Doctoral theses at NTNU, 2024:171

Masab Khalid Annaqeeb

Energy-related occupant behavior in buildings: Approaches for monitoring and modelling

Doctoral thesis

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Engineering
Department of Energy and Process Engineering

NTNU
Norwegian University of
Science and Technology

Masab Khalid Annaqeeb

Energy-related occupant behavior in buildings: Approaches for monitoring and modelling

Thesis for the Degree of Philosophiae Doctor

Trondheim, May 2024

Norwegian University of Science and Technology
Faculty of Engineering
Department of Energy and Process Engineering

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Engineering
Department of Energy and Process Engineering

ISBN 978-82-326-7934-8 (printed ver.)
ISBN 978-82-326-7933-1 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2024:171

Printed by NTNU Grafisk senter

Dedicated to my Baba: His hard work and determination gave the foundation on which his kids could build their dreams. His compassion and kindness made us human so we could dream.

PREFACE

This thesis has been submitted to the Faculty of Engineering Science of the Norwegian University of Science and Technology in partial fulfilment of the requirements for the degree of Doctor of Philosophy (PhD).

This work was carried out at the Department of Energy and Process Engineering, Norwegian University of Science and Technology, Trondheim, Norway, during the period of September 2018 to December 2023. The work was under the supervision of Professor Vojislav Novakovic from the Norwegian University of Science and Technology and the co-supervision from Professor Da Yan from Tsinghua University in Beijing, China, and Professor Arild Gustavsen from the Norwegian University of Science and Technology.

This PhD project was under financial support by Strategic Research Area NTNU Energy through a specific call for collaboration projects within the energy field between NTNU and Shanghai Jiao Tong University or Tsinghua University (of 12. June 2017).

The project was carried out under the scientific umbrella of the Research Centre on Zero Emission Neighborhood in Smart Cities (FME ZEN) realized under financial support of the Research Council of Norway through project number 257660.

Masab Khalid Annaqeeb

Trondheim, December 2023

PREFACE

ACKNOWLEDGMENTS

When I look back on the last five years of work and efforts that went into this doctoral thesis, I see many highs and lows, I see moments of hesitation and uncertainty, I see determination and pursuit of academic knowledge, I see failed experiments and dejection, I see jubilation and successful outcomes. Above all, I see the multitudes of people along the way that contributed to this journey. It would be remiss of me to present this work without thanking all those that made it possible.

I would like to begin by expressing my thanks to my main supervisor, Professor Vojislav Novakovic for his guidance and contributions during my PhD. I am grateful not only for his advice on academic matters, but also his relentless support and encouragement since I started my doctoral studies. Furthermore, the access he provided to a broader network of academics and resources has been invaluable for my academic career. I also appreciate the personal and professional help he provided for me to adjust to life in Norway as a foreigner, especially the assistance in learning the language.

To my co-supervisor Professor Da Yan and his wonderful research group at Tsinghua University, I am highly thankful for all the academic discussions and the insights I gained during our collaborations and my research stay. Professor Yan’s assistance was central in framing the questions that shaped this thesis, and the hospitality that I received in China has given me an experience that I deeply cherish.

In addition, I would like to thank my co-supervisor Professor Arild Gustavsen and the group at FME ZEN that introduced me to several key concepts regarding sustainable neighborhoods. My gratitude extends to Professor Guangyu Cao, who provided the opportunity to connect occupant behavior research to another field, and to Liv-Inger Stensad at St. Olavs, who made it possible. To Associate Professor Laurent Georges, thank you for your course and discussions regarding building simulations, they were greatly helpful for developing the agent-based models used in this work. I would also like to thank the professors and staff at EPT, especially the administration, for helping me navigate the PhD programme. Special thanks also to Eugen and his team for helping me through the various technical challenges I encountered.

ACKNOWLEDGMENTS

One of the highlights during my PhD was the participation in the International Energy Agency's Energy in Buildings and Communities Programme (IEA EBC). I got to meet, learn, and collaborate with researchers from all parts of the world working on the theme of Occupant-centric Building Design and Operation through the Annex 79. I am highly grateful to the operators for fostering such a community of dedicated scientists, all of whom enriched my learning experience. It was through this Annex that I met Professor Laura Arpan, whose knowledge and guidance was vital for understanding and examining the social factors influencing people's behaviors. My heartfelt thanks to you Laura, for always being there to lend your time and advice on various parts of the research, and for your efforts and patience with my work over the past three years.

During the past five years, I had times where I was uncertain of my work and would struggle to find the motivation to continue. The effort took its toll several times, and I was lucky to have an excellent support system of colleagues and friends that tirelessly cheered me on, with their sage advice, academic insights, and helpful conversations.

My sincere thanks to my colleague and friend Anooshmita, whose dedication and work ethic has been a source of inspiration. I love how our discussions started on Indian street foods and resulted in successful research ideas and experiments.

I would like to express my gratitude to all my colleagues at EPT; to Jakub, whose work laid the foundation for this research; to Maria, one of the best officemates I could have, with her keen insights on academic research as well as everyday life; to Yiyu, who I could always count on to meet and unwind with at the brain center; and to Qiaoqiao, Haoran, Elyas, Juan, Xingji, Vegard, Mohammed, Yang, and my fellow DION board members, for their excellent camaraderie.

To my chosen family in Trondheim: Prerna, Samrridhi, and Raj, who gave me a desi home away from home; to Andreas and Alberto, who had to listen to me talk and rant about my research for quite some time, and to all the friends I made here, I am thankful for your friendship and support. To my lovely QRC friends: Aditya, Akanksh, Anandi, Gitanjali, Seff, Sol and Utti, you all brighten up my day with your conversations and consolations and love.

ACKNOWLEDGMENTS

To my amazing boyfriend Pshem, thank you for being who you are. Your dedication to science keeps my hopes in academia alive, and your constant love and support has made a world of a difference to me. I consider myself extremely lucky to have you

And to Vidya, you have been with me during this PhD journey since the day I applied for the position. I cannot find enough words to express my appreciation for all the things you did, the virtual writing sessions, the late-night panicked calls, the random bits of writing I would send to you for feedback, and the consistent emotional and mental support you gave.

I would not be able to be where I am or do what I have done without the constant support of my family. My brother Anas, you have been a constant cheerleader in everything I do, and I feel so blessed to have you in my life. My Farhana phuppu, who is like a second mother to me and whose hard work and care made it possible for me to get this far, I have the deepest appreciation. To my Qalajaan Asia and all my relatives for their care and believing in me, thank you.

Last but not the least, my heartfelt gratitude goes to my parents. My father Khalid Annaqeeb left no stone unturned to prioritize my education, with his constant encouragements to instill ambitions in me, and motivate me to achieve them. And my mother Anjum Sultana, who tirelessly worked days and nights, just to make sure that her children have the best possible care in the world. Her kindness and love shaped me into the person I am today, and for that I am very grateful. I love you both.

ACKNOWLEDGMENTS

ABSTRACT

Buildings have become one of the most energy-intensive sectors globally, contributing to about one-third of the worldwide energy consumption. The response to growing global concerns regarding energy-use has prompted the building sector to develop strategies for energy efficiency and investigate potential areas for optimizing building energy performance. The research efforts have highlighted several underlying factors that are attributed to this performance. Occupant Behavior (OB), or the way people interact with building systems, was found to have significant influence on building's performance, and further investigations revealed the lack of understanding the sector has regarding this phenomenon. Moreover, considerations of building occupants during the design stage rely on simplistic or deterministic approaches that do not provide an accurate representation of occupants. This challenge is also present in building performance simulation tools that are used to develop policies and provide recommendations for enhancing building energy efficiency. The lack of provisions for a dynamic human-building interaction is often attributed to the gap in expected and actual performance of buildings. The difficulties in addressing the topic arise from OB being a complex culmination of several phenomenon, including the occupant's presence, movement, actions, clothing, habits, social parameters and other contextual factors. Consolidated research efforts in this area, such as the International Energy Agency's Annex reports highlight the need for improving the accuracy of monitoring and modeling OB. Key recommendations also include the investigation and inclusion of social factors underlying OB, which are often neglected or underreported in energy research.

This thesis aims to provide a better understanding of OB by analyzing several aspects of it, in diverse contexts and settings. The work carried out can be categorized into three topics: OB Monitoring, OB Profiles, OB Modelling. The first one of these comprises of case studies using diverse methods to monitor and analyze occupants in different built environments. The second one focuses on the preparatory work needed for developing OB models, by creating occupant profiles, databases, and frameworks. The last one aims to provide insight and recommendations regarding the modelling and simulation part. The four research questions framed under these categories and their outputs can be summarized as follows:

ABSTRACT

How to monitor and analyze the impacts of occupant behavior in different spaces: Three distinct case studies were carried out to answer this question. The work consisted of using multiple sensing modalities to monitor occupant presence, energy-use, movement, and activities. Each study provided a unique perspective, with the first one using a combination of PIR and environmental sensors to collect and examine granular device-level, office-level, and occupant-level data in a shared office space. The second one made use of depth registration to monitor occupant influences in an operating room, and the last study focused on capturing an extensive activity-based dataset for smart homes. Since the lack of shared standards and benchmark datasets is another challenge in the development of guidelines and models, one of the goals of these studies was also to provide benchmark datasets that are available for a broader use.

How can the social aspects of OB be evaluated for modelling/simulation purposes: The research question was divided into three tasks, starting with the development of a hypothetical framework and extended Theory of Planned behavior model, which was useful in quantifying the social aspects. The second task put this model in implementation, by collecting data through surveys and analyzing the factors that influence people's behavior. The investigation process involved creating structural equation models, regressions, and path analysis. Regression was found to be the best fit, while path analysis was acceptable. Influence factors for occupant's energy-related behaviors were obtained for each building system. The addition of additional variables improved the predicting power of current models. The framework and evaluation of variables led to the creation of OB profiles, supported by analysis highlighting valid and significant variables. The k-modes clustering technique yielded suitable clusters for modelling purposes.

How can the social aspects of OB be modelled/simulated: An important part of this thesis was incorporating an interdisciplinary approach, connecting social behavioral theories to the engineering aspects of modeling and simulation. Addressing this question required assessing modeling techniques that would be best suited to include the diversity and flexibility required to include social aspects. Agent-based models were used to simulate this aspect, which proved capable of executing the already defined fixed and static models while also accommodating the diversity of OB.

ABSTRACT

What kinds of prerequisites are needed to map environmental layouts around the occupant: This question was part of a broader study aimed at developing a hypothetical OB model that attempts to address multiple modeling requirements for OB, a part of which was to simulate the environment around the occupant. The work carried out in this thesis aimed to map occupant environmental layouts in a room by identifying seven variables for a library and developing a Matlab application for spatial information collection. The library was tested using a sample dataset of 80 offices at NTNU. A database was created to connect the information generated from this study to larger models, which were expanded to include other datasets from previous studies.

By addressing these questions, this thesis was able to contribute to a broader understanding on the subject of occupant behavior, providing insights about the monitoring and modelling process, and highlighting additional challenges in the subject. The findings may contribute to a better design of building operation and management that places its occupants at the center.

ABSTRACT

LIST OF PUBLICATIONS

Paper 1:

Annaqeeb M K, Dziedzic J W, Yan D, Novakovic V. Exploring the tools and methods to evaluate influence of social groups on individual occupant behavior with impact on energy use. *Proceedings of the IOP Conference Series: Earth and Environmental Science*. 2019; 352 (1): 012044.

Paper 2:

Annaqeeb M K, Markovic R, Novakovic V, Azar E. Non-intrusive data monitoring and analysis of occupant energy-use behaviors in shared office spaces. *IEEE Access,* 2020; 8:141246-141257.

Paper 3:

Annaqeeb M K, Azar E, Yan D, Novakovic V. Evaluating occupant perceptions of their presence and energy-use patterns in shared office spaces. *Proceedings of the 16th Conference of the International Society of Indoor Air Quality and Climate: Creative and Smart Solutions for Better Built Environments*, Indoor Air 2020.

Paper 4:

Annaqeeb M K, Das A, Arpan L, Novakovic V. Evaluating and Modeling Social Aspects of Occupant Behavior in Buildings: An Agent-Based Modeling Approach. To be submitted to: Building and Research Information

Paper 5:

Annaqeeb M K, Dziedzic J W, Yan D, Novakovic V. Development of a Library for Building Surface Layout Simulator. *Proceedings of the 11th the International Symposium on Heating, Ventilation and Air Conditioning (ISHVAC).* 2019;1137-1144.

Paper 6:

Annaqeeb M K, Zhang Y, Dziedzic J W, Xue K, Pedersen C, Stenstad L I, Novakovic V, Cao G. Influence of surgical team activity on airborne bacterial distribution in the operating room with a mixing ventilation system: a case study at St. Olavs Hospital. *Journal of Hospital Infection*. 2021; 116:91-98.

Paper 7:

Das A, **Annaqeeb M K**, Azar E, Novakovic V, Kjærgaard M B. Occupant-centric miscellaneous electric loads prediction in buildings using state-of-the-art deep learning methods. *Applied Energy*. 2020; 269:115-135.

Paper 8:

Das A, Dziedzic J W, **Annaqeeb M K**, Novakovic V, Kjærgaard M B. Human Activity Recognition Using Sensor Fusion and Deep Learning Methods. *Submitted to IMWUT.*

Paper 9:

Das A, **Annaqeeb M K**, Schwee J H, Dziedzic J W, Novakovic V, Kjærgaard M B. Sequential Activity Recognition and Privacy Implications Using Fusion and Deep Learning Methods Inside A Smart Living Lab. *Submitted to Information Fusion*.

ABBREVIATIONS

ABM	Agent-Based Model
AoC	Awareness of Consequences
ADL	Activities of Daily Living
AGFI	Adjusted Goodness of Fit Index
AR	Activity Recognition
ASHRAE	American Society of Heating Refrigeration and Air-Conditioning Engineers
BMS	Building Management System
BOT-ABM	Building Occupant Transient-Agent Based Model
BPS	Building Performance Simulation
CFI	Comparative Fit Index
CFU	Colony Forming Units
DBMS	Database Management System
DL	Deep Learning
EBC	Energy in Buildings and Communities
EU	European Union
GFI	Goodness of Fit Index
GRU	Gated Recurrent Unit
HAR	Human Activity Recognition
HVAC	Heating Ventilation and Air-Conditioning
IEA	International Energy Agency
IN	Injunctive Norms
LSTM	Long-Short Term Memory
MAE	Mean Absolute Error
ML	Machine Learning
NILM	Non-Intrusive Load Monitoring
NTNU	Norwegian University of Science and Technology

ABBREVIATIONS

OB	Occupant Behavior
OR	Operating Room
PBC	Perceived Behavioral Control
PIR	Passive Infra-Red
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analysis
RDBMS	Relational Database Management System
RFI	Relative Fit Index
RGB	Red-Green-Blue (refers to traditional cameras)
RMSEA	Root Mean Square Error of Approximation
SEM	Structural Equation Model
SPSS	Statistical Package for Social Sciences
SQL	Structured Query Language
SRMR	Standardized Root Mean Square Residual
TPB	Theory of Planned Behavior
ZEB/nZEB	Zero Energy Building/Nearly Zero Emission Building

ABBREVIATIONS

LIST OF CONTENTS

LIST OF CONTENTS

LIST OF TABLES

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES

1 INTRODUCTION

The rise in global energy-use has led to worldwide concerns and calls for sustainable transitions across several disciplines. The building industry is a significant contributor to this consumption, and as such, sustained effort has been directed to implement energy efficient technologies in the whole life cycle of buildings. An integral part of the building are its users, or occupants, and the interactions between them and the building systems have a major impact on the building's energy consumption. This interaction has been the focus of several research efforts during the last decade, and there still exist several gaps in the understanding of this human-building interaction. This thesis aims to explore aspects of this interaction, to provide understanding of occupant behaviors, and lay the groundwork for developing models of the same.

1.1 Motivation

The last few decades have seen buildings emerge as one of the most energy-intensive sectors with regards to global energy consumption, contributing to about one-third of the total figure [1]. The same sector has also been responsible for 39% of the nationwide consumption in the United States, and 40% on average in countries part of the European Union [2][3]. These numbers are projected to be increasing at a higher rate as well. In the global context, there is a need to shift towards strategies that drive a more sustainable future, which is perhaps best reflected in the United Nations Sustainable Development Goals, that advocate for more energy efficient technologies and responsible consumption [4]. This has led to an increase in efforts to optimize and bolster energy efficiency in buildings, which have taken the form of certification programs, energy codes, financial and other incentives etc. The advent of low-energy or zero-energy buildings contribute to the same cause as well. However, in operation, it has been found that buildings far exceed their expected consumption, and this gap is more pronounced in low-energy buildings [5][6]. In addition, studies have indicated that building energy consumption is driven by several diverse factors, such as engineering technology, geographical context, local climate, cultural background, social equity, occupant behavior (OB), and more [7,8]. In order to advance energy-efficient strategies, it becomes imperative to have an understanding of the underlying factors.

INTRODUCTION

Over the past decade, a central underlying theme has surfaced to the forefront, due to the consolidated efforts of several global researchers and scientists. The theme is that of occupant behavior, that is, the way that users/occupants interact with building systems [9]. The International Energy Agency (IEA) has had multiple annexes devoted to the task, under the Energy in Buildings and Communities (EBC) program, wherein researchers have studied, analyzed, and documented this branch of study in detail. It has been evident that OB plays a defining role in influencing the total consumption of buildings. This role has been documented by several studies that demonstrated the differences in OB that can lead to significant changes in the resulting building energy consumption, whether it is through monitoring different occupants, or developing energy models with different occupant profiles [10][11]. This behavior, however, is multifaceted in terms of the interaction with different kinds of building systems, varied building physics and available technology, and needs to be studied accordingly. Moreover, efforts directed at building energy management have adopted approaches that classify the building into certain seasonal or temporal cycles, or gauge the end-use of different building types, with the goal of reducing or optimizing the energy consumption. Such strategies have been met with counter proposals that argue for a more holistic approach that considers each influential component of the building system to arrive at a sustainable building design. This approach gives space for including OB, which is one of the main sources of uncertainty in building energy consumption [12]. In addition, discrepancies between the expected and actual performance of buildings can often be attributed to oversimplification in occupant's presence and schedules [13].

A significant approach to tackle these issues is by using building performance simulation (BPS) programs to identify the needs and schedules of the users and efficiently improve the design and operation of buildings. Coupled with simulations, recommendations regarding energy saving measures and improving comfort of the users benefit occupants as well as the building facility managers. In addition, the results of these simulations provide better decision support, including policy support, early phase design support and multi-scale approaches from construction detail to district level. These programs include the modelling and evaluation of different systems in buildings, such as the thermal or electric systems [14]. The models in simulation sciences have had considerable advances in the past few decades. Physical phenomena are being represented by increasingly sophisticated algorithmic domain models. Geometrical representation of the building components, elements etc. has achieved substantial progress [15], and so has the representation of dynamic boundary conditions. Compared to these developments, implemented models of people in buildings have been rather simplistic, making use of fixed and static schedules, which are not accurately representative of OB [16]. In addition, the lack of shared methods and standards represents another challenge in this field. These challenges invoke the need for sustained research and development efforts for improving OB models.

As identified in the IEA EBC Annex 66 report, current research needs to be directed at improving the accuracy of monitoring occupants and simulating procedures involving OB [17]. Key directions highlighted by the experts in the community include occupant monitoring and data collection, model development, model evaluation, and integration into building simulation tools. This thesis aims to provide insight into the first two directions. The first of which is to conduct case studies of occupants in different settings of the built environment, such as private/shared office spaces, residential areas, and specialized zones like hospital rooms. These studies use a range of data acquisition techniques to extract information about several different aspects of occupant behavior, ranging from movement, presence, activities, and energy-use behaviors.

The second key direction concerns OB modeling, which has its complications based in the diverse set of actions as well the different aspects of the OB itself. The complexity and uncertainty in this field stem from the fact that OB contains various similitudes in the form of presence, movement, activity level, comfort level, social influences etc., and detailed attention has to be given to each of these in order to construct a complete individual profile. Furthermore, OB models tend to focus on the physical occupant interactions, neglecting the socio-psychological factors that drive them [18]. This thesis aims to provide a bridge between social behavioral theories and OB models, by identifying influencing factors and datasets that need to be used to prepare profiles for a proposed agent-based modelling approach. The studies in this thesis also focus on other aspects of OB such as their movement and surroundings, and the preparatory work and profiles needed in order to construct the models.

This thesis is also a part of the overarching work proposed by Dziedzic [19], wherein a bottom-up approach would be used to combine a series of simulators based on different aspects of OB. Such an undertaking is vast in scope and would need significant effort directed at each relevant topic, taking into consideration the datasets that need to be collected for each OB aspect, the sensing modalities to be used, the privacy implications, suitable data processing techniques, model development and validation, and so on. Through the questions raised and studies undertaken, this thesis also aims to contribute to the body of knowledge required to build such a holistic model.

1.2 Research questions and research tasks

Occupant behavior is a result of diverse, multidisciplinary phenomenon, and the approach to modelling OB needs to consider several different aspects of it. This thesis aims to understand and analyze OB with regards to several diverse contexts and settings. The three main areas explored in the study were related to:

(i) OB monitoring: conducting case studies that use diverse modes of data acquisition to monitor and analyze OB in various settings of the built environment

(ii) OB Profiles: Collecting occupant behavioral data with regards to their environmental layouts, social influences, movement, and energy-use habits, in order to generate profiles for, and conduct preparatory work for OB modeling

(iii) OB Modeling/Simulation: Develop an agent-based model based on (2) to simulate the social aspect of OB.

Figure 1-1 provides a visual description of the three research areas in focus.

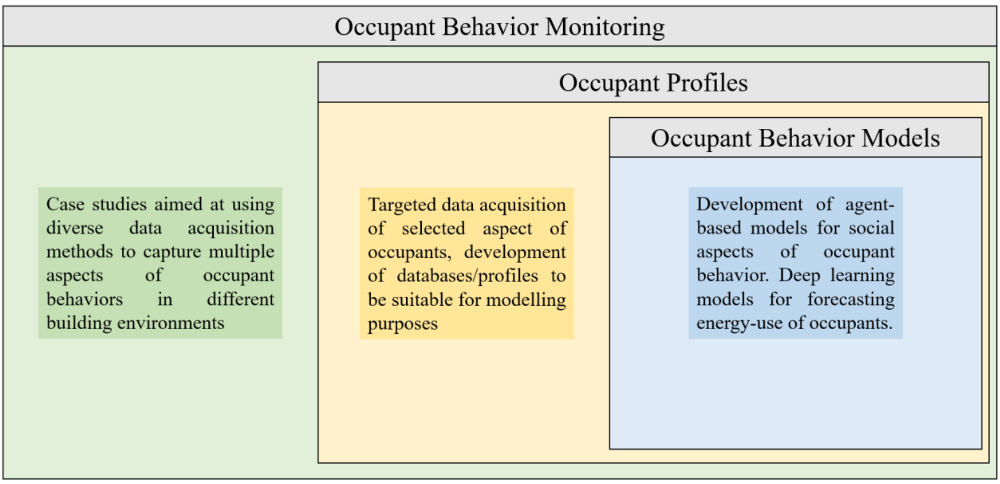

Figure 1-1. Research Areas

The following four questions were considered to achieve the objective of this thesis, each accompanied by a set of tasks outlined to answer them:

Question 1. How to monitor and analyze the impacts of occupant behavior in different spaces?

Task 1.1: Case study monitoring and analyzing occupant presence, and energy-use patterns in shared office spaces

Task 1.2: Case study in Living Lab to collect data regarding activities of daily living

Task 1.3 Case study on influence of OB in operating rooms

Question 2. How can the social aspects of OB be evaluated for modelling/simulation purposes?

Task 2.1: Perform a literature review to understand the extent of social influences and the tools/methodologies that are used to assess it. Develop a framework for measuring these influences in the context of energy-related occupant behavior in buildings.

Task 2.2: Conduct a study to collect data regarding social aspects of OB and analyze it to identify significant factors

Task 2.3 Analyze different factors and clustering strategies to develop behavioral profiles to be used for modelling purposes

Question 3. How can the social aspects of OB be modelled/simulated?

Task 3.1: Develop agent-based models using the profiles generated from (Q2) to conduct a proof-of-concept study for OB modelling.

Task 3.2: Develop ABMs targeted at diversity in OB models

Question 4. What kinds of prerequisites are needed to map environmental layouts around the occupant?

Task 4: Analyze variables required to construct a library for a zone/floor layout simulator. Develop an application for geometric data collection and conduct a study for developing a database for zone layout simulator.

Figure 1-2 categorizes the research questions and task into their respective research areas.

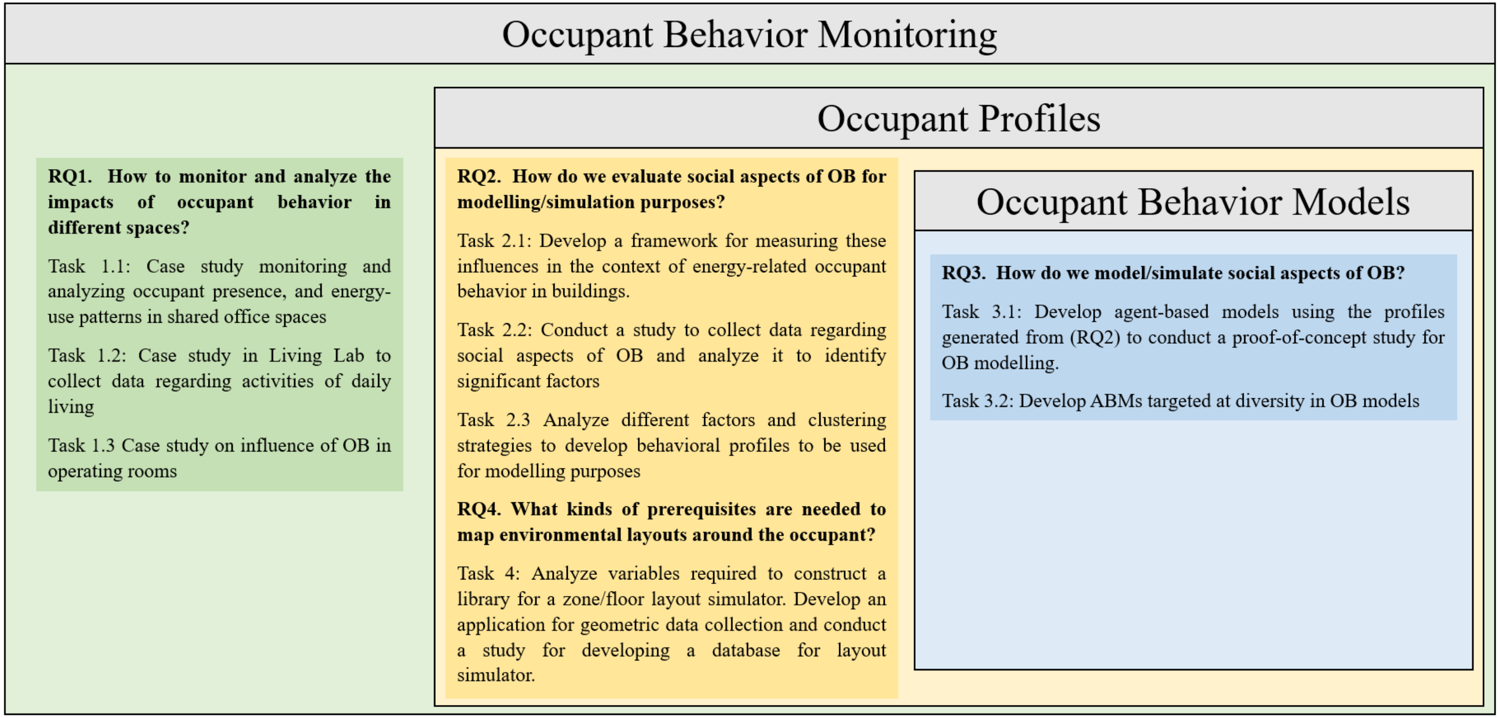

Figure 1-2. Research questions within respective research areas

1.3 Structure of the thesis

The rest of the thesis is structured as follows: Chapter 2 introduces the background of the different topics covered in the study, in an order that explains the progression of the thesis. The next three chapters follow the theme that is highlighted in Figure 1-1. As shown in the figure, there are three main areas of research covered in the study (OB monitoring, OB profiles and preparatory work, OB modelling), and each of the chapters is segregated based on these themes. Chapter 3 explains the methodology for the OB monitoring case studies, the preparatory work for OB models, the investigation and evaluation of the social aspects, and finally the models developed. Chapter 4 then presents the results for each of these sections. Chapter 5 outlines the main conclusions, provides the key limitations, and the considerations for the future work.

1.4 List of publications

The foundation of this thesis is built on a collection of nine papers, six of which were journal articles, and the remaining three were part of conference proceedings. An overview of these articles is presented in Figure 1-3. Here, they are distinguished between *primary papers* and *supporting papers,* wherein the *primary papers* address the key research questions, and *supporting papers* provide preparatory work for the former, along with additional insights around the subject of Occupant Behavior. Each paper addressed one or more of the following five different aspects of Occupant Behavior (OB):

- Presence: Having to do with the occupant's presence within an area
- Social: Effects and influences that the occupant has from their beliefs, perceptions, or other people around them
- Environmental: Interaction of the occupant with their immediate surroundings, or influences that an occupant can have on them.
- Activity: General activities that an occupant undertakes, specifically categorized based on the zone/room they occupy.
- Energy-use: More specific activities that directly impact energy-use, such as plug loads, lighting, heating/cooling, or intention to perform these.

The papers included in the thesis are listed below, along with the author's contributions to each:

Primary Papers

Paper 1

Annaqeeb M K, Dziedzic J W, Yan D, Novakovic V. Exploring the tools and methods to evaluate influence of social groups on individual occupant behavior with impact on energy use. *Proceedings of the IOP Conference Series: Earth and Environmental Science*. 2019; 352 (1): 012044.

Contribution: This paper presented a review of the social features in energy-related OB, and their significance. The author was responsible for preparing the conceptual framework, conducting the literature search, and original draft. Jakub Dziedzic, Da Yan and Vojislav Novakovic reviewed and commented on the article.

- Review of social features in energy-related OB, and their significance
- Developed conceptual framework on tools to assess social aspect of OB
- Preliminary factors identified to investigate topic
- *Aspects of OB addressed*: **Social**

Paper 2

Annaqeeb M K, Markovic R, Novakovic V, Azar E. Non-intrusive data monitoring and analysis of occupant energy-use behaviors in shared office spaces. *IEEE Access*, 2020; 8:141246-141257.

Contribution: This paper developed a framework for monitoring OB with regards to energy-use in shared offices. The conceptualization was done in collaboration with Elie Azar. The author conducted the experiment, data collection and processing, formal analysis, and wrote the original draft. Data visualization was done with Romana Markovic. Elie Azar Vojislav Novakovic reviewed and commented on the work.

- Framework for monitoring OB with regards to energy-use in shared offices
- Case-study using sensors to monitor OB, environmental parameters and Miscellaneous Electric Loads (MELs)
- Analysis of MEL energy-use behaviors
- *Aspects of OB addressed*: **Presence | Environment | Energy-use**
- *Methods*: **Sensors**

Paper 3

Annaqeeb M K, Azar E, Yan D, Novakovic V. Evaluating occupant perceptions of their presence and energy-use patterns in shared office spaces. *Proceedings of the 16th Conference of the International Society of Indoor Air Quality and Climate: Creative and Smart Solutions for Better Built Environments*, Indoor Air 2020.

Contribution: This paper is based on the conceptual framework developed in Paper 1, and parts of its datasets are from the experiment conducted in Paper 2. These were used to develop profiles for occupant presence and energy-use and perform a comparative analysis. The conceptualization was done with Vojislav Novakovic. The author conducted the formal analysis, visualizations, and wrote the original draft. Elie Azar, Da Yan, and Vojislav Novakovic reviewed and commented on the paper.

- Use of conceptual framework from Paper 1
- Developed profiles for occupant presence and energy-use
- Comparative analysis of OB profiles (standard, perceptual, actual)
- *Aspects of OB addressed*: **Presence | Social | Energy-use**
- *Methods*: **Surveys | Sensors**

Paper 4

Annaqeeb M K, Das A, Arpan L, Novakovic V. Evaluating and Modeling Social Aspects of Occupant Behavior in Buildings: An Agent-Based Modeling Approach (To be submitted to Building Research and Information)

Contribution: This paper was an interdisciplinary study connecting social behavioral theories with engineering applications. The idea was developed with Laura Arpan, Anooshmita Das and Vojislav Novakovic. The author, along with Laura Arpan, developed the theoretical model and survey design, who also guided the data analysis and structural equation modeling. The data collection, processing, analysis, modelling, and writing were conducted by the author. Anooshmita Das developed the profiles and assisted in the data collection and analysis. All the authors reviewed and commented on the article.

- Proof-of-concept study in using Agent-Based Models for social models in OB
- Investigating influence factors for social aspects in OB
- Developing profiles for social models
- *Aspects of OB addressed*: **Social | Energy-use**
- *Methods*: **Surveys**

Paper 5

Annaqeeb M K, Dziedzic J W, Yan D, Novakovic V. Development of a Library for Building Surface Layout Simulator. *Proceedings of the 11th the International Symposium on Heating, Ventilation and Air Conditioning (ISHVAC).* 2019;1137-1144.

Contribution: This paper was in conjunction with *Additional Publications: Paper 1*, providing supplementary work to build a simulation for zone layouts. The conceptualization and data collection were done with Jakub Dziedzic. The author was responsible for developing the application and writing the original draft. Da Yan and Vojislav Novakovic reviewed and commented on the paper.

- Developed application for geometric data collection for surface layout simulator
- Database for furniture layouts
- *Aspects of OB addressed*: **Environment**

Paper 6

Annaqeeb M K, Zhang Y, Dziedzic J W, Xue K, Pedersen C, Stenstad L I, Novakovic V, Cao G. Influence of surgical team activity on airborne bacterial distribution in the operating room with a mixing ventilation system: a case study at St. Olavs Hospital. *Journal of Hospital Infection*. 2021; 116:91-98.

Contribution: The concept and experiment design were developed by Jakub Dziedzic, Vojislav Novakovic and Guangyu Cao. Dziedzic also developed the data capture and processing techniques for the experiment. Christoffer Pederson, Kai Xue, and Yixian Zhang assisted in performing the experiment, where Zhang analyzed the bacterial contamination of the samples. Liv Inger Stensad provided technical assistance for the equipment at St. Olavs hospital. The author was responsible for part of data collection, data analysis, visualizations, and writing the original draft. Vojislav Novakovic and Guangyu Cao reviewed and commented on the paper.

- Case study on analyzing human activity on bacterial contamination in ORs
- *Aspects of OB addressed*: **Movement | Activity | Environment**
- *Methods*: **Sensors**

Supporting Papers

Paper 7

Das A, **Annaqeeb M K**, Azar E, Novakovic V, Kjærgaard M B. Occupant-centric miscellaneous electric loads prediction in buildings using state-of-the-art deep learning methods. *Applied Energy*. 2020; 269:115-135.

Contribution: This paper used the dataset from Paper 2 to develop deep learning models to forecast occupant energy-use. Anooshmita Das created the models and the visualizations. The author performed the data collection and assisted in the analysis and writing. Elie Azar, Mikkel Kjærgaard and Vojislav Novakovic reviewed and commented on the paper.

- Same study as used in Paper 2

- Forecasted MELs, focusing on occupants and their energy-use behaviors
- *Aspects of OB addressed*: **Presence | Energy-use**
- *Methods*: **Sensors, Deep Learning**

Paper 8

Das A, Dziedzic J W, **Annaqeeb M K**, Novakovic V, Kjærgaard M B. Human Activity Recognition Using Sensor Fusion and Deep Learning Methods. *Submitted to IMWUT.*

Contribution: This was an extensive case-study conducted in ZEB Living Lab, about human activities inside residential spaces, the results of which were documented in Paper 8 and Paper 9. The experiment design was developed by Jakub Dziedzic and Anooshmita Das. The author assisted in the data collection, processing the skeleton datasets, and writing the original draft. Mikkel Kjærgaard and Vojislav Novakovic reviewed and commented on the paper.

- Case-study of monitoring human activity in a living lab, and develop HAR models
- *Aspects of OB addressed*: **Activity | Energy-use**
- *Methods*: **Sensors, Deep Learning**

Paper 9

Das A, **Annaqeeb M K**, Schwee J H, Dziedzic J W, Novakovic V, Kjærgaard M B. Sequential Activity Recognition and Privacy Implications Using Fusion and Deep Learning Methods Inside A Smart Living Lab. *Submitted to Information Fusion*.

Contribution: This paper was the second part of the results from the case study done at the Living Lab. It focused more on the privacy risk and used additional deep learning models developed by Anooshmita Das. The author contributed to the data collection, analysis, literature search, and writing. Mikkel Kjærgaard and Vojislav Novakovic reviewed and commented on the paper.

- Same as above
- *Aspects of OB addressed*: **Activity | Energy-use**
- *Methods*: **Sensors, Deep Learning**

The abovementioned studies were carried out with the aim of laying the groundwork for OB models. As such, each study was conducted with one or more of the three goals: (i) OB Monitoring (ii) Developing OB Profiles/Databases (iii) OB Models

The following table summarizes the contribution of each paper, along with indicating the goal of the study:

Research Papers	**Occupant Behavior Studied:**					**Studied in relation to:**		
	Presence	Environment	Activity	Social	Energy-use	OB Monitoring	OB Profiles/ Databases	OB Models
Paper 1				✓			✓	
Paper 2	✓	✓			✓	✓		
Paper 3	✓			✓	✓		✓	
Paper 4				✓	✓		✓	✓
Paper 5		✓					✓	
Paper 6		✓	✓			✓		
Paper 7					✓	✓		✓
Paper 8					✓			
Paper 9					✓			✓

Table 1-1. Contribution of research papers in respective areas of research

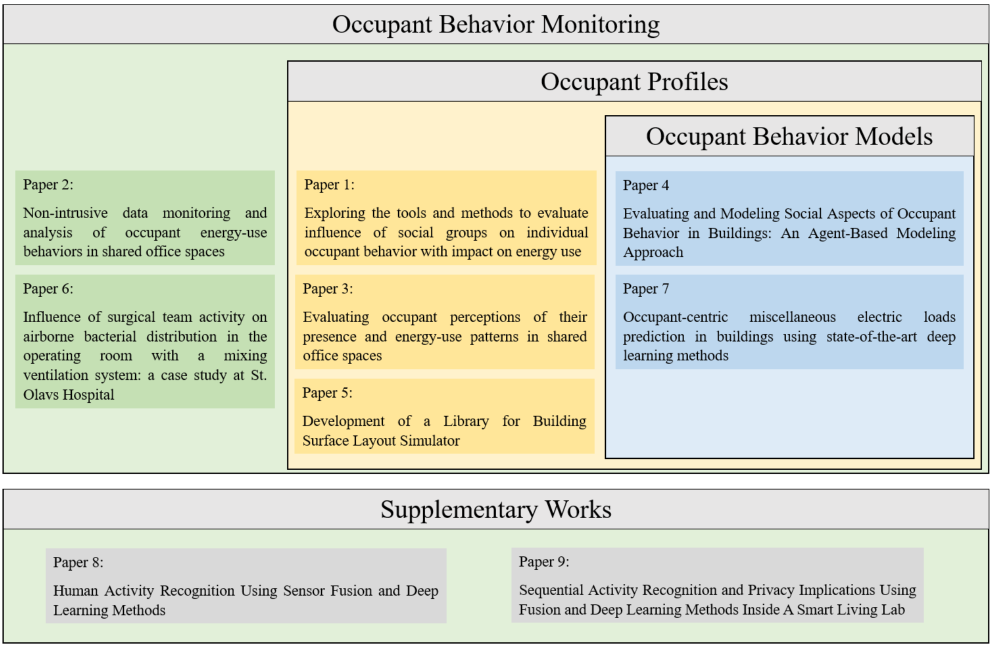

Figure 1-3. Overview of papers considered in this thesis

Additional publications*:*

Collaboration with the IEA Annex-79 and other researchers led to the following additional publications:

Conference contributions:

1. Dziedzic J W, **Annaqeeb M K**, Yan D, Novakovic V. Zone Layout Simulator for energy-related occupant behavior modelling. *Proceedings of the 11th the International Symposium on Heating, Ventilation and Air Conditioning (ISHVAC).* 2019

Other contributions:

2. Carlucci, S., De Simone, M., Firth, S.K., Kjærgaard, M.B., Markovic, R., Rahaman, M.S., **Annaqeeb, M.K**., Biandrate, S., Das, A., Dziedzic, J.W. and Fajilla, G., 2020. Modeling occupant behavior in buildings. *Building and Environment*, *174*, p.106768.

3. Markovic, R., Azar, E., **Annaqeeb, M.K.**, Frisch, J. and van Treeck, C., 2021. Day-ahead prediction of plug-in loads using a long short-term memory neural network. Energy and Buildings, 234, p.110667.

INTRODUCTION

INTRODUCTION

2 BACKGROUND

2.1 Occupant Behavior in Buildings

In a rapidly urbanizing world, humans spend a vast majority of their lifetimes living and working in buildings [20]. They play a central role in the performance of the buildings due to their presence and actions, including the use of heating, cooling, water, electricity, and other interactions with the building systems. Such interactions, collectively termed as Occupant Behavior (OB) include the use of the building's heating, cooling, and ventilation (HVAC) systems to establish thermal comfort, use of shading devices and lighting systems for visual comfort, appliance operation, etc. The first point of consideration of occupants comes from the building design perspective. Appropriate understanding of the buildings' occupants is necessary to define control settings for its systems. One of the first studies considering the active role of an occupant in the building's performance took place in 1979, where Hunt introduced a function for the lighting systems in offices [21]. Subsequent studies explored the role of the occupant with regards to either lighting systems [22,23] or window operation [24–26], providing insights about occupancy control actions.

The significance of OB from these niche studies became more apparent during the last decade, with increasing number of studies documenting the impact that users have on energy-use in buildings [10,27–29]. Masoso et al. pointed out the way in which user's actions would negatively impact energy-use in commercial buildings, leading more energy used during non-working hours than during working hours [30]. In another study, Hong et al. created different profiles for offices and created building energy models based on them. The authors found that mindful behaviors can save up to 50% of the current energy-use and heedless behavior can increase it by 89% [11]. Similar studies conducted in residential buildings reflect the same impact of OB, where the energy-use varies largely based on the occupants, despite the building having the same or comparable physical characteristics and systems [31]. These findings span across different geographic and cultural boundaries as well [32,33].

The impact of building energy-use and its contributions to carbon emissions around the globe has been met with calls for transitions towards low-energy or Zero Energy Buildings (ZEBs or nZEBs). This transition has been adopted as a goal by the European Union (EU) parliament, with several countries following suit [34]. However, while there exist a pronounced gap in the expected and actual energy performance of buildings [35], this gap can be wider in low-energy buildings if the theoretical assumptions about the user behaviors is not accurate. Carpino et al. demonstrated the effects of an occupant's actions on nZEBs, which might generate much higher loads if the assumptions made during the building energy simulations do not prove to be accurate [36]. In fact, a large-scale study in Netherlands considered 200,000 dwellings and found that dwellings that were designed to be energy-efficient ended up using more energy than predicted, whereas dwelling with poor energy labels used less energy than predicted by the label [37]. The implication here is not the shortcomings of technological advancements but highlighting the extended role of OB in these situations. Since the physical characteristics and the building envelope is optimized for a better energy performance, the uncertainty of the loads generated by the occupant now has a greater influence on the outcome.

However, the uncertainties and complexities arising from this phenomenon are not easy to mitigate. The challenges exist because OB is multifaceted in terms of its actions, which span from direct adaptive actions to establish or restore comfort, to indirect and supporting actions such as presence and movement. Furthermore, these are also driven by several underlying factors that vary depending on occupant, building type, contextual factors etc. An accurate representation of an occupant within a building will have to consider several aspects of OB and these aspects, drivers and complexities are detailed in the next section.

2.2 Drivers, aspects, and complexities of OB

The consolidated efforts of scientists related to occupant behavior research have been documented in multiple annexes of the International Energy Agency (IEA) under the program of Energy in Buildings and Communities (EBC). OB was identified as one of the six driving factors of total energy-use in buildings in Annex 53 [38], while the completed and ongoing Annex 66 and Annex 79 (respectively) established benchmarks and frameworks within the topic to cover different aspects of OB, from data acquisition strategies, to modelling and integration with building simulation tools [39,40]. The results from the Annex 66 identified challenges in OB to be a result of three distinct features: stochastics, diversity, and complexity. The first one stem from occupants changing their day-to-day behaviors due to several influencing factors [41,42]. The second feature refers to the differences in individuals' actions even when the external stimuli are the same [43]. The third one revolves around the underlying drivers that influence the first two, which can be contextual, environmental, physiological, or social.

The driving factors can be related to the physical environment, such as the building layout or orientation, or physiological, such as those related to an individual's thermal, visual, or comfort levels. In addition, all behaviors can be linked to social behavioral theories, and this dimension is explored in further detail in section 2.4. The aspects of OB can encompass their movement, presence, habits, activities, social influences, clothing, latent emissions and so on. Resulting energy-use from these aspects occur in the form of their interaction with different building systems, such as operation of windows, shading, lighting, thermostat adjustment, appliance use, etc. These aspects are the author's attempt to group most of the observable and unobservable features of an occupant inside a building but are in no way a final summarization on the topic. Since the subject is vast and usually interdisciplinary, studies conducted in the field tend to focus on one or more aspects of the occupant, such as their presence or habits, and their relevance or influence on one particular building system. This resultant energy-use is the main component of interest, for application in building energy design, optimization, simulation, and operation.

2.3 Modeling Occupant Behavior

The goal of creating OB models is to understand how people use a space and how their actions or presence impact the building energy-use. The impact from occupant presence stems from the release of latent and sensible heat, while their impact from the actions centers on the operation of building systems, and appliances. Existing building simulation tools often fail to effectively represent this phenomenon, and the impact of such discrepancies has already been elaborated in 2.1. OB is also one of the main sources of uncertainties in building energy modeling [12]. Due to these reasons, several studies in the recent years have centered on developing OB models [44]. The objective of these studies is either to apply this knowledge in optimizing the building design, or represent the occupant in building performance simulations, or to predict the behavior for the building control systems [45]. Oftentimes the precursor for model development is creation of datasets by monitoring occupants in different settings. This section provides a brief summary regarding the state-of-the-art and challenges in OB monitoring and modeling.

2.3.1 Monitoring OB

Earlier studies involving occupants would rely on manual measurements or self-reported data. Technical advancements in sensing technologies have led to the deployment of these in the built environment for capturing OB. These sensors or sensing modalities have varying range, cost, accuracy, and applications. Their mode of collection can be based on a fixed sampling time (periodic) or be triggered by certain events (event-based), such as door opening, window interaction etc. A systemic review conducted as part of the activities in the Annex 66 classified the different sensing modalities into seven types: image-based, mechanical or threshold, motion-based, radio signal, mixed, environmental, consumption-based, and human-in-the-loop. The selection of the sensing modality is dependent on the research context, such as the type of data to be collected; whether it is to sense the presence of the occupant, or track their movement through a space, or the interaction with windows, shades, lights, or thermostat. All of these technologies come with their advantages and disadvantages, ranging from their cost, complexity, issues with the privacy of the occupant, under or over-reporting of variables, reliability etc. A comprehensive description of each sensing modality and its associated features is given in [46]. The first research question in this thesis centers around this topic of OB monitoring, and the tasks undertaken to study this phenomenon consists of case studies conducted in diverse settings, using multiple sensing modalities. From section 1.4 (List of Publications), it can be seen that Papers 2, 6, and 9 document three case studies.

One of the ways to capture OB data is through Non-Intrusive Load Monitoring (NILM), an alternative to centralized data collection approaches through a Building Management System (BMS) [47]. This concept includes a monitoring system consisting of multiple sensing technologies to capture data about individual appliance consumption profiles. The main feature of this technology is its non-reliance on existing monitoring infrastructure and the little to no intervention from the occupant-side. However, a consistent lack in NILM in building energy context is the granularity and resolution of occupancy information [48–50], which leads to an open question regarding the combination of NILM with high resolution occupancy data. From Paper 2, the first of the case studies in this thesis attempts to address this gap [51].

The advances in sensing technologies, however, has not been implemented in all building typologies where occupants are involved. Most studies centered on OB take place in offices, commercial, residential, or institutional buildings. While that is necessary for assessing the relevance to the energy performance, there exist other spaces that can benefit from this kind of monitoring. The second of the case studies in this thesis (Paper 6) took such a space into consideration, wherein the occupant's influence were monitored in an Operating Room (OR) of a hospital. The activities and movement of an occupant have an influence on the contamination of the room [52], and the studies measuring these influences have relied on manual or restricted methods [53]. This case study was aimed at using depth registration technology to measure the occupant's influence [54].

One aspect of OB and its data acquisition is sensing and recognition of the daily activities of the occupant. This kind of monitoring uses diverse sensing modalities to collect information about and automate the identification of the occupant's actions, often using machine learning or deep learning models. This field of Human Activity Recognition (HAR or AR) has a wide range of applications, from smart homes and intelligent building systems to safety and surveillance, and adaptive control of building systems. The last of the case studies (Paper 8 and 9) was conducted in a smart Living Lab to capture 22 discrete activities of the occupants using sensor fusion methods with the goal of developing Deep Learning (DL) models. In order to ascertain the trends in the field and gauge the rise of trends in using Machine Learning (ML) and DL models, a review was carried out using the PRISMA methodology [55]. An overview of the review methodology and number of studies pertaining to the use of these models for HAR is given in Figure 2-1.

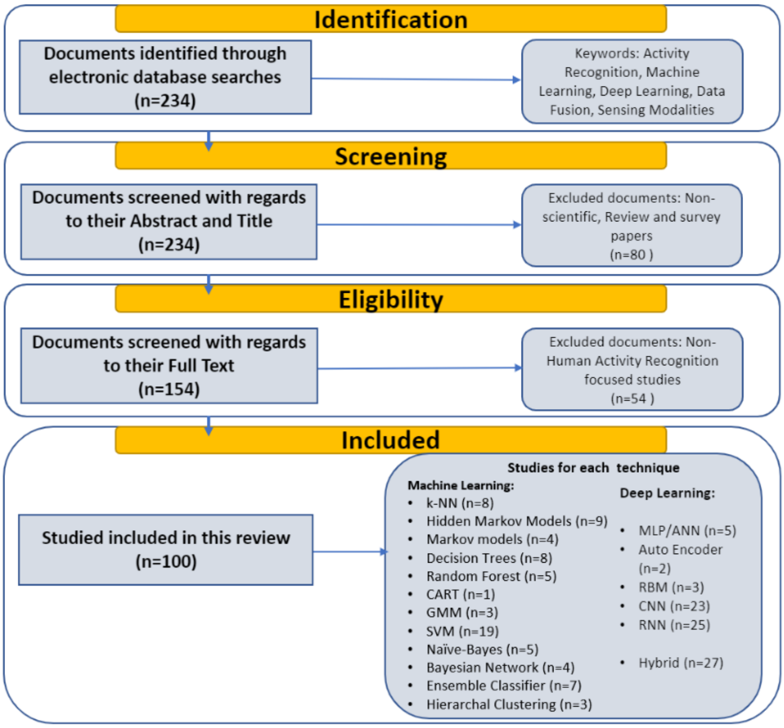

Figure 2-1. From Paper 9: Literature screening procedure using PRISMA framework for the recent work in the field of ML/DL

2.3.2 Current strategies for OB modeling

As described in 2.1, the first few studies involving occupant were centered around light and window operation. These studies made use of probit analysis, based on the correlations observed [22,24,25]. Subsequent simulation or modeling strategies could be broadly classified into two different groups. The first one comprises of models that focus on the systems that the occupant is interacting with, rather than directly with the occupant. These would include linear regressions [56], sub-hourly occupancy-based control models [57] etc. The second group of models deals directly with the occupant and their actions, making use of Agent-Based Models (ABMs) [58], and Markov Chains [59]. A benchmark review on OB modeling by Carlucci et al. had contributions from the author during the course of this thesis work (Additional Publications: Paper 2) [45]. The systematic review selected 278 publications on the subject and grouped the OB models into three different paradigms. The first one was regarding rule-based models which make use of time-dependent user profiles, also known as diversity profiles, such as the ones issued by the American Society of Heating, Refrigeration and Air-conditioning Engineers (ASHRAE) standards [60]. The second paradigm was that of stochastic models, which recognizes that OB is a result of complex relationships and varies over time and context. The third paradigm covers data-driven models, wherein the goal is less to understand OB and more to use ML methods that would replace ‘knowledge-driven’ models that target physical behavior. The work in this thesis concerns the last two paradigms, developing datasets and profiles to enhance knowledge-based occupancy information, and laying the groundwork for and preparing OB models for specific aspects of OB.

The complexity and uncertainty in the field of OB modeling stem from the fact that OB contains various similitudes in the form of presence, movement, activity level, comfort level, social influences etc., and detailed attention has to be given to each of these in order to construct a complete individual profile. Dziedzic et al. proposed a bottom-up approach wherein the collected data from these different fields of simulations could be used to eventually develop a Building Occupant Transient Agent-Based Model (BOT-ABM) [61]. This kind of approach was proposed as a holistic method, where each aspect of OB has its own model, and would be better equipped to represent the diversity in OB (model shown in Figure 2-2).

