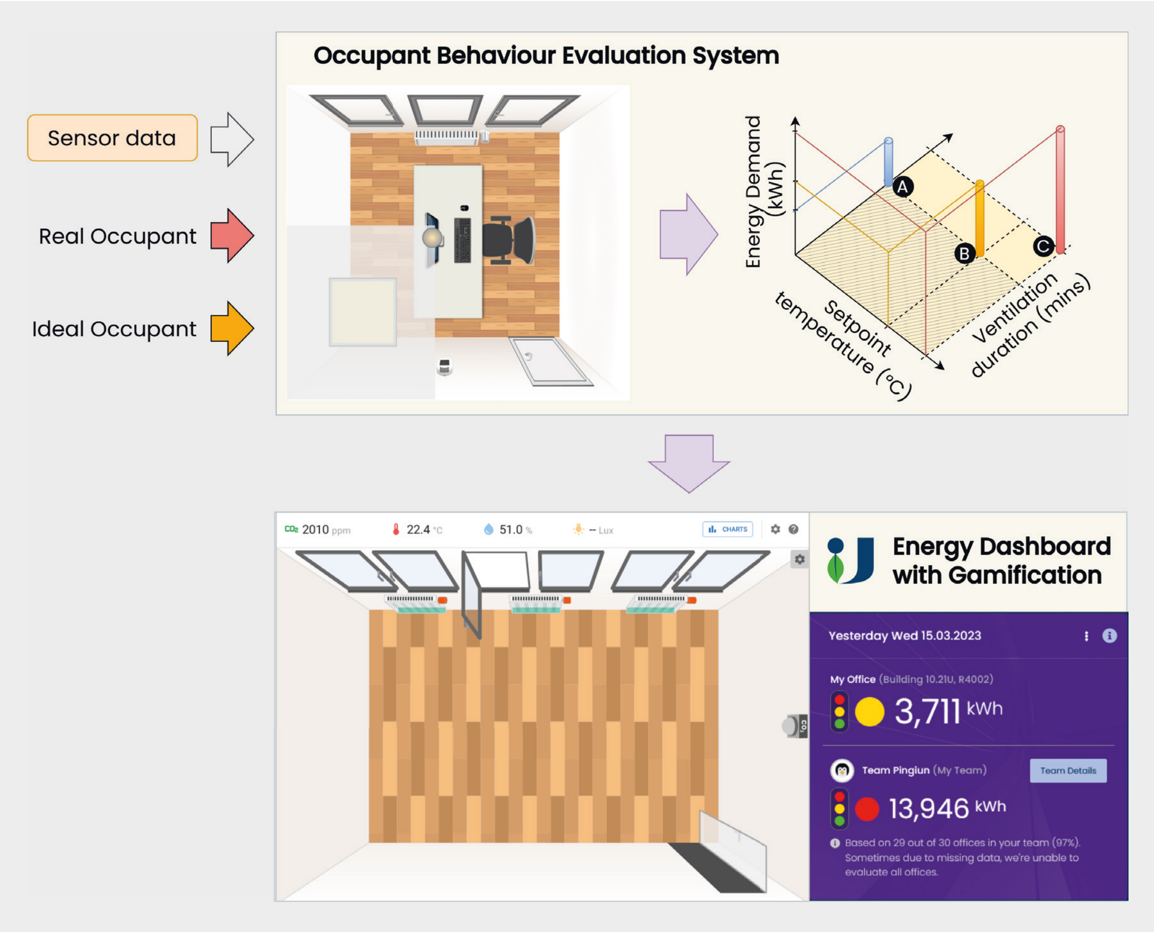


Improving Energy Efficiency of Public Buildings by Influencing Occupant Behaviour using Dashboards and Gamification

Eziama Ubachukwu

Energie & Umwelt / Energy & Environment
Band / Volume 671
ISBN 978-3-95806-840-7

Mitglied der Helmholtz-Gemeinschaft

JÜLICH
Forschungszentrum

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment Band / Volume 671

Forschungszentrum Jülich GmbH
Institute of Climate and Energy Systems (ICE)
Energiesystemtechnik (ICE-1)

Improving Energy Efficiency of Public Buildings by Influencing Occupant Behaviour using Dashboards and Gamification

Eziama Ubachukwu

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment Band / Volume 671

ISSN 1866-1793 ISBN 978-3-95806-840-7

Bibliografische Information der Deutschen Nationalbibliothek.
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte Bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

Herausgeber und Vertrieb:	Forschungszentrum Jülich GmbH Zentralbibliothek, Verlag 52425 Jülich Tel.: +49 2461 61-5368 Fax: +49 2461 61-6103 zb-publikation@fz-juelich.de www.fz-juelich.de/zb
Umschlaggestaltung:	Grafische Medien, Forschungszentrum Jülich GmbH
Druck:	Grafische Medien, Forschungszentrum Jülich GmbH
Copyright:	Forschungszentrum Jülich 2025

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment, Band / Volume 671

D 82 (Diss. RWTH Aachen University, 2025)

ISSN 1866-1793
ISBN 978-3-95806-840-7

Vollständig frei verfügbar über das Publikationsportal des Forschungszentrums Jülich (JuSER) unter www.fz-juelich.de/zb/openaccess.

Acknowledgements

This thesis is the culmination of several years of toil, throughout which I was supported by many people, several of whom I want to acknowledge explicitly.

First, I am grateful to my daily supervisor, Dr.-Ing. André Xhonneux, for his unwavering support, which went above and beyond the tasks of a daily supervisor. I particularly appreciate his efforts in procuring the necessary permissions and workarounds in the face of the policy and legal hurdles that were attendant to my research, and in ensuring that my employment contracts ran smoothly and satisfactorily.

Also, I am grateful to my doctoral father, Univ.-Prof. Dr.-Ing. Dirk Müller, whose leadership, insights, and support enabled me to successfully complete this degree. Likewise, I appreciate Univ.-Prof. Dr.-Ing. Andrea Benigni, who examined my thesis and played a great role in ensuring the defense took place speedily. As the director of ICE-1, he provided a conducive atmosphere for the execution of our various tasks.

To my many colleagues in IEK-10 (now ICE-1), whom I worked with directly and indirectly, and with whom I shared many beautiful everyday moments, I say "thank you", particularly to (in no particular order) Lea, Philipp, Jana, Paul, Alex, Max, Ghadeer, Florian, and several others.

My deep appreciation goes to Dr. Solomon Agbo, through whom the opportunity to do a PhD in Germany came to me in the first place, and who has played a fatherly role for me for several years.

I would like to thank my parents, Prof. Augustine and Prof. Patience Ubachukwu, who set the standards and guided me through life and PhD. They made numerous sacrifices consicienciously to establish me in every aspect and supported me through my hardest moments. To my brothers also, Chibundo, Chidi, Daberechi, and Ugochukwu, who carried my burdens as though they were theirs, and have demonstrated time and again what unity and family really mean, I am eternally grateful.

To my dear wife, Princess, and our kids, who all became an inseparable and indispensable part of my life during this PhD journey, I say thank you for being there and for your love and inestimable value to me.

Finally, to God, the Father of my Lord and Savior Jesus Christ, and to His Holy Spirit, who has loved me with an eternal love and preserved me by His mercies, I owe my entire being now and forever.

Eziama Ubachukwu
May 2025

Abstract

The behaviour of building occupants plays a significant role in the energy-efficient operation of buildings. In fact, occupant behaviour is always implicated as a factor in the performance gap observed between the design performance of a building and its post-occupancy performance. In order to motivate energy-efficient behaviour in building occupants, strategies like eco-visualization and gamification have been successfully employed in the literature. This thesis introduces a suite of web-based software applications that aim to encourage thermal energy-efficient occupant behaviour in an office buildings. Behaviour change motivation is provided through the eco-visualization and gamification with real-time feedback and social competition, in addition to support for occupancy-based heating control. In the process of developing a occupant behaviour evaluation system, systematic analysis of strategies for designing such an evaluation system is developed in this thesis, resulting in the RMM (Rule-Model-Measurement) framework. This framework is then applied to develop the primary behaviour evaluation metric used in the thesis, called *energy penalties*.

An experiment was designed to test the interventions in a real-world setting using naturally ventilated office buildings of Forschungszentrum Jülich, where the focus of the experiment was on the setpoint temperature and ventilation habits of the occupants. The experiments demonstrated that the interventions had largely positive effects on occupant energy efficiency as reflected in ventilation styles and setpoint temperature. The mean daily energy penalties in the ventilation intervention group was 65% lower than that of its control group (1.66 kWh vs 4.67 kWh), with even lower penalties in the "activated" subgroup of the intervention group (0.74 kWh). In another test building that considered both ventilation and setpoint temperature, activated offices had 56% lower daily mean energy penalties than the control (1.91 kWh vs. 4.35 kWh), while in the pilot building, the energy penalties in the activated offices was 40% less than that of its control group (1.61 kWh vs. 2.94 kWh). All these effects were statistically significant and with large effect sizes. Furthermore, year-on-year thermal energy savings of about 18% (11.8 MWh) were realized in the pilot building where occupancy-driven heating was introduced. Furthermore, the results demonstrated superior energy efficiency in offices with full access to the system compared to offices without access as a result of more efficient window ventilation styles in the former. These results demonstrate the effectiveness of the developed systems in improving occupant thermal energy efficiency in public buildings.

Zusammenfassung

Das Verhalten von Gebäudenutzern spielt eine bedeutende Rolle für den energieeffizienten Betrieb von Gebäuden. Tatsächlich wird das Nutzerverhalten stets als ein Faktor betrachtet, der zur Leistungslücke zwischen der geplanten Energieeffizienz eines Gebäudes und seiner tatsächlichen Leistung nach der Nutzung beiträgt. Um energieeffizientes Verhalten von Gebäudenutzern zu fördern, wurden in der Literatur erfolgreich Strategien wie Eco-Visualisierung und Gamification eingesetzt. Diese Arbeit stellt eine Suite von webbasierten Softwareanwendungen vor, die darauf abzielen, das thermische energieeffiziente Verhalten der Nutzer in Bürogebäuden zu fördern. Die Motivation zur Verhaltensänderung wird durch Eco-Visualisierung und Gamification mit Echtzeit-Feedback und sozialem Wettbewerb erreicht, ergänzt durch Unterstützung für eine präsenzbasierte Heizungssteuerung.

Im Rahmen der Entwicklung eines Systems zur Bewertung der Energieeffizienz des Nutzerverhaltens wurde in dieser Arbeit eine systematische Analyse von Strategien zur Gestaltung eines solchen Bewertungssystems durchgeführt, die im RMM-Framework (Regel-Modell-Messung) resultiert. Dieses Framework wurde dann verwendet, um die primäre Verhaltensbewertungsmetrik zu entwickeln, die in der Arbeit als *Energiepenalitäten* bezeichnet wird.

Ein Experiment wurde entworfen, um die Maßnahmen in einer realen Umgebung mit natürlich belüfteten Bürogebäuden des Forschungszentrums Jülich zu testen. Der Schwerpunkt des Experiments lag auf der Solltemperatur und den Lüftungsgewohnheiten der Nutzer. Die Experimente zeigten, dass die Maßnahmen überwiegend positive Auswirkungen auf die Energieeffizienz der Nutzer hatten, insbesondere bei den Lüftungsstilen und der Einstellung der Solltemperatur. Die durchschnittlichen täglichen Energiepenalitäten in der Lüftungs-Interventionsgruppe waren 65% niedriger als in der Kontrollgruppe (1,66 kWh vs. 4,67 kWh), mit noch niedrigeren Werten in der „aktivierten“ Untergruppe der Interventionsgruppe (0,74 kWh). In einem weiteren Testgebäude, das sowohl Lüftung als auch Solltemperatur berücksichtigte, hatten aktivierte Büros 56% niedrigere durchschnittliche tägliche Energiepenalitäten als die Kontrollgruppe (1,91 kWh vs. 4,35 kWh), während im Pilotgebäude die Energiepenalitäten in den aktivierten Büros 40% niedriger waren als in der Kontrollgruppe (1,61 kWh vs. 2,94 kWh). Alle diese Effekte waren statistisch signifikant und wiesen große Effektstärken auf.

Darüber hinaus wurden im Pilotgebäude, in dem eine präsenzbasierte Heizungssteuerung eingeführt wurde, jährliche thermische Energieeinsparungen von etwa 18% (11,8 MWh) erzielt. Die Ergebnisse zeigten außerdem eine überlegene Energieeffizienz in Büros mit vollständigem Zugang zum System im Vergleich zu Büros ohne Zugang, was auf effizientere Fensterlüftungsstile in ersteren zurückzuführen war. Diese Ergebnisse demonstrieren die Wirksamkeit der entwickelten Systeme zur Verbesserung der thermischen Energieeffizienz von Nutzern in öffentlichen Gebäuden.

Table of contents

List of figures

List of tables

Nomenclature

Roman Symbols

A	Area of building envelope	m^2
C_{amb}	CO_2 concentration of ambient (outside) air	ppm
C_{corr}	CO_2 concentration in corridor	ppm
C_{room}	CO_2 concentration in room	ppm
$f_{pen,trickle}$	Trickle ventilation penalty factor	
N_{buf}	Duration of window ventilation "buffer"	min
N_{occ}	Number of occupants in room in CO_2 mass balance equation	–
$N_{vent,eq}$	Equivalent ventilation duration	min
$N_{vent,ref}$	Reference (ideal) ventilation duration	min
$N_{vent,shock}$	Shock ventilation duration	min
$N_{vent,trickle}$	Trickle ventilation duration	min
p_k	Weight of weather cluster k	–
T_{amb}	Ambient (outside) temperature	°C
T_{base}	Base temperature for calculating heating degree days	°C
$T_{room,avg}$	Average room temperature	°C
$T_{sp,ref,occ}$	Reference temperature for setpoint deviation for occupied office	°C
$T_{sp,ref,unocc}$	Reference temperature for setpoint deviation for unoccupied office	°C
T_{sp}	Setpoint temperature	°C
$T_{wall,avg}$	Average wall temperature	°C
U	U-value (thermal transmittance) of building envelope	$kW/m^2{\cdot}K$
$y^{-}_{>ub}$	Indicator variable that checks if CO_2 upper threshold was exceeded in the preceding period	–
$\dot{c}_{pp}$	CO_2 production rate of a single person	ppm/s
$\dot{m}_{airx}$	Mass flow rate of air from room to its surroundings	kg/s
$\dot{m}_{amb}$	Mass flow rate of air from outside into the room	kg/s
$\dot{m}_{corr}$	Mass flow rate of air from corridor into room	kg/s
$\widehat{E}$	Estimated heating demand of building	kWh
U'	Overall heat loss coefficient of building envelope	kW/K

Greek Symbols

ΔT_{cont}	Temperature difference driving the continuous ventilation process in a room	K
ΔT_{short}	Temperature difference driving the short-time ventilation process in a room	K
ρ_{air}	Density of air	kg/m^3

Sets

$\mathcal{N}_d$	Set of time points for day d at which ambient temperature is measured for calculating heating degree days	
$\mathcal{W}$	Weather clusters for deriving energy demand of reference room model.	

Subscripts

sp	Setpoint	–
vent	Ventilation	–

Acronyms / Abbreviations

ADS	Automated Device Specification protocol
API	Application Programming Interface
AUC	Area Under the Receiver Operating Characteristic Curve
AUROC	Area Under Receiver Operating Characteristic curve
BAS	Building Automation System
BMS	Building Management System
CAN	Controller Area Network bus protocol
CV(RMSE)	Coefficient of Variation of the Root Mean Squared Error
DDE	Design, Dynamics, Experience (game design framework)
ECM	Energy Conservation Measure
EEM	Energy-Efficiency Measure
EU	European Union
FPR	False Positive Rate
FZJ	Forschungszentrum Jülich
GMC	Gamification Model Canvas (game design framework)
GPT	Generative Pre-trained Transformer
HDD	Heating-Degree Days
HMI	Human-Machine Interface
HVAC	Heating, Ventilation, and Air Conditioning
IAQ	Indoor Air Quality
IAQ	Indoor Air Quality
IoT	Internet of Things
IP	Internet Protocol
JSON	JavaScript Object Notation
KPI	Key Performance Indicator
LLEC	Living Lab Energy Campus
LLM	Large Language Model
M&V	Measurement and Verification
MBE	Mean Bias Error
MDA	Mechanics, Dynamics, Aesthetics (game design framework)
MIQCP	Mixed Integer Quadratically Constrained Program
MPC	Model-Predictive Control
MQTT	MQ Message Telemetry protocol
NFC	Near-Field Communication
PIR	Passive Infrared
RFID	Radio Frequency Identification

RH	Relative Humidity	
RMM	Rule-Model-Measurement framework	
RMSE	Root Mean Squared Error	
ROC	Receiver Operating Characteristic curve	
SDT	Self-Determination Theory	
SSE	Sum-of-squared-errors	–
SVG	Scalable Vector Graphics	
TCP	Transmission Control Protocol	
TPB	Theory of Planned Behaviour	
TPR	True Positive Rate	
TTM	Transtheoretical Model (of behaviour change)	
UDP	User Datagram Protocol	
VOC	Volatile Organic Compound	

Chapter 1

Introduction

Energy in various forms is used to drive numerous aspects of human life. Many countries, especially in the developed world, spend millions of dollars annually on energy research, ranging from energy generation and distribution to energy use. Among the member countries of the International Energy Agency (IEA) for example, a total of over 20 billion US dollars was spent in 2019, covering research areas such as hydrogen and battery cells, renewable energy sources, and improvement of energy efficiency [1].

The building sector accounts for one-third of the global energy demand and 40% of direct and indirect CO_2 emissions [2, 3]. In the EU, the sector accounts for 40% of energy consumption and over 30% of the CO_2 emissions [4, 5]. Furthermore, it is estimated that about 75% of the buildings in the EU are energy inefficient [6]. The European Commission, under the 2018 Energy Efficiency Directive, requires that EU countries achieve new energy savings of 0.8% each year of final energy consumption for the 2021-2030 period, to reach a target energy savings of at least 32.5% by 2030 [7]. The building sector provides the highest potential for energy savings and therefore is a key target of EU energy efficiency improvement policies [4].

Within the building sector, occupant behaviour has been identified as a key factor in the energy efficiency of buildings and is often implicated in the difference between modelled and actual (post-occupancy) energy consumption of buildings [8–10]. According to the PROBE studies (Post-occupancy Review of Buildings and their Engineering), this difference is usually a factor of two: the actual consumption is twice the modelled consumption [8, 11]. Similar results are also reported by other studies (e.g. as cited in [12]). Furthermore, in one simulation study of energy behaviours of office occupants with profiles classified as one of austerity, standard, or wasteful, it was estimated that the wasteful lifestyle can use up to 90% more energy than the standard energy profile in a one-person office, while the austerity profile can use up to 50% less energy than the standard profile [13]. Additionally, Zhao et al. [14] estimates that technological measures to increase building energy efficiency can only result in 42% improvement, while occupant behaviour is a significant part of the remainder. Clearly, there is potential for the improvement of energy efficiency in buildings through energy-efficient occupant behaviour. This thesis aims to develop and apply methods to improve energy conservation behaviour in occupants of public buildings, using the campus of Forschungszentrum Jülich GmbH (FZJ) as a case study.

1.1 Research Contribution

The main aim of this research is to investigate the effectiveness of heating energy-related occupant behaviour interventions based on the implementation of an energy dashboard and the concepts of *eco-visualization* and *gamification* for occupants in naturally ventilated public buildings. Heating energy efficiency is the focus of this thesis because the potential for energy savings is much higher for thermal demand than for electrical demand in the office buildings involved in the thesis, according to an initial analysis of the available data. Also, the majority of the buildings in the case study are naturally ventilated. In the process of this research, a method is developed and implemented for evaluating the thermal energy efficiency of occupant behaviour. At the same time, a taxonomy is developed and discussed for classifying behaviour-efficiency estimation methodologies. In summary, the main contributions of this thesis are four-fold, namely:

- Development of a gamified energy dashboard that targets the improvement of occupant behaviour in public office buildings, and which integrates with the everyday life of the building occupants.
- Integration of occupant schedules and thermal preferences into the building automation system to achieve higher thermal energy efficiencies.
- Development and characterisation of a novel framework and taxonomy for categorizing energy-related occupant behaviour evaluation methodologies, which is especially useful for gamification, followed by demonstration of the framework by implementing a behaviour evaluation system guided by the framework.
- Real-world experimental investigation of the developed behaviour intervention system, including quantitative and qualitative results and analyses covering the effectiveness of interventions, user feedback, and cost-benefit analysis.

1.2 Structure of this document

The remainder of this document consists of seven chapters. In Chapter 2 the background of the thesis and the review of relevant literature are presented, in order to identify the research gap that this thesis intends to fill. Also, relevant concepts are discussed. Chapter 3 presents the overall system consisting of the software and tools developed to support the energy-related behavioural intervention goals in this thesis, along with the supporting hardware framework and system architectures. The implementation details of the software and tools are also described, with the associated design decisions and theoretical considerations provided as applicable. In Chapter 4, a conceptual analysis of behaviour evaluation methodologies is presented with a bias towards its application in gamification, while Chapter 5 discusses the implementation of the behaviour evaluation system used in this thesis based on the conceptual analysis of Chapter 4. Afterwards, Chapter 6 details the experiment design methodology via which the effectiveness of the intervention measures is tested in the real world. In addition, the chapter presents the methodology employed in this thesis for evaluating the energy performance of buildings. In Chapter 7, the results of the experiment are presented and discussed, while Chapter 8 concludes and provides recommendations for further work.

Chapter 2

Background and Literature Review

This section provides a background on the core concepts pertaining to the thesis and a review of the relevant literature, covering gamification and serious games, and the interplay between occupants and buildings. Afterwards, the research gaps that this thesis fills are identified considering the reviewed literature, followed by an introduction of the general framework under which the thesis is carried out.

2.1 Background

A major difficulty in driving user engagement in public buildings is that the occupants are mostly indifferent to energy consumption efficiency, since they usually are not aware of, or responsible for, the costs [15]. In a 2012 study, it was found that only 25% of employees in a Dutch academic institution cared about the financial cost of their individual energy consumption to the organisation [16, 17]. This is in contrast to residential homes, where saving energy directly translates to lower bills for the occupants. Another difficulty is instrumentation: it is easier to equip a home with smart devices like sensors and actuators (including from legal and policy perspectives), than a public building. Yet again, different societies present different dispositions to such energy-related interventions, especially considering the privacy implications. For example, in the United States instrumentation of offices can be executed in a top-down approach (decided unilaterally by the employer) while in the European Union it tends to be bottom-up (explicitly consented to by the employee) [18]. In Germany in particular, privacy is taken more seriously, and supporting structures like the works council and Data Protection Officers that serve to protect employees from infringements are more vibrant, than in most other European countries [19–22].

Furthermore, the usual aversion of employees to management-led initiatives for fear of the risk of unfavourable working conditions from the employee's perspective, leads to these initiatives being generally viewed with suspicion. Specifically, the risk of personal or potentially privacy-violating observation of work lifestyle, presents difficulties to the adoption of fine-grained smart metering as part of measures that target the improvement of energy efficiency at work. Nevertheless, in order for individuals to improve their energy conservation habits, they need awareness of their current energy performance, as well as knowledge about their potential for improvement, or else they might neither see the need nor the means for changing behaviour in the direction of better energy conservation. In addition, timely feedback to the occupants on their progress towards this goal is required, as well as guidance during the process in the form of actionable advice and tips.

These problems and considerations necessitate the employment of effective approaches to targeting occupant behaviour with respect to energy conservation. As already mentioned, user behaviour is a key component of energy efficiency in buildings, and often leads to sub-performance of buildings compared to design expectations even when the non-human components and systems themselves work as expected.

2.2 A Brief History of Trends in Energy-Related Behaviour Change Programmes

Various types of programmes have been developed over time to induce higher energy efficiency through influencing building occupants and other energy end-users. In the USA, beginning in the 1970s, programmes were initiated targeting reduction in energy use via behaviour change in response to global oil and natural gas shortages during the period [23]. These programmes initially had the form of information campaigns encouraging citizens to conserve energy via e.g. lowering heating setpoints, dressing warm, and weatherproofing homes and businesses [23].

Nowadays, both the motivation for behaviour change programmes and the range of programmes targeting behaviour change have evolved. Globally, climate change is now at the forefront of sustained energy efficiency programmes, although the unexpected invasion of Ukraine by Russia in 2022 led again to drastic energy conservation measures in European countries that depended on gas supply from Russia. On the other hand, technology-based solutions now dominate the energy-related behaviour change landscape. For thermal energy efficiency improvement, for example, connected thermostats are increasingly being adopted, leveraging advances in information technology to offer smart and efficient space-heating management [24].

A key trend in behaviour efficiency programs emerged in the early 2010s with the widespread availability of mobile devices coupled with increasing use of the world-wide web, namely the incorporation of *gamification* and *serious games* as tools for driving behaviour change. The concomitant rise of device connectivity through the Internet of Things (IoT) opened up possibilities for richer, real-time feedback on the state of appliances and environments. As a result, *visualization* of energy-related information through energy dashboards and other Human-Machine Interfaces (HMIs) – collectively known as *eco-feedback* or *eco-visualization* [25, 26] – became a core component of behaviour programs, enabling more efficient and personalized campaigns and user-centric energy system management. Thus, gamification and serious games that are linked with real-time energy system data and which are available on portable devices like laptops and mobile phones naturally gained prominence. Gamification and serious games are discussed in more detail in the next section.

The rise of gamification and serious games targeting behaviour improvement necessarily required more structured approaches to the application of the concepts and in the evaluation of their effectiveness [27]. As a result, behavioural theories from the field of human psychology started finding expression in the design and evaluation of gamified systems and serious games, which then led to the formulation of general guidelines and rules – so-called *game design frameworks* – to aid the development of effective gamification systems and serious games applied in behaviour intervention programmes. The main behavioural theories applied in gamification are discussed in Section 2.3.1, while game design frameworks are discussed later in Section 2.3.2.

2.3 Gamification, Serious Games and Behavioural Change

In this section, gamification and serious games are defined, and some human psychology theories that underpin behaviour change are discussed. Next, game development frameworks, which are formalized processes for the development of games and gamification, are discussed.

Gamification is defined as "the use of game design elements in non-game contexts" [28]. An alternative definition is that gamification is "a process of enhancing a service with affordances for gameful experiences in order to support user's overall value creation" [29]. Here, *affordances* are "qualities of the service system that contribute to the emergence of gameful experience" [29]. Thus, while the first definition approaches gamification from the methods and mechanics perspective, the latter considers the resulting effects on the end-user. However, both definitions have at their core that gamification seeks to create the feeling associated with games, whether from the perspective of the included features in the gamified system, or from the emotional responses evoked in the user of the gamified system. In the literature, a distinction is made between *serious games*, which is the development of full-fledged games for non-entertainment purposes, and gamification, which just involves the use of game elements or "atoms" in an otherwise non-game context [28]. The goal of gamification and serious games is to achieve some positive impact by improving user behaviour whilst providing the fun and engaging experience inherent in games.

Game design elements generally refer to features that are characteristic to games, that is, "elements that are found in most (but not necessarily all) games, readily associated with games, and found to play a significant role in gameplay" [28]. Examples of such elements from the interface design perspective are badges, levels, points, teams, and leaderboards. Other elements include time constraint, limited resources, and turns (or *plies*). Gamified applications combine these elements, depending on business objectives and target audience. A 2017 review of several studies related to gamification and serious games in the area of domestic energy consumption [30] showed that challenges and feedback were the most frequently employed game elements, followed by rewards and sharing on social media. Leaderboards and points were the next in frequency of use, followed in order by tips, levels, rankings, and avatars. At the least-used end of the spectrum were badges and user-generated content.

The notion of the improvement of user behaviour through gamification or serious games derives its basis in certain theories about behaviour change, which enable the streamlining of gamification approaches and the design of effective game mechanics to achieve effective behaviour change. Some of these theories are presented next.

2.3.1 Theories of Behavioural Change

Behaviour change theories have applicability in several areas of human endeavour, and have been extensively applied in the design of games and in the interpretation of the effects of games on players. Understandably, no theory is all-encompassing and able to explain every observation. However, these theories provide a useful lens through which behaviour can be understood and predicted, even if only partly. The following sections provide an overview of a selection of the leading behaviour change theories, as well as comments on their applicability to gamification as appropriate.

Self-efficacy and the Theory of Planned Behaviour

Self-efficacy refers to a person's confidence in his or her ability to take action and to persist in that action in order to achieve a goal despite obstacles or challenges [31, 32]. This theory has been applied in several gamification studies [4, 33–36] and is identified as a factor affecting occupants' ability to conserve energy [33].

Related to the self-efficacy theory, the Theory of Planned Behaviour (TPB), advanced by Ajzen [37], proposes that the driving factors for people's behaviour are their *intention* to perform the action in question, and the degree of control they have over the action [37]. The degree of control in this sense refers to *perceived self-efficacy* – the degree to which the actor believes that they can successfully execute the actions that are required in the given situation [37]. Thus, TPB augments the self-efficacy theory by including the desire to perform an action (intention) as a predictor of behaviour.

Self-Determination Theory

The goal of interventions in user behaviour for improving energy efficiency is two-fold: to provide users with incentives and motivation for using energy efficiently, and to educate and guide them along that path. Ideally, this motivation and knowledge become internalised (*internal motivation*), developing into a permanent lifestyle that lasts beyond the removal of external incentives and motivation. The highly successful *Self-Determination Theory* (SDT) in psychology [38] provides a scientifically proven basis for understanding human motivation. SDT offers a differentiated approach to motivation, identifying different types of motivation in a continuum from *intrinsic motivation* through *extrinsic motivation* to "*amotivation*", along with their associated *self-regulatory* styles, as shown in Figure 2.1. Intrinsic motivation is the "inherent tendency to seek out novelty and challenges, to extend and exercise one's capacities, to explore, and to learn" [38].

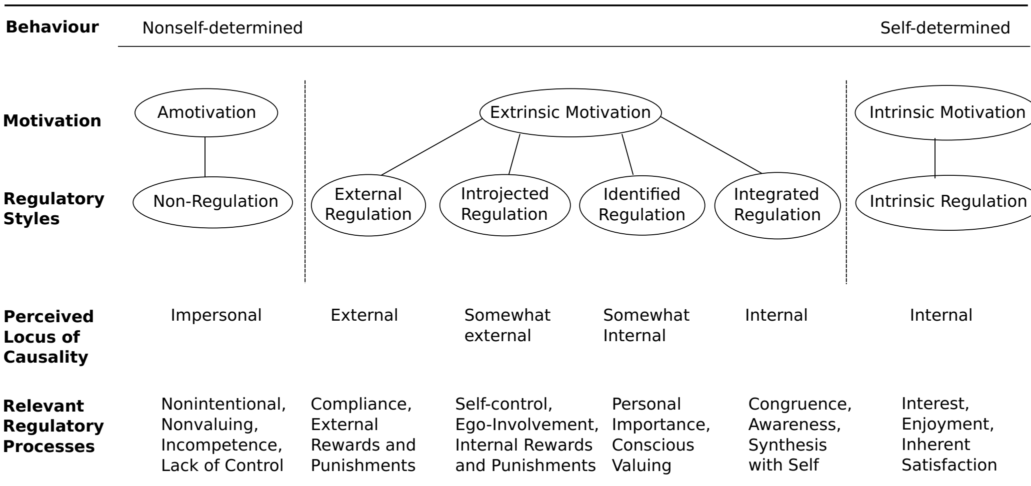


Fig. 2.1 The Self-Determination Continuum showing types of motivation with their regulatory styles, loci of causality, and corresponding processes [38]

The premise of SDT is that every human being is capable of (and is born) being intrinsically motivated, so the theory also deals with the factors that facilitate or forestall intrinsic motivation. Extrinsic motivation, on the other hand, refers to the performance of an activity in order to obtain a separable (from self) outcome.

This kind of motivation forms a continuum that traverses *external* regulation, *introjected* regulation, *identified* regulation, and *integrated* regulation, the latter being closest to intrinsic motivation, as shown in Figure 2.1. "Amotivation" is the state of lacking the intention to act, resulting in not acting at all, or acting without intent (i.e. like a robot). According to previous research, it results from not valuing an activity, not feeling competent to do it, or not expecting it to yield a desired outcome [38].

While intrinsic motivation is the ideal for long-lasting behavioural change, it is important to note that, it is not the only type of self-determined (or autonomous) motivation [38]. In fact, people will be intrinsically motivated only for activities that they are intrinsically interested in [38]. Hence, extrinsic motivation, specifically *identified regulation* and *integrated regulation*, accounts for self-determined motivation that is not intrinsic. Also, although the motivation types form a continuum, people do not necessarily go through each stage in succession, but can begin at any point depending on both prior experiences and situational factors [38]. *Internalisation* and *integration* are the processes through which external regulation becomes more autonomous and central to self.

SDT empirically identifies three psychological needs that form the basis for self-motivation, as well as govern the conditions that foster the process of integration of extrinsic motivation. These needs are *competence*, *autonomy* and *relatedness*. Competence refers to feelings of efficacy and mastery towards an activity. Autonomy reflects the feeling of volition that accompanies an act, in other words, a feeling that the initiative for an action comes from within the individual, and that the individual fully endorses the behaviour. Relatedness is the need to feel belongingness and connectedness with others. Thus, SDT overlaps with the above-mentioned Theory of Self-efficacy and the Theory of Planned Behaviour in that competence and intention are also identified in SDT as drivers for behaviour change.

SDT has been interpreted in the context of gamification and serious games [30]. A 2022 review found it to be the most used behavioural theory in serious games aimed at behaviour change [39]. A fundamental consequence of the application of the theory to gamification is a distinction between *reward-based gamification* (e.g. points, leaderboards, badges) and *meaningful gamification*, which deals with game design elements like play, exposition, choice, information, engagement and reflection [40]. The former is effectively based on extrinsic motivation, tending to produce short-term behavioural changes that disappear when the stimulus is removed, and might even require an ongoing increase in the incentives in order to maintain the user's motivation. The latter targets intrinsic motivation and tends to produce long-lasting behavioural changes. Deterding in [34] underscores this fact, highlighting the central importance of targeting intrinsic motivation in game design.

In both theory and practice, pitfalls have been highlighted related to gamification and serious games. In the context of SDT, the ideal form of motivation is intrinsic motivation; however, there is the risk that the external motivation afforded by gamification and serious games could undermine the ideal goal of intrinsic motivation (see [41]). Nevertheless, as Ryan and Deci [38] points out, intrinsic motivation is *not* the only form of autonomous (self-driven) motivation; therefore, people can still be self-motivated for activities that they are not necessarily intrinsically motivated to do.

Transtheoretical Model of Behaviour Change

The Transtheoretical model (TTM) conceptualises behaviour change as a process that takes place over time and incorporates five stages: *precontemplation*, *contemplation*, *preparation*, *action*, and *maintenance* [42, 43]. Although it was originally applied in the context of health-related behaviour change, it has found validity and application in several other areas like nutrition intervention [31], driving [44], and curbing youth violence [45]. Several gamification studies have equally applied it (see for example [4, 46–48]).

In the *precontemplation* stage, there is no intention to change behaviour on the part of the subject, owing to the subjects not been aware of the need to change in most cases, although people around the them might be fully aware. *Contemplation* stage is the point at which the subjects are aware of the existence of a problem and are seriously thinking about overcoming it, but without having made any commitments to action. At the *preparation* stage, individuals plan to take action in the near future and are reporting small behavioural changes in that direction. *Action* is the stage in which the subjects modify their behaviour, experiences, and/or environment to overcome their problems. Finally, the *maintenance* stage is the period during which the individual keeps up the behavioural change to prevent relapse and consolidate the gains of the previous stage.

While the stages above represent *when* people change, TTM also details the processes of change – the *how*, i.e. "the overt and covert activities that individuals engage in when they attempt to modify problem behaviours" [42]. The model posits that these processes are "differentially effective" in certain stages of change, i.e. particular processes are effective in some stages more than in others. Eight processes of change are implicated in TTM, which are shown in Table 2.1, along with their definitions and their representative interventions.

Table 2.1 Titles, definitions, and representative interventions of the eight processes of change from the Transtheoritical Model of behaviour change. [49]

Process	**Definition: Interventions**
1. Consciousness raising	Increasing information about self and problem: observations; confrontations; interpretations; feedback; bibliotherapy.
2. Self-reevaluation	Assessing how one feels and thinks about oneself with respect to a problem: value clarification; imagery; corrective emotional experience.
3. Emotional arousal (or dramatic relief)	Experiencing and expressing feelings about one's problems and solutions: psychodrama; grieving losses; role playing; journaling.
4. Social liberation	Increasing alternatives for nonproblem behaviours available in society: advocating for rights of repressed; empowering; policy interventions.
5. Self-liberation	Choosing and committing to act or belief in ability to change: decision-making therapy; New Year's resolutions; logotherapy techniques; commitment-enhancing techniques.
6. Counterconditioning	Substituting alternatives for anxiety related behaviours: relaxation; desensitisation; assertion; cognitive restructuring.
7. Stimulus control	Avoiding or countering stimuli that elicit problem behaviours: restructuring one's environment (e.g., removing alcohol or fattening foods); avoiding high-risk cues; fading techniques.
8. Contingency management	Rewarding oneself or being rewarded by others for making changes: contingency contracts; overt and covert reinforcement; self-reward.

2.3.2 Game and Gamification Design Frameworks

In the context of game development, structured approaches exist in the literature and in practice as formalized methodologies that aim to guide the development process to ensure predefined end-results. Some of these *game design frameworks* are especially suitable for the development of full-fledged games with complex mechanics and gameplay, but some are equally applicable to gamification. A few of these game design frameworks are discussed in the following paragraphs at a high level, especially to the extent that they could be applicable to the gamification and serious game development goals of this thesis. (More overviews of game design frameworks are provided in e.g. Villegas et al. [50] and in the thesis of Dormans [51].) It is noteworthy, nevertheless, that there are arguments against using game design frameworks, e.g. as detailed in [51]; in the game development industry, many projects did not use such formalized approaches. However, some form of design documents are in any case helpful for game development.

In a 2015 report, Grossberg et al. [52] developed an analytic game design framework useful for the design, implementation and evaluation of gamified systems. They derived the framework by analysing gamified energy efficiency programmes. The elements of the framework are:

- **Provenance Information** about the origin and stakeholders in the project; nature (pilot or full deployment) and timeline of project
- **Business objectives and desired outcomes** The business case for the deployment of the gamified system.
- **Target audience and their goals** Who is meant to play the game, and what personal goals does the game help them achieve?
- **Target behaviours and metrics for success** What real-world actions does the developer want the players to take? What are the desired quantifiable results?
- **Play space** Real-world or virtual-world play? Mobile phone or computer-based?
- **Progress path, levels** What is the player's progression pathway through levels of increasing challenge and skill development?
- **Triggers** What reminders or calls to action prompt players to continue on their journey?
- **Player engagement model** Inter-player interactions? Teams?
- **Data-based feedback** What quantified data do players receive about their progress? How often?
- **Achievements and rewards** What actual and virtual rewards do players receive? For which achievements?
- **Social dimension** Use of social norms ("peer pressure")? Social media?
- **Intrinsic versus Extrinsic motivation** Are players motivated to change behaviour for intrinsic or extrinsic reasons?
- **Results** What results have been documented? (E.g. energy savings, behavioural changes in players, etc.)

On the other hand, the widely used MDA (Mechanics, Dynamics, and Aesthetics) framework [53] breaks down game development and consumption into interacting sub-components to facilitate reasoning about the effects of design decisions on gameplay and vice versa. The *Mechanics* embodies the particular components of the game at the data and algorithm level, while the *Dynamics* relates to the run-time behaviour of the game, and the *Aesthetics* describes the desirable emotions that the game evokes in its players. The MDA framework identifies the views of the two actors in game production and consumption – the game designer's view and the game player's view. While the game designer thinks from Mechanics to Dynamics to Aesthetics, the player approaches the game from the reverse direction. Hence, the MDA framework espouses thinking like the player in the game development process, and iteratively establishing the lower levels (mechanics and dynamics) from elicitations of the top level (aesthetics). However, the MDA framework has been criticized for being mainly applicable to entertainment games, but insufficient for serious game development or gamification [54]. Another weakness of the framework is that the classification tends to be fuzzy and fails to capture some important aspects of game development like the game's storyline [51, 54].

Another game development framework, called the Gamification Model Canvas (GMC) Evolution framework and based on the MDA framework and behavioural theories, was developed by Escribano [55] as an enhancement to the original (unpublished) Gamification Model Canvas by Sergio Jiménez [55]. At its core, the GMC evolution framework develops dynamic profiles of players and, using relevant theories of behaviour and motivation, connects these profiles to the MDA framework-inspired GMC layers, called *simplicity*, *aesthetics*, *dynamics*, and *components*. It introduces so-called GMC levels to encode how intrinsic or extrinsic a player profile's motivation is, and then uses these levels to appropriately tag various design choices within the GMC layers. These tags then help in creating a compatible mix of game design choices.

Yet again, the more recent Design, Dynamics, Experience (DDE) game development framework by Walk, Görlich, and Barrett [54], was published in 2017 as an advancement over the MDA framework to overcome the weaknesses of the latter. The authors combine ideas from several existing frameworks to extend the *mechanics*, *dynamics*, and *aesthetics* components of the MDA framework, thereby forming three eponymous pillars of the DDE game development framework: *design*, *dynamics*, and *experience*. *Design* encompasses all the aspects of the development that are under the developer's control, including game code and rules, game world and characters, and game data representation, among others. Specifically, *design* is broken into three sub-categories: *blueprint*, *mechanics*, and *interface*. The *blueprint* encompasses the conceptual "setting" of the game, including the character and sound design, the game-world rules, the cultures, the physics, etc. The *mechanics* referring to all the code, data structures, input-output routines and so on that implement the game objects and rules. The *interface*, which refers to what is presented to the user via the game interface, includes look-and-feel, and the sensory feedback or *reporting* that the User Interface provides to the player. This *reporting* can be classified as diegetic, non-diegetic, spatial, or meta [54, 56], as shown in Fig. 2.2. *Diegetic* refers to UI elements that exist in the game narrative as well as in the game space, for example the roads in a racing game. *Non-diegetic* components are neither part of the game narrative nor the game space, for example game menus. *Spatial* refers to elements that are not part of the game narrative, but are still a part of the game space, e.g. visual highlight around the currently active player in soccer games. Finally, *meta* refers to UI components which exist outside the game space but are part of the game narrative, e.g. blood splatter on the screen in a first-person shooter game [56].

The *dynamics* category of DDE framework, similar in spirit to the eponymous concept in MDA framework, refers to all emergent properties of the game which result from the interactions of the game with itself and with the player during gameplay. The dynamics is only indirectly under the control of the game developer, in that the latter specifies e.g. the rules and possibilities within the game but not necessarily all the ways in which these rules can simultaneously interact. In the game development process, the developer can iteratively tune the game to improve the dynamics. Lastly, the *experience* category revises the aesthetics category of MDA, retaining the latter's core meaning but at the same time introducing the *Player-Subject* concept to denote the "subset" of the human player that actually plays the game. Furthermore, in the DDE framework, the challenge attribute of the game is the result of the game presenting a so-called *Antagonist* to the Player-Subject, which in turn evokes emotional, intellectual, and sensory responses in the player.

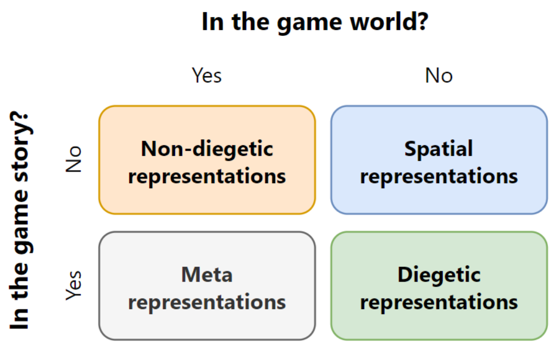


Fig. 2.2 The four types of reporting provided by UI components in games according to the diegetic theory. Source: [56]

Game developers, therefore, can by this token reason about game development narratively, where the game itself is the Antagonist of the Player-Subject, enabling the developer to create engaging storylines. The game narrative has two parts, according to Walk, Görlich, and Barrett [54] and Grave [57] but also reasoned about in DDE: the *embedded* narrative as designed into the game by the developer, and the *emergent* narrative which is "created by the sequence of challenges and other sensations emerging from the game dynamics" [54]. For successful games, these two parts of the game narrative should agree.

2.4 Previous Work on Gamification and Serious Games in Energy Savings

This section discusses the use of gamification and serious games in the literature for triggering behaviour change towards improved energy efficiency. Behavioural theories (such as discussed in the previous section) that are applied in these previous works are also presented where available.

Within the last decade, concerted research efforts have been applied to the concept of gamification as a separate and new phenomenon distinguishable from full-fledged games and from "play", with Deterding et al. [28] providing a formal definition and discourse to situate gamification-related studies on a unified taxonomic foundation. Several of these works have aimed to improve an aspect of energy system efficiency using various user-focused gamification and serious game strategies. In addition, reviews of gamification-related intervention studies are presented in this section.

In 2017, Johnson et al [30] carried out a review of 25 gamified applications and serious games in the domestic energy consumption sector, selected based on a stringent set of criteria, which include large sample sizes, well-described methodologies, presence of controls, presentation of quantitative statistics, use of validated methods to quantify outcomes, and long data collection time frames. From these studies, they identified four classes of outcomes of such behavioural interventions, namely *behavioural*, *cognitive*, *learning and knowledge acquisition*, and *user experience*. Behavioural outcomes dealt with real-world and in-game actions and intentions to save energy. Cognitive outcomes encompassed the affective and motivational aspects, including energy-related attitudes, self-awareness of energy conservation, and motivation to engage in energy-saving activities. Learning and knowledge acquisition covered learning effectiveness and knowledge gains. The User Experience (UX) aspect highlighted attitude of the user towards the game, including engagement, satisfaction and usability. Most of the studies (a total of 17 out of 25) reported more than one outcome category, with the user experience category having the highest frequency, followed by cognitive, real-world behavioural, knowledge, then in-game behavioural categories, in that order. All the outcome classes had mostly positive results across the studies: user experience and in-game behavioural had all-positive results; cognitive, real-world behavioural had mixed (positive and neutral) results; and knowledge had mostly positive, a few mixed, and one negative result. In other words, in nearly all the studies, gamification produced measurable positive effects on the subjects in these four categories. In terms of the effects of the interventions on energy-related outcomes, the results were largely positive, with a few studies having both positive and negative effects.

Regarding the nature of the gamification in these reviewed studies, it was found that roughly half of the studies can be classified as serious games and half were gamified applications/tools. Seven (7) of them were mobile apps, nine (9) were online applications and five (5) were computer games. In terms of level of digitisation and integration into the real world, again roughly half of the games and gamified applications employed significant integrations with the real world, while the other half were entirely digital with no real-world integration.

In AlSkaif et al. [58] a gamification-based conceptual framework for driving customer engagement in residential buildings was presented. Using the widely applied Transtheoretical Model (TTM) for behavioural change [43], they categorised the requirements of the framework and mapped these to high-level gamification objectives. The employed game design elements were categorised into five groups, namely information provision (statistics, messages, tips), rewarding system (electricity bill discounts, virtual currency, prizes/offers/coupons), social connection (competition, collaboration, energy community), user interface (dashboards, leaderboard, progress bar, message box, notifications, degree of control), and performance status (points, badges, levels). Finally, they highlighted the value of the proposed intervention. However, the report had no tangible products (gamified software, results, etc.).

Energy Chickens [59] was a successful gamified intervention in an office building. The health of the chickens represented the energy savings or overspending. Healthy chickens laid eggs daily, but only when the user logged into the app; these eggs were the currency for shopping items in the game store. The game progress path consisted of access to more elaborate and expensive store items, and the ability to view the graph of one's performance and the (chicken) performance of other players. The test consisted of four phases – a baseline, no-intervention phase; two experiment phases of interventions; and, a follow-up phase. Participants decreased their plug load energy consumption by 13% from the pre-game baseline on average, with reduction on non-working days being much stronger than on work days, and 69% of participants indicated an improvement in their energy consciousness due to the game. However, the savings did not persist in the follow-up phase. The game deals only with plug loads, and there was no way to systematically determine what an optimum performance would be: users were made to pledge to reduce consumption by 15% at the beginning of the game, an arbitrarily predetermined goal.

Mindergie [17] was a pervasive game developed in order to foster pro-environmental behaviour at the workplace. It ran for four weeks in an academic institution involving 15 participants. Game elements employed were information, action, badges, quiz, activity, and challenge. Additionally, physical rewards were also given. From the results, the users favoured the active elements of the game (action, activity) above the informational elements (quiz, information on energy systems). Badges had the least impact, and the authors explain that the nature of the badges (not related to any earned skills) was probably inappropriate for a university environment where skills are highly valued. Again, the rewards approach was questionable, since it explicitly motivates extrinsically, and the risk is that the positive behavioural change reverts as soon as the reward is removed. Furthermore, the users demanded for more personalised feedback and more variety in the game mechanics. Mindergie employed game mechanics that were designed to take the users out of their offices/desks to look for clues on the campus, which would not produce sustainable engagement since it detracts from the user's daily work and requires too much effort, although the authors state that the users found them interesting. Also, no quantitative measures regarding energy savings was involved. Finally, the authors point out that a more longitudinal study should be carried out.

The Smart Consumer, Smart Customer, Smart Citizen (S3C) project [60] was an EU project by a group of European companies tagged The S3C Consortium. This project developed tools and guidelines to enable the successful engagement of users in smart grid interventions based on the review of several smart grid projects throughout Europe. The developed tools include a multilingual web-based energy quiz module embeddable in web apps using iFrames [61], an Excel-based tool [61] for categorising the target audience according to their energy-related behavioural profile [62] based on the user segmentation model proposed by Sütterlin, Brunner, and Siegrist [63]. The guidelines cover many topics around customer engagement, including gamification, motivating behavioural change, creating a consumption baseline, and evaluation strategies for such projects [64]. Apart from the quiz mentioned above, no other tangible products related to gamification were developed.

Konstantakopoulos et al. [65] report on a gamification approach targeting dorm residents in a university, in which the occupants were modelled game-theoretically as non-cooperative agents playing a sequential discrete-choice game in order to maximize some individual utility function. A social game experiment was carried out to obtain real data for occupant resource use. In the social game, occupants were awarded points based on energy savings compared to a historical baseline. The game consisted of visual feedback on consumption via a web app for lighting usage, as well as HVAC (heating, ventilation, and air conditioning) systems usage (ceiling fan and air conditioner in this case), with incentives taking the form of lotteries that the occupants can win. Chances of winning the lottery increase with points earned and engagement with the app. The game was run separately in autumn and spring, with results showing significant reduction in the use of the lighting and HVAC appliances measured in minutes per day. In the spring results, reduction in usage ranged from 30% for ceiling lights and ceiling fans, to 76% for desk light usage during weekdays, and 54% to 84% on weekends. For the autumn trial, the reduction for weekdays was 6%, 61%, and 19% for ceiling lights, desk lights, and ceiling fans, respectively. On weekends, more reductions were observed – 38%, 76%, and 60% for ceiling lights, desk lights, and ceiling fans, respectively. Additionally, the data obtained from the social game was used to train the game-theoretic occupant models using deep learning approaches, as well as to forecast energy usage. The results showed reasonable accuracy of predictions of occupant actions and energy usage. Using the AUC (Area Under the Receiver Operating Characteristic Curve) metric, the deep-learning models achieved accuracies of up to 95%.

In addition to the above-mentioned previous work, the European Commission funded a large-scale Research and Innovation programme named Horizon 2020 [66] to the tune of almost 80 million euros. The gamification-related projects described in the remaining paragraphs of this section were partially or fully funded by this programme. These projects are closely related with the general thrust of this thesis and were generally successful; therefore, a more detailed review is provided.

OrbEEt (which stands for Organizational Behaviour Improvement for Energy Efficient AdminisTrative Public Offices) [67] is a major gamification project aimed at developing "an innovative solution to facilitate public and social engagement to action for energy efficiency by providing real-time assessments of the energy impact and energy-related organisational behaviour". It was co-funded by the Horizon 2020 Program with pilot runs in four buildings across four European countries. In the Germany pilot study, which boasted an acceptance level of 80%, overall energy savings of 17% was reported, broken down into 17%, over 20%, and 7% reduction in the energy consumption for heating, lighting, and other electrical load types, respectively [68].

A core contribution of this project is the development of the so-called Systemic Enterprise Operational Rating (SEOR), a building performance certification framework based on operational rating, which integrates business processes model, the occupant behaviour and preference models, and the physical building model [4]. For each of the foregoing aspects, relevant Key Performance Indices (KPIs) are defined, on which basis to judge success and steer enhancement efforts. In the end, buildings are issued an enhanced Display Energy Certificate (eDEC), which, in addition to including organisational aspects, also provides near real-time feedback [4]. It should be noted that the energy performance model does not focus on the structural aspects of buildings, but only addresses the aspects that influence occupancy and the aspects of energy performance, business performance, and individual comfort for the occupants. Similarly, it also does not focus on specific HVAC systems, but on the estimation of the variability of the energy performance due to enterprise- and occupancy-induced factors [4]. Again to point out that the Occupancy Related Indicators are not defined as indicators of the proposed SEOR framework, rather act as enablers for the comprehensive analysis of energy performance indicators [4].

In OrbEEt, detailed and extensive Key Performance Indicators (KPIs) were defined to capture the performance of the facilities under various aspects, namely energy performance, business performance, human comfort, and the behavioural triggering framework [4]. This approach provides several complementary views on the performance, both in terms of energy consumption and in terms of CO_2 emissions. Additionally, the metrics are classified at several granularities across a number of measures, for example at different spatial, temporal, and occupancy-related granularities. Normalisation of these measures is done on the occupancy (e.g. per person), geographical area (e.g. per floor area, per typical room), and time/weather-related (e.g. per season, day-night, seasonal holidays) dimensions. Directions for future research, as outlined in the project summary, include extensions to the OrbEEt framework with an integrated building automation system, and application of the framework to heterogeneous building types.

The EnerGAware project [69] ran from 2015 to 2017 and involved developing a serious game, the Energy Cat game [70]. Pilot studies were conducted in over 50 homes in the United Kingdom. At the end of the interventions, an average electricity saving of 3.46% and average gas saving of 7.48% were achieved, when compared to the baseline period before the interventions [71]. However, the savings were not statistically significant and had a low effect size. Also, the effect of the game was more pronounced in the mid-term evaluation than at the end of the intervention, implying that the effectiveness of the game diminished over time. Also, the authors state that a key reason for the poor results was that participants did not perceive a relationship between the game and their real life.

TRIBE (TRaIning Behaviours towards Energy efficiency: Play it!) [72] was a project aimed at enabling behavioural change in occupants of public buildings towards increased energy conservation. The project analysed and classified energy-related behaviour regarding heating, cooling, lighting and electrical appliances. Five pilot buildings in two European countries (Spain and Turkey) consisting of a public residential building, a university building, and office buildings, were used in the study. A serious game in 3D, called TRIBE, was developed as a mobile app to achieve behavioural intervention. The main objective of the game was to achieve the highest energy savings within the virtual buildings in the game. The winning criterion was that the user achieves 20% energy savings in the virtual building (compared with the original baseline) by year 2020 in terms of the game time progression [73]. In order to measure the effect of energy efficiency measures (EEMs), they applied the building simulation approach. Faced with the difficulty that simulation of a building's energy system for a year or a season on an hourly or sub-hourly basis is usually relatively slow for real-time applications, and that the hardware requirements usually supersede the available resources of mobile phones, the authors first simulated the pilot buildings in detail using EnergyPlus. Inputs to the simulation models include building's structural characteristics, installed heating/cooling equipment, lighting and plug loads, and occupancy profiles. The precompiled simulations were then coupled into an energy simulation engine developed for the game and stored within the game environment [15]. The baseline was obtained via simulation, since at the time of writing the gamification report, enough data was not available from the installed sensors until the end of the project [73].

A total of 250 EEMs – each involving a specific investment cost and leading to specific energy savings – were predefined in the game and grouped in levels: certain EEMs are available only in certain levels of the game. Energy efficiency improvement achieved after the application of each EEM was simulated and fed into the internal game database. For some of the EEMs, these results were modelled as collections of input parameters to the simulated building using EnergyPlus. For others, the results were estimated using other tools (e.g. HOMER for micro-wind turbines installation) or using available methods in the literature where no tools were applicable. Game progression was achieved by unlocking new zones in the virtual buildings as the user's energy decision-making skills improved, and with the new zones came more EEM possibilities. In parallel to the virtual building being changed by the player, a replica of the same virtual building with inputs from the real-time monitoring of the corresponding real-world building was available to the player, so that a comparison was possible between the two virtual representations [73].

In TRIBE, the play space is entirely virtual, and no user actions in the real world are directly accounted for, although the parallel game world that is fed with real data showed the current status of the actual buildings. The virtual building that a player manipulates was also not necessarily the same as where the player was located, so it was not expected that the players should influence the real-world counterparts of the buildings in the virtual world.

ChArGED (CleAnweb Gamified Energy Disaggregation) [74] is a framework that leverages gamification to reduce energy inefficiency and wastage in public buildings. Central to the idea is the incorporation of the possibility for micro-generation (i.e. local energy generation), allowing users to maximise locally generated energy and minimise grid power usage. An intended consequence is the improvement of the predictability of the baseline energy spending. In the project, the energy consumption-related sensors are augmented with location sensors (Near-Field Communication (NFC) and iBeacon devices) to enhance energy disaggregation [74]. The gamified interface was developed for portable/mobile devices and features social interaction aspects like teams and competitions, as well as individual feedback.

2.5 Research Gap and Contribution of Research

Table 2.2 analyzes previous work vis-à-vis the aspects of gamification and user behaviour intervention that are interesting for the current work, thereby providing the basis for identifying the research gap that this work aims to fill. First, as pointed out earlier in the chapter, public office buildings present unique difficulties to motivating energy-efficient occupant behaviour since occupants are not responsible for energy costs. Additionally, in public office buildings especially in Germany, data privacy is enforced more strongly than in most other countries and settings, limiting the range and granularity of measurements that can be taken, as well as the kinds of interventions that can be permitted. Since this work deals with public office buildings, it was important to consider how often previous work dealt with these challenges. Many of the interventions in

Table 2.2 Analysis of previous work with respect to the aspects of gamification and user behaviour intervention that are interesting for the current work.

Relevant features of previous studies						**Previous studies**
Objective benchmark	Cost-benefit analysis	BAS integration	Public building	Living-lab	Thermal energy	
×		×		×	×	[75]
×			×	×		[74, 76]
	×		×	×		[77]
	×			×		[71]
			×	×		[78] [79]
			×			[59]
				×		[65] [80]

public office buildings from previous studies highlighted in Table 2.2, cannot be applied in Germany due to stricter workplace privacy laws and enforcement. For example, requiring that occupant activities and movement in the office are tracked by Bluetooth and Near-Field Communication (NFC) chips beacons as done in the ChArGED project [74] would almost certainly not receive approval from the Workers' Council in public institutions in Germany, especially when the interventions are at scale outside a specialized testing framework.

Related to the foregoing, therefore, is the *living lab* nature of the present work, in which the interventions are integrated into the normal work schedules of occupants of these public office buildings without requiring special business-interrupting energy efficiency campaigns. Indeed, the design of the work in this thesis permits continuous behaviour intervention as part of normal worklife. This characteristic is absent or limited in many of the previous studies analyzed in Table 2.2.

Yet again, in temperate climatic regions like Germany, heating dominates energy consumption in buildings. In the buildings considered in this thesis, thermal energy consumption is often significantly than the electrical energy consumption, and the potential for efficiency improvement is also higher for thermal than electrical energy consumption in these buildings. Since the focus of this thesis is thermal energy efficiency, the contributions of the thesis are mainly focused on user behaviour affecting the thermal energy system. As can be seen from Table 2.2, the majority of listed publications target the consumption of other forms of energy or resources.

Regarding gamification, the performance metric derived in most previous works is based on the *improvement-from-baseline* perspective, lacking an *objective benchmark* that reasonably quantifies and embodies the theoretical potential for improvement given the current scenario (see Table 2.2). Rather, in such an improvement-from-baseline approach, the system disfavours already energy-efficient occupants since there is little room for improvement from the baseline, whilst incentivising high energy-inefficiency during the baselining period. For example, players of Energy Chickens [59] were made to pledge to reduce electricity consumption by 15%, irrespective of where the individual player's current performance lay w.r.t. the ideal consumption. In fact, Johnson et al. [81] clearly point out that it is a common myth to assume that the improvement-from-baseline approach is fair when comparing the performance of multiple actors.

An alternative approach, which is proposed in this thesis, is gamification from the *deviation-from-ideal* perspective, which is expected to have better chances of judging performance fairly. At the heart of the conceptual solution proposed in this thesis is the development of a so-called *oracle* that defines what the ideal behaviour in a given context should be. One advantage of this approach is that the potential for improvement becomes apparent, irrespective of the current baseline behaviour. Hence, the user can be steered towards a realistic ideal behaviour profile, which takes into consideration other balancing factors like maintaining comfort and a healthy office environment. In other words, the proposed evaluation system naturally incorporates comfort-and-air-quality protection when considering heating in naturally ventilated offices, such that extreme energy savings which would likely correspond to deterioration in comfort and indoor air quality is by design not encouraged by the gamification approach. This checks-and-balance system is only present in few studies (e.g. OrbEEt [68], which considers necessary business activities while deciding energy performance); most other studies just assumed that all energy savings, no matter what other values were compromised, was good or rewardable. As would be expected, such interventions tend to have high rates of *rebound*, since the behaviour changes were not by definition sustainable in the first place.

Additionally, included in this work is the integration of interventions with the Building Automation System (BAS), which is not common in the previously analyzed work in Table 2.2. This integration in this thesis allows energy-related occupant actions like specifying temperature setpoints and expected occupancy schedules to be used for automatically heating the buildings. In fact, the use of individual presence schedules for on-demand heating in this work goes beyond the state-of-the-art use of "connected thermostats" in terms of the opportunities to obtain and honour user occupancy schedules, leading to more potential for real-world energy savings, as would be demonstrated in this work.

Furthermore, a novel contribution of this work is the development of a framework for reasoning about the architecture of a behaviour evaluation system. The framework, called the Rule-Model-Measurement (RMM) framework, is established using behaviour-analytic concepts and categorizes the approaches for deriving an energy-related behaviour evaluation metric for use in a reward / penalty intervention system, such as is common in gamification. These approaches are then eloborated with illustrative conceptual applications to the current case study. Also, previous gamification studies having real-world behaviour intervention components are recast into this RMM framework, demonstrating how these works fit into the framework.

Finally, behaviour interventions cost money. Only a handful of studies provided a cost-benefit analysis to determine if the investments in the system were reasonable. In this work, the cost of instrumentation is compared to the whole-building energy savings in the pilot building over an appropriately long time horizon (one year of baseline and one year of intervention). From the cost-benefit analysis, a payback period is also estimated, along with other realistic recommendations and lessons learned that could make similar interventions even more economically rewarding while potentially being equally or more successful.

2.5.1 Research Questions and Research Contributions

Based on the foregoing identifying the gaps in the literature and elaborating the unique nature of the problem at hand in this thesis, the following three main research questions arise, which should be addressed by the thesis.

Q1: What is a systematic methodology for developing an occupant energy-related behaviour evaluation system that is *fair* in that it is applicable to different occupants irrespective of their current point on the energy efficiency spectrum, while taking into consideration occupant comfort and wellbeing? What are the boundary conditions for such a system, including input and output data requirements? What are the characteristics of the system that are relevant to the occupant, when the system is used to drive behaviour change?

Q2: How can occupants of public office buildings, who have no financial incentive to be energy efficient, be motivated to become energy efficient using gamification, while preserving privacy and supporting workplace productivity? In other words, which gamification methodologies should be applied to achieve these aims?

Q3: How effective are such gamification-based interventions in terms of measurable change in behaviour and / or energy efficiency? What are the financial implications of these interventions compared to the achieved energy savings?

The main contributions of this work, therefore, are:

- Development of a gamified energy dashboard that targets the improvement of occupant behaviour in public office buildings, and which integrates with the everyday life of the building occupants.
- Integration of occupant schedules and thermal preferences into the building automation system to achieve higher thermal energy efficiencies.
- Development and characterisation of a novel framework and taxonomy for categorizing energy-related occupant behaviour evaluation methodologies, which is especially useful for gamification, followed by demonstration of the framework by implementing a behaviour evaluation system guided by the framework.
- Real-world experimental investigation of the developed behaviour intervention system, including quantitative and qualitative results and analyses covering the effectiveness of interventions, user feedback, and cost-benefit analysis.

The set of software applications and tools that were developed for the realization of these pillars is referred to in this project as *the Energy Dashboard Suite*. The behavioural interventions in this work are implemented in the style of a *living lab* under the wider Living Lab Energy Campus (LLEC) project. The LLEC project uses the campus of Forschungszentrum Jülich GmbH, (FZJ) located in the city of Jülich in North-Rhine Westphalia, Germany, as a test-bed for innovative and future energy systems. The campus comprises about 164 buildings, with most buildings housing offices and/or laboratories. One building – the Seecasino – serves as the campus restaurant, catering to about two thousand employees and visitors every day. Two other buildings house the Jülich supercomputers, JEWELS and JURECA. In this thesis, only a small subset of the FZJ buildings (12 in total) are considered in detail.

Chapter 3

Methodology

Given the research questions **Q1** to **Q3** raised in the previous chapter, this chapter builds the foundations for addressing the questions. It begins with the development of the methodological basis for evaluating the energy efficiency of building occupants' behaviour, employing behaviour-analytic concepts and considerations. The aim is to include such evaluations in a *fair* gamification system, thus addressing **Q1**. Subsequently, the design of experiment that enables testing the effectiveness of the interventions is established, along with a brief introduction of the developed software to provide some context for the experiment design. In particular, the design of experiment elaborates the hypotheses to be tested and the strategies for testing them, in line with research question **Q3**. (The details of the gamification implementation are discussed in Chapter 4, along with how the implementation addresses **Q2**.) The rest of this chapter covers standard methodology needed for evaluating energy savings at the building level, as well as for deriving inputs to the behaviour evaluation engine.

Conceptually, this work is founded on four foundational concepts related to eco-feedback and gamification, the so-called *pillars* of this thesis, which are designed to work synergetically to achieve improvement in the energy efficiency of occupant behaviour (see Fig. 3.1). The first pillar is **eco-visualization**, in which the state of the surrounding energy system is shown to the occupant at different spatial and temporal resolutions. The second pillar, **control**, provides a human-machine interface (HMI) that enables the occupant to interact with the Building Automation System (BAS) in order to achieve comfort. The third pillar is the concept of a *behaviour evaluation system* for energy-related occupant behaviour, which in turn powers a **gamification** system that providing performance feedback and competition.

3.1 Performance Feedback

As substantiated extensively in the literature, gamification as a behavioural intervention strategy in energy research can be effective to motivate occupants towards improved energy efficiency. *Performance feedback* is a core aspect of gamification. While the term *feedback* is difficult to precisely define from a behaviour analytic perspective [82, 83], a working definition is that feedback is "presentation of an exteroceptive stimulus whose parameters vary as a function of parameters of antecedent responding" [83]. In other words, feedback is external to the recipient and the properties of feedback depend on the recipient's prior responses. However, the implementation of feedback varies widely across studies. The process by which gamification influences behaviour involves *operant conditioning*, which is a behaviour-analytic term that describes the learning process where voluntary behaviour is modified by associating particular behavioural responses with reward

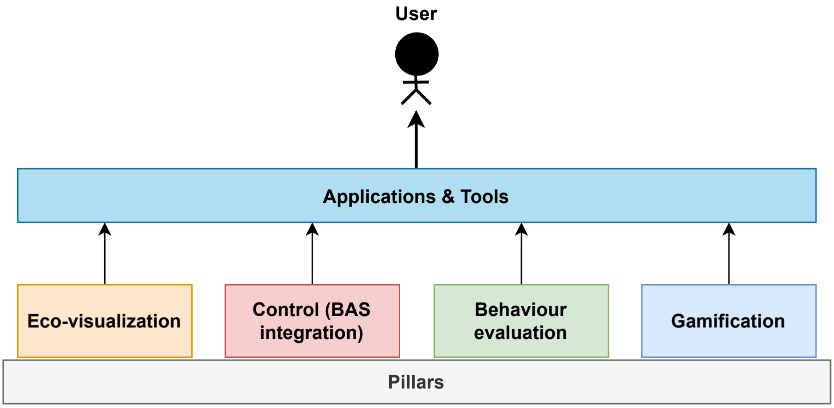


Fig. 3.1 The four main "pillars" of the thesis, embodied within a set of user-centric applications and tools.

or punishment [84]. Although strictly speaking, operant conditioning is distinct from feedback albeit very closely related [83], in this discussion the concepts are sometimes used interchangeably when feedback serves as reinforcement or punishment, since in the application of the concepts to gamification, game elements like points and badges provide performance feedback while simultaneously functioning as rewards and punishments (when points are lost). In this sense, therefore, performance feedback can be used as an operant conditioning procedure where reward *reinforces* some behaviour and *punishment* discourages it.

Broadly speaking, performance feedback can be categorized as *objective* or *evaluative* [82, 83]. *Objective feedback* provides information about performance without comparison to any expected standards, e.g. "You have won 10 points." in a gamified app. *Evaluative feedback*, on the other hand, compares performance to a predefined standard for judging how good or bad the performance is, such as "Excellent!" or "Poor". In a social context, feedback that compares a user's performance with that of others, termed *normative feedback*, is often employed in gamification [47, 85–87]. *Injunctive normative feedback* compares the individual's behaviour to behaviours that are deemed socially acceptable, i.e., "what ought to be", while *descriptive normative feedback* compares the individual's behaviour to what is typical in an environment, irrespective of if these typical behaviours being socially acceptable or not [88–90]. A common drawback of descriptive norms is the so-called "boomerang effect," whereby individuals can be motivated to behave in socially unacceptable ways when they realize that "everybody does it" [85]. Hence, injunctive norms are regarded as a means to prevent the boomerang effect [88, 89, 91] and are proven to be superior in encouraging positive behaviour than descriptive norms [90], although the combination of both is most effective [91].

