FEED-FORWARD NEURAL NETWORK (FFNN) BASED OPTIMIZATION OF AIR HANDLING UNITS: A STATE-OF-THE-ART DATA-DRIVEN DEMAND-CONTROLLED VENTILATION STRATEGY

A Thesis

Submitted to the Faculty

of

Purdue University

by

Mehdi Momeni

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science in Mechanical Engineering

August 2020

Purdue University

Indianapolis, Indiana

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF THESIS APPROVAL

Dr. Ali Razban, Chair

Department of Mechanical and Energy Engineering

Dr. Jie Chen, Member

Department of Mechanical and Energy Engineering

Dr. Eric Adams, Member

Department of Mechanical and Energy Engineering

Approved by:

Dr. Sohel Anwar

Chair of the Graduate Program

Dedicated to Science ...

ACKNOWLEDGMENTS

Foremost, I would like to express my sincere gratitude to my advisors professor Ali Razban and professor Jie Chen for the continuous support of my master's study and my research, for the patience, motivation, enthusiasm, and their immense knowledge.

Besides my advisors, I would like to thank Dr. Eric Adam, my thesis committee member, for his encouragement and insightful comments.

My appreciation to Dr. David Goodman, the IAC director, for his welcoming attitude towards my questions and his precious time.

My sincere thanks also goes to Ms. Holly Thomas, the campus facility manager, for providing the initial research materials and her valuable guidance.

Finally, I thank my fellow labmate, Allen Wu, for the stimulating discussions and sharing his knowledge.

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ASHRAE	American Society of Heating, Refrigerating and Air-Conditioning Engineers
HVAC	Heating, Ventilation and Air Conditioning
AHU	Air Handing Unit
VAV	Variable Air Volume
CAV	Constant Air Volume
BMS	Building Management System
BAS	Building Automation System
BACnet	Building Automation and Control Networks
IoT	Internet of Things
PIR	Passive Infrared Sensors
DCV	Demand-Controlled Ventilation
IAQ	Indoor Air Quality
ML	Machine Learning
ANN	Artificial Neural Network
FFNN	Feed-Forward Neural Network
PMV	Predicted Mean Vote
PPD	Predicted Percentage of Dissatisfied
GPR	Gaussian Process Regression
GRG	Generalized Reduced Gradient
MLR	Multiple Linear Regression
VAC	Volts of Alternating Current

ABSTRACT

Momeni, Mehdi, M.S.M.E., Purdue University, August 2020. Feed-Forward Neural Network (FFNN) Based Optimization Of Air Handling Units: A State-Of-The-Art Data-Driven Demand-Controlled Ventilation Strategy. Major Professor: Ali Razban.

Heating, ventilation and air conditioning systems (HVAC) are the single largest consumer of energy in commercial and residential sectors. Minimizing its energy consumption without compromising indoor air quality (IAQ) and thermal comfort would result in environmental and financial benefits. Currently, most buildings still utilize constant air volume (CAV) systems with on/off control to meet the thermal loads. Such systems, without any consideration of occupancy, may ventilate a zone excessively and result in energy waste. Previous studies showed that CO_2-based demand-controlled ventilation (DCV) methods are the most widely used strategies to determine the optimal level of supply air volume. However, conventional CO_2 mass balanced models do not yield an optimal estimation accuracy. In this study, feed-forward neural network algorithm (FFNN) was proposed to estimate the zone occupancy using CO_2 concentrations, observed occupancy data and the zone schedule. The occupancy prediction result was then utilized to optimize supply fan operation of the air handling unit (AHU) associated with the zone. IAQ and thermal comfort standards were also taken into consideration as the active constraints of this optimization. As for the validation, the experiment was carried out in an auditorium located on a university campus. The results revealed that utilizing neural network occupancy estimation model can reduce the daily ventilation energy by 74.2% when compared to the current on/off control.

1. INTRODUCTION

The optimized performance of heating, ventilation and air conditioning (HVAC) systems is critically important. Studies showed that people in the US and Europe spend on average 85% to 90% of their time indoors [1]. Furthermore, 47.7% and 51% of energy consumption in residential and commercial buildings, respectively, are allocated to HVAC systems. Given the substantial effects of HVAC systems on energy use and human comfort, research efforts have been mainly focused on (1) the minimization of HVAC energy use without sacrificing thermal comfort or (2) the optimization of occupants' thermal comfort [2]. Previous HVAC related studies demonstrated that current system operations suffer from a number of major shortcomings: conditioning unoccupied spaces [3], assuming maximum zone occupancy [4], and over-conditioning of buildings regardless of occupants' perspectives [5]. These sub-optimal circumstances are primarily due to the fact that conventional operating modes of HVAC systems do not take the occupants' dynamic into account. These conventional operating modes include (1) fixed operating schedules assuming full occupancy [6] and (2) a single-point temperature measurement of a thermal zone (i.e. multiple zones conditioned by one AHU) [7]. That is, there has been a lack of context-aware information delivery and means of interaction between occupants and HVAC systems. Thermostats, in commercial buildings, are generally inaccessible for occupants' intervention [8] since facility managers mostly set the temperature setpoints without consideration of occupants' thermal feedback [9]. Post-occupancy evaluation strategies, as a means of integrating occupants' feedback, have been rarely implemented [10]. As a consequence of these limitations, a higher rate of thermal dissatisfaction occurs in commercial sectors compared to that of residential buildings [11].

Overall, previous studies have mainly investigated several physical and data-driven occupancy estimation models, disregarding the implementation outcomes on building

energy management. Added to this, there have been a plethora of research works focusing on DCV strategies in which only CO_2 concentrations were utilized, but not the number of indoor occupants. This study aimed to quantify not only the energy savings but also IAQ and thermal comfort by implementing a dynamic per-occupant DCV strategy, using machine learning techniques based on experimental data. High capacity single-zone areas were the focus of this study since the best potential applications can be found in spaces with highly variable occupancy, namely, auditoriums [12].

1.1 Thesis Structure

Figure 1.1 illustrates the structure of this study. Chapter 1 provides the background information, problem statement and the objectives of this thesis. In Chapter 2, an overview of energy saving opportunities through occupancy prediction and different occupancy estimation methods are discussed. Chapter 3 outlines the rationale behind the methods utilized to solve the problem stated in the first two chapters. It describes the test environment, HVAC system sequence of operation and schedule, data acquisition procedure, and modeling methods. Chapter 4 comprehensively discusses physical and statistical occupancy prediction models, different demand-controlled ventilation strategies, the effects of each strategy on ventilation as well as indoor air quality. Thermal comfort analyses are included in Chapter 5. Chapter 6, finally, documents the energy savings, thorough conclusion of this research, and potential future works.

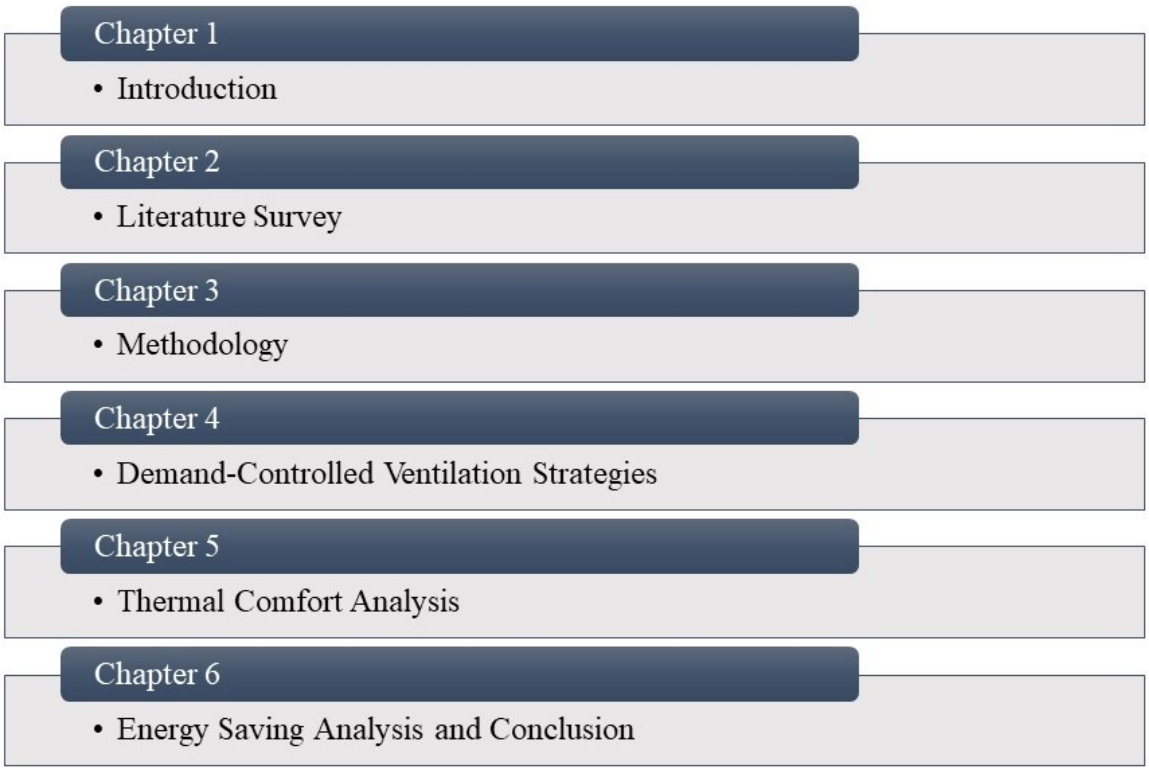

Fig. 1.1.: Thesis structure

2. LITERATURE SURVEY

This chapter gives an overview of the major topics related to the study of dynamic per-occupant demand-controlled ventilation. It is divided into two subsections as follows: **section 1** presents the background and previous studies to prove the necessity of occupancy detection in HVAC energy savings. **Section 2** comprehensively reviews the current techniques for occupancy detection using different sensors and strategies.

2.1 Occupancy-Based Energy Saving Strategies

Building management systems (BMS) can utilize occupancy information to decrease the energy consumption of lighting, HVAC systems, and even other integrated building systems. Occupancy information allows building systems to operate proportional to the number of occupants and consequently to optimize the building energy management as well as thermal comfort by optimal active and passive control of heating, cooling, ventilation, lighting, and other possible systems [13]. Many researchers have been studying occupancy detection methods and its applications in building management. Yang and Becerik-Gerber demonstrated that occupancy-based HVAC schedules could result in energy savings of up to 9% [14]. A low-cost system was proposed by Agarwal et al. to accurately detect the presence and absence of occupants in office areas. They further showed that energy savings between 10% and 15% are achievable with the detected occupancy information [15]. L. James Lo and Atila Novoselac evaluated an occupancy-based air-conditioning system through a case study. Their CFD and energy simulation showed that it is feasible to produce a localized airflow in an open office space and to create relatively isolated zones in order to save approximately 12% of total cooling energy [16]. Varick L. Erickson et al. collected occupancy data using a sensor network. They developed a statistical model of

the building temporal occupancy based on an in-homogeneous Markov chain, which was estimated from the collected occupancy information. Compared to the baseline, the average annual savings of 42% were attained by implementing their real-time occupancy method [17]. Anil Aswani et al. studied the transient and steady-state electrical characteristics of the heat pump in a special platform called BRITE at UC Berkeley. They identified a dynamical model of the system, explained the impact of occupants on the dynamics, and implemented a learning-based model predictive control (MPC) scheme that estimates occupancy using only temperature measurements. Experiment results revealed that learning-based MPC saved an estimated 30% to 70% of energy compared to the two-position control . More advanced occupancy estimation, the authors claimed, would yield further reductions [18]. To find out the energy-saving opportunities, Siddharth Goyal Herbert compared the baseline controller of an HVAC system to three different proposed algorithms, MOBS (Measured Occupancy Based Setback), MOBO (Measured Occupancy Based Optimal) and POBO (Predicted Occupancy Based Optimal). While the proposed MOBS and MOBO controllers depend upon occupancy information, POBO controller requires occupancy predictions. Simulation results showed in their work that all three controllers lead to solid improvements in energy savings (about 50% on average depending on the zone type, weather, climate, etc.) with trivial impacts on IAQ and thermal comfort [19]. Bharathan Balaji and his team designed and implemented an HVAC actuation system that benefits from the existing WiFi infrastructure and tracks the occupants carrying WiFi-enabled smartphones. They proved that this system provides an accuracy of 86% leading to HVAC electrical energy savings of 17.8% in comparison to the static scheduling based control [20]. The authors of [21] investigated the potential of using occupancy information through a case study at ETH Zurich, Switzerland. An MPC controller, which controls the building based on a pre-defined fixed occupancy schedule, was used in their study as a benchmark. A comparison of energy consumption of this benchmark was made to three other control strategies. They first chose the mentioned MPC controller that uses the same

schedule for control as the benchmark. Unlike having a fixed schedule, however, this MPC controller turns off the lights in case of vacancy. Second, the same MPC controller that not only turns off the lights, but also the ventilation while the building is unoccupied; and finally, the same MPC controller by which a perfect prediction of the upcoming occupancy can occur. The simulations with homogeneous occupancy in this study showed a saving potential of up to 34% in case of average vacancy and occupancy intervals of 5 and 10 days, respectively. They finally concluded that such savings, to a large degree, can be accomplished by prediction through instantaneous measurement of the occupancy status along with a schedule. Samuel R. West et al. also presented a novel method of optimizing the operation of commercial HVAC systems using model predictive control, to minimize a weighted combination of operating costs, greenhouse gas emissions and occupant thermal comfort. They put the system to the test in two different buildings in Newcastle and Melbourne over the winter 2011. In comparison with standard building management and control systems, the performance results obtained from the Newcastle Office Wing showed an average energy reduction of 19% during OptiCOOL operation without compromising occupants' comfort. Results from the Melbourne trial, on the other hand, indicated a 32% thermal energy reduction while considering comfort standards [22]. Justin R. Dobbs and Brandon M. suggested an occupancy-predicting control algorithm, which integrates the building thermal properties, local weather predictions, and a self-tuning stochastic occupancy model to reduce energy consumption while maintaining comfort level. They afterwards outlined the results of three scenarios with regards to energy savings and discomfort. (1) An exclusively occupancy-triggered controller, (2) a scheduled controller augmented by occupancy triggering, and (3) an online trained occupancy-predicting controller with one week of pre-training. Their results ended up with up to 19% energy savings compared to the scheduled controller and significantly lower discomfort than the occupancy-triggered controller [23]. Jonathan Brooks et al. developed an occupancy-based feedback control algorithm that is applicable to "under-actuated" cases in which numerous rooms are served by the same HVAC system. They

examined their proposed control logic called $MOBS^{ua}$ in order to quantify the energy savings while maintaining thermal comfort and IAQ metrics. The $MOBS^{ua}$ controller turned out to offer notable energy savings for under-actuated zones, accounting for 29% to 80% [24]. Hencey Gunay et al. suggested two self-adaptive models that can predict the temperature of a perimeter office space. The models were designed to be fed by the outdoor air temperature, indoor light intensity readings, occupancy status, radiant panel heaters' state, discharge airflow rate and temperature. In contrast to the default control scheme, results of this implementation indicated that the duration of the weekday temperature setback periods increased by over 50% for both heating and cooling [25]. In [26], Fulin Wang et al. proposed a predictive control algorithm for building environmental control based on occupant number detected by video data and CO_2 concentration. They showed an energy reduction of approximately 40% for buildings without compromising thermal comfort and air quality. Alfonso Capozzoli et al. used anonymous occupancy data of an office building for a monitoring period of four months to create an occupancy-based schedule. They carried out trials on the resulting schedule through an energy simulation approach by which a model can be calibrated with real energy consumption data. Concerning the three tested thermal zones in the building, 27 hours of the HVAC operation was saved during weekday in comparison to the fixed schedule. In other words, the overall energy savings of 14% were resulted [27]. Another study by Yuzhen Peng and his team at ETH Zurich, Switzerland analyzed a seven-month period of occupancy data, which was based on motion signals collected from six offices with ten occupants. By running an occupancy analysis, a learning-based demand-driven control strategy was proposed for sensible cooling. It could predict occupants' next presence and the presence duration of the remainder of a day by learning human behavior from the past and current days. The predicted occupancy information was then utilized indirectly to infer setback temperature setpoints. During this study, they observed that energy saving potentials in an individual office were inversely correlated to its occupancy rate. An energy saving of 20.3%, finally, was achieved as compared to the benchmark [28]. The authors of [29]

studied the effect of spatial variations in mean radiant temperature (MRT) and occupancy on HVAC energy consumption and thermal comfort. This research was carried out in an open-plan office space with multiple air handling units. A predictive control method was developed that reaches the optimum temperature setpoint by solving an optimization problem. It minimizes HVAC energy consumption through modulating the fan speed such that an acceptable thermal comfort and adequate outside air intake are its constraints. Throughout the entire period, this proposed control reached an average savings of 15% over a PID control that assumes uniform spatial occupancy distribution and 12% over a PID based strategy that uses actual spatial occupancy information.

2.2 Occupancy Detection and Estimation Techniques

In this section, a comprehensive overview of current options for building occupancy estimation and detection with different sensor categories is performed.

2.2.1 Passive Infrared Sensors (PIR)

PIR sensors can detect the infrared light radiations caused by the movement of subjects. Thus, they could be used to detect the motion and presence of occupants [30]. Dodier et al. proposed a PIR sensor network to detect the occupancy [31]. They deployed three independent PIR sensors to monitor the occupancy presence. Bayesian probability theory was then exploit to infer the presence and absence of occupants in a zone. Wahl et al. proposed a distributed PIR-based approach which could detect the moving directions of occupants [32]. A direction-based algorithm along with probabilistic distance-based algorithm were utilized to perceive the moving directions and consequently counting occupants. PIR sensors have some pros and cons; they are low-cost and simple to use in different environments. However, PIR sensors are not able to detect static occupants since they are motion-based sensors [30].

2.2.2 Carbon Dioxide Sensors (CO_2)

Human occupants contribute to rising indoor carbon dioxide (CO_2) level which can also indicate how much ventilation is required [33]. This means that human occupants directly affect the HVAC energy consumption as well as indoor air quality (IAQ). Many studies have reported the use of CO_2 concentrations for occupancy prediction and estimation. Davide Cali et al. proposed an algorithm for occupancy detection based on the mass balance equation of indoor CO_2 concentrations. The algorithm provided correct presence profile up to 95.8% of the time while the exact number of occupants was identified with the maximum accuracy of 80.6% [34]. Chaoyang Jiang et al. developed a dynamic model of the occupancy level with which they could estimate the real-time number of indoor occupants based on the CO_2 measurements. They showed the accuracy of up to 94% with a tolerance of four occupants in an office room [35]. In another study, T. Pedersen et al. proposed a novel plug-and-play occupancy detection method in which they utilized multiple sensory data including CO_2 sensors. Testifying the proposed method in a single room and a three-bedroom apartment resulted in a maximum accuracy of 98% and 78%, respectively [36]. S. Ryu and H. Moon developed a machine-learning occupancy prediction model using indoor and outdoor CO_2 concentration data. By different observation states, they achieved the prediction accuracies ranging from 85% to 93.2% [37]. Many literatures have proven that CO_2-based occupancy detection for demand-controlled ventilation (DCV) systems is a promising approach as CO_2 concentration is a proxy indicator of occupant-related contaminants [38–40]. Further, CO_2 sensors are affordable, compact, non-intrusive and non-terminal-based. As they are also conventionally integrated into standard HVAC systems, almost no additional investments are required in current infrastructures [33].

2.2.3 Camera

Cameras could also be utilized in occupancy detection and estimation due to their high precision. Erickson et al. applied a group of cameras to build an occupancy model [41]. Their camera system captured and processed the image data of individuals in order to determine whether the transition and direction occur. They afterwards implemented the developed occupancy model into building management systems to evaluate energy savings. With a multi-camera people tracking, Fleuret et al. proposed an algorithm to reliably track multiple occupants in a complex environment and provide the location and number of individuals [42]. The authors of [43] came up with a real-time head detection and occupancy estimation algorithm which analyzes the occupants' videos in an office environment. They finally reported the accuracy of up to 95.3% in occupancy estimations. Overall, occupancy detection and estimation using cameras can offer accurate results; however, this method suffers from a number of limitations including human privacy violations, high computational complexity and the effect of environmental illumination in accuracy [30].

2.2.4 Wireless Fidelity

Since WiFi signals are highly likely available in indoor environments, WiFi occupancy tracking method could be a good candidate. Therefore, the number of smartphones carried by occupants can be inferred as the number of occupants with an acceptable accuracy. Balaji et al. designed and implemented an occupancy-based HVAC controller in which the existing WiFi infrastructures within commercial buildings were deployed [44]. To lower inaccuracies in occupancy sensing, they utilized metadata information of the occupants, access points and the HVAC zones in the building. This resulted in the precision of 86% in occupancy detection within the office areas, with merely 6.2% false negative errors. In another research, Wang et al. proposed a WiFi scanning occupancy detection method [45]. They introduced dynamic Markov time-window inference (DMTWI) approach which formulates the

occupancy dynamics as a Markov process. Their work was finally validated through an experiment using WiFi probes and cameras to collect the ground truth information. In WiFi-based occupancy estimation methods, it is presumed that each occupant would carry a smartphone with an enabled WiFi. Yet, this is not always the case in real practice. Occupants might also carry multiple WiFi-enabled devices which causes inaccuracies [30].

2.2.5 Sensor Fusion

Each sensor has a number of unique properties and limitations for occupancy detection and estimation. The fusion of different types of sensors can compensate for the limitations of a standalone sensor and therefore enhances the performance [30]. This study utilized not only CO_2 sensors but also the auditorium occupancy schedule and ventilation rates for both detection and estimation.

2.3 Closure on The Chapter

Section 1 in this chapter summarized the impact of occupancy information in building management systems. The energy saving strategies, which incorporated the occupancy information, were also discussed. It is quite obvious that occupancy detection can lead to considerable energy savings, ranging from 9% to 70% depending on the system settings. Not only could the occupancy detection result in energy savings, but also it can contribute to thermal comfort and indoor air quality as it prevents from over and under-ventilation. Even though some research works did not take thermal comfort into account, most of energy optimization algorithms have been designed with comfort constraints. However, the correlation between supply air flow rate, zone temperature and the number of occupants seem to be missing. The proposed optimization strategy in this research mainly relied on the interactions of mentioned variables in order to generate the trade-off between energy savings, IAQ and human comfort thresholds. Due to the importance of occupancy, section 2 attempted to cover

different major techniques of occupancy detection and estimation. Each method was discussed based on some evaluation metrics including existing infrastructure, cost, privacy issue, detection and estimation accuracy, and limitations. As an enclosure on this chapter, the following table represents a comprehensive comparison on all the discussed occupancy sensors along with . According to the section 2.2.2 and table 2.1, it was decided to exploit CO_2 sensors to predict occupancy in this study.

Table 2.1.: Comparison of different occupancy detection and estimation methods

Sensor	**Existing Infrastructure**	**Cost**	**Detection Accuracy**	**Estimation Accuracy**	**Privacy Issue**	**Limitations**
PIR	No	Low	High	Low	No	Miss static occupants
CO2	Yes	No	Medium	Medium	No	Delay detection
Camera	No	High	High	High	Yes	Illumination condition
WiFi	Yes	No	High	High	Partial	Need to turn on WiFi
Sensor Fusion	Partial	Medium	Very High	Very High	No	Use of different sensors

3. METHODOLOGY

Figure 3.1 represents the methodology used in this research. Each section is discussed in the following chapters.

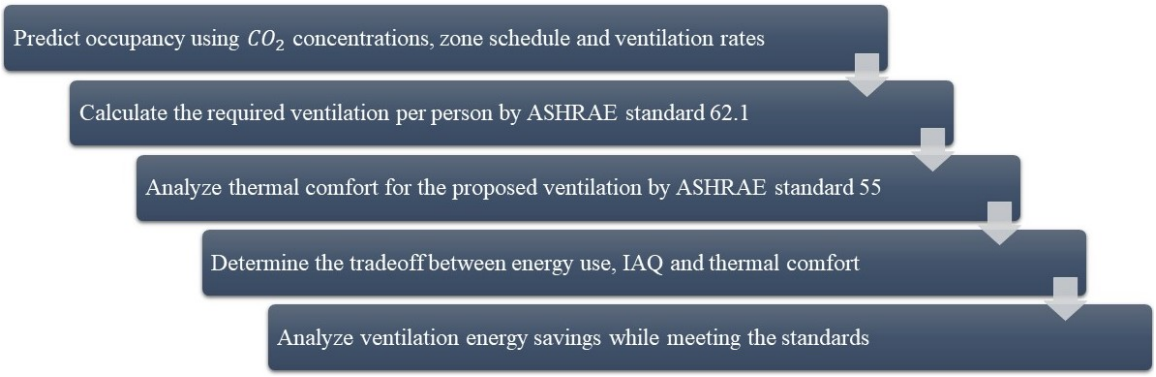

Fig. 3.1.: Methodology

3.1 Test Environment

The test environment used in this study was an auditorium located at the VanNuys Medical Science Building, IUPUI campus. This is a five-story building which includes auditoriums, office areas, and mainly medical laboratories. The targeted zone is situated at the center of this building. Figures 3.2, 3.3, and 3.4 depict the floor layout as well as 3D layouts of the zone with different viewpoints. It is noteworthy that this auditorium is designed for the maximum seating capacity of 182 and mostly holds scheduled lecture classes and seminars. It has the volume and floor area of 1,400 m^3 and 306 m^2, respectively. There are three entrances to the zone, as portrayed in the layouts, whereas the window per wall ratio is zero. The conditioned air is delivered by six supply diffusers mounted on the suspended ceiling while the zone air is returned and relieved by one sidewall along with four ceiling-mounted square extract grilles.

Per ASHRAE standard 62.1-2016 [46], CO_2 concentrations were measured not only in the zone, but inside the supply and return ducts. Red marks shown in the figure 3.3 are the exact location of the CO_2 sensors.