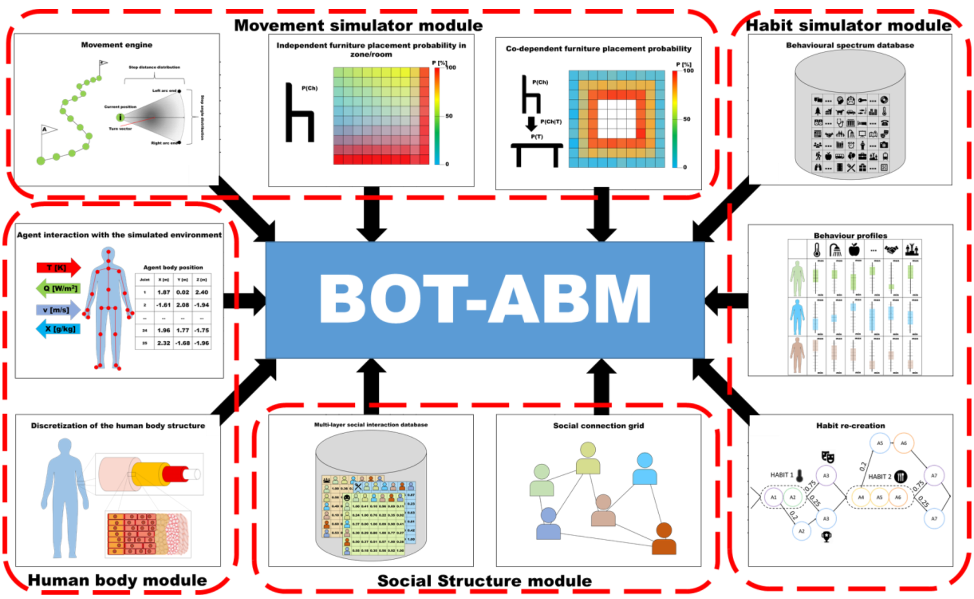

Figure 2-2. Dziedzic (2021): A novel monitoring and modeling technique for energy-related occupant behavior

A large part of simulating OB is modeling the indoor movement and transition of the occupant. Markov chains were used by Wang et al. [62] wherein the movement process was simulated by associating each occupant with a homogeneous Markov matrix. A different form of data collection was used by Martani et al. where the Wi-Fi connections were used as proxy for the occupant sensing [63]. Similar to the occupant monitoring and data collection, this comes with its own set of privacy concerns. To overcome those, another study used a depth registration camera to track and monitor the movement and presence while maintaining a sufficient amount of privacy [64].

A complementary aspect in consideration with the modeling of movement is the simulation of the floor surface layout and the placement of objects/furniture around the occupant. In order to accomplish these simulations, the surface simulator will need access to a database or library of specific information regarding the furniture, as well as the details of the occupant's interaction with it. Current literature does not contain any specifications that can support a surface simulator with that kind of a database. This will have to include the information about the order of importance of the object for the occupant, their access points, area of influence, placement criteria for each, amongst others. The necessity of this information arises from the need to understand the boundaries and potential paths for the occupants' movement, as well as their order of actions with the objects around them. A part of this thesis is centered on the work needed to establish such a library of information for simulation purposes.

2.4 Social aspects of OB

A landmark study on the advances in the field of OB summarized the research gaps and future research directions in the form of ten questions [65]. Two of those questions concern the social aspect of OB, the first one regarding the need of using quantitative methods from social sciences (such as surveys) to understand and provide insight on OB. The point of emphasis was creating appropriate surveys, using suitable social behavioral theories. Despite the limitations of these quantitative methods, surveys are quite useful in assessing latent and unobserved variables that drive behaviors [66,67]. The second of the two questions concern the implication of such an understanding from a social perspective. This understanding, in order to be used for promoting energy-efficient strategies and behaviors, would need interdisciplinary investigations from the physical and social sciences. While there is a marked growth in studying OB in the built environment, technological solutions are often promoted whilst the socio-psychological factors that drive the behavior are ignored[68,69]. Sovacool et al. identified social factors as one of the most under-examined in energy research, especially the factors related to energy-use and decision-making among individuals. It was determined that negligence of these factors during OB or energy modeling would negatively impact the reliability of such models [68].

Quantifying the social drivers of behavior takes place through behavioral theories, and one of the most commonly used ones is the *Theory of Planned Behavior*, that enlists the three components of attitudes, subjective norms, and perceived behavioral control as the main drivers of an individual's behavior. Abrahamse and Steg used the theory of planned behavior to explain subjective norms in residential households, that is, behaviors that would rely on the extent of importance to other members of the household, and the social pressure to carry out the actions [70]. Chen and Knight used the same theory to examine the influence of colleagues among more than 500 employees in electric power companies in China, and the results indicated that the approval or disapproval from the colleagues was the major factor influencing the individual's energy saving actions [71]. Chen et al. also made use of agent-based modeling to build a network level computational model that simulates the decision-making process of individuals under different type of network configurations [72].

A large part of this thesis is dedicating to using this theory to make a framework for assessing and analyzing social aspects of OB, investigating influence factors in that aspect, and use the datasets and results generated for a proof-of-concept model. This work also feeds into the proposed BOT-ABM, with the contribution aimed at a social structure module. The modeling technique considered is described in the next section.

2.5 Agent-based modeling

Agent-based modeling or Individual Based Modeling is a computational modeling technique which rose to prominence during the 2000s in several fields. The core principle of this model is that it consists of autonomous agents, which can be individuals or belong to a collective of individuals. These agents interact with other agents and with the environment, with the objective of measuring their effect on the system as a whole. It is a microscale model, where the high-level system changes come into effect as a result of interactions between the low-level subsystems[73].

When it comes to building performance simulations, ABMs have been used to assess occupant behavior. Erickson et. al used this process to address issues related to oversizing of HVAC equipment, by introducing ABMs to model room occupancy. Their results showed that HVAC consumption could be reduced by 14% by using occupancy simulations[74]. In another study targeted at zones in healthcare facilities, occupancy modeled using ABMs showed that using accurate simulations would lead to changes in the design loads and sizing of HVAC equipment that could produce up to 43% reductions in the energy consumptions [75]. Furthermore, this approach has been used to incorporate different types of occupant behaviors as well, with Lee et. al. testing an agent's consideration of five different factors (ranging from clothing levels, activity level, window use, blind use, and personal heater) to restore their comfort levels [58]. Azar et. al. used a similar approach for incorporating three different kinds of occupant behaviors and their changing energy-use behaviors. The changes were triggered when the agents with less energy-conscious behaviors interacted with agents with more energy-conscious ones, thereby learning by simulated social influences between occupants in a commercial building [76].

The last part of this thesis is based on the last two cited examples, to use ABM as a technique to model occupant behavior and incorporate social aspects in OB models.

3 METHODOLOGY

This chapter describes the methodology of this thesis. Section 3.1 centers on the monitoring of occupant behavior, introducing the different data acquisition strategies and experimental setups for the case studies. These studies were conducted to provide answers to the first of the research questions (Q1) outlined in the beginning of the thesis. In addition, it provides an overview of the data processing undertaken during the process. Section 3.2 explains the groundwork for the development of OB models, describing the preparatory work done in establishing a database for some aspects of OB. Section 3.3 includes a proposal for a framework for assessing social aspects in OB, targeting Q2 and Q4 of the research questions. Section 3.4 aims to answer Q3 and details the methods used for developing the models from the datasets collected in the first two subsections.

3.1 OB Monitoring

One of the main tasks of the studies conducted in this thesis was to contribute to the understanding of OB. This was achieved by collecting data regarding occupant presence, movement, habits, interactions or influences on their surroundings, and energy-use. Building occupants were monitored in diverse settings and spaces, using a variety of sensors and data acquisition strategies. Each setting came with its own set of requirements, challenges, etc. pertaining to the spatial structure, privacy demands, need for accuracy etc. The following subsections delve into each of the individual studies, explaining the selection of the data acquisition and data processing strategies according to the needs of the experimental setup.

3.1.1 Occupant energy-use behaviors in shared office spaces

In order to control energy-use efficiency, costs, and meet the requirements for occupant comfort, building facility managers make use of a system of sensors, networks and controls. Such a system is termed as a Building Management System (BMS), which falls under a centralized process of collecting data about occupants or building systems in (mostly) commercial buildings. While a good number of studies typically gather data through such centralized systems, it is important to note that the applicability of these is limited to buildings with similar monitoring capabilities. The high cost and incompatibility of such monitoring systems with building that are either too old or already have existing BMSs, gives space for other monitoring/data collection alternatives to emerge, and Non-Intrusive Load Monitoring (NILM) is one such methodology [77,78]. This alternative to a centralized BMS is based on a collection of smart sensors or energy meters that separate the aggregated electrical load to individual appliance consumption profiles. The uniqueness of NILM stems from the minimal to no intervention required from the end-user (occupants) and its ability to operate without the need to connect to any existing infrastructure, thereby making it ideal in cases where the application of BMS is either restricted or too expensive.

However, in both these procedures, studies often lack granular information about occupancy, either due to absence of measured occupancy data [79–82] or binary occupancy states about zones [83,84]. Studies that do take into account individual occupancy of a space analyze it with regards to the energy-use of the entire space or as a sum of all used devices. This leaves an open question regarding the potential of identifying energy-saving opportunities by applying NILM methods to end-uses such as plug loads or lighting loads, in combination with granular occupancy information.

The goal of this particular study was then to monitor occupants and their energy-use at an individualistic level, both to investigate energy-saving opportunities and generate datasets for building OB models. The occupants were monitored at the desk level, capturing their presence, as well as the electric loads for specific devices. In addition to quantifying the relationship between individual presence and energy-use and two end-uses (plug-loads and lighting), the data was also used to compare the measured and standard profiles (from ASHRAE) that are generally used during building design and simulation [60]. Additionally, surveys were used to capture the social aspect of the energy-use as well (Subsection 3.3.6).

Figure 3-1. From Paper 2: Area of study and sensor placement

The experiment was conducted in an academic building in Abu Dhabi, UAE, in a shared office space. The area of study consisted of 6 primary workstations, 2 shared workstations, and 1 common table. The layout can be seen in Figure 3-1 (Schematic view and View A). The space was used primarily by students and was subject to change in occupancy every semester. The study was conducted for a period of 8 months, and 8 different occupants were monitored during this time. At a time thought, no more than 6 primary users were present at the office, occupying the workstations. While the operational hours of the academic institution were from 8am-5pm on weekdays, the area was open to access at all times. In addition, the area was not restricted for visitors. Both of these factors were in-lieu with the non-intrusive nature of the proposed study. Access to daylight was limited, which meant that the occupants mainly relied on the artificial lighting system.

Three different types of sensors were employed in this study. The first type was Passive Infra-red based (PIR sensors), to detect the occupancy at each of the workstations, as well as the common table. A total of 9 such sensors were deployed, wherein sensor had a unique ID, and a range of up to 80 meters. These sensors only communicated with the server during a change in occupancy status, reporting a binary measurement (occupied/unoccupied). The position of an individual sensor, as seen in Figure 3-1 (View A), was calibrated based on three factors: detection of actual occupancy over the entire range of the workstation, avoidance of false triggers from passersby, and zero occlusions.

Each workstation had 6 power outlets, and the plug load sensors were installed in each of them, contributing to a total of 48 such sensors. They were installed and configured to each specific device (laptops, monitors, docking stations, desk lamp, misc.). The only intervention from the occupants that occurred during this study was the agreement to use each outlet for a specific device. However, this intervention was not strictly necessary since the reported energy-use could be compared to the power specification of the device that was being monitored. The last set of sensors were for recording the illuminance of the area. The office space was illuminated by 6 fixtures and one sensor was attached to each of them as shown (Figure 3-1: Schematic View). The lighting was motion-controlled, without any manual control options. Based on the monitored level of illuminance, the status of the lighting was inferred. Figure 3-1 shows the placement of all the sensors, occupancy (O), plug-load (P), and lighting (L).

The data collection process was continuous, with each plug-load and lighting sensor reporting measurements at 15-minute intervals. The data from (passive infrared) PIR (occupancy) sensors, however, was event-based, recording 'blank' triggers for an unchanged event. These blank triggers had to be first replicated with the last recorded trigger. After this up-sampling, the data was in accordance with the 15 minute-interval of the other sensors.

The data analysis was conducted in three parts, the first of which was to determine the relationship between the occupancy and energy consumption in the shared office. The consumption included distinct figures for lighting and plug-loads. Box-and-whisker plots were used to visualize the power levels for varying occupancy levels. The second part was to measure the amount of energy consumed when an occupant was not present at their workstation. Two analyses were drawn for this part, the first one targeting office-level, which compared the total office level consumption when occupancy was at zero (vacant), and when at least one occupant was present (occupied). This analysis is the one commonly used in the literature. The second analysis measured the consumption at a device-level, while also distinguishing between each device's consumption when the occupant is present/absent from their workstation. The last part of the data analysis was the development of diversity profiles for occupancy, plug-loads, and lighting patterns in the area of study. These profiles show the intensity of a variable (such as occupancy) throughout the day (marked hourly). The intensities are represented numerically from 0 to 1, 0 being the least possible value and 1 being the highest. These profiles are generally used by energy models, building facility managers etc. while predicting the performance of a building. These were then compared to standard ASHRAE profiles that are used for common building types. More information about the calculation is presented in 3.3.1.

3.1.2 Occupant activity influences in operating rooms

While the perspective of understanding OB relies mostly on its relationship to the building's energy performance, building occupants exert their influence in numerous other ways. Newer and more advanced monitoring systems provide a method to quantify and document these influences. Specialized settings and spaces offer interdisciplinary opportunities of testing such methodologies in conjunction with other fields. One such setting was used in this study, which centered on monitoring occupant influences in an Operating Room (OR). The way occupants' presence and activity play a role in such a specialized space is by being a source of contamination. In fact, in modern hospitals, the most significant source of airborne contamination is from the dispersal of particles from people in the OR [85,86]. This can be caused due to the staff movement that can increase Colony Forming Units (CFUs) by shedding [87], re-suspension of settled particles on the floor or other surfaces into the air, or pumping effects inside clothes [88].

While a number of studies have explored this influence and documented the correlation between occupant activity and bacterial distribution in the OR, the experimental methods used to monitor occupants show potential for improvement. In some cases, the studies would rely on manual observations and door openings [53], while in other cases the air quality would be used to infer the occupancy count [89]. The setting for this study is quite unconventional considering that occupant behavior studies tend to take place in commercial or residential buildings. However, the lack of individual and dynamic monitoring of occupant activity in ORs, combined with developments of sophisticated data acquisition strategies in OB studies, served as an opportunity to employ OB monitoring in such a setting.

This study was conducted in a cardiopulmonary OR located in St. Olav's Hospital in Trondheim, Norway. The OR was used to conduct three mock surgeries, during which the occupants playing the role of the surgical staff would perform a controlled series of actions. The movement and activity of the staff was then monitored using depth registration cameras. The selection of depth registration was based on its capabilities of capturing dynamic geographic information of the occupants. The principle behind such a device is similar to that of a projector, wherein the device emits an incident infrared beam that gets reflected back to the device. An additional sensor measures the time for the reflection of this incident beam, thereby inferring the distance of the objects in its field of view.

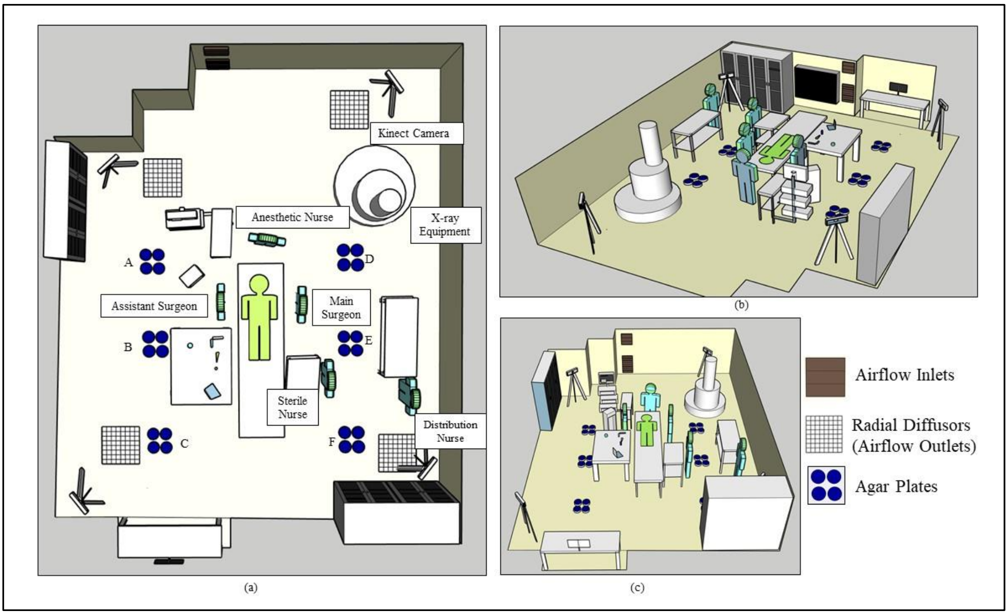

Figure 3-2. From Paper 6: a) Layout of the mock surgery experimental setup (Top View) (b) & (c) Bird's angle view of different perspectives of the setup

The experimental setup is shown in Figure 3-2, wherein 6 participants represented 1 patient and 5 staff members (main surgeon, assistant surgeon, sterile nurse, distribution nurse, anesthetic nurse). In order to assess the influence of the occupant activities on the airborne bacterial contamination, a total of 24 passive agar plates (in groups of 4) were used in 6 different locations around the operating zone. The mock surgeries consisted of 4 different phases, during which the participants performed a designated set of actions for each phase. During the start of these phases, a set of agar plates would be opened to record the contaminations corresponding to the start of the phase till the end of the mock surgery. Here, 'group' represents the 4 agar plates present at each of the locations (A-F) and 'set' is the specific plate from each group that would be opened during the start of a phase. Figure 3-3 provides a visualization of the procedure, depicting the sampling windows.

The occupants were monitored using 4 Kinect devices (cameras) present in each corner of the room. The devices pick up any human body present in their field of view and process it into a skeleton model consisting of 25 points in a 3-dimensional matrix. All of them were capable of registering 6 people simultaneously and the use of 4 of these negated any possibilities of misregistering any activity. Since each device operated on its local coordinate system, the data from the 4 devices had to be adjusted by selecting one of them as a reference point. The details regarding the activities conducted in the mock surgery, along with the reasoning behind them are further elaborated in [54]

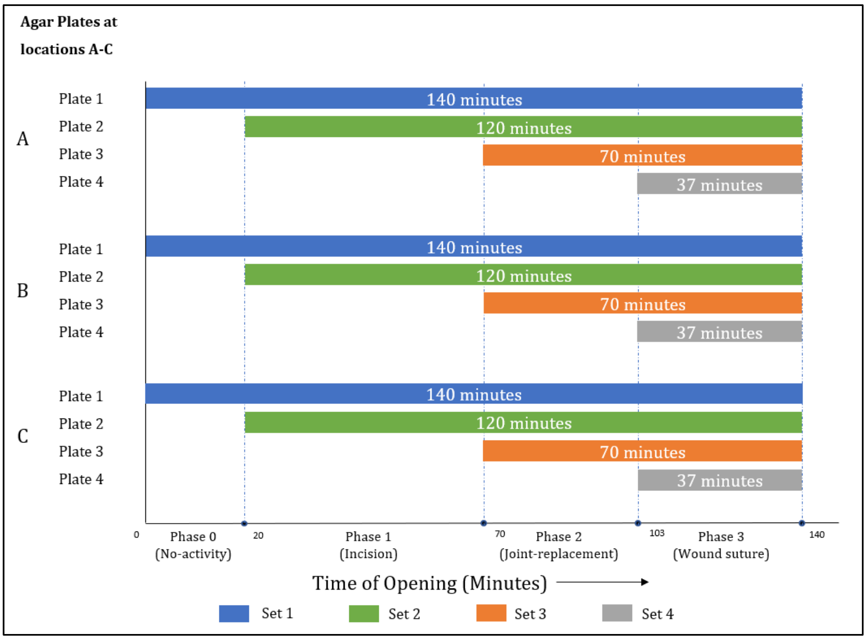

Figure 3-3. From Paper 6: Sampling windows of the agar plates

3.1.3 Occupant activities in residential spaces

Another dimension of monitoring OB that has garnered much attention is the necessity of studying occupant movement and activities to develop Activity Recognition (AR) systems. The applications of AR are extensive, from establishing smart homes to improve the end-user efficiency of the residence [90,91], to healthcare, fitness, safety monitoring and elderly assistance [90–94]. The principle behind AR is to infer an action or activity an occupant is performing based on discrete consecutive actions or movements they take. For example, cooking can be considered an activity that comprises of several atomic actions, such as turning the stove on, taking utensils from a cabinet, pour water from the sink, etc. This way, an activity gets described as an ordered sequence of events.

The advancements in sensing modalities combined with the exponential growth of Artificial Intelligence and Deep Learning (DL) paradigms has transformed the smart environment and made AR a pivotal process to facilitate the ambient intelligence of such environments. Establishments of an AR system, however, is dependent on capturing a significantly large amount of data about each discrete action undertaken by the occupant. This data must then be assigned accurate labels and be granular enough to avoid mislabeling. In addition, real-world data in this context can be generated from multiple sensors. Some of the challenges associated with AR have to do with how similarly the atomic or discrete actions are performed within a time interval, and how quickly occupants move between different activities. Data collection regarding these activities has to be precise enough to address these issues. Depth registration cameras provide a better alternative to traditional RGB (red, green and blue) cameras in this regard, considering their capabilities for background segmentation and object detection. It also offers better possibilities for preserving the privacy of the occupants.

Figure 3-4. From Paper 9: Case scenario and experimental design setups in the smart living lab

Similar to the previous case study (Section 3.1.2), this study utilized Microsoft Kinect cameras in a smart living lab to capture occupant activities typically performed in a residential space. These activities are mundane everyday tasks, collectively termed as Activities of Daily Living (ADL). The living lab is built to resemble a modern home, complete with a bathroom and kitchen. Three different experimental setups were utilized to perform the study. As shown in Figure 3-4, these setups were centered on kitchen (Setup 1), bedroom (Setup 2) with three participants performing activities typically associated with that space. Three depth cameras were used in each setting to avoid occlusion and enhance the accuracy of data acquisition. A total of 14 continuous activities were performed, ranging from ‘cooking a meal’, ‘dishwashing’ in the kitchen, to ‘sleeping’, ‘reading’ in the bedroom. While the participants were given a set of tasks to conduct in a specified order, they performed it with their own interpretation, in order to have subject variability. Each activity was performed ten times, contributing to thirty recordings for each activity from the three participants. These activities were punctuated with a calibration pose to provide segmentation between the recordings, to make it feasible to identify individual activity for training a DL model.

3.2 Developing libraries and databases for OB models

Since the establishment of a holistic OB modeling system such as a BOT-ABM is vast in its undertaking, the studies conducted for this thesis centered on laying the groundwork for modeling social aspects and aid in the surface layout simulator module. This section discusses the preparatory works done in this regard. For the layout simulator, the task consisted of establishing a library of knowledge about indoor office layouts, along with the development of an application for the related data collection. The process for the social aspect was more comprehensive and is presented in section 3.3. It consisted of identifying key concepts regarding OB and social influences and creating a framework which enabled an extensive survey for data collection. The data was used to create path diagrams and regression models, and eventually develop OB profiles.

3.2.1 Library for a floor surface layout simulator

Occupant movement and indoor transitions constitute an important part of OB modelling. Dziedzic et al. developed a movement simulator for the proposed BOT-ABM system [95] which simulated each step of the occupant’s movement based on pre-defined metrics. A complementary aspect for the movement simulator is the simulation of floor surface layouts and the placement of objects in indoor areas, since occupants have to navigate their movement based on these factors. These layouts need not reflect a ‘standard’ room setting, but the lack of information about typical layouts and object placement can be an obstacle for developing future models. Generating such layouts can also be useful for deciding the optimal placement of HVAC installations, by giving the building designers the ability to compare different layouts. More detailed description of Dziedzic’s layout simulator is present in the additional publications. The upscaling of this simulator needed an access to a database or library containing information about object placement, along with data about occupants' interaction with it. To achieve this, the work outlined in Paper 5 was focused on identifying the variables to be incorporated in such a library.

The seven variables identified are shown in Table 3-1. The 'object class' is to denote the type of object/furniture, that would be acquired from a compiled list, whereas the 'room category' would represent the room type. The 'order of importance' would be used to prioritize the placement of objects during the simulation process, while 'placement criteria' would be based on referential distance from doors, openings, edges etc. The 'area of influence', 'access points', 'hinge points' would be used to establish the constraints in modeling the path of an occupant. In order to collect data that would fulfill the requirements of the study, an application was developed using Matlab App Designer toolbox. This application was designed to provide users with an interface consisting of a grid based on the room size, and a geometric layout where they could proceed to select different objects and place them around the two-dimensional grid. Figure 3-5 depicts a sample of the interface. The application also registered the order of importance for each object that the user had to denote before placing the selection. This information would be then processed into a database to be used by the simulator. Further description of the database design is detailed in Section 3.2.2.

Furniture Class	Room Category	Order Of Importance	Placement Criteria	Area Of Influence	Access Points	Hinge Points
	1		Criteria1			
Class1	2		Criteria2			
	3		Criteria3			
	1		Criteria1			
Class2	2		Criteria2			
	3		Criteria3			

Table 3-1. Identified factors for the database

The objectives of this study were threefold; to identify the type of information needed for this library, develop an application for data collection, and collect a sample dataset to create the architecture for a database. The sample dataset was collected at the Norwegian University of Science and Technology, consisting of spatial information of 80 individual offices (private and shared). This dataset contained details about the placement of objects in offices, location of HVAC installations, type of occupancy and number of occupants.

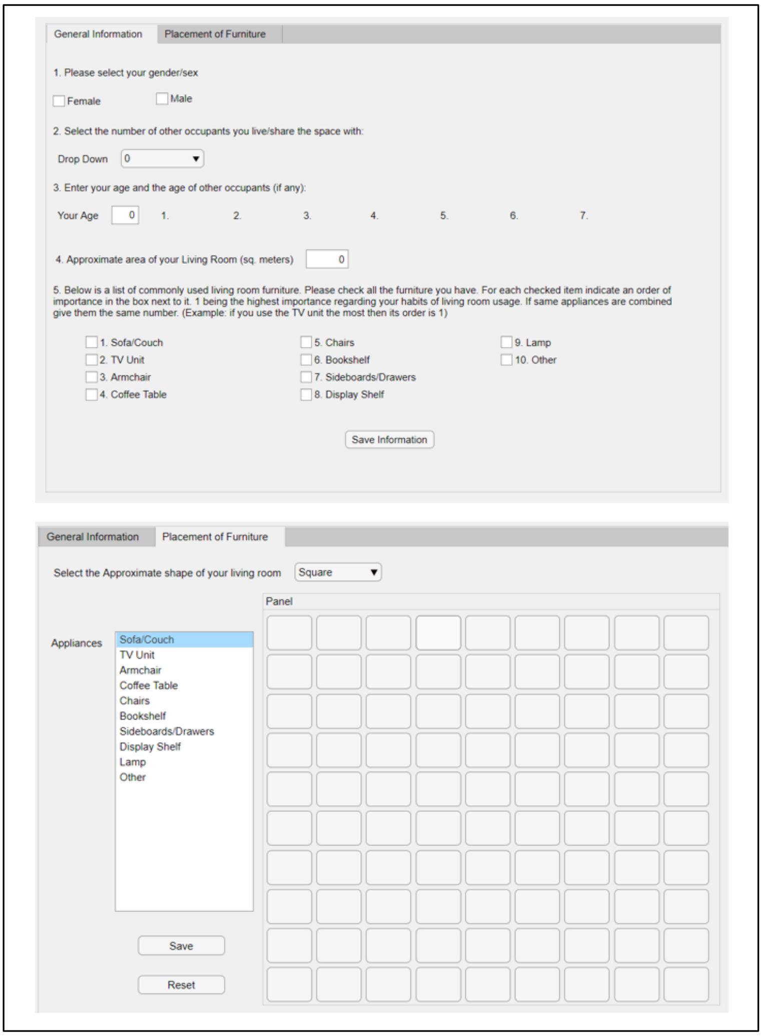

Figure 3-5. Interface of the MATLAB Application

3.2.2 Database design

The studies from this thesis generated multiple datasets, and profiles that were developed to be used in eventual OB models. In addition, the aim to contribute to a system such as the proposed BOT-ABM created a need for having structured databases to be used by simulators. A database can be defined as an organized collection of data that can be accessed and interacted with. This is done through a Database management System (DBMS) by using appropriate queries. A majority oof DBMS use SQL (Structured Query language) to communicate with a database, such as to insert, update, or retrieve data. The DBMS used in this thesis was MySQL, which is a Relational Database Management System (RDBMS). It provides a user interface to perform tasks on the data and organizes it into different tables that are linked together.

The idea behind the creation of these databases is to have an efficient data organizational system that provides a platform for simulators connected on a shared server. This process involves normalizing the data in databases collected during the thesis. Data normalization is done to reduce data redundancy and provide tables of data that only contain relevant information. The benefits of data normalization include reduced storage space, limit anomalies regarding insertion and deletion of new data, and improve query performance. Another reason for establishing these databases through MySQL is to have data that is related to each other. Tables are related to each other through primary and foreign keys, which take the form of one-to-one, one-to-many, or many-to-many relationships. For one-to-one relationships, a single record of data in one table is associated to a single record in a different table, whereas one-to-many denotes that single record is associated with multiple records in another table. The resulting database from this section is described in Results and Discussions (4.2).

3.3 Social aspects of OB

As elaborated in Section 2.4, social aspects tend to be largely neglected and underexamined in energy research, especially concerning the integration of this aspect in modelling and simulation. This provided the motivation for developing a framework to quantify social aspects in OB, with the aim of collecting data based on the framework and generating profiles of occupants to be used in models. Subsections 3.3.1 – 3.3.5 provide the evolution of the framework, data collection, data processing, analysis, and the resulting OB profiles from the study. Subsection 3.3.6 describes an addendum to the case study presented in 3.1.1, where surveys were used along with all the sensing modalities to explore the social aspect.

3.3.1 Addendum to the case study in shared office spaces

In order to explore social aspects of OB, the case study presented in 3.1.1 included a survey that the occupants had to fill in during the course of the experiment. This was done to map an occupant's perception of their own occupancy and energy-consumption patterns, and draw a comparison to standard and observed values. This was only an exploratory study, to gain an insight into the relationships between perceived and actual occupancy and energy-use.

The surveys attempted to create a schedule similar to the ones used in ASHRAE. For this process, the occupants had to denote their occupancy for each hour, ranging from 'always occupied' to 'never occupied'. This was repeated for device-level consumption as well, with options ranging from 'always switched off' to 'always switched on'. The profiles resulting from these surveys were then compared to the data from sensors and the standard ASHRAE profiles.

The diversity profile values were calculated using the equations (1) and (2), which takes the ratio of sensors that indicate 'occupied' status over the total number of occupancy sensors in the study. This was averaged over the period of study, from the first day 'd' to day 'N'.

$$Occupancy\ Diversity = \frac{\sum_{d=1} N \frac{\#Sensors\ Occupied}{\#Sensors\ Total}}{N} \qquad (1)$$

$$Energy\ Diversity = \frac{\sum_{d=1} N \frac{Energy\ Consumed}{Max\ Energy\ Observedl}}{N} \qquad (2)$$

3.3.2 Extended TPB model

The quantification of social aspects occurs through the use of social behavioral theories like the *Social Cognitive Theory* [96], the DNAS framework[97], or the *Theory of Planned Behavior (TPB)* [98]. The last of these, developed in 1991 by Azjen has been used widely in conjunction with social sciences and energy research. The implementation of this theory has been to analyze energy-related behaviors and target factors that are responsible for more environmentally friendly or energy conscious behaviors. According to TPB, the major predictor of human behavior is the intention to perform that behavior, which is in turn influenced by three key components: (i) Attitudes: that describe the beliefs and motivations a person has towards a behavior or action, (ii) Subjective Norms: beliefs about others will respond to their behaviors and, (iii) Perceived Behavioral Control: ideas about the ease of performing an action. Subjective norms are further split into Descriptive Norms and Injunctive Norms, where the former measures perceptions about how others behave, and the latter is about how other might respond to your behavior. Within the field of OB in buildings, this theory was used to confirm that the components of TPB were able to predict up to 81% of the occupant's intention to act in an energy-conscious manner [99]. Greaves et al. also demonstrated that TPB could explain 61- 46% of the variance in occupant energy-related behavior at the workplace [100].

With the rise in the use of TPB in energy-related behavior, researchers have also sought to construct extended TPB models by integrating it with their behavioral theories, adding selected components to it. D'Oca et al. developed an interdisciplinary framework based on TPB, DNAS framework and Social Cognitive Theory for OB in buildings [101]. Chen demonstrated that an extended TPB model can also be more effective in explaining behaviors [102]. An example of extended TPB model can be seen in [103], where Xu et al. added two components to the theory to analyze energy-saving at work.

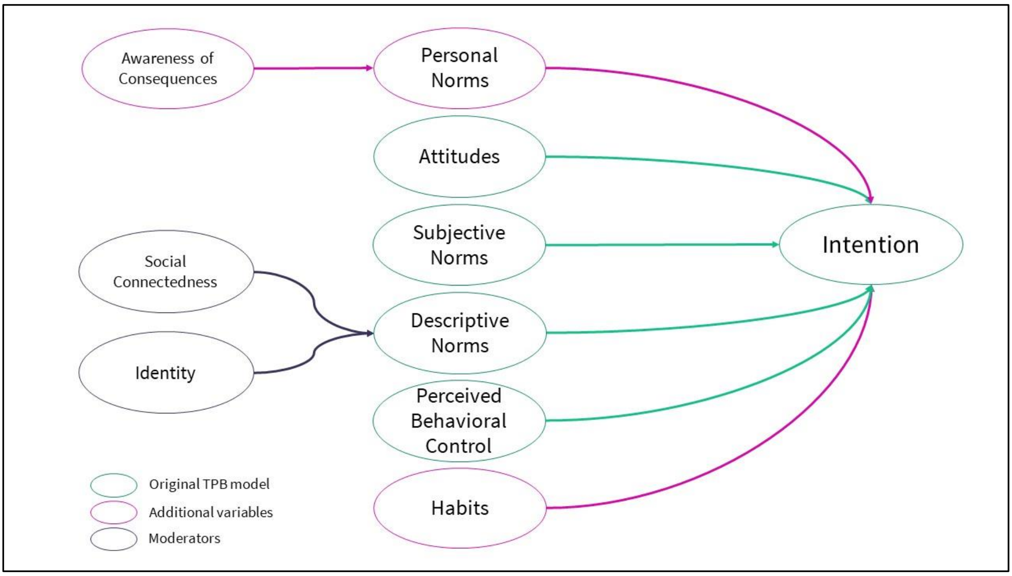

Figure 3-6. From Paper 4: Theoretical extended-TPB model

The study in this thesis utilized an extended TPB model as well. There were two reasons behind this selection. Firstly, creating models with higher number of variables in the hypothesis is helpful because some of the variables might get discarded during the data analysis stage, due to lower reliability. Secondly, components such as habits (past behaviors) or identity have been shown to be strong predictors when added to TPB [104,105]. The addition of habits as a component in an extended TPB model has been applied with successful results in consumer behavior, transportation, etc.[106,107] but the application in energy-related OB is still lacking. In addition to habits, the factors of social connectedness and identity were added, to moderate the effect of Descriptive Norms. Furthermore, to test it out, another component, that of Self-Construals was added, which was identified as a strong potential moderator of pro-environmental or sustainable behavior [108]. This component measures whether the behavior is conducted as an independent or collective decision. The Extended-TPB model is shown in Figure 3-6.

3.3.3 Data collection

This framework was used to create a questionnaire that presented respondents with 40 questions, wherein each variable from the model was measured with two or more questions (items). More items are included to reduce the measurement error or increase the reliability of the data. The questionnaire used a 5-point Likert Scale, a commonly employed instrument for measuring people's attitudes towards behaviors. It records a respondent's reaction to a series of statement on a scale from 'Strongly Agree' to 'Strongly Disagree'. Each response has a corresponding numerical value, the sum of which contributes to an individual's score for that variable. The questions targeted people's attitudes and behaviors at the workplace, and their intentions regarding energy-related behavior, which was in the form of interactions with the building systems. These systems consisted of lighting, thermostat, windows, and appliances. Before conducting the data collection, a pilot study was carried out with a focus group of 10 people to gauge the suitability of the questionnaire.

The feedback from the pilot study was used to improve the questionnaire by adding statements for clarification, adjusting the length etc. The order of questions was also edited to avoid bias in the responses. Since it can be difficult to validate models that are based on self-reported surveys, a section was devoted to asking specific scenario-based questions, presenting the respondents with a set of circumstances in the workplace and asking how likely they would be to perform an action. The idea behind this was to create training and testing data. The occupant profiles would be based on their scores for each of the variables (training) and validate the models with the scenario-based questions. It would involve constructing the circumstances in the model and check whether an occupant in the model would behave the same way as their corresponding responses in the scenarios.

The data collection was carried out using Qualtrics, an online survey platform. Prior to the distribution of the questionnaire, a privacy assessment was conducted by the Norwegian Centre for Research Data, to ensure all compliance with data protection legislation. It was distributed among the employees at the Norwegian University of Science and Technology, collecting 110 responses. Some of these were discarded since they were incomplete, resulting in a final tally of 101 samples.

3.3.4 Data analysis overview

The collected data was processed and analyzed with the statistical software SPSS [109]. The data screening was performed to check for outliers and missing data. A univariate analysis was used for outliers, and both the outliers and missing data were under the ordained limit of 5%. The missing data was handled using replacement methods, based on the average of the nearest corresponding values. An overview of the data analysis is shown Figure 3-7, that consisted of three steps. The first one was measurement modelling, which was concerned with checking the reliability and internal consistency of the data.

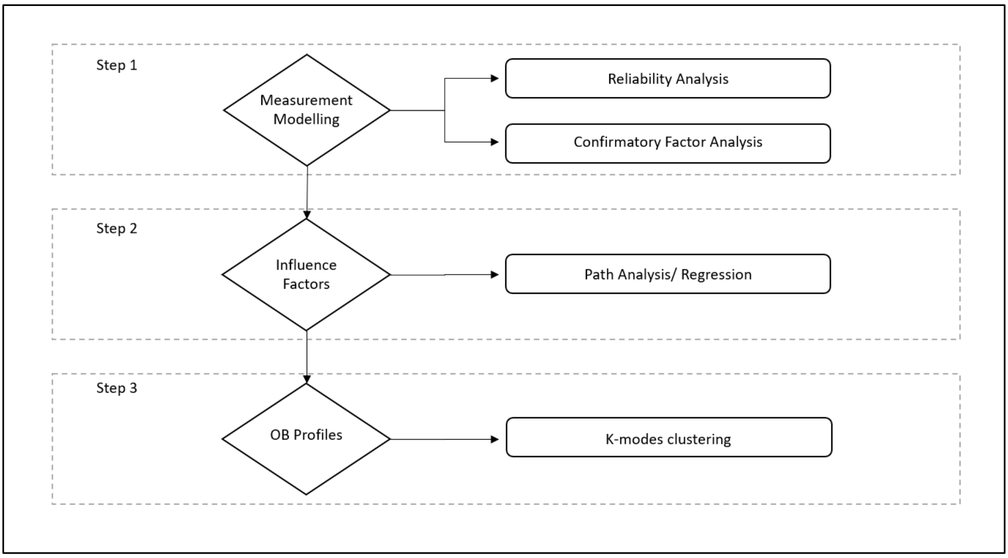

Figure 3-7. From Paper 4: Overview of Data Analysis

The reliability analysis was done on each variable in the questionnaire. Here, variable denotes the factors identified in the extended TPB model, such as ‘Attitudes’ or ‘Perceived Behavioral Control’. Each variable was recorded based on two or more questions. For example, to record people’s attitudes towards energy-saving at work, they would be asked to express how much they agreed or disagreed with the following two statements: ‘I feel like saving energy at work is beneficial’, and ‘I feel like saving energy at work is good’. The answers to both of them should be correlated enough to ensure reliability. Similarly, all the answers that are used to assess a variable should be strongly correlated to classify the data as reliable and consistent.

For variables that had two questions (or items) the parameter used was Pearson’s correlation coefficient, while for variables with more than two items, Cronbach’s alpha was used. For some variables however, this approach was not suitable. Past behavior, or habits, contained four items, each dealing with one building system. These items need not be correlated strongly because an occupant might have different habits for different systems (for example, lighting and windows), that are influenced by the type of installations in the building. In addition, the ease of access for both systems can be different. For these variables, a confirmatory factor analysis was performed. Factor Analysis is a method for modeling observed variables to identify unobserved ‘factors’. Factor Loadings are generated, that show the correlation between the original variables and the factors. The analysis is tested using two parameters, the p-value for significance, and the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO value).

3.3.5 Path analysis, SEM, and regression models

Path analysis, Structural Equation Models (SEM), and regression are all statistical methods to study the relationship between multiple variables. Regression models are suitable to analyze the strength and direction of relationships between one dependent variable and one or more independent variables. SEM, on the other hand, is a broader term that refers to a family of models that are more complex and allow researchers to examine the relationships between multiple variables simultaneously. In SEM, researchers can examine the relationships between multiple observed and unobserved (latent) variables. Latent variables are constructs that are not directly observable, but are inferred from observable measures, such as survey responses or test scores. These can be the underlying factors that influence people's attitudes, beliefs etc. A theoretical model is first hypothesized, that includes both observed and latent variables. This model is then tested using the data to determine if the relationships hypothesized are valid. An important difference between SEM and path analysis is the inclusion of error terms. Error terms represent the amount of unexplained variation in a variable and are included in the model to account for measurement error, unobserved variables, and other sources of variance. In SEM, error terms are often included for each observed variable and for each latent variable. This allows for the estimation of the amount of error associated with each variable, and to assess the overall fit of the model to the data.

In addition, SEM allows for testing moderation effect. Moderation is when the strength or direction of the relationship between an independent variable and a dependent variable depends on the level of a third variable (moderator). It is examined by including interaction terms between the independent variable and the moderator variable in the model.

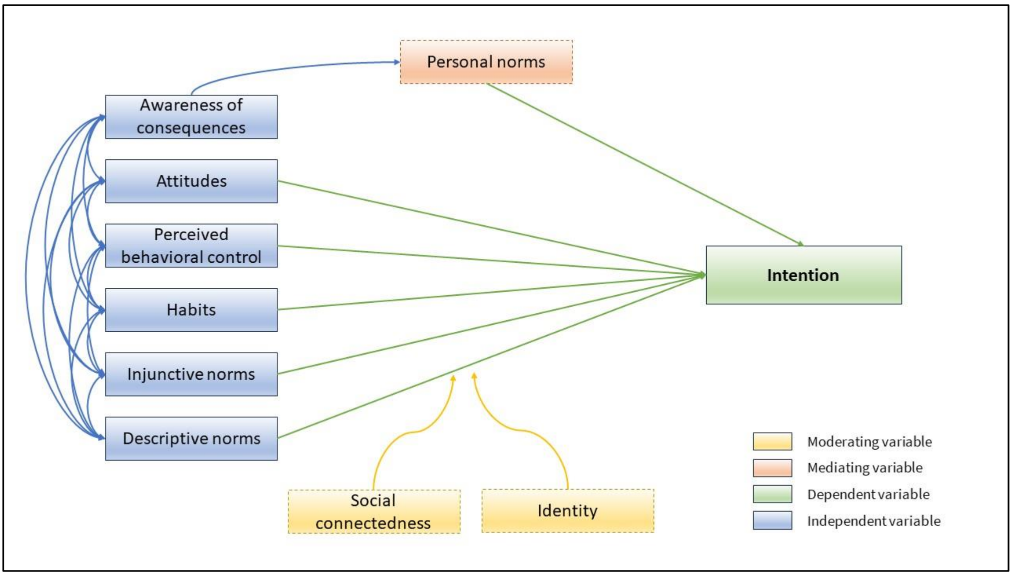

Figure 3-8. From Paper 4: Path analysis model

The data collected from the survey had variables as shown in Figure 3-6, with intention being the dependent variable. While most of the other variables were independent, there were some that were both dependent and independent. For example, 'awareness of consequences' is a predictor for 'personal norms', which is in turn a predictor for intention. Here 'personal norms' becomes a dependent variable for 'awareness of consequences' while being an independent variable for intention. Such a variable is also termed as a mediating variable. Furthermore, variables like 'social connectedness' and 'identity' are theorized to moderate the effect of 'descriptive norms'. To address these complexities and challenges, SEM and path analysis were both included to investigate the relationships between these variables. The models were developed in SPSS Amos [110], a tool for statistical modeling, that works in conjunction with SPSS.

Figure 3-8 depicts one of the path analysis models used. The independent variables were determined from the results of the reliability analyses described in the previous section. After determining that the variables were reliable and consistent based on the survey data, the scores for the questions were averaged to obtain a single score for each independent variable. These were used to draw covariance between each of them. Error terms were only calculated for the dependent variables, of which there were two (personal norms and intentions; error terms e1 and e2, respectively).

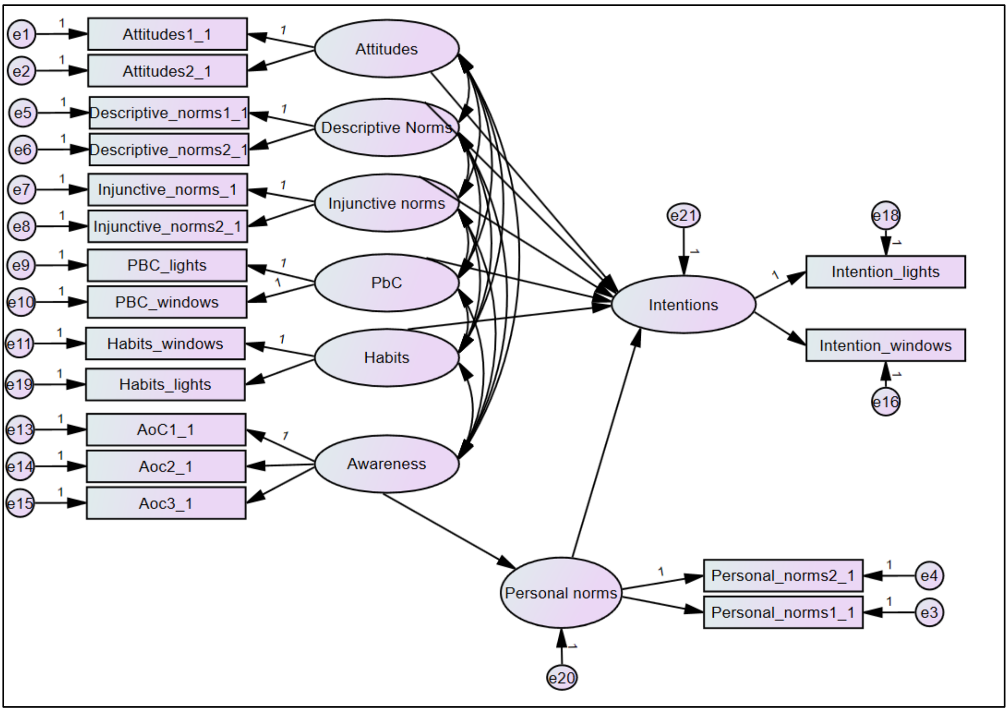

Figure 3-9. Structural Equation Model

The SEMs were more complex, as depicted in Figure 3-9. This model was drawn without taking the average of the scores, but rather taking the raw input from the survey. For example, to determine Attitudes, 2 questions were used in the survey. Here, the scores for each of them were used as observed variable, and following the theory of SEM, a latent variable was drawn for them. These are the inferred, or unobserved variables. Error terms in an SEM are generated for each of the observed variables and dependent variables.

Individual models were developed for each of the building systems, since occupants can have completely different interactions depending on the building system. As depicted in Figure 3-10, 4 regression models were made, one for each building system (windows, lighting, appliances, computers). Similarly, there were 4 SEMs, and 4 path analysis models. It should be noted that thermostat is not available as an individual category, but instead, merged with windows. This was done since the questions in the survey regarding window opening behavior were linked to people's use of thermostat, such as their willingness to regulate the indoor temperature using windows or thermostats. After this step, several model variations were applied, along with the baseline TPB model (which includes Attitudes, Perceived

Behavioral Control, Descriptive Norms, Injunctive norms and Behavioral Intention). These variations took the form of including one or both moderator variables, the additional variables such as Habits and Personal Norms (PN) or those that were not directly predicting the dependent variables (Awareness of Consequences). The fit of the models is assessed using several fit indices, such as the chi-square test, root mean square error of approximation (RMSEA), goodness of fit index (GFI) and comparative fit index (CFI).

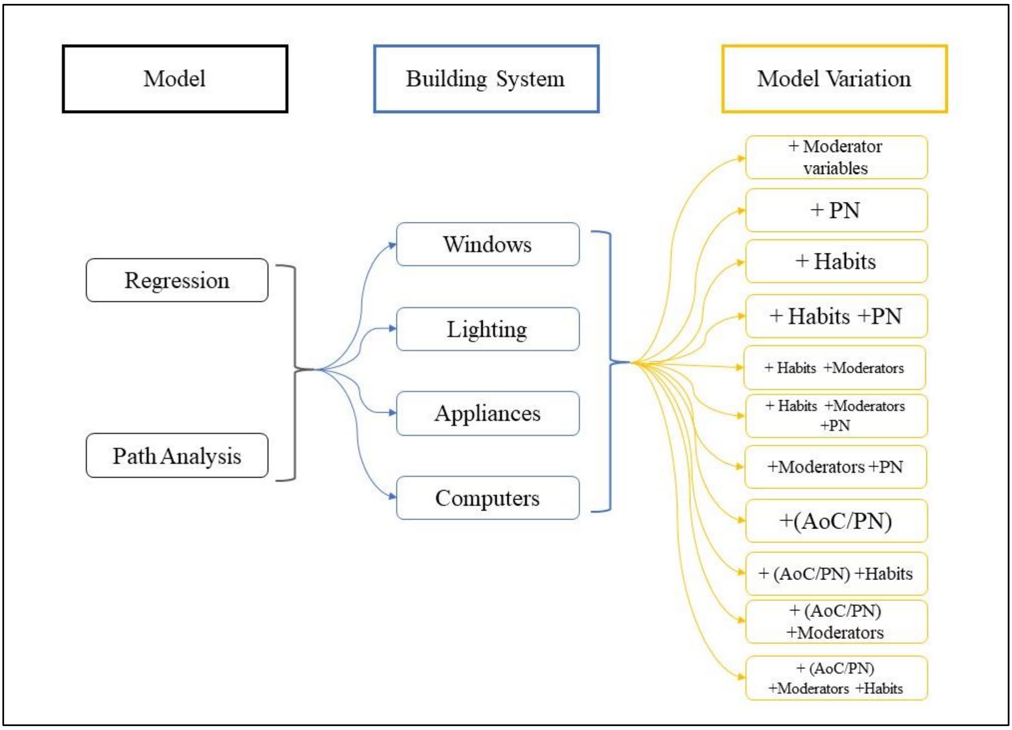

Figure 3-10. From Paper 4: Model Variations

3.3.6 OB profiles

The dataset and resulting influence factors were then used to create occupant profiles by using clustering methodologies. These methods are unsupervised machine learning algorithms that use the similarities or dissimilarities between data groups to categorize them together. The two clustering techniques considered were k-means and k-modes clustering. For k-means, the clusters are based on mathematical measure, i.e., the distance between the centroids of each data group. The centroids are computed by the means of each group. This technique is more suitable for continuous measures, while k-modes is more suited for categorical variables.

Here, the modes (most occurring values) are used to gauge the dissimilarities and mismatches between data points. Both SPSS and the programming language Python were used to perform these clustering analyses

In order to pick the optimal number of clusters (k), an elbow score was used. This score is calculated based on the sum of squared distances from each data point to its centroid (distortion score). A cost curve is then plotted with the number of clusters on the x-axis and the distortion score on the y-axis. The line that is drawn on the chart has a point of inflection (elbow) that indicates the optimal number of clusters.

3.4 Agent-based model

The objective behind this OB model was to develop a proof-of-concept model that can simulate behavior based on the social aspects and incorporate occupant diversity. The building system that gave the most consistent results from the survey was windows, hence window-opening behavior was targeted in this model. The environment used in this project to create agent-based models was Netlogo[111]. It is an open-source programmable environment for simulating natural and social phenomena, particularly well-suited to handle complex models that develop over time. It uses the Logo dialect for directing the agents and runs on the Java Virtual Machine. A review of different agent-based modeling tools took into consideration several aspects of these tools, from the user-interface to their ability to handle complexities. Netlogo was found to have low modelling effort required to deliver models with medium to high-scale complexities [112]. The language and user interface were one of the easiest, and yet capable of simulating very complex environments. For the purposes of this study, it was preferred to develop a model that would be accessible and available to people without requiring expertise in using or changing the model. Hence, Netlogo was considered as the most suitable option. The following sections delve into the creation of the model, based on the findings of the previous sections.

3.4.1 Environment setup and agent creation

Traditional BPS tools, such as eQuest and DesignBuilder allow for the construction of a three-dimensional layout for building models, providing physical boundaries for the model to take place within. One drawback of using an agent-based model is the lack of access to such a tool, which requires the user to use their own interpretation of the modeling environment to generate a building layout. Netlogo consists of two main elements: ‘turtles’ (mobile agents) and a grid of 'patches' (stationary agents), within which the turtles move and perform actions. The grid is adjustable, and a default 32 x 32 grid was utilized here, which contained 1024 distinct patches. Each patch is associated with a label (*plabel*), color (*pcolor*), and is distinguished by the coordinates (*pxcor* and *pycor*) on which it lies. These associated variables for patches were used to construct a building office type layout wherein the agents would associate the patches with specific variables as elements of a building. The sample layout that functioned as an environment for the agents to move around and perform actions in is shown in the Figure 3-11. It consisted of 16 *office-zones* and one *central-zone*. Each *office-zone* was modeled as a shared office, occupied by two or more people. It also consisted of a *workstation* for each person present inside the zone, as well as a *window* for the person to interact with. The *office-zones* were all equipped with a *door* that was connected to a *lobby-zone*. The *central-zone* was created to represent an area that would consist of the entrance, stairs, restrooms etc. These were not vital to for the window-opening simulation, hence confined to a single area.