Therefore, the conceptual development of the evaluation methodologies in this chapter incorporates the idea of injunctive norms by first determining what is *correct* behaviour as regards occupant interactions with the heating-related systems of a building. It is then necessary to quantify these behaviours or their effects in energy terms or in some other equivalent terms like points and badges, so that numerical comparisons can be made between *correct* and *actual* occupant behaviour. These ideas are fleshed out in this chapter at the conceptual level by discussing the theoretical considerations for developing such behaviour evaluation, without regard to the specifics of how it is implemented in practice. Specifically, the method of incentivizing occupants, i.e., use of rewards vs. penalties, is first discussed, followed by a taxonomy for methodologies for developing such a gamification-oriented behaviour evaluation system (rule-based system vs. model-based system), along with detailed comparative analyses of these methodologies. The implementation of the behaviour evaluation system as applied in this thesis based on this chapter is embodied in Juracle and is presented in Chapter 4.

3.2 Operant Conditioning: Reinforcement vs. Punishment

As stated previously, *operant conditioning* is the term used to describe the use of *reinforcements* and *punishments* to modify behaviour [84, 92]. The four styles are *positive/negative reinforcement* and *positive/negative punishment* [93]. In this classification, *positive* represents *adding* something, while *negative* represents *removing* something. *Reinforcement* means *increasing* the likelihood, frequency or degree of a (desirable) behaviour, while *punishment* means doing the opposite for non-desirable behaviour.

Applied to occupant behaviour motivation for improved energy efficiency, the operant conditioning mode used to incentivise behavioural change influences the behaviour evaluation methodology. Example incentivisation scenarios are outlined and classified in Table 3.1. Scenario combinations are also possible.

Table 3.1 Possible incentivisation scenarios for motivating behaviour change and their corresponding operant conditioning classification. Combinations of scenarios are also possible.

Scenario ID	Scenario Description	Operant Conditioning Classification
S1	*Addition* of points for energy savings	Positive reinforcement
S2	*Removal* of points for energy waste	Negative punishment
S3	*Addition* of penalty points for energy waste	Positive punishment
S4	*Removal* of penalty points for energy saving	Negative reinforcement

Each of the scenarios (or combinations thereof) has its own characteristics, which make it differentially applicable in different use cases. Before discussing the operant conditioning modes in details, it is necessary to present the criteria which guide the selection of conditioning modes for this thesis. These criteria are as follows:

- **Intuitiveness and simplicity**: The concept should be easy to appreciate for the users.
- **Ease of retrospection**: The concept should admit an easy retrospection, i.e., the end-result of the evaluation based on the concept should be backward-traceable to the user's behaviour that produced it. In other words, the derivation of the incentive from the deviations from the reference point should be transparent.
- **Provable fairness**: It should be demonstrable that the concept has a reasonable degree of fairness in its application to users with different starting points and environmental characteristics.
- **Compatibility with a serious game**: For possible integration into a serious game centered on reducing energy demand in a virtual environment through in-game purchases, the concept should be compatible with energy or monetary units.

The last requirement biases the choice of operant conditioning mode towards having real-world significance, as against arbitrary points that have no physical meaning.

3.2.1 Analysis of Individual Operant Conditioning Modes

The operant condition modes presented in the previous section of Table 3.1 are analysed in this section. The question being addressed by this analysis is

Scenario S1 (*positive reinforcement*), where rewards are given for behaviour improvement, is commonplace and intuitive. It naturally induces a "deviation-from-worst-case" evaluation approach: for rewards, the natural limiting case is zero reward for zero improvement, with the reward increasing in proportion to the deviation from the reference point (i.e. a behaviour improvement compared to some worst-case reference). Indeed the reverse relationship is possible, where the reward grows with a *decrease* in deviation from a reference point, but this requires that the reference point be defined as an "ideal behaviour", such that the maximum reward coincides with the minimum (zero) deviation. This approach then demands the definition of the maximum reward ab initio as a boundary condition at the ideal reference point. Finding such a maximum that can be justified intuitively is difficult using units compatible with energy or money, and no easy solution was found in this analysis. Therefore, a reward applied in the deviation-from-worst-case sense offers the most natural and intuitive approach for positive reinforcement.

In the case of pure *negative punishment* (scenario S2), some existing asset or advantage is removed in response to energy waste (undesirable behaviour). This approach requires the pre-existence of assets that can be taken away, which does not fit well with the environment of the case study, since there was no previous opportunity to earn assets in the experiment setup prior to the commencement of evaluations. In combination with *positive reinforcement*, however, the assets could be the rewards previously earned by energy savings. Such combinations are discussed in the next section.

Positive punishment (scenario S3), in which penalties are added for undesirable behaviour (*wasted energy*), naturally induces a "deviation-from-ideal" approach. Here, a direct proportional relationship can be defined such that the minimum (zero) deviation coincides with the minimum penalty. Energy wasted in kWh is an example of a penalty, where it is exactly zero when the occupant's behaviour is ideal. Unlike positive reward based on deviation-from-worst-case semantics discussed previously, the typical deviations of penalties are relatively close to the origin (ideal reference point) in practice, so that the magnitudes and spreads of the deviations are easy to appreciate with respect to the reference point.

Scenario S4 is an example of *negative reinforcement*, where an unpleasant effect (e.g. a debt or penalty) is removed in response to positive behaviour in order to reinforce the behaviour. Like the *negative punishment* approach, it requires the pre-existence of the unpleasant effect, and so can only make sense in combination with positive punishment for the generation of the unpleasant effect in our use case.

3.2.2 Combinations of Operant Conditioning Modes

From the foregoing analysis, we find that *positive reinforcement* (addition of rewards) and *positive punishment* (addition of penalties) are the modes that can best be applied *individually* in our case study. For applying reinforcement or punishment in response to desirable or non-desirable behaviour, respectively, within a single setup, *positive reinforcement* paired with *negative punishment* can be considered as two sides of the same coin: the assets created by rewards are reversed by removing the rewards when the behaviour changes. For this combination, it is not possible to have net liabilities (e.g. debts), since it lacks the mechanism to create penalties, but rather only acts on existing rewards. Likewise, *positive punishment* paired with *negative reinforcement* are conceptually related such that the penalties created by the former are reversed by the latter. Again, there is no mechanism to create net rewards. These two combinations have the issue that when the respective commodity is depleted (no existing rewards in the former case, and no existing penalties in the latter case), further activation in the direction of depletion does not produce effect. Hence, the ability to differentiate between the performance of no-commodity entities is lost.

The combination of *positive reinforcement* with *positive punishment* is more complicated, since it requires tracking at least two separate commodities – the assets accrued due to energy savings do not directly interact with the debts accrued due to energy wastage, since there is no mechanism for depleting any one of these commodity classes or for converting one into the other, at least within the definitions of implicated operant conditioning modes.

Finally, the combination of all four modes leads to a continuous spectrum that crosses a neutral point, which represents how many real-world systems work, for example in the financial sector. *Positive reinforcement* creates assets in response to energy savings, which *negative punishment* depletes in response to energy wastage until the neutral point is reached. Subsequently, *positive punishment* moves the operating point into the liability zone (debts) in response to further energy wastage. Energy savings from this point cause first a reversal of the liabilities (debts) via *negative reinforcement* up to the neutral point, before *positive reinforcement* again takes over, completing the cycle. The combination requires, however, without loss of generality, that the reference point is always the last evaluation, so that improvements are measured against this baseline. Applied across entities with different operating points and characteristics, the disadvantage is that standardization of the metrics is difficult and comparison of different occupants is equally hard, thus failing the requirement for fairness.

In conclusion, the **positive punishment** strategy was selected. Here, an energy penalty is *added* to represent wasted energy, in comparison to an *ideal reference point*. The derivation of the ideal reference point is presented in the next section. This strategy fulfils all the requirements outlined above, and also fits perfectly into the environment of the research, given that the grounds for the derivation of the ideal reference point have can be reasonably derived. Instances of other reinforcement modes can also be layered on top of this base mode. For example, ranges of energy penalty can be classified using traffic-light feedback: green for penalty values close to ideal, amber for regions close to average, and red for farther regions. The traffic light feedback can be thought of as *positive reinforcement* when in the green zone, and *positive punishment* when in the red zone. This feedback system is also employed in this work, with the derivation provided in Section 4.5.9.

3.3 Behaviour Evaluation System Development Approaches: The Rule-Model-Measurement Framework

The task of a behaviour evaluation system is to analyse the effects of occupants' direct and indirect interactions with the energy system of the spaces in which the occupants are located (including interactions that persist after the occupant has left the space), such that the output of such an evaluation system indicates the *appropriateness* of such interactions with regard to energy consumption or other relevant key performance index. In behaviour analysis terms, the behaviour evaluation system defines the *contingency* or dependency between user behaviour and consequences. This evaluation output or consequences can be categorical or continuous, having levels that represent *reinforcement*, *neutral*, or *punishment*. In other words, this behaviour evaluation system provides *evaluative feedback*, which equally contains *descriptive feedback*. As previously mentioned, this combination of evaluative and descriptive feedback has been shown in the literature to be more effective than either type of feedback alone [82, 83].

One of the main contributions of this thesis is the classification of methodologies used to develop a behaviour evaluation system for energy-related gamification-like programmes. To the best of the author's knowledge, this classification does not exist so far in the literature, although elements of it are littered throughout gamification and behaviour intervention studies in the field of energy conservation. From the analysis of the task of quantitative behaviour evaluation and existing literature, the following categories were identified in this thesis into which methodologies for deriving energy-related behaviour evaluation metrics can be roughly classified.

- **Rule-based approach**, where particular actions or states are associated with particular points or ratings on the evaluation scale using a pre-defined set of simple rules.
- **Model-based approach**, in which the ratings on the evaluation scale are derived using models that *faithfully* represent the occupant's environment and interactions with the energy system. The model can be physics-based or data-driven, including using statistical models and machine-learning.
- **Measurement-based approach**, in which the evaluation *output* is measured directly *in energy terms*. The specification here that the output be measured directly in energy terms is important, since both the rule-based and model-based approaches almost always require measurements as input.
- **Mixed-mode approach**, which combines elements of the other three approaches.

The above classification is subsequently referred to as the *Rule-Model-Measurement* (RMM) framework. A generic energy-related behaviour evaluation system architecture is shown in Fig 3.2. From the general architecture, specialized variants can be derived for specific system development methodologies corresponding to one of the above-listed RMM approaches. In the figure, the measured actions and (implied) behaviour of the building occupant are fed into the behaviour evaluator, along with other *operational data* required for the proper functioning of the evaluator during behaviour evaluation. The evaluator, in turn, produces a *behaviour evaluation metric* like gamification points or an energy score, which is the evaluative feedback that serves as the reinforcer or punishment within the operant conditioning procedure. For model-based approaches, *calibration data* is required to create and/or calibrate the energy system model, and for a measurement-based approach, the calibration data could be used for *baselining* prior to interventions.

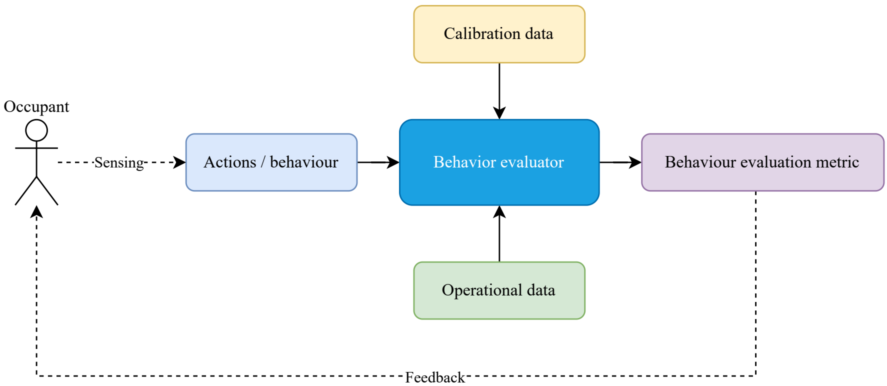


Fig. 3.2 Generic energy-related behaviour evaluation framework architecture.

In Table 3.2, previous energy-related gamification studies are classified based on the Rule-Model-Measurement framework, while highlighting at the same time the operant conditioning mode employed in the study.

Table 3.2 Categorization of previous energy-related gamification work on based on the behaviour evaluation methodology classification developed in this thesis.

Research work	Aim	Derivation of main behaviour evaluation metric	RMM Category	Operant conditioning
MySmartE [75]	Improve proper use of smart thermostat in residential buildings	**Output**: Personal energy score (0 - 100 / worst to best) **Inputs**: System operating mode (heat/cool/off); thermostat setpoint; occupancy state (home/away/sleep) (derived) **Mapping**: Weighted sum (weights are derived from sigmoid curves that send a setpoint to [0, 1] interval based on predefined reference setpoint ranges)	Rule-based	Positive reinforcement
Wemyss et al. [47, 80]	Reduce electricity consumption	**Output**: Energy consumption compared to historical baseline **Inputs**: Room-level electricity meter data (baseline & current) **Mapping**: $Current - Baseline$	Measurement-based	Positive reinforcement / punishment
OrbEEt project [94, 95]	Reduce electricity / heating consumption	**Output**: Display Energy Certificate (A=[0-25] to G>150) **Inputs**: Component-/room-/building-level metering data (electricity, heating); indoor temperature/humidity/luminosity; occupant comfort model ([96]) **Mapping**: Machine-learning model	Model (black-box)	Positive reinforcement/ punishment
Energy Chickens [59]	Reduce electricity consumption	**Output**: Virtual chicken size / health **Inputs**: Component-level electricity meter data (baseline & current) **Mapping***: $f(Current - Baseline)$	Measurement	Positive reinforcement/ negative punishment
ChArGED [74, 76]	Perform energy-saving actions	**Output**: Points **Inputs**: NFC/Bluetooth data, (disaggregated) electricity metering data **Mapping**†: $k * EnergySaved$	Model (black-box)	Positive reinforcement/ punishment
Papaioannou et al. [78]	Perform energy-saving actions	**Output**: Points **Inputs**: NFC/Bluetooth data, electricity metering data to determine if challenge was completed **Mapping***: $f(ChallengeCompleted)$	Rule-based	Positive reinforcement
GReSBAS [77]	Reduce electricity consumption	**Output**: Points **Inputs**: Electricity metering data **Mapping***: $f(ElectricityConsumption)$	Mixed: measurement-/rule-based	Positive/negative reinforcement

† k = constant.
* Mapping function $f(\cdot)$ unspecified.

These approaches, which are discussed in greater detail in the following sections considering the context of this work, have their respective strengths and weaknesses that make them appropriate for different settings. In deciding the methodology to adopt, a key factor to be considered is which of *behaviour modification* and *energy footprint estimation* is to be emphasized. As will be demonstrated shortly, the different approaches outlined above tend to emphasize one or the other. For example, the rule-based approach lends itself better to behaviour modification in general, while the model-based and measurement approaches primarily enable realistic energy footprint estimation. These two emphases are not necessarily compatible. This is because since the relationship between the evaluation system inputs (user actions) and the resulting performance rating becomes more complex and less prone to retrospection as one goes from a rule-based system to a model-based system; the effect of specific user actions becomes increasingly "lost" in the final rating. These considerations are discussed further below in the respective sections dedicated to each of the above-mentioned approaches.

In order to determine the occupant actions to consider in the evaluation system development, we note again that this thesis deals exclusively with *naturally ventilated* buildings, and that the system to be developed relates only to the thermal energy efficiency of occupant behaviour. The interaction of occupants with the thermal energy system of a naturally ventilated building, when considered in terms of energy efficiency, can generally be classified into two main activities:

- control of the heat output of the heating system, and
- space ventilation.

The former is usually effected by means of a **setpoint temperatures** for the heating device (whether provided through manual valves like in conventional non-smart radiators, or electronically as in smart heating systems), while the latter is mainly via the **manual operation of windows** and equivalent fenestration (like openable skylights or external doors). Hence, these interactions naturally form a basis for the evaluation of behavioural energy efficiency in naturally ventilated buildings. Such an evaluation proceeds first with the definition of a reference point which corresponds to an "ideal occupant", followed by the estimation of deviations from this reference, and then translation of this deviation into energy values or some other unified "rating" to be presented to the user. This *deviation-from-ideal* approach fits with the positive punishment operant conditioning approach that is penalty-based, as discussed in the preceding section. Given then that the profile of an ideal occupant is defined in terms of setpoint temperature and window ventilation style, the goal of the evaluation system is to quantify deviations from this ideal profile and to estimate the corresponding penalty. Related to the setpoint temperature is the presence profile of the occupant, which ideally should govern when the given space should be heated, meaning that the setpoint temperature of a space should vary depending on if the space is occupied or not. In our case study, the possibility exists to automatically lower the setpoint temperature in periods of absence for offices that were equipped with smart cloud-controllable radiator valves.

Detailed analysis of these approaches as they could be applied in the context of this work are presented next, while the analysis also strives to highlight generic considerations that apply to other energy-related behaviour evaluation methodologies. In Table 3.2, a categorization of selected previous works according to the proposed scheme is provided.

3.4 Rule-Based Evaluation Approach

In the most basic sense, a rule-based evaluation strategy involves defining a set of simple rules that is used to derive an evaluation metric from predefined evaluation criteria (e.g. setpoint temperature and window ventilation in this thesis). Such rules are based on existing empirical evidence and predefined behaviour aspects to be targeted. As hinted in the previous section, a rule-based system is more effective when applied to target certain specific behavioural patterns, for example discouraging long-term tilted window ventilation or high temperature setpoints during absence, or even promoting the use of an app or other digital interface. This is because the evaluation system is developed starting with the inputs as the focus – the user actions that are targeted by the evaluation system. The progress of the user is then judged in a way that makes it easy to demonstrate the effects that changes in the input user behaviour have on the output evaluation. This condition is explicitly expressed for example in Kim et al. [75]. This retrospective quality of the system demands that the rules relating the input actions to the output evaluation be simple and intuitive, and with minimal interactions between inputs. In the literature, Schakib-Ekbatan et al. [9] applied such a rule-based system for the evaluation of window ventilation based on a predefined optimal ventilation duration, while in the *DataFEE* project,[1] users are awarded an energy behaviour score for predefined user actions, where each action had an arbitrary score attached to it.

In a purely rule-based system using the positive punishment (penalty-based) approach, the deviations from the ideal are separately derived for each evaluated criterion, which deviations can then be shown to users as a collection of ratings – one per evaluated criterion. For example, a separate rating for ventilation efficiency can be shown independent of the rating for setpoint temperature evaluation. Nevertheless, it is preferable to present a unified "rating" to users so that they can keep track of one just quantity. More interested users could then look into the contribution of each criteria to the unified rating. However, deriving such a unified rating means that the constituent ratings should be combined using some rule (e.g. a linear combination using weighting factors). When these contributing ratings represent physical quantities that are related to energy consumption (e.g. manner and duration of window ventilation), then a combination of the ratings would naturally be in units of energy use. Without an appropriate unifying physics-based model, such a combined rating would likely be physically inaccurate in the best case, and in the worst case, could be physically meaningless and potentially not well-behaved (i.e. lead to inconsistent overall ratings). Such a rating without an adequate physical intuition might be difficult to communicate to users. Here lies the main disadvantage of the rule-based system: interpretability of the rating in an intuitive and physically meaningful manner is limited when multiple factors contribute to the rating. Nevertheless, the rule-based system is the simplest to develop, both in terms of its conceptualization and implementation, and in terms of the amount of input data required. The derivation of the inputs for the rule-based approach is discussed next.

3.4.1 Deriving the Setpoint Temperature Deviation

The temperature of an occupied space is the major factor that determines the level of thermal comfort the occupant experiences. To achieve a desired temperature in a heated room during winter, the occupant provides a setpoint temperature to the heating system, which in turn has an inbuilt control system (manual like in non-smart radiators, or electronic as in smart heating systems) that then operates the heating system to achieve and maintain the specified setpoint temperature. The ideal setpoint temperature of an office depends on occupancy, since in order to save energy, unoccupied offices should be heated less than occupied

[1] https://www.ebc.eonerc.rwth-aachen.de/cms/e-on-erc-ebc/forschung/forschungsprojekte2/abgeschlossene-projekte/ bgznzt/-datafee/

offices using a so-called *setback temperature* as setpoint, which is a lower setpoint temperature than for an occupied office. This can be achieved by occupants turning down the radiator valve for manual non-smart radiators, or else some form of smart control system automatically reducing the setpoint temperature during periods of expected or detected non-occupancy.

The derivation of the deviation of the user-configured temperature setpoint from the ideal setpoint is shown pictorially in Fig. 3.3. Succinctly put, the daily deviation is the cumulative of the deviations calculated at each timestep with respect to the ideal setpoint temperature at the timestep, expressed in *degree-minutes*.

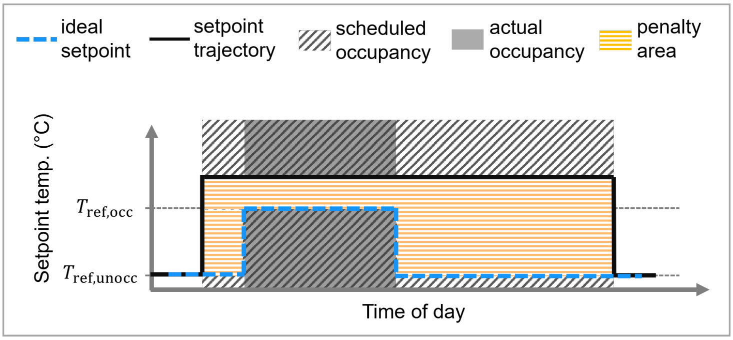


Fig. 3.3 Illustration of the derivation of the setpoint deviation for a hypothetical scenario. Deviations are calculated based on actual presence in the office using a presence-dependent reference setpoint temperature (the "ideal setpoint" temperature in blue dashed lines).

Mathematically, the setpoint deviation $\pi_{\text{sp}}(t)$ at time t, is calculated as:

$$\pi_{\text{sp}}(t) = \max\left(0, T_{\text{sp}}(t) - T_{\text{sp,ref}}(t)\right) \tag{3.1}$$

where $T_{\text{sp,ref}}(t)$ is the ideal setpoint temperature at time t defined as

$$T_{\text{sp,ref}}(t) = y(t) \cdot T_{\text{sp,ref,occ}} + \left(1 - y(t)\right) \cdot T_{\text{sp,ref,unocc}} \tag{3.2}$$

where $T_{\text{sp,ref,occ}}$ and $T_{\text{sp,ref,unocc}}$ are the ideal setpoint temperatures for an occupied and an unoccupied office, respectively, and the $y(t)$ is an occupancy indicator variable defined as

$$y(t) = \begin{cases} 1, & \text{if office is occupied at } t \\ 0, & \text{otherwise} \end{cases}$$

The values of $T_{\text{sp,ref,occ}}$ and $T_{\text{sp,ref,unocc}}$ are derived in Section 4.5.7 of Chapter 4.

The total setpoint deviation for a day (1440 minutes), Π_{sp} in *degree-minutes*, is:

$$\Pi_{\text{sp}} = \sum_{t=1}^{1440} \pi_{\text{sp}}(t) \tag{3.3}$$

The lower bound of zero for each time-point guards against "overcompensation" – an attempt by an occupant to set the setpoint temperature below $T_{\text{sp,ref,occ}}$ in order to gain a possible advantage, invariably possibly violating acceptable comfort standards in the process (e.g. a too low setpoint temperature). The evaluation system is designed to ensure a balance between the competing goals of energy savings and occupant comfort.

3.4.2 Deriving Ventilation Deviation

The main purposes of ventilation are indoor air quality control and temperature control. During the heating season, efficient ventilation seeks to keep the indoor air clean whilst minimizing heat losses. Two possible approaches for deriving an ideal ventilation rule as a basis for estimating deviations are 1) by fixing a daily ventilation quota, and 2) in response to indoor air quality (particularly CO_2 concentration).

The use of "trickle ventilation" (windows opened in a tilted state, or "Kipplüften" in German) during the winter season is discouraged in some previous studies (e.g. [97, 98]), as well as several government campaigns (e.g. [99, 100]). Instead, the so-called shock-ventilation strategy (publicized in German as *Stoßlüften*) is recommended, in which the windows are thrown completely open for a few minutes at a time, several times a day as required [97]. Nevertheless, trickle ventilation is recommended in some cases, for example in bedrooms at night especially when there are several occupants [101]. In reality, the issue with trickle ventilation is *not* that it is inherently wasteful, but rather that in practice people tend to leave the windows in this tilted state for extended periods, while the heating system labours to compensate for the constant influx of cold outside air. Because the heating system is sometimes able to still warm the space to comfortable temperatures under these circumstances, the occupants do not feel the need to close the windows. Also, because such long ventilation durations lead to the walls cooling down, the energy stored in thermal mass of the building is dissipated, leading to more energy being needed to reheat the walls.

Daily Ventilation Quota Approach

One way to derive the ideal ventilation duration for a given day in the heating season is to determine an optimal ventilation duration from models and previous research, such that when occupants ventilate up to this length of time per day, the indoor air quality is maintained without unnecessary heat losses. This type of *quota-based* optimal ventilation duration was applied e.g. by Schakib-Ekbatan et al. [9]. Given the general emphasis on *shock ventilation* in Germany as described in the preceding section, the analysis in this section assumes that the use of shock ventilation to be ideal, and to be preferred to trickle ventilation. However, as will be seen in the following discussion, there is no loss of generality in this assumption, since the analysis in principle applies to a system where shock and trickle ventilation are equally recommended.

Given, then, an ideal ventilation strategy expressed as a quota of $N_{\text{vent,ref}}$ minutes of shock-ventilation per day, the deviation of the occupant ventilation from the ideal strategy is expressed for an entire day in minutes, thus:

$$\Pi_{\text{vent}} = \max\left(0, N_{\text{vent,eq}} - N_{\text{vent,ref}}\right) \tag{3.4}$$

where Π_{vent} is the daily ventilation deviation (in minutes), and $N_{\text{vent,eq}}$ is the number of *equivalent* ventilation minutes by the occupant for the given day. The term *equivalent* indicates that as per the rule-based system prioritizing shock ventilation, a penalty is awarded for the use of trickle ventilation (bottom-hung windows), in which a multiplicity (penalty) factor $f_{\text{pen,trickle}} > 1$ is used in the case of trickle ventilation. Indeed, setting $f_{\text{pen,trickle}} = 1$ makes both shock and trickle ventilation equally recommended, and the possibility also exists in that case to account for the lower air change rate of trickle ventilation compared to shock ventilation by setting $f_{\text{pen,trickle}} < 1$.

Thus, the *equivalent* ventilation minutes of the occupant incorporates the penalty factor, so that:

$$N_{\text{vent,eq}} = N_{\text{vent,shock}} + N_{\text{vent,trickle}} \cdot f_{\text{pen,trickle}} \tag{3.5}$$

where $N_{\text{vent,shock}}$ is the number of minutes of shock ventilation, and $N_{\text{vent,trickle}}$ is the number of minutes of trickle ventilation per day.

An illustration of the ventilation deviation is shown in Fig. 3.4, including the trickle ventilation penalty, $f_{\text{pen,trickle}}$. In the diagram, the first ventilation period used shock (side-hung) ventilation, for which the duration in minutes counts directly towards the daily quota. In the second ventilation period, the trickle (bottom-hung) ventilation is first used, and afterwards shock ventilation. Within this trickle ventilation period, the penalty factor $f_{\text{pen,trickle}}$ is used to accelerate the depletion of the ventilation quota as a deterrent. Beyond this depletion point (the end of the green hashed area in the figure), the ventilation counts as excessive.

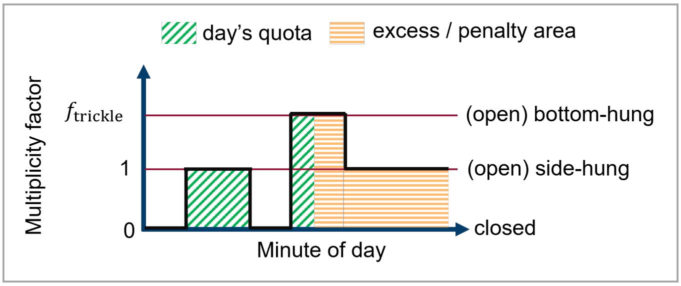


Fig. 3.4 Illustration of the derivation of the ventilation deviation for a hypothetical scenario. The ventilation duration beyond the allowed period is penalized.

The lower bound of zero minutes in Eq. 3.4 guards against overcompensation. In other words, the system does not reward under-ventilation in order to forestall possible deliberate violations of good indoor air quality by occupants in order to get good ratings. However, one apparent disadvantage of the fixed-quota ventilation approach is that it does not behave well when the occupant density in the room is high, e.g. in meeting rooms and high occupant-density multi-person offices. More discussion about this limitation is provided in Section 4.5.6, where this approach is implemented.

Indoor Air Quality Approach

In this approach, the emphasis is on ventilating is such a way as to maintain good indoor air quality (IAQ) while minimizing energy wastage during the heating season. Consequently, the style of ventilation (trickle ventilation or shock ventilation) is immaterial. The basic premise is that energy wastage can only be considered after basic indoor health needs have been met, and that the energy expended in meeting these needs cannot be regarded as wasted. Furthermore, it is assumed that possible differences in the ratio of the *hygienic air change efficiency* to the *energetic air change efficiency* for the two ventilation styles, are ignored (see [102] and Section 4.5.3 for further discussion of hygienic vs. energetic air change). The IAQ can be measured by proxy by CO_2 concentration; therefore, the CO_2 level in the room is used in this analysis for triggering ventilation. Conceptually, the approach works by summing up all the time periods in the day when the windows were open *and* the CO_2 concentration in the room was below a given threshold that indicates good indoor air quality. That is, the approach considers periods when the windows were left open although the room air quality was good as indicated by a low CO_2 concentration. For each transition into this low-CO_2 region (from a higher concentration, e.g. shortly after opening the windows), a *buffer period* of a few minutes is allowed, within which the room freshness is supposed to "settle" before penalties are considered. This approach of determining ventilation deviation was not applied in this thesis as at the time of the experimental run, since the initial goal of the ventilation evaluation in this thesis was to conform to the widespread recommendations for shock ventilation instead of trickle ventilation, which the IAQ approach cannot support. However, the IAQ approach is being implemented for the next version of the behaviour evaluation engine, and is described more explicitly below.

Mathematically, given a lower bound CO_2 concentration, C_{lb} (in ppm), below which open windows should be closed during the heating season, identify m time periods $D_{\mathrm{CO2<lb}} = \{D_1, ..., D_m\}$ in the evaluated day such that for each period $D_j \in D_{\mathrm{CO2<lb}}$, the CO_2 concentration in the room continuously stays below C_{lb} ppm for the entire duration. Specifically, let t be an index over the minutes of the day, i.e. $1 \le t \le 1440$, and let the duration of D_j be d_j minutes starting at time $t = t_j$ up to $t = t_j + d_j$, then for D_j it holds that

$$C_{\mathrm{room}}(t) < C_{\mathrm{lb}} \qquad \forall t \in \{t_j, ..., t_j + d_j\}$$

where $C_{\mathrm{room}}(t)$ is the CO_2 concentration in the room (in ppm) at time t.

Denoting the state of the window for each minute $t \in D_j$ as $y_{\mathrm{win}}(t) = 1$ if the window is open (bottom- or side-hung) at time t and $y_{\mathrm{win}}(t) = 0$ otherwise, and defining a *buffer period* of N_{buf} minutes, the **ventilation deviation** for the period D_j can then be computed as:

$$\Pi_{\mathrm{vent},j} = \max\left(0, \quad \sum_{t=t_j}^{t_j+d_j} y_{\mathrm{win}}(t) \cdot \Delta t \; - \; N_{\mathrm{buf}} \cdot y^{-}_{>\mathrm{ub},j}\right) \tag{3.6}$$

where $\max(0, \cdot)$ indicates that the expression is bounded below by 0 to avoid overcompensation and ensure that periods of correct ventilation cannot be used to offset the penalties from periods of incorrect ventilation, and Δt is the length of each time step in minutes. The indicator variable $y^{-}_{>\mathrm{ub},j}$ factors in the history of CO_2 concentration just before period D_j. Specifically, $y^{-}_{>\mathrm{ub},j}$ indicates if the room CO_2 concentration was above C_{ub} ppm in the N_{buf} minutes before the start of period D_j, where C_{ub} is the CO_2 concentration (in ppm) above which ventilation is recommended, and $C_{\mathrm{ub}} > C_{\mathrm{lb}}$. That is, for period D_j,

$$y^{-}_{>\mathrm{ub},j} = \begin{cases} 1, & \text{if } \exists\, t \in \{t_j - N_{\mathrm{buf}}, ..., t_j - 1\} \text{ such that } C_{\mathrm{room}}(t) \ge C_{\mathrm{ub}} \\ 0, & \text{otherwise.} \end{cases} \tag{3.7}$$

The purpose of $y^{-}_{>\mathrm{ub}}$ is to harden the system against a particular "exploit", where an occupant opens the window when the CO_2 concentration is between C_{lb} and C_{ub} until the concentration has stayed below C_{lb} for *almost* N_{buf} minutes, and then closes it briefly until concentration is slightly above C_{lb} again, and then repeats the cycle. Thus, the occupant can take advantage of the N_{buf} period each time to avoid penalties, without actually needing to have opened the window at all. In other words, using $y^{-}_{>\mathrm{ub}}$ ensures that the upper threshold C_{ub} is used for opening the windows.

A scenario illustrating the indoor air quality approach is shown in Fig. 3.5, also demonstrating the implication of $y^{-}_{>\mathrm{ub},j}$. Specifically, the mitigation of the described exploit scenario is exemplified in Period D_2 of the figure, where the buffer period is *not* considered (i.e. $y^{-}_{>\mathrm{ub}} = 0$), since in the N_{buf} minutes preceding D_2, the CO_2 level was below C_{ub} ppm (red section of the curve). In other words, the window opening at point B in the figure was premature, and $y^{-}_{>\mathrm{ub},j}$ ensures that the penalty starts immediately without a buffer period. In Period D_1, the buffer period is considered given that the CO_2 concentration in the N_{buf} minutes preceding D_1 exceeded C_{ub} (as shown by the blue section of the curve), which confirms that the window opening at point A in the figure was indeed appropriate.

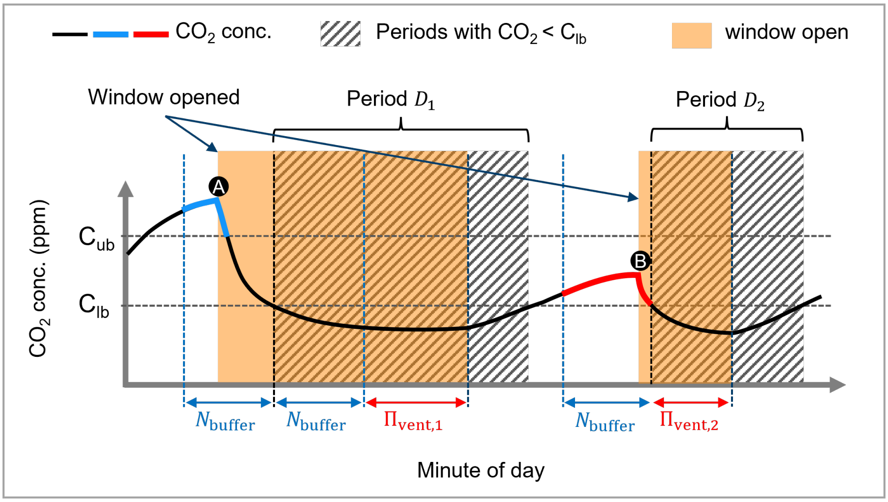


Fig. 3.5 Illustration of the derivation of the ventilation deviation for a hypothetical scenario using the indoor air quality approach. In Period D_1, penalty is applied only if the window is still open beyond N_{buf}, since the CO_2 level in the relevant period before D_1 was above C_{ub} ppm (blue section of curve) before the window was opened (point A). In Period D_2, N_{buf} is not considered, since before D_2, the CO_2 level was less than C_{ub} ppm (red section of curve).

The total ventilation deviation for the given day is then simply the sum of deviations for all m periods in $D_{CO2<lb} = \{D_1, ..., D_m\}$. Mathematically,

$$\Pi_{vent} = \sum_{j=1}^{m} \Pi_{vent,j} \tag{3.8}$$

where Π_{vent} is the total ventilation deviation (in minutes) for the evaluated day.

3.4.3 Weighting of Evaluation Criteria

In order to provide a unified rating, the input deviations (for setpoint temperature and ventilation) can be assumed without loss of generality to be linearly combined with weighting factors w_{sp} and w_{vent} respectively. Hence, we have the daily rating in the rule-based system as

$$\text{Daily Penalty Rating,} \quad \Psi = w_{sp} \cdot \Pi_{sp} + w_{vent} \cdot \Pi_{vent} \tag{3.9}$$

The derivation of the weighting factors w_{sp} and w_{vent} is up to the designer of the system. A possible strategy is to choose the weighting factors such that the magnitude of both deviations in the average case are similar, so that each deviation contributes roughly half of the final penalty value in the average or modal case, effectively normalizing the contributions of each deviation. The recommendation to use the "average" or "modal" case for normalization (i.e. the case most commonly encountered in the real world) instead of selecting the midpoints of the deviation ranges is because the distribution of the ventilation deviation is expected to be different from that of the setpoint temperature deviation in practice. In other words, while the setpoint deviation is expected to have an approximately normal distribution, the ventilation deviation is expected to be skewed with a long tail, i.e. to have a narrow band of a few minutes to a few hours where most deviations are clustered, but some extreme cases where windows are left open for (almost) the entire day. Game-theoretic approaches can be applied also for the derivation of the weights, for example like in Papaioannou and Stamoulis [103].

3.4.4 Data Requirements for Rule-Based Approach

Fig. 3.6 shows the data requirements for the rule-based approach. Few inputs are required, making this approach very simple and attractive. Specifically, the four required input quantities are represented by the boxes with the horizontal dashed "input lines" in the top part of Fig. 3.6, namely: window state, CO_2 concentration (when using the indoor air quality approach for determining deviations, as described in the previous section), presence (i.e. "binary" occupancy), and heating setpoint temperature. Occupancy can be derived from direct sensing (presence detection through PIR sensors) or indirectly (via environmental sensing, in this case CO_2 concentration). Also, only *current* measurement data for these inputs is required for performing the evaluations; there is no need for historical data as is required for calibration. If the modal or average case weighting approach is used when combining the ratings into a single value, it suffices to use generic artificial data obtained through simple experience or else data that is generally available.

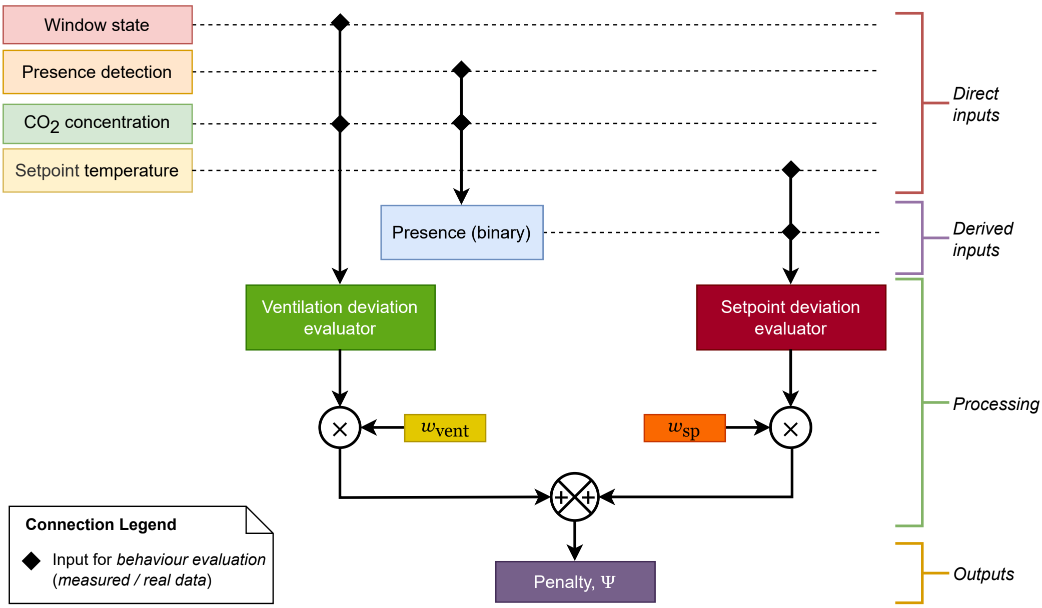


Fig. 3.6 Data requirements for the rule-based evaluation approach. The filled diamond symbols at the intersections with the dashed horizontal "input lines" mean that the corresponding input is required as *measured data* during the *behaviour evaluation phase*.

From the input data, the ventilation and setpoint deviations are calculated in the corresponding blocks of Fig. 3.6, and then combined using their respective weighting factors w_{vent} and w_{sp} to produce the final penalty, Ψ.

3.5 Model-Based Evaluation Approach

In the model-based approach, the behaviour of occupants is evaluated holistically based on the energy implications of their actions. In this case, the environment of the occupant is modelled with reasonable accuracy as a virtual office (essentially a "digital twin"), and the relevant inputs to the model describing both the occupants' actions and the surrounding environment are adequately measured or estimated (see Fig. 3.7). Note that in the RMM classification, the model could be physics-based, data-based or hybrid (see Fig. 3.12 in Section 3.9). At the end of each evaluation cycle, the actual energy consumption of the office is calculated via measurements obtained from the real energy system, which represent the cumulative effect of the occupants on the energy consumption of the room. Alternatively, the actions of the real occupant can be "replayed" within the virtual office, so that the energy consumption of the real office is then approximated by that of the virtual office. This choice is of measured vs. model-calculated demand is shown as a switch in Fig. 3.7.

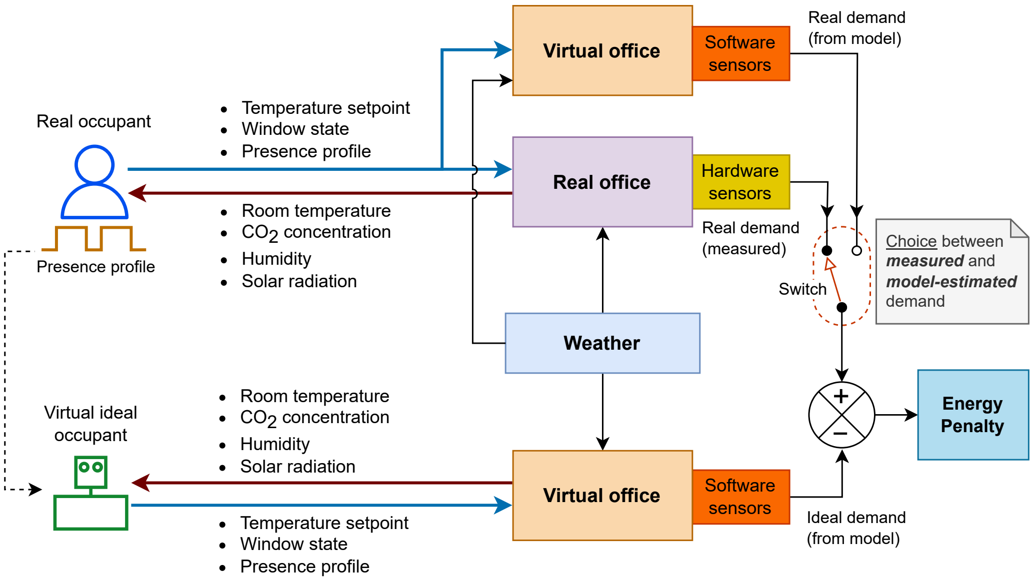


Fig. 3.7 Scheme for model-based occupancy behaviour efficiency evaluation. A virtual occupant acts on a virtual office to emulate ideal actions and consequently ideal thermal energy demand, given the same environmental circumstances as the real occupant. The difference between the real-world energy consumption of the room (either measured or model-estimated) and that of the ideal virtual counterpart provides a behaviour efficiency metric for the real occupant.

In order to derive the equivalent ideal behaviour for the real occupant, a virtual ideal occupant with an identical occupancy profile as the real-world occupant is subjected to the same prevailing real-world "drivers" (weather parameters) in the virtual office model, as depicted in Fig. 3.7. The ideal actions of the virtual occupant in response to the environmental drivers are then introduced into the virtual office as disturbances, whereby the virtual occupant acts to maintain some pre-defined notion of optimum comfort (by adjusting the setpoint temperature and operating the windows). The thermal energy demand of the virtual office is then calculated, and the difference to that of the real-world office represents the energy penalty. The so-called virtual occupant is effectively a controller running within the virtual office, which then controls virtual actuators that mirror the control opportunities available to the real-world occupant.

Papaioannou et al. [76] is a rare example of the application of the model-based approach in a fully gamified setting dealing with direct energy savings, and which mirrors the sequence of the previous paragraph. In that study, at the instant of user action, the analytics engine first estimates the energy footprint of the user's action by disaggregating currently measured electrical power data to identify the change in consumption caused by the action. The action itself is reported to the analytics engine when the user swipes their phone on an NFC device attached to e.g. the appliance the user interacted with. The engine then awards points by comparing the estimated actual energy use with that of a model-based hypothetical ideal scenario using the same boundary conditions as in the real case. The model is data-based and was developed from clustering historical data. Additionally, this deviation-from-ideal point-awarding system was specifically chosen in that study to counter the previously mentioned unfairness inherent in deviation-from-baseline evaluations, so that "green" consumers are not penalized unduly [76].

The model-based approach has the advantage that it provides the highest fidelity in terms of estimating the real-world energy implications of occupant behaviour. It can account for all the relevant factors and their complex interactions, including the effects of solar gains, thereby being able to capture the necessary differential treatment of the sun-side versus the shadow-side of a building for example, unlike the rule-based approach. Additionally, the use of sunblinds can be integrated, along with its effect on the state of the office and on occupant comfort. Furthermore, the developed models can be applied to several other use cases, including for Model-Predictive Control (MPC) of the building energy systems.

On the other hand, the model-based approach suffers from some drawbacks. First is that it requires significant effort to develop the models for the rooms, especially since the models are required to have acceptably high accuracy, so that the measured energy consumption of the real office under consideration can be reliably compared with that of the virtual office. The accuracy requirement can be somewhat relaxed when the virtual office is used as a proxy for calculating the thermal energy demand of the real office, instead of real measurements (as depicted by the "Switch" on the right side of Fig. 3.7). In this case, the virtual office is used for both the real and ideal demands, thereby ensuring compatibility of results and possibly reducing the number of sensors required to be installed in the real building. For calibration purposes, it would still be necessary obtain relevant real-world data in any case. For the 12 buildings covered in this thesis representing nearly 500 rooms, the modelling workload would be significantly high compared to the other approaches.

Another disadvantage is that the model-based approach scales poorly, since each office is essentially unique and has to be modelled and/or parametrized separately. This limits the applicability of the approach in large-scale use-cases. Furthermore, the computational resources that are required to perform the evaluations, in terms of memory, CPU load, and time, are much more significant than for other approaches, especially when evaluation results are required within relatively short time frames. To be factored in, also, is the fact that the input data also needs to be pre-processed before each evaluation run, and with higher data requirements, more computation is required for the pre-processing phase.

In terms of the "ease of retrospection" quality desired of the evaluation approach (see Section 3.2), it is not straightforward to relate the derived energy performance to the contributing occupant-related input actions (ventilation and setpoint temperature). This is because several other factors that have a complex relationship with the occupant input actions vary alongside these actions, for example variations in ambient temperature from one evaluation day to another, time of arrival / departure, and number of occupants performing the actions (having more occupants present in a multi-person office increases the likelihood of more windows being opened at once during a ventilation regime, due to the lower average "distance to control" [104]). A rule-based system, on the other hand, thinks first in terms of the rules, so that whatever the final evaluation becomes, it can be traceable to the rules since it was derived directly from them in a predefined and tractable manner.

Finally, a purely model-based recommendation system cannot easily function as an *online* recommendation system, i.e. a system which provides real-time feedback about energy saving actions based on the current state of the office. To demonstrate limitation, not that in an online recommendation system, it is

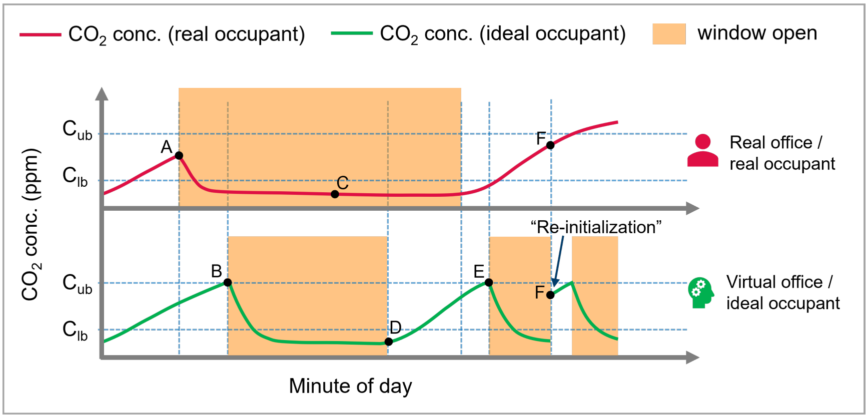


Fig. 3.8 Illustration of the state divergence that makes the model-based approach unsuitable for use as an online recommendation system. The

required that every user action or inaction can be evaluated instantaneously in terms of energy efficiency, for example when a window has been left open for too long, so that feedback about corrective action can be provided. For a purely model-based system, in which the ideal scenario is defined by the actions of the ideal occupant in the virtual office, this means that the virtual office would somehow track the real office in real time and at the same time provide the correct course of action based on the ideal actions of the virtual occupant. However, by definition, such a simultaneous tracking *and* recommendation is difficult if not impossible in practice, since the state of the virtual office depends on the actions of the virtual occupant, which in turn would be different from those of the (non-ideal) real occupant by definition. In other words, the state of the virtual office diverges from that of the real office as a result of different control actions by the ideal occupants than the real occupants, so that continued tracking can only be possible when the state of the virtual office is periodically re-initialized to the current state of the real office. Indeed, a similar kind of difficulty was reported in Papaioannou et al. [76], where the re-initialization approach was taken as a solution.

Fig. 3.8 illustrates the problem, borrowing the indoor air quality-based ideal ventilation strategy of Section 3.4.2. At point A in the figure, the real occupant opens the windows (prematurely), causing a drop in CO_2 concentration. Meanwhile, the ventilation should have been done at B, according to the ideal occupant, all things being equal. Hence at A, the states of the virtual and real offices begin to diverge. If the states continue to diverge as shown in the figure, then at point E, the recommendation of the ideal occupant is to open the windows, but this is inappropriate for the real occupant since the CO_2 level in the real room is well below the ventilation threshold. At point F, the virtual office is re-initialized to the state of the real office. However, it diverges again shortly afterwards as the ideal occupant correctly opens the windows while the real occupant does not. Furthermore, it might seem reasonable to have recommended closing the windows to the real occupant at time D based on the action of the ideal occupant, but this does not work in general since point B could have also come several hours later (due to a very slow CO_2 rise), meaning that the windows opened at point A do not get recommendations for closing, until several hours later when point D is reached in the virtual office. Indeed, the correct recommendation for the real occupant to close the window after the initial opening at A, is at point C. However, no actions of the ideal occupant correspond to this timepoint due to the divergent states. Hence, it is not possible to use the actions of the virtual ideal occupant as source of recommendations for the real occupant. Even the re-initialization could be complicated in this, since the state of the room is not described only by its current state, but by its history. For example, the duration that the windows have been opened, and not just the current CO_2 concentration, is required to know when to shut them.

One way to resolve the issue described above is to incorporate into the model-based system a separate rule-based recommendation sub-system, which monitors the current status of the real office and uses the same decision engine as the ideal occupant, such that the violation of those rules results in corrective recommendations. For example, the recommendation sub-system could have recommended closing the windows at point C in Fig. 3.8, irrespective of the actions of the ideal occupant and the state of the virtual office.

3.5.1 Data Requirements for Model-Based Approach

As depicted in Fig. 3.9, the data requirements for the actualization of the model-based behaviour evaluation approach are extensive. Inputs required for the model development phase and for the behaviour evaluation phase are differentiated in the figure, as well as those required in both phases. Additionally, the required "source" for the data is indicated, either as actual (current/historical) data, or as artificial (generic) data. At the development phase, extensive historical data is needed for calibration of the models, apart from real building construction data like the building geometry, thermal transmittance (U-values of the envelope), air infiltration rates, characteristics of the heat transfer equipment (wall-mounted radiators and floor heaters), and energy supply information like flow and return temperatures for heating. Additionally, accurate occupancy data at the *count* level of occupancy estimation (see Section 3.10) is required, with the attendant need for potentially large-scale calibration of the occupancy estimation models (if environmental sensing based on CO_2 concentration is used, rather than PIR presence sensors or some other direct measurement approach). The models could also be developed such that they are online-calibrated as more data becomes available; however, this requires more complex implementations and a possible "learning stage" until the models are reasonably accurate to be deployed. After the deployment of the model within the evaluation system, the data requirements also remain significant – weather data including solar gains, position of window blinds, and number of occupants per time.

3.6 Measurement-Based Evaluation Approach

The measurement-based approach relies on real-world measurements of relevant parameters in order to evaluate the behavioural efficiency of the occupant. A baseline measurement is usually taken prior to the application of the behaviour interventions, against which subsequent measurements are compared to determine if improvements occurred or not. As previously mentioned, the preceding rule-based and model-based approaches also involve measurements from which the behaviour evaluation metrics are derived. However, while these overlaps exist between the rule-/model-based approaches and the measurement-based approach, the distinguishing factor of the latter is that the resulting measurements are directly the metric of interest (when compared to the rule-based approach that transforms the measurements into a metric), and that these measurements are obtained directly from sensors.

The disadvantage of the purely measurement-based system is that the metric of interest must be instrumented directly, e.g. windows sensors if window ventilation were to evaluated. A model-based system, on the other hand, can derive window interactions via other measured parameters like indoor temperature and

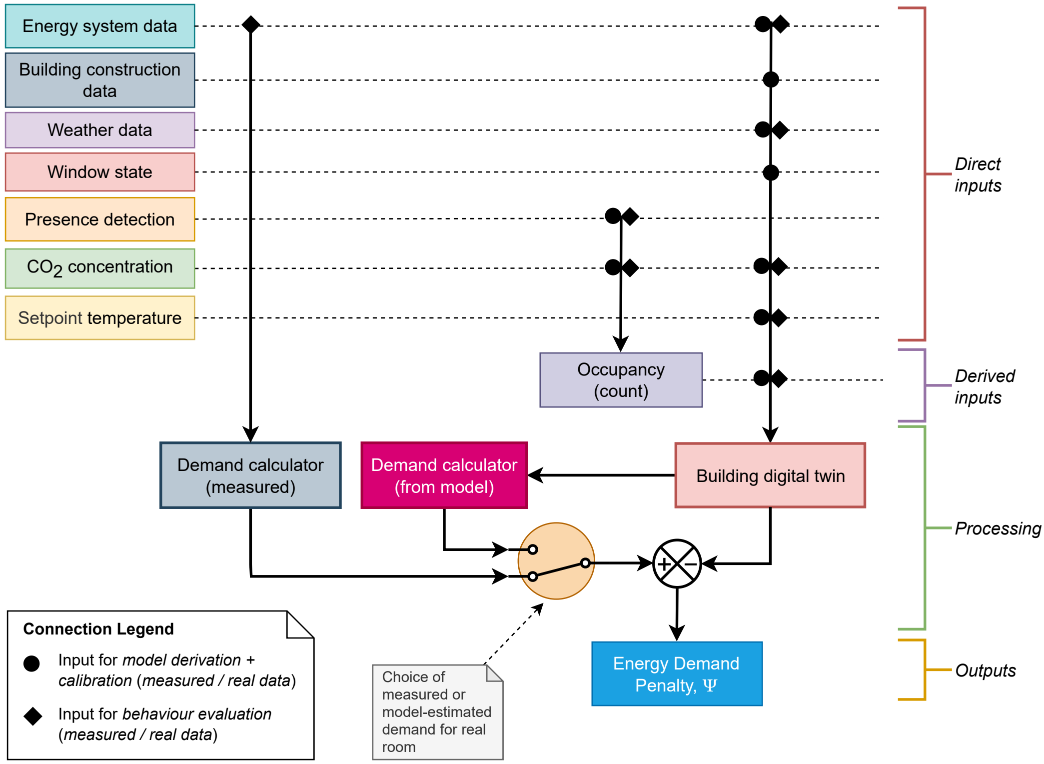


Fig. 3.9 Data requirements for the model-based evaluation approach. The square and diamond symbols at the input intersection lines denote, respectively, the input being required only during the model derivation phase, or only during the behaviour evaluation phase. The filled version of the symbols means that the data comes from actual measurements and not artificially generated data (e.g. for model development and calibration).

CO_2 concentration using the system model. Nevertheless, the measurements of the system could also be derived by so-called *soft sensors* without needing as many real sensors.

Another disadvantage of the measurement-based evaluation approach is that the ideal behaviour is not easily determined, since the ideality of the baseline behaviour is generally not guaranteed. Indeed, the measurement-based approach directly leads to the problematic deviation-from-baseline evaluation that potentially penalizes already-efficient behaviour.

3.7 Mixed-Mode Approach: Model-augmented Rule-based System

The mixed-mode approach seeks a middle ground between the rule-based approach and the model-based approach. Here, the rule-based approach is augmented with information from a physics-based model, such that the rules can be reliably expressed in a physically sensible and relevant manner. In this approach, the physical model of a reference room is first developed, parameterisable by ambient temperature, occupancy schedule, window ventilation, and setpoint temperature (see Fig. 3.10). The thermal energy demand of the reference room is obtained through the model for combinations of relevant input parameters that describe occupant interaction scenarios. The input variable combinations are determined through systematic state-space sampling and the results of simulating each input condition are stored, such that the thermal demand for any arbitrary input combination can subsequently be determined via interpolation of the stored results during the behaviour evaluation phase (an example of such a result in shown in Fig. 4.10).

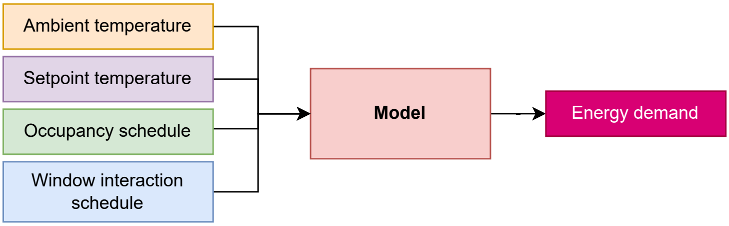


Fig. 3.10 Generic scheme for deriving weighting factors for rule-based evaluation using a simulation model.

Equipped with this state-space sampling result, deviations from the ideal scenario, expressed in terms of deviations in setpoint temperature and ventilation duration respectively, can then be converted into a surplus in energy consumption by summing up the energy consumption difference between the ideal operating point for the given scenario (e.g. 19 °C for temperature setpoint of an occupied office) and the actual operating point of the office for each time step. Thus, the deviations in the two input criteria are derived as in the rule-based system, but the unified "rating" is now reliably and sensibly derived from the physical model, allowing interpretability of the rating and easier communication to users. To bridge the differences between the energy characteristics of different rooms based on their geometrical and construction properties, several such reference models can be derived, with each one representing a set of real-world offices with similar enough geometrical properties. Hence, in the evaluation of each office, the best-fitting reference model is used to derive the energy performance.

In this work, the mixed-mode approach is used, albeit with only one reference model. The implementation details in this thesis are discussed in Chapter 4. The evaluation metric of the hybrid approach has the following physical and mathematical properties.

1. **Over-compensation protection** The reference values for the evaluated criteria (i.e. the values for the ideal occupant) form lower bounds on the criteria. This implies that occupants cannot improve their rating by setting a lower setpoint temperature than the ideal, and / or by ventilating for less than the ideal daily ventilation duration. Hence, occupants are not rewarded for possibly comfort- or air quality-violating actions.

2. **Meaning of derived energy penalty** The derived energy penalty can be easily appreciated conceptually, since it expresses how occupants would naturally expect that their energy efficiency is expressed in energy units.

3. **Non-dependence of ventilation on presence** Since the ideal ventilation duration N_{ref} does not depend on presence, an unoccupied office can theoretically be ventilated for up to N_{ref} minutes *without* accruing penalties, provided the setpoint temperature is ideal (i.e. at $T_{\mathrm{sp,ref,unocc}}$) throughout.

3.7.1 Data Requirements of the Mixed-Mode Approach

The mixed-mode approach requires significantly less data than the model-based approach, both in the specific data required, as well as in the source of the data (actual measurements vs. artificial/generic data) (see Fig. 3.11). The most significant data requirement of the model-based approach which is completely

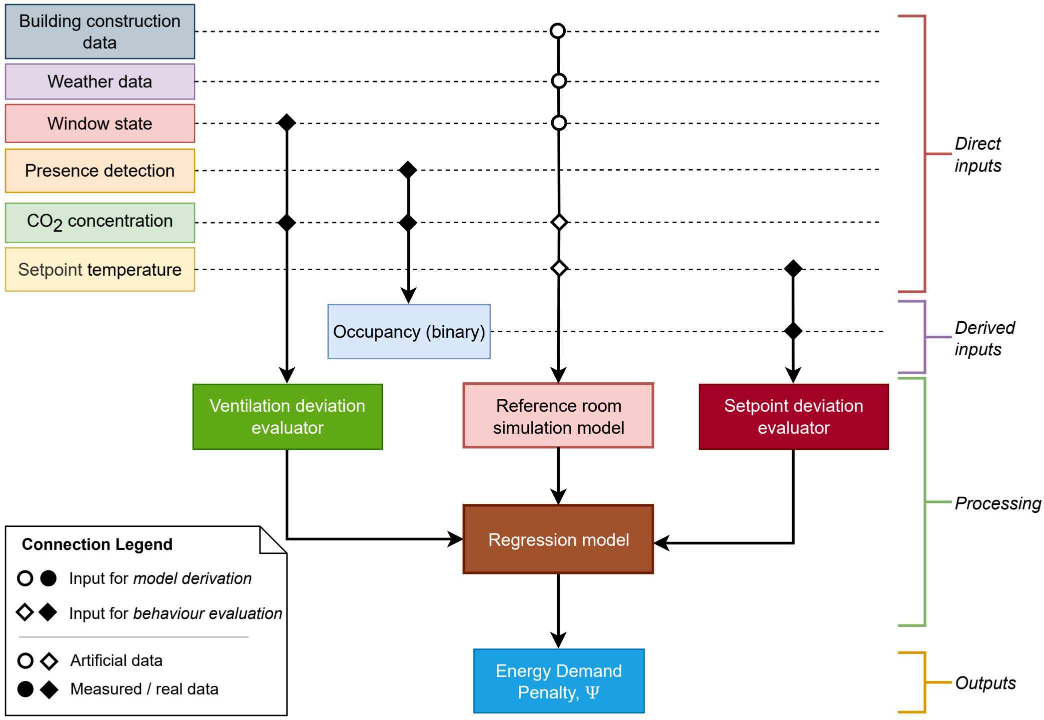


Fig. 3.11 Data requirements for the mixed-mode approach. The square and diamond symbols at the input intersection lines denote, respectively, the input being required only during the model derivation phase, or only during the behaviour evaluation phase. The filled version of the symbols means that the data comes from actual measurements and not artificially generated data (e.g. for determining occupancy).

absent in the hybrid approach is energy system data, since the evaluations in the hybrid approach do not depend on the real energy consumption of individual rooms.

3.8 Design of Experiment

In this section, the experiment hypotheses are presented, followed by the experiment setup that encompasses the buildings involved in the experiment, the experiment variables and groups, and the allocation of buildings and offices to experiment groups. Thus, the means for addressing research question *Q3* of Chapter 2 is elaborated here.

3.8.1 Experimental Variables and Feature Groups

The experiment features which form the basis of the variables that are manipulated in the experiment, are based on the developed software, collectively called the *Energy Dashboard Suite*. A summary of the functions of each of these software applications is given in Table 3.3, while more details are provided in the next chapter.

The test rooms are divided into experiment groups. Each group represents a particular combination of experimental variables. Seven experimental variables are incorporated in the design of the experiment, where the variables correspond to features of the deployed system that are enabled or disabled in each

Table 3.3 Overview of the purpose of the applications in the Energy Dashboard Suite.

Application	UI Type	Purpose
Campus Viewer	Web-based, public	Visualization of the FZJ campus energy demand for heating, cooling, and electricity at the building level.
JuControl	Web-based, public	Room-level visualization and, for selected buildings, possibility to control the heating system of the building. Also serves as the gamification and recommendation platform.
Juracle	None	Energy-related behaviour evaluation engine that powers the gamification aspect of JuControl.
ALICE	Web-based, internal	Mini-language and tool for describing the geometrical aspects of rooms and room components. Generates corresponding interactive room diagram and links sensors to room components.

experiment group. These seven features can be classified into three broad categories: JuControl interaction; evaluation (gamification) and recommendation. The full set of features is as follows.

- **JuControl View**: whether or not the measurements for the room are available in JuControl. This feature is a prerequisite for the availability of several other features, so as will be discussed later, it was enabled for all buildings in the experiment. Note that no behaviour improvement-related functionality such as energy evaluations are included in this version.

- **JuControl Control**: whether or not the heating system in the office can be automatically controlled using the expected presence schedules that occupants supplied in JuControl. This requires that the building be equipped with smart radiators which are cloud-controllable, according to Table 3.6.

- **Setpoint Temperature Evaluation**: whether or not the setpoint temperature for room heating is considered in the room evaluation (as described in Section 3.4.1). The evaluation of setpoint temperature is based on *actual* presence (which could differ from the scheduled presence provided by the user in JuControl), in order to ensure a reality-based evaluation.

- **Window Ventilation Evaluation**: whether or not the window ventilation strategy is considered in the room evaluation (as described in Section 3.4.2). As previously described, in line with recommended practice in Germany, tilt-ventilation is discouraged with using an *arbitrary* penalty factor.

- **Recommendations**: whether or not the office occupants receive real-time recommendations about energy savings via email, specifically about exceeding of alloted setpoint and / or ventilation quotas, and to discourage tilt ventilation. Recommendations were activated after the experiment had already commenced.

Based on these features, four (4) experiment groups were formed as shown in Table 3.4, to which offices would be assigned. These groups were selected in order to optimize the use of the limited sample size and to deal effectively with the most relevant questions for this research. In order to provide *collaborative* as well as *competitive* social elements, offices are first grouped into teams, and then teams are assigned to experiment groups (see Section 3.8.4). The main hypotheses to be tested by the experiment design, as well the strategy for testing the hypotheses using the teams and experiment groups, are discussed in the next subsections.

Table 3.4 List of all experiment groups and their associated feature sets.

Feature Group	Features				
	View	**Control**	**Setpoint Eval.**	**Ventilation Eval.**	**Recommendation**
A	×	×	×	×	×
B	×		×	×	×
C	×			×	×
D	×				

3.8.2 Experiment Hypotheses

The design of the experiment seeks to directly investigate the following hypotheses related to user behaviour and energy performance. The *energy penalty* behaviour evaluation metric developed in Chapter 3 is the basis for judging "energy performance".

H_1 (Effect of evaluations / recommendations) Offices with evaluation (and recommendation) enabled will have better energy performance than those without evaluation / recommendation.

H_2 (Effect of active participation) Within a given evaluation-enabled team, offices in which users actively interacted with the developed behaviour intervention systems will have better energy performance than those in which users did not.

The basis of measuring "active participation" in H_2 is *JuControl-activation* of an office, which is explained in Section 3.8.5. A further hypothesis based on the development of scheduled-based heating control in the pilot building, Building B-01, is proposed below as Hypothesis H_3, in which the performance of the building is expected to improve in 2023 (with automatic heating control), compared to its performance a year prior in 2022 (before the implementation of the automatic control).

H_3 (Effect of automatic heating control) For Building B-01, in which an occupancy schedule-based automatic controller was implemented, the performance of the building in 2023 post-intervention will be better than its previous performance in 2022.

The premise for hypothesis H_3 is that, whilst there were no structured experiments performed in this building in the months before the main experiment (the pilot phase), and neither were the gamification and recommendation functionality present in that period, the occupants of that building were able to leverage the available energy-saving features of JuControl regarding automatic temperature control linked with presence. Specifically, they were motivated to use the JuControl calendar to specify their presence schedule for use by the heating controller, thereby creating the opportunity to save energy when the offices were unoccupied. A detailed discussion of the effect of this leveraging is presented in Section 5.5.

For each of the above-mentioned hypotheses, the corresponding *null hypothesis* asserts that the converse is true. Specifically, for H_1, the null hypothesis is that teams with evaluation will have similar or worse performance than teams without it. For H_2, the null hypothesis asserts that *JuControl-activated* offices within a team with evaluation would perform

Finally, the level of user engagement will be analysed to identify factors that contributed to its improvement.

3.8.3 Experiment Setup

The experiment setup consists of 12 instrumented buildings, in which JuControl was enabled. These buildings range in construction year from 1969 to 2014, and total floor area ranging from 700 m^2 to almost 7000 m^2. The basic construction data of the buildings is presented in Table 3.5 below.

Table 3.5 Building construction data for the buildings of the experiment.