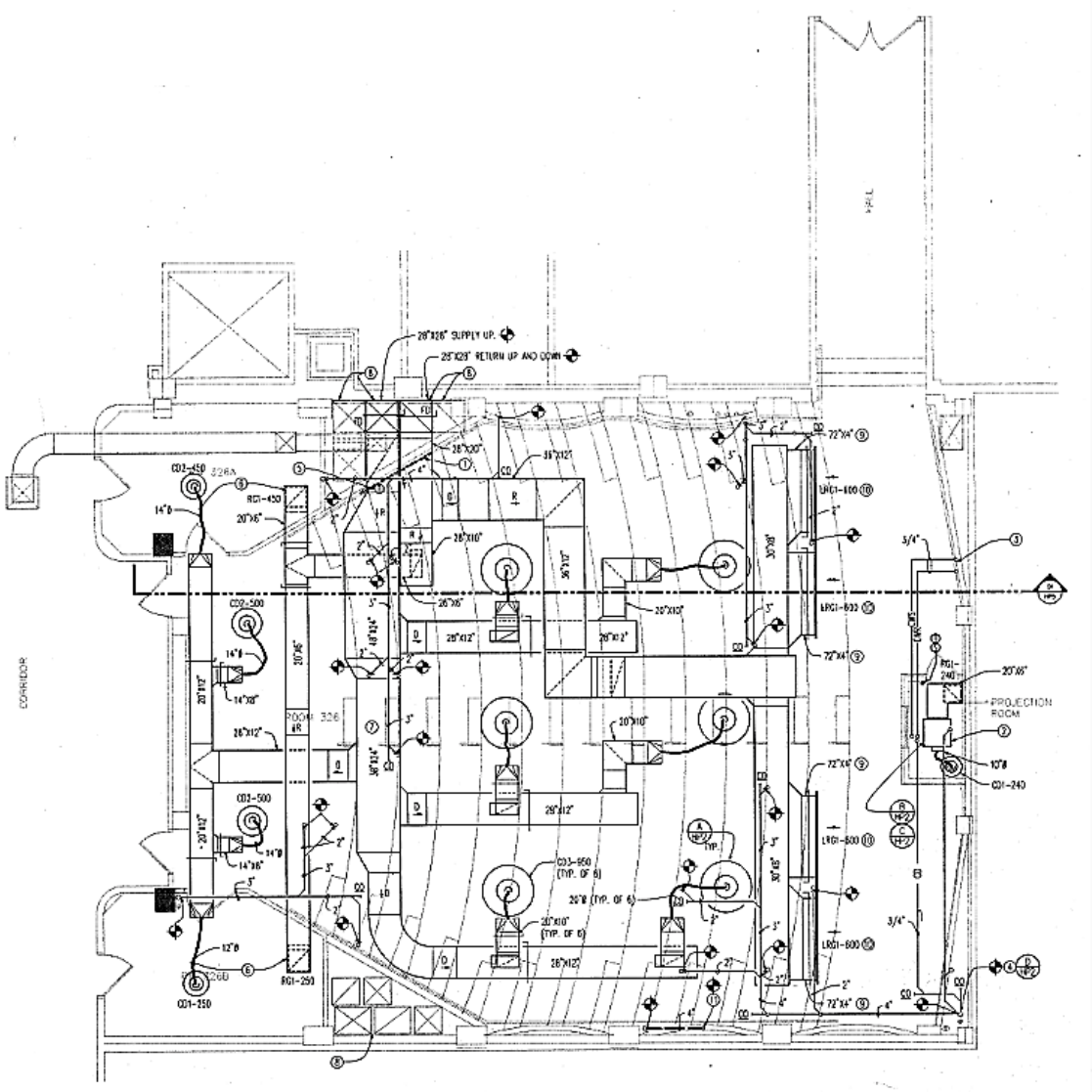

Fig. 3.2.: The Auditorium floor layout

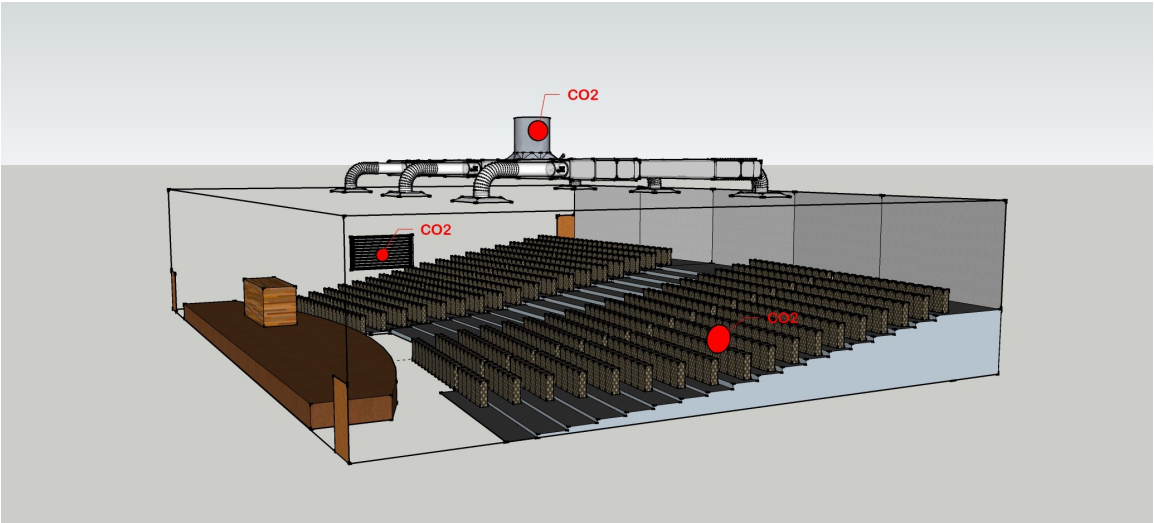

Fig. 3.3.: An overview of the test environment

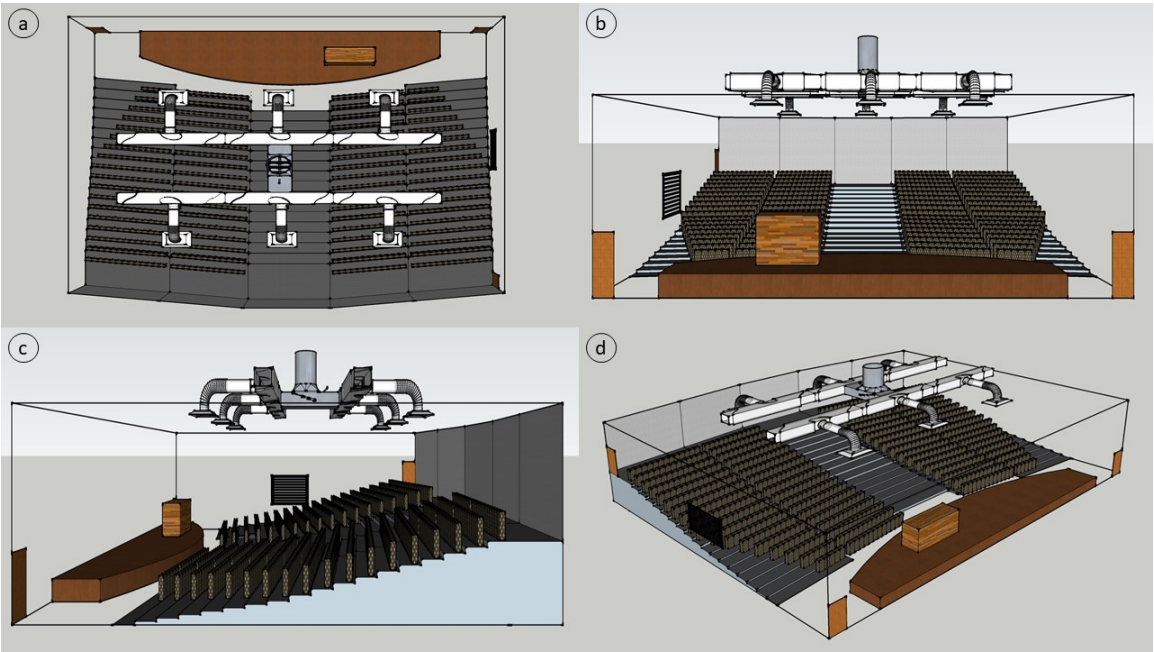

Fig. 3.4.: Different viewpoints of the test environment: a) top, b) front, c) Side, d) iso

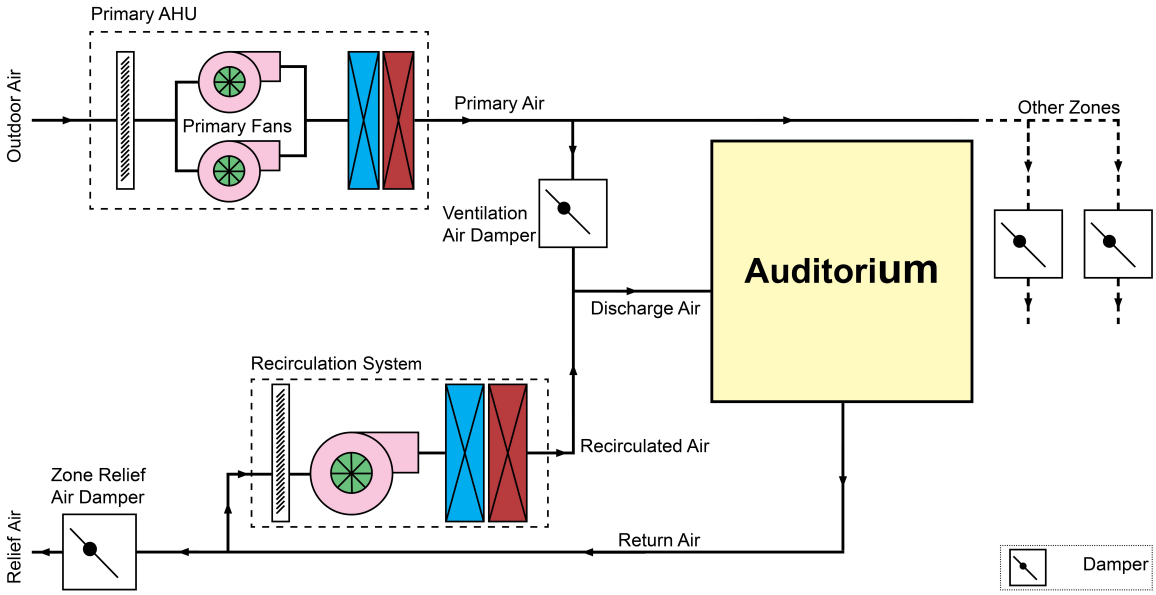

Fig. 3.5.: Schematic of the HVAC system serving the auditorium

3.2 HVAC System Sequence of Operation

Figure 3.5 is the schematic diagram of the HVAC system associated with the auditorium. It comprises a primary air handling unit as well as a recirculation system. The primary AHU is a variable air volume (VAV) system containing two identical supply fans, while the recirculation system is a constant air volume (CAV) type. The primary AHU receives 100% outdoor air and distributes to multiple zones including the auditorium. Due to the presence of laboratories, there is no return duct and fan associated with the primary AHU. In spite of the VAV type, this primary AHU is currently operating between two fixed control positions, occupied and unoccupied mode. The primary air is modulated by a ventilation air damper, supplied to the zone as fresh ventilation air, and then mixed with the conditioned air coming from the recirculation system. The recirculation system fully operates with no shutdown (CAV type), while the primary fans are on full operation during occupied mode and on 70% operation during unoccupied mode. The occupied mode schedule is set from 6:00 a.m. to 9:00 p.m. excluding weekends and holidays. During occupied mode

the ventilation air damper is fully open, introducing about 1,000 L/s of outdoor air; this fraction of outdoor air, then, gets mixed with the recirculated air so that the discharge air volume of 2800 L/s is achieved in the duct down to the auditorium. The zone temperature, in occupied mode, is allowed to float between 21°C and 24°C (campus standard) with a 23°C setpoint. During unoccupied mode the ventilation air damper position goes down to 0% so that no outdoor air is supplied to the zone. Due to mass balance, the zone relief air damper position is proportional to the ventilation air damper position. Thus, the relief damper also closes when unoccupied mode, meaning that only the recirculation system is in the loop.

It is worth noting that there are no boiler and/or chiller linked to the air handlers as steam and chilled water are purchased upon the university policies. This causes a number of limitations on controlling the water and steam supply systems. That is, the boundaries of this research had been defined based on the degrees of freedom on the air handling units; the experiments were conducted on the primary AHU (VAV) as well as zone's adjacent damper, but not the recirculation system (CAV) as it is only on constant operation.

3.3 Data Acquisition System

Data collection was conducted with the aim of analyzing occupancy level as well as HVAC energy performance. The following sections show details of the employed sensors, locations they were installed and the period of their utilization.

3.3.1 Integrated IoT Sensors

The integrated IoT system, as Figure 3.6 represents, consists of one power/energy meter sensor, three CO_2 sensors, an IoT gateway, the cloud platform and its dashboard. All the integrated sensors are capable of wireless communication via Bluetooth low energy (BLE) technology. The power meter is connected to the primary fans , measuring power factor, active power (kW) and energy consumption (kWh). It can

cover the voltage range of 90-600 VAC with 0.5% accuracy. The CO_2 sensors are self-calibrating non-dispersive infrared (NDIR) type, which measure a range of 0 to 5000 *ppm* with $\mp$5% of reading accuracy. The cloud dashboard is capable of showing both historical and real-time data that can also be downloaded for further processing.

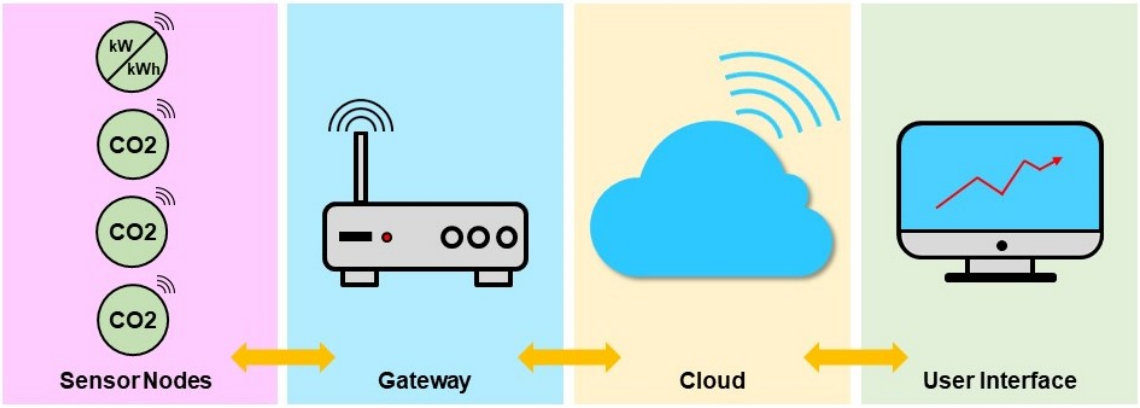

Fig. 3.6.: IoT-based data acquisition system

3.3.2 Building Management System (BMS)

Aside from the IoT platform, the building management system (BMS) was set to store the rest of required data, namely, temperature and airflow rate. As figure 3.7 represents, the utilized building management system is *Metasys* made by *Johnson Controls*. The Metasys user interface provides coordinated control over the commercial building's systems including HVAC, lighting, security and protection systems. As it controls equipment, Metasys also collects critical data, which can be used for decision-making strategies by operators.

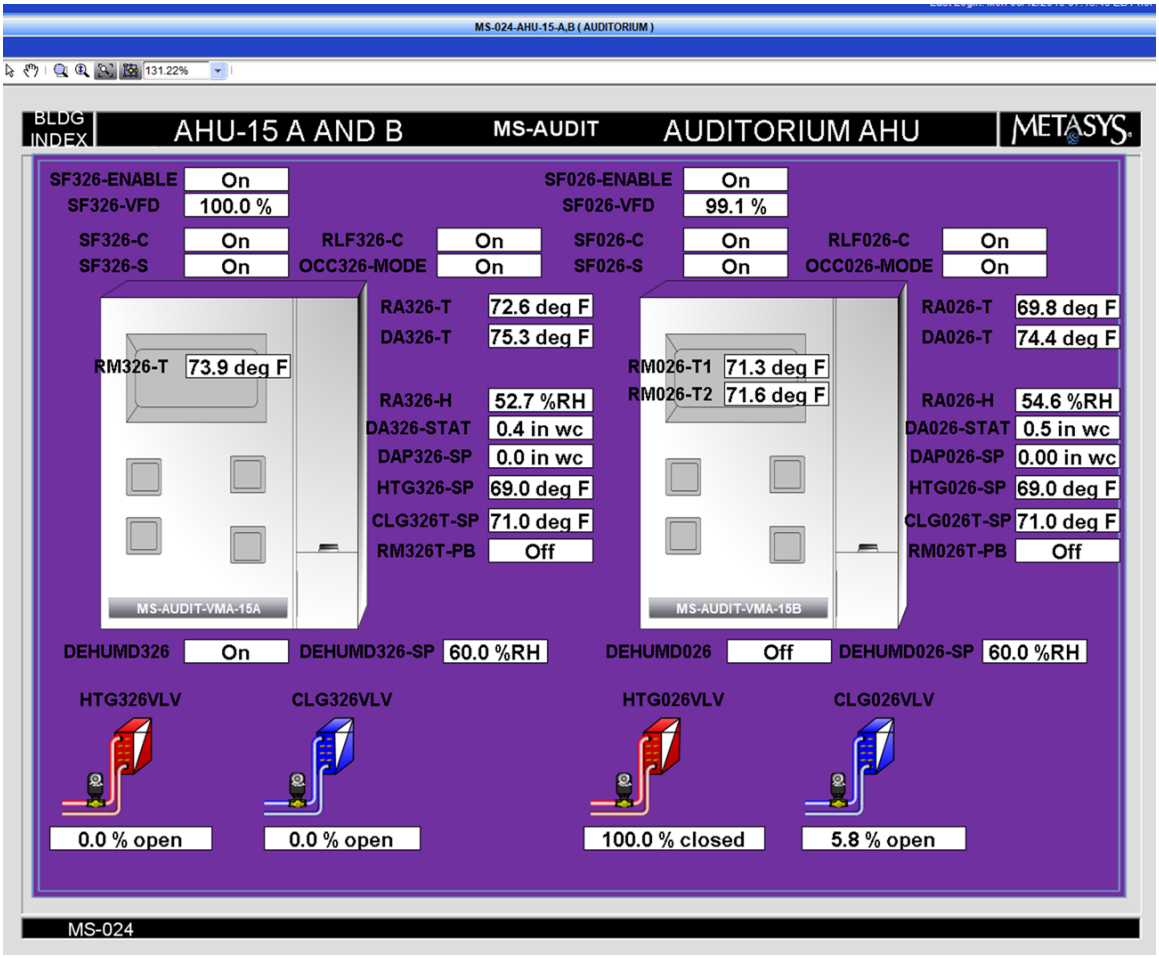

Fig. 3.7.: Metasys software environment

3.3.3 Data Collection

The collected data of both IoT and BMS systems along with the ground truth occupancy readings were eventually lumped together for data analysis. The experimental data were collected and monitored on a five-minute time interval from September 27^{th}, 2019 to January 7^{th}, 2020. The number of occupants in the auditorium was manually counted every 15 minutes, and then cross-checked with the class schedule provided by the university personnel. The time intervals were, indeed, selected based on the variables' rate of change. Table 3.1 indicates the data-set collected for this study.

Table 3.1.: Collected data-set

Data Type	Unit	Sensor Group	Sensor Quantity
CO2	$[ppm]$	IoT	3
Reactive Power	$[kW]$	IoT	1
Airflow Rate	$[L/S]$	BMS	5
Temperature	$[^{\circ}C]$	BMS	7
Damper Position	$[\%]$	BMS	2
Occupancy Count	$[person]$	Manual Observation	-

3.3.4 Sensor Locations

Figure 3.8 shows the location of sensors, both integrated to the system for the sake of this research and the pre-embedded sensors.

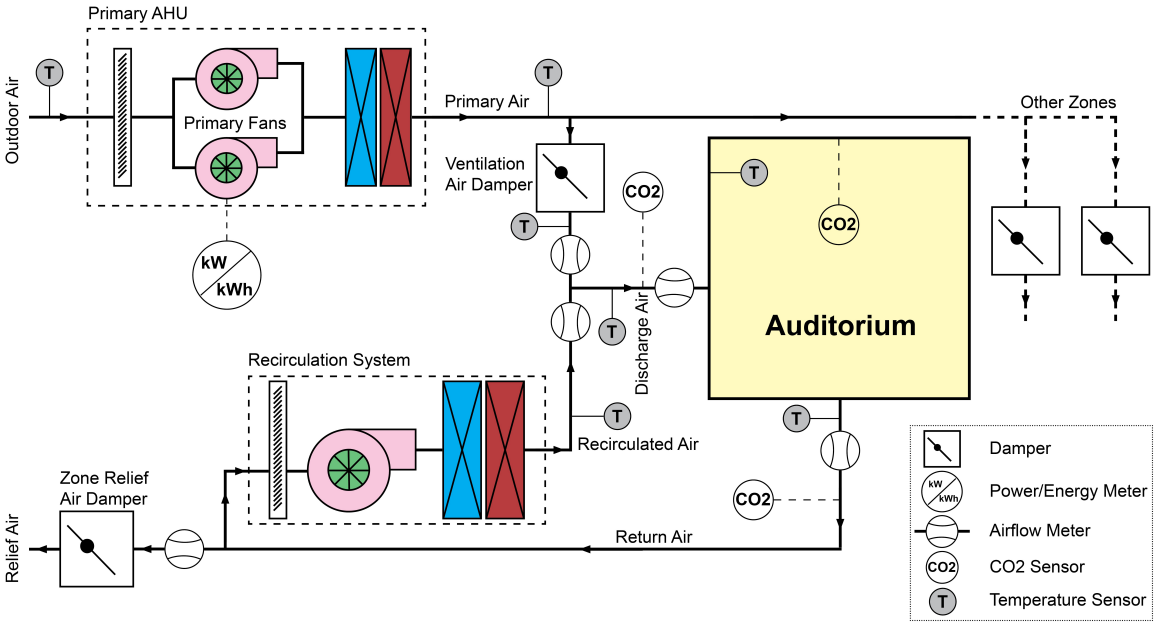

Fig. 3.8.: Sensor locations in the test-bed

3.4 Equipment Specification

As mentioned in section 3.2, the primary AHU has two identical VAV fans which are synced. Tables 3.2 and 3.3 are the the spec sheets of the fans and the electric motors coupled with them.

Table 3.2.: Primary fans' specification

Model	30-AFDW-41
Operating Power	67.9 $[hp]$ / 50.6 $[kW]$
Volume	35,000 $[CFM]$ / 16,518 $[L/s]$
Static Pressure	8.5 $[In.\ WC]$ / 2.1 $[kPa]$
Static Efficiency	69%
RPM	1,868

Table 3.3.: Primary motors' specification

Operating Power	100 $[hp]$ / 74.6 $[kW]$
RPM	1,725
Voltage	460 V
Phase	3
Service Factor	1.15
Efficiency Factor	95.4%
Power Factor	85%
Enclosure	Open Drip Proof (ODP)
Frame	404T

3.5 Modeling Methods

Modeling of HVAC systems is a crucial step towards the appropriate analysis and optimization. HVAC system modeling approaches can be divided into three main categories:

- Physics-based (white box / mathematical / forward)
- Data-driven (black box / empirical / inverse)
- Hybrid (gray box)

These models can be linear or nonlinear, static or dynamic, explicit or implicit, discrete or continuous, deterministic or probabilistic, deductive, inductive or floating [47].

3.5.1 Physics-Based Models

Physics-based models are derived from the fundamental laws, namely, the conservation of energy and mass, heat transfer and momentum. These types of models could be essentially used at the design level where it is requisite to predict and analyze the operation of HVAC system components [47]. The dynamic-based white box models are commonly developed for the processes with higher delay time such as temperature and humidity (e.g. heating/cooling coil model). In contrast, the static models have shown superior performance on the fast moving processes in the HVAC systems (e.g. damper airflow rate model) [48].

3.5.2 Data-Driven Models

Data-driven models are based on the system performance data which can be collected through a real practice in order to establish a relationship between the correlated variables using statistical techniques. Statistical regression and artificial neural

network are the most common examples of data-driven models [47]. Black box modeling is suitable for the existing HVAC systems optimization for which adequate training data are available.

3.5.3 Hybrid Models

Hybrid models could be considered as the combination of physics-based and data-driven models [47]. The structure of a gray box model is formed by physics-based equations while the parameters are estimated from the measured data of the system. This modeling technique is primarily beneficial to the control applications in which a model can be formed as a transfer function or state space [49].

3.6 Closure on The Chapter

This chapter attempted to clearly describe the scope of research and to map out the methodology. The case study along with its specifications were discussed in detail. Data acquisition system, sensors' types and locations were also scrutinized. Modeling methods, as a crucial step towards modeling selection, were explored. It was concluded that Gray-box models have been developed to overcome the shortages of both white-box and black-box models. As for the benefits, higher accuracy, more generalization potential, less complexity and lower computational costs could be named. However, developing hybrid models require more efforts since the process involves the implementation of both governing equations as well as statistical data analysis. Due to the nature of this research, both data-driven and hybrid modeling methods were utilized.

4. DEMAND-CONTROLLED VENTILATION STRATEGIES

Demand-controlled ventilation methods introduced in this chapter can be divided into two categories, dynamic per-occupant and CO_2-based. As the dynamic per-occupant DCV strategies require the number of occupants for implementation, three different occupancy prediction models were first proposed. Proportional control, on the other hand, was also quantified as a CO_2-based DCV strategy. Figure 4.1 represents this division.

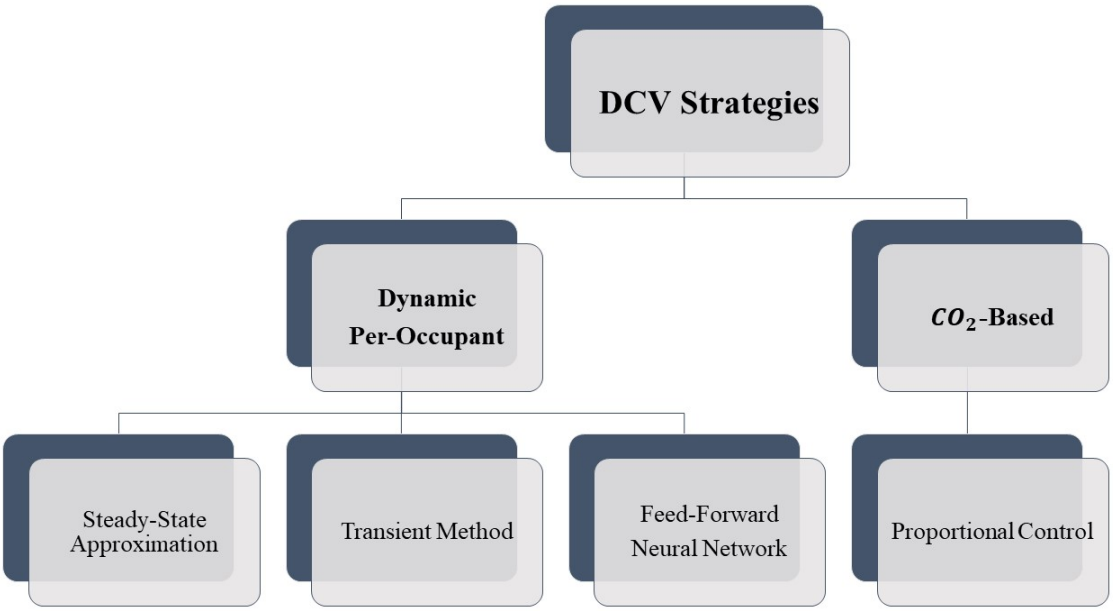

Fig. 4.1.: Demand-controlled ventilation strategy breakdown

With regards to the modeling methods discussed in chapter 3, the steady-state approximation and transient models must be categorized as physics-based models due to the fact that they are directly derived from mass balance equations. In comparison,

the feed-forward neural network model is developed based on the collected data of the system; therefore, it falls under the data-driven modeling approach and is a type of artificial neural networks (ANN).

4.1 Physics-Based Models

For the auditorium space, a mass balance model was used to describe the change of CO_2 within the zone using the equation (4.1) [50].

$$v\frac{dC_z}{dt} = P_zG + V_{dz}C_{dz} - V_rC_r \tag{4.1}$$

where v is the auditorium volume (L), C_z is the CO_2 concentrations in the auditorium $(ppm/10^6)$, P_z is the number of occupants in the auditorium (person), G is the average rate of CO_2 generation per person $(L/s.person)$, V_{dz} is the zone discharge airflow rate to the auditorium (L/s), C_{dz} is the CO_2 concentrations of the zone discharge air $(ppm/10^6)$, V_r is the return airflow rate (L/s), and C_r is the CO_2 concentrations of the return air $(ppm/10^6)$.

Equation (4.2) shows the mass balance at the air handling units.

$$V_{dz}C_{dz} = (V_r - V_e)C_r + V_{oz}C_{oz} \tag{4.2}$$

where V_e is the relief airflow rate (L/s) and V_{oz} is the outdoor airflow rate (L/s). Equation (4.2) assumes the CO_2 concentrations in the relief air equal to that of return air.

Upon substituting equation (4.2) into equation (4.1), the zone CO_2 mass balance equation can be formulated as equation (4.3).

$$v\frac{dC_z}{dt} = P_zG + V_{oz}C_{oz} - V_eC_r \tag{4.3}$$

Although equation (4.3) is solvable via simple integration, it is more convenient to either adopt steady-state approximation or to use the transient method, according to ASHRAE 62.1-2016 [46].

4.1.1 Steady-State Approximation

Assuming that CO_2 concentrations in the auditorium have reached a steady-state, the equation (4.3) drops the derivative term and turns to equation (4.4). The estimated zone occupancy can be calculated by equation (4.5) using steady-state approximation.

$$P_z G + V_{oz} C_{oz} - V_e C_r = o \tag{4.4}$$

$$P_{z,est} = \frac{V_e C_r - V_{oz} C_{oz}}{G} \tag{4.5}$$

4.1.2 Transient Method

The transient method assumes the CO_2 derivative to be the rate of change of concentrations between the current and previous sampling instants, expressed as equation (4.6). The zone estimated occupancy can subsequently be calculated by equation (4.7).

$$\frac{dC_z}{dt} \approx \frac{C_z^i - C_z^{i-1}}{\Delta t} \tag{4.6}$$

$$P_{z,est}^i = \frac{V_e C_r - V_{oz} C_{oz}}{G} + v\frac{C_z^i - C_z^{i-1}}{G\Delta t} \tag{4.7}$$

where Δt is the time-step in second and i is the current step.

4.2 Artificial Neural Network Model

Artificial neural networks or connectionist systems are machine learning tools inspired by the biological neural networks and able to process data same as human brain. ANNs can develop linear and nonlinear models for time series. They are widely accepted as effective tools for fitting functions [51].

This study utilized a two-layer feed-forward neural network (FFNN) to approximate nonlinear relationships between inputs and outputs. The data of the occupancy

model need to move in only one direction (forward) from the input nodes, through the hidden layers and finally to the output node [52]. Since there are no cycles or loops associated with this specific network, FFNN was selected as the best fit. It consists of layers of processing units, denoted as neurons. The basic elements in FFNN are neurons arranged by the inputs, outputs and hidden layers. Input layers read in a signal to the network and hidden layers pass the signal through the network through weighted connections. In this network every hidden neuron receives inputs in the form of weighted signals from the previous layer plus a bias; then, flows to the output layer in one direction. By using the Sigmoidal activation function in equation (4.8), the network output can be described as equation (4.9) [53].

$$f(x) = \frac{1}{1 + e^{-x}} \tag{4.8}$$

and

$$y_k = f\left(\sum_{j=1}^{M} u_{jk}.f\left(\sum_{i=1}^{N} w_{ij}x_i + \Theta j\right)\right) + \Theta_k \tag{4.9}$$

where $f(x)$ is the Sigmoidal activation function, y is the output of the network, w and u are the scalar weights, and Θ is the bias. N and M are the number of inputs and the number neurons in the hidden layer, respectively.

In this study, the neural network model was built using MATLAB deep learning toolbox [54]. The inputs to the algorithm include the CO_2 concentrations of the zone, discharge and return air; The ventilation rate of the discharge, outdoor, and return air; and the schedule of classes specifying when the zone is likely to be occupied. The training function was selected to be Bayesian regularization back-propagation which minimizes the mean squared error (MSE) between the predicted and observed values. The number of hidden layer neurons was determined by the number of inputs using the equation (4.10) [55]. Figure 4.2 portrays the FFNN architecture used in this research.

$$M = 2N + 1 \tag{4.10}$$

where M and N are the number of hidden layer neurons and inputs, respectively.