The layout was loosely based on an office building in NTNU (Norwegian University of Science and Technology) Dragvoll campus, and was selected because it maximized the access to windows for all the people present in the office. It is by no means an extensive representation of an office building layout, but ideal for the scope of simulating the interaction between occupants and windows.

The next step was the creation of 'turtles', or mobile agents, which would execute actions. Similar to patches, agents are associated with distinct variables as well, which included an *id, size, label, heading, coordinates, shape,* and *breed.* One functionality of an agent-based model is that agents can be 'asked' to follow certain rules by assigning them to a particular *breed*. Each breed will have to be first defined with respect to the rules it will have to adhere to. Two distinct types of agents were created in this model: people, and windows. The windows were created as agents so they could record the interactions with occupants and denote themselves as open/closed. Furthermore, there were 8 different types of people created, based on the OB profiles generated in 3.3.6.

Upon creation, a person would be assigned the following attributes: my-type, which would denote the OB profile the person belonged to; my-window, the window assigned to the person; my-office, giving the office number of the person; my-office-mates, the other people sharing the office with them. The windows were assigned two attributes: my-person, and my-state (open/closed).

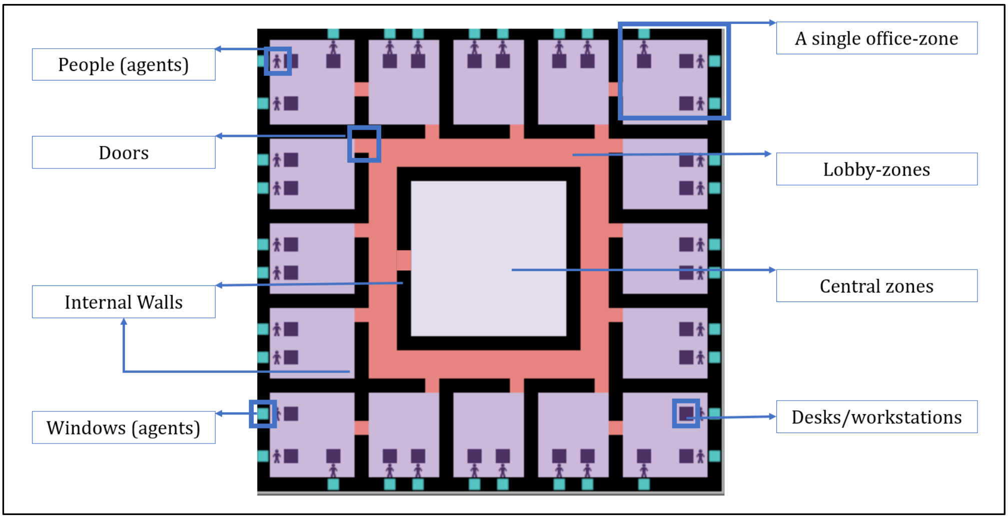

Figure 3-11. Office-floor layout and elements in netlogo model

3.4.2 Adding temporal elements

The need for a temporal element was necessary to dictate when an agent should execute a certain action. Normally, the simulation keeps running till it executes the block of code, or until it is stopped by the user, with no accounting for the number of simulation runs that have passed. However, there is a provision of a feature of 'ticks', where the user can design a tick counter, which gets advanced according to the model design. For this model, the ticks were used to create a timeframe that consisted of hours, *days*, *months*, *seasons*, and *years*. Each advance of the ticks was associated as the passing of one *hour*. After 24 hours, the ticks would be reset to zero and the process would start again. In addition, the passing of 24 hours would denote that a *day* has passed, and a counter for *days* would record this change. Similarly, counters would keep track of *months*, *seasons*, and *years*. The day would also consist of *working-hours*, meaning that all the activity is confined to this time (since the simulation is of an office).

3.4.3 Movement simulation

An additional model generated for a different office layout attempted to simulate movement. The goal was to provide a possibility to cover another aspect of OB with agent-based models. The movement simulation commenced with the occupants generating links with their associated workstations. Upon creation, each occupant already had a link with their associated desks and windows. The *heading* variable of the occupant is a value that denotes the direction that an agent is facing in, and is measured in degrees (0-360), where the North direction would be 0/360. The path selection for the agents was based on a multilayered decision-making process, wherein the agent examined the patch ahead of it before moving forward by one step. If the patch ahead belonged to the *lobby-zones*, it would continue moving forward. If it belonged to the *internal-walls*, it would take a step back, reset its *heading*, check the patch ahead again, and move forward or backward where it could. It would have to seek out a *door* through which it could enter, to get to its destination. A simplified version of the control of flow and logic in this simulation is shown in Figure 3-12.

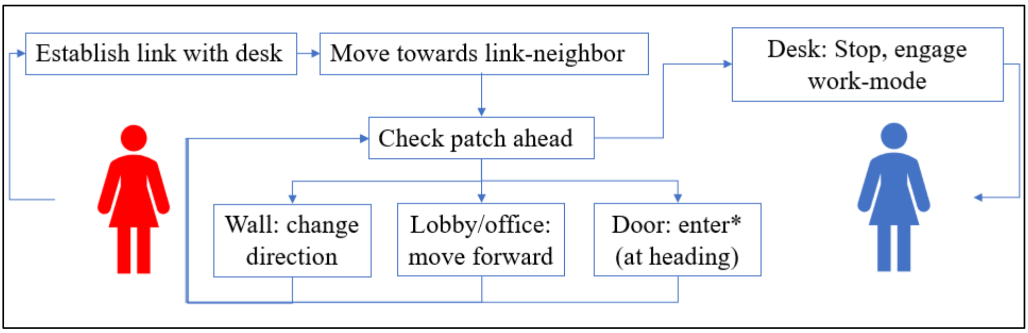

Figure 3-12. Logic behind the movement simulation

3.4.4 Behavioral rules

Once the value of the tick counter was in motion, an agent would be able to use the counter to segregate the *days* and *working-hours*, and perform actions based on it. In accordance with standard ASHRAE schedules, the *working-hours* was set to start between 6:00am – 9:00am. During this temporal range the agents would 'arrive' at their respective workstations, either using the movement rules specified above, or just be generated (or 'spawned') next to their workstations. The next temporal range to take into consideration was the end of *working-hours*, which would assign the agents as 'not working'.

The next consideration was the interaction with the building system, in this case, the windows. This interaction depended on the *season*. The occupants would open windows for a duration that would be suitable based on the season. This meant that windows would be open for longer during *summer* (around 6 hours), and much shorter during *winter* (around an hour). The window-opening duration was selected based on case studies on shared office spaces in different seasons. During the working hours, each agent would ‘open’ the window once till the hourly target based on the season is fulfilled. In order to create a baseline model, all occupants had to act in the same manner, thereby adopting an aggregated behavior pattern where no deviation from the norm was allowed. Realistically, human occupants cannot be expected to behave as a programmed entity, which is why this model was designed for the inclusion of occupant diversity. Within an agent-based model, it is possible to design each agent separately. This model takes into consideration a continuous distribution, wherein the traits of window-opening behaviors are across a range. Occupants would be assigned a ‘type’, and their chances of opening a window for the set duration would depend on their type. Type 1, for example, was the standard one, wherein the occupant would use the rules for the baseline model for their actions. Types 2-6 would have a decreasing chance of adhering to the baseline model, where they would be assigned a probability to act according to its rule (type 2: 80% probability, type 3: 60%, and so on). Types 7 and type 8 were more dynamic ones, which would be based on the actions of their office-mates. Their willingness to act the same as, or opposite, to their peers depended on the strength of the subjective norms that was calculated in the SEM/regression models. The logic of the simulation is visualized in Figure 3-13.

The proportion of each type of occupant was kept flexible, in order to provide the user with control over the simulation. For that purpose, a slider was added to the model, which could be used by the designer to set the composition and vary the amount of any specific personality trait. Extensive testing and analysis can be carried out in this manner, utilizing different compositions to determine occupant-window interactions associated with different occupant diversities.

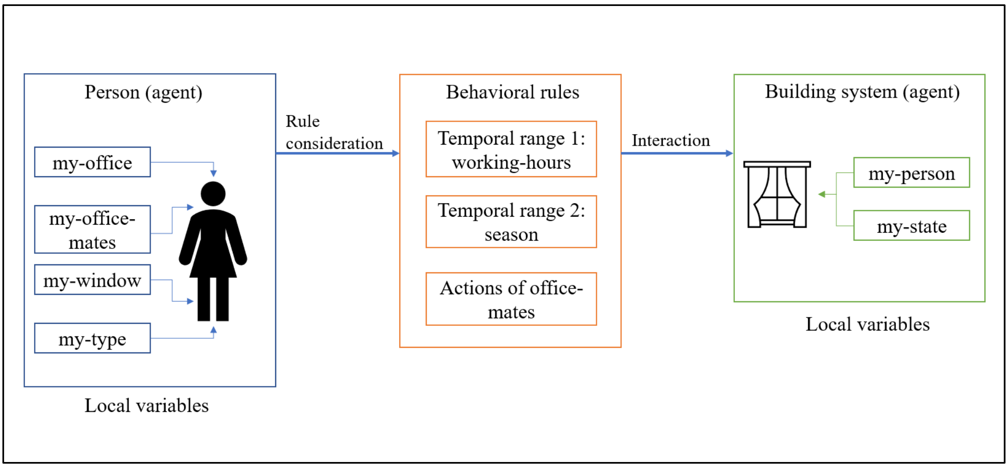

Figure 3-13. From Paper 4: Logic and flow of the agent-based model

4 RESULTS AND DISCUSSION

This chapter describes the results from the studies conducted as part of this thesis, and their association with the research questions raised during the commencement of the work. Section 4.1 compiles the results generated from the case studies used to monitor occupants along with their actions and influences, addressing Q1. Section 4.2 briefly summarizes the outputs from the library and database development and contains insights regarding Q4. Section 4.3 describes the results of the social framework, surveys, and path analysis models targeting Q2. Lastly, section 4.4 describes the proof-of-concept model that was developed using agent-based modelling, exploring the implications of such models and the outputs gained from them (Q3). It also includes a summary of the deep learning models that were based on the datasets from the case-studies.

4.1 OB Monitoring

The case studies for OB monitoring were conducted as described in the previous chapter, in three different settings, each with their own set of goals. The results from the shared offices are presented in 4.1.1 and include the relationship between occupancy and the electric load, energy consumption at occupant and device level, and the diversity profiles. The next subsection details the findings from the sensors used in the operating room at St. Olavs, and its implications regarding occupant movement and activity levels in surgical rooms. Subsection 4.1.3 describes the activity dataset from the living lab and the contribution to benchmark datasets.

4.1.1 Occupant energy-use behaviors in shared office spaces

Box-and-whisker plots were used to analyze the relationship between the occupancy in the shared offices and the electricity consumption Figure 4-1. The central, red-colored lines represent the median consumption, and the blue boxes mark the 25th and 75th quartiles. Along with the red-colored lines outside representing the outliers, the figure shows the distribution of the electric loads measured using the plug-load and lighting sensors.

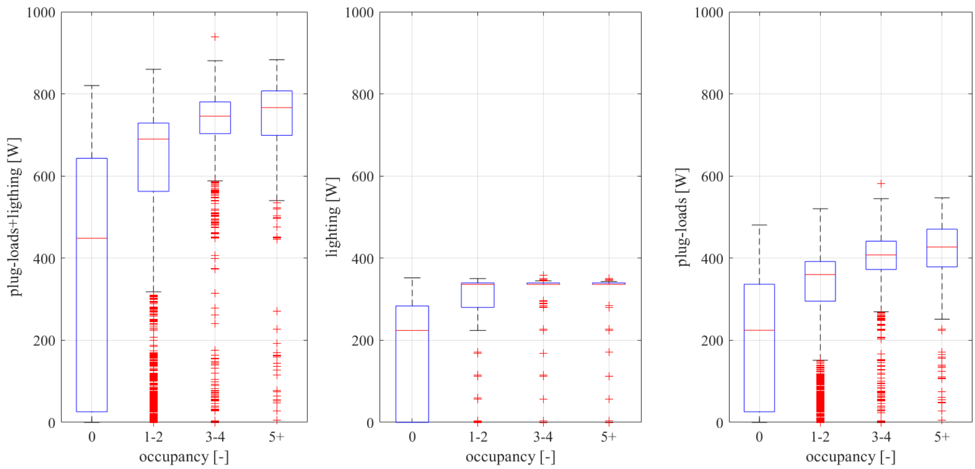

Figure 4-1. From Paper 2: Box-and-whisker plots of electric load function of the number of occupancy sensors that are indicating occupied status

The lighting system in the monitored workspace was found to consume an excessive amount of energy during vacant periods, with mean power values exceeding 200W and 25% and 75% quantiles ranging between 0W and 300W. The system is triggered by motion sensors, which were confirmed to be activated by occupants passing by the monitored workspace. Interestingly, even with five independent occupancy motion sensors controlling the lighting zones, the presence of only 1-2 occupants was sufficient to activate the entire lighting system, with an average power level of approximately 350W. This highlighted the importance of selecting appropriate light control systems and ensuring their proper calibration and maintenance. Similarly, plug-loads were found to exhibit high power levels exceeding 200W during vacant periods, indicating that occupants were leaving equipment running when leaving their desks. However, unlike the lighting system, plug-loads showed a positive relationship between average power levels and occupancy count. Total electric loads were characterized by high consumption values during vacant periods and a moderate positive relationship with occupancy count.

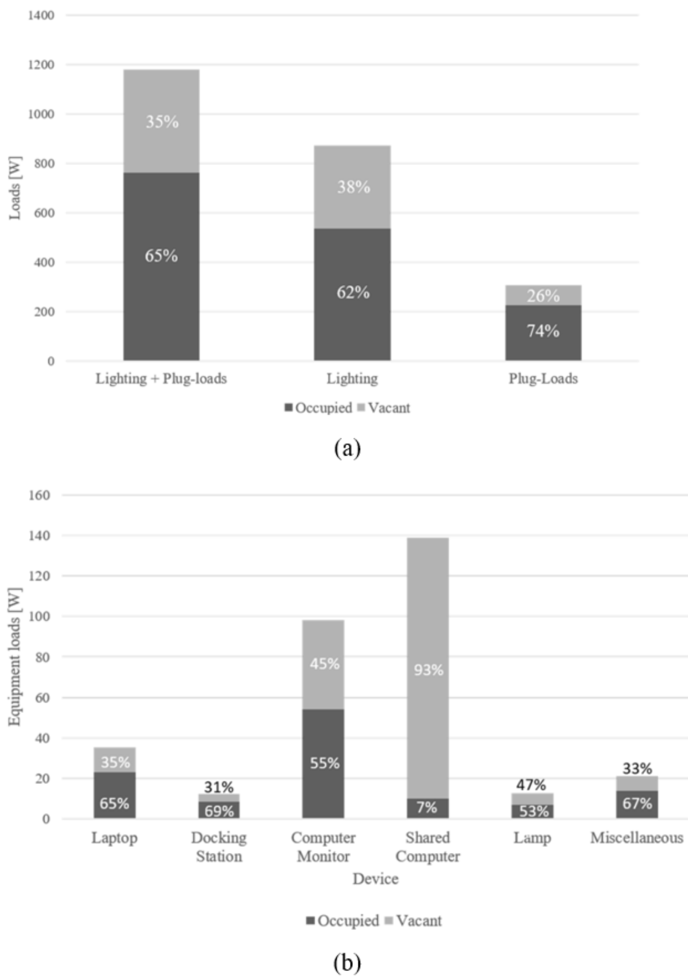

Figure 4-2. From Paper 2: (a) Office-level comparison of energy consumed during vacant and occupied periods. (b) Device-level comparison of energy consumed during vacant and occupied periods

The lab's energy consumption was analyzed during periods of occupancy and vacancy, with results presented in Figure 4-2. The analysis showed that a significant portion of the building's energy systems operate after hours. Specifically, the bar chart on the left side of the figure indicates that 35% of the lab's electric consumption occurs during vacant hours. Further analysis by end-use showed that the lighting system and plug-loads accounted for 38% and 26% of consumption during vacant hours, respectively. However, the observed plug-load values are conservative because the definition of "occupied" period assumes at least one occupant is present. This assumption may have led to overestimation of plug-load energy consumed during operation and underestimation of the energy consumed during vacant periods.

The energy consumption at the device-level granularity was also analyzed, which revealed that 64% of the plug-load energy was consumed when the occupancy status of desks is reported as "vacant", reinforcing the need of occupant intervention in controlling the unnecessary consumption. The shared personal computer was found to be a major contributor to unnecessary consumption with 93% of its energy consumption occurring without the presence of an occupant. The analysis demonstrated that device-wise monitoring can reveal important energy-saving potentials that were not observed when considering occupancy in the office as a binary variable. Overall, the results highlight the need to identify instances of unoccupied consumption, particularly for shared equipment such as printers and appliances, and to curtail energy consumption during vacant periods for monitors and lamps.

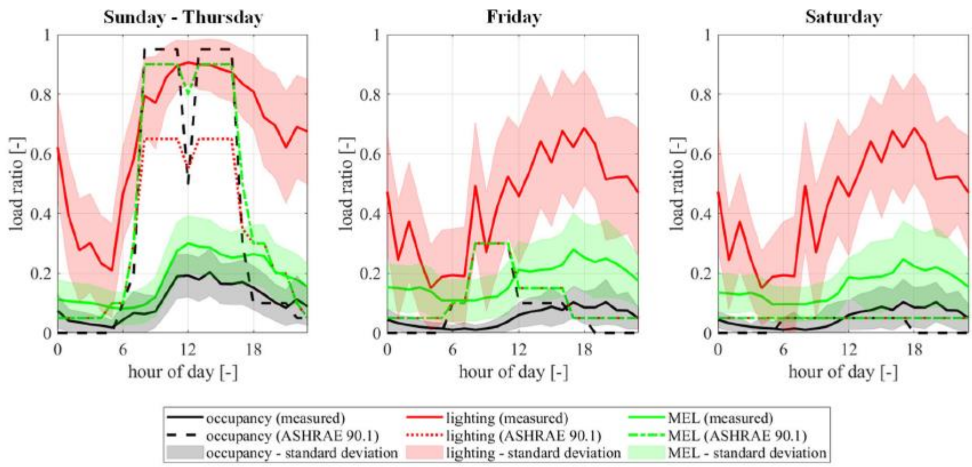

Figure 4-3. From Paper 2: Comparison between the ASHRAE 90.1 profiles with the measured occupancy, plug loads, and lighting consumption

The occupancy and energy consumption patterns of an office building were analyzed and compared to the schedules recommended by ASHRAE 90.1 (Figure 4-3). The study found significant deviations between the standard schedules and the measured occupancy and energy consumption. The absolute hourly error between the theoretical and observed occupancy, lighting, and plug-loads was presented to shed more light on the discrepancies. The error values showed that occupancy and plug-loads were overestimated over the typical working hours of workdays while being underestimated outside of working hours. In contrast, lighting loads were consistently underestimated throughout the week.

The study found that the observed discrepancies could be attributed to several factors. Firstly, the studied area is a shared office space in an educational facility, where the occupants are graduate students with a flexible working schedule. Secondly, the number of occupants studied is relatively low, making the impact of individual behavior significant on the general patterns. Thirdly, schedules, such as ASHRAE's, assume a good correlation between occupancy patterns and the energy use levels of systems, such as lighting and plug-loads. However, there is an important difference in the scale or magnitude of the profiles.

The observed discrepancies can lead to inefficient use of lighting and equipment and contribute to misestimations of the actual energy use levels, resulting in an energy performance gap. This study highlights the importance of considering the unique characteristics of each building and its occupants when designing and implementing energy-efficient systems. A non-intrusive framework with the level of granularity this study demonstrated is beneficial in engaging fault detection in buildings, handling miscommunication from occupants or mismanagement on part of the facility management.

4.1.2 Occupant activity influences in operating rooms

The data collected from the four Kinect cameras was processed and merged to plot an activity heat map. This map had a resolution of cells measuring 5cm by 5cm, and the cells corresponded to geographic locations of the OR. Each cell registers movement in its corresponding location and increments its value by one. No activity is registered for static occupants. The spatial activity distribution map shown in Figure 4-4 was developed in this manner.

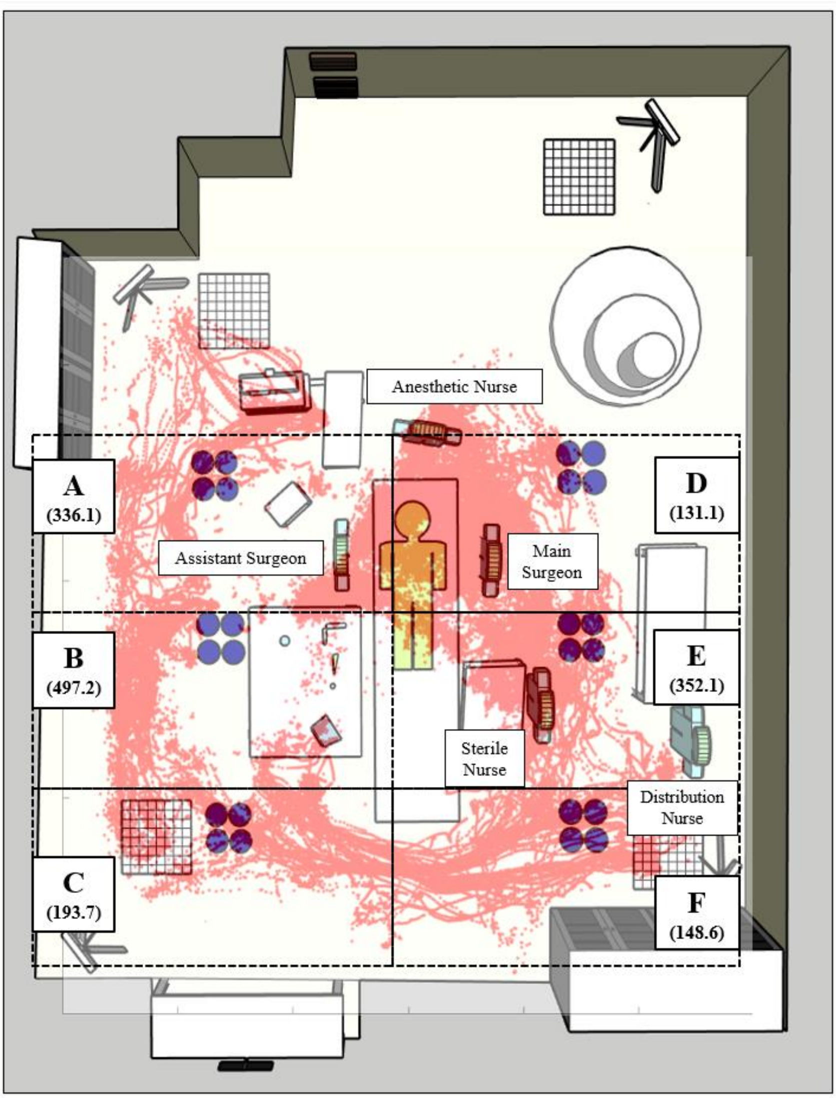

Figure 4-4. From Paper 6: Activity levels around the surgical site

For the data regarding bacterial contamination, the agar plates measured the number of Colony Forming Units (CFUs) that were formed by settling on the exposed surface of the agar plates. These measurements were standardized for comparison in the form of CFU density (CFU/m^2/hour). Table 4-1 summarizes the CFU levels at each of the locations along with the percentage of activity level during a phase. The average CFU levels at each location are also indicated in Figure 4-4.

The results indicated that while activity levels played a role in the distribution of contamination, other factors also had an impact. For instance, location E had consistently high activity levels due to the presence of the Distribution Nurse, who had to move across the room, bringing particles from other areas into that location. Higher CFU densities were found in locations A, B, and E, where the surgical staff were more active, compared to locations C, D, and F, where the anesthetist nurse had less activity. The CFU density in location B was particularly high, which can be attributed to the stronger activity levels of the assistant surgeon. This participant's roster of activities included squatting at the location as part of the mock surgery procedure and was responsible for opening the lid of the agar plates at A, B, and C. Additionally, the obstruction caused by a large table in the area may have obstructed air flow, leading to higher CFU densities.

Experiment Phase	Phase 0-3		Phase 1-3		Phase 2-3		Phase 3	
Agar Plate Location	Activity Level	CFU/ m2/h	Activity Level	CFU/ m2/h	Activity Level	CFU/ m2/h	Activity Level	CFU/ m2/h
A	12.7%	302.15	20.9%	264.34	18 %	302.02	24.7%	**476.03**
B	0.5%	**377.68**	8 %	**469.94**	10.3%	**855.72**	10 %	285.62
C	1.3%	151.07	0.9%	176.23	0.5%	352.35	3.3%	95.21
D	23.9%	50.36	**37 %**	88.11	**36 %**	100.67	**30.1%**	285.62
E	**60 %**	302.15	31.5%	323.08	34 %	402.69	28.9%	380.83
F	1.6%	201.43	1.7%	146.86	1.2%	151.01	3 %	95.21

Table 4-1. From Paper 6: CFU measurements and their corresponding activity levels

However, irregularities were also observed. For example, location D was expected to have high CFU due to its proximity to the main surgeon and sterile nurse, but no distinguishable CFUs were found. This might indicate a lack of sufficiently large agar plates to capture the airborne contamination in the area. The overestimation of CFU counts during the hypothesis of the study proved to be a significant limitation. While it was expected that the highest CFU levels would correspond to the areas with highest activity levels, the results show that this was not always true.

4.1.3 Occupant activities in residential spaces

The case study in the Living Lab resulted in an extensive dataset labelled as *ADL SmartLab Dataset*. The aim of establishing this dataset was to build an Activity Recognition framework that represents activities as 3D trajectories of the 25 body joints of the human torso (also known as a Skeleton Model). The developed framework was able to segment and identify the performed 14 activities (comprised as 22 discrete atomic activities. The deep learning models used for this framework were Long Short-Term Memory (LSTM), Bi-directional Long Short-Term Memory (Bi-LSTM), Gated Recurrent Unit (GRU) and Bi-directional Gated Recurrent Unit (Bi-GRU).

It should be noted that this case study forms a part of the supplementary works as described in the Introduction (Figure 1-3). The resulting deep learning models are described in further detail in the attached papers (Paper 8 and 9 in Supplementary works). The contributions from this study to this thesis comprise of a literature review regarding the use of machine learning and deep learning models for AR (Figure 2-2), and the generation of a dataset for occupant activities.

4.1.4 Answer to Question 1

Based on the above findings, a summary of the case studies conducted is presented in Table 4-2. An answer to Research Question 1 can be formulated as follows:

Question: **How to monitor and analyze the impacts of occupant behavior in different spaces?**

Answer: It is important to note that the question presented here is not unique in its goal. Rather, multiple studies have examined it within their respective contexts. However, despite the generic nature of the question, it still merits discussion due to the variation of its answers depending on the context and setting. The studies presented highlight the importance of considering the unique characteristics of each building/environment and its occupants while designing an OB monitoring system. It also demonstrates that such monitoring systems can vastly benefit from using multiple modalities of sensors, which makes it possible to obtain detailed and individualistic data. The use of the sensors also addressed the need for privacy and non-intrusiveness of the monitoring system.

Each study provided a unique perspective on the answer. The first study achieved a framework capable of examining granular device-level, office-level and occupant-level data in a shared office space while exhibiting its non-intrusiveness. The second study made use of an advanced monitoring system to examine occupant influences in an operating room. The last study made use of a similar monitoring system, but with procedures aimed at capturing an extensive activity-based dataset suited for smart homes and activity recognition systems.

The complexities in developing standard OB models often arise from lack of benchmark datasets, since each study is focused on a specific case and the models are based on its own dataset. This creates difficulties for researchers to compare models or provide meaningful information for the building design stage. One of the objectives of conducting these case studies was to contribute to benchmark datasets that would be available for wider use. Table 4-2 summarizes the datasets resulting from the case studies.

Area of study	OB aspect	Sensing modality	Dataset generated
Shared office	Energy-use, presence	PIR, Plug Load sensors, surveys	Miscellaneous Electric Loads (Occupant-centric)
Operating Room	Activity influence on surroundings	Depth registration	OR Activity Influence
Residential space (Living Lab)	Activities of Daily Living	Depth Registration	ADL dataset

Table 4-2. Summary of case studies

In addition, ASHRAE invited researchers to contribute to a global database for OB to address the same challenges. This database was processed and launched in 2022 in *Nature's* Scientific Data, consisting of 34 datasets that were selected as representative samples for different aspects of OB [93]. The dataset from the study in shared office spaces was selected as one of them and is now available online. The dataset on OB activities has been submitted as an open dataset as well.

4.2 Database for layout simulator and BOT-ABM

This section describes the database that was developed to facilitate exchange of information regarding different aspects of occupant behavior for an OB model. It also provides a sample of the results achieved from the Matlab application built for data collection purposes. The methodology for this part of the thesis consisted of (i) identifying key variables needed in a library for layout simulation, (ii) establishing an application for the collection of such data, (iii) data collection at NTNU. The next step was to create a functional database that could be used by OB models, such as the hypothesized BOT-ABM.

4.2.1 Application for data collection

The application for the spatial data collection of occupant's environment consisted of two pages. The first one asked the general information regarding the room, such as the occupancy type, order of importance for placing the objects around the room, the area etc. The user interface, presented in the previous chapter (Figure 3-5), resulted in a data collection format as shown in Figure 4-5. This interface generated a grid based on the input of the room area, dividing it into cells. The occupant would then have to proceed by highlighting the areas inside their rooms that are occupied by a certain object. The object list provides a list of common objects, specified for the room type. This list was by no means exhaustive, but users had an option to specify what other object was used.

Figure 4-5. Sample output from the Matlab application

4.2.2 Database design

Databases created through MySQL consists of tables that are related to each other via primary and foreign keys. A 'table' consists of rows of information, wherein one type of information is assigned as a 'primary key', unique to the table. In most cases, this is an id for the row of information. For example, in a table containing data about a room, the id could be the room number, and designated as a primary key. Tables consisting of information about other parts of the building can now have a 'foreign key' that is connected to this room number. Use of this foreign key will indicate which table to access to extract information. This helps with eliminating data redundancy, since only the desired set of information is accessed for an operation. Additionally, the keys might be in the form of one-to-one relationships, or one-to-many. The former links a single record in one table to a single record in another, whereas the latter connects a single record to multiple records. The database schema (Figure 4-6) provides a graphical representation of the information present in the tables and the connections between them.

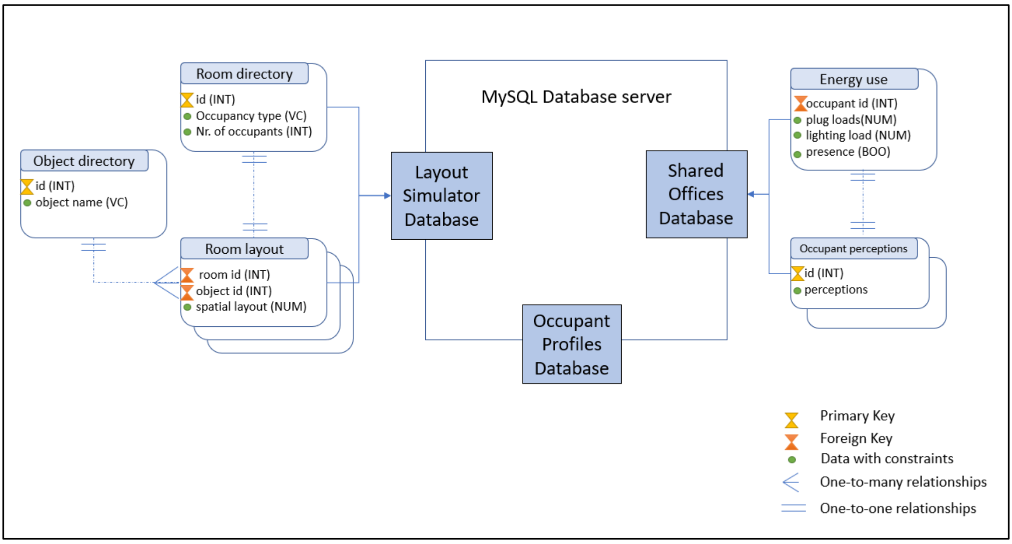

Figure 4-6. Database schema for the collected datasets

The database created for surface layouts consists of three tables wherein the parent table in the 'room directory'. This table contains a generated id for the room, and additional details about the room such as the number of occupants and occupancy type (private/shared). The id in this parent table is connected to a table that displays information about the spatial layout of the room. There are multiple layout tables, and each of them has a one-to-one relationship with the parent table via the room id. It also has the information about the objects located in the room, denoted by specified object id. These ids are connected to another table with information about the objects. Hence, information about objects and rooms has to be entered only once in the relevant table and just connected to the layout, where information is represented with a single- or double-digit id. These connections contribute to avoiding data redundancy and reducing computational power. Instead of having a number of irrelevant details in each file, the relevant information can be accessed via queries.

Similarly, as shown in the schema, the shared offices database contains the dataset from the case study, wherein the tables reflect the actual and perceived energy-use behaviors of occupants.

4.2.3 Answer to question 4

The research question addressed in this section can be summarized as follows:

Question: What kinds of prerequisites are needed to map environmental layouts around the occupant?

Answer: The response to this question took into consideration the movement of an occupant inside a room and their interactions with the objects present inside it. Prerequisites needed to map these occupant environmental layouts were identified in the form of a library that contains sufficient information for a surface layout simulator. The resulting study had three goals, the first one being the identification of seven variables for the library, presented in Table 3.1 in the previous chapter. Secondly, a Matlab application was developed for collecting spatial information about the environment, the interface and sample output of which is described above. These ideas were tested by collecting a sample dataset of 80 offices at NTNU. A subsequent part of this answer was the development of a database to connect the information generated from this study to larger models, which was then expanded to include other datasets from previous studies as well.

4.3 Social Aspects of OB

This section details the results obtained to answer question 3, regarding the social aspects of OB. To that end, the study explored the literature for quantification of social aspects and developed a framework and extended-TPB model to address it. Using these, a questionnaire was created and distributed for data collection. The data analysis (overview in Figure 3-7) consisted of three steps (i) Reliability Analyses, (ii) Path analysis/ SEM/ Regression models, and (iii) OB profiles. The following subsections 4.3.2 – 4.3.4 describe the results obtained for each of these. Subsection 4.3.1 describes the results for the addendum to the case study on shared offices (Paper 2).

4.3.1 Addendum to the case study in shared office spaces - Results

The reason for including this addendum in this section is that the topic revolved around social perception of energy-use behaviors. The study analyzed occupancy schedules and device utilization in a building using sensors and occupant surveys. The actual schedules obtained from the sensors differed significantly from both the perceived and standard schedules, with one of the biggest deviations being a delay in the start of the day. However, occupant perceptions about their presence, obtained from the surveys, showed a smaller gap in this delay and followed the trend of the actual schedules more than the standards. Statistical analysis showed that perceived occupancy had a stronger correlation to actual occupancy, but a significant gap still persisted, as can be seen in Figure 4-7

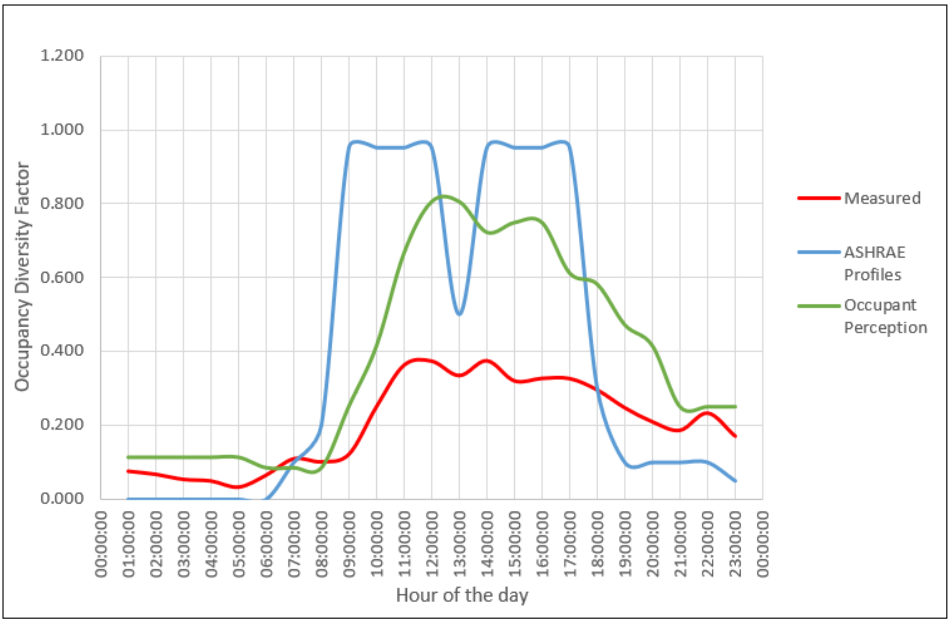

Figure 4-7. From Paper 3: Comparison of Standard, measured and perceived profiles of occupancy

On the other hand, the study found that device utilization did not yield significant correlations. The perceived energy-use of occupants varied greatly from occupant to occupant, as well as device to device. Even devices subjected to frequent use, such as laptops and monitors, showed discrepancies regarding perceptions. For example, laptops were the most consistent, and monitors the least. Occupants that considered themselves to be switching off their monitors while leaving found that their perceptions were inaccurate, according to plug load sensor measurements. These results are highlighted in Paper 3.

4.3.2 Reliability analysis (Measurement modelling)

This section details the results from the first step of the data analysis. As described in the previous chapter (section 3.3.4), the data collection resulted in 101 samples that were deemed suitable for use. A brief look at the demographics indicated that a majority of the respondents were from Norway (33%) and the rest were a mixture of different nationalities. The age groups in the range 21-35 years constituted the vast majority of the respondents, at 85%. Almost all the participants were working in shared offices, which is why the agent-based model was centered on shared offices. Post data processing (for outliers and missing data), the collected data was ready for testing reliability.

A reliability analysis is a test performed on each variable of the extended-TPB model to check whether the theorized model is reliable and valid. For testing this, each variable in the model has minimum two questions. The responses to both needed to be correlated enough, since they measured the same variable. The scores used for this test were *pearsons correlation coefficient* (when the variable had 2 questions) and *cronbach's alpha* (when the variable had more than two questions). Figure 4-8 summarizes the results for this step.

Variables testing for reliability analysis were 'attitudes', 'personal norms', 'awareness of consequences', 'descriptive norms', 'injunctive norms', 'social connectedness', 'identity', and 'self-construals'. Out of these, all except 'self-construals' passed the limit of 0.6 for pearson's coefficient and 0.7 for cronbach's alpha. 'Personal norms' and 'awareness of consequences' scored the highest (0.8 and 0.79 respectively), while 'social connectedness' and 'descriptive norms' scored the lowest (0.61).

Since 'perceived behavioral control (PBC)' and 'habits' targeted different building systems, it was not necessary for them to be highly correlated. These were instead subjected to a confirmatory factor analysis that models observed variables to have unobserved 'factors'. In this test for example, the answers regarding habits had one factor that accounted for lighting, appliances, and computers, and one factor that accounted for windows. This denoted that people's behavior was similar, or stemmed from similar reasons, when it applied to lights, computers and appliances. And this was different from the habits and behaviors while interacting with windows. The scores used to test on confirmatory factor analysis were the p-value for significance, and the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO value). Both these variables scored well to clear the ordained limits of these scores. Extensive results of these tests are attached in the appendix of paper 4.

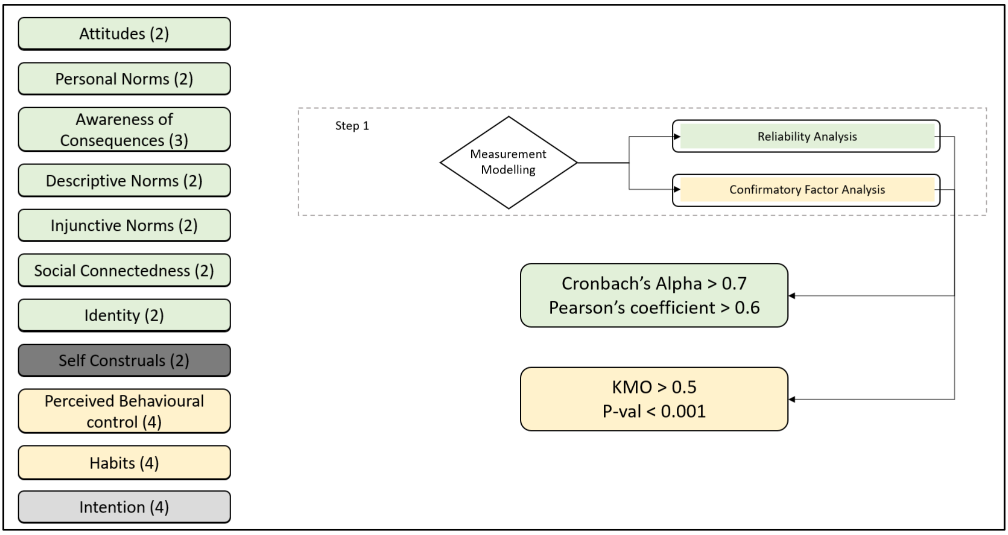

Figure 4-8. Results of reliability analysis

4.3.3 Path analysis/Regression models

Following the reliability analysis, the variables deemed to be valid were then subjected to the second stage of the data analysis, which was aimed at determining the influence of each variable on the intention to perform energy-related behavior. This section summarizes the results obtained from these models. Recalling the variations of the models run (Figure 3-10), each building system had separate models developed to evaluate the behavior associated with it. This was done since the previous steps in the analysis pointed out that the behavior for each building system could have different results and reasoning behind them. In addition, it could provide clearer insights regarding social aspects of OB. However, the results displayed here are those from the path analysis and regression. The SEMs required that the dependent variable (intention to perform an action) would be a latent variable that would be associated with two observed variables. This meant that it was not possible to create a model for each building system. At least two building systems would have to be clubbed together to give a valid output. While this made for interesting observations, it did not fit with the objective of the analysis. However, results from these SEMs are also attached in the appendix (paper 4).

For both path analysis and regression, it is important to assess whether the theoretical model fits with the observed data. To evaluate the fit of a model, several fit indices are used. These indices can be classified into Absolute Fit Indices and Increment Fit Indices. The Absolute Fit Indices provide a measure of how well the specified model fits the observed data without any reference to a comparison model. These include the Chi-square test (χ^2), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR). Increment Fit Indices provide a measure of how well the specified model fits the data compared to a comparison model, such as a null model. Incremental fit indices examine the difference in fit between the specified model and a more restrictive model that assumes no relationships between the variables. These include the Goodness-of-Fit Index (GFI), Adjusted Goodness-of-Fit Index (AGFI), Comparative Fit Index (CFI), and Relative Fit Index (RFI)

A non-significant chi-square (χ^2) value indicates that the observed data fit the specified model well, which is generally desirable. However, it's important to note that the chi-square test is sensitive to sample size, and small or moderate sample sizes can lead to a significant chi-square value even if the model fits the data well. Therefore, in addition to the chi-square test, it's recommended to use multiple fit indices to evaluate the goodness of fit of the model. These fit indices should include both absolute fit of RMSEA and SRMR as well as a minimum of two Increment Fit Indices.

The models were subjected to variations such as testing them with or without the additional variables like moderator variables (Social Connectedness and Identity), Awareness of Consequences (AoC), Habits, or Personal Norms (PN). These variations were conducted since the extended-TPB model was a theoretical one, and it was necessary to assess which parts of the hypothesis rang true. Table 4-3 depicts the best of these results, along with the evaluation metrics (model fit indices). The desired value of these is present in the parentheses beside each of the indices, except the degree of freedom (which needs to be positive but not excessively high). The values in red denote the scores that did not pass the mark for the desired value, whereas the figures in bold represent the best fit achieved.

		Baseline TPB	Model Variations			
			+ AoC/PN + Habits	+Habits + Moderators	+ Habits +PN + Moderators	+Habits + PN
Absolute Fit Indices	GFI (>0.9)	0.81	0.9	0.98	0.98	**0.98**
	CFI (>0.9)	0.12	0.85	0.996	0.99	**0.99**
	AGFI (>0.9)	0.558	0.7	0.87	0.87	**0.9**
	RFI (>0.9)	0.4	0.5	0.91	0.91	**0.086**
Increment Fit Indices	RMSEA (<0.08)	0.02	0.17	0.05	0.05	**0.029**
	SRMR (<0.08)	0.08	0.1	0.18	0.04	**0.059**
	Chi-square (>0.05)	0.000	0.000	0.26	0.26	**0.36**
	Degree of freedom	7.3	4.0	7.6	1.2	**1.08**

Table 4-3. From Paper 4**:** Evaluation of path analysis/ regression models (windows)

All the models shown in the Table 4-3 pertain to windows, since that was the building system that continued giving consistently well-fitted models. This also reflected in the reliability analysis, where the factor analysis confirmed that behaviors regarding windows was different than other building systems. After determining the goodness of fit, a closer look can be taken at the models to understand the influences of each variable on occupants' intention to act. Figure 4-9 shows the results from the two best models, one from regression (+Habits +PN), and one from path analysis (+AoC/PN + Habits). It can be seen from the results that 3-4 of the variables were statistically significant, i.e., had a p-value less than 0.05. This was in line with other models seen in literature, where not all of the variables have statistically significant relationships but are nevertheless important for drawing interpretations. The numbers indicated beside the path line (arrows) of each variable denotes the path coefficient (between -1 to 1 for standardized estimates). This indicates the amount of variance that will occur in the dependent variable when the independent variable changes by one. For example, from Figure 4-9, the path coefficient of 'Attitudes' is 0.03. This means that if the value of attitudes changes by one, the intention will vary by 0.03. Negative signs indicate the direction of the relationship.

The results indicated that around 60% of the variance in the intentions is covered by the independent variables in the model. From all the variables, 'Habits' was the one that had consistently strong influence on intention, in addition to always being statistically significant. 'Injunctive Norms' reported the second highest scores, while 'Personal Norms' displayed a strong influence, albeit in a negative direction. The path analysis also proved the 'AoC' variable to be statistically significant, as well as a strong predictor of 'Personal Norms'. This interpretation was applicable to the windows-related behavior. Since (IN) was a variable that measured people's willingness to act according to the expectations of their peers, this gave an important insight into people's behavior.

Analysis of the models for lighting, appliances, and computers revealed slightly different results. However, 'Habits' was still the major influencer across all the building systems. 'Descriptive Norms' proved to be a good influencing factor for lights, whereas 'Personal Norms' was more strongly associated to appliances. The windows model was also the one with the best fit, and hence was selected as a baseline for developing the agent-based model. The best ones of these models are summarized with the respective coefficients of the variables in table 4-4.

Variables	Building System			
	Windows	Lights	Computers	Appliances
Attitudes	0.03	0.01	0.01	-0.02
Personal norms	**-0.17**	**0.06**	**0.06**	**0.20**
Descriptive norms	0.00	0.10	0.01	-0.01
Injunctive norms	**0.25**	**-0.13**	**-0.01**	**0.02**
Perceived Behavioral Control	0.10	0.06	0.02	0.02
Habits	**0.70**	**0.82**	**0.70**	**0.72**

Table 4-4. From Paper 4: Variable coefficients for different building systems

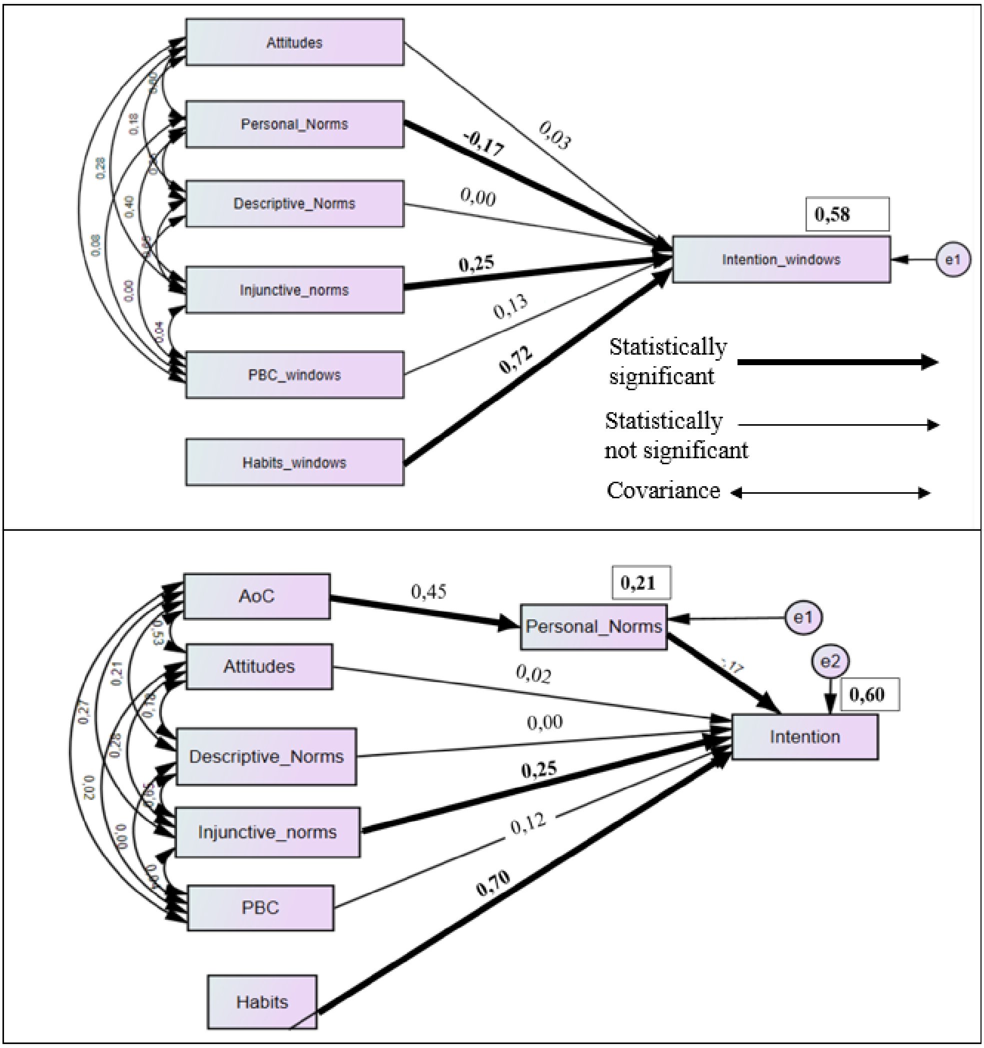

Figure 4-9. From Paper 4: (top) Regression with +PN +Habits (bottom), Path Analysis with +AoC/PN +Habits

4.3.4 OB Profiles

After determining the reliable variables as well as their influences, the processed data was used to generate OB profiles. The goal of developing these profiles was to have an optimal number of occupant types to use in the agent-based model, as well as adjust their defining characteristics based on their profile. The methodology for generating these profiles was k-means and k-modes clustering, the latter of which was more suitable for the dataset used in the study. Establishing an optimal number of clusters was essential, since the data could be adjusted into 5-20 clusters. Elbow score was used in this case, which plots a cost curve as shown in Figure 4-10. The x-axis represents the number of clusters, while the y-axis plots a distortion score. The drawn curve has a point of inflection that indicates the optimal number, which in this case, was 8 clusters.

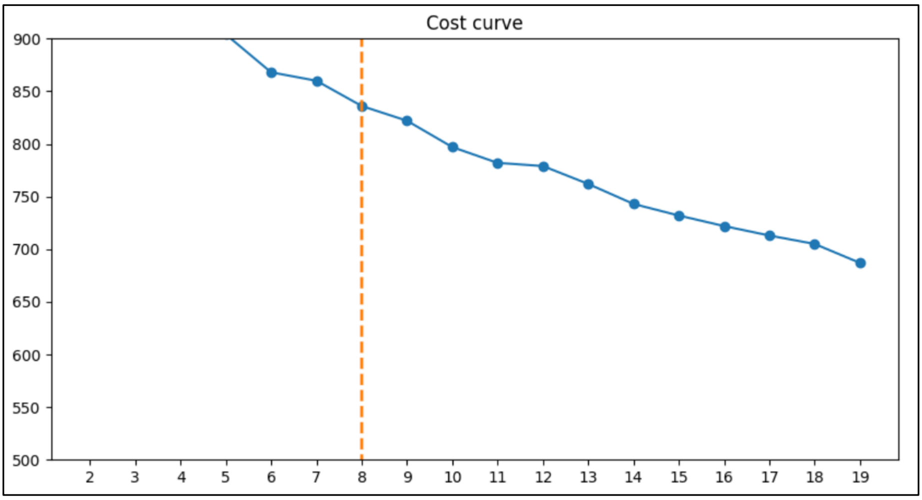

Figure 4-10. From Paper 4: Cost curve to determine optimal number of clusters

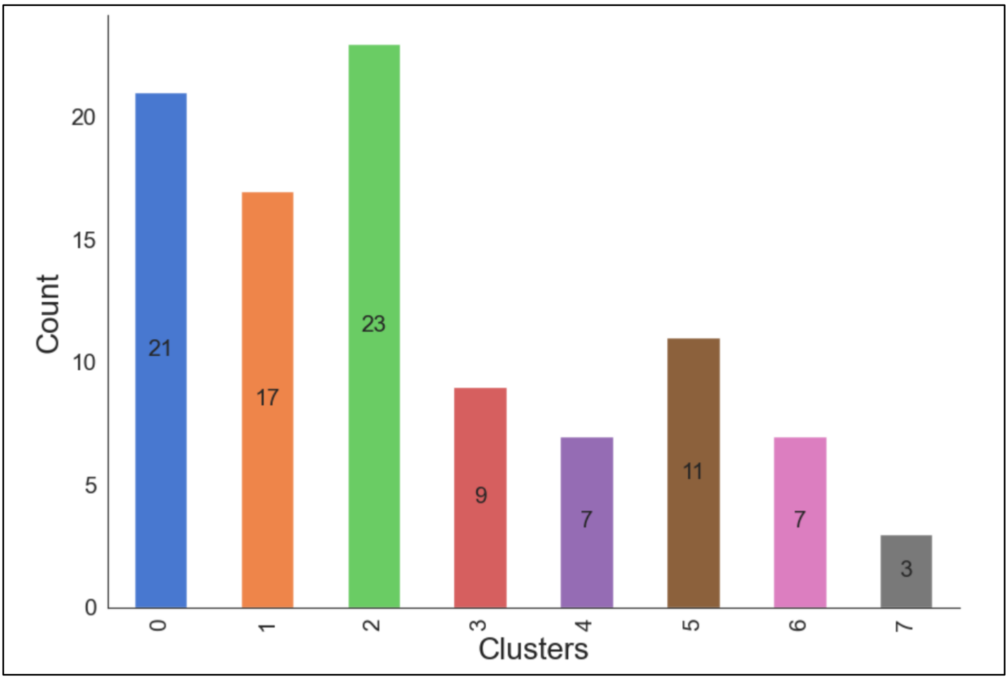

Figure 4-11. Distribution of clusters/ OB profiles

Further analysis of these 8 clusters (or profiles) is shown in Figure 4-11. This represents the amount of people present in one cluster. These were divided based on their responses, where the scores for each response was used to denote a centroid of a cluster. The densest cluster was number 2, with 23 people who gave similar responses, while number 7 was the least dense with just 3 people in it.

