mehrheit der von den installierten Sensoren verwendeten Einheiten liegt demnach außerhalb des definierten Wertebereiches, da der orangene Anteil überwiegt. Dieses sind Werte, die vor einer Analyse gefiltert werden müssen, da diese ansonsten das Ergebnis der Analysen verfälschen. Dies wurde u.a. mit datenbankseitig implementierten Funktionen und Skripten durchgeführt, welche im Kapitel über die eingesetzte Software im Detail beschrieben werden.

	UNIT	MIN-Wert (Vorgabe)	MAX-Wert (Vorgabe)	MIN-Wert (Daten)	MAX-Wert (Daten)
1	%	0	100	-8607309824	8565067776
2	°	0	100	-35	999
3	bar	0	100	0	7
4	°C	0	10000000000	-9563012096	9026141184
5	digi	0	100	-10	76,3671875
6	Digits	0	100	0	32767
7	g/kg	0	100	0	20,59937096
8	hPa	0	100	0	1036,968994
9	J/G	0	100	-17,00938225	85,09215546
10	kg/kg	0	100	-2	1,306243628
11	klux	0	100	0	150
12	kVA	0	500	0	703,0240479
13	kW	0	500	-107479312	125156
14	mbar	0	100	-2065,397217	198283,5
15	Me.Units	0	1000	-107151616	107151616
16	m³/h	0	100	-1298415,875	3248571904
17	mm/min	0	100	0	10
18	m/s	0	100	-1298415,875	35
19	None	0	2000000000	0	1445022976
20	Pa	0	100	0	4000
21	ppm	0	100	0	5000
22	%rF	0	100	-2	2000
23	V	0	20000	0	1199323648
24	W	0	1000000000	0	11025
25	W/m²	0	1000000000	0	1537
26				0	1

Abbildung 23: minimale und maximale Werte pro Einheit

5) In einem weiteren Schritt der Datenvalidierung wurde die Analyse der Werteausprägung zeitlich in einzelne Jahre unterteilt. Hierdurch können Problemfälle ggf. genauer zeitlich eingeordnet werden. Hierfür wurden die Datenwerte ergänzend in die folgenden sechs Wertekategorien eingeteilt:

- 1) Wert  $< \text{min-row}$  : Minimum unterschritten
- 2) Wert  $= \text{min-row}$  : Minimum wird angenommen
- 3) Wert  $> \text{max-row}$  : Maximum überschritten
- 4) Wert  $= \text{max-row}$  : Maximum wird angenommen
- 5) Wert ohne zugehörige Grenzwerte  $\text{min-row}$  und  $\text{max-row}$
- 6)  $\text{min-row} < \text{Wert} < \text{max-row}$  : Wert liegt im gewünschten Bereich

Die folgende Abbildung 24 stellt die Anteile der jeweiligen Kategorien im Laufe für die Jahre 2006 bis 2015 dar. Als Darstellungsform wurde ein Kurvendiagramm gewählt. In diesem zeigt sich, dass der Anteil der validen Werte für den gesamten Zeitraum in einem Bereich zwischen 80 und 90% liegt. Die anderen Wertekategorien stellen jeweils nie mehr als 5% der Punkte dar. Eine Ausnahme bildet hierbei die Kategorie „Wert  $> \text{max-row}$ “. Knapp 20% der Daten aus dem Jahr 2013 überschreiten das vorgegebene Maximum. Die Ursache dieser Beobachtung lässt sich anhand der vorliegenden Daten nicht ermitteln, sondern dient lediglich als Indikator für die Datenqualität. Da die Daten rückwirkend betrachtet werden, wird diese reaktiv und nicht proaktiv gewährleistet. Letzterer Ansatz versucht den technischen Ursachen mangelhafter Datenqualität auf den Grund zu gehen, damit keine Datenpunkte verworfen werden müssen.

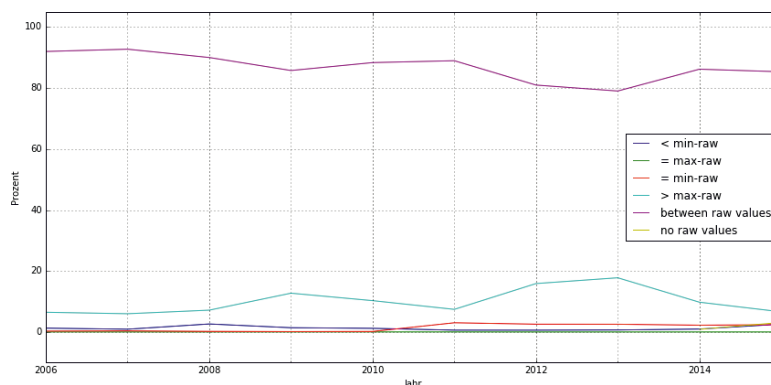


Abbildung 24: Anteile der Datenqualitätskategorien von 2006 bis 2015

Die durchgeführten Validierungen bieten durch ihren Inhalt und ihre Darstellung grundsätzlich eine schnelle Übersicht über wichtige Eckpunkte des zu analysierenden Datensatzes. Diese Informationen sind sowohl für eine Langzeitbeobachtung der Daten durch die GLT-Betreiber von Interesse, als auch für externe Dienstleister, die sich vorab von Analysen einen schnellen und umfassenden Überblick über die jeweils vorliegende Datenlandschaft machen möchten.

## 2.6 Datenverarbeitung und Analyse-Tools

Für die Durchführung des Projektes wurde auf eine Auswahl an bereits existierenden Softwareprodukten und Tools gesetzt. Diese haben sich in bisherigen Projekten im Zusammenspiel bewährt und bilden aktuell gängige Tools zur Analyse von großen Datenbeständen ab. Im folgenden Kapitel wird die eingesetzte Software beschrieben und in ihrer Funktion eingeordnet. Dies ist unter dem Punkt „extern“ zu finden. Der Softwareanteil der Datenbank wird dabei ausgeklammert, da dieser in Ansätzen im Kapitel über die Architektur beschrieben wird. Eine Evaluierung der analytischen

Möglichkeiten der GLT-Software selbst fand in diesem Rahmen ebenfalls nicht statt. Ergänzend hierzu werden auch selbst geschriebene bzw. programmierte Skripte und Funktionen beschrieben. Diese sind zunächst unter dem folgenden Punkt „intern“ beschrieben. Im Allgemeinen wird hierbei auf die wesentlichen Komponenten eingegangen, die für die grundsätzliche Auf- bzw. Verarbeitung der Daten eingesetzt wurden. Sollten für einzelne Analysen spezielle Aufbereitungen notwendig sein, werden diese in den entsprechenden Kapiteln behandelt. Eine Orientierung für eine Einordnung der Software in die Datenverarbeitung bietet die folgende Abbildung:

-----  
Datenaufbereitung  
-----

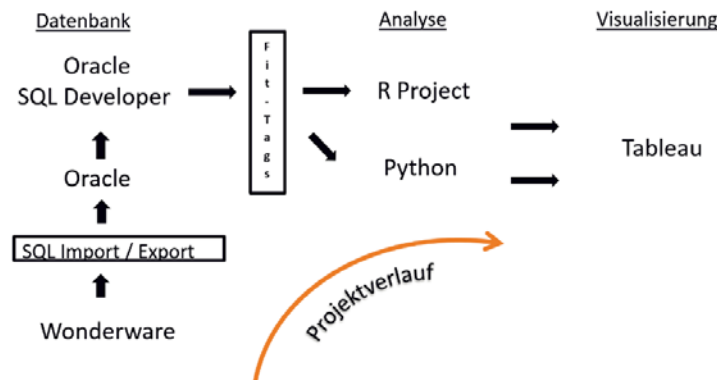


Abbildung 25: Übersicht Software im Projekt

## 2.6.1 Intern

Im Folgenden erfolgt zunächst eine Übersicht der für dieses Projekt eigens programmierten Softwarekomponenten.

### 2.6.1.1 SQL-Export-Import-Skripte

Wie in den Kapiteln 2.1 und 2.3 beschrieben, wurden die Daten zunächst aus der Datenbank hinter der GLT exportiert um diese in die für die Analyse als Basis verwendete Oracle Datenbank zu importieren. Für beide Prozesse wurden jeweils entsprechend Skripte geschrieben und angewendet. Dies sei hier der Vollständigkeit halber erwähnt. Die Details sind den entsprechenden Kapiteln zu entnehmen.

### 2.6.1.2 Die „fit\_tags“-Funktion

Wie im Kapitel über den Datenexport beschrieben, wurden die Daten als zeitliches Delta der Datenpunkte exportiert. Dies bietet zwar die genannten Vorteile (Originalität, Datenmenge), hat jedoch den Nachteil der Asynchronität. Ändern sich die Datenpunkte bei einem Vergleich zweier oder mehr Sensoren nicht zum identischen Zeitpunkt, ist eine unmittelbare Vergleichbarkeit der Sensoren nicht mehr gegeben. Aufgrund dessen wurde eine Funktion auf Datenbankebene geschrieben und implementiert, welche die zeitliche Vergleichbarkeit herstellt:

Die Aufzeichnungszeitpunkte verschiedener Sensoren sind voneinander unabhängig. Dies stellt für analytische Fragestellungen, die auf einer zeitgleichen Betrachtung von mehreren Sensoren beruhen, ein Problem dar, weil manche Sensoren zum Betrachtungszeitpunkt einen undefinierten Wert haben können. Es ist deshalb notwendig, undefinierte Werte aus vorhandenen, zeitlich am nächsten liegenden Sensorwerten zu interpolieren.

Die im Rahmen dieses Projekts programmierte Funktion *fit\_tags* erfüllt den Zweck, eine zeitliche Synchronisierung von Sensoren durch Interpolation von undefinierten Werten zu ermöglichen, so dass die Werte einer Reihe von Sensoren zu vorgegebenen Zeitpunkten gemeinsam betrachtet werden können. Dies ist eine wichtige Vorbedingung, um die Anforderungen von Algorithmen zu erfüllen, die nicht mit fehlenden Werten in Eingabedaten umgehen können.

Für die Anwendung der Funktion muss zunächst der zeitliche Rahmen definiert werden, in dem der Abgleich der Datenpunkte durchgeführt werden soll. Dies geschieht durch die Angabe der Anfangs- und Endzeiten des gewünschten Untersuchungszeitraums. Anschließend erfolgt die Spezifikation der zeitlichen Schrittweite. Dies ist die Zeit, die in der gewünschten Sensortabelle zwischen zwei Datenpunkten liegt. Ein Datenpunkt entspricht einer Zeile in der Ergebnistabelle –jede Spalte enthält die (ggf. interpolierten) Messungen eines der angefragten Sensoren.

Die Interpolation fehlender Werte geschieht nach Maßgabe verschiedener wählbarer Parameter. Diese Parameter werden im Folgenden gelistet. Im Wesentlichen geschieht die Interpolation von Werten zu undefinierten Zeitpunkten entweder durch ein „Fortschreiben“ des letzten bekannten Wertes, oder durch die Mittelung der Werte von dem unmittelbaren zeitlichen Vorläufer und Nachfolger. Mittels der Parameter wird zum Beispiel kontrolliert, wie lange die erlaubte Zeitspanne zwischen Vorläufer und Nachfolger sein darf, damit eine Mittelung noch sinnvoll erscheint (bspw., wenn über mehrere Stunden oder Tage keine Daten erfasst wurden). Sollte die Schrittweite kleiner als das reale Änderungsintervall des jeweiligen Datenpunktes sein, wird der jeweils letzte Wert des Datenpunktes übernommen. Im aktuellen Zustand können mit der Funktion maximal 64 Datenpunkte gleichzeitig zeitlich angepasst und verglichen werden. Dies ist ausreichend für die in den folgenden Kapiteln ausgeführten Analysen

**Parameter:**

1.	Datenquelle	<b>Auswahl der Datenquelle</b>  <i>"uba", "bmg"</i> oder ggf noch weitere
2.	Startzeitpunkt	<b>Ab wann sollen Daten geliefert werden</b>  Format <i>yyyy[-mm[-dd[+hh24[:mi[:ss]]]]]</i> , wobei <b>abgerundet</b> wird. Z.B.  2014 = 2014-01-01+00:00:00 2014-02 = 2014-02-01+00:00:00 2014-02-23+18 = 2014-02-23+18:00:00
3.	Endzeitpunkt	<b>Bis wann sollen Daten geliefert werden</b>  Format <i>yyyy[-mm[-dd[+hh24[:mi[:ss]]]]]</i> , wobei <b>aufgerundet</b> wird. Z.B. 2014 = 2014-12-31+23:59:59 2014-02 = 2014-02-28+23:59:59 2014-02-23+18 = 2014-02-23+18:59:59  Alternativ kann auch "=" angegeben werden, um den gleichen String als Eingabe zu verwenden, der als Startzeitpunkt übergeben wurde.  Eine weitere Alternative besteht in einem Offset, z.B. 2 Tage nach Startzeitpunkt: "+2d": s = second m = minute

		h = hour d = day w = week o = month !!! m wird für Minuten verwendet q = quarter (3 month) y = year
4.	<b>Schrittweite</b>	<b>Spezifikation der Schritte</b>  Entweder fest vorgegeben, z.B. "step 1m" = jede Minute, oder "step 30s" = 30 Sekunden oder an einen der ausgegebenen Sensoren, z.B. "tag 3" oder an einen beliebigen Sensor "<<tagname>>"
5.	<b>Sensoren</b>	<b>Anschließend folgt die Liste der Sensoren, für die Werte ausgegeben werden sollen.</b> Es sind maximal 64 Sensoren möglich. Jeder Sensor kann zusätzliche Parameter erhalten, die dann unmittelbar nach dem Sensornamen im nächsten Parameter angegeben werden.
6.	<b>Weitere Parameter</b>	<b>Parameter zur Kontrolle der Wertebereiche:</b>  Definition erlaubter Werte mittels unterer und oberer Schranke für den Wertebereich Kontrolle der Interpolation: <ul style="list-style-type: none"> <li>• Zwischen Vorgänger und Nachfolger gemittelter Wert</li> <li>• Durchschnittliche Änderung des Werts pro Sekunde</li> <li>• Zuletzt gelieferter Wert, d.h. Fortschreiben des Vorgängers</li> <li>• Obere Schranke für den erlaubten zeitlichen Abstand zwischen Vorgänger und Nachfolger</li> <li>• Mathematische Transformationen der Werte, z.B. Z-Transformation, um Werte auf eine Standardnormalverteilung zu skalieren</li> </ul>

-----  
Datenaufbereitung  
-----

Im Folgenden wird zur Veranschaulichung ein Beispiel für einen Aufruf der Funktion gezeigt:

```
select datetime, num_values, tag_1, tag_2
from table(fit_tags('uba', '2013-01-18', '+2h', 'step 5m',
'D06844UBAU1A422_020P20_05EX01.PValue', 'max_dist_sec=1',
'D06844UBAU1A422_010M77_01SR01.PV', 'max_dist_sec=2h30m',
p_default_param=>'round=2'
))
```

In diesem Beispiel werden zwei Sensoren (AKS "D06844UBAU1A422\_020P20\_05EX01.PValue" und „D06844UBAU1A422\_010M77\_01SR01.PV“) zeitlich aufeinander angeglichen. Die Ergebnistabelle beinhaltet die Werte beider Sensoren für einen Zeitraum zwischen 00.00 Uhr und 02.00 Uhr morgens, bei einer Schrittweite von jeweils 5 Minuten. Die Werte der Sensoren werden interpoliert, sofern es keine zur Schrittweite passend liegenden originalen Werte gibt. Dabei gelten für beide Sensoren jeweils unterschiedliche Einschränkungen bezüglich der „oberen Schranken für den erlaubten zeitlichen Abstand zwischen Vorgänger und Nachfolger“.

Im Folgenden erfolgt eine Übersicht der aus externen Quellen stammenden Softwareprodukte und Tools.

### 2.6.2.1 Tableau

Ein wesentliches Tool zur Unterstützung der Explorationsphase und Visualisierung von (Zwischen-) Ergebnissen ist Tableau. Es wird hauptsächlich im Bereich der visuellen Datenanalyse verwendet und vereinfacht die Datenexploration mithilfe von interaktiven Filtern, die es ermöglichen, die zu betrachtende Datenmenge auf den aktuell relevanten Teil zu reduzieren und somit schneller und flexibler ein besseres Verständnis der Daten zu bekommen. Parameter verschiedener Analyseverfahren können mithilfe von Steuerelementen interaktiv konfiguriert werden. Sind diese Verfahren dabei direkt in Tableau eingebunden, lassen sich die Ergebnisse der entsprechenden Parameterwahl unmittelbar visualisieren.

Die zu analysierenden Daten können in Textdateien oder Datenbanken vorliegen. Dabei spielt es keine Rolle, ob es sich um eine relationale oder nicht-relationale Datenbank handelt. Schließt man die Datenquelle direkt an, wird im sogenannten Live-Modus gearbeitet, was bedeutet, dass die jeweilige Tableau Arbeitsmappe die benötigten Daten mittels SQL-Anfrage aus der Datenbank importiert. Zusätzlich gibt es die Möglichkeit, indizierte Tableau Datenextrakte zu generieren, was die Performanz deutlich erhöht. Hierfür werden die Daten jedoch lokal dupliziert, was in der Gesamtbetrachtung die zu handhabende Datenmenge erhöht. Aus den geladenen Daten können Datenhierarchien aufgebaut werden, die ein Drill-Down und somit eine Datenabstraktion ermöglichen. In der aktuellsten Version von Tableau können gängige Analysesprachen wie R oder Python angebunden werden. Hierdurch können komplexeren Analyseschritte, welche nicht vorab in Tableau integriert sind, in den Workflow eingebunden und verfügbar gemacht werden. Die Ergebnisse werden wiederum an Tableau zurückgegeben und dort interaktiv visualisiert.

Die bereits erwähnten Features, sowie die Möglichkeit aus Visualisierungen ein Dashboard zu bauen und Tableau-Stories zu entwerfen, haben dafür gesorgt, dass Tableau vor allem Einzug in die Bereiche des „Monitoring und Reporting“ gefunden hat.

### 2.6.2.2 Python / R

Die beiden Programmiersprachen Python und R bieten eine große Vielfalt an Bibliotheken für statistische Berechnungen, was sie zum de facto Standard im Bereich der Datenanalyse und des Maschinellen Lernens gemacht hat. Daher wurde diese auch im Rahmen dieses Projektes verwendet. Es handelt sich in beiden Fällen um interpretierte Sprachen, weshalb Befehle Schritt für Schritt hintereinander ausgeführt werden können ohne das Programm nach jeder Änderung neu kompilieren zu müssen (Live Code). Die Ergebnisse jedes Befehles sind direkt – in Form von Zahlen oder Grafiken – sichtbar und können mit erklärendem Text versehen werden. Dabei können auch interaktive Elemente mit einbezogen werden. Das Ergebnis einer Analyse ist also ein Notebook, was als Report in verschiedenen Dateiformaten exportiert werden kann. Es stellt die gewählte Vorgehensweise unter Beibehaltung der Reihenfolge dar und beinhaltet somit auf natürliche Art einen roten Faden.



Nach dem die Daten aufbereitet wurden, gilt es über die Analysen im folgenden Kapitel zunächst es ein tiefer gehendes Verständnis für die Daten aufzubauen. Dies gilt sowohl für die Semantik einzelner Sensordaten, als auch für den Zusammenhang mehrerer Sensoren. Zudem wird die Plausibilität der Sensordaten anhand wichtiger Anlagenteile der GLT überprüft. Nur wenn diese gewährleistet ist, sind Daten grundsätzlich für tiefer gehende Analysen geeignet.

Das UBA verfügt über mehrere Anlagen, die zum Betrieb des Gebäudes eingesetzt werden. Hierzu zählen z.B. die Heizungs-, Lüftungs- oder auch die Solaranlage. Alle Anlagen werden über die GLT gesteuert und erzeugen folglich Datenpunkte, die zu Analysezielen ausgewertet werden können. Innerhalb des AKS werden über die Kostengruppen die einzelnen Anlagentypen kodiert. Eine Übersicht zeigt folgende Abbildung 27. Eine vollständige Auswertung aller Anlagen in jeglicher Detailfülle hätte den Rahmen dieses Projektes gesprengt. Aufgrund dessen wurde entschieden sich bei der Auswertung und Analyse zum Großteil auf bestimmte Anlagen zu konzentrieren. Hierzu zählen hauptsächlich die Heizungs- und Lüftungsanlagen, da diese eine zentrale Rolle im Betrieb des Gebäudes spielen.

Kostengruppe	Kostengruppen-Beschreibung
334	Außentüren und Außenfenster, RWA-Anlagen, Toranlagen
344	Innentüren und Innenfenster, RWA-Anlagen, Toranlagen
411	Abwasseranlagen
419	Abwasser-, Wasser-, Gasanlagen, Sonstiges
421	Wärmeerzeugungsanlage
422	Wärmeverteilnetze
423	Raumheizflächen
429	Wärmeversorgungsanlagen, Sonstiges
431	Lüftungsanlagen
432	Teilklimaanlagen
434	Presslufttechnische Anlage
435	Kälteanlagen
441	Hoch- und Mittelspannungsanlagen
442	Eigenstromversorgungsanlagen
443	Niederspannungsanlagen
444	Niederspannungsinstallationsanlagen
445	Beleuchtungsanlagen
456	Gefahrenmelde- und Alarmanlagen
461	Aufzugsanlage
480	nicht beschrieben im Pflichtenheft
489	Gebäudeautomation, Sonstiges
542	Wasseranlagen
935	nicht beschrieben im Pflichtenheft

Abbildung 27: Kodierung der Anlagentypen des UBA

Im Folgenden wird eine Auswahl an Auswertungen beschrieben, die zur Datenexploration und Überprüfung der Plausibilität durchgeführt wurden. Diese werden anhand einiger Beispiele aus dem Fernwärmeheizkreis aufgezeigt, der Teil der Heizungsanlage ist. Eine zentrale Rolle spielen dabei die Sensorwerte der jeweiligen Soll- und Ist-Temperaturen. Deren Zuordnung wird vorab im folgenden Kapitel beschrieben.

### 3.1 Automatisierte Zuordnungen von Soll- und Ist-Werten

Um allgemeine Aussagen über die systematischen Abweichungen zwischen Soll- und Ist-Werten treffen zu können, benötigen wir eine Zuordnung zwischen korrespondierenden Werten. Die „physikalische Bezeichnung der Funktion“ eines Datenpunkts (Stelle 28 des AKS) gibt darüber Auskunft, ob es sich um einen Sollwert (‘F’, für *Führungswert*) oder einen *Messwert* (‘X’) handelt. Außerdem wird darin die Messgröße codiert (z.B. ‘T’, für Temperatur oder ‘P’, für Druck).



Wir können mittels der Anlagennummer für jede Anlage zwischen Soll- und Ist-Werten unterscheiden. Im Allgemeinen gibt es eine 1:N- (bzw. M:N) Beziehung zwischen Soll- und zugehörigen Ist-Werten. Beispielsweise können einem Temperaturregler mehrere Temperaturfühler zugeordnet sein.

Durch Kombination der AKS-Teilinformationen „Anlagen-Nr.“ und „Kostengruppe“, sowie der „physikalischen Bezeichnung“ eines Datenpunkts können wir die Zuordnung von Soll- und Ist-Werten je Anlage und Kostengruppe einschränken. Dennoch verbleiben schlussendlich Mehrdeutigkeiten, die mittels manueller Durchsicht aufgelöst werden müssen. Dies wird exemplarisch für die „Statische Heizung“ am Beispiel der Temperatursensorik deutlich:

Anlage	Kostengruppe	Betriebsmittel	Betriebsmittelbezeichnung	Datenpunkt-Art	Messgröße
32	422	N50	Vorlauftemperatur Heizung nach Heizkennlinie	Führungswert	Temperatur
32	422	N51	Begrenzungsregler Rücklauftemperatur Heizung	Führungswert	Temperatur

Anlage	Kostengruppe	Betriebsmittel	Betriebsmittelbezeichnung	Datenpunkt-Art	Messgröße
32	422	B04	Außenlufttemperaturfühler	Messwert	Temperatur
32	422	B09	Vorlauftemperaturfühler	Messwert	Temperatur
32	422	B10	Rücklauftemperaturfühler	Messwert	Temperatur

Die korrekten Paarungen (N50, B09) und (N51, B10) lassen sich in diesem Fall nur mittels Vergleich der Betriebsmittelbezeichnung finden. Der Außenlufttemperaturfühler lässt sich als Ist-Wert keinem Soll-Wert zuordnen.

Das Beispiel für Anlage 32 (Statische Heizung) zeigt die Probleme, eine vollständig automatisierte Zuordnung von Ist- und Soll-Werten vorzunehmen. Es ist letztendlich eine manuelle Zuordnung von Nöten.

### 3.2 Exploration des Fernwärmeheizkreises

Wie im Eingangskapitel über das Umweltbundesamt beschrieben, sind dem Heizungssystem definierte Heizkurven hinterlegt, nach der die GLT eine Steuerung der Wärmemenge über die Ventile der jeweiligen Anlage steuert. Demzufolge müssen die Sensordaten diese hinterlegten Heizkurven abbilden. Um dies zu überprüfen, wurde aus den Sensordaten eine empirische Abbildung der Heizkurven durchgeführt. Als relevante Datenpunkte für die Analyse wurden die folgenden Datenpunkte identifiziert. Diese beziehen sich auf die Sensoren zur Steuerung der statischen Heizung und bilden die Datenlage für das Jahr 2014 ab. Dies ist das aktuellste Jahr, welches vollständig in den Daten abgebildet ist:

Temperatur Außen:	D06844UBAU1A422_032B04_01TX01
Temperatur SOLL:	D06844UBAU1A422_032N50_01TF01
	N050 = Vorlauftemperatur Heizung nach Heizkurve Ebene10: T=Temperatur, F=Führungsgröße/Sollwert
Temperatur IST:	D06844UBAU1A422_032B09_01TX01
	B009 = Vorlauftemperaturfühler
Ventil:	D06844UBAU1A422_032Y52_01GR1
	G=Abstand/Länge/Stellung R=Rückmeldung

*Anmerkung: Die Betriebskennnummer (Stelle 28 des AKS) gibt darüber Auskunft, ob es sich um einen Sollwert ('F', für Führungsgröße) oder einen Messwert ('X') handelt.*

Das Ergebnis der Analyse ist in der folgenden Abbildung dargestellt. Die X-Achse stellt dabei die Außentemperatur dar, während die Y-Achse die SOLL-Temperatur der statischen Heizung innen repräsentiert. Jeder Punkt stellt dabei einen Datenpunkt im genannten Zeitraum dar. Die farbliche Kodierung ist jeweils stellvertretend für das Quartal des in den Jahren 2014 und 2015, in dem der Datenpunkt aufgezeichnet wurde.