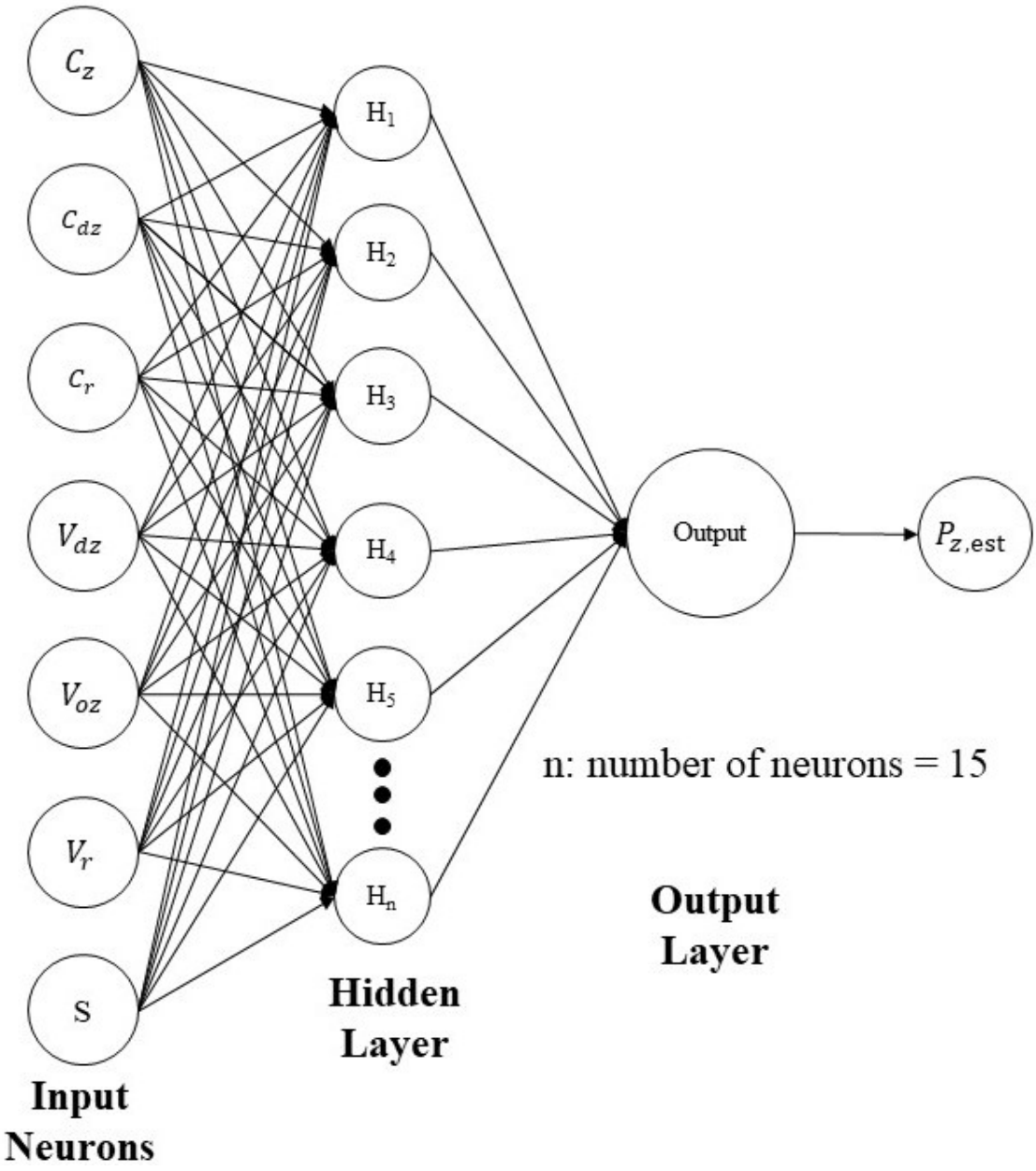

Fig. 4.2.: The feed-forward neural network architecture of the occupancy prediction model

The only output of the model is the estimated number of occupants inside the zone. The algorithm utilized the collected data from September 27^{th} 2019 to January 7^{th} 2020, except for the period of September 30^{th} to October 4^{th} as the validation period. A summary of FFNN design parameters and inputs is listed in Table 1.

Table 4.1.: Summary of FFNN design parameters and inputs

Network Type	FFNN
Input Parameters	CO_2 concentrations $[ppm]$ Ventilation rates $[L/s]$ Class schedule
Target(s)	Number of occupants
Training Algorithm	Bayesian regularization
Number of Neurons	15
Performance	MSE

4.3 The Observed Zone CO2 Concentrations and Occupancy

Figure 4.3 shows the actual zone CO_2 concentrations and the occupancy level in the auditorium over 5 working days, from September 30^{th} to October 4^{th}, 2019. The average breathing zone CO_2 concentrations during occupied mode were 508 ppm. The statistics also revealed that zone CO_2 concentrations while classes were in session and not in session averaged out about 697 ppm and 439 ppm, respectively. The occupancy level peaked at 150 people, occurring on September 30^{th}. The average number of occupants during occupied mode was 47 people while only 0 during unoccupied mode. Over the course of validation, the average occupancy level was observed to be 7.76 people. The average ventilation rate during occupied mode was about 1000 L/s, but 0 L/s when unoccupied mode. Table 4.2 summarizes the statistical information of the zone regarding occupancy level as well as CO_2 concentrations.

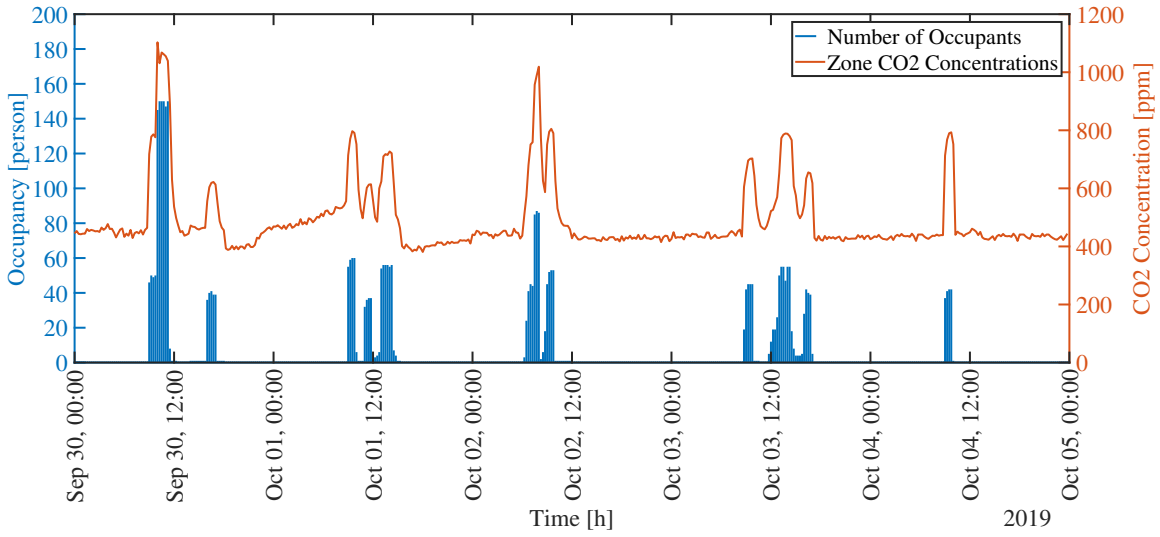

Fig. 4.3.: The auditorium human occupancy and CO_2 concentrations over 5 working days

Table 4.2.: The Auditorium occupancy and CO_2 level

	Occupied Mode	**Unoccupied Mode**	**Validation Period**
Average Number of Occupants	47	0	7.76
Average CO2 Concentrations [ppm]	508	443	490

4.4 Occupancy Prediction Model Performance

The occupancy estimations resulted from the physics-based models [equations (4.5) and (4.7)] and the FFNN model were presented in this section. This study used a typical Monday through Friday school week in 2019 for validation of the models. Both physics-based models predicted the occupancy level using the data obtained within this validation period. The FFNN model was trained using the data of the entire experiment period, excluding the days used for validation.

All models were evaluated using ASTM D5157 Standard Guide for Statistical Evaluation of Indoor Air Quality Models [56]. This standard provides three statistical

tools for evaluating the accuracy of IAQ model predictions as suggested by [33] and [56]. Table 4.3 classifies the values calculated using the equations (4.11) and (4.12).

$$r = \frac{\sum_{i=1}^{N}\left[(P_z^i - \bar{P}_z)(P_{z,est}^i - \bar{P}_{z,est})\right]}{\sqrt{\sum_{i=1}^{N}(P_z^i - \bar{P}_z)^2}\sqrt{\sum_{i=1}^{N}(P_{z,est}^i - \bar{P}_{z,est})^2}} \tag{4.11}$$

$$NMSE = \frac{1}{N\bar{P}_z\bar{P}_{z,est}}\sum_{i=1}^{N}(P_{z,est}^i - P_z^i)^2 \tag{4.12}$$

where P_z is the observed occupancy, $\bar{P}_z$ is the average of the observed occupancy, $P_{z,est}$ is the estimated occupancy, $\bar{P}_{z,est}$ is the average estimated occupancy, and N is the number of observations.

These parameters should be within certain ranges as mentioned below:

1. The correlation coefficient, calculated using equation (4.11), shall be 0.9 or greater.

2. The best fit regression line between the predicted and observed data should have a slope between 0.75 and 1.25, and an intercept less than 25% of the average observed value (7.76 people).

3. The normalize mean square error (NMSE), calculated using equation (4.12) should be less than 0.25.

Table 4.3.: The performance of Occupancy Prediction Models

Modeling Techniques	**r**	**Slope**	**Intercept**	**NMSE**
ASTM Criteria	>0.9	0.75-1.25	Less than 25% of average occupancy (1.94 people)	<0.25
Steady-State Approximation	0.85	1.07	5.56	1.60
Transient Method	0.86	0.85	3.86	1.56
FFNN Model	0.98	1.00	0.97	0.23

The steady-state approximation model established the correlation coefficient of 0.85, with the best fit regression line introducing a slope of 1.07 and an intercept of 5.56. The NMSE was also calculated to be 1.60. The best fit line turned out to be close to unity, revealing that the model could well estimate the overall trend of the change of occupancy. However, an intercept of 5.56, 72% of the average measured occupancy, means that the model tends to underestimate the number of occupants by 5 to 6 people. The correlation coefficient as well did not fall within the recommended guideline. Its NMSE further confirms that the model did not yield a satisfactory result.

The transient model, in comparison, gave the correlation coefficient of 0.86. Its best fit regression line had a slope of 0.86 and an intercept of 3.86. This intercept accounted for 50% of the average measured occupancy. The slope was within the guideline, although both the intercept and NMSE did not meet a satisfactory result. This conveys that the transient model, even though providing a better result compared to the steady-state approximation, is still not an optimal option to estimate the occupancy level.

Contrary to the aforementioned physical models, the FFNN introduced the correlation coefficient of 0.98. The NMSE was as small as 0.23. The slope and intercept of its best fit regression line turned out the numbers of 1.00 and 0.97, respectively. The FFNN intercept, with only 12% of the average measured occupancy, stayed well within the guideline. According to the ASTM criteria, the results suggest that the FFNN model can provide superior performance over the physical models. Figure 4.4 displays the assessment criteria of all three models discussed in this section.

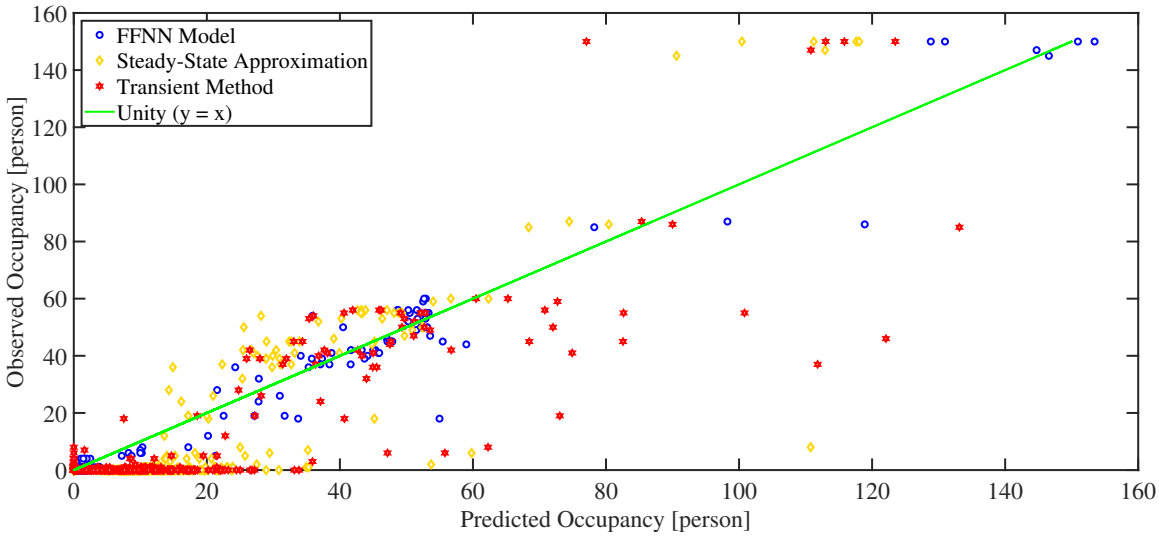

Fig. 4.4.: The occupancy prediction best fitted Models

To emphasize on performance differences, figure 4.5 provides the magnified inset of occupancy prediction models along with the observed zone occupancy.

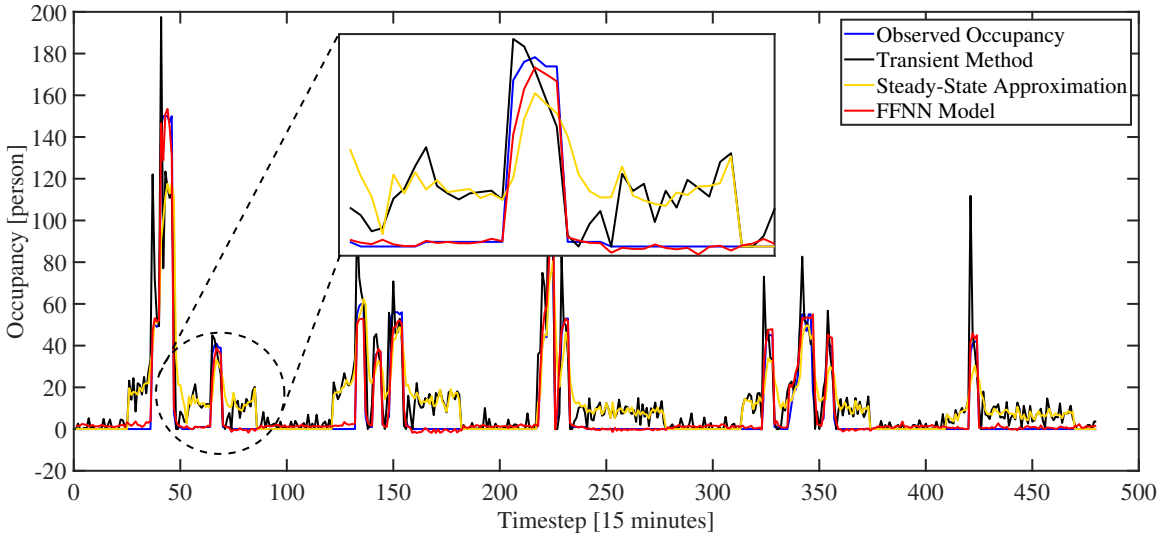

Fig. 4.5.: The magnified inset of occupancy prediction models as well as observed values

Aside from the comparison between these three models, figures 4.6 to 4.11 visualize their prediction performance individually. For each model, the prediction values were

plotted against the ground truth occupancy data over the validation period on 15 minute-timesteps. Each prediction comes with the errors analysed and plotted based on the Pearson residuals function, denoted as the equation (4.13) [57]. The equation (4.14) also simply shows the root mean square deviation.

$$pr_i = \frac{P_z^i - P_{z,est}^i}{RMSE} \tag{4.13}$$

$$RMSE = \frac{1}{n}\sum_{i=1}^{n}(P_z^i - P_{z,est}^i)^2 \tag{4.14}$$

where pr is the Pearson residual error, $RMSE$ stands for the root mean square error, and n is the number of samples.

The rationale behind using Pearson residuals is to check the model fit at each observation step. Positive or negative Pearson residuals indicate whether the observed values at each timestep are higher or lower than the predicted values, and the magnitude indicates the degree of departure.

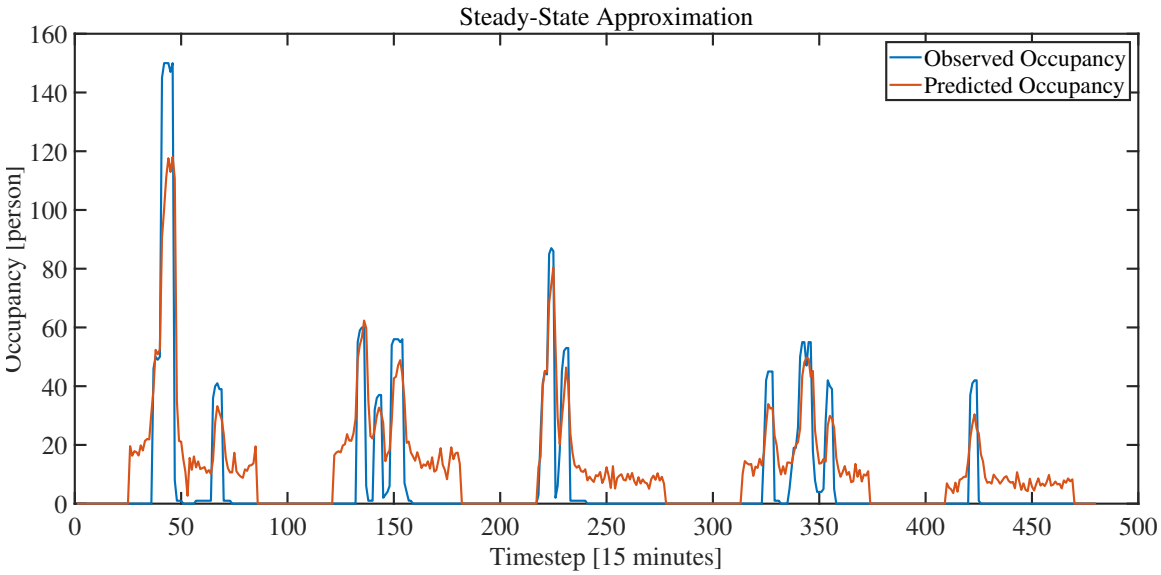

Fig. 4.6.: Predicted vs. observed (steady-sate approximation)

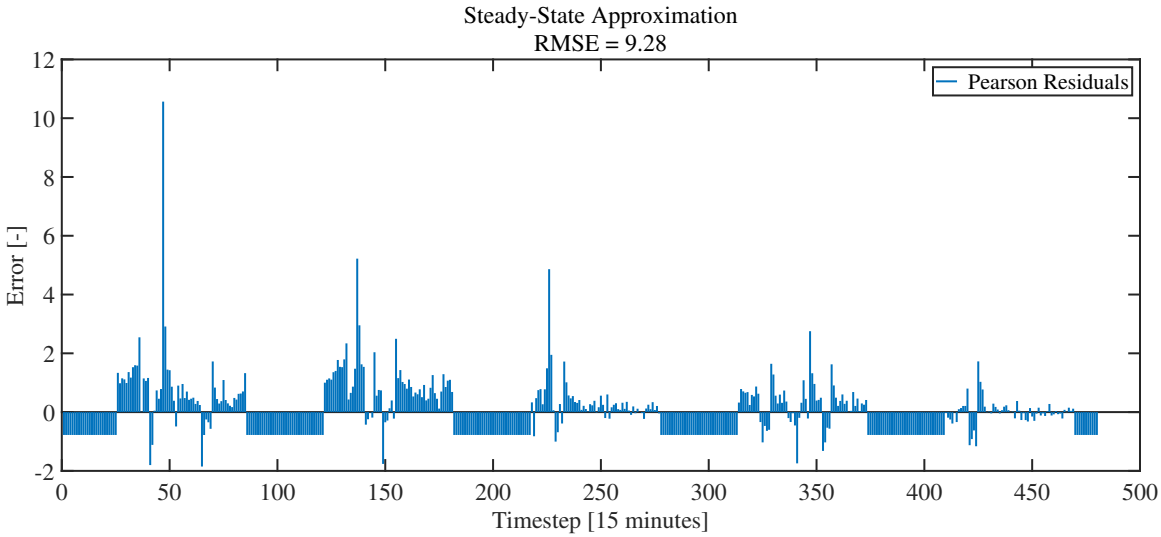

Fig. 4.7.: Pearson Residuals (steady-sate approximation)

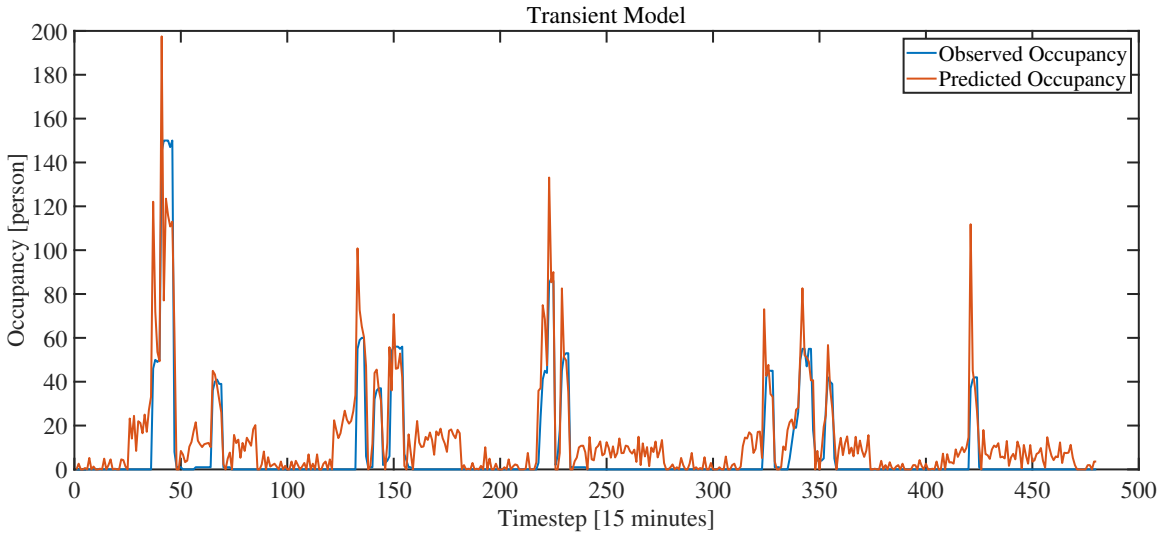

Fig. 4.8.: Predicted vs. observed (transient model)

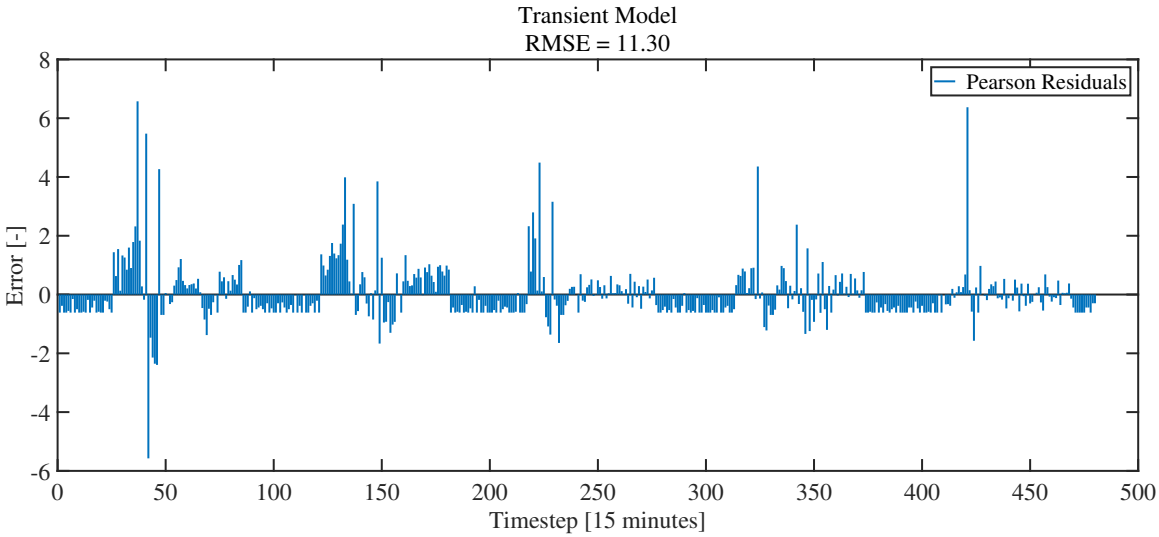

Fig. 4.9.: Pearson Residuals (transient model)

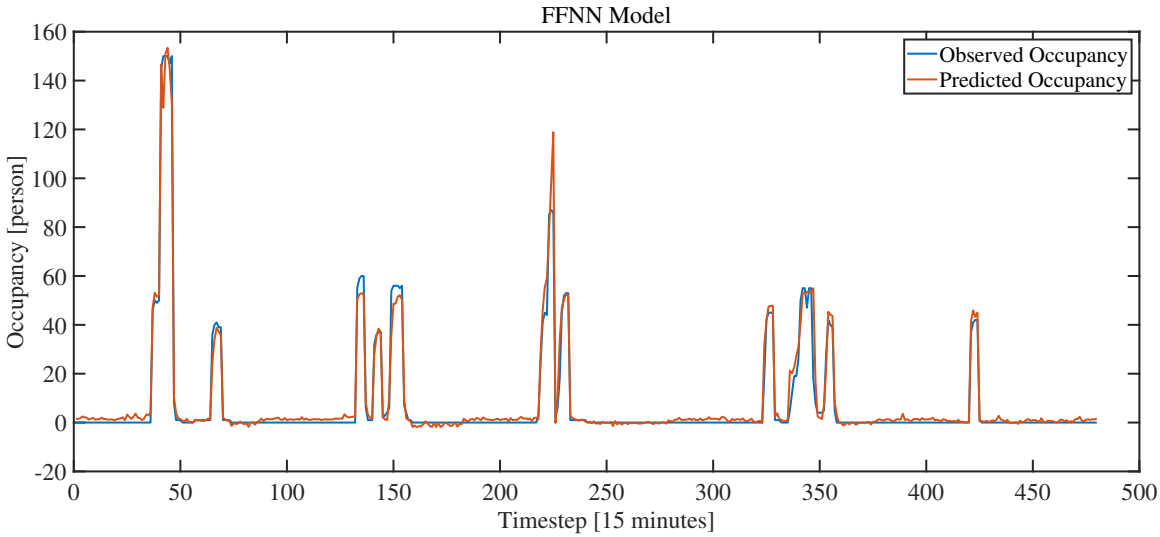

Fig. 4.10.: Predicted vs. observed (FFNN)

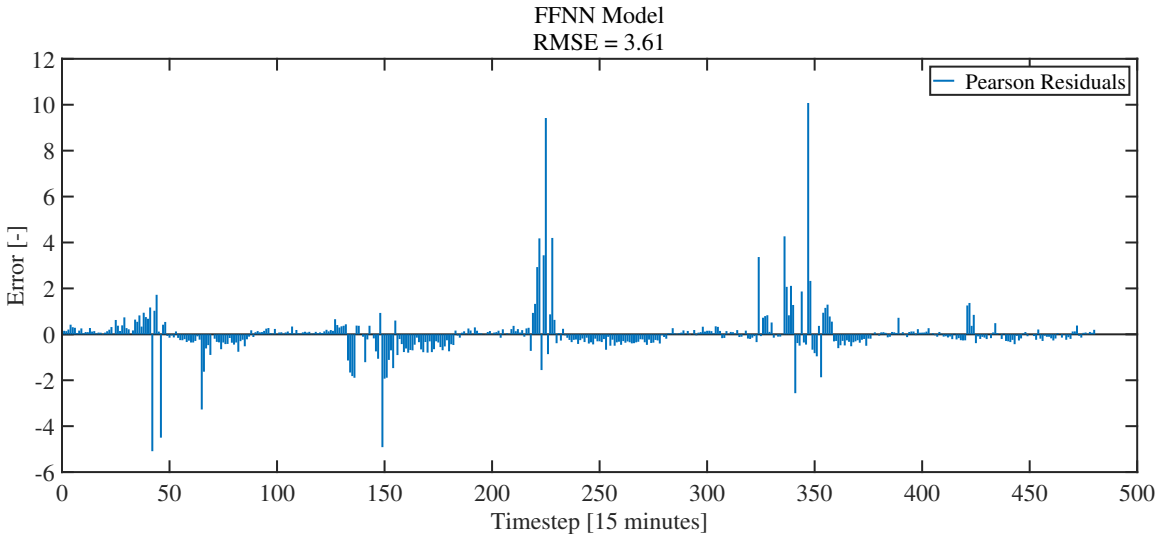

Fig. 4.11.: Pearson Residuals (FFNN)

4.5 Implementation of DCV Using CO_2

Ventilation and IAQ principles approve the utilization of DCV to control outdoor air ventilation rate. Under ASHRAE Standard 62.1-2016, an effective ventilation prerequisite is dependent on the occupancy level and the floor area of a zone. Equation (4.15) gives the minimum required outdoor airflow.

$$V_{bz} = R_p P_z + R_a A_z \tag{4.15}$$

where V_{bz} is the required outdoor fresh airflow in the breathing zone, R_p is the outdoor airflow rate required per person, P_z is the number of people in the ventilated zone during use, R_a is the outdoor airflow rate per unit area, and A_z is the occupiable floor area of the breathing zone.