4.3.5 Answer to question 2

Based on the results obtained from these studies, an answer to research question 2 can be formulated as follows:

Question: How can the social aspects of OB be evaluated for modelling/simulation purposes?

During the formulation of the research questions, this question was divided into three tasks in order to be addressed. The study presented provided answers to each part, starting with a hypothetical framework and extended-TPB model developed to start the process (Task 2.1). This proved meaningful in quantifying the social aspects. The second task (Task 2.2) answered this question by deploying the framework to collect and analyze the social aspects.

The analysis resulted in eliminating some factors, such as ‘self-construals’ from the extended-TPB model. Rest of the variables were proved to be reliable and valid based on reliability analysis and confirmatory factor analysis (Figure 4-8).

The investigation process included the creation of structural equation models, regressions, and path analysis. Regression proved to be the model that provided the best fit, while path analysis was also acceptable (Table 4-3). SEMs did not result in eligible models due to the lack of segregation in building systems. The influence factors for occupant’s energy related behaviors were obtained for each building system, highlighted in Table 4-4. The addition of ‘Habits’ strengthened the model significantly, where the variable was the strongest predictor for all building systems, in all model variations that were used.

Lastly, the construction of the framework and subsequent evaluation of all the variables enable the creation of OB profiles (Task 2.3). These profiles were now backed by analysis that emphasized which variables were valid and significant enough to justify the clustering techniques use. The k-modes clustering technique yielded suitable clusters, elig.ible to be used for modelling purposes.

4.4 Agent-based modelling

This section describes the resulting model using the agent-based modelling in Netlogo. It includes the interface of the model, the applicability, and the simulation of window-opening behavior. It also includes the result from the movement simulation that was modelled as a separate entity, with capabilities for integration with the larger model.

4.4.1 Model interface

The idea behind creating a proof-of concept model was to initiate the process of bridging the social behavioral theories with the engineering methods of modeling OB. The aim was also to provide an open model that would have a degree of flexibility for the user to control and run several iterations of the window-opening behavior. For this to be eligible for use to other researchers, especially those involved in social sciences, the model had to be simplistic and user-friendly, with blocks of code that ran on the press of a button, and the parameters are adjustable in the form of sliders. The interface is presented in Figure 4-12. The *Setup* and *Go* buttons are those common to most agent-based models, the first run setting up the parameters to the desired settings as a starting point, and the second one running the simulation. In this example, the starting point was the generation of the layout, the agents, their personalities, and the distribution of occupants.

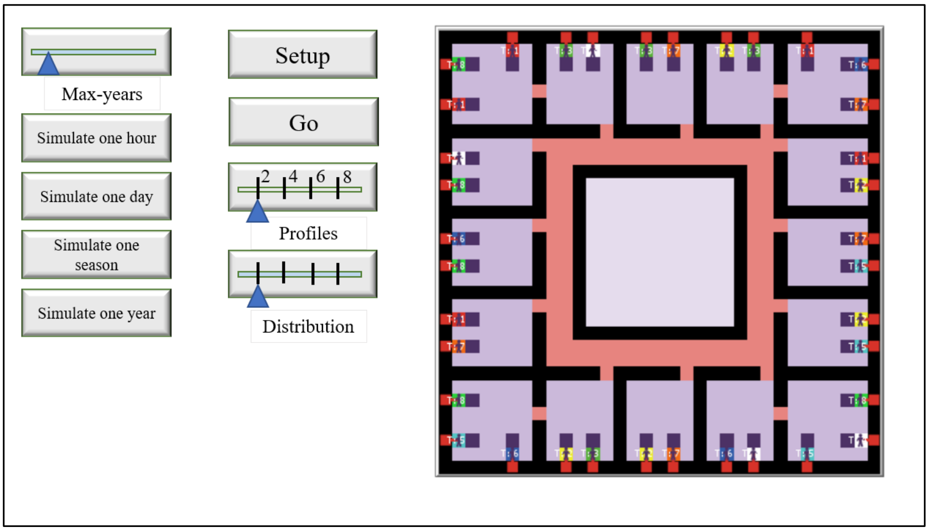

Figure 4-12. From Paper 4: Interface of the agent-based model in Netlogo

Several provisions are present in the model to investigate the window-opening behavior. The resulting profiles from previous steps were modeled into 8 different personalities, or 'types' that each occupant would be assigned. However, the model had a slider to choose the number of profiles to run the simulation on. Moreover, the distribution here was selected to provide a sufficient number of types to run simulations with good amount of occupant diversity. This feature was also kept flexible to a certain extent, where 4 different distribution types were present. This can be expanded and scaled to include more types and distributions as well.

4.4.2 Simulation of window-opening behavior

The window-opening behavior would be carried out by each agent (occupant) in accordance with the rules. These rules were the ones associated with their respective types, as well as general rules about window opening durations during different seasons. The simulation could be done from the period of one hour, up to a period of 10 years. The slider for 'Max-years' gives the user the option to choose the maximum number of years they would like to run the simulation for or set a maximum limit. The user could also choose to run the simulation in increments of one hour, one day, one month, or one season. A season would last for a period of three months, hence one year would have equal distributions for summer, fall, winter, and spring. Each season instructed the agent with a standard opening duration of windows during the workday. The probability of the agent of following the instruction would differ from type to type, which would mimic the OB profiles developed. For example, a person belonging to a cluster that scored higher on window interaction would have more chance to follow the instruction.

Based on these behavioral criteria, the simulations would run for the cycle specified in the model interface. The data would be shown in the form of a graph that summed the number of hours a window was open throughout the day (the increment was originally for each hour but was changed to one day to be more visually evident). Figure 4-13 depicts graphs illustrating the duration of window openings on each day, one for each sort of occupant/agent. Type 1 depicts the typical baseline model, with a graph made up of straight lines. The agents in the baseline model would simply open the window for a set amount of time each day based on the season, resulting in a deterministic model. Winter seasons would be shorter than summer seasons, and so on. The remaining graphs can be observed, with some variation in behavior. The simulation results were for a complete year, with three months between seasons. However, for clarity, the figure only includes one month from each season.

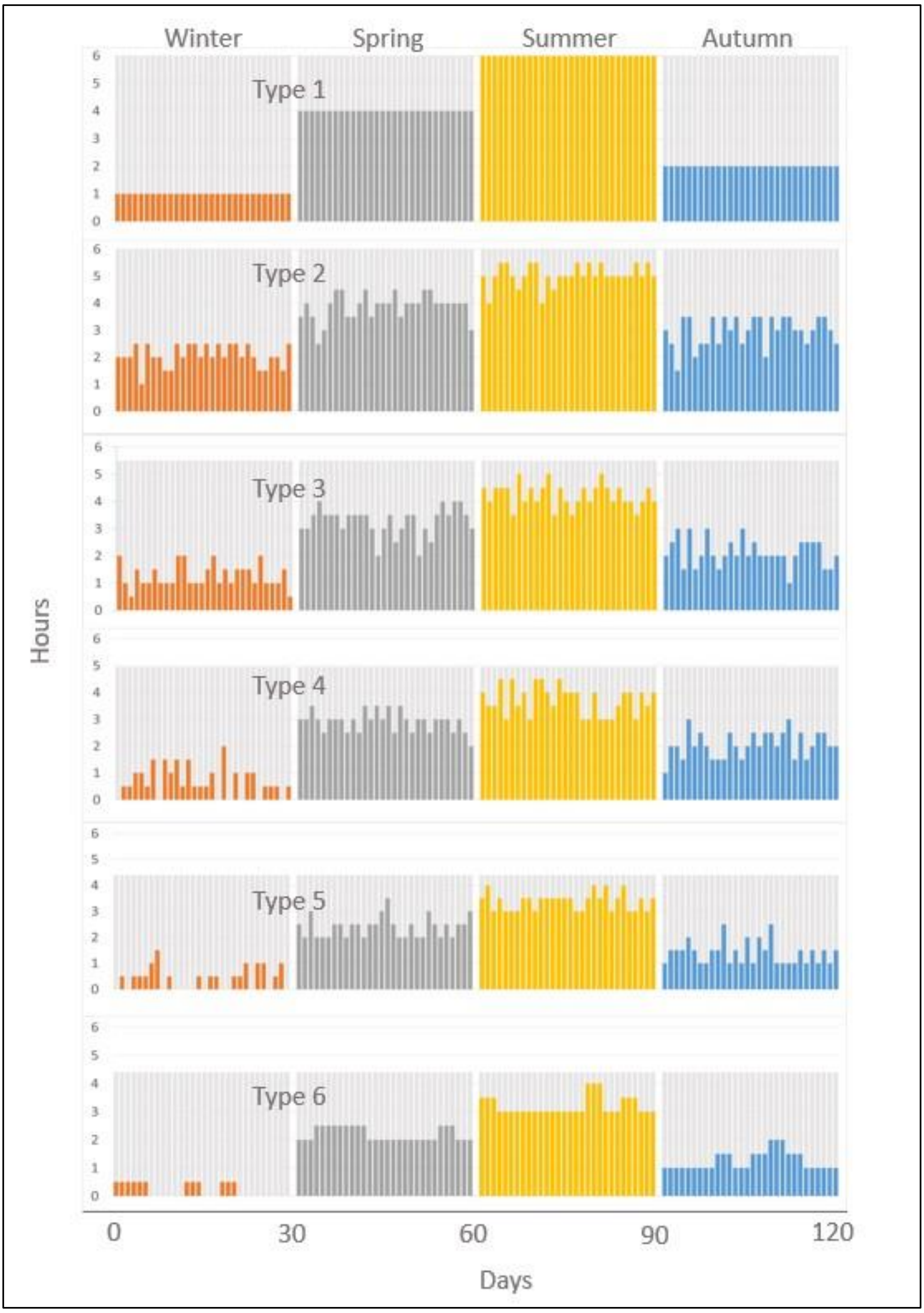

Figure 4-13. Window opening duration of different OB profiles

As can be observed, the graphs progressively demonstrate shorter and shorter window-opening timings, which corresponds to the tapering structure developed when the agent kinds were created. The results demonstrate the ABM's capacity to add personality-specific traits for a set of occupants. This operates within a user-specified range but allows the model to be flexible within that range. For example, because type 3 had a 60% chance of obeying the guidelines, their deviation from the baseline model would be between 0 and 40%. The subsequent versions, which opened the windows for progressively shorter periods of time, yielded similar findings.

The simulation results for types 7 and 8 were particularly noteworthy. These two categories were created to express occupants' Subjective/Injunctive Norms beliefs/concerns by either acting in the same way as their peers or "officemates" or acting in the opposite way. These agents had a greater proclivity to follow routines influenced by those around them. However, it was difficult to plot these since they appeared random without the context of which other type of agent was present in the workplace with Type 7 and Type 8. Some images in the Appendix of Paper 4 exhibit screenshots from the simulations, which demonstrate the office layout and the agents impersonating their coworkers.

4.4.3 Movement simulation results

The movement simulation was carried out based on the logic presented in the previous chapter in Figure 3-13. This simulation was run to mimic a standard day according to ASHRAE profiles, wherein the occupant would enter the office at a designated time, move to their assigned desks/workstation, and leave at the end of the day. As the agents were generated, each agent had a link established to their respective workstation The *heading* of the occupant, that denotes the agent’s direction, would be set based on the angle between the occupant and its destination. A simulation in progress is shown in figure 4-14. An occupant would be denoted as ‘moving’ if the distance between the agent and its destination was larger than one patch (red agents), whereas, when the distance from their target was less than a patch, i.e. the occupant was as closest as it can get to the target, it would be denoted as ‘working’ (blue agents).

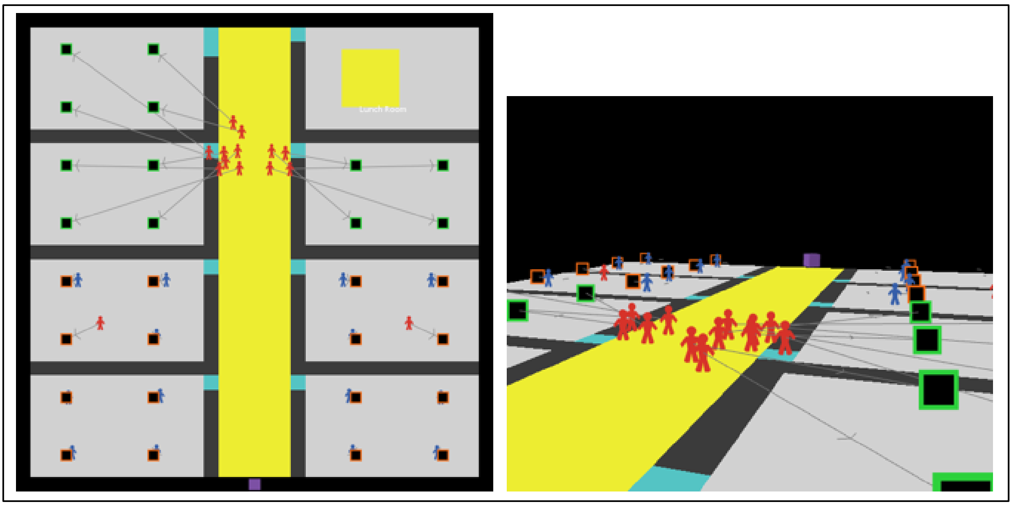

Figure 4-14. Occupants moving in the direction of their linked desk (2-D and 3-D views)

4.4.4 Additional deep learning models

Other OB models contributed from the work of this thesis include support work in developing deep learning models, some of which were mentioned in section 4.1.3, as part of supplementary work. In addition, the study presented in paper 7 revolved around the use of state-of-the-art DL models to predict occupancy and plug loads. It used the dataset from the case study on shared offices, with the objective of developing LSTM, Bi-LSTM and GRU models to assess which ones predict occupant's plug load patterns more accurately. This dataset was dividing into training, test, and validation sets, where the training set was used to develop the models, and predictions were made on the test set. LSTM was selected as the baseline model, and the predictions were carried out for a 'day-ahead' and 'week-ahead' period.

The result was three predictive models developed in the Python language, which were evaluated using the RMSE and Mean Absolute Error (MAE). The methodologies and optimization strategies are better outlined in the attached paper. The plug loads being predicted were for six devices (laptop, two monitors, docking station, desk lamp, and miscellaneous). The miscellaneous plug-loads proved to be the real challenge, since it had unstable consumption patterns. However, the Bi-LSTM model was the best one to forecast these with the least amount of error. All the three models performed well on the day-ahead predictions, capturing the peaks and lows of all devices with minimal deviations. For the week-ahead predictions, it was again the Bi-LSTM model that performed the best. These predictive models are useful for implementing demand-response assessments, as well as providing insights for pricing and tariffing energy-usage.

4.4.5 Answer to question 3

The results obtained from these simulations can be used to formulate an answer to research question 3 as follows:

Question: How can the social aspects of OB be modelled/simulated?

Answer: It can be worthwhile to note that this was the original question considered while formulating the goal of this doctoral study. It is a fairly straightforward question at first glance and addresses the core objective of aiding building energy performance simulations and optimizing energy efficiency of the same. However, as the work progressed, it was evident that there were several questions that lay in the background that had to be addressed first. This question was originally about modelling OB in general, and got modified to center around social aspects of OB. It also proved to be the source of all the other questions considered in the thesis.

The key to answering the question lay in Question 2, about assessing and evaluating social aspects in order to prepare them for simulation purposes. All in all, social aspects in OB can be modelled using diverse modelling strategies, but first require a framework that is suited to the method of modelling selected. Agent-based models were the ones selected in for this task, and were assessed to be a suitable conduit for capturing and channeling the complexities in OB. It had the capability of implementing the already established fixed and static models, while also accommodating the diversity of OB. Furthermore, the simulations could be adjusted using a simple user interface and incorporate flexibility as desired by the user

RESULTS AND DISCUSSION

5 CONCLUSIONS

This chapter summarizes the key takeaways from the studies conducted in this thesis. Section 5.1 provides the main conclusions of the three main parts, or themes that the work in this thesis can be divided into. Section 5.2 talks about the various limitations present in these studies. Section 5.3 presents some considerations for future work.

5.1 Concluding Remarks

Regarding monitoring of occupant behavior

Occupant behavior has a vital influence on the energy performance of buildings, and it is therefore essential to understand and document it to improve building design and operation. The first focus of this thesis, hence, was to build strategies to monitor OB and their influence in diverse settings using an assortment of sensing modalities.

The first case study of OB in shared office spaces yielded interesting insights. The use of a non-intrusive monitoring framework made it eligible for application in different buildings, especially those lacking BMS infrastructure. The insights about energy-related OB after monitoring a shared office for a year showed large amounts of energy being consumed when the occupants were not present in the area. In addition, device-level and occupant-level granularity made it possible to point out the devices that were causing the unnecessary consumption. Lastly, discrepancies were observed between the standard ASHRAE profiles and measured values of occupancy, lighting, and energy-use. These also provided insights for the kind of patterns that occur in academic buildings.

The second study was undertaken inside an operating room of a hospital. The aim of this research was to tackle the issue of limited investigation into human activity in hospital rooms during experiments. The findings indicate a correlation between higher activity levels and increased bacterial contamination in various areas. However, the study also revealed that areas near physical obstructions had the highest CFU densities, indicating that airflow patterns may have a role to play in such environments. Insights gained from studies like this could help to implement infection control measures, such as optimizing airflow and positioning of surgical instruments based on staff activity, as the current indoor environment design does not consider the effects of human activity during real surgical procedures. By dynamically recording human activity and using reproducible techniques to measure airborne contamination and airflow patterns, valuable information could be obtained that could influence the design of operating theatres and working practices.

It is crucial to acknowledge that the outcomes obtained in these studies are not intended to be broadly applicable or extended to conventional buildings. Rather, they demonstrate the effectiveness of the proposed high-resolution occupant monitoring systems in identifying possibly wasteful energy consumption patterns in office spaces, and potential sources of contaminations due to occupant activities in ORs. The results reveal the significant influence of individuals on their environment and can be beneficial including facility managers who maintain and adjust centralized systems.

The case study in residential spaces successfully demonstrated the precision of data fusion in capturing human activities by using multiple depth registration devices. An automated, real-time, and reliable activity recognition framework was able to monitor and assess occupant's activities of daily living in an unrestricted environment. The data collected from three cameras were fused into a single skeletal model and calibrated to synchronize spatial and temporal dimensions. In addition, the study showed the reliability of depth registration in accurately capturing activities, with 22 discrete atomic activities labeled and recognized. The resulting dataset is vital in establishing benchmark datasets for occupant activities.

Regarding OB profiles and preparatory work for OB models

While the first theme of the thesis focused on using sophisticated data acquisition techniques to collect detailed datasets, the second theme revolved around the preparatory work needed for OB models. This was a direct question generated while addressing the complexities of creating holistic models. Since OB has several aspects to it, the chosen ones to address here were the social aspects. In addition, supplementary work was done for generation of surface layouts for an occupant's environment. Both of these also tied into the proposed BOT-ABM model, covering two of the aspects mentioned in the model hypothesis. Lastly, a database to be used for OB models addressing multiple aspects was established.

The support work done for the simulation of the occupant's environmental layout and the database development is more of a methodological approach, making it difficult to draw conclusions on. Since there exist no such application that collect the data or contains variables that were defined, there cannot be a comparative analysis on it either. The main conclusions hence comprise of the methodology. To map occupant environmental layouts, a prerequisite library containing sufficient information for a surface layout simulator was identified. The study had three objectives: first, the identification of seven variables for the library; second, the development of a Matlab application for collecting spatial information about the environment, with the interface and sample output described above; and third, testing the ideas by collecting a sample dataset of 80 offices at NTNU. Additionally, a database was developed to connect the information generated from this study to larger models, which was then expanded to include other datasets from previous studies.

Meanwhile, the preparatory work for the social aspect of OB was much more complete, conducted from start to finish, including the conceptual theories, proposed extended-TPB model framework, data collection, analysis and evaluation. The necessity of using behavioral theories was well highlighted during the survey formulation, since it provided ground for quantification of these behaviors. These theories also assisted in the evaluation of the variables obtained from the data collection, in terms of their reliability and accuracy. The theoretical extended-TBP model finally concluded in one that included *habits* and *injunctive norms* as the best fit, although the models including *awareness of consequences* were not far behind. The addition of moderating variables did not prove to significantly change the outcome of these models. The data analysis to determine the influence of the different variables consisted of using path analysis, regression, and structural equation modelling. The latter of these did not have good capabilities to address the individual building systems, and the models created were for a general behavior pattern. This was not consistent with the objective of the study, hence the focus turned to the path analysis and regression models. Although SEM and path analysis are powerful tools to analyze complex relationships, regression analysis allows for the examination of relationships between a small number of variables, which can be useful when the goal is to identify the most important predictors of a particular outcome.

Habits consistently proved to be the best predictor for behavioral patterns regarding all building systems, while *personal norms*, *injunctive norms*, and *descriptive norms* competed for the second most influential predictors for different systems. Furthermore, the results of these models aided in the construction of OB profiles to be used for models. The optimal number of profiles was calculated using the elbow score in k-modes clustering methods, which turned out to be 8 profiles.

Regarding OB models

While the bulk of this thesis focused on preparatory work for OB models, the end goal was to develop accurate and more representative models, that could be used for building performance simulations. Two of the datasets became available as benchmark datasets and were used for developing machine learning and deep learning models. One of them even became part of the global OB database that has been made available by the IEA's Annex-79. The results from the deep learning models summarized the performance of state-of-the-art DL models with regards to occupancy, plug-loads, and activity recognition. Bi-LSTM models predicted the energy-use patterns most accurately.

The agent-based model was developed to simulate the social aspects of OB, as part of a proof-of-concept study to showcase the abilities of ABMs in capturing the diversity and complexities of this phenomenon. Moreover, an important contribution of this study was to provide an interdisciplinary connection for social sciences and the engineering part of the field. Studies concerning OB often highlight the necessity of incorporating multiple branches of sciences in order to provide an accurate and representative assessment of occupant's energy-use behaviors in buildings. This recommendation extends to it modelling and simulation as well. The agent-based model created in this study was grounded in and based on the social behavioral theories, and then used the assessments and profiles provided by the preparatory step to simulate window-opening behavior. It also provided a simple and accessible user interface and flexibility for simulations and is scheduled to be released for access as a Netlogo model. This can serve as a starting point for building larger and more detailed models and optimize the research methodology on OB.

5.2 Limitations

Occupant behavior is an evolving topic, where the lack of global standards regarding OB modelling and simulation leads to the need for several assumptions. There are also restrictions due to the scale and complexities involved in the topic. This section briefly discusses the key limitations regarding the studies, with more detailed discussion provided in the individual papers.

Regarding monitoring of occupant behavior

The main limitations with regards to OB monitoring studies in this thesis can be summarized in three key points: scale, homogeneity, and lack of additional considerations.

1. Scale: The sample size of the participants in the studies was sufficient enough to test the framework, assess the occupants’ influences etc. However, it was only sufficient to point out the eligibility of either the framework or the methodology. For example, 8 occupants were monitored in the shared office space, and while this was to show the efficiency of a granular monitoring framework, it could benefit vastly from an increased sample size. Similarly, the study conducted in the residential Living Lab could benefit from a larger number of occupants, as well as scalability. All three studies were very specific for their environments and need to incorporate scalability in their approaches.
2. Homogeneity: The homogeneity referred to here is with respect to the building types, building systems or monitoring systems involved in the studies. For instance, passive sampling strategies were used to collect data about the bacterial accumulation in the operating room, which has several limitations in comparison with active sampling strategies. Moreover, the ventilation systems present in the study come with their own set of limitations. Studies need to be conducted in rooms with different ventilation systems to see if the occupant activity levels are impactful enough in other environments. Similarly, the room layout and placement of furniture is always kept constant in order to have a controlled study. However, this can reduce the applicability of the monitoring methodologies, since different environments may contribute to different results. Lastly, additional sensors such as wearable sensors, HVAC monitoring etc would have improved the application of these monitoring systems as well.
3. Lack of additional considerations: Some of the considerations missing in the studies are those of HVAC systems. This is a significant part of energy-use in buildings and could not be considered due to the requirements of the study, either being non-intrusive or due to lack of valid data. Other additional considerations include occupant bias in reporting and the need for strong computing systems for the activity recognition.

Regarding OB profiles and preparatory work for OB models

Both the library for the surface layout simulator and the database developed for connecting different datasets across a platform are in the initial stages. There is a lack of testing and validation involved, due to these being developed for hypothetically larger models. However, this scope was taken into account since works like these would be parts of a larger project. That makes it difficult to run them in a real-world scenario. In addition, there is a computational expense involved that makes it expensive and laborious to implement, thereby requiring a lot more optimization.

For the work regarding the social aspect of OB, the sample size is again one of the major drawbacks. Since this was more of a proof-of-concept study, the aim was to build a functional model and provide an interdisciplinary link between two different sciences concerning OB. For social aspect, larger sample size, different types of buildings and rooms, more investigation around behavior regarding appliances, computers etc. needs to be addressed.

Regarding OB models

The most significant drawback of the agent-based model developed in the study is the lack of validation. While the model offers flexibility and incorporates the diversity in OB, there is no way to gauge how accurately that is done. This problem persists in OB models in general, since behavior is hard to predict, and even harder to validate. Protocols for validation were adopted in the study but fell out of scope. The future work considerations elaborate on the proposed validation strategies. The model is also extremely limited to a specific office layout and cannot be easily transferred to other settings. The agents in the model, representing the occupants, already have ties to the layout, setting, and desks/windows. This limits the transferability of the model. While other layouts can be made keeping the occupants intact, the process is time-consuming and computationally demanding, at least to encompass a bigger range of layouts and environmental settings.

5.3 Considerations for future work

Regarding monitoring of occupant behavior

All the considerations for future work of OB monitoring stem largely from the limitations observed. First and foremost is the scalability. All the case studies mentioned in the thesis can benefit from a larger sample size. More occupants would make it viable to draw standard conclusions regarding OB, as well as test out the applicability of the monitoring frameworks proposed in the individual studies. The diversity of occupants would make it easier to train models on a large scale of data with varying energy-use and activity patterns. Secondly, the monitoring systems can be expanded to be used in different building types, thereby increasing the scope of the studies. Different monitoring strategies can be compared, which take into account the sensing modalities used here and compare it with other sensors. A significant hindrance on setting up such monitoring systems is the cost involved. Economic assessments need to be made about these systems with regards to the capital needed and the improvement in energy efficiency offered. Incorporation of HVAC systems is another goal for the future considerations, which would introduce more complexities and would need more capable monitoring systems to accurately record the use-patterns.

These considerations would provide building facility managers with adequate data about how building energy performance is being affected by occupant behavior, and what optimization strategies could be adopted to improve it. This also helps policymakers provide better guidelines and make decisions in a more informed manner.

Regarding OB profiles and preparatory work for OB models

The developments around these have been all about initial steps. In order to take it forward, the application developed for the surface layout needs to be used in different environments to evaluate its applicability. Future work can also start with testing out the capabilities of the database, by connecting it to multiple models, or simply releasing it for wider use for entering benchmark datasets that are accessible to a larger community.

While a major part of this thesis focused on the social aspects of OB, starting from conceptual frameworks to a working model, similar care and effort has to be given to the other aspects as well. For example, some aspects identified in the BOT-ABM proposal have been researched on and analyzed in a similar manner, whereas others need to be assessed. This has to contribute in building a holistic OB model that can standardize datasets from across disciplines, use profiles for multiple aspects of OB, and have dedicated models for each.

Regarding OB models

In order to proceed with OB models, the question of validation has to be tackled. The selection of agent-based models was also done due to its capacity to program individual agents and monitor the outcome on a larger scale. A validation protocol can hence be established by programming the agents based on the profiles and subjecting them to a scenario where we already know what the outcome is supposed to look like. This protocol was initiated in the study during the survey creation. Two scenarios were presented to the participants, one regarding their interactions with lighting systems and the other regarding appliances. Since all participants are categorized into profiles, agents can be modelled according to the general rules of the profiles. A single agent from a profile can then be placed in a scenario that is the same as presented to the participants during the survey. The actions of these agents would give an idea to the accuracy of the behaviors being simulated. This would be the next step in terms of agent-based modelling.

The DL models provided forecasts and predictions around energy-use of the occupants. The benefit of developing these models would include the awareness of OB and their patterns of energy-use, which could be further utilized by building facility managers to develop energy-efficiency programs. These programs can aim at spreading awareness, identifying the requirements of the occupants through feedback systems, and incentivize reduction in energy consumption.

BIBLIOGRAPHY

[1] IEA, World Energy Outlook 2018, Paris, 2018.

[2] E.I.A. (US), Annual Energy Outlook 2012: With Projections to 2035, 2012.

[3] L. Pérez-Lombard, J. Ortiz, C. Pout, A review on buildings energy consumption information, Energy Build. 40 (2008) 394–398. https://doi.org/10.1016/j.enbuild.2007.03.007.

[4] UN, United Nations Sustainable Development Goals, (n.d.). https://www.un.org/sustainabledevelopment/ (accessed June 15, 2022).

[5] A.C. Menezes, A. Cripps, D. Bouchlaghem, R. Buswell, Predicted vs. actual energy performance of non-domestic buildings: Using post-occupancy evaluation data to reduce the performance gap, Appl Energy. 97 (2012) 355–364. https://doi.org/10.1016/j.apenergy.2011.11.075.

[6] D. Parker, E. Mills, L. Rainer, N. Bourassa, G. Homan, Accuracy of the Home Energy Saver Energy Calculation Methodology, 2012.

[7] A. Mahdavi, L. Lambeva, A. Mohammadi, E. Kabir, C. Pröglhöf, Two case studies on user interactions with buildings' environmental systems, Bauphysik. 29 (2007) 72–75. https://doi.org/10.1002/bapi.200710013.

[8] G. Hitchcock, An integrated framework for energy use and behaviour in the domestic sector, Energy Build. 20 (1993) 151–157. https://doi.org/10.1016/0378-7788(93)90006-G.

[9] D. Yan, T. Hong, B. Dong, A. Mahdavi, S. D'Oca, I. Gaetani, X. Feng, IEA EBC Annex 66: Definition and simulation of occupant behavior in buildings, Energy Build. 156 (2017) 258–270. https://doi.org/10.1016/j.enbuild.2017.09.084.

[10] C.M. Clevenger, J.R. Haymaker, M. Jalili, Demonstrating the Impact of the Occupant on Building Performance, Journal of Computing in Civil Engineering. 28 (2014) 99–102. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000323.

[11] T. Hong, Hung-Wen Lin, Occupant behavior: impact on energy use of private offices, 2013.

[12] I. Gaetani, P.J. Hoes, J.L.M. Hensen, Estimating the influence of occupant behavior on building heating and cooling energy in one simulation run, Appl Energy. 223 (2018) 159–171. https://doi.org/10.1016/J.APENERGY.2018.03.108.

[13] W. O'Brien, I. Gaetani, S. Gilani, S. Carlucci, P.J. Hoes, J. Hensen, International survey on current occupant modelling approaches in building performance simulation†, Https://Doi.Org/10.1080/19401493.2016.1243731. 10 (2016) 653–671. https://doi.org/10.1080/19401493.2016.1243731.

[14] J.L.M. Hensen, R. Lamberts, R. Lamberts, Building Performance Simulation for Design and Operation, Routledge, 2019. https://doi.org/10.1201/9780429402296.

[15] A. Mahdavi, The human dimension of building performance simulation, in: F Building Simulation 2011: 12th Conference of International Building Performance Simulation Association, Sydney, 2011.

[16] P. Hoes, J.L.M. Hensen, M.G.L.C. Loomans, B. de Vries, D. Bourgeois, User behavior in whole building simulation, Energy Build. 41 (2009) 295–302. https://doi.org/10.1016/j.enbuild.2008.09.008.

[17] D. Yan, T. Hong, Definition and Simulation of Occupant Behavior in Buildings Annex 66 Final Report Operating Agents of Annex 66, 2018. www.iea-ebc.org (accessed February 20, 2020).

[18] B.K. Sovacool, S.E. Ryan, P.C. Stern, K. Janda, G. Rochlin, D. Spreng, M.J. Pasqualetti, H. Wilhite, L. Lutzenhiser, Integrating social science in energy research, Energy Res Soc Sci. 6 (2015) 95–99. https://doi.org/10.1016/j.erss.2014.12.005.

[19] J.W. Dziedzic, A novel monitoring and modelling technique for energy-related occupant behaviour, 2021.

[20] N.E. Klepeis, W.C. Nelson, W.R. Ott, J.P. Robinson, A.M. Tsang, P. Switzer, J. V. Behar, S.C. Hern, W.H. Engelmann, The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants, Journal of Exposure Science & Environmental Epidemiology 2001 11:3. 11 (2001) 231–252. https://doi.org/10.1038/sj.jea.7500165.

[21] D.R.G. Hunt, The use of artificial lighting in relation to daylight levels and occupancy, Build Environ. 14 (1979) 21–33. https://doi.org/10.1016/0360-1323(79)90025-8.

[22] C.F. Reinhart, Lightswitch-2002: A model for manual and automated control of electric lighting and blinds, Solar Energy. 77 (2004) 15–28. https://doi.org/10.1016/j.solener.2004.04.003.

[23] S. Pigg, M. Eilers, J. Reed, Behavioral Aspects of Lighting and Occupancy Sensors in Private Offices: A Case Study of a University Office Building, ACEEE 1996 Summer Study on Energy Efficiency in Buildings. (1996) 161–170.

[24] A. Rubin, B. Collins, R. Tibbott, Window Blinds as a Potential Energy Saver-A Case Study, NATIONAL BUREAU OF STANDARDS, Washington D.C, 1978.

[25] M.S. Rea, Window Blind Occlusion: a Pilot Study, 1984.

[26] T. Inoue, T. Kawase, T. Ibamoto, S. Takakusa, Y. Matsuo, The development of an optimal control system for window shading devices based on investigations in office buildings, ASHRAE Trans. 94 (1988) 1034–1049.

[27] E. Delzendeh, S. Wu, A. Lee, Y. Zhou, The impact of occupants' behaviours on building energy analysis: A research review, Renewable and Sustainable Energy Reviews. 80 (2017) 1061–1071. https://doi.org/10.1016/J.RSER.2017.05.264.

[28] J. Schot, L. Kanger, G. Verbong, The roles of users in shaping transitions to new energy systems, Nature Energy 2016 1:5. 1 (2016) 1–7. https://doi.org/10.1038/nenergy.2016.54.

[29] M.A.R. Lopes, C.H. Antunes, N. Martins, Energy behaviours as promoters of energy efficiency: A 21st century review, Renewable and Sustainable Energy Reviews. 16 (2012) 4095–4104. https://doi.org/10.1016/J.RSER.2012.03.034.

[30] O.T. Masoso, L.J. Grobler, The dark side of occupants' behaviour on building energy use, Energy Build. 42 (2010) 173–177. https://doi.org/10.1016/j.enbuild.2009.08.009.

[31] N. Zhou, N. Khanna, W. Feng, H. Lixuan, D. Fridley, J. Creyts, E. Franconi, R. Torbert, Y. Ke, Cost-effective options for transforming the Chinese building sector, ACEEE Summer Study on Energy Efficiency in Buildings. (n.d.).

[32] D. Calì, T. Osterhage, R. Streblow, D. Müller, Energy performance gap in refurbished German dwellings: Lesson learned from a field test, Energy Build. 127 (2016) 1146–1158. https://doi.org/10.1016/J.ENBUILD.2016.05.020.

[33] R.V. Andersen, J. Toftum, K.K. Andersen, B.W. Olesen, Survey of occupant behaviour and control of indoor environment in Danish dwellings, Energy Build. 41 (2009) 11–16. https://doi.org/10.1016/J.ENBUILD.2008.07.004.

[34] E. Recast, Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the energy performance of buildings, Official Journal of the European Union. 18 (2010).

[35] P. De Wilde, The gap between predicted and measured energy performance of buildings: A framework for investigation, Autom Constr. 41 (2014) 40–49. https://doi.org/10.1016/J.AUTCON.2014.02.009.

[36] C. Carpino, D. Mora, N. Arcuri, M. De Simone, Behavioral variables and occupancy patterns in the design and modeling of Nearly Zero Energy Buildings, Building Simulation 2017 10:6. 10 (2017) 875–888. https://doi.org/10.1007/S12273-017-0371-2.

[37] D. Majcen, L.C.M. Itard, H. Visscher, Theoretical vs. actual energy consumption of labelled dwellings in the Netherlands: Discrepancies and policy implications, Energy Policy. 54 (2013) 125–136. https://doi.org/10.1016/J.ENPOL.2012.11.008.

[38] H. Yoshino, T. Hong, N. Nord, IEA EBC annex 53: Total energy use in buildings—Analysis and evaluation methods, Energy Build. 152 (2017) 124–136. https://doi.org/10.1016/J.ENBUILD.2017.07.038.

[39] D. Yan, T. Hong, B. Dong, A. Mahdavi, S. D'Oca, I. Gaetani, X. Feng, IEA EBC Annex 66: Definition and simulation of occupant behavior in buildings, Energy Build. 156 (2017) 258–270. https://doi.org/10.1016/J.ENBUILD.2017.09.084.

[40] W. O'Brien, A. Wagner, M. Schweiker, A. Mahdavi, J. Day, M.B. Kjærgaard, S. Carlucci, B. Dong, F. Tahmasebi, D. Yan, T. Hong, H.B. Gunay, Z. Nagy, C. Miller, C. Berger, Introducing IEA EBC annex 79: Key challenges and opportunities in the field of occupant-centric building design and operation, Build Environ. 178 (2020) 106738. https://doi.org/10.1016/J.BUILDENV.2020.106738.

[41] C. Peng, D. Yan, R. Wu, C. Wang, X. Zhou, Y. Jiang, Quantitative description and simulation of human behavior in residential buildings, Building Simulation 2011 5:2. 5 (2011) 85–94. https://doi.org/10.1007/S12273-011-0049-0.

[42] A. Mahdavi, A. Mohammadi, E. Kabir, L. Lambeva, Occupants' operation of lighting and shading systems in office buildings, J Build Perform Simul. 1 (2008) 57–65. https://doi.org/10.1080/19401490801906502.

[43] X. Feng, D. Yan, C. Wang, H. Sun, A preliminary research on the derivation of typical occupant behavior based on large-scale questionnaire surveys, Energy Build. 117 (2016) 332–340. https://doi.org/10.1016/J.ENBUILD.2015.09.055.

[44] Y. Zhang, X. Bai, F.P. Mills, J.C.V. Pezzey, Rethinking the role of occupant behavior in building energy performance: A review, Energy Build. 172 (2018) 279–294. https://doi.org/10.1016/J.ENBUILD.2018.05.017.

[45] S. Carlucci, M. De Simone, S.K. Firth, M.B. Kjærgaard, R. Markovic, M.S. Rahaman, M.K. Annaqeeb, S. Biandrate, A. Das, J.W. Dziedzic, G. Fajilla, M. Favero, M. Ferrando, J. Hahn, M. Han, Y. Peng, F. Salim, A. Schlüter, C. van Treeck, Modeling occupant behavior in buildings, Build Environ. 174 (2020) 106768. https://doi.org/10.1016/J.BUILDENV.2020.106768.

[46] B. Dong, M.B. Kjærgaard, M. De Simone, H.B. Gunay, W. O'Brien, D. Mora, J. Dziedzic, J. Zhao, Sensing and data acquisition, in: Exploring Occupant Behavior in Buildings: Methods and Challenges, Springer International Publishing, 2017: pp. 77–105. https://doi.org/10.1007/978-3-319-61464-9_4.

[47] G.W. Hart, Nonintrusive Appliance Load Monitoring, Proceedings of the IEEE. 80 (1992) 1870–1891. https://doi.org/10.1109/5.192069.

[48] H.N. Rafsanjani, C. Ahn, Linking Building Energy-Load Variations with Occupants' Energy-Use Behaviors in Commercial Buildings: Non-Intrusive Occupant Load Monitoring (NIOLM), in: Procedia Eng, Elsevier Ltd, 2016: pp. 532–539. https://doi.org/10.1016/j.proeng.2016.04.041.

[49] J. Chen, C. Ahn, Assessing occupants' energy load variation through existing wireless network infrastructure in commercial and educational buildings, Energy Build. 82 (2014) 540–549. https://doi.org/10.1016/j.enbuild.2014.07.053.

[50] O. Ardakanian, A. Bhattacharya, D. Culler, Non-intrusive occupancy monitoring for energy conservation in commercial buildings, Energy Build. 179 (2018) 311–323. https://doi.org/10.1016/j.enbuild.2018.09.033.

[51] M.K. Annaqeeb, R. Markovic, V. Novakovic, E. Azar, Non-Intrusive Data Monitoring and Analysis of Occupant Energy-Use Behaviors in Shared Office Spaces, IEEE Access. 8 (2020) 141246–141257. https://doi.org/10.1109/ACCESS.2020.3012905.

[52] S. Scaltriti, S. Cencetti, S. Rovesti, I. Marchesi, A. Bargellini, P. Borella, Risk factors for particulate and microbial contamination of air in operating theatres, Journal of Hospital Infection. 66 (2007) 320–326. https://doi.org/10.1016/j.jhin.2007.05.019.

[53] A.E. Andersson, I. Bergh, J. Karlsson, B.I. Eriksson, K. Nilsson, Traffic flow in the operating room: An explorative and descriptive study on air quality during orthopedic trauma implant surgery, Am J Infect Control. 40 (2012) 750–755. https://doi.org/10.1016/j.ajic.2011.09.015.

[54] M.K. Annaqeeb, Y. Zhang, J.W. Dziedzic, K. Xue, C. Pedersen, L.I. Stenstad, V. Novakovic, G. Cao, Influence of surgical team activity on airborne bacterial distribution in the operating room with a mixing ventilation system: a case study at St. Olavs Hospital, Journal of Hospital Infection. 116 (2021) 91–98. https://doi.org/10.1016/J.JHIN.2021.08.009.

[55] D. Moher, A. Liberati, J. Tetzlaff, D.G. Altman, Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement, Int J Surg. 8 (2010) 336–341. https://doi.org/10.1016/J.IJSU.2010.02.007.

[56] J. Zhao, B. Lasternas, K.P. Lam, R. Yun, V. Loftness, Occupant behavior and schedule modeling for building energy simulation through office appliance power consumption data mining, Energy Build. 82 (2014) 341–355. https://doi.org/10.1016/j.enbuild.2014.07.033.

[57] D. Bourgeois, C. Reinhart, I. Macdonald, Adding advanced behavioural models in whole building energy simulation: A study on the total energy impact of manual and automated lighting control, Energy Build. 38 (2006) 814–823. https://doi.org/10.1016/j.enbuild.2006.03.002.

[58] Y.S. Lee, A.M. Malkawi, Simulating multiple occupant behaviors in buildings: An agent-based modeling approach, Energy Build. 69 (2014) 407–416. https://doi.org/10.1016/j.enbuild.2013.11.020.

[59] D. Aerts, J. Minnen, I. Glorieux, I. Wouters, F. Descamps, A method for the identification and modelling of realistic domestic occupancy sequences for building energy demand simulations and peer comparison, Build Environ. 75 (2014) 67–78. https://doi.org/10.1016/j.buildenv.2014.01.021.

[60] Amareican Society of Heating Refrigeration and Air Conditioning Engineers, ANSI/ASHRAE/IES Standard 90.1-2016: Energy Standard for Buildings Except Low-rise Residential Buildings, 2016. https://scholar.google.com/scholar?q=ANSIASHRAEIES%20Standard%2090.1-2016:%20Energy%20Standard%20for%20Buildings%20except%20Low-Rise%20Residential%20Buildings (accessed May 29, 2023).

[61] J. Dziedzic, D. Yan, V. Novakovic, Occupant migration monitoring in residential buildings with the use of a depth registration camera, in: Procedia Eng, Elsevier Ltd, 2017: pp. 1193–1200. https://doi.org/10.1016/j.proeng.2017.10.352.

[62] C. Wang, D. Yan, Y. Jiang, A novel approach for building occupancy simulation, Build Simul. 4 (2011) 149–167. https://doi.org/10.1007/s12273-011-0044-5.

[63] C. Martani, D. Lee, P. Robinson, R. Britter, C. Ratti, ENERNET: Studying the dynamic relationship between building occupancy and energy consumption, Energy Build. 47 (2012) 584–591. https://doi.org/10.1016/j.enbuild.2011.12.037.

[64] J.W. Dziedzic, Y. Da, V. Novakovic, Indoor occupant behaviour monitoring with the use of a depth registration camera, Build Environ. 148 (2019) 44–54. https://doi.org/10.1016/j.buildenv.2018.10.032.

[65] T. Hong, D. Yan, S. D'Oca, C. fei Chen, Ten questions concerning occupant behavior in buildings: The big picture, Build Environ. 114 (2017) 518–530. https://doi.org/10.1016/j.buildenv.2016.12.006.

[66] K. Steemers, S. Manchanda, Energy efficient design and occupant well-being: Case studies in the UK and India, Build Environ. 45 (2010) 270–278. https://doi.org/10.1016/J.BUILDENV.2009.08.025.

[67] R. Holopainen, P. Tuomaala, P. Hernandez, T. Häkkinen, K. Piira, J. Piippo, Comfort assessment in the context of sustainable buildings: Comparison of simplified and detailed human thermal sensation methods, Build Environ. 71 (2014) 60–70. https://doi.org/10.1016/J.BUILDENV.2013.09.009.

[68] B.K. Sovacool, S.E. Ryan, P.C. Stern, K. Janda, G. Rochlin, D. Spreng, M.J. Pasqualetti, H. Wilhite, L. Lutzenhiser, Integrating social science in energy research, Energy Res Soc Sci. 6 (2015) 95–99. https://doi.org/10.1016/j.erss.2014.12.005.

[69] N.W. Biggart, L. Lutzenhiser, Economic Sociology and the Social Problem of Energy Inefficiency:, Http://Dx.Doi.Org/10.1177/0002764207299355. 50 (2016) 1070–1087. https://doi.org/10.1177/0002764207299355.

[70] W. Abrahamse, L. Steg, Factors Related to Household Energy Use and Intention to Reduce It: The Role of Psychological and Socio-Demographic Variables, Human Ecology Review. 18 (2011) 30–40.

[71] C.F. Chen, K. Knight, Energy at work: Social psychological factors affecting energy conservation intentions within Chinese electric power companies, Energy Res Soc Sci. 4 (2014) 23–31. https://doi.org/10.1016/j.erss.2014.08.004.

[72] J. Chen, J.E. Taylor, H.H. Wei, Modeling building occupant network energy consumption decision-making: The interplay between network structure and conservation, Energy Build. 47 (2012) 515–524. https://doi.org/10.1016/J.ENBUILD.2011.12.026.

[73] L. Gustafsson, M. Sternad, Consistent micro, macro and state-based population modelling, Math Biosci. 225 (2010) 94–107. https://doi.org/10.1016/j.mbs.2010.02.003.

[74] V.L. Erickson, Y. Lin, A. Kamthe, R. Brahme, A. Surana, A.E. Cerpa, M.D. Sohn, S. Narayanan, Energy efficient building environment control strategies using real-time occupancy measurements, in: BUILDSYS 2009 - Proceedings of the 1st ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings, Held in Conjunction with ACM SenSys 2009, ACM Press, New York, New York, USA, 2009: pp. 19–24. https://doi.org/10.1145/1810279.1810284.

[75] Z. Li, Y. Heo, G. Augenbroe, HVAC design informed by organizational simulation Stochastic optimized chiller operation strategy based on multi-objective optimization View project HVAC DESIGN INFORMED BY ORGANIZATIONAL SIMULATION, 2009.

[76] E. Azar, C.C. Menassa, Agent-Based Modeling of Occupants and Their Impact on Energy Use in Commercial Buildings, Journal of Computing in Civil Engineering. 26 (2012) 506–518. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000158.

[77] G.W. Hart, Nonintrusive Appliance Load Monitoring, Proceedings of the IEEE. 80 (1992) 1870–1891. https://doi.org/10.1109/5.192069.

[78] R. Bonfigli, S. Squartini, M. Fagiani, F. Piazza, Unsupervised algorithms for non-intrusive load monitoring: An up-to-date overview, in: 2015 IEEE 15th International Conference on Environment and Electrical Engineering, EEEIC 2015 - Conference Proceedings, Institute of Electrical and Electronics Engineers Inc., 2015: pp. 1175–1180. https://doi.org/10.1109/EEEIC.2015.7165334.

[79] H. Rafsanjani, C. Ahn, M. Alahmad, A Review of Approaches for Sensing, Understanding, and Improving Occupancy-Related Energy-Use Behaviors in Commercial Buildings, Energies (Basel). 8 (2015) 10996–11029. https://doi.org/10.3390/en81010996.

[80] J. Chen, C. Ahn, Assessing occupants' energy load variation through existing wireless network infrastructure in commercial and educational buildings, Energy Build. 82 (2014) 540–549. https://doi.org/10.1016/J.ENBUILD.2014.07.053.

[81] O. Ardakanian, A. Bhattacharya, D. Culler, Non-intrusive occupancy monitoring for energy conservation in commercial buildings, Energy Build. 179 (2018) 311–323. https://doi.org/10.1016/J.ENBUILD.2018.09.033.

[82] T. Zhang, O. Ardakanian, A domain adaptation technique for fine-grained occupancy estimation in commercial buildings, in: IoTDI 2019 - Proceedings of the 2019 Internet of Things Design and Implementation, Association for Computing Machinery, Inc, New York, NY, USA, 2019: pp. 148–159. https://doi.org/10.1145/3302505.3310077.

[83] Z. Wang, T. Hong, M.A. Piette, Data fusion in predicting internal heat gains for office buildings through a deep learning approach, Appl Energy. 240 (2019) 386–398. https://doi.org/10.1016/j.apenergy.2019.02.066.

[84] C. Lobato, S. Pless, M. Sheppy, P. Torcellini, Reducing plug and process loads for a large scale, low energy office building: NREL's research support facility, No. NREL/CP-5500-49002. National Renewable Energy Lab.(NREL). (2011). https://www.osti.gov/biblio/1009280 (accessed July 7, 2020).

[85] W. Whyte, R. Hodgson, J. Tinkler, The importance of airborne bacterial contamination of wounds, Journal of Hospital Infection. 3 (1982) 123–135. https://doi.org/10.1016/0195-6701(82)90004-4.

[86] A. Tammelin, P. Domicel, A. Hambræus, E. Ståhle, Dispersal of methicillin-resistant Staphylococcus epidermidis by staff in an operating suite for thoracic and cardiovascular surgery: Relation to skin carriage and clothing, Journal of Hospital Infection. 44 (2000) 119–126. https://doi.org/10.1053/jhin.1999.0665.

[87] S. Bhangar, R.I. Adams, W. Pasut, J.A. Huffman, E.A. Arens, J.W. Taylor, T.D. Bruns, W.W. Nazaroff, Chamber bioaerosol study: human emissions of size-resolved fluorescent biological aerosol particles, Indoor Air. 26 (2016) 193–206. https://doi.org/10.1111/ina.12195.

[88] L. Sanzen, M. Walder, S. Carlsson, Air contamination during total hip arthroplasty in an ultraclean air enclosure using different types of staff clothing, Journal of Arthroplasty. 5 (1990) 127–130. https://doi.org/10.1016/S0883-5403(06)80231-7.

[89] R. You, W. Cui, C. Chen, B. Zhao, Measuring the short-term emission rates of particles in the "personal cloud" with different clothes and activity intensities in a sealed chamber, Aerosol Air Qual Res. 13 (2013) 911–921. https://doi.org/10.4209/aaqr.2012.03.0061.

[90] D. Marikyan, S. Papagiannidis, E. Alamanos, A systematic review of the smart home literature: A user perspective, Technol Forecast Soc Change. 138 (2019) 139–154. https://doi.org/10.1016/J.TECHFORE.2018.08.015.

[91] M. Li, W. Gu, W. Chen, Y. He, Y. Wu, Y. Zhang, Smart Home: Architecture, Technologies and Systems, Procedia Comput Sci. 131 (2018) 393–400. https://doi.org/10.1016/J.PROCS.2018.04.219.