			Area (m^2)			
Building ID	**YoC**[a]	**Num. Floors**	**Floor**	**External Walls**	**Roof**	**Windows**
B-01	1976	2	999.4	442.2	383.1	136.7
B-02	2014	4	2,003.1	838.8	767.8	291.4
B-03	1979	2	1,009.0	446.2	386.8	138.1
B-04	2009	1	701.7	796.0	201.7	93.0
B-05	2014	4	3,309.5	1,331.7	951.5	503.4
B-06	2004	3	4,235.5	1,671.2	1,623.6	658.5
B-07	1969	4	3,438.8	2,124.7	1,318.2	524.8
B-08	1990	2	858.6	384.6	329.1	115.8
B-09	2007	2	737.0	334.1	282.5	98.1
B-10	2011	3	1,769.4	748.3	678.3	254.6
B-11	2010	2	1,209.9	527.3	463.8	168.3
B-12	1967	3	6,923.1	2,627.2	2,653.8	1,124.4

[a] Year of construction

Table 3.6 shows the level of instrumentation in the building that was available at the time of the experiment, and a rough timeline for the availability of the necessary instrumentation. It is important to note that some of the buildings have instrumentation that is not accessible via the hardware ICT platform (which was discussed in Section A.1.2), for example due to no existing connection between the building management system and the ICT platform. For the purpose of this thesis (and in the table below), it is assumed that these features do not exist in the said buildings. The indicated dates are also the dates when data from the devices became available in the ICT platform and hence accessible for this thesis, and not necessarily when they were installed in the buildings. Having a "radio-controlled radiator" as indicated in Table 3.6 implies that the setpoint temperature which drives the heating in the offices can be accessed via the ICT platform. The JuControl radiator control column indicates that the automatic controller (reported in [105]) works with JuControl to manage the heating in the building.

The base unit of the experimental setup is the "room", consisting of both meeting rooms and staff offices in the selected buildings. JuControl was available for 496 rooms across 12 buildings. However, two buildings (Building B-11 and Building B-12) are not included in the experiment analysis of this thesis since they were not assigned to experiment groups that fall within the scope of the thesis. Thus, 386 rooms in the remaining 10 buildings were included in the experiment design relevant to the thesis and the subsequent analysis of results. (Note that the quoted number of rooms do not necessarily indicate the total number of rooms in all the buildings, since spaces that are neither offices nor meeting rooms (e.g. laboratories and restrooms) are excluded.) Nevertheless, the two excluded buildings are part of the analysis of JuControl usage presented in the "user engagement" results section (Section 5.2 of Chapter 5: Run of Experiment and Analysis of Results).

Table 3.6 Level of instrumentation of buildings considered in the study, with approximate dates of availability of the features in parenthesis. All the window/door and environmental sensors were installed as part of the LLEC project, along with the radio-controlled radiator valves in Buildings B-01 and B-02.

Building ID	Instrumentation				
	Window/door sensor	**Env. sensor [a] (CO_2, RH, Temp.)**	**Presence detector**	**Radio-controlled radiator[b]**	**JuControl radiator control**
B-01	Yes (09.2021)	Yes (09.2021)	No	Yes (09.2021)	Yes (09.2021)
B-02	Yes (05.2022)	Yes (03.2022)	No	Yes (02.2022)	Yes (02.2022)
B-03	Yes (10.2021)	Yes (10.2021)	No	Yes (03.2023)[‡]	No
B-04	Yes (03.2023)	Yes (03.2023)	Yes (03.2023)[‡]	Yes (03.2023)[‡]	No
B-05	Yes (07.2023)	Yes (07.2022)	No	No	No
B-06	Yes (03.2023)	Yes (01.2023)	No	No	No
B-07	Yes (06.2022)	Yes (06.2022)	No	No	No
B-08	Yes (01.2023)	Yes (01.2023)	No	No	No
B-09	Yes (01.2023)	Yes (01.2023)	No	No	No
B-10	Yes (08.2022)	Yes (08.2022)	Yes (03.2023)[‡]	No	No
B-11	Yes (01.2023)	Yes (01.2023)	No	No	No
B-12	Yes (01.2023)	Yes (01.2023)	No	No	No

[a] Environmental sensors installed. RH = relative humidity; Temp. = temperature.
[b] Cloud-controllable smart radiator installed.
[‡] Part of KNX System.

3.8.4 Experiment Groups and Teams

In the experiment design, rooms were clustered together to form *teams*, which in turn were assigned to experiment groups. Each experiment group can contain multiple teams, but each team belongs to only one experiment group. Thus, the set of features available to a team is determined by the group to which it belongs. In Table 3.7, the assignment of rooms to teams, as well as teams to experimental groups, is shown. Each team consists of between 20 and 44 rooms, and rooms within a team can be drawn from one or more buildings. A total of 12 teams are involved in the experiment, assigned to the four experiment groups.

Table 3.7 Assignment of teams and buildings to experiment groups, along with number of rooms in each feature group. The features enabled in each group are given in Table 3.4.

Feature Group	Teams	Buildings	Num. Offices	Num. Teams
A	T1, T2, T3	B-01, B-02	92	3
B	T4	B-03, B-04	20	1
C	T5, T7, T8, T9, T10, T12, T13	B-05, B-07, B-10, B-06, B-08, B-09	230	7
D	T6	B-05	44	1
Total		10	386	12

The buildings assigned to each team and the total number of offices are shown in Table 3.8, alongside the number of employees in the team, and the absolute and relative number of rooms without occupants (meeting rooms and otherwise no-occupant offices). Three additional teams (Teams T14, T15, and T16) which belong to the two excluded buildings mentioned previously, are included in the table for completeness purposes, but are not analyzed further in the results section except for presentation of user engagement results (Section 5.2).

Table 3.8 The buildings and number of offices assigned to each team with corresponding employee count. Among these offices, some are meeting rooms and otherwise offices without assigned occupants – these are shown as "Unassigned Rooms", with their percentage relative to the total shown in parenthesis.

Team	Building(s)	Num. Employees	Total Num. Offices	Num. Unassigned Offices
T1	B-01	58	39	7 (17.9%)
T2	B-02	33	24	7 (29.2%)
T3	B-02	70	30	3 (10%)
T4	B-03, B-04	37	20	3 (15%)
T5	B-05	63	44	5 (11.4%)
T6	B-05	51	44	3 (6.8%)
T7	B-06, B-07	71	31	5 (16.1%)
T8	B-06, B-07	81	30	2 (6.7%)
T9	B-08	48	29	1 (3.4%)
T10	B-09	56	26	2 (7.7%)
T12	B-10	46	33	0 (0%)
T13	B-10	52	37	2 (5.4%)
T14*	B-11	52	25	3 (12%)
T15*	B-12	60	40	6 (15%)
T16*	B-12	82	43	7 (16.3%)

* Not included in the main experiment, but presented here for use in user engagement results.

3.8.5 Hypothesis Testing Strategy

The strategy for testing the experiment hypotheses H_1 to H_3 using the experiment setup is presented below. It is necessary to clarify the "JuControl-activated" terminology used for the hypothesis testing and in later chapters. "Activation" and "non-activation" of JuControl refers to the availability of JuControl to occupants of a particular office depending on data consent. For privacy reasons and to adhere to legal data protection policies in Forschungszentrum Jülich, JuControl is only "unlocked" for offices where *all* occupants have actively consented to the required data privacy policy that permits other occupants to view potentially sensitive sensor data related to the office. In other words, although JuControl is *available* in principle for occupants of the offices in all the experiment groups, JuControl requires that all occupants of a given office consent to the data policy before it opens up for that office, after first confirming that they are officially assigned to the said office. Therefore, the term *JuControl-activated offices* refers to offices where all officially assigned occupants have granted their consent to the data policy, while *non–JuControl-activated offices* refers to offices where at least one occupant has not yet granted consent, or otherwise has declined granting consent. Nevertheless, as would be seen in the result analysis of Chapter 5, occupants of non–JuControl-activated offices in recommendation-enabled groups (Feature Group A, B, and C) received evaluation summaries and energy savings recommendations via email, just like occupants of JuControl-activated offices.

Table 3.9 shows how the experiment design is to be used for testing each of the previously outlined hypotheses. The design strategy incorporates redundancy in order to improve robustness, such that technical failures during the experiment can possibly be tolerated. Furthermore, it can be observed that many teams are allocated to the evaluation-enabled groups (Groups A, B, C) as against non-evaluation groups (Group D). This is because the teams in the evaluation groups were further split into experiment groups that do not affect the experiment for this thesis.

Table 3.9 Hypothesis testing strategy based on the experiment design and building performance comparison.

Hypothesis	Hypothesis Testing Strategy
H_1 (Effect of evaluations / recommendations)	**H_1-Test-1** Performance comparison of Team T5 vs Team T6. This test checks that Team T5 has a better performance than Team T6 as judged by the mean energy penalty of the offices, since although both are drawn from the same building, T5 has ventilation evaluation as well as recommendations enabled, in addition to JuControl access, while Team T6 has only JuControl access with no behavioural improvement functionality.
H_2 (Effect of active participation)	**H_2-Test-1** Performance comparison between *non–JuControl-activated* rooms and *JuControl-activated* rooms within evaluation- and recommendation-enabled teams (all teams except Team T6). In other words, *JuControl activation* is used as a proxy for judging active participation, since such a participation is only technically possible with unrestricted JuControl access. The test checks if JuControl-activated offices in these teams have better performance than non-activated rooms within the same building or team, under the premise that occupants of JuControl-activated rooms have access to full information about energy evaluations and other contextual information, unlike occupants of non–JuControl-activated offices who can only receive occasional emails.
H_3 (Effect of automatic heating control)	**H_3-Test-1** Comparison of setpoint temperature *deviation-from-ideal* between JuControl-activated offices in Building B-01 and non–JuControl-activated offices. Since in non–JuControl-activated offices, occupants physically turn the smart radiator to specify their setpoint temperatures without intervention from the automatic controller, this test checks if the non–JuControl-activated offices have a significantly worse setpoint deviation than the JuControl-activated offices where the automatic controller manages the setpoint temperature based on occupant-provided presence schedules.
	H_3-Test-2 Whole-building thermal demand before-and-after comparison for Building B-01 (i.e. before and after the implementation of automatic heating control) using the building "energy signature" performance comparison methodology, which was discussed in Section 3.9. This check should show that the improvement in setpoint temperature efficiency for JuControl-activated offices established by **$H_{3\text{-Test-1}}$** above, if any, translates to noticeable energy savings at the building level on a longer time horizon.

For testing Hypothesis H_1, H_2, and the first part of H_3, statistical significance tests are used assuming a significance value $p = 0.05$. Furthermore, the effect size is also tested using Cohen's d as the effect size metric, where $|d| < 0.2$ represents a small effect size, $|d| \approx 0.5$ represents a medium effect size, and $|d| \geq 0.8$ represents a large effect size, and $|d| \geq 1.3$ represents a very large effect size [106]. While the statistical significance only tells us if a difference in energy performance exists between an experimental group and its control group, the effect size provides an estimate of the relative magnitude of the performance difference under consideration.

3.9 Building Performance Evaluation

Several factors determine the energy demand of buildings, among which are: building characteristics; energy system characteristics, control and maintenance; weather parameters; and occupant behaviour [107]. This makes the prediction of building energy demand challenging. Over time, numerous approaches have been established for estimating this demand, which can as well be classified in a few slightly different ways [107, 108]. For this discussion, we adopt the classification of Fumo [107], consisting of the statistical, engineering, and hybrid methods, as shown in Fig. 3.12. In the statistical approach, the mathematical relationships between measured input and output variables of the building are extracted via the analysis of the data, without recourse to the physical modelling of the building itself. In the engineering approach, however,

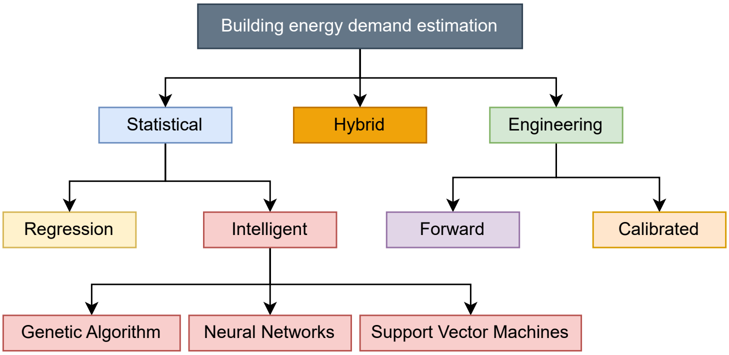


Fig. 3.12 Classification of building demand estimation methods according to Fumo [107].

mathematical equations are used to describe the physical model of the building and the output variables are obtained from the building model with known inputs. The hybrid approach is a mix of statistical and engineering approaches, equivalent to a grey-box approach.

For building performance rating, on the other hand, the International Performance Measurement and Verification Protocol (IPVMP) [109] details general strategies for assessing energy savings resulting from Energy Conservation Measures (ECMs). The protocol was developed by the U.S. Department of Energy in 1994 and is now overseen by the Efficiency Valuation Organization (EVO). Energy Conservation Measures (ECMs) are activities carried out to reduce the energy consumption of a building or facility, for example building envelope retrofits, lighting upgrades, and occupant energy-related behaviour improvement programmes. The IPMVP references industry standards like the U.S. Department of Energy's "M&V Guidelines: Measurement and Verification for Performance-Based Contracts" [110] and ASHRAE Guideline 14, all related to the measurement of energy and demand savings through the application of ECMs. In general, these savings assessment strategies involve *baselining*, which is an activity during which the current operational status of monitored facilities, as well as the human aspects like occupancy and environmental preference, are determined and modelled. This performance determined during the *baseline period* then provides a reference point for assessing the savings achieved through ECMs during the *reporting period*. Energy savings result from two drivers: improvement in performance, and reduction in usage, as shown in Fig. 3.13 [110].

The general M&V (Measurement and Verification) equation according to IPMVP is:

$$\text{Savings} = \text{Baseline Period Energy} - \text{Reporting Period Energy} \pm \text{Adjustments} \tag{3.10}$$

where "Adjustments" are divided into two types: *routine adjustments* that account for energy-governing factors that are expected to change routinely between the baseline and reporting periods, such as weather; and, *non-routine adjustments* which account for energy-governing factors that are usually expected to stay constant between the periods, e.g. the size of the facility, or installed equipment. Savings are usually reported under the conditions of the reporting period, i.e. the baseline conditions are adjusted (or "forecasted") to match the reporting period conditions. These savings are called *avoided energy consumption*, which essentially represents the savings in the reporting period compared to what it would have been without the ECMs. The M&V equation (Eq. 3.10) can now be re-written as:

$$\begin{aligned} \text{Avoided Energy Consumption } &= \text{Adjusted Baseline Energy } - \text{ Reporting Period Energy} \\ &\pm \text{ Non-Routine Adjustments to Reporting Period Conditions} \end{aligned} \tag{3.11}$$

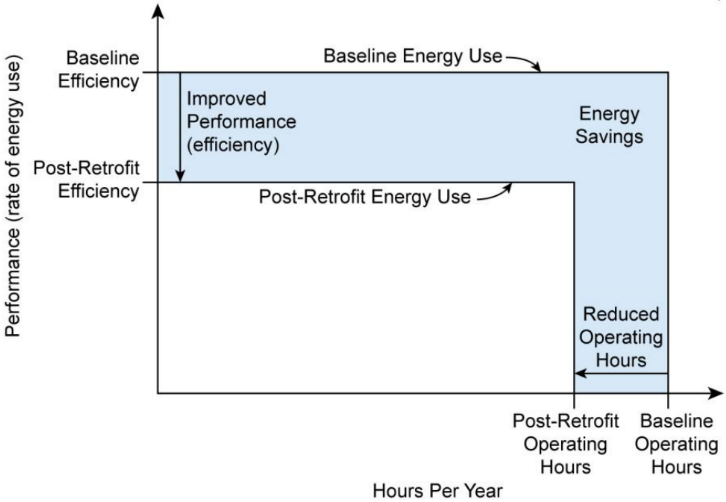


Fig. 3.13 Driving factors that determine energy savings: efficiency and use (from [110])

where the routine adjustments (for the reporting period) are now subsumed within "Adjusted Baseline Energy" through the forecasting of the baseline period conditions into the reporting period. The forecasting of baseline consumption is often through regression analysis, especially when whole-facility savings are being investigated. Validation of such statistical techniques is required in IPMVP, using statistical evaluation indices like Coefficient of Variation of the Root Mean Squared Error (CV(RMSE)), the Mean Bias Error (MBE), amongst others.

For the measurement and verification of the savings through ECMs in buildings, four approaches ("options") with increasing complexity are described in the IPMVP (which also tallies with the classification in the U.S. Department of Energy's "M&V Guidelines"). These are shown in Table 3.10.

Table 3.10 Description of IPMVP measurement and verification options for ECMs in buildings (from [109, 110]).

IPMVP Option	Description	Example Applications
A. Retrofit-Isolation: Key Parameter Measurement	Savings are determined by field measurements of the key parameter(s) that define the energy consumption of the systems affected by the ECM.	Lighting retrofit where the power consumption is the measured key parameter.
B. Retrofit-Isolation: All Parameter Measurement	Savings are determined by field measurement of the energy consumption and/or related independent or proxy variables of the system affected by the ECM.	Installation of a variable-speed drive and associated controls on an electric motor. Electric power is measured with a meter installed on the electrical supply to the motor.
C. Whole-Facility Measurement	Savings are determined by measuring the energy consumption at the level of the whole facility using utility meters.	Multifaceted energy management programs affecting many systems in a facility.
D. Calibrated Computer Simulation	Savings are determined through simulation of the energy consumption of the whole facility or sub-facility.	Multifaceted energy management programs affecting many systems in a facility but without a meter during the baseline period.

In this thesis, whole-facility measurement (Option C in Table 3.10) is used to determine the savings through the behaviour intervention measures and associated ECMs developed and executed in this project, since multi-year fine-grained building-level metering data is available for many of the investigated buildings. Since heating energy is the form of energy investigated in this thesis, the methodology for the building performance evaluation applied this thesis is based on the widely used Degree Days approach [108, 111], which theorises that the space-conditioning (heating or cooling) energy demand of buildings is a linear function of the ambient temperature, when the ambient temperature is below (above) a given threshold for heating (cooling). The underlying idea is that heating (cooling) is only required below (above) the threshold ambient temperature. The main advantages of the Degree Day approach are that it requires few input data, and it produces acceptable results in practice. The Degree Day approach is mainly used for [111]:

- estimating energy consumption of new and existing buildings for space heating and cooling; and,
- continuous monitoring and analysis of existing buildings using historical data

The application in this thesis concerns primarily the second function of historical data-based analysis.

Mathematically, for heating demand, the Heating Degree Days (HDD) for a period D consisting of d days, is [111, 112]:

$$HDD_{\text{period}} = \sum_{d \in D} \left[\frac{1}{N} \sum_{t \in \mathcal{N}_d} \left(T_{\text{base}} - T_{\text{amb}}(t) \right)^+ \right] \quad (\text{K} \cdot \text{day}) \tag{3.12}$$

where $\mathcal{N}_d$ is the set of time points of ambient temperature measurements for day d (e.g. hourly data of 24 time points), $N = |\mathcal{N}_d|$ is the number of measurements per day, and $T_{\text{amb}}(t)$ (°C) is the ambient temperature at time t in $\mathcal{N}_d$. The superscript plus notation $(\cdot)^+$ indicates that only the positive values of the enclosed expression are taken. The *base temperature*, T_{base} (°C) is commonly a standard value for simplicity (for example, 15.5 °C in Europe [113] and 18.3 °C in the U.S.A. [108]). However, T_{base} refers more accurately to the *balance point temperature* of the particular building under analysis [108, 111], and the assumption of a standard value can lead to errors in energy estimation. The balance point temperature is the ambient dry-bulb temperature above which the building needs no heating to maintain thermal comfort within the building. In other words, it is the ambient temperature value at which, for a given interior temperature, the total heat loss through the building envelope is equal to the total heat gain from insolation, occupants, etc. [111, 112]. It is recommended to use building-specific base temperatures [111]. Therefore, for the methodology employed in this thesis, T_{base} refers to the building-specific balance point temperature. The method for the estimation of the balance point temperature from historical data is described in Section 3.9.2 below.

More accurate HDD calculations are achieved with finer-grained timeseries data, ideally with hourly or sub-hourly temperature measurements [111]. However, it can also be applied with adjustments for lower-resolution data, depending on the available data (see e.g. [114] for details; in Germany, the mean daily ambient temperature is standard [111], while at the level of the European Union, daily mean, maximum, and minimum temperatures are generally used [113]). In this thesis, since detailed temperature measurements are available (down to minute-wise resolution and measured on-site), we use the standard formula in Eq. 3.12.

Hence, the thermal energy, $\widehat{E}_{\text{th},d}$, required for space heating in a building for a given day d, can be estimated from the Heating Degree Days (HDD) as follows:

$$\widehat{E}_{\text{th},d} = U' \cdot 24 \cdot HDD_d \quad (\text{kWh}) \tag{3.13}$$

where U' is the proportionality constant representing the overall heat loss coefficient of the building (in kW/K), and HDD_d is the heating-degree day value for day d. The scaling by factor 24 converts the daily estimate to kWh. The overall heat loss coefficient (or *building envelope factor*), U', comprises two main parts. First is the total thermal transmittance of the building envelope in kW/K, computed as $U \cdot A$, where U is the per-unit-area thermal transmittance (U-value) of the building as a whole, i.e. combining the values for the different thermal boundary surfaces of the building (kW/m^2·K), and A is the total external surface area (m^2). The second relates to the air infiltration rate of the building, through which heat is lost according to $\dot{m} \cdot c_{\text{air}} \cdot \Delta T$, where $\dot{m}$ is the mass flow rate of air being exchanged between the building and the ambient (in kg/s), c_{air} is the specific heat capacity of air (in kJ/kg·K), and ΔT is the difference between indoor and ambient temperature (in K). The total thermal transmittance and infiltration losses are usually estimated under simplifying assumptions, including the assumption of a constant infiltration rate [111].

However, where historical building thermal demand data is available, in which case one is interested in the performance of an existing building, the parameter U' in Eq. 3.13 is estimated by regression using data from the baselining period. The resulting linear equation is then applied to the reporting period, providing a "forecast" of the building thermal demand based on the baseline *signature*. The difference between the "forecasted" energy demand and the demand of the reporting period, gives a measure of the performance of the building between the two periods. When the total thermal transmittance of the building envelope remains constant between the two periods, then the performance difference can only be explained by *other* factors such as change in usage patterns (including natural ventilation patterns), upgrade of energy systems, etc. Usually, such analyses require interpretation, considering the on-site conditions.

3.9.1 Building Energy Signature and Performance Line

A general procedure for determining savings due to Energy Conservation Measures in the presence of historical data is described in Kissock, Haberl, and Claridge [112] as follows.

Step 1: Measure energy use and influential variables during the baseline period.

Step 2: Develop a regression model of baseline energy use as function of influential variables.

Step 3: Measure energy use and influential variables during reporting period.

Step 4: Use the values of the influential variables from the reporting period (Step 3) in the baseline model (Step 2) to predict how much energy the building would have used if there had not been any energy conservation measures (ECMs).

Step 5: Subtract measured reporting period energy use (Step 3) from the predicted baseline energy use (Step 4) to estimate savings.

Depending on the kind and resolution of the available data, different linear regression models can be used to represent the historical performance of a building. Where only monthly thermal energy demand data is available, a **two-parameter linear regression model** (2P model) is fitted over the historical data (monthly demand vs. monthly heating degree days) to produce what is termed the building's *energy performance line* [111] (see Fig. 3.14a). Ideally, the historical data would fit an energy performance line of the form:

$$\widehat{Y} = \beta_1 + \beta_2 \cdot X \tag{3.14}$$

where β_1 and β_2 are the regression coefficients. The dependent variable, $\widehat{Y}$, represents the monthly thermal energy demand estimated by the model (in kWh), and X is the independent variable representing the monthly heating degree days (in K·day). The parameter β_2 is equivalent to $U' \times 24$ from Eq. 3.13 (in kWh/K·day), and the intercept β_1 represents the base load (kWh). However, β_1 is only reliable if the building-specific base temperature was used for calculating the HDD [111].

Nevertheless, since the available data in this thesis is fine-grained (minute-wise), a **three-parameter regression model** (3P model) is used on daily data, where the daily thermal energy demand is regressed on the average daily ambient temperature [111, 112]. The regression model is called the *energy signature* of the building and consists of two line segments – a horizontal segment representing the base temperature, and a linear segment representing the dependence of the demand on the heating degree day (see Fig. 3.14b for illustration of the 3P model, and Fig. 3.14c for historical data of one of the buildings considered in the thesis corresponding to the 3P model). The form of this model is:

$$\widehat{Y} = \beta_1 + \beta_2 \cdot \left(\beta_3 - X\right)^+ \tag{3.15}$$

where β_1, β_2, and β_3 are the three regression coefficients, $\widehat{Y}$ is the estimated daily thermal energy demand (dependent variable, in kWh), and X, the independent variable, is the daily average ambient temperature (in °C). β_2 is again equivalent to $U' \cdot 24$ from Eq. 3.13 (kWh/K·day), and β_3 is the building's balance-point temperature (°C). The magnitude of the constant line β_1 represents the base load (kWh). Again, the $(\cdot)^+$ notation means that the parenthesized expression is set to zero if the result is non-positive. The expression $\left(\beta_3 - X\right)^+$ effectively represents the daily HDD.

The 3P regression model is also called a *change point model*, since it enables the determination of the balance temperature, at which the regression line changes [112]. The 3P model can be extended to up to four independent variables, where the other variables are fitted by simple linear functions [112]. An algorithm for finding the change-point (base temperature) for the 3P model is described in the next section.

It is important to check the fitness of the regression models, using the following statistical tests [112]. First is the Root Mean Squared Error (RMSE), which estimates the model residuals, i.e. the "distance" between the modelled data and the actual data. Mathematically,

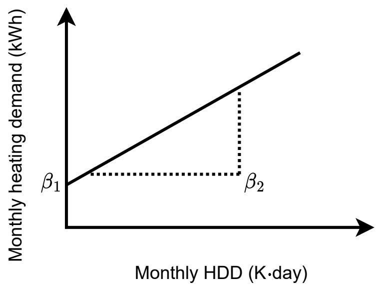
$$RMSE = \sqrt{\frac{\sum(Y - \widehat{Y})^2}{n - p}} \tag{3.16}$$

where n is the number of data points, p is the number of regression coefficients, Y is the set of historical thermal energy demand data points, and $\widehat{Y}$ is the model-estimated demand computed from Eq. 3.14 or Eq. 3.15.

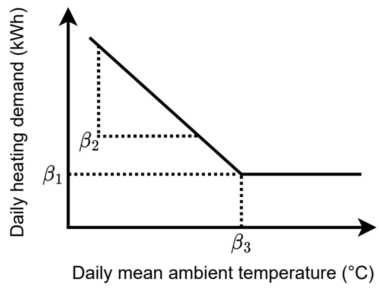
Secondly, the squared correlation coefficient, $R^2 \in [0,1]$, is computed, representing how well the regression model fits the data compared to how well the mean fits the data. Mathematically,

$$R^2 = 1 - \frac{\sum(Y - \widehat{Y})^2}{\sum(Y - \bar{y})^2} \tag{3.17}$$

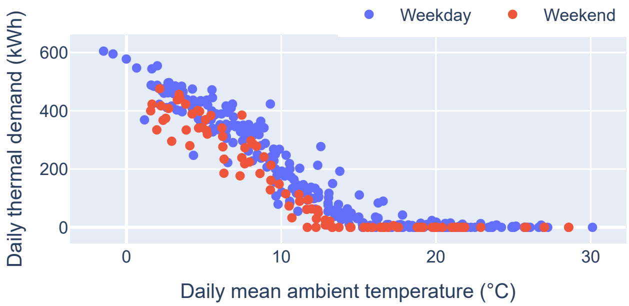
where $\bar{y}$ is the mean of the historical demands, and Y and $\widehat{Y}$ are as defined for Eq. 3.16. An R^2 value of zero means that the regression model does not fit the data any better than the mean does, and $R^2 = 1$ indicates a perfect fit between the data and the model (although this rarely occurs in practice).



(a) Two-parameter regression model



(b) Three-parameter regression model



(c) Example historical data suitable for three-parameter regression model.

Fig. 3.14 Illustration of two- and three-parameter regression models showing energy performance line (a) and building energy signature (b). The labels β_1, β_2, and β_3 are the parameters of the models (as applicable), according to Eq. 3.14 and Eq. 3.15. In (c), example data for determining the "energy signature" for one of the experiment buildings for the year 2022 in terms of the three-parameter regression model is shown.

3.9.2 Estimating the Heating Balance-Point Temperature of a Building

Estimating the heating balance point temperature from the 3P model of Eq. 3.15 simply means finding the value of the regression coefficients β_1, β_2, and β_3 that minimizes the RMSE of the model. A two-stage search algorithm for determining β_3 (also a "change point detection algorithm") is described in the ASHRAE Inverse Modelling Toolkit (IMT) [112], in which the independent variable X is divided into equal-width grids of width dx between the minimum value, $x_{\min}$, and the maximum $x_{\max}$. Then β_3 is incremented from $x_{\min}$ to $x_{\max}$ in steps of dx, and for each value of β_3, the RMSE is determined, retaining the value $\beta_{3,\text{best}}$ that results in the minimum RMSE. Afterwards, a smaller region centred on $\beta_{3,\text{best}}$, i.e. $\beta_{3,\text{best}} \pm \text{dx}$, is again divided into smaller grids and the process is repeated once more, retaining the $\beta_{3,\text{best}}$ that corresponds to the minimum RMSE.

In this thesis, however, a more direct approach is employed, based on solving the problem as an optimization problem using publicly available software libraries. Specifically, the *NumPy* Python library [115] is used to define the piecewise linear function:

$$\hat{Y} = \begin{cases} \beta_1 + \beta_2 \cdot (\beta_3 - X), & \text{if } X \leq \beta_3 \\ \beta_1, & \text{otherwise} \end{cases} \tag{3.18}$$

Subsequently, the *SciPy* Python library [116] is used to optimize the parameters (β_1, β_2, and β_3) of Eq. 3.18 to obtain the values yielding the minimum RMSE. From this, β_3 then corresponds to the building balance-point temperature.

Finally, to determine the savings in the reporting period, the daily mean ambient temperatures for the reporting period are plugged into the 3P regression model of Eq. 3.15 to estimate the "forecasted" consumption of the reporting period given the parameters of the baseline period. The energy savings $E_{\text{th,saved},D}$ (kWh) in the reporting period D is then the difference between the forecasted consumption and the actual consumption for that period. Mathematically, for the reporting period D consisting of d days, the energy saved is:

$$E_{\text{th,saved},D} = \sum_{d \in D} \widehat{E}_{\text{th},d} - E_{\text{th},d} \tag{3.19}$$

where $\widehat{E}_{\text{th},d}$ is the 3P-model estimated thermal energy demand for day d from Eq. 3.15 (in kWh), and $E_{\text{th},d}$ is the actual (historical) thermal demand for the same day (in kWh).

The above-described building performance evaluation methodology is applied in the building performance analysis of Chapter 5 (Section 5.5).

3.10 Occupancy Estimation

As has been previously discussed, the behaviour of the occupant affects the final performance of the building energy-wise, and in this thesis behavioural change towards increased energy efficiency is the overarching objective. To account for the presence profile of occupants in the building, two distinct classes of needs in terms of area of application can be identified. The first class of needs, which I term the *synthesis* need, deals with generating realistic occupancy profiles for the building. These profiles, which generally depend on the intended use of the building (whether public/office, or private/residential), find application in areas such as building design and simulation, and energy budgeting. Solutions can be deterministic or nondeterministic. In the deterministic approach, a fixed profile is given for the building based primarily on intended use, time of day, and day of week (e.g. in [13]). In the nondeterministic approach, the profiles are modelled statistically with inputs based on empirical studies [10, 117, 118]. Markov Chain models appear to be the most widely used in the literature (e.g. in [10, 119–121]).

The other class of needs, which I term the *inference* need, deals with determining the current status of occupancy in a space under consideration, i.e. human sensing. While *synthesis* deals with general trends and occupant profile generation, *inference* deals with particular instances and actual current or historical status. Inference is classified according to the required level of output information of the spatio-temporal properties of the occupancy: *presence*, i.e. whether the space is occupied or not; *count* – the number of people in the space; the location of the person(s) within the space; the track of the person(s) in the space, i.e. the spatio-temporal history of the person(s); and finally, the identity of the person [122]. These information requirements have the cumulative property that knowing one implies knowing all the others below it in the progression. Another classification includes behavioural characteristics of the current activity of the person.

To solve problems in the inference class, two major approaches are used for occupancy detection, namely observational studies and occupant surveys [123]. Observation relies on sensing, where occupancy is derived directly or indirectly through sensors. Observation methods can be classified into six groups: image-based, threshold and mechanical, motion sensing, radio-based environmental, human-in-the-loop, and consumption sensing [124]. The direct observation methods include use of motion sensors (mostly Passive Infrared (PIR) sensors) [125], wearable sensor devices, and image-based methods (using both video and photo cameras) [123, 126, 127]. Indirect sensing methods estimate occupancy using techniques that track quantities that are correlated with occupancy, or, in few cases, using historical data (e.g. in [128]).

The majority of research efforts have used environmental sensors. CO_2 concentration offers the strongest correlation amongst these [129], and has been used as the sole estimation parameter in several studies (e.g. [125, 130–132]). Several other research works have used combinations of CO_2 concentration with relative humidity, Volatile Organic Compounds (VOC) concentrations, indoor temperature, and indoor pressure [125, 133–136].

Other methods include near-field communication (NFC) or Bluetooth beaconing devices (where the occupant is represented by their mobile device) [74], mobile phone Wi-Fi tracking [137]; use of acoustic pressure (in microphones) [138]; use of pressure sensors placed on chairs [139]; and monitoring of the consumption of electrical devices [140]. A review of these methods is presented in [127].

3.10.1 Occupancy Estimation via Environmental Sensing

As mentioned above, typical environmental sensors used in occupancy estimation are CO_2 concentration, indoor/outdoor temperatures, relative humidity, and VOC concentrations. In this section the main methods used for deriving occupancy parameters from individual and combinations of these sensors are discussed.

The techniques used are divided into three broad categories by method: machine learning methods, analytical methods, and statistical methods. Each category has pros and cons. For using machine learning, one requires reliable and extensive historical data. On the positive side, the methods are able to represent complex and hidden relationships between the observed variables and the estimated parameter. Analytical methods use mathematical principles and physical laws to model the relationship between the variables. State estimation techniques are usually applied here. The advantages include the need for significantly less historical data and more general applicability of the models. On the flip side, it requires specialist knowledge and it is more difficult to derive exact relationships.

For the analytic models, the simplest approach for deriving occupancy uses CO_2 concentration gradient [129, 134]. Here the rate of change of CO_2 concentration in the monitored space is analysed and using a rule-based system, occupancy is derived. A more complicated analytical method estimates occupancy using CO_2 mass-balance equations [130, 141, 142]. On the machine learning side, estimators of occupancy based on CO_2 concentration have been studied (e.g. [138, 143, 144], and reviews [127, 145]). In this thesis, occupancy estimation (detection) is done based on the CO_2 mass-balance approach or from PIR sensors, where available.

3.10.2 Privacy Considerations in Occupancy Estimation

While occupancy estimation via direct sensing techniques provides the highest accuracy, one major drawback with the approach, however, is that it is considered the most intrusive approach. This is especially true for detection via cameras. Given the increasingly strict privacy requirements of organisations and governments, especially in the EU, this method has limited application, both in scale and in geographical location. Furthermore, as mentioned in the introduction (Chapter 1), Germany has stricter privacy requirements and policies than other EU countries, so particular care must be taken to diminish potential feelings of privacy violation among occupants. While it is possible in practice to use cameras to detect occupancy in a privacy-preserving manner (for example by automatically obfuscating people with Gaussian noise, as applied in [127]), it is easy to appreciate that this approach does little to allay the fears of occupants. This is because the method can "in principle" detect both identities and activities, requiring (for many occupants, prohibitively high) trust in the researchers (or management) and in the robustness of the system. Beyond this, some camera based systems also suffer from false triggers like moving foliage, lighting variations, and shadows [122].

PIR sensors mitigate this problem to some extent in that they lack the ability to "see" what occupants are doing or to potentially identify them by any means, thereby providing better privacy. However, PIR sensors have other disadvantages, which include inability to count people and failure to detect an occupant in a stationary state, (which state tends to be the default for occupants in an office setting) [127, 129]. Modern PIR sensors are not susceptible to the stationary-occupant false negative, nevertheless. Yet again, PIR sensors can also be triggered by non-human environmental factors like HVAC systems [122]. PIR sensors and other direct sensing methods (camera-based) share the disadvantage that they are prone to "blind spots", i.e. areas of the space occluded from the sensors by furniture or the form factor of the space [129].

On the other hand, indirect sensing using environmental sensors is perceived as less intrusive than direct sensing and is a preferred choice under strict privacy demands. However, one major drawback of indirect sensing via environmental sensors as compared with a direct sensing is the time lag between the occupancy change event and the reflection in the monitored quantities [123, 127]. Another disadvantage, as mentioned above, is that accuracy tends to be higher for PIR-based occupancy detection than for environmental sensing techniques [125].

Other sensing methods include use of ultrasound and radio frequency detection. The former has the disadvantage that it has poor accuracy in locating occupants, and high rates of false positives due to sound from other sources apart from occupants. Furthermore, arrays of sensors have been deployed to improve location accuracy, but suffer from mutual interference [127]. Using radio frequency requires users to carry Radio Frequency Identification (RFID) tags, which has obvious problems with convenience and privacy [127]. These methods are not considered further.

Chapter 4

Implementation of Models and Tools

This chapter first briefly touches on the software tools developed in this work as the means to provide incentives for behaviour change. Afterwards, the implementation of the mixed-mode evaluation system derived in the previous chapter is presented, along with the methodology employed to derive occupancy as an input into the evaluation system. As previously mentioned, *eco-visualization*, *control* (via an HMI), *behaviour evaluation*, and *gamification* are the foundational concepts of the behaviour intervention of this thesis. The developed software are presented next, indicating how they address the above-mentioned concepts.

4.1 Overall Software Framework

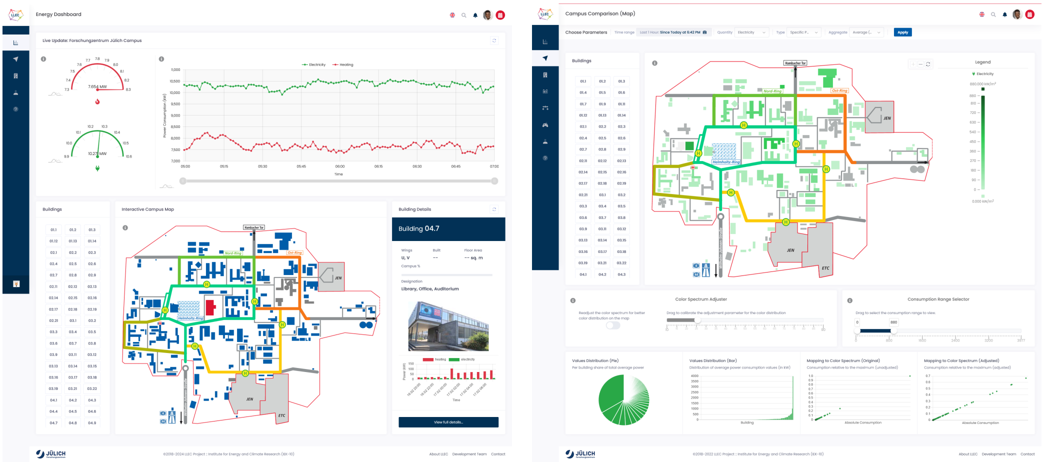
As introduced in the previous chapter, the software developed in this thesis are collectively called the *Energy Dashboard Suite*, consisting of the *Campus Viewer*, *JuControl*, *Juracle*, and *ALICE*. The next sections briefly describe these software applications, while even more details are provided in Appendix A covering the overall system architecture to show how the software applications interact; the supporting hardware platform for data capture, including the utilised sensors, actuators and communication devices; and the strategies employed to deal with the potential privacy issues that are inherent in such user-focused interventions.

4.2 The Campus Viewer: *Eco-Visualization*

The *Campus Viewer* is a web application that visualizes the energy system-related data of the campus at the building level. It features charts showing near real-time, auto-updated heating and electricity demand of the campus as a whole, as well as of individual buildings. Additionally, in the Campus Viewer buildings can be compared visually on a map using colour scales, both for heating and for electricity demand. Screenshots of the Campus Viewer are shown in Fig. 4.1.

4.3 JuControl: *Eco-Visualization, Control, Gamification*

JuControl provides access to room-level real-time and historical data, including indoor conditions like temperature, CO_2 concentration, humidity and luminosity, as well as the status of the windows and doors. It also features a stylized pseudo-3D real-time visualisation of the room, which shows window and door status



(a) Home page

(b) Building demand comparison page

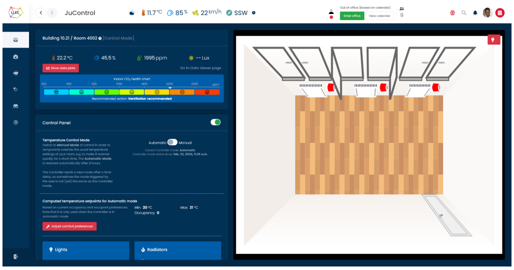
Fig. 4.1 Screenshots of the Campus Viewer. In (a), the live heating and electricity demand of the campus is shown on the home page, and in (b), the building demand comparison for heating is shown. Full-sized screenshots are provided in Appendix E.

visually (see Fig. 4.2a). The 3D diagrams are generated by *ALICE* (more details in Section **??**). Furthermore, JuControl features a building page, where the interactive floor plans of the buildings that the user has access to, are shown, colour-coded by the kind of access the user has to the rooms in the building (Fig. 4.2b). This is a privacy and data access feature of JuControl that is discussed more in Section A.3. These features support the *visualisation* function of JuControl according to the four-pillar classification.

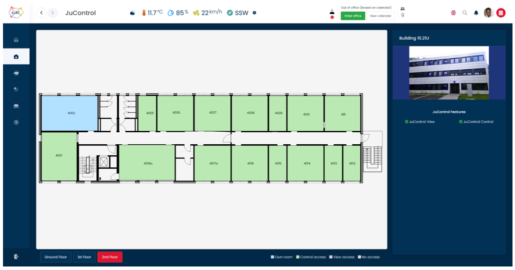
On the *control* pillar side, JuControl provides a Human Machine Interface (HMI) widget in some buildings, by which occupants can specify their preferred heating setpoint temperature. Added to this, JuControl has a highly customizable in-built calendar through which occupants also provide their schedule for being physically present in the office. This schedule is then combined with the heating setpoints to provide on-demand heating via a cloud controller. (The cloud controller [105] (in preparation) is not part of this thesis, but the controller inputs are supplied by JuControl via an API.)

Finally, in cooperation with Juracle (Section 4.5), JuControl implements the *gamification* pillar through the use of ratings, competition, and leaderboards (Fig. 4.5). For enabling competition, sensor-equipped offices are grouped into *teams* comprising several offices, and both teams and offices are rated in terms of energy efficiency. These ratings are then compared within and between teams anonymously (see Appendix A, Section A.3.2 for discussion on privacy).

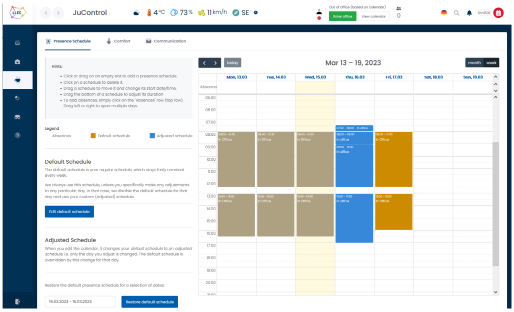
To access JuControl as a room occupant (in other words, to "activate" JuControl for a room), all the occupants of the given room must first grant their consent for data processing and visualization. This "activation" is the only means via which occupants can access JuControl in order to view the status and performance of their room, or to control the heating in their office (depending on available features). The user consent process is started automatically when a user visits any of the JuControl pages. After the user grants their consent, the office mates as determined by a centrally managed FZJ facility allocation database are automatically emailed with a link to also grant theirs. If any occupant declines consent, the room is *not* activated in JuControl, meaning that none of the occupants can view the data.



(a) "My Room" page in JuControl



(b) "My Building" page in JuControl



(c) The presence schedule specification page in JuControl

Fig. 4.2 Screenshots of JuControl showing (a) the user's room view, (b) the user's building floor plan, and (c) the user's JuControl calendar. The pseudo-3D representation of the room in (a) is produced by ALICE. In (b), the colour coding of the rooms in the building floor plan indicates what kind of JuControl access the user has (view access, control access, or no access) to that room. The calendar in (c) enables the user to input their expected presence schedule. Full-sized screenshots are provided in Appendix E.

4.3.1 Coupling JuControl with Building Automation

The presence schedule and comfort preferences of occupants are captured in JuControl and aggregated for each room. JuControl then provides the composite data as a timeseries at any requested time resolution and for any time period to an external heating controller running in the cloud (see Althaus et al. [105]) via a RESTful API. A complex signalling process between JuControl and the cloud controller is used to establish what "control mode" should be activated: *manual setpoint mode* triggered via the physical HMI (i.e. radiator valve) or within JuControl via the setpoint temperature widget, or *automatic setpoint mode* via the occupants' schedules; however, this signalling is not discussed further in this report. The JuControl calendar is used to capture occupant presence schedules, and the temperature setpoint widgets in JuControl capture occupant comfort preferences. It should be reiterated here that the availability of the occupants' presence schedule to the heating controller depends on if the office is *JuControl-activated*, i.e. if all the occupants in the office have agreed to the data privacy terms of the project. It is important to note that for JuControl-activated offices, an 8.5-hour presence schedule with a 30-minute lunch break is enabled *by default* for each day of a working week as a baseline, in order to ensure that at the offices are heated pre-emptively during working hours and avoid cold offices if the occupant never adjusts the calendar. For non–JuControl-activated offices, the heating control is *manual* in the sense that the occupants physically turn the smart radiator valve in order to set a desired temperature. Unlike the traditional radiator valve in which the angle of rotation indicates the current setpoint temperature, the smart radiator valve has a spring-loaded dial which can only be turned a few degrees in either direction and when released, returns to the default position, as illustrated in Fig. 4.3.

Each clockwise or counter-clockwise turn represents a 1 °C raise or lowering of the setpoint temperature, respectively. Since there is no display on the device, without JuControl access this activity is trial-and-error since the user cannot see the temperature setpoint.

Fig. 4.3 Micropelt MVA005 smart radiator valve showing the rotation directions. The rotatable dial returns to the default position after being released. Each rotation changes the setpoint temperature up or down by 1 °C.

The JuControl calendar provides a robust presence scheduling interface, a key piece of the puzzle that allows the automatic control of the room temperature according to occupants' presence schedules. The calendar enables the occupant to input their schedule using an intuitive, drag-and-drop interface. Events in the calendar are classified internally in JuControl into three types of schedules: *default schedule* for regular, weekly repeated events; *adjusted schedule* for one-off events; and an *absence schedule* for temporarily overriding a weekly schedule with e.g. a holiday.

Merging Comfort Preferences for an Office

Each occupant provides an upper and lower setpoint temperature limit in the comfort preferences section of JuControl, in addition to the presence schedule provided via the calendar. This leads to multiple, possibly-conflicting setpoint temperature preferences for a multi-person office. JuControl merges these preferences into a single range using the "min-max / max-min" rules below (exemplified in Fig. 4.4).

- If no conflicts, the intersection of the temperature ranges is selected
- With conflicts, a "dead zone" is selected, which consists of the minimum of all maximum temperatures specified, and the maximum of all minimum temperatures specified. These are swapped if necessary to ensure minimum is not greater than maximum.

In Fig. 4.4, occupant A is alone between 07:00 and 08:00, so their preference is active. When occupant B arrives at 08:00, the preferences of the two occupants can be merged by taking the intersection according to the merging rule, until 09:00 when occupant A leaves. At 13:00, occupants C and D arrive, having non-intersecting (conflicting) temperature preferences. Hence, according to the rule, the maximum temperature for the office becomes D's minimum temperature, and the room minimum temperature is C's maximum.

Mathematically, given desired temperature bounds $\left[T_{\min,i}, T_{\max,i}\right]$ for each occupant i in a room of n occupants present in the office at the given time, the computed setpoint temperature range for the room, $[T_{\min}, T_{\max}]$, is

$$\begin{aligned} T_{\min} &= \min\left(\max\left(T_{\min,i}\right), \min\left(T_{\max,i}\right)\right) \\ T_{\max} &= \max\left(\max\left(T_{\min,i}\right), \min\left(T_{\max,i}\right)\right) \end{aligned} \qquad \forall i \in \{1..n\} \tag{4.1}$$

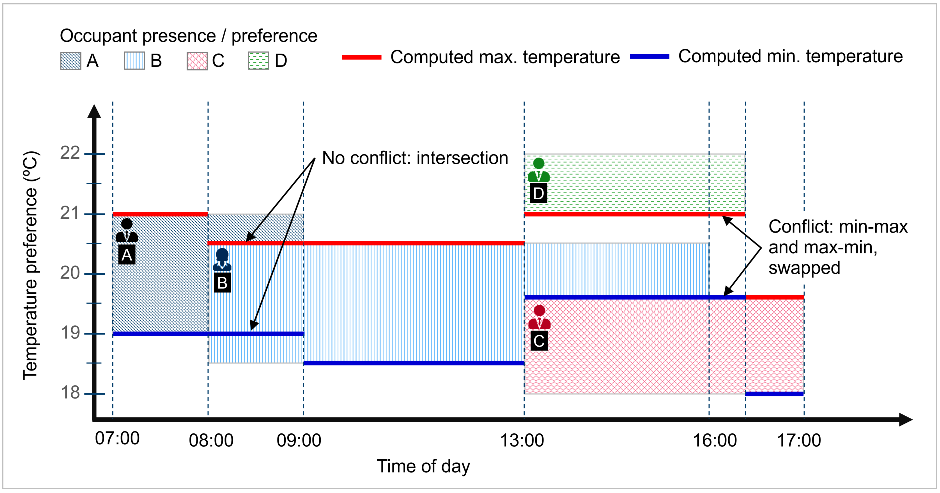


Fig. 4.4 Exemplification of JuControl algorithm for merging temperature preferences in a multi-person office, showing the resolution of conflicting (occupants C and D) and non-conflicting (occupants A and B) temperature preferences. The preference of each occupant is only considered when they are present in the office, according to the schedule.

4.4 ALICE

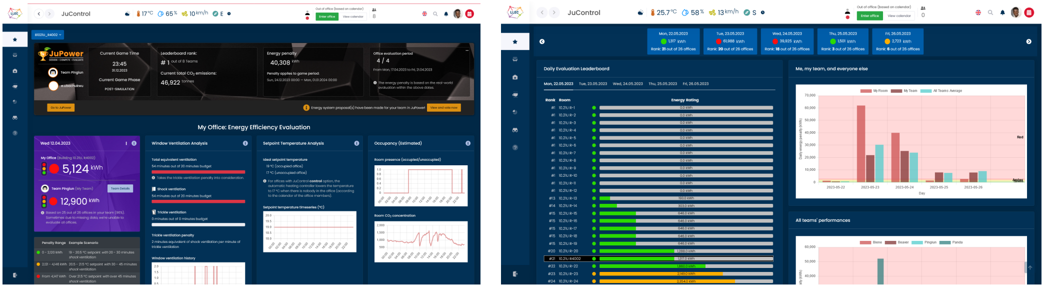
ALICE is both a mini-language and a tool for automatically generating geometrical diagrams of rooms, as well as linking the wireless sensors and actuators to room components, in order to support visualization and control. ALICE was conceptualised, implemented, and deployed within the scope of this thesis. The mini-language is designed to be efficiently used by a human operator to describe (possibly with pen-and-paper) the salient features of a room and its energy-related components (radiators, doors, and windows). In this project, several hundred rooms were thus described in ALICE, requiring two minutes or less to capture a room (for identical rooms, one room can be used as a template for the others). ALICE provides a web-based input editor with syntax analysis and real-time progressive visualisation of the end-result to facilitate the process of data capture. The ALICE parser then processes the descriptions, and the visualizer generates an SVG-based perspective view of the room as seen in Fig. 4.2a. After generating the geometrical representations of the room, ALICE then fetches devices (i.e. sensors and actuators) associated with each room from a device book-keeping tool called WALDO (see Redder et al. [146]), examines the properties of the imported devices, and then automatically associates each device with the appropriate room components the device is physically or conceptually attached to. All the data visualization in JuControl is done on the basis of the room component and devices data that ALICE supplies. Screenshots of the user interface of ALICE are shown in Appendix E.

4.5 Juracle: *Gamification*

Juracle[1] is the tool for estimating the energy performance of offices based on occupant interactions with the room, specifically with respect to heating and ventilation practices. The performance evaluations of the offices are derived as *energy penalties*, expressed in terms of wasted energy in kilowatt-hours (kWh) compared

[1]The name Juracle was derived from "Jülich Oracle", where the "oracle" terminology connotes the ability of Juracle to, like an oracle, "predict" an ideal behaviour profile for occupants.

to a predefined ideal baseline. These evaluations are run automatically at the end of each weekday, and the results and relevant contextual data are stored in a database, from where JuControl fetches and processes them for presentation to the user. Thus, Juracle is built on the *gamification* pillar, since it provides the basis for the introduction of game elements like points and leaderboards in JuControl. Fig. 4.5 shows examples of evaluation information presented to the user via JuControl.



(a) Home page

(b) Comparison of performance of offices user's team.

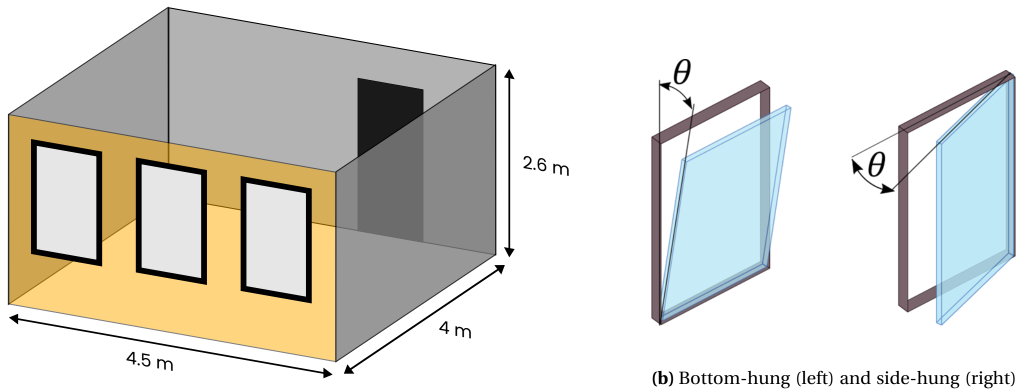
Fig. 4.5 Screenshots of JuControl showing the gamification page powered by Juracle. (a) The office evaluation. (b) The anonymized comparison of offices within a team for a particular evaluation date. Full-sized screenshots are provided in Appendix E.

The implementation of *Juracle* follows the mixed-mode approach discussed in the previous chapter. Here, the model of a reference room is implemented, which relates the investigated occupant actions (**setpoint temperature**, and **style and duration of window ventilation**) to the thermal energy demand of the room. In the following subsections, the implementation of the reference room model is presented first. Afterwards, the general scheme for simulating the model of the reference room to get the thermal demand for the predefined scenarios is described, followed by the results of the simulation and the derivation of the ideal setpoint temperature and ideal ventilation duration. Finally, the method for deriving energy penalties from the simulation results based on the ideal setpoint temperature and ideal ventilation duration is presented, along with the categorization of these penalties as *evaluative feedback* using a traffic-light rating system.

4.5.1 The Reference Room Model

The model of the reference room was developed using *Modelica* [147] (v4.0.0) and the *AixLib* model library [148] (v1.3.0) in the Dymola modelling environment [149]. The reference room, representative of a typical two-person office, has a floor area of 18 m^2 (4.5 m long and 4 m wide, with a height of 2.6 m), one outside wall and three inner walls (see the room geometry diagram in Fig. 4.6a). The outer wall has three windows, having each an area of 1.35 m^2 (1.5 m by 0.9 m). The floor, ceiling and inner walls are assumed to be adiabatic – no heat transfer occurs through them.

A schematic diagram of the high-level Modelica model is shown in Fig. 4.7. The reference room is heated by an ideal heater equipped with a proportional-integral controller, which is an instance of the `HeaterCoolerPI` component in the `AixLib.Utilities.Sources.HeaterCooler` Modelica package. The heater is oversized by design, which then allows the setpoint temperature to always be attained in a relatively short time even with fully opened windows. Thus, by this design decision the thermal energy demand of the reference room calculated as the measured demand of this heater can still be derived without worrying about the heater not being powerful enough to attain the setpoint temperature. In a realistic setting though, the heater in an office may not attain the setpoint temperature when the windows are fully open on a cold winter



(a) Reference room.

(b) Bottom-hung (left) and side-hung (right) windows.

Fig. 4.6 Reference room and window configurations for the derivation of the model parameters for mixed-mode approach. (a) The geometry of the reference room. (b) Tilted or bottom-hung (left) and normally open or side-hung (right) window configurations showing the opening angles.

day. Furthermore, the outer wall of the reference room is exposed to ambient conditions, including solar radiation, while the other walls are assumed to be adiabatic boundaries. Details of the window ventilation model are presented in Section 4.5.3.

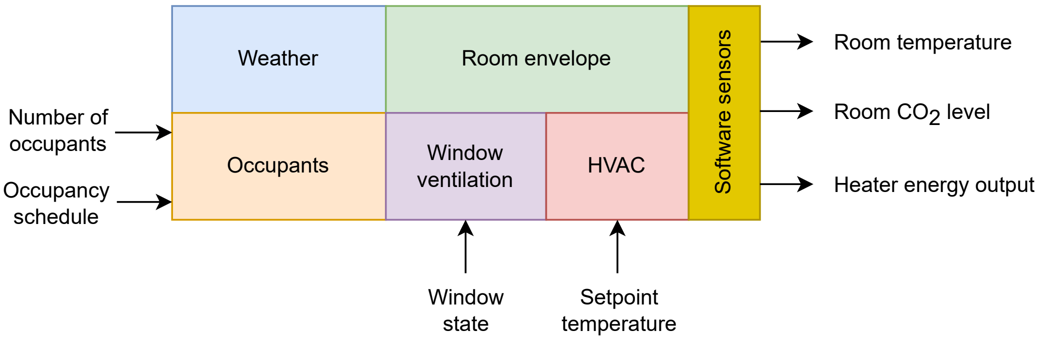


Fig. 4.7 Schematic representation of the Modelica model of the reference room showing the main sub-models with inputs and outputs.

4.5.2 Occupants Model

Occupants are modelled as both heat sources and CO_2 sources using a model that combines `CO2Balance` and `HumanSensibleHeatTemperatureDependent` models available in the `CO2` and `Humans` sub-packages, respectively, of the `AixLib.BoundaryConditions.InternalGains` package. The specific heat output per person was assumed to be 70 W and the activity degree of 1.5 met, using 1 met = 50 W/m^2. The CO_2 production rate for a single occupant is assumed to be 8.67 mg/s.

4.5.3 Window Ventilation Model

The exchange of air through the windows is modelled using the empirical relationships between window arrangement in the room, opening style (whether side-hung, i.e. normal opening, or bottom-hung i.e. tilted opening, see Fig. 4.6b), and the air exchange, derived from Richter et al. [150]. Specifically, in Richter et al. [150], the air change rate in the room is related to the appropriate driving temperature difference for the ventilation process (which depends on the ventilation duration) for different opening angles, styles, and window arrangements in multi-window rooms. Two phases of ventilation were identified Richter et al. [150].

- **Short-time ventilation**: This type of ventilation lasts only a few minutes and the temperature of the building envelope varies negligibly. The air change is then driven by the difference between the mean room air temperature and the ambient temperature, with the boundary conditions of the building envelope being considered constant. That is, the driving temperature difference for short-time ventilation is:

$$\Delta T_{\text{short}} = T_{\text{room,avg}} - T_{\text{amb}} \quad (\text{K}) \tag{4.2}$$

  where $T_{\text{room,avg}}$ is the mean room temperature and T_{amb} is the outside temperature.

- **Continuous ventilation**: As the duration of the ventilation increases, the building envelope cools down, and the ventilation process is then driven by the difference between the instantaneous wall temperature and the ambient temperature. This is because the room temperature would generally have attained some equilibrium with the ambient temperature at this point, due to the lower heat capacity of the air and the active air change that occurred at the short-time ventilation stage.

$$\Delta T_{\text{cont}} = T_{\text{wall}} - T_{\text{amb}} \quad (\text{K}) \tag{4.3}$$

  where T_{wall} is the mean instantaneous wall temperature.

In this thesis, the window model according to Richter et al. [150] was implemented in Modelica as the `WindowVentilationFlow` component (the schematic representation in Dymola is shown in Fig. D.2 of Appendix D). Multiple linearly interpolated lookup tables representing each combination of ventilation phase (short vs. continuous), style of opening (bottom-/side-hung) were implemented based on the `CombiTable2Ds` model of the `Modelica.Blocks.Tables` package. Each table takes as input opening angle and driving temperature difference to output the required air change rate. The active table depends on the ventilation process (short-term or continuous) and the window opening style (side-hung or bottom-hung). As already mentioned, the driving temperature difference used as input depends also on the active ventilation process. The duration of the short-time ventilation phase driven by ΔT_{short} was assumed to be two minutes in this thesis – it was not possible to determine an exact value from the literature. A gradual crossover from the air change rate for the short-time ventilation to that for the continuous ventilation is implemented using the `AixLib.Utilities.Math.Splice` component, which joins two values using a continuously differentiable function. The dynamics of room temperature and CO_2 concentration in response to ventilation in the model was based on the air exchange tables.

4.5.4 General Scheme for Simulation of Reference Model

The general scheme for deriving the relationship between energy consumption of the reference room and the inputs (setpoint temperature and window ventilation duration) is shown graphically in Fig. 4.8.

In this scheme, historical weather data, notably ambient temperature, solar radiation, and wind speed and direction, is first obtained for January to March (inclusive) of the two previous years for the geographic region (Jan. to Mar. 2021, and Jan. to Mar. 2022, respectively). These winter months were selected since the experimental run of the complete *Energy Dashboard Suite* was planned for a similar period. The data is then

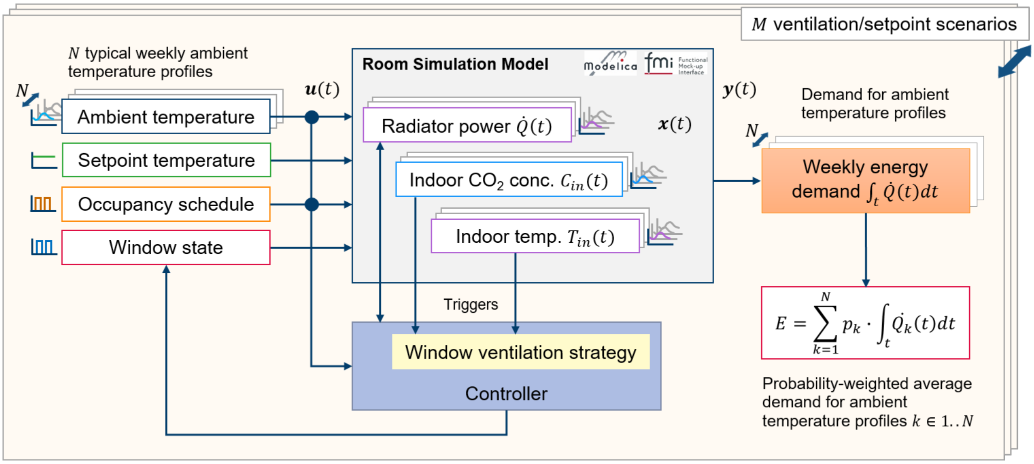
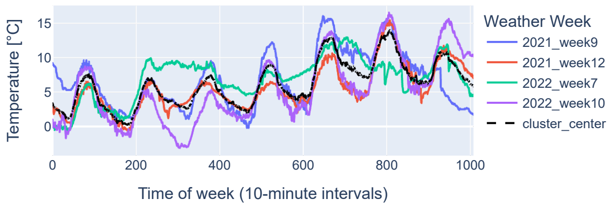


Fig. 4.8 Scheme for deriving weighting factors for rule-based evaluation using a simulation model. The weather is first classified into n profiles.

split into *weeks* and clustered based on ambient temperature using the K-means algorithm. To determine the optimum number of clusters, the *cluster inertia* (sum-of-squared-errors, or SSE) was plotted against the number of clusters in order to identify the "elbow joint" (see Fig. 4.9a). Fifteen (15) clusters were selected as a reasonable threshold. An example cluster with four members is shown in Fig. 4.9b, along with the "cluster center". For each cluster, a *cluster representative* determined as the cluster member with the least deviation from the cluster center, is taken as the representative weather profile for the cluster. An actual member of the cluster is selected as the cluster representative instead of selecting the computed "cluster center", in order to maintain the congruency between the ambient temperature used for clustering, and the other relevant weather measurements like solar radiation that are associated with the chosen ambient temperature profile. Finally, each cluster $k \in \mathscr{W}$ (where $\mathscr{W}$ is the set of weather clusters) is assigned a weight, p_k, proportional to the size of the cluster.

(a) Elbow method



(b) Sample cluster of weekly ambient temperatures.

Fig. 4.9 Weather clustering output. In (a), the cluster performance as measured by the cluster inertia (i.e. sum-of-squared-errors, or SSE) is shown, plotted against number of clusters. In (b), a sample cluster is depicted, with the "cluster center" shown.

The rest of the scheme proceeds as follows. Given the set of investigated setpoint-ventilation scenarios S, obtained by varying the setpoint temperature from 16 °C to 24 °C, and the ventilation duration from 0 minutes (no ventilation) to a full day (using the shock ventilation style where the windows are fully opened), for each scenario $s \in S$, run a simulation for each weekly weather cluster $k \in \mathscr{W}$ represented by the cluster representative, to obtain the thermal energy demand for one week, $E_{\text{th,week},s,k}$ kWh, for scenario s and weather profile k. The total energy demand $E_{\text{th},s}$ for scenario s across all weather profiles **for a single day** is then:

$$E_{\text{th},s} = \frac{\sum\limits_{k \in \mathcal{W}} p_k \cdot E_{\text{th,week},s,k}}{7 \times \sum\limits_{k \in \mathcal{W}} p_k} \qquad \text{(kWh)} \tag{4.4}$$

where p_k is the weight of cluster k, and the sum term in the bottom is the normalization factor. The division by 7 reduces the weekly demand to an average daily demand.

4.5.5 Results of Reference Room Simulation

The simulations of the thermal energy demand $E_{\text{th},s}$ of the reference room using the above-developed room model for each predefined scenario s produces a set of points in 3D space. These points are transformed into a C1-smooth (i.e. one-time continuously differentiable) surface by interpolation. The *interpolator* used is the `CloughTocher2DInterpolator` from the SciPy Python library [116], which triangulates the input data and then constructs a piecewise cubic Bezier polynomial on each triangle [151]. The contour plot from the interpolator for the simulated domain of setpoint temperature and ventilation duration is shown in Fig. 4.10, depicting the thermal demand of the reference room for the input domain for an average winter day. From the figures, the daily thermal demand ranges from less than 10 kWh to almost 60 kWh, and varies more strongly along the ventilation axis than along the setpoint temperature axis within the range of the respective variables, implying that more thermal energy can be potentially lost through wrong ventilation than through wrong setpoint temperature.

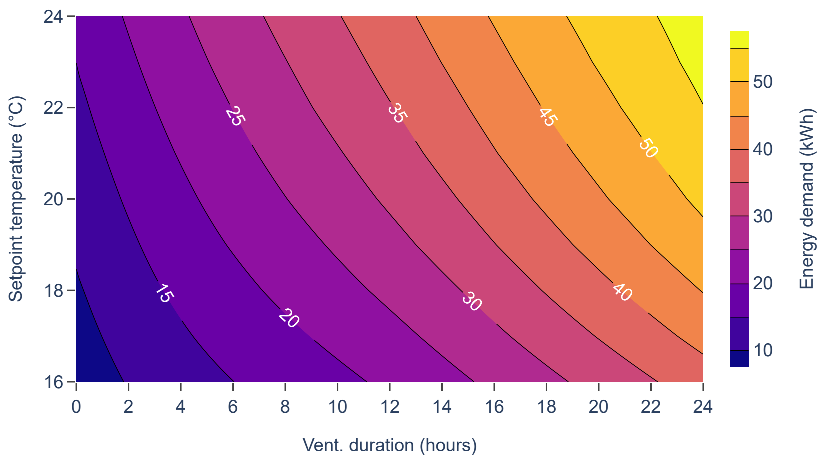


Fig. 4.10 Contour plot showing the simulated thermal energy demand of the reference room as function of ventilation duration and setpoint temperature after interpolation.

4.5.6 Derivation of Ideal Window Ventilation Strategy

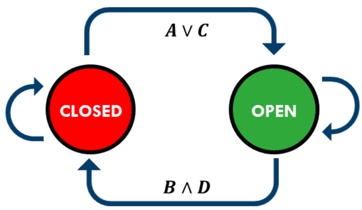
In addition to deriving the energy demand of the room for various combinations of the investigated inputs, the room model was also used to derive the benchmark for ideal ventilation in terms of ideal minutes of shock ventilation, $N_{\text{vent,ref}}$. This provides a soft maximum, above which penalties apply. The underlying assumption that permits the extrapolation of these results to rooms of different sizes is that rooms tend to have a comparable window-to-floor-area ratio to the reference room, since the larger a room is, the more windows are built into it, in general. This implies that comparable air exchange efficiencies can usually be achieved for small rooms as for large rooms, provided that all the available windows are opened during ventilation. Indeed, the simultaneous opening of all windows is well within the control of the room occupant, justifying its consideration in the award of penalties. However, this approach does not consider possibly higher occupancy (per unit floor area) than was accounted for in the models, such as in densely populated offices and meeting rooms. In fact, the rule for occupant density in Germany for offices is that the office floor area is at least $12\ \mathrm{m}^2 + n \times 6\ \mathrm{m}^2$, where n is the number of occupants in the office. Thus, as n increases, the area per occupant approaches $6\ \mathrm{m}^2$/person, which is one-half of the case in the single-person office. (Workarounds for this limitation are addressed in the recommendations section of Chapter 8.)