Im Ergebnis zeigen sich insgesamt vier Muster, welche die hinterlegten vier Heizkurven des UBA repräsentieren. Die beiden oberen und unteren Muster spiegeln jeweils die Heizkurve des inneren und äußeren Heizkreises wieder. Der Unterschied ist durch eine Umstellung der Heizgrenztemperatur entstanden. Diese wurden im betrachteten Zeitraum von 15°C auf 18 °C angehoben. Im Ergebnis konnten somit die erwarteten Heizkurven empirisch nachgebildet werden.

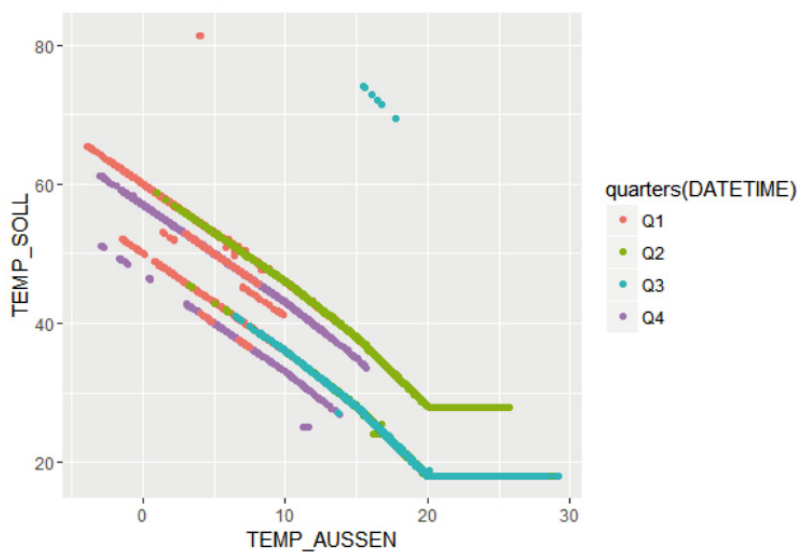


Abbildung 28: Empirische Heizkurven für den Zeitraum Nov. 2014 – Sep. 2015

### 3.2.1 Leistung des Fernwärmeheizkreises

Die Anlage beinhaltet einen Sensor (im Folgenden LFW-Sensor = Leistung Fernwärme genannt), der auf der höchsten Ebene misst, wann welche Verbräuche der gesamten Heizungsanlage des Gebäudes vorliegen. Da dieser Wert den zu reduzierenden Energieverbrauch repräsentiert, ist er besonders wichtig. Um also einerseits die Steuerung der Heizungsanlage zu verstehen und andererseits Muster in den Verbräuchen zu finden, wird im Folgenden näher auf das Verhalten dieses Sensors zu verschiedenen Zeitpunkten eingegangen. Da die Leistung der Fernwärme letztlich ein aggregierter Wert der Verbräuche der einzelnen Heizungsanlagen ist, wird daraufhin auf die Steuerung dieser einzelnen Anlagen eingegangen.

Das Verhalten des LFW-Sensors wurde wochenweise für verschiedene Zeiträume analysiert. Da für die Heizung nur die Zustände an / aus existieren war es das Ziel hierfür zwei Repräsentanten zu wählen, bei denen über die vollständige Woche dieser Zustand permanent gegeben ist. Im Herbst und Frühjahr schwanken die Temperaturen oft um die 15 °C, so dass dieser permanente Zustand innerhalb einer Woche nicht gewährleistet werden kann. Aufgrund dessen wurden hierfür Zeiträume im Winter und Sommer selektiert. Diese Wochen entsprechen jeweils einer Zeit innerhalb und außerhalb der Heizperiode.

### 3.2.1.1 Leistungswerte im Winter

Als Repräsentant für den Winter wurde die Woche vom 1. Dezember 2014 bis zum 7. Dezember 2014 gewählt. Die Abbildung 29 zeigt jeden Tagesverlauf einzeln, wobei jede Kurve ihrem Datum entsprechend eingefärbt ist.

Auf den ersten Blick erkennt man verschiedene Regelmäßigkeiten. Am auffälligsten sind die „Rampen“ jeweils um 06:00 und 18:00 Uhr. An den Wochentagen sind dies die Zeiten, zu denen über die Schaltung innerhalb der GLT die Heizung hoch- bzw. heruntergefahren wird. Beim Hochheizen steigt der momentane Verbrauch um das Dreifache an, woraufhin er nach kurzer Zeit wieder auf das Tagesniveau fällt. Dies entspricht ca. dem Doppelten des nächtlichen Verbrauchs. Bei der Betrachtung der unterschiedlichen Wochentage stellt man fest, dass Aufheizphasen einen „Überschwinger“ im Verbrauch aufweisen. Über den Arbeitstag hinweg bleibt das Leistungsniveau weitestgehend konstant und fällt um 18:00 Uhr rapide auf den Nachtverbrauch. Das Wochenende bleibt über den ganzen Tag hinweg bei knapp 200kW, bis auf einzelne Phasen, in denen die Heizungen kurz hochgefahren werden. Hier ist zu überlegen, ob eine Einstellung der Heizung existiert, die am Wochenende einen konstanten Verbrauch über den gesamten Tag gewährleisten kann. Diese Fragestellung hängt jedoch von vielen externen Faktoren ab und kann daher aktuell nicht beantwortet werden.

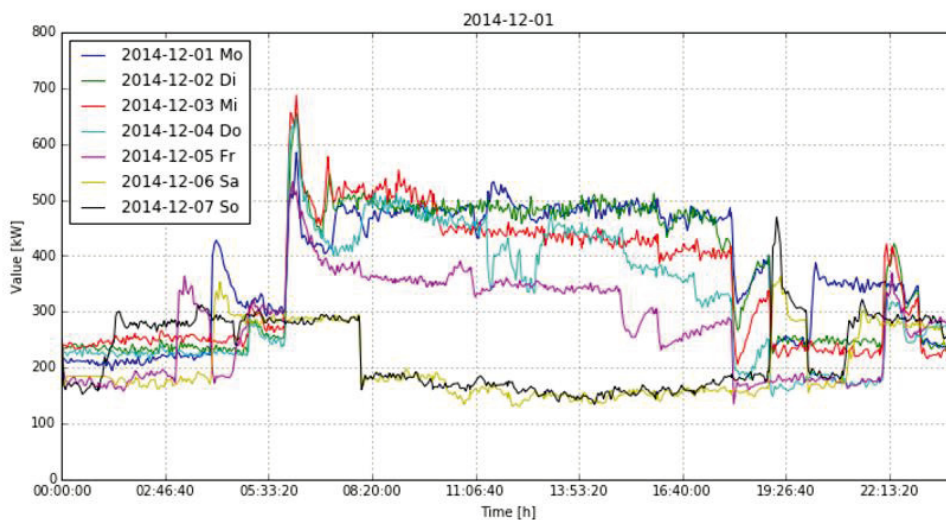


Abbildung 29: Leistung des Fernwärmekreislaufs in der ersten Dezemberwoche

Die pinke Kurve, die den Freitag repräsentiert, zeigt im Vergleich zu den übrigen Wochentagen einen niedrigeren Verbrauch, was sich bei der Betrachtung weiterer Wochen bestätigt. Außerdem ist um 15:00 Uhr an verschiedenen Freitagen ein Leistungsabfall erkennbar, was für ein partielles Herunterfahren der Heizung hin zum Wochenende spricht.

### 3.2.1.2 Leistungswerte im Sommer

Im Sommer wurde die erste Juliwoche 2015 betrachtet. Wie in Abbildung 30 zu sehen bewegt sich die Leistung der Fernwärme für diesen Zeitraum im Bereich unter 100kW und ist somit erwartungsgemäß niedrig. Von 0:00 bis 11:00 Uhr ist jedoch nahezu jeden Tag eine starke Oszillation zu erkennen. Dieses Muster ist unabhängig vom Wochentag und wurde im Rahmen dieses Projektes als nicht relevant erachtet, da die Verbräuche relativ klein ausfallen.

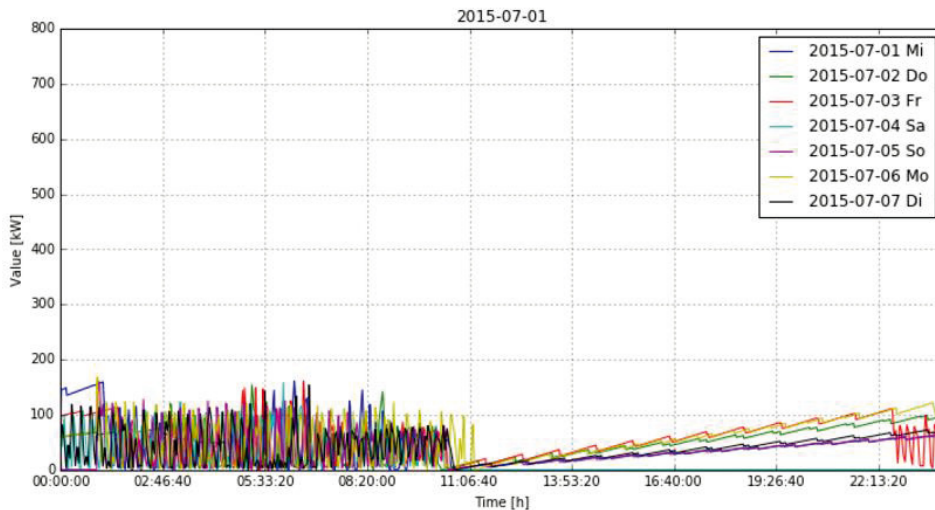


Abbildung 30: Leistung des Fernwärme-Kreislaufs in der ersten Juliwoche

Dasselbe gilt für die „zackenförmig“ ansteigende Leistung zwischen Mittag und Abend. Dieser Anstieg findet sich häufig (und ausschließlich) in der Sommerzeit und ein Rückschluss auf seine Ursache erschien nicht möglich, gerade weil der Zeitpunkt seines Auftretens sehr unterschiedlich war.

### 3.2.2 Steuerung einer Heizungsanlage

Für das Finden von Energieeinsparpotentialen wurde ein besonderer Fokus auf die Sensoren aus dem Bereich Heizung und Lüftung gelegt. Daher folgt hier zunächst eine Aufzählung der Heizungsanlagen des UBA, die im Rahmen der Analysen verwendet wurden.

- Statische Heizung (Heizkörper)
  - Anlage 031: Adsorptionskälte
  - Anlage 032: innerer Heizkreis
  - Anlage 033: äußerer Heizkreis
  - Anlage 034: Hörsaal
  - Anlage 035: Bibliothek
  - Anlage 041: Kantine
  - Wörlitzer Bahnhof
    - Anlage 600: Heizkörper
    - Anlage 601: Fußbodenheizung
- dynamische Heizung (Lüftung)
  - Anlage 042: Kantine
  - Anlage 036: Hauptgebäude

Der jährliche Wertebereich der Sollwerte eines Sensors lässt sich durch die Kennzahlen Minimum, Maximum, Mittelwert und Anzahl voneinander verschiedener Werte im jeweiligen Jahr zusammenfassen. Betrachtet man nun diese Kennzahlen für die vorliegenden Sollwerte (siehe Abbildung 31: Monatliche Kennzahlen eines Sollwert-Sensors von Januar 2014 bis Oktober 2015), so stellt man fest, dass die Anzahl voneinander verschiedener Werte von 2006 bis November 2014 sehr gering ist. Die Sollwerte sind teilweise über Tage hinweg konstant und lassen keine dynamische Anpassung an die Außentemperatur zu. Des Weiteren regelt das System die Ist-Temperatur in diesem Zeitraum nicht so, dass die vorliegende Soll-Temperatur erreicht wird. Stattdessen wird scheinbar eine nicht angegebene Soll-Temperatur verwendet.

Dieser Fall ist beispielhaft in Abbildung 32 zu sehen. Die grüne Kurve stellt die Ist-Temperatur dar, die blaue Kurve die Soll-Temperatur. Die Ist-Temperatur wird im typischen Arbeitszeitraum um etwa 10 Grad angehoben obwohl die Soll-Temperatur über den ganzen Tag hinweg konstant ist.

GRUPPE	TAGNAME	MONAT	GESAMT	DIST_VALUES	MIN_VALUE	AVG_VALUE	MAX_VALUE
heiz	D06844UBAU1A422_034N50_01TF01.PVeff	2014-01	21	1	50	50	50
heiz	D06844UBAU1A422_034N50_01TF01.PVeff	2014-03	1	1	50	50	50
heiz	D06844UBAU1A422_034N50_01TF01.PVeff	2014-04	1	1	50	50	50
heiz	D06844UBAU1A422_034N50_01TF01.PVeff	2014-05	10	1	50	50	50
heiz	D06844UBAU1A422_034N50_01TF01.PVeff	2014-06	2	1	50	50	50
heiz	D06844UBAU1A422_034N50_01TF01.PVeff	2014-07	2	1	50	50	50
heiz	D06844UBAU1A422_034N50_01TF01.PVeff	2014-08	1	1	50	50	50
heiz	D06844UBAU1A422_034N50_01TF01.PVeff	2014-09	6	1	50	50	50
heiz	D06844UBAU1A422_034N50_01TF01.PVeff	2014-10	4	1	50	50	50
heiz	D06844UBAU1A422_034N50_01TF01.PVeff	2014-11	51535	633	42.14120483	47.17	59.85243225
heiz	D06844UBAU1A422_034N50_01TF01.PVeff	2014-12	176462	769	42.85011673	47.74	61.1946106
heiz	D06844UBAU1A422_034N50_01TF01.PVeff	2015-01	171658	699	42.08333588	47.84	61.38062286
heiz	D06844UBAU1A422_034N50_01TF01.PVeff	2015-02	134171	461	43.67477036	48.29	52.84804916
heiz	D06844UBAU1A422_034N50_01TF01.PVeff	2015-03	154394	534	41.07060242	46.37	51.02926636
heiz	D06844UBAU1A422_034N50_01TF01.PVeff	2015-04	147043	884	40	44.97	58.57928467
heiz	D06844UBAU1A422_034N50_01TF01.PVeff	2015-05	105793	597	40	42.88	52.93692398
heiz	D06844UBAU1A422_034N50_01TF01.PVeff	2015-06	85484	345	40	42.05	55
heiz	D06844UBAU1A422_034N50_01TF01.PVeff	2015-07	63200	280	40	41.33	45.03472519
heiz	D06844UBAU1A422_034N50_01TF01.PVeff	2015-08	47287	249	40	41.24	44.47048569
heiz	D06844UBAU1A422_034N50_01TF01.PVeff	2015-09	142215	378	40	42.29	46.81423569
heiz	D06844UBAU1A422_034N50_01TF01.PVeff	2015-10	86584	641	0	45.02	58.40567398

Abbildung 31: Monatliche Kennzahlen eines Sollwert-Sensors von Januar 2014 bis Oktober 2015

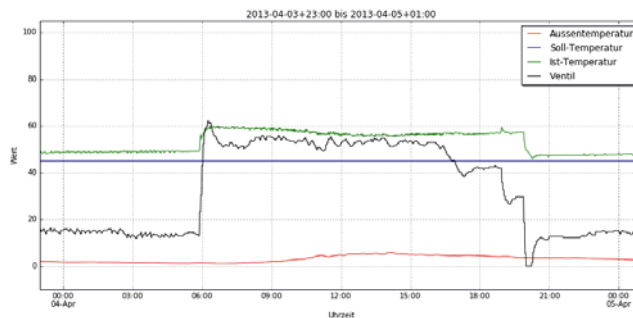


Abbildung 32: Dynamische Regelung des Ist-Werts trotz konstanten Sollwerts (Anfang April 2013)

Ab November 2014 werden die Werte dynamischer und die Kurven lassen Rückschlüsse auf den Zusammenhang zur Außentemperatur schließen. Das Ventil regelt die Ist-Temperatur so, dass die Differenz zur Soll-Temperatur gering ist und zeigt somit das erwartete Verhalten (siehe Abbildung 33).

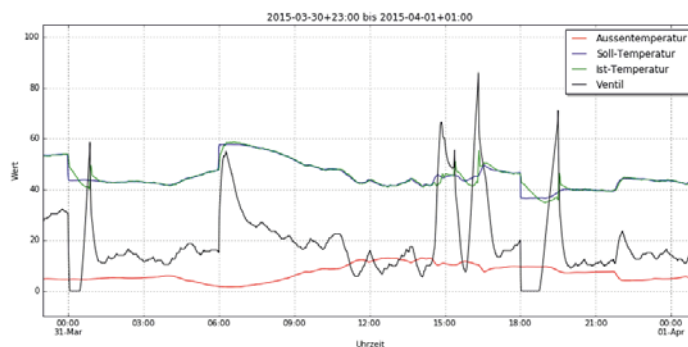


Abbildung 33: Dynamische Ist- und Sollwertregelung (Ende März 2015)

## 4.1 Einführung

Nachdem die Daten in den vorherigen Kapiteln zunächst aufbereitet und auf Ihre grundsätzliche Eignung für komplexeren Analysen validiert wurden, erfolgt in diesem Kapitel die Beschreibung des eigentlichen Analyseteils. Dabei wurde auf Analysen gesetzt, deren Ergebnisse dem Betreiber einer GLT als Entscheidungsgrundlage bei der Identifizierung von Energieeinsparpotentialen dienen können. Die vorgestellten Analysen stützen sich dabei zwar auf konkrete AKS des UBA. Für die Auswahl wurde sich jedoch vorab auf einzelne Sensoren bzw. Sensorgruppen gestützt. Diese sind grundsätzlich nicht UBA spezifisch und lassen sich auch in einer GLT anderer Gebäude finden.

Zur Unterstützung der Datenaufbereitung für die jeweilig Analyse wurde sich, vor allem bei Analysen unter Beteiligung mehrerer Sensoren, auf die `fit_tags` – Funktion gestützt (vgl. Kapitel 2.6.1.2). Diese ermöglicht die zeitliche Anpassung von Sensoren mit unterschiedlichem Zeitstempel. Ein wichtiger Parameter dieser Funktion ist u.a. der zeitliche Abstand zwischen zwei Ereignissen im jeweiligen Datensatz. Die jeweilige Größe dieses Abstandes ist maßgeblich für die zu verarbeitende Datenmenge, die Dauer und die Interpretationsmöglichkeiten der Analyse. Je nach zeitlicher Samplegröße und Betrachtungszeitraum nimmt die Datenmenge schnell an Größe zu (vgl. Abbildung 34).

	In Gigabyte			
	1 Sensor		4 Sensoren	
	1min Sample	1sec Sample	1min Sample	1sec Sample
1 Tag	0.00010368	0.0062208	0.00041472	0.0248832
1 Woche	0.00072576	0.0616896	0.00290304	0.1741824
1 Monat	0.00321408	0.1928448	0.01285632	0.7713792
1 Jahr	0.0378432	2.270592	0.1513728	9.082368

Abbildung 34: Datenmenge in Abhängigkeit der zeitlichen Samplegröße  
(Annahme: 72 Byte pro Eintrag inkl. Zeitstempel)

Für die folgenden Analysen hat sich oft ein Minutenbereich als praktikabler Zeitabstand herausgestellt. Diese Größe kann jedoch nicht verallgemeinert werden, da diese sehr individuell von der jeweils einzelnen Fragestellung abhängt. Im Allgemeinen kann als Anhaltspunkt der zeitliche Abstand in den Originaldaten als Maß herangezogen werden. Für eine Explorationsphase sind jedoch zeitlich Abstände zu empfehlen, die erste Analyseergebnisse in annehmbarer Zeit erlauben.

## Methoden

Grundsätzlich lassen sich Methoden zur analytischen Untersuchung von Sensordaten in unterschiedliche Ansätze einteilen. Diese lassen sich vor allem nach ihrer jeweiligen Komplexität unterscheiden. Eine gängige Unterscheidung und Eingliederung liefert ein in [7] präsentiertes Schema. Nach diesem Schema wurden die im vorliegenden Projekt dargestellten Analysevorhaben ausgesucht und eingeteilt. Ziel war es u.a. exemplarisch für die Ansätze der Analysen ein Beispiel zu liefern. Somit kann eine gewisse Bandbreite an Analytik in diesem Projekt aufgezeigt werden.



Im verwendeten Schema (vgl. Abbildung 35) steigt die Komplexität der Methoden von links nach rechts. Auf der linken Seite sind zunächst alle manuellen Methoden angesiedelt. Diese decken hauptsächlich das visuelle und deskriptive Spektrum an Analytik ab. Die regelbasierten Methoden überprüfen ergänzend, ob in den zu untersuchenden Daten bestimmte Grenzwerte oder Regeln eingehalten werden. Die modellbasierten Ansätze lassen sich wiederum in messdatenbasierte und simulationsbasierte Ansätze unterteilen. Bei den messdatenbasierten Ansätzen werden rein datenbasierte Modelle ohne tiefergehendes physikalisches Wissen über das jeweilige System entwickelt und angewendet. Aufgrund dessen werden diese Modelle auch als Black Box Modell verstanden. Sie führen relativ schnell zu Ergebnissen. Simulationsbasierte Ansätze versuchen dagegen das physikalische System zunächst in hoher Präzision digital abzubilden. Auf der Basis von physikalischen Gesetzen können anschließend Simulationen des Systems unter verschiedenen Bedingungen gefahren werden. Dieser Ansatz benötigt in der Regel wesentlich mehr Zeit um zu ersten Ergebnissen zu gelangen.

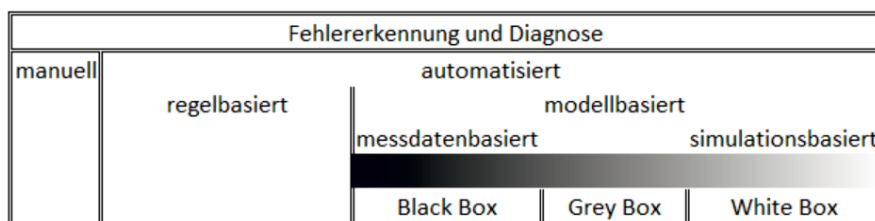


Abbildung 35: Übersicht Methodenansatz Analytik

Die im Folgenden vorgestellten Analysen sind hierarchisch in Anlehnung an das vorgestellte Schema angeordnet. Zu Beginn erfolgen rein deskriptive Analysen, gefolgt von einer regelbasierten Analyse von Schwingungen bis hin zur komplexen Analyse des Wettereinflusses auf den Energieverbrauch. Simulationsbasierte Ansätze sind aufgrund des hohen Zeitaufwandes nicht Teil dieses Projektes.

## 4.2 Fragestellungen

Auf Basis der Durchgeführten Meetings und Telefonkonferenzen (vgl. Kapitel Meetings / Telefonkonferenzen (Auswahl)) konnten mehrere Forschungsfragen herausgestellt werden. Diese bilden die grundlegende Basis für alle folgenden analytischen Schritte des Projektes. Nach der Definition der Fragestellungen wurden diese in die vorab dargestellten drei Teilbereiche nach Ihrer Komplexität eingeteilt. Eine konkrete Übersicht liefert die folgende Auflistung. Diese Liste ist auf Basis der Kenntnisse und Problemstellungen der GLT des UBA in Abstimmung mit den Fachverantwortlichen entwickelt worden, jedoch können diese Fragenstellen auch allgemein bei der analytischen Betrachtung anderen GLT-Systeme in Betracht gezogen werden können. Mit dieser Auswahl wird der gewünschten analytischen Breite des Projektes Rechnung getragen. Ziel bei der Auswahl der Analysen war es somit nicht, einzelne Analysen bis in die Tiefe durchzuführen, sondern eine möglichst breite Auswahl an möglichen Analysen aufzuzeigen.





Wie der Prozess der Erwärmung im Detail funktioniert ist in folgender Abbildung 38 anhand der Belüftungsanlage des UBA beispielhaft demonstriert. In diesem Fall wird ein Luftstrom bei 1 mit einer Temperatur von 9,9 °C in die Anlage eingeleitet. Über die Wärmerückgewinnungsanlage WR erfolgt zunächst eine Erwärmung der Luft auf eine Temperatur von 17,27 °C in 2. Erst danach nimmt die Heizung HZ Einfluss. Dieser erwärmt den bereits vorkonditionierten Luftstrom auf eine Temperatur von 22,5 °C in 3. Bis zur Verbrauchsstelle findet ergänzend eine Erwärmung der Luft auf 23,32 °C statt. Diese ist auf den Ventilator zur Aufrechterhaltung des Luftstroms zurückzuführen. Somit konnte das Erreichen der Solltemperatur erfolgreich durchgeführt werden.

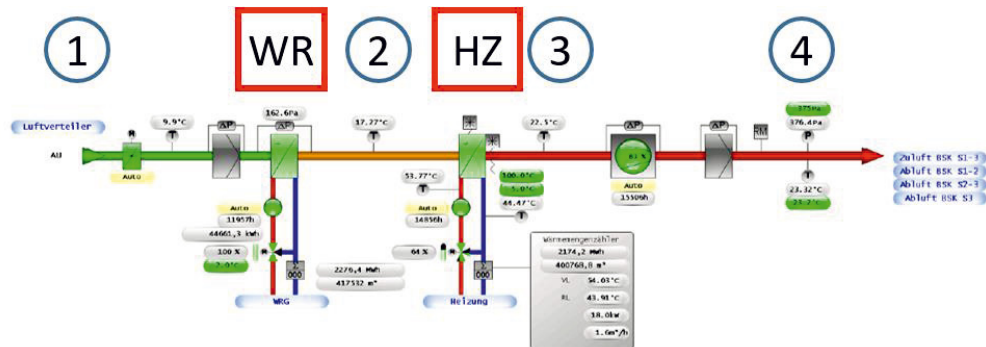


Abbildung 38: Heizung einer Lüftungsanlage

Für die Heizungsanlage des UBA wurde das Verhalten der Anlage bezüglich des Verhältnisses im Einsatz von Primärenergie und Wärmerückgewinnung untersucht. Im Zentrum stand dabei die Fragestellung, ob der Einsatz der Heizungsanlage immer erst dann geschieht, wenn die Energiemenge der Wärmerückgewinnungsanlage vollständig ausgeschöpft wurde. Konkret wurden dabei die Ventilstellungen der beiden Systeme Heizung und Wärmerückgewinnung für die Anlagen 010, 020, 030 und 040 miteinander verglichen. Die Analyse wurde dabei in mehrere Jahre von 2008 bis 2015 unterteilt, um ggf. Auffälligkeiten im zeitlichen Verlauf zu finden. Die einzelnen Datenpunkte wurden über die fit\_tags Funktion datenbankseitig auf ein zeitliches Intervall von 10 Minuten abgebildet und abgeglichen. Die Ergebnisse sind in der folgenden Abbildung komprimiert dargestellt.