[92] A. Murad, J.Y. Pyun, Deep Recurrent Neural Networks for Human Activity Recognition, Sensors 2017, Vol. 17, Page 2556. 17 (2017) 2556. https://doi.org/10.3390/S17112556.

[93] C. Wilson, T. Hargreaves, R. Hauxwell-Baldwin, Benefits and risks of smart home technologies, Energy Policy. 103 (2017) 72–83. https://doi.org/10.1016/J.ENPOL.2016.12.047.

[94] A. Das, F.C. Sangogboye, E.S.K. Raun, M.B. Kjærgaard, HeteroSense: An occupancy sensing framework for multi-class classification for activity recognition and trajectory detection, SocialSense 2019 - Proceedings of the 2019 4th International Workshop on Social Sensing. (2019) 12–17. https://doi.org/10.1145/3313294.3313383.

[95] J.W. Dziedzic, D. Yan, H. Sun, V. Novakovic, Building occupant transient agent-based model – Movement module, Appl Energy. 261 (2020) 114417. https://doi.org/10.1016/J.APENERGY.2019.114417.

[96] C.H.-B. Change, undefined 1988, Social foundations of thought and action: A social cognitive theory, Albert Bandura Englewood Cliffs, New Jersey: Prentice Hall, 1986, xiii+ 617 pp. Hardback. US, Cambridge.Org. (n.d.). https://www.cambridge.org/core/journals/behaviour-change/article/social-foundations-of-thought-and-action-a-social-cognitive-theory-banduraalbertenglewood-cliffs-new-jersey-prentice-hall-1986-xiii-617-pp-hardback-us3950/B601D38456EF1C78547064C6D8C07C2C (accessed December 4, 2022).

[97] T. Hong, S. D'Oca, W.J.N. Turner, S.C. Taylor-Lange, An ontology to represent energy-related occupant behavior in buildings. Part I: Introduction to the DNAs framework, Build Environ. 92 (2015) 764–777. https://doi.org/10.1016/j.buildenv.2015.02.019.

[98] I. Ajzen, The theory of planned behavior, Organ Behav Hum Decis Process. 50 (1991) 179–211. https://doi.org/10.1016/0749-5978(91)90020-T.

[99] F.G. Kaiser^, H. Gutscher, P. Harland, H.J. Staats, The Proposition of a General Version of the Theory of Planned Behavior: Predicting Ecological Behavior1, J Appl Soc Psychol. 33 (2003) 586–603. https://doi.org/10.1111/J.1559-1816.2003.TB01914.X.

[100] M. Greaves, L.D. Zibarras, C. Stride, Using the theory of planned behavior to explore environmental behavioral intentions in the workplace, J Environ Psychol. 34 (2013) 109–120. https://doi.org/10.1016/J.JENVP.2013.02.003.

[101] S. D'Oca, C.F. Chen, T. Hong, Z. Belafi, Synthesizing building physics with social psychology: An interdisciplinary framework for context and occupant behavior in office buildings, Energy Res Soc Sci. 34 (2017) 240–251. https://doi.org/10.1016/j.erss.2017.08.002.

[102] M.F. Chen, Extending the theory of planned behavior model to explain people's energy savings and carbon reduction behavioral intentions to mitigate climate change in Taiwan-moral obligation matters, J Clean Prod. 112 (2016). https://doi.org/10.1016/j.jclepro.2015.07.043.

[103] X. Xu, C.F. Chen, D. Li, C. Menassa, Energy Saving at Work: Exploring the Role of Social Norms, Perceived Control and Ascribed Responsibility in Different Office Layouts, Front Built Environ. 6 (2020). https://doi.org/10.3389/fbuil.2020.00016.

[104] A. Gkargkavouzi, G. Halkos, S. Matsiori, Environmental behavior in a private-sphere context: Integrating theories of planned behavior and value belief norm, self-identity and habit, Resour Conserv Recycl. 148 (2019) 145–156. https://doi.org/10.1016/J.RESCONREC.2019.01.039.

[105] S. Bamberg, I. Ajzen, P. Schmidt, Choice of Travel Mode in the Theory of Planned Behavior: The Roles of Past Behavior, Habit, and Reasoned Action, Http://Dx.Doi.Org/10.1207/S15324834BASP2503_01. 25 (2010) 175–187. https://doi.org/10.1207/S15324834BASP2503_01.

[106] B. Muñoz, A. Monzon, E. López, Transition to a cyclable city: Latent variables affecting bicycle commuting, Transp Res Part A Policy Pract. 84 (2016) 4–17. https://doi.org/10.1016/J.TRA.2015.10.006.

[107] G.J. de Bruijn, S.P.J. Kremers, A. Singh, B. van den Putte, W. van Mechelen, Adult Active Transportation: Adding Habit Strength to the Theory of Planned Behavior, Am J Prev Med. 36 (2009) 189–194. https://doi.org/10.1016/J.AMEPRE.2008.10.019.

[108] S. Saracevic, B.B. Schlegelmilch, The Impact of Social Norms on Pro-Environmental Behavior: A Systematic Literature Review of The Role of Culture and Self-Construal, Sustainability 2021, Vol. 13, Page 5156. 13 (2021) 5156. https://doi.org/10.3390/SU13095156.

[109] SPSS Software | IBM, (n.d.). https://www.ibm.com/spss (accessed April 15, 2023).

[110] SPSS Amos | IBM, (n.d.). https://www.ibm.com/products/structural-equation-modeling-sem (accessed April 15, 2023).

[111] U. Wilensky, Netlogo, (n.d.). http://ccl.northwestern.edu/netlogo/.

[112] S. Abar, G.K. Theodoropoulos, P. Lemarinier, G.M.P. O’Hare, Agent Based Modelling and Simulation tools: A review of the state-of-art software, Comput Sci Rev. 24 (2017) 13–33. https://doi.org/10.1016/J.COSREV.2017.03.001.

BIBLIOGRAPHY

RESEARCH PUBLICATIONS

This section contains the complete results of the work which were disseminated through the following international publications.

RESEARCH PUBLICATIONS

PAPER 1

Annaqeeb M K, Dziedzic J W, Yan D, Novakovic V. Exploring the tools and methods to evaluate influence of social groups on individual occupant behavior with impact on energy use. *Proceedings of the IOP Conference Series: Earth and Environmental Science*. 2019; 352 (1): 012044.

RESEARCH PUBLICATIONS

IOP Conf. Series: Earth and Environmental Science **352** (2019) 012044 doi:10.1088/1755-1315/352/1/012044

Exploring the tools and methods to evaluate influence of social groups on individual occupant behavior with impact on energy use

M K Annaqeeb[1,3], J W Dziedzic[1], D Yan[2] and V Novakovic[1]

[1]Department of Energy and Process Engineering, Norwegian University of Science and Technology, Trondheim Høgskoleringen 1, 7491, Norway

[2]School of Architecture, Tsinghua University, Beijing, 30 Shuangqing Rd, Haidian Qu, Beijing Shi, China, 100091

[3]Author to whom any correspondence should be addressed

masab.k.annaqeeb@ntnu.no

Abstract. One of the key elements in driving the energy performance of buildings has been recognized as occupant behavior (OB). However, available tools for assessing and simulating occupant behavior are based on fixed schedules and aggregated profiles, which fail to capture the diversity of OB. A significant aspect of OB is its relation to social groups and their influence and interdependence on each other. The data regarding the influence of social groups is important to achieve an effective model of OB as it accentuates the individual OB profile based on the influences it can have from the social groups they belong to. This added module is not present in traditional building simulation tools. This study aims to explore the tools and methods to evaluate the factors that are responsible for the influence of social groups on individuals' energy-related behavior. The paper investigates the kind of data sets needed for understanding this interdependence, including the occupant's social group, their standing in the group, and the intent behind different actions and its comparison to the actions the individual would take without any external influence. The results will be used to construct questionnaires, which can prove beneficial in developing social group profiles in OB models.

1. Introduction

Current statistics on worldwide energy consumption indicate that more than one-third of the expenditure comes from residential and commercial buildings [1]. Reportedly, this consumption would have been 40% higher if not for the implementation of energy-efficient technologies, but despite the technological advancement, the challenges in energy reduction are enormous, and these are compounded with the increasing global pressure and concerns to meet the worldwide goals for this reduction [1]. In addition, the disparity between the expected and the actual energy performance [2] [3]of buildings is another cause for concern, and it leads to a need of having a more innovative and holistic approach to bridge this gap and adapt to the energy demands in an efficient manner.

However, occupants are an integral part of the building systems, and due to the dynamic nature of occupant interactions with these systems, building sectors become much more complex in terms of their energy performance. While there can be several other factors involved in the contribution of the performance gap, most notably, faults in the mechanical and electrical systems, malfunctioning of equipment etc. [4], occupant behavior(OB) has been identified as one of the major drivers of this gap [5]. This trend has also been reflected in the fact that, over the last 30 years, the research community has shifted a lot of its focus towards OB [6]. This behavior, then becomes a deciding factor when it comes to the variability in the comfort settings and the implications it has on the energy usage.

In that context, there is a marked dissonance when it comes to the building design and performance simulations as well. Low-energy, or zero-energy buildings, are designed to comply with high-efficiency standards, which are vital for reducing carbon emissions, and these designs are dependent on an active occupant interaction with the building systems. However, the predictions of the energy performances result in an even larger error for these low-energy buildings [7]. This can be attributed to the actuality that the potential for achieving that high level of efficiency in these buildings is reliant on the building being operated in a specific way according to the design. Another side of this problem is also the rigidity in the building performance simulation (BPS) process, wherein OB is most often considered as a fixed or schedule-based model [8]. Earlier BPS tools utilized deterministic inputs to represent OB over the life cycle of the building, which took the form of daily, weekly, and monthly schedules. The main advantage of these inputs was the ease of use and simplicity of the models. However, more and more BPS tools have started incorporating stochastic models to capture the diversity and include the uncertainties which arise in OB [9].

The need to incorporate an interdisciplinary approach is based on this human-building dynamic, and this approach can be viewed at three different levels: the individual occupant, the group behaviors (building level), and the collective behaviors (district scale) [10]. In these terms, a multi-scalar approach, one that deals with different zones and interdisciplinary considerations of occupants' activities (habits, movement, presence, etc.) will be necessary for both the data collection and the simulation processes, and this will involve an amalgamation of engineers, architects, social scientists etc. Along these lines, Dziedzic et al. proposed a Building Occupant Transient – Agent-Based Model, which would contain several different modules for the different aspects of OB [11]. This would include distinct, individual modules in the simulations to account for the movement, habits, surroundings, and social structure.

This added modules for the social structure and influences is often neglected, since a lot of focus tends to be towards measuring the physical occupant interactions, but not the socio-psychological drivers of that behavior. Researchers have often been sensitive to the cost and technological availability, and largely ignored accompanying factors such as social equity, group status, peer influence etc. Sovacool et al. identified three main negative patterns from the energy research literature over the last 15 years and the first of them was that the social dimensions are under-examined, especially those pertaining to perceptions about energy use, and decision-making process among individuals, organizations, and communities [12].

2. Social group influences on individual occupant behavior

2.1 Occupant social groups and networks

Measurement of social influences is often conducted through the evaluation of social groups. These social networks, or groups, comprise of the group of people that occupants have to interact or share the space with. This is applicable to both residential and commercial buildings. It is an umbrella term which includes distinct nodes, where the nodes would represent the number of people in a particular group (size), the strength of the relationship between the members (connection), and in some cases, they represent events and ideas that affect the group as well [13]. These networks have been documented has having a significant effect on the energy-us behaviors of occupants [14] [15] [16]. While the framework of this paper has its focus on commercial buildings and workplaces, the intended scope is to recognize the kind of behaviors needed to be studied to collect the datasets, and that data collection can be adapted to residential buildings as well, by modifying the kinds of questions asked within the data collection surveys.

2.2 Theory of planned behavior

In order to quantify these influences and behaviors, literature in energy research and social science relies heavily on the *theory of planned behavior* [13] [18]. The theory states that there exist three factors behind the behavioral intentions and behaviors; attitudes and beliefs towards the action, beliefs about how significant others may respond to the action (*subjective norms*), and the *perceived behavioral control* (i.e. beliefs about the ease of performing the action). These three factors have been taken as the foundation in this study, to build the framework for the social structure module in OB modeling.

Abrahamse and Steg used the theory of planned behavior to explain subjective norms in residential households, that is, behaviors that would rely on the extent of importance other members of the household, and the social pressure to carry out the actions [19]. For instance, changing the thermostat settings as an energy saving measure might not be carried out if the family members disapproved of those settings. In addition, when it comes to workplaces, or commercial buildings, the utilization of energy is integrated with the group behaviors, since the equipment and facilities are often shared between co-workers. Chen and Knight used the same theory to examine the influence of colleagues among more than 500 employees in electric power companies in China, and the results indicated that the approval or disapproval from the colleagues was the major factor influencing the individual's energy saving actions [18].

Streamlining the study of these influences can also take the form of studying the social networks in relation to energy use. Anderson et al. used an agent-based model to investigate different types of social networks in buildings and the extent of their significance on administrative actions designed to reduce energy use [14]. The results indicated that accurate representations of the social networks, in relation to the historical data of the administrative actions, are vital in predicting energy use. Chen et al. also made use of agent-based modeling to build a network level computational model that simulates the decision-making process of individuals under different type of network configurations [15].

3. Framework for incorporating social influences in OB modeling

While quite a few studies have highlighted the impact of this social influence, not much research has been done on integrating this aspect in the OB simulation process. Anderson et al, did manage to conduct simulations based on occupant's susceptibility to social influences depending on their position in the social network, and the model also took into account the corresponding peer influence on their energy-use behavior [16]. However, this topic warrants a lot more study, especially in the wider context of how it fits into the overall OB profile.

3.1 General Principles

This framework takes into account all the above-mentioned factors that have been determined to be the drivers of the social influences in OB, combined with the principles of the theory of planned behavior, and dictates the datasets required from the framework's perspective. Since a major part of the topic has its focus on the social networks, this framework will involve the creation of a database that details the occupant's involvement and interaction with their respective networks. This would be a representation of the *subjective norms,* and will include a mapping of the network grid around the workplace, the number of people in the network that count as significant enough to influence the behavior, and the rank of the occupant in their networks. Considering the sharing of facilities in workplaces, the database also has to involve these interactions in designated areas, such as the common areas, lunchrooms/breakrooms, shared office spaces etc.

Regarding the *perceived behavioral control* (PBC) and the attitude/beliefs of the occupant towards different actions, gathered data has to include their perceptions about the energy-use, concerns for reducing that usage, and awareness of facilities or obstacles that accommodate a specific action. The collection of these datasets can then be utilized by connecting them to different energy-use behaviors such as thermal control, lighting control, ventilation etc.

3.2 Required Datasets to construct the module

As seen in figure 1, the datasets can be divided according to the principles, and the combined effect be correlated to the energy-related behavior using weighed factors that are derived on the basis of the extent of each parameter's perceived influence. For subjective norms, this influence would be the variability in the action, with and without the external influence. A general description of these is summarized as follows:

- Perceived behavioral control (PBC): Occupants perceptions about their own ability to conduct the action, awareness of the kinds of facilities that are available to ease the performance of the action, or obstacles that may hinder it. This could come in the form of technological aids present at the workplace, administrative actions that incentivize certain behaviors etc.
- Subjective norms: These factors will be tied in with each of the spaces that involve the interaction within the social network, and the factors are as mentioned in 3.1
- Beliefs/Attitudes: These would investigate the occupant's personal beliefs and perceptions about the action.

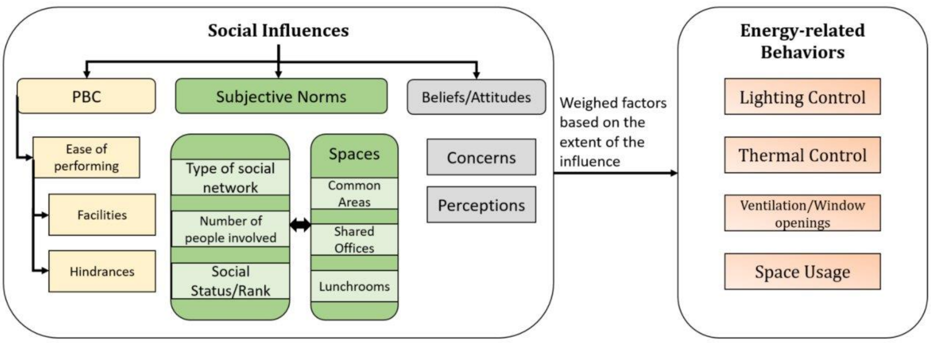

Figure 1. Schematic of datasets required.

3.3 Survey Methodology

There is a growing trend in adapting survey methodology to explore the nuances in OB. The reason behind this is that surveys still remain the most effective method for measuring variables that cannot be monitored or observed directly such as the intentions behind the energy-related behaviors, concerns, and perceived norms (latent variables) [20] [21]. However, surveys carry their own share of risks, and doubts regarding the validity, since discrepancies are often present between the reported and actual behaviors. To minimize these risks, several factors have to be taken into consideration during the design process of the survey/questionnaire. In addition, in order to incorporate the beliefs and motivations of the occupants' behind their actions, quantitative survey methods may not be sufficient [22], and qualitative interviews with the occupants might have to be combined with them to give a better understanding of the behaviour. Since the focus and scope of this work is towards OB modelling, the qualitative step will be taken into consideration after the initial results of the quantitative survey, in order to identify the factors that were insufficiently covered, and achieve better design of the process.

When measuring social and psychological parameters, it is necessary to avoid using absolute responses or brief scales, and instead adopt combined items and contextually large scales based on established theories. Vague quantifiers need to be avoided and equal number of choices should be provided on both the positive and negative sides. Apart from these, other general guidelines for designing questionnaires should be implemented as well, such as the use of simplified words over specialized ones, general aesthetic, and response that are mutually exclusive.

4. Discussions and conclusions

Energy research literature gives us definitive insights about the impact of social influences on occupant behavior, and the theory of planned behavior is an effective tool to quantify these influences. This paper brings out the necessity of considering these factors in conjunction with the energy-use behaviors to obtain a better approach to OB modeling. The guidelines and types of datasets outlined in this paper will be used to construct questionnaires and gather the required data. Results from the data collection will be vital for constructing the social modules in OB modeling, in addition to providing insights about general behavioral intentions and norms. However, most limitations here are due to the discrepancies in the reported behaviour. Special care has to be taken to avoid any kind of sampling errors, response bias, or reliance on occupant's estimations based on long-term memory.

References

[1] International Energy Agency, *World Energy Outlook* 2018 (IEA Publications)

[2] Azar E and Menassa C C 2012 A comprehensive analysis of the impact of occupancy parameters in energy simulation of office buildings *Energy and Buildings* **55** 841-53

[3] Menezes A C, Cripps A, Bouchlaghem D and Buswell R 2012 Predicted vs. actual energy performance of non-domestic buildings: Using post-occupancy evaluation data to reduce the performance gap *Applied energy* **97** 355-64

[4] Turner C and Frankel M 2008 *Energy performance of LEED for new construction buildings* (Washington DC: New Buildings Institute)

[5] Yan D, Hong T, Dong B, Mahdavi A, D'Oca S, Gaetani I and Feng X 2017 IEA EBC Annex 66: Definition and simulation of occupant behavior in buildings *Energy and Buildings* **156** 258-70

[6] Hong T, Taylor-Lange S C, D'Oca S, Yan D and Corgnati S P 2016 Advances in research and applications of energy-related occupant behavior in buildings *Energy and buildings* **116** 694-702

[7] Parker D, Mills E, Rainer L, Bourassa N and Homan G 2012 Accuracy of the home energy saver energy calculation methodology *Proc. of the 2012 ACEEE summer study on energy efficiency in buildings (Florida)* (Lawrence Berkeley National Lab Report)

[8] Hoes P, Hensen J L, Loomans M G, De Vries B and Bourgeois D 2009 User behavior in whole building simulation *Energy and buildings* **41(3)** 295-302

[9] Yoshino H, Hong T and Nord N 2017 IEA EBC annex 53: Total energy use in buildings—Analysis and evaluation methods *Energy and Buildings* **152** 124-36
[10] Hong T, Yan D, D'Oca S and Chen C F 2017 Ten questions concerning occupant behavior in buildings: The big picture *Building and Environment* **114** 518-30
[11] Dziedzic J, Yan D and Novakovic V 2017 Occupant migration monitoring in residential buildings with the use of a depth registration camera *Procedia Engineering: 10th Int. Symp. On heating, ventilation and air conditioning (Janin)* vol 205 (Janin: Elsevier) p 1193-200.
[12] Sovacool B K, Ryan S E, Stern P C, Janda K, Rochlin G, Spreng D, Pasqualetti M J, Wilhite H and Lutzenhiser L 2015 Integrating social science in energy research *Energy Research & Social Science* **6** 95-9
[13] Serrat O 2017 *Social network analysis*. In *Knowledge solutions* (Singapore: Springer) p 39-43
[14] Anderson K, Lee S and Menassa C 2012 Effect of social network type on building occupant energy use *Proc. of the Fourth ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings (Toronto)* (New York: ACM Digital Library) p 17-24
[15] Chen J, Taylor J E, Wei H H 2012 Modeling building occupant network energy consumption decision-making: The interplay between network structure and conservation *Energy and Buildings* **47** 515-24
[16] Anderson K and Lee S 2013 Modeling occupant energy use interventions in evolving social networks *2013 Winter Simulations Conference (Washington DC)* (Washington DC: IEEE) p. 3051-305
[17] Scherbaum C A, Popovich P M and Finlinson S 2008 Exploring Individual-Level Factors Related to Employee Energy-Conservation Behaviors at Work *Journal of Applied Social Psychology* **38(3)** 818-35
[18] Chen C F and Knight K 2014 Energy at work: Social psychological factors affecting energy conservation intentions within Chinese electric power companies *Energy Research & Social Science* **4** 23-31
[19] Abrahamse W and Steg L 2011 Factors related to household energy use and intention to reduce it: The role of psychological and socio-demographic variables *Human Ecology Review* **18(1)** 30-40
[20] Holopainen R, Tuomaala P, Hernandez P, Häkkinen T, Piira K and Piippo J 2014 Comfort assessment in the context of sustainable buildings: Comparison of simplified and detailed human thermal sensation methods *Building and environment* **71** 60-70
[21] Steemers K and Manchanda S 2010 Energy efficient design and occupant well-being: Case studies in the UK and India. *Building and environment* **45(2)** 270-8'
[22] Bryman A 1988 *Quantity and quality in social research* (London: Unwin Hyman)

PAPER 2

Annaqeeb M K, Markovic R, Novakovic V, Azar E. Non-intrusive data monitoring and analysis of occupant energy-use behaviors in shared office spaces. *IEEE Access,* 2020; 8:141246-141257.

IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received July 13, 2020, accepted July 23, 2020, date of publication July 29, 2020, date of current version August 13, 2020.
Digital Object Identifier 10.1109/ACCESS.2020.3012905

Non-Intrusive Data Monitoring and Analysis of Occupant Energy-Use Behaviors in Shared Office Spaces

MASAB KHALID ANNAQEEB[1], ROMANA MARKOVIC[2], VOJISLAV NOVAKOVIC[1], AND ELIE AZAR[3], (Member, IEEE)
[1]Department of Energy and Process Engineering, Norwegian University of Science and Technology, 7491 Trondheim, Norway
[2]E3D-Institute of Energy Efficiency and Sustainable Building, RWTH Aachen University, 52074 Aachen, Germany
[3]Department of Industrial and Systems Engineering, Khalifa University of Science and Technology, Abu Dhabi 127788, United Arab Emirates
Corresponding author: Elie Azar (elie.azar@ku.ac.ae)

ABSTRACT A non-intrusive data collection framework is developed to analyze the desk-level occupancy and energy use patterns of occupants in shared office spaces. The framework addresses the limitations of previous studies in the literature, which either lacked the granularity to study individual occupants' behaviors or relied on data from complex Building Management Systems (BMS). The framework is applied to a shared office space of an academic institution in the United Arab Emirates (UAE), where occupancy, lighting, and plug-load data were collected from individual desks for 6 months. The results highlight weak relationships between the occupancy status and the total electric loads, with 35% of the total electric loads consumed when the area is completely vacant, and 64% of the plug-load energy consumed when the desks were reported as unoccupied. While specific to the studied building, the results highlight the role that a high-resolution data monitoring framework plays in capturing inefficient consumption patterns. The findings also confirm the contribution of occupant behavior (OB) to the energy performance gap commonly observed between predicted and actual energy levels.

INDEX TERMS Buildings, energy conservation, monitoring, occupant behavior, performance analysis, shared office.

I. INTRODUCTION

A. BACKGROUND

Building energy management is an increasingly used process to control energy consumption and costs in buildings while maintaining comfortable indoor environmental conditions for occupants and fully meeting functional needs [1]. In commercial buildings, facility managers are often tasked with managing building energy demands and indoor conditions by continually monitoring, evaluating, and optimizing the operation efficiency of different building systems [2]. A common way to gather the needed data for analysis is through Building Management Systems (BMS). A BMS is a system of sensors, communication networks, and controls, which can be used to monitor the performance of various building systems and control their operation patterns. Benchmarking can then be performed (with the help of additional utility data), which consists of comparing the energy performance of a building to a baseline or "benchmark" [3]. The baseline can be obtained from a group of similar buildings (i.e., cross-sectional benchmarking), or from past performances of the building under study (i.e., longitudinal benchmarking). This method allows facility managers to identify higher than expected energy consumption levels. Along the same lines, Fault Detection and Diagnostics (FDD) is another method that helps identify abnormal operation patterns that can cause excessive energy use levels [4]. Based on the faults detected in different building systems, facility managers can perform corresponding maintenance actions to reduce energy consumption levels [5], [6]. Despite significant advancement in the design of efficient building systems and energy management strategies, important discrepancies are commonly found between the energy levels estimated for buildings during the design phase, and those observed during operation [3], [7]. This is often referred to as the energy performance (or efficiency) gap.

The associate editor coordinating the review of this manuscript and approving it for publication was Young Jin Chun.

B. OCCUPANT BEHAVIOR

Among various contributing factors to the energy performance gap (e.g., faults in systems or weather variations), recent studies indicate that occupant behavior (OB) is one of its major drivers since actions taken – or not taken – by occupants can significantly impact building performance [3], [7]–[10]. In [7], the authors demonstrated that different OB patterns could vary the total energy used by 150% for commercial buildings. In [8], different occupant profiles were created and used as inputs to building energy models. The authors found that occupant behavior can save up to 50% of current energy use levels or increase them by 89%.

The growing body of knowledge on the impact of OB on buildings' energy performance has motivated further research on OB data acquisition [11], analytics [12]–[15], and predictive modeling [12]. Previous studies analyzed occupancy and energy consumption data using different data granularity. For instance, zone-wise energy use data, in combination with estimated occupancy levels, were used in [13]. The authors in [16] fused measured binary zone-wise occupancy with zone-wise energy consumption. In other studies, the energy use per device type was investigated while no data regarding the measured occupancy count was available [17]–[19], or, in conjunction with binary occupancy states (e.g., [20]). In parallel, a higher granularity of the energy consumption data for single occupants has been explored. Here, the energy consumption has commonly been analyzed as a sum of all used devices [12], [21], [22]. Resultantly, the device-wise data, in combination with the occupancy data for each in-situ monitored occupant, has been rarely analyzed in the scope of existing studies. It is also important to note that the majority of studies typically gather their data through a centralized BMS, which makes their analysis replicable only to buildings with similar monitoring capabilities. Such systems are often incompatible with other existing systems (especially for old buildings) or might be too expensive to install.

C. NON-INTRUSIVE LOAD MONITORING

Non-intrusive load monitoring (NILM) is a mature technology [23] that has been widely researched and can offer an alternative to centralized data collection approaches, such as through BMS. The concept of NILM is based on having a monitoring system in the form of energy meters, smart sensors, or other sensing technologies, that separates the aggregated data regarding the electrical consumption in the area of study into individual appliance consumption profiles. The novelty of this method stems from the involvement from the user-end since it requires little to no intervention on behalf of the occupants [23], [24], and its negation of the necessity to connect the monitoring devices or sensors to existing infrastructure. It was acknowledged as a suitable data collection procedure, especially in cases where the connection to a BMS is restricted. NILM was applied to recognize the appliance types in commercial and residential buildings [24]–[28]. For instance, the authors of [26] disaggregated device types collected by NILM using Hidden Markov Model (HMM). Others proposed unsupervised learning for recognizing the device type and load segregation in household energy consumption, based on load classification and source separation [24].

However, the non-intrusive monitoring commonly applied in the context of energy analysis often lacked explicit and granular occupancy information [29]–[32]. As a consequence, the potential of NILM in combination with high-resolution occupancy data remains an open question. As an example of non-intrusive monitoring efforts, the authors of [30] investigated the relationship between the energy consumption and occupancy that was implicitly detected based on WiFi count. They concluded that the number of connected WiFi devices correlated with the resulting energy consumption. However, their scope was limited to zone-level energy load values that were estimated from trends reported by a BMS. In a more recent study, researchers proposed a non-intrusive occupancy count in an effort to optimize Heating, Ventilation, and Air Conditioning (HVAC) energy consumption [31]. They argue that reheat energy use can be reduced by 38% if the HVAC settings are adjusted to reflect actual occupancy levels. While the authors did not explore end-uses such as plug-loads or lighting levels, their findings support the potential of non-intrusive monitoring of OB in identifying energy-saving opportunities.

D. OBJECTIVES

The goal of this study is to develop a non-intrusive data collection framework that can be used to capture the energy use patterns of individual occupants in a shared office space and quantify energy-saving opportunities. The framework is unique in its ability to monitor – at the desk level – the occupancy status and energy consumption levels by type of device (e.g., monitor, laptop, task light), in addition to general lighting loads. Such a level of granularity is rarely achieved in the literature, especially for shared office spaces. Furthermore, the non-intrusive nature of the proposed data collection approach makes the framework completely independent from a BMS, and hence applicable to any building. In this context, the framework is deemed as non-intrusive in terms of the non-invasive nature of data collection to existing building systems and its independence from any BMS infrastructure. The independence from a BMS was a crucial factor in this work, which is why HVAC data was excluded since any kind of HVAC data is reliant on an existing BMS. The framework instead focused on the energy consumption patterns generated directly by the occupant at their desks, with emphasis on granularity and device utilization.

The framework is demonstrated through a case study of a shared office space in an educational facility where data is collected for a duration of 6 months. The data is then used to:

1) Quantify the relationship between the presence of individual occupants at their desks and the amount of energy they consume for two end-uses: plug-loads and lighting.

TABLE 1. Related work.

Author and year	Building type	Region (associated KGCC climate)	Occupant-related data collected	Sensing modality	Granularity
Anand et al. (2019) [33]	Academic	Singapore (Af)	Occupancy, plug-loads	Vision-based sensors, plug-load meters	Office-level
Bennet and O'Brian (2017)[13]	Commercial	Ontario, Canada (Dfb)	Lighting, plug-loads	Electricity meters	Floor-level
Gunay et al. (2016) [12]	Office (private)	Ottawa, Canada (Dfb)	Occupancy, plug-loads	PIR sensors, plug-load meters	Office-level
Gandhi and Brager (2016) [34]	Office (private, shared)	Oakland, USA (Csb)	Plug-loads	Plug-load meters	Desk-level
Rafsanjani and Ahn (2016) [35]	Academic	Lincoln, USA (Dfa)	Occupancy, plug-loads	Wi-fi, plug-load meters	Desk-level
Murtagh et al. (2013) [27]	Academic	Surrey, England, UK (Csb)	Plug-loads	Plug-load meters	Desk-level
Webber et al. (2006) [19]	Commercial (offices, schools, medical)	San Francisco, Pittsburgh, and Atlanta, USA (Csb, Dfb, Cfa)	Equipment status	Energy Audits (field investigations)	Building-level
Kawamoto et al. (2004) [17]	Office	Japan (Cfa)	Plug-loads	Field investigations	Office-Level

2) Estimate the potential energy savings from reducing consumption during unoccupied periods, both at the office level and by device type (e.g., monitors or task lights).
3) Compare the monitored energy consumption profiles of the occupants to the standard profiles (obtained from ASHRAE) that are commonly used by energy modelers when designing and simulating the performance of a similar building environment. Such an assessment can help explain – or contribute to the discussion of – the role of OB in the energy performance gap commonly observed in commercial buildings.

In summary, the main contributions of this work stem from the granularity and resolution of the data collection, coupled with the multiple sensing modalities and data streams effectively capturing the occupant presence status, the environmental parameters, and the associated plug loads of each device. This data was captured at the individual desk level, in a shared office space in Abu Dhabi, UAE, thereby also contributing a new regional perspective to a body of case studies (highlighted in the next section) that commonly originate from western countries.

II. RELATED WORK

This section explores previous research works relevant to this study, particularly OB and NILM applications. The summary is presented in Table 1, organizing the studies according to the type of buildings they covered, the region that they were conducted in, the specific type of occupant data collected, the sensing modalities used for data acquisition, and the granularity of this data. The regional information column is accompanied by the respective Köppen–Geiger climate classification (KGCC), which is one of the most commonly used classifications systems. The first letter of the classification represents one of five broad climate types (A: tropical, B: dry, C: temperate, D: continental, and E: polar). The second and third letters represent subcategories corresponding to seasonal precipitation and heat levels. For instance, 'fb', represents a warm summer humid sub climate.

The following two main observations from Table 1 reconfirm the gap in the literature and the need for the current work: (i) while plug loads have been the primary focus of the listed studies, very few of them consider occupancy parameters, especially at desk level, to better explain the impact of OB on the monitored energy data; (ii) along with the granularity and resolution of the data collection, the lack of studies from the Middle East region (e.g., KGCC of Bwh) is clear, which is another viable gap to address given the complex and case-specific nature of OB.

III. METHODOLOGY

The area of study is an open-office space in an educational building in Abu Dhabi, UAE, as shown in the bottom part of Fig. 1 (View A). It is occupied by graduate students and research employees where an individual desk is assigned to each student. The working hours vary due to the absence of official working hours, and the building operating 24/7. The shared office space consists of 6 individual desks, 2 main computer workstations (designated as shared), and a common table, as shown in Fig. 1. A total of 8 students occupied the space over the evaluation period of six months. A maximum of 6 students were present at the same time since 2 students graduated during the study period and were replaced by 2 new ones. The area is illuminated by motion-controlled area lights, while each desk is equipped with its own manually operated desk light. The area is accessible throughout the week by the employees of the educational facility. It may be noted that this is not a completely controlled environment, since the working hours are flexible, and though there are 6 primary occupants, there are no restrictions for visitors, who often do occupy the common table. This suited the

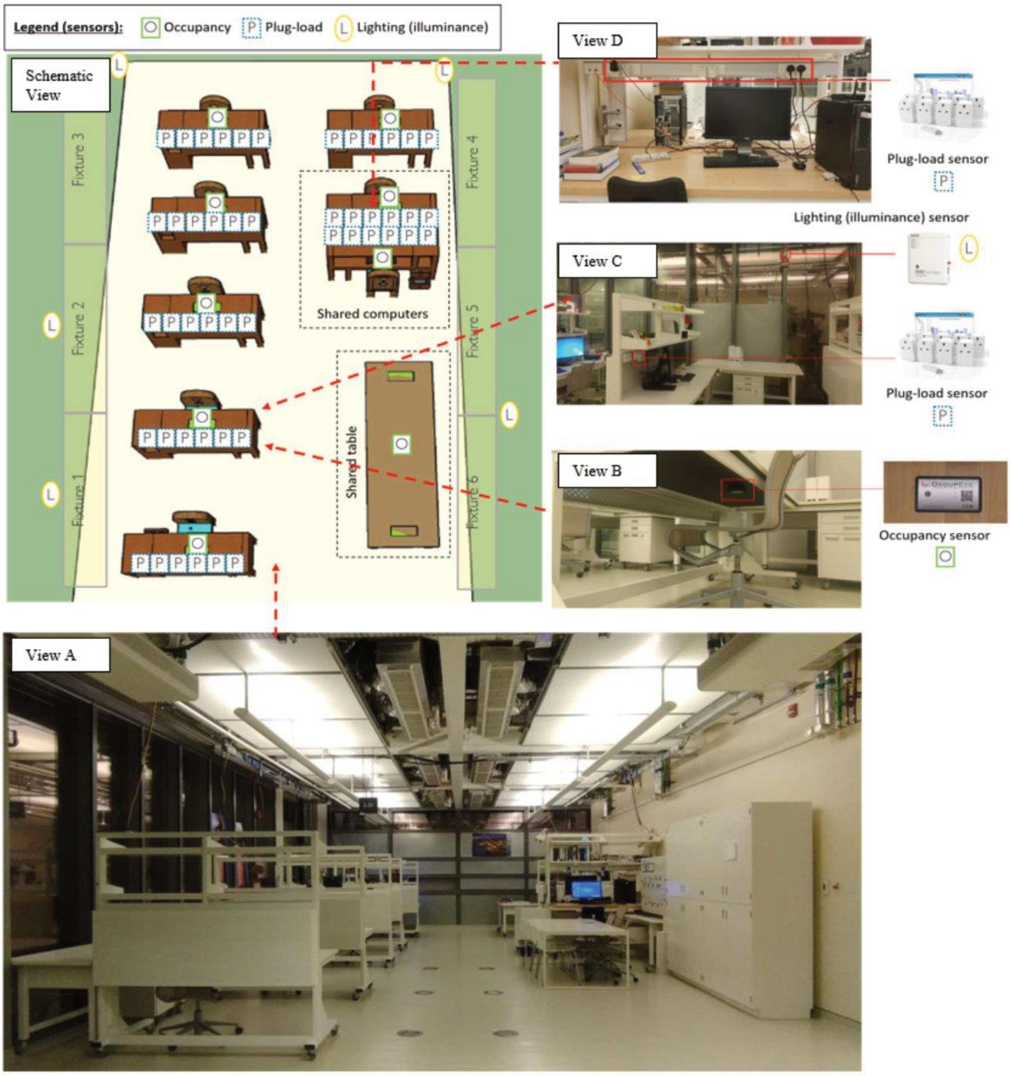

FIGURE 1. **Area of study and sensor placement.**

objectives of the work since the non-intrusive nature of the proposed monitoring and analysis approach signified that the occupants should not be imposed upon by the nature of the study, such as adding rigid working hours or restricting access.

Fig.1 presents a schematic view of the area showing the number and placement of different sensors around each work-station along with four closer looks at representative setups (views A-D). Fig.1 also shows the back-end wall of the office, which is the only façade exposed to the outdoor environment. The window-to-wall ratio (WWR) of the studied area is 26%. Overall, access to daylight was limited in the building due to the low WWR stated above in addition to external shading devices installed on the windows. These shading devices were part of passive cooling strategies meant to reduce the heat gains in the building considering the extremly hot climate of the region. This resulted in low daylight availability, which is further examined in [36], leading to a high reliance on the artificial lighting system. Moreover, the artificial lighting system was motion-based, eliminating the potential for the occupant to control its status, regardless of the amount of daylight available.

A. SENSOR INSTALLATION AND CALIBRATION

In the scope of this study, occupancy, plug-load, and lighting (illuminance) sensors were installed. The latter is used to determine the on-off status of lighting fixtures and calculate their electric energy use accordingly. The sensors were commercially available and required no connection to BMS, which allows the replicability of the study to other built environments.

1) OCCUPANCY SENSORS

A total of 9 PIR-based (passive infrared) occupancy sensors are placed in the study area, including 6 at the individual desks, 2 at the shared computer workstation, and 1 at the shared table (Refer to Fig. 1, Schematic view). Each sensor is installed under the desk it monitors, has a unique ID, and reports the occupancy status of the desk (i.e., occupied or unoccupied) in real-time to a host receiver over WiFi. More specifically, the sensor only communicates with the server (without noticeable delays) whenever there is a change in the occupancy status, such as from 'occupied' to 'unoccupied', and vice-versa. The detection range of the sensor was up to 80 meters, provided there was no obstruction. In order to calibrate the sensors, three factors were taken into account: detection of the occupant at all positions at their desk, avoidance of false triggers from passersby, and no occlusions.

The exact placement of the sensors is an essential factor in the installation/calibration process. At each desk, the sensor needs to register the presence of the occupant in all extents of their position in front of the desk. At the same time, it should avoid false triggers from people passing in the area or sitting at nearby desks. To understand the sensors' radius of influence, sensitivity, and trigger points, a sensor was first placed in a fixed position on the desk while an occupant was seated. The authors then asked the occupant to move between the ends of the desk as well as away from it while recording the occupancy status reported by the sensor. An optimal position was found and shown in Fig. 1 (View B), reporting an 'occupied' status only when the occupant was within the boundaries of the desk and up to half a meter away from it. The process was repeated for all 9 occupancy sensors, which were then connected to the host WiFi receiver. The sensors were evaluated for a period of one week by comparing their measurements to manually recorded occupancy at different times of the day. This process confirmed that the occupancy sensors were calibrated and properly placed. Additional information on the specifications of the occupancy sensors can be found on the manufacturer's website [37].

2) PLUG-LOAD SENSORS

The office space has a total of 48 different power outlets, 6 per desk, which are used by the occupants to connect their plug-loads devices (refer to Fig. 1, Schematic view). A plug-load monitoring device is installed for each power outlet (48 in total), allowing the authors to monitor the energy consumed at each outlet at a 15-min interval. Each sensor has a unique ID and connects to a host receiver over WiFi. Fig. 1 (Views C and D) show sample pictures of the sensors and their placements in the electric sockets.

The authors configured and installed each sensor, connected it to the receiver's network, and labeled it to clarify what type of plug-load it is measuring. This was done in collaboration with the occupants who agreed to connect their devices (e.g., laptops, monitors) in the specific plugs that are labeled for that use. Each occupant had 6 power outlets at their desks, which were used for different devices as follows: 1 desk lamp, 1 docking station, 2 computer monitors, 1 laptop, and 1 miscellaneous. Apart from the necessity of connecting each device to its associated power source, the occupants were free to leave any device switched on during their absence, or take it with them. The overall list of plug-loads that are measured in the shared office are listed next along with their instances: 6 docking stations, 6 laptops/notebooks, 2 desktop computers (shared), 14 computer monitors, 6 desk lamps (task lights), and 14 miscellaneous loads. Following the installation, the authors confirmed the accuracy of the sensors during a one-week evaluation period by comparing their reported energy use to the power specifications of the devices they are monitoring. This concluded the verification process for the plug-load sensors. Additional information on the specifications of the plug-load sensors can be found on the manufacturer's website [38].

3) LIGHTING SENSORS

The area of study is illuminated by 6 ceiling fixtures, which were positioned as shown in Fig. 1 (Schematic view and View A). Each of the fixtures is equipped with and is activated by motion sensors, without any option for manual control and/or dimming. In addition, natural daylight was also available, which led to a low need for using desk lamps. Given the non- intrusive nature of the proposed research approach, it was important to monitor the energy consumption of the lighting system without having to connect to the existing building's infrastructure (e.g., BMS or electric supply lines). This is achieved by installing light illuminance sensors that measure the intensity of the light (in lux) near each lighting fixture at 15-min intervals. Then, based on the monitored levels of illuminance, the on/off status of each lighting fixture can be inferred and used to determine the energy consumed by simply multiplying the duration of use by the power wattage of the fixtures.

The placement of the lighting sensors was carefully done to first, have each lighting fixture monitored by one lighting sensor, and second, minimize the noise in the measurement from neighboring lighting fixtures, or daylight. The final placement of the sensors is shown in Fig. 1 (Schematic view). It should be noted that Fixtures 5 and 6 are triggered by the same motion sensor; hence, one lighting sensor was placed between these fixtures. Following the installation of the sensors, the illuminance output of the sensors was monitored for one week, and thresholds were set to distinguish between the on/off statuses of the fixtures. Put differently, it was important to identify a reference illuminance value (for each light illuminance sensor) above which the neighboring lighting fixture is "on", and below which it is "off". After analyzing the data that was collected, it was noticed that the average value between the lowest and highest illuminance value monitored by each sensor is a good and reliable threshold. In addition, since the sensors were positioned very close to the light sources, the availability of daylighting did not create a significant difference in these illuminance thresholds. Observations at different times of the day confirmed the adequacy of this approach in implying the actual on/off status of each fixture.

Finally, unlike the previous occupancy and plug-load sensors, lighting sensors stored the data locally. Manual data transfers were performed bi-weekly through the USB ports of the sensors and stored on a local computer. Additional information on the specifications of the lighting sensors can be found on the manufacturer's website [39].

B. DATA COLLECTION AND PROCESSING

Following the verification process for the three types of sensors described above, data was collected for a period of 6 months, covering the months of April-May, July-August, and October-November 2017. A routine check (weekly or bi-weekly) was carried out to ensure that the sensors are working properly, restarting, or reconfiguring malfunctioning sensors when needed. All the collected data (over WiFi or manually) were gathered in one CSV Excel spreadsheet file for data processing and analysis.

In this study, the plug-load and lighting sensors reported data in a 15-min interval, while the occupancy sensors only reported data when the occupancy status changes (i.e., a space became 'occupied' or 'unoccupied'). To ensure consistency between the different datasets, the occupancy sensor data was converted to a 15-min format by setting the last recorded trigger type (i.e., occupied/unoccupied) within each 15-min time as the new occupancy status for that period. If no triggers are observed within a period, the state of the preceding period is taken by default. Lighting energy consumption was calculated for each fixture in 15-min intervals. If the fixture was estimated to be "on", then its energy consumption is calculated by multiplying the power wattage of each fixture, which consists of two lamps consuming 28W each, by 15 minutes.

C. VALIDATION PROCESS

As mentioned in the previous section, routine checks were carried out every week to ensure the proper functioning of sensors. This would comprise of manual notation of occupancy triggers and checking them with the reported data, along with checking the functionality of plug-load sensors to ensure that they are representational of the occupant activities. The illuminance sensors were battery-operated and had to be accessed through their respective dashboards to ensure that they had enough power to function for the next few weeks. Written consent was obtained from all the occupants included in the study. The weeks that had a significant amount of missing data, or malfunctioning sensors, were removed before the final data analysis. However, it should be noted that those weeks were mostly at the beginning of the course of study, and after conducting the required adjustments, the study was able to operate smoothly.

D. DATA ANALYSIS

The data analysis consists of three main stages. The first stage is to quantify the relationships between the presence of occupants in the office space and the levels of energy that are consumed. A distinction is made between the energy consumption of plug-loads and that of general lighting loads. Plug-loads are directly controlled by the occupants, which makes it their responsibility to operate these loads efficiently (i.e., turning them off when leaving a space). Lighting loads are triggered by dedicated motion sensors that are installed and maintained by facility managers (FMs). Therefore, it is the responsibility of FMs to ensure the proper calibration and operation of these systems to avoid over-lighting (e.g., passing occupants triggering the lighting system) or under-lighting (e.g., motion sensors failing to capture occupancy presence). The link between occupancy presence and different end-use consumption is studied using box-whisker-plots of the monitored power levels for varying levels of occupancy; the minimum occupancy level being zero occupancy sensors indicating an "occupied" status while the maximum occupancy level being all 9 sensors indicating an "occupied" status. Such a representation helps visualize how dependent energy demand is given a certain level of occupancy in the space.

The second stage consists of quantifying the amount of energy that is consumed while the occupants are away from their desks. In theory, this portion of energy can be considered as unnecessary or as a potential for energy conservation. In practice, some plug-loads such as the shared desktop computers may be used to run experiments or simulations without the presence of occupants at the desk, which may explain a portion of any energy consumed during vacant periods. Therefore, it is essential to make such a distinction before labeling some of its energy consumed as unnecessary. Two particular analyses are presented. The first is an office level analysis comparing the energy consumption of the office during "vacant" periods (i.e., zero occupancy sensors triggered) and "occupied" periods (i.e., at least one occupancy sensor triggered). This approach is commonly used in the literature when occupancy is monitored at the office level rather than the desk and corresponding device (i.e., plug-load level). The second analysis covers the performance of specific device- types (e.g., laptops, monitors, and docking stations) while distinguishing between their energy consumption while occupants were present at or away from their desks. In this stage of the analysis, the "vacant" and "occupied" occupancy statuses refer to the presence and absence of occupants from their desks (respectively), while a specific plug-load was consuming energy.

The last stage of analysis consists of developing diversity profiles (i.e., schedules) of the occupancy, lighting, and

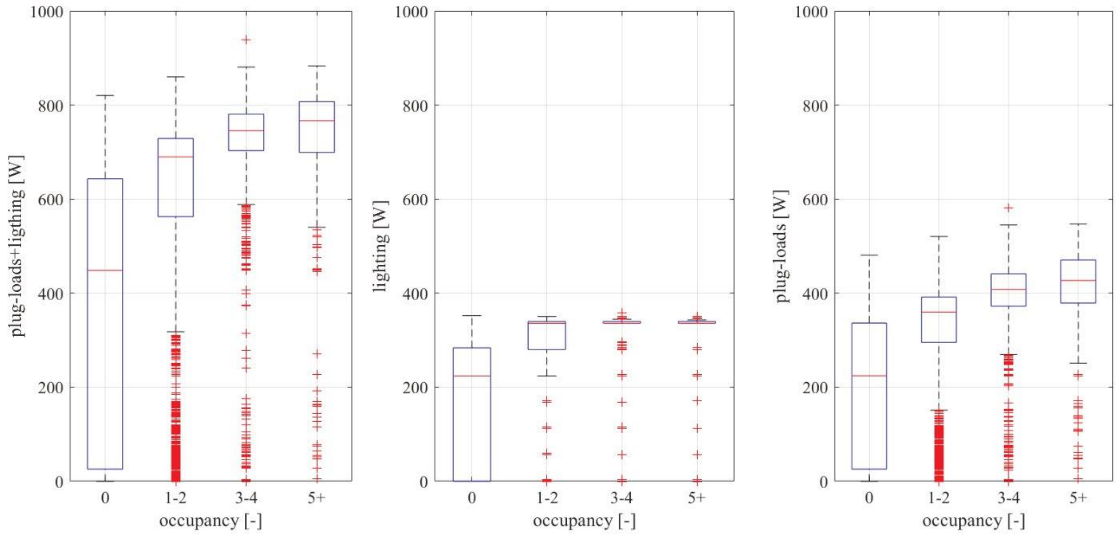

FIGURE 2. Box-and-whisker plots of electric load function of the number of occupants' sensors that are indicating occupied status.

plug- load patterns observed in the studied area. Diversity profiles show the intensity of a variable (i.e., occupancy, lighting, and plug-load) over the 24 hours of a typical day. The intensity is expressed in numerical values from 0 to 1, where 0 represents the minimum possible value (e.g., no occupancy or no energy consumption), and 1 represents the maximum possible value (e.g., maximum occupancy or maximum energy consumption). Diversity profiles are commonly used by energy modelers when simulating/predicting the performance of actual buildings.

The American Society of Heating and Refrigeration Engineers (ASHRAE) has developed profiles for common building types (e.g., office), which are extensively used in the literature (e.g., [40]). However, recent research (e.g., [41]) shows significant discrepancies between ASHRAE profiles and those observed in actual buildings. Moreover, since the developed profiles are for common building types, academic buildings are not explicitly covered; there is often a need to build new schedules for these buildings to achieve an efficient energy management system [42]. This has motivated the current stage of the analysis, where a comparison is conducted between the ASHRAE profiles and those of the studied area. In total, diversity profiles are developed for occupancy levels, lighting, and plug-loads energy consumption. A distinction is also made between weekdays and weekends, which typically witness different occupancy and energy use patterns.

The occupancy diversity value for a typical hour h of the day (e.g., 00:00-01:00 am) is computed using (1), which averages the ratio of sensors with "occupied" status over the total number of sensors over the study period from the first day "d" of the study to day N. Equation (2) is used to compute the diversity factors for plug-loads and lighting energy use. The diversity for the end-uses is defined as the ratio of the observed energy (i.e., monitored) to the maximum energy observed throughout the study. Here again, to obtain the diversity value for an hour h, the values observed for that hour over the N days of the study are averaged.

$$Occupancy\ Diversity = \frac{\sum_{d=1} N \frac{\#Sensors\ Occupied}{\#Sensors\ Total}}{N} \tag{1}$$

$$Energy\ Diversity = \frac{\sum_{d=1} N \frac{Energy\ Consumed}{Max\ Energy\ Observedl}}{N} \tag{2}$$

IV. RESULTS

A. RELATIONSHIPS BETWEEN OCCUPANCY LEVEL AND ELECTRIC LOAD

The relationship between the occupants' count and electricity consumption in the space is analyzed using box (and whisker) plots (Fig. 2). Here, the median consumption (central red-colored line), 25^{th} and 75^{th} quartiles (blue-colored box limits), and the outliers (red-colored crosses outside of the boxes) are used to present the distribution of the measured power loads. The resulting boxplots show that the total lighting and plug-loads power could be correlated to the number of present occupants. However, the baseline energy values corresponding to the null occupancy are higher than expected, as detailed next.

Starting with the lighting system (central box plot in Fig. 2), during vacant periods (i.e., 0 occupants present in the space), the mean power values for the lighting system exceeds 200 W, with 25% and 75% quantiles between approximately 0 W and 300 W, respectively. The lighting system, which is controlled by motion sensors, seems to be triggered

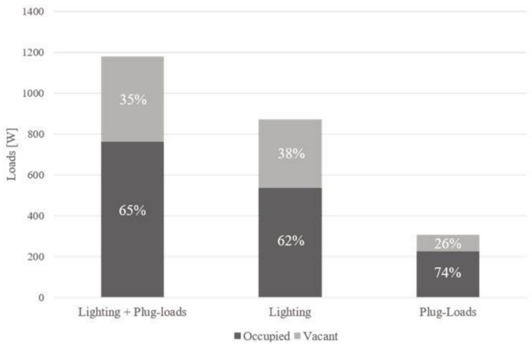

FIGURE 3. Office-level comparison of the energy consumed during vacant and occupied periods. "Vacant" refers to periods where there are no people in the office space. "Occupied" refers to periods with at least one occupant in the office.