In the simulation schema of Fig. 4.8, the Window ventilation strategy block contains a rule-based controller that ventilates in response to indoor CO_2 concentration and ambient temperature. The derivation of the controller logic is shown in Fig 4.11. The CO_2 concentration lower bound was set at 900 ppm, and upper bound at 1500 ppm, where the window is opened when the CO_2 concentration exceeds the upper bound, and closed when the concentration falls below the lower bound. Lower and upper bounds for the ambient temperature are 15 °C and 17 °C, respectively. Apart from the buffer regions bounded by the upper and lower bounds, which help to prevent control output oscillations, the control logic also enforces a minimum open / closed duration of 5 minutes to prevent chattering even with oscillating input signals. In the control logic of Fig. 4.11b, the window is closed if previously open and both CO_2 concentration and ambient temperature are below their respective lower bounds, and is opened if previously closed when either CO_2 concentration or ambient temperature is higher than the upper bound. Under every other condition, the window remains as-is. Naturally, in a real-world setting other relevant factors could come into play, like external noise and very high ambient temperatures; these are not modelled in the controller.

Prev. State	A	B	C	D	Next State	Transition
	CO_2 > UB	CO_2 < LB	T_{amb} > UB	T_{amb} < LB		
0	0	0	0	0	**0**	-
1	0	0	0	0	**1**	-
0	1	x	x	x	**1**	0 → 1
1	1	x	x	x	**1**	-
0	0	1	0	0	**0**	-
1	0	1	0	0	**1**	-
0	x	x	1	x	**1**	0 → 1
1	x	x	1	x	**1**	-
0	0	0	x	1	**0**	-
1	0	0	x	1	**1**	-
0	0	1	x	1	**0**	-
1	x	1	x	1	**0**	1 → 0

1: open/true; 0: closed/false; x: "Don't care"; UB: upper bound; LB: lower bound

(a) Window controller truth table



(b) Window controller state machine

Fig. 4.11 Derivation of the control logic for the window controller used to obtain ideal ventilation duration in response to indoor CO_2 concentration (CO_2) and ambient temperature (T_{amb}). (a) The truth table representing the controller logic. (b) The equivalent state machine of the controller after optimization of the truth table logic.

According to the results of this simulation, for a single occupant in the reference room working a typical 8-hour day, 20 minutes of *shock* ventilation divided into two instances of 10 minutes each was sufficient to keep the CO_2 levels (and by implication the air quality) within the above-mentioned range. Hence, an **ideal ventilation duration** of **20 minutes per working day**, i.e. considering an average of about 8 hours per day in the office, was applied in the evaluation methodology of this thesis. Note that a more stringent air quality standard (e.g. an upper bound of 1200 ppm CO_2 in the model above instead of 1500 ppm) would of course lead to a longer ideal ventilation duration; indeed, Schakib-Ekbatan et al. [9] considered 40 minutes per 8-hour working day as optimal, although as they mention, some other sources recommend much less (e.g. assuming the three minutes of ventilation every hour in [152], this comes to 24 minutes per working day). Nevertheless, the 20 minutes quota per 8-hour working day used in this thesis also serves as a lower bound, so that any shorter ventilation durations are *not* rewarded. Also, as discussed in the traffic light rating system (Section 4.5.9), up to 30 minutes ventilation per work-day still lies within the green zone, provided setpoint temperature is ideal.

4.5.7 Derivation of Ideal Setpoint Temperature

In order to determine an "ideal" setpoint temperature, which forms the basis for benchmarking occupant setpoint temperature efficiency, thermal comfort requirements in accordance with the DIN EN 16798 standard are considered. These standards use both the Predicted Mean Vote (PMV) and Adaptive Comfort models. For offices, the standard recommends the minimum room temperatures shown in Table 4.1 according to the different comfort categories I - IV defined by the standard, assuming *typical conditions*. The typical conditions for the derivation of these temperatures are: clothing = 1.0 clo (typical indoor winter clothes); relative humidity = 40%; activity level = 1.2 met, and air speed < 0.1 m/s.

Table 4.1 Recommended temperatures for winter for offices, according to DIN EN 16798 standard. These recommended temperatures were determined for the following conditions: clothing = 1.0 clo; relative humidity = 40%; activity level = 1.2 met, air speed < 0.1 m/s.

Category	**Minimum temperature**
I	21 °C
II	20 °C
III	19 °C
IV	18 °C

Under the conditions of the DIN EN 16798 stated above, selecting an ideal setpoint temperature of 19 °C office implies Category III compliance, as shown in Table 4.1. However, adjusting the conditions to reflect the local area of Jülich (relative humidity of > 70% in winter) and the insulating property of typical office chairs (+0.15 clo [153]), the 19 °C then falls into Category I, as determined by the CBE Thermal Comfort Tool [154].

In the light of the foregoing, in this thesis the **ideal setpoint temperature** for an **occupied office** was chosen to be:

$$T_{\mathrm{sp,ref,occ}} = 19\,^{\circ}\mathrm{C}$$

In addition to the determination of the ideal setpoint temperature for occupancy using the DIN EN 16798 standard, another point that supports the use of 19 °C as the ideal setpoint temperature is the recent policies of the German government regarding space heating, especially as a result of the recent upheavals in the European energy market due to the ongoing Russia-Ukraine war. Specifically, the German government stipulated some measures to avoid energy wastage, including a maximum setpoint of 19 °C for heating in public buildings [155]. Hence, the choice of 19 °C as the ideal setpoint temperature underscores these policy efforts.

It should be noted that while an ideal setpoint temperature of 19 °C is selected in this work, the evaluation system is designed such that minor violations are also acceptable. The level of acceptability is given as feedback to the user by means of a traffic-light feedback system, the details of which are presented in Section 4.5.9.

For periods of **no occupancy**, the **ideal setpoint temperature**, was chosen in this thesis as:

$$T_{\mathrm{sp,ref,unocc}} = 17\,^{\circ}\mathrm{C}$$

which allows saving energy, but at enables ramping up quickly enough after cold nights and weekends. For rooms where heating can be controlled via the user presence schedule captured by the JuControl calendar, the heating controller algorithm reduces the setpoint temperature to 17 °C when the room is unoccupied. As discussed in Section 4.3.1, JuControl features a calendar through which occupants of selected buildings indicate their presence in the office, and the automatic heating controller uses this presence profile to determine room the setpoint temperature for the room. Hence, the reference temperature for calculating setpoint temperature *deviation* depends on occupancy, which in turn is a function of time. In other buildings with an in-built building management system (BMS) that features a temperature regulation subsystem and a corresponding HMI for manipulating the subsystem (usually present in those buildings as wall-mounted displays), a similar occupancy-dependent ideal temperature setpoint profile is assumed for behaviour evaluations.

Therefore, to obtain the setpoint deviation $\pi_{sp}(t)$ at time t, the setpoint temperatures $T_{\mathrm{sp,ref,occ}} = 19\,^{\circ}\mathrm{C}$ and $T_{\mathrm{sp,ref,unocc}} = 17\,^{\circ}\mathrm{C}$ are substituted into Eq. 3.1. In Fig. 4.12, a scenario is illustrated for which the setpoint deviation is derived, considering that the heating controller in JuControl-*controlled* offices tracks the user's presence schedule, hence applying the user-specified setpoint temperature, T_{sp}, during the periods of occupancy, as dictated by the user-supplied schedule. During periods of absence, the controller falls back to $T_{\mathrm{sp,ref,unocc}} = 17\,^{\circ}\mathrm{C}$. Given the multiple possibilities for adjusting the temperature setpoint in JuControl-controlled offices ("manual override" via the physical radiator valve or the temperature widget in JuControl, or else "automatic mode" relying on presence schedules and personal comfort preferences), deriving the deviation is a little more complex. Normally, where only the automatic schedule-based heating control is used, it suffices to check if the user's schedule was adhered to. However, due to other possibilities for specifying the setpoint temperature, unscheduled presence in the office that necessitates a manual override should also be considered as correct behaviour. Hence, in Fig. 3.3, the deviation immediately after the manually triggered setpoint is judged based on the occupied status (i.e. $T_{\mathrm{sp,ref,occ}} = 19\,^{\circ}\mathrm{C}$ is the reference temperature), even though it is outside the calendar-specified presence schedule.

4.5.8 Deriving Energy Penalties from Deviations

Given the ideal ventilation strategy and setpoint temperature determined as detailed above, the final step in the mixed-mode approach is to derive an energy penalty from the deviations from these ideal cases using the developed reference room model. The general idea is that for each minute of the evaluated day, the energy penalty for that minute is the difference between the energy demand of the reference room computed at two operating points: the *ideal operating point* and the *real operating point*. The *ideal operating point* is specified by the ideal ventilation duration $N_{\mathrm{vent,ref}}$ and the ideal setpoint temperature for that minute $T_{\mathrm{sp,ref}}(t)$ (Eq. 3.2). The *actual operating point*, on the other hand, is specified by the total *equivalent ventilation* duration for the day in the evaluated office $N_{\mathrm{vent,eq}}$ (see Eq. 3.5) and the actual setpoint temperature for the given minute, $T_{\mathrm{sp}}(t)$.

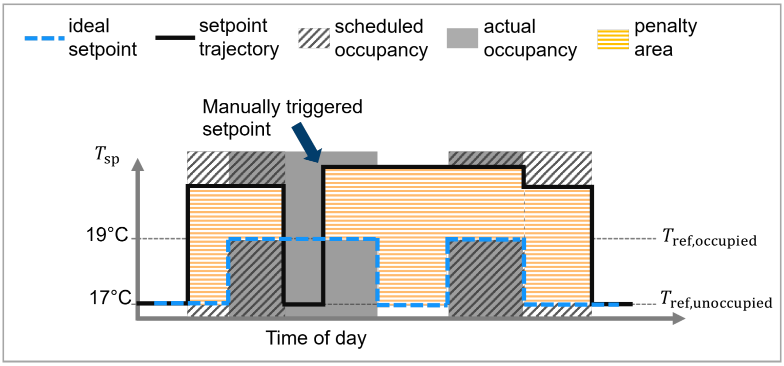


Fig. 4.12 Illustration of the derivation of the setpoint deviation for a hypothetical scenario. The automatic heating controller tracks the presence schedule of the occupant (for JuControl-controlled offices). Deviations are calculated based on actual presence in the office, and not the user-supplied presence schedule (the "ideal setpoint" curve in blue). The "Manually triggered setpoint" reflects manual override of the controller setpoint temperature by the occupant during an out-of-schedule presence.

Computing the energy demand of the reference room at an operating point involves using the interpolator of Section 4.5.5 to determine the energy demand corresponding to the setpoint temperature and ventilation duration specified by the operating point (see Fig. 4.10). Thus, the energy penalty for an evaluated day is given by

$$E_{\text{th,pen}} = \frac{1}{1440} \sum_{t=1}^{1440} E\left(T_{\text{sp}}(t), N_{\text{vent,eq}}\right) - E\left(T_{\text{sp,ref}}(t), N_{\text{vent,ref}}\right) \qquad \text{(kWh)} \tag{4.5}$$

where the $E(x, y)$ notation represents the energy demand of the reference room for the operating point given by (x, y) as described above, T_{sp} and $T_{\text{sp,ref}}$ are the real and ideal setpoint temperatures, and $N_{\text{vent,eq}}$ and $N_{\text{vent,ref}}$ are the real and ideal ventilation durations, as described in the previous paragraph. The division by 1440 scales the demand computed by $E(\cdot,\cdot)$ to one minute, since this energy demand from the reference model is the *daily* energy demand where the given setpoint temperature is assumed constant for the entire day and the given ventilation duration is the total for the whole day. Thus, the energy penalty $E_{\text{th,pen}}$ for a given day is the cumulation of all minute-wise differences between the thermal demand of the real operating point $(T_{\text{sp}}(t), N_{\text{vent,eq}})$ and that of the ideal operating point $(T_{\text{sp,ref}}(t), N_{\text{vent,ref}})$ over the entire day, i.e. for $1 \leq t \leq 1440$. This is illustrated in Fig. 4.13 below.

From the figure, considering a day with 30 minutes of ventilation and a constant setpoint temperature of 20 °C throughout the day (point C), the energy penalty for each minute t in the day when the office is unoccupied is $E_{\text{th,pen}}(t) = E(C) - E(A)$ kWh (i.e. using point A as ideal reference), while for each occupied minute it is $E_{\text{th,pen}}(t) = E(C) - E(B)$ kWh (i.e. using point B as ideal reference). The final daily penalty is then the sum of these respective minute-wise penalties over the entire day. The "forbidden area" of the figure is an area in which no operating point can exist, which is the effect of the over-compensation protection built into the evaluation system that is enforced by the use of limiting constructs like $(\cdot)^+$ in Eq. 3.1 and $\max(0,\cdot)$ in Eq. 3.4. Since the deviations are bounded below by zero using these limiting constructs, it means that it is impossible to have an operating point in which the ventilation duration is less than $N_{\text{vent,ref}} = 20$ min or setpoint temperature less than $T_{\text{sp,ref,unocc}} = 17$ °C for unoccupied periods or $T_{\text{sp,ref,occ}} = 19$ °C for occupied periods.

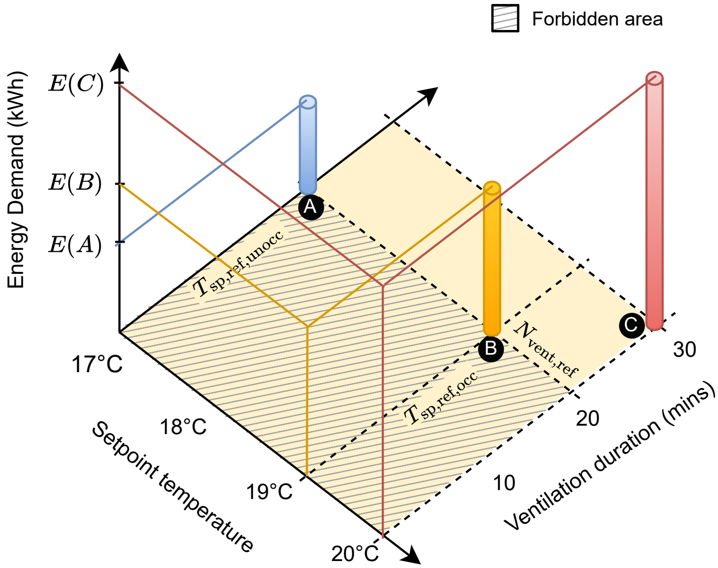


Fig. 4.13 Illustration of the derivation of the energy penalty from the reference room model considering the energy demand $E(C)$ kWh of the real operating point C = (20 °C, 30 min), and those of the ideal operating points $E(B)$ kWh for point B = (19 °C, 20 min) (for occupied office) and $E(A)$ kWh for point A = (17 °C, 20 min) (for unoccupied office).

4.5.9 Traffic Light Rating

A three-colour traffic light rating system was developed based on the energy penalties to categorize penalty values into green, amber, and red zones. The purpose of this rating system is to facilitate the interpretation of the energy penalty and provide clear and intuitive feedback on what is expected of the user. In essence, the traffic-light feedback provides a form of *injunctive normative feedback* that tells the recipient what is acceptable behaviour [88, 91], but with the social basis of comparison in the traditional definition of normative feedback replaced by an empirically-derived benchmark derived by Juracle. From a behavioural science perspective, the energy penalties provide *objective feedback*, while the traffic light system provides *evaluative feedback*; this combination has been shown to be more effective than either type of feedback in isolation [82]. In the area of energy-related gamification, a similar objective-evaluative feedback approach was adopted by Wemyss et al. [47], but with emojis instead of a traffic light system.

The range of penalties assigned to each traffic light colour is shown in Table 4.2, along with the reference scenarios used to derive the ranges. The reference scenarios are arbitrarily predefined values of setpoint temperature and *equivalent* ventilation duration whose energy penalties set the upper bound for the corresponding traffic light. These scenarios were chosen to accommodate in principle individual differences in thermal and indoor air quality preferences without unduly pressurizing occupants. In other words, slightly exceeding the recommended setpoint temperature and ventilation duration still shows green light, indicating to the occupant that they are still within an approved range, although the penalty is non-zero. It should be noted that the reference scenarios were only used as an introspectable mechanism for generating the penalty boundaries for the traffic lights; other combinations of setpoint temperature and ventilation deviation could possibly lead to the same penalty values as the reference scenarios.

Table 4.2 Determination of the upper boundaries of energy penalties for the "traffic light" rating system based on reference scenarios. The penalty values are given both at the scale of the reference room and the FZJ campus.

Traffic light colour	Reference scenario of upper boundary	Daily energy penalty range (kWh)	
		Single office	Campus scale
Green	$N = 30$ min; $T_{occ} = 20.5\,°C$	0 – 0.8683	0 – 2120
Amber	$N = 45$ min; $T_{occ} = 21.5\,°C$	0.8683 – 1.6981	2120 – 4146
Red	$N = 60$ min; $T_{occ} = 22.5\,°C$	More than 1.6981	More than 4146

For each traffic light colour, two penalty ranges are given: the "single office" range is based directly on the energy demand of the reference room, while the "campus scale" ranges are when the penalties derived for the reference room are scaled to the area of the campus. Furthermore, for deriving the energy penalties corresponding to the reference scenarios, an occupancy period of 8 hours as in a typical workday is assumed, so that the given setpoint temperature is applied during these 8 hours, while the rest of the day assumes the unoccupied reference temperature $T_{sp,ref,unocc} = 17\,°C$. By this token, the non-zero setpoint temperature *deviation* caused by the reference scenario only occurs during the period of occupancy when computing the penalty using Eq. 4.5.

4.6 Occupancy Detection

The (binary) presence of occupants, which forms a key input into Juracle for behaviour evaluation, was determined via presence detectors in some buildings where this was available and accessible, and via so-called environmental sensing based on indoor CO_2 concentration for other buildings. (See Table 3.6 for details on the level of instrumentation in each of the considered buildings.)

4.6.1 Use of Presence Detector

For the buildings with presence detectors, the devices are PIR sensors manufactured by MDT Technologies GmbH (Model SCN-P360D3.03), and were installed in the buildings as part of the Building Management System (BMS). They communicate over wired KNX and have a horizontal field detection of 360°, which works by incorporating three sensors at 120° angles to one another, thereby covering the entire horizontal field of view. The sensors have a detection radius of 5 m for presence, and 11 m for movement, when installed at the maximum recommended height of 4 m. They can record data in intervals of as low as one second, although the deployed sensors are configured for higher intervals (about one minute). The PIR sensors in these buildings were configured for movement-detection by Facility Management, making them susceptible to false negatives when the occupant is not moving. To get around this limitation, and to also eliminate false positives, a boundary assumption is made regarding initial arrival time of occupants, and their final departure time for a given day. Under this assumption, the first detection event of the day after 06:00 is the initial arrival time, and the last detection event before 20:00 is the final departure time. Thus, it is assumed that an arrival or departure always triggers the sensors. In-between initial arrival and final departure, there are possible periods of absence. The methodology devised in this thesis to derive the presence of the occupant from these readings is described next.

For the analysis, consider the detection events as constituting a "chain of events" $\{e_1, e_2, ..., e_n\}$ separated by some time interval δ_{ij} minutes between any two consecutive events e_i and e_j, with e_1 and e_n being the initial arrival and final departure events respectively. (The $\{\cdot\}$ notation is to be understood here as representing an ordered set.) To begin, therefore, the time interval from the initial arrival event until the next detection event is classified as presence, as well as every subsequent consecutive interval having duration less than 30 minutes before the next detection event. That is, presence is assumed for the interval spanning the initial chain L_{initial}, given by:

$$L_{\text{initial}} = \{e_1, e_2\} \cup \left\{e_i \mid \delta_{i-1,i} < 30, i \geq 3\right\} \tag{4.6}$$

The same approach is applied starting from e_n and moving backwards to form the final chain, i.e.

$$L_{\text{final}} = \text{reversed}\left(\{e_n, e_{n-1}\} \cup \left\{e_{n-1-j} \mid \delta_{n-1-j,n-j} < 30, j = 1, 2, \ldots\right\}\right) \tag{4.7}$$

The final chain stops when any intersection occurs between the initial and final chains, i.e. as soon as $L_{\text{initial}} \cap L_{\text{final}} \neq \emptyset$. Otherwise, for all other detection events, the algorithm is run in the forward direction repeatedly, starting with the first event not captured by any of the existing chains (initially L_{initial} and L_{final}, but subsequently any new chains formed by the repeated forward runs). The intersection termination criteria applies still. Finally, any chain with only one element is joined to *both* the preceding and the succeeding chain. The intuition is that such long islands without detection should correspond to periods of physical stillness for the occupant, given that the boundaries have been sensibly considered.

By implication from the foregoing, the separating intervals between chains correspond to absences. For the edge case where only one event exists (corresponding to an initial arrival event, e_1 in our model), we assume the final departure event to occur at 17:00, if later than e_1. Otherwise, an absence is assumed for the day. Note also that if the final departure time is encountered in the first run, i.e. $e_n \in L_{\text{initial}}$, this implies presence for the entire period between initial arrival and final departure, and the algorithm terminates. While in practice all the assumptions might not hold, the presence detection accuracy was sufficient for the needs of the project.

4.6.2 Use of Indoor CO_2 Concentration

In this work, the CO_2 mass-balance approach was used to determine occupancy in buildings for which there were no available presence sensor data. The occupancy estimation equation below based on the CO_2 mass balance is used (adapted from [130]). The number of occupants $N_{\text{occ},i}$ for time step i is given by:

$$N_{\text{occ},i} = \frac{\left[\left(1 - \frac{\frac{\dot{m}_{\text{airx},i}}{\rho_{\text{air}}} \cdot \Delta t_i}{V_{\text{room}}}\right) \cdot C_{\text{R},i-1} + \frac{\frac{\dot{m}_{\text{amb},i}}{\rho_{\text{air}}} \cdot \Delta t_i}{V_{\text{room}}} \cdot C_{\text{amb},i} + \frac{\frac{\dot{m}_{\text{corr},i}}{\rho_{\text{air}}} \cdot \Delta t_i}{V_{\text{room}}} \cdot C_{\text{corr},i}\right] - C_{\text{R},i}}{\frac{\dot{c}_{\text{pp},i} \cdot \Delta t_i}{V_{\text{room}}}} \tag{4.8}$$

where N_{occ} is the number of occupants, $\dot{m}_{\text{airx}}$ is the mass flow rate of air exchange between the room and its surroundings, and $\dot{m}_{\text{amb}}$ and $\dot{m}_{\text{corr}}$ are the mass flow rates of air into the room from outside and the corridor, respectively, all in kg/s. C_{room}, C_{corr}, and C_{amb} are the CO_2 concentrations in the room, the corridor, and the outside air (in ppm). V_{room} is the volume of the room (in m^3), ρ_{air} is the density of air (in kg/m^3), Δt is the duration of the time step in seconds, and $\dot{c}_{\text{pp}}$ is the rate of production of CO_2 by an average occupant (in ppm/s). A brief survey was done in Building B-02 in which occupants of three rooms recorded their incoming and outgoing times, in order to calibrate the occupancy model. Afterwards, a Mixed Integer Quadratically Constrained Program (MIQCP) was solved to determine the air infiltration rates, air exchange rates through fully opened windows and doors, and the CO_2 concentration in the corridor.

CO_2 Sensor Calibration

Several CO_2 sensors deployed in the offices required recalibration post-installation for their readings to be meaningful. Although an attempt was made to calibrate each sensor following the manufacturer-recommended approach before installation, in which the sensors are put into "calibration mode" and then set in an environment with ambient conditions, wide discrepancies and offsets still existed in the deployed sensors. This necessitated a data analysis approach for determining the CO_2 sensor baseline corresponding to ambient conditions, using the fact that during long periods of absence, eventually the CO_2 reading of the sensors would approach ambient conditions. These baselines are then stored, and the offset bias of the sensor from an assumed standard ambient CO_2 concentration of 500 ppm, is then actively factored in whenever the sensor data is fetched. An example of the baselines for CO_2 sensors in a small building is shown in Fig. 4.14. As can be seen, many of the sensors were not properly calibrated by the installation technician, all the more justifying the software-based calibration.

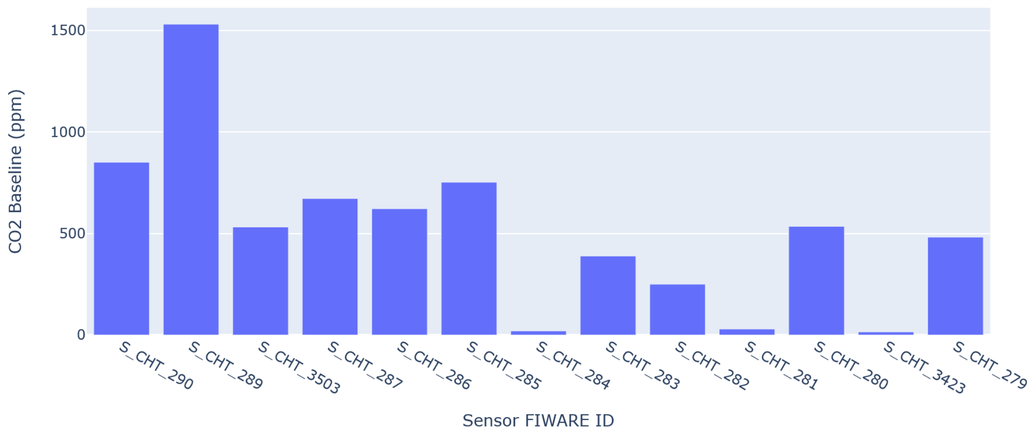


Fig. 4.14 CO_2 sensor baselines (corresponding to ambient concentration assumed to be 500 ppm) for a small building, determined from the analysis of historical data produced by the sensors.

Chapter 5

Run of Experiment and Analysis of Results

In this chapter, the execution of the experiment described in Chapter 3 is first discussed, highlighting the timing and environment of the experiment, and the issues encountered in the process. Afterwards, the results of the experiment are discussed, starting with analysis of user engagement, then followed by the testing of the hypotheses of Chapter 3 along with related results, and concludes with an estimate of energy savings in the pilot building using the energy signature methodology of Section 3.9. Some of the results presented here have been published in Ubachukwu et al. [156].

In the analysis of the results, several factors are involved beyond the experimental variables, and thus have to be considered to some degree since they generally affect the results. The most significant factors are weather (specifically ambient temperature), occupancy (binary presence in the offices), and frequency of user interaction with the developed applications. The frequency of user interaction is difficult to estimate in this work, since this interaction was not tracked, and the web server log files that could have provided some insights were unavailable as at the time of analysis, as the web server had overwritten them as a result of "log file rotation". The user survey responses, discussed in Chapter 7, provide some hints, but are not an accurate estimate of the degree of interaction in general. These factors are analysed alongside the results whenever they are deemed to be significant for the particular result.

Furthermore, as part of the analysis of the evaluation performance of the teams in the following sections, the penalty ratings of the teams are also deconstructed into their constituent factors, in order to examine how the measured behaviour aspects (ventilation pattern and setpoint temperature) affected the penalties in each case. As already established in this thesis, **ventilation pattern** and **setpoint temperature** in offices are the two factors used to evaluate the energy efficiency of occupant behaviour, where these factors are computed as *deviations* from an *ideal* scenario. The mixed-mode approach used in this thesis then combines these factors into a single penalty rating using a physics-based model, as described in Section 3.7. To compute the *deviation* of the observed ventilation from the ideal case, the evaluation methodology of the thesis separates shock ventilation (side-hung opening) from trickle ventilation (bottom-hung opening), and applies an arbitrarily chosen penalty factor $f_{\text{pen,trickle}} = 2$ to discourage trickle ventilation, i.e. each minute of trickle ventilation is billed as equivalent to two minutes of shock ventilation. Thus, **equivalent ventilation deviation** considers *shock ventilation duration*, *trickle ventilation duration*, and the *trickle ventilation penalty*, all measured in minutes (see Eqs. 3.4 - 3.5), while the setpoint temperature deviation, measured in *degree-minutes*, considers *actual* occupancy and the setpoint temperature profile for the office.

5.1 Running the Experiment

The experiments officially started on Monday, 13.03.2023 and ran for seven weeks, until the end of Friday 28.04.2023. Fig. 5.1 provides a pictorial timeline of the major activities that were performed regarding the experiment.

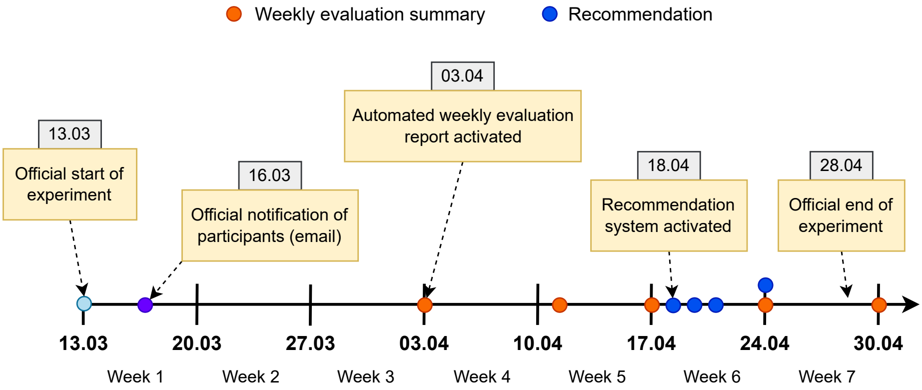


Fig. 5.1 Timeline of activities related to the experiment.

5.1.1 User On-boarding

The occupants of the experimental offices were informed by email on 16.03.2023 – a total of about 870 employees. Each email was tailored to the experimental group and the building where the occupants were located, based on available features in the team (and transitively, the experimental group) to which the room occupants belong, and based on the degree of instrumentation in the building. However, users were not informed about the existence of different experimental groups, and which features are absent from their own experimental group. On the other hand, users were aware of the existence of their and other teams, as this information was contained on the user interfaces and email communications. Furthermore, a poster was displayed at each building (or floor) entrance, customized according to the features available to the teams in that building or floor. The language of the poster (German or English) was also tailored to suit the personnel demographics of the deployment site (depending for example on the field of expertise and percentage of international colleagues). An example deployment poster is shown in Appendix B (Fig B.17).

Furthermore, a presentation regarding the availability of JuControl was made to the institute that make up the bulk of Team T2 on 23.03.2023. Several months earlier (during the pilot phase before the experiment), a similar presentation was made at the institute of Building B-01 (who were assigned to Team T1 for this experiment), along with the release of a "JuControl manual". These activities led to more users granting their consent for JuControl to become available to their office (i.e. more activated offices in JuControl).

Finally, the implementation of automated emails (weekly evaluation summaries, and energy efficiency recommendations) also triggered more participation as a side-effect, since these emails were sent to participants irrespective of whether they had hitherto actively participated in the experiment (e.g. by interacting with any of the apps in the Energy Dashboard Suite) or not. As can be seen in Fig. 5.1, weekly evaluation summary emails were sent every week beginning on 03.04.3023. Likewise, the recommendation system was activated on 18.04.2023 and sent at most one email per day to each participant, as needed. The ability to control the emails (at a fine granularity corresponding to particular types of emails, or using a "master switch" to generally disable emails) was also implemented. Instructions and a link to the email settings page of JuControl were included in the automated emails. The on-boarding effect of these automated emails are discussed further in the results section (Section 5.2).

5.1.2 Structure of Evaluation Summary and Recommendation Emails

Both evaluation summary and recommendation emails contained both English and German text. An evaluation summary email states the weekly average penalty for the office and team, and shows the rank of the recipient's office in the team and the rank of the team amongst other teams (see Fig. B.15 in Appendix B for a sample). It also compares the current performance of the recipient's team with its previous performance. A recommendation email, on the other hand, was sent in response to exceeding Juracle-determined thresholds for setpoint temperature (1.5 °C above ideal) and ventilation (10 min above ideal ventilation duration), and for trickle ventilation (if lasting for more than 5 min). An example email is shown in Fig. B.16 in Appendix B for exceeding the ventilation duration threshold.

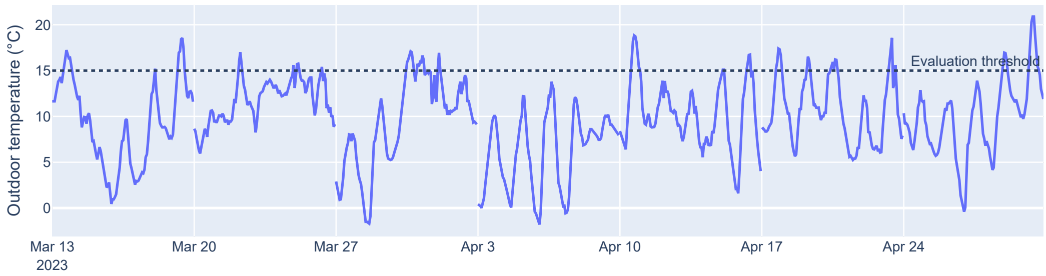
5.1.3 Suspension of Evaluation during Warm Periods

In order to account for the fact that heat losses through ventilation are minimal during warm periods in a day, the evaluation system pauses evaluations when the outside temperature reaches or exceeds 15 °C, since heating may no longer be required. This means that during such periods, the evaluation penalty is kept at zero. Another justification for the suspension of evaluation in war periods is that higher ambient temperatures naturally induce longer ventilation durations to achieve air freshness than colder ambient temperatures, since the temperature gradient-driven air exchange rate is lower for smaller temperature differences between room and outside air. Yet again, the model of the reference room used in the evaluation system was derived for temperatures generally below 15 °C, so the evaluation system was more suitable for colder periods. Nevertheless, evaluations were still run for other periods of the day when the ambient temperature was not above the threshold, which has the drawback that effects like thermal storage capabilities of the building envelope could still make the environment feel warm for the occupant, warranting more justified ventilation which is not necessarily energy-wasting.

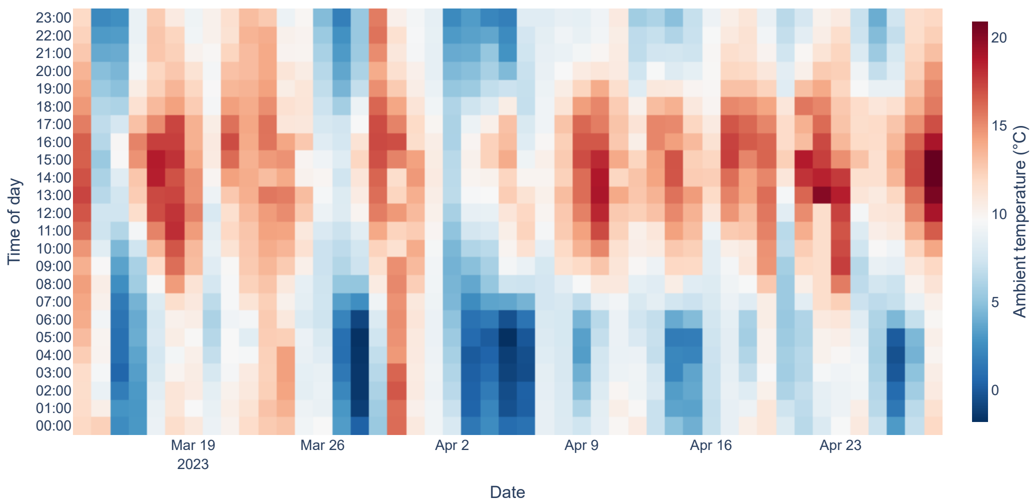
The ambient temperature for the experiment period is shown in Fig 5.2a, with the 15 °C threshold marked with a dashed horizontal line. In Fig 5.2c, the distribution of temperature over each day is shown in a heatmap for the period. The number of minutes during which evaluation was suspended due to high ambient temperatures is shown for each day of experiment in Fig 5.2c. As can be seen, several days within the experiment period had long durations (up to several hours) in which the ambient temperature was at or above 15 °C.

5.1.4 Technical and User Issues

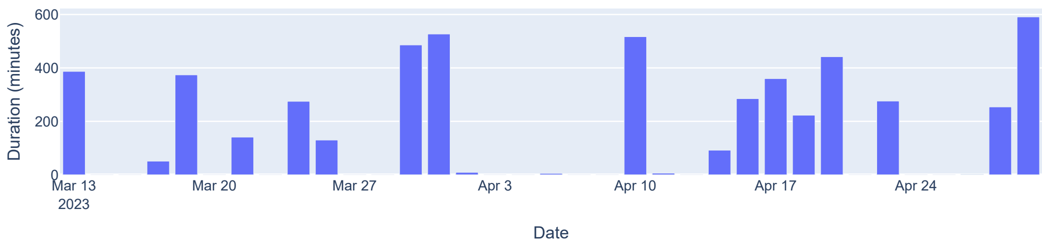
By the beginning of Week 7 of the experiment (on 24.04.2023), a total of 104 offices and meeting rooms had been deactivated from evaluations across 10 buildings, due to hardware and configuration faults. Initially, 84 offices and meeting rooms were deactivated in response to occupant reports and related human analysis. Out of these, about a fifth was directly due to occupant reports and subsequent verification of the reports. The remainder was selected based on human analysis of sensor data that detected patterns in evaluation



(a) Ambient temperature.



(b) Hourly ambient temperature heatmap.



(c) Minutes above evaluation temperature threshold.

Fig. 5.2 Figures characterising ambient temperature during the period of the experiment. In (a), the raw ambient temperature is shown in 15-minute resolution, along with the evaluation threshold mark at 15 °C, above which evaluations are suspended. In (b), the distribution of the temperature over each day is shown as a heatmap. In (c), the number of minutes per day during which evaluation was suspended due to high ambient temperature (at or above 15 °C) during the experiment is shown. Weekends are not included in (a) and (c).

output that were highly probable to be indicative of faults, using consistently high-penalty ratings as a marker. An example of such a pattern was that some windows never reported a closed state, and the records in the database seemed to correspond to the default state recorded by the edge IoT Gateway (Fig. A.2) on restart. In virtually all these cases, the evaluation penalty was consistently high, averaging over 30 MWh per day throughout the evaluation period.

After the initial set of deactivated rooms, another 40 rooms were additionally deactivated when it was discovered that there were fundamental problems with the sensor installations in two buildings. Some of the rooms in these buildings had already been deactivated in the initial phase, bringing the total to all 67 rooms in these two buildings. The deactivation was necessary since the team performance was negatively skewed by these faulty offices, and the occupants also cannot reliably make sense of the evaluations presented. Fig 5.3 shows the percentage of disabled rooms for each team. Teams T7 and T8 constituted the entirety of the two buildings that were fully disabled, so no evaluation data can be derived from these two teams. Team T4 was also significantly affected – half of the 20 rooms in the two buildings that make up Team T4 were disabled due to sensor faults.

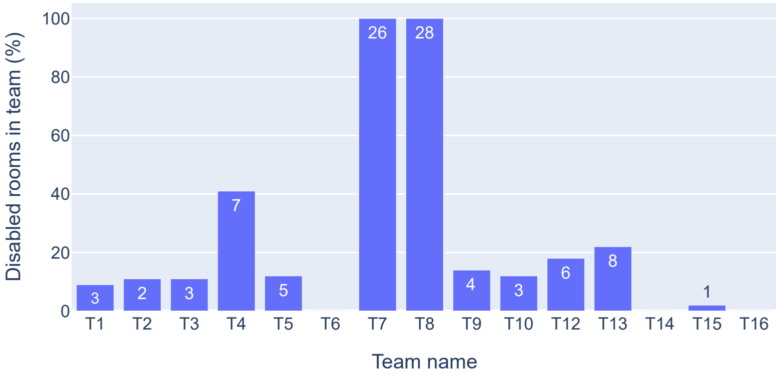


Fig. 5.3 Percentage of offices disabled per team due to faulty sensors. Raw values are placed on the bars. As can be seen, all offices in Buildings B-07 and B-06 were disabled due to an issue with the hardware configuration of the sensors.

Finally, although it was planned that the recommendation should be activated at the beginning of the experiment, it did not go online until Week 6 (on 18.04.2023), due to delays in the implementation and testing. Also, it was planned that recommendations are sent every day if required (but maximum of once per day per office to avoid spamming participants). However, the recommendation system only functioned for four days from April 18 to 20, and 24, due to a tiny bug in the scheduling logic that remained undiscovered until after the experiment.

5.1.5 Poor Heating in Building B-01 and B-02

One of the long-running issues encountered during the project (especially before the experiment) was the poor performance of the heating controller for the pilot building, Building B-01, where JuControl was part of the automated heating system. (The interaction between JuControl and the heating controller was already discussed in Section 4.3.1, "Coupling JuControl with Heating Controller".) Although the heating controller itself was not developed as part of this thesis, the negative effects affected the engagement of users – dissatisfied occupants in the building had dwindling motivation to use JuControl and its associated apps after the issue had persisted for months. This heating issue was also present in Building B-02, where the controller had also been deployed at a much later date. However, the problem did not persist for long, since by the time the automatic control was extended to Building B-02, measures were already being put in place to address the issue.

A primary cause of the poor performance had to do with the source of the temperature feedback for the smart radiator's hardware-integrated controller. The cloud controller send the setpoint temperature to the local hardware-integrated controller of the smart radiator, which in turn provides the actual control functionality. As at the time of the experiment, this hardware-integrated controller was configured to derive its closed-loop feedback temperature from the in-built radiator temperature sensor by default, instead of from the room temperature sensor. Due to local heating effects, the temperature at the radiator was often 1–2 °C higher than the actual room temperature (as measured by the room temperature sensor on the opposite side of the room), which caused the local hardware controller to stop heating before the room attained the actual desired temperature. Nevertheless, as at the time of writing, the hardware-integrated controller now correctly derives its feedback temperature from the room temperature sensor.

Another cause of the poor heating in Building B-01 was the relatively poor insulation in the building, which in some cases required several hours of pre-heating in order to meet the target temperature at the start of work. To compound this issue, the supply temperature of the working fluid was usually reduced at night by Facility Management and only raised again only a short while before the start of work, making it virtually impossible to ramp up fast enough to meet the setpoint temperature in many cases. Yet again, there were issues with low supply temperatures also during the day, which was beyond the control of the researchers. In most cases, these drops in supply temperature occur without notice, and were only identified via post-mortem analysis. Furthermore, there were occasional problems with individual radiators in the building, where the flow rate of the heating fluid was not high enough, even with fully opened valves.

Finally, because there was no adequate monitoring system to proactively (and even pre-emptively) detect these faults and investigate them early enough, there were multiple complaints from the occupants. In Building B-01 in particular, where the issue persisted longer, communications with the occupants revealed that they were not motivated to participate in the experiment, having been generally dissatisfied with the comfort situation in the offices in the prior months. In order to maintain fairness in the evaluation of setpoint temperature during the experiment, Juracle compensated for the heating shortfall by raising the *reference* setpoint temperature for unoccupied offices in the building, $T_{\mathrm{sp,ref,unocc}}$, from 17 °C to 19 °C. This compensation coincided with a corresponding increase of the setback temperature by the heating controller in Building B-01 during unoccupied hours, which was one of the measures taken to ameliorate the problem. A more complete discussion of these measures that were taken over time to mitigate these issues, including the implementation of a monitoring system, is presented in the next chapter (in Section 7.3.7).

In this chapter, the results of the behaviour interventions are presented, starting with an analysis of user engagement with the developed systems. Subsequently, the experiment hypotheses $\mathbf{H_1}$ to $\mathbf{H_3}$ introduced in Chapter 3 (Section 3.8.2), are tested on the basis of the results of the experiment. These results include derived performance measures of teams and buildings based on the behaviour evaluation methodology developed in the thesis.

5.2 User Engagement Results

This section discusses the level of engagement of users with the developed systems. In Fig 5.4, the breakdown of the interaction of employees of FZJ with the Energy Dashboard Suite as at the end of the experiment is depicted. This number includes those who have visited JuControl, since the same authentication system is used for the entire Energy Dashboard Suite. About 1,906 staff members, making up almost 30% of the nominal staff strength of FZJ, have interacted with the Energy Dashboard Suite at least once. Out of these, 43.3% (825 employees) have visited it more than once.

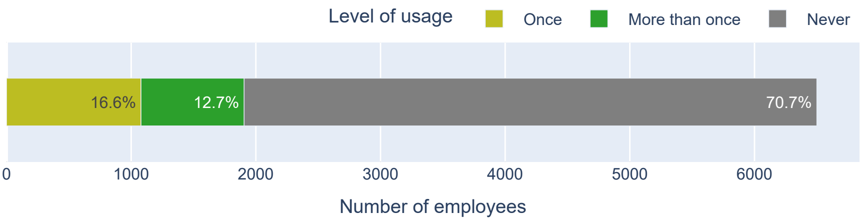


Fig. 5.4 Level of usage of the Energy Dashboard Suite by employees of FZJ.

For JuControl in particular, Fig. 5.5 shows a timeline of first interactions of users with JuControl, as well as the number of offices implicated in these interactions. This first interaction involves the confirmation of the officially assigned office of the Shibboleth-authenticated user, and the agreement to the data privacy notice. However, JuControl is *not* available for an office until *all* occupants of the office have agreed to the data privacy notice. Note that one multi-person office can appear across multiple days, since different occupants of the same office can interact for the first time with JuControl on different days. The graph shows that almost 140 occupants in 95 offices signed up for JuControl on the day of the initial announcement (March 16, 2023), with additional sign-ups average nine persons per day occurring over the next one week. On April 3 and 11, there are spikes in sign-ups in response to automated emails sent to all participants, in which the summary of the evaluation of the previous week was provided. Also, it can be seen that the days on which the recommendation emails were sent (April 18-20, and 24, 2023), saw additional sign-ups as occupants tried to understand the details of the recommendations. A link to JuControl was contained in all the emails.

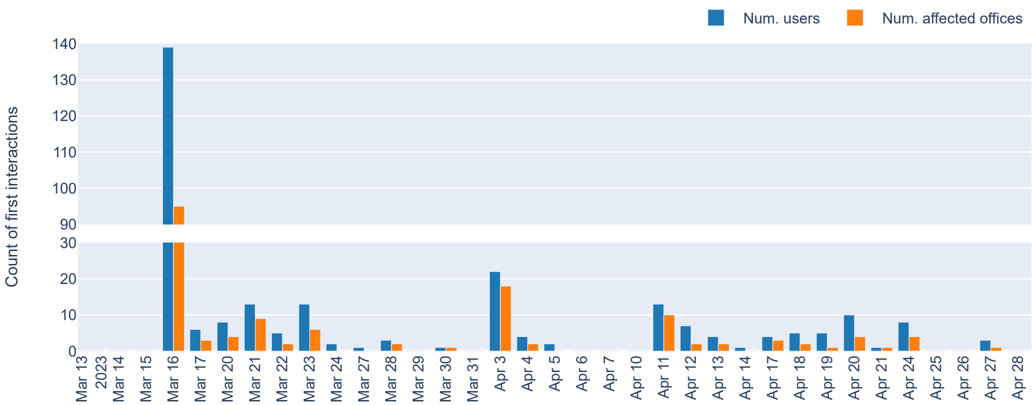


Fig. 5.5 Timeline of first interactions with JuControl over the experiment period. For each day, the number of users who interacted with JuControl for the first time ever is shown, along with the number of unique offices in which the corresponding users sit.

On the other hand, the timeline for the "activation" of JuControl in offices by occupants (i.e. by *all* occupants of each office consenting to the data agreement to enable JuControl for their office) is shown in Fig. 5.6. This activation is the only means by which occupants can access JuControl in order to view the status and performance of their room, or to control the heating in their office (depending on available

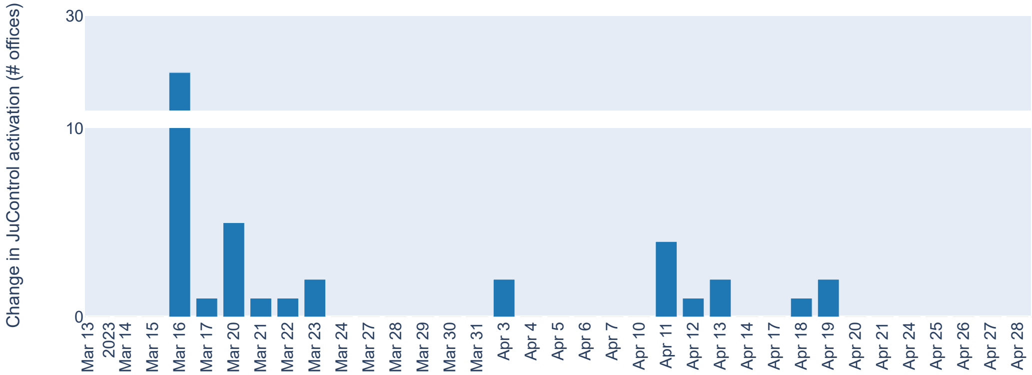


Fig. 5.6 Timeline of activation of offices in JuControl by occupants over the experiment period. The plot shows the change in number of offices activated in JuControl on each day of the experiment.

features). (The receipt of evaluation summaries and recommendations via email, however, does not depend on activation.) Note that it is possible that all occupants in a multi-person office sign up for JuControl. Again, the pattern of activation follows that of the first-interaction scenario discussed in the previous paragraph (see Fig. 5.5). Specifically, for the announcement of March 16, 2023, 49 offices were activated in JuControl on that day. In the one week following the announcement date, another 18 offices were activated. Subsequently, further spikes in activations took place on April 3 and 11, in response to the evaluation summary emails, similar to the first-interaction case already described.

By the end of the experiment on April 28, 2023, a total of 420 users (48.3%) of the approximately 870 potential users involved in the experiment had visited JuControl, including those who only started the consent process without finishing it. These 420 users comprise 283 office occupants who signed up during the experiment, and 137 users who already had access to JuControl in the preceding trial phases. In terms of JuControl "activation" of offices, out of the 494 offices involved in the experiment, 56 offices (about 11.3%) correspond to meeting rooms and offices without officially assigned employees; these are termed "unactivatable" offices (since they cannot be activated in the current setup without having officially-assigned occupants) and are generally excluded from the analysis. From the remaining 439 "activatable" offices, 82 offices (18.6%) were activated in JuControl by the end of the experiment, consisting of 51 offices (11.6%) that were activated during the experiment period itself, and 31 offices (7%) that were already activated during the trial phase prior to the experiment.

Here it should be remarked that due to the policy of all occupants actively granting consent for activation to occur, in most non-activated offices, only one occupant had failed to grant consent, and this was predominantly by omission rather than by actively declining. In fact, 70 offices involved in the experiment had at least one occupant who simply did not click the activation link emailed to them by the JuControl activation system. Out of these offices, 84% (n=59) had only *one* outstanding consent request that was not acted upon, which prevented the office from being activated. Meanwhile, among these 59 offices with one missing consent, only 28 are two-person offices, meaning that in remaining 31 offices, at least two other occupants had granted consent. Note that this analysis excludes offices with only one occupant (since this issue does not arise in the first place), offices where no occupant gave their consent, and offices where there was *consent denial*. Nevertheless, as at the time of writing, a new agreement has been reached with the Workers' Council, which allows offices to be available in JuControl without this consent process, on the condition that no private data is shown to the occupants, especially CO_2 concentration.

The breakdown of missing consents by team is given in Fig. 5.7, where the predominance of a single missing consent is evident across most teams. In terms of activation at the team level, Fig 5.8 shows JuControl activation by team, i.e. the number of activated offices (pre-experiment and during the experiment, depicted by the blue and green bars, respectively) and non-activated offices (red bar) for each team. As can be seen from the plot, Team T5 (comprised of part of Building B-05) had the highest number of activated offices (n=11) corresponding to 28% activation rate relative to the total activatable offices in the team. Most of these offices were activated pre-experiment. Team T7 had the highest relative activation rate (31%, n=8), followed by Teams T4 (n=4) and T15 (n=10) at 29% relative activation rate each. The teams with significant pre-experiment activation were either part of the testing phase (Teams T1 and T3), or had participants who participated in the various LLEC workshops, or else were privy to the existence of JuControl through word-of-mouth. In fact, according to the survey results (analysed in Section 7.2), word-of-mouth information from colleagues was the main means via which JuControl was publicised.

Fig. 5.7 Distribution of missing consents according to team for multi-person offices where the consent process had been started by at least one occupant.

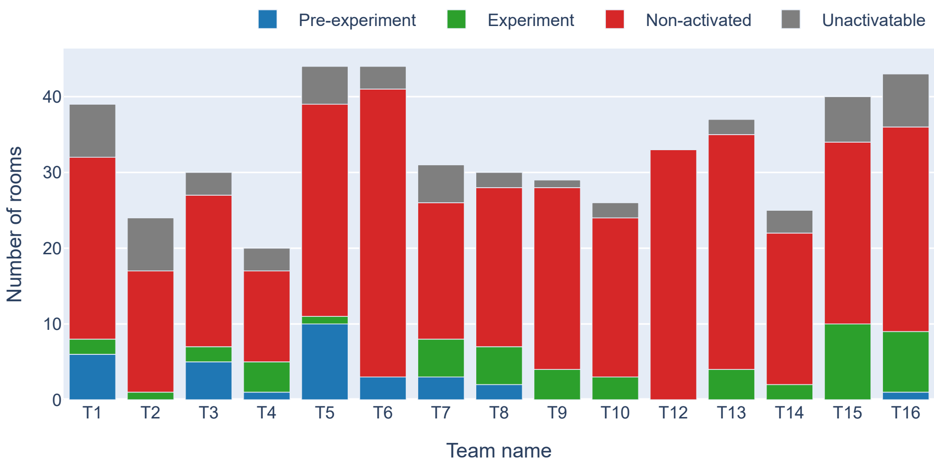


Fig. 5.8 JuControl activation by team as at the end of the experiment, showing number of offices activated pre-experiment (blue) and during the experiment (green), as well as non-activated offices (red). Unassigned offices (including meeting rooms), which by definition cannot be activated, are shown in grey.

In terms of relative activation rate as at the end of the experiment, Team T3 had the highest activation rate relative to "activatable" offices (63%, or n=17), while in terms of absolute numbers, Team T1 has the highest. In several cases there were some offices that were activated pre-experiment without actually being involved in the pre-experiment tests, as can be seen from Fig. 5.8. This was as a result of the occupants of these offices being aware of the existence of JuControl through formal or informal channels prior to the official announcement and start of the experiment, e.g. through participating in the LLEC workshops.

Team T6 had the lowest experiment-period activation rate (n=0) because there were no recommendations or evaluations, since no further prompts were provided to such teams to visit JuControl apart from the initial announcement. Another factor that influenced the activation rate was the kind of work and academic background of the building occupants: teams comprising administrative departments (e.g. Teams T2, T6, T12) tended to have low activation rates. On the other hand, Teams T15 and T16 had the highest experiment-period activations, although they had only *JuControl View* (Team T15) and access to a related serious game (Team T16). The high engagement was because, as anecdotal evidence demonstrated, the participants in these teams were more interested in JuControl due to the participants having a technical background related to ICT and software development, which enabled them to better appreciate JuControl. Team T14, which had only the *JuControl View* feature enabled belonged to a small building having mainly laboratories with few, hardly occupied offices, which explains the low activation. Again, note that within all these teams (including the 0% activated teams), interested occupants attempted to activate JuControl but were prevented by the inaction of their fellow office mates, since all occupants must grant consent before activation. As mentioned above, this issue affected 70 offices during the experiment.

5.2.1 Effects of Automated Emails on User Engagement

During the course of the experiment, two kinds of automated emails were sent to participants, starting first at the beginning of Week 5, as shown Fig. 5.1. First was the evaluation summary emails, which were sent at the beginning of each week summarizing the previous week's evaluation for the office and the team to which the office belongs, and also comparing it with the performance of the week preceding the summarized week. The first of these was sent on Monday of Week 4 (April 3, 2023, see Fig. 5.1), and then every week afterwards. The second kind of automated email was real-time recommendations about actions to perform to maintain a good rating, regarding either reduction in setpoint temperature, or closing / not tilting the windows. The first of these was sent on Tuesday of Week 6, i.e. on 18.04, and then on 19.04, 20.04, and 24.04 (due to a scheduling bug, the original plan of sending everyday afterwards until 28.08 did not work as expected, as discussed in 5.1.4).

Within the experiment design and research hypotheses framework, recommendations were not considered a major experiment variable given the limited sample size for the experiment, and for that reason no hypothesis was dedicated to it. Consequently, none of the teams differed from each other based on the (un)availability of recommendations, meaning that all the offices that had evaluation also had recommendation and vice versa. The implication of this is that, in theory, one cannot test the effect of recommendations on energy performance in isolation, based on the experiment design. However, the recommendations were activated after the experiment had run to about 70% completion (on Day 2 of Week 6, i.e. on 18.04, see Fig. 5.1), making it possible to analyse the effect of recommendations on energy rating along the time axis by considering the periods before and after activation. Additionally, it is possible to gain some insight into the response of individuals to the recommended actions by considering the time duration between the receipt of a recommendation email and the execution of the action recommended in the email. However, this insight can only be approximate since it is not possible to know when the user has read the email. These latter two effects of recommendations are handled in the next sections. The rest of this section, however, discusses the effect of these emails on user engagement.

As already mentioned in the previous paragraph, one of the (intended) "side effects" of these automated emails was the triggering of mass activation of JuControl in offices, since these emails did not consider room JuControl-activation status. The timeline of activation of offices in JuControl shows spikes of activity occurring at precisely the dates of the dissemination of the automated emails, as can be seen in Fig 5.6. Apart from the initial activations on the date of announcement, one-half of all further activations during the experiment period (n=12) occurred on, or immediately following, the dates when the automated emails were sent. Meanwhile, these email dates account for less than 18% of the duration of the experiment (excluding weekends and the day of initial announcement). Also, the first interaction plots (Fig. 5.5) show increased first-time activity on JuControl for these dates. This shows that the automated emails led to a marked increase in the rate of activations for the offices that received them.

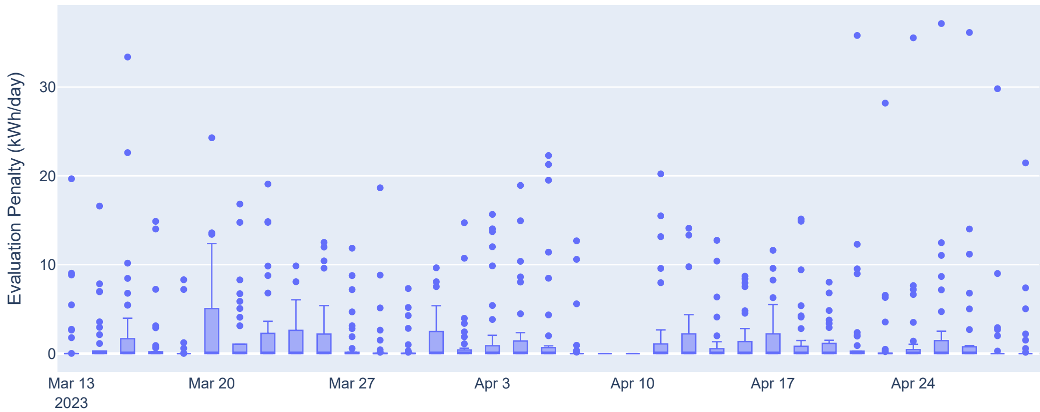
5.3 Results for Hypothesis H_1 (Effect of Evaluations / Recommendation)

Hypothesis H_1 posits that the energy performance of offices with evaluation / recommendation would be better than those of offices without these. To test this hypothesis, we employ the H_1-Test-1 and H_1-Test-2 tests defined in Table 3.9 in the previous chapter. Since the daily *energy penalties* in the teams follow an approximately normal distribution as determined by the Shapiro-Wilk Test (e.g. for Team T5: W=0.977, p=.68; for Team T6: W=0.986, p=.93), and have unequal variances (Team T5: σ^2=0.468; Team T6: σ^2=2.892), we choose Welch's T-test to determine the statistical significance of the performance difference between the two teams, and then we use Cohen's d to estimate the effect sizes.

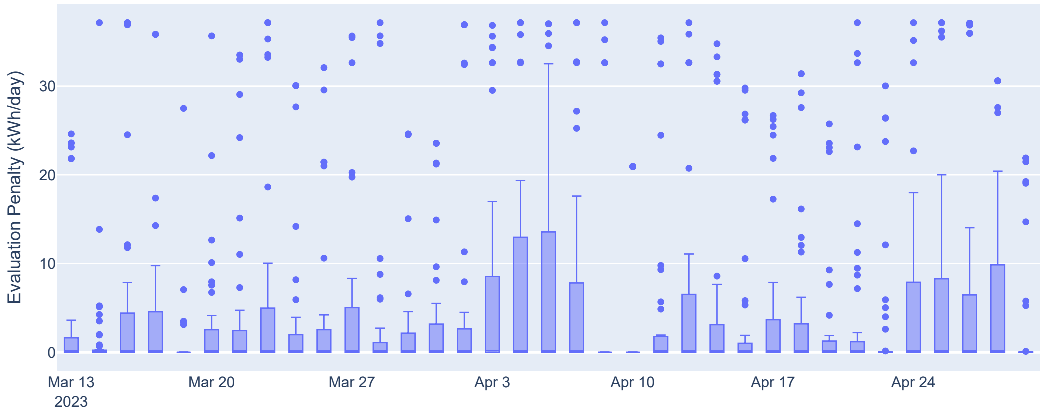
5.3.1 H_1-Test-1: Comparison of Team T5 and Team T6

Teams T5 and T6 were drawn from the four floors of the same building (Building B-05), where Team T5 comprises the lower two floors, and Team T6 the upper two floors. Whilst Team T5 had *JuControl view* with **ventilation evaluation** and **recommendation** enabled, Team T6 only had *JuControl view* (see Table 3.7). In Fig. 5.9, the distribution of energy penalties for Team T5 (Fig 5.9a) and Team T6 (Fig 5.9b) are shown in a boxplot. It can be reasonably assumed that both teams were exposed to similar environmental disturbances and similar office conditions by being in the same building. Mere visual inspection of the boxplots indicates a difference in the distribution of penalties between the two teams, with Team T5 appearing to have generally lower penalties than T6. Applying the afore-mentioned statistical tests of significance and effect size on the results, the average daily *energy penalty* is significantly lower in Team T5 (μ = 1.659 kWh) than in Team T6 (μ = 4.667 kWh) (p<.001) with a very large effect size (d=2.27), showing that offices in Team T5 were more energy efficient than those in Team T6 according to the evaluation methodology established in this work.

To investigate the factors underpinning the difference in performance between Team T5 and Team T6, composite plots depicting JuControl activation, average team penalty, and number of offices in each activation category, are shown analogously in Fig. 5.10 and Fig. 5.11 for Team T5 and Team T6, respectively. Each of these figures comprises three sub-figures categorised based on JuControl activation for the respective team: (a) JuControl activated offices, (b) non–JuControl-activated offices, and (c) all offices. In each activation category, two complementary charts are shown: an evaluation penalty plot (line chart, top), and an office-count plot (bar chart, bottom). The evaluation penalty plot (top) depicts the trend in average team penalty



(a) Team T5 penalty distribution



(b) Team T6 penalty distribution

Fig. 5.9 Box plots of the distribution of the energy penalties for offices in (a) Team T5 and (b) Team T6 showing a stronger trend towards energy efficiency improvement (lower penalties) over the course of the experiment for Team T5 than for Team T6.

over the experiment period, while indicating the traffic-light rating of the penalties (see Section 4.5.9 for details of the traffic light rating system). The office-count plot (bar chart, bottom) shows the number of offices in the JuControl-activation category, with each bar consisting of two vertically stacked bars representing, respectively, the proportion of occupied offices (black), and that of unoccupied offices (grey) on the given date. The presence duration threshold for an office to be considered as occupied in this analysis is three hours of occupancy. Presence was estimated in this building via the CO_2 concentration mass-balance approach described in Section 4.6.2.

From Fig. 5.10 and Fig. 5.11, several deductions can be made. First, the penalties during the Easter holiday period (Apr. 7 and 10, i.e. Good Friday and Easter Monday, respectively) were practically zero for Team T5 and for JuControl-activated offices in Team T6. This confirms the expected baseline that unoccupied offices should have zero penalties for ventilation-evaluated offices, since there are no window interactions. However, in non–JuControl-activated offices in Team T6, even when there was seemingly no occupied office according to Fig. 5.11(c) (on Apr. 10), the penalty was relatively high (red zone), and analysis of the data indicated that at least one window appeared to be left tilted in some offices after occupants had gone for the holidays.

Fig. 5.10 Composite plots for Team T5 showing evaluation penalty (line chart, top) and number of offices (bar chart, bottom), categorized by JuControl activation for offices in the team: (a) JuControl-activated offices, (b) non–JuControl-activated offices, and (c) all offices. Each evaluation penalty plot (top) is divided into horizontal strips corresponding to the traffic light rating of the penalty values, and important dates are marked by vertical lines. Each office-count plot (bottom) shows the number of offices in the respective JuControl-activation category as a composite bar chart, where each bar depicts the proportion of occupied/unoccupied offices for each day of the experiment period.

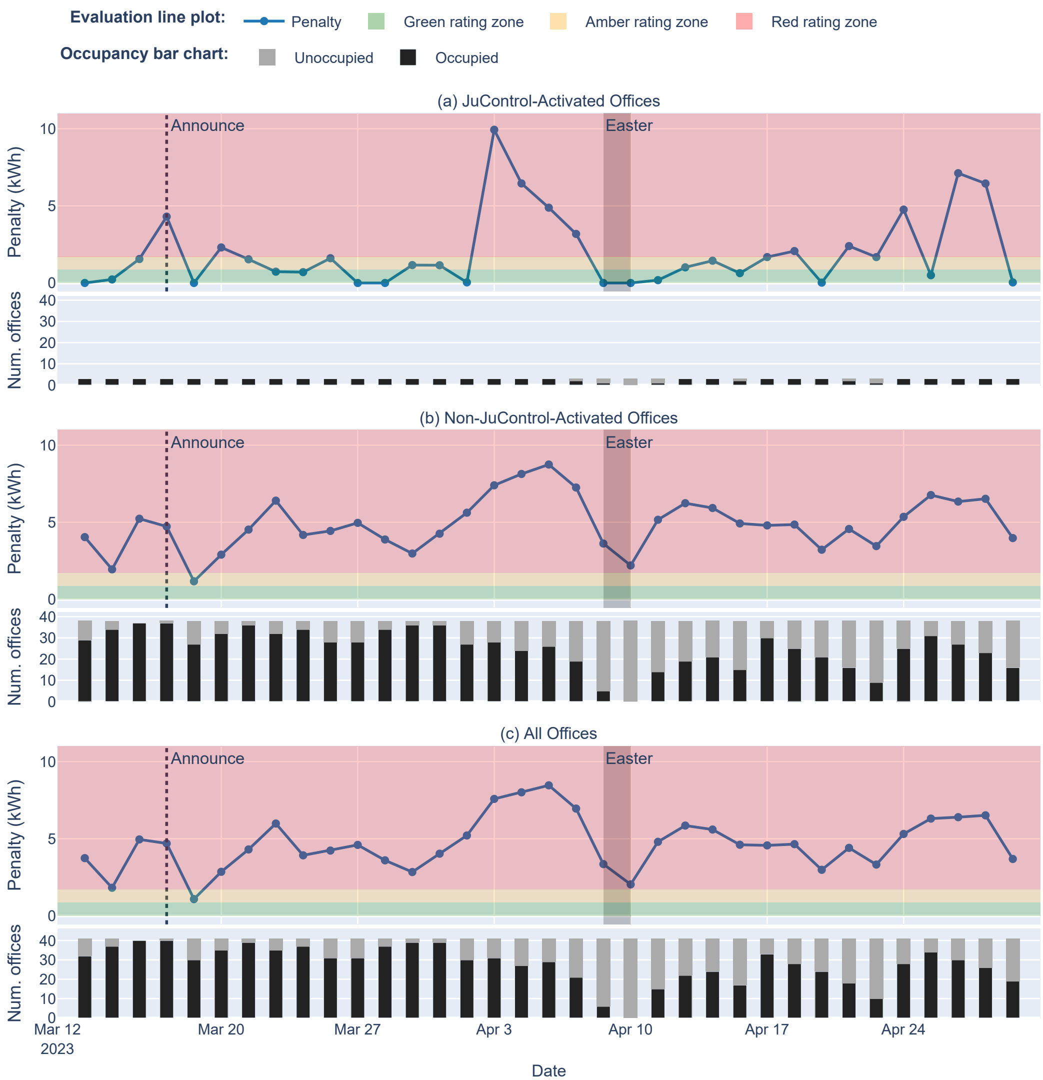


Fig. 5.11 Composite plots for Team T6 showing evaluation penalty (line chart, top) and number of offices (bar chart, bottom), categorized by JuControl activation for offices in the team: (a) JuControl-activated offices, (b) non–JuControl-activated offices, and (c) all offices. Each evaluation penalty plot (top) is divided into horizontal strips corresponding to the traffic light rating of the penalty values, and important dates are marked by vertical lines. Each office-count plot (bottom) shows the number of offices in the respective JuControl-activation category as a composite bar chart, where each bar depicts the proportion of occupied/unoccupied offices for each day of the experiment period.

Furthermore, due to the "activation policy" in JuControl, whereby office occupants are only granted access to JuControl and its features after *all* occupants have agreed to the data consent form, only a subset of offices in each team was "activated". As can be seen from the figures, for Team T5 (Fig. 5.10), there were relatively more offices that were JuControl-activated than in Team T6 (Fig. 5.11), and these JuControl-activated offices had lower penalties than non-activated offices, whilst trying to stay within the "green zone". The initial low energy penalties in Team T6 is an artifact of the nature of the experiment – offices became JuControl-activated throughout the experiment period, and for Team T6, the number of activated offices is too low to be statistically meaningful, and the low penalties can be regarded as coincidental, as the spikes on Apr. 3 and 25 show. Considering Team T5 and analyzing the difference between activated offices (n=9) and non-activated offices (n=24), a significant difference is observed in the average energy penalty with a very large effect size, where the penalty is much lower in activated offices (μ = 0.739 kWh) than in non-activated offices (μ = 1.995 kWh) (p<.001, d=1.4).