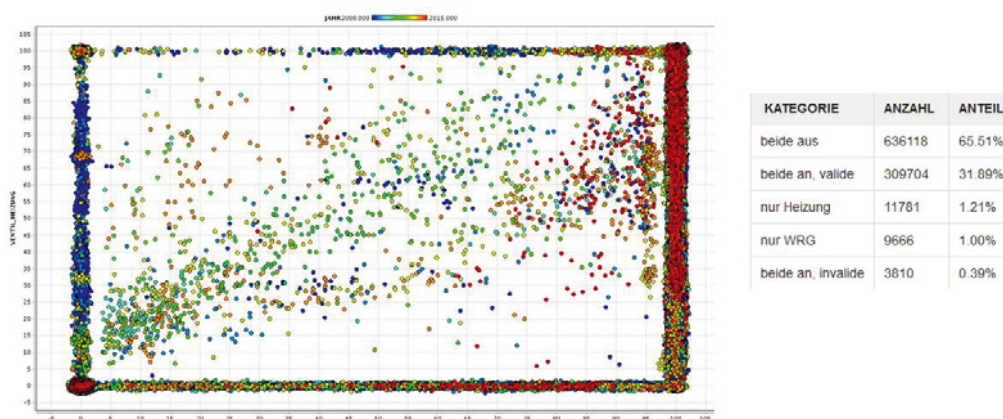


Abbildung 39: Vergleich Ventilstellungen Heizung und Wärmerückgewinnung

Auf der X-Achse wird in der Abbildung die Stellung des Ventils der Wärmerückgewinnung dargestellt, während sich die Y-Achse auf die Ventilstellung der Heizungsanlage bezieht. Der Null-Punkt der Darstellung befindet sich für beide Achsen am linken unteren Ende des Diagramms. Die farbliche Darstellung der Punkte bezieht sich dabei auf die jeweils einzelnen Jahre der Untersuchung (blau = 2008 – rot = 2015). Zur besseren Darstellung und Unterscheidung der Punkte wurden diese mit einem Jitter versehen. Somit können ggf. sich überschneidende Punkte visuell unterschieden werden.

Bei einer ideal eingestellten Anlage ist zu erwarten, dass sich alle Punkte entweder auf der X-Achse oder auf der Y-Achse im rechten Teil der Abbildung befinden. Dies bedeutet, dass entweder keine primäre Heizenergie genutzt wird, sofern die Wärmerückgewinnung noch nicht vollständig ausgenutzt wurde oder aber, dass die primäre Heizenergie erst dann hinzugeschaltet wird, wenn der Wärmegewinn durch die Wärmerückgewinnung für ein Erreichen der Solltemperatur nicht ausreicht. Alle anderen Punktepositionen deuten somit auf eine Verschwendung von primärer Heizenergie hin.

Im Ergebnis der Untersuchung kann festgehalten werden, dass in 1,6% der untersuchten Zeitintervalle vereinzelte Situationen auftreten, in denen das Potential der Wärmerückgewinnung nicht voll ausgenutzt wird. Dies sind solche Fälle, in denen die primäre Energiequelle entweder ausschließlich alleine (1.21%) oder in Kombination (0.39%) mit einer nicht zu 100% ausgelasteten Wärmerückgewinnung zum Einsatz kommt. Visuell betrachtet befinden sich diese Fälle alle außerhalb des im vorherigen Abschnitt beschriebenen Bereichs. Auffällig ist dabei, dass diese Fälle lediglich den jüngeren Jahren des Umweltbundesamtes zuzuordnen sind – erkennbar an der blauen Farbe.

Grundsätzlich lässt sich die Analyse in ihrer dargestellten Form auch auf Sensordaten anderer Gebäude bzw. GLT-Anlagen anwenden, da diese Analyse unabhängig von UBA-spezifischen Aufbereitungen oder Einschränkungen durchgeführt wurde. Voraussetzung ist dabei das Vorhandensein der genannten Sensortypen über einen gegebenen Zeitraum. Im Ergebnis liefert sie dem Betreiber einer GLT Anlage durch die kombinierte Darstellung von Lage und farblicher Kodierung der Datenpunkte eine schnelle Einschätzung der Lage hinsichtlich der Verschwendungspotentiale im Betrieb einer Heizung mit vorgeschalteter Wärmerückgewinnung.

#### **4.3.3 Differenztemperatur zwischen ein- und ausfließendem Wasser der Fernwärmeübergabe**

Im Allgemeinen muss ein Abnehmer von Fernwärmeenergie gewährleisten, dass zwischen primärer Energiequelle des Versorgers hin zum Abnehmer und der Energiequelle des Abnehmers hin zum Versorger eine definierte Differenz besteht. Somit soll gewährleistet werden, dass der Versorger stets als Quelle und der Abnehmer als Senke von Energie in der Gesamtbetrachtung des Systems funktioniert. Die Mengen und Zeiten sind in Form von Regeln mit dem Versorger vertraglich festgehalten. Ein Verstoß gegen die Regeln wird ggf. geahndet.

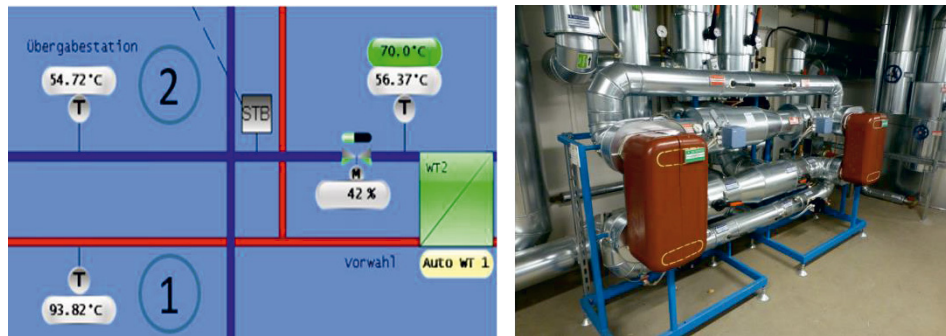


Abbildung 40: Differenztemperatur Fernwärmeübergabe + Wärmetauscher (eigene Aufnahme)

Im konkreten Fall des UBA wird das Einhalten dieser Vorgabe anhand der Differenztemperatur des ein- und ausfließenden Wassers an der Fernwärmeübergabe untersucht. Wie dies anhand des GLT-Abbildes aussieht, zeigt Abbildung 40 einen Teilbereich der Fernwärmeübergabe. Im Punkt 1 ist das einfließende Wasser aus dem Versorgernetz abgebildet. Dieses weist beim Eintritt in das Heizsystem des UBA eine aktuelle Temperatur von 93,82°C auf, bevor es auf den Wärmetauscher 2 (WT2) im rechten Bereich der Abbildung trifft. In Punkt 2 ist die Temperatur des zurückfließenden Wassers aus dem Heizsystem des UBA hin zum System des Versorgers abgebildet. Mit 54,72°C weist diese Temperatur eine um ca. 40°C deutlich niedrigere Temperatur auf. Somit ist eine Differenztemperatur gewährleistet. Ob dies historisch betrachtet auch für die Berücksichtigung von längeren Zeiträumen gilt, zeigt die folgende Analyse.

#### 4.3.3.1 Differenztemperatur für den gesamten Zeitraum

In einer ersten Analyse wurde die Verteilung der Temperaturdifferenzen von ein- und ausfließendem Wasser über den vollständigen Datenraum untersucht. Die Ergebnisse sind in der folgenden Abbildung 41 dargestellt. Als Datengrundlage dienen der Analyse die Jahre von 2008 bis 2015 in jeweils 10-Minuten-Schritten. Dabei werden jedes Mal die zwei Temperaturen mit demselben Zeitstempel verglichen. Die Dauer, die das Wasser benötigt, um den Heizkreis zu durchlaufen, ist hier nicht berücksichtigt.

Im Ergebnis zeigt sich, dass zu 87,44% die Temperatur des einfließenden Wassers vom Versorger mindestens genau so warm ist, wie das ausfließende Wasser nach dem Durchfluss durch das Heizsystem des UBA. In weiteren 11,61% ist die ausfließende Temperatur höchstens 2 Grad wärmer. In nur 0,95% der Fälle ist die ausfließende Temperatur um mehr als 2°C wärmer als die einfließende. Eingeschränkt auf den Normalfall, in dem das einfließende Wasser mindestens so warm wie das ausfließende ist, beträgt der durchschnittliche Temperaturverlust 25°C. Eine Temperaturdifferenz um die 0°C tritt am häufigsten auf.

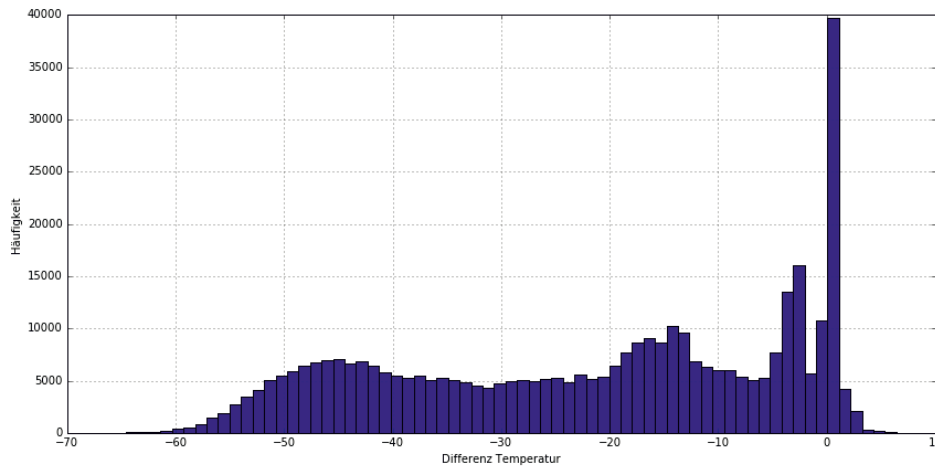


Abbildung 41: Verteilung der Differenztemperatur zwischen ein- und ausfließendem Wasser der Fernwärmeübergabe

Im Ergebnis der Analyse zeigt sich, dass ein geringer Prozentsatz an Fällen in den Daten existiert, bei dem eine Differenztemperatur zwischen ein- und ausfließendem Wasser vom UBA nicht gewährleistet werden kann. Dies muss jedoch nicht auf ein Problem der Energiesteuerung hindeuten, sondern kann auch im Zusammenhang mit den Heizperioden stehen. Aufgrund dessen wird in einer Folgeanalyse eine zeitliche Komponente in die Analyse integriert, um z.B. Unterschiede zwischen Heizperioden und Perioden, in denen die Heizung ausgeschaltet ist, zu identifizieren.

#### 4.3.3.2 Differenztemperatur zeitlich differenziert nach Monaten

Betrachtet man nun dieselben Histogramme aufgeschlüsselt nach den einzelnen Monaten (siehe Abbildung 42), so stellt man fest, dass sie je nach Jahreszeit unterschiedlich aussehen. Dies kann darauf zurückgeführt werden, dass das Heizverhalten stark von der Jahreszeit abhängt. Für die kalten Monate des Jahres liegt die Temperaturdifferenz zwischen ein- und ausfließendem Wasser deutlich im negativen Bereich und für die Sommermonate hauptsächlich im Bereich nahe der Null. Dies legt die Vermutung nahe, dass in den Sommermonaten nicht geheizt und somit nichts von der Energie der Fernwärme abgenommen wird. Um diesen Effekt näher zu analysieren, werden nun die Temperaturdifferenzen größer 0°C bzw. größer -10°C betrachtet, um die Monate zu identifizieren, aus denen die ungewünschten Beiträge resultieren.

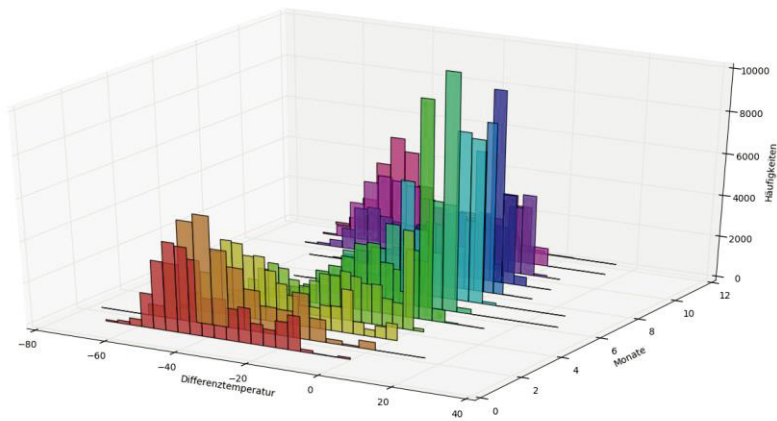


Abbildung 42: Differenztemperatur des Primärkreislaufs aufgeschlüsselt nach jedem Monat

Schließt man nun alle Differenzen kleiner als  $0^{\circ}\text{C}$  aus, so bleiben fast nur noch Beiträge aus den Monaten April bis Oktober übrig, also genau die Monate, in denen wenig bis gar nicht geheizt wird.

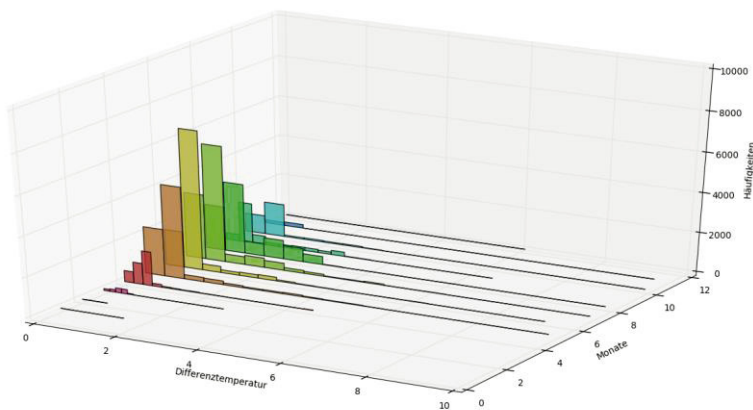


Abbildung 43: Monate, in denen die Differenztemperatur größer als  $0^{\circ}\text{C}$  ist

Dieses Bild ändert sich nur unwesentlich, wenn man den Bereich auf  $> -10^{\circ}\text{C}$  vergrößert. In dem neu hinzugekommenen Bereich liegen verhältnismäßig wenige Werte und wenn ebenfalls nur in den Monaten April bis Oktober.

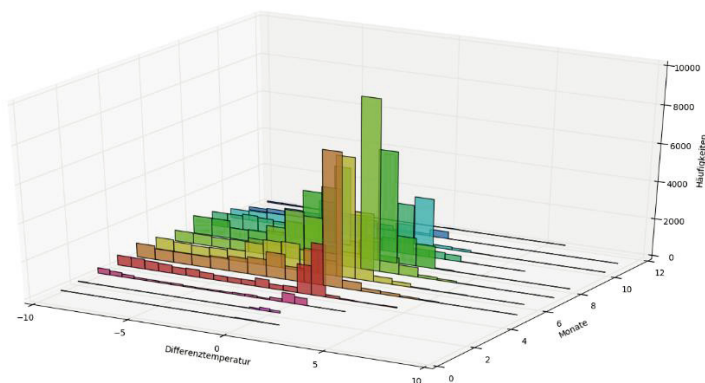


Abbildung 44: Monate, in denen die Differenztemperatur größer als  $-10^{\circ}\text{C}$  ist

Daher kann das Umweltbundesamt für die Wintermonate – und somit die Heizperiode – gewährleisten, dass das ausfließende Wasser 10°C kälter ist als das einfließende. Da im Sommer die Fernwärmeübergabe in der Regel nicht in Betrieb ist, spielt die geringere Differenz zwischen ein- und ausfließendem Wasser in diesen Monaten hier keine Rolle.

#### 4.3.4 Detektion von großen Differenzen zwischen Soll- und Ist-Werten

In der Heizungsanlage steuert die Ventilstellung die Differenz zwischen der gewünschten Temperatur (Soll) und der tatsächlichen Temperatur (Ist) und damit den Energieverbrauch. Die Ist-Temperatur soll durch Ventilöffnung oder -schließung so reguliert werden, dass sie sich der Soll-Temperatur nähert. Eine dauerhaft größere Differenz zwischen Soll- und Ist-Temperatur weist dementsprechend auf einen Fehler hin. Daher wurde für die Heizungsanlagen der Fernwärme ein Verfahren zur Detektion von Differenzen zwischen beiden Werten entwickelt. Für einen gegebenen Zeitraum (z. B. einen konkreten Tag) wird die mittlere Abweichung des Ist-Werts vom jeweiligen Soll-Wert bestimmt. Überschreitet diese einen vorher definierten Grenzwert (z.B. 2°C), so gilt dieses Verhalten als potentiell fehlerhaft.

Das entwickelte Verfahren wurde anschließend auf die Sensorwerte verschiedener Anlagen des UBA angewendet und ausgewertet. Im Ergebnis konnten teilweise große Differenzen zwischen den Soll- und Ist-Temperaturen entdeckt werden. Für die Analyse wurden die folgenden Anlagen im Referenzzeitraum Winter 2014/2015 berücksichtigt:

- Statische Heizung des UBA-Gebäudes:
  - Innerer Heizkreis: 032
  - Äußerer Heizkreis: 033
  - Bibliothek: 035 (statische Heizung)
- Statische Heizung des Wörlitzer Bahnhofs:
  - 600 (Heizkörper)
  - 601 (Fußbodenheizung)

Die verschiedenen Typen der Abweichungen werden in den folgenden Abschnitten vorgestellt.

##### 4.3.4.1 Abweichung aufgrund hoher Außentemperatur und Abschalten der Heizungsanlage

Das Verfahren detektiert in der Anlage 32 „statische Heizung (innen)“ am 10.04.2015 eine große mittlere Differenz zwischen Soll- und Ist-Temperatur. Der Verlauf der Kurven der Außen-, Soll- und Ist-Temperatur sowie der Ventilstellung ist für diesen Tag in Abbildung 45 zu sehen. Bis zur Mittagszeit liegen die gewünschte und die tatsächliche Temperatur nahe beieinander. Danach beginnen sie zu divergieren. Die Ist-Temperatur steigt leicht an und verweilt für den Rest des Tages auf dem Niveau von etwa 37°C. Die Soll-Temperatur fällt und bleibt über zwei Stunden auf einem Plateau von 29°C, was dem Grundsollwert entspricht. Um 15:00 Uhr fällt die Soll-Temperatur um weitere 10°C, da der 10.04.2015 ein Freitag ist und die Wochenendabsenkung in Kraft tritt.



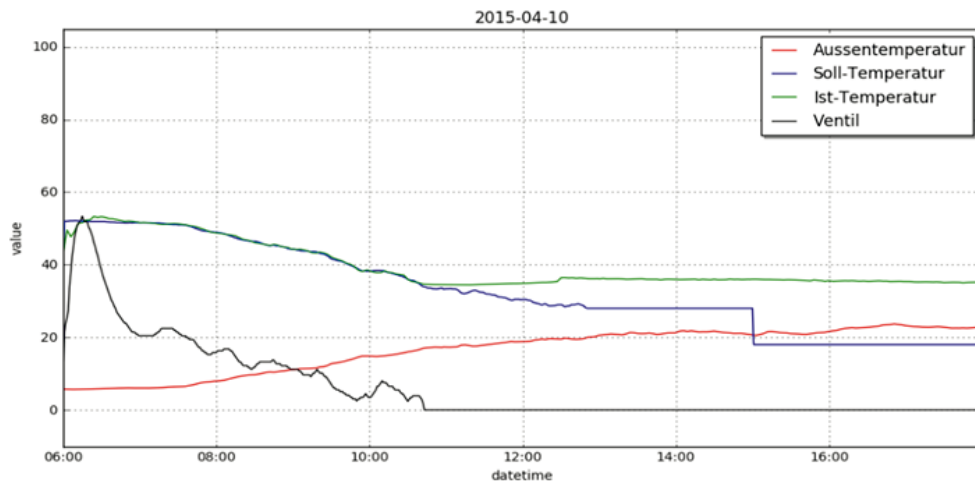


Abbildung 45: Soll- und Ist-Temperatur der Anlage 032 am 10.04.2015

Betrachtet man das Histogramm der Temperaturdifferenz (in der Abbildung 46 oben links) zwischen Soll- und Ist-Temperatur, so stellt man fest, dass die Peaks sich den verschiedenen Phasen des Tages zuordnen lassen. Diese Beziehung ist in der Grafik durch Pfeile dargestellt. Der Peak nahe der Null entspricht der validen Regelung bis zur Mittagszeit. Der zweite Peak bei etwa 8°C gibt das zwischenzeitliche Plateau der Soll-Temperatur bis 15:00 Uhr wieder. Der letzte Peak entspricht der Differenz ab der Wochenendabsenkung.

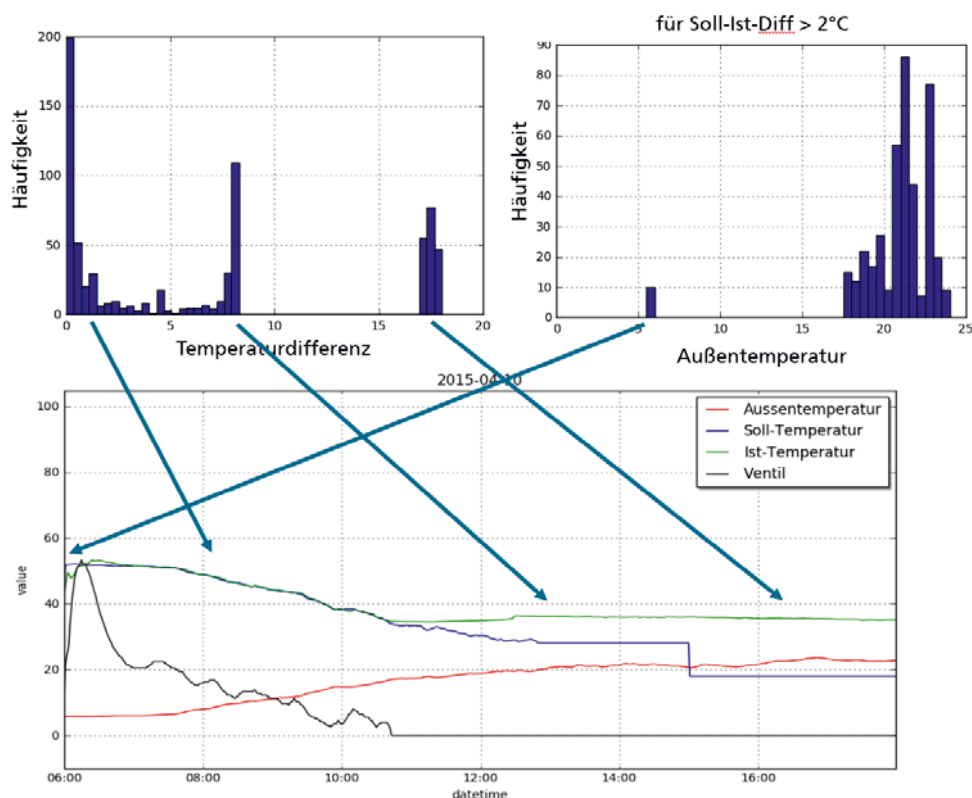


Abbildung 46: Histogramm der Soll-Ist-Temperaturdifferenz und der Außentemperatur für alle Zeitpunkte, an denen die Soll-Ist-Differenz größer als 2°C ist (Anlage 32 am 10.04.2015)



Das in Abbildung 46 oben rechts zu sehende Histogramm der Außentemperatur beinhaltet für alle Zeitpunkte, an denen die Soll-Ist-Differenz den vorgegebenen Schwellwert von 2°C überschreitet. Es zeigt eindeutig den Schwellwert, an dem die Heizung heruntergefahren und die Heizungsventile geschlossen werden. Dieser liegt zu dem gegebenen Zeitpunkt bei 18°C. Die Werte im Histogramm bei ca. 6°C werden von der Aufheizungsphase zu Beginn des Tages verursacht.

Folglich liegt kein Fehler in der Ventilsteuerung vor. Die Differenz kommt dadurch zustande, dass der Sensorwert der Soll-Temperatur aufgrund der Überschreitung des Schwellwertes der Außentemperatur beim Grundsollwert von 29°C liegt und die Ist-Temperatur diesen bereits mit geschlossenem Ventil überschreitet. Andere Beispiele haben gezeigt, dass sich das Ventil wieder öffnet, wenn die Ist-Temperatur unter die Soll-Temperatur fällt.

#### 4.3.4.2 Abweichung aufgrund von Schwingungen

Ein weiterer Abweichungstyp ist am selben Datum in Anlage 600 festzustellen. Dort beginnt die Abweichung ebenfalls zur Mittagszeit, jedoch damit, dass die Ist-Temperatur anfängt, um die Soll-Temperatur periodisch zu schwingen. Die Amplitude beträgt dabei um die 5°C. Diese Schwingung deutet für diesen Fall auf eine ineffiziente Ventilsteuerung oder auf ein ungünstig dimensioniertes Ventil hin. Dies wird auch als Schwingung bezeichnet. Dieses Phänomen wird im Abschnitt „Automatisiert Detektion von Schwingungen“ näher erläutert. In der vorliegenden Heizperiode tritt dies ausschließlich bei der Heizungsanlage 600 auf und dort hauptsächlich, wenn das Ventil weniger als 15% geöffnet ist. Da das Ventil genau dann wenig geöffnet ist, wenn die Außentemperatur nahe der Heizgrenztemperatur ist, scheint es einen Zusammenhang zwischen dem Auftreten von Schwingungen und der Außentemperatur zu geben.