4.5.1 Proportional Control

Based on the zone CO_2 concentrations measurement, the ventilation rate can be adjusted proportionally between the minimum and outdoor value. A proportional control system can be applied and integrated into the outdoor air dampers in order

to modulate the outdoor airflow rate. The equation (4.16) calculates the outdoor airflow rate which needs to be introduced by a proportionally controlled outdoor air damper [50].

$$V_{oz} = \frac{C_z - C_{z,min}}{C_{z,des} - C_{z,min}}(V_{oz,des} - V_{oz,min}) + V_{oz,min} \tag{4.16}$$

where C_z is the measured zone CO_2 concentrations ($ppm/10^6$), $C_{z,min}$ is the indoor CO_2 concentrations when the zone is unoccupied ($ppm/10^6$), $C_{z,des}$ is the indoor CO_2 concentrations when the zone has the design level of occupancy ($ppm/10^6$), $V_{oz,des}$ is the ventilation rate at the design level of occupancy (L/s), and $V_{oz,min}$ is the ventilation rate when the zone is unoccupied (L/s).

4.5.2 Dynamic Per-Occupant Controls

Using either of the physics-based models or FFNN model, the minimum ventilation rate could be achieved by equation (4.15). Both steady-state and transient models calculate the breathing zone population using the air properties that can be easily obtained from data-loggers. The FFNN model requires more computing power, yet it is achievable through implementing a dedicated system.

4.6 Comparison of Ventilation Rates

The CO_2-based DCV strategies were compared with the schedule-based on/off control that is currently being implemented. The proportional control do not consider the occupancy level directly. Instead, it employs the CO_2 concentrations of 1,200 ppm as the reference. The zone occupancy, rather than CO_2, is utilized by the dynamic per-occupant control strategies, following the minimum ventilation rate per person by ASHRAE standard. For a lecture hall, ASHRAE standard 62.1-2016 recommends the ventilation rate of 3.8 L/s per person and 0.3 L/s per square meter of the floor area [46].

A close-up view of the day of October 3^{rd} is shown in figure 4.12. It is evident that CO_2-based DCV strategies proposed lower ventilation rates when the AHUs

were in occupied mode. During unoccupied mode, the current on/off control strategy supplied no ventilation due to the closure of the outdoor air damper. Other strategies, in opposition, are required to introduce a minimum unoccupied ventilation upon ASHRAE standard 62.1 [46].

The proportional control strategy produced the most ventilation airflow amongst the DCV strategies as it continued to ventilate the zone even when there was no occupancy. It also reacted slowly to the change of occupancy when people entered and left the zone, creating a huge saving opportunity due to over-ventilation. The average ventilation rate during occupied and unoccupied mode were, respectively, 246 L/s, and 103 L/s.

The dynamic control strategies reacted quickly to the changes of occupancy. With no zone occupancy, the controls modulate the outdoor air damper to the minimum required. Compared to the on/off and proportional controls, all the dynamic control strategies proved to yield higher ventilation savings. The FFNN model ended up giving the lowest average ventilation rate due to its high prediction capability. The steady-state approximation method came after the FFNN model. The reason, however, is its tendency to underestimate the number of occupants.

The transient model introduced a comparable average ventilation rate to that of the steady-state approximation, on account of their similar occupancy prediction accuracy. The average ventilation rate using the steady-state approximation, the transient method, and the FFNN model, when occupied, were 183 L/s, 187 L/s, and 155 L/s, respectively. During unoccupied mode, these numbers changed to 94 L/s, 100 L/s, and 99 L/s, respectively. Table 4.4 compares the baseline average ventilation against the results of proposed strategies.

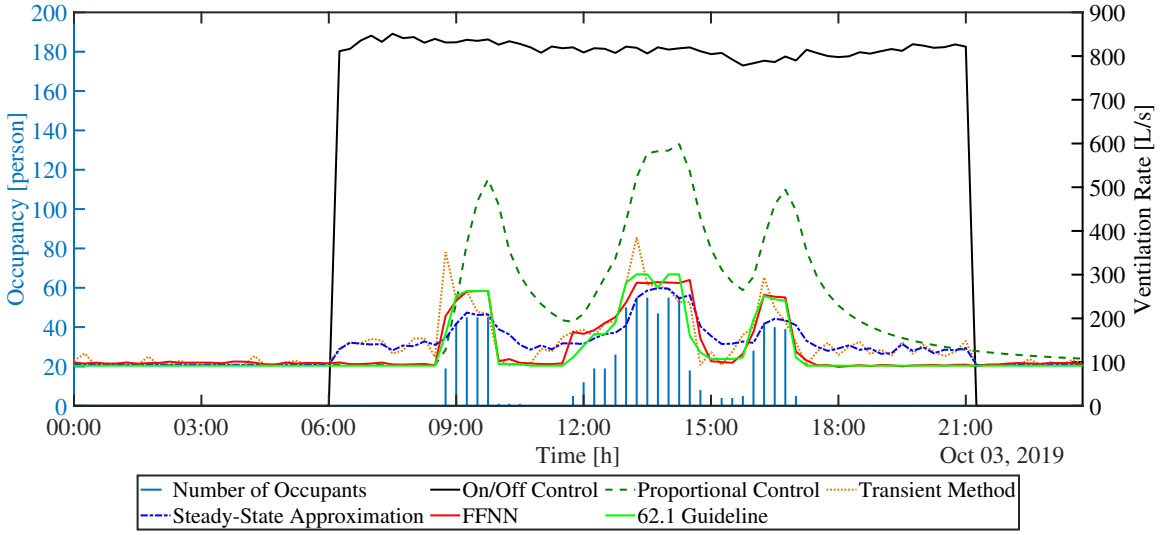

Fig. 4.12.: Simulated ventilation rates of different control strategies , October 3^{rd}

Table 4.4.: Comparison between average ventilation of the baseline and proposed DCV strategies

	Average Ventilation Rate [L/s]				
	Baseline	**Proportional Control**	**Steady-State Approximation**	**Transient Method**	**FFNN Model**
Occupied Mode	1,025	246	183	187	155
Unoccupied Mode	0	103	94	100	99

4.7 Indoor CO2 Concentrations Standards and Guidelines

The Occupational Safety and Health Administration (OSHA) introduced "code of federal regulations, title 29" in which the indoor CO_2 concentrations of 5,000 *ppm* were determined as the threshold to pose health risks [58]. However the ASHRAE Standard 62-1989 [59] recommended indoor CO_2 level of 1,000 *ppm* to satisfy comfort criteria. This recommendation was later dropped in the ASHRAE Standard 62-1999,

eliminating an absolute CO_2 threshold of 1,000 *ppm*. Instead, an indoor/outdoor CO_2 concentrations differential of 700 *ppm* was established [60]. Previous studies have shown that outdoor CO_2 concentrations typically range from 300 to 500 *ppm*. Therefore, indoor CO_2 concentration between 1,000 *ppm* to 1,200 *ppm* for sedentary occupants is considered an acceptable indoor air quality. This study deployed indoor CO_2 concentrations of 1,200 *ppm* as a guideline to build the proposed DCV controls discussed in previous sections.

4.8 The Effects of DCV Strategies on CO2 Concentrations

The impact of implementing different DCV strategies on zone CO_2 concentrations was modeled by equation (4.3). Figure 4.13 illustrates the change of zone CO_2 concentrations on the day of October 3^{rd}. As is clear, CO_2-based DCV strategies increased the overall CO_2 concentrations due to lowering the ventilation rate. None of the DCV strategies caused a peak in CO_2 level to violate the guideline.

The proportional control method resulted in an average of 530 *ppm* during occupied mode, and 408 *ppm* during unoccupied mode. Compared to the current practice, this model increased the average CO_2 level by 2.5% when occupied, but decreased by 8% during unoccupied mode.

The FFNN model led to the highest average CO_2 concentrations among per-occupant DCV strategies, which was expected as it proposed the lowest average ventilation rate. The numbers were 671 *ppm* and 445 *ppm*, 30% higher and 1% lower than the current CO_2 level during occupied and unoccupied mode, respectively.

The steady-state and transient models also showed similar outcomes as they introduced comparable average ventilation rates. During occupied mode the steady-steady and transient models yielded 646 *ppm* and 660 *ppm*, respectively; these numbers stood 25% and 28% higher than the current CO_2 concentrations measurement. During unoccupied mode the numbers were derived to be 425 *ppm* and 426 *ppm*, accounting for 3.6% lower CO_2 concentrations than that of current situation. It is noteworthy

to mention that over the unoccupied hours all aforementioned DCV strategies proposed lower CO_2 concentrations comparad to the current practice due to introducing a minimum fresh airflow.

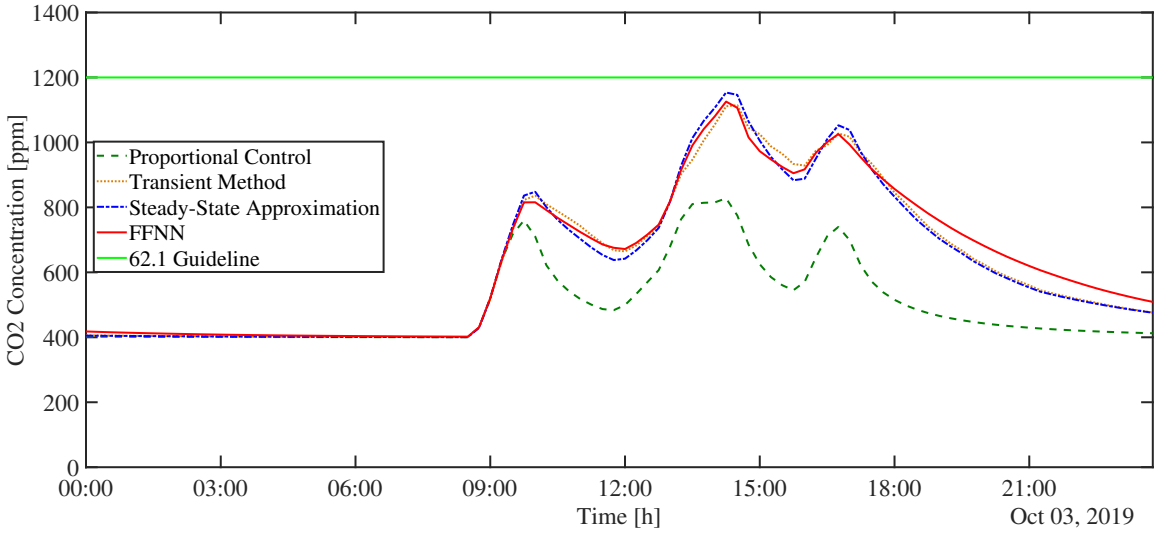

Fig. 4.13.: Zone CO_2 concentrations under different control strategies, compared with the recommended threshold by ASHRAE 62.1

4.9 Closure on The Chapter

So far, the superiority of implementing machine learning algorithms in CO_2-based DCV methods has been explored. The two-layer feed-forward neural network algorithm was exploited to estimate human occupancy level, and then proven through a case study. Drawing a comparison amongst the four examined DCV strategies demonstrated that the application of ANN in per-occupant DCV can result in substantial ventilation savings while considering IAQ as the first constraint. All strategies can well maintain indoor CO_2 concentrations below 1,200 ppm, meaning that they are all effective energy saving strategies. Further research will be included to analyze the effects on thermal comfort when implementing excessive reduction in ventilation airflow.

5. THERMAL COMFORT ANALYSIS

In addition to the consideration of indoor air quality criteria discussed in the previous chapter, this chapter attempts to analyze the zone thermal conditions. ASHRAE standard 55-2017 [61] had defined the thermal comfort as "the condition of mind that expresses satisfaction with the thermal environment and is assessed by subjective evaluation".

5.1 Predicted Mean Vote (PMV)

According to ASHRAE standard 55, PMV is a dimensionless index which predicts the mean value of the thermal sensation votes of a large group of people. The sensation scale is expressed from –3 to +3, corresponding to the categories shown in table 5.1.

Table 5.1.: PMV index scale and corresponding categories

Thermal Sensation Scale	**Category**
+3	Hot
+2	Warm
+1	Slightly warm
0	Neutral
−1	Slightly cool
−2	Cool
−3	Cold

The predicted mean vote (PMV) model employs heat balance principles to relate six primary factors to the average response of people regarding thermal comfort. These key factors are as follows:

- Metabolic rate
- Clothing insulation
- Air temperature
- Radiant temperature
- Air speed
- Humidity

The first two factors are characteristics of the occupants, and the remaining four factors are conditions of the thermal environment.

5.1.1 Modified PMV Model

The PMV model recommended by ASHRAE has caused a number of limitations, namely, the complexity of analysis and expensive sensor integration requirements. The assumptions of metabolic rate and clothing insulation could substantially vary from person to person, arising inaccuracies as well. Therefore, many researches [62–65] have been proposing the modified versions of this model in order to pave the way. In this study, a modified PMV model, by Sheng Zhang et. al. [65], was utilized. In comparison to the ASHRAE PMV model, the results of their proposed modified PMV have shown a negligible mean discrepancy of 0.14 (less than 5%). Unlike PMV_{ASHRAE} that comprises of six factors, this modified PMV model can be quantified by only two factors, the zone temperature and discharge airflow rate. The equations (5.1) and (5.2) show this regression-based model with fixed dimensionless coefficients.

$$PMV_m = -\frac{7}{91,300} N T_z^2 + \frac{19}{50} T_z - \frac{689}{74} \tag{5.1}$$

$$N = 3,600\frac{V_{dz}}{v} \tag{5.2}$$

where PMV_m is the modified PMV, N is the air change per hour (ACH), T_z is the zone air temperature (°C), V_{dz} is the zone discharge airflow rate (L/s), and v is the volume of the zone (L).

5.2 Predicted Percentage of Dissatisfied (PPD)

Added to the PMV index, the ASHRAE standard 55 introduced another index by which a quantitative prediction of the percentage of thermally dissatisfied people can be determined. Indeed, PPD index can be calculated using PMV as denoted in equation (5.3). Figure 5.1 shows the PMV-PPD relationship on the coordinate system.

$$PPD = 100 - 95\mathrm{e}^{\left[-0.03353(PMV)^4 - 0.2179(PMV)^2\right]} \tag{5.3}$$

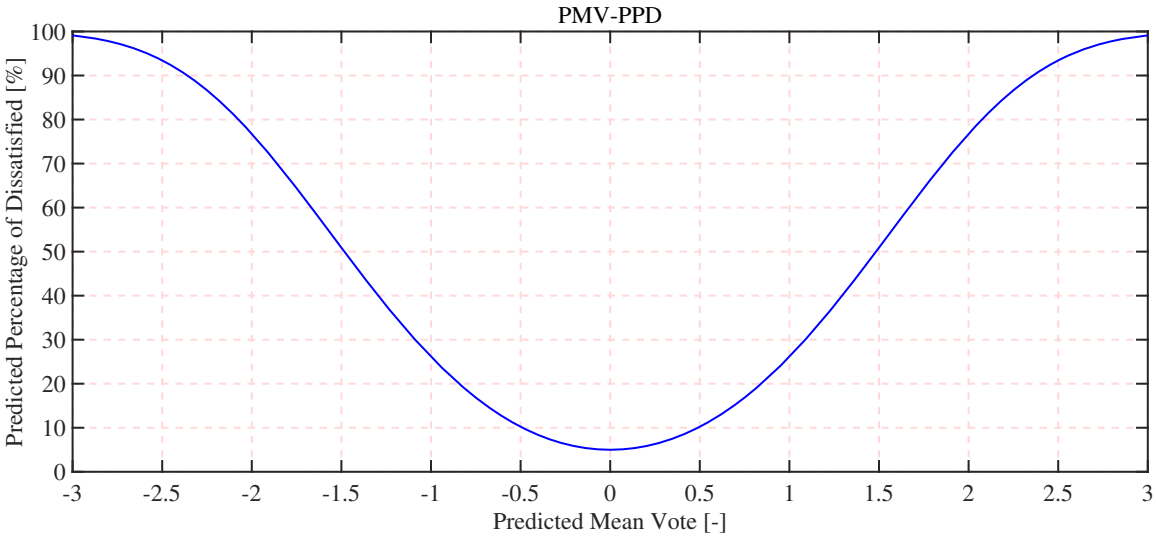

Fig. 5.1.: PPD as a function of PMV

Due to the large variations between occupants, physiologically and psychologically, it is quite challenging to attain 100% thermal satisfaction in a space. ASHRAE standard 55 determined the acceptable thermal environment when more than 80% of

the human occupants perceive the situation thermally comfortable. In other words, the Predicted Percentage of Dissatisfied (PPD index) should be maintained at or less than 20% [61].

5.3 Current Zone Thermal Comfort Analysis

Figure 5.2 shows the actual auditorium air Temperature, outdoor airflow rate and occupancy level against each other over 5 working days, from September 30^{th} to October 4^{th}, 2019. It was observed that the zone air temperature started decreasing due to the ventilation during occupied mode, averaging about $22°C$. This temperature drop continued to the point where it reached the steady-state. During unoccupied mode, as a result of heat transfer from surrounding, the zone air temperature increased gradually with a transient form, until the next day when the ventilation started again. The zone air temperature averaged about $26°C$ when in unoccupied mode. It should also be noted that the occupancy level was the reason for the spikes of temperature due to the fact that occupants release latent and sensible heat. The interaction between temperature and ventilation, however, confirmed the fact that the impact of occupancy in temperature adjustments were far less than that of outdoor airflow rate.

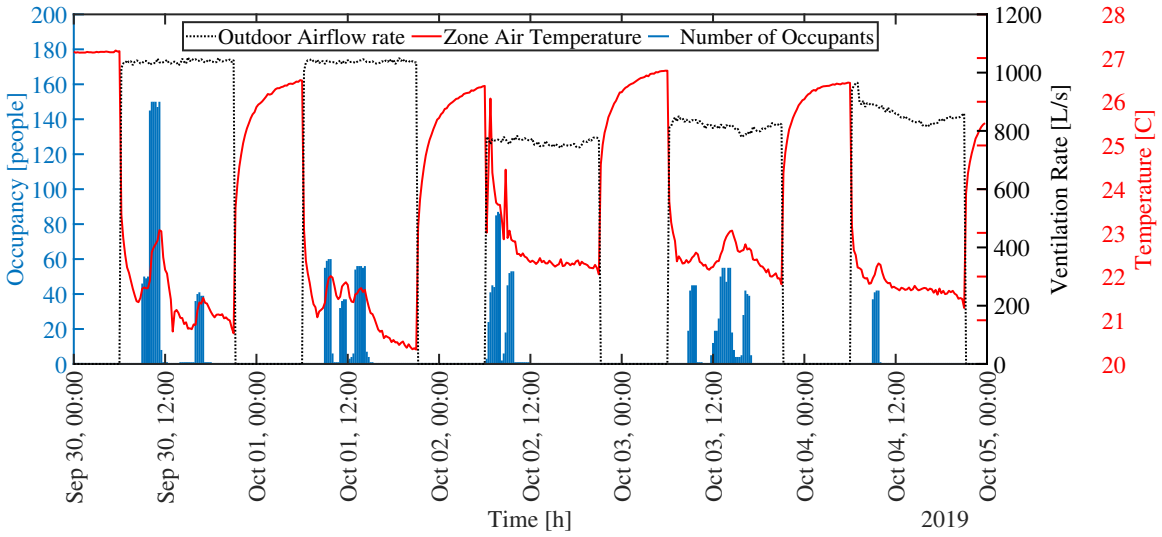

Fig. 5.2.: Observed auditorium air temperature, outdoor airflow rate and occupancy level

5.3.1 Current Zone PMV and PPD Indices

Figure 5.3 shows the current PMV index of the auditorium only during occupied mode since the thermal sensation is associated with occupants' votes. It could be concluded that the current operation is over-ventilating the zone, as the PMV index was lower than -1 almost over the entire course of experiment. As the figure 5.4 portraits the PPD index, this over-ventilation resulted in the thermal dissatisfaction of more than 20% of occupants, reaching as high as 70% at some point. For the second half of this period, the PMV index fell between -1 to 0, showing less dissatisfaction, due to the lower average of ventilation rate. Figure 5.5 visualizes how the PMV and PPD indices are related. It confirms that almost all votes were negative, with rare cases of "neutral" sensation but mostly "slightly cool" and "cool". Table 5.2, also, gives an overview of the current practice.

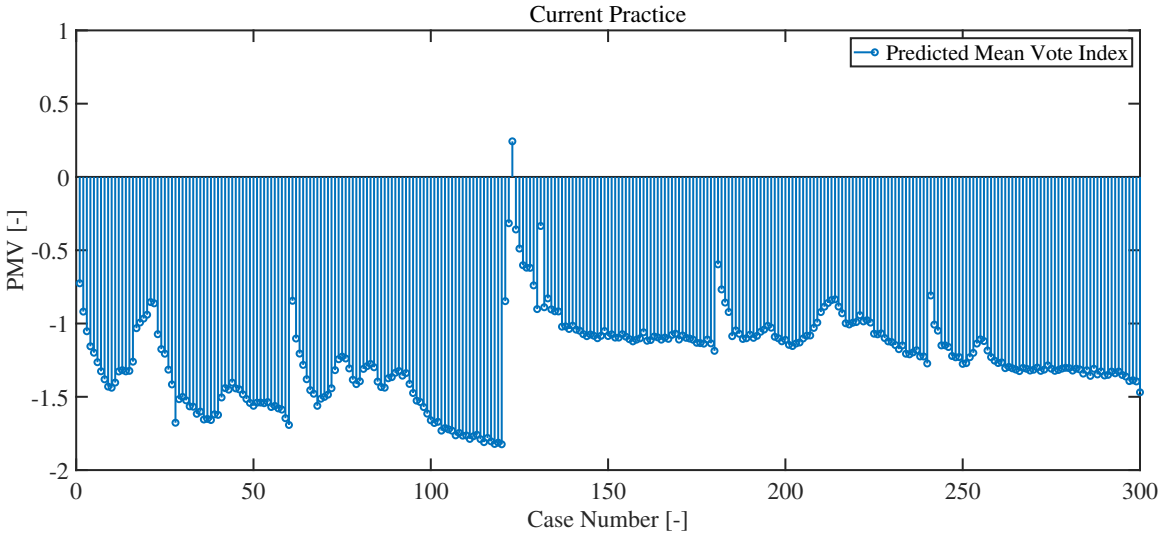

Fig. 5.3.: Auditorium current PMV index

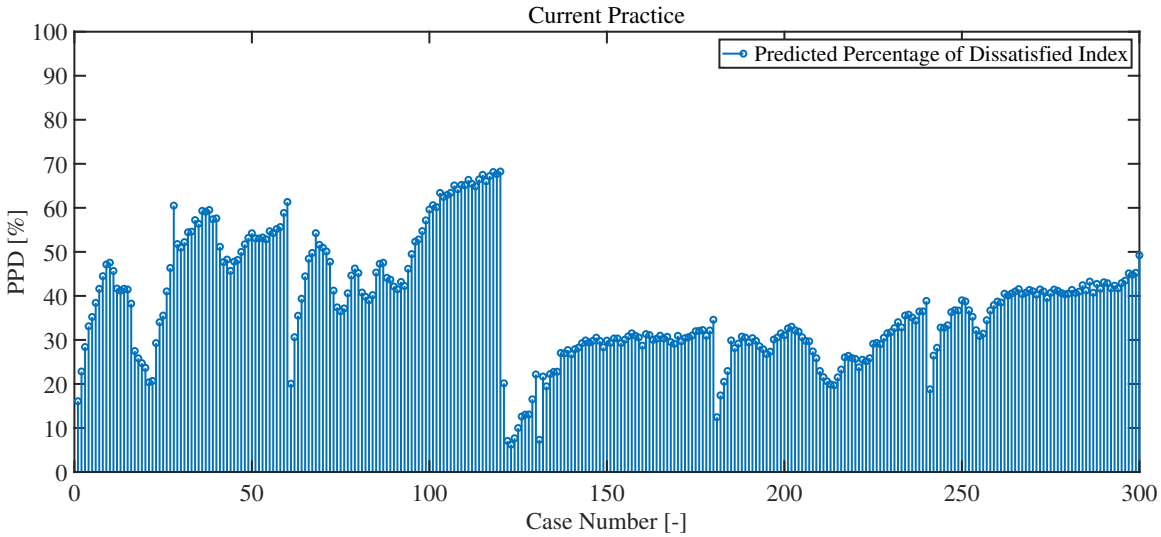

Fig. 5.4.: Auditorium current PPD index

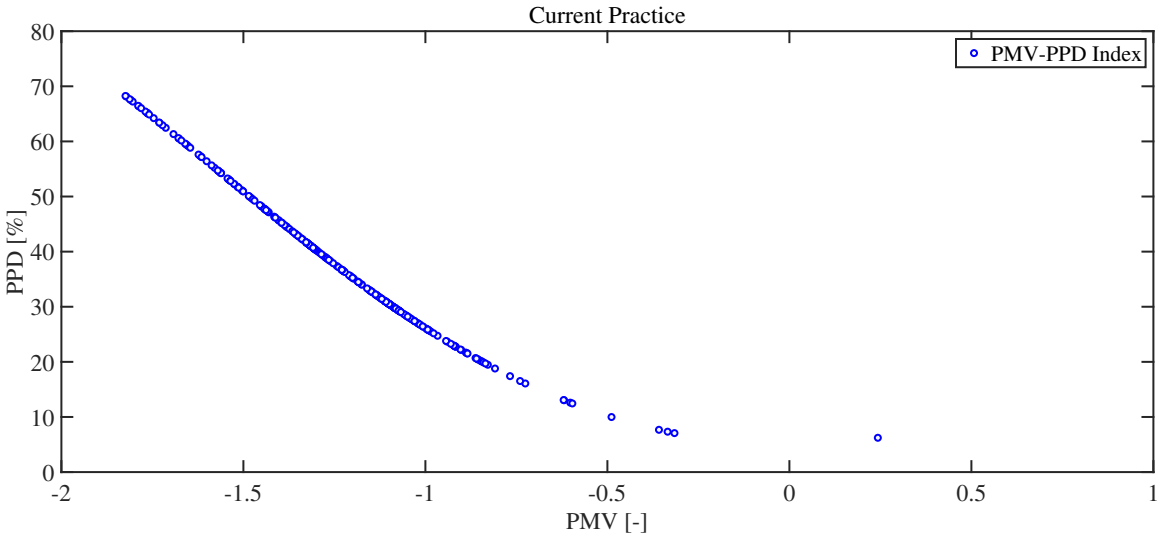

Fig. 5.5.: Auditorium current PMV-PPD index

Table 5.2.: Summary of current temperature, PMV and PPD indices (occupied mode)

	Minimum	**Maximum**	**Average**
Zone Air Temperature [°C]	20.33	26.71	21.95
PMV [-]	−1.82	+0.57	−1.24
PPD [%]	5.00	68.25	38.05

5.4 Zone Air Temperature Model

With the proposed ventilation of the aforementioned models discussed in the previous chapter, the zone air temperature must be expected to change. The number of occupants, as figure 5.2 proved, is also another important factor in affecting temperature. Given the correlations between the zone air temperature, outdoor airflow rate and occupancy, a data-driven model was developed.

5.4.1 Gaussian Process Regression (GPR)

Gaussian Process Regression (GPR) models are non-parametric Bayesian approach models, fitting various functions into the training data until the prediction values reach the nearest possible neighborhood of the observed values [66]. As opposed to this non-linear regression algorithm, the simple linear regression model only allowed a limited flexibility due to the fixed coefficients of the linear function by which a poor prediction was offered. Assuming the training set of $\{(x_i, y_i); i = 1, 2, \ldots, n\}$, the predictive model can be structured as the equation (5.4).

$$y = x^T \beta + \epsilon \tag{5.4}$$

where y is the predicted values, x is the matrix of input variables, x^T is the transpose of the matrix x, β indicates the fitted coefficients, and ϵ is the additive noise between the observed and predicted values. As for developing the GPR model, MATLAB regression learner toolbox was utilized. The inputs of the model were selected to be the outdoor airflow rate (V_{oz}) and the number of occupants in the auditorium (P_z), while the only output was the auditorium air temperature (T_z).

5.4.2 Zone Air Temperature Model Performance

Figure 5.6 shows the observed and predicted values of the auditorium air temperature, over the validation period. It is noticeable that the predicted values closely followed the observed values during the occupied mode. However, the model showed a lack of prediction during the unoccupied mode as it did not take the transient temperature rise into account. The rationale behind this is because of no airflow variations (damper closure during unoccupied mode) while the temperature still fluctuates. The GPR model, overall, was considered to be functional as it was supposed to contribute to the thermal comfort analysis for which only the occupied mode matters.