In conclusion, we can reject the null hypothesis and conclude that having access to recommendations and evaluations, as well as to the relevant contextual performance information in JuControl, led to more energy-efficient behaviour among occupants, than in the control group without these features. Also, the willingness to use the provided behaviour intervention tools was a key factor in causing positive behaviour change. In the following sub-section, further discussion is provided regarding the particular occupant behaviours that influenced the performance of Teams T5 and T6.

Analysis of Ventilation Behaviour for Teams T5 and T6

The relative contributions of window ventilation style (trickle vs. shock) and duration on the penalties of Team T5 and Team T6 are depicted in Fig. 5.12 and Fig. 5.13, respectively. Note that only the raw ventilation durations are given, without the trickle ventilation penalty. From the figures, it can be seen that the use of trickle ventilation is the main factor contributing to penalties, as the daily duration of trickle ventilation is almost always higher than for shock ventilation by a wide margin. On average considering the entire experiment period, in Team T5, trickle ventilation is used almost 5 times as long as shock ventilation across all offices (38.8 min vs. 8 min, respectively) (see Table 5.1). For JuControl-activated offices, trickle ventilation is used only about 70% longer on average than shock ventilation (15.4 min vs. 9.1 min, respectively), while for non–JuControl-activated offices trickle ventilation is used more than 6 times as long as shock ventilation (47.5 min vs. 7.6 min, respectively). Furthermore, it can be observed from Fig. 5.12(a) that towards the end of the experiment period, JuControl-activated offices tended to favour shock ventilation over trickle ventilation, consistent with the nudges from the behavioural intervention measures.

Table 5.1 Comparison of average daily duration of shock vs. trickle ventilation in Teams T5 and T6 for the entire experiment period, showing that trickle ventilation was the main driver for high penalties in both teams, although much more pronounced in Team T6 than in T5.

Team	Shock vent.[a] (minutes)			Trickle vent.[b] (minutes)			Ratio[c] (Trickle:Shock)		
	All[‡]	Act.[‡]	N/Act[‡]	All	Act.	N/Act	All	Act.	N/Act
T5	8.0	9.1	7.6	38.8	15.4	47.5	4.8	1.7	6.2
T6	5.4	14.0	4.7	149.5	39.2	158.2	27.7	2.8	33.5

[a] Mean shock ventilation duration. [b] Mean trickle ventilation duration.
[c] Ratio of average trickle to average shock ventilation duration.
[‡] **All**: All offices; **Act.**: JuControl-activated offices; **N/Act**: non–JuControl-activated offices

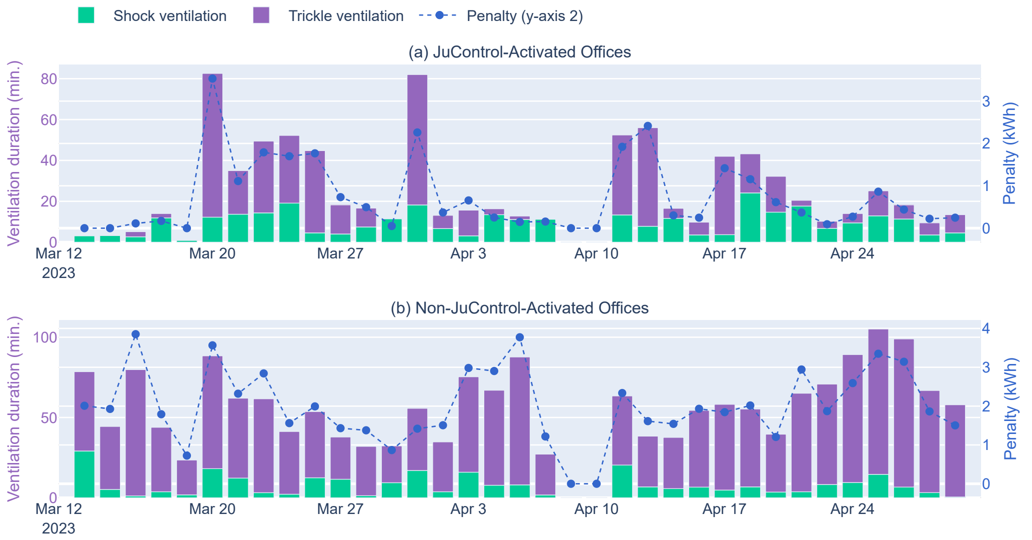


Fig. 5.12 Factors contributing to penalty of **Team T5** for (a) JuControl-activated and (b) non–JuControl-activated offices. The bar charts show the average *equivalent* ventilation duration (primary y-axis). The line chart (secondary y-axis) shows the average penalty. Recall that trickle ventilation penalty = trickle ventilation duration since $f_{\text{pen,trickle}} = 2$.

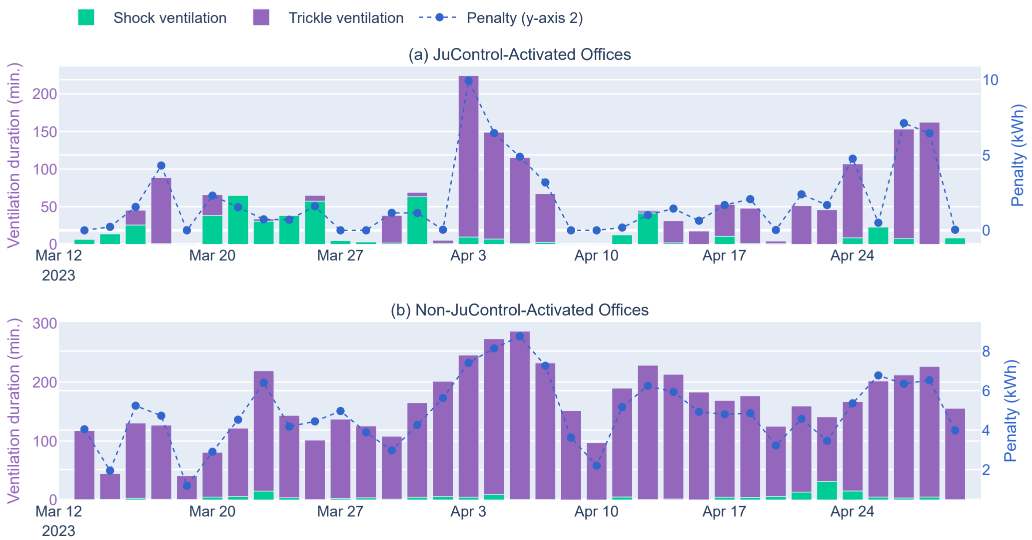


Fig. 5.13 Factors contributing to penalty of **Team T6** for (a) JuControl-activated and (b) non–JuControl-activated offices. The bar charts show the average *equivalent* ventilation duration (primary y-axis). The line chart (secondary y-axis) shows the average penalty. Recall that trickle ventilation penalty = trickle ventilation duration since $f_{\text{pen,trickle}} = 2$.

On the other hand, Team T6 tells a more extreme story: across all offices, trickle ventilation is used more than 27 times as long as shock ventilation on average (149.5 min trickle vs. 5.4 min shock), with non–JuControl-activated offices being the major contributors to trickle ventilation, which is used almost 36 times as long as shock ventilation on average (see Table 5.1). Fig. 5.13 shows that on several days throughout the experiment period, offices in Team T6 used trickle ventilation exclusively and for relatively long periods. Comparing Teams T5 and T6, the average trickle ventilation usage in Team T6 is almost 4 times that of Team T5 for all offices, while average shock ventilation usage in Team T5 is slightly higher than for Team T6 (8 min vs. 5.4 min), although the values are low in absolute terms. This higher use of trickle ventilation in Team T6 in place of shock ventilation contributed to its high penalties.

5.4 Results for Hypothesis H_2 (Effect of Active Participation)

Hypothesis H_2 posits that the energy performance of JuControl-activated offices would be better than that of non–JuControl-activated offices among teams that have evaluation / recommendation enabled. To test this hypothesis, we employ the **H_2-Test-1** test defined in Table 3.9 in the previous chapter.

5.4.1 H_2-Test-1: JuControl-activated vs. non–JuControl-activated Offices

This test compares the performance of JuControl-activated offices to non–JuControl-activated offices within buildings in which evaluation / recommendation is enabled, where the working assumption is that the activation of JuControl by occupants through unanimous data consent affords such occupants the opportunity to fully exploit the system in order to be more energy efficient. Note again that in general, all evaluation- / recommendation-enabled teams received the recommendation and evaluation summary emails across both activated and non-activated offices. This working assumption is all the more justified since anecdotal evidence as well as user survey feedback and data analysis show that emails were only limitedly effective in triggering behaviour change. In fact, some users redirected the evaluation summary and recommendation emails to the spam folder or outrightly blocked the sending email address, meaning that the receipt of these recommendations did not lead to corresponding action in some cases. The reasons for the varied user responses are explored in the analysis of the survey results in Section 7.2.1.

For teams with *JuControl Control* enabled, i.e. Teams T1 to T3, where the automatic heating controller manages the room heating in conjunction with JuControl presence schedules, the difference between JuControl-activated and non-activated offices is expected to reflect directly in the ratings, since non–JuControl-activated offices do not have automatic setpoint temperature regulation. Some evaluation-enabled teams were excluded from this analysis for the following reasons. Teams T2, T12, T13, and T14 are excluded due to low number of JuControl-activated offices, where only three or fewer JuControl-activated offices were successfully evaluated. Also excluded are Teams T7 and T8 for which no analysis is possible since the entire building, Building B-07, was disabled for evaluations due to faulty sensors in the building, as mentioned previously. Hence, only Teams T1, T3, T4, T5, and T9 are analyzed here. Additionally, since Teams T2 and T3 are roughly equal in size and drawn from the same building, the entire T2 (non–JuControl-activated) is also compared with JuControl-activated offices of Team T3.

The results of the comparison between mean daily energy penalties of the above-mentioned "activation groups" is shown in Table 5.2, including the statistical significance and effect sizes. From the results, JuControl-activated offices perform significantly better than non-activated offices in Teams T1, T3, and T5, with medium effect size in T1 and large effect sizes in the other two teams. The comparatively lower effect size in Team

T1 was partly caused by a brief increase in the *reference setpoint temperature* during the experiment to compensate for heating issues in the building, as explained in Section 5.1.5, which diminished the advantage of the JuControl-activated offices over non-activated offices.

Furthermore, comparing JuControl-activated offices in Team T3 with offices of T2 (which are all non-activated) reveals the same significantly better performance of the former, with a large effect size. In fact, a comparison of Team T2 with *non-activated offices* in Team T3 shows no significant difference in the means (p=.201) and the effect size is small ($d = 0.3$). In other words, the non-activated offices in the entire building, irrespective of team / grouping, had similar performance that was significantly worse than that of activated offices. These results agree with hypothesis H_2, thus strengthening the working assumption that actual *and* full use of JuControl contributed to higher energy efficiencies than merely receiving evaluations and occasional recommendations by email.

Table 5.2 Mean daily energy penalties and corresponding comparative statistics for JuControl-activated vs. non–JuControl-activated offices in select teams. The smaller values (more energy-efficient groups) are bolded.

Team	**Mean daily penalty (kWh)**		**p-value**	**Effect size ($\|d\|$)**
	JuControl-activated	Non–JuControl-activated		
T1	**1.61**	2.49	.029	Medium (0.5)
T3	**1.91**	3.82	<.001	Very large (1.4)
T3/T2*	**1.91**	4.36	<.001	Very large (1.5)
T4	6.09	**3.64**	.002	Medium (0.79)
T5	**0.74**	2.0	<.001	Very large (1.4)
T9	1.24	**1.06**	>0.05	Small (0.15)

* T3-JuControl-activated vs. T2 (all offices are non–JuControl-activated).

Team T4 presents an interesting case, in which the opposite trend appears to hold, where JuControl-activated offices perform even worse than non-activated offices. In particular, the mean daily energy penalty in activated offices is significantly higher than that for non-activated offices. In fact, Team T4 is the worst-performing of all the evaluated teams in the experiment, when its average penalties are compared to those of other teams. The reason for this poor performance is that the bulk of this team is made up of offices from Building B-03, in which the occupants reported that there was no possibility for them to control the radiators in several offices, so that the occupants left the windows in the tilted position for long periods to control overheating. Note that the heating in that building is not managed by JuControl with the automatic heating controller; for this project, only a "read" connection was established to the building automation system in order to include setpoint temperature evaluation for that building. The fact that JuControl-activated offices perform worse than non-activated offices is likely because coincidentally the worst affected offices happened to be also activated, as indeed the initial report about the faulty heating control came from a JuControl-activated office where the occupant could see their performance in JuControl.

For Team T9, the performance of non–JuControl-activated offices seems to be statistically similar to that of activated offices. However, a deeper analysis of the results reveals the reason for this: during the first half of the experiment period, only three or less offices out of the 21 evaluated offices in the team were activated, and the performance of one out of these three activated offices was poor throughout the experiment, thereby skewing the average penalty for activated offices due to the low number of offices in that sub-category (see Fig. C.13 in Appendix C). The average penalty for activated offices reduced as more offices became activated, thereby counteracting the effect of the poorly-performing office; however, only 7 offices were activated in total as at the end of the experiment. Thus, beyond the observed mean penalties in the team, the time-dependent nature of JuControl-activation also affected the results.

Binary classification of offices by JuControl-activation and energy penalties

To provide additional insight into the performance difference between activated and non-activated teams, a binary classifier based on logistic regression is trained on the daily penalties for each included team, where the penalties are averaged over all JuControl-activated offices on the one hand (Class 0), and non–JuControl-activated offices on the other hand (Class 1). This classification is also done for all the included teams combined. The purpose is to estimate the likelihood of accurately classifying an office in a given team as JuControl-activated or non–JuControl-activated based on its energy penalty, given the energy penalties of other offices in the team. Two goodness-of-fit metrics are stated in the regression plots. The first metric is the *accuracy score* from *scikit-learn* [157], defined as

$$\text{accuracy}(y, \hat{y}) = \frac{1}{n_{\text{samples}}} \sum_{i=0}^{n_{\text{samples}}-1} 1(\hat{y}_i = y_i) \tag{5.1}$$

where y and $\hat{y}_i$ are the actual and predicted class of item i, respectively, and $1(\cdot)$ is the indicator function which returns 1 if $\hat{y}_i = y_i$, otherwise 0. Note that this accuracy metric only measures how well the regression line fits the given training / prediction data, and does not say anything about whether the data confirms the hypothesis that JuControl-activated offices have less penalties on average than non–JuControl-activated offices. The other metric is the *Area Under Receiver Operating Characteristic* curve (AUROC), which ranges from 0.5 (worst classifier) to 1 (best classifier); values below 0.5 mean that "inverting" the binary classifier would produce better results. The Receiver Operating Characteristic (ROC) curve is a plot of the *True Positive Rate* (TPR) of the predications against the *False Positive Rate* (FPR).

In Fig. 5.14, the logistic regression curves are shown for each analysed team. As can be seen from the figure, Team T3, which is comprised partly of the author's research institute, shows the best fit for the classification with 80% accuracy and AUROC=0.91, demonstrating a marked difference in performance between JuControl-activated and non–JuControl-activated offices. This is followed by Team T5, which has already been analysed in a previous section. Furthermore, Team T1 which makes up Building B-01, shows some distinction between activated and non-activated offices, but only to a small degree. Since Building B-01 was part of the pilot phase, detailed analysis for Team T1 is presented separately in Section 5.5 as part of tests for **Hypothesis H_3**.

To further analyse the performance of Team T4, Table 5.3 shows a comparison of Team T4 (Building B-03) with Teams T2 and T3 (Building B-02) in terms of the contributing factors to the penalties, since setpoint temperature evaluation is available in all three teams. From the table, the average setpoint deviation over the experiment period for all offices in Team T4 is 22% and 53% higher than that of Teams T2 and T3 respectively. For JuControl-activated offices, the difference is more pronounced: more than three times as much setpoint temperature deviation as in Team T3. (Team T2 had only one JuControl-activated office so is not shown.) The same pattern holds for ventilation duration, where it can be seen that in Team T4, the windows are kept tilted several times longer than fully open – one order of magnitude on average for all offices (27.8 min vs. 2.8 min). Compare this with the more even mix of tilt and fully open ventilation styles in Teams T2 and T3. The plots showing the daily breakdown of the penalties for these three teams in terms of ventilation duration and setpoint temperature deviation are not shown in this chapter but rather in Appendix C due to space constraints (Figs. C.4, C.6, and C.8).

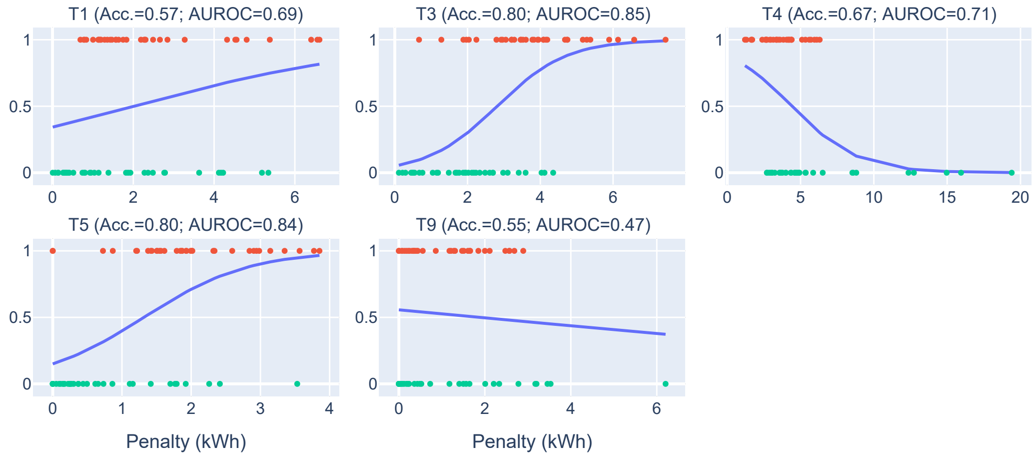


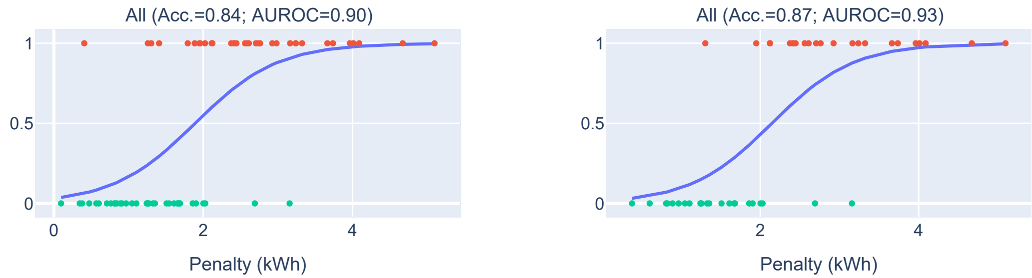
Fig. 5.14 Logistic regression curves based on average daily penalty for evaluation-enabled teams included in the analysis. Offices are categorized as JuControl-activated ($y = 0$, green) vs. non–JuControl-activated ($y = 1$, red). (Acc. = Accuracy; AUROC = Area Under Receiver Operating Characteristic curve.)

Table 5.3 Comparison of the mean deviation of setpoint temperature from ideal for Teams T2, T3 and T4, as well as the mean daily shock and trickle ventilation duration in the teams for the period of the experiment. The offices in each team are additionally categorized into JuControl-activated and non–JuControl-activated offices. The ratio of trickle ventilation to shock ventilation duration is also given, showing the predominant style of ventilation for the teams and categories.

Team	**Setpoint deviation** (degree-minutes)			**Shock vent.**[a] (minutes)			**Trickle vent.**[b] (minutes)			**Ratio**[c] **(Trickle:Shock)**		
	All‡	Act.‡	N/Act‡	All	Act.	N/Act	All	Act.	N/Act	All	Act.	N/Act
T2	3481.7	-	3481.7	48.0	-	48.0	16.2	-	16.2	0.3	-	0.3
T3	2770.0	1598.7	3109.7	17.3	14.3	18.2	22.8	9.4	26.7	1.3	0.7	1.5
T4	4248.1	5384.0	3814.8	2.8	1.0	3.9	27.8	49.0	19.7	9.7	49.0	5.1

[a] Mean shock ventilation duration. [b] Mean trickle ventilation duration.
[c] Ratio of average trickle to average shock ventilation duration.
‡ **All**: All offices; **Act.**: JuControl-activated offices; **N/Act**: non–JuControl-activated offices

Considering the three no-issue teams as discussed above, namely Teams T1, T3, and T5, their combined logistic regression plot is shown in Fig. 5.15a (accuracy=84%, AUROC=0.9). For these teams, it can be concluded that indeed, occupants of JuControl-activated offices were more energy-efficient than those of non–JuControl-activated offices in general. Furthermore, to investigate the sensitivity of the performance difference between JuControl-activated and non–JuControl-activated offices in general to ambient temperature, the days of the experiment in which evaluations were suspended for four hours or more due to high ambient temperatures are excluded in the combined logistic regression for Teams T1, T3, and T5 (see Fig. 5.15b). This test was carried out in order to compensate for the weakness of the developed evaluation system whereby penalties were still computed for ventilation during warm days (although penalties were suspended at those *instances* when the ambient temperature was above 15 °C; see Section 5.1.3 for the discussion). A total of 12 days out of the 35 experiment days were affected. As can be seen, the relative performance of JuControl-activated offices increased slightly (accuracy=87%, AUROC=0.93). This shows that there was some conflict within occupants between being comfortable on a warm day and having good ratings according to the evaluation system, and that in many cases, occupants chose comfort, particularly on the warmest days of the period.



(a) T1, T3, T5 combined **(b)** Excluding days exceeding 4-hour threshold

Fig. 5.15 Logistic regression curves for (a) Teams T1, T3, and T5 combined, and (b) same teams combined but with the exclusion of days having above 4 hours of evaluation suspension. Offices are categorized into JuControl-activated ($y = 0$, green) and non–JuControl-activated ($y = 1$, red). (Acc. = Accuracy; AUROC = Area Under Receiver Operating Characteristic curve.)

5.5 Results for Hypothesis H_3: Effect of Automatic Heating in Building B-01

For the pilot building, Building B-01, where the occupancy scheduled-based automatic heating controller was deployed along with the possibility of setpoint temperature control via JuControl, an analysis of the savings due to this intervention is presented in this section. The automatic heating controller was designed to work only for JuControl-activated offices, since the occupants of such offices have granted access to JuControl and therefore have *implicitly* activated their presence schedules. Nevertheless, in non–JuControl-activated offices in the building, smart radiator valves are also installed, but can only be physically controlled by the occupants with no automation or JuControl involvement. (As at the time writing, a default schedule is now also used for heating non-activated offices, although the upgrade post-dates this analysis.)

As mentioned previously in Section 5.1.5, participants in Team T1, who are the occupants of Building B-01, did not participate in the experiment as much as expected due to dissatisfaction caused by the initial issues with the heating controller that had persisted up to the experiment period (see Section 5.1.5). Nevertheless, the application of the automatic control to space heating in the JuControl-activated offices had positive effects on the energy performance of the building. In any case, the explanatory power of the setpoint deviation analysis in the next section is *not* diminished by these heating issues, since the deviation does not depend on the actual temperature in the room (which in the case of under-heating is lower than the setpoint temperature), but on the setpoint temperature itself, and as will be seen in the analysis, some measures were put in place during the experiment to account for possibly higher setpoint temperatures that occupants could have set in order to compensate for the poor heating. Furthermore, in the calculation of potential savings in Building B-01 due to the use of the automatic controller (Section 5.5.2), the poor heating in the building is compensated for to remove the false advantage that the under-heating would have conferred on the savings.

5.5.1 H_3-Test-1: Analysis of Setpoint Temperature Deviation in Building B-01

The *deviations* of the setpoint temperature from the ideal case for Building B-01 (Team T1), along with the average ventilation durations, are depicted in Fig. 5.16 for JuControl- and non–JuControl-activated offices, showing the relative contributions of these factors to penalties accrued by the team. In each sub-figure, the bar charts, which have values on the primary y-axis, represent the setpoint deviation-from-ideal (orange bar) and ventilation durations (stacked green/purple bar) averaged across all offices in the respective JuControl activation category.

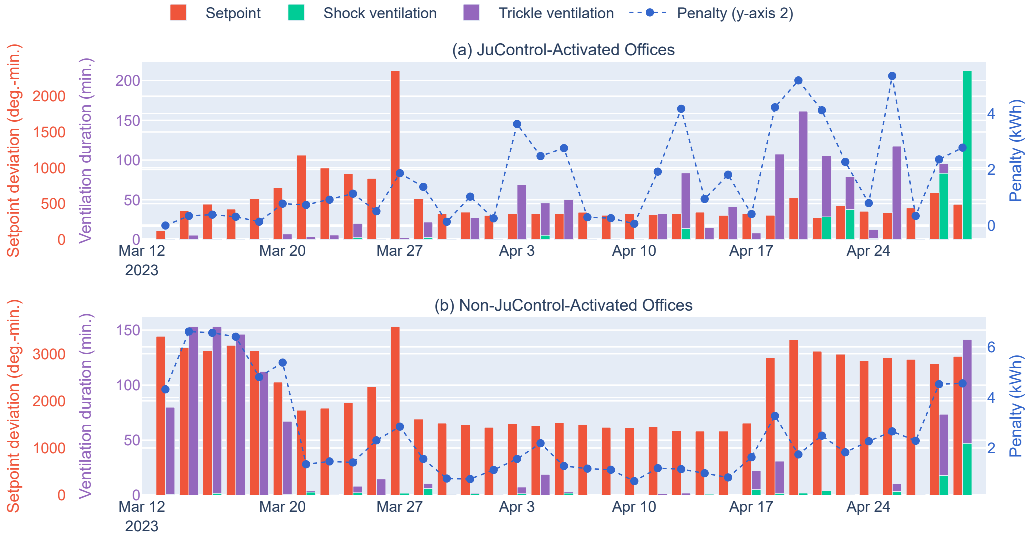


Fig. 5.16 Factors contributing to penalty of **Team T1** for (a) JuControl-activated and (b) non–JuControl-activated offices. The bar charts show the average setpoint deviation and average *equivalent* ventilation duration (primary y-axis). The line chart (secondary y-axis) shows the average penalty for reference purposes.

First, the effect of the compensation in evaluations to account for poor heating in the building by raising the reference setpoint temperatures, as discussed in Section 5.1.5, can be seen between Mar. 28 to Apr. 18 in Fig. 5.16, where the red bars (setpoint temperature *deviation*) are relatively low throughout the period. A difference can be seen, nevertheless, between the setpoint deviation of the activated vs. non-activated offices, where the activated offices averaged 50% or less of the deviation of non-activated offices. According to Table 5.4, which shows the breakdown of setpoint deviation and ventilation durations for the experiment period, on average, the JuControl-activated offices had less than one-quarter of the deviations on non-activated offices (528.9 deg.–min. vs. 2246.2 deg.–min.). More robust statistical analysis shows a significant difference in the daily mean of setpoint deviations across JuControl-activated offices vs. non-activated offices (p<.001) with a very large effect size (d=2.8). This demonstrates the energy-saving effect of the heating controller that managed the heating in the activated offices through the presence schedules of the occupants as capture in JuControl.

Table 5.4 Mean deviation of setpoint temperature from ideal for **Team T1**, as well as the mean daily shock and trickle ventilation duration for the experiment period. The offices are categorized into JuControl-activated and non–JuControl-activated offices. The ratio of trickle ventilation to shock ventilation duration is also given, showing the predominant style of ventilation for Team T1.

Setpoint deviation (degree-minutes)			**Shock vent.**[a] (minutes)			**Trickle vent.**[b] (minutes)			**Ratio**[c] **(Trickle:Shock)**		
All‡	Act.‡	N/Act‡	All	Act.	N/Act	All	Act.	N/Act	All	Act.	N/Act
1971.7	534.3	2240.4	4.5	12.4	2.9	28.7	29.9	28.5	6.4	2.4	9.7

[a] Mean shock ventilation duration. [b] Mean trickle ventilation duration.
[c] Ratio of average trickle to average shock ventilation duration.
‡ **All**: All offices; **Act.**: JuControl-activated offices; **N/Act**: non–JuControl-activated offices

Furthermore, it can be seen that the penalties for JuControl-activated offices were low at the start of the experiment but increased later, while the opposite effect was true for non-activated offices, where the penalties reduced after the experiment started. This was caused mainly by a highly non-performant office that became JuControl-activated and hence contributed to worse penalties in the activated offices. Again, the long periods of ventilation at the end of the experiment in both activated and non-activated groups was discovered to be caused by offices in the buidling which have south-east facing windows. Since those days were warm, a likely explanation, therefore, is that the windows were kept open to counter the high heat gains due to direct solar radiation, coupled with the warm ambient temperatures on those days that resulted in suspension of evaluation for over 4 hours and almost 10 hours on these respective days, as seen from Fig. 5.2. An immediately apparent improvement to the evaluation system is that the state of the heating system should be considered along with ventilation, to know when energy is actually being consumed during the ventilation period. In general, however, trickle ventilation was more predominant in the building.

For ventilation practices in Team T1, from Table 5.4, it can be seen that trickle ventilation was used for similar durations on average between JuControl-activated and non-activated offices, averaging almost 30 minutes per day during the experiment period. However, shock ventilation was used for longer periods in JuControl-activated offices than in non-activated offices, which is then responsible for the similar or sometimes slightly worse penalties of activated offices, despite the latter having better setpoint deviation performance. In general, however, trickle ventilation was used more than six times as long as shock ventilation in the team, with the ratio being higher in non-activated offices due to their lower shock ventilation use. No performance advantage is seen therefore for JuControl-activated offices in terms of ventilation, which was already explained by the non-participation of occupants of the building in the experiment.

In conclusion, the performance improvement in Building B-01 demonstrates that the deployment of the heating controller led to improved setpoint temperature performance, confirming hypothesis $\mathbf{H_3}$ that the deviations of setpoint temperature from ideal would be reduced by the introduction of the JuControl-calendar–managed heating schedule. In the next section, the second test for hypothesis $\mathbf{H_3}$ is carried out to determine energy savings at the building level.

5.5.2 $\mathbf{H_3}$-Test-2: Historical Performance Comparison of Building B-01

The second test for hypothesis H_3, tagged **$\mathbf{H_3}$-Test-2**, compares the annual energy demand of Building B-01 for October 2022 to September 2023 inclusive, a.k.a. the *reporting period*, with that of the same period in the previous year (October 2021 to September 2022 inclusive), a.k.a. the *baseline period*, using the building performance evaluation methodology detailed in Section 3.9. To apply the performance methodology, the buildings *energy signature* (thermal demand regression curve) is first derived using historical data from the baseline period (weekends are excluded). Specifically, the daily total thermal demand data, along with the daily average ambient temperature, is used to obtain the piecewise regression curve of Eq. 5.2 (see the change-point regression equation Eq. 3.15 in Section 3.9).

$$\widehat{E}_{\mathrm{th},d} = 3.21 + 41.86 \times (17.91 - T_{\mathrm{amb},d})^{+} \tag{5.2}$$

where $\widehat{E}_{\mathrm{th},d}$ is the estimated thermal energy demand (kWh) for daily mean ambient temperature $T_{\mathrm{amb},d}$ (°C) on day d. Comparing the above equation with Eq. 3.15, $\beta_1 = 3.21$ kWh is the baseline consumption, $\beta_2 = 41.86$ kWh/K is the marginal energy demand for each Kelvin decrease in temperature, and $\beta_3 = 17.91$ °C is the balance-point temperature of the building. Fig. 5.17 shows the daily mean ambient temperature for the two periods overlaid on each other, while Fig. 5.18 shows the regression line and historical data on the plot of daily demand vs. mean daily ambient temperature for Building B-01, excluding weekends.

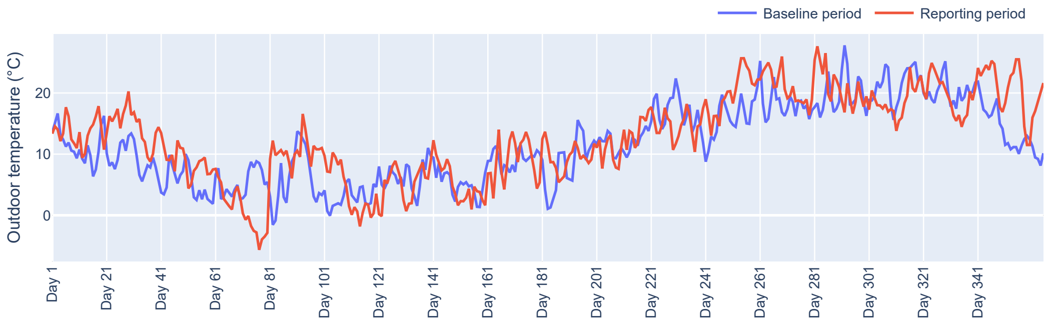


Fig. 5.17 Graphical comparison of ambient temperature for the baseline period (Oct. 2021 to Sep. 2022) and reporting period (Oct. 2022 to Sep. 2023) for historical comparison of Building B-01.

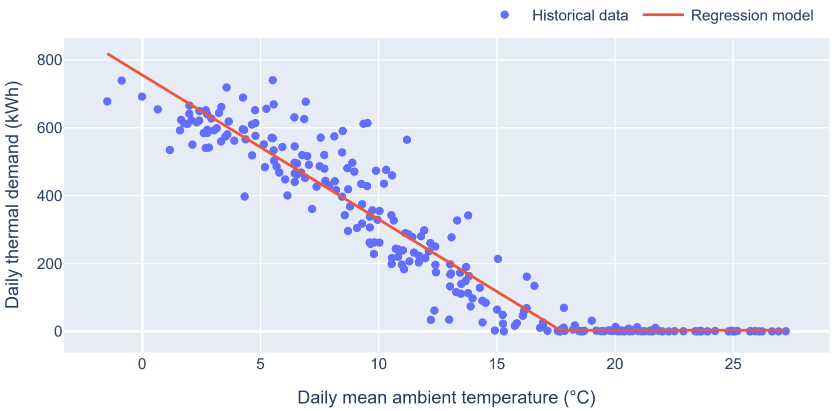


Fig. 5.18 Three-parameter regression model fitted on the historical data of Building B-01 to derive the energy signature for the baseline period. The data is from Oct. 2021 to Sept. 2022 (inclusive), using weekdays only.

The thermal energy savings in the reporting period, D, is therefore the difference between the actual demand and the estimated demand, previously given in Eq. 3.19 but repeated here for convenience:

$$E_{\text{th,saved},D} = \sum_{d \in D} \widehat{E}_{\text{th},d} - E_{\text{th},D}$$

where $E_{\text{th},D}$ is the total historical thermal demand in period D (in kWh).

As previously mentioned, there were issues with the heating controller in Building B-01, in which the desired temperature was usually not attained, especially after weekends and in the early mornings (see Section 5.1.5). Since the temperature shortfall was about 1.5 K on average, this "shortfall" in supply should be accounted for in the estimated savings by subtracting an energy *penalty* proportional to the shortfall from the estimated savings. Specifically, from Eq. 5.2 and Eq. 3.15, the dependence of the daily heating energy on temperature is captured by the slope term, $\beta_2 = 41.86$ (when ambient temperature is below the balance-point temperature). Hence, the savings in the reporting period is adjusted as below:

$$E^{*}_{\text{th,saved},D} = E_{\text{th,saved},D} - 41.86 \times HDD_{\text{shortfall}} \tag{5.3}$$

where $E^*_{\text{th,saved},D}$ is the adjusted thermal energy savings in period D. The shortfall in terms of heating degree days, $HDD_{\text{shortfall}}$, is estimated as follows. From the data, out of about 120 working days in the heating season that had daily mean temperature at least 3 °C below the building balance point temperature, the first 60 days were affected by this issue. For each day, about 4 hours had insufficient heating (1.5 °C below the desired temperature) within the working period (assumed to be 07:00 to 19:00). Thus, the **total heating shortfall** is therefore estimated as $1.5\,\text{K} \times 60\text{days} \times \frac{4}{24} = 15\,\text{K}\cdot\text{day}$.

Based on the fitted model of Eq. 5.2, the performance comparison of Building B-01 between the baseline period and reporting period is detailed in Table 5.5. As can be seen in the table, the model fits the training data with reasonable accuracy in the baseline period ($R^2 = 0.91$; RMSE = 73.68 kWh). Also, Table 5.5 shows that the reporting period achieved a savings of 18.6% (about 12.4 MWh) compared with the baseline. The heating shortfall estimated as 15 K · day above, would have required 41.86 × 30 = 627.9 kWh of additional heating energy, using Eq. 5.3. Hence, the adjusted total savings is 11,133.2 kWh, or approximately 17.66%. The heating shortfall-adjusted savings are shown in parenthesis in Table 5.5.

Table 5.5 Performance comparison of Building B-01 between the baseline period (from Oct. 1, 2021 to Sept. 30, 2022, inclusive) and the reporting period (from Oct. 1, 2022 to Sept. 30, 2023, inclusive). Weekends are excluded from the analysis.

Period	**HDD$_{15.5}$** (K · day)[a]	**Thermal Demand (kWh)**		**Difference** [b]		**Statistics**	
		Model	Actual	Abs. (kWh)	Rel. (%)	RMSE (kWh)	R^2
Baseline	1,333.7	71,928.2	71,916.5	11.8	0.016	73.68	0.91
Reporting	1,207.1	66,599.5	54,210.5 (54,838.5[‡])	12,388.9 (11,761[‡])	18.60 (17.66[‡])	-	-

[a] HDD is calculated using EU-standard base temperature of 15.5 °C for standardization purposes, and only for weekdays.
[b] Difference between model-predicted demand and actual demand.
[‡] Heating shortfall-adjusted values. See Eq. 5.3..

In order to examine how the building energy signature changed between the two periods, the energy signature of the baseline period is compared with that of the reporting period in Fig. 5.19, showing that the energy savings in the reporting period are as a result of lower thermal demand per unit increase in the associated driving temperature difference (equivalent to HDD) compared to the baseline period (β_2 = 38.4 kWh/K in reporting period vs. β_2 = 41.9 kWh/K in baseline period), and a lowering of the balance-point temperature by 1 °C (β_3 = 16.9 °C in reporting period vs. β_3 = 17.9 °C in baseline period).

Since the building envelope and energy systems remained the same throughout the period, the savings can be attributed to more efficient use of the heating system and more efficient building-occupant interactions. From the energy signature comparison curve of Fig. 5.5, the improved performance can be attributed to the schedule-based heating and lower setpoint temperatures as demonstrated in the penalty analysis of the previous section, which reduces the balance point temperature by reducing the overall thermal demand during office hours but especially during periods of absence, including at night. Note that these savings were achieved despite the fact that only about one-fifth of the 32 offices in Building B-01 were JuControl-activated, indicating more potential for energy saving, from both the user and control perspectives. In conclusion, the deployment of the automatic heating controller where heating is determined by presence schedules led to significant energy savings.

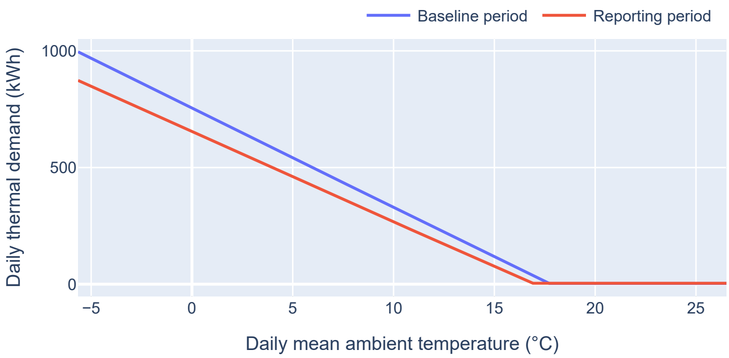


Fig. 5.19 Comparison of "energy signature" regression lines for baseline and reporting periods for Building B-01, showing reduction in balance-point temperature and slope (from β_2 = 38.4 kWh/K in reporting period vs. β_2 = 41.9 kWh/K in baseline period), and a lowering of the balance-point temperature (β_3 = 16.9 °C in reporting period vs. β_3 = 17.9 °C in baseline period).

5.6 Analysis of the Use of JuControl Calendar

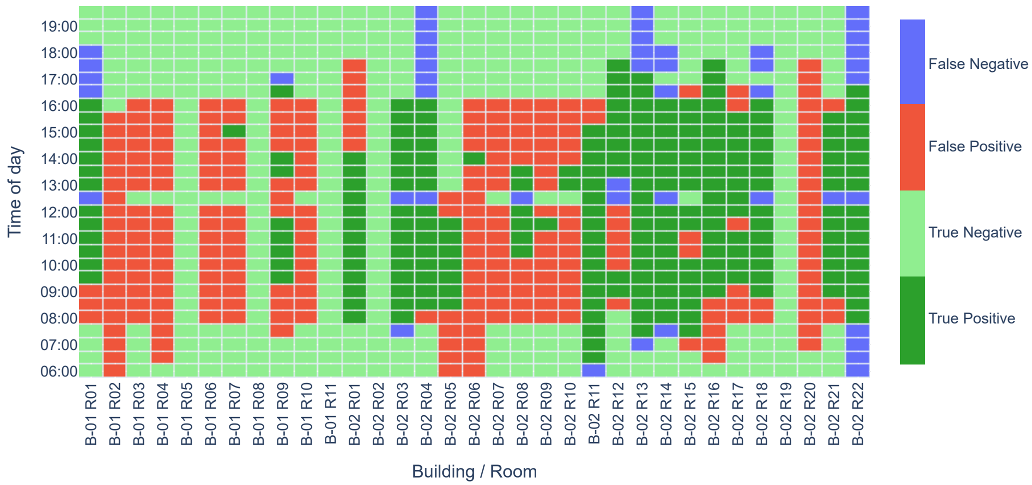
In this section, an analysis of the usage of the JuControl calendar is presented, specifically the agreement or otherwise between the actual presence of occupants in the office and their respective schedules as seen from the JuControl calendar. The analysis here only affects JuControl-activated offices in Buildings B-01 and B-02, where the automatic heating controller is deployed, since the schedule-based heating is only active in JuControl-activated offices. The occupancy detection for these buildings was based on CO_2 mass-balance, as described in Section 4.6, since there were no presence sensors in the two buildings. It should be kept in mind that while it is difficult to estimate the ground truth occupancy at the scale of deployment, the implemented occupancy detection was designed to be more lenient than strict, so that it rather tends to overestimate than underestimate real occupancy. Thus, the calculation of the setpoint temperature *deviation* is more forgiving, since it depends on estimates of real occupancy. As will be seen shortly from the holiday baseline case below, the occupancy detection was good enough to be practical.

The terminology used in this analysis to describe the mismatch between the schedule and real presence is akin to supervised learning terminology, where the calendar schedule can be thought as *predicting* the real presence. For example, when the JuControl calendar *predicts* that the office is occupied but it in reality it is not, then this is a *false positive* (see Table 5.6 for the full terminology).

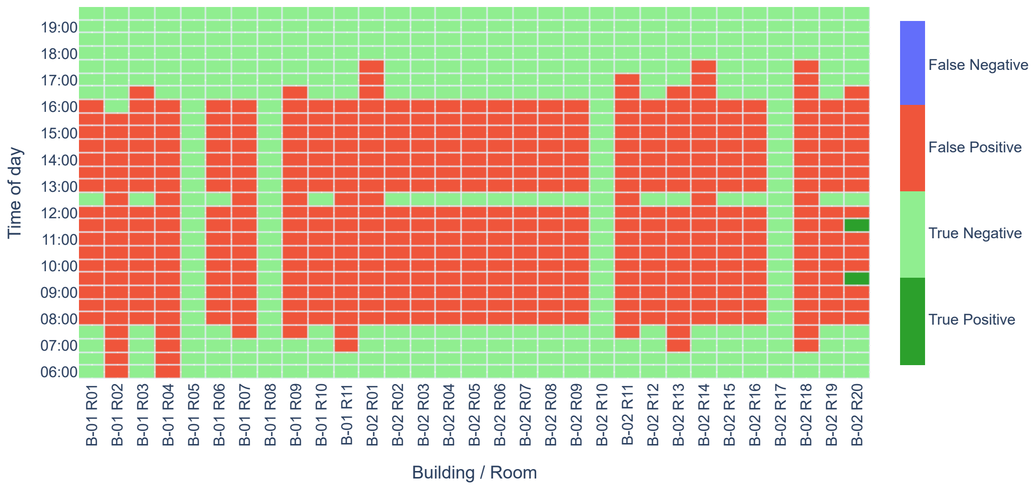
Table 5.6 Terminology for describing agreement or otherwise between occupant-specified JuControl calendar schedules and real presence. The terminology is derived by thinking of the calendar schedule as *predicting* the real presence.

JuControl calendar (*predictor*)	Computed real presence	Terminology for *prediction outcome*
Occupied	Occupied	True positive
Unoccupied	Unoccupied	True negative
Occupied	Unoccupied	False positive
Unoccupied	Occupied	False negative

In Fig.5.20, the degree of agreement between the JuControl calendar and real presence is shown for all JuControl-activated offices for two sample days – one workday and one public holiday, covering from 06:00 to 20:00 each day divided into 30-minute periods. As can be seen for the holiday (Fig. 5.20b), the calendar predicts presence for all the analysed offices, while the estimated real occupancy says the opposite, showing that occupants did not adjust their calendars to reflect the holiday. It is important to note here that for JuControl-activated offices, an 8.5-hour presence schedule with a 30-minute lunch break is enabled *by default* for each day of a working week, in order to ensure that at the offices are heated pre-emptively during working hours and avoid cold offices if the occupant never adjusts the calendar. The result also provides a benchmark for non-occupancy that demonstrates the practicality of the occupancy estimation. For the workday case (Fig. 5.20a), the rate of true positives seem to be higher for Building B-02 than for B-01, although the pattern is not necessarily consistent across all dates.



(a) Normal workday (12.04.2023).



(b) Easter holiday (10.04.2023).

Fig. 5.20 Comparison of JuControl calendar schedules with real occupancy for offices in Buildings B-01 and B-02. Each cell represents a 30-minute period for one office. In (a), a normal working day is shown, while in (b) a public holiday is shown.

Additionally, a comparison of the agreement between the JuControl schedule and real presence is performed for each date of the experiment for all included offices (excluding weekends), covering typical occupancy hours from 06:00 to 20:00 on each day in 30-minute increments. For this comparison, the frequency of the true/false positives/negatives for these 30-minute buckets is shown in Fig. 5.21. This plot shows that across all the analysed offices, employees were not present in the office as much as their JuControl calendar schedules indicated, with the false positive rate (i.e. calendar presence but actual absence) ranging from 25% to over 50%. This trend remained consistent in general throughout the experiment period. There were generally very low false negative rates (periods where the calendar predicted an unoccupied office, but in reality it was occupied). In general, the matches between the calendar schedule and real presence (true negatives + true positives) were almost always above 50% for each day, averaging about 61% for the entire experiment period.

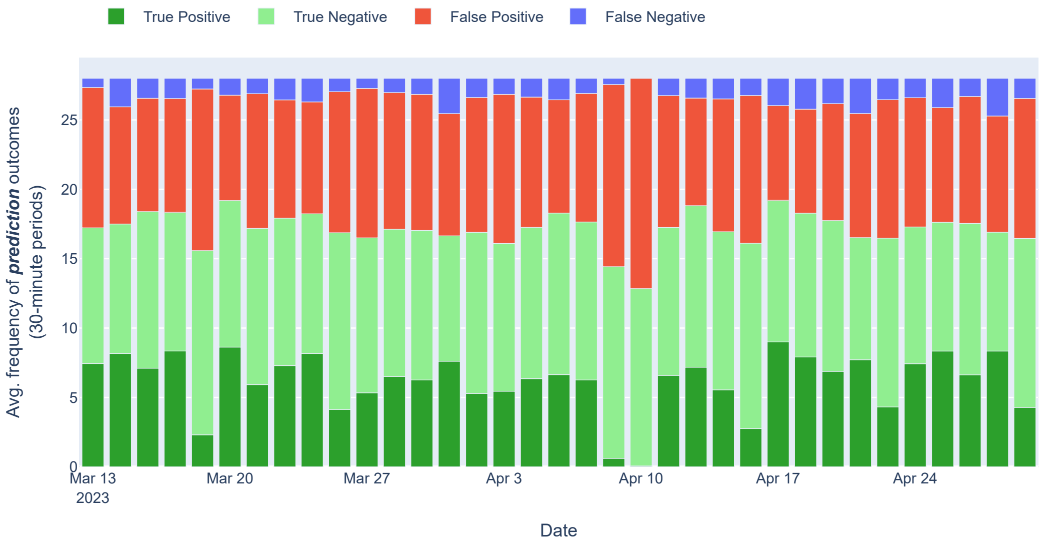


Fig. 5.21 Comparison of agreement between the JuControl schedule and real presence for the experiment period, showing average frequency of *prediction outcomes* for 30-minute buckets during the typical occupancy hours of each day for all analysed offices combined.

5.7 Analysis of Response to JuControl Recommendations

In this section, the effectiveness of recommendations received via email in triggering corrective action is analysed. For the three classes of recommendation message – "Trickle ventilation detected", "Ventilation exceeded", and "Setpoint temperature exceeded" – the triggers and corresponding corrective actions are shown in Table 5.7.

Over the four days during which recommendation was active, recommendations were triggered for offices 239 times in total, discounting those triggered in offices that were (later) disabled due to sensor faults. The sum of the unique email addresses *per day* receiving one or more recommendation emails over the period is 393, since multi-person offices received one email per occupant for each triggered recommendation. Each kind of recommendation message was triggered a maximum of once per day per office (hence also per user), even if the trigger conditions persisted or repeated during that day. The distribution of these 239 recommendation instances over the recommendation days, grouped by JuControl-activation status, is shown in Fig. 5.22.

Table 5.7 Triggers and corresponding corrective action for each class of recommendation email sent to participants.

Recommendation message	Trigger	Corrective action
Trickle ventilation detected	Any window becomes tilted	Fully open the window (side-hung) or close it
Ventilation exceeded	$N_{\text{vent,eq}} > N_{\text{vent,ref}} + 10$ min, i.e. more than 30 min equivalent ventilation	Close the window
Setpoint temperature exceeded	$T_{\text{sp}} > T_{\text{sp,ref,occ}} + 1.5\,°\text{C}$, i.e. more than 20.5 °C setpoint temperature	Reduce the setpoint temperature

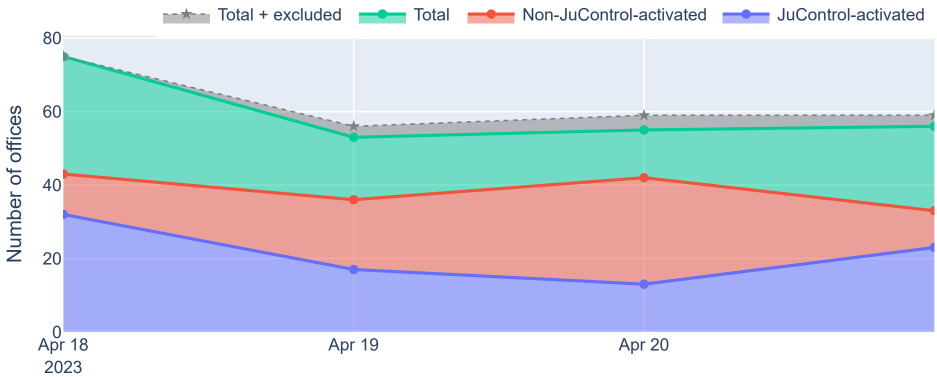


Fig. 5.22 Distribution of recommendations triggered by offices over the recommendation days. The "Total + excluded" plot accounts for offices where *all* occupants disabled email notifications, assuming the worst-case scenario that these offices could have triggered recommendations in subsequent days if they had not been disabled.

Since the recommendation emails contain a link for deactivating receipt of emails, some occupants used this option, resulting in less emails being sent. To account for this, the total number of *offices* that became completely deactivated from receiving emails for each recommendation day, is added to the "Total" plot in Fig. 5.22 for the *following* recommendation day to form the "Total + excluded" plot, under the worst-case assumption that these recommendation-deactivated offices could have triggered recommendations in the subsequent days if they had not been deactivated. Note that in offices where only *some* occupants disabled recommendation emails, the remaining occupant(s) could also in principle perform the recommended actions if on-site; therefore only offices where all occupants turned off emails are considered deactivated for recommendations. As seen from the figure, the corrected total number of recommendations reduced initially, then marginally increased and then remained relatively constant. This seems to suggest that the first recommendation produced the strongest effect, possibly due to its novelty property, as has been identified previously in the literature. Comparing JuControl-activated offices with non-activated offices, there is no consistent difference in trend, although it should be noted that due to these recommendations, a total of nine (9) offices transitioned to become JuControl-activated within this period, as shown previously in Fig. 5.6, which then possibly biased the trend for JuControl-activated offices towards increased triggering of recommendations.

In Fig. 5.23, the distribution of recommendation instances is broken down by team to show the trends amongst teams. The majority of teams exhibit a similar trend as in the totals trend of Fig. 5.22, with an initial reduction in triggered recommendations followed by a slight uptick. Only Teams T1 and T5 either maintained or reduced the number of recommendations triggered throughout the period.

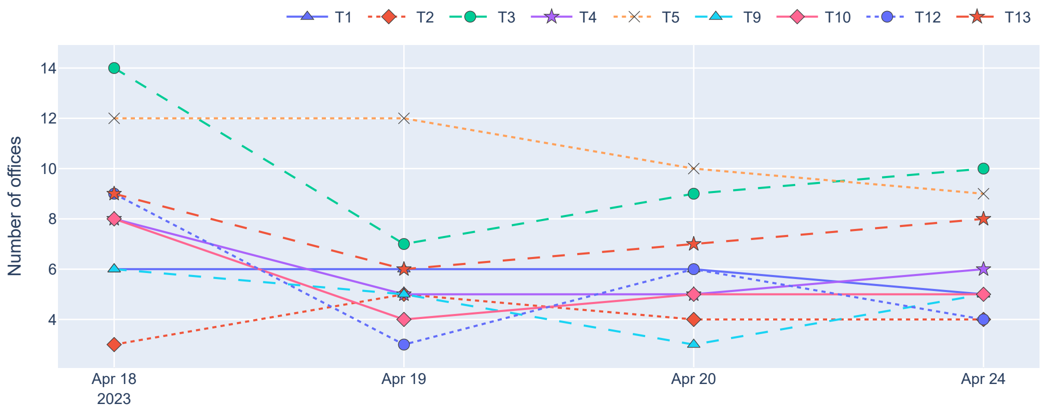
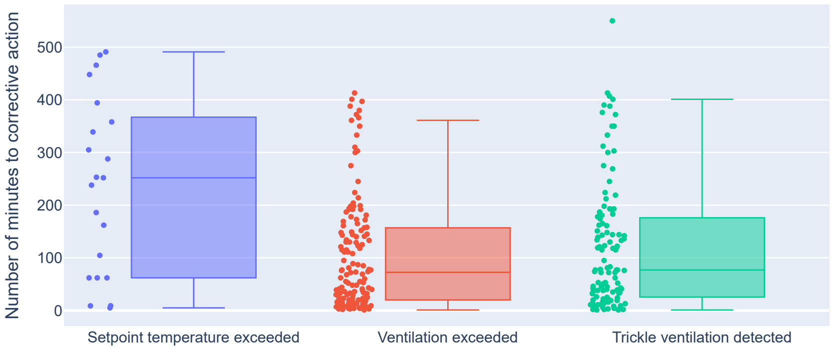


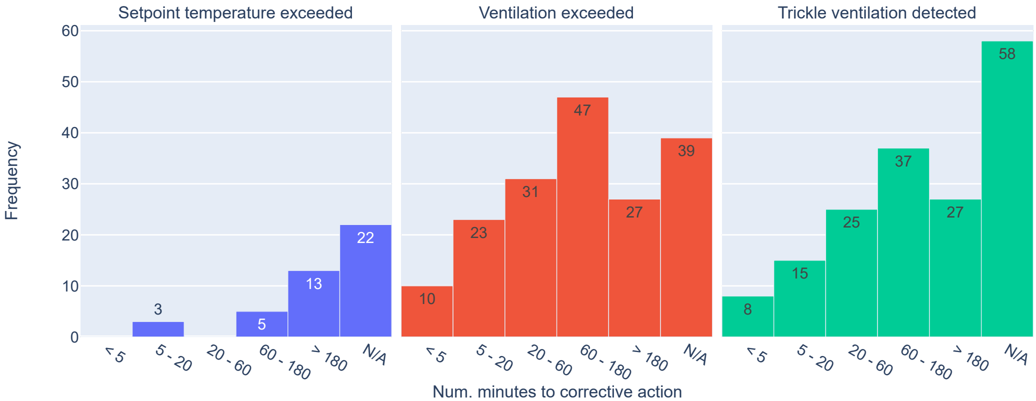
Fig. 5.23 Distribution of recommendations triggered by offices over the recommendation days grouped by team.

5.7.1 Occupant Response to Recommended Actions

To analyse the response of occupants to recommended actions, Fig. 5.24 shows the distribution of the elapsed time in minutes between the receipt of an email recommendation and the performance of a corresponding corrective action for each type of recommendation message for the entire recommendation period. In Fig. 5.24a it is shown as a boxplot, with median values of 238, 76, and 83 minutes respectively for setpoint temperature exceeded, ventilation exceeded, and trickle ventilation detected recommendation messages. The lower median and interquartile range for window changes than for setpoint adjustments is an indirect indication that manipulating the window position is done more frequently than adjusting the setpoint temperature. However, it does not seem to support the claim that most users responded to recommendations for corrective action *immediately*, since the relatively long time lag before corrective action can be reasonably assumed to indicate that the corrective action was not a result of the recommendation. According to Fig. 5.24b, only in a small number of office did occupants respond appropriately to ventilation recommendations within 5 minutes of receiving them (n=10 or 5.6% for ventilation exceeded; n=8 or 4.7% for trickle ventilation). Since it is not possible to know when such emails were read by the occupants, it is difficult to conclude definitively on the reasons for the time-to-corrective-action patterns. The survey analysis in Section 7.2 provides some information about user engagement with these emails.



(a) Boxplots of distribution of time-to-action.



(b) Histogram of distribution of time-to-action.

Fig. 5.24 (a) Boxplots and (b) histogram showing the distribution of the time elapsed (in minutes) before corrective action corresponding to each type of recommendation email was carried out, grouped by type of required action. Only detected corrective actions are represented in the boxplots, and in the histogram, the "N/A" bin indicates that no corrective action was detected during the day of the recommendation.

Chapter 6

Discussion of Results

In this chapter, the results of the previous chapter are discussed in detail. Two main aspects are covered in the discussion: the extent to which the experiment hypotheses could be confirmed or otherwise, and the comparison of obtained results with the literature where available and applicable.

6.1 User Engagement

As is typical in programmes that require voluntary user participation, a good percentage of users would be unable to participate for one reason or another. In this work, the user engagement results can be considered as promising, especially given the fact that there was no top-down or policy-driven push to adopt or engage with the developed systems.

In the present study, almost half of the potential employees of the experiment setup accessed the developed system within or before the experiment. Since there was no explicit logging of interactions, it is hard to say to what extent (in terms of frequency of use) occupants interacted with the systems. The multiplying effect of consent refusal in multi-person offices due to the unique nature of the data privacy policy, led to even higher attrition rates than would otherwise have occurred if JuControl activation were strictly based on individual consent, as discussed previously in Section 5.2. Future experiments using the developed systems are now taking the more favorable approach of limiting access only to provably personal data like CO_2 concentration in case of lack of consent, instead of blocking access to the entire system for all other occupants, including thermal comfort feedback and schedule-based heating.

Indeed, in behaviour intervention programmes, privacy considerations are always central nowadays in Europe especially, given the General Data Protection Regulation (GDPR)[1] policy active in many member states. Like in several other areas such as personalized advertizing and in health-and-wellness apps, there always exists some incompatibility between respect for user privacy and fulfilling business functions. As would be expected, previous behaviour intervention programmes in the energy sector have also been negatively affected. For example, the strict privacy requirements for anonymity and absence of communication among team members in the *Social Power* gamification project [80] detracted from the feeling of a sense of community amongst participants, thus robbing them and the research team the opportunity to exploit and understand the social motivational factors otherwise integral to such interventions. An *attrition rate* of 57% (i.e. participants dropping out of the experiment before it finishes) was observed in *Social Power*, attributable in a large part to the lack of social cohesion due to privacy requirements [80].

[1] https://gdpr-info.eu/

As a general guideline, following the completion of the EnerGAware programme, Casals et al. [71] surmise that experiment participation rate of one-third of initial participants should be expected by the designers ab initio, to account for the effects of attrition. In the present work, no attrition was observed directly by *JuControl deactivation* during the experiment, but this says nothing about the trajectory of the strength of participation of users, which was not directly measured.

6.2 Effect of Behaviour Interventions during the Experiment

Hypothesis H_1 that the energy performance of offices with evaluation / recommendation would be better than those of offices without these, was tested using H_1-Test-1. Hypothesis H_1-Test-1 compared the performance of two equal-sized teams derived from the same building, and a significant difference was found between the mean energy penalties of the test team and those of the control team, with a large effect size. Specifically, the test team had significantly lower energy penalties on average than the control team. When considering the contributing factors to the difference in performance, namely the ventilation style and durations in both teams, it is seen that Team T5 used more of shock ventilation than Team T6, and ventilated for shorter durations on average in each ventilation period.

However, the splitting of the building could have introduced a bias, since from the beginning of the experiment, the mean consumption of Team T5 was generally lower than that of Team T6. One possible bias could be that the occupants in Team T5 have a more energy-efficient disposition due to differences in professional background. Indeed, analysis of the level of engagement with JuControl shows such a skew in favour of research scientists or technicians, especially those with an energy-related background. However, the split appears to be a fairly evenly mixed population of research and management staff in both teams. Another possible explanation for the bias could be the difference in floor level, where Team T5 occupies the bottom two floors, and Team T6 the upper two floors. Again, a few offices in Team T5 already had informal access to JuControl before the actual experiment announcement, although the sensors in the building were only commissioned the week before the experiment. At any rate, there does not appear to be a clearly identifiable composition or environmental differences between the two teams from this post-mortem analysis. At the whole-building level, since Building B-05 was not homogeneously used for energy saving measures, and the duration of the experiment was short compared to the duration of the entire heating season, the effect of the interventions on building thermal energy demand is expected to be negligible.

The second hypothesis regarding the behaviour interventions during the experiment posits that JuControl-activated offices would perform better than non-activated offices within the same team. JuControl activation is used here as a measure of active engagement with the intervention system, since most of the behaviour modifying aspects of the intervention, including eco-feedback and competition, were only available after JuControl activation. This hypothesis was tested by and H_2-Test-1, and as the results demonstrated, in the teams where there was good adoption of the implemented behaviour intervention systems (Teams T1, T3, and T5), barring technical and exceptional issues like discussed for Team T4, or low adoption of the system like in Team T2 and T9, JuControl-activated offices show moderate-to-strong statistically significant improvement in performance when compared to non-activated offices, also with mostly very large effect sizes. The possible bias discussed above due to systematic splitting of Building B-05 does not apply in these cases, since office activation within a building can be seen as random for this purpose. Nevertheless, it is conceivable that occupants of offices physically close to each other might also tend to communicate more than occupants of distant offices, and thus when an office becomes activated, a neighbouring office could come under more social pressure to do the same. In any case, the trend of higher performance of JuControl offices was strongly demonstrated across multiple buildings. In terms of ventilation styles, the predominant style in non-activated offices was trickle ventilation, while activated offices tended to follow the JuControl recommendation of shock ventilation, resulting in a more even mix of ventilation styles. Therefore, the analysis of this section confirms hypothesis $\mathbf{H_2}$ that the interventions had a positive impact on the energy efficiency of occupants, as demonstrated by the superiour performance of JuControl-activated offices, in which occupants opted to take full advantage of the energy-saving behaviour improvement opportunities afforded by JuControl.

In conclusion, the strong results of the tests of the above two hypotheses enable us to reject the respective null hypotheses and conclude that having access to recommendations and evaluations, as well as to the relevant contextual performance information in JuControl, led to more energy-efficient behaviour among occupants, than in the control group without these features, as determined by the performance evaluation methodology developed in the thesis.

Indeed, performing a three-way comparison among Team T6 (without any interventions), and *JuControl-activated* and *non–JuControl-activated* offices in Team T5 (with evaluation / recommendations), we see a progression of improvement corresponding to higher levels of accessibility of the user to the designed intervention measures. Specifically, with Team T6 as the base case having average daily energy penalty of 4.67 kWh, we observe a statistically significant improvement of 57% for non–JuControl-activated offices of Team T5 compared to Team T6 (down to 2.0 kWh), and a *further* 63% improvement for JuControl-activated offices of Team T5 compared to non-activated offices in the same Team (down to 0.74 kWh). These results are testament to the huge potential the developed interventions have to positively affect user behaviour in office buildings in a real-world setting where there are neither financial incentives for employees nor mandatory personal responsibility for energy efficiency.

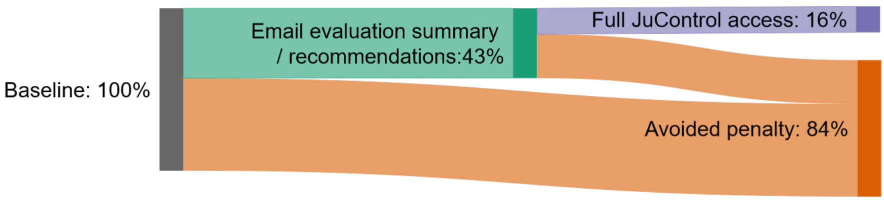


Fig. 6.1 Progression of energy efficiency with increasing exposure to interventions showing incremental improvement in energy performance as derived from the results of testing Hypotheses H_1 and H_2 on Building B-05.

Eco-visualization and feedback have been used widely for engaging building occupants and facilitating energy-related behaviour change. Similar to the interactive room in JuControl and its traffic light energy rating system, Francisco et al. [158] tested occupant responses to a building information modeling–based interface that colors zones based on energy consumption. Their results show that such representations lead to improved user engagement, and the 2D version of the interface was more engaging for users than the 3D version due to its simplicity. Regarding the use of energy penalties instead of rewards, Jain, Taylor, and Peschiera [87] suggests that rewards have more positive effects on users than penalties, although they only tested the initial view the user has on logging into their interface (negatively or positively signed reward points). Nevertheless, like in our case study, they find a positive correlation between interaction with their digital platform and energy savings among the participants.

Also, since JuControl provided more than mere visualization by including traffic-light–based energy ratings, recommendations, and competition, the results agree with existing literature that indicate that visualization alone does not produce behaviour change [26, 75, 159, 160]. For example, Peschiera, Taylor, and Siegel [161] showed that no change in behaviour was noticed when occupants of a school dormitory merely saw their energy consumption, as against those that had the consumption compared to either the building or peer network average.

6.2.1 Impact on Indoor Air Quality

An analysis of the indoor air quality (IAQ) as represented by the CO_2 concentration shows that there were no significant differences in the daily mean concentration of CO_2 in offices of Team T5 (μ = 883 ppm) as compared to offices of Team T6 (μ = 901 ppm) during the experiment period. Here, each data point is the mean when the CO_2 concentration is grouped by datein each team [2]. Since the two teams are drawn from the same building, it is reasonable to assume that air infiltration rate differences in the offices cancel out across the two teams. The same observation of comparable CO_2 concentrations holds for the activated vs. non-activated offices in Team T3, where the test of Hypothesis H_1-Test-2 also showed significantly better performance for activated offices than for non-activated ones in the team.

In Team T1, a significant difference in mean is observed between the CO_2 concentration in activated offices (μ = 905 ppm) and in non-activated offices (μ = 705 ppm) ($p<.001$) with a very large effect size (d=1.8). Similar but less extreme difference is also observed for activated vs. non-activated offices in Team T5 with medium effect size (activated: μ = 954 ppm; non-activated: μ = 859 ppm, p=0.01, d=0.5).

In conclusion, although in some cases the indoor CO_2 concentration was statistically significantly higher in experiment group than in its corresponding control group, these higher concentrations were still below 1000 ppm on average. With the new indoor air quality–based approach to ventilation evaluation implemented in the new version of JuControl, the CO_2 concentrations in the test groups are expected to be lower in the next experiment.

6.3 Effect of Automatic Heating Control

For the whole-building energy savings of nearly 18% realized in Building B-01 for the test year compared to the reference year, previous work indicate similar results, where it has been noted that a significant portion of energy wastage in buildings is due to unnecessary heating or cooling of unoccupied spaces [162, 163]. The analysis carried out by Meyers, Williams, and Matthews [163] estimates potential energy savings of 14–20% due to *not* heating unoccupied rooms in homes in the United States, and even further savings with lowering of excessively high temperature setpoints. Becker et al. [164] investigated the energy savings from simulated occupancy-based heating for several thousand households as compared to the real-world energy consumption without a setback temperature. The results showed average energy savings of 9%, with about 11% and 5% of these households being able to save up to 15% and 20% respectively. Interestingly, the 5% with the highest savings potential share similar characteristics with Building B-01 in our study, i.e. being old and having rooms that are unoccupied for several hours a day. Furthermore, Iria et al. [77] report electricity

[2] "Grouped by date" means that each data point is the average CO_2 concentration across all offices in each team (e.g. T5 / T6) or activation category (e.g. T5-activated / T5–non-activated) for each date of the experiment period. The alternate grouping was also analysed, where each data point consists of the mean of all experiment dates for each office (a.k.a. grouped by office). In all cases, "grouped by office" yielded non-significant differences in all comparisons, so the more varied "grouped by date" is presented.

savings of 20%, and the review of Zhang et al. [165] estimates energy savings of 5–30% in commercial buildings, due to occupant behaviour change. Additionally, Peng et al. [166] reports savings of 7–52% in cooling energy in a commercial building by using machine-learning–based occupancy-driven cooling.