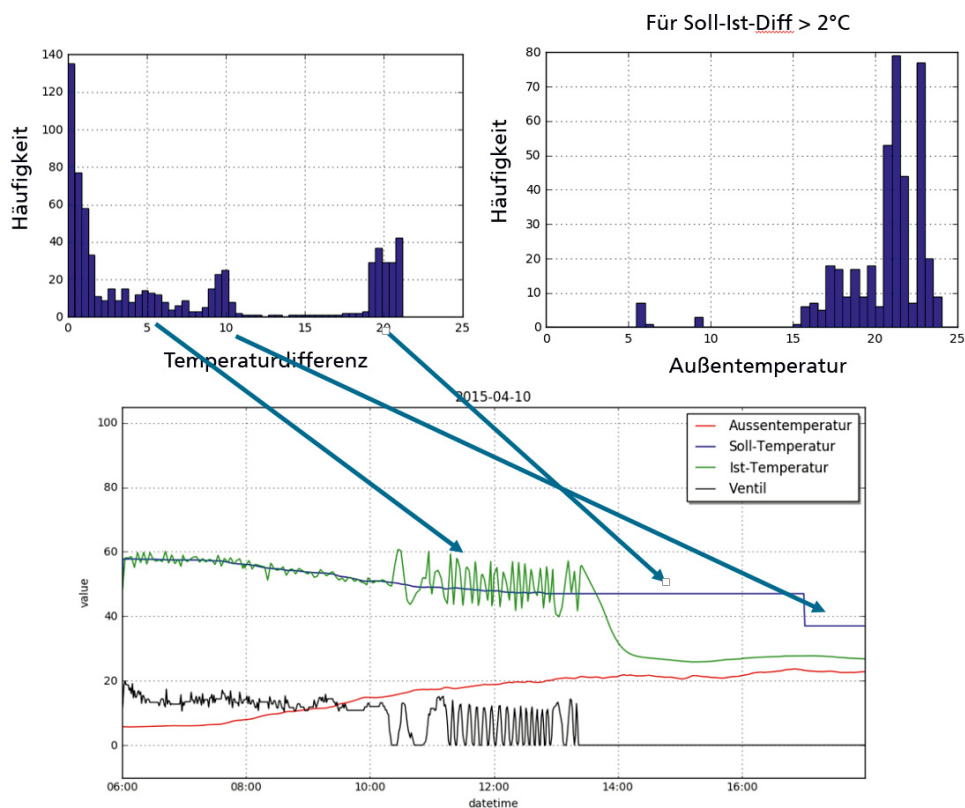


Abbildung 47: Darstellung aus Abbildung 46 für die Anlage 600 an demselben Tag

Betrachtet man nun wie beim vorigen Abweichtungstyp das Histogramm der Temperaturdifferenzen zwischen den Soll- und Ist-Werten, so ergibt sich bei den Schwingungen eine ähnliche Verteilung wie bei der Abweichung wegen des Abschaltens der Heizung. Das Histogramm der Außentemperatur für die Zeitpunkte, an denen der gewünschte und tatsächliche Wert stark voneinander abweichen, zeigt jedoch auch Werte im Bereich unter der Heizgrenztemperatur von 18°C. Dies kommt in dem vorherigen Fall nicht vor, da die Soll- und Ist-Temperaturen nur dann voneinander abweichen, wenn die Heizgrenztemperatur überschritten wurde. Somit kann anhand des Histogramms der Außentemperatur der Abweichtungstyp „Überschreitung der Heizgrenztemperatur“ vom Typ „Schwingung“ unterschieden werden. Im Abschnitt „Automatisierte Detektion von Schwingungen“ wird ein allgemeines Verfahren vorgestellt, das durch seine Parametrisierung in der Lage ist, automatisiert Schwingungen zu erkennen.

### 4.3.5 Analyse der Ventilstellungen

#### 4.3.5.1 Vergleich der Ventilstellung der Heizungsanlage

Wie im vorherigen Kapitel erläutert können Ventile als Ursache für Schwingungen angesehen werden. Im Verlauf der Analyse wurde dies bisher an einem einzelnen Ventil der Anlage 600 demonstriert. Im Folgenden wird sich der Herausforderung gestellt, solche Schwingungen systematisch in den Daten zu finden. Vorab dieser Analyse wird eine Meta-Übersicht über die Ventile erstellt. Diese Analyse bezieht sich wiederum auf die Ventile der zuvor definierten Heizungsanlagen. Eine wichtige Fragestellung ist dabei, ob Ventile existieren, die dauerhaft extreme Werte anzeigen.

In der folgenden Abbildung sind zunächst die durchschnittlichen Ventilstellungen pro Tag für die Heizungsanlagen des UBA für die Heizperiode des Winters 2014 / 2015 dargestellt. Die Darstellung erfolgt differenziert nach der jeweiligen Anlage (031 – 601). Während auf der X-Achse der Zeitraum pro Anlage von Mitte November 2014 bis zum April 2015 erfolgt, zeigt die Y-Achse die jeweilige prozentuale Ventilstellung von 0 bis 100%.

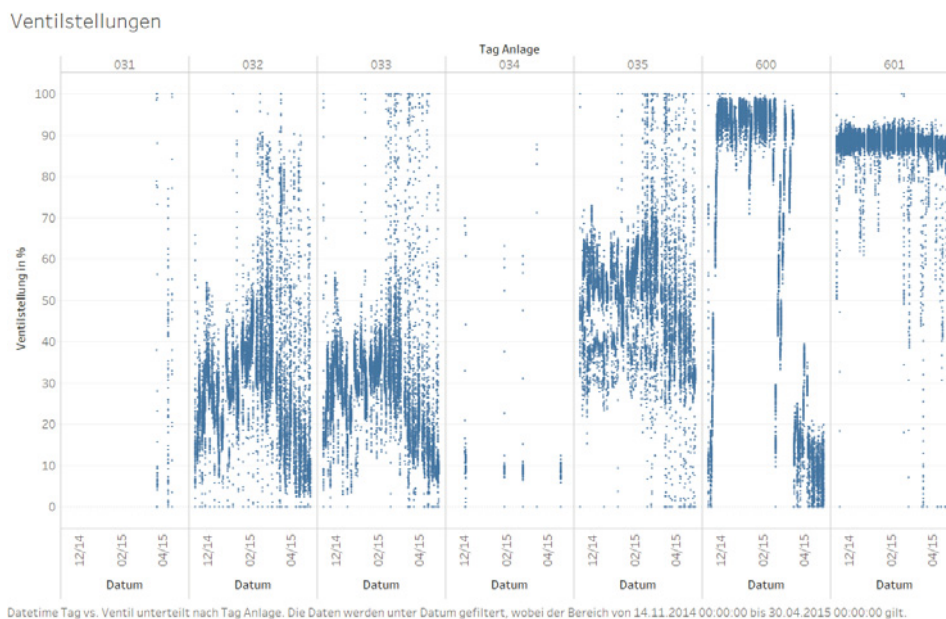


Abbildung 48: Prozentuale Ventilstellungen der Heizungsanlagen 2014 / 2015

Die Anlagen 031 und 034 liefern für den untersuchten Zeitraum zu wenige Daten, um Aussagen über das Verhalten der Ventilstellungen treffen zu können. Die Anlagen 032, 033 und 035 zeigen ein eher typisches Verhalten mit höheren prozentualen Ventilstellungen während der Wintermonate. Im Frühjahr nimmt die prozentuale Ventilstellung im Schnitt wieder ab und weist zudem in dieser Übergangsperiode eine größere Schwankungsbreite auf. Die Ventile der Anlagen 600 und 601 weisen dagegen ein abweichendes Verhalten auf. Während Anlage 601, abgesehen von Ausreißern, eine fast konstant hohe prozentuale Ventilstellung mit wenig Dynamik aufweist, sind bei Anlage 600 zwei Extrema zu erkennen. In den Monaten der Heizperiode ist eine im Vergleich zu den anderen Ventilen relativ hohe prozentuale Ventilstellung nahe dem Maximum zu erkennen, während im Frühjahr in der Übergangszeit das Ventil eine relativ niedrige prozentuale Ventilstellung aufweist. Diese Auffälligkeiten spiegeln somit auch die unterschiedlichen Anlagen wieder. Während die Anlagen 031 bis 035 die großen Heizungsanlagen des UBA-Gebäudes repräsentieren, beziehen sich die Anlagen 600 und 601 auf die kleiner dimensionierten Heizungsanlagen des Bürogebäudes „Wörlitzer Bahnhof“.

Verlauf der Ventilstellungen

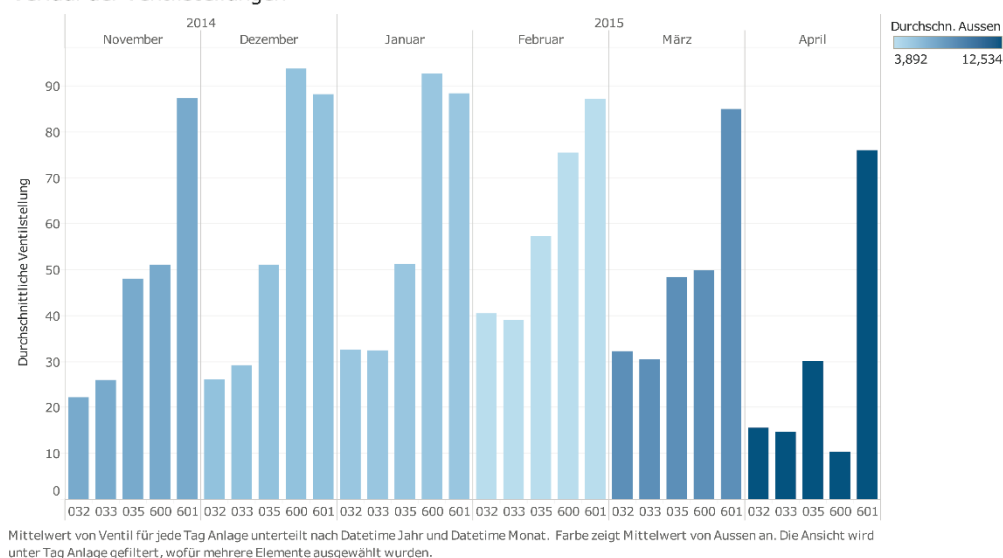


Abbildung 49: Zusammenhang zwischen Außentemperatur und Durchschnittlicher Ventilstellung pro Monat pro Anlage

Im vorigen Absatz wurden die Ventilstellungen einer Anlage über einen Zeitraum von 6 Monaten (November 2014 bis April 2015) betrachtet. Zusätzlich können auch die verschiedenen Anlagen für einen einzelnen Monat betrachtet werden. Dieser Vergleich über mehrere Monate vermittelt ein neues Bild und lässt es zu, dass als zusätzlicher Faktor die Außentemperatur mit einbezogen werden kann. Die Balken in Abbildung 49 spiegeln die durchschnittliche Ventilstellung der Anlage in diesem Monat wieder, wobei ihre Farbe der durchschnittlichen Außentemperatur entspricht.

In der Abbildung ist zu erkennen, dass es – wie zu erwarten – grundsätzlich einen Zusammenhang zwischen Außentemperatur und den Ventilstellungen gibt. Bei niedrigerer Temperatur sind die Ventile weiter geöffnet als bei höherer Temperatur. Dieser Effekt erscheint jedoch im Wesentlichen linear. Das heißt, die Verhältnisse der Anlagen zueinander ändern sich nicht stark, was dafürspricht, dass alle Anlagen ähnlich

auf unterschiedliche Temperaturen reagieren. Hierbei könnte eine vergleichbare Isolation eine Rolle spielen. Dabei bildet die Anlage 600 eine Ausnahme für diese Regel, da ihre Ventilstellungen viel stärker auf eine andere Außentemperatur reagieren. Im Vergleich mit den übrigen Anlagen kann sie abhängig von der Außentemperatur relativ am meisten oder am wenigsten anfordern. Für besonders kalte Zeiträume stellt dies ein Problem dar, weil die höchste Anforderung bestimmt, wie warm das Wasser im Vorlauf des Festwärmekreises aufgeheizt wird. Je heißer dieses Wasser jedoch ist, desto mehr Energie geht jedoch passiv z.B. beim Transport durch die Rohre verloren. Folglich kann eine einzelne Anlage den Energieverlust des gesamten Kreislaufs beeinflussen, weshalb solch ein schwächstes Glied in der Kette vermieden werden sollte.

Bisher wurde oft über die Dynamik bzw. Varianz von Ventilen gesprochen. Die bisherigen Visualisierungen stellen diese jedoch nur indirekt dar. Daher ist in Abbildung 50 ein Boxplot der Ventilstellungen der verschiedenen Anlagen pro Monat zu sehen. Dabei entspricht die hellgraue Box dem Bereich zwischen Median und 75%-Quantil und die dunkelgraue dem Bereich zwischen Median und 25%-Quantil. Die Größe der Box zeigt also ausgehend vom Median die Schwankungsbreite in die jeweilige Richtung. Auffällig erscheint hier – wie zu erwarten war – ebenfalls die Anlage 600, die in den Wintermonaten Dezember und Januar sowie dem Januar wenig schwankt. Dies ist vermutlich darauf zurückzuführen, dass auch Außentemperatur über den Monat hinweg nicht so sehr schwankt wie in den Übergangsmonaten November, Februar und März. Folglich weist dieselbe Anlage in den genannten Übergangsmonaten die größten Varianzen überhaupt auf, was sich aus ebenjenen stärkeren Temperaturschwankungen ergibt.

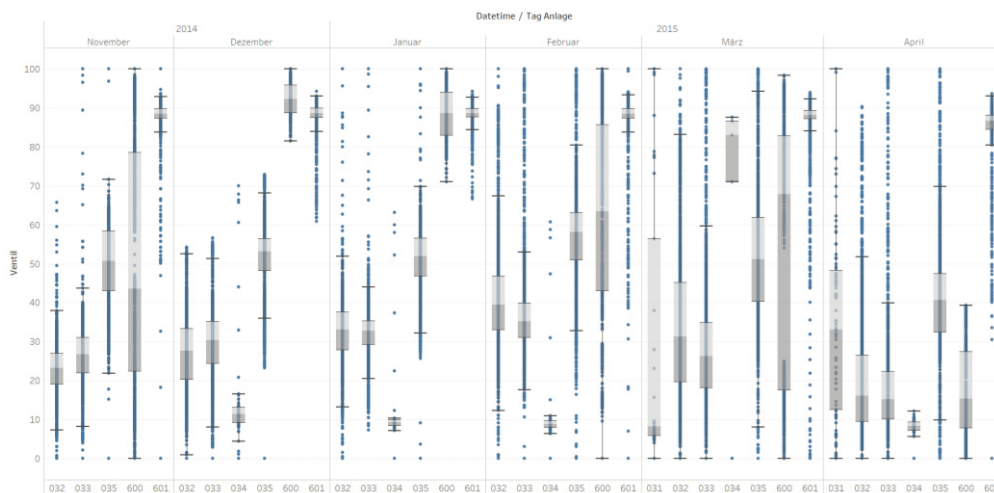


Abbildung 50: Varianz der Ventilstellungen pro Monat pro Anlage

Die Anlage 600 zeichnet sich durch eine dauerhaft niedrige Varianz aus. Für die übrigen Anlagen ist zu erkennen, dass ihre Varianz in den Frühlingsmonaten höher ist als in den Wintermonaten.

Analoge Darstellungen und Analysen lassen sich auch für Ventile anderer Anlagen durchführen. Dieses wurde exemplarisch mit einer Übersicht für die Ventile der Solaranlage des UBA durchgeführt. Das Ergebnis ist in der folgenden Abbildung zu erkennen. Hier sind farblich die prozentualen Ventilstellungen für das Jahr 2015 abgebildet. Dieses ist vergleichbar mit den Analyseergebnissen aus der Abbildung „Prozentuale Ventilstellungen der Heizungsanlagen 2014 / 2015“. Für das vorliegende Projekt wurde die Analyse nicht weiter vertieft, da sich in der Projektlaufzeit auf die Heizungsanlagen konzentriert wurde.

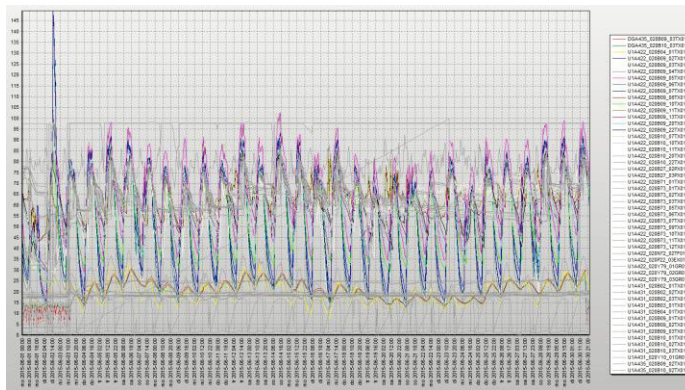


Abbildung 51: Ventilstellungen aller Ventile der Solaranlage

Die in diesem Abschnitt vorgestellten Betrachtungen zielten in erster Linie darauf ab, das Ventilhalten zu verstehen und gegebenenfalls Auffälligkeiten aufzudecken. So konnte erkannt werden, dass die Anlagen unterschiedlich auf externe Faktoren reagieren. Im Beispiel der Anlage 600 sorgen diese dafür, dass das Ventil entweder nahezu komplett geöffnet oder geschlossen ist. In diesem Fall kann es nicht weiter regulieren und erfüllt somit nicht mehr seine Aufgabe. Mithilfe der aufgezeigten Techniken können folglich Auffälligkeiten detektiert werden, indem dieselbe Datenmenge aus verschiedenen Blickwinkeln betrachtet wird. Dabei stellt jeder Blickwinkel eine zu erwartende Beziehung in den Daten dar, die es zu überprüfen gilt.

### 4.3.6 Automatisierte Detektion von Schwingungen

#### 4.3.6.1 Allgemeines

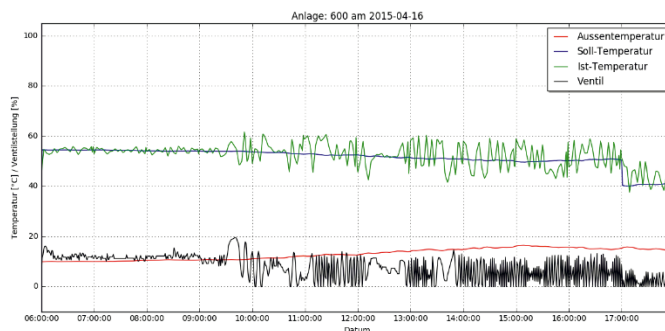


Abbildung 52: Typische Schwingung der Anlage 600

Wie bereits im vorherigen Abschnitt erwähnt und in Abbildung 47 zu sehen ist, stellen Schwingungen eine Möglichkeit der fehlerhaften Ventilsteuerung dar. In dem Fall, dass die Ist-Temperatur niedriger als die Soll-Temperatur ist, öffnet das Ventil zu weit, wodurch die Ist-Temperatur über die Zieltemperatur hinaussteigt und somit anschließend zu hoch ist. Daraufhin schließt das Ventil zu weit, weshalb die Temperatur wiederum zu tief fällt. Dieser Kreislauf setzt sich auf unbestimmte Zeit fort und benötigt trivialerweise mehr Energie, als die Beibehaltung der gewünschten Temperatur. Daher gilt es, solche Fälle im laufenden Gebäudebetrieb zu detektieren, um den unnötigen Energieverbrauch zu reduzieren.



Zuerst erscheint es in diesem Kontext intuitiv sich diesem Thema mit einer Fourier-Analyse auf Basis von Frequenzspektren und anschließender Zeitreihensegmentierung zu nähern. Dabei stößt man jedoch schnell auf das Problem, dass die vorliegenden Schwingungen in Frequenz und Amplitude sehr unregelmäßig sind und zusätzlich auch einen Versatz in y-Richtung haben können. Folglich wurde dieser Ansatz nicht weiterverfolgt und ein neues, einfacheres Verfahren zur Detektion solcher Schwingungen entwickelt, das im Bereich des Monitorings angewandt werden kann.

Für das Auftreten von Schwingungen gibt es verschiedene Ursachen. Die einfachste Ursache ist eine unpassende Parametrisierung in der Ventilsteuerung. Bei gegebener Differenz zwischen Soll- und Ist-Wert wird die Ventilstellung überreguliert, was in einer Schwingung resultiert. Eine andere Ursache liegt in der Größe des Ventils. Die Auflösung eines Ventils, ein Maß dafür, wie genau ein Ventil angesteuert werden kann, ist abhängig von seiner Größe und setzt somit voraus, dass das Ventil dem Kontext entsprechend richtig dimensioniert ist. Die Wahl der richtigen Ventilgröße ist jedoch schwierig und hängt von verschiedenen Parametern ab, die sich zusätzlich über die Zeit ändern können. Im Sinne der Regelfähigkeit ist vor allem wichtig, dass ein Ventil „Handlungsspielraum“ hat, sich also weiter öffnen bzw. schließen kann. Somit kann es auf äußere Einflüsse reagieren und ist möglichst selten komplett geöffnet oder geschlossen.

Des Weiteren beeinflussen sich die verschiedenen Ventile der Fernwärme gegenseitig. So sorgt beispielsweise ein zu klein gewähltes Ventil dafür, dass die Vorlauftemperatur des Sekundärkreislaufs höher wird als nötig, um selbst bei kleinem Ventil den Soll-Wert zu erreichen. Die höhere Vorlauftemperatur beeinflusst aber auch die anderen parallel geschalteten Anlagen, da das dort ankommende Wasser nun aufgrund der Anforderung des „schwächsten Glieds“ wärmer ist als mit einem richtig dimensionierten Ventil und die Ventile dieser Anlagen somit schließen.

Für die Schwingungen sind die zu groß gewählten Ventile von höherer Relevanz, da sie in der Regel eine schlechtere Auflösung haben als kleinere und außerdem immer nur zu einem Bruchteil geöffnet sind. Gerade in diesem Fall kommt es häufig zum Übersteuern bei der Regulierung und somit zu Schwingungen.

Der Bereich, in dem sich die Ventilstellungen eines Ventils bewegen, ist jedoch stark abhängig von externen Faktoren wie zum Beispiel der Außentemperatur. So sind trivialerweise im Winter alle Ventile weiter geöffnet als im Frühling und morgens weiter geöffnet als mittags. Aufgrund von äußeren Einflüssen - wie beispielsweise der unterschiedlichen Isolation einzelner Gebäudeteile - können sich auch die Verhältnisse der Ventilstellungen der Heizungsanlagen zueinander ändern. Daher ist es selbst im Nachhinein schwierig zu evaluieren, welche Ventile zu groß bzw. zu klein sind. Dafür muss entweder ein langer Zeitraum betrachtet werden oder ein offensichtlicher Fall vorliegen.

#### 4.3.6.2 Datenvorbereitung/-auswahl

Das entwickelte Verfahren stellt Anforderungen an die benötigten Daten, weswegen eine Auswahl betrachtet und diese für den Algorithmus vorbereitet wird.

Es werden nur Daten aus dem Zeitraum November 2014 bis April 2015 betrachtet, da erst ab November 2014 sinnvolle Soll-Werte vorliegen und im April 2015 die Heizperiode endet. Dabei werden nur Wochentage betrachtet, die keine Feiertage in Sachsen-Anhalt sind. Der relevante Zeitraum eines Tages reicht von 6:00:00 Uhr morgens bis 17:59:59 Uhr abends.

Für die Detektion von Schwingungen wurde eine Samplingrate von einer Minute verwendet. Für jeden Sensor liegt also jede Minute zeitgleich ein Wert vor. Dies entspricht 720 Werten am Tag. Da die Anlagensteuerung in erster Linie auf Soll-, Ist- und Außentemperatur sowie der Ventilstellung basiert und diese für das Verfahren

vorhanden sein müssen, werden nur Tage berücksichtigt, an denen für jeden dieser Sensoren 720 Werte vorhanden sind. Diese Tage werden *very good days* genannt. Für jede Heizungsanlage wurde die Anzahl von *very good days* bestimmt, was folgendes Ergebnis hatte:

Anlage	Funktion	very good days
031	Statische Heiz. innen	4
032	Statische Heiz. außen	104
033	Statische Heiz. Hörsaal	104
034	Statische Heiz. Bibliothek	6
035	Dynamische Heizung RLT	104
600	Wörlitzer Bahnhof Heizkörper	104
601	Wörlitzer Bahnhof Fußbodenheizung	104

Für fast alle Anlagen stehen 104 zu untersuchende Tag zur Verfügung. Die Anlagen 031 und 034 stellen nur 4 bzw. 6 „very good days“ und werden deshalb bei der Schwingungsdetektion nicht berücksichtigt.

#### 4.3.6.3 Algorithmus

Der folgende Abschnitt beschäftigt sich mit der Beschreibung des Detektionsverfahrens. Der Algorithmus identifiziert Schwingungen der Ist-Temperatur. Dabei sind die detektierten Schwingungen stark abhängig von den verwendeten Parametern. Dadurch ist das Verfahren sehr flexibel und kann mit geeigneter Parametrisierung auf verschiedene Heizungsanlagen angewandt werden. Einstellbare Parameter sind

- der Schwellwert für die Amplitude (z.B. 3,5°C)
- die minimale Anzahl gefundener Schwinger (z.B. 5)
- der maximale Abstand zwischen zwei Schwingern (z.B. 25min)
- die maximale Dauer eines Ausschlags beim Schwingen (z.B. 45min)
- die minimale Standardabweichung des Ventils beim Schwingen (z.B. 4%)

Die Werte in Klammern wurden für die Detektion von Schwingungen in den statischen Heizungsanlagen des UBA verwendet. In Abbildung 53 ist ein Beispielschwingung mit den entsprechenden Schranken für die Amplitude eingezeichnet.

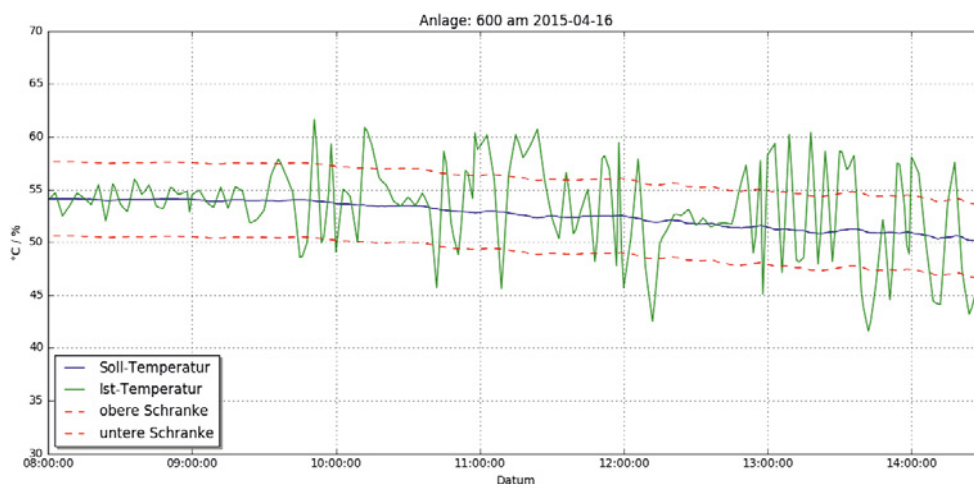


Abbildung 53: Temperaturschranken der Ist-Temperatur um die Soll-Temperatur



Wie auch an den Parametern zu sehen ist, benötigt das Verfahren die Zeitreihen der Ist- und Außentemperatur sowie der Ventilstellung in dem zu betrachtenden Zeitraum. Die Detektion betrachtet dabei jeden Tag von jeder Anlage einzeln. Zu Beginn werden alle Ist-Werte herangezogen, die weiter als z.B. 3,5°C von der Soll-Temperatur entfernt sind, wobei direkt benachbarte Punkte zusammengefasst werden. Die entstehenden Punktmengen repräsentieren einen Ausschlag. Dabei wird darauf geachtet, dass die Dauer eines Ausschlags nicht länger als 45min beträgt. Schließlich werden benachbarte Ausschläge zusammengefasst, sofern diese beiden nicht länger als 25min auseinanderliegen. Wenn nun für eine Vereinigung die vorher definierte minimale Anzahl an Ausschlägen überschritten ist und die Standardabweichung der Ventilstellungen in diesem Zeitintervall einen vorgegebenen Wert übersteigt, liegt eine Schwingung vor.