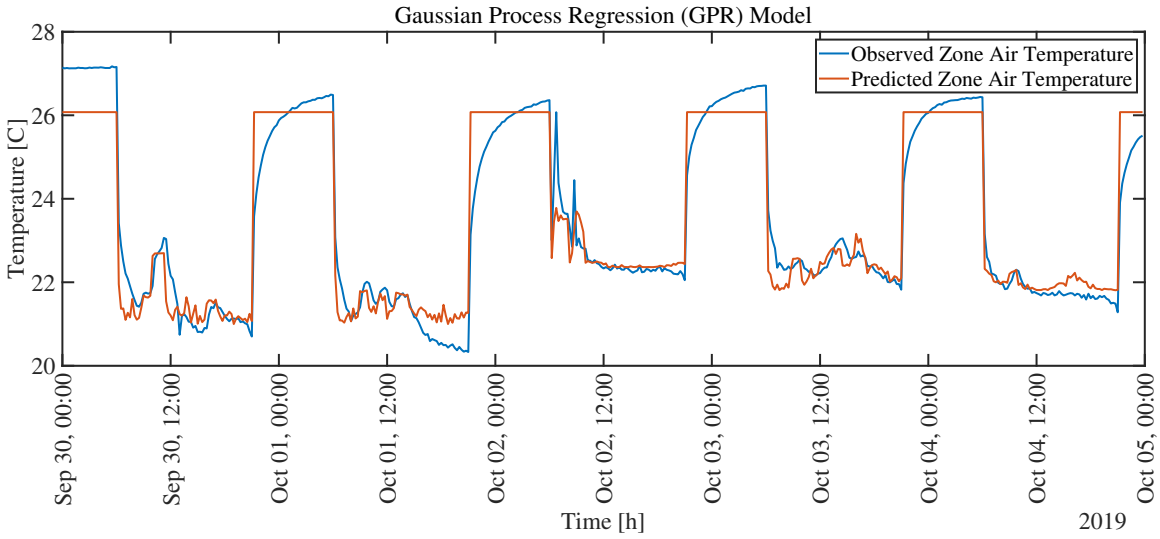

Fig. 5.6.: Predicted vs. observed

Figure 5.7 represents the linear correlation between the observed and predicted values. The proportion of the variance between the two values was measured by the coefficient of determination (R-squared), denoted as the equation (5.5).

$$R^2 = 1 - \frac{\sum_i \left(T_z^i - \hat{T}_z^i\right)^2}{\sum_i \left(T_z^i - \bar{T}_z^i\right)^2} \tag{5.5}$$

where T_z is the observed zone air temperature, $\hat{T}_z$ is the predicted value of zone temperature, and $\bar{T}_z$ is the mean of the observed zone temperature.

Table 5.2 reports this number along with the probability value (p-value) and $RMSE$. The results turned out that the predictive model was fitted with a negligible error and a high significance.

Table 5.3.: Zone air temperature model performance parameters

Statistical Parameters	**Magnitude**
R-squared	0.924
P-value	1.02×10^{-269}
RMSE	0.567

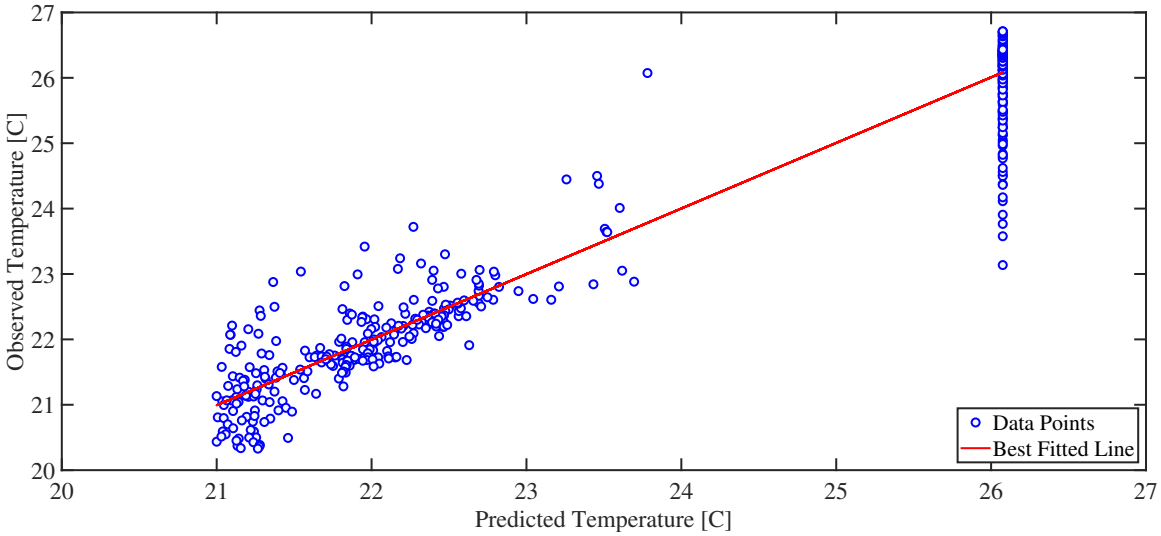

Fig. 5.7.: The GPR best fitted model

Figure 5.8, lastly, shows the Pearson residuals. The noticeable patterns confirm the fact that GPR model could not offer accurate predictions during unoccupied mode; although, it worked highly accurately during the occupied hours.

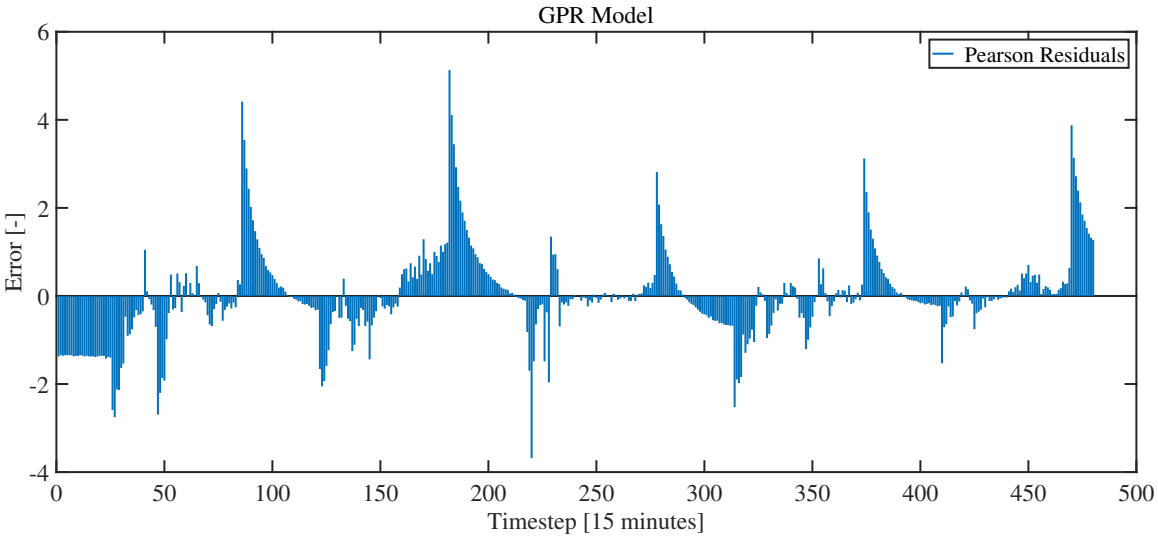

Fig. 5.8.: Pearson residuals (GPR Model)

5.5 Proposed Zone Thermal Comfort Analysis

As discussed in chapter 4, the dynamic per-occupant models estimated number of occupants and subsequently proposed ventilation rates. For each individual model, the estimated number of people along with associated ventilation rate were substituted into the GPR model to predict zone air temperature. As far as proportional control, however, the ground-truth occupancy data were used since this model proposed ventilation rate based on CO_2 concentrations.

As demonstrated by the observed data (section 5.3), the auditorium air temperature was negatively correlated with ventilation rate. Figures 5.9, 5.13, 5.17 and 5.21 show the auditorium air temperature predicted by the estimated occupancy and ventilation rate for each DCV strategy. As they all proposed lower ventilation rates during occupied mode, temperature was predicted to be higher compared to the current practice, averaging about $25°C$. Over unoccupied mode, this average went up by less than $1°C$, staying below that of current practice. This is because DCV strategies introduced a minimum required fresh airflow for unoccupied hours. It should be pointed out that occupied mode in all proposed cases refer to only the time when

there are occupants (as opposed to the current fixed calendar); unoccupied mode, as well, is vice versa. Each model is followed by the PMV, PPD and PMV-PPD graphs; figures 5.10 - 5.12, 5.14 - 5.16, 5.18 - 5.20 and 5.22 - 5.24 represent FFNN model, steady-state approximation, transient method and proportional control, respectively. The noticeable fact is that all DCV strategies contributed positively to the zone thermal condition due to the prevention of over-ventilation. All models considered, the PMV index ranged from -0.8 to 0.3, proposing dissatisfaction percentage as low as 5% and as high as 20%. This predicted range fell within the ASHRAE standard 55 comfort criteria by keeping the PPD index far less than 20%. Table 5.4 draws a comparison by presenting the break-down numbers of each scenario. It reveals that FFNN, steady-state and transient models were closely comparable while the proportional control model resulted in a slightly less average temperature as it proposed higher ventilation amongst others.

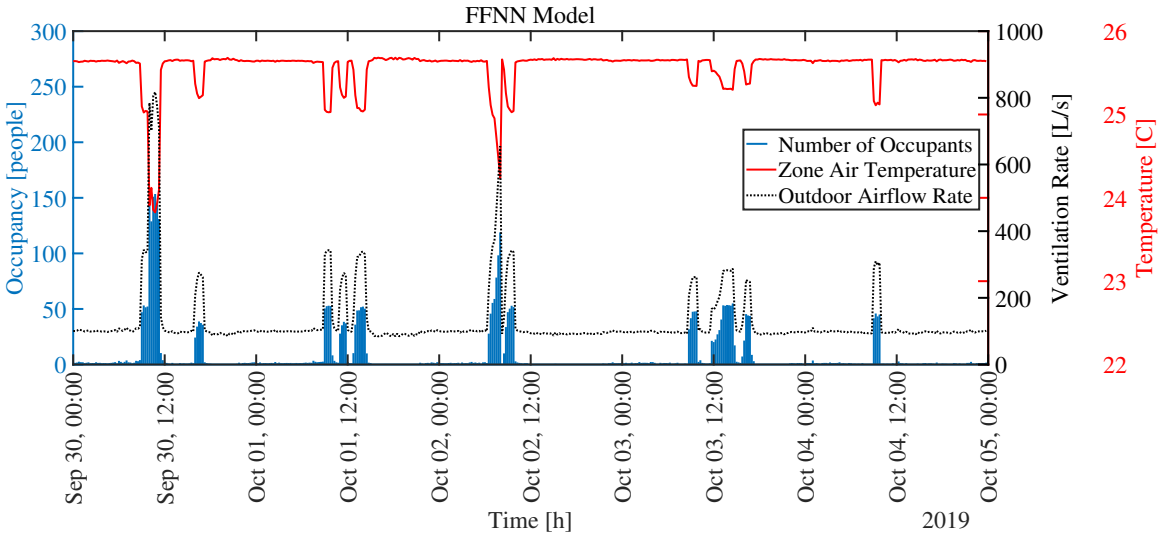

Fig. 5.9.: Proposed temperature, outdoor airflow rate and occupancy (FFNN model)

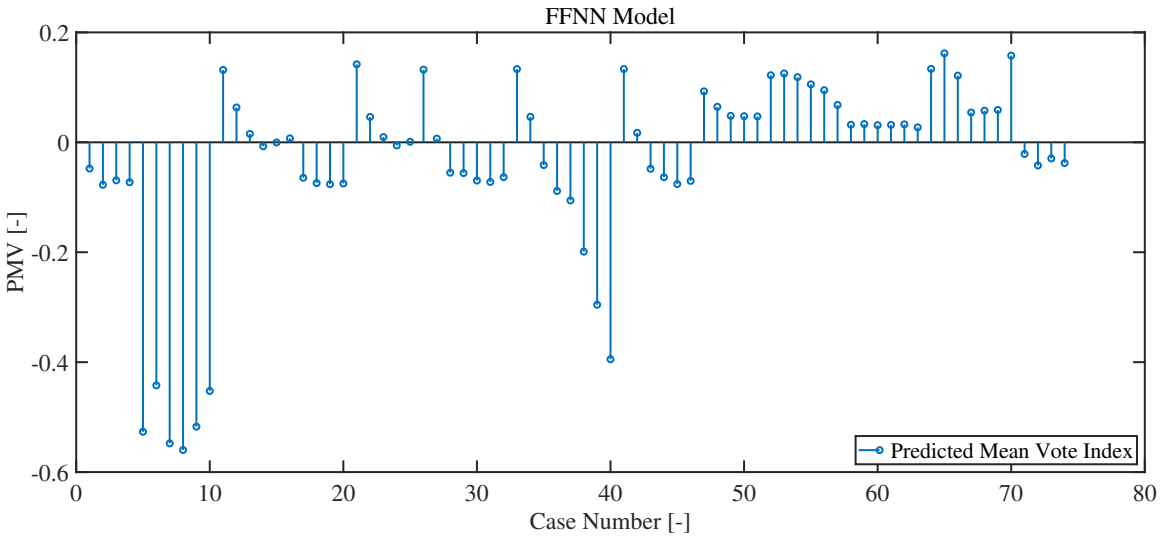

Fig. 5.10.: Proposed PMV index (FFNN model)

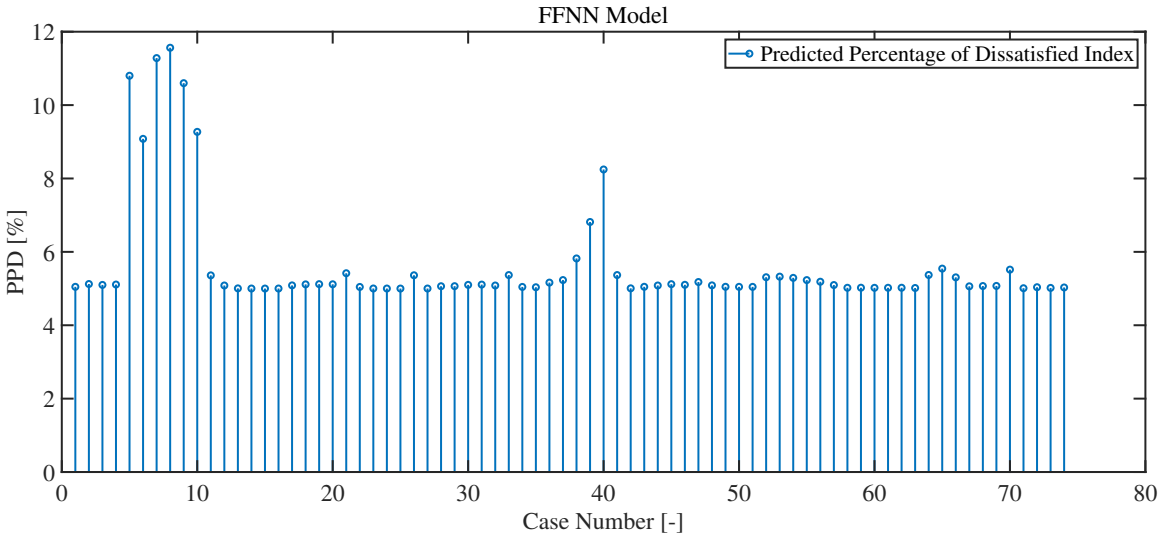

Fig. 5.11.: Proposed PPD index (FFNN model)

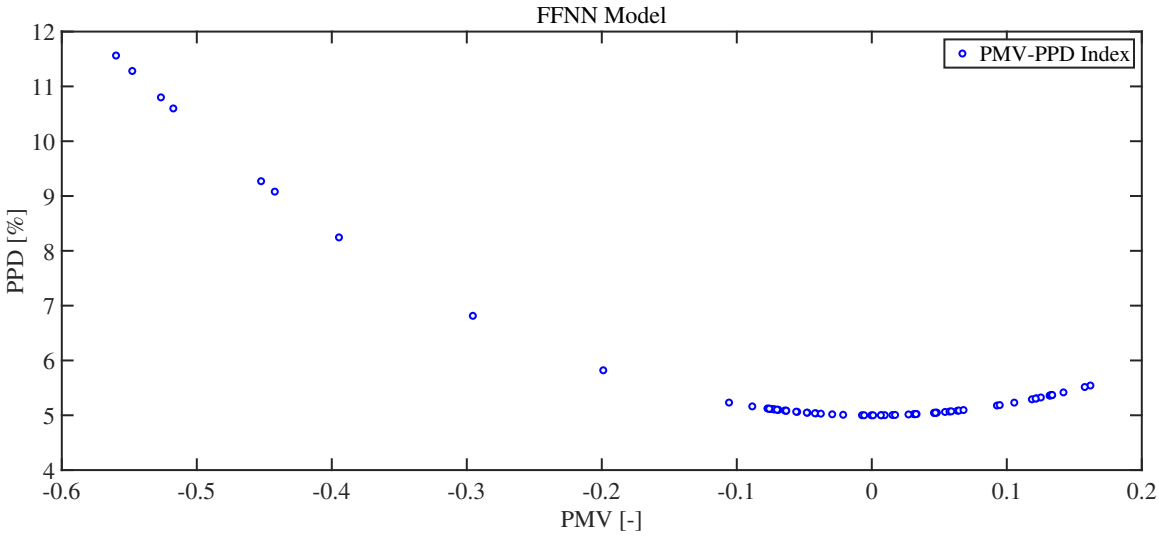

Fig. 5.12.: Proposed PMV-PPD index (FFNN model)

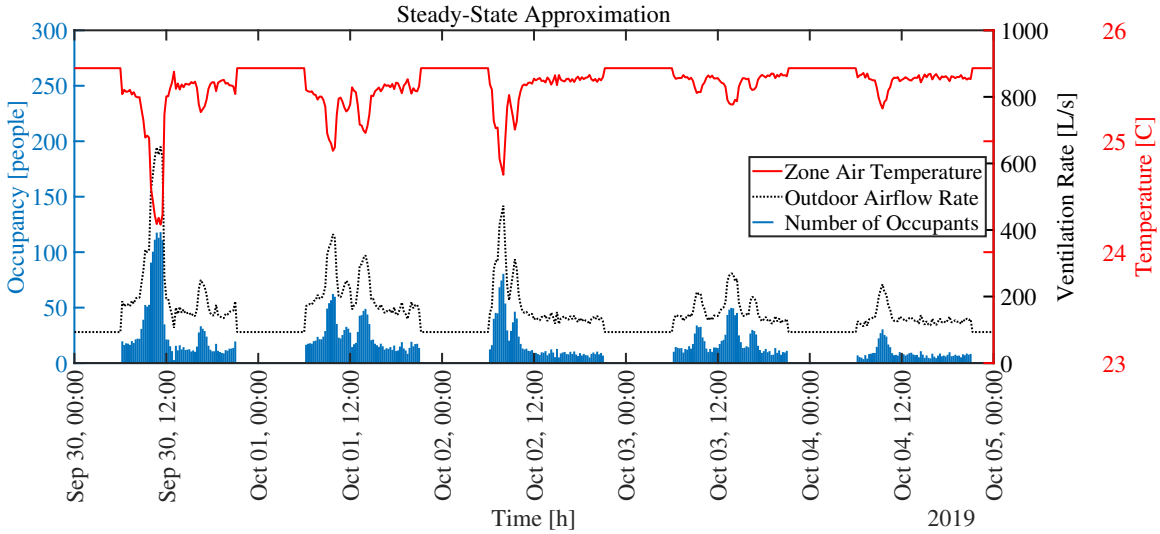

Fig. 5.13.: Proposed temperature, outdoor airflow rate and occupancy (steady-state approximation)

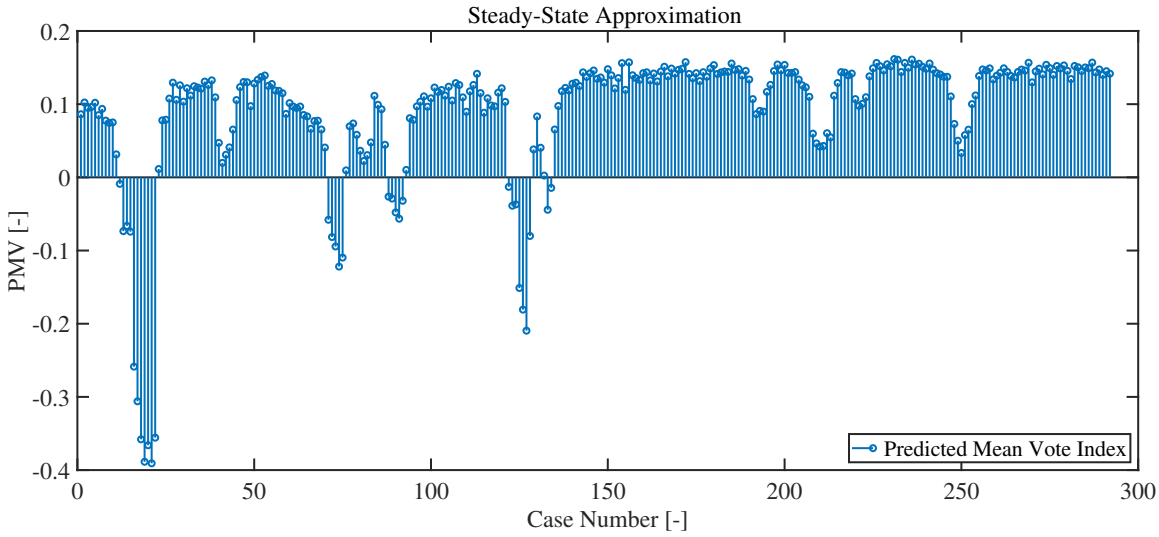

Fig. 5.14.: Proposed PMV index (steady-state approximation)

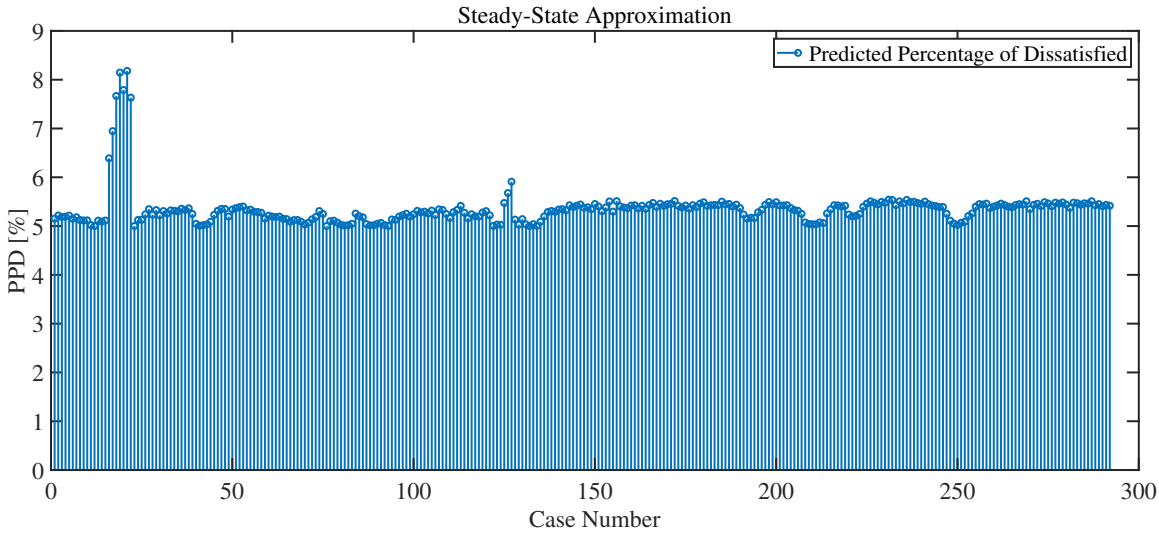

Fig. 5.15.: Proposed PPD index (steady-state approximation)

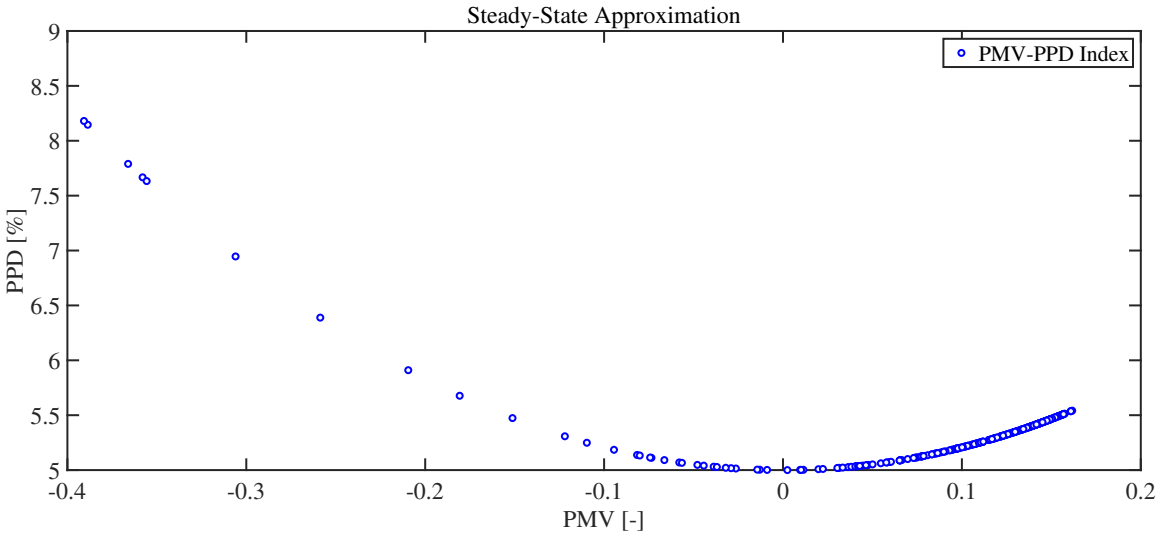

Fig. 5.16.: Proposed PMV-PPD index (steady-state approximation)

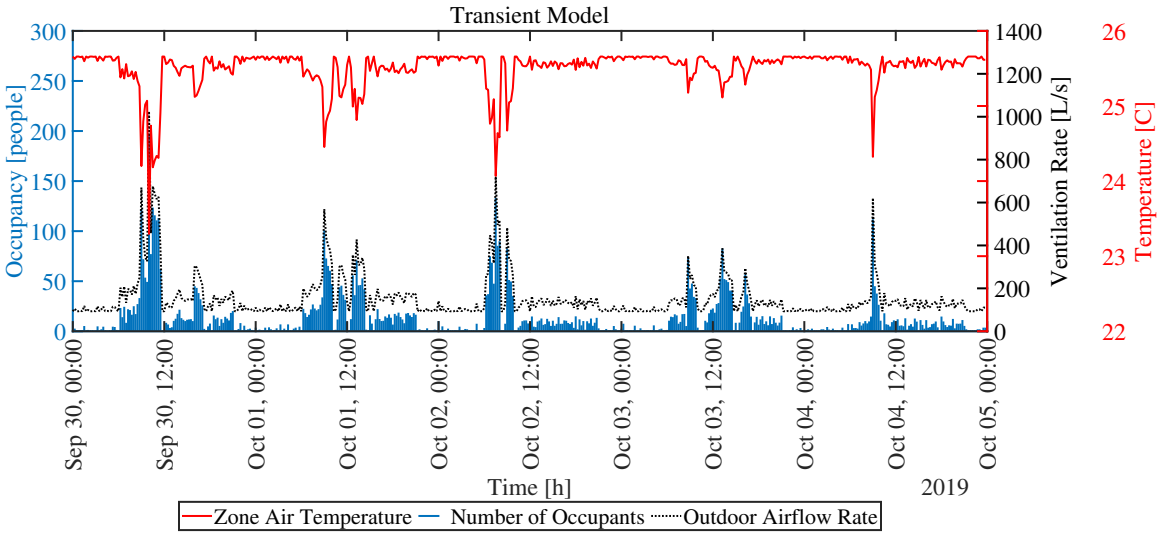

Fig. 5.17.: Proposed temperature, outdoor airflow rate and occupancy (transient model)

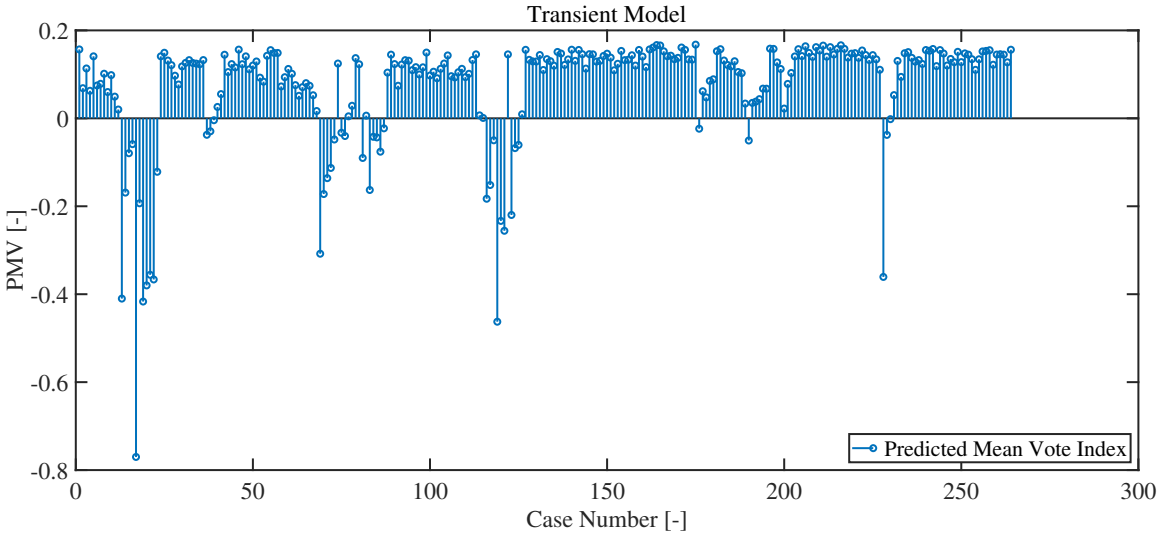

Fig. 5.18.: Proposed PMV index (transient model)