Since in Building B-01, the building envelope and energy systems remained the same during the baseline and test periods, the savings can be attributed to more efficient use of the heating system and more efficient building-occupant interactions. Specifically, from the energy signature comparison curve of Fig. 5.5, the improved performance can be attributed to:

- schedule-based heating and lower setpoint temperatures as demonstrated in the energy penalty analysis of the previous section, which reduces the balance point temperature by reducing the overall thermal demand during office hours but especially during periods of absence, including at night.
- less wasteful ventilation as shown in the previous section, which reduces the overall building heat loss coefficient accounted for by the slope term, β_2 in Eq. 3.15, as well as reduces the balance point temperature, since the balance point temperature depends on the overall heat loss coefficient.

Note that these savings were achieved despite the fact that just 8 out of the 32 offices in Building B-01 were JuControl-activated and hence had automatic setpoint temperature regulation, indicating more potential for energy saving, from both the user and control perspectives.

6.3.1 Cost-Benefit Analysis for Instrumentation

For the pilot building, Building B-01, a total of 183 wireless sensors and valve actuators were installed in 38 spaces, costing €23,330 in total (see Table 6.1). There were no KNX-protocol–based devices in the building. The costs for the installation of sensors and actuators are not considered here, since after the development of a sophisticated workflow for the preparation and installation of these devices, the time required for the equipping further offices is minimal. If only the installation costs are compared with the reduced energy costs due to energy savings which amount to 11.8 MWh, assuming the price for heat of €0.1647/kWh (2022), it would take 12.0 years to fully recover the investment. A few important remarks should be made here. First, only 8 offices were JuControl-activated in the building, meaning that the heating controller managed only these rooms with respect to dedicated user-specific schedules in the reporting period. In a best-case scenario where all offices are activated and occupants specify schedules that are up to 90% in alignment with real presence, the energy savings compared to an uninstrumented baseline would be expected to be above 30%. In this case it would take at most 7.2 years to recover the investment. Indeed, after further consultation with the works council and the data protection officer, in a future upgrade the automatic heating controller would always be enabled, regardless of JuControl activation status. Apart from the energy savings, the installation setup also provides occupants an insight into the indoor air quality and consequently ensures better air quality on average.

As can be seen from the table Table 6.1, the Indoor Air Quality (IAQ) sensors, and sensors for window/door states account for around 50% and 22% of the hardware costs respectively. In a minimal setup where the detection of ventilation patterns is based on software sensors (e.g. based on temperature profile) and the IAQ sensor could be replaced by an EnOcean-based temperature-only sensor (costing approximately €50 instead of €300) the payback period could be reduced to 4.5 years. When switching from EnOcean to LoRaWAN, the IoT gateway and the transceivers could be replaced by a single, significantly cheaper LoRaWAN gateway and the cabling costs would be eliminated. In such a case, the payback period could be decreased further. This effect especially holds for smaller setups with a limited number of rooms like in the pilot building.

Table 6.1 Installed device types and their associated purchase costs.

Device Type	Unit Price	Quantity	Total Price
IoT Gateway	€ 1,500	1	€ 1,500
EnOcean Transceiver	€ 120	8	€ 960
Cabling of Transceivers	€ 650	1	€ 650
IAQ multisensor	€ 300	38	€ 11,400
Window handle	€ 70	54	€ 3,780
Contact sensor	€ 35	38	€ 1,330
Valve actuator	€ 70	53	€ 3,710
Total			€ 23,330

Regarding the estimated payback period of 4.5 to 12 years, Iria et al. [77] estimate 15 years of payback period for electricity savings in a gamified application involving similar instrumentation, and 4-6 years if commercialized. The reported energy savings was 20%.

Rotondo et al. [24] report says up to 30% savings in residential buildings in the United States via connected thermostats.

Chapter 7

User Feedback and Critique of Methodology

This chapter first deals with the feedback of users during the project, which was obtained primarily through Co-Design workshops during development and testing, and a user survey conducted after the main experiment. Additionally, the methodology adopted in the work is critically analysed to identify potential improvements, also considering user feedback. Finally, recommendations are provided based on the foregoing analysis.

7.1 Feedback from Co-Design Workshops

The development of the software and tools in the *Energy Dashboard Suite* followed a co-design strategy, in which volunteers drawn from different departments of the campus contributed to the planning, design, and review of features of these software (similar to [14]). Before these workshops, early access was granted to participants prior to the main release of the Dashboard, and before new major features were made available.

The first co-design workshop, which focused on the evaluation of the first version of the Campus Viewer, took place in February 2020 and involved about 25 participants. The second co-design workshop took place about a year later in January 2021, where the implementation of the outcome of the first workshop was reviewed. Concomitantly, the initial version of JuControl containing the visualization and control features was presented in the same workshop and received feedback from participants. This initial version had been developed solely according to the ideas of the author and project collaborators. A rough roadmap for gamification in JuControl was also presented. About 20 employees participated. After the implementation of gamification in JuControl, no further co-design workshop was carried out for JuControl.

7.2 Post-Experiment User Survey

A user survey was designed and implemented within JuControl after the experiment had ended, in which participants were asked about their experience during the experiment. The goal of the questionnaire was to gain insight into factors affecting the outcomes of the experiment from the users' perspective, and to identify areas of improvement. Thus, the results of the experiment discussed in the previous chapter can be better explained when considering the additional information afforded by the survey. The start of the survey was announced to all participants on June 20, 2023 via email, and feedback was received until October 16, 2023. Since the survey was directly implemented into JuControl, no "activation" was required to access it. Additionally, on the JuControl interface, a popup regularly appeared, which could only be snoozed for 6 hours. The survey garnered 113 responses, whose analysis is presented in the following subsections. The questionnaire was designed to be dynamic – questions and/or options that were displayed to each participant depended on the features enabled for the participant, and the response to one question determines the next question to be displayed. Additionally, relevant images were included that served as memory and visual aids to respondents. An example screenshot of a survey question is shown in Fig. 7.1.

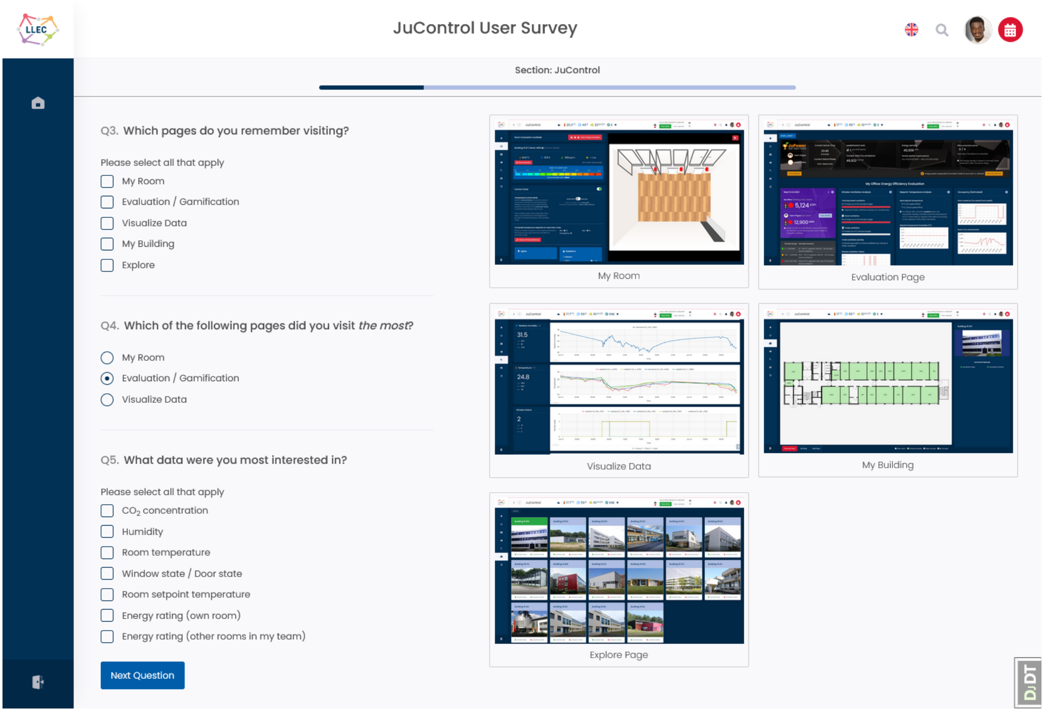


Fig. 7.1 Screenshot of a survey page in JuControl showing pictorial memory-aids associated with the questions.

The results of the post-experiment user survey is presented in this section. The survey was carried out between June 20 and October 16, 2023. In total, there were 113 respondents. However, since the questions were programmed to be context-sensitive, meaning that subsequent questions depend on previous answers and the features available to the respondent's team, some questions were not shown to all 113 respondents if a qualifying condition for displaying the question was not met. The distribution of respondents according to team, categorized by experimental group, is shown in Fig 7.2. The highest number of respondents according to team came from Teams T7 (n=13) and T8 (n=10), which belong to the same building in which faulty sensors were discovered and thus evaluations disabled.

In terms of getting to know about JuControl, Fig. 7.3 shows that the most common response was that users knew about it from colleagues (n=50), although there were several other information dissemination methods attempted, including posters at the building entrance (see Fig. B.17 in Appendix B for a sample poster). In terms of frequency of use of JuControl during the experiment period, out of 103 respondents, about half visited only a few times throughout the experiment period (n=54), although more than a quarter (n=30) visited JuControl several times per week, while 15% (n=15) never visited it and less than one-tenth (n=9) visited it several times a day. Amongst those who visited JuControl at least once (n=88), "My Room" page (see Fig. 4.2a) was the most visited (n=69), followed by the "Data Visualization" page (Fig. E.6 in Appendix E) that shows plots of historical data (n=13). Very few respondents visited the Evaluation/Gamification page (n=6).

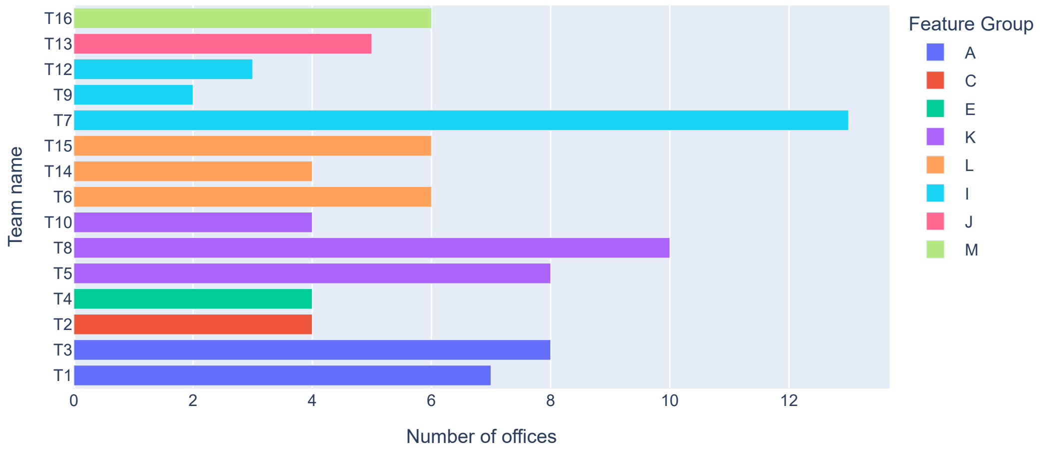


Fig. 7.2 Survey response by team, also indicating the group to which the team belongs.

This suggests that placing basic evaluation results on "My Room" page would likely increase its visibility, while detailed results can be reserved for the "Evaluation" page.

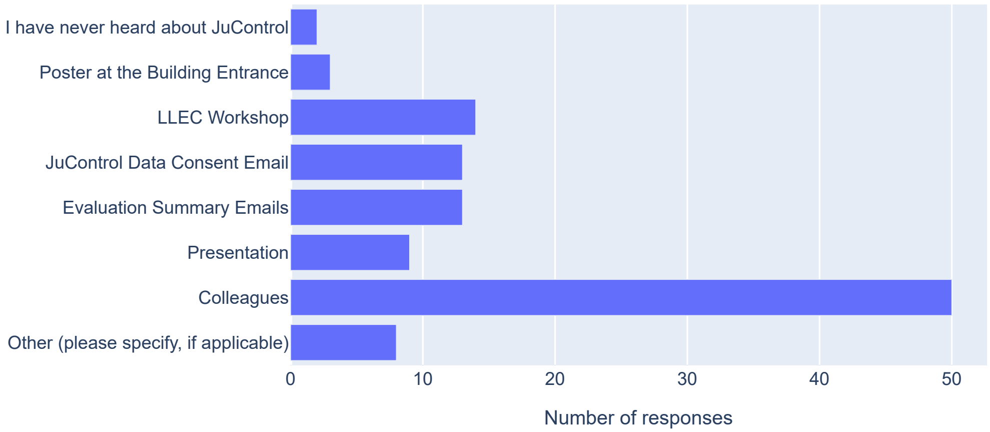


Fig. 7.3 Survey response to "How did you know about JuControl?", showing that word-of-mouth information transmission was the most effective.

When asked about which data they were most interested in (Fig. 7.4), respondents mostly said room temperature (n=38), which is probably because in winter season, thermal comfort is one of the most important needs for building occupants. This was followed by energy rating of own room (n=22) and energy rating of other rooms (n=10). The least common response was CO_2 concentration (n=3), which seems to imply that indoor air quality was not a priority for most respondents. However, since the question allows selecting only one option, it is likely that CO_2 concentration still mattered to the respondents, especially since most respondents viewed "My Room" page most often, which does *not* show energy ratings at all but indoor air quality parameters, setpoint temperature (available to respondents from Building B-01 and B-02 only), and window / door state.

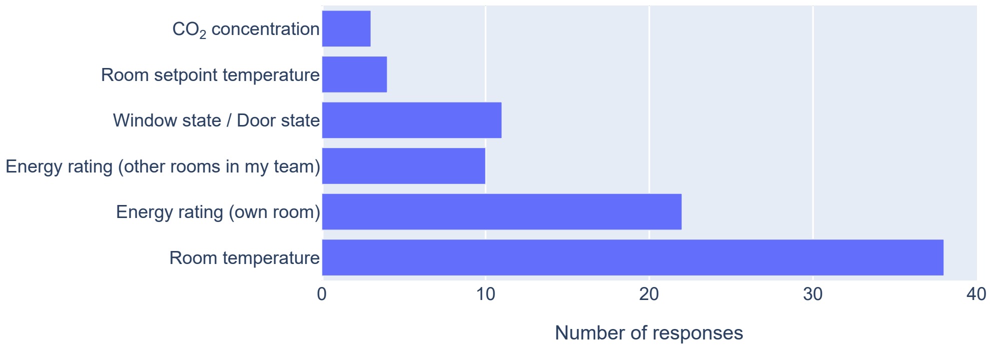
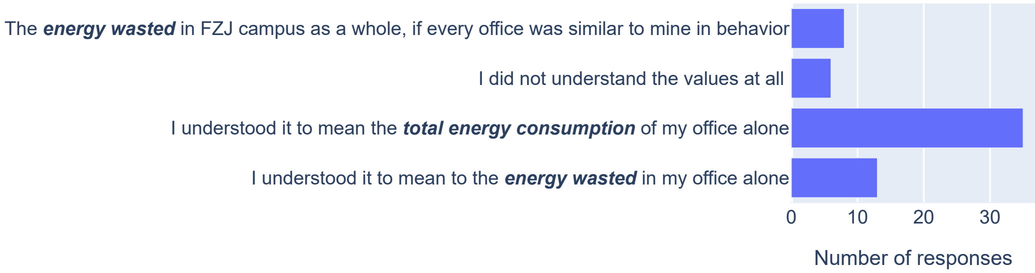


Fig. 7.4 Survey response to "What data were you most interested in?", showing that room temperature was most important to the respondents.

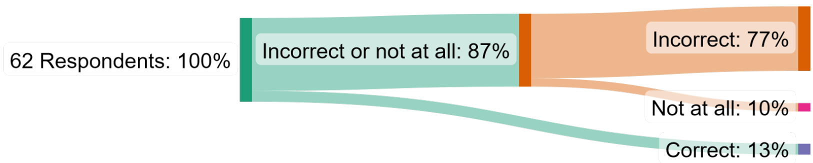
7.2.1 User Response to Evaluations and Recommendations

A number of questions in the survey dealt with user response to evaluations. Regarding evaluation summary and recommendation emails (see Section 5.1.2), out of 85 respondents, two-thirds remembered receiving the emails and read them at least once (n=64), while only one person received it but did not read it; the rest didn't recall receiving it or did not see the emails (only offices with evaluation or recommendation enabled according to the experiment design had access to this question). Among those who recalled receiving the emails (n=65), almost half (n=31) liked the idea and thought the emails were interesting, while more than one-fifth (n=14) found them annoying and uninteresting. Eleven respondents were indifferent, while the remaining nine found them annoying but interesting. Out of those who found the emails annoying (n=23), when asked about the *most* annoying thing about the emails, the top response was that the data in the emails was erroneous or inaccurate (n=9), followed closely by the emails being spam-like (n=8). Most of the respondents that received the emails claimed nevertheless that the emails were quite or totally understandable (n=39), and one-third said it was only a little understandable (n=22). In response to whether the emails affected their behaviour in terms of window use and/or heating setpoint temperature, the majority said "No" (n=39), while the remaining said "Yes" (n=24), which shows that the option of sending emails is effective to an extent. On the general subject of energy ratings as a means of improving their personal behaviour with the assumption that the ratings were "properly implemented", half of the respondents answered that it would only be "a little" effective for them, while more than one-third (n=23) said "totally" effective or "to a large extent". Eight answered in the negative.

In order to further test the respondents' understanding of the evaluation penalties, occupants were asked to select the option that reflected their understanding, as shown in Fig. 7.5a. The responses revealed, however, that penalty values were mostly misunderstood (Fig. 7.5b), which implies that the claim of "erroneous or inaccurate" values above was at least partly due to misunderstanding. One factor that possibly contributed to the misunderstanding was the wording of the first version of the evaluation summary emails, shown in Fig. B.15, which did not clarify that the given penalties were scaled to the size of the campus, although the help files and contextual help buttons in JuControl explained the concepts. Indeed, when respondents were asked in a follow-up question about what they felt about the comparison of their performance with that of other offices in their team, the most common response was that they believed something was technically wrong with the comparisons (n=29), while less than one-quarter said they were motivated to perform better (n=15). One respondent stated they did not understand the comparison, while the rest selected "No thoughts / I did not notice the comparison" (n=17). The implication for the User Experience design is that the most natural understanding of "energy rating" among the target audience was energy consumption in *own* office, unlike the *penalty-scaled-to-campus-size* approach that was used in this thesis. Hence, energy rating of own office would have been a better way to present the evaluation results to users. Again, this underscores the importance of user acceptance tests before deployment of such tools. In Chapter 8, lessons learned are presented, including how to improve the User Experience in the developed tools.



(a) Response to survey question



(b) Breakdown of survey response

Fig. 7.5 Respondents' understanding of energy penalty values. (a) Survey response options and frequency of response. (b) Breakdown of survey responses, showing that most respondents did not understand the meaning of the energy penalties.

Finally, many improvements were suggested by the respondents in the free-text field following the above-analysed questions. The most recurrent theme was about having more "meaningful" penalty values, which relates to the misunderstanding of penalty values discussed above. Other themes included fixing sensor and software issues, having a real-world bonus system that rewards top performers, and performance comparison against own office benchmark instead of against other offices. Additionally, the unavailability of German translations for some parts of JuControl (especially the Evaluation / Gamification part) negatively affected more than one-fifth of the respondents (21 out of 96).

7.3 Critique of Experiment Methodology

All through the planning and development phases of the project, up to and including the testing and feedback rounds, several issues were identified that should be improved upon to achieve more effective systems and better quality results. Some of the proposed solutions are already integrated in the system and would be described in the following subsections.

7.3.1 Improved Publicity

The Energy Dashboard Suite would benefit immensely from better publicity, which potentially improves its effectiveness in at least two ways: a lower psychological barrier-to-entry for the users, since the publicity already primes the user, thereby reducing the foreignness of the products; and, more understanding of the Dashboard ecosystem, since the publicity also provides background and direct information that help reduce the mental burden the users experience when interacting with the apps.

Already, within the scope of this work, several publicity campaigns were carried out, including dedicated information workshops and Co-Design workshops. In addition to these, more avenues should be explored, where the attention of employees of FZJ can be captured, including on the Intranet website, in public screens like at the canteen, and in the internal newsletter. These avenues were discussed during the project, but were not exploited so far.

On the development side of the apps, summary or overview pages should be created that are appropriate for public screens. The corresponding API endpoints for such summary pages should be implemented, as well as authentication models that are not person-based (e.g. non-Shibboleth authentication in the case of the Campus Viewer / JuControl), since such display devices do not represent human users of the app.

Real-world reward tokens for improved publicity and engagement

In order to improve visibility and publicity, as well as user engagement, small physical tokens of reward can be given to participants for feats accomplished across the Energy Dashboard suite. For example, JuControl tracks "achievements" for teams and individuals related to e.g. low CO_2 emissions, some of which are computed on a weekly basis (in game-time). In JuControl also, streaks of energy savings could receive publicity and tokens as rewards.

A referrer-based reward program can also be developed and integrated into the applications to further enhance the social aspects of the apps and tools. Under such a program, users are provided means to invite participants to the experiments. An exemplary means is via customized links that can be shared with colleagues, which then identify the referrer when the invitee joins the system. This is standard practice in referrer-based programs.

7.3.2 Improved Experiment Design

One of the possible flaws of the experiment design employed in the thesis was the method that was used to divide some buildings into experiment groups, in which a set of contiguous rooms were assigned to one group, and the next set of contiguous rooms were assigned to another group (for example, assigning top floor to one group and bottom floor to another group). The flaw in this strategy lies in the fact that it is possible that such a grouping could contain a latent confounding factor which correlates the contiguous rooms and makes it difficult to compare the two groups. These factors include which institutes occupy the rooms (in practice, institutes are usually collocated when multiple institutes share a building, and energy-related attitudes might differ across institutes), and possible differences in building characteristics (e.g. offices in the top floors could experience higher heat losses through the ceiling boundary). Naturally, when the experiment goals require that user behaviour differences according to job description or educational background are investigated, then this collocation approach suits perfectly. A better strategy in general could be to randomly assign the rooms to different experiment groups, or a round-robin assignment, thereby cancelling out such unwanted correlations. This method, nevertheless, has its own drawbacks, including increased risk of "cross-talk", whereby information meant to be siloed in one experiment group leaks out into other groups, especially given that the user survey results indicated that employee-to-employee communication contributed the most to the spread of awareness about JuControl. In any case, the comparison between JuControl-activated offices and non-activated offices in this work was not susceptible to this office assignment bias, as it can be considered close to random assignment.

Furthermore, the next experiment should be structured such that the most attention is paid to the most important hypotheses to test, meaning that offices should be assigned to experiment groups redundantly prioritizing the most important hypotheses. Likewise, the number of hypotheses or effects to be investigated should be appropriate for the sample size to account for *attrition* of participants and possible data collection issues. Again, testing should be carried out in the core of winter to allow simplification of the analysis, since weather would then not be a significant confounding factor. Finally, a longer test period should be chosen, allowing measurements to be taken before official experimentation begins, and after experimentation ends (at least two weeks each of pre-experiment and post-experiment measurements, and at least 8 weeks of the experiment itself). This enables analysis not just by comparing experimental groups, but by comparing particular teams with their historical baseline performance and their performance after the "stimuli" is removed. The latter can answer questions about the ability of the interventions to foster *intrinsic motivation* in participants.

Finally, periodic surveys can be conducted during the experiment, consisting of short questions regarding user experience difficulties followed possibly by quick fixes to improve the systems online. These surveys could be presented to the user as dismissible prompts on the JuControl interface, without pestering users with emails. This provides a more unobtrusive alternative to emails.

7.3.3 Improved In-App Monitoring and Measurements

The analysis in this work would have been easier if the level of interaction of users with the developed tools was measured and stored by instrumenting certain aspects of the applications. The degree of possible instrumentation can range from high-level interactions (page visits) to detailed interactions (click, hover, and scroll interaction measurement). Importantly, the issues around user privacy and consent have to be handled. For JuControl, an appropriate degree of instrumentation should ideally answer the following user engagement questions, so that office evaluations can be better correlated with interactions with the developed system.

- Which pages did the user visit, how often, and how much time did the user spend on each page?
- Which page views are most correlated (i.e. which pages are most likely to be viewed when the user is on a particular page)?
- Which widgets did the user interact with (e.g. the setpoint temperature widget in JuControl's control panel) and how often?
- Which automated emails were read, and which links in the emails were followed by the user (the recommendation emails for the experiment, for example, had links to JuControl for getting more information about the recommendations)?

In order to correctly answer the above questions and avoid false positives, certain technical considerations regarding the implementation should be borne in mind. In particular, logging HTTP requests is not a good way to measure how much time the user spent on a JuControl page, since by design, JuControl makes frequent HTTP requests to update its state, which also happens when the browser is running in the background. Rather, mouse movements can be regarded as being indicative of (active) interaction with JuControl, and these can be harvested through some throttling mechanism to avoid data explosion. Likewise, in order to track page-view correlation, page loads through bookmarks or directly via the browser navigation bar should be differentiated from page loads triggered by clicking on links within a page. The latter kind of page loads can then be stored with the page containing the clicked link as the source page, and the link's target as the target page.

7.3.4 Indoor Air Quality as Focus of Energy Evaluation Strategy

The approach used by Juracle to estimate the energy performance of offices uses a fixed set of criteria that were selected to be reasonably applicable to multiple shapes and sizes of rooms. Particularly, the estimation of ventilation efficiency in this work assumed a distinction between the efficiency of trickle ventilation (windows tilted on bottom hinges) and shock ventilation (windows fully swung open on their vertical hinges). This assumption was supported in the extant literature and in government campaigns in Germany.

However, from the simulation results of the reference models, the difference between the efficiency of trickle ventilation and shock ventilation was not significant for the most common window configuration (all windows on one outside wall). Rather, it seems that the aversion for trickle ventilation stems more from the fact that usually people adopt this ventilation style for extended periods in practice, than that trickle ventilation is inherently wasteful. In other words, the inefficiency relates rather to the duration of ventilation, and not the ventilation style. Since trickle ventilation leads to slow heat losses, the heating system often is able to compensate for the energy losses for long periods, luring the occupants to keep the windows open since they feel no discomfort.

Furthermore, the current approach fails in maintaining good Indoor Air Quality (IAQ), especially when there is a spike in occupancy in rooms with several occupants. In particular, higher ventilation rates by occupants above the reference model assumptions are penalized by the current system. Indeed, this drawback of the system also elicited negative responses from occupants during the experiment.

On the grounds of the foregoing discussion, the next version of Juracle would rather focus on the effect of ventilation on indoor air quality, specifically measured by the CO_2 concentration in the room, irrespective of how the ventilation was carried out. Here, energy wasted can then be estimated as a function of the duration during which the CO_2 concentration stayed below a given predefined lower threshold due to windows being open during the heating season, with the threshold indicating an expected healthy indoor CO_2 concentration for a room with closed windows. The proposed approach, for which details have been provided in Section 3.4.2, has the advantage of being more intuitive than the previous approach, in addition to having the potential for higher acceptability amongst users, since it is more user-centric – the health of the user's working environment is now the focus.

7.3.5 Communication of Energy Evaluations to Users

In the communication of the energy ratings for offices and teams to users, many "numbers" were involved, which was overwhelming for many users, especially those without a background related to energy. Besides the evaluation values, ventilation duration was expressed as three additional numbers – trickle ventilation minutes, shock ventilation minutes, and equivalent ventilation minutes. Indeed, previous studies show that users find it difficult to comprehend direct numbers (energy rating in kWh, CO_2 emissions in tonnes, etc.) (see e.g. the review of [167]). One mitigating strategy is to convert some of the numbers to graphics expressing the same ideas (so-called *eco-visualization*) [85, 168], e.g. showing trees to represent CO_2 emissions, or some chosen home appliances to represent energy consumption. A supporting idea is to show only one or two core numerical metrics, only revealing more numbers and details when the user explicitly requests them.

To add to the difficulties users experienced with numbers, the physical meaning of penalty values was not clear to many users as evidenced in the user survey results of Section 7.2. Two main flaws in the conceptualization of what numbers to display are responsible for this. Firstly, "penalties" – the excess consumption above an ideal demand – were displayed to the user as the energy evaluation of their room. However, users expected the energy evaluation numbers to rather be related to the total consumption of their offices. Therefore, for many users, zero penalties (in the correct interpretation) seemed to them (erroneously) to mean that their offices consumed no energy. This led to email exchanges between the author and some users, who were trying to understand the zero-demand rating of their offices. Secondly, the displayed energy penalties for the office and team were scaled to the size of the campus, but was still shown as being "associated" with the office/team, causing many users to believe that the penalty was indeed for their office alone. For some knowledgeable users, this led to confusion since the values were significantly higher than reasonable for a single office (the author wrote several emails and had at least one online video call in response to this confusion, in order to clarify the meaning of the numbers). Many other knowledgeable users just concluded that the values were wrong, indicating a broken evaluation system that was not worth investing any more attention in.

In a future version of the system, the general concept of penalizing wastage will be retained (subject to the revised methodology based on indoor air quality, as described in the preceding sub-section). However, instead of directly displaying this penalty to users as a stand-alone value scaled to the size of the campus, the displayed value should be the sum of the penalty and an estimate of the ideal consumption of the room in order to represent the total demand of the office itself. The derivation of such an ideal (baseline) demand estimate for the office would be relatively straightforward, since the reference model used to derive the penalties already embodies this information (also expressible as Energy Use Intensity, i.e. in kWh/m^2, and thus easily scalable).

7.3.6 Software and Hardware Testing and Validation Strategy

Many of the bugs that hampered the effectiveness of the developed applications at various levels could have been eliminated with better, more structured testing following software engineering best practices. It is more difficult to test web-based applications than desktop applications due to the additional complexities of the request-response cycle and the difficulty of testing the correctness of text-based HTML pages in response to requests. Nevertheless, there are frameworks like Selenium that aid in this process, which should be employed. Additionally, more backend tests can be performed to test and validate units of functionality.

Furthermore, the co-design strategy, in which volunteers gain early access to test new apps and features, should be increasingly adopted. In the developmental stages of the Energy Dashboard Suite that underwent testing by co-design participants, many bugs and user experience issues were discovered by the testers, and these were fixed prior to the main deployment. However, for some later phases of development, the co-design testing was skipped, resulting in poorer user acceptance overall than for the tested phases.

7.3.7 Performance of the Heating Controller

As already discussed in Section 5.1.5, there were issues with adequately heating Building B-01. The measures taken to improve the situation include setting a higher temperature setpoint than indicated by the occupant, according to the average shortfall by the controller; increasing the pre-heating period; and increasing the setback temperature, i.e. the fallback temperature when there are no occupants. Also, for particularly affected rooms, the pre-heating period could be higher than for other rooms. Also, several on-site inspections were carried out, especially to identify faulty sensors and actuators. These measures greatly reduced the comfort violations.

Additionally and crucially, as at the time of writing this thesis, a monitoring system has now been developed and integrated into JuControl, giving the responsible admins oversight of the performance of the heating controller in the offices having JuControl-assisted heating (i.e. for Buildings B-01 and B-02). The performance oversight provides both predictive diagnosis (based on current heating trend, but before occupants arrive in the affected office) and current status reports. Additionally, automated emails describing these faulty conditions are sent to admins in real-time, both for predicted comfort violations and for existing violations. Screenshots of the heatmap graphs providing a quick visual cue as to current and past performance of heating in the offices of Building B-01 are shown in Fig. 7.6 for different time periods. In the heatmaps, each row represents an office in the building, and each column a one-hour time period. The red colour represents an ongoing comfort violation (office is occupied at that time slot, but temperature setpoint is below target); yellow represents an imminent comfort violation (the office will soon be occupied, but the rate of temperature ramp-up does not seem sufficient to meet the target before the occupant arrives); and, green represents either an unoccupied office, or a correctly heated occupied office. The grey cells are indeterminate, meaning that there was not enough data to determine the status. In Fig. 7.6a, the status of the heating performance is shown for Building B-01 as at when the monitoring system was initially deployed, covering the working week of 0ct. 30 to Nov. 3, 2023, while in Fig. 7.6b, the heating performance is shown for working week of Dec. 11 to 15, 2023. As can be seen, there were less violations in the latter period, and all indeterminate states (grey cells) had been resolved by fixing or replacing faulty sensors and actuators before the end of the period.

In addition to the heatmap visualization, the monitoring system also features tables with detailed information regarding the heating status, including current room temperature, the start time of current or upcoming occupancy, the aggregated temperature preference of the current (if occupied) or next (if to be occupied later) occupants, and statistics regarding the current state of the rooms in the building, amongst others. Thus, the monitoring system enabled targeted resolution of the heating issues in real-time even beyond the initial general resolution measures described in the previous paragraph, thereby minimizing discomfort caused to occupants.

(**a**) Heating performance heatmap for Oct. 30 to Nov. 3, 2023

(**b**) Heating performance heatmap for Dec. 11 to 15, 2023

Fig. 7.6 Heating performance heatmap for Building B-01 for two periods, showing improvement in heating performance due to more accurate monitoring (Red: ongoing comfort violation, i.e. occupied office with below-target room temperature; Yellow: predicted violation, i.e. probably future violation of comfort; Green: correctly-heated occupied office, or unoccupied office; Gray: indeterminate due to missing data). In (a), the status of the building at the initial deployment of monitoring system are shown (Oct. 30 to Nov. 3, 2023), and (b) is the status after many of the pinpointed problem spots had been rectified a few weeks later (Dec. 11 to 15, 2023).

Chapter 8

Conclusion and Outlook

8.1 Conclusion

In this thesis, a set of software applications, tools, and methods targeting building occupants in public buildings was developed, with the overall goal of improving the energy efficiency of the occupants' behaviour within the buildings. These user behaviour interventions were developed based on the concepts of eco-visualization, control, gamification, and serious games, which have been shown in the literature to have a positive effect on user behaviour in general. The set of software applications, collectively called the Energy Dashboard Suite, include the *Campus Viewer*, which deals with visualization of energy consumption at the building and campus level; *JuControl*, which visualizes the indoor conditions and energy systems at the office level, facilitates automatic schedule-based heating, and provides gamification; *Juracle*, an engine for evaluating the energy efficiency of occupant behaviour; and *ALICE*, a tool that supports semi-automatic generation of interactive visualizations of occupant offices for embedding in JuControl.

These above applications were developed considering the three main research questions being addressed by the thesis, which can be rephrased succinctly as follows:

Q1: What is a systematic methodology for developing *fair* occupant energy-related behaviour evaluation system while taking into consideration occupant comfort and wellbeing?

Q2: Which gamification methodologies should be employed to enable occupants of public buildings, who have no financial incentive to be energy efficient, be motivated to become energy efficient without disrupting normal business functions?

Q3: How effective and efficient are such gamification-based interventions in terms of measurable change in behaviour and / or energy efficiency?

To address *Q1*, a novel framework and taxonomy for categorizing energy-related occupant behaviour evaluation methodologies was developed and characterized, called the Rule-Model-Measurement (RMM) framework. The framework categorizes behaviour evaluation systems into rule-based, model-based, measurement-based, or a mixture of these. Based on behaviour-analytic considerations, the development of the framework demonstrated the link between energy-related behaviour evaluation systems and user behaviour modification procedures, also called *operant conditioning*. Operant conditioning modes of *reinforcement* and *punishment* were analysed vis-a-vis the characteristics of the behaviour evaluation system and the compatibility of these modes with energy-related performance feedback. The framework was conceptually applied to the setting of the current study, highlighting input/output requirements, modelling effort, and strengths and weaknesses in specific use cases. For example, for an evaluation system that targets particular pre-defined user actions, a rule-based system is more appropriate, since it can provide the intended type of feedback even when the action has no immediate negative energy consequence. This contrasts with a model-based system that only judges user actions based on their real energy footprint, even when those actions are energy inefficient in the long term, but possibly efficient under niche conditions. As an illustration, consider an occupant who leaves the office window open overnight in autumn and goes home. This is generally an energy inefficient action, but during a warm autumn night where the heating is turned off, the energy impact might be negligible from a modelling perspective, but *as a rule* could be equivalent to doing so on a cold autumn night with the heating on.

On the other hand, the *Energy Dashboard Suite* implements the concepts of *eco-visualization*, *control* (i.e. integration into Building Automation System), *occupant behaviour evaluation*, and *gamification*, thus outlining the approach taken to address research question **Q2**. To deal with privacy, clear data privacy policies were negotiated with the responsible bodies, in addition to carefully selecting the type and granularity of instrumentation in other to achieve the thesis aims while not violating the agreed privacy policies. Furthermore, the implementation of the gamification software ensured seamless integration into the normal office routine of building occupants. For example, user survey results show that many occupants used JuControl CO_2 concentrations to know monitor indoor air quality in their offices. Also, the integration of JuControl into the Building Automation System ensured that JuControl became part of daily life in those buildings, since occupant schedules in JuControl were used for heating the offices. In fact, informal interactions showed that the pre-heating functionality was especially favoured by occupants.

In a seven-week experiment period from March to April 2023, the developed systems were holistically tested in a real-world setting, using selected buildings of the campus of Forschungszentrum Jülich to investigate the effectiveness of the developed systems and methods. During the experiment, buildings were divided into teams, with each team belonging to an experiment group with a predefined set of experimental variables enabled. By the end of the experiment period, almost 2000 employees had accessed the Energy Dashboard Suite at least once, and among the potential 870 employees who were part of the experiment by design, about 50% of them accessed JuControl before or within the experiment period. Nearly one-fifth of the potential 439 offices involved in the experiment were activated in JuControl by all occupants granting consent to the data processing terms. Considering that many willing users could not access JuControl due to the privacy policy, as discussed in Section 5.2, the level of engagement could even easily be higher. In fact, the new privacy policy as at the time of writing does *not* restrict access to JuControl, but only to CO_2 concentration when there is no consensus of consent, giving even more room for engagement.

The experiments demonstrated that the interventions had largely positive effects on occupant energy efficiency as reflected in ventilation styles and setpoint temperature, especially where the level of engagement with the developed systems was reasonably high. The mean daily energy penalties in the ventilation intervention group was 65% lower than that of its control group (1.66 kWh vs 4.67 kWh), with even lower penalties in the "activated" subgroup of the intervention group (0.74 kWh). In another test building that considered both ventilation and setpoint temperature, activated offices had 56% lower daily mean energy penalties than the control (1.91 kWh vs. 4.35 kWh), while in the pilot building, the energy penalties in the activated offices was 40% less than that of its control group (1.61 kWh vs. 2.94 kWh). All these effects were statistically significant and with large effect sizes. Furthermore, year-on-year thermal energy savings of about 18% (11.8 MWh) were realized in the pilot building where occupancy-driven heating was introduced. Accordingly, a preference for shock ventilation was adopted above trickle ventilation in line with the goals of the interventions, demonstrated by the predominant use of trickle ventilation for offices in the control group, as against the use of shock ventilation in the offices with interventions. Furthermore, the results demonstrated superior energy efficiency in JuControl-activated offices compared to non–JuControl-activated offices as a result of more efficient window ventilation styles in JuControl-activated offices than in non-activated offices. These results reflect the potential that the developed system has to improve energy efficiency when used. Nevertheless, an analysis of the use of the JuControl calendar for specifying planned presence indicates that the occupants did not generally update their calendars to reflect their planned presence, with a tendency towards overestimating scheduled presence in the office.

On the other hand, the level of engagement with the developed systems was shown through data and surveys to be dependent on the user acceptance of the developed systems, especially as a function of the perceived correctness of the behaviour evaluation results and the understandability of their presentation. In addition to the seven-week experiment, the results of enabling occupant control of heating via presence schedules and general sensitization in a pilot building showed that effective energy savings of 16.7% was achieved in the building over a full year compared to the baseline case, which was reasonably demonstrated to be as a result of the interventions.

Several issues were highlighted in the experiment methodology and the results analysis, and these are presented as recommendations for future work in the next section, where applicable. Additionally, due to the limited duration of the trial, the effect of the interventions on *intrinsic motivation* of the occupants to act energy-efficiently for which a more longitudinal study and repeated tests are required, could not be established.

8.2 Outlook

The work carried out in this thesis exposes ample opportunity for further research in user engagement and behaviour motivation in the energy sector. The most significant of these are highlighted in this section, along with implications for research and policy.

First, a longer experiment is planned to cover the entire 2024/25 winter season. The results of the upcoming experiment can be subsequently be compared with data from the preceding year at the building level, where no experiments were run for the same setup. By implication, the design of the experiment will permit the investigation of building-wide energy savings due to interventions.

Secondly, indoor air quality (IAQ) will be the major driver for ventilation recommendations in future versions of JuControl, without considering the window opening style (tilted or fully opened), where the goal is to maintain a healthy indoor CO_2 concentration. This approach is expected to be more acceptable by users since it is more occupant-focused. However, how the resulting energy demand compares with that of control groups without intervention is yet to be seen.

Thirdly, a serious game called *JuPower* will be integrated into JuControl. The *JuPower* game story focuses on making a virtual campus of Forschungszentrum Jülich more energy efficient. The integration into could be used as the primary eco-visualization for the occupants, showing the "greenness" of their respective virtual campuses as a function of their real-world energy efficiency.

Finally, innovative integration of Large Language Models (LLMs) through Generative Pre-trained Transformers (GPTs) are planned for JuControl, such that users can query their own data conversationally, as well as request insights into their historical performance visualizations.

References

[1] International Energy Agency. *Energy Technology RD&D Budgets 2020: Overview – Analysis - IEA*. Tech. rep. 2020. URL: https://www.iea.org/reports/energy-technology-rdd-budgets-2020.

[2] International Energy Agency. *Buildings*. 2021. URL: https://www.iea.org/topics/buildings (visited on 05/15/2020).

[3] IEA. *Final energy consumption in the buildings sector*. Tech. rep. Paris, 2021. URL: https://www.iea.org/data-and-statistics/charts/final-energy-consumption-in-the-buildings-sector-2021.

[4] OrbEEt. *D1.2 Specs of SEOR methodology and Enhanced Display Energy Certificates*. Tech. rep. 2017.

[5] European Commission. *Towards reaching the 20% energy efficiency target for 2020, and beyond*. 2017. URL: https://ec.europa.eu/commission/presscorner/detail/en/MEMO%7B%5C_%7D17%7B%5C_%7D162 (visited on 06/15/2020).

[6] European Commission. *COM(2016) 860 ANNEX 1: Accelerating clean energy in buildings*. 2016. URL: https://ec.europa.eu/energy/sites/ener/files/documents/1%7B%5C_%7Den%7B%5C_%7Dannexe%7B%5C_%7Dautre%7B%5C_%7Dacte%7B%5C_%7Dpart1%7B%5C_%7Dv9.pdf (visited on 06/15/2020).

[7] European Commission. *Energy efficiency directive*. 2018. URL: https://ec.europa.eu/energy/topics/energy-efficiency/targets-directive-and-rules/energy-efficiency-directive%7B%5C_%7Den (visited on 06/15/2020).

[8] Anna Carolina Menezes et al. "Predicted vs. actual energy performance of non-domestic buildings: Using post-occupancy evaluation data to reduce the performance gap". In: *Applied Energy* 97 (Sept. 2012), pp. 355–364. ISSN: 03062619. DOI: 10.1016/j.apenergy.2011.11.075. URL: http://linkinghub.elsevier.com/retrieve/pii/S0306261911007811%20http://dx.doi.org/10.1016/j.apenergy.2011.11.075.

[9] Karin Schakib-Ekbatan et al. "Does the occupant behavior match the energy concept of the building? - Analysis of a German naturally ventilated office building". In: *Building and Environment* 84 (Jan. 2015), pp. 142–150. ISSN: 03601323. DOI: 10.1016/j.buildenv.2014.10.018.

[10] H. Burak Gunay, William O'Brien, and Ian Beausoleil-Morrison. "A critical review of observation studies, modeling, and simulation of adaptive occupant behaviors in offices". PhD thesis. The University of Gloucestershire, Dec. 2013, pp. 31–47. DOI: 10.1016/J.BUILDENV.2013.07.020. URL: https://www.sciencedirect.com/science/article/pii/S0360132313002187.

[11] B. Bordass et al. "Assessing building performance in use 3: Energy performance of the Probe buildings". In: *Building Research and Information* 29.2 (2001), pp. 114–128. ISSN: 09613218. DOI: 10.1080/09613210010008036.

[12] Valentina Fabi et al. "Occupants' window opening behaviour: A literature review of factors influencing occupant behaviour and models". In: *Building and Environment* 58 (Dec. 2012), pp. 188–198. ISSN: 03601323. DOI: 10.1016/j.buildenv.2012.07.009. URL: http://dx.doi.org/10.1016/j.buildenv.2012.07.009.

[13] Tianzhen Hong and Hung-Wen Lin. *Occupant Behavior: impact on energy use of private offices*. Tech. rep. January. Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), 2012, p. 8. URL: https://escholarship.org/uc/item/6jp5w8kn%7B%5C#%7Dpage-11.

[14] Dong Zhao et al. "Interaction effects of building technology and resident behavior on energy consumption in residential buildings". In: *Energy and Buildings* 134 (2017), pp. 223–233. ISSN: 03787788. DOI: 10.1016/j.enbuild.2016.10.049. URL: http://dx.doi.org/10.1016/j.enbuild.2016.10.049.

[15] Andrea Conserva et al. “Selection and assessment of a set of 250 energy efficiency measures integrated in a serious game to promote citizens’ behaviour change towards energy efficiency in buildings”. In: *12th Conference on Sustainable Development of Energy, Water and Environment Systems* (2017), pp. 1–15.

[16] Dirk Börner et al. “Energy awareness displays: Motivating conservation at the workplace through feedback”. In: *International Journal of Mobile Learning and Organisation* (2012). ISSN: 1746725X. DOI: 10.1504/IJMLO.2012.050048.

[17] Marco Kalz et al. “Mindergie: A Pervasive Learning Game for Pro-environmental Behaviour at the Workplace”. In: *Seamless Learning in the Age of Mobile Connectivity*. Ed. by Lung-Hsiang Wong, Marcelo Milrad, and Marcus Specht. Singapore: Springer Singapore, 2015, pp. 397–417. ISBN: 978-981-287-113-8. DOI: 10.1007/978-981-287-113-8_20. URL: https://doi.org/10.1007/978-981-287-113-8%7B%5C_%7D20.

[18] Ann Bevitt and Miriam Wugmeister. *Comparing the U.S. and EU Approach to Employee Privacy*. Tech. rep. 2008. URL: https://www.jdsupra.com/legalnews/comparing-the-us-and-eu-approach-to-em-10096/.

[19] Dotmagazine Contributors. *Germany: Land of Data Protection and Security – But Why?* 2017. URL: https://www.dotmagazine.online/issues/security/germany-land-of-data-protection-and-security-but-why (visited on 07/06/2020).

[20] Tatjana Zrinski. *EU GDPR vs. German Bundesdatenschutzgesetz – Similarities and Differences*. URL: https://advisera.com/eugdpracademy/knowledgebase/eu-gdpr-vs-german-bundesdatenschutzgesetz-similarities-and-differences/ (visited on 07/06/2020).

[21] Olga Stepanova and Florian Groothuis. *The Privacy, Data Protection and Cybersecurity Law Review - Edition 6: Germany*. 2019. URL: https://thelawreviews.co.uk/edition/the-privacy-data-protection-and-cybersecurity-law-review-edition-6/1210039/germany (visited on 07/06/2020).

[22] Holger Lutz and Simoe Bach. “Employee Monitoring (Germany)”. In: *Practical Law* (2019).

[23] Jane S. Peters. *A brief history of behaviour in US energy programs: Landscape, integration, and future opportunities*. Elsevier Inc., 2020, pp. 479–503. ISBN: 9780128185674. DOI: 10.1016/b978-0-12-818567-4.00017-x. URL: https://doi.org/10.1016/B978-0-12-818567-4.00017-X.

[24] Julia Rotondo et al. *Overview of Existing Future Residential Use Cases for Connected Thermostats*. Tech. rep. US Department of Energy: Energy Efficiency & Renewable Energy, 2016. URL: https://info.ornl.gov/sites/publications/files/Pub72302.pdf.

[25] L. T. McCalley and G. J.H. Midden. “Computer based systems in household appliances: The study of eco-feedback as a tool for increasing conservation behavior”. In: *Proceedings - 3rd Asia Pacific Computer Human Interaction, APCHI 1998*. 1998, pp. 344–349. ISBN: 0818683473. DOI: 10.1109/APCHI.1998.704455.

[26] Jon Froehlich, Leah Findlater, and James Landay. “The design of eco-feedback technology”. In: *Conference on Human Factors in Computing Systems - Proceedings*. Vol. 3. May. 2010, pp. 1999–2008. ISBN: 9781605589299. DOI: 10.1145/1753326.1753629.

[27] Lennart E. Nacke and Sebastian Deterding. “The maturing of gamification research”. In: *Computers in Human Behavior* 71 (June 2017), pp. 450–454. ISSN: 07475632. DOI: 10.1016/j.chb.2016.11.062.

[28] Sebastian Deterding et al. “From game design elements to gamefulness: Defining "gamification"”. In: *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, MindTrek 2011*. 2011, pp. 9–15. ISBN: 9781450308168. DOI: 10.1145/2181037.2181040.

[29] Kai Huotari and Juho Hamari. “Defining gamification - A service marketing perspective”. In: *Proceedings of the 16th International Academic MindTrek Conference 2012: "Envisioning Future Media Environments", MindTrek 2012* (2012), pp. 17–22. DOI: 10.1145/2393132.2393137.

[30] Daniel Johnson et al. “Gamification and serious games within the domain of domestic energy consumption: A systematic review”. In: *Renewable and Sustainable Energy Reviews* 73.February 2016 (2017), pp. 249–264. ISSN: 18790690. DOI: 10.1016/j.rser.2017.01.134.

[31] Karen Glanz. “Current Theoretical Bases for Nutrition Intervention and Their Uses”. In: *Nutrition in the Prevention and Treatment of Disease*. Academic Press, Jan. 2001, pp. 83–93. DOI: 10.1016/b978-012193155-1/50008-8.

[32] M W Gallagher. *Self-Efficacy*. Ed. by V S B T - Encyclopedia of Human Behavior (Second Edition) Ramachandran. San Diego, 2012. DOI: https://doi.org/10.1016/B978-0-12-375000-6.00312-8. URL: http://www.sciencedirect.com/science/article/pii/B9780123750006003128.

[33] Daniela Pasini, Francesco Reda, and Tarja Häkkinen. "User engaging practices for energy saving in buildings: Critical review and new enhanced procedure". In: *Energy and Buildings* 148 (2017), pp. 74–88. ISSN: 03787788. DOI: 10.1016/j.enbuild.2017.05.010. URL: http://dx.doi.org/10.1016/j.enbuild.2017.05.010.

[34] Sebastian Deterding. *Paideia as Paidia: From Game-Based Learning to a Life Well-Played*. 2012. URL: https://codingconduct.cc/Paideia-as-Paidia.

[35] Jonna Koivisto and Juho Hamari. "The rise of motivational information systems: A review of gamification research". In: *International Journal of Information Management* 45.October 2018 (2019), pp. 191–210. ISSN: 02684012. DOI: 10.1016/j.ijinfomgt.2018.10.013. URL: https://doi.org/10.1016/j.ijinfomgt.2018.10.013.

[36] Spartaco Albertarelli et al. "A survey on the design of gamified systems for energy and water sustainability". In: *Games* 9.3 (2018), pp. 1–34. ISSN: 20734336. DOI: 10.3390/g9030038.

[37] Icek Ajzen. "The Theory of Planned Behavior". In: *Organizational Behavior and Human Decision Processe* 50 (1991), pp. 179–211.

[38] Richard M. Ryan and Edward L. Deci. "Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being". In: *American Psychologist* 55.1 (2000), pp. 68–78. ISSN: 0003066X. DOI: 10.1037/0003-066X.55.1.68.

[39] Ramy Hammady and Sylvester Arnab. "Serious Gaming for Behaviour Change: A Systematic Review". In: *Information (Switzerland)* 13.3 (2022). ISSN: 20782489. DOI: 10.3390/info13030142.

[40] Scott Nicholson. "A RECIPE for meaningful gamification". In: *Gamification in Education and Business*. Springer International Publishing, Jan. 2015, pp. 1–20. ISBN: 9783319102085. DOI: 10.1007/978-3-319-10208-5_1.

[41] Juho Hamari, Jonna Koivisto, and Harri Sarsa. "Does gamification work? - A literature review of empirical studies on gamification". In: *Proceedings of the Annual Hawaii International Conference on System Sciences*. IEEE Computer Society, 2014, pp. 3025–3034. ISBN: 9781479925049. DOI: 10.1109/HICSS.2014.377.

[42] John C. Norcross, Paul M. Krebs, and James O. Prochaska. "Stages of Change". In: *Journal of Clinical Psychology* 67.2 (Feb. 2011), pp. 143–154. ISSN: 00219762. DOI: 10.1002/jclp.20758.

[43] J.O. Prochaska. *Transtheoretical model of behavior change*. 2013.

[44] David W. Eby, Lisa J. Molnar, and Joseph M. Pellerito. "Driving cessation and alternative community mobility". In: *Driver Rehabilitation and Community Mobility*. Mosby Inc., Jan. 2006, pp. 425–454. ISBN: 9780323029377. DOI: 10.1016/B978-032302937-7.50027-7.

[45] Kenneth R. Ginsburg. "Approaching Youth Violence in a Clinical Setting". In: *Adolescent Medicine*. Elsevier, Jan. 2008, pp. 28–33. DOI: 10.1016/b978-032304073-0.10005-6.

[46] Katie Seaborn and Deborah I. Fels. "Gamification in theory and action: A survey". In: *International Journal of Human Computer Studies* 74 (2015), pp. 14–31. ISSN: 10959300. DOI: 10.1016/j.ijhcs.2014.09.006.

[47] Devon Wemyss et al. "Does it last? Long-term impacts of an app-based behavior change intervention on household electricity savings in Switzerland". In: *Energy Research and Social Science* 47.August 2018 (2019), pp. 16–27. ISSN: 22146296. DOI: 10.1016/j.erss.2018.08.018. URL: https://doi.org/10.1016/j.erss.2018.08.018.

[48] J. Novak et al. "Integrating behavioural change and gamified incentive modelling for stimulating water saving". In: *Environmental Modelling and Software* 102 (2018), pp. 120–137. ISSN: 13648152. DOI: 10.1016/j.envsoft.2017.11.038. URL: https://doi.org/10.1016/j.envsoft.2017.11.038.

[49] James O. Prochaska, John C. Norcross, and Carlo C. DiClemente. "Applying the Stages of Change". In: *Psychotherapy in Australia* 19.2 (2013), pp. 177–181. DOI: 10.1093/med:psych/9780199845491.003.0034.

[50] Eva Villegas et al. "Design Thinking and Gamification: User Centered Methodologies". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 11590 LNCS (2019), pp. 115–124. ISSN: 16113349. DOI: 10.1007/978-3-030-21814-0_10.

[51] J. Dormans. “Engineering emergence: applied theory for game design”. PhD thesis. Universiteit van Amsterdam, 2012. ISBN: 9789461907523.

[52] Frederick Grossberg et al. *Gamified Energy Efficiency Programs*. Tech. rep. February. 2015, p. 73. URL: http://aceee.org/research-report/b1501.

[53] Robin Hunicke, Marc Leblanc, and Robert Zubek. “MDA: A formal approach to game design and game research”. In: *AAAI Workshop - Technical Report* (2004), pp. 1–5.

[54] Wolfgang Walk, Daniel Görlich, and Mark Barrett. “Design, Dynamics, Experience (DDE): An Advancement of the MDA Framework for Game Design”. In: *Game dynamics: Best practices in procedural and dynamic game content generation*. Ed. by Oliver Korn and Newton Lee. Springer, 2017. ISBN: 9783319530871. DOI: 10.1007/978-3-319-53088-8.

[55] Flavio Escribano. “Gamification Model Canvas Evolution for Design Improvement: Player Profiling and Decision Support Models”. In: *Fundación Iberoamericana del Conocimiento* (2017), pp. 1–6. URL: https://gecon.es/wp-content/uploads/2017/07/GMC-Evolution%7B%5C_%7DvDef.pdf.

[56] Dave Russell. *Video game user interface design: Diegesis theory*. 2011. URL: http://devmag.org.za/2011/02/02/video-game-user-interface-design-diegesis-theory/ (visited on 02/12/2024).

[57] Gerben Grave. *Emergent narratives in games*. 2015. URL: https://multiverse-narratives.com/2015/05/07/emergent-narratives-in-games/ (visited on 02/12/2024).

[58] Tarek AlSkaif et al. “Gamification-based framework for engagement of residential customers in energy applications”. In: *Energy Research and Social Science* 44.April (2018), pp. 187–195. ISSN: 22146296. DOI: 10.1016/j.erss.2018.04.043. URL: https://doi.org/10.1016/j.erss.2018.04.043.

[59] Brian Orland et al. “Saving energy in an office environment: A serious game intervention”. In: *Energy and Buildings* 74 (May 2014), pp. 43–52. ISSN: 03787788. DOI: 10.1016/j.enbuild.2014.01.036.

[60] S3C Consortium. *S3C Smartgrid Engagement Toolkit*. 2015. URL: https://www.smartgrid-engagement-toolkit.eu/home/ (visited on 06/09/2020).

[61] S3C Consortium. *TOOL: WEB-BASED ENERGY QUIZ*. Tech. rep. URL: http://www.smartgrid-engagement-toolkit.eu/exploring/incentives-.

[62] S3C Consortium. *GUIDELINE: USING SEGMENTATION TO BETTER TARGET USER GROUPS*. Tech. rep. 2015.

[63] Bernadette Sütterlin, Thomas A. Brunner, and Michael Siegrist. “Who puts the most energy into energy conservation? A segmentation of energy consumers based on energy-related behavioral characteristics”. In: *Energy Policy* 39.12 (2011), pp. 8137–8152. ISSN: 03014215. DOI: 10.1016/j.enpol.2011.10.008.

[64] S3C Consortium. *LIST OF GUIDELINES AND TOOLS*. URL: https://www.smartgrid-engagement-toolkit.eu/learning/list-of-guidelines-and-tools/.

[65] Ioannis C. Konstantakopoulos et al. “A deep learning and gamification approach to improving human-building interaction and energy efficiency in smart infrastructure”. In: *Applied Energy* 237.September 2018 (2019), pp. 810–821. ISSN: 03062619. DOI: 10.1016/j.apenergy.2018.12.065. URL: https://doi.org/10.1016/j.apenergy.2018.12.065.

[66] European Commission. *Horizon 2020*. URL: https://ec.europa.eu/programmes/horizon2020/en.

[67] OrbEEt. *Organizational behaviour improvement for energy efficient administrative public offices (OrbEEt)*. 2017. URL: http://orbeet.eu/.

[68] OrbEEt. *OrbEEt in Erlangen*. Tech. rep. 2017. URL: http://orbeet.eu/wp-content/uploads/2018/04/OrbEEt%7B%5C_%7DDemo%7B%5C_%7DGermany%7B%5C_%7DFinal.pdf.

[69] Universitat Politecnica de Catalunya. *Energy Game for Awareness of energy efficiency in social housing communities*. 2018. URL: http://www.energaware.eu/.

[70] EnerGAware. *Energy Cat: The House of Tomorrow*. URL: http://www.energycatgame.com/.

[71] Miquel Casals et al. “Assessing the effectiveness of gamification in reducing domestic energy consumption: Lessons learned from the EnerGAware project”. In: *Energy and Buildings* 210 (2020), p. 109753. ISSN: 03787788. DOI: 10.1016/j.enbuild.2019.109753. URL: https://doi.org/10.1016/j.enbuild.2019.109753.

[72] TRIBE consortium. *TRIBE*. 2017. URL: http://tribe-h2020.eu/.

[73] TRIBE consortium. *Deliverable 7.1 Game Vision*. Tech. rep. 2016. URL: http://tribe-h2020.eu/wp-content/uploads/2017/03/TRIBE%7B%5C_%7DDLV-7.1-Game%7B%5C_%7DVision-FV2.pdf.

[74] Thanasis G Papaioannou et al. "IoT-enabled gamification for energy conservation in public buildings". In: *GIoTS 2017 - Global Internet of Things Summit*. 2017. ISBN: 9781509058730. DOI: 10.1109/GIOTS.2017.8016269.

[75] Huijeong Kim et al. "MySmartE – An eco-feedback and gaming platform to promote energy conserving thermostat-adjustment behaviors in multi-unit residential buildings". In: *Building and Environment* 221.May (2022), p. 109252. ISSN: 03601323. DOI: 10.1016/j.buildenv.2022.109252. URL: https://doi.org/10.1016/j.buildenv.2022.109252.

[76] Thanasis G Papaioannou et al. "An IoT-based gamified approach for reducing occupants' energy wastage in public buildings". In: *Sensors (Switzerland)* 18.2 (2018). ISSN: 14248220. DOI: 10.3390/s18020537.

[77] José Iria et al. "A gamification platform to foster energy efficiency in office buildings". In: *Energy and Buildings* 222 (2020). ISSN: 03787788. DOI: 10.1016/j.enbuild.2020.110101.

[78] Thanasis G. Papaioannou et al. "A sensor-enabled rule engine for changing energy-wasting behaviours in public buildings". In: *2018 IEEE International Energy Conference, ENERGYCON 2018* (2018), pp. 1–6. DOI: 10.1109/ENERGYCON.2018.8398784.

[79] Piero Fraternali et al. "A Socio-Technical System Based on Gamification Towards Energy Savings". In: *2018 IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom Workshops 2018* (2018), pp. 59–64. DOI: 10.1109/PERCOMW.2018.8480405.

[80] Devon Wemyss et al. "Examining community-level collaborative vs. competitive approaches to enhance household electricity-saving behavior". In: *Energy Efficiency* 11.8 (2018), pp. 2057–2075. ISSN: 15706478. DOI: 10.1007/s12053-018-9691-z.

[81] P.M. Johnson et al. "Beyond kWh: Myths and fixes for energy competition game design". In: *Proceedings of Meaningful Play* (2012), pp. 1–10. URL: http://csdl.ics.hawaii.edu/techreports/12-08/12-08.pdf.

[82] Douglas A Johnson. "A Component Analysis of the Impact of Evaluative and Objective Feedback on Performance". In: *Journal of Organizational Behavior Management* 33.2 (2013), pp. 89–103. ISSN: 1540-8604. DOI: 10.1080/01608061.2013.785879. URL: https://www.tandfonline.com/action/journalInformation?journalCode=worg20.

[83] Kathleen A Mangiapanello and Nancy S Hemmes. "An Analysis of Feedback from a Behavior Analytic Perspective". In: *Behavior Analyst* 38.1 (2015), pp. 51–75. ISSN: 21968918. DOI: 10.1007/s40614-014-0026-x.

[84] Cyrus Wahome. *What Is Operant Conditioning?* 2022. URL: https://www.webmd.com/mental-health/what-is-operant-conditioning.

[85] Kyle Anderson and Sang Hyun Lee. "An empirically grounded model for simulating normative energy use feedback interventions". In: *Applied Energy* 173 (2016), pp. 272–282. ISSN: 03062619. DOI: 10.1016/j.apenergy.2016.04.063. URL: http://dx.doi.org/10.1016/j.apenergy.2016.04.063.

[86] Richard M Tetlow et al. "Simple prompts reduce inadvertent energy consumption from lighting in office buildings". In: *Building and Environment* 81 (2014), pp. 234–242. ISSN: 03601323. DOI: 10.1016/j.buildenv.2014.07.003. URL: http://dx.doi.org/10.1016/j.buildenv.2014.07.003.

[87] Rishee K. Jain, John E. Taylor, and Gabriel Peschiera. "Assessing eco-feedback interface usage and design to drive energy efficiency in buildings". In: *Energy and Buildings* 48 (2012), pp. 8–17. ISSN: 03787788. DOI: 10.1016/j.enbuild.2011.12.033. URL: http://dx.doi.org/10.1016/j.enbuild.2011.12.033.

[88] Joanne R. Smith et al. "Congruent or conflicted? The impact of injunctive and descriptive norms on environmental intentions". In: *Journal of Environmental Psychology* 32.4 (2012), pp. 353–361. ISSN: 02724944. DOI: 10.1016/j.jenvp.2012.06.001.

[89] Hunt Allcott. "Social norms and energy conservation". In: *Journal of Public Economics* 95.9-10 (2011), pp. 1082–1095. ISSN: 00472727. DOI: 10.1016/j.jpubeco.2011.03.003. URL: http://dx.doi.org/10.1016/j.jpubeco.2011.03.003.

[90] Robert B Cialdini. "Crafting Normative Messages to Protect the Environment". In: *Current Directions in Psychological Science* 12.4 (2003), pp. 105–110.

[91] P Wesley Schultz et al. "The Constructive, Destructive, and Reconstructive Power of Social Norms". In: *Psychological Science* 18.5 (May 2007), pp. 429–434. ISSN: 0956-7976. DOI: 10.1111/j.1467-9280.2007.01917.x. URL: https://doi.org/10.1111/j.1467-9280.2007.01917.x.

[92] J E R Staddon and D T Cerutti. "Operant conditioning." eng. In: *Annual review of psychology* 54 (2003), pp. 115–144. ISSN: 0066-4308 (Print). DOI: 10.1146/annurev.psych.54.101601.145124.

[93] Lumen Learning. *Reinforcement and Punishment.* URL: https://pressbooks.online.ucf.edu/lumenpsychology/chapter/operant-conditioning/.

[94] Konstantinos Tsatsakis, Anastasios Tsitsanis, and Christos Malavazos. "Enhanced human-centric building energy performance rating: Operational rating and behavioral change in the scope of OrbEEt project". In: *2017 Global Internet of Things Summit (GIoTS)*. 2017, pp. 1–6. DOI: 10.1109/GIOTS.2017.8016268.

[95] Stuart O Connor et al. *Developing Gamified Elements to Influence Positive Behavioural Change towards Organisational Energy Efficiency*. Tech. rep. 2017.

[96] Christos Malavazos, Konstantinos Tsatsakis, and Anastasios Tsitsanis. "Future Demand Side Management Practices driven by Ambient Services and Consumer Profiling". In: (2014).

[97] Ray Galvin. "Impediments to energy-efficient ventilation of German dwellings: A case study in Aachen". In: *Energy and Buildings* 56 (2013), pp. 32–40. ISSN: 03787788. DOI: 10.1016/j.enbuild.2012.10.020. URL: http://dx.doi.org/10.1016/j.enbuild.2012.10.020.

[98] N Hansmeier and E Matthies. *Energiebewusste RUB–Richtig Heizen und Lüften. Ergebnisse einer umweltpsychologischen Intervention zur Förderung energieeffizienten Verhaltens an der Ruhr-Universität Bochum.* 2007.

[99] Stadt Münster. *Richtiges Heizen für Münsters Klimaneutralität.* 2020. URL: https://www.muenster.de/pressemeldungen/web/frontend/output/index.php?offset=/search/1/design/standard/page/1582/show/1055831 (visited on 02/15/2024).

[100] Die Landesregierung Nordrhein-Westfalen. *Minister Krischer: Gesundheitsschutz hat Vorrang.* 2022. URL: https://www.land.nrw/pressemitteilung/minister-krischer-gesundheitsschutz-hat-vorrang (visited on 02/15/2024).

[101] Verbraucherzentrale NRW. *Heizen und Lüften: So geht's richtig.* 2023. URL: https://www.verbraucherzentrale.nrw/wissen/energie/heizen-und-warmwasser/heizen-und-lueften-so-gehts-richtig-10426.

[102] W. Richter et al. *Bestimmung des realen Luftwechsels bei Fensterlüftung aus energetischer und bauphysikalischer Sicht.* Tech. rep. Technische Universität Dresden, 2003.

[103] Thanasis G. Papaioannou and George D. Stamoulis. "Teaming and competition for demand-side management in office buildings". In: *2017 IEEE International Conference on Smart Grid Communications, SmartGridComm 2017.* Vol. 2018-Janua. Institute of Electrical and Electronics Engineers Inc., July 2017, pp. 332–337. ISBN: 9781538640555. DOI: 10.1109/SmartGridComm.2017.8340734.

[104] Laura Marín-Restrepo, Maureen Trebilcock, and Mark Gillott. "Occupant action patterns regarding spatial and human factors in office environments". In: *Energy and Buildings* 214 (2020). ISSN: 03787788. DOI: 10.1016/j.enbuild.2020.109889.

[105] Philipp Althaus et al. "Cloud-based controller architecture for the testing of conventional and model predictive room heating controllers in a real-world environment". In: *2024 10th International Conference on Control, Decision and Information Technologies (CoDIT).* 10th International Conference on Control, Decision and Information Technologies, Valletta (Malta), 1 Jul 2024 - 4 Jul 2024. IEEE, July 2024, pp. 1792–1797. ISBN: 979-8-3503-7397-4. DOI: 10.1109/CoDIT62066.2024.10708077. URL: https://juser.fz-juelich.de/record/1029191%20https://ieeexplore.ieee.org/document/10708077/.

[106] Gail M Sullivan and Richard Feinn. "Using Effect Size-or Why the P Value Is Not Enough." eng. In: *Journal of graduate medical education* 4.3 (Sept. 2012), pp. 279–282. ISSN: 1949-8349 (Print). DOI: 10.4300/JGME-D-12-00156.1.

[107] Nelson Fumo. "A review on the basics of building energy estimation". In: *Renewable and Sustainable Energy Reviews* 31 (2014), pp. 53–60. ISSN: 13640321. DOI: 10.1016/j.rser.2013.11.040. URL: http://dx.doi.org/10.1016/j.rser.2013.11.040.