#### 4.3.6.4 Evaluation und Anwendungsbeispiele

Im Folgenden sollen die mit dem Verfahren erzielten Ergebnisse auf Basis der vorigen Datenauswahl dargestellt, evaluiert und interpretiert werden.

In Abbildung 54 ist ein typischer Stellvertreter der detektierten Schwingungen zu sehen. Die Ist-Temperatur schwingt – mit kurzer Unterbrechung – über sieben Stunden hinweg.

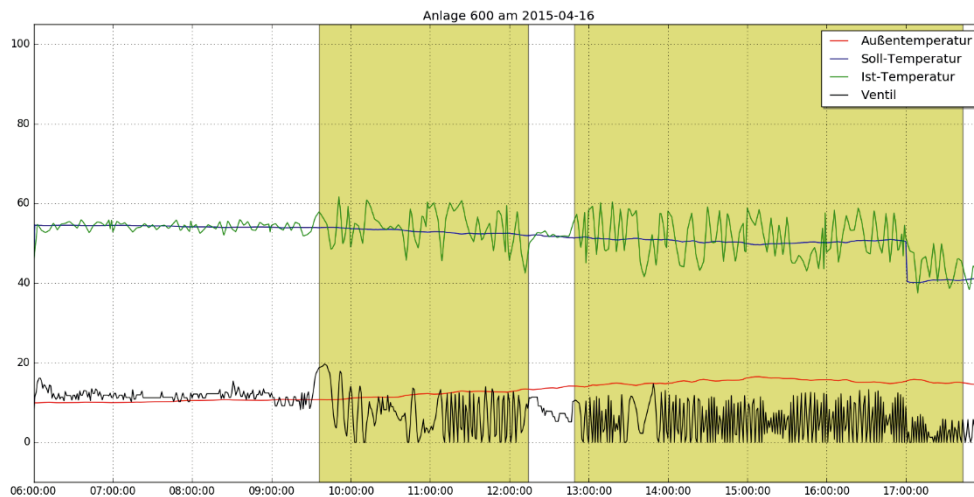
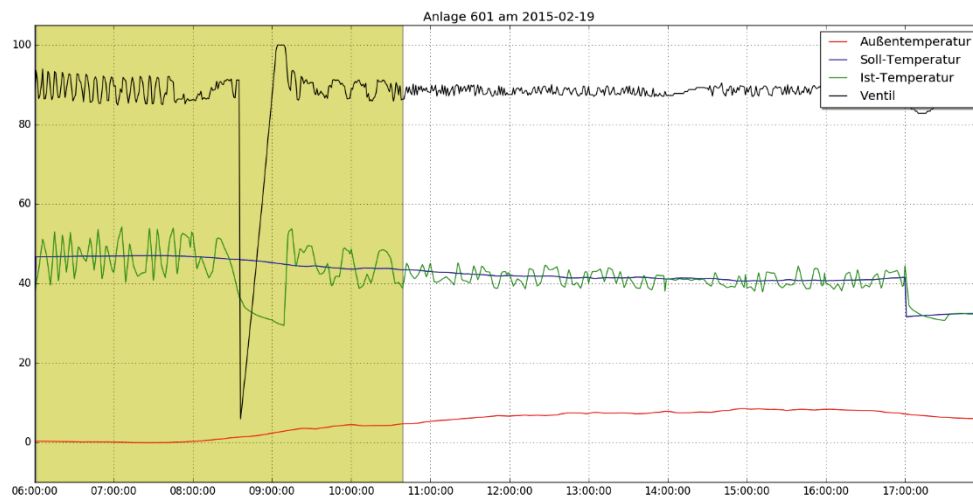


Abbildung 54: Detektierte Schwingungen der Anlage 600 bei fast geschlossenem Ventil

Währenddessen schwingt das Ventil zwischen 0 und 10% und zeigt somit nur Dynamik im unteren Bereich. Dies ist darauf zurückzuführen, dass sich die Außentemperatur an diesem Tag an der Grenze zur Heizgrenztemperatur bewegt und die Heizung somit dauerhaft kurz vor dem Herunterfahren steht. Allgemein wurden ausschließlich Schwingungen in den Heizungsanlagen 600 und 601 detektiert, also denen des Wörlitzer Bahnhofs. Dabei weist die Anlage 600 fast nur Schwingungen im März und im April auf und dann auch nur an warmen Tagen, an welchen sich die Heiztemperatur nahe der Heizgrenztemperatur befindet und die Ventilstellung entsprechend niedrig ist. Folglich scheint die Ventilsteuerung in diesem Grenzbereich nicht richtig zu funktionieren. Bei Anlage 601 (Fußbodenheizung des Wörlitzer Bahnhofs) ist zu beobachten, dass sich das Ventil über den gesamten betrachteten Zeitraum im Bereich zwischen 90 und 100% bewegt und diese Anlage somit wesentlich die Vorlauf-Soll-Temperatur der Fernwärme bestimmt. Im Vergleich zur Anlage 600, in der an 16 Tagen Schwingungen festgestellt wurden, zeigt die Anlage 601 nur an einem Tag eine Schwingung, welche in Abbildung 55 zu sehen ist. Dieser tritt bei weit geöffnetem Ventil und einer verhältnismäßig niedrigen Außentemperatur auf. Da es sich jedoch um ein einmaliges Phänomen handelt, wurde dieser nicht weiter interpretiert.



Datenanalytik

Abbildung 55: Detektierte Schwingung der Anlage 601 bei weit geöffnetem Ventil

#### 4.3.7 Wetteranalyse

Die Klimatisierung bzw. Beheizung und Lüftung von Gebäuden ist grob betrachtet von zwei Faktoren abhängig. Dies sind die jeweils aktuelle Bewirtschaftung und die äußeren Wetterbedingungen. In Kombination sind diese für die Leistungsverbräuche eines Gebäudes verantwortlich. Aktuell reagieren GLT-Systeme auf beide Einflussfaktoren durch eine Regelung nach den jeweils vorliegenden IST-Bedingungen. Weitere Zeiträume fließen aktuell nicht in die Bewertung der Leistungsverbräuche ein. Für die Bewertung und Regelung der Leistungsverbräuche ist jedoch auch eine Betrachtung der zukünftigen Bedingungen förderlich. Mit dieser ist es möglich, vorhandene Energie effizienter zu nutzen. Dies ist beispielsweise dann der Fall, wenn nach einem warmen Sonnentag am Folgetag ein Kälteeinbruch zu erwarten ist und somit die bereits im Gebäude befindliche Energie statt nachts aus dem Gebäude zu transportieren im Gebäude gehalten wird. Aufgrund dessen ist zum Beispiel eine prädiktive Heizungssteuerung ein wichtiges Thema, das es im Rahmen einer Energiebewertung und –steuerung zu betrachten gilt.

Für das Gebäude des Umweltbundesamtes ergibt sich diese Anforderung nicht nur für den Außenbereich, welcher unmittelbar der Witterung ausgesetzt ist, sondern auch für den Innenbereich des Atriums und des Forums. Wie im Eingangskapitel über das UBA erläutert, sind im darüber liegenden Sheddach diverse Klappfenster integriert, welche zur Klimatisierung des Innenbereiches eingesetzt werden. Der Einsatz der Klappfenster erfolgt sowohl im Sommer als auch im Winter. Im Sommer erfolgt dies hauptsächlich zur Kühlung, im Winter dagegen zur Entfeuchtung. Diese ist notwendig, da sich aufgrund der verwendeten Verglasung unter bestimmten Witterungsbedingungen Kondenswasser im Bereich der Fenster bilden kann. Dieses führt zu Wasserschäden auf der innen liegenden Holzfassade in Verbindung mit der Brandschutzversiegelung (vgl. Abbildung 56). Somit würde auch die Bewirtschaftung des Innenteils des UBA von einer prädiktiven Heizungssteuerung profitieren.



Abbildung 56: Dachfensterkonstruktion des UBA (mittig) und Wasserschäden aufgrund Kondenswasser entlang der Holzvertäfelung im Atrium

Damit verbunden ist zunächst die Frage, welchen Einfluss verschiedene Wetterfaktoren auf den Leistungsverbrauch eines Gebäudes besitzen. Aktuell sieht die GLT des UBA jedoch lediglich eine Steuerung nach der Außentemperatur vor. Diese wird durch einen einzelnen Sensor auf dem Dach des UBA erfasst. Neben der Außentemperatur lassen sich dagegen noch weitere Faktoren, wie z.B. die Globalstrahlung, der Niederschlag oder die Bewölkung identifizieren, die ggf. Einfluss auf die klimatischen Bedingungen innerhalb eines Gebäudes besitzen. Deren Einflüsse gilt es im Folgenden zu quantifizieren.

### **Übergeordnete Fragestellung - Wie wirken sich einzelne Einflussfaktoren des Wetters auf den Energieverbrauch aus?**

Der Wärmeverbrauch des UBA wird - basierend auf der hinterlegten Heizkurve - in Abhängigkeit von der Außentemperatur gesteuert. Es ist jedoch anzunehmen, dass die an einem einzelnen Messpunkt bestimmte Außentemperatur die Umgebungsbedingungen, die thermisch auf das UBA von allen Seiten her einwirken, nur begrenzt abbilden kann. Beispielsweise hängt die tatsächliche Wärmebelastung der Mitarbeiter in den Büros weniger von der im Schatten gemessenen Außentemperatur ab, falls eine intensive Globalstrahlung auf die Außenseite der Fassade des UBA trifft.

In diesem Zusammenhang wird häufig von Verantwortlichen für die Gebäudesteuerung der Wunsch geäußert, die Regelung der Wärmeverbrauchsanlagen enger an die Witterungsbedingungen anzukoppeln. Beispielsweise könnte man die Steuerung der Anlagen anhand der Wetterprognose für die kommenden Tage ausrichten, um den Energieverbrauch an Tagen mit günstigen klimatischen Verhältnissen zu reduzieren. Diese so genannte „Prädiktive Steuerung anhand von Wetterprognosen“ wirkt auf den ersten Blick vielversprechend, ist aber in der technischen Umsetzung aufwendig. Bevor man sich der technischen Umsetzung widmet, gilt es zunächst zu untersuchen, ob sich die externen Wetterfaktoren nicht ohnehin schon auf den Wärmeverbrauch eines Gebäudes auswirken, denn für diese Annahme gibt es gute Gründe. Die Heizkörper in den Büroräumen sind mit Thermostaten ausgestattet. Wenn also zum Beispiel durch vermehrte Sonneneinstrahlung zusätzliche Energie in die Räume eingetragen wird, könnte aufgrund der Thermostatregelung eine negative Rückkopplung auf den Wärmeverbrauch stattfinden.

Ziel der folgenden Analyse ist es, den Einfluss von Witterungsfaktoren (wie z.B. Sonneneinstrahlung, Niederschlagsmenge, Windstärke bzw. -richtung) auf den Wärmeverbrauch des UBA zu untersuchen und zu quantifizieren. Dazu betrachten wir im Folgenden die historischen Aufzeichnungen der Wettersensorik und des Wärmeverbrauchs des UBA.

Zwei gegensätzliche Szenarien können das Resultat einer solchen Untersuchung sein:

- 1.) Es gibt in den historischen Daten signifikante Abhängigkeiten zwischen Wetterfaktoren und dem Wärmeverbrauch.
- 2.) Es gibt in den historischen Daten keinen erkennbaren Zusammenhang zwischen Wetterfaktoren und dem Wärmeverbrauch.

Sollten sich anhand der Daten signifikante Einflüsse der Wetterfaktoren auf den Wärmeverbrauch zeigen (Szenario 1), käme es auf eine Einzelbetrachtung der Faktoren an. Der Faktor Sonnenstrahlung beispielsweise sollte sich positiv auf die Energiebilanz auswirken, wohingegen Witterungseinflüsse durch Wind und Regen sich negativ auswirken können. Szenario 1 würde darauf hindeuten, dass die vorhandene Heizungssteuerung am UBA bereits eine Rückkopplung der thermischen Einflüsse externer Wetterfaktoren, z.B. über Thermostate, ermöglicht. In diesem Falle ist es also fraglich, ob sich der zusätzliche Aufwand für die Implementierung einer prädiktiven Steuerung basierend auf Echtzeitwetterdaten bzw. Wetterprognosen lohnen würde.

Szenario 2 bedeutet, dass sich die Wetterfaktoren nicht in den historischen Aufzeichnungen des Wärmeverbrauchs des UBAs niederschlagen. Dies wäre ein datengetriebenes Indiz dafür, dass die Thermostatsteuerung in den Räumen keinen ausschlaggebenden Effekt auf die Energiebilanz zeigt. Unter der Prämisse, dass der Energieeintrag durch Sonneneinstrahlung in das Gebäude aber idealerweise eine Verringerung des inneren Wärmeverbrauchs ermöglichen sollte, lohnt es sich in Szenario 2 über gezielte Maßnahmen zur Kopplung von Wetterfaktoren und der Heizungssteuerung nachzudenken. Beispielsweise durch eine automatisierte Regelung der Heizungsanlage in Abhängigkeit der Sensorik für die Sonneneinstrahlung, oder durch die Verwendung von Wetterprognosen bei der Ausgestaltung von Aufheiz- und Abkühlphasen des Gebäudes.

#### 4.3.7.1 Methodik

##### Explorative Datenanalyse

Im Rahmen der explorativen Datenanalyse geht es um die Auswahl der relevanten Sensordatenpunkte und eine erste Charakterisierung der Beziehungen zwischen einzelnen Faktoren, z.B. paarweise Korrelationen.

##### Regressionsmodelle

Von Interesse sind statistisch relevante und interpretierbare Zusammenhänge zwischen dem Wärmeverbrauch im UBA und verschiedener Messgrößen, wie z.B. Außentemperatur, Sonnenstrahlung, Niederschlagsmenge und Windstärke. Lineare Regressionsmodelle erfüllen diese Voraussetzungen und erlauben es insbesondere, die einzelnen Einflussfaktoren gegenseitig abzuwägen.

Ein lineares Regressionsmodell erklärt den Wert einer Zielvariablen (hier des Wärmeverbrauchs) anhand der gewichteten Summe einzelner Faktoren (in diesem Falle: in Abhängigkeit der Wetterfaktoren).

Die Gewichte der Faktoren können aus historischen Daten (Aufzeichnungen der Sensoren und des zugehörigen Wärmeverbrauchs) abgeleitet werden. Zwar lassen sich die Gewichte – wegen möglicher Korrelationen zwischen den Faktoren – nicht direkt interpretieren, jedoch kann die Bedeutung der Faktoren aus ihren so genannten P-Values abgeleitet werden. P-Values korrespondieren zu dem Signifikanzniveau der aus den Daten geschätzten Parameter. Als Gütemaß für das aus den historischen Daten abgeleitete Modell im Ganzen wird das so genannte Bestimmtheitsmaß (auch  $R^2$ -Wert genannt) verwendet. Das Bestimmtheitsmaß ist ein Wert zwischen 0 und 1.0. Mittels des Maßes können verschiedene Regressionsmodelle untereinander verglichen werden, wobei höhere Werte eine bessere Übereinstimmung zwischen den Vorhersagen des Modells und den Eingabedaten bedeuten.

Wegen der Trägheit des Heizungssystems in einem Gebäude ist hier nicht sinnvoll, den Wärmeverbrauch zu einem bestimmten Zeitpunkt als Funktion der aktuell vorliegenden Wetterbedingungen zu modellieren. Stattdessen wird der Ansatz verfolgt, die Zusammenhänge zwischen Durchschnittswerten (während unterschiedlicher Tageszeiten) zu betrachten.

Tabelle 3 gibt ein Beispiel für die Datensatzbeschreibung, die als Eingabe für die Analysen dieses Kapitels dient. Abbildung 57 veranschaulicht anhand des Außentemperatursensors, wie Durchschnittswerte für Tageszeitabschnitte berechnet wurden. Abhängig von der Art des Sensors kann anstatt des Durchschnittswerts auch die akkumulierte Summe der Werte eines Sensordatenpunkts während des Zeitabschnitts betrachtet werden, um einen aggregierten Messpunkt zu erhalten. Beispielsweise wurde hier der Wärmeverbrauchswert als akkumulierte Summe betrachtet.

Aus den Spalten der Tabelle lassen sich paarweise Korrelationen zwischen den einzelnen Faktoren berechnen, oder aber Regressionsmodelle ableiten. In dem konkreten Beispiel von Tabelle 3 würde man mittels eines Regressionsmodells die individuellen Zusammenhänge zwischen den beiden ersten Spalten (mittlere Außentemperatur, mittlere Sonnenstrahlung) und der dritten Spalte (Wärmemengenverbrauch als Zielvariable) analysieren können.

Tabelle 3: Datensatzbeschreibung (mit einer beispielhaften Sensorauswahl)

	Durchschnittliche Außentemperatur	Durchschnittliche Sonnenstrahlung	Wärmemengenverbrauch (akkumuliert)
<b>Montag, 01.01 (11-15 Uhr)</b>	8 °C	50 W/m <sup>2</sup>	20 Mw
...			
<b>Freitag, 05. 01 (11-15 Uhr)</b>	9 °C	320 W/m <sup>2</sup>	17.5 Mw
<b>Montag, 08. 01 (11-15 Uhr)</b>	7.5 °C	231 W/m <sup>2</sup>	15 Mw
...			

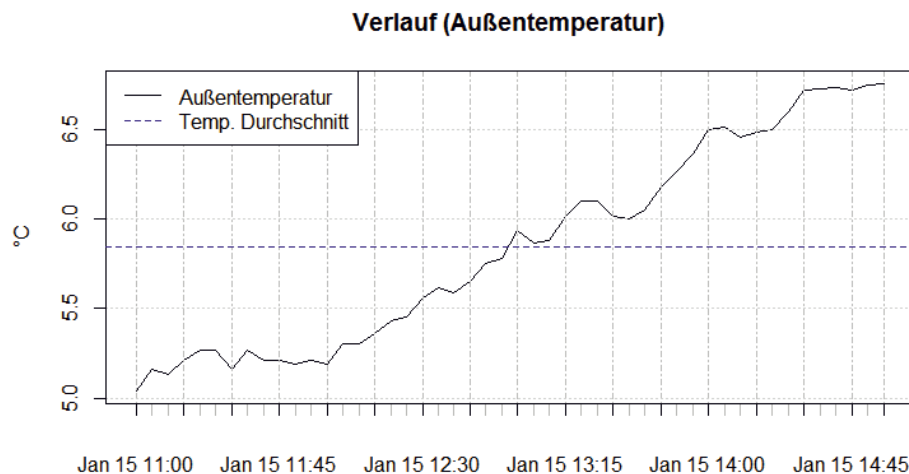


Abbildung 57: Temperaturverlauf (werktags, zwischen 11 und 15 Uhr)

### Auswahl des betrachteten Zeitabschnitts

Der stärkste Effekt auf den Gebäudewärmeverbrauch wird durch die Sonnenstrahlung vermutet, die auf Teile des Gebäudes trifft. Deren Effekt ist während der Mittagszeit am größten, wenn die Sonne am intensivsten scheint. Aus diesem Grund wurde die Datenauswahl für die nachfolgenden Analysen gemäß der Beschreibung in Tabelle 3 zunächst auf das Zeitfenster „werktags, zwischen 11 und 15 Uhr“ festgesetzt. Die Zeilen in Tabelle 3 fassen somit verschiedene aggregierte Sensorwerte für den Mittagszeitraum einzelner Werktage zusammen.

Um die Effekte einzelner Witterungsbedingungen auf den Wärmeverbrauch besonders gut untersuchen zu können, wurden die Tage in der Datenauswahl auf Tage beschränkt, deren mittlere Außentemperatur während der Mittagszeit zwischen 0° und maximal 15° lag (vergleiche zur Veranschaulichung Abbildung 57). Die Wahl dieser Schwellwerte dient der Einschränkung auf die Heizperioden während der „Übergangszeiten“ eines Jahres. Bei Temperaturen unter 0°C überwiegt der Kälteeinfluss der Außentemperatur sehr wahrscheinlich alle sonstigen Witterungseffekte. Außerdem ist die Betrachtung der Niederschlagsmenge bei Temperaturen unter 0° C schwierig. Bei Temperaturen über 15° C wird die Heizung im UBA deaktiviert.

Da es in der GLT-Steuerung des UBA in der Vergangenheit häufiger Wechsel in der Parametrisierung gab, wird den historisch jüngeren Datensätzen stärker vertraut als den zeitlich weiter zurückliegenden Sensordatenaufzeichnungen. Die Datenauswahl wurde daher auf die Jahre neueren Datums beschränkt, für die vollständige Aufzeichnungen vorlagen. Insgesamt wurde der 3-Jahreszeitraum der Jahre 2012 bis 2014 betrachtet.

Im Endeffekt umfasst das so ausgewählte und gefilterte Datenset insgesamt 353 Werktage, für einen Zeitraum von drei Jahren (2012—2014), an denen die durchschnittliche Außentemperatur im Bereich zwischen 0° und 15°C lag.

## Möglichkeiten zur späteren Erweiterung des gewählten Modellierungsansatzes

Der gewählte Modellierungsansatz bedingt, dass in Abhängigkeit der Tageszeit (vormittags, mittags, nachmittags, abends, nachts) bzw. gemäß der Unterscheidung von Werktagen und Wochenenden jeweils potenziell unterschiedliche Regressionsmodelle aus den Daten hergeleitet werden müssen. Um Vorhersagen auf neuen Daten zu machen, muss dementsprechend ein passendes Regressionsmodell ausgewählt werden.

In diesem Zusammenhang spielen so genannte „Model trees“ eine Rolle. Dies sind Meta-Modelle, welche die einzelnen Regressionsmodelle in die Struktur eines Entscheidungsbaumes einbetten (Abbildung 58). Auf diese Weise können die einzelnen Regressionsmodelle zu einem umfassenden Modell zusammengefasst werden.

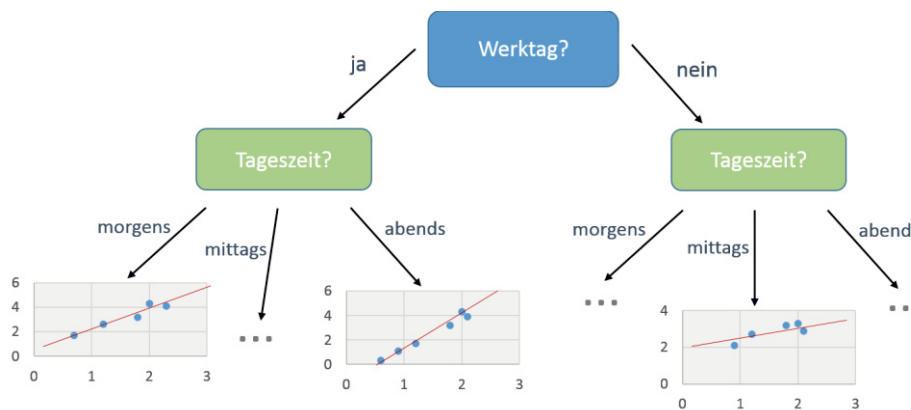


Abbildung 58: Model Tree für lineare Regressionsmodelle

### 4.3.7.2 Explorative Datenanalyse

#### Auswahl der Wetterfaktoren

Die Werte der Anlagen und einzelne Datenpunkt-AKS wurden hauptsächlich manuell dem "Wetterstationen"-Screen der GLT-GUI (Wonderware) entnommen (Abbildung 59).

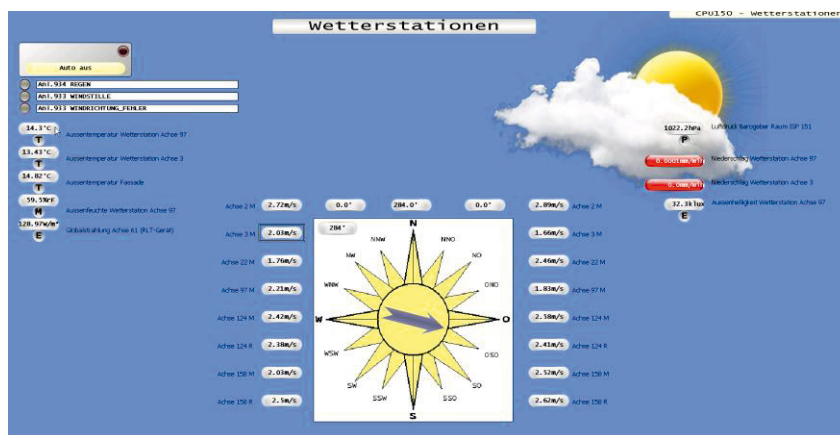


Abbildung 59: GLT-Ansicht der Daten der Wetterstationen



Die Sensordatenpunkte ließen sich grob in die Kategorien „Verbrauchsfaktoren“, „Temperaturfaktoren“, „Windfaktoren“, „Sonnenfaktoren“, „Referenzraumsensorik“ und „Sonstige“ (z.B. Außenfeuchte, Luftdruck, Niederschlag) einteilen. Insgesamt ergab sich eine Anzahl von weit mehr als 30 Sensoren.

Viele der untersuchten Sensoren innerhalb einzelner Kategorien liefern erwartungsgemäß nahezu identische Werte. Ihre paarweisen Korrelationen sind stark ausgeprägt. Dieses ist in Abbildung 60 abgebildet. Die in der Abbildung grün markierten Bereiche weisen auf die jeweilige Redundanz in der Temperatur- und Globalstrahlungssensorik hin. Die Färbung der Zellen gibt den Betrag der Korrelation an, d.h. rot gefärbte Zellen weisen auf eine ausgeprägte Korrelation (positiver oder negativer Art) hin. So gibt es beispielsweise mehrere Außentemperatursensoren und mehrere Sensoren für die globale Sonnenstrahlung in der Wetterstation und der Solaranlage des UBA. Gleiches gilt für die Wind- und Niederschlagssensorik. Zur Vereinfachung des Datensets wurden deshalb für die Faktoren „Außentemperatur“, „Sonnenstrahlung“, „Außenfeuchte“, „Niederschlagsmenge“, „Windrichtung“ und „Windgeschwindigkeit“ einzelne AKS ausgewählt.