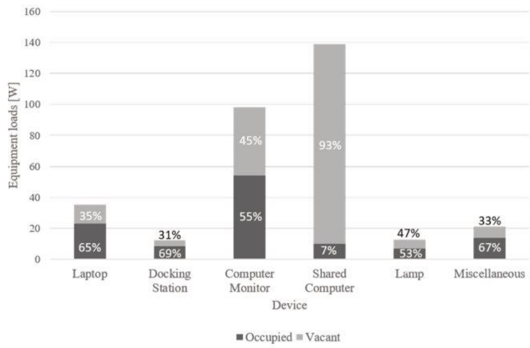

FIGURE 4. Device-level comparison of the energy consumed during vacant and occupied periods. "Vacant" refers to periods where a specific device was running without the presence of an occupant at the desk. "Occupied" refers to periods where an occupant was present at the desk during the operation of the device.

by occupants passing by the monitored workspace. Such a scenario was tested by the researchers who confirmed that walking in the hallway near the office space was triggering the lighting system. Additionally, for the three higher ranges of occupancy (i.e., 1-2, 3-4, and 5+ occupants), the average power level is almost the same (~350W). The results indicate that the presence of 1-2 occupants was sufficient to activate the lighting for the whole workspace, despite having 5 lighting zones that are controlled by independent occupancy motion sensors. These results pointed out the importance of the choice of suitable light control systems and the crucial role of the building operation management to ensure the proper calibration and maintenance of the systems.

The box-plots of the plug-loads, shown on the right side of Fig. 2, reveal similar inefficient patterns as with the lighting system. High power levels are observed during vacant periods, here again exceeding 200 W in average value. The results imply that occupants are constantly leaving plug-load equipment running when leaving their desks, which is further explored in the upcoming sections. However, unlike the lighting systems, there is a positive relationship between the average power levels and the occupancy count bins of 1- 2, 3-4, and 5+ occupants. Such a trend was expected as an office space with a higher number of occupants is expected to consume more, and vice-versa.

Finally, the patterns of total electric loads shown on the left side of Fig. 2 are simply an addition of the lighting and plug-load power values. The box plots are characterized by high consumption values during vacant periods (average power levels exceeding 400 W) and a moderate positive relationship between occupancy count and total electric loads.

B. ENERGY CONSUMPTION DURING "OCCUPIED" AND "VACANT" PERIODS

1) OFFICE-LEVEL ANALYSIS

The energy consumption of the lab was analyzed for the periods of occupancy (i.e., with 1 occupant or more) and vacancy (i.e., 0 occupants). The results are presented in Fig. 3. The bar chart on the left side shows that 35% of the lab's electric consumption occurs during the vacant hours. A similar trend could be observed when energy consumption was separately analyzed by end-use. Here, 38% and 26% of the consumption occurred during vacant hours for the lighting system and plug-loads, respectively. These results confirmed the findings of previous studies (e.g. [12], [43], [44]) that have observed a significant proportion of building energy systems running after hours.

It is important to note that the observed values for the plug-loads are conservative due to the definition of the office-level "occupied" period that was used; the space is considered occupied if there is at least one occupant present. Such an assumption might lead to an overestimation of the plug-load energy consumed during operation and an underestimation of the portion consumed during "vacant" periods. For instance, in a scenario where occupant A is the only occupant present in the shared office while the computers and monitors of all occupants are running during that period, the energy consumed by all occupants would be labeled as energy during an "occupied" period. Hence, the results presented in Fig. 3 (right side) can be considered conservative, and in reality, more plug-load energy is consumed when individuals are away from their desks, as shown in the next subsection.

2) DEVICE-LEVEL ANALYSIS

Eventually, the energy consumption for occupied and absent periods was analyzed at the device-level granularity. In this section, the occupancy status of each desk was evaluated separately and was used to classify the energy consumption of associated plug-loads between "vacant" and "occupied". As presented in Fig. 4, between 31% and 93% of the energy consumed by each type of the plugged-in device was consumed while no occupancy was detected at the respective desk. Hence, by recalculating the total plug-load energy consumed during "vacant" and "occupied" periods at the desk-level, it is seen that 64% of the plug-load energy is actually

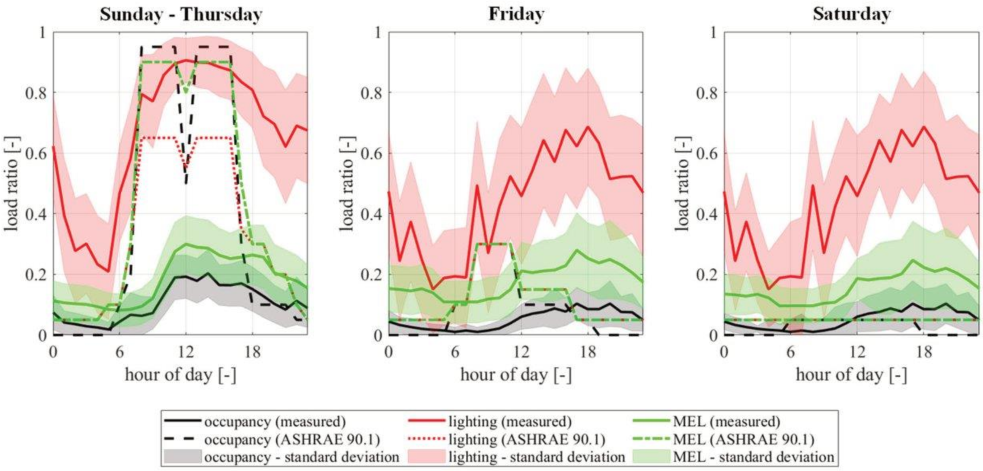

FIGURE 5. Comparison between the ASHRAE 90.1 profiles with the measured occupancy, plug loads, and lighting energy consumption.

consumed when the occupancy status of the desks is reported as 'vacant'. As such, even though all the 8 desks cannot be occupied at the same time, the occupants are responsible for switching off relevant equipment when leaving the workstation, which is why any consumption at a desk when it is unoccupied is classified as 'unnecessary consumption'.

A major contributor to this misbalance is the use of the shared personal computer, with 93% of its energy consumption occurring without the presence of an occupant. This can be caused by the use of computers to run experiments scheduled by occupants for unoccupied periods, and, resultantly, could not be identified as a clear saving potential. While such kind of shared computer is specific to this particular academic setting, shared equipment (e.g., printers and appliances) is an integral part of office spaces and commercial buildings. Therefore, the need to identify instances of unoccupied consumption can be beneficial in cases wherein they do represent a clear saving potential.

In parallel, end-uses, such as monitors and lamps, also show important energy consumption during vacant periods and can be easy targets of energy curtailment efforts. In summary, the adopted device-wise monitoring granularity revealed important energy-saving potentials that were not observed when considering occupancy in the office as a binary variable, as in the previous section.

C. DIVERSITY PROFILES AND COMPARISON TO ASHRAE

Eventually, the mean daily occupancy and energy consumption patterns were analyzed and compared to the schedules recommended by ASHRAE 90.1 [40] for office buildings (Fig. 5). The figure illustrates the difference between the measured loads and the ASHRAE-proposed profiles for each operating schedule, including workdays and the two profiles defined for weekends. The workdays, as defined by ASHRAE, were the typical working days (i.e., Sunday to Thursday in Abu Dhabi). Consequently, the ASHRAE 90.1 schedules proposed for Saturday and Sunday were compared to the data collected on Fridays and Saturdays, respectively.

The results showed that the schedules proposed by the guidelines could not realistically depict the magnitude nor the course of the occupancy and energy consumption in the building in question. Significant deviations are observed between the standard schedules and the measured occupancy and energy consumption. To shed more light on the discrepancies, Fig. 6 presents the absolute hourly error between the theoretical and observed occupancy, lighting, and plug-loads. The error values show that occupancy and plug-loads were (for the most part) overestimated over the typical working hours of workdays while being underestimated outside of working hours. In contrast, lighting loads were consistently underestimated throughout the week. The following are potential contributing factors to the observed discrepancies.

Firstly, the studied area is a shared office space in an educational facility, where the occupants are graduate students. Unlike a traditional office space with clear working hours, the studied environment provides the researchers with a flexible working schedule that allows them to attend classes, events, or work remotely. The students often choose to work over the weekends, which is less common in traditional office environments. Secondly, the number of occupants studied is relatively low, which makes the impact of individual behavior significant on the general patterns shown in Fig. 5. Such an effect will be less significant if the office has a higher number of occupants. Thirdly, schedules, such as ASHRAE's, assume a good correlation between occupancy patterns and the energy use levels of systems, such as lighting

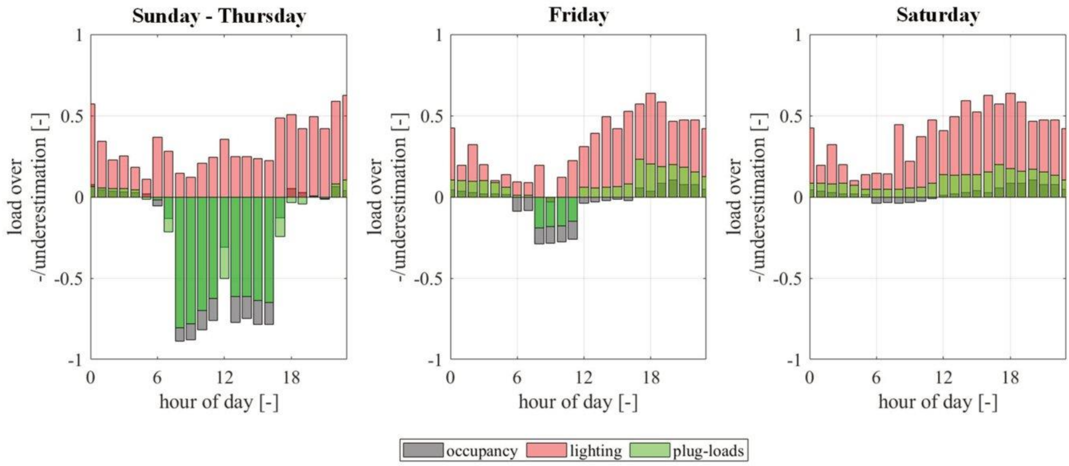

FIGURE 6. Error bars of the measured and standard diversity factors for occupancy, plug loads, and lighting energy consumption. The colors overlap in cases where the direction of over/underestimation of the variables is the same at a particular hour of the day.

and plug-loads. The results from Fig. 5 show that the lighting and plug-load patterns seem to follow the occupancy patterns to a good degree. However, there is an important difference in the scale or magnitude of the profiles. From an operation perspective, the observed gap reconfirms the inefficient use of the lighting and equipment, where a low occupancy level can lead to electric consumption levels near their maximum possible values. From a design perspective, the observed discrepancies are difficult to account for during the design stage and could contribute to misestimations of the actual energy use levels (i.e., energy performance gap).

V. DISCUSSION, LIMITATIONS, AND FUTURE WORK

The framework presented in this paper evaluated and quantified the weak relationships between occupancy patterns and electric loads in the studied space. More importantly, the granularity of analysis ensured the identification of the causes of those discrepancies. As discussed earlier, one primary source of unnecessary lighting consumption at the office level was attributed to the improper calibration and maintenance of the motion sensors activating the lights. This can be an indication of mismanagement by facility managers, as well as miscommunication from the occupants who failed to report inefficiencies in building systems. A non-intrusive framework with high levels of granularity has the potential to identify such discrepancies and bring it to the notice of the facilities management, so that appropriate action can be taken to avoid more unnecessary consumption. As such, this framework can be vital to the FDD process as well. It can also motivate the need for more effective communication channels between FM and occupants.

It is important to highlight some assumptions that were made in the study, along with its limitations. The first assumption was regarding lighting consumption. In order to maintain the framework's non-reliance on a functioning BMS, the lighting consumption was calculated based on the nominal wattage and the light status (i.e., on or off), which works on the assumption that the nominal wattage is a good approximation of the actual consumption. While such an approach has been used in the literature [45], [46], deviations from the actual levels can be observed due, for instance, to inefficiencies in electronic components.

Another limitation of the framework is that it does not currently measure or estimate HVAC loads. Such an addition is essential to make the proposed framework comprehensive and provide a holistic evaluation of building energy performance. Another potential expansion of the work can include indoor environmental factors and various metrics of occupants' comfort (e.g., thermal, visual, acoustic, etc.). This will allow understanding and capturing adaptive actions that occupants may take to maximize their comfort, which in turn affect building performance (e.g., window opening).

When it comes to the ''non-intrusive'' description of the framework, the term in this paper referred to the concept of minimal deployment of sensors on the property [23], [25], with a special emphasis put on the independence from existing BMS infrastructure. This was rarely achieved in similar studies in the literature. However, ''non-intrusive'' can also refer to data collection methods that protect and anonymize information collected from individuals. While important steps were taken in this study to protect the occupants' information, this process can be further developed and standardized as part of future research and before any deployment of the framework at larger scales.

Finally, an important limitation pertaining to the case study is the small sample size used. However, it is important to note that the aim of the case study was to illustrate and validate the capabilities of the proposed framework, which, in the opinion of the authors, was well achieved. Nonetheless, applications of the framework on a large number of occupants can further confirm the observed findings and draw general conclusions on the role of occupants in achieving low-energy building performance.

VI. CONCLUSION

This paper proposed a non-intrusive data collection and framework that was used to capture the energy use patterns of individual occupants in a shared office space and identify areas for energy savings. The framework is characterized by its ability to capture occupancy presence and plug-load usage at the desk level. This helped identify energy-saving opportunities that could not be captured with a more general office-level analysis of energy use. Moreover, the non-intrusive characteristics of the framework make it independent of a BMS infrastructure and easily applicable to other building environments.

The application of the framework to a shared office space of an educational facility confirmed the capabilities of the framework by observing: (1) large amounts of energy being observed when no occupants are in the office; (2) 64% of the energy consumed by plug-load devices occurring when the desks are unoccupied; and (3), large discrepancies between the observed occupancy and energy consumption profiles on the one hand, and those provided by ASHRAE, on the other.

It is important to note that the specific results that were obtained are not meant to be generalized nor extrapolated to typical academic buildings. Rather, they serve to showcase the capabilities of the proposed high-resolution data monitoring infrastructure in identifying unexpected (and potentially wasteful) energy consumption patterns in office spaces. More specifically, the findings highlight the important impact of people on building energy performance. These include facility managers, in their role of maintaining and calibrating centralizing systems (e.g., lighting), and occupants, in their role of operating end-uses that they control (e.g., plug-loads).

Moreover, the difference observed in the diversity profiles (Fig. 5 and Fig. 6) from ASHRAE contributes to the growing body of literature on the drivers of the energy performance gap commonly observed between predicted and actual energy use levels. While predicting occupancy patterns and behaviors during the design phase is a highly complex – if not impossible – task, designers can apply methods, such as uncertainty analysis and parametric variation, to quantify the performance risk and then apply robust design practices to mitigate it. In parallel, large-scale data collection efforts can help refine the diversity profiles used in building standards, adjusting for different building types, levels of automation, and geographical locations.

REFERENCES

[1] *Energy Efficiency in Existing Buildings: Applications, Updates and Plans*, Standard ANSI/ASHRAE/IES Standard 100-2015, 2015.
[2] M. Donnelly, *Building Energy Management: Using Data as a Tool*. Cork, Ireland: Johnson Control, 2012.
[3] J. Granderson, M. A. Piette, and G. Ghatikar, "Building energy information systems: User case studies," *Energy Efficiency*, vol. 4, no. 1, pp. 17–30, Feb. 2011.
[4] X. Li, C. P. Bowers, and T. Schnier, "Classification of energy consumption in buildings with outlier detection," *IEEE Trans. Ind. Electron.*, vol. 57, no. 11, pp. 3639–3644, Nov. 2010.
[5] G. P. Henze, G. S. Pavlak, A. R. Florita, R. H. Dodier, and A. I. Hirsch, "An energy signal tool for decision support in building energy systems," *Appl. Energy*, vol. 138, pp. 51–70, Jan. 2015.
[6] S. Katipamula and M. R. Brambley, "Review article: Methods for fault detection, diagnostics, and prognostics for building systems—A review—Part I," *HVAC R Res.*, vol. 11, no. 1, pp. 3–25, 2005.
[7] C. M. Clevenger, J. R. Haymaker, and M. Jalili, "Demonstrating the impact of the occupant on building performance," *J. Comput. Civil Eng.*, vol. 28, no. 1, pp. 99–102, Jan. 2014.
[8] T. Hong, Y. Chen, Z. Belafi, and S. D'Oca, "Occupant behavior models: A critical review of implementation and representation approaches in building performance simulation programs," *Building Simul.*, vol. 11, no. 1, pp. 1–14, Feb. 2018.
[9] D. Yan, W. O'Brien, T. Hong, X. Feng, H. Burak Gunay, F. Tahmasebi, and A. Mahdavi, "Occupant behavior modeling for building performance simulation: Current state and future challenges," *Energy Buildings*, vol. 107, pp. 264–278, Nov. 2015.
[10] K. Schakib-Ekbatan, F. Z. Çakıcı, M. Schweiker, and A. Wagner, "Does the occupant behavior match the energy concept of the building?—Analysis of a german naturally ventilated office building," *Building Environ.*, vol. 84, pp. 142–150, Jan. 2015.
[11] B. Dong, M. B. Kjærgaard, M. De Simone, H. B. Gunay, W. O'Brien, D. Mora, J. Dziedzic, and J. Zhao, "Sensing and data acquisition," in *Exploring Occupant Behavior in Buildings*. Cham, Switzerland: Springer, 2017, pp. 77–105.
[12] H. B. Gunay, W. O'Brien, I. Beausoleil-Morrison, and S. Gilani, "Modeling plug-in equipment load patterns in private office spaces," *Energy Buildings*, vol. 121, pp. 234–249, Jun. 2016.
[13] I. E. Bennet and W. O'Brien, "Office building plug and light loads: Comparison of a multi-tenant office tower to conventional assumptions," *Energy Buildings*, vol. 153, pp. 461–475, Oct. 2017.
[14] E. Azar and C. C. Menassa, "A comprehensive framework to quantify energy savings potential from improved operations of commercial building stocks," *Energy Policy*, vol. 67, pp. 459–472, Apr. 2014.
[15] K. Sun and T. Hong, "A framework for quantifying the impact of occupant behavior on energy savings of energy conservation measures," *Energy Buildings*, vol. 146, pp. 383–396, Jul. 2017.
[16] Z. Wang, T. Hong, and M. A. Piette, "Data fusion in predicting internal heat gains for office buildings through a deep learning approach," *Appl. Energy*, vol. 240, pp. 386–398, Apr. 2019.
[17] K. Kawamoto, Y. Shimoda, and M. Mizuno, "Energy saving potential of office equipment power management," *Energy Buildings*, vol. 36, no. 9, pp. 915–923, Sep. 2004.
[18] L. Moorefield, B. Frazer, and P. Bendt, "Office plug load field monitoring report," ECOS, Durango, CO, USA, Tech. Rep., 2008.
[19] C. Webber, J. Roberson, M. Mcwhinney, R. Brown, M. Pinckard, and J. Busch, "After-hours power status of office equipment in the USA," *Energy*, vol. 31, no. 14, pp. 2823–2838, Nov. 2006.
[20] C. Lobato, S. Pless, M. Sheppy, and P. Torcellini, "Reducing plug and process loads for a large scale, low energy office building: NREL's research support facility," Natl. Renew. Energy Lab.(NREL), Golden, CO, USA, Tech. Rep., NREL/CP-5500-49002, 2011.
[21] A. Mahdavi, F. Tahmasebi, and M. Kayalar, "Prediction of plug loads in office buildings: Simplified and probabilistic methods," *Energy Buildings*, vol. 129, pp. 322–329, Oct. 2016.
[22] Y. Agarwal, B. Balaji, R. Gupta, J. Lyles, M. Wei, and T. Weng, "Occupancy-driven energy management for smart building automation," in *Proc. 2nd ACM Workshop Embedded Sens. Syst. Energy-Efficiency Building BuildSys*, 2010, pp. 1–6.
[23] G. W. Hart, "Nonintrusive appliance load monitoring," *Proc. IEEE*, vol. 80, no. 12, pp. 1870–1891, Dec. 1992.
[24] R. Bonfigli, S. Squartini, M. Fagiani, and F. Piazza, "Unsupervised algorithms for non-intrusive load monitoring: An up-to-date overview," in *Proc. IEEE 15th Int. Conf. Environ. Electr. Eng. (EEEIC)*, Jun. 2015, pp. 1175–1180.
[25] A. Zoha, A. Gluhak, M. Imran, and S. Rajasegarar, "Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey," *Sensors*, vol. 12, no. 12, pp. 16838–16866, Dec. 2012.

[26] O. Parson, S. Ghosh, M. Weal, and A. Rogers, "Non-intrusive load monitoring using prior models of general appliance types," in *Proc. 26th AAAI Conf. Artif. Intell.*, Jul. 2012, pp. 1–7.

[27] N. Murtagh, M. Nati, W. R. Headley, B. Gatersleben, A. Gluhak, M. A. Imran, and D. Uzzell, "Individual energy use and feedback in an office setting: A field trial," *Energy Policy*, vol. 62, pp. 717–728, Nov. 2013.

[28] S. Makonin, F. Popowich, I. V. Bajic, B. Gill, and L. Bartram, "Exploiting HMM sparsity to perform online real-time nonintrusive load monitoring," *IEEE Trans. Smart Grid*, vol. 7, no. 6, pp. 2575–2585, Nov. 2016.

[29] H. Rafsanjani, C. Ahn, and M. Alahmad, "A review of approaches for sensing, understanding, and improving occupancy-related energy-use behaviors in commercial buildings," *Energies*, vol. 8, no. 10, pp. 10996–11029, Oct. 2015.

[30] J. Chen and C. Ahn, "Assessing occupants' energy load variation through existing wireless network infrastructure in commercial and educational buildings," *Energy Buildings*, vol. 82, pp. 540–549, Oct. 2014.

[31] O. Ardakanian, A. Bhattacharya, and D. Culler, "Non-intrusive occupancy monitoring for energy conservation in commercial buildings," *Energy Buildings*, vol. 179, pp. 311–323, Nov. 2018.

[32] T. Zhang and O. Ardakanian, "A domain adaptation technique for fine-grained occupancy estimation in commercial buildings," in *Proc. Int. Conf. Internet Things Design Implement.*, Apr. 2019, pp. 148–159.

[33] P. Anand, D. Cheong, C. Sekhar, M. Santamouris, and S. Kondepudi, "Energy saving estimation for plug and lighting load using occupancy analysis," *Renew. Energy*, vol. 143, pp. 1143–1161, Dec. 2019.

[34] P. Gandhi and G. S. Brager, "Commercial office plug load energy consumption trends and the role of occupant behavior," *Energy Buildings*, vol. 125, pp. 1–8, Aug. 2016.

[35] H. N. Rafsanjani and C. Ahn, "Linking building energy-load variations with Occupants' energy-use behaviors in commercial buildings: Non-intrusive occupant load monitoring (NIOLM)," *Procedia Eng.*, vol. 145, pp. 532–539, Jan. 2016.

[36] M. S. Andargie and E. Azar, "An applied framework to evaluate the impact of indoor office environmental factors on occupants' comfort and working conditions," *Sustain. Cities Soc.*, vol. 46, Apr. 2019, Art. no. 101447.

[37] *Home Page—OccupEye | IoT*, Occupeye, Blackburn, U.K., 2019.

[38] *Plugwise o Intelligent Heating Control*. Accessed: Jul. 7, 2020. [Online]. Available: https://www.plugwise.com/nl_NL/

[39] *HOBO Data Loggers for all Measurements or Applications*. Accessed: Jul. 7, 2020. [Online]. Available: https://www.onsetcomp.com/products/

[40] *Energy Standard for Building Except Low-Rise Residential Buildings*, Standard ANSI/ASHRAE/IES Standard 90.1-2019, 2019.

[41] C. Duarte, K. Van Den Wymelenberg, and C. Rieger, "Revealing occupancy patterns in an office building through the use of occupancy sensor data," *Energy Buildings*, vol. 67, pp. 587–595, Dec. 2013.

[42] J. A. Davis and D. W. Nutter, "Occupancy diversity factors for common university building types," *Energy Buildings*, vol. 42, no. 9, pp. 1543–1551, Sep. 2010.

[43] E. Azar and C. C. Menassa, "Optimizing the performance of energy-intensive commercial buildings: Occupancy-focused data collection and analysis approach," *J. Comput. Civil Eng.*, vol. 30, no. 5, Sep. 2016, Art. no. C4015002.

[44] O. T. Masoso and L. J. Grobler, "The dark side of occupants' behaviour on building energy use," *Energy Buildings*, vol. 42, no. 2, pp. 173–177, Feb. 2010.

[45] M. Sepehr, R. Eghtedaei, A. Toolabimoghadam, Y. Noorollahi, and M. Mohammadi, "Modeling the electrical energy consumption profile for residential buildings in iran," *Sustain. Cities Soc.*, vol. 41, pp. 481–489, Aug. 2018.

[46] D. Bonino, F. Corno, and L. De Russis, "Poweront: An ontology-based approach for power consumption estimation in smart homes," in *Internet of Things. User-Centric IoT* (Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering), vol. 150. Cham, Switzerland: Springer, 2015, pp. 3–8.

MASAB KHALID ANNAQEEB received the bachelor's degree in civil engineering from Osmania University, Hyderabad, India, in 2016, and the M.S. degree in sustainable critical infrastructure from Khalifa University, Abu Dhabi, United Arab Emirates, in 2018. He is currently pursuing the Ph.D. degree in energy and process engineering with the Norwegian University of Science and Technology. His research interests include study of occupant behavior in buildings and modeling of simulation of the same, investigation of different aspects of occupant behavior, and relevant techniques to simulate them.

ROMANA MARKOVIC received the M.S. degree in civil engineering from RWTH Aachen University, Germany, in 2016, where she is currently pursuing the Ph.D. degree with the E3D-Institute of Energy Efficiency and Sustainable Building. She is also a Research Assistant with the E3D-Institute of Energy Efficiency and Sustainable Building, RWTH Aachen University. Her research interests include occupant behavior modeling and simulation, with a focus on generic occupant behavior modeling in commercial buildings.

VOJISLAV NOVAKOVIC was a Researcher with Stiftelsen for Industriell og Teknisk Forskning (SINTEF), from 1983 to 1994. He has been with the Norwegian University of Science and Technology, Trondheim, Norway, since 1990, where he is currently a Professor with the Department of Energy and Process Engineering. His current research interest includes monitoring and simulation of energy-related occupant behavior in buildings.

ELIE AZAR (Member, IEEE) received the bachelor's degree in mechanical engineering from the Ecole Polytechnique de Montreal and the M.S. and Ph.D. degrees in civil and environmental engineering from the University of Wisconsin-Madison. He is currently an Associate Professor in industrial and systems engineering with the Khalifa University of Science and Technology, Abu Dhabi, United Arab Emirates. He has authored more than 60 publications in peer-reviewed journals and refereed conference proceedings and editing two volumes on *Sustainability* and *Smart Cities in the Gulf*. His research interest includes enhance the performance of the built environment through occupant- building interactions.

PAPER 3

Annaqeeb M K, Azar E, Yan D, Novakovic V. Evaluating occupant perceptions of their presence and energy-use patterns in shared office spaces. *Proceedings of the 16th Conference of the International Society of Indoor Air Quality and Climate: Creative and Smart Solutions for Better Built Environments*, Indoor Air 2020.

The 16th Conference of the International Society of Indoor Air Quality & Climate
ONLINE | From November 1, 2020
Paper ID ABS-0967

Evaluating Occupant Perceptions of Their Presence and Energy-use Patterns in Shared Office Spaces

Masab K. Annaqeeb[1*], Elie Azar[2], Da Yan[3], and Vojislav Novakovic[1]

[1]Department of Energy and Process Engineering, Norwegian University of Science and Technology, Norway
[2]Department of Industrial and Systems Engineering, Khalifa University of Science and Technology, United Arab Emirates
[3]School of Architecture, Tsinghua University, China

[*] *Corresponding email: masab.k.annaqeeb@ntnu.no*

SUMMARY
The understanding of occupant behavior (OB) is vital to improve building energy performance, and this understanding needs an interdisciplinary approach that incorporates a holistic view of the topic. However, social dimensions of this behavior tend to be neglected. Based on the Theory of Planned Behavior, this study assesses the perceptions of occupants about their own presence and device utilization. This is done in the form of hourly schedules for presence, and energy-use patterns of plug loads in a shared office space, which were gained through a survey. The perceptions are then compared to the actual measurements, which were derived from monitoring these occupants with regards to their presence, environmental parameters, and plug loads for a period of six months. This enabled a comparative analysis that evaluates the correlations between perceived presence, actual presence, and standard schedules from ASHRAE. The results from this study can be used for developing social models for OB, by defining driving factors for different parameters of social influences.

KEYWORDS
Occupant behavior, Perceived Behavior, Energy-use Patterns, Occupant Presence.

1 INTRODUCTION
Occupant Behavior (OB) in buildings is a complex phenomenon, and has several facets, ranging from occupant comfort, presence, movement, habits, and the associated energy-use patterns of each. OB has also been identified as a driving factor in building energy performance, and the research in the last few decades has highlighted the importance of improving the understanding of OB, in order to enhance energy efficiency of buildings, along with occupant comfort (Hong et al., 2016; Yan et al., 2017). Consideration of OB has to be inclusive of all its aspects, and the accompanying models should reflect this comprehensiveness.

A significantly neglected part within this field of study is the lack of investigation of the social aspect of OB, comprising of their intentions behind the interactions with the building systems, the social influences, perceptions etc. A review of energy research literature over the past two decades highlighted the main negative patterns, the first of which was the underexploration of the social dimensions, especially those concerning energy-use (Sovacool et al., 2015).

The quantification of these dimensions often happens through the Theory of Planned Behavior, which considers three underlying factors behind behaviors and intentions: the beliefs and motivations regarding the action; peer influences (how significant others may respond to the action); and perceptions regarding the performance of the action. While some studies have investigated the first and second factor (Abrahamse & Steg, 2011; Chen & Knight, 2014), occupant perceptions remain an unexplored topic. An understanding of the occupant's own perception regarding their actions, and the associated effects on their behavior, is necessary to construct a complete profile of their social structures (Annaqeeb et al., 2019).

The main objective of the paper is to commence addressing these gaps, by assessing occupant perceptions regarding their own presence, and energy-use behaviors with regards to their device utilization. These perceptions are evaluated by drawing a comparative analysis with the actual occupancy and energy-use patterns, to measure the influence that these parameters have on OB.

2 METHODOLOGY

2.1 Area of Study

The study was part of a larger experiment aimed at monitoring and understanding occupant behavior in shared office spaces. It took place in a graduate research facility in Abu Dhabi, United Arab Emirates, and lasted for a duration of eight months, from April to November 2017. The office space consisted of eight individual desks (single occupancy), six of which were designated for specific individuals, and the other two were assigned as shared workstations. Since all the occupants taking part in the study were students at the research facility, the specific occupancy of each desk could change every semester. Over the course of the study, eight different occupants were included. Each desk was equipped with six power outlets, wherein each of the power outlet was assigned for a particular appliance (a desk lamp, two monitors, a docking station, a laptop, and miscellaneous), and the entire area was lit by six lighting fixtures. While the office was operational from 8am to 5pm on weekdays, it remained accessible to the staff at all hours throughout the week.

2.2 Data Collection

The data collection was conducted on two fronts: quantitative surveys and sensor-based measurements. The surveys were conducted to measure each occupant's perceptions about their own presence, energy-use patterns, and intents behind their actions. The sensor-based measurements recorded the actual presence, energy consumption, as well as the environmental parameters such as lighting status, temperature, and humidity.

For the surveys, it was vital to know the perceived schedule at every hour of the day, in order to have a comparison with standard schedules of shared offices from ASHRAE (American Society of Heating, Refrigeration, and Air-conditioning Engineers) (Abushakra et al., 2000). The occupants had to select an option from a scaled range of typical hourly occupancy patterns, ranging from 'never occupied' to 'always occupied'. This was done for both weekdays and weekends. In addition, their behavior with regards to unnecessary energy consumption was recorded for each device, where the occupant had another scale of options to denote the status of each appliance when the occupant left their workstation, ranging from ‘always switched off’ to ‘always switched on’. Other questions in the survey focused on the personal beliefs and energy-saving motivations of the occupant. As for the sensor-based measurements, three different kinds of sensors were included: PIR-based occupancy sensors for each desk, plug load sensors for each power outlet, and six environmental sensors placed at appropriate locations. Each of the sensors was installed and calibrated accordingly with their purpose. Routine checks were conducted to ensure the functionality of the sensors and the compliance of the participants with the experimental setup. A more detailed description of the experimental setup, calibration and data acquisition process is presented in this work (Das et al., 2020)

2.3 Data Processing and Analysis

The data from the occupancy sensors was event-based (recording data whenever the presence status changed), while the plug load and environmental sensors reported at a frequency of fifteen-minute intervals. Since the occupancy data was event-based, it lacked consistency in the temporal resolution. The first step in processing the data was the temporal alignment of the input from all the three sensors, followed by upsampling. This led to a consistent datasheet that consisted of four readings every hour for each sensor, conforming to a uniform temporal resolution of fifteen minutes. The consequent step was to address missing values, which were caused due to sensor malfunction in the early stages of the study. The nature of the study made in necessary to have consistent data for each occupant and their associated device, and weeks with missing values were excluded.

The first part of the data analysis involved a comparison of the standard ASHRAE schedules with the measurements from the sensors and the schedules obtained from the survey. In order to construct schedules from the sensor-based measurements, occupancy diversity factors were used. Diversity Factors denote the intensity of a variable throughout the day, ranging from 0 to 1. In this case, it is the ratio of occupants present in an area to the maximum possible occupancy, for every hour of the day. Equation (1) was used to determine the diversity factors, wherein *h* denotes the hour of the day. The value obtained expresses the ratio of sensors which had an ‘occupied’ status over the total sensors present in the area of study, averaged over the *N* days of the study, starting from the first day *d*. The occupancy diversity factor was determined for the entire office space.

$$Occupancy\ diversity\ factor_h = \frac{\sum_{d=1}^{N} \frac{\text{number of sensors in 'occupied' mode}}{Total\ number\ of\ sensors}}{N} \tag{1}$$

The data from the surveys was used to calculate the occupancy in a similar manner, by obtaining the diversity factors from the response of each occupant regarding their presence at a particular hour. The schedules obtained from surveys and ASHRAE were each subjected to a regression analysis with the actual schedules, to gauge their respective correlations.

The second part of the analysis focused on the perceived energy-use behaviors regarding the appliances available to the occupant. Each occupant expressed their perception about the manner in which they would leave their appliances when they left their workstations, whether they were switched off to avoid unnecessary energy consumption or left switched on. These perceptions were then compared to the actual energy consumption data, obtained from the plug load sensors. The ‘unnecessary consumption’ was determined by summing up the plug loads during periods of long absence, as denoted by the occupancy status of the workstation.

2.4 Privacy Handling

All the participants were involved on a voluntary basis and were provided with a consent form approved by the Human Research Ethics committee of the research facility, that detailed the data that was to be collected. In addition, to avoid recording of any personal details, each occupant was associated with an ID. No other personal details of the occupant were documented.

3 RESULTS

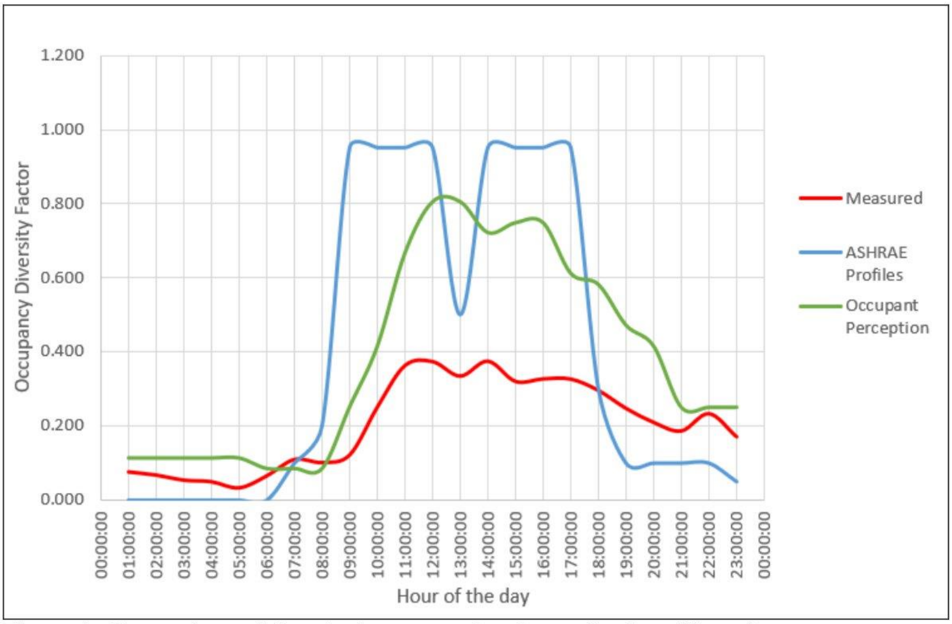

Figure 1. Comparison of Standard, measured and perceived profiles of occupancy

The occupancy schedules were plotted according to the hours and their corresponding average diversity factors. As seen in Figure 1, the actual schedules obtained from the sensors had significant differences from both the perceived and standard schedules. One of the biggest deviations was the delay in the start of the day, which is an important piece of information for the consideration of the building facility managers. However, the occupant perceptions about their presence, obtained from the surveys, had a smaller gap in this delay, and the schedules followed the trend of the actual ones more than the standards. These results were corroborated by a statistical analysis, the results of which are displayed in Table 1. The perceived occupancy shows a stronger correlation to the actual occupancy, and the higher R-square value signifies the dependency of the variable, and its predictability using the occupant's perceptions. The reported value of correlation is high, but it must be noted that the cause of this is the similarity in following the trend. Nevertheless, a significant gap still persists, as seen in the Figure 1.

The results for the device utilization, however, did not yield any significant correlations. Figure 2 displays the results for one of those devices (Laptop). The actual energy consumption is displayed with regards to the perceived energy-use. Occupant 1, 3, and 6 had all reported the device to be frequently switched on while leaving the office, while Occupant 4 and 5 always left it on, and Occupant 2 rarely left it switched on. However, this consistency between the reported and actual behavior was limited to very few devices. The perceived energy-use of the occupants varied greatly from occupant to occupant, as well as device to device. Even devices subjected to frequent use (such as laptops and monitors), showed discrepancies regarding the perceptions. For example, laptops were the most consistent, and monitors the least. Occupants that considered themselves to be switching off their monitors while leaving found that their perceptions were quite inaccurate, according to the measurements from the plug load sensors.

Table 1. Statistical correlations of standard and perceived schedules with measured occupancy

Statistical Analysis	Standard Occupancy Schedules	Perceived Occupancy Schedule
Correlation	0.76	**0.95**
R-squared (Linear Regression)	0.58	**0.91**
Covariance	0.04	**0.03**

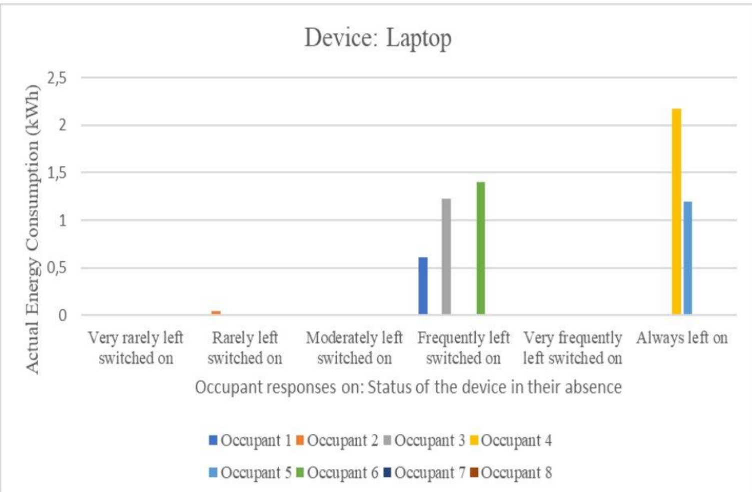

Figure 2 (a) Perceptions about device utilization: Laptops

4 DISCUSSION

It is important to note that due to the small sample size of both the occupants, and the building type, the objective of this study is not to draw conclusions regarding the occupancy schedules and energy-use patterns, but to highlight the importance of including social dimensions while trying to understand occupant behavior. Constructing an appropriate smart environment in office spaces is expensive, and several buildings do not have access to a technologically advanced Building Management Systems. In such cases, standard schedules are adopted to regulate the energy performance of the building. However, Post Occupancy Evaluation techniques can suit this task better, as occupant perceptions are more attuned to their presence patterns, which was also corroborated in the results of this study. Moreover, considering the gap between the perceived and actual energy consumption, the perception accuracies can still be improved further, by employing suitable feedback systems that keep occupants better informed about their energy-use behaviors in the office.

5 CONCLUSIONS

This work conducted an assessment of occupant perceptions regarding their own presence and energy-use patterns in shared office spaces. The data collection and consequent analysis showed the extent of influence perceived behaviors have on actual ones, and the respective correlation between the actual, standard, and perceived schedules. Results of the device utilization, however, showed no significant correlations in this regard. However, future works need to expand on these social dimensions by conducting such experiments at a much larger scale, and incorporate the occupants perceived behavior within the social modules of occupant behavior patterns.

6 ACKNOWLEDGEMENT

The authors are members of IEA-EBC Annex-79: Occupant-Centric Building Design and Operation, and this research work has been motivated, in part, by the activities within the Annex-79.

7 REFERENCES

Abrahamse, W., & Steg, L. (2011). Factors Related to Household Energy Use and Intention to Reduce It: The Role of Psychological and Socio-Demographic Variables. *Human Ecology Review*, *18*(1), 30–40. Retrieved from www.jstor.org/stable/24707684

Abushakra, B., Sreshthaputra, A., Haberl, J. S., David, P. E., & Claridge, E. (2000). *Compilation of Diversity Factors and Schedules for Energy and Cooling Load Calculations, ASHRAE Research Project 1093-RP, Final Report*. Energy Systems Laboratory (http://esl.tamu.edu), Texas A&M University.

Annaqeeb, M. K., Dziedzic, J. W., Yan, D., & Novakovic, V. (2019). Exploring the tools and methods to evaluate influence of social groups on individual occupant behavior with impact on energy use. *IOP Conference Series: Earth and Environmental Science*, *352*(1). https://doi.org/10.1088/1755-1315/352/1/012044

Chen, C. F., & Knight, K. (2014). Energy at work: Social psychological factors affecting energy conservation intentions within Chinese electric power companies. *Energy Research and Social Science*, *4*(C), 23–31. https://doi.org/10.1016/j.erss.2014.08.004

Das, A., Annaqeeb, M. K., Azar, E., Novakovic, V., & Kjærgaard, M. B. (2020). Occupant-centric miscellaneous electric loads prediction in buildings using state-of-the-art deep learning methods. *Applied Energy*, *269*, 115135. https://doi.org/10.1016/j.apenergy.2020.115135

Hong, T., Taylor-Lange, S. C., D'Oca, S., Yan, D., & Corgnati, S. P. (2016, March 15). Advances in research and applications of energy-related occupant behavior in buildings. *Energy and Buildings*, Vol. 116, pp. 694–702. https://doi.org/10.1016/j.enbuild.2015.11.052

Sovacool, B. K., Ryan, S. E., Stern, P. C., Janda, K., Rochlin, G., Spreng, D., … Lutzenhiser, L. (2015). Integrating social science in energy research. *Energy Research and Social Science*, *6*, 95–99. https://doi.org/10.1016/j.erss.2014.12.005

Yan, D., Hong, T., Dong, B., Mahdavi, A., D'Oca, S., Gaetani, I., & Feng, X. (2017). IEA EBC Annex 66: Definition and simulation of occupant behavior in buildings. *Energy and Buildings*, *156*, 258–270. https://doi.org/10.1016/j.enbuild.2017.09.084

PAPER 4

Annaqeeb M K, Das A, Arpan L, Novakovic V. Evaluating and Modeling Social Aspects of Occupant Behavior in Buildings: An Agent-Based Modeling Approach. To be submitted to: Building and Research Information

This paper is submitted for publication and is therefore not included.

PAPER 5

Annaqeeb M K, Dziedzic J W, Yan D, Novakovic V. Development of a Library for Building Surface Layout Simulator. *Proceedings of the 11th the International Symposium on Heating, Ventilation and Air Conditioning (ISHVAC).* 2019;1137-1144.

Development of a Library for Building Surface Layout Simulator

Masab K. Annaqeeb, Jakub W. Dziedzic, Da Yan and Vojislav Novakovic

Abstract Available building simulation tools resort to using fixed schedules for modeling occupant behavior (OB), which does not accurately capture its nature. A significant aspect of OB is the movement and sequence of actions with regards to their surroundings. This requires some coherence about the surface layout, including the placement of furniture and the occupant's interaction with it. There is a need for understanding vital information about the different attributes of the furniture, such as the placement and order of importance. Until now, there exists no such library with this kind of granularity in information. This paper explores the questions with regard to the development of such a library. This includes the description of the type of variables associated with different kinds of furniture, along with the occupant interaction under typical scenarios. The results from this study can be used to integrate the resulting library with building simulation tools and to better understand and develop occupant behavior models.

Keywords Building performance simulation · Occupant behavior · Data mining · Building energy management

M. K. Annaqeeb (✉) · J. W. Dziedzic · V. Novakovic
Department of Energy and Process Engineering, Norwegian University of Science and Technology, Trondheim, Norway
e-mail: masab.k.annaqeeb@ntnu.no

J. W. Dziedzic
e-mail: jakub.w.dziedzic@ntnu.no

V. Novakovic
e-mail: vojislav.novakovic@ntnu.no

D. Yan
School of Architecture, Tsinghua University, Beijing, China
e-mail: yanda@tsinghua.edu.cn

Z. Wang et al. (eds.), *Proceedings of the 11th International Symposium on Heating, Ventilation and Air Conditioning (ISHVAC 2019)*, Environmental Science and Engineering, https://doi.org/10.1007/978-981-13-9528-4_115

1 Introduction

Global energy trends have indicated building energy consumption to be emerging as one of the most energy-intensive sectors, with more than 55% of the global electricity usage belonging to commercial and residential buildings [1]. Recent reports estimate this sector being responsible for 39% and 40% of the nationwide energy consumption in the USA and EU countries, respectively [2]. Several efforts have been directed to increase energy efficiency in buildings in the form of incentives, certification programs, building codes, etc., but despite the advent of these efforts, buildings continue to show a large variation in their consumption patterns, with respect to the expected performance [3, 4].

This performance gap can be attributed to several different causes, pertaining to the mechanical and electrical faults within the buildings, to the weather and climatic variations, or architectural design [5]. However, the past decade has seen a lot of research efforts focus on one particular aspect: occupant behavior (OB). Through the efforts of several global researchers and scientists in the "Energy in Buildings and Communities (EBC)" program of the International Energy Agency (IEA), the way occupants interact with the building systems have been identified as one of the major drivers of a building's energy performance [6]. Clevenger et al. demonstrated that occupant behavior could vary the total energy use by 150% for commercial buildings [7]. Hong et al. created different occupant profiles and ran building energy models. The authors found that occupant behavior can save up to 50% of current energy use levels or increase them by 89% [8]. One particular approach to tackle this issue is using building performance simulation (BPS) programs to efficiently improve the design and operation of buildings. These programs include the modeling and evaluation of different systems in buildings, such as thermal or electrical, and are vital for drafting energy-saving recommendations [9]. Even so, BPS programs often lack accurate OB models, with most of the traditional BPS tools using fixed or pre-loaded schedules [10].

Within the field of OB modeling, the current simulation strategies can be broadly classified into two different groups. The first one comprises models that focus on the systems that the occupant is interacting with, rather than directly with the occupant. These would include linear regressions [11], sub-hourly occupancy-based control models [12], etc. The second group of models deals directly with the occupant and their actions, making use of agent-based models [13] and Markov chains [14]. However, the application of these models is often limited to one particular function (e.g., window opening and lighting control). In addition, their dependence on an aggregated model ignores the diversity and inhibits the accuracy in simulating the OB.

OB modeling has its complications based in the diverse set of actions as well as the different aspects of the OB itself. The complexity and uncertainty in this field stem from the fact that OB contains various similitudes in the form of presence, movement, activity level, comfort level, social influences, etc., and detailed attention has to be given to each of these in order to construct a complete individual profile. Dziedzic et al. proposed a bottom-up approach wherein the collected data from these different fields of simulations could be used to eventually develop a building occupant transient agent-based model (BOT-ABM) [15].

A large part of simulating OB is modeling the indoor movement and transition of the occupant. Markov chains were used by Wang et al. [16] wherein the movement process was simulated by associating each occupant with a homogeneous Markov matrix. A different form of data collection was used by Martani et al. where the Wi-Fi connections were used as proxy for the occupant sensing [17]. Similar to the occupant monitoring and data collection, this comes with its own set of privacy concerns. To overcome those, another study used a depth registration camera to track and monitor the movement and presence while maintaining a sufficient amount of privacy [18].

A complementary aspect in consideration with the modeling of movement is the simulation of the floor surface layout and the placement of objects/furniture around the occupant. In order to accomplish these simulations, the surface simulator will need access to a database or library of specific information regarding the furniture, as well as the details of the occupant's interaction with it. Current literature does not contain any particular specifications that can support a surface simulator with that kind of a database. This will have to include the information about the order of importance of the object for the occupant, their access points, area of influence, placement criteria for each, among others. The necessity of this information arises from the need to understand the boundaries and potential paths for the occupants' movement, as well as their order of actions with the objects around them. The next section describes the development of this database, definition and properties of each variable, and their necessity for the simulator.

2 Database Description

The library is intended to provide necessary information about the furniture and its placement to be used for a floor surface layout. The furniture would be the one typically used in residential buildings and will consist of different variables associated with each object, the information about which will be determined by collecting data from the occupants themselves. Each variable is selected based on its connection to the way the object influences or hinders the actions of the occupant with regard to their location and movement. The description of each of these is as follows:

- Furniture class: This variable contains the description regarding the type of furniture. The different classes will be procured from a compiled list of the typical furniture used in different rooms.
- Room category: Represents the type of room (bedroom, living room, etc.). This variable would further influence the order of importance, since the objects having the same furniture class can have different significance depending on the room category. For instance, the order would be different for a table in a study (where it might be prioritized higher) and in a bedroom (where its importance will be relatively lower).

- Order of importance: This variable describes the order of the object's importance to the occupant based on the frequency of use within a particular room category. It also indicates the rank of this furniture when it comes to the placement. The simulator will be using the allotted ranks to generate the sequence of each simulated object within the room. This will have to work in accordance with the next variable, wherein objects with higher rank will be prioritized and placed according to their placement criteria and those criteria will be re-evaluated for the next object, without disturbing the placement of the preceding object.
- Placement criteria: This will ascertain the typical factors that influence the occupant while placing the furniture with reference to the distance from the corners, edges, doors, windows, etc. The information will be significant in generating a surface layout based on the floor map of the rooms and will further ease modeling the path of movement that the occupant will be taking. It will also enhance understanding about the intent behind the occupant's actions, and these preferences can also be used as input for building habit profiles for the occupant.
- Area of influence: This variable reflects what kind of constraints and influences the particular furniture creates for the occupants' movement around them. It also constitutes how it affects the placement of the other objects. Along with hinge points, this serves as a decisive factor for the path simulation.
- Access points: The position respective to the object through which the occupant would be interacting with a particular furniture class.
- Hinge points: These points would indicate the corners or edges of the furniture and will form the basis for movement simulation, as the distances from different hinge points will reflect a range of the potential path the occupant could take.

Table 1 consists of the seven different but interconnected variables for the database and is meant to showcase how the library is structured. Variables like these will be necessary to act as trigger points for further actions, and as specified for some of them, the use of them goes beyond floor surface simulation. One instance of this is the placement of outlets for the HVAC design. Better surface layout simulations can help adjust the outlet placement in accordance with the occupant's thermal comfort needs and their surroundings.

Table 1 Variables to be used in the database

Furniture class	Room category	Order of importance	Placement criteria	Area of influence	Access points	Hinge points
Class 1	1		Criteria 1			
	2		Criteria 2			
	3		Criteria 3			
Class 2	1		Criteria 1			
	2		Criteria 2			
	3		Criteria 3			

3 Preliminary Survey

The survey to gather the necessary input from occupants to create an adequate database was designated to be set in an online format. The main factor responsible for the design format was the need for different possible combinations and scenarios with regard to the occupants' preferences. The database needed a large number of layout preferences while keeping the models realistic by using different constraints. It would be extremely difficult to accomplish this using the traditional experimental data collection platforms. Fortunately, the advent of massive online experiments (MOEs) has made possible conducting studies with large-scale participation, an exponential combination of different variables, while retaining the control on the experimenter's side. These web-based experiments provide specific advantages over the lab-based ones in terms of collecting larger sample sets at a much lower cost.

Additionally, the main requirement of this study was to have sufficient features that enable the occupant to assemble their own layout and provide feedback regarding that preference, which is possible through these MOEs. The development platform chosen for this was Meteor, because of the template-based structure it offered, in addition to its useful packages and dynamic scripting.