[108] ASHRAE. "2009 ASHRAE Handbook. Fundamentals. SI edition." In: (2009).

[109] Efficiency Valuation Organization. "IPMVP Core Concepts 2016". In: October (2016). URL: https://evo-world.org/en/products-services-mainmenu-en/protocols/ipmvp.

[110] U.S. Department of Energy. *M&V Guidelines: Measurement and Verification for Performance-Based Contracts.* Tech. rep. November. 2015.

[111] Tony Day. *Degree-days: theory and application.* London: The Chartered Institution of Building Services Engineers, 2006, p. 106. ISBN: 9781903287767.

[112] John Kelly Kissock, Jeff S. Haberl, and David E. Claridge. “Inverse modeling toolkit: Numerical algorithms”. In: *ASHRAE Transactions* 109 PART 2 (2003), pp. 425–434. ISSN: 00012505.

[113] Alice Crespi et al. *Climate-related hazard indices for Europe.* Tech. rep. 2020, p. 64. DOI: 10.25424/cmcc/climate. URL: https://www.eionet.europa.eu/etcs/etc-cca/products/etc-cca-reports/climate-related-hazard-indices-for-europe.

[114] Degree Days.net. *Calculating Degree Days.* URL: https://www.degreedays.net/calculation.

[115] Charles R Harris et al. “Array programming with NumPy”. In: *Nature* 585.7825 (Sept. 2020), pp. 357–362. DOI: 10.1038/s41586-020-2649-2. URL: https://doi.org/10.1038/s41586-020-2649-2.

[116] Pauli Virtanen et al. “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”. In: *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

[117] Wen Kuei Chang and Tianzhen Hong. “Statistical analysis and modeling of occupancy patterns in open-plan offices using measured lighting-switch data”. In: *Building Simulation* 6.1 (2013), pp. 23–32. ISSN: 19963599. DOI: 10.1007/s12273-013-0106-y.

[118] Christopher M. Stoppel and Fernanda Leite. “Integrating probabilistic methods for describing occupant presence with building energy simulation models”. In: *Energy & Buildings* 68 (2014), pp. 99–107. ISSN: 03787788. DOI: 10.1016/j.enbuild.2013.08.042. URL: http://dx.doi.org/10.1016/j.enbuild.2013.08.042.

[119] J. Page et al. “A generalised stochastic model for the simulation of occupant presence”. In: *Energy and Buildings* 40.2 (Jan. 2008), pp. 83–98. ISSN: 03787788. DOI: 10.1016/j.enbuild.2007.01.018. URL: https://linkinghub.elsevier.com/retrieve/pii/S037877880700031X.

[120] Philip Delff Andersen et al. “Dynamic modeling of presence of occupants using inhomogeneous Markov chains”. In: *Energy and Buildings* 69 (2014), pp. 213–223. ISSN: 03787788. DOI: 10.1016/j.enbuild.2013.10.001.

[121] Ian Richardson, Murray Thomson, and David Infield. “A high-resolution domestic building occupancy model for energy demand simulations”. In: *Energy and Buildings* 40.8 (2008), pp. 1560–1566. ISSN: 03787788. DOI: 10.1016/j.enbuild.2008.02.006.

[122] Thiago Teixeira, Gershon Dublon, and Andreas Savvides. *A survey of human-sensing: Methods for detecting presence, count, location, track, and identity.ENALAB Technical Report Yale University.* Tech. rep. 1. Yale University, 2010, pp. 1–41. URL: https://www.researchgate.net/publication/319791520.

[123] Da Yan et al. “Occupant behavior modeling for building performance simulation: Current state and future challenges”. In: *Energy and Buildings* 107 (2015), pp. 264–278. ISSN: 03787788. DOI: 10.1016/j.enbuild.2015.08.032. URL: http://dx.doi.org/10.1016/j.enbuild.2015.08.032.

[124] Da Yan et al. “IEA EBC Annex 66: Definition and simulation of occupant behavior in buildings”. In: *Energy and Buildings* 156 (2017), pp. 258–270. ISSN: 03787788. DOI: 10.1016/j.enbuild.2017.09.084. URL: http://dx.doi.org/10.1016/j.enbuild.2017.09.084.

[125] Ebenezer Hailemariam et al. “Real-time occupancy detection using decision trees with multiple sensor types”. In: *Procs. of the 2011 Symposium on Simulation for Architecture and Urban Design* January (2011), pp. 141–148.

[126] Huang Chia Shih. “A robust occupancy detection and tracking algorithm for the automatic monitoring and commissioning of a building”. In: *Energy and Buildings* 77.May 2013 (2014), pp. 270–280. ISSN: 03787788. DOI: 10.1016/j.enbuild.2014.03.069. URL: http://dx.doi.org/10.1016/j.enbuild.2014.03.069.

[127] J. Ahmad et al. “Occupancy detection in non-residential buildings – A survey and novel privacy preserved occupancy monitoring solution”. In: *Applied Computing and Informatics* xxxx (2018), pp. 1–9. ISSN: 22108327. DOI: 10.1016/j.aci.2018.12.001. URL: https://doi.org/10.1016/j.aci.2018.12.001.

[128] Ardeshir Mahdavi and Farhang Tahmasebi. “Predicting people’s presence in buildings: An empirically based model performance analysis”. In: *Energy and Buildings* 86 (2015), pp. 349–355. ISSN: 03787788. DOI: 10.1016/j.enbuild.2014.10.027. URL: http://dx.doi.org/10.1016/j.enbuild.2014.10.027.

[129] Guillaume Ansanay-Alex. “Estimating occupancy using indoor carbon dioxide concentrations only in an office building: a method and qualitative assessment”. In: *REHVA World Congress on Energy efficient, smart and healthy buildings (CLIMA)* April (2013), pp. 1–8.

[130] Davide Calì et al. “CO2 based occupancy detection algorithm: Experimental analysis and validation for office and residential buildings”. In: *Building and Environment* 86 (Apr. 2015), pp. 39–49. ISSN: 03601323. DOI: 10.1016/j.buildenv.2014.12.011. URL: http://dx.doi.org/10.1016/j.buildenv.2014.12.011%20https://linkinghub.elsevier.com/retrieve/pii/S0360132314004223.

[131] Luis M. Candanedo, Véronique Feldheim, and Dominique Deramaix. “A methodology based on Hidden Markov Models for occupancy detection and a case study in a low energy residential building”. In: *Energy and Buildings* 148 (Aug. 2017), pp. 327–341. ISSN: 03787788. DOI: 10.1016/j.enbuild.2017.05.031. URL: https://linkinghub.elsevier.com/retrieve/pii/S0378778817304152.

[132] S Wang and X Jin. “CO₂-Based Occupancy Detection for On-Line Outdoor Air Flow Control”. In: *Indoor and Built Environment* 7.3 (1998), pp. 165–181. ISSN: 1420-326X. DOI: 10.1159/000024577. URL: https://www.karger.com/DOI/10.1159/000024577.

[133] Andrzej Szczurek, Monika Maciejewska, and Tomasz Pietrucha. “Occupancy determination based on time series of CO2 concentration, temperature and relative humidity”. In: *Energy and Buildings* 147 (July 2017), pp. 142–154. ISSN: 03787788. DOI: 10.1016/j.enbuild.2017.04.080. URL: https://linkinghub.elsevier.com/retrieve/pii/S0378778817302396.

[134] Theis Heidmann Pedersen, Kasper Ubbe Nielsen, and Steffen Petersen. “Method for room occupancy detection based on trajectory of indoor climate sensor data”. In: *Building and Environment* 115 (Apr. 2017), pp. 147–156. ISSN: 03601323. DOI: 10.1016/j.buildenv.2017.01.023. URL: https://linkinghub.elsevier.com/retrieve/pii/S0360132317300367.

[135] Khee Poh Lam et al. “Occupancy detection through an extensive environmental sensor network in an open-plan office building”. In: *IBPSA 2009 - International Building Performance Simulation Association 2009* (2009), pp. 1452–1459. URL: https://pdfs.semanticscholar.org/8ee8/e6650e2b22e8dbd3ec7ba75971d53bc pdf.

[136] Zhenghua Chen, Mustafa K. Masood, and Yeng Chai Soh. “A fusion framework for occupancy estimation in office buildings based on environmental sensor data”. In: *Energy and Buildings* 133 (2016), pp. 790–798. ISSN: 03787788. DOI: 10.1016/j.enbuild.2016.10.030. URL: http://dx.doi.org/10.1016/j.enbuild.2016.10.030.

[137] Han Zou et al. “Non-intrusive occupancy sensing in commercial buildings”. In: *Energy and Buildings* 154 (2017), pp. 633–643. ISSN: 03787788. DOI: 10.1016/j.enbuild.2017.08.045. URL: https://doi.org/10.1016/j.enbuild.2017.08.045.

[138] Manar Amayri et al. “Estimating occupancy in heterogeneous sensor environment”. In: *Energy and Buildings* 129 (2016), pp. 46–58. ISSN: 03787788. DOI: 10.1016/j.enbuild.2016.07.026. URL: http://dx.doi.org/10.1016/j.enbuild.2016.07.026.

[139] Anindya Nag and Subhas Chandra Mukhopadhyay. “Occupancy Detection at Smart Home Using Real-Time Dynamic Thresholding of Flexiforce Sensor”. In: *IEEE Sensors Journal* 15.8 (2015), pp. 4457–4463. ISSN: 1530437X. DOI: 10.1109/JSEN.2015.2421348.

[140] Wilhelm Kleiminger et al. “Occupancy Detection from Electricity Consumption Data”. In: (2013), pp. 1–8. DOI: 10.1145/2528282.2528295.

[141] Felix Nienaber et al. “Validation, optimisation and comparison of carbon dioxide-based occupancy estimation algorithms”. In: *Indoor and Built Environment* 29.6 (2020), pp. 820–834. ISSN: 14230070. DOI: 10.1177/1420326X19871332.

[142] Mattias Gruber, Anders Trüschel, and Jan Olof Dalenbäck. “CO2 sensors for occupancy estimations: Potential in building automation applications”. In: *Energy and Buildings* 84 (2014), pp. 548–556. ISSN: 03787788. DOI: 10.1016/j.enbuild.2014.09.002. URL: http://dx.doi.org/10.1016/j.enbuild.2014.09.002.

[143] Chaoyang Jiang et al. “Indoor occupancy estimation from carbon dioxide concentration”. In: *Energy and Buildings* 131 (2016), pp. 132–141. ISSN: 03787788. DOI: 10.1016/j.enbuild.2016.09.002. arXiv: 1607.05962. URL: http://dx.doi.org/10.1016/j.enbuild.2016.09.002.

[144] Mustafa K. Masood, Yeng Chai Soh, and Victor W.C. Chang. “Real-time occupancy estimation using environmental parameters”. In: *Proceedings of the International Joint Conference on Neural Networks* 2015-Septe (2015), pp. 1–8. DOI: 10.1109/IJCNN.2015.7280781.

[145] Zhenghua Chen, Chaoyang Jiang, and Lihua Xie. “Building occupancy estimation and detection: A review”. In: *Energy and Buildings* 169 (2018), pp. 260–270. ISSN: 03787788. DOI: 10.1016/j.enbuild.2018.03.084.

[146] Florian Redder et al. “Information and Communication Technologies (ICT) for Holistic Building Energy System Operation in Living Labs: Conceptualization, Implementation, Evaluation [under review]”. In: *Applied Energy* (2024).

[147] Peter Fritzson. “Modelica—A language for equation-based physical modeling and high performance simulation BT - Applied Parallel Computing Large Scale Scientific and Industrial Problems”. In: ed. by Bo Kågström et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 1998, pp. 149–160. ISBN: 978-3-540-49261-0.

[148] D. Müller et al. “AixLib - An Open-Source Modelica Library within the IEA-EBC Annex 60 Framework”. In: *BauSIM 2016*. 2016.

[149] Dassault Systèmes. *Dymola*. 2020. URL: https://www.3ds.com/products/catia/dymola (visited on 01/15/2024).

[150] W. Richter et al. *Determination of the Actual Air Change Rate for Natural Ventilation with Regard to Energy Efficiency and Building Physics*. Tech. rep. Technische Universität Dresden, 2003.

[151] SciPy 1.0 Contributors. *scipy.interpolate.CloughTocher2DInterpolator*. URL: https://docs.scipy.org/doc/scipy/reference/generated/scipy.interpolate.CloughTocher2DInterpolator.html (visited on 01/06/2023).

[152] Arbeitssischerheit.de. *Richtig lüften*. 2011. URL: https://www.arbeitssicherheit.de/themen/sonstiges/detail/richtig-lueften.html (visited on 02/15/2024).

[153] Richard J. de Dear and Gail Schiller Brager. “Developing an Adaptive Model of Thermal Comfort and Preference”. In: *ASHRAE Transactions* 104.Pt 1A (1998), pp. 145–167. ISSN: 00012505.

[154] Federico Tartarini et al. “CBE Thermal Comfort Tool: Online tool for thermal comfort calculations and visualizations”. In: *SoftwareX* 12 (2020), p. 100563. ISSN: 23527110. DOI: 10.1016/j.softx.2020.100563. URL: https://doi.org/10.1016/j.softx.2020.100563.

[155] Federal Cabinet of the German Federal Government. *Additional energy saving measures approved*. 2022. URL: https://www.bundesregierung.de/breg-en/issues/energy-saving-measures-2078510 (visited on 09/12/2022).

[156] Eziama Ubachukwu et al. “User engagement for thermal energy-efficient behavior in office buildings using dashboards and gamification”. In: *Applied Thermal Engineering* 266 (2025), p. 125598. ISSN: 1359-4311. DOI: https://doi.org/10.1016/j.applthermaleng.2025.125598. URL: https://www.sciencedirect.com/science/article/pii/S1359431125001899.

[157] F Pedregosa et al. “Scikit-learn: Machine Learning in Python”. In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[158] Abigail Francisco et al. “Occupant perceptions of building information model-based energy visualizations in eco-feedback systems”. In: *Applied Energy* 221.December 2017 (2018), pp. 220–228. ISSN: 03062619. DOI: 10.1016/j.apenergy.2018.03.132. URL: https://doi.org/10.1016/j.apenergy.2018.03.132.

[159] Nick Verkade and Johanna Höffken. “Is the Resource Man coming home? Engaging with an energy monitoring platform to foster flexible energy consumption in the Netherlands”. In: *Energy Research and Social Science* 27 (2017), pp. 36–44. ISSN: 22146296. DOI: 10.1016/j.erss.2017.02.015. URL: http://dx.doi.org/10.1016/j.erss.2017.02.015.

[160] Stephanie N. Timm and Brian M. Deal. “Effective or ephemeral? the role of energy information dashboards in changing occupant energy behaviors”. In: *Energy Research and Social Science* 19 (2016), pp. 11–20. ISSN: 22146296. DOI: 10.1016/j.erss.2016.04.020. URL: http://dx.doi.org/10.1016/j.erss.2016.04.020.

[161] Gabriel Peschiera, John E Taylor, and Jeffrey A Siegel. “Response-relapse patterns of building occupant electricity consumption following exposure to personal, contextualized and occupant peer network utilization data”. In: *Energy and Buildings* 42.8 (2010), pp. 1329–1336. ISSN: 03787788. DOI: 10.1016/j.enbuild.2010.03.001. URL: http://dx.doi.org/10.1016/j.enbuild.2010.03.001.

[162] Kaiyu Sun and Tianzhen Hong. “A simulation approach to estimate energy savings potential of occupant behavior measures”. In: *Energy and Buildings* 136 (2017), pp. 43–62. ISSN: 03787788. DOI: 10.1016/j.enbuild.2016.12.010. URL: http://dx.doi.org/10.1016/j.enbuild.2016.12.010.

[163] Robert J Meyers, Eric D Williams, and H Scott Matthews. “Scoping the potential of monitoring and control technologies to reduce energy use in homes”. In: *Energy and Buildings* 42.5 (2010), pp. 563–569. ISSN: 03787788. DOI: 10.1016/j.enbuild.2009.10.026. URL: http://dx.doi.org/10.1016/j.enbuild.2009.10.026.

[164] Vincent Becker et al. “Estimating the savings potential of occupancy-based heating strategies”. In: *Energy Informatics* 1.Suppl 1 (2018), pp. 35–54. ISSN: 25208942. DOI: 10.1186/s42162-018-0022-6.

[165] Yan Zhang et al. “Rethinking the role of occupant behavior in building energy performance: A review”. In: *Energy and Buildings* 172 (2018), pp. 279–294. ISSN: 03787788. DOI: 10.1016/j.enbuild.2018.05.017. URL: https://doi.org/10.1016/j.enbuild.2018.05.017.

[166] Yuzhen Peng et al. “Using machine learning techniques for occupancy-prediction-based cooling control in office buildings”. In: *Applied Energy* 211.December 2017 (2018), pp. 1343–1358. ISSN: 03062619. DOI: 10.1016/j.apenergy.2017.12.002. URL: https://doi.org/10.1016/j.apenergy.2017.12.002.

[167] Antonio Paone and Jean Philippe Bacher. “The impact of building occupant behavior on energy efficiency and methods to influence it: A review of the state of the art”. In: *Energies* 11.4 (2018). ISSN: 19961073. DOI: 10.3390/en11040953.

[168] Tiffany Grace Holmes. “Eco-visualization: Combining art and technology to reduce energy consumption”. In: *Creativity and Cognition 2007, CC2007 - Seeding Creativity: Tools, Media, and Environments* (2007), pp. 153–162. DOI: 10.1145/1254960.1254982.

[169] Philipp Althaus et al. “Enhancing Building Monitoring and Control for District Energy Systems: Technology Selection and Installation within the Living Lab Energy Campus”. In: *Applied Sciences* 12.7 (2022). ISSN: 2076-3417. DOI: 10.3390/app12073305. URL: https://www.mdpi.com/2076-3417/12/7/3305.

[170] Shibboleth Consortium. *Shibboleth Consortium*. URL: https://www.shibboleth.net/ (visited on 09/04/2023).

Appendix A

Software and Hardware Details

This appendix provides more details on the implementation of the software, as well as the supporting hardware framework consisting of necessary sensors, actuators, and hardware communication platforms is presented.

A.1 Overall System Design

A.1.1 System Architecture

An architectural overview of the system developed in this thesis is depicted in Fig. A.1, showing the main components of the overall system, their interrelationships, and the supporting databases that provide the necessary data for occupant behaviour evaluation.

As shown in Fig. A.1, JuControl is at the heart of the system, interfacing with the user on the one hand, and with other parts of the system on the other hand, including:

- Juracle, via the exchange of evaluation results via the Postgres database;
- ALICE, to retrieve room geometry data and images, and room energy system components and their associated sensors and actuators;
- WALDO (a device management tool used in the LLEC project) to obtain metadata on installed sensors and actuators.

Furthermore, JuControl reads sensor and actuator timeseries data from the CrateDB database, where the timeseries data from room-level devices and gateways is stored using so-called adapter scripts that bridge the data formats required by the different parts of the data capture system. The data capture system is described in Section A.1.2 below. The devices which use the KNX hardware communication protocol are not stored in the WALDO device book-keeping tool, but are included into JuControl via data files that contain the group addresses of the devices. The user interacts directly not just with JuControl, but also with the Campus Viewer.

A.1.2 ICT Platform for Hardware and Data Capture

In this section, the ICT platform for hardware and the communication protocols that facilitate the capture of the field-level sensor data used in the thesis are presented. For the research undertaken in this thesis, the following data was available and utilized:

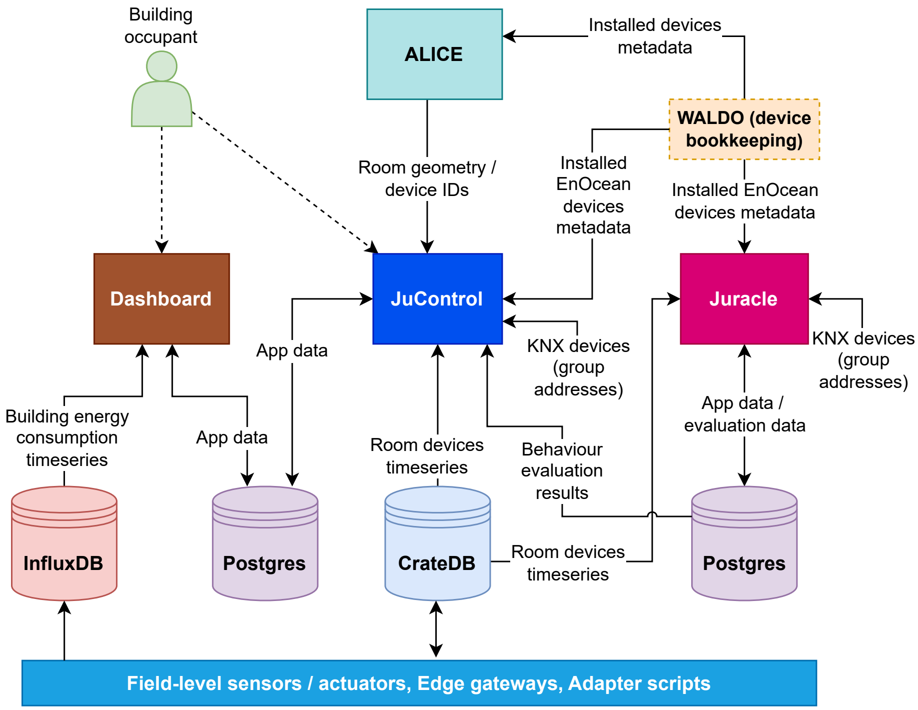


Fig. A.1 Overall system architecture of the Energy Dashboard Suite showing databases and high-level interaction amongst components.

- building-level energy demand data for heating and electricity;
- indoor air quality and environmental data at room level, specifically indoor temperature, CO_2 concentration, and relative humidity;
- radiator setpoint temperature, local temperature around the radiator as measured by the radiator, and valve position for offices in selected buildings;
- passive infra-red presence sensor data for offices in selected buildings;
- window and door status data; and
- weather data, specifically ambient temperature and solar radiation.

A simplified diagram showing the hardware setup and communication protocols is given in Fig. A.2. More details about the hardware setup are provided in Althaus et al. [169] and Redder et al. [146].

Building metering data is captured within a proprietary system developed by an external contractor and managed by the Facility Management Department of FZJ. For research purposes, a custom plugin was deployed by the external contractor to publish the metering data to an MQTT broker. The *MQTT-to-InfluxDB Adapter*, which subscribes to the MQTT broker, was developed to transfer the metering data published on the MQTT broker to the InfluxDB timeseries database. The metering data thus transferred include heating and electrical power demand at the building level for most of the buildings in the FZJ campus, currently at minute-wise resolution. A total of 153 buildings are thus metered and the data stored in the InfluxDB database. Some of these buildings are divided into multiple wings that are either conjoined or separate, in which case the metering is at the building wing level, bringing the total number of metered building wings to 326.

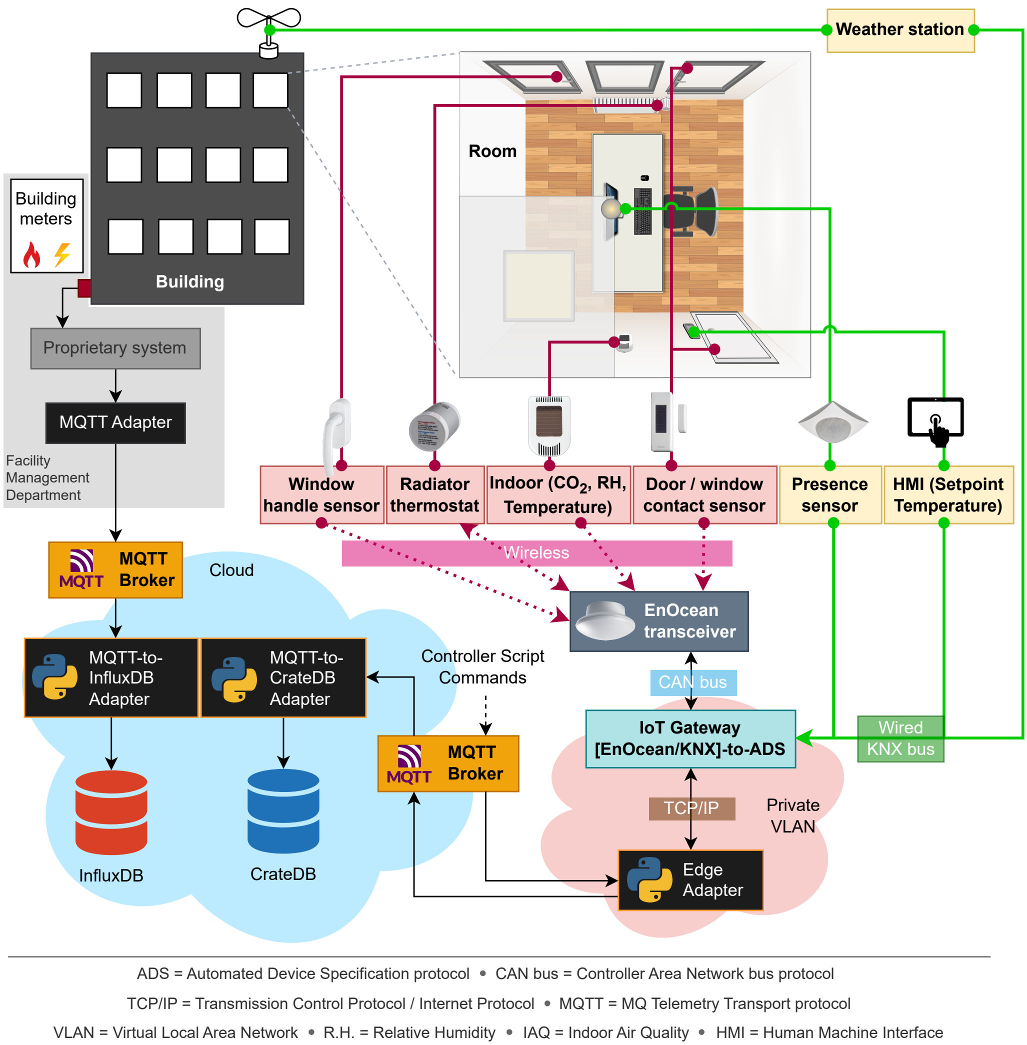


Fig. A.2 Hardware architecture and communication protocols for the sensors and actuators used to gather data for the thesis and to control devices.

At the room level, two major communication protocols are involved in the hardware setup, namely the EnOcean wireless protocol, and the KNX-TP (i.e. KNX twisted pair) protocol. The EnOcean protocol is used by the indoor environment sensors (room temperature, CO_2 concentration, and relative humidity), window and door state sensors, and the cloud-controllable radiator valve available in two buildings, while the KNX-TP protocol is used by the presence and luminosity sensors, as well as by the heating-related systems in buildings in which the building management system (BMS) manages the heating. Some other protocols serve niche purposes at this level, e.g. the CAN bus protocol. The manufacturers and models of the installed hardware devices for room-level data capture as it relates to this thesis is provided in Table A.1. As at the time of writing this thesis, about 12 buildings are equipped with these room-level sensors and / or actuators. (The full details for the installations in these buildings are provided in Table 3.6 of Chapter 3, where the experiment setup for testing the developed systems is described.) Since JuControl deals with the room level, only these 12 building are currently available in JuControl.

Table A.1 Installed hardware and their associated protocols for measuring data at the room level, for actuation, and for communication. (Adapted from [169].)

Device Type	Manufacturer	Device Model	Comm. Protocols[a]	Purpose
IoT Gateway	Beckhoff	CX5130	CAN bus, IP[b]	Collect data from multiple sources and transmit over IP to cloud
EnOcean Transceiver	Beckhoff	KL6583	EnOcean, CAN bus	Receive and send EnOcean telegrams
Air multi-sensor	Pressac	60.CO2 SLR TMP HUM.868	EnOcean	Measure CO_2 conc., relative humidity, and temperature of room air
Window handle	Thermokon	SRG02	EnOcean	Measure window state (open/closed/tilted)
Contact sensor	Eltako	FTKB	EnOcean	Measure window / door state (open/closed)
Valve actuator	Micropelt	MVA005	EnOcean	Control radiator valve position; local heating controller; measure radiator temperature and valve position
Presence sensor	MDT	SCN-P360D3.03	KNX	Detect presence in room
Luminosity sensor	-	-	KNX	Measure the luminosity in room

[a] Communication protocols
[b] Internet Protocol

As already mentioned, most of the devices installed at the room level (except the presence detector, luminosity sensor, and wall-mounted HMI for setpoint temperature) use the wireless EnOcean protocol to send data to, and receive data from, EnOcean transceivers, which in turn communicate with the edge device (IoT Gateway) via the CAN bus protocol. The IoT Gateway provides the connection to the cloud via the Internet Protocol using the connection-oriented Transmission Control Protocol (TCP) or the connectionless User Datagram Protocol (UDP). It transfers data to and from the cloud using another protocol built on top of the IP/TCP protocol, called the Automated Device Specification (ADS) protocol. The *Edge Adapter* program sitting in the cloud then transforms the ADS-encoded data into JSON and publishes it to the MQTT broker on predefined MQTT topics. Another adapter program (the *MQTT-to-CrateDB Adapter*) listens for data on the relevant topics on the MQTT broker and then writes the data to the CrateDB database.

A.2 JuControl System Architecture

The high-level architecture of JuControl is shown in Fig. A.3. The *Privacy / Access Manager* enforces user access restrictions to building and room information, and to sensor data. The room geometry and sensor information is in turn managed by the *Building, Room, and Devices Manager*. The Gamification manager oversees the experiment setup and uses evaluation data from Juracle to provide gamification functionality to the user. The *Occupant Schedule & Comfort Preference Manager* is responsible for sourcing occupant thermal preferences and their presence schedules via the JuControl calendar. It then supplies the aggregated thermal preferences to the external heating controller via an API. In the following subsection, the coupling of JuControl with the heating controller is discussed.

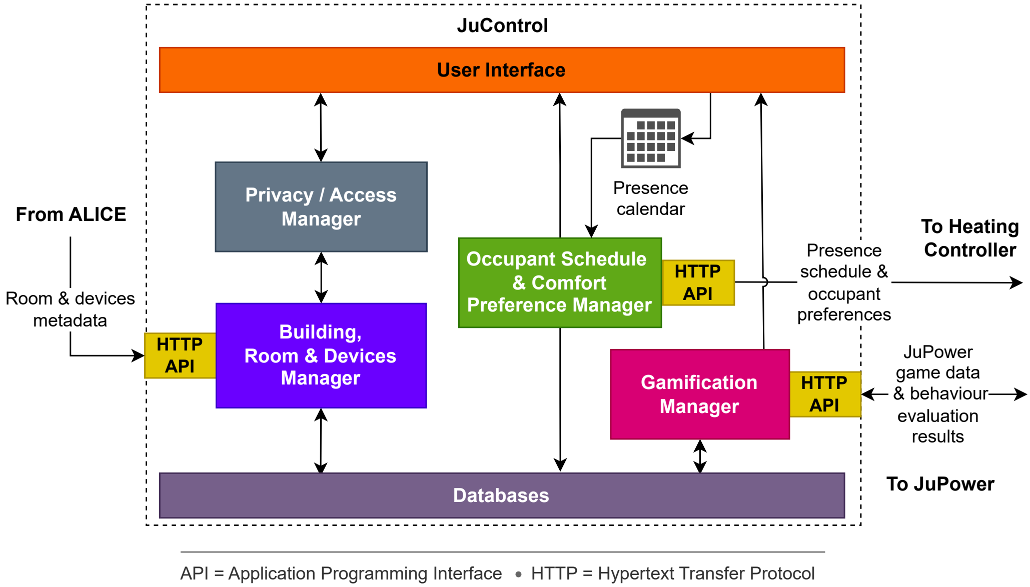


Fig. A.3 JuControl system architecture showing the main components.

A.3 Dealing with User Privacy and Data Security

As discussed in Chapters 1 and 2, the issue of privacy requires special consideration in a research work such as this. Consequently, several design decisions are based on considerations for privacy preservation, in addition to factoring in legal and policy limitations.

A.3.1 Legal and Policy Compliance

In collaboration with the data privacy officer for Forschungszentrum Jülich and the Works Council (*Betriebsrat* in German), the policy-related compliance issues were ironed out. In particular, the following measures were taken in this thesis (and all related research activities) to meet the data privacy requirements as agreed with the relevant stakeholders.

- **Data privacy course** All the developers having access to one or more project databases completed a data privacy course.
- **Data agreement by users** In JuControl, where data that can be used to indirectly track user activities is available, a data agreement form has to be consented to by all the occupants of the room before the information is shown to any of the room's occupants.
- **Data for research** For research purposes, measurement data can be processed anonymously independent of the user data consent in JuControl.

A.3.2 Data Isolation Measures

One of the requirements for privacy policy compliance in this work is that data should only be accessed by authorised persons. The implication of this is that users should only see their own data. This was particularly of importance to the Works Council, and such special data isolation schemes were developed in this thesis to fulfil this policy requirement.

First, in the context of competitions and comparisons of office performance, since these performance metrics are considered private to the offices to which they apply, the scheme shown in Fig. A.4 was developed. In this scheme, actual office performance data is only visible to the occupants of the office, who are required to all agree to the data access before room-related data is shown to any occupant. As mentioned previously, the gamification aspect of JuControl necessitated the formation of teams comprising multiple offices per team (described in detail in Chapter 3). At the team level, the ratings of other offices are ranked and shown to members of a team anonymously, so that the user knows only the identify of their own office in the ranking (see Fig. 4.5b). Above the team level, only team-related performance is shown to members of other teams. The membership of offices to teams is also protected, such that each team's members only know the offices within their own team (both JuControl-activated and non-activated offices). Additionally, a statistically significant number of offices are assigned to each team (between 20 and 44 offices per team), so that the identity of the offices that contribute to the team statistics cannot be retroactively derived from the statistics, thus preserving privacy.

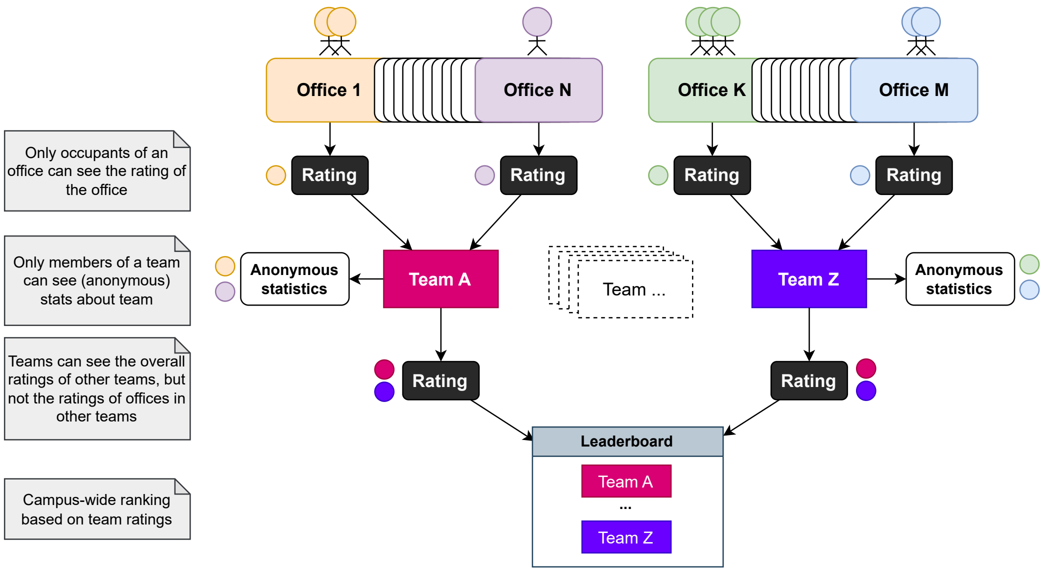


Fig. A.4 Proof-of-concept depiction of privacy-compliant strategy for displaying and comparing behaviour evaluation ratings at the office, team, and campus levels.

Secondly, a scheme was developed to track the official allocation of staff to offices such that relocation of staff is accounted for automatically. A script runs periodically on the JuControl server to check for changes in office allocation in the centrally managed facility allocation database of FZJ, and when changes occur that add new occupants to an office, JuControl is deactivated for the office automatically and the new occupants are presented with the data access agreement form. Each data consent that is granted by an occupant is stored for the office for which the user granted it, so that when the occupant moves to a new office, they need to grant consent for the new office separately. Furthermore, the new occupants, on consenting to the data agreement, can only view historical room data which begins at the time-point at which they joined the office. In this scheme also, when an occupant relocates away from an office, they lose access to the data of the office from which they left.

A.3.3 Authentication and Data Access

In order to reduce the risk of data leaks and security compromises, an early decision in the development of the Energy Dashboard suite was to "outsource" authentication by using the Shibboleth federated authentication system [170] managed by the IT services department of the campus. By this token, users are not required to input their passwords at any point in any app in the Suite, and the burden of password encryption and management is totally avoided. It also serves to reduce scepticism about the trustworthiness of the development team to manage such sensitive information, since some users would probably share the same passwords across many sensitive applications.

Furthermore, besides human users of the Energy Dashboard Suite, other systems ("clients") also consume data from the applications ("servers") via Application Programming Interface (API) endpoints. Specific instances of such API data consumption include:

- The Campus Viewer supplies building and room location information, via an API endpoint, to an "external" book-keeping tool (called *WALDO*). WALDO manages the wireless sensors and actuators installed within the wider LLEC project. The Campus Viewer also provides the same information to *ALICE* in a similar manner.
- The cloud-based room heating controller available in some buildings pulls presence and comfort preference data from a JuControl API endpoint.

Access to these API endpoints is validated using short-lived access tokens that are provided to the API clients when they are successfully authenticated by the respective API server.

Appendix B

User Notifications and User Survey Questionnaire

This appendix presents the full results of the post-experiment user survey, followed by samples of the communication sent automatically via email to the participants as evaluation summary and as recommendation.

B.1 User Survey Questionnaire

The questions asked in the post-experiment user survey are shown below, along with the options and frequency of responses. The sub-texts accompanying the questions and the visual aids are omitted, and some question and option texts have been edited to shorten them for presentation below or to remove unnecessary details. In total, 113 respondents completed the survey.

Q1. How did you know about JuControl? Please select all that apply.

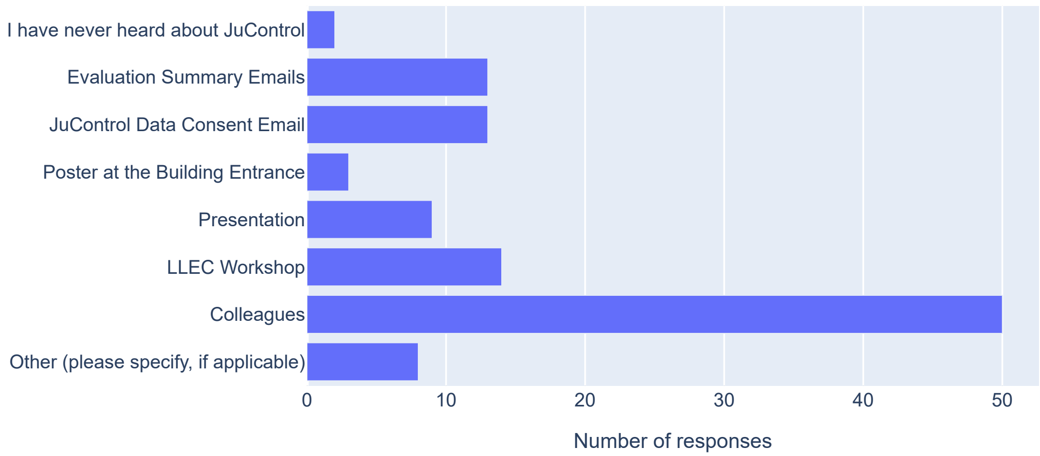


Fig. B.1 Response to survey question Q1

Q2. How often did you visit JuControl in general during the evaluation period (March to April 2023)?

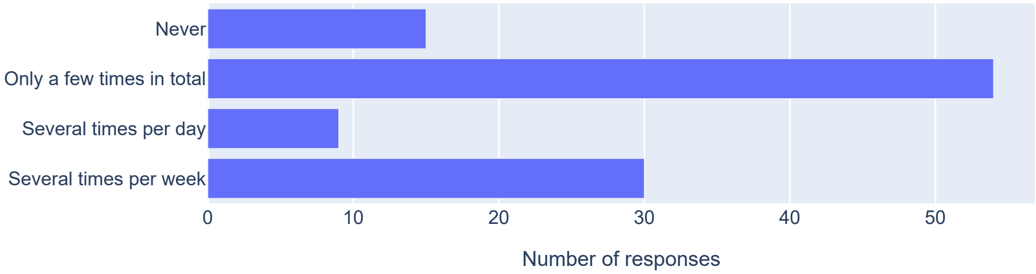


Fig. B.2 Response to survey question Q2

Q3. Which pages do you remember visiting? Please select all that apply.

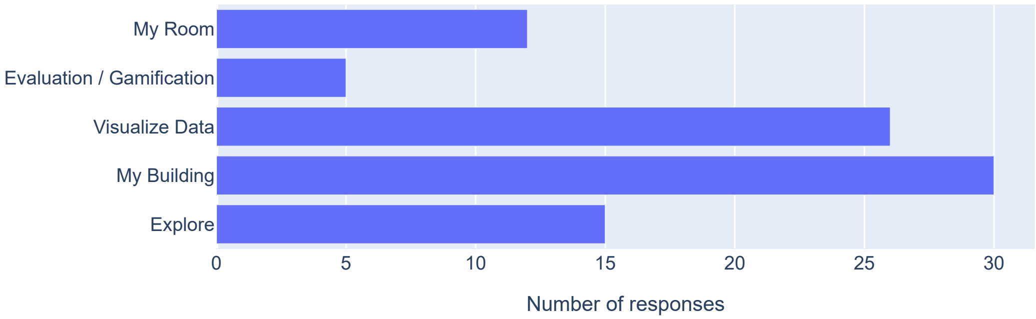


Fig. B.3 Response to survey question Q3

Q4. Which of the following pages did you visit the most?

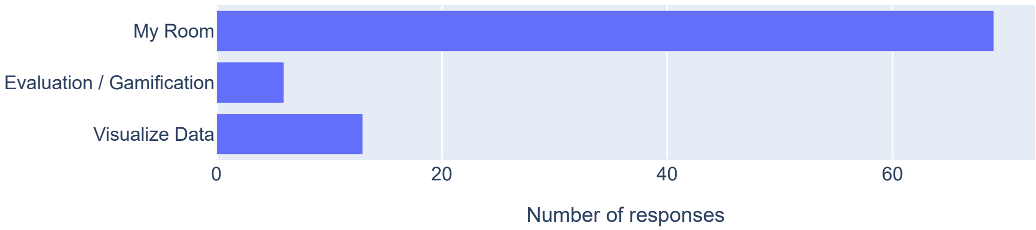


Fig. B.4 Response to survey question Q4

Q5. What data were you most interested in?

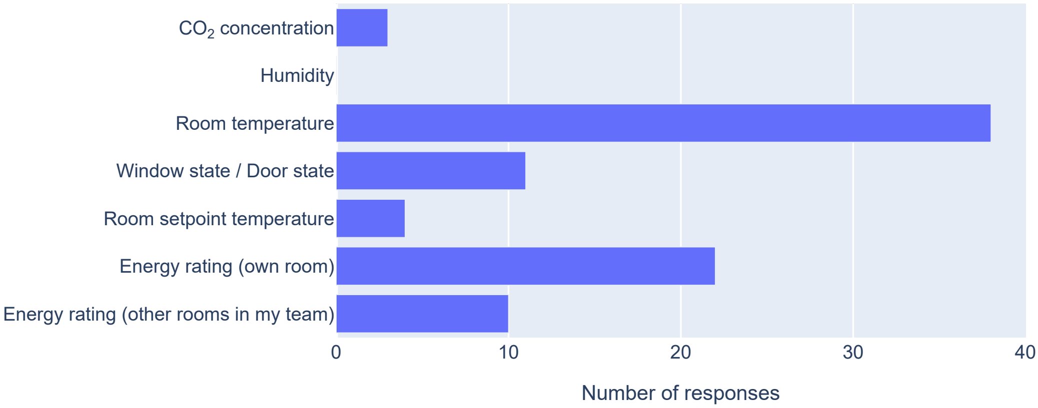


Fig. B.5 Response to survey question Q5

Q6. Did you receive any automated emails with subject similar to any of the following?

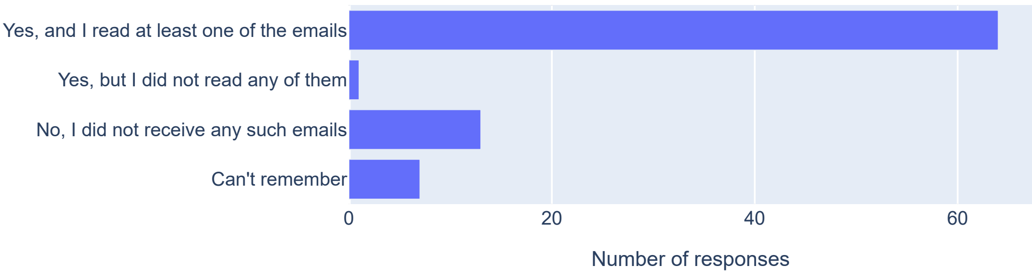


Fig. B.6 Response to survey question Q6

Q7. If yes to the above question, which of the following sums up your general reaction to the automated emails?

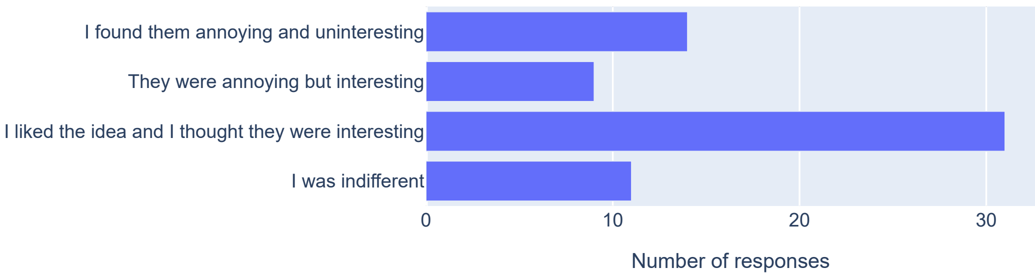


Fig. B.7 Response to survey question Q7

Q8. What annoyed you the most about the automated emails?

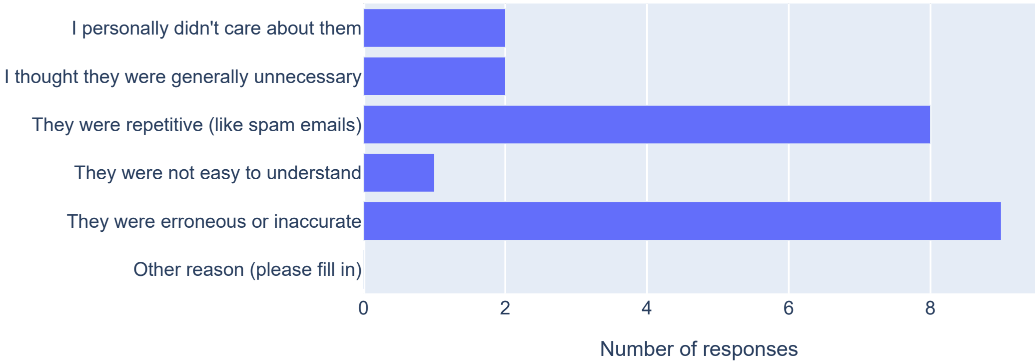


Fig. B.8 Response to survey question Q8

Q9. Were the automated emails understandable to you?

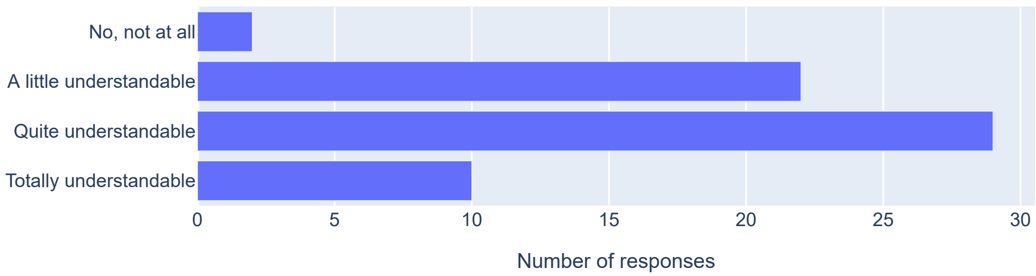


Fig. B.9 Response to survey question Q9

Q10. Did they affect your behaviour with regards to window use / heater setpoint temperature?

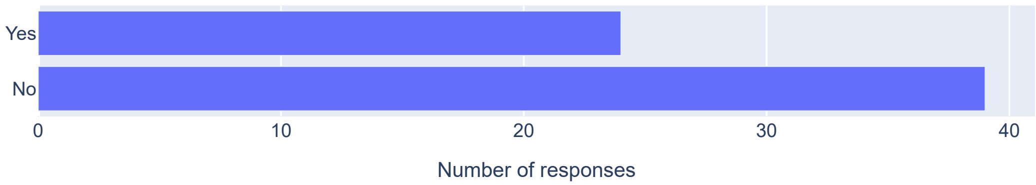


Fig. B.10 Response to survey question Q10

Q11. Do you have any suggestions for improvement, or issues you encountered?

Free-text responses.

Q12. Which of the following options most accurately describes your understanding of the energy values shown on the Evaluation page, or as mentioned in the Evaluation Summary email?

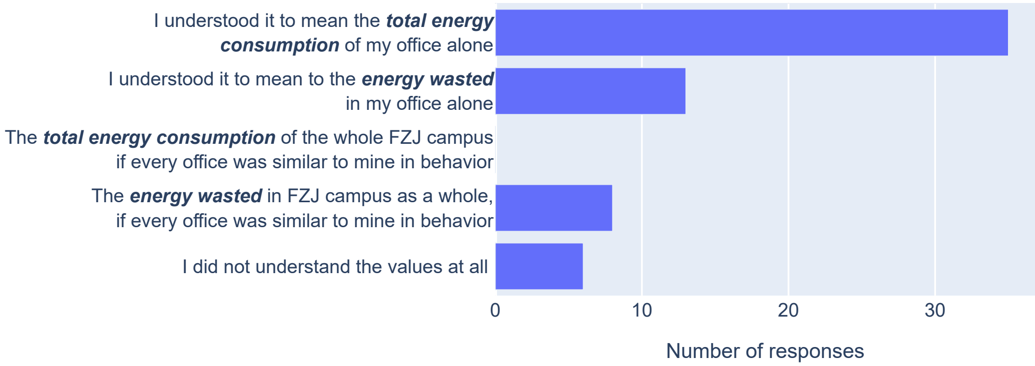


Fig. B.11 Response to survey question Q12

Q13. What did you feel about the performance (rating) of your office compared to other offices in your team?

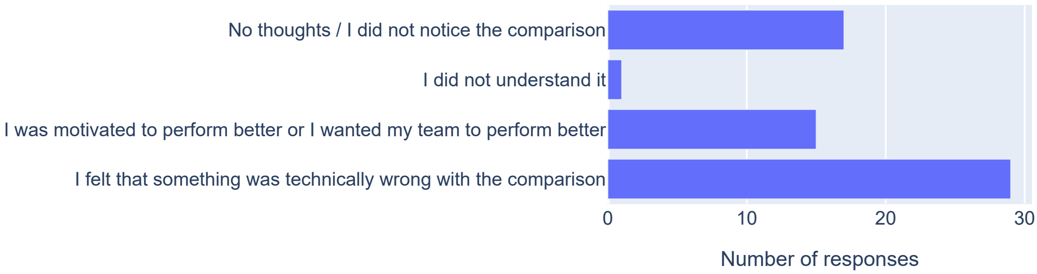


Fig. B.12 Response to survey question Q13

Q14. Do you believe that such energy rating (if properly implemented) can lead to more energy-efficient behaviour for you personally?

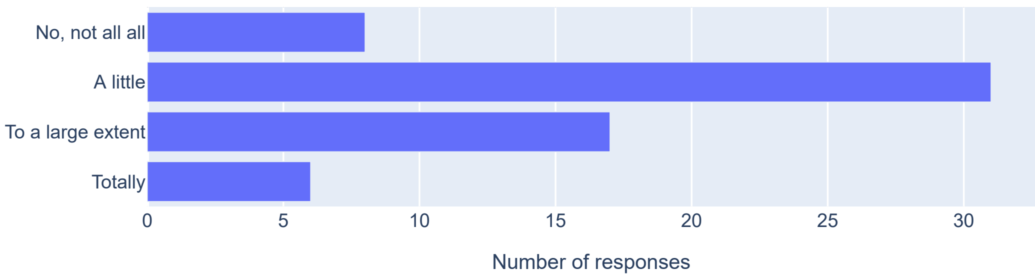


Fig. B.13 Response to survey question Q14

Q15. What was confusing / needed improvement about the energy evaluation (optional)?

Free-text responses.

Q22. Did the unavailability of the German translation of some parts of JuControl negatively affect your use and/or understanding of the app?

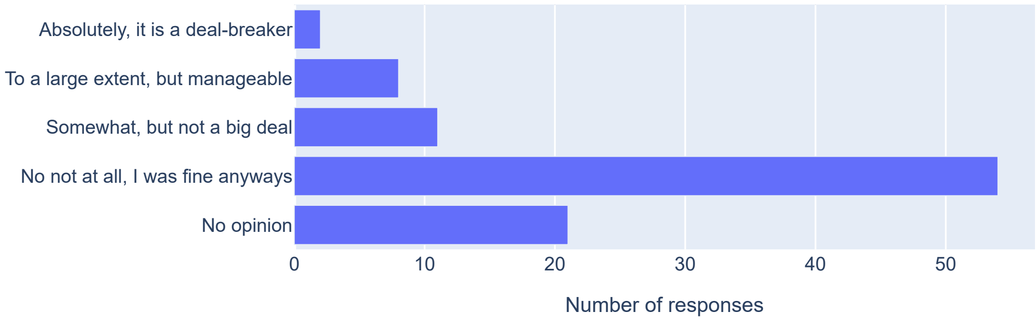


Fig. B.14 Response to survey question Q22

Q23. Final comments

Free-text responses.

B.2 Evaluation Summary and Recommendation Email Samples

Examples of the English version automatic emails sent to the users are shown below: weekly evaluation summary email (Fig. B.15), and recommendation email after exceeding the ideal ventilation duration (Fig. B.16).

JuControl Week 3 (27.03.2023 - 31.03.2023): Zusammenfassung Energiebewertung der Woche // Week Energy Evalua...

FE [redacted]

To Ubachukwu, Eziama

Reply | Reply All | Forward | ...

Wed 19/04/2023 22:18

===== English version =====

Dear JuControl User,

The daily average energy evaluation for your office (Building [redacted]) for **JuControl Evaluation Week 3** (from **March 27, 2023** to **March 31, 2023**), is

6,153.00 kWh

which is a measure of the **excess energy** consumed *per day on average* due to **non-ideal** occupant interactions with the energy system for the above week.

Your office ranks **#28 out of 39 offices** in your team.

Compared to the previous week's evaluation (from March 20, 2023 to March 24, 2023), this is **an increase** of **11%**.

Your team, **Team** [redacted], has an average evaluation of **2,005.00 kWh** for the week. Compared to the last evaluation for your team, this is **a decrease** of **26%**.

Fig. B.15 English version of evaluation summary email. A newer version clarifies that the given energy penalty was w.r.t. the entire campus, not just the user's office.

2023-04-19 JuControl: Energie-Empfehlung (Lüftung) // Energy Recommendation (Ventilation)

FE [redacted]

To Ubachukwu, Eziama

Reply | Reply All | Forward | ...

Wed 19/04/2023 22:04

Dear JuControl User,

The recommended duration for ventilation has been exceeded in your office (Building [redacted]). The recommended duration is **20 minutes** of *shock ventilation*.

The total "equivalent" ventilation duration for your office so far is **429.0 minutes** (429.0 minute(s) of *shock ventilation* and 0.0 minute(s) of *trickle ventilation*.

Note that this value takes into "trickle ventilation minutes" are counted as equivalent to 2 minutes of shock ventilation for each minute of trickle ventilation, in order to discourage the less energy-efficient trickle ventilation.

Consider opening all the windows fully at once for a few minutes, in order to minimize energy wastage while ensuring maximum refreshing of the room air.

This evaluation is suspended when the outside temperature is greater than 15 °C. You can view details in JuControl: https://energydashboard.fz-juelich.de/jucontrol/gamification/

Fig. B.16 English version of an example recommendation email in response to the ideal ventilation duration being exceeded.

B.3 Deployment poster

A sample deployment poster for a building with both JuControl and a serious game's features available is shown in Fig. B.17.

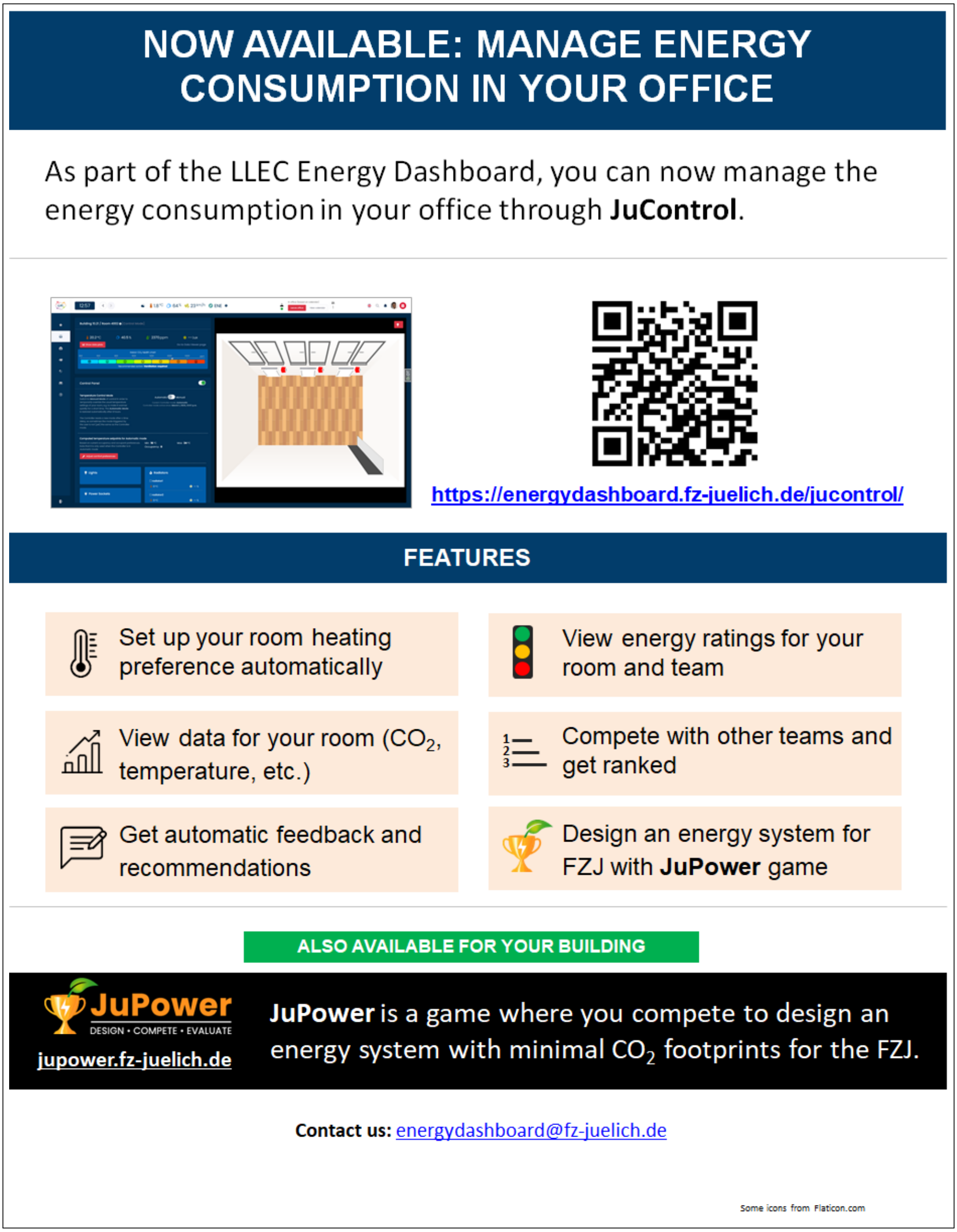


Fig. B.17 An example of a deployment poster.

Appendix C

Additional Result Plots

In this appendix, additional result plots are presented for each team with valid results, specifically Teams T1, T2, T3, T4, T5, T6, T9, T10, T12, T13 and T14. For each team, two plots are presented. First is the composite penalty plot showing daily penalties for JuControl-activated, non–JuControl-activated, and all offices in the team, in addition to the number of offices in each activation group. The second is the breakdown of penalties for these evaluated teams in terms of setpoint deviation (for teams with *Setpoint Temperature Evaluation* enabled) and ventilation duration.

C.1 Result Plots for Team T1

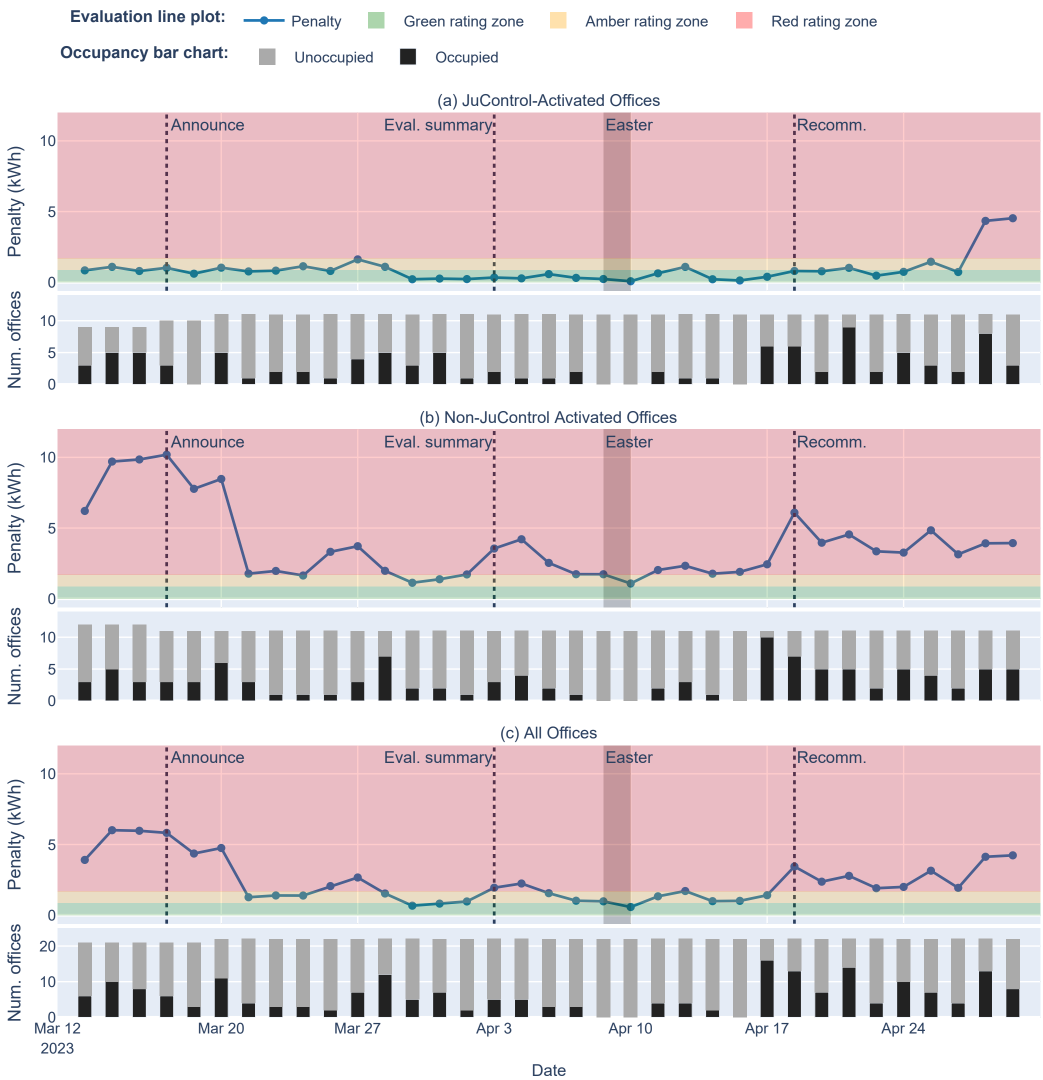


Fig. C.1 Composite plots for Team T1 showing evaluation penalty (line chart, top) and number of offices (bar chart, bottom), categorized by JuControl activation for offices in the team: (a) JuControl-activated offices, (b) non–JuControl-activated offices, and (c) all offices. Each evaluation penalty plot (top) is divided into horizontal strips corresponding to the traffic light rating of the penalty values, and important dates are marked by vertical lines. Each office-count plot (bottom) shows the number of offices in the respective JuControl-activation category as a composite bar chart, where each bar depicts the proportion of occupied/unoccupied offices for each day of the experiment period.

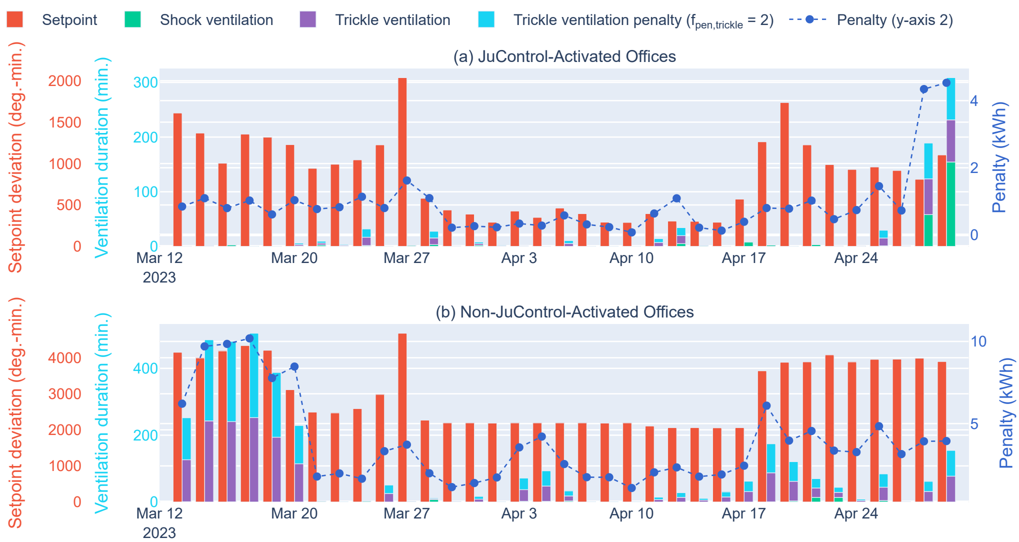


Fig. C.2 Factors contributing to penalty of **Team T1** for (a) JuControl-activated and (b) non–JuControl-activated offices. The bar charts show the average setpoint deviation and average *equivalent* ventilation duration (primary y-axis). The line chart (secondary y-axis) shows the average penalty.

C.2 Result Plots for Team T2

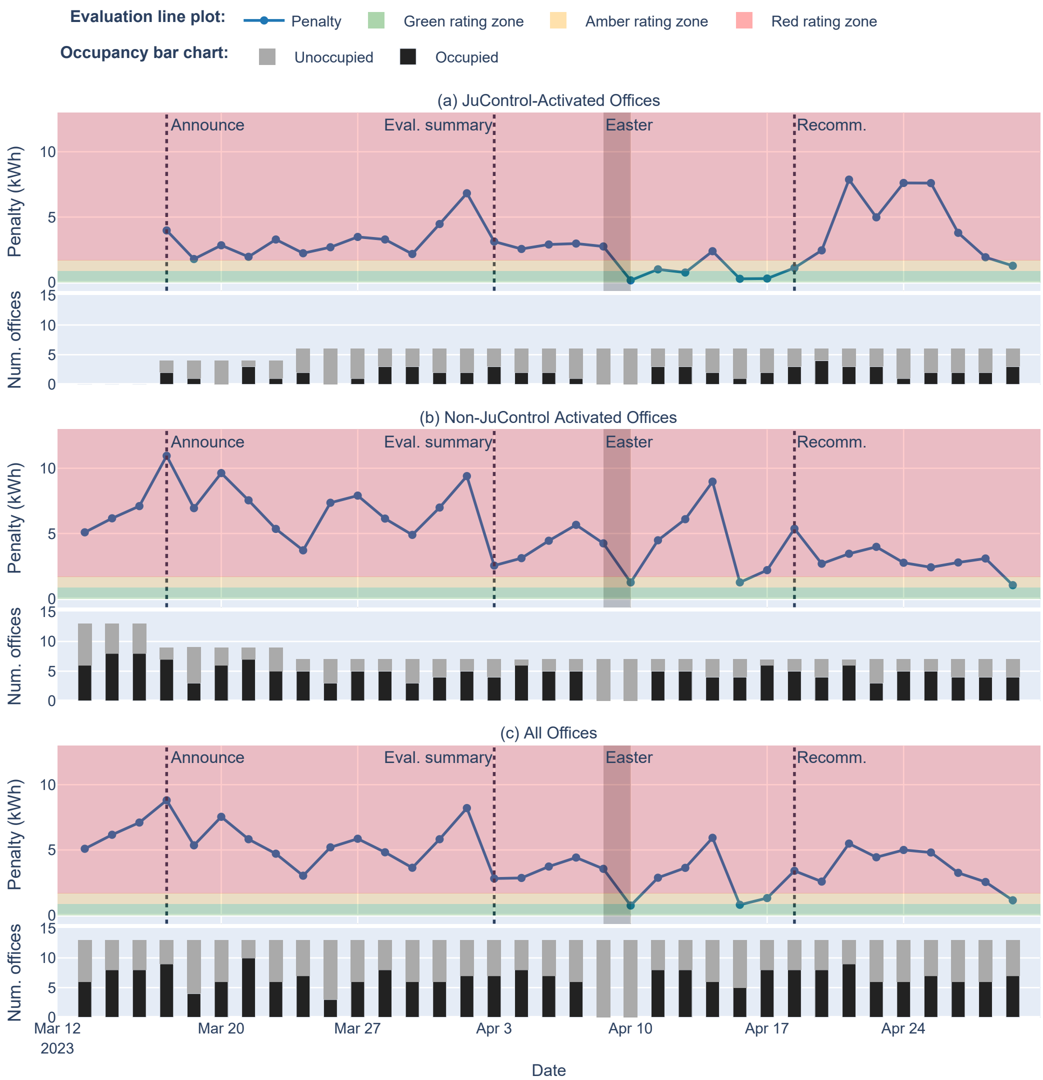


Fig. C.3 Composite plots for Team T2 showing evaluation penalty (line chart, top) and number of offices (bar chart, bottom), categorized by JuControl activation for offices in the team: (a) JuControl-activated offices, (b) non–JuControl-activated offices, and (c) all offices. Each evaluation penalty plot (top) is divided into horizontal strips corresponding to the traffic light rating of the penalty values, and important dates are marked by vertical lines. Each office-count plot (bottom) shows the number of offices in the respective JuControl-activation category as a composite bar chart, where each bar depicts the proportion of occupied/unoccupied offices for each day of the experiment period.