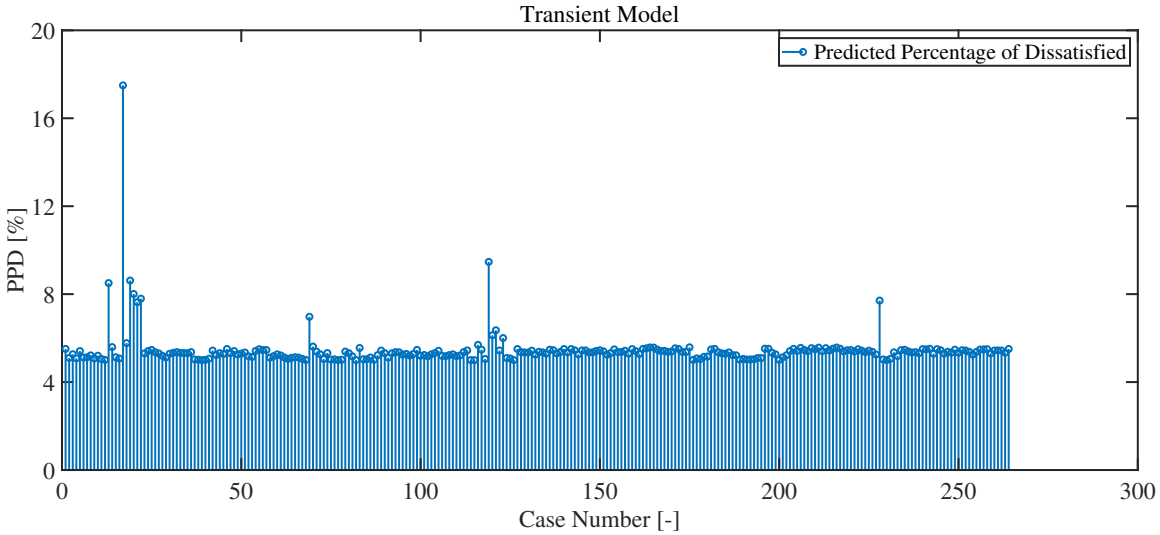

Fig. 5.19.: Proposed PPD index (transient model)

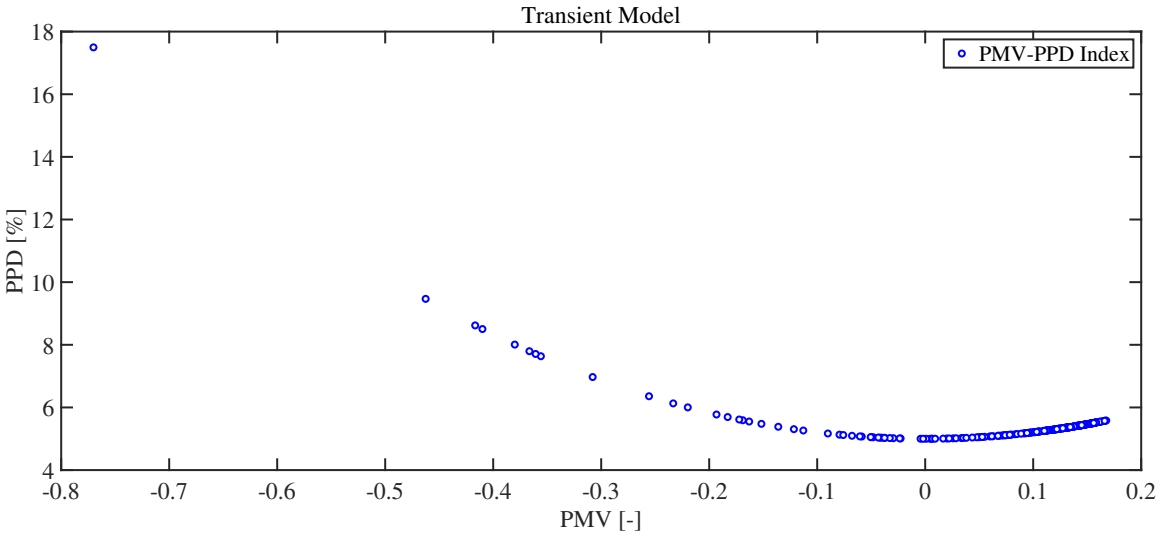

Fig. 5.20.: Proposed PMV-PPD index (transient model)

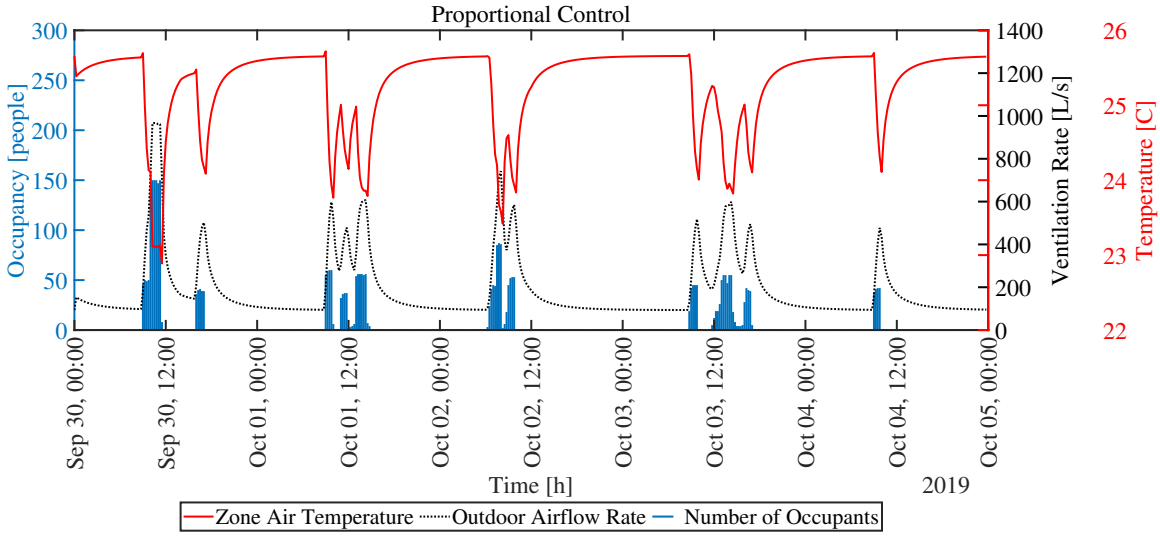

Fig. 5.21.: Proposed temperature, outdoor airflow rate and occupancy (proportional control)

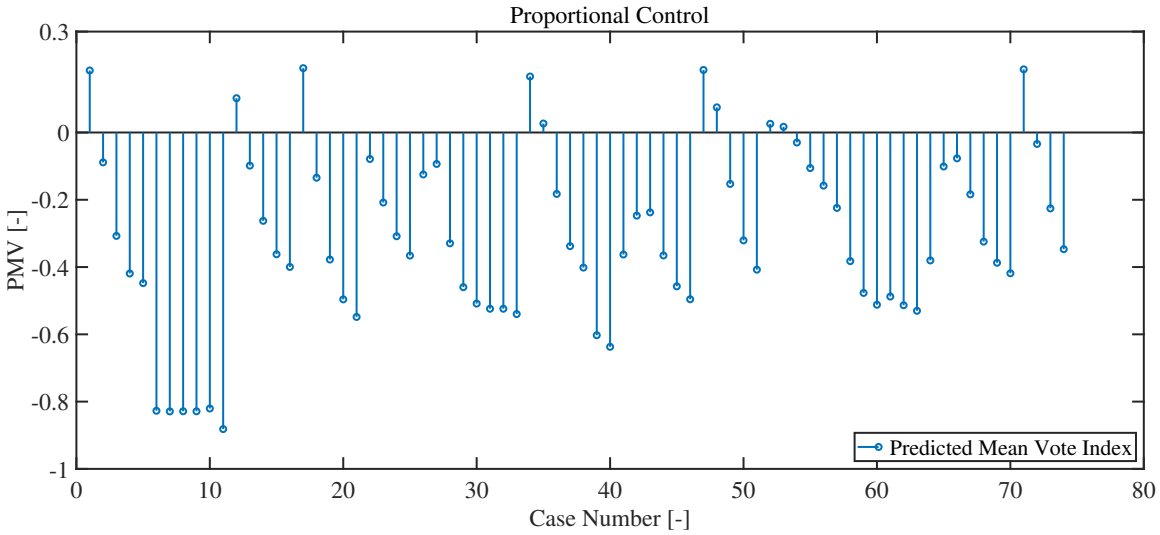

Fig. 5.22.: Proposed PMV index (proportional control)

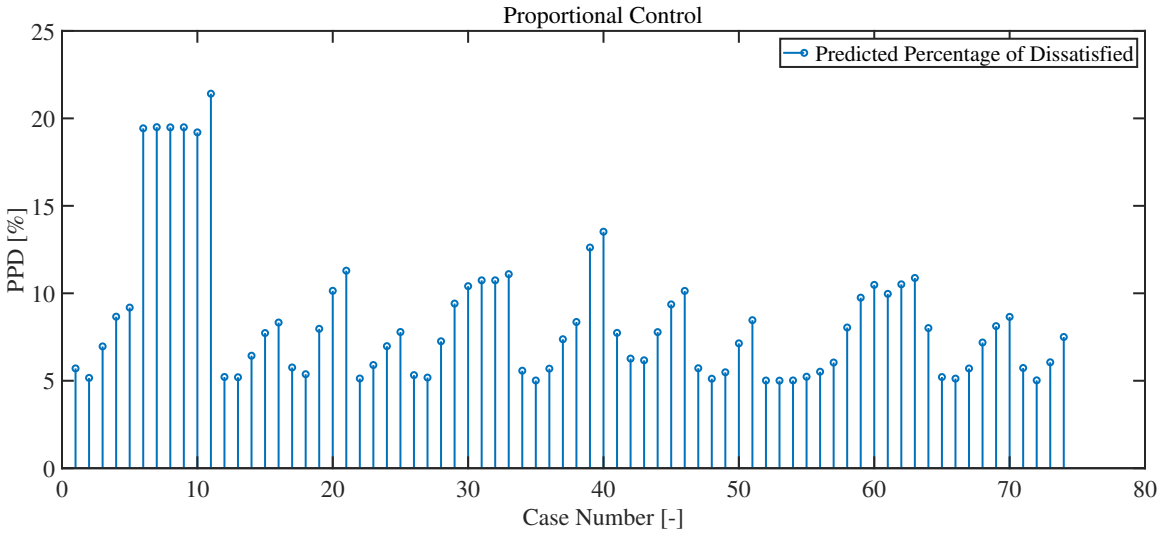

Fig. 5.23.: Proposed PPD index (proportional control)

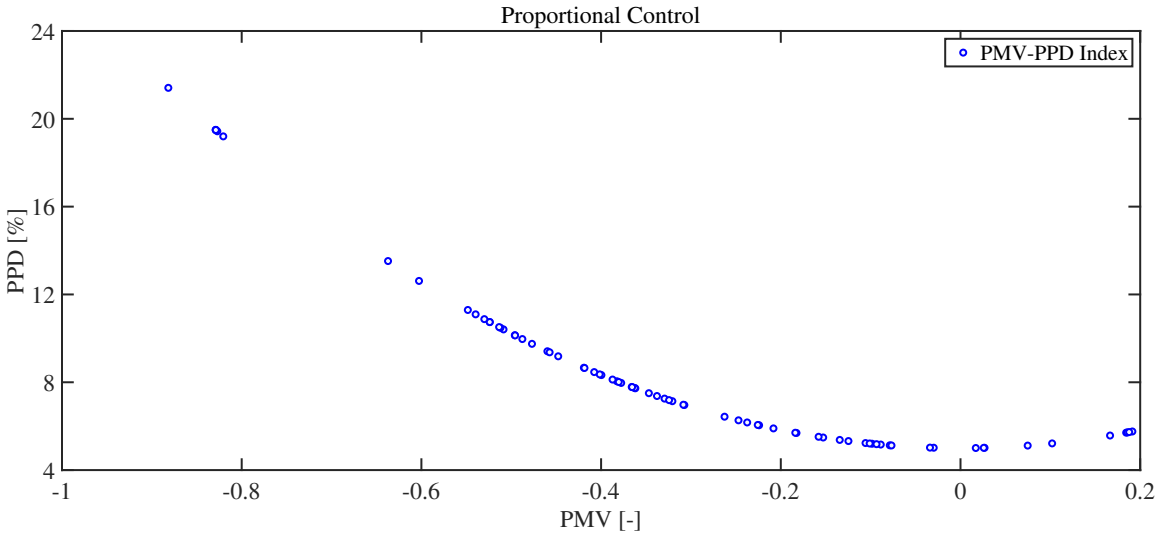

Fig. 5.24.: Proposed PMV-PPD index (proportional control)

The observed and proposed zone air temperature values compiled together, figure 5.25 provides the magnified inset for a better representation.

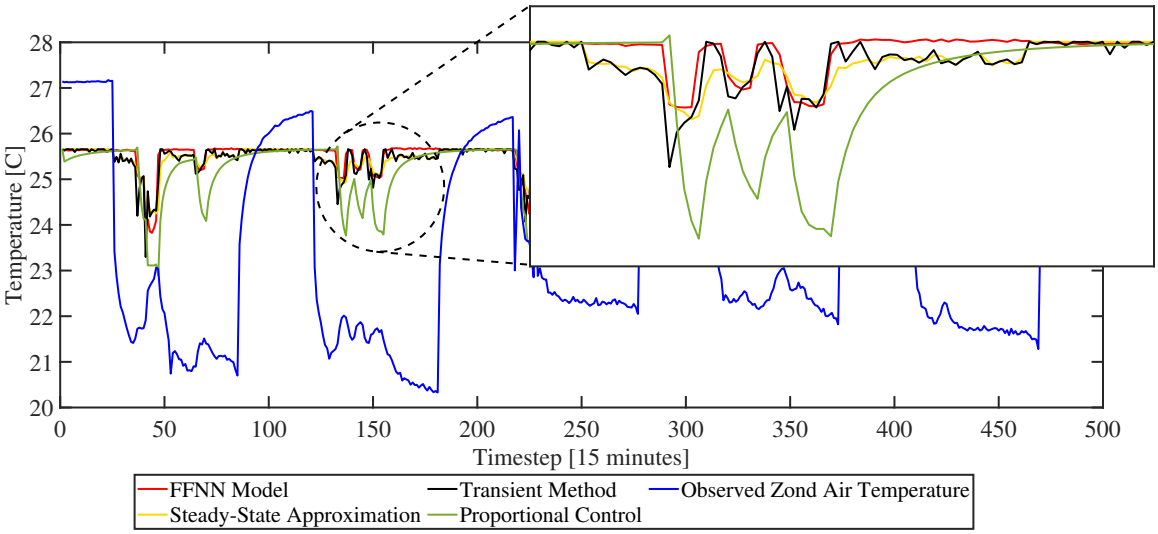

Fig. 5.25.: The magnified inset of temperature values under different scenarios

Table 5.4 draws a comparison by presenting the break-down numbers of each scenario. It reveals that FFNN, steady-state and transient models were closely comparable while the proportional control model resulted in a slightly less average tem-

perature as it proposed higher ventilation amongst others. Figures 5.26, 5.27 and 5.28 also visualize the reported average values of zone air temperature, PMV and PPD indices.

Table 5.4.: Summary of the current and predicted temperature, PMV and PPD

	Zone Air Temperature [°C]			PMV Index [-]			PPD Index [%]		
	Minimum	Maximum	Average	Minimum	Maximum	Average	Minimum	Maximum	Average
Baseline	20.33	26.71	21.95	−1.82	+0.57	−1.24	5.00	68.25	38.05
FFNN	23.82	25.68	25.12	−0.56	+0.16	−0.036	5.00	11.56	5.62
Steady-State	24.24	25.65	25.43	−0.39	+0.14	+0.09	5.00	8.18	5.35
Transient	23.29	25.65	25.39	−0.77	+0.15	+0.074	5.00	17.49	5.45
Proportional	22.88	25.72	24.41	−0.88	+0.19	−0.31	5.00	21.41	8.44

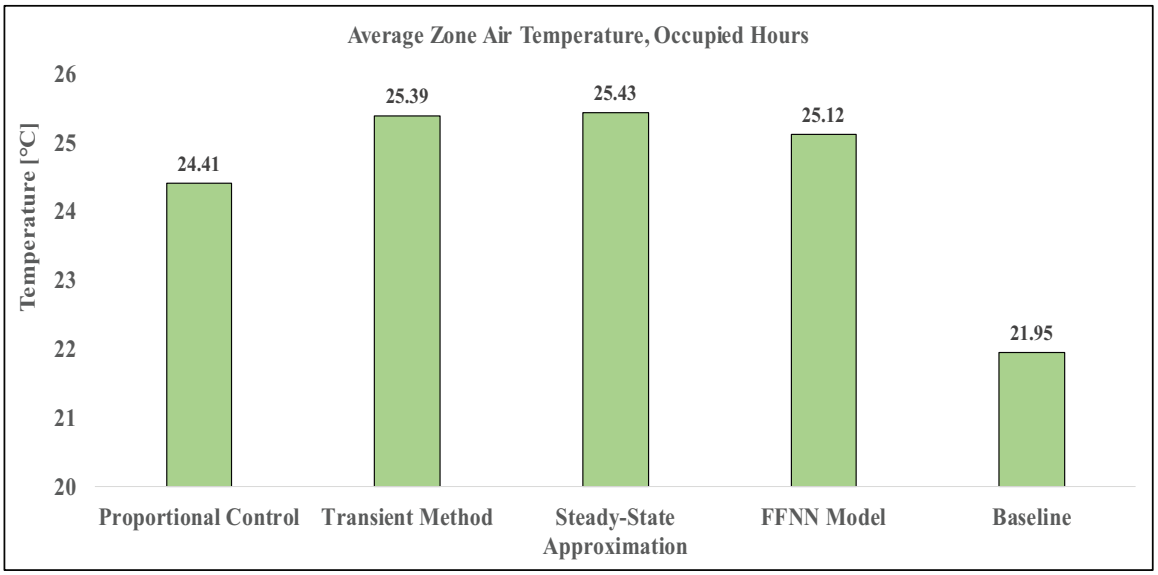

Fig. 5.26.: Comparison of average zone air temperature under different scenarios

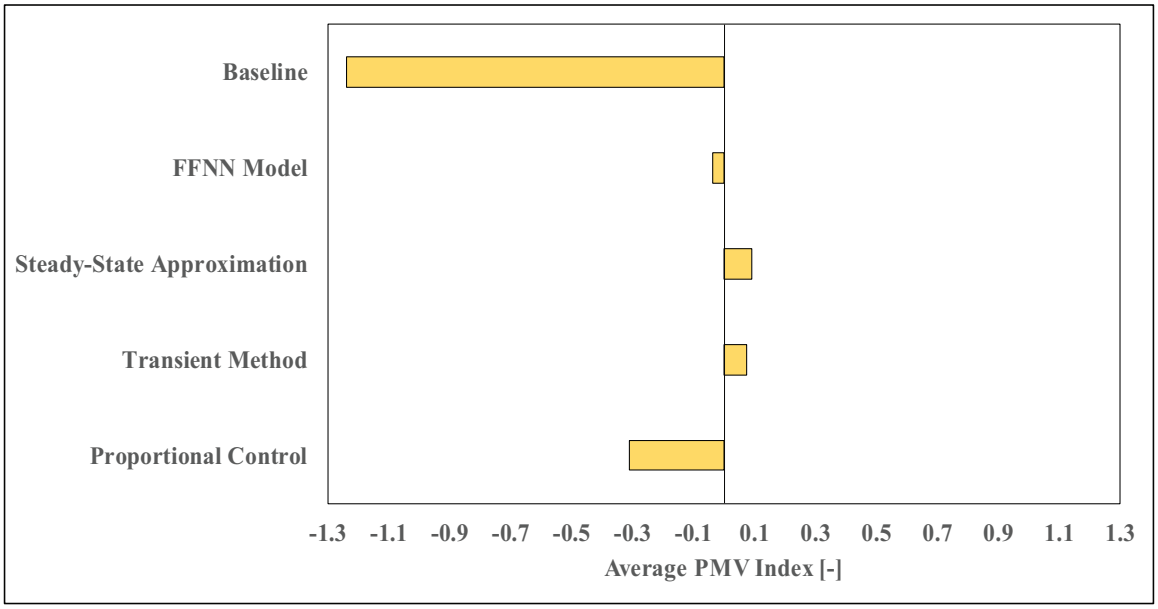

Fig. 5.27.: Comparison of average PMV indices under different scenarios

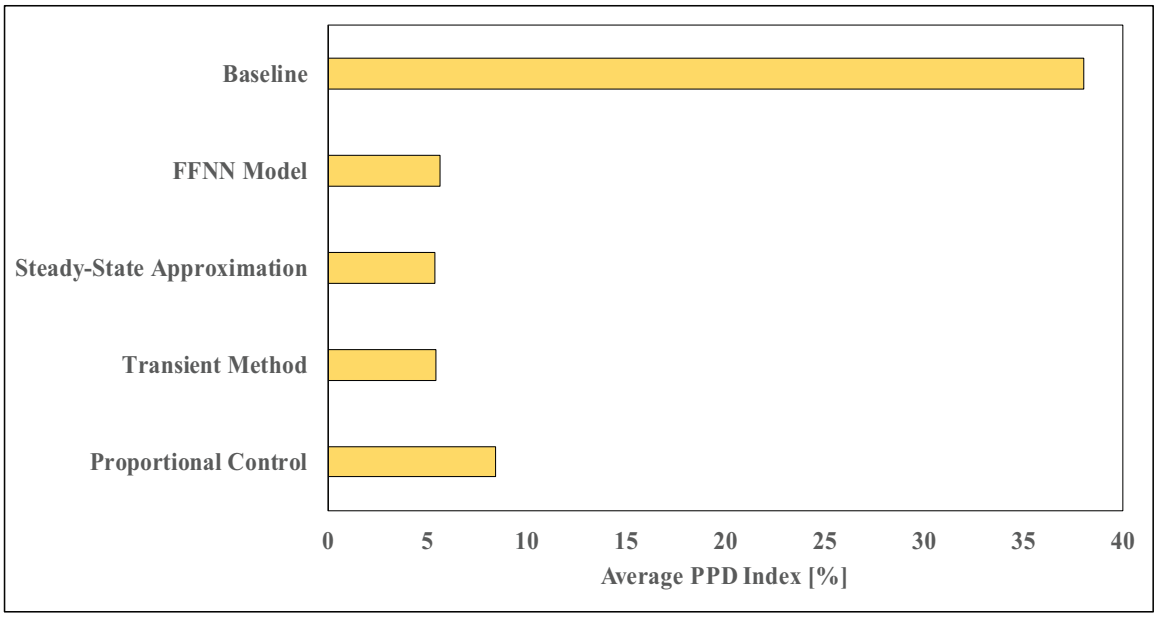

Fig. 5.28.: Comparison of average PPD indices under different scenarios

5.6 Intended PMV Index

The predicted PMV index for all four DCV strategies ranged from -0.6 to $+0.3$ in almost every case number. Although this range is absolutely acceptable according to ASHRAE standard 55 [61], the intended PMV index was defined to be 0 (the most optimal point) only as a demonstration of the concept.

In case of positive pmv indices, the intended pmv index of 0 could be achieved by increasing the outdoor airflow rate (to lower the zone air temperature). As for negative magnitudes of pmv, however, only the heating and cooling coils can be adjusted since further reduction of outdoor airflow rate violates the IAQ standards. Thus, only the predicted pmv indices of above 0 was targeted.

5.6.1 Generalized Reduced Gradient Optimization (GRG Nonlinear)

Generalized Reduced Gradient methods are algorithms for solving smooth nonlinear functions [67]. The modified PMV model, used in this research, is a second-order polynomial function which tends to show smooth behaviour. Figure 5.29 is the optimization flowchart representing the logic behind this mathematical problem. The objective of the algorithm was to solve the PMV_m equation for magnitudes of 0 or closest possible to 0. The algorithm was subjected to a constraint by which the proposed outdoor airflow rate cannot exceed that of baseline (average of 1000 L/s). The Analytic Solver Toolbox was utilized for this optimization.

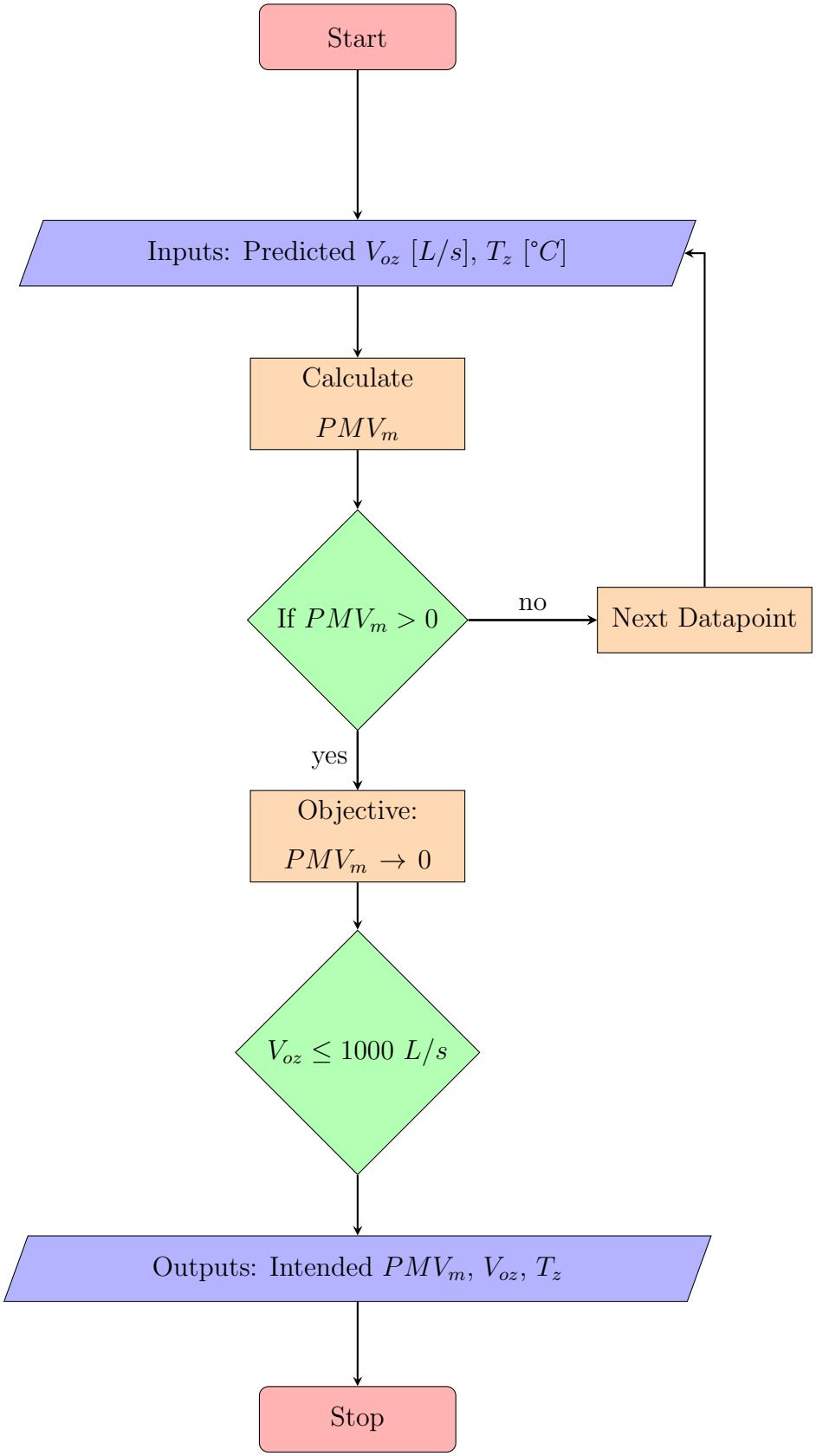

Fig. 5.29.: The procedure on how to achieve the intended PMV index

5.6.2 Optimization Results

Figures 5.30, 5.31, 5.32 and 5.33 show the intended PMV-PPD index for FFNN, steady-state, transient and proportional control models. It is quite obvious that all positive data points moved to 0 where the dissatisfaction percentage is at the lowest level of 5%.