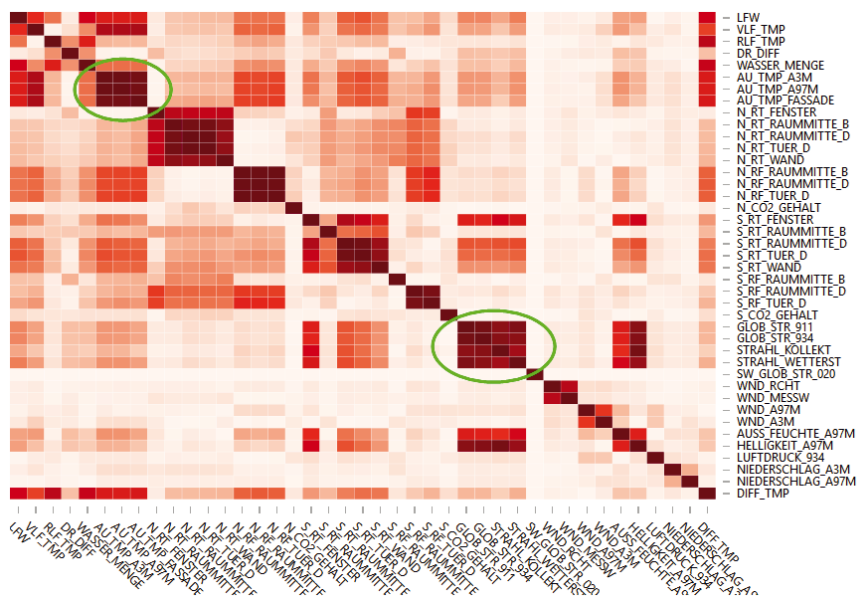


Abbildung 60: Heatmap paarweiser Korrelationen der untersuchten Witterungsfaktoren

Als Faktor zur Messung des Wärmeverbrauchs wurde gemäß seiner zentralen Bedeutung in der GLT-Verbrauchsübersicht der Verbrauchsmesser mit der AKS D06844UBAU1A422\_020P20\_05EX01 (LFW – „Leistungsmesser Fernwärme“) ausgewählt.

Tabelle 4 listet die getroffene Auswahl an Sensordatenpunkten auf, die als relevant betrachtet wurden, und gibt darüber Auskunft, wie die einzelnen Sensordatenpunkte über den jeweiligen Betrachtungszeitraum aggregiert wurden. Dabei wird zwischen Aufsummierung aller Werte während des Betrachtungszeitraums und einer Durchschnittsbildung unterschieden.

Tabelle 4: Übersicht Sensordatenpunkte

Kurzbezeichnung	Beschreibung laut GLT	AKS	Gemessene Einheit	Aggregationsform (Betrachtungszeitraum)	Bezeichnung (Aggregierter Wert)
LFW	Anl. 020; Leistung Fernwärme	D06844UBAU1A422_020P20_05EXD	kw	Kumulierte Summe	CUM_LFW
AU_TMP_A97M	Anl. 934; Außentemperatur (Achse A97M)	D06844UBADGA489_934B04_01TXD	°C	Durchschnitt	AVERAGE_AU_TMP
N_RT_RAUMMITTE_D	Deckentempertursensor (Raum Nordseite)	D06844UBAD1A489_916B01_13TXD	°C	Durchschnitt	AVERAGE_N_RT
S_RT_RAUMMITTE_D	Deckentempertursensor (Raum Südseite)	D06844UBA03A489_924B01_08TXD	°C	Durchschnitt	AVERAGE_S_RT
STRAHL_WETTERST	Anl. 020 Strahlung Wetterstation	D06844UBAU1A422_020B46_02EXD	W/m <sup>2</sup>	Durchschnitt	AVERAGE_STRAHLUNG
WND_RCHT	Anl. 933 Windrichtung	D06844UBAD1A489_933A08_12EXD	°	Durchschnitt	AVERAGE_WND_RCHT
WND_A97M	Anl. 934 Windsensor (Achse A97M)	D06844UBADGA489_934B76_08EXD	m/s	Durchschnitt	Siehe nachfolgender Wert „AVERAGE_WND“
WND_A3M	Anl. 934 Windsensor (Achse A3M)	D06844UBADGA489_934B76_07EXD	m/s	Durchschnitt	AVERAGE_WND (Es wird das jeweilige Maximum der
AUSS_FEUCHTE_A97M	Anl. 934 Aussenfeuchte (Achse A97M)	D06844UBADGA489_934B14_01MXD	%F	Durchschnitt	AVERAGE_AUSSENFEUCHTE
NIEDERSCHLAG_A97M	Anl. 934 Niederschlag (Achse A97M)	D06844UBADGA489_934B48_01EXD	mm/min	Kumulierte Summe	CUM_NIEDERSCHLAG

Datenanalytik

## Globale Zusammenhänge der Wetterfaktoren

Abbildung 61 stellt die gemittelten Jahresverläufe von Außentemperatur und Wärmeverbrauch des UBA über 10 Jahre im Vergleich dar. Der Wärmeverbrauch wurde für diese Abbildung als "Differenz zwischen primärer Vor- und Rücklauftemperatur" berechnet.

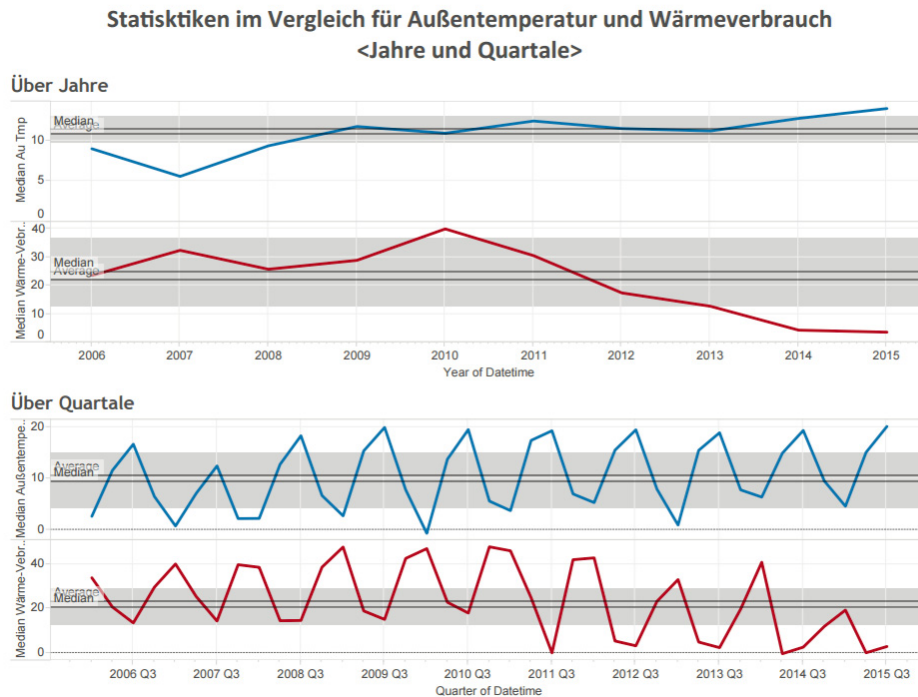


Abbildung 61: Anti-korreliertes Muster von Außentemperatur und Wärmeverbrauch

Auf Quartalsbasis sind deutlich die gegenläufigen Schwingungen zu erkennen, d.h. niedrige Außentemperaturen korrespondieren zu höheren Verbrauchswerten und umgekehrt.

Als nächstes betrachten wurde der gefilterte Referenzdatensatz für werktägliche Mittagszeiträume in den Jahren 2012 bis 2014 betrachtet. Erwartungsgemäß ergibt sich hier ebenfalls eine starke Antikorrelation (Pearson-Korrelationswert: -0.77, Pvalue < 0.001) für den Zusammenhang der mittleren Außentemperatur und des kumulierten Wärmeverbrauchs (Abbildung 62). In diesem Ergebnis spiegelt sich die Tatsache, dass der Wärmeverbrauch des UBA hauptsächlich in Abhängigkeit der gemessenen

Außentemperatur geregelt wird. Dennoch ist die Korrelation nicht perfekt, d.h. es gibt eine Reihe weiterer Einflussfaktoren. Dazu gehören beispielsweise die Gebäudeauslastung durch die Anwesenheit von Mitarbeitern in den Büros und andere Verbrauchseinheiten im Anlagensystem des UBA. Einen kleinen Anteil des Effekts könnten aber auch externe Witterungsbedingungen spielen, wie noch im Folgenden untersucht wird.

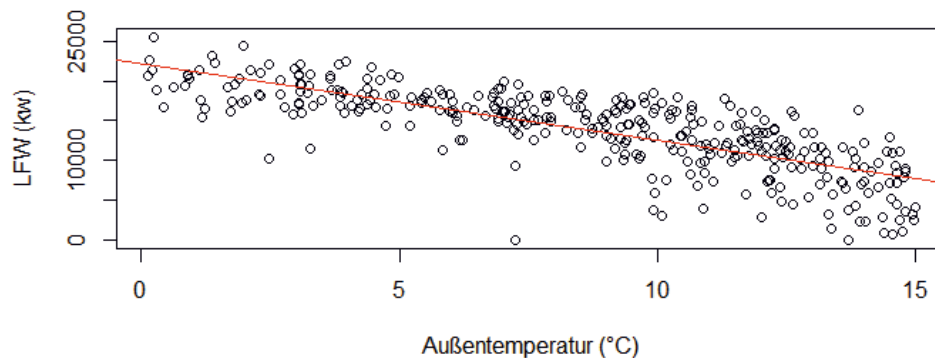


Abbildung 62: Anti-Korrelation von Außentemperatur und LFW (zur Mittagszeit)

Interessant ist auch ein Vergleich des Wärmeverbrauchs zwischen Sommer- und Wintermonaten für das UBA. Abbildung 63 stellt den mittleren Wärmeverbrauch im Vergleich zur Jahresdurchschnittstemperatur über 10 Jahre auf Quartalsbasis dar. Man kann sehen, dass in den beiden "Sommerquartalen" Q2 und Q3, die Verbräuche geringer sind als in den beiden „Winterquartalen“ Q1 und Q4. Im Verlauf der letzten 4-5 Jahre ist der Wärmeverbrauch des UBA während der Sommermonate stark vermindert worden (Abbildung 63).



Abbildung 63: Wärmeverbrauchsübersicht über 10 Jahre

### Korrelationen zwischen Wetterfaktoren und dem Wärmeverbrauch

Ein erster Ansatz, um die Beziehungen der Witterungsfaktoren untereinander – und in Bezug auf den Wärmeverbrauch (LFW) – zu charakterisieren, besteht darin, die paarweisen Pearson-Korrelationen der verschiedenen Faktoren zu betrachten, wie sie in der folgenden Heatmap-Grafik dargestellt sind (Abbildung 64). Die Färbung der Zellen gibt den Betrag der Korrelation an, d.h. rot gefärbte Zellen weisen auf eine ausgeprägte Korrelation (positiver oder negativer Art) hin, wohingegen helle Zellen auf eine lineare Unabhängigkeit der Faktoren hindeuten. Als Datengrundlage dienen hier erneut die aggregierten Werte (Werktage, Mittagszeitraum).

Die stärksten Korrelationen treten zwischen LFW und Außentemperatur (-0.77), sowie zwischen LFW und der Sonnenstrahlung (-0.51) auf. Die Korrelation zwischen Außentemperatur und Sonnenstrahlung liegt bei 0.39, ist also verhältnismäßig schwach. In Kombination gibt dies einen ersten Hinweis auf einen möglichen Einfluss der Sonnenstrahlung auf den Wärmeverbrauch. Interessant ist aber auch die schwach positive Korrelation von Außenfeuchte und Wärmeverbrauch (0.35). Der Faktor Außenfeuchte könnte also ebenfalls in Bezug auf den Wärmeverbrauch einen geringen Einfluss ausüben.

Bemerkenswerterweise gibt es nur eine sehr schwache Korrelation (0.2) zwischen Außenfeuchte und kumuliertem Niederschlag. Allgemein sind die Werte des Niederschlagsensors sehr klein und eher selten von Null verschieden, weshalb dieser Sensor als weniger verlässlich eingeschätzt werden kann.

Alle genannten Korrelationen sind signifikant (Pvalue < 0.001).

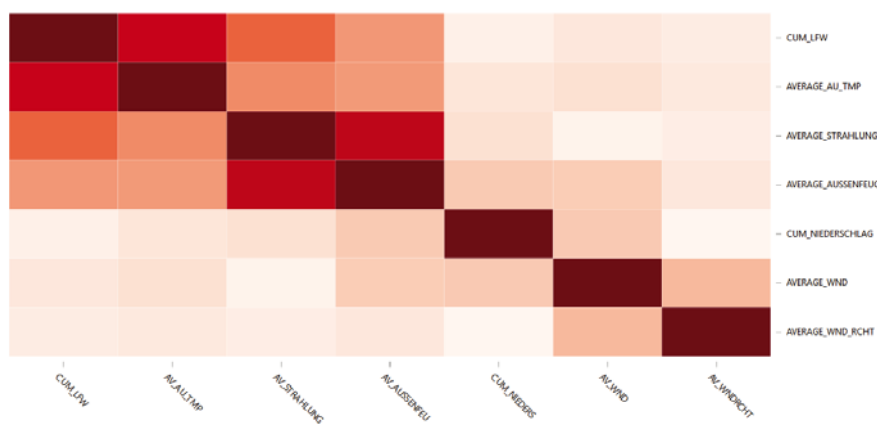


Abbildung 64: Heatmap Visualisierung paarweiser Pearson-Korrelationen

Insgesamt bedeuten die Ergebnisse eine Bestätigung der Annahmen, insofern als dass der Faktor Sonnenstrahlung einen positiven und der Faktor Außenfeuchte einen negativen Effekt auf die Verringerung des Wärmeverbrauchs zeigen.

### Einfluss der Sonnenstrahlung

Wie im vorherigen Abschnitt beschrieben, weisen die paarweisen Korrelationen auf einen besonderen Einfluss der Sonnenstrahlung auf den Wärmeverbrauch des UBA hin. Dies bestätigt diese Erwartungen, denn theoretisch sollte es aufgrund der Thermostatsteuerung der Heizkörper zu einer negativen Rückkopplung auf den Wärmeverbrauch kommen, wodurch sich der Energieverbrauch verringert.

Es ist davon auszugehen, dass die Wärmebelastung der Mitarbeiter in vielen Büros stark von der Sonneneinstrahlung abhängt, sofern die Isolierung der Außenflächen des Gebäudes nicht ganz perfekt ist. Die Sonnenstrahlung wirkt sich daher sehr wahrscheinlich auch auf die Lüftungssysteme des UBA aus.

Allerdings betrifft der Faktor Sonnenstrahlung nicht alle Gebäudeteile des UBA gleichermaßen. Die Sonne bescheint das Umweltbundesamt von der Südseite aus, so dass die Büros im Norden des Gebäudes ganzjährig im Schatten liegen (vgl. Abbildung 65).

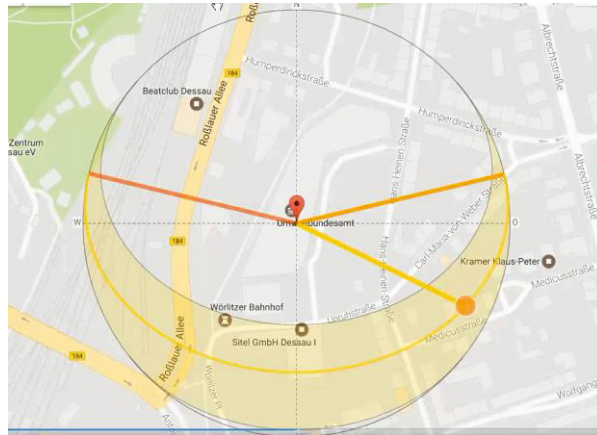


Abbildung 65: Sonnenstandverlauf am Untersuchungsobjekt (UBA).  
(Quelle: <http://www.sonnenverlauf.de> [8])

Mittels eines vom UBA zur Verfügung gestellten Gebäudeplans und einer Excel-Liste von sieben Referenzräumen mit umfangreicher Sensorausstattung konnten Räume in der Außenlage für die beiden Himmelsrichtungen Nord (Raum 0.109) und Süd (Raum 3.015) ausgewählt werden (siehe Abbildung 69). Damit stellen diese Räume die einzig möglichen repräsentativen Messpunkte für die sonnenbeschienenen bzw. –abgewandten Gebäudeteile dar. Während sich Raum 0.109 im Erdgeschoss befindet, ist Raum 3.015 auf der dritten und zugleich obersten Etage des Umweltbundesamtes zu finden. Beides sind als Einzelbüro genutzte Büroräume. Abbildung 66 zeigt ein Foto des Raums 3.015. Raum 0.109 war zum Zeitpunkt des Besuchs beim UBA aufgrund einer Nutzung nicht betretbar.



Abbildung 66: Raum 3.015 des UBA

Zur Untersuchung des Unterschieds zwischen Nord- und Südseite wurden die Korrelationen des Sensors für die Sonnenstrahlung im Verhältnis zu den Raumtemperaturen in den beiden Büros mit Nord- bzw. Südlage betrachtet. Dabei wurde die Raumtemperatur in beiden Räumen in der Mitte des Raumes an der Decke erfasst. Die Ergebnisse bedeuten eine signifikant bessere Korrelation für den Raum in Südlage (Süd: Pearson-Korrelationswert 0.61 (Pvalue < 0.001) versus Nord: Pearson-Korrelationswert 0.15 (Pvalue: 0.004)). Zusätzlich wird dieser Unterschied in Abbildung 67 veranschaulicht.

In diesem Zusammenhang ist anzumerken, dass die vermeintliche Redundanz der Sensormessungen von Boden- und Deckensensoren in den Räumen wertvolle Zusatzinformationen geliefert hätten. Bei der Begehung der Gebäudeteile wurde aber im Nachhinein deutlich, dass die Bodensensoren als nicht zuverlässig gelten, weil sie mittels Kabeln von der Decke bis zum Boden geführt wurden und frei im Raum hängen. Aus diesem Grunde wurden die Bodensensoren der verschiedenen Räume von den Mitarbeitern individuell „beiseite geräumt“ (Abbildung 68). Dieser Fall ist ein Beispiel dafür, dass der Gebäudesensorik keinesfalls blind vertraut werden darf, sondern immer der Kontext hinterfragt werden muss.



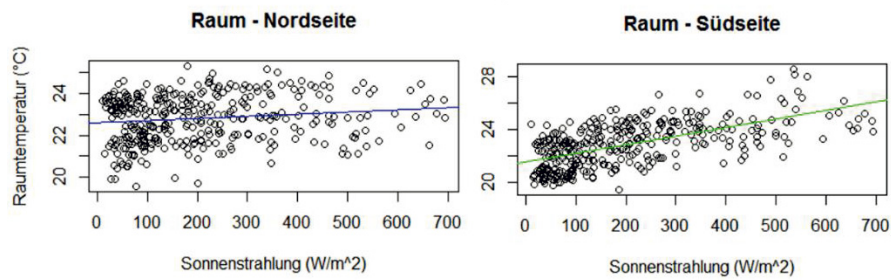
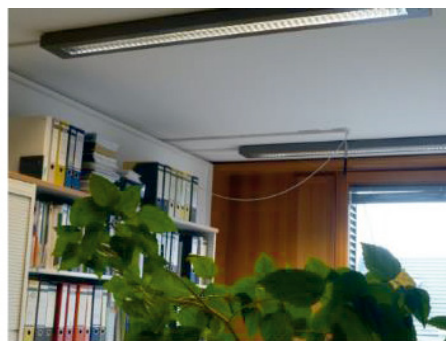


Abbildung 67: Einfluss der Sonneneinstrahlung auf die Raumtemperaturen

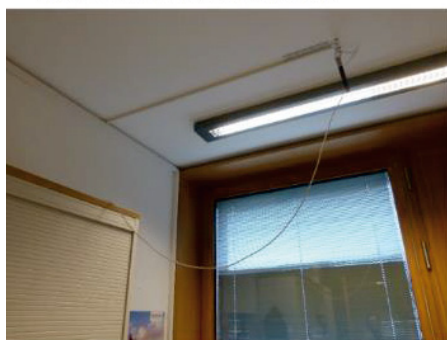
Insgesamt wird also deutlich, dass entsprechend des Sonnenstandverlaufs am Standort des UBA (Abbildung 65) die Raumtemperaturen der Büros auf der Südseite stark von der Sonneneinstrahlung beeinflusst sind, während dies für sonnenabgewandte Gebäudeteile nicht gilt. Exemplarisch zeigt sich das anhand der Temperaturkurven in den Räumen für ausgewählte Tage mit einer sehr hohen (Abbildung 70) bzw. sehr niedrigen Globalstrahlung (Abbildung 71). In beiden Fällen handelt es sich um Montage, betrachtet wird die Mittagszeit. Als Kriterium für die jeweilige Auswahl des hohen ( $563.3 \text{ W/m}^2$ ) bzw. niedrigen ( $22.6 \text{ W/m}^2$ ) Globalstrahlungswertes diente das Verteilungshistogramm der Strahlungswerte (Abbildung 72).



Raum 3.015: auf Schreibtisch



Raum 3.148: hinter Aktenschrank



Raum 0.033: hinter Aktenschrank



Raum 2.211: auf Schreibtisch

Abbildung 68: Auswahl Lage Sensoren Bodentemperatur

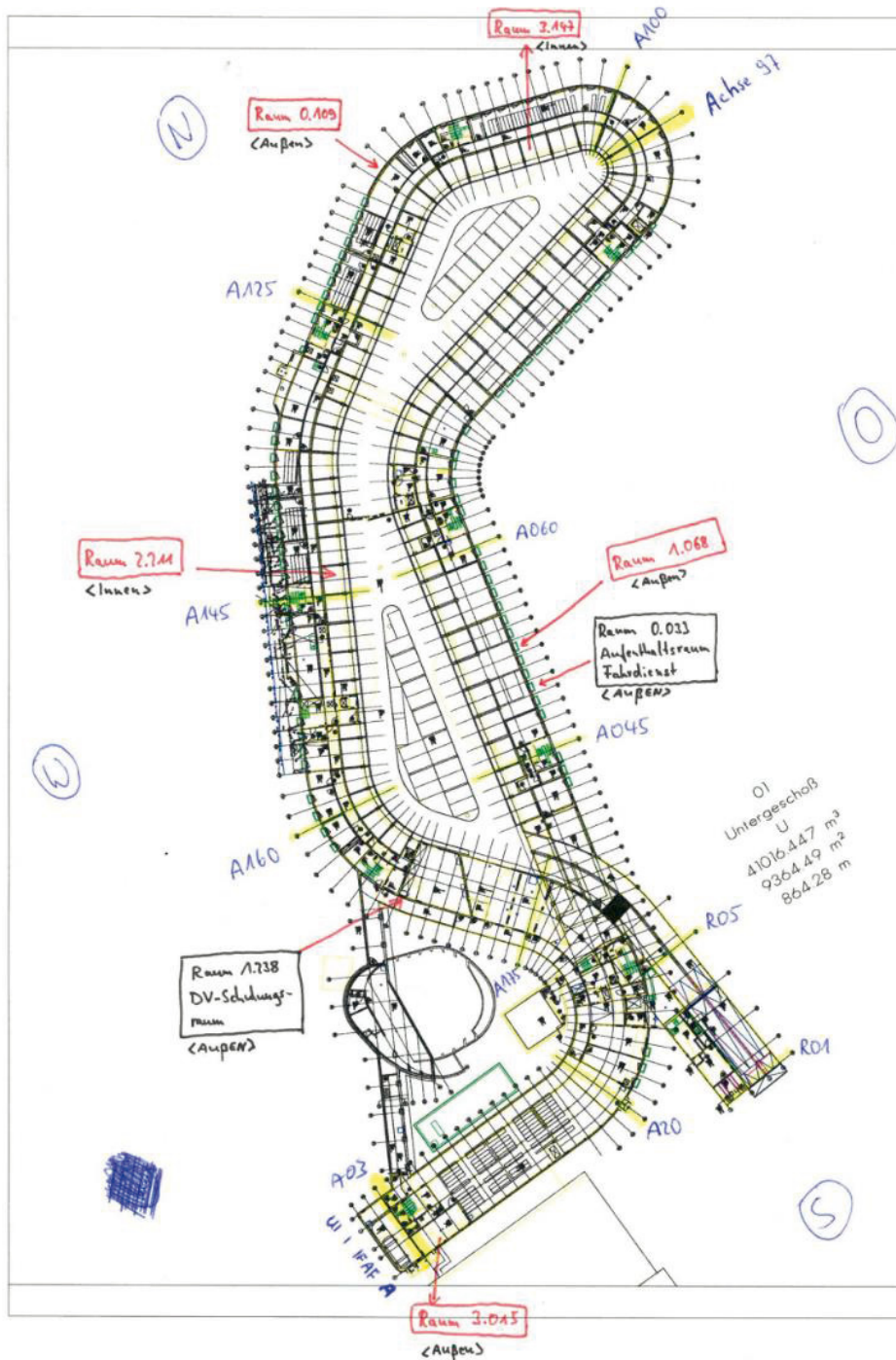


Abbildung 69: Gebäudeplan des UBA und Räume mit Sensorausstattung



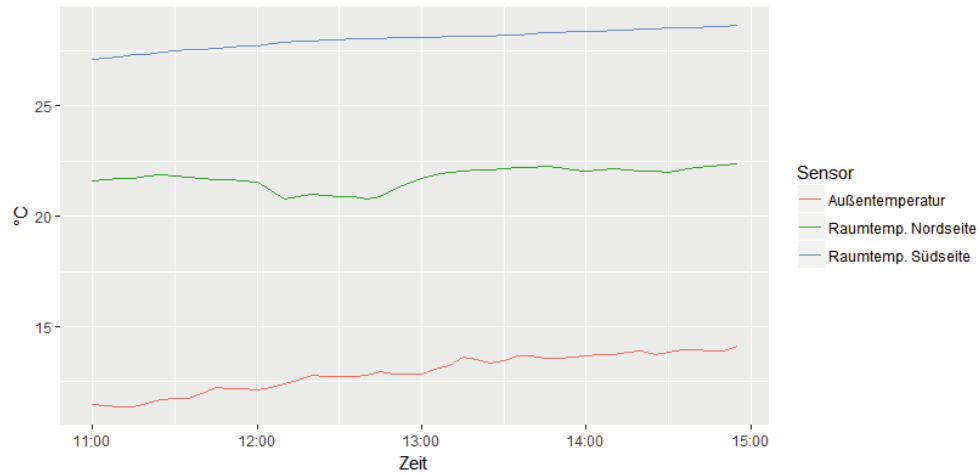


Abbildung 70: Temperaturkurven im Vergleich (Hohe Globalstrahlung)

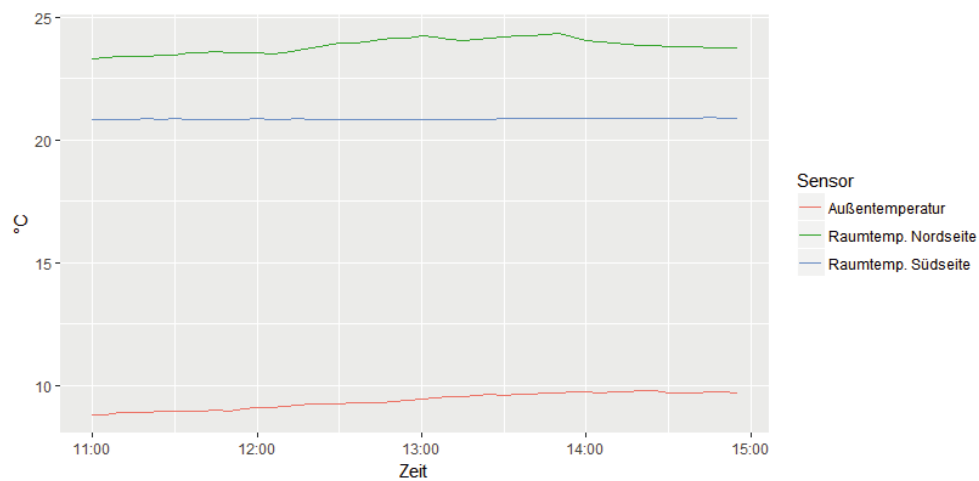


Abbildung 71: Temperaturkurven im Vergleich (Niedrige Globalstrahlung)

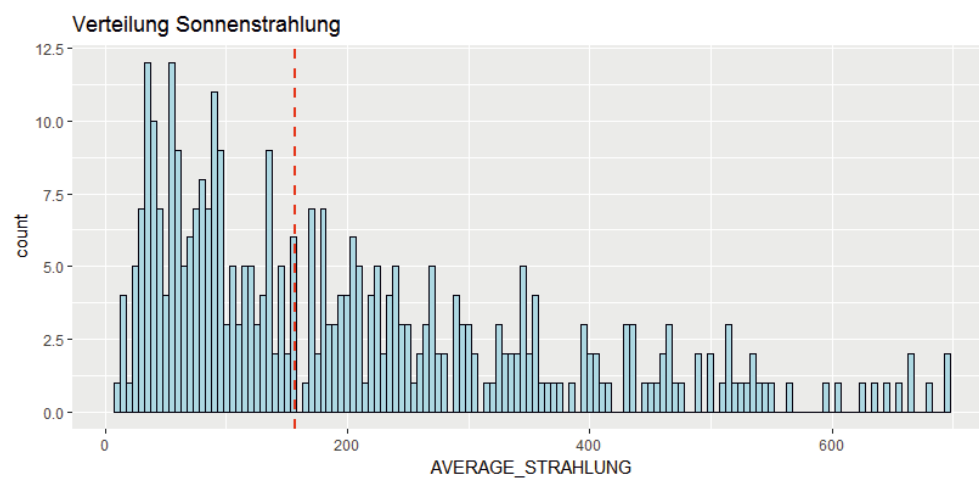


Abbildung 72: Histogramm der mittleren Sonnenstrahlung (11-15 Uhr)

Bisher wurde der individuelle Einfluss verschiedener Witterungsfaktoren auf den Wärmeverbrauch betrachtet. Dabei stellen die paarweisen Korrelationen aber nur Einzelbetrachtungen dar. Aufgrund von Korrelationen der Faktoren untereinander ist damit nicht gesagt, dass es tatsächlich additive Effekte gibt. Beispielsweise korrelieren die Faktoren Außentemperatur und Sonnenstrahlung recht stark miteinander (0.39, Pvalue < 0.001), was auf Überlapp ihres Informationsgehaltes hindeutet. Es stellt sich also die Frage, inwiefern die Kombinationen von Faktoren additive Effekte erbringen.