However, a preliminary survey was conducted to test out the feasibility of the concept and gather feedback from the occupants. The survey was designed to investigate the home space usage of kitchens in residential buildings. The scope was to seek out information regarding the different kitchen appliances and how they are placed around by the occupants. This was done to generate a sufficient database for a kitchen layout simulator. It explored the number of occupants and their demographics. These occupants were then provided with a questionnaire along with a list of typical appliances found in a kitchen, with the option to add any that were not present in the compilation.

The room category in this case would remain fixed, as the study was still in its preliminary stages and would not extend beyond the kitchen. There were in total 18 furniture classes in the compiled list provided for the participants to choose from (Fig. 1). The occupants had to denote the presence of the appliance and provide the order of importance for each. In addition, they were also asked to mark down the position of windows, doors, orientation, and the approximate shape of the kitchen. In order to have better insights, they were asked to mark down this information on a grid, as shown in Fig. 2.

The collected information would then be used to generate appropriate hinge points to determine the movement path. Other variables such as the area of influences and placement criteria were not added at that stage. They were to be approximated based on the layout given but later added in the main questionnaire to remove the need for any assumptions. It was due to the ease of incorporating the feature of having a base layout that can be modified by the participants that the choice was made to shift the process online.

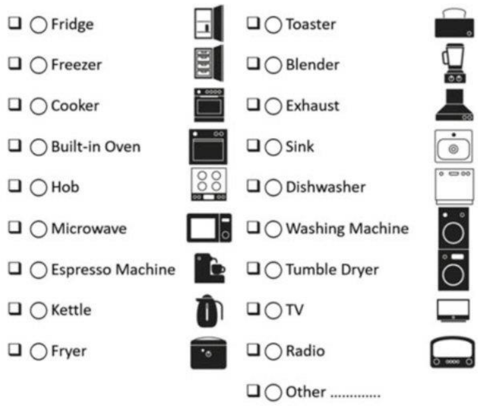

Fig. 1 Compilation of furniture class

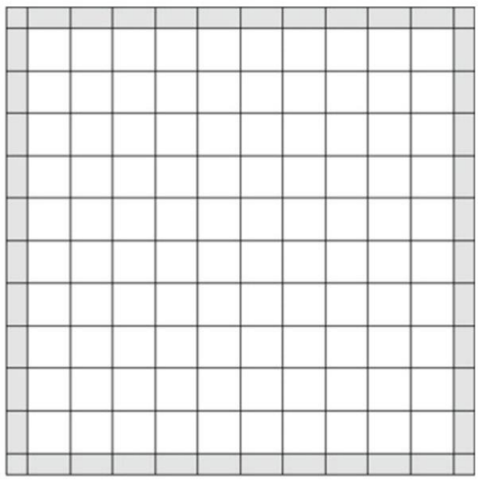

Fig. 2 Grid to record layout and preferences

4 Discussion and Conclusion

As seen from the literature review, there is still a significant potential for improving building performance simulations through OB modeling. Considering the diverse aspects of OB, this study supports the bottom-up approach and highlights the need to consider a more dynamic process for delivering the surface layout simulation. It should be noted that the primary purpose of this database is to provide a library for the layout simulation. Furthermore, the layout simulator is intended to act as a building component for an eventual BOT-ABM.

As mentioned previously, some of the variables can prove useful for other purposes, such as HVAC design. However, that is beyond the scope of this particular study. Future research can be directed toward incorporating this kind of a database and surface simulator in traditional BPS programs.

References

1. IEA: World Energy Statistics, International Energy Agency, pp. 32–51 (2017)
2. U.S. Energy Information Administration (EIA)—Total Energy Annual Data (2018)
3. Azar, E., Menassa, C.C.: A comprehensive analysis of the impact of occupancy parameters in energy simulation of office buildings. Energy Build. **55**, 841–853 (2012)
4. Menezes, A.C., Cripps, A., Bouchlaghem, D., Buswell, R.: Predicted vs. actual energy performance of non-domestic buildings: using post-occupancy evaluation data to reduce the performance gap. Appl. Energy **97**, 355–364 (2012)
5. Turner, C., Frankel, M.: Energy performance of LEED for new construction buildings. New Build. Inst. **4**, 1–42 (2008)
6. Yan, D., Hong, T., Dong, B., Mahdavi, A., D'Oca, S., Gaetani, I., Feng, X.: IEA EBC Annex 66: definition and simulation of occupant behavior in buildings. Energy Build. **156**, 258–270 (2017)
7. Clevenger, C.M., Haymaker, J.R., Jalili, M.: Demonstrating the impact of the occupant on building performance. J. Comput. Civ. Eng. **28**(1), 99–102 (2013)
8. Hong, T.: Occupant behavior: impact on energy use of private offices. In: ASim 2012-1st Asia conference of International Building Performance Simulation Association, Shanghai, China, 25 Nov 2012–27 Nov 12 (2014)
9. Yan, D., O'Brien, W., Hong, T., Feng, X., Gunay, H.B., Tahmasebi, F., Mahdavi, A.: Occupant behavior modeling for building performance simulation: current state and future challenges. Energy Build. **107**, 264–278 (2015)
10. Hoes, P., Hensen, J.L.M., Loomans, M.G.L.C., De Vries, B., Bourgeois, D.: User behavior in whole building simulation. Energy Build. **41**(3), 295–302 (2009)
11. Zhao, J., Lasternas, B., Lam, K.P., Yun, R., Loftness, V.: Occupant behavior and schedule modeling for building energy simulation through office appliance power consumption data mining. Energy Build. **82**, 341–355 (2014)
12. Bourgeois, D., Reinhart, C., Macdonald, I.: Adding advanced behavioural models in whole building energy simulation: a study on the total energy impact of manual and automated lighting control. Energy Build. **38**(7), 814–823 (2006)
13. Aerts, D., Minnen, J., Glorieux, I., Wouters, I., Descamps, F.: A method for the identification and modelling of realistic domestic occupancy sequences for building energy demand simulations and peer comparison. Build. Environ. **75**, 67–78 (2014)
14. Lee, Y.S., Malkawi, A.M.: Simulating multiple occupant behaviors in buildings: an agent-based modeling approach. Energy Build. **69**, 407–416 (2014)
15. Dziedzic, J., Yan, D., Novakovic, V.: Occupant migration monitoring in residential buildings with the use of a depth registration camera. Procedia Eng. **205**, 1193–1200 (2017)
16. Wang, C., Yan, D., Jiang, Y.: A novel approach for building occupancy simulation. In: Building Simulation, pp. 149–167 (2011)

17. Martani, C., Lee, D., Robinson, P., Britter, R., Ratti, C.: ENERNET: studying the dynamic relationship between building occupancy and energy consumption. Energy Build. **47**, 584–591 (2012)
18. Dziedzic, J.W., Da, Y., Novakovic, V.: Indoor occupant behaviour monitoring with the use of a depth registration camera. Build. Environ. **148**, 44–54 (2019)

PAPER 6

Annaqeeb M K, Zhang Y, Dziedzic J W, Xue K, Pedersen C, Stenstad L I, Novakovic V, Cao G. Influence of surgical team activity on airborne bacterial distribution in the operating room with a mixing ventilation system: a case study at St. Olavs Hospital. *Journal of Hospital Infection*. 2021; 116:91-98.

Journal of Hospital Infection 116 (2021) 91–98

Available online at www.sciencedirect.com

Journal of Hospital Infection

journal homepage: www.elsevier.com/locate/jhin

Influence of surgical team activity on airborne bacterial distribution in the operating room with a mixing ventilation system: a case study at St. Olavs Hospital

M.K. Annaqeeb [a], Y. Zhang [b,*], J.W. Dziedzic [a], K. Xue [c], C. Pedersen [d], L.I. Stenstad [e], V. Novakovic [a], G. Cao [a]

[a] *Department of Energy and Process Engineering, Norwegian University of Science and Technology, Trondheim, Norway*
[b] *College of Civil Engineering and Architecture, Hainan University, Haikou, China*
[c] *School of Civil Engineering, Chongqing University, Chongqing, China*
[d] *MultiConsult Norge AS, Seksjon VVS Tromsø, Norway*
[e] *St. Olavs Hospital, Operating Room of the Future, Trondheim, Norway*

ARTICLE INFO

Article history:
Received 28 March 2021
Accepted 9 August 2021
Available online 14 August 2021

Keywords:
Surgical site infection
Hospital operating room
Hospital-associated infection
Human activity

SUMMARY

Background: Operating rooms (ORs) have strict requirements regarding cleanliness. While existing standards concerning the ventilation and staff guidelines are theoretically sufficient to subvert the threats posed by micro-organisms within the room, there exist potential sources of contamination due to human activity around the area. Studies exploring this influence of human activity on distribution of micro-organism contamination in ORs have relied on manual observations, or indirect methods such as number of door openings.
Aim: To utilize depth registration sensing technology to identify the activities of surgical staff and investigate their effect on the distribution of airborne micro-organism contamination in ORs.
Methods: A mock surgical experiment was performed using a depth registration technique for the dynamic capturing of human presence and activity levels. Field measurements were carried out in one real OR to analyse its influence on the bacterial distribution in ORs with mixing ventilation system. *Findings*: Bacterial contamination levels tended to correlate with higher activity levels, albeit with some inconsistencies. The highest activity levels were around the surgical bed when the patient was placed, and around the instrument table during the surgical procedure. Locations with obstructions had the highest cfu densities, indicating that airflow patterns are important in such spaces.
Conclusion: Our activity monitoring methods demonstrate a novel means of studying the influences of human activities in hospital rooms.

* Corresponding author. Address: College of Civil Engineering and Architecture, Hainan University, Haikou, 570228, China.
E-mail address: zhangyixian902@163.com (Y. Zhang).

https://doi.org/10.1016/j.jhin.2021.08.009

Introduction

Operating rooms (ORs) in hospitals have to uphold the highest standards of cleanliness [1]. Introduction of bacterial contamination into the surgical wound, through direct airborne transmission or indirectly, e.g., through airborne contamination of instruments that then enter the wound, can cause infection [1,2]. Theoretically, the installed ventilation system is sufficient to protect the surgical zone from any potential sources of contamination, with staff inside the surgical zone having limited contact with those outside the zone. However, in practice, the dynamic of the environment may change during the surgery, especially if there is poor compliance with OR discipline. Even when OR personnel are fully compliant, the transitions and movement inside the whole room can still be a source of potential contamination [3].

In modern hospitals, the most important source of airborne contamination is related to the dispersal of particles from persons present in the OR and their movements [4,5]. A high volume of staff movement and activities could play a role in higher risk of surgical site infections (SSIs) [6]. Generally, staff movement can increase the colony forming unit (cfu) level by three means: (1) clothes rub against the skin, leading to increased shedding [7]; (2) a pumping effect inside clothes that creates air streams that can transport skin scales into the OR air through pores in the fabric structure or from openings (such as the wrists, neck, ankles and the waist) [8]; and (3) movement may cause settled particles on the floor and other surfaces to be re-suspended into the air [7].

A number of studies have explored influencing factors using various experimental methods. For example, You *et al.* utilized the emission rates of particles in a sealed chamber [9], whereas Scaltriti *et al.* focused on a recently built operating theatre using measurement of microbiological and dust contamination to assess the influence of human activity [3]. Andersson *et al.* investigated 24 orthopaedic operations in Sweden and concluded that different activity intensities highly influence the cfu level [10]. However, these studies investigated the influence of human activity by relying on air-quality data, correlating those with the occupancy, and the type of clothing worn, or the traffic flow within the operating room. In some cases, manual observations and door openings served as the basis for this comparison [10]. The occupants in the room were not monitored individually, regarding their locations or movement trajectories. Developments within the field of occupant monitoring in indoor spaces has allowed for much more sophisticated techniques of data acquisition [11], which have not yet been widely utilized for experimental procedures in operating rooms.

The objective of this study was to use of depth registration sensing as a tool to investigate the effect of surgical activities on the distribution of airborne micro-organism contamination in ORs.

Methods

Three mock-up surgeries were performed with a controllable series of actions. The main monitoring parameters were dynamic registration of staff movement and passive air sampling, using downward displacement of air and enumeration of bacterial cfu, where these settle on to the surface of exposed agar plates, at floor height. Bacterial contamination was established by cumulatively exposing agar plates positioned inside the surgical room and the surgical staff were monitored with the use of a depth registration camera. This method made it possible to assess the influence of medical personnel's activity, especially their movement, on the potential contamination.

Experimental setup

Three mock surgeries were performed in the cardiopulmonary OR with a mixing ventilation system in St. Olavs Hospital, Trondheim, Norway. There were six participants in the study, one of whom represented the patient and the rest carried out the roles of five staff members (main surgeon, assistant surgeon, sterile nurse, distribution nurse and anaesthetic nurse). The layout of the room and the respective positions of each participant is shown in Figure 1. A total of 24 passive agar plates (internal diameter 85 mm) were used in six locations (A–F), placed around the surgical bed (Figure 1).

The experiment consisted of four different phases, and each agar plate group consisted of four agar plates. Each set of the agar plates was opened at the start of the different phases and kept open until the end of the experiment. 'Group' refers to the four agar plates present at each of the six locations, and 'set' refers to the plates from each group that would be opened at the start of specific phases. The first set of agar plates were open from Phases 0 to 3 (140 min), the second set were opened from Phases 1 to 3 (120 min), and so on (Figure 2). The durations of the phases were based on the timings of a typical hip arthroplasty procedure. The sampling windows for the agar plate groups is depicted in further detail in Figure 2. All of the locations (A–F) followed the same pattern. Agar plates incubated for 48 h at 35 $\pm$ 2°C, followed by 24 h at room temperature, before counting.

The clothing used by the staff was provided by and adhered to the standards of the hospital. The staff wore an EN 13796-compliant two-piece disposable non-woven polypropylene suit. The main surgeon, assistant surgeon and sterile nurses wore surgical gowns outside the two-piece disposable non-woven suit, which were made of nonwoven polyester/polyethylene and approved according to the EN13795: 2011 standard. The surgical masks worn by the staff were approved according to EN 14683 type II, which were the double band, tie-on type with an integrated adjustable nose clamp. Latex gloves, surgical caps and hoods were also used to cover hands and the exposed parts of face. The patient wore a two-piece disposable non-woven suit and a surgical cap.

Monitoring and processing movement of participants

Participants were monitored using Microsoft Kinect Xbox devices, which use depth registration to capture geometric information about human activity [12]. The principle of depth registration is similar to a traditional projector, but visible light is replaced by an infrared beam. An additional sensor measures the time it takes for this incident ray to be reflected back to the camera, and this is used to gauge the distance of nearby objects. These measurements are recorded in a matrix array, which makes it possible to recreate a three-dimensional representation of the observed surface. If the human body is present within its field of view (FoV), it gets processed with a

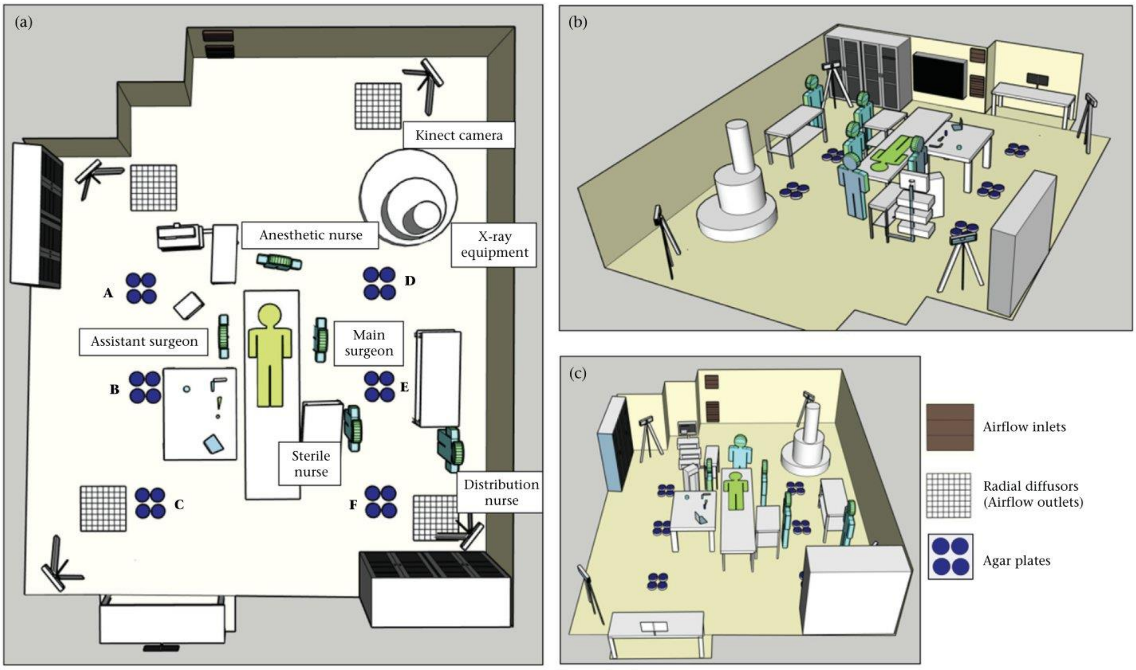

Figure 1. (a) Layout of the mock surgery experimental setup (top view). (b) and (c) Bird's angle view of different perspectives of the set-up.

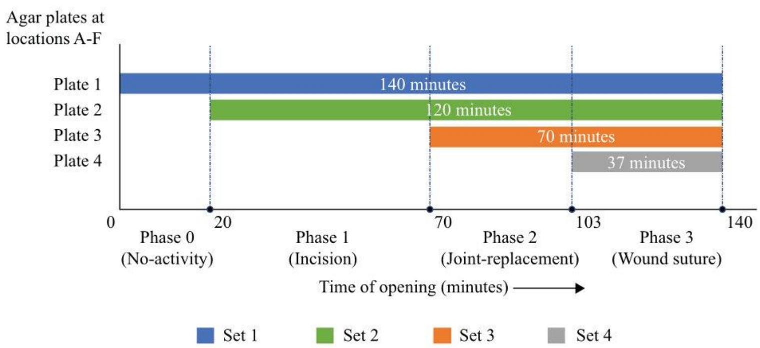

Figure 2. Sampling windows of the agar plates.

skeleton model (SM), with 25 joints representing the human body in a three-dimensional matrix. For this trial, four cameras were used, each having a registration angle of 46° in a 5-m radius, and a 30-Hz sampling rate. To cover the OR entirely and avoid occlusion, each device was placed in a corner of the room (Figure 1). Because all the devices have a capability of capturing up to six persons simultaneously, there was no possibility of misregistering any participant's activity.

Data processing consisted of unification of planes, data fusion, and plotting an activity heat map. Each device captured information according to its own local coordinate system. This had to be synchronized by selecting one device as referential and changing its horizontal plane to be parallel to the floor. The data from other devices were then accordingly adjusted and combined to create the final map. Based on previous studies [13], another significant component to validate the synchronization of the data is by using the most stable joints from the SM. The plotted map has a resolution of 5 $\times$5 cm. Each time the movement readings indicate a position in a particular cell of the heat map, it increases the value of this cell by adding one. After all the movement recordings are processed in that manner, a developed figure can show a spatial activity distribution.

Table I
Movement and action plan for the surgical team

Activity	Main surgeon	Assistant surgeon	Sterile nurse	Distribution nurse	Anesthetic nurse
1	Body and arm movement: towards and away from sterile nurse (10 times)	Arm movement: continuous circular motions with one hand close to the wound (1 min)	Arm movement: towards and away from main surgeon (10 times)	Body and arm movement: Towards and away the sterile nurse (10 times)	Sit still (1 min)
2	Hand and arm movement: continuous random finger motions close to the wound (1 min)	Body and arm movement: Side to side (10 times)	Body and arm movement: towards and away from distribution nurse (10 times)	Body movement: walking to a cabinet and back again	NA (no activity)
3	Arm movement: fast up and down movement of arm (10 times)	Hand movement: holding a steady hand close to the wound (1 min)	Arm movement: shaking arms continuously (1 min)	Arm movement: twisting of hands in front of the chest (1 min)	NA
4	Arm movement: shaking arms close to the wound continuously (1 min)	Body movement: squatting – three times	NA	NA	NA
5	Resting position (1 min)	Resting position (1 min)	Resting position (1 min)	Resting position (1 min)	NA

Static persons do not register any readings. With such filtering, it was possible to obtain a heat map of the staff activity in the OR during the experiments. This heat map represents the concentration of the activities within a particular cell.

Mock surgery procedure

Based on a typical hip arthroplasty procedure, a definite movement and action plan was formulated for each participant

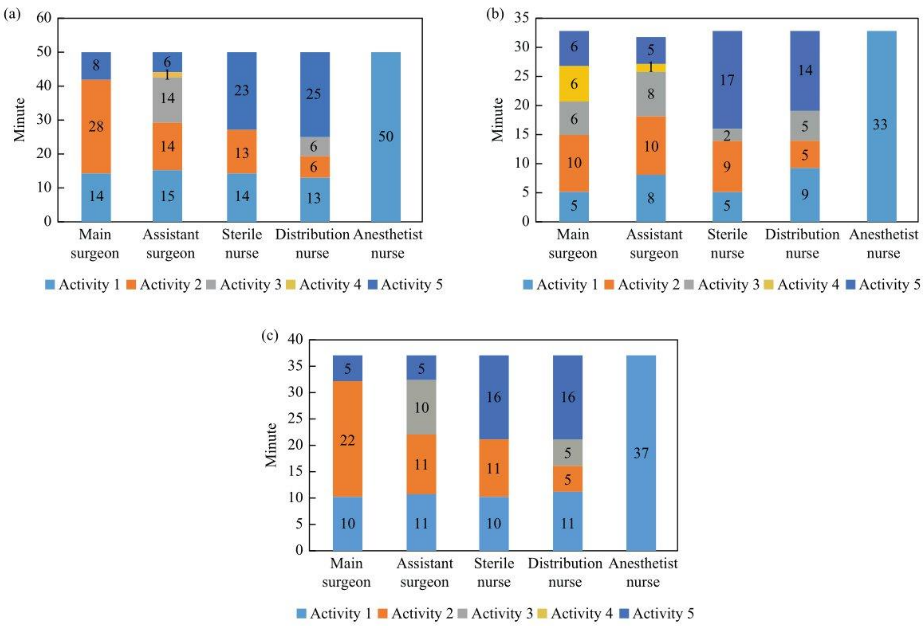

Figure 3. Distribution of activities within each phase, for each staff member.

Table II
cfu Counts of three experiments

Agar plate location	Experiment phase	Phases 0–3	Phases 1–3	Phase 2–3	Phase 3
	Experiment no.	cfu			
A	Experiment 1	7	2	2	3
	Experiment 2	2	1	1	1
	Experiment 3	3	6	3	1
B	Experiment 1	5	3	6	1
	Experiment 2	2	5	3	0
	Experiment 3	8	8	8	2
C	Experiment 1	1	0	1	0
	Experiment 2	2	2	1	0
	Experiment 3	3	4	5	1
D	Experiment 1	1	1	0	1
	Experiment 2	1	2	0	0
	Experiment 3	0	0	2	2
E	Experiment 1	3	2	2	1
	Experiment 2	3	5	2	1
	Experiment 3	6	4	4	2
F	Experiment 1	3	2	2	1
	Experiment 2	4	2	0	0
	Experiment 3	1	1	1	0

during the mock surgeries in the experiment. The mock surgeries (repeat thrice) were divided into three main phases: phase 1 – incision (50 min); phase 2 – joint replacement (33 min), phase 3 – wound suture (37 min). In addition, 20 min of non-activity and non-speaking phase (Phase 0) was added before the start of the mock-surgery to establish the difference in cfu levels of activity compared with non-activity of the surgical team.

A detailed description of the activities performed by each participant during the mock surgeries is given in Table I and Figure 3. In addition to considering the body movement, the participants were also required to speak (by reciting the alphabet out loud every 7th minute), because speech can disseminate respiratory tract bacteria including important pathogens such as *Staphylococcus aureus* [14].

Because the measuring time is highly dependent on the type of surgery, specifically the operating time, it is hard to compare the results with other different cfu standards by directly listing the cfu counts captured on the agar plates. Therefore, the measured data were normalized by transferring them to cfu density, which is formulized as cfu counts/agar plate area per measuring time (cfu/m^2 per h).

Results

Measured cfu levels with different activities

The cfu counts of three experiments are shown in Table II. The observed amount of cfu counts were limited, with the maximum count being eight. Locations A, E and F showed higher average counts with the first set of plates (open from Phases 0–3), while the location B and C showed the highest counts in the third set of plates (open during Phase 2–3).

The mean values for the normalized cfu density for each agar plate during different phases of the mock surgery are illustrated in Table III. The cfu/m^2 per h at locations B, E, A were higher than those at location C, F and D (Figure 4). The highest cfu/m^2 per h was measured in location B during Phases 2 and 3. The highest value was observed as 1208.07. Very high values were observed in other repeated experiments, with 906.05 in the first experiment and 453.03 in the second experiment. The lowest cfu/m^2 per h was measured several times at locations C, D and F.

Mapping of human activities

The human activity from the depth registration measurements was in the form of a spatial activity distribution map (Figure 4).

Discussion

Considering the activity levels associated with the cfu measurements in each location, it is clear that activity levels are not the only influencing factor on the distribution of airborne micro-organism contamination. However, the area around location E had consistently high activity levels in different phases, a trend that was also reflected in the cfu measurements. It might also be noteworthy that the Distribution Nurse was stationed near that area, and was the only one of the surgical staff that had to move across the room (Activity 2: walking to the cabinet and back), which would introduce particles from other areas of the room into the area near E, thus resulting in consistently higher cfu levels.

Table III
cfu Measurements and their corresponding activity levels

Experiment phase	Phases 0–3		Phases 1–3		Phase 2–3		Phase 3	
Agar plate location	Activity Level	cfu/m^2 per h	Activity Level	cfu/m^2 per h	Activity Level	cfu/m^2 per h	Activity Level	cfu/m^2 per h
A	12.7%	302.15	20.9%	264.34	18 %	302.02	24.7%	**476.03**
B	0.5%	**377.68**	8 %	**469.94**	10.3%	**855.72**	10 %	285.62
C	1.3%	151.07	0.9%	176.23	0.5%	352.35	3.3%	95.21
D	23.9%	50.36	**37 %**	88.11	**36 %**	100.67	**30.1%**	285.62
E	**60 %**	302.15	31.5%	323.08	34 %	402.69	28.9%	380.83
F	1.6%	201.43	1.7%	146.86	1.2%	151.01	3 %	95.21

In bold: The highest recorded values of activity level and CFU/m^2 per h in each phase.

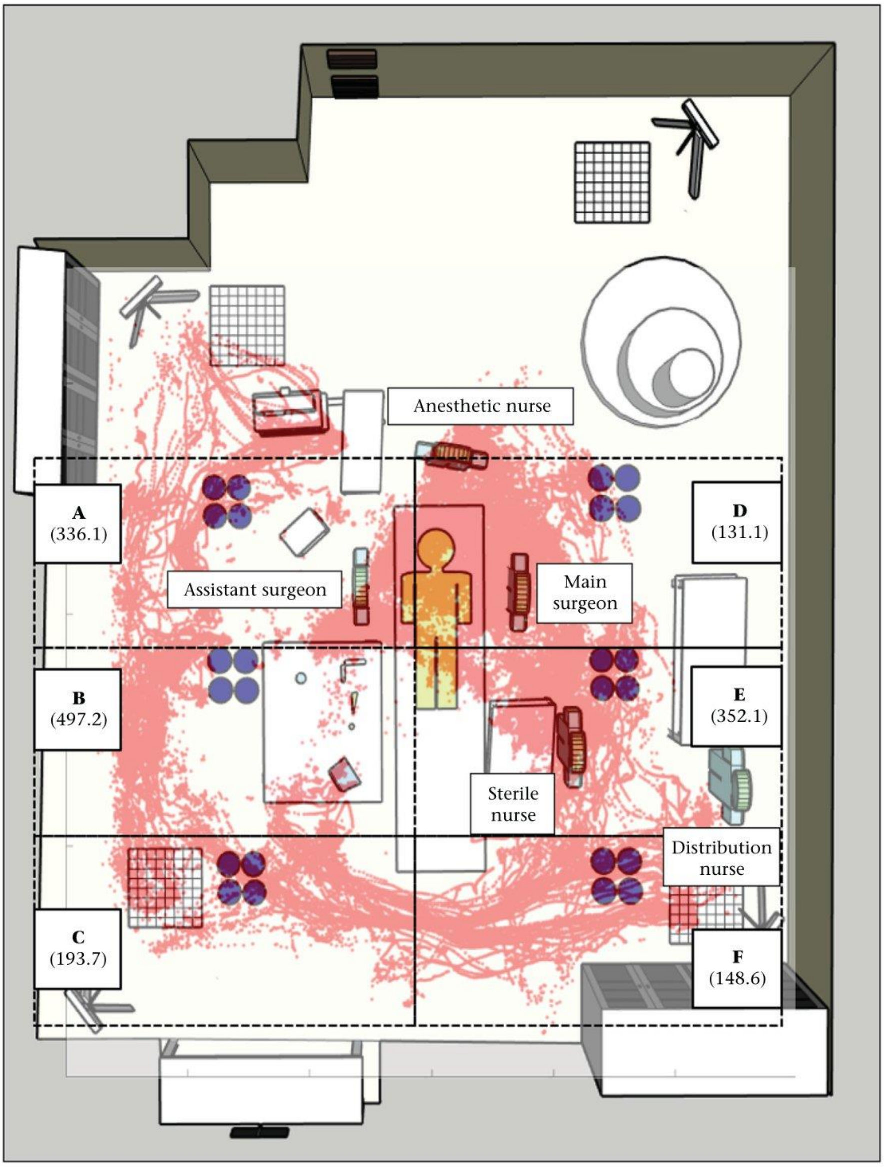

Figure 4. Activity levels around the surgical site. Each pink dot represents a single recorded activity. The average cfu measurements for each test location are also indicated.

Measurements in locations A, B and E reported higher cfu densities than C, D and F, which was consistent with the surgical staff within the respective region (main surgeon, assistant surgeon, sterile nurse) being more active than the anaesthetist nurse. It can also be seen that there is a strong body movement (Activity 4: Squatting — three times) of the assistant surgent in Phase 1 and Phase 2, which agreed well with high significance of cfu density for Phases 0—3, Phases 1—3 and Phase 2—3 in location B. In addition, the assistant surgeon during Phase 2 was responsible for opening the lid of the agar plates at locations A, B and C. And the agar plates at location B were placed closer to the assistant surgeon. These reasons might be the cause of having extremely high cfu densities in location B. Another possible explanation for the high cfu densities in agar plate location B is the obstruction caused by the amount of objects in that area. Most of the area around B is occupied by a large table, thereby obstructing the flow of fresh filtered air to that region. As such, measured cfu results may be influenced by both activity levels and airflow patterns within the OR.

However, there were unexpected findings. For example, it was expected that the cfu in location D would be high due to its close distance to the main surgeon and the sterile nurse. In addition, there was a walking movement by the distribution nurse with a path bypassing location D. It may be that different results would be obtained with a more sensitive bacterial sampling technique.

A significant limitation during this study was the overestimation of the cfu counts in the experiments, especially those measured at the beginning of the experiment with no activity. While designing the experiment, it was expected that these cfu counts would be highly correlated to the activity level at a location, and it can be seen from the results that this was not entirely the case. Furthermore, the conducted surgery was a controlled imitation, which will undoubtedly contain significant differences from a real surgery.

Several studies have examined the efficiency of both active and passive air sampling, with variable results. Napoli *et al.* compared the results from active and passive sampling in 32 ORs with turbulent flows and concluded that both methods are applicable for monitoring of air contamination. However, it is pointed out that passive sampling, as we used, is more suitable in studies designed to monitor the risk of microbial wound contamination, whereas active sampling is more suitable for investigating the concentration of all inhalable particles [15]. In similar comparisons, other studies recommended the use of passive sampling for evaluating airborne risk of contamination, due to its relevance, simplicity and economy of use [16–18]. However, this technique also has its own limitations. Firstly, it may collect more relatively larger particles that are settled down mainly due to gravity, which has more influence than the indoor turbulent air. Secondly, ORs have surgical tables, X-ray equipment, and other facilities that pose as physical obstructions to the airflow and can have an impact on the cfu measurements. As mentioned previously, the second effect can be a significant factor in this study, considering the measured high cfu concentration around location B.

In order to overcome these limitations, future studies might include both active and passive techniques in order to provide a better comparison. The active sampler can also be programmed to be activated remotely, which would remove the need to physically interact with it, thereby minimizing additional contamination during measurements.

Because the vast majority (80–90%) of postsurgical contaminants in wounds have their sources in ORs [19], it is imperative to discuss the significance of the ventilation system present in the study. The mock surgeries were performed in an OR with turbulent mixing ventilation (MV). Typically, the main types of ventilation systems employed to reduce the airborne bacterial load are laminar airflow (LAF) and MV. While many national standards considered LAF to be superior to MV in reducing the bacterial load, many recent studies have contested this claim [20,21]. A systematic review of studies from 1990 to 2016 showed that LAF did not reduce the risk of SSIs in comparison with MV [20], which became the basis for new World Health Organization (WHO) guidelines recommending against employing LAF systems after total joint arthroplasty [22]. More recently however, these guidelines were again contested by several studies [23–25], providing the evidence that showed otherwise. In light of these studies, it is even more important to have a detailed understanding of the influence that movement and activity can have on SSI. Taking these differences into account, future expansions of this study could include experiments performed in different ventilation systems to obtain a more detailed insight into the correlation between activity, airborne contamination and risk of SSI.

In conclusion, this study was designed to address the lack of dynamic capturing of human activity in experiments conducted in hospital rooms, which can be vital in investigating the standards of cleanliness, and other influences that human activity might have in such spaces. The results highlight trends of bacterial contamination in different locations. In general, higher activity levels correlated with higher cfu densities, but we also observed that locations near physical obstructions had the highest cfu densities, suggesting that airflow patterns might play a role in such spaces. Results from studies such as this might be used for implementing infection control practices concerning the staff activity and positioning of surgical instruments to optimize the airflow within the OR, because the current indoor environment design does not take into account the effect of human activities during real surgical procedures. We believe that dynamic recording of human activity, together with reproducible techniques to measure airborne contamination and airflow patterns can provide valuable information that could change operating theatre design and/or working practices.

Acknowledgements

The authors are grateful for the support provided by the Department of Energy and Process Engineering at the Norwegian University of Science and Technology (NTNU). We greatly appreciate the collaboration with the Operating Room of The Future (FOR) – St. Olavs hospital.

Funding sources
None.

Conflict of interest statement
None declared.

References

[1] Zamuner N. Operating room environment with turbulent airflow. ASHRAE Trans 1986;92:343–9. 2(A).
[2] Robson MC. Wound infection: a failure of wound healing caused by an imbalance of bacteria. Surg Clin North Am 1997;77(3):637–50.
[3] Scaltriti S, Cencetti S, Rovesti S, Marchesi I, Bargellini A, Borella P. Risk factors for particulate and microbial contamination of air in operating theatres. J Hosp Infect 2007;66(4):320–6.
[4] Whyte W, Hodgson R, Tinkler J. The importance of airborne bacterial contamination of wounds. J Hosp Infect 1982;3 (2):123–35.
[5] Tammelin A, Domicel P, Hambræus A, Ståhle E. Dispersal of methicillin-resistant Staphylococcus epidermidis by staff in an operating suite for thoracic and cardiovascular surgery: relation to skin carriage and clothing. J Hosp Infect 2000;44(2):119–26.
[6] Sadrizadeh S, Tammelin A, Ekolind P, Holmberg S. Influence of staff number and internal constellation on surgical site infection in an operating room. Particuology 2014;13(1):42–51.
[7] Bhangar S, Adams RI, Pasut W, Huffman JA, Arens EA, Taylor JW, et al. Chamber bioaerosol study: human emissions of size-resolved fluorescent biological aerosol particles. Indoor Air 2016;26(2):193–206.

[8] Sanzen L, Walder M, Carlsson S. Air contamination during total hip arthroplasty in an ultraclean air enclosure using different types of staff clothing. J Arthroplasty 1990;5(2):127—30.
[9] You R, Cui W, Chen C, Zhao B. Measuring the short-term emission rates of particles in the 'personal cloud' with different clothes and activity intensities in a sealed chamber. Aerosol Air Qual Res 2013;13(3):911—21.
[10] Andersson AE, Bergh I, Karlsson J, Eriksson BI, Nilsson K. Traffic flow in the operating room: an explorative and descriptive study on air quality during orthopedic trauma implant surgery. Am. J. Infect. Control 2012;40(8):750—5.
[11] Dong B, Kjaergaard MB, De Simone M, Gunay HB, O'Brien W, Mora D, et al. Sensing and data acquisition. In: Exploring occupant behavior in buildings: methods and challenges. Springer International Publishing; 2018. p. 77—105.
[12] Zhang Z. Microsoft kinect sensor and its effect. IEEE Multimedia 2012;19(2):4—10.
[13] Dziedzic JW, Da Y, Novakovic V. Indoor occupant behaviour monitoring with the use of a depth registration camera. Build Environ 2019;148:44—54.
[14] Wertheim HF, Melles DC, Vos MC, van Leeuwen W, van Belkum A, Verbrugh HA, et al. The role of nasal carriage in Staphylococcus aureus infections. Lancet Infect Dis 2005;5(12):751—62. Elsevier.
[15] Napoli C, Marcotrigiano V, Montagna MT. Air sampling procedures to evaluate microbial contamination: A comparison between active and passive methods in operating theatres. BMC Public Health 2012;12(1):1—6.
[16] Viani I, Colucci ME, Pergreffi M, Rossi D, Veronesi L, Bizzarro A, et al. Passive air sampling: The use of the index of microbial air contamination. Acta Biomed 2020;91(Suppl 3):92—105.
[17] Napoli C, Tafuri S, Montenegro L, Cassano M, Notarnicola A, Lattarulo S, et al. Air sampling methods to evaluate microbial contamination in operating theatres: results of a comparative study in an orthopaedics department. J Hosp Infect 2012;80(2):128—32.
[18] Saha S Agarawal, Khan AM. Air sampling procedures to evaluate microbial contamination: a comparison between active and passive methods at high-risk areas in a Tertiary Care Hospital of Delhi. J Patient Saf Infect Control 2017;5(1):18.
[19] Howorth FH. Prevention of airborne infection during surgery. Lancet 1985;325(8425):386—8.
[20] Bischoff P, Kubilay NZ, Allegranzi B, Egger M, Gastmeier P. Effect of laminar airflow ventilation on surgical site infections: a systematic review and meta-analysis. Lancet Infect Dis 2017;17(5):553—61.
[21] Gastmeier P, Breier AC, Brandt C. Influence of laminar airflow on prosthetic joint infections: a systematic review. J Hosp Infect 2012;81(2):73—8.
[22] World Health Organization. Global guidelines for the prevention of surgical site infection. 2016. Geneva.
[23] Whyte W, Lytsy B. Ultraclean air systems and the claim that laminar airflow systems fail to prevent deep infections after total joint arthroplasty. J Hosp Infect 2019;103(1). e9—e15.
[24] Alsved M, Civilis A, Ekolind P, Tammelin A, Erichsen Andersson A, Jakobsson J, et al. Temperature-controlled airflow ventilation in operating rooms compared with laminar airflow and turbulent mixed airflow. J Hosp Infect 2018;98(2):181—90.
[25] Knudsen RJ, Knudsen SMN, Nymark T, Anstensrud T, Jensen ET, La Mia Malekzadeh MJ, et al. Laminar airflow decreases microbial air contamination compared with turbulent ventilated operating theatres during live total joint arthroplasty: a nationwide survey. J Hosp Infect 2021;113:65—70.

PAPER 7

Das A, **Annaqeeb M K**, Azar E, Novakovic V, Kjærgaard M B. Occupant-centric miscellaneous electric loads prediction in buildings using state-of-the-art deep learning methods. *Applied Energy*. 2020; 269:115-135.

Applied Energy 269 (2020) 115135

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Occupant-centric miscellaneous electric loads prediction in buildings using state-of-the-art deep learning methods

Anooshmita Das*, Masab Khalid Annaqeeb, Elie Azar, Vojislav Novakovic, Mikkel Baun Kjærgaard

Mærsk Mc-Kinney Møller Institute, University of Southern Denmark, Odense 5230 M, Denmark
Department of Energy and Process Engineering, Norwegian University of Science and Technology, Trondheim NO-7491, Norway
Khalifa University of Science and Technology, PO Box 127788, Abu Dhabi, United Arab Emirates

HIGHLIGHTS

- Case-study to capture individualistic plug load patterns of occupants.
- Forecast energy consumption for occupant-centric miscellaneous electric loads.
- Compare state-of-the-art prediction models for forecasting.
- Evaluation with device-utilization patterns of multiple occupants.

ARTICLE INFO

Keywords:
Buildings
Consumption patterns
Occupant behavior
Miscellaneous Energy Loads (MEL)
Prediction models
Plug loads
Deep learning

ABSTRACT

Buildings have emerged as one of the dominant sectors when it comes to worldwide energy consumption. While a large portion of this consumption is due to the Heating, Ventilation, and Air Conditioning (HVAC) loads, a significant portion is contributed through the use of standard equipment, also known as Miscellaneous Electric Loads (MEL). It is necessary to understand the consumption patterns to optimize the MELs of the occupants using the building and conduct accurate forecasts for building energy management. One of the methods to achieve that purpose is the employment of Deep Learning (DL) methods. This study provides an analysis using Long Short-Term Memory (LSTM) model as a baseline for predicting MELs. The predictions were conducted for a day-ahead and a week-ahead period. Furthermore, the results from the baseline model were then used in a comparative analysis with two other state-of-the-art DL models; Bidirectional Long Short-Term Memory (Bi-LSTM) and Gated Recurrent Units (GRU).

The results from this study showed that both the Bi-LSTM and GRU models were significantly better than the LSTM model, especially when the prediction horizon was longer. The conclusions obtained can help implement these models in building energy management systems to draft strategic responses and schedules for more efficient energy usage.

1. Introduction

The building sector is a significant contributor to the massive and expanding energy demands which are responsible for greenhouse gases and carbon emissions [1]. Worldwide, this sector accounts for 35% to 40% of total energy consumption [2]. This ratio is even higher in countries with extreme climatic weather such as the United Arab Emirates (UAE), where buildings devour more than 70% of the power produced [3,4]. While a significant portion of that consumption is due to the cooling loads, around 17% constitutes the load consumption from standard equipment in buildings. It is crucial to have an understanding of consumption patterns to achieve energy savings in office buildings or shared spaces [4]. To mitigate the growing energy demands of the building sector, the impact of occupant behavior has become a focus of recent studies. Various studies have investigated the impact of occupants on the energy consumption in buildings to qualitatively and quantitatively comprehend occupant behavior, foster energy efficiency, and minimize the gap between the actual and predicted energy

* Corresponding author.
E-mail addresses: adas@mmmi.sdu.dk (A. Das), masab.k.annaqeeb@ntnu.no (M.K. Annaqeeb), elie.azar@ku.ac.ae (E. Azar), vojislav.novakovic@ntnu.no (V. Novakovic), mbkj@mmmi.sdu.dk (M.B. Kjærgaard).

https://doi.org/10.1016/j.apenergy.2020.115135
Received 17 February 2020; Received in revised form 27 April 2020; Accepted 2 May 2020

consumption [5,6].

Miscellaneous Electric Loads (MELs) are a critical hindrance for creating low energy buildings, and are a significant contributor to the building's energy load. This has been documented by Roth et al. [7] in the case of residential buildings located in the United States of America (US), and in studies regarding US office buildings [8]. These loads can contribute up to 20% of the total building energy consumption, and are set to increase by 40% in the next two decades [9]. MELs in buildings get referred to as the diverse electric loads emanating from electronic devices not responsible for Heating, Ventilation, Air Conditioning (HVAC), or lighting [10,11]. Syed and Hachem [3] considered MELs as a factor in a simulation study of a greenhouse-retail complex and found that MELs accounted for 23% of the building energy load. Burgett and Chini [12] conducted an analysis regarding the improvement of MEL prediction using occupant-centric methods. MELs are gaining attention as electronic devices get accumulated inside offices or buildings, and have become more sophisticated, thus generating upsurge in the miscellaneous electric usage [12,13]. MELs are one of the fastest expanding loads and are evolving to become one of the dominant load categories [3]. This growth reveals the fact that personal computers (PCs) and other office devices are penetrating office buildings, creating a sizeable base of installed computing equipment [1]. MELs comprise the vast majority of office equipment, while a fundamental part of them is plug loads related to Information and Communication Technologies (ICTs), such as desktops, monitors, and printers [14]. Plug load disaggregation is imperative for evaluating and investigating the underlying causes of energy wastage and developing strategies for energy reduction inside buildings [1,3].

1.1. Challenges of occupant-centric miscellaneous electric loads

The challenges in occupant-centric MELs energy research in buildings can get summarized as (1) unexplored and understudied as compared to design-focused research studies; (2) combined effects of uncertainty in numerous parameters responsible for the upsurge in MELs are neglected; and (3) particular types of buildings (e.g., educational facilities), or buildings exposed to extreme climatic settings are not considered, or inspected thoroughly.

Another challenge is that MELs are difficult to predict because individual occupants are in control of the electronic devices in a shared office or building. Burgett et al. [12] used an occupant based operational model to predict MELs and found that the results improved significantly when compared to the standardized occupant behavior models.

1.2. Benefits of predicting MELs

Considering the significance of MELs in the contribution to building energy, it is not only necessary to have an understanding of the consumption patterns, but also reliable estimates of those MELs. These estimates are important for decision-making processes [15], and establishing predictive control [16]. Also, accurate estimations are beneficial for predicting internal heat gains [16], and more importantly, for building performance simulations [17]. Building facility managers can benefit from an understanding of these patterns, and then look ahead into the future, for recommending energy-saving strategies, and determine the period of the peak loads.

However, these predictions often rely on typical load profiles and schedules, which are obtained from published benchmarks. A review of these benchmarks for offices in United Kingdom (UK) highlighted their inaccuracy, stemming from these benchmarks being outdated and unrepresentative of the equipment currently used in buildings (that contribute to the MELs) [18]. Only a handful of studies employ more sophisticated models for such predictions [19,20]. For example, Menezes et al. [19] introduced two methodologies for predicting MELs, one of which made use of detailed monitored data, and the other was independent of it.

The use of deep learning models is rare, but some studies have made use of the Long Short-Term Memory (LSTM) algorithms to forecast electricity consumption in commercial and residential buildings. Wang et al. [16] applied an LSTM model to predict internal load in buildings, in order to facilitate predictive HVAC control. Marino et al. [21] employed the same technique in building energy load forecasting, while Rahman et al. [22] used it to predict electricity consumption. The motivation for using LSTM models is based on the transient nature of electricity consumption patterns. Arahal et al. [23], and Fan et al [24] explored this temporal dependency. Another problem with such predictions is the long term dependencies, which is not accounted for in vanilla Recurrent Neural Networks (vRNNs). Hochreiter and Schmidhuber [25] recommended LSTM algorithms to address both these issues. Based on these studies, the baseline model in this paper was taken as LSTM. Besides, newer models like Gated Recurrent Units (GRUs), or variants of LSTM (such as Bi-directional LSTM) have not yet been tested for these purposes, which motivated the comparative analysis in this research work.

1.3. Research contributions

This research work enables us to delve into the limitations in the previous studies by proposing a comparison between state-of-the-art prediction models to quantify the impact of occupants on MEL energy consumption. The prediction models apply time-series data analysis methods to capture potential synergetic repercussions of occupant actions and consumption patterns on building performance. In this paper, a case study is presented on a typical educational building located in the extreme hot climate of Abu Dhabi, United Arab Emirates (UAE). Such climates have not yet been the target of studies focusing on MEL predictions, signifying a notable gap in the literature.

In this research work, we try to find a solution to the following questions:

(1) LSTM, Bi-LSTM, or GRU? Which model yields better results and achieve greater prediction accuracy on extrapolating consumption patterns for plug load devices using heterogeneous sensors?
(2) How can we make accurate predictions on the number of occupants and energy consumption patterns in a non-intrusive and reliable manner? Does occupancy duration have more influence on prediction accuracy compared to the occupancy ratio? Does the duration of prediction (look ahead into the time horizon) have an impact on the prediction accuracy of the different models?

The significant contributions of this research work:

(1) Present predictive models to forecast energy consumption more transparently and consistently and try to find out where the energy consumption is coming from, measuring in detail its energy footprint for the different plug load devices. Thus, it is interesting to identify and present state-of-the-art techniques regarding efforts to characterize, analyze, measure, and reduce the consumption of MELs.
(2) Propose a comparison between state-of-the-art prediction models. Experimental results highlight that the Bi-LSTM model is slightly more accurate than the GRU and the LSTM (baseline model). However, the GRU and Bi-LSTM converges around the same number of epochs than the baseline LSTM model.
(3) Device-utilization patterns for multiple occupants get extrapolated inside the multi-utility test-space. A comparison between ground-truth and predicted patterns was carried out to demonstrate which model yields higher accuracy. Also, proposed reasonable strategies for the reduction of MELs inside buildings.

Besides the upsurge in energy consumption, MELs also influence the power quality of the network, generally by implanting harmonics and transients in the voltage signal. However, this is not the prime focus of this research work; the practitioners and model developers should consider it as another essential aspect that affects the power quality inside a building that requires further investigation [3].

The rest of the work in this paper is structured as follows: Section 2 highlights the related work in this field. Section 3 describes the problem formulation. Section 4 explicitly describes the case scenario and the experimental set-up. Section 5 illustrates the data acquisition to analysis methods in detail. Section 6 explicitly describes the deep neural network models as well as the implementation steps of the prediction models. Section 7 describes the hyper-parameter tuning and optimization strategies. Section 8 describes the experimental results for the case study, and Section 9 discusses the model evaluation results. The summary of the lessons learned is presented as a discussion in Section 10. Section 11 elucidates the applications of the prediction framework and also includes open research questions for future research directions. Finally, the conclusion is presented in Section 12.

2. Related work

This section explores the related works found in literature, which includes the use of MEL for whole-building energy simulations, behavioral impacts on energy savings, and control systems for building energy management.

A significant amount of these studies obtained the loads from regional or global design standards [26,4], regional surveys [12,27,26] or electricity meters [15,13,28,29] installed in buildings. A broad range of building types are covered as well, most common ones being offices or commercial ones, but there is a notable dearth of academic buildings in these studies. The key findings of these related works, and the associated Köppen–Geiger climate classifications (KG-CC) of the cities they were conducted in, are presented in Table 1. The Köppen climate classification (KG-CC) system is one of the most commonly used systems, wherein the regions are represented by letters that indicate the amount of precipitation and the normal temperatures that region experiences. The regions are first divided into five main climate types; A (tropical), B (dry), C (temperate), D (continental), and E (polar). The second letter indicates the seasonal precipitation, while the third letter indicates the heat levels [30]. It can be identified from the table that there are various prediction methods used for the consumption patterns, including LSTM [16]. However, it can be noted that no works are making the use of Bi-LSTM or GRU.

3. Problem formulation

The prediction of an occupant's energy consumption into the future at any time instant t gets represented by the occupant's presence status S and consumption patterns C. Occupancy varies throughout the week within the monitored space, and each occupant is perceived as a decision-making agent who considers multitudes of different environmental and contextual parameters before deciding which plug load devices to use. At any time instant t, the n^{th} occupant represented as (O_n) has a fixed desk position D, assigned to the occupant and plug load devices deployed at each desk P, where $P = p_0,\ldots,p_n - 1$. This task gets perceived as a regression problem. The time-series data we observe of the occupant for different plug load devices are from time-stamp T_0 to $T_{obs} - values$. And, the predictions made from $T_{obs} - values + h$ to $T_{predict}$ with h as the desirable horizon ahead of the current time instant. h is called the windowing method, where the size of the window is a parameter that gets used to predict the miscellaneous load in the next time-stamp $t + 1$, based on the current time-stamp t and past time-stamp $t - 1$ and $t - 2$. An occupant spawns an input consumption sequence that corresponds to the values observed, and the task of the prediction models is to generate an output sequence for the different MEL plug load devices predicting the occupant's future consumption

Table 1
Related work.