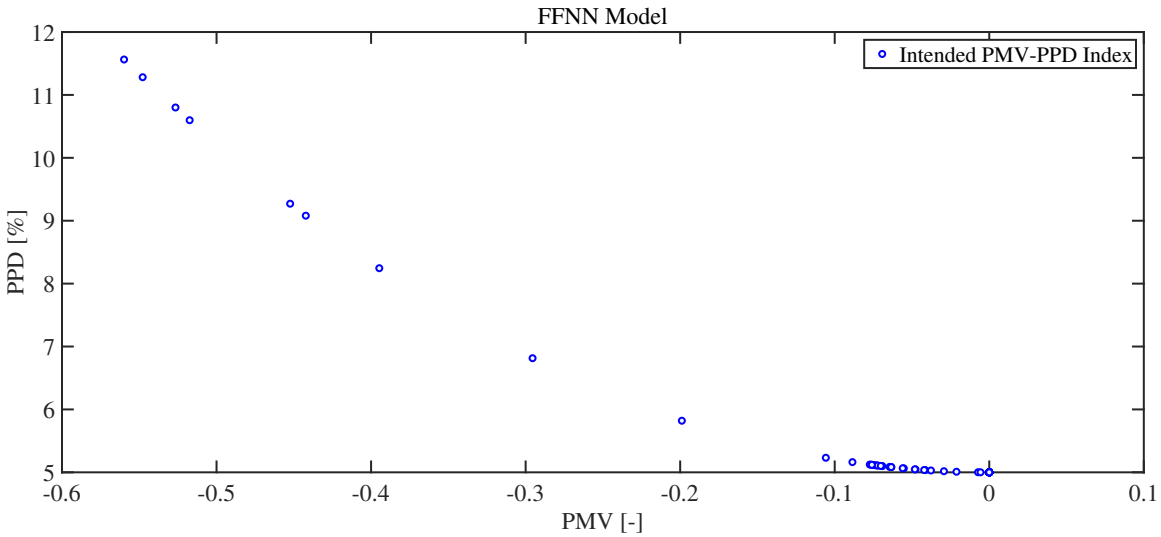

Fig. 5.30.: The intended PMV-PPD index (FFNN model)

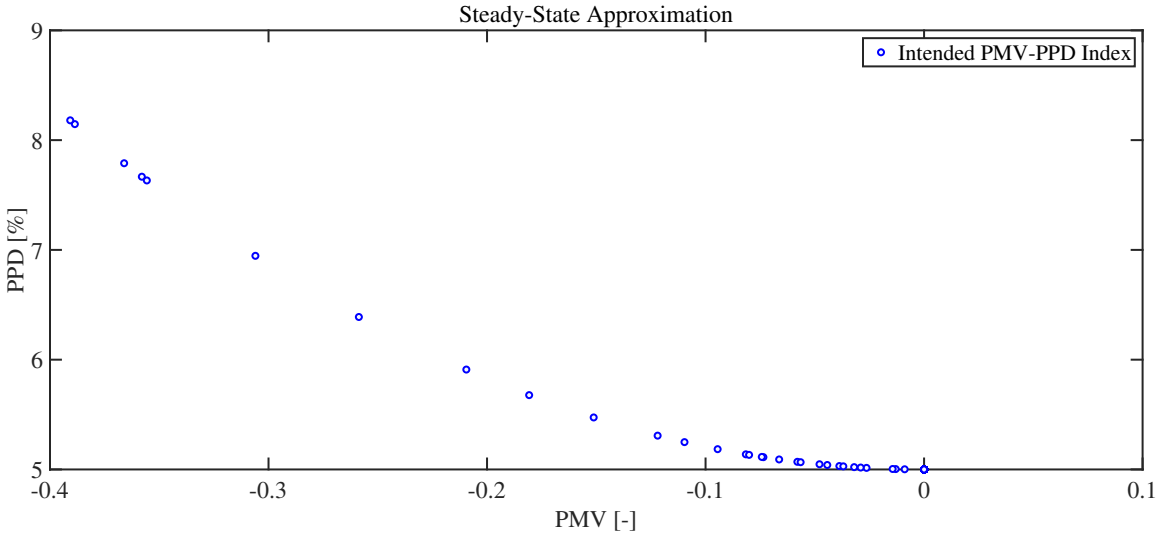

Fig. 5.31.: The intended PMV-PPD index (steady-state approximation)

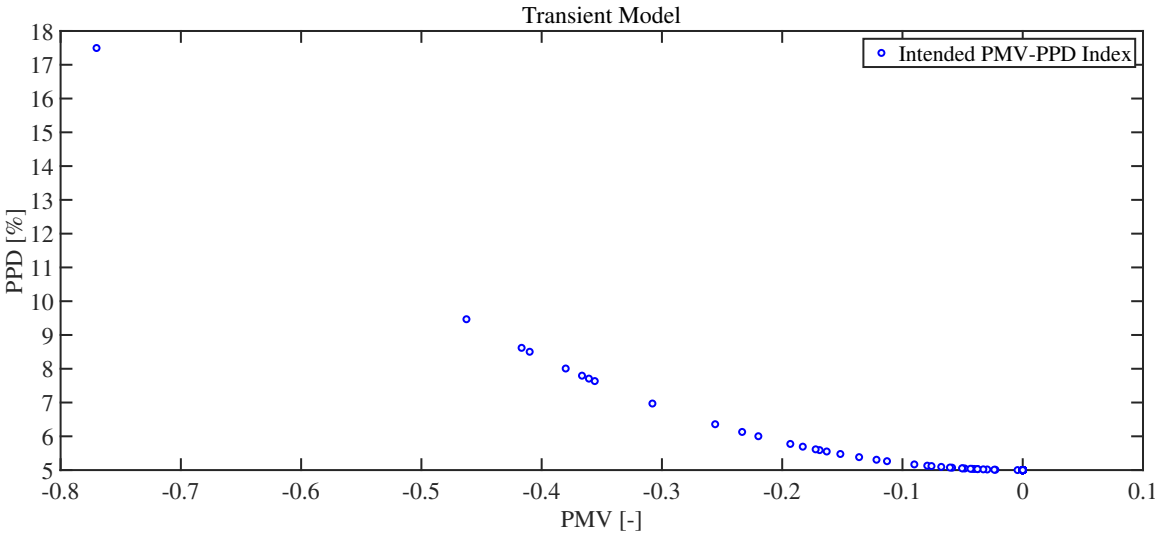

Fig. 5.32.: The intended PMV-PPD index (transient model)

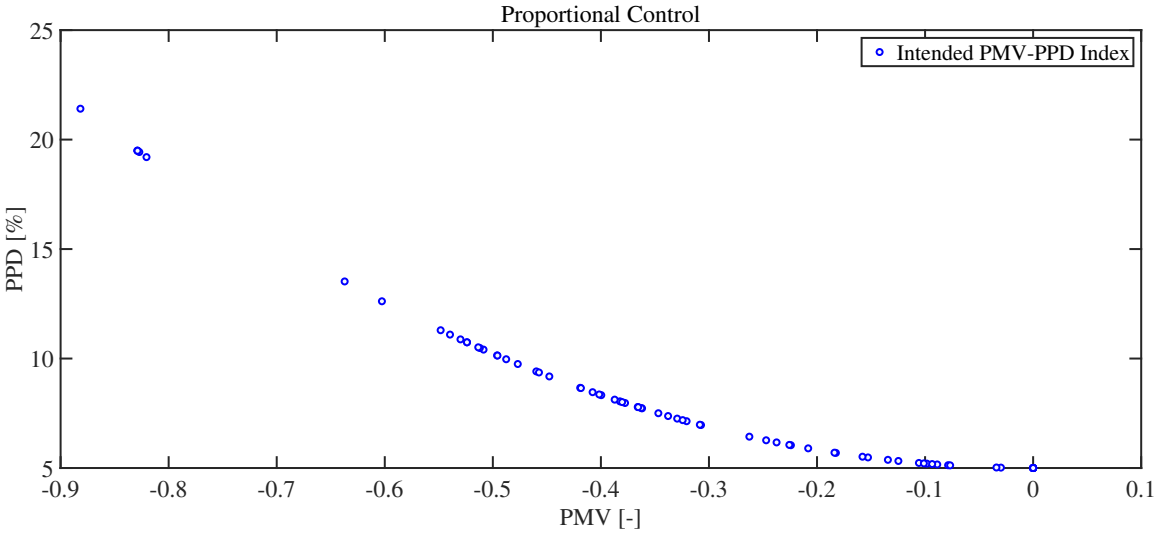

Fig. 5.33.: The intended PMV-PPD index (proportional control)

5.7 The Proposed vs. Intended Ventilation Rates

Figure 5.34 shows the difference between predicted and intended average values of outdoor airflow rate. The intended values are in fact considered as the trade-off between energy saving, IAQ and thermal comfort. As expected, the outdoor

airflow rate showed slightly higher numbers due to considering the auditorium thermal condition; this means less energy savings could be proposed while maintaining both IAQ and thermal comfort constraints. The proportional control and FFNN model had as low as 0.5% and 2.3% increase, respectively. In contrast, the transient method experienced higher rise of 18.7% and the highest of 22.8% for steady-state model. As shown in chapter 4, the steady-state model tended to underestimate the occupancy level and accordingly introduced less ventilation rate; this is the reason why it showed the highest discrepancy percentage.

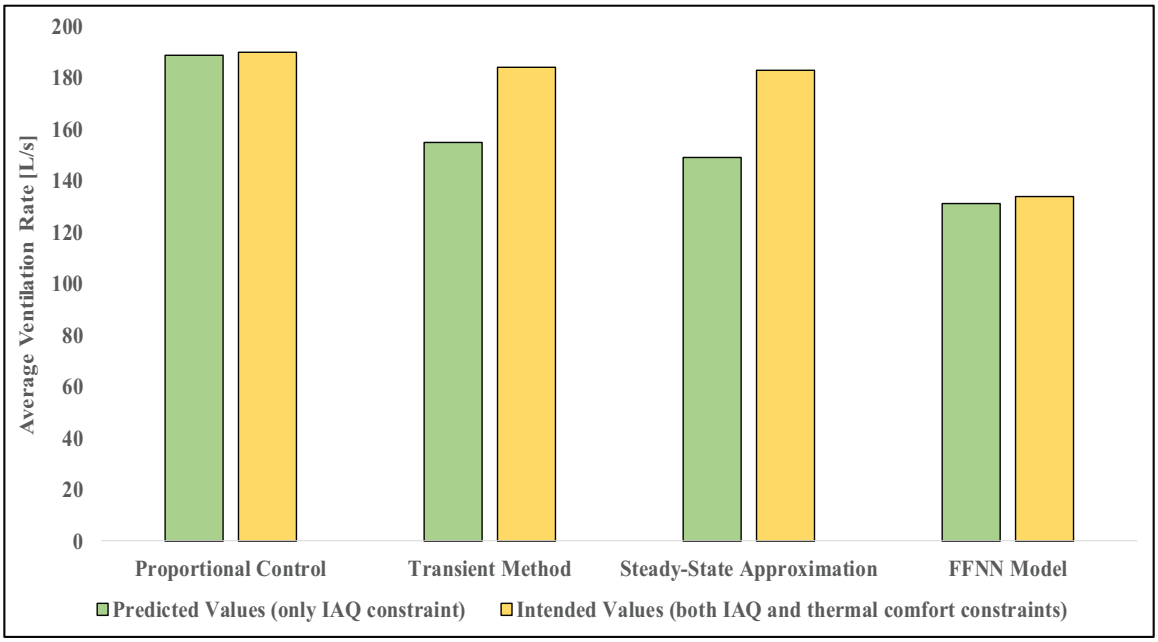

Fig. 5.34.: Comparison between ventilation rates considering IAQ and thermal comfort constraints under different DCV strategies

5.8 Closure on The Chapter

This chapter analyzed the impact of ventilation rate as well as occupancy on the zone air temperature. Considering thermal comfort as the second constraint of energy analysis, it turned out that slightly more ventilation could be introduced to the zone to achieve higher thermal satisfaction. This procedure was done through

optimizing the PMV index using generalized reduced gradient algorithm. By solving PMV for intended values of 0, the new ventilation rate and zone temperature as trade-off values were derived from the equation (5.1). In the next chapter, energy savings will be quantified using the trade-off values; This will ensure that IAQ and thermal comfort constraints are taken into consideration.

6. ENERGY SAVING ANALYSIS AND CONCLUSION

Energy savings can be calculated by converting ventilation reductions from the proposed strategies into electric power. The following model describes this process.

6.1 Power-Airflow Rate Model

The auditorium ventilation is directly provided by the primary fans. Thus, reducing ventilation rate results in lower fans' power consumption and obviously energy savings. To quantify that, there must be a model that relates the primary fans' power to the primary airflow rate. Since fans are powered by electric motors, it would be more convenient to establish the correlation between the primary motors' input power and primary airflow rate. As discussed earlier, the primary airflow is being provided by two identical fans. The airflow rates associated with the primary fans were selected as an inputs while output is the accumulated power of the two electric motors coupled with fans.

6.1.1 Data Preprocessing

During data collection several operation and maintenance issues were reported; therefore, the dataset including motors' input power and primary fans' airflow rates were filtered to remove the outliers. Since the motors' input power data also seemed to be noisy, simple moving mean was deployed over 7 datapoints to smooth out short-term fluctuations and highlight longer-term trends [68]; figure 6.1 represents the actual and smoothed motors' power data for 1000 timesteps.

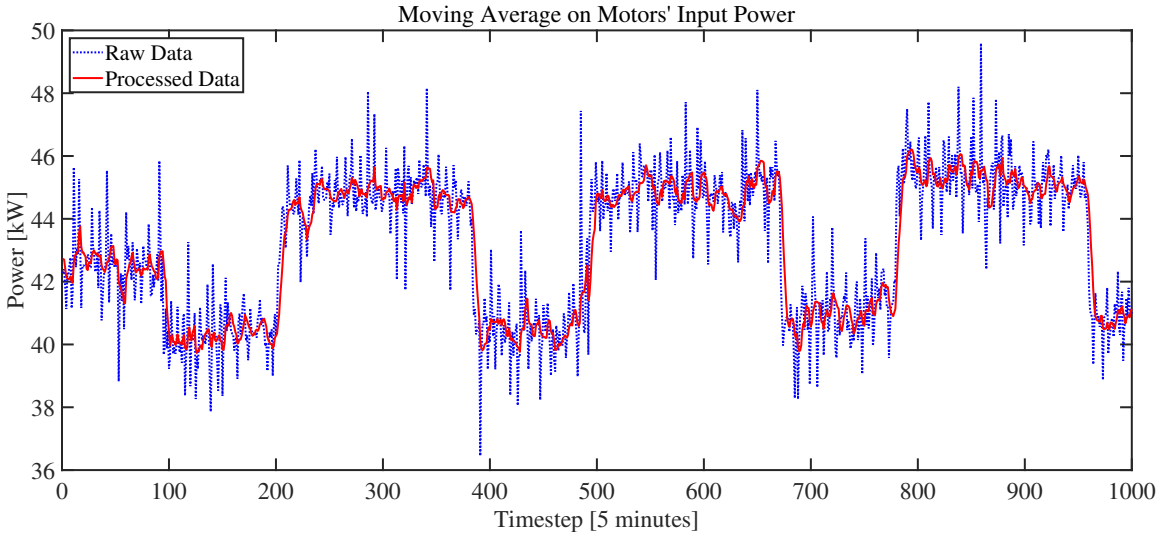

Fig. 6.1.: Moving average on motors' input power

6.1.2 Multiple Linear Regression (MLR)

Multiple linear regression is a linear approach to modeling the relationship between a scalar response (motors' input power) and more than one explanatory variable (primary airflow rates) [69]. In this case, MLR was chosen as a suitable method because of the observed linearity between power and airflow rates. The whole dataset was utilized as the training set, excluding 8 days of data from November 22^{nd} to 29^{th} 2019 for the sake of validation. As for this specific model, a different cross-validation period was selected due to the discontinuity caused by data filtration. The model was developed using MATLAB regression learner toolbox. The equation (6.1) denotes the structure of MLR model.

$$P_m = \Phi_1 V_{p_1} + \Phi_2 V_{p_2} + \Phi_3 V_{p_1} V_{p_2} + \Psi \tag{6.1}$$

where P_m is the electric motors' power, Φ_1, Φ_2 and Φ_3 are fixed coefficients, Ψ is the intercept, V_{p_1} and V_{p_2} are airflow rates of each fan. Table 6.1 shows the magnitude of coefficients as well as intercept.

Table 6.1.: The MLR model parameters

Parameter	Magnitude
Φ_1	-0.0045
Φ_2	0.0098
Φ_3	-3.55×10^{-7}
Ψ	20.38

6.1.3 MLR Model Performance

Figure 6.2 shows the observed and predicted values of motors' power consumption. The predictive MLR model followed the actual data consistently and closely throughout the whole period.

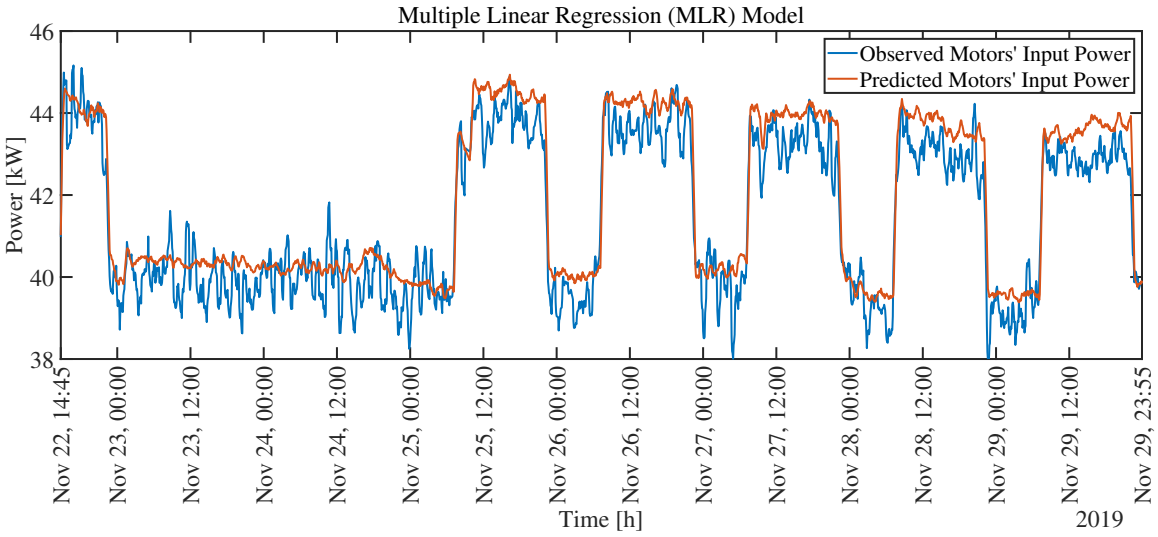

Fig. 6.2.: Predicted vs. observed

Figure 6.3, in addition, represents the linear correlation between the observed and predicted values. Table 6.2 reports the statistical parameters by which the MLR model was evaluated.

Table 6.2.: The MLR model performance parameters

Statistical Parameters	**Magnitude**
R-squared	0.918
P-value	3.9199×10^{-22}
RMSE	0.542

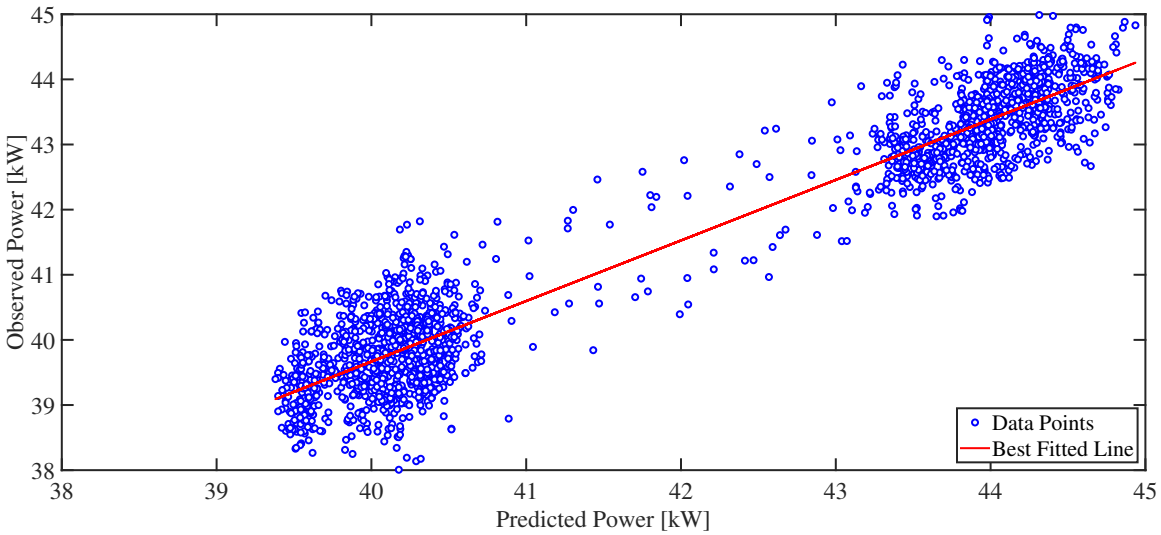

Fig. 6.3.: The MLR best fitted model

In order to portray the discrepancy between actual and predicted values, Pearson residuals were plotted as figure 6.4. It proves the fact that predicted values showed a consistent behaviour towards the observed values.

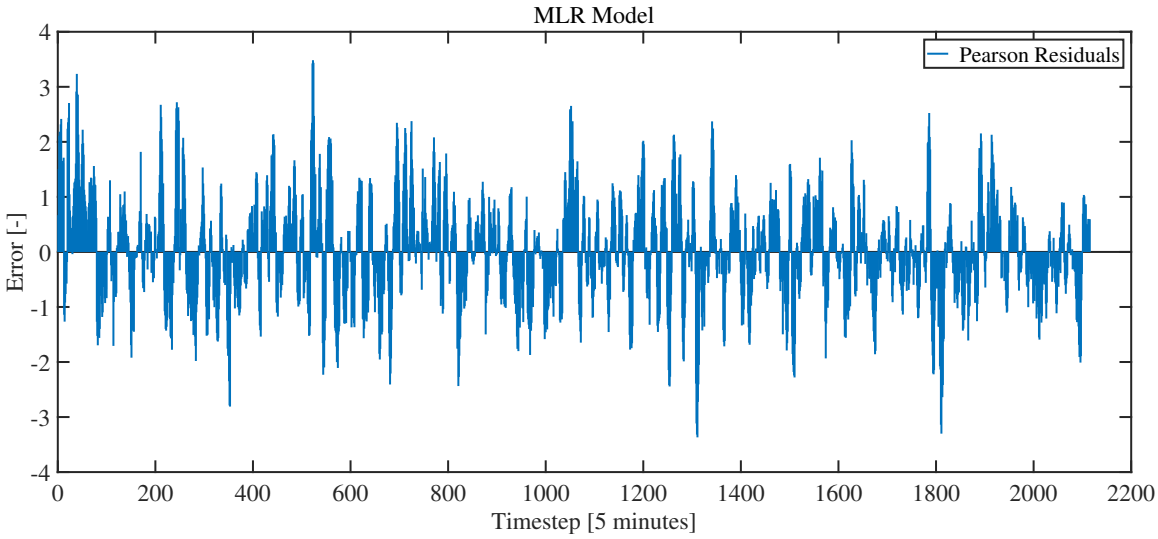

Fig. 6.4.: Pearson residuals (MLR Model)

6.2 Energy Saving Analysis

As discussed in chapter 3, the primary fans can be modulated by adjusting the ventilation air damper; therefore, proposing new ventilation rates by DCV strategies results in new primary airflow rates. Due to mass balance, the auditorium damper ventilation reduction equals to the summation of airflow rate reduction in each primary fan. This can be denoted as the equation (6.2) considering that losses in the ducts are negligible.

$$\Delta V_{oz}^{B-P} \approx \Delta V_{p_1}^{B-P} + \Delta V_{p_2}^{B-P} \tag{6.2}$$

where ΔV_{oz}^{B-P} shows the damper ventilation reduction, $\Delta V_{p_1}^{B-P}$ and $\Delta V_{p_2}^{B-P}$ are the airflow rate reduction of each primary fan. Since both primary fans are identical and synchronised, their airflow rate reduction was observed to be similar. Thus, the equation (6.2) can be reformed as follow:

$$\Delta V_{p_1}^{B-P} = \Delta V_{p_2}^{B-P} \approx \frac{\Delta V_{oz}^{B-P}}{2} \tag{6.3}$$

Using the power-airflow rate model [equations (6.1)] and equation (6.3), the equation (6.4) could be generated to calculate the proposed power consumption of DCV strategies.

$$P_m^P = \Phi_1[V_{p_1} - \frac{\Delta V_{oz}^{B-P}}{2}] + \Phi_2[V_{p_2} - \frac{\Delta V_{oz}^{B-P}}{2}] + \Phi_3[V_{p_1} - \frac{\Delta V_{oz}^{B-P}}{2}][V_{p_2} - \frac{\Delta V_{oz}^{B-P}}{2}] + \Psi \tag{6.4}$$

It is obvious that the subtraction of baseline power consumption from the proposed one gives the power savings. Subsequently, the energy savings were calculated by the equation (6.5).

$$ES = [P_m^B - P_m^P].\Delta t \tag{6.5}$$

where ES stands for energy savings (kWh), P_m^B and P_m^P are the baseline and proposed power consumption (kW), respectively, and Δt is the time interval (h).

6.2.1 Results

Considering the trade-off values of ventilation rate introduced by DCV strategies, the results of energy savings are summarized in table 6.3. Figure 6.5, in addition, portrays this summary. The maximum energy saving was achieved by utilizing FFNN for occupancy detection in dynamic per-occupant control, accounting for 74.2%. Steady-state and transient models introduced almost equal savings of 65.6% and 65.3%, respectively. It should be noted that the steady-state model tended to underestimate the occupancy. This is the main reason that it resulted in a slightly higher energy saving. Proportional control, however, did not take the occupancy estimation into account, and therefore 64.2% reduction of energy was achieved.

Table 6.3.: Summary of ventilation energy savings under different control strategies

Control Strategy	Average Fan Energy Consumption [kWh/day]	Saving Compared to Baseline [%]
Baseline	1,002.4	-
Proportional Control	358.8	64.2%
Steady-State Approximation	344.8	65.6%
Transient Method	347.8	65.3%
FFNN Model	258.6	74.2%

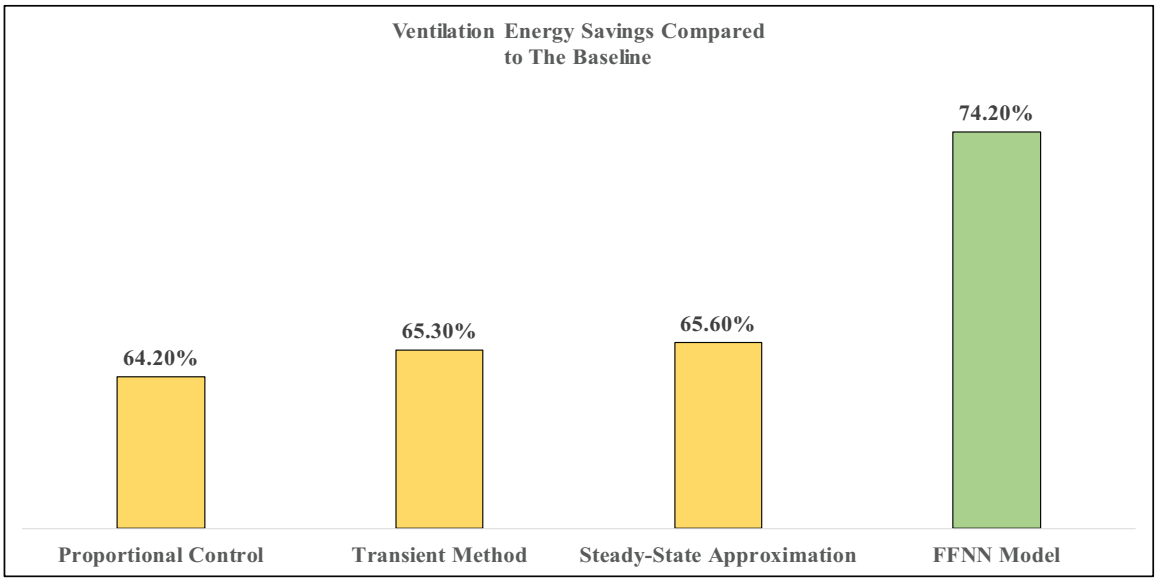

Fig. 6.5.: Ventilation energy savings of the proposed DCV strategies, compared to baseline

6.2.2 Energy Saving Rationale

This research was conducted from September 2019 to January 2020. Throughout this period, the outdoor air temperature ranged from $-11°C$ to $32°C$, meaning that

both cooling and heating seasons were included. Figure 6.6 displays the outdoor versus zone air temperature for this period. As is clear, temperature inside the auditorium remained consistent with a constant variation. In previous chapters, it was shown that ventilation operates with a two-position control between occupied and unoccupied modes to maintain inside temperature. Besides ventilation, the cooling and heating coils must consume energy to compensate for temperature fluctuations. This study only focused on the ventilation side of HVACs. Thus, it can be concluded that the reported ventilation energy savings could be generalized for all year round with an acceptable accuracy.

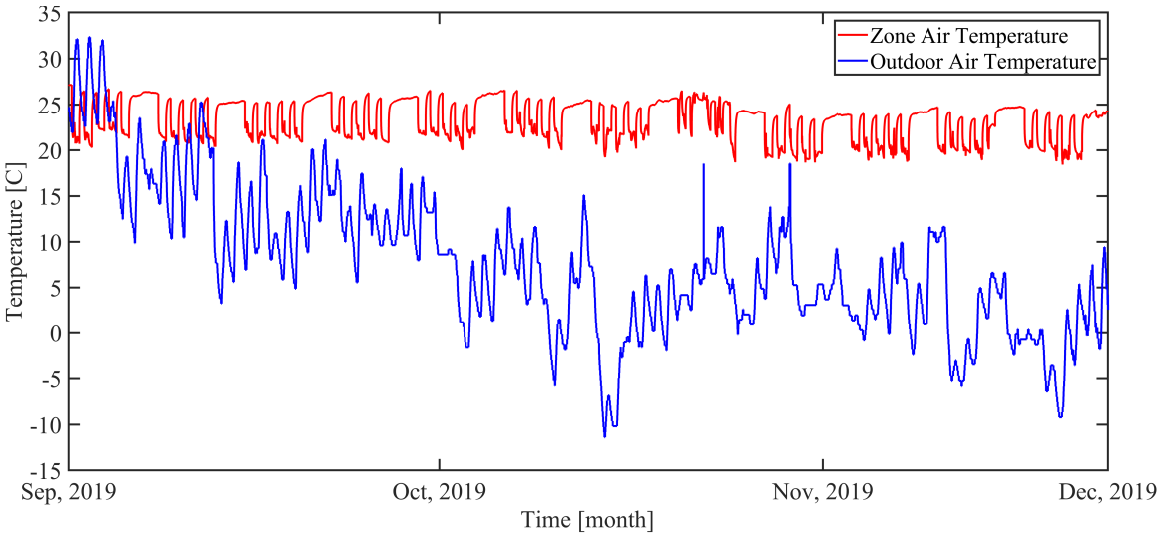

Fig. 6.6.: The auditorium vs. outdoor air temperature

6.3 Discussion and Conclusion

This research attempted to explore the applicability of demand-controlled ventilation strategies and their potentials for saving energy on HVAC systems considering indoor air quality and thermal comfort standards. DCV strategies were categorized as dynamic per-occupant and CO_2-based methods. With respect to the former, feed-forward neural network, steady-state and transient occupancy prediction models were evaluated whilst proportional control as the latter.