Folgend wird das zuvor beschriebene Datenset aus aggregierten Sensordatenpunkten verwendet (siehe dazu die Datensatzbeschreibung in Tabelle 3, sowie die Auswahl der Sensoren in Tabelle 4).

Basierend darauf wurden drei verschiedene Regressionsmodelle (M1—M3) erstellt. Alle drei Modelle enthalten als Zielvariable den Faktor CUM\_LFW, d.h. den kumulierten Wärmeverbrauch. Außerdem basieren alle drei Modelle auf der unabhängigen Variable AVERAGE\_AU\_TMP, d.h. der mittleren Außentemperatur. Das Modell M1 bezeichnen wir als Referenzmodell, weil es nur den Faktor Außentemperatur beinhaltet. Die Modelle M2 und M3 dienen uns dazu die additiven Effekte weiterer Witterungsfaktoren abzuschätzen, d.h. im Falle von M2 der mittleren Sonnenstrahlung (AVERAGE\_STRAHLUNG) und im Falle von M3 der Kombination aus mittlerer Sonnenstrahlung, kumuliertem Niederschlag, mittlerer relativer Außenfeuchte, sowie mittlerer Windgeschwindigkeit und Windrichtung.

Tabelle 5: Übersicht der Regressionsmodelle

	<b>Abhängige Variable (Zielvariable)</b>	<b>Unabhängige Variablen</b>
<b>M1 (Referenz)</b>	CUM_LFW	<b>AVERAGE_AU_TMP</b>
<b>M2</b>	CUM_LFW	<b>AVERAGE_AU_TMP</b> + AVERAGE_STRAHLUNG
<b>M3</b>	CUM_LFW	<b>AVERAGE_AU_TMP</b> + AVERAGE_STRAHLUNG + CUM_NIEDERSCHLAG + AVERAGE_AUSSENFEUCHTE + AVERAGE_WND_RCHT + AVERAGE_WND

Wie zuvor im Methodenabschnitt des Wetterkapitels beschrieben, wird der  $R^2$ -Wert zur Abschätzung der Vorhersagekraft der einzelnen Modelle verwendet. Der  $R^2$ -Wert eines Modells quantifiziert die Vorhersagegüte eines Regressionsmodells. Er nimmt Werte zwischen 0 und 1 an, wobei höhere Werte besser sind. Man spricht beim  $R^2$ -Wert von einem Maß dafür, wie viel Prozent der Varianz der Daten durch das Modell erklärt werden.

Die relative Verbesserung des  $R^2$ -Werts (für zwei Modelle mit den Werten „res1“ und „res2“) lässt sich gemäß folgender Formel berechnen:

$$\text{Verbesserung} = \frac{\text{res2} - \text{res1}}{(1.0 - \text{res1})} \cdot 100\%$$

Dazu ein Beispiel: Ein Modell mit dem  $R^2$ -Wert 0.3 kann nur 30% der Varianz in den Daten erklären, es bleiben also 70% der Varianz durch das Modell unerklärt. Ein konkurrierendes Modell für dieselben Daten mit dem  $R^2$ -Wert 0.35 erklärt hingegen 35% der Varianz. Damit sind zusätzliche 5% der Varianz erklärt worden, was bezogen auf die zuvor nicht erklärbare 70% der Varianz eine relative Verbesserung von  $5/70 = 7.1\%$  bedeutet.

Das betrachtete Datenset ist mit 353 Datenpunkten nicht sehr umfangreich. Dennoch ist es wichtig, dass man die Vorhersagen der Modelle nicht auf den zum Training verwendeten Datenpunkten testet, weil komplexere Modelle (mit mehr freien Variablen) zu einer Überanpassung (so genanntem *Overfitting*) an die Daten neigen. Aus diesem Grunde wurden alle Modelle mittels 10-facher Kreuzvalidierung evaluiert, wobei die Datenpunkte in statistisch akkurater Art und Weise sowohl zum Training als auch zum Testen verwendet werden können. Dazu wird die Kreuzvalidierung in 10 Iterationen ausgeführt, wobei jeweils 90% der Daten zum Training und 10% der Daten zum Testen verwendet werden. Am Ende wird auf diese Art jeder Datenpunkt einmalig evaluiert. Abbildung 73 zeigt in einem Plot die 10 linearen Regressionsgraden für das Referenzmodell M1.

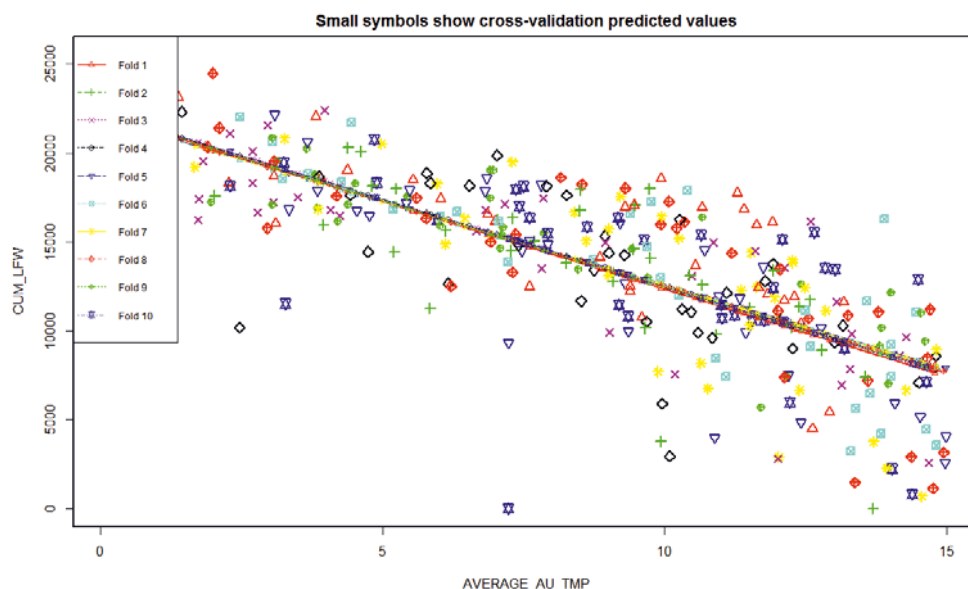


Abbildung 73: Lineare Regressionsgraden in 10-facher Kreuzvalidierung

## Ergebnisse

Die Außentemperatur hat aufgrund der Steuerung über die Heizkurve einen wesentlichen Einfluss auf den Wärmeverbrauch im UBA ( $R^2$ -Wert: 0.59). Sie erklärt den Wärmeverbrauch aber nur teilweise - es muss noch weitere externe Faktoren geben.

Tabelle 6 zeigt die Kreuzvalidierungsergebnisse für die Modelle M1–M3 in einer Übersicht. Anhand der Steigerung der  $R^2$ -Werte lässt sich der Effekt der Hinzunahme von Faktoren erkennen. Faktoren, deren Effekt in den Modellen als signifikant identifiziert wurde ( $P\text{-Value} \leq 0.001$ ), sind farblich grün in der Tabelle markiert. Insgesamt entspricht die Verbesserung des Modells M3 im Vergleich zu M1 einer Verringerung der unerklärten Varianz um 19%.

Auch der Faktor Sonneneinstrahlung für sich alleine genommen nimmt einen signifikanten Einfluss auf den Wärmeverbrauch in unseren Regressionsmodellen, wenn die Faktoren Außentemperatur und Sonnenstrahlung in Kombination betrachtet werden. (Modell M2). Dies zeigt sich anhand der Signifikanz der Koeffizienten in den linearen Regressionsmodellen ( $P\text{-Value} \leq 0.001$ ), wie auch in einer Verbesserung des in Kreuzvalidierung gemessenen  $R^2$ -Werts von 0.59 (nur Außentemperatur) auf 0.64 (Außentemperatur plus Sonnenstrahlung). Abbildung 74 zeigt eine grafische Visualisierung der zweidimensionalen Regressionsebene von Modell M2, in Analogie zu den zuvor beschriebenen eindimensionalen Regressionsgraden aus den Iterationen der Kreuzvalidierung von Modell M1 (Abbildung 73).

Tabelle 6: Ergebnisübersicht der Regressionsmodelle (10-fach kreuzvalidiert)  
In grün: signifikante Faktoren; in Rot: Faktoren, die nicht signifikant sind.

	Unabhängige Variablen	$R^2$
<b>M1 (Referenz)</b>	<b>AVERAGE_AU_TMP</b>	<b>0.59</b>
<b>M2</b>	<b>AVERAGE_AU_TMP</b> + <b>AVERAGE_STRAHLUNG</b>	<b>0.64</b>
<b>M3</b>	<b>AVERAGE_AU_TMP</b> + <b>AVERAGE_STRAHLUNG</b> + <b>CUM_NIEDERSCHLAG</b> + <b>AVERAGE_AUSSENFEUCHTE</b> + <b>AVERAGE_WND_RCHT</b> + <b>AVERAGE_WND</b>	<b>0.67</b>

Insgesamt zeigen die Faktoren Sonnenstrahlung und Außenfeuchte damit additive Effekte, wenn auch die Außentemperatur aufgrund der sehr hohen Korrelation ( $-0.77$ ) mit dem Wärmeverbrauch den Hauptteil der  $R^2$ -Werte ausmacht.

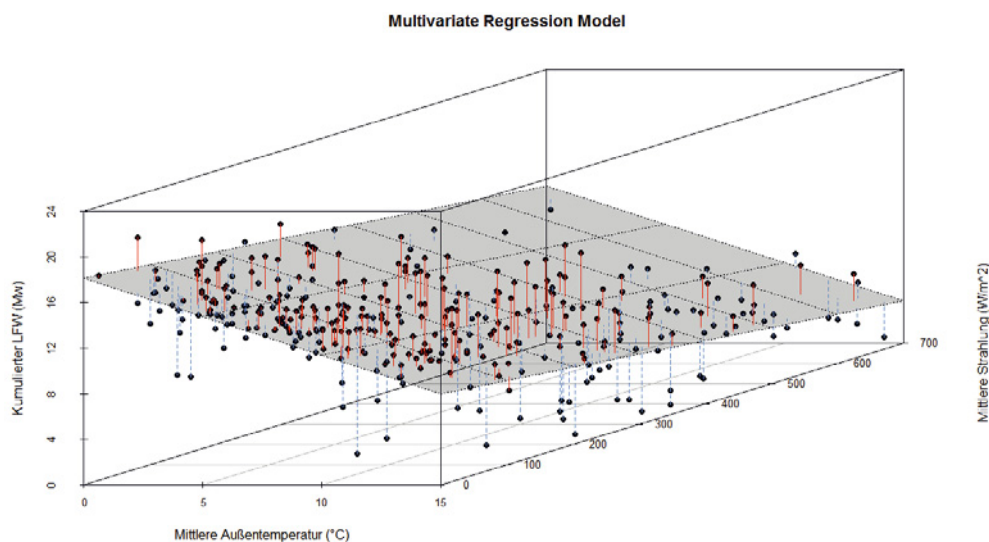


Abbildung 74: Multivariates Regressionsmodell (Wärmeverbrauch in Abhängigkeit zweier Faktoren)

In diesem Kapitel wurde der Einfluss von verschiedenen Witterungsfaktoren (Sonnenstrahlung, Außenfeuchte und Wind) auf den gemessenen Wärmeverbrauch des UBA im Zeitraum 2012 – 2014 für die Mittagszeit zwischen 11 und 15 Uhr untersucht. Dabei zeigte sich, dass es in den historischen Datenaufzeichnungen klar erkennbare Korrelationen zwischen dem Wärmeverbrauch einerseits und der Sonnenstrahlung beziehungsweise Außenfeuchtigkeit andererseits gibt. Dies sind Hinweise darauf, dass der Wärmeverbrauch des UBA, der maßgeblich von dem betrachteten Außentemperatursensor über die hinterlegte Heizkurve gesteuert wird, zusätzlich auch von den Witterungsfaktoren abhängig ist. Der Einfluss der Sonnenstrahlung wirkt sich positiv, der Einfluss der Außenfeuchtigkeit negativ auf den Wärmeverbrauch aus, womit sich die gängigen Thesen bezüglich dieser Zusammenhänge aus den Daten heraus eindeutig bestätigen.

Zwar korreliert der Faktor Sonnenstrahlung mit der Außentemperatur, was eine gewisse Redundanz dieser beiden Informationsquellen andeutet; dennoch zeigen sich in unseren Regressionsmodellen deutliche additive Effekte, wenn man zur Vorhersage des Wärmeverbrauchs während der Mittagszeit neben der Betrachtung der Außentemperatur zusätzlich auch die Faktoren Sonnenstrahlung und Außenfeuchte mit einbezieht.

Aufgrund dieser Ergebnisse erfüllen sich somit die Bedingungen des zu Beginn dieses Kapitels beschriebenen Szenario 1, das heißt, es findet in der Gebäudesteuerung des UBA eine Rückkopplung der Witterungsbedingungen auf den Wärmeverbrauch statt. Die gemessenen Effekte sind deutlich und signifikant. Wir können allerdings die Größenordnung der dadurch eingesparten Energie an dieser Stelle nicht weiter bilanzieren. Anders als in Szenario 2, ist es im Falle des UBA demnach schwierig abzuschätzen, ob sich der finanzielle und technische Aufwand der Entwicklung einer „prädiktiven Heizungssteuerung“ (in Abhängigkeit von Wetterprognosen) lohnen würde, da sich die Witterungsfaktoren auch in der bestehenden Anlagenkonstellation bereits bemerkbar machen.

Die Ergebnisse sprechen dafür, dass die Thermostatsteuerungen der Heizkörper im UBA sich tatsächlich positiv auf den Energieverbrauch auswirken. In Anbetracht der Tatsache, dass nur Teile des Gebäudes der Sonnenstrahlung ausgesetzt sind, ist es bemerkenswert, dass sich die Effekte auf den Gesamtenergieverbrauch auswirken. Insbesondere auch, weil es beim UBA keine getrennten Heizkreise für die unterschiedlich von der Sonne beschienenen Gebäudeteile in Nord- und Südlage gibt.

Für das UBA speziell wären weitere Untersuchungen notwendig, da es zum Beispiel ein großflächiges Atrium mit weitem Glasdach und einen inneren und äußeren Heizkreis gibt. Es handelt sich um ein Gebäude mit sehr anspruchsvoller Architektur. Der von uns betrachtete LFW-Sensor misst somit nur den mittleren Einfluss der beiden Heizkreisläufe. Einzelbetrachtungen wären ergänzend sinnvoll.

Untersuchungen des vorgestellten Typs können auch für andere Gebäude durchgeführt werden, um die Rückkopplung von Witterungseinflüssen auf den Energieverbrauch datengetrieben zu evaluieren. Damit kann ganz konkret überprüft werden, ob sich als positiv angenommene Effekte, wie z.B. aufgrund von Thermostatsteuerungen, auch in der Praxis niederschlagen. Zu beachten ist, dass die Verhältnisse jedes Standorts und jedes Gebäudes (z.B. die Ausrichtung von Gebäudeteilen, die Isolierung der Fassade, etc.) individuelle Einflüsse bedeuten und somit auch individuell in Analysen beachtet werden müssen.

Im Rahmen dieses Projektes wurden verschiedene Ansätze zur Visualisierung der GLT-Daten genutzt. Dies gilt sowohl für die Visualisierung der Rohdaten, als auch der Analyseergebnisse. Für die Erstellung der Visualisierung kamen größtenteils hochgradige Expertentools wie R und Python zum Einsatz. Diese sind hauptsächlich bei komplexeren Analysen relevant. Ergänzend hierzu existieren Tools, die sich durch eine interaktive Visualisierung auszeichnen. Bei diesen liegt der Schwerpunkt vor allem auf der Schnittstelle zwischen Anwender und Daten. Über die Schnittstelle ist eine Interaktion in beide Richtungen möglich und erlaubt auch die Manipulation der Daten.

Die Grundidee hinter diesem Vorgehen stammt aus dem Bereich der Visual Analytics. Ihr liegt zu Grunde, dass ein Mensch im allgemeinen sehr gut komplexe Muster und Trends visuell erfassen kann. Das Ziel von Visual Analytics ist folglich, den Menschen in den Prozess der Datenanalyse mit einzubeziehen. Somit kann der Prozess der Datenanalyse effektiver und effizienter gestaltet werden. Die beiden „Welten“ Daten und Mensch werden dabei in der Regel durch Mensch-Maschine-Interfaces und Dialogen verknüpft. Hierdurch ist es möglich, interaktive Visualisierungen in die Datenanalyse mit einzubeziehen, die nicht nur in der Lage sind die Ergebnisse darzustellen, sondern auch die Eingaben der Analysemethoden beeinflussen können. Dadurch kann das Tuning der Parameter einer Analyse beispielsweise anhand eines visuellem Feedback erfolgen.

Ergänzend hat diese Vorgehensweise den Vorteil, dass das tacite Wissen von Fachleuten berücksichtigt wird, da die Darstellungen um die Anforderungen der Fachleute herum entworfen werden. Daher können diese einerseits direkt die Ergebnisse von Analyseverfahren visualisieren, andererseits aber auch die Daten in einer aufbereiteten Form präsentieren, auf deren Basis der Anwender selbst Entscheidungen treffen kann. Dabei liegt der Mehrwert in der Aufbereitung und im Visualisierungsentwurf, der bestimmte Aspekte für den Menschen sichtbar hervorhebt.

Zwei solcher Beispiele werden im folgenden Kapitel vorgestellt. Durch die Beispiele wird demonstriert, wie durch eine interaktive Visualisierung das Monitoring und auch die Fehlersuche im Speziellen unterstützt werden kann und einem Entscheider somit eine bessere Entscheidungsgrundlage biete. Diese Beispiele wurden mit der in Abschnitt „Tableau“ vorgestellten Software realisiert.

### 5.1 Übersicht aller statischen Heizungsanlagen

Für die Analyse von Heizungsanlagen ist es wichtig keine Anlage isoliert von den übrigen zu betrachten, da bestimmte Phänomene oft durch externe Faktoren entstehen. Um z.B. ein bestimmtes Verhalten gleichzeitig bei mehreren Anlagen festzustellen ist eine anlagenübergreifende Darstellungsform hilfreich. Dementsprechend stellt die Visualisierung, die in Form eines Screenshots in folgender Abbildung zu sehen ist, einen Überblick über die statischen Heizungsanlagen zu einem gegebenen Zeitpunkt dar. Wie in den Plots der Heizungsanlagen üblich, sind die Außen-, Ist- und Solltemperatur sowie die Ventilstellung für jede Anlage zu sehen.





Visualisierung

Abbildung 75: Interaktive Anlagenübersicht

Hier erlauben Slider jeweils die Auswahl des Datums. Über Checkboxes können die zu betrachtenden Anlagen ausgewählt werden. Am vorliegenden Tag zeigt sich sowohl in Anlage 600 als auch in Anlage 601 eine Schwingung. Außerdem ähneln sich die Verläufe der Ist-Temperaturen über den Tag hinweg. Ausgehend von einem Fund in einer Anlage können also die übrigen Anlagen betrachtet werden. Eine ähnliche Übersicht existiert auch als Zeitverlauf einer einzelnen Anlage, um das Auftreten einer Anomalie in einen zeitlichen Kontext der betroffenen Anlage setzen zu können.

## 5.2 Interaktives Dashboard

Visualisierungen wie aus dem vorigen Abschnitt geben einen guten Überblick über Teilaspekte der Daten. Möchte man jedoch einen Überblick über möglichst alle relevanten Aspekte haben, ist ein Gesamtüberblick notwendig. Die hier eingesetzte Software Tableau ermöglicht dies über sogenannten Blätter, die nebeneinander platziert werden können. Die zugehörige Funktion wird vom Tableau-System als Dashboard bereitgestellt. Ein Beispiel hierfür ist in Abbildung 76 zu sehen. Dort sind verschiedene Statistiken zu Ventilstellungen und einzelnen Heizungsanlagen abgebildet und beleuchten diese aus unterschiedlichen Blickwinkeln, die für Betrachter relevant sein könnten. Das Dashboard ist weiterhin interaktiv und unterstützt mit seiner Übersicht das Monitoring von großen Datenmengen. Beispielsweise sind in dieser Übersicht gleichzeitig der Zeitverlauf einer Anlage und eine Anlagenübersicht nebeneinander zu sehen und können gemeinsam evaluiert werden.

Die dargestellten Statistiken können auch im Bereich des Reporting zur Kommunikation von Ergebnissen bzw. Funden verwendet werden. Grundsätzlich dienen diese jedoch vorwiegend zur allgemeinen Entscheidungsfindung.

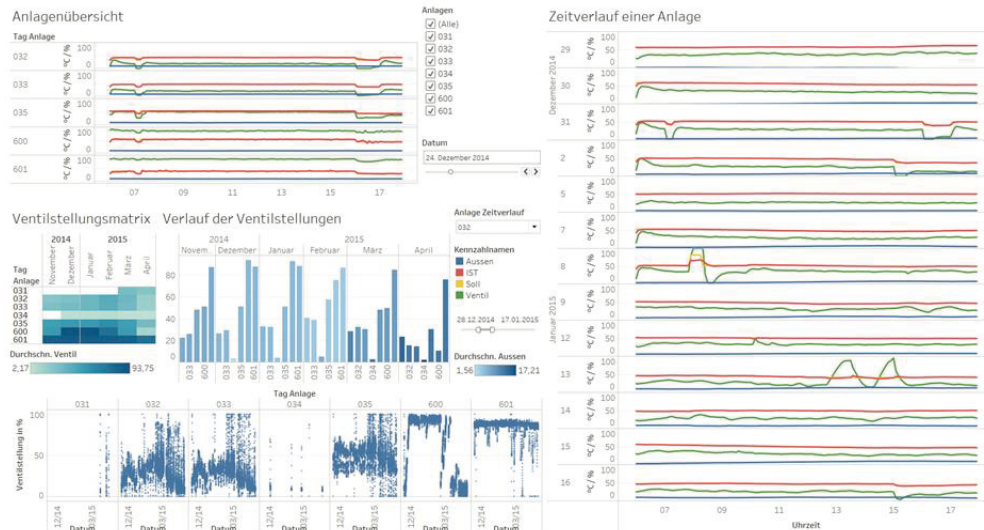


Abbildung 76: Interaktives Dashboard

## 6 Zusammenfassung

Das vorliegende Forschungsprojekt MONALisa hatte das Ziel einen wesentlichen Beitrag zur Optimierung des Energie- und Gebäudemanagements zu leisten. Für die Erschließung wurde prototypisch ein Prozess aufgezeigt, der auf der Basis bewährter Algorithmen der Statistik, des Data Mining und der Zeitreihenanalyse historische Sensordaten der GLT analytisch auswertet. Allgemeines Ziel war es, den Prozess in seiner vollständigen Breite von der Erschließung der Rohdaten bis hin zu den Analyseergebnissen verständlich aufzuzeigen. Dabei galt es insbesondere die jeweiligen Herausforderungen herauszustellen. Im Zentrum standen dabei Anwendungsbeispiele, die auf der Basis von Sensordaten der GLT des Umweltbundesamtes (UBA) entwickelt wurden.