Author and Year	Data Used	Building Type	KG-CC	Methods	Results
Afshari et. al. 2014	Typical building profiles and data from Urban Planning Council	Mixed-use office	BWh	Life cycle cost analysis (LCA) Marginal abatement cost curves (MACC) EnergyPlus models	Proposed retrofits for a bussiness-as-usual (BAU) building
Burgett et. al. 2014	MEL data extracted from RECS surveys 12,000 households (US EIA 2011)	Residential	Cfa/Aw	Regression model to predict MELs	Occupant characteristics predictors of MELs than building characteristics Achieved accuracy of 79%
Wang et. al. 2019	MELs, lighting, occupant counts, Wi-Fi connection counts	Office	Csb/Csa	LSTM	Reduced prediction errors compared to ASHRAE schedules
Mahdavi et. al. 2016	Occupancy (PIR) Plug loads (energy meters)	Office	Cfb	Aggregate estimation and stochastic model	Achieved NRMSE range of 12.6–20.3 % with stochastic models and 12–14% with simplified models.
Mahdavi et. al. 2017	Occupancy (PIR) Plug loads (energy meters)	Office	Cfb	Simplified and probabilistic models (using Weibull distributions)	Reduced previous RMSE values by including occupant diversity
Wang et. al. 2018	Electricity data (energy meters)	Office	Dfa	LMS nLMS RLS GMMR	RMSE (kw) for different models: LMS − 124.8 nLMS − 129.6 RLS − 122.0 GMMR − 45.2
Yalcintas et. al. 2008	Occupancy Hourly electricity measurements	Hotel	BSh	ANN	RMSE ranged from 6.81 to 16.4%
Lee et. al. 2001	Design criteria from ASHRAE Energy end-use survey by SRCI	Commercial	Cwa	Comparative analysis of design criteria and surveyed values	Recommended realistic design criteria, saving estimates of 6–22% of electricity consumption
Sarfraz et. al. 2018	Power consumption profiles Load factor profiles	Office	NA	Combined diversity factors and load fators	Recommended using office level load factors for overall load calculation
Ren et. al. 2013	Time use survey Metering data	Residential	Cfa/Cfb	AccuRate	Prediction error 6.5%
Christiansen et. al. 2015	Plug load measurements	Hospital	Cfb	Multiple regression	Error of less than 6%

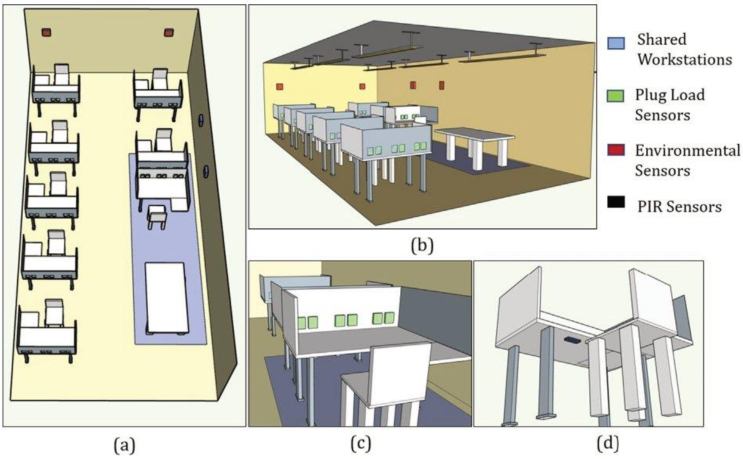

Fig. 1. Experimental Set-up (a) Bird's view angle of the space layout. (b) Environmental sensor placement in the space layout (c) Six Plug-load Sensors on each desk (Miscellaneous, Dock, Monitor 1, Monitor 2, Laptop, Lamp). (d) PIR sensors on each occupant desk.

patterns. The horizon of the prediction required (look ahead into the future) is flexible for an extension. It also depends on the demands which can range from an hour to day-ahead or week-ahead prediction. In this paper, we have a 24 h-ahead (day-ahead) and week-ahead prediction from the different deep learning models.

4. Experimental set-up and design scenario

The data collection was performed at a research facility located in the city of Abu Dhabi, UAE. The area used for data collection served as a shared work-space consisting of eight individual desks, which would be occupied by up to 6 graduate students at a time. The remaining two desks were designated as shared desks for the occupants. Also, another common table was present for shared use, as shown in Fig. 1. Since students used the test-space, the specific occupancy of each desk was subject to change every semester, with incoming and outgoing students.

In total, eight different occupants were included in the study over the entire course of the data collection, which spanned over a period of eight months during April - November in the year 2017. Each desk contained a personal lamp for illumination, controlled by the occupant, while the area itself was illuminated by six different lighting fixtures, each controlled by localized motion sensors. Besides, each desk was equipped with six power outlets, including the power source for the lamp. While the university was operational from 8 am to 5 pm on weekdays, the area was open for access throughout the week. Moreover, it should be noted that the area of study was not a fully controlled environment, which meant that the occupants had flexible hours for work, and while there were six primary occupants in the area, there were no restrictions for visitors to access the area. Such visitors occasionally occupied the shared table. However, all these visitors were students or staff members of the institute.

5. Approach

In this section, we explicitly describe the methodology implemented in this research work. The overview is also highlighted in Fig. 2.

5.1. Sensor placement strategy and calibration

The occupants were monitored with regards to their presence, device consumption patterns (via plug loads), and the associated environmental parameters such as illuminance, temperature, and relative humidity. This was achieved with the use of three different kinds of sensing modalities: plug load sensors, PIR (Passive Infra Red) sensors for the occupant presence, and environmental sensors to record the illuminance (lighting status), temperature, and humidity. For more information about the specific sensors, the individual datasheet can be referred to [31,32]. In total, nine PIR sensors from OccupEye [31] were deployed in the area, eight on the work desks, and one on the common table. The sensors were placed under the work desks, as shown in Fig. 1(d), and no other objects were placed in the area beneath the desk, so as to avoid occlusion and have a better range of detection. To find the optimal position of the PIR sensor at each desk, three factors were

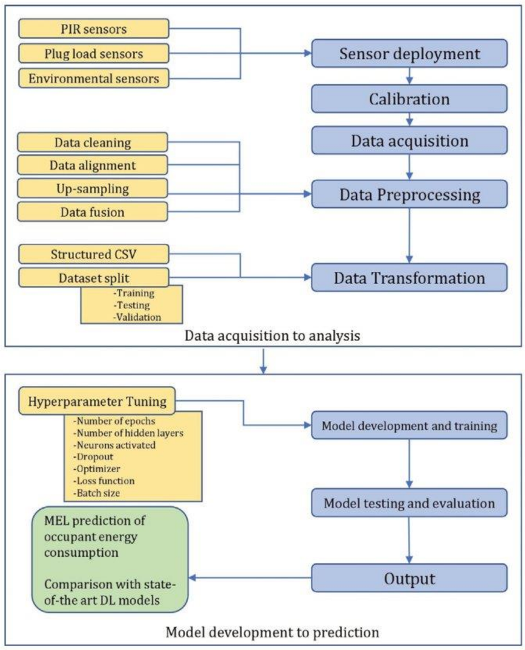

Fig. 2. Overview of the methodology.

taken into consideration; first, the sensors needed to be able to detect the occupant's presence at all positions at the desk, and second, the optimal position should enable the sensor to avoid any triggers from passers-by near the desk. Thirdly, the sensors are mounted to ensure there is no occlusion in its Field of View (FoV). Different positions and scenarios were performed, recording the triggers from the sensor to determine the optimal position with respect to the three conditions.

For the plug loads, each desk was equipped with six sensors, and each power outlet was designated to be used by a specific device, as seen in Fig. 1 (c). This was done in order to label and segment the plug load consumption according to the devices plugged in. The devices used for each socket and sensor were kept constant throughout the data collection process. To ensure the due procedure was followed, weekly checks were conducted.

As for the environmental sensors, one of the objectives of the study was to be independent of any Building Management System (BMS), which is why the data was not extracted from any BMS. Instead, the environmental sensors were used locally to monitor the environmental conditions around the occupant. The sensors were placed at different heights and positions near the lighting fixtures in order to accurately capture the luminance levels for each, without any significant interference from the others. Hence, six such sensors were used, one for each lighting fixture.

While the data collection included input from all the sensors, it should be noted that the environmental data was not used in the deep learning models in this research work. The environmental data was collected to understand the relationships and dependencies of occupants' visual and thermal comfort with the energy consumption patterns, which was not included in the scope of this study. However, there is a possibility to extend this work in that direction in the future. With regards to the data from the two remaining sensing modalities, the application of the PIR data was for the authors to understand the occupant presence status within the domain and context of the space. The deep learning models were modeled for the weeks wherein the authors had determined that the occupants were present in the area of study, and tested the models in the week with maximum occupancy. The plug load data represented the MEL consumption data for each device, and was the dataset used to develop the deep learning models.

5.2. Data acquisition

The occupants were monitored for a period of 8 months (April – November 2017) in total. However, due to the malfunction of devices and the associated missing data, the time period taken into consideration was for around 6 months, or 167 days precisely. The data were individually collected from each sensor's respective data storage platform in the form of comma-separated values (CSV) files. Along with the routine checks to ensure the functionality of sensors, this procedure was conducted at weekly intervals. Among the sensors deployed, both PIR and environmental sensors were battery-operated and could collect data for up to 6 months at a stretch. The plug load sensors used the power source they were attached to for functioning.

The data collection process for the environmental and plug load sensors was continuous, reporting the associated device consumption and environmental parameters at 15-min intervals. The PIR sensors, however, were event-based, recording triggers as soon as the occupant enters the detection range of the sensor. The sensor also registered the absence as soon as the occupant was out of the range. However, to account for cases where the occupant may have momentarily left the area, on occupancy duration of five minutes was selected. The five minutes was a wait-period before reporting for an 'Absent' reading; i.e., the sensor would have to detect the absence of the occupant for a continuous five minutes.

5.3. Data pre-processing and data transformation

In this section, we explicitly describe the pre-processing steps for the collected data. Since the PIR sensor was event-based, it was not aligned to a consistent temporal resolution. Also, those sensors recorded 'blank' triggers when there was no change in the previously reported presence status. To counter these issues, all the 'blank' triggers had to be first substituted with the occupancy triggers they were indicative of. The data was then processed to sort and replicate the last recorded trigger for every fifteen minutes. After up-sampling, each sensor now had four readings every hour (one reading every fifteen minutes), totaling 96 readings per day.

For anomaly detection, the routine checks involved determining if data was being continuously reported by each sensor. In some cases, there was some missing data, frequently attributed to sensor malfunction. The malfunctions included sensors running out of battery and removal of sensors from designated positions by maintenance personnel.

Considering the nature of the study, it was necessary to have consistent data for each device and its user, and in order to maintain that consistency, weeks with missing data were discarded. That was the reason for the difference between the original monitored time of 8 months and the final duration of 6 months after data pre-processing. The final data set consisted of 16,032 rows (96 readings $\times$ 167 days) of data for each occupant. After the pre-processing step, the data was transformed into a structured and consistent format to work with and build the prediction models.

The sufficient amount of data needed for modeling depends both on the complexity of the problem at hand and on the complexity of the chosen algorithm. However, one heuristic about the relation between the number of features and training samples is that the number of training data should be 10 times more than the number of features. Our MEL dataset comprises of consumption data (via plug loads) for 6 devices (Miscellaneous, Monitor1, Monitor2, Docking Station, Laptop and Lamp).

The dataset was split into training, testing and validation sets in the ratio 60%: 20%: 20% respectively, which consists of 57,715 training samples, 19,238 testing samples and 19,239 validation samples. There are 6 features in the dataset. Thus, it is ensured that the training data, with approximately around 60,000 samples and 6 features in the dataset is sufficient to build the predictive models.

5.4. Privacy handling and data suppression

This section details the steps that were taken to ensure the privacy of the occupants that were monitored. Each occupant was monitored on a voluntary basis, and was provided with a consent form detailing the data that was being collected, and the duration and aims of the study. This consent form was approved by the Human Research Ethics Committee of the host university. In order to avoid revealing the identification of the occupant, the data was labeled as Occupant 1, 2, etc. No names or other personal details of the occupants were recorded.

The most sensitive part of this dataset is the data collected outside of the opening hours since those readings are most likely collected from employees working in the monitored area. This part of the data can be used by the employer to estimate the work performance of the employees in the area. Data suppression is generally referred to as the process of withholding or ad-hoc removal of a selected piece of information from the data to protect the identity, privacy, and revealing confidential information of any individual occupant. This is a crucial step when sharing data publicly or with third parties to protect the privacy of each occupant. Due to this, for the future release of the datasets, we will not release any of the readings collected outside of the opening hours. Furthermore, the occupants in the monitored area also have the right to privacy. We have decided to protect the identity of the days, by not including the dates as part of the dataset. Furthermore, re-

Table 2
The attributes included in the dataset (which is collected for a period 8 months).

Attributes	Description
Time	Time-stamp when the entries were collected.
Time-Interval	Every 15 min (up-sampled).
Day ID	Randomly assigned number which holds the ID for the specific day in the dataset.
Week ID	Randomly assigned number which holds the ID for the specific week in the dataset.
Workday	Boolean, telling if the entries were collected on a workday.
Weekend	Boolean, telling if the entries were collected on a weekend.
Holiday	Boolean, telling if the entries were collected on a national holiday.
Occupant ID	Each head-count is assigned a unique Occupant-ID, in number.
Presence Status	The occupancy status in a boolean format (0/1) from PIR sensors.
No. of occupants	8 occupants observed (in-situ)
Consumption Data	MEL Devices: Miscellaneous, Dock, Monitor 1, Monitor2, Laptop, Lamp

ordering the days by mapping the date component, thus creating a random permutation of the days.

Similarly, for the weeks, we have a Week ID indicator. These precautions make it significantly more difficult for adversaries to perform data linkage attacks upon the released data, and hence identifying/revealing the identity of the occupants or the location of the monitored area. In the dataset, we have introduced a workday indicator that accounts for weekends and national holidays. Furthermore, we included a national holiday indicator. The time-stamp remains in the up-sampled form (15 min). Table 2, presents an overview of all the attributes in the dataset.

Later, we want to release the full dataset, which can serve for benchmarking and foster data-driven research in occupant centric MEL prediction. This dataset used in this paper would help to extract the inter-relationships of different devices influencing occupant behavior and utilization patterns.

6. Models used

Initially, the time-series dataset gets split into training, test, and validation set. The prediction models have been developed using the training sets, and the predictions are made on the test set, which is unseen by the model. The prediction model used for comparison are - Long Short-Term Memory (LSTM) [baseline model], Bidirectional Long Short-Term Memory (Bi-LSTM), and Gated Recurrent Unit (GRU).

The next step is transforming the time-series data into a supervised learning problem, i.e., the data gets organized into input and output patterns where the observation at the previous time step is fed to the network as an input to predict the observation at the current time step. Another step would be that the observations get transformed to have a specific scale, i.e., re-scale the data values to lie between -1 and 1. The re-scaling gets done to meet the default hyperbolic tangent activation function of the model. These transforms are inverted on predictions to revert them into their original scale before calculating an error score.

LSTM was proposed by Hochreiter et al. in 1997 [25], and the major motivation behind building the model using LSTM for MEL predictions on plug load devices for multiple occupants is that the model can account for energy savings and efficient device utilization within the test zone. Also, to comprehend the significant factors influencing decision-making for different device choices and individual consumption patterns in the multi-utility test zone. This could further help to perform intelligent building operations and curtail energy wastage. Bidirectional LSTMs (Bi-LSTMs) proposed by Graves et al. in 2005 [33] are an augmented version of conventional LSTMs that can boost up the performance of the model on prediction problems. Cho et al. in 2014 proposed GRU and it got successfully implemented for sequence prediction tasks [34].

For implementing the prediction models, we have used Python language to evaluate and process the time-series data. For the data analysis, we developed prototypes using Scikit-Learn, Scipy, Pandas, Numpy, Seaborn, Matplotlib libraries. We also used sequential models in Keras (compatible with python 3.6 version) under Tensor-flow as back-end. We used a workstation with an Intel i7-8850H with a base frequency of 2.6 GHz and a turbo boost to a decent 4.3 GHz. Our algorithm utilizes Tensor-flow as the back-end to evaluate the overall model performance and computational efficiency. The explicit description of how each model gets implemented follows in the next subsections:

6.1. Long Short-Term memory (LSTM) [Baseline Model]

The advantage of using LSTM as a baseline model over others is that LSTM's can randomize the order dependencies, possess memory blocks across observations in the input, which Multi-Layer Perceptron (MLP's) lack. The gates involved in implementing the LSTM and the information-flow is described below:

1) Forget Gate(Ft) decides which information to delete that is not important from the previous time-stamp. The unnecessary parts of the previous cell state are forgotten. To decide which information to be omitted to form the cell in that particular time step, it is decided by the sigmoid function σ. It looks at the previous state $h(t-1)$ and the memory of the previous unit represented as m_{t-1} and the current input X_t then compares the function. The weights get represented by w. The weights in the LSTM module is updated using Back-Propagation through Time (BPTT). This enables stability in the model. This can be given by the Eq. (1):

$$Ft = \sigma(wX_t + wh_{t-1} + wm_{t-1} + bias(Ft)) \tag{1}$$

2) The second layer comprises 2 parts i.e., input gate (It) and cell state (St). There are two activation functions, one is the sigmoid function, and the other is tanh. The sigmoid function then decides which values to let through (0 or 1). The tanh function gives the weightage to the values which are passed, deciding their level of importance (-1 to 1). Cell state (St) can selectively update cell state values and decide what part of the current cell state makes it to the output and gets defined by the Eqs. (2) and (3):

$$It = \sigma(wX_t + wh_{t-1} + wm_{t-1} + bias(It)) \tag{2}$$

$$St = \sigma(wX_t + wh_{t-1} + wm_{t-1} + bias(St)) \tag{3}$$

3) The third step is to decide what will be the output. It gets governed by the output gate (Ot), and Ht captures the entire process, defined in Eqs. (4) and (5):

$$Ot = \sigma(wX_t + wh_{t-1} + wm_t + bias(Ot)) \tag{4}$$

$$Ht = (Ot) * tanh(St) \tag{5}$$

6.2. Bidirectional Long Short-Term memory (Bi-LSTM)

Bi-LSTMs train on two rather than one LSTM on the input time-series data. BiLSTMs runs the inputs in two approaches, one from past to future and one from future to past, i.e., the first on the input sequence as it is, and the second is a reversed copy of the input sequence. However, what varies this approach from others is that in the LSTM that runs backward, the information from the future gets retained. Using the hidden states in combination enables the preservation of information at any point in time from both past and future. The implementation steps are defined by Eqs. (6)–(8).

$$h_t^f = f_h(wX_t,\ h_{t-1}^f)\cdots Forward \tag{6}$$

$$h_t^b = f_h(wX_t,\ h_{t+1}^b)\cdots Backward \tag{7}$$

$$h_t = f_o(h_t^f,\ h_t^b)\cdots OutputLayer \tag{8}$$

Bi-LSTMs can provide additional context to the neural network and result in faster convergence and even fuller learning on the prediction problem, although it depends on the task. The structure of Bi-LSTM allows having both backward and forward information about the sequence of consumption patterns at every time step. By using the information from the future, it becomes easier for the network to predict the consumption patterns efficiently.

6.3. Gated Recurrent Unit (GRU)

In this section, we describe the execution of the prediction model, using Gated Recurrent Units (GRU). GRU performs better since it has a less complicated structure and is computationally less expensive. Also, the GRU training phase is faster than RNN or LSTMs on limited training data.

Another advantage of GRU is that it solves the problem of vanishing gradient, which generally occurs with vanilla RNN. Vanishing Gradient problem occurs when the gradient shrinks as it back-propagates through time. If the gradient value becomes too small, it doesn't contribute much in the learning phase.

To solve the vanishing gradient problem and short-term memory, the gates in GRU help to regulate the flow of information and handle which data in the sequence is essential to retain and others to throw away. By continuing this process, the relevant information is passed along the sequence-chain and makes accurate predictions.

The GRU discloses the memory content at each time-stamp between the previous and the upcoming next memory content with the help of an update gate. The update gate governs how much of the previous memory state should get forgotten and how much new content should pass into the future. The update gate is analogous to the forget and input gates of an LSTM model. GRU has fewer parameters and train a bit faster and also need fewer data to generalize. GRU's got rid of the cell state value and directly use the hidden layers to transfer the information ahead. The update gate z_t at time-stamp t gets defined in the Eq. (9):

$$z_t = \sigma(w^z X_t + U^z h_{t-1}) \tag{9}$$

X_t is the input that gets multiplied with its weight w^z. Similarly, h_{t-1}, which holds the information regarding the previous time-stamp t-1, is multiplied by its own weight U^z. The results added together, and a sigmoid activation function gets implemented to squish the results between 0 and 1. The reset gate decides how much of the past information to forget. It can be defined as given in Eq. (10):

$$r_t = \sigma(w^r X_t + U^r h_{t-1}) \tag{10}$$

The current memory content h_t, which uses the reset gate to store the necessary information from the previous time-stamp, is given in Eq. (11). An element-wise product between the reset gate r_t and Uh_{t-1} is performed. This step handles what information is needed to be removed from the previous time-stamps and apply the tanh activation function. The tanh activation function squishes the values between −1 and 1, thus regulating the output while performing the parameter tuning through successive runs.

$$h_t = tanh(wX_t + r_t \odot Uh_{t-1} \tag{11}$$

6.4. Model evaluation metrics used

For the evaluation, a rolling-forecast gets implemented, also known as walkforward model validation. Every time step of the test set gets escorted one at a time. The prediction models are used to make a forecast on the MEL consumption for the time step, and then the actual expected value from the test dataset gets captured and made accessible to the model for the prediction on the next time step. This walk-forward model validation mimics a sophisticated real-world setting where training data would be accessible, and test data gets used in forecasting (in a oneshot method), i.e., the energy consumption of the occupants. All forecasts on the test dataset get accumulated; thus, an error score gets calculated to encapsulate the skill of the prediction model.

For the evaluation of the models, the first metric used is Root Mean Squared Error (RMSE), and it quantifies the amount by which the estimator deviates from the targeted output. The RMSE gets used as the evaluation metric as it penalizes significant errors, and the outcome is a score that is in the same units as the forecasted data, i.e., occupants plug load energy consumption.

Another metric used is Mean Absolute Error (MAE), which is the average of the absolute values of the differences between predicted and the corresponding observation in the plug load consumption data. The evaluation results are elucidated in Section 9.

7. Hyper-parameter tuning and optimization strategies

The hyper-parameter for a model is a configuration that is external to the model and whose value cannot get approximated from the data. The hyper-parameters are often practiced to help estimate model parameters and are usually specified heuristically. The hyper-parameters get tuned for a given prediction problem. The best value for a parameter tuning on a given problem is not known; however, we often use different rules of thumb or explore for the best value by trial and error. The hyper-parameters for the prediction model gets discussed below:

The *number of training epochs* is the number of times that the entire training dataset gets demonstrated to the network during training. The *batch size* of a model can be referred to as iterative gradient descent in the patterns exhibited and defining what patterns to read at one go and restore in-memory and discard others before the weights get updated in the model [35]. *Dropout layer* was added to prevent overfitting. *Mean squared error (MSE)* is used as the loss function as it helps to evaluate the accuracy of the model in predicting the test data.

Optimization Strategies:. We implemented a grid search for hyper-parameter optimization. Grid search helps to find the optimal hyper-parameters of the predictive model, which results in better predictions.

Table 3
Hyper-parameters selected for the three different deep learning models.

Hyper-parameters	LSTM	Bi-LSTM	GRU
1) No. of Training Epochs	100	100	100
2) No. of Neurons Activated	64	64	64
3) No. of Layers	3	2	3
4) Batch-size	64	256	256
5) Dropout	0.1	0.1	0.1
6) Learning Rate	0.001	0.001	0.001
7) Activation Function	tanh	tanh	tanh
8) Optimizer	Adam	Adam	Adam
9) Loss Function	MSE	MSE	MSE
10) Evaluation Metric	MSE MAE MAPE	MSE MAE MAPE	MSE MAE MAPE

Table 3 highlights the selected hyperparameters for the three different deep learning models. Also, we choose Adam because it can get perceived as a merged version of RMSprop and Stochastic Gradient Descent (SGD) with momentum [36]. The advantage of using Adam is that it is computationally efficient and straightforward to implement for big datasets; also, hyper-parameters have intuitive interpretation capabilities and require less parameter tuning. Adam is also appropriate for noisy datasets or a sparse gradient.

Batch Normalization (BN) is another optimization technique where the distribution of each layer in the input of the deep learning model can change quite consistently. As the input changes, the model parameters keep changing too, during the training phase. Lower learning rates can be set to deal with the dynamic parameters. However, lower learning rates slow down convergence and therefore make the learning process slower. There seems to be a trade-off between dynamic parameters and learning rates. The trade-off gets more emphasized with saturating nonlinearities across different layers in the deep learning model. It has been demonstrated how by initializing values of parameters with zero mean and unit variance, also known as normalization. By updating parameters as training progresses for every mini-batch, and then back-propagating through time (BPTT), it is viable to use higher learning rates, and to pay less attention to the initial values of the parameters. This process not only makes learning more robust but also speeds up the training process, as higher learning rates need fewer epochs to converge. Notably, batch normalization acts as a regularizer, requiring less dropout and discouraging overfitting [37], normalizing the loss, i.e., sum the loss terms along with the sequences and divide by the maximum length of the sequence. In that way, it becomes easier to reuse the hyper-parameters between multiple case scenarios.

Another strategy is early pruning of the training phase to avoid overfitting and to evade from training a neural network more than needed. Detecting when a model offsets to overfit the data is a challenge and one of the few methods to discover when the network is not learning anything new about the data comprises of investigating the validation loss, which gets calculated on the validation dataset. If the validation loss does not improve, it indicates that further training doesn't add anything new to the current parameters of the model. Early pruning is decided based on the patience parameter in Keras. In our case, a value of 30 epochs was used for the early pruning of the models if the validation loss doesn't decrease and remains almost stagnant for 30 epochs).

A good fit represents a case where the performance of the deep learning model is extremely well on both the training and validation sets. This can get diagnosed from a plot where the train and validation loss decrease and stabilize around the same number of epochs. We have highlighted diagnostic plots for the three deep learning models, see Fig. 3 for LSTM, Bi-LSTM, and GRU, respectively. The Fig. 3 shows that the models stabilize around the same point indicating a well-fitted model. The loss function used is Mean Squared Error (MSE).

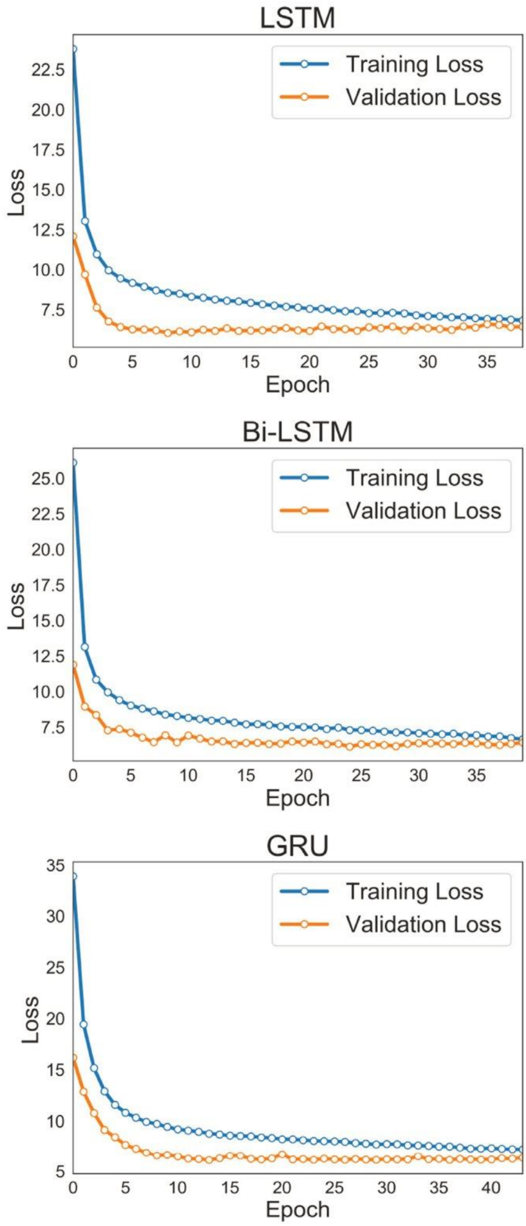

Fig. 3. Training Loss vs. Validation Loss for the LSTM, Bi-LSTM and GRU Model respectively, using Mean Squared Error as the Loss Function. The unit is in kW.

8. Experimental results

This section highlights the day-ahead and week-ahead prediction for the different plug load devices in Figs. 4 and 5, respectively, for Occupant No. 6. The week chosen for the week-ahead prediction was selected based on its high occupancy. In order to compare the occupancy between different weeks, diversity factors were used. Diversity factors represent the ratio of actual occupancy to the maximum possible occupancy of an hour. These factors form the basis for generating standard occupancy schedules in offices [38,39]. The week selected was the one with the highest diversity factor.

The six plug load devices are - Miscellaneous (no fixed device label i.e., occupants had the flexibility to plug in any device), dock charging station, monitor 1, monitor 2, laptop, and desk lamp.

From Fig. 4, it is evident that the traffic of occupants starts at around 9 am. Based on the prediction result analysis, Occupant No. 6 has unstable patterns for the Miscellaneous plug load. Miscellaneous is hard to predict, and the reason is that there is less training data, and it becomes hard for the models to learn and capture the consumption patterns for it. However, the Bi-LSTM model was the best one to capture the miscellaneous plug load patterns with the least RMSE and MAE. The lower value of RMSE and MAE interprets that the predictions were closer to the groundtruth data.

As seen in Fig. 4, dock consumption had quite a consistent pattern

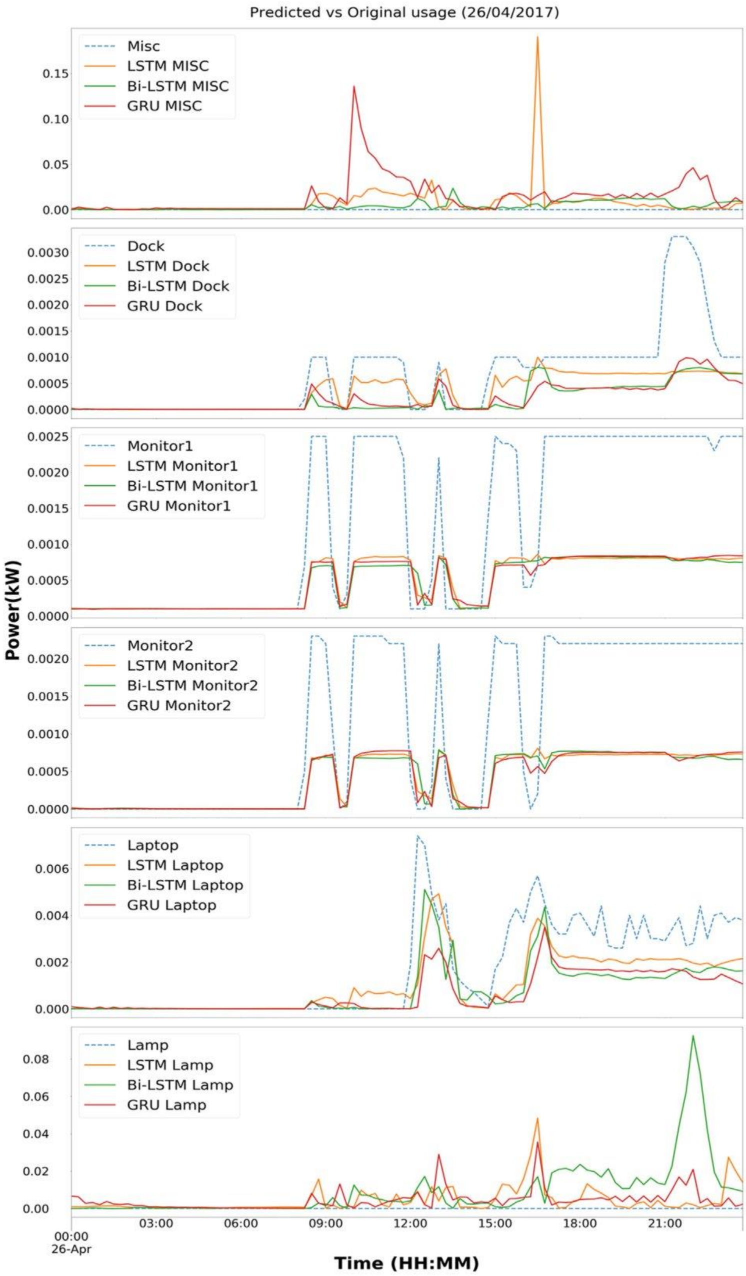

Fig. 4. Day-ahead Prediction for Occupant 6 shows comparison between the prediction models (LSTM (baseline), Bi-LSTM and GRU with the actual MEL consumption (dashed) for the six plug load devices [in kW].

during office hours.

However, it can be noted that there is a peak right before midnight, which coincides with the deviation in the models' prediction (LSTM, Bi-LSTM, and GRU). This peak is again repeated in the monitor and laptop consumption as well. This would constitute a scenario where the desk occupant has stayed much longer than the regular office hours. Nevertheless, the deviation of the prediction is less than two decimal points, indicating the accuracy and reliability of such a prediction. The peak consumption of the monitors starts with the regular office hours, taking a dip at noon, which represents the occupants leaving for lunch. The similarity in the consumption patterns for the monitors is corroborated by the fact that both of the monitors are simultaneously

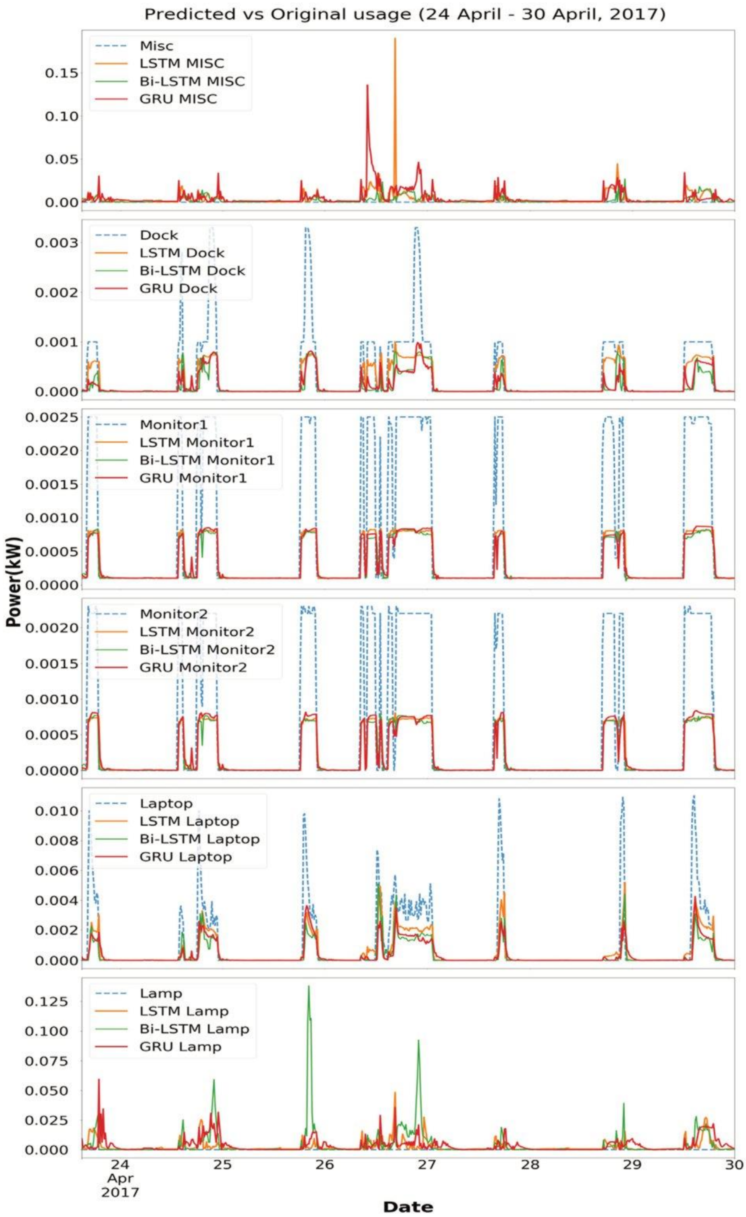

Fig. 5. Week-ahead Prediction for Occupant 6 shows comparison between the prediction models (LSTM (baseline), Bi-LSTM and GRU) with the actual MEL consumption (dashed) for the six plug load devices [in kW].

switched on and used by the occupant.

The peak load for laptop consumption occurs from 12 pm to 1 pm, indicating a lag in the initiation of the consumption compared to the monitors and the dock. In comparison to the rest of the devices, the lamp is rarely used by the occupant. The area of study was well lit by the motion-controlled lights, which were complemented with natural daylight during the day. Hence, the need to use the lamps did not often arise, which can be seen in both Figs. 4 and 5.

From the week-ahead predictions in Fig. 5, it can be seen that the largest deviations occur in the miscellaneous plug load and lamp consumption. Both of these categories had the least amount of training data since the occupants rarely used the lamps, and miscellaneous plug load had no consistent device attached to it. Fig. 6 highlights validation loss [in kW] for the different prediction models. It shows that the Bi-LSTM model is more stable compared to GRU and the baseline LSTM model.

The monitors consumed extra power, and the least power was devoured by the lamp, out of the six plug load devices. After the comparison of the different models, we can see that the LSTM predictions

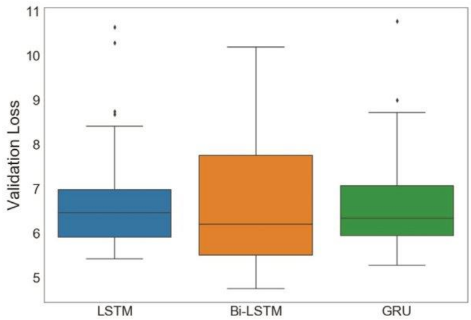

Fig. 6. Validation loss [in kW] for the three different prediction models.

were better for day-ahead predictions. However, the LSTM model performance dropped significantly for the week-ahead predictions, and the comparison is illustrated in Figs. 4 and 5. On the contrary, GRU did not perform very well on the day-ahead predictions; however, it captured the future consumption patterns more competently in the week-ahead predictions. Furthermore, when comparisons are made between Bi-LSTM and GRU, Bi-LSTM is slightly better than GRU.

9. Model evaluation results

In this section, we explicitly describe how we evaluated the model. As mentioned in Section 6.4, the evaluation metrics used are RMSE and MAE. The RMSE is directly interpretable, making it a better measure of *goodness of fit* rather than using a correlation coefficient. Furthermore, the errors are squared before they are averaged, thus applying a relatively high weight to significant errors. In this case, RMSE is an appropriate metric since it provides the benefit of penalizing large errors. Table 4 highlights the average RMSE [in kW] from each plug load device for the different prediction models for the day-ahead and week-ahead predictions.

The RMSE values for the GRU model for the plug load devices were closer to the groundtruth, except miscellaneous plug load, see Fig. 7 in subplot A and C, for comparison. Finally, we can interpret from the plots that the miscellaneous plug load has a vital role in the MEL prediction, and it can negatively impact the performance of the prediction models.

Another evaluation metric used is Mean Absolute Error (MAE) for evaluating the three different deep learning models. The Mean Absolute Error is the average of all absolute errors between actual and predicted values. Table 5 highlights the MAE values from each plug load device for the different prediction models for the day-ahead and week-ahead predictions. The MAE values for the GRU model mirrored the values from the RMSE evaluation, by being closer to the groundtruth except

Table 4
Average RMSE Comparison [in kW] between the different prediction model on the plug- load devices [Bold indicates lowest RMSE out of the three prediction models for individual plug load devices].

Prediction	Day-Ahead Prediction			Week-Ahead Prediction		
Devices/Model	LSTM	Bi-LSTM	GRU	LSTM	Bi-LSTM	GRU
RMSE_Misc	0.7527	**0.3960**	1.3805	3.1515	**1.5172**	2.0673
RMSE_Dock	**0.0281**	0.0430	0.0413	0.1302	0.1387	**0.0874**
RMSE_Monitor1	**0.0741**	0.0760	**0.0749**	0.2388	0.2454	**0.2355**
RMSE_Monitor2	**0.0667**	0.0672	0.0674	**0.2116**	0.2189	0.2142
RMSE_Laptop	**0.0600**	0.0855	0.0969	0.3613	0.3843	**0.2860**
RMSE_Lamp	0.1020	0.0765	**0.0651**	0.2353	**0.2123**	0.3107

for the miscellaneous plug load. The comparison can be seen in Fig. 7 between subplot B and D.

10. Discussion and lessons learned

A significant conclusion from this state-of-the-art overview is that user behavior has a significant impact on energy consumption and device utilization in commercial buildings, and institutional buildings reflect the same pattern.

The influence of user behavior is a challenge to quantify for methodological interpretation. The decision-making process of end-user is multi-factorial and complex; thus, factors influencing behavior are also numerous and varied. The dynamic nature of occupant's energy behavior is hard to comprehend, and multi-disciplinary approaches and meticulous investigations are required to contribute new insights into the energy-use domain.

To determine building occupant behaviors, a scientific study that interprets the dominant factors that are involved in energy behaviors has to be conducted with the users. Since the users do not always make rational decisions, the manner of presenting the choice itself becomes determinant in adopting energy-efficient behaviors. Also, the energy conservation measures introduced without taking into account user comfort and satisfaction can often have negative impacts and be counter-productive because users are likely to try to adapt to their environment to attain satisfactory conditions.

The prediction of energy-use complements the process of understanding and optimizing building energy consumption by providing valuable insights regarding the occupant's consumption patterns and device-utilization. These predictive models used in this research work become vital for implementing demand-response assessments, as well as providing pathways for valid pricing and tariffing for energy usage. Moreover, the benefits of such models can get amplified through the use of software tools to achieve these objectives.

11. Application and future research directions

An essential application of this work is the light it sheds on the accuracy of MEL prediction using deep learning models. These estimates through software tools can be beneficial to building owners and facility managers in design and decision-making processes to recommend energy-saving control strategies and identifying energy footprints from each device. Another important application is the usage of these predictions to identify periods of peak loads for equipment to further recommend energy retrofit measures more efficiently.

Studies investigating the effect of feedback on user consumption patterns highlight that it has indeed been a successful strategy [40]. The results of this work are useful in highlighting the diversity of occupant energy-use patterns with regards to MELs, and illustrates the subsequent opportunities that arise for contributing to such potential feedback systems.

An example of inducing behavioral changes can be in the form of gamification approaches in offices/mixed-use building, wherein occupants are provided with incentives for adopting energy-saving measures. However, case studies in this regard are rare and warrant further research and evidence. Some potential applications of MEL predictions to incorporate energy-saving measures can be:

1. Developing occupant awareness programs, coupled with awards or financial ncentives for reducing their energy usage.
2. Identifying the requirements of occupants through feedback systems and assess the trade-off between those requirements and actual energy usage.
3. Investigating heterogeneous sensing modalities to adopt the best metering technique for extrapolating such data.

For future studies, attention has to be paid to the interaction

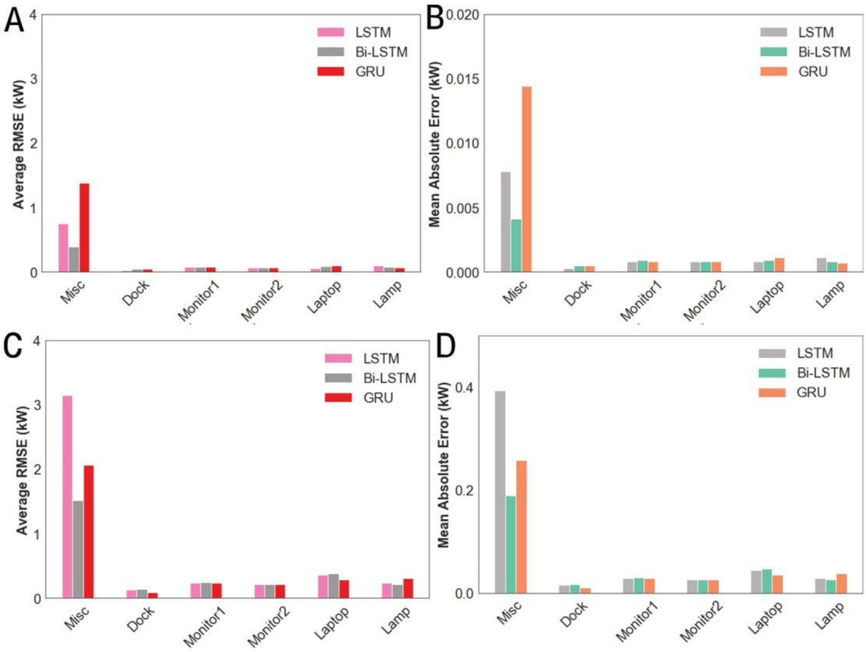

Fig. 7. Average RMSE and MAE comparison [in kW] between the different prediction models: LSTM, Bi-LSTM and GRU from the six plug load devices for day-ahead predictions in subplot: A and B and week-ahead predictions in subplot: C and D.

Table 5
Mean Absolute Error (MAE) Comparison [in kW] between the different prediction model on the plug load devices [Bold indicates lowest MAE out of the three prediction models for individual plug load devices].

Prediction		Day- Ahead Prediction			Week-Ahead Prediction	
Devices/Model	LSTM	Bi-LSTM	GRU	LSTM	Bi-LSTM	GRU
MAE_Misc	0.0078	**0.0041**	0.0144	0.3939	**0.1897**	0.2584
MAE_Dock	**0.0003**	0.0005	0.0005	0.0163	0.0173	**0.0109**
MAE_Monitor1	**0.0008**	0.0009	**0.0008**	0.0298	0.0307	**0.0294**
MAE_Monitor2	**0.0008**	**0.0008**	**0.0008**	**0.0265**	0.0274	0.0268
MAE_Laptop	**0.0008**	0.0009	0.0011	0.0452	0.0480	**0.0357**
MAE_Lamp	0.0011	0.0008	**0.0007**	0.0294	**0.0265**	0.0388

between the occupants' preferences and advanced building automation systems. Another aspect could be the comparison of predictions when inputs from the HVAC systems are available. There are still some open research questions and scope for further investigation about the design of such building control systems regarding occupant behaviors and preferences. One aspect is the synergy between the building control systems, with the occupant preferences regarding their comfort and energy usage. Another aspect is the flexibility or robustness of those control systems towards occupant behaviors, and how well they can adapt to those behaviors.

12. Conclusion

This work was successful in performing a case study focused on MEL prediction through deep learning methods. The data comprised of the MEL consumption of each occupant and the associated device-utilization. This data was used to develop a baseline LSTM model for MEL predictions, which was suitably efficient for predicting consumption patterns over a shorter period. Furthermore, the study compared the results from this baseline model with two state-of-the-art deep learning models (Bi-LSTM and GRU) and highlighted the evaluation results of each model. The experimental results on this new dataset collected demonstrate that considering device-utilization patterns and occupant interactions with these devices is fundamental for miscellaneous loads prediction.

In this paper, our comparative analysis results show that in terms of RMSE and MAE, all three deep learning models can achieve the best results depending on the considered devices and also the duration of the prediction required, i.e., short-term or long-term prediction. The low RMSE and MAE values indicate that the deviation of the predicted values from the measured values was quite low. Besides, Bi-LSTM proved to be the marginally more stable model out of those three, in both day-ahead and week-ahead predictions. Usually, due to GRUs less complicated structure and fewer gates, the training phase is faster and converges faster than other models. But, in our case, GRU and Bi-LSTM converge around the same number of epochs. As discussed in Sections 10 and 11, the accuracy of these prediction models can be beneficial to building management systems and can enable the BMS authorities to perform intelligent energy management strategies.

CRediT authorship contribution statement

Anooshmita Das: Conceptualization, Methodology, Investigation, Writing - original draft, Visualization, Validation, Writing - review & editing, Formal analysis, Writing - original draft, Writing - review & editing. **Masab Khalid Annaqeeb:** . **Elie Azar:** Supervision, Writing - review & editing. **Vojislav Novakovic:** Supervision, Writing - review & editing. **Mikkel Baun Kjærgaard:** Supervision, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The University of Southern Denmark supports this work for the HBODEx project (64018-0558). The authors are participating in IEA EBC Annex 79 and acknowledge support from EUDP (64 018-0558).

References

[1] Kamilaris A, Kalluri B, Kondepudi S, Wai TK. A literature survey on measuring energy usage for miscellaneous electric loads in offices and commercial buildings. Renew Sustain Energy Rev 2014;34:536–50.
[2] World energy outlook 2018 – analysis - iea, https://www.iea.org/reports/ world-energy-outlook-2018 [accessed on 04/03/2020].
[3] Syed AM, Hachem C. Net-zero energy design and energy sharing potential of retail-greenhouse complex. J Build Eng 2019;24:100736.
[4] Afshari A, Nikolopoulou C, Martin M. Life-cycle analysis of building retrofits at the urban scale—a case study in united arab emirates. Sustainability 2014;6(1):453–73.
[5] Yan D, Hong T, Dong B, Mahdavi A, D'Oca S, Gaetani I, et al. Iea ebc annex 66: Definition and simulation of occupant behavior in buildings. Energy Build 2017;156:258–70.
[6] Zhang Y, Bai X, Mills FP, Pezzey JC. Rethinking the role of occupant behavior in building energy performance: a review. Energy Build 2018;172:279–94.
[7] Roth K, Mckenney K, Paetsch C, Ponoum R. Us residential miscellaneous electric loads electricity consumption. In: ACEEE summer study on energy efficiency in buildings; 2008.
[8] B. E. D. Book, Us department of energy-office of energy efficiency and renewable energy; 2010.
[9] Ghatikar G, Cheung I, Lanzisera S, Wardell B, Deshpande M, Ugarkar J. Miscellaneous and electronic loads energy efficiency opportunities for commercial buildings: a collaborative study by the United States and India, Tech. rep., Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States); 2013.
[10] Raviraj K, Gupta N, Shet HN. Analysis of measures to improve energy performance of a commercial building by energy modeling. 2016 Online international conference on green engineering and technologies (IC-GET). IEEE; 2016. p. 1–4.
[11] King DJ. Residential dc power bus. ASHRAE J 2010;52(9):73.
[12] Burgett JM, Chini AR. Using building and occupant characteristics to predict residential residual miscellaneous electrical loads: a comparison between an asset label and an occupant-based operational model for homes in florida. J Build Perform Simul 2016;9(1):84–100.
[13] A. Mahdavi, F. Tahmasebi, Predictive models of electrical energy use in office buildings due to plug loads.
[14] Hargan MR. Ashrae standards committee; 2001–2002.
[15] Mahdavi A, Tahmasebi F, Kayalar M. Prediction of plug loads in office buildings: simplified and probabilistic methods. Energy Build 2016;129:322–9.
[16] Wang Z, Hong T, Piette MA. Data fusion in predicting internal heat gains for office buildings through a deep learning approach. Appl Energy 2019;240:386–98.
[17] Yan D, O'Brien W, Hong T, Feng X, Gunay HB, Tahmasebi F, et al. Occupant behavior modeling for building performance simulation: Current state and future challenges. Energy Build 2015;107:264–78.
[18] Menezes AC, Cripps A, Buswell RA, Bouchlaghem D. Benchmarking small power energy consumption in office buildings in the United Kingdom: a review of data published in cibse guide f. Build Serv Eng Res Technol 2013;34(1):73–86.
[19] Menezes A, Cripps A, Buswell RA, Wright J, Bouchlaghem D. Estimating the energy consumption and power demand of small power equipment in office buildings. Energy Build 2014;75:199–209.
[20] Gunay HB, O'Brien W, Beausoleil-Morrison I, Gilani S. Modeling plug-in equipment load patterns in private office spaces. Energy Build 2016;121:234–49.
[21] Marino DL, Amarasinghe K, Manic M. Building energy load forecasting using deep neural networks. In: IECON 2016-42nd annual conference of the IEEE industrial electronics society. IEEE; 2016. p. 7046–51.
[22] Rahman A, Srikumar V, Smith AD. Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks. Appl Energy 2018;212:372–85.
[23] Arahal MR, Cepeda A, Camacho EF. Input variable selection for forecasting models. IFAC Proc Volumes 2002;35(1):463–8.
[24] Fan C, Xiao F, Zhao Y. A short-term building cooling load prediction method using deep learning algorithms. Appl Energy 2017;195:222–33.
[25] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997;9(8):1735–80.
[26] Lee W, Yik F, Jones P, Burnett J. Energy saving by realistic design data for commercial buildings in Hong Kong. Appl Energy 2001;70(1):59–75.
[27] Ren Z, Foliente G, Chan W-Y, Chen D, Ambrose M, Paevere P. A model for predicting household end-use energy consumption and greenhouse gas emissions in australia. Int J Sustain Build Technol Urban Develop 2013;4(3):210–28.
[28] Wang L, Kubichek R, Zhou X. Adaptive learning based data-driven models for predicting hourly building energy use. Energy Build 2018;159:454–61.
[29] Yalcintas M. Energy-savings predictions for building-equipment retrofits. Energy Build 2008;40(12):2111–20.
[30] Kottek M, Grieser J, Beck C, Rudolf B, Rubel F. World map of the köppengeiger climate classification updated. Meteorol Z 2006;15(3):259–63.
[31] Home page - occupeye® | iot, https://www.occupeye.com/ [accessed on 01/02/2020].
[32] Plugwise • plug, https://www.plugwise.com/en_US/products/plug?fbclid=IwAR3M8MVzhDjdzcATHEg2NijameXfjPAcLZ0M-U58u0GQAAkJPs7sCKqXkvE [accessed on 01/03/2020].
[33] Graves A, Schmidhuber J. Framewise phoneme classification with bidirectional lstm networks. In: Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005, vol. 4. IEEE; 2005. p. 2047–52.
[34] Chung J, Gulcehre C, Cho K, Bengio Y. Gated feedback recurrent neural networks. International conference on machine learning. 2015. p. 2067–75.
[35] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res 2014;15(1):1929–58.
[36] Kingma DP, Ba J. Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980.
[37] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift, arXiv preprint arXiv:1502.03167.
[38] Abushakra B, Sreshthaputra A, Haberl J, Claridge D. Compilation of diversity factors and schedules for energy and cooling load calculations. Ashrae research project 1093-rp, final report.
[39] Davis III JA, Nutter DW. Occupancy diversity factors for common university building types. Energy Build 2010;42(9):1543–51.
[40] Jain RK, Taylor JE, Culligan PJ. Investigating the impact eco-feedback information representation has on building occupant energy consumption behavior and savings. Energy Build 2013;64:408–14.

PAPER 8

Das A, Dziedzic J W, **Annaqeeb M K**, Novakovic V, Kjærgaard M B. Human Activity Recognition Using Sensor Fusion and Deep Learning Methods. *Submitted to IMWUT.*

This paper is submitted for publication and is therefore not included.

PAPER 9

Das A, **Annaqeeb M K**, Schwee J H, Dziedzic J W, Novakovic V, Kjærgaard M B. Sequential Activity Recognition and Privacy Implications Using Fusion and Deep Learning Methods Inside A Smart Living Lab. *Submitted to Information Fusion*.

This paper is submitted for publication and is therefore not included.

ISBN 978-82-326-7934-8 (printed ver.)
ISBN 978-82-326-7933-1 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

NTNU
Norwegian University of
Science and Technology