The case study was a lecture hall as a densely populated single zone. That is, the results of aforementioned DCV strategies can only be generalized to the similar situations.

Compared to the conventional physics-based DCV strategies, the benefits of the newly introduced ANN model was investigated. It was proven that the two-layer feed-forward neural network algorithm is able to predict human occupancy with the highest accuracy. Comparison amongst the four tested methods demonstrated that the application of machine learning algorithms resulted in 74.2% energy savings against the baseline, about 28% against the proportional control, and 25% against steady-state and transient models. This reveals the fact that higher precision in occupancy prediction leads to higher ventilation energy savings.

Transient method, in predicting occupancy and subsequently introducing ventilation rate, seemed to overshoot from vacancy to occupancy. However, steady-state approximation undershot almost over the entire period. Overall, the average of introduced ventilation in both models turned out to be nearly identical.

Proportional control offered the lowest saving. Direct use of CO_2 concentrations rather than predicting the number of occupants created saving opportunities. This is because of its slow reaction to the zone occupancy level and overestimating the ventilation rate.

All the proposed DCV strategies maintained indoor CO_2 concentrations within the threshold recommended by ASHRAE 62.1 standard. As for indoor thermal condition, the DCV methods showed positive contribution when compared to the current practice; this is because of regulating current over-ventilation proportional to the number of indoor occupants. Despite their positive impacts, the thermal comfort was further optimized to the most optimum point (PMV=0). This, as the second constraint, required slightly higher ventilation to be introduced into the zone. The results of ventilation, when meeting both IAQ and thermal comfort, were labelled as intended or trade-off values.

6.4 Recommendations

The following recommendations may potentially pave the way for the research scopes of same nature:

- As mentioned in chapter 2, a combination of sensors were employed to achieve more accurate occupancy predictions. The FFNN algorithm utilized CO_2 concentrations, the associated ventilation rates and the zone schedule. Since updating the zone schedule might be a hassle in a continuous operating DCV-based system, automatic occupancy detection sensors (e.g., motion detectors) will be better options.

- Although the FFNN model resulted in a better energy saving, its implementation requires more complex programming on the BMS communication protocol (BACnet). Yet, the steady-state and transient models offered highly acceptable results with a less complexity of implementation.

- To better comprehend the operation of the case-study, this research exploited a higher quantity of temperature sensors than necessary (7 sensors). For the same type of projects, it may go down to 4 sensors, measuring the temperature of outdoor air, zone, supplied and return air.

6.5 Future Scope

The present work attempted to achieve energy savings by optimizing only ventilation rates through indoor occupancy prediction. IAQ and thermal comfort, as two crucial factors, were considered a well. The energy saving results would be further enhanced if including the water side of the HVAC system. The heating and cooling coils could be modeled and attached to the optimization modules. Having all modules lumped together, the trade-off between savings and constraints will become more precise.

6.6 Closure on the Chapter

Chapter 6 introduced the power-airflow rate model for analyzing the ventilation energy savings proposed by the DCV strategies. Comparison of the proposed savings was drawn against the baseline and against each other. The overall conclusion of this study along with future works were also discussed.

REFERENCES

REFERENCES

[1] S. Salimi and A. Hammad, "Critical review and research roadmap of office building energy management based on occupancy monitoring," *Energy and Buildings*, vol. 182, pp. 214–241, 2019.

[2] W. Jung and F. Jazizadeh, "Human-in-the-loop HVAC operations: A quantitative review on occupancy, comfort, and energy-efficiency dimensions," *Applied Energy*, vol. 239, pp. 1471–1508, 2019.

[3] V. L. Erickson, Y. Lin, A. Kamthe, R. Brahme, A. Surana, A. E. Cerpa, M. D. Sohn, and S. Narayanan, "Energy efficient building environment control strategies using real-time occupancy measurements," *BUILDSYS 2009 - Proceedings of the 1st ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings, Held in Conjunction with ACM SenSys 2009*, pp. 19–24, 2009.

[4] V. L. Erickson and A. E. Cerpa, "Occupancy based demand response HVAC control strategy," *BuildSys'10 - Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*, pp. 7–12, 2010.

[5] G. Brager, H. Zhang, and E. Arens, "Evolving opportunities for providing thermal comfort," *Building Research and Information*, 2015.

[6] Y. Agarwal, B. Balaji, S. Dutta, R. K. Gupta, and T. Weng, "Duty-cycling buildings aggressively: The next frontier in HVAC control," in *Proceedings of the 10th ACM/IEEE International Conference on Information Processing in Sensor Networks, IPSN'11*, 2011.

[7] C. Lin, C. C. Federspiel, and D. M. Auslander, "Multi-sensor single-actuator control of HVAC systems," *International Conference for Enhanced Building Operations*, 2002.

[8] S. Karjalainen and O. Koistinen, "User problems with individual temperature control in offices," *Building and Environment*, 2007.

[9] F. Jazizadeh, A. Ghahramani, B. Becerik-Gerber, T. Kichkaylo, and M. Orosz, "Human-building interaction framework for personalized thermal comfort-driven systems in office buildings," *Journal of Computing in Civil Engineering*, 2014.

[10] C. Huizenga, K. Laeser, and E. Arens, "A web-based occupant satisfaction survey for benchmarking building quality Author:," *Indoor Air*, 2002.

[11] S. Karjalainen, "Thermal comfort and use of thermostats in Finnish homes and offices," *Building and Environment*, 2009.

[12] M. J. Brandemuehl and J. E. Braun, "Impact of demand-controlled and economizer ventilation strategies on energy use in buildings," in *ASHRAE Transactions*, 1999.

[13] W. Shen, G. Newsham, and B. Gunay, "Leveraging existing occupancy-related data for optimal control of commercial office buildings: A review," *Advanced Engineering Informatics*, vol. 33, pp. 230–242, 2017.

[14] Z. Yang and B. Becerik-Gerber, "The coupled effects of personalized occupancy profile based HVAC schedules and room reassignment on building energy use," *Energy and Buildings*, 2014.

[15] Y. Agarwal, B. Balaji, R. Gupta, J. Lyles, M. Wei, and T. Weng, "Occupancy-driven energy management for smart building automation," in *BuildSys'10 - Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*, 2010.

[16] L. J. Lo and A. Novoselac, "Localized air-conditioning with occupancy control in an open office," *Energy and Buildings*, vol. 42, no. 7, pp. 1120–1128, 2010.

[17] V. L. Erickson, M. Carreira-Perpiñán, and A. E. Cerpa, "OBSERVE: Occupancy-based system for efficient reduction of HVAC energy," *Proceedings of the 10th ACM/IEEE International Conference on Information Processing in Sensor Networks, IPSN'11*, pp. 258–269, 2011.

[18] A. Aswani, N. Master, J. Taneja, D. Culler, and C. Tomlin, "Reducing transient and steady state electricity consumption in HVAC using learning-based model-predictive control," *Proceedings of the IEEE*, vol. 100, no. 1, pp. 240–253, 2012.

[19] S. Goyal, H. A. Ingley, and P. Barooah, "Occupancy-based zone-climate control for energy-efficient buildings: Complexity vs. performance," *Applied Energy*, vol. 106, pp. 209–221, 2013.

[20] B. Balaji, J. Xu, A. Nwokafor, R. Gupta, and Y. Agarwal, "Sentinel: Occupancy based HVAC actuation using existing wifi infrastructure within commercial buildings," *SenSys 2013 - Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*, 2013.

[21] F. Oldewurtel, D. Sturzenegger, and M. Morari, "Importance of occupancy information for building climate control," *Applied Energy*, 2013.

[22] S. West and J. Ward, "Trial results from a model predictive control and optimisation system for commercial building HVAC," *Energy and Buildings*, vol. 72, pp. 271–279, 2014.

[23] J. R. Dobbs and B. M. Hencey, "Model predictive HVAC control with online occupancy model," *Energy and Buildings*, vol. 82, pp. 675–684, 2014.

[24] J. Brooks, S. Kumar, S. Goyal, R. Subramany, and P. Barooah, "Energy-efficient control of under-actuated HVAC zones in commercial buildings," *Energy and Buildings*, vol. 93, pp. 160–168, 2015.

[25] H. B. Gunay, W. O'Brien, I. Beausoleil-Morrison, P. Bisaillon, and Z. Shi, "Development and implementation of control-oriented models for terminal heating and cooling units," *Energy and Buildings*, vol. 121, pp. 78–91, 2016.

[26] F. Wang, Q. Feng, Z. Chen, Q. Zhao, Z. Cheng, J. Zou, Y. Zhang, J. Mai, Y. Li, and H. Reeve, "Predictive control of indoor environment using occupant number detected by video data and CO 2 concentration," *Energy and Buildings*, 2017.

[27] A. Capozzoli, M. S. Piscitelli, A. Gorrino, I. Ballarini, and V. Corrado, "Data analytics for occupancy pattern learning to reduce the energy consumption of HVAC systems in o ffi ce buildings," *Sustainable Cities and Society*, vol. 35, no. February, pp. 191–208, 2017.

[28] Y. Peng, A. Rysanek, Z. Nagy, and A. Schlüter, "Occupancy learning-based demand-driven cooling control for office spaces," *Building and Environment*, vol. 122, pp. 145–160, 2017.

[29] S. Nagarathinam, H. Doddi, A. Vasan, V. Sarangan, P. Venkata Ramakrishna, and A. Sivasubramaniam, "Energy efficient thermal comfort in open-plan office buildings," *Energy and Buildings*, vol. 139, pp. 476–486, 2017.

[30] Z. Chen, C. Jiang, and L. Xie, "Building occupancy estimation and detection: A review," *Energy and Buildings*, vol. 169, pp. 260–270, 2018.

[31] R. H. Dodier, G. P. Henze, D. K. Tiller, and X. Guo, "Building occupancy detection through sensor belief networks," *Energy and Buildings*, 2006.

[32] F. Wahl, M. Milenkovic, and O. Amft, "A distributed PIR-based approach for estimating people count in office environments," in *Proceedings - 15th IEEE International Conference on Computational Science and Engineering, CSE 2012 and 10th IEEE/IFIP International Conference on Embedded and Ubiquitous Computing, EUC 2012*, 2012.

[33] M. S. Zuraimi, A. Pantazaras, K. A. Chaturvedi, J. J. Yang, K. W. Tham, and S. E. Lee, "Predicting occupancy counts using physical and statistical Co2-based modeling methodologies," *Building and Environment*, vol. 123, pp. 517–528, 2017.

[34] D. Calì, P. Matthes, K. Huchtemann, R. Streblow, and D. Müller, "CO2 based occupancy detection algorithm: Experimental analysis and validation for office and residential buildings," *Building and Environment*, vol. 86, pp. 39–49, 2015.

[35] C. Jiang, M. K. Masood, Y. C. Soh, and H. Li, "Indoor occupancy estimation from carbon dioxide concentration," *Energy and Buildings*, vol. 131, pp. 132–141, 2016.

[36] T. H. Pedersen, K. U. Nielsen, and S. Petersen, "Method for room occupancy detection based on trajectory of indoor climate sensor data," *Building and Environment*, vol. 115, pp. 147–156, 2017.

[37] S. H. Ryu and H. J. Moon, "Development of an occupancy prediction model using indoor environmental data based on machine learning techniques," *Building and Environment*, vol. 107, pp. 1–9, 2016.

[38] T. Lu, X. Lü, and M. Viljanen, "A novel and dynamic demand-controlled ventilation strategy for CO 2 control and energy saving in buildings," *Energy and Buildings*, vol. 43, no. 9, pp. 2499–2508, 2011.

[39] M. B. Schell, S. C. Turner, and R. O. Shim, "Application of CO2-based demand-controlled ventilation using ASHRAE Standard 62: optimizing energy use and ventilation," pp. 1213–1225, 1998.

[40] Y. P. Ke and S. A. Mumma, "Using carbon dioxide measurements to determine occupancy for ventilation controls," *ASHRAE Transactions*, vol. 103, no. pt 2, pp. 365–374, 1997.

[41] V. L. Erickson, M. Carreira-Perpiñán, and A. E. Cerpa, "OBSERVE: Occupancy-based system for efficient reduction of HVAC energy," *Proceedings of the 10th ACM/IEEE International Conference on Information Processing in Sensor Networks, IPSN'11*, pp. 258–269, 2011.

[42] F. Fleuret, J. Berclaz, R. Lengagne, and P. Fua, "Multicamera people tracking with a probabilistic occupancy map," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008.

[43] J. Zou, Q. Zhao, W. Yang, and F. Wang, "Occupancy detection in the office by analyzing surveillance videos and its application to building energy conservation," *Energy and Buildings*, 2017.

[44] B. Balaji, J. Xu, A. Nwokafor, R. Gupta, and Y. Agarwal, "Sentinel: Occupancy based HVAC actuation using existing wifi infrastructure within commercial buildings," *SenSys 2013 - Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*, 2013.

[45] W. Wang, J. Chen, and X. Song, "Modeling and predicting occupancy profile in office space with a Wi-Fi probe-based Dynamic Markov Time-Window Inference approach," *Building and Environment*, 2017.

[46] S. T. Taylor, A. K. Persily, and D. C. Burge, "Ventilation for Acceptable Indoor Air Quality," vol. 2016, 2016.

[47] Z. Afroz, G. M. Shafiullah, T. Urmee, and G. Higgins, "Modeling techniques used in building HVAC control systems: A review," *Renewable and Sustainable Energy Reviews*, vol. 83, no. June 2017, pp. 64–84, 2018.

[48] A. Afram and F. Janabi-Sharifi, "Review of modeling methods for HVAC systems," *Applied Thermal Engineering*, vol. 67, no. 1-2, pp. 507–519, 2014.

[49] C. Ghiaus, A. Chicinas, and C. Inard, "Grey-box identification of air-handling unit elements," *Control Engineering Practice*, vol. 15, no. 4, pp. 421–433, 2007.

[50] Shengwei Wang and Xinqiao Jin, "CO 2-Based Occupancy Detection for On-Line Outdoor Air Flow Control," *Indoor and Built Environment*, 1998.

[51] L. I. Zheng, Z. Cheng, L. I. Xu, and L. I. Tonghua, "Nonlinear Fitting by Using a Neural Net Algorithm," *Analytical Chemistry*, 1993.

[52] A. Zell, N. Mache, M. Hüttel, and M. Vogt, "Simulation Neuronaler Netze auf Massiv Parallelen Rechnern," 1993.

[53] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: A review and evaluation," *IEEE Transactions on Power Systems*, 2001.

[54] P. Kim, "Machine Learning," in *MATLAB Deep Learning: With Machine Learning, Neural Networks and Artificial Intelligence*. Berkeley, CA: Apress, 2017, pp. 1–18.

[55] R. Hecht-Nielsen, "Theory of the backpropagation neural network." Publ by IEEE, 1989, pp. 593–605.

[56] S. J. Emmerich, C. Howard-Reed, and S. J. Nabinger, "Validation of multizone IAQ model predictions for tracer gas in a townhouse," *Building Services Engineering Research and Technology*, 2004.

[57] V. Barnett, J. Neter, and W. Wasserman, "Applied Linear Statistical Models." *Journal of the Royal Statistical Society. Series A (General)*, 1975.

[58] C. T. Washington, "1910.1000, air contaminants," *Code of federal regulations. Title 29, Labor*, 1983.

[59] ASHRAE, "ASHRAE Standard 62-1989 Ventilation for Acceptable Indoor Air Quality," *American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Inc.*, 1990.

[60] M. Schell and D. Int-Hout, "Demand control ventilation using CO2," *ASHRAE Journal*, 2001.

[61] ANSI/ASHRAE, "ANSI/ASHRAE Standard 55-2017 : Thermal Environmental Conditions for Human Occupancy," *ASHRAE Inc.*, vol. 2017, p. 66, 2017.

[62] S. Zhang, Y. Cheng, M. Olaide Oladokun, Y. Wu, and Z. Lin, "Improving predicted mean vote with inversely determined metabolic rate," *Sustainable Cities and Society*, vol. 53, no. October 2019, p. 101870, 2020.

[63] H. Han, J. Lee, J. Kim, C. Jang, and H. Jeong, "Thermal comfort control based on a simplified predicted mean vote index," *Energy Procedia*, vol. 61, pp. 970–974, 2014.

[64] J. Gao, Y. Wang, and P. Wargocki, "Comparative analysis of modified PMV models and SET models to predict human thermal sensation in naturally ventilated buildings," *Building and Environment*, vol. 92, pp. 200–208, 2015.

[65] S. Zhang, Y. Cheng, Z. Fang, C. Huan, and Z. Lin, "Optimization of room air temperature in stratum-ventilated rooms for both thermal comfort and energy saving," *Applied Energy*, vol. 204, no. October, pp. 420–431, 2017.

[66] C. E. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning. 2006*, 2006, vol. 38, no. 2.

[67] L. S. Lasdon, R. L. Fox, and M. W. Ratner, "Nonlinear Optimization Using The Generalized Reduced Gradient Method." *Rev Fr Autom Inf Rech Oper*, 1974.

[68] B. N. Bauer and Y.-l. Chou, "Statistical Analysis: With Business and Economic Applications." *Journal of the American Statistical Association*, 1971.

[69] D. A. Freedman, *Statistical models: Theory and practice*, 2009.

VITA

VITA

EDUCATION

- M.Sc., Mechanical Engineering, IUPUI, 2017-2020
- B.Sc., Mechanical Engineering, Shahid Beheshti University, 2011-2015

AWARDS AND ACHIEVEMENTS

- Graduate Research Assistantship Award, IUPUI, 2018-2020
- Fellowship Award, IUPUI, 2017-2018
- Elite Student Award (ranked 2^{nd}), Shahid Beheshti University, 2015

WORK EXPERIENCE

- Energy Engineer, Industrial Assessment Center, Sponsored by D.O.E, 2017-2020
- Energy Engineering Intern/Project Manager, Telamon Corporation, 2019

CERTIFICATIONS

- Energy Management and Assessment, Purdue University, 2019
- Building Management Systems Training, Delta Controls, 2019

PUBLICATIONS

- **Mehdi Momeni**, Da-Chun Wu, A. Razban, J. Chen, *"Data-driven Demand Control Ventilation Using Machine Learning CO2-based Occupancy Detection Method"*, ECOS 2020, Accepted.
- K. Nagasaka, A. Amini, **Mehdi Momeni**, *"WinDam: A Novel Airborne Wind Turbine"*, Journal of Clean Energy Technologies, vol. 5, no. 3, 2016, pp. 243-247.

APPENDICES

A. GPR MODEL

The GPR model was optimized based on the least mean square error (MSE) over 30 iterations. Figure A.1 shows this process.

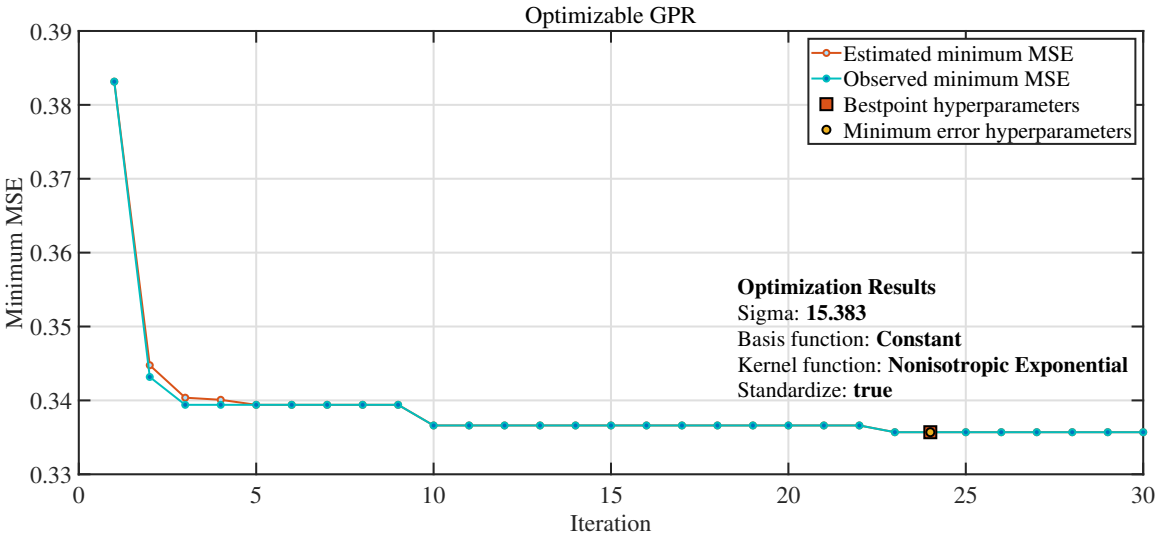

Fig. A.1.: The GPR model optimization process

The MATLAB code used for this model is as follow:

```
1 function [trainedModel, validationRMSE] = ...
      trainRegressionModel(trainingData)
2 % [trainedModel, validationRMSE] = trainRegressionModel(trainingData)
3 % returns a trained regression model and its RMSE. This code ...
      recreates the
4 % model trained in Regression Learner app. Use the generated code to
5 % automate training the same model with new data, or to learn how to
6 % programmatically train models.
7 %
```

```
8  %  Input:
9  %      trainingData: a table containing the same predictor and ...
       response
10 %       columns as imported into the app.
11 %
12 %  Output:
13 %      trainedModel: a struct containing the trained regression ...
       model. The
14 %       struct contains various fields with information about the ...
       trained
15 %       model.
16 %
17 %      trainedModel.predictFcn: a function to make predictions on ...
       new data.
18 %
19 %      validationRMSE: a double containing the RMSE. In the app, the
20 %       History list displays the RMSE for each model.
21 %
22 % Use the code to train the model with new data. To retrain your ...
       model,
23 % call the function from the command line with your original data ...
       or new
24 % data as the input argument trainingData.
25 %
26 % For example, to retrain a regression model trained with the ...
       original data
27 % set T, enter:
28 %   [trainedModel, validationRMSE] = trainRegressionModel(T)
29 %
30 % To make predictions with the returned 'trainedModel' on new ...
       data T2, use
31 %   yfit = trainedModel.predictFcn(T2)
32 %
33 % T2 must be a table containing at least the same predictor ...
       columns as used
```

```
% during training. For details, enter:
%   trainedModel.HowToPredict

% Auto-generated by MATLAB on 06-Apr-2020 14:03:03

% Extract predictors and response
% This code processes the data into the right shape for training the
% model.
inputTable = trainingData;
predictorNames = {'Voz', 'Pz'};
predictors = inputTable(:, predictorNames);
response = inputTable.Tz;
isCategoricalPredictor = [false, false];

% Train a regression model
% This code specifies all the model options and trains the model.
regressionGP = fitrgp(...
    predictors, ...
    response, ...
    'BasisFunction', 'constant', ...
    'KernelFunction', 'exponential', ...
    'Standardize', true);

% Create the result struct with predict function
predictorExtractionFcn = @(t) t(:, predictorNames);
gpPredictFcn = @(x) predict(regressionGP, x);
trainedModel.predictFcn = @(x) ...
    gpPredictFcn(predictorExtractionFcn(x));

% Add additional fields to the result struct
trainedModel.RequiredVariables = {'Pz', 'Voz'};
trainedModel.RegressionGP = regressionGP;
trainedModel.About = 'This struct is a trained model exported ...
    from Regression Learner R2019b.';
```

```
66 trainedModel.HowToPredict = sprintf('To make predictions on a new ...
       table, T, use: \n  yfit = c.predictFcn(T) \nreplacing ''c'' ...
       with the name of the variable that is this struct, e.g. ...
       ''trainedModel''. \n \nThe table, T, must contain the ...
       variables returned by: \n  c.RequiredVariables \nVariable ...
       formats (e.g. matrix/vector, datatype) must match the original ...
       training data. \nAdditional variables are ignored. \n \nFor ...
       more information, see <a ...
       href="matlab:helpview(fullfile(docroot, ''stats'', ...
       ''stats.map''), ''appregression_exportmodeltoworkspace'')">How ...
       to predict using an exported model</a>.');
67
68 % Extract predictors and response
69 % This code processes the data into the right shape for training the
70 % model.
71 inputTable = trainingData;
72 predictorNames = {'Voz', 'Pz'};
73 predictors = inputTable(:, predictorNames);
74 response = inputTable.Tz;
75 isCategoricalPredictor = [false, false];
76
77 % Compute resubstitution predictions
78 validationPredictions = predict(trainedModel.RegressionGP, ...
       predictors);
79
80 % Compute validation RMSE
81 validationRMSE = sqrt(resubLoss(trainedModel.RegressionGP, ...
       'LossFun', 'mse'));
```

B. MLR MODEL

The MATLAB code used for this model is as follow:

```
function [trainedModel, validationRMSE] = ...
    trainRegressionModel(trainingData)

% [trainedModel, validationRMSE] = trainRegressionModel(trainingData)
% returns a trained regression model and its RMSE. This code ...
    recreates the model trained in Regression Learner app. Use the ...
    generated code to automate training the same model with new ...
    data, or to learn how to programmatically train models.

%  Input:
%      trainingData: a table containing the same predictor and ...
    response columns as imported into the app.

%  Output:
%      trainedModel: a struct containing the trained regression ...
    model. The struct contains various fields with information ...
    about the trained model.

%      trainedModel.predictFcn: a function to make predictions on ...
    new data.

%      validationRMSE: a double containing the RMSE. In the app, ...
    the History list displays the RMSE for each model.

% Use the code to train the model with new data. To retrain your ...
    model, call the function from the command line with your ...
    original data or new data as the input argument trainingData.
```

```
17
18 % For example, to retrain a regression model trained with the ...
      original dataset T, enter:
19 %   [trainedModel, validationRMSE] = trainRegressionModel(T)
20
21 % To make predictions with the returned 'trainedModel' on new ...
      data T2, use yfit = trainedModel.predictFcn(T2)
22
23 % T2 must be a table containing at least the same predictor ...
      columns as used during training. For details, enter:
24 %   trainedModel.HowToPredict
25
26 % Auto-generated by MATLAB on 06-Apr-2020 14:21:44
27
28
29 % Extract predictors and response
30 % This code processes the data into the right shape for training the
31 % model.
32 inputTable = trainingData;
33 predictorNames = {'flow-a', 'flow-b', 'int'};
34 predictors = inputTable(:, predictorNames);
35 response = inputTable.kw;
36 isCategoricalPredictor = [false, false, false];
37
38 % Train a regression model
39 % This code specifies all the model options and trains the model.
40 concatenatedPredictorsAndResponse = predictors;
41 concatenatedPredictorsAndResponse.kw = response;
42 linearModel = fitlm(...
43     concatenatedPredictorsAndResponse, ...
44     'linear', ...
45     'RobustOpts', 'off');
46
47 % Create the result struct with predict function
48 predictorExtractionFcn = @(t) t(:, predictorNames);
```

```
49 linearModelPredictFcn = @(x) predict(linearModel, x);
50 trainedModel.predictFcn = @(x) ...
       linearModelPredictFcn(predictorExtractionFcn(x));
51
52 % Add additional fields to the result struct
53 trainedModel.RequiredVariables = {'flow-a', 'flow-b', 'int'};
54 trainedModel.LinearModel = linearModel;
55 trainedModel.About = 'This struct is a trained model exported ...
       from Regression Learner R2019b.';
56 trainedModel.HowToPredict = sprintf('To make predictions on a new ...
       table, T, use: \n  yfit = c.predictFcn(T) \nreplacing ''c'' ...
       with the name of the variable that is this struct, e.g. ...
       ''trainedModel''. \n \nThe table, T, must contain the ...
       variables returned by: \n  c.RequiredVariables \nVariable ...
       formats (e.g. matrix/vector, datatype) must match the original ...
       training data. \nAdditional variables are ignored. \n \nFor ...
       more information, see <a ...
       href="matlab:helpview(fullfile(docroot, ''stats'', ...
       ''stats.map''), ''appregression_exportmodeltoworkspace'')">How ...
       to predict using an exported model</a>.');
57
58 % Extract predictors and response
59 % This code processes the data into the right shape for training the
60 % model.
61 inputTable = trainingData;
62 predictorNames = {'flow-a', 'flow-b', 'int'};
63 predictors = inputTable(:, predictorNames);
64 response = inputTable.kw;
65 isCategoricalPredictor = [false, false, false];
66
67 validationPredictFcn = @(x) linearModelPredictFcn(x);
68
69 % Compute resubstitution predictions
70 validationPredictions = validationPredictFcn(predictors);
71
```

```
% Compute validation RMSE
isNotMissing = ¬isnan(validationPredictions) & ¬isnan(response);
validationRMSE = sqrt(nansum(( validationPredictions - response ...
    ).^2) / numel(response(isNotMissing) ));
```