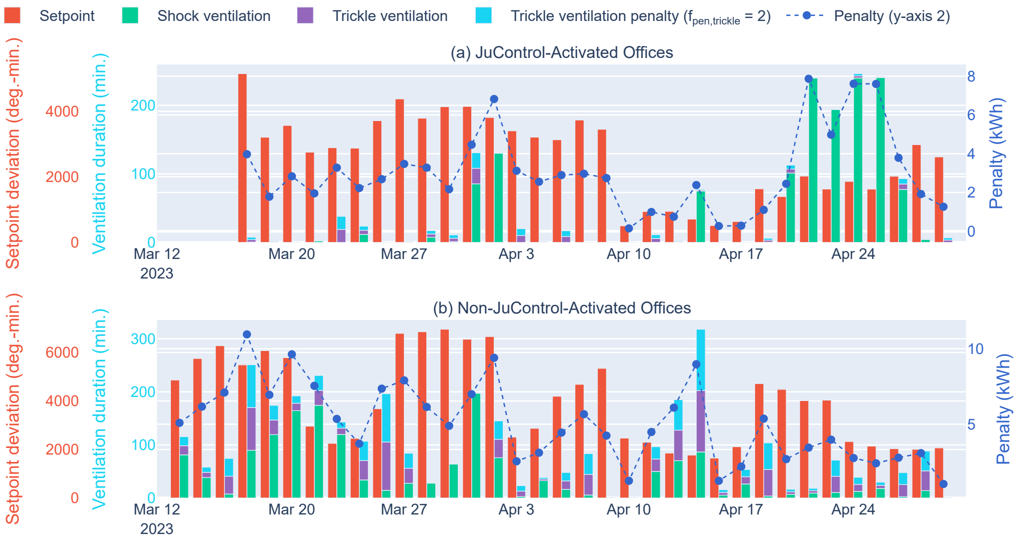


Fig. C.4 Factors contributing to penalty of **Team T2** for (a) JuControl-activated and (b) non–JuControl-activated offices. The bar charts show the average setpoint deviation and average *equivalent* ventilation duration (primary y-axis). The line chart (secondary y-axis) shows the average penalty.

C.3 Result Plots for Team T3

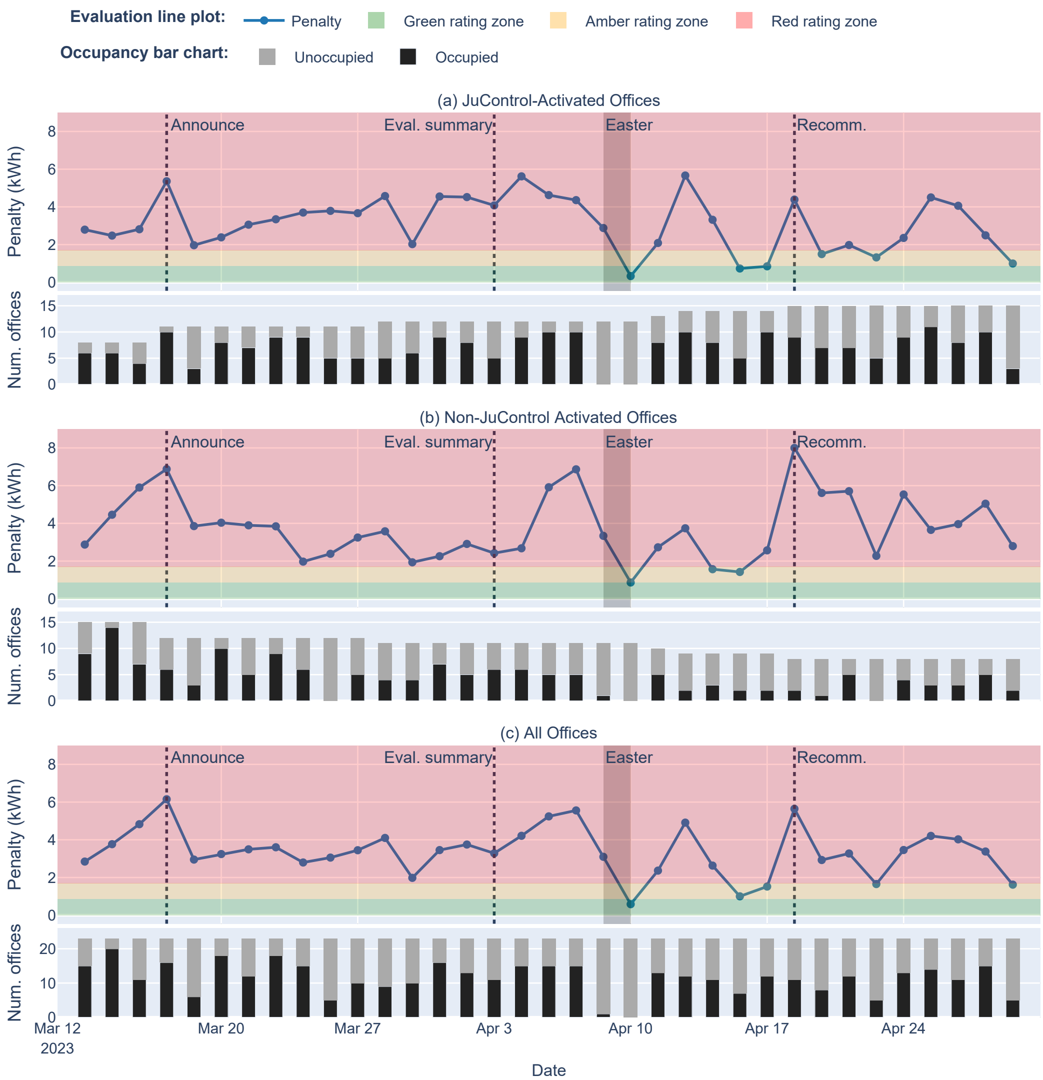


Fig. C.5 Composite plots for Team T3 showing evaluation penalty (line chart, top) and number of offices (bar chart, bottom), categorized by JuControl activation for offices in the team: (a) JuControl-activated offices, (b) non–JuControl-activated offices, and (c) all offices. Each evaluation penalty plot (top) is divided into horizontal strips corresponding to the traffic light rating of the penalty values, and important dates are marked by vertical lines. Each office-count plot (bottom) shows the number of offices in the respective JuControl-activation category as a composite bar chart, where each bar depicts the proportion of occupied/unoccupied offices for each day of the experiment period.

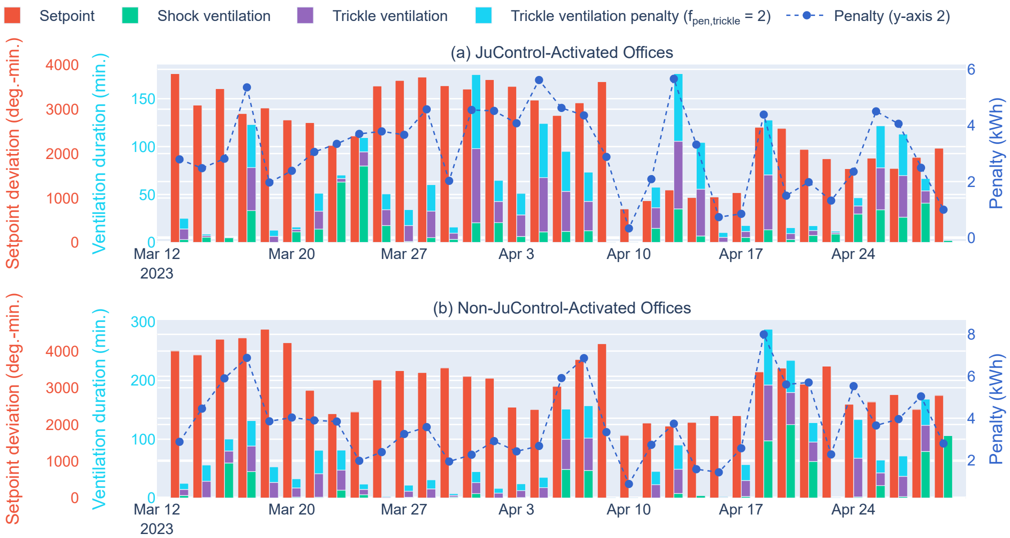


Fig. C.6 Factors contributing to penalty of **Team T3** for (a) JuControl-activated and (b) non–JuControl-activated offices. The bar charts show the average setpoint deviation and average *equivalent* ventilation duration (primary y-axis). The line chart (secondary y-axis) shows the average penalty.

C.4 Result Plots for Team T4

Fig. C.7 Composite plots for Team T4 showing evaluation penalty (line chart, top) and number of offices (bar chart, bottom), categorized by JuControl activation for offices in the team: (a) JuControl-activated offices, (b) non–JuControl-activated offices, and (c) all offices. Each evaluation penalty plot (top) is divided into horizontal strips corresponding to the traffic light rating of the penalty values, and important dates are marked by vertical lines. Each office-count plot (bottom) shows the number of offices in the respective JuControl-activation category as a composite bar chart, where each bar depicts the proportion of occupied/unoccupied offices for each day of the experiment period.

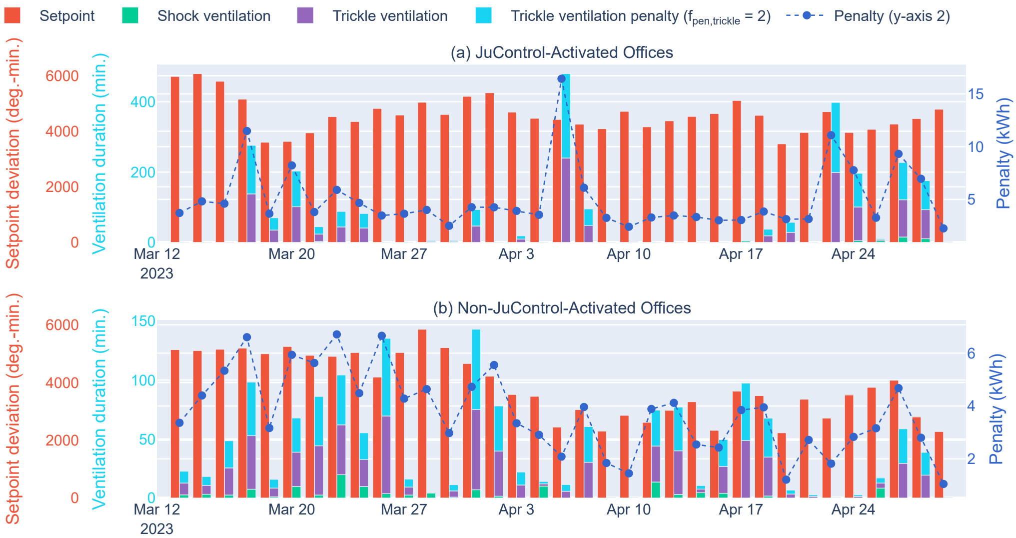


Fig. C.8 Factors contributing to penalty of **Team T4** for (a) JuControl-activated and (b) non–JuControl-activated offices. The bar charts show the average setpoint deviation and average *equivalent* ventilation duration (primary y-axis). The line chart (secondary y-axis) shows the average penalty.

C.5 Result Plots for Team T5

Fig. C.9 Composite plots for Team T5 showing evaluation penalty (line chart, top) and number of offices (bar chart, bottom), categorized by JuControl activation for offices in the team: (a) JuControl-activated offices, (b) non–JuControl-activated offices, and (c) all offices. Each evaluation penalty plot (top) is divided into horizontal strips corresponding to the traffic light rating of the penalty values, and important dates are marked by vertical lines. Each office-count plot (bottom) shows the number of offices in the respective JuControl-activation category as a composite bar chart, where each bar depicts the proportion of occupied/unoccupied offices for each day of the experiment period.

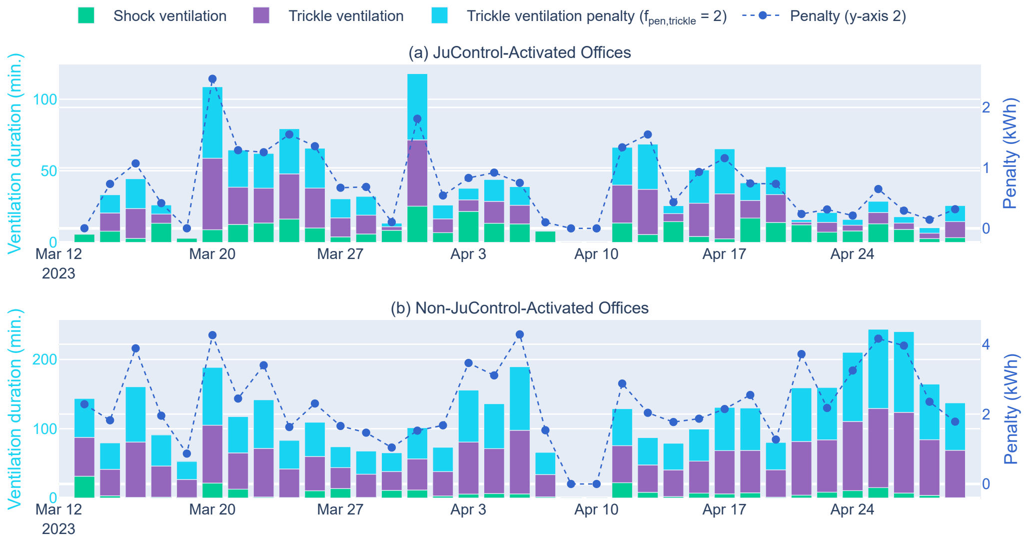


Fig. C.10 Factors contributing to penalty of **Team T5** for (a) JuControl-activated and (b) non–JuControl-activated offices. The bar charts show the average setpoint deviation and average *equivalent* ventilation duration (primary y-axis). The line chart (secondary y-axis) shows the average penalty.

C.6 Result Plots for Team T6

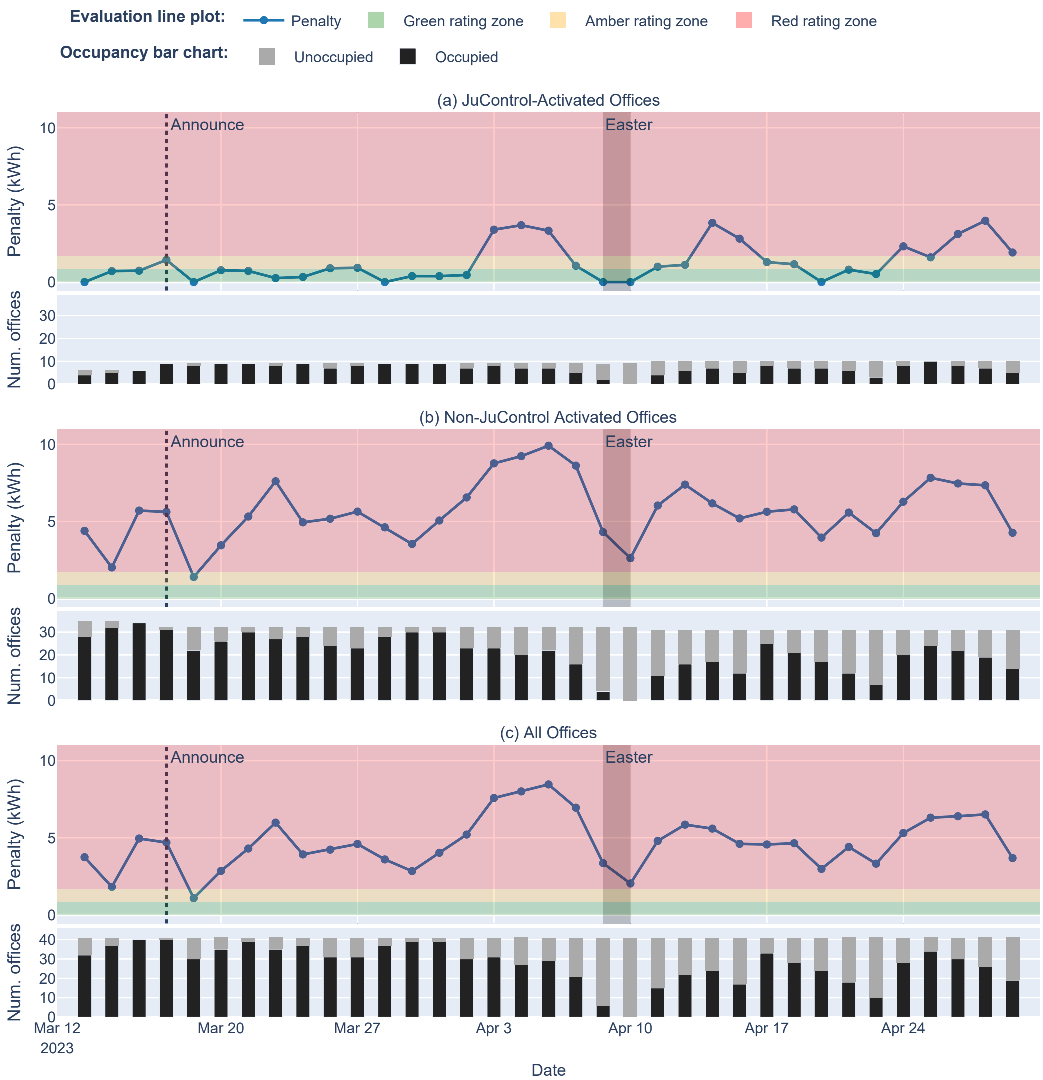


Fig. C.11 Composite plots for Team T6 showing evaluation penalty (line chart, top) and number of offices (bar chart, bottom), categorized by JuControl activation for offices in the team: (a) JuControl-activated offices, (b) non–JuControl-activated offices, and (c) all offices. Each evaluation penalty plot (top) is divided into horizontal strips corresponding to the traffic light rating of the penalty values, and important dates are marked by vertical lines. Each office-count plot (bottom) shows the number of offices in the respective JuControl-activation category as a composite bar chart, where each bar depicts the proportion of occupied/unoccupied offices for each day of the experiment period.

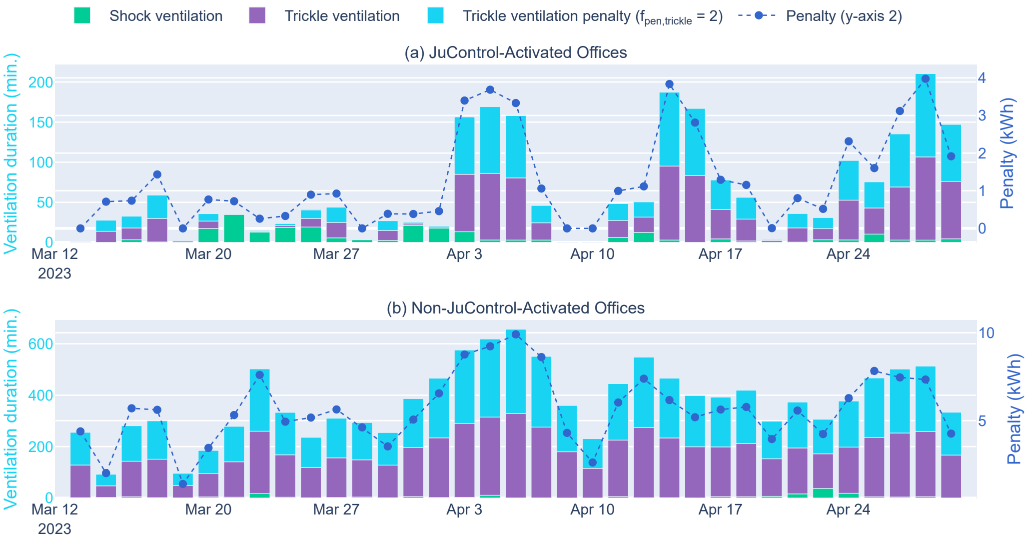


Fig. C.12 Factors contributing to penalty of **Team T6** for (a) JuControl-activated and (b) non–JuControl-activated offices. The bar charts show the average setpoint deviation and average *equivalent* ventilation duration (primary y-axis). The line chart (secondary y-axis) shows the average penalty.

C.7 Result Plots for Team T9

Fig. C.13 Composite plots for Team T9 showing evaluation penalty (line chart, top) and number of offices (bar chart, bottom), categorized by JuControl activation for offices in the team: (a) JuControl-activated offices, (b) non–JuControl-activated offices, and (c) all offices. Each evaluation penalty plot (top) is divided into horizontal strips corresponding to the traffic light rating of the penalty values, and important dates are marked by vertical lines. Each office-count plot (bottom) shows the number of offices in the respective JuControl-activation category as a composite bar chart, where each bar depicts the proportion of occupied/unoccupied offices for each day of the experiment period.

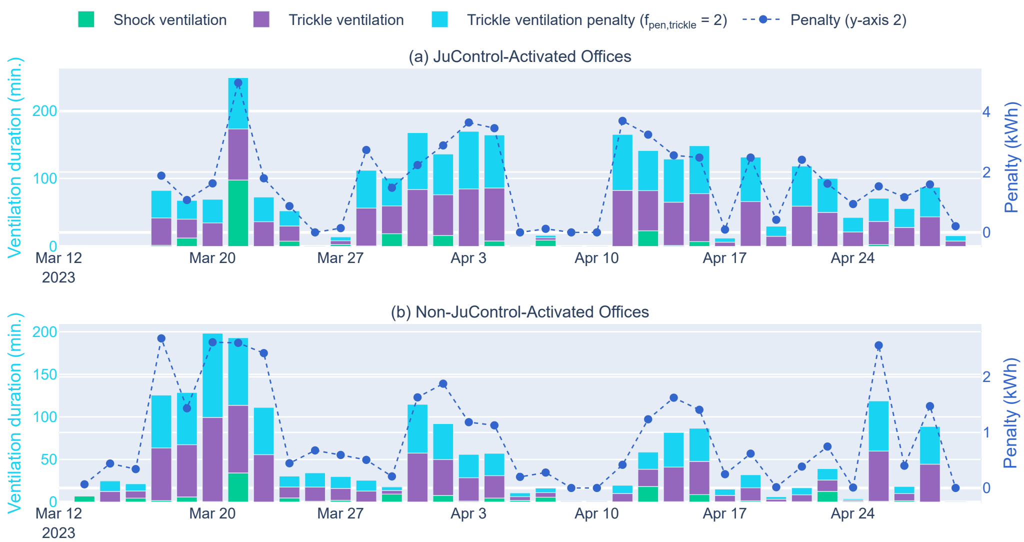


Fig. C.14 Factors contributing to penalty of **Team T9** for (a) JuControl-activated and (b) non–JuControl-activated offices. The bar charts show the average setpoint deviation and average *equivalent* ventilation duration (primary y-axis). The line chart (secondary y-axis) shows the average penalty.

C.8 Result Plots for Team T10

Fig. C.15 Composite plots for Team T10 showing evaluation penalty (line chart, top) and number of offices (bar chart, bottom), categorized by JuControl activation for offices in the team: (a) JuControl-activated offices, (b) non–JuControl-activated offices, and (c) all offices. Each evaluation penalty plot (top) is divided into horizontal strips corresponding to the traffic light rating of the penalty values, and important dates are marked by vertical lines. Each office-count plot (bottom) shows the number of offices in the respective JuControl-activation category as a composite bar chart, where each bar depicts the proportion of occupied/unoccupied offices for each day of the experiment period.

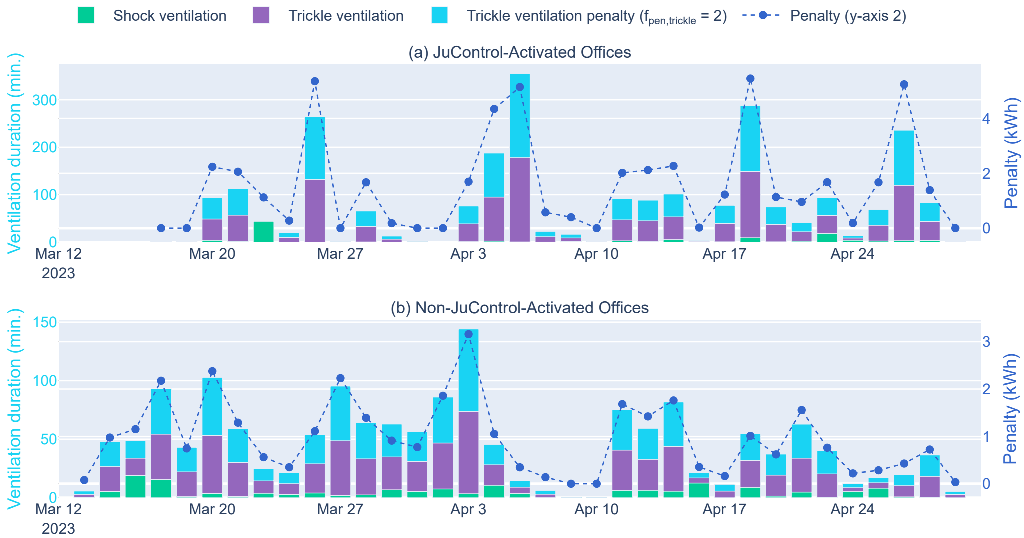


Fig. C.16 Factors contributing to penalty of **Team T10** for (a) JuControl-activated and (b) non–JuControl-activated offices. The bar charts show the average setpoint deviation and average *equivalent* ventilation duration (primary y-axis). The line chart (secondary y-axis) shows the average penalty.

C.9 Result Plots for Team T12

Fig. C.17 Composite plots for Team T12 showing evaluation penalty (line chart, top) and number of offices (bar chart, bottom), categorized by JuControl activation for offices in the team: (a) JuControl-activated offices, (b) non–JuControl-activated offices, and (c) all offices. Each evaluation penalty plot (top) is divided into horizontal strips corresponding to the traffic light rating of the penalty values, and important dates are marked by vertical lines. Each office-count plot (bottom) shows the number of offices in the respective JuControl-activation category as a composite bar chart, where each bar depicts the proportion of occupied/unoccupied offices for each day of the experiment period.

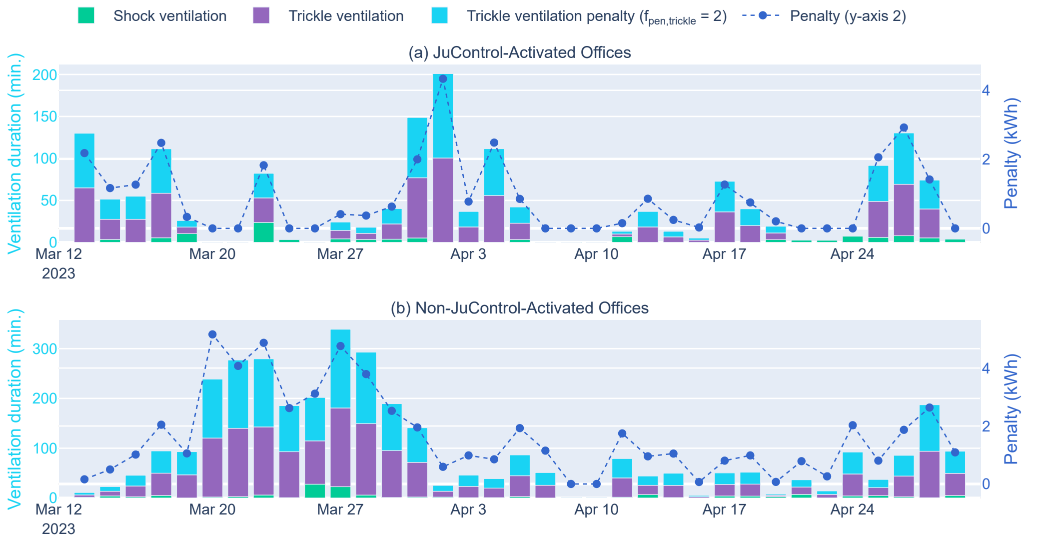


Fig. C.18 Factors contributing to penalty of **Team T12** for (a) JuControl-activated and (b) non–JuControl-activated offices. The bar charts show the average setpoint deviation and average *equivalent* ventilation duration (primary y-axis). The line chart (secondary y-axis) shows the average penalty.

C.10 Result Plots for Team T13

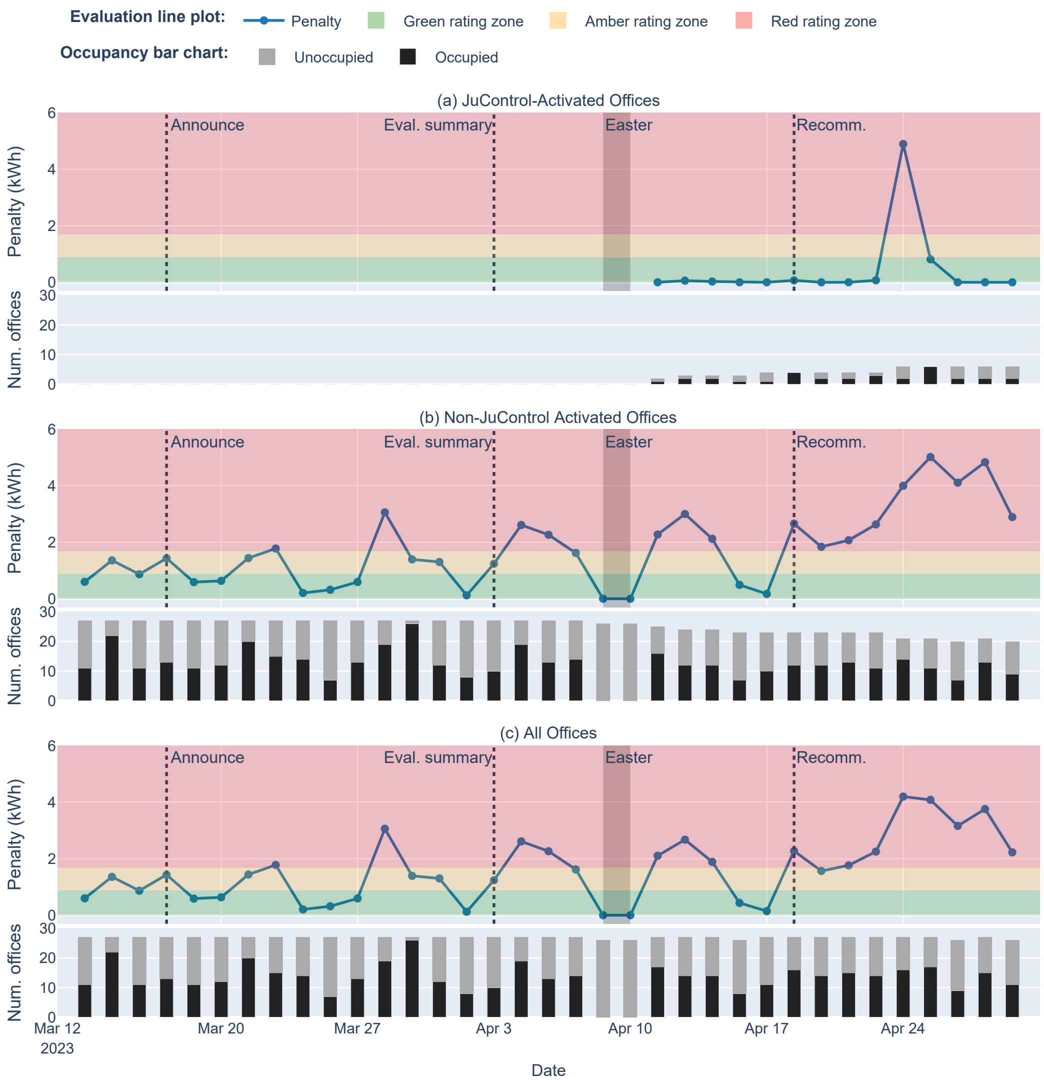


Fig. C.19 Composite plots for Team T13 showing evaluation penalty (line chart, top) and number of offices (bar chart, bottom), categorized by JuControl activation for offices in the team: (a) JuControl-activated offices, (b) non–JuControl-activated offices, and (c) all offices. Each evaluation penalty plot (top) is divided into horizontal strips corresponding to the traffic light rating of the penalty values, and important dates are marked by vertical lines. Each office-count plot (bottom) shows the number of offices in the respective JuControl-activation category as a composite bar chart, where each bar depicts the proportion of occupied/unoccupied offices for each day of the experiment period.

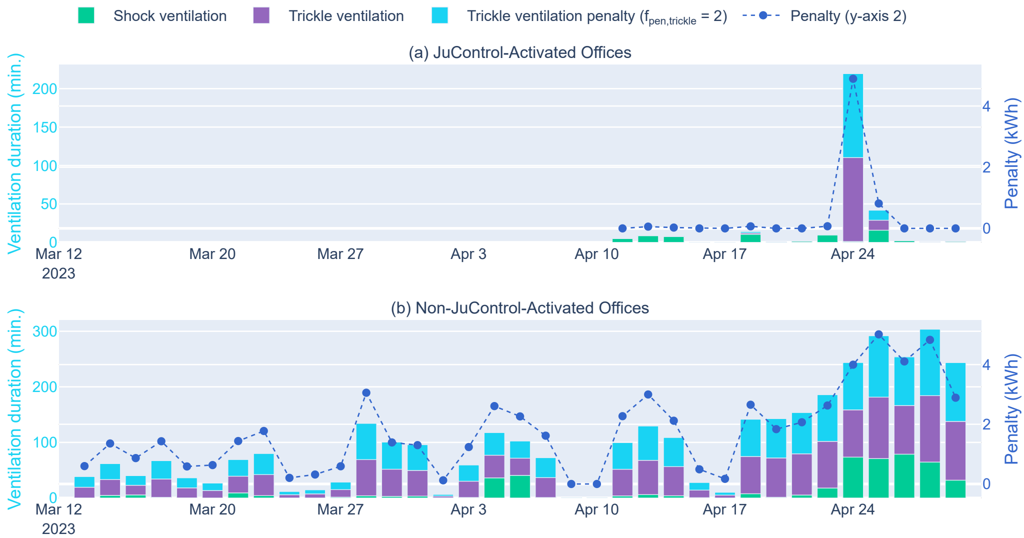


Fig. C.20 Factors contributing to penalty of **Team T13** for (a) JuControl-activated and (b) non–JuControl-activated offices. The bar charts show the average setpoint deviation and average *equivalent* ventilation duration (primary y-axis). The line chart (secondary y-axis) shows the average penalty.

C.11 Result Plots for Team T14

Fig. C.21 Composite plots for Team T14 showing evaluation penalty (line chart, top) and number of offices (bar chart, bottom), categorized by JuControl activation for offices in the team: (a) JuControl-activated offices, (b) non–JuControl-activated offices, and (c) all offices. Each evaluation penalty plot (top) is divided into horizontal strips corresponding to the traffic light rating of the penalty values, and important dates are marked by vertical lines. Each office-count plot (bottom) shows the number of offices in the respective JuControl-activation category as a composite bar chart, where each bar depicts the proportion of occupied/unoccupied offices for each day of the experiment period.

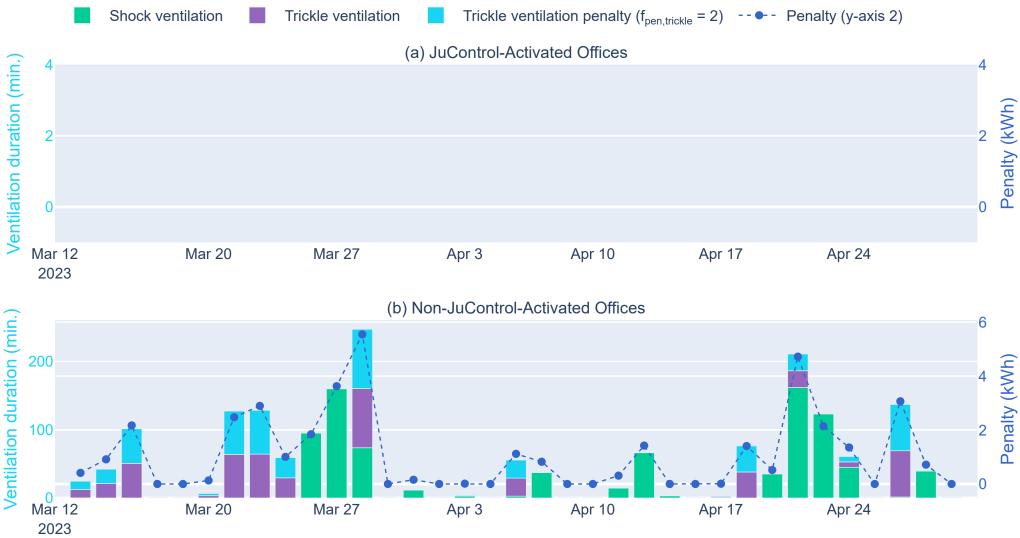


Fig. C.22 Factors contributing to penalty of **Team T14** for (a) JuControl-activated and (b) non–JuControl-activated offices. The bar charts show the average setpoint deviation and average *equivalent* ventilation duration (primary y-axis). The line chart (secondary y-axis) shows the average penalty.

Appendix D

Reference Model Derivation Details

This appendix presents some details regarding the development of the model of the reference room. Specifically, it shows the model schema of the reference room and window model, and the weather clusters determined as input for the simulation of the reference room.

D.1 Modelica Model Diagrams

The schema of the Modelica model of the reference room as depicted in Dymola is shown in Fig. D.1 below, while the Dymola representation of the implemented empirical window ventilation model based on Richter et al. [150] is shown in Fig. D.2.

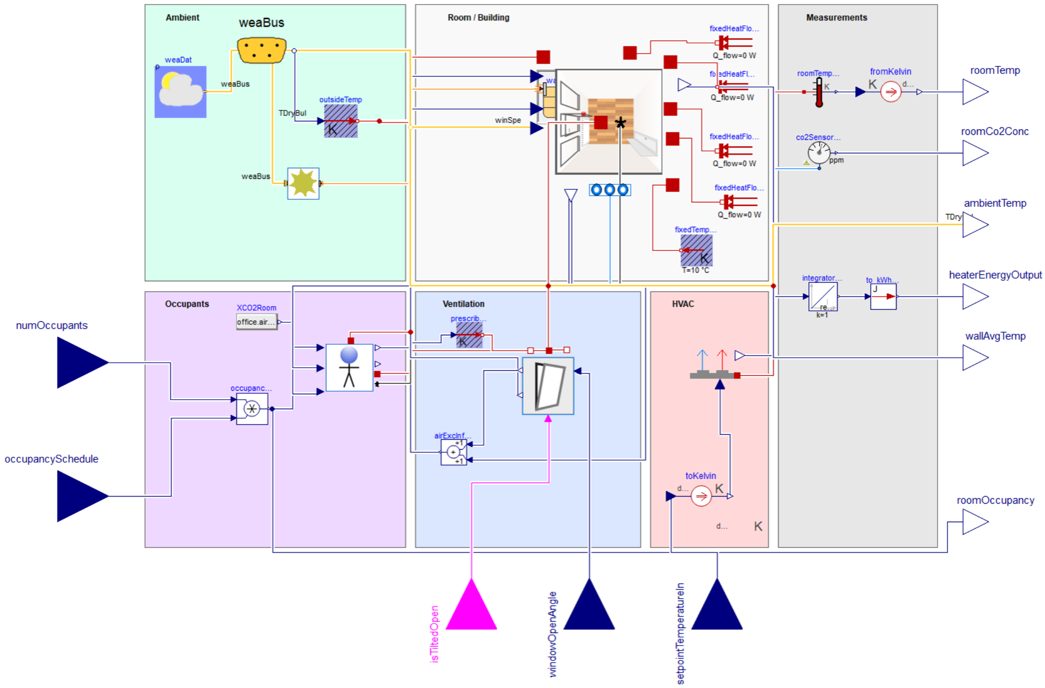


Fig. D.1 Schematic representation of the Modelica model of the reference room.

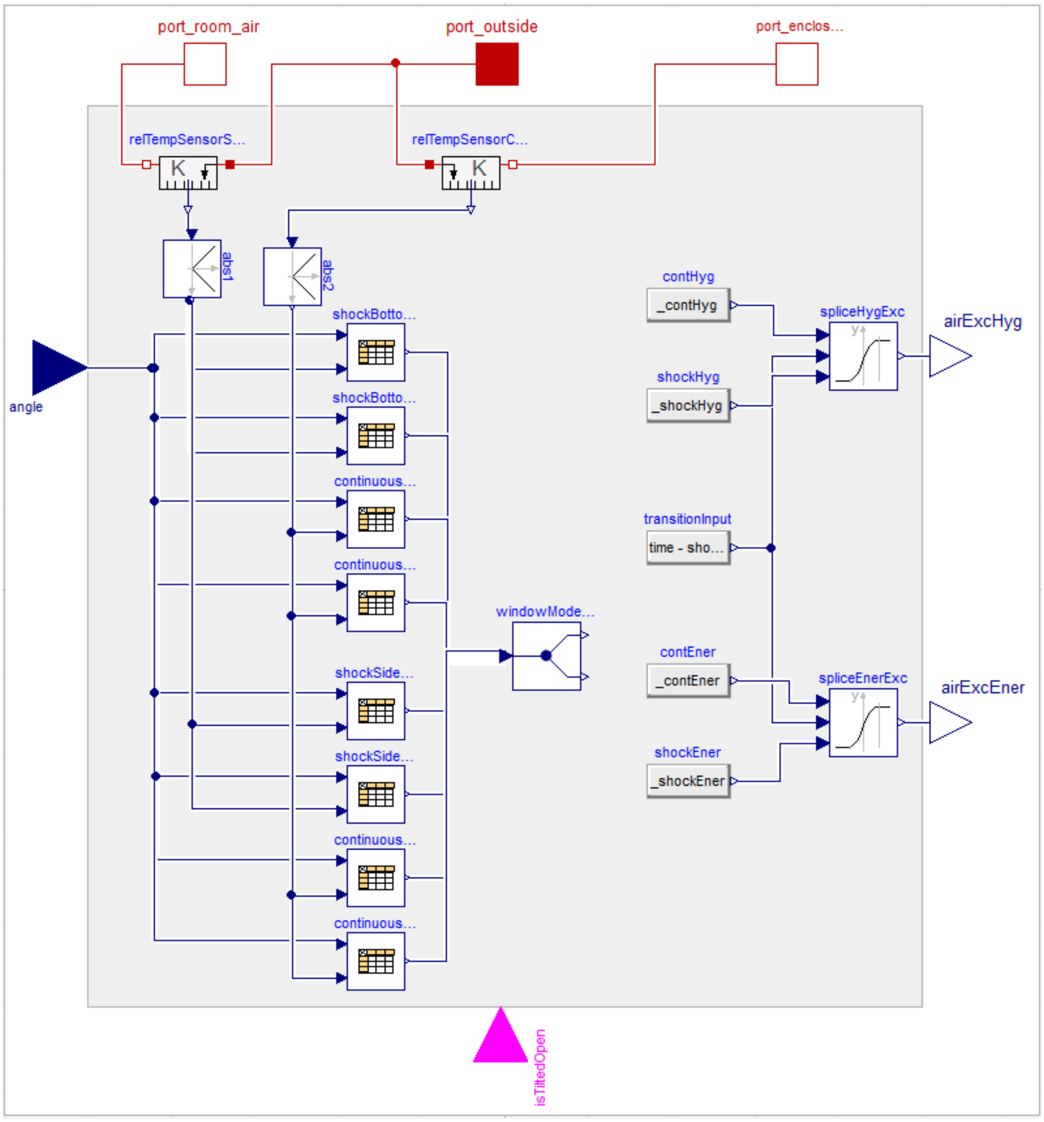


Fig. D.2 Implementation of the empirical window ventilation model of Richter et al. [150] in Modelica.

D.2 Weather Clustering

The full output of the weather clustering that formed an input into the simulation of the reference room is presented in Fig. D.3.

Fig. D.3 Weather clusters

Appendix E

Screenshots of the Energy Dashboard Suite

E.1 Campus Viewer Screenshots

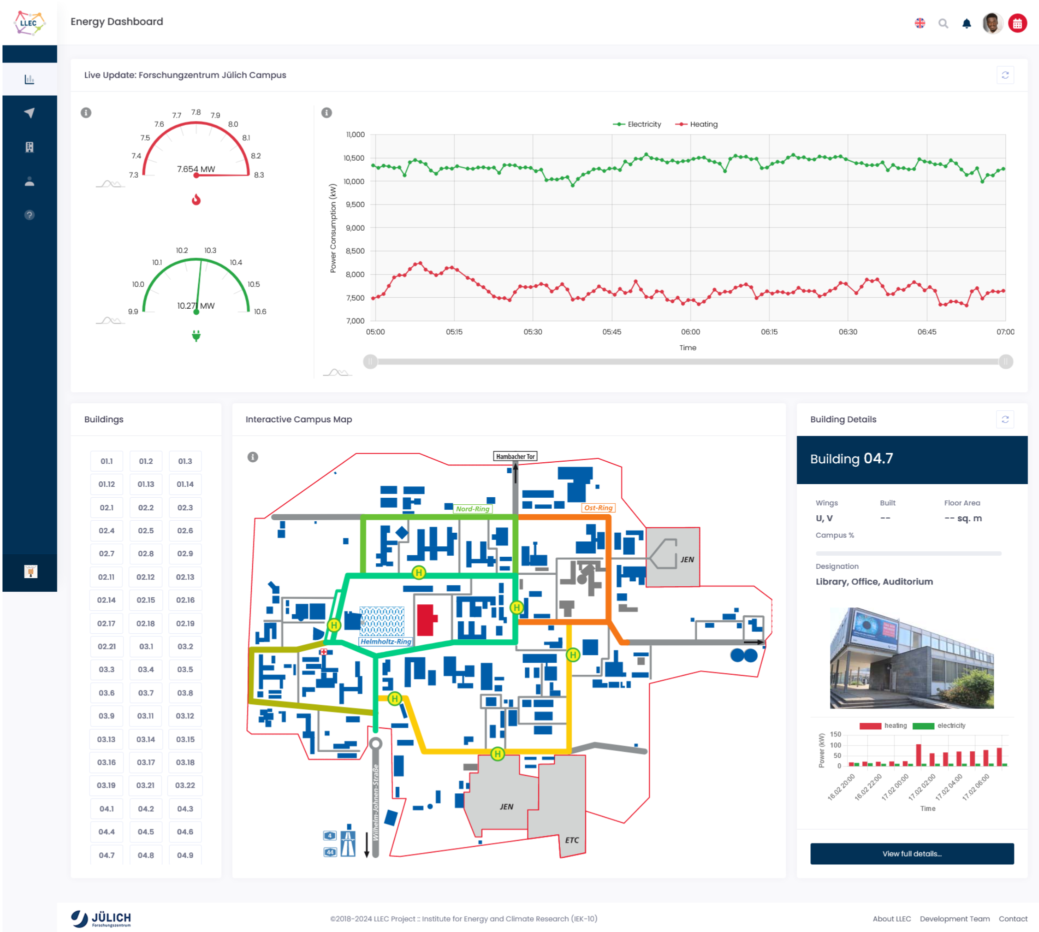


Fig. E.1 Campus Viewer home page.

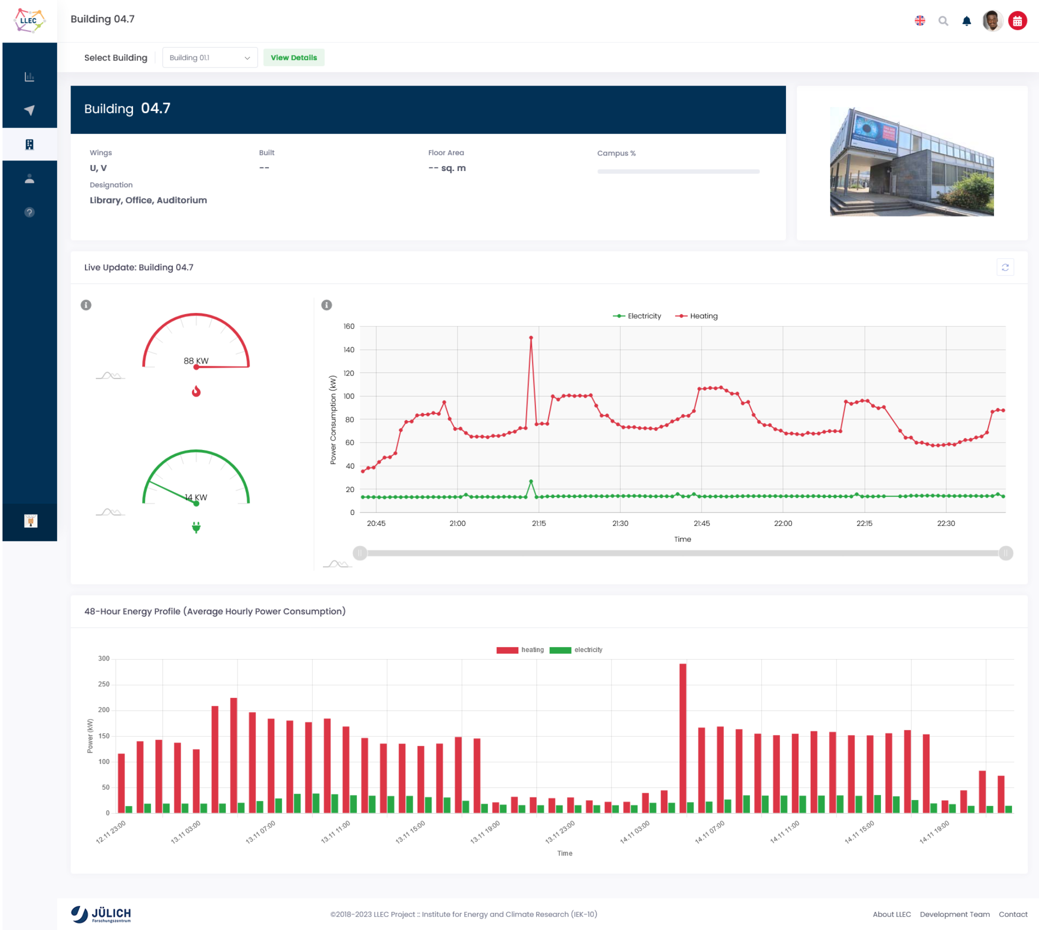


Fig. E.2 Campus Viewer building details page.

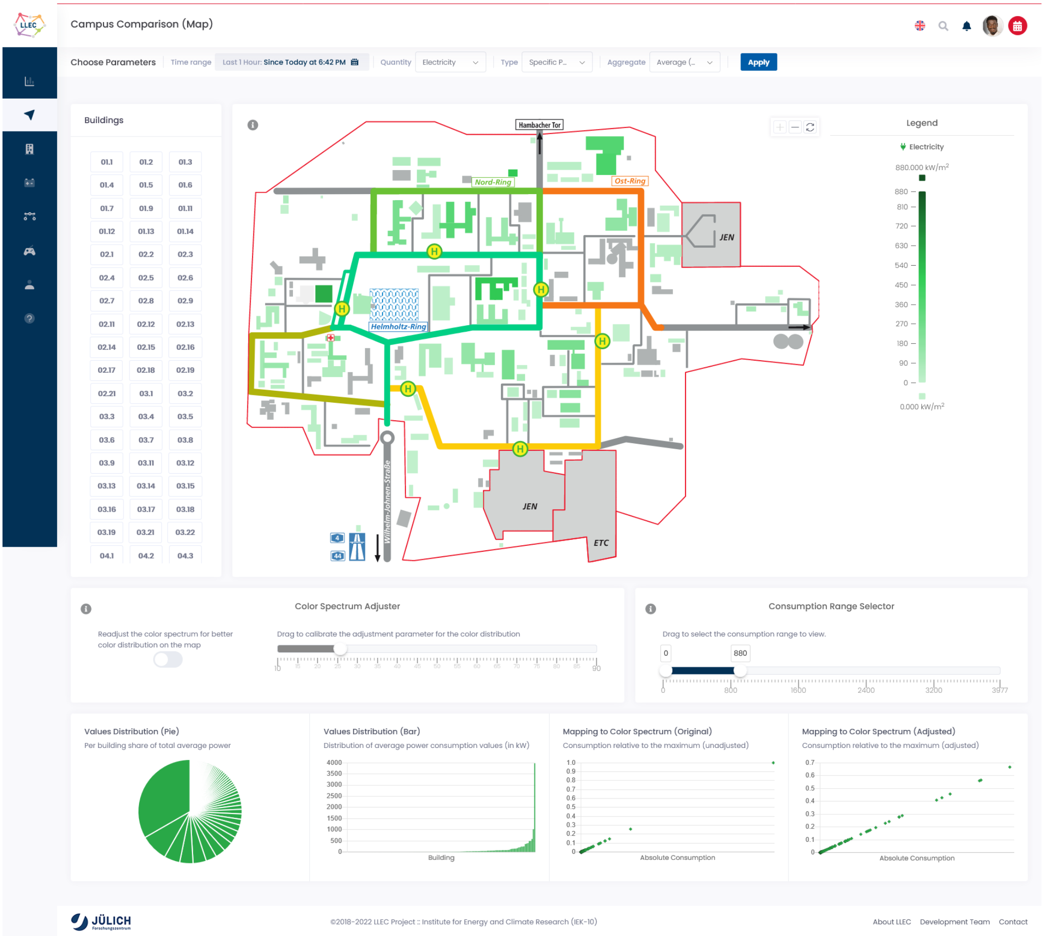


Fig. E.3 Campus Viewer building demand comparison page.

E.2 JuControl Screenshots

Fig. E.4 JuControl "My Building" page.

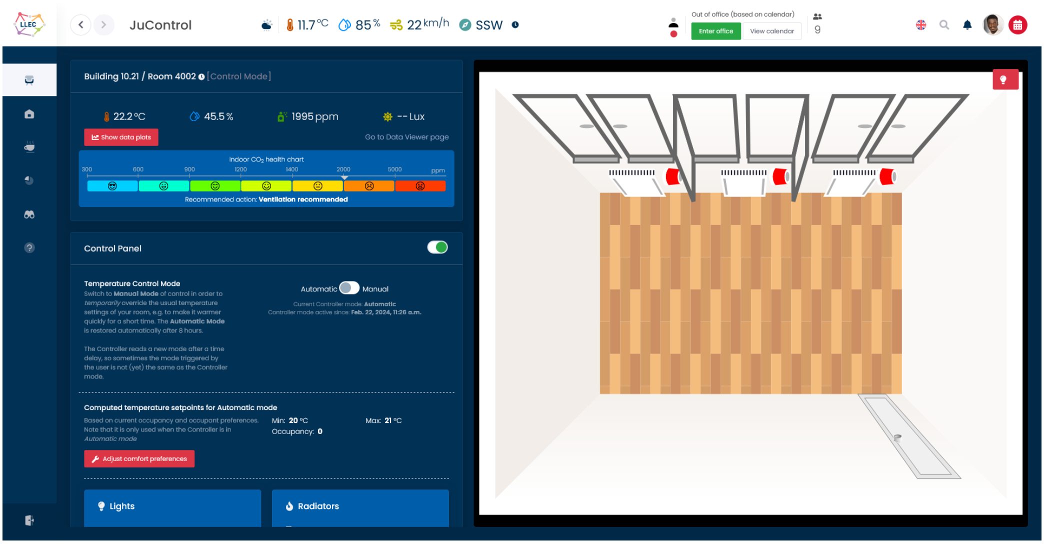


Fig. E.5 JuControl "My Room" page.

Fig. E.6 JuControl data visualization page.

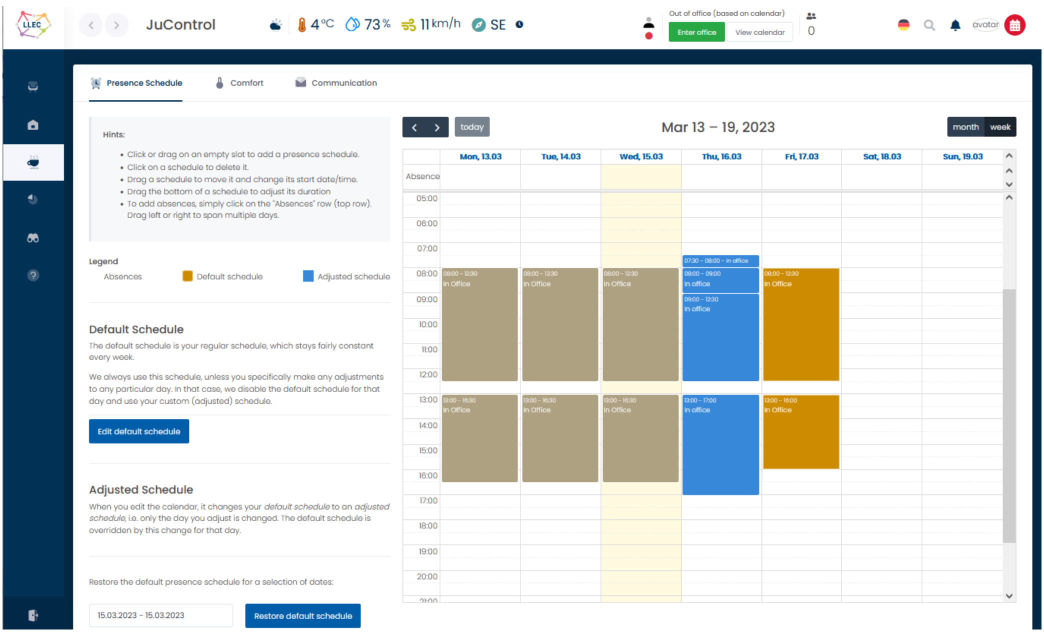


Fig. E.7 JuControl calendar page.

E.3 Screenshots of Juracle Results in JuControl

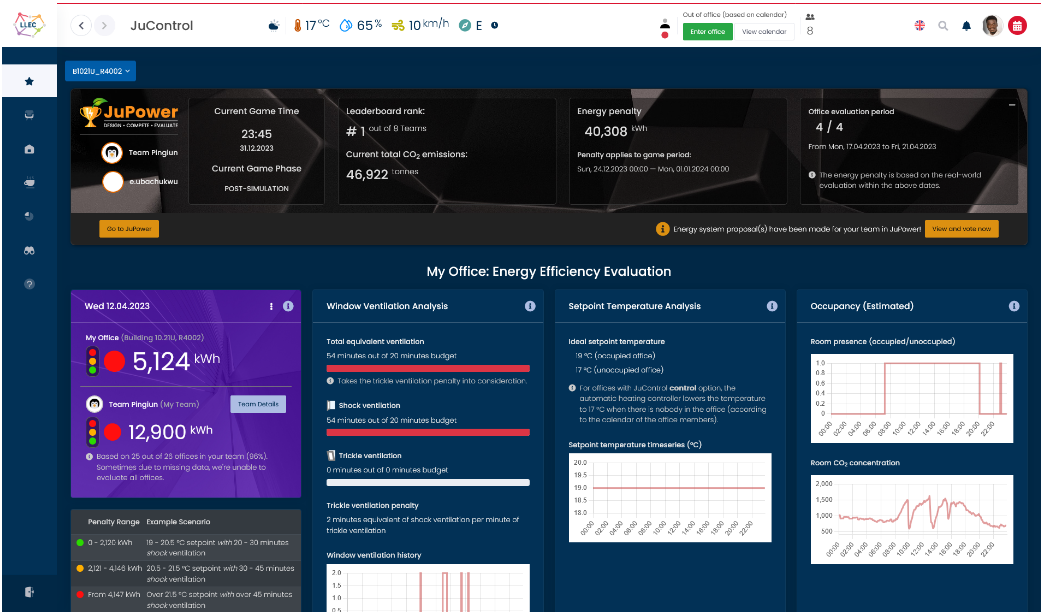


Fig. E.8 JuControl behaviour evaluation page showing Juracle-derived penalties.

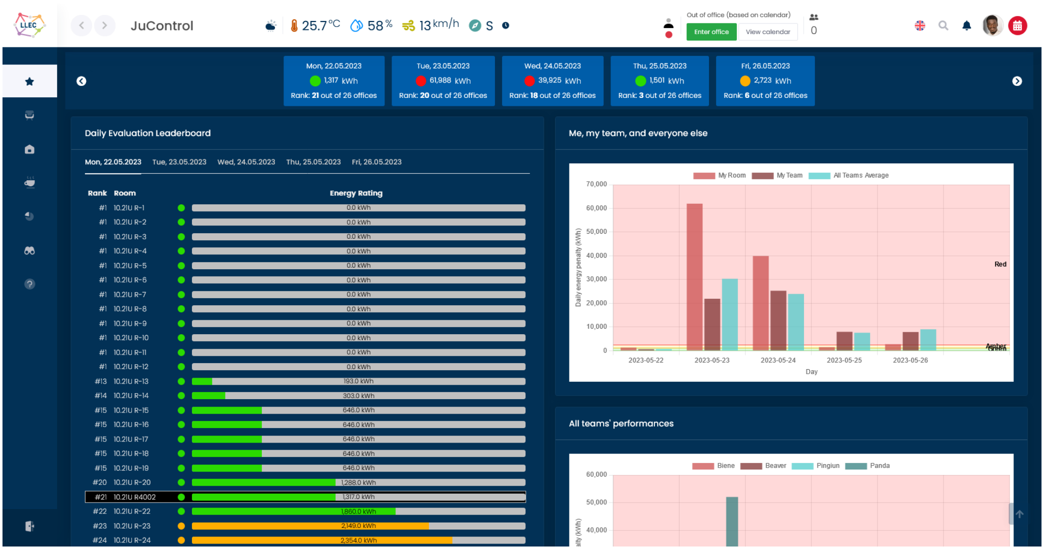


Fig. E.9 JuControl gamification page showing ranking of offices in a team.

E.4 ALICE Screenshots

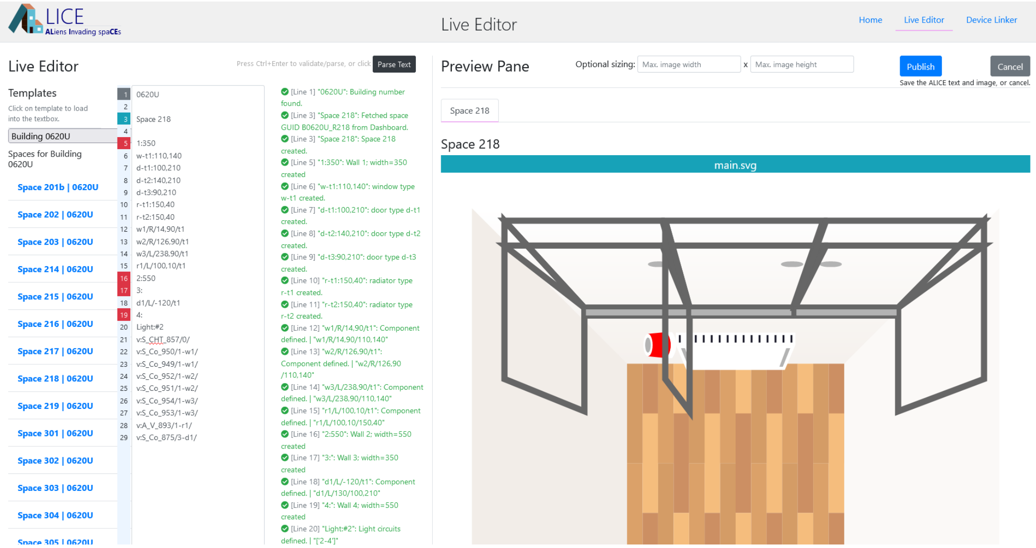


Fig. E.10 ALICE Editor page showing real-time validation of input and real-time visualization.

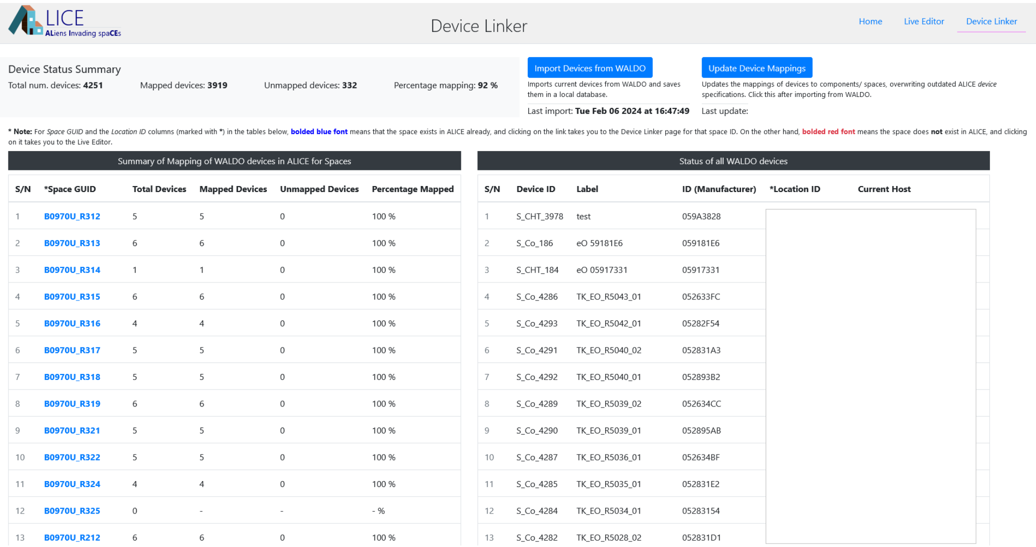


Fig. E.11 ALICE device linking page.

Band / Volume 657
Investigation of current and future anthropogenic chemical regimes in simulation chamber experiments
M. Färber (2025), 213 pp
ISBN: 978-3-95806-809-4

Band / Volume 658
Dynamischer Betrieb von Polymer-Elektrolyt-Membran Wasserelektrolyseuren
E. Rauls (2025), XIV, 239 pp
ISBN: 978-3-95806-811-7

Band / Volume 659
Pore-scale reactive transport modeling in cementitious materials: Development and application of a high-performance computing code based on the Lattice-Boltzmann method
S. Rohmen (2025), X, 295 pp
ISBN: 978-3-95806-812-4

Band / Volume 660
Recyclingmöglichkeiten für die Keramikkomponenten einer Festoxidzelle
S. Sarner (2025), VIII, 122 pp
ISBN: 978-3-95806-816-2

Band / Volume 661
Methodological Approach Enabling the Two-phase Flow Investigation in Alkaline Electrolysis under Demanding Conditions
S. Renz (2025), IX, 252 pp
ISBN: 978-3-95806-821-6

Band / Volume 662
Variable renewable energy potential estimates based on high-resolution regional atmospheric modelling over southern Africa
S. Chen (2025), XIII, 141 pp
ISBN: 978-3-95806-822-3

Band / Volume 663
Advances in Understanding Nitrate Aerosol Formation and the Implications for Atmospheric Radiative Balance
A. Milousis (2025), 195 pp
ISBN: 978-3-95806-823-0

Band / Volume 664
Optimization of NaSICON-type lithium- ion conductors for solid-state batteries
A. Loutati (2025), viii, 104 pp
ISBN: 978-3-95806-824-7

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment

Band / Volume 665
Innovative Plasma Sprayed Thermal Barrier Coatings for Enhanced Flexibility in Gas Turbine Operation
J. Igel (2025), V, 153, XXXVI pp
ISBN: 978-3-95806-827-8

Band / Volume 666
Techno- ökonomisches Potenzial dezentraler und autarker Energiesysteme
S. K. A. Risch (2025), xxiii, 210 pp
ISBN: 978-3-95806-829-2

Band / Volume 667
Reactive Field Assisted Sintering of Novel Rare Earth Garnets for Plasma Etching Applications
C. Stern (2025), VII, 101, XXVIII pp
ISBN: 978-3-95806-833-9

Band / Volume 668
Effects of mucilage and extracellular polymeric substances on soil gas diffusion
A. Haupenthal (2025), v, 99 pp
ISBN: 978-3-95806-834-6

Band / Volume 669
Quantifying Recombination Losses and Charge Extraction in Halide Perovskite Solar Cells
L. Krückemeier (2025), vi, 286 pp
ISBN: 978-3-95806-835-3

Band / Volume 670
Investigation of Dynamic Material Changes During the Preparation of ZnPd Nanoparticles Supported on ZnO and their Catalytic Application in Methanol Steam Reforming on the Atomic Level
A. Meise (2025), xviii, 175 pp
ISBN: 978-3-95806-838-4

Band / Volume 671
Improving Energy Efficiency of Public Buildings by Influencing Occupant Behaviour using Dashboards and Gamification
E. Ubachukwu (2025), xxi, 191 pp
ISBN: 978-3-95806-840-7

Weitere ***Schriften des Verlags im Forschungszentrum Jülich*** unter
http://wwwzb1.fz-juelich.de/verlagextern1/index.asp

Energie & Umwelt / Energy & Environment
Band / Volume 671
ISBN 978-3-95806-840-7

JÜLICH
Forschungszentrum