Der im Projekt beschriebene Prozess wurde in mehrere Phasen unterteilt. Dabei wurde sich am CRISP-DM Prozess orientiert, einem gängigen Standard für Analyseprojekte. Die Phasen bauen hierarchisch aufeinander auf, jedoch erlaubt der Prozess auf Rückschritte:

1) Zu Beginn des Projektes wurde zunächst ein allgemeines Domänenwissen sowohl über die GLT als auch das Untersuchungsobjekt selbst aufgebaut. Hierzu zählt z.B. das Wissen über Prozesse innerhalb der GLT, Funktionsweisen einzelner Sensoren aber auch historische Informationen zum Untersuchungsobjekt. Letztere haben sich als unerlässlich für die zielführende Interpretation der Analyseergebnisse herausgestellt. Auf der Basis dieser Informationen wurden parallel dazu Forschungsfragen formuliert, die es in Phase 4 zu untersuchen galt.

2) Vor diesem Hintergrund wurde in der zweiten Phase die Datenaufbereitung durchgeführt. Hierzu wurden die Daten zunächst in einem aufwendigen Prozess aus der GLT des UBA exportiert und anschließend in eine Analysedatenbank importiert. In der Datenbank wurden die Daten anschließend für die Analysen vorbereitet. Hierzu wurden die Daten z.B. mit einer Semantik versehen und einer klassischen Indexierung unterzogen. Zudem wurde über diverse Skripte eine zeitliche Vergleichbarkeit einzelner Sensordaten hergestellt. Dies war notwendig, da die Sensoren mit unterschiedlichen Zeitstempeln vorlagen.

3) Während der anschließenden Exploration galt es ein tiefer gehendes Verständnis für die Daten aufzubauen. Dies galt in zweifacher Hinsicht. Zum einen galt es die Semantik einzelner Sensordaten zu verstehen, aber auch den Zusammenhang mehrerer Sensordaten zu erschließen. Dies wurde durch diverse Skripte datenbankseitig unterstützt. Zudem anderen wurde die Plausibilität der Sensordaten anhand wichtiger Anlagenteile der GLT überprüft. Nur wenn diese gewährleistet ist, sind Daten grundsätzlich für tiefer gehende Analysen geeignet.

4) Die eigentliche Datenanalyse bildete das Kernstück des Projektes. Als Basis wurde hierzu ein Ansatz gewählt, der verschiedene Analysemethoden zunächst in unterschiedliche Komplexitätsstufen einteilt. Hierzu zählen manuelle, regelbasierte oder automatisierte Methoden. Entsprechend dieser Stufen wurden die Forschungsfragen aus Phase 1 den Komplexitätsstufen zugewiesen. Ziel war es dabei, jede Komplexitätsstufe mit einer Forschungsfrage zu versehen. Hintergrund jeder Forschungsfrage bildeten dabei Energieeinsparpotentiale. Dieses Thema galt es in einer Breite an vielfältigen Analysen darzustellen. Dabei wurde sich auf diverse Aspekte der Heizungs- und Lüftungsanlage des Umweltbundesamtes konzentriert. Zu den Analysen zählen beispielsweise die Ventilstellungen bei Heizung und Wärmerückgewinnung, die Differenztemperatur zwischen ein- und ausfließendem Wasser der Fernwärmeübergabe oder die Analyse des Energieverbrauchs in Abhängigkeit der Wetterlage.

5) Im Abschluss des Projektes wurden exemplarisch Methoden der Visualisierung vorgestellt. Dabei wurde auf gängige Software gesetzt, die im Umfeld von datengetriebenen Analysen zum Einsatz kommt und bereits mehrfach vorab in Projekten erprobt wurde. Die Visualisierung unterstützt dabei nicht nur bei der reinen Darstellung der Analyseergebnisse. Sie ist auch Teil der Analyse selbst, da teilweise erste durch die Visualisierung selbst Muster in den Daten sichtbar werden.

## 6.1 Weitere Erkenntnisse aus dem Forschungsprojekt

Die Bezeichnungen Industrie 4.0 und Internet-of-Things (IoT) sind heute viel diskutierte Begriffe. Sie stehen beide als Synonym für die Digitalisierung der Industrie. Der Betrieb eines Gebäudes durch eine sensorgesteuerte GLT ist ein Teil dieser Entwicklung. In dieser Konstellation ist die Kombination von Anlagen, Sensoren und GLT auch als Cyber-physisches System zu bezeichnen. Neben der Vernetzung und Datenerfassung tritt bei diesen Systemen immer mehr die Analytik der Daten in den Vordergrund. Über die Komplexität der Analysen gewinnen diese Systeme immer mehr Wert entlang der Wertschöpfungskette.

Aus den bisherigen Erfahrungen aus dem vorliegenden Projekt decken gängige GLT-Systeme jedoch nur einen Teil der möglichen analytischen Fähigkeiten ab bzw. werden durch den Anwender nicht vollends ausgenutzt. Diese Fähigkeiten sind jedoch notwendig um einen Beitrag zur Erschließung von niedrig investiven Energiesparpotentialen zu leisten. Für den Großteil der Auswertungen muss das analytische Spektrum der verwendeten IT somit um komplexere Analyse- und Visualisierungsmöglichkeiten erweitert werden. Hierzu bieten sich grundsätzlich zwei Ansätze an:

a) Im ersten Ansatz wird die Optimierung intern von den jeweiligen Betreibern der GLT durchgeführt. Neuartige Analysemöglichkeiten werden in das System vorab integriert und von den Mitarbeitern vor Ort durchgeführt. Dies setzt jedoch voraus, dass die Mitarbeiter die jeweilige Kompetenz zur Analyse vorweisen. Sollte diese nicht vorhanden sein, bieten aktuell zum Beispiel Schulungen zum Data Scientist „Smart Buildings“ die Möglichkeit diese Kompetenzen berufsbegleitend zu ergänzen. Zudem muss dem Mitarbeiter ein gewisses Zeitkontingent während seiner täglichen Arbeit für die Analysen

offenstehen. Dies ist zum Teil schwierig, da die Analytik die bestehenden Aufgaben ergänzt. In diesem Kontext können deshalb Kosten / Nutzen Rechnungen die Entscheidung unterstützen, ob die Schaffung eines neuen Arbeitsplatzes zur reinen Analytik der GLT-Daten durch die resultierenden Kosteneinsparungen durch den geringeren Energieaufwand rechtfertigen.

b) Der zweite Ansatz sieht die Verlagerung der Analyse und Optimierung an einen externen Dienstleister vor. Die Optimierung und Analytik geschieht somit nicht intern beim Betreiber der GLT, sondern wird extern außerhalb des Systems durch eine Kopplung durchgeführt. Die internen Aufwände für Personal und Aufbau an Kompetenzen können somit reduziert werden. Die entstehenden Kosten für die externe Dienstleistung können beispielsweise an die resultierenden Kosteneinsparungen gekoppelt werden. Dies setzt jedoch voraus, dass die jeweiligen GLT Systeme geeignete und offene Schnittstellen aufweisen, die eine Kopplung an einen externen Dienstleister ermöglichen. Nur durch diese kann ein effizientes Arbeiten an den Systemen ermöglicht werden. In diesem Kontext muss jedoch die Sicherheit der GLT Systeme im Vordergrund stehen. Nur dadurch kann verhindert werden, dass Unbefugten Zugang zu dem System ermöglicht wird, so wie es in [9] leidliche Erfahrungen gemacht werden musste. Diese können zum Teil erhebliche Konsequenzen nach sich ziehen, sofern, wie im geschilderten Fall, die GLT Teil eines Krankenhauses darstellt.

Welcher der beiden Punkte sich letztendlich als wirtschaftlich praktikabel herausstellt, kann an dieser Stelle nicht verallgemeinernd beantwortet werden. Dies wird immer eine individuelle Entscheidung eines jeden GLT-Betreibers sein. Auf das vorliegende Projekt bezogen, lag in diesem Fall der zweite Ansatz vor. Die Schnittstellen stellten sich jedoch schon zu Beginn des Projektes als erhebliche Hürden dar. Dies ist auf die in den anfänglichen Kapiteln beschriebenen IT-technischen und lizenztechnischen Probleme bzw. Herausforderungen zurückzuführen. Es konnte somit nicht, wie im beschriebenen Idealfall auf eine geeignete und offene Schnittstelle zurückgegriffen werden, sondern es musste ein aufwendiger Workaround geschaffen werden.

Neben einem freien Zugang zu den Sensordaten müssen für tiefer greifende Analysen ergänzende Metadaten zu den jeweiligen Datensätzen zur Verfügung stehen. Diese beschreiben zum Beispiel ergänzende semantische Aspekte der vorliegenden Datensätze. Da in einem System die datenproduzierenden Sensoren in der Regel nie isoliert betrachtet werden können, zählen zur Semantik unter anderem auch die Information über die Zusammenhänge der Sensoren in ihrem jeweiligen Netzwerk. Deren Bedeutung wurde zum Beispiel bei der kombinierten Auswertung von Soll-, Ist- und Außentemperatur in Kombination mit der jeweiligen Ventilstellung in den Heizungsanlagen deutlich. Das Netzwerk kann in diesem Fällen durch Informationen über die Topologie bzw. eine Grafenstruktur abgebildet werden. Für das vorliegende Projekt lagen diesen Informationen leider nicht vor. Somit mussten diese Zusammenhänge durch zu programmierenden Skripte aufwendig konstruiert werden. Dabei wäre es möglich diese Informationen bereits aus der Planung der GLT heraus digital abzulegen. Die aktuelle Entwicklung im Bereich des Building Information Modeling (BIM) birgt für diese Informationen ggf. zukünftig eine neue Quelle.

Zu weiteren Metadaten zählen aber auch Informationen über Wartungsarbeiten an einer Anlage oder der Software. Hierbei sind zum Beispiel ausgetauschte Geräte, die Reinigung von Geräten und Updates in der Software zu nennen. Auf der Basis dieser Informationen lassen sich Auffälligkeiten und Muster in Daten besser beschreiben und schneller deuten. Im vorliegenden Projekt traten bei den Ergebnissen vieler Analysen Auffälligkeiten zu Tage, die sich ursächlich als Ersatz eines Gerätes oder als Softwareupdate herausstellten. Dies wird zum Beispiel bei der Analyse der Soll-Temperaturen deutlich, deren Auffälligkeit nicht auf einen offensichtlichen Fehler, sondern auf eine Aktualisierung der Software zurückgeführt werden konnte. Zudem helfen Metainformationen, um neue

Felder von Analysen zu öffnen. Hier ist der gesamte Themenkomplex der Predictive Maintenance zu nennen. Dieser beschäftigt sich mit datengetriebenen Vorhersage von Fehlern und Ausfällen im System schon bevor diese auftreten. Als Datengrundlage dienen hierzu lange Zeitreihen von Sensordaten in Kombination mit Informationen darüber, zu welchen Zeitpunkten die jeweiligen Gerätschaften einem Ausfall oder Fehler unterlegen waren. Solche Analysen waren für das vorliegende Projekt ursprünglich vorgesehen. Die Praxis hat jedoch gezeigt, dass die hierfür notwendigen Metadaten dem Projekt nicht zugeführt werden konnten. Somit musste der Themenkomplex der Predictive Maintenance aus dem Analysespektrum ausgeklammert werden.

Ergänzend ist für den zweiten Ansatz zu beachten, dass für die Auswertung und Interpretation der Sensordaten einer GLT ein entsprechendes Hintergrundwissen über die Historie des Gebäudes und der einzelnen Anlagen unerlässlich ist. Es stellt sich somit nicht als sinnvoll heraus externe Dienstleister ohne dieses Hintergrundwissen mit den Daten zu konfrontieren. Es bedarf immer einer intensiven Kommunikation mit dem Betreiber der GLT um dieses Wissen offenzulegen und bei Bedarf abfragen zu können. Eine Analyse der Gebäudedaten ist demnach nur möglich, wenn alle notwendigen Kompetenzen gebündelt vorliegen in einem Team vorliegen.

Daraus resultiert jedoch, dass jedes Gebäude mit seiner GLT und den entsprechenden Eigenschaften einzigartig ist. Dies lässt eine einfache Übertragung allgemein entwickelter Methoden nur in einem begrenzten Maße zu. Alle methodischen Schritte müssen grundsätzlich immer in Anbetracht der Historie eines Gebäudes hinterfragt und ggf. individuell angepasst werden.

## 7.1 Meetings / Telefonkonferenzen (Auswahl)

Datum	Teilnehmer	Inhalt
22.02.2016	Speelmanns, Bretschneider	GLT am Beispiel des UBA
24.02.2016	Walther	Datenpunktliste / AKS
25.02.2016	Bretschneider	Heizkurzen und Außentemperatur
10.03.2016	Speelmanns, Bretschneider	Heizkreislauf und Temperaturwerte
08.04.2016	Bretschneider	Prädiktive Wartung und Wetterdaten
08.04.2016	Speelmanns	Prädiktive Wartung
03.06.2016	Krings	GLT auf dem Campus Birlinghoven
08.06.2016	Bretschneider	Prädiktive Steuerung
20.06.2016	Bretschneider	Sensorik in Einzelräumen
26.01.2017	Bretschneider	Wettereinfluss auf Wärmeregulierung
02.02.2017	Bretschneider	Temperaturen, Ventile, Heizung
03.02.2017	Speelmanns	Ventile, Energiesparmaßnahmen
10.02.2017	Speelmanns	Definition Fragestellungen Heizkurve
15.02.2017	Speelmanns	Wettereinfluss auf Wärmeregulierung
30.03.2017	Bretschneider	Besichtigung vor Ort / Besuch beim UBA

## 7.2 Außendarstellung

Datum	Veranstaltung	Inhalt
07.10.2015	GLT-Anwendertagung	Big Data – vom Sensor zur Geschäftsidee
02.03.2016	Projekttag Bau	Vorstellung erster Zwischenbericht beim BBSR
07.03.2017	Projekttag Bau	Vorstellung zweiter Zwischenbericht beim BBSR



Sofern nicht anders angegeben wurden die Abbildungen vom UBA zur Verfügung gestellt oder im Rahmen des Projektes selbst erzeugt.

ABBILDUNG 1: AUFBAU DES BERICHTES .....	5
ABBILDUNG 2: CRISP-DM MODELL (QUELLE: WIKIPEDIA [5]) .....	6
ABBILDUNG 3: ÜBERSICHT ENERGIESPARMAßNAMEN FÜR GEBÄUDE .....	6
ABBILDUNG 4: UNTERSUCHUNGSOBJEKTE.....	8
ABBILDUNG 5: LAGEPLAN GEBÄUDE DES UMWELTBUNDESAMTES (EIGENE AUFNAHME) .....	8
ABBILDUNG 6: SHEDDACH DES UBA VON INNEN (LINKS) UND AUßEN (RECHTS).....	9
ABBILDUNG 7: INNENFLUR DER SCHLEIFE (LINKS), ATRIUM (MITTE) UND BEISPIEL-BÜRO DES UBA (EIGENE AUFNAHMEN) .....	9
ABBILDUNG 8: ACHSEN DES UBA .....	9
ABBILDUNG 9: WETTERSTATIONEN UND AUßENTEMPÉRATURSENSOR DES UBA (EIGENE AUFNAHMEN) .....	10
ABBILDUNG 10: HEIZKREISE DES UBA (EIGENE AUFNAHMEN) .....	11
ABBILDUNG 11: ZEICHNUNG DER HEIZKURVE DES UBA.....	11
ABBILDUNG 12: GLT-ABBILD STATISCHE HEIZUNG INNEN.....	12
ABBILDUNG 13: GLT-ABBILD RLT-ZENTRALE 3.....	12
ABBILDUNG 14: GLT-ABBILD FERNWÄRMEÜBERGABE .....	13
ABBILDUNG 15: BEISPIEL TÜRMEMBRAN (LINKS) UND WÄRMERÜCKGEWINNUNG ABLUFT (RECHTS).....	13
ABBILDUNG 16: WORKAROUND FÜR DIE BEREITSTELLUNG DER DATEN.....	15
ABBILDUNG 17: GRUNDAUFBAU DER LAMBDA-ARCHITEKTUR .....	17
ABBILDUNG 18: AKS VORLAUF-TEMPERATUR SEK. KREISLAUF .....	20
ABBILDUNG 19: ANLAGENKENNZEICHNUNGSSCHLÜSSEL .....	21
ABBILDUNG 20: SENSOREN - BEGINN DER AUFZEICHNUNG .....	21
ABBILDUNG 21: SENSOREN – DATUM ENDE DER AUFZEICHNUNG .....	22
ABBILDUNG 22: ANZAHL SENSOREN PRO MONAT .....	22
ABBILDUNG 23: MINIMALE UND MAXIMALE WERTE PRO EINHEIT .....	23
ABBILDUNG 24: ANTEILE DER DATENQUALITÄTSKATEGORIEN VON 2006 BIS 2015 .....	24
ABBILDUNG 25: ÜBERSICHT SOFTWARE IM PROJEKT.....	25
ABBILDUNG 26: AUSZUG AUS REMOTE DESKTOP DER GLT (UBA) .....	29
ABBILDUNG 27: KODIERUNG DER ANLAGENTYPEN DES UBA.....	30
ABBILDUNG 28: EMPIRISCHE HEIZKURVEN FÜR DEN ZEITRAUM NOV. 2014 – SEP. 2015 .....	32
ABBILDUNG 29: LEISTUNG DES FERNWÄRMEKREISLAUFS IN DER ERSTEN DEZEMBERWOCHE .....	33
ABBILDUNG 30: LEISTUNG DES FERNWÄRMEKREISLAUFS IN DER ERSTEN JULIWOCHE.....	34
ABBILDUNG 31: MONATLICHE KENNZAHLEN EINES SOLLWERT-SENSORS VON JANUAR 2014 BIS OKTOBER 2015 .....	35
ABBILDUNG 32: DYNAMISCHE REGELUNG DES IST-WERTS TROTZ KONSTANTEN SOLLWERTS (ANFANG APRIL 2013) .....	35
ABBILDUNG 33: DYNAMISCHE IST- UND SOLLWERTREGELUNG (ENDE MÄRZ 2015) .....	35
ABBILDUNG 34: DATENMENGE IN ABHÄNGIGKEIT DER ZEITLICHEN SAMPLEGRÖßE (ANNAHME: 72 BYTE PRO EINTRAG INKL. ZEITSTEMPEL) .....	36
ABBILDUNG 35: ÜBERSICHT METHODENANSATZ ANALYTIK.....	37
ABBILDUNG 36: FRAGESTELLUNGEN NACH METHODENANSATZ.....	38
ABBILDUNG 37: DEFEKTER TEMPÉRATURSENSOR IN DER ABLUFT .....	38
ABBILDUNG 38: HEIZUNG EINER LÜFTUNGSANLAGE .....	39
ABBILDUNG 39: VERGLEICH VENTILSTELLUNGEN HEIZUNG UND WÄRMERÜCKGEWINNUNG .....	39
ABBILDUNG 40: DIFFERENZTEMPÉRATUR FERNWÄRMEÜBERGABE + WÄRMETAUSCHER (EIGENE AUFNAHME) .....	41

ABBILDUNG 41: VERTEILUNG DER DIFFERENZTEMPERATUR ZWISCHEN EIN- UND AUSFLIEßENDEM WASSER DER FERNWÄRMEÜBERGABE .....	42
ABBILDUNG 42: DIFFERENZTEMPERATUR DES PRIMÄRKREISLAUFS AUFGESCHLÜSSELT NACH JEDEM MONAT .....	43
ABBILDUNG 43: MONATE, IN DENEN DIE DIFFERENZTEMPERATUR GRÖßER ALS 0°C IST .....	43
ABBILDUNG 44: MONATE, IN DENEN DIE DIFFERENZTEMPERATUR GRÖßER ALS -10°C IST .....	43
ABBILDUNG 45: SOLL- UND IST-TEMPERATUR DER ANLAGE 032 AM 10.04.2015 .....	45
ABBILDUNG 46: HISTOGRAMM DER SOLL-IST-TEMPERATURDIFFERENZ UND DER AUßENTEMPERATUR FÜR ALLE ZEITPUNKTE AN DENEN DIE SOLL-IST-DIFFERENZ GRÖßER ALS 2°C IST (ANLAGE 32 AM 10.04.2015) .....	45
ABBILDUNG 47: DARSTELLUNG AUS ABBILDUNG 46 FÜR DIE ANLAGE 600 AN DEMSELBEN TAG .....	46
ABBILDUNG 48: PROZENTUALE VENTILSTELLUNGEN DER HEIZUNGSANLAGEN 2014 / 2015 .....	47
ABBILDUNG 49: ZUSAMMENHANG ZWISCHEN AUßENTEMPERATUR UND DURCHSCHNITTLICHER VENTILSTELLUNG PRO MONAT PRO ANLAGE .....	48
ABBILDUNG 50: VARIANZ DER VENTILSTELLUNGEN PRO MONAT PRO ANLAGE .....	49
ABBILDUNG 51: VENTILSTELLUNGEN ALLER VENTILE DER SOLARANLAGE .....	50
ABBILDUNG 52: TYPISCHE SCHWINGUNG DER ANLAGE 600 .....	50
ABBILDUNG 53: TEMPERATURSCHRANKEN DER IST-TEMPERATUR UM DIE SOLL-TEMPERATUR .....	52
ABBILDUNG 54: DETEKTIERTE SCHWINGUNGEN DER ANLAGE 600 BEI FAST GESCHLOSSENEM VENTIL .....	53
ABBILDUNG 55: DETEKTIERTE SCHWINGUNG DER ANLAGE 601 BEI WEIT GEÖFFNETEM VENTIL .....	54
ABBILDUNG 56: DACHFENSTERKONSTRUKTION DES UBA (MITTIG) UND WASSERSCHÄDEN AUFGRUND KONDENSWASSER ENTLANG DER HOLZVERTÄFELUNG IM ATRIUM .....	55
ABBILDUNG 57: TEMPERATURVERLAUF (WERKTAGS, ZWISCHEN 11 UND 15 UHR) .....	58
ABBILDUNG 58: MODEL TREE FÜR LINEARE REGRESSIONSMODELLE .....	59
ABBILDUNG 59: GLT-ANSICHT DER DATEN DER WETTERSTATIONEN .....	59
ABBILDUNG 60: HEATMAP PAARWEISER KORRELATIONEN DER UNTERSUCHTEN WITTERUNGSFAKTOREN .....	60
ABBILDUNG 61: ANTI-KORRELIERTES MUSTER VON AUßENTEMPERATUR UND WÄRMEVERBRAUCH .....	61
ABBILDUNG 62: ANTI-KORRELATION VON AUßENTEMPERATUR UND LFW (ZUR MITTAGSZEIT) .....	62
ABBILDUNG 63: WÄRMEVERBRAUCHSÜBERSICHT ÜBER 10 JAHRE .....	62
ABBILDUNG 64: HEATMAP VISUALISIERUNG PAARWEISER PEARSON-KORRELATIONEN .....	63
ABBILDUNG 65: SONNENSTANDVERLAUF AM UNTERSUCHUNGSOBJEKT (UBA). (QUELLE: <a href="http://www.sonnenverlauf.de">HTTP://WWW.SONNENVERLAUF.DE</a> [8]) .....	64
ABBILDUNG 67: RAUM 3.015 DES UBA .....	64
ABBILDUNG 67: EINFLUSS DER SONNENSTRAHLUNG AUF DIE RAUMTEMPERATUREN .....	65
ABBILDUNG 68: AUSWAHL LAGE SENSOREN BODENTEMPERATUR .....	65
ABBILDUNG 69: GEBÄUDEPLAN DES UBA UND RÄUME MIT SENSORAUSSTATTUNG .....	66
ABBILDUNG 70: TEMPERATURKURVEN IM VERGLEICH (HOHE GLOBALSTRAHLUNG) .....	67
ABBILDUNG 71 TEMPERATURKURVEN IM VERGLEICH (NIEDRIGE GLOBALSTRAHLUNG) .....	67
ABBILDUNG 72 HISTOGRAMM DER MITTLEREN SONNENSTRAHLUNG (11-15 UHR) .....	67
ABBILDUNG 73: LINEARE REGRESSIONSGRADEN IN 10-FACHER KREUZVALIDIERUNG .....	69
ABBILDUNG 74: MULTIVARIATES REGRESSIONSMODELL (WÄRMEVERBRAUCH IN ABHÄNGIGKEIT ZWEIER FAKTOREN) .....	70
ABBILDUNG 75: INTERAKTIVE ANLAGENÜBERSICHT .....	73
ABBILDUNG 76: INTERAKTIVES DASHBOARD .....	74

- [1] O. Opel und F. Lüdeke-Freund, „Energie,“ in *Nachhaltigkeitswissenschaften*, H. Heinrichs und G. Michelsen, Hrsg., Berlin – Heidelberg, 2014, p. 434.
- [2] „European Association for External Thermal Insulation Composite Systems (EAE),“ [Online]. Available: <http://www.ea-etics.eu/views/etics/benefits/index.html>. [Zugriff am 16 07 2016].
- [3] M. Chobot, „Automated Buildings,“ [Online]. Available: <http://www.automatedbuildings.com/news/sep12/articles/fmsystems/120829122606fmsystems.html>. [Zugriff am 16 07 2016].
- [4] „Gebäudeautomation,“ [Online]. Available: <https://de.wikipedia.org/wiki/Geb%C3%A4udeautomation>. [Zugriff am 16 07 2016].
- [5] „CRISP-DM,“ [Online]. Available: [https://en.wikipedia.org/wiki/Cross\\_Industry\\_Standard\\_Process\\_for\\_Data\\_Mining](https://en.wikipedia.org/wiki/Cross_Industry_Standard_Process_for_Data_Mining). [Zugriff am 22 04 2017].
- [6] N. M. w. J. Warren, *Big Data - Principles and best practices of scalable real-time data systems*, Manning Publications & Co., 2015.
- [7] T. Müller-Eping, „GLT Anwendertagung 2016,“ [Online]. Available: [https://www.glt-anwendertagung.de/images/2016/Vortraege/16\\_10\\_06\\_Donnerstag/Mueller\\_Eping\\_Fraunhofer\\_ISE.pdf](https://www.glt-anwendertagung.de/images/2016/Vortraege/16_10_06_Donnerstag/Mueller_Eping_Fraunhofer_ISE.pdf). [Zugriff am 16 03 2017].
- [8] T. Hoffmann, „Sonnenverlauf,“ [Online]. Available: <http://www.sonnenverlauf.de>. [Zugriff am 25 04 2017].
- [9] H. Gierow, 28 02 2017. [Online]. Available: <https://www.golem.de/news/gebaeudesteuerung-luxusklinik-vergass-it-im-netz-1702-126362.html>. [Zugriff am 14 03 2017].
- [10] „Jupyter,“ [Online]. Available: <http://jupyter.org>. [Zugriff am 25 04 2017].